

Assistenzrobotik für die Gesundheitsassistenz - ein Beitrag zur Evaluierung der Praxistauglichkeit am Beispiel eines mobilen Reha-Roboters¹

Assistive Robotics for Health Assistance - a Contribution towards Evaluating the Practicability by Example of a Mobile Rehab Robot

Horst-Michael Gross, Andrea Scheidig, Markus Eisenbach, Thanh Quang Trinh, Tim Wengefeld
 Technische Universität Ilmenau, Fachgebiet Neuroinformatik und Kognitive Robotik
 PF 100565, 98684 Ilmenau Email: horst-michael.gross@tu-ilmenau.de

Kurzfassung

Dieser Beitrag stellt für das Gebiet der sozialen Assistenzrobotik für die Gesundheitsassistenz einen kompakten Leitfaden für eine systematische Bestandsaufnahme implementierter Roboterservices und der zugrundeliegenden HRI- und Navigationsbasisleistungen im Rahmen von Funktions- und technischen Nutzertests zur Verfügung. Basierend auf diesem Leitfaden werden anschließend zwei Fallbeispiele aus dem Bereich der robotischen Gesundheitsassistenz bezüglich ihrer Praxistauglichkeit für den Einsatz in realen Seniorenwohnungen und in einem klinischen Umfeld evaluiert. Dabei wird insbesondere das BMBF-Projekt ROREAS (Interaktiver **RO**botischer **RE**ha-Assistent für das Lauftraining von Patienten nach Schlaganfällen) selbstkritisch analysiert, um die Stärken und Schwächen der entwickelten Assistenzlösung bezüglich ihrer erreichten Praxistauglichkeit im klinischen Einsatzfeld aus technischer Sicht deutlich werden zu lassen und damit eine ehrliche Bestandsaufnahme des Erreichten möglich zu machen.

Abstract

For the research field of socially assistive robotics for health assistance, this paper provides a compact guideline for a systematic evaluation of implemented services and underlying basic functionalities for HRI and navigation as part of function tests and technical user trials. Based on this guideline, subsequently two case studies from the field of robotic health assistance are evaluated with respect to their practical suitability for domestic and clinical application. In particular, the project ROREAS (**RO**botic **RE**habilitation Assistant for walking training of **St**roke patients) is analyzed in detail, in order to demonstrate the strengths and weaknesses of the developed assistive solution regarding its practicability for the clinical use from technical point of view to make an honest inventory of the achievements possible.

1 Motivation

In den letzten zehn Jahren gab es weltweit eine Fülle von Projekten im Bereich der sozialen Assistenzrobotik für Alltagsunterstützung und Gesundheitsassistenz. Dabei ist es selbst für "Insider" häufig schwierig, ein aussagekräftiges und realistisches Gesamtbild über das Funktions- und Servicespektrum und die Praxistauglichkeit der entwickelten Robotersysteme samt ihrer Servicefunktionen zu erhalten, da in wissenschaftlichen Publikationen meist zwar Angaben zu den methodisch-technischen Teilleistungen zu finden sind, hingegen nur selten oder gar nicht konkrete quantitative Ergebnisse von Funktions- und Nutzertests im realen Einsatzfeld mit echten Nutzern präsentiert werden. Häufig werden in den Publikationen sogar nur ausgewählte Teilfunktionen ohne die dazu erforderlichen robotischen Basisleistungen präsentiert, wie z.B. Roboter als "Vorturner" für Bewegungsübungen oder als "Medikamentenerinnerer", aber ohne die dazu notwendige, alltagstaugliche Navigation und Mensch-Roboter Interaktion (HRI). Zudem wird in Nut-

zerstudien durch "Wizard of Oz"-Experimente häufig eine Autonomie simuliert, die in dieser Form noch gar nicht existiert, was in der Öffentlichkeit oft zu erheblichen Fehleinschätzungen bezüglich der bereits erreichten Praxistauglichkeit führt. Gleichzeitig ist es auch auffallend, dass zwar häufig ein detailliertes technisches Benchmarking von Teilfunktionen (z.B. zur Nutzerwahrnehmung oder zur Hindernisvermeidung) vorgenommen wird, eine Evaluierung des Roboter-Gesamtsystems inklusive der angebotenen Services unter realen Einsatzbedingungen meist aber noch die seltene Ausnahme ist. Dies führt vor allem bei Außenstehenden zu einer bedenklichen Situation, die von deutlichen Zweifeln an der Robotik bis zur völligen Überschätzung ihrer derzeitigen Möglichkeiten geprägt ist. Ursache dafür ist der nicht immer aufrichtige Umgang der Entwickler bezüglich der Autonomie und der erreichten Robustheit und Praxistauglichkeit ihrer Lösung als Ganzes sowie der Systemstabilität im Realeinsatz, oft gepaart mit fehlendem Benchmarking und unzureichender Systemevaluierung in Funktionstests mit realen Endnutzern.

¹Gefördert im Projekt ROREAS durch das Programm "IKT 2020 - Forschung für Innovationen" des BMBF (Förderkennzeichen 16SV6133)

2 Evaluierung der Praxistauglichkeit von Assistenzrobotern

Ziel dieses Beitrages ist es, für zukünftige Publikationen im Bereich der sozialen Assistenzrobotik für den häuslichen und klinischen Einsatz einen kompakten Leitfaden für eine systematische und ehrliche Bestandaufnahme getesteter Roboterservices und der zugrundeliegenden Basisleistungen zur Verfügung zu stellen. Dieser Leitfaden erfragt möglichst präzise Angaben zur eigentlichen Funktionalität des Roboters und zur Durchführung der Funktions- und Nutzertests unter Realwelt-Bedingungen. Dabei lassen sich vier Themenkomplexe zur Bewertung der Praxistauglichkeit definieren:

Komplex S - Servicespektrum:

- S1:** Welche *Serviceleistungen* für die Nutzer sind bereits verfügbar? Welche davon a) völlig autonom auf dem Roboter, b) nur mit geeigneter Peripherie (z.B. externe Sensoren, Kameras oder Mikrofone), c) nur über Fernsteuerung (remote control) durch einen telepräsenten Operator?
- S2:** Welche *Basisleistungen* der Navigation und HRI stehen dem Roboter dafür in welcher Autonomiestufe zur Verfügung?
- S3:** Wie sieht dafür die notwendige *IT-Infrastruktur* vor Ort aus? Welche Konsequenzen ergeben sich daraus für den späteren Praxiseinsatz?

Komplex R - Reifegrad:

- R1:** Wie ist der *Reifegrad* des getesteten Robotersystems - handelt es sich noch um ein Labormuster/einen Demonstrator, bereits einen Prototypen oder schon um ein am Markt verfügbares Produkt?
- R2:** Gab es bereits *Funktionstests* außerhalb des Labors (wann, wo, mit wem, wie oft, wie lange, unter welchen Bedingungen)?
- R3:** Gab es bereits *Nutzertests* mit den Endnutzern in der finalen Einsatzumgebung (wann, wo, mit wem, wie oft, wie lange, unter welchen Bedingungen)?
- R4:** War dabei technisches oder sozialwissenschaftliches *Begleitpersonal* zugegen, und wo hielt es sich während der Tests auf?
- R5:** Über welchen *Zeitraum* war der Roboter für den Nutzer verfügbar, wie war die Nutzungsquote?

Komplex F - Funktionstests von Basisleistungen: Hier werden die szenariospezifisch notwendigen Einzelleistungen, die Basisleistungen der Navigation und der Mensch-Roboter-Interaktion (HRI), funktionell getestet und möglichst auch quantitativ bewertet. Leistungsvergleiche der Einzelleistungen (Benchmarks) unter Laborbedingungen und in realer Einsatzumgebung sind anzustreben, um die Besonderheiten und Schwierigkeiten der Einsatzumgebung zu beurteilen.

