

THORSTEN STEIN, ANDREAS FISCHER, INGO BOESNACH,
DIRK GEHRIG, HILDEGARD KÖHLER, HERMANN
SCHWAMEDER

Kinematische Analyse menschlicher Alltagsbewegungen für die Mensch-Maschine-Interaktion

1. Einleitung

Das Thema „Humanoide Roboter“ erlebte aufgrund der technischen Fortschritte in den letzten 15 Jahren eine Renaissance (Becher et al., 2004). Neben der Erforschung der menschlichen Fortbewegung konzentriert sich die Forschung auf die Entwicklung humanoider, multifunktionaler Servicemaschinen. Der Terminus „humanoide“ steht für eine annähernd menschliche Gestalt und zumindest im Ansatz vergleichbare sensorische und motorische Eigenschaften. Dabei ist die Anzahl der Freiheitsgrade sowie Art und Anordnung der Sensoren dem Vorbild Mensch nachempfunden (Asfour, 2003). In der Scientific Community ist man sich einig, dass Serviceroboter unabhängig von ihrem Einsatzbereich eine humanoide Bauform aufweisen sollten (Khatib et al., 2004). Ziel ist es, die Maschinen in den Alltag des Menschen zu integrieren, sie sollen mit den Menschen koexistieren. Dies setzt voraus, dass sich die Roboter die Lebensräume des Menschen erschließen können. Durch die (annähernd) humanoide Bauform kann der Roboter dem Menschen überall hin folgen und die Umgebung, die in aller Regel auf die Anatomie des Menschen ausgelegt ist, muss nicht an den Roboter angepasst werden. Ein weiteres Argument für eine humanoide Bauform und Bewegungscharakteristik ist, dass die Bewegungen des Roboters für den Menschen leichter vorherzusagen sind und somit die Interaktion zwischen Mensch und Maschine erleichtert wird. Die potentiellen Einsatzszenarien sind vielfältiger Natur und reichen von Serviceleistungen in Werkstätten bis zum privaten Wohnbereich (Tanie, 2003).

Vor diesem Hintergrund wurde an der Universität Karlsruhe (TH) im Juli 2001 von der Deutschen Forschungsgemeinschaft der interdisziplinäre Sonderforschungsbereich 588 „Humanoide Roboter – Lernende und kooperierende multimodale Roboter“ eingerichtet. Das Projekt ist auf 12 Jahre angelegt und darauf ausgerichtet, Konzepte, Methoden und konkrete mechatronische Komponenten für einen humanoiden Serviceroboter zu entwickeln (Abb. 1), der im direkten Kontakt mit dem Menschen ein hilfreicher Assistent im Alltag sein soll (Becher et al., 2004).

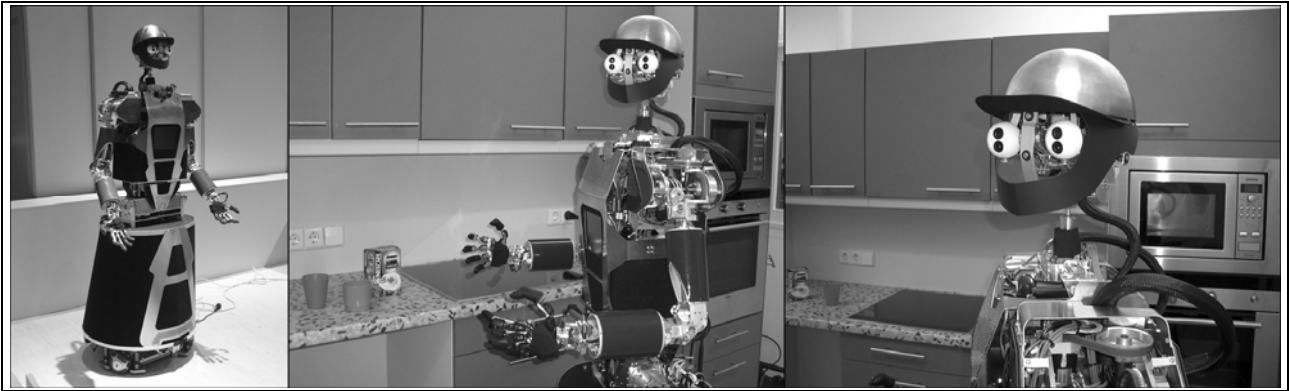


Abb. 1: Teilanthropomorphes Robotersystem ARMAR IIIA (Sonderforschungsbereich 588 "Humanoide Roboter", 2007).

