

Institut für Visualisierung und Interaktive Systeme

Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Diplomarbeit Nr. 3559

**2D und 3D Gesten-Interaktion mit
einem Assistenzsystem am
Arbeitsplatz für
leistungsgeminderte Arbeiter**

Alexander Martin

Studiengang:	Informatik
Prüfer/in:	Prof. Dr. Albrecht Schmidt
Betreuer/in:	Dipl.-Inf. Markus Funk

Beginn am: 01. August 2013

Beendet am: 31. Januar 2014

CR-Nummer: H.5.2, H.5.m

Kurzfassung

Arbeitnehmer werden aufgrund des demografischen Wandels immer älter. Dies hat direkte Auswirkungen auf die Wirtschaft. Die Produktivitätsreserven einer Gesellschaft müssen mobilisiert werden, um dieser Entwicklung entgegenzutreten. Leistungsgewandelte und leistungsgeminderte Menschen können mit Hilfe von Assistenzsystemen am Arbeitsplatz ein selbstbestimmtes und produktives Arbeitsleben führen. Dabei stellt die Benutzerinteraktion mit Assistenzsystemen einen Kernpunkt für die Akzeptanz dieser Systeme dar. In der vorliegenden Diplomarbeit werden 2D und 3D Gesteneingaben an einem Assistenzsystem am Arbeitsplatz für leistungsgewandelte und leistungsgeminderte Arbeiter untersucht. 2D und 3D Gesten-Sets werden entwickelt und an einem Gesten-Interaktionssystem evaluiert. Ziel dieser Arbeit ist die Ableitung von Design Richtlinien für die Entwicklung von Gesten und die Feststellung einer geeigneten Gestenform für Assistenzsysteme am Arbeitsplatz im Kontext leistungsgewandelter und leistungsgeminderter Arbeiter.

Workers are aging due to demographic changes. This has a direct impact on the economy. The productivity reserves of a society must be mobilized to oppose this changes. Older and impaired people can live an independent and productive working life with the help of assistance systems in the workplace. The user interaction with assistance systems is a centerpiece of the acceptance of these systems. In the present thesis 2D and 3D gesture inputs are investigated on an assistance system in the workplace for older and impaired workers. 2D and 3D gestures sets are developed and evaluated in a gesture interaction system. The aim of this work is the derivation of design guidelines for the development of gestures and the finding of a suitable gesture form for assistance systems in the workplace in the context of older and impaired workers.

Inhaltsverzeichnis

1. Einleitung	11
1.1. Einführung	11
1.2. Assistierende Arbeitsplätze in der Industrie	13
1.2.1. Arbeiter	15
1.2.2. Grundlegende Schwierigkeiten	17
1.3. Problemstellung	17
1.4. Anmerkungen	19
1.5. Gliederung	20
2. Forschungsstand	21
2.1. Assistenzsysteme	21
2.1.1. Realitätserweiternde Systeme	22
2.1.2. Zusammenfassung	28
2.2. Gesten und Gestenerkennung	29
2.2.1. Konzeptionell	29
2.2.2. Technisch	35
2.2.3. Zusammenfassung	38
2.3. Fazit	39
3. Assistenzsystem	41
3.1. Beschreibung	42
3.2. Anwendungsfälle	43
3.3. Aktionsmöglichkeiten	45
3.3.1. Trainer	45
3.3.2. Arbeiter	47
4. Vorstudie	51
4.1. Konzeption der Durchführung	51
4.2. Teilnehmer	52
4.3. Apparatur	53
4.4. Durchführung	54
4.5. Resultate und Diskussion	54
4.5.1. Identifikation relevanter Gesten	56
5. Interaktionssystem	73
5.1. Konzeption	73
5.1.1. Anforderungen	73

5.2. Architektur	74
5.3. Implementierung	76
5.3.1. GestureRecognitionManager und GestureManager	77
5.3.2. Logik der Gestenerkennung	77
5.3.3. Verwendete Technologien	79
5.3.4. Verwendete Algorithmen	80
5.3.5. Fazit	81
6. Evaluation	83
6.1. Evaluationsdurchführung — Usability Studie	83
6.1.1. Verwendetes Vorgehensmodell	83
6.1.2. Teilnehmer	84
6.1.3. Apparatur	84
6.1.4. Durchführung	84
6.1.5. Darstellung der Evaluationsergebnisse	85
6.2. Schlussfolgerungen und Diskussion	89
6.2.1. Design Richtlinien für Gesteneingaben	91
7. Zusammenfassung und Ausblick	95
7.1. Zusammenfassung der Arbeit	95
7.1.1. Interpretation der Resultate	96
7.1.2. Auswirkungen der Ergebnisse	96
7.2. Ausblick	96
A. Anwendungsfälle	99
Literaturverzeichnis	101

Abbildungsverzeichnis

1.1. Einsatz von Assistenzsystemen in der Wirtschaft	13
1.2. Gliederung assistierender Systeme am Arbeitsplatz	15
2.1. ASLM System und erweitertes ASLM System	24
2.2. Gamification-Elemente am Assistenzsystem	25
2.3. LGS Assistenzsystem OPS Solutions	27
2.4. Schnaithmann Montage-Assistenzsystem auf Basis von Bewegungserkennung	28
2.5. ShowMe Montage-Assistenzsystem	29
2.6. Gestenerfassung durch künstliche neuronale Netze	38
2.7. Darstellung von Systemen zur Gestenerfassung	39
3.1. motionEAP Montagetisch und Aufbau	42
3.2. motionEAP Prototyp Assistenzsystem	43
3.3. Montage-Szenario Schraubzwinde	45
3.4. Arbeitsablauf erstellt durch Trainer	46
3.5. Ausführung von Arbeitsschritten durch Arbeiter	48
4.1. Benutzerstudie zur Identifikation von Gesten	53
4.2. Ermittlung Gewinnergesten aus Benutzerstudie zur Identifikation von Gesten	55
4.3. Range of motion Normwerte	61
4.4. 2D Gesten Taxonomie Verteilung	69
4.5. 3D Gesten Taxonomie Verteilung	70
4.6. Hybrid Gesten Taxonomie Verteilung	71
5.1. Konzeptionelle Architektur des motionEAP Systems	76
5.2. Gesten-Interaktionssystem Admin-View, Erkennung einer Handgeste	77
5.3. Gesten-Interaktionssystem Gestenerkennung 2D Geste	78
5.4. Gesten-Interaktionssystem Gestenerkennung 3D Geste	79
5.5. Gesten-Interaktionssystem Veranschaulichung des \$ Algorithmus	81
5.6. Gesten-Interaktionssystem Beispiel-Code des \$P Algorithmus	82
6.1. Benutzerstudie zur Usability von Gesten	85
6.2. Auswertung Benutzerstudie zur Usability, Gesten zu Aktionen	86
6.3. Auswertung Benutzerstudie zur Usability, Schweregrad von Gesten	87
6.4. Auswertung Benutzerstudie zur Usability, Prozesssicherheit von Gesten	88
A.1. Use case, Arbeiter	99
A.2. Use case, Trainer	100

Tabellenverzeichnis

2.1. 2D-Gesten Taxonomie von Wobbrock	32
2.2. 3D-Gesten Taxonomie von Ruiz	33
3.1. Aktionen des Trainers	47
3.2. Aktionen des Arbeiters	49
4.1. 2D-Gesten Taxonomie für Assistenzsysteme und spezielle Arbeiter	58
4.2. 3D- und Hybrid-Gesten Taxonomie für Assistenzsysteme und spezielle Arbeiter	59
4.3. 2D-Gesten für Assistenzsysteme Teil 1	62
4.4. 2D-Gesten für Assistenzsysteme Teil 2	63
4.5. 3D-Gesten für Assistenzsysteme Teil 1	64
4.6. 3D-Gesten für Assistenzsysteme Teil 2	65
4.7. 3D-Gesten für Assistenzsysteme Teil 3	66

Abkürzungsverzeichnis

ANN	Artificial neural network
AR	Augmented reality
ASLM	Assistenzsysteme für leistungseingeschränkte Mitarbeiter
CAD	Computer-aided design
CMOS	Complementary Metal Oxide Semiconductor
CSV	Comma-separated values
DLP	Digital Light Processing
HCI	Human-computer interaction
HGR	Hand gestures recognition
HMI	Human-machine interaction
JSON	JavaScript Object Notation
LCD	Liquid-crystal display
LED	Light-emitting diode
LGS	Light Guide System

NUI Natural User Interfaces
PCA Principal Component Analysis
PDA Personal Data Assistant
RITE Rapid iterative testing and evaluation
ROM Range of motion
SDK Software Development Kit
WPF Windows Presentation Foundation

1. Einleitung

1.1. Einführung

Der demographische Wandel in Deutschland hat nicht nur Auswirkungen auf den gesellschaftlichen Zusammenhalt, die sozialen Sicherungssysteme, die Politik, sondern auch auf die Wirtschaft. Das Wissenschaftsjahr 2013 des Bundesministeriums für Bildung und Forschung zeigt mit dem Überbegriff – Die demografische Chance – die aktuelle und steigende Bedeutung dieses Themas für die Gesellschaft.

Die Altersverteilung der Bevölkerung verschiebt sich seit dem ersten Drittel des 20. Jahrhunderts, genauer seit dem ersten Weltkrieg, zugunsten der alten und sehr alten Jahrgänge. In mehreren Abwärtswellen setzte ein irreversibler Geburtenrückgang ein, der zu der heutigen geringen Geburtenrate führte. Lag der Altenquotient, der dem Quotient aus der Zahl der über 65-jährigen und den erwerbsfähigen Menschen unter 65 Jahren entspricht, um 1900 noch bei 8,1, so lag er 2011 bei 33,7. Bevölkerungsvorausberechnungen sagen weiterhin einen tendenziell deutlichen Anstieg des Altenquotienten voraus. Bis zum Jahr 2060 könnte dieser einen Wert von über 67 erreichen (Bundesinstitut für Bevölkerungsforschung¹). Diese Entwicklung bedeutet also, dass der Altersdurchschnitt der Bevölkerung bis 2060 bei mehr als 50 Jahren liegen wird, wobei 2003 der Altersdurchschnitt bei etwa 41 Jahren lag. Schon 2025 dürfte etwa jeder vierte Erwerbsfähige älter als 55 Jahre sein. Deutschland ordnet sich somit mit seiner demographischen Entwicklung in den Trend der entwickelten Länder Europas ein.

Aus wirtschaftlicher Sicht hat diese Entwicklung weitreichende Folgen, bietet Gefahren und Chancen. Um diese Folgen aufzufangen brauche es gesellschaftliche Reformen, Innovationen und Ideen, mit denen sich die Produktivitätsreserven innerhalb einer Gesellschaft mobilisieren lassen. Unter diesem Aspekt liegt hier ein besonderer Fokus auf leistungsgewandelte, als auch leistungsgeminderte Menschen.

In den nächsten zehn Jahren werden die geburtenstarken Nachkriegsjahrgänge aus dem Arbeitsleben ausscheiden. Als direkte Folge wird der Fachkräftemangel noch weiter zunehmen. Die zunehmende Alterung der Bevölkerung und die geringere Anzahl an neuen Erwerbstätigen stellt die klassische Dreiteilung des Lebens in Ausbildung, Erwerbsleben und Ruhestand in ihrer gegenwärtigen Form auf den Prüfstand. Auch im Alter müsse es normal werden, zu arbeiten und sich weiterzubilden. Unternehmen müssen stärker auf die Erfahrungen leistungsgewandelter Arbeitnehmer bauen, die im normalen Alterungsprozess körperliche

¹siehe <http://www.bib-demografie.de>

und geistige Einbußen erlitten haben. Obwohl der Anteil der leistungsgewandelten Arbeitnehmer in Deutschland stetig wächst, betrug 2012 die Beschäftigungsrate zwischen den 55- bis 64-jährigen nur 62% (vgl. [SBKK13, S. 6]). Obwohl Deutschland im europäischen Vergleich damit im vorderen Drittel liegt, haben vor allem nordische Länder weiterhin deutlich höhere Beschäftigungsraten. Arbeitnehmer werden in Zukunft ein längeres Erwerbsleben haben und müssen auch im höheren Alter aktiv an diesem teilnehmen können. Arbeitgeber müssen darauf reagieren und Arbeitstätigkeiten und Arbeitsbedingungen an ihre Mitarbeiter anpassen.

Auch leistungsgeminderte Menschen haben das Recht im Arbeitsalltag integriert zu werden und eine an ihre Gegebenheiten angepasste Arbeit zu verrichten. Unternehmen, die mehr als 20 Beschäftigte haben, werden per Gesetz verpflichtet fünf Prozent der Belegschaft mit Menschen mit Behinderungen zu besetzen (vgl. [NPMP10, 2. Kapitel Beschäftigungspflicht]). Diese Verpflichtung kann in einer Kooperation mit Werkstätten für behinderte Menschen eingelöst werden. Durch die Individualisierung von Produkten gibt es im produzierenden Gewerbe einen Trend zur Kleinserie. Themen wie Variantenvielfalt und die Veredelung durch Handarbeit sind aktuell. In der hochautomatisierten Massenfertigung ist die Produktion von Kleinserien, als auch die Auslagerung dieser ins Ausland, für Unternehmen jedoch unwirtschaftlich und wird häufig an lokale Werkstätten für behinderte Menschen vergeben (vgl. [Kor12]). Leistungsgeminderte Arbeiter bedürfen jedoch einer Führung und einer Aufsicht durch einen Betreuer. Diese individuelle menschliche Assistenz ist jedoch nicht wirtschaftlich.

Auf Grund der gesellschaftlichen Entwicklung gibt es ein besonderes Interesse leistungsgewandelten, als auch leistungsgeminderten Menschen ein weitgehend selbstbestimmtes und motivierendes Arbeitsumfeld zu schaffen. Assistenzsysteme sollen diese Aufgabe übernehmen, Arbeitnehmer unterstützen und fördern. Dabei stellt die Interaktion mit diesen Systemen einen Kernpunkt für die Akzeptanz und Annahme dieser Systeme dar. Die Interaktion sollte möglichst natürlich, intuitiv sowie leicht erlernbar sein (vgl. [KBS⁺12]). Sie umfasst die direkte Bedienung eines Systems durch Wischen, Tippen, Berühren, Gesten oder Sprache. Durch neue Entwicklungen in Technologien wie der Projektionstechnik, der Sensorik und Kameratechnik entwickeln sich Assistenzsysteme weiter und können vielfältig eingesetzt werden.

Smartphones und Tablets sind zum Alltagsgegenstand geworden, Gestensteuerung durch Berührung von Displays ist allgegenwärtig. Durch die Spieleindustrie und Nintendos² Wii wurde die Bewegungssteuerung im dreidimensionalen Raum anwendergerecht. Seit November 2013 kann Microsofts³ neue Spielekonsole Xbox One vollständig durch dreidimensionale Gesten gesteuert werden. Microsofts Xbox Kinect ermöglicht eine kostengünstige Gestensteuerung im dreidimensionalen Raum. Im Bereich der Unterhaltungselektronik wird die Interaktion mit Geräten durch zwei- und vor allem dreidimensionale Gesten immer populärer. Auch in der Industrie wird die Interaktion mit der Arbeitsumgebung durch Gesten in Zukunft an Bedeutung gewinnen.

²siehe <http://www.nintendo.com>

³siehe <http://www.microsoft.com>

Die Erforschung gestenbasierter Benutzerschnittstellen wurde im Bereich der assistierenden Systeme im Kontext eines Arbeitsplatzes für spezialisierte Arbeiter bislang weitgehend vernachlässigt. Diese Arbeit will sich in diesem Kontext mit der Gesteninteraktion innerhalb dieser Systeme befassen, zentrale Fragen nach dem Nutzen von Gesteneingaben beantworten sowie abschließend Richtlinien für den Designprozess von Gesten für die zuvor erwähnten Systeme und dessen Anwender ableiten.

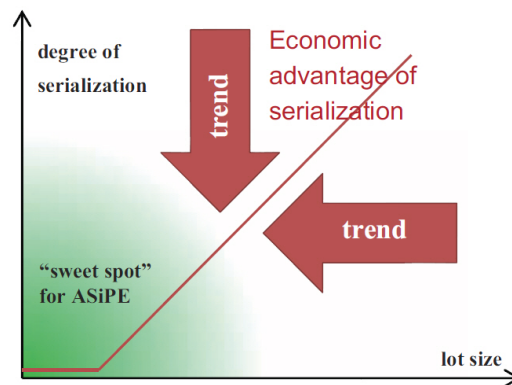


Abbildung 1.1.: Bei kleineren Produktionsmengen (x-Achse) und steigenden Produktvarianten (y-Achse) ist der Einsatz von realitätserweiternden Assistenzsysteme (engl. Assistive Systems in Production Environments, kurz ASiPE) wirtschaftlich. (vgl. [KSH12])

1.2. Assistierende Arbeitsplätze in der Industrie

Assistierende Systeme am Arbeitsplatz, welche den Arbeiter aktiv, oder auch passiv unterstützen sollen, werden in der Industrie seit Jahren eingesetzt. Vor allem finden in der produzierenden Industrie diese Systeme Verbreitung. Mit dem Fortschritt der computerbasierten Systeme ist in den letzten 20 Jahren ein Trend zu hochentwickelten assistierenden Systemen zu beobachten (vgl. [SPH10, BK12]). Von einfachen Bilderkarten entwickelten sich diese Systeme bis hin zu fernüberwachten Robotersystemen. Die Anforderungen der Industrie an diese Systeme sind zum einen flexible Systeme die verschiedene Aufgaben lösen können, zum anderen Systeme die einfach und schnell zu konfigurieren sowie leicht zu erlernen sind. Aus Arbeitersicht ergeben sich weitere Anforderungen, wie die mögliche Skalierbarkeit der Systeme an die Fähigkeiten des Arbeiters (Komplexität der Arbeitsschritte), die Bereitstellung motivierender Elemente und die Unterstützung des Arbeiters durch kontextsensitive Hilfe. Das Design dieser Systeme stellt eine Herausforderung dar. Es soll Arbeiter unterstützen und entlasten, ihre Möglichkeiten fördern und die Produktivität steigern. Gleichzeitig soll die Arbeit interessanter werden und mehr Freude machen. Im Ganzen soll der Einsatz dieser Systeme die Wirtschaftlichkeit des Produktionsprozesses steigern.

Assistierende Systeme am Arbeitsplatz können in Systeme zu Lern- und Übungszwecken und Systeme zur Verbesserung und Übernahme von Arbeitstätigkeiten eingeteilt werden. Die letzteren können in drei grobe Bereiche gegliedert werden (vgl. [BKE⁺13]). Die Abbildung 1.2 zeigt diese Bereiche schemenhaft. Erstens durch Roboter unterstützte Systeme (engl. robotic co-worker systems), zweitens fernüberwachte Robotersysteme (engl. tele-supervised robotics) und drittens realitätserweiternde Systeme (engl. augmentation systems).

Roboter unterstützte Systeme ermöglichen es dem Arbeiter die Genauigkeit und Kraft eines Roboters gezielt zur Lösung von Problemen sowie zur Herstellung von Produkten einzusetzen. Hierbei leitet der Arbeiter den Roboter und trifft Entscheidungen, zum Beispiel wie eine schwere Autotür zur Montage gedreht werden soll. Roboter unterstützte Systeme können an verschiedene Anforderungen im Produktionsprozess angepasst werden und vereinigen die Genauigkeit, Kraft und Geschwindigkeit einer Maschine mit der Flexibilität, Kreativität und Entscheidungsfindung eines Menschen.

Fernüberwachte Robotersysteme geben dem Arbeiter die Möglichkeit Aufgaben durch einen fernkontrollierten Roboter ausführen zu können. Der Arbeiter muss nicht am Ausführungs-ort anwesend sein und ist somit keinen gefährlichen Bedingungen und Umwelteinflüssen ausgesetzt. Fernüberwachte Robotersysteme erlauben das Ausführen von Aufgaben auf sehr kleinem und großem Raum, wobei der Arbeiter normalerweise auf diesen keinen Zugriff hätte.

Realitätserweiternde Systeme erleichtern dem Arbeiter während der Produktion, der Montage oder der Qualitätssicherung die Durchführung von Aufgaben. Sensoren zur Objekt- und Personenerkennung unterstützen hierbei diese Systeme. Projektionen in den Arbeitsbereich erweitern die Realität und bieten kontextsensitive Hilfe sowie aufgabenspezifische Informationen. Durch diese Erweiterungen können die Fähigkeiten, nicht nur von uneingeschränkten, sondern auch von leistungsgewandelten und leistungsgeminderten Arbeitern gesteigert werden (vgl. [KSH13b, RKWA10]). Defizite leistungsgewandelter und leistungsgeminderter Arbeiter, wie das Nachlassen des Kurzzeitgedächtnisses und die damit einhergehende Verschlechterung der Lernfähigkeit (vgl. [BK12]), als auch geringere kognitive Aufnahmefähigkeiten können verringert werden. Durch die leichtere Durchführung von Aufgaben und das frühzeitige Erkennen von potenziellen Fehlern werden Arbeiter befähigt Produkte in höherer Qualität zu fertigen. Gamification-Elemente wie Fortschrittsanzeigen, Levels, Punktesysteme sowie akustische und visuelle Feedbacks können zur Verbesserung der Qualität, Geschwindigkeit und der Zufriedenheit der Mitarbeiter beitragen und so die Produktivität erhöhen (vgl. [KBS⁺12, KSH12]).

Aus Sicht der Arbeits- und Organisationspsychologie ist das Ziel der Entwicklung eines Assistenzsystems die Ausführbarkeit von Aufgaben, beeinträchtigungsfreie Arbeitsbedingungen und die Entwicklungsförderung des Arbeiters. Pädagogisch gesehen steht die Aufgabe der Zusammenführung von Arbeiter und Assistenzsystem im Vordergrund. Psychische Belastungen im Arbeitsprozess sollen Arbeiter nicht negativ beanspruchen. Unter diesen Belastungen versteht man die Gesamtheit aller erfassbaren Einflüsse, die von außen auf den Menschen zukommen und psychisch auf ihn einwirken (vgl. [FKB⁺13]). In diesem Kontext wären diese Belastungen die Arbeitsaufgabe, die Arbeitsmittel, die Arbeitsumgebung, die Arbeitsorganisation und der Arbeitsplatz. Neben den psychischen Belastungen wird die

psychische Beanspruchung wie folgt definiert. Sie bezeichnet die unmittelbare (nicht langfristige) Auswirkung der psychischen Belastung im Individuum in Abhängigkeit von seinen jeweiligen überdauernden und augenblicklichen Voraussetzungen, einschließlich der individuellen Bewältigungsstrategien (vgl. [FKB⁺13]). Psychische Beanspruchung entsteht somit aus der Wechselwirkung zwischen psychischen Belastungen und der Fähigkeit des Arbeiters diese Belastungen zu bewältigen. Positive psychische Beanspruchung äußert sich in der Aktivierung körperlicher und geistiger Fähigkeiten, während sich negative Beanspruchung in psychischer Ermüdung, Stress, Unzufriedenheit, Depression und Leistungsminderung niederschlägt. Durch den Einsatz von Assistenzsystemen soll es zur Optimierung der psychischen Belastungen kommen, mit dem Ziel den Arbeiter weder zu unterfordern noch zu überfordern und ihn positiv psychisch zu beanspruchen.

Die Verwendung assistierender Systeme wirft aufgrund von Überwachungsmöglichkeiten auch ethische Fragen auf. Aus negativen ethischen Aspekten entstehen Gründe, die gegen einen Einsatz von Assistenzsystemen am Arbeitsplatz sprechen, aus positiven Aspekten entstehen Gründe für einen Einsatz. Das Für und Wider für den Einsatz eines Assistenzsystems muss abgewogen werden. Allgemeine Standards, an denen heute ein guter Arbeitsplatz gemessen wird, können in ihrer Struktur auf Arbeitsplätze mit Assistenzsystemen übertragen werden. Diese Standards tragen die Vorstellung von „guter“ Arbeit in sich, bei der Bedürfnisse nach Selbstbestimmung, Anerkennung und Beteiligung erfüllt werden.

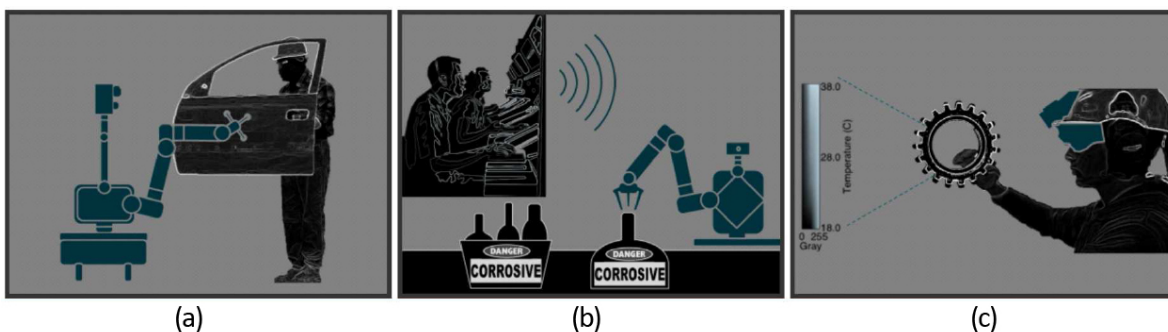


Abbildung 1.2.: Einordnung assistierender Systeme am Arbeitsplatz. Bild (a) zeigt durch Roboter unterstützte Systeme, Zusammenarbeit mit Robotern. (b) zeigt fernüberwachte Robotersysteme, Arbeitsverrichtung durch Roboter. Bild (c) zeigt realitätserweiternde Systeme, Arbeitserleichterung für den Menschen. (vgl. [BKE⁺13])

1.2.1. Arbeiter

Assistierende Systeme können die Fähigkeiten von uneingeschränkten, als auch leistungsgewandelten und leistungsgeminderten Arbeitern steigern. Die Systeme haben das Ziel vor allem leistungsgewandelten und leistungsgeminderten Arbeitern ein möglichst autonomes

Arbeitsleben zu ermöglichen. Assistenzsysteme können diesen Benutzergruppen dementsprechend einen Einstieg in Arbeitsbereiche ermöglichen, die diesen sonst verschlossen blieben, oder zu viel Betreuungsaufwand bedeuten würden.

Leistungsgewandelte Arbeiter sind Arbeitnehmer, die im normalen Alterungsprozess körperliche und geistige Einbußen erlitten haben. Diese Arbeitnehmer weisen neben größerer Arbeitserfahrung ein besseres strategisches Denken, verbesserte Ausdrucksfähigkeiten, ein größeres Engagement sowie eine höhere Motivation auf (vgl. [DBB⁺12, Ilm05]). Jedoch lässt im Alter die physische Muskelkraft nach, Bewegungen können nicht mehr so schnell durchgeführt werden. Ebenfalls nimmt die Reaktionszeit ab. Visuell tritt zwischen dem 40. und 50. Lebensjahr eine Altersweitsichtigkeit (Presbyopie) auf. Obwohl es unter älteren Arbeitnehmern zu gleich vielen Arbeitsunfällen wie unter jüngeren Arbeitnehmern kommt, dauert nach einem Unfall die Erholungsphase von älteren Arbeitnehmern länger. Neben physischen Einschränkungen stellen sich im Alter kognitive Einschränkungen ein, zum Beispiel lässt das Kurzzeitgedächtnis, als auch die Konzentrationsfähigkeit nach. Dies hat die Verschlechterung der Lernfähigkeit zur Folge (vgl. [BK12]) und wirkt sich direkt auf die Fehlerrate in gestellten Produktionsaufgaben aus. Die Konsequenz aus diesen durch den Alterungsprozess bedingten körperlichen und geistigen Folgen ist eine frühzeitige Rente, die oft ungewollt und zu früh in Anspruch genommen werden muss (vgl. [Kor12]). Aufgrund des demographischen Wandels wird sich jedoch das Erwerbsleben von Arbeitnehmern verlängern und das Arbeiten im höheren Alter zur Notwendigkeit.

Leistungsgeminderte Arbeiter sind Arbeitnehmer, die eine körperliche oder geistige Behinderung haben. Behinderungen können angeboren sein, durch Unfälle oder Krankheiten hervorgerufen werden. Leistungsgeminderte und leistungsgewandelte Arbeiter haben Gemeinsamkeiten bei physischen und kognitiven Einschränkungen (vgl. [KBS⁺12]). Beide Gruppen leiden unter generellem Muskelabbau, Koordinationsschwierigkeiten und Herz-Kreislauf-Problemen. Die Senso- und Feinmotorik kann zum Beispiel unter Umständen unter der Parkinsonkrankheit leiden. Das Kurzzeitgedächtnis, als auch die Konzentrationsfähigkeit lassen nach. Der psychische, als auch physische Zustand kann innerhalb weniger Stunden stark variieren. Weitere Einschränkungen leistungsgeminderter Arbeiter können durch geistige Behinderungen, Autismus, das Asperger-Syndrom und andere tiefgreifende Entwicklungsstörungen hervorgerufen werden. Diese Einschränkungen können Sprach- und Verständnisschwierigkeiten, Problemlösungsdefizite, Schwierigkeiten mit Veränderungen der Umgebung, als auch repetitive Körperbewegungen und andere stereotype Verhaltensweisen sein. Werkstätten für behinderte Menschen definieren leistungsgeminderte Arbeiter als Personen, die aufgrund ihrer körperlichen, geistigen oder psychischen Beeinträchtigungen oder Besonderheiten noch nicht oder noch nicht wieder auf dem allgemeinen Arbeitsmarkt erwerbstätig werden können (vgl. [NPMP10, § 136 – Begriff und Aufgaben der Werkstatt für behinderte Menschen]). Dies schließt teilweise auch Menschen mit einer Lernbehinderung oder Körperbehinderung ohne das Merkmal einer geistigen Behinderung ein.

Uneingeschränkte Arbeiter stellen den durchschnittlichen gesunden Arbeitnehmer dar und bedürfen keiner näheren Definition.

1.2.2. Grundlegende Schwierigkeiten

Die grundlegenden Schwierigkeiten im Zusammenhang mit assistierenden Systemen am Arbeitsplatz können aus den Anforderungen an diese Systeme abgeleitet werden. Die folgenden allgemeinen Anforderungen an diese Systeme sind:

- **Flexibilität der Systeme**, d.h. das System soll möglichst für verschiedene Aufgaben einsetzbar sein und diese lösen können.
- **Skalierbarkeit der Systeme**, d.h. das System soll sich möglichst an die Fähigkeiten des Arbeiters anpassen und die Schwierigkeit der Aufgabe passend wählen und dementsprechend präsentieren.
- **Vereinfachte Benutzerschnittstellen**, d.h. das System soll einfach und schnell zu bedienen und zu konfigurieren sein.
- **Mechanismen zur Motivierung**, d.h. motivierende Elemente wie Fortschrittsanzeigen und Punktesysteme sollen den Nutzer motivieren und dabei den Arbeitsfluss fördern.

Die Schwierigkeiten die sich aus diesen Anforderungen ergeben, sind in erster Linie Designprobleme des Systems. Folgende Fragen müssen zur Lösung dieser Probleme gestellt werden. In welchem Kontext soll das assistierende System Verwendung finden? Welche Aufgaben sollen mit Hilfe dieses Systems gelöst werden können? Welche Fähigkeiten und Einschränkungen haben Arbeiter, die dieses System nutzen werden? Anhand der Antworten dieser Fragen können Designentscheidungen getroffen und die Erkenntnisse aus Forschungsergebnissen der Mensch-Computer-Interaktion einbezogen werden.

Um die Anforderungen an assistierende Systeme am Arbeitsplatz zu erfüllen, wird ein nutzerorientierter Designansatz (engl. *user-centered design*) vorgeschlagen (vgl. [KSH12, BK12, KBS⁺12]). Der Benutzer steht in diesem Ansatz im Mittelpunkt. Das System reagiert und passt sich an die Fähigkeiten, Bedürfnisse, Einschränkungen und Absichten des Nutzers an. Der Benutzer ist nicht gezwungen sich an ein System anpassen zu müssen.

1.3. Problemstellung

Im weiteren Verlauf konzentriert sich diese Arbeit auf die im vorhergehenden Kapitel beschriebenen realitätserweiternden assistierenden Systeme am Arbeitsplatz, weiterhin kurz „Assistenzsysteme“ genannt. Die Nutzer sollen leistungsgeminderte, als auch leistungsgewandelte Arbeiter sein. Die Entwicklung dieser Assistenzsysteme stellt Systemdesigner und Entwickler vor Schwierigkeiten, die aus den Anforderungen an diese Systeme abgeleitet werden.

Korn et al. identifizieren in den Arbeiten [Kor12, KSH12, KSH13b] zusammenfassend die folgenden Anforderungen an Assistenzsysteme:

- **Steigerung der Prozessorientierung**, d.h. das System soll möglichst im Hintergrund arbeiten und nur bei potenziellen Fehlern und falschen Aktionen unterstützen und eingreifen. Ein durchgängiger Arbeitsfluss soll gefördert werden.
- **Vereinfachung der Benutzerschnittstellen**, d.h. eine möglichst natürliche Bedienung des Systems soll möglich sein. Benutzerschnittstellen sollen dem Nutzer eine direkte Interaktion mit dem System durch Wischen, Tippen, Berühren, Gesten oder Sprache ermöglichen.
- **Bereitstellung von Mechanismen zur Motivierung**, d.h. Gamification-Elemente wie Fortschrittsanzeigen, Levels, Punktesysteme sowie akustische und visuelle Feedbacks sollen den Nutzer motivieren und dabei den Arbeitsfluss fördern.

Aus diesen Anforderungen ergibt sich eine Vielzahl von Fragen und Problemen, deren Erforschung nicht umfassend ist oder noch aussteht. Korn et al. untersuchen zum Beispiel in [Kor12] die Integration von Gamification-Elementen in Assistenzsysteme zur Motivierung leistungsgeminderter und leistungsgewandelter Arbeiter. Die Erforschung von Benutzerschnittstellen in Verbindung mit Assistenzsystemen im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter blieb bis heute aus.

Diese Arbeit möchte diese Benutzerschnittstellen untersuchen. Dabei liegt der Fokus primär auf einer Benutzerschnittstelle, welche die Bedienung eines Assistenzsystems durch zweidimensionale, als auch dreidimensionale Gesten ermöglichen soll.

Wie kann eine Benutzerschnittstelle eines Assistenzsystems entworfen und vereinfacht werden, sodass eine intuitive und natürliche Interaktion möglich wird? Wie sehen momentan genutzte Bedienkonzepte aus? Können zwei- und dreidimensionale Gesten eine natürliche Interaktion fördern? Welche Unterschiede, Vor- und Nachteile gibt es zwischen diesen Gestenarten? Welche Probleme können bei leistungsgeminderten und leistungsgewandelten Arbeitern im Hinblick auf die Gestenausführung entstehen? Welche Gesten können für diese Arbeiter gewählt werden? Diese Fragen sollen durch diese Arbeit beleuchtet und beantwortet werden.

Zwei Studien werden zur Identifikation von Gesten und zur Evaluierung dieser Gesten durchgeführt. In diesen Studien werden empirische Daten erfasst. Als Ergebnis der Evaluierung werden Designrichtlinien für den Gestenentwurf und die Gesteneingabe für leistungsgeminderte und leistungsgewandelte Arbeiter aufgestellt. Im Vorfeld dieser Studien werden Aktionen an einem Assistenzsystem identifiziert, welche den Benutzergruppen zugeordnet sind. Anhand der Ergebnisse der ersten Studie werden zu diesen Aktionen Gesten entwickelt, welche die Aktionen repräsentieren. Ein Gestenerfassungssystem wird für ein Assistenzsystem implementiert und unterstützt zwei- und dreidimensionale Gesten. Als Erweiterung werden ebenfalls sogenannte Hybridgesten entwickelt, implementiert und bei der Untersuchung miteinbezogen. Im Vorfeld haben sich Hybridgesten als interessante Alternative zu zwei- und dreidimensionale Gesten herausgestellt. In der zweiten Studie wird die Benutzerfreundlichkeit, die Gebrauchstauglichkeit der verschiedenen Gestenarten untersucht. Abschließend wird eine Empfehlung für die Nutzung von Gesten in Assistenzsystemen für leistungsgeminderte und leistungsgewandelte Arbeiter gegeben.

Die technische Herausforderung aus Sicht der Informatik stellt in dieser Arbeit die Handgestenerkennung dar (engl. HGR). Gestenerkennung ist ein multidisziplinäres Forschungsgebiet und fällt in den Bereich der *Computer Vision und Mustererkennung*. Techniken zur Erkennung von Gesten umfassen *statistische Modelle, Bild- und Mustererkennung, Bildverarbeitung* und *konnectionistische Systeme*, wie *künstliche neuronale Netze* und andere. Aufgrund der Komplexität der meisten Verfahren wird in dieser Arbeit ein erweitertes Verfahren der geometrischen Mustererkennung angewandt.

In Folge einer Kooperation mit dem motionEAP-Projekt⁴ (vgl. [FKB⁺13]) dient ein Prototyp eines Assistenzsystems als Forschungsgrundlage. Der nähere Aufbau dieses Systems wird in Kapitel 3 beschrieben.

