

Modellierung von peripersonalem
Raum und Interaktionsraum durch
ein Körperschema für humanoide
Agenten

Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

Nhung Nguyen

Modellierung von peripersonalem Raum und Interaktionsraum durch ein Körperschema für humanoide Agenten

Nhung Nguyen
AG Wissensbasierte Systeme (Künstliche Intelligenz)
Technische Fakultät
Universität Bielefeld
Postfach 10 01 31
D-33501 Bielefeld
Germany
E-Mail: nnguyen@techfak.uni-bielefeld.de

Von der Technischen Fakultät der Universität Bielefeld zur Erlangung des Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Dissertation.

Dekan der Fakultät: Prof. Dr. Jens Stoye
Erster Gutachter: Prof. Dr. Ipke Wachsmuth
Zweiter Gutachter: Prof. Dr. Helge Ritter

Einreichung der Arbeit: 28. September 2012
Tag der Disputation: 19. Dezember 2012

Die offizielle Druckversion wurde auf alterungsbeständigem Papier gemäß DIN-ISO 9706 gedruckt.

Danksagung

Diese Arbeit entstand im Rahmen meiner Tätigkeit als Wissenschaftliche Mitarbeiterin im Sonderforschungsbereich 673 – Ausrichtung in der Kommunikation und in der Arbeitsgruppe Wissensbasierte Systeme (Künstliche Intelligenz) der Technischen Fakultät der Universität Bielefeld. Dieses Umfeld machte diese Arbeit erst möglich.

Mein besonderer Dank gilt meinem Betreuer Ipke Wachsmuth für die wertvollen Ratschläge, die jahrelange Unterstützung und für viele bestärkende Worte. Helge Ritter danke ich herzlich für die Begutachtung der Arbeit.

Ich danke allen ehemaligen und aktuellen Mitgliedern der Arbeitsgruppe Wissensbasierte Systeme für die konstruktiven Anregungen und die gute Zusammenarbeit. Den Hilfskräften möchte ich für ihr Engagement danken. Besonders danke ich meinem Binom Hana, den Nerdettes Alexa und Maha für den Zusammenhalt und für eine unvergessliche Zeit.

Ich danke meinen Eltern und beiden Brüdern für all den Rückhalt und die liebevolle Unterstützung einer manchmal nicht ganz einfachen Tochter und Schwester. Ich danke Christoph für das Mitleiden und Mitfreuen. Ich danke allen Mitwohnis und Freund*innen, die mich immer daran erinnern haben, worauf es wirklich ankommt.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	3
1.2	Problemstellungen: Virtuelle Menschen in Nahraumaufgaben	3
1.3	Ziele und Lösungsansätze	6
1.4	Struktur der Arbeit	8
2	Der peripersonale und interpersonale Raum – theoretische Konzepte	11
2.1	Individuelle Körperräume	12
2.1.1	Referenzsysteme für Raumrepräsentationen	12
2.1.2	Transformation zwischen Referenzsystemen	14
2.1.3	Raummodelle für Nah- und Fernräume	15
2.1.4	Greifbewegungen im Nahraum	20
2.2	Eigenschaften peripersonaler Raum- und Körperrepräsentation	21
2.2.1	Peripersonaler Raum	23
2.2.2	Körperschema	25
2.2.3	Modulierung von peripersonalem Raum und Körperschema	27
2.3	Interpersonaler Raum	29
2.3.1	Repräsentation interpersonaler Räume	30
2.3.2	Interpersonale Raummodelle	31
2.4	Zusammenfassung	33
3	Humanoide Agenten in Nahrauminteraktion – der Stand der Forschung	35
3.1	Bewegungsplanung und -kontrolle	35
3.1.1	Grundlagen zur Bewegungsanimation virtueller Menschen	36
3.1.2	Ansätze zur Armbewegungsplanung	41
3.1.3	Ansätze zur Modulation der Bewegungsqualität	44
3.1.4	Resümee	45

3.2	Raum- und Körperrepräsentationen	46
3.2.1	Peripersonaler Raum für humanoide Agenten	46
3.2.2	Körperrepräsentationen für humanoide Agenten	51
3.2.3	Ansätze zur Generierung eines künstlichen Körperschemas	54
3.2.4	Vergleich und Analyse	56
3.3	Interpersonale Körperräume	60
3.4	Nahraumaufgaben mit humanoiden Agenten	63
3.5	Zusammenfassung	66
4	Entwurf eines peripersonalen und Interaktionsraums für virtuelle Menschen	69
4.1	Technische Interpretation der theoretischen Konzepte	69
4.1.1	Ein peripersonaler Raum für humanoide Agenten	70
4.1.2	Ein künstliches Körperschema	72
4.1.3	Kooperatives Verhalten im interpersonalen Raum	73
4.2	Max – ein virtueller Mensch	74
4.3	Sensorische Voraussetzungen für virtuelle Menschen	77
4.4	Lösungsansatz	81
5	Ein adaptives künstliches Körperschema für virtuelle Agenten	83
5.1	Problemstellung	83
5.1.1	Vergleich funktionaler und geometrischer Körperschemaeigenschaften	84
5.1.2	Anforderungen einer peripersonalen Raumrepräsentation	85
5.1.3	Ein Körperschema-Lernverfahren	86
5.2	Taktiler Lernen des Körperschemas	87
5.2.1	Initialisierung	87
5.2.2	Taktiler Lernprozess	89
5.3	Evaluation	95
5.3.1	Visueller und taktiler Lernprozess	97
5.3.2	Taktile Body-Babbling-Strategie	98
5.3.3	Startwerte über Tastrezeptorposition	99
5.4	Diskussion	99
5.5	Zusammenfassung und Resümee	103

6	Ein Modell des peripersonalen Raums für virtuelle Agenten	105
6.1	Voranalyse	105
6.2	Strukturierung des peripersonalen Raums	107
6.3	Technische Umsetzung für virtuelle Menschen	110
6.3.1	Peripersonale Unterräume für den virtuellen Agenten Max . . .	111
6.3.2	Arm- und Torsobewegungsplanung	114
6.4	Zusammenfassung und Diskussion	121
 7	 Von peripersonalem Raum zu kooperativem Agentenverhalten	 123
7.1	Ein Modell des Interaktionsraums	123
7.1.1	Räumliches Partnermodell	124
7.2	Konstruktionsaufgabe in kooperativer Interaktion	126
7.2.1	Technischer Aufbau	126
7.2.2	Spielregeln	128
7.2.3	Szenario-Komponenten	129
7.2.4	Spielabläufe und interne Verarbeitung	130
7.3	Empathie als Basis für räumlich-kooperatives Verhalten	136
7.3.1	Modulierung von Hilfsaktionen über soziale Faktoren	137
7.3.2	Kostenfunktion	138
7.3.3	Modulierung der Hilfsbewegung über den Grad der Empathie .	139
7.3.4	Spielabläufe und interne Verarbeitung	144
7.4	Zusammenfassung und Diskussion	152
 8	 Resümee	 155
8.1	Beiträge zum Stand der Forschung	157
8.2	Ausblick	160
 Literaturverzeichnis		 163

1 Einleitung

„Robotic manipulators, including humanoid upper torsos, can be dangerous to humans and in most cases are not recommended for interaction within distances at which physical contact is possible.“(Brooks und Arkin, 2007, S. 57).

Würde in der Robotik-Forschung dieser Hinweis von Brooks und Arkin befolgt werden, so wäre nur schwer vorstellbar, wie die viel prophezeiten Haushaltsroboter uns in Zukunft bei täglichen Arbeiten Hilfe leisten sollen. Der Gedanke, dass ein elektronischer Assistent uns im Alltag zur Seite steht, ist ein häufiges Motiv in der Entwicklung sogenannter künstlicher verkörperter Agenten. Diese sollen uns manuelle Tätigkeiten demonstrieren oder uns in diesen sogar unterstützen können. Interaktionen auf nahe Distanzen mit solchen verkörperten Agenten können für bestimmte Aufgaben notwendig sein. Welche Konsequenzen dies für die Entwicklung von geeigneten räumlichen Fähigkeiten und für die Eigenschaften von verkörperten Agenten allgemein hat, soll im Verlauf dieser Arbeit deutlich gemacht werden.

Verkörperte Agenten sind intelligente Maschinen oder Computerprogramme, die über einen künstlichen Körper verfügen, mit dem sie in ihrer Umgebung Handlungen durchführen und auf menschenähnliche Weise kommunizieren können. Die Überlegung dahinter ist, dass Computersysteme, ausgestattet mit einem künstlichen Körper, ausreichender Sensorik und Wahrnehmung, Menschen direkt an Ort und Stelle bei Aufgaben unterstützen sollen. Ein weiterer Gedanke ist, dass verkörperte Agenten, die sich natürlich und auf eine dem Menschen vertraute Art verhalten, die Interaktion für den Menschen erleichtern können. Beispiele für solche verkörperten Agenten sind menschenähnliche, physische Roboter, wie sie im Eingangszitat von Brooks und Arkin erwähnt werden, sowie computergrafisch simulierte Agenten mit **humanoidem**, also menschenähnlichem Erscheinungsbild, sogenannte **virtuelle Menschen** (engl. Virtual Humans).

Roboter können eingesetzt werden, um Menschen bei Hausarbeiten zur Seite zu



Abbildung 1.1: Linke Seite: Kochroboter Motoman SDA10 (Yaskawa, 2008), hergestellt von dem Unternehmen YASKAWA, Japan (Bilder von Tsuno, 2008). Rechte Seite: Roboter JAST (Foster und Matheson, 2008) von der Universität München leitet eine Konstruktionsaufgabe an.



Abbildung 1.2: v.l.n.r.: Steve zeigt die Bedienung von Maschinen. Max assistiert bei der Konstruktion eines Baufix-Flugzeugs. Ritchie leitet eine anatomiebezogene Aufgabe an.

stehen oder sie bei Konstruktionstätigkeiten anzuleiten (siehe Abbildung 1.1). Virtuelle Menschen werden unter anderem zu Schulungs- und Trainingszwecken in virtuellen Umgebungen (Virtuelle Realität) eingesetzt, um zum Beispiel simulierte manuelle Planungs- und Konstruktionsaufgaben vorzuführen und anzuleiten (siehe Abbildung 1.2). Hier bietet die Virtuelle Realität häufig eine Möglichkeit, Trainingsumgebungen bereitzustellen, die in der Realität kostenintensiv sind, wie zum Beispiel bei medizinischen Operationen, oder schwer darzustellen, wie zum Beispiel bei Gebäudebauplanungen. Der Mensch kann über Ausgabegeräte in Form von speziellen Brillen oder Helmen die dreidimensionale virtuelle Umgebung sehen und über Aufnahme- oder Eingabegeräte, zum Beispiel in Form von Datenhandschuhen, in der virtuellen Umgebung und mit dem virtuellen Menschen interagieren.

1.1 Motivation

Angesichts der Entwicklungen in der Haushalts- und Unterhaltungselektronik werden wir in Zukunft immer mehr mit physisch und virtuell verkörperten Agenten konfrontiert sein. Die Einführung von 3-D-fähigen Fernsehgeräten und Videospiele-Konsolen für den Heimgebrauch legt jedoch die Vorstellung nah, dass wir in vielen Lebensbereichen virtuellen Menschen noch vor den Haushaltsrobotern begegnen werden. 3-D-fähige Videospiele-Konsolen haben durch niedrige Anschaffungskosten inzwischen weite Verbreitung gefunden. Es ist nun nicht nur möglich, Inhalte über 3-D-Fernseher zu betrachten, sondern diese auch über 3-D-Kameras mit natürlicher Eingabe über Körperposition und -bewegungen zu steuern. Virtuelle Realität ist somit bereits mit den heutigen Mitteln der Unterhaltungselektronik nicht länger nur in groß angelegten Assistenz- und Tutorssystemen für den professionellen Gebrauch, sondern auch für den Privatgebrauch nutzbar. Mit diesen Mitteln wachsen auch die Möglichkeiten, wie und wo Menschen mit virtuellen Menschen interagieren können.

Neben den zahlreichen Spieleanwendungen, in denen virtuelle Charaktere bereits heute eingesetzt werden, sind virtuelle Menschen als interaktive Tutoren in 3-D-Lern- und Bildungssoftware denkbar. Experimente für den Chemie- und Physikunterricht sind nur ein Beispiel, wie praktische Tätigkeiten in virtueller Realität gemeinsam mit virtuellen Menschen durchgeführt werden könnten.

1.2 Problemstellungen: Virtuelle Menschen in Nahraumaufgaben

Wie die genannten Anwendungsbeispiele zeigen, erfordern manuelle Aufgaben, bei denen bestimmte Abläufe kooperativ durchgeführt werden müssen, zusätzlich räumliche Nähe zwischen den Partnern. Im Folgenden werden diese Aufgaben als **Nahraumaufgaben** bezeichnet. In dieser Arbeit werden sie durch folgende Eigenschaften charakterisiert: Die am häufigsten auftretende Objektmanipulation ist das Umplatzen von Objekten. Diese Manipulation wird für die Konstruktion, für die Reorganisation, aber auch für das Anreichen von Objekten an einen Partner benötigt. Die Objekte befinden sich in Distanzen, die sich ungefähr in Reichweite der Arme bzw. des Oberkörpers befinden. Aus diesem Grund erfordern Nahraumaufgaben Arm- und Torsobewegungen.

Sollen uns verkörperte Agenten gerade bei Nahraumaufgaben eine Hilfe sein, scheinen Interaktionen auf nahe Distanz also unvermeidlich. Es stellt sich daher die Frage, ob verkörperte Agenten mit ihren bisherigen räumlichen Interaktionsfähigkeiten den Anforderungen gewachsen sind, Menschen auch in solchen Situationen kooperativ zu unterstützen. Ein Beispiel für eine Nahraumaufgabe mit virtuellen Menschen könnte ein gemeinsam durchzuführendes chemisches Experiment sein. Hier könnte der virtuelle Mensch einen Glaskolben halten, während der Mensch gleichzeitig eine Flüssigkeit einfüllt. Solche Szenarien, die in virtueller Realität leicht denkbar sind, können in der Realität zwischen physischen Robotern und Menschen nicht leicht umgesetzt werden, da in der Regel Sicherheitsmaßnahmen beachtet werden müssen, um die Verletzungsgefahr bei Kollisionen zu vermeiden. Verdeutlicht wird dies durch das einführende Zitat dieses Kapitels von Brooks und Arkin. In Szenarien mit Robotern wird daher immer ein gewisser Sicherheitsabstand zum Menschen eingehalten, sodass natürliche und kooperative Interaktionen, die manuelle Handlungen in naher Distanz zwischen den Partnern erfordern, in der Robotik nicht zu finden sind.

Obwohl sich virtuelle Menschen als proaktive Helfer gerade für Szenarien mit geringer Distanz zum Menschen anbieten, ähneln viele Anwendungsumgebungen, in denen sie bisher eingesetzt werden, kaum natürlichen Interaktionsumgebungen. Sei es der virtuelle Anatomie-Assistent Ritchie (Wiendl, Dorfmueller-Ulhaas, Schulz und André, 2007), der Menschen bei der Platzierung von Organen in ein menschliches Skelett anleitet, oder Steve (Johnson und Rickel, 1997), der als Tutor Studierenden in virtueller Realität Handgriffe an virtuellen Maschinen vorführt, oder Max (Kopp, Jung, Lessmann und Wachsmuth, 2003), der mit einem menschlichen Partner an einem Tisch steht und einen Propeller baut (siehe Abbildung 1.2). Die räumlichen Distanzen sind auch in solchen Szenarien häufig so gewählt, dass Wechselwirkungen oder gar Beeinträchtigungen der Bewegungen zwischen den Partnern ausgeschlossen sind. Die individuellen und gemeinsamen Arbeitsbereiche werden genau festgelegt und getrennt, wodurch natürliche Interaktionen und möglicherweise notwendige kooperative Handlungen erschwert werden. Für künstliche Agenten, unabhängig von ihrer physischen oder virtuellen Verkörperung, bergen solche Nahraumaufgaben, in denen räumliche Interaktion auf begrenztem Platz stattfindet, viele Herausforderungen. Im Folgenden werden diese genauer beleuchtet, um die verschiedenen Anforderungen an virtuelle Menschen zu identifizieren, aus denen sich die Fragestellungen dieser Arbeit ableiten.

Die Probleme der Kollisionserkennung, -vermeidung und Bewegungsplanung für In-

dustrieroboter sind zwar weitgehend gut erforschte Gebiete, die Arbeitsbereiche sind jedoch häufig so konzipiert, dass der gesamte Bereich möglichst vollständig durch Kameras oder ähnliche Aufnahmetechniken verfolgt werden muss. Meist wird die Szene aus einer Vogelperspektive aufgenommen (z. B. Projekt JAST. Vgl. Rickert, Foster, Giuliani, By, Panin und Knoll, 2007; Foster und Matheson, 2008). Solch ein Gesamtüberblick über alle Arbeitsbereiche ist in natürlicher Interaktion kaum vorauszusetzen, sodass ein virtueller Mensch mit der ihm verfügbaren sensorischen Wahrnehmung den Arbeitsbereich im Überblick behalten muss. Diese Anforderung muss also in der **Wahrnehmung und Repräsentation des Nahraums** für einen virtuellen Menschen berücksichtigt werden.

Bei manuellen Aufgaben, in denen mehrere Objekte zum Einsatz kommen und im Arbeitsbereich verteilt sind, haben Menschen ein intuitives Gefühl dafür, welche Objekte sie leicht und welche sie nur durch zusätzliches Vorlehnen des Oberkörpers erreichen können. Solch ein Verhalten stellt Anforderungen an die **Raumrepräsentation** und gleichzeitig an die **räumlich-motorischen Fähigkeiten** eines virtuellen Menschen. Das automatische Wissen darüber, wie weit der Körper eines virtuellen Menschen reicht, ist eine Anforderung an seine **Körperrepräsentation**. Nicht zuletzt muss ein virtueller Mensch in natürlichen Nahraumumgebungen neben dem Umgang mit körpernahen Gegenständen einen weiteren Faktor berücksichtigen: den Kooperationspartner. Ein virtueller Mensch sollte daher mit räumlichen Verhaltensweisen ausgestattet sein, die nicht nur in seinem individuellen Arbeitsbereich, sondern auch in Interaktion mit einem Partner einsetzbar sind. Dies bedeutet zum Beispiel, dass der virtuelle Mensch bei eigenen Armbewegungen mögliche Armbewegungen des Partners mitberücksichtigt, um diesen nicht zu behindern. Auch diese Vorgänge bewältigen Menschen fließend, mit einem intuitiven Gefühl dafür, wie weit sie selbst und das Gegenüber mit den Armen und dem Oberkörper reichen können. Die Anforderungen an die **Partnerinteraktion** sollten daher zusätzlich in der Raum- und Körperrepräsentation und den räumlich-motorischen Fähigkeiten des virtuellen Menschen berücksichtigt werden.

Obwohl manuelle Handlungen in geringen Distanzen zwischen den Partnern unvermeidbar scheinen und die genannten räumlichen Verhaltensweisen die Interaktionsfähigkeiten verkörperter Agenten in Nahraumaufgaben verbessern könnten, finden sie in bisher existierenden Anwendungsszenarien keine ausreichende Berücksichtigung.

1.3 Ziele und Lösungsansätze

Räumliche Interaktion in Nahraumaufgaben, die Menschen scheinbar mühelos meistern, stellt, wie in Abschnitt 1.2 deutlich macht, besondere Anforderungen an die **räumliche Perzeption**, an die **Körper- und Raumrepräsentation** und an die **räumlich-motorischen Fähigkeiten** eines virtuellen Menschen. Ziel dieser Arbeit ist es daher, ein Modell für den Raum, der den virtuellen Menschen unmittelbar umgibt, zu entwickeln und dieses Modell prototypisch in ein Interaktionsszenario einzubetten. Dieses Raummodell soll verschiedene künstliche sensorische Modalitäten zur Raumwahrnehmung einbeziehen, um eine möglichst flächendeckende Nahraumrepräsentation zu bilden, die die Planung und Generierung weitgehend domänenunabhängiger Arm- und Torsobewegungen in Nahraumaufgaben mit und ohne Partner erlaubt. Der virtuelle Mensch soll also mit räumlichen Verhaltensweisen und Repräsentationen ausgestattet werden, die nicht nur in seinem individuellen Arbeitsbereich, sondern auch in Interaktion mit einem Partner einsetzbar sind. Das gezeigte Verhalten soll dabei weitgehend natürliche und kooperative Partner-Interaktionen in Nahraumaufgaben ermöglichen. In der Partner-Interaktion bezeichnen wir den Raum, den die Partner gemeinsam teilen, als **Interaktionsraum**. In dieser Arbeit befassen wir uns mit den folgenden zentralen Fragen, um das gesetzte Ziel zu erreichen:

Welche Erkenntnisse über den menschlichen Nahraum und die Körperrepräsentation können die räumlichen Fähigkeiten eines virtuellen Menschen verbessern?

Welche Eigenschaften hat der Interaktionsraum, den ein virtueller Mensch mit einem Partner teilt?

Wie kann ein virtueller Mensch diese Räume und Distanzen darin wahrnehmen und repräsentieren?

Wie können seine Bewegungen in diesen Räumen geplant und generiert werden?

Die in Abschnitt 1.2 identifizierten Anforderungen werfen folgende Teilprobleme auf, die in den nächsten Kapiteln behandelt werden. Zwar wurden die Lösungsansätze für den virtuellen Menschen Max realisiert, die Anforderungen werden jedoch zunächst generell für humanoide Agenten formuliert. Sind die Ansätze nicht oder nur schwer auf physische Roboter übertragbar, so wird dies im Verlauf der Arbeit deutlich gemacht.

P1 – Wahrnehmung und Repräsentation des Nahraums. Für humanoide Agenten stellt sich die Frage, wie sie den umgebenden Raum mit den ihnen verfügbaren Sensormodalitäten so gut wie möglich wahrnehmen und repräsentieren

können. Ihre Fähigkeiten, in der nahen Umgebung zu agieren, hängen stark von einer strukturierten und leicht zugänglichen Raumrepräsentation ab. Dafür könnte sich ein Nahraummodell anbieten, das diese Repräsentation liefert. Besitzt der humanoide Agent eingeschränkte sensorische Systeme, so muss ein Modell damit umgehen können.

P2 – Repräsentation des Körpers. Das Wissen darüber, wie groß der peripersonale Raum ist, hängt damit zusammen, wie weit der virtuelle Mensch mit Armen und Oberkörper reichen kann. Dieses Wissen ist wiederum von einer geeigneten Körperrepräsentation abhängig. Mit dieser soll es insbesondere möglich sein, auf einfache und automatisierte Weise räumliche Distanzen zwischen Körper und umgebenden Gegenständen zu ermitteln.

P3 – Intelligentes räumliches Verhalten. Damit sind Motoraktionen im Nah- und Interaktionsraum gemeint. Die räumlichen Verhaltensweisen, wie das Umplatzen von Objekten, sollen für individuelle und Partner-Interaktionen geeignet sein. Dabei soll das räumliche Verhalten anpassungsfähig und für den Partner intuitiv verständlich sein.

Für die Lösung dieser Teilprobleme wird in der vorliegenden Dissertation ein interdisziplinärer Ansatz verfolgt, bei dem Wissen über kognitive Mechanismen und Verhaltensweisen von Menschen in den technischen Modellierungsprozess einbezogen werden. Ein für diese Arbeit zentrales Konzept ist der Raum, der den menschlichen Körper unmittelbar umgibt und den wir mühelos mit unseren Armen erreichen können. Dieser Raum wird vom Gehirn als zum Körper gehörend repräsentiert und in der Psychologie und den Neurowissenschaften als **peripersonaler Raum** bezeichnet. Ergebnisse über dessen Struktur und Repräsentation können eine wertvolle Unterstützung bilden in der technischen Modellierung des Nahraums für künstliche verkörperte Agenten. So hängt beispielsweise der peripersonale Raum eng mit der mutmaßlich bei Menschen existierenden Körperrepräsentation, dem sogenannten **Körperschema**, zusammen. Ob diese theoretischen Konzepte bzw. Teilaspekte dieser Konzepte sich möglicherweise zur Lösung der Teilprobleme eignen, wird diskutiert. Zudem wird ein Bezug hergestellt zu existierenden technischen Lösungsansätzen. Nach eingehender Analyse werden für die Teilprobleme P1 und P2 die Konzepte des peripersonalen Raums und des Körperschemas als Modellierungsgrundlage gewählt, und es wird eine technische

Interpretation entwickelt. Auf dieser Basis werden die technischen Lösungen für die Teilprobleme vorgestellt.

Forschungskontext der Arbeit

Diese Arbeit entstand im Forschungskontext des Sonderforschungsbereiches 673 „Ausrichtung in der Kommunikation“ (engl. Alignment in Communication), gefördert von der Deutschen Forschungsgemeinschaft. Der Sonderforschungsbereich untersucht Phänomene, bei denen menschliche Gesprächspartner sich interaktiv aneinander ausrichten. Damit sind Prozesse gemeint, die in der menschlichen Kommunikation dazu führen, dass Dialogpartner ihre mentalen linguistischen Repräsentationen automatisch und ohne explizite Aushandlung aneinander angleichen. Die Forschungsziele werden interdisziplinär und mit dem Fokus auf dem Ansatz der **interaktiven Ausrichtung** verfolgt. Dieser Ansatz wird als bedeutsam für die Erforschung erfolgreicher Kommunikation angesehen, bei der die weitgehend automatisierten Angleichungsprozesse bei Menschen ressourcensparend ablaufen (Pickering und Garrod, 2004). In diesem Zusammenhang untersucht das Teilprojekt *C1 - Interaktionsraum* die räumliche Angleichung von künstlichen Agenten an ihre Interaktionspartner in Manipulationsaufgaben, die auf räumlich begrenztem Platz, zum Beispiel einem Tisch, stattfinden. Ein Motiv dieses Teilprojekts ist die Frage, inwiefern die Angleichung von räumlichen Repräsentationen der Interaktionspartner die gesamte Interaktion beeinflusst. Die hier vorgestellte Arbeit ist mit dem Projektteil *Künstliche Intelligenz* verknüpft, in dem zwei Agenten (menschliche und künstliche) sich einen gemeinsamen sogenannten **Interaktionsraum** teilen. Dieser entsteht durch die zwei überlappenden peripersonalen Räume. In diesem Szenario werden die Konsequenzen aus der räumlichen Perspektivenübernahme des Partners bei gemeinsamen Manipulationsaufgaben in virtueller Realität untersucht.

1.4 Struktur der Arbeit

Die für diese Arbeit relevanten theoretischen Konzepte aus den Kognitions- und Neurowissenschaften werden in Kapitel 2 eingeführt und in Kapitel 3 im Hinblick auf ihre Anwendung in technischen Modellen diskutiert. Auf Basis der theoretischen und technischen Grundlagen wird in Kapitel 4 der Gesamtlösungsansatz für die drei auf-

geworfenen Teilprobleme entwickelt. In Kapitel 5 wird eine Lösung des Teilproblems P2 dargestellt. Diese beinhaltet ein Lernverfahren, mit dem es möglich ist, über taktile Sensordaten eine Körperrepräsentation eines humanoiden Agenten zu ermitteln. In Kapitel 6 wird eine Lösung des Teilproblems P1 vorgestellt. Die Lösung besteht aus einem Nahraummodell basierend auf kognitiven und neuropsychologischen Befunden. Es ermöglicht eine umfassendere Raumabdeckung und Strukturierung der Aktionen des humanoiden Agenten. In Kapitel 7 fließen die Lösungen der Teilprobleme P1 und P2 ein und bilden damit eine Lösung des Teilproblems P3. Die Lösung wird durch ein interpersonales Raummodell vorgeschlagen, mit dem ein humanoider Agent kooperatives Verhalten in räumlicher Interaktion mit einem menschlichen Partner zeigen kann. Im Abschlusskapitel 8 folgt eine Zusammenfassung der wichtigsten Erkenntnisse aus den einzelnen Kapiteln. Die einzelnen Beiträge zum Stand der Forschung wird identifiziert und ein Ausblick auf weiterführende Aspekte vorgestellt.

Teile dieser Dissertation wurden bereits in folgenden Arbeiten vorveröffentlicht: Nguyen und Wachsmuth (2009a), Nguyen und Wachsmuth (2009b), Nguyen und Wachsmuth (2011), Boukricha, Nguyen und Wachsmuth (2011), Nguyen und Wachsmuth (2013).

2 Der peripersonale und interpersonale Raum – theoretische Konzepte

In diesem Kapitel werden theoretische Konzepte über den Raum eingeführt, der den menschlichen Körper unmittelbar umgibt, soweit sie Implikationen für diese Arbeit beinhalten. Diese Implikationen werden im weiteren Verlauf durch Verweise zu den in der Einleitung unter Abschnitt 1.3 definierten Teilproblemen deutlich gemacht. Neben dem individuellen Raum und dem Zusammenhang zwischen Körper- und Raumrepräsentationen wird der interpersonale Raum zwischen Interaktionspartnern behandelt. Die für diese Arbeit wichtigsten Begriffe und Konzepte sowie geeignete Definitionen sollen zunächst eingeführt und vorgestellt werden. Für den theoretischen Hintergrund werden verschiedene Forschungsgebiete einbezogen, die die einzelnen Argumente für das in dieser Arbeit entwickelte theoretische Modell und damit auch für die technische Umsetzung liefern. Hierbei sind unterschiedliche Terminologien und Paradigmen aus den kognitiven Neurowissenschaften und der Neuropsychologie und -physiologie zu berücksichtigen. Im weiteren Verlauf wird auch die Bezeichnung *Körperraum* verwendet, wenn der Raum, der den menschlichen oder humanoiden Körper unmittelbar umgibt, allgemein benannt werden soll.

Betrachtet man relevante Arbeiten, so ist festzustellen, dass der individuelle Körperraum häufig unter dem Gesichtspunkt sensomotorischer und perzeptueller Charakteristiken erforscht wird. Dieser Blickwinkel wird nicht nur in den Neuro-, den Kognitionswissenschaften und in der Biologie, sondern auch in den technischen Wissenschaften eingenommen. Ein dort häufig auftretender Begriff ist der des **peripersonalen Raums**. In Interaktion mit Partnern werden die individuellen Räume üblicherweise als soziale Räume verstanden und unter dem Gesichtspunkt sozialer Zusammenhänge, abhängig von der Körperpostur und -orientierung, analysiert. Aktuelle Arbeiten zeigen jedoch eine Entwicklung, bei der der Zusammenhang zwischen gemeinsam geteilten Räumen und kognitiven Prozessen untersucht wird.

2.1 Individuelle Körperräume

In der Regel haben wir Menschen einen konsistenten Sinneseindruck von unserem Körperraum. Wir nehmen wahr, wo sich unser Körper im Raum befindet, und können weitgehend den Überblick über die Objekte um uns herum behalten. Dieser Eindruck hängt nicht nur von den Objekten in unserer Nähe ab, sondern auch unser Körper und seine Struktur haben einen Einfluss darauf, auf welche Weise wir räumliche Beziehungen wahrnehmen. Im Folgenden soll zunächst geklärt werden, in welcher Form Raumrepräsentationen möglicherweise im menschlichen Gehirn kodiert werden. Dieser Frage wird im Abschnitt 2.1.1 nachgegangen. Darauf aufbauend werden im Unterabschnitt 2.1.3 Raummodelle eingeführt, die den dreidimensionalen Raum um den Menschen herum strukturieren.

2.1.1 Referenzsysteme für Raumrepräsentationen

Um die konsistente räumliche Wahrnehmung der Umgebung aufbauen und aufrechterhalten zu können, muss das menschliche Gehirn Informationen darüber haben, wo sich die Objekte befinden. Die Frage, wie bzw. in welcher Form das Gehirn unseren Körperraum repräsentiert, ist Gegenstand vieler neuropsychologischer und -physiologischer Arbeiten. Eine zentrale Sichtweise ist, dass das Gehirn räumliche Informationen und Bewegungen des eigenen Körpers in verschiedenen **Referenzsystemen** kodiert. Ein Referenzsystem ist dabei als Koordinatensystem zu verstehen, in dem die Koordinatenachsen von einem bestimmten Ursprung ausgehen und in dem Objektpositionen bezüglich dieses Ursprungs eindeutig beschrieben werden können (Holmes und Spence, 2004; Vallar und Maravita, 2009).

Wichtige Fragen sind erstens, welche Referenzsysteme im Gehirn existieren, um Positionen von Objekten in der Umgebung zu kodieren, und zweitens, inwiefern diese tatsächlich für eine räumliche Repräsentation verantwortlich sind. Die Möglichkeiten, die dabei häufig gegenübergestellt werden, sind die **egozentrischen** und die **allozentrischen** Referenzsysteme (Lacquaniti, 1997; Vallar und Maravita, 2009). In einem allozentrischen Referenzsystem werden Objektpositionen unabhängig von einem Betrachter kodiert, zum Beispiel relativ zueinander (der Stuhl steht neben dem Tisch) bzw. relativ zu bestimmten Objekteigenschaften (der Stuhl steht an der Längsseite des Tisches). Ein egozentrisches Referenzsystem hat den Beobachter als Bezugspunkt. Hier

werden Objekte relativ zum menschlichen Körper kodiert (der Stuhl steht rechts von mir, der Tisch vor mir). Nach Vallar und Maravita (2009) bieten sich allozentrische Referenzsysteme zur Objektidentifikation und räumlichen Navigation an, wohingegen egozentrische Referenzsysteme bei zielgerichteten Bewegungen, zum Beispiel beim Greifen von Objekten und bei der Hindernisvermeidung, von Vorteil sind. Sie nehmen zudem an, dass die Umgebung in diesen Fällen im menschlichen Gehirn in egozentrischen Referenzsystemen repräsentiert wird. Da der Fokus der vorliegenden Arbeit nicht auf Fortbewegung und Navigation, sondern auf Nahraumaufgaben liegt, sollen in erster Linie egozentrische Referenzsysteme in den Blick genommen werden. Bevor in den nächsten Unterabschnitten vorgestellt wird, in welchen Referenzsystemen Körperräume repräsentiert werden und welche sensorischen Modalitäten beteiligt sind, werden hier zunächst die relevanten egozentrischen Referenzsysteme eingeführt. Der Begriff des egozentrischen Referenzsystems umfasst alle Referenzsysteme, die den gesamten menschlichen Körper oder Teile dessen als Bezugspunkt nehmen. In der Literatur werden folgende Referenzsysteme häufig untersucht und diskutiert (vgl. Lacquaniti, 1997; Cohen und Andersen, 2002; Holmes und Spence, 2004; Vallar und Maravita, 2009): **augenzentrierte**, **kopfzentrierte**, **gliedmaßenzentrierte** sowie **körperzentrierte** Systeme.

In einem **augenzentrierten** Referenzsystem werden Objekte abhängig von der Augenposition im Kopf und bei **kopfzentrierten** Referenzsystemen abhängig von der Kopfposition kodiert. Diese Referenzsysteme repräsentieren räumliche Informationen demnach unabhängig vom restlichen Körper. Von **gliedmaßenzentrierten** (engl. limb-centered) Referenzsystemen wird gesprochen, wenn der Bezugspunkt sich auf einzelne Gliedmaßen bezieht. Cohen und Andersen verwenden den Begriff zum Beispiel für Referenzsysteme, die die Hand oder den ganzen Arm als Bezugspunkt haben. Im Zusammenhang mit gliedmaßenzentrierten Referenzsystemen stößt man häufig auf die Bezeichnung **körperteilzentrierte** (engl. body-part-centered) Referenzsysteme. Der Begriff ist dabei als Oberkategorie für bestimmte Referenzsysteme zu verstehen, die nicht nur die gliedmaßenzentrierten beinhalten. Maravita, Spence und Driver (2003) zum Beispiel schließen auch weitere Körperteile wie den Kopf als Ursprung mit ein. Darunter fallen aber auch die Schultern oder, wie bei Graziano und Gross (1998), der gesamte Oberkörper als Kombination von einzelnen Körperteilen.

Der Begriff der **körperzentrierten** Referenzsysteme wird sehr uneinheitlich gebraucht und teilweise sehr weit und allgemein gefasst. In einigen Arbeiten werden

zur besseren Abgrenzung von allozentrischen Referenzsystemen körperzentrierte mit egozentrischen Referenzsystemen gleichgesetzt (Grön, Wunderlich, Spitzer, Tomczak und Riepe, 2000; Kaas und van Mier, 2006). In anderen Arbeiten wiederum werden Referenzsysteme auch als körperzentriert bezeichnet, wenn generell ausgedrückt werden soll, dass es sich um einen Bezugspunkt handelt, der nicht den Kopf oder die Augen, sondern Gliedmaßen und die übrigen Körperteile betrifft (Cohen und Andersen, 2002). Holmes und Spence hingegen vermuten eine differenzierte Repräsentation in verschiedenen Gehirnregionen, die räumliche Informationen, vergleichbar mit einer topografischen Karte der gesamten Körperoberfläche, kodiert. Dieses Referenzsystem verstehen sie als körperzentriert, und es ist laut Bremner, Mareschal, Lloyd-Fox und Spence (2008) theoretisch unabhängig von den Gliedmaßenpositionen. Um Missverständnisse zu vermeiden, wird im Folgenden der Begriff **körperzentriert** als Oberbegriff für alle körperteil- und gliedmaßenzentrierten Referenzsysteme verwendet. Ist ein spezifisches Referenzsystem gemeint, das als Unterkategorie unter den Begriff *körperzentriert* fällt, so wird es explizit als gliedmaßenzentriertes (z. B. handzentriertes) bzw. körperteilzentriertes (z. B. oberkörperzentriertes) Referenzsystem bezeichnet.

2.1.2 Transformation zwischen Referenzsystemen

Um Objekte in der unmittelbaren Umgebung greifen und manipulieren zu können, muss das Gehirn mit Sensorinformationen umgehen, die in verschiedenen Referenzsystemen kodiert sind. Während wir unsere Hand in Richtung eines Objektes bewegen, ändert sich möglicherweise die visuelle Position durch Kopf- oder Augenbewegungen. Das Gehirn benötigt daher mindestens ein festes Referenzsystem, das invariant gegenüber solchen Änderungen ist. Für Greifbewegungen wäre ein stabiles Referenzsystem zum Beispiel eines am Torso. Um das Objekt in Relation zum Torso kodieren zu können, müsste das Gehirn die Position in eine Reihe weiterer Referenzsysteme überführen. Es muss eine sogenannte *Koordinatentransformation* durchgeführt werden. Um die Objektposition von einem augen- in ein kopfzentriertes Referenzsystem zu transformieren, müssten zum Beispiel die Augenorientierung und -position bezüglich des Kopfes verwendet werden. Vom kopfzentrierten Referenzsystem könnte das Objekt dann in ein torsozentriertes überführt werden. Es wurden im Gehirn Regionen gefunden, die für diese Art der Koordinatentransformation verantwortlich gemacht werden können (Thier, 2006). Eine wichtige Beobachtung ist, dass diese Hirnregionen

scheinbar verschiedene Referenzsysteme nutzen. Betrachtet man das genannte Beispiel der torsozentrierten Greifbewegung, so erscheint es plausibel, dass für unterschiedliche Bewegungen auch unterschiedliche Referenzsysteme sinnvoll sein können.

2.1.3 Raummodelle für Nah- und Fernräume

Eine wichtige Unterscheidung zur Beschreibung von Körperräumen in der Neuropsychologie ist die zwischen Nah- und Fernräumen. Bei Menschen (Halligan und Marshall, 1991) sowie bei Affen (Mountcastle, 1976) konnten Unterschiede in der neuronalen Repräsentation, in der Funktion, in der Aufmerksamkeit und im Verhalten zwischen Fern- und Nahräumen festgestellt werden. Für den Fernraum ist der Begriff des **extrapersonalen Raums** (engl. extrapersonal space) gebräuchlich. Für den Nahraum findet sich die erstmalige Erwähnung des Begriffs **peripersonaler Raum** (engl. peripersonal space; griech. Vorsilbe: peri = um, herum) in der Arbeit von Rizzolatti, Scandolara, Matelli und Gentilucci (1981), S. 150):

„[...] was formed by neurons that required stimuli presented in the space immediately around the animal. We will refer to this space as peripersonal space [...].“

Ursprünglich definiert der Begriff einen Raum, der Makak-Affen unmittelbar umgibt. In den Folgejahren wurde die Bezeichnung auch auf den Menschen übertragen. Vor dieser Arbeit war (Semmes, Weinstein, Ghent und Teuber, 1963) und ist teilweise noch (Vallar und Maravita, 2009) in Abgrenzung zum extrapersonalen Raum der Begriff des **persönlichen Raums** (engl. personal space) üblich. Dieser Begriff bezeichnet zum einen neurowissenschaftliche Inhalte, wie sie in den Arbeiten von Rizzolatti u.a. unter dem Begriff peripersonaler Raum untersucht werden. Zum anderen betrifft der persönliche Raum aber auch Konzepte, wie sie in Arbeiten von Hall (1969) angewendet werden. Gegenstand der Fragestellungen ist dort, wie persönliche Räume, die Menschen umgeben, in der nonverbalen Kommunikation eingesetzt werden. Im weiteren Verlauf dieser Arbeit werden der *extra- und der peripersonale Raum* unter dem Blickwinkel neurowissenschaftlicher und insbesondere sensomotorischer Aspekte betrachtet. Der *persönliche Raum* hingegen wird als sozialer Raum interpretiert. Hierdurch können sensomotorische und soziale Gesichtspunkte detaillierter diskutiert und

voneinander abgegrenzt werden. Holmes und Spence (2004) definieren den peripersonalen und extrapersonalen Raum folgendermaßen (S. 94):

„Objects within peripersonal space can be grasped and manipulated; objects located beyond this space (in what is often termed ‘extrapersonal space’) cannot normally be reached without moving toward them [...].“

Nach ihrer Auffassung sind die neuronalen Repräsentationen des peri- und extrapersonalen Raums unterschiedlich, weil Objekte in nächster Nähe zum Körper potenziell mehr Auswirkungen auf diesen haben könnten als weiter entfernte Objekte. Bevor im nächsten Abschnitt erläutert werden kann, wie der peripersonale Raum im menschlichen Gehirn repräsentiert wird und welche kognitiven Mechanismen unsere Fähigkeiten in diesem Raum erklären, muss zunächst geklärt werden, welche Funktion und Ausdehnung der peripersonale Raum haben könnte. Im Folgenden werden daher Modelle zur Unterteilung des dreidimensionalen menschlichen Körperraums vorgestellt. Von den zahlreichen Raummodellen sollen drei Modelle, darunter die zwei am häufigsten diskutierten, einen Überblick darüber schaffen, wie dieser Raum prinzipiell strukturiert sein könnte. Das älteste hier vorgestellte Modell wurde von Grüsser (1983) eingeführt und ist teilweise durch neuropsychologische Befunde motiviert. Das Modell von Previc (1998) ist vollständig durch neuropsychologische und -kognitive Studien untermauert und kann als das umfassendste und differenzierteste Modell betrachtet werden. Eine ebenso fundierte Analyse existierender neurophysiologischer Befunde führt in der Arbeit von Cardinali, Brozzoli und Farnè (2009) zu einem anderen Raummodell, das als aktuellstes Modell denen von Grüsser und Previc gegenübergestellt wird.

Raummodell nach Grüsser: Grüsser (1983) verwendet in seinem Modell den Begriff des persönlichen Raums als Abgrenzung zum extrapersonalen Raum. Seine Definition des persönlichen Raums umfasst zum einen den *Ego-Raum* (engl. ego-space), der die Wahrnehmung des eigenen Körpers durch die sogenannten *inneren Sinne* (engl. inner senses) beinhaltet und dessen Grenzen durch den *Körper-Raum* (engl. body space) bestimmt sind (nicht zu verwechseln mit dem in dieser Arbeit allgemein verwendeten Begriff des Körperraums). Der extrapersonale Raum ist in vier Unterräume unterteilt und umfasst den *Greif-Raum* (engl. grasping space), den *Nah-Aktionsraum* (engl. near-distant action space), den *Fern-Aktionsraum* (engl. far-distant action space) und den *visuellen Hintergrund* (engl. visual background). Die vier Räume haben als Zentrum den menschlichen Körper und verlaufen teilweise radial um den Menschen (siehe

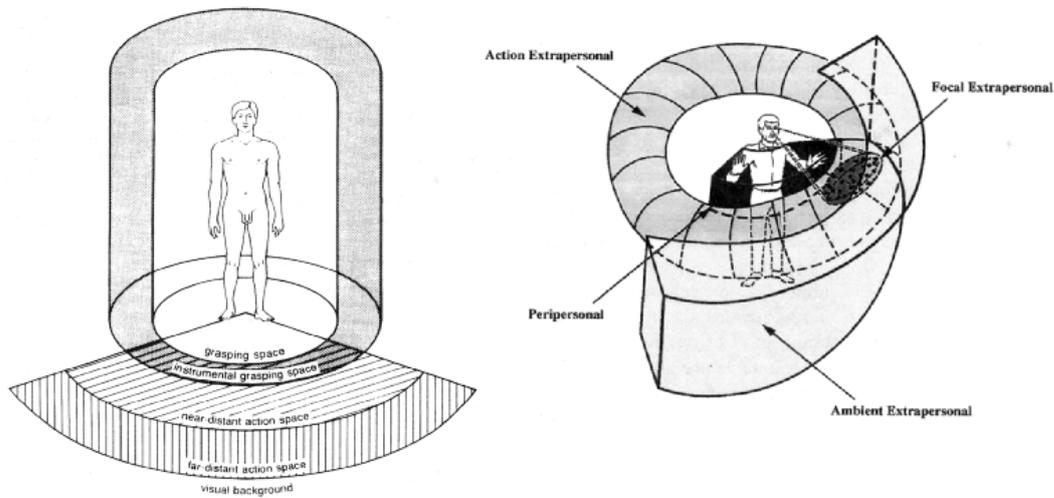


Abbildung 2.1: Rechte Seite: Grüssers Raummodell mit vier Unterräumen (Grüsser, 1983; Fig. 18-1, S. 328). Linke Seite: Previc's Raummodell mit vier Unterräumen (Previc, 1998; Figure 3, S. 126)

Abbildung 2.1). Mit größerer Distanz nimmt die Anzahl der einbezogenen Sensormodalitäten ab. Grüssers Greif-Raum ist der Raum, der den Körper unmittelbar umgibt. In diesem findet überwiegend Kontakt zwischen Objekten und der Körperoberfläche statt. Objekte haben nach Grüsser multimodale Eigenschaften und werden aus verschiedenen sensorischen Signalen ermittelt. Die Eigenschaften bleiben weitgehend invariant in den Modalitäten und der Intensität. Das bedeutet, dass ein Objekt, das vor uns auf dem Tisch liegt, für uns weiter existiert, auch wenn wir es nicht mehr sensorisch wahrnehmen, zum Beispiel beim Blick in eine andere Richtung. Der Greif-Raum ist mit visueller und taktiler Objektmanipulation und -erkennung verbunden. Der Nah-Aktionsraum ist bestimmt durch die Entfernung, in der sich ein Mensch ohne visuelle Signale noch räumlich orientieren kann. Der Radius ist wenige (6–15) Schritte weit und hängt von der Körpergröße und der Gehgeschwindigkeit ab. Der Raum kann über den Bewegungs- (kinästhetischen), Gleichgewichts- (vestibulären), visuellen, auditiven und den Geruchssinn wahrgenommen werden. Der Nah-Aktionsraum geht 6–8 Metern in den Fern-Aktionsraum über. Dieser wird hauptsächlich durch visuelle und teilweise durch vestibuläre und auditive Signale strukturiert.

Raummodell nach Previc: Das Modell von Previc (1998) ist eine Verbindung vorhandener Modelle zur Beschreibung von menschlicher Interaktion in einer dreidimensionalen Umgebung. Previc setzt sich unter anderem auch mit dem Modell von

Grüsser auseinander. Das Modell umfasst vier Unterräume (siehe Abbildung 2.1): einen peripersonalen (kurz PrP), einen fokalen extrapersonalen (kurz FcE, engl. focal extrapersonal), einen aktions-extrapersonalen (kurz AcE, engl. action extrapersonal) und einen ambient-extrapersonalen Raum (kurz AmE, engl. ambient extrapersonal space). Previc's Charakteristiken für die einzelnen Räume beinhalten die Funktion, den Ort und die Ausmaße im dreidimensionalen Raum entlang der Körperachsen sowie das vorherrschende Referenzsystem, die vorherrschenden Sensormodalitäten und die vorherrschenden Motorsysteme. Für den weiteren Verlauf sind der peripersonale und der fokale extrapersonale Raum von zentraler Bedeutung. Diese werden in diesem Abschnitt eingeführt. Die neuropsychologischen Hintergründe werden gesondert im nächsten Abschnitt mit aktuellen Arbeiten verknüpft.

Laut Previc entspricht sein **PrP** zwar Grüssers Greif-Raum, stützt sich aber auf neuropsychologische Befunde und nicht auf pragmatische Anforderungen. Die Hauptfunktionen umfassen den Verzehr von Nahrung sowie Objektgreifen und -manipulation. Objekterkennung ist in diesem Raum nicht bedeutend, da nach Meinung des Autors nur Objekte in den PrP gebracht werden, die bereits erkannt wurden. Der PrP hat eine radiale Reichweite von ungefähr zwei Metern. Die laterale Ausdehnung entspricht ungefähr der maximalen stereoskopischen Sicht von 60° , die für die Hand-Augen-Koordination wichtig ist, da ein zu erreichendes Objekt lateral zu jeder Seite mit circa 30° abgedeckt ist. Das Aufmerksamkeitssystem des PrPs hat eine vertikale Tendenz (Bias) in Richtung des Feldes unterhalb der visuellen Achse. Das vorherrschende Referenzsystem ist körperzentriert. Previc gibt die Positionen des Torsos und der Schultern als Bezugspunkte an. Die Wahrnehmung im PrP umfasst hauptsächlich die Erkennung von globalen Form-, Tiefen- und Bewegungsinformationen über das visuelle System. Positionen der Arm-, der Hand- und der weiteren muskuloskelettalen Systeme werden über die somatosensorischen und propriozeptiven Systeme bereitgestellt. Die zwei vorherrschenden Motorsysteme im PrP sind die zur Arm- und Augenbewegungskontrolle.

Der **FcE** ist der Raumbereich, auf den unsere Blickaufmerksamkeit gerichtet ist und den wir mit den Augen fixieren. Dieser Raum hat keine genaue Entsprechung im Modell von Grüsser. Die Hauptfunktion des FcEs ist die Objektsuche und -erkennung. Da unsere Aufmerksamkeit sich immer wieder verschiebt, verschiebt sich der FcE auch im Raum und kann daher nicht als ein fester Bereich um den menschlichen Körper definiert werden. Die radiale Reichweite beginnt bei den meisten Menschen bei einer minimalen Distanz von ungefähr 10–20 cm, innerhalb derer sie aber üblicherweise nicht

agieren. Die maximale Distanz ist die, hinter der ein Ziel nicht mehr visuell auflösen ist. Diese Distanz kann gegebenenfalls sehr weit sein. Die laterale Ausdehnung gibt Previc mit $20 - 30^\circ$ an. Die vorherrschende Sensormodalität ist die visuelle, da sie als einzige Modalität effektive Wahrnehmung von Objekteigenschaften mit einer geeigneten Auflösung auf räumlicher Ferne bietet. Im Gegensatz zum PrP beinhaltet die visuelle Erkennung im FcE eher lokale Informationen wie Formen und Farben. Der AcE sowie der AmE sind als ausschließlich extrapersonale Räume für uns von geringerer Relevanz und werden an dieser Stelle daher nicht weiter berücksichtigt.

Raummodell nach Cardinali, Brozzoli und Farnè: Cardinali, Brozzoli und Farnè (2009) definieren ein Raummodell, das aus drei Bereichen besteht: dem *persönlichen Raum* (engl. personal space) bzw. Perikörper (engl. peribody), dem *extrapersonalen Raum* und dem *erreichbaren Raum* (engl. reaching space). Der persönliche Raum wird hauptsächlich aus propriozeptiven und taktilen Sensorinformationen gebildet und kann zusätzlich visuelle Informationen über die Körperteile einbeziehen. Dieser Raum entspricht ungefähr Grüssers Ego- und Körper-Raum. Der persönliche Raum liegt im extrapersonalen Raum, der im Wesentlichen durch visuelle und auditorische Stimuli gebildet wird. Die Repräsentation des extrapersonalen Raums enthält Informationen über den weiter entfernten Raum. Dieser enthält einen Unterraum, den erreichbaren Raum, der sich nah am Körper befindet und funktional durch die Distanz definiert ist, innerhalb derer mit der Hand Objekte erreicht werden können, ohne den Torso zu bewegen. Der extrapersonale Raum dieses Modells entspricht somit Grüssers extrapersonalem Raum, der einen ähnlich definierten Unterraum enthält, Grüssers sogenannten Greif-Raum.

Resümee

In den vorgestellten Raummodellen ist eine Gemeinsamkeit, die egozentrisch definierten Räume, was durch die ringförmige Struktur der Unterräume, um den menschlichen Körper herum, ausgedrückt wird. Zudem wurde in den Modellen versucht, zu jedem Unterraum die dominierenden Sensormodalitäten und Funktionen zu identifizieren. Die Zuordnung zu Nah- und Fernräumen ist jedoch nicht deckungsgleich. So zählen Grüsser sowie Cardinali und Kollegen den Greif-Raum bzw. den persönlichen Raum zum extrapersonalen Raum, wohingegen Previc den peripersonalen von dem extrapersonalen Raum entkoppelt. Bei Grüsser und Cardinali u.a. wird die Körperoberfläche,

die durch Tastsinn und Propriozeption wahrgenommen wird, als Unterraum von dem extrapersonalen Raum abgegrenzt. Diese Einteilung reflektiert die Vorstellung, dass die Körperwahrnehmung eine entscheidende Rolle in der Strukturierung des uns unmittelbar umgebenden Raums einnimmt. Diese Vorstellung steht im direkten Bezug zu der noch nicht abschließend geklärten Frage, wie eine Körperrepräsentation mit der peripersonalen Raumrepräsentation zusammenhängt. In Abschnitt 2.2.3 wird diese Diskussion behandelt.

Um, wie in der Einleitung beschrieben, humanoide Agenten in Nahraumaufgaben zu verbessern, können die Erkenntnisse dieser theoretischen Modelle wertvolle Hinweise zur Lösung der in der Einleitung aufgeworfenen Teilprobleme geben. Die Raummodelle werden in späteren Kapiteln relevant sein für die Lösung der Teilprobleme zur **Wahrnehmung und Repräsentation des Nahraums (P1)** und zur Entwicklung von **intelligentem räumlichem Verhalten im Nah- und Interaktionsraum (P3)**.

2.1.4 Greifbewegungen im Nahraum

Eine wesentliche Eigenschaft der bisher vorgestellten Raummodelle bei Menschen ist, dass räumliche Distanzen mit bestimmten Funktionen belegt sein können. In diesem Abschnitt werden Befunde zu Funktionen von Armbewegungen und im Speziellen von Greifbewegungen innerhalb des menschlichen Nahraums beleuchtet. Die Wahrnehmung, ob ein Objekt in der unmittelbaren Greifreichweite liegt, entwickelt sich bei Menschen schon im Kindesalter. McKenzie, Skouteris, Day, Hartman und Yonas (1993) zeigten in ihren Studien, dass Kinder bereits ab einem Alter von 8 Monaten erkennen können, dass das Vorlehnen des Oberkörpers ihre effektive Reichweite für das Ergreifen von Objekten erweitert. Ab einem Alter von 10 Monaten können sie abschätzen, ob ein Vorlehnen zum Objektgreifen erfolgreich sein wird, und ab einem Alter von 12 Monaten erkennen sie, wie ihre Reichweite durch ein zusätzliches Werkzeug erweitert werden kann.

Mark, Nemeth, Gardner, Dainoff, Paasche, Duffy und Grandt (1997) untersuchten, ab welcher räumlichen Distanz vom Körper ausgehend Menschen Vorlehnbewegungen einsetzen. Sie unterscheiden dabei zwei mögliche Grenzen, an denen der Übergang von einem Aktionsmodus in einen anderen stattfinden könnte. Die *absolute kritische Grenze* (engl. absolute critical boundary) zwischen einfachen Armbewegungen zu Armbewegungen mit zusätzlichem Vorlehnen des Oberkörpers wird definiert als die Ma-

ximallänge des Armes. Im Gegensatz dazu beschreibt die *präferierte kritische Grenze* (engl. preferred critical boundary) eine Grenze, an der ein anderer Aktionsmodus im Vergleich zum aktuellen, komfortabler wird. In Studien, bei denen Versuchspersonen an einem Tisch sitzend Objekte in verschiedenen Distanzen frei greifen sollten, konnte gezeigt werden, dass der Übergang zwischen einfacher Armgreifbewegung und Armgreifbewegung mit zusätzlichem Vorlehnen bereits vor der absoluten kritischen Grenze stattfindet. Die durchschnittliche präferierte kritische Grenze zwischen diesen beiden Bewegungen liegt bei einem Anteil von 85 % der Distanz für die absolute kritische Grenze.

Sartori, Becchio und Castiello (2011) zeigten, dass Menschen in der Lage sind, Intentionen aus Greifbewegungen zu erkennen. In ihren Studien beobachteten Versuchspersonen Videoclips, in denen ein Mensch lediglich eine Greifbewegung mit der Hand durchführt, bei der ein vor der Person liegender Holzklötzchen in die Mitte des Tisches gestellt wird. Die vordefinierten Greifbewegungen wurden nach kinematischen Parametern modelliert, die in einer Studie von Georgiou, Becchio, Glover und Castiello (2006) für kooperative, kompetitive und individuell ausgeführte Greifaktionen spezifisch ermittelt wurden. Die Versuchspersonen konnten ohne weiteres entscheiden, ob die Bewegung mit oder ohne soziale Intention (kooperativ oder kompetitiv) ausgeführt wurde. Sie konnten dabei, ohne den Kontext der Situation zu sehen, den Unterschied zwischen einer kooperativen und langsamen bzw. einer kompetitiven und schnellen Greifbewegung erkennen. Diese Arbeit zeigt, dass unterschiedliche kinematische Eigenschaften von Greifbewegungen Intentionen und soziale Signale vermitteln können.

