

Bachelorarbeit

Identifikation, Bewertung und Umsetzung von Körpergesten zur Steuerung eines barrierefreien Assistenzroboters

eingereicht von

Vincenz Arlt

geboren am 02.02.1995 in Dresden

Technische Universität Dresden

Fakultät Informatik

Institut für Angewandte Informatik

Lehrstuhl Mensch-Computer-Interaktion



Betreuer:

David Gollasch, M. Sc.

Hochschullehrer:

Prof. Dr. rer. nat. habil. Gerhard Weber

Eingereicht am 27. Dezember 2018

Aufgabenstellung für die Bachelorarbeit

Name des Studenten: **Vincenz Arlt**
Immatrikulationsnummer: **4081473**

Thema:

Identifikation, Bewertung und Umsetzung von Körpergesten zur Steuerung eines barrierefreien Assistenzroboters

Identification, Analysis and Realisation of Body Gestures for Controlling an Accessible Assistive Robot

Zielstellung

Kontext. Verschiedentliche Motivationen begründen einen sich erhaltenden Bedarf an smarterer, assistiver Technologie in Form autonomer Assistenzroboter. Ein gesellschaftlicher Druck geht dabei besonders von der alternden Bevölkerung in Verbindung mit dem angespannten Pflegesektor aus. Hier sind die Hoffnungen groß, das Leben von alten Menschen sowie von Pflegekräften deutlich zu erleichtern, indem Assistenzroboter zukünftig ein breites Spektrum von Unterstützung in unterschiedlichen Alltagssituationen bieten können. Zwar sind die Entwicklungen in der Robotik rasant schnell, aber reale Anwendungsfälle beschränken sich bislang eher auf industrielle Fertigungsroboter. Ein Knackpunkt im Bereich der Assistenzroboter ist der Mangel an Funktionsvielfalt, welche einen Roboter überhaupt erst interessant macht.

Ein Lösungsansatz ist hier das Bereitstellen einer gemeinsamen, erweiterbaren Entwicklungsplattform. Der Segway Robotics „Loomo“ bietet hierfür eine Basis bestehend aus Hardware- und Softwareplattform. Die Hardware besteht aus einem Segway/Ninebot Self-Balancing Vehicle in Kombination mit einer Intel-Atom-basierten Recheneinheit und verschiedentlichen Sensoren und Aktuatoren zur Wahrnehmung und Interaktion mit der Umgebung. Die Softwareplattform bildet Android in Verbindung mit einem SDK zur Ansteuerung der Sensorik und Aktorik. Die Implementierung von Funktionen für spezifische Anwendungsfälle erfolgt in Form von Android-Apps mit Zugriff zum Steuerungs-SDK. Weiterhin erlaubt Loomo auch die Erweiterung der Hardware mittels Erweiterungskupplung bestehend aus belastbarer Metallaufhängung, USB-Anbindung und zusätzlicher Stromversorgung.

(Fortsetzung Rückseite)

Betreuer: David Gollasch, M.Sc.
Beginn am: 12.10.2018
Einzureichen am: 28.12.2018 (11 Wochen)

Dresden, 08.10.2018

Prof. Dr. rer. nat. habil. Gerhard Weber
verantwortlicher Hochschullehrer

Fortsetzung Zielstellung

Projektziel. Eine wichtige Problemstellung, um einen nützlichen Assistenzroboter mit Funktionalitäten auszustatten, ist die Unterstützung unterschiedlicher Eingabemodalitäten zur Steuerung des Systems. Dies wird im Sinne der Barrierefreiheit wie auch der intuitiven Bedienung benötigt. Neben der Implementierung einer Sprachsteuerung wirkt hierbei die Umsetzung einer Steuerung mittels Körpergesten vielversprechend und sinnvoll. Dabei ist von Interesse, welche Steuerungsbefehle umgesetzt werden sollten, welche Gesten sich wann für den Einsatz eignen und wie zuverlässig diese Eingabemodalität funktioniert. Die Zuverlässigkeit meint dabei die Erkennungsgenauigkeit des Systems einerseits, aber auch die Praktikabilität und Benutzbarkeit durch den Anwender andererseits. Hierfür muss das geplante System prototypisch umgesetzt und überprüft werden.

Schwerpunkte

- Einarbeitung in die folgenden Themengebiete inkl. Analyse des aktuellen Forschungsstandes:
 - Entwicklung von Roboter-Apps auf Basis von Java und Android für Segway Robotics Loomo
 - Körpergesten als Eingabemodalität im Kontext von Assistenzrobotern
 - Einsatz unterschiedlicher Sensoren zur Gestenerkennung (Kamera, 3D-Kamera)¹
- Umsetzung eines vielversprechenden Konzeptes zur Verwendung von Körpergesten zur Steuerung eines Roboters.
 - Identifikation benötigter Steuerbefehle des Roboters
 - Identifikation geeigneter Gesten; Unterscheidbarkeit von Gesten
 - Analyse der Erkennungsgenauigkeit der Gesten mittels der unterschiedlichen Sensoren des Roboters
 - Analyse der Benutzbarkeit der Gesten auf Anwenderseite unter Berücksichtigung besonderer Bedürfnisse und Präferenzen
 - Zusammenfassung der Ergebnisse zu einem Konzept zur Körpergestensteuerung sowie Bewertung des Konzepts
- Prototypische Implementierung und Test einer Softwarelösung für den beschriebenen Roboter (Loomo)
- Evaluation und Auswertung des erarbeiteten Verfahrens auf angemessene, wissenschaftliche Weise
- Dokumentation der Ergebnisse in geeigneter, wissenschaftlicher Form

¹ Blank, C. (2014). *Gesten im dreidimensionalen Raum*. Hamburg, Germany.

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit mit dem Titel *Identifikation, Bewertung und Umsetzung von Körpergesten zur Steuerung eines barrierefreien Assistenzroboters* selbständig unter Angabe aller Zitate angefertigt und dabei ausschließlich die aufgeführte Literatur und genannten Hilfsmittel verwendet habe.

Dresden, 27. Dezember 2018

Vincenz Arlt

Kurzfassung

Es wird eine auf Körpergesten basierende Steuerung für den Segway Robotics Loomo vorgestellt. Dabei wird zuerst analysiert, welche Möglichkeiten die Sensorik des Roboters und Gesten als Steuerungskonzept bieten. Auf dieser Grundlage wird ein Gestensystem aus statischen Handpositionen entworfen, das die Funktionalitäten des Roboters dem Nutzer über Gesten zugänglich macht. Die Implementierung auf dem Loomo mithilfe von Android und OpenCV wird danach beschrieben. Zum Test der Anwendung wurden Nutzer durch einen Hindernisparcour mit dem Roboter geführt. Es hat sich gezeigt, dass die entwickelte Körpergestensteuerung eine geeignete Möglichkeit darstellt, mit einem Roboter zu interagieren.

Abstract

A control scheme based on body gestures for the segway robotics loomo is presented. First the possibilities of Loomos sensors and controlling him through gestures are analyzed. Based on this, a gesture system consisting of static hand positions is created, which enables the user to utilize the robots features. Next, the implementation with Android and usage of OpenCV is described. To test the application, users have been send through a obstacle course to use Loomo. It has been shown that the developed gesture control system enables to successfully interact with a robot.

Inhaltsverzeichnis

Kurzfassung	I
1 Einleitung	1
2 Grundlegende Konzepte und Arbeiten zum Umgang mit Assistenzrobotern	3
2.1 Einführung in Assistenzroboter und Gestenerkennung	3
2.2 Verwandte Arbeiten	5
2.2.1 Verwandte Forschungsansätze	5
2.2.2 Aktueller Forschungsstand verwandter Techniken	6
2.3 Beschreibung des Segway Robotics Loomo	8
2.4 Weitere gängige Roboter	8
2.5 Zusammenfassung	9
3 Analyse der vorgestellten Roboter- und Gestenkonzepte für die Entwicklung einer umfassenden Gestensteuerung eines Assistenzroboters	11
3.1 Einordnung des Segway Robotics Loomo in den Kontext anderer Assistenzroboter	11
3.2 Typische Aufgaben und Ziele in Bezug auf den Loomo Roboter	12
3.3 Gestenanalyse	13
3.4 Zusammenfassung der aus der Analyse gewonnenen Erkenntnisse	13
4 Konzeptionierung einer Gestensteuerung für den Segway Robotics Loomo	15
4.1 Sensorische Konzeptionierung bei Umgang mit dem Loomo	15
4.2 Befehlsauswahl anhand der betrachteten Anwendungsszenarien	17
4.3 Genaue Gestenauswahl und Gestenerkennung	17
4.4 Geplante Methodik der Mensch-Roboter-Interaktion	22
4.5 Zusammenfassung des Gestenkonzepts	22
5 Prototypische Implementierung der Gestensteuerung	25
5.1 Technische Grundlagen	25
5.2 Programmaufbau	25
5.3 Zusammenfassung	27
6 Evaluation der entwickelten Steuerung	29
6.1 Planung des praktischen Steuerungstests	29
6.2 Durchführung der Evaluation	30
6.3 Gewonnene Werte aus der Evaluation	31
6.4 Zusammenfassung des Nutzertests	31
7 Abschluss	33
7.1 Zusammenfassung des entwickelten Systems	33

7.2	Diskussion der erlangten Ergebnisse	33
7.3	Ausblick auf weitere Möglichkeiten zur Verwendung von Gestensteuerungen	34
7.4	Fazit	35
Abkürzungsverzeichnis		i
Abbildungsverzeichnis		iii
Tabellenverzeichnis		v
Literatur		vii

1 Einleitung

In der heutigen Gesellschaft ist der Bedarf an Pflege und Assistenz stetig steigend. Viele Menschen benötigen Hilfe und Unterstützung im Alltag, sei es gelegentlich oder ganztägig. Daher sind Assistenzroboter eine betrachtenswerte Möglichkeit um hier Versorgungslücken zu schließen. Sie können einige der alltäglichen Aufgaben übernehmen, für die sonst menschliche Arbeitskraft erforderlich wäre. Diese menschliche Arbeitsleistung kann dann für schwierigere, wichtigere, oder kurz, wertvollere Tätigkeiten aufgebracht werden.

Mit diesen Zielen für Assistenzroboter gehen aber auch bestimmte Anforderungen einher. Sie müssen dabei vor allem zugänglich sein, um von den Nutzern mit ihren speziellen Bedürfnissen überhaupt erst eingesetzt werden zu können. Die Nutzer müssen mit ihnen interagieren können, um einerseits die Roboter zu steuern. Andererseits sorgt die Interaktivität aber auch dafür, dass der Roboter ihre Bedürfnisse erkennen und erfüllen kann. Einige der vorhandenen Interaktionstechniken, wie etwa Sprache, Touchoberflächen oder Wearables, lassen dabei einige Wünsche offen. Bei einer Sprachsteuerung wird es schwierig, gleichzeitig ein Gespräch mit seinen Mitmenschen und dem Roboter zu halten. Muss der Roboter gezielt berührt werden, um ihn zu steuern, ist seine Reichweite eingeschränkt und seine Bewegungsfähigkeit vom Anwender abhängig. Bei am Körper getragenen Steuerelementen müssen diese stets gehalten werden und in der Nähe sein. Zudem müssen diese Steuergeräte für jeden einzelnen Nutzer des Roboters zur Verfügung stehen.

Eine alternative Steuerungsmöglichkeit soll hier vorgestellt und analysiert werden: eine Assistenzrobotersteuerung über Gesten. Dies würde die vorher beschriebenen Probleme lösen. Bei einer Interaktion über Gesten lassen sich die Roboter mit weniger Aufwand oder Ablenkung von anderen Tätigkeiten steuern. Dabei stellen sich aber Fragen darüber, welche Befehle an den Roboter über Gesten sinnvoll umsetzbar sind. Außerdem sind die genauen Gesten auszuwählen. Nur wenn diese für den Nutzer anwendbar sind und vom Roboter gut erkannt werden können, ist eine Gestensteuerung verwendbar. Zusammengefasst lässt sich so die Forschungsfrage wie folgt formulieren:

Wie sollte eine Körpergestensteuerung gestaltet sein, um in unterschiedlichen Szenarien eine geeignete Steuerung des Roboters zu erlauben?

Dabei entstehen Teilfragen, die zur Beantwortung dieser Frage zuerst geklärt werden müssen.

1. Welche Möglichkeiten hat der Assistenzroboter in Bezug auf Sensoren und Freiheitsgraden den Nutzer und seine Umwelt (visuell) wahrzunehmen und mit ihm zu interagieren? Welche Einschränkungen sind zu berücksichtigen?
2. In welchen Anwendungsszenarien kann der Roboter eingesetzt werden, um mittels seiner Interaktionsmöglichkeiten die in den Szenarien und ihrem Umfeld typischen Aufgaben sinnvoll zu erfüllen?

3. Welche Gesten sollen unter Beachtung der Sensorik des Roboters und der Nutzbarkeit für den Anwender ausgewählt werden?
4. Welche Aufgaben sollen den Gesten zugewiesen werden, so dass der Nutzer auf das nutzbare Potential an Aktionen des Roboters zugreifen kann?

Wurden die Teilfragen gelöst, so entsteht eine zur Steuerung des betrachteten Roboters geeignete Auswahl an Körpergesten. In dieser Arbeit soll auf diese Art und Weise eine Gestensteuerung für einen Roboter erdacht und umgesetzt werden. Es wird analysiert, welche Fortschritte es bei Assistenzrobotern und der Gestenerkennung gibt. Aus den daraus gewonnenen Erkenntnissen wird im Anschluss eine Auswahl an geeigneten Gesten und Befehlen getroffen. Die Befehle werden prototypisch mithilfe des Segway Robotics Loomo umgesetzt, einem auf einem Segway basierenden Roboter mit Android-Betriebssystem. Zum Schluss wird die Implementierung auf ihre Praktikabilität als Gestensteuerung getestet.

Diese Arbeit soll die beschriebene Entwicklung eingrenzen und dokumentieren. Es werden zuerst in Kapitel 2 die grundlegenden Konzepte vorgestellt, etwa wie eine Gestensteuerung in der Theorie aufgebaut ist. Im Anschluss folgen verwandte Arbeiten, in denen bereits an durch Gesten gesteuerten Robotern gearbeitet wurde. Zudem wird in diesem Abschnitt auch der aktuelle Forschungsstand aufgegriffen. Dabei ist es wichtig, die aktuell verwendeten Roboter, ihre Interaktionsmethoden und eingebaute Technik zu betrachten. Auch der verwendete Roboter wird in diesem Kapitel beschrieben.

Im Anschluss findet in Kapitel 3 die Analyse der erfassten Arbeiten statt. Die gesammelten Informationen sollen so verdichtet werden, dass eine Basis für ein Steuerungskonzept entsteht. Hier soll unter anderem auf die von Robotern erfüllbaren Aufgaben, bereits bestehende Steuerungssysteme und Besonderheiten in der Verwendung von Gesten eingegangen werden. Auch soll hier der Loomo mit bereits etablierten Robotern verglichen werden.

Damit kann das Konzept für die Gestensteuerung im Kapitel 4 erstellt werden. Aus den Möglichkeiten des Loomo und den analysierten Voraussetzungen wird ein neues Steuerungssystem entwickelt. Dieses besteht aus Befehlen zur Bedienung des Roboters und den jeweils den Befehlen zugeordneten Gesten. Hier soll auch erklärt werden, warum bestimmte Gesten ausgewählt wurden und wie über die Zuordnungen von Gesten und Befehlen an den Roboter entschieden wurde.

Als nächstes wird die prototypische Implementierung dieses Konzepts in Kapitel 5 erläutert. Es soll deutlich werden, wie das Programm aufgebaut ist. Es wird Wert darauf gelegt, wie die Gesten erfasst werden und wie die Befehlszuweisung im Detail funktioniert.