1.4. Anmerkungen

Ein Teil der verwendeten Literatur liegt ausschließlich in englischer Sprache vor. Für die Nutzung und Referenzierung solcher englischen Originaltexte verwendet der Verfasser dieses Dokuments eigenhändig angefertigte Übersetzungen. Diese wurden nach bestem Wissen und Gewissen vorgenommen. Alle vorkommenden Fachbegriffe werden, sofern dies möglich ist und es sich nicht um Eigennamen handelt, durch entsprechende deutsche Begriffe ersetzt.

Anerkannte Technologien und Standards, aber auch bekannte Markennamen, werden standardmäßig durch die Angabe ihrer offiziellen Web-Präsenz in einer Fußnote referenziert.

Das erste Vorkommen neu eingeführter Fachbegriffe und Akronyme wird vom restlichen Text durch eine kursive *Schriftauszeichnungsart* hervorgehoben.

Alle besprochenen Grundlagen und Hintergründe erheben keinen Anspruch auf Vollständigkeit. Sie beschränken sich lediglich auf die zum Verständnis des vorliegenden Sachverhalts erforderlichen Themen.

⁴siehe <http://www.motioneap.de>

1.5. Gliederung

Die Arbeit ist in folgender Weise gegliedert:

Kapitel 2 – Forschungsstand: In Kapitel 2 wird der aktuelle Stand der Wissenschaft und der Technik betrachtet. Eine detaillierte Übersicht existierender Arbeiten in den Bereichen der Assistenzsysteme, der Gestenforschung und der Gestenerkennung bildet die Grundlage dieser Arbeit.

Kapitel 3 – Assistenzsystem: Dieses Kapitel beschreibt das zur Untersuchung der Problemstellung eingesetzte Assistenzsystem und das motionEAP Projekt. Es werden Anwendungsfälle innerhalb dieses Assistenzsystems identifiziert. Hierbei werden die Benutzer des Systems in Gruppen gegliedert und näher erläutert. Aus den beschriebenen Anwendungsfällen werden Aktionsmöglichkeiten der Benutzergruppen hergeleitet und gruppiert.

Kapitel 4 – Vorstudie: Kapitel 4 geht auf die Vorstudie zur Identifikation relevanter Gesten und die Zuordnung dieser Gesten zu den zuvor bestimmten Aktionsmöglichkeiten ein. Im Rahmen der Studie werden Gesten von Teilnehmern gesammelt und im Hinblick auf die Verwendung im Assistenzsystem bewertet. Der Ablauf der Studie wird beschrieben. Abschließend erfolgt die statistische Analyse der Studienergebnisse. Die Auswahl relevanter Gesten erfolgt aufgrund dieser Ergebnisse, wobei ebenfalls bestimmte anhand der Zielgruppe (Arbeiter) festgelegte Kriterien berücksichtigt werden.

Kapitel 5 – Interaktionssystem: Im Mittelpunkt von Kapitel 4 steht die Konzeptions-, Entwurfs- und Entwicklungsphase des Gesten-Interaktionssystems. Neben einer Anforderungsanalyse findet eine detaillierte Beschreibung der Architektur, des logischen Datenmodells und der einzelnen Applikationskomponenten statt. Folgend fließen die bisherigen theoretischen Erkenntnisse in die Entwicklungsphase ein. Es werden verwendete Technologien und Datenformate beschrieben. Die Funktionsweise und Verwendung von Algorithmen wird detailliert erklärt.

Kapitel 6 – Evaluation: Kapitel 6 beschreibt die Durchführung einer Usability Studie. Anhand dieser wird die Benutzerfreundlichkeit der getroffenen Gestenauswahl und somit der Gestenarten ermittelt. Der Ablauf der Studie wird durch das verwendete Vorgehensmodell erläutert. Es folgt die statistische Analyse der Studienergebnisse und die Präsentation dieser. Anschließend werden detaillierte Schlussfolgerungen zur Nutzbarkeit von zwei- und dreidimensionalen Gesten aus den Ergebnissen gezogen und Design Richtlinien für den Gestenentwurf und die Gesteneingabe aufgestellt.

Kapitel 7 – Zusammenfassung und Ausblick: Das letzte Kapitel fasst die Ergebnisse der Arbeit zusammen und stellt Anknüpfungspunkte vor. Unter anderem werden in diesem Kapitel die erreichten Resultate mit Blick auf die Problemstellung interpretiert sowie die Auswirkungen der gefundenen Ergebnisse auf den Bereich der Assistenzsysteme vorgestellt.

2. Forschungsstand

Dieses Kapitel gibt einen Einblick in den aktuellen Stand der Forschung. Verwandte Arbeiten und kommerzielle Systeme werden vorgestellt. Alle zum Verständnis dieser Arbeit notwendigen Grundlagen werden erläutert. Anschließend folgt ein Fazit über den momentanen Stand der Forschung sowie eine Abgrenzung zur vorliegenden Arbeit.

Die Erforschung assistierender Systeme am Arbeitsplatz gewinnt aufgrund des demografischen Wandels und wirtschaftlicher Interessen in den letzten Jahren an Bedeutung. Der Bereich der Assistenzsysteme ist aufgrund der vielfältigen Einsatzmöglichkeiten dieser Systeme ein multidisziplinäres Forschungsgebiet. Die Entwicklung roboterunterstützter assistierender Systeme, fernüberwachter assistierender Robotersysteme, als auch realitätserweiternder assistierender Systeme bedarf der Forschung in den Bereichen der Mensch-Computer-Interaktion (engl. HCI), der Mensch-Maschine-Interaktion (engl. HMI), der Computer Vision und Mustererkennung, der Computervisualisierung, der Computergraphik, der Robotik und anderen. Im Hinblick auf diese Arbeit und die untersuchten realitätserweiternden assistierenden Systeme am Arbeitsplatz sind besonders zwei Bereiche von Bedeutung, die Mensch-Computer-Interaktion und die Computer Vision und Mustererkennung. Die Mensch-Computer-Interaktion befasst sich mit der benutzergerechten Gestaltung von interaktiven Systemen. Mensch-Maschine-Schnittstellen stellen eine Verbindung zwischen Mensch und Maschine her und setzen zur Entwicklung Kenntnisse über die Maschine und das menschliche Verhalten voraus. Erkenntnisse über das menschliche Verhalten aus Psychologie, Arbeitswissenschaft, Kognitionswissenschaft, der Ergonomie und der Soziologie werden mit Erkenntnissen der Informatik verbunden. Die Gestenforschung, als Teilgebiet der Psychologie, Linguistik, Semiotik und Kognitionswissenschaft, wird mit dem Forschungsgebiet der Handgestenerkennung (engl. HGR), dem Informatik-Teilgebiet der Computer Vision und Mustererkennung, verbunden.

2.1. Assistenzsysteme

Assistierende Systeme am Arbeitsplatz werden seit Jahren in der Industrie und manuellen Montage eingesetzt. Die Aufgaben sind vielfältig an den Arbeitsplätzen, die durch manuelle Tätigkeiten geprägt sind. Konfektionieren, Entgraten, Bestücken und Montieren sind hier nur einige Beispiele. Von einfachen Systemen entwickelten sich diese Assistenzsysteme zu hochtechnologischen Systemen. Vor der rasanten Entwicklung des Computers und elektronischer Geräte konnte der Arbeiter mit Zeichnungen, Bilderkarten und Fotos, die am Arbeitsplatz in Buchform vorlagen oder an Wänden befestigt wurden, unterstützt werden. In den Folgejahren konnten gesprochenen Anweisungen zur Unterstützung herangezogen

werden, Handgeräte wie PDAs und Monitore konnten Bilder und Instruktionen wiedergeben (vgl. [SPH10]). Heute sind roboterunterstützte und fernüberwachte Systeme, als auch realitätserweiternde Systeme in der Industrie und manuellen Produktion im Einsatz (siehe Kapitel 1, Assistierende Arbeitsplätze in der Industrie). Sauer et al. stellen in [SPH10] fest, dass hochtechnologische assistierende Systeme gegenüber einfachen Systemen positive Effekte auf die Benutzer haben. Diese Effekte sind eine schnellere und präzisere Ausführung von Aufgaben sowie die gesteigerte Fähigkeit neue Aufgaben auszuführen. Nicht nur bei uneingeschränkten, sondern auch bei leistungsgewandelten und leistungsgeminderten Arbeitern sind diese Effekte zu beobachten (vgl. [KSH13b]). 87 Prozent befragter Unternehmen würden assistierende Systeme zur Steigerung des Wohlbefindens und der Leistungsmöglichkeiten von Arbeitern einsetzen (vgl. [Kor12]).

Die Entwicklung aktueller assistierender Systeme konzentriert sich auf die Ergebniskontrolle im Produktionsprozess, wobei die Interaktionsmöglichkeiten zwischen Mensch und Maschine bleiben weitgehend rudimentär (vgl. [KSH13a]). Die visuelle Überwachung und Anleitung der Bauteilentnahme (engl. *pick-by-light*) ist die am häufigsten vorkommende Interaktionsmöglichkeit in assistierenden Systemen zur manuellen Montage. Wie das Fraunhofer-Institut für Arbeitswirtschaft und Organisation in einer Studie feststellt (vgl. [BPHH11]), ist der *Touchscreen* die einzige modernere Interaktionstechnik die Einzug in die Arbeitsplätze der produzierenden Industrie gehalten hat. Häufig sind die genutzten Benutzerschnittstellen und Interaktionsmöglichkeiten für leistungsgewandelte und leistungsgeminderte Arbeiter nicht geeignet und erfüllen die speziellen Anforderungen dieser nicht, oder nur zum Teil.

2.1.1. Realitätserweiternde Systeme

Korn et al. stellen in ihrer Arbeit (vgl. [KSH13a]) fest, dass assistierende Systeme am Arbeitsplatz realitätserweiternde Techniken einsetzen müssen, um die Leistungsfähigkeit von Arbeitern zu steigern. Realitätserweiternde Technologien (engl. *Augmented reality*, kurz AR) können zur Lösung von Problemen in vielen Bereichen eingesetzt werden. Neben der Medizin, auch in der Unterhaltungsindustrie und der Fertigung. Unter realitätserweiternd wird die computergestützte Erweiterung der Realitätswahrnehmung verstanden. Diese Erweiterung der Wahrnehmung kann durch visuelle, als auch andere Sinnesreize wie Vibration oder Audiosignale erfolgen. Zum Beispiel können Videobilder einer Kamera visuell um computergenerierte Inhalte wie Animationen erweitert werden. In der manuellen Produktion, Fertigung und Montage werden zur Visualisierung von erweiternden Informationen hauptsächlich Datenbrillen und Bildschirme eingesetzt (vgl. [OYN08]). Neben Datenbrillen und Bildschirmen können auch Projektoren zur Darstellung von computergenerierten Informationen, welche die Realität erweitern sollen, eingesetzt werden. Im weiteren soll der Aufbau und der aktuelle Stand der Technik und Forschung projizierender Systeme diskutiert werden.

in-situ Projektion

Diese realitätserweiternden Systeme kombinieren Computervisualisierung und Projektionstechnik um eine immersive Arbeitsumgebung zu schaffen. Durch die direkte Projektion von

Informationen in den Arbeitsbereich (in-situ, lat. am Platz) wird es dem Arbeiter ermöglicht gezielt Arbeitsschritte auszuführen. Daten oder Informationen werden an genau dem Ort, an dem sie von Interesse sind, visualisiert. Lichtstarke Projektoren und neue Projektionstechniken, welche Reflexionen und Verzerrungen verhindern, werden eingesetzt.

Eine der ersten Arbeiten auf dem Gebiet der in-situ Projektion ist die Arbeit von Claudio Pinhanez aus dem Jahr 2001 (vgl. [Pin01]). Ein rotierender Spiegel lenkt die Projektion eines LCD/DLP Projektors auf verschiedene Oberflächen wie Möbel, Wände und Tische. Über Videokameras und Bilderkennungstechniken wird eine Interaktion mit den projizierten Bildern möglich. Die Gruppenarbeit in Konferenzräume sollte durch diese Technik verbessert werden. 2009 stellen Kane et al. in [KAW⁺09] eine robuste in-situ Projektionstechnik mit Interaktionsmöglichkeit vor. Micro-Projektoren projizieren hierbei an beide Seiten eines Notebooks interaktive Inhalte auf die Tischoberfläche. Micro-Kameras nehmen Handbewegungen auf, durch Bilderkennungstechniken sind *Multi-Touch* Gesten und Interaktionen mit projizierten Objekten möglich. Diese Arbeiten zur interaktiven Projektion zielen auf den Einsatz in Büros, der Unterhaltungsindustrie und mobilen Endgeräten. 2011 entwickeln Hörz et al. im Zuge des Forschungsprojekts ASLM¹ der Hochschule Esslingen ein realitätserweiterndes assistierendes System zur manuellen Montage in der Industrie für leistungsgewandelte und leistungsgeminderte Arbeiter (siehe Abbildung 2.1 (a)). In diesem Projekt werden technische Verfahren wie die Projektion und Bewegungserkennung eingesetzt, systematisch bewertet und in den Entwicklungs- und Erprobungsprozess integriert. Das System zeigt einen Montagearbeitsplatz, der mittels Bewegungserkennung die anwendungsorientierte Informationsbereitstellung von Text, Bild und Video direkt im Arbeitsbereich ermöglicht. Informationen werden hierbei mit Hilfe eines Projektors in den Arbeitsbereich projiziert. Neben einer anleitenden Funktion ist eine Qualitätskontrolle über die Erfassung von Arbeitsabläufen, Bewegungsvorgängen sowie der Bauteilpositionierung und -orientierung möglich. Die Raumüberwachung funktioniert auf Basis von Lichtgittern, diese werden mit Hilfe einer Microsoft² Kinect simuliert. Anhand der Auswertung erfasster 3D-Tiefendaten können Bewegungen des Arbeiters, als auch Bauteile sowie deren Position und Orientierung erkannt werden. Das System ist für die ergonomische Ausführung vor allem für leistungsgeminderte Mitarbeiter abgestimmt. 2013 stellen Korn et al. in ihren Arbeiten [KSH13b], [KSH13a] eine Weiterentwicklung des ALSM Systems vor. Dieses Prototypsystem ermöglicht dem Benutzer einfache Touch-Interaktionen innerhalb der Projektionsfläche. Hierbei berührt der Benutzer projizierte Objekte und löst zum Beispiel die Bestätigung eines Montagearbeitsschrittes aus, siehe Abbildung 2.1 (b).

Realitätserweiternde Systeme, die in-situ Projektionen einsetzen verbessern im ganzen die Fähigkeiten von uneingeschränkten Arbeitern (vgl. [RKWA10]). Korn et al. stellen jedoch in ihren Arbeiten [KSH13b], [KSH13a] einen katalytischen Effekt auf leistungsgeminderte Arbeiter fest. Die Arbeiter, die ein höheres kognitives Potential aufweisen, zeigen eine Steigerung der Leistungsfähigkeit und lösen eine Aufgabe der manuellen Montage schneller und

¹siehe <http://www.hs-esslingen.de/de/hochschule/fakultaeten/maschinenbau/forschung-und-transfer/forschungsprojekte/forschungsprojekt-aslm.html>

²siehe <http://www.microsoft.com>

2. Forschungsstand

qualitativ besser. Arbeiter, deren kognitives Auffassungsvermögen unter dem Durchschnitt liegt, können Projektionen nicht einordnen und zeigen eine schlechtere Leistung bei der Ausführung einer Aufgabe. Es ist anzunehmen, dass Arbeiter mit einem höheren kognitiven Potential schneller lernen und andere Arbeiter mit mehr Übung ebenfalls positivere Resultate erzielen können. Aufgrund dieser Ergebnisse kommen Korn et al. zur Erkenntnis, dass eine Skalierbarkeit des assistierenden Systems auf die Fähigkeiten des Arbeiters nötig ist (vgl. [KBS⁺12]).

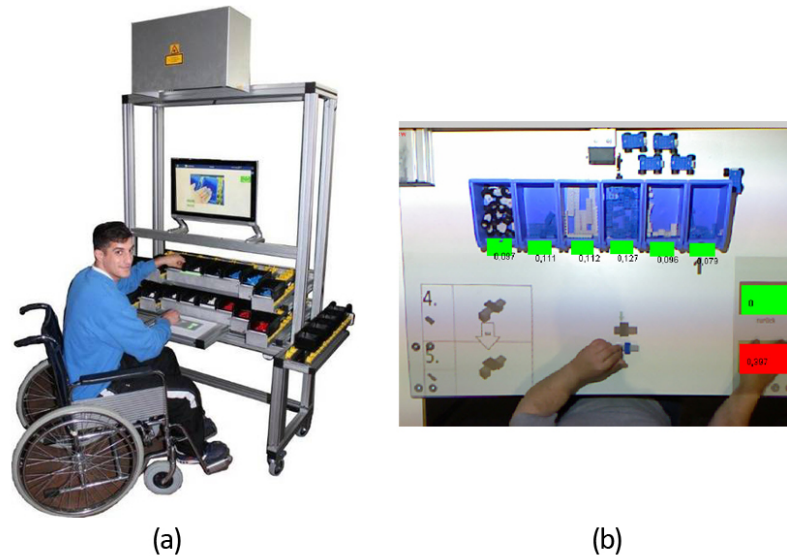


Abbildung 2.1.: Darstellung des ASLM Systems in Bild (a) und der erweiterten Interaktionsmöglichkeiten in Bild (b) des erweiterten ASLM Systems. Das rechte rote und grüne Rechteck reagieren auf Berührungen. (vgl. Forschungsprojekt ASLM und [KSH13b])

Gamification

In [Kor12] untersuchen Korn et al. Mechanismen zur Motivierung von leistungsgewandelten und leistungsgeminderten Arbeitern. Studien zeigen (vgl. [Kor12]), dass wiederkehrende und monotone Tätigkeiten, als auch die nicht optimale Auslastung von Arbeitern zu Langeweile, dem Gefühl geringer Wertschätzung, sinkender Konzentrationsfähigkeit und Motivation führen. Diese Folgen haben als Ergebnis eine steigende Fehlerrate in der manuellen Montage und einen allgemeinen Leistungsverlust. Gamification-Elemente können diese Tatsachen mildern und zu einer steigenden Motivation und weniger Fehlern in der Aufgabenausführung führen. Arbeiter werden aufmerksamer, konzentrieren sich und haben mehr Freude an einer Tätigkeit, dies hat positive Auswirkungen auf die Qualität der Produktion. Gamification-Elemente können aus Fortschrittsanzeigen, Levels, Punktesystemen sowie akustischen und visuellen Feedbacks bestehen. In [KBS⁺12] ähnelt die Gamification-Funktionalität dem Spieleklassiker

Tetris³ von Alexei Paschitnow, wobei die Spielesteine die auszuführenden Aufgaben symbolisieren und ihre Geschwindigkeit und Farbe der Leistung des Arbeiters entsprechen (siehe Abbildung 2.2). Gamification ist ein Überbegriff für die Nutzung von Videospielelementen zur Verbesserung der Benutzerfreundlichkeit und Benutzeraufmerksamkeit außerhalb des Spielbereichs. Das pädagogische Konzept des „Lernens durch Spielen“ wird hier auf die Durchführung von Arbeit übertragen und kann als „Arbeit als Spiel“ bezeichnet werden. Produktdesigns die zum spielen und intuitiven interagieren anregen sind im Alltag angekommen. Spiele sind ein Teil der Gesellschaft und finden sich in vielen Bereichen wieder. Leistungsgewandelte Menschen können durch ein spielerisches körperliches Training ihr aktives und gesundes Leben bis ins hohe Alter aufrecht erhalten. Leistungsgeminderte Menschen können durch ein spielerisches Training kognitive Fähigkeiten festigen und verbessern (vgl. [KBS⁺12]). Durch das Einsetzen von spielerischen Elementen zur Identifikation des Leistungsniveaus eines Arbeiters kann die Skalierbarkeit eines assistierenden Systems verbessert werden. Die Stärken eines Arbeiters, aber auch Schwankungen in dessen Tagesform können so vom System erkannt und genutzt werden.



Abbildung 2.2.: Darstellung von Gamification in einem Assistenzsystem. (vgl. [KSH13b])

Natürliche Bedienung

Korn et al. stellen in [KBS⁺12] fest, dass ein durchgängiger Arbeitsfluss, in dem die Fähigkeiten des Arbeiters sich den Schwierigkeiten der gestellten Aufgaben annähern, durch vier Faktoren beeinflusst wird. Durch klare Ziele und Aufgabenbeschreibungen, durch ein Gleichgewicht zwischen der Aufgabe und den eigenen Fähigkeiten, durch ein klares und direktes Feedback und durch das Gefühl, dass jede Aktivität lohnend ist und somit als leichter empfunden wird. Ein direktes Feedback hat die Steigerung der Prozessorientierung zur Folge, das System arbeitet im Hintergrund und greift bei Fehlern und bestimmten Aktionen unterstützend ein. Dieses Verhalten trägt ebenfalls zur Skalierbarkeit des Systems bei.

Voraussetzung für ein direktes Feedback des Systems auf eine Aktion des Benutzers ist eine Bewegungserfassung in Echtzeit. In [KSH12] erfasst der Prototyp eines assistierenden Systems über eine Microsoft Kinect Tiefen-Kamera die Bewegungen des Arbeiters im Arbeitsbereich.

³siehe <http://www.tetris.com>

Die Kinect Kamera ist im 40 Grad Winkel über dem Arbeitstisch angebracht und erfasst den Oberkörper. Obwohl die Kamera wegen ihrer geringen Tiefen-Auflösung nur grobmotorische Bewegungen und keine Hände und Finger erkennen kann, ist es durch die Softwareerweiterung OpenCV⁴ möglich Handbewegungen zu erfassen. Korn et al. stellen in dieser Arbeit fest, dass eine natürliche und intuitive Bedienung Voraussetzung für die Vereinfachung einer Benutzerschnittstelle ist. Ein *Natural User Interface*, kurz NUI, beschreibt eine Benutzerschnittstelle, die für den Benutzer natürlich wirkt. Dies kann durch Wischen, Tippen, Berühren, Gesten oder Sprache erreicht werden. Entgegengesetzt rein beobachtender Funktion (traditionelle Eingabemethoden wie Maus und Tastatur) bietet diese Benutzerschnittstelle die Möglichkeit der direkten Interaktion. Durch den Einsatz des Körpers als „Eingabegerät“ ermöglichen NUIs eine einfache und leicht erlernbare Bedienung, mit besonderem Potential für leistungsgewandelte und leistungsgeminderte Arbeiter. Hierbei steht vor allem die Bedienung durch Gesten und Körperbewegungen im Vordergrund.

Kommerzielle Systeme

Es gibt eine Vielzahl kommerzieller Systeme im Bereich der assistierenden Systeme für die manuelle Montage. Drei dieser Systeme stellen den letzten Stand der Technik dar und bieten eine Praxisanwendung.

Das „Light Guide System“, kurz LGS, vom US-amerikanischen Unternehmen OPS Solutions⁵ stellt eine Industrielösung zur manuellen Montage dar und zeigt ein assistierendes System mit in-situ Projektion (siehe Abbildung 2.3). Das LGS verwendet zur Visualisierung von Informationen einen industriellen DLP-Projektor mit Weitwinkellinse und einen Monitor. Das System ist modular aufgebaut und lässt sich an nahezu allen Arbeitsplätzen einsetzen. Informationen werden direkt in den Arbeitsbereich projiziert, dieser kann ein Arbeitstisch oder jede andere Oberfläche sein. Die Projektionsfläche kann dabei auf bis zu 17 Quadratmeter durch mehrere Projektoren erweitert werden. Informationen können Anweisungen, Texte, CAD-Bilder, Konstruktionszeichnungen, Videos mit Audioausgabe, Animationen, Timer und grafische Elemente wie Pfeile, Punkte und Zielscheiben sein. Zusätzliche Informationen können Zyklus-Zeiten und Schrittnummern sein. Die Prozessüberwachung kann manuell oder automatisiert erfolgen. Die automatische Überwachung kann mittels Drehmomentwerten, an den zum Einsatz kommenden Hilfswerkzeugen und Hilfsaggregaten, Lichtschranken sowie mittels Bilderkennungstechniken erfolgen. Bei manueller Überwachung kann nach Ausführung eines Arbeitsschrittes dieser durch ein Fußpedal, einen Schalter, einen Drehknopf oder per externem Touchbildschirm bestätigt werden. Erst durch eine korrekte Ausführung eines Arbeitsschrittes leitet das System den Benutzer zum nächsten Arbeitsvorgang. Anwendungsszenarien können vom Benutzer über eine spezielle Software programmiert werden. Das System kann zur Unterstützung der manuellen Montage, der Qualitätssicherung und zu Lern- und Übungszwecken eingesetzt werden. Es bietet nur eingeschränkte Interaktionsmöglichkeiten, diese beschränken sich auf externe Komponenten wie Schalter, Pedale, Maus,

⁴siehe <http://opencv.org>

⁵siehe <http://www.ops-solutions.com>

Tastatur und Touchbildschirme. Der Benutzer muss durch externe Bedienkomponenten mit dem System interagieren.

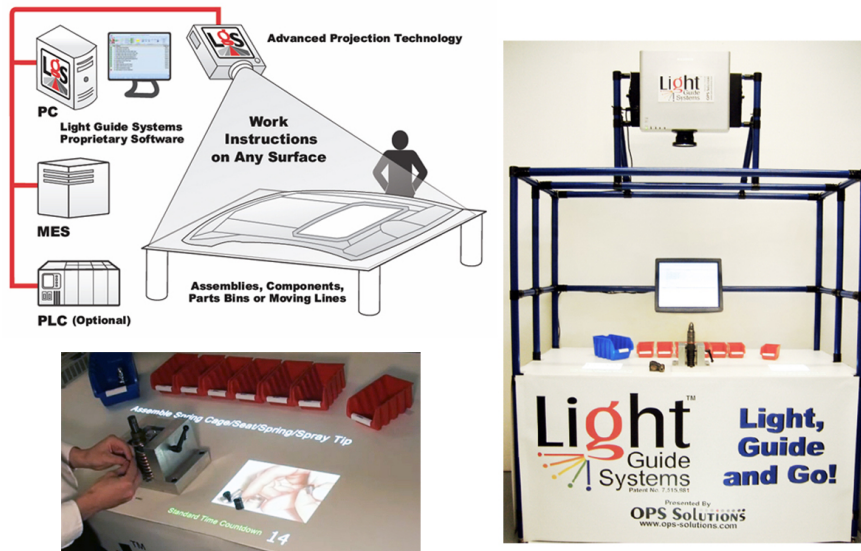


Abbildung 2.3.: LGS Assistenzsystem der Firma OPS Solutions. Das System kann in einem Produktionsleitsystem (engl. Manufacturing Execution System, kurz MES) betrieben werden, als auch speicherprogrammierbare Steuerung (engl. Programmable Logic Controller, kurz PLC) unterstützen. (vgl. Webauftritt OPS Solutions)

Das Schnaithmann⁶ Montage-Assistenzsystem auf Basis von Bewegungserkennung wurde anhand der Ergebnisse des ASLM Forschungsprojekts der Hochschule Esslingen entwickelt (siehe Abbildung 2.4). Das Montage-Assistenzsystem verwendet in-situ Projektionen und eine Bewegungserkennung zur frühzeitigen Fehlererkennung und Fehlervermeidung. Die Visualisierung von Informationen zur Aufgabenausführung erfolgt mittels hybrider Laser-LED-Technologie. Informationen werden direkt in den Aufgabenbereich projiziert, weitere Informationen werden zusätzlich auf einem Monitor gezeigt. Die Bauteilentnahme aus bis zu 16 Kleinteilebehältern wird während eines Arbeitsschrittes anhand von Lichtschranken (pick-by-light) erkannt. Zeitgleich wird die Position, Orientierung sowie die Identität des richtigen Bauteils im Montagebereich geprüft. Diese Prüfung geschieht durch die Erfassung und Auswertung von 3D-Tiefendaten und RGB-Werten. Optische und akustische Signale werden bei Fehlern und falschen Bewegungen innerhalb eines Arbeitsschrittes ausgegeben und fordern zum wiederholen der Bewegung oder des Arbeitsschrittes auf. Die Interaktion mit dem System erfolgt bei der Programmierung von Anwendungsszenarien über Maus und Tastatur. Bei der Aufgabenausführung erkennt das System den Abschluss einer Aufgabe

⁶siehe <http://www.schnaithmann.de/news/news-uebersicht/montage-assistenzsysteme>

2. Forschungsstand

automatisch. Der Benutzer wird durch die Arbeitsschritte geführt und hat keine Möglichkeit mit dem System aktiv zu interagieren.

Das ShowMe System der Firma PROFACTOR⁷ ist ein Montage-Assistenzsystem und verwendet in-situ Projektionen und eine Bewegungserkennung von Arbeitern (siehe Abbildung 2.5). Das System ist in Kooperation mit der Autoindustrie entwickelt worden, alleine bei einem einzigen Motorenmodell können zehn Produktvarianten mit jeweils unterschiedlichen Montagevorgaben vorhanden sein. Hier ist das ShowMe System für vielfältige Aufgaben wie Lern- und Übungszwecke, manuelle Montage von Motoren und Qualitätssicherung geeignet. Informationen werden mit Hilfe eines Projektors direkt in den Arbeitsbereich und auf Werkstücke projiziert. Eine Bewegungserkennung zur frühzeitigen Fehlererkennung und Fehlervermeidung wird durch zwei Kameras, welche 3D-Tiefendaten erfassen, realisiert. Die 3D-Tiefendaten der Kameras werden in Echtzeit mit protokollierten Arbeitsschrittdaten verglichen, optische und akustische Signale weisen den Arbeiter auf Fehler in der Ausführung eines Arbeitsschrittes hin. Bei der Aufgabenausführung erkennt das System den Abschluss einer Aufgabe automatisch. Der Benutzer wird durch die Arbeitsschritte geführt und hat keine Möglichkeit mit dem System aktiv zu interagieren. Anwendungsszenarien können vom Benutzer über eine spezielle Software programmiert werden.

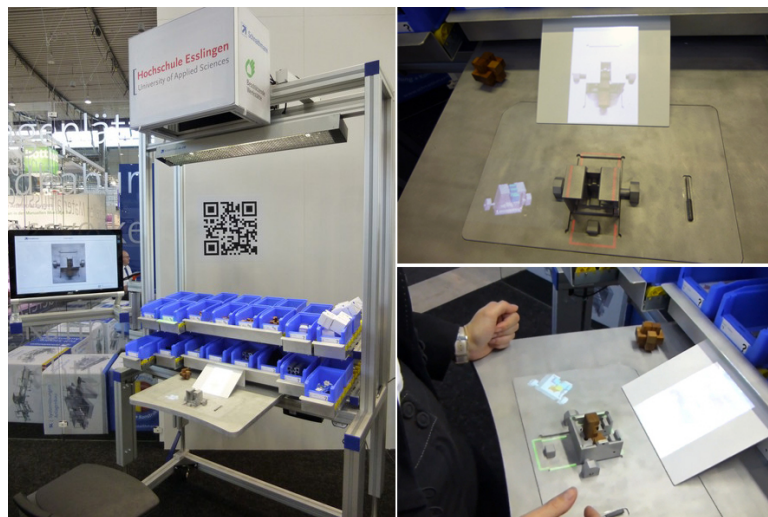


Abbildung 2.4.: Montage-Assistenzsystem auf Basis von Bewegungserkennung der Firma Schnaithmann. (vgl. Webauftritt Schnaithmann Maschinenbau GmbH)

2.1.2. Zusammenfassung

Assistierende Systeme mit in-situ Projektion verbessern die Leistungsfähigkeit nicht nur uneingeschränkter, sondern auch leistungsgewandelter und leistungsgeminderter Arbeiter.

⁷siehe <http://www.profactor.at>

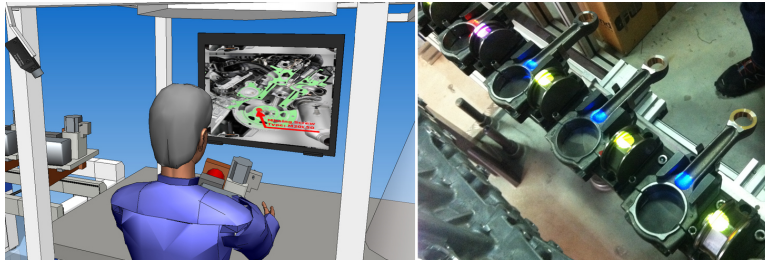


Abbildung 2.5.: Montage-Assistenzsystem der Firma PROFACTOR. (vgl. Webauftritt PROFACTOR)

Zusammenfassend werden folgende Anforderungen an diese Systeme gestellt: Die Steigerung der Prozessorientierung, vereinfachte Benutzerschnittstellen und die Bereitstellung von Mechanismen zur Motivierung. Aktuell erforschte und kommerzielle Systeme können diese Anforderungen nur teilweise erfüllen. Vor allem im Bereich der Benutzerschnittstellen und der Mensch-Computer-, Mensch-Maschine-Interaktion liegen diese Systeme zurück. Die Bedienung der Systeme erfolgt weitgehend durch externe Komponenten. Als einziges System erlaubt das weiterentwickelte ASLM System von Korn et al. (vgl. [KSH13a]) die Interaktion durch Berührung innerhalb der Projektionsfläche. Jedoch lässt dieses System nur eine rudimentäre Bedienung zu, es unterstützt keine Gesten wie Wischen und Ziehen. Eine Untersuchung natürlicher Interaktionsformen im Kontext assistierender Systeme mit in-situ Projektion in der manuellen Montage für leistungsgewandelte und leistungsgeminderte Arbeiter blieb bis heute aus.

2.2. Gesten und Gestenerkennung

Der Begriff der Geste ist ein sehr allgemeiner. Dieses Kapitel grenzt diesen Begriff ein, gibt einen Überblick wie Gesten entstanden sind, fasst Gestenformen zusammen und geht auf zwei- und dreidimensionale Gesten ein. Die Interaktion mit einem System durch Gesten macht eine Gestenerkennung und Auswertung unverzichtbar. Techniken der Gestenerkennung und Auswertung werden erklärt und beleuchtet. Abschließend wird auf die in dieser Arbeit eingesetzte Hardware zur Gestenerkennung eingegangen.

2.2.1. Konzeptionell

McNeill beschreibt eine Geste als „spontane oder bewusst eingesetzte Bewegung des Körpers, besonders der Hände und des Kopfes, die jemandes Worte begleitet oder ersetzt und eine bestimmte innere Haltung ausdrückt“ (vgl. [McN92]). In der Literatur wird der Begriff der Geste im Kontext der Mensch-Computer-Interaktion als allgemeingültiger Oberbegriff für „eine Bewegung des Körpers, die eine Information beinhaltet“ verwendet (vgl. [KH90]). Neben Handgesten, existieren Kopf- und Körpergesten. Anhand der zur Durchführung eingesetzten

Körperteile, kann eine Unterscheidung erfolgen. Kopfgesten können ein Nicken, ein Drehen, oder zur Seite Lehnen des Kopfes sein. Körpergesten können zum Beispiel eine Verbeugung sein. In Japan gibt es drei Arten der Verbeugung zur Begrüßung im Stehen und neun Arten der Verbeugung im Sitzen. Aufgrund der Relevanz von Handgesten für diese Arbeit werden im folgenden Verlauf diese ausführlich diskutiert.

Gesten und Sprache

Die Entstehung von Sprache ist Forschungsgebiet der Linguistik, der Wissenschaft um die menschliche Sprache. Semiotik ist Teilgebiet der Linguistik und die Wissenschaft, die sich mit Zeichensystemen aller Art befasst. Eine der populärsten Theorien der Linguistik über die Entstehung und Entwicklung von Sprache stellt Gesten an den Anfang der sprachlichen Evolution. Die fehlende Sprachfähigkeit affenähnlicher Vorfahren heutiger Menschen hatte die Kommunikation durch Hände und Urlaute zur Folge. Aufgrund der Anatomie des Sprachtraktes und ungeeigneter Formen von Mund, Zunge und Rachen entwickelte sich eine Gestensprache. Im Laufe der Evolution entstand aus dieser das heute gesprochene Wort. Die menschliche Hand ist ein flexibles Werkzeug und machte diese Entwicklung möglich.

Alle Menschen jeden Alters gestikulieren wenn sie sprechen oder etwas erklären. Handgesten sind natürliche, allgegenwärtige und bedeutsame Teile einer Sprache. Die enge Verbindung zwischen Sprache und Gesten ist so natürlich und umfassend, dass Wissenschaftler verschiedener Gebiete, von der Linguistik, der Psychologie, bis hin zur Neurowissenschaft sich eines integrierten Systems sicher sind. Während der Spracherzeugung und dem Sprachverständnis formt sich demnach ein enges einheitliches System. McNeill stellt fest, dass sich Sprache und Gesten zeitlich überlappen, jedoch Informationen auf zwei sehr verschiedenen Wegen transportieren (vgl. [McN92]). Sprache ist willkürlich und stilisiert, wobei Gesten bildhaft und charakteristisch sind. Diese zwei Arten Informationen in unterschiedlicher Form auszudrücken, reflektieren verschiedene Aspekte einheitlicher zugrunde liegender kognitiver Prozesse. Die Bedeutung einer Aussage besteht aus der Kombination von Sprache und Geste. Wird ein Teil dieser Aussage, die Sprache oder die Geste nicht beachtet, so ändert sich auch die Bedeutung der Aussage. Wird zum Beispiel versucht einem Kind nur durch gesprochene Anweisungen das Schnüren seiner Schuhe beizubringen, so funktioniert diese sehr schlecht. Werden Gesten zu gesprochenen Anweisungen gemacht, so ist die Aufgabe leicht.

