

Otto-von-Guericke-Universität Magdeburg
Fakultät für Informatik

Masterarbeit



Entwicklung einer multimodalen Benutzerschnittstelle zur
berührungslosen Steuerung eines Roboterarms im medizinischen
Interventionsraum

von: Tonia Mielke (Matrikelnr. 215169)
geboren am 08. August 1997
in Wittingen

Erstgutachter: Prof. Dr. Christian Hansen
Zweitgutachterin: PD Dr. Sylvia Saalfeld

Kurzreferat

In der interventionellen Radiologie können unter medizinischer Bildgebung minimalinvasive Eingriffe durchgeführt werden. Um bei nadelgestützten Eingriffen die Instrumente sicher zu den Zielstrukturen zu navigieren, kann für die Visualisierung die Kombination aus Computertomographie und Ultraschall verwendet werden. Ultraschall bietet den Vorteil Bildgebung ohne ionisierende Strahlung zu ermöglichen. Jedoch müssen Radiolog*innen während der Intervention gleichzeitig die Sonde und die Nadel platzieren, woraus nicht ergonomische Positionen und Ungenauigkeiten resultieren können. Um diese Herausforderungen zu überwinden, kann ein Roboterarm verwendet werden, der die Ultraschallsonde positioniert. Während einige Aufgaben, wie die Visualisierung der Nadelspitze, perspektivisch autonom erfolgen könnten, gibt es weiterhin Aufgaben, die manuell erfolgen müssen. Somit wird eine Schnittstelle zur direkten Kommunikation mit dem Roboter benötigt, die beispielsweise die manuelle Visualisierung der Zielläsion oder die Korrektur der automatischen Robotersteuerung ermöglicht.

In dieser Arbeit wird eine multimodale Schnittstelle für die Mensch-Roboter-Interaktion mit einem Roboterarm zur Ultraschallbildgebung entwickelt. Dazu wurden im ersten Schritt verschiedene Interaktionskonzepte mit unterschiedlichen Interaktionsmodalitäten konzeptioniert, umgesetzt und evaluiert. Im Rahmen einer Nutzerstudie konnte gezeigt werden, dass für diskrete Interaktionen, wie die Auswahl von Modi, in Bezug auf Usability, die Interaktion über Sprache besser geeignet ist als die Blickinteraktion. Für Navigationsinteraktionen konnten durch die Steuerung über Handbewegung eine geringere Dauer in Form des Task-Completion-Times als über Kopfbewegungen erreicht werden. Basierend auf diesen Ergebnissen wurde im zweiten Schritt ein finales Interaktionskonzept bestehend aus Sprach- und Handinteraktion entwickelt. Dieses Konzept wurde abschließend im Rahmen einer qualitativen Nutzerstudie evaluiert. Dadurch konnte festgestellt werden, dass durch das System die nutzerfreundliche manuelle Positionierung des Roboters für die Ultraschallbildgebung ermöglicht wird. Weiterhin konnten Funktionen identifiziert werden, die zukünftig zur Verbesserung des Systems beitragen könnten.

Abstract

In interventional radiology, minimally invasive procedures can be performed using medical imaging. To safely navigate the instruments to target structures during needle-guided procedures, the combination of computed tomography and ultrasound can be used for visualization. Ultrasound offers the advantage of imaging without ionizing radiation. However, radiologists must simultaneously position the probe and needle during the procedure, which can lead to unergonomic postures and inaccuracies. To overcome these challenges, a robotic arm can be used to position the ultrasound probe. While some tasks such as visualization of the needle tip could be performed autonomously in the future, there are still tasks that need to be performed manually. Therefore, an interface for direct communication with the robot is required, which facilitates, for example, manual visualization of the target lesion or correction of automatic robot control.

In this work, a multimodal interface for human-robot interaction with a robotic arm for ultrasound imaging is developed. For this purpose, different interaction concepts with different interaction modalities were developed and evaluated. In a user study, we were able to show that for discrete interactions such as mode selection, interaction via speech is more suitable than gaze interaction. For navigation interactions, control via hand movements resulted in a lower task completion time than via head movements. Based on these results, a final interaction concept consisting of speech and hand interaction was developed. This concept was finally evaluated in a qualitative user study. As a result, it was determined that the system enables user-friendly manual positioning of the robot for ultrasound imaging. Furthermore, additional functions could be identified that could contribute to the overall improvement of the system.

Eidesstattliche Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig und unter ausschließlicher Verwendung der angegebenen Literatur und Hilfsmittel erstellt zu haben.

Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Magdeburg, den 30.02.2022

Tonia Mielke

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Zielsetzung	2
1.3	Aufbau der Arbeit	2
2	Grundlagen	3
2.1	Medizinische Grundlagen	3
2.2	Roboter im medizinischen Interventionsraum	4
2.3	Roboterprogrammierung	6
2.4	Mensch-Roboter-Interaktion	7
2.4.1	Interaktionsparadigmen	7
2.4.2	Interaktionsmodalitäten	9
2.5	Registrierung	12
2.6	Stand der Technik	13
2.6.1	Robotergestützter Ultraschall	13
2.6.2	Mensch-Roboter-Interaktion	14
3	Anforderungsanalyse	19
3.1	User Research	19
3.1.1	Workflowanalyse	19
3.1.2	Interaktionsaufgabe	20
3.2	Funktionale Anforderungen	21
3.3	Nichtfunktionale Anforderungen	22
4	Konzeption	25
4.1	Konzeption der Rahmenbedingungen	25
4.1.1	Multimodale Schnittstelle	25
4.1.2	Registrierung	28
4.2	Konzeption der Interaktion	28
4.2.1	Diskrete Interaktionen	28
4.2.2	Navigationsinteraktion	29
4.3	Konzeption Interaktionsaufgabe	33
4.3.1	Interaktionsaufgabe initiale Studie	33
4.3.2	Interaktionsaufgabe finale Studie	35

5	Umsetzung	37
5.1	Rahmenbedingung	37
5.1.1	Entwicklungsumgebung	37
5.1.2	Robotersteuerung	37
5.1.3	Registrierung	39
5.2	Initialer Prototyp	41
5.2.1	Diskrete Interaktion	41
5.2.2	Navigationsinteraktion	42
5.3	Finaler Prototyp	46
5.3.1	Diskrete Interaktion	46
5.3.2	Navigationsinteraktion	47
5.4	Interaktionsaufgabe	51
5.4.1	Initiale Studie	51
5.4.2	Finale Studie	52
6	Evaluierung	53
6.1	Evaluierung der Interaktionsmodalitäten	53
6.1.1	Ziel und Fragestellung	53
6.1.2	Methodisches Vorgehen	53
6.1.3	Ergebnisse	56
6.1.4	Diskussion	61
6.1.5	Schlussfolgerungen	63
6.2	Evaluierung des Gesamtkonzepts	65
6.2.1	Ziel und Fragestellung	65
6.2.2	Methodisches Vorgehen	65
6.2.3	Ergebnisse	67
6.2.4	Diskussion	70
6.2.5	Schlussfolgerungen	73
6.3	Auswertung der Anforderungsanalyse	74
7	Abschluss	77
7.1	Zusammenfassung	77
7.2	Ausblick	78
	Literaturverzeichnis	81
A	Anhang	97
A.1	Fragebögen	97
A.1.1	Fragebogen Experteninterview	97
A.1.2	Nasa TLX	98

A.1.3	System Usability Scale	99
A.2	Ergebnisse Experteninterviews	100
A.3	Ergebnisse initiale Nutzerstudie	101
A.3.1	Quantitative Ergebnisse	101
A.3.2	Lernkurven	104

Symbolverzeichnis

a, A	Skalar
\vec{a}, \vec{A}	Vektor
\vec{T}	Translation
R	Rotation
q	Rotation als Quaternion
${}^A_B R, {}^A_B \vec{T}$	Rotation bzw. Position des Koordinatensystems A im Koordinatensystem B
R_n^a, \vec{T}_n^a	Rotation bzw. Position von a zum Zeitpunkt n
$\perp \vec{a}$	Perpendikularer Vektor zu \vec{a}

Abkürzungsverzeichnis

2D	Zweidimensional
3D	Dreidimensional
AR	Augmented Reality
CT	Computertomographie
DoF	Freiheitsgrade engl. Degree of Freedom
GUI	Grafische Benutzeroberfläche, engl. Graphical User Interface
HMD	Head-Mounted-Display
HRI	Mensch-Roboter-Interaktion engl. Human-Robot-Interaction
IMU	Inertiale Messeinheit engl. Intertial Measurement Unit
KOS	Koordinatensystem
LIN	Lineare Roboterbewegung
OP	Operationssaal
PETG	Polyethylen Terephthalat Glykol
PTP	Point-to-Point Roboterbewegung
SUS	System Usability Scale
TCP	Tool Center Point
TCT	Task-Completion-Time
UDP	User Data Protocol
US	Ultraschall
VR	Virtual Reality

Abbildungsverzeichnis

2.1	Übersicht über Mensch-Roboter-Interaktionen (nach [40, 41])	8
2.2	Dimensionen zur Unterscheidung verschiedener mehrdimensionaler Manipulationsmethoden (nach [44])	8
2.3	Überblick über Ablauf und die benötigten Schritte der Mensch-Roboter-Interaktion durch Gesten [55]	11
2.4	Schematischer Überblick über Rotationen und Translationen des Kopfes und Beispielbewegungen [59]	12
3.1	Arbeitsablauf während ultraschallgestützter Nadelführung	20
4.1	Überblick über die Kopfinteraktionen für die Navigation des Roboters . . .	31
4.2	Gomperzfunktion als Transferfunktion zur Abbildung der Änderung der normierten Kopforientierung δ_n auf die normierte Geschwindigkeit v_n der Bewegung des Roboters	31
4.3	Überblick über die Handinteraktion für die Navigation des Roboters	32
5.1	Übersicht über die für die Registrierung relevanten Koordinatensysteme. Dabei beschreiben die Indizes c das KameraKOS, w das WeltKOS, m das MarkerKOS und r das RoboterKOS	39
5.2	Beispielhafte Darstellung der positionsabhängigen Auswahl der Achsen für die Roboterbewegung	44
5.3	Logik der Interaktionskonzepte für Kopfinteraktion	45
5.4	Logik der Interaktionskonzepte für Handinteraktion	46
5.5	Für die Interaktion genutzte Freiheitsgrade der Hand	48
5.6	Überblick über die finale Handinteraktion für die Rotation des Endeffektors	49
5.7	Logik der Rotationsinteraktionen für finale Navigationsinteraktion	50
5.8	Visuelles Feedback für die Rotation	50
5.9	Darstellung der Sonde als Hologramm auf der HoloLens als Zielposition für die Primäraufgabe der initialen Studie	51
5.10	Darstellung des Balkens der Sekundäraufgabe mit Regler und Zieltargets .	52
6.1	Übersicht über den Versuchsaufbau der initialen Nutzerstudie	55
6.2	Darstellung des mittleren Translationsfehlers mit Standardfehler	57

6.3	Darstellung des mittleren Rotationsfehlers für die unterschiedlichen Interaktionsmethoden unterteilt nach den Teilgruppen mit Standardfehler	57
6.4	Darstellung der mittleren Task-Completion-Time mit Standardfehler	58
6.5	Darstellung des mittleren Raw Nasa TLX Scores mit Standardfehler	58
6.6	Nasa TLX Bewertungen der unterschiedlichen Dimensionen nach Interaktionskombination und Versuchsgruppe	59
6.7	Darstellung der durchschnittlichen Fehlerquoten für die verpassten Targets und die Fehlklicks der Sekundäraufgabe	59
6.8	Aufbau des für die finale Nutzerstudie verwendeten Ultraschallphantoms .	66
A.1	Darstellung des Translationsfehlers unter allen Teilnehmenden und mit allen Interaktionskombinationen für die verschiedenen Durchläufe	104
A.2	Darstellung des Rotationsfehlers unter allen Teilnehmenden und mit allen Interaktionskombinationen für die verschiedenen Durchläufe	104
A.3	Darstellung des TCT unter allen Teilnehmenden und mit allen Interaktionskombinationen für die verschiedenen Durchläufe	105
A.4	Darstellung der verpassten Targets und Fehlklicks unter allen Teilnehmenden der Teilgruppe A und mit allen Interaktionskombinationen für die verschiedenen Durchläufe	105

Tabellenverzeichnis

2.1	Übersicht über in der Literatur verwendete Eingabemodalitäten für natürliche Schnittstellen zur Mensch-Roboter-Interaktion	17
3.1	Übersicht über funktionale Anforderungen	22
3.2	Übersicht über nicht funktionale Anforderungen	23
5.1	Übersicht über verwendete Sprachbefehle für den finalen Prototypen	48
6.1	Statistisch signifikante Ergebnisse der initialen Nutzerstudie	56
6.2	Zusammenfassung der qualitativen Ergebnisse der finalen Nutzerstudie in den verschiedenen Kategorien. Die ID gibt dabei die entsprechenden Proband*innen an	68
6.3	Zusammenfassung über die Auswertung der funktionalen Anforderungen	74
6.4	Zusammenfassung über die Auswertung der nichtfunktionalen Anforderungen	75
A.1	Zusammenfassung der Ergebnisse der Experteninterviews zu dem Use Case der nadelgestützten Interventionen unter Ultraschall und CT Bildgebung. Die ID gibt dabei die entsprechenden Expert*innen an	100
A.2	Ergebnisse der Teilgruppe mit Sekundäraufgabe (Teilgruppe A)	101
A.3	Ergebnisse der Teilaufgabe ohne Sekundäraufgabe (Teilgruppe B)	102
A.4	Zusammenfassung der ANOVA-Ergebnisse ($\alpha > 0.05$) für die Parameter der Sekundäraufgabe	102
A.5	Zusammenfassung der ANOVA-Ergebnisse ($\alpha > 0.05$) des Translationsfehlers, des Rotationsfehlers, des TCTs und des Raw Nasa TLX Scores	103

1 Einleitung

1.1 Motivation

In der Radiologie werden sowohl diagnostische als auch therapeutische Verfahren durch bildgestützte minimalinvasive Eingriffe ermöglicht. Die Interventionen minimieren dabei die Gewebeverletzungen, indem der Zugang durch Nadeln oder nadelartige Instrumente erfolgt. Dadurch haben die Eingriffe im Gegensatz zu konventionellen Operationen den Vorteil die postoperativen Schmerzen und die Genesungszeit zu reduzieren, während kosmetische Ergebnisse und die Kosten-Effektivität gesteigert werden [1]. Da in diesen Eingriffen keine direkte Sicht auf die Zielstrukturen möglich ist, wird zur Unterstützung der Navigation medizinische Bildgebung verwendet, sodass die Nadel ohne Verletzungen von Risikostrukturen zur Zielstruktur geführt werden kann. Dabei ist eine präzise Bildgebung entscheidend für den Erfolg der Eingriffe.

Für die Bildgebung wird oftmals Computertomographie (CT) verwendet, jedoch entsteht durch das CT eine Strahlenbelastung für Patient*innen und medizinisches Personal. Um die Strahlenexposition zu reduzieren kann ergänzend Ultraschall (US) verwendet werden. US ermöglicht eine Echtzeit-Bildgebung mit gutem Weichteilkontrast ohne ionisierende Strahlung zu verursachen [2]. Durch die Kombination von US und CT können die Interventionen schneller und einfacher durchgeführt werden [3].

Jedoch hat US den Nachteil, dass bei nadelbasierten Interventionen Radiolog*innen gleichzeitig die Nadel führen und mit der anderen Hand die Sonde positionieren müssen. Daraus können Fehler und Ungenauigkeiten resultieren [4]. Zusätzlich müssen für die US-Bildgebung teilweise ermüdende Kräfte angewendet und die Sonde in unangenehmen Positionen gehalten werden, um gute Bilder erhalten zu können. Aus dieser wiederholten Belastung können Verletzungen für Radiolog*innen entstehen [5,6]. Um eine ergonomische Arbeitsumgebung zu ermöglichen kann robotergestützter Ultraschall verwendet werden. Dabei wird die US-Sonde durch einen Roboterarm gehalten. Die Bewegung des Roboters kann autonom oder manuell erfolgen. Jedoch kann es selbst bei der autonomen Robotersteuerung durch Fehler oder verdeckende Strukturen, wie Knochen, notwendig sein, die Sondenposition manuell zu korrigieren. Somit wird eine Schnittstelle für die Mensch-Roboter-Interaktion benötigt, die die Kommunikation mit dem Roboter und somit dessen Steuerung ermöglicht. Dabei ist für die erfolgreiche Nutzung eines solchen Systems eine intuitive Interaktionsschnittstelle essentiell [7].

Da die Interaktion mit dem Roboter während der Intervention erfolgt, muss darauf geachtet werden, dass die Sterilität erhalten bleibt. Dies kann durch eine berührungslose Interaktionsschnittstelle realisiert werden. Diese Schnittstelle kann multimodal, also unter Verwendung mehrerer Modalitäten umgesetzt werden, um die Interaktion natürlicher und intuitiver zu gestalten [8].

1.2 Zielsetzung

Das Ziel dieser Arbeit ist die Konzeption und Umsetzung eines multimodalen Konzepts zur Mensch-Roboter-Interaktion mit einem Roboterarm im medizinischen Interventionsraum. Dazu soll der Nutzungskontext analysiert und basierend auf einem abstrahierten Workflow die Interaktionsaufgaben spezifiziert werden. Auf Grundlage der daraus resultierenden Anforderungen sollen geeignete Interaktionsmodalitäten identifiziert und ein Interaktionskonzept entworfen werden. Das Interaktionskonzept soll abschließend im Rahmen einer Nutzerstudie evaluiert werden.

1.3 Aufbau der Arbeit

Die Arbeit wird in die folgenden Kapitel gegliedert.

In Kapitel 2 werden die für die Arbeit benötigten Grundlagen über den medizinischen Kontext, Robotik in der Medizin und Mensch-Roboter-Interaktion vorgestellt. Zusätzlich wird im Unterkapitel *Stand der Technik* auf verwandte Arbeiten eingegangen.

Kapitel 3 stellt die Anforderungen an das zu entwickelnde System vor. Dabei werden zur Anforderungsanalyse die Ergebnisse von Experteninterviews präsentiert.

Aufbauend auf den beiden vorhergehenden Kapiteln wird in Kapitel 4 die Konzeption der Mensch-Roboter-Interaktion vorgestellt. Diese umfasst die Konzeption der Rahmenbedingungen, die der Interaktion und die Konzeption von abstrahierten Interaktionsaufgaben für die Evaluierung des Systems.

In Kapitel 5 wird die Realisierung der zuvor erarbeiteten Konzepte beschrieben. Dabei wird die Umsetzung eines initialen und eines darauf basierenden, verbesserten finalen Prototyps eingegangen.

Kapitel 6 beschreibt die Evaluierung der erarbeiteten und umgesetzten Konzepte im Rahmen von zwei Nutzerstudien. Dabei dient die Erste der Bestimmung eines geeigneten Interaktionskonzepts, während die zweite Studie das verbesserte Konzept unter Verwendung der Ultraschallbildgebung qualitativ evaluiert.

Abschließend werden in Kapitel 7 die Forschungsergebnisse zusammengefasst und ein Ausblick über weitere mögliche Arbeitsschritte gegeben.

2 Grundlagen

Im Folgenden werden die Grundlagen für die Entwicklung der Anforderungen und Konzepte dieser Arbeit vorgestellt. Dazu wird auf die medizinischen Grundlagen der interventionellen Radiologie sowie auf die Thematik der Roboter im medizinischen Interventionsraum eingegangen. Anschließend werden Möglichkeiten der Roboterprogrammierung vorgestellt. Eine Form der Roboterprogrammierung ist die Mensch-Roboter-Interaktion, für die Interaktionsparadigmen sowie Interaktionsmodalitäten vorgestellt werden. Abschließend wird ein Überblick über den aktuellen Forschungsstand bezüglich robotergestütztem Ultraschall und Mensch-Roboter-Interaktion gegeben.

2.1 Medizinische Grundlagen

Bildgestützte minimalinvasive Verfahren, die unter dem Begriff interventionelle Radiologie zusammengefasst werden können, sind zu einem festen Bestandteil der klinischen Praxis geworden [9]. Dabei werden bildgebende Verfahren nicht nur zur Diagnose, sondern auch zur Führung von Instrumenten verwendet. Dadurch werden Eingriffe ermöglicht, die nur ein lokal begrenztes Trauma verursachen. Zur Bildgebung werden hauptsächlich Ultraschall (US), Computertomographie (CT), Magnetresonanztomographie (MRT) und Angiographietechniken verwendet, wobei komplexe Verfahren auch hybrid unter Kombination verschiedener Bildgebungsmethoden stattfinden können [10]. Zu den Anwendungsgebieten der interventionellen Radiologie zählen vaskuläre Interventionen (lokale Lysetherapien, gefäßweiternde und gefäßverschließende Maßnahmen [11]), invasive diagnostische Verfahren (Biopsien [12]), onkologische Verfahren (ablative Therapien [13]) und Schmerztherapie.

Für diese Verfahren muss ein Instrument perkutan in eine kleine Struktur im Körper eingeführt werden. Dabei ist die Bildgebung und die Auswahl der optimalen diagnostischen und interventionellen Modalitäten entscheidend für die sichere und wirksame Durchführung [14]. Bildgebung wird dabei für verschiedene Schritte der Intervention benötigt. Zum einen für die Planung, ob Patient*innen für das Verfahren geeignet sind, zum anderen für das Targeting, also die Platzierung des Applikators. Dazu werden meistens CT und US verwendet. Die Bildgebung sollte dabei idealerweise 3D-Echtzeit-Darstellung der Eingriffsinstrumente, der umgebenden Anatomie sowie eine klare Abgrenzung der Zielregion ermöglichen und multiplaneare und interaktive Fähigkeiten besitzen [15]. Die derzeitigen

Bildgebungssysteme bieten zwar einige dieser Funktionen, aber keine davon kann alle bieten [14].

Viele Interventionen werden unter CT-Bildgebung durchgeführt. Dazu wird CT-Fluoroskopie verwendet, wodurch nahezu eine Echtzeit-Bildgebung möglich ist, die jedoch aufgrund der Strahlenbelastung für Patient*innen und Personal nur begrenzt eingesetzt werden kann [16]. Außerdem fehlt eine mögliche 3D-Rekonstruktion [17]. Zusätzlich entstehen verursacht durch die schwere Schutzbekleidung und der ungünstigen Körperhaltung aufgrund des Aufbaus des CTs während der Intervention ergonomische Herausforderungen für die Radiolog*innen [18].

Im Gegensatz dazu bietet Ultraschall den Vorteil, keine ionisierende Strahlung zu verursachen und einen guten Weichteilkontrast zu ermöglichen [2]. Somit kann durch Ultraschall eine Echtzeit-Bildgebung zur konstanten Überwachung der Nadelposition [19] und von beispielsweise durch die Atmung bewegten Körperteilen wie der Niere [7] ermöglicht werden. Weitere Vorteile sind die Verfügbarkeit, die Geschwindigkeit und die geringen Kosten [20]. Jedoch hat Ultraschall den Nachteil, dass die Bildgebung durch Knochen und Luft limitiert ist [19].

Somit kann eine Kombination beider Modalitäten sinnvoll sein, da die anatomischen Details des CTs mit der Echtzeit-Bildgebung ohne Strahlung des Ultraschalls kombiniert werden können. Für die Pfadplanung können somit die 3D-CT-Bilder genutzt und die Durchführung des Plans mithilfe von Ultraschall überwacht werden [21].

2.2 Roboter im medizinischen Interventionsraum

Zur Unterstützung von Chirurg*innen während medizinischer Interventionen können Roboter verwendet werden. Durch die hohe Genauigkeit, Geschicklichkeit und Beweglichkeit können Roboter als sehr vielseitiges Werkzeug zur Erweiterung der menschlichen Fähigkeiten dienen. Dabei sind die Roboter auf bestimmte Aufgaben innerhalb der vorgesehenen Verfahren spezialisiert, während weiterhin viele Aufgaben durch den Menschen übernommen werden [22]. Daher werden Robotersysteme oft als chirurgischen Assistenten bezeichnet, die mit Chirurg*innen zusammenarbeiten [23].

Roboter im medizinischen Interventionsraum können mithilfe unterschiedlicher Kriterien klassifiziert werden. Dazu zählen das Design des Manipulators (z.B. Kinematik, Antriebstechnik), der Zielanatomie oder -technik (z.B. kardial, intravaskulär, perkutan, laparoskopisch), der Autonomiegrad (z.B. vorprogrammiert, teleoperativ oder eingeschränkte kooperative Steuerung) oder die vorgesehene Betriebsumgebung (z.B. im CT-Scanner, im konventionellen OP) [23].

Eine mögliche Anwendung der chirurgischen Assistenzrobotik liegt im robotischen Ultraschall. Dieser kann genutzt werden, um die Ultraschallsonde während des Eingriffs präzise zu steuern und dadurch die kognitive und physische Belastung der Nutzenden während der Bildgebung zu verringern [24]. Dabei wird die Ultraschallsonde am Endeffektor des Roboters befestigt, der teleoperativ, kollaborierend oder autonom gesteuert werden kann [25]. Die Art und der Umfang, in dem mit dem Roboter interagiert wird, kann dabei entsprechend der von Yang et al. [26] vorgestellten Levels der Autonomie klassifiziert werden.

Das niedrigste Level der Autonomie entspricht dem Level 0, das keinen autonomen Anteil der Robotersteuerung bedeutet, sodass die explizite Steuerung durch Chirurg*innen erfolgt. Dies kann entweder durch eine direkte Handführung oder durch Teleoperation realisiert werden. Bei der Handführung wird von Chirurg*innen das am Roboter befestigte Instrument direkt gehalten. Durch einen Kraftsensor können die gewünschten Bewegungsrichtungen erfasst und die Roboterbewegungen entsprechend angepasst werden [23]. In der teleoperativen Steuerung findet die Steuerung physisch getrennt vom Roboter statt. Durch diese physische Trennung der Steuerung wird die Bedienung als teleoperativ bezeichnet, selbst wenn sich die steuernde Person im OP befindet [27]. Diese Systeme bestehen aus drei Komponenten: der Steuerungskonsole, an der direkte Eingaben getätigt werden, der Robotersteuerung, die die Eingaben automatisch in explizite Roboterbefehle umwandelt und der Kommunikationsverbindung, die die Kommunikation zwischen Steuerungskonsole und Robotersteuerung ermöglicht. Dabei wird jede Bewegung, die an der Steuerungskonsole eingegeben wird, an den Roboter übertragen. Für die Steuerung bieten Systeme in der Regel eine oder mehrere direkte Steuerungsschnittstellen wie Joysticks, Trackingsysteme, Sprachsteuerung oder andere Eingabeschnittstellen für die Mensch-Roboter-Interaktion [23]. Die dadurch erfassten Bewegungen können dann beispielsweise skaliert auf den Roboter übertragen werden, um präzise Bewegungen zu ermöglichen [28]. Um die Bewegung des Roboters in einem festen Koordinatensystem zu ermöglichen, werden Sensoren in den Roboter Gelenken gemeinsam mit dem kinematischen Modell des Manipulators verwendet [29]. Diese Systeme werden meist genutzt, um menschliche Bewegungen nachzuahmen oder Bilder zu erfassen [30].

Level 1 oder auch Roboterunterstützung, beschreibt Mensch-Roboter-Kooperation. Dadurch können Bediener*innen entlastet werden, da eine Arbeitsteilung, bei der nur einige Freiheitsgrade von der bedienenden Person und die restlichen vom Roboter kontrolliert werden können, ermöglicht wird [31]. Eine andere Möglichkeit der Mensch-Roboter-Kollaboration besteht darin, dass Verfahren assistiert ausgeführt werden [32]. Dafür kann beispielsweise der Roboter die Handlungen der Chirurg*innen überwachen und Kräfte anwenden, um Bewegungen in bestimmte vordefinierte Regionen zu vermeiden [33]. Die Assistenz des Roboters kann auch darin bestehen, die Fähigkeiten der Chirurg*innen zu erweitern, indem sie beispielsweise helfen, das Zittern der Hand zu beseitigen [23].

Ab Level 2 handelt es sich um autonome Robotersysteme, wobei die meisten Systeme sich zu Level 2 (Aufgabenautonomie) und Level 3 (bedingte Autonomie) zuordnen lassen. Dabei wird bei Level 2 die Ultraschallbildgebung autonom, entlang eines manuell geplanten Pfades ausgeführt, während in Level 3 sowohl die Planung als auch die Durchführung autonom stattfinden. Autonome Eingriffe können basierend auf präoperativen Informationen oder auf Grundlage von Echtzeit-Bildgebungsdaten durchgeführt werden. Bei Eingriffen auf Grundlage präoperativer Informationen wird ein aus diesen Daten gewonnenes patientenspezifisches Modell verwendet. Dieses Modell dient als Vorlage für die chirurgische Planung und wird von dem Roboter während des Eingriffs als Leitfaden für die autonome Durchführung genutzt [30]. Für die auf Echtzeitdaten basierende Robotersteuerung kann sogenanntes visuelle Servoing durchgeführt werden. Dabei wird die digitale Analyse von Bildgebungsdaten (wie US-Bildern) verwendet, um eine Roboterreaktion zu erzeugen [4]. Somit können gewünschte Bildmerkmale automatisch verfolgt und unerwünschte Bewegungen von Patient*innen kompensiert werden.

2.3 Roboterprogrammierung

Um Bewegungen des Roboters zur Erfüllung vordefinierter Aufgaben zu ermöglichen, müssen Informationen von Nutzer*innen auf den Roboter übertragen werden. Diese Informationsübertragung wird als Roboterprogrammierung bezeichnet [34]. Programmiermethoden unterscheiden sich dabei hinsichtlich der Form der benötigten Informationseingabe, die zu Roboterbefehlen führt.

Ein Kompromiss zwischen autonomen und kontinuierlich gesteuerten Systemen kann durch instruktive Systeme erreicht werden. Diese ermöglichen die Kommunikation zwischen Mensch und Roboter in Echtzeit. Dazu erhalten die Systeme während der Programmausführung eine Folge einfacher Anweisungen, die automatisch in Bewegungsbefehle des Roboters umgewandelt werden [35]. Somit wird eine sogenannte Mensch-Roboter-Kollaboration möglich, sodass Mensch und Roboter gleichzeitig am gleichen Arbeitsplatz dasselbe Ziel verfolgen und zum Erreichen dieses Ziels miteinander kommunizieren [36]. Dadurch kann im Gegensatz zu autonomen Systemen die Flexibilität erhöht und im Vergleich zu kontinuierlich gesteuerten Systemen die für die Nutzung erforderlichen Informationen reduziert werden [36]. Die für die Instruktion benötigte Kommunikation kann durch Methoden der Mensch-Roboter-Interaktion realisiert werden.

Wenn die Roboterprogrammierung für die Vorgabe von Roboterbewegungen verwendet wird, können Zielvorgaben entweder im kartesischen Raum oder im Konfigurationsraum erfolgen. Angaben im kartesischen Raum sind durch die Position und Ausrichtung des Endeffektors, also des am Roboter befestigten Werkzeugs, in einem festen Koordinatensystem definiert. Im Gegensatz dazu sind Angaben im Konfigurationsraum durch die

Achsenwinkel der verschiedenen Robotergelenke gegeben [37]. Die Bestimmung der Pfade, in denen sich der Roboter zur Zielposition bzw. Zielkonfiguration bewegt, kann ebenfalls im Konfigurationsraum oder im kartesischen Raum erfolgen. Pfade, die im kartesischen Raum interpoliert werden, werden als LIN-Bewegungen bezeichnet. Durch LIN-Bewegung erfolgt ein geometrisch linearer Pfad im kartesischen Raum. Im Gegensatz dazu sind Point-to-Point (PTP)-Bewegungen im Konfigurationsraum definiert. Dabei wird für die Bewegung nicht der Pfad des Endeffektors beachtet, sondern die Achsenwinkel der einzelnen Robotergelenke interpoliert [38].

2.4 Mensch-Roboter-Interaktion

Die Mensch-Roboter-Interaktion (engl. Human-Robot-Interaction (HRI)) kann als Unter-einheit der Mensch-Maschinen-Interaktion verstanden werden und befasst sich mit der Gestaltung von Robotersystemen für den Einsatz durch oder mit Menschen. Dafür ist die Kommunikation zwischen Mensch und Roboter erforderlich, die auf unterschiedliche Weise erfolgen kann [8]. Die Interaktionsmöglichkeiten unterscheiden sich dabei hinsichtlich der Interaktionsparadigmen, Interaktionsmodalitäten, Interaktionsschnittstelle und dem Interaktionsgrad [39]. Durch Interaktionsparadigmen wird der Stil bzw. das Modell der Interaktion beschrieben. Die Interaktionsmodalitäten definieren die Kanäle zur Interaktionseingabe, während die Schnittstelle die entsprechende Realisierung beschreibt. Der Interaktionsgrad legt die Aufgabenverteilung zwischen Mensch und Roboter fest.

2.4.1 Interaktionsparadigmen

Nach Pavlovic et al. [40] lassen sich HRI in die in Abbildung 2.1 dargestellten Kategorien unterteilen. Die Interaktionen können beispielsweise in Form einer Geste beabsichtigt oder auch unabsichtlich erfolgen. Die beabsichtigten Interaktionen können in manipulative und kommunikative Intentionen unterteilt werden, wobei manipulative Interaktionen Einfluss auf Objekte einer Umgebung haben. Die Manipulation der Objekte kann in intuitiver Roboterprogrammierung in Steuerbefehle übersetzt werden [41]. Kommunikative Gesten haben hingegen einen inhärenten kommunikativen Zweck und dienen hauptsächlich als Anweisungen für die Bewegungssteuerung des Roboters [41]. Dabei können kommunikative Gesten entweder Handlungen oder Symbole sein. Handlungen sind Interaktionen, die direkt mit der Interpretation einer Bewegung zusammenhängen. Da sie sich unmittelbar auf die beabsichtigte Bedeutung beziehen, sind sie transparent und können ohne weiteres Vorwissen verstanden werden [42]. Dabei können sie bestimmte Handlungen imitieren (mimetisch) oder etwas zeigen (deiktisch). Zeigende Gesten können zur Spezifikation durch eine lokale Referenz in Form von Richtungen, Posen oder Pfade dienen [41]. Symbole

hingegen sind Bewegungsabkürzungen, die nicht ohne vorheriges Lernen verwendet werden können [42]. Symbole können implizit oder explizit sein.

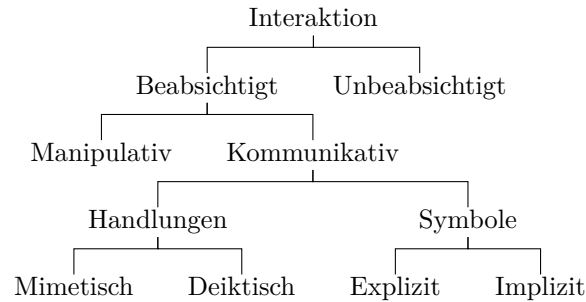


Abbildung 2.1: Übersicht über Mensch-Roboter-Interaktionen (nach [40, 41])

Zusätzlich kann zwischen diskreten und kontinuierlichen Gesten unterschieden werden [43]. Diskrete Gesten sind dabei symbolische Gesten, nach deren vollständiger Durchführung Befehle ausgeführt werden. Bei kontinuierlichen Gesten handelt es sich um manipulative Gesten, die durchgängig interpretiert werden.

Wenn Gesten für die Manipulation von Freiheitsgraden verwendet werden, können diese zusätzlich in die nach Milgram und Zhai [44] definierten drei Dimensionen unterschieden werden (siehe Abbildung 2.2). Diese Dimensionen beschreiben die Abbildungsbeziehung, den Erfassungsmodus und den Grad der Integration.

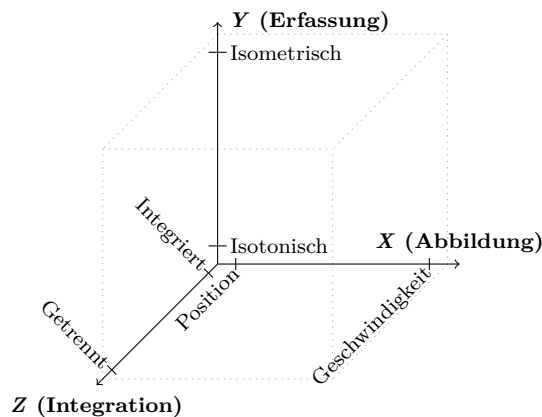


Abbildung 2.2: Dimensionen zur Unterscheidung verschiedener mehrdimensionaler Manipulationsmethoden (nach [44])

Die Abbildungsbeziehungen, die entlang der x-Achse dargestellt sind, beschreiben das Kontinuum der Beziehung zwischen Geste und Auswirkung auf die manipulierten Freiheitsgrade. Nahe des Ursprungs befindet sich die exakte Transformation, dabei werden räumliche Transformationen der Eingabe direkt auf die Transformation des zu manipulierende Objekts abgebildet, wodurch eine 1:1-Kontrolle ermöglicht wird. Dadurch kann eine reale Manipulation simuliert werden. Diese direkten Eingaben können zusätzlich linear oder nichtlinear skaliert werden, um die Genauigkeit zu erhöhen oder einen größeren Bereich

abdecken zu können [45]. Diese Art der Abbildung kann als Positionssteuerung beschrieben werden. Am äußeren Ende der x-Achse befindet sich die Zeitintegration erster Ordnung, sodass durch Gesten eine Geschwindigkeitssteuerung entlang der zu manipulierenden Freiheitsgrade ermöglicht wird.