Bei der Evaluation in Kapitel 6 des Prototyps wird untersucht, wie gut das umgesetzte Konzept die Anforderungen erfüllt hat. Dabei sind Genauigkeit der Erkennung und Eingängigkeit der Gesten besonders wichtig. Hierfür soll ein Nutzertest durchgeführt und ausgewertet werden, der die entwickelte Steuerung praktisch zeigen soll.

Als Abschluss findet in Kapitel 7 eine Zusammenfassung der bisher erfolgten Schritte statt. Das System sowie seine Nutzbarkeit werden noch einmal in kurzer Form beschrieben. Darauf folgt eine Diskussion der Ergebnisse der Arbeit. Es soll klar werden, wie gut die entwickelte Steuerung die Aufgabenstellung erfüllt hat. Wo gab es Schwierigkeiten, wo zeigen sich besonders gelungene Aspekte? Wie konnte die Forschungsfrage beantwortet werden? Es folgt nach dem Ausblick auf weitere Forschungsfragen und -möglichkeiten ein Fazit der Arbeit und des entwickelten Prototypen.

2 Grundlegende Konzepte und Arbeiten zum Umgang mit Assistenzrobotern

Zu Beginn soll ein Überblick über die Grundlagen gegeben werden auf denen diese Arbeit basiert. Diese sind die Mensch-Roboter-Interaktion im Kontext der Assistenzroboter und dabei insbesondere die Verwendung von Gesten.

2.1 Einführung in Assistenzroboter und Gestenerkennung

An dieser Stelle soll eine kurze Einführung in die wichtigsten Konzepte erfolgen, auf denen die weitere Arbeit beruht. Diese sind die Assistenzroboter, ihre Einsatzmöglichkeiten und Fähigkeiten, sowie die Gestensteuerung von der Aufnahme über verschiedene Sensoren bis hin zur Ausführung.

Assistenzroboter stellen eine Form der Medizinroboter dar. Ihre Aufgabe ist es im Allgemeinen Assistenzleistungen, etwa in der Pflege, zu bieten. Darüber hinaus sind sie aber vielseitig einsetzbar. Ihre genauen Möglichkeiten hängen stark von ihrer Form und ihrem Einsatzort ab. Ein paar Möglichkeiten zur Unterteilung, wie sie auch in der Abbildung dargestellt werden, sind zum Beispiel:

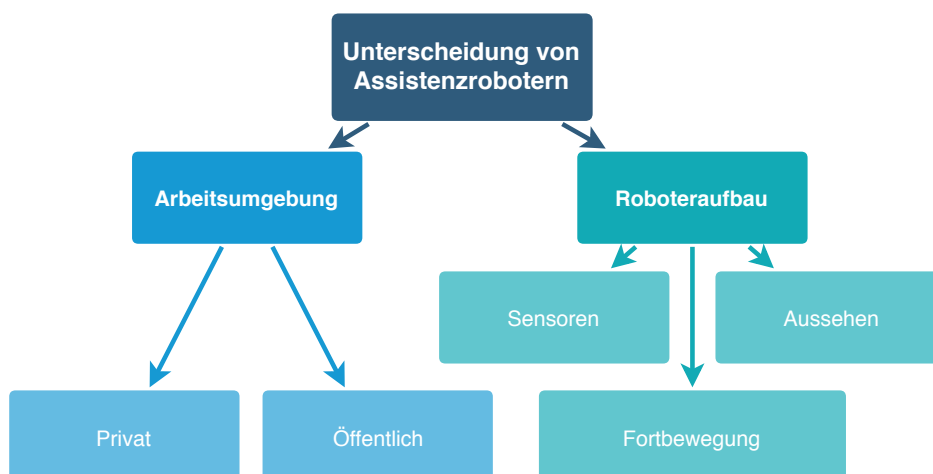


Abbildung 2.1 – mögliche Roboterunterscheidungen

Arbeitsumgebung Hier kann besonders zwischen privatem Gebrauch und Einsatz in öffentlichen Gebäuden getrennt werden. Wird der Roboter privat eingesetzt, etwa in der Wohnung des Nutzers, kann die Assistenz stark auf ihn spezialisiert werden. Dabei kann der Roboter und seine Aufgaben auf die Besonderheiten des Nutzers abgestimmt werden. Er kann etwa demente Nutzer

an Dinge erinnern [Beg+13] oder mit Kindern mit Autismus interagieren [Sha+12]. Auch in Gesundheitsfragen kann ein Assistenzroboter den Nutzer unterstützen [Gro+15].

Im öffentlichen Raum dagegen wird der Roboter meist dazu verwendet, um bei wechselnden, kurzzeitigen Nutzern Orientierungshilfen zu bieten. Eine Individualisierung ist schlecht möglich, da der Roboter feste Aufgaben für im voraus nicht bekannte Nutzer erfüllen soll. Mehrfach wurden beispielsweise Roboter als Museumsführer eingesetzt, etwa der etwas ältere Minerva [Thr+99]. Auch andere Roboter wie Pepper sind in Museen einsetzbar [All+18]. Auch der Einsatz in Krankenhäusern fällt in diese Kategorie [MSW17].

Roboteraufbau Hier hängt es individuell von der im Roboter vorhandenen Technik ab, wie er eingeordnet werden kann. Manche, wie der Segway Robotics Lomo, sind komplett eigenständig und besitzen neben Antrieb und Motoren auch weitere Hilfsmittel, Kameras oder Sensoren. Andere hingegen können auch nur aus einer an einem Rollstuhl angebrachten Oberfläche für eine Kamera bestehen, um diesen über Handgesten zu bewegen [ZSW12]. Häufiger sind die Roboter aber menschenähnlich gestaltet, verfügen etwa über einen Körper, Kopf mit Gesicht und Gliedmaßen. Dies trifft unter anderem auf Pepper [Rob18c] und Nao [Rob18b] zu.

Gesten sind, nach Definition von Mitra et al. [MA07], gezielt ausgeführte Körperbewegungen mit der Aufgabe, Informationen zu übertragen oder mit der Umgebung zu interagieren. Dies reicht von einem Winken als Begrüßung, Zeigehandlungen als Richtungsangaben bis hin zu einer ganzen Zeichensprache.

Für die Gestenerkennung werden üblicherweise Videokameras und Sensoren eingesetzt, um die Geste visuell zu erfassen. Das dabei eingesetzte Vorgehen kann in einer dreistufigen Pipeline ausgedrückt werden [Cha+18], die im folgenden detaillierter beschrieben wird.

Signalverarbeitung: Das eingehende Signal wird ausgewertet und die Geste herausgefiltert. Dabei findet auch eine Vorverarbeitung statt, indem etwa Hintergrund, Beleuchtung und Verdeckungen bestimmt werden.

Filterung und Segmentierung: Es wird ein im voraus gewähltes Modell eingesetzt, das dazu dient, die eingehende Geste zu repräsentieren. Wichtige Parameter dabei können bestimmte Formen, Positionen oder Bewegungen sein. Für zweidimensionale Gesten sind farbbasierte oder Silhouettenmodelle üblich, für dreidimensionale dagegen beispielsweise Gitternetze oder Skelette.

Gestenerkennung: Die modellierte Geste wird klassifiziert. Dabei werden die erfassten Merkmale dazu eingesetzt, die erkannte Geste einer Klasse zuzuordnen. Anhand dieser Klasse ist es dann möglich, der Geste eine mögliche Aktion zuzuweisen, um mit dem gestenaufnehmenden Roboter zu interagieren.

Es kann zwischen zwei- und dreidimensionalen Gesten unterschieden werden [Cha+18]. Zweidimensionale Gesten benötigen keine Tiefeninformationen und können auch mit 2D-Kameras erfasst werden. Mithilfe einer Stereokamera (häufig wird beispielsweise eine Microsoft Kinect verwendet) lassen sich auch dreidimensionale Gesten erfassen. Manche Systeme setzen auch eine virtuelle Ebene ein, auf der die Gesten zweidimensional erfasst werden.

Auch der erfasste Körperbereich muss beachtet werden. Eine Möglichkeit der Unterteilung ist zum Beispiel die in Handgesten, Körpergesten oder Kopfgesten und Mimik [MA07], die hier kurz

vorgestellt wird. Die meisten menschlichen Gesten werden mit den Händen ausgeführt. Aufgrund der Komplexität und den Freiheitsgraden von Handbewegungen sind diese aber schwieriger zu erfassen. Sie erfordern eine genaue, hochauflösende Kamera sowie ein komplexeres Modell. Weitere Gesten und Bewegungen des Körpers werden unter Körpergesten zusammengefasst. Ihre Aufnahme erfordert es, den gesamten Körper mit der Kamera abbilden zu können. Auch Kopfbewegungen und Mimik können als Gesten erkannt und eingesetzt werden. Ihre Aufnahme erfordert eine höhere Auflösung, damit auch feinere Bewegungen in der Gesichtsmuskulatur erkennbar und auswertbar werden.

Wie diese Grundlagen in aktuellen Forschungsarbeiten verwendet werden und sich in ähnlichen Forschungsfragen wiederfinden lassen, wird im folgendem beschrieben.

2.2 Verwandte Arbeiten

An dieser Stelle soll auf Arbeiten mit verwandter Forschungsfrage oder auf solche mit ähnlichen Techniken und Arbeitsschritten eingegangen werden. Arbeiten mit einem verwandtem Ansatz beschäftigen sich mit der Gestensteuerung eines Roboters und gehen dabei auf die verwendeten Roboter, insbesondere auf die nötigen Sensoren, und die Implementierung der Gesten ein. Daher werden sie hier gesondert betrachtet. In weiteren Arbeiten werden gezielt bestimmte Aspekte wie Aufgabenbereiche von Robotern oder Gestensteuerungen beleuchtet. An dieser Stelle werden auch kurz verschiedene Robotermodelle, unter ihnen der Segway Robotics Lomo, vorgestellt.

2.2.1 Verwandte Forschungsansätze

Ein beliebtes System zur Erkennung von Gesten stellt die Kinect dar. Mit verschiedenen Sensoren für Farbwerte, Tiefen und Infrarot sowie einer eigenständigen skelettbasierten Modellierung der Nutzer ist sie auch für die Interaktion mit Robotern geeignet. Wie diese spezifisch in Innenräumen ablaufen kann, wurde von Zhang et al. beschrieben [Qi+14]. Hierfür haben sie einen mobilen Roboter mit Kinect 2.0 ausgestattet. Zur Steuerung hat der Nutzer bestimmte Handbewegungen durchzuführen: Hochhalten, nach links und nach rechts bewegen. Unter Verwendung beider Hände wurden damit sechs Befehle realisiert, das Starten und Stoppen des Roboters sowie die Bewegung in vier Richtungen(vor, zurück, links, rechts). Zur Verfolgung des Nutzers durch den Roboter greifen sie auf die Schnittstelle der Kinect zurück.

Cheng et al. haben sich ebenfalls mit der Mensch-Roboter-Interaktion mithilfe von Kinect auseinandergesetzt [Liy+12]. Dafür griffen sie auf den humanoiden Roboter Nao zurück. Er wird unter anderem als Fußball spielender Roboter eingesetzt und verfügt über die dafür geeignete Beweglichkeit, Freiheitsgrade und Sensoren. Die von der Kinect erkannten Gesten werden drahtlos an den verwendeten Nao übertragen. Die Benutzung von Kinect hat dabei einige Einschränkungen aufgeworfen. So darf die Distanz zur Kinect 0,8 Meter nicht unterschreiten und der Nutzer muss gerade vor der Kinect stehen. Cheng et al. haben dabei auf Körpergesten zurückgegriffen, die den ganzen Körper ausnutzen: vom Winken mit der linken Hand zum Zurückwinken des Roboters, über Kniebeugen, damit sich der Roboter setzt, bis zum Anheben der Knie, um eine Laufbewegung zu imitieren und beim Roboter zu initiieren.

Die Arbeiten zeigen, dass es mithilfe der Kinect und der durch sie möglichen Skelettmodellierung gut möglich ist, Gesten zu erkennen. Besonders ist bei Zhang et al., wie die Gesten durch die asymmetrische Verwendung der Hände realisiert werden. So werden mit drei Gesten durch

Erfassung von zwei ausführenden Gliedmaßen sechs Aktionen realisiert. Cheng et al. dagegen zeigen den Versuch, die Gesten intuitiv mit der auszuführenden Aktion zu verbinden. Ihr Roboter spiegelt die ausgeführte Geste wieder. Winkt der Nutzer, winkt der Roboter zurück, geht er in die Knie, setzt er sich.

2.2.2 Aktueller Forschungsstand verwandter Techniken

Des Weiteren gibt es in der aktuellen Forschung andere Arbeiten, unter denen sich die hier vorliegende Arbeit einordnen soll. Wie diese miteinander in Verbindung stehen, wird anhand von verschiedenen Teilaspekten der Roboterinteraktion festgelegt.

Die Einsatzbereiche der Roboter sind von Interesse, um die Aufgaben bestimmen zu können, die über eine gestengesteuerte Interaktion gelöst werden sollen.

Eine Nutzergruppe, die in den eigenen vier Wänden Unterstützung gebrauchen kann, sind Menschen mit Demenz. Ein Roboter, der ihnen als Gedächtnisstütze und Hilfe bei alltäglichen Aufgaben dienen soll, wurde von Begum et al. entworfen [Beg+13]. Dafür haben sie sich auf bestimmte Aktivitäten des täglichen Lebens konzentriert. Diese wurden von ihnen in Teilaufgaben zerlegt, worauf für jede Teilaufgabe entsprechende Hilfestellungen generiert wurden. Der Roboter besteht hier aus einer mobilen Plattform mit angebrachtem Display sowie Lautsprechern, Mikrofon und Kameras. Um den demenzkranken Nutzer zu unterstützen, beobachtet der Roboter ihn bei der Durchführung der Alltagsaufgaben. Kommt er nicht weiter oder fragt er direkt den Roboter, hilft dieser mit Hinweisen. Beispielsweise zeigt er beim Tee kochen dorthin, wo sich die Teebeutel befinden, wenn der Nutzer diese nicht findet. Reichen die Hilfen nicht aus, werden sie spezifischer und reichen bis zu einem kommentierten Videoclip der durchzuführenden Aufgabe.

Ein selbstständiger Roboter wird von Yamazaki et al. gezeigt [Yam+12]. Sie geben ihm dabei Aufgaben wie das eigenständige Reinigen des Bodens oder das Verschieben von Tischen und Stühlen. Für ihren Roboter haben sie sich dabei auf die Umsetzung von zwei grundlegenden Funktionen konzentriert, das Halten von Objekten sowie das Bedienen dieser. Weiter entwickelt, wäre ihr Roboter als Haushaltshilfe verwendbar, um dem Nutzer einige körperlichen Arbeiten abzunehmen.

Kinder haben ebenfalls besondere Ansprüche, können aber im Vergleich zu Erwachsenen oft unvoreingenommener an Neues herangeführt werden. Ein für Kinder wichtiger Kontext ist die Schule und das Lernen. Wie sich dabei ein Roboter nutzbringend einsetzen lässt, wurde von Tanaka et al. gezeigt [Tan+15]. Sie verwenden den menschenähnlichen Pepper Roboter als Unterstützung im Lernprozess von Kindern. Dabei lernt der Roboter gemeinsam und durch die Kinder den Lehrstoff nach dem Care Receiving Robot (CRR) Prinzip. Entworfen wurde er für das Erlernen der englischen Sprache von vier- bis fünfjährigen japanischen Kindern. Die Verwendung von Pepper als CRR setzt einen begleitenden Lehrer voraus. Tanaka et al. haben drei Lernprogramme entworfen: Farbe, Körper und ein praktisches Spiel. Beim Farbprogramm fragt Pepper beispielsweise danach, ihm rot beizubringen, indem ihm rote Objekte gezeigt werden und das englische Wort dafür gesagt wird. Beim Körperprogramm werden Körperteile am Roboter berührt, um einen ähnlichen Lerneffekt zu erzielen. Beim praktischen Programm wurden Wörter wie Flugzeug verwendet, wobei die Kinder im Anschluss mit gehobenen Armen wie ein Flugzeug herum toben sollten. Bei allen Programmen konnte über Sprache, Berührungen und das bei Pepper verwendete Tablet mit ihm interagiert werden.