Um alltägliche Kooperationen zwischen Mensch und Maschine zu erleichtern, sollte nicht nur das Design des Bewegungsapparates, sondern auch die Bewegungscharakteristik des Roboters am Menschen orientiert sein.

Menschliche Bewegungen zeigen einerseits konstante (qualitativ-topologische) äußere Merkmale, die es erlauben, sie einer bestimmten Klasse von Bewegungen zuzuordnen (z.B. Greifen oder Werfen). Dabei können die konstanten äußeren Merkmale auf unterschiedlichen internen Prozessen (z.B. Muskelaktivierungen) basieren. Andererseits ergeben sich innerhalb einer Bewegungsklasse sogar unter konstanten Ausführungsbedingungen immer auch variable (quantitativ-metrische) äußere Merkmale (Wiemeyer, 1994). Die Bewegungsvariabilitäten sind u.a. auf die enorme Komplexität des biologischen Bewegungssystems zurückzuführen. So verfügt der Mensch auf der Ebene der kortikalen Neuronen, der Muskeln und der Gelenke über eine unvorstellbar große Anzahl an Bewegungsfreiheitsgraden (Turvey, 1990). Eine koordinierte Bewegung des Menschen zur Lösung einer Alltagsaufgabe besteht also letztendlich darin, das Problem der Freiheitsgrade zu lösen (Bernstein, 1967). Die Bewegungen des Robotersystem ARMAR werden durch Motoren erzeugt, wodurch die Anzahl der zu kontrollierenden Freiheitsgrade im technischen System deutlich geringer ist als im biologischen Originalsystem. Vor diesem Hintergrund erfordert die Übertragung menschlicher Bewegungsmuster auf die Maschine eine Analyse der Kinematik menschlicher Alltagsbewegungen.

Es wurden verschiedene Alltagsbewegungen, wie z.B. das Beladen einer Spülmaschine, das Aufschrauben einer Kaffeedose oder das Einschenken einer Flüssigkeit in ein Glas kinematisch vermessen. Im Kontext der Mensch-Maschine-Interaktion sind Zeigegesten von besonderem Interesse, da Mensch und Maschine mit einer solchen Geste den Interaktionspartner auf aktuell wichtige Dinge in der unmittelbaren Umgebung aufmerksam machen können. Anhand von Zeigegesten wird in diesem Beitrag die Übertragung menschlicher Bewegungen auf ein humanoides Robotersystem skizziert.

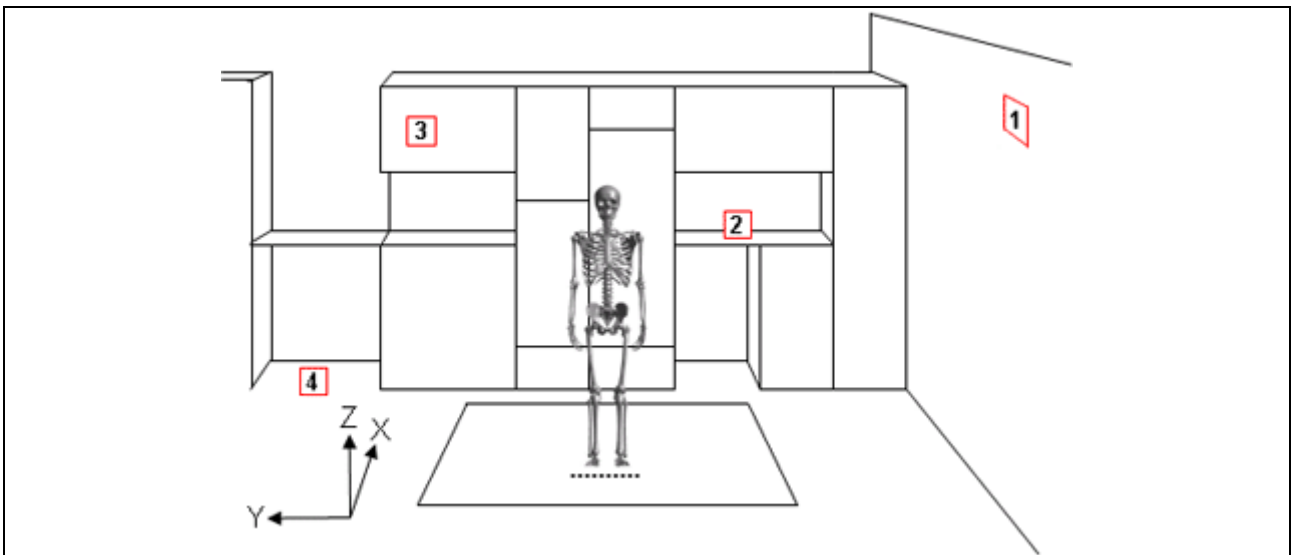


Abb. 2: Ausgangsposition aller Probanden in der SFB-Küche. Die Nummern an der Wand zeigen die Ziele an, auf die die Probanden zeigen mussten.