- F1:** Welche *Navigationsleistungen* wurden getestet und unter welchen konkreten Testbedingungen?
- F2:** Sind dabei *Navigationsprobleme* aufgetreten und wie wurden sie quantitativ erfasst, z.B. als Anzahl der Verklemmungen, Kollisionen, Beinahekollisionen, Fehllokalisationen, Verletzungen des Personal Space, etc. ?
- F3:** Welche *HRI-Leistungen* wurden getestet und unter welchen konkreten Testbedingungen?
- F4:** Sind dabei *HRI-Fehlfunktionen* (z.B. der Personendetektion, des Personentrackings, der Nutzerwiedererkennung, etc.) aufgetreten?
- F5:** Welche *Erfolgsraten* der Basisleistungen wurden dabei als quantitative Gütemaße ermittelt (z.B. Zielerreichung, Lokalisierungsgenauigkeit, Nutzerdetektion, Nutzersuche, Nutzerwiedererkennung, etc.)?
- F6:** Gab es notwendige *manuelle Eingriffe* vor und während der Tests (z.B. Markierung von No-Go Areas, Präparierung von Hindernissen, Not-Stopps, Ablaufänderungen, etc.)?
- F7:** Wurde eine quantitative Bewertung der *Komplexität* der Testumgebung vorgenommen?

Komplex N - Nutzertests auf Applikationsebene: Hier interessieren der Grad der Autonomie und die Praxistauglichkeit der Anwendung im Zusammenspiel der einzelnen Basisleistungen. Wir haben dazu drei Fehlermaße für die Qualitätsbeschreibung der Tests auf Applikationsebene definiert:

- N1:** *Unkritische Fehler:* Fehler, die autonom durch die Applikation selbst behandelt werden können, ggf. auch unter Beteiligung des Endnutzers (z.B. Anfahren eines Ruhepunktes bei einem Kontaktverlust zum Nutzer oder selbständiges Beenden eines Bumper-Stopps bei Wissen über Dauer und Lage des auslösenden Ereignisses)
- N2:** *Kritische Fehler:* Fehler, die nur durch das Eingreifen eines Operators behebbar sind (z.B. die Korrektur einer fälschlichen Personenhypothese)
- N3:** *Sehr kritische Fehler:* Fehler, welche nicht durch den Eingriff eines Operators behebbar sind, wie z.B. Sensorausfälle

Die Einstufung der Fehler muss immer im Kontext der szenariotypischen Anforderungen erfolgen.

Ein wichtige Voraussetzung für Funktions- und Nutzertests unter Realwelt- und Praxisbedingungen ist zudem eine objektive Bewertung der Komplexität der Einsatzumgebung bezüglich ihrer Größe, der zeitlich variablen Präsenz von Personen, der vorherrschenden Lichtverhältnisse, der Möblierung und der damit verbundenen "Beengtheit" mit Konsequenzen für die Navigation sowie Interaktion im verfügbaren Freiraum (↗ F7). Allerdings werden dazu in vielen Publikationen bislang höchstens subjektive, qualitative Einschätzungen, wie z.B. "sehr

eng, kaum Freiraum”, “kaum Tageslicht”, “stark bevölkert”, jedoch keine objektiven, quantitativen Kennzahlen verwendet. Ein erster Ansatz zur Objektivierung der Tests und zur quantitativen Bewertung der Komplexität der Testumgebungen wurde für das häusliche Einsatzfeld bereits in [1, 2] präsentiert, mit dem Ziel zukünftig auch vergleichende Benchmarks in verschiedenen Einsatzumgebungen möglich zu machen.

Im nachfolgenden Beitrag wird es nach den entsprechenden Erläuterungen jeweils direkte Verweise auf diesen Fragenkatalog geben, wie zum Beispiel (↗ S1).

3 Eigene Fallbeispiele

3.1 Gesundheitsassistenz SERROGA



Abbildung 1: Gesundheitsassistent “Max” in Interaktion mit einer Seniorin in deren privater Wohnung im Rahmen des SERROGA-Projektes.

Basierend auf diesem Leitfaden wurden bereits die im Rahmen der Forschergruppe SERROGA (SERVICE-ROBOTIK für die GesundheitsAssistenz) [3] entwickelte robotergestützte Gesundheitsassistenz bezüglich ihrer Praxistauglichkeit für den Einsatz in realen Seniorenwohnungen evaluiert. Die im Rahmen von SERROGA entwickelte Herangehensweise an die Durchführung von technischen Nutzertests in Seniorenwohnungen wurde erstmals in [4] und in erweiterter Form in [1] vorgestellt.

In Vorbereitung der mehrtägigen Nutzertests mit dem entwickelten Demonstrator (↗ R1) erfolgten zunächst in einer wohnungsähnlichen Laborumgebung und darauf aufbauend in Mitarbeiterwohnungen Funktionstests zu den notwendigen robotischen Basisleistungen (↗ R2). Dabei wurden solche Verhaltensweisen wie die autonome Zielanfahrt, das automatische Andocken an die Ladestation, das Dem-Nutzer-Folgen und das Suchen des Nutzers bewertet (↗ F1, F3). Neben dem Nachweis der grundsätzlichen Funktionsfähigkeit wurden auch objektive Bewertungsmaße für die Güte dieser Basisleistungen bestimmt (↗ F5).

Bei den Funktionstests in verschiedenen Wohnumgebungen war zunächst zu klären, wie realistisch dort vorhandene Bedingungen sind, und ob die erreichten Ergebnisse auch auf andere Wohnungen übertragbar sind. Daher wurden zunächst die räumlichen Eigenschaften der Wohnungen anhand von navigationsrelevanten Kenngrößen

gegenübergestellt (↗ F7). Diese Kenngrößen ließen sich unmittelbar aus den vorher gelernten Navigationskarten ermitteln. Exemplarisch seien hier nochmals einige der in [4] vorgestellten Kenngrößen genannt: a) Grundfläche A: gesamte von Wänden eingeschlossene Fläche der durch den Roboter befahrbaren Räume; b) Freifläche F: Grundfläche A ohne Möbel; c) befahrbare Fläche B: durch den Roboter kollisionsfrei erreichbare Positionen auf der Freifläche, d) Flächenabdeckung B/A: Verhältnis von befahrbarer Fläche zur Grundfläche als Maß, wie zugestellt eine Umgebung ist; e) Formfaktor Z: Verhältnis des Umfangs der befahrbaren Fläche B zur Wurzel dieser Fläche als Maß für die Konvexität bzw. Zerklüftetheit; f) Wandabstand W: mittlerer Abstand zum nächsten Hindernis über der gesamten befahrbaren Fläche; g) Durchfahrtsbreite D: mittlerer Abstand von der Mitte des befahrbaren Raumes oder einer Durchfahrt zum nächsten Hindernis. Für die nutzerzentrierten Navigationsverhalten Folgen und Suchen war auch die Beleuchtungssituation der Wohnung zu berücksichtigen, da die Nutzerwahrnehmung u.a. bildbasiert über Detektionen im frontalen Fischaugenkamerabild der SERROGA-Roboter erfolgt. In den Tests wurden daher auch erstmals die wohnungsspezifischen Beleuchtungsbedingungen erfasst, um deren Einfluss auf die Erkennungsleistungen des Roboters zu objektivieren und vergleichbar zu machen.