Entlang der y-Achse werden die Arten der Gestenerfassung dargestellt. Diese werden dabei in die beiden Extreme der isometrischen und isotonischen Erfassung unterteilt.

Die dritte Dimension entlang der z-Achse unterteilt den Grad der Integration. Dabei kann zwischen einer vollständig integrierten und einer vollständig getrennten Steuerung unterschieden werden. In einer vollständig integrierten Steuerung können alle Freiheitsgrade simultan manipuliert werden, während in einer getrennten Steuerung alle Freiheitsgrade einzeln angepasst werden. Zwischen den beiden Extremen befindet sich die Option, dass translatorische und rotatorische Freiheitsgrade jeweils integriert manipuliert werden, jedoch die Modi an sich voneinander getrennt sind, sodass zwei 3-DoF-Steuerungen möglich sind.

2.4.2 Interaktionsmodalitäten

Die Interaktion kann durch unterschiedliche Interaktionsmodalitäten realisiert werden. Allgemein kann zwischen haptischen und berührungslosen Eingabemethoden unterschieden werden. Haptische Eingaben können dabei direkt am Roboter oder indirekt über externe haptische Eingabemodalitäten erfolgen. Die Erfassung der direkten Interaktionen können beispielsweise durch taktile Oberflächen und Kraftsensoren am Roboter realisiert werden. Dadurch können symbolische Gesten auf der Roboterhaut oder eine Handführung ermöglicht werden. Dabei kann Handführung unter Impedanzsteuerung des Roboters mit sehr hoher Nachgiebigkeit (zero gravity control) stattfinden, sodass der Roboter sich nur durch direkten Input des Nutzens bewegt [46]. Für die externe Eingabe können Eingabemodalitäten wie Touchscreens, Maus/Tastatur, Trackballs, Joysticks oder weitere haptische Eingabegeräte verwendet werden [47]. Jedoch konnte nachgewiesen werden, dass berührungsbasierte Schnittstellen zum Auftreten von Infektionen im OP führen können, sodass berührungslose Eingaben besser geeignet sein können [48].

Berührungslose Eingabemethoden können eine natürliche Schnittstelle ermöglichen. Dabei beschreibt eine natürliche Schnittstelle nach Preim und Dachselt [43] ein Interaktionssystem, das Nutzenden die Interaktion mittels intuitiver und direkter Eingaben ermöglicht, die einen klaren Bezug zu natürlichen realen Verhaltensweisen haben.

Wenn zwei oder mehr Eingabemodalitäten parallel oder sequenziell zur Interaktion genutzt werden, wird die Interaktion als multimodal bezeichnet [43]. Diese Kombination an Eingabemodalitäten kann zur Vereinfachung komplexer Aufgaben und zur Erhöhung der Wahrnehmbarkeit und Genauigkeit führen [49]. Dabei können multimodale Interaktionen

komplementär (indem beispielsweise auf ein Objekt gezeigt wird und dazu der Sprachbefehl „nimm das“ gesagt wird) oder redundant (indem beispielsweise „stopp“ gesagt wird und gleichzeitig eine Stopp-Geste gemacht wird) sein [50]. Durch die komplementären Informationen können verlässlichere semantische Bedeutungen gewonnen werden, da die sich ergänzenden Eingaben vage Informationen spezifizieren können [51]. Redundante Informationen können die Robustheit gegenüber Störeinflüssen wie Hintergrundgeräuschen oder Lichtverhältnissen erhöhen, wodurch die Effizienz des Systems gesteigert werden kann [50].

Durch multimodale Interaktionen können somit die Vorteile verschiedener Interaktionsmodalitäten miteinander kombiniert werden, um mögliche Schwächen der einzelnen Methoden zu minimieren.

Während Sprache effektive und natürliche symbolische Eingaben ermöglichen kann, hat sie den Nachteil, dass durch den fehlenden räumlichen Bezug, bestimmte Kommandos wie Positionsangaben nicht kurz und intuitiv angegeben werden können. Durch die Kombination von Sprache mit anderen Eingabemodalitäten können die Sprachbefehle jedoch kürzer und expliziter werden [52]. Somit kann in einer multimodalen Interaktion Sprache für symbolische Angaben wie Kommandos, Parametereinstellungen oder Moduswechsel genutzt werden, während räumliche Angaben über eine andere Modalität realisiert werden [43].

Bei der Spracheingabe kann zwischen Kommandosprache und Interaktion über natürliche Sprache unterscheiden werden [53]. Für Kommandosprache muss eine synthetische Sprache erlernt werden, die eindeutige Eingaben ermöglicht. Im Gegensatz dazu verwendet natürliche Sprache die alltägliche Sprache, wie sie in zwischenmenschlichen Dialogen auftreten würde.

Gesten sind als Bewegung von Händen, Augen, Kopf, Gesicht, Fuß oder weiteren Körperteilen zur kommunikativen Absicht definiert [43]. Im Gegensatz zur Sprache können durch Gesten intuitiv geometrische räumliche Angaben gemacht werden [54]. Es besteht zwar auch die Möglichkeit einfacher symbolischer Gesten (beispielsweise für Start und Stopp), jedoch wird oftmals eine Schnittstelle benötigt, die komplexere und vielfältigere symbolische Eingaben ermöglicht [12]. Somit sind Gesten besser für räumliche Befehle wie Positionsangaben geeignet als für symbolische Eingaben.

Wie in Abbildung 2.3 dargestellt, müssen die von Nutzenden ausgeführten Gesten erfasst, identifiziert und im Fall von dynamischen Gesten getrackt werden. Anschließend müssen die Gesten klassifiziert und entsprechend der zuvor beschriebenen Abbildungsbeziehungen in Roboterbefehle übersetzt und Feedback an Nutzende gegeben werden [55].

Um die Gesten erfassen zu können, müssen die Positionen, Orientierungen und Bewegungen des Körpers aufgenommen werden [56]. Dies kann entweder durch am Körper befestigte Sensoren oder bildbasiert, durch Kameras und Computer Vision ermöglicht werden. Als tragbare Sensoren können beispielsweise Magnetfeld-Tracker, Datenhandschuhe oder

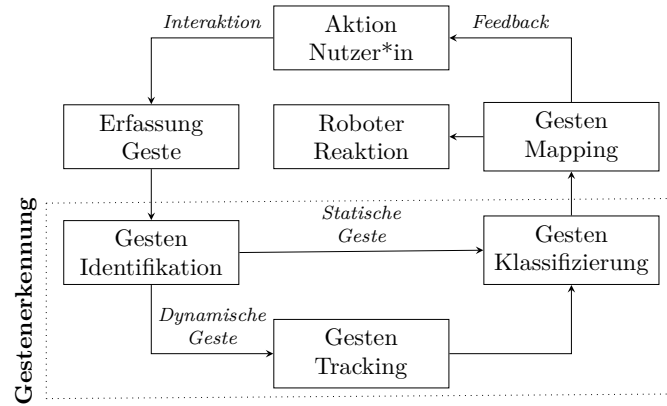


Abbildung 2.3: Überblick über Ablauf und die benötigten Schritte der Mensch-Roboter-Interaktion durch Gesten [55]

Körperanzüge verwendet werden. Durch unhandliche tragbare Sensoren kann jedoch die Nutzerfreundlichkeit und Natürlichkeit der Interaktion eingeschränkt werden. Bildbasierte Trackingverfahren können durch Marker, Tiefensensoren, Stereokameras oder einzelne Kameras ermöglicht werden [24].

Freihandgesten (auch *natural gestures* oder *bare handed gestures*) ermöglichen die Eingabe von Gesten durch die freie Bewegung einer oder beider Hände im Raum. Freihandgesten können bei längerer Nutzung körperlich anstrengend und ermüdend sein, sodass auf Gesten mit minimalem Aufwand geachtet werden und ein schneller Wechsel zwischen Ruhephasen und Interaktion möglich sein sollte [52].

Eine weitere Interaktionsmöglichkeit bilden kontrollierte Augenbewegungen. Diese können in Augengesten, die durch eine vordefinierte Abfolge an Augenbewegungen definiert sind und explizite Blicksteuerung unterteilt werden. In der expliziten Blicksteuerung wird die bewusste Bewegung der Augen für die Interaktion verwendet. Der Vorteil kann darin liegen, dass der visuelle Fokus automatisch auf relevante Objekte gelegt wird, ohne dass ein bewusster Aufwand betrieben werden muss [43]. Dies kann jedoch auch zu Problemen führen: dadurch, dass die Augen immer aktiv sind, da sie in erster Linie zur Wahrnehmung und nicht zur bewussten Manipulation dienen, kann dies zu ungewollt ausgelösten Handlungen führen [57]. Dieses Problem wird das Midas-Touch-Problem genannt. Daher sollten zusätzliche Augengesten wie Zwinkern, Blinzeln oder eine bestimmte Verweilzeit eingeführt werden, um kontrolliertere Eingaben zu ermöglichen [52].

Ähnlich zu der Interaktion über die Bewegung der Augen kann die Orientierung des Kopfes in Form eines vom Kopf ausgehenden Sichtstrahls für die Interaktion verwendet werden [58]. Eine andere Möglichkeit besteht darin, die einzelnen Freiheitsgrade des Kopfes zu nutzen. Dabei können wie in Abbildung 2.4 dargestellt, Rotationsbewegungen (Nicken (engl. pitch), zur Seite Drehen (eng. yaw) oder das seitliche Neigen (engl. roll))

oder Translationsbewegungen in die verschiedenen Raumrichtungen für die Interaktion verwendet werden [43].



Abbildung 2.4: Schematischer Überblick über Rotationen und Translationen des Kopfes und Beispielbewegungen [59]

Bei Gesichtsgesten können verschiedene gestische Gesichtsausdrücke zur Eingabe verwendet werden. Beispiele dafür sind das Heben und Senken der Augenbrauen oder das Rümpfen der Nase [43].

Wenn die Hände bereits mit anderen Aufgaben beschäftigt sind und der visuelle Fokus in einem bestimmten Gebiet liegen muss, sodass weder Hand- noch Kopf- oder Augengesten realisierbar sind, kann der Fuß als Interaktionsmodalität dienen. Dabei können sowohl im Sitzen als auch im Stehen verschiedene Freiheitsgrade des Fußes für die Eingabe verwendet werden [43]. Dafür kann entweder die Position des Fußes oder die ausgeübte Kraft gemessen werden [60]. Jedoch sollte beachtet werden, dass wie auch bei den Freihandgesten, große Bewegungen zu Ermüdungen führen können. Bei der Kraftmessung hingegen können beispielsweise durch Sensoren in den Schuhen feine Veränderungen erkannt werden [60].

2.5 Registrierung

Damit unabhängig von der Position der interagierenden Person die Befehle korrekt ausgeführt werden können, müssen sowohl die Interaktion als auch die ausgeführte Bewegung des Roboters in ein gemeinsames Koordinatensystem transformiert werden.

Eine Methode, die Transformation in das Roboterkoordinatensystem zu ermöglichen, besteht darin, den Roboter zu beliebigen bekannten Positionen im Weltkoordinatensystem zu fahren und dann mithilfe einer Least-Square-Minimization die Position des Roboters zu bestimmen [61, 62]. Ähnlich dazu kann für die Registrierung ein Marker, der in einer bekannten Position im Roboter-Koordinatensystem platziert wird, verwendet werden [63, 64]. Eine weitere Möglichkeit, um die fehlende Transformation zu bestimmen, besteht in der Nutzung einer am Endeffektor montierten Kamera. Dazu können Hand-Eye-Kalibrierungsmethoden verwendet werden, wie beispielsweise die von Tsai et al. [65] vorgestellte. Hierbei

kann, durch die Platzierung eines Markers, der aus verschiedenen Roboterpositionen aufgenommen wird, die Roboterposition und die Trackingdaten aufgenommen werden. Diese ermöglichen dann die Berechnung des Hand-Eye-Algorithmus [66]. Ein ähnlicher Ansatz wäre viele Marker im Raum zu positionieren, deren Positionen durch eine am Endeffektor montierte Kamera aufgenommen werden. Dadurch kann anschließend, durch Bündelblockausgleich, die Position des Roboters bestimmt werden [67].

2.6 Stand der Technik

2.6.1 Robotergestützter Ultraschall

Die häufigste Form der rein manuellen Robotersteuerung ist die teleoperative Steuerung [68]. Dabei wird in der Literatur die Eingabe beispielsweise durch einen Joystick [69], eine 3D-Maus [70], eine Tastatur [71] oder einen Dummy [72] realisiert. Um zusätzliche Rückmeldung über die auf die Ultraschallsonde wirkende Kraft zu erhalten, werden haptische Eingabegeräte verwendet (beispielsweise [68, 73, 74]). Dabei wird ein sogenannter Forward Flow Haptic Control verwendet, indem Bewegungsbefehle durch ein haptisches Eingabegerät aufgenommen, an den Roboter übertragen und von dem Roboter Krafrückmeldungen an das Eingabegerät gesendet werden.

Für die manuelle Robotersteuerung durch Demonstration stellen Janvier et al. [75] ein Teach/Replay System zur 3D-Ultraschalluntersuchung der Gefäße der unteren Gliedmaße vor. Dazu wird in der Lernphase die Ultraschallsonde physisch geführt und die Bewegungen der Sonde aufgezeichnet. Anschließend kann der Roboter in der Replay Phase automatisch entsprechend der demonstrierten Bahnen die Untersuchung durchführen.

Für die Mensch-Roboter-Kooperation verwenden Salcudean et al. [76, 77] und Abolmaen-sumi et al. [78] visuelles Servoing zur Anpassung von bis zu drei Freiheitsgraden, während die weiteren Freiheitsgrade durch Arbeitsteilung gesteuert werden. Die Manipulation der verbleibenden Freiheitsgrade kann durch einen Handcontroller teleoperativ erfolgen. Als Eingabe können dabei ähnliche Möglichkeiten wie bei der Teleoperation genutzt werden (beispielsweise 3D-Maus, Joystick oder haptisches Eingabegerät). Eine weitere Möglichkeit der Kooperation wurde von Fang et al. [79] vorgestellt. Sie verwenden ein duales Kraft-sensorsystem, um eine stabile admittanzgesteuerte Kontaktkraft aufzubringen, während Bediener*innen die Bewegungen anpassen. Dadurch kann die körperliche Belastung während der Bildgebung reduziert werden.

Ein System, das der Zwischenstufe zwischen Mensch-Roboter-Interaktion und autonomen Robotersystemen zugeordnet werden kann, wurde von Kaschwich et al. [80] vorgestellt. Sie verwenden visuelles Servoing um Blutgefäßen zu folgen. Doch bevor die autonome

Bewegung des Systems erfolgen kann, wird eine manuelle Initiierung benötigt, die eine Startposition spezifiziert.

Chatelain et al. [81] stellen ein System für autonomen robotergestützten Ultraschall vor. Dazu haben sie eine ultraschallgestützte Nadelerkennung und -verfolgung in Echtzeit entwickelt, sodass durch eine visuelle 3-DoF Servosteuerung der Roboter die Nadelachse automatisch auf der Längsachse der Sonde ausrichten kann. Langsch et al. [82] realisieren eine autonome Katheter- und Zielverfolgung, mithilfe eines Verfolgungsalgorithmus und einer Kraftsteuerung. Diese basieren auf der Registrierung von preoperativen Bilddaten, die sowohl eine Trajektorie für das Tracking als auch eine visuelle Rückmeldung über Echtzeit-Katheterpositionen ermöglichen. Neben der Verfolgung von medizinischen Instrumenten wie Kathetern und Nadeln können jedoch auch anatomische Strukturen wie Tumore erkannt werden. Dazu haben beispielsweise Mebarki et al. [83] eine Methode für visuelles Servoing basierend auf 2D-US-Bildern entwickelt, die dazu dient Objekte ohne vorheriger Kenntnis ihrer Form zu identifizieren und auf dem US-Bild zu zentrieren.

2.6.2 Mensch-Roboter-Interaktion

Die Mensch-Roboter Interaktion bei Ultraschallrobotern erfolgt in der Literatur hauptsächlich durch haptische Eingabemethoden. Da in dieser Arbeit eine multimodale Interaktion realisiert werden soll, werden nachfolgend verwandte Arbeiten der allgemeinen Mensch-Roboter-Interaktion vorgestellt. Darüber hinaus soll die Schnittstelle über eine natürliche Interaktion ermöglicht werden, sodass weiterhin auf Arbeiten eingegangen wird, die die intuitive Steuerung über Sprache, Hand-, Kopf-, Augen- oder Fußbewegungen ermöglichen. Ein Überblick über die verwandten Arbeiten wird in Tabelle 2.1 dargestellt.

Ferre et al. [84] haben die Eignung von Sprache für die Telemanipulation eines Roboters untersucht. Dabei wurde festgestellt, dass für Low-Level-Kommandos, wie die direkte Steuerung eines Roboters, Handbewegungen besser geeignet sind als Sprache. Dies wird darauf zurückgeführt, dass es komplexer und zeitaufwendiger ist, ein Wort auszusprechen, als eine Hand zu bewegen. Im Gegensatz dazu konnten durch die Erkennung einzelner verbaler Befehle zur Ausführung diskreter, komplexer Befehlen gute Ergebnisse erzielt werden. Dabei wurde festgestellt, dass die Verwendung von Menüs, die durch einzelne Kommandos ausgelöst werden, intuitiver und einfacher zu Nutzen sind als natürliche Sprache.

Um mit dem Roboter Bewegungen der Hand nachzuahmen, tracken Lambrecht et al. [85] Handbewegungen durch einen Kinect Sensor und bilden Veränderungen der Hand kontinuierlich auf Roboterbewegungen ab. Die Robotergeschwindigkeit wird während dieser Bewegung zur Arbeitssicherheit reduziert.

Zusätzlich zur Positionsvorgabe wird von Gîrbacia et al. [86] und Bassily et al. [87] durch die Bewegung der Hand die Rotation des Endeffektors angepasst. Dabei wird die Hand mithilfe eines Leap Motion Sensors getrackt und sowohl Veränderungen der Position als auch der Orientierung direkt auf den Roboter übertragen. Weiterhin verwenden Bassily et al. [87] eine GUI, die die Eingabe komplexer Befehle wie das Fahren zu vordefinierten Positionen oder die Aktivierung/Deaktivierung der Gestenerkennung ermöglichen.

Neben der direkten Übertragung von Bewegungen der Hand stellen Peppoloni et al. [88] die Möglichkeit vor, durch eine Geste einzelne Wegpunkte mithilfe der aktuellen Handposition zu definieren. Aus den Wegpunkten wird ein Pfad gebildet, der anschließend von dem Roboter abgefahren wird.

Quintero et al. [89] verwenden für die Definition diskreter Endeffektorpositionen Zeigegesten, wobei die Zielpositionen durch den 3D-Schnittpunkt mit Objekten definiert werden. Zusätzlich werden symbolische Gesten für das Bestätigen und das Verwerfen von Roboterbewegungen implementiert.

Um neben der diskreten Zielposition auch die Zielorientierung anzupassen, verwenden Hügler et al. [90] eine aus drei Schritten bestehende Interaktion, die durch Augmented Reality unterstützt wird. Für das Festlegen der Position wird eine Zeigegeste verwendet, während die Freiheitsgrade der Orientierung in zwei Schritten durch die Ausrichtung zweier Achsen definiert werden. Für die nachträgliche Feinanpassung nach dem Anfahren des Ziels wird eine Kraftsteuerung implementiert.

Bereits von Bolt [91] konnte gezeigt werden, dass die Kombination von Gesten und natürlicher Sprache zu natürlicheren Mensch-Maschinen-Schnittstellen führt.

Perzanowski et al. [92] nutzen eine Kombination aus natürlichen Sprachbefehlen, Gesten und Eingaben auf einem Handheld Device. Dabei sind redundante Variationen zwischen den Eingaben möglich, wie z.B. der Sprachbefehl „geh nach da“ und das Zeigen auf eine Position, in Kombination mit dem Sprachbefehl auf eine Position auf dem Handheld klicken oder den entsprechenden Befehl auf dem Handheld auswählen und auf die Position zeigen. Auch Rogalla et al. [93] und Maurtua et al. [50] verwenden die Kombination von Gesten und Sprache. Dabei wird durch eine Geste die gewünschte Endeffektorposition definiert und Sprache verwendet, um die Aktion festzulegen.

Iba et al. [54] stellen ein Sprach- und Gesten-Vokabular vor, dass redundante Eingaben für symbolische Befehle wie Stopp ermöglichen. Zusätzlich zu einhändigen Gesten können auch beidhändige Gesten verwendet werden, um relative Positionen zu spezifizieren.

Du et al. [94] kombinieren Sprache und Gesten durch audiovisuelle Fusion basierend auf Text. Dadurch werden beispielsweise statische Gesten mit Sprachbefehlen (beispielsweise in eine Richtung zeigen und sagen „Beweg dich 3 mm in die Richtung“) oder dynamische Gesten in Kombination mit Basisbefehlen (beispielsweise: „Folge meiner Hand“) möglich. Durch die statischen Gesten mit Anweisung können somit feine Bewegungen und durch die dynamischen Gesten mit Grundbefehl grobe Bewegungen ermöglicht werden.

Stiefelhagen et al. [95] verwenden zusätzlich zur Kombination von Sprache mit Gesten die Blickrichtung, indem die Orientierung des Kopfes getrackt wird. Dabei dient die Sprache als Haupteingabemodalität, während die Hand- und Kopfgesten zur Spezifizierung uneindeutiger Befehle dienen.

Krupke et al. [58] vergleichen die Verwendung von Zeigegesten mit Hand und Kopf in Kombination mit Sprache für einen Pick-and-Place Task. Dabei werden die Zeigegesten verwendet, um Objekte, die von dem Roboter gegriffen werden sollen, auszuwählen. Es konnte gezeigt werden, dass die Interaktion über die Kopforientierung präziser und schneller ist. Zusätzlich wurde von Teilnehmenden während der Nutzung der Interaktionsmodalitäten durch die Interaktion mit der Kopforientierung weniger körperliche und geistige Anstrengung empfunden.

Zeigegesten können nicht nur für die Auswahl diskreter Objekte (wie bei Pick-and-Place) Tasks verwendet werden, sondern auch um beliebige Positionen im Raum zu spezifizieren. Dazu verwenden Tostado et al. [96] die Blickrichtung. Durch die Bestimmung des 3D-Fokuspunktes des Auges wird die Zielposition des Roboterendeffektors festgelegt.

Wöhle et al. [97] kombinieren darüber hinaus die Blickrichtung mit der Ausrichtung des Kopfes. Sie verwenden ebenfalls den dreidimensionalen Blickpunkt als Input für die Positionsbestimmung, jedoch wird zusätzlich die Anpassung der Orientierung des Endeffektors durch Veränderung der Kopforientierung ermöglicht. Um zwischen den Modi für die Translation und Rotation zu wechseln, wird das Zwinkern des rechten Auges verwendet.

Williams et al. [98] verwenden die Rotation des Kopfes, um die Endeffektorposition eines Roboters zu manipulieren. Dabei wird die Änderung des Pitch-Winkels auf Bewegungen nach oben und unten, die des Yaw-Winkels für Bewegungen von links nach rechts und die des Roll-Winkels für Bewegungen in die Tiefe abgebildet.

Um nicht nur die Translation, sondern auch die Rotation des Endeffektors anzupassen, führen Rudigkeit et al. [99] unterschiedliche Modi ein, die aktiviert werden können, indem durch den Sichtstrahl des Kopfes ein Slider-Button ausgewählt und verschoben wird. Im Gegensatz zu Williams et al. [98] unterteilen sie die Translation in zwei Modi. Dabei kann in beiden Modi die Bewegung nach links und rechts durch Änderung des Roll-Winkels manipuliert werden. Zusätzlich dazu wird der Pitch-Winkel genutzt, um im vertikalen Modus die Höhe und im horizontalen Modus die Tiefe der Endeffektorposition anzupassen. Im Orientierungsmodus werden die Änderungen der Kopforientierung auf analoge Rotationen des Endeffektors übertragen. Um aus der Interaktion resultierende Roboterbewegung skalieren zu können, wird eine Transferfunktion verwendet, die Auslenkungen des Kopfes auf Robotergeschwindigkeiten abbildet.

Neben der Aktivierung der Modi durch das Verschieben eines Slider-Buttons werden auch das lange Fokussieren eines durch die Kopfrichtung gesteuerten Cursor auf einen Button [100] oder Kopfgesten wie Nicken oder Kippen des Kopfes [101] verwendet.

Ein anderer Ansatz zur Verwendung der Blickrichtung ist, den Fokus des Blicks auf verschiedene dynamische Felder für die Interaktion zu nutzen. Dadurch kann der visuelle Fokus auf verschiedene interaktive Schaltflächen zu Veränderungen der Freiheitsgrade des Roboters führen. Beispielsweise Alsharif [102] verwendet verschiedene Modi (Translation X-Y, Translation Z, Rotation X-Y und Rotation Z), die durch eine Blicksequenz (beispielsweise: Blick nach links-rechts-links) und Zwinkern aktiviert werden können. Innerhalb dieser Modi werden um den Endeffektor des Roboters verschiedene dynamische Kommandobereiche initiiert. Wenn der Blick für eine bestimmte Zeit auf diese Bereiche fokussiert ist, wird der entsprechende Befehl ausgeführt.

Park et al. [103] verwenden zusätzlich Augmented Reality, um die Nutzeroberfläche zu visualisieren. Dabei werden nicht direkt die Freiheitsgrade des Endeffektors manipuliert, sondern virtuelle Objekte an den gewünschten Roboterpositionen platziert. Dazu können über die Blickrichtung verschiedene Modi ausgewählt werden, wobei sowohl die grobe Anpassung der Freiheitsgrade durch den Blick auf entsprechende Felder als auch die feine Anpassung durch die Orientierung des Kopfes möglich ist. Für die grobe Ausrichtung wird durch den Blick auf entsprechende dargestellte Felder das virtuelle Objekt entweder um 10 cm oder 90° bewegt. Für die feine Anpassung wird zwischen Translation X-Z, Translation X-Y, Rotation um Y und Rotation um die Achsen des Kamerakoordinatensystems unterschieden, in denen durch Veränderungen des Pitch, Roll, oder Yaw-Winkels des Kopfes die Manipulation der Freiheitsgrade ermöglicht wird.

Huang et al. [60] verwenden die Orientierung des Fußes, um vier Freiheitsgrade eines Roboters zu manipulieren. Dazu wird ein aus Federn und Kraftsensoren bestehendes Interface verwendet, das die Abbildung der Rotation und Vorwärtsbewegung des Fußes auf die Rotation und die Vor- bzw. Rückwärtsbewegung des Roboters ermöglicht.

Kawai et al. [104] nutzen eine Kraftsensorplatte, die temporäre Veränderungen des Fußschwerpunkts misst. Dadurch werden verschiedene Fußbewegungen genutzt, die in Rotation und Vor- und Rückwärtsbewegung eines chirurgischen Roboters umgewandelt werden.

Tabelle 2.1: Übersicht über in der Literatur verwendete Eingabemodalitäten für natürliche Schnittstellen zur Mensch-Roboter-Interaktion

Modalität	Beschreibung	Verwendung
Sprache	Kommandosprache	[58], [84], [86], [94], [103]
	Natürliche Sprache	[50], [84], [91], [92], [93], [94], [95]

Fortsetzung auf nächster Seite

Tabelle 2.1: Fortsetzung

Modalität	Beschreibung	Verwendung
Hand	Symbolische Gesten	[50], [54], [89], [93]
	Direkte Übertragung von Handbewegung auf Robotertranslation	[85], [86], [87], [88], [94]
	Definition von Zielposition durch Zeigegeste	[47], [50], [58], [89], [90], [91], [92], [94]
	Translation über relative Position von zwei Händen	[54]
	Translation über Definition von Wegpunkten	[88]
	Direkte Übertragung von Handorientierung auf Roboterrotation	[86], [87]
	Rotation durch Anpassung einzelner Achsen	[90]
Augen	Augengesten	[96], [97], [102]
	Blickrichtung zur Auswahl von Modi	[103]
	3D-Blickpunkt für Zieldefinition	[96], [97]
	Translation und Rotation durch Blick auf Schaltflächen	[102], [103]
Kopf	Kopfgesten	[101]
	Kopfausrichtung (Sichtstrahl) zur Auswahl von Modi	[100], [105]
	Kopfausrichtung (Sichtstrahl) für Zieldefinition	[58], [95]
	Robotertranslation durch Kopfrotation	[98], [100], [103], [105]
	Roboterrotation durch Kopfrotation	[97], [100], [105], [103]
Fuß	Translation und Rotation durch isometrische Erfassung	[60]
	Translation und Rotation durch isotonische Erfassung	[104]
Multimodal	Kombination von Handgesten und Sprache	[50], [58], [91], [92], [93], [94]
	Kombination Kopfinteraktion und Sprache	[58]
	Kombination Handgesten, Sprache und Kopfbewegungen	[95]
	Kombination Kopf- und Blickinteraktion	[97]

3 Anforderungsanalyse

In dieser Arbeit soll eine Anwendung zur Unterstützung von Radiolog*innen während nadelgestützter Interventionen, die unter kombinierter Bildgebung von CT und Ultraschall stattfinden, entwickelt werden. Dazu soll der Ultraschall robotergestützt durchgeführt werden, indem die Ultraschallsonde an dem Endeffektor eines Roboters befestigt wird. Die Anwendung soll die Interaktion mit dem Roboter zur manuellen Steuerung der Bildgebung ermöglichen. Die zu erfüllenden Anforderungen an ein entsprechendes Robotersystem und die zugehörige Mensch-Roboter-Schnittstelle werden nachfolgend erarbeitet.

3.1 User Research

Für die Identifikation der Handlungsabläufe und der Nutzungsbedürfnisse wurde ein Experteninterview mit vier Radiolog*innen des Universitätsklinikums Magdeburg durchgeführt. Die Teilnehmenden wurden zum aktuellen Workflow bei nadelgestützten Interventionen unter kombinierter CT- und Ultraschall-Bildgebung, möglichen Use Cases und Anforderungen an robotergestützten Ultraschall befragt. Der dazu verwendete Fragebogen befindet sich in Anhang A.1.1. Aus den von Expert*innen während des Interviews mehrfach getätigten Aussagen (siehe Anhang A.2) geht der im Folgenden beschriebene Workflow hervor.

3.1.1 Workflowanalyse

Nadelgestützte Interventionen unter kombinierter Bildgebung von CT und Ultraschall werden insbesondere bei Läsionen in der Niere und der Leber angewendet, da sich zum einen die Bildgebungsmodalitäten für die Visualisierung dieser Organe eignen und zum anderen die Läsionen gut erreichbar sind.

Vor der Intervention müssen Patient*innen vorbereitet werden, indem sie steril abgedeckt und entsprechend der Voraufnahmen positioniert werden. Die meisten Interventionen finden unter lokaler Anästhesie statt. Anschließend wird ein Probe-CT des gesamten Ziel-Volumens ohne Kontrastmittel aufgenommen. Basierend auf dieser Aufnahme kann der Einstichpunkt und die Richtung des Nadelpfades bestimmt werden. Der Beginn der Intervention findet unter Fluoroskopie statt, somit werden einzelne axiale Schichten aufgenommen. Da anguliert interveniert wird, sich der Pfad der Nadel also nicht auf der

Schichtführung der CT-Aufnahmen befindet, wird immer die axiale Schicht aufgenommen, in der sich die Nadelspitze befindet. Um mit den Aufnahmen dem Nadelvorschub folgen zu können, kann entweder der Tisch oder die Gantry (bei iCT) verschoben werden. Ultraschall ist für die Navigation zu Beginn der Intervention nicht geeignet, da dieser durch Luft, Fett und Knochen limitiert ist, wodurch für die ersten Zentimeter unter der Hautoberfläche keine gute Aufnahmequalität möglich ist. Da im CT feine Strukturen, wie Risikostrukturen, die sich auf dem Weg zur Läsion befinden, nicht erkennbar sind und die Navigation auf den axialen Schichten herausfordernd ist, wird für die Feinjustierung zu Ultraschallbildgebung gewechselt. Dabei ist, wie zuvor im CT, der Fokus der Bildgebung auf der Nadelspitze. Da die Ultraschallbildgebung jedoch nicht auf axiale Schichten begrenzt ist, ist es möglich zusätzlich den Bereich vor der Nadelspitze zu visualisieren. Daher wird durch den Ultraschall der Winkel der Nadel nachgeahmt, sodass sowohl die Nadelspitze als auch davor liegende Risikostrukturen visualisiert werden können. Somit kann der Ultraschall als eine Art Taschenlampe dienen. Dadurch kann in-plane gestochen werden, ohne dass die Ebene axial ist. Neben der Visualisierung der Nadelspitze kann auch die Visualisierung der Läsion nötig sein, um zu überprüfen, ob sich die Nadel auf dem richtigen Pfad befindet. Wenn die Nadel erfolgreich in der Zielläsion platziert ist, wird ein Kontroll-CT aufgenommen, um die finale Positionierung zu überprüfen.

3.1.2 Interaktionsaufgabe

Somit ergibt sich während der Ultraschallbildgebung der in Abbildung 3.1 dargestellte Arbeitsablauf. Zur Unterstützung während dieser Schritte kann ein Roboterarm verwendet werden, der die Ultraschallsonde positioniert. Dabei könnte die kontinuierliche Visualisierung der Nadelspitze autonom durchgeführt werden. Neben dieser automatischen Steuerung gibt es jedoch weitere Arbeitsschritte, die manuell durch die Kommunikation zwischen Mensch und Roboter erfolgen müssen. Um für diese Interaktion eine geeignete Schnittstelle entwickeln zu können, müssen die aus dem Arbeitsablauf resultierenden Interaktionsaufgaben identifiziert werden.

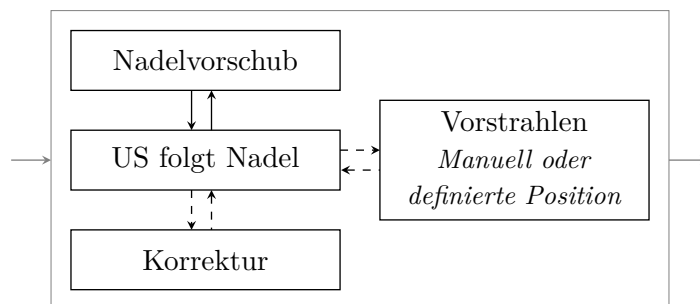


Abbildung 3.1: Arbeitsablauf während ultraschallgestützter Nadelführung

Bei dem Wechsel zwischen CT- und US-Bildgebung muss zunächst der Roboter aktiviert werden. Die US-Sonde soll mithilfe des Roboters die Fähigkeit haben, automatisch der Nadelspitze zu folgen und dabei ihren Winkel nachzuahmen. Dazu muss die automatische Nachführung aktiviert und deaktiviert werden können.

Da bei der autonomen Roboterbewegung Fehler auftreten oder Strukturen wie Knochen die Nadel verdecken können, kann eine Korrektur der Ultraschallposition notwendig sein. Dazu muss die manuelle Einstellung der Freiheitsgrade durch Nutzende möglich sein, sodass die Translation und Rotation der Ultraschallsonde angepasst werden können.

Die manuelle Navigation der Ultraschallsonde kann auch beim Vorstrahlen notwendig sein. Dies dient zur Visualisierung von Strukturen, die auf dem Weg zur Läsion liegen oder zur Darstellung der Läsion an sich. Neben der manuellen Navigation der Sonde zu diesen Strukturen sollte es auch möglich sein, diese Sondenpositionen zu speichern, um die Positionen zu einem späteren Zeitpunkt erneut betrachten zu können. Abschließend soll zusätzlich die Möglichkeit bestehen, den Roboterarm und die Sonde aus dem Arbeitsbereich zu bewegen, beispielsweise um das Abschluss-CT aufzunehmen und den Roboter zu deaktivieren.

3.2 Funktionale Anforderungen

Im Folgenden werden die funktionalen Anforderungen festgelegt, die beschreiben, was das System leisten soll. Eine Übersicht dazu ist in Tabelle 3.1 dargestellt. Dabei soll das System eine multimodale Interaktion mit dem Roboter ermöglichen, der während der nadelgestützten Intervention die Ultraschallsonde positioniert. In dieser Arbeit wird die autonome Roboterbewegung zur Darstellung der Nadel vernachlässigt und der Fokus auf die manuelle Bewegung des Roboters zum Vorstrahlen und der Korrektur des autonomen Systems gelegt. Somit soll ein System ohne autonome Anteile (Level 0, vgl. Abschnitt 2.2) entwickelt werden, das in der Anwendung zur Unterstützung eines autonomen Systems (Level 2 oder 3) dienen soll.

Daraus ergeben sich folgende technische Anforderungen: Zur Kommunikation mit dem System soll die multimodale Interaktion möglich sein. Dazu müssen die über verschiedene Kommunikationskanäle erfolgenden Interaktionen mit dem System erkannt und in Befehle umgewandelt werden. Diese müssen an den Roboter gesendet, vom Roboter interpretiert und ausgeführt werden. Um räumliche Interaktionen in korrekte Reaktionen des Roboters übertragen zu können, müssen die Interaktionen in Zusammenhang mit dem Roboterkoordinatensystem gebracht werden. Durch die Interaktion soll als Grundfunktionalität die Aktivierung und Deaktivierung unterschiedlicher Modi ermöglicht werden, die beispielsweise die Aktivierung des Roboters, des Nadeltrackings oder die manuelle Navigation umsetzen. Für diese Navigation des Roboters werden weitere Interaktionen

benötigt, die die Anpassung der Freiheitsgrade für die Translation und die Rotation der Ultraschallsonde ermöglichen. Zusätzlich soll es eine Schnittstelle geben, die das Anfahren von zuvor gespeicherten und vordefinierten Positionen realisiert.