Ein anderer spezieller Anwendungskontext ist die Unterstützung von Kindern mit Autismus,

dem sich Shamsuddin et al. zugewandt haben [Sha+12] und der im folgenden beschrieben wird. Hierbei wurde auf das freundliche und niedlich gestaltete Äußere des Roboters Nao zurückgegriffen. Damit soll es ermöglicht werden, den Kindern die Interaktion mit dem Roboter so leicht wie möglich zu machen. In einem Experiment mit einem verborgenen Kontrolleur wurde der Roboter im Umgang mit Kindern erprobt. Dabei wurde in mehreren Phasen gearbeitet, jeweils durch kurze Pausen getrennt. Dabei wurden die Aktivitäten des Roboters so behutsam wie möglich eingeführt. Auf Inaktivität folgen nacheinander Kopfbewegungen, Aufleuchten der Augen, Sprache, Musik und Handbewegungen.

Ein anderer Einsatzkontext im öffentlichen Raum wird von Allegra et al. behandelt [All+18]. Sie beschreiben die Planung zum Einsatz des Pepper-Roboters als Museumsführer. Er soll Besuchern folgen, die Exponate beschreiben, weiterführende Hinweise zu kulturellen Angeboten geben und Aufgaben eines Museumsmitarbeiters übernehmen können. Dafür haben sie ebenfalls Pepper verwendet und ihn mit einer Anwendung ausgestattet, die die Basisaufgaben implementiert. Diese sind das Begrüßen und Erkennen von Besuchern, das Planen und Organisieren von Führungen sowie das Wiedergeben von Informationen über Sprache, Bilder oder Videos.

Neben den Einsatzgebieten für die Roboter ist auch wichtig, wie in diesen die Interaktion genau umgesetzt wird. Dabei sind verschiedene Steuerungsmodelle und -techniken verwendbar. Im Kontext dieser Arbeit sind vor allem Steuerungen mittels Gesten relevant.

Eine weitere Lösung zur Gestensteuerung mittels Kinect, die hier erläutert werden soll, wird von Lai et al. geliefert [Lai+16]. Sie konzentrieren sich dabei auf Zeigegesten. Über die Sensoren der Kinect wird zuerst das Skelettmodell des Nutzers erfasst und aufgebaut. Um im Anschluss die gewünschten Zeigegesten von anderen Bewegungen zu trennen, geben sie bestimmte Regeln für ihre Gesten an. Sie beschränken sich darauf, nur dann eine Zeigegeste zu registrieren, wenn sich die Hand auf Schulterhöhe befindet. Eine Besonderheit ist, dass die beiden Arme verschiedene Aufgaben erhalten. Während einer der Arme die Richtung vorgibt, steuert der Nutzer mit dem anderen die Geschwindigkeit des Roboters. Dafür wird die vertikale Distanz zwischen der anderen Hand und dem Kopf verwendet. Darüber hinaus wurde von ihnen ein Test der Genauigkeit der Kinect durchgeführt. Im Testszenario baut sich der Nutzer vor dem Roboter auf und zeigt in verschiedene Richtungen. Dabei wurden Reaktionszeiten der Kinect und Richtungsfehler (Differenzen zwischen gezeigter und berechneter Richtung) in Grad betrachtet. Die durchschnittliche Reaktionszeit lag bei 0,39 Sekunden, der durchschnittliche Richtungsfehler bei einem Winkel von 3,59 Grad. In einem weiteren Anwendungstest sollte der Roboter vom Nutzer mittels Zeigegesten durch einen verwinkelten Gang geführt werden. Die minimale Anzahl an benötigten Gesten lag dabei bei 10 Gesten, da der Roboter um 10 Ecken geführt werden musste. Im Schnitt benötigten die Nutzer rund 18 Gesten, um den Roboter zu korrigieren, etwa wenn dieser gegen eine Wand gefahren war.

Wenn Handgesten verwendet werden sollen, können Probleme auftreten. Die geringe Größe der Hand und die vielen Bewegungen, die mit ihr durchgeführt werden können, können bei der korrekten Aufnahme herausfordernd sein. Wie dennoch eine Gestensteuerung mittels Handgesten umgesetzt werden kann, demonstrieren Van den Bergh et al. mithilfe der Kinect [Van+11]. Sie beginnen, nach einer Kalibrierung der Kinect-Sensoren, mit der Erkennung eines Nutzers. In einem 60 Grad Winkel vor dem Roboter wird geprüft, ob sich Hindernisse vor ihm befinden und ob diese die Form eines Beinpaars aufweisen. Ist dies der Fall, wird ein Haut- und ein Gesichtserkennungsalgorithmus verwendet, um den Nutzer zu registrieren. Für die Handerkennung wurde ein spezielles Feature der Kinect verwendet. Erkennt der Tiefensensor der Kinect ein sich bewe-

gendes Objekt (wie eine winkende Hand), kann dieses als Hand erkannt und verfolgt werden. Wird das Objekt verloren, gibt das System ein Signal, dass erneut zu winken ist. Ein weiteres Signal tritt auf, wenn erfolgreich eine Hand erkannt wurde und das Winken beendet werden kann. Diese Signale werden vom Roboter während der Interaktion dem Nutzer mitgeteilt. Die erkannte Hand wird segmentiert, indem die Hand in ein Rechteck eingepasst wird. Wo dieses den Arm schneidet, wird das Handgelenk vermutet. Aus dem Mittelpunkt des Rechtecks, der sich in der Hand befindet, und dem Handgelenkspunkt kann die Orientierung der Hand ermittelt werden. Zur Erkennung bestimmter Gesten wird ein Klassifizierer mit Schlüsselgesten trainiert. Im Anwendungsfall werden die erkannten Gesten mit einer Datenbank, bestehend aus den Schlüsselgesten, verglichen. Van den Bergh et al. haben eine geöffnete Hand, eine Zeigegeste mit ausgestrecktem Zeigefinger und eine geschlossene Hand verwendet. Für die Zeigegeste wird die Richtung dabei wie die Handorientierung berechnet und als Vektor vom Handgelenk in Richtung der Hand verwendet.

Bei verschiedenen Arbeiten, wie zum Beispiel der von Lai et al. [Lai+16], wurden einige Einschränkungen bei der Verwendung von Kinect aufgezeigt. So kann die Kinect nichts registrieren, wenn der Arm oder die Hand vom Körper verdeckt wird. Ein weiteres Problem der Kinect-Sensoren ist die Verwendung von Infrarot-Lasern. Diese erschweren den Einsatz in Außenbereichen. Außerdem schränken sie den Nutzer in der Wahl der Kleidung ein, etwa wenn diese Licht reflektiert. Zudem kann die Skelettaufnahme gestört werden, wenn sich weitere Personen zwischen dem Nutzer und der Kinect bewegen.

2.3 Beschreibung des Segway Robotics Loomo

An dieser Stelle soll konkret auf den Segway Robotics Loomo eingegangen werden [Rob18a]. Sein Einsatz bietet sich für die Entwicklung des Prototypen an, weshalb er näher beleuchtet werden soll. Der Loomo besteht aus einem Segway-Fahruntersatz mit einem im Mittelteil angebrachten Roboterkörper und -kopf. In diesen sind Sensoren sowie ein LCD Display angebracht. Über dieses kann mithilfe eines auf Android basierenden Betriebssystems auf verschiedene Apps zur Bedienung zurückgegriffen werden. Im aktuellen Stand umfassen diese unter anderem Sprachsteuerung und Fernsteuerung über Android und iOS Smartphones. Neben der Verwendungsmöglichkeit als Segway kann der Roboter eigenständige Bewegungen und Tätigkeiten ausführen.

Für Entwickler wird ein Software Development Kit (SDK) zur Verfügung gestellt. Dieses bietet Zugriff auf die Kopf- und Fahrwerkmodule sowie die vorhandenen Sensoren und Sprachsteuerung. Der Loomo verwendet zur Wahrnehmung des Nutzers die Intel Realsense für RGB- und Tiefenbilder. Des Weiteren verfügt er über eine HD-Kamera mit einem Aufnahmewinkel von 104 Grad. Zur Orientierung und Hindernisumgehung wird ein Ultraschallsensor eingesetzt. Ein besonderer Punkt ist die Erweiterbarkeit des Loomo. Über ein offenes Erweiterungssegment können weitere Bauteile via USB an den Hauptteil des Roboters angehängt werden. Damit kann der Roboter beliebig um weitere Funktionalitäten erweitert werden, besitzt standardmäßig aber keine Gliedmaßen.

2.4 Weitere gängige Roboter

Der Segway Robotics Loomo ist noch nicht offiziell veröffentlicht. Es gibt aber weitere nutzbare Assistenzroboter, deren Einsatz schon wissenschaftlich in verschiedenen Situationen erprobt

wurde. Zwei davon - die Roboter Pepper und Nao - folgen an dieser Stelle.

Der Nao ist am bekanntesten durch seinen Einsatz beim Robocup, einer Art von Roboterfußball, bei dem die Roboter mithilfe von künstlicher Intelligenz ein Sportspiel bestreiten müssen. Vertrieben wird er heute von Softbank Robotics [Rob18b].

Er besteht aus einem vollständigen Körper mit Beinen, Armen und Kopf. Mit insgesamt 25 Freiheitsgraden in der Bewegung sind auch komplexe Bewegungen möglich. Des Weiteren besitzt Nao mehrere Berührungssensoren an den Extremitäten sowie ein Sonar und Gleichgewichtssensoren zur Orientierung im Raum. Sprachsteuerung und Sprachausgabe werden durch Mikrofon und Lautsprecher realisierbar. Zur Objekt- und Personenerkennung sind zwei 2D-Kameras im Roboter eingebaut.

Als robotischer und persönlicher Begleiter wurde der Roboter Pepper von Aldebaran Robotics und Softbank konzipiert [Rob18c]. Besonders wurde hierbei auf kindliche Niedlichkeit geachtet. Pepper ist klein, besitzt überdimensionierte Augen und gibt Sprache in einer hohen, kindlichen Stimmlage wieder. In diesem Zusammenhang ist auch der Fokus auf die Erkennung von Gesichtern, Ton- und Stimmlage sowie Mimik und Gestik zu sehen. Der Roboter ist darauf ausgelegt, aus der Analyse des Menschen und seines Verhaltens dessen Emotionen zu erkennen und darauf entsprechend zu reagieren.

Pepper besteht aus einem Fahrwerk, auf dem der Roboterkörper mit Armen und Kopf aufgesetzt ist. Auf der Brust befindet sich ein Tablet zur direkten Interaktion mit dem Roboter. Er verfügt über 20 Freiheitsgrade, Berührungssensoren, LEDs und Mikrofone zur multimodalen Interaktion mit dem Nutzer. Außerdem besitzt er Infrarot- und Gleichgewichtssensoren, Sonar und sowohl 2D- als auch 3D-Kameras.

Beiden Robotern gemein ist die offene Plattform zur Programmierung. Damit ist es möglich, beliebig Anwendungen für die Roboter zu implementieren. Damit einher geht der Einsatz dieser Roboter für Forschung und Lehre, aber auch ein praktischer Einsatz in der Industrie oder im Gesundheitswesen ist möglich.

2.5 Zusammenfassung

Der Forschungsbereich der Arbeit mit und der Steuerung von (Assistenz-)robotern ist groß und vielseitig. Neben den verschiedensten Robotertypen, von kleinen Lösungen für spezielle Probleme bis hin zu fertigen, praxistauglichen Varianten wie Nao oder Pepper, sind auch die Einsatzmöglichkeiten imposant. Es sind eine Vielzahl von Anwendungen denkbar und auch umgesetzt, von eigenständigen Haushaltshilfen über Navigations- und Orientierungshelfer bis hin zu Begleitern für Menschen mit allerlei speziellen Einschränkungen, Bedürfnissen und Ansprüchen.

Eine Auflistung der Möglichkeiten allein ist daher auch nicht ausreichend, um die Forschungsfrage zur Umsetzung einer Gestensteuerung mittels des Loomo zufriedenstellend zu beantworten. Die gesichteten Arbeiten und Konzepte müssen in Bezug auf die eigene Arbeit analysiert werden, um ein zielgerichtetes Gestenkonzept entwickeln zu können.

3 Analyse der vorgestellten Roboter- und Gestenkonzepte für die Entwicklung einer umfassenden Gestensteuerung eines Assistenzroboters

Die zu entwickelnde Robotersteuerung besteht aus der Verknüpfung von zwei Faktoren: einer Auswahl von geeigneten Gesten und den ermittelten Kernaufgaben. Diese Komponenten sollen im Folgenden analysiert werden, um ein verwendbares Konzept zur Steuerung zu erhalten. Des Weiteren findet in diesem Kapitel eine Einordnung des Loomo innerhalb anderweitig eingesetzter Roboter statt, da das Steuerungskonzept mit ihm umgesetzt wird.

3.1 Einordnung des Segway Robotics Loomo in den Kontext anderer Assistenzroboter

Wie bereits erwähnt, konnte der Loomo noch wenig für wissenschaftliche Arbeiten verwendet werden, da er erst in letzter Zeit entwickelt wurde und sich in einer Testphase befindet. Daher müssen die Anwendungsmöglichkeiten dieses Roboters anhand seiner praktischen Möglichkeiten und anhand des Vergleichs mit bereits erprobten Robotern ausgelotet werden.

Der Vergleich mit anderen Robotern stützt sich hier vor allem auf die Roboter Pepper und Nao. Dies liegt vor allem an einigen Gemeinsamkeiten in Vertrieb und in der Arbeitsweise mit den Robotern. Alle drei stellen fertige Komplettlösungen dar, Roboter, die in einem bereits verwendbaren Zustand geliefert und für einfache Anwendungen bereits ohne Mehraufwand einsetzbar sind. Über offene Entwicklungsplattformen und Zugriff auf die eingebauten Sensoren können aber auch weitere Anwendungen implementiert und getestet werden.

Dabei zeigen sich im Detail aber verschiedene Schwerpunkte, die die einzelnen Roboter verfolgen.

Pepper ist auf eine möglichst menschenähnliche Interaktion spezialisiert, besonders in der Rolle eines Kindes. Die geringe Größe, kindlichen Proportionen und freundlich wirkendes Äußeres sind darauf zurecht geschnitten. In der vorhandenen Technik wird dieser Fokus ebenfalls deutlich. Dank der Möglichkeit, die Stimmung der Nutzer aus der Kombination verschiedener Sensoren zu ermitteln, kann er dafür eingesetzt werden, sich speziell auf die besonderen Bedürfnisse bestimmter Nutzer einzustellen. Beispielsweise eignet er sich für den Einsatz mit Kindern.

Der Nao besticht durch seine beeindruckende Beweglichkeit. Als einziger der hier betrachteten Roboter besitzt er kein Fahrwerk oder Räder, sondern zwei Beine. Das ist eine zwingende Notwendigkeit, wenn er für Roboterfußball eingesetzt werden soll, bietet aber auch andere Anwendungsmöglichkeiten, wo freiere Bewegungen erforderlich sind. Ebenfalls kindlich aufgebaut und von geringer Größe, kann er teilweise dieselben Aufgaben wie Pepper übernehmen.

Im Vergleich dazu erscheint der Loomo als eine Art Segway, dem im Nachhinein ein Roboter hinzugefügt wurde. Er ist in erster Linie ein Transportmittel mit zusätzlichen Robotereigenschaften. Ohne weitere Gliedmaßen fehlen Möglichkeiten, Dinge zu greifen oder mit ihnen zu interagieren. Die Sensorik (Ultraschall, Realsense, HD-Kamera) wird im ausgelieferten Zustand vor allem zur Orientierung des Roboters im Raum und zum Verfolgen des Nutzers eingesetzt, um sich zum Beispiel wieder zu ihm zu bewegen. Der kindliche Eindruck, mit dem die anderen Roboter den Einstieg in die Interaktion erleichtern, ist beim Loomo geringer ausgeprägt. Als Segway bleibt er relativ klein, allerdings besitzt er anstelle eines Kopfes mit Gesicht ein Display.