Nachdem mit den durchgeführten Funktionstests nachgewiesen wurde, dass sich die SERROGA-Roboter stabil autonom in Wohnungen auch im Zusammenspiel mit Personen bewegen konnten, fanden im Zeitraum Februar bis April 2015 mit den beiden in SERROGA entwickelten mobilen Assistenzrobotern Max und Tweety abschließende Mehrtagesnutzertests mit vier Senioren der AWO Servicewohnanlage “Am Krämpferufer” und mit fünf Senioren der ARTIS Servicewohnanlage Erfurt im Alter von 68 bis 92 Jahren in deren eigenen Wohnungen statt (↗ R3). Mit einer Gesamttestzeit von 220 Stunden ohne Anwesenheit von Fremdpersonen in den Wohnungen gilt dies als einer der bislang längsten Praxistests von autonomen mobilen Assistenzrobotern bei Senioren in deren privaten Wohnungen (↗ R4, R5) (Stand 4/2015). Diese Nutzertests bestätigten, dass die in SERROGA entwickelten Assistenzroboter trotz sehr unterschiedlicher Einsatzbedingungen über einen mehrtägigen Zeitraum stabil funktionieren und durch die Senioren gut zu bedienen waren. Es zeigt sich aber auch, dass die Länge der Tests von ein bis drei Tagen noch nicht ausreichte, um eine gewisse Alltagsnormalität zu erreichen. Nutzungsintensität, Nutzungsmuster und die Bewertung des Nutzungserlebnisses mussten daher im Kontext eines außer-alltäglichen Erlebnisses interpretiert werden und ließen nur Tendenzen für eine reale Alltagsnutzung erkennen. Detaillierte Ergebnisse der Funktions- der Nutzertests sind [1, 4] zu entnehmen.

3.2 Rehabilitationsassistent ROREAS

Basierend auf dieser Herangehensweise soll nachfolgend eines unserer Robotikprojekte aus dem Gebiet der kli-

nischen sozialen Assistenzrobotik, das ROREAS-Projekt [5, 2, 6, 7]), so aufgearbeitet werden, dass die Stärken und Schwächen der entwickelten Assistenzlösung bezüglich ihrer erreichten Praxistauglichkeit im klinischen Einsatzfeld aus technischer Sicht deutlich werden. Dabei kommen auch wiederum Gütemaße zur Bewertung des gezeigten Navigations- und Interaktionsverhaltens der Roboter zum Einsatz.

In [2] wurde das seit Mitte 2013 laufende BMBF-Projekt ROREAS (Interaktiver ROBotischer REHa-Assistent für das Lauf- und Orientierungstraining von Patienten nach Schlaganfällen) vorgestellt. Dessen Zielstellung ist die Entwicklung eines robotischen Reha-Assistenten für ein Eigentaining von Patienten zur Durchführung eines Lauf- und später auch Orientierungstrainings in der klinischen Schlaganfallnachsorge. Von Projektbeginn an wurde die Technikentwicklung durch die Einbindung der Techniknutzer - in diesem Fall Schlaganfallpatienten - begleitet. Unter Berücksichtigung medizinischer, technischer und sozialwissenschaftlicher Aspekte wurden für den robotischen Reha-Assistenten unterschiedliche Szenarien definiert, von denen bislang der Laufcoach zum Lauftraining praktisch realisiert und in Funktions- und Nutzertests bewertet wurde. Der Laufcoach wurde dabei durch solche Reha-Patienten genutzt, die bereits die Erlaubnis erhalten hatten, mithilfe eines Laufhilfsmittels selbständig zu gehen. Insbesondere sollte er ängstliche, antriebsgestörte oder selbstunsichere Patienten unterstützen, sich zunächst auf dem Gang und dann in einem immer größer werdenden Umfeld in der Klinik sicher zu bewegen (↗ S1). Details zum Ablauf einer typischen Trainingseinheit mit dem Laufcoach wurden bereits in [2] vorgestellt, in Kurzform findet sich dies auch noch einmal im folgenden Abschnitt.

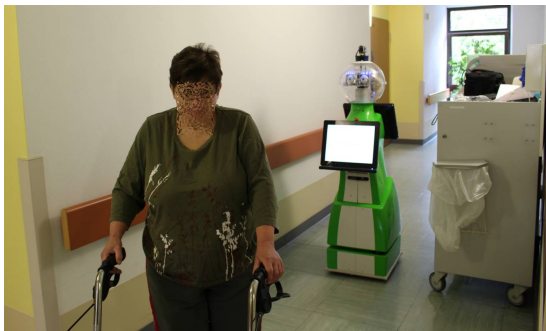


Abbildung 2: Robotischer Reha-Assistent in der Funktion eines Laufcoach-Demonstrators (↗ R1) basierend auf einem SCITOS G3 Roboter der Firma MetraLabs GmbH Ilmenau während eines Gangtrainings in der m&i-Fachklinik Bad Liebenstein (Deutschland).

4 Wie praxistauglich ist ROREAS?

4.1 Dreistufige Teststrategie

Im Rahmen des ROREAS-Projektes wurde für den entwickelten Demonstrator (↗ R1) eine dreistufige Teststrategie angewandt, mit der die nötigen Funktions- und Nutzertests in der Einsatzumgebung der Reha-Klinik unter realen Alltagsbedingungen konzipiert und durchgeführt wurden. Aus Sicherheitsgründen erfolgte in *Stufe 1* noch keine Interaktion des Laufcoachs mit realen Patienten vor Ort, stattdessen wurden sie zur Untersuchung aller notwendigen Basisverhalten und Skills (siehe Abb. 3) ersetzt durch eingewiesene Mitarbeiter unseres Robotik Labs. Darauf aufbauend wurde in *Stufe 2* der Laufcoach mit sogenannten Patienten-Doubles, Projektfremde oder klinische Mitarbeiter, die das Laufverhalten von Schlaganfallpatienten imitierten, getestet (↗ R2, R3). Erst nachdem dies erfolgreich war und sich ROREAS als stabil und genügend zuverlässig erwiesen hatte, wurden in *Stufe 3* Nutzertests mit freiwilligen Schlaganfall-Patienten begonnen.

Im Rahmen dieses Beitrages soll anhand der erreichten methodisch-technischen Ergebnisse der Nutzertests die Frage diskutiert werden, ob der Reha-Assistent Laufcoach aus technischer Sicht im klinischen Einsatzfeld bereits autonom agieren kann. Möglich wird dies durch das stabile Funktionieren wesentlicher funktionaler Elemente des Trainingsablaufs, unabhängig von der Komplexität des Trainingsprogramms. Ein typischer Trainingsablauf mit dem Laufcoach umfasst folgende Teilabschnitte (↗ S2):

Im Rahmen dieses Beitrages soll anhand der erreichten methodisch-technischen Ergebnisse der Nutzertests die Frage diskutiert werden, ob der Reha-Assistent Laufcoach aus technischer Sicht im klinischen Einsatzfeld bereits autonom agieren kann. Möglich wird dies durch das stabile Funktionieren wesentlicher funktionaler Elemente des Trainingsablaufs, unabhängig von der Komplexität des Trainingsprogramms. Ein typischer Trainingsablauf mit dem Laufcoach umfasst folgende Teilabschnitte (↗ S2):

- TA1:** Fahrt des Roboters zum Patientenzimmer auf eine geeignete Warteposition und Warten auf die Kontaktaufnahme durch den Patienten
- TA2:** Kontaktaufnahme des Patienten mit dem Roboter durch Hinsetzen auf die Startposition (ein Stuhl vor dem Zimmer)
- TA3:** Heranfahren des Roboters an den Patienten zur Initiierung der Interaktion
- TA4:** Dialog zwischen Roboter und Patient und Auswahl des Trainingsprogramms, Lernen des aktuellen Nutzermodells nach Aufstehen vom Stuhl
- TA5:** Begleiten des Patienten (Folgen) bis zur erneuten Kontaktaufnahme durch den Patienten durch Hinsetzen (an Ruhepunkten oder am Ziel)
- TA6:** Interaktion mit dem Patienten an den Ruhepunkten und am Ziel getriggert durch Sitzenderkennung
- TA7:** Auswertung der Trainingsergebnisse und Abschlussdialog vor dem Zimmer des Patienten

Zur Umsetzung dieses Trainingsablaufs werden zahlreiche robust funktionierende Basisleistungen benötigt, in der finalen Ausbaustufe müssen alle völlig autonom funktionieren. Die wichtigsten von ihnen wurden im Rahmen der Funktions- und Nutzertests zwecks kontinuierlicher Optimierung immer wieder technisch evaluiert, um eine zunehmend größere Autonomie zu erreichen. Hierzu zählen (siehe Abb. 3, orange hinterlegt):