Tabelle 3.1: Übersicht über funktionale Anforderungen

Anforderung	Beschreibung
Technische Rahmenbedingungen	
Multimodale Schnittstelle	Interaktionen müssen erkannt, getrackt, in Roboterbefehle umgewandelt und an den Roboter gesendet werden
Registrierung	Räumliche Interaktionen müssen in korrekte Roboter-Reaktionen transformiert werden können
Grundfunktionalitäten	
Aktivierung Modi	Unterschiedliche Modi wie automatische Nadelführung oder Modi zur Navigation müssen aktiviert/deaktiviert werden
Navigation Roboter	
Anpassung Ausrichtung	Rotation der Ultraschallsonde in drei Raumrichtungen
Anpassung Position	Translation der Ultraschallsonde in drei Raumrichtungen
Gespeicherte Positionen	Speicherung aktueller Roboterposition und späteres Anfahren dieser Position
Vordefinierte Positionen	Roboter muss, z.B. nach vollendeter Interaktion, in eine vordefinierte, nicht störende Position gefahren werden

3.3 Nichtfunktionale Anforderungen

Durch die nichtfunktionalen Anforderungen wird definiert, wie ein System arbeiten soll, somit in welcher Qualität die funktionalen Anforderungen zu erbringen sind. Diese sind in Tabelle 3.2 dargestellt.

Die erste Anforderung ist dabei die Gebrauchstauglichkeit oder auch Usability. Diese beschreibt „Das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen“ [106]. Die Effektivität beschreibt dabei, wie gut das Ziel erreicht werden kann, also inwiefern durch die Interaktion mit dem System die Bildgebung

der relevanten Bereiche ermöglicht wird. Dies soll effizient möglich sein, also unter geringem kognitiven, körperlichen und zeitlichen Aufwand. Zu der Effizienz zählt zusätzlich, dass das System schnell erlernbar sein soll. Weiterhin sollte die Nutzerzufriedenheit gegeben sein, dazu sollte das System natürlich und intuitiv sein, wodurch es von Nutzenden akzeptiert wird.

Tabelle 3.2: Übersicht über nicht funktionale Anforderungen

Anforderungen	Beschreibung
Gebrauchstauglichkeit	
Effektivität	Durch das System soll Nutzenden eine genaue und sichere Nadelnavigation möglich sein, indem die Ultraschallsonde präzise zu für die Visualisierung relevanten Positionen navigiert werden kann
Effizienz	Die Interaktion mit dem System soll mit geringem kognitiven, körperlichen und zeitlichen Aufwand zu präzisen Ergebnissen führen und schnell erlernbar sein
Zufriedenheit	Das System sollte natürlich und intuitiv sein und dadurch von Nutzenden akzeptiert werden
Leistung	
Robustheit	Die Erkennung der Interaktionen soll stabil und fehlertolerant erfolgen, dabei sollen Fehleingaben sowie unabsichtliches Auslösen vermieden und gegebenenfalls korrigiert werden
Performance	Interaktion soll in Echtzeit möglich sein
Klinische Anforderungen	
Sterilität	Das System darf keine hygienischen Herausforderungen verursachen
Arbeitsumfeld	Das Arbeitsumfeld sollte beachtet werden, sodass vorhandene Arbeitsabläufe durch die Interaktion mit dem System nicht gestört und der Zugang zu Patient*innen nicht beschränkt werden
Feedback	
Rückmeldung Systemzustand	Der Systemzustand sollte kontinuierlich sichtbar sein und Nutzende Feedback zu den Interaktionen erhalten

Die zweite Anforderung umfasst die Leistungen, die von dem System erbracht werden sollen. Das System sollte robust sein, sodass Interaktionen zuverlässig erkannt und mögliche Fehleingaben sowie unabsichtliches Auslösen vermieden oder gegebenenfalls korrigiert werden können. Zusätzlich ist die Performance wichtig, sodass das System in Echtzeit

erfolgen kann. Dies ist notwendig, da Verzögerung der Erkennung der Eingaben sich auf die Gebrauchstauglichkeit auswirken könnten.

Weiterhin bestehen an das System klinische Anforderungen, die durch das Anwendungsumfeld entstehen. Das System muss steril sein, sodass keine hygienischen Herausforderungen durch die Verwendung entstehen. Zusätzlich muss die Arbeitsumgebung beachtet werden, sodass der Zugang zu Patient*innen nicht beschränkt wird, um weiterhin eine effiziente Intervention ohne ergonomische Herausforderungen zu ermöglichen. Zusätzlich soll das System möglichst wenig invasiv in die bestehenden Arbeitsabläufe sein, sodass die Intervention erfolgreich verlaufen kann.

Abschließend ist ein kontinuierliches Feedback über den Systemzustand notwendig, um Nutzende bei der Verwendung der unterschiedlichen Modi zu unterstützen und somit eine effiziente Interaktion ermöglichen zu können.

4 Konzeption

Im Folgenden werden basierend auf den vorgestellten Grundlagen und den spezifizierten Anforderungen verschiedene Interaktionskonzepte für die Mensch-Roboter-Interaktion entwickelt. Zusätzlich werden für die Evaluierung dieser Konzepte Interaktionsaufgaben entworfen.

4.1 Konzeption der Rahmenbedingungen

4.1.1 Multimodale Schnittstelle

Für die Auswahl einer geeigneten Schnittstelle zur Realisierung der Interaktion mit dem Roboter müssen zunächst die Interaktionsmodalitäten festgelegt werden. Basierend auf den ausgewählten Interaktionsmodalitäten kann anschließend eine geeignete Schnittstelle für die Erkennung und das Tracking der entsprechenden Interaktionen identifiziert werden.

Interaktionsmodalitäten

Für die Auswahl der Modalitäten wird sich auf die in den zuvorgehenden Abschnitten beschriebene Literaturrecherche gestützt. Die in Abschnitt 3.1.2 identifizierten Interaktionsaufgaben lassen sich in diskrete Interaktionen und Navigationsinteraktionen unterteilen, die durch verschiedene Interaktionsmodalitäten erfüllt werden können.

Diskrete Interaktion

Wie aus Abschnitt 3.1.2 hervorgeht, wird eine Schnittstelle für die Eingabe diskreter Bewegungsbefehle sowie für die Aktivierung und Deaktivierung unterschiedlicher Modi benötigt. Zu den diskreten Bewegungsbefehlen zählen dabei das Fahren zu vordefinierten Punkten sowie die Speicherung von Roboterpositionen und das Anfahren dieser gespeicherten Positionen. Somit wird eine Eingabemodalität benötigt, die die intuitive Eingabe verschiedener diskreter Kommandos ermöglicht.

Wenn symbolische Gesten für die Eingabe von Kommandos verwendet werden, sollte darauf geachtet werden, dass diese für Nutzende leicht zu merken sind und durch das System robust erkannt werden [43]. Daher sollte nur eine geringe Anzahl unterschiedlicher Gesten eingesetzt werden, um Funktionen auf intuitive und bedeutsame Gesten abbilden zu können [107]. Da für die angestrebte Interaktion eine große Auswahl diskreter Befehle

benötigt wird, wird sich gegen die Verwendung von Gesten entschieden.

Kontrollierte Augenbewegungen können für die Eingabe von Kommandos verwendet werden, indem verschiedenen Regionen im Blickfeld Funktionen zugeordnet werden. Dies hat den Vorteil, dass die Bewegung der Augen natürlich ist, wenig bewusste Anstrengung erfordert und das zeitgleiche Ausführen anderer Aufgaben mit den Händen ermöglicht [108]. Dabei kann die Auswahl der Zielregionen unter der Voraussetzung, dass die Zielregionen nicht zu klein sind, sehr schnell sein [109].

Wie in Abschnitt 2.4.2 beschrieben, ist Spracherkennung eine verbreitete Eingabemodalität für symbolische Kommandos, da sie insbesondere in multimodalen Interaktionen gut geeignet ist, um die effiziente Übermittlung komplexer Befehle zu ermöglichen, ohne die Hände verwenden zu müssen [110].

Da sowohl durch die Verwendung von blickgesteuerten Buttons als auch von Sprache die Eingabe einer Vielzahl verschiedener Befehle ermöglicht werden, kann ohne räumliche Interaktionen zu verlangen, die die Ergonomie einschränken könnten, wurde sich für die Implementierung dieser beiden Interaktionsmodalitäten für die diskrete Interaktion entschieden.

Navigationsinteraktion

Die Navigationsinteraktion wird, wie in Abschnitt 3.1.2 beschrieben, benötigt, um den Roboter manuell in beliebige Positionen und Orientierungen zu bewegen. Dazu müssen drei rotatorische und drei translatorische Freiheitsgrade manipuliert werden. In der Literatur werden dazu zum einen kontinuierliche Bewegungsvorgaben und zum anderen diskrete Zielvorgaben verwendet. Da zur Verifikation der Position der Ultraschallsonde das Ultraschallbild ausgewertet werden muss und Bewegungsvorgaben sukzessiv auf Veränderungen des Ultraschallbildes aufbauen, ist eine kontinuierliche Bewegungsvorgabe notwendig. Dabei soll die Eingabe der Freiheitsgrade möglichst direkt erfolgen, sodass auf die Verwendung virtueller und realer Schaltflächen, die zur Eingabe von Manipulationsbefehlen dienen, verzichtet wird. Daher werden, um eine mindestens teilweise integrierte Manipulation der Freiheitsgrade zu ermöglichen, Eingabemodalitäten benötigt, die die kontinuierliche Eingabe von mindestens drei Freiheitsgraden unterstützen. Da Gesten von sich aus einen direkten räumlichen Bezug haben, können sie für die Manipulation von Freiheitsgraden verwendet werden [43]. Dabei sind insbesondere Hand-, Kopf- und Fußgesten geeignet, da diese die Eingabe von mindestens drei unterschiedlichen Freiheitsgraden ermöglichen.

Fußgesten, die im Stehen durchgeführt werden, sind durch das Halten des Gleichgewichts nur eingeschränkt einsetzbar [111]. Da für die Vorgabe von Bewegungen die Gesten gegebenenfalls über längere Zeiträume gehalten werden müssen, bis der Roboter die gewünschte Position erreicht hat, können ergonomische Schwierigkeiten resultieren. Zusätzlich ist die Präzision von Fußinteraktionen begrenzt [43].

Im Gegensatz dazu bieten die Hände und der Kopf den Vorteil, dass sie ohne das Gleichge-

wicht zu beeinträchtigen, in beliebigen Positionen gehalten werden können. Da durch die Verwendung des Roboters die körperliche Anstrengung der Chirurg*innen gesenkt werden soll, wurde sich daher für die Interaktion mit Hand und Kopf entschieden.

Schnittstelle

Für die Interaktion mit dem Roboter müssen somit Nutzereingaben über Sprache, Blick, Hand- und Kopfbewegungen aufgenommen, in Befehle transformiert und an den Roboter gesendet werden, der diese interpretiert und entsprechend reagiert. Insbesondere für die diskreten Befehle wird Feedback benötigt, ob diese erkannt und somit Modi aktiviert worden sind. Für die Interaktion mittels Blick werden zusätzlich interaktive Schaltflächen im Blickfeld benötigt, die die Auswahl der Modi ermöglichen. Für die Navigationsinteraktion muss neben der robusten Erkennung der Position und Orientierung der Hand und des Kopfes auch die Lokalisierung dieser im Verhältnis zum Roboter ermöglicht werden, um die Bewegungen korrekt abbilden zu können.

Diese vielfältigen Anforderungen könnten durch die Verwendung unterschiedlicher, auf die Anforderungen spezialisierter Hardware realisiert werden. Somit könnten beispielsweise die Hand, die Augen und der Kopf durch jeweils externe Sensoren getrackt werden. Jedoch wurde sich für die Realisierung des Prototypen für die Auswahl einer einzelnen Hardware, die die Erfassung aller für die Interaktion benötigter Daten ermöglicht.

Interaktive Head-Mounted-Displays (HMDs) ermöglichen sowohl die Eingabe verschiedener Interaktionen als auch das Anzeigen visueller Informationen über ein mobiles, am Kopf angebrachten Systems [112]. Sie bestehen aus einem See-Through Display und unterschiedlichen integrierten Sensoren. Bei dem ersten kommerziell erhältlichen und populärsten HMD handelt es sich um die Microsoft HoloLens [113]. Dabei bietet die zweite Generation der HoloLens einige Verbesserungen gegenüber der ersten Generation, unter anderem ein größeres Sichtfeld und vollständiges Kopf- und Augen-Tracking [114]. Die HoloLens 2 verfügt über Kopftracking durch vier sichtbare Lichtkameras, Augentracking durch zwei Infrarotkameras, einen Time-of-Flight-Tiefensensor, eine Trägheitsmesseinheit (Inertiale Messeinheit (IMU), mit Beschleunigungsmesser, Gyroskop und Magnetometer), eine RGB-Kamera, ein Mikrofon-Array und Lautsprecher [115]. Dabei wird die Position der Nutzenden im Raum über Inside-Out-Tracking bestimmt, indem über zwei Kameras Features in der Umgebung beobachtet und mit Informationen der IMU verbunden werden. Somit können mithilfe der HoloLens die benötigten Daten für die Sprach-, Blick-, Hand- und Kopffinteraktion aufgenommen werden. Zusätzlich können durch das semitransparente Display visuelle Informationen angezeigt und in Kombination mit den Lautsprechern Feedback gegeben werden.

4.1.2 Registrierung

Um die Interaktionen, die mithilfe der Kamera oder den Bewegungssensoren der HoloLens aufgenommen werden, korrekt auf Bewegungen des Roboters abbilden zu können, muss die Transformation zwischen diesen Koordinatensystemen (KOS) bestimmt werden, sie müssen somit registriert sein.

Für die Lokalisierung der HoloLens im Raum können optische Marker verwendet werden, die durch die RGB-Kamera der HoloLens identifiziert und somit das räumliche Verhältnis bestimmt werden kann. Um das KameraKOS zusätzlich mit dem RoboterKOS registrieren zu können, muss außerdem die Position und Ausrichtung des Roboters im Verhältnis zum Marker bekannt sein. Dies kann wie in Abschnitt 4.1.2 beschrieben, optisch oder mechanisch bestimmt werden. Da in der vorgesehenen Anwendung keine Kamera am Endeffektor montiert wird, wird von optischen Lokalisierungsmethoden abgesehen. Stattdessen kann das Verhältnis zwischen dem Marker und dem Roboter mechanisch bestimmt werden, indem der Endeffektor zu im MarkerKOS bekannten Positionen gefahren wird.

4.2 Konzeption der Interaktion

4.2.1 Diskrete Interaktionen

Blickinteraktion

Um bei der blickbasierten Betätigung interaktiver Schaltflächen zur Auswahl von Kommandos das Midas-Touch-Problem zu vermeiden, können wie in Abschnitt 2.4.2 beschrieben, zusätzliche Eingaben wie Handgesten, Blinzeln, Zwinkern oder eine Verweilzeit verwendet werden.

Handgesten haben den Nachteil, dass sie durch Fehler in der Gestenerkennung frustrierend und sowohl physisch als auch mental anstrengend sein können [116]. Zusätzlich wird dazu die Bewegung der Hand benötigt, die in dem vorgestellten Interaktionskonzept bereits für die Navigationsinteraktion verwendet wird.

Im Gegensatz dazu kann durch Blinzeln oder Zwinkern eine freihändige Eingabe ermöglicht werden. Jedoch muss für die Abgrenzung zwischen natürlichem und beabsichtigtem Blinzeln ein verlängertes Blinzeln verwendet werden [117], woraus eine fehlende Blickkontrolle für die möglicherweise gleichzeitig stattfindende Roboterbewegung resultieren kann. Während Zwinkern die kontinuierliche Blickkontrolle ermöglicht, kann die Umsetzung jedoch für einige Menschen physisch anstrengend sein [118].

Die Auswahlmöglichkeit, den Blick für eine bestimmte Verweilzeit auf die entsprechende Schaltfläche zu richten, bietet den Vorteil, wenig fälschliche Aktivierung zu verursachen [109]. Dafür nimmt diese Bestätigungsart mehr Zeit in Anspruch [116, 119].

Um eine robuste und freihändige Blickinteraktion zu ermöglichen, die für Nutzende unter geringem physischen und mentalen Aufwand umsetzbar ist und auf das längere Schließen der Augen verzichtet, wird sich für die Implementierung der Bestätigung der Selektion der Schaltflächen über eine Verweilzeit entschieden.

Sprachinteraktion

Bei der Interaktion mittels Sprache kann zwischen natürlicher und Kommandosprache unterschieden werden. Während bei der natürlichen Sprache der Lernaufwand geringer ist, ist Kommandosprache effizienter und kann robuster erkannt werden [53]. Zusätzlich kann die Flüssigkeit natürlicher Sprache dadurch beeinträchtigt werden, dass für die Erkennung jedes Wort deutlich ausgesprochen werden muss. Dadurch kann die Verwendung von Kommandosprache, in der einzelne verbale Kommandos eindeutig interpretiert werden, intuitiver und einfacher sein [84].

Da für die diskrete Interaktion robust eindeutige Befehle erkannt werden müssen, wird für die Implementierung Kommandosprache verwendet.

Für die Spracherkennung können kommerzielle Spracherkennungssysteme wie die Microsoft Speech API oder die Google Speech API oder Open-Source-Spracherkennungssysteme wie Sphinx-4 verwendet werden. Im Vergleich der drei zuvor genannten konnte durch die Google Speech API und die Microsoft Speech API geringere Wort-Fehler-Raten erzielt werden, wobei die Google Speech API die geringste Wort-Fehler-Rate hatte [120]. Jedoch hat die Google Speech API den Nachteil, dass die Verarbeitung über die Cloud erfolgt, sodass die Spracherkennung nicht offline möglich ist [121]. Da zusätzlich die HoloLens eine Schnittstelle zur Microsoft Speech API bereitstellt, wurde sich für die Verwendung dieses Spracherkennungssystems entschieden.

4.2.2 Navigationsinteraktion

Wie zuvor beschrieben, soll die Navigationsinteraktion zur Manipulation der Translation und der Rotation des Roboters verwendet werden. Da bei komplexen Manipulationsaufgaben die Translation und Rotation von Nutzenden meistens nicht gleichzeitig durchgeführt werden [122], wurde sich sowohl bei der Hand- als auch bei der Kopfeingabe für eine Trennung der translatorischen und rotatorischen Freiheitsgrade entschieden. Durch diese Trennung können zusätzlich Fehleingaben wie leichte Rotationen bei der Translation vermieden werden [43, 123]

Kopferinteraktion

Für die Interaktion durch Kopferotation muss die Änderung der Ausrichtung des Kopfes erfasst werden. Dazu kann die Tatsache genutzt werden, dass ein Sensor, der an dem Kopf

befestigt wird, sich während der Rotation des Kopfes auf einer sphärischen Oberfläche bewegt [124]. Wenn die Ausrichtung der Yaw-Achse dieses Sensors mit der des Kopfes übereinstimmt, sind Änderungen der Kopf- und Sensorausrichtung identisch [99]. Dadurch wird ermöglicht, dass die von der HoloLens bereitgestellten Informationen über die Ausrichtung des Kamerakoordinatensystems für die Erfassung der Kopffrotation verwendet werden können. Dabei muss lediglich beachtet werden, dass die Ausrichtung des Kamerakoordinatensystems im Gegensatz zu den in Abschnitt 2.4.2 vorgestellten Achsen der Kopforientierung rotiert ist, da die Blickrichtung der Kamera durch die z-Achse definiert wird.

Sowohl für die Umsetzung der Rotation als auch für die der Translation besteht die Möglichkeit, die verschiedenen Freiheitsgrade getrennt oder kontinuierlich abzubilden. Williams et al. [98] haben in einer Studie die Rotation des Kopfes für die Translation eines Roboters verwendet, dabei wurde festgestellt, dass Teilnehmende die Freiheitsgrade getrennt und nacheinander ausgeführt haben. Dies führten sie auf die Anforderung zurück, die 3D-Bewegung des Roboters durch Interaktionen, dessen Bewegungsrichtungen nicht direkt auf die des Roboters übertragbar sind, zu kontrollieren. Ähnliche Beobachtungen wurden auch von Tijmsa et al. [125] bei der Steuerung eines Roboters durch einen Joystick berichtet. Da strikte Trennung der Freiheitsgrade zusätzlich eine robuste Erkennung der jeweiligen Interaktionen ermöglicht und kleine Fehlbewegungen nicht direkt in Änderungen der Rotations- bzw. Translationsrichtung resultieren [45], wurde sich für die Implementierung von getrennten Freiheitsgraden entschieden.

Da die Bewegungsfreiheit des Kopfes limitiert ist, wird ein Remapping durchgeführt. Somit werden Positions- bzw. Rotationsänderungen des Kopfes nicht direkt auf die Position und Orientierung des Roboters übertragen, sondern auf Geschwindigkeiten abgebildet. Eine Auslenkung des Kopfes aus der Ruheposition kann somit auf eine Robotergeschwindigkeit entlang einer bestimmten Richtung übersetzt werden.

Dazu werden die in Abbildung 4.1 dargestellten Kopfbewegungen verwendet. Für die Interaktion dienen hauptsächlich die Rotationen des Kopfes. Bei der Translation besteht jedoch die Schwierigkeit, dass sowohl die Bewegung in die Tiefe (entlang der x-Achse) als auch die vertikale Bewegung (z-Achse) sich am besten durch eine Pitch-Rotation des Kopfes (Nicken) abbilden lassen. Daher wurde für die Translation in die Tiefe eine zusätzliche Interaktion eingeführt, indem die Bewegung des Kopfes entlang der x-Achse verwendet wird, die durch das nach vorne und hinten Lehnen des Oberkörpers manipuliert werden kann. Somit kann die Translation in x-Richtung durch das nach vorne/hinten Lehnen, in y-Richtung durch das Neigen des Kopfes (Roll) und in z-Richtung durch das Nicken (Pitch) manipuliert werden. Für die Rotationen resultiert das Neigen des Kopfes (Roll) in Rotation um die x-Achse, Nicken (Pitch) in Rotationen um die y-Achse und Kopfschütteln (Yaw) in Rotationen um die z-Achse.

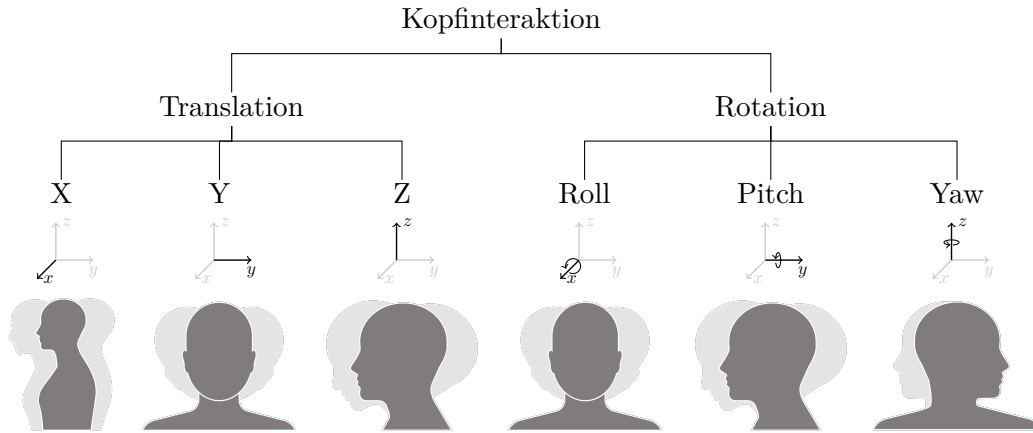


Abbildung 4.1: Überblick über die Kopfinteraktionen für die Navigation des Roboters

Um die Abbildung von Änderungen der Rotation des Kopfes δ auf Geschwindigkeiten des Roboters v zu ermöglichen, wird eine Transferfunktion benötigt. Dazu kann die in Formel 4.1 und Abbildung 4.2 dargestellte Gompertzfunktion verwendet werden [101, 105]. Dabei beschreibt der Parameter a die Maximalgeschwindigkeit des Roboters, b die Verschiebung entlang des Kopfwinkels und c die Wachstumsrate.

$$v_n(\delta) = \begin{cases} a \cdot e^{b \cdot e^{c\delta_n}} & , \delta_n \geq 0 \\ -a \cdot e^{b \cdot e^{-c\delta_n}} & , \text{sonst} \end{cases} \quad (4.1)$$

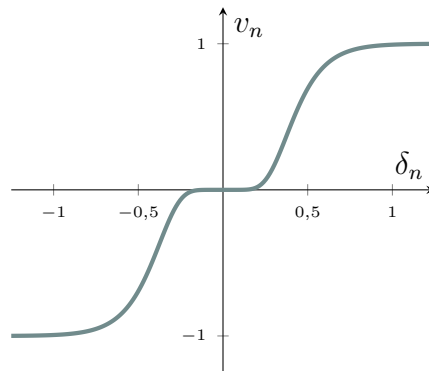


Abbildung 4.2: Gompertzfunktion als Transferfunktion zur Abbildung der Änderung der normierten Kopforientierung δ_n auf die normierte Geschwindigkeit v_n der Bewegung des Roboters

Die Gompertzfunktion bietet den Vorteil, einen *Toten-Bereich* in der Nähe des Nullpunktes zu haben, sodass durch kleine Bewegungen des Kopfes keine Bewegung des Roboters ausgelöst werden. Ab einer bestimmten maximalen Auslenkung des Kopfes bleibt die Geschwindigkeit der Roboterbewegung konstant, sodass die gewünschte Maximalgeschwindigkeit nicht überschritten wird.

Die normierte Kopfrotation δ_n lässt sich dabei durch $\delta_n = \frac{\delta}{\delta_{max}}$ mit der maximalen Rotation

δ_{max} bestimmen. Dadurch wird die Anpassung der Geschwindigkeiten auf die unterschiedlichen Kopfrotationen möglich, sodass für jede Rotationsrichtung ein eigener Maximalwert festgelegt werden kann. Für die Translation in x-Richtung wird nicht die Änderung der Kopfausrichtung δ verwendet, sondern diese durch die Änderung der Translation des Kopfes substituiert.

Handinteraktion

Handinteraktionen können einhändig oder beidhändig realisiert werden. Die Interaktion soll den Workflow während der Intervention möglichst wenig stören. Da die dominante Hand gegebenenfalls für Aufgaben zur Durchführung der Intervention benötigt wird, wird eine einhändige Interaktion mit der nicht-dominanten Hand umgesetzt.

In Virtual Reality werden für die Manipulation virtueller Objekte Freihandgesten oft exakt übertragen. Durch die direkte Abbildung werden physische Interaktionen nachgeahmt, wodurch eine natürliche Interaktion ermöglicht wird [45]. Zusätzlich stellt die direkte Positionsübertragung den Kontrast zu der bei der Kopfinteraktion verwendeten Geschwindigkeitsabbildung dar, sodass durch die Realisierung der Positionsabbildung das in Abschnitt 2.4.1 vorgestellte Abbildungskontinuum abgedeckt werden kann.

Durch die skalierte Umsetzung der Eingabe könnte ein größerer Bereich abgedeckt werden. Jedoch kann dadurch die Übertragbarkeit und somit die Nachvollziehbarkeit der Auswirkung der Interaktion auf die resultierende Bewegung reduzieren werden. Insbesondere die skalierte Rotation wird oft als verwirrend empfunden [123]. Aus diesen Gründen wurde sich für die Implementierung einer unskalierten Abbildung der Handtranslation und -rotation auf die Position und Ausrichtung des Roboters entschieden.

Die Steuerung des Roboters soll somit wie in Abbildung 4.3 dargestellt, über die Bewegung der Fingerspitze und die Rotation der Handfläche ermöglicht werden, wobei diese nicht auf diskrete Achsen reduziert wird, sondern in alle Raumrichtungen möglich ist.

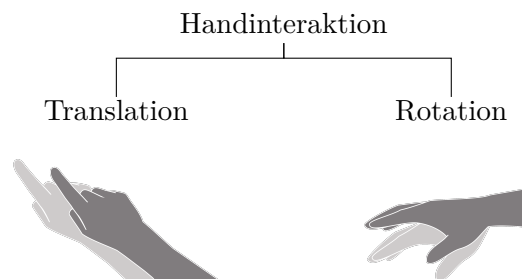


Abbildung 4.3: Überblick über die Handinteraktion für die Navigation des Roboters

4.3 Konzeption Interaktionsaufgabe

Für die Entwicklung der berührungslosen Steuerung des Ultraschallroboters müssen die verschiedenen vorgestellten Interaktionskonzepte evaluiert werden. Dafür soll zunächst im Rahmen einer initialen Nutzerstudie die am besten geeignete Kombination an Interaktionsmodalitäten für die Aktivierungs- und Navigationsinteraktion ermittelt werden. Unter Berücksichtigung aus dieser Studie hervorgehender Erkenntnisse soll ein finales Interaktionskonzept entwickelt und qualitativ evaluiert werden. Für diese beiden Studien werden unterschiedliche Interaktionsaufgaben benötigt, die die Anwendung und Evaluierung der Interaktionskonzepte ermöglichen.

4.3.1 Interaktionsaufgabe initiale Studie

Zur Identifikation der geeigneten Aktivierungs- und Navigationsinteraktions-Modalitäten wird eine quantitative Nutzerstudie durchgeführt. Um dabei sowohl die Interaktionsaufgabe als auch das Multitasking für während der Intervention parallel ablaufende kognitive Aufgaben abbilden zu können, werden eine Primär- und eine Sekundäraufgabe entwickelt.

Primäraufgabe

Die Primäraufgabe soll dazu dienen, in einer abstrahierten Aufgabe die verschiedenen Modalitäten zur Aktivierungs- und Navigationsinteraktion zu testen. Wie in Abschnitt 3.1.2 beschreiben, wird die Navigationsinteraktion benötigt, um vorzustrahlen. Dazu wird die am Roboter befestigte Ultraschallsonde manuell zu einer bestimmten Position gefahren, um beispielsweise die Zielstruktur zu visualisieren. Während der Navigationsinteraktion wird die diskrete Interaktion benötigt, um die unterschiedlichen Navigationsmodi zu aktivieren. Somit ist die isolierte Aufgabe des Vorstrahlens für die Evaluierung von sowohl der diskreten als auch der Navigationsinteraktion geeignet.

Um die Interaktion für die Studie ohne medizinisches Vorwissen zu ermöglichen, soll auf die Auswertung von Ultraschallbildern zur Navigation des Roboters verzichtet werden. Somit muss das zu erreichende Ziel, zu dem Ultraschallsonde bewegt werden soll, anders definiert und visualisiert werden. Damit für die benötigte Bewegung sowohl die Interaktion zur Translation als auch die zur Rotation verwendet werden muss, muss das Ziel durch eine Position und eine Orientierung definiert werden.

Da für die Interaktionen bereits eine HoloLens verwendet wird, die die Möglichkeit bereitstellt, semitransparente Inhalte in Augmented Reality darzustellen, kann dies zur Festlegung des Ziels genutzt werden. Um die angestrebte Position und Ausrichtung der Ultraschallsonde zu visualisieren, kann dabei ein 3D-Modell der Sonde dargestellt werden.

Somit ist das Ziel der Primäraufgabe die Bewegung der am Roboter befestigten Ultraschallsonde, sodass die reale Ultraschallsonde durch Anpassung der Position und Ausrichtung bestmöglich die virtuelle Sonde überlagert.

Sekundäraufgabe

Da die Primäraufgabe nicht das Multitasking und somit nicht die kognitiven Anforderungen während der Intervention abbilden kann, wird eine Sekundäraufgabe eingeführt. Diese wird parallel zur Primäraufgabe durchgeführt und kann als Indikator für die kognitive Anstrengung dienen. Unter der Annahme, dass die kognitive Kapazität von Nutzenden begrenzt ist, steht bei hoher Beanspruchung durch die Primäraufgabe weniger kognitive Kapazität für die Sekundäraufgabe zur Verfügung. Die Performance der Sekundäraufgabe kann somit Aufschluss über die kognitive Anstrengung der Primäraufgabe geben [126]. Dadurch kann der Einfluss der Primäraufgabe auf die Sekundäraufgabe und somit auch der mögliche Einfluss auf den Workflow, bestimmt werden. Unter Verwendung einer Kontrollgruppe, die die Sekundäraufgabe nicht durchführt, kann zusätzlich auch der Einfluss der Sekundäraufgabe auf die Primäraufgabe untersucht werden.

Zusätzlich zur fehlenden Abbildung des kognitiven Aufwands wird durch die Primäraufgabe auch eine motorische Aufgabe vernachlässigt. Während der Intervention würde auf einem Monitor das Ultraschallbild angezeigt und zur Navigation verwendet werden. Zur Verifikation der Sondenposition würde somit in regelmäßigen Abständen auf den Monitor geblickt werden. Um den Einfluss von dieser Bewegung auf die Interaktionen bzw. der Interaktion auf die Bewegung evaluieren zu können, sollte die Sekundäraufgabe diesen motorischen Aufwand abbilden.

Ein sekundärer Task kann beispielsweise die Reaktion auf visuelle, taktile oder auditive Reize durch Betätigung eines Knopfes sein [127]. Um den motorischen Aufwand durch die Sekundäraufgabe zu modellieren, sollten die Reize erst nach einer erforderlichen Bewegung für Nutzende zugänglich sein. Daher werden visuelle Reize verwendet, die auf einem Monitor dargestellt werden. Der Blick auf den Monitor würde während der Intervention nicht zufällig stattfinden, sondern hauptsächlich wenn signifikante Veränderungen des Ultraschallbilds zu erwarten sind. Daher sollte die Sekundäraufgabe nicht die kontinuierliche Überwachung des Monitors, sondern die regelmäßige Kontrolle in vorhersehbaren Abständen erfordern.

Um eine zyklische Überwachung der Sekundäraufgabe zu ermöglichen, wird ein Regler verwendet, der sich in einer konstanten Geschwindigkeit entlang eines Balkens verschiebt. Auf diesem Balken befinden sich an zufälligen Positionen platzierte Zielbereiche (Targets). Die Sekundäraufgabe besteht darin, einen Knopf zu betätigen, wenn sich der Regler innerhalb eines Targets befindet. Durch die konstante Geschwindigkeit des Reglers kann die Entfernung des Reglers zum Target und somit die nötigen Zeitabstände zur Kontrolle grob eingeordnet werden. Dadurch, dass der Kopf somit in regelmäßigen Abständen zur

Überprüfung der Reglerposition zum Bildschirm gedreht und die mögliche Position des Reglers innerhalb des Targets rechtzeitig erkannt werden muss, können sowohl die kognitiven als auch die motorischen Herausforderungen durch die Sekundäraufgabe abgebildet werden.

4.3.2 Interaktionsaufgabe finale Studie

In der finalen Studie soll das Gesamtsystem evaluiert werden. Somit soll das finale Interaktionskonzept nicht isoliert, sondern unter Berücksichtigung der Ultraschallbildgebung betrachtet werden. Dabei sollen die Funktionen der Roboternavigation, die Speicherung von Punkten und das Anfahren bekannter bzw. gespeicherter Punkte durch die Aufgabe gefordert werden.

Die Roboternavigation wird benötigt, um bestimmte Zielstrukturen zu suchen und zu visualisieren. Damit dazu nicht nur die Translation der Sonde benötigt wird, werden störende Strukturen benötigt, die den direkten Blick auf die Zielstruktur verdecken, sodass die Struktur nur durch zusätzliche Rotation visualisiert werden kann.

Um während der Aufgabe auch das Anfahren gespeicherter Punkte zu verwenden, muss durch die Aufgabe die Rückkehr zu bereits visualisierten Strukturen erfordert werden. Dies könnte der Fall sein, wenn die Aufgabe aus dem Anfahren verschiedener Zielstrukturen in einer vordefinierten Reihenfolge besteht. Wenn somit während der Suche nach der aktuell verlangten Struktur eine andere Struktur gefunden wird, kann diese gespeichert werden. Durch die Speicherung kann somit die Suche nach dieser Struktur zum entsprechenden Zeitpunkt vereinfacht werden, indem lediglich der zuvor gespeicherte Punkt angefahren wird.

5 Umsetzung

Nachfolgend wird die Umsetzung der im vorhergehenden Kapitel erarbeiteten Konzepte beschrieben. Dabei wird auf die Umsetzung eines initialen und eines verbesserten finalen Prototyps eingegangen. Das Konzept des finalen Prototyps basiert dabei auf Erkenntnissen der in 6.1 vorgestellten Nutzerstudie. Zusätzlich zu der Umsetzung der Interaktionsprototypen wird auch die Umsetzung der Interaktionsaufgaben für die in Kapitel 6 vorgestellten Nutzerstudien beschrieben.