Die Bewegungen wurden mit Hilfe eines IR-Tracking Systems vermessen und Gelenkwinkel in Matlab rekonstruiert (Kap. 2). In einem ersten Schritt wird nach allgemeinen Bewegungsprinzipien gesucht, die Hinweise liefern, nach welchen Regeln der Mensch im Alltag seine Bewegungsfreiheitsgrade einschränkt (Kap. 3). Diese Informationen sollen Informatikern und Ingenieuren helfen, biologisch inspirierte Regelungs- u. Steuerungsalgorithmen zu entwickeln. Das akquirierte Datenmaterial wird in einem weiteren Schritt für das Training adaptiver stochastischer Modelle herangezogen, die die Bewegungsplanung nach menschlichem Vorbild verbessern und eine Erkennung menschlicher Bewegungen ermöglichen sollen (Kap. 4).

2. Methode

An der Studie nahmen 20 Studierende der Universität Karlsruhe (16 Männer u. 4 Frauen) im Alter zwischen 20 und 25 Jahren ($M = 22,2$; $SD = 1,3$) teil. Ihre Körpergröße lag zwischen 160 und 189 cm ($M = 177$; $SD = 9$). Die Zeigebewegungen wurden in der Demonstrationsküche des SFB 588 am Institut für Technische Informatik aufgenommen. Alle Probanden standen an derselben Stelle in der Küche mit Blick auf eine Robotergrafik an der gegenüberliegenden Wand. Es wurden vier Tafeln mit Ziffern auf den Küchenmöbeln, der Wand sowie dem Boden angebracht (Abb. 2). Der Versuchsleiter gab z.B. die Anweisung „Nummer 1“, woraufhin die Testperson auf die entsprechende Nummer zeigte. Die Testpersonen wurden vor der Versuchsdurchführung instruiert, die Bewegungen in üblicher Geschwindigkeit und Präzision durchzuführen. Darüber hinaus wurde es der Testperson freigestellt, mit welcher Hand die Zeigegeste ausgeführt wird. Die Startposition und die Reihenfolge der Nummernansage waren dagegen standardisiert, wobei die Probanden über die

