



Hochschule für Angewandte Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*

# Bachelorarbeit

Krystian Graczyk

**Eignungsuntersuchung von 2D und 3D Gesten für Kollaboration**

*Fakultät Technik und Informatik  
Studiendepartment Informatik*

*Faculty of Engineering and Computer  
Science  
Department of Computer Science*

Krystian Graczyk

**Eignungsuntersuchung von 2D und 3D Gesten für  
Kollaboration**

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung

im Studiengang Bachelor of Science Angewandte Informatik  
am Department Informatik  
der Fakultät Technik und Informatik  
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Prof. Dr. Kai von Luck  
Zweitgutachter: M.Sc. Tobias Eichler

Eingereicht am: 07. November 2016

**Krystian Graczyk**

**Thema der Arbeit**

Eignungsuntersuchung von 2D und 3D Gesten für Kollaboration

**Stichworte**

2D-Interaktion, 3D-Interaktion, Gestensteuerung, kollaborativer Bildschirm, CSCW

**Kurzzusammenfassung**

Die Arbeitsmethoden und Produktionsprozesse in der Wirtschaft werden immer komplexer. Dies führt dazu, dass auch Teamarbeit stetig relevanter wird, um das Unternehmensziel effektiv verfolgen zu können. Ziel dieser Arbeit ist es das Zusammenspiel von 2D und 3D Gesten, im Hinblick darauf zu evaluieren, inwiefern sie es einem Team ermöglichen kollaborativ eine Aufgabe an einem Anzeigemedium zu bewältigen. Nach durchgeführter Eignungsuntersuchung stellt sich heraus, dass beide Steuerungstypen ihre Stärken und Schwächen haben und die Kombination beider Varianten die jeweiligen Stärken verbindet. Doch auch die Schwächen, wie Ungenauigkeit, müde werdende Arme und mangelnde Übung im Umgang mit 3D Gesten bleiben bestehen. Dennoch zeigte sich die Kombination beider Steuerungstypen als vielversprechend für die Arbeit in Gruppen.

**Krystian Graczyk**

**Title of the paper**

Eligability test of 2D and 3D gestures for collaboration

**Keywords**

2D interaction, 3D interaction, gesture control, collaborative display, group work, CSCW

**Abstract**

Work methods and production processes are getting more complex in the economy. As a result group work gets continuously more important to effectively pursue the business goal. This paper evaluates, if and how well 2D and 3D gestures are able to enable a team to collaboratively resolve a task at a single display. The eligibility tests showed that 2D as well as 3D gestures have their own strengths and weaknesses. The combination of both gesture controls merges those strengths. However, it has the same weaknesses like inaccuracy, tiring arms and the lack of practice with 3D gestures. Nevertheless the combination of both gesture types looked promising for the work in groups.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Motivation . . . . .	2
1.2	Themenbeschreibung . . . . .	3
1.3	Gliederung . . . . .	3
<b>2</b>	<b>Grundlagen</b>	<b>5</b>
2.1	Gruppe . . . . .	5
2.2	Human Computer Interaction . . . . .	6
2.3	Ubiquitous Computing . . . . .	8
2.4	Computer Supported Cooperative Work . . . . .	9
2.5	Mental Model . . . . .	11
2.6	Gesten . . . . .	14
2.6.1	Kulturelle Unterschiede . . . . .	15
2.6.2	Taxonomie . . . . .	16
2.6.3	Unintentional Gestures . . . . .	19
2.6.4	2D Gesten . . . . .	20
2.6.5	3D Gesten . . . . .	22
<b>3</b>	<b>Analyse</b>	<b>25</b>
3.1	Szenarien . . . . .	25
3.2	Anwendungsfälle . . . . .	27
3.3	Anforderungsanalyse . . . . .	29
3.3.1	funktionale Anforderungen . . . . .	30
3.3.2	nicht funktionale Anforderungen . . . . .	30
3.4	Laborumgebung . . . . .	33
3.5	Fazit . . . . .	36
<b>4</b>	<b>Entwurf</b>	<b>37</b>
4.1	2D Gesten . . . . .	37
4.2	3D Gesten . . . . .	39
4.2.1	Registrierung . . . . .	39
4.2.2	Orientierung auf dem Display . . . . .	40
4.2.3	Cursor bewegen . . . . .	43
4.2.4	Objektverschiebung . . . . .	43
4.3	Benutzeranzahl . . . . .	46

4.4	Benutzeroberfläche . . . . .	46
4.5	Architektur . . . . .	48
4.6	Fazit . . . . .	50
<b>5</b>	<b>Evaluation</b>	<b>52</b>
5.1	Implementierbarkeit . . . . .	52
5.2	Eignungsuntersuchung . . . . .	56
5.2.1	Aufbau . . . . .	56
5.2.2	Testgruppen . . . . .	58
5.2.3	Aufgabenstellung . . . . .	59
5.2.4	Ablauf . . . . .	59
5.2.5	Pretest . . . . .	60
5.3	Ergebnisdarstellung . . . . .	61
5.3.1	Datenaufbereitung . . . . .	62
5.3.2	Demografischer Hintergrund . . . . .	63
5.3.3	Relevante Vorkenntnisse . . . . .	64
5.3.4	Bewertung der Steuerungskonzepte für Kollaboration . . . . .	68
5.3.5	Fazit . . . . .	74
<b>6</b>	<b>Fazit</b>	<b>81</b>
6.1	Zusammenfassung . . . . .	81
6.2	Ausblick . . . . .	82
<b>7</b>	<b>Materialien der Untersuchung</b>	<b>84</b>
7.1	Fragebogen vor dem Experiment . . . . .	85
7.2	Fragebogen nach dem Experiment . . . . .	87

# Listings

# 1 Einleitung

Die Wirtschaft ist stetig im Wandel und so auch die Anforderungen, die an die Arbeitnehmer in einem Unternehmen gestellt werden. Arbeitsprozesse werden immer komplexer und sind kaum noch überschaubar. Dies hat zur Folge, dass in den wenigsten Fällen einzelne Mitarbeiter noch effektiv in der Wirtschaft agieren können, sondern das Zusammenspiel einer Vielzahl von unterschiedlichen Kompetenzbereichen gefordert ist. Als Reaktion darauf setzen Unternehmen seit längerem immer mehr auf kleine selbständig agierende Projektteams, um gemeinsam unterstützt von Computertechnologien das entsprechende Unternehmensziel zielorientiert verfolgen zu können (vgl. Sch95, S.19). Dementsprechend können Menschen mit verschiedenen Kompetenzbereichen und Stärken enger miteinander zusammenarbeiten, wodurch ein besserer Austausch über die Gesamtsituation innerhalb des Projektes statt finden kann.

Eine andere beobachtbare Veränderung, die jedoch keinen direkten Zusammenhang mit der Ersten hat, ist, dass Bildschirme immer größer und funktionsreicher werden, die Preise zeitgleich aber zunehmend sinken und sie somit immer attraktiver für den breiten Markt werden. Bildschirme in einer Größenordnung an denen ein einzelner Mensch nicht optimal interagieren kann sind keine Seltenheit mehr. Den Zuwachs dieser Bildschirme (siehe Abbildung 1.1) kann man auch an der Anzahl der Medienfassaden feststellen, die vor allem zunehmend an Flughäfen, U-Bahn Stationen und größeren Einkaufspassagen angebracht werden und dem Menschen das Wetter, die nächsten Transportmöglichkeiten oder Ähnliches anzeigen. Bis heute wird ein Großteil dieser Medienfassaden nur für passive nicht interaktive

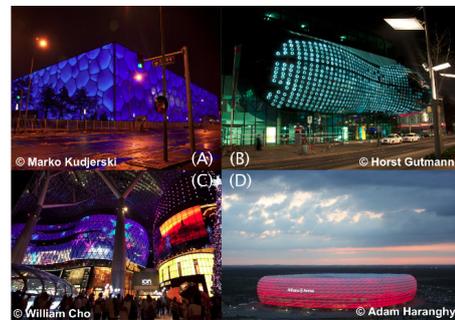


Figure 2. Media façades of different sizes and form factors: (A) The National Aquatics Center, Beijing, China (Photo: Marko Kudjerski, 2012), (B) Kunsthau Graz in Graz, Austria (Photo: Horst Gutmann, 2011), (C) the ION Orchard building in Singapore (Photo: William Cho, 2009) and (D) the Allianz Arena in Munich, Germany (Photo: Adam Haranghy, 2009).

Abb. 1.1: Medienfassaden Quelle:  
(GHLK13)

Zwecke angebracht, doch Vertreter unterschiedlichster Disziplinen versuchen eine neue Art der Kommunikation und Interaktion mithilfe dieser zu realisieren<sup>1</sup>. Diese großen Bildschirme schaffen nicht nur an öffentlichen Plätzen, sondern auch im privaten Bereich für Familien und Wohngemeinschaften sowie in Unternehmen neue Möglichkeiten. Kleinere Projektteams können so gemeinsam an einem Anzeigemedium arbeiten, um bei Meetings kollaborativ Entscheidungen zu treffen, Aufgaben zu verteilen oder zu bewältigen.

### 1.1 Motivation

Diese großen Bildschirme oder auch Medienfassaden bieten neue Möglichkeiten der Interaktion als Team an einem Anzeigemedium (siehe Abbildung 1.2). Dabei sollen 2D Gesten sowie auch 3D Gesten für die Interaktion verwendet und somit ein Informationsaustausch aus jeglicher Entfernung ermöglicht werden, um den Bereich aus dem das Anzeigebild erkennbar ist, effektiv für mehrere Personen zu nutzen und das Team bei Konferenzen oder sonstigen kollaborativen Sitzungen visuell und funktionell zu unterstützen und den gesamten Workflow positiv zu beeinflussen.

Der Grund warum gerade Gestensteuerungskonzepte so interessant für diesen Bereich sind und von typischen Konzepten wie der Maus und Tastatur abgesehen wird, ist, dass bei den üblichen Konzepten eine zusätzliche Abstraktionsebene vor dem Senden der eigentlichen Information steckt. Möchte man ein Objekt auf dem Bildschirm beispielsweise markieren, muss man erst einmal eine Maus, die auf einem Tisch also einer horizontalen Fläche liegt, bewegen und zwar so, dass sich ein, auf einem senkrecht stehenden Bildschirm befindlicher Zeiger, auf das entsprechende Objekt bewegt. Dabei stimmt in den seltensten Fällen auch die Distanz, die die Maus auf dem Tisch zurücklegen muss mit der Distanz, die der Zeiger zurück legt überein, sondern es Bedarf dabei einer weiteren Abstraktion. Gegebenenfalls muss man sogar mehrfach eine Mausbewegung ausführen um mit dem Zeiger zur vorgesehenen Position zu gelangen. Gesten sind hingegen intuiti-



Abb. 1.2: computer supported collaborative work Quelle: (BDHM11)

---

<sup>1</sup>Beispiel: <http://www.mediafacades.eu/>

ver und direkter. Man zeigt beispielsweise mit seinem Finger auf das gesuchte Objekt und es ist selektiert. Je einfacher die Bedienung und Informationsübergabe ist, umso schneller geht die Arbeit von statten und umso mehr können sich die Benutzer auch dem Informationsaustausch untereinander widmen.

### 1.2 Themenbeschreibung

Ziel dieser Arbeit soll es sein 2D und 3D Gesten sowie deren Kombination zu evaluieren und auf ihre Tauglichkeit hin zu überprüfen, einem Team zu ermöglichen gemeinsam an einem Anzeigemedium eine gestellte Aufgabe zu bewältigen. Weiter wird evaluiert für welche Art von Interaktion jeweils 2D oder 3D Gesten bevorzugt werden, inwiefern die Kombination beider Steuerungstechniken das Arbeiten als Team unterstützt, wie sie von den Benutzergruppen aufgenommen werden und inwieweit Transitionen von einer Steuerungstechnik zur anderen stattfinden. Es wird dabei nicht nur ein Augenmerk auf die dafür nötige Computer - Mensch Interaktion, sondern auch Mensch - Mensch Interaktion gelegt. Hierfür wird ein geeignetes Umfeld geschaffen um entsprechende Tests durchführen zu können. Als TestszENARIO wird hierbei ein Team simuliert und betrachtet, welches gemeinsam an der Zusammenstellung einer neuen Bildercollage arbeiten soll. Das Team besteht aus drei Personen, die entweder einige Meter vor dem Bildschirm mit Handgesten interagieren oder die anderen Personen, die direkt vor dem Bildschirm stehen und mit Touchgesten interagieren, dirigieren können.

### 1.3 Gliederung

Die Arbeit ist in sechs Kapitel gegliedert.

Zunächst wurden in der Einleitung (siehe Kapitel 1) einige praxisrelevante technologische Änderungen aufgegriffen, die die Motivation darstellen, die zur Entwicklung dieser Arbeit führten. Das Thema wurde beschrieben und der grobe Aufbau der Arbeit erläutert.

Im Grundlagenkapitel (siehe Kapitel 2) werden für diese Arbeit relevante Begriffe und ihre Zusammenhänge zueinander erläutert. Darunter fallen unter anderem HCI (Human Computer Interaction), CSCW (Computer Supported Collaborative Work),

Ubiquitous Computing und Gesten.

Nach den Grundlagen folgt das Analysekapitel (siehe Kapitel 3), in dem die, an die notwendige Software gestellten, Anforderungen entwickelt und analysiert werden. Bei den Anforderungen handelt es sich zum einen um bedienungsbezogene Anforderungen, auf kurze sowie lange Distanz zum Anzeigemedium für mehrere zeitgleiche Benutzer. Zum anderen handelt es sich um Anforderungen, die eine Software zur Erstellung einer Bildercollage mindestens erfüllen sollte. Auch Anforderungen an die Benutzeroberfläche, die durch die Benutzung von mehreren parallelen Anwendern und die enorme Größe des Displays spezielle Kriterien erfüllen sollte, werden betrachtet. Weiter wird sich für die notwendige Hardware entschieden.

Als nächstes folgt das Entwurfskapitel (siehe Kapitel 4), in dem sich bei genauerer Betrachtung der Anforderungen für eine Softwarearchitektur entschieden wird. In diesem Kapitel wird auf die einzelnen Softwarekomponenten, ihre Aufgaben und Schnittstellen eingegangen, sowie gewählte Programmierparadigmen begründet.

Die Implementierbarkeit des Entwurfs und anschließend auch Nutzbarkeit der Steuerungsmethoden werden in der Evaluation (siehe Kapitel 5.1) betrachtet. Bei der Betrachtung der Implementierbarkeit werden zunächst die einzelnen Komponenten und verwendeten Libraries vorgestellt. Anschließend werden, bei der Evaluation der Nutzbarkeit, praxisnahe Testszenarien entwickelt, mithilfe von Freiwilligengruppen durchgeführt und die dabei erhobenen Daten ausgewertet. Anhand dieser Experimente werden neue Erkenntnisse, in Bezug auf die Kombination von 2D und 3D Gesten zur Erfüllung einer Gruppenaufgabe, herauskristallisiert.

Zum Schluss (siehe Kapitel 6) folgt eine Zusammenfassung der innerhalb dieser Arbeit erlangten Kenntnisse und ein Ausblick auf die zukünftige Entwicklung des Themenbereichs und Fragen, die es noch zu beantworten gilt.

## 2 Grundlagen

Im folgenden Kapitel werden für die Arbeit relevante Begriffe und Forschungsgebiete erläutert, miteinander verknüpft und in den Kontext dieser Arbeit eingeordnet.

### 2.1 Gruppe

Der Begriff Gruppe hat unzählige unterschiedliche Definitionen in der Literatur, die sich jeweils auf verschiedene Aspekte einer Gruppe fokussieren. Forsyth hat, in seinem Buch *Group Dynamics* (vgl. [For05](#)), einen Teil der Definitionen in einer übersichtlichen Tabelle zusammengetragen. Innerhalb dieser Arbeit wird sich auf folgende Definition einer Gruppe bezogen:

„Gruppen sind soziale Systeme von zwei oder mehr Personen, die durch gegenseitige Kenntnis und potentielle gegenseitige Interaktion miteinander verbunden sind.“ ([Sch95](#), S.127)

Diese Definition ist an eine, von McGrath 1984 verfasste, Definition angelehnt die in der bisherigen Forschung zum Thema Computerunterstützung von Gruppenarbeiten mit am häufigsten verwendet wird. Ob zwei oder drei Personen die minimale Größe einer Gruppe ausmachen wird bis heute diskutiert. Auch die Obergrenze ist schwer festzumachen. Die Gruppe muss laut Definition klein genug sein, dass es jedem Mitglied möglich ist mit jedem Anderen in Interaktion zu treten. Innerhalb der Gruppe können den Mitgliedern verschiedene Rollen, Pflichten und Zuständigkeiten zugeteilt werden oder sich aus der Gruppendynamik selbst herauskristallisieren. So kann man beim Beobachten von Gruppen häufig Rollen, wie den Gruppenleiter, Moderator oder Protokollant auffinden (vgl. [Sch95](#), S.127-128).

### **Arbeitsgruppe**

„Eine Arbeitsgruppe ist eine Gruppe innerhalb einer Organisation, die gemeinsam an einer Aufgabe oder an dem Erreichen eines gemeinsamen Ziels arbeitet“ (Sch95, S.129). Sie bearbeiten die Aufgabenstellung mit Hilfe von Kommunikation und Kooperation. Ohne Kommunikation, also eine vernünftige Absprache, können die Gruppenmitglieder keinen Plan formulieren den es zu verfolgen gilt und ohne Kooperation wären weder ineinander greifende Handlungen unterschiedlicher Personen, noch ein koordinierter Informations- und Meinungs austausch in Bezug auf die Aufgabe, möglich (vgl. BG10). Gruppen werden also in den Kontext von Arbeit gesetzt und so werden Familien oder Freundeskreise überwiegend ausgeschlossen. Arbeitsgruppen haben im Normalfall einen geringeren emotionalen Zusammenhalt als andere Gruppen. Sie bekommen als Einheit eine Aufgabe zugeordnet, die es zu erfüllen gilt und unterliegen häufig formalen Strukturen innerhalb der Gruppe, die von der Organisation vorgegeben werden (vgl. Sch95, S.129).

### **Gruppenarbeit**

Als Gruppenarbeit wird der Prozess bezeichnet, bei dem Arbeiter gemeinsam ein Produkt, gemäß einem gemeinsamen Ziel oder einer gemeinsamen Aufgabenstellung, erstellen. Komponenten einer Gruppenarbeit sind zum einen die vollziehenden Arbeiter, die angewendeten Arbeitsmittel und Materialien, das Produkt welches als Ergebnis der Gruppenarbeit entsteht, der Arbeitsprozess an sich und die entsprechende Aufgabe, die es zu erfüllen gilt (vgl. Sch95, S.129-130).

## **2.2 Human Computer Interaction**

HCI, oder auch Human-Computer Interaction genannt, ist ein wissenschaftlicher Teilbereich, der sich mit dem Design, der Evaluation und der Implementation von interaktiven Computer Systemen für die Verwendung durch Menschen, und alle sonstigen wesentlichen Teilaspekte dessen, befasst.

HCI lässt sich in eine Vielzahl an Disziplinen unterteilen, die von großer Relevanz für den Forschungsbereich sind. Abgesehen von der Informatik, die oftmals ihr Hauptaugenmerk auf die Software Architektur und technische Umsetzung solcher Computer Systeme

wirft, wirken auch die Bereiche Ingenieurwesen, Psychologie, Ergonomiewissenschaft, Soziologie, sowie jegliche Formen an Designwissenschaften, wesentlich auf den Bereich HCI ein (siehe Abbildung 2.1) (vgl. [Bad15](#), S.80). Psychologie sowie Soziologie spielen eine wichtige Rolle, da HCI die Beziehung zwischen Menschen und Computern studiert und dementsprechend gleichermaßen ein tiefgründiges Wissen von Computern wie auch Menschen von Nöten ist.

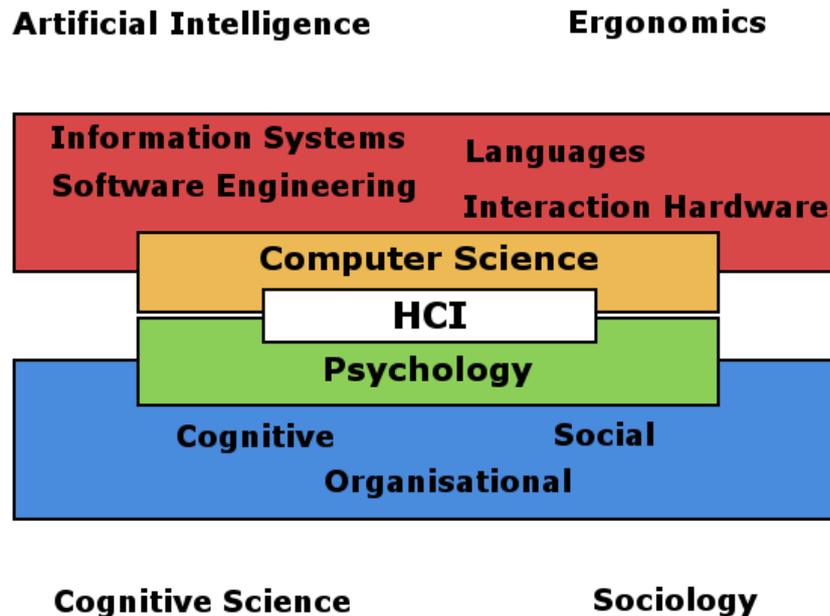


Abb. 2.1: Verschiedene Disziplinen im Bereich HCI. Quelle: (vgl. [Bad15](#), S.81)

Die von HCI verfolgten Ziele lassen sich mit Sicherheit, Funktionalität, Effektivität, Effizienz und Attraktivität des Systems zusammenfassen. Das wichtigste Ziel des HCI ist die Benutzbarkeit beziehungsweise Usability des Systems sicherzustellen. Leichte Erlernbarkeit, sowie leichte Bedienbarkeit sind wichtige Leitfäden für Systeme, um schnelle Erledigung von Aufgaben zu gewährleisten und das Wohlgefallen der Benutzer zu erlangen. Das System muss den Bedürfnissen und Anforderungen des Benutzers entsprechen (vgl. [III15](#), S.24). Wege um dies sicherzustellen, sind, das System iterativ weiter zu entwickeln, es regelmäßig zu testen, die Meinung der zukünftigen Anwender, sowie jeglicher an dem Forschungsgebiet HCI beteiligten Disziplinen einzuholen und das System entsprechend anzupassen (vgl. [III15](#), S.25).

## 2.3 Ubiquitous Computing

Ubiquitous Computing, häufig auch als Ambient Intelligence oder Pervasive Computing betitelt, ist ein Forschungsgebiet des HCI, welches das Ziel verfolgt die typischen stationären Computer an Schreibtischen durch integrierte, für den Menschen unsichtbare, Computer in der Umgebung zu ersetzen. Der Mensch soll in der Lage sein mit jeglichen Gegenständen in seinem Umfeld interagieren zu können, die mit kleinen Computern ausgestattet sind und drahtlos miteinander kommunizieren und ihn somit im Alltag unterstützen. Als Ubiquitous bezeichnet man außerdem die dritte Computer-Ära, in welcher Computer allgegenwärtig sind und der Mensch jederzeit auf eine Vielzahl dieser Zugriff hat und sie nutzen kann. Die erste Computerer-Ära war die Mainframe Ära, in der ein Computer jeweils für mehrere Personen zur Verfügung stand, gefolgt von der Personal Computer Ära, in der jede Person ihren persönlichen Computer zur Verfügung hatte (siehe Abbildung 2.2) (vgl. [Bad15](#), S.80).

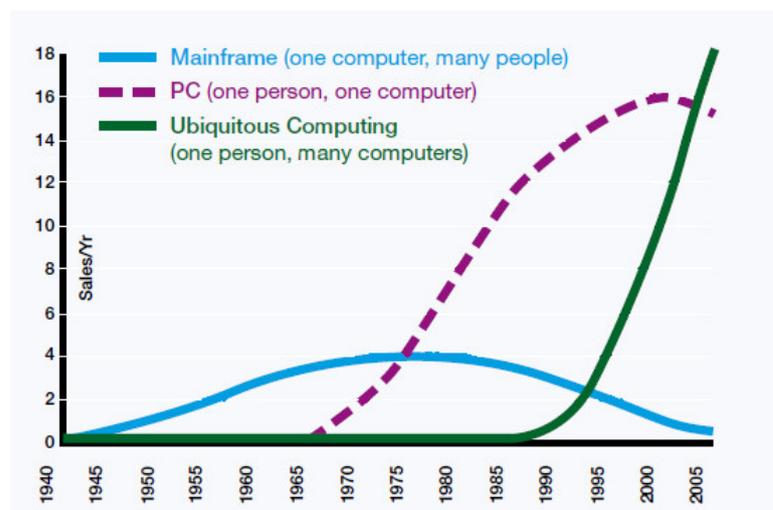


Abb. 2.2: Evolution des Ubiquitous Computing. Quelle: (vgl. [Riv05](#), S.5)

Die Vision des Ubiquitous Computing wurde erstmals von Mark Weiser (vgl. [Wei91](#)) zu einer Zeit formuliert, in der die Technologie noch nicht weit genug war das Vorhaben zu unterstützen. Doch in den letzten Jahrzehnten haben Technologien wie Handheld- und Wearable-Computer, drahtlose Netzwerkverbindungen, und Sensoren wie beispielsweise die Kinect oder Leap Motion vieles geändert (vgl. [Sat01](#), S.1). Das Internet der Dinge, also eine Vernetzung jeglicher Alltagsgegenstände, die Informationen aus der Umgebung

aufgreifen, miteinander teilen und diese verarbeiten und interpretieren ist schon lange keine Zukunftsvision mehr. Mit dem Wegfall des typischen stationären Arbeitscomputers, und somit der Tastatur und Maus, müssen sich auch die Eingabemethoden evolvieren. So werden auch Arbeitsplätze in größeren Unternehmen einem stetigen Wandel unterliegen. In diesem Zusammenhang spielt der Begriff CSCW (siehe Kapitel 2.4) eine wichtige Rolle.

### 2.4 Computer Supported Cooperative Work

Computer-Supported Cooperative Work (CSCW) ist ein Forschungsgebiet der Informatik und Teilbereich des HCI (siehe Kapitel 2.2), welches sich mit der Nutzung einer Vielzahl von Computer-, Netzwerk- und Telekommunikationstechnologien beschäftigt, die die Arbeitsweise einer Gruppe von Menschen an einer gemeinsamen Aufgabe verbessern soll. Dabei soll die Arbeit einfacher gestaltet und die Effizienz erhöht werden.

Doch auch andere Disziplinen befassen sich mit dem Forschungsgebiet, wie die Soziologie, Psychologie, Arbeits- und Organisationswissenschaften, Ethologie, Anthropologie sowie Wirtschaftswissenschaften. Auch verschiedenste Bereiche der Informatik wie Usability, Verteilte Systeme, Künstliche Intelligenz und Informationssysteme sind an CSCW beteiligt (vgl. Ger07, S.143-146). Die häufigste in der Literatur auffindbare Definition ist von Paul Wilson:

CSCW is a generic term that combines the understanding of the way people work in groups with the enabling technologies of computer networking and associated hardware, software, services and techniques (Wil91, S.1).

Ansgar Gerlicher unterteilt die CSCW Systeme in drei Anwendungsbereiche sowie Synchronisationstypen, die in der Abbildung (siehe Abbildung 2.3) visualisiert sind. Die eindeutige Trennung ist in einigen Fällen jedoch nicht möglich und ein System kann Teil mehrerer dieser Kategorien sein (vgl. Ger07, S.143-146).



Abb. 2.3: Kategorisierung von Groupware / CSCW Systemen. Quelle: (vgl. Ger07, S.144)

Eine weitere Unterteilung der CSCW Systeme, oder auch Groupware Systeme, ist mithilfe der Raum-Zeit-Matrix (siehe Abbildung 2.4) möglich. Diese unterteilt Software zum Einen in asynchrone und synchrone Systeme und zum Anderen in collocated und remote Systeme ein. Das Display, an dem ein Projektteam gemeinsam eine Aufgabe erledigen soll, wäre beispielsweise ein synchrones und kollaboratives System. Displays, an denen Notizen für die nächste Schicht hinterlassen werden, können, würden unter asynchrone, kollaborative Systeme fallen, da die Arbeit zwar am gleichen Ort, jedoch zeitverschoben stattfindet. Systeme die eine Konferenz in einer virtuellen Umgebung ermöglichen sollen, würden unter die synchronen, örtlich getrennte Systeme fallen und als letzte Kategorie existieren noch Systeme wie E-Mail Verteiler, Versionierungssoftware oder Ähnliches, die asynchron wie auch ortsunabhängig sind. Den Grundstein dieser Taxonomie haben De Sanctis und Gallupe (1987) gelegt. Johansen (1988) hat ihre Arbeit daraufhin zu der häufig in der Literatur genutzten unten dargestellten 2x2 Matrix verfeinert (vgl. Bae14, S.742).

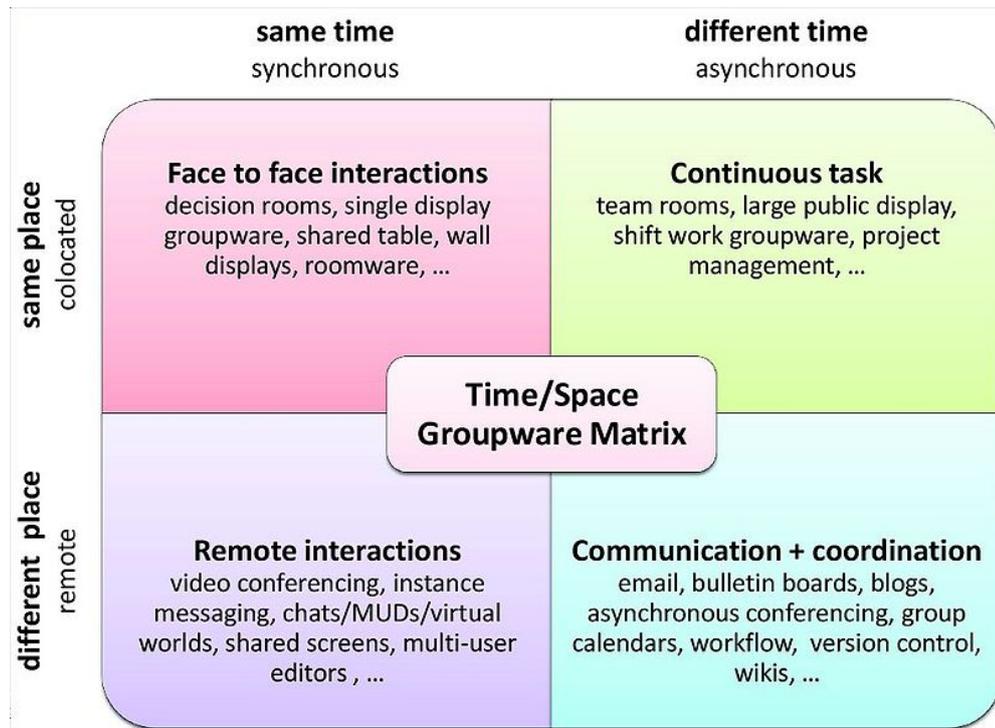


Abb. 2.4: Groupware Time / Space Matrix Quelle: (Wik16)

## Groupware

Während man mit CSCW das Forschungsfeld beschreibt, welches sich mit den theoretischen Grundlagen befasst, sind unter Groupware, die dabei entwickelten Softwarelösungen zu verstehen. Das Wort setzt sich aus „Group“, also Gruppe, sowie Software zusammen. Es umschreibt jegliche Systeme, die eine Gruppe durch Kommunikations- und Koordinationsmittel dabei unterstützen zusammen zu arbeiten (vgl. Ger07, S.148).