Gesten sind die Gedanken eines Redners. Kleinkinder gestikulieren im Alter von 18 Monaten, bevor sie lernen zu sprechen. Blinde Kinder zeigen beim Erlernen einer Sprache Gesten, obwohl sie diese nie gesehen haben. Defizite in Wahrnehmung und Sprache äußern sich in direkter Verbindung zur Fähigkeit Gesten einzusetzen. Obwohl leistungsgeminderte Kinder mit Entwicklungsstörungen Schwierigkeiten haben beim gestikulieren, können Gesten kognitive und sprachliche Defizite ausgleichen. Kinder mit Downsyndrom und spezifischen Sprachentwicklungsstörungen zeigen häufiger und mehr Gesten als normal entwickelte Kinder und können so sprachliche Defizite ausgleichen. Gesten werden auch zur Diagnose von Autismus herangezogen. Schon Kleinkinder im Alter von zwölf Monaten zeigen bei Veranlagung zu Autismus eine veränderte Gestikulation. Bemerkenswert ist die Tatsache, dass dieser

Indikator sechs Monate bevor die Kinder mit dem Sprechen beginnen, den Hinweis auf den Schweregrad der autistischen Ausprägung gibt.

Gesten können die kognitiven Fähigkeiten verbessern und Verständnisprobleme verringern. Teilnehmer einer Studie erinnern sich an gemerkte japanische Vokabeln besser, wenn sie während der Lernphase zu den Vokabeln von der Bedeutung her passende Gesten zeigen (vgl. [KMR08]). Dieser Effekt geht jedoch verloren, wenn der semantische Kontext einer Geste nicht zum Kontext von Wörtern passt.

Gestenformen in Linguistik und Semiotik

Nach McNeill (vgl. [McN92]) werden Gesten in zwei Kategorien unterteilt. Die erste Kategorie bezeichnet **„nebensprachliche“ Gesten** (engl. *co-speech gestures*), die zweite Kategorie **lexikalisierte, oder „Emblem“ Gesten** (engl. *emblem gestures*). Nebensprachliche Gesten werden unbewusst gezeigt, während gesprochen wird und sind charakteristisch für die sprechende Person. Emblem Gesten sind Handstellungen, die je nach kulturellem Hintergrund verschiedene Bedeutungen haben und Aussagen ergänzen oder substituieren. Diese Gesten transportieren Informationen unabhängig von der Sprache.

Nebensprachliche Gesten werden in vier Typen unterteilt:

- **Ikonische Gesten** (engl. *iconic gestures*), bilden Objekteigenschaften, Handlungen und räumliche Beziehungen ab. Die Referenzierung kann hierbei auch auf nicht reale Objekte und Metaphern erfolgen. Objekteigenschaften können Formen und Umrisse sein. Beispielsweise formen Hände eine nach oben offene Halbkugel, als Repräsentation einer Schale. Handlungen können aus dem Blättern eines Buches bestehen und durch ein Aufklappen beider Hände repräsentiert werden. Beispiele räumlicher Beziehungen können Etagen eines Gebäudes sein und durch übereinander gehaltene Hände repräsentiert werden.
- **Deiktische Gesten** (engl. *deictic or pointing gestures*), verbinden einen Aspekt einer gesprochenen Aussage mit einer anderen Idee, einer Meinung, einem Objekt, einer Handlung oder einem Ort. Zeigegesten sind deiktische Gesten. Als Beispiel kann das Zeigen auf ein reales oder imaginäres Objekt genannt werden.
- **Metaphorische Gesten** (engl. *metaphoric gestures*), transportieren eine abstrakte Idee in konkreter Art und Weise. Als Beispiel kann hier die Metapher „ein Licht geht auf“ und die zugehörige Hand- und Armbewegung genannt werden. Die Bewegung enthält das Halten eines Fingers am Kopf mit dem darauffolgenden hoch halten des Fingers in die Luft und über den Kopf.
- **Rhythmische Gesten** (engl. *beat gestures*), enthalten den Rhythmus einer gesprochenen Aussage. Diese Gesten transportieren keine Informationen, verbinden jedoch Teile von Aussagen. Rhythmische Gesten betonen und unterstreichen. Beispielsweise wenn Eltern einem Kind zum wiederholten Mal etwas sagen und den Finger dabei rhythmisch hoch und runter bewegen.

Gesten Taxonomien in der Mensch-Computer-Interaktion

In der Linguistik und Semiotik bewegt sich die Klassifizierung von Gesten meist auf einem hohen Abstraktionslevel und nicht im Kontext der Mensch-Computer-Interaktion. Gestenformen der Linguistik haben zum Beispiel keine Eigenschaft für die Geschwindigkeit der Ausführung einer Geste. Eventuell ist diese aber für die Ausführung einer Geste in einem System wichtig. Ohne Taxonomien, die solche Eigenschaften einbeziehen, fehlen Klassifizierungsgrundlagen für Gesten in der Mensch-Computer-Interaktion. Karam et al. beziehen sich als erste in ihrer Arbeit [KS05] auf eine Klassifizierung von Gesten im Sinne der Mensch-Computer-Interaktion.

In der Mensch-Computer-Interaktion gibt es eine Reihe von Taxonomien, die Gestenformen der Linguistik erweitern, ausbauen und an verschiedene Systeme anpassen. Eine Taxonomie stellt eine Reihe von Gestentypen vor und ordnet diesen bestimmte Eigenschaften zu. Mit Hilfe dieser Eigenschaften können Gesten der Mensch-Computer-Interaktion entwickelt und zu festgelegten Gestentypen zugeordnet werden. Unter Beachtung von Taxonomien können Gestenerkennungstechniken entwickelt und verbessert werden.

Wobbrock et al. stellen in [WMW09] eine **Taxonomie für zweidimensionale Gesten** (kurz 2D-Gesten, engl. *surface gestures*) für die Bedienung von berührungsempfindliche Oberflächen (engl. *surface computing*) vor. Diese Taxonomie teilt Gesten in vier Dimensionen und verschiedene Kategorien ein, siehe Tabelle 2.1.

Taxonomie zweidimensionaler Gesten		
Form	statische Pose	Handpose gehalten auf einer Position
	dynamische Pose	Handpose ändert sich auf einer Position
	statische Pose auf Pfad	Handpose gehalten bei Bewegung auf Pfad
	dynamische Pose auf Pfad	Handpose ändert sich bei Bewegung auf Pfad
	ein-Punkt Berührung	Statische Pose mit einem Finger
	ein-Punkt Pfad	Statische Pose auf Pfad mit einem Finger
Natur	symbolisch	Geste wird als Symbol erkannt
	physikalisch	Geste agiert physisch mit Objekt
	metaphorisch	Geste zeigt eine Metapher
	abstrakt	Zuordnung Geste zu Aktion ist abstrakt
Bindung	objekt-zentriert	Position-Def. bzgl. Objekteigenschaft
	umgebungs-abhängig	Position-Def. bzgl. Umgebung
	umgebungs-unabhängig	Position-Def. kann Umgebung ignorieren
	gemischte Abhängigkeit	umgebungs-unabhängig plus weitere
Fluß	diskret	Reaktion erfolgt nach Benutzeraktion
	kontinuierlich	Reaktion erfolgt während Benutzeraktion

Tabelle 2.1.: 2D-Gesten Taxonomie von Wobbrock et al. (vgl. [WMW09])

Die Form Dimension (engl. *form*) bezieht sich auf Gesten mit einer Hand, wobei die Kategorien „ein-Punkt Berührung und Pfad“ (engl. *one-point touch, path*) auch auf Berührungen

eines Punktes mit mehreren Fingern zutreffen. Die Kategorie „symbolisch“ (engl. *symbolic*) der Natur Dimension (engl. *nature*) bezieht sich auf Symbole wie das „Ok“ mit Daumen und Zeigefinger und gespreizten restlichen Fingern. „Metaphorische“ (engl. *metaphorical*) Gesten zeigen Metaphern. Zum Beispiel formt die Hand ein Lupe als Vergrößerungsgeste und bewegt sich dementsprechend. „Abstrakte“ (engl. *abstract*) Gesten haben keine andere Kategoriezugehörigkeit in der Natur Dimension. Zum Beispiel ist das mehrmalige Berühren (engl. *tap*) eines Objektes um es zu löschen abstrakt. Die Dimension der Bindung (engl. *binding*) fasst Gesten zusammen, die Informationen im Bezug auf ein Objekt oder die Umgebung zur Ausführung brauchen. Eine Beispiel für „objekt-zentrierte“ (engl. *object-centric*) Gesten wäre das Vergrößern von Objekten mit Daumen und Zeigefinger durch entgegengesetzte Bewegung dieser (engl. *pinch to zoom*). „Umgebungs-unabhängige“ (engl. *world-independent*) Gesten brauchen zur Ausführung keine Informationen der Umgebung und können überall stattfinden. Gesten mit „gemischten Abhängigkeiten“ (engl. *mixed dependencies*) sind zum Beispiel mit zwei Händen ausgeführte Gesten (engl. *bimanual gestures*), wobei eine Hand „umgebungs-unabhängig“ und die andere „objekt-zentriert“ wäre. Eine „diskrete“ (engl. *discrete*) Geste ist zum Beispiel die Zeichnung eines Fragezeichen als Aktion für Hilfe.

Taxonomie dreidimensionaler Gesten		
Gesten Mapping		
Natur	metaphorisch	Geste eine Metapher eines physischen Objekts
	physikalisch	Geste agiert physisch mit Objekt
	symbolisch	Geste wird als Symbol erkannt
	abstrakt	Zuordnung Geste zu Aktion ist abstrakt
Kontext	im-Kontext	Geste fordert Kontext
	nicht-Kontext	Geste fordert kein Kontext
Temporal	diskret	Reaktion erfolgt nach Benutzeraktion
	kontinuierlich	Reaktion erfolgt während Benutzeraktion
Physikalische Eigenschaften		
Kinematisch	niedrig	Geste mit Änderung der Geschw. unter $3m/s^3$
	mittel	Geste mit Änderung der Geschw. zw. $3m/s^3$ und $6m/s^3$
	hoch	Geste mit Änderung der Geschw. über $6m/s^3$
Dimension	1-Achse	Bewegung um eine Achse
	3-Achsen	Bewegung enthält Translation oder Rotation, nicht beide
	6-Achsen	Bewegung enthält Translation und Rotation
Komplexität	einfach	Geste enthält nur einzelne Geste
	verbunden	Geste kann in einfache Geste geteilt werden

Tabelle 2.2.: 3D-Gesten Taxonomie von Ruiz et al. (vgl. [RLL11])

In der Arbeit [RLL11] stellen Ruiz et al. eine **Taxonomie für dreidimensionale Gesten** (kurz 3D-Gesten, engl. *motion, mid-air gestures*) für die Bedienung von *Smartphones* vor. Die Taxonomie teilt Gesten in sechs Dimensionen ein und gruppiert diese sind in zwei Klassen (siehe Tabelle 2.2).

Die Klasse des Gesten Mappings (engl. *gesture mapping*) beschreibt wie Benutzer die Zuordnung von Gesten zu Aktionen wählen. Die Klasse der physikalischen Eigenschaften (engl. *physical characteristics*) beschreibt technische und physikalische Eigenschaften von Gesten. Die Natur Dimension entspricht Wobbrocks Definition in der Taxonomie für 2D-Gesten. Eine Geste, die der Kategorie „im-Kontext“ (engl. *in-context*) zugeordnet wird, benötigt einen Kontext zur Ausführung. Zum Beispiel löst das Führen eines Smartphones zum Kopf die Aktion des Wählens aus. Eine Geste, die „nicht-Kontext“ (engl. *no-context*) behaftet ist, ist zum Beispiel das Schütteln des Smartphones zum Zurückkehren in das Hauptmenü. Die Kinematische Dimension (engl. *kinematic impulse*) gibt die Änderungsrate der Geschwindigkeit einer Geste bei Ausführung an. Die Geste eines kräftigen Schüttelns würde in die Kategorie „hoch“ (engl. *high*) eingeordnet. Die Dimension Dimension (engl. *dimension*) beschreibt die genutzten Achsen, die während der Ausführung einer Geste gebraucht werden. Ist in einer Geste eine Rotation des Smartphones vorgesehen, so ordnet sich diese Geste in die Kategorie „3-Achsen“ (engl. *tri-axis*) ein.

Eine weitere Taxonomie für dreidimensionale Gesten geben Aigner et al. in ihrer Arbeit [AWB⁺12] an. Diese teilt Gesten in fünf Dimensionen ein: Zeigegesten (engl. *pointing*), Zeichen- oder Symbolgesten (engl. *semaphoric*), pantomimische Gesten (engl. *pantomimic*), ikonische Gesten (engl. *iconic*) und manipulative Gesten (engl. *manipulation*). Zeichen- oder Symbolgesten werden in statische („Ok“ Zeichen), dynamische (Kreisbewegung der Hand für eine Rotation) und Wisch- und Streich-Gesten (einfaches Wischen mit dem Finger oder der Hand) eingeteilt. Pantomimische Gesten ahmen eine Handlung nach, wie zum Beispiel das Füllen eines Wasserglases. Ikonische Gesten entsprechen der Definition der Gestenform der Linguistik. Manipulative Gesten definieren Gesten, die während der Ausführung eine Teilaktion auslösen.

Das Feld der Taxonomien für Gesten ist breit gefächert. Von Taxonomien für Gesten zur Bedienung von „*Pubic Displays*“ bis hin zu Taxonomien für Gesten, die im Auto eingesetzt werden, gibt es eine Fülle von Klassifizierungen. Unterschiede in diesen Klassifizierungen ergeben sich aufgrund von vielen Faktoren, wie dem Einsatzgebiet, den Benutzern und den zu lösenden Aufgabentypen. Taxonomien für 2D-Gesten können nicht einfach auf 3D-Gesten übertragen werden. Aufgrund verschiedener Ausführung und Ausführungsmöglichkeiten sowie verschiedener Einsatzgebiete dieser Gestentypen entstehen unterschiedliche Taxonomien. Zu bemerken ist, dass keine Arbeiten zu Taxonomien für Hybrid-Gesten, der Verbindung von 2D- und 3D-Gesten, existieren.

Natural User Interfaces

Unter einem Interaktionsdesign wird nicht das Erscheinungsbild, sondern das Verhalten eines Produktes bei seiner Benutzung verstanden (vgl. [Saf09]). Natural User Interfaces, kurz NUIs, beschreiben Benutzerschnittstellen, die Benutzern das Gefühl geben sollen eine natürliche Person zu sein. Nicht die Schnittstellen sind natürlich, sondern die Benutzer sollen sich im Umgang mit diesen natürlich fühlen (vgl. [WW10], S. 14). Benutzer sollen sich nicht in Technologie verfangen sehen, sondern leicht und intuitiv Schnittstellen benutzen und erlernen

können. Sprache, Berührungen und Gesten machen den Benutzer zum Eingabegerät und lassen ihn ein Teil eines Ganzen werden.

Paradigmen und Metaphern herkömmlicher Benutzerschnittstellen und Interaktionssysteme müssen überdacht und erneuert werden. Die Interaktion mit dem Körper, vor allem der menschlichen Hand, stellt ein Benutzerschnittstellendesign vor Herausforderungen. Die Aufmerksamkeit sollte auf die Art der Interaktion, die technischen und physischen Möglichkeiten ein System zu bedienen, liegen. Die Erweiterung der Interaktionsmöglichkeiten um die dritte Dimension lässt ein System für den Benutzer greifbar werden.

Die Entwicklung von NUIs beinhaltet die Erforschung von Gesten, da diese ein wichtiges Interaktionsinstrument darstellen. Die Entwicklung von Gesten und Gesten-Sets für verschiedene Gebiete der Nutzung befindet sich in den Anfängen. Im Kontext der Mensch-Computer-Interaktion haben Gesten großes Potential den Umgang mit Technologie positiv zu verändern.

Die Arbeiten zur Erforschung und Entwicklung von Gesten und Gesten-Sets lassen sich in zwei Lager teilen. Einmal in die Sichtweise, dass Gesten-Sets aus dem Verhalten der Benutzer durch Studien hergeleitet werden sollen (vgl. [WMW09], [RLL11]). Hierbei entstehen sogenannte vom Benutzer abgeleitete Benutzerschnittstellen (engl. *User-Derived Interfaces*). Aus anderer Sichtweise wird die Herleitung von Designprinzipien für Gesten und Gesten-Sets aus der Untersuchung bestehender Eingabegeräte und deren Rahmenbedingungen vorgeschlagen (vgl. [WW10]). Die physischen und psychischen Auswirkungen während der Nutzung dieser Eingabegeräte werden erfasst und aus diesen Schlussfolgerungen für die Entwicklung von Gesten und Gesten-Sets gezogen. Nicht nur die Benutzer werden in den Designprozess mit einbezogen, sondern auch Faktoren um das System und die Umgebung. Durch dieses Vorgehen entstehen Gesten-Sets die bestimmte Kriterien und Anforderungen der Nutzer erfüllen, leicht zu erlernen und intuitiv zu nutzen sind. Somit eröffnet Gestensteuerung Möglichkeiten, die andere Arten der Nutzerinteraktion nicht leisten können.

Gesten leistungsgeminderter und leistungsgewandelter Menschen

Die vorgestellten Taxonomien beziehen sich auf Nutzer ohne körperliche Einschränkungen. Bei Menschen mit körperlichen Einschränkungen müssen Gesten und Klassifizierungen dieser nach anderen Kriterien erfolgen. Leistungsgeminderten Menschen fehlen oft Gliedmaßen, sie haben durch Spasmen entstellte Hände oder führen Bewegungen mit einem zittern durch. Diese Behinderungen müssen bei der Konzeption von Gesten berücksichtigt werden. Die Arbeit von Aly (vgl. [Aly13]) zeigt Interaktionsschwierigkeiten leistungsgeminderter Arbeiter an einem Assistenzsystem mit berührungsempfindlicher Oberfläche. Hier treten Probleme bei Handstellungen sowie der Ausführung von feinmotorischen Bewegungen auf.

2.2.2. Technisch

Die technische Umsetzung einer Gestensteuerung ist Gegenstand einer Vielzahl von Arbeiten. Eine Ursache hierfür kann in der Diversität und dem Potential der zugrundeliegenden

Technologien gesehen werden, deren Entwicklung nur schwer abzuschätzen ist. Aus der Tatsache, dass sich die Mehrzahl der Arbeiten auf spezielle Applikationen beschränkt, lässt sich ableiten, dass auch verschiedenen Gestenerkennungstechniken eingesetzt werden.

Techniken und Technologien zur Gestenerkennung bieten verschiedene Möglichkeiten und besitzen Einschränkungen. Im Folgenden werden verschiedenen Techniken und Technologien der Gesten- und Handgestenerkennung vorgestellt. Weiter wird ein Überblick über die eingesetzte Hardware zur Gestenerkennung gegeben.

Verfahren zur Gestenerfassung

Markergestützte Systeme erfassen Bewegungen durch angebrachte Sensoren am menschlichen Körper. Das *Motion Capture* Verfahren ist vielleicht das bekannteste Verfahren unter den markergestützten Systemen. Bewegungen werden in diesem Verfahren durch aufgebrachte Punkte am Körper erfasst. Kameras nehmen diese Punkte auf, Bilderkennungstechniken extrahieren Bewegungsdaten aus diesen Aufnahmen und übertragen diese auf ein computer-generiertes Modell. Es können auch am Körper angebrachte Sensoren wie Gyroskope und Accelerometer zur Erfassung von Bewegungen genutzt werden. Markergestützte Systeme haben den Nachteil, dass der Benutzer Sensoren tragen muss und eventuell durch diese in seiner Bewegung eingeschränkt wird. Datenhandschuhe verdeutlichen diese Diskrepanz.

Bilderfassungssysteme nehmen Bewegungen des menschlichen Körpers in Form von Bildern auf und leiten diese zur Bewegungserkennung weiter. Diese Systeme setzen Kameras ein. Auch werden Tiefenblinder zur Erkennung von Bewegungen genutzt. Der Vorteil dieser Systeme ist das Wegfallen von Sensoren die am Körper getragen werden müssen. Dies macht eine natürlichere Interaktion möglich. Nachteilig ist jedoch, dass Störungen im Bild oder der Bildaufnahme negative Auswirkungen auf spätere Bewegungserkennung durch Bilderkennungstechniken haben.

Verfahren zur Gestenerkennung

Aufgrund dem Fokus auf Handgesten in dieser Arbeit werden im weiteren Verlauf Verfahren der Handgestenerkennung (engl. *Hand gestures recognition*, kurz HGR) diskutiert. Handgestenerkennung ist ein sehr aktives Forschungsfeld in der Computer Vision und Mustererkennung. Gesten können alle Handbewegungen von einem Menschen sein, jedoch haben nur bestimmte Bewegungen eine Bedeutung. Die menschliche Hand kann nahezu in jede beliebige Richtung verdreht und um fast jeden beliebigen Winkel gestreckt werden. Finger können eingeklappert oder nur halb angewinkelt sein. Diese Eigenschaften machen die Hand zu einem der kompliziertesten Körperteile des Menschen.

Verfahren der Handgestenerkennung umfassen statistische Modelle, Bild- und Mustererkennung, Bildverarbeitung und konnektionistische Systeme. Statistische Modelle sind Hidden Markov Modelle (engl. *hidden Markov model*), Hauptkomponentenanalysen (engl. *Principal*

Component Analysis, kurz PCA) und Partikel-Filter (engl. *particle filters*). Konnektionistische Systeme sind künstliche neuronale Netze (engl. *artificial neural networks*, kurz ANNs), *genetische Algorithmen* und *Fuzzy Logic*. Bildverarbeitungstechniken zielen auf die Bildsegmentierung (engl. *image segmentation*) sowie die Analyse und Erkennung von Formen, Oberflächen, Bewegungen und Konturen. Bild- und Mustererkennungstechniken werden zur Extraktion von Bildeigenschaften, Objekterkennung, Clusteranalyse und Klassifikation von Mustern eingesetzt. Zu diesen Techniken gehören Verfahren wie das *Template Matching*, was zu Deutsch der Vorlagenübereinstimmung entspricht.

Template Matching Verfahren beruhen auf Wahrscheinlichkeiten, Statistiken und Signalverarbeitung. Grob erklärt, wird anhand einer Vorlage, bestehend aus einer Aufzeichnung von Daten, versucht diese in einem Strom von Daten zu erkennen. Die Erkennung von statischen Gesten kann durch Template Matching Verfahren und künstliche neuronale Netze erfolgen (vgl. [MA07]). Wobbrock et al. stellen in [WVA12] ein Template Matching Verfahren zur Erkennung von zweidimensionalen Gesten vor. Dieses nutzt eine aufgenommene Geste als Vorlage, bestehend aus einer Punktwolke von Bewegungsdaten eines Fingers, und erkennt anhand dieser Daten die aufgenommene Geste in einer Bewegung. Dieses Verfahren ist Teil der \mathcal{S} -Verfahren (vgl. [WVA12]), einer Familie von Template Matching Verfahren, welche leicht zu verstehen und zu implementieren sind und auf geometrischen Berechnungen beruhen.

Die Erkennung von dynamischen Gesten kann durch *Dynamic time warping* Verfahren, Hidden Markov Modelle und künstliche neuronale Netze erfolgen. Dynamic time warping Verfahren ermöglichen die Erkennung von Mustern in zeitlichen Sequenzen von Daten. Hidden Markov Modelle, als auch künstliche neuronale Netze bilden etablierte, jedoch auch sehr komplexe Verfahren zur Gestenerkennung. Das Hidden Markov Modell ist die einfachste Form eines dynamischen bayesschen Netzes (engl. *dynamic bayesian network*). Bayes-Netze bilden ein Problem durch eine Menge von Zufallsvariablen ab. Jede Zufallsvariable korrespondiert mit genau einem Netzknoten. Kausale Zusammenhänge zwischen den einzelnen Zufallsvariablen werden mit Kanten zwischen den Knoten modelliert. Für jeden Knoten wird eine bedingte Wahrscheinlichkeitsverteilung hinterlegt, welche die Berechnungsvorschrift für die korrespondierende Zufallsvariable aus den benachbarten Netzknoten liefert. Somit sind Bayes-Netze eine graphische Notation für Aussagen über bedingte Unabhängigkeit und damit ein praktischer Weg, gemeinsame Wahrscheinlichkeitsverteilungen zu spezifizieren. Dynamischen Bayes Netze beinhalten zeitabhängige Variablen und es werden kausale Abhängigkeiten über mehrere Zeitschritte modelliert. Künstliche neuronale Netze übernehmen das biologische Modell neuronaler Netze im menschlichen Gehirn. Die Struktur eines künstlichen neuronalen Netzes ist durch die Anzahl der Neuronen, durch eine besondere Topologie und durch die Art der Aktivierungsfunktion der Neuronen definiert. Faktoren wirken auf Neuronen ein und aktivieren diese je nach Bedingung. Künstliche neuronale Netze werden trainiert, sodass eine Verkettung von Faktoren, Konstanten und Gewichten einen Zusammenhang zwischen Neuronen entstehen lässt. Durch dieses trainierte Wissen lassen sich Probleme verschiedener Art lösen.

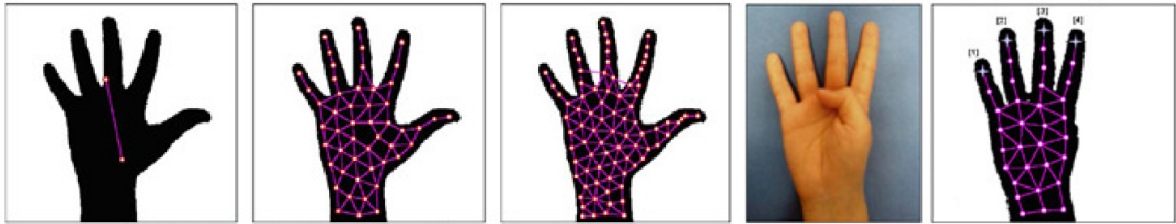


Abbildung 2.6.: Gestenerfassung durch künstliche neuronale Netze. Beginn einer Handgestenerfassung mit zwei Neuronen, 45 Neuronen, 83 Neuronen für die ganze Hand und die Erkennung der Geste mit vier Fingern. (vgl. [CRDR11])

Hardware zur Gestenerfassung

Microsofts⁸ Kinect Tiefenkamera erfasst neben Farbwerten auch Tiefeninformationen. Ein Infrarotsensor projiziert ein großflächiges Infrarot-Punkt-Muster in den Raum, ein schwarzweiß CMOS Bildsensor empfängt reflektierte Infrarotstrahlen. Aus der Laufzeit jedes reflektierenden Punktes werden Tiefenwerte errechnet. Jedes Pixel eines Bildes erhält einen Entfernungswert. Stark reflektierende oder absorbierende Flächen können zu Fehlmessungen führen. Ein weiterer Nachteil der Kinect Kamera ist, dass die Erfassung von Tiefenwerten erst ab einer Entfernung von 50 Zentimeter möglich ist. Dieser Nachteil lässt eine Bewegungserkennung direkt an der Kamera nicht zu.

Die Leap Motion⁹ ist ein Gerät zur hochauflösenden und schnellen Handerkennung. Das Gerät enthält zwei Kameras mit monochromen CMOS Bildsensoren, welche Infrarotstrahlen empfangen, sowie drei Infrarot LEDs. Diese Infrarot LEDs spannen durch Projektion ein fast halbkugelförmiges Infrarot-Punkt-Muster mit einem Meter Radius auf. Reflektierende Punkte innerhalb dieser Halbkugel werden durch die zwei Kameras mit 290 Bildern pro Sekunde erfasst. Durch Bildererkennungstechniken werden diese Tiefenbilder verglichen und ein dreidimensionales Modell einer Hand erstellt. Im Schnitt werden diese Handdaten 215 Mal pro Sekunde aktualisiert, wobei eine Bewegungsgenauigkeit von einem Zehntelmillimeter erreicht wird. Nachteilig können sich jedoch stark reflektierende oder absorbierende Flächen, als auch helle Lichtquellen auswirken und zu Fehlmessungen führen.

2.2.3. Zusammenfassung

Menschliche Gesten sind komplexe Gebilde aus physischen und psychischen Komponenten. In der Linguistik und der Mensch-Computer-Interaktion müssen Gesten auf verschiedene Art und Weise klassifiziert werden. Je nach System, Benutzer und Rahmenbedingung fallen Taxonomien für Gesten unterschiedlich aus. Natürliche Interaktion umfasst Gesteninteraktion,

⁸siehe <http://www.microsoft.com>

⁹siehe <https://www.leapmotion.com>

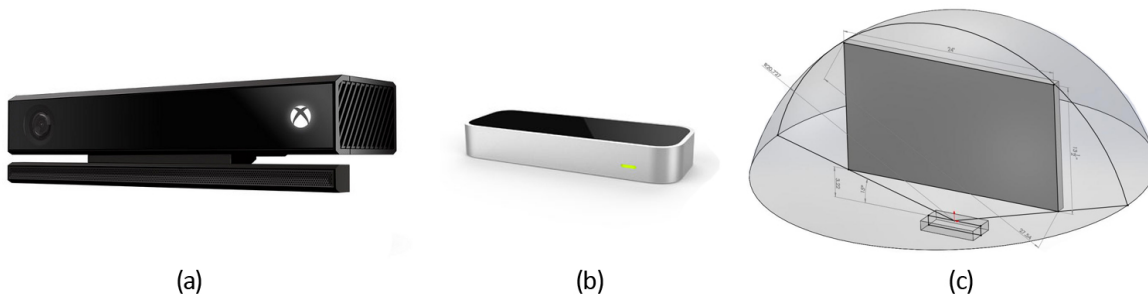


Abbildung 2.7.: Darstellung von Systemen zur Gestenerfassung. Links die Microsoft Kinect Kamera (a), in der Mitte die Leap Motion (b) und rechts eine Skizze des Erfassungsbereichs der Leap Motion. (vgl. Webauftritt Microsoft, Leap Motion)

wobei Gesten und Gesten-Sets durch verschieden Ansätze aus dem Verhalten von Benutzern und den Rahmenbedingungen der untersuchten Systeme extrahiert werden. Leistungsgeminderte und leistungsgewandelte Menschen haben bei der Ausführung von Gesten verschiedene Schwierigkeiten. Die Konzeption von Geste und Gesten-Sets sollte die motorischen Fähigkeiten von Benutzern mit einbeziehen. Etablierte technische Verfahren zur Gestenerkennung sind komplex und in der Realisierung aufwendig. Neben diesen existieren Template Matching Verfahren der δ -Familie, die in ihrer Anwendung mit vertretbarem Aufwand auskommen. Im Verlauf dieser Arbeit wird auf das Template Matching Verfahren von Wobbrock et al. (vgl. [WVA12]) näher eingegangen.

2.3. Fazit

Bestehende Assistenzsysteme in der manuellen Montage erfordern vereinfachte Benutzerschnittstellen. Besonders leistungsgewandelten und leistungsgeminderten Arbeitern soll durch diese der Zugang zum System erleichtert werden. Eine einfache und natürlich Interaktion kann durch Gesten erfolgen. Die Entwicklung dieser muss auf die Gegebenheiten der Arbeiter und deren Bedürfnisse abgestimmt sein. Die ausführliche Analyse bestehender Assistenzsysteme in der manuellen Montage und die Betrachtung von Gesten, Gesten-Sets sowie Taxonomien lassen den folgenden Schluss zu. Bis auf wenige Ansätze gibt es keine ausführlichen Untersuchungen zur Interaktion durch Gesten an assistierenden Systemen am Arbeitsplatz im Kontext leistungsgewandelter und leistungsgeminderter Arbeiter. Der Einsatz von Gesten in diesen Systemen und die Auswirkungen dieser auf die Leistungsfähigkeit von Arbeitern sind nicht erforscht.

3. Assistenzsystem

Das motionEAP¹ Projekt hat das Ziel ein System zur Effizienzsteigerung und zur Assistenz bei Produktionsprozessen in Unternehmen auf Basis von Bewegungserkennung und Projektion zu entwickeln. Durch eine anwenderorientierte Konzeption, prototypische Umsetzung und Evaluation soll ein neuartiges prozessorientiertes Assistenzsystem für Produktionsprozesse entstehen. Das motionEAP Projekt ist aus dem ASLM Forschungsprojekt (siehe Kapitel 2) der Hochschule Esslingen hervorgegangen. Das motionEAP Assistenzsystem soll Arbeiter aktiv und prozessorientiert unterstützen. Das System richtet sich vor allem an leistungsgeminderte und leistungsgewandelte Arbeiter.

Der Arbeiter wird am Assistenzsystem unterstützt, indem jeder Arbeitsschritt am Arbeitstisch erkannt und vom System mit einer logischen Aktion verbunden wird. Das System ist immer über den aktuellen Vorgang informiert und kann ein in-situ Feedback bezüglich des Arbeitsschrittes direkt in die Arbeitsfläche einblenden.

Das motionEAP Assistenzsystem realisiert das Konzept der Programmierung durch Demonstration. Arbeitsabläufe werden in einem Lernmodus dem System angelernt. Vorarbeiter oder erfahrene Arbeiter entwickeln direkt am Arbeitsplatz Schritt für Schritt einen Arbeitsablauf, welcher vom System erfasst und gespeichert wird. In einem Ausführungsmodus können diese Arbeitsabläufe unter kontextsensitiver Hilfe von unerfahrenen und zu unterstützenden Arbeitern abgerufen und durchgeführt werden. Aufgrund der Leistung des Arbeiters wird vom System ein kontinuierliches und personalisiertes Feedback in visueller oder akustischer Form gegeben. Das System soll die Überbelastung des Benutzers erkennen können und seine Leistungsfähigkeit und kognitiven Fähigkeiten richtig einordnen. Mit Hilfe von Gamification-Elementen wird dies ermöglicht (vgl. [FKB⁺13]). Metriken, welche die Zeit oder die Fehler pro Arbeitsschritt messen, können zur Motivation und Qualitätskontrolle eingesetzt werden.

Als Weiterentwicklung des ASLM Assistenzsystems soll das motionEAP Assistenzsystem folgende Anforderungen und Verbesserungen erfüllen (vgl. [FKB⁺13]):

- Steigerung der Prozessorientierung
- Vereinfachung der Bedienoberflächen und die Unterstützung natürlicher Interaktion
- Integration von Mechanismen zur Steigerung der Arbeitszufriedenheit und Motivation
- Integration von Mechanismen zur Fehlervermeidung und -früherkennung
- Anpassung an die Bedürfnisse von leistungsgeminderten Mitarbeitern

¹siehe <http://www.motioneap.de>

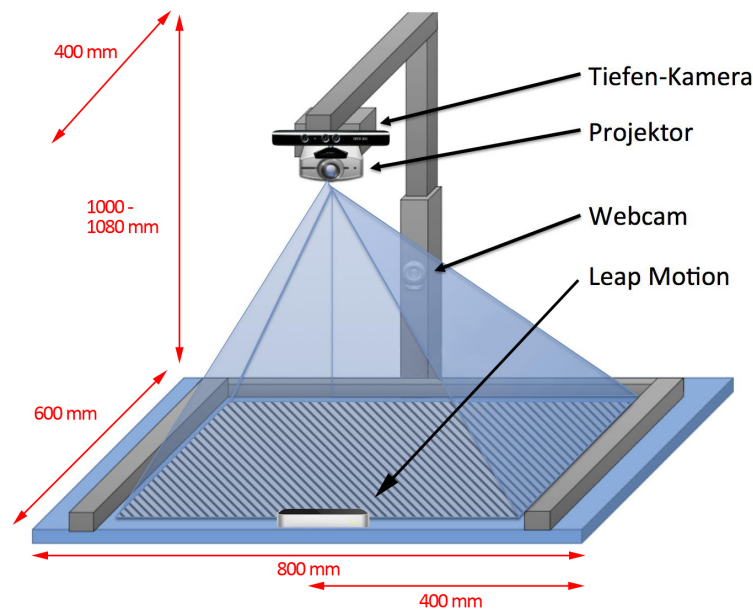


Abbildung 3.1.: motionEAP Montagetisch und Aufbau.