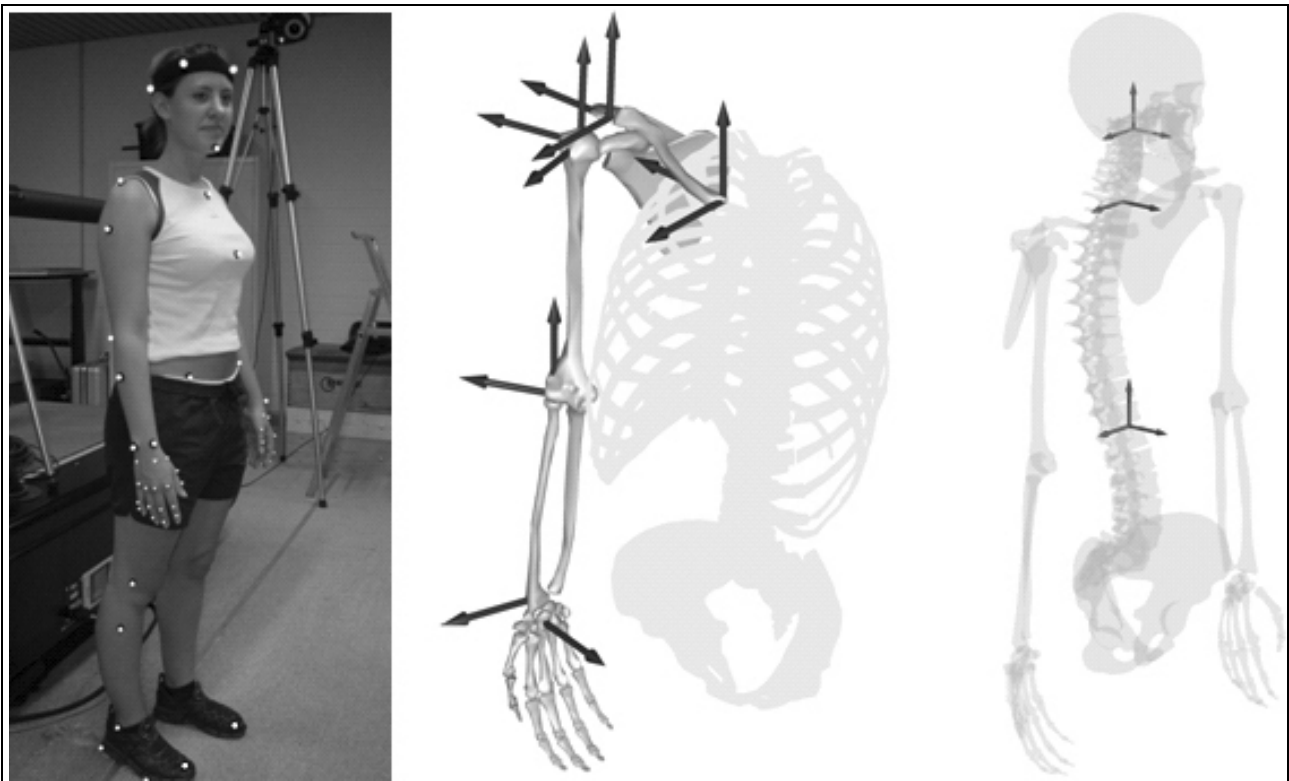


Abb. 3: Markerset (links), Drehachsen des Schulter-Arm-Systems (Mitte) und der Wirbelsäule (rechts).

Reihenfolge der Nummeransage nicht informiert wurden. Jede Nummer wurde fünfmal aufgerufen, sodass jeder Proband insgesamt 20 Zeigegesten ausführen musste.

Alle Versuche wurden mit einem IR-Tracking-System der Firma Vicon (MX 13) mit einer Bildfrequenz von 120 Hz aufgenommen. Es wurde ein eigenes Markerset entwickelt, das aus 80 Markerpunkten bestand (Abb. 3). Um die 3D-Koordinaten der Markerpunkte berechnen zu können, mussten für jeden Marker zu jedem Zeitpunkt die 2D-Koordinaten aus der Sicht von mindestens zwei Kameras zur Verfügung stehen. Die 3D-Rekonstruktion der Markerpunkte erfolgte mit Hilfe der Software „Workstation“ (V. 4.6). Die Daten wurden mit einem Splineverfahren nach Woltring (1986) mit MSE Werten von bis zu 15 geglättet.

Die einzelnen Körpersegmente wurden als starr angenommen (Cappozzo et al., 2005). In Anlehnung an die Kinematik des Menschen wurde ein Menschmodell aus 52 Starrkörpern entwickelt (Stelzner et al., 2007). Das Modell verfügt über 40 Freiheitsgrade ohne und 84 Freiheitsgrade mit Berücksichtigung der Hände (Abb. 3). Die mathematische Beschreibung erfolgt über ein rekursives Verfahren in Minimalkoordinaten (relative Gelenkkoordinaten) und erlaubt in Zusammenhang mit dem entwickelten Markerset die Rekonstruktion der Modellfreiheitsgrade. Die Übertragung der gemessenen Bewegung erfolgt über eine nichtlineare Optimierung im Sinne eines „Least-Squares“-Ansatzes.

3. Ergebnisse

Im Folgenden werden am Beispiel des Zeigens auf das Ziel 3 (Abb. 2) ausgewählte Befunde veranschaulicht.

Bis auf einen Versuch eines Probanden zeigten alle Versuchsteilnehmer auf das Ziel 3 mit der rechten Hand. Dieses befand sich nicht im Sichtfeld der Probanden und konnte durch ein ausschließliches Bewegen der Augen nicht lokalisiert werden. Obgleich die Probanden wussten, wo sich das Ziel im Raum befand, konnte bei allen Probanden eine Bewegung des Kopfes zum Ziel hin beobachtet werden. Dabei verließen vier der 20 Probanden ihre Startposition. Bei den 16 Versuchsteilnehmern, die die Ausgangsstellung nicht verließen, wurde in der Mehrzahl der Fälle die Bewegung durch eine Kopfdrehung eingeleitet (Abb. 4). Zu dem Zeitpunkt, zu dem das Ziel durch die Kopfdrehung in das Blickfeld des Probanden kam, war die Zeigegeste bei allen Probanden noch nicht beendet (Abb. 4). Befunde von Slater-Hammel (1960) deuten darauf hin, dass Bewegungen, die länger als 150-200 ms andauern, einer Feedbackregelung grundsätzlich zugänglich sind. Da es sich im vorliegenden Fall um eine zielgerichtete Bewegung handelt, die deutlich länger als 200 ms andauert, erscheint es plausibel, dass die Probanden das visuelle Feedback für die Feinabstimmung der Bewegung im Endpunkt nutzen.

