

3D-Modellierung mit interaktiven Oberflächen

Dem Fachbereich 03 - Mathematik/Informatik der
Universität Bremen
zur Erlangung des akademischen Grades eines
Dr.-Ing.

eingereichte Dissertation

von
Marc Herrlich

Datum der Einreichung: 2013/03/28

1. Gutachter: Prof. Dr. Rainer Malaka
2. Gutachter: Prof. Dr. Antonio Krüger

Tag der mündlichen Verteidigung: 2013/04/17

Für meine Eltern

Für Antje

Danksagung

Vor allen anderen möchte ich meinem Betreuer, Chef und Kollegen Rainer Malaka dafür danken, dass er diese Arbeit überhaupt erst ermöglicht hat. Danke für die Unterstützung und den „Schubs“ zum Fertigmachen. Danke auch für das in mich gesetzte Vertrauen und den Freiraum, den ich bei der Bearbeitung dieses Themas stets genießen durfte.

Danken möchte ich auch meinen Gutachtern, insbesondere Antonio Krüger, der selbst vom anderen Ende der Welt aus, ohne zu zögern, die Begutachtung dieser Arbeit übernommen hat. Danke für die faire und fristgemäße Begutachtung und das Feedback.

Weiterhin möchte ich den übrigen Mitgliedern des Prüfungsausschusses danken: Gabriel Zachmann, Ute Bormann, Benjamin Walther-Franks und Dirk Wenig.

Dank gebührt natürlich auch den Studenten, die ich während meiner Promotionszeit betreuen durfte und deren Arbeit maßgeblich zum Gelingen dieser Dissertation beigetragen hat (in alphabetischer Reihenfolge): Markus Aust, Andreas Braun, Jan Holthusen und Jannis Stoppe.

Ein großes Dankeschön möchte ich auch an die gesamte Arbeitsgruppe Digitale Medien (Ehemalige eingeschlossen) richten. Ich habe es stets als Privileg empfunden, zu dieser Gruppe zu gehören und möchte mich für die vielen inspirierenden Besprechungen, Flur- und Messagespräche und die großartige Unterstützung in jeder Form bedanken. Meinen besonderen Dank möchte ich richten an (in alphabetischer Reihenfolge): Hidir ‘van Helsing’ Aras, Lutz Dickmann, Tanja Döring, Martin Faust, Sebastian Feige, Gesa Friederichs-Büttner, Christoph Grimmer-Dietrich, Markus Krause, Meike Langer, Irmgard Laumann, Frederic Pollmann, Robert Porzel, Nina Runge, Roland Schröder-Kroll, Lasse Schwarten, Jan Smeddinck, Aneta Takhtamysheva, Jens Teichert, Gerald Volkmann, Benjamin Walther-Franks, Insa Warms und Dirk Wenig.

Ich danke meinen Korrekturlesern, insbesondere meinen Eltern, meinem Bruder Bastian, sowie Dominique Jessen und Sascha Werhahn für ihr hilfreiches Feedback.

Darüber hinaus danke ich allen, die direkt oder indirekt zum Gelingen dieser Arbeit beigetragen haben. Über die Jahre durfte ich viele interessante und inspirierende Menschen kennenlernen und ich fürchte, den einen oder anderen Namen habe ich beim Aufstellen dieser Liste sicher übersehen.

Ganz besonders möchte ich meiner Familie und meiner Frau Antje danken. Ihr habt mich stets unterstützt und an mich geglaubt. Danke für die Geduld und das Verständnis, wenn ich andere Dinge mal wieder auf „nach meiner Arbeit“ verschoben habe.

Abstract

3D models are at the core of many important applications from industry, science, and also entertainment. The creation of 3D models is a complex and time consuming process. Current modeling tools are hard to learn and require a deep understanding of the underlying mathematical models. Furthermore, established input devices like the mouse and keyboard do not utilize the full interaction potential – especially regarding bimanual control – of the human hand. The growing interest and the commercial breakthrough of multi-touch displays and interactive surfaces raises questions about their potential in the context of 3d modeling, which are thoroughly discussed and evaluated in this work.

The presented approach is closely aligned to the whole processing chain for multi-touch applications, starting with the hardware and tracking issues, continuing with fundamental design discussions and operations like selection and 3D manipulation of objects and finishing with complex modeling techniques and metaphors.

In regard to hardware and tracking, a robust illumination setup for the diffuse illumination technique is presented along with two extensions of this approach, i.e., hover detection and hand distinction. The design space is organized into specific design dimensions characterized by extremal positions to allow a better overview of design choices and a classification of existing and future systems. Fundamental techniques for selection and integrated 3D manipulation with six degrees of freedom are presented and empirically evaluated. Finally, two established modeling techniques – implicit surfaces and virtual sculpting – are extended and evaluated for multi-touch input.

Keywords: 3d modeling, multi-touch, selection, interaction, implicit surfaces, virtual sculpting

Zusammenfassung

3D-Modelle bilden die Grundlage vieler Anwendungen aus Industrie, Wissenschaft und der Unterhaltungsbranche. Die Erstellung von 3D-Modellen ist ein aufwendiger und zeitraubender Prozess. Aktuelle Modellierungswerkzeuge sind schwierig zu lernen und erfordern ein tiefes Verständnis der zu Grunde liegenden mathematischen Modelle. Außerdem schöpfen die etablierten Eingabegeräte wie Maus und Tastatur die Interaktionsmöglichkeiten der Hände – insbesondere der beidhändigen Interaktion – noch unzureichend aus. Mit dem wachsenden Interesse und dem kommerziellen Durchbruch von Multi-touch-Bildschirmen und interaktiven Oberflächen stellt sich die Frage, ob diese neuen Eingabegeräte vorteilhaft im Kontext der 3D-Modellierung eingesetzt werden können. Im Rahmen der vorliegenden Arbeit wird diese Frage eingehend betrachtet und untersucht.

Das Vorgehen orientiert sich dabei an der gesamten Verarbeitungskette für Multi-touch-Anwendungen, angefangen bei der Hardware und der Erkennung über eine grundsätzliche Gliederung und Diskussion des Gestaltungsraums und elementare Operationen für Selektion und 3D-Manipulation von Objekten bis hin zu komplexen Modellierungstechniken und -metaphern.

Im Bereich der Hardware und der Erkennung wird ein robuster Beleuchtungsaufbau für die Diffuse Illumination-Technik vorgestellt. Außerdem werden zwei Erweiterungen auf Basis dieser Technik für die Nahbereichserkennung über der Oberfläche (Hovering) und die Unterscheidung der Hände vorgestellt. Der Gestaltungsraum wird anhand von Gestaltungsdimensionen, welche durch bestimmte Extrempositionen charakterisiert sind, geordnet, um so einen besseren Überblick der notwendigen Designschwerpunkte bzw. -kompromisse und eine Klassifizierung existierender Techniken zu ermöglichen. Als elementare Operationen werden Techniken für die Selektion und integrierte Manipulation in sechs Freiheitsgraden präsentiert und empirisch evaluiert. Schließlich werden zwei etablierte Modellierungsansätze – implizite Oberflächen bzw. virtuelles Sculpting – für den Einsatz auf Multi-touch-Bildschirmen erweitert und ebenfalls empirisch evaluiert.

Schlagwörter: 3D-Modellierung, Multi-touch, Selektion, Interaktion, Implizite Flächen, Virtuelles Sculpting

Inhaltsverzeichnis

1	Einleitung	1
1.1	Fragestellung	3
1.2	Eigene Beiträge	4
1.3	Aufbau dieser Arbeit	5
1.4	Veröffentlichungen	6
2	Stand der Technik	9
2.1	Mathematische Modelle für 3D-Objekte	9
2.2	Werkzeuge für die 3D-Modellierung	13
2.3	Einordnung dieser Arbeit	16
3	Multi-touch-Hardware und -Software	17
3.1	Einleitung	17
3.2	Stand der Technik	18
3.2.1	Hardware	19
3.2.2	Software	25
3.3	Diffuse Illumination-Beleuchtungskonzept	29
3.4	Hover-Detektion	38
3.4.1	Ansatz und Implementierung	39
3.4.2	Testlauf	41
3.4.3	Ergebnisse	43
3.4.4	Diskussion und Ausblick	44
3.5	Handunterscheidung	46
3.5.1	Verwandte Arbeiten	46
3.5.2	Eigenschaften der menschlichen Hand	47
3.5.3	Empirisches Modell	48

3.5.4	Implementierung	49
3.5.5	Testlauf	50
3.5.6	Ergebnisse	51
3.5.7	Diskussion und Ausblick	52
3.6	Fazit	53
4	Interfacegestaltungsraum	55
4.1	Einleitung	55
4.2	Stand der Technik	56
4.2.1	Klassifikation von Eingabegeräten	56
4.2.2	Klassifikation von Gesten	60
4.2.3	Interfacegestaltung	60
4.3	Designdimensionen	63
4.3.1	Multi-point vs. Multi-touch	63
4.3.2	Statischer vs. dynamischer Kontext	67
4.3.3	Integrierte vs. separierte Freiheitsgrade	73
4.3.4	Direkte vs. indirekte Interaktion	76
4.3.5	Mehrhändige Interaktionsmuster	80
4.4	Feedback	86
5	Objektmanipulation in 3D-Szenen	89
5.1	Selektion	89
5.1.1	Stand der Technik	90
5.1.2	Proxybasierte Selektion	91
5.1.3	Experiment	95
5.1.4	Ergebnisse	96
5.1.5	Diskussion	100
5.1.6	Fazit	103
5.2	Manipulation in sechs Freiheitsgraden	104
5.2.1	Stand der Technik	106
5.2.2	Interaktionstechniken	107
5.2.3	Experiment	110
5.2.4	Ergebnisse	113
5.2.5	Diskussion	116
5.2.6	Fazit	117

6	Modellierungsansätze	119
6.1	Implizite Flächen	119
6.1.1	Stand der Technik	122
6.1.2	Metaball-Metapher	124
6.1.3	Interaktionsdesign und Gesten	125
6.1.4	System und Softwareintegration	132
6.1.5	Experiment	134
6.1.6	Ergebnisse	136
6.1.7	Diskussion	137
6.1.8	Fazit	139
6.2	Virtuelles Sculpting	140
6.2.1	Stand der Technik	142
6.2.2	Interaktionsdesign und Gesten	143
6.2.3	Experiment	145
6.2.4	Ergebnisse	149
6.2.5	Diskussion	152
6.2.6	Fazit	153
7	Fazit	155
7.1	Diskussion der eigenen Beiträge	156
7.2	Ergebnis bzgl. der wissenschaftlichen Fragestellung	157
7.3	Offene Fragen und Ausblick	157

Abbildungsverzeichnis

1.1	Graphische Benutzerschnittstelle der freien 3D-Modellierungssoftware Blender – Auswahl und Anwendung verschiedener Werkzeuge erfolgen über eine Vielzahl von Buttons. Es stehen verschiedene Ansichten und Perspektiven eines Objektes zur Verfügung.	2
1.2	Schematische Übersicht zur Vorgehensweise dieser Arbeit – Ansatz ist die Behandlung verschiedener wichtiger Themen entlang der gesamten Verarbeitungskette von der Hardware über elementare Operationen bis zu komplexen Modellierungsansätzen.	6
2.1	Beispiel für die Modellierung von Objekten mit polygonalen Netzen – Das dargestellte Modell ist ein Standard-Objekt aus dem 3D-Modellierungstool Blender und verwendet sowohl Dreiecke (z.B. für die Augen) als auch Vierecke zur Approximation der Objektoberfläche.	10
2.2	Metaballs oder Blobbies sind eine Anwendung von impliziten Funktionen bzw. Oberflächen in der 3D-Modellierung – Der Modellierung liegt kein physikalisches Modell zu Grunde, dennoch weist die Modellierung und die Darstellung Ähnlichkeiten zu der Arbeit mit Knete oder Ton in der realen Welt auf. Es ist zu sehen, wie die einzelnen Metaballs in Abhängigkeit zur Entfernung <i>verschmelzen</i> und auf diese Weise ein komplexeres Modell bilden.	11
2.3	Beispiel für Free Form Deformation – Das komplexe polygonale Netz des Zielobjektes (vgl. Abbildung 2.1) kann über die acht Kontrollpunkte des umgebenden Quaders auf einfache Weise manipuliert werden.	14

2.4	Virtuelles oder digitales Sculpting in Blender – Es stehen verschiedene virtuelle Werkzeuge z.B. zum Eindrücken oder Herausziehen der Oberfläche zur Verfügung (links oben im Bild).	15
3.1	Prinzip der Frustrated Total Internal Reflection für die Nutzung in Multi-touch-Bildschirmen (Abbildung aus Han (2005)) – Durch den Oberflächenkontakt mit dem Finger kann das Infrarotlicht austreten und als heller Bereich durch eine entsprechende Kamera auf der Unterseite detektiert werden.	21
3.2	Prinzip der Diffuse Illumination (Abbildung aus Schöning u. a. (2008)) – Bedingt durch den Diffusor auf der Oberfläche, nimmt die Menge des reflektierten Lichtes mit Abstand zur Oberfläche schnell ab. Dadurch erscheinen Objekte direkt auf der Oberfläche sehr viel heller und schärfer und können als hellere Gebiete vor dem Hintergrund erkannt und isoliert werden.	23
3.3	Ablauf der Kommunikation zwischen Hard- und Softwarekomponenten eines Multi-touch-Systems	26
3.4	Bildverarbeitungs pipeline für Diffuse Illumination – Die Reihenfolge der Schritte variiert zum Teil zwischen verschiedenen Implementierungen. Der Verarbeitungsschritt <i>Frequenzfilter</i> ist optional und nicht in allen Erkennern vorhanden.	28
3.5	Fotografie des mrT (Abbildung aus Teichert u. a. (2010))	30
3.6	Schematische Darstellung des Aufbaus des mrT (Abbildung aus Teichert u. a. (2010)) – Die Projektion bzw. der Sichtbereich der Kamera werden über drei optische Spiegel gelenkt. Die Infrarotbeleuchtung ist in den oberen Bereich des Gehäuses integriert und beleuchtet jeweils die gegenüberliegende Seite der Oberfläche.	32
3.7	Originalbild (Rohdaten) der Interaktionsoberfläche aufgenommen durch die interne Infrarotkamera (Abbildung aus Teichert u. a. (2010))	33
3.8	Hände auf der Oberfläche aufgenommen durch die interne Infrarotkamera (Abbildung aus Teichert u. a. (2010))	33
3.9	Extraktion verschiedener Features durch die Anwendung unterschiedlicher Schwellwerte auf Abbildung 3.8 (Abbildung aus Teichert u. a. (2010))	34

3.10	Schematische Darstellung der IR-Scheinwerferschaltung (Abbildung aus Teichert u. a. (2010))	36
3.11	Spiegel an den kurzen Seiten verlängern optisch die Scheinwerferzeilen (Abbildung aus Teichert u. a. (2010)).	36
3.12	Die Anordnung der Oberfläche, der IR-Beleuchtung, der Kamera und des Projektors muss die grundlegenden Anforderungen wie im Schema dargestellt einhalten, um ungewollte Reflexionen zu verhindern (Abbildung aus Teichert u. a. (2010)).	37
3.13	Das beschriebene Verfahren zur Hover-Detektion wurde in einem eigenständigen Erkennen zu Testzwecken implementiert (Abbildung aus Stoppe (2009)). Im Foto ist zu erkennen, dass die Fingerberührung links und der nicht ganz aufgesetzte Finger rechts korrekt erkannt und klassifiziert werden.	41
3.14	Bildschirmfoto der Testapplikation (Abbildung aus Stoppe (2009)) – Die kreisförmigen Testobjekte sollen per Hovering von der linken in die rechte Zone bewegt werden.	42
3.15	Erzielte Lernkurven für die Handunterscheidung (Abbildung aus Walther-Franks u. a. (2011a))	52
4.1	Zustandsmodelle für Touch- und Mausinteraktion nach Buxton bzw. Hinckley und Wigdor (Abbildung übersetzt aus Hinckley und Wigdor (2012))	58
4.2	Schematische Darstellung aller diskutierten Designdimensionen und ihrer Beziehungen untereinander in Auswahl und Verarbeitung der Eingabedaten	64
4.3	Schematische Darstellung zu der Gestaltungsdimension <i>Multi-point vs. Multi-touch</i>	66
4.4	Schematische Darstellung zu der Gestaltungsdimension <i>statischer vs. dynamischer Kontext</i>	69
4.5	Schematische Darstellung zu der Gestaltungsdimension <i>integrierte vs. separierte Freiheitsgrade</i>	73
4.6	Schematische Darstellung zu der Gestaltungsdimension <i>direkte vs. indirekte Interaktion</i>	77
4.7	Schematische Darstellung zu der Gestaltungsdimension <i>mehrhändige Interaktionsmuster</i>	83

5.1	Screenshots der proxybasierten Techniken für verdeckte (links) und dynamische Objekte (rechts) – Die Bewegungsrichtung der Proxyobjekte ist durch rote Pfeile angezeigt. In den dargestellten Beispielen ist <i>Farbübereinstimmung</i> und <i>Bewegungsübereinstimmung</i> aktiviert (Abbildung aus Herrlich u. a. (2011b)).	95
5.2	Mittelwerte der Selektionszeiten je Technik und Szenario	97
5.3	Mittelwerte der Fehleranzahlen je Technik und Szenario	99
5.4	Kombinierter Mittelwert der Selektionszeiten aller Proxytechniken im Vergleich zur Referenztechnik ohne Proxies	101
5.5	Kombinierter Mittelwert der Fehleranzahlen aller Proxytechniken im Vergleich zur Referenztechnik ohne Proxies	102
5.6	Gesten für die 3D-Translation und -Rotation mit den vorgestellten Techniken. Alle Zweifingergesten können auch mit mehr als zwei Fingern ausgeführt werden. Es wird nur die einhändige Bedienung gezeigt, die Bedienung mit beiden Händen ist aber ebenfalls möglich (Abbildung aus Herrlich u. a. (2011a)).	105
5.7	Ausschnitt aus einem Bildschirmfoto der Versuchsumgebung – Das Tetraeder im Vordergrund muss unter Anwendung der jeweiligen Technik mit dem (leicht transparenten) Zieltetraeder möglichst genau in Deckung gebracht werden. Schatten auf der Bodenebene dienen als visuelle Hilfen für die korrekte Platzierung in der Tiefendimension.	112
5.8	Mittelwerte der Abschlusszeiten der Aufgaben für beide Techniken	114
5.9	Mittelwerte der Abschlusszeiten der Aufgaben für unterschiedliche Zielpositionen für PieRotate	115
6.1	Indirekte beidhändige Objektmanipulation mit <i>auto-constraints</i> (die Constraint-Achse wurde für die Abbildung farblich hervorgehoben)	126
6.2	Indirekte gleichzeitige beidhändige Kamerakontrolle und Objektmanipulation mit <i>object pinning</i>	126
6.3	Links: Direkte Selektion und Manipulation; Rechts: Indirekte Manipulation eines selektierten Objektes	127
6.4	Links: Kameraschwenk (zwei Finger) Mitte: Kamerarotation (drei Finger) Rechts: Kamerazoom (vier Finger)	129

6.5	Direkte beidhändige Kamerakontrolle und Objektmanipulation (<i>object pinning</i>)	131
6.6	Direkte beidhändige Objektmanipulation mit <i>auto-constraints</i> (für die Abbildung farblich zusätzlich hervorgehoben)	132
6.7	Die für den Versuch verwendete Modellierungsumgebung: Im oberen und unteren Bereich der 3D-Ansicht ist die Palette der positiven bzw. negativen Metaballs, die den Teilnehmern zur Verfügung stand, zu erkennen.	135
6.8	Beispiele für die Ergebnisse, die von den Teilnehmern während des Versuchs erstellt wurden	136
6.9	Screenshot zur Illustration von virtuellem Sculpting (Abbildung aus Herrlich u. a. (2012a)) – Abgebildet ist die Modellierungs- umgebung, wie sie dem Benutzer präsentiert wird. Das Ziel- objekt befindet sich in der Mitte des Bildschirms; auf der lin- ken Seite kann der Benutzer verschiedene Werkzeuge („Brushes“) auswählen. Modelliert wird durch das (wiederholte) Anwenden der aktiven Brush an den gewünschten Stellen des Zielobjektes.	141
6.10	Foto des experimentellen Aufbaus (anonymisiert)	147
6.11	Verschiedene Ansichten des Modells, welches den Teilnehmern für das Experiment als Ziel vorgegeben wurde (Abbildung aus Herrlich u. a. (2012a)).	148

Tabellenverzeichnis

3.1	Versuchsergebnisse aus mehreren Durchläufen mit einem Benutzer der Testapplikation (Tabelle aus Stoppe (2009))	43
5.1	Einordnung der proxybasierten Selektionstechniken gemäß den Gestaltungsdimensionen aus Kapitel 4	103
5.2	Einordnung der Techniken PieRotate und Turn&Roll gemäß den Gestaltungsdimensionen aus Kapitel 4	117
6.1	Einhändige Objektselektion und -manipulation	128
6.2	Gesten für die einhändige Kamerakontrolle	130
6.3	QUESI Ergebnisse pro Teilnehmer – Skala von 1,0 (niedrig) bis 5.0 (hoch)	138
6.4	Einordnung der Metaball-Interaktionstechniken gemäß den Gestaltungsdimensionen aus Kapitel 4	139
6.5	Mehrfingergesten für das Sculpting (Tabelle aus Herrlich u. a. (2012a))	145
6.6	Interfaceeinstellungen für Stift-Tablet und Tastatur. Links: Die auszuführende Operation; Rechts: Stift-/Tastatur-Aktion, um die jeweilige Operation auszuführen (Tabelle aus Herrlich u. a. (2012a))	147
6.7	Einordnung der Virtual Sculpting-Interaktionstechniken gemäß den Gestaltungsdimensionen aus Kapitel 4	153

Kapitel 1

Einleitung

Die 3D-Computergrafik ist heute weder aus dem Alltag noch aus vielen Bereichen der Wissenschaft und Industrie wegzudenken. Geeignete 3D-Modelle bilden dabei die Grundlage der meisten Anwendungen und somit haben geeignete Techniken und Werkzeuge für die 3D-Modellierung über die letzten Jahre zunehmend an Bedeutung gewonnen.

Die Anwendungsbereiche von 3D-Computergrafik bzw. 3D-Modellen sind vielfältig. Diese reichen von Geoinformationssystemen, CAD/CAM, allgemeinen Visualisierungsanwendungen, Animationsfilmen, Spezialeffekten für Film und Fernsehen, Computer- und Videospiele sowie Anwendungen im Web und für virtuelle Welten bis hin zu 3D-Druckern und Simulationsanwendungen. Ebenso vielfältig und unterschiedlich wie die Anwendungen sind die Anforderungen an die Modelle und die eingesetzten Werkzeuge und Techniken.

Die intensive Forschung in den letzten 40 Jahren hat vielfältige Ansätze und Werkzeuge hervorgebracht, von denen einige den Weg in aktuelle 3D-Modellierungsanwendungen gefunden haben. Trotz dieser enormen Entwicklung und der großen Verbreitung von 3D-Modellen, gehören die 3D-Modellierung und die zugehörigen Werkzeuge aus Anwendersicht zu den komplexesten Softwareanwendungen, die es gibt.

Etablierte Softwarewerkzeuge wie 3ds Max, Maya oder die freie Software Blender (vgl. Abbildung 1.1) bieten zwar eine Vielzahl unterschiedlicher Werkzeuge und Modellierungsmethoden, erfordern aber insgesamt einen großen Lernaufwand und viel Hintergrundwissen und Verständnis des zu Grunde liegenden mathematischen Modells. Der Benutzer muss sich z.B. mit der Topologie

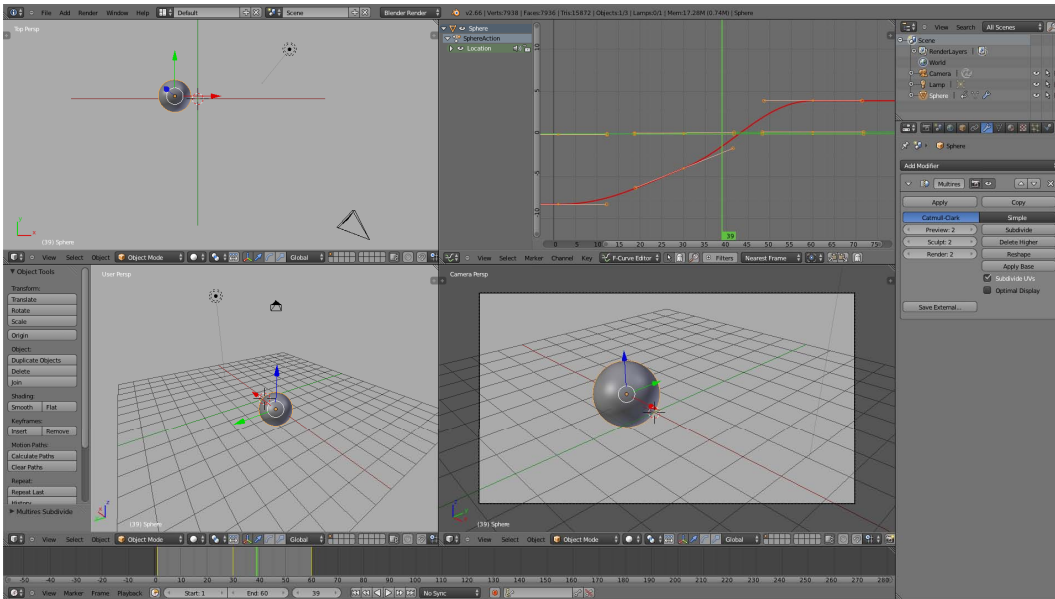


Abbildung 1.1: Graphische Benutzerschnittstelle der freien 3D-Modellierungssoftware Blender – Auswahl und Anwendung verschiedener Werkzeuge erfolgen über eine Vielzahl von Buttons. Es stehen verschiedene Ansichten und Perspektiven eines Objektes zur Verfügung.

polygonaler Meshes oder ähnlich abstrakten Themen auseinander setzen. Darüber hinaus sind aktuelle Softwareanwendungen nach dem WIMP-Paradigma gestaltet und machen kaum von den Möglichkeiten von Multi-touch (MT)-Eingabegeräten bzw. interaktiven Oberflächen Gebrauch.

Die Arbeitsweise bei der Modellierung ist durch einen hohen Grad der Serialisierung geprägt. Ein typisches Interaktionsmuster ist beispielsweise ein ständiger Wechsel zwischen verschiedenen Operationen und Werkzeugen mit der Maus und die Anwendung der Werkzeuge in kleinen Inkrementen, z.B. Rotieren der Kamera, Manipulation des Objektes, Rotieren der Kamera, Manipulation des Objektes usw. Außerdem ist die Interaktion von einer einhändigen Arbeitsweise geprägt, da diese in der Regel nur mit dem Stift oder der Maus mit einer Hand ausgeführt werden können. Die gleichzeitige Verwendung der zweiten Hand, z.B. zum Betätigen von bestimmten Funktionstasten auf der Tastatur hilft zwar, den Wechsel zwischen verschiedenen Operationen zu beschleunigen, ändert aber nichts an der grundlegend seriellen Ausführung von Operationen.