3.1. Beschreibung

Das untersuchte motionEAP Assistenzsystem befindet sich im Stadium eines Prototyps. Der technische Aufbau des Systems besteht aus einem Aluminiumrahmen, welcher über einem Montagetisch positioniert wird (siehe Abbildung 3.1). Dieser Rahmen hält in einem Aluminiumkäfig in einer variablen Höhe von 100 bis 180 Zentimetern eine Projektionseinheit, die einem Acer Projektor K330 entspricht, und eine Microsoft Kinect Tiefenkamera. Zusätzlich wird eine Leap Motion in einem vertikalen Winkel von 35 Grad an der vorderen Kante des Montagetisches platziert, wobei das Gerät in einer späteren Version im Tisch versenkt wird. Eine Webkamera und andere Sensoren können in zusätzlicher Form zur Überwachung und Objekterkennung eingesetzt werden. Der Projektor projiziert vertikal ein Bild auf den Montagetisch, wobei die gesamte Arbeitsfläche sowie die Werkstücke auf dem Tisch mit einer Projektion überlagert werden können. Pfeile, Umrandungen und andere grafischen Elemente können direkt auf ein Werkstück projiziert werden und wichtige Hinweise zur Ausführung eines Arbeitsschrittes geben. Der Projektor als auch die Tiefenkamera sind getrennt voneinander verstell- und fixierbar, dies erlaubt eine exakte und permanente Kalibrierung. Die Kalibrierung des Projektors und der Tiefenkamera wird durch die Erkennung von projizierten Kreisen im Bild der Kamera ermöglicht. Die Tiefeninformationen der Kinect werden zur Positionsbestimmung der Hand oberhalb der Arbeitsfläche genutzt. Durch diese Daten können auch Berührungen des Montagetisches festgestellt werden. Durch das open-source

Framework Candescent NUI² können Hände, einzelne Finger eines Benutzers sowie die Abstände dieser Finger zur Kinect erkannt werden. Mit Hilfe dieser Abstände werden durch das Ubidisplays Toolkit³ Berührungen der Arbeitsfläche durch einzelne Finger erkannt (vgl. [HA12]). Durchstößt ein Finger einen 15 Millimeter hohen Raum über der Arbeitsfläche, so wird ein „Touchevent“ ausgelöst. Somit kann jede Arbeitsfläche unter dem motionEAP Gerüst zu einem Multitouch-Tisch werden. Neben der Erkennung zweidimensionaler Berührungen des Montagetisches können mit Hilfe der Leap Motion dreidimensionale Handbewegungen und Gesten erfasst werden.

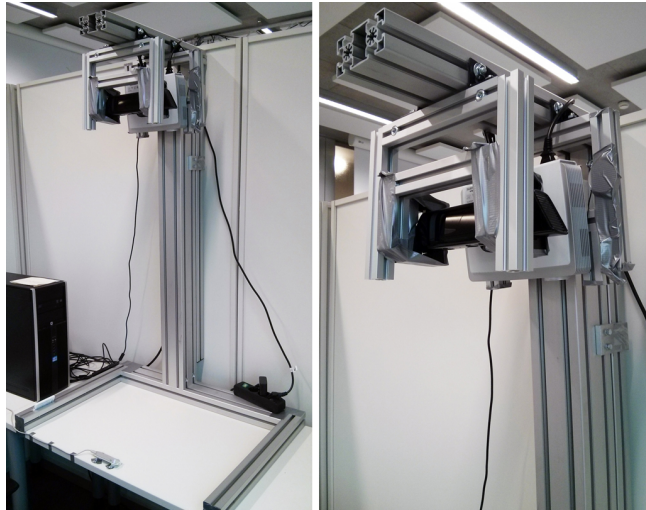


Abbildung 3.2.: motionEAP Prototyp Assistenzsystem.

3.2. Anwendungsfälle

Das motionEAP Assistenzsystem kann in den folgenden hier beschriebenen Szenarien eingesetzt werden (vgl. [FKB⁺13]). Besonders die Szenarien des Trainings und der Fortbildung, als auch des Montagetisches sind hinsichtlich dieser Arbeit von Bedeutung. Anhand dieser Szenarien werden Aktionsmöglichkeiten beteiligter Arbeiter und Trainer identifiziert und Gesten für diese Aktion entwickelt und bewertet.

Szenario des Trainings und der Fortbildung

In einer Trainingsumgebung eingesetzt, kann das System zur Schulung von Arbeitern und zur Demonstration neuer Arbeitsschritte eingesetzt werden. Hierbei können neue Arbeitsab-

²siehe <http://candescentnui.codeplex.com>

³siehe <https://code.google.com/p/ubidisplays>

läufe konzipiert und programmiert werden. Das System wird außerhalb seines industriellen Einsatzortes eingesetzt, simuliert jedoch zukünftige Arbeitsabläufe durch verschiedene Trainingsszenarien. Auftretende Fehler in Arbeitsabläufen werden identifiziert und korrigiert. Hinweise zu diesen Korrekturen werden in den Arbeitsabläufen hinterlegt. Dieses Vorgehen erlaubt prozesssichere Arbeitsabläufe für zukünftige Serienanwendungen. Weiter sollen Hinweise zur Verbesserung von Bewegungsabläufen und zu ergonomischen Aspekten in den Arbeitsabläufen hinterlegt werden können. Die Trainingsumgebung ermöglicht eine stressfreie Einlernphase an System und Montagetätigkeit ohne Zeit- und Leistungsdruck. Die Abbildung des Anhangs A.2 zeigt den Anwendungsfall (engl. *use case*) des Trainers und mögliche Aktionen dieser Benutzergruppe.

Szenario des Montagetisches

Der in der Industrie verwendete Montagetisch stellt den primären Einsatzort des Assistenzsystems dar. In einem Produktionsprozess wird eine Montageanleitung in den Arbeitsbereich projiziert, Hinweise zu falsch ausgeführten Arbeitsschritten gegeben und Korrekturen vorgeschlagen. Das System muss flexibel auf den konstruktiven Einsatzort, als auch auf eine hohe Variantenanzahl in der Produktion konzipiert sein. Die Abbildung des Anhangs A.1 zeigt den Anwendungsfall (engl. *use case*) des Arbeiters und mögliche Aktionen dieser Benutzergruppe.

Szenario der Kommissionierung

In diesem Szenario wird das Assistenzsystem zur Anleitung und Kontrolle bei der Vorkommissionierung verwendet. Hierbei sollen die verbreiteten Systeme des „Pick-by-Light“ und „Put-to-Light“ durch das Assistenzsystem ersetzt werden. Anstatt den Entnahmeort sowie den Ablageort eines Werkteiles oder Produktes durch Lichtsignale zu kennzeichnen, werden diese mit Hilfe von Projektionen angezeigt. Die Kombination aus Bewegungserkennung und Projektion gibt nicht nur eine Anleitung zur Kommissionierung, sondern stellt auch eine Kontrolle der Entnahme und Platzierung von Teilen sicher.

Szenario der Montagezelle

Mehrere Montagetische können zu einer Montagelinie zusammengenommen werden. Die nahtlose Kombination von mehreren Assistenzsystemen bildet hierbei eine Herausforderung. Tiefenkameras und Sensoren wie die Leap Motion müssen ohne Probleme zusammenspielen. In Montagezellen werden komplexere und umfangreichere Montageabläufe abgebildet, dies hat eine hohe Produktionsflexibilität zur Folge. Der Einlern- und Trainingsaufwand, als auch die Fehlerquote von Arbeitern steigt. Die anleitende und überprüfende Funktion des Assistenzsystems soll den Einstiegsaufwand und die Fehlerrate neuer Mitarbeiter stark reduzieren.

3.3. Aktionsmöglichkeiten

Aus den Anwendungsfällen des Trainings, der Fortbildung sowie des Montagetechnischen lassen sich eine Reihe von Aktionen an dem motionEAP Assistenzsystem herleiten. Diese Aktionen können von den Personen, die am Assistenzsystem beteiligt sind, ausgeführt werden. Die Benutzer sind in erster Linie der Vorarbeiter, Meister oder Trainer und der Arbeiter.

Der Trainer, welcher Arbeitsabläufe am System entwickelt und plant, ist meistens eine uneingeschränkte Person. Der Arbeiter, welcher Montagetätigkeiten am System ausführt und durch Arbeitsabläufe in mehreren Arbeitsschritten geführt wird, ist meistens eine leistungsgeminderte oder leistungsgewandelte Person. Beide Benutzergruppen verwenden das System und müssen verschiedene Aufgaben lösen.

Zur Analyse verschiedener Aktionen wurde ein Montage-Szenario für das Assistenzsystem entwickelt, dieses enthält die manuelle Montage einer einfachen Schraubzwinde der Firma Bessey⁴ bestehend aus fünf Werkzeugteilen (siehe Abbildung 3.3). Im Folgenden wird dieses Szenario für den Trainer und den Arbeiter beschrieben und aus diesem die Aktionen im System hergeleitet.

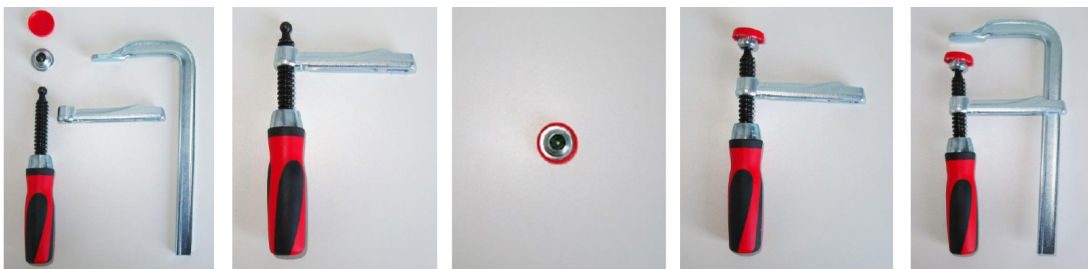


Abbildung 3.3.: Montage-Szenario der fünfteiligen Schraubzwinde.

3.3.1. Trainer

Der Trainer kann Arbeitsabläufe erstellen und speichern, Abbildung 3.4 zeigt die Erstellung schematisch. Arbeitsabläufe bestehen aus Arbeitsschritten, diese bestehen aus Masken, Hinweistexten, Fotos von Werkstücken, einfachen Grafiken wie Pfeilen, Kreisen, Rechtecken und Umrandungen.

Die Erstellung eines Beispiel Arbeitsablaufs, anhand der oben erwähnten Schraubzwinde, würde im System wie folgt ablaufen. Der Trainer erfasst alle Teile der Schraubzwinde, erstellt Fotos von jedem Teil und beschriftet alle Teile. Es folgt eine Anordnung der Teile nach Montagereihenfolge und die Erstellung der Arbeitsschritte:

⁴siehe <http://www.bessey.de>

3. Assistenzsystem

- **Schritt 1 anlegen:** Anfangszustand und Endzustand als Foto einfügen. Position der Teile festlegen. Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: Griff in Backe eindrehen (GB1).
- **Schritt 2 anlegen:** Anfangszustand und Endzustand als Foto einfügen. Position der Teile festlegen. Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: Schutzkappe auf Klemmauflage drücken (SK).
- **Schritt 3 anlegen:** Anfangszustand und Endzustand als Foto einfügen. Position der Teile festlegen. Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: SK auf GB1 drücken (mit Gummihammer) (SKGB1).
- **Schritt 4 anlegen:** Anfangszustand und Endzustand als Foto einfügen. Position der Teile festlegen. Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: Backe2 in SKGB1 einführen.
- **Schritt 5 anlegen:** Endzustand als Foto einfügen. Als Endsituation definieren und Erfolgsnachricht festlegen.
- **Schritt 6:** Montagereihenfolge Schraubzwinde abspeichern.

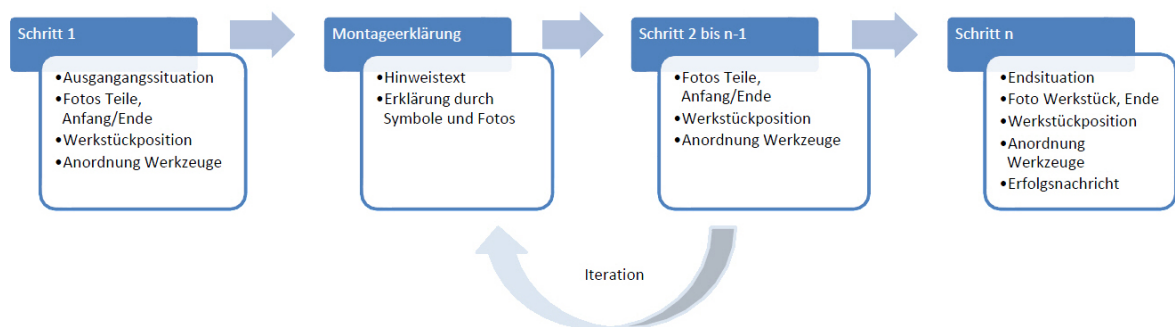


Abbildung 3.4.: Die Erstellung eines allgemeinen Arbeitsablaufs durch den Trainer.

Aktionen des Trainers

Aus dem obigen Szenario können folgende allgemeinen Aktionen des Trainers ermittelt werden, siehe Tabelle 3.1. Jede Aktion wird kurz beschrieben und ist eindeutig mit einer ID gekennzeichnet, dies erlaubt eine spätere Zuordnung zu bestimmten Gesten.

Rot hervorgehobene Aktionen können vom Trainer, als auch vom Arbeiter ausgeführt werden. Dies betrifft vor allem überlappende Tätigkeiten wie „in Hinweistexten scrollen“, die Schrift vergrößern, Objekte auswählen und „Videos von vorne spielen lassen“. In erster Linie werden diese Tätigkeiten jedoch dem Trainer bei der Erstellung von Arbeitsabläufen unterstellt.

	Aktion	Beschreibung	ID
01	Position von Werkstück/ Werkstückteilen/Werkzeugen festlegen, markieren	Festhalten der Position eines Teiles, Ablage- ort bestimmen.	01
02	Foto von Werkstück/ Werkstückteilen machen	Ein Foto eines Teiles erstellen.	02
03	Maske/Hinweistext Position ändern	Die Position eingeblendeter Masken ändern.	03
04	Maske/Hinweistext Größe ändern	Die Größe eingeblendeter Masken ändern.	04
05	Maske/Hinweistext an bestimmter Position neu einfügen/ aufmachen	Eine Maske an einer neuen Position aufma- chen.	06
06	In Masken/Hinweistexten scrollen	Im Text einer Maske scrollen.	08
07	In Masken/Hinweistexten Schrift vergrößern	Den Text in einer Maske vergrößern.	09
08	Objekt auswählen	Ein Objekt, wie einfache Grafiken, Videos, Fotos etc. auswählen.	12
09	Objekt rotieren	Ein Objekt, wie einfache Grafiken, Videos, Fotos etc. rotieren.	14
10	In Video/Animation Kamera Position ändern	Die Kamerasicht auf eine Szene in einer Ani- mation ändern.	15
11	11. In Video/Animation Kamera Zoom ändern	Den Kamera Zoom in einer Animation än- dern.	16
12	Video von vorne spielen lassen	Die Wiedergabe eines Videos von vorne be- ginnen.	19

Tabelle 3.1.: Identifikation der Aktionen des Trainers

3.3.2. Arbeiter

Der Arbeiter führt eine manuelle Montage aus und wird aktiv durch einen Arbeitsablauf geleitet (siehe Abbildung 3.5). Jeder Arbeitsschritt enthält Montageerklärungen in Form von Hinweisen, Fotos, Videos, Animationen und einfachen Grafiken wie Pfeilen, Kreisen, Rechtecken und Umrandungen.

Die Durchführung eines Arbeitsablaufs, anhand der anfangs erwähnten Schraubzwinde, würde im System wie folgt aussehen. Alle Werkteile der Schraubzwinde liegen griffbereit neben oder auf dem Werk Tisch. Vor dem Einsatz und der Entnahme für den jeweiligen Arbeitsschritt erfolgt eine Kennzeichnung der Teile. Weiter läuft die Montage in fünf Schritten ab:

3. Assistenzsystem

- **Schritt 1:** Montage nach Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: Griff (G) in Backe1 (B1) eindrehen. Ergebnis: Griff in Backe1 (GB1), Arbeitsschritt bestätigen.
- **Schritt 2:** Montage nach Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: Schutzkappe (S) auf Klemmauflage (K) drücken. Ergebnis: Schutzkappe auf Klemmauflage (SK), Arbeitsschritt bestätigen.
- **Schritt 3:** Montage nach Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: SK auf GB1 drücken (mit Gummihammer). Ergebnis: SK auf GB1 (SKGB1), Arbeitsschritt bestätigen.
- **Schritt 4:** Montage nach Montageerklärung mit Hinweistext, Symbolen, Fotos, Videos. ToDo: B2 in SKGB1 einführen. Ergebnis: B2 in SKGB1 (B2SKGB1), Arbeitsschritt bestätigen.
- **Schritt 5:** Ende der Montage bestätigen.

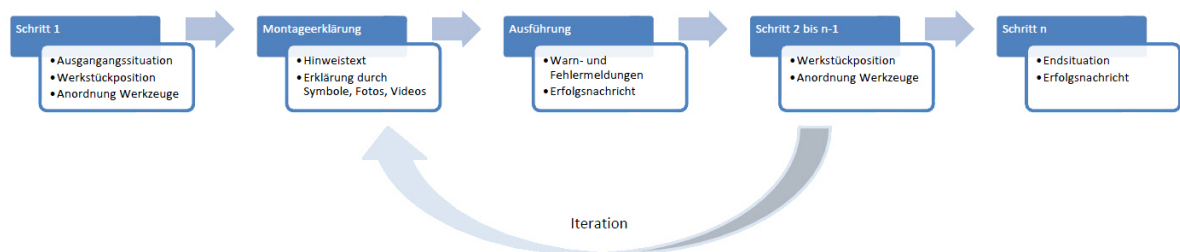


Abbildung 3.5.: Die Ausführung von allgemeinen Arbeitsschritten durch den Arbeiter.

Aktionen des Arbeiters

Aus dem obigen Szenario können folgende allgemeinen Aktionen des Arbeiters ermittelt werden, siehe Tabelle 3.2.

Aufgrund gleicher Funktionalität können verschiedene Aktionen wie das Wechseln zwischen Masken und das Wechseln zwischen Objekten gruppiert werden. Gleiches gilt für das Löschen und Schließen von Objekten, Dialogen und Masken. Ein Dialog enthält zum Beispiel eine Meldung vom System und kann bestätigt oder abgelehnt werden. Rot hervorgehobene Aktionen können vom Arbeiter, als auch vom Trainer ausgeführt werden. Dies betrifft in erster Linie übergeordnete Aktionen wie einen „Dialog bestätigen“, nicht bestätigen und die spezielle Aktion „zwischen Masken wechseln“.

	Aktion	Beschreibung	ID
01	Ausführung des Arbeitsschrittes bestätigen	Den erfolgreichen Abschluss eines Arbeitsschrittes bestätigen.	22

02	Ausführung des Arbeitsschrittes nicht bestätigen	Den Abschluss eines Arbeitsschrittes nicht bestätigen.	23
03	Arbeitsschritt wiederholen	Einen Arbeitsschritt wiederholen.	24
04	Zwischen Maske/Hinweistexten, Objekten/Arbeitsschritten wechseln/navigieren	Zwischen Objekten, Masken und Arbeitsschritten wechseln.	26
05	Dialog bestätigen, Video abspielen	Einen Dialog bestätigen oder ein Video abspielen lassen.	27
06	Maske/Hinweistext schließen, Objekt löschen/schließen, Dialog nicht bestätigen, Video anhalten	Objekte löschen, Masken schließen, Videos anhalten und Dialoge schließen.	28

Tabelle 3.2.: Identifikation der Aktionen des Arbeiters

4. Vorstudie

In diesem Kapitel werden zu den im Vorfeld bestimmten Benutzeraktionen Gesten identifiziert. Hierbei handelt es sich um zweidimensionale (2D), dreidimensionale (3D) und hybrid Gesten. Hybrid Gesten setzen sich aus 2D und 3D Gesten zusammen. Die Entwicklung dieser Gesten bezieht eine durchgeführte Studie, als auch verschiedene Erkenntnisse zum Benutzerverhalten leistungsgeminderter und leistungsgewandelter Arbeiter, nicht nur an Assistenzsystemen, ein.

Im Rahmen der Studie werden Gesten von Teilnehmern gesammelt und im Hinblick auf die Verwendung im Assistenzsystem bewertet. Die Auswahl relevanter Gesten erfolgt aufgrund dieser Ergebnisse, wobei ebenfalls bestimmte anhand der Zielgruppe festgelegte Kriterien berücksichtigt werden. Somit verfolgt diese Arbeit das Ziel der Herleitung von Gesten mit Hilfe von Studien, jedoch auch aufgrund der Untersuchung bestehender Eingabemethoden und deren Rahmenbedingungen. Ein ähnliches Vorgehen zur Entwicklung von Natural User Interfaces schlagen auch Wigdor et al. in [WW10] (S. 207 und 219) vor. Über Iterationen im Designprozess von Benutzerschnittstellen und der Zuhilfenahme von Ergebnissen aus Benutzerstudien können intuitive und natürliche Schnittstellen zwischen Mensch und Maschine entstehen.

Im Folgenden wird der Ablauf der Studie beschrieben. Hierzu zählen das Durchführungskonzept, die Beschreibung und Auswahl der Teilnehmer und die verwendeten technischen Hilfsmittel zur Ausführung. Abschließend erfolgt die statistische Analyse der Studienergebnisse und die Präsentation relevanter Gesten.

4.1. Konzeption der Durchführung

Die Methodik der Studie ist in Anlehnung an die Methodologie von Wobbrock in den Arbeiten [WRM05], [WMW09] und Grandhi in [GJM11b] konzipiert.

Ziel der Studie ist die Erfassung qualitativer Daten zu möglichen Gesten für festgelegte Aktionen. Zwölf Teilnehmer werden eingeladen. Jeder Teilnehmer benennt zu insgesamt 25 zuvor festgelegten Aktionen jeweils eine 2D, 3D und hybrid Geste und muss diese vorführen. Zusätzlich erläutert der Teilnehmer während der Vorführung, nach dem *Thinking Aloud* Prinzip, laut seine Gedankengänge und begründet seine getroffene Auswahl. Jeder Teilnehmer wird zu Analysezwecken auf Video aufgenommen und muss am Ende des Interviews einen Fragebogen zur Person und getroffener Gestenauswahl beantworten. Während des

Interviews werden Notizen zur Ausführung und den Gedankengängen des Teilnehmers gemacht, diese erlauben das Festhalten von Besonderheiten im Auswahlprozess. Die technische Realisierbarkeit wird vom Teilnehmer bei der Auswahl von Gesten nicht beachtet.

Die Auswertung der gesammelten Daten verläuft wie folgt, alle verschiedenen Gesten einer Gestenart werden identifiziert, drei Listen aller nicht redundanten Gesten werden erstellt. Zu jeder Geste einer Liste werden die Aktionen zugeordnet, für die die Geste gezeigt wurde. Es entsteht somit eine Konfliktgruppe von Aktionen für die selbe Geste. Durch mehrmaliges zuordnen der gleichen Aktion zu einer Aktionengruppe erhöht sich ein Zähler für diese Aktion. Dieser Zähler wird *score* genannt. Die Aktion in der Aktionengruppe mit dem höchsten *score* gewinnt die Geste. Eine Geste wird auch als *symbol* bezeichnet. Die Gleichung 4.1 berechnet den *score* einer Aktion. Dieser ist die Anzahl gleicher vorgeschlagener Gesten für eine Aktion.

$$(4.1) \text{ score} = |\text{symbols}|$$

Durch dieses Vorgehen soll ein Gesten-Set mit einer hohen Übereinstimmung zu der Auswahl aller Teilnehmer entstehen. Diese Übereinstimmung eines Gesten-Sets wird als *guessability* bezeichnet. Die Gleichung 4.2 berechnet die *guessability*, wobei S die Anzahl aller Aktionen die gewonnen haben darstellt. $|P_s|$ stellt die Anzahl gleicher vorgeschlagener Gesten für Aktion s , also den *score* der Aktion s dar. $|P|$ sind alle vorgeschlagenen Gesten für alle Aktionen. Zusammenfassend sagt die *guessability* aus, zu wie viel Prozent sich ein ermitteltes Gesten-Set mit den vorgeschlagenen Gesten der Teilnehmer deckt.

$$(4.2) \text{ G} = \frac{\sum_{s \in S} |P_s|}{|P|} \cdot 100\%$$

Weitere qualitative Daten aus dem Fragebogen zum Interview, als auch aus den Notizen zur Ausführung und den Gedankengängen der Teilnehmer fließen in die spätere Diskussion zur Studie mit ein.

4.2. Teilnehmer

Die zwölf Teilnehmer der Studie setzen sich aus zehn männlichen und zwei weiblichen Personen zusammen. Das Durchschnittsalter liegt bei 27 Jahren ($M = 26.67, SD = 5.18$), wobei 75% der Teilnehmer Rechtshänder und 25% Linkshänder sind. Bei den Teilnehmern handelt es sich um Studenten der Universität Stuttgart, vorwiegend aus dem Bereich der Informatik. Die Teilnahme ist freiwillig. Alle Teilnehmer haben Erfahrungen mit 2D Gesten, vor allem aus der alltäglichen Nutzung von Smartphone und Tablet. Nur 50% der Teilnehmer haben Erfahrungen mit 3D Gesten. Es werden vor allem Erfahrungen in dem Unterhaltungsbereich genannt. Microsofts Kinect und Nintendos Wii werden hier zu Spielzwecken genutzt. Ein Teilnehmer

gibt die Bedienung eines Fernsehers an. Zur Ermittlung qualitativer Daten werden für diese Studie keine leistungsgeminderten und leistungsgewandelten Personen vorausgesetzt.

4.3. Apparatur

Die technischen Hilfsmittel, die in dieser Studie benutzt werden, können in drei Bereiche eingeteilt werden. Technisches Equipment, Fragebogen und Gegenstände zur Simulation des zu untersuchenden Assistenzsystems. Der Fragebogen enthält neben Fragen nach Alter, Geschlecht, Händigkeit, auch Fragen zu Erfahrungen mit 2D und 3D Gesten. Zum Abschluss wird noch gefragt, zu welchen Aktionen es am schwierigsten ist Gesten zu finden und für welche Gestenart dies der Fall ist. Der Fragebogen wird mit Hilfe von Google Forms¹ erstellt und per Webmaske den Teilnehmern zur Verfügung gestellt. Ebenfalls erlaubt der Google Dienst das Herunterladen der Antworten in Form einer CSV-Datei. Die Gegenstände zur Simulation bestehen aus einem Arbeitstisch mit abgeklebter Arbeitsfläche, einer einfachen Schraubzwinge und verschiedenen aus Pappe hergestellten Beispielmasken zur Veranschaulichung einer grafischen Oberfläche. Zum technischen Equipment gehören zwei Notebooks zum Erfassen des Fragebogens, als auch von Notizen. Weiter wird der Teilnehmer mit zwei Weitwinkelkameras aufgenommen.

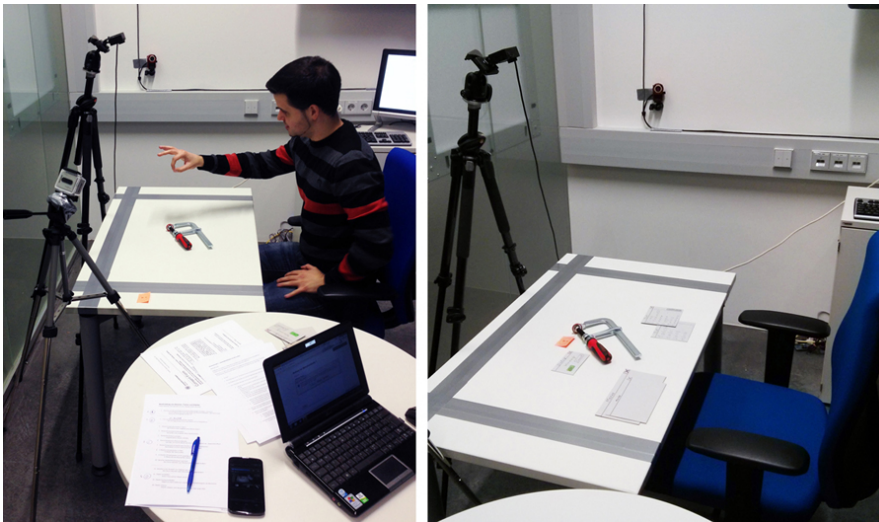


Abbildung 4.1.: Technische Hilfsmittel der Benutzerstudie zur Identifikation von Gesten.

¹siehe <http://www.google.com/drive/apps.html>

4.4. Durchführung

An zwei Tagen nahmen zwölf Teilnehmer an der Studie teil. Jedes Interview dauerte eine halbe Stunde. Jedem Teilnehmer wurde der Hintergrund der Studie erklärt und das Ziel kurz zusammengefasst. Eine Zustimmungserklärung zur Verwendung von Daten wurde vorgelegt und vom Teilnehmer unterzeichnet. Alle Gestenarten wurden erklärt sowie die möglichen Unterschiede dieser gezeigt. Die Hilfsmittel zur Simulation einer manuellen Montage wurden ebenfalls erklärt und gezeigt. Daraufhin wurde jede Aktion beschrieben und der Effekt, welcher von dieser Aktion ausgelöst wird, gezeigt. Der Teilnehmer wiederholte den Effekt und zeigte die von ihm gewählte 2D, 3D und Hybrid Geste. Gesten konnten mit der rechten, als auch mit der Linken Hand ausgeführt werden. Während der Vorführung dieser Gesten erläuterte der Teilnehmer seine Gedanken zum Auffinden einer Geste. Der Oberkörper und die Hände des Teilnehmers wurden dabei auf Video aufgenommen. Weiter wurden zur jeder Aktion Notizen zur Ausführung der Gesten gemacht. Die Reihenfolge der abgefragten Aktionen wurde anhand eines bestimmten lateinischen Quadrates (vgl. [FH10], S. 84) gewählt, um Nebeneffekte in den Gedankengängen der Teilnehmer zu reduzieren. Für diese Zwecke wurde das „*Balanced Latin Square*“ genutzt, dieses wurde in sechs auf sechs Felder unterteilt. Gruppen von jeweils zwei Teilnehmern bekamen somit verschiedene Reihenfolgen von Aktionen. Am Ende eines jeden Interviews füllte jeder Teilnehmer einen Fragebogen zu seiner Person und zu Schwierigkeiten im Verlauf der Gestenauswahl aus.

4.5. Resultate und Diskussion

Insgesamt wurden 900 Gesten von zwölf Teilnehmern gesammelt und ausgewertet. Das entspricht 25 Gesten für 25 Aktionen pro Gestenart und pro Teilnehmer. Somit entstanden 300 Gesten pro Gestenart. Aus diesen 300 Gesten blieben nach Abzug redundanter und ähnlicher Gesten 40 2D Gesten, 77 3D Gesten und 31 Hybrid Gesten übrig.

Eine Auswertung der Ergebnisse zeigt Redundanzen in den vormals festgelegten Aktionen. Von ursprünglich sieben Aktionsbereichen und 25 Aktionen haben sich neun Aktionen als redundant erwiesen. Dies betrifft vor allem Aktionen, die im Kontext das selbe aussagen. Aktionen wie das Bestätigen eines Dialogs oder das Abspielen eines Videos können zusammengefasst werden, da sie Bestätigungen darstellen. Gleiches gilt für Aktionen wie dem Schließen einer Maske, dem Löschen eines Objektes, dem Anhalten eines Videos und dem „nicht bestätigen“ eines Dialogs. Diese Aktionen können zusammengefasst werden, da sie ein Löschen darstellen. Ebenfalls können Aktionen wie das Wechseln zwischen Masken, Objekten und Arbeitsschritten zusammengefasst werden. Eine ursprüngliche Trennung der Aktionen nach Masken, Objekten, interaktiven Inhalten wie Videos und übergeordneten Dialogen erwies sich als nicht genau genug. Von 25 Aktionen bleiben aufgrund von neuen Gruppierungen 18 Aktionen bestehen, diese Aktionen können in Kapitel 3 eingesehen werden.

Die Auswertung der Teilnehmergesten zeigt ein „Gesten-zu-Aktionen Verhältnis“ bei 2D Gesten von zwei zu eins, bei 3D Gesten von vier zu eins und bei Hybrid Gesten von beinahe

zwei zu eins. Nach Anwendung des im obigen Kapitel beschriebenen Verfahrens zum Feststellen von Gewinnergesten ergeben sich folgende Beobachtungen. Die alleinige Identifizierung finaler Gesten-Sets für 2D, 3D und Hybrid Gesten aufgrund von Benutzervorschlägen und „User-derived“ Gesten funktioniert für spezielle Systeme nicht. Die Verhältnisse der gezeigten Gesten zu den Aktionen und die Auswertung der Gewinnergesten zeigen, dass es für jede Aktion eine Vielzahl von unterschiedlichen Gesten gibt. Vor allem im dreidimensionalen und hybriden Bereich existieren wenige Übereinstimmungen zwischen den vorgeschlagenen Gesten der Teilnehmer. Für bestimmte Aktionen gibt es zu viele oder keine Gesten. Dies hat zur Folge, dass sich für Aktionen mehrere Gewinnergesten ergeben. Die Abbildung 4.2 zeigt die Ermittlung von Gewinnergesten, wobei die Zeilen sieben und zehn zwei Gewinnergesten für die gleiche Aktion darstellen (Aktion 3). In Spalte C können die scores der jeweiligen Aktionen in Klammern gesehen werden.

	A	B	C	D
1	Nr.	3D Gesten aus Benutzerstudie 1	Aktionen (wieviel mal genutzt)	Gewinner-Aktion
2				
3	1	Mit vier Fingern, Dauem Zeigefinger horizontal ein Rechteck zeigen	2(8)	2
4	2	Mit vier Fingern, Dauem Zeigefinger horizontal ein Rechteck zeigen, Finger bewegen sich synchron kurz zu/von einander	1(1), 2(3)	2
5	3	Enlarge/Shrink - Mit vier Fingern, Dauem Zeigefinger horizontal ein Rechteck zeigen, Finger bewegen sich synchron zu/von einander	4(1), 7(2), 9(1)	7
6	4	Hand horizontal statisch	7(1), 12(3), 25(1), 1(4), 6(1)	1 (auch 12)
7	5	Hand horizontal statisch mit Bewegung auf Pfad	3(5)	3
8	6	Hand horizontal statisch mit Bewegung auf Pfad in Aussenbereich	13(4), 25(2), 5(2), 21(1), 22(1)	13
9	7	Hand vertikal statisch mit Bewegung auf Pfad in Aussenbereich	5(1), 13(1)	5,13
10	8	Faust horizontal statisch mit Bewegung auf Pfad	3(3)	3
11	9	Faust horizontal statisch mit Bewegung auf Pfad in Aussenbereich	13(1)	13
12	10	Enlarge/Shrink - Daumen Zeigefinger von/zueinander bewegen	4(2), 6(1), 9(1), 16(2)	4,16
13	11	Enlarge/Shrink - Mit zwei Händen horizontal halten von/zueinander bewegen	1(1), 4(8), 6(2), 7(1), 9(2), 12(2), 16(2)	4
14	12	Enlarge/Shrink - Mit zwei Zeigefingern horizontal zeigen und von/zueinander bewegen	1(1), 4(6), 9(2), 12(1), 16(2)	4
15	13	Swipe Beide Hände horizontal von aussen nach innen bewegen	24(1)	24
16	14	Swipe mit Handfläche horizontal nach rechts	10(1), 11(2), 21(2), 22(1), 23(1), 25(1)	10,21

Abbildung 4.2.: Ein Ausschnitt aus der Ermittlung der 3D Gewinnergesten.

Die Auswahl der Gewinnergesten kann somit nicht vollständig aus vorgeschlagenen Gesten der Teilnehmer erfolgen. Diese Feststellung zeigt, dass es so etwas wie natürliche Gesten nicht gibt. Gesten sind so vielfältig, dass die Teilnehmer diese schwer Beurteilen und zu Aktionen zuordnen können. Dies gilt vor allem für 3D Gesten und besonders für Hybrid Gesten. Aufgrund dieser Feststellung müssen weitere Kriterien zur Identifikation der finalen Gesten-Sets hinzugenommen werden.

Nach Auswahl aller vorläufigen Gewinnergesten ergibt sich eine guessability von 50.49% für das 2D Gesten-Set und 25.49% für das 3D Gesten-Set. Eine guessability für das Hybrid Gesten-Set kann aufgrund lückenhafter Gestenvorschläge der Teilnehmer nicht angegeben werden. Diese Ergebnisse unterstreichen nochmals die Uneinigkeit der Teilnehmer in Bezug auf die Gestenwahl und die Vielfalt möglicher Gesten.

4.5.1. Identifikation relevanter Gesten

Elf von zwölf Studienteilnehmern sehen Hybrid Gesten als die schwierigsten Gesten bei der Gestenauswahl an. Die Kombination von 3D und 2D Geste wird oft als unnötig für die Ausführung einer Aktion erachtet, vor allem wenn eine 3D Geste eine Aktion ausreichend beschreibt. Sechs von zwölf Befragten können keine Beispiele für Hybrid Gesten zeigen, sie haben schlichtweg keine Vorstellung welche Kombination einer 2D und 3D Geste zu einer Aktion passen könnte. Jedoch werden Hybrid Gesten auch als Hilfe bei mehreren zur Auswahl stehenden Masken und Objekten gesehen. Fünf von zwölf Teilnehmern setzten als 2D Gestenanteil einer Hybrid Geste den einfachen 2D „Tap“ als Selektionshilfe bei Masken und Objekten ein. In Verbindung mit dieser Selektion werden Hybrid Gesten als intuitiv und in einigen Aktionen als sinnvoll und vereinfachend angesehen. Hybrid Gesten laufen sequentiell ab, da ein gleichzeitiges Zeigen einer 2D und 3D Geste motorisch schwierig ist.