Diese Erkenntnisse werden in späteren Kapiteln Hinweise auf die Modellierung des Teilproblems zu **intelligentem räumlichem Verhalten im Nah- und Interaktionsraum (P3)** humanoider Agenten geben.

2.2 Eigenschaften peripersonaler Raum- und Körperrepräsentation

Wenn wir unaufmerksam an einem Kinderspielplatz vorbeigehen und ein Ball fliegt plötzlich und mit hoher Geschwindigkeit nah auf unseren Körper zu, dann versuchen wir vielleicht den Ball aufzufangen oder ihm auszuweichen. In beiden Fällen würden wir jedoch innerhalb kürzester Zeit mit einer Körperbewegung reagieren, um mit der

unerwarteten Ballsituation umzugehen. Läge der Ball allerdings nur auf unserem Weg und wir liefen unaufmerksam die Straße entlang, könnte es sein, dass wir den Ball übersehen.

Diese Beispiele veranschaulichen, wie unterschiedlich wir auf Ereignisse reagieren, abhängig davon, wie nah und potenziell gefährlich Dinge für unseren Körper sein könnten. Wie gerade die automatischen Reaktionen auf Ereignisse im peripersonalen Raum produziert werden, ist ein Gegenstand neurophysiologischer und -psychologischer Untersuchungen. In den vorangegangenen Abschnitten lag der Fokus auf der Unterteilung der Körperräume sowie den räumlichen Referenzsystemen, in denen diese im Gehirn kodiert sein könnten. Die Tatsache, dass wir trotz verschiedener Referenzsysteme eine einheitliche räumliche Vorstellung unserer Umgebung und der Objekte in ihr haben, ist eine Leistung des menschlichen Gehirns. Im Folgenden soll nun der Frage nachgegangen werden, wie diese einheitliche räumliche Wahrnehmung aufrechterhalten wird und warum wir uns im Normalfall nicht an Objekten stoßen, wenn wir unseren Körper in der Umgebung bewegen. Diese scheinbar einfachen Fähigkeiten, die den Eindruck vermitteln, wir verfügten über eine Art *räumliches Körpergefühl* für unsere Bewegungen, erfordern jedoch komplexe Prozesse im Gehirn. Das Gehirn muss dabei sensorische Signale einer sich ständig ändernden Umgebung über verschiedene Sinnesmodalitäten verarbeiten und gegebenenfalls gleichzeitig die nächsten Körperbewegungen planen.

Ein zentraler Gedanke ist, dass wir zwei Arten von Informationen benötigen, um erfolgreich im Raum agieren zu können. Um zum Beispiel Greifbewegungen im peripersonalen Raum durchführen zu können oder um Hindernisse zu vermeiden, ist es nicht nur wichtig zu wissen, *wo* relevante *Objekte* im Raum sind, sondern auch, *wo* die *eigenen Körperteile* sich relativ zu diesen befinden, wie also die aktuelle Position und Orientierung der Gliedmaßen sind. In welcher Form und in welchen Gehirnregionen unsere Umgebung bzw. die Position von Objekten in der Umgebung dargestellt wird, ist Gegenstand des Abschnittes 2.2.1. Hier werden Erkenntnisse und Eigenschaften des **peripersonalen Raums** diskutiert. Wie unsere eigenen Körperteile repräsentiert werden, ist Gegenstand von Abschnitt 2.2.2, in dem das sogenannte **Körperschema** vorgestellt wird. Beide Repräsentationen sind nötig, um die räumliche Position der Umgebung zu den Positionen der Körperteile in Beziehung zu setzen, um Körperbewegungen überhaupt planen zu können. In dieser Beziehung wird bereits die enge Wechselwirkung zwischen Körper und dem ihn umgebenden Raum deutlich. Inwiefern diese Wechselwirkung auch in den darunterliegenden Körper- und Raumrepräsentation

tionen besteht, ist nach dem derzeitigen Stand der Forschung noch nicht abschließend geklärt. Zu beobachten ist jedoch, dass neuropsychologisch motivierte Modelle des peripersonalen Raums in jüngster Zeit immer stärker die Verbindung zwischen Raum- und Körperrepräsentationen betonen.

2.2.1 Peripersonaler Raum

Holmes und Spence (2004) geben in einem Forschungsüberblick eine Zusammenfassung relevanter Resultate und Analysen aus der Neuropsychologie und -physiologie zu Körperrepräsentationen- und peripersonalen Raumrepräsentationen bei Affen und Menschen. Sie schätzen die Ausmaße des menschlichen peripersonalen Raums auf einen Raum, der die Körperoberfläche bei Menschen mit ungefähr 70 cm und bei Affen mit 20 – 40 cm Abstand umgibt. Die Autoren vertreten die Ansicht, dass die Verarbeitung von Objekten im peripersonalen Raum gegenüber der im extrapersonalen Raum genauer und komplexer ist und mehr Sensormodalitäten umfasst. Mehrere Areale im menschlichen Gehirn werden für die Kodierung von im peripersonalen Raum wahrgenommenen Stimuli verantwortlich gemacht. Da Einzelzelluntersuchungen bisher nur mit invasiven Methoden möglich sind, werden die Hypothesen hauptsächlich durch Studien an Affen, meist an Makaken, gestützt. Einige Befunde aus solchen Makak-Studien konnten in den vergangenen Jahren auch bei speziell hirngeschädigten Patienten und teilweise bei gesunden Probanden verifiziert werden. Im weiteren Verlauf wird daher der Begriff der Primaten verwendet, um die Befundlage bei Menschen und Affen zusammenzufassen.

Im Folgenden werden neurowissenschaftliche Befunde für drei Eigenschaften der peripersonalen Raumrepräsentation erläutert, die für dessen technische Umsetzung von Bedeutung sind. Erstens: Informationen verschiedener Sensormodalitäten wie der somatosensorischen Sinne (Tast-, Bewegungs- [Kinästhetik] und Gelenkstellungssinne [Propriozeption]), aber auch des Seh- und Hörsinnes tragen zu einer sogenannten **multisensorischen** peripersonalen Raumrepräsentation bei. Zweitens: Objekte innerhalb dieses Raums werden in **körperzentrierten** Koordinaten repräsentiert. Drittens: Vorhandene Objekte bleiben weitgehend **invariant** in den Sensormodalitäten. Das bedeutet, dass sie zum Beispiel auch bei Dunkelheit noch **weiterrepräsentiert** werden.

Peripersonaler Raum bei Primaten

Eine wichtige Eigenschaft für die Wahrnehmung des peripersonalen Raums ist die Zusammenführung von teilweise gleichzeitig wahrgenommenen Stimuli von verschiedenen Körpersinnen. Diese *multisensorische* Integration von Stimuli, die nah am Körper wahrgenommen werden, ist in zahlreichen Makak-Studien untersucht worden (Thier, 2006; Holmes und Spence, 2004). Dafür verantwortlich sind sogenannte *bimodale* Neurone, die bei Makak-Affen untersucht wurden (Duhamel, Colby und Goldberg, 1998, Graziano, Yap und Gross, 1994). In Experimenten beobachteten Duhamel u. a. (1998), dass die bimodalen Neurone nicht nur reagierten, wenn zum Beispiel Gliedmaßen berührt wurden, sondern auch dann, wenn visuelle Stimuli in der Nähe der Berührungsstelle präsentiert wurden. Neben visuotaktilen bimodalen Neuronen fanden Graziano, Reiss und Gross (1999) trimodale Neurone, die zudem auch auf auditorische Stimuli reagieren. Studien von Graziano u. a. (1994) legen eine Repräsentation des peripersonalen Raums bei Makak-Affen nahe, die nicht nur multisensorisch, sondern zudem *körperteilzentriert* ist. In einer weiteren Studie fanden Graziano, Hu und Gross (1997), dass Teile der bimodalen Neurone für gesehene Objekte auch bei Dunkelheit weiterfeuerten. Dieses Ergebnis zeigt, dass die Gegenwart der Objekte, unabhängig von der Sinnesmodalität, in der sie ursprünglich wahrgenommen wurden, *weiterrepräsentiert* wurde, auch wenn das Objekt nicht kontinuierlich sichtbar war. Die Autoren nehmen an, dass dieser Effekt für das Ergreifen von Objekten wichtig ist, die nicht permanent wahrnehmbar sind, sondern zum Beispiel durch ein Hindernis visuell verdeckt werden.

Bestimmte Mechanismen und Eigenschaften, die bei Affen gefunden wurden, bilden mutmaßlich auch die Grundlage der menschlichen peripersonalen Raumrepräsentation. Einige Ergebnisse aus den Makak-Studien wurden bei hirngeschädigten Patienten und bei gesunden Probanden näher untersucht. Dabei wurden auch im menschlichen Gehirn Hinweise auf eine *multisensorische* und *körperteilzentrierte* peripersonale Raumrepräsentation gefunden. Eine Vielzahl an Studien zur Untersuchung der visuotaktilen Repräsentation des peripersonalen Raums wurde von Làdavas u.a. an rechtsseitig hirngeschädigten Patienten durchgeführt (vgl. Làdavas, 2002). Patienten mit dieser Art von Schädigung eignen sich, um zu untersuchen, ob Neurone im menschlichen Gehirn den peripersonalen Raum multisensorisch repräsentieren. Tatsächlich zeigten unter anderem Di Pellegrino, Làdavas und Farné (1997) Ergebnisse, die darauf hinweisen, dass auch bei Menschen der peripersonale Raum multisensorisch repräsentiert wird. Hier-

bei werden nicht nur visuelle und somatosensorische, sondern auch auditorische (Farnè und Làdavas, 2002) Signale integriert. Di Pellegrino u.a. stellten außerdem Ergebnisse vor, die darauf hinweisen, dass die Stimuli auch bei Menschen *körperteilzentriert* repräsentiert werden. Auf Grundlage der bisher vorgestellten Befunde definieren wir den peripersonalen Raum wie folgt:

Definition 1 *Der **peripersonale Raum** ist der Raum, der mit den Armen erreicht werden kann. Die Repräsentation des Raums ist **multisensorisch**, **körperzentriert** und in den Sensormodalitäten **invariant**.*

Diese Definition bildet eine erste Grundlage für die Lösung des in der Einleitung aufgeworfenen Teilproblems der **Wahrnehmung und Repräsentation des Nahraums (P1)** für humanoide Agenten.

2.2.2 Körperschema

Die bisher genannten Studien bezogen sich auf die Kodierung von Stimuli um den Körper herum, also die neuronale Repräsentation des peripersonalen Raums. Graziano u.a. führten jedoch auch Studien durch, die Einblicke darin geben, wie die Position und Orientierung der Körperteile selbst neuronal kodiert sein könnten. Es konnte beobachtet werden, dass Makak-Gehirnareale nicht nur von visuellen und propriozeptiven Stimuli des eigenen Armes beeinflusst wurden, sondern auch von visuellen Stimuli einer Makak-Armattrappe (Graziano, 1999, Graziano, Cooke und Taylor, 2000). Je mehr die Lage der Armattrappe von der echten Armlage abwich, oder wenn stattdessen ein rechteckiges Papier in der Größe des Armes präsentiert wurde (Graziano, Cooke und Taylor, 2000), desto weniger wurden die entsprechenden Gehirnareale beeinflusst. Die Autoren kommen daher zu dem Schluss, dass das Makak-Gehirn nicht nur die Position des Armes durch visuelle und propriozeptive Informationen kodiert, sondern auch, ob der präsentierte Stimulus Ähnlichkeiten mit dem echten Arm hat.

Farnè, Pavani, Meneghello und Làdavas (2000) zeigten, dass neben dem realistischen Erscheinungsbild eine plausible Armorientierung und -position auch für Menschen wichtig sind. Thier (2006) sieht die maßgebliche Funktion der multisensorischen Repräsentation des Körpers in der Generierung eines sogenannten **Körperschemas** (engl. body schema). Der Autor bezeichnet damit eine Repräsentation des Körpers, in

der Körperteile und deren Stellung zueinander kodiert werden. Ein enger Zusammenhang besteht dabei zu der *Transformation* zwischen verschiedenen *Referenzsystemen* (vgl. Abschnitt 2.1.2). Eine weitere wichtige Funktion sieht er in Bezug zum Nahraum. Hier versteht er das Körperschema als ein Abbild des Körpers in Abgrenzung zu Objekten in der räumlichen Umgebung. Das Körperschema dient nicht nur der Berechnung der Ausmaße und der räumlichen Position des Körpers bzw. seiner Gliedmaße relativ zu Umgebungsobjekten, sondern ist generell für die *Bewegungsplanung* wichtig. Hierbei werden Informationen verschiedener Modalitäten miteinander kombiniert. So trägt es dazu bei, dass wir nicht bewusst nachdenken müssen, um zu erkennen, ob wir Objekte einfach erreichen können oder uns vorlehnen müssen (Cardinali u. a., 2009). Unter den verschiedenen Repräsentationsmodellen für den Körper ist die Beschreibung des Körperschemas von Head und Holmes (1911) die früheste, die dem nahekommt, was bei Menschen möglicherweise neuronal repräsentiert wird. Es wird vermutet, dass der Erwerb des Körperschemas bei Menschen im Säuglingsalter geschieht. So vermutete Rochat (1998) eine visuo-propriozeptive Kalibrierung des Körpers, wenn Säuglinge in den ersten Lebensmonaten viel Zeit damit verbringen, ihre eigenen Körperteile zu betrachten und sich zu berühren. In den Folgejahren wurde der Prozess auch *Body Babbling* genannt. Übersetzt bedeutet es soviel wie mit dem Körper *plappern*. Eine Annahme ist, dass Säuglinge auf diese Weise die sensorischen und motorischen Zusammenhänge ihrer Körperbewegungen lernen, die im Körperschema repräsentiert werden. Gallagher (2005) arbeitete das Konzept des Körperschemas mit neueren Forschungsergebnissen weiter heraus. Eine wichtige begriffliche Unterscheidung, die Gallagher dabei vorgenommen hat, ist die zwischen dem Konzept des Körperschemas und des *Körperbildes* (engl. *body image*). Diese beiden Begriffe werden in Forschungsarbeiten häufig synonym oder ungenau verwendet. Das Körperbild beinhaltet Einstellungen, Überzeugungen und teilweise die bewusste Wahrnehmung, die Menschen von ihrem eigenen Körper und dessen Erscheinungsbild haben. Im Gegensatz dazu handelt es sich bei dem *Körperschema* um eine vorwiegend *unbewusste* und *sensomotorische Repräsentation* des Körpers. In dieser Arbeit wird die Bezeichnung des Körperschemas verwendet und wie folgt definiert:

Definition 2 *Das Körperschema ist eine unbewusste neuronale Repräsentation des Körpers, die durch Informationen verschiedener Sensormodalitäten gebildet wird. Die Repräsentation ist anpassungsfähig und kann sich mit Änderungen des Kör-*

*pers ebenfalls ändern. Sie bildet die Grundlage für die **Transformation zwischen Referenzsystemen** und für die **Bewegungsplanung**.*

Diese Definition bildet eine Grundlage für die Lösung des in der Einleitung aufgeworfenen Teilproblems der **Repräsentation des Körpers (P2)** für humanoide Agenten.

2.2.3 Modulierung von peripersonalem Raum und Körperschema

Obwohl die Erforschung der Funktion und der Repräsentation des peripersonalen Raums bereits tiefere Erkenntnisse erzielt hat, ist die räumliche Grenze zwischen peripersonalem und extrapersonalem Raum neuropsychologisch nicht eindeutig bestimmt. Caggiano, Fogassi, Rizzolatti, Thier und Casile (2009) zum Beispiel charakterisieren den peripersonalen Raum als den räumlichen Bereich, der durch die maximale Reichweite beider Arme begrenzt wird. Longo und Lourenco (2007) merken zwar an, wie Holmes und Spence (2004) sowie Halligan und Marshall (1991), dass der peripersonale und extrapersonale Raum verschieden repräsentiert werden, stellen aber keine feste Grenze zwischen diesen Räumen fest. Sie finden in ihrer Arbeit vielmehr Hinweise darauf, dass der peripersonale Raum zum einen in Anteile der Armreichweite eingeteilt ist und zum anderen graduell in den extrapersonalen Raum übergeht (Longo und Lourenco, 2006). Die Autoren heben somit eine weniger starre Definition der räumlichen Ausmaße des peripersonalen Raums hervor. Dieser ändert sich ihnen zufolge mit Bewegungen des Torsos und mit der Verwendung von Werkzeugen.

Plastizität des peripersonalen Raums und des Körperschemas

Ein Faktor, der bei der Größe des peripersonalen Raums berücksichtigt werden muss, ist das Körperwachstum. Werden bei Menschen oder Affen die Arme über Jahre der körperlichen Entwicklung länger, so wird auch die Greifreichweite größer. Die Repräsentationen des Körpers und des Raums müssen sich an diese Umstände anpassen können. Maravita, Spence, Sergent und Driver (2002) zeigten, dass visuelle und taktile Stimuli innerhalb kürzester Zeit den eigenen Körperteilen zugerechnet und dem visuo-taktilen peripersonalen Raum um die Hand zugeordnet werden, obwohl sie zuvor im

extrapersonalen Raum wahrgenommen wurden. Durch sensomotorische Manipulationen konnte bei Menschen gezeigt werden, dass innerhalb kurzer Zeit die Wahrnehmung des peripersonalen Raums beeinflusst wird, sodass selbst entfernt gesehene Körperteile, auch wenn sie nur realistische Attrappen sind, in den Bereich des peripersonalen Raums integriert werden. Russo u. a. (2006) zeigten, dass die Körperrepräsentation sich selbst an rapides Körperwachstum anpassen kann. Die Anpassung geschieht unbewusst und stellt die Informationen über aktuelle Körperausmaße, z. B. Gliedmaßenlängen, sowie die aktuelle Körperpostur bereit. Das Körperschema bildet auf diese Weise eine fortwährend aktualisierte Grundlage für die effektive Planung und Ausführung von Körperbewegungen.

Zusammenhang zwischen peripersonalem Raum und Körperschema

Das Problem, den eigenen Körper ohne Schaden durch seine Umgebung zu führen, ist, wie bisher dargelegt, zum einen abhängig von der peripersonalen Raumrepräsentation, mit der Objektpositionen im Körperraum bestimmt werden. Zum anderen ist es abhängig von der Körperrepräsentation, mit der die räumlichen Ausmaße und der Ort des Körpers bzw. der Körperteile relativ zu Objekten ermittelt werden können. Clark (2007) stellt folgenden Zusammenhang zwischen Nahraum und Körperschema her (S. 272):

„The body schema is a suite of neural settings that implicitly (and non-consciously) define a body in terms of its capabilities for action, for example, by defining the extent of 'near space' for action programs.“

Die Resultate zur Untersuchung von Körper- und Raumrepräsentationen sind nicht immer eindeutig zu interpretieren und in den Neurowissenschaften Gegenstand von Diskussionen. In den Auseinandersetzungen ist eine Sichtweise zu erkennen, die den Zusammenhang beider Repräsentationen betont. Es ist insbesondere ein Wechsel von rein perzeptuellen Modellen des Raums und des Körpers zu einer aktionsgerichteten und aus mehreren Komponenten bestehenden Repräsentation zu beobachten (vgl. Maravita und Iriki, 2004). Die Ähnlichkeit in den Befunden für das Körperschema und für die peripersonale Raumrepräsentation ließ in der Vergangenheit viele Interpretationen zu. Die Plastizität und die Beteiligung an der Motorkontrolle führten immer wieder zu der Vermutung, dass es sich bei den beiden Repräsentationen um überlappende oder

sogar identische Konzepte handelt. Cardinali u. a. (2009) halten die überlappenden Eigenschaften von peripersonalem Raum und Körperschema für starke Hinweise auf die Sicht, dass es sich um dasselbe Konzept handelt. Der derzeitige Stand der Forschung lässt bisher darauf schließen, dass das Körperschema in erster Linie eine Repräsentation des Körpers ist und der peripersonale Raum eine Repräsentation des Raums um diesen Körper. Gestützt durch diese Befunde formulieren wir daher als Arbeitshypothese eine Definition, die das Verhältnis zwischen Körperschema und peripersonaler Raumrepräsentation wiedergibt:

Definition 3 *Eine **peripersonale Raumrepräsentation** kodiert den Raum um den Körper in verschiedenen Körperbezugssystemen. Die Informationen über die Bezugspunkte am Körper und die Ausmaße des peripersonalen Raums werden durch das **Körperschema** geliefert.*

Die Teilprobleme **Wahrnehmung und Repräsentation des Nahraums (P1)** und **Repräsentation des Körpers (P2)** wurden in der Einleitung zunächst aus pragmatischen Gründen aufgeworfen. In Nahraumaufgaben sollte eine Nahraum-Repräsentation es einem humanoiden Agenten möglich machen, wie bei Menschen, auf einfache Weise räumliche Beziehungen zwischen Körper und umgebenden Objekten festzustellen. Eine Körperrepräsentation sollte es ihm zudem möglich machen, zu wissen, wie weit er reichen kann. Die Definition 3 bietet somit nicht nur den zentralen Zusammenhang zwischen den Konzepten des peripersonalen Raums und des Körperschemas, sondern auch zwischen den technischen Teilproblemen P1 und P2.

2.3 Interpersonaler Raum

Nachdem im vorangegangenen Abschnitt die individuellen Körperräume eingeführt wurden, soll im vorliegenden Abschnitt dargelegt werden, wie Räume in interpersonaler Interaktion beschrieben werden können. Zunächst sollen kognitive und neuronale Mechanismen dargestellt werden, die bei der Bildung und Repräsentation interpersonaler Räume möglicherweise eine Rolle spielen. Im Anschluss daran werden existierende Modelle zur Beschreibung interpersonaler Räume vorgestellt. Hierbei wird deutlich, dass zwischen den Befunden zu den kognitiven Mechanismen und den entwickelten Verhaltensmodellen noch wenige Verknüpfungen bestehen. Die Arbeiten dieses Ab-

schnitts dienen in späteren Kapiteln als Grundlage zur Lösung des Teilproblems des **intelligenten räumlichen Verhaltens im Nah- und Interaktionsraum (P3)**.

2.3.1 Repräsentation interpersonaler Räume

In Nahraum-Interaktionen sind wir nicht nur in der Lage, über unsere peripersonale Raumrepräsentation festzustellen, wo Objekte um uns herum relativ zu unserem Körper stehen und ob sie in Greifnähe sind. Wir nutzen auch Wissen darüber, ob unser Interaktionspartner dieselben Objekte erreichen kann oder ob diese aus dessen Sicht verdeckt sind.

Welche mentalen Prozesse die Grundlage solcher Fähigkeiten sind, ist noch nicht vollständig geklärt. Diese Art der Perspektivenübernahme erfordert neben höheren und bewussten Schlussfolgerungsprozessen auch kognitive Prozesse auf niedrigerer Ebene, die automatisiert ablaufen. Solch ein Prozess könnte die räumliche Perspektivenübernahme (engl. spatial perspective taking, SPT) sein. Kessler und Thomson (2010) fanden in ihrer Untersuchung Hinweise darauf, dass SPT in körperlichen Repräsentationen verwurzelt sein könnte und Teile des Körperschemas benötigt, um mentale Rotationen durchführen zu können. In ihrer Studie wurden Probanden an einem Bildschirm Bilder von einer Person gezeigt, die an verschiedenen Positionen eines runden Tisches saß. Die Personen auf den Bildern hatten dabei Körperorientierungen, die von denen der Probanden abwichen. Die Probanden sollten so schnell wie möglich, also ohne lange nachzudenken, die Körperorientierung der Person auf dem Bild einnehmen. Die Probanden führen in dieser Aufgabe eine mentale Körperrotation, also eine imaginierte Körperbewegung durch. Je ähnlicher die Körperorientierung der Person auf dem Bild der des Probanden war, desto schneller wurde die Aufgabe erfüllt. Die mentale Körperrotation fiel den Probanden in diesen Fällen leichter. Die Autoren sehen SPT daher als einen verkörperten kognitiven Prozess an (engl. embodied cognitive process).

Dieser *Verkörperungseffekt* (engl. embodiment effect) hängt möglicherweise mit einem Mechanismus zusammen, den Gallese (2005) auch als **verkörperte Simulation** (engl. embodied simulation) bezeichnet. Dahinter steckt der Gedanke, dass bei Menschen die eigenen sensorischen Systeme und die eigene Körperrepräsentation verwendet wird, um gerichtete Körperbewegungen anderer, also deren Bewegungen und Aktionen in der Welt, zu verstehen. Diese verkörperte Simulation ist im Forschungsfeld der Philosophie des Geistes den Simulationstheorien zuzuordnen. Es handelt sich dabei um

Wahrnehmungstheorien darüber, welche kognitiven Mechanismen dafür verantwortlich sind, dass Menschen Handlungen, Intentionen sowie Emotionen anderer verstehen können. Die Simulationstheorien stehen dabei den Theorie-Theorien gegenüber. Hinter diesen steckt der Gedanke, dass die Wahrnehmung von anderen Menschen mit der Zuschreibung mentaler Zustände einhergeht. Übertragen auf die verkörperte Simulation würde eine Theorie-Theorie postulieren, dass Menschen auf Basis von Schlussfolgerungen und der individuellen körperlichen Konfiguration des Partners eine Theorie über dessen Körperbewegungen und Intentionen bilden, um dessen Aktionen zu verstehen. Gallese sieht als Beleg für die Existenz der verkörperten Simulation die sogenannten Spiegelneurone. Diese Art der Neurone hat die Eigenschaft, nicht nur zu feuern, wenn ein Mensch bestimmte Aktionen ausführt, zum Beispiel beim Greifen eines Objektes, sondern ebenfalls, wenn ein anderer Mensch bei dieser Aktion lediglich beobachtet wird (Rizzolatti, Fadiga, Matelli, Bettinardi, Paulesu, Perani und Fazio, 1996). Zahlreiche Forscher sehen seit der Entdeckung der Spiegelneurone Hinweise auf Simulationsmechanismen, die es Menschen ermöglichen, Handlungen anderer zu verstehen. Die verkörperte Simulation in interpersonaler räumlicher Interaktion ist noch kaum erforscht (Lloyd, 2009). Kessler und Thomson (2008) nehmen an, dass in der Evolution die körperliche Angleichung räumlicher Perspektiven ursprünglich die Grundlage für Angleichungen von sozio-kognitiven Perspektiven gewesen sein könnten. Lloyd (2009) hält den Raum, der sich zwischen Partnern befindet, für wichtig bei der Bildung und Definition sozialer Dynamiken.

2.3.2 Interpersonale Raummodelle

Vor dem Hintergrund, dass die neurowissenschaftliche Analyse des peripersonalen Raums und die Forschung an interpersonalem Verhalten üblicherweise isoliert voneinander betrachtet werden, schlägt Lloyd (2009) ein neurophilosophisches Framework vor, das die beiden Felder miteinander verbindet. Das Ziel soll sein, die neuronalen Mechanismen des sozialen Raums ('neural mechanisms of social space', Lloyd, 2009, S. 298) zu untersuchen und zu interpretieren. In ihrer Hypothese geht die Autorin davon aus, dass die Mechanismen, die erklären, wie unbelebte Objekte den Raum um den Körper beeinflussen, auf Interaktionen zwischen menschlichen Partnern übertragen werden können. Existierende Modelle interpersonaler Räume, also Räume, die sich zwischen (Interaktions-)Partnern befinden, wurden hauptsächlich im Hinblick auf soziale

Funktionen und Verhaltensweisen und weniger auf die darunterliegenden kognitiven Prozesse entwickelt. Ein prominentes Modell zur Beschreibung interpersonaler Räume ist das **Proxemikmodell** (engl. proxemics) von Hall (1969). Es beschreibt räumliche Distanzen, die Menschen in Interaktion, teilweise unbewusst, zueinander einhalten. Die einzelnen Distanzen können kulturell bestimmte Funktionen tragen und unterschiedliche soziale Signale aussenden. Halls Modell beinhaltet vier Distanzen: intime, persönliche, soziale und öffentliche Distanz (engl. intimate, personal, social, public). Er unterteilt jede einzelne Distanz zusätzlich in eine nahe und eine ferne. Die Distanzen beginnen direkt am Körper und gehen von wenigen Zoll-Distanzen (6–18 Zoll, ca. 15–45 cm) über in Distanzen, die mehr als 25 Fuß (ca. 7,5 Meter) umfassen. Der peripersonale Raum fällt ungefähr in den Bereich des nahen persönlichen Raums (ca. 2,5 Fuß bzw. 76 cm). Diesen Bereich bezeichnet Hall auch als Tast-Distanz (engl. touching distance). Der Wert von ungefähr 76 cm entspricht etwa dem Bereich, den auch Holmes und Spence (2004) für den peripersonalen Raum bei Menschen annehmen (vgl. Abschnitt 2.2.1). Halls Proxemikmodell bildet interpersonale Distanzen auf menschliche soziale Beziehungen ab, hat aber nicht zum Ziel, die kognitive Relevanz dieser Räume zu analysieren. Ein weiteres bekanntes Modell ist das **F-Formationen-System** (engl. F-formation system) von Kendon (1990). Das Modell beschreibt die Körperorientierung und die Gruppierung von miteinander interagierenden Menschen. Eine F-Formation (alternativ: facing-formation, S. 249) entsteht, wenn zwei oder mehrere Personen sich auf eine gemeinsame Interaktion konzentrieren, zum Beispiel wenn sie ein Gespräch führen. Kendon definiert:

„An F-formation arises whenever two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct, and exclusive access.“ (S. 209)

Drei F-Formationen, die dabei typischerweise eingenommen werden, sind die *vis-a-vis*-, *L*- und *side-by-side*-Anordnung (siehe Abbildung 2.2). Kendon führt einen Aktivitätsraum (engl. activity space) ein, der sich vor einzelnen Interaktionspartnern befindet und den er als *transaktionales Segment* (engl. transactional segment) bezeichnet. Dieser Raum entspricht örtlich in etwa dem peripersonalen Raum. Bei räumlichen Anordnungen, in denen diese transaktionalen Segmente überlappen, entsteht eine gemeinsame räumliche Region: der *O-Raum* (engl. o-space). Kendon führt zwar den O-Raum und das transaktionale Segment ein, geht aber nicht im Detail darauf ein.

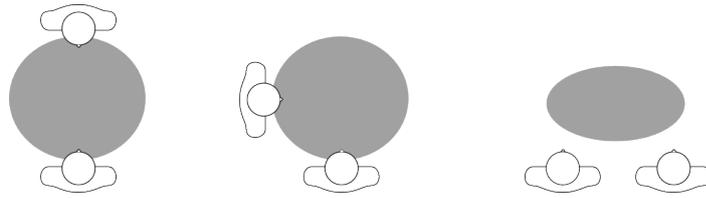


Abbildung 2.2: Kendons F-Formationen mit O-Räumen (grau unterlegt). Anordnungen v.l.n.r.: *vis-a-vis*, *L*- und *side-by-side*.

Sowohl Halls als auch Kendons Raummodelle betrachten die Räume zwischen Interaktionspartnern. In beiden Modellen wird der individuelle Aktionsraum gesondert konzipiert, und er weist Parallelen zu den Funktionen des peripersonalen Raums auf. Diese Modelle können geben auf die räumliche Struktur und die Konsequenzen für die Interaktion nützliche Hinweise geben, wenn die peripersonalen Räume zweier Interaktionspartner sich überlappen.

2.4 Zusammenfassung

In diesem Kapitel wurden theoretische Grundlagen des individuellen Körperraums und der interpersonalen Räume bei Menschen vorgestellt. Für die individuellen Körperräume wurden zunächst verschiedene Begriffe zur Charakterisierung von Referenzsystemen für Raumrepräsentationen eingeführt und ein möglicher Mechanismus zur Transformation zwischen diesen Referenzsystemen dargelegt. Anschließend wurden drei Raummodelle für nahe und ferne Distanzen um den menschlichen Körper vorgestellt, in denen körperzentrierte Referenzsysteme definiert werden. Diese gaben Aufschluss darüber, wie der Raum um Menschen herum strukturiert und mit welchen Eigenschaften und Funktionen Unterräume verknüpft sein könnten. Die für den individuellen Körperraum zentralen Konzepte des peripersonalen Raums und des Körperschemas wurden durch Studien aus verschiedenen Disziplinen vorgestellt und ihr Zusammenhang erläutert. Auf Grundlage der Befunde wurden drei Arbeitsdefinitionen vorgeschlagen, die die für diese Arbeit wichtigsten Charakteristiken zusammenfassen. Die Arbeitsdefinitionen bilden den theoretischen Ausgangspunkt für die Folgekapitel.

Im Gegensatz zum individuellen Raum existieren wenige Befunde darüber, wie Menschen den interpersonalen Raum repräsentieren und welche kognitiven Mechanismen darunterliegen. Die mutmaßlichen Mechanismen, die für diese Dissertation relevant

sind, sind die verkörperte Simulation und die räumliche Perspektivenübernahme. Weit besser erforscht sind Raummodelle, die das soziale Verhalten von Menschen im interpersonalem Raum abbilden. Als bekannteste Vertreter dieser Modelle wurden das Proxemikmodell von Hall und das F-Formationen-Modell von Kendon beschrieben.

3 Humanoide Agenten in Nahrauminteraktion – der Stand der Forschung

Wenn humanoide Agenten effektiv an Nahraumaufgaben teilnehmen sollen, so müssen die in Abschnitt 1.3 identifizierten Teilprobleme durch eine Reihe technischer Aufgaben gelöst werden. Zu den Teilproblemen zählen die **Wahrnehmung und Repräsentation des Nahraums (P1)**, die **Repräsentation des Körpers (P2)** sowie ein **intelligentes räumliches Verhalten im Nah- und Interaktionsraum (P3)**.

Dieses Kapitel gibt daher einen Überblick über die relevanten Arbeiten, die mit der Lösung der Teilprobleme zusammenhängen. Es werden zunächst in Abschnitt 3.1 die technischen Herausforderungen dargestellt, die in der Bewegungs- und Verhaltensmodellierung humanoider Agenten immer wiederkehren. Der Stand der dazugehörigen Lösungsmethoden wird ebenfalls aufgearbeitet. Der Abschnitt legt damit Grundlagen für Entscheidungen, die das Teilproblem P3 betreffen. Anschließend werden die Arbeiten vorgestellt, die für die Teilprobleme P1 und P2 relevant sind. Diese zeigen den Forschungsstand bei der technischen Umsetzung der in Kapitel 2 eingeführten theoretischen Konzepte. Dafür werden Arbeiten diskutiert, in denen Körperrepräsentationen (Abschnitt 3.2.2) sowie der individuelle (Abschnitt 3.2.1) und der interpersonale Raum (Abschnitt 3.3) für humanoide Agenten technisch realisiert wurden. Zum Abschluss werden Nahrauminteraktionsszenarien (Abschnitt 3.4) vorgestellt, in denen die Herausforderungen bezüglich des Teilproblems P3 herausgestellt werden.

3.1 Bewegungsplanung und -kontrolle

Der individuelle Körperraum eines humanoiden Agenten wird sowohl bei Robotern als auch bei virtuellen Menschen als eine gesondert zu untersuchende Region betrachtet. Dieser Raum ist nicht mit dem peripersonalen Raum des neuropsychologischen Kon-

zepts gleichzusetzen. Es ist nicht explizit ein Raum in bestimmter Distanz um den Torso gemeint, der durch verschiedene Sensormodalitäten abgedeckt wird, sondern der Raum, den ein Roboterarm mit Greifern oder künstlichen Händen, den sogenannten **Endeffektoren**, erreichen kann. Der Raum wird auch als **Arbeitsraum** (engl. workspace) bezeichnet. Die **Bewegungsplanung** (engl. motion planning) ist in der Robotik ein zentrales Problem. Damit ist die Fähigkeit eines technischen Systems gemeint, seine Bewegungen automatisch durch Berechnung zu planen. Die geplante Bewegung wird dann durch die **Bewegungskontrolle** erzeugt. In dieser Arbeit bezieht sich die Bewegungsplanung auf die Arme eines humanoiden Agenten im Nahraum bzw. in seinem Arbeitsraum. Fortbewegungsplanung wird nicht betrachtet, da sie sich von dem Problem der Bewegungsplanung in Nahrauminteraktion unterscheidet: Die Kontrolle einzelner Gelenkmanipulatoren bringt andere Herausforderungen mit sich als die Navigation eines mobilen Roboters durch ein weitläufiges Gelände. Detaillierte Angaben zu den verwendeten Notationen und Konventionen aus dem Bereich der Robotik finden sich in Spong, Hutchinson und Vidyasagar (2005). Für die Bewegungskontrolle virtueller Menschen, wie sie in dieser Arbeit durchgeführt wird, bildet die Arbeit von Kopp eine wichtige Grundlage. Da der Ursprung des Problems der Bewegungsplanung in der Robotik liegt, finden sich seine Lösungsmethoden auch in der Computeranimation von virtuellen Menschen wieder. Aus diesem Einsatz entwickelten sich darüber hinaus eigenständige Techniken für Anwendungen in interaktiven Computerspielen oder auch für sogenannte Off-Line-Medienproduktionen, wie zum Beispiel computeranimierte Filme. In den folgenden Abschnitten werden daher neben den aus der Robotik bekannten Ansätzen diejenigen Techniken der Bewegungsplanung vorgestellt, die speziell für virtuelle Menschen entwickelt wurden.

3.1.1 Grundlagen zur Bewegungsanimation virtueller Menschen

Ein beweglicher Greifarm muss nicht zwingend dem humanoiden Arm eines Roboters oder eines virtuellen Menschen entsprechen, sondern wird allgemeiner definiert. Die geometrische Struktur eines humanoiden Agenten wird allgemein als Menge von **Segmenten** l_i (engl. Links) und **Gelenken** j_i (engl. Joints) beschrieben. Die üblicherweise starren Segmente sind durch die Gelenke verbunden und dadurch beweglich. Diese miteinander verbundenen Elemente werden auch **Manipulatoren** oder, aufgrund

ihrer Beweglichkeit, als **kinematische Ketten** (engl. kinematic chains) bezeichnet. Die Beschreibung der kinematischen Kette wird für die physische Bewegungskontrolle von humanoiden Robotern, aber auch für die grafische Animation virtueller Menschen verwendet.

In der Computeranimation wird diese als **Skelett-Animation** (engl. skeletal animation) bezeichnet. Gegenüber alternativen Bewegungssteuerungen, wie der sogenannten **Morph-Target-Animation**, müssen bei der Skelett-Animation weniger Charakteristiken des Körpers angesteuert werden. Bei der Skelett-Animation wird manchmal von einer angenäherten Bewegungsanimation gesprochen, da die computergrafische Körpergeometrie durch Diskretisierungen über Gelenke und Segmente angesteuert wird. Im Vergleich dazu ist die Morph-Target-Animation eine genauere Methode, da sie auf der Ebene von computergrafischen Grundelementen, den sogenannten Vertexen, stattfindet. Durch die feinere Auflösung und die größere Kontrollmöglichkeit wird diese Methode vorwiegend in der Gesichtsanimation eingesetzt. Die Genauigkeit dieser Methode geht jedoch auf Kosten eines im Vergleich zur Skelett-Animation höheren Rechenaufwands. Die Skelett-Animation gilt bei der Bewegungskontrolle des restlichen Charakterkörpers, sowohl von virtuellen Menschen als auch von Computerspielcharakteren oder 3-D-Trickfilmfiguren, als Standard-Methode. Diese Charaktere werden auch als **artikulierte Figuren** bezeichnet. Die darunterliegende gelenkige Struktur ist das sogenannte **kinematische Skelett**. Es besteht aus einer teilweise verzweigten Anordnung kinematischer Ketten. Ein **Wurzelgelenk** j_1 verbindet dabei alle kinematischen Ketten miteinander und bestimmt damit die räumliche Position und Orientierung des gesamten kinematischen Skeletts. Im Folgenden legen wir den Fokus auf die Methode der Skelett-Animation. Die hier verglichenen Animationsmethoden werden jedoch in Abschnitt 4.3 wieder aufgegriffen.

Kinematik artikulierter Figuren

Zur Vereinheitlichung von Körpermodellen für artikulierte Figuren ist der sogenannte H-Anim-Standard (Humanoid Animation Standard)¹ weit verbreitet. Dieser beschreibt die Skeletthierarchie und die Bezeichnung einzelner Gelenke und Segmente virtueller humanoider Figuren. Das damit beschriebene kinematische Skelett soll als Grundlage dienen, menschenähnliche Bewegungen für virtuelle Figuren zu modellieren. Die Wahl

¹<http://www.h-anim.org/> ; Letzter Zugriff: 27. Sep 2012

der Segmentlängen und geeigneter bzw. möglicherweise biologisch plausibler Gelenkwinkelgrenzen ist dem Entwickler überlassen.

Kinematische Ketten werden formal als nacheinander ausgeführte homogene **Rotationen** \mathbf{R}_i und homogene **Translationen** \mathbf{T}_i definiert. Die Länge k einer kinematischen Kette wird durch die Anzahl der Segmente, die darin miteinander verbunden sind, bestimmt. Der Index i der Gelenke j_i wird aufsteigend von dem proximalen Gelenk, also dem Anfang der Kette, hin zum distalen, dem Ende der Kette durchnummeriert. Jedem Gelenk j_i ist eine Rotation \mathbf{R}_i und jedem Segment l_i eine Translation \mathbf{T}_i zugeordnet. Jede Translation \mathbf{T}_i beschreibt den Übergang zwischen einem Gelenk j_{i-1} zu einem Gelenk j_i . Wir folgen der Konvention, demnach die Translation in Richtung der z -Achse eines Gelenkes j_i entlang eines Vektors \mathbf{z}_i geht. Das dadurch beschriebene Segment l_i ist somit die Vektordifferenz zwischen den Gelenken j_{i-1} und j_i .

Folgt einem Gelenk kein weiteres, so ist das letzte Element einer kinematischen Kette erreicht. Dieses wird als **Endeffektor** bezeichnet. Bei humanoiden Agenten können es Greifer oder künstliche Hände sein. Der **Arbeitsraum** des Manipulators ist dann definiert als das Gesamtvolumen, das durch den Endeffektor aufgespannt wird, wenn er alle ihm möglichen Bewegungen ausführt. Die **Konfiguration** eines Manipulators ist die aktuelle Beschreibung aller Positionen und Orientierungen seiner Gelenke und Segmente. Die Menge aller Konfigurationen eines Manipulators heißt **Konfigurationsraum**.

Die Aufgabe der **Bewegungsplanung** ist es nun, von der gegebenen Ausgangsstellung eines Armes bzw. des Endeffektors einen Weg zu einer gegebenen Zielstellung zu finden. Es wird also ein Weg von einer bestimmten Ausgangs- zu einer Zielkonfiguration eines Manipulators gesucht. Der Weg besteht aus einer kontinuierlichen Folge von Zwischenkonfigurationen. Bei der Planung muss zudem festgestellt werden, ob Hindernisse auf dem Weg liegen und ob diese durch einen alternativen Weg umgangen werden können bzw. ob ein geeigneter Weg zur Zielkonfiguration existiert. Die Aufgabe der **Motorkontrolle** ist es, solch eine artikulierte Figur in Bewegung zu setzen. Dabei werden die Gelenkrotationen \mathbf{R}_i verändert, die zu Veränderungen der miteinander verbundenen Segmentpositionen führen. Dadurch kann ein Endeffektor in eine, zum Beispiel durch die Bewegungsplanung vorgegebene, gewünschte Position gebracht werden. Gelenkrotationen werden häufig als Rotationswinkel bezüglich festgelegter Rotationsachsen notiert. Hier wird vorausgesetzt, dass in einem Gelenk j_i drei Rotationsachsen liegen, die orthogonal zueinander sind und sich in einem Punkt im Gelenk schneiden.

Die Konfiguration des Gelenkes j_i kann dann durch eine Rotationsmatrix beschrieben werden. Bei der Darstellung ist zu beachten, dass es verschiedene Reihenfolgen für die Rotationen um die drei Achsen gibt. Eine übliche Art der Beschreibung stellen die **Eulerschen Winkel** dar, bei der eine Konvention gewählt werden kann, die festlegt, wie die drei Winkel (meist ϕ, θ, ψ) nacheinander um welche Basisachsen gedreht werden. Jedoch kann auch bei der Festlegung der Achsenreihenfolge keine Eindeutigkeit in der Beschreibung gesichert werden. So ist auch mit mehreren Winkelkombinationen dieselbe Orientierung möglich. Um die Position eines Gelenkes eindeutig zu beschreiben, werden im kartesischen Raum \mathbb{R}^3 sechs voneinander unabhängige Parameter benötigt. Dies sind die drei Winkelwerte für die Rotation und die drei Werte für die Translation. Hat ein Manipulator mehr als sechs solcher sogenannten **Freiheitsgrade** (engl. degrees of freedom bzw. **DoF**), so bezeichnet man ihn als kinematisch **redundant**. Eine weitere Art der Notation für Gelenkrotationen ist die Achsel/Winkel-Notation bei der lediglich ein Rotationswinkel θ und eine Rotationsachse \mathbf{k} zur Beschreibung benötigt wird. Die Umrechnung einer Rotationsmatrix in die Achsen/Winkel-Notation und umgekehrt ist problemlos möglich.

Kinematische Funktionen

Die räumliche Position eines Gelenkes j_i kann im Koordinatensystem eines anderen Gelenkes, also beispielsweise relativ zu einem Wurzelgelenk j_1 , ausgedrückt werden. Koordinatensysteme werden auch mit FoR (engl. Frame of Reference) abgekürzt. Jedes Gelenk beschreibt also ein eigenes FoR. Die Position wird durch eine **Transformation** \mathcal{T} berechnet, in der alle Rotationen und Translationen berücksichtigt werden, die zwischen den beiden Gelenken liegen.

$$\mathcal{T} := \prod_{j=1}^i \mathbf{T}_j \circ \mathbf{R}_j = \mathbf{T}_1 \circ \mathbf{R}_1 \circ \dots \circ \mathbf{T}_i \circ \mathbf{R}_i \quad (3.1)$$

Die räumliche Position eines Gelenkes hängt somit nicht nur vom Wurzelgelenk ab. Wie die Transformation \mathcal{T} zeigt, hängt sie auch von den in der kinematischen Kette vorhergehenden Gelenkpositionen ab. Die Gelenke sind hierarchisch angeordnet, das bedeutet, dass jede Bewegung eines Gelenkes j_i alle Gelenke $j_{i+1..k}$ bewegt, die darunter liegen. Bei den Verzweigungen in kinematischen Skeletten ist zusätzlich zu beachten, dass unterschiedliche Transformationen durchgeführt werden müssen, je

nachdem ob zum Wurzelgelenk hin- oder von diesem wegtransformiert wird. Wandert man sozusagen vom Wurzelgelenk weg in Richtung Endeffektor, so gelten die Transformationen wie in Gleichung 3.1. Wandert man jedoch in die Gegenrichtung, also zum Wurzelgelenk hin, so sind statt der direkten Transformationen \mathbf{T}_i und \mathbf{R}_i die inversen Transformationen \mathbf{T}_i^{-1} und \mathbf{R}_i^{-1} zu berücksichtigen. Bei den meisten Manipulatoren ändert sich die Segmentlänge nicht, sodass die Werte für die Translation normalerweise bereits festgelegt und bekannt sind. Somit wird die Konfiguration eines Manipulators durch die Anzahl der Parameter zur Beschreibung der Rotationen bestimmt. Bei k Gelenken in einer kinematischen Kette wären es also $3k$ Freiheitsgrade. Man sagt auch, eine kinematische Kette hat n Freiheitsgrade, wenn ihre Konfiguration minimal mit n Parametern beschrieben werden kann.

Die momentane Konfiguration eines Manipulators mit n DoF kann durch die Menge aller Gelenkwinkel beschrieben werden. Diese Werte q_1, \dots, q_n mit $q_i = \theta_i$ können als Elemente eines Vektors \mathbf{q} notiert werden. Die Berechnung der Endeffektorposition im Koordinatensystem eines Basisgelenks bei gegebener Manipulatorkonfiguration \mathbf{q} wird als **Vorwärtskinematik** bezeichnet. Es handelt sich also um eine Transformation F dieser Form: $F(\mathbf{q}) = \mathbf{v}$. Bei gegebener Konfiguration q_1, \dots, q_n eines jeden Gelenks und der Geometrie des Manipulators (d. h. bei gegebenen Segmentlängen) ist es also problemlos möglich, die Position \mathbf{v} eines Endeffektors zu ermitteln. Die umgekehrte Problemstellung zur Vorwärtskinematik lautet **inverse Kinematik** (IK). Diese Transformation versucht eine Konfiguration \mathbf{q} zu finden, bei der eine gegebene bzw. gewünschte Endeffektorposition \mathbf{v} erreicht werden soll. Sie ist definiert durch: $F^{-1}(\mathbf{v}) = \mathbf{q}$. Die Aufgabe, für eine vorgegebene Endeffektorposition geeignete Gelenkwinkelrotationen zu finden, ist zentral bei der Bewegungskontrolle, gestaltet sich jedoch schwieriger als das Problem der Vorwärtskinematik. Bei redundanten kinematischen Ketten gibt es üblicherweise keine eindeutige Lösung der inversen Kinematik. Verschiedene Gelenkwinkelrotationen führen dann zu derselben Lösung. Bei zusätzlichen Einschränkungen (engl. constraints) wie Kollisionsvermeidung oder dem Einhalten vorgegebener Kostenfunktionen kann dies von Vorteil sein. Es kann dann im Lösungsraum nach einer Lösung gesucht werden, die alle Bedingungen erfüllt.

3.1.2 Ansätze zur Armbewegungsplanung

Für virtuelle Charaktere und physische Roboter wurden zahlreiche Arbeiten präsentiert, die sich den verschiedenen Bewegungen und ihrer Planung annehmen. In diesem Abschnitt soll der Fokus auf Armbewegungen gelegt werden, die speziell für Nahraumaufgaben relevant sind. Es folgt zunächst ein Überblick über existierende Verfahren zur *Armbewegungsplanung*, die speziell für virtuelle Menschen geeignet sind. In dieser Arbeit wird kein solches Verfahren entwickelt. Es soll vielmehr die Frage geklärt werden, welches Verfahren für die zu generierenden Bewegungen in Nahraumaufgaben geeignet ist. Unter derselben Fragestellung werden daher in dem folgenden Unterabschnitt Methoden zur *Hindernisvermeidung* diskutiert. Diese spielt insbesondere dann eine Rolle, wenn ein humanoider Agent sich einen Interaktionsraum mit einem Partner teilt. Die relevanteste Armbewegung in Nahrauminteraktion, die *Umplatzierbewegung*, wird in dem darauffolgenden Unterabschnitt definiert, zudem wird ihre Planung erläutert. Alle hier vorgestellten Ansätze beinhalten wichtige Überlegungen zur Lösung des Teilproblems des **intelligenten räumlichen Verhaltens im Nah- und Interaktionsraum (P3)**.

Ansätze zur Arm-Animation in der Computergrafik

Pettré, Kallmann und Lin (2008) geben einen Überblick über die Vielzahl an Methoden in der Bewegungsplanung mit speziellem Fokus auf Forschungsarbeiten zu virtuellen Charakteren bzw. autonomen virtuellen Menschen. In der Computeranimation unterscheidet Kallmann (2008) verschiedene Ansätze zur *Bewegungsplanung bei Objektmanipulationen*. Im Folgenden werden die zwei für diese Arbeit relevantesten verglichen: die globale Bewegungsplanung und die inverse Kinematik.

Bei der **globalen Bewegungsplanung** wird der gesamte Konfigurationsraum des Arms exploriert. Auf diese Weise können auch komplexe Lösungen gefunden werden. In niedrig-dimensionalen Räumen können Methoden zum Einsatz kommen, die den gesamten Raum in Zellen zerlegen und über eine Graph-Suche nach einer Lösung suchen. In hochdimensionalen Räumen bieten sich die auf der Sampling-Methode basierenden Ansätze an. Diese gelten als aktuellster Stand der Technik. Bei der Methode wird eine Karte aller gültigen Konfigurationen aufgebaut. Einzelne Konfigurationen werden daraufhin getestet, ob sie bestimmte Randbedingungen erfüllen, zum Beispiel kollisionsfrei sind. Die Methode gilt als wahrscheinlichkeits-vollständig. Das bedeu-

tet, dass die Wahrscheinlichkeit, dass eine Lösung gefunden wird, mit der Anzahl der Rechenschritte steigt. Ein Nachteil ist, dass die Methode nicht ermitteln kann, ob keine Lösung existiert. Auf der Sampling-Methode basierende Bewegungsplaner werden sowohl in der Robotik für Roboterarme (Behnisch, Haschke und Gienger, 2010; Berenson, Srinivasa, Ferguson, Collet und Kuffner, 2009) als auch für Objektmanipulationen virtueller Menschen (Shapiro, Kallmann und Faloutsos, 2007) verwendet.

Bewegungsplaner, die auf der **inversen Kinematik** beruhen, werden verwendet, um eine präzise Kontrolle der kinematischen Ketten zu erreichen. IK-basierte Planer gelten als die etabliertesten und die am häufigsten verwendeten. Da sie jedoch lokale Planer sind, können sie keine globale Suche nach einer kollisionsfreien Lösung durchführen. IK-Lösungsverfahren unterteilen sich in **analytische** und **iterative** Methoden. Die analytischen Lösungsansätze sind bei nicht redundanten bzw. mit zusätzlichen Einschränkungen, bei wenig redundanten kinematischen Ketten anwendbar. Die Gelenkwinkelrotationen werden dabei direkt ermittelt, das bedeutet, dass direkt versucht wird, die Funktion F^{-1} zu finden. Bei höher redundanten kinematischen Ketten sind jedoch iterative Methoden unerlässlich. Diese numerischen Verfahren konvergieren dann iterativ gegen eine Lösung. Im Berechnungsschritt der Bewegungsplanung besteht das übliche Vorgehen darin, zunächst eine Trajektorie, also einen vordefinierten Bewegungspfad, für den Endeffektor zu planen. Es wird dann iterativ versucht, die Gelenkwinkelrotationen zu finden, die den Endeffektor von einem Trajektorienpunkt zum nächsten bringen. Der Nachteil dieser Methode ist jedoch der hohe Berechnungsaufwand. Unter den IK-Lösungsansätzen stellt das *IKAN-Modell* (Tolani, Goswami und Badler, 2000) ein analytisch-hybrides Verfahren für kinematische Ketten zur Verfügung, die mit 7 DoF parametrisiert sind. Diese Art der Kette entspricht bei den meisten humanoiden Agenten dem Arm mit drei Gelenken. Bei diesem Verfahren werden einzelne Teilprobleme analytisch gelöst, und dann wird mit iterativer Suche versucht, die Komplettlösung zu finden.

Ansätze zur Hindernisvermeidung in der Bewegungsplanung

Für die Bewegungsplanung mit Hindernisvermeidung unterscheidet Latombe (1991) *globale* und *lokale* Lösungsansätze. Methoden die auf sogenannten Roadmaps und Zell-Zerlegung basieren sind globale Verfahren, da sie den gesamten Konfigurationsraum in zerlegen und durchsuchen. Ansätze, die Potentialfelder verwenden, sind wiederum

lokale Methoden, da nur die Konfigurationen des Roboters bzw. virtuellen Agenten berücksichtigen, die in der Nachbarschaft liegen. Das Prinzip der künstlichen Potentialfelder (Khatib, 1986) funktioniert über vordefinierte Kraftfeldfunktionen, die einen Punkt im Raum, zum Beispiel einen Roboter als Massepunkt betrachtet, beeinflussen. Von Hindernissen wirken abstoßende, von einem Zielpunkt wirken anziehende Kräfte. Der Roboter sucht dann einen Pfad durch die Kraftfelder, der den Abstand zu den Hindernissen maximiert und gleichzeitig die Entfernung zum Ziel minimiert. Stilman, Schamburek, Kuffner und Asfour (2007) präsentierten einen Ansatz zur Manipulationsplanung unter Berücksichtigung beweglicher Hindernisse für physische Roboterarme. Das Verfahren basiert auf der im vorherigen Abschnitt beschriebenen Sampling-Methode und ist daher als global einzustufen. Liu (2003) präsentierte eine Methode für virtuelle Charaktere, bei der die Kollisionsvermeidung bei Greifbewegungen über einen Potentialfeld-Ansatz geschieht. Die Planung der Bewegungen finden dort auf Grundlage des IKAN-Modells statt und ist daher als lokal einzuordnen.

Planung von Objektmanipulationsbewegungen

Für Nahrauminteraktionen, wie wir sie in dieser Arbeit behandeln, ist die häufigste Objektmanipulation mit den Armen das **Umplatzieren** von Objekten. Daher wird im Folgenden die Umplatzierbewegung definiert. Für das Umplatzieren von Objekten müssen folgende Teilbewegungen geplant werden: 1. *Reichen* (engl. reaching), 2. *Greifen* (engl. grasping), 3. *Umplatzieren* (engl. relocation) und 4. *Loslösen* (engl. release).