Betrachtet man die komplexen Interaktionsmuster von Menschen in der physischen Welt, so könnte man sagen, dass etablierte Modellierungssoftware

im Vergleich *mit einer Hand auf dem Rücken* bedient wird. Hauptmotivation und Anliegen dieser Arbeit ist es, diese Arbeitsweise zu verändern. Hierzu wird vor allem von den Möglichkeiten interaktiver Oberflächen bzgl. der direkten beidhändigen Interaktion Gebrauch gemacht.

Ein weiterer wichtiger Impuls zum Verfassen dieser Arbeit ist die Faszination von den Möglichkeiten und den Entwicklungen neuer Eingabegeräte. Das Feld der 3D-Modellierung erscheint als aussichtsreicher Kandidat, um die Möglichkeiten und die Nutzererfahrung dieser Geräte auszuloten.

Diese Arbeit ist durch einen breiten Ansatz geprägt. D.h. ausgehend von den Möglichkeiten und Techniken für den Bau von MT-Hardware und -Software, wird der Gestaltungsraum für MT-Interfaces zunächst allgemein, wenn auch immer im Hinblick auf die 3D-Modellierung diskutiert. In einem nächsten Schritt werden grundlegende Operationen näher betrachtet und empirisch untersucht und schließlich werden zwei verschiedene Modellierungstechniken in einer komplexen Modellierungsumgebung im Rahmen von Fallstudien mit Probanden getestet.

1.1 Fragestellung

Die übergeordnete wissenschaftliche Fragestellung dieser Arbeit lautet: Kann MT gewinnbringend für die 3D-Modellierung eingesetzt werden und falls ja, wie?

Aus dieser Fragestellung leiten sich eine Reihe von untergeordneten Detailfragen ab, die im Rahmen dieser Arbeit sowohl theoretisch als auch praktisch näher untersucht und soweit möglich geklärt werden sollen:

- Was ist der Stand der Technik bzgl. Hardware und Software im Bereich MT? Welche Hardwareplattformen erscheinen nach Abwägung der Vor- und Nachteile für die geplanten Untersuchungen am besten geeignet?
- Welche Erweiterungen wären nach dem derzeitigen Stand der Technik insbesondere im Hinblick auf Werkzeuge für die 3D-Modellierung wünschenswert? Diese sollen ggfs. als Proof of Concept umgesetzt werden.
- Was unterscheidet MT wesentlich von der Interaktion mit etablierten Eingabegeräten wie Maus und Tastatur? Wie kann dieser komplexe Ge-

staltungsraum geordnet und organisiert werden, um das Interaktionsdesign zu unterstützen bzw. die einzugehenden Kompromisse explizit deutlich zu machen, um diese in geeigneter Weise abwägen zu können?

- Wie können grundlegende Operationen der Modellierung wie Selektion und Objektmanipulation mit geeigneten MT-Techniken ausgeführt werden? Profitieren diese Techniken von beidhändiger Interaktion?
- Welche grundsätzlichen Modellierungstechniken/-metaphern können auf MT erfolgreich eingesetzt werden? Welche Ergänzungen erlaubt MT für erweiterte Techniken? Können diese insbesondere von beidhändiger Interaktion profitieren? Welchen Einfluss hat MT auf die Nutzererfahrung und/oder -präferenz? Muss MT-Interaktion in diesem Zusammenhang erlernt werden oder hat diese eher Walk-up-and-use-Charakter?

1.2 Eigene Beiträge

In dieser Arbeit wird das Feld der 3D-Modellierung auf interaktiven Oberflächen entlang der gesamten Verarbeitungskette von den Eingabegeräten, über den Gestaltungsraum und elementare Operationen bis hin zur Untersuchung verschiedener Modellierungsmetaphern betrachtet.

Die Hauptbeiträge dieser Arbeit zum derzeitigen Stand der Technik sind:

- Die Konstruktion und Diskussion eines MT-Tisches auf Diffuse Illumination (DI)-Basis, mit empirischen Beiträgen zur Verbesserung der Ausleuchtung.
- Grundlegende Ansätze, um die Interaktionsmöglichkeiten auf Basis der DI-Technik in den Bereichen Hovering und Handunterscheidung zu erweitern.
- Eine systematische Gliederung des umfangreichen Gestaltungsraums und der Forschungsergebnisse in diesem Bereich anhand von Gestaltungsdimensionen, um Designentscheidungen besser einzuordnen und zu verstehen.

- Elementare Techniken für die Selektion von Objekten mit Hilfe von Proxies und für die integrierte 3D-Manipulation in sechs Freiheitsgraden auf interaktiven Oberflächen.
- Beidhändige Interaktionstechniken für zwei verschiedene Modellierungsansätze (Metaballs und Virtual Sculpting).

1.3 Aufbau dieser Arbeit

Diese Arbeit ist wie folgt strukturiert (siehe auch Abbildung 1.2):

- Kapitel 2 diskutiert computergrafische Grundlagen und geht kurz auf alternative Ansätze ein, die nicht unmittelbar im Fokus dieser Arbeit stehen, aber relevant für die Einordnung der folgenden Beiträge sind.
- Kapitel 3 diskutiert zunächst ausführlich den Stand der Technik bzgl. MT-Hard- und -Software, um dann Ideen und Ergebnisse für die Konstruktion eines verbesserten MT-Tisches auf DI-Basis vorzustellen. Außerdem werden zwei Erweiterungen auf Basis der DI-Technik vorgestellt: Hover-Detektion und Handunterscheidung.
- Kapitel 4 erörtert den Stand der Technik im Bereich der Interfacegestaltung mit dem Fokus auf Interaktionstechniken und Gestaltungsmöglichkeiten für interaktive Oberflächen und verwandte Eingabegeräte. Dies geschieht anhand von Designdimensionen, welche durch ihre Extremwerte aufgespannt werden. Innerhalb jeder Designdimension stehen dem Interfacedesigner verschiedene Abwägungen offen, deren Vor-/Nachteile in diesem Abschnitt analysiert werden.
- Kapitel 5 befasst sich mit neuen Interaktionstechniken für elementare Operationen der 3D-Modellierung: Selektion von Objekten und Manipulation in sechs Freiheitsgraden.
- Kapitel 6 beschäftigt sich mit Ansätzen für zwei verschiedene Modellierungstechniken (Metaballs und Virtual Sculpting), um diese auf interaktiven Oberflächen einzusetzen. Im Fokus stehen dabei insbesondere beidhändige Interaktionstechniken.

3D-Modellierung mit interaktiven Oberflächen

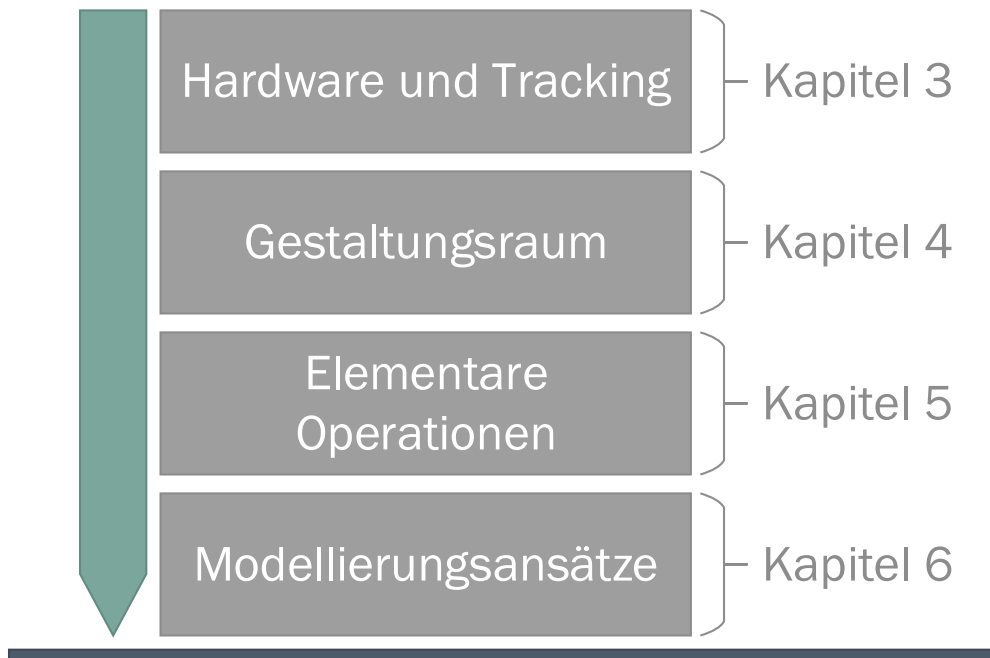


Abbildung 1.2: Schematische Übersicht zur Vorgehensweise dieser Arbeit – Ansatz ist die Behandlung verschiedener wichtiger Themen entlang der gesamten Verarbeitungskette von der Hardware über elementare Operationen bis zu komplexen Modellierungsansätzen.

- Kapitel 7 zieht schließlich ein Fazit, fasst die Ergebnisse der Arbeit zusammen und diskutiert offene Fragen und mögliche zukünftige Entwicklungen.

1.4 Veröffentlichungen

Einige der Ideen und Ergebnisse, die im Rahmen dieser Arbeit entstanden sind, wurden (meist in gekürzter Form) als Papiere oder im Rahmen von Vorträgen auf wissenschaftlichen Konferenzen, in wissenschaftlichen Magazinen oder als begutachtete Buchkapitel veröffentlicht. Außerdem sind einige Ideen und Ansätze in Zusammenarbeit mit Kollegen aus der Arbeitsgruppe (AG) Digitale

Medien oder anderen Arbeitsgruppen in diesem Bereich und/oder mit durch den Autor betreuten Studenten entstanden bzw. evaluiert worden:

- Die Ideen und Verbesserungen der Beleuchtungsanlage für einen DI-Tisch und die Diskussion zur Konstruktion des mrT wurden veröffentlicht als Teichert u. a. (2010) in Kooperation mit Jens Teichert und weiteren Kollegen.
- Die Erweiterungen zur Hover-Detektion sind im Rahmen der durch den Autor betreuten Masterarbeit von Jannis Stoppe (Stoppe, 2009) entstanden.
- Die Erweiterungen zur Handunterscheidung sind im Rahmen der durch den Autor betreuten Masterarbeit von Markus Aust entstanden und als Walther-Franks u. a. (2011a) in Kooperation mit Benjamin Walther-Franks und weiteren Kollegen veröffentlicht worden.
- Einige Gedanken und Ideen zum Gestaltungsraum für interaktive Oberflächen wurden veröffentlicht als Herrlich u. a. (2012b) in Kooperation mit Benjamin Walther-Franks und Rainer Malaka.
- Die proxybasierten Selektionstechniken sind im Rahmen der durch den Autor betreuten Diplomarbeit von Jan Holthusen entstanden und als Herrlich u. a. (2011b) in Kooperation mit Benjamin Walther-Franks, Roland Schröder-Kroll und weiteren Kollegen veröffentlicht worden.
- Die Interaktionstechniken für die integrierte Manipulation in sechs Freiheitsgraden wurden veröffentlicht als Herrlich u. a. (2011a) in Kooperation mit Benjamin Walther-Franks.
- Einige Teile und Konzepte der Modellierungstechniken für implizite Flächen wurden veröffentlicht als Herrlich u. a. (2008) und Walther-Franks u. a. (2011b) in Kooperation mit Benjamin Walther-Franks und Rainer Malaka.
- Das Interaktionsdesign und die Untersuchung zu Virtual Sculpting mit MT entstanden im Rahmen der durch den Autor betreuten Diplomarbeit von Andreas Braun und wurden veröffentlicht als Herrlich u. a. (2012a) in Kooperation mit Andreas Braun und Rainer Malaka.

- Darüber hinaus wurden einzelne Gedanken und Ansätze (teilweise mit anderem Fokus) in den folgenden Publikationen veröffentlicht und/oder im Rahmen von Vorträgen präsentiert: Krause u. a. (2008), Walther-Franks u. a. (2008), Walther-Franks und Herrlich (2010), Walther-Franks u. a. (2010), Aras u. a. (2011), Herrlich u. a. (2011c).

Kapitel 2

Stand der Technik

Das Feld der 3D-Modellierung wurde in den letzten 30 Jahren intensiv erforscht. Eine vollständige Aufzählung und Diskussion aller Ansätze würde daher den Rahmen dieser Arbeit sprengen. In diesem Kapitel wird die Darstellung somit auf eine kurze Zusammenfassung der wichtigsten Grundlagen beschränkt, mit dem Ziel eine Verortung dieser Arbeit in dem großen Feld der 3D-Modellierung vorzunehmen. Die Diskussion der spezifischen relevanten Arbeiten zu den im weiteren Verlauf dieser Arbeit näher untersuchten Ansätzen findet sich an geeigneter Stelle in den jeweiligen Kapiteln. Außerdem erfolgt eine ausführliche Diskussion relevanter Arbeiten unter Interaktionsgesichtspunkten in Kapitel 4. Daher wird die Darstellung in diesem Kapitel im Wesentlichen auf die Grundlagen und alternativen Ansätze und Techniken aus dem Bereich der Computergrafik beschränkt, welche im weiteren Verlauf der Arbeit nicht im Detail untersucht werden, aber für das Verständnis und der Vollständigkeit halber zumindest kurz angesprochen werden müssen. Für eine detailliertere Darstellung der computergrafischen Grundlagen und Modellierungstechniken wird der Leser auf existierende Lehrbücher (Foley u. a., 1995; Erleben u. a., 2005; Akenine-Moller u. a., 2008) und Review-Papiere (Gibson und Mirtich, 1997; Jones u. a., 2006) in diesem Bereich verwiesen.

2.1 Mathematische Modelle für 3D-Objekte

3D-Objekte können auf verschiedene Arten durch ein mathematisches Modell repräsentiert werden mit spezifischen Vor- und Nachteilen in Bezug auf die Mo-

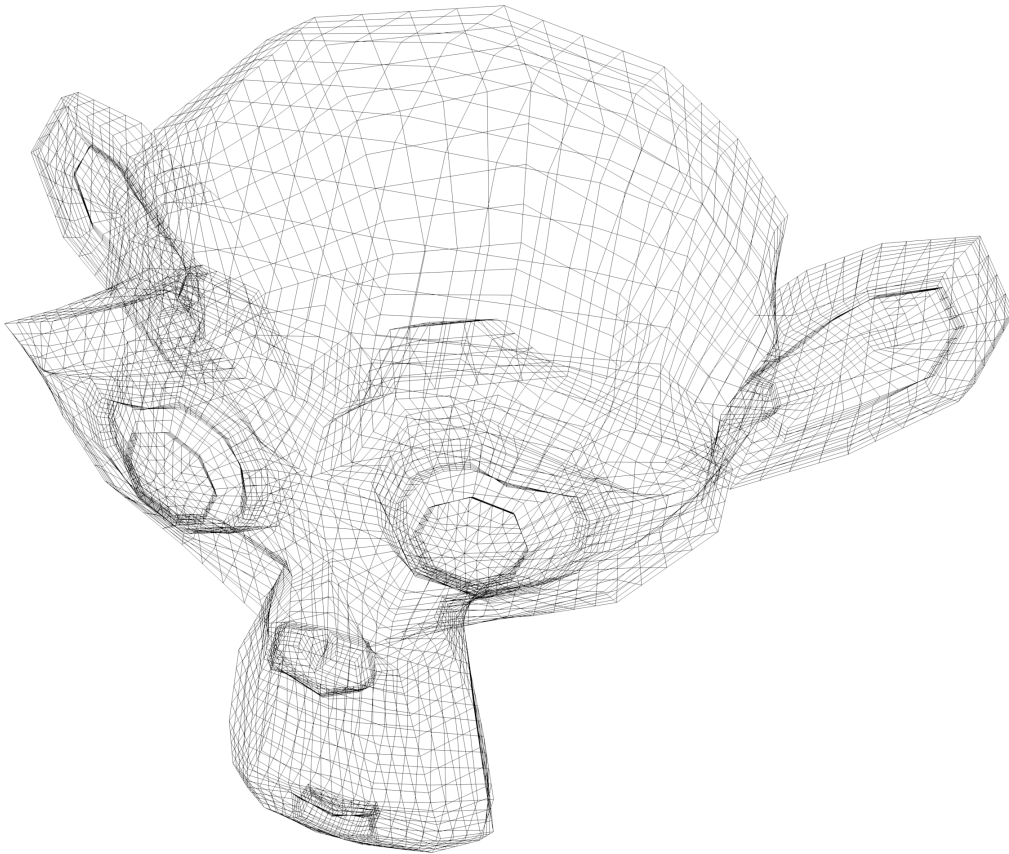


Abbildung 2.1: Beispiel für die Modellierung von Objekten mit polygonalen Netzen – Das dargestellte Modell ist ein Standard-Objekt aus dem 3D-Modellierungstool Blender und verwendet sowohl Dreiecke (z.B. für die Augen) als auch Vierecke zur Approximation der Objektoberfläche.

dellierungstechniken, Darstellungs- und Interaktionsmöglichkeiten. Für einige Anwendungen, z.B. ist dies in Computerspielen häufig der Fall, muss lediglich die Oberfläche des Objektes modelliert bzw. repräsentiert werden (Boundary Representation) für andere ist dagegen auch das Objektinnere wichtig und muss dementsprechend im Modell vorhanden sein (Volume Representation).

Polygonale Gitternetze (Meshes) (vgl. Abbildung 2.1) sind eine Möglichkeit, um die Oberfläche eines Objektes zu repräsentieren (Foley u. a., 1995). Die gekrümmte Oberfläche wird durch ein Netz aus Polygonen approximiert. Der einfachste Fall ist ein Netz aus Dreiecken. In der Modellierung sind außerdem Netze aus Vier-, Fünf-, oder Sechsecken gebräuchlich, da diese in bestimmten Fällen topologische Vorteile bieten. Auch eine Kombination aus verschie-

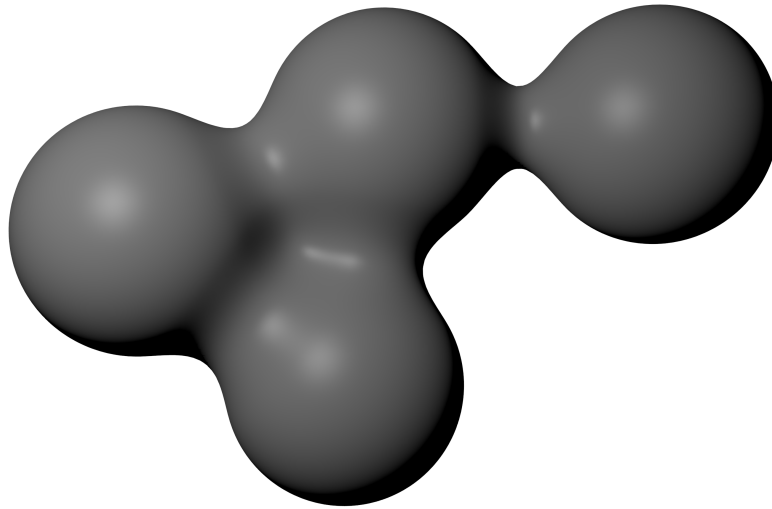


Abbildung 2.2: Metaballs oder Blobbies sind eine Anwendung von impliziten Funktionen bzw. Oberflächen in der 3D-Modellierung – Der Modellierung liegt kein physikalisches Modell zu Grunde, dennoch weist die Modellierung und die Darstellung Ähnlichkeiten zu der Arbeit mit Knete oder Ton in der realen Welt auf. Es ist zu sehen, wie die einzelnen Metaballs in Abhängigkeit zur Entfernung *verschmelzen* und auf diese Weise ein komplexeres Modell bilden.

denen Polygontypen ist möglich. Auf Grund der einfachen Struktur und der Hardwareunterstützung (Akenine-Moller u. a., 2008) beim Rendering haben polygonale Netze über die letzten Jahre eine weite Verbreitung erfahren und sind derzeit die Standarddarstellung in den meisten interaktiven grafischen 3D-Anwendungen.

Bezierkurven und -flächen sowie verschiedene Varianten von Splinekurven und -flächen werden ebenfalls zur Oberflächenmodellierung eingesetzt (Foley u. a., 1995; Gibson und Mirtich, 1997). Diese erlauben eine beliebig präzise Approximation von gekrümmten Flächen und garantieren (je nach Splinevariante) dabei bestimmte Glättkriterien, z.B. C1- oder C2-Kontinuität. Algorithmisch sind die Verfahren leicht zu beherrschen, erfordern aber ein gewisses mathematisches Verständnis beim Benutzer. Vor der Darstellung muss in gängigen Renderingpipelines eine Umwandlung in ein polygonales Netz erfolgen.

Implizite Flächen liegen als Lösung bzw. Isoflächen einer impliziten Gleichung der Form $f(x, y, z) = 0$ vor (Foley u. a., 1995). Diese Darstellung wird

auch für die in Abschnitt 6.1 beschriebenen Metaball-Techniken (vgl. Abbildung 2.2) eingesetzt. Implizite Flächen wurden erfolgreich für einige intuitive Modellierungsmethoden, speziell für organische Modelle eingesetzt und sind sehr einfach in der Handhabung (Bloomenthal und Wyvill, 1990). Allerdings erfordert die Extraktion (Bloomenthal, 1988, 1994) der resultierenden Oberfläche Rechenaufwand und feine Details sind mit dieser Darstellung nur schwer zu modellieren.

Eine Möglichkeit zur Modellierung eines Objektes als Volumen stellt die Voxel-Repräsentation (Foley u. a., 1995) dar. Dabei kann als Voxel entweder ein Polyeder (meist ein Quader) oder ein Punkt in einem Gitter verstanden werden. Jedem Voxel können Eigenschaften zugeordnet werden, im einfachsten Fall einfach die binäre Information voll oder leer. Aus der Menge der vollen Voxel ergibt sich dann das Objekt. Die Oberfläche des Objektes muss dabei für die Darstellung als Isofläche mittels eines geeigneten Verfahrens aus der Menge der Voxel extrahiert werden. Das Voxelmodell kann somit auch als gleichmäßig im Raum angeordnetes Skalar- oder Vektorfeld betrachtet werden. Einige Ansätze für die Repräsentation von virtuellem Ton (Virtual Clay) verwenden ebenfalls Skalar- oder Vektorfelder, welche nicht notwendigerweise gleichmäßig im Raum angeordnet sind (Dewaele und Cani, 2004). Ein weiterer alternativer Ansatz ist der Einsatz von Partikeln, um Oberflächen und Objekte zu repräsentieren (Markosian u. a., 1999; Rungjiratananon u. a., 2008).

Bei Constructive Solid Geometry (CSG) (Foley u. a., 1995) werden elementare Flächen (2D) oder Körper (3D) über Mengenoperationen wie Vereinigung und Schnitt kombiniert, um ein Objekt zu modellieren. Da die elementaren Körper in der Regel so gewählt werden, dass diese analytisch exakt beschrieben werden können, ist damit auch das resultierende Gesamtobjekt analytisch exakt zu beschreiben. Diese hat vor allem Vorteile für den Einsatz in automatischen Fräsen oder ähnlichen Systemen.

Eine Konvertierung zwischen den beschriebenen Modellen ist grundsätzlich möglich, allerdings nicht immer ohne Informationsverlust. Für einige Operationen kann außerdem auch die Darstellung als Distanzfeld (Jones u. a., 2006) Vorteile bieten.

2.2 Werkzeuge für die 3D-Modellierung

Basierend auf den verschiedenen Modellen bzw. Datenstrukturen wurden in den letzten Jahren unterschiedliche Ansätze verfolgt, um die Modellierung für den Benutzer zugänglicher, präziser und effektiver zu gestalten. Es ist zu beachten, dass häufig auch unterschiedliche Repräsentationen innerhalb derselben Anwendung und für dieselben Objekte eingesetzt werden können, z.B. eine bestimmte Repräsentation, die es dem Benutzer erlaubt, Änderungen durchzuführen und eine andere Repräsentation für das Rendering (Foley u. a., 1995). In aktuellen Rastergrafikpipelines (Akenine-Moller u. a., 2008) wird zur Darstellung letztlich immer ein polygonales Mesh verwendet. Insgesamt ist die Wechselwirkung und die Zahl der Kombinationsmöglichkeiten unterschiedlicher Modelle und Interaktionstechniken kaum überschaubar, daher werden hier exemplarisch nur einige der wichtigsten Vertreter diskutiert.

Zunächst besteht immer die Möglichkeit, die verwendete Datenstruktur direkt zu manipulieren, d.h. entsprechende Werkzeuge für den Benutzer anzubieten. Je nach verwendeter Repräsentation erlaubt dies bereits eine mehr oder weniger intuitive Modellierung. Parametrische Kurven und Flächen beispielsweise kann der Benutzer über Kontrollpunkte und Handles für die Tangenten manipulieren (Foley u. a., 1995). Dies ist für einfache Anwendungen recht intuitiv, erfordert aber vom Benutzer eine eingehende Auseinandersetzung mit dem zu Grunde liegenden mathematischen Modell.