## 2.5 Mental Model

Unter Mental Model versteht man ein Modell, welches ein Benutzer aufstellt um sich fremdartige Dinge zu erklären. Stellt man bezogen auf diese Arbeit einen Anwender vor einen ihn unbekanntem Bildschirm, entsteht in ihm ein mentales Modell, welches ihm eine Erklärung liefert worum es sich dabei handelt und wie er damit zu interagieren hat. Vorherige Erfahrungen des Anwenders haben starken Einfluss auf den Aufbau dieser

mentalen Modelle. Menschen assoziieren fremdartige Dinge automatisch mit Dingen, die einen gewissen Grad an Ähnlichkeit aufweisen und ihnen bekannt vorkommen. Deshalb versuchen sich Entwickler bei dem Design von Interfaces an Standards zu orientieren, die den Benutzern bereits bekannt vorkommen. Auch das Feedback des Systems an den Anwender ist ein wichtiges Werkzeug, um das sein mentales Modell zu beeinflussen und in eine gewisse Richtung zu lenken (vgl. Jac12, S.33). Laut Jacko (vgl. Jac12) sind die mentalen Modelle dabei viel fragmentierter als man annehmen würde. So entsteht für jede einzelne atomare Operation ein eigenes mentales Modell im Menschen statt ein Modell für das System als Ganzes. Möchte jemand beispielsweise ein Objekt auf dem Bildschirm verschieben, kann es für einige nahe liegen nach dem Objekt in der Luft zu greifen und es zu verschieben, da in dem Fall die Analogie zu realen Objekten im Raum greift. Möchte er wiederum eine Seite auf dem Bildschirm umblättern, könnte er die seit den Smartphones eingeführte Swipe Geste verwenden wollen, da ihm in dem Fall schneller die Analogie zu einem Smartphone als einem realen Buch in den Sinn kommen würde. Eine wichtige und schwierige Aufgabe im HCI Bereich ist es also Systeme und Steuerungen zu entwickeln, deren Modelle eine größtmögliche Überschneidungen mit den mentalen Modellen der Anwender haben. Gelingt dieser Schritt spricht man von Seamless Interaction.

Ein interessantes Beispiel ist die „slide to unlock“ Funktion, welche bei den ersten Smartphone Modellen, wie auch dem ersten iPhone, auf dem Startbildschirm (siehe Abbildung 2.5) zu finden war, um das Gerät zu entsperren.



Abb. 2.5: Sperrbildschirm der ersten iPhone Generation. Quelle: Eigene Darstellung

Die Besitzer eines solchen Gerätes hatten damals keine Vorerfahrungen und würden ohne die Aufschrift nicht wissen was von ihnen erwartet wird. Es bestand in den meisten Fällen kein anwendbares mentales Modell, auf welches sie sich hätten berufen können. Doch über die Jahre hat sich diese Form der Interaktion in der Gesellschaft etabliert und so konnte man die Aufforderung zur Slide Bewegung immer dezenter werden lassen. Zunächst wurde die textuelle Aufforderung immer häufiger durch eine symbolische Aufforderung in Form eines Pfeils oder Schlosses ersetzt und ab einem gewissen Zeitpunkt wurde jegliche Aufforderung weggelassen (siehe Abbildung 2.6). Die Gesellschaft hat sich so stark an touchfähige Geräte gewöhnt, dass es selbstverständlich geworden ist ein Gerät ohne anderweitige Sperre, in Form eines Passworts oder Musters, durch eine Swipebewegung über den Bildschirm zu entsperren.



Abb. 2.6: Sperrbildschirm des Samsung Galaxy S5. Quelle: Eigene Darstellung

### Shared Mental Model

“Wie der Name vermuten lässt, handelt es sich bei shared mental models um Überlappungen in den individuellen Erfahrungen und den Erwartungen” (Jac12, S.48) bei einer Gruppe von Menschen. Es wird vermutet, dass das Mitteilen von relevanten Erfahrungen und Erwartungen über die entsprechende Situation oder Aufgabe, der sich die Gruppe entgegengestellt sieht, die Performanz der Gruppe steigert (vgl. Jac12, S.48).

### Seamless Interaction

Wie bereits angesprochen (siehe Kapitel 2.3), werden Computertechnologien immer günstiger und kleiner und daraus folgend stetig allgegenwärtiger. Diese Allgegenwärtigkeit ist es, die Seamless Interaction zu einem so wichtigen Thema werden lässt. Immer mehr Gegenstände im Alltag haben kleine Computer verbaut und sollen sich ohne große Mühe und bestenfalls ohne Bedienungsanleitung ansteuern lassen. Seamlessness ist für die Akzeptanz der Systeme von größter Bedeutung. Die Themengebiete Smart Home und Internet of Things, die in der Forschung ebenfalls viel Aufmerksamkeit in den letzten Jahren bekamen, treiben diese Entwicklung noch weiter voran. Damit Interaktion ohne großen Aufwand möglich ist, muss ein entsprechend gut passendes mentales Modell gefunden werden, sodass die sie ein intuitives Gefühl vermittelt. Hierfür sollte eine größtmögliche Überschneidung, zwischen dem mentalen Modell des Entwicklers eines Produktes, sowie des letztendlichen Anwenders, angestrebt werden.

## 2.6 Gesten

Innerhalb des Bereiches HCI (siehe Kapitel 2.2) gibt es eine Vielzahl unterschiedlichster Definitionen für Gesten. Viele beziehen sich in dem Zusammenhang auf Kurtenbach und Hulteen, die Gesten wie folgt definieren:

A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed (KH90).

Doch auch diese Definition ist laut Caroline Hummels, Gerda Smets und Kees Overbeeke ausbaufähig. Sie bringen als Beispiel einen Designer an, welcher mit einem Stift Linien zieht. Laut der Definition von Kurtenbach und Hulteen würde es sich hierbei nicht um eine Geste handeln, da in diesem Fall die Linie und nicht die Hände und ihre Bewegungen im Fokus liegen. Würde der Designer nun aber einen virtuellen Stift oder seinen Finger zum Zeichnen in einer virtuellen Umgebung verwenden, wäre es wieder eine Geste nach ihrer Definition, da nun die Finger im Fokus der Gestenerkennung liegen (vgl. HSO97, S.2). Des Weiteren geben sie eine eigene Definition an:

A gesture is a movement of one's body that conveys meaning to oneself or to a partner in communication (HSO97, S.2).

Sie fügen weiter an, dass der Kommunikationspartner in dem Fall sowohl menschlich als auch ein Computer sein kann. Harling und Edwards verstehen wiederum unter Gesten nicht nur Bewegungen des Körpers, sondern auch statische Handhaltungen (vgl. HE97, S. 2), wie es beispielsweise bei der im Entwurf (siehe Abschnitt **imaginäre Fläche**) festgelegten Selektionsgeste der Fall ist.

Deshalb ist im Zusammenhang mit dieser Arbeit unter einer Geste eine Bewegung oder statische Haltung des Körpers oder eines Körperteils zu verstehen, welche eine Bedeutung an einen selbst oder einen beliebigen Empfänger überliefern soll.

Gesten sind ein intuitives und direktes Verfahren um Informationen auszutauschen. Möchte man beispielsweise eine Seite umblättern, braucht man nur eine Swipe Bewegung über die entsprechende Seite in die gewollte Richtung tätigen in die auch umgeblättert werden soll, statt auf einen entsprechenden Knopf zum umblättern zu drücken oder den Befehl „Blättere um“ einzutippen oder sonstige Verfahren mit einer oder mehreren Indirektionsebenen anzuwenden. Solche Interaktionen fühlen sich natürlicher an, da man mit realen Objekten auf eine sehr ähnliche Art und Weise interagieren würde. Im Idealfall entfallen wie in der Motivation (siehe Kapitel 1.1) angerissen, jegliche Abstraktionsebenen zwischen dem Ziel eine bestimmte Information zu übermitteln und der eigentlichen Übermittlung.

### 2.6.1 Kulturelle Unterschiede

Dennoch ist gerade bei Gesten Vorsicht geboten. Denn anders als vielleicht angenommen „sind Gesten definitiv keine universale Sprache“ (Arc97, S.79). Je nach der Kultur, in der man sich aufhält, kann eine Geste eine völlig andere Bedeutung besitzen. Genauso wie man nicht erwarten kann, dass ein deutsches Wort weltweit gleich verstanden wird, kann auch nicht erwartet werden, dass eine Geste überall die selbe Bedeutung inne hält. Wie die Sprache hat sich auch die Gestik über die Jahrhunderte entwickelt und je nach Kultur ihre eigene Bedeutung erlangt. Wenn man beispielsweise im Iran von einem Einheimischen einen Daumen hoch gezeigt bekommt, könne man meinen, die Person wünscht einem Glück oder ihm gefällt etwas was man getan hat. Man sollte diese Geste jedoch keinesfalls erwidern, da sie dort eine völlig andere Bedeutung hat und mit zu der aggressivsten Form von „Scher-dich-zum-Teufel“ Gestik gehört (vgl. Arc97, S.80-81).

### 2.6.2 Taxonomie

Es gibt eine Vielzahl verschiedenster Taxonomien für Gesten, die über die letzten Jahrzehnte entwickelt wurden. Ein ausschlaggebender Grund hierfür ist, dass je nach Definition einer Geste, die abhängig vom Kontext unterschiedlich sein kann, auch die Klassifikation unterschiedlich ausfallen kann. Mit einer der bekanntesten ist die Taxonomie von Mc Neill (vgl. [McN92](#)), welcher Gesten in vier verschiedene Kategorien unterteilt:

1. **Iconic gestures**, die ein konkretes Element der Sprache umschreiben.
2. **Metaphoric gestures**, die ein abstraktes Element der Sprache umschreiben.
3. **Deictic gestures**, die zum Zeigen auf bestimmte Objekte der Realität verwendet werden.
4. **Beat gestures**, die unterstützend zur Sprache rhythmisch oder in bestimmten Momenten eingefügt werden um dem Gesagten mehr Bedeutung zuzuweisen.

Da der Bereich der Human Computer Interaction jedoch einen etwas anderen Blickwinkel auf die Thematik Gestik benötigt, hat sich Vafaei in seiner Masterarbeit ([Vaf13](#)) dem Thema angenommen. Vafaei hat viele frühe Taxonomien und auch Taxonomien aus den letzten Jahren thematisiert, evaluiert und basierend auf zwei bestimmten Arbeiten von Ruiz ([RLL11](#)) und Wobbrock ([WMW09](#)) eine eigene an die HCI angepasste Taxonomie für Gesten entwickelt, die in der folgenden Tabelle (siehe Abbildung [2.7](#) und [2.8](#)) visualisiert wird.

<b>Group 1: Gesture Mapping</b>		
<b>Nature</b>	<i>Manipulative</i>	Gesture directly manipulates the object
	<i>Pantomimic</i>	Gesture imitates a real meaningful action
	<i>Symbolic</i>	Gesture visually depicts a sign
	<i>Pointing</i>	Gesture points to a specific location
	<i>Abstract</i>	Gesture mapping is arbitrary
<b>Form</b>	<i>Static</i>	No motion/change is in gesture
	<i>Dynamic</i>	Motion/change occurs in gesture
	<i>Stroke</i>	Gesture consists of tap/flick(s)
<b>Binding</b>	<i>Object-Centric</i>	Location defined w.r.t. object features
	<i>World-Dependent</i>	Location defined w.r.t. world (app. environment) features
	<i>World-Independent</i>	Gesture can occur anywhere. Location can ignore world features
<b>Temporal</b>	<i>Continuous</i>	Action/task is performed during gesture; user can see impact of gesture simultaneously
	<i>Discrete</i>	Action/task is performed after completion of gesture; user can see impact after gesture is done
<b>Context</b>	<i>In-Context</i>	Gesture requires specific context
	<i>No-Context</i>	Gesture does not require specific context

Abb. 2.7: Eigenschaften einer Geste - 1. Quelle: (Vaf13, S.14-15)

<b>Group 2: Physical Characteristics</b>		
<b>Dimensionality</b>	<i>Single-Axis</i>	Gesture occurs around a single axis
	<i>Double-Axis</i>	Gesture occurs on a surface, a 2D plane
	<i>Tri-Axis</i>	Gesture involves either translational or rotational motion, not both, in a 3D space
	<i>Six-Axis</i>	Gesture occurs around both rotational and translational axes, in a 3D space
<b>Complexity</b>	<i>Simple</i>	Gesture consists of one atomic gesture
	<i>Compound</i>	Gesture can be decomposed into two or more simple gestures
<b>Body Part</b>	<i>Hand</i>	Arm is fixed, but palm or fingers move
	<i>Arm</i>	Arm moves (hand moves as well)
	<i>Head</i>	Gesture is performed by head movement
	<i>Shoulder</i>	Gesture is performed by shoulder movement
	<i>Foot</i>	Gesture is performed by foot movement
<b>Handedness</b>	<i>Dominant</i>	Gesture is performed by the user's dominant hand/arm
	<i>Non-Dominant</i>	Gesture is performed by the user's non-dominant hand/arm
	<i>Bi-Manual</i>	Gesture is performed using both hands/arms
<b>Hand Shape</b>	<i>Flat, Open, Bent, Curved, Index Finger, Fist, ASL Shapes ...</i>	
<b>Range of Motion (ROM)</b>	<i>Small</i>	Joint rotation is less than 50% of its normal ROM.
	<i>Large</i>	Joint rotation equals or is more than 50% of its normal ROM

Abb. 2.8: Eigenschaften einer Geste - 2. Quelle: (Vaf13, S.14-15)

Vafaei unterteilt hierbei die Gesten nach zwei verschiedenen Grundaspekten. Zum einen dem Gesture Mapping, welches die Gesten und ihre inhaltlichen Bedeutungen verbinden sollen und zum anderen die Physical Characteristics, welche die physikalischen Attribute der Geste an sich umschreiben sollen.

Das **Gesture Mapping** lässt sich wiederum in Nature, Form, Binding, Temporal und Context unterteilen.

Das **Nature** Attribut soll den Zweck der Geste umschreiben, ob eine Geste beispielsweise ein Objekt manipulieren, auf etwas Zeigen oder Ähnliches soll.

Das Attribut **Form** spiegelt wieder ob eine Geste eine Bewegung beinhaltet oder statisch ist.

Anhand des **Binding** Attributs wird dargestellt, ob die Geste an ein bestimmtes Objekt oder einen Ort gebunden ist und nur in diesem Zusammenhang existieren kann.

Das **Temporal** Attribut gibt an, ob der Effekt der Geste zeitgleich mit der Ausführung dieser geschieht oder erst nach Beendigung der Geste eine Auswirkung zu sehen ist.

Und das **Kontext** Attribut legt fest ob eine Geste einen gewissen Kontext benötigt.

Die **physikalischen Attribute** unterteilen sich wiederum in Dimensionality, Complexity, Body Part, Handedness, Hand Shape und Range of Motion.

**Dimensionality** legt fest wie viele Dimensionen für die Interpretation der Geste benötigt werden.

**Complexity** gibt an ob die Geste aus mehreren Untergesten besteht oder atomar ist.

**Body Part** zählt die Körperteile auf, die an der Ausführung der Geste beteiligt sind.

**Handedness** stellt dar mit welcher Hand die Geste jeweils ausgeführt werden kann.

Das **Handshape** Attribut umschreibt die Form in der die gestikulierende Hand während der Ausführung gehalten wird und **Range of Motion** die Reichweite der Gestenbewegung beziehungsweise die Rotationsreichweite an den Gelenken .

### 2.6.3 Unintentional Gestures

Weiter gibt es unabhängig von dieser Taxonomie auch unintentional gestures also unbeabsichtigte Gesten, die getätigt werden können. Ein Großteil dieser Gesten wird durch Handbewegungen relativ zum Gesicht, wie das Augen reiben, das Kopfkratzen oder das Kopfstützen (siehe Abbildung 2.9) ausgelöst. Diese Gesten sind meist unbeabsichtigt, können aber dennoch nützliche Meta Informationen über den Gemütszustand des Anwenders zum jeweiligen Zeitpunkt liefern. (vgl. AAU08, S.355-362) Ein Beispiel könnte ein frustrierter Anwender darstellen, der Probleme mit der Gestensteuerung hat und sich fragend den Kopf kratzt. Erkennt und interpretiert das System solche Gesten richtig, könnte es in der Lage sein entsprechend passende Hilfestellungen zu leisten.



Abb. 2.9: Während eines Experiments observierte ungewollte Gesten. Quelle: (AAU08, S.358)

Die Anzahl an unbeabsichtigten Gesten, die teils auch dem selben Zweck dienen ist enorm, wie man bereits an der Vielfalt der Kopfstützgesten in der nächsten Abbildung (siehe Abbildung 2.10) erkennen kann.



Abb. 2.10: Verschiedene Gestenvariationen beim Abstützen des Kinns. Quelle: (AAU08, S.362)

Diese Gesten bei der Gestenerkennung zu filtern und nicht als beispielsweise manipulative Geste fehlzuinterpretieren gestaltet sich als sehr komplex und stellt eine von vielen Hürden der 3D Gestenerkennung dar.

#### 2.6.4 2D Gesten

2D Gesten, alternativ auch Touchgesten genannt, sind Gesten, die mit einem oder mehreren Fingern, Stiften oder Ähnlichem auf einer ebenen Fläche ausgeführt werden (siehe Abbildung 2.11). Sogenannte Multi-Touch Geräte sind über eine Vielzahl verschiedenster Technologien (Resistance Based Touch Surfaces, Capacitance Based

Touch Surfaces, Surface Wave Touch Surfaces (SAW), Frustrated Total Internal Reflection, Diffuse Illumination, ...) in der Lage die Koordinaten, sowie teils auch weitere Metadaten der Berührungspunkte zu registrieren. Die Technologie ist bereits seit Jahrzehnten in stetiger Entwicklung und erreichte erstmals mit den ersten Multi-Touch Smartphones erfolgreich den globalen Markt (vgl. [Lin10](#), S.1-2); (vgl. [SBD<sup>+</sup>08](#), S.1-2).



Abb. 2.11: 2D Gesten auf einer Multitouch Oberfläche. Quelle: ([Lin10](#))

Ein großer Vorteil von Multi-Touch Systemen ist, dass sie durch entsprechende Anpassungen am Design der Anwendung, das Lösen jeglicher erdenklicher Probleme ermöglichen. Es ist beispielsweise möglich eine Tastatureingabe durch eine entsprechend abgebildete Tastatur auf dem Bildschirm nachzusimulieren. Außerdem steigert die Möglichkeit mehrerer Berührungspunkte die Produktivität und Effizienz der Interaktion, da eine Person durch die Verwendung mehrerer Finger komplexere Interaktionen oder mehrere kleine simple Interaktionen zur selben Zeit ausführen kann. Weiter sind durch die multiplen Berührungspunkte mehrere Anwender einsetzbar, die kollaborativ arbeiten können, da nicht unbedingt alle Berührungspunkte von der selben Person getätigt werden müssen. Auch die restlichen Informationen, die die Sensoren unter Umständen liefern können, wie

Winkel, Druckkraft oder Körpertemperatur können sinnvoll innerhalb der Anwendung verwendet werden. Doch Multi-Touch Systeme haben auch ihre Nachteile. Zum Einen können die Hände und Finger, die für die Interaktion benötigt werden den in Interaktion befindlichen Bereich des Bildschirms verdecken, zum Anderen lässt die Präzision der Steuerung bei dickeren Fingern entsprechend nach (vgl. [Lin10](#), S.1-2). Außerdem müssen Touchgesten, vor allem wenn sie nicht intuitiv erscheinen erst erlernt werden, bevor sie effektiv genutzt werden können. Je nach Anwendungsszenario können noch weitere Nachteile einer Touchgestensteuerung zum Vorschein kommen. Betrachtet man beispielsweise eine große Multi-Touch Oberfläche, wie den BenQ RP 790 (siehe Kapitel [3.4](#)), wird schnell deutlich, dass unter anderem ein Problem mit der Reichweite des Anwenders und der Interaktion mit weiter entfernten Bereichen der Oberfläche hinzukommt. Zusätzlich stellt die notwendige Nähe zum Anzeigegerät, welches zeitgleich als Eingabegerät fungiert, ein Problem dar, da das Sichtfeld des Anwenders eingeschränkt wird und er unter Umständen nicht mehr die gesamte Oberfläche betrachten kann.

### 2.6.5 3D Gesten

3D Gesten sind Gesten, die mit dem ganzen Körper oder Teilen des Körpers in einem dreidimensionalen Raum ausgeführt werden. Mit dem Zeigefinger auf etwas zeigen, mit der Hand nach etwas greifen oder auch das bloße Anheben eines Beines können als 3D Gesten angesehen werden. Möglichkeiten, diese Gesten zu erkennen gibt es, genauso wie bei den 2D Gesten, viele. Die ersten Erfolge wurden mit gerätebasierten Erkennungsmethoden verbucht. Dabei wurden in den Geräten verbaute Beschleunigungs- sowie Positionssensoren verwendet. Beispiele für diese Herangehensweise sind eine Vielzahl an Datenhandschuhen, der Controller der Nintendo Wii und die BlueWand (vgl. [FKO03](#)). Doch auch der Großteil der Smartphones, sowie neue Konsolencontroller beinhalten die entsprechende Sensorik und könnten dementsprechend für die Gestenerkennung verwendet werden. 3D Gestenerkennungssysteme die mit Tiefenbildern arbeiten kamen erst etwas später hinzu und haben erstmals mithilfe der Microsoft Kinect (siehe Kapitel [3.4](#)) den globalen Markt erreicht. Daraufhin folgten viele weitere Geräte für die 3D Gestenerkennung, wie unter anderem die Leap Motion (siehe Kapitel [5.3.5](#)).

Es wurden viele unterschiedliche Algorithmen für die visuelle Erkennung und das maschinelle Erlernen von Gesten aus Bildmaterial entwickelt, wie das Dynamic Time Warping und das Hidden Markov Modell, auf die innerhalb dieser Arbeit jedoch nicht tiefer eingegangen wird.

Außerdem wird an schallbasierten Gestenerkennungsmethoden, wie dem SoundWave (GMPT12) und SoundSense (YTCYPH+12), sowie Systemen, die mithilfe der von Menschen erzeugten, elektrischen Felder, (LGTIK14) Gesten erkennen, gearbeitet.

Auch 3D Gesten haben ihre Vor- und Nachteile. Ein großer Vorteil ist bei den meisten Systemen die Unabhängigkeit von Eingabegeräten. Die Person kann interagieren ohne ein Gerät mitführen oder aufsuchen zu müssen. Außerdem stehen ohne Eingabegerät auch beide Hände während der Interaktion frei zur Verfügung. Man kann auf flexible Entfernung zu dem interpretierendem System interagieren und muss nicht genau an der Maus sitzen oder sich direkt an den Multi-Touch Bildschirm stellen. Durch die Möglichkeit auch auf höhere Entfernungen mit dem System zu interagieren ist beispielsweise bei größeren Displays auch ein besserer Überblick über die Gesamtsituation auf dem Display möglich und der Benutzer ist in der Lage mit Objekten zu interagieren, die so weit von seinen derzeitigen Standort entfernt positioniert sind, dass er für eine Interaktion mittels Touchgeste erst seinen gesamten Körper umpositionieren müsste um es zu erreichen.

Gerade bei der Entwicklung von Smart Environments (Ber11) oder zukünftigen Automodellen (BMW15) wird der 3D Gestensteuerung hohe Aufmerksamkeit zuteil, da sie dort entsprechend einen enormen Mehrwert beitragen kann. So ist man beispielsweise in der Lage in Smart Homes mit einer einfachen Wischgeste Richtung Vorhang diesen zu schließen, mit einer kleinen vorkonfigurierten Bewegung den Fernseher einzuschalten oder im Auto die Lautstärke des Radios anzupassen, ohne die Augen von der Straße richten zu müssen und nach dem jeweiligen Knopf zu suchen.

Doch 3D Gesten haben auch ihre Schattenseite. Ein großes Problem ist die Akzeptanz dieses Steuerungskonzepts, da es noch in Kinderschuhen steckt und bisher keine einheitlichen Gestensets existieren, wie es bei den Touchgesten, nach jahrelangen Werbemaßnahmen, der Fall ist. Gerade ältere Testpersonen äußerten in den Experimenten von Hans Comtet zusätzlich häufiger ihre Skepsis gegenüber 3D Gesten und dem Einfluss den diese auf ihren Alltag haben könnten (vgl. Com13, S.61).

3D Gesten leiden außerdem an erhöhter Ungenauigkeit, gerade auf größere Entfernungen und bei kleinen für die Interaktion vorgesehenen Objekten. Dies liegt an dem, bei einigen Menschen mehr bei Anderen weniger ausgeprägten, natürlichen Händezittern (vgl. JJBH15, S.457).

Ein weiteres großes Problem ist das Gorilla Arm Syndrome. Durch das langfristige Wegstrecken der Hände von dem restlichen Körper ohne jegliche Stütze treten schnell Müdigkeitserscheinungen in den Armen auf, die längeres interagieren auf diese Art und

Weise schwierig gestalten. Ein weiterer Grund für diese Ermüdungserscheinungen ist das fehlende haptische Feedback beim Gestikulieren in der Luft. Auch hier arbeitet man stetig an neuen Ansätzen, wie zum Beispiel an den Händen befestigbare Geräte, die durch leichten Druck in entsprechend passenden Momenten haptisches Feedback nachsimulieren sollen (KKC<sup>+15</sup>).

## 3 Analyse

Dieses Kapitel dokumentiert die Analyse des Systems, die einen wichtigen Bestandteil des Softwareentwicklungsprozesses darstellt. Dabei werden die an das System gestellten Anforderungen ermittelt, strukturiert und geprüft.

### 3.1 Szenarien

Im folgenden Abschnitt werden zwei realistische Szenarien vorgestellt, die eine optimale Möglichkeit bieten ein System mit einer 2D sowie 3D Steuerung zu verwenden und es daraufhin auf die Probe zu stellen. Weiter wird erläutert auf welches Szenario die endgültige Entscheidung fiel und was der ausschlaggebende Aspekt dafür war.

#### **Redaktion beim gemeinsamen Zusammenstellen der nächsten Zeitungsausgabe**

Ein Redaktionsteam aus drei bis fünf Mitarbeitern befindet sich in einem Konferenzsaal und bespricht gemeinsam das zukünftige Layout ihrer neuen Zeitungsausgabe. Als visuelles und funktionales Hilfsmittel steht ihnen ein 79 Zoll großes Display, auf dem ein Tool verwendbar ist, welches Ihnen ein leeres Zeitungslayout anbietet, zur Verfügung. Außerdem ist es möglich die Seitenanzahl und das Layout jeder einzelnen Seite anzupassen. Weiter wird ihnen von dem Tool ermöglicht, auf ihren Mobiltelefonen erreichbare, Artikel auf das Display zu schieben und mit ihnen zu interagieren. Man kann die hinzugefügten Artikel frei auf dem oder um das Zeitungslayout herum positionieren und die Seiten samt der bereits auf ihr positionierten Artikeln umblättern. Zu jedem Zeitpunkt besteht die Möglichkeit die entstandene Zeitung als PDF Datei zu exportieren und auszudrucken. Die Redaktionsmitarbeiter können sich beim Interagieren, mit den auf dem Display hinterlegten Artikeln, frei in einem gewissen Radius um die Frontseite des Displays

bewegen und je nach Entfernung zum Display mit Touchgesten oder gestikulieren in der Luft agieren. Ein solches Redaktionsteam könnte nun beispielsweise aus einem Chefredakteur bestehen, welcher einige Meter von dem Display entfernt steht, um einen Überblick über das Gesamtgeschehen auf der Oberfläche zu haben, Artikel für die weiter vorne stehenden Redakteure zu markieren, bestimmte Anweisungen weiter zu geben oder auch bereits grobe Arbeiten durch Gesten in der Luft selbst durchzuführen. Die vorne in Reichweite zum Bildschirm stehenden Redakteure können dementsprechend dem Chefredakteur Folge leisten und per Touchgesten feingranulare Verschiebungen, Skalierungen oder Ähnliches an den Artikeln vornehmen. Ein ähnliches Szenario, bei dem die Redaktion jedoch an einem digitalen Tabletop, statt herkömmlichen Display, arbeitet, wurde auch von Lorenz Barnkow in seiner Masterarbeit betrachtet. Schwerpunkte seiner Arbeit lagen auf dem Territorialverhalten, der Überwindung großer Distanzen sowie der Ausrichtung von Objekten an einem Tabletop als Gruppe. (vgl. Bar12)

#### **Projektteam beim gemeinsamen Erstellen einer Bildercollage**

Dieses Szenario ist dem Ersten in vielerlei Hinsicht sehr ähnlich. Auch hier findet sich eine Gruppe aus drei bis fünf Menschen zusammen, jedoch muss es sich in diesem Fall nicht zwingend um ein Redaktionsteam handeln, sondern es kann eine Gruppe von Arbeitskollegen, eine Gruppe von Freunden oder auch eine Familie zu Hause an ihrem Fernseher sein, welche gemeinsam eine Bildercollage erstellen möchte (siehe Abbildung 3.1). Genauso wie im vorherigen Szenario bietet ein auf einem 79 Zoll laufendes Tool alle zuvor beschriebenen Funktionen, ein veränderbares Layout, das Transferieren von Bildern von den Smartphones der Anwender auf das Display, die Interaktion und Transformation dieser auf dem Display liegenden Bilder, sowie das Umblättern und Exportieren der Bildercollage. Der große Unterschied zwischen den beiden Szenarien liegt hauptsächlich darin, dass das Erstellen einer Bildercollage eine breiter gefächerte Benutzergruppe anspricht und dass im Kontext mit Bildern statt Zeitungsartikeln die Transformation, also Skalierung und Rotation der Elemente mehr Sinn ergibt. Dies ist war der ausschlaggebende Grund weshalb die Entscheidung auf dieses Szenario fiel. Die Interaktionsmöglichkeiten mit den Bildern erlauben eine größere Anzahl an hilfreichen Gesten, deren Anwendung im Rahmen dieser Arbeit evaluiert werden kann.