2D Gesten werden aus bekannten Bereichen, wie der Nutzung von Smartphones und Tablets, übernommen. Auch Metaphern aus dem *Desktop* Bereich sind zu erkennen. Die Gestenausführung mit zwei Händen wird oft vernachlässigt. Ebenfalls werden Gesten bevorzugt mit einem Finger ausgeführt, wobei für Tap Gesten auch bis zu drei Finger verwendet werden. Gesten mit fünf Fingern oder der ganzen Hand sind selten. Eine Gestenunterscheidung anhand der Fingeranzahl erfolgt nicht. Griffpunkte von Masken und Objekten sind in dem meisten Fällen Eckpunkte. Der äußere Bereich der Arbeitsfläche wird als Papierkorb oder Ablage benutzt. Gesten, die eine längere Ausführung erfordern, werden als leichter empfunden. Die Geste eines Lassos ist eindeutig zur Aktion der Markierung einer Werkstückposition zugeordnet worden. Einfache Aktionen werden durch physische Gesten und komplexere Aktionen durch metaphorische Gesten abgebildet. Die Aktion des Auslösens eines Fotos wird durch vier Finger und eine Auslösebewegung abgebildet, dies erinnert an einen gehaltenen Fotoapparat. 17.50% der ausgeführten Gesten sind zweihändig, 82.50% aller Gesten werden mit einer Hand ausgeführt.

Obwohl der Mensch, wie in Kapitel 3 beschrieben, sehr viel gestikuliert, werden Gesten im Alltag kaum wahrgenommen. Alle Studienteilnehmer haben Schwierigkeiten 3D Gesten für bestimmte Aktionen festzulegen. Der Begriff einer Geste ist nicht scharf definiert. Metaphern zur Nutzung von 3D Gesten in Assistenzsystemen existieren nicht, auch können keine Metaphern aus bekannten Bereichen übernommen werden, da dreidimensionale Interaktionsformen rar sind. Dreidimensionale Gesten unterliegen mehr Freiheitsgraden, dies hat eine größere Anzahl an Gestenmöglichkeiten zur Folge. Somit werden verschiedene Handbewegungen gezeigt. 7.79% der Studienteilnehmer zeigen statische Gesten und 92.21% dynamische Gesten. Dynamische Gesten bilden die Aktion mit einem Gegenstand oder die Nutzung dieses ab. Die Selektion von Masken und Objekten gestaltet sich als schwierig, die Teilnehmer benutzen den Finger als Mausersatz. Auch wird ein Greifen und Loslassen zur Selektion gezeigt. Die Aktionen des Bestätigens oder des nicht-Bestätigens werden mit symbolischen Gesten, wie dem Daumen hoch, Daumen runter und der Stop Handbewegung, gezeigt. 15.58% der ausgeführten Gesten sind zweihändig, 84.42% aller Gesten werden mit einer Hand ausgeführt.

Bei der Gestenausführung erfolgt bei allen Gestenarten keine Unterscheidung nach Kontext, für ähnliche Aufgaben werden gleiche oder ähnliche Gesten gezeigt.

Kriterien für leistungsgeminderte und leistungsgewandelte Arbeiter

Die Erforschung von Gesten in der Mensch-Computer-Interaktion bezieht sich in viele Arbeiten vorwiegend auf uneingeschränkte Personen. Die Untersuchung der Interaktion leistungsgeminderter und leistungsgewandelter Menschen mit einem System durch Gesten zeigt folgende hier beschriebene Probleme. Diese Probleme treten gleichermaßen bei zwei- und dreidimensionalen Gesten auf, falls nicht eine Gestenart explizit erwähnt wird. Bei leistungsgewandelten, als auch leistungsgeminderten Menschen treten mitunter die selben motorischen und kognitiven Beschränkungen auf. Dabei stellt der Schweregrad der Beschränkung den signifikanten Unterschied zwischen den beiden Gruppen dar.

Leistungsgeminderte und leistungsgewandelte Menschen zeigen eine reduzierte Geschwindigkeit in der Ausführung von Gesten. Ebenfalls sind Verzögerungen in der Ausführung komplexer Gesten zu beobachten. Dies gilt auch für die Ausführung einhändiger Gesten. Besonders bei fein- und grobmotorischen Behinderungen ist eine verlangsamte Ausführung von zweidimensionalen Gesten zu beobachten.

Weiter ist eine reduzierte Genauigkeit in der Ausführung von Gesten zu erkennen. Gesten, die bei der Ausführung mehrere Finger und besonders den Daumen involvieren, sind ungenau. Generell werden bei der Ausführung von Gesten weniger Finger eingesetzt. Komplexe Gesten werden als schwierig empfunden. Dies hat den Hintergrund, dass leistungsgeminderte Menschen eingeschränkt in Sprache und Kognition sein können. Da es zwischen Sprache und Gesten eine Korrelation gibt, hat dies Auswirkungen auf die Gestik (vgl. [KMR08]). Es folgt der Schluss, dass Gesten so einfach wie möglich sein sollten.

Bei leistungsgewandelten, als auch leistungsgeminderten Menschen ist eine reduzierte Koordination festzustellen. Das ruhige Halten von Händen bereitet Schwierigkeiten. Ein Beispiel wären zitternde Hände (vgl. [Aly13]). Leistungsgeminderte Menschen mit Dyspraxie haben Schwierigkeiten mit grob- und feinmotorischen Bewegungen. Unwillkürliche, plötzliche, rasche und unregelmäßige Bewegungen können auftreten. Hände werden immer wieder in ungewöhnliche Winkel gebracht. Die Folge sind weniger präzise zwei- und dreidimensionale Gesten. Grobmotorische Behinderungen führen zu längeren Haltezeiten bei Touch-Events und einer größeren Krafteinwirkung auf die berührungssensitive Oberfläche (vgl. [AKF13]). Einschränkungen in der Feinmotorik und nicht geschlossene Hände können bei Berührungsgesten Fehleingaben auslösen. Biswas et al. zeigen (vgl. [BL12]), dass die fehlende Koordination in einer Bewegung innerhalb der Anfangs und Endphase einer Geste zu beobachten ist.

Studien zeigen, dass leistungsgeminderte und leistungsgewandelte Menschen metaphorische Gesten für komplexere Aktionen des Alltages besser verstehen (vgl. [GJM11b]). Gesten müssen somit zu festgelegten Aktionen passen, dann tritt ein verbesserter Lerneffekt ein. Die Merkbarkeit von Gesten und Aktionen steigt. Zu Bemerken ist, dass geistige Behinderungen leistungsgeminderter Menschen sich negativ auf motorischen Fähigkeiten auswirken.

Wohingegen leistungsgeminderte Menschen mit psychische Störungen bessere motorische und kognitive Fähigkeiten zeigen (vgl. [Aly13]). Starke grob- und feinmotorische Einschränkungen, wie starkes Zittern, können die Nutzung von zwei- und dreidimensionalen Gesten erschweren oder verhindern. Biswas et al. zeigen jedoch (vgl. [BL12]), dass zweidimensionale Gesten mit Hilfe eines Stiftwerkzeuges auch unter starkem Zittern durchgeführt werden können.

Gesten Taxonomie

Aus den Ergebnissen der Benutzerstudie und den festgestellten Kriterien, welche für leistungsgeminderte und leistungsgewandelte Arbeiter im Hinblick auf die Ausführung von Gesten gelten, werden folgende Taxonomien für 2D, 3D und Hybrid Gesten vorgeschlagen.

Eine Taxonomie für 2D Gesten, im Bezug auf leistungsgeminderte und leistungsgewandelte Arbeiter am Assistenzsystem, sieht wie folgt aus:

Taxonomie zweidimensionaler Gesten		
Gesten Mapping		
Form	statische Pose	Handpose gehalten auf einer Position
	dynamische Pose	Handpose ändert sich auf einer Position
	statische Pose auf Pfad	Handpose gehalten bei Bewegung auf Pfad
	dynamische Pose auf Pfad	Handpose ändert sich bei Bewegung auf Pfad
	ein-Punkt Berührung	Statische Pose mit einem Finger
	ein-Punkt Pfad	Statische Pose auf Pfad mit einem Finger
Natur	symbolisch	Geste wird als Symbol erkannt
	physikalisch	Geste agiert physisch mit Objekt
	metaphorisch	Geste zeigt eine Metapher
	abstrakt	Zuordnung Geste zu Aktion ist abstrakt
Bindung	objekt-zentriert	Position-Def. bzgl. Objekteigenschaft
	umgebungs-abhängig	Position-Def. bzgl. Welteigenschaft
	umgebungs-unabhängig	Position-Def. kann Welteigenschaften ignorieren
	gemischte Abhängigkeit	umgebungs-unabhängig plus weitere
Fluß	diskret	Reaktion erfolgt nach Benutzeraktion
	kontinuierlich	Reaktion erfolgt während Benutzeraktion
Physikalische Eigenschaften		
Händig-keit	einhändig	Geste wird mit einer Hand ausgeführt
	beidhändig	Geste wird mit zwei Händen ausgeführt

Tabelle 4.1.: 2D-Gesten Taxonomie für Assistenzsysteme und leistungsgeminderte und leistungsgewandelte Arbeiter

Bis auf die physikalische Eigenschaft der Händigkeit einer Geste bleibt die Taxonomie gegenüber der von Wobbrock für 2D Gesten aus Kapitel 2 (vgl. [WMW09]) gleich. Die Händigkeit gibt die Anzahl der an der Geste beteiligten Hände an und stellt eine physikalische Eigenschaft der Geste dar.

Eine Taxonomie für 3D Gesten wird in Tabelle 4.2 angegeben. Ebenfalls gilt diese für Hybrid Gesten. Als Referenz gelten die in Kapitel 2 beschriebenen Taxonomien für dreidimensionale Gesten von Ruiz et al. (vgl. [RLL11]) und Aigner et al. (vgl. [AWB⁺12]).

Taxonomie dreidimensionaler und hybrider Gesten		
Gesten Mapping		
Form	statisch	Keine Bewegung oder Änderung in Geste
	dynamisch	Bewegung oder Änderung in Geste
	getippt, gestrichen	Geste aus einem Tip oder Strich
Natur	symbolisch	Geste wird als Symbol erkannt
	physikalisch	Geste agiert physisch mit Objekt
	pantomimisch	Geste imitiert eine reale Aktion
	deiktisch	Geste zeigt auf ein Objekt, Ort
Bindung	abstrakt	Zuordnung Geste zu Aktion ist abstrakt
	objekt-zentriert	Position-Def. bzgl. Objekteigenschaft
	umgebungs-abhängig	Position-Def. bzgl. Umgebung
	umgebungs-unabhängig	Position-Def. kann Umgebung ignorieren
Fluß	diskret	Reaktion erfolgt nach Benutzeraktion
	kontinuierlich	Reaktion erfolgt während Benutzeraktion
Physikalische Eigenschaften		
Händigkeit	einhändig	Geste wird mit einer Hand ausgeführt
	beidhändig	Geste wird mit zwei Händen ausgeführt
Komplexität	einfach	Geste aus einfachen Geste
	zusammengesetzt	Geste zusammengesetzt
Körperteil	Hand	Arm fixiert, Hand, Finger in Bewegung
	Arm	Arm in Bewegung
Dimension	1-Achse	Bewegung um eine Achse
	2-Achsen	Geste auf Oberfläche, 2D Ebene
	3-Achsen	Bewegung enthält Translation oder Rotation, nicht beide
	6-Achsen	Bewegung enthält Translation und Rotation
ROM	gering	Gelenkrotation weniger als 50% der Norm
	groß	Gelenkrotation mehr als 50% der Norm

Tabelle 4.2.: 3D- und Hybrid-Gesten Taxonomie für Assistenzsysteme und leistungsgeminderte und leistungsgewandelte Arbeiter

Die Dimensionen des Gesten Mappings entsprechen den Dimensionen der 2D Taxonomie von Wobbrock (vgl. [WMW09]). Getippte oder gestrichene Gesten in der Form Dimension sind einfache Tipp- und Streichgesten. Die Abgrenzung zu dynamischen Gesten ist die Anzahl von Tipps und Strichen (engl. *swipes*). Ein Doppeltipp im dreidimensionalen Raum ähnelt einem Doppelklick auf dem Desktop und wäre eine getippte Geste. Eine einfache Wischbewegung in den Außenbereich wäre eine gestrichene Geste. Eine physikalische Geste manipuliert ein Objekt bei Ausführung. Eine pantomimische Geste entspricht einer metaphorischen Geste, jedoch zeigt diese nicht nur Metaphern, sondern auch Aktionsabläufe. Dies kann eine Geste, die das Greifen und Wegwerfen von etwas zeigt, sein. Deiktische Gesten sind Zeigegesten, welche auf einen Ort oder ein Objekt zeigen. Objekt zentrierte Gesten benötigen Informationen zur Objektposition. Zum Beispiel stellt die Selektion eines Objektes, durch einen einfachen 3D Tap, eine Objekt zentrierte Geste dar.

Physikalische Eigenschaften von Gesten werden durch fünf Dimensionen beschrieben. Die Händigkeit gibt die Anzahl der an der Geste beteiligten Hände an. Die Komplexität erfasst Gesten, die aus mehreren Teilgesten bestehen. Die oben beschriebene „Wegwerfgeste“ wäre eine zusammengesetzte Geste. Eine „Arm“ Geste wäre eine Geste, in der die Bewegung des Armes mit einbezogen ist. Die Dimension einer Geste beschreibt die genutzten Achsen, die während der Ausführung einer Geste gebraucht werden. Eine „2-Achsen“ Geste entspricht einer 2D Geste. Wird eine Geste beidhändig ausgeführt und eine Hand zeigt eine Teilgeste auf der 2D Ebene, so beschreibt dies eine Hybrid Geste. Das Bewegungsausmaß, kurz ROM (engl. *range of motion*), beschreibt in der Physiotherapie die durchschnittliche Beweglichkeit der Gelenke eines Erwachsenen und somit den möglichen Winkel, welchen ein Gelenk maximal ausführen kann (vgl. [KEP96]). In der Physiotherapie kann das Bewegungsausmaß mit einem Goniometer ermittelt werden, mit dem bloßen Auge ist eine visuelle Approximation möglich. Die ROM Eigenschaft einer Geste zeigt den approximierten Winkel des Gelenks, welches an der Ausführung der Geste beteiligt ist. In erster Linie sind dies das Hand- und Schultergelenk, siehe Abbildung 4.3. Für das Schultergelenk interessieren die Bewegungen der Flexion und des horizontalen Heran- und Wegführens des Armes (horizontale Adduktion und Abduktion). Die Normwerte der Bewegung sind hier 180 Grad für die Flexion und 165 Grad für die horizontale Adduktion (45 Grad) und Abduktion (110 Grad). Für das Handgelenk sind das die Supination/Pronation mit ± 90 Grad, die Radialabduktion/Ulnarabduktion mit 30/-40 Grad und die Dorsalflexion/Palmarflexion mit ± 60 Grad. Die Supination/Pronation beschreibt eine Drehung der Handfläche auf den Handrücken, die Radialabduktion/Ulnarabduktion beschreibt das Schwenken der Hand über den Daumen, den kleinen Finger (speichenwärts, ellenwärts) und die Dorsalflexion/Palmarflexion beschreibt eine Handbewegung handrückenwärts und hohlhandwärts. Über diese Bewegungsausmaße kann somit die Beweglichkeit des Armes und der Hand eines Menschen beschrieben werden. Eine Geste, in der die Handfläche im Raum in einem großen Kreis bewegt wird, besitzt einen großen ROM Wert. Eine Geste, in der die Handfläche statisch gehalten wird, besitzt einen geringen ROM Wert.

Die Taxonomie für 3D und Hybrid Gesten (siehe Tabelle 4.2) beschreibt somit Gesten die im Umfeld eines Assistenzsystems und durch leistungsgeminderte und leistungsgewandelte Arbeiter durchgeführt werden können. Physische Merkmale von Gesten können erfasst und auf die Benutzer abgestimmt werden.

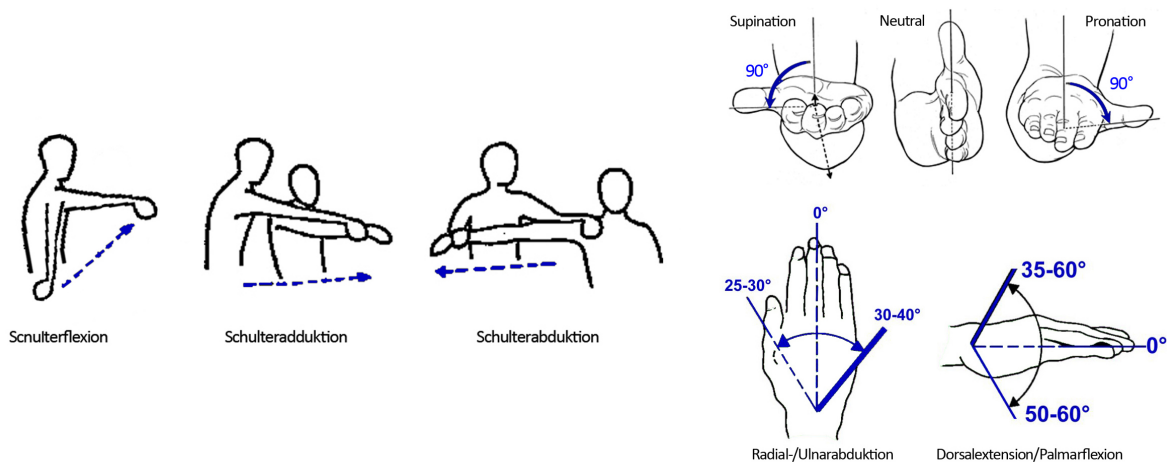


Abbildung 4.3.: ROM für Schulter (links) und Hand (rechts). (Bilder zum Teil aus [KEP96])

Gesten Präsentation

Folgende Schlussfolgerungen können aus den Ergebnissen der Benutzerstudie und den festgestellten Kriterien, welche für leistungsgeminderte und leistungsgewandelte Arbeiter im Hinblick auf die Ausführung von Gesten gelten, gezogen werden. Diese Schlussfolgerungen gelten gleichermaßen für zwei- und dreidimensionale Gesten, falls nicht eine Gestenart explizit hervorgehoben wird.

Es werden weniger Gesten mit zwei Händen ausgeführt. Eine Gestenunterscheidung anhand der Fingeranzahl wird vermieden. Eine Grenze für 2D Gesten scheinen vier Finger zu sein, 2D Tap Gesten können mit bis zu drei Fingern getätigt werden. Gesten werden zur bessern Lern- und Merkbarkeit in ähnlichen Kontexten wiederholt. Es wird versucht Gesten zu eindeutigen mentalen Modellen zuzuordnen, wie zum Beispiel das Vergrößern und Verkleinern von Masken und Objekten über die gegenläufige Bewegung von Finger (zu- und voneinander). Gesten mit hoher physischer Bewegung und maximalen Gelenkwinkelstellungen werden vermieden. 3D Gesten die eine Manipulation eines Objektes darstellen, sollen mit zwei Händen ausgeführt werden. Die Geschwindigkeit der Gestenausführung soll keine Rolle spielen. Eine langsame Ausführung der Gesten sollte möglich sein. 2D Gesten sollen eine Toleranz gegenüber längeren Berührungszeiten haben.

Gesten werden mit Aktionen so klar wie möglich verbunden. 3D Gesten, in welchen Körperteile bei der Ausführung Objekte darstellen, sind nicht intuitiv. 3D Gesten, in welchen imaginäre Objekte gehalten und pantomimisch benutzt werden, werden als intuitiv angesehen (vgl. [GJM11a]). 3D Gesten, in denen Werkzeuge gezeigt werden, sollen diese Werkzeuge durch Körperteile darstellen. Zum Beispiel kann die Aktion Ausschneiden durch die Geste einer Schere, welche etwas schneidet, dargestellt werden. Aktionen sollen häufiger mit einfachen dynamischen Gesten ausgeführt werden. Kontinuierliche Gesten wie das „scrolling“ sollen in

4. Vorstudie

diskreten Schritten eine Reaktion auslösen, dieses Vorgehen erlaubt geringere feinmotorische Fähigkeiten des Benutzers bei der Ausführung einer Geste.

Aus diesen Ergebnissen können folgende 2D, 3D und Hybrid Gesten präsentiert werden:


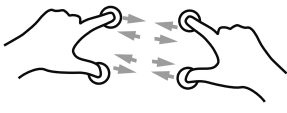
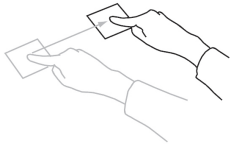
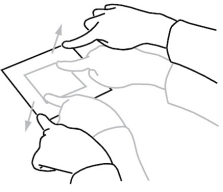

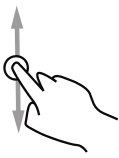
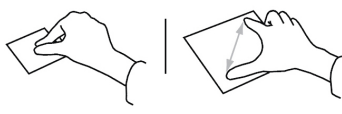

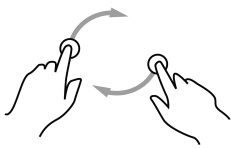
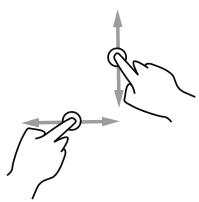
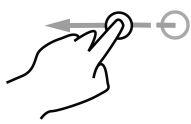

2D-Gesten für Assistenzsysteme		
Position von Werkstück markieren (ID 1)	Ein Foto von Werkstück machen (ID 2)	Position von Maske od. Hinweistext ändern (ID 3)
		
Maske/Hinweistext Größe ändern (ID 4)	Maske/Hinweistext an Position einfügen (ID 6)	In Masken/Hinweistexten scrollen (ID 8)
		
Textschrift vergrößern (ID 9) Kamera Zoom (ID 16)	Objekt auswählen, Dialog bestätigen, Video anhalten/abspielen (ID 27)	Objekt rotieren (ID 14)
		
In Video/Animation Kamera Position ändern (ID 15)	Video von vorne spielen lassen (ID 19)	Ausführung eines Arbeitsschrittes bestätigen (ID 22)
		

Tabelle 4.3.: Erster Teil der 2D-Gesten für Assistenzsysteme und leistungsgeminderte, leistungsgewandelte Arbeiter

Die Bildquellen für die Skizzen der 2D Gestenbeispiele stammen zum Teil aus der Arbeit [WMW09] und der freien Gestenbibliothek Gestureworks der Firma Ideum².


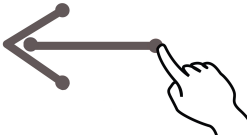


2D-Gesten für Assistenzsysteme – Fortsetzung		
Arbeitsschritt nicht bestätigen (ID 23)	Arbeitsschritt wiederholen (ID 24)	Zwischen Objekten und Masken wechseln (ID 26)
		
Maske/Hinweistext schließen, Objekt löschen/schließen, Dialog nicht bestätigen (ID 28)		
		

Tabelle 4.4.: Zweiter Teil der 2D-Gesten für Assistenzsysteme und leistungsgeminderte, leistungsgewandelte Arbeiter

Eine Beschreibung der 16 unterschiedlichen 2D Gesten ermöglicht ein genaueres Verständnis dieser. Die Aktion des „Fotografierens“ (ID 2) wird durch ein Rechteck aus vier Fingern und einer kurzen Auslösebewegung zum, vom Objekt dargestellt. Zu der Aktion „Maske vergrößern“ (ID 9) wird die „*enlarge/shrink*“ (deu. vergrößern, verkleinern) Geste mit zwei Zeigefingern/Händen von der Mehrheit der Studienteilnehmer benutzt. Die Größe der Projektionsfläche und somit größere grafische Masken und Objekte lassen den Schluss zu, dass bei größeren Objekten gerne beidhändige Gesten verwendet werden. Aufgrund dieser Tatsache wird die Aktion 9 (ID 9) mit der *enlarge/shrink* Geste mit zwei Zeigefingern/Händen verknüpft. Als Gegenstück werden Inhalte in Masken, wie Text (ID 9) oder der Kamerazoom (ID 16), direkt innerhalb der Maske mit Daumen und Zeigefinger vergrößert und verkleinert. Eine Objektrotation (ID 14) wird aufgrund der möglichen Größe von Objekten ebenfalls mit zwei Zeigefingern/Händen ausgeführt. Die Änderung der Kameraposition in einer Animation oder einem Video (ID 15) erfolgt durch ein vertikales, horizontales Tippen und ziehen (scrollen) an dem Video Objekt. In vertikaler Richtung ändert sich der Neigungswinkel der

²siehe <http://ideum.com>

4. Vorstudie

Kamera auf die Szene, in horizontaler Richtung rotiert die Kamera um die Szene herum. Eine einfache Wischbewegung (engl. „swipe“) nach links direkt am Video Objekt startet ein Video neu (ID 19), diese Geste symbolisiert das Springen an den Anfang. Das Wechseln zwischen Masken und Objekten (ID 26) erfolgt durch eine einfache Kreisgeste an den Objekten.

Bis auf vier Gesten nutzen alle 2D Gesten nur einen Finger zur Ausführung von Aktionen. Gezeichnete Gesten bestehen aus einfachen und bekannten Symbolen und werden passend zur Aktion festgelegt. Die Ausführung der Gesten ist leicht und erfordert keine größeren physischen Bewegungen, der Nutzer kann sich die Funktion schnell einprägen.

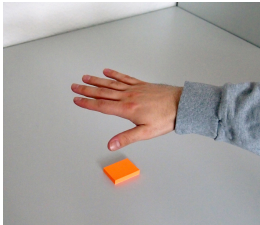
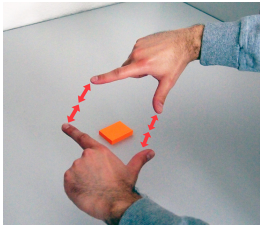
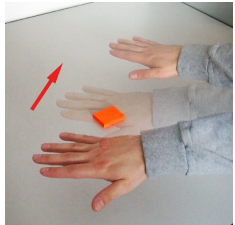
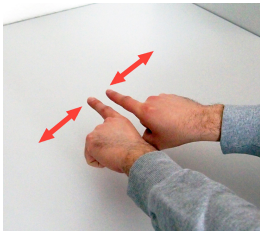
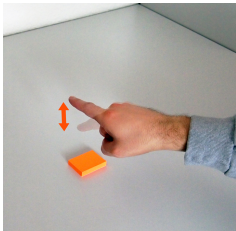
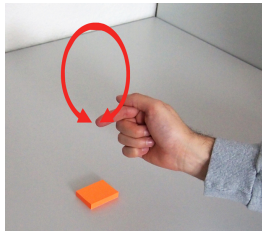
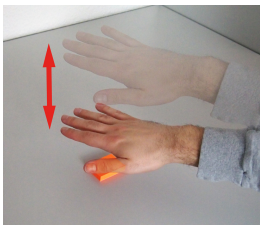
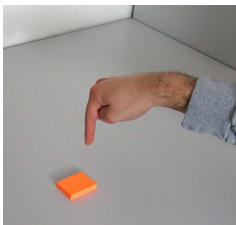
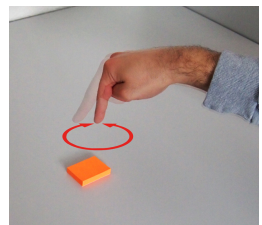
3D-Gesten für Assistenzsysteme		
Position von Werkstück markieren (ID 1)	Ein Foto von Werkstück machen (ID 2)	Position von Maske od. Hinweistext ändern (ID 3)
		
Maske/Hinweistext Größe ändern (ID 4)	Maske/Hinweistext an Position einfügen (ID 6)	In Masken/Hinweistexten scrollen (ID 8)
		
Textschrift vergrößern (ID 9) Kamera Zoom (ID 16)	Objekt auswählen (ID 12)	Objekt rotieren (ID 14)
		

Tabelle 4.5.: Erster Teil der 3D-Gesten für Assistenzsysteme und leistungsgeminderte, leistungsgewandelte Arbeiter

Alle abgebildeten 3D Gesten werden im dreidimensionalen Raum gezeigt.

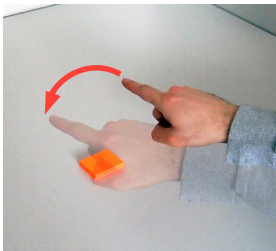
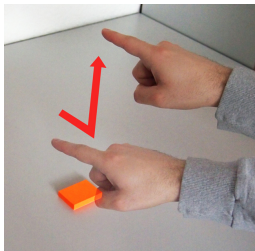
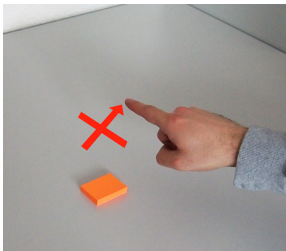
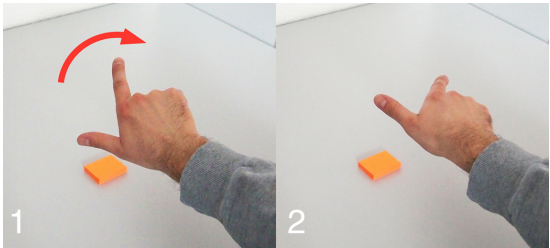
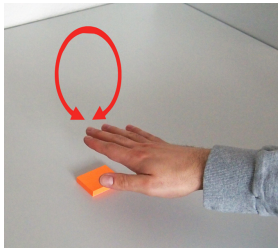
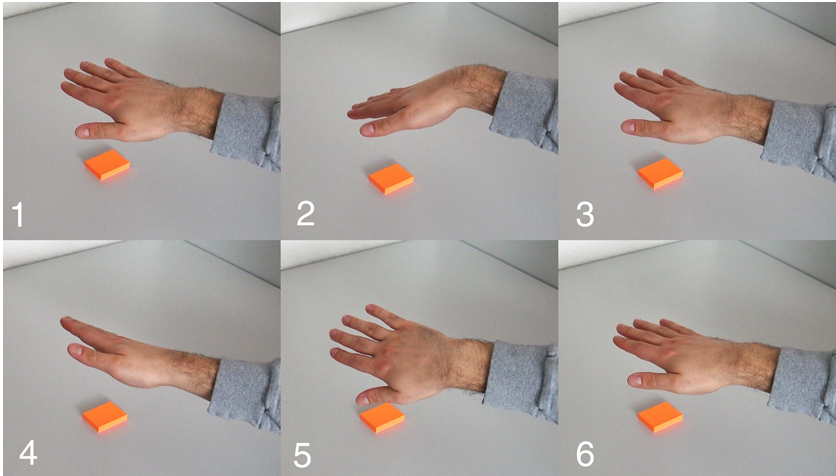
3D-Gesten für Assistenzsysteme – Fortsetzung		
Video von vorne spielen lassen (ID 19)	Ausführung eines Arbeitsschrittes bestätigen (ID 22)	Arbeitsschritt nicht bestätigen (ID 23)
		
Zwischen Maske und Objekten wechseln (ID 26)		Arbeitsschritt wiederholen (ID 24)
		
In Video/ Animation Kamera Position ändern (ID 15)		
		

Tabelle 4.6.: Zweiter Teil der 3D-Gesten für Assistenzsysteme und leistungsgeminderte, leistungsgewandelte Arbeiter

4. Vorstudie

Fortsetzung der 3D Gesten. Ebenfalls wird eine Auswahl von Hybrid Gesten gezeigt.

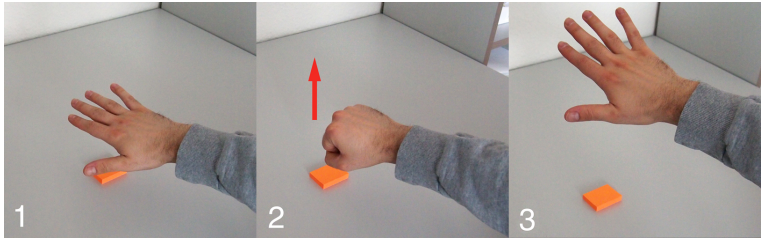
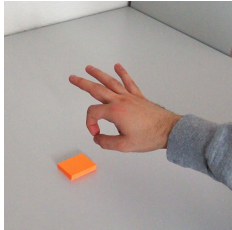
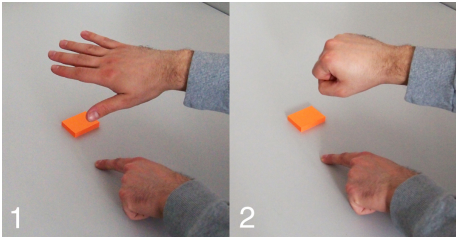
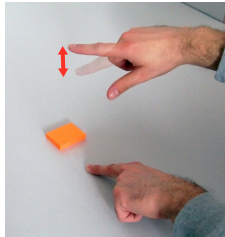
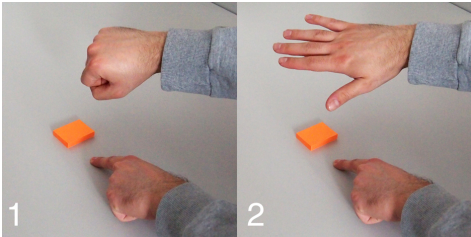
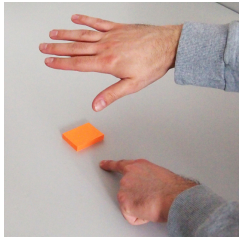
3D-Gesten für Assistenzsysteme – Fortsetzung	
Maske/Hinweistext schließen, Objekt löschen/schließen, Dialog nicht bestätigen (ID 28)	
	
Dialog bestätigen/ Video abspielen (ID 27)	
	
Hybrid-Gesten für Assistenzsysteme	
Position von Werkstück markieren (ID 1)	Foto machen (ID 2)
	
Maske/Hinweistext an Position einfügen (ID 6)	Pos. Maske ändern (ID 3)
	

Tabelle 4.7.: Dritter Teil der 3D-Gesten für Assistenzsysteme und leistungsgeminderte, leistungsgewandelte Arbeiter

Eine Beschreibung der 17 unterschiedlichen 3D Gesten ermöglicht eine genauere Vorstellung dieser. Das Markieren einer Position von Werkstücken (ID 1) erfolgt über die Geste des Haltens der Handfläche über dem Werkstück. Dabei wird die Hand für ein bis zwei Sekunden an der Position gehalten. Die Aktion des „Fotografierens“ (ID 2) wird durch ein Rechteck aus vier Fingern und einer kurzen Auslösebewegung dargestellt. Die Aktion 3 (ID 3), die Positionsänderung von Masken, erfolgt durch die Bewegung der statisch gehaltenen Hand auf einem Pfad. Dabei wird eine selektierte Maske in die Richtung der Handbewegung verschoben. Das Ändern der Maskengröße (ID 4) erfolgt durch zwei Zeigefinger/Hände und die Bewegung dieser in entgegengesetzter Richtung. Die Größenänderung erfolgt proportional zur Bewegung und erhält die Seitenverhältnisse. Ein Einfügen einer neuen Maske (ID 6) erfolgt durch einen 3D Tap über der Zielposition. Ein 3D Tap entspricht einem „Klick“ im dreidimensionalen Raum. Das scrollen in Hinweistexten (ID 8) erfolgt durch das Zeigen eines Fingers horizontal nach rechts oder links und der Bewegung der Hand vertikal im Kreis nach oben und unten. Es wird also ein Kreis mit der Ausrichtung nach rechts oder links gezeichnet. Die scroll Richtung wird durch die Richtung des Kreises bestimmt. Die Schriftvergrößerung (ID 9) und der Kamerazoom (ID 16) werden durch die horizontale auf und ab Bewegung der statisch gehaltenen Hand symbolisiert. Eine Objektauswahl (ID 12) erfolgt durch das vertikale Zeigen auf eine Objekt, wobei die Objektrotation (ID 14) durch eine horizontale Kreisbewegung über dem Objekt stattfindet. Das Zurückspringen in einem Video (ID 19) wird über einen Sprung des Zeigefingers im Halbkreis nach links ausgelöst, dabei befindet sich die Hand über dem Video. Die Bestätigung (ID 22) und nicht Bestätigung (ID 23) eines Arbeitsschrittes wird einmal durch eine Geste des „Abhakens“ und einmal durch eine Geste eines „Kreuzes“ symbolisiert. Beide Gesten werden vertikal, als ob an eine senkrechte Wand, in den Raum gezeichnet. Das Wechseln zwischen Masken (ID 26) wird durch eine „snap“ Geste des Zeigefingers mit abgespreiztem Daumen aus dem Handgelenk ausgelöst. Der ROM Wert dieser Geste ist groß, jedoch würde anstatt dieser Geste eine einfache Wischbewegung nicht robust genug bei der Ausführung sein. Die Wiederholung eines Arbeitsschrittes (ID 24) erfolgt durch einen vertikalen Kreis, dabei wird dieser mit der statisch gehaltenen Hand durchgeführt. Die Änderung der Kameraposition (ID 15) erfolgt durch den Nick- und Rollwinkel (engl. „pitch and roll angle“) der Hand. Wird die Hand nach vorne geneigt (pitch) so ändert sich der Neigungswinkel der Kamera auf die Szene, wird die Hand zur Seite gedreht (roll) so rotiert die Kamera um die Szene herum in Richtung der Drehung der Hand. Die Aktion des Bestätigens von Dialogen und des Abspielens von Videos (ID 27) erfolgt durch die symbolische Geste des „OK“ Zeichens. Aufgrund von visueller Überdeckung (Daumen wird nicht gesehen) werden die Gesten des „Daumens hoch“ und „Daumens runter“ für diese Aktion aus technischen Gründen nicht genommen. Die Aktion 28 (ID 28), das Schließen und Löschen von Masken und Objekten sowie die Nichtbestätigung von Dialogen, wird durch eine „wegwerf“ Geste dargestellt. Dabei wird mit der offenen Hand gegriffen und die geschlossene Hand nach oben geführt und wieder aufgemacht.