Auch für polygonale Netze werden häufig Werkzeuge eingesetzt, die vom Benutzer eine gute Kenntnis der Eigenschaften und Funktionsweise des mathematischen Modells verlangen. Typisch sind beispielsweise Werkzeuge zum Hinzufügen, Löschen und Verschieben einzelner Eckpunkte, Kanten und Flächen. Erweiterte Direct Manipulation- oder Free Form Deformation (FFD)-Ansätze erlauben es dem Benutzer, z.B. über den Einsatz von Kontrollpunkten, sich von der Manipulation einzelner Ecken und Kanten zu lösen und sich mehr auf die gewünschte Zielform zu konzentrieren. Bei der FFD ist das Objekt in einen umgebenden Raum eingebettet (Barr, 1984; Sederberg und Parry, 1986; Hsu u. a., 1992; MacCracken und Joy, 1996), welcher auf verschiedene Weisen manipuliert werden kann, entweder über definierte mathematische Operationen wie Rotationen oder auch über Kontrollpunkte. Das Objekt wird somit indirekt

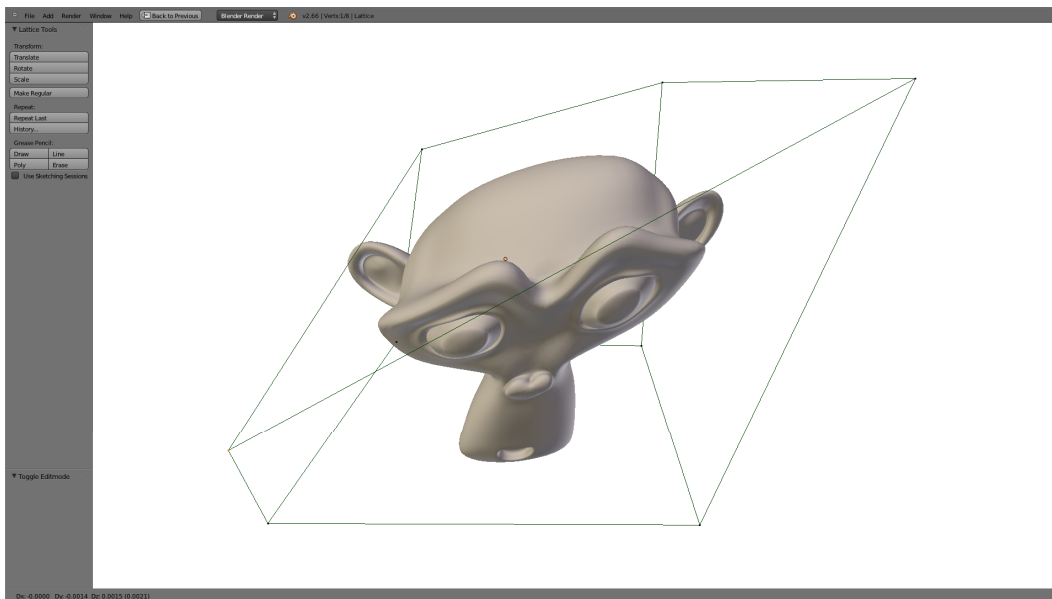


Abbildung 2.3: Beispiel für Free Form Deformation – Das komplexe polygonale Netz des Zielobjektes (vgl. Abbildung 2.1) kann über die acht Kontrollpunkte des umgebenden Quaders auf einfache Weise manipuliert werden.

manipuliert, es ist für den Benutzer jedoch einfacher, globale Deformationen korrekt durchzuführen. Um ein Objekt z.B. zu verdrehen oder anderweitig zu verformen, ist es im Allgemeinen notwendig, eine Vielzahl von Eckpunkten und Kanten zu manipulieren und auf einander abzustimmen. FFD-Techniken erlauben es dem Benutzer, hier zu abstrahieren und beispielsweise nur wenige Kontrollpunkte zu manipulieren (vgl. Abbildung 2.3), um ggf. tausende von Punkten des Objektes wie gewünscht zu verändern, ohne jeden dieser Punkte individuell zu kontrollieren.

Noch einen Schritt weiter geht die Manipulation von polygonalen Netzen über virtuelles Sculpting (Williams, 1990; Galyean und Hughes, 1991; Bill und Lodha, 1994). Hier muss der Benutzer gar nicht mehr an Punkte oder Kanten denken, auch wenn diese dem Modell zu Grunde liegen. Stattdessen modelliert der Benutzer durch den Einsatz virtueller Werkzeuge, welche stark an physische Werkzeuge eines Bildhauers angelehnt sind (vgl. Abbildung 2.4 und Abschnitt 6.2).

In Kombination mit geeigneten Echtzeit-Renderingverfahren und direkter Manipulation können implizite Flächen sich ähnlich verhalten wie virtueller Ton oder Knetmasse. Einfache Grundkörper (Metaballs, Blobbies oder Soft

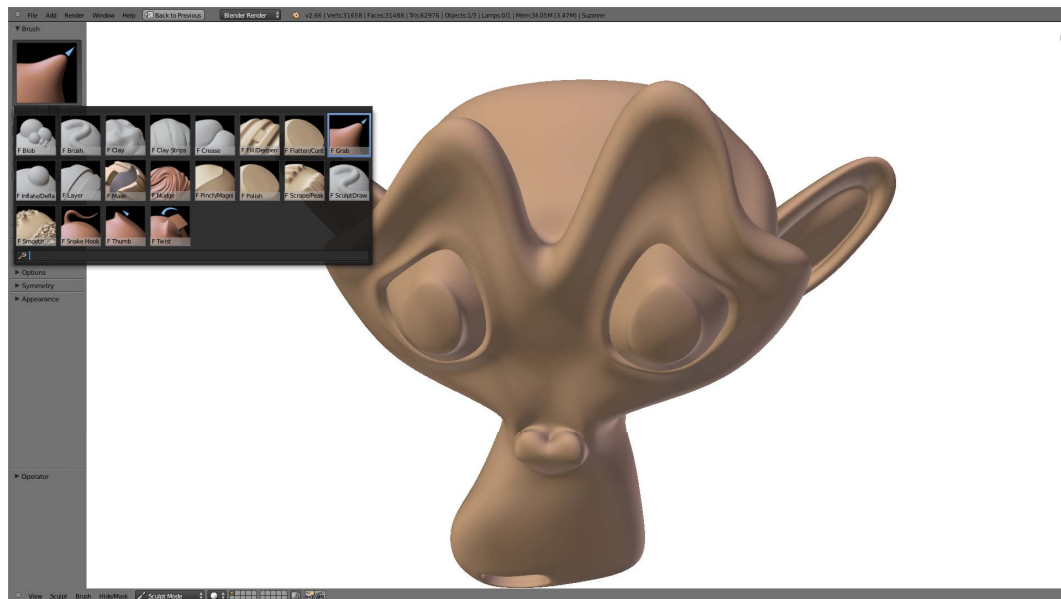


Abbildung 2.4: Virtuelles oder digitales Sculpting in Blender – Es stehen verschiedene virtuelle Werkzeuge z.B. zum Eindrücken oder Herausziehen der Oberfläche zur Verfügung (links oben im Bild).

Objects genannt) verschmelzen dabei scheinbar zu komplexeren organischen Formen (Blinn, 1982; Wyvill u. a., 1986b,a; Bloomenthal und Wyvill, 1990) (vgl. Abbildung 2.2).

Die Grundidee, virtuelle Werkzeuge und den Prozess der Modellierung virtueller 3D Objekte nach dem Vorbild physischer Werkzeuge zu gestalten und auf diese Weise das Wissen und die Erfahrung der Benutzer aus der physischen Welt zu nutzen, macht den Bereich der sog. physikbasierten Ansätze (Physics-based Modeling) aus (Gibson und Mirtich, 1997; Erleben u. a., 2005). Oftmals wird dies über die Simulation einfacher mechanischer Zusammenhänge erreicht. Es existieren wiederum verschiedene Ansätze, um die zunächst rein geometrischen Modelle, z.B. polygonale Netze, mit physikalischen Eigenschaften zu versehen, um eine physikalisch plausible Deformation der Objekte zu ermöglichen (sog. Soft Bodies). Die Bandbreite der Ansätze reicht hier von sehr einfachen Masse-Feder-Modellen, die zwar physikalisch plausibel, aber nicht unbedingt korrekt sind, bis hin zu aufwändigen numerischen Verfahren wie der Methode der finiten Elemente (FEM) (Terzopoulos u. a., 1987; Erleben u. a., 2005).

Sketch-based Modeling-Ansätze (Sutherland, 1964; Zeleznik u. a., 1996; Igarashi u. a., 1999) erlauben es dem Benutzer, durch das Zeichnen mit einem Stylus Modellierungsoperationen durchzuführen. Hier gibt es eine Reihe verschiedener Varianten und Techniken. Eine Möglichkeit ist der Einsatz symbolischer Gesten, um bestimmte Operationen auszuwählen. Ein andere Ansatz ist den Benutzer die gewünschte Kontur zeichnen zu lassen und das Objekt dann so gut wie möglich an diese anzupassen (Nealen, 2007; Nealen u. a., 2009).

Auch Kombinationen aus den verschiedenen oben genannten Ansätzen sind möglich. Des Weiteren können unterschiedliche Interfaces gewählt werden. Aktuelle Modellierungssoftware folgt größtenteils dem etablierten WIMP-Paradigma und stellt grafische Tools in Form von Widgets und Gizmos zur Verfügung, die mit der Maus oder einem Stift eingesetzt werden können. In allen dem Autor bekannten Programmen besteht aber auch die Möglichkeit, Transformationen direkt durch die numerische Eingabe von Winkeln oder Positionen zu bestimmen.

2.3 Einordnung dieser Arbeit

Im Fokus dieser Arbeit liegt die Kombination existierender Modellierungstechniken mit den erweiterten Möglichkeiten von MT-Interaktion. Dabei ist nicht von vornherein klar, welche Repräsentation oder Modellierungstechnik hier eine optimale Kombination darstellt. Stattdessen werden verschiedene Ansätze betrachtet, die auf Grund ihrer Anlehnung an physische Modellierungstechniken als aussichtsreiche Kandidaten für die MT-Manipulation erscheinen: die Modellierung mit impliziten Flächen bzw. Metaballs und das virtuelle Sculpting von Objekten. Die Entwicklung neuer Repräsentationen ist dagegen nicht Teil dieser Arbeit. Die Anwendung von MT-Interaktion für die 3D-Modellierung erfordert eine ganzheitliche Betrachtung der Stärken und Schwächen von MT-Interaktion. Dies schließt auch Möglichkeiten und Grenzen der eingesetzten Hardware ein, welche im folgenden Abschnitt diskutiert werden.

Kapitel 3

Multi-touch-Hardware und -Software

Dieses Kapitel dient der Darstellung und Einordnung aktueller und etablierter Ansätze im Bereich Multi-touch (MT)-Hard- und -Software. Das existierende MT-Hard- und -Softwareumfeld ist äußerst heterogen, da verschiedene technische Ansätze und Umsetzungen nebeneinander existieren, welche durch unterschiedliche Interaktionseigenschaften gekennzeichnet sind. Eine Kenntnis und Einordnung verschiedener Ansätze ist daher für die Diskussion der in dieser Arbeit vorgestellten Beiträge äußerst wichtig. Es werden verschiedene MT-Detektionsansätze vorgestellt und diskutiert und die gängigen Softwareframeworks kurz beleuchtet. Darüber hinaus werden die eigenen Beiträge in diesem Gebiet, d.h. die Verbesserung existierender Diffuse Illumination (DI)-Ansätze und Detektionserweiterungen zu Hovering und Handunterscheidung auf Basis dieses Ansatzes vorgestellt.

3.1 Einleitung

Hauptziel und Eigenschaft von MT-Eingabegeräten ist die zuverlässige Erkennung von Fingerberührungen zur Interaktion mit Softwareanwendungen. Bei MT-Eingabegeräten ist dabei die Erkennung von mindestens zwei oder mehr Fingereingaben gemeint. Bei kamerabasierten Eingabegeräten ist die Anzahl der erkennbaren bzw. verarbeitbaren Berührungen nur durch die Leistungsfähigkeit des Gesamtsystems in Bezug auf CPU und Speicher begrenzt, aber

keiner prinzipiellen Beschränkung unterworfen, wie dies beispielsweise beim kapazitiven oder resistiven Tracking der Fall ist.

MT-Eingabegeräte unterscheiden sich auf vielfältige Art und Weise, z.B. in der genauen Zahl der Berührungen, welche ein Gerät zu unterscheiden in der Lage ist, aber auch in Genauigkeit und Größe des Eingabebereiches oder des möglichen Formfaktors. Außerdem ist es entscheidend, ob das Gerät lediglich zur Eingabe dient oder Eingabe und Ausgabe in einem Gerät vereint (MT-Bildschirm). Einige Geräte unterstützen darüber hinaus nicht nur Fingereingaben, sondern auch die Verwendung von speziellen Markern oder physischen Objekten (Tangibles).

Neben der Diskussion existierender Techniken als wichtige Grundlage des Interaktionsdesigns und der umgesetzten Techniken und Experimente, sind im Rahmen dieser Arbeit auch Beiträge zur Verbesserung und Erweiterung der Eingabegeräte selbst entstanden, welche in diesem Kapitel ebenfalls vorgestellt werden.

3.2 Stand der Technik

Erste Ansätze zur Forschung im Bereich MT reichen zurück bis in die frühen 1980er Jahre (Buxton, 2007). Allerdings haben erst neuere Entwicklungen das Interesse an der Forschung in diesem Bereich verstärkt. Insbesondere die Beiträge von Jeff Han auf der TED-Konferenz sowie der kommerzielle Erfolg des Apple iPhone haben das Interesse an MT-Technologie in Forschung und Industrie stark gesteigert.

Dieser Abschnitt gliedert sich in zwei Teile. Zunächst werden verschiedene Hardwareansätze und grundsätzliche Systemansätze zur Detektion von multiplen Fingereingaben vorgestellt und kurz Vor- und Nachteile sowie grundsätzliche Implikationen für das Software- und Interaktionsdesign diskutiert. Im zweiten Teil wird die Softwaresituation beleuchtet, soweit diese für das weitere Verständnis der Arbeit notwendig erscheint.

Insbesondere bei kamerabasierten Systemen stehen verschiedene Trackinglösungen zur Verfügung. Auch ist eine geeignete einheitliche Schnittstelle zur Verarbeitung der Fingereingaben notwendig. Hier hat sich im Forschungsbereich das TUIO-Protokoll etabliert, welches kurz beschrieben wird. Daneben

existieren einige kommerzielle Frameworks. MT-Eingaben (für Geräte mit entsprechenden Treibern) werden seit Windows 7 auch direkt vom Betriebssystem bzw. über eine Erweiterung der Windows-API unterstützt.

3.2.1 Hardware

Im Bereich der MT-Hardware existieren verschiedene Ansätze mit spezifischen Vor- und Nachteilen für die Erkennung und Verarbeitung von Mehrfingereingaben. Eine Vielzahl von Forschungsprototypen arbeitet auf Basis von optischem Tracking mit Hilfe einer oder mehrerer Kameras (Krueger u. a., 1985; Wilson, 2004, 2005; Han, 2005; Jordà u. a., 2005; Hodges u. a., 2007; Izadi u. a., 2007; Teichert u. a., 2010). Das Tracking findet meist nicht im sichtbaren Lichtspektrum, sondern im Nahinfrarotbereich statt, um Wechselwirkungen mit dem normalen Umgebungslicht oder der Ausgabe über einen Projektor zu reduzieren. Dazu kommen entsprechende optische Filterkombinationen zum Einsatz. Andere Ansätze basieren auf der Messung bestimmter elektromagnetischer Eigenschaften, z.B. Kapazitäten oder Widerstände, die direkt oder indirekt durch die Interaktion mit der Hand bzw. den Fingern verändert und daher gemessen werden können (Westerman, 1999; Patten u. a., 2001; Dietz und Leigh, 2001; Rekimoto, 2002; Rosenberg und Perlin, 2009). Die genaue Lokalisierung und Unterscheidung der Finger ist insbesondere für letztere Systeme eine Herausforderung, während die eigentliche Berührungsdetektion zuverlässig funktioniert. Dafür ist die Anzahl der gleichzeitigen Berührungen, die unterschieden werden können, je nach Technik beschränkt. Eine Detektion von Gegenständen oder Markern ist nur begrenzt oder nicht möglich. Der optische Ansatz ist dagegen sehr flexibel, allerdings auf Grund des *Computer Vision*-Ansatzes mit größeren Ungenauigkeiten bei der Unterscheidung zwischen Berührung und Nicht-Berührung behaftet, welche nur durch ein sehr kontrolliertes Beleuchtungsumfeld ausgeglichen werden können.

Auf Grund der Vielzahl der kommerziellen und individuellen Lösungen im Bereich der MT-Hardware können nicht alle existierenden Geräte im Detail besprochen werden. Daher erfolgt die Diskussion der spezifischen Vor- und Nachteile bestimmter Trackingtechnologien exemplarisch an ausgewählten Vertretern des jeweiligen Ansatzes. Für weitere Details wird verwiesen auf den Technical Report von Schöning u. a. (2008). Eine historische Übersicht findet

sich außerdem auf der Webseite von Bill Buxton (Buxton, 2007).

Zu den Systemen, welche Objekte und/oder Fingerberührungen auf Basis elektromagnetischer Eigenschaften erkennen, zählen SenseTable (Patten u. a., 2001), DiamondTouch (Dietz und Leigh, 2001), SmartSkin (Rekimoto, 2002) und das UnMousePad (Rosenberg und Perlin, 2009).

Der DiamondTouch von Dietz und Leigh (2001) funktioniert mittels kapazitiver Kopplung. Die interaktive Oberfläche besteht im Prinzip aus einer Matrix von Antennen, die über ein spezifisches Signal individuell und eindeutig kodiert werden können. Allerdings ist diese Technik bisher nicht mit transparenten Oberflächen kompatibel, weshalb eine Rückprojektion nicht möglich ist und es daher zu Verschattungen bei der Interaktion kommt. Besonderes Kennzeichen und de-facto immer noch ein Alleinstellungsmerkmal des DiamondTouch ist die Möglichkeit der spezifischen Zuordnung von Berührungspunkten zu Benutzern bzw. Stühlen, was dieses System insbesondere für den Einsatz in kooperativen Szenarien äußerst interessant macht. Allerdings funktioniert das System nur, solange eine spezielle Bodenplatte oder der Stuhl vom Benutzer berührt wird und ist daher schon aus Platzgründen auf relativ wenige Nutzer gleichzeitig beschränkt. Auch wird nicht der tatsächliche Nutzer erkannt, sondern lediglich der Stuhl. Daher eignen sich die Identifizierungsdaten nur für die Interaktion, also z.B. für die automatische Orientierung von GUI-Elementen, aber nicht für Authentifizierungszwecke.

Rekimotos SmartSkin (Rekimoto, 2002) basiert ebenfalls auf kapazitivem Tracking. Über eine Matrix von Antennen kann die Änderung des elektrischen Feldes gemessen werden, die durch die Annäherung der Hand bzw. des Fingers induziert wird. Die Fingerdetektion erfolgt über eine Suche von Extremwerten in dem resultierenden Potentialbild. Dadurch können auch mehrere Finger ohne Probleme gleichzeitig erkannt werden. Außerdem kann innerhalb gewisser Grenzen auch die relative Entfernung der Hand zur Oberfläche gemessen werden, was sowohl zur Touch-, als auch zur Hover-Detektion (vgl. Abschnitt 3.4) genutzt werden kann. Ähnlich wie bei DiamondTouch ist aber auch bei SmartSkin auf Grund der nicht transparenten Oberfläche lediglich eine Aufprojektion mit den entsprechenden Nachteilen bzgl. Verschattung möglich. Die Detektion von Objekten ist über den Einsatz spezieller kapazitiver Tags möglich, deren Herstellung und Einsatz allerdings wesentlich aufwändiger ist als der einfacher

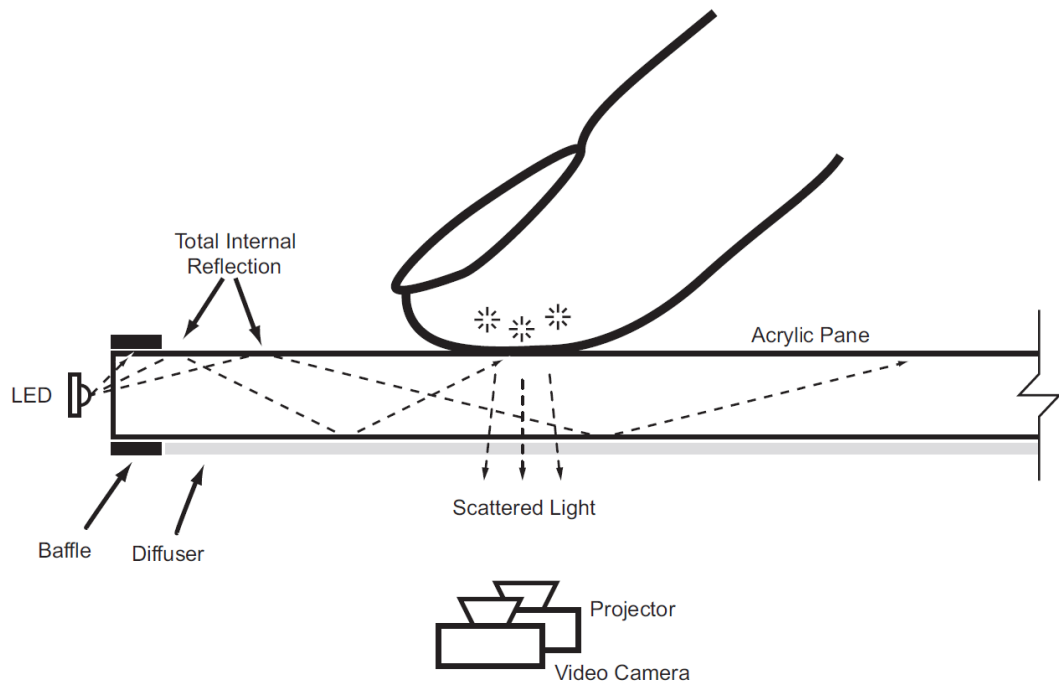


Abbildung 3.1: Prinzip der Frustrated Total Internal Reflection für die Nutzung in Multi-touch-Bildschirmen (Abbildung aus Han (2005)) – Durch den Oberflächenkontakt mit dem Finger kann das Infrarotlicht austreten und als heller Bereich durch eine entsprechende Kamera auf der Unterseite detektiert werden.

optischer Marker. Auf Grund des einfachen und regelmäßigen Hardwareaufbaus in Kombination, sind mit dieser Technik große interaktive Oberflächen realisierbar.

Zu den frühesten Systemen auf Basis von optischem Tracking zählen Kruegers Arbeiten zu Videoplace aus den Jahren 1983 bis 1985. Krueger entwirft einige Anwendungsszenarien, die auch heute noch Bestand haben (Krueger u. a., 1985). Das System kombiniert konzeptionell horizontale und vertikale Flächen und eine Gestensteuerung. Unter anderem schlägt Krueger die Nutzung eines solchen Systems zur Interaktion mit grafischen Objekten und typische Operationen wie Translation und Skalierung vor.

Bis heute haben sich insbesondere im akademischen Bereich zwei Standardtechniken (neben einer Vielzahl von Varianten) für das optische Tracking für MT-Bildschirme etabliert: Frustrated Total Internal Reflection (FTIR) und Diffuse Illumination (DI).

Die Nutzung des FTIR-Effekts für die Konstruktion von MT-Oberflächen hat, angeregt durch die Veröffentlichung von Jefferson Han aus dem Jahre 2005 (Han, 2005) und dessen Auftritt auf der TED-Konferenz, große Verbreitung gefunden. Besonderer Vorteil dieser Technik ist die kostengünstige und einfache Konstruktion, die im Prinzip wenig mehr als eine Infrarot-Lichtquelle, eine Kamera, einen Projektor und eine Acrylglasplatte benötigt und somit in Interaktionslaboren und bei interessierten Hobbyisten fast aus *Hausmitteln* zusammengesetzt werden kann. Wenn auch einfach im Grundaufbau, so ist es im konkreten Einsatz dennoch eine Herausforderung eine robuste und ausgewogene Erkennung, eine haptisch angenehme Oberfläche und eine gute Abstimmung mit der Projektion zu erreichen und Erfahrungswerte hierzu werden in diversen Internetforen und Blogs gehandelt. Zusätzlich existieren inzwischen eine Reihe von Abwandlungen und verschiedene hybride Ansätze, insbesondere zwischen FTIR und DI.

Der FTIR-Effekt war auch vor der Veröffentlichung von Han gut bekannt. Han hat diesen Aufbau weiter für die Konstruktion interaktiver Oberflächen wie z.B. Touch-Tische verfeinert und mit einer Rückprojektion als MT-Bildschirm oder MT-Tisch (Tabletop) nutzbar gemacht. Das Grundprinzip nutzt den Effekt der Total Reflection. D.h. am Übergang zu einem Medium mit niedrigerem Brechungsindex hängt der Anteil des reflektierten Lichts vom Eintrittswinkel ab (vgl. (Han, 2005)) und ab einem bestimmten kritischen Winkel findet Totalreflexion statt (das Licht verbleibt nahezu vollständig im Medium). An optischen Übergängen zu anderen Medien allerdings, z.B. einem Finger auf einer Acrylglasplatte, wird das Licht *frustrated* und tritt dort nahezu im 90 Grad Winkel aus. Um diesen Effekt für die MT-Interaktion zu nutzen, wird eine Lichtquelle verwendet, die vom Rand in eine Acrylglasplatte einstrahlt und zwar in einem Winkel, bei dem nahezu Totalreflexion auftritt. Berührt nun ein Finger die Oberfläche, kann das an dieser Stelle austretende Licht von einer unter der Oberfläche befindlichen Kamera detektiert werden und die Finger erscheinen als helle Blobs, welche mit Computer Vision (CV)-Algorithmen weiterverarbeitet werden können (vgl. Abbildung 3.1).

Durch die Verwendung von Infrarotlicht und den Einsatz entsprechender wechselseitiger Filter für Kamera und Projektor kann im sichtbaren Bereich ein Bild auf eine geeignete Folie auf der Interaktionsoberfläche projiziert wer-

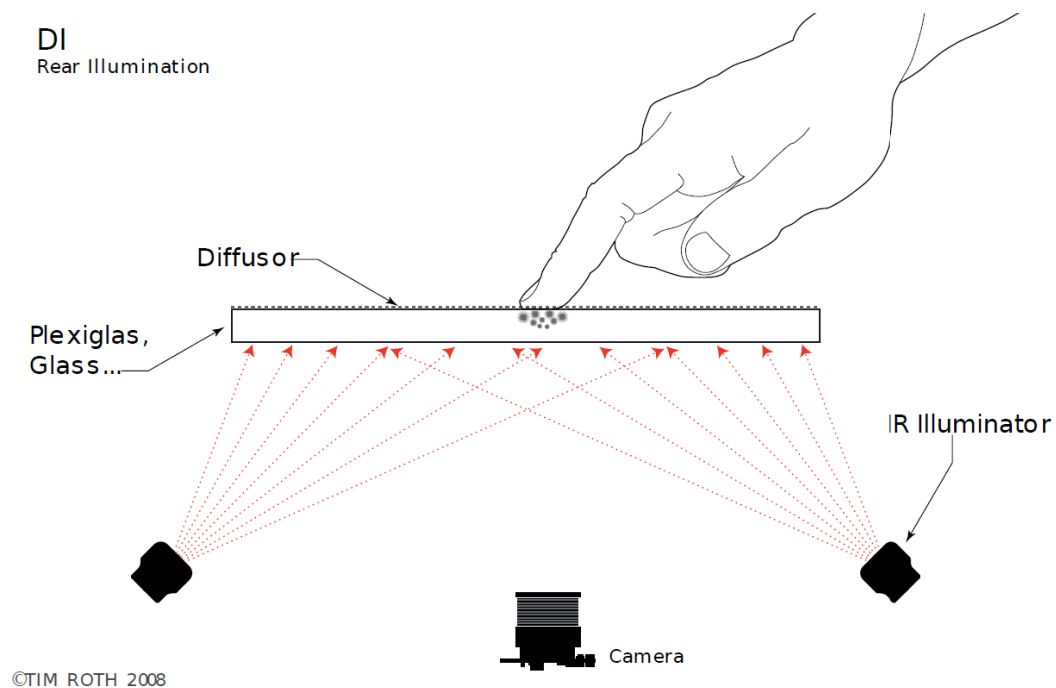


Abbildung 3.2: Prinzip der Diffuse Illumination (Abbildung aus Schöning u. a. (2008)) – Bedingt durch den Diffusor auf der Oberfläche, nimmt die Menge des reflektierten Lichtes mit Abstand zur Oberfläche schnell ab. Dadurch erscheinen Objekte direkt auf der Oberfläche sehr viel heller und schärfer und können als hellere Gebiete vor dem Hintergrund erkannt und isoliert werden.

den, ohne dass die Erkennung der Finger beeinflusst wird. Neben der Detektion von Fingern ist auch die Erkennung von anderen Objekten auf der Oberfläche möglich, hängt aber von deren optischen Eigenschaften ab. Im Gegensatz zur DI-Technik können keine Details, etwa Schrift, erkannt werden. Die Beschaffenheit der Haut (Feuchtigkeit usw.) hat Einfluss auf den Effekt. Um die Bewegung auf der trockenen Oberfläche für den Benutzer angenehmer zu gestalten, werden in der Regel noch zusätzliche Stoffe wie Silikon auf die Oberfläche aufgebracht. Geeignete Stoffe werden häufig experimentell per Trial-and-Error ermittelt und in Nutzerforen ausgetauscht. Han selbst empfiehlt Rosco Gray #02105 (Han, 2005) als Projektionsfolie auf der Acrylglasplatte einzusetzen. In gewissen Grenzen ist es möglich, eine Kraft aus Größe und Verteilung der resultierenden Blobs abzuleiten (Smith u. a., 2007).