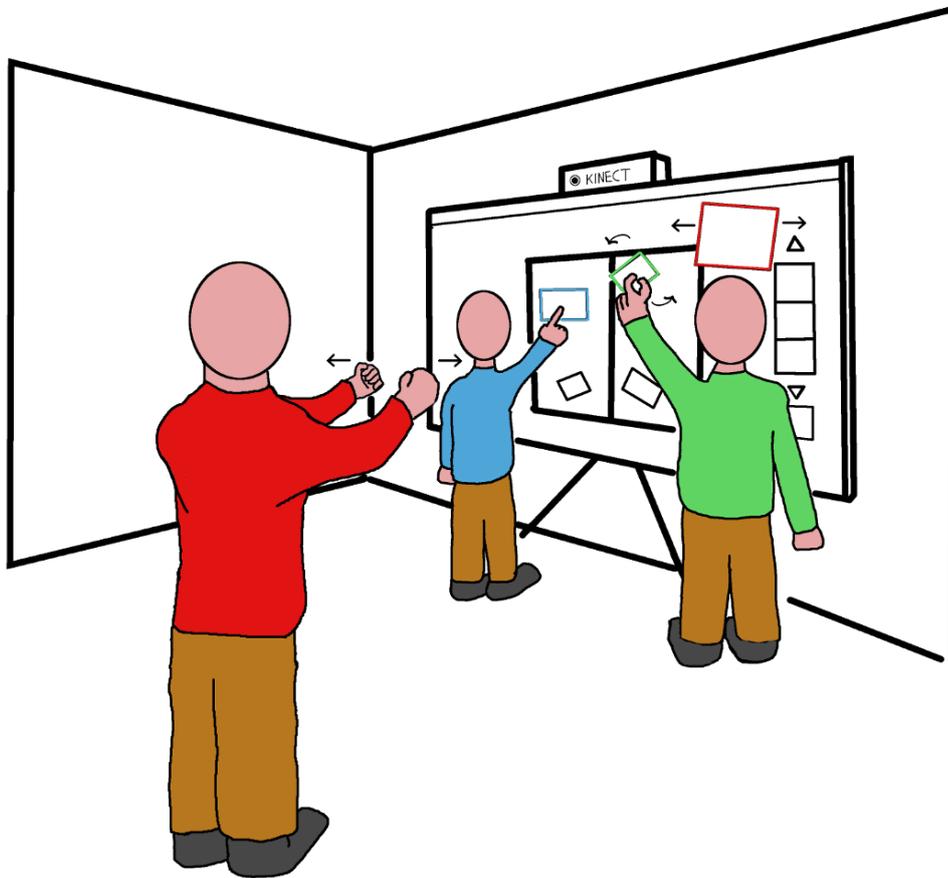


Abb. 3.1: Skizze eines mit dem Bildschirm interagierenden Teams. Quelle: eigene Darstellung

## 3.2 Anwendungsfälle

In den unten folgenden Tabellen (siehe Abbildung 3.2 und 3.3) sind, die sich aus dem zweiten Szenario ergebenden, Anwendungsfälle aufgelistet. Die Tabelle enthält die jeweilige Bezeichnung, eine knappe Beschreibung, die Akteure, die den Anwendungsfall jeweils auslösen können, sowie die für die Durchführung benötigten anderen Anwendungsfälle. Außerdem werden die Vor- und Nachbedingungen sowie die Invarianten der jeweiligen Anwendungsfälle aufgelistet, falls solche Bedingungen vorhanden sind.

### 3 Analyse

Name	Beschreibung	Akteure	verwendete Anwendungsfälle	Vorbedingungen	Invarianten	Nachbedingungen
Benutzerregistrierung	Ein Anwender muss sich mit parallel zum Boden von dem Körper weg ausgestreckten Armen registrieren, um sich als solcher kenntlich zu machen.	3D Anwender		Der Benutzer wird von der Tiefenbildkamera gesehen und erkannt.	Die Kamera verliert Benutzer nicht aus dem Bild.	Die Gesten des Benutzers werden interpretiert.
Seite umblättern	Der Anwender blättert von der jetzigen aufgeschlagenen Seite entweder eine Seite vor oder zurück.	2D Anwender, 3D Anwender		Für 3D Anwender: Der Benutzer muss registriert sein und die Hände müssen gut erkennbar sein		Die im Fokus liegenden Seiten ändern sich zu denen auf die geblättert wurde.
Bild/Text von Smartphone auf Display schieben	Der Anwender schiebt ein von seinem Smartphone aus erreichbares Bild oder einen kurzen Text auf das Display.	2D Anwender, 3D Anwender		Smartphone und Display müssen im selben Netzwerk sein.	Die Verbindung bricht nicht ab während das Bild hochgeladen wird.	Das neue Bild ist nun an erster Stelle im Bildslider.
Durch Bilderslider scrollen	Der Anwender scrollt entweder hoch oder herunter durch den Bilderslider.	2D Anwender, 3D Anwender		Es stehen mehr Bilder im Bilderslider zur Verfügung als auf einmal auf dem Display angezeigt werden Für 3D Anwender: Benutzer muss registriert sein und Hände gut erkennbar sein.		Die auf dem Display zu sehenden Bilder wurden je nachdem in welche Richtung gescrollt wurde durch die Nächsten oder Zuvorkommenden ersetzt und die vorher zu sehenden werden versteckt.
Bild dem Bilderslider entnehmen	der Anwender schiebt ein Bild zur weiteren Interaktion aus dem Slider heraus auf die restliche Benutzeroberfläche	2D Anwender, 3D Anwender	Bild umpositionieren	Es steht mindestens ein Bild im Bildslider zur Verfügung . Für 3D Anwender: Der Benutzer muss registriert sein und seine Hände müssen gut erkennbar sein.		Das Bild wird entnommen und alle im Bildslider weiter unten befindlichen Bilder werden eine Position höher geschoben.
Bild auf den Bilderslider schieben	der Anwender schiebt ein vorerst nicht mehr benötigtes Bild zurück auf den Bildslider	2D Anwender, 3D Anwender	Bild umpositionieren	Für 3D Anwender: Benutzer muss registriert sein und Hände gut erkennbar sein		Das Bild wird an vorderste Stelle zum Bildslider hinzugefügt und andere Bilder werden, falls vorhanden, eine Position herunter geschoben.
Bild löschen	Der Anwender schiebt ein nicht mehr benötigtes Bild in die „Deletion Area“ und entfernt es somit von dem Display.	2D Anwender, 3D Anwender	Bild umpositionieren	Es ist mindestens ein Bild auf dem Display vorhanden Für 3D Anwender: Der Benutzer muss registriert sein und seine Hände müssen gut erkennbar sein.		Das Bild wird endgültig von dem Display entfernt und kann nicht mehr verwendet werden.

Abb. 3.2: Die benötigten Anwendungsfälle Tabelle 1. Quelle: eigene Darstellung

### 3 Analyse

Name	Beschreibung	Akteure	verwendete Anwendungsfälle	Vorbedingungen	Invarianten	Nachbedingungen
Hilfscursor bewegen	Der Anwender bewegt durch seine Handposition einen Hilfscursor zur besseren Orientierung im Interaktionsbereich.	3D Anwender		Benutzer muss registriert sein und Hände gut erkennbar sein		Der Hilfscursor wird in Relation zur Handposition an der entsprechenden Stelle auf dem Display angezeigt.
Bild selektieren	Der Anwender selektiert ein bestimmtes Bild auf dem Display.	2D Anwender, 3D Anwender		Für 3D Anwender: Der Benutzer muss registriert sein und seine Hände müssen gut erkennbar sein.		Das Bild ist selektiert und wird entsprechend visuell gekennzeichnet.
Bild umpositionieren	Der Anwender schiebt ein auf dem Display befindliches Bild von seiner jetzigen Position auf eine neue Position.	2D Anwender, 3D Anwender		Für 3D Anwender: Der Benutzer muss registriert sein und seine Hände müssen gut erkennbar sein.		Das Bild wird an der neuen Position angezeigt.
Bild skalieren	Der Anwender vergrößert oder verkleinert ein bestimmtes Bild.	2D Anwender, 3D Anwender		Das Bild ist selektiert. Für 3D Anwender: Der Benutzer muss registriert sein und seine Hände müssen gut erkennbar sein.		Das Bild wird in der neuen Größe angezeigt.
Bild rotieren	Der Anwender rotiert ein bestimmtes Bild entweder links oder rechts herum um den Bildmittelpunkt.	2D Anwender, 3D Anwender		Bild selektiert Für 3D Anwender: Der Benutzer muss registriert sein und seine Hände müssen gut erkennbar sein.		Das Bild wird rotiert angezeigt.
Bildercollage exportieren/drucken	Der Anwender exportiert oder druckt den Ist-Zustand der Collage.	2D Anwender				Eine der Collage entsprechende PDF Datei oder bedruckte Zettel entstehen.
Seite hinzufügen	Der Anwender fügt der bisherigen Collage eine neue leere Seite hinzu.	2D Anwender				Die Collage hat eine weitere Seite.
Seite entfernen	Der Anwender entfernt die letzte Seite der Collage, dabei werden alle unter Umständen auf der Seite befindlichen Bilder zurück an den Anfang des Bildersliders geschoben.	2D Anwender				Die Collage hat eine Seite weniger und alle auf der letzten Seite befindlichen Bilder sind wieder im Bildslider.
Seitenhintergrund von Smartphone auf Display schieben	Der Anwender stellt dem Display ein von seinem Smartphone aus erreichbares Hintergrundbild als Auswahl zur Verfügung.	2D Anwender, 3D Anwender		Smartphone und Display müssen im selben Netzwerk sein.	Die Verbindung bricht nicht ab während das Bild hochgeladen wird.	Der neue Seitenhintergrund ist nun in der Auswahl an Hintergründen mit aufgelistet und kann ausgesucht werden.
Seitenhintergrund ändern	Der Anwender ändert den Seitenhintergrund der zu dem Zeitpunkt zu sehenden Seiten.	2D Anwender				Das Bild wird an der neuen Position angezeigt.

Abb. 3.3: Die benötigten Anwendungsfälle Tabelle 2. Quelle: eigene Darstellung

### 3.3 Anforderungsanalyse

Im folgenden Abschnitt werden die funktionalen sowie nicht funktionalen Anforderungen an die Software betrachtet. Diese ermöglichen es im weiteren Verlauf der Softwareentwicklung gezielt, auf ein den Bedürfnissen entsprechendes, System hinzuarbeiten.

### 3.3.1 funktionale Anforderungen

Die funktionalen Anforderungen werden durch die oben aufgelisteten Anwendungsfälle (siehe Abbildung 3.2 und 3.3) dargestellt. Das System muss dem Anwender ermöglichen jeden dieser Anwendungsfälle durchzuführen. Je nachdem ob der Anwendungsfall für nah oder weit vom Bildschirm entfernte Akteure gedacht ist, muss das System entsprechend eine Touch oder 3D Geste für die Ausführung bereitstellen. Das System soll theoretisch, so vielen Benutzern wie möglich, die parallele Interaktion mit dem Display ermöglichen. Praktisch wird die Anzahl durch die gewählte Hardware und die zur Verfügung stehenden Schnittstellen eingeschränkt (siehe Kapitel 4.3).

### 3.3.2 nicht funktionale Anforderungen

In diesem Abschnitt werden erwartete Randbedingungen und Qualitätseigenschaften des Systems definiert.

#### Bedienbarkeit

Die Gesten des Systems müssen so intuitiv wie möglich sein um eine optimale Bedienbarkeit zu ermöglichen und eine mögliche Alternative zu den gängigen Bedienmethoden wie beispielsweise der Maus und Tastatur zu bieten. Weiter muss bei den Gesten auf eine international vertretbare und universal verständliche Auswahl geachtet werden. Denn viele Gesten haben je nach Kultur, in der man sich aufhält, eine völlig unterschiedliche Bedeutung, wie bereits beschrieben wurde (siehe Kapitel 2.6.1). Auf die jeweils gewählten Gesten wird konkret im Kapitel Entwurf (siehe Kapitel 4) Bezug genommen.

Ein weiterer Aspekt der für die Bedienbarkeit eine wichtige Rolle spielt ist die grafische Benutzeroberfläche. Gerade bei größeren Anzeigegeräten und Anzeigegeräten, die für mehrere parallele Benutzer gedacht sind, muss auf bestimmte Dinge geachtet werden, die bei herkömmlichen Bildschirmen keine Relevanz haben. Nutsi und Koch (vgl. NK15) haben wichtige Leitlinien für diesen Anwendungsfall zusammengetragen:

- Man soll keine Audiowiedergabe verwenden, da dies zu Verwirrungen führen kann, wenn man nicht selber, sondern ein anderer paralleler Anwender, der Auslöser der Wiedergabe ist.

- Beziehe Benutzer, die nicht direkt mit dem Bildschirm interagieren, trotzdem mit ein und gebe ihnen einen Sinn.
- Ziehe immer die verschiedenen Anwendungszonen in Betracht und passe beispielsweise die Schriftgrößen und -arten entsprechend an, um die Lesbarkeit sicherzustellen.
- Falls persönliche Zonen auf dem Bildschirm eingerichtet werden, sollen sie kleinstmöglich gehalten werden, um die anderen Anwender nicht abzulenken oder ihren Inhalt zu verdecken.
- Überfülle nicht den gemeinsamen Bereich auf dem Bildschirm, halte ihn stattdessen immer sauber und entferne ungenutzte Objekte.
- Achte auf die Erreichbarkeit beim Platzieren von Knöpfen oder Ähnlichem.
- Überlasse einem Anwender nicht die Kontrolle darüber die gesamte Sicht für alle zu verändern.
- Biete allen Anwendern, gleichermaßen alle Funktionsmöglichkeiten an.
- Es wird außerdem empfohlen Contextmenüs zu verwenden, die überall aufklappbar sind, statt fest positionierten Menüs.

### **Performance**

Das System darf bei Ausführung der oben aufgeführten funktionalen Anforderungen keine spürbare Verzögerung aufweisen, da es sich bei den meisten Interaktionen um eine direkte Manipulation mit den, auf dem Display dargestellten, Objekten handelt und sich auch nur kleine Verzögerungen unnatürlich anfühlen und die Bedienbarkeit negativ beeinflussen. Das System kann relativ schnell den Eindruck erwecken abgestürzt zu sein und das Frustrationslevel der Anwender deutlich steigern (vgl. [CLB<sup>+</sup>04](#), S.5). Es ist also unabdingbar, dass das System in Echtzeit reagiert. Es müssen ausreichend Ressourcen bereit gestellt werden, sodass auch mehrere Benutzer parallel mit den dargestellten Objekten, ohne Ruckeln oder Verzögerung, interagieren können. Konkrete Richtlinien hierfür zu finden erweist sich als schwierig. Studien, die sich mit der Akzeptanz von Latenz bei Anwendung eines Touch Interfaces befassen, liefern relativ unterschiedliche Ergebnisse. Was sich jedoch festhalten lässt, ist, dass Menschen unterschiedliche Schwellenwerte diesbezüglich haben. Weiter können einige Menschen bereits einen spürbaren

Unterschied bei Latenzen unter 10ms feststellen, wirkliche Performanceeinbußungen können jedoch erst ab einer Latenz von 25ms aufwärts festgestellt werden. Die spürbare Verzögerung soll auch eine direkte Abhängigkeit mit der jeweiligen Geste haben, so liegt der Schwellenwert beispielsweise bei „Tap Gesten“ eher bei 40ms (vgl. [JNDW13](#), S.2298-2299). Die Akzeptanz der Latenz soll bei 3D Gesten laut Starner, mit 75-100ms, etwas höher liegen (vgl. [SLM+03](#), S.2). All diese Aspekte führen zu einem zu großen Aufwand bei der Berechnung und Bewertung der Latenz. In Folge dessen wird sich im Rahmen dieser Arbeit auf das eigene Gefühl und die Meinung befragter Kommilitonen bezüglich der Performance verlassen.

#### **Zuverlässigkeit und Korrektheit**

Die Gesten, die zur Interaktionen verwendet werden, müssen so zuverlässig und präzise wie möglich funktionieren, da auch hier ansonsten die Bedienbarkeit stark beeinträchtigt wäre. Ein Mensch hat eine gewisse Anpassungsfähigkeit und versucht bei Problemen die Ausführung seiner Geste mehrfach deutlicher darzustellen und sich den Anforderungen des Systems anzunähern, sobald es zu Bedienungsproblemen kommt. Doch auch diese Anpassungsfähigkeit hat seine Grenzen und sollte nicht überstrapaziert werden. Eine falsch erkannte Geste mehrfach auszuführen kann auf den Benutzer frustrierend wirken (vgl. [CLB+04](#), S.4) oder auch zu Missverständnissen oder ungewollten Falscheingaben führen. Es muss also darauf geachtet werden, dass alle verwendbaren Gesten stark voneinander differenziert werden und mögliche Gestenverwechslungen minimiert werden. Weiter muss auch darauf geachtet werden, dass eine beabsichtigte Geste überhaupt als solche erkannt wird, da das System sonst den Eindruck erwecken kann nicht ansprechbar zu sein. Gleichzeitig muss die Rate an Gesten, die unfreiwillig ausgeführt worden sind und erkannt werden so gering wie möglich gehalten werden. Dies ist eine große, wenn nicht unmögliche Herausforderung, der man sich nur mit kleinen Schritten annähern kann und für die es viele verschiedene Ansätze und Verfahren gibt. Eine sich positiv auf die Erkennungsrate auswirkende Technologie ist zum Beispiel die kompetitive Sensorfusion (siehe Abschnitt [Sensorfusion](#)).

#### **Aussehen**

Die Benutzeroberfläche muss so simpel und leicht zugänglich wie möglich aufgebaut sein, um dem Anwender die Orientierung zu erleichtern. Gerade bei Mehrbenutzeranwen-

dungen kann schnell ein gewisses Chaos auf der Benutzeroberfläche entstehen, da viele unvorhergesehene Interaktionen der anderen Teammitglieder parallel auf dem Bildschirm dargestellt werden. Eine weitere Hilfe, um diesem Problem entgegenzuwirken, wäre visuell darzustellen, welches Teammitglied wann für welche Informationsänderung auf dem Bildschirm verantwortlich ist, in welchem Bereich und mit welchen Informationen er gerade interagiert.

#### **Skalierbarkeit**

Das System soll die Möglichkeit bieten ohne enormen Aufwand weitere Gesten zu implementieren, um weitere vielleicht in der Zukunft notwendige Funktionen anstoßen zu können.

#### **Portierbarkeit**

Ein gewisser Grad an Portierbarkeit wäre erwünscht, um sich nicht vollkommen von einer bestimmten Hardware abhängig zu machen und die Anwendung auf jeglichen Betriebssystemen und Anzeigemedien ausführen zu können.

### **3.4 Laborumgebung**

Es werden ein Rechner und minimal zwei weitere Hardware Komponenten benötigt. Zum einen eine Hardware Komponente, welche uns ermöglicht 2D Gesten beziehungsweise Touch Input auszulesen und zum anderen eine Komponente, die in der Lage ist die Bewegungen, der Nutzer, vor dem Anzeigemedium zu erkennen. Verwendet wird ein Windows 10 Rechner, welcher eine Intel Core i7-5820K, 16GB RAM und eine Geforce GTX 980Ti verbaut hat. Als kollaborative Arbeitsplattform wurde ein 79 Zoll großer Multi-Touch Bildschirm (siehe Abschnitt [Benq](#)) ausgewählt. Die Informationen über die Bewegungen der Nutzer vor dem Bildschirm, die als 3D Gesten interpretiert werden sollen, werden von einer Kinect v2 (siehe Abschnitt [Kinect](#)) verfolgt. Die notwendige Hardware sowie der Arbeitsplatz werden von dem CSTI Labor<sup>1</sup> an der HAW (siehe Abbildung [3.4](#) und [3.5](#)) bereitgestellt.

---

<sup>1</sup><http://creative-space.haw-hamburg.de/>

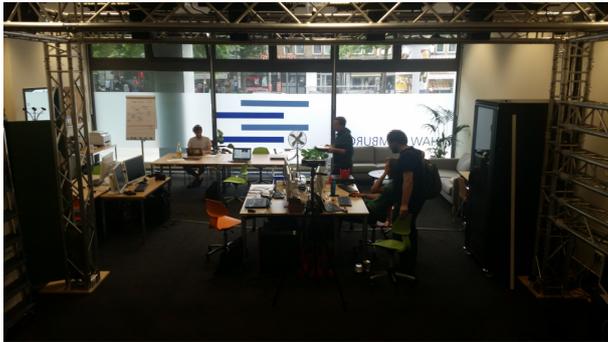


Abb. 3.4: Arbeitsplätze des CSTI Labors.  
Quelle: Eigene Darstellung

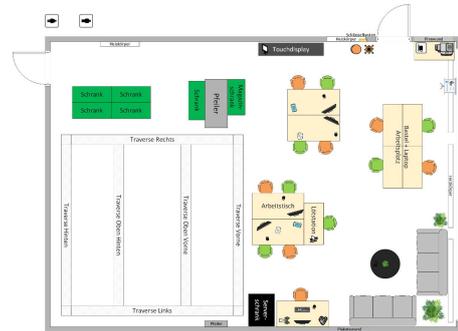


Abb. 3.5: Raumplan des CSTI Labors.  
Quelle: Eigene Darstellung

### Multitouch Display BenQ RP 790

Der BenQ RP 790 (siehe Abbildung 3.6) ist ein 79 Zoll großer Multi-Touch Bildschirm. Er umfasst also 200,7cm in der Diagonalen, die mit einer 4k (3840 x 2160) Auflösung ausgefüllt werden. Die hohe Auflösung sowie der Blickwinkel von 178° ermöglichen angenehmes interaktives Arbeiten aus jeglichen Positionen zum Bildschirm. Die IR-Touch Technologie ermöglicht 10 gleichzeitig registrierbare Berührungspunkte und somit paralleles kollaboratives Arbeiten mehrerer Anwender. Hierfür sind um das gesamte Display herum, Infrarot LEDs und Empfänger in den Rahmen eingelassen, die die Fläche direkt vor dem Display beleuchten. Wird nun ein oder mehrere Finger auf das Display gelegt, wird das Infrarot Licht an der jeweiligen Stelle unterbrochen und das Touch Panel kann durch das Rastermuster der Infrarot LEDs, die Koordinaten der Berührungspunkte ablesen (siehe Abbildung 3.7).

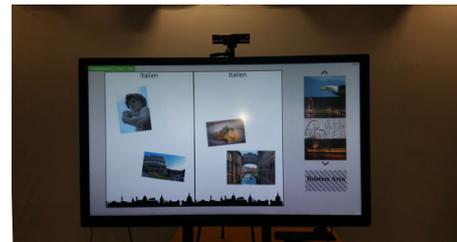


Abb. 3.6: BenQ RP790. Quelle: Eigene Darstellung

## Infrared Touch Panel

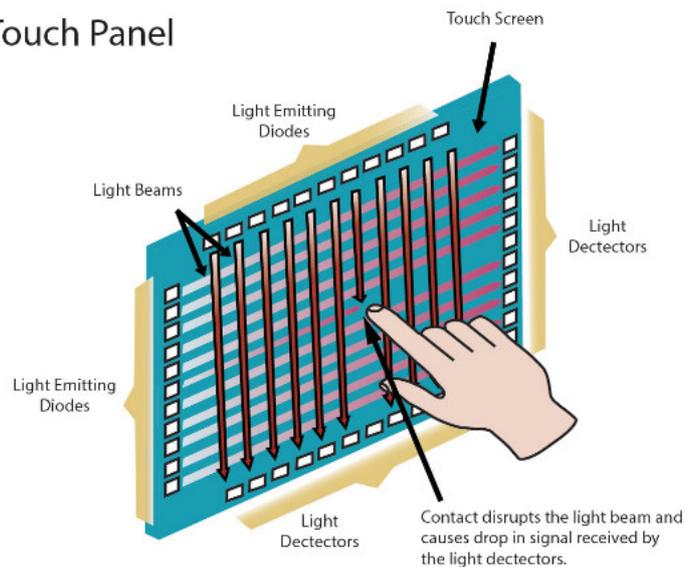


Abb. 3.7: Funktionsweise eines IR-Panels aus (Dix13)

## Kinect

Die Kinect v2 (siehe Abbildung 3.9) ist eine von Microsoft überwiegend für die Gestensteuerung entwickelte Kamera und der Nachfolger der Kinect. Die Kinect v2 enthält eine RGB Kamera, eine Tiefenkamera sowie ein Mikrofon Array. Die Tiefenkamera der Kinect v2 arbeitet nach dem time of flight Prinzip. Ein IR-Emitter sendet ein Infrarot Signal aus und die Kamera misst für jeden einzelnen der  $512 \times 484$  Pixel (siehe Abbildung 3.8) die Zeit die das Signal benötigt, um reflektiert an den Objekten wieder zurück zur Kamera zu gelangen. Da die Lichtgeschwindigkeit eine konstante Einheit ist, lässt sich daraus die Entfernung der jeweiligen Objekte ermitteln. Die Kinect kann auf diese Art und Weise

Feature	Kinect for Windows 1	Kinect for Windows 2
Color Camera	640 x 480 @30 fps	1920 x 1080 @30 fps
Depth Camera	320 x 240	512 x 424
Max Depth Distance	~4.5 M	~4.5 M
Min Depth Distance	40 cm in near mode	50 cm
Horizontal Field of View	57 degrees	70 degrees
Vertical Field of View	43 degrees	60 degrees
Tilt Motor	yes	no
Skeleton Joints Defined	20 joints	26 joints
Full Skeletons Tracked	2	6
USB Standard	2.0	3.0
Supported OS	Win 7, Win 8	Win 8
Price	\$299	TBD

Abb. 3.8: Kinect v2 Spezifikationen im Vergleich. Quelle: msdn.microsoft.com

ein 3D Modell der Welt vor der Kinect rekonstruieren und mithilfe weiterer Algorithmen Skelette der Anwender auf bis zu circa 4,5 Meter Entfernung erkennen.



Abb. 3.9: Kinect Sensoren. Quelle: msdn.microsoft.com

### 3.5 Fazit

Zusammengefasst muss die zu entwickelnde Software jegliche aus den Anforderungen ermittelte Funktionen bereitstellen. Sie muss Mehrbenutzertauglichkeit aufweisen und sämtliche Funktionen mit der 2D sowie 3D Gestensteuerung ausführbar machen. Es muss sich für ein intuitives und klar voneinander differenzierbares Gestenset entschieden werden, welches global akzeptierbar ist und nicht auf kulturelle Missverständnisse stößt. Weiter sollten die von Nutsi und Koch aufgestellten Leitlinien (siehe Kapitel 3.3.2) bestmöglich eingehalten werden. Das System sollte darüber hinaus performant, zuverlässig und korrekt reagieren um das Frustrationslevel der Anwender bei der Interaktion so gering wie möglich zu halten. Die Benutzeroberfläche sollte simpel gestaltet und in der Interaktion befindliche Objekte visuell abgehoben und den jeweiligen Akteuren zuordenbar sein. Weiter sollte das System leicht erweiterbar und portierbar sein, um eine flexible Anwendung zu ermöglichen. An Hardware werden ein Rechner, ein Smartphone, ein BenQ RP790 Display sowie eine Kinect benötigt.

## 4 Entwurf

Das folgende Kapitel stellt den nächsten Schritt des Softwareentwicklungsprozesses dar, in dem sich für eine Architektur entschieden wird, welche die zuvor entwickelten Anforderungen erfüllt. Die hier getroffenen Entscheidungen beeinflussen maßgeblich den Ablauf der Realierungsphase.

### 4.1 2D Gesten

Bei den Touchgesten (siehe Abbildung 4.1) wurde versucht sich an die gängigen Touchgesten, die sich in den letzten Jahren bereits auf Smartphones, Tablets und anderen Geräten durchgesetzt haben, zu halten. Der Leitgedanke dabei war, somit nicht allzu weit von dem wahrscheinlichsten Mental Modell der Anwender abzuweichen, sodass sie auch ohne Anweisungen schnell mit dem Bildschirm interagieren können sollten.

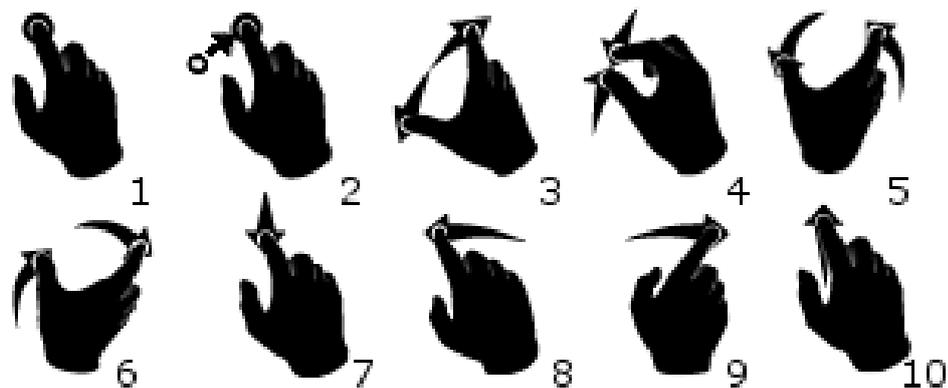


Abb. 4.1: Bildliche Darstellung aller Touchgesten. Quelle: Eigene Darstellung auf Grundlage der unter <http://www.mobiletuxedo.com/touch-gesture-icons/> zur freien Verfügbarkeit bereitgestellten Icons

Die Select Geste<sup>1</sup> lässt sich durch das Halten eines Fingers auf einem bestimmten Objekt auslösen. Durch die Geste wird das jeweilige Objekt als selektiertes Objekt markiert und ermöglicht weitere Interaktion mit diesem.

Wird ein Finger auf ein Objekt gelegt und daraufhin bewegt, bewegt sich das selektierte Objekt im selben Verhältnis auf der Benutzeroberfläche mit dem Finger mit. Diese Art der Geste wird auch als Drag Geste<sup>2</sup> bezeichnet. Wird der führende Finger wieder von der Benutzeroberfläche entfernt, wird auch das Objekt an der entsprechenden Stelle belassen. Dies wird auch als Drop bezeichnet.

Werden zwei Finger benutzt und beide jeweils an der Stelle an der ein Objekt auf der Benutzeroberfläche abgebildet ist aufgelegt, kann man daraufhin durch das Vergrößern des Abstandes zwischen den beiden interagierenden Fingern proportional das entsprechende Objekt vergrößern. Wird beispielsweise der anfängliche Abstand zwischen den Finger verdoppelt wird auch das skalierte Objekt auf die doppelten Maße vergrößert. Diese Geste wird auch als Spread Geste<sup>3</sup> bezeichnet.