- Zielfahrten

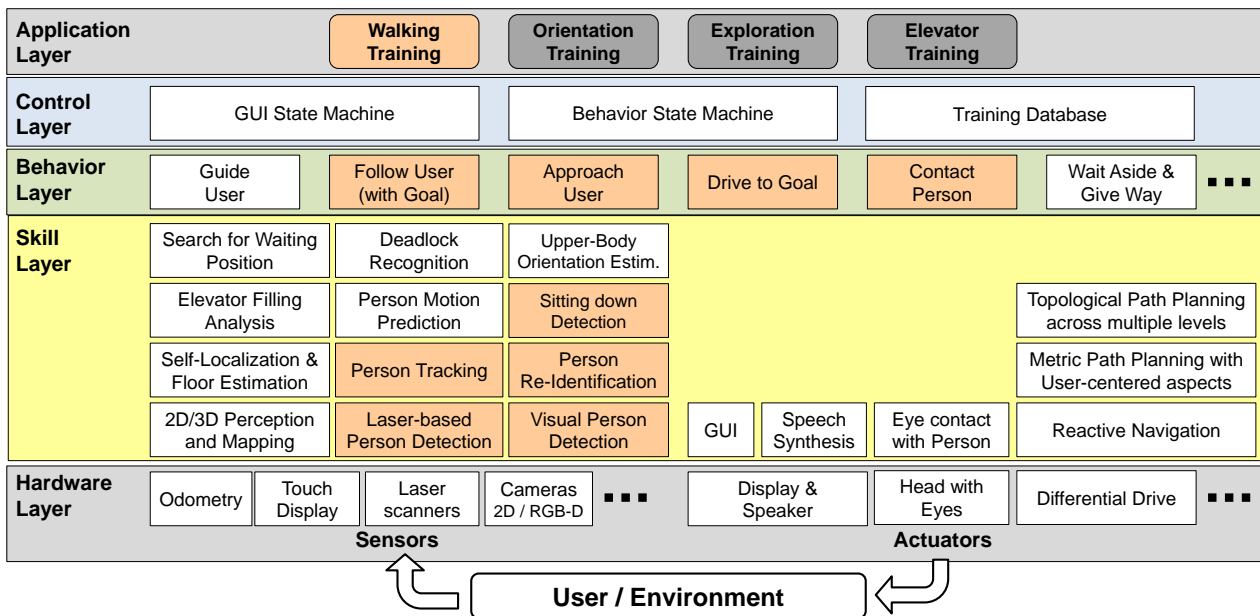


Abbildung 3: Übersicht über die benötigten Basisleistungen des Demonstrators “Laufcoach” (aus [5, 2, 6]). In der Application Layer grau hinterlegt sind weitere mögliche Demonstratoren, die bislang aber noch nicht realisiert wurden. Der Laufcoach wird durch zwei State Machines (Control Layer) und kombinierte Basisverhalten (Behavior Layer) umgesetzt, die dazu eine Vielzahl von robust funktionierenden Erkennungs-, Navigations- und Interaktionsleistungen benötigen (Skill Layer) (↗ S2). Sowohl die Behaviors als auch die Skills werden nachfolgend als *Basisleistungen* bezeichnet, die in Funktions- und technischen Nutzertests auf ihre Praxistauglichkeit bewertet werden müssen.

- Erkennung des Hinsetzens des Patienten
- Heranfahen an den Patienten zwecks Erreichbarkeit des Touch-Displays
- Personentracking und Personenwiedererkennung
- Begleiten des Patienten durch Hinterherfahren ohne Unterschreiten eines minimalen und Überschreiten eines maximalen Begleitabstandes
- höfliche Navigation gegenüber Passanten ohne Verletzung ihres Personal Space

Zur Korrektur noch fehlerhafter Entscheidungen der Basisleistungen und damit zur möglichst unterbrechungsfreien Durchführung der Nutzertests wurde im Rahmen von ROREAS ein tabletbasiertes Korrekturinterface (Control Tablet) entwickelt, das einem Testbeobachter manuelle Eingriffsmöglichkeiten bietet. Das über WLAN angebundene Control Tablet erlaubt ihm, mit ausreichend Abstand zum Roboter (>5 m) die Tests zu begleiten (↗ R4), so dass er das Verhalten des Roboters nicht beeinflusst und trotzdem auf Ereignisse und beobachtbare Probleme reagieren kann. Dadurch konnte noch vor Abschluss der Algorithmenentwicklung und der Funktionstests und damit deutlich frühzeitiger als sonst üblich mit Nutzertests in der realen Umgebung begonnen werden, was für den engen zeitlichen Projektlauf wichtig war. Außerdem erhielten die Entwickler durch die mittels des Control Tablets protokollierten Eingriffe des Beobachters ein sehr spezifisches und objektives Feedback, in welcher Situation die Basisleistungen noch Probleme haben.

Gleichzeitig stand so ein weiteres Gütemaß für das autonome Agieren des Reha-Assistenten zur Verfügung, in dem die Anzahl notwendiger Ferneingriffe pro Wegstrecke erfasst wurde (↗ F6). Für die spätere Auswertung wurden in allen Tests auf dem Roboter zudem Logdaten mitgeschnitten. Diese speicherten den inneren Zustand des Roboters und ausgewählte Sensordaten, um ein Fehlverhalten im Nachgang nachvollziehen und interpretieren zu können.

4.2 Stufe 1: Funktionstests mit Projektmitarbeitern

Bevor der Laufcoach in Nutzertests mit Patienten evaluiert werden konnte, mussten zunächst alle grundlegende Basisverhalten in der klinischen Einsatzumgebung getestet und bewertet werden. Erste umfangreiche Funktionstests erfolgten mit ROREAS in der m&i Fachklinik in Bad Liebenstein im Februar 2015 über 4 Tage, wobei ca. 15.000 m Fahrtstrecke zurückgelegt wurden. Die Funktionstests wurden in verschiedenen Etagen der Klinik und zu unterschiedlichen Tageszeiten durchgeführt, um die Basisverhalten unter möglichst vielfältigen Klinikbedingungen zu bewerten, wie z.B. räumliche Charakteristik, Beleuchtungsverhältnisse oder Personendichte auf den Gängen. Ein detaillierter Überblick über die während der Funktionstests getesteten Basisverhalten autonome Zielanfahrt, Lotsen und Folgen einer Person im Kontext der Nutzung der Skills 2D-Umgebungserfassung, 3D-Umgebungserfassung, Personal Space, Engstellenbehandlung, Rechtsfahren und Personenwiedererkennung

(↗ F1, F3) und die ermittelten Erfolgsraten (↗ F5) wurde bereits in [2] gegeben und soll daher hier nicht wiederholt werden. Manuelle Eingriffe per Control Tablet erfolgten bei diesen Tests nur bei der Wiedererkennung und bei Not-Stopps vor möglichen Kollisionen (↗ F6). Im Unterschied zur SERROGA (siehe Abschnitt 3.1) konnte wegen der Weitläufigkeit und der Erscheinungsvielfalt der Klinik bislang allerdings noch keine quantitative Bewertung der Komplexität der Einsatzumgebung vorgenommen werden (↗ F7).