5.1 Rahmenbedingung

5.1.1 Entwicklungsumgebung

Der Prototyp für die Mensch-Roboter-Interaktion wird mithilfe einer Microsoft HoloLens 2 und einem KUKA LBR iiwa 14 R820 umgesetzt. Die HoloLens Applikation wird über eine Universal Windows Plattform App für Windows 10 realisiert. Aufgrund des vorhandenen HoloLens Supports wird die Anwendung mithilfe der Entwicklungsumgebung Unity 2020.3.17f [128] umgesetzt. Dabei erfolgt die Implementierung unter Verwendung von C# und Visual Studio 2019 [129]. Zur Unterstützung der Entwicklung für die HoloLens wird das Microsoft Mixed Reality Toolkit 2.5.1 (MRTK) [130] verwendet. Für die Erkennung der optischen Marker wird die Vuforia SDK 8.6.7 [131] genutzt.

Die Robotersteuerung wird mithilfe von KUKA Sunrise.OS 1.16 [132] und der KUKA Sunrise Workbench 1.16 in Java implementiert. Für die Kommunikation zwischen HoloLens und Roboter wird das Netzwerkprotokoll UDP (User Data Protocol) verwendet. Da die Kommunikation mit der Robotersteuerungseinheit nur über LAN möglich ist, werden die Daten zunächst an einen Desktop-PC gesendet, der über ein WLAN Netzwerk mit der HoloLens verbunden ist und die Daten entsprechend weiterleitet. Diese Schnittstelle auf dem Desktop-PC wird ebenfalls in Unity umgesetzt.

5.1.2 Robotersteuerung

Für die Steuerung des Roboters wird das DirectServo Interface des KUKA Sunrise.Servoing Pakets verwendet, das verschiedene Bewegungsklassen für die Realisierung nichtdetermi-

nistischer Echtzeitanwendungen zur Verfügung stellt. Das DirectServo Interface ermöglicht kontinuierliche und hochfrequente Vorgaben achsenspezifischer und kartesischer Zielpunkten. Diese werden durch eine geschwindigkeitsbegrenzte PTP-Bewegung entlang des schnellstmöglichen Bewegungspfades angefahren. Dadurch, dass die Servobewegungen durch asynchrone Bewegungen umgesetzt werden, werden Zielpunkte während des Anfahrens verworfen, wenn ein neuer Zielpunkt gesetzt wird. Für visuelles Feedback am Roboter können zwei separat ansteuerbare LED-Ringe verwendet werden.

Bewegungsvorgaben in Form kartesischen Zielpunkte können durch die Angabe der Zielposition im Roboterkoordinatensystem oder einem fest damit verbundenen Koordinatensystem erfolgen. Aus Performancegründen wurde sich für die Angabe der Koordinaten im Roboterkoordinatensystem entschieden. Da die aktuelle Position und Ausrichtung des Roboters durch die Abfrage der letzten Zielposition ermittelt werden kann, kann die absolute Zielvorgabe zusätzlich für relative Bewegungsvorgaben genutzt werden.

Die Bewegungsbefehle werden als Byte codiert per UDP-Verbindung an die Robotersteuerung gesendet, dabei beschreibt der erste Teil der Nachricht die Funktion (beispielsweise absolute oder relative Bewegung oder die Anfrage der aktuellen Roboterposition) und der zweite Teil gegebenenfalls Zusatzinformationen wie die Koordinaten der angestrebten relativen oder absoluten Zielposition. Bewegungsbefehle beziehen sich dabei immer auf den Tool Center Point (TCP), der fest mit dem Roboterflansch verbunden ist, wobei die relative Position des TCP zum Flansch manuell eingemessen werden kann.

Für die Arbeitssicherheit wird ein Schutzraum definiert, aus dem der Endeffektor des Roboters nicht hinaus bewegt werden kann. Um Nutzenden Feedback zu liefern, wann die Grenzen des Schutzraumes erreicht sind, wird der blaue LED-Ring verwendet. Wenn ein Ziel außerhalb des Schutzraumes gesetzt wird, wird dieses Ziel verworfen und es erfolgt keine Roboterbewegung.

Zusätzlich wird zur Kollisionsvermeidung die Kraft in z-Richtung, also die senkrechte Krafteinwirkung auf den Roboter begrenzt. Durch diese Einschränkung soll die Ausübung von starkem Druck mit der Ultraschallsonde auf Oberflächen vermieden werden. Mithilfe eines Condition Observers wird bei zu starker Krafteinwirkung die Roboterbewegung gestoppt. Zur Visualisierung wird der rote LED-Ring des Roboters aktiviert. Um die Bewegung fortsetzen zu können, muss im initialen Prototypen der Roboter angetippt werden, was durch die Krafteinwirkung in y-Richtung erkannt wird. Dadurch wird der Roboter um 1 cm in z-Richtung bewegt, wodurch die Krafteinwirkung reduziert und die Bewegung fortgesetzt werden kann. Um im finalen Prototypen eine flüssige Interaktion zu ermöglichen, wird hier nach Krafteinwirkung die Roboterbewegung gestoppt und der Roboter direkt um 1 cm nach oben gefahren. Als Signal zur Fortsetzung der Interaktion wird lediglich die Deaktivierung des aktiven Interaktionsmodus verwendet. Somit kann die Interaktion ohne zusätzliche Befehle wie das Berühren des Roboters fortgesetzt werden.

5.1.3 Registrierung

Um die HoloLens mit dem Roboter zu registrieren, wird ein optischer Marker verwendet. Die Koordinatensysteme (KOS), zwischen denen für die Registrierung die Transformationen bestimmt werden müssen, sind in Abbildung 5.1 dargestellt. Dabei wird das WeltKOS von Unity beim Programmstart initiiert, die Position und Ausrichtung der HoloLens und somit auch des KameraKOS im WeltKOS sind dabei bekannt. Um die fehlende Transformation zwischen RoboterKOS und KameraKOS zu bestimmen, wird ein MarkerKOS eingeführt, das durch einen optischen Marker definiert ist.

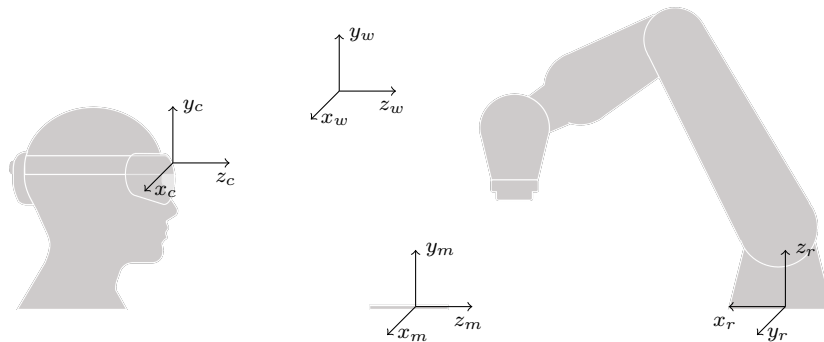


Abbildung 5.1: Übersicht über die für die Registrierung relevanten Koordinatensysteme. Dabei beschreiben die Indizes c das KameraKOS, w das WeltKOS, m das MarkerKOS und r das RoboterKOS

Zur Bestimmung des Verhältnisses zwischen KameraKOS und MarkerKOS wird die Vuforia Augmented Reality SDK verwendet. Diese ermöglicht mithilfe von Computer Vision das Echtzeittracking von planaren Bildern. Aus Performancegründen wird das MarkerKOS bei Start der Applikation lokalisiert und das Markertracking anschließend deaktiviert. Um die Robustheit der Lokalisierung zu erhöhen, wird die Position und Ausrichtung des Markers drei Sekunden lang aufgenommen und anschließend die Werte gemittelt. Dadurch, dass das Verhältnis zwischen KameraKOS und WeltKOS durch das interne Inside-Out-Tracking der HoloLens bekannt ist, ist auch die Position und Ausrichtung zwischen MarkerKOS und WeltKOS bekannt. Somit kann ohne aktives Markertracking weiterhin das Verhältnis zwischen KameraKOS und MarkerKOS bestimmt werden. Zur Verifikation wird das MarkerKOS in Augmented Reality dargestellt, sodass Ungenauigkeiten des internen HoloLens Trackings erkannt und gegebenenfalls korrigiert werden können.

Da sowohl der Ursprung des MarkerKOS, der sich im Mittelpunkt befindet, als auch die Ausrichtung des MarkerKOS auf dem Marker bekannt ist, kann das MarkerKOS vom Roboter eingelesen werden. Die Software des Roboters ermöglicht dazu das Einlesen von Frames, indem mit einem zuvor kalibrierten Tool der Ursprung eines Koordinatensystems und ein auf der x- sowie ein auf der y-Achse liegender Punkt angefahren wird. Dadurch ist die Position des Markers im RoboterKOS und daraus resultierend auch die Position

des Roboters im MarkerKOS bekannt. Somit kann wie folgt die Orientierung ${}^R_W R$ als auch die Position ${}^R_W P$ des RoboterKOS im Weltkoordinatensystem bestimmt werden. Die Rotationen R werden dabei für die Berechnung als Quaternionen q dargestellt.

$${}^R_W R = {}^M_W q \cdot {}^M_R q^{-1} \quad (5.1)$$

$${}^R_W P = {}^M_W P - {}^R_W R \cdot {}^M_R P \quad (5.2)$$

Dabei muss jedoch beachtet werden, dass das RoboterKOS als rechtshändiges KOS und die Koordinatensysteme in Unity als linkshändige KOS definiert sind. Um die Koordinaten entsprechend zu transformieren, können diese 90° um die x-Achse rotiert und anschließend die Werte der y- und z-Koordinaten vertauscht werden.

Zusätzlich ist die Ausrichtung des MarkerKOS im RoboterKOS ${}^M_R R$ in Euler-Winkeln gegeben. Die Drehfolge des Roboters ist dabei die intrinsische Reihenfolge Z-Y'-X". In Unity wird hingegen die extrinsische Drehfolge Z-X-Y verwendet, sodass die Winkel nicht direkt übertragen werden können. Die benötigten Winkel können jedoch über die Multiplikation von Quaternionen, die jeweils um die einzelnen Achsen rotieren, bestimmt werden. Dabei muss erneut beachtet werden, dass durch das rechts- bzw. linksdrehende KOS die Vorzeichen der Winkel angepasst werden müssen.

$${}^M_R R = {}^M_R q_Z \cdot {}^M_R q_{-Y} \cdot {}^M_R q_{-X} \quad (5.3)$$

Somit kann das MarkerKOS als gemeinsames KOS verwendet werden und die benötigten Parameter des Kamera- und des RoboterKOS ineinander umgewandelt werden.

Neben der Umrechnung von Winkeln aus dem Roboterkoordinatensystem in das Kamerakoordinatensystem kann es auch notwendig sein, Orientierungen, die in Unity als Quaternionen $q = [q_0, q_1, q_2, q_3]^T$ gegeben sind, in Eulerwinkel für die Robotersteuerung, also in der intrinsischen Reihenfolge Z-Y'-X" umzuwandeln. Dabei entspricht φ Roll, also der Rotation um die x-Achse, θ Pitch, also der Rotation um die y-Achse und ψ Yaw, also der Rotation um die z-Achse (siehe Gleichung 5.4) [133].

$$\begin{pmatrix} \varphi \\ \theta \\ \psi \end{pmatrix} = \begin{pmatrix} \arctan \frac{2(q_0 q_1 + q_2 q_3)}{1 - 2(q_1^2 + q_2^2)} \\ \arcsin(2(q_0 q_2 - q_3 q_1)) \\ \arctan \frac{2(q_0 q_3 + q_1 q_2)}{1 - 2(q_2^2 + q_3^2)} \end{pmatrix} \quad (5.4)$$

5.2 Initialer Prototyp

5.2.1 Diskrete Interaktion

Die diskrete Interaktion wird unabhängig von der Interaktionsmodalität über Buttons umgesetzt. Diese dienen zum einen dazu, verfügbare Befehle zu visualisieren und zum anderen um visuelles Feedback über den Systemzustand zu geben. Wenn eine Eingabe zur Aktivierung eines Modus erkannt wird, wird der entsprechende Button hervorgehoben. Zusätzlich ertönt ein Signalton über die Lautsprecher der HoloLens. Bei Deaktivierung der Modi wird der entsprechende Button wieder in der ursprünglichen Farbe dargestellt und es wird erneut ein Signalton abgespielt. Zusätzlich zu dem visuellen und auditiven Feedback auf der HoloLens, wird auch direkt am Roboter durch den LED-Ring der Systemzustand visualisiert. Solange wie ein Navigationsmodus aktiv ist, leuchtet der LED-Ring grün. Dadurch kann überprüft werden, dass die Aktivierung des entsprechenden Modus korrekt an den Roboter weitergeleitet wurde und der Roboter somit in Bewegungsbereitschaft ist. Zusätzlich kann dadurch, dass die Darstellung des aktiven Modus und die der Kollisionserkennung auf dem gleichen LED-Ring erfolgt, erkannt werden, ob während einer Kollision die Navigationsinteraktion noch aktiv ist. Wenn der LED-Ring somit orange leuchtet, muss die Navigationsinteraktion beendet werden, um die Kollisionskorrektur vornehmen zu können.

Die diskrete Interaktion wird verwendet, um den Rotations- und den Translationsmodus zu aktivieren, sowie um die aktuelle Interaktionsaufgabe zu beenden. Da das Beenden der Interaktionsaufgabe keine direkte Visualisierung benötigt, wird dieser Befehl bei der Sprachinteraktion nicht als Button dargestellt. Bei der Blickinteraktion wird der Button, um die freie Sicht möglichst wenig zu beeinträchtigen, am oberen rechten Rand des Sichtfeldes platziert. Die Buttons für die Navigationsmodi werden für beide Interaktionsmodalitäten zentral am oberen Rand des Sichtfeldes platziert.

Für die Blickinteraktion muss der Blick auf die Buttons gerichtet werden. Diese verändern sich während der Verweilzeit farblich, um Feedback über den erkannten Fokus zu geben. Die Blickrichtung wird nicht über einen Cursor visualisiert, damit es im normalen Sichtfeld nicht zu Störungen kommt. Um das Midas-Touch-Problem zu vermeiden, wird der Blick auf den Button und somit der Beginn der Verweilzeit erst nach 0.2 Sekunden erkannt. Die Verweilzeit beträgt 0.75 Sekunden. Da durch das generelle Zittern der Blickrichtung und durch Ablenkungen der Fokus der Blickrichtung sich kurz vom Button entfernen kann, wird eine Recovery Zeit von 0.2 Sekunden eingeführt, in der der Fokus auf den Button wieder aufgenommen und die Verweilzeit fortgesetzt werden kann.

Bei der Sprachinteraktion werden einzelne Kommandos verwendet, um die Modi zu aktivieren. Diese werden auf den Buttons dargestellt. Die verwendeten Sprachbefehle

sind: *Translation*, *Rotation*, *Stopp* und *Beenden*. Die Befehle *Translation* und *Rotation* dienen zur Aktivierung der entsprechenden Navigationsmodi. Für Parallelität zwischen den diskreten Interaktionen kann wie bei der Blickinteraktion, bei der ein erneuter Blick auf den Button deaktiviert, auch der Befehl für den aktiven Modus erneut gesagt werden, um diesen zu deaktivieren. Zusätzlich besteht die Möglichkeit, den Modus zu deaktivieren, indem *Stopp* gesagt wird. Um die Interaktionsaufgabe zu beenden, kann das Kommando *Beenden* gesagt werden.

5.2.2 Navigationsinteraktion

Durch die Navigationsinteraktion soll eine relative Bewegung des Roboters, also die Translation bzw. Rotation aus der aktuellen Position in eine bestimmte Richtung/um eine bestimmte Achse realisiert werden. Die Robotersteuerung ermöglicht jedoch nur die Vorgabe einer absoluten Zielposition, sodass relative Angaben umgerechnet werden müssen. Für die Translation kann die relative Zielposition durch einen Richtungsvektor angegeben werden, der an den Roboter gesendet und anschließend zur Positionsvorgabe auf die aktuelle Roboterposition addiert werden kann. Dabei wird der Richtungsvektor basierend auf der zuvor beschriebenen Registrierung (siehe Abschnitt 5.1.3) in Roboterkoordinaten umgerechnet.

Im Gegensatz dazu kann die Zielorientierung nicht einfach durch Addition einer relativen Rotation bestimmt werden, da die Linearkombination von Eulerwinkeln nicht möglich ist. Der Grund dafür ist, dass die Eulerwinkel nur Repräsentanten der Rotationsmatrizen und somit von der Drehreihenfolge abhängig sind. Für die Anwendung einer relativen Rotation wird daher eine Umrechnung in die Matrizendarstellung oder in Quaternionen benötigt. Da die Unity Engine eine Schnittstelle zur Umrechnung von Eulerwinkeln in Quaternionen und für Rechenoperationen mit Quaternionen bietet, wird die Berechnung der Zielrotation in Unity umgesetzt. Dazu muss zu Beginn der Rotation die aktuelle Ausrichtung des Endeffektors R_n^R an die HoloLens übergeben werden. Nachdem basierend auf der Interaktion die angestrebte relative Rotation $R_{rel} = (q_{n-1})^{-1} \cdot q_n$ bestimmt wird, kann unter der Verwendung von Quaternionen die vorherige Roboterausrichtung entsprechend rotiert und somit die Zielausrichtung $R_{n+1}^R = q_{rel} \cdot q_n$ bestimmt werden. Die daraus resultierende Zielrotation kann anschließend entsprechend Formel 5.4 in Eulerwinkel umgerechnet und an den Roboter gesendet werden.

Durch die Festlegung der Zielposition wird implizit die Geschwindigkeit der Roboterbewegung gesetzt. Dabei wird in der Robotersteuerung eine minimale Ausführungszeit für Pfade und eine Maximalgeschwindigkeit und -beschleunigung festgelegt. Dadurch werden Zielvorgaben mit geringer Distanz zur aktuellen Position in der vorgegebenen minimalen Ausführungszeit und weiter entfernte Zielvorgaben unter Berücksichtigung der

Maximalgeschwindigkeit zurückgelegt. Bei Zielvorgaben, die gegeben werden, bevor die vorherige Zielposition erreicht wurde, wird die vorherige Zielvorgabe verworfen und die neue Zielposition gesetzt. Wenn die neue Zielvorgabe nicht erfolgt, bevor der Roboter die Zielposition erreicht hat, wird die Bewegung gestoppt, bis die neue Zielvorgabe erfolgt. Dadurch kann der sogenannte Sekundenzeigereffekt auftreten, der das ruckartige Fahren des Roboters durch das Stoppen zwischen Erreichen des Ziels und neuer Zielvorgabe beschreibt. Daher ist die Abstimmung zwischen Zielvorgabe und Robotersteuerung essentiell. Ein weiterer zu beachtender Faktor bei der Zielvorgabe ergibt sich durch die Angabe der Zielposition als relative Zielvorgabe in Bezug zur aktuellen Position. Um die aktuelle Roboterposition $\vec{T}R_n$ abzufragen, ermöglicht die Robotersteuerung den Zugriff auf die letzte Zielposition $\vec{T}R_{z-1}$, die bei abgeschlossener Bewegung $\vec{T}R_n$ entspricht. Wenn jedoch zum Zeitpunkt der neuen relativen Zielvorgabe $\Delta\vec{T}R$ das Ziel $\vec{T}R_{z-1}$ noch nicht erreicht ist, entsteht ein Offset, da $\vec{T}R_z \neq \vec{T}R_n + \Delta\vec{T}R$ sondern $\vec{T}R_z = \vec{T}R_n + \Delta\vec{T}R + (\vec{T}R_{z-1} - \vec{T}R_n)$. Dieser Offset kann bei längeren Bewegungen immer größer werden, was dazu führt, dass der Roboter sich weiter als von Nutzenden erwartet bewegt und die Bewegung sogar fortgesetzt werden könnte, obwohl keine neuen Zielvorgaben erfolgen.

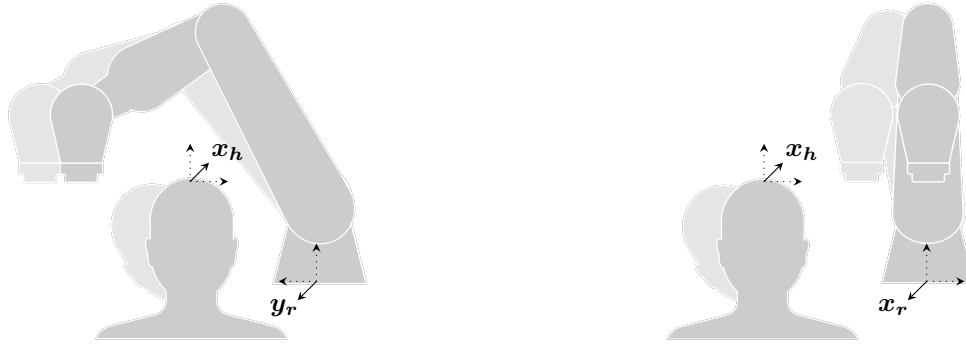
Um sowohl den Sekundenzeigereffekt als auch das nachträgliche Weiterbewegen des Roboters zu vermeiden, müssen somit die Länge und Frequenzen der vorgegebenen Bewegungsabschnitte auf die durch die Robotersteuerung vorgegebene Robotergeschwindigkeit, -beschleunigung und Synchronisationszeit angepasst werden. Nur so kann eine flüssige und kontrollierte Roboterbewegung ermöglicht werden.

Kopfinteraktion

In der Kopfinteraktion werden Änderungen der Kopfposition und -ausrichtung im Bezug zur Ruhelage verwendet, um kontinuierliche Translation und Rotation des Roboters zu erzielen. Die Ruhelage wird dabei durch die Kopfposition und -orientierung bei der Aktivierung des Navigationsmodus festgelegt.

Da die einzelnen Freiheitsgrade getrennt manipuliert werden, müssen die Rotationsrichtungen des Kopfes den diskreten Rotationsrichtungen zugeordnet und in die korrekten Bewegungsbefehle umgewandelt werden. Um die Rotationsrichtungen des Kopfes zu erkennen, wird die relative Rotation im Bezug zur Ruhelage bestimmt und in die Angle-Axis Darstellung konvertiert. Der daraus resultierende Winkel, der die Auslenkung der Kopfrotaion beschreibt, wird der Transferfunktion übergeben, um die Robotergeschwindigkeit und somit die benötigte Distanz der Zielposition zu bestimmen. Aus der Rotationsachse kann bestimmt werden, ob es sich um das Nicken, Drehen oder Neigen des Kopfes handelt. Um eine positionsabhängige Manipulation der Roboterfreiheitsgrade zu ermöglichen, wird zur Ermittlung der Achse der resultierenden Roboterbewegung, die am ähnlichsten zur Rotationsachse des Kopfes ausgerichtete Achse des RoboterKOS bestimmt (siehe Abbildung 5.2). Diese entspricht im Rotationsmodus der resultierenden Rotationsachse des

Roboters. Im Modus für die vertikale und horizontale Translation entspricht die um 90° in der vertikalen Ebene rotierte Achse der Translationsrichtung. Somit werden abhängig von der Ausrichtung der Nutzenden zum Roboter verschiedene Manipulationsrichtungen gewählt, sodass unabhängig von der Position der Nutzenden die gleiche Rotation des Kopfes in die gleiche Manipulation des Roboters aus der Sicht der Nutzenden resultiert.



- (a) Bei seitlicher Positionierung zum Roboter wird Roll-Rotation des Kopfes auf die Translation in x-Richtung abgebildet (b) Bei frontaler Positionierung zum Roboter wird Roll-Rotation des Kopfes auf die Translation in y-Richtung abgebildet

Abbildung 5.2: Beispielhafte Darstellung der positionsabhängigen Auswahl der Achsen für die Roboterbewegung

Im Folgenden wird schematisch die Logik der Kopffinteraktionen für die Manipulation der Freiheitsgrade beschrieben. Dazu wird zur Vereinfachung die Transformation in Roboterkoordinaten vernachlässigt, stattdessen werden die resultierenden Roboterbewegung in denselben Koordinatensystem wie die Kopffrotationen und -translationen dargestellt.

Das Vorgehen für die vertikale und horizontale Translation wird in Abbildung 5.3a dargestellt. Für die vertikale Translation wird die Änderung des Pitch-Winkels (Nicken) und für die horizontale Translation die Änderung des Roll-Winkels (Neigen) verwendet. Diese Rotationsachsen werden unter a zusammengefasst. Wenn die Auslenkung der aktuellen Kopffrotation $R_{n,a}^K$ in Bezug zur Ruhelage $R_{0,a}^K$ größer als ein Schwellenwert S ist, bewegt sich der Roboter in die entsprechende Richtung $\perp \vec{a}$, um die durch die Transferfunktion bestimmte Distanz b .

Da die Translation des Roboters in die Tiefe, über das nach vorne und hinten Lehnen manipuliert wird, wird hierfür die Translation des Kopfes in x-Richtung verwendet (siehe Abbildung 5.3b). Wenn die Änderung der x-Position zwischen der aktuellen Position $\vec{T}_{n,x}^K$ und der Ruheposition $\vec{T}_{0,x}^K$ größer als ein Schwellenwert S ist, wird der Roboter entsprechend der durch die Transferfunktion bestimmten Distanz b in der Tiefe bewegt.

Die Berechnung der Rotation erfolgt analog zur vertikalen und horizontalen Translation, mit dem Unterschied, dass die Änderung der Ausrichtung des Kopfes $R_{n,a}^K - R_{0,a}^K$ mit $a \in \{x, y, z\}$ auf die Rotation um die Achse a um den durch die Transferfunktion bestimmten Winkel b abgebildet wird (siehe Abbildung 2.4).

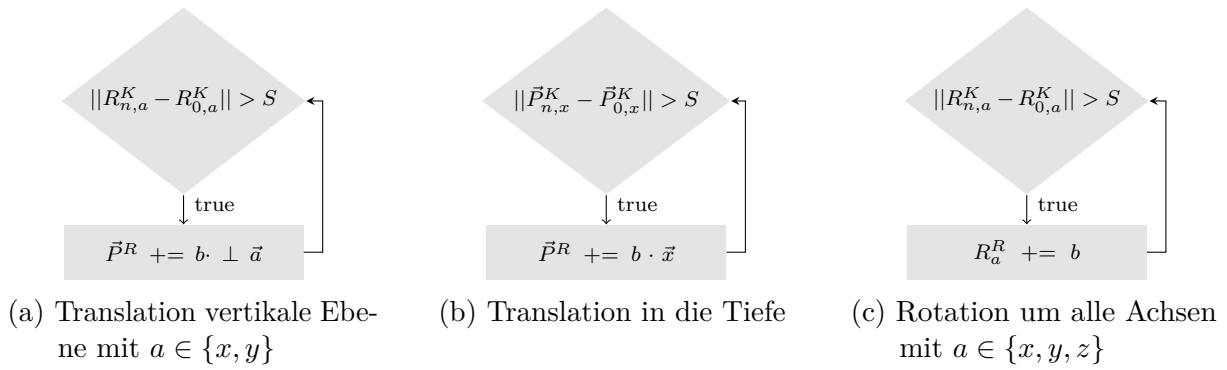


Abbildung 5.3: Logik der Interaktionskonzepte für Kopfinteraktion

Handinteraktion

Bei der Handinteraktion ist eine nur teilweise getrennte Manipulation der Freiheitsgrade möglich, sodass die drei translatorischen und die drei rotatorischen Freiheitsgrade jeweils gleichzeitig angepasst werden können. Somit wird die Bewegung nicht wie bei der Kopfinteraktion auf diskrete Achsen begrenzt. Um trotzdem die vom Standort der Nutzenden abhängige Interaktion zu ermöglichen, werden die Translationsvektoren und Rotationsachsen der Handinteraktion in das RoboterKOS umgewandelt. Diese Umwandlung wird in der folgenden schematischen Beschreibung zur Vereinfachung vernachlässigt und die resultierenden Roboterbewegungen im gleichen Koordinatensystem wie die Handbewegungen dargestellt.

Grundsätzlich wird in der Handinteraktion die relative Translation und Rotation der Hand unskaliert auf Bewegungen des Roboters übertragen. Für die Berechnung der relativen Bewegung der Hand kann aufgrund der Linearität die Differenz zwischen aktueller Position und Ausrichtung und einem Referenzwert verwendet werden. Als dieser Referenzwert, wird die letzte Position und Ausrichtung der Hand verwendet, die zu einer Handbewegung über dem Schwellenwert und somit zu einer Roboterbewegung geführt hat. Der Referenzwert wird somit immer nach Handpositionen und -ausrichtungen, die zu Roboterreaktionen führen, aktualisiert.

Dabei ist die direkte Übertragung der Bewegungen durch die maximale Robotergeschwindigkeiten limitiert. Somit kann für schnelle Handbewegungen die Zeit, die vom Roboter benötigt wird, um die relative Bewegungen der Hand umzusetzen, größer sein, als die der Handbewegung an sich. Da zusätzlich nur Bewegungen über einem bestimmten Schwellenwert an den Roboter gesendet werden, ist die Minimalgeschwindigkeit der Roboterbewegung ebenfalls limitiert.

Der Ablauf für die Translation durch die Handinteraktion wird in Abbildung 5.4a dargestellt. Dabei wird die Handposition durch die Position der Spitze des Zeigefingers der nichtdominanten Hand der Nutzenden beschrieben. Diese wird durch Nutzende zu Beginn der Interaktion festgelegt. Wenn die Änderung der Handposition, von der aktuellen Position

\vec{T}_n^H zur letzten an den Roboter übertragenen Position \vec{T}_{n-1}^H größer als ein Schwellenwert S ist, wird diese Positionsänderung $\Delta\vec{T}^H = \vec{T}_n^H - \vec{T}_{n-1}^H$ auf die Position des Roboters \vec{T}^R übertragen. Dadurch bewegt sich der Endeffektor des Roboters in die gleiche Richtung, um die gleiche Distanz wie die Fingerspitze.

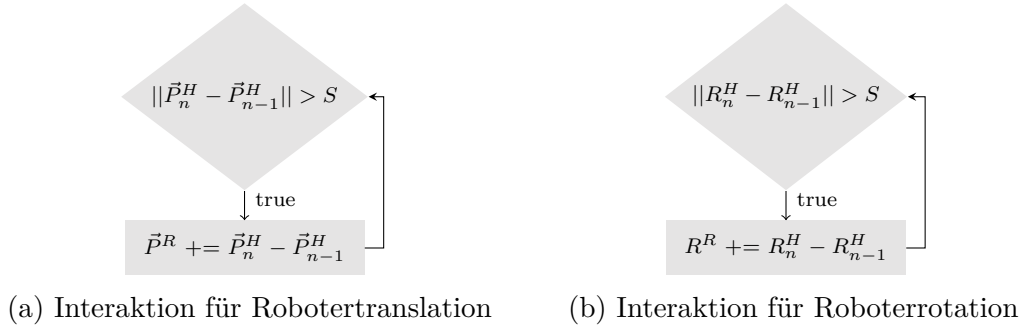


Abbildung 5.4: Logik der Interaktionskonzepte für Handinteraktion

Die Funktionsweise der Rotation wird in Abbildung 5.4b dargestellt. Die Rotation der Hand wird durch die Rotation des senkrecht auf der Handfläche stehenden Vektors bestimmt. Wenn die aktuelle Orientierung der Handfläche R_n^H im Vergleich zur letzten auf den Roboter übertragenen Rotation R_{n-1}^H um ein $\Delta R^H = R_n^H - R_{n-1}^H$ größer als der Schwellenwert S verändert wird, wird die Rotation ΔR^H erneut exakt auf die Rotation des Endeffektors R^R übertragen.

5.3 Finaler Prototyp

Wie aus Abschnitt 6.1.5 hervor geht, wird für den finalen Prototypen die diskrete Interaktion über Sprache und die Navigationsinteraktion durch Handgesten realisiert. Die Interaktionsmethoden werden dazu basierend auf den Ergebnissen der Nutzerstudie verbessert. Zusätzlich werden insbesondere bei der diskreten Interaktion Features hinzugefügt, die für die Umsetzung der finalen Interaktionsaufgabe benötigt werden.

5.3.1 Diskrete Interaktion

Die diskrete Interaktion wird über Sprache umgesetzt. Zusätzlich zur Sprachinteraktion werden symbolische Aktivierungsgesten der Hand zum Starten und Stoppen der Navigationsinteraktion ergänzt. Diese symbolischen Gesten werden während der Ausführung der Navigation gehalten, sodass durch die Ausführung der Geste implizit die Aktivierung und Deaktivierung des Modus erfolgt. Für die Translation wird das Greifen mit Daumen und Zeigefinger (auch als Pinch-Geste bezeichnet) verwendet. Die Rotation erfolgt durch das Bilden einer Faust. Zusätzlich wird zur Erkennung dieser Gesten die Erkennung der

flachen Hand mit ausgestreckten Fingern ergänzt, die die Erkennung des Beendens der Interaktion unterstützen soll.

Um die unerwünschte Erkennung der symbolischen Gesten und somit Fehleingaben zu vermeiden, wird die Bereitschaft der Erkennung symbolischer Gesten durch einen Sprachbefehl initiiert. Neben den symbolischen Gesten können weiterhin die Sprachbefehle *Translation* und *Rotation* für die Aktivierung und *Stopp* für die Deaktivierung der Navigationsmodi verwendet werden.

Zusätzlich zu den bei dem initialen Prototypen verwendeten Sprachbefehlen werden Befehle für die Speicherung und das Anfahren von Roboterpositionen benötigt. Da diese Befehle keinen Modus aktivieren, dessen Systemzustand kontinuierlich ersichtlich sein muss, werden sie nicht als Buttons visualisiert.

Um die Speicherung unterschiedlicher diskreter Positionen zu ermöglichen, werden neben dem Speicher-Befehl weitere Befehle für die Zuordnung der Speicherung benötigt. Um die Auswahlmöglichkeiten zu visualisieren, werden Nutzenden nach dem Speichern-Befehl die Zuordnungsoptionen in Form von Buttons angezeigt. Um diese Buttons wieder auszublenken, ohne einen Punkt zu speichern, kann der *Abbrechen*-Befehl verwendet werden. Analog dazu wird bei dem Befehl zum Anfahren der gespeicherten Punkte die Auswahl gespeicherter Punkte dargestellt.

Für die Speicherung und das Anfahren der diskreten Punkte wird die aktuelle Zielposition des Roboters an die HoloLens übergeben. Die Koordinaten können anschließend, wenn die entsprechende Position angefahren werden soll, an den Roboter als absolute Zielvorgabe gesendet werden. Dabei gilt zu beachten, dass im Gegensatz zu den Navigationsinteraktionen eine größere Strecke durch eine Zielvorgabe überbrückt werden muss. Die Systemsteuerung des Roboters limitiert die Distanz, die durch eine Zielvorgabe überbrückt werden kann, jedoch auf eine maximale Veränderung der Achsenwinkel um 5° . Daher muss in der Robotersteuerung eine zusätzliche Funktion implementiert werden, die das Erreichen der aktuellen Zielvorgabe überprüft und bei Nichterreicherung repetitiv neue Zielvorgaben tätigt, bis das gewünschte Ziel erreicht ist.

Zusammenfassend werden die für die diskrete Interaktion verwendeten Sprachbefehle in Tabelle 5.1 dargestellt.

5.3.2 Navigationsinteraktion

Für die Navigationsinteraktion wird die Hand verwendet. Dabei wird weiterhin zwischen einer Interaktion zur Translation und einer zur Rotation unterschieden. Für die Translationsinteraktion werden keine Änderungen an der in Abschnitt 5.2.2 beschriebenen Interaktion vorgenommen, sodass relative Bewegungen der Hand unskaliert auf Bewegungen des Endeffektors übertragen werden. Das Konzept der Rotationsinteraktion wird

Tabelle 5.1: Übersicht über verwendete Sprachbefehle für den finalen Prototypen

Befehl	Beschreibung
Translation	Start des Translationsmodus
Rotation	Start des Rotationsmodus
Stopp	Beenden der aktuellen Interaktion und der Erkennung der symbolischen Gesten
Start	Aktivierung der Erkennung der symbolischen Gesten
Speichern	Öffnen der möglichen Positionszuordnung zur Speicherung
Anfahren	Öffnen der möglichen Positionen zum Anfahren
Kugel, Pyramide, Würfel	Auswahl der entsprechenden Positionszuordnung zur Speicherung und zum Anfahren
Abbrechen	Schließen der Visualisierung zur Positionszuordnung
Beenden	Beenden der aktuellen Versuchsdurchführung

hingegen angepasst, indem anstatt der direkten Übertragung der relativen Handausrichtung das Konzept der Kopfrotation auf die Hand übertragen wird. Somit werden die Änderungen der Handausrichtung auf die Rotation des Endeffektors um diskrete Rotationsachsen in bestimmten Geschwindigkeiten abgebildet.