Das Planen des Ablaufes der einzelnen Teilbewegungen wird auch als **Sequenzierungsplanung** (engl. sequencing planning, vgl. Kallmann, 2008) bezeichnet. Welche Körperteile dafür eingesetzt werden müssen, ist abhängig von der Entfernung des Objekts. So stellen Huang, Eliëns und Visser (2003) eine Arbeit vor, bei der drei verschiedene Strategien für Reich-Bewegungen eines virtuellen Charakters modelliert werden. Je nachdem wo sich das zu erreichende Objekt befindet, werden verschiedene Aktionen durchgeführt und gegebenenfalls kombiniert. Handbewegungen werden als *Reach by hand*-Aktionen und Torsobewegungen als *Reach by body*-Aktionen bezeichnet und über eine Entscheidungskomponente gelöst. Fortbewegungen im Raum sind sogenannte *Reach by move*-Aktionen und werden als Suchproblem behandelt. Bei virtuellen Menschen kann das Reichen und Umplatzieren gleich behandelt werden. Das umzuplatzierende Objekt wird nach Ablauf der Reich-Bewegung an den Endeffektor

angehängt. Die Generierung glaubhafter und natürlicher Hand-Greifbewegungen ist eine anspruchsvolle Aufgabe, die sowohl bei virtuellen Menschen, als auch bei Robotern (Röthling, Haschke, Steil und Ritter, 2007) untersucht wird. Verschiedene Faktoren beeinflussen die Bewegung. Das Greifen mit der Hand kann dabei von der Form und dem Material des Objektes abhängig sein, das Loslösen von der Fläche, auf die das Objekt gestellt werden soll. Diese Bewegung wird hier nicht im Detail behandelt. Es wird stattdessen eine vordefinierte Greif- und Loslösebewegung vorausgesetzt. Besonders die von Huang u. a. (2003) beschriebene *Reach by body*-Aktion ist in Nahraumaufgaben zu berücksichtigen.

3.1.3 Ansätze zur Modulation der Bewegungsqualität

In Abschnitt 2.2.3 wurden Studien vorgestellt, die zeigen, dass Menschen in der Lage sind, über die Art, wie eine Greifbewegung durchgeführt wird, die soziale Intention dahinter verstehen. Dafür sind bestimmte kinematische Eigenschaften der Armbewegungen verantwortlich. Durch die technische Modulierung solcher Eigenschaften ist es möglich, Bewegungen für humanoide Agenten zu generieren, die eine bestimmte Bedeutung tragen. Chi, Costa, Zhao und Badler (2000) entwickelten das EMOTE-System (**E**xpressive **M**OTION **E**ngine), mit dem es möglich ist, die Bewegungsqualität von Gesten virtueller Charaktere zu modulieren. Es werden zwei Komponenten der *Laban* Bewegungsanalyse zur qualitativen Beschreibung von Bewegungen, variiert. Zum einen die Anstrengung (engl. effort) und zum anderen die Form (engl. shape) einer Geste. Basierend auf diesen Parametern können Bewegungsanimationen für virtuelle Charaktere generiert werden, die auf Betrachter natürlicher, expressiver und glaubhafter wirken. Cassell, Vilhjálmsón und Bickmore (2001) stellten mit dem BEAT-System (**B**ehavior **E**xpression **A**nimation **T**oolkit) ebenfalls ein Software-Werkzeug zur Generierung expressiver Verhaltensweisen für virtuelle Charaktere vor. Bei diesem System liegt der Fokus auf der Erzeugung von nonverbalem Verhalten zu Text, der von der virtuellen Figur verbal geäußert wird. Das nonverbale Verhalten wird aus linguistischen Informationen sowie Kontextwissen gezogen. Es werden auf dieser Grundlage Blickbewegungen, Gesten und weitere Verhaltensweisen erzeugt, die für menschliches, konversationales Verhalten angemessen sind. Breazeal und Fitzpatrick (2000) beschreiben in ihrer Arbeit, wie der Roboter Kismet durch das Zeigen verschiedener, nonverbaler Verhaltensweisen Beobachtern den Eindruck vermittelt, soziale Signale zu geben. So

beobachten die Autoren, dass das Vor- und Zurücklehnen des Kopfes, abhängig von der räumlichen Distanz des menschlichen Interaktionspartners, bei diesem den Eindruck vermitteln konnte, der Roboter hätte die Absicht, mit dem Menschen in Austausch zu treten. Diesen Effekt bezeichnen die Autoren als soziale Verstärkung (engl. social amplification). Für künstliche, verkörperte Agenten ist dieser Effekt ein wirksames Mittel, um über einfache Motoraktionen, Verhalten zu zeigen, von dem menschliche Partner intuitiv und auf natürliche Weise Intentionen ablesen.

3.1.4 Resümee

Ansätze zur Arm-Animation: Für die **Bewegungsplanung** der humanoiden Charaktere, mit denen sich diese Arbeit beschäftigt, wird die IKAN-Methode als die sinnvollste angesehen. Diese Methode gilt gegenüber anderen IK-Methoden als besonders performant in Echtzeitumgebungen (Liu, 2003). Die Bewegungen des in dieser Arbeit eingesetzten virtuellen Menschen Max werden mit der Softwarebibliothek *Articulated Communicator Engine* (ACE) kontrolliert (Kopp und Wachsmuth, 2004). Die ACE-Bibliothek dient zur Generierung multimodaler Äußerungen für virtuelle Menschen und verwendet für die Motorplanung die IKAN-Softwarebibliothek. Da zeitliche Performanz in Nahrauminteraktion von besonders großer Bedeutung ist, wird die bestehende Motorplanung mit der IKAN-Methode weiterhin eingesetzt. Mit dieser Entscheidung ist der Einsatz von Potentialfeldern besonders geeignet, um Bewegungsplanung bei räumlichen Einschränkungen zu realisieren. Die Methode eignet sich zudem für ein Nahraummodell eines Agenten sowie für den Interaktionsraum, da Potentialfelder für einzelne Raumbereiche definiert werden können.

Planung von Objektmanipulationsbewegungen: Bei Bewegungen des Umplatziens von entfernteren Objekten, sind nicht nur die Arme, sondern auch der Torso nötig. Das Vorlehnen, um ein Objekt zu erreichen, wird in dieser Arbeit daher im Zusammenhang mit der Bewegungsplanung ebenfalls behandelt. Relevante Befunde, wie sie bei Menschen beobachtet wurden, wie die in Abschnitt 2.1.3 vorgestellten Raummodelle und das Vorlehnverhalten bei Greifbewegungen aus Abschnitt 2.2.3, bieten dabei nützliche Hinweise zur technischen Modellierung. Die in Abschnitt 3.1.3 beschriebene *Modulation der Bewegungsqualität* wird in dieser Dissertation für die Bewegungsmerkmale des Umplatziens realisiert. Das Vermitteln von sozialen Signalen über die Armbewegungen ist besonders für die Modellierung des Verhaltens eines humanoiden Agenten im

Interaktionsraum sinnvoll.

3.2 Raum- und Körperrepräsentationen

Um humanoide Agenten in Nahrauminteraktion mit Fähigkeiten auszustatten, die Menschen mühelos beherrschen, wurden im vorangegangenen Kapitel 2 Hinweise aus nichttechnischen Disziplinen präsentiert, die bei der Lösung der technischen Teilprobleme **Wahrnehmung und Repräsentation des Nahraums (P1)** und **Repräsentation des Körpers (P2)** von Nutzen sein können. Im nächsten Abschnitt diskutieren wir die Modellierung und Repräsentation des Nahraums um einen humanoiden Körper. Anhand einer Analyse alternativer Modellierungsansätze wird eine Entscheidung für eine geeignete Raumrepräsentationsform getroffen. Relevante technische Arbeiten hierzu werden im Anschluss vorgestellt, und es werden Schlussfolgerungen gezogen. In Abschnitt 3.2.2 werden mögliche Formen der Körperrepräsentation verglichen, und es wird ein geeignetes Verfahren zur Generierung einer solchen Körperrepräsentation vorgeschlagen.

3.2.1 Peripersonaler Raum für humanoide Agenten

Soll ein humanoider Agent sich natürlich und intelligent im Nahraum verhalten, so ist die Frage, wie er diesen Raum wahrnehmen und was von diesem Raum repräsentiert werden soll. Es stellt sich hier also die Frage nach einem Raummodell eines humanoiden Agenten, das eine Lösung der Teilprobleme der **Wahrnehmung und Repräsentation des Nahraums (P1)** und des darin gezeigten **intelligenten räumlichen Verhaltens (P3)** bietet. Zwar haben wir mit den theoretischen Befunden aus Kapitel 2 Nahraumdefinitionen 1 und 3 entwickelt, das bedeutet jedoch nicht, dass eine bei Menschen vorzufindende Raumrepräsentation in derselben Form auch für humanoide Agenten sinnvoll ist. Um existierende technische Arbeiten zur Nahraumrepräsentation auswählen und auf für diese Arbeit geeignete Eigenschaften prüfen zu können, wird zunächst eine Voranalyse erstellt.

Voranalyse

Eine Möglichkeit, wie sie in technischen Systemen häufig vorkommt, besteht darin, durch Kameras oder andere technische Mittel eine gesamte Szene, zum Beispiel den

unmittelbaren Raum um den humanoiden Agenten, aus der Vogelperspektive (Gray, Breazeal, Berlin, Brooks und Lieberman, 2005) und teilweise aus weiteren Perspektiven (Rickert, Foster, Giuliani, By, Panin und Knoll, 2007) zu erfassen. Auf diese Weise kann eine Repräsentation aller Objekte erstellt werden, die wir als **allozentrisch** bezeichnen. Diese bietet den Vorteil, dass ein Roboter oder ein virtueller Mensch, der Zugriff auf diese Repräsentation besitzt, Schlussfolgerungen ziehen und Aktionen auf allen Objekten durchführen kann. Ein Nachteil ist, dass es oftmals nur mit erheblichem Aufwand möglich ist, eine umfassende Abdeckung der gesamten Szene zu bewerkstelligen. Bei mobilen Agenten ist diese Art der Szenenerfassung fast unmöglich.

Für intelligente Agenten, wie sie in dieser Dissertation betrachtet werden, ist ein zentrales Merkmal die Ausführung autonomer, also selbstständiger, Aktionen in ihrer Umgebung (siehe Definition nach Wooldridge, 2009, S. 21). Mit Umgebungskameras wäre diese Fähigkeit nicht gegeben. Bei humanoiden Robotern gibt es daher die Möglichkeit, Kameras in den Roboterkörper (z. B. in den Kopf) einzubauen und nur einen Teil des umgebenden Raums zu erfassen. Diese Komplexitätsreduktion stellt jedoch besondere Anforderungen an das Raummodell der unmittelbaren Umgebung und an die Verhaltensstrategien, um effektiv und sinnvoll im Nahraum interagieren zu können. Die *menschenähnliche* Sensorik des humanoiden Agenten muss dann durch verschiedene Sensormodalitäten abgedeckt werden, um eine *robuste* Wahrnehmung zu erzielen. Für die in dieser Arbeit verfolgten Ziele bietet sich daher eine Raumrepräsentation an, die Sensorik und Motorik einbezieht. Dieser in der sogenannten *körperintegrierten Künstlichen Intelligenz* (engl. Embodied Artificial Intelligence) verfolgte Ansatz bietet gegenüber traditionellen allozentrischen Raumrepräsentationen eine flexiblere, robustere und anpassungsfähigere Möglichkeit, den Raum um einen humanoiden Agenten zu modellieren (Butz, 2008). Dieser Ansatz ist besonders gut vereinbar mit dem Konzept des peripersonalen Raums, wie es in Definition 3 gegeben ist. Wie in Abschnitt 2.1.3 eingeführt, können Menschen sich sehr gut in ihrer Umgebung zurechtfinden, auch ohne über eine umfassende Gesamtübersicht über die Umgebung zu verfügen. Unsere sensorische Wahrnehmung des unmittelbaren Raums ist **egozentrisch** repräsentiert und teilweise sogar eingeschränkt. Diese Überlegungen führen zu der Entscheidung, als technische Modellierungshilfe den in der Definition 1 in Abschnitt 2.2.1 vorgestellten **biologischen peripersonalen Raum** zu nutzen. Der biologische peripersonale Raum wurde für technischen Systeme erstmals in der Arbeit von Goerick, Wersing, Mikhailova und Dunn (2005) wahrgenommen, als nützliches Konzept angesehen und

eingesetzt. Unabhängig davon erschienen wenige weitere technische Arbeiten mit diesem Ansatz. Die vorhandenen Modelle sind vorwiegend in der Robotik zu finden. Deutlich ist zu erkennen, dass es noch keine allgemein verbreitete technische Interpretation des peripersonalen Raums gibt. Entsprechend unterschiedlich ist die Umsetzung der interdisziplinären Konzepte. Im Folgenden werden daher Arbeiten vorgestellt und diskutiert, die wichtige Aspekte mit dem in dieser Dissertation entwickelten peripersonalen Raummodell teilen.

Sensorische Voraussetzungen humanoider Agenten

Zur Wahrnehmung des umgebenden Raumes benötigen künstliche Agenten Sensoren. Insbesondere bei humanoiden Agenten ist die Sensorik und Perzeption häufig der menschlichen angelehnt. So wird die visuelle Wahrnehmung bei physischen Roboterplattformen, wie im vorangegangenen Abschnitt beschrieben, üblicherweise über Kamerasysteme im Roboterkörper oder Umgebungskameras gelöst. Abhängig von der Gelenkart werden die Gelenkstellungen als propriozeptive Signale verwendet (für Manipulatoren vgl. einschlägige Robotikliteratur wie Spong u. a. (2005)). Taktile Sensoren, die an die menschliche Haut angelehnt sind, sind bei physischen Robotern aufgrund des technischen Aufwands seltener zu finden. Beispiele sind Roboterhände, die über taktile Sensoren, menschenähnliche Handgreifbewegungen nachmodellieren (Jockusch, Walter und Ritter, 1997; Steffen, Haschke und Ritter, 2007). Ein humanoide Roboter, dessen gesamter Körper wie eine Haut mit taktilen Sensoren bedeckt, wurde für den Roboter CB2 von Minato, Yoshikawa, Noda, Ikemoto, Ishiguro und Asada (2007) vorgestellt. Auch für virtuelle Menschen sind Sensorik-Architekturen vorgestellt worden, die die propriozeptive und visuelle (Conde und Thalmann, 2006) Wahrnehmung beinhalten. Eine virtuelle Haut für den virtuellen Menschen Max wurde von Nguyen (2007) realisiert.

Technische Umsetzung des peripersonalen Raums

Die im Folgenden vorgestellten Arbeiten werden in die Kategorien *explizite* und *implizite* Repräsentation eingeordnet. Explizit bedeutet, dass die Repräsentation Objekteigenschaften, z. B. die Lage im Raum, vorhält. Implizite Repräsentationen lassen es erst durch weitere Verarbeitungsschritte zu, räumliche Objekteigenschaften zu ermitteln. Darüber hinaus werden die Modelle daraufhin geprüft, ob sie *biologisch motiviert*

sind oder ein *technisches System verbessern* sollen.

Hersch (2009) sieht sein Modell der Körperrepräsentation gleichzeitig als Modell des peripersonalen Raums an. Nach unserer Interpretation in Definition 3 fehlt in Herschs Auslegung jedoch die explizite Kodierung des Raums bzw. von Gegenständen. Daher verstehen wir seine Arbeit als Mittel zur Modellierung einer peripersonalen Raumrepräsentation. Auf seine Arbeit wird in Abschnitt 3.2.4 im Kontext des Teilproblems P2 näher eingegangen.

Fuke, Ogino und Asada (2009) präsentieren ein Modell, das es einem Roboter ermöglichen soll, ein kopfzentriertes Koordinatensystem für visuelle Stimuli und eine cross-modale Repräsentation seines Gesichtes zu erlernen. Das Modell stützt sich direkt auf die in Abschnitt 2.2.1 vorgestellten bimodalen Neuronen von Makak-Gehirnen, die der peripersonalen Raumrepräsentation zugerechnet werden. Fuke u. a. versuchen sich mit ihrem funktionalgetreuen Neuronenmodell der anspruchsvollen Aufgabe zu stellen, eine peripersonale Raumrepräsentation zu erlernen. Ihr Modell ist jedoch nur auf den peripersonalen Raum vor dem Gesicht beschränkt und wird nicht explizit als Raum modelliert, der mit den Armen erreichbar ist. Die Repräsentation des peripersonalen Raums ist somit als *implizit* einzuordnen. Die vorgestellte Arbeit bezieht für den peripersonalen Raum wichtige Sensormodalitäten (taktile, visuelle und propriozeptive) mit ein. Das Modell ist als *biologisch motiviert* einzustufen. Der Fokus liegt auf der Erlernbarkeit der bimodalen Neuronenfunktionen, um kognitive und neuropsychologische Prozesse bei Menschen besser verstehen zu können. So bleibt beispielsweise offen, welchen technischen Vorteil humanoide Agenten von den erlernten Neuronenfunktionen haben könnten. Der Aspekt der Adaptierbarkeit, abhängig von den Gliedmaßenlängen, wurde ebenfalls nicht einbezogen.

Antonelli, Chinellato und del Pobil (2011) schlagen ein Modell vor, bei dem eine sensomotorische Karte gelernt wird, die es dem Roboter ermöglicht, Blickverfolgungs- und Greifbewegungen durchzuführen. Die Autoren bilden die neuropsychologische Funktion der Koordinatentransformation (vgl. Abschnitt 2.1.2) für visuelle und propriozeptive Daten ab. Der Roboter lernt zum Beispiel, seine Armbewegungen mit dem Blick zu verfolgen. Diese Bewegungen werden verwendet, um sein Visuomotorwissen über die nahe Umgebung zu aktualisieren. Dieses Wissen bezeichnen die Autoren als peripersonalen Raum. Diese Repräsentation liegt in Form eines sogenannten *Visuomotorbewusstseins* (engl. visuomotor awareness) des umgebenden Raums vor. Nach den Autoren emergiert dieses Bewusstsein aus der Interaktion mit dem peripersona-

len Raum und ist daher als *implizite Repräsentation* und nicht als Umgebungskarte aufzufassen. So interpretieren sie den peripersonalen Raum weder als explizite Karte des Raumes und der darin enthaltenen Objekte, noch als einen Raum, den ein humanoider Agent innerhalb bestimmter Distanzen erreichen kann. Das Resultat stellt lediglich eine Fähigkeit dar, in der nahen Umgebung zu interagieren. Der Fokus der Arbeit liegt auf der Neuromodellierung und nicht auf der Verbesserung technischer Systeme, durch die Entwicklung räumlicher Fähigkeiten in konkreten Anwendungen.

Goerick u. a. (2005) benutzen den peripersonalen Raum als Mittel, um die Mensch-Roboter-Interaktion zu strukturieren. Die Autoren machen sich Aspekte des peripersonalen Raums sowie des menschlichen Aufmerksamkeitsverhaltens zunutze. Der Sichtbereich eines Stereokamerasystems wird hierbei in zwei Räume geteilt. Er besteht zum einen aus dem gesamten dreidimensionalen Sichtfeld und zum anderen aus einem darin liegenden Unterraum. Letzterer ist ein Raumbereich, der sich direkt vor dem Kamerasystem befindet. Dieser wird als *körperzentrierte Zone* (engl. *body-centered zone*) betrachtet und bildet somit den peripersonalen Raum. Je nachdem, in welchem der beiden Raumbereiche visuelle Signale wahrgenommen werden, werden andere Verhaltensweisen in dem angeschlossenen technischen System aktiviert. Werden außerhalb des peripersonalen Raums visuelle Bewegungen wahrgenommen, so verfolgt das System das visuelle Signal. Es richtet gewissermaßen seine Blickrichtung auf die gesehene Bewegung. Wird das verfolgte Signal als Winkbewegung erkannt, die sich gleichzeitig der Kamera nähert, also in den peripersonalen Raum tritt, wird das Signal automatisch als Handposition weiterverarbeitet. Hier modellieren die Autoren das natürliche Verhalten von Menschen, in Interaktionen die Aufmerksamkeit von Partnern auf Dinge zu lenken. Signale im peripersonalen Raum dominieren das technische Verhaltenssystem, das bedeutet, dass das System nun die Hand und gegebenenfalls in der Hand gehaltene Objekte fixiert. Mit diesem einfachen Verfahren ist es also möglich, das Aufmerksamkeitsverhalten des Systems durch eine räumliche Strukturierung automatisch zu steuern. Goerick und Kollegen haben in ihrer Arbeit nicht den Anspruch, den peripersonalen Raum in all seinen Facetten auf neuronaler Basis technisch nachzuempfinden, sondern lassen sich vielmehr von ein paar Aspekten des peripersonalen Raums inspirieren. Ihr Modell nimmt im Gegensatz zu den bisher vorgestellten Arbeiten eine tatsächliche Raumaufteilung für das Robotiksystem vor. Die Repräsentation ist also als *explizit* einzuordnen. Im Modell von Goerick u. a. wird bisher jedoch nur die visuelle Modalität behandelt. Darüber hinaus handelt es sich um ein Stereokamerasystem

und nicht um einen humanoiden Agenten. Die Größe des peripersonalen Raums ist daher nicht als Greifraum definiert, abhängig von der Länge der Arme, sondern fest vorgegeben.

Resümee

Da nur wenige Arbeiten zum peripersonalen Raum in technischen Systemen existieren, kann bisher keines der bestehenden Modelle die Teilprobleme dieser Dissertation vollständig lösen. Einige Modellierungsaspekte können jedoch nützlich sein. Wie Antonelli u. a. (2011) behandeln auch wir den Aspekt unterschiedlicher Koordinatensysteme einzelner Sensormodalitäten und der Transformation zwischen diesen. In der Arbeit von Fuke u. a. (2009) ist der Einsatz von taktilen Sensordaten von besonderem Interesse, da er eine Alternative zu der visuellen Modalität bietet, um den peripersonalen Raum wahrzunehmen. Alternative Sensormodalitäten könnten zu einer robusteren Wahrnehmung des Nahraums beitragen. Goerick u. a. (2005) unterteilen den gesamten wahrgenommenen Raum und verknüpfen die Teilräume mit unterschiedlichen Funktionen. Mit dieser Vorgehensweise ist ihnen eine sinnvolle und geeignete Lösung für die Mensch-Roboter-Interaktion gelungen. Diese Aspekte sollten bei der Lösung der Teilprobleme der **Wahrnehmung und Repräsentation des Nahraums (P1)** und des darin gezeigten **intelligenten räumlichen Verhaltens (P3)** mitberücksichtigt werden.

3.2.2 Körperrepräsentationen für humanoide Agenten

Wenn ein humanoider Agent ein Werkzeug hält, so erweitert sich auch der räumliche Bereich, den er mit den Armen erreichen kann, und somit sein peripersonaler Raum. Diese Eigenschaft ist entscheidend für die gesuchte Körperrepräsentation. Diese findet sich in der Arbeitsdefinition 3 aus Abschnitt 2.2.3 wieder. Die Definition verbindet die technisch motivierten Teilprobleme P1 und P2 und führt zu der folgenden **Modellierungsentscheidung**: Die gesuchte Körperrepräsentation für einen humanoiden Agenten kann als *Körperschema* interpretiert werden. Eine weitere Modellierungsentscheidung lautet: Das Körperschema eines humanoiden Agenten soll die *Ausmaße* und *Körperbezugspunkte* für die peripersonale Raumrepräsentation beinhalten.

Die Definition bildet somit ein hilfreiches Kriterium für technische Modellierungsentscheidungen, die im weiteren Verlauf diskutiert werden. Wie in Abschnitt 1.2 bereits

beschrieben, sind technische Arbeiten zur Wahrnehmungs- und Verhaltensmodellierung humanoider Agenten noch nicht ausreichend auf Interaktion auf nahe Distanzen ausgerichtet. Eine weitere Grundlage für die Modellierungsentscheidungen bilden daher die technischen Arbeiten, die relevante Teilaspekte behandeln. In diesem Abschnitt wird zunächst ein Überblick über die Eigenschaften des biologischen Körperschemas gegeben, die in technischen Systemen umsetzbar und sinnvoll sind. Die Eigenschaft der Integration von Werkzeugen in das Körperschema findet sich üblicherweise in Körperschema-Lernverfahren wieder. Relevante Beispiele für existierende Ansätze zur Generierung einer Körperschemarepräsentation, die biologische Eigenschaften für humanoide Agenten nachmodellieren, werden in Abschnitt 3.2.3 vorgestellt. Es werden anschließend Anforderungen an eine Körperschemarepräsentation formuliert, die geeignet wäre, um als Grundlage für eine peripersonale Raumrepräsentation zu dienen. Auf Basis dieser Anforderungen werden die in Abschnitt 3.2.3 vorgestellten Ansätze verglichen, analysiert und ein geeigneter Ansatz vorgeschlagen, der im Anschluss daran im Detail erläutert wird.

Eigenschaften eines künstlichen Körperschemas

Viele Forschungsarbeiten in der Robotik wurden durch das Konzept des Körperschemas (Abschnitt 2.2.2) inspiriert. Das biologische Vorbild soll zum einen dazu dienen, die Motorkontrolle humanoider Agenten zu *verbessern*, und zum anderen sollen sie sich dadurch an Änderungen ihrer Körperstruktur *anpassen* können. Für den Einsatz in künstlichen Agenten besteht die große Herausforderung grundsätzlich darin, relevante biologische bzw. neuropsychologische Befunde in ein technisches Modell zu überführen, da die Ergebnisse oftmals nicht konkret genug sind. Dies hat zur Folge, dass verschiedene technische Interpretationen des Körperschemas und seiner Eigenschaften existieren. Das Körperschema, wie es in Definition 2 in Abschnitt 2.2.2 eingeführt wurde, bezeichnen wir im weiteren Verlauf als **biologisches Körperschema**, um es von dem Körperschema, wie es in technischen Systemen umgesetzt wird, zu unterscheiden. Gerade technische Einschränkungen können zu einer Interpretation führen, die nicht immer eine exakte Nachbildung des biologischen Körperschemas ist. In technischen Arbeiten werden zumeist einzelne Eigenschaften des biologischen Körperschemas nachgebildet, die für humanoide Agenten sinnvolle Erweiterungen darstellen könnten. Diese werden im folgenden Überblick zusammengefasst, und es werden zudem Verweise zu den in

Kapitel 2 beschriebenen biologischen Eigenschaften gegeben. Ein nach dem aktuellen Stand umfassender Bericht über Körperschemaarbeiten in der Robotik findet sich in Hoffmann, Marques, Arieta, Sumioka, Lungarella und Pfeifer (2010). Ein technisch interpretiertes Körperschema definieren wir als **künstliches Körperschema**, wenn mindestens eine der folgenden Eigenschaften darin umgesetzt wird:

1. Adaptivität (vgl. Abschnitt 2.2.3): Diese Eigenschaft bezieht sich auf die Anpassung an Körperänderungen und an Werkzeuggebrauch. Die Plastizität des biologischen Körperschemas wird üblicherweise als Lernmechanismus interpretiert. Solche Mechanismen sollen Entwicklern in erster Linie Arbeit bei der Definition der Körperstruktur bzw. bestimmter Körpereigenschaften humanoider Agenten abnehmen. Ändert sich die Körperstruktur, so hat dies Konsequenzen für die Koordinatentransformationen und die Motorkontrolle. Da zum Beispiel die kinematischen Funktionen von Segmentlängen und Gelenkpositionen abhängen, müssten die Entwickler die Kontrollprogramme an die neue Körperstruktur anpassen. Hierfür sind adaptive Verfahren, die die Körpereigenschaften automatisch lernen, von großem Vorteil. Die adaptiven Mechanismen werden zudem häufig verwendet, um den Werkzeuggebrauch für artikulierte Figuren zu erlernen. Der Ausgangspunkt ist die Annahme, dass bei Werkzeugnutzung das biologische Körperschema so erweitert wird, als sei es eine Verlängerung des Armes. Dieser Effekt soll es Robotern ermöglichen, ein Werkzeug in die eigene künstliche Körperschemarepräsentation zu integrieren, um es wie einen eigenen Arm zu verwenden.

2. Koordinatentransformationen (vgl. Abschnitt 2.1.2): Koordinatentransformationen entsprechen aus technischer Sicht der Transformation, wie sie in der Formel 3.1 in Abschnitt 3.1.1 beschrieben ist. Transformationen zwischen verschiedenen Referenzsystemen, wie sie in dem Greifbeispiel in Abschnitt 2.1.2 eingeführt wurden, werden dabei explizit berechnet und nacheinander ausgeführt. Hierfür müssen zum Beispiel die korrekten Segmentlängen und Gelenkpositionen der artikulierten Figur bekannt sein. Zudem sind sie für die Vorwärts- und Rückwärtskinematik artikulierter Figuren, wie sie ebenfalls in Abschnitt 3.1.1 beschrieben wurden, unerlässlich.

3. Motorkontrolle (vgl. Abschnitt 2.2.2): Mit dieser Eigenschaft ist die Beziehung zwischen dem Körperschema und der Motorkontrolle gemeint. Arbeiten, die sich auf diese Beziehung konzentrieren, interpretieren das biologische Körperschema als Basis für Motorbewegungen. Soll diese Beziehung gelten, so muss das künstliche Körperschema Parameter enthalten, die für die Motorkontrolle wichtig sind. Dies können zum Beispiel die Segmentlängen und Gelenkorientierungen des kinematischen Skeletts sein, die

nötig sind, um Roboter manipulieren in Bewegung zu setzen. Eine weitere Möglichkeit ist, aus Körperbewegungen die resultierenden Zusammenhänge zwischen sensorischen Zuständen (z. B. propriozeptive und visuelle) und Körperposturen zu ermitteln. Die technische Formalisierung dieser Zusammenhänge wird auch als Sensor-Motor-Karte (engl. sensorimotor map) bezeichnet. Auch diese kann in der Motorkontrolle eingesetzt werden.

4. Sensorkombination (Vgl. Abschnitt 2.2.2): Diese Eigenschaft bezeichnet die Berücksichtigung verschiedener Sensormodalitäten im Körperschema. Liegen Sensorinformationen unterschiedlicher Modalitäten in unterschiedlichen Referenzsystemen vor, zum Beispiel visuelle in augenzentrierten und taktile in gliedmaßenzentrierten, so kann man die in Punkt 2. beschriebene Koordinatentransformation auch als Beziehung zwischen den verschiedenen Sensormodalitäten interpretieren. Eine weitere Art, die Integration von Sensormodalitäten im Sinne des biologischen Körperschemas in technischen Arbeiten umzusetzen, ist die Einbeziehung mehr als einer Sensormodalität. Hierbei werden zumeist visuelle und propriozeptive Sensorinformationen genutzt, um das künstliche Körperschema zu erlernen oder zu repräsentieren. Im Fall der Repräsentation kann die bereits beschriebene Sensor-Motor-Karte als Integration verschiedener Sensormodalitäten interpretiert werden. Ein übliches Vorgehen ist, den Roboter zunächst einen wie bei Menschen vermuteten Body-Babbling-Prozess (siehe Abschnitt 2.2.2) durchlaufen zu lassen. Hiermit können die Zusammenhänge zwischen den Sensorinformationen und den Körperbewegungen gelernt werden, die einen Teil des Körperschemas repräsentieren sollen. Dieses Vorgehen ist aus dem Forschungsfeld des sogenannten Imitationslernens bekannt (vgl. Breazeal, Buchsbaum, Gray, Gatenby und Blumberg, 2005), bei dem humanoide Agenten durch Nachahmung Motorbewegungen lernen sollen (vgl. Kopp, Wachsmuth, Bonaiuto und Arbib, 2008).

3.2.3 Ansätze zur Generierung eines künstlichen Körperschemas

Sinnvolle Einteilungen der existierenden technischen Arbeiten zur Bildung einer Körperschemarepräsentation für humanoide Agenten wurden in dem Überblicksartikel von Hoffmann u. a. (2010) vorgestellt. Die in Abschnitt 3.2.2 erläuterte *Adaptivität* des künstlichen Körperschemas ist die zentrale Eigenschaft, die in jeder technischen Körperschemaarbeit behandelt wird. Wie bereits beschrieben, wird diese Eigenschaft

technisch als Lernproblem interpretiert und ist auf diese Weise einfacher formalisierbar. Die *expliziten* und *impliziten* Modelle unterscheiden sich darin, welche Aspekte des Körperschemas repräsentiert werden und in welcher Form. Eine weitere Unterscheidung betrifft die Motivation der technischen Modelle. Hier können zum einen die *Verbesserung des Roboterverhaltens* oder aber die *Modellierung nach einem biologischen Vorbild* im Vordergrund stehen. Stellvertretend für die einzelnen Kategorien werden in diesem Abschnitt relevante Arbeiten vorgestellt und auf Eigenschaften des in Abschnitt 3.2.2 definierten *künstlichen Körperschemas* überprüft.

Explizite Modelle: In diesen Modellen liegen die Beschreibungen üblicherweise als explizite Formeln vor. Diese können die Abbildungen zwischen den sensomotorischen Informationen, aber auch Körperstrukturen wiedergeben. Sie sind dazu geeignet, das Verhalten eines Systems durch Datenextrapolation auch für nicht vorher gelernte Ereignisse zu berechnen. Explizite Modelle wurden zum Beispiel von Sturm, Plagemann und Burgard (2009) sowie von Hersch, Sauser und Billard (2008) vorgestellt. Sturm u. a. verwenden ein sogenanntes Bayes-Netzwerk, in dem die Knoten die Segmente eines Roboterarmes und die dazugehörigen Kontrollsignale repräsentieren. Diese Arbeit setzt somit die Eigenschaften der *Adaptivität* und *Motorikontrolle* des künstlichen Körperschemas um. Das Modell von Hersch u. a. wird in Abschnitt 3.2.4 genauer erläutert.

Implizite Modelle: In diesen Modellen können vorgelernte sensomotorische Informationen zum Beispiel in Form von neuronalen Netzen vorliegen. Solche Modelle eignen sich für Eigenschaften, die aufgrund hoher Komplexität nicht durch ein explizites Modell erfasst werden können. Das Verhalten des Systems kann dann durch Dateninterpolation beschrieben werden. Das Körperschema liegt implizit vor, das bedeutet, dass die Körperstrukturen bzw. Körperteile keine direkte Entsprechung in der Repräsentation finden. Zu den impliziten Modellen zählen u. a. die Arbeiten von Fuke, Ogino und Asada (2007), von Stoytchev (2003) und von Yoshikawa, Kawanishi, Asada und Hosoda (2002). Bei Fuke u. a. lernt ein Roboter durch Betasten des eigenen Gesichtes dessen Repräsentation. Es werden dafür taktile, propriozeptive und visuelle Sensordaten über neuronale Netze integriert. Bei Stoytchev stehen die Erweiterung des Roboter-Körperschemas zur *Werkzeugnutzung* sowie die Erkennung der eigenen Körperteile durch Abgleich der visuellen Sensordaten im Fokus. Bei Yoshikawa u. a. lernt ein Roboter die Repräsentation seiner Körperoberfläche, wenn er seine eigenen Bewegungen beobachtet. Die sensorischen Zusammenhänge liegen dann in einer so-

Tabelle 3.1: Übersicht repräsentativer Ansätze zur Generierung eines künstlichen Körperschemas und ihre Eigenschaften.

	Koordinaten- transf.	Erweite- rung	Motor- kontrolle	Sensor- modalität	Reprä- sentation	Motiv
Sturm u. a. (2009)	✓	✓	✓	visuell	explizit	techn.
Hersch u. a. (2008)	✓	✓	✓	visuell, propriozeptiv	explizit	techn.
Fuke u. a. (2007)	-	-	-	taktil,visuell, propriozeptiv	implizit	biol.
Stoytchev (2003)	-	✓	✓	visuell, propriozeptiv	implizit	biol.
Yoshikawa u. a. (2002)	-	-	-	taktil,visuell, propriozeptiv	implizit	biol., techn.

nannten cross-modalen Karte vor. In allen drei Arbeiten findet sich die Eigenschaft der *Sensorkombination* des künstlichen Körperschemas wieder. Die Eigenschaft der *Adaptivität* wird bei Stoytchev zusätzlich durch die Werkzeugnutzung erweitert.

Motivation: Zu den Arbeiten, die zum Ziel haben, durch das Konzept des Körperschemas die Fähigkeiten und das Verhalten von Robotern zu *verbessern*, zählen die expliziten Modelle von Sturm u. a. und Hersch u. a. sowie das implizite Modell von Yoshikawa u. a. Die ersten beiden Modelle können direkt in der Motorkontrolle verwendet werden. Durch das Modell von Yoshikawa u. a. ist es einem Roboter möglich, Objekte in der Welt von den eigenen Roboterkörperteilen zu unterscheiden. Stoytchev und Fuke u. a. wiederum stellen implizite Modelle vor, die zum Ziel haben, neurowissenschaftliche bzw. *biologische Mechanismen* nachzumodellieren, um sie zu verstehen. Diese Arbeiten folgen dem Forschungsparadigma der *Künstlichen Intelligenz*, technische Modelle zu entwickeln, um menschliche Intelligenz bzw. menschliches Verhalten zu verstehen. Technische Vorteile spielen zunächst eine untergeordnete Rolle. Deutlich wird dies auch im Vergleich zu den expliziten Modellen, die sich durch transparente Repräsentationen häufig besser eignen, das Verhalten der Agenten direkt zu beeinflussen.

3.2.4 Vergleich und Analyse

Die in den Abschnitten 3.2.2 und 3.2.3 eingeführten Eigenschaften und Kategorien für künstliche Körperschemata können nun verwendet werden, um Anforderungen an ein künstliches Körperschema zu formulieren, das als Wissensbasis für den peripersona-

len Raum eingesetzt werden kann. Um festzustellen, welche der vorgestellten Modelle geeignet wären, um ein Körperschema zu generieren, das Definition 3 entspricht, wurden in Tabelle 3.1 die technischen Arbeiten mit ihren Eigenschaften zusammengefasst. Da alle vorgestellten Ansätze mit Änderungen der Körpergeometrie umgehen können, wurde die Eigenschaft der Adaptivität nicht separat aufgeführt. Für den peripersonalen Raum ist jedoch eine damit zusammenhängende Eigenschaft (vgl. Abschnitt 2.2.3) zentral. Sie ist in der Tabelle als *Erweiterung* aufgeführt und bezeichnet die Eigenschaft des künstlichen Körperschemas, sich an Werkzeugnutzung anzupassen und somit den peripersonalen Raum erweitern zu können. Für unser gesuchtes künstliches Körperschema ist diese Eigenschaft entscheidend. Sowohl Sturm u. a. (2009), Hersch u. a. (2008) als auch Stoytchev (2003) können dieses Phänomen nachbilden. Eine weitere entscheidende Eigenschaft des gesuchten künstlichen Körperschemas ist, dass die Ausmaße des peripersonalen Raums einfach aus der Repräsentation zu ermitteln sind. Zudem ist wichtig, dass die Körperrepräsentation die Ursprünge der Referenzsysteme für eine mögliche körperzentrierte peripersonale Raumrepräsentation beinhaltet. Beide Eigenschaften hängen von der Repräsentationsart ab. Obwohl implizite Modelle den Vorteil haben, Eigenschaften mit hoher Komplexität erfassen zu können, entsteht für unsere Anforderungen ein Nachteil gegenüber expliziten Modellen. Die Grenze des peripersonalen Raums ist nur ermittelbar, wenn die Gliedmaßenlängen des Körpers explizit repräsentiert sind. Dies ist nur mit expliziten Modellen möglich, in denen die Repräsentation transparent ist. Ein weiterer Vorteil expliziter Repräsentationen ist, dass die Leistungsfähigkeit des Modells und mögliche Fehler leichter zu erkennen sind. Um humanoide Agenten in Nahrauminteraktion mit der Fähigkeit auszustatten, erkennen zu können, ob sie benötigte Gegenstände erreichen können oder nicht, entscheiden wir uns zugunsten einer *expliziten Repräsentation*.

Somit bleiben als geeignete Ansätze aus Tabelle 3.1 die Arbeiten von Sturm u. a. (2009) und Hersch u. a. (2008). Beide Ansätze haben zum Ziel, humanoide Agenten technisch zu verbessern, und mit beiden Modellen ist es durch die explizite Repräsentation möglich, Koordinatentransformationen durchzuführen. Diese Eigenschaft ist zentral, um Gegenstände, die in verschiedenen Sensormodalitäten wahrgenommen werden und somit in verschiedenen Referenzsystemen vorliegen, ineinander zu überführen. Eine weitere nützliche Eigenschaft ist, dass die Repräsentation direkt in der Motorkontrolle verwendet werden kann. Das einzige Merkmal, in denen sich die Modelle unterscheiden, sind die Sensormodalitäten, die für das Lernen verwendet werden. Sturm

u. a. (2009) nutzen ausschließlich visuelle, Hersch u. a. (2008) verwenden noch zusätzlich propriozeptive Sensordaten. Generell sind Lernverfahren mit visueller Eingabe darauf angewiesen, dass die beobachtbaren Stimuli durch die Kameras erfasst werden können. Bezieht ein Lernverfahren seine Informationen nur aus einer Modalität, können ungenaue bzw. fehlerhafte Sensordaten nicht durch alternative kompensiert werden und zu fehlerhaften oder ungenauen Ergebnissen führen. Diese Schwierigkeiten sprechen daher gegen das von Sturm u. a. (2009) vorgeschlagene Verfahren. Hersch (2009) selbst sieht sein Körperschemamodell als Modell für den peripersonalen Raum an, da das Körperschema die Raumwahrnehmung strukturiert und mit der Raumrepräsentation eng verbunden ist. Sein Ansatz bietet eine gute Voraussetzung für die technische Realisierung räumlicher Wahrnehmung und räumlicher Fähigkeiten humanoider Agenten, wie wir sie für Nahrauminteraktion anstreben. Aus diesem Grund halten wir Herschs Methode für geeignet, eine Grundlage zur Lösung des Teilproblems der **Repräsentation des Körpers (P2)** zu bilden. Im nächsten Unterabschnitt wird sein Verfahren genauer dargestellt.

Lernen des Körperschemas nach Hersch

Hersch (2009) fasst sein Körperschemamodell als eine Hierarchie von Koordinatensystemtransformationen auf. Die baumartige Hierarchie wird dabei durch das kinematische Skelett eines humanoiden Agenten aufgespannt. Entlang der darin verbundenen Gelenke und starren Segmente werden die Koordinatentransformationen durchgeführt. Die Transformationen repräsentieren einen *sensorischen Zusammenhang* zwischen propriozeptiven und visuellen Informationen, und das kinematische Skelett gibt die *Gliedmaßenstruktur* des Körpers wieder. Herschs Verfahren besteht aus einem Rekalibrierungsprozess, in dem für eine vorgegebene Skeletthierarchie eines Roboters versucht wird, Annahmen für die Gelenkpositionen und Gelenkachsen zu treffen.

Adaptierungsalgorithmus: Um die korrekten Koordinatentransformationen für ein kinematisches Skelett zu finden, entwickelten Hersch u. a. einen Algorithmus, der an dieser Stelle zunächst durch ein vereinfachtes Beispiel verständlich gemacht werden soll. Der Algorithmus behandelt das Problem, wie es in Abbildung 3.1 für den einfachsten Fall eines Manipulators mit einem einzelnen Segment dargestellt ist. Die Herleitung des Algorithmus ist in der Arbeit von Hersch (2009) im Detail dargelegt und wird hier zur besseren Übersicht durch die weiter unten aufgeführten Gleichungen

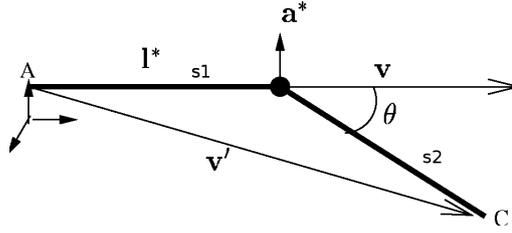


Abbildung 3.1: Einfachster Fall eines Einsegment-Manipulators. A ist der Ursprung, C der Endeffektor, s1 das proximale und s2 das distale Segment. Eine starre Transformation ist durch die Parametervektoren \mathbf{l} und \mathbf{a} beschrieben (Bild angepasst, entnommen aus Hersch [2009], Abb. 3.2, S. 40).

zusammengefasst. Es wird angenommen, dass durch diesen Manipulator eine Transformation \mathcal{T}_{orig} mit einem bekannten Rotationswinkel θ durchgeführt wurde. Es wird weiterhin angenommen, dass \mathcal{T}_{orig} dabei einen Vektor \mathbf{v}_{orig} in den transformierten Vektor \mathbf{v}_{trans} überführt hat, die beide bekannt sind. Der Vektor \mathbf{v}_{orig} liegt dabei im Bezugssystem des distalen Segments vor (siehe s2 in Abb. 3.1), Vektor \mathbf{v}_{trans} im Bezugssystem des proximalen Segments (siehe s1 in Abb. 3.1). Die Transformation \mathcal{T}_{orig} kann durch drei Parameter gekennzeichnet werden: einen Vektor \mathbf{l} , der die Gelenkposition angibt, eine Einheitsrotationsachse \mathbf{a} und den Rotationswinkel θ entlang der Achse \mathbf{a} . Sind die Vektoren \mathbf{l} und \mathbf{a} jedoch unbekannt, so ist es möglich, diese so lange anzupassen, bis sie der ursprünglichen Transformation \mathcal{T}_{orig} entsprechen. Die Idee dahinter ist, verschiedene Schätzungen für die Vektoren \mathbf{l} und \mathbf{a} zu treffen und eine Transformation $\mathcal{T}_{schätz}$ mit den bekannten Werten für den Rotationswinkel θ und den Vektor \mathbf{v}_{orig} durchzuführen. Der durch $\mathcal{T}_{schätz}$ transformierte \mathbf{v}_{orig} resultiert dann in einen Vektor $\mathbf{v}_{schätz}$. Entspricht $\mathbf{v}_{schätz}$ dem bekannten Vektor \mathbf{v}_{trans} , so wird davon ausgegangen, dass die geschätzten Werte für \mathbf{l} und \mathbf{a} richtig gewählt wurden. Um geeignete Schätzungen zu finden, wird ein Gradientenabstieg über den Fehler zwischen \mathbf{v}_{trans} und $\mathbf{v}_{schätz}$ durchgeführt. Dieses Prinzip kann auf Manipulatoren mit mehreren Segmenten übertragen werden. Für einen **Manipulator mit n Gelenken** wird der Gradientenabstieg dann über folgende Fehler ausgeführt:

$$\Delta \mathbf{l}_i = -\epsilon \frac{\partial}{\partial \mathbf{l}_i} \frac{1}{2} \|\mathbf{v}_{trans,n} - \mathcal{T}_{schätz}(\mathbf{v}_{orig,n})\|^2 \quad (3.2)$$

$$\Delta \mathbf{a}_i = -\epsilon \frac{\partial}{\partial \mathbf{a}_i} \frac{1}{2} \|\mathbf{v}_{trans,n} - \mathcal{T}_{schätz}(\mathbf{v}_{orig,n})\|^2 \quad (3.3)$$

Wie bereits in Abschnitt 3.1 als Notation eingeführt, bezeichnet \mathbf{l}_i den Translati-

onsvektor, der das Segment i beschreibt und Gelenk j_i zugeordnet ist. Der Wert ϵ ist ein kleiner positiver Skalar, der den Lernschritt festlegt. Der Vektor \mathbf{a}_i bezeichnet die Einheitsrotationsachse im Gelenk j_i , um den die Rotation \mathbf{R}_i mit dem Winkel θ_i durchgeführt wird. Für die allgemeine Form der Transformationen, wie sie in Gleichung 3.1 definiert ist, ergibt sich nach partieller Ableitung für Gleichung 3.2:

$$\Delta \mathbf{l}_i = \epsilon (\mathbf{v}_{trans,n} - \mathbf{v}_{schätz,n})^T \prod_{j=1}^{i-1} \mathbf{R}_j \quad (3.4)$$

Für die Gleichung 3.3 ergibt sich:

$$\Delta \mathbf{a}_i = \epsilon (\mathbf{v}_{trans,n} - \mathbf{v}_{schätz,n})^T \left(\left(\prod_{j=1}^{i-1} \mathbf{R}_j \right) \frac{\partial}{\partial \mathbf{a}_i} (\mathbf{R}_i (\mathbf{T}_{i+1} \circ \mathbf{R}_{i+1} \dots \circ \mathbf{T}_n \circ \mathbf{R}_n (\mathbf{v}_{orig,n}))) \right) \quad (3.5)$$

Über die Gleichungen 3.4 und 3.5 wird versucht, iterativ die Schätzungen der Rotationsachsen und Translationsvektoren zu verändern. Als Eingabe werden unterschiedliche Werte von \mathbf{v}_{orig} benötigt. Trotz guter Ergebnisse in der schnellen Berechnung eines funktional adäquaten Körperschemas bildet das resultierende Körperschema aber nicht immer die tatsächliche Körpergeometrie ab. Der Grund ist, dass nur die Endeffektoren bei der Bewegung beobachtet werden. Es kann dadurch mehrere kinematische Funktionen geben, die die gegebenen Sensorkombinationen erfüllen. Die Autoren bezeichnen das Ergebnis daher als **subjektives Körperschema** (engl. subjective body schema).

3.3 Interpersonale Körperräume

Interpersonale Körperräume werden sowohl im Bereich der Robotik als auch bei virtuellen Menschen im Zusammenhang mit der räumlichen Verhaltenskontrolle untersucht und modelliert. Diese Modelle kommen in Szenarien zum Einsatz, die folgende Eigenschaften aufweisen: Der Roboter oder der virtuelle Mensch muss sich in einer Umgebung fortbewegen, die durch Wände oder Einrichtungsgegenstände beschränkt ist. In dieser Umgebung befinden sich zudem menschliche Partner oder andere humanoide Agenten. Die Herausforderung liegt oftmals in der Navigation und Orientierung in der beschränkten Umgebung unter Berücksichtigung der räumlichen Distanzen zu den Interaktionspartnern. Im weiteren Verlauf werden relevante Arbeiten vorgestellt,

die interpersonale Räume für verkörperte Agenten modellieren und in Interaktionsszenarien einsetzen. Wir unterscheiden zwei Arten von Modellen. Modelle, mit denen hauptsächlich das Distanzverhalten künstlicher Agenten kontrolliert wird, bezeichnen wir im Folgenden als Modelle zur **Proxemikkontrolle**. Betrachten wir Modelle, die primär Gruppenformationen und die Körperorientierung der Agenten berücksichtigen, so sind dies Modelle zur **Formationenkontrolle**. Technische Raummodelle beziehen sich häufig auf die in Kapitel 2.3.2 beschriebenen interpersonalen Raummodelle. Zur Proxemikkontrolle von mobilen Agenten wird oft das *Proxemikmodell von Hall* verwendet. Arbeiten zur Formationenkontrolle verwenden meist das *F-Formationen-System von Kendon*. Der Grund für diese Wahl kann in den konkreten Definitionen und der einfachen technischen Umsetzung gesehen werden. Die Motivation hinter diesen Arbeiten besteht darin, räumliches Verhalten zu generieren, das sozialen Regeln der menschlichen Interaktion nachmodelliert ist. Diese Regeln sind menschlichen Partnern intuitiv verständlich und sollen ihnen den Umgang mit technischen Systemen erleichtern. Ein Teilaspekt sozialen Verhaltens ist *kooperatives Verhalten*. Ein umfassendes Modell für Kooperationsfähigkeiten künstlicher Agenten wurde von Pfeiffer-Lessmann (2011) vorgestellt, in dem das Konzept der *gemeinsamen kooperativen Aktivität* von Bratman (1992) eine wichtige Rolle einnimmt. Bratman charakterisiert das Konzept unter anderem über *gemeinsame Aktivität* und *gegenseitige Unterstützung*. In dieser Arbeit untersuchen wir die räumlichen Aspekte dieser Charakteristika und entwickelten daraus eine Definition *räumlich-kooperativen Verhaltens*.

Proxemikkontrolle

Das von **Holthaus, Pitsch und Wachsmuth (2011)** vorgestellte Modell dient der Steuerung der Roboteraufmerksamkeit für Personen. Es handelt sich dabei um ein distanzbasiertes System, unter Verwendung von Halls Proxemikmodell. Der stationär montierte Robotertorso BARTHOC mit dem Roboterkopf Flobi soll als Rezeptionist an einem öffentlichen Ort Besuchern Wegbeschreibungen geben können. Der Roboter soll, je nach Distanz des Besuchers, verschiedene Posturen zeigen. Befindet sich zum Beispiel ein Mensch im nahen öffentlichen Raum, verfolgt Flobi den Menschen nur mit den Augen. Nähert sich die Person, so verwendet er nicht nur die Augen, sondern auch den Kopf und den gesamten Rumpf, um sich der Person zuzuwenden. In einer Evaluationsstudie konnte gezeigt werden, dass menschliche Beobachter Flobi

mehr Interesse am Menschen zugeschrieben, wenn diese Strategie eingesetzt wurde. Zudem beurteilten die Beobachter, dass Flobi mit dieser Strategie seine Aufmerksamkeit und seine Intentionen früher und zu einem höheren Grad signalisierte, als mit anderen Strategien. Die von Holthaus u. a. (2011) vorgestellte Arbeit hat einen besonders engen Bezug zu dieser Dissertation, da sie ebenfalls im Rahmen des Projektes *C1 – Interaktionsraum* des SFB 673 entstanden ist. Wie in Abschnitt 1.3 bereits erläutert, bildet das Projekt C1 den Forschungskontext der vorliegenden Arbeit. Das Konzept des Interaktionsraums ist in dieser Arbeit auf Nahrauminteraktion eingegrenzt, im Projektteil *Kognitive Robotik* wird der Interaktionsraum hingegen auch für größere Interaktionsdistanzen definiert.

Das Modell von **Sisbot, Marin, Alami und Simeon (2006)** basiert nicht auf dem Proxemikmodell von Hall, sondern auf einem eigenen Ansatz. In dieser Arbeit soll ein mobiler Roboter auf eine Art durch den Raum navigieren, die speziell an die Anwesenheit von Menschen angepasst ist. So wird das Abstandsbedürfnis zwischen Roboter und Mensch nicht starr definiert, sondern an die jeweilige Interaktion angepasst. Eine Überlegung ist zum Beispiel, dass das Abstands- bzw. Sicherheitsbedürfnis unterschiedlich sein kann, je nachdem ob der Mensch sitzt oder steht. Eine weitere Einschränkung in der Abstandseinhaltung ist zum Beispiel, dass der Roboter keine Bewegungsrouten einschlägt, die sich hinter dem Rücken eines Menschen befindet, da das plötzliche Erscheinen des Roboters zu Unbehagen führen könnte. Im technischen Modell werden bestimmte Raumregionen mit unterschiedlichen Kostenfaktoren belegt. Die Kosten drücken z. B. ein erhöhtes Sicherheitsbedürfnis oder Hindernisse aus. Der Pfad mit den geringsten Kosten durch den Raum wird dann durch A*-Suche ermittelt.

Formationenkontrolle

Ein räumliches Navigationsmodell für virtuelle Charaktere wurde von **Pedica und Vilhjálmsson (2009)** vorgestellt. Die Grundlage ist hier das F-Formationen-System von Kendon. Das Ziel der Arbeit ist, das Gruppenverhalten von virtuellen Charakteren so zu modellieren, dass es einen gewissen Grad an sogenannter *sozialer Präsenz* (engl. social presence) vermittelt. Diese soziale Präsenz soll zum Beispiel in virtuellen Umgebungen von Computerspielen zum Einsatz kommen und das Spielerlebnis positiv beeinflussen. Das Verhaltensmodell steuert hierbei die Gruppenformationen der virtuellen Charaktere. Bewegt sich ein virtueller Charakter beispielsweise auf eine

vorhandene Gruppe von zwei anderen Charakteren zu, so formieren sie sich automatisch in gleichen Abständen um einen virtuellen O-Raum (siehe S.33). Die Gesichter der Charaktere zeigen dabei in die Mitte des O-Raums, sodass für einen Beobachter der Situation der Eindruck entsteht, die Charaktere seien einander zugewandt und gehörten zu einer gemeinsamen Gruppe.

Yamaoka, Kanda, Ishiguro und Hagita (2010) präsentierten eine Arbeit, in der ein fahrender Roboter die am besten geeignete Position und Orientierung im Raum einnimmt, um KundInnen bestimmte Produkte bzw. MuseumsbesucherInnen Exponate zu präsentieren. Der Ausgangspunkt des Modells ist auch hier Kendons F-Formationen-System. Das Modell wurde jedoch erweitert, um den zu präsentierenden Gegenstand in die Formation mit einzubeziehen. Aus einer empirischen Studie wurden Daten über Körperorientierungen und -positionen von menschlichen Präsentierern gesammelt. Aus den Daten wurde, zusammen mit dem erweiterten F-Formationen-System, ein Modell zur Formationenkontrolle entwickelt. In einer weiteren Studie, in der der Roboter mit diesem Modell gesteuert wurde, konnte gezeigt werden, dass das entstandene Verhalten von menschlichen Interaktionspartnern positiv aufgenommen wurde.

Resümee

Obwohl in dieser Arbeit keine Fortbewegung im Raum, sondern fixe Umgebungen betrachtet werden, stellt die Proxemik- und Formationenkontrolle im Nahraum einen wichtigen Faktor dar, der auch in einem Nahraummodell wiederzufinden sein sollte. Sozial wahrgenommene räumliche Verhaltensweisen könnten dazu führen, dass Menschen den künstlichen Agenten als Partner und nicht als tutorielles, technisches System wahrnehmen. Bei der Lösung des Teilproblems des **intelligenten räumlichen Verhaltens im Interaktionsraum (P3)** sollten diese Aspekte daher berücksichtigt werden.

3.4 Nahraumaufgaben mit humanoiden Agenten

In den vorangegangenen Abschnitten wurden technische Arbeiten vorgestellt, in denen sich die Interaktionspartner im interpersonalen Raum fortbewegen konnten. Die Distanzen überstiegen somit die Körperräume der beteiligten Partner. In diesem Ab-

schnitt liegt der Fokus auf Nahraumaufgaben, wie sie in der Einleitung eingeführt wurden. Diesen Szenarien ist gemein, dass die Interaktionen auf kurzen Distanzen stattfinden, innerhalb derer keine Fortbewegung nötig ist. Die hier vorgestellten Arbeiten beinhalten daher Manipulationsaufgaben, die von zwei Partnern gelöst werden sollen. Ein Partner ist meist der Mensch und der andere ein Roboter oder ein virtueller Mensch. Die Szenarien werden anhand der folgenden, für eine Nahraumaufgabe wünschenswerten Kriterien beleuchtet.