4.3 Stufe 2: Nutzertests mit Patienten-Doubles

Nach erfolgreichem Abschluss der Funktionstests mit Projektmitarbeitern wurde der Laufcoach erstmals im Mai 2015 und später erneut vor jedem Nutzertest im Rahmen von simulierten Nutzertests mit Patienten-Doubles getestet (↗ R2, R3). Diese Tests unter Einbeziehung von Klinikmitarbeitern und Projektfremden, die das Laufverhalten der Patienten imitierten und dabei auch Laufhilfsmittel benutzten, waren sozusagen die Generalprobe für die noch ausstehenden Nutzertests. Getestet wurden dabei neben der Stabilität der benötigten HRI- & Navigationsbasisleistungen (↗ F1, F3) vor allem die eigentliche Trainingsapplikation und die Schlüssigkeit und Verständlichkeit des Trainingsablauf sowie die Notwendigkeit manueller Eingriffe durch den Beobachter mittels des Control Tablets (↗ F6). Im Fokus dieser Tests standen damit erstmals der Grad der Autonomie und die Praxistauglichkeit der Anwendung im Zusammenspiel der einzelnen Basisleistungen mit Blick auf die Qualitätsbeschreibung der Tests auf Applikationsebene. Wie bei den Funktionstests war außerdem ein externer Versuchsbeobachter (Observer) anwesend (↗ R4), der mit ausreichend Abstand zum Roboter die Tests begleitete und die noch aufgetretenen technischen Probleme dokumentierte. Eine detaillierte quantitative Erfassung der Erfolgsraten der Basisleistungen (↗ F5) und der ermittelten Fehler auf Applikationsebene (↗ N1-N3) erfolgte aus Beherrschbarkeitsgründen auf dieser Stufe jedoch nicht sondern war Gegenstand der folgenden Nutzertests.

4.4 Stufe 3: Nutzertests mit Patienten

Aufbauend auf den Tests mit Patienten-Doubles wurden im weiteren Projektverlauf im Zeitraum von Juni 2015 bis März 2016 fünf Kampagnen von Nutzertest durchgeführt (↗ R3). Während die Nutzertests im Juni und September 2015 noch kürzere und fest vorgegebene Laufwege mit einer Trainingsdauer von ca. 20 Minuten umfassten, wurde der Laufcoach für die Nutzertests im November 2015 und im Januar 2016 so erweitert, dass auch beliebige, durch den Nutzer frei wählbare Laufstrecken möglich waren und eine Trainingseinheit, je nach Fitness des Patienten, bis zu einer Stunde dauern konnte (↗ R5). In allen Nutzertests hatte der Testbeobachter die Möglichkeit, per Ferneingriff noch fehlende Basisleistungen (z.B. Sitzenderkennung) zu emulieren oder noch fehlerhafte Erkennungsleistungen zu korrigieren. Ziel war es, von Nutzer-
test zu Nutzertest die Anzahl der Korrekturingriffe zu reduzieren und die Autonomie des Systems zu erhöhen. Daher wird in den nachfolgend beschriebenen Nutzertests der Fokus der Auswertung auf die jeweils noch fehlende Autonomie, d.h. auf die noch notwendigen Ferneingriffe, gelegt, um eine ehrliche Problemdiskussion zu ermöglichen. In allen Nutzertests kamen als freiwillige Tester nur solche Schlaganfallpatienten zum Einsatz, die von ihrem behandelnden Arzt bereits die Freigabe erhalten hatten, in Eigenverantwortung ein Lauftraining als Eigen-
training zu betreiben. Auf die Details der Einweisung der Patienten am Tage der Nutzertests und die Durchführung von Probedurchläufen (Instruktionsgänge) soll hier allerdings nicht weiter eingegangen werden.

Leistungen zu korrigieren. Ziel war es, von Nutzer-
test zu Nutzertest die Anzahl der Korrekturingriffe zu reduzieren und die Autonomie des Systems zu erhöhen. Daher wird in den nachfolgend beschriebenen Nutzertests der Fokus der Auswertung auf die jeweils noch fehlende Autonomie, d.h. auf die noch notwendigen Ferneingriffe, gelegt, um eine ehrliche Problemdiskussion zu ermöglichen. In allen Nutzertests kamen als freiwillige Tester nur solche Schlaganfallpatienten zum Einsatz, die von ihrem behandelnden Arzt bereits die Freigabe erhalten hatten, in Eigenverantwortung ein Lauftraining als Eigen-
training zu betreiben. Auf die Details der Einweisung der Patienten am Tage der Nutzertests und die Durchführung von Probedurchläufen (Instruktionsgänge) soll hier allerdings nicht weiter eingegangen werden.

4.4.1 Nutzertest Nr. 1 im Juni 2015

In diesem ersten Nutzertest mit freiwilligen Patienten waren bereits alle Teilabschnitte des Trainingsablaufs (siehe Abschnitt 4.1) zu absolvieren, allerdings mit unterschiedlicher Autonomie und Unterstützung durch Ferneingriffe (↗ S1, F6). Dabei ergab sich folgender Stand:

Teilabschnitt	Erreichte Autonomie
TA1: Fahrt zum Zimmer	100% autonom
TA2: Sitzenderkennung	100% per Ferneingriff
TA3: Heranfahen an Patienten	100 % autonom
TA4: Dialog und Auswahl	nur 1 Trainingsstrecke
TA5: Begleiten des Patienten	teilautonom, tw. mit Ferneingriffen
TA6: Interakt. an Ruhepunkten	100% per Ferneingriff
TA7: Auswertung/Verabschied.	100% autonom

Außerdem wurden die Testbedingungen bewusst noch vereinfacht und in die ruhigen Nebenzeiten (nachmittags und abends) gelegt, um die Störungen durch unbeteiligte Passanten zu minimieren. Die Trainingsapplikation war noch stark eingeschränkt, und die Trainingsstrecken waren kurz. Wichtige Angaben zu den Ergebnissen dieses technischen Nutzertests sind der nachfolgenden Tabelle zu entnehmen (↗ R5):

Kriterien	Angaben
Zeitraum:	6/2015, 2 Tage
Anzahl Patienten:	5
Anzahl Durchläufe:	11 (Gesamt)
Verwend. Gehhilfsmittel:	Rollatoren
Gefahrene Gesamtstrecke:	873 m
Trainingszeit:	62 min
Mittlere Dauer pro Pat.:	5 min 30 s
Anzahl Passanten:	78
Eingriffe per Contr. Tablet:	19

Navigationsprobleme (↗ F2): Während dieses Tests berührte der Roboter 3 mal Hindernisse und musste durch einen manuellen Eingriff aus dieser Situation befreit werden (↗ N3). Gründe für diese Form der Kollision waren Drehbewegungen des Roboters in der Nähe von Wänden und Berührungen der dort verankerten Handläufe, die von den Laserscanner oder den 3D-Kameras des Roboters aufgrund ihrer Befestigungshöhe nicht erfasst wer-

den konnten. Diese Probleme waren in den vorangegangenen Funktionstests nicht aufgetreten und kamen daher überraschend. 2 mal fuhr der Roboter sehr nah an Personen heran (< 15 cm) und verletzte damit ihren Personal Space. Gründe dafür waren plötzlich aus Zimmern heraustretende Personen und eng überholende Passanten. 19 mal musste der Beobachter per Ferneingriff reagieren (↗ F6), um unsichere oder fehlerhafte Hypothesen des Wiedererkennungsmoduls bei Abrissen des Personentrackings zu bestätigen bzw. zu korrigieren (↗ N2). Nur in zwei Fällen waren es falsche Hypothesen, alle anderen waren zu unsicher und erforderten lediglich Bestätigung. Hier wurde deutlich, dass die kleidungsbasierte Wiedererkennung von Patienten gegenüber den Tests mit Mitarbeitern und Patienten-Doubles eine höhere Schwierigkeit aufwies.

Die Ursachenforschung für die beobachteten Fehler ergab weiterhin, dass bestimmte Fehler im Ablauf der Applikation noch nicht korrigiert werden konnten (↗ N3), die Erfassungsbereiche der bislang verwendeten Panoramakamera auf dem Kopf zu gering waren, so dass als Zwischenlösung für die nächsten Nutzertests eine zusätzliche Frontkamera am Display montiert werden sollte, und dass das Heranfahren an Patienten noch suboptimal war, da die Abstände für eine gute Bedienbarkeit des Touch-Displays zu groß waren.

Zwischenfazit: Es wurde in diesem Test sehr deutlich, dass das Trainingsszenario unbedingt erweitert werden muss im Sinne längerer Trainingsstrecken und dem Bereitstellen von Auswahlmöglichkeiten, da der Ablauf für die Patienten zu wenig herausfordernd war. Außerdem sollten weitere Korrekturmöglichkeiten in der Ablaufsteuerung hinzugefügt werden, um auftretende Fehler gleich während des Tests per Ferneingriff auch korrigieren zu können. Zudem sollten sensorisch noch nicht erfassbare kritische Bereiche, wie z.B. Handläufe in den Gängen der Klinik, in der Navigationskarte als NoGo-Gebiete eingezeichnet werden (↗ F6).