Die Einschränkungen, denen der Loomo aufgrund von in diesem Vergleich fehlenden Teilen unterliegt, können durch optionale Erweiterungen in Angriff genommen werden. Diese sind aber im Rahmen dieser Betrachtung nicht verfügbar und können für die Einordnung des Loomo nicht benutzt werden.

Dafür ist der Loomo am einfachsten als Begleiter im Alltag zu integrieren. Da er dem Nutzer eigenständig folgen kann, wenn er nicht gar selbst als Transportmittel eingesetzt wird, kann er sich ständig in der Nähe des Nutzers aufhalten. Ob zum Tragen von Lasten, als selbstständiges Fahrzeug oder als Roboter, der Segway Robotics Loomo muss seinem Nutzern nur selten von der Seite weichen.

Welche Aufgaben der Roboter nun in der Nähe seines Nutzers erfüllen kann und sollte, wird im folgenden Abschnitt betrachtet.

3.2 Typische Aufgaben und Ziele in Bezug auf den Loomo Roboter

Die in den vorherigen Kapiteln aufgezeigten Aufgaben für Roboter sind vielschichtig und hängen unter anderem vom Nutzer oder der Nutzergruppe, vom Ort der Verwendung und von den Möglichkeiten des Roboters ab.

Nutzer des Loomo müssen in der Lage sein, ihn bedienen zu können. Das heißt, dass sie den Loomo anschalten, seinen Modus zwischen Fahr- und Robotermodus wechseln und Anwendungen auf dem Display starten kann. Alle weiteren Funktionen können anschließend wahlweise durch direktes Berühren und Führen des Loomo, über Sprache, Fernsteuerung oder, wie in dieser Arbeit geplant, auch über Gestensteuerung verwendet werden. Die Anforderungen an den Nutzer sind damit gering, auch da mehrere multimodale Interaktionsmethoden zur Verfügung stehen.

Die Betrachtung der Verwendungsorte offenbart eine große Stärke des Loomo. Dank der Möglichkeit, ihn als leichtes Transportmittel zu verwenden, und der dabei gegebenen Reichweite von bis zu 35 Kilometern kann er nicht nur in verschiedenen Umgebungen eingesetzt werden, sondern auch Nutzern folgen und in anderen Anwendungsszenarien eingesetzt werden. Er kann als direkter Begleiter oder etwa als Assistent und Führer im öffentlichen Raum Verwendung finden.

Die weiteren Möglichkeiten des Roboters sind aber zum Teil eingeschränkt. Für eigenständige Handlungen und Interaktionen mit der Umgebung fehlen Gliedmaßen wie Arme, wie sie zum Beispiel Pepper und Nao besitzen. Das neutraler gestaltete Aussehen könnte die Hemmschwelle zur Interaktion bei bestimmten Nutzergruppen vergrößern. Das Grundprogramm, welches der Loomo standardmäßig besitzt, ist zudem wenig umfangreich und muss für weitere Einsatzbereiche um neue Anwendungen ergänzt werden. Fairerweise muss gesagt werden, dass dies meist auch für andere, nicht selbst entworfene Roboter gilt. Werden eigene Mittel und Komponenten eingesetzt, kann dies den Loomo sogar deutlich stärker erweitern, als dies bei anderen Robotern möglich wäre. Dies liegt in der Erweiterungsschnittstelle begründet, mit der der Loomo um

weitere Komponenten erweitert werden kann. Roboterarme oder Greifwerkzeuge könnten seine Aktionsmöglichkeiten erweitern und neue Sensoren andere Interaktionsmethoden erlauben, etwa mithilfe von Berührungssensoren.

Davon abgesehen ist die Sensorik des Loomo vielseitig und erlaubt vielfältige Anwendungen. Die Orientierung im Raum ist mittels Ultraschall und der Intel Realsense in verschiedenen Umgebungen möglich, egal ob innen oder außen. Die Kamera sowie leichter Zugriff auf den benutzten Algorithmus erlaubt das Verfolgen der Nutzerposition und das Folgen von diesem.

Diese Grundlagen erlauben es damit auch, Gesten wahrnehmen zu können.

3.3 Gestenanalyse

So vielfältig die Roboteraufgaben, so unterschiedlich sind auch die Gestensysteme, die zur Bearbeitung dieser eingesetzt werden können.

Die meisten Gestensysteme beschäftigen sich mit der Erfüllung einer bestimmten Aufgabe, beispielsweise Zeigegesten zur Positionsbestimmung oder Handgesten. Dabei wird oft versucht, Gesten zu verwenden, die der auszuführenden Handlung ähneln. Eine Zeigegeste ähnelt dem Befehl, sich an den angezeigten Punkt zu bewegen, das Bewegen der Beine ist eine Aufforderung zu laufen und so weiter. Dies ist ebenfalls mit den Möglichkeiten des Roboters verknüpft. Zu einer Steuerung für den Nao, der sich mit vielen Freiheitsgraden flexibel bewegen kann, passt zum Beispiel eine Gestensteuerung, die den ganzen Körper einsetzt, mit Laufbewegungen, Hinhocken, Strecken und anderen Bewegungen.

Bei dem Loomo, der in den meisten Fällen Befehle zur Bewegung erhält, wären demnach Richtungsgesten oder Anweisungen zum Folgen besonders wichtig.

Das verwendete Körpermodell hängt meist von den verwendeten Sensoren ab. So benutzen viele bei der Kinect die von der Software gegebene skelettbasierte Modellierung, um den Körper und seine Gesten zu berechnen. Zum Vergleich: die für die Realsense angebotene Software bietet im wesentlichen Zugriff und Kalibrierung der Kameradaten, aber keine vorgefertigten Modelle für diese. Zur Verwendung der Daten muss auf zusätzliche Bibliotheken zurückgegriffen werden.

3.4 Zusammenfassung der aus der Analyse gewonnenen Erkenntnisse

Aus der Analyse anderer Roboter und ihrer Verwendung konnte ermittelt werden, in welchen Bereichen ein Segway Robotics Loomo eingesetzt werden kann und wo seine Stärken und Schwächen liegen. Die verwendete Sensorik und der Vergleich mit anderen Gestensystemen bietet die Grundlage für ein Gestensystem, mit dem der Loomo gesteuert und bewegt werden kann. Dieses Gestenkonzept wird im nun folgenden Abschnitt vorgestellt und erläutert.

4 Konzeptionierung einer Gestensteuerung für den Segway Robotics Loomo

Nachdem die Grundlagen an umzusetzenden Aktionen und möglichen Gesten gelegt wurden, kann ein konkretes Gestensteuerungssystem aufgestellt werden. Das Ziel ist es, unter Verwendung geeigneter Sensoren die Möglichkeiten des Loomo so einzusetzen, dass eine zur praktischen Verwendung geeignete Steuerung entsteht. Hierfür müssen die ausgewählten Sensoren beschrieben, die verwendeten Gesten ausgewählt und begründet und schließlich die eingeplanten Aktionen gezeigt werden.

4.1 Sensorische Konzeptionierung bei Umgang mit dem Loomo

Beim Loomo muss auf einer geringen Nutzfläche eine große Zahl nötiger Sensoren untergebracht werden. Viel Platz geht für Fahrwerk und Trittflächen verloren. Der Roboterkörper ist verglichen dazu kleiner, um das Gefährt nicht zu unhandlich werden zu lassen. Die für die Gestensteuerung am wichtigsten Sensoren sind die Kameras am drehbaren Kopf: die HD-Kamera im neig- und kippbaren Oberteil, die Realsense Farb- und Tiefenkamera im unteren Übergang zum Körper des Loomo. Egal, welche Kamera genau verwendet werden soll, es entstehen besondere Herausforderungen.

HD-Kamera Die HD-Kamera befindet sich direkt im Kopf des Loomo. Das bedeutet, sie ist von allen Sensoren am höchsten Punkt innerhalb des Roboters angebracht. Dies ist mit einer Höhe von knapp 60 Zentimetern aber dennoch relativ niedrig. Zur Erfassung des Nutzers ist dieser Teil des Kopfes kippbar. Die Kamera wird dabei nach oben gerichtet, um auch Oberkörper und Kopf aufzunehmen. Nachteil dieser Methode sind perspektivische Verzerrungen, wenn der gesamte Körper betrachtet werden soll. Wird dieses Kamerabild für die Gestenerkennung eingesetzt, ist zwar ein hochauflösendes Bild verfügbar, aber es fehlen Tiefeninformationen, um auch dreidimensionale Gesten auszuwerten. Dies kann die Erkennung und Unterscheidbarkeit erschweren, gerade bei komplexeren Gesten oder einer größeren Anzahl von zu erkennenden Gesten.

Realsense Die Realsense befindet sich in einer Höhe von ungefähr 50 Zentimetern am Loomo im drehbaren, aber nicht kippbaren Kopfteil des Roboters. Das bedeutet, die Kamera hat stets einen senkrechten Blickwinkel, kann aber andererseits nur schlecht höher gelegene Objekte aufnehmen. Für hoch gelegene Körperteile muss der Nutzer sich weit vom Roboter entfernen, damit dieser sie mit der Realsense erfassen kann. In einem Selbsttest wurden folgende Distanzen zwischen Mensch und Roboter und die jeweils neuen erkennbaren Körpersegmente ermittelt:

- 50 Zentimeter: Beine und Hände
- 100 Zentimeter: Unterkörper, Unterarme

- 150 Zentimeter: Oberkörper, Schultern
- 200 Zentimeter: Kopf

Sollen mit der Realsense Gesten registriert werden, muss also der Nutzer entweder weiter weg stehen, wobei die Qualität des von ihm aufgenommenen Bildes sinkt, oder er muss Gesten mit gesenkten Armen oder mit den Beinen ausführen. Der Raum für die Gesten ist also entweder in der Höhe oder der Detailgenauigkeit eingeschränkt. Dafür können die Tiefendaten eingesetzt werden, um andere Gesten verstehen und unterscheiden zu können.

Unter Betrachtung der Sensoren sind an dieser Stelle drei Konzepte denkbar. Eine Kombination, etwa von HD-Kamera und Realsense, wäre zwar auch denkbar, aber da sie sich an unterschiedlichen Positionen im Roboter befinden, wird die Angleichung und Umrechnung zwischen ihren Bildräumen schwierig. Daher wird sich im folgenden auf Varianten mit einzelnen Sensoren konzentriert.

HD-Kamera-Gestenerkennung Gesten werden über die Kamera erkannt und über zweidimensionale Bilder ausgewertet. Die Geste muss über andere Werte als die Tiefe vom Hintergrund getrennt werden. Es müssen Verdeckungen und Verzerrungen beachtet werden.

Realsense, weit entfernt Der Nutzer befindet sich in einer Distanz von 150 bis 200 Zentimetern von Loomo entfernt, welcher die Gesten über die Realsense mit Farb- und Tiefeninformationen aufnimmt. Es muss darauf geachtet werden, dass keine Hindernisse zwischen Nutzer und Roboter auftauchen. Zudem muss mit kleineren und weniger gut unterscheidbaren Bildwerten gearbeitet werden. Dafür wird es möglich und auch nötig, den ganzen Körper und größere Gesten und Bewegungen einzuplanen.

Realsense, gering entfernt Zwischen Nutzer und Loomo bleibt ein Abstand von ungefähr einem Meter. Die Gesten können nur aus Bewegungen der Beine oder der Unterarme und Hände bestehen. Werden die Arme gehoben, verschwinden sie aus dem von der Kamera erfassten Bereich.

Bei allen Konzepten muss darauf geachtet werden, dass der Roboter den Nutzer verfolgen und im Kamerabild halten kann, egal wie sich Mensch oder Roboter bewegen. Der Loomo besitzt standardmäßig einen Algorithmus zum Erfassen des Nutzers durch Kopfbewegungen und zum Verfolgen und hinterher Fahren. Bei diesem hält er einen Abstand von ungefähr einem Meter.

Eines dieser Konzepte zum Einsatz der Sensoren wird nun genauer verfolgt und für die Gestensteuerung eingesetzt. Dieses ist das Konzept zum Einsatz der Realsense bei geringer Distanz. Es bietet den Vorteil, weniger stark von der Umgebung beeinflusst zu sein. Bei der HD-Kamera kann Beleuchtung und Blickwinkel problematisch werden, bei größerer Distanz können Hindernisse sowie der allgemein große Abstand die Interaktion stören. Diese Auswahl bedeutet, dass sich die Gestenauswahl auf einfache Handbewegungen oder auf Beinbewegungen beschränken muss. Dafür bietet sich die Möglichkeit, dank der Verwendung von Tiefendaten unabhängig von Beleuchtung und Umgebungshelligkeit zu sein und auch die Bewegung nach vorne und hinten als Gesten in Betracht zu ziehen.

Aufgrund dieser Auswahl wird nun beschrieben, wie trotz der geringen Kamerahöhe bei geringer Distanz eine zufriedenstellende Gestenauswahl aussehen kann.

4.2 Befehlsauswahl anhand der betrachteten Anwendungsszenarien

In den verschiedenen Umgebungen, in denen Assistenzroboter allgemein und insbesondere der Loomo eingesetzt werden können, sind unterschiedliche Anforderungen an den Roboter typisch. Im folgenden sollen daher über verschiedene Anwendungsfälle die für den Prototypen wichtigen Aktionen gefunden werden.

Im privaten Einsatz muss der Loomo dem Nutzer nur über geringe Distanzen folgen. Wichtiger ist es, ihm im Auge zu behalten und auf sein Verhalten zu achten. Auch die Wiedergabe von Informationen, beispielsweise über Sprache, ist von Bedeutung. Außerdem ist häufig der Platz, den sich Nutzer und Roboter teilen, gering bemessen. Steht der Roboter im Weg, sollte der Nutzer diesen dazu bewegen können, in eine bestimmte Richtung auszuweichen oder sich zu drehen. Die Befehle und Aktionen sind also:

- Verfolgen des Nutzers im Raum
- Sprachwiedergabe
- Ausweichen in mehreren Richtungen
- Drehung des Roboters

Im öffentlichen Raum soll der Loomo dem Nutzer über längere Strecken folgen oder ihn führen. Eine Wiedergabe von Informationen über Sprache ist auch in diesem Kontext nützlich. Je nach Umgebung kann es aber auch nötig sein, die Sprachausgabe zu unterbrechen und den Roboter stumm zuschalten, falls Ruhe gewünscht ist. Zu der Befehlsmenge kommen dazu:

- Folgen des Nutzers und Hinterherfahren
- Lautstärke regeln

Des weiteren treten noch Befehle auf, die einen verwaltenden Charakter haben. So müssen Folge- und Verfolgungsaktionen nicht nur gestartet, sondern auch gestoppt werden. Natürlich muss auch die Gestensteuerung selbst gestartet, gestoppt oder fortgesetzt werden können. Damit kommen folgende Befehle hinzu:

- Folgen und Verfolgen stoppen
- Gestensteuerung stoppen

4.3 Genaue Gestenauswahl und Gestenerkennung

Die verwendeten Gesten können aus Bewegungen der Arme und Hände bestehen. Beinbewegungen wären von der vorhandenen Szene her ebenfalls verwendbar. Diese Möglichkeit zur gesten-gesteuerten Interaktion wird beispielsweise von Alexander et al. [Ale+12] oder Fan et al. [Fan+17] beschrieben. Diese im Kontext des Loomo einzusetzen hätte aber den Nachteil, dass sie dann vom Nutzer nicht für andere Bewegungen benutzt werden könnten. Dies würde die Freiheit des Nutzers einschränken. Außerdem kann es vorkommen, dass sich der Nutzer in der Umgebung bewegt, während er mit dem Roboter interagiert. Folgt der Roboter dabei dem Nutzer, werden

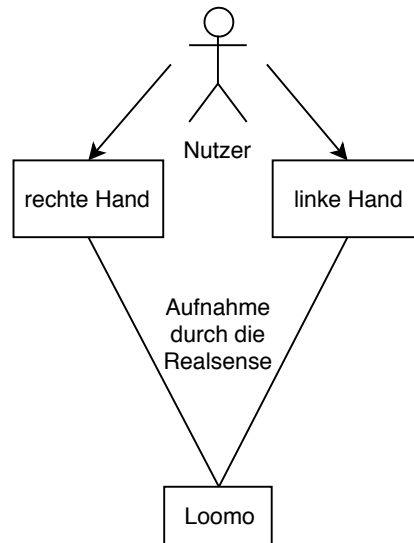


Abbildung 4.1 – Schema der Interaktion zwischen Nutzer und Loomo

diese Bewegungen dazu verwendet, die Roboterposition anzupassen. Würden die Beinbewegungen in diesem Fall als Gesten interpretiert, würde dies den Bewegungsablauf stören. Aus diesen Gründen wird auf die Erfassung der Beine als gestenausführende Gliedmaßen verzichtet wird.