Der Aspekt der *Zugänglichkeit* beschreibt, ob der interpersonale Raum beiden Partnern zugänglich ist. Diese bedingt dabei die Art der Interaktionen, die zwischen den Partnern möglich sind, und kann als Voraussetzung für *Hilfsaktionen* und *kooperatives* Verhalten gesehen werden. Hilfsaktionen beschreibt die Fähigkeit des Agenten, den Menschen verbal oder durch Motoraktionen zu unterstützen. Der Aspekt der *Kooperation* bezeichnet die Rolle des künstlichen Agenten. Fungiert dieser als gleichberechtigter Partner, oder nimmt er die Funktion eines anleitenden Tutors ein? Als Mechanismus, um den gemeinsamen Raum zu modellieren oder zu repräsentieren, spielt das *Partnermodell* eine zentrale Rolle. Unter einem Partnermodell verstehen wir ein Modell von den Intentionen, Zielen und Annahmen des Partners über die jeweilige Umgebung. Unter die Annahmen fällt auch das Wissen über und die Wahrnehmung von Objekten aus der räumlichen Perspektive des Partners. Diese Art der Annahmen fassen wir unter dem *räumlichen Perspektivenmodell* zusammen. Die aufgeführten Kriterien sind als Herausforderungen des Teilproblems **P3 - intelligentes räumliches Verhalten im Interaktionsraum** zu verstehen. In späteren Kapiteln wird das in dieser Arbeit entwickelte Modell des Nah- und Interaktionsraums ebenfalls nach diesen Kriterien geprüft und diskutiert.

Max: Ein Szenario, in dem eine Nahraumaufgabe mit dem virtuellen Menschen Max zu lösen ist, wurde von **Kopp, Jung, Lessmann und Wachsmuth (2003)** vorgestellt. Max und ein menschlicher Partner stehen in dieser Interaktion um einen virtuellen Tisch in einem virtuell projizierten Raum. Die Aufgabe besteht darin, gemeinsam einen Propeller aus virtuellen Baufix-Elementen zusammenzusetzen. Zwar teilen sich beide Partner einen Tisch und können mit denselben Bauteilen operieren, direkte Objektmanipulationen sind jedoch nicht möglich. Stattdessen zeigen die Partner auf die gewünschten Objekte, um eine Aktion mit den Baufix-Elementen anzustoßen. Max hat in diesem Szenario die Rolle des *Tutors*. So kann er *verbale Hilfe* anbieten oder den Propeller *selbst zusammenbauen*, wenn der menschliche Partner kei-

ne eigenen Aktionen durchführen will. Hier wurde zunächst kein explizites Wissen aus der Partnerperspektive aufgebaut, allerdings wurde es in weiterentwickelter Form in der Arbeit Pfeiffer-Lessmann (2011) eingeführt.

Jast: Eine ähnliche, aber aktuellere Ausführung des Max-Baufix-Szenarios wurde mit dem Roboter Jast entwickelt (vgl. Foster und Matheson, 2008). In diesem Szenario sollen der Mensch und der Roboter, ebenfalls mit Baufix-Elementen, Manipulationen durchführen. Der *Raum* zwischen den beiden Partnern ist aber explizit *getrennt* in einen Bereich, in dem der Roboter, und einen Bereich, in dem der Mensch agieren darf. Geplant ist jedoch, dass der Roboter dem Menschen benötigte Gegenstände *überreicht*. In diesem Szenario nimmt Jast nicht die Rolle eines Tutors ein, sondern soll mit dem Partner die Aufgabe *gemeinsam* lösen.

Leonardo: Gray, Breazeal, Berlin, Brooks und Lieberman (2005) stellen mit dem Roboter Leonardo ein sehr fortgeschrittenes Modell zur Perspektivenübernahme, allerdings nicht der räumlichen, vor. Im Szenario soll ein Mensch eine Aufgabe mit Bauklötzen lösen, die von Leonardo beobachtet wird. Beide sitzen gemeinsam an einem Tisch. Zwar haben beide Zugang zum Tisch, da Leonardo jedoch keine humanoide Körperform besitzt, sondern bewusst mit tierähnlichen und kleinen Körpermaßen entwickelt wurde, ist davon auszugehen, dass der Mensch einen besseren Zugang zum gemeinsamen Raum hat. Zudem stehen die Objektmanipulationen in diesem Szenario nicht im Vordergrund. Ein benötigtes Objekt ist durch einen Sichtschutz für den menschlichen Partner verdeckt, aus Leonardos räumlicher Perspektive jedoch sichtbar. Leonardo bildet, während der Partner versucht, die Aufgabe zu lösen, ein Modell über das Aufgabenwissen und über die Pläne des Menschen. Auf dieser Basis inferiert der Roboter, ob der Mensch das verdeckte Objekt berücksichtigen kann. Ist dies nicht der Fall, zeigt Leonardo dem Partner das Objekt. Das Wissen über das verdeckte Objekt wird bei Leonardo bereits vorgegeben und nicht über räumliche Perspektivenübernahme erzeugt. Der Fokus liegt in diesem Szenario auf der Intentionzuschreibung und der Modellierung von mentaler Perspektivenübernahme. Auf Grundlage dieses Wissens führt Leonardo Schlussfolgerungen durch, um dem menschlichen Partner zu *helfen*.

3.5 Zusammenfassung

In diesem Kapitel wurde der Stand der Forschung zu drei Bereichen aufgearbeitet. Der erste Bereich umfasste die technischen Grundlagen der Bewegungsplanung und -kontrolle humanoider Agenten. Der zweite Forschungsbereich beinhaltete Ansätze zur technischen Modellierung und Generierung von individuellen Raum- und Körperrepräsentationen, die teilweise auf den in Kapitel 2 erläuterten theoretischen Konzepten basieren. Der dritte Bereich konzentrierte sich auf die technische Modellierung interpersonaler Räume für Interaktionen mit humanoiden Agenten.

Innerhalb des ersten Forschungsbereichs wurde zunächst die Bewegungsanimation virtueller Menschen beleuchtet, die im weiteren Verlauf verwendete Notation eingeführt und wichtige Begriffe erläutert. Für Armbewegungen wurden relevante Animationsansätze, Ansätze zur Hindernisvermeidung und die Planung von Objektmanipulationsbewegungen beschrieben. Diese Ansätze bilden den Stand der Forschung zur Lösung des Teilproblems des **intelligenten räumlichen Verhaltens im Nah- und Interaktionsraum (P3)**. Ein für das Teilproblem wichtiger Aspekt stellt die Modulation der Armbewegungsqualität dar, die es möglich macht, soziale Signale zu vermitteln. Aus den unterschiedlichen Ansätzen wurden die lokale Bewegungsplanung über das IKAN-Modell und der lokale Hindernisvermeidungsansatz der Potentialfeldmethode als geeignet betrachtet.

Aus dem zweiten Forschungsbereich wurde nach einer Voranalyse, das theoretische Konzept des peripersonalen Raums als geeignetes Konzept zur Lösung des Teilproblems der **Wahrnehmung und Repräsentation des Nahraums (P1)** identifiziert und technische Modelle dieses Raums vorgestellt. Für die Lösung des Teilproblems der **Repräsentation des Körpers (P2)** wurde eine Lösung gesucht, die gleichzeitig Anforderungen an die peripersonale Raumrepräsentation eines humanoiden Agenten erfüllt. Hierfür wurde das Konzept des Körperschemas als geeignete Modellierungsgrundlage gewählt und nach einem Vergleich existierender Arbeiten der Generierungsansatz von Hersch vorgeschlagen.

Im dritten Forschungsgebiet wurden Arbeiten vorgestellt, die Ergebnisse zur Lösung des Teilproblems des **intelligenten räumlichen Verhaltens im Interaktionsraum (P3)** bieten können. Unter diese Arbeiten fallen diejenigen zur Proxemik- und Formationenkontrolle humanoider Agenten. Die daraus resultierenden Erkenntnisse zu sozialer Rauminteraktion zwischen menschlichen und humanoiden Partnern sollten

auch in Nahrauminteraktion berücksichtigt werden. Relevante Beispielszenarien, in denen humanoide Agenten mit menschlichen Partnern in Nahrauminteraktion treten, wurden in einem Überblick vorgestellt und Kriterien vorgeschlagen (*Zugänglichkeit, Kooperation, Partnermodell*), nach denen diese Szenarien diskutiert werden können.

4 Entwurf eines peripersonalen und Interaktionsraums für virtuelle Menschen

Im Verlauf des 2. Kapitels konnten wir mit der interdisziplinären Herangehensweise ein besseres Verständnis über die im Einleitungskapitel aufgeworfenen Teilprobleme erlangen. Zudem halfen uns die theoretischen Konzepte dabei, im 3. Kapitel den Stand der Forschung einzugrenzen und erste Modellierungsentscheidungen zu treffen. Dieses Kapitel gibt nun einen Überblick über den Gesamtentwurf der in dieser Dissertation vorgestellten Lösungen für die Teilprobleme. Die Lösungen basieren zum einen auf Analysen der technischen Arbeiten, die im Stand der Forschung in Kapitel 3 aufgearbeitet wurden. Zum anderen basieren sie auf technischen Interpretationen der in Kapitel 2 eingeführten biologischen Konzepte. Die technischen Interpretationen werden in diesem Kapitel entwickelt.

4.1 Technische Interpretation der theoretischen Konzepte

Die drei Teilprobleme dieser Arbeit beruhen in erster Linie auf technischen Anforderungen. Diese entstanden aus der Frage, wie ein humanoider Agent so mühelos mit Gegenständen oder Interaktionspartnern auf nahe Distanz zum Körper interagieren kann wie der Mensch. Die Teilprobleme sind:

P1 – Wahrnehmung und Repräsentation des Nahraums: Dieses Teilproblem ist besonders für autonome, humanoide Agenten mit menschenähnlichen und daher zum Teil eingeschränkten sensorischen Systemen (z. B. nur mit einer Stereokamera im Kopf) relevant. Humanoide Agenten sollen ihre unmittelbare Umgebung auch ohne umfassende Sensorabdeckung wahrnehmen und strukturieren können.

P2 – Repräsentation des Körpers: Dieses Teilproblem beschreibt die Fähigkeit, festzustellen wie weit der humanoide Agent mit den Armen reichen kann. Die Fähig-

keit hängt auch mit derjenigen zusammen, festzustellen, welche Körperbewegungen nötig sind, um Objekte zu ergreifen – ob zum Beispiel einfache Arm- oder zusätzliche Torsobewegungen nötig sind. Hält der Agent ein Werkzeug in der Hand oder ändert sich seine Körperstruktur, so erweitert sich auch sein erreichbarer Raum.

P3 – Intelligentes räumliches Verhalten im Nah- und Interaktionsraum: Die räumlichen Verhaltensweisen, wie das Umplatzen von Objekten, soll für individuelle und Partner-Interaktion geeignet sein. Dabei soll das räumliche Verhalten anpassungsfähig und für den Partner intuitiv verständlich sein.

4.1.1 Ein peripersonaler Raum für humanoide Agenten

Mit dem Teilproblem P1 stellt sich die Frage, wie der Nahraum für einen humanoiden Agenten strukturiert und repräsentiert werden kann. In Abschnitt 3.2.1 kamen wir zu dem Ergebnis, dass eine egozentrische Raumrepräsentation für den Nahraum geeignet ist. Es wurde eine Betrachtung der Eigenschaften des biologischen peripersonalen Raums für eine technische Modellierung vorgeschlagen. Technische Arbeiten, die den peripersonalen Raum so modellieren, wie wir es für die Lösung des Teilproblems P1 anstreben, existieren nach dem aktuellen Stand der Forschung noch nicht. Für den biologischen peripersonalen Raum haben wir in Abschnitt 2.2.1 folgende Definition 1 aufgestellt:

Der peripersonale Raum ist der Raum, der mit den Armen erreicht werden kann. Die Repräsentation des Raums ist multisensorisch, körperzentriert und in den Sensormodalitäten invariant.

Für humanoide Agenten interpretieren wir die Eigenschaften wie folgt: Die Grenze des peripersonalen Raums für humanoide Agenten wird über die Länge der Arme und die maximale Vorlehnbewegung des Torsos definiert. Die Repräsentation des Raums wird über die taktile und visuelle Sensormodalität gebildet. Gegenstände im peripersonalen Raum werden in einem expliziten Speicher registriert. Ihre Positionen werden körperteilzentriert abgelegt. Welcher Körperteil das Ursprungskoordinatensystem vorgibt, hängt von der Sensormodalität ab, mit der der Gegenstand wahrgenommen wurde. Dieser wird so lange im Speicher vorgehalten, solange er mindestens in einer Sensormodalität noch wahrgenommen werden kann. Auf Grundlage dieser technischen Interpretation und der in Abschnitt 2.1.3 vorgestellten Raummodelle wird in Kapitel 6 ein peripersonales Raummodell entwickelt.

Carello, Grosfoky, Reichel, Solomon und Turvey (1989) definieren bei Menschen die Schwierigkeit, ein Objekt als erreichbar zu erkennen, als Berechnungsproblem mit drei Schritten. Der erste Schritt ist die Berechnung der Distanz des Zielobjektes. Der zweite Schritt ist die Berechnung der weitesten möglichen Ausdehnung der Körpergliedmaßen. Der dritte Schritt ist das Vergleichen der beiden Werte. Für die erfolgreiche Ausführung des zweiten Schrittes ist nach Meinung der Autoren ein Körperschema nötig. Aus den in Kapitel 2 vorgestellten Befunden bezüglich des Körperschemas bei Primaten entwickelten wir Definition 3 in Abschnitt 2.2.1:

Eine peripersonale Raumrepräsentation kodiert den Raum um den Körper in verschiedenen Körperbezugssystemen. Die Informationen über die Bezugspunkte am Körper und die Ausmaße des peripersonalen Raums werden durch das Körperschema geliefert.

Dieser biologische Zusammenhang ist für die vorliegende Arbeit von zentraler Bedeutung, da er die technischen Teilprobleme P1 und P2 miteinander verknüpft. Mit diesem Zusammenhang kommen wir der Frage näher, wie es einem humanoiden Agenten möglich ist, aus verschiedenen sensorischen Eindrücken eine sinnvolle Repräsentation des eigenen Körpers im Nahraum zu bilden. Wir interpretieren den Zusammenhang aus Definition 3 für humanoide Agenten daher wie folgt:

Definition 4 : *Eine peripersonale Raumrepräsentation für einen humanoiden Agenten kodiert den Raum um dessen Körper in körperteilzentrierten Koordinatensystemen. Das künstliche Körperschema ist die Informationsbasis, über die die Ursprungspositionen der Koordinatensysteme bereitgestellt werden. Über die Gliedmaßenlängen wird die Grenze des peripersonalen Raums berechnet. Ändert sich die Körperstruktur des Agenten oder wird ein Werkzeug verwendet, erweitert sich automatisch die Grenze des peripersonalen Raums.*

Insbesondere der Aspekt der Werkzeugnutzung wird in technischen Arbeiten, wie sie in Abschnitt 3.2.2 beschrieben wurden, üblicherweise dem *Körperschema* zugeschrieben. Wie in Abschnitt 2.2.3 erläutert, ist diese Vermutung in den Neurowissenschaften jedoch umstritten. Eine alternative Auffassung ist dort, dass die dem Körperschema zugerechnete Erweiterung durch Werkzeuge eigentlich eine Erweiterung der *Repräsentation des peripersonalen Raums* ist. Diese Vermutung erscheint plausibel, da nach den vorgestellten Befunden die räumliche Plastizität den Eigenschaften des periper-

sonalen Raums am ehesten entspricht. Wir vertreten in dieser Arbeit deswegen diese Sichtweise und berücksichtigen sie in der technischen Realisierung.

4.1.2 Ein künstliches Körperschema

Wie bereits im vorherigen Abschnitt 4.1.1 erläutert, eignet sich für die Lösung des Teilproblems P2 eine Körperrepräsentation, die Eigenschaften des biologischen Körperschemas nachmodelliert. Diese Modellierungsentscheidung wird durch die in Abschnitt 3.2 vorgeschlagene Lösung in Form eines künstlichen Körperschemas gestützt. Da die Körperschema-Lernverfahren sich sehr stark unterscheiden, ist es wichtig, zunächst festzustellen, welches künstliche Körperschema geeignet wäre. Nach Analyse der existierenden Ansätze (vgl. Abschnitt 3.2.4) sollte ein für diese Arbeit geeignetes künstliches Körperschema explizit repräsentiert und direkt in der Motorkontrolle einsetzbar sein. Die vorgestellten expliziten Modelle liefern eine Körperrepräsentation, die ein *kinematisches Skelett* (vgl. Abschnitt 3.1.1) beinhaltet. Darüber hinaus soll die explizite Repräsentation Koordinatentransformationen bewerkstelligen und Werkzeuge integrieren können. Für einen weniger fehleranfälligen Lernprozess sollten mindestens zwei verschiedene Arten von Sensorinformationen verwendbar sein. Fehlerhafte Sensordaten können somit durch alternative Modalitäten kompensiert werden. Je mehr Modalitäten einem Adaptiermechanismus zur Verfügung stehen, desto robuster kann auch die Adaptierfähigkeit des humanoiden Agenten sein. Diese Eigenschaft entspricht auch dem biologischen Vorbild, das laut Definition 2 aus mehreren Sensormodalitäten gebildet wird.

Das Verfahren von Hersch u. a. (2008) ist in der Lage, ein solches künstliches Körperschema zu generieren, und wird daher als geeignete Grundlage zur Lösung des Teilproblems P2 angesehen. Eine Schwierigkeit liegt in der Generierung eines subjektiven Körperschemas. Dieses resultiert aus zu wenig Referenzpunkten für den von Hersch vorgestellten visuellen Lernprozess. Der große Nachteil dieses Ergebnisses ist, dass zwar die kinematischen Funktionen für die Endeffektoren korrekt angenähert werden, die einzelnen Gliedmaßenlängen jedoch stark variieren können. Eine darauf aufbauende peripersonale Raumrepräsentation ist aber auf korrekte Gliedmaßenlängen angewiesen. Basierend auf Herschs Lernverfahren schlagen wir daher in Kapitel 5 einen alternativen Lernprozess über *taktile Sensordaten* vor, der mehr Referenzpunkte und somit akkuratere Ergebnisse für das resultierende Körperschema liefert.

4.1.3 Kooperatives Verhalten im interpersonalen Raum

Bezüglich des Teilproblems P3 stellt sich die Frage, wie das räumliche Verhalten eines humanoiden Agenten im interpersonalen Raum modelliert werden soll. Eine weitere Frage ist, wie der interpersonale Nahraum, den sich ein humanoider Agent mit einem Partner teilt, strukturiert und repräsentiert werden kann.

Wir wenden uns zuerst der räumlichen Verhaltensmodellierung zu. Wie in Abschnitt 3.3 beschrieben, verfolgen viele technische Systeme das Ziel, räumliches Verhalten nach sozialen Interaktionsregeln bei Menschen zu generieren. Dieses soll für den Menschen intuitiv verständlich sein. Für uns liegt der Fokus im Besonderen auf dem *räumlich-kooperativen Verhalten* (engl. spatial cooperative behavior). Kamashima, Kanda, Imai, Ono, Sakamoto, Ishiguro und Anzai (2004) sehen darin einen wichtigen Aspekt für die natürliche Kommunikation mit verkörperten Agenten. In ihrer Arbeit wird dieses Verhalten bei Robotern dadurch ausgedrückt, dass sie in Wegbeschreibungssituationen Richtungsangaben mit dem Arm zeigen können.

Kooperatives Verhalten ist ein Teilaspekt sozialen Verhaltens. Aus dem Modell von Bratman (1992) stammt das Konzept der *gemeinsamen kooperativen Aktivität*. Zwei Charakteristiken seines Modells sind für die Lösung des Teilproblems P3 relevant. Zum einen die Verpflichtung zur gemeinsamen Aktivität (commitment to joint activity) und zum anderen die Verpflichtung zur gegenseitigen Unterstützung (commitment to mutual support). Aus diesen beiden Charakteristiken formulieren wir unsere Definition von räumlich-kooperativem Verhalten für einen humanoiden Agenten wie folgt:

Definition 5 *Räumlich-kooperatives Verhalten:* *Räumlich-kooperatives Verhalten zeichnet sich durch folgende Charakteristiken aus: Die Verpflichtung zur gemeinsamen räumlichen Aktivität in einem mit einem Partner geteilten Raum. Die Verpflichtung zur gegenseitigen Unterstützung in den räumlichen Aktivitäten.*

In Abschnitt 2.3.1 wurden die bei Menschen vermuteten kognitiven Mechanismen der *räumlichen Perspektivenübernahme* und die darunterliegende *verkörperte Simulation* vorgestellt. Übertragen wir die Konzepte auf einen humanoiden Agenten, so könnte dieser das eigene künstliche Körperschema und den damit zusammenhängenden peripersonalen Raum verwenden, um mentale Rotationen durchzuführen und die räumliche Perspektive eines Partners zu ermitteln und zu verstehen. Auf Grundlage dieses Gedankens entwickeln wir folgenden technischen Modellierungsansatz: ein humanoider Agent verwendet sein eigenes künstliches Körperschema, um eine Hypothese

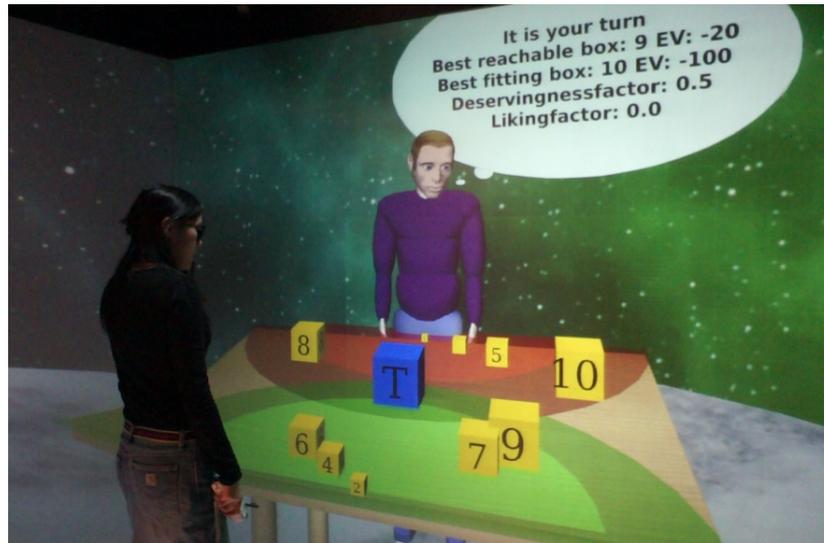


Abbildung 4.1: Interaktionsraum-Szenario

über die körperlichen Eigenschaften eines Partners automatisch zu generieren. Dieser Prozess entspricht somit der verkörperten Simulation. Um eine Hypothese über die räumliche Perspektive des Partners bilden zu können, schreibt der künstliche Agent seinen eigenen peripersonalen Raum dem Partner zu. Dieser Prozess entspricht der räumlichen Perspektivenübernahme. Auf dieser Grundlage entwickeln wir für interpersonale Interaktion die Definition eines Interaktionsraums wie folgt:

Definition 6 *Interaktionsraum:* *Der Interaktionsraum ist die überlappende Raumregion zweier peripersonaler Räume.*

Der Einsatz der hier vorgestellten Konzepte wird in Kapitel 7 in einer Nahraumaufgabe demonstriert (siehe Abbildung 4.1).

4.2 Max – ein virtueller Mensch

Die in dieser Arbeit vorgeschlagenen Lösungen für die Teilprobleme wurden für den virtuellen Menschen Max umgesetzt. Max besteht aus verschiedenen Softwarekomponenten. Für die Integration der in dieser Dissertation vorgestellten Ansätze wurden bestehende Komponenten erweitert und neue Komponenten in das Max zugrunde liegende technische Framework integriert. An dieser Stelle sollen daher zur besseren

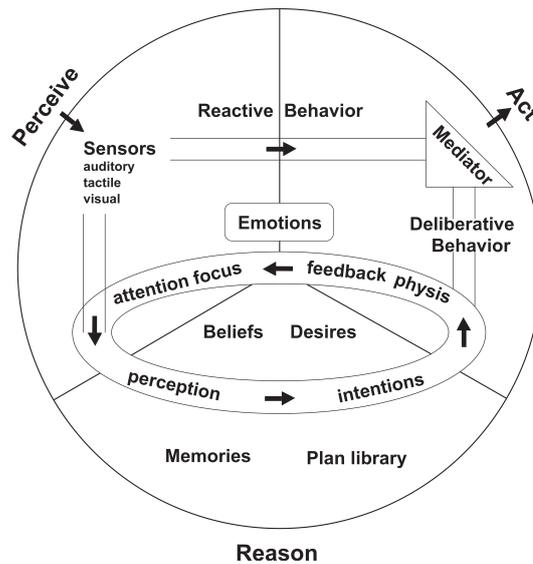


Abbildung 4.2: (Quelle: Lessmann, Kopp und Wachsmuth, 2006, S. 302)

Einordnung zunächst die für diese Arbeit relevanten existierenden Komponenten beschrieben werden.

Der virtuelle Mensch Max wurde in der Arbeitsgruppe *Wissensbasierte Systeme (Künstliche Intelligenz)* an der Universität Bielefeld – ursprünglich im Kontext des Sonderforschungsbereiches 360 *Situierte künstliche Kommunikatoren* – als multimodaler Konstruktionsexperte (**M**ultimodal **A**ssembly **eX**pert, MAX) entwickelt. In Abschnitt 3.4 wurde das im Sonderforschungsbereich entwickelte Anwendungsszenario vorgestellt. Der virtuelle Mensch wurde seit seiner Entstehung in zahlreichen Forschungsprojekten eingesetzt und weiterentwickelt. Jede Gliedmaße in Max' virtuellem Körper entspricht einer computergrafisch dargestellten 3-D-Geometrie, und Körperbewegungen werden über Skelettanimation mit einem kinematischen Skelett (s. Abbildung 4.3) gesteuert. Mit der Arbeit von Kopp wurde die natürliche Synchronisation zwischen Max' Sprachäußerungen und der Gestenanimation realisiert. Mit der Arbeit von Pfeiffer-Lessmann (2011) wurde eine umfangreiche **kognitive Architektur** für Max geschaffen. Max wird über diese Architektur kontrolliert (s. Abbildung 4.2). Die für kognitive Agenten üblicherweise angenommene Verarbeitungsschleife von Wahrnehmen (perceive), anschließendem Schlussfolgern bzw. Denken (reason) und dem resultierendem Handeln (act) wurde für Max übernommen. Die Verarbeitungsschleife wird jedoch nicht nacheinander abgearbeitet, sondern ist so konzipiert, dass alle Komponenten parallel ablaufen. So können schnelle, reaktive Verhaltensweisen (z. B. über

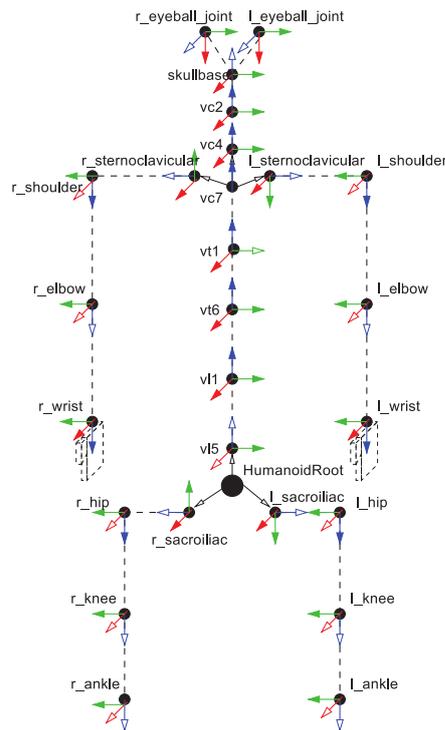


Abbildung 4.3: (Quelle: Kopp, Abb. 4.3, S. 100)

direkte Verknüpfungen von wahrgenommenen Stimuli mit Aktionen) und langsamere Verhaltensweisen auf Basis von Schlussfolgerungen für Max gleichzeitig generiert werden. Die Komponente der Architektur, die für das Schlussfolgern zuständig ist, basiert auf dem sogenannten Belief-Desire-Intention-Ansatz (BDI-Ansatz) (Rao und Georgeff, 1991), der es erlaubt, intentionale Aktionen zu modellieren und mit komplexen Aufgaben unter speziellen Kontextbedingungen umzugehen. Die Idee dahinter ist, den internen Zustand eines Agenten durch eine Menge von mentalen Kategorien zu beschreiben. Zu diesen Kategorien gehören die sogenannten *beliefs*, die das Wissen des Agenten über die Welt beinhalten, die *desires*, die die Zustände in der Welt darstellen, die der Agent erreichen will und die *intentions*, die jeweils aus einem aktuellen Ziel und einem zugehörigen Plan bestehen, um das aktuelle Ziel zu erreichen. Becker-Asano (2008) erweiterte Max um die Fähigkeit, eigene **Emotionen** zu simulieren. Basierend auf sowohl domänenspezifischen als auch domänenunabhängigen Bewertungsmechanismen (engl. appraisal mechanisms) werden emotionale Valenzwerte aus der Deliberationskomponente des BDI-Moduls abgeleitet. So wird zum Beispiel das Erreichen eines Ziels mit positiven Werten *belohnt*. Die Valenzwerte treiben Max' Emotionsdy-

namiken über die Zeit an. Das in dieser Arbeit entwickelte Körperschema ist in der kognitiven Architektur in der *act*-Komponente anzusiedeln, die peripersonale und die Interaktionsraum-Repräsentation in der *reason*-Komponente.

4.3 Sensorische Voraussetzungen für virtuelle Menschen

In Abschnitt 3.2.3 wurden Ansätze zur Generierung von künstlichen Körperschemata vorgestellt. Diese wurden ursprünglich für Roboterplattformen entwickelt. So wurde die Rekalibrierungsmethode von Hersch u. a. (2008) beispielsweise für visuelle und propriozeptive Sensordaten entwickelt und für die Hoap3-Roboterplattform realisiert. Die Übertragbarkeit solcher Methoden auf andere Roboter oder gar auf virtuelle Menschen muss daher genauer analysiert werden. Weitaus gravierendere Konsequenzen kann zudem die Übertragung auf andere Sensormodalitäten mit sich bringen. Die folgenden Überlegungen beschränken sich auf die relevantesten expliziten Körperschema-Lernverfahren (vgl. Abschnitt 4.1.2). Bevor also virtuelle Menschen wie Max von diesen Ansätzen profitieren können, müssen zunächst drei Fragen geklärt werden.

1. Die erste Frage betrifft, inwiefern es möglich ist, ein künstliches *Körperschema* nicht nur für physische Roboter, sondern auch *für virtuelle Menschen* zu ermitteln. Um diese zu untersuchen, muss das in Abschnitt 4.1.2 beschriebene künstliche Körperschema hinzugezogen werden. Dieses beinhaltet durch die explizite Repräsentation die Definition eines *kinematischen Skeletts*. Wie in Abschnitt 3.1 beschrieben, kann die Bewegungssteuerung virtueller Menschen auf verschiedene Arten erfolgen. Eine Voraussetzung für das Lernen eines künstlichen Körperschemas für virtuelle Menschen muss also eine auf der *Skelettanimation* beruhende Bewegungssteuerung über ein kinematisches Skelett sein. Da diese Methode weit verbreitet ist, erfüllt ein Großteil virtueller Menschen, so auch Max, die Voraussetzung, um ein künstliches Körperschema, zum Beispiel mit dem Verfahren von Hersch, zu ermitteln.

2. Eine weitere Frage ist, wie sich *virtuelle von physischen Sensordaten* unterscheiden. Anhand der visuellen Modalität können die wesentlichen Unterschiede zwischen Sensoren bei Robotern und bei virtuellen Menschen erläutert werden. Sensoren dienen grundsätzlich der Erfassung der Umgebung, in der sich ein humanoider Agent befindet. Die Roboter-Umgebung gibt bereits physikalische Gesetzmäßigkeiten und Beziehun-

gen zwischen Objekten vor. Die Erfassung erfolgt über Aufnahmegeräte wie Kameras, die die Auflösung der Sensordaten bestimmt. Es stellt sich dabei häufig das Problem verrauschter oder ungenauer Daten. Auf diesen müssen Bilderkennungsverfahren angewendet werden, um sinnvolle Informationen aus der physischen Welt zu extrahieren. Dieses Problem stellt sich in der virtuellen Welt üblicherweise nicht. Im Gegensatz zur physischen Realität, ist die virtuelle Realität eine künstliche Welt, in der die Schwierigkeit darin besteht, alle Objekte und ihre Beziehungen zueinander vordefinieren zu müssen. In der virtuellen Realität, wie die, in der Max eingebettet ist, liegen grafische Objekte in der Welt als dreidimensionale Gittermodelle vor. Alle grafischen Elemente, die in der virtuellen Welt sichtbar sind, werden intern in einem sogenannten Szenengraphen repräsentiert. Damit ein virtueller Mensch mit diesen Elementen interagieren kann, müssen Effekte, wie Kollisionen zwischen Objekten, vordefiniert werden. Im Gegensatz zur realen Welt finden physikalisch korrekte Ereignisse nicht automatisch statt, sondern müssen entweder durch entsprechende Softwaremodule, zum Beispiel über einen Kollisionserkennung oder eine sogenannte Physik-Engine erfasst werden. Dadurch entsteht der Vorteil, dass die Genauigkeit virtueller Sensordaten vom gewählten Detailgrad der grafischen Modellierung abhängt.

3. Die Frage, die sich abschließend stellt, ist, welche *Konsequenzen* diese *Unterschiede* für das *Lernverfahren* haben. Da virtuelle Sensoren genauere Informationen liefern können, ist anzunehmen, dass Lernverfahren, die auf diesen Daten beruhen, weniger fehleranfällig sind und genauere Ergebnisse erzielen können. Der Nachteil besteht jedoch grundsätzlich in der technischen Realisierung geeigneter Sensoren. Im Folgenden werden daher Max' taktile, propriozeptive und visuelle Wahrnehmung vorgestellt.

Virtueller Tastsinn

Max' virtuelle Tastsensoren und seine taktile Wahrnehmung wurden ursprünglich im Rahmen einer Diplomarbeit konzipiert (Nguyen, 2007; Nguyen u. a., 2007). Die Tastrezeptoren wurden für Max' gesamten virtuellen Körper entwickelt und bilden somit eine virtuelle Haut. Befunde zu dem menschlichen Tastsinn und der menschlichen Haut waren die Grundlage für die Entwicklung des virtuellen Tastsinns. Die künstliche Haut und der künstliche Tastsinn wurden dabei nicht nur im Hinblick auf den Einsatz bei virtuellen Menschen, sondern für den Einsatz bei humanoiden Agenten im Allgemeinen entworfen. Der Fokus wurde auf die sogenannte passive taktile Wahrnehmung gelegt.



Abbildung 4.4: Beispiele unterschiedlich großer, virtueller Tastrezeptoren für Max' virtuelle Haut.

Max konnte mit dieser Art der taktilen Wahrnehmung unterscheiden, welche Qualität (schlagen oder tippen) taktile Stimuli haben. Die Arbeit wurde ursprünglich für das 3-D-Computergrafik-Framework Avango für die virtuelle Realität umgesetzt. Mit der technischen Umstellung auf das modernere VR-Framework InstantReality¹ musste Max' virtueller Tastsinn für diese Arbeit in das neue Framework überführt werden. Max' Körperoberfläche ist mit einer virtuellen Haut bedeckt, die aus rechteckigen, flachen Geometrien gebildet wird. Die Geometrien bestehen aus einer geringen Anzahl von Polygonen, um eine einfache und schnelle grafische Darstellung und Verarbeitung zu gewährleisten. Jede Geometrie stellt einen Tastrezeptor dar. Durch variierende Größen kann die Anzahl an Tastrezeptoren an bestimmten Körperstellen vergrößert oder verringert werden, um eine geringere oder höhere taktile Auflösung zu erzielen. So ist bei Menschen die taktile Auflösung in den Fingern und Händen höher als auf dem Oberkörper (s. Abbildung 4.4). Wie in Abschnitt 2.1.1 beschrieben, werden nach Holmes und Spence (2004) die somatosensorischen Modalitäten im Menschen vermutlich in körperteilzentrierten Referenzsystemen kodiert. Dieser Aspekt ist für die Tastrezeptoren nachmodelliert worden. Das bedeutet, dass jeder virtuelle Tastrezeptor sowohl einer eindeutigen Gliedmaße als auch dem im kinematischen Skelett darunterliegenden Segment zugeordnet ist. Zudem ist jedem virtuellen Tastrezeptor eine eindeutige Bezeichnung (ID) zugeordnet.

Statt der verschiedenen Tastrezeptoren, wie sie in der menschlichen Haut existieren, ist es für die virtuelle Haut ausreichend, nur eine Art von virtuellem Tastrezeptor zu modellieren, da mit diesen bereits verschiedene Arten von taktilen Stimulationen unterschieden werden können. Jedes Objekt, das in der virtuellen Realität grafisch

¹<http://www.instantreality.org/> ; Letzter Zugriff: 27. Sep. 2012

dargestellt wird, kann so definiert werden, dass es einen taktilen Stimulus auf Max' virtueller Haut auslöst. In der hier vorgestellten Arbeit wurde die im VR-Framework InstantReality integrierte Kollisionsüberprüfung verwendet. Objekte, mit denen Max taktil in Berührung kommen kann, wurden darüber vordefiniert. Dadurch kann ein Kontakt zwischen einem Objekt und einem virtuellen Tastrezeptor erfasst werden und als taktiler Stimulus interpretiert werden. Kollisionen zwischen zwei Tastrezeptoren können auf diesem Wege ebenfalls erkannt werden. Dies ist notwendig, um taktilen Selbstkontakt für Max zu identifizieren. Für jeden taktilen Stimulus wird für den Zeitpunkt des Kontaktes, die eindeutige ID der beteiligten virtuellen Tastrezeptoren, ihre Zuordnung zu einer Gliedmaße, sowie die lokale Position im Koordinatensystem der Gliedmaße erfasst und an eine Softwarekomponente weitergeleitet, die alle Informationen über taktile Stimuli sammelt. Die Komponente kann als Max' *virtueller Tastsensor* bezeichnet werden.

Virtuelle Propriozeption und visuelle Wahrnehmung

Bei Menschen beschreibt die Propriozeption den Sinn für die Orientierung und die Positionen der Gliedmaßen im Raum. Dieser Sinn ist für die Wahrnehmung der eigenen Körperbewegungen und Körperposturen wichtig. Für humanoide Agenten ist es üblich, die Gelenkwinkelkonfiguration als propriozeptive Information zu interpretieren. Max' anthropomorphes kinematische Skelett besteht aus 57 Gelenken und 103 Freiheitsgraden (siehe Abbildung 4.3). Bei jeder Bewegung, die Max durchführt, werden die aktuellen Gelenkwinkel der beteiligten Gelenke ausgegeben.

In dieser Dissertation wird Max' virtuelles Sichtfeld dem menschlichen stereoskopischen Sichtbereich nachgeahmt. Dies entspricht ungefähr einem Öffnungswinkel von 60° von den Augen ausgehend (vgl. Abschnitt 2.1.3). Da unser Fokus nicht auf der visuellen Modalität liegt, werden keine Objekterkennungsmethoden benötigt. Zur Vereinfachung werden grafisch dargestellte Objekte als Punkte im Raum aufgefasst. Fallen diese in der virtuellen Realität in das virtuelle Sichtfeld, wird ihre Position in kopfzentrierten Koordinaten kodiert und als visueller Stimulus registriert.

4.4 Lösungsansatz

Das übergeordnete Ziel dieser Arbeit ist, humanoiden Agenten die Fähigkeit zu geben, auf natürliche und intelligente Weise Aufgaben im Nahraum mit menschlichen Partnern zu lösen. Auf Grundlage der in diesem Kapitel vorgeschlagenen technischen Interpretationen und Lösungen der einzelnen Teilprobleme kann der Gesamtlösungsansatz für das übergeordnete Ziel nun zusammengefasst werden. Abbildung 4.5 zeigt schematisch die Zusammenhänge zwischen den Teillösungen. Im Flussdiagramm sind auf der linken Seite die Teilprobleme aufgeführt. Die Pfeile zwischen den Komponenten beschreiben, wie die Teilprobleme gelöst werden. Die *Körperrepräsentation* wird über Sensorinformationen *gelernt*. Das Ergebnis sind die Gliedmaßenlängen und Gelenkpositionen. Diese Komponente wird konzeptuell als (künstliches) Körperschema zusammengefasst und bildet die Lösung zu Teilproblem P2. Die *Nahraum-Repräsentation und deren Wahrnehmung* wird aus der Körperrepräsentation *ermittelt*. Das Ergebnis sind das Wissen über die Greifreichweite und die Repräsentation von Objekten in körperzentrierten Koordinaten. Diese Komponente wird konzeptuell als peripersonaler Raum zusammengefasst und bildet die Lösung zu Teilproblem P1. Das *Interaktionsverhalten* wird mit Hilfe der Nahraum-Repräsentation und deren Wahrnehmung *generiert*. Das Ergebnis sind das Wissen über die Grenzen des geteilten Raums, eine Hypothese über die räumliche Perspektive des Partners und darauf basierendes räumlich-kooperatives Verhalten. Diese Komponente wird konzeptuell als Interaktionsraum zusammengefasst und bildet die Lösung zu Teilproblem P3.

Eine abstraktere Zusammenfassung der Zusammenhänge kann auf der rechten Seite des Flussdiagramms entnommen werden. Die Konzepte hängen somit hierarchisch voneinander ab. In dieser Arbeit wird ein künstliches Körperschema für einen humanoiden Agenten gelernt. Dieses bestimmt seinen peripersonalen Raum. Der peripersonale Raum wiederum wird verwendet, um den Interaktionsraum zu bilden.

5 Ein adaptives künstliches Körperschema für virtuelle Agenten

Die Körperrepräsentation ist die Informationsbasis für die peripersonale Raumrepräsentation und deren Wahrnehmung; diese Basis wird in diesem Kapitel über ein Lernverfahren ermittelt. Das Lernverfahren basiert auf dem von Hersch u. a. (2008) vorgestellten Rekalibrierungsverfahren. Herschs visueller Lernprozess wird in der hier vorgeschlagenen Lösung durch einen taktilen Lernprozess ersetzt. Dabei werden sensorische Informationen aus dem Tastsinn und der Propriozeption dazu verwendet, Max' kinematische Skelettstruktur zu lernen. Dieses Vorgehen bietet eine Lösung zur Problemstellung **P2 – Repräsentation des Körpers** und wird in Abschnitt 5.2 erläutert. In Kapitel 4 wurden bereits die Definition des in dieser Arbeit verfolgten künstlichen Körperschemas sowie die Methode, die dieses Körperschema erzeugen soll, eingeführt. Im Folgenden werden auf Grundlage der in Abschnitt 3.2.4 beschriebenen Lernmethode nach Hersch die Grenzen der visuell-propriozeptiven Rekalibrierung diskutiert und mit den Anforderungen an ein künstliches Körperschema verglichen, das als eine geeignete Basis für den peripersonalen Raum genutzt werden kann. In Abschnitt 5.2 wird dann das Lernverfahren entwickelt, das diesen Anforderungen entspricht. Anschließend werden die Resultate in Abschnitt 5.3 vorgestellt und in Abschnitt 5.4 diskutiert. Im letzten Abschnitt 5.5 werden die in diesem Kapitel gewonnenen Erkenntnisse zusammengefasst und ein Abschlussresümee gegeben.

5.1 Problemstellung

Um den Beitrag herausstellen zu können, den diese Arbeit zur Beantwortung der Frage nach einem adäquaten Lernverfahren für ein künstliches Körperschema leistet, ist es unumgänglich, die Leistung und die Grenzen der visuell-propriozeptiven Rekalibrierung im Detail zu beleuchten. Für ein Lernverfahren, das nicht nur technische Probleme, sondern gleichzeitig biologisch relevante Prozesse behandelt, müssen die

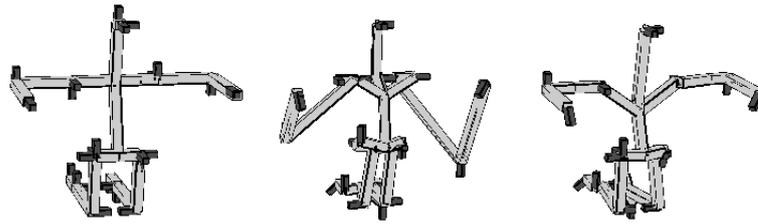


Abbildung 5.1: Die virtuelle Simulation der von Hersch verwendeten Roboterplattform Hoap3. Links das *tatsächliche* Körperschema, in der Mitte das subjektive Körperschema nach visueller Verfolgung der Endeffektoren und auf der rechten Seite ein subjektives Körperschema, wenn zusätzliche Referenzpositionen (Schultern, Knie, Ellbogen und Hüfte) simuliert wurden. (Quelle: Hersch u. a. (2008), Abb. 3.14, S. 49.)

Eigenschaften aus beiden Blickwinkeln analysiert werden.

5.1.1 Vergleich funktionaler und geometrischer Körperschemaeigenschaften

Es soll nun geklärt werden, wie es einem humanoiden Agenten möglich ist, aus verschiedenen sensorischen Eindrücken eine sinnvolle Repräsentation des eigenen Körpers im Raum, genauer im peripersonalen Raum, zu bilden. Die erste Anforderung an diese Repräsentation ist, die räumliche Umgebung in Referenzsystemen der verschiedenen Körperteile ausdrücken zu können. Mit Herschs visuell-propriozeptiver Rekalibrierung wird dies ermöglicht und erlernt. Eine weitere Anforderung an die Repräsentation ist, die Grenzen dieses Raums zu kennen. Für Letztere ist vor allen Dingen eine angemessene Repräsentation der Gliedmaßenlängen nötig. An diesem Punkt tritt in Herschs Modell eine Schwierigkeit auf.

Ist die Anzahl der Referenzpunkte für den Lernalgorithmus zu gering, können trotz korrektem Körperschema Körpergeometrien entstehen, die nicht den realen Geometrien des Roboters entsprechen. Bei Hersch werden zum Beispiel in einem Experiment nur die vier Endeffektoren (zwei Hände, zwei Füße) des Roboters visuell verfolgt. Für unterschiedliche Lerndurchläufe kann das Verfahren zwar möglicherweise die korrekten kinematischen Funktionen finden, die geometrischen Eigenschaften zwischen den Referenzpunkten (z. B. zwischen den beiden Händen) können sich aber voneinander unterscheiden. Abbildung 5.1 zeigt ein Beispiel von Hersch u. a. (2008), in dem die einzelnen Gliedmaßenlängen nicht korrekt abgebildet werden. Das geometrische Resultat

des Lernverfahrens kann demnach erheblich von den tatsächlichen Körpercharakteristiken des echten Roboters abweichen. Je mehr Referenzpositionen einbezogen werden, desto mehr repräsentiert das resultierende Körperschema den echten Roboterkörper. Hersch definiert für dieses funktional korrekte, aber geometrisch variierende Ergebnis den Begriff des **subjektiven Körperschemas**. Dieses soll hauptsächlich die Koordinatentransformationen zwischen den sensorischen Informationen korrekt abbilden. Liegen beispielsweise propriozeptive Informationen vor, so können mit Hilfe des subjektiven Körperschemas die dazugehörigen visuellen Informationen vorherbestimmt werden.

Die gängige Formulierung des biologischen Körperschemas beinhaltet üblicherweise neben den funktionalen Eigenschaften, wie der Koordinatentransformation, auch die Abbildung der Körpercharakteristiken, wie Gliedmaßenlängen und Positionen der Gelenke. Dieses bezeichnet Hersch daher als das **objektive Körperschema**. Das subjektive Körperschema behandelt demnach nur indirekt die realen Körpercharakteristiken und hängt stark von dem sensorischen System des humanoiden Agenten ab. Für die Bewegungssteuerung ist das subjektive Körperschema zwar ausreichend, für eine peripersonale Raumrepräsentation ist es jedoch notwendig zu wissen, welche Länge die einzelnen Gliedmaßen haben.

5.1.2 Anforderungen einer peripersonalen Raumrepräsentation

Zwar gelingt es Hersch u. a., mit dem Lernverfahren einen funktionalen Zusammenhang zwischen dem subjektiven visuellen und dem subjektiven propriozeptiven Raum (engl. *subjective sensory space*) herzustellen, die taktile Modalität wird jedoch nicht behandelt. Diese Modalität ist für das biologische Körperschema ebenso zentral wie die Propriozeption, bei humanoiden Agenten ist sie jedoch selten vorzufinden. Hersch selbst weist darauf hin, dass taktiler Lernen mit dem Adaptierungsalgorithmus möglich sein müsste. Jedoch standen auch ihm nur die visuelle und die propriozeptive Sensormodalität für seine Roboterplattform zur Verfügung. Bei physischen Robotern wie auch beim Menschen ergibt sich das Problem, dass der Roboter aufgrund des eingeschränkten visuellen Sichtfeldes nicht alle Endeffektorpositionen sehen kann, weil sie in bestimmten Konfigurationen durch andere Körperteile verdeckt sind (Hersch u. a., 2008). Dies erschwert die Erfassung der Sensordaten und somit den Lernprozess selbst. Eine künstliche taktile Haut, die den gesamten humanoiden Körper bedeckt, wäre von

diesem Problem weniger betroffen.

Eine weitere Schwierigkeit des visuellen Lernens ist die Art, wie visuelle Referenzpunkte am Körper bestimmt werden. Visuelle Referenzpositionen müssen entweder manuell am Körper über Marker festgelegt oder aber über aufwendigere Bildsegmentierungsverfahren ermittelt werden, um festzustellen, welche Bildpositionen zum Körper gehören und visuell verfolgt werden müssen. Daher ist der wohl wichtigste Grund, den Einsatz taktiler Daten in den Lernprozess einzubinden, die Möglichkeit, eine große Anzahl über den Körper verteilter Referenzpunkte nutzen zu können. Über taktile Daten wäre es also möglich, das *subjektive* Körperschema dem *objektiven* anzunähern und eine geeignete Repräsentation des peripersonalen Raums zu erreichen.

Im Folgenden wird vorgestellt, welche Änderungen an Herschs Lernverfahren vorgenommen wurden, zudem werden die Ergebnisse und der geleistete Beitrag zum Stand der Forschung diskutiert.

5.1.3 Ein Körperschema-Lernverfahren

Zum besseren Verständnis von Herschs Rekalibrierungsmethode unterteilen wir diese **Lernmethode** im Folgenden in einen **Lernprozess** und in den darin ausgeführten **Lernalgorithmus**.

In unserem Fall ist die Lernmethode das Rekalibrierungsverfahren von Hersch. Die Lernmethode benötigt die Skeletthierarchie des humanoiden Agenten, die Möglichkeit, eine geschlossene kinematische Kette darin zu finden, und eine Position innerhalb der geschlossenen kinematischen Kette, die in verschiedenen Referenzsystemen dargestellt ist. Das Verfahren kann über verschiedene Sensormodalitäten Informationen annehmen. Der in dem Lernverfahren stattfindende Lernprozess nimmt die Sensordaten an und verarbeitet sie weiter. Lernprozesse innerhalb der Methode unterscheiden sich darin, wie die Sensordaten erhoben und weiterverarbeitet werden. In unserem Fall ist dies die Unterscheidung zwischen dem **visuellen** und dem **taktilen Lernprozess**. Je nachdem, welche Sensordaten also in der Methode vorkommen, werden unterschiedliche Lernprozesse angestoßen. Beim visuellen Lernprozess werden die relevanten Positionen, wie Körperteile oder Objekte, im Kamerabild aufgenommen. Beim taktilen Lernprozess werden die Daten aufgenommen, die bei einem taktilen Selbstkontakt zustande kommen, das heißt, wenn der humanoide Agent sich selbst berührt hat. Die einzelnen Lernprozesse verarbeiten diese Sensordaten und berechnen damit

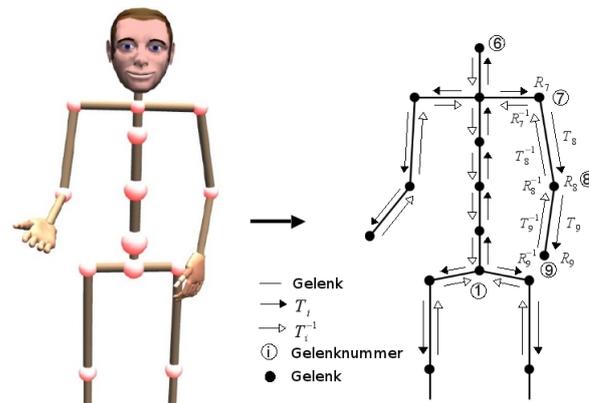


Abbildung 5.2: Der kinematische Baum (rechte Seite), der Max' kinematisches Skelett (linke Seite, Quelle: AG Wissensbasierte Systeme) repräsentiert. Die Gelenkziffern sind willkürlich gewählt. Ein Beispiel für eine Koordinatentransformation ist die folgende Komposition, die ein Referenzsystem an Gelenk 7 in ein Referenzsystem an Gelenk 6 transformiert (Gelenk 5 liegt zwischen ihnen) $\mathbf{R}_6^{-1} \circ \mathbf{T}_6^{-1} \circ \mathbf{T}_7 \circ \mathbf{R}_7$.

ein künstliches Körperschema mit Hilfe eines Lernalgorithmus. In diesem werden je nach Lernprozess unterschiedliche Formeln und Berechnungsschritte angewendet.

In dieser Arbeit wird auf Grundlage von Herschs Lernmethode demnach der **taktile Lernprozess** als Alternative zum **visuellen Lernprozess** realisiert; darüber hinaus werden Erweiterungen für den **Lernalgorithmus** vorgeschlagen und umgesetzt.

5.2 Taktiler Lernen des Körperschemas

Unter Beachtung der im vorherigen Abschnitt diskutierten Konsequenzen der visuell-propriozeptiven Rekalibrierung wird nun vorgestellt, wie ein Körperschema für den virtuellen Menschen Max gelernt wird. Wir zeigen, wie durch den Einsatz einer alternativen Sensormodalität und unter gezielten Erweiterungen des Lernverfahrens die Leistung der Rekalibrierungsmethode verbessert wird. In das Verfahren fließen sowohl taktile als auch propriozeptive Sensorinformationen von Max' virtuellem Körper ein. Dadurch werden akkuratere Resultate für das resultierende Körperschema erzielt.

5.2.1 Initialisierung

Es wird zunächst festgelegt, welche Teile des kinematischen Skeletts gelernt werden sollen. Abbildung 5.2 zeigt schematisch den gewählten kinematischen Baum $KineBaum_{Max}$.

Dieser bildet nur die Gelenkstruktur einer Teilmenge aus Max' tatsächlichem kinematischem Skelett ab (s. Abb. 4.3). Da das künstliche Körperschema eine Grundlage für Max' peripersonalen Raum bilden soll, genügt es, die Oberkörpergelenke zu wählen. Dazu gehören elf Gelenke: diejenigen Gelenke, die die beiden Endeffektoren miteinander verbinden, sowie die Gelenke bis zur Hüfte. Die kinematischen Ketten innerhalb des kinematischen Baums enden somit jeweils an den beiden Handgelenken und am Hüftgelenk, das auch gleichzeitig das Wurzelgelenk von Max' kinematischem Skelett bildet. Die Gelenke in den Fingern wurden nicht mit einbezogen. Diese umfassen pro Hand 16 Gelenke – mehr als die restlichen Oberkörpergelenke zusammen. Ihre Berechnung würde einen erheblichen Anteil der Berechnungsiterationen ausmachen, aber nur einen verhältnismäßig geringen Anteil zur Ermittlung der Gesamtlänge eines Armes beitragen. Um diese ineffektiven Berechnungen zu vermeiden, wird ein Standardwert für die Länge der Hand angenommen. Sie entspricht ungefähr der Länge einer geschlossenen Hand.

Kinematischer Baum für virtuelle Tastrezeptoren

Im kinematischen Baum $KineBaum_{Max}$ entspricht jeder Knoten einem Gelenk in Max' kinematischem Skelett, und jede Kante entspricht einem Körpersegment zwischen zwei Gelenken (s. Abb. 5.2). Der kinematische Baum beschreibt also im Unterschied zum kinematischen Skelett nur die Anzahl der Gelenke, ihre hierarchische Reihenfolge sowie die Anzahl der Körpersegmente, kurz Segmente. Die Gelenkpositionen und -orientierungen bildet er nicht ab. Die virtuellen Tastrezeptoren sind an die Gliedmaßen angefügt (s. Abschnitt 4.3). Ihre Positionen sind im lokalen Referenzsystem des jeweiligen Körperteils repräsentiert. Innerhalb des kinematischen Baums sind die Tastrezeptoren also entlang der Kanten repräsentiert.

Wie in Abschnitt 3.1.1 im Detail erläutert wurde, beschreibt jedes Gelenk in einer kinematischen Kette ein eigenes Koordinatensystem. Im kinematischen Baum kann daher jeder Übergang zwischen zwei Gelenken durch eine Koordinatentransformation definiert werden, die das Koordinatensystem des einen Gelenks in das des anderen transformiert. Deshalb beschreibt jeder Pfad innerhalb des kinematischen Baums, der ein Gelenk mit einem anderen verbindet, eine kinematische Kette. Da die Tastrezeptoren an den Gliedmaßen anliegen, können die Positionen der Tastrezeptoren, die im Referenzsystem der jeweiligen Gliedmaße vorliegen, in eine Position eines beliebigen

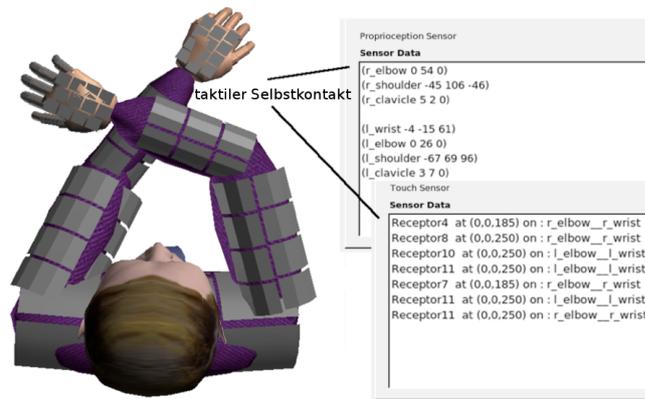


Abbildung 5.3: Taktiler Körperschema-Lernen: Bei jedem taktilen Selbstkontakt werden die sensorischen Informationen der Sensormodalitäten ausgegeben. Der virtuelle Tastsinn stellt für jeden aktivierten Tastrezeptor eine eindeutige ID, die Gliedmaße, an der er angebracht ist, und die Position im Referenzsystem der Gliedmaße aus. Winkelwerte für die relevanten Gelenke werden als propriozeptive Informationen ausgegeben.

anderen Referenzsystems einer Gliedmaße transformiert werden. Max' kinematisches Skelett legt dabei die Hierarchie der Transformationen fest. Dadurch wird bestimmt, ob eine direkte oder eine inverse Koordinatentransformation entlang der Kette durchgeführt werden muss. Abbildung 5.2 zeigt die Transformationshierarchie des kinematischen Baums mit dem Hüftgelenk als Wurzelgelenk. Folgen wir einer Kante in Richtung des Wurzelgelenks (Gelenk 1), so muss eine Translation \mathbf{T}_i und eine Rotation \mathbf{R}_i für das entsprechende Gelenk i durchgeführt werden. Folgen wir der Kante in die entgegengesetzte Richtung, müssen die inverse Translation \mathbf{T}_i^{-1} und die inverse Rotation \mathbf{R}_i^{-1} durchgeführt werden.