4.4.2 Nutzertest Nr. 2 im September 2015

Im Nachgang von Nutzertest 1 wurde eine zusätzliche Kamera zur Erkennung von sitzenden Personen am Roboter installiert, außerdem wurde der Algorithmus für das Heranfahren an sitzende Personen verbessert (neuer zweistufiger Ansatz: erst Heranfahren an eine Position mit optimalem Abstand, dann Anpassung der Orientierung des Roboters). Damit konnte im August 2015 nochmals ein Nutzertest mit Patienten-Doubles durchgeführt werden, ehe es im September 2015 zum zweiten Nutzertest mit freiwilligen Patienten kam (↗ R2, R3). Im Unterschied zum Nutzertest 1 wurde hier ein anspruchsvolleres Setting gewählt - das Training erfolgte auch während der Stoßzeiten, zudem konnten die Patienten längere Trainingsrouten wählen. Immer noch vorhandene Einschränkungen dieses Tests waren aber (↗ S1): keine vollständige Autonomie (die Sitzenderkennung war zwar aktiviert zur Funktionsprüfung, aber noch nicht in die Applikation integriert und musste daher noch manuell per

Ferneingriff getriggert werden), die Trainingsapplikation hatte nur eingeschränkte Wahlmöglichkeiten (der Patient konnte wahlweise nur die linke oder rechte Teilstrecke bis zum Ende des Ganges wählen). Bezüglich der erreichten Autonomie ergab sich damit folgender Stand:

Teilabschnitt	Erreichte Autonomie
TA1: Fahrt zum Zimmer	100% autonom
TA2: Sitzenderkennung	autonom im Test. noch nicht integriert 100% per Ferneingriff
TA3: Heranfahren an Patienten	100% autonom
TA4: Dialog und Auswahl	2 Trainingsstrecken
TA5: Begleiten des Patienten	teilautonom, tw. mit Ferneingriffen
TA6: Interakt. an Ruhepunkten	siehe TA2
TA7: Auswertung/Verabschied.	100% autonom

Die Angaben zur Testdurchführung sind wiederum der nachfolgenden Tabelle zu entnehmen (↗ R5):

Kriterien	Angaben
Zeitraum:	9/2015, 2 Tage
Anzahl Patienten:	7
Anzahl Durchläufe:	21 (Gesamt)
Verwend. Gehhilfsmittel:	3 x Rollatoren 2 x Gehstöcke 1 x 4-Punkt-Gehstock 1 x ohne
Gefahrene Gesamtstrecke:	2109 m
Trainingszeit:	142 min
Mittlere Dauer pro Pat.:	6 min 45 s
Anzahl Passanten:	353 16,8 pro Durchlauf 16,7 pro 100 m 2,5 pro Minute
Eingriffe per Contr. Tablet:	56

Navigationsprobleme (↗ F2): Während dieses Tests berührte der Roboter nur 1 mal ein Hindernis (ein herausstehendes Rad eines Transportwagens) und musste durch einen manuellen Eingriff aus dieser Situation befreit werden (↗ N3). Grund dafür war ein Ausfall des 3D-Sensors, so dass das Rad nicht wahrgenommen wurde. Einmal fuhr der Roboter zu nah an eine in einem Rollstuhl sitzende Person heran, weil die Personendetektion diese Person übersehen hatte (↗ F4).

56 mal musste der Beobachter per Ferneingriff reagieren, um unsichere oder fehlerhafte Hypothesen des Wiedererkennungsmoduls bei Abrissen des Personentrackings zu bestätigen bzw. zu korrigieren (↗ N2). In neun Fällen waren es wiederum falsche Hypothesen, alle anderen Fälle waren wieder zu unsicher und erforderten lediglich Bestätigung. Wesentliche Ursachen hierfür waren die größeren Distanzen zwischen Roboter und Patient und stärker wechselnde Beleuchtungsbedingungen auf den längeren Laufstrecken dieser Testreihe. Bezogen auf die Anzahl der Passanten lag die Erkennungsleistung aber in der gleichen Größenordnung wie die während der Stufe 1, den Funktionstests mit Mitarbeitern.

Zwischenfazit: Insgesamt waren immer noch recht viele Korrekturingriffe nötig (↗ N2) und der Abstand zwischen Roboter und Patient war teilweise noch zu groß.

Während dieses zweiten Tests wurde zudem deutlich, dass das Trainingsszenario noch stärker erweitert werden muss durch längere Trainingsstrecken und mehr Auswahlmöglichkeiten. Außerdem waren weitere Verbesserungen an den Basisleistungen erforderlich, so z.B. bei der Wiedererkennung des aktuellen Nutzers, weshalb zusätzlich Sensorik zur Wiedererkennung eingesetzt werden sollte (auf Sonar basierende Peilung des Patienten).

4.4.3 Nutzertest Nr. 3 im November 2015

Bei diesem Test im November 2015 lag der Schwerpunkt auf dem erzielbaren Begleitabstand zwischen Patient und Roboter sowie auf noch immer vorhandenen Navigationsproblemen beim Trainingsablauf mit Patienten (↗ F2). Außerdem wurden während dieses Tests die neuen Komponenten zur Sitzenderkennung und zur peilungsbasierten Wiedererkennung auf ihre Funktionalität getestet, aber noch nicht in die Applikation integriert (↗ F3). Bezüglich der Autonomie ergab sich jetzt folgender Stand:

Teilabschnitt	Erreichte Autonomie
TA1: Fahrt zum Zimmer	100% autonom
TA2: Sitzenderkennung	autonom im Test. noch nicht integriert 100% per Feineingriff
TA3: Heranfahren an Patienten	100% autonom
TA4: Dialog und Auswahl	längere Trainingsstrecken
TA5: Begleiten des Patienten	teilautonom, tw. mit Feineingriffen
TA6: Interakt. an Ruhepunkten	siehe TA2
TA7: Auswertung/Verabschied.	100% autonom

Angaben zur Testdurchführung sind wiederum tabellarisch aufgelistet (↗ R5):

Kriterien	Angaben
Zeitraum:	11/2015, 2 Tage
Anzahl Patienten:	4
Anzahl Durchläufe:	15 (Gesamt) 4x Instruktionsgang 11x Eigentraining
Verwend. Gehhilfsmittel:	Rollatoren
Gefahrene Gesamtstrecke:	2714,2 m
Trainingszeit:	3h 21 min
Mittlere Dauer pro Pat.:	50,2 min
Anzahl Passanten:	311 20,7 pro Durchlauf 11,4 pro 100 m 1,54 pro Minute

Begleitabstand: Der Abstand zwischen dem Laufcoach und den Patienten lag während des Trainings mit den Patienten bei durchschnittlich 2,6 m (minimal: 1 m, maximal 8 m) (↗ F5). Der wesentliche Grund für den zu hohen Maximalabstand ist folgender: Nachdem der Patient sich am Ruhepunkt hinsetzt oder sich am Anfang einloggt, wartet der Roboter bis der Patient vorbei läuft und eine festgelegte Mindestdistanz zum Roboter erreicht ist, bevor er dem Patienten folgt. Dieses Verhalten ist nötig, damit der Roboter den Patienten beim Passieren nicht auf seinem Weg behindert. Wenn der Patient aber

in dieser Phase des Vorbeilaufens kurzzeitig nicht beobachtbar ist (z.B. durch eine fehlerhafte Personendetektion oder Probleme der Wiedererkennung durch vorbeilaufende Passanten), können große Abstände entstehen, bis der Roboter das Vorbeilaufen auch erkannt hat und sich in Bewegung setzt, um dem Patienten zu folgen.