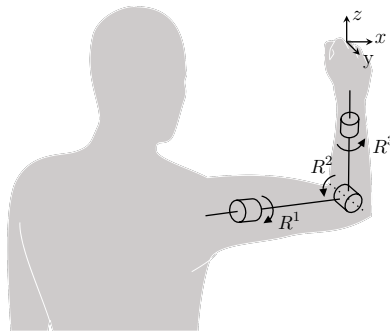


Abbildung 5.5: Für die Interaktion genutzte Freiheitsgrade der Hand

Um die Übertragbarkeit zwischen der Ausrichtung der Hand und der Ausrichtung der Ultraschallsonde zu erleichtern, soll die aufgestellte Hand für die Interaktion verwendet werden (vergleiche Abbildung 5.5). Die Ausrichtung der Hand kann dabei entweder durch die Rotation des Handgelenks oder durch Rotationen des Schulter- und Ellenbogengelenks, die zu Rotationen des gesamten Unterarms führen, verändert werden. Die Bewegung des gesamten Unterarms bietet den Vorteil, dass daraus größere Bewegungen der Hand resultieren. Dadurch können zum einen feiner skalierte Bewegungen und zum anderen eine robuste Erkennung ermöglicht werden. Daher soll die für die Interaktion benötigte Handrotation durch Änderung der in Abbildung 5.5 dargestellten Winkel realisiert werden. Somit wird

der Yaw-Winkel durch die Pronation und Supination der Hand (R^3), der Pitch-Winkel durch die Flexion und Extension des Ellenbogengelenks (R^2) und der Roll-Winkel durch die Abduktion und Adduktion der Schulter in der Sagittalachse (R^1) umgesetzt. Dieser Zusammenhang zwischen Änderung der Handausrichtung und resultierender Rotation wird in Abbildung 5.6 dargestellt.

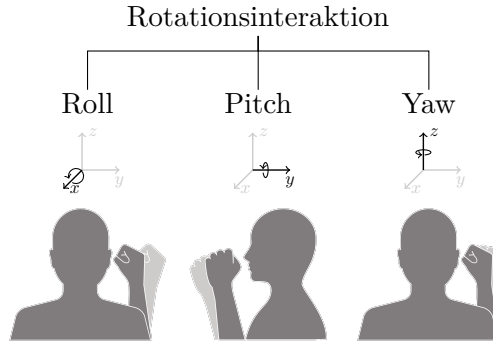


Abbildung 5.6: Überblick über die finale Handinteraktion für die Rotation des Endeffektors

Um die Ausrichtung der Hand zu bestimmen, könnte äquivalent zum initialen Prototypen die Rotation der Handfläche verwendet werden. Jedoch hat sich für den Roll- und Pitch-Winkel gezeigt, dass diese bei aufgestellter Hand nicht zuverlässig erkannt werden können. Da diese beiden Rotationen zusätzlich jedoch auch zu einer Translation der Hand führen, kann die Bewegung der Hand als Parameter zur Erfassung der Rotation dienen. Somit kann die Änderung des Pitch-Winkels durch die Translation in x-Richtung und die Änderung des Roll-Winkels durch die Translation in y-Richtung identifiziert werden. Die Funktionsweise dieser Interaktionserkennung ist in Abbildung 5.7a dargestellt. Solange die Translation der Hand \vec{T}_n^H im Vergleich zur Ruhelage \vec{T}_0^H in die Richtungen $a \in \{x, y\}$ größer als der Schwellenwert S ist, rotiert der Roboter um die dazu orthogonale Achse um die durch die Transferfunktion bestimmte Distanz b . Für die Erkennung dieser Bewegungen wird die Position des Handgelenks getrackt, da das Handgelenk auf der Yaw-Achse liegt, sodass durch Veränderung des Yaw-Winkels keine Translation der getrackten Position erfolgt. Für die Manipulation der Rotation um die z-Achse (Yaw) wird die Rotation des Handgelenks getrackt und die resultierende Roboterbewegung entsprechend der in Abbildung 5.7b dargestellten Logik umgesetzt.

Für die Interaktion werden relative Rotationen im Verhältnis zu einer Ruhelage \vec{T}_0^H und R_0^H verwendet. Die aufgestellte Hand hat jedoch viele Freiheitsgrade ohne einen festen räumlichen Bezugspunkt, wodurch im Gegensatz zu Kopf keine präzise natürliche Ruhelage vorhanden ist. Dies kann zu Schwierigkeiten führen, da zum Stoppen von Rotationen die Hand in die Ruhelage bewegt werden muss. Durch den fehlenden räumlichen Bezugspunkt kann es vorkommen, dass bei Rückkehr in die Ruhelage zwar die Ruhelage der zuletzt manipulierten Rotationsrichtung erreicht wird, diese jedoch nicht der Ruhelage der anderen Winkel entspricht. Somit würden in der vermeintlichen Ruhelage fälschlicherweise

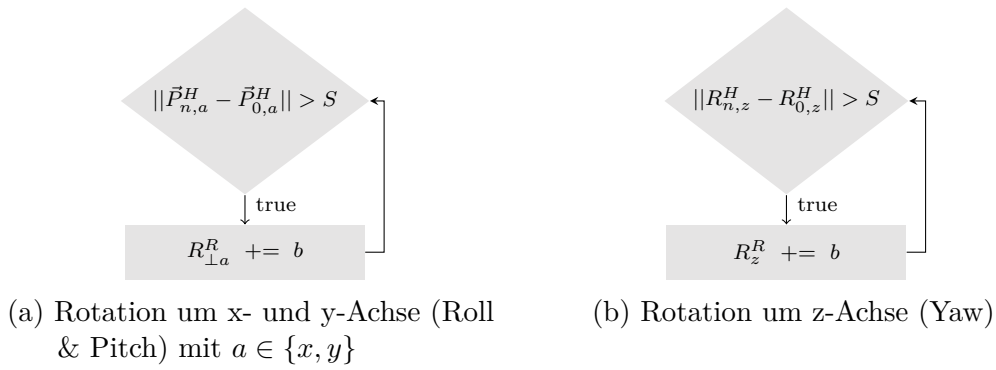


Abbildung 5.7: Logik der Rotationsinteraktionen für finale Navigationsinteraktion

Interaktionen erkannt und somit rotatorische Freiheitsgrade manipuliert werden. Um dies zu verhindern, wird die Ruhelage dynamisch definiert. Unter der Annahme, dass Nutzende nach der Rotation um eine Achse zuerst in die neutrale Position zurückkehren, bevor andere Rotationsrichtungen manipuliert werden, kann bei Rückkehr in die Ruhelage der aktuellen Rotationsrichtung die Ruhelage neu definiert werden. Dabei werden in der neudefinierten Ruhelage nur die Parameter für die zuvor nicht manipulierten Rotationen aktualisiert.

Um Nutzende bei der Interaktion zusätzlich zu unterstützen, wird visuelles Feedback über die Rotationsrichtung gegeben. Dazu wird das in Abbildung 5.8 dargestellte Widget verwendet. Dieses wird zum Start der Rotationsinteraktion als Hologramm auf der Ultraschallsonde angezeigt. Bei Änderung der Handrotation über dem Schwellenwert, die somit zu Roboterbewegungen führen, wird die aus der Interaktion resultierende Rotationsrichtung des Endeffektors in dem Widget durch einen roten Pfeil hervorgehoben. Dadurch kann schon bei kleinen Auslenkungen, die nur zu langsamen Roboterbewegungen führen, direkt die resultierende Rotationsrichtung des Roboters erkannt werden.

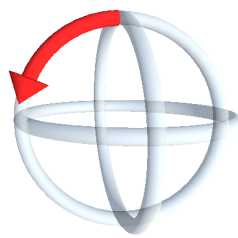


Abbildung 5.8: Visuelles Feedback für die Rotation

5.4 Interaktionsaufgabe

5.4.1 Initiale Studie

Primäraufgabe

Die Primäraufgabe besteht darin, die am Endeffektor befestigte Ultraschallsonde in eine vordefinierte Position zu bewegen. Die Zielposition wird dabei in Form eines semitransparenten Hologramms in AR dargestellt (siehe Abbildung 5.9). Die Festlegung der Zielposition erfolgt über den Desktop-PC, der für die Kommunikation zwischen HoloLens und Roboter verwendet wird. Mithilfe einer einfachen GUI können vordefinierte Sondenpositionen ausgewählt und per UDP an die HoloLens gesendet werden, sodass die Sonde an der angegebenen Position dargestellt wird.

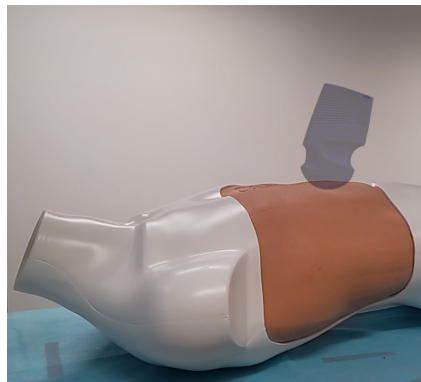


Abbildung 5.9: Darstellung der Sonde als Hologramm auf der HoloLens als Zielposition für die Primäraufgabe der initialen Studie

Sekundäraufgabe

Die Sekundäraufgabe wird ebenfalls in Unity implementiert und auf dem Bildschirm eines PCs ausgeführt. Zur Interaktion mit der Sekundäraufgabe wird eine Fernbedienung verwendet, die über Bluetooth mit dem PC verbunden ist. Für die Interaktion mit dem System werden zwei Buttons benötigt. Um die Aufgabe zu starten, muss ein Button auf der Fernbedienung betätigt werden. Anschließend muss der andere Button einmalig betätigt werden, wenn sich der Regler innerhalb eines Targets befindet. Die Sekundäraufgabe wird automatisch durch Beenden der Primäraufgabe gestoppt.

Die Darstellung der Sekundäraufgabe wird in Abbildung 5.10 dargestellt. Die Länge des Balkens ist dabei 35 cm und für die Bewegung des Reglers entlang des gesamten Balkens werden 20 Sekunden benötigt. Die Targets werden an zufälligen Positionen in der zweiten Hälfte des Balkens initiiert. Die Länge der Targets variiert zufällig zwischen 2 bis 5 cm. Wie in Abbildung 5.10a dargestellt, ist das Target in einem dunkleren Farbton als der

Balken dargestellt. Wenn der Regler sich innerhalb des Targets befindet, wird dieses wie in Abbildung 5.10b dargestellt farblich hervorgehoben.



(a) Darstellung des Zieltargets, wenn sich Regler außerhalb des Targets befindet



(b) Darstellung des Zieltargets, wenn sich Regler innerhalb des Targets befindet

Abbildung 5.10: Darstellung des Balkens der Sekundäraufgabe mit Regler und Zieltargets

5.4.2 Finale Studie

Für die Studie wird ein Phantom benötigt, das mehrere klar differenzierbare Zielstrukturen und zusätzliche Störstrukturen enthält. Da kommerziell verfügbare Phantome wie beispielsweise das Triple Modality 3D Abdominal Phantom Model 057A vom Hersteller CIRS [134] komplexe, teilweise schwer differenzierbare Zielstrukturen beinhalten, wird sich für die Verwendung eines selbst hergestellten simplifizierten Phantoms entschieden.

Um geeignete Materialien für das Phantom zu finden, wurden verschiedene Phantome hergestellt. Für die Grundsubstanz wurden Kerzen-Gel [135], Gelatine und Agar Agar [136] in verschiedenen Konzentrationen getestet. In diese Grundsubstanzen wurden verschiedene Objekte als Ziel- und Störstrukturen eingegossen. Als Zielstrukturen wurden Gummikugeln und Plastikfiguren getestet, während die mögliche Darstellung störender Strukturen zum einen durch Gefäße (in Form von wassergefüllten Luftballons [137] und Latexschläuchen [138]) und zum anderen durch Knochen (3D-gedruckt aus Polyethylen Terephthalat Glykol (PETG) [139]) erfolgte. Die besten Ergebnisse für die Grundsubstanz konnten durch die Verwendung von Agar Agar mit Isopropanol und Wasser erreicht werden (im Verhältnis 1 g : 4.84 ml : 30.30 ml). Als Zielstrukturen können dabei Gummikugeln am besten von der Umgebung abgegrenzt werden, während für die Störstrukturen die 3D-gedruckten Knochen aus PETG am besten geeignet sind. Um verschiedene Zielstrukturen zu ermöglichen, wurden die Gummikugeln in verschiedene Formen zurechtgeschnitten, sodass neben der runden Form auch ein Würfel und eine Pyramide dargestellt werden.

6 Evaluierung

Die im Folgenden beschriebene Evaluierung besteht aus zwei Teilen. Im ersten Teil wird in einer initialen Nutzerstudie anhand einer abstrahierten Interaktionsaufgabe überprüft, welche Interaktionsmodalitäten für die Mensch-Roboter-Interaktion am besten geeignet sind. Im zweiten Teil wird die Anwendbarkeit des finalen Interaktionskonzepts in einer finalen Nutzerstudie unter Verwendung der Ultraschallbildgebung qualitativ evaluiert.

6.1 Evaluierung der Interaktionsmodalitäten

6.1.1 Ziel und Fragestellung

Im Rahmen dieser Nutzerstudie soll identifiziert werden, welche Interaktionsmodalitäten generell für die multimodale Steuerung des Roboters geeignet sind. Dazu wird eine abstrahierte Interaktionsaufgabe verwendet. Dabei sollen die unterschiedlichen Kombinationen der Interaktionsmodalitäten miteinander vergleichen und ihr potentiellen Einfluss auf den Workflow untersucht werden.

6.1.2 Methodisches Vorgehen

Studiendesign

Die Nutzerstudie untersucht drei unabhängige Variablen. Um den Einfluss parallel ablaufender kognitiver Aufgaben auf die Interaktion und den Einfluss der Interaktion auf diese Aufgaben evaluieren zu können, wird von einer Teilgruppe zusätzlich zur Interaktionsaufgabe eine Sekundäraufgabe durchgeführt. Diese Variable wird in einem between-groups Studiendesign untersucht, sodass von Teilnehmenden entweder alle Interaktionen in Kombination mit der Sekundäraufgabe (im folgenden Teilgruppe A) oder lediglich die Interaktionsaufgabe (Teilgruppe B) durchgeführt werden.

Bei den zwei weiteren unabhängigen Variablen handelt es sich um die diskrete und die Navigationsinteraktion. Bei den diskreten Interaktionsmodalitäten wird zwischen Sprache und Blick und bei den Navigationsmodalitäten zwischen Hand- und Kopfingeraktion unterschieden. Diese Variablen werden in einem inbetween-subject Design analysiert, sodass alle Teilnehmenden alle Kombinationen der Interaktionsmodalitäten testen. Die Kombinationen

der Interaktionsmodalitäten werden jeweils an drei verschiedenen Zielpositionen getestet. Die Zuordnung zu den Teilgruppen, die Reihenfolge der Interaktionsmodalitäten und die Reihenfolge der Zielposition erfolgen randomisiert.

Als drei abhängige Variablen werden die Genauigkeit der Positionierung des Roboters, die Dauer für das Abschließen jeder Aufgabe und das subjektive Nutzerfeedback gemessen. Zusätzlich werden zur Evaluierung der in der Sekundäraufgabe gebrachten Leistung die Anzahl der verpassten Targets und der Fehleingaben aufgezeichnet. Für weiteres qualitatives Feedback wird ein semi-strukturiertes Interview durchgeführt.

Erhebungsinstrumente

Für die Bestimmung der Genauigkeit wurde der Rotations- und Translationsfehler gemessen. Dafür wurde als Referenzpunkt der Tool Center Point (TCP) des Roboters verwendet, da dieser als Bezugspunkt der Robotersteuerung genutzt wird. Der TCP wurde vor dem Versuch eingemessen und befindet sich an der Spitze der Ultraschallsonde. Für die Ground-Truth Positionen wurde der Roboter in die entsprechenden Positionen bewegt und die Translation und Rotation gespeichert. Dadurch konnte die Genauigkeit für die Translation über den euklidischen Abstand und für die Rotation über den Winkel zwischen Soll- und Ist-Rotation bestimmt werden.

Die Dauer wurde in Form des Task-Completion-Times (TCT) angegeben, was die Zeit beschreibt, die zum vollenden der Aufgabe benötigt wird. Dazu wurde die Zeit zwischen dem manuellen Start und Beenden der Aufgabe gemessen.

Um das subjektive Nutzerfeedback aufnehmen zu können, wurde der NASA TLX Fragebogen [140] verwendet (siehe Anhang A.1.2). Dieser soll dazu dienen, die empfundene mentale, physiologische und zeitliche Beanspruchung der Teilnehmenden zu erfassen. Der Fragebogen wurde ungewichtet als sogenannter Raw TLX, ausgewertet.

Die in Teilgruppe A aufgenommenen verpassten Targets und Fehleingaben werden als Fehlerquoten relativ im Verhältnis zu jeweils korrespondierenden korrekten Eingaben angegeben.

Für die Evaluierung wurden der Mittelwert für die Dauer, den Translations- und den Rotationsfehler der Teilnehmenden über die drei Durchläufe der Interaktionskombinationen verwendet. Für die Analyse der Dauer, der NASA TLX Bewertung, des Translations- und des Rotationsfehlers wurde eine dreifaktorielle Varianzanalyse (ANOVA) durchgeführt. Für Teilgruppe A wurden die Ergebnisse der Sekundäraufgabe (verpasste Targets und Fehleingaben) für die Durchläufe der einzelnen Interaktionskombinationen gemittelt und eine zweifaktorielle Varianzanalyse durchgeführt.

Das während des semistrukturierten Interviews aufgenommene Feedback wurde von einer Person geclustert und qualitativ ausgewertet.

Stichprobe

Da für die Evaluierung der Interaktionsmodalitäten eine abstrahierte Interaktionsaufgabe verwendet wurde, die kein medizinisches Fachwissen erfordert, gab es hinsichtlich der Fachkenntnisse der Teilnehmenden keine Einschränkungen.

Studienaufbau

Der Studienaufbau ist schematisch in Abbildung 6.1 dargestellt. Die Teilnehmenden wurden gebeten, sich innerhalb eines auf dem Boden markierten Rechtecks zu positionieren, sodass sie gerade vor dem Roboter standen. An dem Flansch des Roboters wurde ein 3D-gedrucktes Modell einer Ultraschallsonde befestigt. Vor den Teilnehmenden war auf einem Tisch ein Torso-Phantom platziert, auf dessen Oberfläche in Augmented Reality die Zielpositionen angezeigt wurden. Zusätzlich war auf dem Tisch der für die Kalibrierung benötigte Marker befestigt. Rechts neben dem Roboter war der Monitor für die Sekundäraufgabe platziert.

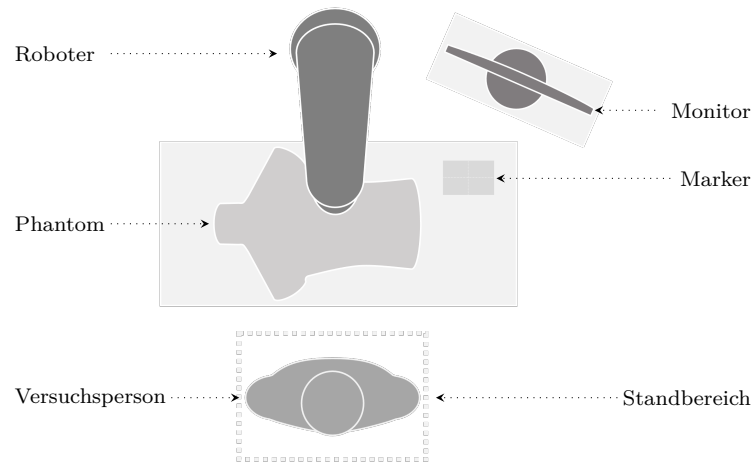


Abbildung 6.1: Übersicht über den Versuchsaufbau der initialen Nutzerstudie

Studiendurchführung

Zu Beginn der Studie wurden die Teilnehmenden begrüßt und gebeten, die Einverständniserklärung zur Teilnahme und zu den Hygienemaßnahmen sowie einen Fragebogen zu den demographischen Daten auszufüllen. Anschließend wurden die Interaktionsaufgabe und die verschiedenen Interaktionsmodalitäten vorgestellt. Zusätzlich erhielt die Kontrollgruppe A eine Erklärung zur Funktionsweise der Sekundäraufgabe.

Vor dem Start der Interaktionsaufgabe musste das System kalibriert werden, indem der Marker für drei Sekunden aus leicht unterschiedlichen Blickwinkeln betrachtet wurde. Zusätzlich wurde mithilfe der HoloLens eine für das Augentracking erforderliche Augenkalibrierung durchgeführt.

Darauf folgten die Blöcke der verschiedenen Interaktionen. Diese bestanden jeweils aus

Trainings- und Evaluierungsdurchläufen. Für das Training konnten die Interaktionsmodalitäten an 2-3 Testzielpositionen getestet werden. Wenn die Teilnehmenden angaben, sich sicher in der Interaktion zu fühlen, wurden die drei Durchläufe durchgeführt, bei denen die Daten für die Evaluierung aufgenommen werden. Zum Abschluss jedes Blockes wurde der NASA TLX Fragebogen ausgefüllt.

Im Anschluss an die vier Blöcke wurde das semi-strukturiertes Interview durchgeführt. Die Studie dauerte pro Proband*in circa 60 Minuten.

6.1.3 Ergebnisse

Demographie der Teilnehmenden

An der Studie haben 16 Proband*innen (6 weiblich, 9 männlich, 1 divers) teilgenommen. Die Teilnehmenden waren zwischen 23 und 30 Jahren alt (Median = 25 Jahre). Bei den Teilnehmenden handelte es sich um Studierende (8), Dokotrand*innen (5), wissenschaftliche Mitarbeiter*innen (2) und Ingenieur*innen (1). Acht der Teilnehmenden hatten keine Sehschwäche, sieben eine korrigierte und eine Person eine unkorrigierte Fehlsichtigkeit. Da die Interaktion mit der nicht-dominanten Hand ausgeführt werden sollte und das System dementsprechend angepasst werden musste, wurde die Händigkeit der Teilnehmenden erfragt. Dabei gaben 17 der Teilnehmenden an rechtshändig und eine Person linkshändig zu sein. Um Komplikationen mit der Sprachsteuerung zu prüfen, wurde die Muttersprache der Teilnehmenden abgefragt, wobei alle bis auf eine Person Deutsch als Muttersprache angaben. Zusätzlich wurde eine Selbsteinschätzung zur technischen Affinität und Erfahrungen mit berührungslosen Interaktionen auf einer Likert Skala erfragt. Die Teilnehmenden gaben eine technische Affinität zwischen 4 und 5 (1: niedrig, 5: hoch, Median=5) und eine Erfahrung mit berührungslosen Interaktionen zwischen 1 und 5 an (1: gar keine, 5: sehr viel, Median=3).

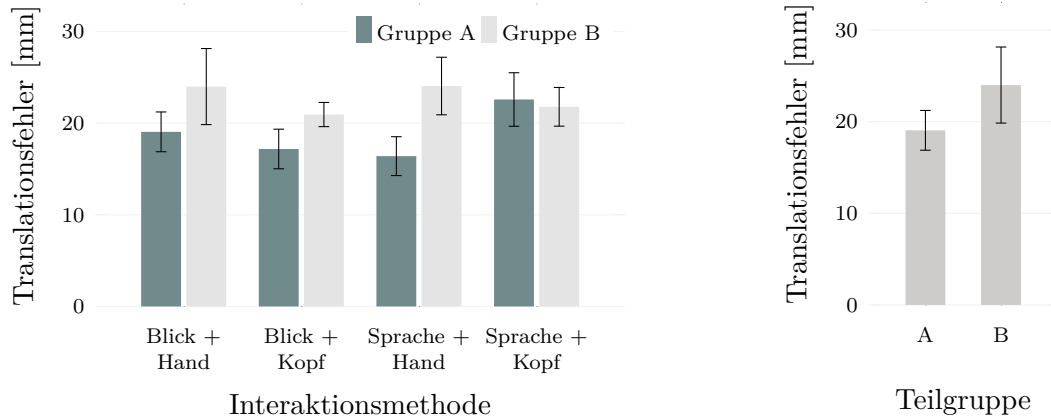
Quantitative Ergebnisse

Die aus der Varianzanalyse hervorgehenden statistisch signifikanten Ergebnisse werden in Tabelle 6.1 dargestellt. Die Lernkurven der Teilnehmenden über alle Durchläufe und über die Durchläufe pro Interaktionskombination sind Anhang A.3.2 zu entnehmen.

Tabelle 6.1: Statistisch signifikante Ergebnisse der initialen Nutzerstudie

Abhängige Variable	Unabhängige Variable	F	p	partielles η^2
Translationsfehler	Sekundäraufgabe	4.33	0.0419	0.072
Dauer	Manipulationsinteraktion	4.57	0.0370	0.075
NASA TLX	Diskrete Interaktion	4.15	0.0463	0.069

Die durchschnittlichen Translationsfehler werden in Abbildung 6.2a und die entsprechenden Lernkurven in Abbildung A.1 dargestellt. Es kann ein signifikanter Einfluss mit mittlerem Effekt der Anwesenheit der sekundär Aufgabe auf den Translationsfehler festgestellt werden (Abbildung 6.2b). Der gesamte durchschnittliche Translationsfehler lag bei 20.75 ± 0.95 mm.



(a) Mittlere Translationsfehler für unterschiedliche Interaktionsmethoden und Teilgruppen (b) Mittlerer Translationsfehler nach Teilgruppen

Abbildung 6.2: Darstellung des mittleren Translationsfehlers mit Standardfehler

In Abbildung 6.3 werden die mittleren Rotationsfehler und in Abbildung A.2 die Lernkurven dargestellt. Dabei konnte kein signifikanter Einfluss der unabhängigen Variablen festgestellt werden. Jedoch konnte ein geringer Effekt von der Kombination der Aktivierungs-, Navigationsinteraktion und Teilgruppe auf den Rotationsfehler festgestellt werden (partielles $\eta^2 = 0.055$). Der gesamte mittlere Rotationsfehler ist $13.58 \pm 0.61^\circ$.

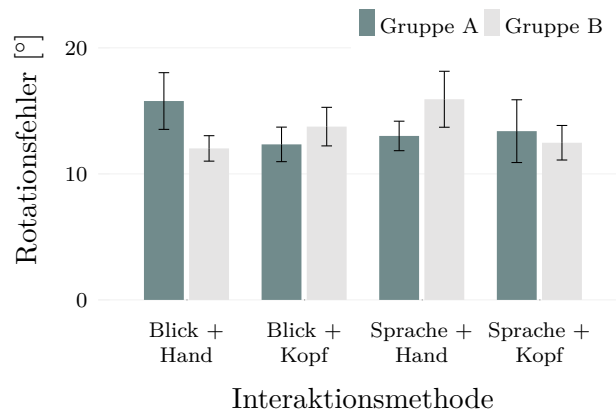
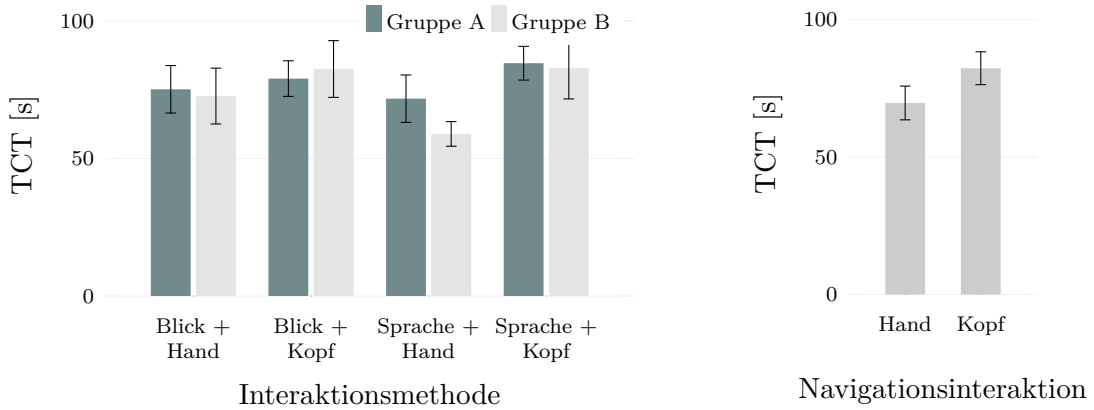


Abbildung 6.3: Darstellung des mittleren Rotationsfehlers für die unterschiedlichen Interaktionsmethoden unterteilt nach den Teilgruppen mit Standardfehler

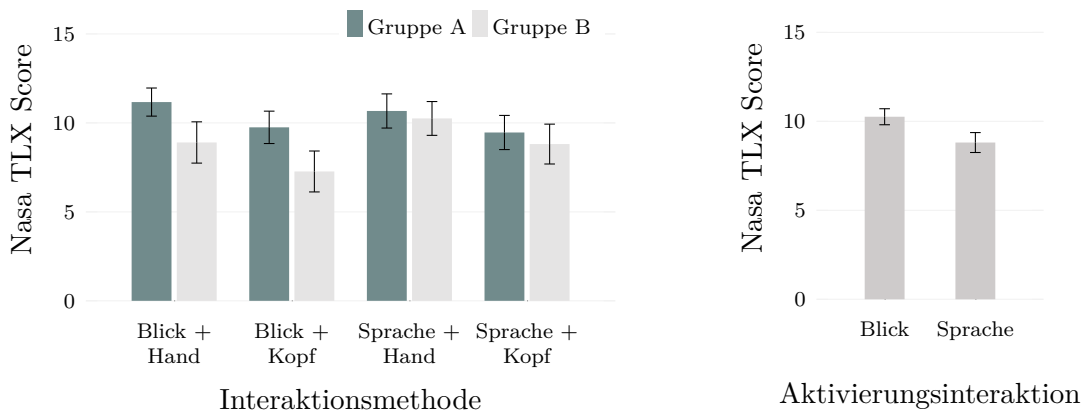
Der durchschnittliche TCT wird in Abbildung 6.4a und die entsprechenden Lernkurven in Abbildung A.3 dargestellt. Für die Dauer konnte ein signifikanter Einfluss mit mittlerem Effekt der Navigationsinteraktion auf den TCT festgestellt werden (Abbildung 6.4b). Der mittlere TCT für alle Teilnehmenden und Interaktionsmethoden beträgt 75.98 ± 3.01 s.



(a) Mittlerer TCT für unterschiedliche Interaktionsmethoden und Teilgruppen (b) Mittlerer TCT nach Navigationsinteraktion

Abbildung 6.4: Darstellung der mittleren Task-Completion-Time mit Standardfehler

Der durchschnittliche Raw TLX ist in Abbildung 6.5 dargestellt. Die durchschnittliche Bewertungen der unterschiedlichen Dimensionen sind Abbildung 6.6 zu entnehmen. Auf das Ergebnis des Scores konnte ein signifikanter Einfluss mit mittlerem Effekt der verwendeten diskrete Interaktion gefunden werden (Abbildung 6.5b). Die durchschnittliche TLX Bewertung der Teilnehmenden beträgt 9.53 ± 0.37 .



(a) Mittlere Nasa TLX Score für unterschiedliche Interaktionsmethoden und Teilgruppen (b) Mittlerer TLX Score nach Aktivierungsinteraktion

Abbildung 6.5: Darstellung des mittleren Raw Nasa TLX Scores mit Standardfehler

Die gemessenen Kennwerte für die Evaluierung der Sekundäraufgabe werden in Abbildung 6.7 und die Lernkurven in Abbildung A.4 dargestellt. Dabei kann kein direkter signifikanter Einfluss der unabhängigen Variablen auf die verpassten Targets und die Fehlclicks festgestellt werden. Jedoch kann bei den verpassten Targets ein Trend festgestellt werden, dass die Performance der Interaktion mit Sprache und Hand besser ist, als die der anderen Interaktionskombinationen. Insbesondere im direkten Vergleich zwischen Sprache und Hand und Blick und Hand lässt sich ein signifikanter Effekt ($p=0.048$, $\eta^2 = 0.24$) nachweisen.

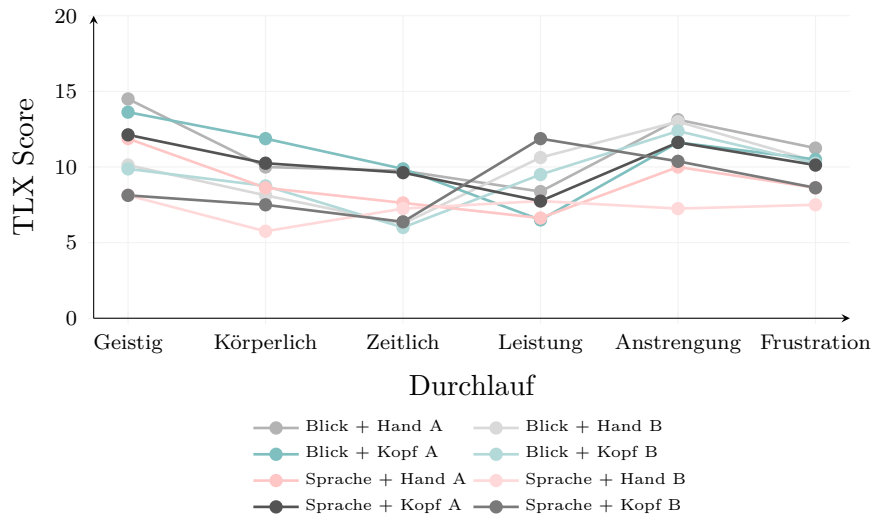


Abbildung 6.6: Nasa TLX Bewertungen der unterschiedlichen Dimensionen nach Interaktionskombination und Versuchsgruppe

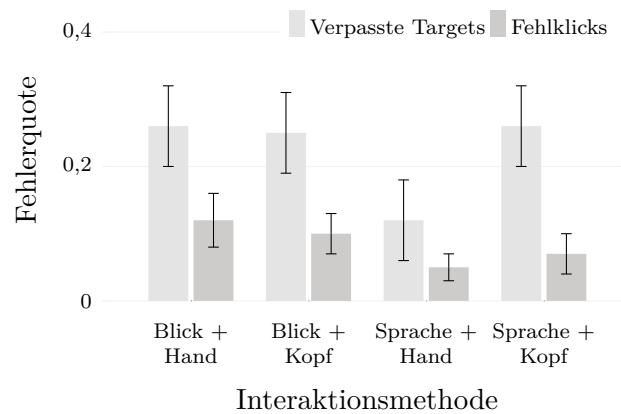


Abbildung 6.7: Darstellung der durchschnittlichen Fehlerquoten für die verpassten Targets und die Fehlclicks der Sekundäraufgabe

Um Einflüsse der unterschiedlichen Zielpositionen auf die Ergebnisse ausschließen zu können, wurde zusätzlich die Abhängigkeit zwischen Zielposition und abhängigen Variablen untersucht. Dabei konnten keine signifikanten Effekte festgestellt werden.

Qualitatives Feedback

Nachfolgend wird das aus dem semistrukturierten Interview hervorgehende qualitative Feedback beschrieben. Die in Klammern angegebenen Zahlenwerte weisen auf Häufigkeiten bestimmter Aussagen hin. Dabei sollte jedoch beachtet werden, dass aus diesen Häufigkeiten keine Schlussfolgerungen gezogen werden sollten, da es sich um qualitative Daten handelt. Zusätzlich sind diese Zahlen aufgrund des Interviewformats nicht zwangsläufig repräsentativ.

Bei der diskreten Interaktion durch den Blick wurde die Verweilzeit, die zum Auslösen

der Buttons benötigt wurde, von vier Teilnehmenden als lang empfunden. Dies führe insbesondere bei dem Stoppen von Interaktionen zu Schwierigkeiten, da laut vier Teilnehmenden in der Phase, während der Blick erkannt wurde, ungewollte Interaktionen stattfinden konnten. Eine weitere Schwierigkeit bestand laut vier Teilnehmenden darin, dass die Buttons versehentlich aktiviert werden konnten, wenn diese sich aufgrund der Kopfausrichtung im Zielbereich befanden. Eine ähnliche Schwierigkeit wurde von vier Teilnehmenden in Bezug zur Sekundäraufgabe festgestellt, da hierbei durch Blick auf den Bildschirm Buttons versehentlich aktiviert werden konnten. Generell sei für die Interaktion mit den Buttons laut drei Teilnehmenden Konzentration erforderlich.

Im Gegensatz dazu wurden die Sprachbefehle als einfacher (3) und weniger ablenkend (3) beschrieben. Laut drei Teilnehmenden war die Erfüllung der Sekundäraufgabe unter der Verwendung von Sprache einfacher. Zusätzlich sei die Aktivierung durch Sprache schneller (3), was insbesondere bei dem *Stopp*-Befehl wichtig sei. Drei Teilnehmende merkten daher an, dass es hilfreich wäre, bei der Blicksteuerung auch die Option des *Stopp*-Sprachbefehls zu verwenden. Jedoch beschrieben im Gegensatz dazu drei Teilnehmende Schwierigkeiten mit Latenz der Spracherkennung, die die Interaktion erschwert hätte.

Die Navigationsinteraktion mit der Hand wurde von den Teilnehmenden generell als intuitiv empfunden. Dabei wurde insbesondere die Translation als einfach (3), schnell (3) und intuitiv (5) beschrieben. Im Gegensatz dazu wurden für die Rotation Schwierigkeiten geschildert. Es wurde von sieben Teilnehmenden angemerkt, dass bei großen Rotationen der Hand die Bewegungen des Roboters nicht nachvollziehbar wären. Eine weitere Herausforderung bestand laut vier Teilnehmenden darin, dass die Hand anders ausgerichtet sei als die Sonde, sodass die Übertragung der Ausgangsorientierung der Sonde auf die Ausgangsorientierung der Hand schwierig sei. Zusätzlich sei die Rotation laut drei Teilnehmenden durch die Anatomie der Hand eingeschränkt, sodass die Rotation in kleinere Rotationen zerlegt werden müsse. Aufgrund dieser Schwierigkeiten beschrieben sieben Teilnehmende, dass sie sich eine Verbesserung der Handrotation vorstellen könnten, wenn diese nicht direkt übertragen, sondern wie bei der Kopfrotation Auslenkungen der Hand aus der Ausgangslage in kontinuierliche Rotationen transformiert würden.