Um das künstliche Körperschema lernen zu können, müssen Anfangswerte für die zu adaptierenden Parameter gewählt werden. Alle Gelenkpositionen \mathbf{l}_i werden auf den Nullvektor gesetzt, das heißt, alle i Gelenke liegen zu Anfang in einem Punkt. Die Rotationsachsen \mathbf{a}_i werden mit zufälligen Werten initialisiert.

5.2.2 Taktiler Lernprozess

Das initiale künstliche Körperschema, wie es im vorherigen Abschnitt 5.2.1 definiert und initialisiert wurde, kann nun durch den Lernalgorithmus angepasst werden. Zuvor müssen jedoch geeignete Sensordaten bereitgestellt werden. Einen Überblick über den gesamten taktilen Lernprozess gibt der Algorithmus 1.

Die Sensordaten werden in einem Body-Babbling-Schritt erzeugt, bei dem Max zufällige Bewegungen mit seinen Armen durchführt (Zeile 2–4). Die erzeugten taktilen Sensordaten werden zusammen mit den propriozeptiven Daten zur Adaptierung der Gelenkachsen und -positionen aktiver kinematischer Ketten innerhalb des kinematischen Baums $KineBaum_{Max}$ verwendet (Zeile 5–15). Die Adaptierung der Gelenkpositionen erfolgt dabei zunächst über den Gradientenabstieg, wie Hersch ihn vorschlägt (Zeile 18–21). Darüber hinaus kann optional eine Nach-Adaptierung durch eine in dieser Arbeit entwickelte Adaptierungsmethode erfolgen. Ob Letztere verwendet werden kann, hängt von den Tastreferenzpositionen ab (Zeile 22–27).

Der gesamte Algorithmus wird bei jedem taktilen Selbstkontakt einmal ausgeführt. Dies wird so lange iterativ durchgeführt, bis der Fehler $\| \mathbf{v}_{trans,n} - \mathcal{T}_{schätz}(\mathbf{v}_{orig,n}) \|$ aller Gelenkachsen und -positionen des gesamten kinematischen Baums $KineBaum_{Max}$ über eine gewählte Anzahl Iterationen genügend klein geblieben ist (Zeile 28).

Erfassung der Sensordaten

Der erste Schritt im taktilen Lernprozess ist zunächst die Datenerfassung aus den Sensormodalitäten (Zeile 2–12). Sobald mindestens zwei Tastrezeptoren gleichzeitig einen Stimulus registrieren, wird die Bewegung angehalten. Dies verhindert, dass Max durch die eigenen virtuellen Gliedmaßen hindurchfasst. Die Tastrezeptoren arbeiten passiv und lokal und können selbst keine Informationen darüber liefern, was die Quelle des Stimulus war. Daher wird die Annahme getroffen, dass, wenn Tastrezeptoren zu einem gleichen Zeitpunkt $t_{SelbstKontakt}$ aktiviert sind, sie sich gegenseitig berührt haben. Wir gehen in diesem Fall von einem taktilen Selbstkontakt aus. Mögliche externe Quellen werden also nicht berücksichtigt. Bei mehr als zwei aktivierten Tastrezeptoren wird pro Zeitpunkt immer nur ein Paar, bestehend aus den Tastrezeptoren $\{TR_1; TR_2\}$, zufällig ausgewählt. Für das Paar $\{TR_1; TR_2\}$ muss ermittelt werden, welche kinematische Kette innerhalb des kinematischen Baums $KineBaum_{Max}$ es miteinander verbindet. Da im Unterschied zum visuellen Lernprozess die zu verfolgenden Referenzpunkte nicht vorher festgelegt werden, variieren die zu adaptierenden kinematischen Ketten während des Lernprozesses. Die kinematische Kette wird als aktuell ermittelte kinematische Kette $kineChain_{aktiv}$ vorgehalten. Abbildung 5.4 zeigt zwei aktivierte kinematische Ketten mit unterschiedlicher Länge bei taktilen Selbstkontakt. Der virtuelle Tastsinn gibt für die Tastrezeptoren eine eindeutige Kennnummer (ID) aus,

Algorithmus 1 Pseudo-Code: Taktiler Lernprozess

```

1: repeat
2:   for all  $\{TorsioGelenke\}$  do
3:     generiere zufällige Bewegungen
4:   end for
5:   if  $\{TR_k, TR_l\} \leftarrow$  taktilerSelbstkontakt then
6:     BewegungStoppen
7:      $pos_k \leftarrow$  HoleTastrezeptorPosition( $TR_k$ )
8:      $pos_l \leftarrow$  HoleTastrezeptorPosition( $TR_l$ )
9:      $j_n \leftarrow$  HoleGelenkVonRezeptor( $TR_k$ )
10:     $j_m \leftarrow$  HoleGelenkVonRezeptor( $TR_l$ )
11:     $kineChain_{aktiv} \leftarrow$  HoleKinematischeKette ( $StartKnoten \leftarrow j_m, EndKnoten$ 
     $\leftarrow j_n$ )
12:   end if
13:   for  $g = startnode$  to  $endnode$  do
14:      $\{aktuelleWinkel\} \leftarrow$  HoleWinkelAusGelenk( $g$ )
15:   end for
16:   Setze Transformation  $T \leftarrow kineChain_{aktiv}$ 
17:    $pos_l = T ( pos_k )$ 
18:   for  $k = startnode$  to  $endnode$  do
19:      $\Delta l_i \leftarrow$  BerechneAdaptierungGradientenabstieg ( $pos_k, pos_l, kineChain_{aktiv}$ )
20:      $\Delta a_i \leftarrow$  BerechneAdaptierungGradientenabstieg ( $pos_k, pos_l, kineChain_{aktiv}$ )
21:   end for
22:   for  $tr = \{TR_k, TR_l\}$  do
23:     if  $Länge(tr) > Länge(l_{segmentRezeptor+1})$  then
24:        $segmentRezeptor \leftarrow$  HoleSegmentVonRezeptor( $tr$ )
25:        $\Delta l_{segmentRezeptor+1} \leftarrow$  BerechneAdaptierungTaktileKomponente()
26:     end if
27:   end for
28: until ( $pos_j - T(pos_i)$ )  $< SchwellwertKlein_{Fehler}$ 

```

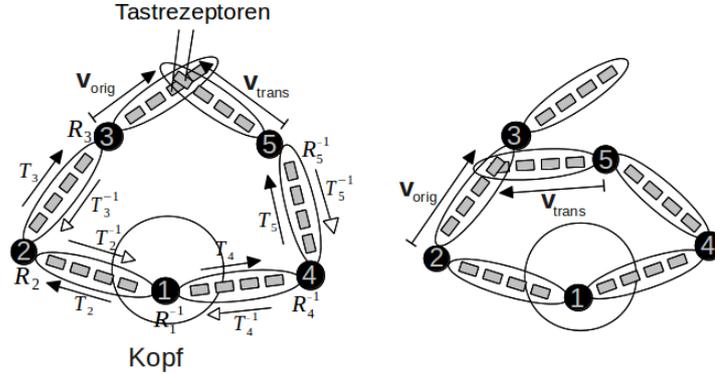


Abbildung 5.4: Schematische Darstellungen von taktilem Selbstkontakt an Max' Körper (Notation wie in Abb. 5.2). Linke Seite: Über eine Koordinatentransformation wird die Position \mathbf{v}_{orig} , die im Koordinatensystem von Gelenk 3 vorliegt, in das Koordinatensystem des Gelenks 5 transformiert. Diese Beispieltransformation ist gegeben durch: $\mathbf{R}_5^{-1} \circ \mathbf{T}_5^{-1} \circ \mathbf{R}_4^{-1} \circ \mathbf{T}_4^{-1} \circ \mathbf{T}_2 \circ \mathbf{R}_2 \circ \mathbf{T}_3 \circ \mathbf{R}_3$. Rechte Seite: Beispiel eines taktilem Selbstkontakts, bei dem die kinematische Kette kürzer ist.

die Gliedmaße, an denen der Rezeptor angefügt ist, sowie die Position im Koordinatensystem der Gliedmaße. Die Abbildung 5.4 (linke Seite) zeigt ein Beispiel für eine kinematische Kette, die zwei taktile Sensoren miteinander verbindet. Koordinatentransformationen entlang dieser Kette überführen eine Tasterezeptorposition \mathbf{v}_{orig} , die im Referenzsystem von Gelenk 3 vorliegt, in das Referenzsystem von Gelenk 5. Wird diese kinematische Kette in die entgegengesetzte Richtung verfolgt, so wird die Tasterezeptorposition \mathbf{v}_{trans} , die im Referenzsystem von Gelenk 5 vorliegt, in das von Gelenk 3 transformiert. Für die Messung der propriozeptiven Sensorinformationen wird die zum Zeitpunkt $t_{SelbstKontakt}$ resultierende Körperpostur verwendet. Hierfür werden alle Gelenkwinkel der kinematischen Kette $kineChain_{aktiv}$ abgefragt. Die propriozeptiven Informationen müssen für den Algorithmus in Achsen-Winkel Darstellung vorliegen.

Adaptierungsmethode über Gradientenabstieg

Nach der Datenerfassung findet die eigentliche Adaptierung über den Lernalgorithmus statt (ab Zeile 13). Nachdem die Sensordaten, bestehend aus den lokalen Positionen der Tasterezeptoren sowie den aktuellen Gelenkwinkelwerten, für den taktilem Selbstkontakt erfasst wurden, findet die eigentliche Adaptierung über den Lernalgorithmus statt (ab Zeile 13). Dabei werden die Gelenkachsen \mathbf{a}_i und die Gelenkpositionen \mathbf{l}_i der aktiven kinematischen Kette $kineChain_{aktiv}$ verändert. Aus den Sensordaten wer-

den mit den von Hersch entwickelten Formeln 3.5 und 3.4 über das Gradientenabstiegsverfahren Änderungswerte berechnet. Welche der beiden lokalen Positionen (im Pseudocode1 $pos_{k,l}$) der Tastrezeptoren als \mathbf{v}_{orig} und welche als \mathbf{v}_{trans} in die Formeln eingehen, ist dabei nicht wichtig und kann zufällig gewählt werden. Mit diesen Werten wird versucht, den aktuellen Fehler der Transformationen zu verkleinern, indem die aktuellen Schätzungen der Gelenkachsen und -positionen damit adaptiert werden. Die aktualisierte $kineChain_{aktiv}$ bildet dann die Grundlage für die Transformationen im nächsten Iterationsdurchlauf mit neuen Sensordaten.

Adaptierungsmethode für Gelenkpositionen-Startwerte

Nachdem die Gelenkachsen und -positionen bereits einmal adaptiert wurden, kann je nach Rezeptorposition eine Nach-Adaptierung erfolgen.

Analyse: Betrachtet man die Möglichkeiten des taktilen Selbstkontakts, so besteht ein Unterschied darin, ob die Rezeptoren nah an einem Gelenk liegen oder weiter davon entfernt sind. Abbildung 5.5 illustriert die Fälle. In Abbildung 5.5 (linke Seite) besteht der taktile Selbstkontakt zwischen zwei Tastrezeptoren $\{TR_1; TR_2\}$, die auf zwei Gelenken sitzen. Dies ist die minimale lokale Position, die ein Tastrezeptor einnehmen kann. Es ist somit nicht möglich, Tastrezeptorpositionen am Gelenk zu verwenden, um die nächste Gelenkposition zu bestimmen. Zudem gibt es kein Wissen darüber, ob ein Tastrezeptor auf einem Gelenk sitzt oder nicht, da nur die lokale Position bekannt ist. In Abbildung 5.5 (Mitte) besteht der taktile Selbstkontakt zwischen zwei Tastrezeptoren $\{TR_1; TR_2\}$, die jeweils entlang der Gliedmaße angebracht sind. Die Tastrezeptoren liegen also mit gewissem Abstand vom Koordinatenursprung auf den Gliedmaßen. Die maximale Position, die ein Tastrezeptor annehmen kann, liegt am Ende des Segments, kurz vor dem nächsten Gelenk in der kinematischen Kette.

Da Tastrezeptoren auf der Oberfläche der virtuellen Gliedmaßen angebracht sind, stellt ihre lokale Position im Ursprungskoordinatensystem des jeweiligen darunterliegenden starren Segments immer einen Beitrag zur Länge dieses Segments dar. Abbildung 5.6 zeigt die Anordnung der virtuellen Tastrezeptoren auf Max' Gliedmaßen. Sie liegen nicht genau auf den Segmenten, sondern bedecken die Gliedmaßen, die formal durch die Segmente beschrieben werden. Das bedeutet, dass nicht die gesamte Position des Tastrezeptors einen Beitrag zum Segment erbringt, sondern lediglich helfen kann, eine Schätzung für die nächste Gelenkposition zu finden. Im Gegensatz zur Ad-

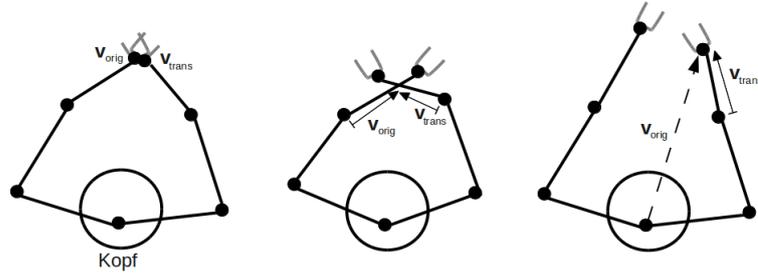


Abbildung 5.5: Verschiedene Selbstkontakte und die möglichen lokalen Tastrezeptorpositionen (linke Seite und Mitte) im Vergleich zu visuellen Beiträgen der Referenzpositionen (rechte Seite). Linke Seite: Tastrezeptoren liegen in der minimalen Position ihres lokalen Koordinatensystems. Mitte: Tastrezeptoren liegen mit gewissem Abstand zum Ursprung ihres lokalen Koordinatensystems entlang des jeweiligen Segments.

aptierungsmethode über Gradientenabstieg, wie sie von Hersch vorgeschlagen wird, resultiert die Schätzung nicht aus der Verkleinerung eines Transformationsfehlers. Die hier vorgeschlagene Adaptierungsmethode nutzt stattdessen Sensorinformationen über die tatsächlichen Körpercharakteristiken. Die Tastrezeptoren enthalten diese Informationen, weil sie die Körperoberfläche abdecken und zum Körper gehören.

Berechnung der Nach-Adaptierung: Für die Tastrezeptoren liegen nur die lokalen Positionen im Koordinatensystem der Gliedmaßen vor. Nehmen wir an, dass bei einem Selbstkontakt für einen taktilen Rezeptor TR_i die Position als \mathbf{tr}_i vorliegt. Der reelle Beitrag einer Tastrezeptorposition \mathbf{tr}_i zu einem Segment \mathbf{l}_{i+1} ist die orthogonale Projektion $\mathbf{tr}_{i\mathbf{l}_{i+1}}$ (siehe Abbildung 5.6, gestrichelte Linie). Das Segment \mathbf{l}_{i+1} ist jedoch nicht bekannt und soll durch gute Schätzungen adaptiert werden. Da wir kein Wissen über das kinematische Skelett, sondern lediglich Vorwissen über den kinematischen Baum voraussetzen, können wir keine Vorannahmen über die Koordinatenachsen treffen. Wir wissen demnach nur, dass sowohl \mathbf{l}_{i+1} als auch \mathbf{tr}_i in demselben Koordinatensystem dargestellt sind und dass \mathbf{tr}_i einen Beitrag zu Segment \mathbf{l}_{i+1} darstellt, aber nicht welchen. Der Vektor \mathbf{tr}_i kann daher als neuer Startwert $\mathbf{tr}_{\text{schätz}}$ für die Segmentadaptierung in Richtung der nächsten Gelenkposition verwendet werden. Da der Schätzwert eine Annäherung für die Absolutposition des nächsten Gelenkes darstellt, muss vor jeder Adaptierung geprüft werden, ob diese einen Zuwachs für die aktuell geschätzte Gelenkposition darstellt. Dies verhindert, dass die Schätzung durch die Nach-Adaptierung in die Richtung des Ursprungs koordinatensystems gesetzt wird. Innerhalb des Lernalgorithmus werden also zusätzlich die lokalen Rezeptorpositionen des Tastrezeptorenpaars $\{TR_1; TR_2\}$ auf ihre Länge überprüft (Zeile 22). Ist diese

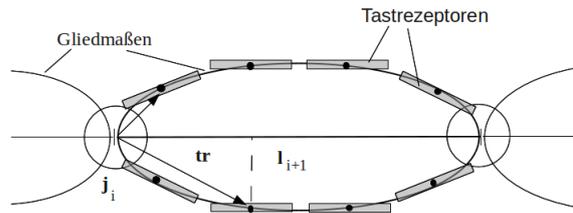


Abbildung 5.6: Lage von Tastrezeptoren auf virtuellen Gliedmaßen. Die Tastrezeptorpositionen \mathbf{tr} werden durch einen gewählten Mittelpunkt spezifiziert (schwarze Punkte in Tastrezeptoren) und liegen im Koordinatensystem des zugeordneten Gelenks j_i vor. Jede \mathbf{tr} besitzt einen Anteil an der nächsten Position von Gelenk j_{i+1} , gegeben durch den Vektor \mathbf{l}_{i+1} (Segment). Die orthogonale Projektion von \mathbf{tr} auf das Segment ist gestrichelt dargestellt.

größer als die aktuelle Segmentlänge, wird die Tastrezeptorposition \mathbf{tr}_1 bzw. \mathbf{tr}_2 als neuer Wert übernommen. Dadurch findet die Nach-Adaptierung für die Position von Gelenk j_{i+1} statt, die in der kinematischen Kette auf das Gelenk j_i folgt. In Gelenk j_i liegt der Koordinatenursprung des aktuellen Schätzwertes. Es ist wichtig zu beachten, dass diese Adaptierungsmethode nur eine Ergänzung zur Adaptierungsmethode von Hersch darstellt. Es ist nicht möglich, mit dieser Methode Schätzungen für die Gelenkachsen vorzunehmen. Mit dieser Methode kann der Schätzwert einmalig pro Segment als neuer Initialwert für die Adaptierung verwendet werden. Wie in Abschnitt 5.2.1 beschrieben, werden bei der Initialisierung alle Gelenkpositionen auf den Nullvektor gesetzt. Eine bessere Schätzung für das initiale künstliche Körperschema könnte auf diese Weise schneller zu Ergebnissen führen. Als Alternative kann auch bei jedem Iterationsdurchlauf eine Nach-Adaptierung durchgeführt werden. Die Ergebnisse werden im nächsten Abschnitt gezeigt und diskutiert.

5.3 Evaluation

Die Adaptierungsformeln, wie sie in Abschnitt 3.2.4 vorgestellt wurden, konnten wie beschrieben implementiert und in den Lernalgorithmus übernommen werden. Dies verdeutlicht die gute Übertragbarkeit von Herschs Arbeit. Der vorgeschlagene taktile Lernprozess bringt verschiedene Vorteile mit sich. So können über die Tastrezeptoren unterschiedlich lange kinematische Ketten im Lernprozess gleichzeitig adaptiert werden, ohne sie, wie bei dem visuell-propriozeptiven Lernprozess, vorannotieren zu müssen (vgl. Abbildung 5.1). Der visuelle Lernprozess berechnet bei wenigen Referenzpositionen ein Körperschema mit Gelenkpositionen, die teilweise von den echten

Körpergeometrien des humanoiden Agenten abweichen. Die Anzahl der Referenzpositionen über taktile Sensoren, wie sie bei Max vorhanden sind, ist dagegen höher. Diese Anzahl visuell zu simulieren ist nur mit hohem Aufwand möglich. Ein weiterer Punkt ist, dass Referenzpositionen, die visuell, zum Beispiel aufgrund von Verdeckung, nicht erfasst werden können, durch Tastrezeptoren wahrgenommen werden können.

Diese Vorteile werfen folgende Hypothesen auf:

1. Je mehr taktile Referenzpositionen verfügbar sind, desto *akkurater* wird das resultierende Ergebnis für die Gelenkpositionen.
2. Im taktilen Lernprozess generierte kurze kinematische Teilketten verringern die Anzahl der Iterationsdurchläufe.

Für die Evaluation des taktilen Lernprozesses wurden verschiedene Berechnungsarten miteinander verglichen und auf zwei Eigenschaften evaluiert: zum einen die *Iterationsanzahl* bis zur Konvergenz und zum anderen die *Akkuratheit* des resultierenden künstlichen Körperschemas im Vergleich zum tatsächlichen Körperschema. Je näher die berechneten Werte für die Gelenkpositionen an den tatsächlichen liegen, desto akkurater ist das Ergebnis. Im Folgenden bezeichnen wir diese Abweichung auch als eine Abweichung des *subjektiven* von dem *objektiven* Körperschema. Diese wird als Summe der Abweichungen aller Gelenkpositionen in Millimetern dargestellt.

Jede Berechnungsart wurde in jeweils 500 Simulationen getestet und mit 500 Simulationen des visuellen Lernprozesses verglichen. Es wurde bei jeder Simulation die längste mögliche kinematische Kette in Max' Torso adaptiert, und dieselben ϵ -Werte wurden für die Achsen- bzw. Gelenkpositionsadaptierung verwendet. Beginnend von Max' Wurzelgelenk (*vl_5*) bis zu einem der Handgelenke (*l_wrist* bzw. *r_wrist*) wurden sieben Gelenke und Segmente gelernt. Die Messbasis für jede der taktilen Berechnungsarten ist Herschs Ergebnis für die Berechnung von kinematischen Ketten mit sieben Freiheitsgraden. Im Durchschnitt benötigt der visuelle Lernprozess über 70.000 Iterationen. Eine Simulation gilt als *konvergiert*, wenn bei über 500 hintereinander folgenden Iterationen mit unterschiedlichen Konfigurationen ein Schwellwert für den Fehler $\| \mathbf{v}_{trans,n} - \mathcal{T}_{schätz}(\mathbf{v}_{orig,n}) \|$ unterschritten wird. Um die Vergleichbarkeit mit dem visuellen Lernprozess zu gewährleisten, wurde als Schwellwert 1% der Gesamtlänge der zu lernenden kinematischen Kette gewählt.

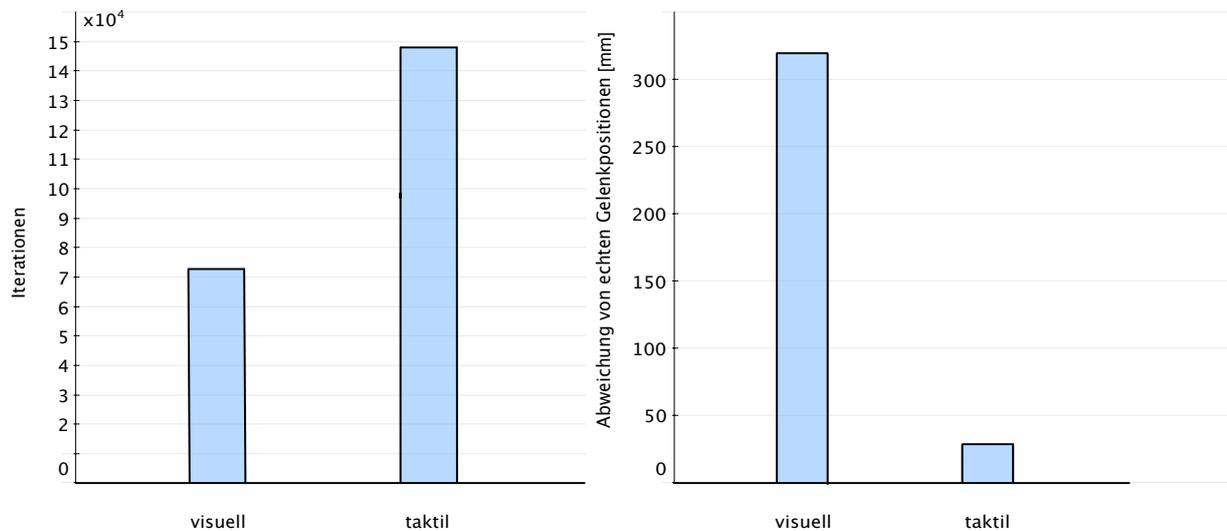


Abbildung 5.7: Linke Seite: Durchschnittliche Iterationszahl für visuellen und taktilen Lernprozess wie in Algorithmus 1 beschrieben, ohne Nach-Adaptierung der Startwerte. Rechte Seite: Durchschnittliche Abweichung des *subjektiven* vom *objektiven* Körperschema.

5.3.1 Visueller und taktiler Lernprozess

In einem ersten Evaluationstest wurde der visuelle mit dem taktilen Lernprozess verglichen. Letzterer wurde in Algorithmus 1 dargestellt. Die Nach-Adaptierung der Startwerte (Zeilen 22–27) für die Gelenkpositionen wurde hierbei noch nicht berücksichtigt, sondern einzeln evaluiert (siehe Abschnitt 5.3.3). Die beiden Lernprozesse unterscheiden sich in der Anzahl der Referenzpositionen sowie in den Längen der kinematischen Teilketten, die adaptiert werden. Bei dem taktilen Lernprozess werden zufällige Positionen am Körper erfasst, die zu zufälligen Längen der zu adaptierenden kinematischen Teilkette führen. In Abbildung 5.7 sind die Ergebnisse für die Iterationszahl und die Akkuratheit des berechneten künstlichen Körperschemas dargestellt. Die durchschnittliche Anzahl der benötigten Iterationen ist fast doppelt so hoch wie beim visuellen Lernprozess. Das taktile Lernverfahren ist somit deutlich langsamer. Für die durchschnittliche Abweichung der berechneten von den tatsächlichen Gelenkpositionen ergibt sich für den visuellen Lernprozess ein höherer Wert als für den taktilen Lernprozess. Das taktil erlernte künstliche Körperschema kommt dem *objektiven Körperschema* somit näher.

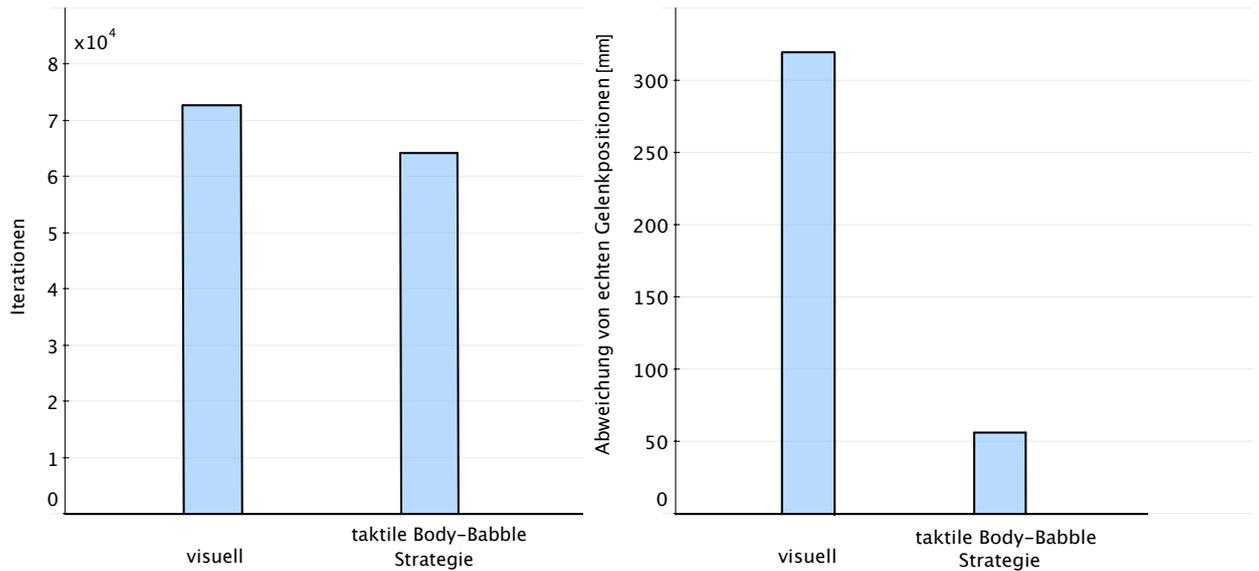


Abbildung 5.8: Linke Seite: Durchschnittliche Iterationszahl für visuellen und taktile Lernprozess mit taktile Body-Babbling-Strategie. Rechte Seite: Durchschnittliche Abweichung des *subjektiven* vom *objektiven* Körperschema für visuelle und taktile Berechnungsart.

5.3.2 Taktile Body-Babbling-Strategie

Ein Vorteil des taktile Lernprozesses ist die variierende Länge der kinematischen Teilketten. Je kürzer die zu adaptierende Kette ist, desto geringer wird die Anzahl der zu adaptierenden Parameter und die Berechnungskomplexität. Da im normalen Lernprozess taktile Selbstkontakte zufällig entstehen, kann der Vorteil kürzerer Teilketten nicht optimal ausgenutzt werden. In einem zweiten Experiment evaluierten wir daher den taktile Lernprozess mit einer *taktile Body-Babbling-Strategie*. In dieser werden zuerst diejenigen Selbstkontakte zur Adaptierung verwendet, die die kürzesten kinematischen Ketten erzeugen. Es werden dafür nacheinander kinematische Ketten mit der Länge 2, 3, 5 und 7 adaptiert. Diese entsprechen einem Selbstkontakt zwischen Max' Hand und seiner Schulter (Gelenk $r_shoulder$ bzw. $l_shoulder$), dem Hals (Gelenk vc_7), dem Bauch (Gelenk vt_6) und der Hüfte (Gelenk vl_5). Abbildung 5.8 zeigt die Ergebnisse. Gegenüber dem visuellen Lernverfahren ist die Iterationszahl um 12,5 % gesunken, und die Akkuratheit des berechneten Körperschemas ist, wie bei dem normalen taktile Lernprozess, höher als beim visuellen Lernprozess.

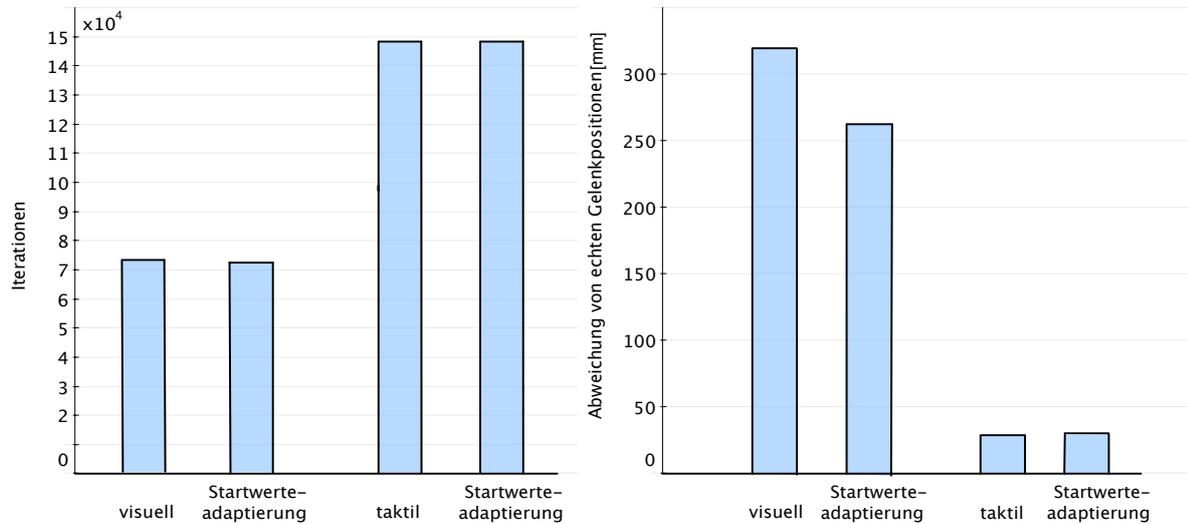


Abbildung 5.9: Linke Seite: Durchschnittliche Iterationszahl für visuellen und taktilen Lernprozess mit Adaptierung der Startwerte. Rechte Seite: Durchschnittliche Abweichung des *subjektiven* vom *objektiven* Körperschema für visuelle und taktile Berechnungsart mit Startwert-Adaptierung.

5.3.3 Startwerte über Tastrezeptorposition

In weiteren Simulationen wurde die Adaptierung über die Tastrezeptorpositionen evaluiert. Hierfür wurden sowohl der visuelle als auch der taktile Lernprozess als Basismessung gewählt. Verglichen wurden die Ergebnisse aus Abschnitt 5.3.1 jeweils mit der einmaligen Nach-Adaptierung der Startwerte für die Gelenkpositionen. Die Ergebnisse sind in Abbildung 5.9 zu sehen. Diese Adaptierungsmethode erzielt weder im visuellen noch im taktilen Lernprozess eine Änderung der Iterationszahlen. Die Abweichungen zwischen subjektivem und objektivem Körperschema blieben für den taktilen Lernprozess gleich. Für den visuellen Lernprozess konnte eine durchschnittliche Verringerung der Abweichung um 18 % erreicht werden. Wird die Nach-Adaptierung für alle Iterationen eingesetzt, so steigt die Anzahl an Berechnungsschritten erheblich.

5.4 Diskussion

Im Folgenden werden zuerst die vorgestellten Simulationsergebnisse diskutiert und Schlussfolgerungen aus diesen gezogen. Da das Lernverfahren dem Konzept des biologischen Körperschemas nachempfunden ist, ist es jedoch nötig, auch die biologische Adäquatheit der technischen Eigenschaften zu betrachten.

Vergleich des taktilen und visuellen Lernprozesses

Obwohl die Nach-Adaptierung der **Startwerte für die Tastrezeptorposition** vielversprechend erschien, erbrachte sie keinen nennenswerten Mehrertrag in den Iterationszahlen. Für den taktilen Lernprozess konnten auch in der Genauigkeit keine Unterschiede festgestellt werden. Nur die Genauigkeit unter dem visuellen Lernprozess konnte verbessert werden. Die Verbesserung kann jedoch die weiterhin bestehende Abweichung im Vergleich zu den anderen Berechnungsarten nicht mindern. Dies lässt folgende Schlüsse zu: Erstens scheint die Gelenkpositionenadaptierung über Formel 3.4 (s. Seite 60) einen geringen Einfluss auf die Gesamtanzahl der Iterationen zu haben; tatsächlich wird der Großteil der Iterationszahlen über die Achsenadaptierung mit der Formel 3.5 erzeugt. Zweitens scheinen die initialen Werte der Gelenkpositionen keinen deutlichen Einfluss auf ihre Adaptierung zu haben. Dieses Ergebnis lässt Schlussfolgerungen zum Design von Tastrezeptoren für humanoide Agenten zu. Wenn die genaue Position der Tastrezeptoren entlang eines Segments keinen nennenswerten Einfluss auf ein Rekalibrierungsverfahren wie dieses hat, dann ist möglicherweise keine große Anzahl an Rezeptoren nötig, die eng und flächendeckend über die Segmente verteilt sind.

Im **Vergleich zum visuellen Lernprozess** konnte im taktilen Lernprozess, durch mehr Referenzpositionen über die Tastrezeptoren ein akkurateres Ergebnis für die Gelenkpositionen erzielt werden. In der Vergleichsevaluation in Abschnitt 5.3.1 konnte Hypothese 1 somit bestätigt werden. Obwohl innerhalb des taktilen Lernprozesses eine Vielzahl von kurzen kinematischen Ketten erzeugt wird, benötigt dieser dafür fast doppelt so viele Berechnungsschritte wie der visuelle Lernprozess. Hypothese 2 konnte also mit dieser Berechnungsart nicht bestätigt werden. Eine Erklärung dafür ist, dass die kurzen kinematischen Teilketten in zufälliger Reihenfolge erzeugt werden. Diese können erst einen Beitrag zu längeren Teilketten erbringen, wenn sie vollständig konvergiert sind. Diese Bedingung ist nicht gegeben, da Selbstkontakte mit kurzen Teilketten in zufälliger Reihenfolge stattfinden. Zudem können immer nur zwei Tastrezeptoren nacheinander als Sensorinformation zur Adaptierung verwendet werden.

Diese Eigenschaft wurde in den Evaluationstests in Abschnitt 5.3.2 über eine **taktile Body-Babbling-Strategie** geprüft. Diese führte gegenüber dem visuellen Lernprozess tatsächlich zu geringeren Iterationszahlen. Die Genauigkeit der Ergebnisse wurde ebenfalls verbessert. Mit dieser Berechnungsmethode konnten Hypothese 1 und 2

bestätigt werden. Bei diesen Ergebnissen gibt es jedoch bei genauerer Prüfung Kritikpunkte. Für den Lernalgorithmus muss Vorwissen bereitgestellt werden. Die Body-Babbling-Strategie geht davon aus, dass der humanoide Agent bereits weiß, welche Körperteile zuerst für den Selbstkontakt exploriert werden müssen, um die kürzeren kinematischen Ketten zu adaptieren. Diese Art der Strategie entspricht nicht dem biologischen Vorbild des zufälligen Body-Babblings. Im Gegensatz zum taktilen Lernprozess ist es bei der visuellen Strategie möglich, mehrere Referenzpositionen gleichzeitig zu beobachten. Es ist aber auch hier Vorwissen darüber nötig, welche Bildbereiche des Kamerabildes zum Körper gehören und um welches Körperteil es sich handelt. Auch hier stellt sich die Frage, ob das gezielte visuelle Verfolgen mehrerer Körperteile zur gleichen Zeit eine natürliche Body-Babbling-Strategie darstellt.

Es stellt sich die generelle Frage, welches Vorwissen und welche Annahmen für technische Systeme, die von biologischen Prozessen motiviert sind, vorausgesetzt werden können und sinnvoll sind. Für jede technische Umsetzung muss der Aufwand, Vorwissen bereitzustellen, dem Mehrgewinn durch geringere Berechnungszeit gegenübergestellt werden. Im hier vorgestellten Fall hat sich die Iterationsanzahl mit taktiler Body-Babbling-Strategie um über die Hälfte reduziert, die Akkuratheit ist jedoch gleich hoch geblieben. Mit dem heutigen Stand der Computerhardware sind beide Iterationszahlen für das gegebene Problem innerhalb kürzester Zeit zu berechnen. Der einfache taktile Lernprozess ist zwar langsamer in der Berechnung, liefert aber akkurate Ergebnisse mit einer kleinen Menge an dafür zu bereitzustellendem Vorwissen.

Biologische Adäquatheit des Lernalgorithmus

Für die hohe Effektivität des Lernverfahrens können zwei wesentliche Faktoren bestimmt werden, die einen entscheidenden Unterschied zu bereits existierenden Körperschemalernverfahren darstellen. Der erste Faktor ist das *Vorwissen*, genauer die Anzahl der Freiheitsgrade und Gelenke sowie die Struktur des kinematischen Skeletts, die beide bereits vor dem Lernen bekannt sind und sowohl in den visuellen als auch in den taktilen Lernprozess einfließen. Für viele der in Abschnitt 3.2.3 vorgestellten Lernverfahren ist ein Motiv für das Lernen eines künstlichen Körperschemas, es einem Roboter zu ermöglichen, sich automatisch an an- bzw. abmontierte Körperteile anzupassen und diese zu repräsentieren. Dieses kann mit der vorgeschlagenen Methode nicht vollständig automatisiert gelöst werden. Die kinematische Struktur müsste manuell vor

Eingabe in den Algorithmus um die Gelenkzahl und die Stelle in der Skeletthierarchie angepasst werden. Solch eine Vormodellierung wäre aus Herschs Sicht jedoch vermutlich biologisch angemessen, da er der Auffassung ist, dass in der menschlichen kinematischen Struktur Gliedmaßen wachsen, während der Entwicklung im Normalfall aber keine neuen entstehen. Der zweite Faktor ist die Vorannahme, dass es sich bei allen Gelenken um *Scharniergelenke* handelt, deren Rotationen im Modell durch die Achsen-Winkel-Repräsentation dargestellt werden (siehe Abschnitt 3.1.1). Diese Annahme bildet jedoch nicht alle biologisch relevanten Gelenke ab. In der menschlichen Schulter befinden sich zum Beispiel Kugelgelenke. Hersch vertritt hier die Ansicht, dass Gliedmaßenbewegungen immer starr sind und sich daher aus jedem propriozeptiven Sensorsignal etwas ableiten lässt, das einem Rotationswinkel um eine bestimmte Achse entspricht.

Biologische Sicht auf die Sensordatenerfassung

Eine besondere Stellung nimmt in Herschs Arbeit der **Online**-Aspekt der Lernphase ein. Diese steht hier im Gegensatz zu dem sogenannten Batch-Lernen. Letzteres ist in Computerlernverfahren sehr üblich und besteht aus zwei Phasen. In der ersten Phase, der sogenannten Trainingsphase, werden Daten gesammelt, und in der zweiten Phase, der sogenannten Lernphase, wird ein Modell an diese Daten angepasst. Beim Online-Lernen fällt die Trainingsphase, in diesem Fall die für das Körperschemalernen übliche Body-Babbling-Phase, mit der Lernphase zusammen. Während der humanoide Agent sich bewegt, werden die Sensordaten direkt in den Lernalgorithmus eingegeben. Das Online-Lernen ist somit gegenüber dem Batch-Lernen nicht nur technisch vorteilhafter, sondern biologisch auch adäquater, da das Lernen über die Sensordaten kontinuierlich vorangetrieben wird. Hersch sieht in dieser graduellen Anpassung eine Ähnlichkeit zu den Prozessen, wie sie auch in der Natur angenommen werden.

Für den taktilen Lernprozess gibt es die Möglichkeit, die biologisch möglicherweise inadäquate taktile Body-Babbling-Strategie *natürlicher* zu modellieren. Führt man in den taktilen Lernprozess eine Trainingsphase ein, so könnte eine große Anzahl von Selbstkontakten zufällig erzeugt werden, was dem biologischen Body-Babbling näher käme. In einer darauffolgenden Lernphase könnten die Daten nach Länge der kinematischen Ketten sortiert und erst die kürzeren adaptiert werden. Das Ausnutzen kürzerer kinematischer Ketten würde somit vom Datenerfassungsprozess in den ei-

gentlichen Lernprozess verschoben werden. Hierfür müsste jedoch der Online-Aspekt des gesamten Lernverfahrens aufgegeben werden. Zudem würde die Zeitersparnis bei der Adaptierung vermutlich durch die zwei getrennten Lernphasen kompensiert werden.

5.5 Zusammenfassung und Resümee

In diesem Kapitel wurde eine Lösung des Teilproblems P2 (Repräsentation des Körpers) dargestellt. Es wurden die Grenzen der visuell-propriozeptiven Rekalibrierung diskutiert und mit den Anforderungen an ein künstliches Körperschema verglichen. Als Konsequenz wurde ein taktile Lernprozess vorgeschlagen, der keine visuellen Sensordaten benötigt. In einer Evaluation wurden aus der Analyse der Problemstellung zwei Hypothesen gebildet, die zur Verringerung der Iterationszahlen und der Verbesserung der Akkuratheit führen können. Die Simulationen des visuellen Lernprozesses wurden mit drei verschiedenen Berechnungsarten des taktilen Lernprozesses verglichen. Aus den Evaluationen wurden drei wesentliche Erkenntnisse gewonnen:

- Der taktile Lernprozess steigert die durchschnittliche Iterationsanzahl, vergrößert aber die Akkuratheit des Lernverfahrens.
- Eine taktile Body-Babbling-Strategie verringert die durchschnittliche Iterationsanzahl und vergrößert gleichzeitig die Akkuratheit des Lernverfahrens. Der Nachteil ist die geringe biologische Adäquatheit.
- Die zunächst vielversprechend wirkende Startwerte-Adaptierung bringt keine nennenswerte Verbesserung der Lernverfahren mit sich. Die Tatsache führt jedoch zu der Erkenntnis, dass eine hohe Abdeckung des humanoiden Körpers mit Tastrezeptoren keine große Rolle für das hier vorgestellte Körperschemalernverfahren spielt.

Mit diesen Ergebnissen konnten wir bessere Erkenntnisse zur Rolle und Funktion taktiler Sensorinformationen im Aufbau der Körperrepräsentation erlangen. Die wichtigste Erkenntnis, die wir mit dem Einsatz des taktilen Lernprozesses gewinnen können ist, dass die taktile Modalität gegenüber der visuellen Modalität ein korrekteres Abbild des tatsächlichen Körpers liefern konnte. Mit dem Einsatz der taktilen Sensormodalität

wurde eine zentrale Eigenschaft des biologischen Körperschemas technisch umgesetzt. Die Umsetzung hebt die Bedeutung der taktilen Modalität nicht nur für Menschen, sondern auch für humanoide Agenten hervor. Die intuitive Vermutung, dass die taktile Modalität bessere Informationen über den Körper liefert als die visuelle, kann mit diesen Ergebnissen bestätigt werden.

In Abschnitt 4.3 wurde die Frage diskutiert, ob es möglich ist, ein für physische Roboter entwickeltes Körperschemalernverfahren auf virtuelle Menschen zu übertragen. In diesem Kapitel konnte dies erfolgreich gezeigt werden. Es stellt sich nun umgekehrt die Frage, ob der taktile Lernprozess auf physische Roboter übertragbar ist. Ein grundsätzliches Problem stellt die Verfügbarkeit taktiler Sensoren dar. Die hier erzielten Ergebnisse profitierten von dem Aufwand, der in der Arbeit steckt, einen biologisch adäquaten Tastsinn für humanoide Agenten zu realisieren. Ist die Verfügbarkeit eines solchen Tastsinns nicht gegeben, muss generell die Frage gestellt werden, ob sich der Entwicklungsaufwand lohnt. Bei virtuellen Menschen ist dieser überschaubar, für physische Roboter stellt diese Entwicklung jedoch eine erhebliche Arbeit dar. Sind taktile Sensoren für einen physischen Roboter verfügbar, stellt sich das Problem des physischen Selbstkontakts. Diese sind nur durch präzise Motorkontrolle möglich, um Beschädigungen des Roboterkörpers zu vermeiden. Der Body-Babbling-Prozess muss sehr langsam durchgeführt werden, was die Gesamtlernzeit erheblich steigert. Diese wird noch zusätzlich dadurch gehoben, da eine genügend große Anzahl zufällig entstehender Selbstkontakte generiert werden müssen. Da virtuelle Menschen grafische Simulationen darstellen, ist die Simulation der Selbstkontakte wesentlich einfacher und sie verläuft ohne nennenswerte zeitliche Begrenzung.

6 Ein Modell des peripersonalen Raums für virtuelle Agenten

Im vorangegangenen Kapitel wurde das Lernverfahren dargestellt, das es ermöglicht, ein künstliches Körperschema für Max aufzubauen. Dieses bildet die Grundlage für zwei wichtige Eigenschaften des biologischen peripersonalen Raums. Die erste ist die Repräsentation der räumlichen Umgebung in körperteilzentrierten Referenzsystemen. Die zweite ist das Wissen über die räumlichen Grenzen des peripersonalen Raums. Aufbauend auf diesen Eigenschaften wird in diesem Kapitel ein Modell eines peripersonalen Raums für Max vorgestellt. Dieses stellt eine Lösung für das **Teilproblem der Wahrnehmung und Repräsentation des Nahraums (P1)** dar.

In diesem Kapitel wird zunächst in Abschnitt 6.1 analysiert, wie ein für ein technisches System sinnvolles peripersonales Raummodell entwickelt werden kann. Darauf aufbauend wird in Abschnitt 6.2 das peripersonale Raummodell vorgestellt, das es Max ermöglichen soll, seine nahe Umgebung sinnvoll wahrzunehmen und zu repräsentieren. Es baut auf den Raummodellen für Menschen auf, die in Abschnitt 2.1.3 beschrieben wurden. In Abschnitt 6.3 wird die technische Umsetzung des peripersonalen Raummodells und die darin realisierten Arm- und Torsobewegungen vorgestellt. Im Anschluss daran, wird in Abschnitt 6.4 eine Zusammenfassung des Kapitels gegeben und ein Resümee gezogen.

6.1 Voranalyse

Die Idee der Übertragung von Raummodellen auf technische Systeme wird in den Kognitionswissenschaften bereits gedanklich verfolgt. Holmes und Spence (2004, S. 94) schlagen zum Beispiel die folgende Umsetzung vor:

If an engineer were to design a robot that could move through the world, selecting and grasping objects as it went, the best use of its limited compu-

tational resources might well be to plan only grasping movements to those objects of interest within direct reach, and to plan only locomotive movements to those objects situated at a distance. Similarly, the avoidance of objects that are of potential harm must be a primary goal for all organisms fighting for their survival. Those objects that are of most immediate threat are those that are closest to, and moving most rapidly toward, the body [...].

Holmes und Spence verdeutlichen, wie zur Ressourceneinsparung in künstlichen Systemen räumliche Distanzen mit unterschiedlichen Funktionen verknüpft werden können. Je näher sich ein Objekt am Körper befindet, desto mehr Aufmerksamkeit und mögliche Handlungsoptionen müssen für das künstliche System geplant werden. Ihr Gedanke kann dem Prinzip des sogenannten Angebotscharakters, besser bekannt als **Affordance** zugeordnet werden (Gibson, 1977). Dieses Konzept wurde in jüngster Zeit insbesondere auf künstliche Systeme übertragen und beschreibt eine Eigenschaft von Gegenständen oder der Umgebung, bestimmte Aktionsmöglichkeiten vorherzubestimmen. Zum Beispiel könnte ein Knopf zum Drücken oder eine Treppe zum Aufsteigen auffordern.

Übertragen auf den Gedanken von Holmes und Spence stellt sich die Frage, ob innerhalb des peripersonalen Raums ein **räumlicher Angebotscharakter** bestimmt werden kann. Hier wäre zu untersuchen, welche räumlichen Distanzen mit welchen Aktionen verknüpft sein könnten. So beschreiben Mark u. a. (1997) (vgl. Abschnitt 2.1.4) den Übergang zwischen einfachen Armgreifbewegungen zu Armgreifbewegungen mit Vorlehnen als Vorgang, bei dem die *Affordance* für den aktuellen Aktionsmodus nicht mehr gegeben ist, sich also der Aktionsmodus nicht mehr anbietet und der nächste Aktionsmodus eingeleitet wird. Obwohl die absolute Grenze der einfachen Armgreifbewegungen durch maximales Strecken noch nicht überschritten ist, wird dann eine Greifbewegung mit Vorlehnen gewählt, weil diese möglicherweise komfortabler oder mit weniger körperlichem Aufwand (engl. effort) und Energie verbunden ist.

Mit dem hier vorgestellten peripersonalen Raummodell werden diese Aspekte aufgegriffen; dies führt zu einem Modell mit unterschiedlichen *Unterräumen*. Diese bieten für technische Systeme zwei wesentliche Vorteile.

- Die Unterräume entsprechen den Befunden bei Menschen, die zeigen, dass der peripersonale Raum multi-sensorisch repräsentiert ist (vgl. Definition 1). In ei-

nem technischen System, in dem verschiedene Sensormodalitäten nicht zwingend dieselben Raumregionen abdecken, bieten solche Befunde die Möglichkeit einer umfassenderen und möglicherweise robusteren Repräsentation des peripersonalen Raums.

- Ein weiterer Vorteil besteht in der Zuordnung von **Aufmerksamkeitsräumen**. Bestimmte Unterräume können für vordefinierte Berechnungsprozesse besonders wichtig sein. So ist z. B. in einem Raumbereich, in dem eine Objektmanipulation mit den Armen stattfindet, die Wahrscheinlichkeit größer, dass Kollisionen zu vermeiden sind, wenn der Raumbereich mit einem Partner geteilt wird.

6.2 Strukturierung des peripersonalen Raums

Damit ein humanoider Agent seinen peripersonalen Raum geeignet repräsentieren kann, um in diesem zu interagieren, ist es sinnvoll, zunächst die künstliche Wahrnehmung vorzustrukturieren. Dies kann durch ein Raummodell erreicht werden, das als Umgebungsfilter betrachtet werden kann. Solch ein Filter hilft einem technischen System, aus der Vielzahl an Umgebungsstimuli die relevanten zu ermitteln. Wir beginnen daher mit der Entwicklung eines Raummodells. Das vorgeschlagene Modell sollte technisch realisier- und nutzbare Eigenschaften des biologischen peripersonalen Raums abbilden, gleichzeitig aber die Gegebenheiten des sensorischen Systems eines humanoiden Agenten, in diesem Fall Max, berücksichtigen. In den Nahraumszenarien nehmen wir an, dass peripersonale Raum-Interaktion mit Objekten üblicherweise eine horizontale Fläche beinhaltet, die sich vor Max' Körper befindet. Dies kann zum Beispiel ein einfacher Tisch sein. Um die Komplexität des Modells einzugrenzen, konzentrieren wir uns auf den peripersonalen Raum, der auf einer zweidimensionalen Transversal-Ebene vor Max' Oberkörper projiziert ist.

In Abschnitt 2.1 wurden bereits verschiedene Raummodelle diskutiert. Das extrapersonale Raummodell von Previc (1998) erwies sich als das umfassendste Modell, das zudem Aspekte von Grüssers Raummodell (Grüsser, 1983) beinhaltet. Da Previc hauptsächlich den extrapersonalen Raum untersuchte, fand der peripersonale Raum zwar Erwähnung, wurde jedoch nicht explizit ausgearbeitet. Cardinali u. a. (2009) schlagen ein peripersonales Raummodell vor, das zwar auf aktuelleren Studien beruht, jedoch weit weniger umfassend und informeller ist als Prevics Modell. In dieser Arbeit

bilden daher diese Modelle die Grundlage für den peripersonalen Raum. Unser Modell stützt sich auf Ergebnisse über den biologischen peripersonalen Raum, wie er von Cardinali u.a. diskutiert wird. Die gewählte Modellierungsweise stützt sich auf Prevics Arbeit.

Unterräume im peripersonalen Raum

Nach Prevics Raummodell ist jeder Raumbereich, der einen Menschen umgibt, mit einer vorherrschenden Verhaltensweise verknüpft. Zum Beispiel ist visuo-motorische Objektmanipulation im peripersonalen Raum vorherrschend, wohingegen Fortbewegung in Prevics Modell an den aktions-extrapersonalen Raum geknüpft ist. Er definiert für sein extrapersonales Raummodell jeweils für jeden Raumbereich sensorisch-perzeptuelle Operationen, Motoroperationen sowie ein vorherrschendes Referenzsystem.

In dieser Arbeit wurde dieser Modellierungsansatz auf den peripersonalen Raum übertragen. Dieser wurde in drei Hauptbereiche unterteilt, denen jeweils eine Sensormodalität zugeordnet werden kann. Es handelt sich um visuelle, taktile und propriozeptive Modalitäten. Jeder Raumbereich ist durch seine spezifische Raumgrenze, ein vorherrschendes Referenzsystem und vorherrschende Motoraktionen gekennzeichnet. Die Sensormodalitäten, von denen angenommen wird, dass sie zum biologischen peripersonalen Raum beitragen, beeinflussen in unserem Modell die drei Unterräume. In Grüssers Raummodell ist der *Greifraum* mit visueller und taktiler Objektmanipulation und -erkennung verbunden (vgl. Abschnitt 2.1.3). Im hier vorgestellten Modell werden diese Aktionen getrennt und zwei unterschiedlichen Räumen zugeordnet. Die visuelle Modalität kommt im peripersonalen Raum hauptsächlich bei der Objektsuche und bei visuellen Manipulationsaktionen mit den Händen zum Einsatz. Daher legt diese Modalität einen sogenannten **visuellen Aufmerksamkeitsraum** (engl. visual attention space) fest. Die taktile Modalität kommt hauptsächlich in der Objektmanipulation und bei Greifbewegungen mit den Händen zum Einsatz. Mit dieser Modalität ist daher ein spezieller **Taktilraum** (engl. touch space) verknüpft. Die Propriozeption wird im gesamten peripersonalen Raum eingesetzt, spielt jedoch eine besondere Rolle beim Platzieren und Greifen von Objekten, die sich am Rande des peripersonalen Raums befinden. Hier wird ein höherer Gelenkaufwand benötigt, um sich vorzulehnen. Dieser Aspekt wird durch den sogenannten **Vorlehnraum** (engl. lean-forward space)

Tabelle 6.1: Charakteristiken der peripersonalen Unterräume

	visueller Aufmerksamkeitsraum	Taktilraum	Vorlehnraum
Funktion	visuelle Suche & Kontrolle	Greifen, Platzieren, Manipulation	Greifen, Platzieren
2D Ort, Ausmaße		Projektion auf frontale 2D-Ebene	
vertikal			
Ursprung	Kopf	Schulter, Oberkörper	Schulter, Oberkörper
lateral	zentrale 60°	360°	frontale 180°
radial	0-2m	Länge: Schultergelenke bis Handflächen	Länge: Hüftgelenk bis Handflächen
Referenzsystem	kopfzentriert	gliedmaßenzentriert	gliedmaßenzentriert
Motoraktion	Kopf- und Augenbewegungen	Armbewegungen	Oberkörperbewegungen

abgebildet. In Tabelle 6.1 sind die Charakteristiken der drei Unterräume aufgeführt.