Selbiger Vorgang, wie bei der Spread Geste ist auch in die entgegengesetzte Richtung möglich. Vergrößert man den Abstand zwischen den beiden auf dem Objekt liegenden Fingern nicht sondern verkleinert ihn, wird auch das Objekt zur gleichen Proportion verkleinert. Diesen Vorgang nennt man Pinch<sup>4</sup>.

Legt man jedoch zwei Finger auf die Abbildung eines Objektes ab ohne den Abstand zwischen diesen, sondern die Position dieser entweder im Uhrzeigersinn oder gegen den Uhrzeigersinn zu verändern, wird auch das Objekt entsprechend rotiert<sup>5-6</sup>.

Die Skalierung und Rotation eines Objektes wird in unserem Fall ohne erneutes Ansetzen der Finger möglich sein. Wenn sich der Abstand sowie die Position der Finger im Verhältnis zum Mittelpunkt beider Finger verändern sollten, werden beide Operationen parallel ausgeführt.

Bewegt man einen Finger während er den Bildschirm berührt mit einer gewissen Geschwindigkeit und über eine etwas größere Entfernung in eine der vier Himmelsrichtungen wird die Swipe Geste<sup>7-10</sup> in diese Richtung ausgelöst und ermöglicht somit das Umblättern von Seiten oder das Scrollen durch die Bildersammlung.

## 4.2 3D Gesten

Auch bei den 3D Gesten, also räumlichen Gesten die mit den Händen in der Luft ausgeübt werden, wurde versucht sich an dem Bewegungsablauf der gängigen Touchgesten zu orientieren und eine Analogie zu diesen zu bilden, um sich auch hier nicht zu weit von den erwarteten Mental Models der Anwender zu entfernen. Dadurch sollte eine leichtere Erlernbarkeit erreichbar und eine höhere Intuitivität ermöglicht sein. Die Unterschiede zwischen beiden Gestentypen liegen somit in der Oberfläche auf der sie ausgeführt werden und der Größe der Gesten beziehungsweise an der Ausführung beteiligten Körperteile. Während die Touchgesten auf der glatten Bildschirmoberfläche ausgeführt werden, werden 3D Gesten frei im Raum ausgeführt. Weiter werden nicht wie bei Touchgesten ausschließlich die Finger verwendet, sondern die gesamten Arme stellen die nötigen Bewegungsabläufe nach.

### 4.2.1 Registrierung

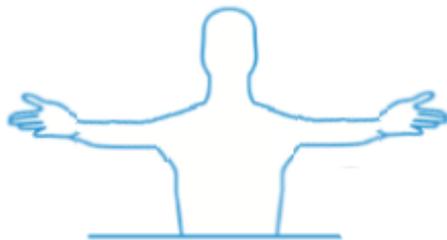


Abb. 4.2: Registrierung mit nach außen ausgestreckten Armen. Quelle: Eigene Darstellung

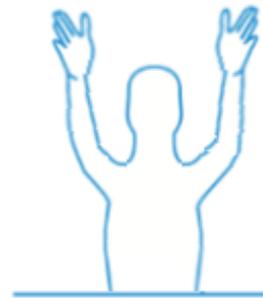


Abb. 4.3: Registrierung mit nach oben ausgestreckten Armen. Quelle: Eigene Darstellung

Die Kinect Kamera verfolgt alle in ihrem Sichtfeld erkannten Skelette und liefert die entsprechenden Daten. Um unbeabsichtigten Benutzerinput (siehe Kapitel 2.6.3) von Personen und fälschlicherweise erkannten Objekten herauszufiltern, müssen sich Benutzer, die mit 3D Gesten interagieren wollen vorerst über eine spezielle Geste registrieren. Diese Geste sollte im Alltag eher selten Anwendung finden und gestaltet somit eine versehentliche Registrierung als unwahrscheinlich. Zunächst fiel die Entscheidung auf die

T-Pose (siehe Abbildung 4.2), bei der der Anwender für kurze Zeit beide Arme parallel von dem Körper auf Schulterhöhe weg strecken muss. Doch nach ersten Versuchen mit einer größeren Gruppe musste festgestellt werden, dass der Platz den die Kinect vor dem Bildschirm auswertet zu beschränkt ist und das Ausbreiten der Arme zu den Seiten erwies sich als hinderlich. Stattdessen müssen die Arme für die Registrierung nun senkrecht nach oben, über den Kopf, angehoben werden (siehe Abbildung 4.3). Beim Ausführen dieser Geste wird auch die Armlänge errechnet und für die in (siehe Kapitel 4.2.2) erwähnten Maße der imaginären Flächen verwendet, um somit sicherzustellen, dass von jeglichen Benutzern unabhängig von der Armlänge, jeder Punkt der Fläche erreichbar ist. Erst nach erfolgreicher Ausführung dieser Geste werden weitere Gesten des Skelettes erkannt und weitergeleitet.

### 4.2.2 Orientierung auf dem Display

Das erste große Hindernis ist es dem Nutzer zu ermöglichen bestimmte Objekte auf der Bildschirmoberfläche für die weitere Interaktion zu selektieren ohne diese tatsächlich zu berühren. Es gibt mehrere Ansätze diese Aufgabe umzusetzen.

#### Raycasting

Eine der am häufigsten verwendeten Methoden ist zum einen Raycasting (vgl. CQG<sup>+11</sup>) (vgl. AA09) bei der ursprünglich ein Sensor an der Hand angebracht wurde der einen Strahl in Richtung Bildschirm schießt. Diese Technik wäre prinzipiell auch mit den 3D Kamera Informationen einer Kinect möglich.

Dabei zieht man eine imaginäre Linie von den Augen des Anwenders über den zeigenden Finger bis hin zur Bildschirmoberfläche (siehe Abbildung 4.4). Dort wo der Strahl die Bildschirmoberfläche schneidet ist der Punkt, den der Benutzer selektieren möchte. Diese Methode führt jedoch gewisse Probleme mit sich. Zum einen kann einem die Kinect von Haus aus nicht die genaue Position der Augen, sondern nur des Kopfmittelpunktes liefern. Weiter ist auch die Fingererkennung der Kinect bisher

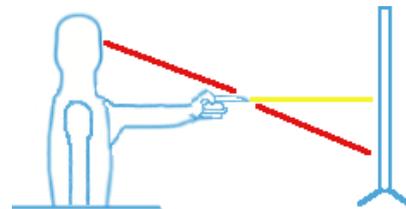


Abb. 4.4: Raycasting - rote Linie zeigt tatsächlich selektierten Punkt auf dem Display. Quelle: Eigene Darstellung

noch nicht genau genug, um feingranulare Bewegungen wie sie bei der Selektion von Nöten wären auszuführen. Außerdem muss bei der Methode zunächst kalibriert werden wo sich die Bildschirmoberfläche im Raum und in Relation zum Anwender befindet. Dafür hätte man den Benutzer beispielsweise auf einige vordefinierte Punkte zeigen lassen können. Eine andere Möglichkeit wäre eine Installation mit mehreren Kameras, die über Sensorfusion verbunden sind, wobei eine in Richtung des Bildschirms zeigt und diesen als solchen erkennt und somit die Position dessen bereitstellen kann, aufzubauen.

### Selektieren durch Blickrichtung

Eine weitere Methode ist das Selektieren ausschließlich anhand der Augenbewegungen. Zum einen ist diese Methode bereits unabhängig von technischen Ungenauigkeiten durch das natürliche Augenzittern sehr ungenau zum anderen werden so ungewollt, durch das immerwährende Verfolgen der Augenbewegungen, Funktionen ausgelöst (vgl. [VTA<sup>+</sup>15](#), S.4).

### imaginäre Fläche

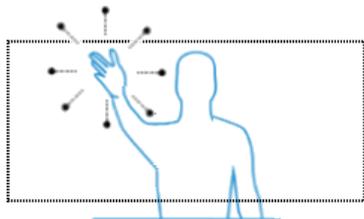


Abb. 4.5: imaginäre Fläche - frontal.  
Quelle: Eigene Darstellung

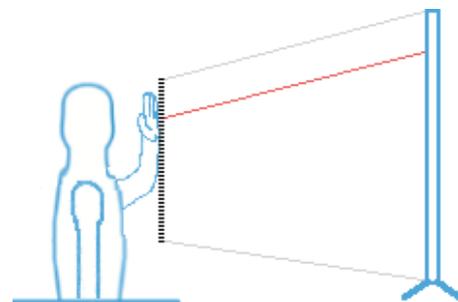


Abb. 4.6: imaginäre Fläche - seitlich  
Quelle: Eigene Darstellung

Schlussendlich wurde sich für eine Variante mit einer imaginären Fläche vor dem Benutzer (siehe Abbildung [4.5](#) und [4.6](#)) entschieden. Bei dieser Methode fällt zum einen der Kalibrierungsschritt weg, zum anderen wirkt das natürliche Händezittern bei dieser Methode weniger stark als bei dem Raycasting, da die Hände nicht so weit von dem Körper weggestreckt werden müssen, also stabiler sind und Ermüdungserscheinungen

dementsprechend auch langsamer eintreten. Vor dem Oberkörper des Anwenders wird eine imaginäre Fläche aufgestellt, deren Dimensionen so eingestellt sind, dass der Anwender mit seiner Armlänge an jeden Punkt dieser Fläche gelangen kann. Das heißt die Diagonale von dem oberen linken Eckpunkt bis zum unteren rechten Eckpunkt darf nicht länger sein als die Reichweite des Anwenders von einer nach links ausgestreckten Hand bis zu einer nach rechts ausgestreckten Hand. Je nachdem, wo der Anwender nun seine Hand vor seinem Oberkörper positioniert, wird diese Position auf die entsprechende Position auf dem Bildschirm umgerechnet und abgebildet. Diese Variante wurde im Verlauf der Implementation insoweit modifiziert, dass nun zwei kleinere Flächen statt einer großen Fläche verwendet wird (siehe Abbildung 4.7). Die Gründe für diese Änderung werden im Kapitel 5.2.5 weiter erläutert.

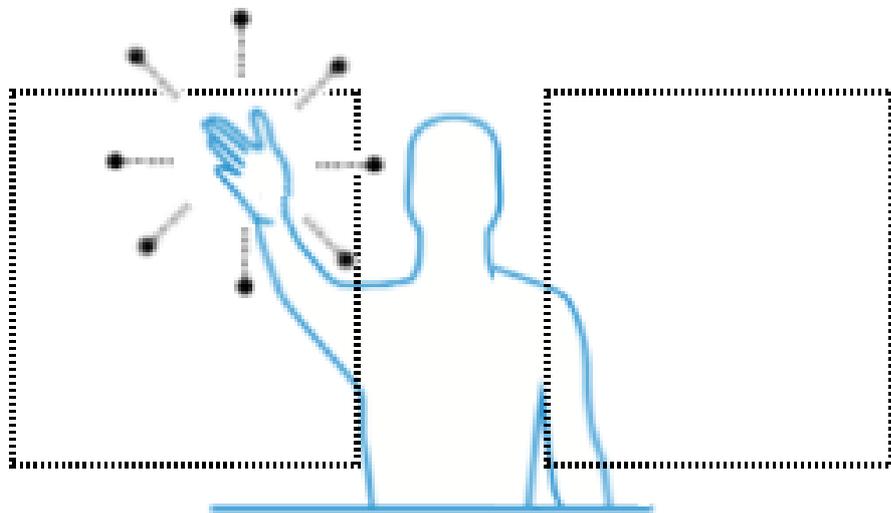


Abb. 4.7: Zwei voneinander unabhängige imaginäre Flächen vor dem Anwender deren Koordinaten sich auf die Koordinaten des Bildschirms abbilden lassen. Quelle: Eigene Darstellung

Die zwei Flächen sind so positioniert, dass ihr vertikaler Mittelpunkt jeweils auf Schulterhöhe ist. Eine der senkrechten Seiten verläuft ebenfalls an der Schulter, während die entgegengesetzte senkrechte Seite auf Höhe der fast völlig ausgestreckten Hand verläuft. Auch bei dieser Variante wird jeder Punkt im Koordinatensystem der Flächen entsprechend proportional auf die Koordinaten des Bildschirms abgebildet. Bewegt der Anwender seine Hand an eine Position auf der imaginären Fläche und hält sie für kurze Zeit so ruhig wie möglich wird der entsprechende Punkt auf dem Bildschirm selektiert.

### 4.2.3 Cursor bewegen

Um einen Cursor zu bewegen und sich anhand des visuellen Feedbacks auf der Bildschirmoberfläche zu orientieren, werden ebenfalls die imaginären Flächen angewandt. Anders als bei der Selektion wird hier die Handfläche nicht ruhig auf einen bestimmten Punkt der Fläche gelegt, sondern der Benutzer ist in der Lage mit geöffneter Hand Bewegungen über diese Flächen auszuführen (siehe Abbildung 4.8 und 4.9), die dann auf den Bildschirm abgebildet werden. Wichtig ist hierbei, dass nur eine Hand zur Zeit vor dem Körper positioniert ist, da durch das Hinzuziehen der zweiten Hand Objekttransformationen ausgeführt werden (siehe Kapitel 4.2.4).

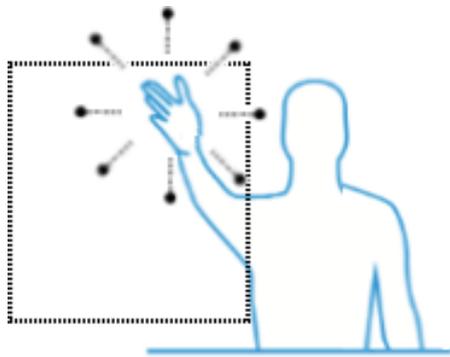


Abb. 4.8: Nutzer bewegt Cursor mit rechter Hand. Quelle: Eigene Darstellung

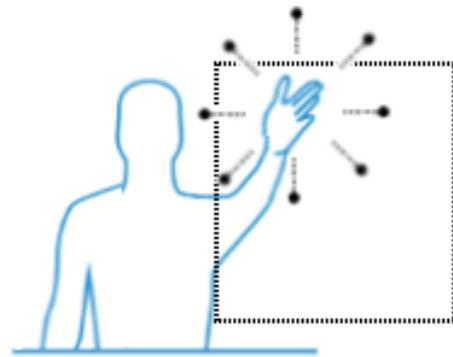


Abb. 4.9: Nutzer bewegt Cursor mit linker Hand. Quelle: Eigene Darstellung

### 4.2.4 Objektverschiebung

Ist der Cursor über einem Objekt positioniert worden, kann man es, durch schließen der Hand, ähnlich einer Greifbewegung, in Bewegung versetzen. Bewegt man die geschlossene Hand, genauso wie bei dem Bewegen des Cursors, auf der imaginären Fläche (siehe Abbildung 4.10 und 4.10) bewegt man im selben Verhältnis auch das selektierte Objekt. Auch hier muss darauf geachtet werden nur eine Hand gleichzeitig vor dem Körper zu führen.

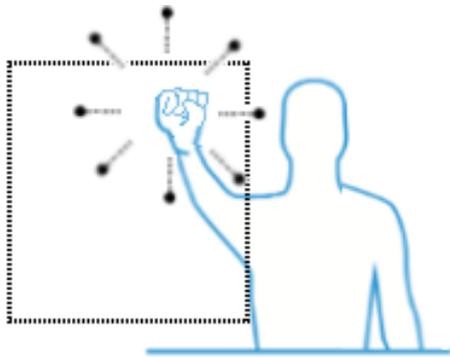


Abb. 4.10: Nutzer bewegt Objekt mit geschlossener rechter Hand. Quelle: Eigene Darstellung

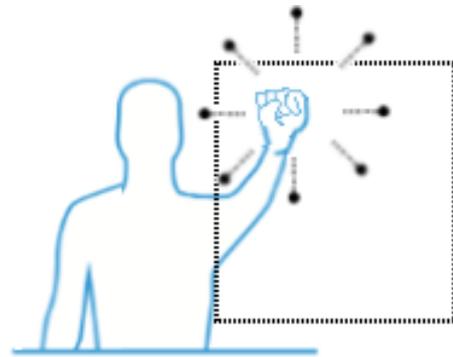


Abb. 4.11: Nutzer bewegt Objekt mit geschlossener linker Hand. Quelle: Eigene Darstellung

### Objekttransformation

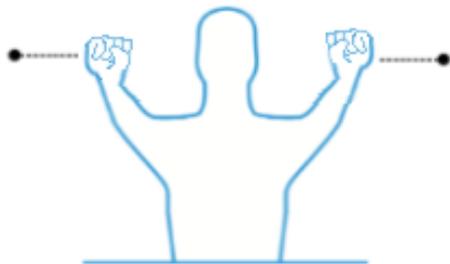


Abb. 4.12: Nutzer vergrößert Objekt mit Spread Geste. Quelle: Eigene Darstellung

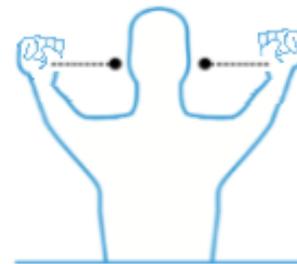


Abb. 4.13: Nutzer verkleinert Objekt mit Shrink Geste. Quelle: Eigene Darstellung

Zieht der Anwender nun die zweite Hand hinzu und positioniert diese ungefähr gleich weit entfernt vom Körper, ist er im Objekttransformationsmodus. Wenn zuvor ein Objekt selektiert wurde, kann er es nun durch das Auseinanderziehen der Hände (siehe Abbildung 4.12) vergrößern oder durch das Zusammenziehen der Hände (siehe Abbildung 4.13) verkleinern. Die Größe des Objektes ändert sich hier, wie auch bei der 2D Variante, proportional zur Handentfernung. Wird der Abstand zwischen den Händen im Verhältnis zur Ursprungsentfernung verdoppelt, wird auch das selektierte und in Transformation befindliche Objekt auf die doppelte Größe skaliert.

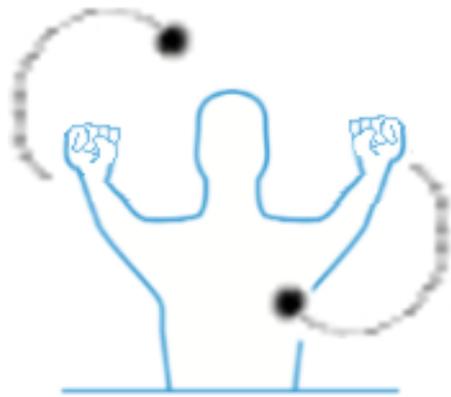


Abb. 4.14: Nutzer rotiert Objekt gegen den Uhrzeigersinn. Quelle: Eigene Darstellung

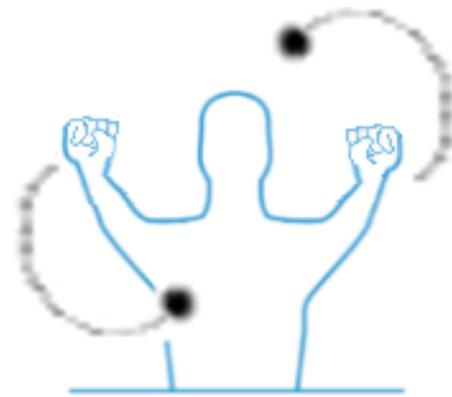


Abb. 4.15: Nutzer rotiert Objekt mit dem Uhrzeigersinn. Quelle: Eigene Darstellung

Rotiert man die Hände wiederum im (siehe Abbildung 4.15) oder gegen den Uhrzeigersinn (siehe Abbildung 4.14) wird das Objekt entsprechend im oder gegen den Uhrzeigersinn rotiert.

## Swipe

Eine Swipebewegung lässt sich auch hier ausführen, indem nur eine Hand nach vorne von dem Körper weggestreckt und geschlossen mit einer gewissen Geschwindigkeit in eine der vier Himmelsrichtungen bewegt wird (siehe Abbildung 4.16). Führt man eine Swipe Bewegung in eine horizontale Richtung aus wird das Magazin in die entsprechende Richtung umgeblättert. Swiped der Benutzer wiederum in vertikale Richtung wird der Slider auf der rechten Seite der Bildschirmoberfläche um ein Bild aufwärts oder abwärts je nach Richtung bewegt und ermöglicht, dadurch das Durchsuchen der auf den Bildschirm hochgeladenen Inhalte. Auch diese Geste ist beliebig mit der linken oder rechten Hand ausführbar.

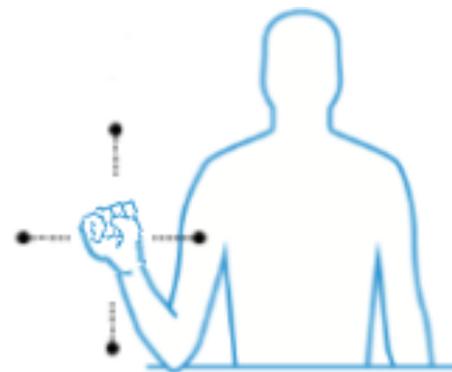


Abb. 4.16: Nutzer swiped in eine der vier Himmelsrichtungen. Quelle: Eigene Darstellung

### 4.3 Benutzeranzahl

Die Benutzeranzahl wird in unserem Fall zum einen durch das Skeleton Tracking, der verwendeten Kinect v2, auf sechs gleichzeitig verfolgbare Skelette eingeschränkt und zum anderen durch den BenQ RP790 Multitouch Bildschirm mit zehn zeitgleichen Berührungspunkten. Da Touchgesten mit einem und auch mit zwei benötigten Fingern implementiert werden sollen, sind, abhängig von der gewählten Touchgeste, zwischen fünf und zehn gleichzeitig interagierende Benutzer möglich. Dabei gehen wir davon aus, dass die Benutzer jeweils mit nur einem Objekt zur Zeit interagieren. Versucht ein Benutzer, jedoch anders als angenommen, parallel mit mehreren Objekten zu interagieren, kann er bis zu zehn Berührungspunkte verwenden und die maximale Anzahl an Touchgesten-Anwendern sinkt auf einen. Um sicherzustellen, dass es während der Verwendung nicht zu Komplikationen durch die flexible maximale Benutzeranzahl kommt, sollte ein festes Limit von fünf parallelen Benutzern gesetzt werden. Während der ersten Experimente wurde jedoch beobachtet, dass die Kinect v2 zwar sechs komplette Skelette verfolgen kann, jedoch nur die Handzustände der jeweils zwei am nächsten stehenden Skelette verfolgt. Da die Handzustände (offen/geschlossen) jedoch für das greifen und wieder loslassen der Objekte notwendig ist, wird die Benutzeranzahl der 3D Gesten-Anwender auf zwei zur Zeit eingeschränkt. Möchte ein anderer Anwender mit 3D Gesten interagieren muss er einen Schritt vor gehen oder einer der zuvor in Interaktion befindlichen Anwender einen Schritt zurück gehen. Theoretisch ist das System dennoch in der Lage bis zu sechs über 3D Gesten interagierende Benutzer und bis zu zehn mit jeweils einem Touchpunkt interagierende Benutzer zu verarbeiten. Ansätze um die Handzustände von mehr als zwei Anwendern parallel abzufragen werden im Kapitel Ausblick (siehe Abschnitt [Handzustände](#)) angerissen.

### 4.4 Benutzeroberfläche

Es wird versucht sich an so viele Empfehlungen bezüglich der Bedienbarkeit (siehe Kapitel [3.3.2](#)) zu halten wie möglich. Auf Audiowiedergabe wird gänzlich verzichtet. Die Schriftgrößen werden so gewählt, dass auf maximale Interaktionsreichweite, die von der Kinect begrenzt ist, noch alles leserlich und erkennbar ist. Auf persönliche Zonen für die Anwender wird verzichtet stattdessen hat jeder Anwender die Möglichkeit flexibel die gesamte Bildschirmoberfläche zu verwenden. Durch die typische Platzierung des

Menüs an dem oberen Bildschirmrand ist auch dessen Erreichbarkeit und Auffindbarkeit weitestgehend gesichert. Auf Kontextmenüs wird in dem Fall verzichtet, da versucht wird die Steuerung mittels Touchgesten so simpel wie möglich zu halten und Korrelationen mit versehentlichen Kontextmenüaufrufen zu verhindern. Die Menüpunkte sollen stattdessen jederzeit sichtbar und leicht verständlich sein. Alle, ob via 2D oder 3D interagierende, Anwender können den vollen Funktionsumfang verwenden dadurch wird keine der Steuerungstechniken durch einen höheren Umfang vorgezogen. Einzige Ausnahme ist hierbei das am oberen Bildschirmrand angelegte Menü. Dieses lässt sich ausschließlich über Touchgesten bedienen, was die Experimente kaum beeinflussen sollte, da der Fokus hierbei auf der Objektinteraktion und nicht der Menüführung liegen wird. Leider hat auch jeder Benutzer die volle Kontrolle über die Sicht auf die Benutzeroberfläche. Dies lässt sich nicht verhindern, da zum einen durch das Magazin geblättert werden muss und zum anderen durch den Bilder und Text Slider gescrollt werden muss und somit die zu sehenden Bilder angepasst werden.

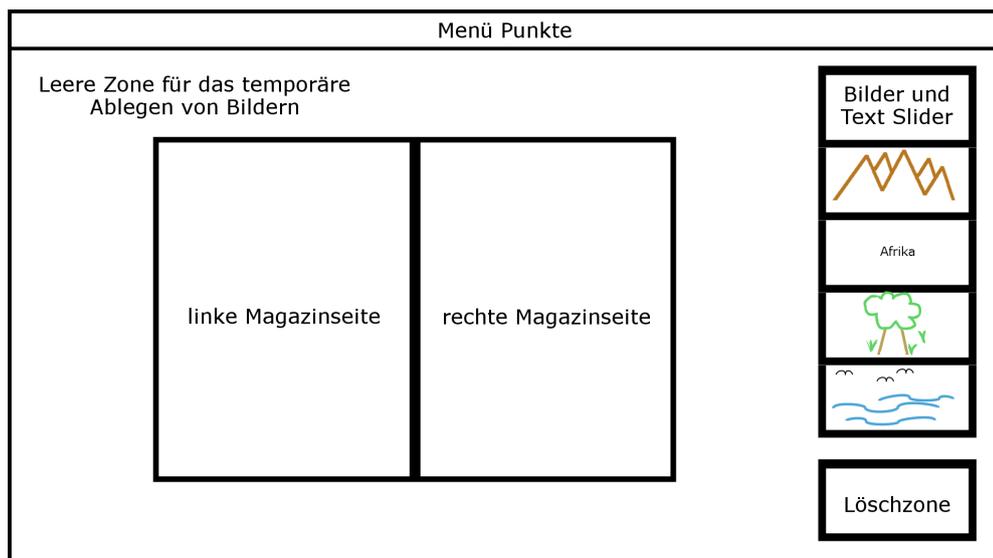


Abb. 4.17: GUI mit allen Elementen und Zonen. Quelle: Eigene Darstellung

In der oberen Grafik (siehe Abbildung 4.17) wird eine Skizze der geplanten Benutzeroberfläche dargestellt, die den Aufbau veranschaulichen soll. Am oberen Bildschirmrand wird leicht ersichtlich das Menü mit zusätzlichen Funktionen, wie das Verändern der Hintergrundbilder oder das Hinzufügen und Entfernen von Seiten des Magazins, aufzufinden sein. Die Navigation wird hierbei sehr schlicht gehalten und auf Unterpunkte

wird weitestgehend verzichtet. Zentral auf der Oberfläche wird das leere umblätterbare Magazin liegen. Der leere Bereich um das Magazin dient zur vorübergehenden Zwischenablage von Bildern und Texten, während man beispielsweise blättern oder andere Bilder und Texte zuerst auf einer Seite arrangieren möchte.

Auf der rechten Seite befindet sich ein Slider für Bilder und Texte, den man durch vertikale swipe Gesten in Bewegung versetzen und somit durch die Inhalte rotieren kann. Bilder lassen sich aus dem Slider per Touch, oder auch eine räumliche Drag Geste, entnehmen und genauso wieder zurück auf den Slider legen. Senkrecht darunter befindet sich eine kleine Zone, die Bilder und Texte welche in ihr losgelassen werden löscht.

Auch die zweite GUI (siehe Abbildung 4.18), die für das Hinzufügen von Elementen auf den Slider benötigt wird, wird sehr schlicht gehalten. Sie besteht nur aus drei farblich getrennten und eindeutig beschrifteten Knöpfen, sowie einem kleinen Eingabefeld für kurze Texte, die mittig auf der Oberfläche positioniert werden. Diese Knöpfe ermöglichen das Hinzufügen von Bildern, Hintergrundbildern und Texten.

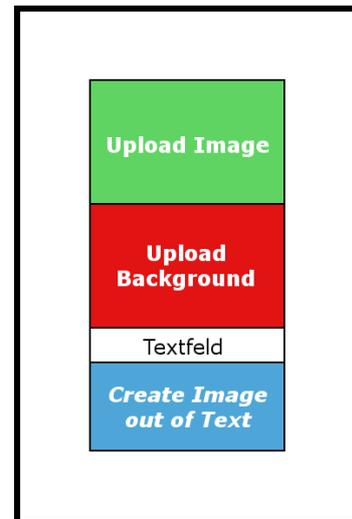


Abb. 4.18: GUI zum Hinzufügen von Elementen.  
Quelle: Eigene Darstellung

## 4.5 Architektur

Im unteren Diagramm (siehe Abbildung 4.19) ist der Aufbau der einzelnen Systemkomponenten, sowie ihr Zusammenspiel abgebildet. Informationen erhält das System über Smartphones oder Tablets, den BenQ RP790 sowie eine Kinect v2. Über das Smartphone können Bilder ausgewählt oder Texte, die in SVGs umgewandelt werden, eingegeben werden, die daraufhin auf einen Server hochgeladen, und von der GUI aus einsehbar sind. Der BenQ Bildschirm liefert die Koordinaten aller Touchpunkte und die Kinect die aktuellen Bildinformationen und Metadaten, der sich vor dem Bildschirm abspielenden Ereignisse. Ein von Tobias Eichler<sup>1</sup> in Scala entwickelter Skeleton Detection Agent verarbeitet die Bildinformationen und liefert die entsprechenden Skelettdaten weiter. Diese Skelettdaten werden dann von einer 3D Gesture Detection Komponente

---

<sup>1</sup>Zweitgutachter und wissenschaftlicher Mitarbeiter im CSTI Labor

weiter verarbeitet und entsprechend passende Gesten werden an eine Gruppe Gestures veröffentlicht. Der Skeleton Detection Agent hält jedoch keine Informationen über die Hand Posen. Da die Information, ob die Hände der Benutzer geöffnet oder geschlossen sind, auch benötigt werden, um eine Art Greifgeste erkennen zu können, benötigen wir eine weitere Komponente, die im Diagramm (siehe Abbildung 4.19) als Hand State Detection bezeichnet wird. Diese schickt die Informationen, der Kinect C# SDK, über die aufgezeichneten Hand Posen an den 3D Gesture Detection Agenten weiter. Mithilfe der Skelettdaten sowie Hand Posen lassen sich nun alle nötigen 3D Gesten erkennen. Die GUI zeigt dann in Echtzeit die Auswirkungen der an die Gestures Gruppe veröffentlichten 3D Gesten sowie 2D Gesten an.