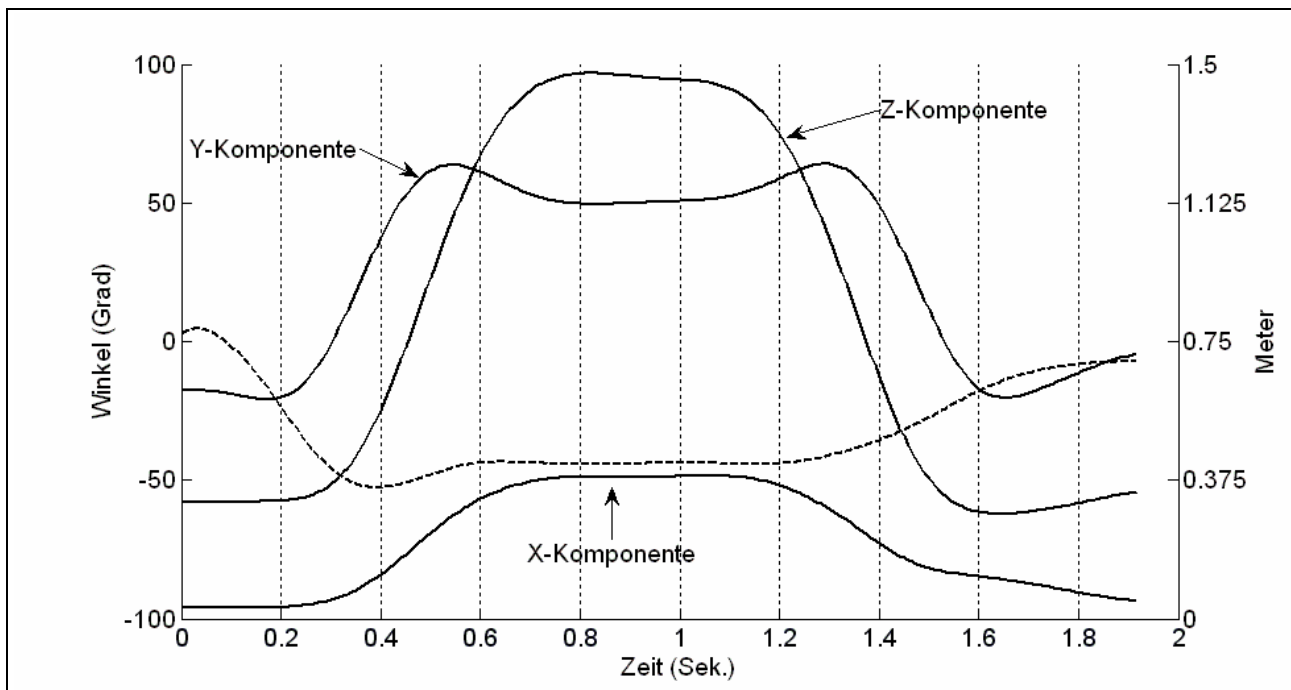


Abb. 4: Typische Weg- u. Winkel-Zeit-Verläufe einer Probandin. Dabei ist die Drehung des Kopfes nach rechts in Richtung des Ziels 3 gestrichelt und die Bewegung der Hand entlang der X-, Y- u. Z-Komponente des globalen Koordinatensystems jeweils mit durchgezogener Linie dargestellt.

Die Bewegungsbahn des Endeffektors verlief bei allen Probanden entlang einer relativ geradlinigen bzw. leicht gekrümmten Bahn mit einem glockenförmigen Geschwindigkeitsprofil (Abb. 5). Diese Befunde decken sich mit den Resultaten von Analysen zielgerichteter menschlicher Bewegungen (Goodman & Gottlieb, 1995).