Jeder definierte Unterraum ist mit einer vorherrschenden Motoraktion verknüpft, die hauptsächlich in dem spezifischen Unterraum ausgeführt wird. Wie zu Beginn des Kapitels erläutert, werden die Unterräume auf eine angenommene zweidimensionale Ebene, möglicherweise ein Tisch vor Max' Oberkörper, projiziert. Daher werden die vertikalen räumlichen Ausdehnungen jedes Unterraums auf einen 180° Halbkreis, der radial auf der 2-D-Ebene aufgespannt ist, projiziert (s. Abbildung 6.2).

Der Ursprung des **visuellen Aufmerksamkeitsraums** liegt in der Mitte des Kopfes. Seine laterale Ausdehnung wird auf den Taktilraum und auf den Vorlehnraum projiziert. Stimuli, die innerhalb eines 60° breiten Sichtwinkels wahrgenommen werden, sind in einem kopfzentrierten Referenzsystem repräsentiert. Die Ausdehnung des Unterraums stützt sich auf Previc's PrP-Raum, in dem angenommen wird, dass das maximale stereoskopische Sichtfeld bei Menschen ungefähr 60° beträgt.

Die Grenzen des **Taktilraums** sind durch die maximale Länge der Armgliedmaßen festgelegt. Die Längen wurden über das künstliche Körperschema ermittelt. Laut Mark u. a. (1997) definiert diese Länge bei Menschen die *absolute kritische Grenze* zwischen einfachen Armbewegungen und Armbewegungen mit zusätzlichem Vorlehnen (vgl. Abschnitt 2.1.4). Der Taktilraum hat seinen Ursprung in der Mitte zwischen den Schultergelenken und wird durch die Distanz zwischen Schultergelenken und Handflächen begrenzt. Vom Ursprung ausgehend deckt diese Distanz eine radiale Fläche ab, die lateral um den Torso herumgeht. Die laterale Ausdehnung beträgt somit 360°, da taktile Stimuli auch am Rücken wahrgenommen werden können. In den folgenden

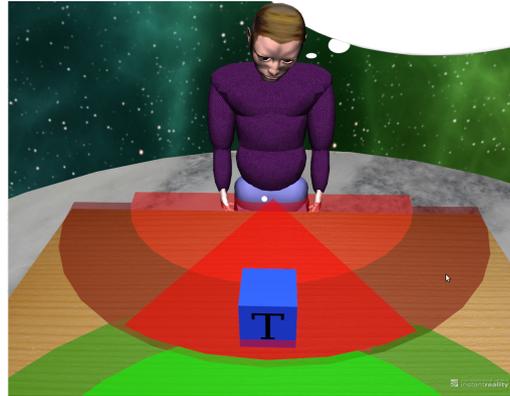


Abbildung 6.1: Visualisierung in der virtuellen Realität der Unterräume des peripersonalen Raums von Max mit Taktilraum (kleine rote Halbscheibe), Vorlehnraum (größere rote Halbscheibe) und dem visuellen Aufmerksamkeitsraum (roter Winkelabschnitt).

Szenarien wird dieser Fall jedoch nicht betrachtet. Zur Vereinfachung werden daher nur die frontalen 180° des Taktilraums berücksichtigt.

Der **Vorlehnraum** wird durch die maximale Reichweite des Oberkörpers begrenzt, wenn dieser sich vorlehnt. Über das Körperschema wird die maximale Distanz ermittelt, die von den Armgliedmaßen zusammen mit den Wirbelsäulengelenken erreicht werden kann. Dieser Raum erweitert somit den Taktilraum. Objekte bzw. Stimuli, die in beiden Unterräumen wahrgenommen werden, sind in einem gliedmaßenzentrierten Referenzsystem repräsentiert. Im Gegensatz zum Taktilraum ist die Funktion der Objektmanipulation im Vorlehnraum nicht vorherrschend. Eine natürliche Aktion wäre es, zunächst das Objekt aus dem Vorlehnraum näher in den Taktilraum zu bringen und erst dort Manipulationsaktionen durchzuführen.

6.3 Technische Umsetzung für virtuelle Menschen

In diesem Abschnitt wird die technische Umsetzung des peripersonalen Raummodells aus Abschnitt 6.2 beschrieben. Die einzelnen Komponenten werden hier zur besseren Übersicht eingeführt und mit den Befunden aus den bisher diskutierten Forschungsarbeiten motiviert. Für die technische Realisierung wurde der peripersonale Raum in die drei bereits eingeführten Unterräume aufgeteilt. Da die Aufteilung des peripersonalen Raums anhand der in den Unterräumen vorherrschenden Sensormodalität bestimmt wird, werden diese Unterräume im Folgenden auch als *sensorische Unter-*

räume bezeichnet, um sie von weiteren möglichen Unterraumdefinitionen abzugrenzen. In Abbildung 6.1 ist eine Visualisierung der Unterräume in der virtuellen Umgebung zu sehen.

6.3.1 Peripersonale Unterräume für den virtuellen Agenten

Max

Für den *Taktilraum* des virtuellen Menschen Max wurde als Ursprungsreferenzsystem der Mittelpunkt zwischen den beiden Schultergelenken gewählt (Gelenk *vc7*, siehe Abbildung 4.3). Steht Max an einem Tisch, so wird der Raum als Halbscheibe auf den Tisch projiziert. Die Halbscheibe bewegt sich mit den Schultergelenken (Gelenke *l_shoulder* und *r_shoulder*) auf der projizierten Tischfläche mit. Lehnt sich Max mit seinem Oberkörper vor, wird die Halbscheibe mitgeführt. In Abbildung 6.5 ist eine Visualisierung der mitgeführten Halbscheibe in der virtuellen Umgebung zu sehen.

Der *Vorlehnraum* wird im Gegensatz zum Taktilraum nur dann mit Max' Körper mitgeführt, wenn Max sich im Raum fortbewegt. Der Vorlehnraum wird über das Wurzelgelenk von Max' kinematischem Körpermodell (Gelenk *vl5*) bewegt. Das Gelenk ist gleichzeitig der Ursprung des Referenzsystems für den Vorlehnraum. Der projizierte Taktilraum kann nicht über den Vorlehnraum hinausbewegt werden, da dieser den größten Unterraum des peripersonalen Raums darstellt. In Nahrauminteraktion bleibt der Vorlehnraum somit fest, und der Taktilraum bewegt sich innerhalb dessen Grenzen.

Der *visuelle Aufmerksamkeitsraum* hat als Ursprung seines Referenzsystems einen Punkt, der zwischen den Augen liegt (Gelenk *skullbase*). Der visuelle Aufmerksamkeitsraum bewegt sich mit diesem Gelenk. Die Mittelachse des visuellen Aufmerksamkeitsraums richtet sich frontal von Max' Kopf weg. Statt eines Kegels, wie er bei Sichtwinkeln üblich ist, wird zur Vereinfachung ein sagittaler Querschnitt des Raums auf den Tisch projiziert. Das in diesem Kapitel vorgeschlagene peripersonale Raummodell wird sowohl für humanoide physische als auch virtuelle Agenten entwickelt. Die visuelle Objekterkennung stellt jedoch ein Forschungsproblem aus dem Gebiet des maschinellen Sehens über Kamerasysteme dar. Diese sind für physische Robotersysteme von Nutzen. In der virtuellen Realität stellt sich dieses Forschungsproblem aufgrund der bereits in Grafikmodellen repräsentierten Welt üblicherweise nicht. Es wird daher eine Beispielanwendung des visuellen Aufmerksamkeitsraums für virtuelle Menschen gezeigt, jedoch liegt der Fokus im Folgenden auf der Nutzung des Taktill- und des

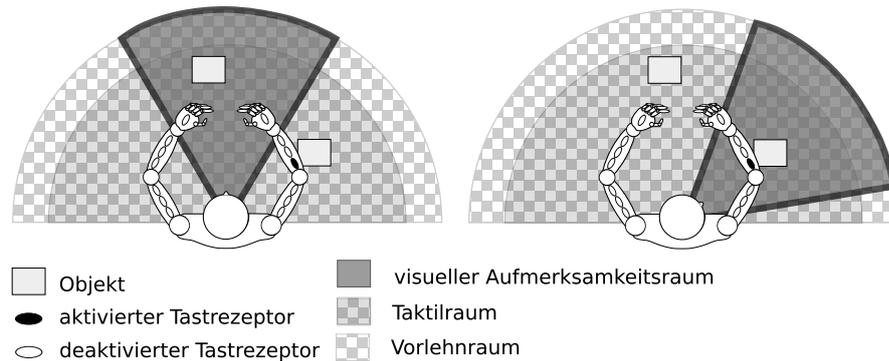


Abbildung 6.2: Schematische Übersicht über die sensorischen Unterräume. Der virtuelle Mensch richtet seine visuelle Aufmerksamkeit auf ein taktil wahrgenommenes Objekt. Linke Seite: Der virtuelle Mensch nimmt ein Objekt im Takttilraum, aber außerhalb seines visuellen Aufmerksamkeitsraums wahr. Rechte Seite: Prüft, ob das Objekt vorhanden ist. Eine Kopfbewegung zum Tastrezeptor wird durchgeführt, um das Objekt visuell zu erfassen und ggf. in der visuellen Objektkarte zu registrieren.

Vorlehnraums.

Sensorhierarchie

Jedes im peripersonalen Raum wahrgenommene Objekt wird in *Objekt-Raum-Karten* verwaltet, die den Sensormodalitäten entsprechen, mit denen die Objekte wahrgenommen wurden. Diese Karten dienen als Gedächtnis, in das Objektinformationen wie die Position im jeweiligen Referenzkoordinatensystem des wahrgenommenen Stimulus abgelegt werden. Auf Basis dieses Wissens kann der Agent seine nächsten Bewegungen planen, wie zum Beispiel einfaches Greifen eines Gegenstandes oder zusätzliches Vorlehnen des Oberkörpers. Visuell erfasste Objekte liegen in kopfzentrierten, taktil wahrgenommene in körperteilzentrierten Koordinaten vor.

Da die Objekte über verschiedene Sensormodalitäten wahrgenommen werden können, kann der Fall auftreten, dass sie sich in verschiedenen Unterräumen befinden. Ein solcher Fall wird in Abbildung schematisch 6.2 gezeigt. In dem Beispiel berührt der virtuelle Mensch Max aus Versehen ein virtuelles Objekt mit dem Ellbogen. Er kann es jedoch nicht sehen, weil sein visueller Aufmerksamkeitsraum zu diesem Zeitpunkt nicht die Objekte hinter dem Arm abdeckt. Um die Objekte innerhalb des peripersonalen Raums im Überblick zu behalten, müssen die Sensormodalitäten die Objekte abdecken. Dies geschieht über eine vordefinierte Sensorhierarchie. Üblicherweise ist es wichtiger, die Objekte in der nächsten Umgebung zu sehen, als sie taktil

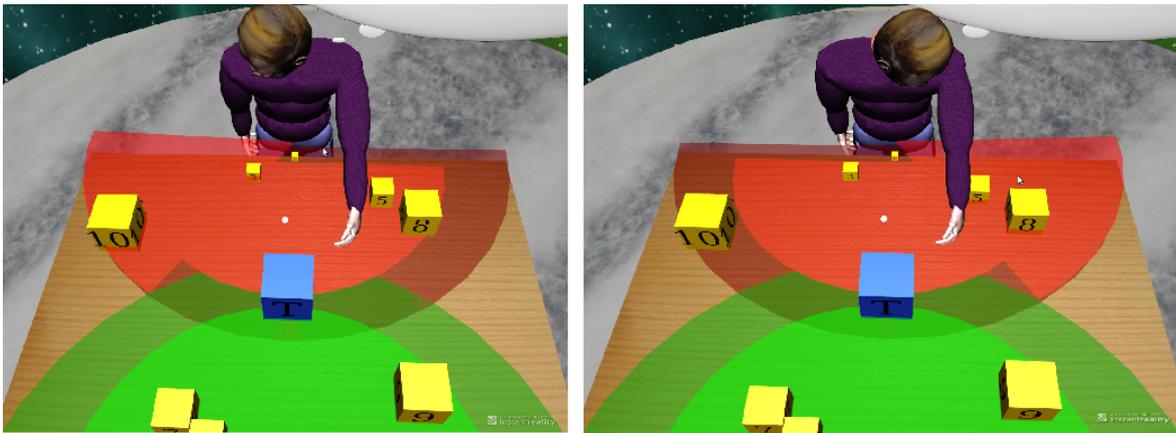


Abbildung 6.3: Linke Seite: Max' visueller Aufmerksamkeitsraum ist auf Objekt 10 gerichtet. Rechte Seite: Ein taktile Kontakt mit Objekt 5 und Max' Arm findet statt. Eine Kopfbewegung zum Tastrezeptor wird durchgeführt. Objekt 5 liegt nun im visuellen Aufmerksamkeitsraum.

wahrzunehmen oder zu greifen. Aus diesem Grund wird die visuelle Suche über die Objektmanipulation gestellt und diese wiederum über die Vorlehnbewegung. Das Objekt wäre somit nur in der Takttilraum-Karte gelistet, jedoch nicht in der visuellen oder der Vorlehnraum-Karte. Aufgrund der Sensorhierarchie wird eine Motoraktion aufgerufen, die zur Folge hat, dass Max das Objekt visuell wahrnehmen kann. Sein Kopf wird zur Objektposition gedreht, sodass der visuelle Aufmerksamkeitsraum das Objekt abdeckt (vgl. Abbildung 6.3). Ist das Objekt noch vorhanden, so wird es zusätzlich in die visuelle Karte eingeführt. Dieses Vorgehen bildet den neurowissenschaftlichen Befund ab, dass die Gegenwart von Objekten im peripersonalen Raum repräsentiert wird, auch wenn die Objekte nicht kontinuierlich sichtbar sind (vgl. Studien in Abschnitt 2.2.1). Das Prinzip wird als wichtig angenommen, um das Ergreifen von Objekten zu ermöglichen, auch wenn diese durch Hindernisse visuell nicht wahrnehmbar sind. Zudem entspricht die technische Realisierung der Eigenschaft des Greif-Raums in Grüssers Raummodell (siehe Abschnitt 2.1.3). Für diesen definiert er, dass die Objekte in den Sensormodalitäten *invariant* bleiben. Objekte existieren für uns weiter, auch wenn wir sensorisch unsere Aufmerksamkeit auf einen anderen Ort lenken. Nach Graziano u.a. wird diese Invarianz erst aufgehoben, wenn die Objekte bei erneutem Wahrnehmen tatsächlich nicht mehr präsent sind. Für physische Roboter wäre das Verschieben des visuellen Aufmerksamkeitsraums auf das taktile wahrgenommene Objekt der Zeitpunkt, an dem ein Prozess zur visuellen Objekterkennung stattfinden könnte. Wie im vorherigen Abschnitt beschrieben, ist dies in der virtuellen Realität nicht üblich.

6.3.2 Arm- und Torsobewegungsplanung

Im Takttilraum sind die vorherrschenden Motoraktionen Greifen, Platzieren und gegebenenfalls weitere Objektmanipulationen. Das bedeutet, dass vorwiegend Armbewegungen in diesem geplant werden. Im Vorlehnraum werden Objekte üblicherweise nicht manipuliert, sondern hinein- oder hinausplatziert. Im Folgenden wird dargestellt, wie Armbewegungen zum Umplatzen von Objekten und zusätzliches Vorlehnen mit dem Oberkörper umgesetzt wurden. Bei der Modellierung wurde insbesondere die Möglichkeit zur Parametrisierung der Bewegungsqualität berücksichtigt. Der Einsatz wird in Kapitel 7 näher beschrieben. Darüber hinaus wird gezeigt, wie der peripersonale Raum verwendet werden kann, sodass Max' Armbewegungen den peripersonalen Raum seines Partners berücksichtigen und vermeiden.

Bevor Max' Arm- und Torsobewegungen im peripersonalen Raum erläutert werden können, wird hier zunächst ein Überblick über Max' Gesten- und Motorplanung gegeben, da die für diese Dissertation realisierten Motoraktionen in die bestehende Softwarebibliothek *Articulated Communicator Engine* (ACE) integriert (vgl. Abschnitt 3.1.4) wurden. Äußerungen, die Max zeigen soll, werden in der Beschreibungssprache *MURML* (**M**ultimodal **U**tterance **R**epresentation **M**arkup **L**anguage) repräsentiert. Mit MURML ist es möglich, die äußere Form einer multimodalen Äußerung formal zu spezifizieren. In der ACE-Bibliothek dient die Spezifikation der automatischen Generierung von zeitlich koordinierter Sprache und Geste.

In Abbildung 6.4 ist das von Kopp vorgeschlagene Gesamtmodell zur Planung und Ausführung von gestischen Bewegungen dargestellt. Aus der MURML-Spezifikation wird im Gestenplaner über verschiedene Verarbeitungsschritte ein Gestenplan erzeugt. Dieser wird im Motorplaner in Unterbewegungen aufgeteilt und die sogenannten lokalen Motorprogramme (LMPs) aktiviert. Es existieren zum Beispiel für die Hände und für die Handgelenke eigene LMPs. Diese werden miteinander in Motorkontrollprogrammen (MKP) kombiniert, um Gesten zu erzeugen. So wird die Armsteuerung über ein MKP kontrolliert, indem es die aktivierten LMPs verwaltet. Die Bewegungstrajektorie, die ein Arm mit dem Handgelenk ausführen soll, wird über die inverse Kinematik nach der IKAN-Methode gelöst. Die Bewegungstrajektorien bestehen aus einer Sequenz von sogenannten *guiding strokes*, die gekrümmt oder linear sein können. Die Trajektorienformation wird durch ein LMP unter anderem daraufhin geprüft, ob ein *guiding stroke* Kollisionen mit dem Torso erzeugt, und ersetzt diese gegebenenfalls

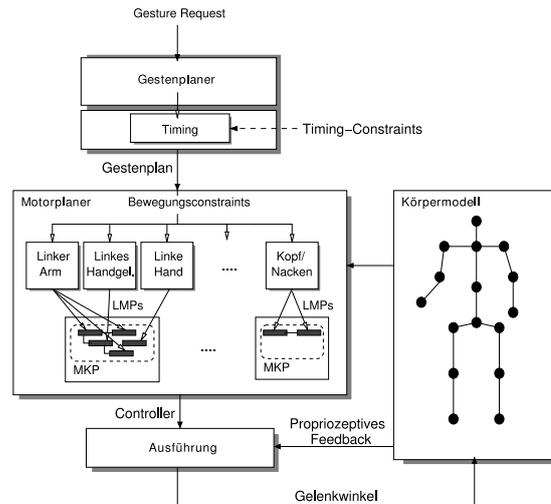


Abbildung 6.4: Ablauf der Gesten- und Motorplanung für Max. Abbildung angepasst, entnommen aus Kopp, Abb. 6.18, S. 176.

durch alternative guiding strokes.

Umplatzieren von Objekten

Die Umplatzier-Bewegung besteht aus den in Abschnitt 3.1.2 beschriebenen Teilen. Bei der Sequenzierungsplanung werden daher die folgenden Bewegungen hintereinander ausgeführt: eine *Reich-Bewegung*, bei der sich die Hand der Objektposition nähert, eine *Greif-Bewegung*, bei der die Finger eine Schließpostur an der Objektposition einnehmen, eine anschließende *Umplatzier-Bewegung*, bei der sich die Hand mit dem Objekt zu einer Zielposition bewegt, und eine *Loslöse-Bewegung*, bei der sich die Finger an der Zielposition wieder öffnen. Um eine Umplatzier-Bewegung durchführen zu lassen, werden folgende Werte benötigt:

- die eindeutige *Objekt-ID* des umzuplatzierenden Objekts,
- eine *Endposition*, zu der das Objekt hinbewegt werden soll, sowie
- ein *Geschwindigkeitsfaktor*.

Die **Objekt-ID** wird zunächst in den Objektkarten gesucht und daraufhin geprüft, ob sie der Sensorhierarchie folgend mindestens in der visuellen Karte registriert wurde. Ist dies nicht der Fall, so muss das Objekt in den übrigen Räumen visuell gesucht werden. In den Interaktionsszenarien, die im Folgenden behandelt werden, liegt der Fokus

nicht auf der visuellen Suche. Es wird stattdessen davon ausgegangen, dass bereits ein initialer Eintrag für die Objekte mit Objekt-ID und initialer Position besteht.

Es gibt mehrere Möglichkeiten, Endpositionen anzugeben. Eine Möglichkeit ist, globale Weltkoordinaten anzugeben. Diese werden dann in ein körperzentriertes Koordinatensystem umgerechnet, das als Ursprung das Wurzelgelenk *vl_5* hat. Dies ist das Koordinatensystem des Vorlehnraums. Eine weitere Möglichkeit besteht darin, keine spezifische Position für das Objekt anzugeben, sondern eine neue Position für das Objekt berechnen zu lassen. Eine andere Möglichkeit ist in speziellen Interaktionsszenarien, wie dem in Kapitel 7 beschriebenen, gegeben. Hier kann als Endposition eine zweite Objekt-ID angegeben werden. Max platziert das Objekt auf ein anderes, was dem *Stapeln* von Objekten entspricht.

Der **Geschwindigkeitsfaktor** ist ein Wert zwischen *0* und *1*, der die Bewegungsgeschwindigkeit der gesamten Umplatzierbewegung moduliert. Jede Umplatzierbewegung hat eine vordefinierte *minimale* und *maximale* Durchführzeit, innerhalb derer die gesamte Bewegung ausgeführt sein soll. Der Wert *0* steht dabei für die *langsamste* Durchführung, die Max für das Umplatzieren benötigen soll. Steigt der Wert, so wird die Durchführzeit linear immer kürzer, sodass Max' Bewegung schneller durchgeführt wird. Der Wert *1* entspricht somit der schnellsten vordefinierten Umplatzierbewegung. Auf diese Art kann Max' Bewegungsqualität auf einfache Weise verändert und in entsprechenden Szenarien eingesetzt werden. In Kapitel 7 werden die Effekte der Bewegungsgeschwindigkeit in kooperativer Interaktion vorgestellt.

Für die Integration in Max' Gesten- und Motorplanung wurde für die Umplatzierbewegung eine sogenannte *autonome* Geste definiert. Dieser Spezifikationstyp stellt eine Erweiterung von MURML dar (vgl. Kopp, S. 134). Autonome Gesten unterliegen keiner zeitlichen Bindung zu simultaner Sprache. Die Spezifikation für die Umplatzierbewegung durchläuft den in Abbildung 6.4 gezeigten Gestenplaner. In Kapitel 7 wird gezeigt, wie Max' Umplatzierbewegungen in einem Interaktionsszenario in der virtuellen Realität eingesetzt werden (vgl. Abbildung 7.10 und 7.14). Zur Laufzeit werden dafür automatisch MURML-Spezifikationen generiert. Eine Beispielspezifikation wird im Folgenden aufgeführt:

```
<constraints SkipRetraction="true" scope="right_arm">
  <sequence scope="right_arm">
    <parallel scope="right_arm">
```

```
<static end="3.50" mode="pointTo" scope="right_arm"
slot="HandLocation" start="0.50" value="830.00 -30.00 223.00"/>
  <static scope="right_arm" slot="HandShape" value="BSfist"/>
</parallel>
</sequence>

<sequence end="4.50" scope="right_arm" start="4.50">
  <parallel end="7.0" scope="right_arm" start="4.50">
    <static end="7.0" scope="right_arm"
slot="HandShape" start="4.50" value="BSflat"/>
    <static end="7.0" mode="pointTo" scope="right_arm"
slot="HandLocation" start="4.50" value="649.00 -200.00 -25.00"/>
  </parallel>
</sequence>
</constraints>
```

Die Dauer der Bewegung wird über die Attribute `start` und `end` definiert, die die Start- und Endzeit in Sekunden bestimmt. Diese Zeit wird über den beschriebenen Geschwindigkeitsfaktor moduliert. Diese Werte können über Konfigurationsparameter geändert werden. Als Standardwerte wurden für die langsamste Durchführung eine Dauer von *11.0* Sekunden und für die schnellste Durchführung eine Dauer von *3.0* Sekunden gewählt. Für die Beispielspezifikation wurde der Geschwindigkeitsfaktor auf den Wert von *0.5* gesetzt. Es ergibt sich dadurch eine Gesamtzeit von *7.0* Sekunden. Das Attribut `scope` zeigt an, mit welcher Hand das Objekt gegriffen wird. Max kann mit beiden Händen eine Umplatzierbewegung ausführen. Die Wahl für die Hand hängt davon ab, ob sich das Objekt auf der linken oder rechten Seite der Sagittalebene befindet. Die *Reich-Bewegung* und die *Umplatzier-Bewegung* werden durch den Wert `pointTo` für das `mode`-Attribut gekennzeichnet und mit der Position für das Objekt und einer gewählten Endposition übergeben. Das Attribut `slot="HandShape"` hat für die Reich-Bewegung den Wert `value="BSfist"`. Mit diesem Wert schließen sich die Finger am Ende der Reich-Bewegung, um eine Handgreifbewegung anzudeuten. Am Ende der Umplatzier-Bewegung öffnet sich die Hand wieder. Der Wert ist hier `value="BSflat"`.

Tabelle 6.2: Winkelgrenzen und Freiheitsgrade der Gelenke für Max' Vorlehnbewegung.

Gelenk	DOF	Winkelgrenzen
vl5 (5. Lendenwirbel)	2	[-10,10] Lateralflexion [-15,40] Flexion/Extension Torsion gesperrt
vl1 (1. Lendenwirbel)	3	[-10,10] Lateralflexion [-10,30] Flexion/Extension [-10,10] Torsion
vt6 (6. Brustwirbel)	3	[-20,20] Lateralflexion [-10,10] Flexion/Extension [-40,40] Torsion

Oberkörper-Vorlehnen

Im Vorlehnraum werden die beschriebenen Umplatzier-Bewegungen mit den Torsogelenken kombiniert. Liegt ein Objekt innerhalb des Vorlehnraums, so muss eine Armgreifbewegung mit den Wirbelsäulengelenken kombiniert werden. Je weiter ein Objekt vom Ursprung des Vorlehnraums entfernt liegt, desto stärker werden die Gelenke beansprucht. Die für die Vorlehnbewegung aktivierten Gelenke in Max' kinematischem Skelett sind *vt6*, *vl1* und *vl5* (vgl. Abb. 4.3). In Abschnitt 2.1.4 fanden Mark u. a. (1997), dass Menschen sich bereits bei einem Anteil von 85 % der Distanz für die *absolute kritische Grenze* vorlehnen. Diese Grenze bildet die *präferierte kritische Grenze*. Für Max' Vorlehnbewegung bedeutet dies, dass er sich nicht erst vorlehnt, wenn ein Objekt sich außerhalb der Grenze des Taktilraums befindet, sondern sich bereits bei einer Objektdistanz vorlehnt, die 85 % des Radius des Taktilraums überschreitet.

Sobald sich ein zu greifendes Objekt außerhalb von Max' präferierter kritischer Grenze befindet, werden die genannten Torsogelenke aktiviert. Jedes der Gelenke hat vordefinierte, biologisch motivierte Gelenkwinkelgrenzen (vgl. Tabelle 6.2). In der Anatomie wird üblicherweise von der Flexion gesprochen, wenn die Beugung nach vorne gemeint ist. Die Extension ist die Streckung nach hinten. Durch diese Bewegungen der Gelenke wird der Oberkörper in der Sagittalebene bewegt. Für das Vorlehnen in der Sagittalebene wird zwischen den neutralen Gelenkwinkelwerten und den obersten Winkelgrenzen für die Extension linear interpoliert. Für die Seitneigung des Oberkörpers in der Frontalebene werden abhängig von der Objektposition, zwischen den neutralen Gelenkwinkelwerten und den lateralen Flexionsgrenzen zur rechten oder linken Körperseite linear interpoliert. Torsionsbewegungen, die den Oberkörper drehen,

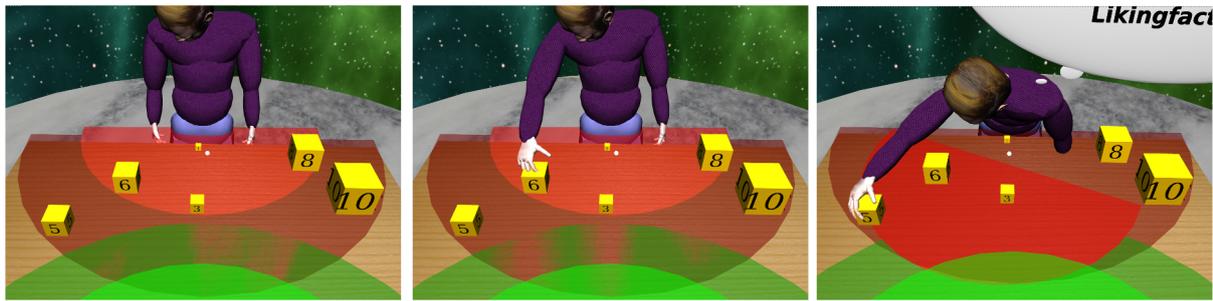


Abbildung 6.5: Visualisierung von Max' Taktilraum (kleine rote Halbscheibe) und Vorlehnraum (größere rote Halbscheibe) in der virtuellen Realität. Linke Seite: Max in neutraler Körperposition vor einem Tisch mit Objekten. Mitte: Max greift mit Arm nach Objekt 6. Rechte Seite: Max greift mit Arm nach Objekt 5 und lehnt sich vor. Der Taktilraum wird innerhalb des Vorlehnraums mitgeführt.

werden in der aktuellen Realisierung der Vorlehnbewegung noch nicht berücksichtigt. Die minimale Vorlehnbewegung findet demnach für Objektdistanzen auf der präferierten kritischen Grenze statt. Je weiter das Objekt vom Ursprung des peripersonalen Raums entfernt liegt, desto größer werden die Gelenkwinkelwerte, bis sie am Rand des Vorlehnraums die oberste Winkelgrenze erreicht haben. Auf diese Weise werden zur Objektdistanz passende Zielwerte für die Gelenkwinkel berechnet.

Für die Vorlehn-Bewegung werden in Max' Motorplanung einzelne Bewegungspositionen über sogenannte *Keyframes* generiert. Hierfür werden die Zielwerte für die Gelenkwinkel verwendet und dann Schlüsselpositionen, den Keyframes, zwischen der aktuellen Körperpostur und der Zielpostur berechnet. Eine vordefinierte Endzeit bestimmt, wie lange Max für die Vorlehnbewegung benötigt. Die Vorlehnbewegung wird in der Ablaufsteuerung der ACE-Bibliothek mit der aus der Gestenplanung generierten Umplatzierbewegung kombiniert und gleichzeitig animiert. Auf diese Weise kann Max, während er eine Umplatzierbewegung durchführt, sich zu einem Objekt vorlehnen (s. Abbildung 6.5).

Ausweichbewegungen

Sobald ein Mensch oder ein weiterer virtueller Mensch in Max' Nähe tritt, wird angenommen, dass der Partner ebenfalls von einem peripersonalen Raum umgeben ist. Die räumliche Überlappung zwischen den peripersonalen Räumen wird als neuer Unterraum registriert. Dies ist der in Abschnitt 4.1.3 definierte *Interaktionsraum*. Für den Partner wird dabei dieselbe peripersonale Raumstruktur angenommen wie für

den virtuellen Menschen Max. Die Objekte werden demnach nicht nur für Max in verschiedenen Objektkarten, sondern parallel auch in einer separaten Repräsentation für den Partner dargestellt. Der Interaktionsraum wird in Kapitel 7 im Detail beschrieben. Der neu definierte Unterraum hat Konsequenzen für Max' Bewegungsplanung. Durch den zusätzlichen Unterraum ist es möglich, den peripersonalen Raum für Kollisionsvermeidung vorzustrukturieren. Die Verknüpfung zwischen peripersonalem Raum und unterschiedlichen Berechnungsprozesse wurde auch in der Arbeit von Goerick u. a. (2005) (vgl. Abschnitt 3.2.1) vorgeschlagen. In dem hier vorgestellten Anwendungsbeispiel wird für Max' peripersonalen Raum, bestehend aus dem Taktil- und Vorlehnraum, zunächst keine Kollisionsüberprüfung bei Umplatzierbewegungen durchgeführt. Erst wenn ein Interaktionsraum registriert wird, wird die Bewegungsplanung für mögliche Kollisionen vorbereitet. Die Überlegung dahinter ist, dass wenn Max mit seinen Händen ein Objekt in den Interaktionsraum platziert, die Wahrscheinlichkeit für Kollisionen mit Armen oder Händen steigt. Statt jedoch im gesamten Interaktionsraum eine Kollisionsvermeidung durchzuführen, wird zwischen zwei Bereichen unterschieden: dem hypothetischen Taktil- und dem Vorlehnraum des Partners.

Nehmen wir an, dass Max ein Objekt mit seinen Händen den Interaktionsraum platziert, sodass der Partner das Objekt erreichen kann. Um die Wahrscheinlichkeit möglicher Kollisionen gering zu halten, wird eine Endposition innerhalb des hypothetischen Vorlehnraums des Partners gewählt. Endpositionen, die innerhalb des hypothetischen Taktilraums des Partners liegen, werden vermieden, da die Wahrscheinlichkeit von Kollisionen mit dessen Händen hier am höchsten ist. Sobald der Partner sich jedoch vorlehnt, verschiebt sich sein Taktilraum innerhalb des Vorlehnraums. Liegt die ursprünglich ermittelte Endposition durch die Vorlehnbewegung des Partners nun im Taktilraum, muss eine neue Endposition für Max' Umplatzierbewegung berechnet und somit eine neue Trajektorie errechnet werden. Hat Max mit seiner Umplatzierbewegung bereits begonnen, muss die Trajektorie an eine neue Endposition *umgelenkt* werden. Max führt damit eine Handbewegung aus, die dem Taktilraum des Partners ausweicht.

Die Trajektorienformation, wie sie für Max' Motorplanung beschrieben wurde, muss somit neu berechnet werden. Ähnlich wie in Max' Motorplanung Kollisionen zwischen guiding strokes mit dem eigenen Torso geprüft werden, können Kollisionen mit dem Taktilraum des Partners überprüft werden. Hierfür bietet sich die in Abschnitt 3.1.2 vorgestellte Methode der künstlichen Potentialfelder an. Für jede Position P_t

der geplanten Trajektorie wird der Einfluss des folgenden abstoßenden Potentialfelds $\mathbf{F}_{takilPartner}$ geprüft.

$$\mathbf{F}_{takilPartner}(\mathbf{p}) = \begin{cases} \xi \left(\frac{1}{\|\mathbf{p}\|} - \frac{1}{r_{takilPartner}} \right) \frac{\mathbf{p}}{\|\mathbf{p}\|^3} & \|\mathbf{p}\| \leq r_{takilPartner}, \\ 0 & sonst \end{cases} \quad (6.1)$$

Der Vektor \mathbf{p} beschreibt dabei den Vektor zwischen P_t und dem Ursprung $P_{takilPartner}$ des Partner-Taktilraums. Mit $\mathbf{F}_{takilPartner}$ kann ein Feldvektor $\mathbf{v}_{partnerKraft}(\mathbf{p})$ bestimmt werden, der die Kraft beschreibt, die auf den Vektor \mathbf{p} wirkt. Der Parameter ξ ist ein positiver Skalarwert, der die Feldkraft, somit die Länge des resultierenden Feldvektors bestimmt. Die Feldvektoren zeigen radial vom Ursprung des Partner-Taktilraums weg und wirken auf alle Positionen P_t , die sich innerhalb des Partner-Taktilraums mit dem Radius $r_{takilPartner}$ befinden. Es ist zu beachten, dass die Taktilräume in unseren Szenarien Halbkreise sind. Die Kraft auf Positionen P_t außerhalb des Taktilraums beträgt 0. Mit $\mathbf{v}_{partnerKraft}$ kann eine alternative Position für die Trajektorie vorgeschlagen werden, die über die inverse Kinematik für die Armkontrolle gelöst wird.

6.4 Zusammenfassung und Diskussion

In diesem Kapitel wurde ein peripersonales Raummodell für humanoide Agenten entwickelt. In einer Voranalyse wurde zunächst diskutiert, welche Eigenschaften ein peripersonales Raummodell besitzen soll, das in einem technischen System eingesetzt wird. In Anlehnung an das Prinzip der *Affordance* wurde der Begriff des *räumlichen Angebotscharakters* eingeführt, der impliziert, dass unterschiedliche räumliche Distanzen unterschiedliche Aktionen erfordern. Dieser Aspekt schlägt sich in einem peripersonalen Raummodell mit drei unterschiedlichen Unterräumen nieder. Die Unterräume sind ein Taktilraum, ein Vorlehnraum und ein visueller Aufmerksamkeitsraum, die in einer Sensorhierarchie eingeordnet sind. Objektpositionen werden in körperzentrierten Koordinaten abgelegt. Der Entwurf des peripersonalen Raummodells stützt sich auf kognitive Raummodelle für Menschen und Befunde aus der Neuropsychologie, die in der technischen Interpretation des peripersonalen Raums (vgl. Abschnitt 4.1.1) zusammengefasst wurden. Der peripersonale Raum wird über verschiedene Sensormodalitäten abgedeckt und bietet gerade für autonome humanoide Agenten eine robustere

und umfassendere Raumrepräsentation. Zudem bieten die Unterräume mögliche Ausgangspunkte zur Strukturierung der Aufmerksamkeit eines humanoiden Agenten und eine bessere Planung der Armbewegungen.

Für die technische Umsetzung des peripersonalen Raums für den virtuellen Menschen Max wurden die Arm- und Torsobewegungen für den peripersonalen Raum definiert. Als wichtigste Manipulationsaktion wurde die Umplatzierbewegung realisiert. Über Parameter kann die Bewegungsqualität der Bewegung moduliert werden. Für den Vorlehnraum ist die wichtigste Bewegung das Vorlehnen mit dem Torso. Die Vorlehnbewegung wurde nach Befunden beim Menschen modelliert, die nahelegt, dass Menschen eine bevorzugte Greifdistanz haben, ab der sie sich zur Armgreifbewegung zusätzlich vorlehnen. Für den Interaktionsraum wurde der Raum in Raumbereiche geteilt, in denen die Wahrscheinlichkeit von Kollisionen geringer ist als in anderen. Max plant dabei seine Umplatzierbewegungen um, sobald der hypothetische Takttilraum des Partners seine Armtrajektorie behindert.

Das hier vorgestellte peripersonale Raummodell wurde zunächst für den virtuellen Menschen Max und dessen biologisch motiviertes sensorisches System entworfen. Das Raummodell setzt eine egozentrische visuelle und taktile Wahrnehmung sowie die Möglichkeit des Vorlehns des Oberkörpers für humanoide Agenten voraus. Der visuelle Aufmerksamkeitsraum wurde, wie bereits erläutert, für den virtuellen Menschen Max nicht weiter verfolgt. Dieser kann aber besonders für Robotersysteme mit Kameras und Objekterkennung sinnvoll sein. Ein weiterer Aspekt, dem nicht weiter nachgegangen wurde, ist die Adaptierungseigenschaft durch Werkzeugnutzung. Die Grenzen der peripersonalen Unterräume werden über das Körperschema ermittelt. Änderungen des peripersonalen Raums bei Werkzeugnutzung ist somit möglich, wurde jedoch nicht technisch eingesetzt. Dennoch kann für die Frage, ob die Adaptierungseigenschaft durch Werkzeugnutzung dem biologischen Körperschema oder dem biologischen peripersonalen Raum zugeschrieben werden muss (vgl. Abschnitt 4.1.1), eine technische Antwort vorgeschlagen werden: Die Änderung des Körpers wird durch das künstliche Körperschema registriert und das Ergebnis schlägt sich erst in der Erweiterung des technisch umgesetzten peripersonalen Raums nieder.

7 Von peripersonalem Raum zu kooperativem Agentenverhalten

Im vorangegangenen Kapitel befassten wir uns mit der Modellierung und technischen Realisierung des individuellen peripersonalen Raums für einen virtuellen Menschen. In diesem Kapitel schlagen wir in Abschnitt 7.1 ein Modell für den Raum zwischen einem virtuellen Menschen und seinem Interaktionspartner vor. Dieses wird in einem Nahraumszenario in der virtuellen Realität eingesetzt. Es handelt sich dabei um eine Konstruktionsaufgabe, die durch kooperative Interaktion zwischen einem menschlichen Partner und Max gelöst werden soll. Die Realisierung der Konstruktionsaufgabe wird in Abschnitt 7.2 vorgestellt. Um das räumlich-kooperative Verhalten von Max sozialer zu gestalten, wurde die Konstruktionsaufgabe und Max' Verhalten mit einem Empathiemodell verknüpft. Das so erweiterte Spielszenario wird in Abschnitt 7.3 beschrieben. Im Anschluss daran wird in Abschnitt 7.4 eine Zusammenfassung und Diskussion gegeben.

7.1 Ein Modell des Interaktionsraums

Wie wir im theoretischen Überblick über interpersonale Räume (Abschnitt 2.3.2) in menschlicher Interaktion feststellen konnten, dient die Modellierung dieser Räume üblicherweise dem Zweck, soziale Interaktion abzubilden bzw. zu verstehen. Für technische Systeme bilden diese von menschlicher Interaktion abgeleiteten Modelle eine wesentliche Unterstützung zur Generierung sozial akzeptierten Verhaltens (vgl. Abschnitt 3.3). Welche kognitiven Prozesse bei Menschen dazu führen, dass sie Proxemiken und Formationen einhalten können und sich mit Partnern ohne Aushandeln im gemeinsam geteilten Raum einigen, und welche internen Repräsentationen dafür möglicherweise benötigt werden, ist für den interpersonalen Raum noch wenig erforscht. In Abschnitt 2.3 wurden sowohl soziale Verhaltensmodelle für den interpersonalen Raum als auch eine Reihe möglicher kognitiver und neuronaler Mechanismen, die dem

zugrunde liegen, vorgestellt.

Für die technische Modellierung des interpersonalen Raums zwischen einem humanoiden Agenten und einem menschlichen oder humanoiden Partner sind beide Aspekte nützlich. Unser Ziel ist es daher, einem humanoiden Agenten die Fähigkeit zu geben, das in Abschnitt 4.1.3 definierte **räumlich-kooperative Verhalten im Interaktionsraum** einzusetzen. Damit soll es dem humanoiden Agenten möglich sein

1. soziales Verhalten im Kontext des interpersonalen Raums zu generieren und
2. die räumliche Perspektive des Partners zu verstehen und diesen gegebenenfalls zu unterstützen.

Aus dem F-Formationen-System von Kendon Kendon (1990) (vgl. Abschnitt 2.3.2) wissen wir, dass Menschen, die in gemeinsame Interaktion treten, sich auf bestimmte Weise körperlich zueinander orientieren. Kendon bezeichnet den Raum zwischen Interaktanten als O-Raum (vgl. Abbildung 2.2). Unser in Kapitel 6 vorgestelltes Modell des peripersonalen Raums erlaubt es uns, diesen O-Raum näher zu spezifizieren und für ein technisches System umsetzbar zu machen. Mit dem in Definition 6 formulierten Interaktionsraum können die O-Räume in den F-Formationen wie in Abbildung 7.1 dargestellt werden. Die Definition des Interaktionsraum kann somit erweitert werden. Laut Kendon kooperieren Menschen miteinander, um den gemeinsamen Raum aufrechtzuerhalten, zu dem alle Interaktanten *direkten, exklusiven* und *gleichen Zugang* haben.

Wir beschränken unsere Betrachtung auf Formationen um einen vorgegebenen Tisch. Der Fokus liegt auf der Vis-a-Vis-Formation, da sie den größten Interaktionsraum ermöglicht. Nahraumaufgaben stellen, wie bereits in vorangegangenen Kapiteln erwähnt, eine besondere Herausforderung aufgrund des gemeinsam geteilten Raums dar. Dieser Raum wird nun explizit als für beide Partner zugänglicher Raum modelliert, und er soll damit eine Grundlage für das im Teilproblem P3 formulierte intelligente räumliche Verhalten im Interaktionsraum schaffen.

7.1.1 Räumliches Partnermodell

Wenn ein virtueller oder menschlicher Partner den peripersonalen Raum des virtuellen Agenten betritt, nimmt der Agent an, dass dieser Partner ebenfalls von einem peripersonalen Raum umgeben ist. Es wird angenommen, dass der peripersonale Raum

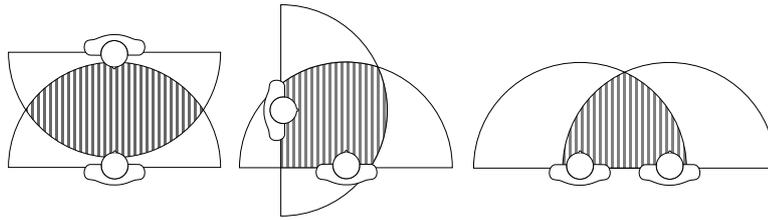


Abbildung 7.1: Kendons F-Formationen durch Interaktionsräume erweitert. Anordnungen v.l.n.r.: *vis-a-vis*, *L-* und *side-by-side*.

des Partners genauso groß ist wie der des Agenten. Die überlappenden peripersonalen Räume bilden einen Schnittraum, den *Interaktionsraum*, der für beide Partner gleich erreichbar ist. In kooperativer Interaktion wird dieser Raum als neuer Unterraum innerhalb des peripersonalen Raums definiert. Der virtuelle Agent richtet seinen Aufmerksamkeitsfokus auf diesen neuen Bereich, um zum Beispiel benötigte Gegenstände, die außerhalb der Reichweite der übrigen Partner liegen, in den neuen Unterraum zu setzen und damit die Interaktion zu unterstützen.

Max schreibt dem Partner sein eigenes Körperschema zu. Über die gesamte Interaktion wird somit der peripersonale Raum des Partners simuliert, indem Max' peripersonaler Raum auf den Partner projiziert wird. Diesen bezeichnen wir im Folgenden als *partner-peripersonalen Raum* oder als *simulierten peripersonalen Raum*. Dieser Prozess läuft somit automatisiert und ohne weitere Schlussfolgerung ab. Damit werden die Eigenschaften der *verkörperten Simulation* bei Menschen nachgeahmt.

Auf Grundlage dieser automatisierten Simulation können Hypothesen darüber gebildet werden, wie weit der Partner mit dem Oberkörper reichen kann. Im nächsten Schritt wird der Prozess der *räumlichen Perspektivenübernahme* durchlaufen. In diesem, bei Menschen vermuteten, räumlichen Kognitionsprozess werden die Objekte, die Max wahrnimmt, daraufhin überprüft, ob sie im simulierten partner-peripersonalen Raum liegen. Die mentale Rotation, die Menschen dabei durchführen, wird für Max über die Umrechnung der Objektpositionen in das Ursprungskordinatensystem des partner-peripersonalen Raums realisiert. Die Objekte werden in einer *simulierten peripersonalen Raumkarte* registriert. In dieser befinden sich Max' Hypothesen darüber, welche Objekte der Partner erreichen kann, welche nicht und für welche er sich vorlehnen muss.



Abbildung 7.2: Linke Seite: Mensch in virtuellem Raum mit 3-D-Brille und Wii-Eingabegerät. Rechte Seite: Visualisierung des Taktilraums bewegt sich mit Position der 3-D-Brille.

7.2 Konstruktionsaufgabe in kooperativer Interaktion

Um den Einsatz des peripersonalen Raum-Modells und Interaktionsraum-Modells zu zeigen, wurde ein Szenario entworfen, in dem der virtuelle Mensch Max und ein menschlicher Partner eine Konstruktionsaufgabe zu bewältigen haben. Die Nahraum-Aufgabe soll in Kooperation gelöst werden und ist als Spielszenario angelegt.

7.2.1 Technischer Aufbau

In einem CAVE-ähnlichen virtuellen Raum sieht der menschliche Partner mit einer 3-D-Brille die dreidimensionale Projektion eines virtuellen Tisches vor sich. Auf dem Tisch sieht der Partner direkt vor sich die zwei peripersonalen Unterräume, die Max für seinen Partner simuliert (siehe Abschnitt 7.1.1). Abbildung 7.2 zeigt das gesamte Setup. Die Position der 3-D-Brille wird über das ART-Tracking-System¹ verfolgt. Bewegt sich der menschliche Partner mit dem Oberkörper nach vorne, so bewegt sich, wie bei Max' Vorlehnbewegungen auch, die Halbscheibe für den Taktilraum vom Körper des Partners weg. Da die Simulation des partner-peripersonalen Raums nur eine Hypothese ist, kann bei Menschen, die nicht Max' Körperlängen besitzen, die Visualisierung des Taktilraums über den Vorlehnraum hinaustreten oder nicht bis zu dessen Rand gelangen (siehe Abbildung 7.2, rechte Seite). Der Vorlehnraum des Menschen wird

¹<http://www.ar-tracking.com> ; Letzter Zugriff: 27. Sep. 2012

im entworfenen Szenario nicht verschoben, wenn der Mensch sich im virtuellen Raum fortbewegt, da Fortbewegung in der Konstruktionsaufgabe nicht modelliert wurde.

Der Mensch kann mit einem Handeingabegerät virtuelle Objekte auf dem Tisch frei bewegen. Zentral in diesem Szenario ist, dass das Eingabegerät eine Objektmanipulation vorgibt, bei der der menschliche Partner die Hand nah an das Objekt bringen muss. So erfordert zum Beispiel ein weiter entfernt gesehenes Objekt einen höheren Einsatz von Arm- und Torsobewegungen des Menschen. Zwei Arten der Objektmanipulation wurden für das Spielszenario umgesetzt: einmal über ART-Fingertracker und einmal über eine Wiimote. Die genutzten Fingertracker sind drei Fingerschienen für Daumen, Zeige- und Mittelfinger, die über Marker von dem ART-Tracking-System aufgenommen werden. Da die Orientierung der Hand sowie die Fingerpositionen verfolgt werden, können auf diese Weise menschliche Handposturen, z. B. Greifposturen, in der virtuellen Realität gemessen werden. Der Mensch muss für das Greifen die Hand nah an das virtuelle Objekt bringen, eine Schließbewegung mit den Fingern durchführen und kann das Objekt dann bewegen. Soll es abgesetzt werden, reicht eine Öffnungsbewegung mit den Fingern. Diese Eingabemethode hat den Vorteil, natürliche Greifgesten in der Objektmanipulation zur Verfügung zu stellen. Die Kalibrierung für verschiedene Fingergrößen sowie Greifarten ist jedoch aufwendig und muss individuell vor jeder Interaktion durchgeführt werden. Zudem ist diese Eingabe für Gestenfehlererkennung anfällig.

Um zum einen einfache und robuste Interaktionen in der virtuellen Umgebung für nicht versierte Partner zu ermöglichen, gleichzeitig aber natürlich wirkende Objektmanipulationen nachzustellen, wurde als Eingabegerät die Nintendo Wiimote gewählt. Der menschliche Partner hält die Wiimote wie eine Fernbedienung in der Hand. So wird die Position der Hand ebenfalls über das ART-Tracking-System verfolgt. Möchte der Partner ein virtuelles Objekt *greifen*, so muss er mit der Wiimote so nah wie möglich in die virtuelle Objektprojektion hineinreichen. Drückt er einen vordefinierten Knopf, wird das Objekt mit der Position der Wii nah an der Hand mitgeführt. Bei wiederholtem Knopfdruck wird die Objektverbindung gelöst, und das virtuelle Objekt fällt auf den virtuellen Tisch oder kann auf weitere virtuelle Objekte gesetzt werden (siehe Abbildung 7.3). An der gegenüberliegenden Seite steht Max. Auf dem Tisch sind ebenfalls seine peripersonalen Unterräume, wie sie in Kapitel 6 beschrieben wurden, projiziert. Die obere Tischkante ist auf Max' Hüfthöhe. Der Interaktionsraum beider Partner befindet sich zu Beginn in der Mitte des Tisches. Wenn die Konstruk-

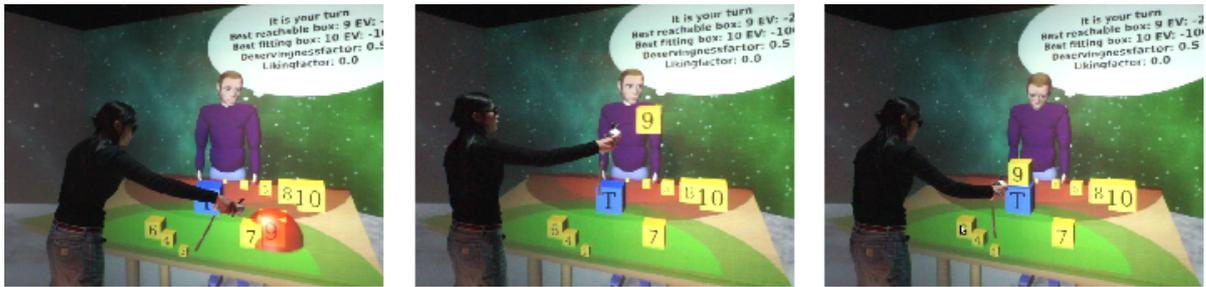


Abbildung 7.3: Menschlicher Partner greift mit der Wii ein Objekt und setzt es auf den Turm.

tionsaufgabe gestartet wird, werden virtuelle quaderförmige Objekte zufällig auf dem Tisch verteilt. Über Max' Kopf ist eine computergrafisch visualisierte *Denkblase* zu sehen (s. Abbildung 7.3). In dieser wird pro Spielzug Max' Wissen über den Stand der Spielparameter angezeigt.

7.2.2 Spielregeln

Das Ziel der Konstruktionsaufgabe ist es, in Kooperation mit einem Partner einen Turm zu bauen. Die Partner stapeln dabei nacheinander jeweils einen virtuellen Block auf einen anderen. Es darf kein Block gesetzt werden, der größer ist als der oberste Block auf dem Turm. Diese Regel ist auch in dem aus der Mathematik bekannten Problem der *Türme von Hanoi* bekannt. Die virtuellen Blöcke sind mit unterschiedlichen Ziffern versehen und unterscheiden sich in ihrer Größe. Die Ziffern steigen mit der Größe der Blöcke, das bedeutet, der größte Block hat die höchste Ziffer. Bei der Initialisierung des Spielszenarios ist bereits der größte Block in der Mitte des Interaktionsraums vorgegeben, um deutlich zu machen, wo die restlichen virtuellen Blöcke aufeinandergestapelt werden sollen. Die übrigen virtuellen Blöcke werden bei der Initialisierung zufällig in den beiden peripersonalen Räumen verteilt. Die Anzahl der Blöcke, die in den einzelnen peripersonalen Räumen liegen, kann unterschiedlich sein. Es ist jedoch immer eine Mindestanzahl an Blöcken für die Partner erreichbar.

Das übergeordnete Ziel für beide Partner ist es, den Turm so hoch wie möglich zu stapeln. Die Partner müssen dabei folgende Regeln beachten. Erstens: Die Blöcke können nach ihrer direkten Ziffernreihenfolge gestapelt werden, zum Beispiel *Block-2* auf *Block-3*. Werden alle Blöcke auf diese Weise gestapelt, entsteht der höchste zu erreichende Turm. Zweitens: Die Blöcke können unabhängig von der direkten Ziffernfolge nach Größe gestapelt werden, zum Beispiel *Block-2* auf *Block-4*. *Block-3* wurde da-

bei ausgelassen und kann nicht mehr gesetzt werden, was zu einem niedrigeren Turm führt. Um den höchsten Turm stapeln zu können, muss jeder Partner den besten Block auf den Turm setzen. Der beste Block ist immer der nächste in der Ziffernfolge. Es kann aber der Fall auftreten, dass der beste Block nicht erreichbar ist für den Spieler, der im aktuellen Spielzug an der Reihe ist. Das heißt, der Block liegt nicht in dessen peripersonalen Raum. In diesem Fall muss der Spieler den nächstbesten Block setzen, der, wie zuvor beschrieben, zu einem niedrigeren Turm führt. Ungültige Blöcke können nicht auf den Turm gesetzt werden und fallen automatisch auf den Tisch. Ein Spielzug ist beendet, wenn ein gültiger Block auf den Turm gesetzt wurde. Das gesamte Spiel endet, wenn der kleinste Block auf den Turm gesetzt wurde.

7.2.3 Szenario-Komponenten

Das Interaktionsraum-Spielszenario kann in drei zentrale Komponenten eingeteilt werden: die *Visualisierungskomponente*, die *Spielkontrolle* und die *interne Repräsentation*.

In der *Visualisierungskomponente* fassen wir die grafische Darstellung aller Objekte in der virtuellen Welt, einschließlich der Visualisierung des virtuellen Menschen Max und der Steuerung der Eingabegeräte (Abschnitt 7.2.1), zusammen. Das 3-D-Computergrafik-Framework InstantReality ermöglicht es, die virtuelle Szene über Softwareskripte zu definieren. Auf diese Weise wurden die Interaktionsmöglichkeiten mit den Szenario-Objekten für menschliche Partner und den virtuellen Menschen Max realisiert. Dabei mussten für menschliche Partner die Eingabemethoden über die Wiimote, aber auch die Manipulationsmöglichkeiten für den virtuellen Menschen Max definiert werden. Die visuell dargestellte Interaktionsszene besitzt Ein- und Ausgabeschnittstellen zur Spielkontroll-Komponente.

Die *Spielkontrolle* bildet die Schnittstelle zwischen der Visualisierungskomponente und den übrigen Softwarekomponenten, die nicht grafisch dargestellt werden. Die Spiellogik im BDI-Modul beinhaltet die Spielregeln und hält das Wissen über den Spielzustand bereit. Max' Wissen und seine Hypothesen über den Spielstand werden für jeden Spielzug in der *Denkblase* visualisiert. Auf Basis der Spielstände werden in der Spiellogik neue Spielzüge generiert. Die Spielkontrolle greift auf die interne Repräsentation zu, um nach außen sichtbares Verhalten für Max zu generieren. Interagiert der menschliche Partner in der virtuellen Szene, werden die Eingaben an die Spielkontrolle weitergegeben, und gegebenenfalls wird ein neuer Spielstand in der Spiellogik

berechnet.