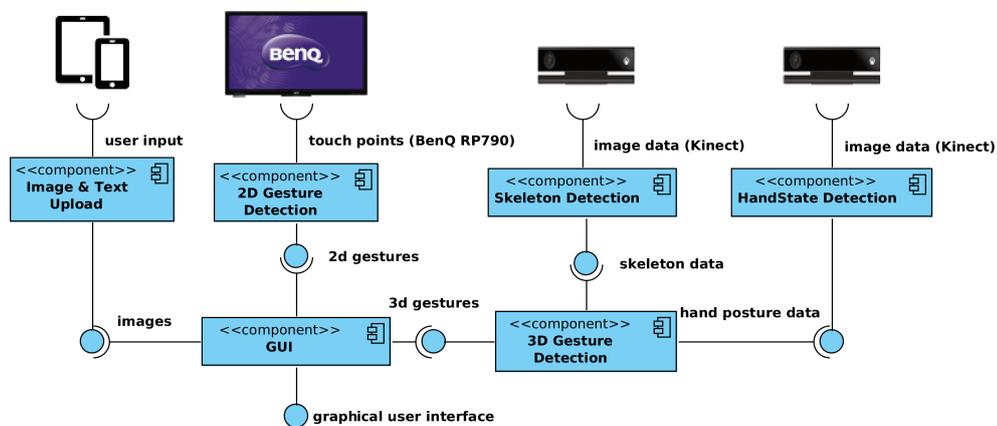


Abb. 4.19: Diagramm zeigt den Aufbau der einzelnen Systemkomponenten

## Lose Kopplung & starke Kohäsion

Ein großer Vorteil dieser komponentenbasierten Architektur ist die lose Kopplung der einzelnen Systembausteine sowie der hohe Grad an Kohäsion. Jede Komponente hat ihren klar definierten Aufgabenteil und Änderungen des Systems beschränken sich meist auf ausschließlich eine Komponente. Die vorgenommenen Änderungen haben, bei Einhaltung der definierten Schnittstellen zwischen den Komponenten, auch keine Auswirkungen auf die anderen Komponenten. Bei der Definition der Schnittstellen und dem Nachrichtenaustausch nach dem Publish & Subscribe Prinzip unterstützt die im CSTI Labor<sup>1</sup> verwendete Middleware (vgl. Eic14). Ein positiver Nebeneffekt der starken Kohäsion ist,

<sup>1</sup><http://creative-space.haw-hamburg.de/>

dass Code-Duplizierungen vermieden und das DRY-Prinzip weitestgehend eingehalten wird. Weiter ermöglicht die Architektur auch Komponenten gänzlich auszutauschen oder einzelne Komponenten in anderen Projekten wiederzuverwenden. So könnte die Handstate Komponente darüber hinaus beispielsweise Anwendung in einem im Living Place <sup>2</sup> ansässigen Szenario und dem dafür zuständigen System finden.

### 4.6 Fazit

Der aufgestellte Entwurf sollte den Anforderungen der Analyse (siehe Kapitel 3) bestmöglich gerecht werden. Die funktionalen Anforderungen werden durch 2D sowie 3D Gesten umsetzbar gemacht. Bei den 2D Gesten (siehe Kapitel 4.1) wird versucht sich an die gängigen Touchgesten, wie man sie bereits von jeglichen Smartphones und Tablets kennt, zu halten. Die 3D Gesten (siehe Kapitel 4.2) sind von dem Bewegungsablauf stark an die bekannten Touchgesten angelehnt. Der einzige große Unterschied ist, dass im Gegensatz zu Touchgesten, bei denen ausschließlich die Finger des Anwenders zum Einsatz kommen, die gesamten Arme teil der Gestenbewegung darstellen. Die Idee hinter den ausgewählten Gesten ist, dass sie so intuitiver erscheinen und leichter erlernbar sein sollten.

Die Gesten sind klar voneinander differenzierbar, sodass bei korrekter Ausführung eine gewisse Korrektheit und Zuverlässigkeit des Systems gegeben sein sollte. Bei dem Aussehen der Benutzeroberfläche und seiner Bedienbarkeit wurde großer Wert auf die von Nutsi und Koch (siehe Kapitel 3.3.2) aufgestellten Empfehlungen gelegt. Das System hat keine Audiowiedergabe, die zu Verwirrungen führen könnte, Schriftgrößen und Symbole sind so gewählt, dass sie bis zur maximalen Interaktionsreichweite klar lesbar sind und das Menü ist typischerweise am oberen Bildschirmrand. Auf persönliche Zonen wird gänzlich verzichtet, sodass jeder Anwender Kontrolle über die gesamte Interaktionsfläche hat. Weiter fanden ebenfalls Kontextmenüs keine Verwendung, um so die Gesten so simpel wie möglich zu halten und keine versehentlichen Kontextmenüaufrufe so provozieren, die dann im schlimmsten Fall noch den Interaktionsbereich der anderen Teilnehmer blockieren könnten.

Die Interaktionsfläche ist sauber strukturiert und übersichtlich gestaltet (siehe Abbildung 4.17). Es existieren keine überlappenden Elemente und, bis auf eine Ausnahme, auch keine Menüunterpunkte. Der Anwender hat somit von Anfang an einen Überblick über

---

<sup>2</sup><http://livingplace.informatik.haw-hamburg.de>

den vollen Funktionsumfang des Systems.

Die komponentenbasierte Architektur sowie die schmalen Schnittstellen sorgen für eine lose Kopplung und starke Kohäsion (siehe Abbildung 4.19). Die Architektur ermöglicht hierdurch eine einfache Skalierbarkeit, da man durch das Erweitern der richtigen Schnittstellen und eine Änderung in der jeweils zuständigen Komponente simpel neue Gesten oder Ähnliches hinzufügen kann. Auch die Portierbarkeit ist, dadurch dass man jederzeit bestimmte Komponenten austauschen kann, gewährleistet, solange die Schnittstellen eingehalten werden. So ist man weder fest an eine spezifische Hardware noch Software gebunden. Bei der Realisierung wird man außerdem auf die eingesetzten Algorithmen achten müssen, um die Performanz des Systems nicht negativ zu beeinflussen und das Gesamtgefühl bei der Interaktion aufrecht zu erhalten.

# 5 Evaluation

## 5.1 Implementierbarkeit

Wie bereits in der Themenbeschreibung angerissen wird als Teil dieser Arbeit eine Software entwickelt, die der Evaluation der Kombination von Gesten, anhand eines bestimmten praxisnahen Beispiels, dient. Dennoch werden auch Probleme und Erkenntnisse, auf die im Laufe der Umsetzung gestoßen wurde, erläutert, sowie einige alternative Umsetzungsmöglichkeiten und Verbesserungsvorschläge angesprochen. Das folgende Kapitel beschreibt die Implementierbarkeit anhand des Ablaufs der Realisierung und der Funktionsweise der einzelnen Softwarekomponenten und verwendeten Libraries.

### Von dem Entwurf zur Implementierung

Als erstes sollte das Herzstück der Arbeit, also die 2D sowie 3D Gestensteuerung implementiert werden. Zunächst wurde sich mit den Informationen, die die Kinect und der Skelett Detection Agent liefert vertraut gemacht und es wurden erste 3D Gesten implementiert. Daraufhin wurden Touch Events implementiert. Um die Funktionstüchtigkeit der Gesten zu testen wurde anfänglich ein geometrisches Objekt auf einer leeren Oberfläche verwendet mit dem interagiert werden konnte. Sobald die Steuerung zufriedenstellend realisiert worden war, wurde die GUI Stück für Stück erweitert. Als erstes wurde mithilfe von turnJS das zentral auf der GUI positionierte umblätterbare Buch realisiert, welches ermöglichte Objekte auf den einzelnen Seiten abzulegen und sie wieder zu entfernen. Der nächste Schritt war es den mit slickJS implementierten Slider hinzuzufügen und dafür zu sorgen, dass Elemente, welche über ein Smartphone hochgeladen werden, in diesem Slider angezeigt werden. Die Elemente des Sliders müssen flexibel entfernbar und wieder hinzufügbare sein. Daraufhin wurde die Löscher Zone implementiert, die Elemente welche auf ihr abgelegt werden wieder von der GUI entfernt. Das letzte noch fehlende GUI Element war die Menüleiste. Es wurde eine Funktion zum

Ändern der einzelnen Buchseitenhintergründe sowie zum Hinzufügen und Entfernen von Buchseiten implementiert. Nach der GUI musste nur noch die Upload-GUI realisiert und ein Webserver mit allen nötigen PHP Funktionen aufgesetzt werden um so zu ermöglichen, dass von den Nutzern ausgewählte Daten auf der Interaktions-GUI zur Verfügung gestellt werden können. Zum Schluss musste dann noch an allen Ecken und Kanten etwas geschliffen werden, um die Gesten und das Gesamtgefühl so intuitiv wie möglich zu gestalten.

### **Skeleton Detection Agent**

Der Skeleton Detection Agent ist ein von Tobias Eichler<sup>1</sup> in Scala programmierter Agent, welcher auf Basis der J4K Java Library<sup>2</sup> aufgebaut ist und alle entsprechend gelieferten Informationen der Kinect an eine Gruppe nach dem Publish & Subscribe Prinzip überträgt und sie so für andere Agenten bereitstellt.

### **3D Gesture Detection Agent**

Das im Rahmen dieser Arbeit entstehende Programm, soll auf der Middleware (**Eic14**) des CSTI Labors<sup>3</sup> laufen und auf den Skelettdaten des Skelett Detection Agenten von Tobias Eichler<sup>1</sup> arbeiten können. Diese beiden Voraussetzungen führten zu dem Entschluss auch den 3D Gesture Detection Agenten in Scala zu programmieren und dabei auf die agentenorientierte Programmierung zurückzugreifen. Die agentenorientierte Programmierung bietet sich hier außerdem an, da die 3D Gesture Detection Komponente Daten mithilfe von Publish & Subscribe empfängt, verarbeitet und darauf aufbauende Informationen wie erwünscht für die GUI wieder zur Verfügung stellen kann.

### **Hand State Detection**

Die Handstate Detection Komponente wird wiederum in C# programmiert, da sie auf den von der Kinect C# SDK bereitgestellten Informationen arbeitet. Diese werden via UDP Pakete an die 3D Gesture Detection Komponente gesendet. Da die Skelett und Handposen Informationen getrennt gesendet werden, müssen die Handposen sobald sie

---

<sup>1</sup>Zweitgutachter und wissenschaftlicher Mitarbeiter im CSTI Labor

<sup>2</sup><http://research.dwi.ufl.edu/ufdw/j4k/>

<sup>3</sup><http://creative-space.haw-hamburg.de/>

von dem 3D Gesture Detection Agenten erhalten worden sind auf die Hände der jeweils richtigen Skelette anhand ihrer Koordinaten abgebildet werden.

### GUI Komponente

Die GUI (siehe Abbildung 5.1) wird in HTML5 aufgebaut, da somit eine hohe Portierbarkeit bei der Entwicklung, sowie der späteren Anwendung sichergestellt werden kann. HTML5 Oberflächen können auf jedem System, ohne zusätzlich notwendige Softwareinstallationen, angezeigt werden, solange es über einen Browser verfügt. So lässt sich die GUI auf einem Webserver abspeichern und auf jeglichen Geräten anzeigen. Ein weiterer Vorteil ist, dass Javascript, eine ursprünglich ausschließlich für HTML entwickelte Sprache, Touch Events interpretieren kann. So kann man auch die 2D Gestenerkennung simpel ohne zusätzlich benötigte Schnittstellen implementieren. Auch die Schnittstelle zwischen dem 3D Gesture Detection Agenten und der GUI lässt sich mithilfe eines Websockets, der die an die Gesture Gruppe gepublizierten Informationen an die GUI sendet, realisieren.

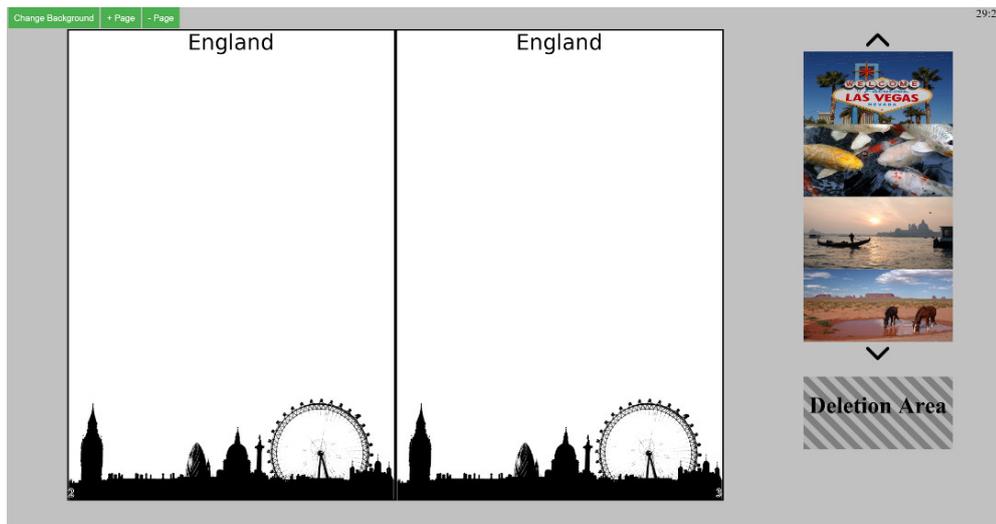


Abb. 5.1: Screenshot der GUI. Quelle: Eigene Darstellung

### Image & Text Upload Komponente

Die GUI der Image und Text Upload Komponente (siehe Abbildung 5.2) wird ebenfalls in HTML5 entwickelt und sendet die ausgewählten Objekte an einen PHP Server, der

die für die GUI benötigten Bilder, Texte und Hintergrundbilder beinhaltet und auf Abruf bereitstellen kann.

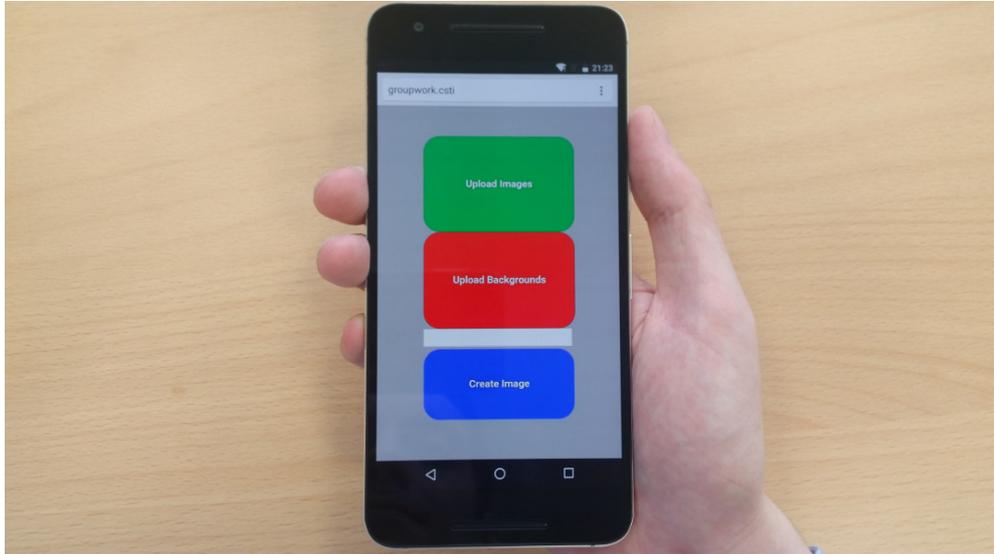


Abb. 5.2: Screenshot der Upload-GUI. Quelle: Eigene Darstellung

### **InteractJS**

InteractJS<sup>1</sup> ist eine JavaScript Library, welche das Ziehen und Ablegen, Vergrößern, Rotieren von Objekten, die Trägheit und das physikalisch realistische Verhalten von Objekten ermöglicht. Mithilfe dieser Library wurde die Arbeit mit 2D Gesten und ihre Auswirkungen auf die Bildobjekte umgesetzt.

### **TurnJS**

TurnJS<sup>2</sup> ist eine Library in Javascript, die das Erstellen realistisch aussehender Umblättereffekte für virtuelle Zeitschriften oder Bücher ermöglicht.

---

<sup>1</sup><http://interactjs.io/>

<sup>2</sup><http://www.turnjs.com/>

## SlickJS

SlickJS<sup>3</sup> ermöglicht die einfache Integration des in dieser GUI verbauten Sliders, der die Bilder und Texte hält.

In diesem Kapitel wird der Aufbau und der Ablauf der durchgeführten Experimente beschrieben. Die entwickelte Anwendung wird im Rahmen einer Laboruntersuchung durch mehrere Gruppen an Testpersonen erprobt. Daraufhin werden die dabei entstandenen Beobachtungen während der Experimente sowie Befragungen der Testpersonen in schriftlicher und mündlicher Form ausgewertet. Aus den erhobenen Daten werden entsprechende Schlüsse gezogen.

## 5.2 Eignungsuntersuchung

Zunächst wird die Eignungsuntersuchung strukturiert beschrieben, bevor in den darauf folgenden Abschnitten die erhobenen Daten ausgewertet und analysiert werden.

### 5.2.1 Aufbau

Das Testszenario, für die Eignungsuntersuchung von 2D sowie 3D Gesten im Zusammenhang mit kollaborativer Arbeit, wird im Creative Space for Technical Innovations (CSTI)<sup>1</sup> Labor der HAW aufgebaut und durchgeführt. Hierfür wird der BenQ RP 790 Multi-Touch Monitor mit der auf ihr montierten Kinect unter ein in dem Labor existierendes quadratisches Traversensystem mit einer Größe von 5x5m Innenmaß gestellt (siehe Abbildung 5.3 und 5.4). Alle Seiten des Traversensystems können beliebig mit weißen Planen verhängt werden und somit zum einen den Anschein eines kleinen Konferenzraumes verstärken und zum anderen visuelle Artefakte bei der Datenauswertung der Kinect senken, die sich negativ auf die Versuchsdurchführung auswirken könnten. Der gesamte Versuchsablauf wird visuell von zwei in der Traverse auf einer Höhe von 2,7m montierten Mako G503C Kameras und akustisch von dem Mikrofon Array der Kinect selbst aufgezeichnet. Die Mako G503C liefert 14 Bilder pro Sekunde in einer Auflösung von 2592x1944. Diese ermöglichen die Versuche in Ruhe mehrfach aus verschiedenen Perspektiven mit anzusehen und unter Umständen Auffälligkeiten zu erkennen, die

---

<sup>3</sup><http://kenwheeler.github.io/slick/>

<sup>1</sup><http://creative-space.haw-hamburg.de/>

einem ansonsten entgangen wären. Dabei ist eine Kinect frontal auf die Testpersonen ausgerichtet (siehe Abbildung 5.5), um die genaue Gestik und Mimik der Personen einzufangen, während die zweite Kamera auf der anderen Seite hinter den Testpersonen im Traversensystem montiert ist. Diese Kamera zeichnet somit unter anderem auf was sich parallel auf der Benutzeroberfläche des Monitors abspielt.

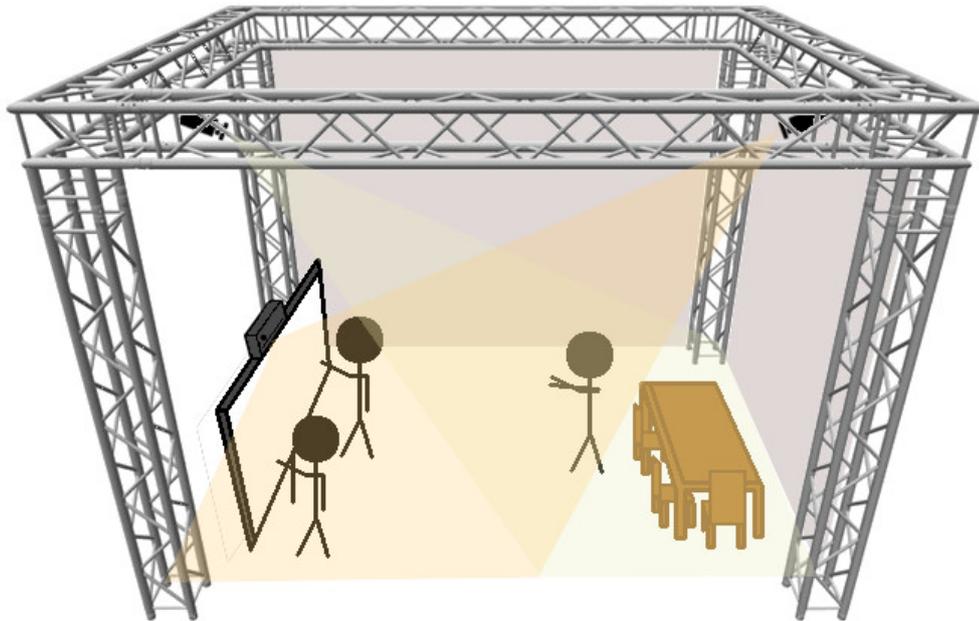


Abb. 5.3: Skizzierter Aufbau der Testumgebung. Quelle: Eigene Darstellung

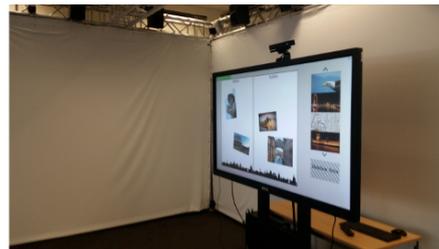


Abb. 5.4: Aufbau der Testumgebung. Quelle: Eigene Darstellung

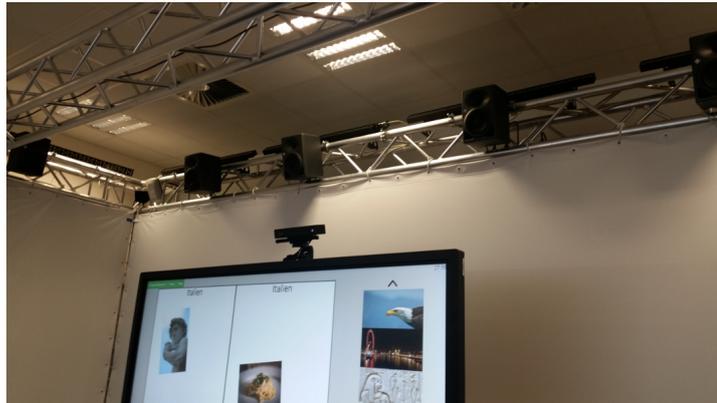


Abb. 5.5: Kamera in der Traverse. Quelle: Eigene Darstellung

### 5.2.2 Testgruppen

Ein Testdurchlauf erfordert jeweils eine Gruppe von drei Personen (siehe Abbildung 5.6). Der Grund hierfür ist, dass drei Personen noch nebeneinander an den BenQ RP 790 passen und einigermaßen effektiv zeitgleich mit Touchgesten interagieren können (siehe Abbildung 5.7). Würde man den Versuch mit größeren Gruppen durchführen, würde die Wahrscheinlichkeit erheblich steigen, dass ein Teil der Testpersonen überhaupt nicht aktiv an den Interaktionen teilnehmen könnte. Außerdem wurde davon ausgegangen, dass bei steigender Mitgliederanzahl die Absprache innerhalb einer Gruppe deutlich sinkt. Um Kommunikationsschwierigkeiten innerhalb der Gruppe weitestgehend zu vermeiden wurde explizit nach Dreiergruppen Ausschau gehalten, die sich bereits vor dem Experiment kannten. Ob es eine Gruppe von Freunden, Kommilitonen oder Familienmitgliedern war spielte in diesem Fall keine Rolle.



Abb. 5.6: Testgruppe diskutiert gemeinsam welche Bilder sie hochladen möchten.  
Quelle: Eigene Darstellung



Abb. 5.7: Testgruppe befüllt erste Seite der Collage mit 2D Gesten. Quelle: Eigene Darstellung

### 5.2.3 Aufgabenstellung

Die Aufgabe der drei Testpersonen besteht darin gemeinsam eine Bildercollage eines fiktiven Urlaubs zusammenzustellen. Die hierfür notwendigen Bilder sind teils bereits auf dem Display im rechts positionierten Slider, teils noch auf zwei für diesen Versuch vorbereiteten Smartphones auffindbar. Die Collage besteht aus einem Titelblatt sowie jeweils drei Doppelseiten mit einer bestimmten Thematik. Anhand der Überschrift sowie dem allgemeinen Layout der Doppelseiten soll deutlich gemacht werden, welche Urlaubsbilder jeweils auf welche Seite passen könnten. Jede dieser Doppelseiten soll nun auf eine andere Art und Weise von den Testgruppen befüllt werden. Die erste Seite soll ausschließlich anhand von Touchgesten zusammengestellt werden. Die zweite Seite wird daraufhin ausschließlich mithilfe von 3D Gesten mit Bildern bestückt und auf der letzten Seite kann mit beiden Varianten interagiert werden. So sollen die Testpersonen den direkten Vergleich zwischen Touch und 3D Gesten sowie das Gefühl beide Varianten gleichzeitig zur Verfügung gestellt zu haben, erfahren. Ab Beginn des Experiments läuft auf der Benutzeroberfläche ein 30 minütiger Timer ab, der die Testgruppe unter eine Art zeitlichen Druck setzen soll. Dieser soll die Auswahl für 3D Gesten aus purem Interesse an der Technik reduzieren und sie im Zweifelsfall dazu motivieren die ihrer Meinung nach besser passende Art der Interaktion zu wählen.

### 5.2.4 Ablauf

Zunächst werden die Testpersonen begrüßt und darum gebeten den zweiseitigen Fragebogen (siehe Kapitel 7.1) mit Fragen zu ihren demographischen Merkmalen sowie Vorkenntnissen die Einfluss auf das Experiment haben können in Ruhe auszufüllen.

Daraufhin bekommen die Testpersonen eine Einführung in das System und dessen Steuerung. Sie bekommen auch die Möglichkeit sich an die Steuerung zu gewöhnen und alle dazugehörenden Elemente auszuprobieren. Weiter wird ihnen die auf sie zukommende Aufgabe erläutert.

Als nächstes haben die Testgruppen jeweils maximal 30 Minuten Zeit die Aufgabe auszuführen (siehe Abbildung 5.8 und 5.9) und sich ein eigenes Bild über das System, seine Vor- und Nachteile sowie ihre eigene Meinung dazu zu machen.



Abb. 5.8: Testgruppe befüllt Collage mit 3D Gesten. Quelle: Eigene Darstellung



Abb. 5.9: Testgruppe befüllt Collage mit 3D Gesten. Quelle: Eigene Darstellung

Danach werden die Testpersonen ein weiteres mal gebeten einen zweiseitigen Fragebogen (siehe Kapitel 7.2) auszufüllen. In diesem Bogen werden ihnen nun bestimmte Fragen zu einzelnen Aspekten des Experiments gestellt und die Möglichkeit gegeben in einem Freitext weitere Auffälligkeiten oder sonstige Punkte anzumerken.

Zu guter Letzt folgt noch eine kurze Diskussion in der die Testpersonen frei miteinander über das Experiment und ihre Meinung zu 2D und 3D Gesten im Kontext kollaborativer Arbeit preisgeben können. Durch die Dynamik einer Diskussion können wichtige und interessante Punkte aufgegriffen werden, die in dem Fragebogen so nicht erfragt worden sind.

### 5.2.5 Pretest

Vor der tatsächlichen Eignungsuntersuchung wurde ein Pretest mit zwei bis drei Personen durchgeführt, um bereits hierbei aufspürbare Schwierigkeiten zu beseitigen. In diesem Abschnitt werden nach dem Pretest getroffene Änderungen und Abweichungen von dem Entwurf aufgelistet und begründet.

### Vereinigung der Skalierungs- und Rotationsgeste

Ursprünglich sollten die Skalierungs- und Rotationsgesten in 2D sowie 3D strikt voneinander getrennt und als eigenständige Gesten erkannt werden. Es war jedoch nicht möglich mit dieser Trennung eine flüssige Transformation der Objekte auszuführen. Oftmals wies eine Skalierungsgeste bei der Ausführung auch leichte Rotationen und

umgekehrt auf. Dadurch wurden abwechselnd kleine ruckartige Transformationen beider Typen ausgeführt. Erst durch das Zusammenführen beider Gestentypen und das Übertragen aller Metadaten, die Bestandteil dessen sind, also des Rotationswinkels, der Richtung, sowie des Skalierungskoeffizienten und der parallelen Ausführung beider Transformationen ergab sich ein natürliches Gefühl bei der Interaktion.

### **Anzahl der imaginären Flächen**

Die Anzahl der Flächen wurde von einer großen Fläche, die die Reichweite beider Arme des Benutzers umfasst und ihre Koordinaten auf die Bildschirmoberfläche abbilden lässt, auf zwei kleinere logisch getrennte Flächen erhöht. Der Grund für diese Entscheidung ist, dass die Kinect Handpositionen vor dem eigenen Körper schlechter erkennt, als Hände, die an den Körperseiten im leeren Raum gehalten werden. Diese Beobachtung hat sich durch starkes Zittern bei dem Bewegen des Cursors im mittleren Bereich des Bildschirms bemerkbar gemacht. Der zentrale Bereich ist jedoch der Bereich, in dem die meisten Interaktionen stattfinden, da dort das umblätterbare Magazin, welches mit Elementen bestückt werden soll, positioniert ist. Durch das Auslagern des zentralen Bereichs auf die zwei getrennten Flächen an den Körperseiten wird die Zone vor dem Körper nicht mehr für feingranulare Selektierungsarbeiten benötigt. Ein weiterer Vorteil ist, dass auf diese Art und Weise jeder beliebige Punkt auf dem Monitor über eine Hand erreichbar ist, ohne die Hand quer über den Körper legen zu müssen oder mitten in der Bewegung die führende Hand auszutauschen, um die entgegengesetzte Bildschirmseite zu erreichen. Als positiver Nebenaspekt werden dadurch resultierende Ermüdungserscheinungen der Arme merklich eingedämmt. Zum einen sind kürzere Handbewegungen zurückzulegen und zum anderen kann beliebig die Hand gewechselt werden.

### **5.3 Ergebnisdarstellung**

In diesem Abschnitt wird zunächst die Herangehensweise beim Aufbereiten und Visualisieren der zusammengeführten Daten erläutert. Daraufhin werden die Daten entsprechend analysiert und hilfreiche Schlüsse gezogen. In einem Fazit werden die wichtigsten Erkenntnisse noch einmal kompakt zusammengetragen.