Damit stehen die Hände und Arme zur Ausführung der Gesten fest. Nun ist zu bestimmen, welche Gesten mit den Händen durchzuführen sind und wie diese strukturiert und dem Nutzer vermittelt werden sollen. Einerseits sollte auf die Ergonomie geachtet werden. Wird länger und mit komplizierten Bewegungen gestikuliert, wie es beispielsweise für die Gebärdensprache erforderlich ist, kann dies zu Schmerzen und Beschwerden führen[RCL14]. Diese Erkenntnis muss in der Gestenentwicklung berücksichtigt werden. Daher wird auf schwierige Verrenkungen oder komplizierte Bewegungen verzichtet. Es bietet sich an, eine neutrale Ruheposition zu verwenden, von der aus mit einfachen Bewegungen Gesten ausgeführt werden können. Des weiteren bieten sich zur Vereinfachung der Erkennung statische Gesten an. Das bedeutet, der Raum wird in räumliche Zonen unterteilt, anhand derer eine Geste klassifiziert wird. Diese können als Bezugspunkt den Körper des Nutzers verwenden.

Über die dreidimensionalen Positionen von linker und rechter Hand im Vergleich zum Körper sollen die verschiedenen Gesten erkannt werden. Hierbei gilt für jede Hand jeweils die neutrale Position als Basis, in der sie sich zu Beginn der Interaktion befinden. Diese Zone wird seitlich am Körper positioniert. Sie wäre auch vor dem Körper einplanbar, so dass die Gesten zwischen Körper des Nutzers und Loomo ausgeführt werden. Allerdings kann es, unter Berücksichtigung der Distanz zum Loomo, dazu kommen, dass dabei die minimale verwendbare Distanz zur Realsense unterschritten wird. Um dies zu vermeiden, wird die Ruheposition an die jeweilige Körperseite gelegt. Im Anschluss können die Hände in sechs Richtungen (oben, unten, links, rechts, vor, zurück) bewegt werden. Vertikale Bewegungen sollen allerdings nicht berücksichtigt werden. Sie könnten problematisch werden, etwa wenn die Hände am oberen Rand des Kamerabereiches erfasst werden. Dann ließe sich aus der neutralen Position heraus die obere Richtung nicht verwenden. Für die Verwendung der unteren Richtung müsste der Nutzer häufiger in die Knie gehen oder die Arme gezielt nach unten ausstrecken. Beides sind auf Dauer anstrengende Bewegungen. Daher werden oben und unten als Richtungen ebenfalls verworfen. Damit verbleiben vorne, hinten,

links und rechts als Bewegungsrichtungen, um verschiedene Reaktionen auszulösen. Als zusätzliche Zustände wird Neutral für die mittlere Ruheposition sowie ein Nicht-Zustand im Falle des Nichterkennens einer Richtung oder von einer Hand benutzt. Zum Beginn der Anwendung und bei Fehlern in der Handerkennung, zum Beispiel wenn der Nutzer die Arme hochhebt und somit aus dem erfassten Bereich entfernt, sollen sich die Hände in diesem Zustand befinden. In einem solchen Fall soll keine Aktion ausgeführt werden. Die Abbildung 4.1 stellt dar, wie die Interaktion zwischen Nutzer und Loomo geplant ist.

Mit fünf verwendbaren Handpositionen je Hand sind prinzipiell 25 verschiedene Kombinationen für verschiedene Aktionen möglich. Für die Auswahl, welche Gesten den einzelnen Aktionen zugewiesen werden sollen, sind möglichst intuitive Zuordnungen zu finden. Blackler et al. haben grundlegende Prinzipien für eine intuitive Interaktion aufgestellt, die an dieser Stelle kurz eingeführt werden sollen [AVD06]:

1. Sind im aktuellen Kontext dem Nutzer bereits Funktionen bekannt, können diese zur Erklärung ähnlicher Funktionen eingesetzt werden.
2. Gibt es in einem anderen Kontext ähnliche Funktionen, können diese zur Erklärung auf den aktuellen Kontext angewendet werden. Nötigenfalls kann dabei auf Metaphern zurückgegriffen werden, wie etwa die Desktop-Metapher.
3. Redundanzen in der Beschreibung von Funktionen und Konsistenz der Funktionen in verschiedenen Zuständen der Anwendung erleichtern die Zugänglichkeit.

Die Funktionen des Loomo können dem Nutzer bereits bekannt sein, allerdings ist es schwierig, dieses Wissen auf Gesten anzuwenden. Metaphern können helfen, den Gesten bestimmte Bedeutungen zuzuweisen. Die Aktionen, die der Roboter durchführt, sollten redundant auf mehreren Wegen dem Nutzer vermittelt werden. Außerdem müssen die Gesten auch in verschiedenen Positionen von Nutzer und Roboter und in unterschiedlichen Steuerungsmodi erfasst werden können. Zusammengefasst lassen sich auf Grundlage der Prinzipien für eine intuitive Interaktion mehrere Regeln für die Gestensteuerung des Loomo aufstellen:

- Die Aktionen sollen gruppiert werden, so dass sich ähnelnde Gesten zu ähnlichen Reaktionen führen.
- Starten und Beenden einer kontinuierlichen Aktion, wie Folgen, benötigen zwei Gesten und werden durch Umkehr der jeweils anderen Geste ausgedrückt.
- Bei Bewegungen entspricht die Richtung der Geste der Bewegungsrichtung des Roboters.

Die linke Hand soll eine Kontrollfunktion übernehmen. Mit ihr soll zwischen verschiedenen Gruppen von Gesten gewechselt werden können. Mit ihr wird damit grob bestimmt, welcher Art die folgende Aktion des Roboters sein soll. Die rechte Hand dient dann zur genauen Auswahl innerhalb einer dieser Funktionsgruppen. Bei Bewegungen ist dies zum Beispiel die Bewegungsrichtung. Somit sind anhand der linken Position fünf Gruppen von Aktionen möglich: neutral, vorn, hinten, links, rechts.

Dieses Schema gibt auch die Schrittfolge für den Nutzer vor. Er beginnt nach Positionierung im Kamerafeld und Erfassung des Körpers und der Hände mit der Bewegung der linken Hand. Im Anschluss folgt die rechte Handbewegung und die folgende Aktion des Roboters. Die fünf Richtungsgruppen geben dabei vor, dass es entsprechend maximal fünf Gruppen von Aktionen geben

kann. Ansonsten käme es zu Problemen, weil sich die Gesten nicht eindeutig der Art der Aktion zuordnen lassen. Allerdings ist in der neutralen Gruppe bereits ein Zustand belegt. Im Startzustand, direkt nachdem die Hände erfasst wurden, sind beide Hände im neutralen Zustand. Damit lassen sich diesem keine Aktionen zuteilen, weil sonst die Unterscheidung von ruhender Haltung vor einer Geste und einer tatsächlichen Geste nicht mehr möglich ist. Es widerspräche zudem der Metapher der Ruheposition, wenn bei Bewegung der rechten Hand Gesten ausgeführt werden, obwohl die linke Hand neutral gehalten wird. Befindet sich mindestens eine Hand im neutralen Zustand, soll daher keine Aktion ausgeführt werden.

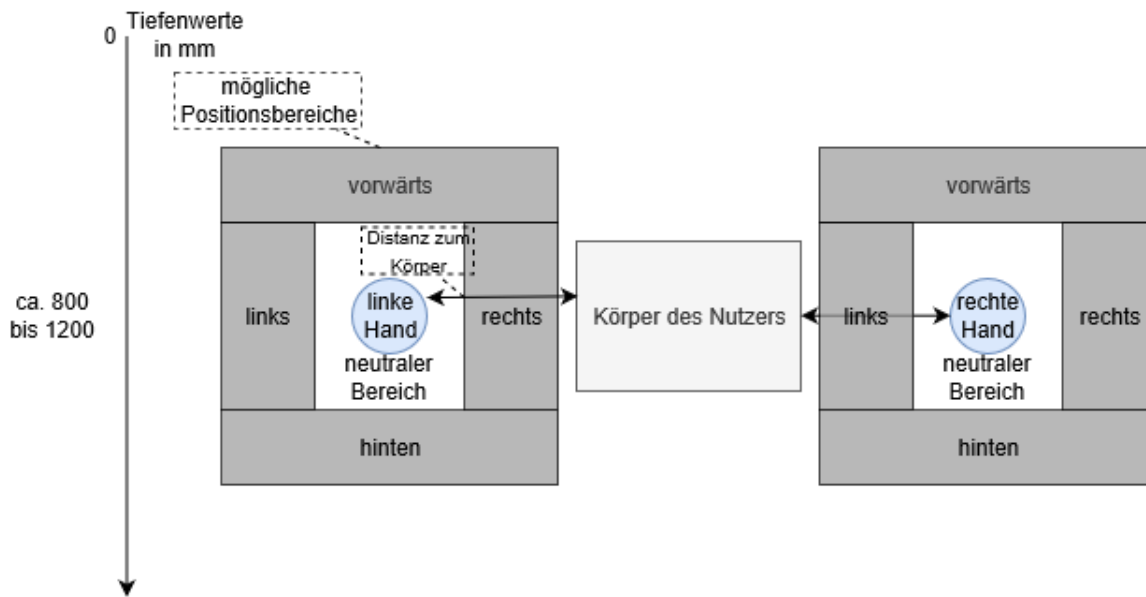


Abbildung 4.2 – Schema der Handverarbeitung und möglichen Positionen

Die Abbildung 4.2 gibt schematisch die möglichen Handpositionen wieder. Die tatsächlich verwendeten Gruppen mit ihren jeweiligen Metaphern, sortiert nach den durch die linke Hand gegebenen Gruppen, sind:

vorn Eine Bewegung nach vorn erhält in diesem System eine höhere Bedeutung. Sie hat etwas Wichtiges an sich, da sie sich näher am Interaktionspartner befindet. Daher sollten sich die wichtigsten Funktionen des Loomo in dieser Gruppe befinden. Diese sind das Verfolgen des Nutzers und das Hinterherfahren. Das Verfolgen ist für die gesamte gestengesteuerte Interaktion wichtig, um den Nutzer im Blickfeld der Kamera zu halten. Das Fahren erhält vor allem in der Nutzung der Beweglichkeit des Roboters seine Wichtigkeit. Die Gesten sind wie folgt der rechten Handausrichtung zugeordnet:

- vorn: Folgen starten (Nutzer hinterher fahren)
- hinten: Folgen wieder beenden
- rechts: Verfolgen starten
- links: Verfolgen wieder beenden

links Der Bewegung nach vorne untergeordnet, allerdings immer noch in mittlerer Distanz, befindet sich die Bewegung nach links. Als Bewegung der linken Hand zeigt sie vom Körper weg. Ihr werden Aktionen zugeordnet, die Bewegungen des Roboters umfassen, etwa zur Seite oder vom Nutzer weg. Diese sind weniger wichtig, wenn einem der Roboter direkt folgen soll, sondern dienen eher für kleinere Bewegungen und Anpassungen der Position des Roboters.

- vorn: Der Roboter entfernt sich vom Nutzer. Die Bewegungsrichtung, bei der sich die Hand vom Körper entfernt, entspricht der Richtung der Geste.
- hinten: Der Roboter nähert sich.
- rechts: Der Roboter weicht nach rechts aus. Er dreht sich, bewegt sich und dreht sich zurück in Richtung des Nutzers.
- links: Der Roboter weicht nach links aus.

rechts Die Bewegung nach rechts bedeutet, dass die linke Hand an den Körper herangeführt wird. Damit werden für diese Gruppe eher kleinere oder persönlichere Aktionen eingesetzt. Dies sind die Ausgabe von Sprache, etwa um Informationen wieder zu geben. Auch die Unterbrechung der Gestenerkennung wird hier eingebunden.

- vorn: Gestensteuerung unterbrechen
- hinten: Gestensteuerung fortsetzen
- rechts: Sprachausgabe, Position des Nutzers wiedergeben
- links: Sprachausgabe, Nutzer begrüßen

hinten Aktionen, die eher nebensächlich sind und vermutlich seltener eingesetzt werden, sollen in dieser Kategorie Platz finden. Diese Aktionen dienen dazu, den Anforderungen für speziellere Kontexte zu dienen, etwa wenn Ruhe gewünscht ist oder ein bestimmtes Objekt vom Roboter erfasst werden soll. Daher kann in dieser Gruppe die Lautstärke und Drehung des Kopfes beeinflusst werden.

- vorn: Roboter stumm schalten
- hinten: Sprachausgabe wieder aktivieren
- rechts: Kopfbewegung nach rechts
- links: Kopfbewegung nach links

Mit diesem System sind mit vier Bewegungsrichtungen somit insgesamt 16 verschiedene Aktionen möglich. Die Aktionen sind in der Tabelle 4.1 aufgelistet. Dabei werden die Bewegungstechniken des Roboters eingesetzt, sowohl von der Basis als auch vom Kopf. Zur Ausgabe weiterer Informationen wird die Ausgabe von Sprache mit zwei verschiedenen Nachrichten verwendet, einmal als feste Begrüßung, einmal mit Daten aus der Erkennung des Nutzers.

Linke Hand	Rechte Hand	Aktion
neutral	beliebig	keine Aktion
beliebig	neutral	keine Aktion
links	links	nach links ausweichen
links	rechts	nach rechts ausweichen
links	vorne	zurück bewegen
links	hinten	vorwärts bewegen
rechts	links	Nutzer begrüßen
rechts	rechts	Distanz wiedergeben
rechts	vorne	Gestensteuerung unterbrechen
rechts	hinten	Gestensteuerung fortsetzen
vorne	links	Nutzer folgen stoppen
vorne	rechts	Nutzer folgen starten
vorne	vorne	Verfolgen starten
vorne	hinten	Verfolgen stoppen
hinten	links	nach links drehen
hinten	rechts	nach rechts drehen
hinten	vorne	stumm schalten
hinten	hinten	laut schalten

Tabelle 4.1 – Auflistung der Gesten und zugehörigen Aktionen

4.4 Geplante Methodik der Mensch-Roboter-Interaktion

Der Ablauf der Interaktion soll einer bestimmten Schrittfolge folgen, damit das Steuern des Roboters einheitlich und systematisch möglich ist.

Zu Beginn wird die Anwendung zur Gestensteuerung gestartet. Sobald der Zugriff auf die Realsense-Kamera gegeben ist, wird der Nutzer vom Loomo verfolgt und im Aufnahmebereich der Kamera verarbeitet. Der Nutzer soll eine neutrale Haltung einnehmen, in der Körper und Arme voneinander getrennt erkannt werden können. Diese Position im Vergleich zum Körper wird als neutraler Zustand zur Kalibrierung eingesetzt. Im Anschluss daran kann mit den Gesten begonnen werden. Sobald die Entfernung der Hand zur neutralen Pose einen Schwellwert überschreitet, wird der Zustand von neutral in den der erkannten Richtung überführt. Der Nutzer bewegt erst die linke Hand und den Arm, um die Art der Aktion zu bestimmen, die ausgeführt werden soll. Mit der Bewegung der rechten Hand und Arm wird daraufhin die konkrete Aktion ausgeführt.