Der Zugriff auf die *interne Repräsentation* des peripersonalen und des Interaktionsraums geschieht über die Spielkontrolle. Diese Komponente fasst das Wissen über folgende Aspekte zusammen: Max' *peripersonaler Raum* mit dem Interaktionsraum als dessen Unterraum, die *peripersonalen Raumkarten*, die die Objektpositionen beinhalten, und für den menschlichen Partner eine *Simulation der peripersonalen Raumkarten* mit Hypothesen über die Objektpositionen. Die Basis für die simulierten peripersonalen Raumkarten ist dabei Max' peripersonaler Taktil- und Vorlehnraum.

7.2.4 Spielabläufe und interne Verarbeitung

Die unterschiedlichen Spielzüge der beiden Partner und die dabei stattfindende interne Verarbeitung werden in den nächsten Abschnitten anhand von Flussdiagrammen deutlich gemacht. In den Flussdiagrammen werden drei Ablaufebenen dargestellt: das beobachtbare Verhalten der Interaktionspartner, die internen Prozesse in der Spielkontrolle und die Verarbeitung in der internen Repräsentation. Die grau unterlegten Kästen zeigen, welche Informationen in Max' Denkblase angezeigt werden. Da in diesem Szenario die Objektsuche nicht im Fokus steht, werden bereits zu Beginn die initialen Objekte in den peripersonalen Raumkarten registriert. Ändern sich die Objektpositionen, zum Beispiel wenn der menschliche Partner ein Objekt verschiebt, so werden diese automatisch aktualisiert.

Spielzug von Max

Wenn Max an der Reihe ist (Max hat *TURN*), wird sein Spielzug (siehe Flussdiagramm in Abbildung 7.4) über die Spielkontrolle abgearbeitet. In Max' Denkblase wird angezeigt, dass er an der Reihe ist (*It's my turn.*). Er blickt zuerst auf zufällige Blöcke, um dem Partner zu signalisieren, dass er mit seinem Zug beginnt. In der Spiellogik wird der beste Block ermittelt, eine Blickbewegung zu diesem wird ausgelöst und die Ziffer des besten Blocks in der Denkblase angezeigt (*best fitting box*). Die Spielkontrolle prüft anhand Max' peripersonaler Raumkarten, ob der Block für ihn erreichbar ist. Ist dies der Fall, so schaut er erst den Partner an, und die Spielkontrolle löst dann eine Umplatzierungsbewegung (bzw. Replatzierungsbewegung), wie sie in Abschnitt 6.3.2 beschrieben wurde, aus. Kann Max den besten Block nicht erreichen, ermittelt die Spiellogik anhand der peripersonalen Raumkarten den nächstbesten Block, der in

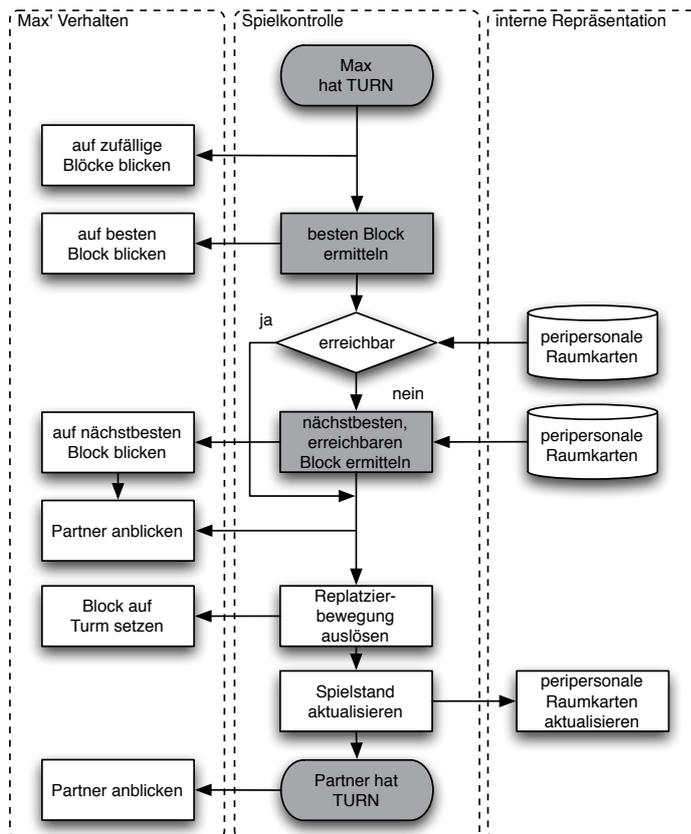


Abbildung 7.4: Spielzug von Max.

Max' peripersonalem Raum liegt. Es ist zu beachten, dass nicht geprüft wird, ob der Block am nächsten zu Max' Körper liegt, sondern, ob der nächstbeste Block im peripersonalen Raum liegt oder nicht. Die Ziffer des nächstbesten Blocks wird ebenfalls in der Denkblase angezeigt (*best reachable box*). Ist der nächstbeste Block ermittelt, blickt Max auf diesen, und die Umplatzierungsbewegung wird mit dem ermittelten Block durchgeführt. Die Parameter sind die eindeutige Objekt-ID des besten Blocks und die Objekt-ID des Blocks, der an der Spitze des Turms steht. Für die Umplatzierungsgeschwindigkeit wird ein Standardwert von 0.5 gewählt. Dieser entspricht der mittleren Geschwindigkeit zwischen schnellster und langsamster Bewegung. Nachdem der Block auf den Turm gesetzt wurde, werden der Spielstand in der Spiellogik und Max' peripersonale Raumkarten aktualisiert. Die Spiellogik meldet das Ende des Spielzuges und löst eine Blickbewegung von Max zum Partner aus, um diesem mitzuteilen, dass er als Nächstes an der Reihe ist.

Spielzug des Partners

Wenn der menschliche Partner an der Reihe ist, wechselt der Turn-Status in Max' Denkblase (*It's your turn.*). Innerhalb dieses Spielzuges sind alle Parameter in der Denkblase Hypothesen von Max über die Spielmöglichkeiten des Partners. Der Spielzug verläuft demnach ähnlich wie bei Max (vgl. Abbildung 7.5 und 7.4). Die Spiellogik ermittelt über die bereits auf den Turm gesetzten Blöcke den aktuell besten Block im Spiel. Die dazugehörige Ziffer wird in Max' Denkblase angezeigt. Im nächsten Schritt wird in der Spielkontrolle versucht, Wissen für Max darüber zu ermitteln, ob der Partner diesen Block erreichen kann. Dieses Wissen kann Max jedoch nicht über das verfügbare Faktenwissen aus der Umgebung oder über die Spielregeln ermitteln. An diesem Punkt kommt das in Abschnitt 7.1.1 definierte räumliche Partnermodell zum Einsatz. Dieses beinhaltet zum einen eine Hypothese des partner-peripersonalen Raums, die über den Mechanismus der verkörperten Simulation gebildet wurde. Zum anderen beinhaltet es die dazugehörigen simulierten peripersonalen Raumkarten, die die Objekte relativ zum Körper des Partners darstellen. Diese wurden über die räumliche Perspektivenübernahme ermittelt.

Im Spielablauf wird über die simulierten peripersonalen Raumkarten, wie bei Max, geprüft, ob der beste Block erreichbar ist, und gegebenenfalls wird ein nächstbester ermittelt. Beide Parameter zeigt Max in seiner Denkblase an. Nimmt der Partner einen

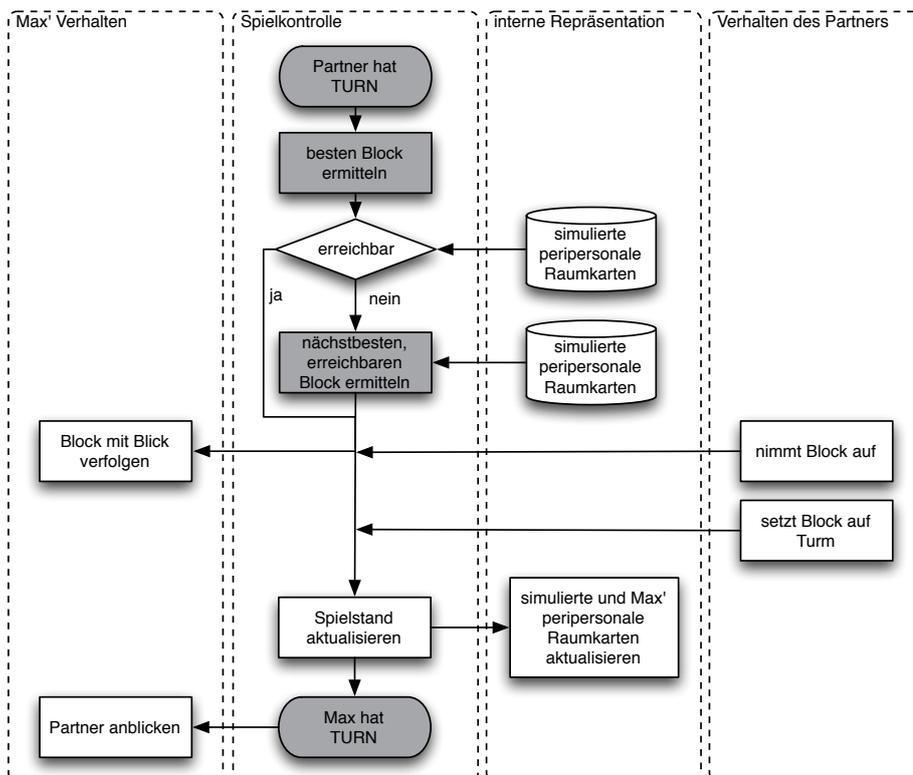


Abbildung 7.5: Spielzug des Partners aus Sicht von Max.

Block auf, so verfolgt Max den Block mit dem Blick. Der vom Partner gewählte Block kann dabei den Hypothesen von Max widersprechen. Hat der Partner einen Block auf den Turm gesetzt, ist sein Spielzug zu Ende, und Max schaut den Partner an. Der Spielstand in der Spiellogik, die simulierte partner-peripersonale Raumkarte und Max' peripersonale Raumkarte werden aktualisiert.

Max zeigt räumlich-kooperatives Verhalten

Es ist im Spiel erlaubt, sich gegenseitig zu helfen. Für Max wird dabei eine *Hilfsaktion* geplant. Diese ist definiert als das Platzieren eines besten Blocks in einen Raumbereich, der für den Partner erreichbar ist. Sie besteht aus einer *Hilfsbewegung*, einem zu bewegendem *Hilfsobjekt* und einer *Hilfsposition*. Die Hilfsbewegung ist eine Umplatzierungsbewegung, mit der Max ein Objekt, das der Partner benötigt (das Hilfsobjekt), zu einer Zielposition (der Hilfsposition) stellt, die der Partner erreichen können soll. Dieses Verhalten bezeichnen wir als *räumlich-kooperatives Verhalten*, wie es in Abschnitt 4.1.3 in Definition 5 formuliert wurde. Max und der Partner teilen sich den Interaktionsraum, damit ist die erste Eigenschaft von Definition 5 erfüllt: Die Verpflichtung zur gemeinsamen räumlichen Aktivität in einem mit einem Partner geteilten Raum. Max zeigt durch die räumliche Perspektivenübernahme Verhalten, das eine räumliche Hürde für den Partner überwindet. Der Partner wird dadurch unterstützt, seine Aufgabe selbst zu lösen. Damit ist die zweite Eigenschaft räumlich-kooperativen Verhaltens erfüllt: Die Verpflichtung zur gegenseitigen Unterstützung in den räumlichen Aktivitäten.

Die Hilfsaktion kann geplant werden, wenn der Partner an der Reihe ist. In Abbildung 7.6 wurde das Flussdiagramm für den Spielzug des Partners um die Hilfsaktion erweitert. Diese setzt ein, wenn Max über die räumliche Perspektivenübernahme annimmt, dass der Partner den besten Block nicht erreichen kann. Der beste Block ist das Hilfsobjekt, das Max für den Partner erreichbar machen möchte. Dafür muss zunächst eine Zielposition ermittelt werden. Da Max annimmt, dass der Interaktionsraum für beide Partner erreichbar ist, wird eine zufällige freie Position innerhalb dieses Raums als Hilfsposition ausgewählt. Die Umplatzierungsbewegung wird dann mit dem Hilfsobjekt und der Hilfsposition ausgelöst. Für den Geschwindigkeitsparameter der Umplatzierungsbewegung wird ein Wert von 0.75 gewählt. Diese Bewegung ist um 25 % schneller als die Umplatzierungsbewegung, die Max durchführt, wenn er während

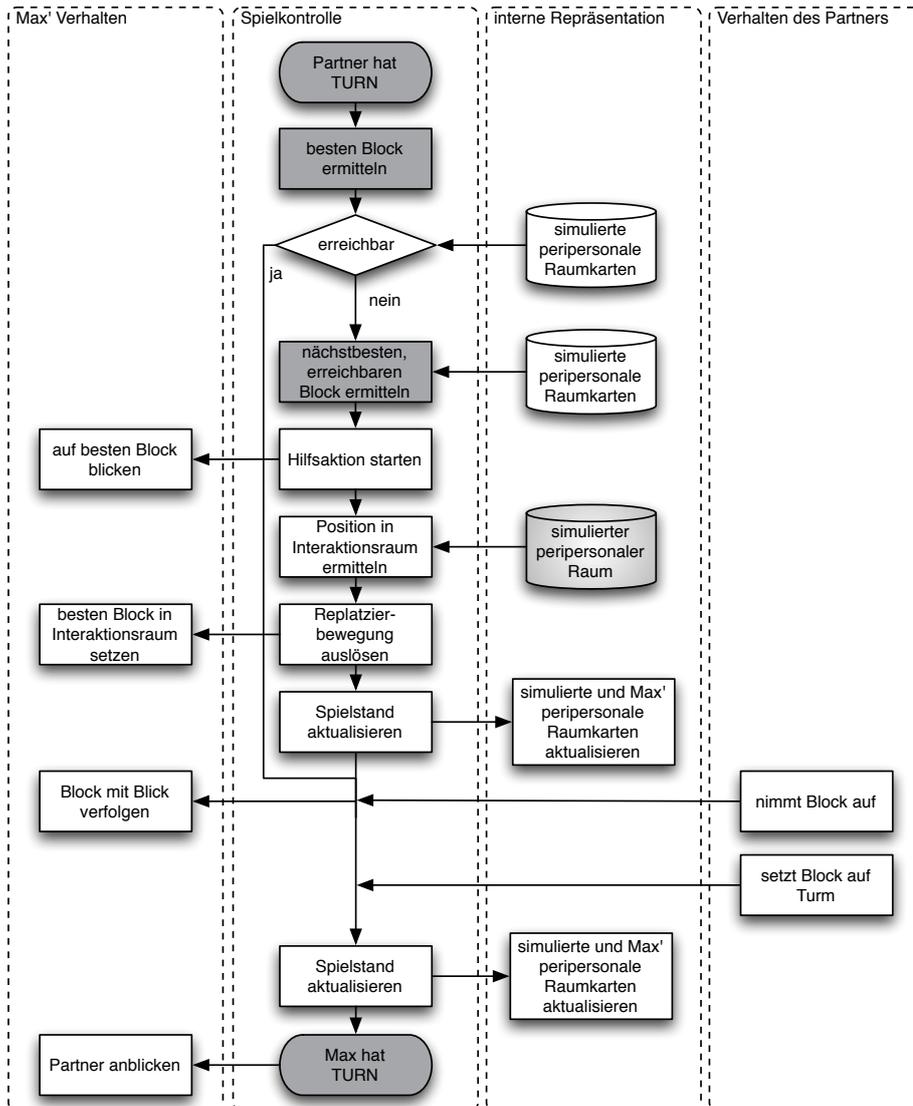


Abbildung 7.6: Max hilft dem Partner.

seines eigenen Spielzuges einen Block auf den Turm setzt. Mit dieser erhöhten Bewegungsgeschwindigkeit soll dem menschlichen Partner eine größere Unterstützungsbereitschaft signalisiert werden. Ist die Hilfsaktion beendet, werden die peripersonalen Raumkarten von Max und dem Partner aktualisiert, und Max blickt den Partner an. Der Partner kann nun seinen Spielzug durchführen. Dabei kann es sein, dass der Partner nicht das Hilfsobjekt wählt. Zum einen kann es sein, dass das Objekt auch nach der Hilfsaktion nicht erreichbar ist, da es sich bei der Hilfsposition nur um Max' Hypothese handelt. Darüber hinaus ist es möglich, dass der Partner sich auch bei einem erreichbaren besten Block entscheidet, einen anderen Block zu wählen. Sobald der Partner einen Block aufnimmt, verfolgt Max den Block mit seinem Blick. Setzt der Partner diesen auf den Turm und ist der Block gültig, so werden der Spielstand und alle peripersonalen Raumkarten aktualisiert.

7.3 Empathie als Basis für räumlich-kooperatives Verhalten

Das bisher vorgestellte Spielszenario zeigt, wie die Modellierung des Interaktionsraums verwendet werden kann, um Max's Hilfsaktion in einer Nahraumaufgabe auszulösen. Max nutzt bisher nur das Wissen über die räumliche Perspektive des Partners. Darüber ermittelt Max, ob ein von dem Partner benötigtes Objekt erreichbar ist, und er kann es ihm an eine Position setzen, die der Partner erreichen kann. Diese Hilfsaktion bezeichnen wir als räumlich-kooperatives Verhalten. Max führt dabei eine Hilfsbewegung immer dann aus, wenn der Partner das Objekt nicht erreichen kann. Die Hilfsposition ist ein beliebiger freier Platz innerhalb des Interaktionsraums. Mit unserer Definition 5 aus Kapitel 4 wurde räumlich-kooperatives Verhalten als ein Aspekt von sozialem Verhalten eingeführt. Die Frage, die sich stellt, ist nun, welche Hilfsstrategie, die Max in der Nahraumaufgabe verfolgt, bei menschlichen Interaktionspartnern tatsächlich als kooperativ oder sozial wahrgenommen wird. Wir betrachten im Folgenden drei verschiedene Hilfsstrategien:

1. Max hilft über eine zufällige Hilfsposition: Die Hilfsposition ist eine *zufällige* Position innerhalb des Interaktionsraums. Dies ist die bisherige Hilfsstrategie.
2. Max hilft über die maximale Hilfsposition: Die Hilfsposition ist die Position

innerhalb des Interaktionsraums, mit der kürzesten Distanz zum Partner. Max führt dabei die *maximale Vorlehnbewegung* aus.

3. Max hilft über eine Hilfsposition, die von sozialen Faktoren abgeleitet wird: Die Hilfsposition ist eine Position innerhalb des Interaktionsraums, die über *soziale Mechanismen* ermittelt wird.

Es ist zu vermuten, dass eine Hilfsbewegung, bei der sich Max bis an seine Körpergrenzen vorlehnt (Hilfsstrategie 2), auf den menschlichen Partner nicht sozial oder kooperativ wirkt. Stattdessen könnte diese als verwirrend oder gar devot interpretiert werden. Die dritte Hilfsstrategie ist aus menschlicher sozialer Interaktion motiviert und Gegenstand der folgenden Abschnitte.

7.3.1 Modulierung von Hilfsaktionen über soziale Faktoren

Die in den folgenden Abschnitten vorgestellte Arbeit ist das Ergebnis der Zusammenarbeit zwischen den SFB-Teilprojekten A1 und C1. Letzteres bildet, wie in der Einleitung erwähnt, den Forschungskontext dieser Dissertation. Unter der Leitung von Prof. Dr. Ipke Wachsmuth konnte das Interaktionsraum-Szenario mit dem von der Projektpartnerin Hana Boukricha entwickelten Empathiemodell (Boukricha und Wachsmuth, 2011) verknüpft werden. Die Verknüpfung selbst wurde in Boukricha, Nguyen und Wachsmuth (2011) vorveröffentlicht.

Aus menschlicher sozialer Interaktion wissen wir, dass Empathie als Motivationsbasis für pro-soziales und kooperatives Verhalten angenommen wird (Hoffman, 2000). Es wird außerdem angenommen, dass Empathie eine große Rolle bei moralischen Vorgängen wie Mitgefühl, Fürsorge oder Helfen spielt. Menschen empathisieren jedoch miteinander in verschiedenen Graden. Die Empathie mit jemandem hängt dabei von unterschiedlichen Modulierungsfaktoren ab. Der Grad der Empathie kann zum Beispiel von der sozialen Beziehung, von der eigenen Stimmung oder vom Situationskontext abhängen. Das Ziel ist für uns daher, Hilfsstrategie 3 umzusetzen. Max soll über seinen *Empathiegrad* eine geeignete Hilfsposition innerhalb des Interaktionsraums ermitteln. Der Empathiegrad wird sowohl aus räumlichen als auch aus festgelegten sozialen Faktoren bestimmt.

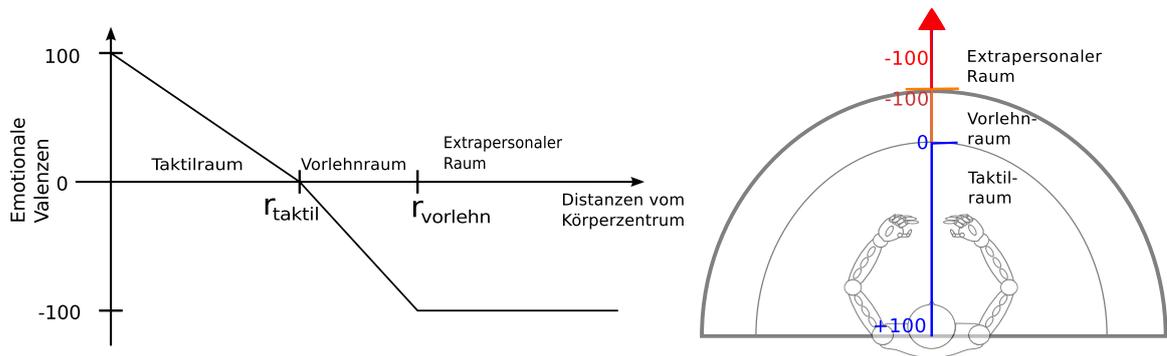


Abbildung 7.7: Kostenfunktion, die räumliche Distanzen, beginnend von der Mitte des peripersonalen Raums, auf emotionale Valenzen abbildet.

7.3.2 Kostenfunktion

Um Hilfsstrategie 3 zu realisieren, wurde zunächst eine Kostenfunktion definiert (siehe Abbildung 7.7). Diese verknüpft Distanzen innerhalb des peripersonalen Raums mit emotionalen Valenzwerten. Über die Verknüpfung wird eine emotionale Bewertung der räumlichen Distanzen vorgenommen, die als Aufwand für bestimmte Motoraktionen aufzufassen ist. Max' Spielzüge werden über eine Sequenz von Plänen durchgeführt, die in seinem BDI-Modul (siehe Abschnitt 4.2) definiert sind. Wenn Max an der Reihe ist, einen Block auf den Turm zu setzen, sucht er zunächst nach dem besten Block. Abhängig von der Position des Blocks wird ein emotionaler Valenzwert ausgelöst. Der Wert wird über die Kostenfunktion bestimmt. Blöcke, die in Max' Taktierraum liegen, sind einfach mit den Armen zu erreichen. Er benötigt dafür wenig motorischen Aufwand, der als geringere *Kosten* dargestellt wird. Daher werden Objektpositionen innerhalb des Taktierraums mit positiven emotionalen Valenzwerten im Wertebereich $[0,100]$ assoziiert. Der höchste Valenzwert liegt im Ursprung des Taktierraums. Je näher ein Objekt also an Max' Körper liegt, desto höher steigt der emotionale Valenzwert. Blöcke, die im Vorlehnraum positioniert sind, sind zwar für Max mit den Armen erreichbar, erfordern jedoch mehr *Kosten* und sind daher mit negativen emotionalen Valenzwerten belegt. Der Wertebereich ist $[-100,0]$. Am Rand des Taktierraums liegt der Wert bei 0 und sinkt bis auf einen Wert von -100, je näher ein Objekt am Rand des Vorlehnraums positioniert ist. Blöcke, die im extrapersonalen Raum liegen, sind für Max nicht erreichbar und werden daher mit einem negativen emotionalen Valenzwert von -100 bewertet. Wenn während Max' Spielzug der beste Block nicht erreichbar ist, weicht er auf den nächstbesten und für ihn erreichbaren Block aus und setzt diesen

auf den Turm. Dieser Spielzug ist als *Goal-Success* im BDI-Modul definiert und löst einen positiven Valenzwert von *60* aus. Wenn Max keinen Block setzen kann, setzt sein Spielzug aus. Dies ist als *Goal-Failure* definiert und löst einen negativen Valenzwert von *-100* aus.

Max verwendet diese Kostenfunktion nicht nur, um seine eigenen peripersonalen Raumdistanzen zu bewerten, sondern auch um seinem Partner dieselbe Funktion zuzuschreiben. Dieser Prozess ist daher dem kognitiven Mechanismus der verkörperten Simulation zuzuordnen. Der Simulation des partner-peripersonalen Raums (siehe Abschnitt 7.2.4) wird demnach noch eine weitere Hypothese hinzugefügt.

7.3.3 Modulierung der Hilfsbewegung über den Grad der Empathie

Es werden nun die sozialen Faktoren erläutert, die den *Empathiegrad* modulieren. Die Modulierung findet über drei Werte statt: den *Mood*-, den *Liking*- sowie den *Deservingness*-Faktor. Um die Konsistenz mit verwandten Arbeiten und den Vorveröffentlichungen zu bewahren, verwenden wir im weiteren Verlauf die englischen Begriffe.

Mood: Dieser Wert beschreibt die emotionale Stimmung als einen PAD-Wert. Dieser wird im Emotionsmodul von Max simuliert (siehe Abschnitt 4.2).

Liking: Dieser Faktor beschreibt, wie sehr bzw. bis zu welchem Grad Max seinen Interaktionspartner mag. Dieser Faktor wird aus dem simulierten Empathiegrad des Partners ermittelt. Es handelt sich somit um eine Hypothese darüber, was Max glaubt, wie viel Empathie der Interaktionspartner für Max aufbringt. Diese Schlussfolgerung wird aus der Hilfsposition des Partners ermittelt. Je mehr der Partner aufwendet, um Max zu helfen, das heißt, je näher die Hilfsposition an Max liegt, desto höher schätzt Max den Empathiegrad des Partners ein. Damit steigt der Liking-Faktor, und Max mag den Partner mehr.

Deservingness: Der Deservingness-Faktor beschreibt, wie sehr Max meint, dass der Partner die Situation, in der er sich befindet, verdient hat. Dieser Wert ist das Verhältnis zwischen der Anzahl der Blöcke, die in Max peripersonalem Raum liegen und noch im Spiel eingesetzt werden können, und der Gesamtanzahl der Blöcke, die im ganzen Spiel noch eingesetzt werden können. Der Wert liegt somit zwischen *0* und *1*. Das heißt, je mehr erreichbare und setzbare Blöcke Max in seinem eigenen Taktiraum hat, desto höher steigt der Deservingness-Faktor und desto mehr schätzt Max die

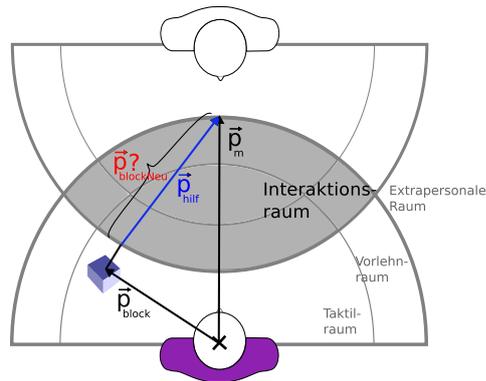


Abbildung 7.8: Ermittlung der Hilfsposition.

Situation des Partners als *unberechtigt* bzw. *unverdient* ein.

Ermitteln des Empathiegrades

Für die Verknüpfung mit dem Interaktionsraum-Szenario ist wichtig zu wissen, welche Eingaben das Empathiemodul benötigt und was es zurückliefert. Die genaue Berechnung des Empathiegrades wird hier nicht im Detail erläutert. Details zum Empathiemodell und zur Berechnung des Empathiegrades sind in den erwähnten Veröffentlichungen von Boukricha und Wachsmuth (2011) und Boukricha u. a. (2011) beschrieben. Als Ausgabewert liefert das Empathiemodul einen Wert zwischen 0 und 1. Als Eingabewerte werden zum einen der *Deservingness*-Faktor, ein möglicher *Liking*-Faktor und emotionale Valenzwerte aus der Kostenfunktion übergeben.

Berechnung des Hilfsgrades für Max

Sobald eine Hilfsaktion ausgelöst wird, setzt Max den besten Block an eine Position, die der Partner erreichen kann, das heißt an eine Position innerhalb des Interaktionsraums. Bisher konnte dies eine beliebige freie Position innerhalb des Interaktionsraums sein. Mit der Modulierung der Hilfsaktion über soziale Faktoren wird die Hilfsposition nun über den Empathiegrad berechnet. Die Grenzen des Interaktionsraums sind durch zwei Kreisbögen festgelegt. Diese werden durch die beiden Vorlehnräume der Interaktionspartner aufgespannt. Den Kreisbogen, der von Max aufgespannt wird, bezeichnen wir als $\text{vorlehnArc}_{\text{Max}}$ und den vom Partner aufgespannten als $\text{vorlehnArc}_{\text{partner}}$. Die im weiteren Verlauf definierten Vektoren haben als Ursprung Max' peripersonalen Raum. Die maximale Hilfsposition ist die räumlich nächste Position P_m zum Part-

ner, die Max noch erreichen kann. Sie ist gegeben durch den Schnittpunkt zwischen $vorlehnArc_{Max}$ und der Strecke, die die beiden Ursprünge miteinander verbindet. Der Ortsvektor von P_m ist gegeben durch \mathbf{p}_m . Abbildung 7.8 zeigt schematisch, wie die Hilfsposition des Blockes ermittelt wird.

Abhängig von Max' Empathiegrad $degEmp$ wird seine Hilfsbewegung moduliert. Das bedeutet, dass $degEmp$ bestimmt, wie nah ein benötigter Block mit Vektor \mathbf{p}_{block} an den Partner herangestellt wird. Mit der folgenden Formel wird die Hilfsposition $\mathbf{p}_{blockNeu}$ des Blockes errechnet:

$$\mathbf{p}_{blockNeu} = \mathbf{p}_{block} + (degEmp * \mathbf{p}_{hilf}), \text{ mit } \mathbf{p}_{hilf} = \mathbf{p}_m - \mathbf{p}_{block} \quad (7.1)$$

Die Formel 7.1 gilt nur, wenn sich die Blockposition vor der Hilfsaktion innerhalb des Interaktionsraums befindet. Ist dies nicht der Fall, wird für die Berechnung \mathbf{p}_{block} an einen neuen Punkt verschoben. Dieser wird berechnet durch den Schnittpunkt von $vorlehnArc_{partner}$ und der Strecke zwischen P_m und \mathbf{p}_{block} . Dies ist die minimale Hilfsposition für einen außerhalb des Interaktionsraums liegenden Block. Diese minimale Hilfsposition wird dann als neuer Referenzpunkt für \mathbf{p}_{block} in die Formel 7.1 eingesetzt. Dieses Vorgehen für Blöcke außerhalb des Interaktionsraums stellt sicher, dass der Empathiegrad $degEmp$ nur den Teil von \mathbf{p}_{help} moduliert, der sich innerhalb der Interaktionsraumgrenzen befindet. Nur dieser Anteil trägt in unserer Definition zu einer Hilfsaktion bei. Die Überlegung ist, dass die räumliche Strecke, die während der Hilfsbewegung außerhalb des Interaktionsraums überbrückt wird, für den Partner nicht erreichbar ist und daher nicht zu einer Hilfsaktion beitragen kann.

Die berechnete Hilfsposition sagt also etwas über den Hilfsgrad der Umplatzierungsbewegung aus und bewegt sich zwischen der minimalen und maximalen Hilfsposition. Die Hilfsbewegung ist die in Abschnitt 6.3.2 beschriebene Umplatzierungsbewegung. Die Übergabeparameter sind die Objekt-ID des Blocks, mit dem Max dem Partner hilft, und die berechnete Hilfsposition $\mathbf{p}_{blockNeu}$. Die Bewegungsqualität der Umplatzierungsbewegung kann zudem über den $degEmp$ moduliert werden. Da zwischen einer geringsten und höchsten Geschwindigkeit der Umplatzierungsbewegung gewählt werden kann, kann der Wert von $degEmp$, der zwischen 0 und 1 liegt, einfach als Geschwindigkeitsparameter übergeben werden. Je höher also Max' Empathie für den Menschen ist, desto höher ist sein Hilfsgrad. In diesem Fall stellt er den benötigten Block näher und schneller zum Partner. Je näher Max den Block zur maximalen

Hilfsposition setzt, desto mehr muss er sich vorlehnen und desto mehr Aufwand kostet ihn die Hilfsbewegung. Da er diese Bewegung nur bei sehr hohem Empathiegrad durchführen würde, könnte sie für einen menschlichen Partner als sozial plausibel wahrgenommen werden.

Berechnung des Hilfsgrades für den Partner

Entscheidet sich der menschliche Partner, Max zu helfen, wird für ihn ein Hilfsgrad berechnet. Für diesen werden dieselben Berechnungsgrundlagen, wie für den Hilfsgrad von Max verwendet. Es wird also angenommen, dass die Hilfsposition, die der Mensch bei einer Hilfsaktion für Max wählt, etwas über seinen Empathiegrad aussagt. Je näher der Mensch ein Objekt an Max heranstellt, je mehr dieser sich also vorlehnt, um Max zu helfen, desto höher schätzt Max den Empathiegrad des Menschen ein. Der Hilfsgrad bewegt sich genau wie für Max zwischen einer hypothetischen minimalen und maximalen Hilfsposition. Diese wird auf dieselbe Weise berechnet wie für Max, also auf Grundlage des simulierten peripersonalen Raums. Zur Berechnung werden folgende Parameter benötigt bzw. als Hypothese errechnet:

- Die *Hilfsposition* des Blocks im Ursprungskoordinatensystem des simulierten peripersonalen Raums des Partners.
- Die *Ursprungsposition* des Blocks, um eine Hypothese über die *minimale Hilfsposition* zu errechnen. Dies geschieht über den simulierten peripersonalen Raum des Partners.
- Eine Hypothese über die *maximale Hilfsposition*. Diese wird ebenfalls über den simulierten peripersonalen Raum des Partners errechnet.

Die Annahmen sind demnach genau die gleichen, wie in Abbildung 7.8 gezeigt, nur als Hypothesen über den Partner. Die Formel 7.1 kann mit dieser Parameterliste nach $degEmp$ aufgelöst werden und beinhaltet dann den Hilfsgrad. Geht der Wert für die tatsächliche Hilfsposition über die vermutete maximale Hilfsposition hinaus, wird nur der Anteil bis zur maximalen Hilfsposition einbezogen. Der Hilfsgrad wird demnach auf einen Wert zwischen 0 und 1 normiert. Hilft der Mensch Max über die maximale Position hinaus, wächst der errechnete Hilfsgrad somit nicht. Der errechnete Wert wird dem Empathiemodul als *Liking*-Wert des Partners für Max übergeben.

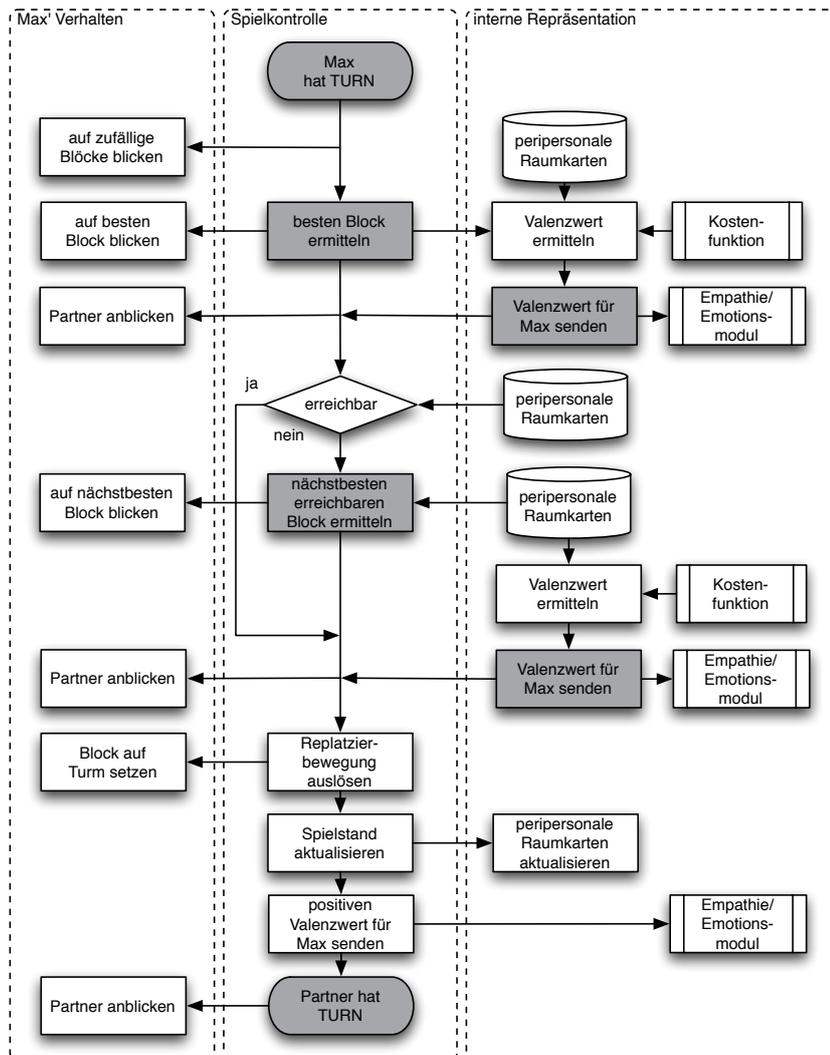


Abbildung 7.9: Max' Spielzug mit Valenzwerte-Berechnung.

7.3.4 Spielabläufe und interne Verarbeitung

Die empathiebasierten Spielzüge der Partner werden in den nächsten Abschnitten wieder in Flussdiagrammen schematisch dargestellt. Im weiteren Verlauf werden nur die Unterschiede zwischen den regulären und den empathiebasierten Spielzügen erläutert. Die Ablaufebene der internen Repräsentation wird dabei um die in den vorherigen Abschnitten erläuterten Komponenten erweitert. Die Kostenfunktion ist für beide Partner dieselbe. Das Empathiemodul und das Emotionssystem werden als *Black Box* behandelt und als eine gemeinsame Komponente im Flussdiagramm dargestellt. Die grau unterlegten Kästen zeigen auch hier wieder an, welche Informationen in Max' Denkblase angezeigt werden.

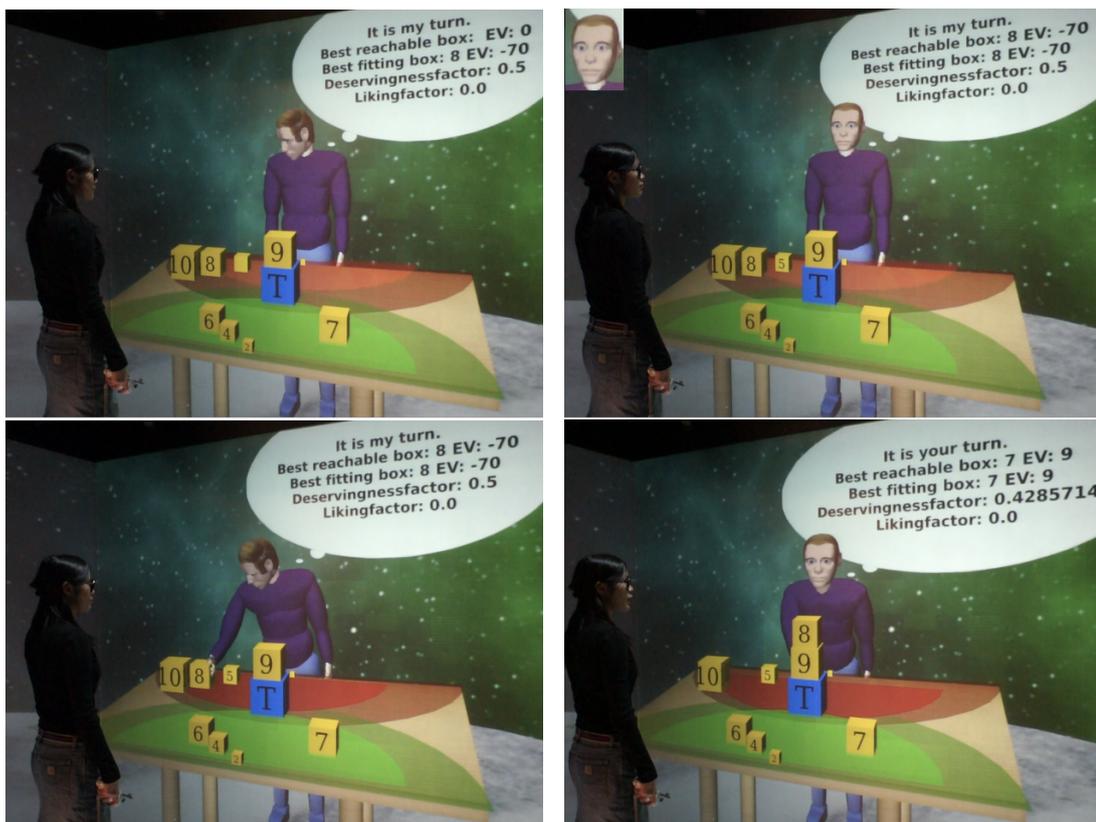


Abbildung 7.10: Max' Spielzug mit Valenzwerte-Berechnung in der virtuellen Realität.

Spielzug von Max

In Abbildung 7.9 wurde das Flussdiagramm für Max' Spielzug um die Berechnung der Valenzwerte erweitert. Der Unterschied zum regulären Spielzug ist, dass an maximal drei Stellen des Spielzuges ein Valenzwert für Max an das Emotionssystem gesendet wird. In dem Spielzug ermittelt die Spiellogik wieder den besten Block und Max richtet seinen Blick auf den Block. Diesmal wird zudem ein Valenzwert anhand der Kostenfunktion und der Distanz in Max' peripersonalem Raum berechnet (vgl. Abschnitt 7.3.2). Der Valenzwert wird für Max an das Emotionssystem gesendet und in Max' Denkblase neben der Blocknummer für den besten Block angezeigt. Anschließend schaut Max zum Partner. In Max' Gesicht kann der Partner möglicherweise einen Emotionsausdruck erkennen, der durch den Valenzwert verändert wurde. Ist der Block nicht erreichbar, schaut Max auf den nächstbesten und es wird wie zuvor, ein Valenzwert für den Block ermittelt. Max schaut den Partner wieder an und zeigt möglicherweise einen Emotionsausdruck. Hat Max einen Block auf den Turm gesetzt wird automatisch ein Goal-Success aus der Spiellogik gemeldet und ein positiver Valenzwert von $+40$ für Max an das Emotionssystem gesendet. Dieser Wert durch experimentelles Testen des Systems als geeigneter Parameter ermittelt.

Abbildung 7.10 zeigt einen Spielzug von Max in der virtuellen Realität. Er schaut auf Block 8, der in seinem Vorlehnraum liegt. Es wird ein Valenzwert von -70 über die Kostenfunktion ermittelt. Max schaut seinen Partner an und zeigt einen negativen Emotionsausdruck im Gesicht. Er greift das Objekt, setzt es auf den Turm und erhält einen positiven Valenzwert von $+40$.

Spielzug des Partners

In Abbildung 7.11 wurde das Flussdiagramm für den Spielzug des Partners um die Berechnung der Valenzwerte erweitert. Der Unterschied zum regulären Spielzug ist, dass an zwei Stellen des Spielzuges ein simulierter Valenzwert für den Partner an Max' Emotionssystem gesendet wird. Hier ist zu beachten, dass Max mit den Valenzwerten eine Hypothese über die Emotionen des Partners bilden kann. Max' Emotionen werden von den simulierten Valenzwerten für den Partner an dieser Stelle nicht beeinflusst. Sobald der Partner an der Reihe ist, wird für ihn wieder über die Spiellogik ein bester Block ermittelt. Für den Partner wird auf dieselbe Weise wie für Max ein Valenzwert über die Kostenfunktion und über die hypothetische Distanz in dessen

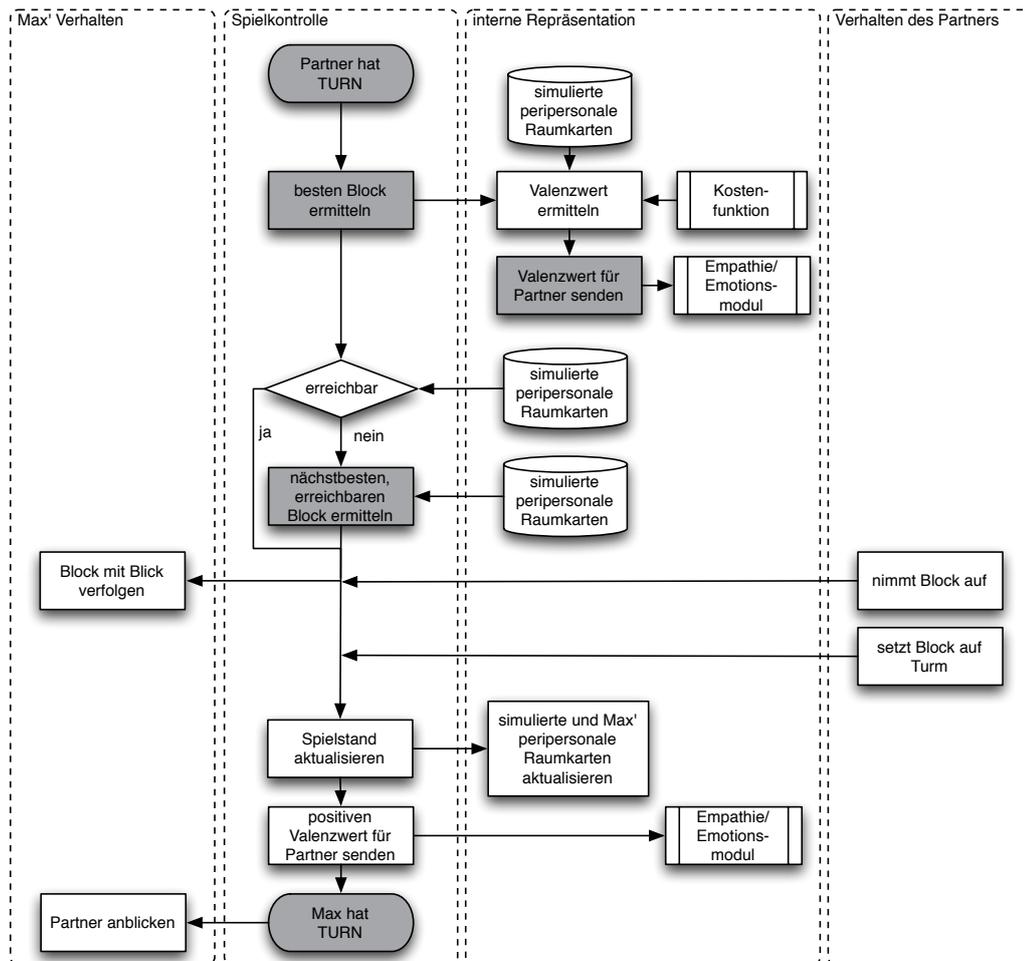


Abbildung 7.11: Spielzug des Partners mit simulierten Valenzwerten.

simuliertem peripersonalen Raum berechnet. Der Valenzwert wird dann als Hypothese an das Emotionssystem gesendet. Es ist zu beachten, dass für den Partner kein Valenzwert für den nächstbesten Block errechnet und simuliert wird. Diese Hypothese entspräche in Max' Emotionssystem einer Erwartung darüber, welchen Block der Partner als nächstes spielen und welche Emotion er dabei fühlen wird. In dem hier vorgestellten Interaktionsszenario werden zum jetzigen Zeitpunkt keine Erwartungshypothesen gebildet. Sobald der Partner einen Block auf den Turm gesetzt hat wird für ihn ein Goal-Success aus der Spiellogik gemeldet und ein positiver Valenzwert von $+40$ für den Partner an Max' Emotionssystem gesendet.

Abbildung 7.3 zeigt einen Beispiel-Spielzug einer menschlichen Partnerin in der virtuellen Realität. Sie greift Block 9 und setzt ihn auf den Turm. In Max' Denkblase sind ihr bester Block und eine Hypothese über den Valenzwert zu sehen. Sobald der Block auf den Turm gesetzt wird, wird ein positiver Valenzwert von $+40$ für den Partner gesendet.

Der Partner hilft Max

Abbildung 7.12 zeigt das Flussdiagramm für den Spielzug von Max, wenn der Partner ihm hilft. Dieser Spielzug ist mit der empathiebasierten Hilfsstrategie neu hinzugekommen. Max' Spielzug verläuft zu Anfang wie bereits erläutert. Zur besseren Übersicht wurde im Flussdiagramm nur der Fall schematisch aufgeführt, wenn Max den besten Block nicht erreichen kann. Der menschliche Partner hat dann die Möglichkeit, Max zu helfen. Sobald der Partner mit der Wii ein Objekt anhebt, wird Max' Spielzug pausiert. Max wartet so lange, bis der Partner das Objekt wieder absetzt. Setzt der Partner das Objekt in den Interaktionsraum, werden alle peripersonalen Karten aktualisiert und ein Hilfsobjekt in der Spiellogik registriert. Für diesen wird mit der beschriebenen Methode der Hilfsgrad für den Partner berechnet, dieser legt den Liking-Faktor fest und wird an das Empathiemodul gesendet. Danach wird für Max ein Valenzwert über die Kostenfunktion berechnet und an sein Emotionssystem gesendet. Max blickt den Partner an und zeigt eine mögliche Änderung seines Emotionsausdrucks an. Anschließend greift er das Hilfsobjekt, setzt es auf den Turm und der dritte Valenzwert von $+40$ wird für das Goal-Success an das Emotionssystem gesendet.

Abbildung 7.13 zeigt einen Beispiel-Spielzug eines menschlichen Partners in der virtuellen Realität. Der menschliche Partner ist Max bei seinem Spielzug behilflich.

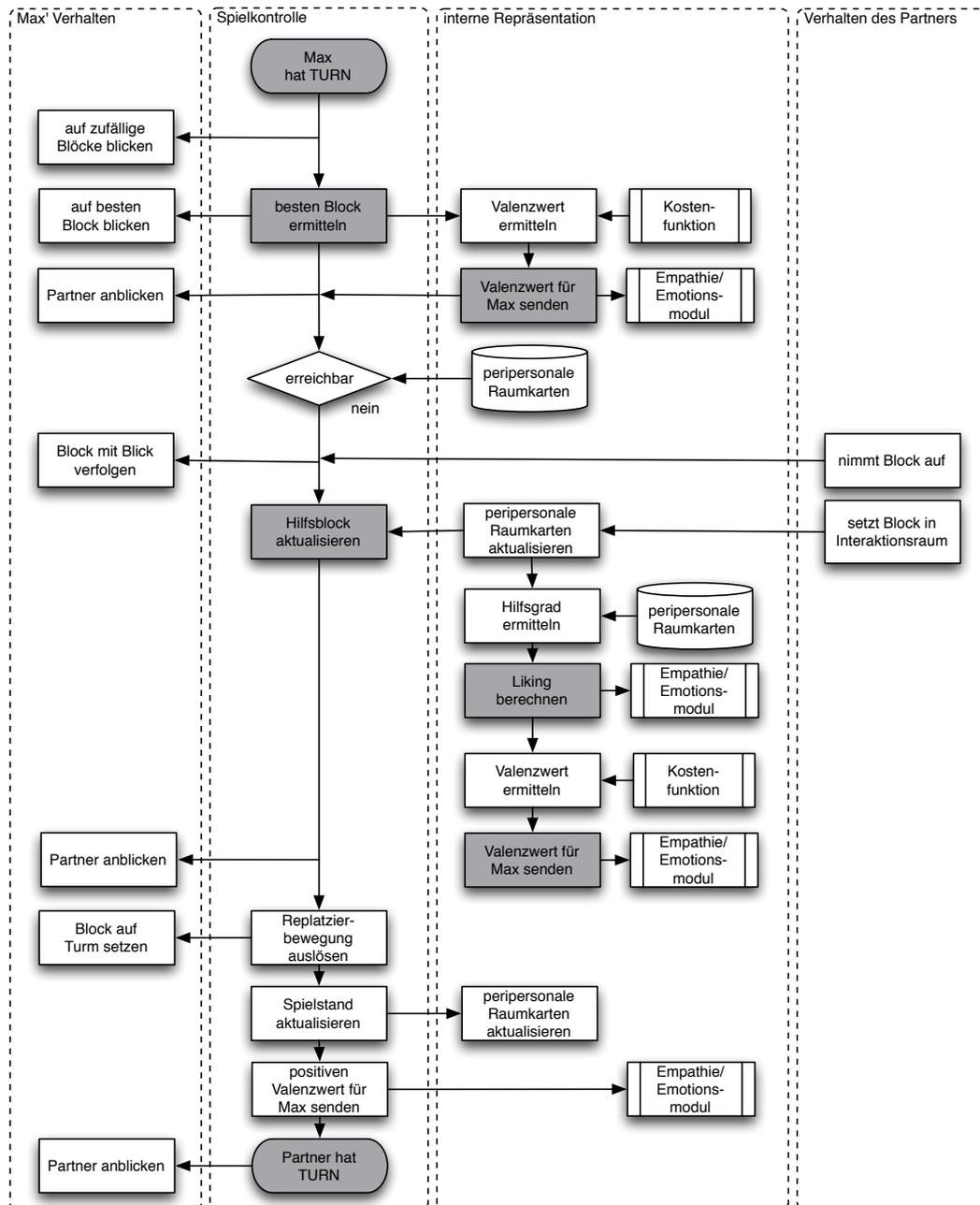


Abbildung 7.12: Spielzug von Max mit Berechnung der Parameter für die empathiebasierte Hilfsstrategie, wenn der Partner Max hilft.

7.3 Empathie als Basis für räumlich-kooperatives Verhalten



Abbildung 7.13: Der menschliche Partner hilft Max in der virtuellen Realität.

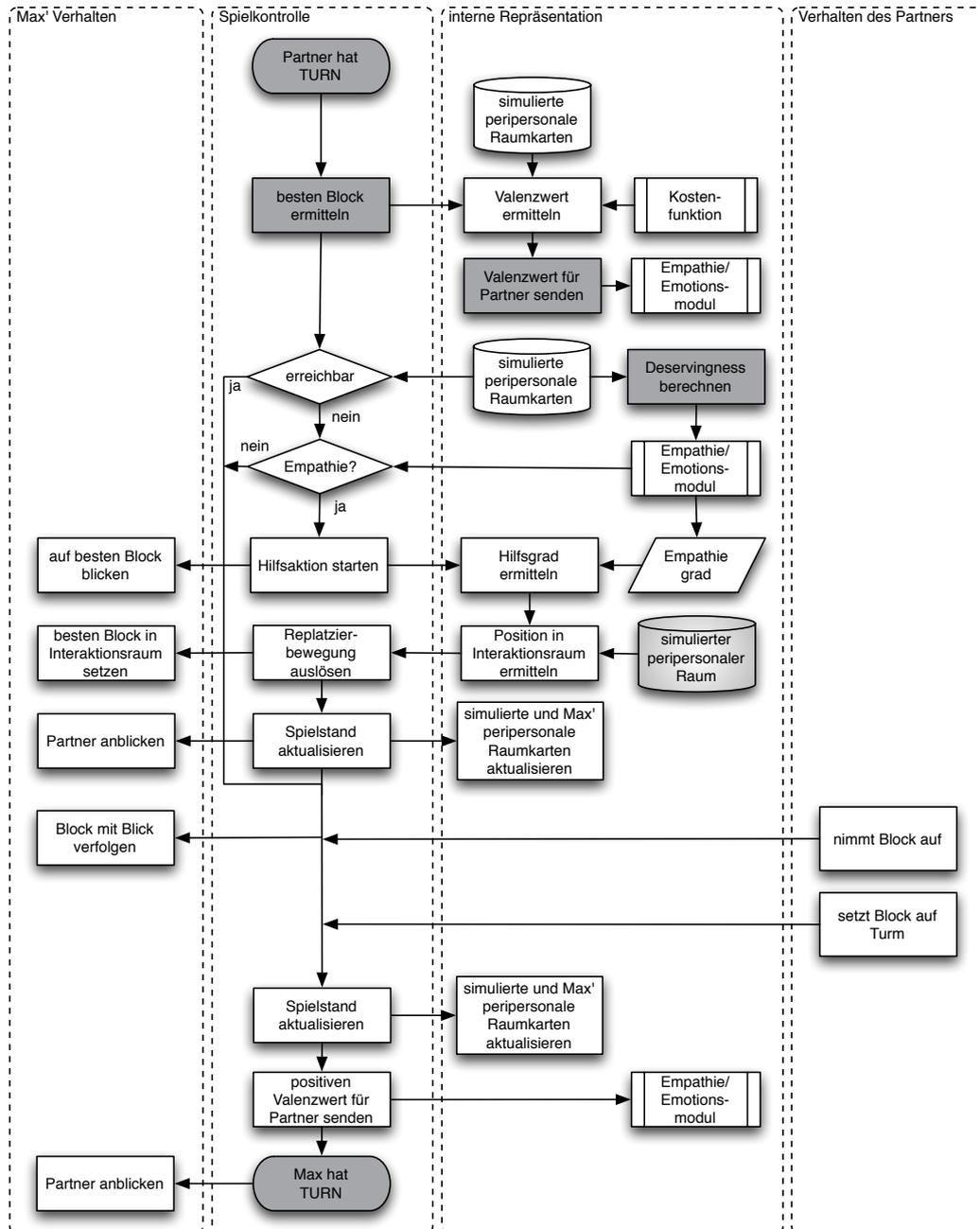


Abbildung 7.14: Max' hilft dem Partner durch empathiebasiertes räumlich-kooperatives Verhalten.