Die Navigationsinteraktion mit dem Kopf wurde von fünf Teilnehmenden als körperlich anstrengend beschrieben. Die Translation wurde aufgrund der Zerlegung der Freiheitsgrade als langsam empfunden. Insbesondere die Steuerung über das nach vorne und hinten Lehnen wurde von drei Teilnehmenden als gewöhnungsbedürftig beschrieben, da viel Bewegung für die Steuerung erforderlich sei. Im Gegensatz dazu wurde die Rotation als einfach (5), intuitiv (3) und nachvollziehbar (3) beschrieben. Eine Schwierigkeit der Kopfbewegungen bestand laut sieben Teilnehmenden darin, dass durch große Rotationen des Kopfes der Zielbereich und das Hologramm nicht mehr gut sichtbar seien. Insbesondere bei Nicken des Kopfes sei dies laut vier Teilnehmenden ein Problem. Im Bezug auf die Sekundäraufgabe wurde von vier Teilnehmenden beschrieben, dass durch die Kopfrotation

bei Blick auf den Bildschirm ungewollte Interaktionen ausgelöst und somit unerwünschte Roboterbewegungen entstehen können.

Eine generelle Schwierigkeit bestand in dem räumlichen Verständnis der Proband*innen. Laut drei Teilnehmenden sei die Tiefe des Zielhologramms schwierig erkennbar. Da auch das räumliche Verständnis in Bezug auf die Roboterbewegungen schwierig sei, merkten fünf Teilnehmende an, dass zur Unterstützung während der Interaktionen eine Visualisierung, um welche Achse der Roboter sich aktuell bewegt, hilfreich sein könnte.

6.1.4 Diskussion

Bezüglich der diskreten Interaktion konnte ein signifikanter Einfluss auf die NASA TLX Scores nachgewiesen werden. Dabei war der Raw TLX für die diskrete Interaktion mittels Blick höher als der mittels Sprache. Dieser zeigte sich insbesondere in der Dimension der Anstrengung, die benötigt wird, um das Ziel zu erreichen. Dies deckt sich auch mit dem Feedback, das von den Teilnehmenden im Interview gegeben wurde. Teilnehmende beschreiben, dass für die Blickaktivierung mehr Konzentration erforderlich sei. Zusätzlich müssen nicht nur bei der Auswahl die Buttons bewusst fokussiert werden, sondern diese auch bewusst vermieden werden, wenn der Blick auf etwas anderes gerichtet wird.

Für die verwendete Navigationsinteraktion konnte ein signifikanter Einfluss auf den TCT gefunden werden. Dabei war die Dauer unter der Verwendung der Handinteraktion signifikant kürzer als unter Verwendung der Kopfingeraktion. Im Interview gaben Teilnehmende an, dass insbesondere die Translation mit der Hand intuitiv und schnell ausführbar sei, was sich auf den TCT auswirken könnte. Dadurch, dass die Translation der Hand direkt auf die des Roboters übertragen werden kann, ist die parallele Manipulation der translatorischen Freiheitsgrade möglich, was im Gegensatz zu den getrennten Freiheitsgraden bei der Kopfingeraktion zu einer kürzeren TCT führen könnte.

Im Interview wurden jedoch Schwierigkeiten hinsichtlich der Rotation mit der Hand beschrieben. Dabei wurde insbesondere darauf eingegangen, dass große Rotationen mit der Hand zu nicht immer vorhersehbaren Rotationen des Roboters führen würden. Eine Ursache für diese Schwierigkeiten wurde von den Teilnehmenden darin identifiziert, dass die Ausgangslage der Hand (flache, vertikal ausgestreckte Hand) schwieriger auf die Ausgangslage der Sonde (senkrecht aufgestellt) übertragen werden kann als die des Kopfes. Eine weitere Ursache lässt sich auf die Bewegungsart des Roboters zurückführen. Bei großen Rotationen der Hand wird die neue Soll-Rotation an den Roboter gesendet, jedoch nicht der Pfad. Der Roboter setzt dann eine PTP-Bewegung um, die zu der entsprechenden Rotation führt. Dazu werden nicht die euklidischen Rotationen interpoliert, sondern die Achsenwinkel des Roboters. Dadurch resultiert eine Bewegung, die im kartesischen Raum nicht zwangsläufig linear ist. Somit ist der Pfad, der insbesondere bei großen Rotationen

gefahren wird, nicht immer für Nutzende vorhersehbar und kann zu Verwirrungen führen. Denn selbst wenn die Endposition korrekt wäre, konnte während der Studie beobachtet werden, dass Nutzende bei unerwarteten Bewegungen die Rotation der Hand veränderten, wodurch die Zielposition nicht erreicht werden konnte. Diese Problematik konnte in der Rotation durch Kopfsteuerung umgangen werden, da dort kontinuierlich sehr kurze Strecken vorgegeben werden, woraus ein nahezu linearer Pfad im kartesischen Raum resultieren kann, der besser nachvollzogen werden kann.

Es konnte ein signifikanter Einfluss der Anwesenheit der Sekundäraufgabe auf den Translationsfehler gefunden werden. Dieser ist signifikant höher, wenn parallel keine Sekundäraufgabe durchgeführt wird. Die unterschiedlichen Genauigkeiten könnten sich möglicherweise auf die Zusammensetzung der Gruppen zurückführen lassen, die randomisiert erfolgte. Dem demographischen Feedback lassen sich Unterschiede hinsichtlich der mittleren Selbsteinschätzung zur technischen Affinität (Gruppe A: 4.75, Gruppe B: 4.5 (1:niedrig, 5:hoch)) und zu Erfahrungen mit berührungslosen Interaktionen (Gruppe A: 3.75, Gruppe B: 2.99 (1: gar keine, 5: sehr viel)) entnehmen. Diese sind zwar nicht signifikant, zeigen jedoch trotzdem quantitative Unterschiede, insbesondere bezüglich der Erfahrung mit berührungslosen Interaktionen. Durch eine Gruppeneinteilung, die zu einer gleichmäßigen Verteilung von Expertise in den unterschiedlichen Teilgruppen führt, könnten möglicherweise andere Ergebnisse erzielt werden.

Für den Rotationsfehler konnten keine signifikanten Unterschiede gefunden werden. Jedoch lässt sich ein Trend erkennen, dass sowohl bei der Kombination Blick und Hand als auch bei der Kombination Sprache und Kopf der Rotationsfehler jeweils geringer wird, wenn keine Sekundäraufgabe parallel ausgeführt werden muss. Dass dieser Effekt nicht signifikant ist, könnte an der geringen Studiengröße liegen. Dies würde sich mit den qualitativen Ergebnissen decken, in denen von Teilnehmenden insbesondere der Einfluss der Sekundäraufgabe auf Kopf- und Blickinteraktionen hervorgehoben wurde, da der Blick auf den Monitor zu fälschlichen Interaktionen geführt habe.

Bezüglich der Kennwerte für die Performance der Sekundäraufgabe konnten keine signifikanten Abhängigkeiten von den Interaktionskombinationen festgestellt werden. Dies könnte sich zum einen auf die geringe Stichprobengröße zurückführen lassen, zum anderen handelt es sich um diskrete aufgenommene Werte. Bei einer durchschnittlichen Durchlaufdauer von 75.98 s und einem Abstand der Targets von 10-20 s mussten die Teilnehmenden auf 3-7 Targets reagieren. Da ein Target nur entweder verpasst oder nicht verpasst werden kann, hat ein Fehler direkt einen sehr großen Einfluss auf das Ergebnis. Kontinuierliche Werte, wie beispielsweise über das Gedrückthalten einer Taste, während sich der Regler innerhalb eines Targets befindet, könnten feiner skalierten Aufschluss über die Performance geben. Insbesondere für den Vergleich der Fehlerquote der verpassten Targets könnte dies aufschlussreich sein, da dort bereits ein Trend erkennbar ist, dass insbesondere die

Interaktion mit Hand und Sprache zu einer geringeren Fehlerquote führen kann.

In der Lernkurve ist für die Fehlerquote der verpassten Targets ein abfallender Trend zu erkennen. Somit verbesserten Teilnehmende sich sowohl im Laufe der Nutzerstudie als auch über die jeweiligen drei Durchläufe für die jeweiligen Interaktionskombinationen in der Sekundäraufgabe. Dies könnte darauf hinweisen, dass die Trainingsdurchläufe vor Beginn der Aufzeichnung nicht ausreichend waren und dass durch mehr Training konstantere und geringere Fehlerquoten erreicht werden könnten.

Die Lernkurven für den TCT, den Translations- und Rotationsfehler, weisen starke Schwankungen auf, wodurch kein eindeutiger Trend erkennbar ist. Das könnte erneut ein Hinweis auf unzureichendes Training sein, sodass durch eine längere Lernphase die Ergebnisse möglicherweise stabilisiert werden könnten.

Für die Studie sind weitere Limitationen zu beachten. Die Studie wurde unter Laborbedingungen durchgeführt. Dadurch können keine externen ablenkenden Faktoren auf die Interaktionen betrachtet werden. Beispielsweise konnte der Einfluss von Gesprächen und Störgeräuschen im OP auf die Sprachsteuerung somit nicht evaluiert werden. Zusätzlich wurde die Interaktionsaufgabe für die Studie abstrahiert. Im normalen Workflow würde anhand des Ultraschallbilds beurteilt werden, wie der Roboter gesteuert werden soll. Die Genauigkeit der Positionierung der Ultraschallsonde wird somit nicht an der Position der Ultraschallsonde an sich, sondern an dem Ultraschallbild bemessen. Dies hat zum einen Einfluss auf die Bewertung der Genauigkeitsdaten, da die exakte Positionierung während des realen Workflows eine untergeordnete Rolle spielt. Zum anderen verändert die abstrahierte Aufgabe das Verhältnis zum Monitor. Während der realen Intervention würde der Monitor als Positionierungshilfe dienen. Das bedeutet, dass das Blicken auf den Bildschirm zur Verifikation der aktuellen Position direkt mit der Roboterbewegung zusammenhängt. Die in der Studie verwendete Aufgabe verlief jedoch unabhängig zur Roboterbewegung, sodass in regelmäßigen Abständen auf den Monitor geblickt werden musste. Durch die Asynchronität der beiden Aufgaben wird somit der Schwierigkeit der Aufgabe erhöht, was besonders in die Betrachtung der qualitativen Ergebnisse einbezogen werden muss.

6.1.5 Schlussfolgerungen

Die Nutzerstudie sollte dazu dienen, eine geeignete Kombination an Modalitäten für die Robotersteuerung für die Aktivierungs- und Navigationsinteraktion zu identifizieren.

Für die diskrete Interaktion konnte gezeigt werden, dass das Feedback über das subjektive Empfinden der Nutzenden sowohl im NASA TLX Fragebogen als auch im Interview für die Sprachsteuerung positiver ausgefallen ist, als das für die Blicksteuerung. Als Vorteile der Sprachsteuerung wurde genannt, dass die Interaktion dadurch schneller, einfacher und

unter weniger Konzentration möglich und dadurch weniger frustrierend sei. Daher wird sich für die Weiterentwicklung der Interaktion für die Verwendung von Sprache entschieden. Eine Schwierigkeit der Sprachinteraktion war jedoch die teilweise auftretende Latenz, die insbesondere beim Stoppen der Interaktion zu Problem führen könnte. Daher könnte es sinnvoll sein, die Eingabemöglichkeiten zur Aktivierung und Deaktivierung der Navigationsinteraktion durch andere Eingabemodalitäten zu ergänzen, um im Fall von auftretender Latenz der Spracherkennung weiterhin eine robuste Interaktion zu ermöglichen.

Bei der Navigationsinteraktion konnte gezeigt werden, dass die Interaktionsaufgabe unter der Verwendung von Handinteraktion signifikant schneller absolviert werden konnte als mit Interaktion durch Kopfbewegungen. Somit wird für die Weiterentwicklung des Konzepts die Hand für die Navigationsinteraktion verwendet. Jedoch sollten insbesondere bei der Rotationsinteraktion Verbesserungen umgesetzt werden, die zu einer intuitiven und besser nachvollziehbaren Interaktion führen. Ein möglicher Ansatz wäre es, das von den Teilnehmenden in Interview positiv hervorgehobene Rotationskonzept des Kopfes auf die Hand zu übertragen. Indem die Auslenkung der aufgestellten Hand aus einer Ruheposition auf kontinuierliche Rotationen des Roboters übertragen werden würde, könnte die Übertragbarkeit der Handrotation auf die Roboterrotation und die Vorhersehbarkeit der Roboterrotation verbessert werden. Zusätzlich könnte visuelles Feedback über die aktiven Rotationsachsen als Unterstützung für Nutzende ergänzt werden.

6.2 Evaluierung des Gesamtkonzepts

6.2.1 Ziel und Fragestellung

Die finale Nutzerstudie soll zur qualitativen Evaluierung des finalen Interaktionskonzepts dienen. Dabei soll die Eignung des Systems zur Steuerung der Ultraschallbildgebung untersucht werden. Dazu wird das Interaktionskonzept unter Verwendung der Ultraschallbildgebung getestet, um Feedback zu dem Gesamtsystem erhalten zu können.

6.2.2 Methodisches Vorgehen

Studiendesign

In der Nutzerstudie soll die Nutzerfreundlichkeit des Systems untersucht werden. Dazu werden während der Studiendurchführung nach dem sogenannten Think Aloud Protokoll [141] Daten aufgenommen. Teilnehmende sollen dafür während der Versuchsdurchführung permanent ihre Gedanken aussprechen, wodurch direktes Nutzerfeedback erhalten werden kann. Basierend auf den somit getätigten Aussagen können anschließend weiterführende Fragen gestellt werden. Dies erfolgt in einem semi-strukturierten Interview. Da neben allgemeinen Fragen die während der Versuchsdurchführung erfassten Anmerkungen einbezogen werden, variieren die gestellten Fragen zwischen den Teilnehmenden. Zusätzlich wird quantifizierbares Feedback zur Nutzerfreundlichkeit durch einen standardisierten Fragebogen aufgenommen.

Erhebungsinstrumente

Zur Kategorisierung der einzelnen während der Studie getätigten Aussagen werden diese den jeweiligen Versuchspersonen zugeordnet in eine Tabelle übertragen. Nach der Einteilung der Aussagen in unterschiedliche Kategorien, werden Überschneidungen zwischen Teilnehmenden identifiziert und bei mindestens zweifach auftretenden Aussagen Cluster gebildet. Für diese Cluster wird abschließend eine zusammenfassende Aussage formuliert.

Für die Aufnahme des subjektiven Nutzerfeedbacks wird die System Usability Scale (SUS) von Brooke [142] verwendet (siehe Anhang A.1.3). Diese ermöglicht durch die Beantwortung von 10 Fragen auf einer Likertskala die effiziente Sammlung statistischer Daten über die empfundene Usability.

Stichprobe

Um die Interaktionsaufgabe absolvieren zu können, müssen Ultraschallbilder von den Teilnehmenden ausgewertet werden. Somit sind fachspezifische Vorkenntnisse erforderlich. Daher werden als Teilnahme Kriterien ein medizinischer Hintergrund und Erfahrungen in Ultraschallbildgebung festgelegt. Als Aufwandsentschädigung haben die Proband*innen 30€ für ihre Teilnahme erhalten.

Studienaufbau

Der Aufbau der Studie erfolgt entsprechend dem in Abschnitt 6.1.2 beschriebenen Studienaufbau. Dabei wird der dort für die Sekundäraufgabe verwendete Bildschirm durch ein Apple iPad Pro (12.9 Zoll, 2. Generation) zur Darstellung der Ultraschallbilder ersetzt. Am Endeffektor des Roboters wird statt dem 3D-gedruckten Sondenmodell eine Clauritius Scanner C3 Ultraschallsonde befestigt, die über ein lokales Netzwerk die aufgenommenen Bilder an das iPad sendet.

Für die Studie werden zwei Phantome benötigt, eins für das Training und eins für die Versuchsdurchführung. Die Bestandteile dieser Phantome sind dabei identisch, wobei das Trainingsphantom 20 x 15 x 5 cm und das Evaluierungsphantom 26.9 x 19.8 x 5 cm groß ist. In den Phantomen befinden sich drei unterschiedlich geformte Gummikugeln (Pyramide, Würfel und Kugel) und eine PETG Rippe. Bei der Rippe handelt es sich um die 12. Rippe eines weiblichen Brustkorbes, die aus den DICOM Daten des Human Visual Project segmentiert wurde. Die Anordnung der Elemente wird schematisch in Abbildung 6.8 dargestellt. Dabei wird die würfelförmige Gummikugel in einer Schicht unterhalb der Rippe platziert, sodass er bei senkrechter Betrachtung im Schatten der Rippe liegt.

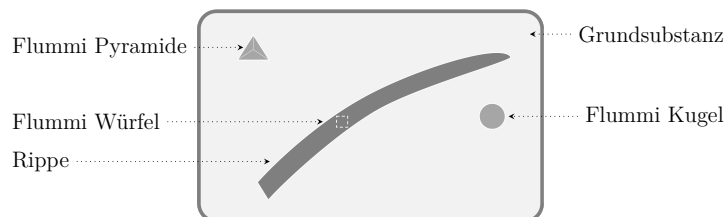


Abbildung 6.8: Aufbau des für die finale Nutzerstudie verwendeten Ultraschallphantoms

Studiendurchführung

Für die Teilnahme an der Studie musste von den Teilnehmer*innen die Einverständniserklärung zur Teilnahme und den Hygienemaßnahmen sowie ein Fragebogen zur Erfassung demographischer Daten ausgefüllt werden. Daraufhin wurde eine Einführung in die Thematik gegeben, die eine Beschreibung des potentiellen Anwendungsfelds sowie eine Beschreibung der Interaktionsaufgabe umfasste.

Um Teilnehmende sowohl mit der Mensch-Roboter-Interaktion als auch mit der verwendeten Ultraschallsonde vertraut zu machen, wurden vor der Durchführung der Interaktionsaufgabe zwei Trainingsaufgaben durchgeführt. Die erste Aufgabe sollte die Identifikation der Zielobjekte auf dem Ultraschallbild unterstützen. Dazu wurde ein Testphantom verwendet, in dem die gleichen Zielobjekte wie im Studienphantom platziert waren. Diese sollen unter Handführung der Ultraschallsonde lokalisiert werden.

Die zweite Trainingsaufgabe diente zur Einarbeitung in die Robotersteuerung. Dazu wurde eine Aufgabe ähnlich zu der in der initialen Studie verwendeten Aufgabe durchgeführt (vgl. Abschnitt 4.3.1), indem in AR Endeffektorpositionen und -ausrichtungen in Form eines Zielhologramms vorgegeben wurden, in die die Ultraschallsonde bewegt werden sollte. Um zusätzlich die Speicherfunktionen zu trainieren, wurde unterhalb der Hologramme jeweils eine geometrische Form dargestellt, sodass die Sondenpositionen an den entsprechenden Speicherorten gespeichert werden konnten. Es wurden in einer randomisierten Reihenfolge drei Sondenpositionen vorgegeben, die jeweils gespeichert werden sollten. Anschließend wurde eine Sondenposition erneut gezeigt, sodass die zuvor gespeicherte Position für das Anfahren genutzt werden konnte. Wenn Teilnehmende nach dem Training angaben, sich sicher in der Handhabung der Interaktion und der Auswertung der Ultraschallbilder zu fühlen und keine weiteren Fragen zu haben, wurde die Studienaufgabe durchgeführt.

Die Studienaufgabe bestand wie in Abschnitt 4.3.2 beschrieben aus der Lokalisierung drei verschiedener Objekte in dem Ultraschallphantom in einer vordefinierten Reihenfolge. Nachdem das dritte Objekt lokalisiert wurde, wurde die Versuchsdurchführung beendet. Anschließend wurde von den Teilnehmenden der SUS-Fragebogen ausgefüllt und das semi-strukturierte Interview durchgeführt. Die Durchführung der Studie nahm pro Teilnehmer*in circa 45 Minuten in Anspruch.

6.2.3 Ergebnisse

Demographie der Teilnehmenden

An der Studie haben acht Versuchsperson*innen (6 weiblich, 2 männlich) im Alter zwischen 24 und 29 Jahren (Median = 26.5) teilgenommen. Davon waren sieben Medizinstudierende zwischen dem 6. und 15. Semester (Median = 10) und eine Person Assistenzärzt*in. Sieben der Teilnehmenden gaben an, dass Deutsch ihre Muttersprache sei. Bezüglich des Sehvermögens gaben fünf Teilnehmende an, keine Einschränkungen zu haben und drei korrigierte Einschränkungen zu haben. Die technische Affinität der Teilnehmenden wurde zwischen 2 und 5 eingeschätzt (1: niedrig, 5: hoch, Median = 4). Bezüglich Erfahrung im Umgang mit berührungslosen Interaktionen gaben die Teilnehmenden Werte von zwischen 2 und 5 (1: keine, 5: viel, Median = 2.5) und im Umgang mit Ultraschall zwischen 3 und 5 (1: keine, 5: viel, Median = 3) an.

Quantitative Ergebnisse

Die Usability des Systems konnte durch den SUS-Fragebogen mit einem mittleren Score von 69.94 (max. = 100, Standardabweichung = 17.95) bewertet werden.

Qualitative Ergebnisse

In der Studie konnten 217 individuelle Statements erfasst werden. Davon konnten 104 Aussagen in 38 Statements zusammengefasst und in 13 Kategorien eingeteilt werden. Diese lassen sich in generelles Feedback und Feedback zur diskreten und Navigationsinteraktion unterteilen. Das generelle Feedback umfasst die Funktionalität, die Modalität, das Phantom sowie das durch das System erhaltene Feedback. Bezüglich der diskreten Interaktion wurde Feedback über die Sprachbefehle, das Speichern und Anfahren von Punkten sowie die Aktivierungsgeste gegeben. Das Feedback zur Navigationsinteraktion lässt sich in allgemeines Feedback und Feedback über die Translation, die Rotation und den Interaktionsraum unterteilen. In allen drei Hauptkategorien wurden weiterhin Aussagen bezüglich möglicher Zusatzfunktionen getätigt. Die kategorisierten Aussagen können Tabelle 6.2 entnommen werden.

Tabelle 6.2: Zusammenfassung der qualitativen Ergebnisse der finalen Nutzerstudie in den verschiedenen Kategorien. Die ID gibt dabei die entsprechenden Proband*innen an

Kategorie	Statement	ID
Generell		
Funktionalität	Reaktionen des Systems waren vorhersehbar	1,2,3
	Automatisches Stoppen des Roboters sehr gut	6,8
Modalität	HoloLens ist angenehm zu tragen	3,7
	Man muss sich an die HoloLens gewöhnen	6,7
Phantom	Im Phantom fehlen feste Orientierungspunkte, die sonst durch Anatomie vorhanden	4,5,6
Feedback	LED-Ring am Roboter ist hilfreich	3,4,8
	LED-Ring am Roboter ist bei Blick auf Monitor außerhalb des Sichtfeldes	3,5,6,7
	Es fehlt das haptische Feedback	1,2,8
	Schwierig einzuschätzen, wie viel Druck aufgewendet wird	1,4,5
Zusatzfunktionen	Genauerer Feedback über aufgewendeten Druck	1,6,8
	Zusatzsignal bei zu starkem Aufdrücken	3,5,6,8
	Zusätzliche Unterstützung zur Orientierung (z.B. miniatur Karte)	3,6,8

Fortsetzung auf nächster Seite

Tabelle 6.2: Fortsetzung

Kategorie	Statement	ID
Diskrete Interaktion		
Sprachbefehle	Spracherkennung hat gut funktioniert	3,4,5,7
	Sprachbefehle sind intuitiv	1,3,7
	<i>Anfahren</i> Befehl ist nicht intuitiv	1,3
	Sprache dauert für Auswahl der Navigationsmodi zu lang	3,5
Speichern und Anfahren	Punkte speichern und anfahren hat gut funktioniert	1,2,3
	Punkte speichern und anfahren praktisch	5,8
Aktivierungsgeste	Aktivierungsgesten zu ähnlich	2,3,5
	Alternative Interaktion zur Aktivierung der Navigationsmodi wären hilfreich	3,5
Zusatzfunktionen	Befehl, um große Rotationen durchzuführen (beispielsweise 45° oder 90°)	5,6
	Befehle um Ultraschallfunktionen zu bedienen	5,7,8
Navigationsinteraktion		
Allgemein	Gesten für die Interaktion sind intuitiv	1,2,4,6,7,8
	Interaktion ist schnell erlernbar	2,7
	Roboterbewegungen sind nachvollziehbar	4,8
	Roboterbewegungen sind sehr präzise	4,8
	Sehr schnelle Übertragung von Geste auf Roboterbewegung	6,7,8
	Aktuelle Position und Orientierung von Sonde schwierig einzuschätzen, wenn Blick auf Monitor gerichtet ist	2,3,4,8
Translation	Bewegung auf einer Ebene schwierig	1,2,7
Rotation	Rotation manchmal durch anatomische Limitationen des Arms herausfordernd	3,4,5
	Übung benötigt, um zu lernen, dass Startorientierung des Arms für Rotation angepasst werden kann	3,4,8
	Hilfreich, dass während der Rotation die Position konstant gehalten wird	5,8
Interaktionsraum	Es muss sich bewusst an das Sichtfeld der Kamera erinnert werden	1,3
	Sichtfeld der Kamera zur Gestenerkennung zu klein	3,4,5
	Es ist körperlich anstrengend, dass aufgrund des Interaktionsbereichs die Hand hochgehalten werden muss	3,4
	Handposition wäre stabiler, wenn die Hand nicht frei in der Luft wäre	1,8
Zusatzfunktionen	Möglichkeit, Achsen zu Sperren um die Bewegung auf Ebene zu ermöglichen	2,7
	Automatisches Halten von konstantem Druck auf Gewebe	3,7

6.2.4 Diskussion

Zur Einordnung des erreichten SUS-Scores kann der von Sauro [143] vorgestellte Richtwert verwendet werden. Dieser definiert 68 als den Durchschnittsscore, der aus der Auswertung von 500 Studien hervorging. In dieser Studie konnte ein SUS-Score von 69.94 erreicht werden, der sich somit als überdurchschnittlich einordnen lässt und für eine gute Usability steht. Dies spiegelt sich auch in den Ergebnissen des Interviews wieder, in denen die Sprachbefehle und die Gesten als intuitiv beschreiben wurden. Jedoch geht aus den Interviews auch Verbesserungspotential hervor, das im Folgenden diskutiert wird.

Das Feedback durch den LED-Ring am Roboter und das automatische Stoppen bei zu starkem Gegendruck wurde von den Teilnehmenden als hilfreich empfunden. Jedoch gaben Teilnehmende an, dass sich der LED-Ring am Roboter teilweise außerhalb ihres Sichtfeldes befand. Dies lässt sich darauf zurückführen, dass während der Studie der Blick eher auf den Monitor zur Darstellung der Ultraschallbilder gerichtet war als auf den Roboter. Dadurch konnte teilweise nicht direkt erkannt werden, wenn der Roboter aufgrund einer Kollision gestoppt und der entsprechende LED-Ring geleuchtet hat. Um dies zu verbessern, könne zusätzliches akustisches, haptisches oder visuelles Feedback verwendet werden. Das visuelle Feedback sollte sich dabei entweder immer im Sichtfeld befinden (z.B. über die Darstellung auf der HoloLens) oder in der Nähe bzw. direkt auf dem Bildschirm zur Darstellung der Ultraschallbilder angezeigt werden.

Weiterhin gaben die Teilnehmenden an, dass es schwierig sei einzuschätzen, wo sich die Sonde auf dem Phantom befindet, wenn der Blick auf den Monitor gerichtet ist. Dies führen sie zum einen auf fehlende anatomische Strukturen zurück, durch die Anhaltspunkte, die im menschlichen Körper zur Orientierung dienen, fehlen. Zum anderen fehle der direkte Kontakt zur Sonde, durch den bei der Handführung eine grobe Einschätzung über die aktuelle Position und Orientierung möglich sei. Somit sollte evaluiert werden, ob diese Schwierigkeiten weiterhin auftreten, wenn das Phantom anatomische Strukturen abbilden würde. Weiterhin könnte es sinnvoll sein, im Sichtfeld Informationen über die aktuelle Position und Ausrichtung der Sonde anzuzeigen.

Sowohl die fehlende Wahrnehmung des Feedbacks auf dem LED-Ring als auch die fehlende Orientierung bezüglich aktueller Position und Orientierung der Sonde resultieren daraus, dass während der Betrachtung der Ultraschallbilder auf dem Monitor der Endeffektor des Roboters sich außerhalb des Sichtfeldes befindet. Neben der Möglichkeit, Feedback in der Nähe des Monitors anzuzeigen, könnte jedoch auch die Position der Darstellung des Ultraschallbildes angepasst werden. Dazu könnte einerseits der Bildschirm anders positioniert werden, oder andererseits Fähigkeiten der HoloLens genutzt werden, um die Ultraschallbilder positionsunabhängig holographisch im Sichtfeld darzustellen.

Eine weitere Schwierigkeit bestand laut den Teilnehmenden darin, dass ihnen generell haptisches Feedback fehlen würde. Dadurch sei es schwierig einzuschätzen, wie viel Druck

von dem Roboter auf das Phantom angewendet wird. Um dies zu verbessern, könnten zur Unterstützung neben dem Feedback über zu starken Druck genauere Angaben über die aktuell ausgeübte Kontaktkraft gegeben werden. Eine Möglichkeit könnte dabei visuelles Feedback beispielsweise in Form eines Balkens, der relative Angaben über die ausgeübte Kontaktkraft macht, sein.

Bezüglich der Eingabemodalität wurde das Tragen der HoloLens als angenehm beschrieben. Jedoch gaben die Teilnehmenden auch an, dass sie sich an das Tragen der Brille gewöhnen mussten. Dies lässt sich möglicherweise auf das Gewicht und das leicht veränderte Sichtfeld zurückführen.

Für die diskrete Interaktion wurde angemerkt, dass die Spracherkennung gut funktioniert habe und die Sprachbefehle intuitiv seien. Die einzige Ausnahme habe dabei der Befehl zum Anfahren der Punkte gebildet, der von zwei Teilnehmenden als nicht intuitiv empfunden wurde. Daher könnte es sinnvoll sein, für diese Funktion einen anderen geeigneten Sprachbefehl zu identifizieren.

Generell habe das Speichern und Anfahren der Punkte jedoch gut funktioniert und sei praktisch gewesen. Alle Teilnehmenden haben die entsprechenden Befehle verwendet, wenn während der Versuchsdurchführung ein Zielobjekt gefunden wurde, das nicht dem aktuell gesuchten entsprach. Somit konnte diese Funktionalität erfolgreich für die Erfüllung der Interaktionsaufgabe genutzt werden.

Einige der Teilnehmenden kritisierten, dass sich die Gesten zur Aktivierung der Navigationsmodi, zu ähnlich seien. Dadurch wurde manchmal durch die Gesten versehentlich der falsche Modus aktiviert. Daher könnte die Ergänzung alternativer Eingabemöglichkeiten hilfreich sein. Sprache sei dabei keine geeignete Alternative, da die Aussprache der Befehle zu lange dauern würde. Somit könnte es sinnvoll sein, entweder die Erkennung der aktuell verwendeten Aktivierungsgesten zu verfeinern, sodass weniger versehentliche Aktivierungen zustande kommen oder andere Handgesten zu verwenden. Eine weitere Möglichkeit wäre die Ergänzung einer noch nicht verwendeten Eingabemodalität, wie etwa Pedale.

Zusätzlich könnte es hilfreich sein, weitere Funktionen, die durch die diskrete Interaktion gesteuert werden, zu ergänzen. Zum einen könnte die Steuerung von Funktionen der Ultraschallbildgebung wie die Aufnahme von Screenshots oder der Wechsel zu anderen Aufnahmemodi ergänzt werden. Zum anderen könnten Befehle zur Unterstützung der Navigationsinteraktion hilfreich sein, beispielsweise um die Sonde direkt um größere Winkel zu rotieren. Dies könnte beispielsweise die Darstellung bestimmter Strukturen aus unterschiedlichen Perspektiven unterstützen, sodass eine Struktur jeweils aus der gleichen Position in perpendikularen Schichten visualisiert werden könnte.

Für die Navigationsinteraktion wurden die Gesten als intuitiv empfunden, wodurch sie schnell erlernbar seien. Die Abbildung auf die Roboterbewegungen sei dabei nachvollziehbar, präzise und schnell. Während der Translation sei eine Schwierigkeit die Bewegung auf

einer Ebene. Dies führe insbesondere bei der Bewegung auf der Oberfläche des Phantoms zu Schwierigkeiten, da dort bei versehentlichen vertikalen Bewegungen entweder durch den Kontaktverlust die Bildgebung schlechter wird oder es durch zu hohen Druck zu Kollisionen kommt. Dies kann sich darauf zurückführen lassen, dass die Hand keinen festen Bezugspunkt hat, sodass es zu Fehlbewegungen kommen kann. Eine Möglichkeit, dies zu verhindern, könnte das Sperren bestimmter Achsen sein, sodass Bewegungen der Hand in diese Richtung ignoriert und nur die verbleibenden Freiheitsgrade auf den Roboter abgebildet werden. Da dieses Problem jedoch hauptsächlich bei der Bewegung auf einer Oberfläche auftritt und diese in der Praxis nicht zwangsläufig einer geraden Ebene entspricht, könnte es auch sinnvoll sein, die Position über den aufgewendeten Druck zu regulieren. Somit könnte, beispielsweise durch eine initiale Positionierung, eine Soll-Kontaktkraft festgelegt werden. Anschließend könnte der Roboter weiterhin in zwei Raumrichtungen verfahren werden, während die Höhe der Sonde automatisch durch die Einhaltung der Soll-Kontaktkraft bestimmt wird.

Für die Rotation bestand eine Herausforderung darin, dass die Freiheitsgrade des menschlichen Arms anatomisch begrenzt sind, sodass der Arm aus bestimmten Ausgangslagen in bestimmte Richtungen nicht ausreichend weit gedreht werden kann. Dieses Problem kann überwunden werden, indem die Ausgangslage angepasst wird. Jedoch merkten die Teilnehmenden an, dass sie einige Zeit benötigt haben, um die Anpassung der Ausgangslage zu erlernen. Eine Ursache hierfür könnte sein, dass der Arm intuitiv entsprechend der Orientierung der Sonde ausgerichtet wird. Daher gilt zu evaluieren, ob dieses Problem nach einer längeren Trainingsphase weiterhin auftritt. Eine Möglichkeit zur Überwindung dieses Problems könnte die Reduktion des *Toten-Bereichs* sein, in dem durch die Bewegung der Hand keine Interaktion ausgelöst wird. Dadurch könnte die Rotation des Roboters schon durch kleinere Rotationen der Hand angepasst werden. Hierbei müsste jedoch darauf geachtet werden, dass der *Tote-Bereich* weiterhin groß genug ist, um die Erkennung ungewollter Interaktionen zu vermeiden.

Eine allgemeine Schwierigkeit wurde durch den Raum verursacht, in dem die Gesten erkannt wurden, der durch das Sichtfeld der Kamera der HoloLens definiert ist. Das Sichtfeld sei zu klein und Teilnehmende gaben an, sich bewusst an das Sichtfeld erinnern zu müssen, damit die Gesten erkannt werden. Da sich die Kamera auf Höhe der Stirn befindet, ist das Sichtfeld der Kamera relativ hoch. Somit muss während der Interaktion die Hand hochgehalten werden, was laut den Teilnehmenden körperlich anstrengend sei und zur Instabilität der Handposition geführt habe. Daher könnte es sinnvoll sein, alternative Schnittstellen für die Aufnahme der Handgesten einzusetzen. Dadurch könnte der Einfluss des Interaktionsbereiches auf die Anstrengung und die Genauigkeit während der Interaktion bestimmt werden.

Für die durchgeführte Studie ergeben sich weiterhin allgemeine Limitationen. In der Studie wurde die isolierte Aufgabe des Vorstrahlens für die manuelle Steuerung des Roboters

gewählt, da dadurch die Verwendung der diskreten Befehle und der Navigationsinteraktion abgedeckt werden könnte. Jedoch können dadurch noch keine Erkenntnisse über die Integrierbarkeit in den Workflow und über das Gesamtsystem für den robotergestützten Ultraschall gewonnen werden. Dazu müsste eine Studie durchgeführt werden, die neben dem manuellen Vorstrahlen, auch die Nadelführung und die dafür verwendete autonome Robotersteuerung beinhaltet. Dabei könnte weiterhin auch der Aufbau angepasst werden, sodass durch das Phantom die Anatomie eines Menschen abgebildet wird, wodurch Erkenntnisse über die Orientierung und die Steuerung in einer realistischeren Umgebung getroffen werden könnten.

Eine weitere Limitation resultiert aus der Stichprobe. Die Teilnehmenden der Studie haben angegeben, eine mittlere Erfahrung im Umgang mit Ultraschall zu haben (Median der Selbsteinschätzung der Ultraschallerfahrung = 3 (1: keine, 5: viel)). Dadurch konnten zwar die gesuchten Strukturen vollständig lokalisiert werden, jedoch könnte sich eine höhere Erfahrung auf den Umgang mit dem robotergestützten Ultraschall auswirken. Dies könnte beispielsweise Einflüsse auf die Orientierung während der Bildgebung sowie der Umgang mit dem fehlenden haptischen Feedback haben. Somit sollten in zukünftigen Evaluierungen Einschätzungen von Ärzt*innen verschiedener Erfahrungsgrade betrachtet werden, um umfassenderes Feedback über die klinische Eignung und Verbesserungsmöglichkeiten zu erhalten.