Der ReacTable (Jordà u. a., 2005) wird beispielhaft als Vertreter für die DI-Technik genannt. Der grundlegende Aufbau und das Tracking sind ähnlich

wie bei FTIR. Eine Kamera unter der Tischoberfläche filmt im Infrarotbereich und gleichzeitig projiziert ein Beamer von unten ein Bild auf eine spezielle Projektionsfolie, welche gleichzeitig als Diffusor dient. Im Unterschied zu FTIR strahlt die Lichtquelle allerdings nicht seitlich in die Oberflächenplatte, sondern befindet sich unterhalb (oder oberhalb) der Oberfläche. Durch den Diffusor wird erreicht, dass das Licht beim Passieren der Oberfläche stark gestreut wird, was zur Folge hat, dass die Intensität mit Entfernung zur Oberfläche stark abnimmt. Dies bedeutet, Dinge in der Nähe bzw. auf der Oberfläche erscheinen als deutlich hellere Punkte als der übrige Hintergrund und können mit CV-Algorithmen extrahiert werden (vgl. Abbildung 3.2).

Kritische Punkte bei einem DI-Aufbau sind die Beleuchtungsquelle, die Ausrichtung von Kamera und Projektor sowie die Oberflächenbeschaffenheit und die Vermeidung von ungewollten Reflexionen. Details und eigene Beiträge hierzu werden in Abschnitt 3.3 besprochen. Ein Vorteil des DI-Ansatzes ist die Möglichkeit, einfache optische Marker auf der Oberfläche zu verfolgen sowie eine Robustheit gegenüber Bewegungen, da anders als bei FTIR auch ein kurzes ungewolltes Abheben der Finger nicht zu einem Aussetzen der Erkennung führen muss. In gewissen Grenzen ist ein Tracking über der Oberfläche möglich (siehe Abschnitt 3.4 und 3.5) und kann die Interaktionsmöglichkeiten erweitern. Gleichzeitig kann dieses Verhalten auch nachteilig im Vergleich zu der FTIR-Technik wirken. Eine exakte Bestimmung, ob tatsächlich ein Oberflächenkontakt vorliegt oder nur ein beinahe Kontakt, ist vergleichsweise schwieriger und in der Praxis ist bisher eine größere Ungenauigkeit hinzunehmen. Auch ist es nicht möglich, aus der Berührung realistische Informationen bzgl. Druck und Kraft abzuleiten.

Neben der eingesetzten Trackingtechnologie mit ihren spezifischen Vor- und Nachteilen, spielen Faktoren wie Größe, Ausrichtung, Form und Ergonomie einer interaktiven Oberfläche eine bedeutende Rolle für das Interaktionsdesign und mögliche Einsatz- und Anwendungsszenarien. Auf einige dieser Punkte wird in den entsprechenden Abschnitten Bezug genommen, soweit sie Einfluss auf das Interaktionsdesign oder andere untersuchte Aspekte haben. Die Untersuchung verschiedener Formfaktoren und ergonomische Erwägungen sind jedoch keine Kernthemen dieser Arbeit. Neben horizontalen und vertikalen MT-Bildschirmen gibt es eine Reihe von Entwicklungen zu hybriden Formen

oder völlig neuen Gestaltungskonzepten (z.B. Curved Displays). An dieser Stelle sollen der Vollständigkeit halber einige Projekte kurz erwähnt werden, ohne die Diskussion der Vor- und Nachteile hier zu vertiefen.

Grossman u. a. (2004) untersuchen die Interaktion mit einer sphärischen (volumetrischen) Anzeige. Das Tracking erfolgt über ein externes Motion-Capture-System und spezielle Marker an den Fingern. Dies erlaubt sowohl eine freie Gesteninteraktion als auch die Simulation von Touch-Eingaben. Die Besonderheiten dieses Systems sind die Interaktion mit einer gekrümmten Kugeloberfläche und die holografische 3D-Darstellung. Das System von Hahne u. a. (2009) ist ein Beispiel für die Kombination von MT mit Stifteingabegeräten.

3.2.2 Software

Zum besseren Verständnis werden zunächst die Rollen der Hardware und der Software im Rahmen eines MT-Gesamtsystems abgegrenzt und erklärt (vgl. Abbildung 3.3). Die Funktionsweise verschiedener Hardwareansätze wurde im vorherigen Abschnitt beschrieben. Aufgabe der Hardware ist es demnach, über einen Sensor Messungen vorzunehmen und diese zur Verarbeitung durch eine Softwareanwendung zur Verfügung zu stellen. Die Software nutzt dann wiederum andere Hardwarekomponenten (Grafik-/Soundkarte) zur Ausgabe. Die Software ist zu unterteilen in den Hardware- oder systemnahen Teil und die eigentliche Anwendungssoftware. Jede Hardware muss über entsprechende Software (Treiber) initialisiert und angesprochen werden. Eine Tracking-Software (Erkenner) erzeugt aus den Rohdaten der Hardware sinnvolle Eingabeereignisse für die Weiterverarbeitung in der Anwendung. Bei optischem Tracking bedeutet dies die Abarbeitung einer Bildverarbeitungs pipeline (vgl. Abbildung 3.4).

Für jede Fingerberührung werden als Daten mindestens eine eindeutige ID und die 2D-Position übertragen. Vielfach werden aber auch noch weitere Informationen, die für Anwendungen nützlich sein könnten, durch den Erkenner zur Verfügung gestellt, z.B. Zeitstempel, Beschleunigungen, Flächenangaben, Druck oder der zeitliche Verlauf jeder Fingerberührung. Einige dieser Daten können auch in der Anwendung nachträglich berechnet werden. Andere Daten, wie z.B. zu Druck, Fläche oder Form einer Berührung können nicht berechnet werden, sondern sind ggf. nicht vorhanden. Dies bedeutet, dass Anwendungen

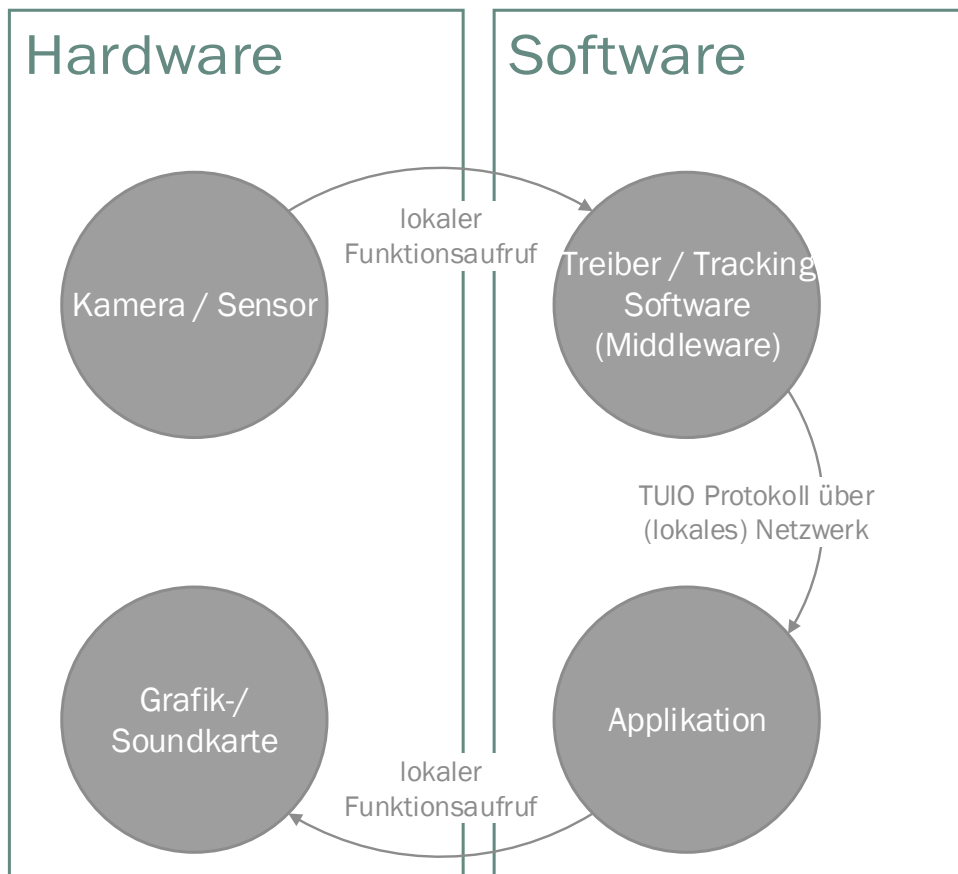


Abbildung 3.3: Ablauf der Kommunikation zwischen Hard- und Softwarekomponenten eines Multi-touch-Systems

im MT-Bereich sich entweder auf bestimmte Daten beschränken müssen oder nur auf wenigen Hardwareplattformen funktionieren.

Im Bereich der Software (besonders im Forschungsumfeld) konnte sich das TUIO-Protokoll (Kaltenbrunner u. a., 2005) als Schnittstelle zwischen Hardware bzw. Erkennen und Anwendungsprogrammen etablieren (vgl. Abbildung 3.3). TUIO stammt ursprünglich aus dem Umfeld des `reactIVision Frameworks` (Kaltenbrunner und Bencina, 2007).

Das TUIO-Protokoll wurde von Kaltenbrunner u. a. (2005) spezifiziert, um eine implementierungs- und plattformunabhängige Datenübertragung zwischen Tracker und Anwendungssoftware über ein Netzwerkprotokoll zu ermöglichen. Zusätzlich erlaubt die Übertragung per Netzwerkschnittstelle auch den einfachen Aufbau einer verteilten Infrastruktur, etwa die Verteilung von Tracking- und Anwendungsaufgaben auf verschiedene Rechner. Das TUIO-Protokoll wurde seit der ersten Veröffentlichung mehrfach aktualisiert. TUIO setzt seinerseits auf `Open Sound Control (OSC)` und `UDP` auf, um Echtzeitnachrichten zu übertragen.

Insbesondere für optische MT-Eingabegeräte, d.h. FTIR, DI bzw. hybride oder verwandte Ansätze, existieren eine Reihe von kommerziellen und frei verfügbaren Bibliotheken, welche die Bildverarbeitungspipeline implementieren. Einige Anbieter kommerzieller Systeme liefern ihre eigene proprietäre Bibliothek bzw. API mit. Da Details zu diesen Bibliotheken nicht in allen Fällen öffentlich gemacht werden, kann an dieser Stelle nicht näher auf diese eingegangen werden. Es ist aber davon auszugehen, dass diese im Wesentlichen den gleichen Prinzipien folgen, wie die frei verfügbaren Frameworks.

Eine Übersicht zu freien Bibliotheken auf verschiedenen Plattformen und eine Analyse zu Performanz und Features findet sich bei Schöning u. a. (2008), weshalb im Folgenden nur die für das weitere Verständnis der eigenen Beiträge wesentlichen Punkte der Verarbeitungspipeline kurz erläutert werden.

Die Verarbeitungspipeline für die DI-Technik sieht prinzipiell aus wie in Abbildung 3.4 dargestellt (vgl. (Schöning u. a., 2008)). Die Rohdaten der Kamera werden in einem Vorverarbeitungsschritt für den weiteren Prozess aufbereitet. Je nach Kamera kann es notwendig sein, das Bild zunächst in ein Grauwertbild zu konvertieren. Dann folgt die Entfernung des Hintergrundes, z.B. durch Vergleich mit einem Referenzbild oder durch Durchschnittsbildung über mehrere

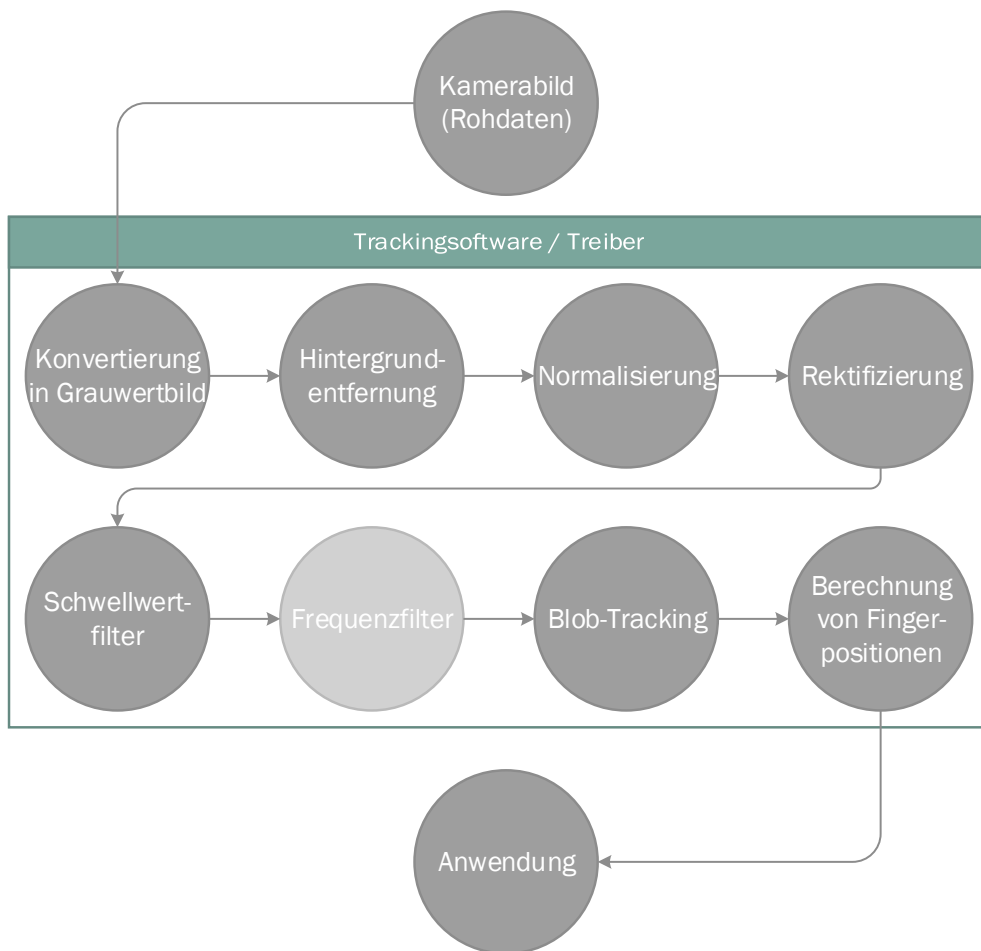


Abbildung 3.4: Bildverarbeitungs-Pipeline für Diffuse Illumination – Die Reihenfolge der Schritte variiert zum Teil zwischen verschiedenen Implementierungen. Der Verarbeitungsschritt *Frequenzfilter* ist optional und nicht in allen Erkennern vorhanden.

Frames. Die Vorverarbeitung schließt ebenfalls die Normalisierung der Helligkeit bzw. des Kontrastes und eventuell die Rectification, d.h. die Bereinigung des Bildes von etwaigen Verzerrungen durch die Optik der Kamera mit ein. In manchen Frameworks kann dieser auch am Ende der Verarbeitung stattfinden. In einem DI-Tracker werden in dem Bild dann über einen Schwellwert die hellsten Objekte herausgefiltert. Bei manchen Erkennern stellt der Schwellwert den Einzigen wesentlichen Filtermechanismus dar und Ergebnis der Filterung ist ein Schwarzweißbild, in welchem dann die zu erkennenden Objekte gesucht werden. Andere Tracker nutzen aus, dass durch den Diffusor die Unschärfe der Objekte mit dem Abstand von der Oberfläche stark zunimmt und setzen einen Highpass-Filter ein, um Objekte nahe der Oberfläche zu identifizieren. In jedem Fall folgt dann die Analyse der resultierenden Blobs anhand definierter Heuristiken. Beispielsweise wird nach elliptischen Objekten gesucht, die eine bestimmte Minimalgröße aufweisen und eine festgelegte Maximalgröße nicht überschreiten. Fortgeschrittene Verfahren versuchen noch bessere Ergebnisse zu erzielen, indem zeitliche und örtliche Zusammenhänge zu bereits erkannten Berührungen ausgenutzt werden.

Ein als Finger identifizierter Blob wird in vielen Systemen im Folgenden auf einen einzigen Punkt reduziert. Dies kann der geometrische Mittelpunkt oder auch ein *spitzes* Ende des Blobs sein. Diese Koordinaten werden dann als Berührung an die Anwendung weitergereicht. Dies stellt eine starke Vereinfachung dar und reduziert die Eingabemöglichkeiten von MT auf Multi-point (MP)-Interaktion. Die erweiterten Möglichkeiten, z.B. mit anderen Teilen der Hand oder des Armes zu interagieren oder Gesten zu benutzen, die sich nicht aus einzelnen Punkten herleiten lassen, gehen verloren. Auch die Form und Größe der Blobs oder ihre Orientierung könnten vielfach in die Interaktion mit einbezogen werden. Die aktuelle Spezifikation des TUIO-Protokolls bietet hierzu einige Optionen an.

3.3 Diffuse Illumination-Beleuchtungskonzept

Im Rahmen dieser Arbeit ist ein MT-Tisch – im Folgenden Mixed Reality Table (mrT) genannt (vgl. Abbildung 3.5) – entwickelt worden. Ziel war die Verbesserung des DI-Aufbaus im Hinblick auf eine möglichst gleichmäßige Be-



Abbildung 3.5: Fotografie des mrT (Abbildung aus Teichert u. a. (2010))

leuchtung, hohe Auflösung und Praktikabilität für den Forschungseinsatz. Der Hardwarebau wurde zusammen mit weiteren Kollegen der AG Digitale Medien durchgeführt. Eigene Beiträge des Autors zum mrT bestanden in: Definition der Anforderungen und Auswahl des DI-Konzeptes, Auswahl und Test von Infrarotfiltern, Hilfe bei Planung und Optimierung des Beleuchtungskonzeptes und Versuchsaufbauten zur optimalen Anordnung der Lichtquellen, der Kamera und des Projektors. Der mrT stellt mit seinem Beleuchtungskonzept einen Beitrag zum Stand der Technik für den DI-Hardwarebau dar (Teichert u. a., 2010).

Die Entscheidung für ein Tracking auf DI-Basis wurde getroffen, um ein Höchstmaß an Flexibilität zu gewährleisten. Insbesondere die Möglichkeiten, Marker und Objekte zu erkennen, waren von großer Bedeutung. Das grundsätzliche Prinzip der DI-Technik wurde bereits in Abschnitt 3.2.1 erläutert und gilt auch für den mrT (vgl. Abbildung 3.2).

Design und Konstruktion

Der mrT ist 80 cm breit, 130 cm lang und 112 cm hoch. Die Projektionsoberfläche hat eine Diagonale von 130 cm, was einer Fläche von ca. $0,72 m^2$ entspricht. Der Tisch verfügt über eine Full-HD-Projektion (1920 x 1080 Pixel). Die Umrandung der Oberfläche wurde soweit wie möglich verringert. Der Tisch besteht im Wesentlichen aus zwei übereinander liegenden Quadern, wobei der Fußteil kleiner ist als der Kopfteil, um den Nutzern mehr Beinfreiheit zu gewähren (vgl. Abbildung 3.5 und 3.6).

Abbildung 3.7 zeigt das Originalbild, aufgenommen von der eingesetzten Kamera bei Tageslicht. Die gute und konstante Ausleuchtung des mrT erlaubt es, die eingesetzten Bildverarbeitungsverfahren stark zu vereinfachen. Dazu wird einmal nach dem Einschalten eine Kalibrierungsprozedur durchgeführt. Für jedes Kamerapixel wird ein Standardkontrastumfang bestimmt, in dem die Helligkeit bei maximaler Reflexion (mit einer weiß reflektierenden Pappe auf dem Tisch) und bei minimaler Reflexion (ohne Pappe oder Gegenstände auf dem Tisch) in einer manuellen Kalibrierungsprozedur gemessen wird. Aus diesen Minima und Maxima kann dann für jedes Pixel ein Normalisierungsfaktor bestimmt werden. Dieser kann im Rahmen des ohnehin durchzuführenden Schwellwertverfahrens auf jedes aufgenommene Kamerabild angewendet wer-

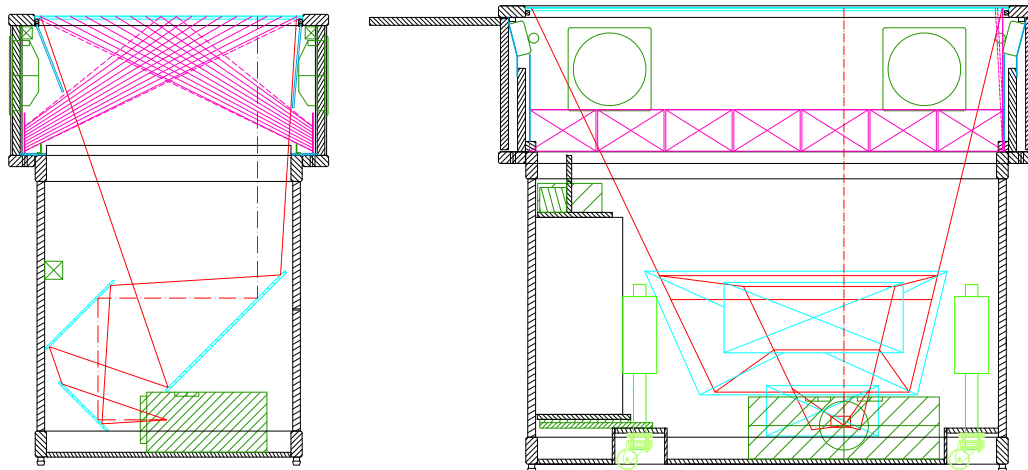


Abbildung 3.6: Schematische Darstellung des Aufbaus des mrT (Abbildung aus Teichert u. a. (2010)) – Die Projektion bzw. der Sichtbereich der Kamera werden über drei optische Spiegel gelenkt. Die Infrarotbeleuchtung ist in den oberen Bereich des Gehäuses integriert und beleuchtet jeweils die gegenüberliegende Seite der Oberfläche.

den. Dies erlaubt für den Rest der Verarbeitungspipeline die Festsetzung von konstanten Schwellwerten, um Objekte bzgl. ihrer Entfernung von der Oberfläche zu segmentieren (unter der Annahme einer gleichbleibenden Beleuchtung und dass alle Objekte das IR-Licht ungefähr in gleichem Maße reflektieren). Die Abbildungen 3.8 und 3.9 zeigen das normalisierte Bild und die beispielhaft durchgeführte Extraktion verschiedener Bildteile wie Finger, Hände und Arme anhand verschiedener Schwellwerte.