Nach Ausführung der Aktion wird, falls nötig, der Nutzer wieder verfolgt und im Kamerabild gehalten. Danach kann mit der nächsten Geste fortgefahren werden.

4.5 Zusammenfassung des Gestenkonzepts

Der Nutzer erhält die Möglichkeit, über Handbewegungen mit dem Loomo zu interagieren. Mit der Gruppierung anhand der linken Hand soll dabei die Erlernbarkeit verbessert werden. Für die genaue Zuordnung der Gesten zu einzelnen Aktionen werden vorwiegend Metaphern und Regeln

für eine hohe Intuitivität eingesetzt. Insgesamt 16 Aktionen sollen möglich sein, mit der Option, diese Menge zu erweitern, falls neue Aktionen nötig sein sollten.

Damit wurde festgelegt, wie der Nutzer mit dem Roboter interagieren soll, wie er von ihm wahrgenommen wird und wie der Loomo auf die einzelnen Gesten reagiert. Diese Steuerung gilt es nun praktisch mithilfe eines Prototypen umzusetzen.

5 Prototypische Implementierung der Gestensteuerung

Es wurde ein Konzept entworfen, mit dem sich ein Segway Robotics Loomo über Gesten steuern lässt. Dieses gilt es nun praktisch umzusetzen. Dafür soll auf Basis des gegebenen Loomos eine prototypische Anwendung entwickelt werden, die das Steuerungskonzept umsetzt. Der Aufbau und die Funktionsweise dieses Prototyps werden im folgenden beschrieben.

5.1 Technische Grundlagen

Die Anwendungen, die für den Loomo umgesetzt werden können, basieren auf dem Android-Betriebssystem und werden daher in Java geschrieben. Zum Zugriff auf die Grundfunktionen des Roboters dient das Loomo-SDK. Dieses bietet auch den Zugriff auf die Kameradaten der Intel Realsense, für den sonst die zusätzliche Einbindung des Realsense-SDK von Intel nötig wäre. Die nötigen Funktionen sind in verschiedene Services unterteilt, von denen jeweils eine Instanz zur Laufzeit an die Anwendung gebunden werden muss. Für die Gestensteuerung werden die folgenden Services für die beschriebenen Funktionen verwendet:

- Vision: Aufnahme der Kameradaten und Verfolgungsalgorithmen
- Base: Steuerung der Motoren und Räder
- Head: Steuerung des Kopfes
- Speaker: Sprachausgabe via Text to Speech (TTS)

Zur Verarbeitung der erhaltenen Bilder ist eine zusätzliche Bibliothek nötig. Dafür wird hier auf OpenCV zurückgegriffen. Diese bietet Funktionen zum Trennen der nötigen Bilddaten vom Rest des Bildes, Berechnung wichtiger Werte wie die Flächen bestimmter Bildbereiche sowie deren Schwerpunkte.

5.2 Programmaufbau

Die prototypische Anwendung besteht aus drei Hauptklassen: die MainActivity, die das User Interface (UI) aufbaut und die Services bindet, der DepthPresenter, der die Tiefeninformationen verarbeitet und die Handdaten verwaltet und die Actions-Klasse, die die vom Roboter auszuführenden Aktionen enthält. Das Klassendiagramm 5.1 zeigt die drei Klassen und ihre wichtigsten Attribute und Methoden.

Die MainActivity hat zwei Aufgaben: Das Aufbauen des UI und das Binden der benötigten Services. Da der Nutzer meist aus der Distanz nur wenige Details des Bildschirms ausmachen kann, ist das UI einfach gehalten. Es besteht aus einem Button, mit dem die Services gebunden und der

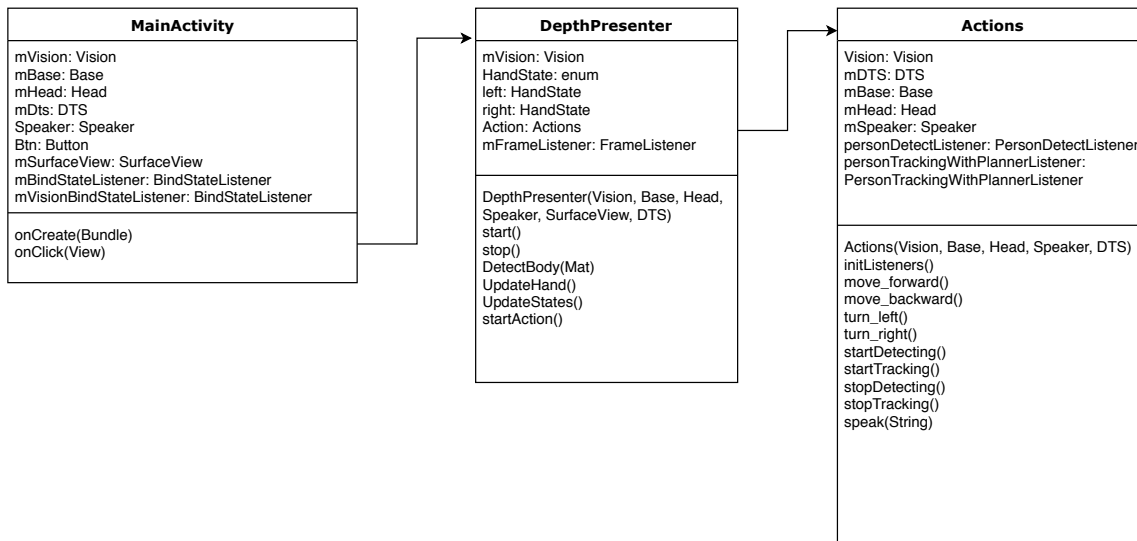


Abbildung 5.1 – Klassendiagramm der Anwendung

DepthPresenter gestartet werden kann. Zudem enthält das UI eine Vorschau auf das von der Realsense aufgenommene Tiefenbild. Die Abbildung 5.2, die das UI während der Anwendung zeigt, befindet sich am Ende des Kapitels.

Der DepthPresenter wird von der MainActivity erzeugt und aufgerufen. In dieser Klasse wird die Tiefenkamera der Realsense über den Vision Service aufgerufen und jeder Bildframe als Tiefenbild entnommen. Er dient dazu, die aufgenommenen Tiefendaten mithilfe von OpenCV zu verarbeiten. Hierfür werden in mehreren Schritten Signalverarbeitung, Filterung und Segmentierung und Gestenerkennung vorgenommen. Zuerst werden als Signalverarbeitung Körper und Hände aus dem erzeugten Bild entnommen. Die Hände werden bei der Filterung ausgewertet und ihre Tiefe und Distanz zum Körper als Geste gespeichert. Die Geste wird erkannt, indem die Tiefen- und Distanzwerte einer Richtung zugeordnet werden.

Der Körper wird aus der Szene extrahiert, indem diese auf ein bestimmtes Band an Tiefenwerten zurechtgeschnitten wird. Dieses liegt anfangs bei 700 bis 1700 Millimetern. Wurde bereits ein Körper gefunden, wird das Band auf die Tiefe des Körpers +/- 200mm reduziert. In der so reduzierten Szene wird nun die größte wahrgenommene Fläche auf ihre Größe geprüft. Ist sie im festgelegten Bereich, wird der Körper als gefunden vermerkt und gespeichert. Wurde in folgenden Frames bereits ein Körper gefunden, wird die neue Szene darauf überprüft, ob sich weiterhin eine als Körper erkennbare Fläche im Bild befindet.

Sobald ein Körper gefunden wurde, wird die Szene auf mögliche Arme und Hände untersucht. Es wird dabei davon ausgegangen, dass der Nutzer mittig im Bild steht. Ausgehend vom bereits durch das Finden des Körpers erzeugten, reduzierten Tiefenbild wird nun nach den Händen gesucht. Dafür wird zuerst der Bereich, in dem sich der Körper befindet, aus dem Bild entfernt, indem seine Werte auf Null reduziert werden. Im Anschluss werden die beiden größten Flächen im restlichen Bild darauf hin untersucht, ob sie von Größe und Position im Raum her als Hände in Betracht gezogen werden können. Ist dies der Fall, werden die Hände als gefunden vermerkt und ihre nötigen Positionsdaten gespeichert. Diese sind neben dem Tiefenwert auch der Schwerpunkt der Fläche. In folgenden Frames wird (nach Entfernen der Körperfläche) ermittelt, ob es weitere Flächen gibt, die Hand- und Armflächen sein können. Liegt der Schwerpunkt in einer der geeig-

neten Flächen, wird dieser als Handfläche verwendet.

Danach werden die Hände einem Positionszustand zugeordnet. Dieser ist direkt nach der Erkennung der neutralen, bei dem keine Geste ausgeführt werden soll. Die weiteren Zustände sind als Relationen zum Körper und seiner Position definiert. Entfernt sich die Hand nach links oder rechts, wird der Zustand auf den entsprechenden Seitenzustand geändert. Findet aber auch eine Änderung der Tiefe statt, erhält diese Vorrang.

Der DepthPresenter erzeugt eine Instanz der Actions-Klasse. Wurden beide Hände gefunden und befinden sich in einem nicht neutralen Zustand, wird anhand der Richtungen und des Schemas eine Aktion ausgeführt.

In der Actions-Klasse sind alle vom Loomo über Gesten ausführbaren Aktionen hinterlegt. Die Klasse wird erzeugt, indem die gebundenen Services aus der MainActivity erst an den DepthPresenter und im Anschluss an die Actions übergeben werden. Die Aktionen lassen sich in drei grundlegende Arten unterteilen: einfache Bewegungen von Kopf oder Loomo selbst, Sprachausgabe und das Verfolgen.

Für Bewegungen werden die entsprechenden Services (Base oder Head) angesprochen. Für die genauen Basisbewegungen werden die Vorwärts- oder Seitwärtsgeschwindigkeiten angepasst. Für einfache Bewegungen genügt dies, während für das Abbiegen eine Verkettung nötig ist. Eine Bewegung zur Seite besteht aus einer Drehung in die gewünschte Richtung, einer Vorwärtsbewegung und einer Drehung zurück in die umgekehrte Richtung, um den Nutzer wieder zu finden. Für den Kopf wird der Kipp- und Neigungswinkel manipuliert, worauf der Kopf sich in die entsprechende Position begibt.

Zur Sprachausgabe wird ein Listener initialisiert, der auf den Sprachausgabebefehl reagiert. Die Sprachaktion enthält die Sprachnachricht in Textform, welche danach vom Speaker-Service mithilfe des Listeners und der TTS-Funktion auditiv umgesetzt wird.

Für die Verfolgung wird auf den Detection and Tracking System (DTS) Algorithmus zurückgegriffen, der als Teil des Vision-Service zur Verfügung steht. Dieser ermittelt in dem von der normalen HD-Kamera aufgenommenen Bild eine zu verfolgende Person. Danach kann je nach gewünschter Aktion der Kopf des Loomo der Person folgen oder der Person hinterher gefahren werden. Dabei spielen auch die Ansteuerungsmodi eine Rolle. Für die einfachen Bewegungen befinden sich Base und Head in dem rohen Steuerungsmodus. Bevor auf DTS zurückgegriffen werden kann, müssen die Modi angepasst werden. Außerdem darf während des Laufens einer DTS-Funktion keine der rohen Aktionen aktiv werden. Diese dürfen erst wieder aktiviert werden, wenn die Verfolgungsaktion durch eine Stopaktion beendet wird.

5.3 Zusammenfassung

Der Prototyp erfüllt drei Aufgaben. Er gibt dem Nutzer ein Interface vor, welches ihm bei der Interaktion Hilfestellung leistet. Er benutzt die erfassten Tiefendaten und verarbeitet sie und implementiert die Aktionen, die der Roboter als Reaktion auf bestimmte Gesten durchzuführen hat.

Dieser Prototyp gibt eine Möglichkeit vor, wie das entwickelte Gestenkonzept praktisch umgesetzt werden kann. Diese Umsetzung soll nun von Anwendern getestet und in einer Praxissituation evaluiert werden.

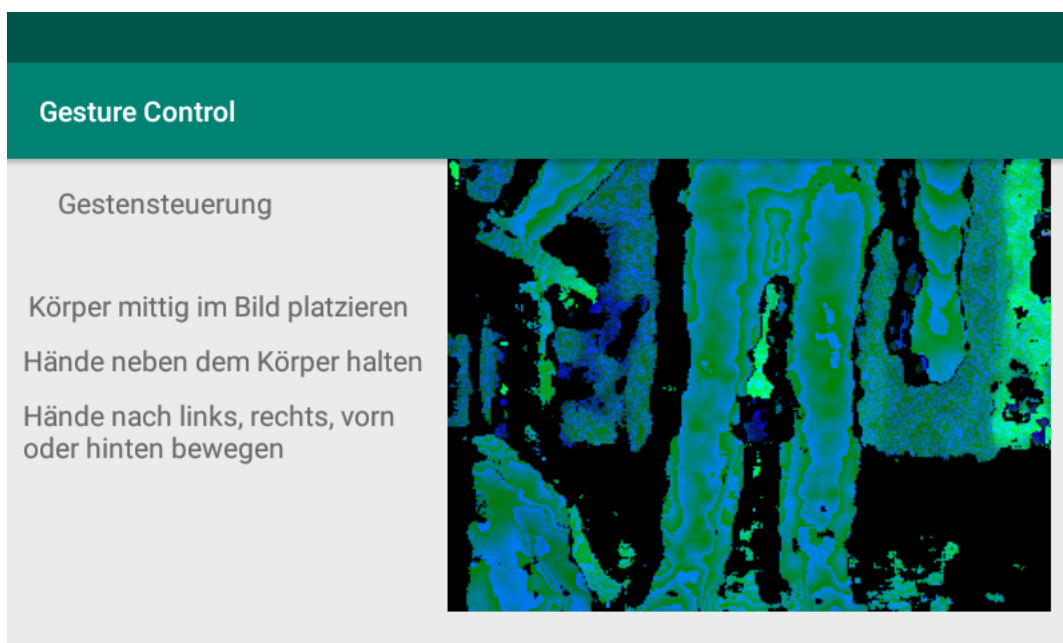


Abbildung 5.2 – User Interface

6 Evaluation der entwickelten Steuerung

Die entwickelte Steuerung muss noch daraufhin untersucht werden, ob sie die Aufgabenstellung erfüllen und die Forschungsfrage zufriedenstellend beantworten kann. Dafür geeignet ist eine praktische Erprobung der Steuerung durch verschiedene Probanden. Der Ablauf und die Erhebung dieser Evaluation sind der Inhalt des folgenden Kapitels.

6.1 Planung des praktischen Steuerungstests

Der Test der entwickelten Steuerung erfolgt in Form eines Hindernisparkour. In diesem soll der Loomo, von einem Tester mittels Gesten gesteuert, einem bestimmten Pfad folgen und verschiedene Aufgaben absolvieren. Damit wird die Beweglichkeit, Geschwindigkeit und die Genauigkeit der Gestenerkennung überprüft. Das Ziel des Tests ist es zu ermitteln wie gut die Gestensteuerung sinnvoll verwendet werden kann. Dafür ist die Zeit von Bedeutung, die für einzelne Aktionen und für die gesamte Strecke benötigt wird. Ist diese zu hoch, etwa weil der Roboter Gesten nicht oder falsch erkennt oder weil der Nutzer die falschen Gesten verwendet, ist die Steuerung wenig praktikabel. Um die Gründe für eine zu lange Zeit genauer untersuchen zu können, werden die Fehler gezählt. Diese sind einerseits falsch erkannte Gesten, wenn der Roboter die Geste anders umsetzt, als der Nutzer sie wünscht. Dies lässt darauf schließen, dass die implementierte Erkennung und Auswertung der Geste Probleme aufweist. Andererseits kann es sein, dass der Nutzer die falsche Geste auswählt. In diesem Fall liegt das Problem vermutlich daran, dass der Nutzer das Gestensystem nicht oder falsch in Erinnerung hat. Sind diese Fehler auffällig, sollte die Gesten- und Aktionsauswahl überdacht werden. Es könnte beispielsweise zu viele mögliche Gesten geben oder die Gesten selbst schlecht zu merken sein. Gemessen werden soll also die benötigte Zeit, um die Strecke zu absolvieren, sowie die Anzahl der Gesten, die dafür notwendig war. Des Weiteren sollen auch die nicht oder falsch erkannten Gesten gezählt werden, entweder als falsch verarbeitete oder falsch ausgeführte. Dies soll es erlauben, Fehler in der Anwendung (Handposition falsch zugewiesen, Hand oder Körper falsch erkannt...) von Nutzerfehlern, wenn etwa die falsche Geste für die gewünschte Aktion benutzt wird, zu trennen.