Hybrid Gesten setzen sich aus einer 2D Tap und einer 3D Geste zusammen. Insgesamt gibt es 15 unterschiedliche Gesten. Bis aus fünf Gesten gilt dies für alle Aktionen, somit wird immer ein 2D Tap und eine bereits oben definierte 3D Geste kombiniert. Die fünf sich unterscheidende Gesten sind die Gesten für Aktion eins, zwei, drei, sechs und 27. Die Position eines Werkstücks (ID 1) wird durch einen 2D Tap und ein Greifen markiert. Ein Foto (ID 2)

wird durch einen 2D Tap an die Position des Werkstücks und eine Auslösegeste (Hand hält imaginären Fotoapparat und Zeigefinger löst aus) gemacht. Eine Maske wird durch einen 2D Tap und das Öffnen einer Faust neu aufgemacht (ID 6), wobei der 2D Tap die Position der neuen Maske markiert. Die Position einer Maske (ID 3) wird durch einen 2D Tap auf die Maske und das Anzeigen der Zielposition mit der Handfläche geändert. Das Auswählen von Objekten, das Bestätigen von Dialogen und das Abspielen/Anhalten von Videos (ID 12, ID 27) wird nur durch einen 2D Tap ausgelöst. Die Hinzunahme einer 3D Geste macht für diese Aktionen keinen sinnvollen Einsatz möglich.

3D Gesten enthalten keine Bewegungen die eine Schütteln der Hand beinhalten. Gesten enthalten eine statische Handhaltung, ein Zittern kann hier ignoriert werden, solange die Handposition in einem bestimmten Bereich bleibt. Bis auf vier Gesten, werden diese bevorzugt mit einem Finger ausgeführt. Die Gesten werden passend zur Aktion gewählt, wobei sich parallelen zu 2D Gesten zeigen. Bis auf vier Gesten (ID 8, 15, 26, 28) können alle Gesten ohne größeren physischen Aufwand ausgeführt werden. Die ROM Werte der fünf Gesten sind höher im Vergleich zum Rest und die Ausführung etwas schwieriger. Jedoch fiel die Wahl auf diese Gesten, da diese zu den Aktionen passend sind und so die Merkbarkeit unterstützt wird. Gezeichnete Gesten bestehen aus einfachen und bekannten Symbolen.

Hybrid Gesten setzen eine Kombination von 2D und 3D Gesten ein. Zur Selektion von Objekten und Masken ist der 2D Tap die schnellste und einfachste Geste. Je nach Aktion muss bewertet werden, ob eine Hybrid Geste nötig ist. Für die Aktionen mit der ID 22, 23, 24, und 26 kann je nach grafischer Benutzeroberfläche eine 2D Selektion überflüssig sein. Hier könnte der Einsatz reiner 2D oder 3D Gesten sinnvoll sein.

Die Abbildung 4.4 zeigt die Verteilung der bestimmten 2D Gesten auf die 2D Taxonomie. Abbildung 4.5 und 4.6 zeigen die Verteilung der bestimmten 3D und Hybrid Gesten auf die 3D/Hybrid Taxonomie. Die Verteilung der Gesten zeigt eine Übereinstimmung mit den oben genannten Faktoren zur Auswahl relevanter Gesten.

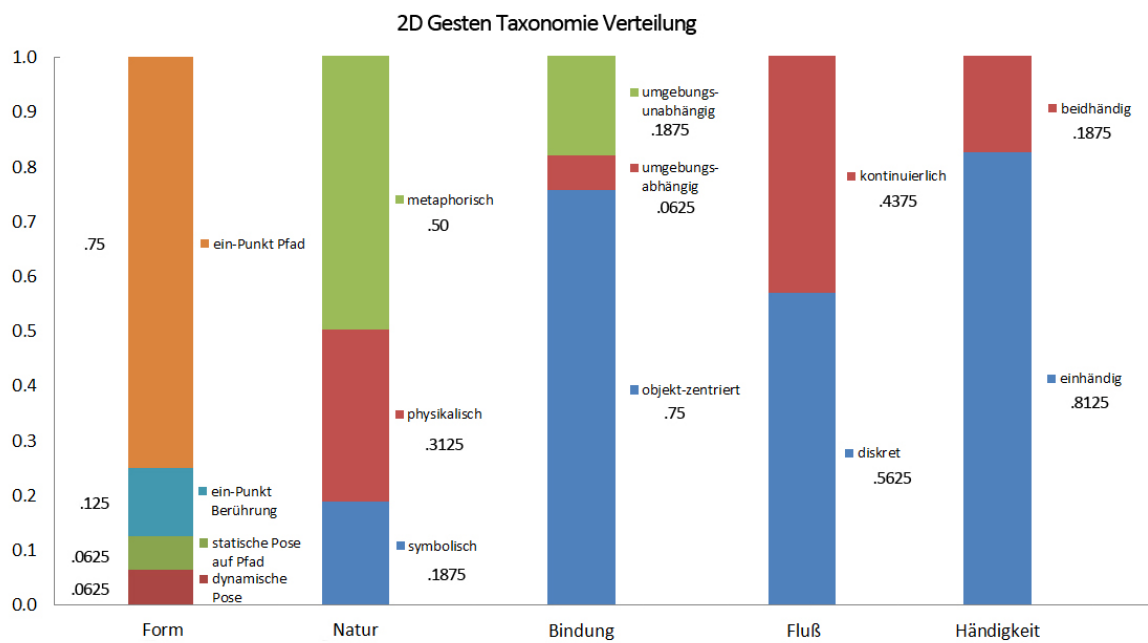


Abbildung 4.4.: Die Verteilung der bestimmten 2D Gesten auf die 2D Taxonomie.

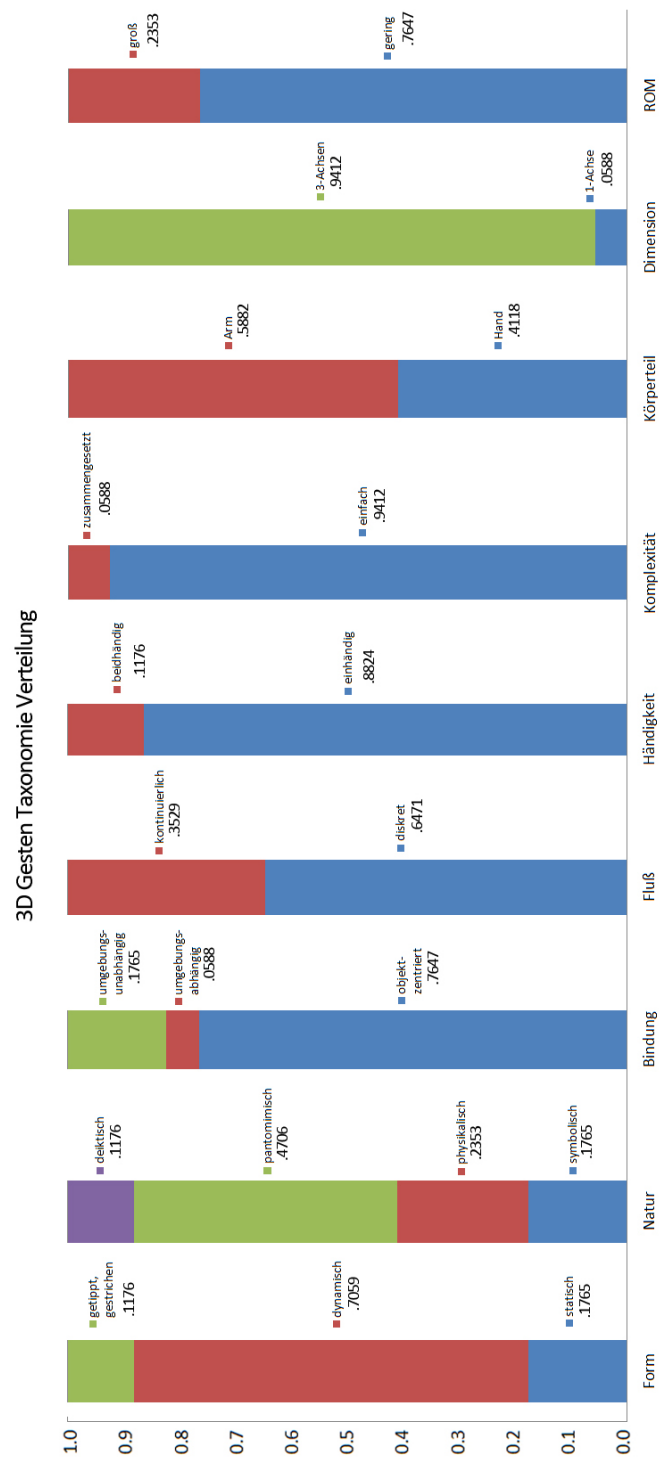


Abbildung 4.5.: Die Verteilung der bestimmten 3D Gesten auf die 3D Taxonomie.

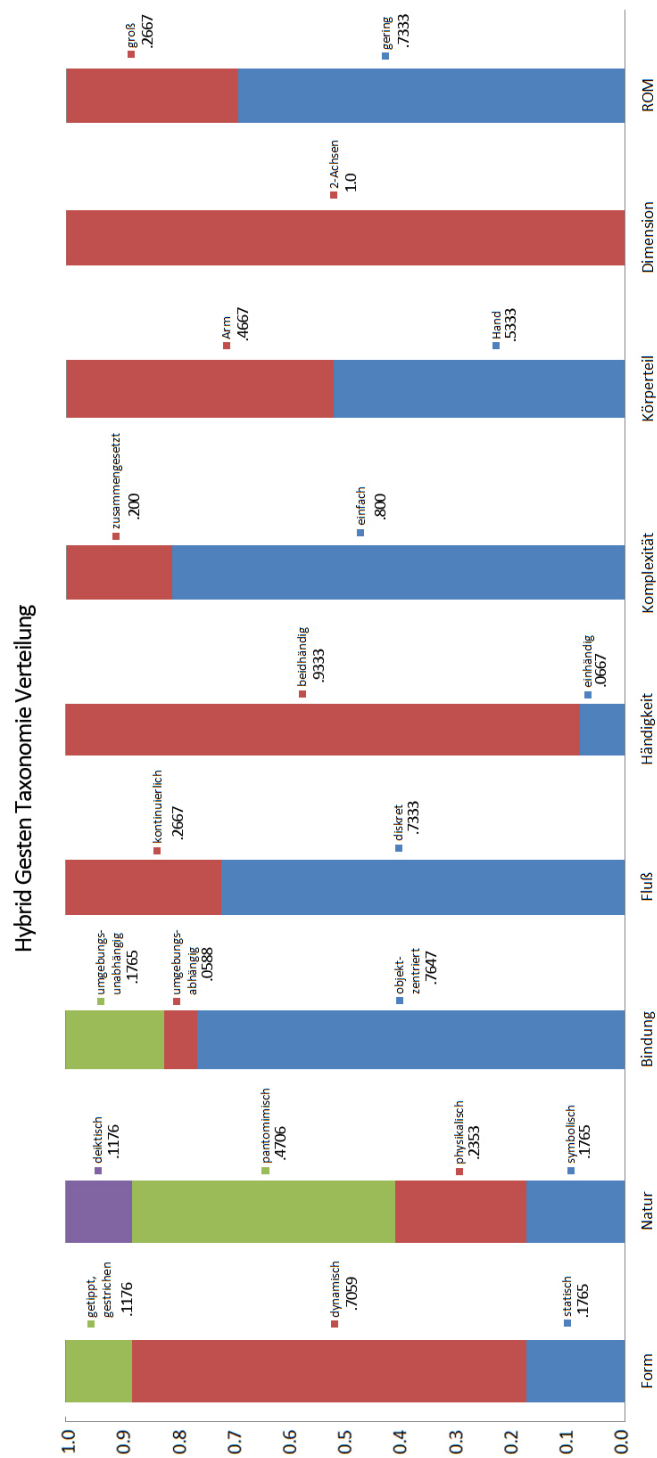


Abbildung 4.6.: Die Verteilung der bestimmten Hybrid Gesten auf die 3D/Hybrid Taxonomie.

5. Interaktionssystem

Dieses Kapitel befasst sich mit den Details des Entwurfs und der Beschreibung der konzipierten Architektur des geplanten Gesten-Interaktionssystems. Zuerst werden die genauen Anforderungen dieser Aufgabe definiert, bevor im nächsten Schritt das entwickelte Lösungskonzept und die einzelnen Komponenten der Architektur untersucht werden können.

5.1. Konzeption

Um die Entscheidungen bezüglich der getroffenen Modellierung detailliert begründen zu können, ist eine Einordnung der Arbeit in den Gesamtzusammenhang von Vorteil. Zu diesem Zweck werden im Folgenden alle Kopplungen der Gesten Schnittstelle zum bestehenden motionEAP Projektumfeld beschrieben. Zudem muss eine bestimmte Menge aller gegebenen Anforderungen diskutiert werden.

5.1.1. Anforderungen

Eine Software bzw. Applikation muss im Allgemeinen gewisse Mindestvoraussetzungen erfüllen, um sinnvoll eingesetzt werden zu können. Diese umfassen neben der Änderbarkeit und Robustheit unter anderem auch die Belastbarkeit und Zuverlässigkeit des Systems. Der im Umfang dieser Diplomarbeit zu implementierende Prototyp eines Gesten-Interaktionssystems muss zudem die Voraussetzung der Austauschbarkeit von Programmkomponenten erfüllen.

Die folgende Liste von Anforderungen ist von allgemeiner Natur und bezieht sich unter anderem auf die Mindestvoraussetzungen von Software im Allgemeinen.

- **Änderbarkeit:** Damit die entwickelte Software zu einem späteren Zeitpunkt erweitert werden kann, soll bei der Planung und Implementierung der Software auf diesen Aspekt geachtet werden. Der Quellcode soll richtig und sinnvoll strukturiert, sowie einfach gestaltet sein. Außerdem soll der Quellcode keine schwer verständlichen Konstrukte beinhalten. Besonders Wert soll auf die Dokumentation des Codes gelegt werden, damit spätere Erweiterung und Wartung der Software keine Probleme mit nicht dokumentierten Schnittstellen bringt.

- **Robustheit:** Da die Software in der Praxis von Menschen eingesetzt wird, die diese Software nicht selbst entwickelt haben, ist es unvermeidlich, dass ein Benutzer die Applikation falsch verwendet und fehlerhafte Eingaben durchführt. Daher soll die Software fehlerhafte Eingaben abfangen und entsprechend darauf reagieren.
- **Zuverlässigkeit:** Die Software soll für den Benutzer eine zuverlässige Unterstützung seiner Arbeit sein. Daher soll die Software korrekt und genau funktionieren. Zu beachten ist, dass es sich bei der Applikation um eine Form der Machbarkeitsstudie handelt.
- **Bedienbarkeit (Usability):** Bei dieser Software soll auf eine gute Bedienbarkeit geachtet werden. Dies bezieht sich nicht nur auf die direkte Benutzungsoberfläche, sondern auch auf funktionale Möglichkeiten bestimmte Prozesse auszuführen.

Auf Grundlage der Anforderungen an Software im Allgemeinen werden in der folgenden Liste die wichtigsten Anforderungen an den Prototypen detaillierter beschrieben.

- **Austauschbarkeit der Programmkomponenten:** Die Module der Gestenerkennung sind austauschbar. Die Logik für die Erkennung jeder Geste wird in separaten Modulen implementiert.
- **Robustheit in der Gestenerkennung:** Die Gestenerkennung soll tolerant gegenüber unwillkürlichen Bewegungen sein. Die Erkennung falscher Gesten soll nicht vorkommen.
- **Zuverlässigkeit in der Gestenerkennung:** Die gleiche Bewegung für eine Geste soll immer die zugeordnete Aktion auslösen. Falsche Bewegungen sollen auch ein Feedback geben.
- **Bedienbarkeit (Usability):** Die Gesteneingabe soll intuitiv und einfach sein. Gesten sollen klare und eindeutige Metaphern für Aktionen bereitstellen.

Die hier diskutierte Anforderungsmenge stellt in erster Linie einen konzeptionellen Rahmen dar, auf dem im Verlauf der Arbeit aufgebaut werden soll. Die genauen Details der Anforderungsumsetzung werden im Folgenden besprochen.

5.2. Architektur

In diesem Kapitel wird die Architektur des Gesten-Interaktionssystems betrachtet. Dazu sieht man in Abbildung 5.1 die konzeptionelle Architektur der Applikation.

Das motionEAP System ist eine C# WPF Applikation, welche auf dem .NET 4.5 Framework aufbaut und die Grafikbibliothek *Media3D* zur Erstellung der Administrations- und Benutzeroberfläche nutzt. Die projizierte Benutzeroberfläche wird mit Hilfe des *Ubi Display* Toolkits eingerichtet. Das *BackendControl* ist die Hauptkomponente der Applikation und steuert alle Funktionalitäten. Der *InputManager* verwaltet und koordiniert die Weitergabe aller Sensordaten, die erfasst werden. Der *GestureRecognitionManager* ist für die Erkennung und Verwaltung von Gesten zuständig und bildet die Implementierung, die für diese Arbeit erstellt wurde. Als Entwurfsrichtung wurde das „Top-Down“ Prinzip gewählt. Aus den Anforderungen wird

ein abstraktes Modell eines Gesten-Interaktionssystems entwickelt. Schrittweise werden dann einzelne Bereiche konkretisiert bis hin zur Programmierung einzelner Module.

Als Architekturmuster wird für das motionEAP System das Model-View-Controller-Architekturmuster verwendet. Auch die Entwicklung dieser Arbeit reiht sich in dieses Architekturmuster ein. Logisch wird die Applikation in das Datenmodell (engl. Model), die Oberfläche (engl. View) und die Steuereinheit (engl. Controller) unterteilt. Dabei wird der Großteil aller ausführbaren Aktionen an die Steuereinheit weitergeleitet und von dieser ausgeführt. Die Komponente Model stellt das benötigte Datenmodell zur Abbildung und Speicherung der verwendeten Daten bereit. Die Komponente View dient dazu die Benutzeroberfläche der Applikation zu realisieren sowie alle vom Benutzer kommenden Eingaben entgegenzunehmen und an die Komponente Controller zu leiten. Die Komponente Controller implementiert die Logik für die Umsetzung aller Funktionalitäten, einschließlich der Gestenerkennung.

Model

Die Model Komponente, welche im Projekt als *Model* bezeichnet wird, bildet Prozessabläufe ab und modelliert Gesten und Hände, die im System zur Eingabe erfasst werden. Sichtbare Hände beinhalten die Definitionen von Finger- und Punkt-Objekten. Punkt-Objekte enthalten zwei- und dreidimensionale Punkte, diese können Berührungspunkte auf der Touch-Oberfläche oder dreidimensionale Punkte im Raum sein. Ebenfalls wird ein Zustand (*State*) modelliert, welcher den aktuellsten Satz von sichtbaren Händen und 2D Berührungspunkten enthält. Das Modell einer Geste enthält eine eindeutige Gestenbeschreibung. Alle Komponenten und Zustandsbeschreibungen werden objektorientiert entworfen.

View

Die Komponente View, welche im Projekt als *SceneManager* bezeichnet wird, dient zur Realisierung der Benutzeroberfläche. Nach dem MVC-Architekturmuster besteht eine strikte Trennung zwischen Benutzereingaben und Ausführung der programminternen Logik. Die Komponente SceneManager nimmt nur Benutzereingaben entgegen und führt keine Logik aus.

Controller

Die Komponente Controller, welche im Projekt als *BackendControl* bezeichnet wird, dient als allgemeine Steuerung der Applikation und stellt die Verbindung zwischen User-Interface, Datenmodell und externen Sensoren und Komponenten her. Die Unterkomponente *GestureRecognitionManager* der BackendControl Komponente dient der Erzeugung der benötigten Gesten Objekte, welche aus dem Datenmodell hervorgehen, und der Ausführung der Gestenerkennung.

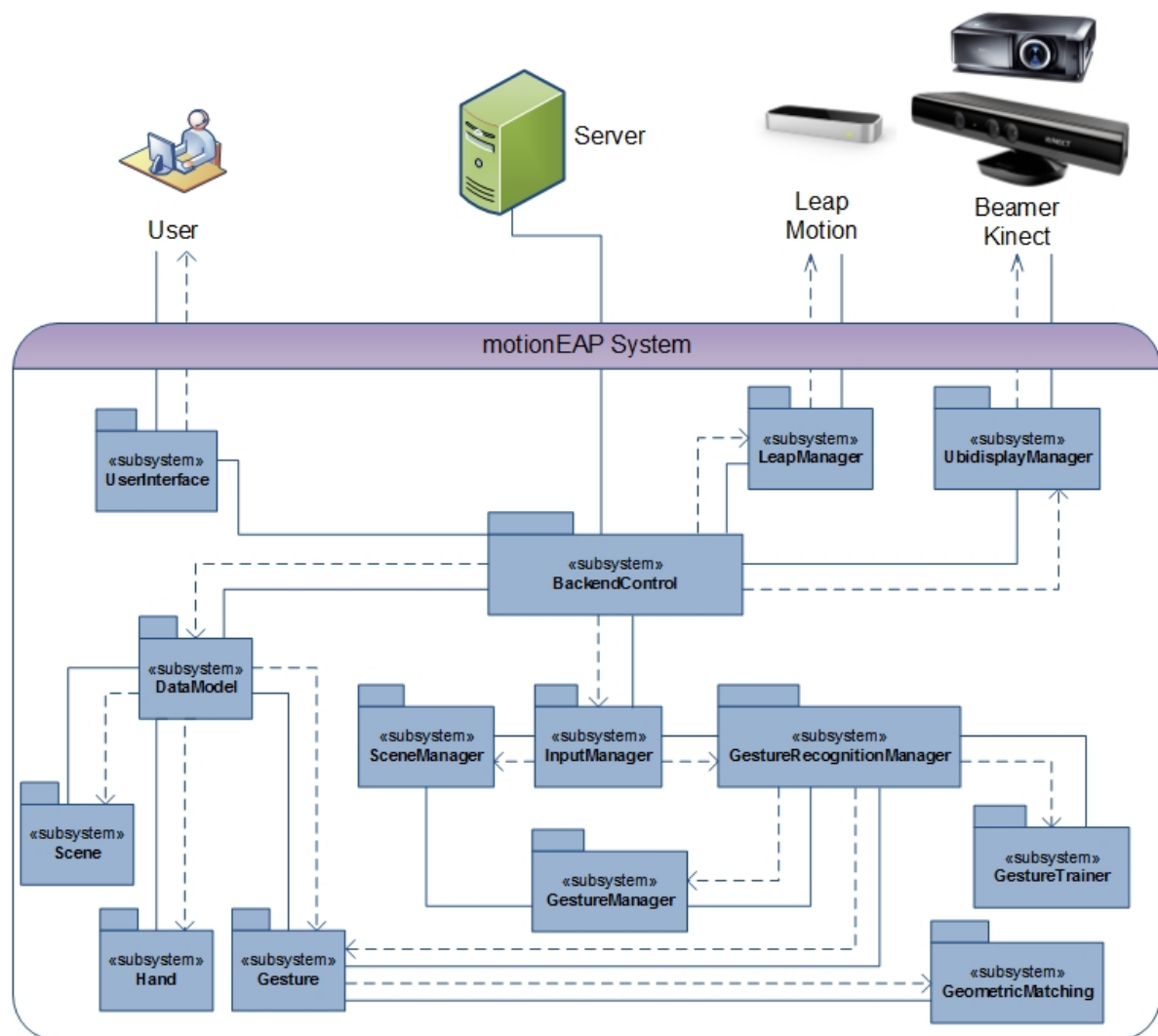


Abbildung 5.1.: Die konzeptionelle Architektur des motionEAP Systems.

5.3. Implementierung

Das Kapitel der Implementierung beschreibt die konkrete Umsetzung einer Gestenerkennung am Assistenzsystem. Die theoretischen Erkenntnisse der letzten Kapitel werden betrachtet und hinsichtlich ihrer praktischen Umsetzung durchleuchtet. Zuerst wird das grobe Gerüst der Implementierung der Gestenerkennung diskutiert. Daraufhin folgt eine detaillierte Betrachtung der Gestenerkennungslogik. Abschließend werden die verwendeten Technologien und Algorithmen innerhalb der Gestenerkennung erklärt.

5.3.1. GestureRecognitionManager und GestureManager

Die Komponente BackendControl startet den InputManager und initialisiert den Empfang von Sensordaten. Diese Sensordaten werden in einem Zustandsobjekt abgelegt und enthalten die sichtbaren Hände über der Arbeitsfläche, als auch Berührungspunkte der Finger mit dieser. Die sichtbaren Hände werden durch einen Array von Hand-Objekten beschrieben. Ein Hand-Objekt enthält eine Liste von Fingern dieser Hand, die dreidimensionale Position der Handfläche im Raum und einen Geschwindigkeitsvektor, welcher die Bewegung der Hand beschreibt. Der GestureRecognitionManager erhält vom InputManager die aktuellsten Eingabedaten in Form eines Zustandsobjekts. Der GestureRecognitionManager erzeugt alle benötigten Gesten Objekte und legt diese in eine Liste ab. Ein Gesten Objekt enthält einen Gestennamen, eine Gesten-ID, den Gestentyp, eine Gestenzustands-ID und Platzhalter für die Speicherung von Zuständen. Diese Zustände können Eintritts-Zustände in eine Geste sein und enthalten, wie oben beschrieben, Sensordaten. Jedes Gesten Objekt enthält die Logik für die Erkennung der Geste. In der Methode *recognizeGestures* des GestureRecognitionManagers übergibt dieser das aktuelle Zustandsobjekt an das jeweilige Gesten Objekt. Alle Gesten Objekte werden iterativ durchlaufen und auf eine Gestenerkennung geprüft. Wird eine Geste erkannt, wird ein Event ausgelöst und an den GestureManager weitergegeben. Der GestureManager verknüpft per „push“ oder „pull“ Methode dieses Gesten Event mit einer bestimmten zuvor festgelegten Aktion. Abbildung 5.2 zeigt die Erkennung einer Handgeste.

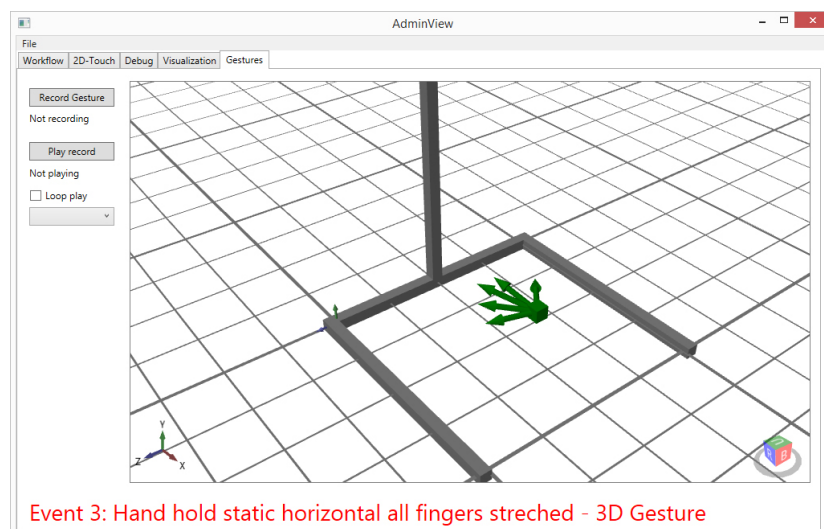


Abbildung 5.2.: Die Erkennung einer Handgeste, gezeigt in der Administrationsmaske des motionEAP Systems.

5.3.2. Logik der Gestenerkennung

Jedes Gesten Objekt enthält die Logik für die Gestenerkennung. 2D Gesten werden, wie in Abbildung 5.3 zu sehen, mit Hilfe von vier Zuständen erkannt. Befindet sich ein Finger im

Raum oberhalb der Arbeitsfläche, wird keine Aktion und 2D Gestenerkennung durchgeführt. Berührt der Finger die Arbeitsfläche wird eine Gestenerkennung durchgeführt und gegebenenfalls eine Aktion ausgelöst. Wird eine Geste erkannt, löst der *GestureRecognitionManager* ein Event aus und gibt diesen mit bestimmten zweidimensionalen Positionsdaten (zum Beispiel die Fingerposition bei einem 2D Tap) an den *GestureManager* weiter. Dieser verknüpft das Event mit einer Aktion (ein 2D Tap ergibt eine Selektion). Die Erkennung komplexer 2D Gesten wird mit Hilfe des *\$P* Algorithmus (vgl. [WVA12]) durchgeführt. Komplexe 2D Gesten sind zum Beispiel gezeichnete Symbole. Hierzu gibt das Gesten Objekt die aktuellen zweidimensionalen Sensordaten an das *GeometricMatching* Modul weiter. Durch einen geometrischen Vergleich der Eingabe mit bereits erfassten Gestenvorlagen kann eine 2D Geste identifiziert und erkannt werden. Eine detaillierte Erklärung des Algorithmus erfolgt im späteren Verlauf dieses Kapitels. Mit Hilfe des Moduls *GestureTrainer* können Gestenvorlagen erstellt und in einem JSON Format gespeichert werden.

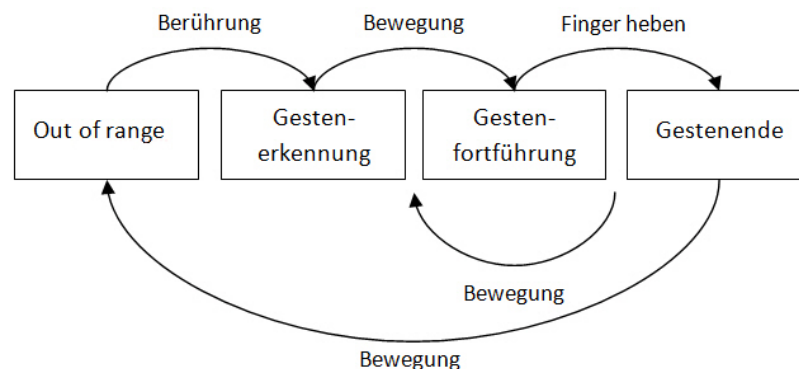


Abbildung 5.3.: Vier Zustände einer 2D Gestenerkennung.

Die Erkennung von 3D Gesten findet ebenfalls in jedem 3D Gesten Objekt statt. Auch hier wird ein Event im *GestureRecognitionManager* bei Erkennung einer Geste ausgelöst und mit bestimmten Daten, wie dreidimensionalen Positionsdaten oder skalaren Werten, an den *GestureManager* übergeben. Dreidimensionale Positionsdaten sind Finger- oder Handpositionen, skalare Werte sind Werte über den Fortschritt einer Geste (zum Beispiel der Fortschritt eines gezeichneten Kreises). Ein großes Problem in der Erkennung von 3D Gesten ist die schwierige Abgrenzung von Gesten zu unwichtigen Bewegungen. Hierbei stellt die Identifikation des Gesteneintritts eine Schwierigkeit dar. Durch die Geschwindigkeit in der Ausführung einer 3D Geste wird der Gesteneintritt ermittelt. Abbildung 5.4 zeigt die Gestenerkennung mit Hilfe von fünf Zuständen. Die Erkennung komplexer 3D Gesten wird mit Hilfe eines um die dritte Dimension erweiterten *\$P* Algorithmus (vgl. [WVA12]) durchgeführt. Komplexe 3D Gesten sind gezeichnete Symbole. Weiter wird der *\$P* Algorithmus zur Aufnahme und Erkennung von Geschwindigkeitsprofilen modifiziert. Ruckartige Bewegungen können so in Gesten identifiziert werden (zum Beispiel das Auslösen in der Foto-Geste). Das Gesten Objekt gibt die aktuellen dreidimensionalen Sensordaten an das *GeometricMatching* Modul weiter. Durch einen geometrischen Vergleich der Eingabe mit bereits erfassten Gestenvorlagen und Geschwindigkeitsprofilen kann eine 3D Geste identifiziert und erkannt werden. Mit Hilfe

des GestureTrainers können Gestenvorlagen erstellt und in einem JSON Format gespeichert werden.

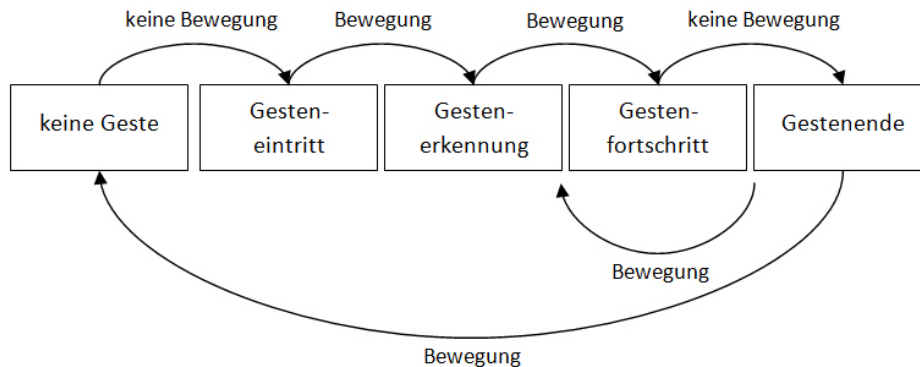


Abbildung 5.4.: Fünf Zustände einer 3D Gestenerkennung.

Hybrid Gesten bauen auf der Implementierung dreidimensionaler Gesten auf und ergänzen diese um eine zweidimensionale Selektionen, einen 2D Tap. Der Gesteneintritt erfolgt nach einem 2D Tap und einer Geschwindigkeitsänderung der Hand, die einen minimalen Wert hat (Hand wird in einer Position gehalten). Die Übergabe von Daten an den GestureManager erfolgt in der gleichen Weise wie im dreidimensionalen Fall und wird um die Positionsdaten der 2D Selektion ergänzt.

Verwendete Datenformate

Die *JavaScript Object Notation*, kurz JSON, ist ein einfach lesbares und kompaktes Datenformat in Textform zum Zweck des Datenaustauschs zwischen Anwendungen. Aufgenommene Gestenvorlagen und Geschwindigkeitsprofile werden in diesem Format geschrieben. Die Gestenvorlagen enthalten eine Liste von aufgezeichneten Gesten. Neben der Bezeichnung der Geste werden aufgezeichnete Positionsdaten der Hand- oder Fingerbewegung festgehalten. Für zweidimensionale Gesten werden die Positionsdaten in Form von X-Y-Koordinaten festgehalten. Für den dreidimensionalen Fall entsprechend in X-Y-Z-Koordinaten. Geschwindigkeitsprofile enthalten in der X-Koordinate einen Zeitschritt und in der Y-Koordinate die Geschwindigkeit der Hand oder des Fingers.

5.3.3. Verwendete Technologien

Ubi Displays

Der Einsatz des Ubi Displays Toolkits (vgl. [HA12]) ermöglicht es jede flache nicht spiegelnde Oberfläche in einen berührungsempfindlichen Bildschirm zu verwandeln. Hierzu werden ein Projektor und eine Kinect Tiefenkamera verwendet. Das Toolkit übernimmt auch die

Kalibrierung der Hardware bei schräg platzierter Aufstellung des Projektors oder der Kinect Kamera und einer nicht ebenen Projektion. Die Einrichtung und Kalibrierung erfolgt schnell und einfach (vgl. [Aly13]). Die Erkennung von Berührungen der Oberfläche erfolgt durch die Auswertung von Tiefendaten der berührenden Hände. Bis zu zehn Finger können auf einer Oberfläche erkannt werden.

Leap Motion

Das Leap Motion SDK ermöglicht eine genaue Erkennung von Händen und Fingern. Positionsdaten von Fingern und maximal zwei Handflächen werden zurückgegeben. Die Erkennung von rudimentären Gesten wird unterstützt. Kreis- und Wischgesten können mit wenig Implementierungsaufwand erkannt werden. 3D Tap Gesten werden in vertikaler oder horizontaler Richtung erkannt. In vertikaler Richtung entspricht ein 3D Tap einer Klickbewegung mit dem Finger. In horizontaler Richtung wird ein Tap ausgelöst, indem ein Finger durch eine definierte vertikale Eben stößt. Durch die Bereitstellung vielfältiger Hand- und Fingerdaten ist eine Erkennung auch von komplexeren Handgesten möglich.

5.3.4. Verwendete Algorithmen

\$P

Der \$P Algorithmus ist ein Template Matching Verfahren zur Erkennung von zweidimensionalen Gesten (vgl. [WVA12]). Das Verfahren ist Teil der \$-Verfahren, einer Familie von Template Matching Verfahren, welche leicht zu verstehen und zu implementieren sind. Template Matching Verfahren beruhen auf Wahrscheinlichkeiten, Statistiken und Signalverarbeitung. Anhand einer Datenvorlage wird in einem neuen Satz von Daten versucht diese Vorlage zu erkennen. Eine aufgenommene Geste wird als Punktwolke gespeichert. Diese Punktwolke enthält die Positionsdaten eines Fingers oder einer Hand, die während der Gestenbewegung erfasst wurden. Die Aufnahme und Erkennung einer Geste findet erst statt, wenn ein Finger oder eine Hand auf einer Position verharren und erst dann die Gestenbewegung durchführen. Die Eingabe von Strichen einer komplexen Geste (zum Beispiel der Zeichnung eines Sternes) ist unabhängig von der Reihenfolge und Richtung (siehe Abbildung 5.5). Dies erlaubt eine im Vergleich zu anderen Verfahren robustere und schnellere Erkennung von Gesten.