Oberfläche und Beleuchtung

Einige DI-Systeme verwenden Projektoren mit speziellen Weitwinkelobjektiven, um auch aus geringer Entfernung eine große Projektionsdiagonale zu erreichen. Diese führen aber oftmals zu Verzerrungen oder Farbrändern. Ein Standard-HD-Projektor benötigt eine Entfernung von ca. 150 cm für ein scharfes 16:9 Bild mit einer Diagonale von 130 cm. Um scharfe Full-HD-Bilder ohne Bildartefakte auf die Interaktionsoberfläche zu projizieren, wurde ein S-förmiges Umlenkensystem aus drei Spiegeln konstruiert. Das verwendete Modell vom Typ Panasonic PT-AE2000E verfügt über eine zweidimensionale mechanische Lin-

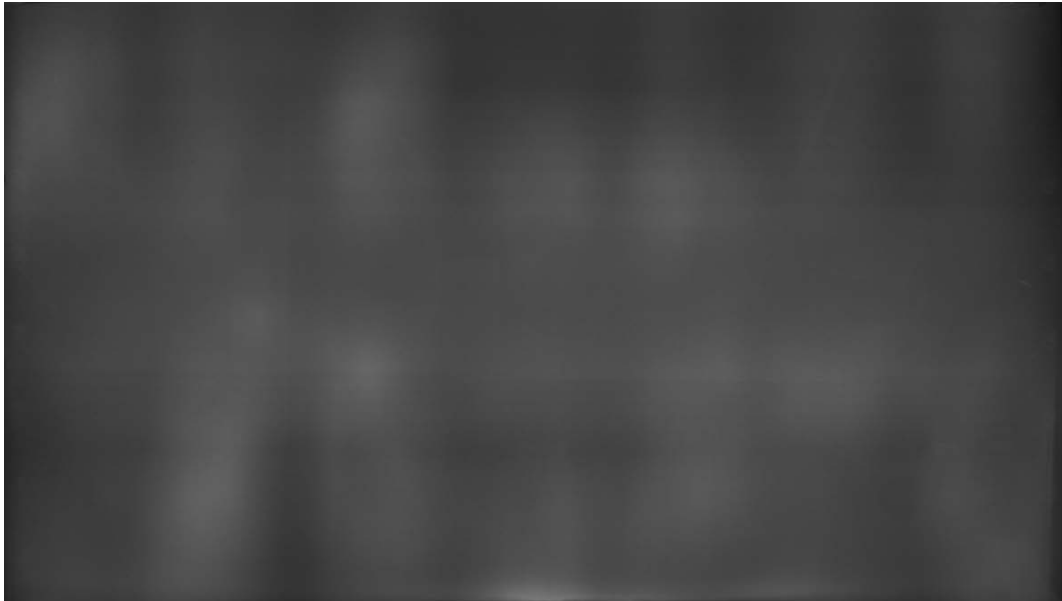


Abbildung 3.7: Originalbild (Rohdaten) der Interaktionsoberfläche aufgenommen durch die interne Infrarotkamera (Abbildung aus Teichert u. a. (2010))

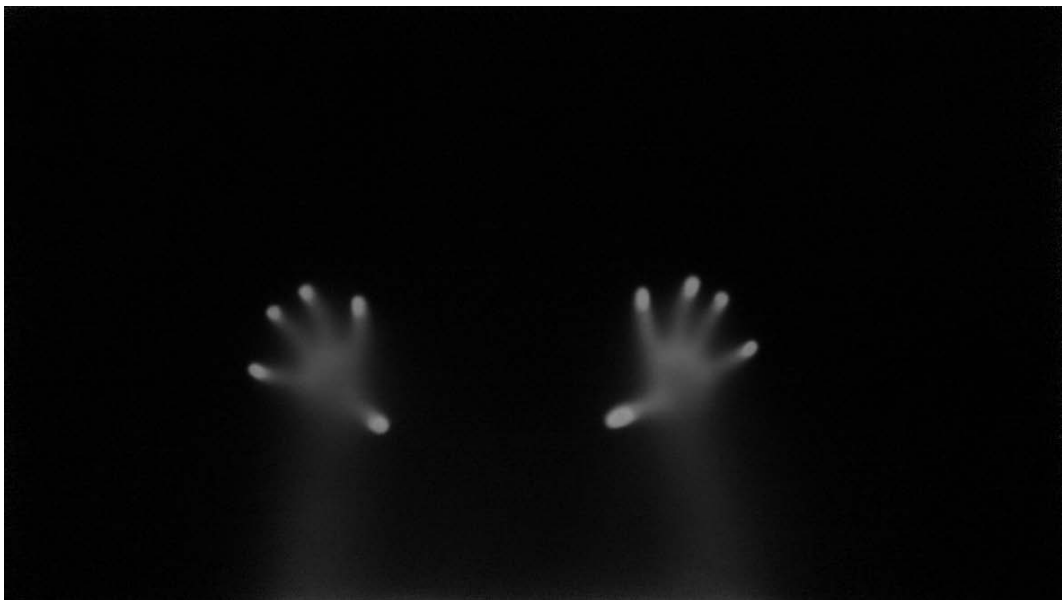


Abbildung 3.8: Hände auf der Oberfläche aufgenommen durch die interne Infrarotkamera (Abbildung aus Teichert u. a. (2010))

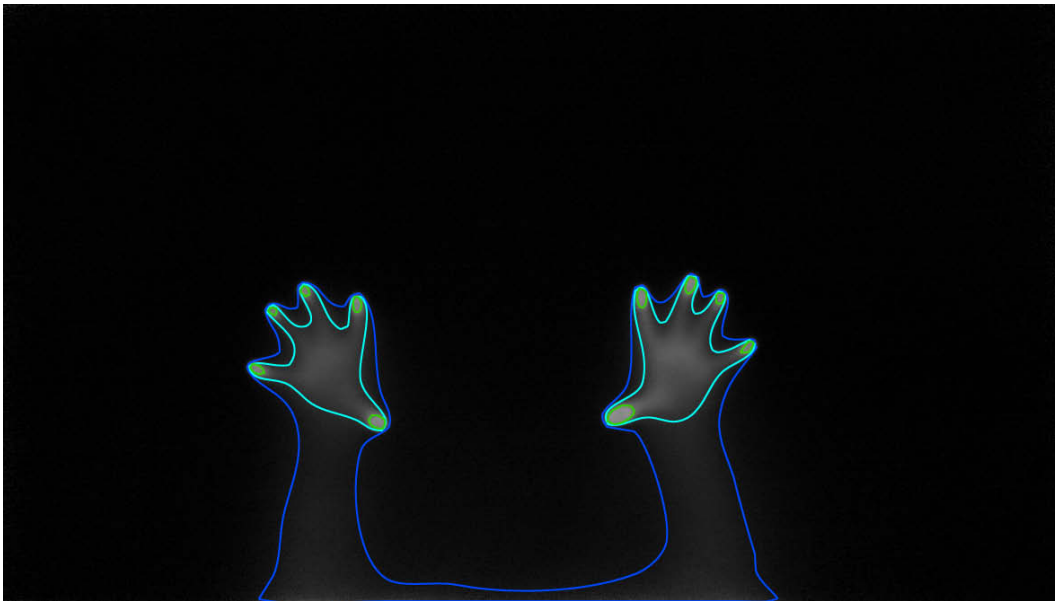


Abbildung 3.9: Extraktion verschiedener Features durch die Anwendung unterschiedlicher Schwellwerte auf Abbildung 3.8 (Abbildung aus Teichert u. a. (2010))

sensteuerung, die eine Feinjustierung der Projektion erlaubt und ohne Qualitätsverlust eine leichte Asymmetrie bzw. Faltung des Bildes ermöglicht, so dass der Projektor nicht exakt zentriert im Gehäuse untergebracht werden muss und genug Platz für den integrierten PC bleibt.

Ein Drei-Spiegelsystem ist zu einem gewissen Grad anfällig für Vibrationen. Außerdem können das Eigengewicht und die thermische Belastung zu Verformungen der optischen Spiegel führen, welche in einer ersten Version auftraten. Daher wurde in einer verbesserten Version die Aufhängung über Metallrahmen mit entsprechenden Dämpfern realisiert, an denen die Spiegel an mehreren Punkten befestigt sind.

Die eingesetzte Kamera ist eine μ Eye USB Kamera mit einer maximalen Auflösung von 1280x1024 Pixeln. Das Sichtfeld der Kamera wird ebenfalls über das Spiegelsystem umgelenkt. Ein Lensagon Zoomobjektiv vom Typ CVM45100 erlaubt eine optimale Abstimmung des Sichtfeldes auf die Projektionsoberfläche. Die Kameraauflösung kann flexibel angepasst werden und erlaubt so eine Balance zwischen Präzision und Aufnahmegeschwindigkeit.

Je nach Schwerpunkt der eingesetzten Anwendung wird eine Auflösung von

1280x750 Pixeln bei ca. 30 Frames pro Sekunde oder eine Auflösung von 1024x570 bei ca. 50 Frames pro Sekunde verwendet. Für die Fingerdetektion ist die niedrigere Auflösung ausreichend und erlaubt schnellere Bewegungen. Für die Detektion von Markern ist allerdings die Präzision wichtiger. Daher erfordert diese eine höhere Auflösung. Bei einer Auflösung von 1280x750 Pixeln wird über die gesamte Interaktionsoberfläche im Mittel eine Trackingauflösung von ca. 30 ppi erreicht.

Die Interaktionsoberfläche besteht aus einem zweischichtigen *Glassandwich*. Zwischen den beiden Glasscheiben befindet sich die Diffusions- bzw. Projektionsfolie. Um die optische Parallaxe zwischen Finger und Oberfläche bzw. Projektion so gering wie möglich zu halten, ist die obere Glasplatte lediglich 2 mm dick und dient hauptsächlich zum Schutz der Folie und zur Verbesserung der Haptik. Hierfür wurde aufgerautes Glas eingesetzt, welches sich angenehm anfühlt und ungewollte Spiegelungen auf der Oberfläche minimiert. Tests mit glattem Glas zeigten zwar ein etwas schärferes Bild, dafür waren aber störende Reflexionen vorhanden. Die untere Glasscheibe ist 6 mm dick und trägt den Hauptteil des Eigengewichtes. Die Diffusionsfolie dient gleichzeitig als Diffusor für das Infrarotlicht und als Projektionsfolie für das sichtbare Licht des Projektors. Nach Testläufen mit verschiedenen Folien und Filtern hat die verwendete LEE Hollywood Frost 255 schließlich den besten Kompromiss dargestellt.

Im mrT kommen insgesamt 14 Infrarotscheinwerfer zum Einsatz, die jeweils mit 180 Infrarot-LEDs bestückt sind. Eingesetzt wird ein LED-Typ mit besonders kleinem Abstrahlwinkel (Typ SFH-4550), um das Licht effektiver zu fokussieren. Auf jedem Scheinwerfer sind die LEDs in neun Reihen angeordnet. Die Reihen werden jeweils individuell im Winkel angepasst, um die Abstrahlung des Scheinwerfers entsprechend aufzufächern. Die Schaltung für die Scheinwerfer ist in Abbildung 3.10 dargestellt. Insgesamt hat die Beleuchtung eine Leistung von 160 Watt.

Um ungewollte Reflexionen zu minimieren, wurde in empirischen Tests nach den optimalen Winkeln gesucht. Während eine Anstrahlung in spitzem Winkel die Lichtleistung besser ausnutzt, hat sich nach Tests mit verschiedenen LEDs (z.B. mit breitem Abstrahlwinkel) und zusätzlichen Diffusoren gezeigt, dass die ungewollten Reflexionen auf diese Weise nicht zu kontrollieren sind. Durch den spitzen Anstrahlwinkel wird eine relativ starke Reflexion nahezu direkt in

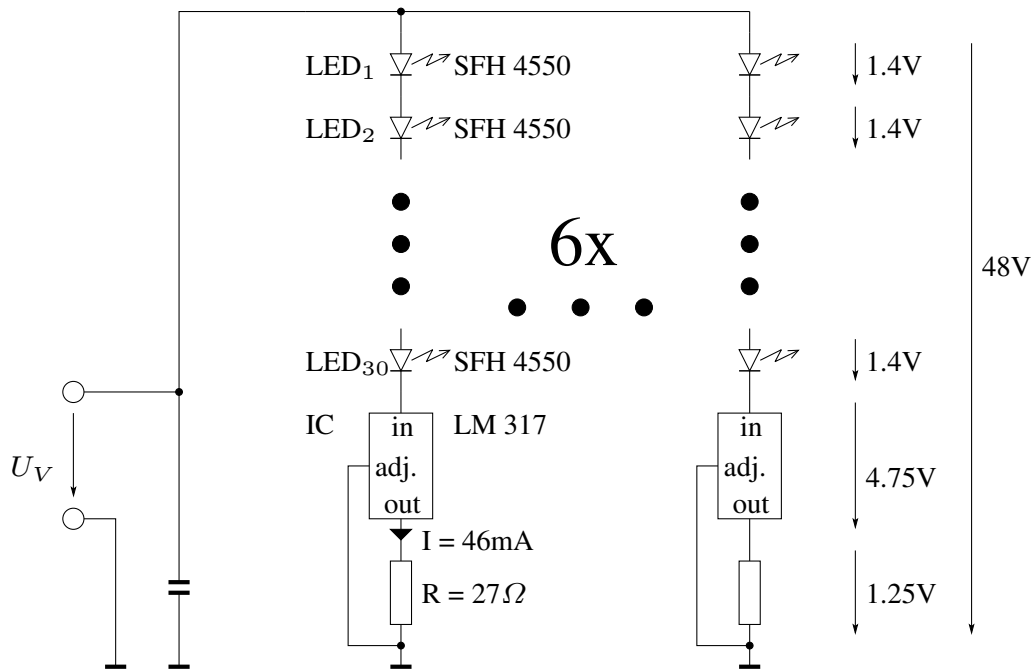


Abbildung 3.10: Schematische Darstellung der IR-Scheinwerferschaltung (Abbildung aus Teichert u. a. (2010))

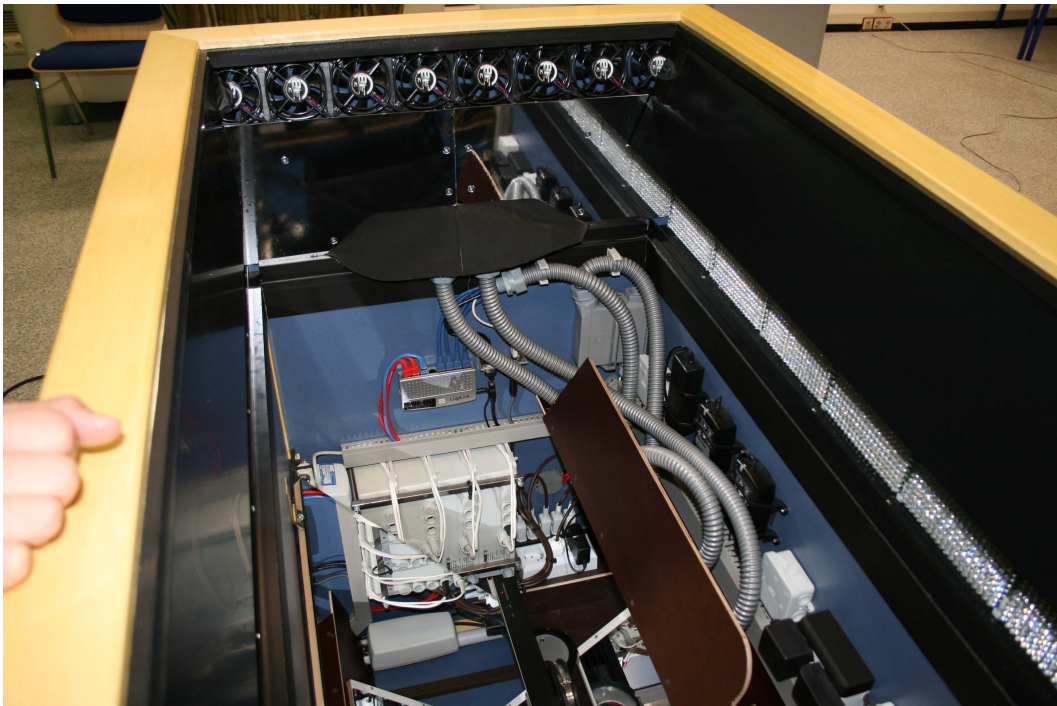


Abbildung 3.11: Spiegel an den kurzen Seiten verlängern optisch die Scheinwerferzeilen (Abbildung aus Teichert u. a. (2010)).

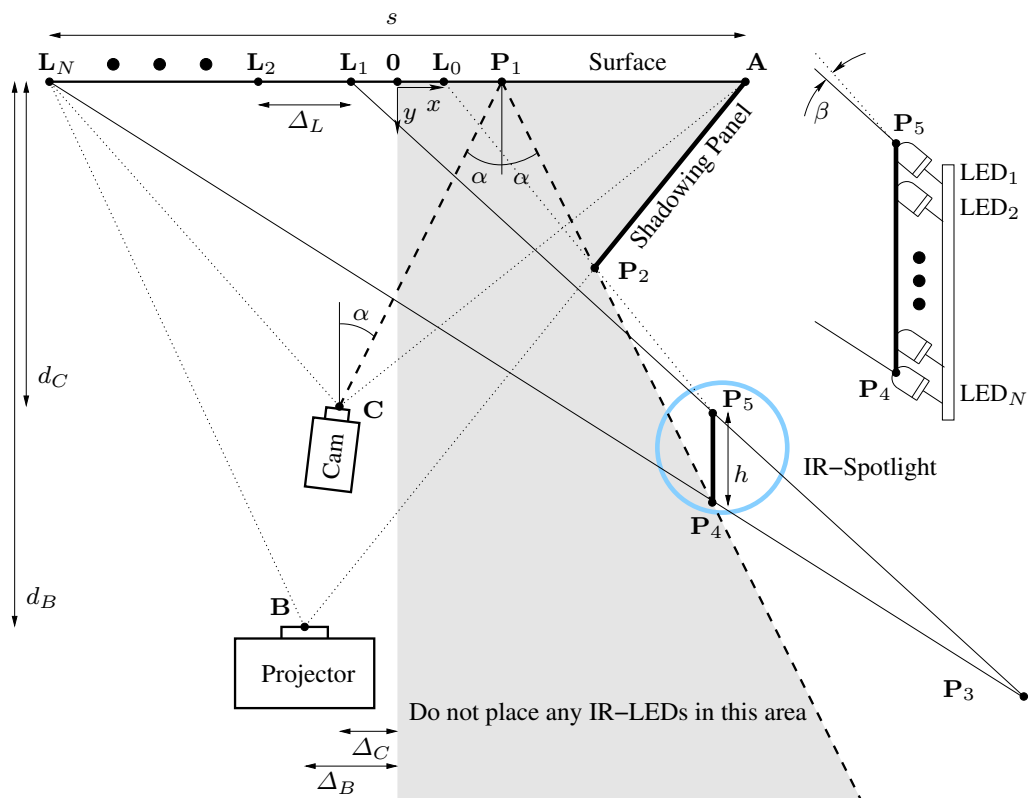


Abbildung 3.12: Die Anordnung der Oberfläche, der IR-Beleuchtung, der Kamera und des Projektors muss die grundlegenden Anforderungen wie im Schema dargestellt einhalten, um ungewollte Reflexionen zu verhindern (Abbildung aus Teichert u. a. (2010)).

Richtung der Kamera erzeugt. Daraus ergibt sich, dass ein flacher Anstrahlwinkel notwendig ist, weshalb die Scheinwerfer überkreuz angeordnet wurden. Die Scheinwerfer auf der linken Seite beleuchten die rechte Oberflächenhälfte und umgekehrt. Verbleibende Reflexionen sind durch Abschattungspaneel effektiv zu kontrollieren. Neben ungewollten Reflexionen entstehen an den schmalen Seiten der Interaktionsoberfläche dunkle Bereiche, da hier keine Scheinwerfer vorhanden sind. Diese können durch weitere Spiegel (vgl. Abbildung 3.11) an den Seiten ausgeglichen werden.

Das Gesamtkonzept der Beleuchtung ist in Abbildung 3.12 dargestellt. Durch die individuelle Ausrichtung der LED-Reihen wird sichergestellt, dass diese in gleichen Abständen auf die Oberfläche projizieren. Abstand und Winkel der Lichtprojektion gleichen sich dabei so aus, dass die Lichtdichte insgesamt über

die Oberfläche nahezu konstant und homogen ist. Die Projektion des LED-Lichtkegels auf die Oberfläche sollte eine bestimmte Größe haben. Diese sollte so bemessen sein, dass einzelne LEDs keinen zu stark fokussierten Spot erzeugen und andererseits die Beleuchtung präzise ausgerichtet werden kann, um ungewollte Reflexionen zu verhindern. Es sollte ein LED-Typ gewählt werden, so dass in Abhängigkeit von Projektionswinkel und -abstand der Radius der Projektion ungefähr dem Abstand zwischen den individuellen Projektionszentren der LEDs entspricht. Eine exakte Berechnungsvorschrift für die spezifischen Maße und Aufbauten des mrT findet sich bei Teichert u. a. (2010).

3.4 Hover-Detektion

Die DI-Technik ermöglicht grundsätzlich die Erkennung von Objekten in einem begrenzten Bereich auch über der Oberfläche. Durch die Diffusionsschicht werden diese Objekte mit zunehmendem Abstand von der Oberfläche jedoch sehr schnell unscharf und dunkler. Daher ist zu untersuchen, ob und mit welcher Zuverlässigkeit Finger über der Oberfläche erkannt werden können. Ein erster Schritt zur Klärung dieser Fragen ist im Rahmen dieser Arbeit geleistet worden. Die Implementierung und der Testlauf wurden von Jannis Stoppe im Rahmen seiner durch den Autor betreuten Masterarbeit (Stoppe, 2009) ausgeführt.

Eine zuverlässige Hover-Detektion würde die Interaktionsmöglichkeiten erheblich erweitern. Das System könnte bereits reagieren bevor der Benutzer die Oberfläche berührt. Dadurch wäre es möglich, einen der wesentlichen Vorteile der Mausinteraktion (vgl. Kapitel 4, Abbildung 4.1) in die Touch-Interaktion zu integrieren. Es könnte z.B. zwischen aktiver Manipulation und explorierender Interaktion unterschieden werden. Weit verbreitet in klassischen WIMP-Anwendungen sind beispielsweise Tooltips, die einen Hilfetext beim Hovering über einem GUI-Element bereitstellen.

Es stellt sich außerdem die Frage, ob eine Lösung zur Hover-Detektion prinzipiell in die DI-Bildverarbeitungspipeline von existierenden Erkennern zu integrieren ist. Im Gegensatz zu dem Ansatz von Hilliges u. a. (2009) soll auf ein zweites Beleuchtungssystem und zusätzliche Spezialkameras verzichtet werden. Als Testhardware wird der mrT als Vertreter für DI-Systeme verwendet, der

grundsätzliche Ansatz lässt sich damit auf andere DI-Systeme übertragen.

3.4.1 Ansatz und Implementierung

Eine wichtige Voraussetzung für die Hover-Detektion ist die Normalisierung der Intensitätswerte im aufgenommenen Kamerabild, um Rückschlüsse über die Distanzen ziehen zu können. Implementiert wird die Normalisierung, indem die minimalen und maximalen Helligkeiten gemessen werden. Dazu werden jeweils Bilder der leeren Oberfläche und Bilder einer weißen Pappe auf der Oberfläche aufgenommen und verglichen. Mittels linearer Interpolation kann nun ein normalisierter Helligkeitswert je Kamerapixel errechnet werden.

Für die Hover-Detektion muss das übliche Schwellwertverfahren, welches bei der DI-Technik zum Einsatz kommt, erweitert werden, so dass verschiedene Schwellwerte entsprechend verschiedener Entfernungen des Fingers von der Oberfläche einbezogen werden können. Ergebnis dieser Filterstufe sind 1 bis n Binärbilder entsprechend der 1 bis n verwendeten Schwellwerte. Eine Normierung der Intensitätswerte durch die Kalibration ist wie eingangs erwähnt Voraussetzung, um auf Basis der Helligkeitswerte Rückschlüsse über die Entfernung eines Objektes von der Oberfläche ziehen zu können. Weitere Annahmen bzw. Einschränkungen müssen getroffen werden, um Störeinflüsse zu begrenzen:

Eine **gleichmäßige diffuse Beleuchtung** der Objekte ist gegeben. Durch die Beleuchtung verursachte lokale Ungleichmäßigkeiten in der Intensität, welche fälschlich als Finger erkannt werden könnten, sind nicht vorhanden bzw. werden durch die Beleuchtungskonstruktion minimiert.

Nur **diffus bzw. matt reflektierende Objekte** sollen erkannt werden. Stark oder ungleichmäßig reflektierende Objekte weisen lokale Helligkeitsunterschiede auf, die zu Fehlerkennungen führen.

Die Objekte weisen **ähnliche Reflexionseigenschaften** auf, so dass Helligkeitswerte miteinander verglichen werden können.

Es existieren grundsätzlich zwei Möglichkeiten, multiple Schwellwerte auf die Eingangsbilddaten anzuwenden: *exklusiv* oder *inklusive*. *Exklusiv* bedeutet, dass für jedes Schwellwertpaar genau ein Bild erzeugt wird und lediglich dieser Wertebereich verwendet wird, um das resultierende Binärbild zu erzeugen. *Inklusiv* dagegen bedeutet, dass nicht mit einem Schwellwertpaar, sondern nur

mit einem Schwellwert nach unten verglichen wird und jedes erzeugte Bild auch alle helleren Bereiche enthält bzw. diese mit in das resultierende Binärbild einbezogen werden. Dadurch entstehen keine *Löcher*, eventuell ist aber redundante Information vorhanden. Beide Varianten werden umgesetzt und verglichen (vgl. Abschnitt 3.4.2).

Das Zusammenspiel der einzelnen Stufen der Bildverarbeitungspipeline ist komplex, da die Einstellungen einzelner Stufen für eine gute Gesamtleistung im Allgemeinen eine Änderung oder Justierung anderer Stufen erfordern, so dass es schwer ist, einzelne Stufen *konstant* zu halten, um das globale Optimum zu finden. Zusätzlich zu dem Schwellwertverfahren werden daher verschiedene Varianten vor- und nachgelagerter Stufen implementiert.

Als Erkennungsverfahren wird als eine Variante auf das etablierte Verfahren aus dem reacTIVision Framework (Kaltenbrunner und Bencina, 2007) zurückgegriffen und als weiteres Verfahren eine selbstentwickelte Heuristik eingesetzt. Der Einsatz von Hough-Transformationen wurde bereits in Vortests auf Grund der hohen Zahl an Fehldetektionen verworfen.

Der reacTIVision-Algorithmus wird als Vertreter eines etablierten und zuvor auf dem mrT eingesetzten Erkenners gewählt. Als Alternative wird eine ähnliche, schneller zu berechnende Heuristik verwendet. Diese ergibt sich aus Breiten- und Höhentests gegen vom Benutzer festgelegte Grenzen und einem Vergleich des Flächeninhalts mit der eines idealen Kreises entsprechender Größe unter Einbeziehung einer Toleranz (in allen durchgeführten Tests wurde 30% Toleranz verwendet).

Redundante Daten im Fall der inklusiven Methode werden reduziert, indem erkannte Finger in Bildern mit einem größeren Schwellwert Priorität vor erkannten Fingern in Bildern mit niedrigerem Schwellwert in der gleichen Bildregion bekommen (diese werden entsprechend verworfen).

Implementiert werden außerdem der Einsatz eines Tiefpassfilters zur Noise-Reduzierung sowie das Caching von nicht mehr erkannten Fingern für einen Frame, um zu untersuchen, ob kurzzeitige Aussetzer das Ergebnis beeinflussen. Automatische Schwellwertverfahren werden nach informellen Tests im Vorfeld zunächst ausgeschlossen und stattdessen a priori festgelegte konstante Schwellwerte verwendet.

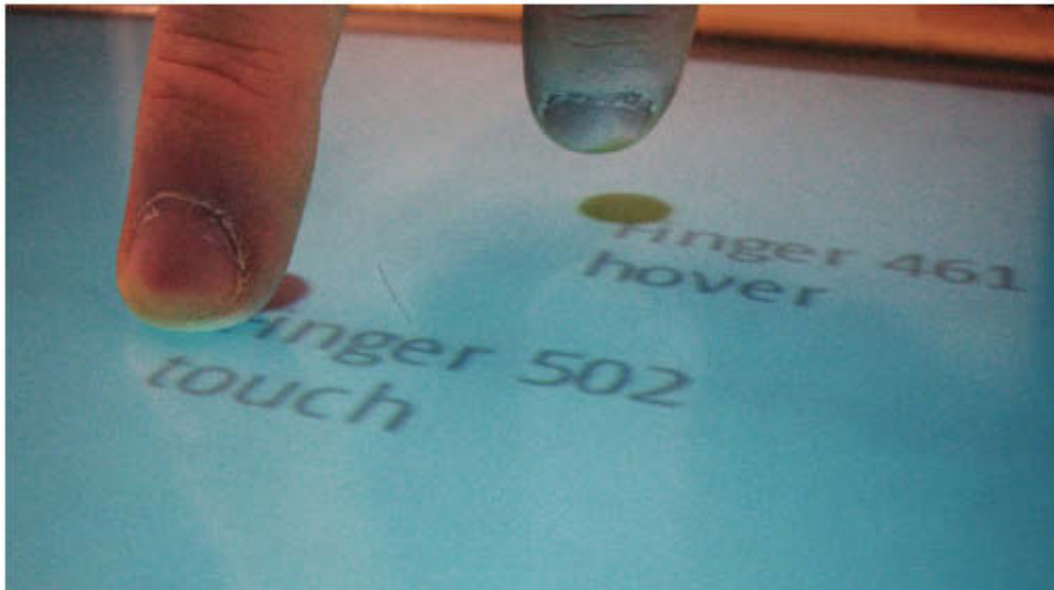


Abbildung 3.13: Das beschriebene Verfahren zur Hover-Detektion wurde in einem eigenständigen Erkennen zu Testzwecken implementiert (Abbildung aus Stoppe (2009)). Im Foto ist zu erkennen, dass die Fingerberührung links und der nicht ganz aufgesetzte Finger rechts korrekt erkannt und klassifiziert werden.