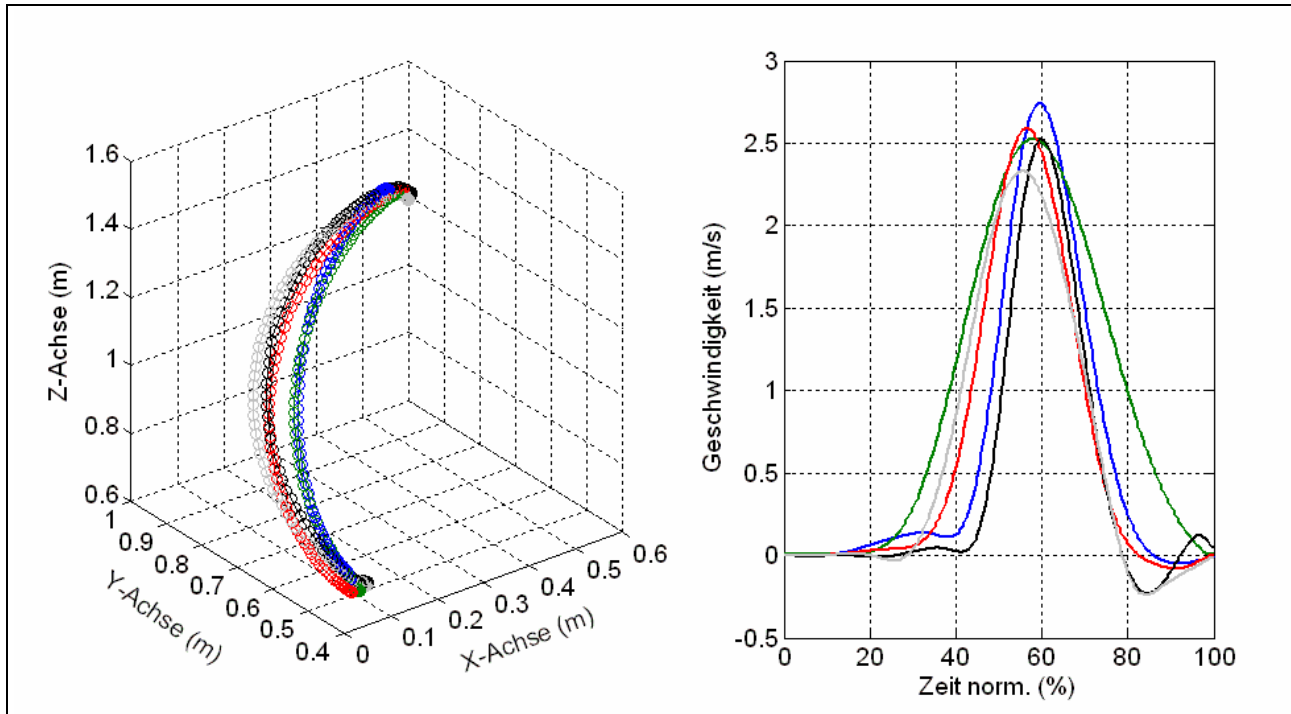


Abb. 5: Fünf Bewegungsbahnen der rechten Hand einer Probandin (links) und die entsprechenden Geschwindigkeitsprofile der Hand in der Z-Komponente (rechts).

4. Modellierung menschlicher Bewegungstrajektorien

4.1 Hidden Markov Modelle

Zur Erkennung menschlicher Bewegungen und Erzeugung menschenähnlicher Bewegungstrajektorien auf Basis der aufgezeichneten und berechneten Größen werden Hidden Markov Modelle (HMM) eingesetzt.

HMM sind adaptive, stochastische Modelle. Es wird angenommen, dass es in dem beobachteten Prozess (z.B. Zeigegeste) versteckte Zustände gibt, die nicht direkt bestimmt werden können. Diese Zustände können nur aufgrund der Beobachtungen (z.B. Gelenkwinkelverläufe) geschätzt werden. Um dies zu ermöglichen, wird jedem HMM eine Menge von Bewegungstrajektorien zugewiesen, die zu einer Kategorie von Bewegungen gehört. Anhand dieser Bewegungstrajektorien werden die Wahrscheinlichkeiten für mögliche Beobachtungen in einem Zustand und die Wahrscheinlichkeiten für einen Übergang von einem Zustand in einen anderen berechnet. Während der Trai-

ningsphase werden die Parameter iterativ angepasst, um die Wahrscheinlichkeit der Bewegungstrajektorien für das Modell zu erhöhen. Eine ausführliche Darstellung des Ansatzes der HMM findet man bei Rabiner (1989).

Jede der Bewegungstrajektorien besteht aus einer Sequenz von elementaren Bewegungsphasen, die in verschiedenen zeitlichen Reihenfolgen vorkommen. Deshalb ist es nicht angebracht, für jede Bewegungstrajektorie ein HMM zu trainieren, sondern es wird jede elementare Bewegung durch ein HMM modelliert und diese zu komplexeren Modellen kombiniert.

4.2 Modellierung von elementaren Bewegungen

Im ersten Schritt müssen die Übergänge zwischen den elementaren Bewegungen gefunden werden. Bisher wurde dies manuell anhand von Videoaufzeichnungen durchgeführt (Beth et al., 2003), wird aber durch eine automatische Vorverarbeitung erweitert werden. Die Parameter der einzelnen elementaren HMMs werden zuerst initialisiert und dann durch den Baum-Welch-Algorithmus (Rabiner, 1989) iterativ an die Trainingsdaten der entsprechenden elementaren Bewegung angepasst.