Die Erkennung von Gesten erfolgt mit Hilfe der *euklidischen Distanz* zwischen den gespeicherten Gestenpunkten und den aktuellen Bewegungsdaten (siehe Abbildung 5.5). Dabei wird der Abstand zwischen den gespeicherten Punkten und den neu erfassten Punkten untersucht. Diese Untersuchung ergibt ein Zuordnungsproblem, welches als ein diskretes Optimierungsproblem aus der Graphentheorie bekannt ist. Gespeicherte Gestenpunkte sollen den erfassten Bewegungspunkten, unter der Bedingung des kleinsten Abstandes zwischen diesen, zugewiesen werden. Durch die Anwendung des *ungarischen Algorithmus* kann dieses Zuordnungsproblem gelöst werden. Eine detaillierte Erklärung des ungarischen Algorithmus kann in einschlägiger Literatur zur Graphentheorie eingesehen werden.

Die Komplexität des \$P Algorithmus liegt bei $O(n^{2.5} \cdot T)$, wobei n die Anzahl der aktuellen Punkte für die Erkennung und T die Anzahl der Gestenvorlagen für eine Geste darstellen. Um eine über 90 prozentige Erkennungsrate zu gewährleisten sollten mindestens fünf verschiedene Gestenvorlagen einer Geste von fünf verschiedenen Personen eingespielt werden. Dabei sollte die Anzahl der geprüften aktuellen Punkte für eine Erkennung bei 32 Punkten liegen, um eine schnellste Ausführung des Algorithmus sicherzustellen.

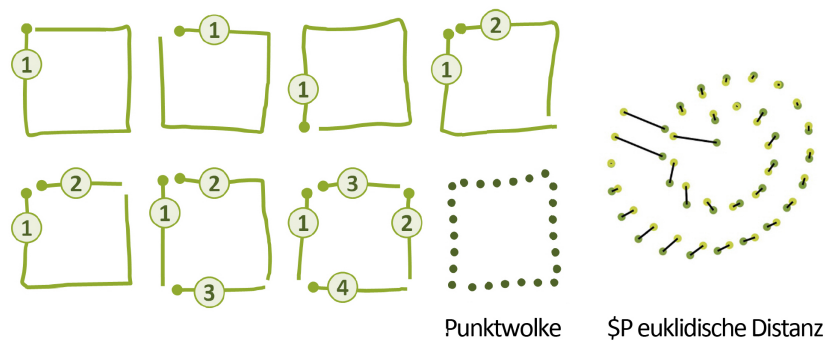


Abbildung 5.5.: Möglichkeiten eine Quadrat-Geste zu zeichnen und die Zuordnung euklidischer Distanzen. (vgl. [WVA12])

Die Erweiterung des \$P Algorithmus um die dritte Dimension ergibt die Untersuchung und den Vergleich von dreidimensionalen Punkten. Der Beispiel-Code in Abbildung 5.6 zeigt die Methode *GestureRecognition* zur Initialisierung einer Gestererkennung. Die Gestenvorlagen werden in der Datei „gesture-lib.json“ gehalten.

5.3.5. Fazit

Die Erkennung dreidimensionaler Gesten ist komplex. Ein großes Problem ist die schwierige Abgrenzung von Gesten zu unwichtigen Bewegungen im Raum. „False positive“ Fehler sind die Folge. Bewegungen werden als Gesten erkannt, wobei diese Bewegungen keine Gesten sind. Durch die Hinzunahme bestimmter Merkmale einer Geste, wie einer besonderen Ausführungsrichtung, können diese Fehler reduziert werden. „False negative“ Fehler treten ebenfalls auf. Diese beschreiben ausgeführte Gesten die nicht erkannt werden. Gründe dieser Fehler liegen meistens nicht in der Erkennung der Gesten, sondern dass der Benutzer nicht weiß wann, wie und wo er Gesten einsetzen kann und sollte. Das Erfassen und Vermeiden von false positive Fehlern trägt maßgeblich zum Erfolg eines Gesten-Interaktionssystems bei. Die Balance aus wenigen false positive und false negative Fehlern ist entscheidend. Durch den Einsatz hybrider Gesten können beide Fehlerarten gezielt verringert werden. Ein Feedback vom System bei false positive und negative Fehlern würde den Benutzer sensibilisieren.

Die Erkennung zwei- und dreidimensionaler Gesten durch den \$P Algorithmus kann bei komplexen Gesten optimiert werden. Je mehr Gestenvorlagen verschiedener Benutzer vorliegen desto höher kann die Erkennungsrate ausfallen. Mehrere Gestenvorlagen einer Geste

5. Interaktionssystem

vom gleichen Benutzer haben keine positiven Auswirkungen auf die Erkennung. Bekannte Gesten wie Buchstaben, Zahlen und Symbole erhöhen die Erkennungsrate. Dies gilt ebenfalls für einfache Gesten, die aus wenigen Strichen bestehen.

```
1 private List<Point> points;
2 private int strokeID = 0;
3
4 private void GestureRecognition(Frame frame, Hand hand, Finger finger)
5 {
6     // Current finger-tip position and stroke
7     Point point = new Point(finger.Position.x, finger.Position.y, finger.Position.z, this.strokeID);
8
9     if (this.points.Count >= 1)
10    {
11        if (Geometry.SqrtEuclideanDistance(point, this.points[this.points.Count - 1]) > 0.5)
12        {
13            if (!this.stroking)
14            {
15                this.stroking = true;
16                this.points[this.points.Count - 1].StrokeID++;
17                this.strokeID++;
18            }
19            point.StrokeID = this.strokeID;
20            this.points.Add(point);
21        }
22        else
23        {
24            if (this.stroking)
25            {
26                // Try to recognize a gesture
27                if (this.points.Count >= 32)
28                {
29                    // Try to recognize a gesture
30                    List<KeyValuePair<string, float>> result = PointCloudRecognizer.Classify(
31                        new GestureData(this.points, "live"),
32                        GestureData.ReadGesturesFromJson(@"\gesture-lib.json"));
33                    if (result[0].Value < 0.2)
34                    {
35                        Console.WriteLine("Das war eine {0} Geste!", result[0].Key);
36                    }
37                }
38                this.stroking = false;
39                this.points.Clear();
40            }
41        }
42    }
43    else
44    {
45        this.points.Add(point);
46    }
47    if (this.points.Count > 1000)
48    {
49        this.points.RemoveRange(0, 10);
50    }
51 }
```

Abbildung 5.6.: Methode zur Initialisierung einer 3D Gestenerkennung mit Hilfe des \$P Algorithmus.

6. Evaluation

Dieses Kapitel beschreibt die Durchführung einer Usability Studie, anhand welcher die Benutzerfreundlichkeit der getroffenen Gestenauswahl und somit der Gestenarten am Assistenzsystem ermittelt wird. Der Ablauf der Studie wird durch das verwendete Vorgehensmodell erläutert. Hierzu werden die Teilnehmer, die verwendeten technischen Hilfsmittel zur Ausführung, als auch die Durchführung beschrieben. Es folgt die statistische Analyse der Studienergebnisse und die Präsentation dieser. Anschließend werden detaillierte Schlussfolgerungen zur Nutzbarkeit von zwei- und dreidimensionalen Gesten an Assistenzsystemen aus den Ergebnissen gezogen und Design Richtlinien für den Gestenentwurf und die Gesteneingabe aufgestellt.

6.1. Evaluationsdurchführung — Usability Studie

6.1.1. Verwendetes Vorgehensmodell

Die Methodik der Studie ist in Anlehnung an die Methodologie von Wobbrock in der Arbeit [WMW09] konzipiert. Die Durchführung orientiert sich an den Richtlinien zu Usability Tests (vgl. [RCS08]).

Ziel der Studie ist die Erfassung qualitativer und quantitativer Daten zur Benutzerfreundlichkeit der festgelegten Gesten-Sets und Gestenarten für Assistenzsysteme im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter. Zwölf Teilnehmer werden eingeladen. Jedem Teilnehmer werden 18 festgelegte Aktionen erklärt. Zu diesen 18 Aktionen werden jeweils eine 2D, 3D und Hybrid Geste gezeigt und die Ausführungsmerkmale beschrieben. Der Teilnehmer führt zu jeder Aktion die 2D, 3D und Hybrid Geste am System aus und erhält ein subjektives Feedback. Anschließend werden die ausgeführten Gesten bewertet. Hierzu füllt der Teilnehmer einen Fragebogen pro Aktion aus. In diesem wählt er eine passende Geste für die Aktion aus und gibt den Schwierigkeitsgrad der Ausführung an. Die Bewertung der Ausführungsschwierigkeit erfolgt anhand einer fünfstufigen Likert-Skala (vgl. [FH10], S. 45), mit der ersten Stufe mit der Bedeutung leicht und der fünften Stufe mit der Bedeutung schwer. Weiter kann er einen Grund für eine leichte oder schwere Ausführung nennen und die prozesssicherste Geste wählen. Eine prozesssichere Geste wird hier als eine Geste bezeichnet, welche am wenigsten Fehleingaben produziert. Während der Ausführung der Gesten werden Notizen zu Besonderheiten in der Ausführung gemacht. Zu Beginn der Studie füllt der Teilnehmer einen Fragebogen zur Person aus. Am Ende des Usability Tests füllt der Teilnehmer einen weiteren Fragebogen zu seiner abschließenden Gestenpräferenz aus.

Die Auswertung der ermittelten Daten erfolgt statistisch. Weitere qualitative Daten aus den Fragebögen zur Bewertung der Gesten, als auch aus den Notizen zur Ausführung fließen in die Evaluationsergebnisse mit ein.

6.1.2. Teilnehmer

Die zwölf Teilnehmer der Studie setzen sich aus neun männlichen und drei weiblichen Personen zusammen. Das Durchschnittsalter liegt bei 29 Jahren ($M = 29.25$, $SD = 9.81$), wobei zehn der zwölf Teilnehmer Rechtshänder und zwei Teilnehmer Linkshänder sind. Bei den Teilnehmern handelt es sich um zehn Studenten der Universität Stuttgart, vorwiegend Informatiker und aus zwei Angestellten und Selbständigen. Aufgrund dem frühen Entwicklungsstand des montionEAP Assistenzsystems kann das System in keiner Werkstätte für Behinderte Menschen getestet werden. Zur Ermittlung von Daten werden für diese Studie somit leistungsgewandelte und uneingeschränkte Personen befragt. Die Teilnahme an der Studie ist freiwillig. Alle Teilnehmer haben Erfahrungen mit 2D Gesten, vor allem aus der alltäglichen Nutzung von Smartphone und Tablet. Ein Teilnehmer gibt die Bedienung eines Fernsehers und die Nutzung von 2D Gesten auf einer Fernbedienung an. 58% der Teilnehmer haben Erfahrungen mit 3D Gesten. Vor allem wird der Unterhaltungsbereich genannt. Microsofts Kinect, Nintendos Wii und Sonys¹ EyeToy werden zu Spielzwecken genutzt.

6.1.3. Apparatur

Die technischen Hilfsmittel, die in dieser Studie benutzt werden, können in drei Bereiche eingeteilt werden. Technisches Equipment, Fragebogen und Gegenstände zur Hilfestellung für ein simuliertes Gesten-Interaktionssystem. Der Eingangsfragebogen enthält neben Fragen nach Alter, Geschlecht, Händigkeit, auch Fragen zu Erfahrungen mit 2D und 3D Gesten. Die weiteren Fragebögen enthalten Fragen zur Bewertung von Gesten, einmal für jede Aktion und für alle Aktionen. Die Fragebögen werden mit Hilfe von Google Forms² erstellt und per Webmaske den Teilnehmern zur Verfügung gestellt. Die Gegenstände zur Hilfestellung für das simulierte Gesten-Interaktionssystem bestehen aus einem Arbeitstisch mit ab geklebter Arbeitsfläche, einer einfachen Schraubzwinge und verschiedenen aus Pappe hergestellten Beispielmasken zur Veranschaulichung einer grafischen Oberfläche. Zum technischen Equipment gehören zwei Notebooks zum Erfassen von Fragebögen, als auch von Notizen. Die Abbildung 6.1 zeigt den Testaufbau.

6.1.4. Durchführung

An zwei Tagen nahmen zwölf Teilnehmer an der Studie teil. Jeder Teilnahme dauerte eine halbe Stunde. Jedem Teilnehmer wurde der Hintergrund der Studie erklärt und das Ziel kurz

¹siehe <http://www.sony.com>

²siehe <http://www.google.com/drive/apps.html>



Abbildung 6.1.: Technische Hilfsmittel der Benutzerstudie zur Usability von Gesten.

zusammengefasst. Eine Zustimmungserklärung zur Verwendung von Daten wurde vorgelegt und vom Teilnehmer unterzeichnet. Ein Eingangsfragebogen zur Person und zu Erfahrungen mit Gesten wurde ausgefüllt. Das Testszenario wurde beschrieben. Daraufhin wurde dem Teilnehmer jede Aktion beschrieben und die zugeordneten 2D, 3D und Hybrid Gesten gezeigt. Der Teilnehmer wiederholte die Gesten und bewertete diese im Fragebogen zu der Aktion. Dies wurde 18 mal für 18 Aktionen wiederholt. Weiter wurden zur jeder Aktion Notizen zur Ausführung der Gesten gemacht. Die Reihenfolge der abgefragten Aktionen wurde anhand eines bestimmten lateinischen Quadrates (vgl. [FH10], S. 84) gewählt, um Nebeneffekte in der Ausführung der Gesten zu reduzieren. Für diese Zwecke wurde das „Balanced Latin Square“ genutzt, dieses wurde in sechs auf sechs Felder unterteilt. Somit bekam jeweils eine Gruppe von zwei Teilnehmern eine verschiedene Reihenfolge von Aktionen. Am Ende eines Tests füllte jeder Teilnehmer einen Fragebogen zur persönlichen gesamtheitlichen Gestenpräferenz aus.

6.1.5. Darstellung der Evaluationsergebnisse

Nach Aufbereitung und Auswertung der gesammelten Daten bevorzugen in der Gesamtheit 75% der Teilnehmer 2D Gesten, 17% der Teilnehmer Hybrid Gesten und nur 8% der Teilnehmer 3D Gesten zur Interaktion mit einem Assistenzsystem.

Zweidimensionale Gesten werden einheitlich als intuitiv angesehen. Das instinktive Anfassen möglicher grafischer Elemente wie Menüs und einfacher Objekte sowie das Zeichnen von Symbolen auf einer Oberfläche wird als einfach und natürlich beschrieben. Das einhändige Ausführen der meisten 2D Gesten wird als nicht physisch anstrengend empfunden. Die Zuordnung von 2D Gesten zu den Aktionen wird mehrheitlich als passend, eindeutig und nicht verwirrend gesehen. 2D Gesten sind aus dem täglichen Leben bekannt. Einige 2D Gesten werden jedoch auch als zu umständlich bewertet.

6. Evaluation

Dreidimensionale Gesten werden als fehleranfällig und unpräzise beschrieben. Die Manipulation von zweidimensionalen Menüs, Masken und Objekten mit 3D Gesten wird als nicht natürlich empfunden. Die Selektion von zweidimensionalen Objekten durch 3D Gesten wird als schwierig und verwirrend gesehen. Die Zuordnung einiger 3D Gesten zu Aktionen wird als nicht passend beschrieben. 3D Gesten umschreiben Aktionen nicht deutlich genug. Handstellung werden vergessen. 3D Gesten werden sehr kontextabhängig bewertet. Ist die Arbeitsfläche beispielsweise großflächig mit Werkstücken überdeckt, werden 3D Gesten als funktional angesehen. Das Gleiche gilt beispielsweise für verschmutzte Hände, oder getragene Handschuhe. 3D Gesten benötigen keine freie Arbeitsfläche und können je nach Arbeitsablauf schneller und besser ausgeführt und in diesen integriert werden.

Hybrid Gesten werden als prozesssicher beschrieben. In der Kombination eines 2D Taps und einer 3D Geste wird in dem 2D Tap eine einfache Selektionsmöglichkeit gesehen. Die Selektion bei mehreren Objekten fällt leicht, jedoch erfordert eine Hybrid Geste zur Ausführung von Aktionen zwei Hände. Bei der Ausführung einer Hybrid Geste wird ein Wechsel der Hände bei den Teilnehmern festgestellt. Die Selektion erfolgt mit der dominanten Hand, wobei die 3D Geste mit der anderen Hand ausgeführt wird.

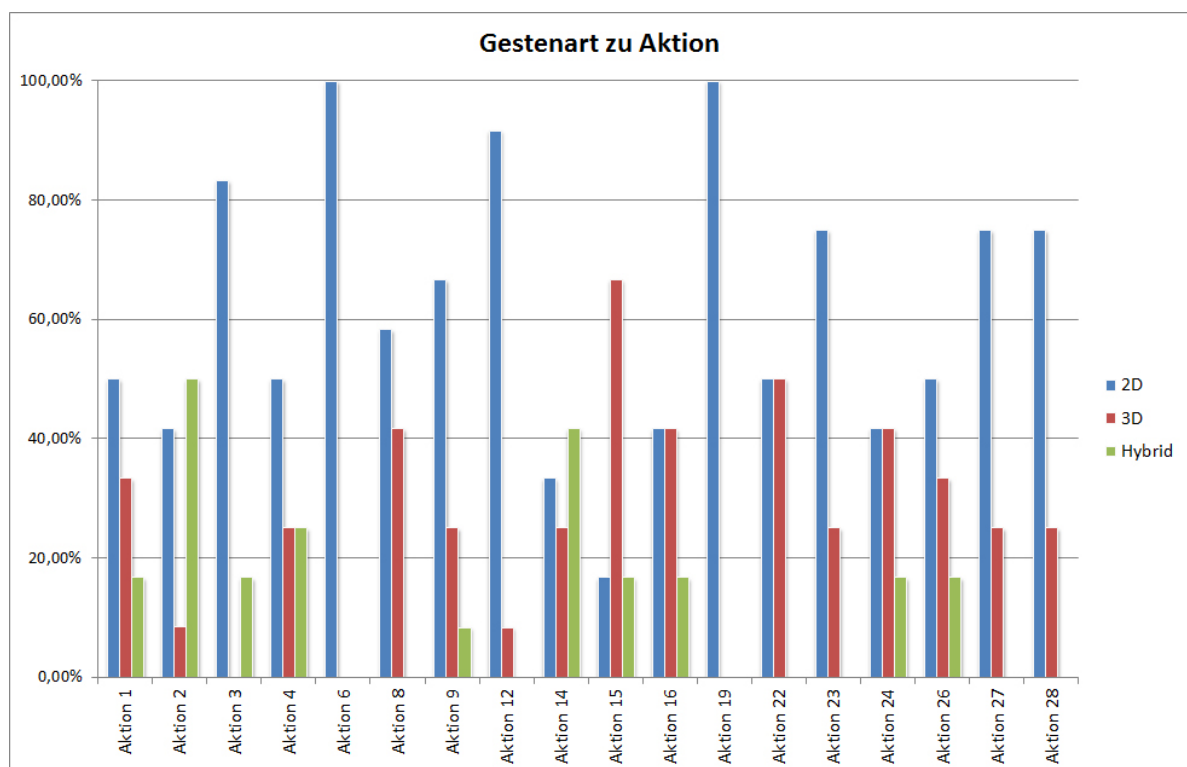


Abbildung 6.2.: Studien-Ergebnisse. Die Auswahl von Gesten zu Aktionen.

Abbildung 6.2 zeigt die mehrheitliche Wahl von 2D Gesten für fast alle Aktionen. Auffällig sind besonders die Aktionen drei, sechs, zwölf, 19, 27 und 28. Über dreiviertel aller Teilnehmer wählen für diese Aktionen eine 2D Geste. Die Aktionen haben gemeinsame Nenner, sie

manipulieren die Position von Masken, Objekten und bestätigen Dialoge oder lehnen diese ab. Alle 2D Gesten für diese Aktionen verwenden ein einfaches Ziehen, Tippen oder Wischen mit einer Hand. Die Aktionen acht und neun manipulieren die Position und Größe von Schrift, auch hier bestehen die 2D Gesten aus einem einfachen scrollen und der enlarge/shrink Geste mit Daumen und Zeigefinger. Beide Gesten werden mit einer Hand ausgeführt und kommen ohne große physische Bewegung der Hand aus. Die 2D Gesten der Aktionen eins, vier, 22, 23, und 26 sind mit größerer physischer Bewegung verbunden. Die 2D Geste für Aktion vier wird mit zwei Händen ausgeführt, die 2D Gesten für Aktion 22, 23 und 26 beinhalten das Zeichnen von Symbolen. Die Gesten für die Aktionen 16 und 24 sind nicht eindeutig einer Gestenart zuzuordnen. Aktion 15, das Ändern der Kameraposition, wird eindeutig der 3D Geste zugeordnet. Die 3D Geste der Aktion 22 ist gleichwertig mit der 2D Geste. 3D Gesten für die Aktionen acht, 16, 22, 24 und 26 erreichen eine höhere Akzeptanz. Die Aktion 2 und die Aktion 14 werden mit Mehrheit durch eine Hybrid Geste ausgeführt.

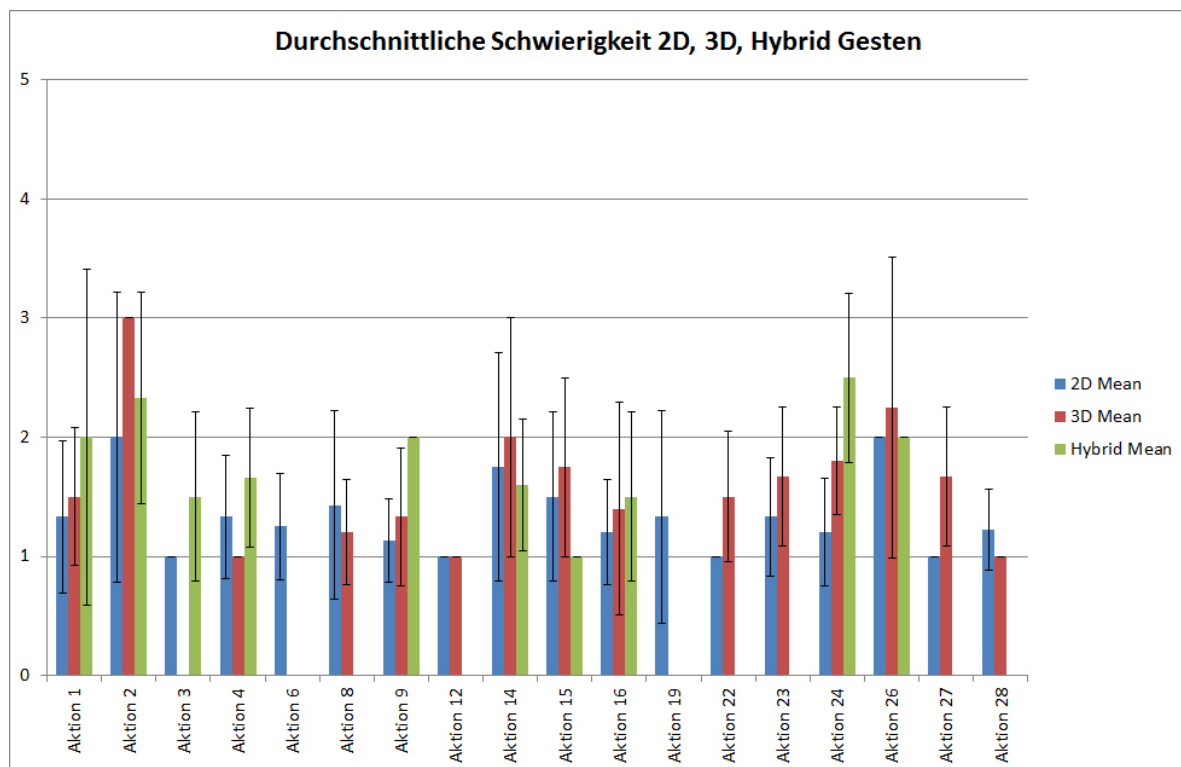


Abbildung 6.3.: Studien-Ergebnisse. Durchschnittliche Schwierigkeit der Ausführung von Gesten. (Angabe sind der Durchschnitt und die Standardabweichung)

Abbildung 6.3 zeigt die Auswertung der Schwierigkeit der Ausführung der gewählten Geste. Zu erkennen ist, dass 2D Gesten in fast jeder Aktion als leichter ausführbar eingestuft werden, wobei der Wert eins als am leichtesten und der Wert fünf als am schwierigsten definiert ist. Zu beachten ist, dass nicht jede Gestenart in jeder Aktion bewertet ist. Nur die gewählten Gesten werden bewertet.

6. Evaluation

Folgende Erkenntnisse zu den einzelnen Gesten können somit bestätigt werden. Die 2D Geste des Markierens durch ein Lasso für Aktion eins ist bei größeren Objekten physisch anstrengend. Die 2D Geste des Fotos für Aktion zwei ist zu komplex. Die Geschwindigkeit des 2D Doppel Tap für Aktion sechs wird in der Ausführung variiert. Die 2D Geste des scrollens für Aktion neun verursacht Verdeckungsprobleme. Die synchrone Handbewegung der 2D Rotationsgeste für Aktion 14 ist bei größeren Objekten schwierig. Die 3D Geste des Fotos für Aktion zwei ist zu komplex. Die 3D Geste für die Aktion acht des scrollens wird als komplex und physisch anstrengend beschrieben. Ein Zielen in der 3D Geste für Aktion zwölf ist ungenau. Die Ausführung der 3D Rotationsgeste für Aktion 14 ist bei kleinen Objekten schwierig. Das Ändern der Kameraposition durch die 3D Geste in Aktion 15 erfordert feinmotorische Fähigkeiten und setzt eine Einarbeitung voraus. Die 3D Gesten für Aktion 26 (wechseln zwischen Masken) und Aktion 28 (Masken schließen und Objekte löschen) werden als physisch anstrengend beschrieben. Aufgrund der Kombinierten 2D und 3D Gesten ergeben sich bei Hybrid Gesten die selben Erkenntnisse wie bei 3D Gesten. Allgemein kann festgehalten werden, dass bei Ausführung einer Geste nicht beteiligte und ungenutzte Finger nicht zur Faust geballt werden.

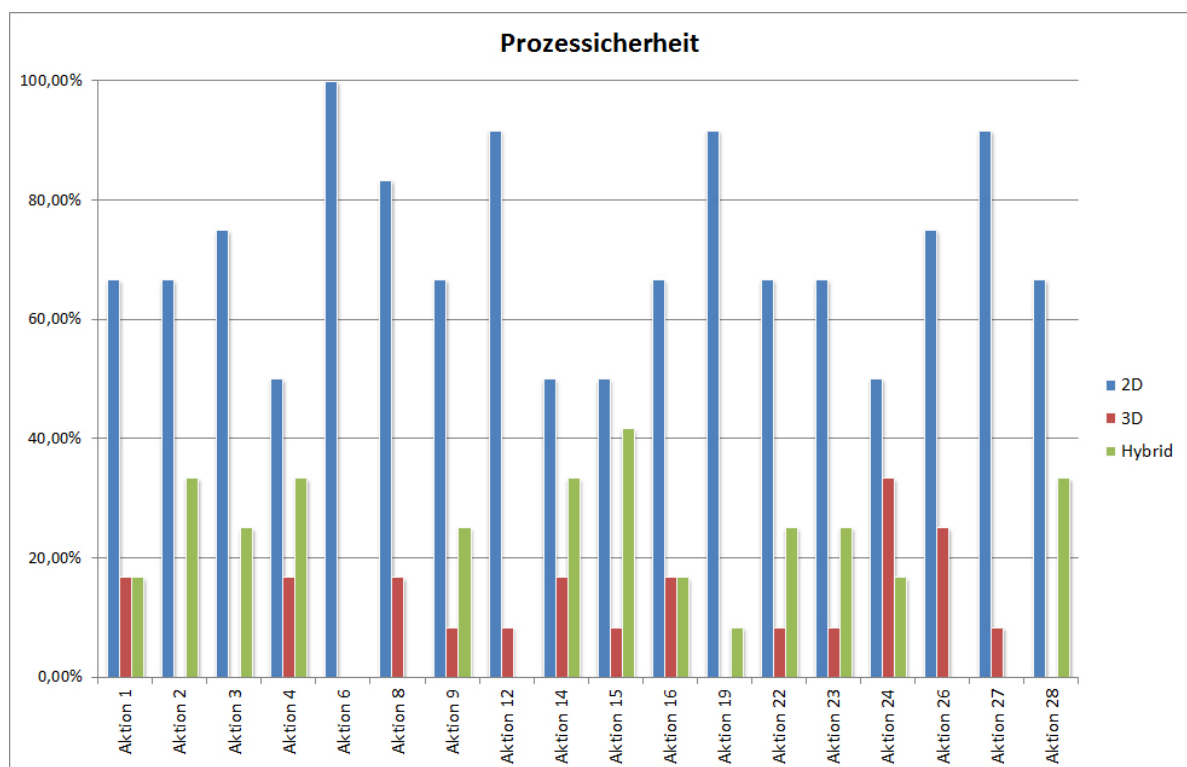


Abbildung 6.4.: Studien-Ergebnisse. Prozesssicherheit von Gesten.

Abbildung 6.4 zeigt die Wahl der prozesssichersten Geste zur Aktion. Eine Geste wird als prozesssicher bezeichnet, wenn sie am wenigsten Fehleingaben auslöst. Auffällig ist eine ähnliche Verteilung der 2D Gesten zur 2D Gestenpräferenz aus Abbildung 6.2. Mehrheitlich

werden 2D Gesten als prozesssicher gekennzeichnet. Hybrid Gesten sind prozesssicherer als 3D Gesten.

6.2. Schlussfolgerungen und Diskussion

Aus den Ergebnissen der Usability Studie lässt sich eine klare Tendenz zur Interaktion durch zweidimensionale Gesten ablesen. Die Manipulation von zweidimensionalen Menüs, Masken und Objekten durch 3D Gesten wird in den meisten Fällen als nicht intuitiv empfunden. Der Kontext der Aktionen passt zu den Metaphern der 2D Gesten. Zweidimensionale Masken und Objekte werden bevorzugt mit 2D Gesten bearbeitet. Die Ergebnisse der Usability Studie zeigen, dass 2D Gesten das Anfassen und Manipulieren von grafischen Objekten ermöglichen. Berührungen sind die einfachste Form natürlicher Interaktion. Gesten, die mit einem Finger und wenig Bewegung des Armes und der Hand ausgeführt werden, werden als leicht und intuitiv angesehen. 2D Gesten der Dimension Form, die statische Posen mit einem Finger zeigen, oder einen Pfad abfahren, werden hier genannt. Metaphorische und physikalische, als auch objekt-zentrierte und umgebungsabhängige Gesten der Dimensionen Natur und Bindung gehören dazu. Je komplexer eine 2D Geste in der Ausführung und Bewegung ist, desto schwieriger wird diese empfunden. Die physikalische Manipulation von zweidimensionalen Masken und Objekten, wie die Änderung von Größe und Positionen, wird bevorzugt mit einer Hand ausgeführt. 2D Gesten, welche zwei Hände zur Ausführung benötigen, werden als schwieriger eingestuft. Durch komplexere Aktionen entstehen komplexere Metaphern für die 2D Gesten. Die geringere Akzeptanz einiger 2D Gesten (für die Aktionen zwei, 14, 16, 24) könnte ein Indikator für eine nicht passende 2D Metapher für die jeweilige Aktion sein. Vor allem an der 2D Geste der Aktion 15 kann dies gesehen werden. Das Ändern der Kameraposition, wird eindeutig der 3D Geste zugeordnet. Dies hat den Hintergrund, dass für diese Aktion keine 2D Geste und Metapher Anwendung findet. Die Positionsänderung der Kamera in einer dreidimensionalen Szene ist mit einer 2D Geste umständlich. Durch eine nutzerorientierte Überarbeitung der entsprechenden schwierigeren 2D Gesten kann eine Verbesserung in der Akzeptanz eintreten. Leistungsgeminderte, als auch leistungsgewandelte Arbeiter profitieren von 2D Gesten. Grafische Objekte werden durch eine Berührung direkt mit der Aktion verknüpft. Bekannte und physisch nicht anstrengende 2D Gesten ermöglichen eine intuitive Bedienung. Diese Erkenntnisse bestätigen Untersuchungen, in welchen berührungsempfindliche Bildschirme als die beste Methode der Eingabe für Menschen mit motorischen Behinderungen festgestellt werden (vgl. [BL12]).

Dreidimensionale Gesten sind fehleranfällig und werden als unpräzise beschrieben. Die physische Manipulation von zweidimensionalen Masken und Objekten durch 3D Gesten erweist sich als schwierig. Die Selektion von Masken und Objekten ist durch ein Zielen und Zeigen auf ein Objekt physisch und kognitiv anstrengend. Bei grafischen zweidimensionalen Objekten auf einem Arbeitstisch wird tendenziell versucht die Objekte zu berühren, eine 3D Geste wird hierbei als unnötig empfunden. 3D Gesten für die Aktionen acht, 16, 22, 24 und 26 erreichen eine höhere Akzeptanz. Hier passt die 3D Metapher zur 2D Aktion. Zu bemerken ist, dass diese 3D Gesten dynamisch, symbolisch, physikalisch, pantomimisch, objekt-zentriert und einhändig sind. Ebenfalls enthalten diese 3D Gesten größtenteils Armbewegungen, wobei die

ROM Werte teils gering und teils größer sind. Dies lässt den Schluss zu, dass besonders 3D Gesten, in welchen Bewegungen des Armes vorkommen, als natürlicher angesehen werden. Dies könnte mit der Natur der Gestik begründet sein. Dynamische Gesten, die Armbewegungen und nicht nur Handzeichen enthalten, werden meist unterbewusst ausgeführt und sind intuitiv. An Aktion 15 kann gesehen werden, dass das Ändern der Kameraposition in einer dreidimensionalen Szene durch eine 3D Geste als passend empfunden wird. Die 3D Metapher der 3D Geste beschreibt die Aktion treffend. Die Beschreibung zweidimensionaler Aktionen durch 3D Metaphern dreidimensionaler Gesten erweist sich jedoch als problematisch. Physisch setzen 3D Gesten neben größerer Beweglichkeit von Arm und Hand auch feinmotorische Fähigkeiten voraus. Leistungsgeminderte und leistungsgewandelte Arbeiter können diese Voraussetzungen nicht immer erfüllen.

Hybrid Gesten sind eine Erweiterung von 3D Gesten um eine 2D Selektion. Die Selektion von Masken und Objekten ist intuitiv und vermeidet Fehleingaben. Eine physische Manipulation von zweidimensionalen Masken und Objekten durch Hybrid Gesten erweist sich ebenfalls als schwierig. Für den 3D Gestenanteil einer Hybrid Geste gelten die gleichen Schlussfolgerungen wie für 3D Gesten. Hybrid Gesten werden mit zwei Händen ausgeführt, da eine 2D Selektion mit einer Hand und eine 3D Geste mit der anderen Hand getätigt werden. Reine 3D Gesten sind laut Definition einer Hybrid Gesten nicht vorgesehen. Die Aktion zwei wird mit Mehrheit durch eine Hybrid Geste ausgeführt, da der 3D Gestenanteil mit einer Hand getätigt werden kann. Dies hat eine geringere physische Belastung zur Folge. Auch die Aktion 14 wird bevorzugt mit einer Hybrid Geste ausgeführt, da die Selektion einer Rotationsachse möglich wäre. Es werden bevorzugt dynamische, physikalische, pantomimische, objekt-zentrierte und kontinuierliche Hybrid Gesten zur Ausführung von Aktionen eingesetzt. Gegenüber leistungsgeminderten und leistungsgewandelten Arbeitern erweist sich eine einfache Selektion in Hybrid Gesten als Vorteil. Die klare Zuordnung des 3D Gestenanteils einer Hybrid Geste zu einem Objekt ist vereinfachend.