### 5.3.1 Datenaufbereitung

Sowohl der erste Fragebogen mit den Fragen zu dem demografischen Hintergrund und den jeweiligen Vorkenntnissen, als auch der zweite Fragebogen, der die Meinung der Testperson zu dem Thema 2D und 3D Gesten bei Kollaboration erfassen soll, besteht überwiegend aus standardisierten Fragen, die Entscheidungen oder Bewertungen einholen sollen. Bei den Bewertungen wird hierbei auf ein Likert-Skala ähnliches Schulnotensystem von eins bis sechs gesetzt. Bei den Fragen handelt es sich um Aussagen, wie das exemplarische Beispiel „Das Wetter heute ist...“, die von den Testpersonen durch einen Haken in der Skala von 1 („sehr gut“) bis 6 („sehr schlecht“) ergänzt werden. Jede Frage bietet bei Bedarf auch die Möglichkeit sich kommentarlos zu enthalten. Der zweite Fragebogen bietet zum Schluss einen teilstandardisierten Bereich für einen Freitext zur Thematik an.

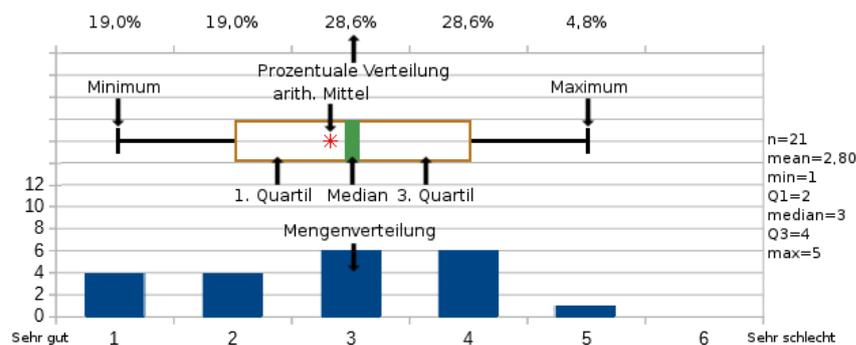


Abb. 5.10: Antwortenverteilung: 1. "Das Wetter heute ist..."(Beispieldiagramm). Quelle: Eigene Darstellung

Bei der Aufbereitung der Daten werden das arithmetische Mittel, der Median, das Minimum und Maximum sowie das erste und dritte Quartil berechnet und zur besseren Veranschaulichung in entsprechenden Boxplot-Diagrammen (siehe Diagrammbeispiel 5.10) dargestellt. Diese sind jeweils in ein herkömmliches Säulendiagramm eingebettet, welches die Verteilung der Antworten in Höhe der entsprechenden Balken präzise visualisiert. An der sekundären X Achse ist die Verteilung der Antworten zusätzlich in prozentualer Form veranschaulicht. Die entsprechende Fragestellung lässt sich dem Titel der Diagramme entnehmen und die beiden Pole der Bewertungsskala sind unter der X Achse auf der linken sowie rechten Seite zu finden. Die Antworten der Entschei-

dungsfragen werden im Gegensatz zu den Bewertungsfragen in einem segmentierten Kreisdiagramm veranschaulicht.

### 5.3.2 Demografischer Hintergrund

An den Versuchen nahmen insgesamt 21 Personen, unterteilt in 7 Gruppen, teil. 12 dieser Teilnehmer (61,9%) waren zum Zeitpunkt der Experimente zwischen 21 und 25 Jahre alt. Dementsprechend liegt auch das Durchschnittsalter bei „Mitte-Ende-Zwanzig“. Das Alter der anderen 9 Teilnehmer (38,1%) liegt verstreut zwischen 19 und über 46 Jahren (siehe Abbildung 5.11). Unter den Teilnehmern befanden sich 9 weiblichen Geschlechts (42,3%) und 12 männlichen Geschlechts (57,7%) (siehe Abbildung 5.12). Durch die recht gleichmäßige Verteilung lässt sich ein besserer Blick auf die Meinungsunterschiede zu den Fragestellungen je nach Geschlecht werfen.

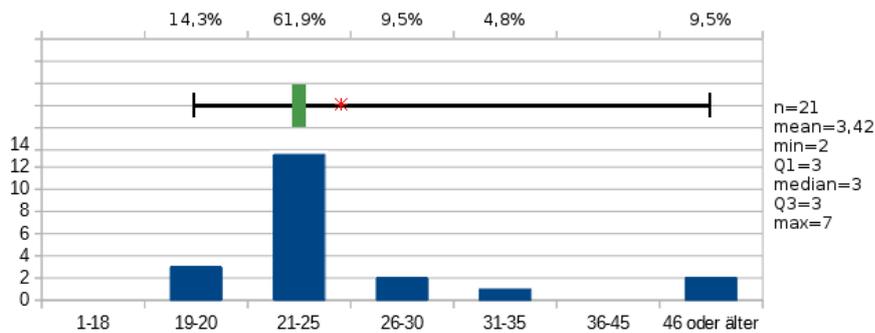


Abb. 5.11: Antwortenverteilung: 1. Bitte geben Sie ihr Alter an. Quelle: Eigene Darstellung

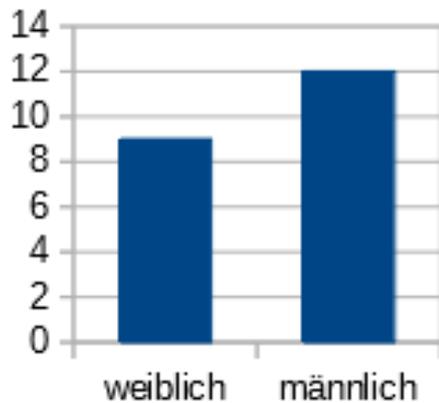


Abb. 5.12: Antwortenverteilung: 2. Bitte geben Sie ihr Geschlecht an. Quelle: Eigene Darstellung

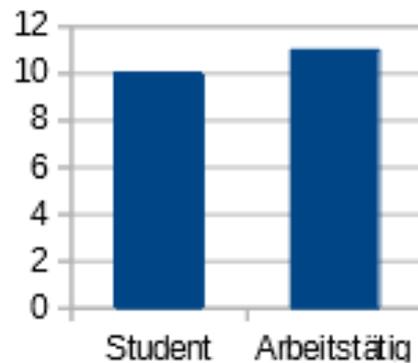


Abb. 5.13: Antwortenverteilung: 3. Sind sie zum Zeitpunkt dieser Umfrage Student einer Hochschule oder Universität? Quelle: Eigene Darstellung

Weiter gaben 10 der Teilnehmer (47,6%) an zum Zeitpunkt der Experimente Studierende zu sein während die restlichen 11 Teilnehmer (52,4%) arbeitstätig waren (siehe Abbildung 5.13). Die Studierenden befinden sich hierbei gleichmäßig verteilt zwischen dem ersten und neunten Semester und studieren Biotechnik, International Management, Information Medien Bibliothek, Informatik, BWL, Kulturwissenschaften sowie Fahrzeugbau. So ließen sich entsprechend Meinungen unterschiedlichster Fachkompetenzen einholen.

### 5.3.3 Relevante Vorkenntnisse

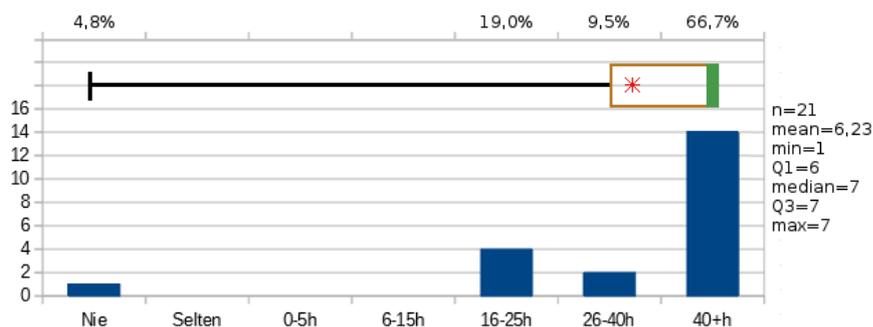


Abb. 5.14: Antwortenverteilung: 4. Wie häufig verwenden Sie einen Computer? Quelle: Eigene Darstellung

14 der Teilnehmer (66,7%) gaben an wöchentlich über 40 Stunden am Computer zu arbeiten. Auch die restlichen Teilnehmer (28,5%), bis auf einen (4,8%) der nie einen Computer verwendet, gaben an zwischen 16 und 40 Stunden wöchentlich am Computer zu verbringen (siehe Abbildung 5.14).

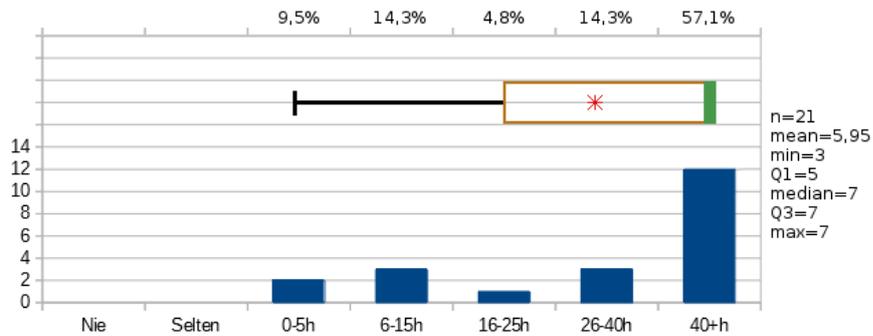


Abb. 5.15: Antwortenverteilung: 5. Wie häufig verwenden Sie einen Touchscreen?  
 Quelle: Eigene Darstellung

Bei der Verwendung von Touchscreens, sieht es seit dem Boom des Smartphones kaum anders aus. So gaben 12 Teilnehmer (57,1%) an über 40 Stunden die Woche einen Touchscreen zu verwenden. Die Antworten der restlichen 9 Teilnehmer (42,9%) erstrecken sich von einem Minimum bei 0-5 Stunden bis hin zu 26-40 Stunden (siehe Abbildung 5.15). Gerade seit dem zum Zeitpunkt der Evaluation vorherrschendem Pokemon GO<sup>1</sup> Phänomen sei die Verwendungsdauer vieler Teilnehmer, laut eigener Aussage, weiter angestiegen. Das arithmetische Mittel liegt der Verteilung zufolge bei 26-40 Stunden.

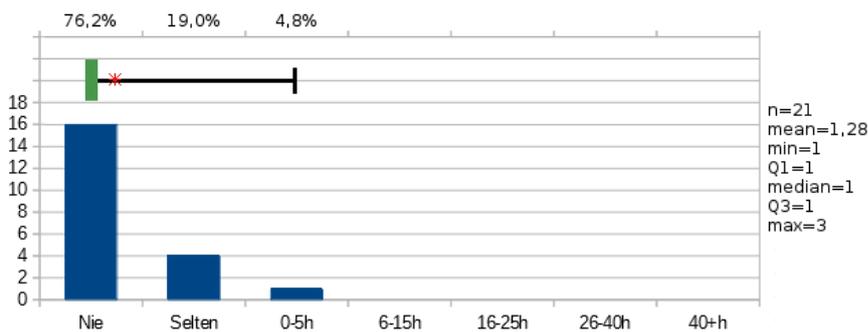


Abb. 5.16: Antwortenverteilung: 6. Wie häufig verwenden Sie eine 3D Gestensteuerung (Kinect oder Ähnliches)? Quelle: Eigene Darstellung

<sup>1</sup>AR Handyspiel unter <http://www.pokemongo.com/>

Die Kinect oder ähnliche 3D Gestenerkennungshardware scheinen im Vergleich eher selten zur Verwendung zu kommen. Insgesamt gaben 16 der 21 Teilnehmer (76,2%) an vorher noch nie mit 3D Gesten interagiert zu haben. Nur 4 der Teilnehmer (19,0%) kamen zumindest selten in Kontakt mit dieser Technik. Auf Nachfrage hin gaben die meisten an, dass es sich bei ihren Erlebnissen um die Wii Konsole von Nintendo gehandelt habe, womit sie bereits eine gestenähnliche Steuerung verwendet haben. Nur eine Person (4,8%) gab an 0-5h in der Woche 3D Gesten anzuwenden. Somit hat der Großteil keine vergleichbaren Vorerfahrungen mit dem Steuerungstyp gemacht (siehe Abbildung 5.16). Durch die hohen Computerkenntnisse sowie Kenntnisse im Umgang mit Touchscreens, hat sich ein Großteil der Teilnehmer recht schnell mit dem System zurechtgefunden. Lediglich bei der Einführung in die 3D Gestensteuerung musste ein höherer Aufwand für die Erläuterung aufgebracht werden.

Auf die Frage hin, wie häufig die Teilnehmer kollaborativ an einem Bildschirm arbeiten oder bereits arbeiten mussten, fielen die Antworten um einiges verteilter aus. Das arithmetische Mittel sowie der Median fallen bei der Verteilung auf 0-5 Stunden die Woche. 4 Personen (19,0%) bildeten mit ihrer Antwort bei „Nie“ das Minimum, während 2 Teilnehmer (9,5%) angaben zwischen 26 und 40 Stunden die Woche kollaborativ an einem Bildschirm zu arbeiten und somit das Maximum bildeten (siehe Abbildung 5.17). Dabei handelte es sich auf spätere Nachfrage hin überwiegend, um das gemeinsame Lernen oder Erarbeiten von Praktikumsaufgaben.

Aufgrund der Tatsache, dass es sich bei den Gruppen jeweils ausschließlich um reale Gruppen handelte, die sich bereits vor den Experimenten kannten, und die Kenntnisse mit kollaborativer Arbeit doch recht ausgeprägt waren, lies sich davon ausgehen, dass Kommunikationsschwierigkeiten bei der Mensch-Mensch-Interaktion auf einem Minimum gehalten werden. So sollte ein stärkerer Fokus auf die Mensch-Computer-Interaktion möglich sein.

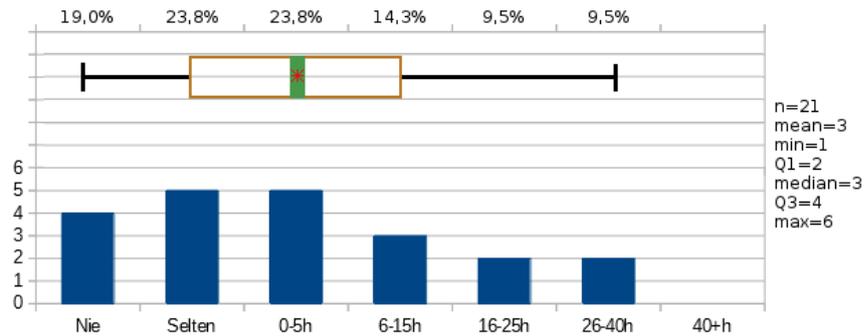


Abb. 5.17: Antwortenverteilung: 7. Wie häufig arbeiten Sie kollaborativ mit anderen Menschen an einem Anzeigemedium (zusammen an einem Bildschirm eine Aufgabe lesen, zusammen lernen, jemandem etwas erklären oder Ähnliches)? Quelle: Eigene Darstellung

Bei der Frage nach Vorerfahrungen bei der Zusammenstellung von Fotocollagen war die Antwortenverteilung wieder sehr einseitig (siehe Abbildung 5.18). Keiner der Teilnehmer hat bereits regelmäßiger eine Collage erstellt. Nur 9 der Teilnehmer (42,9%) haben zumindest 1 bis 2 mal, beispielsweise zum Jahrestag ihrer Eltern, eine solche Collage zusammengestellt.

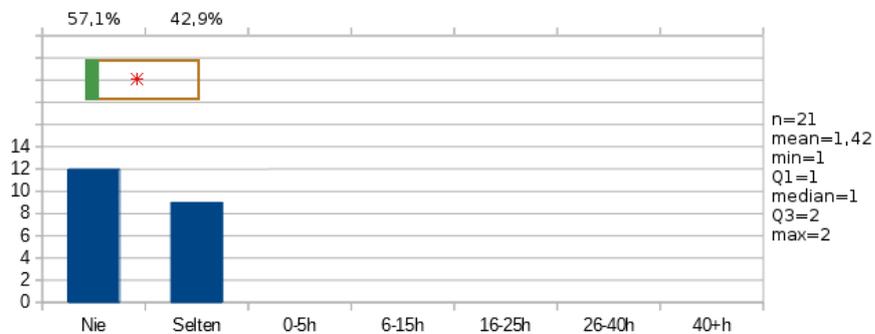


Abb. 5.18: Antwortenverteilung: 8. Wie häufig haben Sie bereits eine Bildercollage zusammengestellt? Quelle: Eigene Darstellung

### 5.3.4 Bewertung der Steuerungskonzepte für Kollaboration

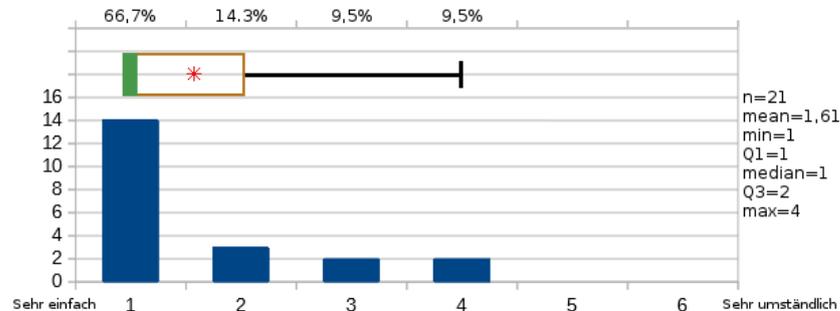


Abb. 5.19: Antwortenverteilung: 1. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage ausschließlich mit Touchgesten? Quelle: Eigene Darstellung

Die ausschließliche Touchsteuerung haben die Teilnehmer sehr positiv aufgenommen (siehe Abbildung 5.19) und das arithmetische Mittel der Bewertungen liegt bei 1,61. Dadurch dass die gängigen Touchgesten, die auch bei Smartphones und Tablets angewandt werden, benutzt wurden, war jeder Teilnehmer sofort mit den Interaktionsmöglichkeiten vertraut und empfand die Steuerung als äußerst intuitiv. So stimmten auch 14 der 21 Teilnehmer (66,7%) mit einer 1 ab. 3 Teilnehmer (14,3%) bewerteten das System mit einer 2 und jeweils 2 weitere Teilnehmer (9,5%) mit einer 3 und 4. Die etwas schlechteren Bewertungen wurden von 2 Teilnehmern durch die etwas andere Haptik der Bildschirmoberfläche, als sie es von ihren Smartphones gewohnt waren, begründet. Die Teilnehmer haben manchmal mehr Druck auf die Touchoberfläche ausgeübt als nötig wäre daraus resultierend empfanden sie die etwas rutschfeste Oberfläche des BenQ RP790 als hinderlich bei Objektverschiebungen. Das Infrarotgitter hätte die Finger auch ohne Druckausübung erkannt. Zwei weitere Teilnehmer fanden es unübersichtlich als Gruppe gemeinsam an einem Bildschirm zu interagieren. So konnte die linke Person teilweise nicht sehen was die rechte Person tut und umgekehrt, da die jeweilige Interaktionszone von dem dritten in der Mitte stehenden Teilnehmer verdeckt wurde.

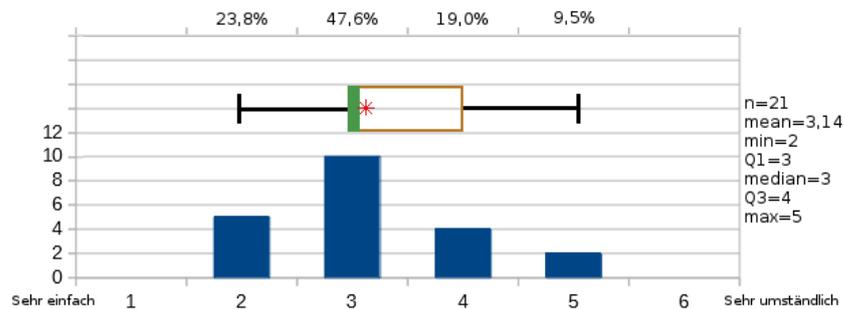


Abb. 5.20: Antwortenverteilung: 2. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage ausschließlich mit 3D Gesten? Quelle: Eigene Darstellung

Bei den 3D Gesten im Zusammenhang mit Kollaboration fallen die Bewertungsergebnisse im Vergleich schlechter aus und ergeben ein arithmetisches Mittel von 3,14. Der Großteil der Teilnehmer (47,6%) bewertete die Steuerung mit einer 3. Die restlichen Testpersonen (52,4%) haben gestreut eine Wertung zwischen 2 und 5 abgegeben. Aus dem Freitext auf der letzten Seite des zweiten Fragebogens lassen sich bei vielen Teilnehmern Gründe für die schlechter ausfallende Bewertung entnehmen. Zum einen wurde von 2 Teilnehmern eine zu hohe Latenz und von 6 Teilnehmern ein zu starkes Zittern des Cursors bemängelt, die bei feinen Positionierungen oder Skalierungen als störend empfunden wurden. Hier musste festgestellt werden, dass weitere Versuche nötig gewesen wären um einen guten Kompromiss bei der Anzahl der zu betrachteten vergangenen Bilder, die den Mittelwert bilden, welcher für die Cursorposition verwendet wird, zu finden. Benutzt man zu wenige vergangene Frames für die Position macht sich schnell ein Zittern in der Position bemerkbar, da zum einen immer ein natürliches Händezittern vorliegt, zum anderen die Kinect nicht immer vollkommen akkurate Informationen über die Handpositionen liefert. Verwendet man wiederum zu viele vergangene Frames, wird die Position zu langsam aktualisiert und es macht sich eine Art Latenz durch das Hinterherziehen des Cursors bemerkbar. Weiter haben 3 Teilnehmer angemerkt, dass durch die ungewohnte Steuerungstechnik mehr Konzentration für die Interaktion erforderlich war und so die Kommunikation mit den restlichen Teammitgliedern darunter litt. Ein weiterer gravierender Minuspunkt für die Teilnehmer war das bereits angesprochene Gorilla Arm Syndrome (siehe Abschnitt [3D Gesten](#)). Insgesamt beklagten sich 7 der 21 Teilnehmer über müde und schwer werdende Arme. Außerdem wurde von 6 Teilnehmern angemerkt, dass sie etwas Probleme mit dem Greifen von Elementen oder dem Verkleinern von Elementen über einen gewissen Punkt hinaus hatten. Weiter konnte man bei den

Versuchen regelmäßig beobachten wie bei den Teilnehmern ungewollt die zweite Hand mit vor den Körper „rutschte“ und sie somit unbeabsichtigt in den Transformationsmodus gelangten. Diese Körperhaltung lies sich vorallem bei den weiblichen Teilnehmern wiederholt beobachten. Positiv sei hier jedoch angemerkt, dass insgesamt 7 Teilnehmer angaben, mehr Übung könne viele der Probleme beseitigen und sie würden Potenzial in dieser Technik sehen. Außerdem wurde von mehreren Teilnehmern der durch 3D Gesten zusätzlich entstehende Platz, beziehungsweise Interaktionsradius, und das freie Sichtfeld auf die gesamte Interaktionsfläche festgestellt und als hilfreich befunden.

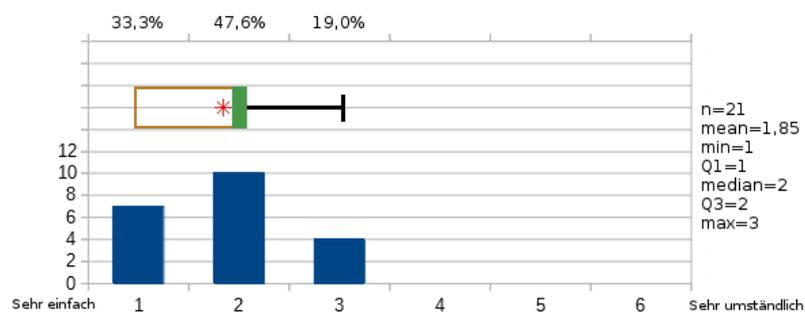


Abb. 5.21: Antwortenverteilung: 3. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage mit der Kombination beider Steuerungstypen? Quelle: Eigene Darstellung

Die Kombination beider Steuerungstypen für kollaboratives Arbeiten reiht sich den Bewertungen nach zwischen der Touch- und 3D Gestensteuerung ein (siehe Abbildung 5.21). Hier liegt das arithmetische Mittel bei einer 1,85 und die Verteilung der Einzelbewertungen erstreckt sich von 1 bis 3. Bei der Kombination beider Steuerungen wurde im Freitext ausschließlich von 2 Teilnehmern bemängelt, dass die Touchgesten Anwender den 3D Gesten Anwendern die Sicht auf den Interaktionsbereich versperrten. Die anderen Teilnehmergruppen haben sich in der Hinsicht besser abgesprochen und standen beim Interagieren per Touch seitlich neben dem Bildschirm um dieses Problem zu umgehen.

Bei der Beobachtung der Gruppen ließen sich, aber weitere interessante Punkte feststellen. Es gab Gruppen bei denen die Kommunikation zwischen den vorne am Bildschirm stehenden Touch Anwendern und den hinten stehenden 3D Gesten Anwendern, vollkommen abbrach. Durch die räumliche Entfernung bildeten sich so zwei unabhängig voneinander interagierende Gruppen. Doch genauso gab es auch Gruppen, die strukturierter an die Aufgabe heran gegangen sind und sofort die Vorteile beider Steuerungstypen erkannten.

So haben einige Gruppen jedem Mitglied eine feste Rollen zugeteilt. Beispielsweise stand einer der Teilnehmer rechts von dem Monitor und schob die entsprechend passenden Bilder von dem Slider in Richtung der Collage während die anderen beiden Teilnehmer jeweils für eine Seite der Doppelseiten zuständig waren und die Bilder optisch ansehnlich per 3D Gesten auf ihrer Seite positionierten und dabei untereinander Ideen für die Verteilung austauschten. Eine weitere Gruppe wiederum verteilte die Rollen genau entgegengesetzt. Eine hinten stehende Person hatte den Gesamtüberblick über das Aussehen der Collage und zog seiner Meinung nach passende Bilder aus dem Slider, die ihm daraufhin auf halben Weg von den mit Touchgesten interagierenden Teilnehmern abgenommen und mit Absprache präzise positioniert und transformiert wurden. Ihnen ist nämlich schnell bewusst geworden, dass 3D Gesten für feine Arbeiten eher weniger geeignet sind und so arrangierten sie sich entsprechend um das Optimum aus den gegebenen Werkzeugen heraus zu holen.

Einen Wechsel von Touch auf 3D Gesten, oder umgekehrt, nahm kaum einer der Teilnehmer bei der Einrichtung der dritten Doppelseite der Collage vor. Diese Beobachtung könnte jedoch an der Kürze der Aufgabe liegen. Interessant wäre zu sehen ob bei verschiedenen komplexen und zeitintensiven Aufgaben häufiger Wechsel zwischen den Steuerungstypen vorgenommen werden würden. Weiter wäre es auch interessant mit verschiedenen Gruppengrößen zu arbeiten und zu sehen wie sich dies auf die Arbeit und das Verhalten der Teilnehmer auswirken würde.

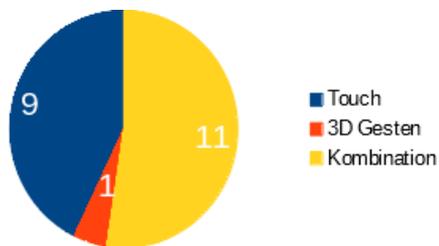


Abb. 5.22: Antwortenverteilung: 4. Welche Steuerung würde Ihrer Meinung nach für die größte Übersichtlichkeit sorgen? Quelle: Eigene Darstellung

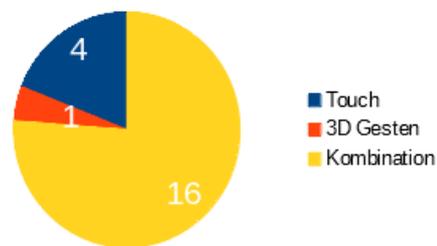


Abb. 5.23: Antwortenverteilung: 5. Welche Steuerung, würde Ihrer Meinung nach für die größte Produktivität sorgen? Quelle: Eigene Darstellung

Obwohl die Bewertung der Kombination beider Steuerungstypen im Schnitt schlechter ausfiel als die der bloßen Touchsteuerung, schnitt die Kombination sowohl bei der Frage nach der Steuerung, welche die größte Übersichtlichkeit bietet als auch bei der

Frage nach der Steuerung, welche die höchste Produktivität verspricht, am besten ab. Bei der Übersichtlichkeit entschieden sich 11 Teilnehmer (52,4%) für die Kombination beider Steuerungstypen, knapp gefolgt von 9 Teilnehmern (42,9%), die die Arbeit nur mit Touchgesten als übersichtlicher empfanden. Nur ein Teilnehmer (4,8%) stimmte in dieser Hinsicht für die bloße 3D Gestensteuerung ab. Bei der Produktivität waren sich die Teilnehmer einiger. 16 Teilnehmer (76,2%) stimmten bei dieser Frage für die Kombination ab. Von den restlichen 5 Teilnehmern fanden 4 Teilnehmer (19,0%) die Arbeit mit Touchgesten und 1 Teilnehmer (4,8%) die Arbeit mit 3D Gesten am produktivsten. Aus dieser Statistik kann man wieder das Potenzial der Kombination beider Steuerungen erkennen, wenn man es schafft die bei 3D Gesten aufgefallenen Probleme zu minimieren.

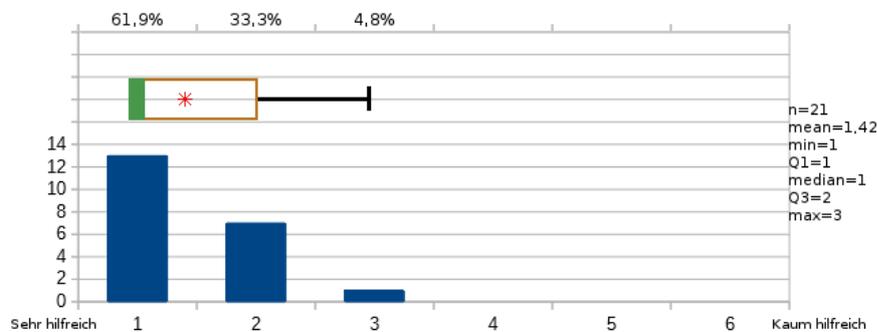


Abb. 5.24: Antwortenverteilung: 6. Wie fanden Sie die farbliche Zuordnung der Zeiger sowie der gerade in Interaktion befindlichen Bilder? Quelle: Eigene Darstellung

Die farbliche Zuordnung der Zeiger und der jeweils in Interaktion befindlichen Elemente zu den entsprechenden Teilnehmern, bei Touch- sowie 3D Gesten Interaktion, wurde als sehr hilfreich wahrgenommen (siehe Abbildung 5.24). 13 Teilnehmer (61,9%) stimmten mit einer 1 ab und bis auf einen Teilnehmer (4,8%), der die farbliche Zuordnung mit einer 3 bewertete, gaben die restlichen 7 Teilnehmer (33,3%) eine 2 als Bewertung ab. Durch die Zuordnung wusste jeder Teilnehmer zu jedem Zeitpunkt wer gerade die Kontrolle über welches Objekt hält. Einziges Manko hierbei war, dass sich die Teilnehmer merken mussten welche Farbe ihnen bei der Registrierung zugeordnet wurde. So kam es häufiger vor, dass sich die Situation ergab in der die Testgruppe per Absprache Einen nach dem Anderen Bewegungen ausführen ließ um im Nachhinein noch einmal festzustellen, welcher farblich hervorgehobene Cursor wem gehört.