6.2.5 Schlussfolgerungen

Ziel der Nutzerstudie war es, das finale Interaktionskonzept hinsichtlich Nutzerzufriedenheit zu evaluieren und qualitatives Feedback zu erhalten. Dabei hat sich gezeigt, dass durch das System eine gute Nutzerfreundlichkeit erreicht werden konnte. Zusätzlich konnten bestehende Herausforderungen und Unterstützungsmöglichkeiten identifiziert werden. Verbesserungen, die in künftigen Arbeitsschritten umgesetzt werden könnten, umfassen dabei insbesondere geeignetes Feedback im Sichtfeld über den aufgewendeten Druck und die Position und Orientierung der Sonde, gut abgrenzbare Aktivierungsgesten, Unterstützung bei der Bewegung auf einer Ebene bzw. unter konstantem Druck, Anpassung des Interaktionsraums sowie die Implementierung von Zusatzfunktionen wie die Rotation um große Winkel und den Zugriff auf Ultraschallfunktionen.

6.3 Auswertung der Anforderungsanalyse

Als abschließender Teil der Evaluierung muss überprüft werden, ob die in Abschnitt 3 vorgestellten funktionalen und nichtfunktionalen Anforderungen durch das System erfüllt werden konnten.

Durch die funktionalen Anforderungen wurde festgelegt, welche grundlegenden Fähigkeiten durch das System ermöglicht werden sollen. Im Rahmen dieser Arbeit wurde ein System entwickelt, das mithilfe einer HoloLens die multimodale Interaktion zur Steuerung eines Roboters für die Ultraschallbildgebung ermöglicht. Dabei konnten, wie Tabelle 6.3 zu entnehmen ist, alle der funktionalen Anforderungen erfüllt werden.

Tabelle 6.3: Zusammenfassung über die Auswertung der funktionalen Anforderungen

Anforderungen	Status	Begründung
Technische Rahmenbedingungen		
Multimodale Schnittstelle	Erfüllt	Interaktion mit dem Roboter über Sprache und Handgesten durch HoloLens umgesetzt
Registrierung	Erfüllt	Kamerakoordinatensystem und somit auch Interaktion sind mit dem Roboterkoordinatensystem über einen optischen Marker registriert
Grundfunktionalitäten		
Aktivierung Modi	Erfüllt	Durch Sprachbefehle können unterschiedliche Modi aktiviert/deaktiviert werden
Navigation Roboter		
Anpassung Ausrichtung	Erfüllt	Rotationen der Hand können auf Roboter Geschwindigkeiten zur Anpassung der Ausrichtung abgebildet werden
Anpassung Position	Erfüllt	Translation der Hand kann direkt auf Roboterbewegungen abgebildet werden
Gespeicherte Positionen	Erfüllt	Positionen können durch Sprachbefehle gespeichert und angefahren werden
Vordefinierte Positionen	Erfüllt	Bei Beenden des Programms wird der Roboter automatisch in eine vordefinierte Position gefahren

Durch die nichtfunktionalen Anforderungen werden die zu erfüllenden Qualitätsanforderungen festgelegt. Diese konnten, wie Tabelle 6.4 zu entnehmen ist, überwiegend erfüllt

werden. Eine Ausnahme bilden dabei die klinischen Anforderungen, deren Erfüllung im Rahmen dieser Arbeit noch nicht vollständig überprüft werden konnte. Zur Gewährleistung der Sterilität wurde zwar eine berührungslose Eingabemethode gewählt, jedoch wird für die Interaktion die HoloLens benötigt. Daher müsste zusätzlich überprüft werden, ob diese steril in das Arbeitsumfeld integriert werden kann.

Hinsichtlich des Einflusses der Interaktion auf bestehende Arbeitsabläufe wurde in der initialen Nutzerstudie bei der Auswahl der Interaktionsmodalitäten mithilfe der Sekundäraufgabe darauf geachtet, dass die Interaktion möglichst geringen Einfluss auf parallel ablaufende kognitive Aufgaben hat. Jedoch müsste zur Evaluierung der tatsächlichen Integrierbarkeit in den Workflow eine weitere Studie im Zusammenhang mit dem bestehenden Workflow mit medizinischem Fachpersonal durchgeführt werden.

Tabelle 6.4: Zusammenfassung über die Auswertung der nichtfunktionalen Anforderungen

Anforderungen	Status	Begründung
Gebrauchstauglichkeit		
Effektivität	Erfüllt	Durch die Interaktion konnte die Ultraschallsonde zur Visualisierung der Zielstrukturen positioniert werden und in der finalen Studie alle Zielstrukturen lokalisiert werden
Effizienz	Erfüllt	Durch die initiale Studie konnten Eingabemodalitäten gewählt werden, die die Interaktion unter geringem kognitiven körperlichen und zeitlichen Aufwand ermöglichen
Zufriedenheit	Erfüllt	In der Studie konnte durch die System Usability Scale ein Score von 69.94 erreicht werden, was über dem von Sauro [143] vorgestellten Referenzwert liegt
Leistung		
Robustheit	Teilweise Erfüllt	Bei der Interaktion mit dem System traten keine systematischen Fehler oder Störungen auf. Die Aktivierungsgesten konnten jedoch teilweise nicht eindeutig identifiziert werden, sodass es zu fälschlichen Aktivierungen der Modi kam
Performance	Erfüllt	Interaktion und entsprechende Roboterreaktionen erfolgen in Echtzeit

Fortsetzung auf nächster Seite

Tabelle 6.4: Fortsetzung

Anforderungen	Status	Begründung
Klinische Anforderungen		
Sterilität	Teilweise Erfüllt	Für die Interaktion müssen keine Eingabegeräte mit den Händen berührt werden, jedoch müsste Sterilität der HoloLens gewährleistet werden
Arbeitsumfeld	Nicht Evaluiert	Die Integrierbarkeit in bestehende Arbeitsabläufe und das Arbeitsumfeld konnte aufgrund der Stichprobe und der Isolation der Interaktionsaufgabe noch nicht evaluiert werden
Feedback		
Rückmeldung Systemzustand	Erfüllt	Aktive Modi und Systemzustand sind durch visuelles und auditives Feedback erkennbar, könnte jedoch zukünftig erweitert werden

7 Abschluss

Das Ziel dieser Arbeit war die Konzeption, Entwicklung und Evaluierung einer multimodalen Benutzerschnittstelle für die Interaktion mit einem Roboterarm zur Ultraschallbildgebung im medizinischen Interventionsraum. Im Folgenden werden die gewonnenen Erkenntnisse zusammenfasst und ein Überblick über weitere mögliche Arbeitsschritte gegeben.

7.1 Zusammenfassung

Zu Beginn wurden die für die Arbeit notwendigen medizinischen und technischen Grundlagen erläutert. Dabei wurde insbesondere auf Eingriffe der interventionellen Radiologie, Roboter im medizinischen Interventionsraum und Mensch-Roboter-Interaktion eingegangen. Weiterhin wurden verwandte Arbeiten in Bezug auf robotergestützten Ultraschall und die multimodale Mensch-Roboter-Interaktion vorgestellt.

Zur Analyse des Nutzungskontexts nadelgestützter Interventionen unter kombinierter Ultraschall- und CT-Bildgebung wurden Experteninterviews durchgeführt. Basierend auf den daraus resultierenden Aufgaben für einen Roboterarm, der während der Intervention die Ultraschallsonde positioniert, konnten somit Anforderungen an eine Schnittstelle zur Mensch-Roboter-Interaktion identifiziert werden.

Zur Erfüllung der Nutzungsanforderungen wurden verschiedene multimodale Interaktionskonzepte zur Kommunikation mit dem Roboter entwickelt. Als Schnittstelle für die Interaktion wurde eine HoloLens verwendet, die mithilfe eines optischen Markers mit dem Roboterkoordinatensystem registriert wurde. Basierend auf den zuvor definierten Anforderungen konnten für die Interaktion Modalitäten für diskrete Interaktionen wie die Auswahl unterschiedlicher Modi und für Navigationsinteraktionen, also zur direkten Steuerung des Roboters, identifiziert werden. Dabei wurden für die diskrete Interaktion Sprache und Blick und für die Navigationsinteraktion Bewegungen des Kopfes und der Hand ausgewählt. Zusätzlich wurden abstrahierte Interaktionsaufgaben zur Evaluierung der Konzepte vorgestellt.

Die Evaluierung erfolgte in zwei Schritten. Im ersten Schritt wurde eine Nutzerstudie durchgeführt, um die am besten geeignete Kombination an Interaktionsmodalitäten zu identifizieren. Dazu wurden die Konzepte hinsichtlich Genauigkeit, Dauer, Nutzerzufriedenheit und Einflüssen einer Sekundäraufgabe verglichen. Dabei stellte sich heraus, dass durch die diskrete Interaktion über Sprache eine signifikant höhere Nutzerzufriedenheit und bei der Navigationsinteraktion über die Hand eine signifikant niedrigere Dauer zur Erfüllung der Interaktionsaufgabe erreicht werden konnte. Jedoch konnte auch Verbesserungspotential, insbesondere hinsichtlich der Interaktion zur Rotation festgestellt werden.

Basierend auf den Erkenntnissen dieser Studie wurde ein verbessertes finales Interaktionskonzept entwickelt, das im zweiten Schritt der Evaluierung in Form einer qualitativen Nutzerstudie untersucht wurde. Im Rahmen dieser Studie konnte gezeigt werden, dass das vorgestellte System für die nutzerfreundliche Erfüllung der funktionalen Anforderung geeignet ist.

Zusammenfassend lässt sich festhalten, dass im Rahmen dieser Arbeit eine nutzerfreundliche multimodale Schnittstelle für die Steuerung eines Roboters zur Ultraschallbildgebung entwickelt werden konnte. Durch eine Nutzerstudie konnte die Kombination aus Sprache und Freihandgesten als geeignete Eingabemodalitäten identifiziert werden. Im Rahmen einer zweiten Nutzerstudie konnte die Nutzerfreundlichkeit des Systems bestätigt und Verbesserungspotentiale identifiziert werden. Basierend auf dem während der finalen Studie erhaltenen Feedback konnten mögliche zukünftige Arbeitsschritte abgeleitet werden.

7.2 Ausblick

Der Diskussion der Evaluierung des Gesamtkonzepts in Abschnitt 6.2.4 lassen sich bereits weitere mögliche Arbeitsschritte entnehmen. Diese umfassen unter anderem die Entwicklung von geeignetem Feedback über die aufgewendete Kontaktkraft des Roboters auf das Gewebe und die Position und Orientierung der Sonde sowie die Unterstützung bei der Bewegung der Sonde auf einer Ebene bzw. unter konstanter Kontaktkraft.

In dieser Arbeit wurde eine Robotersteuerung ohne autonome Anteile vorgestellt. Dabei sollte die Steuerung zur Unterstützung eines autonomen Systems zur automatischen Nadelverfolgung verwendet werden. Somit sollte die Interaktion in Kombination mit den autonomen Anteilen getestet werden, um die Kompatibilität zu überprüfen und gegebenenfalls fehlende Funktionen identifizieren zu können.

Weiterhin könnte es sinnvoll sein, auch weitere Grade der Autonomie in das System zu integrieren. So könnte durch Roboterunterstützung die manuelle Interaktion möglicherweise erleichtert werden, indem einige Freiheitsgrade autonom durch den Roboter gesteuert werden. Dazu könnte ein System, ähnlich zu dem von Fang et al. [79] vorgestellten, entwickelt

werden, sodass die Kontaktkraft durch den Roboter autonom gesteuert wird, während die Bewegung in die anderen Raumrichtungen weiterhin manuell erfolgen können. Dies könnte die Bewegung der Sonde auf Oberflächen ohne haptisches Feedback unterstützen.

Diese Arbeit hat sich lediglich mit der Aufnahme der Ultraschallbilder durch die Positionierung der Sonde befasst. Dabei wurde die Darstellung dieser Bilder vernachlässigt, sodass die aufgenommenen Bilder auf einem fest positionierten Bildschirm dargestellt wurden. Um die Ergonomie während des Eingriffs weiter zu erhöhen, könnten die Bilder mithilfe von Augmented Reality direkt im Sichtfeld positioniert werden. Die Darstellung der Ultraschallbilder könnte dabei beispielsweise in einem festen Abstand zur Ultraschallsonde [144] oder an der Position der visualisierten Strukturen [145] erfolgen. Dazu müssten geeignete Modalitäten und Visualisierungskonzepte identifiziert werden.

Weiterhin könnte es sinnvoll sein, Befehle für die Bedienung von Funktionen der Ultraschallbildgebung zu ergänzen. So könnten beispielsweise Spracheingaben und Gesten für die Aufnahme von Screenshots und die Markierung von Punkten im Ultraschallbild verwendet werden [146].

Um die Eignung der Interaktion mit dem Roboter für die Ultraschallbildgebung während nadelgestützter Interventionen zu überprüfen, müssten zukünftig qualitative Studien in klinischer Umgebung durchgeführt werden. Dabei sollte das System von Radiolog*innen mit Erfahrung in nadelgestützten Interventionen an einer realistischen Interaktionsaufgabe durchgeführt werden. Diese könnten Aufschluss über die klinische Anwendbarkeit und den Einfluss auf bestehende Arbeitsabläufe geben. Zusätzlich könnten dadurch auch weitere benötigte Funktionen identifiziert werden. Im Rahmen einer solchen Studie müsste weiterhin überprüft werden, ob die HoloLens für den Einsatz während der Interventionen als Eingabemodalität geeignet ist.

Literaturverzeichnis

- [1] DARZI, Ara ; MUNZ, Yaron: The impact of minimally invasive surgical techniques. In: *Annu. Rev. Med.* 55 (2004), S. 223–237
- [2] LEVINE, Deborah: Ultrasound versus magnetic resonance imaging in fetal evaluation. In: *Topics in Magnetic resonance imaging* 12 (2001), Nr. 1, S. 25–38
- [3] DODD 3RD, GD ; ESOLA, Christine C. ; MEMEL, David S. ; GHIATAS, Abraham A. ; CHINTAPALLI, Kedar N. ; PAULSON, Erik K. ; NELSON, Rendon C. ; FERRIS, James V. ; BARON, Richard L.: Sonography: the undiscovered jewel of interventional radiology. In: *Radiographics* 16 (1996), Nr. 6, S. 1271–1288
- [4] PRIESTER, Alan M. ; NATARAJAN, Shyam ; CULJAT, Martin O.: Robotic ultrasound systems in medicine. In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 60 (2013), Nr. 3, S. 507–523
- [5] CRAIG, Marveen: Sonography: An occupational health hazard? In: *Journal of Diagnostic Medical Sonography* 1 (1985), Nr. 3, S. 121–126
- [6] CRAIG, Marveen: Occupational hazards of sonography: an update. In: *Journal of Diagnostic Medical Sonography* 6 (1990), Nr. 1, S. 47–50
- [7] SCHREITER, Josefine ; JOERES, Fabian ; MARCH, Christine ; PECH, Maciej ; HANSEN, Christian: Application Potential of Robot-Guided Ultrasound During CT-Guided Interventions. In: *International Workshop on Advances in Simplifying Medical Ultrasound* Springer, 2021, S. 116–125
- [8] GOODRICH, Michael A. ; SCHULTZ, Alan C.: *Human-robot interaction: a survey*. Now Publishers Inc, 2008
- [9] MUKUND, Amar ; BHARDWAJ, Krishna ; MOHAN, Chander: Basic interventional procedures: Practice essentials. In: *Indian Journal of Radiology and Imaging* 29 (2019), Nr. 02, S. 182–189
- [10] BERGER, Hermann: Interventionelle Radiologie. Version: Dezember 2012. <http://dx.doi.org/10.1007/978-3-642-23804-8-15>. In: *Chirurgie Basisweiterbildung*. Springer Berlin Heidelberg, Dezember 2012. – DOI 10.1007/978-3-642-23804-8-15, 133–143

- [11] VOGL, Thomas J.: *Diagnostische und Interventionelle Radiologie*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2011 (SpringerLink Bücher). <http://dx.doi.org/10.1007/978-3-540-87668-7>. <http://dx.doi.org/10.1007/978-3-540-87668-7>. – ISBN 9783540876687. – herausgegeben von Thomas J. Vogl, Wolfgang Reith, Ernst J. Rummeny
- [12] BERG, Julia ; LU, Shuang: Review of interfaces for industrial human-robot interaction. In: *Current Robotics Reports* 1 (2020), Nr. 2, S. 27–34
- [13] AHMED, Muneeb ; BRACE, Christopher L. ; LEE JR, Fred T. ; GOLDBERG, S N.: Principles of and advances in percutaneous ablation. In: *Radiology* 258 (2011), Nr. 2, S. 351–369
- [14] SOLOMON, Stephen B. ; SILVERMAN, Stuart G.: Imaging in interventional oncology. In: *Radiology* 257 (2010), Nr. 3, S. 624–640
- [15] GOLDBERG, S N. ; CHARBONEAU, J W. ; DODD III, Gerald D. ; DUPUY, Damian E. ; GERVAIS, Debra A. ; GILLAMS, Alice R. ; KANE, Robert A. ; LEE JR, Fred T. ; LIVRAGHI, Tito ; MCGAHAN, John P. u. a.: Image-guided tumor ablation: proposal for standardization of terms and reporting criteria. In: *Radiology* 228 (2003), Nr. 2, S. 335–345
- [16] SILVERMAN, Stuart G. ; TUNCALI, Kemal ; ADAMS, Douglass F. ; NAWFEL, Richard D. ; ZOU, Kelly H. ; JUDY, Philip F.: CT fluoroscopy-guided abdominal interventions: techniques, results, and radiation exposure. In: *Radiology* 212 (1999), Nr. 3, S. 673–681
- [17] DE MEY, Jan ; DE BEECK, B O. ; MEYSMAN, M ; NOPPEN, Marc ; DE MAESENEER, Michel ; VANHOEY, M ; VINCKEN, Walter ; OSTEАUX, Michel: Real time CT-fluoroscopy: diagnostic and therapeutic applications. In: *European Journal of Radiology* 34 (2000), Nr. 1, S. 32–40
- [18] CORNELIS, Francois H. ; RAZAKAMANANTSOA, Leo ; AMMAR, Mohamed B. ; LEHRER, Raphael ; HAFFAF, Idriss ; EL-MOUHADI, Sanaa ; GARDАVAUD, Francois ; NAJDАWI, Milan ; BARRAL, Matthias: Ergonomics in interventional radiology: Awareness is mandatory. In: *Medicina* 57 (2021), Nr. 5, S. 500
- [19] GAZELLE, G S. ; HAAGA, JR: Guided percutaneous biopsy of intraabdominal lesions. In: *American Journal of Roentgenology* 153 (1989), Nr. 5, S. 929–935
- [20] MCGAHAN, John P. ; DODD III, Gerald D.: Radiofrequency ablation of the liver: current status. In: *American Journal of Roentgenology* 176 (2001), Nr. 1, S. 3–16

- [21] WOOD, Bradford J. ; LOCKLIN, Julia K. ; VISWANATHAN, Anand ; KRUECKER, Jochen ; HAEMMERICH, Dieter ; CEBRAL, Juan ; SOFER, Ariela ; CHENG, Ruida ; MCCREEDY, Evan ; CLEARY, Kevin u. a.: Technologies for guidance of radiofrequency ablation in the multimodality interventional suite of the future. In: *Journal of vascular and interventional radiology* 18 (2007), Nr. 1, S. 9–24
- [22] HOWE, Robert D. ; MATSUOKA, Yoky: Robotics for surgery. In: *Annual review of biomedical engineering* 1 (1999), Nr. 1, S. 211–240
- [23] TAYLOR, Russell H. ; MENCIASSI, Arianna ; FICHTINGER, Gabor ; FIORINI, Paolo ; DARIO, Paolo: Medical robotics and computer-integrated surgery. In: *Springer handbook of robotics*. Springer, 2016, S. 1657–1684
- [24] LI, Keyu ; XU, Yangxin ; MENG, Max Q-H: An overview of systems and techniques for autonomous robotic ultrasound acquisitions. In: *IEEE Transactions on Medical Robotics and Bionics* (2021)
- [25] HAXTHAUSEN, Felix von ; BÖTTGER, Sven ; WULFF, Daniel ; HAGENAH, Jannis ; GARCÍA-VÁZQUEZ, Verónica ; IPSEN, Svenja: Medical robotics for ultrasound imaging: current systems and future trends. In: *Current Robotics Reports* (2021), S. 1–17
- [26] YANG, Guang-Zhong ; CAMBIAS, James ; CLEARY, Kevin ; DAIMLER, Eric ; DRAKE, James ; DUPONT, Pierre E. ; HATA, Nobuhiko ; KAZANZIDES, Peter ; MARTEL, Sylvain ; PATEL, Rajni V. u. a.: *Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy*. 2017
- [27] SHERIDAN, Thomas B.: *Telerobotics, automation, and human supervisory control*. MIT press, 1992
- [28] WEBER, AS: Emerging medical ethical issues in healthcare and medical robotics. In: *Int J Mech Eng Rob Res* 7 (2018), Nr. 6, S. 604–607
- [29] CRAIG, John J.: *Introduction to robotics: mechanics and control*. Pearson Educacion, 2005
- [30] PRIESTER, Alan M. ; NATARAJAN, Shyam ; CULJAT, Martin O.: Robotic ultrasound systems in medicine. In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 60 (2013), Nr. 3, S. 507–523
- [31] SWERDLOW, Daniel R. ; CLEARY, Kevin ; WILSON, Emmanuel ; AZIZI-KOUTENAEI, Bamshad ; MONFAREDI, Reza: Robotic arm–assisted sonography: Review of technical developments and potential clinical applications. In: *American Journal of Roentgenology* 208 (2017), Nr. 4, S. 733–738

- [32] TROCCAZ, Jocelyne ; PESHKIN, Michael ; DAVIES, Brian: The use of localizers, robots and synergistic devices in CAS. In: *CVRMed-MRCAS'97* Springer, 1997, S. 725–736
- [33] HO, SC ; HIBBERD, RD ; DAVIES, BL: Robot assisted knee surgery. In: *IEEE Engineering in Medicine and Biology Magazine* 14 (1995), Nr. 3, S. 292–300
- [34] KOENIG, Nathan ; TAKAYAMA, Leila ; MATARIĆ, Maja: Communication and knowledge sharing in human–robot interaction and learning from demonstration. In: *Neural Networks* 23 (2010), Nr. 8-9, S. 1104–1112
- [35] BIGGS, Geoffrey ; MACDONALD, Bruce: A survey of robot programming systems. In: *Proceedings of the Australasian conference on robotics and automation* Bd. 1, 2003, S. 1–3
- [36] SCHLEICHER, Tim: *Kollaborierende Roboter anweisen*. Springer Fachmedien Wiesbaden, 2020. <http://dx.doi.org/10.1007/978-3-658-29051-1>. <http://dx.doi.org/10.1007/978-3-658-29051-1>
- [37] NAGRATH, Mittal: *Robotics and Control* -. New Delhi : Tata McGraw-Hill Education, 2003. – ISBN 978–0–070–48293–7
- [38] SICILIANO, Bruno ; SCIAVICCO, Lorenzo ; VILLANI, Luigi ; ORIOLO, Giuseppe: *Robotics - Modelling, Planning and Control*. Berlin Heidelberg : Springer Science & Business Media, 2010. – ISBN 978–1–846–28641–4
- [39] ZIMMERMANN, Markus ; BORTOT, Dino ; BENGLER, Klaus: Allgemeine Interaktionsprinzipien für kooperative Mensch-Maschine-Systeme. In: *Gestaltung nachhaltiger Arbeitssysteme: 58. Kongress der Gesellschaft für Arbeitswissenschaft*, 2012
- [40] PAVLOVIC, V.I. ; SHARMA, R. ; HUANG, T.S.: Visual interpretation of hand gestures for human-computer interaction: a review. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (1997), Nr. 7, S. 677–695. <http://dx.doi.org/10.1109/34.598226>. – DOI 10.1109/34.598226
- [41] WANG, Lihui ; LIU, Sichao ; LIU, Hongyi ; WANG, Xi V.: Overview of human-robot collaboration in manufacturing. In: *Proceedings of 5th international conference on the industry 4.0 model for advanced manufacturing* Springer, 2020, S. 15–58
- [42] COLECA, Foti ; MARTINETZ, Thomas ; BARTH, Erhardt: Gesture interfaces with depth sensors. In: *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications*. Springer, 2013, S. 207–227

- [43] PREIM, Bernhard ; DACHSELT, Raimund: *Interaktive Systeme - Band 2: User Interface Engineering, 3D-Interaktion, Natural User Interfaces*. 2015. <http://dx.doi.org/10.1007/978-3-642-45247-5>. <http://dx.doi.org/10.1007/978-3-642-45247-5>. – ISBN 978-3-642-45246-8
- [44] ZHAI, Shumin ; MILGRAM, Paul: Human performance evaluation of manipulation schemes in virtual environments. In: *Proceedings of IEEE Virtual Reality Annual International Symposium IEEE*, 1993, S. 155–161
- [45] MENDES, D. ; CAPUTO, F. M. ; GIACHETTI, A. ; FERREIRA, A. ; JORGE, J.: A Survey on 3D Virtual Object Manipulation: From the Desktop to Immersive Virtual Environments. In: *Computer Graphics Forum* 38 (2019), Nr. 1, 21–45. <http://dx.doi.org/https://doi.org/10.1111/cgf.13390>. – DOI <https://doi.org/10.1111/cgf.13390>
- [46] MIHELJ, Matjaž ; BAJD, Tadej ; UDE, Aleš ; LENARČIČ, Jadran ; STANOVNIK, Aleš ; MUNIH, Marko ; REJC, Jure ; ŠLAJPAH, Sebastjan: Collaborative robots. In: *Robotics*. Springer, 2019, S. 173–187
- [47] KARPOV, AA ; YUSUPOV, RM: Multimodal interfaces of human–computer interaction. In: *Herald of the Russian Academy of Sciences* 88 (2018), Nr. 1, S. 67–74
- [48] HARTMANN, Bernd ; BENSON, Matthias ; JUNGER, Axel ; QUINZIO, Lorenzo ; RÖHRIG, Rainer ; FENGLER, Bernhard ; FÄRBER, Udo W. ; WILLE, Burkhard ; HEMPELMANN, Gunter: Computer keyboard and mouse as a reservoir of pathogens in an intensive care unit. In: *Journal of clinical monitoring and computing* 18 (2004), Nr. 1, S. 7–12
- [49] VILLANI, Valeria ; PINI, Fabio ; LEALI, Francesco ; SECCHI, Cristian ; FANTUZZI, Cesare: Survey on Human-Robot Interaction for Robot Programming in Industrial Applications. 51 (2018), Nr. 11, 66–71. <http://dx.doi.org/10.1016/j.ifacol.2018.08.236>. – DOI 10.1016/j.ifacol.2018.08.236
- [50] MAURTUA, Inaki ; FERNANDEZ, Izaskun ; TELLAEACHE, Alberto ; KILDAL, Johan ; SUSPERREGI, Loreto ; IBARGUREN, Aitor ; SIERRA, Basilio: Natural multimodal communication for human–robot collaboration. In: *International Journal of Advanced Robotic Systems* 14 (2017), Nr. 4, S. 1729881417716043
- [51] AKAN, Batu ; AMERI, Afshin ; CÜRÜKLÜ, Baran ; ASPLUND, Lars: Intuitive industrial robot programming through incremental multimodal language and augmented reality. In: *2011 IEEE International Conference on Robotics and Automation IEEE*, 2011, S. 3934–3939

- [52] ELEPFANDT, Monika ; SÜNDERHAUF, Marcelina: Multimodal, touchless interaction in spatial augmented reality environments. In: *International Conference on Digital Human Modeling* Springer, 2011, S. 263–271
- [53] PREIM, Bernhard ; DACHSELT, Raimund: *Interaktive Systeme*. Springer Berlin Heidelberg, 2015. <http://dx.doi.org/10.1007/978-3-642-45247-5>. <http://dx.doi.org/10.1007/978-3-642-45247-5>
- [54] IBA, Soshi ; PAREDIS, Christiaan J. J. ; KHOSLA, Pradeep K.: Interactive Multimodal Robot Programming. 24 (2005), Januar, Nr. 1, 83–104. <http://dx.doi.org/10.1177/0278364904049250>. – DOI 10.1177/0278364904049250
- [55] LIU, Hongyi ; WANG, Lihui: Gesture recognition for human-robot collaboration: A review. In: *International Journal of Industrial Ergonomics* 68 (2018), S. 355–367
- [56] MITRA, Sushmita ; ACHARYA, Tinku: Gesture recognition: A survey. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37 (2007), Nr. 3, S. 311–324
- [57] ISTANCE, Howell ; BATES, Richard ; HYRSKYKARI, Aulikki ; VICKERS, Stephen: Snap clutch, a moded approach to solving the Midas touch problem. In: *Proceedings of the 2008 symposium on Eye tracking research & applications*, 2008, S. 221–228
- [58] KRUPKE, Dennis ; STEINICKE, Frank ; LUBOS, Paul ; JONETZKO, Yannick ; GÖRNER, Michael ; ZHANG, Jianwei: Comparison of multimodal heading and pointing gestures for co-located mixed reality human-robot interaction. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* IEEE, 2018, S. 1–9
- [59] WAGNER, Petra ; MALISZ, Zofia ; KOPP, Stefan: *Gesture and speech in interaction: An overview*. 2014
- [60] HUANG, Yanpei ; BURDET, Etienne ; CAO, Lin ; PHAN, Phuoc T. ; TIONG, Anthony Meng H. ; PHEE, Soo J.: A Subject-Specific Four-Degree-of-Freedom Foot Interface to Control a Surgical Robot. 25 (2020), April, Nr. 2, 951–963. <http://dx.doi.org/10.1109/tmech.2020.2964295>. – DOI 10.1109/tmech.2020.2964295
- [61] GASCHLER, Andre ; SPRINGER, Maximilian ; RICKERT, Markus ; KNOLL, Alois: Intuitive robot tasks with augmented reality and virtual obstacles. In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, S. 6026–6031
- [62] HO, Tzu-Hsuan ; SONG, Kai-Tai: Supervised Control for Robot-Assisted Surgery Using Augmented Reality. In: *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, 2020, S. 329–334

- [63] GRADMANN, Michael ; ORENDT, Eric M. ; SCHMIDT, Edgar ; SCHWEIZER, Stephan ; HENRICH, Dominik: Augmented reality robot operation interface with Google Tango. In: *ISR 2018; 50th International Symposium on Robotics VDE*, 2018, S. 1–8
- [64] CAI, Yi ; WANG, Yi ; BURNETT, Morice: Using augmented reality to build digital twin for reconfigurable additive manufacturing system. In: *Journal of Manufacturing Systems* 56 (2020), S. 598–604
- [65] TSAI, Roger Y. ; LENZ, Reimar K. u. a.: A new technique for fully autonomous and efficient 3 d robotics hand/eye calibration. In: *IEEE Transactions on robotics and automation* 5 (1989), Nr. 3, S. 345–358
- [66] BISCHOFF, Rainer ; KAZI, Arif: Perspectives on augmented reality based human-robot interaction with industrial robots. In: *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)* Bd. 4 IEEE, 2004, S. 3226–3231
- [67] SATOH, Kiyohide ; TAKEMOTO, Kazuki ; UCHIYAMA, Shinji ; YAMAMOTO, Hiroyuki: A registration evaluation system using an industrial robot. In: *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2006, S. 79–87
- [68] MATHIASSEN, Kim ; FJELLIN, Jørgen E. ; GLETTE, Kyrre ; HOL, Per K. ; ELLE, Ole J.: An ultrasound robotic system using the commercial robot UR5. In: *Frontiers in Robotics and AI* 3 (2016), S. 1
- [69] HUANG, Qinghua ; LAN, Jiulong: Remote control of a robotic prosthesis arm with six-degree-of-freedom for ultrasonic scanning and three-dimensional imaging. In: *Biomedical Signal Processing and Control* 54 (2019), S. 101606
- [70] DE CUNHA, D ; GRAVEZ, P ; LEROY, C ; MAILLARD, E ; JOUAN, J ; VARLEY, P ; JONES, M ; HALLIWELL, M ; HAWKES, D ; WELLS, PNT u. a.: The MIDSTEP system for ultrasound guided remote telesurgery. In: *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)* Bd. 3 IEEE, 1998, S. 1266–1269
- [71] GUAN, Xiaolong ; WU, Haotian ; HOU, Xilong ; TENG, Qing ; WEI, Shiyu ; JIANG, Tao ; ZHANG, Jiantao ; WANG, Ban ; YANG, Jialin ; XIONG, Linfei: Study of a 6DOF robot assisted ultrasound scanning system and its simulated control handle. In: *2017 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, 2017, S. 469–474

- [72] COURREGES, Fabien ; VIEYRES, Pierre ; ISTEPANIAN, RSH: Advances in robotic tele-echography services-the OTELO system. In: *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* Bd. 2 IEEE, 2004, S. 5371–5374
- [73] GONZALES, Adriana V. ; CINQUIN, Philippe ; TROCCAZ, Jocelyne ; GUERRAZ, Agnès ; HENNION, Bernard ; PELLISSIER, Franck ; THOREL, Pierre ; COURREGES, Fabien ; GOURDON, Alain ; POISSON, Gérard u. a.: TER: a system for robotic tele-echography. In: *Lecture Notes in Computer Science* (2001), S. 326–334
- [74] ZANDSTEEG, CJ ; BRUIJNEN, DJH ; MOLENGRAFT, MJG Van d.: Haptic tele-operation system control design for the ultrasound task: A loop-shaping approach. In: *Mechatronics* 20 (2010), Nr. 7, S. 767–777
- [75] JANVIER, Marie-Ange ; DURAND, Louis-Gilles ; CARDINAL, Marie-Hélène Roy ; RENAUD, Isabelle ; CHAYER, Boris ; BIGRAS, Pascal ; DE GUISE, Jacques ; SOULEZ, Gilles ; CLOUTIER, Guy: Performance evaluation of a medical robotic 3D-ultrasound imaging system. In: *Medical image analysis* 12 (2008), Nr. 3, S. 275–290
- [76] SALCUDEAN, Septimiu E. ; BELL, Gordon ; BACHMANN, Simon ; ZHU, Wen-Hong ; ABOLMAESUMI, Purang ; LAWRENCE, Peter D.: Robot-assisted diagnostic ultrasound—design and feasibility experiments. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention* Springer, 1999, S. 1062–1071
- [77] SALCUDEAN, Septimiu E. ; ZHU, Wen H. ; ABOLMAESUMI, P ; BACHMANN, Simon ; LAWRENCE, Peter D.: A robot system for medical ultrasound. In: *Robotics Research*. Springer, 2000, S. 195–202
- [78] ABOLMAESUMI, Purang ; SALCUDEAN, Septimiu E. ; ZHU, Wen-Hong ; SIROUSPOUR, Mohammad R. ; DiMAIO, Simon P.: Image-guided control of a robot for medical ultrasound. In: *IEEE Transactions on Robotics and Automation* 18 (2002), Nr. 1, S. 11–23
- [79] FANG, Ting-Yun ; ZHANG, Haichong K. ; FINOCCHI, Rodolfo ; TAYLOR, Russell H. ; BOCTOR, Emad M.: Force-assisted ultrasound imaging system through dual force sensing and admittance robot control. In: *International journal of computer assisted radiology and surgery* 12 (2017), Nr. 6, S. 983–991
- [80] KASCHWICH, M. ; HAXTHAUSEN, F. von ; AUST, T. ; ERNST, F. ; KLEEMANN, M.: Roboterbasierte Ultraschallsteuerung. In: *Gefäßchirurgie* 25 (2020), August, Nr. 5, 345–351. <http://dx.doi.org/10.1007/s00772-020-00670-z>. – DOI 10.1007/s00772-020-00670-z

- [81] CHATELAIN, Pierre ; KRUPA, Alexandre ; MARCHAL, Maud: Real-time needle detection and tracking using a visually servoed 3D ultrasound probe. In: *2013 IEEE International Conference on Robotics and Automation* IEEE, 2013, S. 1676–1681
- [82] LANGSCH, Fernanda ; VIRGA, Salvatore ; ESTEBAN, Javier ; GÖBL, Rüdiger ; NAVAB, Nassir: Robotic ultrasound for catheter navigation in endovascular procedures. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* IEEE, 2019, S. 5404–5410
- [83] MEBARKI, Rafik ; KRUPA, Alexandre ; CHAUMETTE, François: 2-D ultrasound probe complete guidance by visual servoing using image moments. In: *IEEE Transactions on Robotics* 26 (2010), Nr. 2, S. 296–306
- [84] FERRE, M ; MACIAS-GUARASA, J ; ARACIL, R ; BARRIENTOS, A: Voice command generation for teleoperated robot systems. In: *7th IEEE International Workshop on Robot and Human Communication 1998 (RO-MAN'98)* Bd. 2, 1998, S. 679–685
- [85] LAMBRECHT, Jens ; KLEINSORGE, Martin ; KRÜGER, Jörg: Markerless gesture-based motion control and programming of industrial robots. In: *ETFA2011* IEEE, 2011, S. 1–4
- [86] GÎRBACIA, Florin ; POSTELNICU, Cristian ; VOINEA, Gheorghe-Daniel: Towards using natural user interfaces for robotic arm manipulation. In: *International Conference on Robotics in Alpe-Adria Danube Region* Springer, 2019, S. 188–193
- [87] BASSILY, D ; GEORGOULAS, C ; GUETTLER, J ; LINNER, Thomas ; BOCK, T: Intuitive and adaptive robotic arm manipulation using the leap motion controller. In: *ISR/Robotik 2014; 41st International Symposium on Robotics* VDE, 2014, S. 1–7
- [88] PEPPOLONI, Lorenzo ; BRIZZI, Filippo ; AVIZZANO, Carlo A. ; RUFFALDI, Emanuele: Immersive ROS-integrated framework for robot teleoperation. In: *2015 IEEE Symposium on 3D User Interfaces (3DUI)* IEEE, 2015, S. 177–178
- [89] QUINTERO, Camilo P. ; TATSAMBON, Romeo ; GRIDSETH, Mona ; JÄGERSAND, Martin: Visual pointing gestures for bi-directional human robot interaction in a pick-and-place task. In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* IEEE, 2015, S. 349–354
- [90] HÜGLE, Johannes ; LAMBRECHT, Jens ; KRÜGER, Jörg: An integrated approach for industrial robot control and programming combining haptic and non-haptic gestures. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* IEEE, 2017, S. 851–857