Die Gesteninteraktion ist eng mit dem Design der grafischen Benutzeroberfläche des Assistenzsystems verbunden. Die grafische Benutzeroberfläche zeigt die Auswirkungen ausgeführter Aktionen am System. Die Frage nach dem Design der Benutzeroberfläche muss die Aktionen und die Metaphern der Gesteninteraktion einschließen und berücksichtigen. Werden für zweidimensionale Benutzeroberflächen Metaphern dreidimensionaler Gesten gefunden, so kann eine 3D Gesteninteraktion intuitiv gestaltet werden. Ist dies nicht möglich, ergibt sich eine Beschreibung zweidimensionaler Aktionen durch unpassende Metaphern dreidimensionaler Gesten. Diese Beschreibung fühlt sich für den Benutzer nicht harmonisch, natürlich oder intuitiv an. Um Metaphern dreidimensionaler Gesten für zweidimensionale Benutzeroberflächen zu nutzen, muss die Z-Achse einer 3D Geste sinnvoll mit einer Eigenschaft eines 2D Objektes oder einer Aktion verknüpft werden. Zweidimensionale Benutzeroberflächen müssen in die dritte Dimension erweitert werden, Masken und Objekte könnten zum Beispiel dreidimensional gedreht werden. Jede Interaktionsmetapher beginnt mit einer physischen Manipulation und muss weiter ausgebaut werden. Eine zweidimensionale Benutzeroberfläche mit 2D Gesteninteraktion ermöglicht eine natürliche Bedienung. 3D und Hybrid Gesten bieten jedoch Vorteile und können eine 2D Gesteninteraktion unterstützen. Folgend kann der Einsatz mehrerer Gestenarten als vorteilhaft angesehen werden. Zu beachten ist, dass fensterbasierte grafische Benutzeroberflächen eine Hierarchie aufweisen. Die intuitive zwei-

oder dreidimensionale Interaktion mit Masken und Objekten innerhalb einer Hierarchie ist jedoch schwer zu realisieren. Das Design der grafischen Benutzeroberfläche sollte dies beachten und passend gewählt werden. Die grafische Benutzeroberfläche eines Assistenzsystems sollte nach Natural User Interface Standards entwickelt werden. Auch die Möglichkeiten einer Texteingabe innerhalb einer zweidimensionalen grafischen Benutzeroberfläche sollte gegeben sein.

Je nach Design der grafischen Benutzeroberfläche können Hybrid Gesten zur Optionsausführung eingesetzt werden. Es ergibt sich ein Mehrwert bei Optionsauswahl. Können in einer Maske, an einem Objekt oder in einem Menü mehrere Aktionen durchgeführt werden, so kann mit einem 2D Tap eine übergeordnete Aktion ausgewählt und mit einer 3D Geste eine Unteraktion ausgeführt werden. Die grafische Benutzeroberfläche könnte eine Art Kreismenü bei einem 2D Tap öffnen, anstatt der 2D Selektion weiterer Untermenüs und Aktionen könnten 3D Gesten Aktionen direkt auslösen. Weiterhin könnten auch 2D Gesten eingesetzt werden. Dieses Design einer Benutzeroberfläche und Gesteninteraktion würde Platz auf der Arbeitsfläche sparen, durch weniger Masken und Objekte. Erfahrene Nutzer, der Trainer oder Vorarbeiter können 3D Gesten als Schnellzugriff nutzen. Ebenfalls können 3D Gesten bei verschmutzten Händen oder getragenen Handschuhen ausgeführt werden. Die berührungsempfindliche Arbeitsfläche würde nur durch einen 2D Tap mit der verschmutzten Hand in Kontakt kommen.

Ein weiteres Interaktionskonzept könnte den umgekehrten Zusammenschluß von 2D und 3D Geste in der Hybrid Geste darstellen. Eine symbolische 3D Geste könnte zum Beispiel ein Menü öffnen, folgend könnte eine 2D Geste eine Auswahl in diesem Menü treffen. Ebenfalls könnte eine 2D Geste als Schnellzugriff für die Durchführung von Aktionen genutzt werden.

Die Ergebnisse der durchgeführten Studie zeigen Vor- und Nachteile der Gesteninteraktion durch verschiedene Gestenarten. Eine Kombination von 2D und Hybrid Gesten ist der richtige Weg zu einer natürlichen, intuitiven und einfachen Bedienung eines Assistenzsystems durch leistungsgeminderte und leistungsgewandelte Arbeiter.

6.2.1. Design Richtlinien für Gesteneingaben

Ziel dieser Richtlinien ist ein effektives Design einer Benutzerschnittstelle zur Interaktion mit einem Assistenzsystem durch Gesten. Diese Gesten sollen durch ihren Angebotscharakter (engl. „*affordance*“) und passendem Feedback zur wiederholten Ausführung anregen und ein schnelles Erlernen ermöglichen. Der Angebotscharakter eines Gegenstands ist die offensichtlich vorhandene oder tatsächlich gegebene angebotene Gebrauchseigenschaft (vgl. [WW10], S. 30). Die Heranführung von Benutzern an ein Gesten-Interaktionssystem kann durch eine Vorführung von Gesten erfolgen. Als Hilfe können Bilder pantomimischer Aktionen im Arbeitsablauf gezeigt werden (vgl. [GJM11a]).

Aufgrund zweier Studien, einmal zur Identifikation relevanter Gesten und einmal der Evaluierung dieser Gesten, können folgende Design Richtlinien für Assistenzsysteme im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter aufgestellt werden.

Design Richtlinien für zweidimensionale Gesten

Form und Durchführung

- Wisch-, Strich- und Tipp-Gesten sind zu bevorzugen (Gesten der Dimension Form und der Eigenschaft „ein-Punkt Pfad/Berührung“).
- Gesten sollten objekt-zentriert sein (Gesten der Dimension Bindung und der Eigenschaft „objekt-zentriert“).
- Einhändige Gesten sind zu bevorzugen. Aufgrund der Größe von Masken und Objekten kann die Manipulation der Größe mit zwei Händen ausgeführt werden. Beidhändige Gesten sollten jedoch im allgemeinen vermieden werden.
- Eine Gestenunterscheidung anhand der Fingeranzahl sollte vermieden werden. 2D Tap Gesten können mit bis zu drei Fingern getätigt werden, eine Grenze sind vier Finger.
- Gesten sollen in ihrer inversen Form ausführbar sein und die gegenläufige Aktion auslösen.
- Die Geschwindigkeit der Gestenausführung sollte keine Rolle spielen, eine langsame Ausführung einer Gesten sollte möglich sein.
- Gesten sollten eine Toleranz gegenüber längeren Berührungszeiten (Touch-Events) haben. Ein Finger kann länger auf einer Position verharren, ohne mehrere Male auszulösen.
- Kontinuierliche Gesten wie das „scrolling“ sollten in diskreten Schritten eine Reaktion auslösen, dieses Vorgehen erlaubt geringere feinmotorische Fähigkeiten des Benutzers bei der Ausführung.
- Die Erkennung von Gesten sollte trotz nicht geballter Faust möglich sein. An einer Geste nicht beteiligte und ungenutzte Finger dürfen nicht stören.
- Gezeichnete Gesten sollten bekannte Symbole und weniger geometrisch komplexe Formen (weniger Striche und Windungen) enthalten. Gesten in einem Strich werden bevorzugt.

Inhaltlich

- Metaphorische Gesten sind für Aktionen zu wählen (Gesten der Dimension Natur und der Eigenschaft „metaphorisch“).
- Gesten sollten zu eindeutigen mentalen Modellen zugeordnet sein. Klare und eindeutige Metaphern beschreiben Aktionen. Einfache und bekannte Gesten sollten für allgemein bekannte Aktionen verwendet werden.
- Gesten sollten zur besseren Lern- und Merkbarkeit in ähnlichen Kontexten und Aktionen wiederholt werden.

Umgebung

- Menüs, Masken und Objekte grafischer Benutzeroberflächen sollten in der Größe angepasst sein. Größere Formen erhöhen die Auswahlgenauigkeit.
- Größere Projektionsflächen und zentrierte Objekte erhöhen die Ausführungsgenauigkeit. Die Bewegung der Finger erhält mehr Freiraum, somit wird mehr Fläche für eine Geste bereitgestellt. Die Auswahl von zentrierten Objekten fällt leichter.
- Griffpunkte für Gesten an Masken und Objekten sollten vorhanden und leicht hervorgehoben sein. Besonders liegen Ecken von Masken und Objekten für eine Manipulation im Fokus.
- Zur Gesteneingabe soll ein direktes und schnelles Feedback erfolgen. Dies kann durch eine Hervorhebung von selektierten Masken und Objekten geschehen. Auch falsche Eingaben sollen ein Feedback geben.

Design Richtlinien für dreidimensionale und Hybrid Gesten

Form und Durchführung

- Dynamische Gesten sollten eingesetzt werden (Gesten der Dimension Form und der Eigenschaft „dynamisch“).
- Pantomimische Gesten sollten Aktionen beschreiben (Gesten der Dimension Natur und der Eigenschaft „pantomimisch“). Gesten, in welchen imaginäre Objekte gehalten und pantomimisch genutzt werden, sind intuitiv. Gesten, in denen Werkzeuge gezeigt werden, sollten diese Werkzeuge durch Körperteile darstellen.
- Gesten sollten objekt-zentriert und umgebungs-unabhängig sein (Gesten der Bindung Natur und den Eigenschaften „objekt-zentriert“, „umgebungs-unabhängig“).
- Gesten sollten möglichst einfach und nicht zusammengesetzt sein (Gesten der Dimension Komplexität und der Eigenschaft „einfach“).
- Gesten sollten die Bewegung des Armes und weniger der Hand beinhalten (Gesten der Dimension Körperteil und der Eigenschaft „Arm“). Handbewegungen erfordern feinmotorische Fähigkeiten, an einer statisch gehaltenen Hand können Zitterbewegungen berücksichtigt werden.
- Gesten mit hoher physischer Bewegung und maximalen Gelenkwinkelstellungen sollen vermieden werden (Gesten der Dimension ROM und der Eigenschaft „groß“).
- Einhändig ausgeführte Gesten sollten bevorzugt werden. Aufgrund der Größe von Masken und Objekten kann die Manipulation der Größe mit zwei Händen sinnvoll sein.
- Eine Gestenunterscheidung anhand der Fingeranzahl sollte vermieden werden.
- Gesten sollen in ihrer inversen Form ausführbar sein und die gegenläufige Aktion auslösen.

6. Evaluation

- Die Geschwindigkeit der Gestenausführung sollte keine Rolle spielen. Eine langsame Ausführung einer Geste sollte möglich sein.
- Kontinuierliche Gesten wie die Rotation sollten in diskreten Schritten eine Reaktion auslösen, dieses Vorgehen erlaubt geringere feinmotorische Fähigkeiten des Benutzers bei der Ausführung.
- Die Erkennung von Gesten sollte trotz nicht vollständig geballter Faust möglich sein. An einer Geste nicht beteiligte und ungenutzte Finger dürfen nicht stören.
- Gezeichnete Gesten sollten bekannte Symbole und weniger geometrisch komplexe Formen (weniger Striche und Windungen) enthalten. Gesten in einem Strich werden bevorzugt.

Inhaltlich

- Kulturelle Unterschiede in der allgemeinen Bedeutung von Gesten sollten beachtet werden. Dies gilt vor allem für symbolische Gesten.
- Gesten sollten zur besseren Lern- und Merkbarkeit in ähnlichen Kontexten und Aktionen wiederholt werden.
- Gesten sollten Aktionen durch klare und eindeutige Metaphern beschreiben. Die Z-Achse einer 3D Geste muss sinnvoll mit einer Eigenschaft eines 2D Objektes oder einer Aktion verknüpft werden. Zweidimensionale Benutzeroberflächen müssen in die dritte Dimension erweitert werden. Einfache und bekannte Gesten sollten für allgemein bekannte Aktionen verwendet werden.

Umgebung

- Menüs, Masken und Objekte grafischer Benutzeroberflächen sollten in der Größe angepasst sein. Größere Formen erhöhen die Auswahlgenauigkeit. Dies trägt zur Robustheit der Eingabe bei.
- Zur Gesteneingabe soll ein direktes und schnelles Feedback erfolgen. Dies kann durch eine Hervorhebung von selektierten Masken und Objekten geschehen. Auch falsche Eingaben sollen ein Feedback geben.
- Griffpunkte für Hybrid Gesten an Masken und Objekten sollten vorhanden und leicht hervorgehoben sein. Besonders liegen Eck- und Achsenpunkte von Masken und Objekten für eine Manipulation im Fokus.

7. Zusammenfassung und Ausblick

Diese Diplomarbeit untersucht zwei- und dreidimensionale Gesteninteraktion an Assistenzsystemen im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter. Das letzte Kapitel fasst die theoretischen und praktischen Erkenntnisse zusammen, die während der Arbeit gesammelt worden sind, und gibt einen Ausblick für zukünftige Untersuchungen im Bezug auf die Problemstellung dieser Arbeit. Im Folgenden wird eine Zusammenfassung der bisherigen Kapitel gegeben. Hierbei wird unter anderem auf die gewonnenen Ergebnisse mit Blick auf die Problemstellung eingegangen. Es werden die Auswirkungen der Ergebnisse auf Assistenzsysteme herausgearbeitet und die Vor- und Nachteile des Einsatzes von Gesten zusammengefasst. Abschließend wird ein Ausblick über die möglichen Potentiale der gefundenen Erkenntnisse für die zukünftige Forschung und Anwendung im Bereich der Assistenzsysteme im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter gegeben.

7.1. Zusammenfassung der Arbeit

Unsere Gesellschaft wird immer älter, der demografische Wandel wirkt sich nicht nur auf den gesellschaftlichen Zusammenhalt, sondern auch auf die Wirtschaft aus. Die zunehmende Alterung der Bevölkerung trifft auch Arbeitnehmer. Bis zum Jahr 2025 dürfte jeder vierte Erwerbstätige älter als 55 Jahre sein. Somit müsse es auch im Alter normal werden zu arbeiten und sich weiterzubilden. Die Produktivitätsreserven einer Gesellschaft müssen mobilisiert werden, um dieser Entwicklung entgegenzutreten. Nicht nur bei leistungsgewandelten, sondern auch bei leistungsgeminderten Menschen liegen Potentiale. Assistenzsysteme am Arbeitsplatz sollen leistungsgeminderte und leistungsgewandelte Arbeiter unterstützen und zum selbstständigen und wirtschaftlichen Arbeiten befähigen. Dabei stellt die Benutzerinteraktion mit Assistenzsystemen einen Kernpunkt für die Akzeptanz dieser Systeme dar.

Mit Hilfe zweier Studien werden zwei- und dreidimensionale Gesten für Assistenzsysteme im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter untersucht, entwickelt und evaluiert. Zuvor werden vorhandene Interaktionsmöglichkeiten untersucht und mögliche Aktionen der Arbeiter am Assistenzsystem identifiziert. Mit Hilfe der ersten Studie und der Untersuchung der Aktionsmöglichkeiten leistungsgeminderter und leistungsgewandelter Arbeiter werden zwei- und dreidimensionale Gesten entwickelt, wobei sich eine dritte Gestenform als interessante Ergänzung zu dreidimensionalen Gesten erweist. Diese Gestenform besteht aus Hybrid Gesten, die einen Zusammenschluss von 2D und 3D Gesten darstellen. In der zweiten Studie werden die Gestenformen und die entwickelten Gesten am System evaluiert. Die Frage nach den Vor- und Nachteilen des Einsatzes einer Gesteninteraktion am

Assistenzsystem wird beantwortet. Abschließend werden Design Richtlinien für die Gestenentwicklung am Assistenzsystem im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter aufgestellt.

7.1.1. Interpretation der Resultate

Die Evaluation der entwickelten Gesten und Gestenformen zeigt eine eindeutige Präferenz zweidimensionaler Gesten. 2D Gesten werden als bekannt und einfach angesehen. Die Fehleranfälligkeit bei Nutzung durch leistungsgeminderte und leistungsgewandelte Arbeiter ist unter Beachtung aufgestellter Design Richtlinien für diese Gesten gering. Auch die Merkbarkeit zweidimensionaler Gesten ist aufgrund der Einfachheit dieser höher. Dreidimensionale Gesten sind im Kontext leistungsgeminderter und leistungsgewandelter Arbeiter fehleranfällig und setzen oft feinmotorische und koordinative Fähigkeiten voraus. Leistungsgeminderte und leistungsgewandelte Arbeiter können diesen Voraussetzungen nicht immer gerecht werden. Hybrid Gesten erweitern dreidimensionale Gesten um eine intuitive Selektionsmöglichkeit, das einfache antippen von Objekten reicht für eine Selektion aus. Hybrid Gesten sind robuster im Vergleich zu dreidimensionalen Gesten. Eine Kombination von 2D und Hybrid Gesten ist der richtige Weg zu einer natürlichen, intuitiven und einfachen Bedienung eines Assistenzsystems durch leistungsgeminderte und leistungsgewandelte Arbeiter.

7.1.2. Auswirkungen der Ergebnisse

Die Kombination von zweidimensionalen und hybriden Gesten vereint die Vorteile beider Gestenformen und macht die Gesteninteraktion für leistungsgeminderte, leistungsgewandelte, als auch uneingeschränkte Arbeiter einfach und intuitiv. Erfahrenen Benutzer haben die Möglichkeit schnell und effizient Aufgaben am Assistenzsystem zu entwickeln und auszuführen. Arbeiter können direkt über die Arbeitsfläche mit dem System kommunizieren und müssen keine externen Eingabegeräte und Bildschirme beachten. Der Arbeitsfluss wird durch eine Gesteninteraktion nicht unterbrochen, ein Schnellaufgriff auf bestimmte Funktionen des Systems kann durch Gesten ermöglicht werden. Eine natürliche Bedienung wirkt sich positiv auf die Leistungsfähigkeit leistungsgeminderter und leistungsgewandelter Arbeiter aus. Die Akzeptanz für realitätserweiternde assistierende Systeme am Arbeitsplatz wird durch eine intuitive Bedienung bestärkt.

Aufgestellte Design Richtlinien für die Entwicklung von Gesten am Assistenzsystem im Hinblick auf die Bedürfnisse und Anforderungen leistungsgeminderter und leistungsgewandelter Arbeiter ermöglichen die Konzeption intuitiver und einfacher Gesten-Sets.

7.2. Ausblick

Die Entwicklung von intuitiven Interaktionsformen ist komplex. Die Realisierung einer Interaktion mit einem System durch Gesten kann konzeptionell und technisch sehr anspruchs-

voll werden. Das Einsatzgebiet eines Systems, welches Gesteninteraktion ermöglichen soll, muss analysiert werden. Die Benutzergruppen, welche mit dem System interagieren, müssen ebenfalls studiert werden. Die Aktionen am System müssen verstanden werden, bevor eine Eingabe durch Gesten konzipiert werden kann. Die Gestik eines Menschen beinhaltet verschiedene Facetten und unterliegt mit der Sprache des Menschen dem gleichen integrierten System. Taxonomien für Gesten versuchen Gesteneigenschaften zu beschreiben. Einheitliche Taxonomien für zwei- und dreidimensionale Gesten im Kontext leistungsgeminderter und leistungsgewandelter Menschen existieren nicht. Zwei- und dreidimensionale Gesten für leistungsgeminderte und leistungsgewandelte Arbeiter finden bis heute keine Anwendung in Assistenzsystemen am Arbeitsplatz. Diese Arbeit präsentiert Taxonomien für zwei- und dreidimensionale Gesten für leistungsgeminderte und leistungsgewandelte Arbeiter und stellt Design Richtlinien für die Entwicklung von Gesten vor.

Es liegt nahe durch weitere Benutzerstudien mit leistungsgeminderten und leistungsgewandelten Arbeitern weitere Messwerte von Hand- und Armbewegungen zu erfassen und die bestehenden Taxonomien durch Zuhilfenahme dieser Messwerte zu verfeinern. Durch das RITE Modell (engl. *rapid iterative testing and evaluation*) kann ein Interaktionsdesign iterativ entworfen und evaluiert werden. Die aufgestellten Gesten-Sets und entwickelten Gesten-Erkennungsmethoden können mit Hilfe verfeinerter Taxonomien verbessert werden. Weitere Benutzerstudien zur erneuten Evaluierung der verbesserten 2D und Hybrid (3D) Gesten können weitere Erkenntnisse zur Verfeinerung der Gesten-Sets liefern. Treffende und unverwechselbare Metaphern können so für Aktionen iterativ gefunden werden. Abschließend sollten die aufgestellten Design Richtlinien für die Entwicklung von Gesten wiederholt geprüft und neu bewertet werden.

A. Anwendungsfälle

In diesem Anhang werden Anwendungsfälle, Use cases des Trainers und des Arbeiters am motionEAP Assistenzsystem gezeigt.

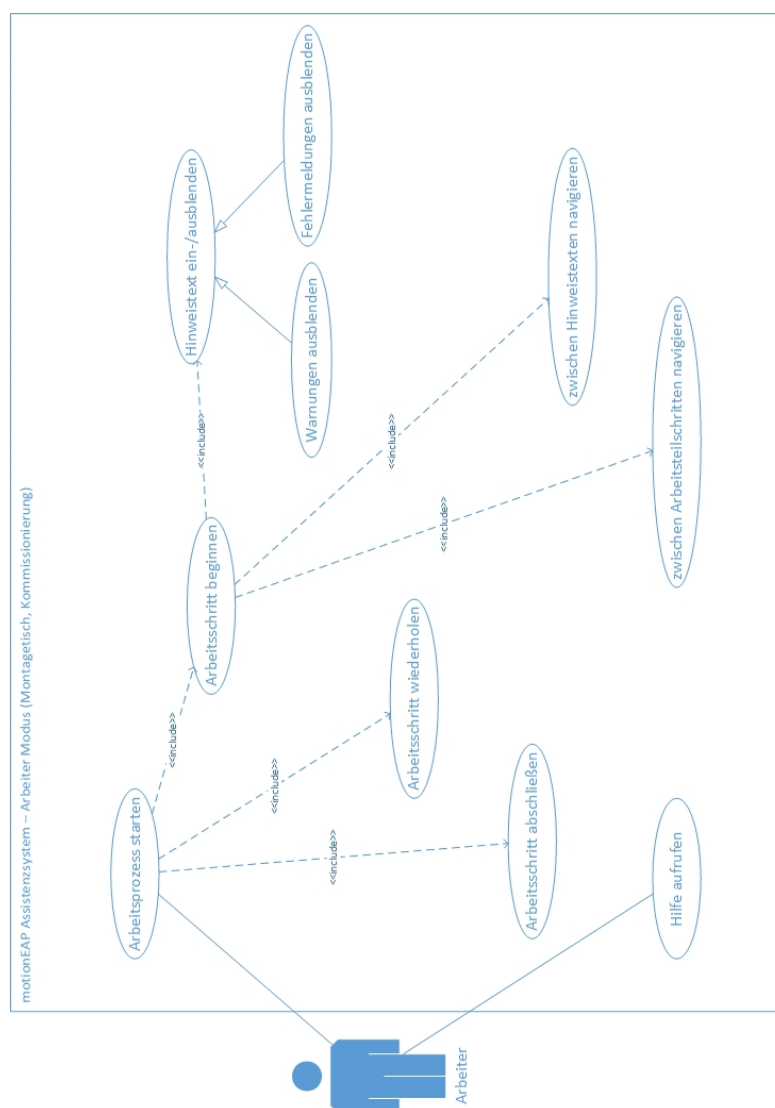


Abbildung A.1.: Use case, Arbeiter

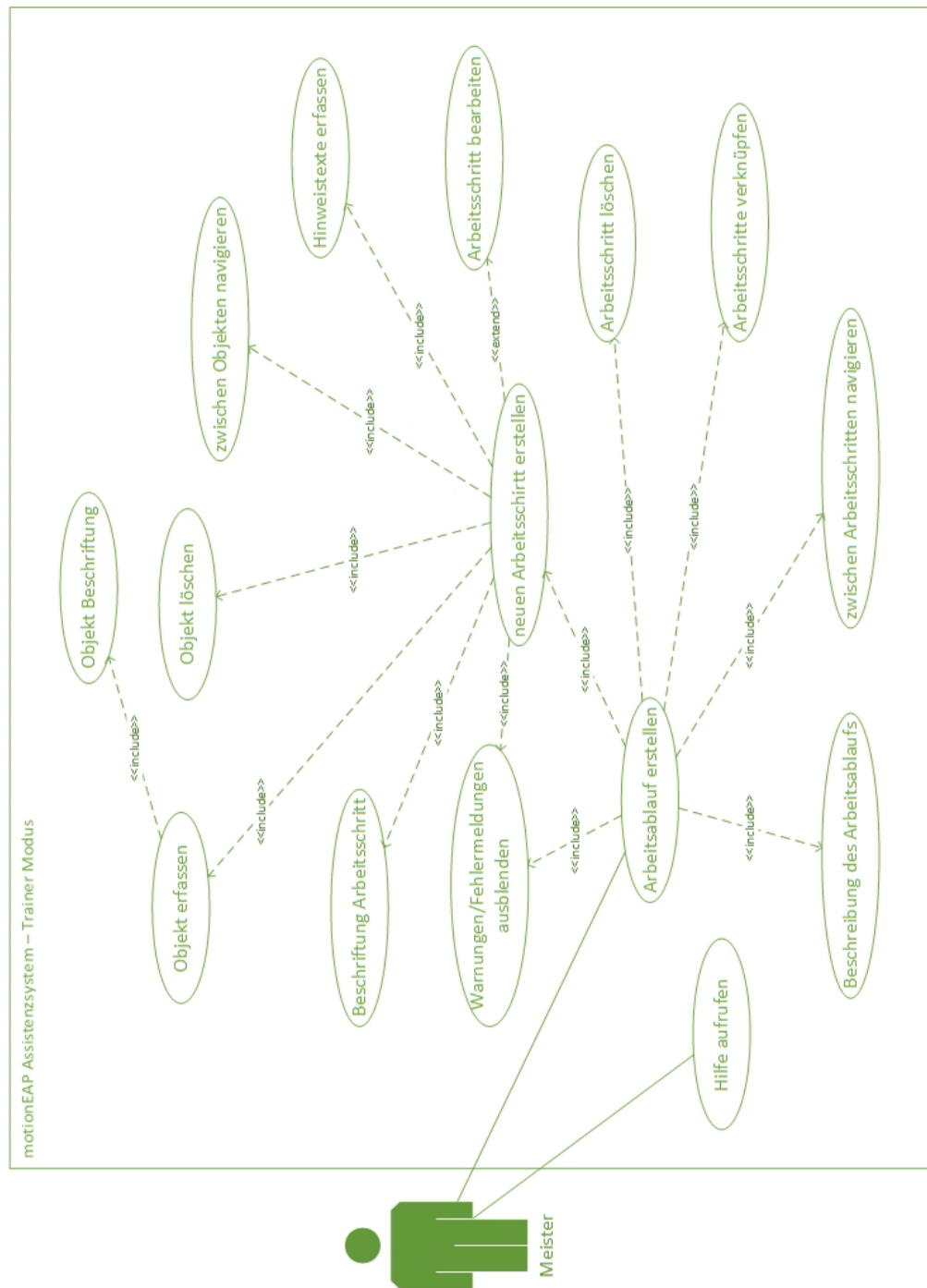


Abbildung A.2.: Use case, Trainer

Literaturverzeichnis

- [AKF13] L. Anthony, Y. Kim, L. Findlater. Analyzing user-generated youtube videos to understand touchscreen use by people with motor impairments. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI'13*, S. 1223–1232, 2013. (Zitiert auf Seite 57)
- [Aly13] M. Aly. *Development and Evaluation of an Assistive System using Hand Recognition and Motivating Elements*. Diplomarbeit, University of Stuttgart, Germany, Institut of Visualization and Interactive Systems, 2013. (Zitiert auf den Seiten 35, 57, 58 und 80)
- [AWB⁺12] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, V. Koh. Understanding Mid-Air Hand Gestures: A Study of Human Preferences in Usage of Gesture Types for HCI. Technischer Bericht, Microsoft Research, 2012. (Zitiert auf den Seiten 34 und 59)
- [BK12] M. Brach, O. Korn. Assistive Technologies at Home and in the Workplace - A Field of Research for Exercise Science and Human Movement Science. *European Review of Aging and Physical Activity*, 9:1–4, 2012. (Zitiert auf den Seiten 13, 14, 16 und 17)
- [BKE⁺13] E. Brea, P. Kambouris, A. Elfes, E. Duff, M. Bick, A. Bonchis, A. Tews, L. Lopes. Lightweight Assistive Manufacturing Solutions - Improving Australia's Manufacturing Competitiveness. White paper. CSIRO, the Commonwealth Scientific and Industrial Research Organisation, Australia. *Future Manufacturing Flagship, Agile Manufacturing Technologies*, 2013. (Zitiert auf den Seiten 14 und 15)
- [BL12] P. Biswas, P. Langdon. Developing Multimodal Adaptation Algorithm for Mobility Impaired Users by Evaluating Their Hand Strength. *International Journal of Human-Computer Interaction*, 28:576–596, 2012. (Zitiert auf den Seiten 57, 58 und 89)
- [BPHH11] J. Bierkandt, M. Peissner, C. Hipp, F. Hermann. *Usability und human-machine interfaces in der Produktion - Studie Qualitätsmerkmale für Entwicklungswerkzeuge*. Fraunhofer Verlag, 2011. (Zitiert auf Seite 22)
- [CRDR11] A. Chaudhar, J. L. Raheja, K. Das, S. Raheja. Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way: A Survey. *International Journal of Computer Science & Engineering Survey*, 2:122–133, 2011. (Zitiert auf Seite 38)

- [DBB⁺12] J. Dul, R. Bruder, P. Buckle, P. Carayon, P. Falzon, W. Marras, J. Wilson, B. van der Doelen. A strategy for human factors/ergonomics: developing the discipline and profession. *Ergonomics*, 55:377–395, 2012. (Zitiert auf Seite 16)
- [FH10] A. Field, G. Hole. *How to Design and Report Experiments*. SAGE Publications, 2010. (Zitiert auf den Seiten 54, 83 und 85)
- [FKB⁺13] M. Funk, O. Korn, H. Behrendt, M. Kölz, A. Bächler, L. Bächler, A. Schmidt, T. Hörz, C. Misselhorn, T. Heidenreich. motionEAP: An assistive system for the workplace. Unpublished, 2013. (Zitiert auf den Seiten 14, 15, 19, 41 und 43)
- [GJM11a] S. Grandhi, G. Joue, I. Mittelberg. Understanding Naturalness and Intuitiveness in Gesture Production: Insights for Touchless Gestural Interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI'11*, S. 821–824, 2011. (Zitiert auf den Seiten 61 und 91)
- [GJM11b] S. Grandhi, G. Joue, I. Mittelberg. What's in a Gesture: Exploring the potential of gesture-based interaction for low literacy user groups. *ACM 2011 Conference on Human Factors in Computing Systems (CHI), Austin, Texas.*, 2011. (Zitiert auf den Seiten 51 und 57)
- [HA12] J. Hardy, J. Alexander. Toolkit Support for Interactive Projected Displays. *MUM '12 Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, 2012. (Zitiert auf den Seiten 43 und 79)
- [Ilm05] J. Ilmarinen. *Towards a longer worklife - Ageing and the quality of worklife in the European Union*. Finnish Institute of Occupational Health, Ministry of Social Affairs and Health, 2005. (Zitiert auf Seite 16)
- [KAW⁺09] S. Kane, D. Avrahami, J. O. Wobbrock, B. Harrison, A. D. Rea, M. Philipose, A. LaMarca. Bonfire: A Nomadic System for Hybrid Laptop-Tabletop Interaction. *UIST '09, Proceedings of the 22nd annual ACM symposium on User interface software and technology*, NewYork, S. 129–138, 2009. (Zitiert auf Seite 23)
- [KBS⁺12] O. Korn, M. Brach, A. Schmidt, T. Hörz, R. Konrad. Context-Sensitive User-Centered Scalability: An Introduction Focusing on Exergames and Assistive Systems in Work Contexts. *E-Learning and Games for Training Education, Health and Sports*. S. Göbel, W. Müller, B. Urban, and J. Wiemeyer, eds. Springer Berlin Heidelberg, 7516:164–176, 2012. (Zitiert auf den Seiten 12, 14, 16, 17, 24 und 25)
- [KEP96] B. Kolster, G. Ebel-Paprotny. *Leitfaden Physiotherapie. Befund, Techniken, Behandlung, Rehabilitation*. Urban & Fischer Verlag, 1996. (Zitiert auf den Seiten 60 und 61)
- [KH90] G. Kurtenbach, E. Hulteen. *Gestures in Human-Computer Communication*. In: *The Art of Human-Computer Interface Design*. Addison-Wesley Publishing Co., 1990. (Zitiert auf Seite 29)

- [KMR08] S. D. Kelly, S. M. Manning, S. Rodak. Gesture Gives a Hand to Language and Learning: Perspectives from Cognitive Neuroscience, Developmental Psychology and Education. *Language and Linguistics Compass*, 2:569–588, 2008. (Zitiert auf den Seiten 31 und 57)
- [Kor12] O. Korn. Industrial Playgrounds: How Gamification Helps to Enrich Work for Elderly or Impaired Persons in Production. *EICS '12 Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*, S. 313–316, 2012. (Zitiert auf den Seiten 12, 16, 17, 18, 22 und 24)
- [KS05] M. Karam, M. C. Schraefel. A taxonomy of gestures in human computer interactions. Technischer Bericht, University of Southampton, 2005. (Zitiert auf Seite 32)
- [KSH12] O. Korn, A. Schmidt, T. Hörz. Assistive systems in production environments: exploring motion recognition and gamification. *PETRA '12 Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments*, S. 9:1–9:5, 2012. (Zitiert auf den Seiten 13, 14, 17 und 25)
- [KSH13a] O. Korn, A. Schmidt, T. Hörz. Augmented Manufacturing: A Study with Impaired Persons on Assistive Systems Using In-Situ Projection. *PETRA'13, Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments, New York*, 2013. (Zitiert auf den Seiten 22, 23 und 29)
- [KSH13b] O. Korn, A. Schmidt, T. Hörz. The Potentials of In-Situ-Projection for Augmented Workplaces in Production - A Study with Impaired Persons. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, S. 979–984, 2013. (Zitiert auf den Seiten 14, 17, 22, 23, 24 und 25)
- [MA07] S. Mitra, T. Acharya. Gesture Recognition: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 37:311 – 324, 2007. (Zitiert auf Seite 37)
- [McN92] D. McNeill. *Hand and mind: what gestures reveal about thought*. The University of Chicago Press, 1992. (Zitiert auf den Seiten 29, 30 und 31)
- [NPMP10] D. Neumann, R. Pahlen, M. Majerski-Pahlen. *Sozialgesetzbuch IX - Rehabilitation und Teilhabe behinderter Menschen*. C.H.Beck, 2010. (Zitiert auf den Seiten 12 und 16)
- [OYN08] S. K. Ong, M. L. Yuan, A. Y. C. Nee. Augmented reality applications in manufacturing: a survey. *International Journal of Production Research*, 46:10:2707–2742, 2008. (Zitiert auf Seite 22)
- [Pin01] C. Pinhanez. The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. *Proceedings of the 3rd international conference on Ubiquitous Computing, London UK*, 2201:315–331, 2001. (Zitiert auf Seite 23)
- [RCS08] J. Rubin, D. Chisnell, J. Spool. *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*. Wiley Publishing, 2008. (Zitiert auf Seite 83)

- [RKWA10] S. Rosenthal, S. K. Kane, J. O. Wobbrock, D. Avrahami. Augmenting on-screen instructions with micro-projected guides: when it works, and when it fails. *Ubicomp '10 Proceedings of the 12th ACM international conference on Ubiquitous computing*, S. 203–212, 2010. (Zitiert auf den Seiten 14 und 23)
- [RLL11] J. Ruiz, Y. Li, E. Lank. User-defined motion gestures for mobile interaction. *CHI '11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, S. 197–206, 2011. (Zitiert auf den Seiten 33, 35 und 59)
- [Saf09] D. Saffer. *Designing Gestural Interfaces*. O'Reilly Media, 2009. (Zitiert auf Seite 34)
- [SBKK13] S. Sievert, U. Berger, S. Kröhnert, R. Klingholz. *Produktiv im Alter - Was Politik und Unternehmen von anderen europäischen Ländern lernen können*. Berlin-Institut für Bevölkerung und Entwicklung, 2013. (Zitiert auf Seite 12)
- [SPH10] A. L. Sauer, A. Parks, P. C. Heyn. Assistive technology effects on the employment outcomes for people with cognitive disabilities - a systematic review. *Informa Healthcare, Disability and Rehabilitation: Assistive Technology*, 5:377–391, 2010. (Zitiert auf den Seiten 13 und 22)
- [WMW09] J. O. Wobbrock, M. R. Morris, A. D. Wilson. User-Defined Gestures for Surface Computing. *CHI '09 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, S. 1083–1092, 2009. (Zitiert auf den Seiten 32, 35, 51, 59, 60, 63 und 83)
- [WRM05] H. H. Wobbrock, J. O. and Aung, B. Rothrock, B. A. Myers. Maximizing the Guessability of Symbolic Input. *Proceeding CHI'05 Extended Abstracts on Human Factors in Computing Systems*, S. 1869–1872, 2005. (Zitiert auf Seite 51)
- [WVA12] J. O. Wobbrock, R. D. Vatavu, L. Anthony. Gestures as Point Clouds: A \$P Recognizer for User Interface Prototypes. *ICMI '12 Proceedings of the 14th ACM international conference on Multimodal interaction*, S. 273–280, 2012. (Zitiert auf den Seiten 37, 39, 78, 80 und 81)
- [WW10] D. Wigdor, D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan Kaufmann Publishers, 2010. (Zitiert auf den Seiten 34, 35, 51 und 91)

Alle URLs wurden zuletzt am 30.01.2014 geprüft.

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben. Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommene Aussagen als solche gekennzeichnet. Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens. Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht. Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

Neuhausen a. d. F., den 30. Januar 2014

Alexander Martin