Probleme: 7 mal fuhr der Roboter sehr nah an Personen heran (< 15 cm) (↗ F4). Gründe dafür waren wiederum plötzlich aus Zimmern heraustretende Personen und Passanten, die den Roboter sehr dicht überholten.

Zwischenfazit: Abgesehen von den noch in die Applikation zu integrierenden Funktionalitäten (Sitzenderkennung, peilungsbasierte Wiedererkennung) erwies sich der Laufcoach schon relativ robust unter den spezifischen Klinikbedingungen. Die Trainingsservices konnten durch die Patienten genutzt werden, für die nächsten Tests sollte den Patienten eine vollständige Wahlfreiheit ihrer Trainingsstrecke auf der gesamten Ebene überlassen werden. Allerdings erwies sich der Begleitabstand zwischen Patient und Roboter beim Training als noch zu variabel, was die Patienten verunsicherte.

4.4.4 Nutzertest Nr. 4 im Januar 2016

Bei diesem Test bildeten die bisherigen, aber auch eine ganze Reihe neuer Fragestellungen den Fokus der Praxistauglichkeitsuntersuchungen: Begleitabstand Patient-Roboter beim Training, Qualität der Nutzerwiedererkennung nach Installation einer neuen hochauflösenden Panoramakamera (bestehend aus 6 Kameras) mit großem vertikalen Öffnungswinkel für die Sitzenderkennung, Rolle der peilungsbasierten Wiedererkennung, Einfluss der Nutzung von Gehhilfsmitteln auf die ansichtsbasierte Wiedererkennung, Erkennung sitzender und aufstehender Personen sowie das Anfahren sitzender Personen.

Von besonderer Bedeutung war bei diesem Test die Fragestellung, wie sich das Gesamtverhalten des Laufcoachs verändert, wenn die Korrekturen durch den Beobachter per Feineingriff gegenüber einer vorsichtigen, frühen Korrektur nur noch in wirklich kritischen Ausnahmefällen erfolgt (späte Korrektur), in allen anderen Fällen aber die Autonomie aller Basisleistungen des Roboters gefragt ist (↗ F6). In den bisherigen Nutzertests wurde die frühe Korrektur angewandt. Mit Hinblick auf den finalen Nutzertest im März 2016, bei denen der Laufcoach möglichst autonom agieren soll, war diese Untersuchung besonders wichtig. Folgender Status wurde damit erreicht:

Teilabschnitt	Erreichte Autonomie
TA1: Fahrt zum Zimmer	100% autonom
TA2: Sitzenderkennung	66% autonom 34% per Feineingriff
TA3: Heranfahren an Patienten	100% autonom
TA4: Dialog und Auswahl	freie Auswahl
TA5: Begleiten des Patienten	teilautonom, tw. mit Feineingriffen
TA6: Interakt. an Ruhepunkten	80% autonom 20% per Feineingriff
TA7: Auswertung/Verabschied.	100% autonom

Angaben zur Testdurchführung sind wiederum tabella-

risch aufgelistet (↗ R5):

Kriterien	Angaben
Zeitraum:	1/2016, 2 Tage
Anzahl Patienten:	3
Anzahl Durchläufe:	6 (Gesamt) 2 pro Patient, je 1x mit früher Korrektur 1x mit später Korrektur
Verwend. Gehhilfsmittel:	2x Gehstock, 1x Vierpunktstock
Gefahrene Gesamtstrecke:	1.966 m
Trainingszeit:	1h 54 min
Mittlere Dauer pro Pat.:	38 min
Anzahl Passanten:	235 39,2 pro Durchlauf 12 pro 100 m 2 pro Minute
Eingriffe per Contr. Tablet:	36 (27 früh, 9 spät)

Begleitabstand: Der Abstand zwischen dem Laufcoach und den Patienten lag während des Trainings mit dem langsam laufenden Patienten bei durchschnittlich 1,2 m (minimal: 0,94 m, maximal 3,0 m) und mit den beiden schnell laufenden Patienten bei durchschnittlich 2,0 m im Bereich von 1,1 bis 4,4 m.

Nutzerwiedererkennung: Im Vergleich zu den vorangegangenen Nutzertests war das Personenaufkommen bei diesem Test sehr hoch mit 39,2 Passanten pro Durchlauf bzw. 12 pro 100 m bzw. 2,0 pro Minute. Es zeigte sich, dass die peilungsbasierte Wiedererkennung vorrangig für die die Wiedererkennung verantwortlich war, während die kleidungsbasierte Wiedererkennung immer dann entschied, wenn keine Peilung möglich war (durch Verdeckungen, Reflexionen, etc.). Im Falle der *frühen Korrekturingriffe* (am ersten Tag) waren in 13 Situationen Eingriffe nötig (↗ N2). In 9 der 13 Fälle wurden Personen nicht wiedererkannt (sowohl peilungsbasiert als auch kleidungsbasiert), es lag aber auch keine Verwechslung vor. Gründe dafür waren fehlende Detektionen, zeitweise vollständige Verdeckungen der Patienten nach Abbiegung um eine Ecke, Nichterkennung nach Hinsetzen. In einem Fall kam es zu einer Verwechslung mit einer nahe vorbeigehenden Person, in den restlichen 3 Fällen war bei der Auswertung kein ersichtlicher Grund für den frühen und damit vorschnellen Eingriff des Testbeobachters erkennbar. Außerdem kam es durch die peilungsbasierte Wiedererkennung auch zu kurzzeitigen Verwechslungen mit nahestehenden Personen, die jedoch den Ablauf nicht negativ beeinflussten.

Im Vergleich dazu waren bei den *späten Korrekturingriffen* (am zweiten Tag) in 9 Situationen Korrekturen notwendig (↗ N2). In den 9 Fällen wurde die Person nicht wiedererkannt (peilungs- und kleidungsbasiert), es lag aber auch keine Verwechslung vor. Auch hier waren die wesentlichen Gründe für die Nichtwiedererkennung das Hinsetzen, zeitweise vollständige Verdeckungen nach Abbiegung um eine Ecke und fehlende Detektionen vom Personendetektor. In einem Fall hatte der Roboter zwischenzeitlich fälschlicherweise den Durchlauf beendet.

Navigationsprobleme: 16 mal fuhr der Roboter zu nah an

Personen heran (< 15 cm), ohne sie zu berühren (↗ F2). Gründe dafür waren ausbleibende Detektionen (meist bei Rollstuhlfahrern oder teilweisen Verdeckungen durch andere Gegenstände), plötzlich auftauchendes Personal aus Zimmern und Passanten, die den Roboter sehr dicht überholten.

Erkennung Hinsetzen & Aufstehen: Eine robuste autonome Sitzenderkennung ist für den Trainingsbeginn (Initialsituation) und für ungeplante Pausen während des Trainings erforderlich. Von den 6 Initialsituationen (3 an jedem Tag) wurden 4 korrekt erkannt (davon 2x leicht verzögert mit 10 und 26 s Verzögerung), zwei wurden nicht erkannt wegen fehlender Personendetektion und mussten per Ferneingriff getriggert werden (↗ N2). Bei der Sitzenderkennung während der Pausen waren 24 Erkennungen notwendig, davon wurden 20 erkannt (8x sofort, 8x verzögert > 5 sec, 4x deutlich verzögert), 4 mussten per Ferneingriff getriggert werden. Grund für die deutlichen Verzögerungen sind die bewusst späten Eingriffe des Beobachters am zweiten Tag (↗ N2), dadurch kam es vor, dass Personen nicht wiedererkannt wurden, während sie sich hinsetzten. Gründe dafür sind zeitweise ausbleibende Personendetektionen durch atypische Ansichten beim Hinsetzen. In drei Fällen wurden sitzende Personen als stehend erkannt, dann wieder als sitzend durch Fehler bei der Wiedererkennung. Das Aufstehen wurde in 20 von 28 Fällen korrekt erkannt, davon 4x länger als 5 Sekunden verzögert. Gründe für die beobachteten Fehler sind wiederum fehlerhafte Wiedererkennung oder ausbleibende Personendetektionen, in einem Fall beendete der Roboter zwischenzeitlich fälschlicherweise den Durchlauf.