Innerhalb der Strecke sollen möglichst viele Aktionsmöglichkeiten des Loomo getestet werden. Er soll durch Gesten bewegt werden, sowohl direkt als auch dem Nutzer folgend. Wenn dazu aufgefordert, führt er Sprachwiedergaben aus und dreht den Kopf in Richtung des Nutzers. Über die Sprachausgabe erhält der Nutzer außerdem Rückmeldung über seine Position und die vom Loomo ausgeführten Aktionen. Mit dieser Aktionsauswahl soll erreicht werden, dass möglichst viele der verwendbaren Gesten und damit Handpositionen in ihrer Erkennbarkeit von den Testern geprüft werden.

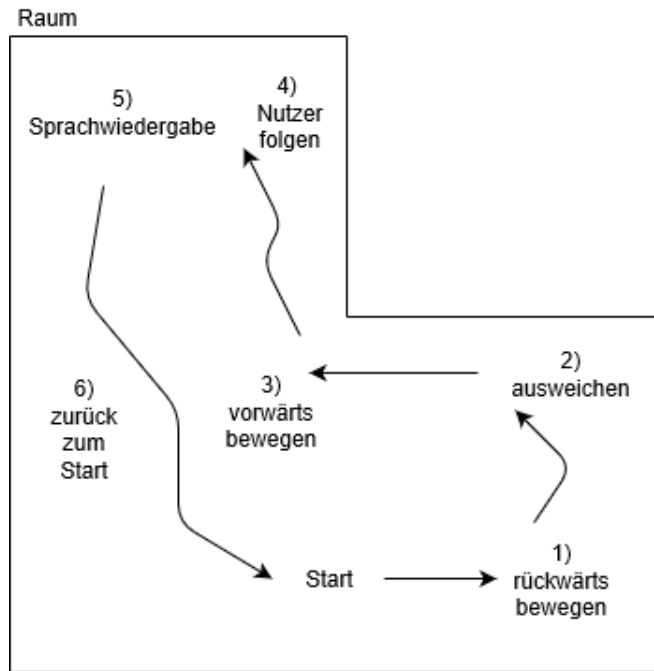


Abbildung 6.1 – Schema der Evaluationsumgebung

6.2 Durchführung der Evaluation

Der Testlauf findet in einem geschlossenen Raum statt. In ihm befinden sich der Tester mit dem Loomo sowie der Evaluierende. Der Tester erhält eine kurze Einführung in den Loomo und Demonstration der Gestensteuerung. Im Anschluss wird ihm die Aufgabe gestellt, den Loomo einmal durch den Raum zu navigieren. Der Evaluierende notiert sich die verwendeten Gesten, stoppt die Zeit und setzt im Notfall den Roboter zurück. Ist eine der Stationen erreicht, werden weitere Aufgaben gestellt. Nach Abschluss der Teststrecke verlässt der Tester den Raum, der Roboter wird zurück an den Startpunkt gebracht und der nächste Tester beginnt.

Der in der Abbildung 6.1 dargestellte Ablauf findet wie folgt statt. Er beginnt nach Einweisung und Demonstration mit den grundlegenden Bewegungsgesten. Zuerst wird der Roboter dazu aufgefordert, sich rückwärts vom Nutzer zu entfernen. Anschließend wird er durch seitliches Ausweichen und Vorwärtsbewegung in Richtung Nutzer durch den rechten Bereich des Raumes gelotst. Anschließend wird die Funktion zum Folgen des Nutzers angewendet, um dem Nutzer frei zur nächsten Station hinterher zu fahren. Dort soll der Nutzer eine Sprachwiedergabe vom Loomo auslösen und ihn danach zurück zum Startpunkt führen. Die bei der Evaluation getesteten Aktionen sind somit vorwärts, rückwärts und seitlich bewegen, Nutzer verfolgen, verfolgen stoppen und Nutzer begrüßen. Von den 16 möglichen werden dabei sieben Aktionen aus drei verschiedenen Gestengruppen getestet

6.3 Gewonnene Werte aus der Evaluation

Die Evaluation wurde mit vier Probanden durchgeführt. Ihre Zeiten, benötigten Gesten, Fehler bei Ausführung der Gesten und Erkennung durch den Loomo und Kommentare sind in der Tabelle 6.1 aufgelistet. Insgesamt zeigt sich, dass deutlich mehr Gesten benötigt wurden, als theoretisch nötig wären. Im Idealfall werden aufgrund der Abmessungen der Testumgebung rund 20 Gesten zum Abschließen der Strecke benötigt. Die Tester haben die doppelte bis dreifache Menge an Gesten eingesetzt. Die Zeiten waren neben der Menge der eingesetzten Gesten auch stark davon beeinflusst, mit welcher Herangehensweise die Nutzer an die Interaktion herangingen. Einige waren vorsichtiger als andere, haben langsamer reagiert oder erst nachgefragt, welche Geste verwendet werden sollte. Zudem haben sich, je nach Nutzer, verschieden starke Störungen bei der Ausführung und Erkennung der Gesten ergeben. Diese lassen sich zum Teil durch Unterschiede in Größe und Körperbau der Teilnehmer erklären. Diese körperlichen Unterschiede haben Einfluss auf die Erfassung von Körper und Armen genommen. Allgemein ist die Anzahl von der Anwendung falsch registrierter Gesten größer als die Anzahl falsch vom Nutzer ausgeführter Gesten. Auffällig war, dass mehrere Tester anfangs davon ausgingen, die Gesten würden vor dem Körper ausgeführt werden, nicht seitlich. Dies entspricht einem der vorher geprüften und verworfenen Konzepte. Im Vergleich zur Erkennung war die Ausführung aber danach weniger fehlerbehaftet.

Proband	Zeit	Gestenanzahl	falsch ausgeführt	falsch erkannt	Kommentare
1	9 Minuten	44	6	19	vorne und Seiten einfach, hinten anstrengend
2	15 Minuten	51	8	16	auf Dauer anstrengend, anfangs verwirrend
3	17,5 Minuten	60	11	25	Gestenauswahl in Ordnung, Runterbeugen schlecht, stimmige Steuerungsidee
4	16 Minuten	45	7	11	störende einzuhaltende Abstände

Tabelle 6.1 – Ergebnisse der Evaluation

6.4 Zusammenfassung des Nutzertests

Das entwickelte und prototypisch umgesetzte System wurde mithilfe eines Nutzertests in Form einer Teststrecke evaluiert. Mit mehreren Teilnehmern wurden typische Befehle und Anforderungen an den Roboter geprüft. Ermittelt wurden je nach Nutzer unterschiedliche Zeiten und Erkennungsfehler. Die Ausführung der Gesten verlief nach anfänglichen Schwierigkeiten leichter. Damit wurde das angefertigte Programm zur Gestensteuerung eines Assistenzroboters evaluiert. Diese Evaluation kann dazu verwendet werden, mit ihren Ergebnissen die Nützlichkeit der Anwendung zu bewerten und zu prüfen, ob die anfängliche Forschungsfrage zufriedenstellend beantwortet werden konnte.

7 Abschluss

Die entworfene Gestensteuerung wurde praktisch umgesetzt und mit mehreren Nutzern evaluiert. Nun gilt es zu prüfen, wie dabei die Forschungsfrage beantwortet werden konnte. Dafür wird in Kurzform noch einmal die bisherige Arbeit zusammengefasst und neben weiteren Verwendungsmöglichkeiten für Gestensteuerung schließlich die Forschungsfrage beantwortet.

7.1 Zusammenfassung des entwickelten Systems

In der vorliegenden Arbeit wurde sich mit der Frage auseinandergesetzt, wie eine Gestensteuerung für einen Assistenzroboter zu entwerfen und umzusetzen sei. Dafür war es zuerst nötig, die Grundlagen in Bezug auf Gesten und Gestensteuerungen und an Assistenzrobotern und ihrer Verwendung bereitzustellen. Mit dieser Basis war es möglich, bestehende Steuerungen darauf zu untersuchen, wie sie auf die Arbeit mit einem Assistenzroboter angewendet werden können. Ebenso wichtig war es zu erfahren, wie und in welchem Umfeld Roboter eingesetzt werden. Welche Aufgaben haben sie und wie werden diese umgesetzt?

Neben der Analyse dieser Anfangsdaten wurde auch der praktisch verwendete Segway Robotics Loomo betrachtet. Er wurde zusammen mit den Gesten und ermittelten Aufgaben dahingehend untersucht, wie er mit Gesten gesteuert werden kann und welche Aufgaben für ihn sinnvoll sind.

Mithilfe dieser Erkenntnisse, wie der Loomo praktisch mit Gesten verwendet werden könnte, wurde ein Konzept zur Gestensteuerung entworfen. Dieses besteht aus der Bewegung von linker und rechter Hand. Über Bewegungen in vier verschiedene Richtungen sind insgesamt 16 Aktionen des Roboters verwendbar. Diese umfassen unter anderem verschiedene Bewegungen, das Verfolgen des Nutzers und eine Sprachausgabe.

Daraufhin wird beschrieben, wie das Programm, mit dem die Gestensteuerung umgesetzt wurde, geschrieben und implementiert wurde. Über die Verarbeitung der von einer Realsense-Tiefenkamera aufgenommenen Bilddaten und die Verarbeitung der Tiefendaten mittels OpenCV war es möglich, Handpositionen und damit Gesten zu unterscheiden und in verschiedenen Aktionen umzusetzen.

Darauf folgend wurde die Evaluation von Anwendung und Gestensteuerung beschrieben. In einem Nutzertest wurden grundlegende Aktionen in einer Hindernisstrecke geprüft. Dabei wurde ermittelt, dass die Anwendung unterschiedlich auf verschiedene Anwender reagiert. Nachdem die Nutzer sich auf das System und den Roboter einstellen konnten, gelang es ihnen aber dennoch, die Strecke zu absolvieren.

Die Erkenntnisse der Gestensteuerung sollen im folgenden diskutiert werden, um mit ihnen prüfen zu können, wie die Forschungsfrage im Rahmen dieser Arbeit beantwortet werden kann.

7.2 Diskussion der erlangten Ergebnisse

Um die Forschungsfrage zu beantworten, ist vor allem die praktische Evaluation von Bedeutung. Sie zeigt, wie das aus der Analyse der Anforderungen an eine Gestensteuerung für Assistenzrobo-

ter entwickelte Steuerungskonzept ebenjene Anforderungen erfüllen konnte. Dafür soll an dieser Stelle detailliert auf die einzelnen Probanden, ihre Ergebnisse und Verhaltensweisen eingegangen werden.

Zuerst einmal konnten alle Probanden mit dem Loomo und der Gestensteuerung umgehen. Anfangs sind mehrere von ihnen von einer anderen Funktionsweise ausgegangen. Sie haben anders als geplant gestikuliert und andere Positionen eingenommen. Nach einer Einstellung auf die umgesetzte Steuerung gelang es, den Roboter über Gesten zu bewegen und zu steuern. Dabei kam es aber in der Verarbeitung der Gesten zu Fehlern, was die Arbeit mit dem Roboter für die Nutzer erschwerte.

Mit nur neun benötigten Minuten hat der erste Proband die schnellste Zeit abgeliefert. Er hat sich ohne besonders lange Einweisung mit dem Roboter beschäftigt und die Gesten schnell und präzise ausgeführt. Dennoch kam es auch bei ihm zu mehreren Erkennungsfehlern. Besonders bei den nach hinten auszuführenden Gesten gab es Schwierigkeiten.

Am längsten hat der dritte Proband benötigt. Dieser hat zwischendurch die Gesten pausiert, um den Roboter genauer zu betrachten und Zwischenfragen gestellt. Dieses größere Interesse an der Funktionsweise des Roboters schlägt sich in einer längeren Gesamtzeit nieder. Auch bei den Gesten selbst hat er mehr ausprobiert als die anderen Tester. So hat er etwa mit verschiedenen Körperhaltungen experimentiert und den von der Tiefenkamera erfassten und vom Programm verwendeten Bereich mit seinen Bewegungen ausgereizt. Dabei kam es zu der für diesen Test größten Anzahl an Ausführungs- und Erkennungsfehlern. Trotz dieser Fehlerwerte konnte er aber mit dem Roboter arbeiten, sobald er sich auf die Aufgabe konzentrierte. Einziges Manko: er musste sich aufgrund seiner Größe generell stärker herunterbeugen, um in den erfassten Kamerabereich zu gelangen.

Die geringste Summe an Fehlern trat beim vierten Probanden auf. Dieser hat sich bei der Steuerung des Roboters durch den Kurs ebenfalls Zeit gelassen. Er hat sich nur so viel wie nötig bewegt und versucht, die Gesten so klar wie möglich auszuführen. Dieses vorsichtige Vorgehen hat für eine lange Zeit und zu guten Erkennungswerten geführt. Bei ihm war die Einhaltung der Abstände das größte Problem. Vermutlich aufgrund seiner geringeren Körpergröße und um die Erkennung der Gesten gut verfolgen zu können, hat er sich sehr nah am Loomo positioniert. Dies sorgte dafür, dass er über Sprachausgabe mehrfach aufgefordert wurde, sich anders zu positionieren.

Allgemein war es, gerade für einen solchen längeren Test, anstrengend, Körper und Arme in den benötigten Positionen zu halten. Dies hat zusammen mit den anfänglich nötigen Erklärungen und falsch erkannten Gestenbefehlen die benötigten Zeiten stark erhöht. Abschließend aber war es allen Testern möglich, des Parkour zu beenden. Sie haben dabei auch bekannt gegeben, dass sie die erforschte Gestensteuerung allgemein und in dieser Form für eine nützliche Alternative und Ergänzung zu den anderen mit dem Loomo möglichen Steuerungsmethoden halten.

7.3 Ausblick auf weitere Möglichkeiten zur Verwendung von Gestensteuerungen

Die ursprüngliche Forschungsfrage beschäftigt sich mit der Umsetzung einer allgemeinen Körpergestensteuerung mithilfe und für einen Assistenzroboter. Auf Basis der dabei entwickelten Lösung und Antwort auf die Forschungsfrage lassen sich weitere Fragen formulieren, die im folgenden kurz aufgezeigt werden sollen.

So wäre es interessant zu betrachten, wie sich die Steuerung in konkreten Anwendungsszenarien einsetzen lässt. Beispielsweise lässt sich auf Basis der hier implementierten Gestensteuerung eine Lösung für einen privaten oder öffentlichen Anwendungsraum erstellen. Dabei kann untersucht werden, wie die Steuerung genau auf die Herausforderungen der verschiedenen Kontexte angepasst werden muss, wie etwa Spezialisierung auf einen einzelnen Nutzer für private Nutzung oder Einstellen auf stets wechselnde Anwender in öffentlichen Gebäuden. Womöglich wären auch Erweiterungen des Loomo verwendbar. Mit diesen lässt sich der Funktionsumfang und damit die über Gesten theoretisch steuerbare Aktionsmenge erhöhen.

Dabei könnte auch auf besondere Nutzergruppen eingegangen werden. Krankheiten, Verletzungen, Behinderungen und andere Einschränkungen werfen besondere Herausforderungen für die Mensch-Roboter-Interaktion auf. Auch hier wäre die Nutzung von Gestensteuerungen in speziellen Anwendungsfällen untersuchbar.