Durch eine Komposition der HMMs zur Modellierung von komplexen Bewegungen basierend auf den elementaren HMMs ist es möglich, größere HMMs für die Erkennung von längeren Bewegungen zu erzeugen (Moldenhauer et al., 2006). Während der Komposition der elementaren HMMs zu einem komplexen HMM werden die Zustandsübergangsmatrizen der elementaren HMMs zu einer Gesamtmatrix kombiniert. Die initialen Zustandswahrscheinlichkeiten und die Parameter der Wahrscheinlichkeitsfunktionen sind durch die Komposition der Parameter der elementaren HMMs definiert. Das komplexe HMM benötigt kein weiteres Training, da die elementaren HMMs bereits trainiert worden sind. Das HMM ermöglicht die Erkennung einer Bewegung mit einer fest vorgegebenen Reihenfolge der elementaren Bewegungen.

Die erweiterte Version des Kompositionsalgorithmus ist in der Lage, komplexe HMMs mit alternativen Sequenzen von Bewegungsphasen zu konstruieren (Moldenhauer et al., 2006). Dies kann dazu benutzt werden, verschiedene mögliche Realisierungen derselben elementaren Bewegung in einer komplexen Bewegung zu modellieren.

4.3 Erkennung menschlicher Bewegungen

Für die Interaktion eines Roboters mit seiner Umgebung sollte der Roboter in der Lage sein, geeignet auf die nicht-verbale Eingaben des Benutzers zu reagieren. Deshalb ist es notwendig, dass der Roboter die Bewegungen eines Menschen erkennt.

In einer unserer früheren Arbeiten (Moldenhauer et al., 2005), wurden elementare HMMs mit Bewegungstrajektorien trainiert, die aus 8-dimensionalen Vektoren von Gelenkwinkeln für jeden Zeitpunkt bestanden. Die Bewegungen

gliederten sich in 13 Phasen, wie z.B. das Decken eines Tisches mit einer Tasse oder einer Untertasse. Entsprechend des vorherigen Kapitels wurde ein komplexes HMM für die komplexe Bewegung erzeugt.

Durch das Training mit menschlichen Gelenkwinkeln haben die HMMs die Charakteristika menschlicher Bewegungen gelernt und dienen als online Klassifikatoren. Anhand der Gelenkwinkelverläufe menschlicher Bewegungen wird das komplexe HMM dazu benutzt, mit dem Viterbi-Algorithmus (Rabiner, 1989) die Grenzen zwischen den elementaren Bewegungen zu berechnen. Bei den beschriebenen Bewegungstrajektorien wurden die 13 Bewegungsphasen mit einer Wahrscheinlichkeit von 84% richtig erkannt (Moldenhauer et al., 2005). Drei der 13 Bewegungen, die zum Einschenken mit einer Kanne in eine Tasse gehören, wurden in zwei verschiedenen Varianten durchgeführt, mit voller und mit fast leerer Kanne. Mit einem entsprechenden komplexen HMM war es möglich, die Bewegungen mit voller Kanne von denen mit fast leerer Kanne zu unterscheiden, d.h. ein komplexes HMM kann neben dem Finden von Phasengrenzen auch zwischen verschiedenen Bewegungen unterscheiden (Moldenhauer et al., 2005).

4.4 Menschenähnliche Roboterbewegungen

Ein weiteres Ziel ist die Erzeugung menschenähnlicher Roboterbewegungen aus den aufgenommenen menschlichen Bewegungstrajektorien. Der klassische Ansatz der Bewegungsgenerierung für Roboter besteht aus einem Bewegungsplaner, einem Bewegungsgenerator und einer Bewegungskontrolleinheit. Basierend auf diesem Schema gibt es einige Ansätze, um aufgabenorientierte Bewegungen menschenähnlicher zu machen. Damit kann jedoch bestenfalls die Integration einiger zusätzlicher Randbedingungen in die inverse Kinematik erreicht werden und es bleibt zu hoffen, dass die resultierende Bewegung menschenähnlich aussieht.

Die wünschenswerten Bewegungen eines zukünftigen Roboters sind kontextorientierte Bewegungen, d.h. Bewegungen, die eine bestimmte Aufgabe in einem Bewegungskontext lösen, wobei die Trajektorie so menschenähnlich wie möglich aussehen sollte.