Heranfahren an Sitzende: 28 mal musste an die Patienten herangefahren werden, 2 mal wurde durch den Roboter abgebrochen, da der Patient nicht mehr erreichbar war (nicht mehr sichtbar wegen fehlender Wiedererkennung). In 26 der erreichten Anfahrpositionen konnten die Patienten das Display gut bedienen. Der Winkel zum Patienten war in 25 von 26 Fällen kleiner als 45°, was Voraussetzung für eine gute Bedienbarkeit des Touch-Displays ist. Der Abstand war in 15 von 26 Fällen kleiner als 30 cm, so dass das Display ohne bzw. mit wenig Vorbeugen gut bedienbar war. Bei Abständen größer als 30 cm mussten sich die Patienten allerdings weit vorbeugen.

5 Fazit und Ausblick

Nach vier zeitintensiven Nutzertests konnte ein Stand erreicht werden, in dem alle Komponenten integriert waren und autonom liefen und die meisten Basisleistungen auch vollständig ohne Korrekturen per Ferneingriff funktionieren. Einige wenige, aber wichtige Basisleistungen, wie die Sitzenderkennung oder die Patientenwiedererkennung, erforderten während des Trainings gelegentlich aber noch immer Korrekturingriffe, wenn in bestimmten Situationen Erkennungen falsch waren oder sogar ausblieben. Hier hat auch die peilungsbasierte Wiedererkennung mittels des vom Patienten zu tragendem Sonar-Gerätes keine hundertprozentige Robustheit erbracht, weil sich durch Eigenverdeckungen und akusti-

sche Einflüsse häufig Fehlpeilungen oder Ausfälle ergaben, die dann per Ferneingriff korrigiert werden mussten. Bei der kleidungs-basierten Wiedererkennung wurden technische Unzulänglichkeiten der verbauten Panoramakamera deutlich, deren unzureichende Farbtreue bei starken Änderungen der Beleuchtungsbedingungen gelegentlich zu Problemen der Wiedererkennung während des Laufens führte, die dann ebenfalls per Ferneingriff korrigiert werden mussten. Konkrete Prozentangaben sind hier nicht möglich, da die Anzahl der korrekten Wiedererkennungen nicht mitprotokolliert wurden.

Bei der kamerabasierten Sitzenderkennung traten in 20% der Fälle Defizite auf, die sich sowohl mit den technischen Unzulänglichkeiten der Panoramakamera als auch mit der teilweise unzureichenden Erkennung von sitzenden Personen auf Basis des Größenunterschiedes zwischen Stehen und Sitzen erklären lassen. Für einen vollständig autonomen praktischen Einsatz reicht diese Erkennungsrate aber noch nicht aus. Abhilfe wird hier vermutlich nur eine echte ansichtsbasierte Erkennung von Sitzposen in 2D oder 3D erbringen. Angesichts der extremen Vielfalt von möglichen Sitzhaltungen und Ansichtstörungen durch Gehhilfsmittel, Bekleidung und mitgeführten Utensilien stellt dies aber ein sehr anspruchsvolles, neues Erkennungsproblem dar, das bis zum Abschluss des Projektes nicht mehr in Angriff genommen werden kann.

Angesichts dieser erkannten Probleme ist für Nutzertest Nr. 5 im März 2016 keine hundertprozentige Fehlerfreiheit und damit keine vollständige Autonomie mehr realisierbar. Statt aber auf eine hundertprozentige Fehlerfreiheit aller Basisleistungen zu setzen (\nearrow N1), sollte in der verbleibenden Zeit und in Folgeprojekten nach intelligenten Strategien zur autonomen Behandlung von ausbleibenden Erkennungen (z.B. bei der Wiedererkennung oder der Sitzenderkennung), von neu auftretenden, unerwarteten Situationen oder von Fehlern im Ablauf des Trainingszyklus gesucht werden. Beispielhaft könnte bei einem Kontaktverlust zum Patienten wegen fehlerhafter Wiedererkennung eine Notfallstrategie vereinbart werden, nach der der Patient am nächsten Ruhepunkt auf den Roboter wartet, so dass dieser dort und auf dem Weg dorthin gezielt nach ihm suchen kann. Mit Hinblick auf die angestrebte vollständige Autonomie als Voraussetzung für eine echte Praxistauglichkeit, muss ehrlich eingeschätzt werden, dass in ROREAS bis zum Projektende im März 2016 eine technische Fehlerfreiheit und damit eine *wirkliche Praxistauglichkeit nicht erreicht* werden kann.

Durch die Möglichkeit der Ferneingriffe ergibt sich für sozialwissenschaftliche Untersuchungen der Nutzertests aber zumindest ein vollwertig nutzbarer Trainingsassistent. Eine Zielstellung der Nutzertests war aber neben der Bewertung der Praxistauglichkeit aus technischer Sicht auch die sozialwissenschaftliche Bewertung des Gesamtszenarios Laufcoach hinsichtlich des Trainingsablaufs und der Art der Vermittlung von Informationen zwischen Patient und Roboter. Darüber hinaus bestand eine weitere Zielstellung in der Bewertung einzelner Komponenten der robotischen Basisleistungen aus so-

zialwissenschaftlicher Sicht. Betrachtete Fragestellungen umfassten z.B. die Situation des Wartens vor dem Patientenzimmer hinsichtlich der Störung bzw. Beeinflussung unbeteiligter Patienten oder des Personals, aber auch die Erreichbarkeit des Roboters durch den Patienten zum Start des Trainingsprogramms. Beim Begleiten des Patienten wurde die Entfernung des Roboters zum Patienten durch diesen bezüglich Zufriedenheit und Sicherheit bewertet. Es soll an dieser Stelle aber ausdrücklich betont werden, dass diese sozialwissenschaftliche Bewertung der Nutzertests nicht Gegenstand dieses Beitrages war, ausgewählte Ergebnisse dazu werden in [8] vorgestellt.

Danksagung: Die Autoren möchten allen Projektpartnern des ROREAS-Verbundes ganz herzlich für die vertrauensvolle und inspirierende Kooperation danken - der MetraLabs GmbH Ilmenau, der "m&i Fachklinik" Bad Liebenstein, dem SIBIS Institut für Sozialforschung Berlin und der Krankenkasse Barmer GEK.

Literatur

- [1] Gross, H.-M., Mueller, S., Schroeter, Ch. et al. (2015). Robot companion for domestic health assistance: Implementation, test and case study under everyday conditions in private apartments. In *Proc. Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 5992–5999
- [2] Scheidig, A., Einhorn, E., Weinrich, Ch. et al. (2015). Robotischer Reha-Assistent zum Lauftraining von Patienten nach Schlaganfall: Erste Ergebnisse zum Laufcoach. In *Proc. 8th German AAL Conference (AAL)*, pp. 436-445
- [3] Projektwebseite: www.serroga.de
- [4] Scheidig, A., Debes, K., Müller et al. (2015). SERROGA: Funktions- und Nutzertests Herangehensweise und Ergebnisse. In *Proc. 8th German AAL Conference (AAL)*, pp. 34-43
- [5] Gross, H.-M., Debes, K., Einhorn, E. et al. (2014). Mobile robotic rehabilitation assistant for walking and orientation training of stroke patients: A report on work in progress. In *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*, pp. 1880–1887
- [6] Gross, H.-M., Scheidig, A., Debes, K. et al. (2016). ROREAS - Robot coach for walking and orientation training in clinical post-stroke rehabilitation: Prototype implementation and evaluation in field trials. to appear: *Autonomous Robots (AR)*
- [7] Projektwebseite: www.roreas.org
- [8] Meyer, S., Fricke, Ch. (2016). Patientenzentrierte Roboterentwicklung - Schlüssel zur Nutzerakzeptanz? Akzeptanzuntersuchungen mit 80 Patienten in der neurologischen Rehabilitation. In *Proc. 9th German AAL Conference (AAL)*