Neben diesen Erweiterungen oder Anpassungen für das entwickelte Konzept sind auch Umsetzungen der anderen während der Konzeptionierung angeschnittenen Entwicklungsansätze denkbar. Auch diese könnten eine verwertbare Antwort auf die ursprüngliche Frage liefern, wie eine nutzbringende Gestensteuerung zu gestalten wäre. Wie, zum Beispiel, würde sich die Interaktion verändern, wenn zwischen Mensch und Roboter eine Distanz von mehr als zwei Metern befindet? Wie ließe sich der Roboter durch Fußbewegungen steuern? Lassen sich durch Hinzunahme von Farbwerten der Realsense bessere Ergebnisse erzielen? Wäre eine Steuerung auf Basis der HD-Kamera denkbar? Solche und ähnliche Fragen können mit Abwandlungen des entwickelten Konzeptes und mithilfe des Loomo bearbeitet werden.

Neben dem Loomo gibt es aber auch andere Assistenzroboter, die in der Praxis eingesetzt werden können. Lässt sich das entwickelte Steuerungskonzept auch mit diesen umsetzen? Hierfür müsste geprüft werden, ob auch unter Berücksichtigung der Eigenschaften und Fähigkeiten eines anderen Roboters eine auf Tiefenwerten und Handbewegungen basierende Körpergestensteuerung umsetzbar ist. Beispielsweise wäre eine Umsetzung auf Basis einer Kinect oder eines Roboters wie Pepper oder Nao denkbar. Wird dies umgesetzt, wäre auch hier eine Erweiterung oder Konkretisierung der Forschungsfrage auf bestimmte Anwender oder Anwendungsgebiete denkbar.

Die hier bearbeitete Forschungsfrage lässt sich also in verschiedene Richtungen fortsetzen und erweitern. Wie die konkrete Antwort auf die Frage aber lautet, ist Inhalt des letzten Abschnitts.

7.4 Fazit

Die Forschungsfrage lautete, wie eine Körpergestensteuerung gestaltet sein sollte, um in unterschiedlichen Szenarien eine geeignete Steuerung des Roboters zu erlauben. Diese Frage soll nun abschließend beantwortet werden, indem die einzelnen Schritte der Gestaltung und schließlich die evaluierte Implementierung beschrieben werden. Die Gestaltung der Steuerung entstand aus der Sensorik des für die Beantwortung verwendeten Roboters, dem Segway Robotics Loomo, heraus. Von diesem zeigte sich vor allem die Verwendung der Realsense-Tiefenkamera als sinnvoll. Die für eben diesen Roboter denkbaren Anwendungsszenarien sind vielfältig, da dieser neben der Bewegung auch Möglichkeiten zur Orientierung im Raum, Verfolgung des Nutzers und Interaktion über Sprache und ein UI bietet. Daher wurde ein breit gefächertes Spektrum an während der Gestensteuerung abrufbaren Handlungen entworfen.

Für diese Steuerung wurde aufgrund der ausgewählten Tiefensensoren eine Gestensteuerung

mittels statischer Handpositionen entwickelt. Diese löst die Einschränkungen, die aufgrund des Roboters berücksichtigt werden mussten. Dass der Nutzer sich nahe am Roboter aufhalten muss und nur bestimmte Körperbereiche erfasst werden, wird durch die auf die Arme und Hände spezialisierte Erfassung gelöst. Die Zusammensetzung aus diskreten Handpositionen und möglichst intuitiver Gruppierung der auslösbaren Handlungen soll die Nutzbarkeit für den Anwender optimieren, indem er so einfach wie möglich auf sämtliche möglichen Aktionen zurückgreifen kann. Das dabei entwickelte System aus Gesten als Handpositionen und zugewiesenen Aktionen des Roboters wurde mithilfe des Loomo praktisch als Android-Anwendung umgesetzt.

Wie gelungen diese prototypisch implementierte, mit Blick auf Einfachheit und Intuitivität gestaltete Umsetzung ist, wurde mithilfe einer Evaluation getestet. Bei dieser wurde eine breite Auswahl an Aktionen in einem Hinderniskurs als Anwendungsfall geprüft. Es hat sich gezeigt, dass abzüglich einiger durch die prototypische Implementierung gegebener Fehler die Gesten der Anwender erkannt und ausgeführt werden konnten. Die Nutzer haben eine kurze Einführung benötigt, konnten danach aber frei mit dem Roboter interagieren und die ihnen gestellten Anwendungsaufgaben erfüllen.

Damit bestätigt die Evaluation, dass die Forschungsfrage mit der hier konzeptionierten und umgesetzten Steuerung erfolgreich beantwortet werden konnte. Die verschiedenen getesteten Aktionen erlauben es, den Loomo in unterschiedlichen Szenarien zu verwenden. Aufgrund der Ergebnisse, die beim Test entstanden, wird klar, dass diese Steuerungsmethode auch in der Praxis geeignet sein kann. Damit stellt diese gestaltete Körpergestensteuerung eine geeignete Robotersteuerung dar.

Allgemein lässt sich also sagen, dass eine Körpergestensteuerung so gestaltet sein muss, dass sie sowohl die Möglichkeiten des Roboters ausschöpft als auch anwendbar für den Nutzer ist, um mit dem Roboter zu interagieren. Die Möglichkeiten des Roboters müssen berücksichtigt werden, um eine Basis für die Interaktion zu schaffen. Die Gesten müssen geeignet ausgewählt und sinnvoll ausgewertet werden. Schließlich muss es dem Nutzer leicht gemacht werden, mit dem Roboter umzugehen. So ist es möglich, die Gestensteuerung als Ergänzung zu anderen Steuerungskonzepten anzubieten. Ist dies geschafft, kann die Zugänglichkeit der Assistenzroboter gerade für die Personen, die sie am meisten benötigen, gewährleistet und verbessert werden.

Abkürzungsverzeichnis

CRR	Care Receiving Robot.....	6
DTS	Detection and Tracking System.....	27
SDK	Software Development Kit.....	8
TTS	Text to Speech.....	25
UI	User Interface.....	25

Abbildungsverzeichnis

2.1	mögliche Roboterunterscheidungen	3
4.1	Schema der Interaktion zwischen Nutzer und Loomo	18
4.2	Schema der Handverarbeitung und möglichen Positionen	20
5.1	Klassendiagramm der Anwendung	26
5.2	User Interface	28
6.1	Schema der Evaluationsumgebung	30

Tabellenverzeichnis

4.1	Auflistung der Gesten und zugehörigen Aktionen	22
6.1	Ergebnisse der Evaluation	31

Literatur

- [Ale+12] Jason Alexander, Teng Han, William Judd, Pourang Irani und Sriram Subramanian. “Putting your best foot forward: investigating real-world mappings for foot-based gestures”. en. In: *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. Austin, Texas, USA: ACM Press, 2012, S. 1229. DOI: 10.1145/2207676.2208575. URL: <http://dl.acm.org/citation.cfm?doid=2207676.2208575> (besucht am 14. 12. 2018) (siehe S. 17).
- [All+18] Dario Allegra, Francesco Alessandro, Corrado Santoro und Filippo Stanco. “Experiences in Using the Pepper Robotic Platform for Museum Assistance Applications”. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. Athens, Greece: IEEE, Okt. 2018, S. 1033–1037. DOI: 10.1109/ICIP.2018.8451777. URL: <https://ieeexplore.ieee.org/document/8451777/> (besucht am 09. 11. 2018) (siehe S. 4, 7).
- [AVD06] Alethea L. Blackler, Vesna Popovic und Douglas P. Mahar. “Towards a design methodology for applying intuitive interaction”. In: Lisbon, 2006 (siehe S. 19).
- [Beg+13] Momotaz Begum, Rosalie Wang, Rajibul Huq und Alex Mihailidis. “Performance of daily activities by older adults with dementia: The role of an assistive robot”. In: *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*. Seattle, WA: IEEE, Juni 2013, S. 1–8. DOI: 10.1109/ICORR.2013.6650405. URL: <http://ieeexplore.ieee.org/document/6650405/> (besucht am 23. 09. 2018) (siehe S. 4, 6).
- [Cha+18] Biplab Ketan Chakraborty, Debajit Sarma, M.K. Bhuyan und Karl F MacDorman. “Review of constraints on vision-based gesture recognition for human-computer interaction”. en. In: *IET Computer Vision* 12.1 (Feb. 2018), S. 3–15. DOI: 10.1049/iet-cvi.2017.0052. URL: <http://digital-library.theiet.org/content/journals/10.1049/iet-cvi.2017.0052> (besucht am 13. 09. 2018) (siehe S. 4).
- [Fan+17] Mingming Fan, Yizheng Ding, Fang Shen, Yuhui You und Zhi Yu. “An empirical study of foot gestures for hands-occupied mobile interaction”. en. In: *Proceedings of the 2017 ACM International Symposium on Wearable Computers - ISWC '17*. Maui, Hawaii: ACM Press, 2017, S. 172–173. DOI: 10.1145/3123021.3123043. URL: <http://dl.acm.org/citation.cfm?doid=3123021.3123043> (besucht am 14. 12. 2018) (siehe S. 17).
- [Gro+15] Horst-Michael Gross, Steffen Mueller, Christof Schroeter, Michael Volkhardt, Andrea Scheidig, Klaus Debes, Katja Richter und Nicola Doering. “Robot companion for domestic health assistance: Implementation, test and case study under everyday conditions in private apartments”. In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Hamburg, Germany: IEEE, Sep. 2015, S. 5992–5999. DOI: 10.1109/IROS.2015.7354230. URL: <http://ieeexplore.ieee.org/document/7354230/> (besucht am 14. 12. 2018) (siehe S. 4).

- [Lai+16] Yuhui Lai, Chen Wang, Yanan Li, Shuzhi Sam Ge und Deqing Huang. “3D pointing gesture recognition for human-robot interaction”. In: *2016 Chinese Control and Decision Conference (CCDC)*. Yinchuan, China: IEEE, Mai 2016, S. 4959–4964. DOI: 10.1109/CCDC.2016.7531881. URL: <http://ieeexplore.ieee.org/document/7531881/> (besucht am 13.09.2018) (siehe S. 7 f.).
- [Liy+12] Liying Cheng, Qi Sun, Han Su, Yang Cong und Shuying Zhao. “Design and implementation of human-robot interactive demonstration system based on Kinect”. In: *2012 24th Chinese Control and Decision Conference (CCDC)*. Taiyuan, China: IEEE, Mai 2012, S. 971–975. DOI: 10.1109/CCDC.2012.6242992. URL: <http://ieeexplore.ieee.org/document/6242992/> (besucht am 18.10.2018) (siehe S. 5).
- [MA07] Sushmita Mitra und Tinku Acharya. “Gesture Recognition: A Survey”. In: *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 37.3 (Mai 2007), S. 311–324. DOI: 10.1109/TSMCC.2007.893280. URL: <http://ieeexplore.ieee.org/document/4154947/> (besucht am 13.09.2018) (siehe S. 4).
- [MSW17] Tobias Mettler, Michaela Sprenger und Robert Winter. “Service robots in hospitals: new perspectives on niche evolution and technology affordances”. en. In: *European Journal of Information Systems* 26.5 (Sep. 2017), S. 451–468. DOI: 10.1057/s41303-017-0046-1. URL: <https://www.tandfonline.com/doi/full/10.1057/s41303-017-0046-1> (besucht am 14.12.2018) (siehe S. 4).
- [Qi +14] Qi Zhang, Baopu Li, Guo-qing Xu, Yimin Zhou, Ming Wang und Max Q.-H Meng. “Indoor environment applications for mobile robots using Kinect2.0”. In: *Proceeding of the 11th World Congress on Intelligent Control and Automation*. Shenyang, China: IEEE, Juni 2014, S. 1462–1466. DOI: 10.1109/WCICA.2014.7052934. URL: <http://ieeexplore.ieee.org/document/7052934/> (besucht am 18.10.2018) (siehe S. 5).
- [RCL14] David Rempel, Matt J. Camilleri und David L. Lee. “The design of hand gestures for human-computer interaction: Lessons from sign language interpreters”. en. In: *International Journal of Human-Computer Studies* 72.10-11 (Okt. 2014), S. 728–735. DOI: 10.1016/j.ijhcs.2014.05.003. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1071581914000706> (besucht am 14.12.2018) (siehe S. 18).
- [Rob18a] Segway Robotics. *Segway Robotics Loomo*. Englisch. Nov. 2018. URL: <http://www.loomo.com/en/> (siehe S. 8).
- [Rob18b] Softbank Robotics. *Softbankrobotics Nao*. Englisch. Nov. 2018. URL: <https://www.softbankrobotics.com/emea/en/nao> (siehe S. 4, 9).
- [Rob18c] Softbank Robotics. *Softbankrobotics Pepper*. Englisch. Nov. 2018. URL: <https://www.softbankrobotics.com/emea/en/pepper> (siehe S. 4, 9).
- [Sha+12] Syamimi Shamsuddin, Hanafiah Yussof, Luthffi Ismail, Fazah Akhtar Hanapiah, Salina Mohamed, Hanizah Ali Piah und Nur Ismarrubie Zahari. “Initial response of autistic children in human-robot interaction therapy with humanoid robot NAO”. In: *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*. Malacca, Malaysia: IEEE, März 2012, S. 188–193. DOI: 10.1109/CSPA.2012.6194716. URL: <http://ieeexplore.ieee.org/document/6194716/> (besucht am 09.11.2018) (siehe S. 4, 7).

-
- [Tan+15] Fumihide Tanaka, Kyosuke Isshiki, Fumiki Takahashi, Manabu Uekusa, Rumiko Sei und Kaname Hayashi. "Pepper learns together with children: Development of an educational application". In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. Seoul, South Korea: IEEE, Nov. 2015, S. 270–275. DOI: 10.1109/HUMANOIDS.2015.7363546. URL: <http://ieeexplore.ieee.org/document/7363546/> (besucht am 09.11.2018) (siehe S. 6).
- [Thr+99] S. Thrun, M. Bennewitz, W. Burgard, A.B. Cremers, F. Dellaert, D. Fox, D. Hahnel, C. Rosenberg, N. Roy, J. Schulte und D. Schulz. "MINERVA: a second-generation museum tour-guide robot". In: *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C)*. Bd. 3. Detroit, MI, USA: IEEE, 1999, S. 1999–2005. DOI: 10.1109/ROBOT.1999.770401. URL: <http://ieeexplore.ieee.org/document/770401/> (besucht am 23.09.2018) (siehe S. 4).
- [Van+11] Michael Van den Bergh, Daniel Carton, Roderick De Nijs, Nikos Mitsou, Christian Landsiedel, Kolja Kuehnlenz, Dirk Wollherr, Luc Van Gool und Martin Buss. "Real-time 3D hand gesture interaction with a robot for understanding directions from humans". In: *2011 RO-MAN*. Atlanta, GA, USA: IEEE, Juli 2011, S. 357–362. DOI: 10.1109/ROMAN.2011.6005195. URL: <http://ieeexplore.ieee.org/document/6005195/> (besucht am 18.10.2018) (siehe S. 7).
- [Yam+12] Kimitoshi Yamazaki, Ryohei Ueda, Shunichi Nozawa, Mitsuharu Kojima, Kei Okada, Kiyoshi Matsumoto, Masaru Ishikawa, Isao Shimoyama und Masayuki Inaba. "Home-Assistant Robot for an Aging Society". In: *Proceedings of the IEEE 100.8* (Aug. 2012), S. 2429–2441. DOI: 10.1109/JPROC.2012.2200563. URL: <http://ieeexplore.ieee.org/document/6218159/> (besucht am 23.09.2018) (siehe S. 6).
- [ZSW12] Jinhua Zeng, Yaoru Sun und Fang Wang. "A Natural Hand Gesture System for Intelligent Human-Computer Interaction and Medical Assistance". In: *2012 Third Global Congress on Intelligent Systems*. Wuhan, China: IEEE, Nov. 2012, S. 382–385. DOI: 10.1109/GCIS.2012.60. URL: <http://ieeexplore.ieee.org/document/6449559/> (besucht am 13.09.2018) (siehe S. 4).