3.4.2 Testlauf

Um die Wechselwirkungen der verschiedenen im vorherigen Abschnitt genannten Parameter näher zu untersuchen und die grundsätzliche Machbarkeit des gewählten Ansatzes mittels multipler Schwellwerte nachzuweisen, wurde ein Testlauf durchgeführt. Neben der grundsätzlichen Machbarkeit sollte anhand quantitativer Kennzahlen ein erster Eindruck der Stärken und Schwächen dieses Ansatzes erlangt werden, um weitere Forschung in diesem Bereich gezielter durchführen zu können. Dieses Ziel ist bei der Betrachtung der Ergebnisse zu bedenken und eine zu tiefgreifende Interpretation der Ergebnisse auf Grund der Beschränkungen des Versuchs zu vermeiden. Dieser ist als Vortest für eine umfangreichere Untersuchung einzustufen.

Für die Untersuchung wurde eine Testapplikation entwickelt (vgl. Abbildung 3.13 und 3.14). Diese beinhaltet eine Start- und eine Zielzone. Mittels Hovering können kreisförmige Objekte von der Start- in die Zielzone bewegt werden. Dazu nimmt der Benutzer per Hovering Objekte in der Startzone auf,



Abbildung 3.14: Bildschirmfoto der Testapplikation (Abbildung aus Stoppe (2009)) – Die kreisförmigen Testobjekte sollen per Hovering von der linken in die rechte Zone bewegt werden.

indem er den Finger über dem Objekt in Hover-Distanz hält und den Finger dann über die Zielzone bewegt. Durch ein Anheben des Fingers über die maximale Hover-Distanz hinaus, kann das Zielobjekt dann in der Zielzone abgelegt werden. Während der Interaktion werden durch die Applikation automatisch folgende Daten erfasst und aufgezeichnet: Finger außerhalb der Zielobjekte, Finger auf den Zielobjekten, Zielobjekte, die außerhalb von Start- und Zielzone *fallen* gelassen werden, sowie Zielobjekte, die korrekt innerhalb der Zielzone abgelegt werden. Erkannte Finger bzw. Hovers außerhalb von Zielobjekten werden als *False Positive* (FP) und Zielobjekte außerhalb der Zielzone als *False Negative* (FN) gezählt. Dem zu Grunde liegt die Annahme, dass der Benutzer die Aufgabe präzise durchführt und Fehler der Erkennung anzulasten sind. Daher werden diese beiden Maße als Kriterien für die Beurteilung der Erkennung verwendet.

Gezählt wurden für jeden Durchlauf jeweils alle entsprechenden Ereignisse bis zum erfolgreichen Abschluss der Aufgabe, d.h. bis alle vier Kreisobjekte erfolgreich in die Zielzone transportiert worden waren. Dies bedeutet, dass bis zum Abschluss der Aufgabe alle Hovers, die nicht über einem Zielobjekt detek-

Regionen	Heuristik	Tiefpass	Verz. Entf.	Falsch Pos.	Falsch Neg.
Exklusiv	Eigene	ja	ja	79	34
Exklusiv	Eigene	ja	nein	61	30
Exklusiv	Eigene	nein	ja	201	102
Exklusiv	Eigene	nein	nein	373	207
Exklusiv	Reactivation	ja	ja	16	13
Exklusiv	Reactivation	ja	nein	21	16
Exklusiv	Reactivation	nein	ja	16	10
Exklusiv	Reactivation	nein	nein	19	9
Inklusiv	Eigene	ja	ja	54	11
Inklusiv	Eigene	ja	nein	35	4
Inklusiv	Eigene	nein	ja	41	5
Inklusiv	Eigene	nein	nein	56	13
Inklusiv	Reactivation	ja	ja	13	10
Inklusiv	Reactivation	ja	nein	21	12
Inklusiv	Reactivation	nein	ja	20	19
Inklusiv	Reactivation	nein	nein	20	15

Tabelle 3.1: Versuchsergebnisse aus mehreren Durchläufen mit einem Benutzer der Testapplikation (Tabelle aus Stoppe (2009))

tiert wurden, als FP gezählt wurden und mit jedem *Verlieren* eines Zielobjektes außerhalb der Zielzone der Zähler für FN erhöht wurde.

Für den Versuch wurden mit Hilfe der Testapplikation und einem freiwilligen Benutzer wiederholt Messungen vorgenommen, um jede der unterschiedlichen Varianten mindestens zweimal zu überprüfen. Als unabhängige Variablen wurden angenommen: das Verfahren zur Schwellwertbildung (inklusiv vs. exklusiv), die Erkennungsheuristik (reactIVision vs. flächenbasiert), die Filterung (Tiefpassfilter vs. kein Filter) und das Caching von nicht mehr erkannten Fingern (Caching über ein Frame vs. kein Caching). Durchgeführt wurde der Versuch auf dem mrT bei künstlicher Raumbeleuchtung ohne Tageslichteinwirkung.

3.4.3 Ergebnisse

Die Versuchsergebnisse sind in Tabelle 3.1 wiedergegeben. Wie eingangs erwähnt, bestand das Versuchsziel zunächst darin, eine Aussage über die grundsätzliche Machbarkeit des Ansatzes zu treffen und ggf. darüber hinaus eine

Zielrichtung für weitere Verbesserungen und Untersuchungen abzuleiten. Aus Zeitgründen konnte der Versuch lediglich mit einem Testnutzer durchgeführt werden. Bedingt durch die Art des Versuchsaufbaus und die Zielsetzung ist eine detaillierte statistische Auswertung daher nur eingeschränkt möglich bzw. aussagekräftig. Um trotzdem Anhaltspunkte zu gewinnen, wurden die in Tabelle 3.1 dargestellten Ergebnisse nach den unabhängigen Variablen gruppiert (Regioning, Heuristic, Filter, Delay) und für die abhängigen Variablen (FP, FN) wurden jeweils statistische Tests auf signifikante Unterschiede durchgeführt. Als Annahmen wurden zu Grunde gelegt, dass nicht von einer Normalverteilung ausgegangen werden kann und dass es sich um abhängige Gruppen handelt. Daher wurde für die Signifikanzprüfung der Wilcoxon-Test für abhängige Gruppen ausgewählt. Die Messungen wurden so gruppiert, dass jeweils nur die zu untersuchende unabhängige Variable für den Test abwich.

Lediglich die verwendete Heuristik hatte signifikanten Einfluss auf die FP. Die reactIVision Heuristik schnitt hier signifikant besser ab als die Flächenheuristik ($p = 0,012$). Bei den FN konnte dagegen keine Signifikanz nachgewiesen werden. Auch konnte sonst für die anderen unabhängigen Variablen weder für die FN noch für die FP ein statistisch signifikanter Unterschied ermittelt werden.

Betrachtet man dagegen nur die absoluten Werte, so war die Kombination aus exklusiver Schwellwertverarbeitung, Flächenheuristik und keiner Filterung besonders ungünstig und erzeugte mehr FP und FN als alle anderen Kombinationen. Betrachtet man das Mittel aus FP und FN, so ist die Kombination aus inklusiver Schwellwertverarbeitung, reactIVision Heuristik, Filterung und Caching Sieger des Versuches ($M = 11,5$).

3.4.4 Diskussion und Ausblick

Die dargestellten Ergebnisse lassen noch keine endgültige Beurteilung des Ansatzes zu. Eine vorläufige Aussage bezüglich der grundsätzlichen Machbarkeit des Ansatzes ist aber möglich und es lassen sich einige Hinweise zu grundsätzlichen Einschränkungen und zukünftigen Verbesserungen ablesen. Die prinzipielle Machbarkeit unter den beschriebenen Randbedingungen (DI, Integration in die vorhandene Pipeline, Echtzeitanforderungen usw.) konnte durch das Experiment belegt werden. Damit stellt der entwickelte Ansatz eine realistische

Möglichkeit zur Erweiterung bestehender MT-Systeme um Hover-Interaktion dar. Die im Experiment erreichte Erkennungsrate würde für einen Produktiveinsatz noch der weiteren Entwicklung bedürfen bzw. auf Seiten der Applikation eine entsprechende Fehlertoleranz voraussetzen.

Das zweite Versuchsziel bestand in der Einschätzung verschiedener Variationen der Bildverarbeitungspipeline. Hier lieferten die Ergebnisse, auf Grund des begrenzten Versuchsumfangs, nur eingeschränkte Resultate. Ein erweitertes Experiment wäre hier notwendig. Am Ende dieses Abschnittes wird daher kurz ein möglicher Versuchsaufbau skizziert.

Unter Berücksichtigung dieser Vorbehalte zeigt sich, dass das reacTIVision-Templatemarking in diesem Versuch der Flächenheuristik überlegen war. Es fanden sich Indizien dafür, dass Filterung und Caching positive Auswirkungen auf die Erkennungsleistung haben.

Einen Beitrag hat diese Arbeit in Bezug auf das Design einer Testapplikation geleistet. Diese könnte in verallgemeinerter Form zur Beurteilung von Erkennen und Bildverarbeitungspipelines eingesetzt werden. Bisher fehlen hier systematische Vergleichsmöglichkeiten.

Der grundsätzliche Versuchsaufbau, d.h. die Messung mittels einer Testapplikation in der Objekte von Start nach Ziel bewegt werden müssen, wäre auch für weiterführende Experimente geeignet. Allerdings sollte die Erfassung der Daten erweitert werden und nach Möglichkeit jede Zustandsänderung und die Positionen der Objekte sowie der Berührungen aufgezeichnet werden, um eine lückenlose Auswertung zu ermöglichen. Über den technischen Versuchsaufbau hinaus, war der beschriebene Versuch vor allem durch die Zahl der Testpersonen begrenzt. Da bei ähnlichen Interaktionstechniken oft große Varianzen zwischen Versuchsteilnehmern beobachtet werden konnten, wäre ein *within-subjects*-Design angemessen. Außerdem könnten andere Faktoren, wie Größe der Zielobjekte und insbesondere Ort und Distanz (auf dem Tisch) zwischen Start und Ziel weitere wichtige Faktoren sein. Daher sollten aus diesen Faktoren verschiedene Testszenarien bzw. -szenen für weitere Experimente erstellt werden.

3.5 Handunterscheidung

Vielen im weiteren Verlauf dieser Arbeit vorgestellten Techniken liegt als wesentliches Merkmal die Möglichkeit der beidhändigen Interaktion zu Grunde. Diese kann in verschiedenen Formen stattfinden. Einige Interaktionstechniken sind so ausgelegt, dass mit der zweiten Hand zusätzliche Operationen durchgeführt werden können, z.B. Objektmanipulation oder Kamerakontrolle (vgl. Kapitel 6). Bei Anderen wird eine Operation mit beiden Händen zusammen durchgeführt (vgl. Kapitel 4), z.B. eine Zoom-Geste mit den Zeigefingern beider Hände (Kin u. a., 2009).

Derzeit müssen beidhändige Interaktionstechniken so ausgelegt sein, dass diese keine Annahmen über die eingesetzte Hand machen oder der Benutzer muss dem System diese Information explizit mitteilen. Dies begrenzt die Anzahl der Funktionen und möglichen Abbildungen (Mappings). Wären Berührungen eindeutig einer Hand zuzuordnen und/oder eine Unterscheidung der starken und schwachen Hand bzw. der rechten und linken Hand möglich, so könnten beidhändige Interaktionstechniken dieses Kontextwissen verwenden. Im Rahmen dieser Arbeit wurde ein vorbereitender Schritt geleistet, um eine solche Erkennung zu realisieren.

Die untersuchte Methode basiert auf Erkenntnissen des im vorherigen Abschnitt beschriebenen Ansatzes zur Hover-Detektion. Es gelten technisch ähnliche Voraussetzungen. Die Implementierungsarbeit und der Testlauf wurden von Markus Aust im Rahmen seiner durch den Autor betreuten Masterarbeit (Aust, 2010) ausgeführt.

Ziel der Untersuchung ist, die Machbarkeit zu überprüfen. Inspiriert durch die Arbeit von Dang u. a. (2009), welche sich mit der Zuordnung von Berührungen zu Händen beschäftigt und der vorgestellten Ergebnisse zur Hover-Detektion, wird ein Ansatz auf Basis von DI vorgestellt, der keine zusätzliche Hardware (vgl. Hilliges u. a. (2009)) oder Instrumentation (vgl. Marquardt u. a. (2010)) des Benutzers voraussetzt.

3.5.1 Verwandte Arbeiten

Dem in den folgenden Abschnitten vorgestellten Ansatz am nächsten steht die Arbeit von Dang u. a. (2009). Diese untersuchen einfache Heuristiken auf

Basis anatomischer Gegebenheiten der Hand, um festzustellen, ob verschiedene Berührungspunkte von derselben Hand stammen. Als Heuristik werden dafür hauptsächlich typische Winkel und Längenverhältnisse zwischen den Fingern verwendet. Sofern die Finger relativ flach auf der Interaktionsfläche aufliegen, d.h. eine elliptische Berührung ergeben, kann außerdem diese Orientierung der Fingerellipse in die Erkennung mit einbezogen werden. Im Gegensatz zu der vorliegenden Arbeit beschäftigen sich Dang u. a. (2009) mit der Zuordnung von Fingern zu Händen, aber nicht mit der Identifikation bzw. Unterscheidung der Hände selbst.

Marquardt u. a. (2010) verwenden Handschuhe mit optischen Markern, mit denen sich sowohl die Hände als auch Teile der Hände wie Knöchel oder Finger explizit unterscheiden lassen. Da hierbei der direkte, nicht instrumentierte Charakter der MT-Interaktion verloren geht, ist dieser Ansatz eher als Übergangslösung zu beurteilen.

Im verwandten Bereich der Nutzerunterscheidung gibt es einige Arbeiten, deren Ansatz sich eventuell auch auf eine Handunterscheidung erweitern ließe, etwa der DiamondTouch (Dietz und Leigh, 2001), welcher mittels spezieller elektrischer Signale in der Lage ist, Berührungen und Benutzer einander zuzuordnen (vgl. Abschnitt 3.2.1).

Besonderes Merkmal des im Folgenden vorgestellten Ansatzes ist demnach der Verzicht auf zusätzliche Sensorik und Instrumentation des Nutzers.

3.5.2 Eigenschaften der menschlichen Hand

Das der Erkennung zu Grunde liegende empirische Interaktionsmodell bzw. dessen Annahmen beruhen auf den anatomischen Eigenschaften der menschlichen Hand (vgl. Schmidt und Lanz (2003)). Die Fähigkeiten und Grenzen der Beweglichkeit der menschlichen Hand bilden die Basis für grundsätzliche geometrische Bedingungen.

Die Finger (mit Ausnahme des Daumens) verfügen über einen Interaktionsbereich, der einem umgedrehten Dreieck gleicht, während der Daumen einen elliptischen Interaktionsbereich besitzt. Der durch den Daumen generierte Blob besitzt wegen des anderen Winkels eine etwas andere Form als der der übrigen Finger. Dieser ist normalerweise größer und die elliptische Form ist weniger symmetrisch.

Die Finger können ausgestreckt oder gebeugt werden, was unterschiedliche Reflexionen und entsprechend sichtbare Blobs zur Folge hat. Das Handgelenk kann ca. 30° nach links und rechts abgelenkt werden und ca. 60° nach oben und unten. Entsprechend kann die Lage der Fingerberührungen variieren. Auch bei aufgesetzten Fingern kann durch die große Flexibilität der Finger und vor allem des Handgelenks, die Hand eher flach über der Oberfläche gehalten werden oder eher senkrecht von oben. Dadurch kann diese entweder als großer Blob neben den Fingern sichtbar sein oder völlig aus dem Bild verschwinden. Das Handgelenk kann außerdem über der Oberfläche gehalten oder bei der Interaktion aufgesetzt werden.

3.5.3 Empirisches Modell

Basierend auf den beschriebenen anatomischen Eigenschaften der menschlichen Hand und des Interaktionsverhaltens liegen dem empirischen Modell fünf Hauptannahmen zu Grunde:

Ohne Berührung findet keine Interaktion statt. Dies bedeutet, dass eine Handunterscheidung grundsätzlich nur bei wenigstens einer erfolgten Berührung stattfinden kann und nur dann die Verarbeitungspipeline ein sinnvolles Ergebnis generieren kann. Dies ist wichtig, da die weiteren Annahmen auf bestimmten Gegebenheiten bezüglich der relativen Positionen von Arm, Hand und Fingern beruhen, die andernfalls nicht als gegeben vorausgesetzt werden können.

Ein heller Finger-Blob befindet sich stets in der Umgebung eines dunkleren Blobs der restlichen Hand-/Armstruktur über der Oberfläche. Dies bedeutet, dass ähnlich wie bei der Hover-Detektion in Abschnitt 3.4 zwischen Fingern und Hand/Arm mittels verschiedener Schwellwerte unterschieden werden kann und dass darüber eine direkte Zuordnung der Finger zu der gemeinsamen Hand möglich ist. Wird die Hand wie im vorherigen Abschnitt kurz erwähnt, annähernd senkrecht bei der Interaktion gehalten oder die Finger sehr weit ausgestreckt, wird diese Annahme offensichtlich verletzt und solche Fälle können durch den beschriebenen Ansatz daher nicht korrekt erkannt werden. Eine Kombination mit den Heuristiken von Dang u. a. (2009) könnte hier eventuell eine Lösung bieten.

Für das vorliegende Modell wird die Annahme getroffen, dass *Fingerberüh-*

rungen keine verlässliche elliptische Orientierung aufweisen, da die Form der Finger-Blobs zwischen kreisförmig und elliptisch variieren kann. In einem erweiterten Modell könnte in Ergänzung zur zweiten Annahme die Orientierung des gesamten Fingers mit einbezogen oder durch eine zeitübergreifende Erkennung Datenlücken ergänzt werden, wodurch diese Einschränkung aufgehoben werden könnte.

Es wird außerdem angenommen, dass *die umliegenden Hand- bzw. Arm-Blobs stets eine zuverlässige Orientierung aufweisen*. Auf Grund von Form und Größe von Hand und Arm weisen diese Blobs eine deutliche elliptische Form und eindeutige Orientierung auf und sind nicht kreisförmig.

Finger-Blobs befinden sich stets am Rande des korrespondierenden umgebenden Hand-/Arm-Blobs. Diese Annahme erlaubt die Eliminierung von Fehldetektionen, die sich zu weit entfernt oder an der falschen Stelle in Relation zur Hand befinden.

3.5.4 Implementierung

Die Klassifizierung der Berührungen erfolgt mittels Entscheidungsbäumen, welche mit verschiedenen Hand-/Fingerpositionen trainiert werden. Die Verarbeitungspipeline entspricht weitgehend dem DI-Standard (vgl. Abschnitt 3.2.1). Wie bei der Hover-Detektion werden verschiedene Schwellwerte eingesetzt, um anhand der Intensität verschiedene Tiefenbereiche im Kamerabild zu unterscheiden. In informellen Tests hat sich Zweifach-Tresholding als ausreichend erwiesen.

Um die Daten mittels der Entscheidungsbäume zu klassifizieren, muss aus diesen ein geeigneter Eingabevektor gebildet werden. Als Eingabedaten werden die Größen und Positionen der Blobs, sowie die Orientierung des Arm-Blobs genutzt. Diese müssen in geeigneter Weise normiert und im Eingabevektor angeordnet werden. Außerdem muss die Erkennung robust mit fehlenden Daten umgehen können.

Das Modell basiert auf der Annahme, dass Arm, Hand und Finger bei der Interaktion mit der Oberfläche typische geometrische Konstellationen bilden (vgl. vorheriger Abschnitt). Orientierung und Position des Arm-Hand-Blobs sind daher Hauptanhaltspunkte und geben ein Referenzkoordinatensystem für die Normierung der Finger-Blobs vor. Als Entscheidungskriterien von hoher

Priorität bilden diese somit die ersten Einträge im Eingabevektor und zwar in absoluten Koordinaten. Die Koordinaten der Finger-Blobs werden dagegen relativ zu dem durch den Arm-Hand-Blob aufgespannten Koordinatensystem in den Eingabevektor aufgenommen. Dieses Koordinatensystem ist wie folgt definiert: der Ursprung wird definiert durch die äußerste Ecke des Hand-Blobs und die Koordinatenachsen entsprechen den Hauptachsen des elliptischen Arm-Blobs. Innerhalb des Eingabevektors werden die Finger-Blobs nach ihrer Position von rechts nach links sortiert, um eine eindeutige Anordnung zu gewährleisten. Fehlende Daten werden mit Nullen aufgefüllt. Für die Implementierung wurde eine Standardimplementierung von Entscheidungsbäumen aus der Weka-Bibliothek¹ verwendet.

In der aktuellen Version findet keine zeitliche Betrachtung über mehrere Frames statt. Daher führt das Hinzufügen und Entfernen von Fingern unter Umständen dazu, dass sich die Reihenfolge der Finger im Eingabevektor ändert. Dies ist nicht optimal, da es die Güte der Finger-Blobs als Entscheidungskriterien senkt und entsprechend die Erkennungsleistung des Entscheidungsbaums. Eine Betrachtung über die Zeit, z.B. in dem bereits in vorherigen Frames erkannte Zuordnungen der selben Blobs bestehen bleiben, könnte die Erkennungsleistung deutlich steigern.

3.5.5 Testlauf

Ziel des Versuches war, die Erkennungsleistung zu quantifizieren und zu überprüfen, ob das Verfahren grundsätzlich umsetzbar ist. Als Kriterium hierfür wurde eine Erkennungsleistung von besser als 50% angesetzt.

Vorab muss eingeräumt werden, dass die im Rahmen dieser Arbeit durchgeführte Untersuchung eher den Charakter einer informellen Vorstudie hat. Obwohl daher noch viel Raum für Verbesserungen bleibt und eine formellere Untersuchung zur Sicherung der Daten unbedingt notwendig ist, geben die Ergebnisse zumindest einen Eindruck von der prinzipiellen Machbarkeit.

Um die Leistung zu überprüfen, wurde eine Testmenge von Gesten erstellt. Dazu wurden verschiedene Fingergesten und -positionen durch einen Testnutzer ausgeführt und aufgezeichnet. Die Auswahl der Testgesten selbst erfolgte

¹<http://www.cs.waikato.ac.nz/ml/weka/>

informell. Ziel war es, typische Gesten in verschiedenen Ausführungen abzubilden.

Der resultierende Testkorpus bestand aus insgesamt fünf verschiedenen Gesten: Berührung mit dem Zeigefinger, Berührung mit Zeigefinger und Daumen, Berührung mit Zeige- und Mittelfinger, Berührung mit Zeige-, Mittel- und Ringfinger und Berührung mit allen fünf Fingern.

Die Interaktion mit dem Zeigefinger ist die am häufigsten beobachtete Interaktionsform und wurde deshalb in den Testkorpus aufgenommen. Zeigefinger und Daumen werden häufig für die bekannte Pinch-Geste verwendet. Die übrigen Gesten wurden in den Testkorpus aufgenommen, da diese sich bereits in einigen Applikationen als Standardgesten etablieren konnten (z.B. unter Apple MacOS) und auch in einigen der im folgenden vorgestellten Anwendungen (vgl. Kapitel 5 und 6) Verwendung finden.

Jede dieser Grundgesten wurde zudem mit aufliegendem und angehobenem Handgelenk durchgeführt und dies insgesamt an zwölf gleichmäßig über die Interaktionsfläche verteilten Positionen und für beide Hände. Um den Testkorpus zu vergrößern, wurde jede Gestenkombination zweimal unabhängig voneinander aufgenommen. Der Gesamtkorpus umfasste damit insgesamt $5 * 2 * 12 * 2 * 2 = 480$ einzelne Testfälle.

In drei Durchläufen wurde der Testkorpus jeweils neu zufällig sortiert und die ersten 40 Datensätze wurden als Testmenge verwendet. Die übrigen Daten wurden zum Training der Entscheidungsbäume benutzt. Dazu wurde die Anzahl der verwendeten Trainingsdatensätze schrittweise inkrementiert und jeweils ein neuer Baum generiert, der dann gegen die Testmenge geprüft wurde.

3.5.6 Ergebnisse

In Abbildung 3.15 sind die Ergebnisse von drei unabhängigen Testläufen dargestellt. Wie beschrieben, basieren alle Tests grundsätzlich auf dem gleichen Korpus. Unabhängig voneinander wurde dieser Korpus für jeden Durchlauf neu randomisiert und jeweils 40 Beispiele als Testmenge entnommen. Die restlichen Beispiele wurden für das Training verwendet.

Wie deutlich zu sehen ist, schwankt die Erkennungsleistung zu Anfang sehr stark, überschreitet bei ca. 160 Trainingsbeispielen aber in allen drei Durchläu-

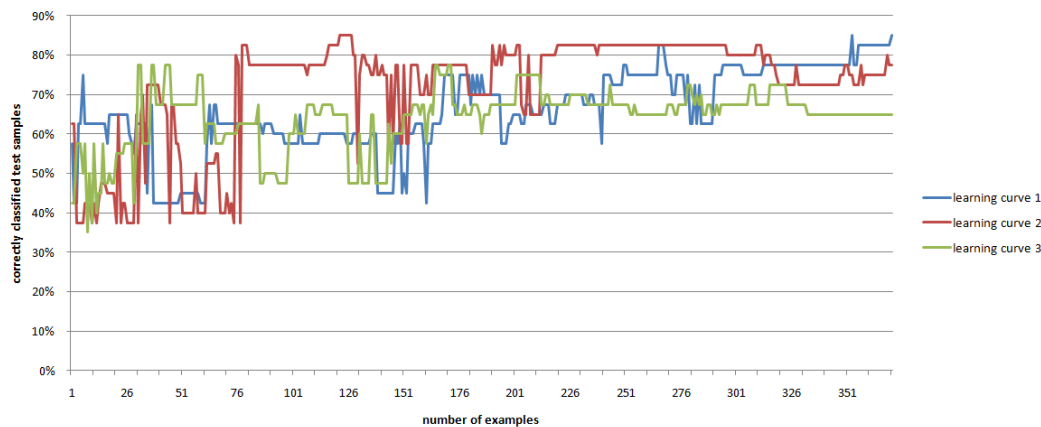


Abbildung 3.15: Erzielte Lernkurven für die Handunterscheidung (Abbildung aus Walther-Franks u. a. (2011a))

fen die 50% Marke. Schließlich wird im besten Durchlauf eine Erkennungsrate von ca. 85% und im schlechtesten Durchlauf von ca. 65% nach ungefähr 370 Trainingsbeispielen erreicht.