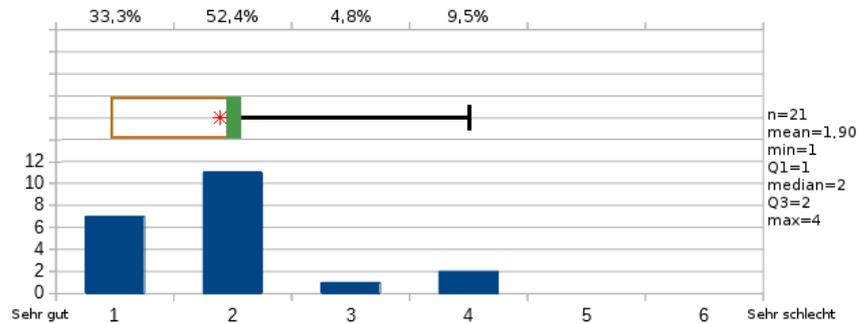


Abb. 5.25: Antwortenverteilung: 7. Wie fanden Sie die Arbeit als Team zusammen an einem Anzeigemedium? Quelle: Eigene Darstellung

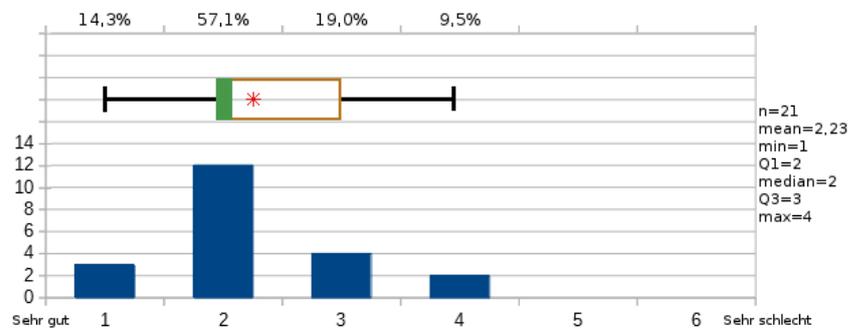


Abb. 5.26: Antwortenverteilung: 8. Wie gut könnten sie sich vorstellen ein solches System welches Touch- sowie 3D Gesten unterstützt auch für andere Gruppentätigkeiten und Aufgaben zu verwenden? Quelle: Eigene Darstellung

Die Kollaboration an einem großen Anzeigemedium wurde mit einem arithmetischen Mittel von 1,9 als sehr gut bewertet (siehe Abbildung 5.25). 7 Teilnehmer (33,3%) haben die gemeinsame Arbeit mit einer 1 bewertet, 11 weitere Teilnehmer (52,4%) mit einer 2, nur 1 Teilnehmer (4,8%) mit einer 3 und die 2 übrigen Teilnehmer (9,5%) mit einer 4. Insgesamt gaben 10 Testpersonen im Freitext an, dass es ihnen sehr viel Spaß gemacht hat und sie gerne häufiger so zusammenarbeiten würden und am liebsten immer auf diese Weise Fotocollagen zusammenstellen würden. 4 der Teilnehmer, bei denen es sich ausschließlich um Frauen handelte, gaben weiter an, dass die Interaktion mit 3D Gesten das Gefühl von Superkräften vermittelte. Weiter wurde auch erwähnt, dass die Möglichkeit eigene Texte in die Collage einzubauen die Kreativität steigern und die Applikation auf den Smartphones zum Hochladen der Elemente eine hilfreiche Erweiterung wäre.

Nichts desto trotz wurde auch eine Vielzahl von konkreten Verbesserungsvorschlägen genannt. Einige Teilnehmer wünschten sich eine Erweiterung um Funktionen wie „Undo“ und „Redo“, wieder andere Teilnehmer äußerten den Wunsch nach einem Berechtigungssystem beziehungsweise einer administrativen Rolle, die eine übergeordnete Kontrolle den restlichen Benutzern gegenüber hat, um so mehr Struktur in die Kollaboration zu bringen. Leitgedanke war bei diesem Wunsch unter anderem das gleichzeitige Umblättern der Collage durch mehrere Anwender zu verhindern. Zwei weitere geäußerte Vorschläge waren das Löschen des Elements durch Greifen und Schütteln und die Möglichkeit die entstandene Collage abzuspeichern oder auszudrucken.

Auch auf die Frage hin ob sich die Teilnehmer eine Kombination aus 2D und 3D Gesten zur Steuerung in anderen Aufgabenbereichen vorstellen könnten war der Großteil sehr positiv gestimmt. So haben 15 der Teilnehmer (71,4%) auf die Frage hin mindestens mit einer 2 abgestimmt. Von den restlichen 6 Teilnehmern konnten es sich 4 Testpersonen (19,0%), die mit einer 3 abgestimmt haben, zumindest noch einigermaßen vorstellen und nur 2 Testpersonen (9,5%), die mit einer 4 abgestimmt haben eher weniger vorstellen.

### 5.3.5 Fazit

Zusammenfassend lässt sich sagen, dass die Teilnehmermenge etwas versierter sein hätte könnten, was die Altersgruppen betrifft. Dafür ist ein ausgeglichenes Verhältnis zwischen männlichen und weiblichen Teilnehmern festzustellen. Genauso ausgewogen ist das Verhältnis zwischen studierenden und berufstätigen Teilnehmern. Die Studierenden decken außerdem ein großes Spektrum an verschiedenen Fachrichtungen ab und befinden sich gleichmäßig verteilt in den Semestern 1 bis 9. Die meisten Teilnehmer sind was Computergrundkenntnisse und Touchgesten betrifft gut geübt. Auch das kollaborative Arbeiten an einem Bildschirm ist vielen der Teilnehmer nicht fremd gewesen. Dafür mangelt es dem Großteil der Testpersonen an Kenntnissen im Umgang mit 3D Gesten und dem Zusammenstellen von Fotocollagen.

Die Touchsteuerung wurde, wie erwartet, sehr positiv aufgenommen. Sie wurde, auf die Frage hin wie die Teilnehmer die Zusammenarbeit an der Collage mit dem Steuerungskonzept fanden (siehe Abbildung 5.19), am besten bewertet, auch wenn sie in Anbetracht der größten Übersichtlichkeit (siehe Abbildung 5.22) und Produktivität (siehe Abbil-

dung 5.23) etwas schlechter abschnitt als die Kombination beider Steuerungsmethoden. Die Teilnehmer bemängelten die gewöhnungsbedürftige Haptik des Bildschirms sowie die Blockierung des Sichtfeldes auf die Interaktionsfläche und die anderen Teilnehmer durch das dichte nebeneinander stehen der Gruppenmitglieder. Dennoch hielten sich die Platzprobleme bei der Gruppen- und Bildschirmgröße in Grenzen. Mit steigender Gruppen- oder Bildschirmgröße wird bei diesem Steuerungskonzept ein Abwärtstrend der Bewertungen erwartet. Da es ab einer gewissen Anzahl an Gruppenmitgliedern unmöglich wird parallel an einem begrenzt großen Bildschirm zu interagieren. Erhöht sich wiederum die Bildschirmgröße, so wird es zunehmend schwerer mit Touchgesten an jeden Punkt der Benutzeroberfläche zu gelangen.

Die 3D Gestensteuerung wurde, obwohl sie als spaßig und interessant angesehen wurde, bei der selben Frage (siehe Abbildung 5.20) erkennbar schlechter bewertet. Auch bei der Frage nach der größten Übersichtlichkeit (siehe Abbildung 5.22) oder Produktivität (siehe Abbildung 5.23) stimmte kaum ein Teilnehmer für dieses Konzept ab. Dafür hat die Steuerung trotz erkennbarer Vorteile, wie der höheren Bewegungs- sowie Sichtfreiheit, zu viele Probleme, die es noch zu lösen gilt. Im Gegensatz zur Touchgestensteuerung wird hier bei steigender Gruppen- oder Bildschirmgröße ein Aufwärtstrend bei den Bewertungen erwartet, da sich zunehmend die Vorteile gegenüber der Touchgestensteuerung herauskristallisieren sollten.

Die Kombination von 2D und 3D Gesten bietet für das Arbeiten von Teams an einem Anzeigemedium klare Vorteile, hat jedoch einige Hürden, unter anderem bei der Integration von 3D Gesten, die es zu überwinden gilt. Die Kombination beider Steuerungskonzepte wurde von den Teilnehmern mit einem Schnitt von 1,85 (siehe Abbildung 5.21) nur knapp hinter der bloßen Touchsteuerung, die einen Bewertungsschnitt von 1,61 erhielt (siehe Abbildung 5.19), eingestuft. Sie wurde jedoch im Hinblick auf Produktivität (siehe Abbildung 5.23) sowie Übersichtlichkeit (siehe Abbildung 5.22) bei Kollaboration häufiger als das bessere Konzept im Vergleich zur reinen Touch- oder 3D Gestensteuerung bewertet. Die Teilnehmer sahen also Potenzial in der Kombination und empfanden sie als das angebrachtere Konzept. Auch bei der Beobachtung der Versuche wurde deutlich, dass die Kombination beider Steuerungsmethoden die jeweiligen Stärken verbindet. Entsprechend waren sämtliche Gruppen bei Verwendung der Kombination am schnellsten mit dem Versuch fertig. Mit der Touchsteuerung haben sie etwas länger gebraucht und die 3D Gestensteuerung hat deutlich die meiste Zeit in Anspruch genom-

men. Man kann jedoch nicht klar sagen, ob die Reihenfolge in denen die Steuerungen von den Gruppen erprobt wurden, einen Einfluss auf die Ergebnisse hatten. Da vor der Kombination beider Steuerungstypen die bloße 3D Gestensteuerung getestet wurde, wäre es möglich, dass durch den daraufhin erhöhten Frustrationsfaktor und die bereits ermüdeten Arme, auch die Bewertung der Kombination beeinflusst wurde.

Die wichtigsten Gründe für die schlechtere Bewertung waren Aspekte, die das Gesamtgefühl getrübt haben, jedoch überwiegend beseitigt werden können. So wurde eine nicht ganz optimierte Balance zwischen Latenz und Genauigkeit der 3D Gesten, sowie vorgekommene Aussetzer bei dem Greifen von Objekten bemängelt. Diese Probleme könnten durch weitere Optimierungen bei der Programmierung oder das Hinzuziehen weiterer Sensoren behoben werden (siehe Abschnitt [Sensorfusion](#)). Ein anderer Grund, welcher wiederum nur schwer anzugehen ist, ist das Gorilla Arm Syndrome. Hierfür gibt es dem Wissensstand des Verfassers dieser Arbeit zufolge bisher keine anwendbare Lösung. Die Teilnehmer gingen jedoch davon aus, dass sich dieses Problem durch eine regelmäßige Anwendung des Systems und nur sporadischen Einsatz von 3D Gesten mildern ließe. Ein weiteres Problem war das Versperren des Sichtfeldes von 3D Gesten Anwendern durch Touchgesten Anwender und die sinkende Kommunikation, die gleichfalls, laut Teilnehmer, durch mehr Übung und Routine im Umgang mit einem solchen System schnell beseitigbar wären.

Klare Vorteile, die von den Teilnehmern bei der Kombination beider Steuerungsmethoden festgestellt wurden, waren der erhöhte Interaktionsradius, da man sich frei im Bereich vor dem Bildschirm bewegen kann und das damit einhergehende größere Sichtfeld, soweit nicht anderweitig versperrt. Weiter wurde die schnellere Interaktionsmöglichkeit angesprochen, wenn man beispielsweise an einem Konferenztisch sitzt und schnell etwas auf dem Bildschirm hervorheben möchte ohne den gesamten Weg auf ihn zu zugehen.

Letztendlich konnten sich die meisten Teilnehmer vorstellen häufiger mit der Kombination beider Steuerungsmethoden zu arbeiten (siehe [Abbildung 5.25](#)) und auch andere Aufgabenfelder auf diese Weise kollaborativ zu bearbeiten (siehe [Abbildung 5.26](#)).

In den folgenden Abschnitten werden Aspekte thematisiert, die das System weiter verbessern und die Akzeptanz der Anwender erhöhen könnten.

### **Magazin abspeichern**

Eine mögliche Erweiterung wäre gewesen, wie anfangs auch vorgehabt, den Ist-Zustand des Magazins zu jedem Zeitpunkt abzuspeichern und gegebenenfalls auch ausdrucken zu können. Aus Zeitmangel ist diese Funktion entfallen, die für die eigentliche Eignungsuntersuchung der 2D und 3D Gesten für Kollaboration keinerlei wesentliche Rolle spielt. Sie hätte vielmehr als zusätzliche Motivation für die Testpersonen fungieren können, indem die aufgewandte Arbeit in einem erhalten bleibenden Produkt endet und sie das Ergebnis behalten hätten können.

### **Größerer und leichter zugänglicher Funktionsumfang**

Wie in der Auswertung der Evaluation bereits angesprochen, hätten sich einige Probanden leicht zugängliche Schnellfunktionen, wie das Rückgängig machen der zuletzt ausgeführten Aktion oder das Kopieren von Elementen gewünscht. Bei weiterer Entwicklung des Systems könne man dementsprechend über ein durch eine gewisse Geste an der Position der Hand auflappbares Dropdown Menü nachdenken. Man müsse bei dem Design einer solchen Funktion nur beachten, dass das Menü nicht die Arbeit der anderen Gruppenmitglieder negativ beeinflusst wird. Im Ausblick werden Erweiterungen und Änderungen angesprochen, die in Zukunft an dem System vorgenommen werden könnten. Außerdem werden auf Grundlage der stattgefundenen Experimente und erlangten Erkenntnisse weitere interessante Fragestellungen für den Bereich HCI und CSCW erarbeitet, die es unter Umständen in weiteren Arbeiten zu beantworten gilt.

### **Handzustände von mehr als zwei Anwendern erfassen**

Im Laufe der ersten Versuche stellte sich heraus, dass die Kinect C# SDK nur die Handzustände von den zwei nächsten Skeletten erfasst. Um der Anwendung zukünftig zu ermöglichen die Handgesten von bis zu sechs Nutzern zu erfassen, kamen folgende Ansätze in den Sinn:

- Eine eigene Handzustandserkennung anhand der Tiefenbildinformationen programmieren.
- Im Round-Robin Algorithmus über alle erkannten Skelette iterieren. Nacheinander jeweils zwei Skelette aus der Gesamtmenge auswählen, die die Skelette sein sollen

deren Handzustände für das nächste Einzelbild verfolgt werden. So sinkt die Aktualisierungsfrequenz der Handzustände von den einzelnen Skellete doch dafür werden alle erfasst.

- Man kann mehrere Kinects aufbauen und jeder eine andere Untermenge an zu verfolgenden Handzuständen zuordnen. Der Overhead für diese Variante ist jedoch zu enorm.

### Erhöhung der Genauigkeit & des Interaktionsradius

Anfänglich war geplant eine 3D Gestenerkennung auf kleinere Distanzen zu dem Bildschirm zu realisieren. Entweder sollten über- oder unterhalb des Bildschirms Leap Motion Sensoren (siehe Abschnitt [Leap Motion](#)) angebracht werden, welche auf den Bereich frontal vor der Bildschirmfläche ausgerichtet sind oder weitere Kinect Kameras, welche den Bereich direkt vor dem Monitor abfangen und ihre Informationen mittels Sensorfusion (siehe Abschnitt [Sensorfusion](#)) teilen. Doch auch diese Aufgabe hätte den zeitlichen Rahmen dieser Arbeit überspannt. Das Hinzuziehen und komplementäre oder kompetitive Fusionieren (siehe Abschnitt [Sensorfusion](#)) weiterer Sensoren würde eine höhere Genauigkeit und Zuverlässigkeit der Daten und bei durchdachter Positionierung der Sensoren, außerdem einen größeren Interaktionsraum ermöglichen. Durch das umfangreichere Gesamtbild der Sensoren würden auch leicht hintereinander stehende und sich überschneidende Körper kein Problem mehr darstellen. Wenn der, auf dem Display positionierte, Sensor beispielsweise die Handbewegung eines Benutzers nicht mehr erkennen kann, da sie von einem Anderen verdeckt wird, kann unter Umständen ein hinten, seitlich oder oberhalb der Gruppe positionierter Sensor die entsprechenden Daten liefern.

### Leap Motion

In der Leap Motion sind 2 Kameras sowie 3 Infrarot-LEDs verbaut. Die Infrarot-LEDs bestrahlen einen bis zu ungefähr 80 cm hohen kegelförmigen Bereich oberhalb der Leap Motion. Die beiden Kameras nehmen separat jeweils ein Graubild der ange-



Abb. 5.27: Das Innere einer Leap Motion.  
Quelle: ([Col14](#))

strahlten Objekte innerhalb des Kegels auf.

Als nächstes werden mehrere Algorithmen aufeinanderfolgend angewandt, die eine

3D Repräsentation der Daten rekonstruieren, daraufhin werden die Informationen herausgefiltert, die getrackt werden sollen, wie Finger oder Werkzeuge. Die Leap Motion wird aufgrund der von der Lichtausbreitung eingeschränkten Reichweite meist für die Erkennung von Handgesten angewandt (vgl. Col14).

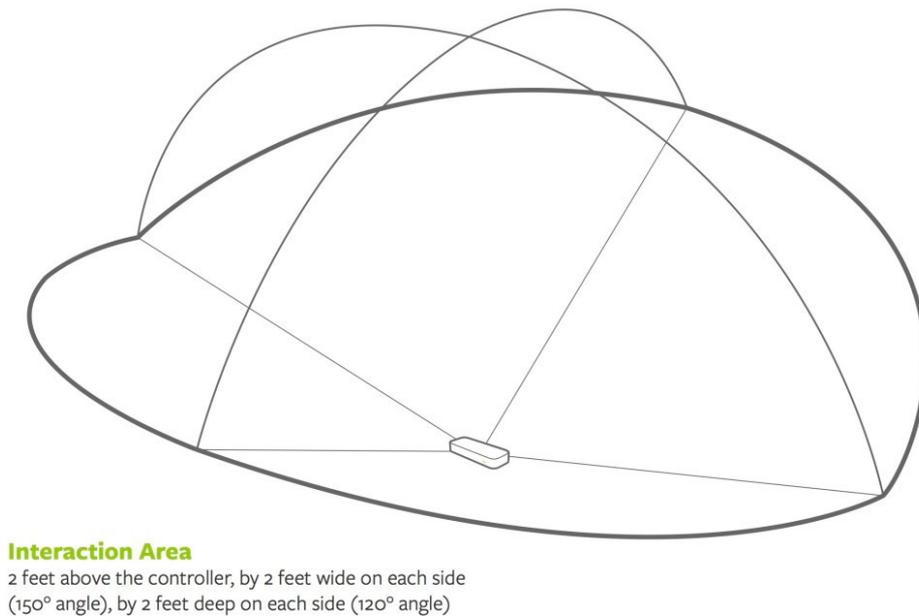


Abb. 5.28: Reichweite einer Leap Motion. Quelle: (Col14)

### Sensorfusion

Sensorfusion setzt sich aus den Teilbegriffen Sensor und Fusion, also der Verschmelzung der Sensoren beziehungsweise der Teilinformationen, die diese Sensoren liefern, zusammen. Sensorfusion besitzt eine Vielzahl in der Literatur häufig synonym verwendeter Bezeichnungen wie Data Fusion und Information Fusion, die in einigen Fällen jedoch auch eine etwas andere Bedeutung haben können. Auch an Definitionen stößt man auf unzählig unterschiedliche. Elmenreich definiert Sensorfusion folgendermaßen:

„Sensor Fusion is the combining of sensory data or data derived from sensory data such

that the resulting information is in some sense better than would be possible when these sources were used individually“ (Elm02, S.8).

Ein Aufbau mit mehreren Kinect Kameras, die auf eine Szenerie ausgerichtet werden und deren Daten kombiniert einen Mehrwert für den Anwender bieten, würde man demnach als Sensorfusion bezeichnen. Genauso könne man die Daten, die die Kinect liefert mit Daten der Leap Motion kombinieren, um eine größere Abdeckung oder höhere Genauigkeit des im Fokus liegenden Bereiches zu erreichen.

Man kann Sensorfusion in direkte sowie indirekte Fusion untergliedern. Wobei unter direkter Fusion das Zusammenführen von Daten homogener oder heterogener Sensoren, Softsensoren oder auch in der Vergangenheit von einem Sensor aufgenommenen Daten zu verstehen ist. Unter indirekter Fusion ist unter anderem das Hinzuziehen menschlichen Inputs zu verstehen (vgl. Elm02, S.9).

Weiter existieren mehrere Kategorisierungsmethoden für Sensorfusion. Eine dieser Methoden teilt sie weiter in komplementäre, kompetitive und kooperative Fusion ein.

**komplementäre Fusion:** Die Fusion disjunkter Informationen, die von voneinander unabhängigen Sensoren stammen, jedoch kombiniert werden können um ein vollständigeres Gesamtbild über das im Fokus liegende Objekt zu erhalten. Ein Beispiel wäre das Aufstellen mehrerer Kameras mit dem Ziel mehr von dem im Fokus liegenden Bereich abzudecken. Man kann durch diese Art der Fusion also der Datenunvollständigkeit entgegenwirken.

**kompetitive Fusion:** Die Fusion unabhängiger Daten eines oder mehrerer Sensoren über die selben Eigenschaften des im Fokus liegenden Objektes. Durch diese Art der Fusion wird die Ausfallsicherheit und Robustheit des Systems erhöht.

**kooperative Fusion:** Durch die Fusion mehrerer unabhängiger Sensordaten wird ein Informationsgewinn erzeugt, der durch die getrennten Daten der einzelnen Sensoren nicht zur Verfügung stehen würde (vgl. Elm02, S.15-17).

## 6 Fazit

Dieses Kapitel unterteilt sich in eine Zusammenfassung und einen Ausblick. In der Zusammenfassung werden zunächst die wichtigsten Punkte noch einmal zusammengetragen und die zu beantwortete Fragestellung der Arbeit behandelt. In dem Ausblick werden daraufhin einige offen gebliebene Fragen angerissen, ein Blick auf die Zukunft des Forschungsgebietes CSCW und die Generalisierbarkeit sowie Übertragbarkeit der Steuerungsmethoden geworfen.

### 6.1 Zusammenfassung

Die grundlegende Frage, die es im Laufe dieser Arbeit zu klären galt, war, inwiefern sich 2D-, 3D- und eine Kombination von 2D und 3D Gesten dazu eignen eine Gruppe von Menschen bei der kollaborativen Arbeit an einem Anzeigemedium zu unterstützen (siehe Kapitel 1.2). Die Arbeit ist, da hierbei die Mensch-Computer-Interaktion eine wichtige Rolle spielt, oder noch genauer, da es sich hierbei um das durch Computertechnologien unterstützte Arbeiten in Gruppen handelt, im Bereich CSCW (siehe Kapitel 2.4), einem Teilbereich des HCI (siehe Kapitel 2.2), angesiedelt. Die resultierenden Erkenntnisse können, durch die rapide Entwicklung des Ubiquitous Computing (siehe Kapitel 2.3), jedoch auch für viele weitere Bereiche von Interesse sein.

Um der Beantwortung der Frage einen Schritt näher zu kommen wurden alle drei Steuerungsmethoden nacheinander in einem konkreten Testszenario (siehe Abschnitt Szenario), in dem ein Team versucht gemeinsam eine Fotocollage zusammenzustellen, angewandt und verglichen. Für die Auswertung wurden die Ergebnisse, der von den Testpersonen ausgefüllten, Fragebögen (siehe Kapitel 7.1 und 7.2) sowie eigene, während der Versuche durchgeführten, Beobachtungen verwendet.

Bei der Durchführung der Versuche (siehe Kapitel 5.2) mit einer Gruppengröße von drei Mitgliedern und einem Anzeigemedium mit einer Diagonale von 79 Zoll wurde die Kombination beider Steuerungsmethoden was die Übersichtlichkeit und Produktivität

bei Kollaboration betrifft am häufigsten als die beste Variante gewählt. Dicht gefolgt von der bloßen Touchsteuerung, welche auf die Frage hin wie die kollaborative Arbeit im Allgemeinen mit ihr gefiel, dennoch knapp besser abschnitt als die Kombination beider Methoden. Dies liegt an mehreren Schwierigkeiten, die die 3D Gestensteuerung mit sich bringt, welche auch Grund dafür sind dass die bloße 3D Gestensteuerung weit abgeschlagen hinter den anderen beiden Konzepten platziert ist (siehe Kapitel 5.3). Einige der Gründe können jedoch angegangen werden und so die 3D Gestensteuerung sowie die Kombination beider Steuerungen noch attraktiver werden lassen (siehe Kapitel 5.3.5).

Insgesamt waren die meisten Teilnehmer dennoch positiv von dieser Art der Kollaboration angetan und konnten sich vorstellen zukünftig auch in anderen Bereichen mit der Kombination beider Steuerungsmethoden als Gruppe zusammenzuarbeiten.

## 6.2 Ausblick

Es ist schwierig auszumachen in welche Richtung sich der Bereich CSCW und die Arbeit in Projektteams entwickeln wird. Ein guter Ort um sich über die neuesten Erkenntnisse und Entwicklungen in diesem Bereich zu informieren ist die jährlich stattfindende ACM Conference on Computer Supported Cooperative Work & Social Computing<sup>1</sup>. Auch inwiefern sich Single Display Groupware durchsetzen wird lässt sich kaum vorhersagen. Was jedoch dafür spricht sind die immer weiter sinkende Kosten großer Displays und der immer größer werdende Funktionsumfang. Dabei sollte man nicht aus dem Auge verlieren, dass Single Display Groupware nicht nur für Unternehmen immer interessanter werden sollte, sondern auch das Bildungssystem langsam in die digitale Richtung umschwankt. So werden in vielen Klassenräumen bereits heute interaktive Whiteboards verwendet, um den Unterricht zu unterstützen. Diese werden jedoch häufig mit herkömmlichen Inputgeräten wie der Maus und Tastatur gesteuert. Interessant wäre hier zu sehen wie sich intuitive Interaktionsformen wie 2D und 3D Gesten auf das Lernen auswirken würden. So könnten Lehrer und Schüler problemlos gemeinsam den Inhalt der „digitalen Tafel“ beeinflussen. Weiter kann auch für den häuslichen Gebrauch ein solches System von Interesse sein. Eine Familie wäre in der Lage von ihrem Sofa aus kollaborativ eine Fotocollage zusammen zu stellen, gestengesteuerte Spiele zu spielen oder anderen gemeinsamen Tätigkeiten nachzugehen. Auch öffentliche Orte dürfen hier keinesfalls

---

<sup>1</sup><http://cscw.acm.org/2017>

vergessen werden. Dort stehen bereits seit längerem große nicht interaktive Bildschirme zur Vermittlung von Informationen bereit, sei es um die Anzahl der freien Parkplätze, den aktuelle Bus- und Bahnfahrplan, eine digitale Werbetafel oder vieles anderes darzustellen. Es wird auch versucht solche Bildschirme interaktiver zu gestalten. Zum einen um mehr Aufmerksamkeit und Interesse zu erregen zum anderen um öffentliche Orte lebendiger zu gestalten.

Weiter ist nicht vollkommen klar inwiefern sich Single Display Groupware Systeme überhaupt auf die kollaborative Arbeit auswirken. Es bleiben Fragen offen wie, wie sich die Gruppengröße oder Bildschirmgröße jeweils auf die Kollaboration auswirken könnten.

Auch ob sich Probleme wie die Ungenauigkeit von 3D Gesten und das bei der Verwendung eintretende Gorilla Arm Syndrome zukünftig lösen lassen werden, kann man zum heutigen Zeitpunkt nicht wirklich sagen.

Was sich wiederum sagen lässt ist, dass Computer uns, wie Mark Weiser (vgl. [Wei91](#)) schon vor Jahren erahnte, inzwischen überall umgeben und die Tendenz steigend ist. Für viele dieser Computer müssen neue Eingabemethoden entwickelt werden. Herkömmliche Methoden eignen sich oftmals nicht mehr, stattdessen ist Seamless Interaction gefragt. Touchgesten, die seit dem Erfolg der Smartphones und Tablets, gravierend an Akzeptanz gewannen, waren ein Schritt in die richtige Richtung, sind aber dennoch nicht in jeglichen Situationen die Lösung. Möchte man einer elektronischen Gardine beispielsweise mitteilen sie solle sich schließen wäre die einfachste Methode eine gerätelose 3D Geste in der Luft die einer Schließbewegung ähnelt, statt auf einem Gerät eine entsprechende Taste drücken zu müssen.

## **7 Materialien der Untersuchung**

## 7.1 Fragebogen vor dem Experiment



# Fragebogen

vor dem Experiment

Hochschule für Angewandte Wissenschaften Hamburg

Hamburg University of Applied Sciences

Bitte kreuzen Sie für jede der folgenden Fragen nur eine Antwort an. Falls Sie eine Frage nicht beantworten möchten, kreuzen Sie bitte „Enthaltung“ an. Ihre Antworten werden vertraulich behandelt und anonymisiert. Sie werden ausschließlich für statistische Erhebungen innerhalb meiner Bachelorarbeit benötigt.

1. Bitte geben Sie Ihr Alter an:

1-18 <input type="radio"/>	19-20 <input type="radio"/>	21-25 <input type="radio"/>	26-30 <input type="radio"/>	31-35 <input type="radio"/>	36-45 <input type="radio"/>	46 oder älter <input type="radio"/>	Enthaltung <input type="radio"/>
-------------------------------	--------------------------------	--------------------------------	--------------------------------	--------------------------------	--------------------------------	--	-------------------------------------

2. Bitte geben Sie Ihr Geschlecht an:

Männlich <input type="radio"/>	Weiblich <input type="radio"/>	Enthaltung <input type="radio"/>
-----------------------------------	-----------------------------------	-------------------------------------

3. Sind sie zum Zeitpunkt dieser Umfrage Student einer Hochschule oder Universität?

Ja <input type="radio"/>	Nein <input type="radio"/>	Enthaltung <input type="radio"/>
-----------------------------	-------------------------------	-------------------------------------

Falls Sie 3. mit Ja beantwortet haben:

a) Welchem Studiengang gehören Sie an? (z.b. Informatik, Maschinenbau, Kunst, etc.)