Im Gegensatz zur klassischen Methode mit drei Komponenten besteht unser Ansatz zur Erzeugung kontextorientierter Bewegungen aus vier Teilen (Boesnach et al., 2006). Zunächst muss der Bewegungskontext (Aufgabe, Pose und Stil der Bewegung) definiert werden. Anhand des Bewegungskontextes erzeugt der Bewegungsgenerator eine Menge von Trajektorien, welche die Aufgabe erfüllen. Jede Trajektorie kann daher vom Roboter benutzt werden, aber es gibt keine Information darüber, welche Trajektorie ein Mensch am ehesten ausführen würde. Zu diesem Zweck wird die Menge an Trajektorien vom Bewegungsklassifikator verarbeitet. Die Klassifikation der Menge von Bewegungen bezüglich eines Bewegungskontextes ist der wichtigste Teil

für die Generierung einer kontextorientierten Bewegung. Wir benutzen die im vorigen Kapitel beschriebenen HMMs, um diese Trajektorien zu klassifizieren.

Autoren

THORSTEN STEIN, ANDREAS FISCHER U. PROF. DR. HERMANN SCHWAMEDER: Universität Karlsruhe (TH), Institut für Sport und Sportwissenschaft, {stein, fischer, schwameder}@sport.uka.de

INGO BOESNACH, HILDEGARD KÖHLER U. DIRK GEHRIG: Universität Karlsruhe (TH), Institut für Algorithmen u. Kognitive Systeme, {boesnach, koehler, dgehrig}@ira.uka.de

Danksagung

Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft im Rahmen der Projekte M2 und M3 innerhalb des Sonderforschungsbereichs 588 „Humanoide Roboter – Lernende und kooperierende multimodale Roboter“ unterstützt.

Literatur

- Asfour, T. (2003). *Sensomotorische Bewegungskoordination zur Handlungsausführung eines humanoiden Roboters*. Herdecke: GCA-Verlag.
- Becher, R., Steinhaus, P. & Dillmann, R. (2004). The Collaborative Research Center 588: "Humanoid Robots - Learning and Cooperating Multimodal Robots". *International Journal on Humanoid Robotics* 1 (3), 429–448.
- Bernstein, N.A. (1967). *The Coordination and Regulation of Movements*. Oxford: Pergamon Press.
- Beth T., Boesnach I., Haimerl M., Moldenhauer J., Bös K. & Wank V. (2003). Characteristics in Human Motion – From Acquisition to Analysis. *Proceedings of the IEEE International Conference on Humanoid Robots*, 56-74.
- Boesnach I., Moldenhauer J., Fischer A. & Stein T. (2006). Context Oriented Motion Generation: A New Scheme for Humanoid Robot Control. *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication*, 304-308.
- Cappozzo, A., Croce, U., Leardini, A. & Chiari, L. (2005). Human Movement Analysis Using Stereophotogrammetry: Part 1: Theoretical Background. *Gait and Posture* 21, 186-196.

- Goodman, S.R. & Gottlieb, G.L. (1995). Analysis of kinematic invariances of multijoint reaching movement. *Biological Cybernetics* 73, 311-322.
- Khatib, O., Sentis, L., Park, J. & Warren, J. (2004). Whole body dynamic behaviour and control of human-like robots. *International Journal of Humanoid Robotics* 1 (1), 29-43.
- Moldenhauer J., Boesnach I., Beth T., Wank V. and Bös K. (2005). Analysis of Human Motions for Humanoid Robots. *Proceedings of the IEEE International Conference on Robotics and Automation*, 312-317.
- Moldenhauer J., Boesnach I., Stein T. and Fischer A. (2006). Composition of Complex Motion Models from Elementary Human Motions. *Proceedings of the IV Conference on Articulated Motion and Deformable Objects*, 68-77.
- Rabiner, L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE* 77 (2), 257-286.
- Slater-Hammel, A. T. (1960). Reliability, accuracy and refractoriness of a transit reaction. *Research Quarterly* 31, 217-228.
- Stelzner, G., Simonidis, C., Seemann, W., Boesnach, I., Köhler, H., Gehrig, D., Stein, T., Fischer, A. & Schwameder, H. (2007). Analysis and Modeling of Human Motions for Human-Robot Interaction (eingereicht bei IEEE Trans. Robotics)
- Tanie, K. (2003). Humanoid Robot and its Application Possibility. *Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, 213-214.
- Turvey, M.T. (1990). Coordination. *American Psychologist* 45, 938-953.
- Wiemeyer, J. (1994). Perspektiven der Motorikforschung: Kritische Reflexionen zentraler Positionen und ausgewählter Probleme der psychologisch orientierten Motorikforschung. *Spectrum der Sportwissenschaften* 6 (1), 5-26.
- Woltring, H.J. (1986). A FORTRAN package for generalized, cross-validatory spline smoothing and differentiation. *Advances in Engineering Software* 8, 104-113.