Max' bester Block ist der mit der Nummer 6. Der Block ist für ihn nicht erreichbar und wird über die Kostenfunktion mit einem Valenzwert von -100 bewertet. Max blickt die Partnerin an und zeigt einen negativen Emotionsausdruck im Gesicht. Die Partnerin setzt Objekt 6 in den Interaktionsraum. Da das Hilfsobjekt nun in Max' Vorlehnraum liegt, wird für Max ein Valenzwert von -30 an das Emotionssystem gesendet. Nachdem Max den Block auf den Turm gesetzt hat wird für ihn der Wert von $+40$ gesendet. Er blickt die Partnerin an und zeigt einen positiven Gesichtsausdruck.

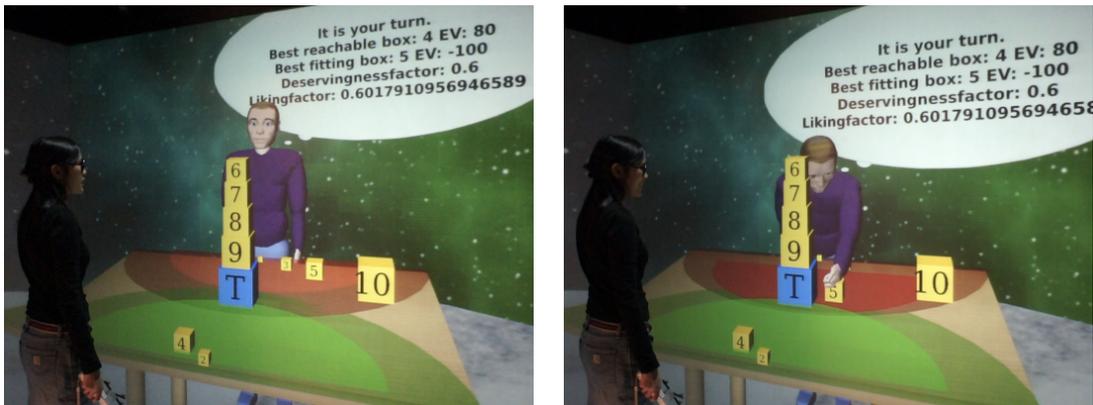


Abbildung 7.15: Max hilft dem menschlichen Partner in der virtuellen Realität.

Max hilft dem Partner

Abbildung 7.14 zeigt das Flussdiagramm für den Spielzug des Partners, wenn Max ihm hilft. In diesem Spielzug zeigt Max empathiebasiertes räumlich-kooperatives Verhalten. Der Unterschied zu Flussdiagramm 7.11 ist der Ablauf, wenn der Partner den besten Block nicht erreichen kann. Dann wird ermittelt, ob Max Empathie mit dem Partner hat. Hat Max Empathie, wird Max' Hilfsaktion gestartet. Auf Grundlage des Deservingness- und Liking-Faktors wird im Empathiemodul der aktuelle Empathiegrad ermittelt. Dieser wird verwendet, um den Hilfsgrad, die Hilfsposition und die Geschwindigkeit, mit der geholfen wird, zu berechnen. Mit der Hilfsposition führt Max eine Umplatzierungsbewegung mit dem besten Block aus und setzt ihn in den Interaktionsraum. Der Partner kann wählen, ob er die Hilfe entgegen nimmt und den Hilfsblock oder einen anderen Block auf den Turm setzt.

Abbildung 7.15 zeigt einen Beispiel-Spielzug von Max in der virtuellen Realität. Die Partnerin ist an der Reihe. Der beste Block ist Block 5. In Max' Denkblase sind

ihr bester Block und eine Hypothese über den Valenzwert von -100 zu sehen. Max ist empathisch mit der Partnerin und hilft ihr bei ihrem Spielzug. Der Liking- und Deservingness-Faktor modulieren die Geschwindigkeit, mit der seine Hilfsbewegung ausgeführt wird, und die Hilfsposition, mit der er ihr hilft. Da Max in dieser Spielphase mit seiner Spielpartnerin empathisch ist, kann es sein, dass er während der Hilfsaktion einen negativen Emotionsausdruck im Gesicht zeigt. Dieses Verhalten spiegelt die Empathie-Emotion wider, die Max seiner Partnerin zuschreibt. Er fühlt sozusagen mit seiner Spielpartnerin mit.

7.4 Zusammenfassung und Diskussion

In diesem Kapitel wurde ausgehend von der in Abschnitt 4.1.3 gebildeten Definition 6 ein Modell des Interaktionsraums zwischen einem virtuellen Menschen und seinem Interaktionspartner eingeführt. Das Modell basiert zum einen auf dem neurologischen Mechanismus der verkörperten Simulation und zum anderen auf dem kognitiven Mechanismus der räumlichen Perspektivenübernahme. Diese Mechanismen werden bei Menschen in interpersonaler Interaktion vermutet und tragen mutmaßlich dazu bei, dass wir unsere Partner und ihre Aktionen verstehen können. Das Interaktionsraummodell für humanoide Agenten bildet zusammen mit der Definition des räumlich-kooperativen Verhaltens eine Lösung für das Teilproblem des intelligenten räumlichen Verhaltens im Interaktionsraum (P3).

Das Modell wurde in einer Nahraumaufgabe in der virtuellen Realität mit dem virtuellen Menschen Max und einem menschlichen Partner erfolgreich eingesetzt. Die Nahraumaufgabe besteht aus einem Spiel, in dem Max und sein Partner in der virtuellen Realität in Kooperation einen Turm aus Blöcken zusammenbauen. Durch das Interaktionsraummodell bildet Max eine Hypothese der räumlichen Perspektive des Partners. Auf dieser Basis kann Max räumlich-kooperatives Verhalten gegenüber seinem Partner zeigen. Das Verhalten wird in unserem Szenario durch eine Hilfsaktion ausgedrückt, mit der Max ein für den Partner benötigtes Objekt in den Raum stellt, der für beide erreichbar ist, das bedeutet, in den Interaktionsraum.

In einer Erweiterung dieses Szenarios wurde die kooperative Konstruktionsaufgabe mit einem Empathiemodell verknüpft. In diesem Szenario wird Max' räumlich-kooperatives Verhalten durch den Empathiegrad moduliert. Max' Hilfsaktionen kann-

ten mit dem empathiebasierten räumlich-kooperativen Verhalten über soziale Faktoren beeinflusst werden. Abhängig davon wie sehr der menschliche Partner Max in einem Spielzug geholfen hat, wird Max' Empathie mit dem Partner beeinflusst. Dieses resultiert in einem Empathiegrad, der wiederum Max' Hilfsaktionen moduliert. Durch diese wechselseitige Beeinflussung bestimmt das Kooperationsverhalten der einzelnen Partner den Ausgang der gelösten Aufgabe. Je kooperativer die Partner sind, desto höher kann der Turm gestapelt werden.

In Abschnitt 3.4 wurden technische Arbeiten zu humanoiden Agenten in Nahraumaufgaben vorgestellt und anhand von Kriterien beleuchtet, die für Nahraumaufgaben wünschenswert sind. Diese sollen nun an das hier vorgestellte Szenario angelegt werden. Die Kriterien waren:

- **Zugänglichkeit:** Ist der interpersonale Raum beiden Partnern zugänglich?
- **Hilfe:** Unterstützt der künstliche Partner den Menschen und wenn ja, in welcher Modalität geschieht das? Verbal oder durch eine eigene (räumliche) Aktion?
- **Partnermodell:** Welches (räumliche) Wissen hat der künstliche Partner von seinem menschlichen Partner? Wie wird es aufgebaut und wofür nutzt der künstliche Partner es?
- **Kooperation:** Welche Rolle hat der humanoide Agent? Fungiert dieser als gleichberechtigter Partner oder nimmt er die Funktion eines anleitenden Tutors ein?

In Abschnitt 7.1 wurde unsere Interaktionsraum-Definition mit Aspekten aus Kendons F-Formationen-System erweitert. Der Interaktionsraum ist der Raum, in dem alle Interaktanten kooperieren, um diesen aufrechtzuerhalten und zu dem alle *direkten*, *exklusiven* und *gleichen Zugang* haben. Die räumliche *Zugänglichkeit* zum gemeinsamen Raum ist also ein zentraler Aspekt des in dieser Arbeit entworfenen Interaktionsraums. Da er über die peripersonalen Räume der Partner gebildet wird, ist seine wesentliche Eigenschaft, dass alle Partner ihn mit ihren Armen erreichen können. Der Aspekt der *Hilfe* ist in unserem Szenario durch die Definition räumlich-kooperativen Verhaltens und der von Max ausgeführten Hilfsaktion gegeben. In dem mit Empathie erweiterten Szenario beeinflusst die Hilfe des menschlichen Interaktionspartners, das Verhalten von Max. Das *Partnermodell* wird im Interaktionsraumszenario über die räumliche Perspektivenübernahme gebildet, um eine Hypothese über die Spielzüge des menschlichen

Partners zu generieren. Das räumliche Partnermodell ist eine Basis für Max' Hilfsaktionen. Der Aspekt der *Kooperation* drückt sich direkt im räumlich-kooperativen Verhalten von Max aus. Im Spielszenario fungieren beide Spieler als gleichberechtigte Partner, da keiner von beiden über mehr Wissen oder weitere Vorteile verfügt.

Das vorgestellte Interaktionsraummodell konnte somit allen gewählten Kriterien entsprechen. Damit bildet es eine geeignete Grundlage, humanoiden Agenten die Fähigkeit zu geben, auf natürliche und intelligente Weise Aufgaben im Nahraum mit menschlichen Partnern zu lösen. Ob dieses Verhalten tatsächlich von menschlichen Partnern angenommen wird, muss in Nutzerstudien überprüft werden. Eine weitere Frage, die sich stellt ist, wie übertragbar das Modell auf andere humanoide Agenten ist. Dies sind offene Fragen, die in Kapitel 8 im Ausblick wieder aufgegriffen werden.

8 Resümee

Das gesetzte Ziel war es, humanoiden Agenten die Fähigkeit zu geben auf natürliche und intelligente Weise, Aufgaben im Nahraum mit menschlichen Partnern zu lösen. Hierfür wurden drei Teilprobleme formuliert, deren Lösungen in den vorangegangenen Kapiteln entwickelt und vorgestellt wurden. Für die Lösungen wurde ein interdisziplinärer Ansatz verfolgt, bei dem neben der Aufarbeitung existierender technischer Lösungen auch biologische Befunde beim Menschen Hinweise zur technischen Modellierung geben sollen. Im Folgenden werden zunächst die wichtigsten Erkenntnisse aus den einzelnen Kapiteln rekapituliert. Im Anschluss daran werden die Beiträge der Lösungsansätze zum Forschungsstand präsentiert und die vorliegende Arbeit in ihren Forschungskontext eingeordnet. Zum Abschluss wird ein Ausblick auf Aspekte gegeben, die noch offen geblieben sind oder durch diese Arbeit neu aufgeworfen wurden.

In Kapitel 2 wurden theoretische Grundlagen des individuellen Körperraums und der interpersonalen Räume bei Menschen vorgestellt. Es wurden Raummodelle für nahe und ferne Distanzen um den menschlichen Körper vorgestellt, die Aufschluss darüber geben, wie der Räume um Menschen herum strukturiert und mit welchen Eigenschaften und Funktionen sie verknüpft sein könnten. Es zeigte sich, dass die biologischen Erklärungen zur menschlichen Körper- und Raumrepräsentation noch nicht abschließend geklärt und Gegenstand von Diskussionen sind. Für den peripersonalen Raum, das Körperschema und für den Zusammenhang zwischen beiden wurden daher auf Grundlage der biologischen Befunde Arbeitsdefinitionen vorgeschlagen. Diese bildeten den theoretischen Ausgangspunkt für technische Modellierungsentscheidungen. Für die interpersonale Raumrepräsentation beim Menschen wurden die verkörperte Simulation und die räumliche Perspektivenübernahme als Mechanismen vorgestellt. Die vorgestellten sozialen Raummodelle für den interpersonalen Raum waren das Proxemikmodell von Hall und das F-Formationen-Modell von Kendon.

In Kapitel 3 wurde der Stand der Forschung zu drei Bereichen aufgearbeitet. Bei den technischen Grundlagen der Bewegungsplanung- und kontrolle humanoider Agen-

ten wurde nach dem Vergleich verschiedener Ansätze die ACE-Bibliothek als geeigneter Animationsansatz vorgeschlagen und dazu passend, die Methode der künstlichen Potentialfelder zur Hindernisvermeidung gewählt. Für die Planung von Umplatzierbewegungen wurden Modelle zur Modulierung der Bewegungsqualität vorgestellt, die soziale Signale vermitteln können. Aus dem zweiten Forschungsbereich wurde nach einer Voranalyse das Konzept des peripersonalen Raums als geeignete Basis zur Lösung des Teilproblems der Wahrnehmung und Repräsentation des Nahraums (P1) identifiziert. Aus den existierenden technischen Arbeiten zum peripersonalen Raum wurden folgende Aspekte als relevante Eigenschaften betrachtet: die Koordinatensystemtransformation, der Einsatz von taktilen Sensordaten und die Verknüpfung von Teilräumen mit unterschiedlichen Berechnungsprozessen. Für die Lösung des Teilproblems der Repräsentation des Körpers (P2) wurde das Konzept des Körperschemas als geeignete Modellierungsgrundlage gewählt, da es gleichzeitig Anforderungen an die peripersonale Raumrepräsentation erfüllt. Nach der Diskussion existierender Arbeiten zur Generierung eines künstlichen Körperschemas wurde der Ansatz von Hersch vorgeschlagen. Im Anschluss daran wurden Arbeiten vorgestellt, die Hinweise zur Lösung des Teilproblems des intelligenten räumlichen Verhaltens im Interaktionsraum (P3) bieten. Arbeiten zur Proxemik- und Formationenkontrolle gaben Einblicke in soziale Rauminteraktion zwischen menschlichen und humanoiden Partnern, die auch in Nahrauminteraktion relevant sind. Nahraumszenarien mit humanoiden Agenten und menschlichen Partnern wurden vorgestellt und Diskussionskriterien für diese vorgeschlagen.

Der Gesamtlösungsansatz für die drei aufgeworfenen Teilprobleme wurde in Kapitel 4 entwickelt. Auf Grundlage der in Kapitel 2 vorgeschlagenen Arbeitsdefinitionen und den in Kapitel 3 getroffenen Modellierungsentscheidungen wurden technische Interpretationen der biologischen Konzepte entwickelt. Zudem wurde für den interpersonellen Raum das Konzept des Interaktionsraums gebildet und darin gezeigtes räumlich-kooperatives Verhalten definiert. Im Anschluss daran wurden die für die Umsetzung des Gesamtlösungsansatzes relevanten Komponenten des virtuellen Menschen Max vorgestellt.

In Kapitel 5 wurde eine Lösung des Teilproblems P2 dargestellt. Zunächst wurden die Grenzen der visuell-propriozeptiven Rekalibrierung diskutiert und mit den Anforderungen an ein künstliches Körperschema verglichen. Als Konsequenz der Grenzen wurde ein taktiles Lernprozess vorgeschlagen, der keine visuellen Sensordaten benö-

tigt. In einer Evaluation wurden Simulationen des taktilen und des visuellen Lernprozesses miteinander verglichen. Es konnte gezeigt werden, dass der takile Lernprozess im Durchschnitt zwar eine höhere Anzahl an Iterationen benötigt bis das künstliche Körperschema gelernt wurde, die Abweichung vom tatsächlichen Körperschema aber geringer ist. Dieses Ergebnis hebt nicht nur die Bedeutung der taktilen Modalität bei humanoiden Agenten hervor, sondern gibt auch Hinweise für deren technische Konzipierung.

In Kapitel 6 wurde die Lösung des Teilproblems P1 vorgestellt. Für die Lösung wurde ein peripersonales Raummodell für humanoide Agenten entwickelt. Der Entwurf des peripersonalen Raummodells stützt sich auf kognitive Raummodelle für Menschen und Befunde aus der Neuropsychologie. Er wird über verschiedene Sensormodalitäten abgedeckt und bietet für autonome humanoide Agenten eine robustere und umfassendere Raumrepräsentation. Zudem konnte gezeigt werden, dass die Unterräume eine Strukturierung der Aufmerksamkeit eines humanoiden Agenten und eine bessere Planung der Armbewegungen ermöglichen.

In Kapitel 7 wurde eine Lösung des Teilproblems P3 vorgeschlagen. Ein Modell des Interaktionsraums zwischen einem virtuellen Menschen und seinem Interaktionspartner wurde eingeführt. Das Modell basiert auf dem Mechanismus der verkörperten Simulation und dem Mechanismus der räumlichen Perspektivenübernahme. Diese tragen mutmaßlich dazu bei, dass Menschen ihre Partner und ihre Aktionen verstehen können. Das Modell konnte in einer Nahraumaufgabe in der virtuellen Realität mit dem virtuellen Menschen Max und einem menschlichen Partner erfolgreich eingesetzt werden. Max kann durch das Modell seinem Partner bei Spielzügen helfen, die er allein nicht ausführen kann. In einer Erweiterung dieses Szenarios wurde die kooperative Konstruktionsaufgabe mit einem Empathiemodell verknüpft. Dadurch kann Max' räumlich-kooperatives Verhalten über soziale Faktoren beeinflusst werden, die möglicherweise die Interaktion für menschliche Partner intuitiver gestaltet.

8.1 Beiträge zum Stand der Forschung

Um die einzelnen Beiträge zum Stand der Forschung identifizieren zu können, werden die Ergebnisse den aufgeworfenen Teilproblemen zugeordnet.

P1 – Wahrnehmung und Repräsentation des Nahraums:

- Es wurde ein peripersonales Raummodell für humanoide Agenten entwickelt, das in Unterräume unterteilt ist. Das Modell basiert auf kognitiven Raummodellen für Menschen und Befunden aus der Neuropsychologie. Die erfolgreiche Umsetzung theoretischer Konzepte kann als Beitrag zum besseren Verständnis menschlicher Kognition gesehen werden.
- Das peripersonale Raummodell ermöglicht einem humanoiden Agenten eine robuste und umfassende Wahrnehmung und Repräsentation des peripersonalen Raums durch die Raumabdeckung über verschiedene Sensormodalitäten. Dies ist ein Beitrag zur Erhöhung der Autonomie humanoider Agenten, um sich trotz limitierter sensorischer Ausstattung in der Umgebung zurechtzufinden.
- Die Unterräume des peripersonalen Raummodells ermöglichen die Strukturierung der Aufmerksamkeit eines humanoiden Agenten. Dies stellt einen Beitrag zur sinnvollen Filterung von Umgebungsinformationen und zu einer verbesserten Aktionsplanung dar.

P2 – Repräsentation des Körpers:

- Es wurde ein taktiler Lernprozess für das von Hersch u. a. (2008) entwickelte Körperschema-Rekalibrierungsverfahren vorgeschlagen und umgesetzt. Dieser Lernprozess benötigt keine visuellen Sensordaten. Der Aspekt stellt einen Beitrag zum Lernen der Körperrepräsentation mit alternativen Modalitäten dar.
- Der taktile Lernprozess konnte gegenüber dem visuellen Lernprozess akkuratere Ergebnisse für die Gelenkpositionen des Körperschemas ermitteln. Dies ist ein Beitrag zur Verbesserung des Lernens expliziter Körperrepräsentationen.
- Mit dem Einsatz der taktilen Sensormodalität wurde eine zentrale Eigenschaft des biologischen Körperschemas technisch umgesetzt. Die Ergebnisse der Umsetzung heben die Bedeutung der taktilen Modalität nicht nur für Menschen, sondern auch für humanoide Agenten hervor. Es konnten dadurch Erkenntnisse zur Rolle und Funktion taktiler Sensorinformationen im Aufbau von Körperrepräsentation gewonnen werden.

P3 – Intelligentes räumliches Verhalten:

- Es wurde das Konzept des Interaktionsraums entwickelt. Es handelt sich um ein interpersonales Raummodell für humanoide Agenten, das auf kognitiven und neurologischen Befunden bei Menschen beruht und durch die peripersonalen Räume gebildet wird. Das Modell bildet einen Beitrag zur Beschreibung und zum Verständnis interpersonaler Räume, die bislang wenig erforscht sind.
- Der Interaktionsraum ermöglicht es, Aktionen im interpersonalen Raum zu strukturieren. Da der Interaktionsraum ein Unterraum des peripersonalen Raums ist, stellt der Interaktionsraum ebenfalls einen Beitrag zur sinnvollen Filterung von Umgebungsinformationen und zur verbesserten Bewegungsplanung dar.
- Durch den Einsatz des Interaktionsraums bildet ein humanoider Agent eine Hypothese über die räumliche Perspektive des Partners. Dieses Wissen kann er nutzen, um räumlich-kooperatives Verhalten zu zeigen, um zum Beispiel dem Partner in einer Nahraumaufgabe zu helfen. Dies ist ein Beitrag zur Verbesserung der Mensch-Maschine-Interaktion.
- Das räumlich-kooperative Verhalten konnte durch eine empathiebasierte Hilfsstrategie sozialer gestaltet werden. Auch dieser Aspekt stellt einen Beitrag zur Verbesserung der Mensch-Maschine-Interaktion dar.
- Der peripersonale Raum und der Interaktionsraum wurden erfolgreich in Nahrauminteraktion in der virtuellen Realität eingesetzt. Damit wurde ein Beitrag geleistet zur Modellierung von Interaktionen mit humanoiden Agenten auf nahe Distanzen.

Einordnung in den Forschungskontext

In Abschnitt 1.3 der Einleitung wurde der Forschungskontext dieser Arbeit eingeführt. Es folgt daher eine Einordnung in die Ziele des Sonderforschungsbereichs 673. Der Sonderforschungsbereich verfolgt als übergeordnetes Ziel, eine einheitliche Theorie der Kommunikation als Alternative zu bestehenden Theorien zu entwickeln. Ausgehend von der Annahme von Pickering und Garrod (2004), dass *interaktive Ausrichtung* nicht nur durch die Angleichung von linguistischen Repräsentationen befördert wird,

sondern auch durch nichtlinguistische Repräsentationen, ist ein weiteres Forschungsziel die Übertragung des Ansatzes der interaktiven Ausrichtung auf die multimodale Kommunikation und die Mensch-Maschine-Interaktion.

Analog zur Angleichung von mentalen linguistischen Repräsentationen bei Dialogpartnern ist unser Ausgangspunkt, die räumliche Repräsentation des humanoiden Agenten automatisch und ohne explizite Aushandlung an den Partner anzugleichen. Einen Hintergrund für diesen Ansatz bilden die psycholinguistischen Arbeiten von Watson, Pickering und Branigan (Watson, Branigan und Pickering, 2004; Watson, Pickering und Branigan, 2009). Mit Bezug auf die Theorie der interaktiven Ausrichtung von Pickering und Garrod untersuchten Watson u. a. die Angleichung von räumlichen, also nichtlinguistischen Repräsentationen bei Dialogpartnern. Der Begriff der *Perspektivenangleichung* (engl. Perspective Alignment) bezeichnet dort das Phänomen der Produktion von sprachlichen Äußerungen und Beschreibungen aus der räumlichen Perspektive des Gesprächspartners (zum Beispiel „Die Tasse rechts von dir.“). Die Autoren nehmen an, dass dieses Phänomen den kognitiven Aufwand des Dialogpartners minimiert.

Die Angleichung räumlicher Repräsentationen wird zum jetzigen Zeitpunkt nicht zur Produktion sprachlicher Äußerungen verwendet, sondern soll es dem humanoiden Agenten ermöglichen, zum einen die räumliche Perspektive und zum anderen die räumliche Greifreichweite des Partners zu verstehen, um die eigenen Aktionen erfolgreich an die Aktionen des Partners anzupassen. Wir schlagen bezüglich des Konzeptes der interaktiven Ausrichtung somit ein Modell zur räumlichen Aktionsangleichung eines virtuellen Menschen an seinen Interaktionspartner vor. Der erfolgreiche Einsatz dieses Modells in einem prototypischen Anwendungsszenario konnte somit die Wirksamkeit der interaktiven Ausrichtung für ein technisches System zeigen.

8.2 Ausblick

Eine Erweiterung für das Interaktionsraummodell stellt die Simulation des visuellen Aufmerksamkeitsraums für den menschlichen Partner dar. Im bisherigen Szenario wird eine Hypothese über den Vorlehnraum und den Takttilraum des Partners gebildet. Dadurch wird zwar simuliert, ob der Partner Objekte erreichen kann oder nicht, es wird jedoch nicht berücksichtigt, ob der Partner das Objekt bereits visuell wahrgenommen

hat. Eine weitere Ergänzung zur räumlichen Perspektivenübernahme ist die Bildung von Hypothesen darüber, ob Gegenstände aus der Perspektive des Partners verdeckt sind. Der humanoide Agent könnte das Objekt in einen Bereich stellen, in dem es für das Gegenüber sichtbar ist.

Eine wesentliche Einschränkung der aktuellen räumlichen Interaktion ist die Abwesenheit von verbalen Äußerungen. Dieser für die Kommunikation wichtige Faktor könnte genutzt werden, um statt Motoraktionen verbale Hinweise zu generieren, die der humanoide Agent dem Partner mitteilen könnte. Ein gerade für mobile humanoide Agenten weiterer sinnvoller Aspekt ist die Modellierung von Formations- und Proxemikverhalten in Kombination mit dem Interaktionsraum. Gerade die Körperpostur und -orientierung sind wichtige kommunikative Signale in menschlicher Interaktion, die ein humanoider Agent beherrschen sollte.

Damit wurde eine geeignete Grundlage für die Fähigkeit humanoider Agenten geschaffen, auf natürliche und intelligente Weise Aufgaben im Nahraum mit menschlichen Partnern zu bewältigen. Die Erweiterung durch eine empathiebasierte Hilfsstrategie scheint dabei ein vielversprechender Ansatz zu sein, menschlichen Partnern intuitiven Umgang mit humanoiden Agenten zu ermöglichen.

Um jedoch abschließende Erkenntnisse darüber zu gewinnen, ob die modellierten Verhaltensweisen zu besserer Mensch-Maschine-Interaktion führen, müssten Nutzerstudien durchgeführt werden. In diesen könnten alternative Hilfsstrategien, wie sie in Abschnitt 7.3 aufgeführt wurden, evaluiert und verglichen werden.

Eine Frage, die sich für alle in dieser Arbeit vorgestellten Lösungen stellt ist, wie übertragbar sie auf andere humanoide Agenten ist. Physische Roboterplattformen die von den Lösungen profitieren könnten wären z. B. der iCub-Roboter, der einen menschenähnlichen Torso besitzt. Angesichts der zunehmenden Erforschung von räumlichen Interaktionsfähigkeiten humanoider Agenten, kann die Warnung von Brooks und Arkin im Eingangszitat des Einleitungskapitels entkräftet werden. Räumlich-kooperatives Verhalten auf nahe Distanzen ist eine zu berücksichtigende Anforderung für humanoide Agenten, um, wie in der Einleitung skizziert, der Entwicklung von Heimrobotern und virtuellen 3-D-Charakteren entgegenzutreten zu können.

Literaturverzeichnis

- [Antonelli u. a. 2011] ANTONELLI, M. ; CHINELLATO, E. ; POBIL, A. P.: Implicit Mapping of the peripersonal space of a humanoid robot. In: Proceedings of the IEEE SSCI2011 - Symposium Series on Computational Intelligence, Workshop on Brain Computer Interfaces, Neural Modeling and Robotics. Paris, 2011, S. 1–8
- [Becker-Asano 2008] BECKER-ASANO, C.: WASABI: Affect Simulation for Agents with Believable Interactivity, Faculty of Technology, University of Bielefeld, Diss., 2008. – IOS Press (DISKI 319)
- [Behnisch u. a. 2010] BEHNISCH, M. ; HASCHKE, R. ; GIENGER, M.: Task space motion planning using reactive control. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)., 2010, S. 5934–5940
- [Berenson u. a. 2009] BERENSON, D. ; SRINIVASA, S. S. ; FERGUSON, D. ; COLLET, A. ; KUFFNER, J. J.: Manipulation planning with Workspace Goal Regions. In: Proceedings of the IEEE International Conference on Robotics and Automation i (2009), S. 618–624
- [Boukricha u. a. 2011] BOUKRICH, H. ; NGUYEN, N. ; WACHSMUTH, I.: Sharing Emotions and Space – Empathy as a Basis for Cooperative Spatial Interaction. In: VILHJÁLMSSON, H. (Hrsg.) ; KOPP, S. (Hrsg.) ; MARSELLA, S. (Hrsg.) ; THÓRIS-SON, K. (Hrsg.): Intelligent Virtual Agents Bd. 6895. Springer Berlin Heidelberg, 2011, S. 350–362
- [Boukricha und Wachsmuth 2011] BOUKRICH, H. ; WACHSMUTH, I.: Empathy-Based Emotional Alignment for a Virtual Human: A Three-Step Approach. In: KI – Künstliche Intelligenz 25 (2011), Nr. 3, S. 195–204
- [Bratman 1992] BRATMAN, M. E.: Shared Cooperative Activity. In: The Philosophical Review 101 (1992), Nr. 2, S. 327–341

- [Breazeal u. a. 2005] BREAZEAL, C. ; BUCHSBAUM, D. ; GRAY, J. ; GATENBY, D. ; BLUMBERG, B.: Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots. In: Artif. Life 11 (2005), Januar, Nr. 1-2, S. 31–62
- [Breazeal und Fitzpatrick 2000] BREAZEAL, C. ; FITZPATRICK, P.: That Certain Look: Social Amplification of Animate Vision. In: In Proceedings AAAI Fall Symposium, Socially Intelligent Agents - The Human in the Loop, 2000, S. 3–5
- [Bremner u. a. 2008] BREMNER, A. J. ; MARESCHAL, D. ; LLOYD-FOX, S. ; SPENCE, C.: Spatial Localization of Touch in the First Year of Life: Early Influence of a Visual Spatial Code and the Development of Remapping Across Changes in Limb Position. In: Journal of Experimental Psychology: General 137 (2008), Nr. 1, S. 149 – 162
- [Brooks und Arkin 2007] BROOKS, A. G. ; ARKIN, R. C.: Behavioral overlays for non-verbal communication expression on a humanoid robot. In: Autonomous Robots 22 (2007), January, S. 55–74
- [Butz 2008] BUTZ, M. V.: Sensomotorische Raumrepräsentationen. In: Informatik-Spektrum 31 (2008), S. 237–240
- [Caggiano u. a. 2009] CAGGIANO, V. ; FOGASSI, L. ; RIZZOLATTI, G. ; THIER, P. ; CASILE, A.: Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. In: Science (New York, N.Y.) 324 (2009), April, Nr. 5925, S. 403–406
- [Cardinali u. a. 2009] CARDINALI, L. ; BROZZOLI, C. ; FARNÈ, A.: Peripersonal Space and Body Schema: Two Labels for the Same Concept? In: Brain Topography 21 (2009), S. 252–260
- [Carello u. a. 1989] CARELLO, C. ; GROSOFSKY, A. ; REICHEL, F. D. ; SOLOMON, H. Y. ; TURVEY, M.: Visually Perceiving What is Reachable. In: Ecological Psychology 1 (1989), Nr. 1, S. 27–54
- [Cassell u. a. 2001] CASSELL, J. ; VILHJÁLMSSON, H. H. ; BICKMORE, T.: BEAT: the Behavior Expression Animation Toolkit. In: Proceedings of the 28th Annual

- Conference on Computer Graphics and Interactive Techniques. New York, NY, USA : ACM, 2001 (SIGGRAPH '01), S. 477–486
- [Chi u. a. 2000] CHI, D. ; COSTA, M. ; ZHAO, L. ; BADLER, N.: The EMOTE model for effort and shape. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 2000 (SIGGRAPH '00), S. 173–182
- [Clark 2007] CLARK, A.: Re-Inventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind. In: Journal of Medicine and Philosophy 32 (2007), Nr. 3, S. 263–282
- [Cohen und Andersen 2002] COHEN, Y. E. ; ANDERSEN, R. A.: A common reference frame for movement plans in the posterior parietal cortex. In: Nature Reviews Neuroscience 3 (2002), Nr. 7, S. 553–562
- [Conde und Thalmann 2006] CONDE, T. ; THALMANN, D.: An integrated perception for autonomous virtual agents: active and predictive perception. In: Journal of Visualization and Computer Animation 17 (2006), Nr. 3-4, S. 457–468
- [Di Pellegrino u. a. 1997] DI PELLEGRINO, G. ; LÀDAVAS, E. ; FARNÈ, A.: Seeing where your hands are. In: Nature 388 (1997), Nr. 6644, S. 730
- [Duhamel u. a. 1998] DUHAMEL, J.-R. ; COLBY, C. L. ; GOLDBERG, M. E.: Ventral intraparietal area of the macaque : congruent visual and somatic response properties. In: Journal of Neurophysiology 79 (1998), S. 126–136
- [Farnè u. a. 2000] FARNÈ, A. ; PAVANI, F. ; MENEGHELLO, F. ; LÀDAVAS, E.: Left tactile extinction following visual stimulation of a rubber hand. In: Brain 123 (Pt 11) (2000), November, S. 2350–2360
- [Farnè und Làdavas 2002] FARNÈ, A. ; LÀDAVAS, E.: Auditory peripersonal space in humans. In: Journal of Cognitive Neuroscience 14 (2002), Nr. 7, S. 1030–1043
- [Foster und Matheson 2008] FOSTER, M. E. ; MATHESON, C.: Following Assembly Plans in Cooperative, Task-Based Human-Robot Dialogue. In: Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue (Londial 2008). London, jun 2008, S. 1–8

- [Fuke u. a. 2007] FUKE, S. ; OGINO, M. ; ASADA, M.: Body Image Constructed from Motor and Tactile Images with Visual Information. In: International Journal of Humanoid Robotics (IJHR) 4 (2007), Nr. 2, S. 347–364
- [Fuke u. a. 2009] FUKE, S. ; OGINO, M. ; ASADA, M.: Acquisition of the Head-Centered Peri-Personal Spatial Representation Found in VIP Neuron. In: IEEE Transactions on Autonomous Mental Development 1 (2009), Nr. 2, S. 131–140
- [Gallagher 2005] GALLAGHER, S.: How the Body Shapes the Mind. Oxford : Clarendon Press, 2005. – 284 S. : Ill., graph. Darst.
- [Gallese 2005] GALLESE, V.: Embodied simulation: From neurons to phenomenal experience. In: Phenomenology and the Cognitive Sciences 4 (2005), Nr. 1, S. 23–48
- [Georgiou u. a. 2006] GEORGIU, I. ; BECCHIO, C. ; GLOVER, S. ; CASTIELLO, U.: Different action patterns for cooperative and competitive behaviour. In: Cognition 102 (2006), März, Nr. 3, S. 415–433
- [Gibson 1977] GIBSON, J.: The theory of affordances. In: SHAW, R. (Hrsg.) ; BRANSFORD, J. (Hrsg.): Perceiving, Acting and Knowing. New York: Wiley, 1977, S. 67–82
- [Goerick u. a. 2005] GOERICK, C. ; WERSING, H. ; MIKHAILOVA, I. ; DUNN, M.: Peripersonal space and object recognition for humanoids. In: Proceedings of the IEEE/RSJ International Conference on Humanoid Robots (Humanoids 2005), Tsukuba, Japan, IEEE Press, 2005, S. 387–392
- [Gray u. a. 2005] GRAY, J. ; BREAZEAL, C. ; BERLIN, M. ; BROOKS, A. ; LIEBERMAN, J.: Action parsing and goal inference using self as simulator. In: Proceedings of the Fourteenth IEEE Workshop on Robot and Human Interactive Communication, Ro-Man05, IEEE, 2005, S. 202–209
- [Graziano 1999] GRAZIANO, M. S.: Where Is My Arm? The Relative Role of Vision and Proprioception in the Neuronal Representation of Limb Position. In: Proceedings of the National Academy of Science 96 (1999), Aug., S. 10418–10421
- [Graziano u. a. 2000] GRAZIANO, M. S. ; COOKE, D. F. ; TAYLOR, C. S.: Coding the location of the arm by sight. In: Science 290 (2000), Nr. 5497, S. 1782–1786

- [Graziano und Gross 1998] GRAZIANO, M. S. ; GROSS, C. G.: Spatial maps for the control of movement. In: Current Opinion in Neurobiology 8 (1998), Nr. 2, S. 195–201
- [Graziano u. a. 1999] GRAZIANO, M. S. ; REISS, L. A. ; GROSS, C. G.: A neuronal representation of the location of nearby sounds. In: Nature 397 (1999), Feb., S. 428–430
- [Graziano u. a. 1997] GRAZIANO, M. S. A. ; HU, X. T. ; GROSS, C. G.: Coding the locations of objects in the dark. In: Science (New York, N.Y.) 277 (1997), Jul., Nr. 5323, S. 239–241
- [Graziano u. a. 1994] GRAZIANO, M. S. A. ; YAP, G. S. ; GROSS, C. G.: Coding of Visual Space By Premotor Neurons. In: Science 266 (1994), S. 1054–1057
- [Grön u. a. 2000] GRÖN, G. ; WUNDERLICH, A. P. ; SPITZER, M. ; TOMCZAK, R. ; RIEPE, M. W.: Brain activation during human navigation: gender-different neural networks as substrate of performance. In: Nature Neuroscience (2000), Apr., S. 404–408
- [Grüsser 1983] GRÜSSER, O.-J.: Multimodal Structure of the Extrapersonal Space. In: HEIN, A. (Hrsg.) ; JEANNEROD, M. (Hrsg.): Spatially Oriented Behavior. Springer, 1983, S. 327–352
- [Hall 1969] HALL, E. T.: The Hidden Dimension. New York : Anchor Books, 1969
- [Halligan und Marshall 1991] HALLIGAN, P. W. ; MARSHALL, J. C.: Left neglect for near but not far space in man. In: Nature 350 (1991), Nr. 6318, S. 498–500
- [Head und Holmes 1911] HEAD, H. ; HOLMES, G.: Sensory disturbances from cerebral lesions. In: Brain (1911), Nr. 34, S. 102–254
- [Hersch 2009] HERSCHE, M.: Adaptive sensorimotor peripersonal space representation and motor learning for a humanoid robot. Ph.D. Thesis EPFL Lausanne, 2009
- [Hersch u. a. 2008] HERSCHE, M. ; SAUSER, E. ; BILLARD, A.: Online Learning of the Body Schema. In: International Journal of Humanoid Robotics 5 (2008), Nr. 2, S. 161–181

- [Hoffman 2000] HOFFMAN, M. L.: Empathy and Moral Development. Cambridge University Press, 2000
- [Hoffmann u. a. 2010] HOFFMANN, M. ; MARQUES, H. ; ARIETA, A. H. ; SUMIOKA, H. ; LUNGARELLA, M. ; PFEIFER, R.: Body schema in robotics: a review. In: IEEE Trans. Auton. Mental Develop. 2 (2010), December, Nr. 4, S. 304–324
- [Holmes und Spence 2004] HOLMES, N. P. ; SPENCE, C.: The body schema and multisensory representation(s) of peripersonal space. In: Cognitive Processing 5 (2004), S. 94–105
- [Holthaus u. a. 2011] HOLTHAUS, P. ; PITSCH, K. ; WACHSMUTH, S.: How Can I Help? - Spatial Attention Strategies for a Receptionist Robot. In: International Journal of Social Robots 3 (2011), S. 383–393
- [Huang u. a. 2003] HUANG, Z. ; ELIËNS, A. ; VISSER, C. T.: Is It Within My Reach? - An Agents Perspective. In: Proceedings of the Intl. Conf. on Intelligent Virtual Agents, IVA 2003, 2003, S. 150–158
- [Jockusch u. a. 1997] JOCKUSCH, J. ; WALTER, J. ; RITTER, H.: A tactile sensor system for a three-fingered robot manipulator. In: Proceedings of the IEEE International Conference on Robotics and Automation Bd. 4, 1997, S. 3080–3086 vol.4
- [Johnson und Rickel 1997] JOHNSON, W. L. ; RICKEL, J.: Steve: an animated pedagogical agent for procedural training in virtual environments. In: SIGART Bulletin 8 (1997), December, S. 16–21
- [Kaas und van Mier 2006] KAAS, A. ; MIER, H. van: Haptic spatial matching in near peripersonal space. In: Experimental Brain Research 170 (2006), Nr. 3, S. 403–413
- [Kallmann 2008] KALLMANN, M.: Autonomous object manipulation for virtual humans. In: ACM SIGGRAPH 2008 classes. New York, NY, USA : ACM, 2008 (SIGGRAPH '08), S. 47:1–47:97
- [Kamashima u. a. 2004] KAMASHIMA, M. ; KANDA, T. ; IMAI, M. ; ONO, T. ; SAKAMOTO, D. ; ISHIGURO, H. ; ANZAI, Y.: Embodied cooperative behaviors by an autonomous humanoid robot. In: Proceedings of the IEEE/RSJ International

- Conference on Intelligent Robots and Systems. (IROS 2004) Bd. 3, 2004, S. 2506 – 2513 vol.3
- [Kendon 1990] KENDON, A.: Conducting Interaction. London : Cambridge University Press, 1990
- [Kessler und Thomson 2010] KESSLER, K. ; THOMSON, L.: The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. In: Cognition 114 (2010), Nr. 1, S. 72–88
- [Kessler und Thomson 2008] KESSLER, K. ; THOMSON, L. A.: I see the world through your eyes: The embodied nature of spatial perspective taking. In: Third international conference on cognitive science, Moscow, Russia Bd. 1, 2008, S. 80–82
- [Khatib 1986] KHATIB, O.: Real-time obstacle avoidance for manipulators and mobile robots. In: Int. J. Rob. Res. 5 (1986), Nr. 1, S. 90–98
- [Kopp] KOPP, S.: Synthese und Koordination von Sprache und Gestik für virtuelle multimodale Agenten. Infix Akademische Verlagsgesellschaft (DISKI, 265). – I–XIV, 1–260 S.
- [Kopp u. a. 2003] KOPP, S. ; JUNG, B. ; LESSMANN, N. ; WACHSMUTH, I.: Max - A Multimodal Assistant in Virtual Reality Construction. In: KI-Künstliche Intelligenz 4/03 (2003), S. 11–17
- [Kopp und Wachsmuth 2004] KOPP, S. ; WACHSMUTH, I.: Synthesizing multimodal utterances for conversational agents. In: Comput. Animat. Virtual Worlds 15 (2004), Nr. 1, S. 39–52
- [Kopp u. a. 2008] KOPP, S. ; WACHSMUTH, I. ; BONAIUTO, J. ; ARBIB, M.: Imitation in Embodied Communication – From Monkey Mirror Neurons to Artificial Humans. In: WACHSMUTH, I. (Hrsg.) ; LENZEN, M. (Hrsg.) ; KNOBLICH, G. (Hrsg.): Embodied Communication. Oxford : Oxford University Press, 2008, S. 357–390
- [Lacquaniti 1997] LACQUANITI, F.: Frames of reference in sensorimotor coordination. In: FRANCOIS BOLLER, J. G. (Hrsg.): Handbook of Neuropsychology Bd. 11. Elsevier Science, 1997, S. 27–64

- [Latombe 1991] LATOMBE, J.-C.: Robot Motion Planning. Kluwer Acad. Publ., 1991 (The Kluwer international series in engineering and computer science ; 124). – XVIII, 651 S. : graph. Darst.
- [Lessmann u. a. 2006] LESSMANN, N. ; KOPP, S. ; WACHSMUTH, I.: Situated interaction with a virtual human - perception, action, and cognition. In: RICKHEIT, G. (Hrsg.) ; WACHSMUTH, I. (Hrsg.): Situated Communication. Berlin : Mouton de Gruyter, 2006, S. 287–323
- [Liu 2003] LIU, Y.: Interactive reach planning for animated characters using hardware acceleration. Ph.D. Thesis University of Pennsylvania, 2003
- [Lloyd 2009] LLOYD, D. M.: The space between us: A neurophilosophical framework for the investigation of human interpersonal space. In: Neuroscience & Biobehavioral Reviews 33 (2009), Nr. 3, S. 297–304
- [Longo und Lourenco 2006] LONGO, M. ; LOURENCO, S.: On the nature of near space: Effects of tool use and the transition to far space. In: Neuropsychologia 44 (2006), Nr. 6, S. 977–981
- [Longo und Lourenco 2007] LONGO, M. ; LOURENCO, S.: Space perception and body morphology: extent of near space scales with arm length. In: Experimental Brain Research 177 (2007), S. 285–290
- [Làdavvas 2002] LĀDAVAS, E.: Functional and dynamic properties of visual peripersonal space. In: Trends in Cognitive Sciences 6 (2002), Nr. 1, S. 17 – 22
- [Maravita und Iriki 2004] MARAVITA, A. ; IRIKI, A.: Tools for the body (schema). In: Trends in Cognitive Sciences 8 (2004), Nr. 2, S. 79–86
- [Maravita u. a. 2003] MARAVITA, A. ; SPENCE, C. ; DRIVER, J.: Multisensory integration and the body schema: close to hand and within reach. In: Current Biology 13 (2003), Nr. 13, S. R531 – R539
- [Maravita u. a. 2002] MARAVITA, A. ; SPENCE, C. ; SERGENT, C. ; DRIVER, J.: Seeing Your Own Touched Hands in a Mirror Modulates Cross-Modal Interactions. In: Psychological Science 13 (2002), jul, Nr. 4, S. 350–355

- [Mark u. a. 1997] MARK, L. S. ; NEMETH, K. ; GARDNER, D. ; DAINOFF, M. J. ; PAASCHE, J. ; DUFFY, M. ; GRANDT, K.: Postural dynamics and the preferred critical boundary for visually guided reaching. In: J. Exp. Psychol. Hum. Percept. Perform. 23 (1997), Nr. 5, S. 1365–79
- [McKenzie u. a. 1993] MCKENZIE, B. E. ; SKOUTERIS, H. ; DAY, R. ; HARTMAN, B. ; YONAS, A.: Effective action by infants to contact objects by reaching and leaning. In: Child Development (1993), Nr. 64, S. 415–29
- [Minato u. a. 2007] MINATO, T. ; YOSHIKAWA, Y. ; NODA, T. ; IKEMOTO, S. ; ISHIGURO, H. ; ASADA, M.: CB2: A child robot with biomimetic body for cognitive developmental robotics. In: 7th IEEE-RAS International Conference on Humanoid Robots, 2007, S. 557–562
- [Mountcastle 1976] MOUNTCASTLE, V. B.: The world around us: Neural command functions for selective attention. In: Neurosciences Research Program Bulletin 14 (1976), S. 1–47
- [Nguyen 2007] NGUYEN, N.: Modeling and Simulation of Tactile Perception for Artificial Virtual Agents, University of Bielefeld, Faculty of Technology, Diplomarbeit, 2007
- [Nguyen und Wachsmuth 2009a] NGUYEN, N. ; WACHSMUTH, I.: Modeling peripersonal action space for a virtual human by learning a tactile body schema. In: KI 2009: Advances in Artificial Intelligence. Paderborn, Germany, Berlin: Springer (LNAI 5803), 2009, S. 516–523
- [Nguyen und Wachsmuth 2009b] NGUYEN, N. ; WACHSMUTH, I.: Modeling Peripersonal Action Space for Virtual Humans Using Touch and Proprioception. In: RUTTKAY, Z. (Hrsg.) ; KIPP, M. (Hrsg.) ; NIJHOLT, A. (Hrsg.) ; VILHJALMSSON, H. H. (Hrsg.): Proceedings of the 9th Conference on Intelligent Virtual Agents. Berlin : Springer (LNAI 5773), 2009, S. 63–75
- [Nguyen und Wachsmuth 2011] NGUYEN, N. ; WACHSMUTH, I.: From body space to interaction space: modeling spatial cooperation for virtual humans. In: The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume

3. Richland, SC : International Foundation for Autonomous Agents and Multiagent Systems, 2011 (AAMAS '11), S. 1047–1054
- [Nguyen und Wachsmuth 2013] NGUYEN, N. ; WACHSMUTH, I.: A computational model of cooperative spatial behavior for virtual humans. In: TENBRINK, T. (Hrsg.) ; WIENER, J. (Hrsg.) ; CLARAMUNT, C. (Hrsg.): Representing Space in Cognition: Behavior, Language, and Formal Models. Oxford University Press, 2013 (in press)
- [Nguyen u. a. 2007] NGUYEN, N. ; WACHSMUTH, I. ; KOPP, S.: Touch perception and emotional appraisal for a virtual agent. In: Proceedings Workshop Emotion and Computing - Current Research and Future Impact, in conj. with 30th Annual German Conference on Artificial Intelligence (KI 2007). Osnabrück, Germany, 2007, S. 17–22
- [Pedica und Vilhjálmsson 2009] PEDICA, C. ; VILHJÁLMSSON, H.: Spontaneous Avatar Behavior for Human Territoriality. In: RUTTKAY, Z. (Hrsg.) ; KIPP, M. (Hrsg.) ; NIJHOLT, A. (Hrsg.) ; VILHJÁLMSSON, H. (Hrsg.): Intelligent Virtual Agents Bd. 5773. Springer Berlin / Heidelberg, 2009, S. 344–357
- [Pettré u. a. 2008] PETTRÉ, J. ; KALLMANN, M. ; LIN, M. C.: Motion planning and autonomy for virtual humans. In: ACM SIGGRAPH 2008 classes. New York, NY, USA : ACM, 2008 (SIGGRAPH '08), S. 42:1–42:31
- [Pfeiffer-Lessmann 2011] PFEIFFER-LESSMANN, N.: Kognitive Modellierung von Kooperationsfähigkeiten für einen künstlichen Agenten, Universität Bielefeld, Diss., 2011. – Bielefeld Universitätsbibliothek
- [Pickering und Garrod 2004] PICKERING, M. J. ; GARROD, S.: Toward a mechanistic psychology of dialogue. In: Behavioral and Brain Sciences 27 (2004), Nr. 02, S. 169–190
- [Previc 1998] PREVIC, F. H.: The neuropsychology of 3-D space. In: Psychological Bulletin 124 (1998), Nr. 2, S. 123–164
- [Rao und Georgeff 1991] RAO, A. S. ; GEORGEFF, M. P.: Modeling Rational Agents within a BDI-Architecture. In: ALLEN, J. (Hrsg.) ; FIKES, R. (Hrsg.) ; SANDEWALL, E. (Hrsg.): Proceedings of the 2nd International Conference on Principles of

- Knowledge Representation and Reasoning (KR '91), Morgan Kaufmann, 1991, S. 473–484
- [Rickert u. a. 2007] RICKERT, M. ; FOSTER, M. ; GIULIANI, M. ; BY, T. ; PANIN, G. ; KNOLL, A.: Integrating Language, Vision and Action for Human Robot Dialog Systems. In: STEPHANIDIS, C. (Hrsg.): Universal Access in Human-Computer Interaction. Ambient Interaction Bd. 4555. Berlin, Heidelberg : Springer Berlin / Heidelberg, 2007, Kapitel 108, S. 987–995
- [Rizzolatti u. a. 1996] RIZZOLATTI, G. ; FADIGA, L. ; MATELLI, M. ; BETTINARDI, V. ; PAULESU, E. ; PERANI, D. ; FAZIO, F.: Localization of grasp representations in humans by PET: 1. Observation versus execution. In: Experimental Brain Research 111 (1996), Nr. 2, S. 246–252
- [Rizzolatti u. a. 1981] RIZZOLATTI, G. ; SCANDOLARA, C. ; MATELLI, M. ; GENTILUCCI, M.: Afferent properties of periarculate neurons in macaque monkeys. II. Visual Responses. In: Behavioural Brain Research 2 (1981), Nr. 2, S. 147–163
- [Rochat 1998] ROCHAT, P.: Self-perception and action in infancy. In: Experimental Brain Research 123 (1998), S. 102–109
- [Röthling u. a. 2007] RÖTHLING, F. ; HASCHKE, R. ; STEIL, J. ; RITTER, H.: Platform portable anthropomorphic grasping with the bielefeld 20-DOF shadow and 9-DOF TUM hand. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007. IROS 2007., 2007, S. 2951–2956
- [Russo u. a. 2006] RUSSO, F. D. ; COMMITTERI, G. ; PITZALIS, S. ; SPITONI, G. ; PICCARDI, L. ; GALATI, G. ; CATAGNI, M. ; NICO, D. ; GUARIGLIA, C. ; PIZZAMIGLIO, L.: Cortical plasticity following surgical extension of lower limbs. In: NeuroImage 30 (2006), Nr. 1, S. 172 – 183
- [Sartori u. a. 2011] SARTORI, L. ; BECCHIO, C. ; CASTIELLO, U.: Cues to intention: The role of movement information. In: Cognition 119 (2011), Mai, Nr. 2, S. 242–252
- [Semmes u. a. 1963] SEMMES, J. ; WEINSTEIN, S. ; GHENT, L. ; TEUBER, H.-L.: Correlates of Impaired Orientation in Personal and Extrapersonal Space. In: Brain (1963), Nr. 86, S. 747–772

- [Shapiro u. a. 2007] SHAPIRO, A. ; KALLMANN, M. ; FALOUTSOS, P.: Interactive motion correction and object manipulation. In: Proceedings of the 2007 Symposium on Interactive 3D Graphics and Games. New York, NY, USA : ACM, 2007 (I3D '07), S. 137–144
- [Sisbot u. a. 2006] SISBOT, E. A. ; MARIN, L. F. ; ALAMI, R. ; SIMEON, T.: A mobile robot that performs human acceptable motion. In: Proceedings of the (IEEE/RSJ) International Conference on Intelligent Robots and Systems, 2006, S. 1811 – 1816
- [Spong u. a. 2005] SPONG, M. W. ; HUTCHINSON, S. ; VIDYASAGAR, M.: Robot Modeling and Control. New York : Wiley, 2005
- [Steffen u. a. 2007] STEFFEN, J. ; HASCHKE, R. ; RITTER, H.: Experience-based and tactile-driven dynamic grasp control. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007. IROS 2007., 2007, S. 2938–2943
- [Stilman u. a. 2007] STILMAN, M. ; SCHAMBUREK, J. ; KUFFNER, J. ; ASFOUR, T.: Manipulation planning among movable obstacles. In: In IEEE Int'l Conf. on Robotics and Automation (ICRA'07), 2007, S. 3327 – 3332
- [Stoytchev 2003] STOYTCHEV, A.: Computational Model for an Extendable Robot Body Schema / Georgia Institute of Technology, College of Computing. 2003 (GIT-CC-03-44). – Forschungsbericht. – 15 S.
- [Sturm u. a. 2009] STURM, J. ; PLAGEMANN, C. ; BURGARD, W.: Body schema learning for robotic manipulators from visual self-perception. In: Journal of Physiology - Paris 103 (2009), Nr. 3–5, S. 220 – 231
- [Thier 2006] THIER, P.: Anatomie und Physiologie des parietalen Kortex. In: KARNATH, H.-O. (Hrsg.) ; THIER, P. (Hrsg.): Neuropsychologie. Springer, 2006, S. 160–176
- [Tolani u. a. 2000] TOLANI, D. ; GOSWAMI, A. ; BADLER, N. I.: Real-time inverse kinematics techniques for anthropomorphic limbs. In: Graph. Models Image Process. 62 (2000), Sept., Nr. 5, S. 353–388
- [Tsuno 2008] TSUNO, Y.: Motoman SDA10 kocht. <http://www.afpbb.com/article/environment-science-it/science-technology/2543168/3561420>.
Version: November 2008. – letzter Zugriff am 15. Februar 2012

- [Vallar und Maravita 2009] VALLAR, G. ; MARAVITA, A.: Personal and Extrapersonal Spatial Perception. In: Handbook of Neuroscience for the Behavioral Sciences. John Wiley & Sons, Inc., 2009, S. 322–336
- [Watson u. a. 2004] WATSON, M. E. ; BRANIGAN, H. P. ; PICKERING, M. J.: Alignment of reference frames in dialogue. In: Proceedings of the 26th Annual Conference of the Cognitive Science Society, 2004, S. 1434–1440
- [Watson u. a. 2009] WATSON, M. E. ; PICKERING, M. J. ; BRANIGAN, H. P.: Why Dialogue Methods are Important for Investigating Spatial Language. In: Spatial Language and Dialogue 1 (2009), Nr. 9, S. 8–23
- [Wiendl u. a. 2007] WIENDL, V. ; DORFMÜLLER-ULHAAS, K. ; SCHULZ, N. ; ANDRÉ, E.: Integrating a Virtual Agent into the Real World: The Virtual Anatomy Assistant Ritchie. In: Proceedings of the 7th International Conference on Intelligent Virtual Agents. Berlin, Heidelberg : Springer-Verlag, 2007 (IVA '07), S. 211–224
- [Wooldridge 2009] WOOLDRIDGE, M.: An Introduction to MultiAgent Systems. 2nd. Wiley Publishing, 2009
- [Yamaoka u. a. 2010] YAMAOKA, F. ; KANDA, T. ; ISHIGURO, H. ; HAGITA, N.: A model of proximity control for information-presenting robots. In: Trans. Rob. 26 (2010), Nr. 1, S. 187–195
- [Yaskawa 2008] YASKAWA, C.: Roboter Motoman SDA10. <http://www.afpbb.com/article/environment-science-it/science-technology/2543168/3561420>. Version: 2008. – letzter Zugriff am 15. Februar 2012
- [Yoshikawa u. a. 2002] YOSHIKAWA, Y. ; KAWANISHI, H. ; ASADA, M. ; HOSODA, K.: Body Scheme Acquisition by Cross Modal Map Learning among Tactile, Visual, and Proprioceptive Spaces. Lund University Cognitive Studies, 2002, S. 181–184