\_\_\_\_\_

b) Im wievielten Studiensemester befinden Sie sich?

\_\_\_\_\_

Enthaltung <input type="radio"/>
-------------------------------------



4. Wie häufig verwenden Sie einen Computer?

Nie <input type="radio"/>	Selten <input type="radio"/>	0-5h / w <input type="radio"/>	5-15h / w <input type="radio"/>	16-25h / w <input type="radio"/>	26-40h / w <input type="radio"/>	40h+ / w <input type="radio"/>	Enthaltung <input type="radio"/>
------------------------------	---------------------------------	-----------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-----------------------------------	-------------------------------------

5. Wie häufig verwenden Sie einen Touchscreen?

Nie <input type="radio"/>	Selten <input type="radio"/>	0-5h / w <input type="radio"/>	5-15h / w <input type="radio"/>	16-25h / w <input type="radio"/>	26-40h / w <input type="radio"/>	40h+ / w <input type="radio"/>	Enthaltung <input type="radio"/>
------------------------------	---------------------------------	-----------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-----------------------------------	-------------------------------------

6. Wie häufig verwenden Sie eine 3D Gestensteuerung (Kinect oder Ähnliches)?

Nie <input type="radio"/>	Selten <input type="radio"/>	0-5h / w <input type="radio"/>	5-15h / w <input type="radio"/>	16-25h / w <input type="radio"/>	26-40h / w <input type="radio"/>	40h+ / w <input type="radio"/>	Enthaltung <input type="radio"/>
------------------------------	---------------------------------	-----------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-----------------------------------	-------------------------------------

7. Wie häufig arbeiten Sie kollaborativ mit anderen Menschen an einem Anzeigemedium?  
(zusammen an einem Bildschirm eine Aufgabe lesen, zusammen lernen, jemandem etwas erklären oder Ähnliches)

Nie <input type="radio"/>	Selten <input type="radio"/>	0-5h / w <input type="radio"/>	5-15h / w <input type="radio"/>	16-25h / w <input type="radio"/>	26-40h / w <input type="radio"/>	40h+ / w <input type="radio"/>	Enthaltung <input type="radio"/>
------------------------------	---------------------------------	-----------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-----------------------------------	-------------------------------------

8. Wie häufig haben Sie bereits eine Bildercollage zusammengestellt?

Nie <input type="radio"/>	Selten <input type="radio"/>	0-5h / w <input type="radio"/>	5-15h / w <input type="radio"/>	16-25h / w <input type="radio"/>	26-40h / w <input type="radio"/>	40h+ / w <input type="radio"/>	Enthaltung <input type="radio"/>
------------------------------	---------------------------------	-----------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-----------------------------------	-------------------------------------

**Vielen Dank für die Beantwortung des ersten Fragebogens! Bitte füllen Sie den Fragebogen eigenständig ohne Austausch mit anderen Teilnehmern aus und geben Sie ihn ausgefüllt wieder an mich zurück.**

## 7.2 Fragebogen nach dem Experiment



# Fragebogen

nach dem Experiment

Hochschule für Angewandte Wissenschaften Hamburg

Hamburg University of Applied Sciences

Bitte kreuzen Sie für jede der folgenden Fragen nur eine Antwort an. Falls Sie eine Frage nicht beantworten möchten, kreuzen Sie bitte „Enthaltung“ an. Außerdem bitte ich Sie sich die Zeit zu nehmen und bei Punkt 9 & 10 alles niederzuschreiben, was Ihnen jeweils dazu einfällt. Ihre Antworten werden vertraulich behandelt und anonymisiert. Sie werden ausschließlich für statistische Erhebungen innerhalb meiner Bachelorarbeit benötigt.

1. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage ausschließlich mit Touchgesten?

Sehr einfach	1 O	2 O	3 O	4 O	5 O	6 O	Sehr umständlich	Enthaltung O
--------------	--------	--------	--------	--------	--------	--------	------------------	-----------------

2. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage ausschließlich mit 3D Gesten?

Sehr einfach	1 O	2 O	3 O	4 O	5 O	6 O	Sehr umständlich	Enthaltung O
--------------	--------	--------	--------	--------	--------	--------	------------------	-----------------

3. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage mit der Kombination beider Steuerungstypen?

Sehr einfach	1 O	2 O	3 O	4 O	5 O	6 O	Sehr umständlich	Enthaltung O
--------------	--------	--------	--------	--------	--------	--------	------------------	-----------------

4. Welche Steuerung würde Ihrer Meinung nach für die größte Übersichtlichkeit sorgen?

Nur Touchgesten O	Nur 3D Gesten O	Kombination O	Enthaltung O
----------------------	--------------------	------------------	-----------------

5. Welche Steuerung, würde Ihrer Meinung nach für die größte Produktivität sorgen?

Nur Touchgesten O	Nur 3D Gesten O	Kombination O	Enthaltung O
----------------------	--------------------	------------------	-----------------

6. Wie fanden Sie die farbliche Zuordnung der Zeiger sowie der gerade in Interaktion befindlichen Bilder?

Sehr hilfreich	1 O	2 O	3 O	4 O	5 O	6 O	Kaum hilfreich	Enthaltung O
----------------	--------	--------	--------	--------	--------	--------	----------------	-----------------

7. Wie fanden Sie die Arbeit als Team zusammen an einem Anzeigemedium?

Sehr gut	1 O	2 O	3 O	4 O	5 O	6 O	Sehr schlecht	Enthaltung O
----------	--------	--------	--------	--------	--------	--------	---------------	-----------------

8. Wie gut könnten sie sich vorstellen ein solches System welches Touch- sowie 3D Gesten untestützt auch für andere Gruppentätigkeiten und Aufgaben zu verwenden?

Sehr gut	1	2	3	4	5	6	Sehr schlecht	Enthaltung
	<input type="radio"/>		<input type="radio"/>					

9. Anmerkungen und Verbesserungsvorschläge zur Steuerung oder einzelnen Steuerungselementen:

10. Sonstige Anmerkungen:

**Vielen Dank für die Beantwortung des zweiten Fragebogens! Bitte füllen Sie den Fragebogen eigenständig ohne Austausch mit anderen Teilnehmern aus und geben Sie ihn ausgefüllt wieder an mich zurück.**

# Abbildungsverzeichnis

1.1	Zeigt einige Medienfassaden. Quelle:( <a href="#">GHLK13</a> ) . . . . .	1
1.2	Zeigt eine computerunterstützte Konferenz. Quelle: ( <a href="#">BDHM11</a> ) . . . . .	2
2.1	Verschiedene Disziplinen im Bereich HCI. Quelle: (vgl. <a href="#">Bad15</a> , S.81) . . .	7
2.2	Evolution des Ubiquitous Computing. Quelle: (vgl. <a href="#">Riv05</a> , S.5) . . . . .	8
2.3	Kategorisierung von Groupware / CSCW Systemen. Quelle: (vgl. <a href="#">Ger07</a> , S.144) . . . . .	10
2.4	Groupware Time / Space Matrix Quelle: ( <a href="#">Wik16</a> ) . . . . .	11
2.5	Sperrbildschirm der ersten iPhone Generation. Quelle: Eigene Darstellung	12
2.6	Sperrbildschirm des Samsung Galaxy S5. Quelle: Eigene Darstellung . .	13
2.7	Eigenschaften einer Geste - 1. Quelle: ( <a href="#">Vaf13</a> , S.14-15) . . . . .	17
2.8	Eigenschaften einer Geste - 2. Quelle: ( <a href="#">Vaf13</a> , S.14-15) . . . . .	18
2.9	Während eines Experiments observierte ungewollte Gesten. Quelle: ( <a href="#">AAU08</a> , S.358) . . . . .	20
2.10	Verschiedene Gestenvariationen beim Abstützen des Kinns. Quelle: ( <a href="#">AAU08</a> , S.362) . . . . .	20
2.11	2D Gesten auf einer Multitouch Oberfläche. Quelle: ( <a href="#">Lin10</a> ) . . . . .	21
3.1	Skizze eines mit dem Bildschirm interagierenden Teams. Quelle: eigene Darstellung . . . . .	27
3.2	Die benötigten Anwendungsfälle Tabelle 1. Quelle: eigene Darstellung .	28
3.3	Die benötigten Anwendungsfälle Tabelle 2. Quelle: eigene Darstellung .	29
3.4	Arbeitsplätze des CSTI Labors. Quelle: Eigene Darstellung . . . . .	34
3.5	Raumplan des CSTI Labors. Quelle: Eigene Darstellung . . . . .	34
3.6	BenQ RP790. Quelle: Eigene Darstellung . . . . .	34
3.7	Funktionsweise eines IR-Panels aus ( <a href="#">Dix13</a> ) . . . . .	35
3.8	Kinect v2 Spezifikationen im Vergleich. Quelle: msdn.microsoft.com . .	35
3.9	Kinect Sensoren. Quelle: msdn.microsoft.com . . . . .	36

4.1	Test . . . . .	37
4.2	Registrierung mit nach außen ausgestreckten Armen. Quelle: Eigene Darstellung . . . . .	39
4.3	Registrierung mit nach oben ausgestreckten Armen. Quelle: Eigene Darstellung . . . . .	39
4.4	Raycasting - rote Linie zeigt tatsächlich selektierten Punkt auf dem Display. Quelle: Eigene Darstellung . . . . .	40
4.5	imaginäre Fläche - frontal. Quelle: Eigene Darstellung . . . . .	41
4.6	imaginäre Fläche - seitlich Quelle: Eigene Darstellung . . . . .	41
4.7	Zwei voneinander unabhängige imaginäre Flächen vor dem Anwender deren Koordinaten sich auf die Koordinaten des Bildschirmes abbilden lassen. Quelle: Eigene Darstellung . . . . .	42
4.8	Nutzer bewegt Cursor mit rechter Hand. Quelle: Eigene Darstellung . . . . .	43
4.9	Nutzer bewegt Cursor mit linker Hand. Quelle: Eigene Darstellung . . . . .	43
4.10	Nutzer bewegt Objekt mit geschlossener rechter Hand. Quelle: Eigene Darstellung . . . . .	44
4.11	Nutzer bewegt Objekt mit geschlossener linker Hand. Quelle: Eigene Darstellung . . . . .	44
4.12	Nutzer vergrößert Objekt mit Spread Geste. Quelle: Eigene Darstellung . . . . .	44
4.13	Nutzer verkleinert Objekt mit Shrink Geste. Quelle: Eigene Darstellung . . . . .	44
4.14	Nutzer rotiert Objekt gegen den Uhrzeigersinn. Quelle: Eigene Darstellung . . . . .	45
4.15	Nutzer rotiert Objekt mit dem Uhrzeigersinn. Quelle: Eigene Darstellung . . . . .	45
4.16	Nutzer swiped in eine der vier Himmelsrichtungen. Quelle: Eigene Darstellung . . . . .	45
4.17	GUI mit allen Elementen und Zonen. Quelle: Eigene Darstellung . . . . .	47
4.18	GUI zum Hinzufügen von Elementen. Quelle: Eigene Darstellung . . . . .	48
4.19	Diagramm zeigt den Aufbau der einzelnen Systemkomponenten . . . . .	49
5.1	Screenshot der GUI. Quelle: Eigene Darstellung . . . . .	54
5.2	Screenshot der Upload-GUI. Quelle: Eigene Darstellung . . . . .	55
5.3	Skizzierter Aufbau der Testumgebung. Quelle: Eigene Darstellung . . . . .	57
5.4	Aufbau der Testumgebung. Quelle: Eigene Darstellung . . . . .	57
5.5	Kamera in der Traverse. Quelle: Eigene Darstellung . . . . .	58
5.6	Testgruppe diskutiert gemeinsam welche Bilder sie hochladen möchten. Quelle: Eigene Darstellung . . . . .	58

5.7	Testgruppe befüllt erste Seite der Collage mit 2D Gesten. Quelle: Eigene Darstellung . . . . .	58
5.8	Testgruppe befüllt Collage mit 3D Gesten. Quelle: Eigene Darstellung .	60
5.9	Testgruppe befüllt Collage mit 3D Gesten. Quelle: Eigene Darstellung .	60
5.10	Antwortenverteilung: 1. "Das Wetter heute ist..."(Beispieldiagramm). Quelle: Eigene Darstellung . . . . .	62
5.11	Antwortenverteilung: 1. Bitte geben Sie ihr Alter an. Quelle: Eigene Darstellung . . . . .	63
5.12	Antwortenverteilung: 2. Bitte geben Sie ihr Geschlecht an. Quelle: Eigene Darstellung . . . . .	64
5.13	Antwortenverteilung: 3. Sind sie zum Zeitpunkt dieser Umfrage Student einer Hochschule oder Universität? Quelle: Eigene Darstellung . . . . .	64
5.14	Antwortenverteilung: 4. Wie häufig verwenden Sie einen Computer? Quelle: Eigene Darstellung . . . . .	64
5.15	Antwortenverteilung: 5. Wie häufig verwenden Sie einen Touchscreen? Quelle: Eigene Darstellung . . . . .	65
5.16	Anwortenverteilung: 6. Wie häufig verwenden Sie eine 3D Gestensteuerung (Kinect oder Ähnliches)? Quelle: Eigene Darstellung . . . . .	65
5.17	Antwortenverteilung: 7. Wie häufig arbeiten Sie kollaborativ mit anderen Menschen an einem Anzeigemedium (zusammen an einem Bildschirm eine Aufgabe lesen, zusammen lernen, jemandem etwas erklären oder Ähnliches)? Quelle: Eigene Darstellung . . . . .	67
5.18	Antwortenverteilung: 8. Wie häufig haben Sie bereits eine Bildercollage zusammengestellt? Quelle: Eigene Darstellung . . . . .	67
5.19	Antwortenverteilung: 1. Wie fanden Sie die Zusammenarbeit als Grup- pe an der Bildercollage ausschließlich mit Touchgesten? Quelle: Eigene Darstellung . . . . .	68
5.20	Antwortenverteilung: 2. Wie fanden Sie die Zusammenarbeit als Grup- pe an der Bildercollage ausschließlich mit 3D Gesten? Quelle: Eigene Darstellung . . . . .	69
5.21	Antwortenverteilung: 3. Wie fanden Sie die Zusammenarbeit als Gruppe an der Bildercollage mit der Kombination beider Steuerungstypen? Quelle: Eigene Darstellung . . . . .	70
5.22	Antwortenverteilung: 4. Welche Steuerung würde Ihrer Meinung nach für die größte Übersichtlichkeit sorgen? Quelle: Eigene Darstellung . . . . .	71

5.23	Antwortenverteilung: 5. Welche Steuerung, würde Ihrer Meinung nach für die größte Produktivität sorgen? Quelle: Eigene Darstellung . . . . .	71
5.24	Antwortenverteilung: 6. Wie fanden Sie die farbliche Zuordnung der Zeiger sowie der gerade in Interaktion befindlichen Bilder? Quelle: Eigene Darstellung . . . . .	72
5.25	Antwortenverteilung: 7. Wie fanden Sie die Arbeit als Team zusammen an einem Anzeigemedium? Quelle: Eigene Darstellung . . . . .	73
5.26	Antwortenverteilung: 8. Wie gut könnten sie sich vorstellen ein solches System welches Touch- sowie 3D Gesten untestützt auch für andere Gruppentätigkeiten und Aufgaben zu verwenden? Quelle: Eigene Darstellung	73
5.27	Das Innere einer Leap Motion. Quelle: (Col14) . . . . .	78
5.28	Reichweite einer Leap Motion. Quelle: (Col14) . . . . .	79

## Literaturverzeichnis

- [AA09] ARGELAGUET, Ferran ; ANDUJAR, Carlos: Efficient 3D Pointing Selection in Cluttered Virtual Environments. In: *IEEE Comput. Graph. Appl.* 29 (2009), November, Nr. 6, 34–43. <http://dx.doi.org/10.1109/MCG.2009.117>. – DOI 10.1109/MCG.2009.117. – ISSN 0272–1716
- [AAU08] ABBASI, Abdul R. ; AFZULPURKAR, Nitin V. ; UNO, Takeaki: Exploring Un-Intentional Body Gestures for Affective System Design. In: *Affective Computing, InTech* (2008). <http://dx.doi.org/10.5772/6170>. – DOI 10.5772/6170
- [Arc97] ARCHER, Dane: Unspoken Diversity: Cultural Differences in Gestures. In: *Qualitative Sociology* 20 (1997), march, Nr. 1, S. 79–105. – ISSN 0162–0436
- [Bad15] BADHITI, Dr. Kezia R.: HCI - Ubiquitous Computing and Ambient Technologies in the Universe. In: *International Journal of Advance Research in Computer Science and Management Studies* 3 (2015), nov, Nr. 11, 80-85. <http://www.ijarcsms.com/docs/paper/volume3/issue11/V3I11-0020.pdf>. – [Online; accessed 20-June-2016]
- [Bae14] BAECKER, R.M.: *Readings in Human-Computer Interaction: Toward the Year 2000*. Elsevier Science, 2014 (Interactive Technologies). <https://books.google.de/books?id=KA2jBQAAQBAJ>. – ISBN 9780080515748
- [Bar12] BARNKOW, Lorenz: *Kollaboration an Multitouch-Tischen: Visualisierung von Besitz bei Gruppenarbeit*. Hamburg, Germany, Hochschule für Angewandte Wissenschaften Hamburg, Diplomarbeit, 2012. <http://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/master/barnkow.pdf>. – [Online; accessed 21-July-2016]

- [BDHM11] BRAGDON, Andrew ; DELINE, Rob ; HINCKLEY, Ken ; MORRIS, Meredith R.: Code Space: Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings. In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. New York, NY, USA : ACM, 2011 (ITS '11). – ISBN 978–1–4503–0871–7, 212–221
- [Ber11] BERNIN, Arne: *Einsatz von 3D-Kameras zur Interpretation von räumlichen Gesten im Smart Home Kontext*. Hamburg, Deutschland, Hochschule für Angewandte Wissenschaften Hamburg, Diplomarbeit, 2011. <http://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/master/bernin.pdf>. – [Online; accessed 07-July-2016]
- [BG10] BRODBECK, F. C. ; GUILLAUME, Y. R. F.: Arbeit in Gruppen.[Working in groups]. In: *Enzyklopädie der Psychologie, Band Arbeitspsychologie.[Encyclopedia of psychology: Work psychology]* (2010), S. 383–443
- [BMW15] BMW: *BMW Group auf der CES 2016 in Las Vegas. Mit Air-Touch präsentiert BMW das Prinzip des berührungslosen Touchscreens*. <https://www.press.bmwgroup.com/deutschland/article/attachment/T0247964DE/343397>. Version: 2015. – [Online; accessed 07-July-2016]
- [CLB<sup>+</sup>04] CEAPARU, Irina ; LAZAR, Jonathan ; BESSIERE, Katie ; ROBINSON, John ; SHNEIDERMAN, Ben: Determining causes and severity of end-user frustration. In: *International journal of human-computer interaction* 17 (2004), Nr. 3, S. 333–356
- [Col14] COLGAN, Alex: *How Does the Leap Motion Controller Work?* <http://blog.leapmotion.com/hardware-to-software-how-does-the-leap-motion-controller-work/>. Version: 2014. – [Online; accessed 07-July-2016]
- [Com13] COMTET, Hans: *Acceptance of 3D-gestures based on age, gender and experience*. Gjøvik, Norway, Department of Computer Science and Media Technology Gjøvik University College, Diplomarbeit, 2013. <https://brage.bibsys.no/xmlui/bitstream/id/101213/HComtet.pdf>. – [Online; accessed 07-July-2016]

- [CQG<sup>+</sup>11] COCKBURN, A. ; QUINN, P. ; GUTWIN, C. ; RAMOS, G. ; LOOSER, J.: Air Pointing: Design and Evaluation of Spatial Target Acquisition with and Without Visual Feedback. In: *Int. J. Hum.-Comput. Stud.* 69 (2011), Juni, Nr. 6, 401–414. <http://dx.doi.org/10.1016/j.ijhcs.2011.02.005>. – DOI 10.1016/j.ijhcs.2011.02.005. – ISSN 1071–5819
- [Dix13] DIXON, Derek: *THE TECHNOLOGY WITHIN TOUCH PANELS*. <http://www.epectec.com/downloads/Article-The-Technology-Within-Touch-Panels.pdf>. Version: 2013. – [Online; accessed 07-July-2016]
- [Eic14] EICHLER, Tobias: *Agentenbasierte Middleware zur Entwicklerunterstützung in einem Smart-Home-Labor*. Hamburg, Germany, Hochschule für Angewandte Wissenschaften Hamburg, Diplomarbeit, 2014. <http://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/master/eichler.pdf>. – [Online; accessed 04-August-2016]
- [Elm02] ELMENREICH, Wilfried: *Sensor Fusion in Time-Triggered Systems*. Vienna, Austria, Technische Universität Wien, Institut für Technische Informatik, Diss., 2002. [https://mobile.aau.at/~welmenre/papers/elmenreich\\_Dissertation\\_sensorFusionInTimeTriggeredSystems.pdf](https://mobile.aau.at/~welmenre/papers/elmenreich_Dissertation_sensorFusionInTimeTriggeredSystems.pdf). – [Online; accessed 11-July-2016]
- [FKO03] FUHRMANN, T. ; KLEIN, M. ; ODENDAHL, M.: The BlueWand as interface for ubiquitous and wearable computing environments. In: *Personal Mobile Communications Conference, 2003. 5th European (Conf. Publ. No. 492)*, 2003. – ISSN 0537–9989, S. 91–95
- [For05] FORSYTH, Donelson: *Group Dynamics*. Cengage Learning, 2005 (International student edition). <https://books.google.ca/books?id=jXTa7Tbkpf4C>. – ISBN 9780534368227
- [Ger07] GERLICHER, Ansgar: Computer-Supported Cooperative Work (CSCW)—kollaborative Systeme und Anwendungen. In: *Kompendium Medieninformatik*. Springer, 2007, S. 143–195
- [GHLK13] GEHRING, Sven ; HARTZ, Elias ; LÖCHTEFELD, Markus ; KRÜGER, Antonio: The Media Facade Toolkit: Prototyping and Simulating Interaction with Media Facades. In: *Proceedings of the 2013 ACM International*

- Joint Conference on Pervasive and Ubiquitous Computing*. New York, NY, USA : ACM, 2013 (UbiComp '13). – ISBN 978-1-4503-1770-2, 763-772
- [GMPT12] GUPTA, Sidhant ; MORRIS, Daniel ; PATEL, Shwetak ; TAN, Desney: SoundWave: Using the Doppler Effect to Sense Gestures. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2012 (CHI '12). – ISBN 978-1-4503-1015-4, 1911-1914
- [HE97] HARLING, Philip A. ; EDWARDS, Alistair D. N.: Hand Tension As a Gesture Segmentation Cue. In: *Proceedings of Gesture Workshop on Progress in Gestural Interaction*. London, UK, UK : Springer-Verlag, 1997. – ISBN 3-540-76094-6, 75-88
- [HSO97] HUMMELS, Caroline ; SMETS, Gerda ; OVERBEEKE, Kees: An Intuitive Two-Handed Gestural Interface for Computer Supported Product Design. In: *Gesture and Sign Language in Human-Computer Interaction, International Gesture Workshop, Bielefeld, Germany, September 17-19, 1997, Proceedings*, 1997, 197-208
- [II15] ISSA, Tomayess ; ISAIAS, Pedro: *Sustainable Design: HCI, Usability and Environmental Concerns*. 1st. Springer Publishing Company, Incorporated, 2015. – ISBN 144716752X, 9781447167525
- [Jac12] JACKO, Julie A.: *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, Third Edition*. 3rd. Boca Raton, FL, USA : CRC Press, Inc., 2012. – ISBN 1439829438, 9781439829431
- [JJBH15] JAKOBSEN, Mikkel R. ; JANSEN, Yvonne ; BORING, Sebastian ; HORN-BÆK, Kasper: Should I stay or should I go? Selecting between touch and mid-air gestures for large-display interaction. In: *Human-Computer Interaction* Springer, 2015, S. 455-473
- [JNDW13] JOTA, Ricardo ; NG, Albert ; DIETZ, Paul ; WIGDOR, Daniel: How fast is fast enough?: a study of the effects of latency in direct-touch pointing tasks. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* ACM, 2013, S. 2291-2300

- [KH90] KURTENBACH, Gordon ; HULTEEN, Eric A.: Gestures in human-computer communication. In: *The art of human-computer interface design* (1990), S. 309–317
- [KKC<sup>+</sup>15] KIM, Kwangtaek ; KIM, Joongrock ; CHOI, Jaesung ; KIM, Junghyun ; LEE, Sangyoun: Depth Camera-Based 3D Hand Gesture Controls with Immersive Tactile Feedback for Natural Mid-Air Gesture Interactions. In: *Sensors* 15 (2015), Nr. 1, 1022-1046. <http://dx.doi.org/10.3390/s150101022>. – DOI 10.3390/s150101022. – ISSN 1424–8220
- [LGTIK14] LE GOC, Mathieu ; TAYLOR, Stuart ; IZADI, Shahram ; KESKIN, Cem: A Low-cost Transparent Electric Field Sensor for 3D Interaction on Mobile Devices. In: *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2014 (CHI '14). – ISBN 978–1–4503–2473–1, 3167–3170
- [Lin10] LINDEMANN, Patrick: *A Short Report on Multi-Touch User Interfaces*. [https://www.medien.ifi.lmu.de/lehre/ws1011/mmi2/mmi2\\_uebungsblatt1\\_loesung\\_lindemann.pdf](https://www.medien.ifi.lmu.de/lehre/ws1011/mmi2/mmi2_uebungsblatt1_loesung_lindemann.pdf). Version: 2010. – [Online; accessed 25-June-2016]
- [McN92] MCNEILL, D.: *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press, 1992 <https://books.google.de/books?id=3ZZAfNumLvwC>. – ISBN 9780226561325
- [NK15] NUTSI, Andrea ; KOCH, Michael: Multi-User Usability Guidelines for Interactive Wall Display Applications. In: *Proceedings of the 4th International Symposium on Pervasive Displays*. New York, NY, USA : ACM, 2015 (PerDis '15). – ISBN 978–1–4503–3608–6, 233–234
- [Riv05] RIVA, G.: *Ambient Intelligence: The Evolution of Technology, Communication and Cognition Towards the Future of Human-computer Interaction*. IOS, 2005 (Emerging communication). <https://books.google.de/books?id=sACug4nZDmUC>. – ISBN 9781586034900
- [RLL11] RUIZ, Jaime ; LI, Yang ; LANK, Edward: User-defined Motion Gestures for Mobile Interaction. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2011 (CHI '11). – ISBN 978–1–4503–0228–9, 197–206

- [Sat01] SATYANARAYANAN, M.: Pervasive Computing: Vision and Challenges. In: *IEEE Personal Communications* 8 (2001), S. 10–17
- [SBD<sup>+</sup>08] SCHÖNING, Johannes ; BRANDL, Peter ; DAIBER, Florian ; ECHTLER, Florian ; HILLIGES, Otmar ; HOOK, Jonathan ; LÖCHTEFELD, Markus ; MOTAMEDI, Nima ; MULLER, Laurence ; OLIVIER, Patrick u. a.: Multi-touch surfaces: A technical guide. In: *IEEE Tabletops and Interactive Surfaces* 2 (2008), S. 11
- [Sch95] SCHWABE, Gerhard: *Objekte der Gruppenarbeit: Ein Konzept für das Computer Aided Team*. Deutscher Universitätsverlag, 1995 (Informationsmanagement und Computer Aided Team). <https://books.google.de/books?id=aar0BgAAQBAJ>. – ISBN 9783663011804
- [SLM<sup>+</sup>03] STARNER, Thad ; LEIBE, Bastian ; MINNEN, David ; WESTYN, Tracy ; HURST, Amy ; WEEKS, Justin: The perceptive workbench: Computer-vision-based gesture tracking, object tracking, and 3D reconstruction for augmented desks. In: *Machine Vision and Applications* 14 (2003), Nr. 1, S. 59–71
- [Vaf13] VAFAEI, Fereydoon: *TAXONOMY OF GESTURES IN HUMAN COMPUTER INTERACTION*. Fargo, North Dakota, North Dakota State University of Agriculture and Applied Science, Diplomarbeit, 2013. [http://library.ndsu.edu/tools/dspace/load/?file=/repository/bitstream/handle/10365/23110/Vafaei\\_Taxonomy%20of%20Gestures%20in%20Human%20Computer%20Interaction.pdf?sequence=1](http://library.ndsu.edu/tools/dspace/load/?file=/repository/bitstream/handle/10365/23110/Vafaei_Taxonomy%20of%20Gestures%20in%20Human%20Computer%20Interaction.pdf?sequence=1). – [Online; accessed 29-May-2016]
- [VTA<sup>+</sup>15] VELLOSO, Eduardo ; TURNER, Jayson ; ALEXANDER, Jason ; BULLING, Andreas ; GELLERSEN, Hans: An empirical investigation of gaze selection in mid-air gestural 3D manipulation. In: *Human-Computer Interaction* Springer, 2015, S. 315–330
- [Wei91] WEISER, Mark: The computer for the 21st century. In: *Scientific american* 265 (1991), Nr. 3, S. 94–104
- [Wik16] WIKIPEDIA: *Computer-supported cooperative work* — *Wikipedia, The Free Encyclopedia*. <https://en.wikipedia.org/w/index.php?>

- [title=Computer-supported\\_cooperative\\_work&oldid=720647603](#).  
Version: 2016. – [Online; accessed 24-June-2016]
- [Wil91] WILSON, P.: *Computer Supported Cooperative Work:: An Introduction*. Springer Netherlands, 1991 <https://books.google.de/books?id=-EMqvaITnEOC>. – ISBN 9780792314462
- [WMW09] WOBROCK, Jacob O. ; MORRIS, Meredith R. ; WILSON, Andrew D.: User-defined Gestures for Surface Computing. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2009 (CHI '09). – ISBN 978-1-60558-246-7, 1083–1092
- [YTCYPH<sup>+</sup>12] YEN-TING, Liu ; CHE-YANG, Wu ; PI-HSUN, Shih ; HSIANG-SHENG, Liang ; MIKE, Y. C.: *SoundSense: 3D Gesture Sensing using Ultrasound on Mobile Devices*. <http://mrorz.github.io/files/soundsense.pdf>.  
Version: 2012. – [Online; accessed 5-July-2016]

*Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.*

Hamburg, 07. November 2016

---

Krystian Graczyk