3.5.7 Diskussion und Ausblick

Die Ergebnisse des Experiments zeigen, dass die Erkennung grundsätzlich funktioniert. Wie sich die Erkennungsleistung bei der Verwendung durch verschiedene Personen entwickeln würde, kann daraus bisher nicht zwingend abgeleitet werden.

Da der Algorithmus immer noch mit einigen Einschränkungen behaftet ist, ist eine maximale Erkennungsleistung zwischen 65% und 85% im Versuch vielversprechend für dieses frühe Stadium.

Als grundsätzliches Problem hat sich während des Versuchs gezeigt, dass bei der Interaktion in den Randbereichen eine der Annahmen des Modells verletzt wird. Dann kann es passieren, dass nur wenig vom Arm zu sehen ist und dieser keine eindeutige Orientierung mehr aufweist. Was nicht zu sehen ist, kann in der Folge nicht erkannt werden. Dies könnte bedeuten, dass dieser Ansatz nur für Anwendungen Verwendung finden kann, bei denen sich die Interaktion auf zentrale Bereiche der Interaktionsfläche beschränkt. Eine andere Möglichkeit wäre, den optischen Trackingbereich größer zu gestalten als den Projektions- bzw. Anzeigebereich.

Zusätzlich könnte eine Betrachtung über die Zeit, d.h. über mehrere Frames, hilfreich sein und die Ergebnisse erheblich verbessern, indem bereits richtig zugeordnete Finger-Blobs diese Zuordnung behalten, auch wenn sie in folgenden Frames nicht eindeutig klassifiziert werden können.

Die Erweiterung des Algorithmus und die Überprüfung dieser These ist eine Aufgabe für zukünftige wissenschaftliche Arbeiten. Aus Anwendersicht wäre außerdem eine Integration des bisherigen Ansatzes in existierende Erkenner (z.B. CCV) wünschenswert, um diesen einer großen Zahl von Anwendungsentwicklern zugänglich zu machen.

3.6 Fazit

In diesem Kapitel wurden die Grundlagen und der aktuelle Stand der Technik im Bereich MT-Hard- und -Software präsentiert. Es wurden wesentliche Prinzipien bestimmter Hardwareansätze und deren spezifische Vor- und Nachteile im Vergleich diskutiert sowie grundsätzliche Anforderungen und Einschränkungen besprochen. Außerdem wurde das typische Softwareumfeld für die Entwicklung von Anwendungen in diesem Bereich erklärt. Eine Kenntnis von Hardware und Software in diesem Bereich ist wesentlich für das Verständnis der folgenden konzeptuellen und praktischen Überlegungen und Ergebnisse.

Als eigene Beiträge zum Stand der Technik wurden ein robustes DI-Beleuchtungskonzept sowie Ansätze zur Hover-Detektion und Handunterscheidung für DI-Hardware vorgestellt und diskutiert.

Insgesamt verdeutlicht dieser Abschnitt die Komplexität von neuen Benutzungsschnittstellen wie MT im Zusammenspiel von Hard- und Software. In den folgenden Abschnitten bildet dieses Kapitel die technische Grundlage für konzeptuelle Überlegungen zum Interaktionsdesign für 3D-Modellierung und für weitere im Rahmen dieser Arbeit durchgeführte praktische Versuche.

Kapitel 4

Interfacegestaltungsraum

Dieses Kapitel zeigt grundlegende und weiterführende Dimensionen für die Gestaltung von Multi-touch (MT)-Interfaces auf. Existierende Ansätze aus der Literatur werden vorgestellt und zusammen mit eigenen Überlegungen und Konzepten vor dem Hintergrund des Anwendungsfeldes 3D-Modellierung diskutiert.

Der folgende Beitrag ordnet existierende Designdimensionen bzw. Konzepte und führt bisher getrennte Taxonomien und Sichtweisen aus verschiedenen Bereichen der Forschung – z.B. HCI und Computergrafik – zusammen. Einige der Konzepte stammen aus dem HCI-Umfeld, wurden aber bisher hauptsächlich auf sehr fokussierte Testapplikationen angewandt und nicht auf die Eignung im speziellen Feld der 3D-Modellierung untersucht. Andere Konzepte stammen aus der Computergrafik, wurden aber bisher nicht in Zusammenhang mit MT-Schnittstellen bzw. generischen Konzepten aus der HCI gebracht. Die Betrachtung verschiedener Zielsetzungen bzw. Ausprägungen der Anwendungs- und Forschungsentwicklung im Bereich der HCI bzw. MT, d.h. inkrementelle Weiterentwicklung (*Emulation*) von Schnittstellen gegenüber der optimalen Ausnutzung neuer Eingabegeräte (*Paradigmenwechsel*) kommt hinzu.

4.1 Einleitung

Jede Eingabe- bzw. Ausgabeschnittstelle verfügt über spezifische Eigenschaften, die es als Interfacedesigner möglichst optimal zu nutzen gilt. Dies bedeutet, die Stärken der Schnittstelle zu nutzen und die Schwächen zu umgehen. Dies

gilt für etablierte Eingabegeräte wie Maus und Tastatur und auch für Eingabegeräte wie MT-Bildschirme bzw. interaktive Oberflächen. Buxton (2007) beschreibt diese grundsätzliche Herausforderung beim Interfacedesign mit dem Satz „everything is best for something and worst for something else“.

4.2 Stand der Technik

In der Literatur finden sich verschiedene Taxonomien, um die Vielzahl der Eingabegeräte und die Ausprägungsarten und Eigenschaften von Benutzerschnittstellen zu klassifizieren und zu ordnen. Dabei existieren verschiedene Ansätze bzw. Perspektiven, z.B. aus Sicht der Hardwareeigenschaften verschiedener Eingabegeräte oder bezogen auf mögliche Eingabegesten. Einige relevante Taxonomien decken allgemein den Raum der 3D-Eingabegeräte ab, andere beziehen sich auf interaktive Oberflächen oder auf bestimmte 3D-Interaktionstechniken. Daneben gibt es eine Reihe von allgemeinen Guidelines, Best Practices oder Normen aus dem Bereich der Usability und dem nutzerzentrierten Entwurf.

4.2.1 Klassifikation von Eingabegeräten

Die Taxonomie von Milgram und Kishino (1994) klassifiziert Mixed Reality (MR)-Anwendungen und Ein-/Ausgabegeräte. Milgram und Kishino (1994) beschreiben den kontinuierlichen Übergang von der realen Welt über Augmented Reality (AR) und Augmented Virtuality (AV) zu immersiver Virtual Reality (VR) und bezeichnen dies als „Virtuality Continuum“. Um diesen Bereich zu ordnen, klassifiziert die Taxonomie Interaktionssysteme anhand von drei Kriterien bzw. Dimensionen: „Extend of World Knowledge“, „Reproduction Fidelity“ und „Extend of Presence Metaphor“.

Interaktive Oberflächen können je nach Ausprägung (z.B. stereoskopisches 3D) und Anwendung bzgl. der ersten beiden Dimensionen nahezu das gesamte Spektrum abdecken. Bzgl. der dritten Dimension befinden sich diese vielfach im Bereich der *Large Screens* und damit zwischen klassischen Desktops und immersiven VR-Umgebungen.

Hinckley und Sinclair (1999) teilen Eingabetechniken nach der Art des Kontaktes (Berührung oder Druck sowie kein Kontakt) und der Eingabedaten (dis-

kret oder kontinuierlich) ein. Ihre Taxonomie bezieht sich lediglich auf die Eingabe und unterscheidet nicht zwischen integrierten Ein-/Ausgabegeräten (interaktiven Oberflächen) und reinen Eingabegeräten wie Touchpads.

Nach Grossman und Wigdor (2007) kann eine taxonomische Einordnung von Anwendungen und Interaktionstechniken auch auf räumlichen Kriterien bzw. der Dimensionalität der Interaktion beruhen. Sie unterscheiden Darstellungseigenschaften, Eingabeeigenschaften und physikalische Eigenschaften.

Wigdor und Wixon (2011) schlagen eine taxonomische Ordnung entsprechend der erkannten Objekte, der erkannten Information und der *Directness* des Eingabegeräts für den Bereich des Surface Computing vor. Sie diskutieren außerdem das *Mode-and-Flow*-Modell für Gesten, also die Modellierung von Touchinteraktion mittels endlicher Automaten (vgl. dazu das Modell von Buxton bzw. Hinckley und Wigdor (2012); Abbildung 4.1).

Eine Eigenschaft von MT bzw. *Natural User Interfaces* (NUIs) ist nach Wigdor und Wixon (2011) eine eingeschränkte Verallgemeinerbarkeit im Vergleich zu etablierten Eingabegeräten, da sich die Hardware in Features, Form und Einsatzzweck stark unterscheidet. Ähnliches findet sich bei Hinckley und Wigdor (2012). Diese klassifizieren Eingabegeräte und -modalitäten anhand von technischen Parametern wie der detektierten (physikalischen) *Eigenschaft*, der Zustände, die erkannt werden, der Anzahl der Freiheitsgrade bzw. Eingabedimensionen und der *Device Acquisition Time*. Sie beschreiben ebenfalls Zustandsmodelle für Touch- und Mausinteraktion (vgl. Abbildung 4.1). Diese zeigen wesentliche Unterschiede in den Zuständen und der Art der gemessenen Daten zwischen Maus und Touch auf, welche das Design von unabhängigen Interaktionstechniken erschweren (vgl. Hinckley u. a. (1998a)).

Die eingeschränkte Übertragbarkeit bzw. Abstrahierbarkeit von Gestaltungsregeln ist eine Motivation für die Bildung von abstrakteren Kategorien bzw. Gestaltungsdimensionen im folgenden Abschnitt. Auf diese Weise besteht die Möglichkeit, allgemeine Gestaltungsprinzipien und -muster auszudrücken.

Hinckley und Wigdor (2012) betonen, dass Interaktionstechniken nur im Kontext beurteilt und evaluiert werden können, da die Nutzbarkeit einer Technik stark von anderen verwendeten Techniken und den speziellen Anwendungszielen abhängt. Umso wichtiger erscheint es, die Möglichkeiten von MT nicht nur generell sondern auch anwendungsbezogen im Bereich der 3D-Modellierung

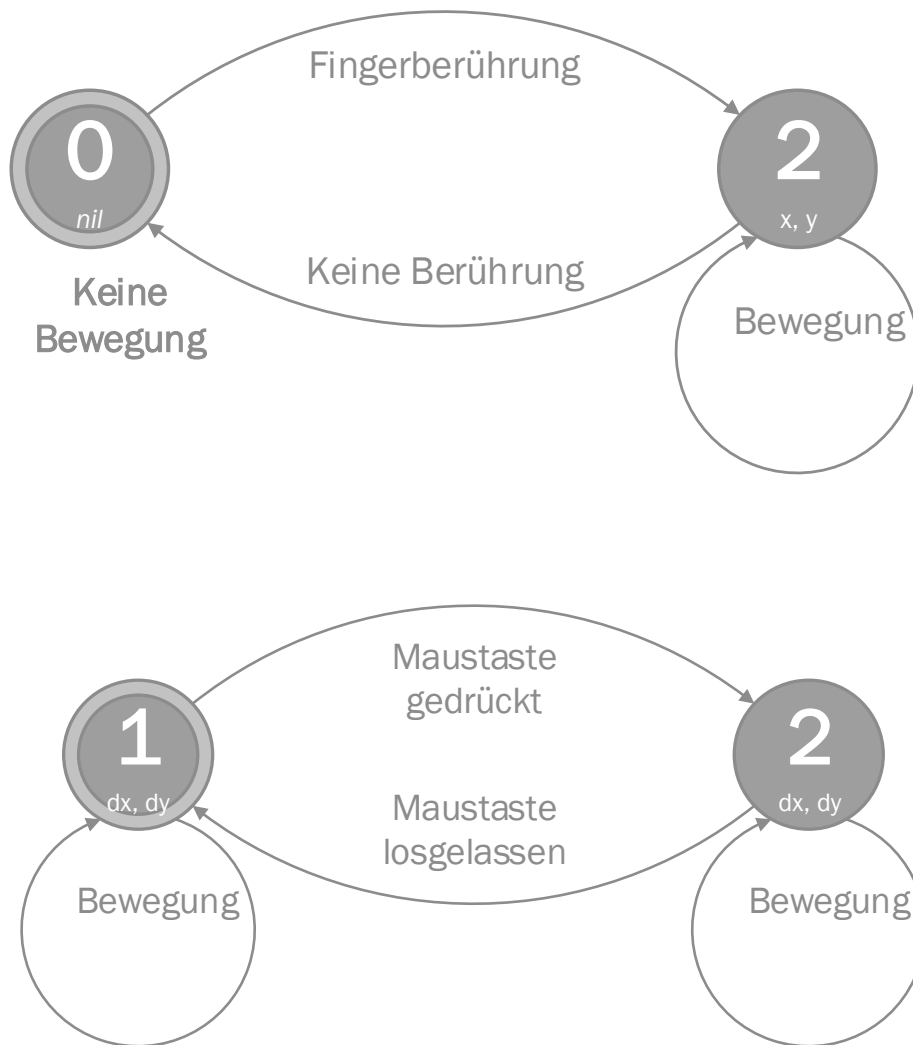


Abbildung 4.1: Zustandsmodelle für Touch- und Mausinteraktion nach Buxton bzw. Hinckley und Wigdor (Abbildung übersetzt aus Hinckley und Wigdor (2012))

zu erörtern. Daher folgt diese Arbeit dem Prinzip, sowohl fokussiert elementare Techniken zu betrachten, als auch die Integration in ein komplexes existierendes Anwendungsumfeld zu untersuchen. Nach Hinckley und Wigdor (2012) ist eine kompetitive Sicht auf verschiedene Eingabegeräte grundsätzlich falsch, vielmehr sollte eine komplementäre Sichtweise eingenommen werden. Je nach Anwendungsfall, Benutzer- und Benutzungskontext kann ein bestimmtes Eingabegerät seine Vorteile ausspielen oder können sich verschiedene Eingabegeräte ergänzen (vgl. Abschnitt 4.1).

Bei Bowman u. a. (2004) findet sich eine umfangreiche Übersicht zu weiteren etablierten Taxonomien für 3D-Eingabegeräte und 3D-Eingabetechniken. Sie zitieren die Taxonomie nach Buxton, welche 3D-Eingabegeräte anhand folgender Attribute ordnet: die Anzahl der Freiheitsgrade (1, 2, 6), die durch das Eingabegerät detektierte Eigenschaft (Druck, Bewegung, Position) und ob es sich dabei um ein direktes Toucheingabegerät handelt oder ob es eine indirekte Vermittlung zwischen dem Benutzer und dem Sensor gibt (Bowman u. a., 2004, Seite 130). Eine erweiterte Taxonomie unterscheidet zwischen Position, Bewegung, Kraft und Deltakraft sowie zwischen linearer Bewegung und Rotation (Bowman u. a., 2004, Seite 131).

Außerdem finden sich bei Bowman u. a. (2004) weitgehend anwendungsunabhängige Klassifikationen für Interaktionstechniken zur Selektion, Travel und Navigation und für die Manipulation von Objekten (Bowman u. a., 2004, Seite 149, 150, 190, 191). Die taxonomische Einordnung bzw. Klassifikation erfolgt alternativ entweder über eine Zerlegung in Subaufgaben oder über die verwendete Metapher. Selektion kann beispielsweise in Indikation des Objektes, Bestätigung der Selektion und Feedback (Bowman u. a., 2004, vgl. Seite 149) aufgeteilt werden. Jedes dieser Konzepte kann dann wiederum feiner zerlegt werden.

Für Manipulationstechniken geben Bowman u. a. (2004) eine taxonomische Ordnung über die Metapher an. Manipulationstechniken können in exo- und egozentrische Metaphern unterteilt werden. Die egozentrischen Metaphern wiederum können nach virtuellen Hand- bzw. Zeigetechniken (Bowman u. a., 2004, vgl. Seite 150) unterschieden werden. Beide Formen der Einteilung können dabei helfen, die Entwicklung oder die Evaluation eines 3D-Interfaces zu unterstützen.

Die im nächsten Abschnitt diskutierten Designdimensionen folgen im Gegensatz zu den Taxonomien aus Bowman dem Ziel, beim Designer ein Verständnis für die Qualität, die Bandbreite und die Zusammenhänge zwischen den Ausprägungsmöglichkeiten fokussiert auf 3D-Modellierung und interaktive Oberflächen zu unterstützen.

4.2.2 Klassifikation von Gesten

Karam und Schraefel (2005) präsentieren eine taxonomische Einordnung von Gesten als Eingabetechnik im Bereich der HCI. Sie klassifizieren gestenbasierte Techniken nach vier Hauptkriterien: Anwendungsdomäne, zu Grunde liegende Eingabetechnologie, Systemantwort und Gestentyp.

Wobbrock u. a. (2009) entwickeln eine Gestentaxonomie basierend auf ihrem Ansatz für die partizipative Entwicklung von MT-Gesten (User-generated Gestures) auf interaktiven Oberflächen. Hauptkategorien und Unterscheidungsmerkmale sind die Form (statisch, dynamisch, *one-point touch* usw.), der Typ der Geste (symbolisch, physikalisch, metaphorisch oder abstrakt), *Binding* (objektzentriert, weltabhängig/-unabhängig oder gemischt) und *Flow* (diskret oder kontinuierlich).

Freeman u. a. (2009) fügen der Taxonomie von Wobbrock u. a. (2009) weitere Kriterien hinzu. Grundsätzlich unterscheiden sie Gesten, welche die Interaktion einleiten, von Gesten, welche die Interaktion fortsetzen. Außerdem differenzieren sie zwischen verschiedenen Bewegungsarten der Hand.

4.2.3 Interfacegestaltung

In seinem grundlegenden Artikel beschreibt Shneiderman (1983) die Vorteile von direkter Interaktion bzw. Manipulation anhand typischer Beispiele und theoretisiert erste Ansätze, warum direkte Interaktion besser funktioniert und von den Benutzern besser angenommen wird.

Grundlegende Prinzipien, die Shneiderman hervorhebt, schließen die Wahl einer geeigneten visuellen oder physischen Repräsentation und das kontinuierliche Feedback für den Benutzer ein. Dadurch kann der Benutzer in intuitiver Weise die Auswirkungen einer Aktion abschätzen, welche ihm dann ggf. auch unmittelbar angezeigt werden und tendenziell wird nicht nur ein minimaler

Ausschnitt herausgehoben (z.B. einzelne Zeilen in einem Dokument), sondern diese werden im Kontext angezeigt.

Shneiderman und Plaisant (2010) geben eine gute Übersicht bzw. Einführung in die generellen Prinzipien von Interfacedesign und nutzerzentriertem Entwurf. Ihre Beispiele stammen oft aus dem Bereich der Webusability. Sie entwickeln die acht „Goldenen Regeln des Interfacedesigns“ (Shneiderman und Plaisant, 2010, vgl. Seite 88/89): „strive for consistency“, „cater to universal usability“, „offer informative feedback“, „design dialogs to yield closure“, „prevent errors“, „permit easy reversal of actions“, „support internal locus of control“ und „reduce short-term memory load“.

Grundsätzliche Konzepte zur *Affordanz* von *Dingen* und Schnittstellen finden sich bei Norman (2002). Er betont die Wichtigkeit des konzeptuellen bzw. mentalen Modells. Bereits durch die äußere Erscheinung sollte der Benutzer möglichst eindeutig auf Verwendungsmöglichkeiten eines Objektes bzw. einer Schnittstelle schließen können. Dazu muss er in die Lage versetzt werden, sich ein korrektes konzeptuelles bzw. mentales Modell der Benutzung zu bilden.

Norman (2002) beschreibt universelle Prinzipien, welche auch für die Gestaltung von Schnittstellen auf interaktiven Oberflächen Gültigkeit haben, allerdings müssen bei digitalen Schnittstellen teilweise Kompromisse bzgl. der Affordanz eingegangen werden. Bei einfachen Gebrauchsgegenständen und Vorgängen lässt sich die Affordanz z.T. gut auf Grund einfacher mechanischer bzw. physikalischer Mechanismen, die der Benutzer aus seiner Alltagswelt kennt, herleiten. Eine solche bei Bowman u. a. (2004) auch als *isomorph* bezeichnete direkte Korrespondenz ist bei komplexen, abstrakten Vorgängen aber nicht immer möglich und skaliert schlecht. Zudem kann diese Sicht den Designer auch einschränken (vgl. nächster Abschnitt).

Hinckley u. a. (1994) geben einen umfassenden Überblick und Designhilfen zu 3D-Interfaces, wobei sie sich speziell auf freie 3D-Eingaben beziehen, welche sie als räumliche Eingabe bezeichnen. Obwohl Hinckley u. a. (1994) in der Diskussion hauptsächlich auf 3D-Eingabegeräte abheben, lässt sich ein Teil der identifizierten Probleme und Guidelines auch auf interaktive Oberflächen übertragen. Die Feststellung, dass Menschen 3D weniger „explizit verstehen“ als vielmehr „erfahren“, scheint auch für interaktive Oberflächen relevant. Außerdem seien räumliche Referenzobjekte bei der Interaktion, insbesondere der

eigene Körper, hilfreich. Beidhändige Interaktion könne nicht nur die Effizienz verbessern, sondern auch das Verständnis und die Wahrnehmung unterstützen. Physikalisch-inspirierte Interaktion bzw. physikalisch-inspiriertes Verhalten der Interaktionsobjekte könne ein zusätzliches wichtiges und nützliches Feedback für die Interaktion bereitstellen. Die Struktur des wahrgenommenen Raumes solle mit der des Eingaberaumes möglichst übereinstimmen. Parameter, die vom Benutzer als abhängig wahrgenommen würden, sollten auch gemeinsam manipuliert werden können (vgl. nächster Abschnitt). Die gleichzeitige präzise Kontrolle mehrerer räumlicher Freiheitsgrade sei dagegen auf Grund motorischer Beschränkungen oft schwierig.

Einige der von Hinckley u. a. (1994) angeführten Probleme sind bei interaktiven Oberflächen nicht im gleichen Maße ausgeprägt. Beispielsweise sind physische Begrenzungen und taktiles Feedback durch die Interaktionsoberfläche für einige Anwendungen ein Vorteil gegenüber der 3D-Interaktion im freien Raum.

Neben wissenschaftlichen Papieren existieren auch relevante Normen und Standards. Die ISO-Norm 9241.10 führt beispielsweise folgende Grundsätze für die Dialoggestaltung in Benutzeroberflächen auf:

- Aufgabenangemessenheit
- Selbstbeschreibungsfähigkeit
- Erwartungskonformität
- Fehlertoleranz
- Steuerbarkeit
- Individualisierbarkeit
- Lernförderlichkeit

Zusätzlich finden sich in der ISO-Norm 9241.11 folgende Leitsätze für die Gestaltung interaktiver Systeme:

- Leicht zu erlernen
- Intuitiv zu benutzen

- Geringe Fehlerrate
- Zufriedenheit sicherstellen

4.3 Designdimensionen

Die möglichen Eigenschaften, die beim Design eines MT-Interfaces variiert werden können, sind vielfältig und zunächst nur schwer zu ordnen. Um diesen Bereich für die weitere Diskussion besser zu strukturieren, wird das gesamte Gestaltungsfeld in einzelne Designdimensionen (vgl. Abbildung 4.2) geordnet, jeweils charakterisiert durch ihre Extrempositionen. In ihrer Gesamtheit spannen diese den Raum der Gestaltungsmöglichkeiten für das Anwendungsfeld 3D-Modellierung auf interaktiven Oberflächen auf.

Die Auswahl der Designdimensionen wird hergeleitet, einerseits aus grundlegenden Hardwareeigenschaften von verschiedenen MT-Eingabegeräten, welche z.T. bereits in Kapitel 3 vorgestellt worden sind. Andererseits beruht die vorgeschlagene Ordnung auf grundlegenden Konzepten von 3D-Modellierungsanwendungen bzw. Schnittstellengestaltung im Allgemeinen, wie sie in der Literatur zu finden sind. Eine dritte Gruppe von Einflüssen stammt aus Untersuchungsergebnissen zu Nutzererfahrung (User Experience), Usability oder auch sozialen Faktoren in der Benutzung von interaktiven Oberflächen, welche im Rahmen dieser Arbeit bezogen auf 3D-Modellierung verwendet werden. Zusätzlich fließen eigene experimentell gewonnene Erfahrungen mit ein.