- [91] BOLT, RA: Pu-That There: Voice and Gesture at the Graphics Interface. In: *Computer Graphics* 14 (1980), S. 3
- [92] PERZANOWSKI, Dennis ; SCHULTZ, Alan C. ; ADAMS, William ; MARSH, Elaine ; BUGAJSKA, Magda: Building a multimodal human-robot interface. In: *IEEE intelligent systems* 16 (2001), Nr. 1, S. 16–21
- [93] ROGALLA, Oliver ; EHRENMANN, Markus ; ZOLLNER, R ; BECHER, Regine ; DILLMANN, Rüdiger: Using gesture and speech control for commanding a robot assistant. In: *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication* IEEE, 2002, S. 454–459
- [94] DU, Guanglong ; CHEN, Mingxuan ; LIU, Caibing ; ZHANG, Bo ; ZHANG, Ping: Online robot teaching with natural human–robot interaction. In: *IEEE Transactions on Industrial Electronics* 65 (2018), Nr. 12, S. 9571–9581
- [95] STIEFELHAGEN, Rainer ; FUGEN, Christian ; GIESELMANN, R ; HOLZAPFEL, Hartwig ; NICKEL, Kai ; WAIBEL, Alex: Natural human-robot interaction using speech, head pose and gestures. In: *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)* Bd. 3 IEEE, 2004, S. 2422–2427
- [96] TOSTADO, Pablo M. ; ABBOTT, William W. ; FAISAL, A A.: 3D gaze cursor: Continuous calibration and end-point grasp control of robotic actuators. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)* IEEE, 2016, S. 3295–3300
- [97] WÖHLE, Lukas ; GEBHARD, Marion: Towards robust robot control in cartesian space using an infrastructureless head-and eye-gaze interface. In: *Sensors* 21 (2021), Nr. 5, S. 1798
- [98] WILLIAMS, Matthew R. ; KIRSCH, Robert F.: Evaluation of head orientation and neck muscle EMG signals as three-dimensional command sources. In: *Journal of neuroengineering and rehabilitation* 12 (2015), Nr. 1, S. 1–16
- [99] RUDIGKEIT, Nina ; GEBHARD, Marion: AMiCUS—A head motion-based interface for control of an assistive robot. In: *Sensors* 19 (2019), Nr. 12, S. 2836
- [100] RUDIGKEIT, Nina ; GEBHARD, Marion ; GRÄSER, Axel: Towards a user-friendly AHRs-based human-machine interface for a semi-autonomous robot. In: *Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2014), Chicago, IL, USA, 2014*, S. 14–18

- [101] JACKOWSKI, Anja ; GEBHARD, Marion ; GRÄSER, Axel: A novel head gesture based interface for hands-free control of a robot. In: *2016 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* IEEE, 2016, S. 1–6
- [102] ALSHARIF, Shiva: *Gaze-Based Control of Robot Arm in Three-Dimensional Space*, Universität Bremen, Diss., 2018
- [103] PARK, Kyeong-Beom ; CHOI, Sung H. ; LEE, Jae Y. ; GHASEMI, Yalda ; MOHAMMED, Mustafa ; JEONG, Heejin: Hands-free human–robot interaction using multimodal gestures and deep learning in wearable mixed reality. In: *IEEE Access* 9 (2021), S. 55448–55464
- [104] KAWAI, Toshikazu ; FUKUNISHI, Masanori ; NISHIKAWA, Atsushi ; NISHIZAWA, Yuji ; NAKAMURA, Tatsuo: Hands-free interface for surgical procedures based on foot movement patterns. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* IEEE, 2014, S. 345–348
- [105] RUDIGKEIT, Nina ; GEBHARD, Marion ; GRÄSER, Axel: Evaluation of control modes for head motion-based control with motion sensors. In: *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings* IEEE, 2015, S. 135–140
- [106] *DIN EN ISO 9241-11:2018-11, Ergonomie der Mensch-System-Interaktion - Teil - 11: Gebrauchstauglichkeit: Begriffe und Konzepte (ISO_9241-11:2018), Deutsche Fassung EN_ISO_9241-11:2018*. <http://dx.doi.org/10.31030/2757945>
- [107] NORMAN, Donald A.: Natural user interfaces are not natural. In: *interactions* 17 (2010), Nr. 3, S. 6–10
- [108] SIBERT, Linda E. ; JACOB, Robert J.: Evaluation of eye gaze interaction. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 2000, S. 281–288
- [109] WARE, C: An Evaluation of an Eye Tracker as a Device for Computer Input; in Human Factors in Computing Systems. In: *CHI+ GI'87 Conference Proceedings*, 1987
- [110] GRASSO, Michael A. ; EBERT, David S. ; FININ, Timothy W.: The integrality of speech in multimodal interfaces. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 5 (1998), Nr. 4, S. 303–325
- [111] VELLOSO, Eduardo ; SCHMIDT, Dominik ; ALEXANDER, Jason ; GELLERSEN, Hans ; BULLING, Andreas: The feet in human–computer interaction: A survey of foot-based interaction. In: *ACM Computing Surveys (CSUR)* 48 (2015), Nr. 2, S. 1–35

- [112] DÖRNER, Ralf (Hrsg.) ; BROLL, Wolfgang (Hrsg.) ; GRIMM, Paul (Hrsg.) ; JUNG, Bernhard (Hrsg.): *Virtual und Augmented Reality (VR / AR)*. Springer Berlin Heidelberg, 2013. <http://dx.doi.org/10.1007/978-3-642-28903-3>. <http://dx.doi.org/10.1007/978-3-642-28903-3>
- [113] SUTHERLAND, Justin ; BELEC, Jason ; SHEIKH, Adnan ; CHEPELEV, Leonid ; ALTHOBAITY, Waleed ; CHOW, Benjamin J. ; MITSOURAS, Dimitrios ; CHRISTENSEN, Andy ; RYBICKI, Frank J. ; LA RUSSA, Daniel J.: Applying modern virtual and augmented reality technologies to medical images and models. In: *Journal of digital imaging* 32 (2019), Nr. 1, S. 38–53
- [114] UNGUREANU, Dorin ; BOGO, Federica ; GALLIANI, Silvano ; SAMA, Pooja ; DUAN, Xin ; MEEKHOF, Casey ; STÜHMER, Jan ; CASHMAN, Thomas J. ; TEKIN, Bugra ; SCHÖNBERGER, Johannes L. u. a.: Hololens 2 research mode as a tool for computer vision research. In: *arXiv preprint arXiv:2008.11239* (2020)
- [115] MICROSOFT: *HoloLens 2-Hardware*. <https://docs.microsoft.com/de-de/hololens/hololens2-hardware>, . – Accessed: 19.02.2022
- [116] MUTASIM, Aunnoy K. ; BATMAZ, Anil U. ; STUERZLINGER, Wolfgang: Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In: *ACM Symposium on Eye Tracking Research and Applications*, 2021, S. 1–7
- [117] GRAUMAN, Kristen ; BETKE, Margrit ; LOMBARDI, Jonathan ; GIPS, James ; BRADSKI, Gary R.: Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces. In: *Universal Access in the Information Society* 2 (2003), Nr. 4, S. 359–373
- [118] MAJARANTA, Päivi ; RÄIHÄ, Kari-Jouko ; HYRSKYKARI, Aulikki ; ŠPAKOV, Oleg: Eye movements and human-computer interaction. In: *Eye Movement Research*. Springer, 2019, S. 971–1015
- [119] CHOE, Mungyeong ; CHOI, Yeongcheol ; PARK, Jaehyun ; KIM, Hyun K.: Comparison of gaze cursor input methods for virtual reality devices. In: *International Journal of Human-Computer Interaction* 35 (2019), Nr. 7, S. 620–629
- [120] KĚPUSKA, Veton ; BOHOUTA, Gamal: Comparing speech recognition systems (Microsoft API, Google API and CMU Sphinx). In: *Int. J. Eng. Res. Appl* 7 (2017), Nr. 03, S. 20–24
- [121] GOOGLE: *Google Speech API*. <https://cloud.google.com/speech-to-text/docs/reference/rest>, . – Accessed: 20.02.2022

- [122] MASLIAH, Maurice R. ; MILGRAM, Paul: Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2000, S. 25–32
- [123] MENDES, Daniel ; RELVAS, Filipe ; FERREIRA, Alfredo ; JORGE, Joaquim: The benefits of dof separation in mid-air 3d object manipulation. In: *Proceedings of the 22nd ACM conference on virtual reality software and technology*, 2016, S. 261–268
- [124] ZHANG, L.: *Investigation of Coupling Patterns of the Cervical Spine*, University of Dundee, Dundee, UK, Diplomarbeit, 2014
- [125] TIJMSMA, Hylke A. ; LIEFHEDBER, Freek ; HERDER, Just L.: Evaluation of new user interface features for the manus robot arm. In: *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*. IEEE, 2005, S. 258–263
- [126] Secondary Task Techniques. Version: 2012. <http://dx.doi.org/10.1007/978-1-4419-1428-6-2391>. In: *Encyclopedia of the Sciences of Learning*. Springer US, 2012. – DOI 10.1007/978-1-4419-1428-6-2391, 2983–2983
- [127] STOJMEANOVA, Kristina ; SODNIK, Jaka: Detection-response task—uses and limitations. In: *Sensors* 18 (2018), Nr. 2, S. 594
- [128] UNITY TECHNOLOGIES: *Unity*. <https://unity3d.com/get-unity/download/archive>. Version: 2020.3.17f. – Accessed: 20.03.2022
- [129] MICROSOFT: *Visual Studio*. <https://visualstudio.microsoft.com/de/vs/older-downloads/>. Version: 2019. – Accessed: 20.03.2022
- [130] MICROSOFT: *Mixed Reality Toolkit*. <https://www.microsoft.com/en-us/download/details.aspx?id=102778>. Version: 2.5.1. – Accessed: 20.03.2022
- [131] PTC: *Vuforia Engine*. <https://developer.vuforia.com/downloads/sdk>. Version: 8.6.7. – Accessed: 20.03.2022
- [132] KUKA: *Sunrise.OS*. <https://www.kuka.com/de-de/produkte-leistungen/robotersysteme/software/systemsoftware/sunriseos>. Version: 1.16. – Accessed: 20.03.2022
- [133] BLANCO, Jose-Luis: A tutorial on SE(3) transformation parameterizations and on-manifold optimization. In: *University of Malaga, Tech. Rep 3* (2010), S. 6
- [134] CIRS TISSUE SIMULATION & PHANTOM TECHNOLOGY: *RIPLE MODALITY 3D ABDOMINAL PHANTOM, MODEL 057A*. <https://www.cirsinc.com/products/ultrasound/zerdine-hydrogel/triple-modality-3d-abdominal-phantom/>. – Accessed: 20.05.2022

- [135] MANEAS, Efthymios ; XIA, Wenfeng ; NIKITICHEV, Daniil I. ; DAHER, Batol ; MANIMARAN, Maniragav ; WONG, Rui Yen J. ; CHANG, Chia-Wei ; RAHMANI, Benyamin ; CAPELLI, Claudio ; SCHIEVANO, Silvia u. a.: Anatomically realistic ultrasound phantoms using gel wax with 3D printed moulds. In: *Physics in Medicine & Biology* 63 (2018), Nr. 1, S. 015033
- [136] CULJAT, Martin O. ; GOLDENBERG, David ; TEWARI, Priyamvada ; SINGH, Rahul S.: A review of tissue substitutes for ultrasound imaging. In: *Ultrasound in medicine & biology* 36 (2010), Nr. 6, S. 861–873
- [137] RIPPEY, James C. ; BLANCO, Pablo ; CARR, Peter J.: An affordable and easily constructed model for training in ultrasound-guided vascular access. In: *The Journal of Vascular Access* 16 (2015), Nr. 5, S. 422–427
- [138] GRICE, JV ; PICKENS, DR ; PRICE, RR: A new phantom design for routine testing of Doppler ultrasound. In: *Medical Physics* 43 (2016), Nr. 7, S. 4431–4434
- [139] ENGEL, Katja: *Entwicklung eines MRT-kompatiblen und beweglichen Abdomenphantoms zur Durchführung von Punktionsversuchen unter simulierten Atemzyklen*. 4 2022
- [140] HART, Sandra G. ; STAVELAND, Lowell E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: *Advances in psychology* Bd. 52. Elsevier, 1988, S. 139–183
- [141] VAN SOMEREN, Maarten ; BARNARD, Yvonne F. ; SANDBERG, J: The think aloud method: a practical approach to modelling cognitive. In: *London: Academic Press* 11 (1994)
- [142] BROOKE, John: Sus: a “quick and dirty” usability. In: *Usability evaluation in industry* 189 (1996), Nr. 3
- [143] SAURO, Jeff: *A practical guide to the system usability scale: Background, benchmarks & best practices*. Measuring Usability LLC, 2011
- [144] HAXTHAUSEN, Felix von ; BÖTTGER, Sven ; KLEEMANN, Markus ; ERNST, Floris ; SCHWEIKARD, Achim: Robotics from the bench: research for ultrasound automation with augmented reality visualization. In: *Proceedings on Minimally Invasive Surgery* 1 (2019), Nr. 1, S. 219–219
- [145] SAUER, Frank ; KHAMENE, Ali ; BASCLE, Benedicte ; SCHINUNANG, L ; WENZEL, Fabian ; VOGT, Sebastian: Augmented reality visualization of ultrasound images: system description, calibration, and features. In: *Proceedings IEEE and ACM International Symposium on Augmented Reality* IEEE, 2001, S. 30–39

- [146] HERNICZEK, Simon K. ; LASSO, Andras ; UNGI, Tamas ; FICHTINGER, Gabor:
Feasibility of a touch-free user interface for ultrasound snapshot-guided nephrostomy.
In: *Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and
Modeling* Bd. 9036 SPIE, 2014, S. 628–633

A Anhang

A.1 Fragebögen

A.1.1 Fragebogen Experteninterview

Workflow: Generell zu onkologisch-therapeutische Interventionen unter kombinierter Bildgebung (CT & Ultraschall)

1. Wie würde die Intervention aussehen?
 - a) Kurzer Ablauf/ Überblick über die wichtigsten Schritte
 - b) Potentieller Use-Case für die Intervention
2. Wo liegen die größten Herausforderungen der Intervention?
 - a) Kognitiv
 - b) Motorisch

Ausblick: Robotergestützter Ultraschall

Angenommen Sie könnten bei der Intervention durch einen Roboter, der die Ultraschallsonde hält unterstützt werden:

3. Welche Anforderungen würde es an den Roboter geben?
 - a) Welche Funktionen sollten automatisch möglich sein?
 - b) Was sollte manuell erfolgen können?
4. Welche Eingabemethoden könnten Sie sich dafür vorstellen?

A.1.2 Nasa TLX

Geben Sie jetzt für jede der unten stehenden Dimensionen an, wie hoch die Beanspruchung der jeweiligen Aufgabe war. Markieren Sie dazu bitte auf den folgenden Skalen, in welchem Maße Sie sich in den sechs genannten Dimensionen von der Aufgabe beansprucht oder gefordert gesehen haben.

Geistige Anforderungen

Wie viel geistige Anstrengung war bei der Informationsaufnahme und -verarbeitung notwendig? (z.B. Entscheiden, Denken, Erinnern, Hinsehen, Suchen...) War die Aufgabe leicht oder anspruchsvoll, einfach oder komplex?



Körperliche Anforderungen

Wie hoch war der Grad der physischen Aktivität? War die Aufgabe leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?



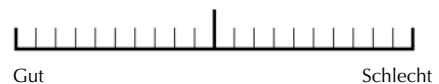
Zeitliche Anforderungen

Wie viel Zeitdruck empfanden Sie hinsichtlich der Häufigkeit oder dem Takt, mit dem Aufgaben auftraten? War die Abfolge langsam und geruhsam oder schnell und hektisch?



Leistung

Wie erfolgreich haben Sie ihrer Meinung nach die vom Untersuchungsleiter (oder Ihnen selbst) gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung bei der Verfolgung dieser Ziele?



Anstrengung

Wie hart mussten Sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen?



Frustration

Wie unsicher, entmutigt, irritiert, gestresst oder verärgert fühlten Sie sich während der Aufgabe?



A.1.3 System Usability Scale

Geben Sie bitte für jede der unten stehenden Aussagen an, wie sehr sie den jeweiligen Aussagen zustimmen, bzw. wie sehr sie diese Aussagen ablehnen.

	Starke Ablehnung				Starke Zustimmung
1. Ich glaube, ich würde dieses System häufig benutzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Ich finde das System unnötig komplex.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Ich finde das System einfach zu benutzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Ich glaube, dass ich technische Hilfe bräuchte, um das System zu benutzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Ich finde, dass die verschiedenen Funktionen des Systems gut integriert sind.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Ich finde, dass es in dem System zu viel Inkonsistenz gibt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Ich glaube, dass die meisten Menschen die Benutzung des Systems schnell erlernen würden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Ich finde das System sehr aufwendig zu benutzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Ich habe mich bei der Benutzung sehr sicher gefühlt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Ich musste viel lernen, bevor ich anfangen konnte, das System zu benutzen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A.2 Ergebnisse Experteninterviews

Tabelle A.1: Zusammenfassung der Ergebnisse der Experteninterviews zu dem Use Case der nadelgestützten Interventionen unter Ultraschall und CT Bildgebung. Die ID gibt dabei die entsprechenden Expert*innen an

Kategorie	Statement	ID
Arbeitsschritte	Vorbereitung Patient*in	1,2
	Probeaufnahme in CT	1,2,4
	Markierung Einstichpunkt	1,2,4
	Beginn Nadelführung unter CT	1,2,4
	US am Anfang nicht für Visualisierung geeignet	2,4
	US für Feinjustierung	2,4
	Mit US Nadel folgen	1,2,3,4
	Mit US Läsion visualisieren	1,3,4
	Kontroll-CT nach Intervention	3,4
Anwendungsgebiete	Niere	1,2,3,4
	Leber	1,2,3,4
Herausforderungen	Axiale CT-Schichten Herausfordernd	1,3
	Komplexe Nadelführung	1,3
	US Einschränkungen durch Materie	1,4
	Positionierung US mit nichtdominanter Hand	3,4
	Schlechte Visualisierung CT	1,2,4
	Ergonomie CT	1,2,4
	Schutzkleidung CT	1,2,4
	Fusion und Wechsel Bildgebung	2,3
	Schichtführung US einfacher	3,4
Autonome Roboterfunktionen	Fokus Läsion US	1,3
	Nadelsteuerung US	1,2,3,4
Manuelle Roboterfunktionen	Fokus Läsion US	1,3,4
	Freie Bewegung	1,3,4
Eingabemodalitäten	Gesten durch Handfreiheit möglich	1,4
	Gesten für Bewegungsbefehle möglich	1,4
	Haptik wichtig	3,4
	Joystick als Eingabe	1,3,4
	Sprache schwierig weil viele diskrete Befehle benötigt	3,4

A.3 Ergebnisse initiale Nutzerstudie

A.3.1 Quantitative Ergebnisse

Tabelle A.2: Ergebnisse der Teilgruppe mit Sekundäraufgabe (Teilgruppe A)

P	Aktivierung	Manipulation	Translation	Rotation	TCT	Raw TLX	Verpasste Targets	Fehlclicks
1	Blick	Hand	22.66	10.71	39.93	13.17	0.56	0.00
	Blick	Kopf	29.01	17.37	60.47	14.33	0.08	0.15
	Sprache	Hand	23.54	7.01	46.03	11.00	0.17	0.00
	Sprache	Kopf	35.48	28.21	60.78	15.33	0.57	0.17
3	Blick	Hand	14.00	21.33	42.67	8.33	0.11	0.00
	Blick	Kopf	18.36	12.28	71.64	8.50	0.15	0.00
	Sprache	Hand	12.90	10.82	45.24	4.67	0.00	0.00
	Sprache	Kopf	33.31	14.29	81.16	11.00	0.05	0.08
5	Blick	Hand	19.67	21.58	79.71	12.67	0.33	0.08
	Blick	Kopf	20.36	12.45	66.69	7.50	0.58	0.11
	Sprache	Hand	19.73	15.51	65.61	6.33	0.11	0.00
	Sprache	Kopf	25.78	16.87	113.03	12.33	0.34	0.04
7	Blick	Hand	28.56	6.20	78.77	7.83	0.24	0.07
	Blick	Kopf	13.41	15.54	102.15	11.33	0.17	0.10
	Sprache	Hand	5.77	9.71	84.79	9.50	0.07	0.00
	Sprache	Kopf	22.62	14.03	83.21	9.17	0.17	0.25
9	Blick	Hand	13.04	24.55	103.16	13.00	0.39	0.22
	Blick	Kopf	11.06	16.83	62.03	13.50	0.42	0.00
	Sprache	Hand	17.53	13.81	98.59	14.17	0.11	0.11
	Sprache	Kopf	16.86	12.66	74.20	7.67	0.22	0.00
11	Blick	Hand	13.12	11.07	107.60	11.50	0.25	0.27
	Blick	Kopf	18.56	10.53	82.88	7.83	0.11	0.26
	Sprache	Hand	14.43	12.29	73.80	5.17	0.08	0.11
	Sprache	Kopf	13.49	11.82	99.54	9.50	0.39	0.00
13	Blick	Hand	26.11	14.59	75.16	13.17	0.05	0.23
	Blick	Kopf	9.55	11.28	76.69	13.00	0.31	0.17
	Sprache	Hand	23.68	11.97	50.81	11.00	0.22	0.00
	Sprache	Kopf	15.98	8.14	68.98	10.17	0.08	0.00
15	Blick	Hand	15.22	16.18	74.42	9.67	0.15	0.07
	Blick	Kopf	17.16	7.80	110.01	9.33	0.18	0.05
	Sprache	Hand	13.60	15.07	109.40	9.33	0.21	0.17
	Sprache	Kopf	17.15	21.32	96.47	6.83	0.23	0.06

Tabelle A.3: Ergebnisse der Teilaufgabe ohne Sekundäraufgabe (Teilgruppe B)

P	Aktivierung	Manipulation	Translation	Rotation	TCT	Raw TLX
2	Blick	Hand	22.25	12.08	53.19	6.83
	Blick	Kopf	24.41	5.99	61.63	10.50
	Sprache	Hand	19.22	13.83	45.61	2.67
	Sprache	Kopf	26.66	13.55	49.99	7.33
4	Blick	Hand	11.92	6.07	77.23	11.50
	Blick	Kopf	23.02	10.23	71.10	8.50
	Sprache	Hand	11.02	13.34	70.57	9.00
	Sprache	Kopf	15.61	7.28	77.50	10.67
6	Blick	Hand	17.47	14.64	96.97	12.50
	Blick	Kopf	15.27	11.52	74.51	5.67
	Sprache	Hand	14.47	19.16	53.93	8.33
	Sprache	Kopf	32.22	13.81	94.91	10.17
8	Blick	Hand	45.11	10.48	131.21	11.50
	Blick	Kopf	26.04	15.27	129.91	10.67
	Sprache	Hand	27.24	10.26	80.63	8.83
	Sprache	Kopf	22.64	7.97	126.93	8.83
10	Blick	Hand	13.53	9.55	65.34	12.17
	Blick	Kopf	18.76	8.34	50.67	12.67
	Sprache	Hand	22.57	9.06	47.25	12.83
	Sprache	Kopf	25.10	10.49	65.93	10.50
12	Blick	Hand	29.01	16.25	46.15	10.00
	Blick	Kopf	18.19	8.83	58.49	7.33
	Sprache	Hand	30.69	21.17	59.93	3.83
	Sprache	Kopf	18.55	12.21	61.57	3.17
14	Blick	Hand	16.82	11.60	51.80	7.33
	Blick	Kopf	18.66	19.71	95.05	13.17
	Sprache	Hand	29.98	10.25	47.69	7.17
	Sprache	Kopf	18.16	15.57	55.26	6.33
16	Blick	Hand	35.81	18.08	59.86	6.17
	Blick	Kopf	23.14	27.20	119.12	7.17
	Sprache	Hand	37.24	12.89	66.05	5.50
	Sprache	Kopf	15.27	18.89	130.82	13.50

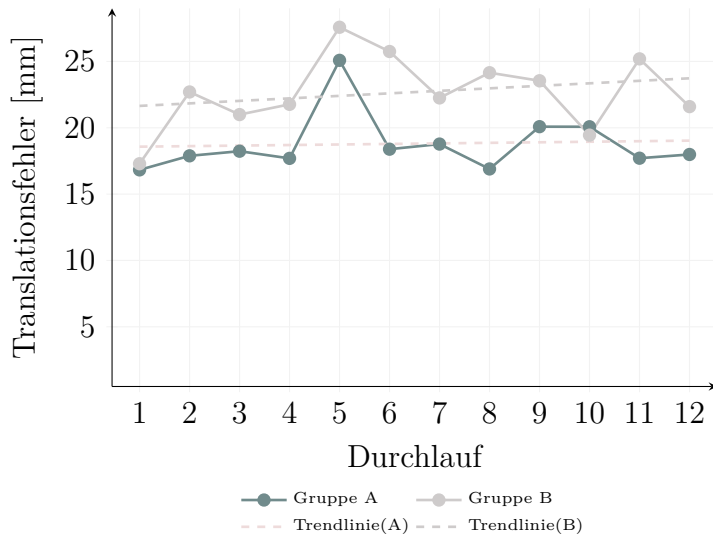
Tabelle A.4: Zusammenfassung der ANOVA-Ergebnisse ($\alpha > 0.05$) für die Parameter der Sekundäraufgabe

Variable	Effekt	F	p	η^2
Verpasste Targets	Haupteffekt diskrete Interaktion	1.535	0.226	0.052
	Haupteffekt Navigationsinteraktion	1.357	0.254	0.046
	Interaktionseffekt	1.758	0.196	0.059
Fehlclicks	Haupteffekt diskrete Interaktion	2.367	0.135	0.078
	Haupteffekt Navigationsinteraktion	0.036	0.851	0.001
	Interaktionseffekt	0.395	0.535	0.014

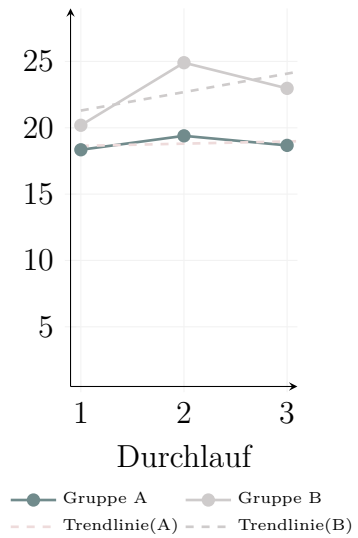
Tabelle A.5: Zusammenfassung der ANOVA-Ergebnisse ($\alpha > 0.05$) des Translationsfehlers, des Rotationsfehlers, des TCTs und des Raw Nasa TLX Scores

Variable	Effekt	F	p	η^2
Translationsfehler	Haupteffekt diskrete Interaktion	0.240	0.626	0.004
	Haupteffekt Navigationsinteraktion	0.018	0.893	0.000
	Haupteffekt Gruppe	4.334	0.042	0.072
	Interaktionseffekt diskrete & Navigationsinteraktion	1.398	0.242	0.024
	Interaktionseffekt diskrete Interaktion & Gruppe	0.061	0.805	0.001
	Interaktionseffekt Navigationsinteraktion & Gruppe	1.672	0.201	0.029
	Interaktionseffekt	0.948	0.334	0.017
Rotationsfehler	Haupteffekt diskrete Interaktion	0.005	0.943	0.000
	Haupteffekt Navigationsinteraktion	0.033	0.858	0.001
	Haupteffekt Gruppe	0.926	0.340	0.016
	Interaktionseffekt diskrete & Navigationsinteraktion	0.767	0.385	0.014
	Interaktionseffekt diskrete Interaktion & Gruppe	0.072	0.789	0.001
	Interaktionseffekt Navigationsinteraktion & Gruppe	0.075	0.785	0.001
	Interaktionseffekt	3.269	0.076	0.055
TCT	Haupteffekt diskrete Interaktion	0.226	0.636	0.004
	Haupteffekt Navigationsinteraktion	4.566	0.037	0.075
	Haupteffekt Gruppe	0.331	0.568	0.006
	Interaktionseffekt diskrete & Navigationsinteraktion	0.951	0.334	0.017
	Interaktionseffekt diskrete Interaktion & Gruppe	0.439	0.510	0.008
	Interaktionseffekt Navigationsinteraktion & Gruppe	0.515	0.476	0.009
	Interaktionseffekt	0.046	0.831	0.001
TLX	Haupteffekt diskrete Interaktion	4.151	0.046	0.069
	Haupteffekt Navigationsinteraktion	0.544	0.464	0.010
	Haupteffekt Gruppe	3.974	0.051	0.066
	Interaktionseffekt diskrete & Navigationsinteraktion	1.671	0.201	0.029
	Interaktionseffekt diskrete Interaktion & Gruppe	0.024	0.879	0.000
	Interaktionseffekt Navigationsinteraktion & Gruppe	0.019	0.890	0.000
	Interaktionseffekt	0.000	0.994	0.000

A.3.2 Lernkurven

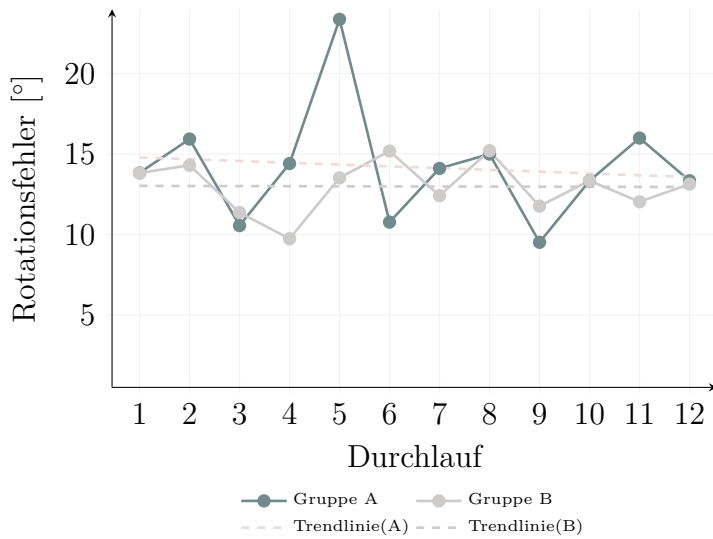


(a) Für alle Durchläufe (drei Durchläufe mal vier Kombinationen)

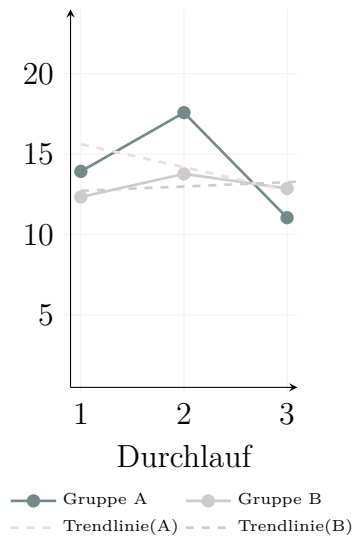


(b) Für alle Durchläufe pro Interaktionskombination

Abbildung A.1: Darstellung des Translationsfehlers unter allen Teilnehmenden und mit allen Interaktionskombinationen für die verschiedenen Durchläufe

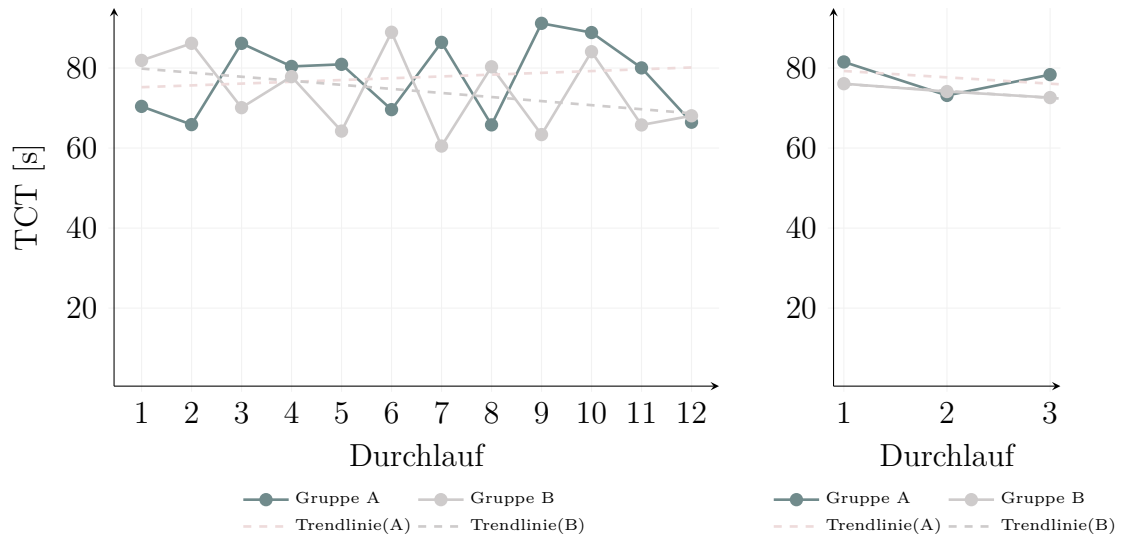


(a) Für alle Durchläufe (drei Durchläufe mal vier Kombinationen)



(b) Für alle Durchläufe pro Interaktionskombination

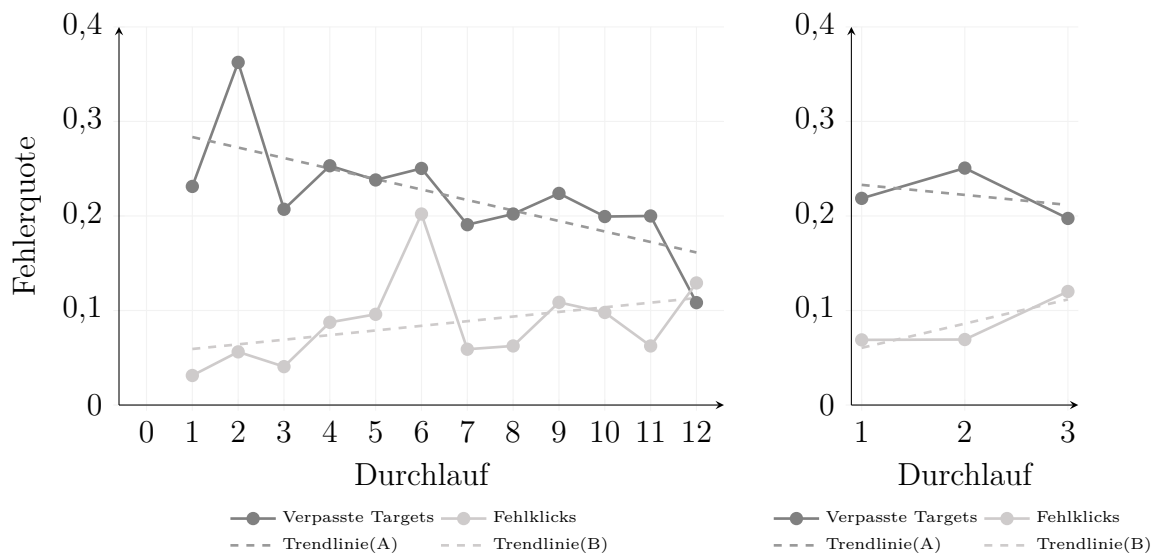
Abbildung A.2: Darstellung des Rotationsfehlers unter allen Teilnehmenden und mit allen Interaktionskombinationen für die verschiedenen Durchläufe



(a) Für alle Durchläufe (drei Durchläufe mal vier Kombinationen)

(b) Für alle Durchläufe pro Interaktionskombination

Abbildung A.3: Darstellung des TCT unter allen Teilnehmenden und mit allen Interaktionskombinationen für die verschiedenen Durchläufe



(a) Für alle Durchläufe (drei Durchläufe mal vier Kombinationen)

(b) Für alle Durchläufe pro Interaktionskombination

Abbildung A.4: Darstellung der verpassten Targets und Fehlclicks unter allen Teilnehmenden der Teilgruppe A und mit allen Interaktionskombinationen für die verschiedenen Durchläufe