4.3.1 Multi-point vs. Multi-touch

Die Vorsilbe *multi* suggeriert bereits, dass es in jedem Fall um die Detektion von mehr als einer Fingerberührung geht. Dies wird im Kontext dieser Arbeit vorausgesetzt, muss aber im Allgemeinen natürlich zu Single-point- bzw. Single-touch-Systemen abgegrenzt werden, welche lediglich eine Berührung zuverlässig erkennen können. Wie in Kapitel 3 bereits angesprochen, variiert die genaue Anzahl der maximal unterstützten gleichzeitigen Berührungen von Gerät zu Gerät sehr stark. Einige wichtige Grenzen können hier definiert werden, da sie die Interaktionseigenschaften maßgeblich beeinflussen. Eine untere Grenze ist die Unterstützung von genau zwei gleichzeitigen Berührungen. Erstens

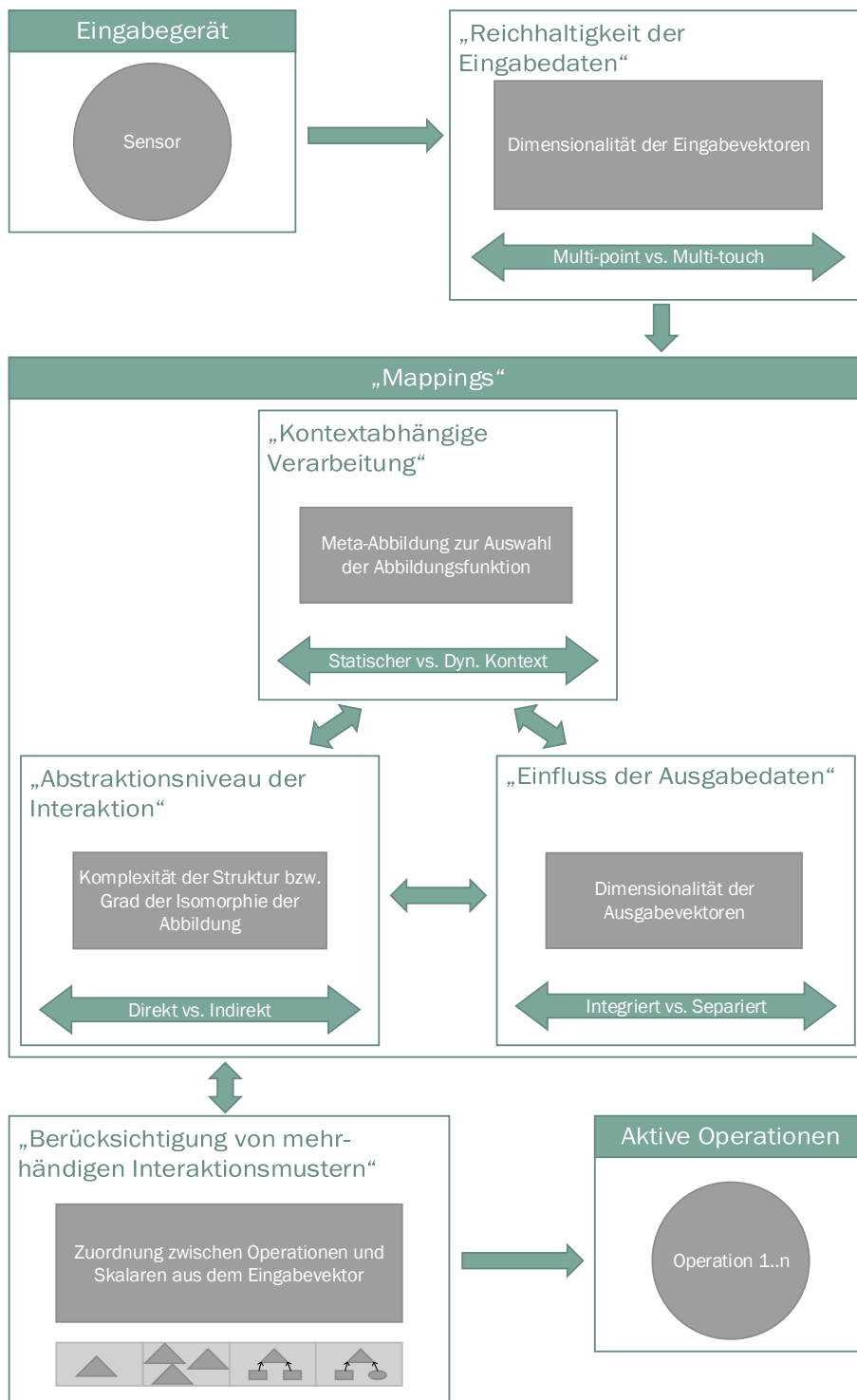


Abbildung 4.2: Schematische Darstellung aller diskutierten Designdimensionen und ihrer Beziehungen untereinander in Auswahl und Verarbeitung der Eingabedaten²

kann an dieser Stelle überhaupt erst von MT gesprochen werden, zweitens können ab zwei Punkten die Berührungen verschiedene Rollen übernehmen und wechselseitig wirken. So kann z.B. eine der Berührungen die Rolle eines Referenzpunktes übernehmen und es können relative Maße wie der Abstand der Punkte zueinander bestimmt werden. Es ist sowohl die Ausführung von Zweifingergesten mit einer Hand als auch mit einem Finger beider Hände möglich. Dies unterscheidet Dual-touch-Systeme qualitativ wesentlich von Single-touch-Eingabegeräten und erlaubt bereits die Ausführung einer großen Menge von MT-Interaktionstechniken, die im Wesentlichen Dual-touch voraussetzen, z.B. die Pinch-Geste oder einige Techniken von Moscovich und Hughes (2008) oder Joshi u. a. (2010). Es konnte außerdem beobachtet werden, dass die Benutzer in bestimmten Kontexten vorrangig mit dem Zeigefinger und einem weiteren Finger oder dem Daumen interagieren (Epps u. a., 2006). Die nächste interessante Grenze liegt bei vier unterstützten Berührungen. Diese ermöglicht es einem Benutzer, zwei Finger beider Hände gleichzeitig einzusetzen und somit zweihändig effektiv zu interagieren. Werden mindestens 10 Finger gleichzeitig unterstützt, kann der Benutzer beide Hände, vollständig zum Einsatz bringen. Entsprechend ermöglichen Vielfache dieser Grenzen die gleichzeitige Interaktion von mehreren Benutzern.

Die Dimension Multi-point vs. Multi-touch (vgl. Abbildung 4.3) bezieht sich auf die Reichhaltigkeit und Vielfältigkeit der zur Verarbeitung in der Anwendung vorhandenen Daten (Wigdor und Wixon, 2011). Das eine Ende des Spektrums markieren Eingabegeräte, welche sämtliche Informationen über Fingerberührungen auf einen einzigen Punkt verdichten (vgl. Kapitel 3 sowie Schöning u. a. (2008); Hinckley und Wigdor (2012)). Dies geschieht oft unabhängig davon, welche Informationen die Hardware tatsächlich zur Detektion und Unterscheidung der Fingerberührungen verwendet. Solche Geräte stellen

²Die Eingabedaten des Sensors können auf Eingabevektoren unterschiedlicher Dimensionalität abgebildet werden (z.B. nur Punktinformationen). Die schließlich wirksamen Abbildungen (Mappings) der Eingabedaten, gehen aus dem Zusammenspiel der drei zentralen Eingabedimensionen (siehe Kasten) hervor. Diese sind nicht vollkommen unabhängig voneinander, sondern repräsentieren verschiedene wichtige Schwerpunkte bei der Gestaltung der Abbildungen. Die Berücksichtigung mehrhändiger Interaktionsmuster wirkt wie ein zusätzlicher Filter für die Zuordnung von Eingabedaten zu Operationen bzw. beschränkt ggf. die Anzahl der aktiven Operationen.



Abbildung 4.3: Schematische Darstellung zu der Gestaltungsdimension *Multi-point vs. Multi-touch*

im kommerziellen Bereich derzeit nach wie vor die Mehrheit dar. Dem gegenüber stehen am anderen Ende des Spektrums Eingabegeräte, welche mehr als nur Punktinformationen zu den Fingerberührungen verarbeiten können. Der theoretisch maximale Fall in dieser Hinsicht wäre die vollständige Weitergabe der Rohdaten (je nach Trackingtechnologie) bis in die Endanwendung. Im Fall von optischem Tracking wären dies im Wesentlichen die Rohdaten der Kamera. Zwischen diesem Extrem auf der einen und der starken Reduktion der Daten auf der anderen Seite, liegt die Teilinterpretation und Reduktion der Daten durch die Hardware bzw. die vorgeschaltete Softwarepipeline, um der Anwendung einerseits nicht die vollständige Verarbeitung der Daten aufzubürden und andererseits eine reichhaltigere Interaktion zu ermöglichen. Das TUIO-Protokoll (vgl. Kapitel 3) beispielsweise sieht die Übertragung von zusätzlichen Daten wie Geschwindigkeiten und Beschleunigungen, aber auch die geometrische Blobform und die elliptische Orientierung vor. Die Windows 7 Touch-API unterstützt die Übertragung eines Flächeninhalts zu jeder Berührung.

Mehr und reichhaltigere Daten ermöglichen entsprechend komplexere Interaktionstechniken (Hinckley und Wigdor, 2012; Krueger u. a., 1985; Roudaut u. a., 2009; Wobbrock u. a., 2009; Wang u. a., 2009; Moscovich, 2009; Cao u. a., 2008; Wu und Balakrishnan, 2003; Freeman u. a., 2009), wobei die Nutzung

der Daten und deren Auswahl komplex sind. Da die Verarbeitung redundanter Daten Aufwand generiert, ist es in der Regel auch ein Ziel der Forschung festzustellen, ob schon mit weniger Daten gleich (oder fast gleich) mächtige Interaktionsmuster ermöglicht werden können. Es folgt nicht automatisch, dass reichhaltigere Daten vom Interfacedesigner auch effektiv genutzt werden können oder gewünscht sind. Die Reduktion der Daten auf der Skala von MT nach Multi-point kann die Robustheit und Präzision der Interaktion auch verbessern, da es sich letztlich um eine Form der Filterung handelt. Ebenso spielt die Verbreitung der Hardware eine Rolle und sollte im Rahmen von angewandter Forschung nicht außer Acht gelassen werden.

Verschiedene Arbeiten (Wilson u. a., 2008; Wilson, 2009; Cao u. a., 2008) zeigen das Potenzial von *echter* MT-Interaktion für die 3D-Modellierung mit physikbasierter Interaktion auf.

Bezogen auf die in der Einleitung angesprochene Klassifizierung bzgl. Emulation vs. Paradigmenwechsel ist eine klare Zuordnung der Extrempositionen dieser Dimension möglich. Multi-point ermöglicht die Emulation (insbesondere der Maus; vgl. Matejka u. a. (2009); Esenther und Ryall (2006); Walther-Franks u. a. (2011b)) und iterative Entwicklung neuer Interfaces, während *echtes* MT völlig neue Paradigmen und Techniken erfordert.

Mathematisch kann diese Designdimension auch als die Dimensionalität des Eingabevektors verstanden werden. Sei \vec{r} der Rohdatenvektor des Eingabegeräts und \vec{i} der resultierende Eingabevektor für die weitere Verarbeitung. Dann kann die Extremposition Multi-point als

$$\dim \vec{i} = 2 \tag{4.1}$$

und die Extremposition Multi-touch als

$$\dim \vec{i} = \dim \vec{r} \tag{4.2}$$

definiert werden.

4.3.2 Statischer vs. dynamischer Kontext

Die je nach Hardware- und Softwarepipeline unterschiedlich vielfältigen MT-Eingabedaten bzw. Freiheitsgrade sind nicht beliebig sinnvoll kombinierbar,

da Merkfähigkeit (Shneiderman und Plaisant, 2010; Norman, 2002) und Feinmotorik von menschlichen Benutzern natürlichen Grenzen unterliegen. Auch sind die Bewegungen von Fingern und Hand anatomischen Einschränkungen unterworfen (Sturman, 1992; Bowman u. a., 2004). Dies hat wesentliche Folgen für das gesamte Interaktionsdesign, aber insbesondere für die folgenden drei Dimensionen *statischer vs. dynamischer Kontext*, *integrierte vs. separierte Freiheitsgrade* und *direkte vs. indirekte Interaktion*.

Geeignete *Mappings* zu finden, ist eine der Hauptschwierigkeiten beim Design eines MT-Interfaces (Hinckley und Wigdor, 2012; Wigdor und Wixon, 2011; Bowman u. a., 2004). Diese und die folgenden beiden Gestaltungsdimensionen entsprechen verschiedenen Teilaspekten bzw. Sichtweisen solcher *Mappings*, d.h. der Abbildung von Eingabedaten auf Ausgabedaten bzw. Aktionen und Operationen. Diese sind daher nicht völlig orthogonal bzw. überdeckungsfrei diskutierbar. Sinn dieser Aufteilung ist die Möglichkeit der Priorisierung bei der Gestaltung eines System für einen speziellen Anwendungsfall. Auch wenn diese und die beiden folgenden Gestaltungsdimensionen nicht unabhängig variiert werden können, so kann es für ein System wichtiger sein, zunächst über die Implikationen der Dimension *integrierte vs. separierte Freiheitsgrade* nachzudenken und dann passend dazu die Dimensionen *statischer vs. dynamischer Kontext* bzw. *direkte vs. indirekte Interaktion* im Rahmen des Möglichen anzupassen.

Gerade komplexe Anwendungsbeispiele wie 3D-Modellierung benötigen eine Vielzahl von Operationen. Hier eine geeignete Abbildung zu finden, welche sämtliche Operationen abdeckt ist im Allgemeinen schwierig. Daher sind die Eingabefreiheitsgrade nicht dauerhaft bestimmten Operationen zugeordnet, wie dies in einfachen Anwendungen etwa dem bekannten „Bilder verschieben“ möglich ist. Grundsätzlich stehen also zwei Ansätze zum Umgang mit solchen Situationen zur Verfügung, welche die definierenden Extrempositionen dieser Designdimension ausmachen (vgl. Abbildung 4.4).

Einerseits können aus dem verwendeten mathematischen Modell und aus dem Anwendungsfall Heuristiken abgeleitet werden oder das Modell so angepasst werden, dass alle notwendigen Operationen zum richtigen Zeitpunkt zur Verfügung stehen oder es sich dem Benutzer, durch das Befolgen physikalischer Gesetzmäßigkeiten direkt erschließt, wie er die Objekte manipulieren

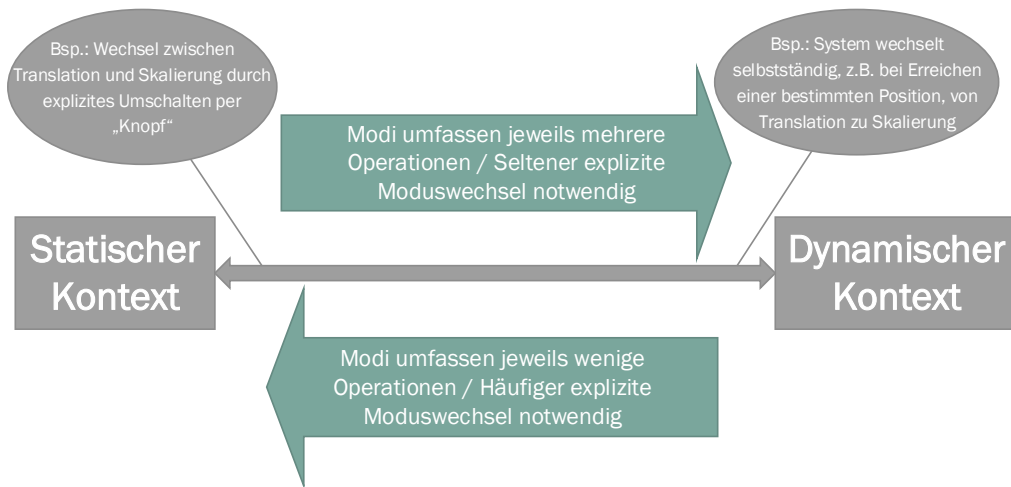


Abbildung 4.4: Schematische Darstellung zu der Gestaltungsdimension *statischer vs. dynamischer Kontext*

muss, um bestimmte Teiloperationen anzuwenden. Eine einfache Heuristik für den Wechsel zwischen verschiedenen Abbildungen könnte der Abstand zwischen zwei Objekten darstellen. Ein weiteres Beispiel ist die Abwandlung der Operation je nach dem an welcher Stelle am oder um das Objekt eine Interaktion stattfindet, sofern der Übergang für den Benutzer „fließend“ abläuft. Einige physikalische oder zumindest physikalisch-inspirierte Modelle fallen in diese Kategorie (Kruger u. a., 2005; Wilson u. a., 2008; Wilson, 2009; Hancock u. a., 2009a). Sämtliche Operationen sind für den Benutzer scheinbar gleichzeitig verfügbar bzw. in ein Modell integriert. Dies gelingt mit Hilfe mathematischer Approximations- bzw. Optimierungsverfahren (etwa Least-Squares, PCA oder Energiegleichungen), sofern die Verfahren alle gewünschten Operationen direkt in einem Modell abdecken (Arun u. a., 1987; Moscovich und Hughes, 2006; Reisman u. a., 2009; Herrlich u. a., 2011a). Die Unterscheidung bzw. Kombination verschiedener Transformationen, z.B. Translation plus Skalierung plus Drehung sowie das individuelle Gewicht bzw. die Rolle der einzelnen Berührungen oder Teile der Hand ergeben sich dann direkt aus dem mathematischen Modell, ggf. aus Stabilitätsgründen leicht gerundet oder gefiltert. Es bestehen jedoch Abhängigkeiten zwischen dieser Gestaltungsdimension und den beiden folgenden Dimensionen. Die Integration verschiedener Operationen in ein Modell kann ebenfalls die Separabilität verschiedener Freiheitsgrade sowie den

Grad der Direktion bzw. Indirektion beeinflussen.

Die andere Extremposition ist durch wesentlich einfachere mathematische Modelle gekennzeichnet. Es werden keine aufwendigen Approximationsverfahren eingesetzt, sondern lediglich Differenzen über der Zeit zwischen alter und neuer Position einer Fingerberührung gebildet. Diese Vektoren zusammen mit einfachen Metriken wie Länge und Winkel stellen die Eingabedaten für das System dar (Hancock u. a., 2006, 2007). Ob diese Eingabedaten dann vom System als Translation oder Skalierung zu werten sind bzw. welche Auswirkungen die resultierenden Transformationen haben, etwa eine Wirkung auf das gesamte Objekt, Objektteile oder die virtuelle Kamera, muss vom Benutzer manuell bestimmt werden. Das bedeutet, während der Interaktion müssen explizite Moduswechsel vorgenommen werden. Dies kann entweder durch MT-basierte Interaktionstechniken geschehen, z.B. durch Mehrfinger- oder Chording-Techniken (Matejka u. a., 2009; Herrlich u. a., 2011a; Aras u. a., 2011; Esenther und Ryall, 2006; Wigdor und Wixon, 2011) oder durch herkömmliche Interfacelemente wie Schaltflächen, Slider, Buttons, Auswahlboxen oder bestimmte Widgets (Bier u. a., 1993; Everitt u. a., 2005).

Beide Extrempositionen lassen sich auch über den jeweiligen Kontext charakterisieren. Mit Kontext ist hier der programmatische Kontext für die Ausführung einer Operation gemeint, d.h. der Zustand des Systems. Die direkte Ableitung von Transformationen kann als dynamischer oder „intelligenter“ Kontext verstanden werden. Teil des Kontextes der Operationen sind die Ergebnisse heuristischer Bewertungen, mathematischer Optimierungsverfahren und KI-Techniken. Das zu Grunde liegende Interaktionsmodell bzw. die Metapher kann auf die Operation *rückkoppeln*. Die Extremposition Moduswechsel dagegen kann so beschrieben werden, dass der Kontext der Operationen von außen manuell und explizit festgelegt werden muss und ansonsten statisch ist. Eine kontinuierliche Rückkopplung/Bewertung findet nicht statt. Hier besteht eine Gemeinsamkeit mit etablierten GUI-Elementen in 3D-Modellierungsanwendungen wie Handles, Gizmos oder auch Proxies. Solche Elemente bewirken einen expliziten Moduswechsel.

Zwischen diesen beiden Extrempositionen liegt ein Kontinuum von Mischformen, in welchen die Rückkopplung verschieden stark ausgeprägt ist bzw. in denen Heuristiken und explizite Festlegungen kombiniert werden.

Bzgl. der Bewertung Emulation vs. Paradigmenwechsel ist festzustellen, dass erst die zusätzlichen Freiheitsgrade von MT-Interaktion einige Approximationsverfahren anwendbar machen. Diese erlauben eine erweiterte Form der Interaktion ohne „mode lock-in“. Mehrfinger- und Chording-Techniken dagegen wurden erfolgreich zur Emulation von Mauseingaben für Anwendungen eingesetzt (Matejka u. a., 2009). Andererseits sind auch etablierte grafische Anwendungen heute bereits nach dem „direct manipulation“-Ansatz gestaltet (Shneiderman, 1983). Vielmehr ändert sich also die Qualität des Moduswechsels, indem dieser einen räumlichen Charakter annimmt (Everitt u. a., 2005; Shen, 2006).

Beide Extrempositionen bieten bezogen auf das Interaktionsdesign unterschiedliche Vor- und Nachteile. So benötigt der Ansatz des dynamischen Kontextes im Extremfall keine weiteren GUI-Elemente außer dem Zielobjekt selbst und erlaubt daher die Bildschirmfläche vollständig für die Interaktion mit dem Objekt zu nutzen. Falls ein entsprechendes Interaktionsmodell für die Operation gefunden werden kann, erlaubt dieser Ansatz eine sehr direkte, effiziente und intuitive Interaktion und kann von Interaktionskenntnissen aus der realen Welt profitieren (North u. a., 2009; Wilson u. a., 2008; Micire u. a., 2009). Im Allgemeinen ist aber keineswegs sicher, dass in jedem Fall ein solches Modell gefunden werden kann (Wobbrock u. a., 2009; Epps u. a., 2006). Passt das Modell nicht, so kann dies kontra intuitiv und ineffizient sein. In komplexeren Anwendungen besteht zudem das Problem der Trennschärfe zwischen den Kontexten und Operationen. Es ist mit einer Vielzahl von Operationen zunehmend unwahrscheinlich, dass für jede Operation ein geeignetes Modell gefunden werden kann und die Kontexte vom System dabei zuverlässig generiert, d.h. die Operationen zuverlässig unterschieden werden können (Wigdor und Wixon, 2011). Dies bedeutet, dass sich dieser Ansatz grundsätzlich eher für einige wenige Operationen eignet, also für ein System, welches eher wenige Operationen in ähnlichen Kontexten benötigt. Der Rechenaufwand auf Grund der eingesetzten Methoden ist groß. Außerdem können Wechselwirkungen mit den Designdimensionen *integrierte vs. separierte Freiheitsgrade* und *direkte vs. indirekte Interaktion* bestehen.

Im Gegensatz dazu ist der Ansatz des statischen Kontextes/Moduswechsels enger verwandt mit dem klassischen WIMP-Paradigma. Wie bereits erwähnt,

können viele GUI-Elemente in 3D-Anwendungen dieser Funktion zugeordnet werden. Durch den manuellen bzw. expliziten Kontextwechsel besteht kein Problem, auch eine große Anzahl ähnlicher Operationen eindeutig und zuverlässig zu trennen (Wigdor und Wixon, 2011; Everitt u. a., 2005; Matejka u. a., 2009). Daher ist die Wechselwirkung mit anderen Gestaltungsdimensionen hier einfacher kontrollierbar, da ggf. einfach entsprechende separierte Operationen bzw. Constraints hinzugefügt werden können. Der Moduswechsel bringt sowohl in der Darstellung als auch in der Interaktion einen gewissen Overhead mit sich, da dieser neben der Interaktion mit dem Zielobjekt auch explizit vom Benutzer bedienbar sein muss, was Zeit und/oder Bildschirmfläche kostet. Auf Grund der Vertrautheit vieler Benutzer mit dem etablierten WIMP-Paradigma (North u. a., 2009; Wobbrock u. a., 2009; North u. a., 2009; Terrenghi u. a., 2007; Epps u. a., 2006) kann ein solches Bedienungskonzept intuitiv wirken, allerdings ist es möglich, dass das Interaktionspotenzial von MT hier nicht zur Gänze ausgeschöpft wird.

Werden die Eingabefreiheitsgrade eindeutig und vollständig (bijektiv) auf die zu kontrollierenden Freiheitsgrade abgebildet, so wird von einer *isomorphen* Abbildung gesprochen (Bowman u. a., 2004). Dies impliziert somit auch eine Gleichzeitigkeit bzw. Integration der Interaktion mit allen Freiheitsgraden. Wie bereits erörtert, skaliert dies außerhalb von sehr begrenzten Testfällen nicht, und im Allgemeinen muss entweder die Eindeutigkeit oder die Vollständigkeit (oder beides) aufgehoben werden. Ein Moduswechsel kann als Meta-Abbildung betrachtet werden, mit welcher eine beliebige Anzahl von Abbildungen „gewechselt“ wird, welche in ihrer Vereinigungsmenge sämtliche Ausgabefreiheitsgrade abdecken kann. Somit kann also Vollständigkeit für die Meta-Abbildung garantiert werden. Die Teilfunktionen können entsprechend weniger komplex ausfallen und es ist einfacher deren Eigenschaften zu kontrollieren.

Sei \vec{i} der Eingabevektor und \vec{o} der resultierende Ausgabevektor und sei die resultierende Abbildungsfunktion (Mapping) definiert als $M(\vec{i}) = \vec{o}$. Der Vektor \vec{k}_t beschreibe den Anwendungskontext zu einem Zeitpunkt t . Dann lässt sich die Rolle dieser Gestaltungsdimension beschreiben als eine Funktion,

$$K(\vec{i}, \vec{k}_t) = k_{t+1} = (k_j), \quad j = 1 \dots n \quad (4.3)$$

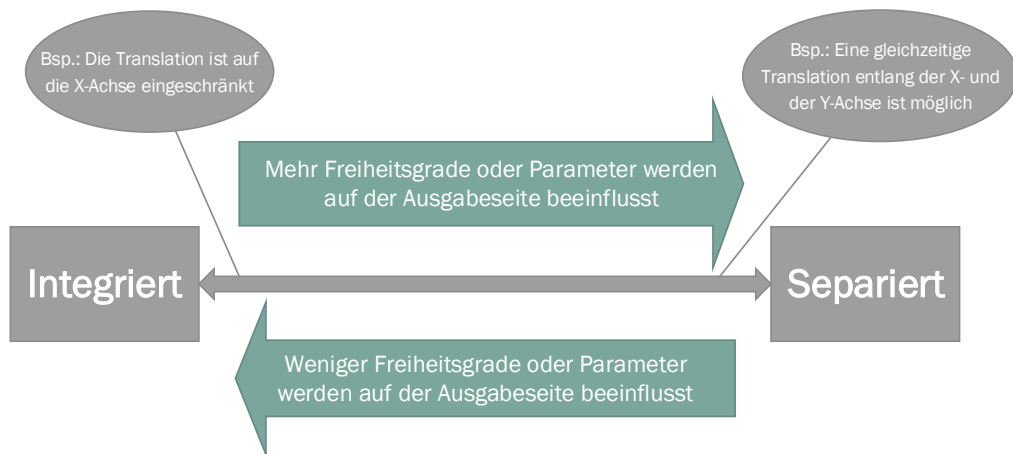


Abbildung 4.5: Schematische Darstellung zu der Gestaltungsdimension *integrierte vs. separierte Freiheitsgrade*

welche die Menge der aktiven Teilabbildungen M_{aktiv}

$$M_{\text{aktiv}} = \{m_k \mid m_k \in M_{\text{System}} \wedge \exists j = 1 \dots n : k = k_j\} \quad (4.4)$$

aus der Menge aller im System vorhandenen Abbildungen M_{System} bestimmt.

Hierbei sei der Einfachheit halber angenommen, dass alle m_k paarweise linear unabhängig seien. Durch

$$M(\vec{i}) = \sum_{M_{\text{aktiv}}} m_k(\vec{i}) = \vec{o} \quad (4.5)$$

ergibt sich dann die Abbildungsfunktion.

4.3.3 Integrierte vs. separierte Freiheitsgrade

Die Gestaltungsdimension *integrierte vs. separierte Freiheitsgrade* wird im Kontext dieser Arbeit meist direkt auf räumliche Freiheitsgrade angewendet. Im Allgemeinen ist diese Dimension jedoch keineswegs nur auf räumliche Operationen beschränkt, auch wenn diese für den Anwendungsfall der Modellierung eine hohe Priorität einnehmen, sondern kann wesentlich weiter gefasst werden. Charakterisiert ist diese Gestaltungsdimension durch die auf der Ausgabeseite beeinflussten Freiheitsgrade der Abbildungen (Mappings) der Eingabedaten (vgl. Abbildung 4.5).

Die Extremposition *separierte Freiheitsgrade* ist dadurch gekennzeichnet, dass immer nur genau ein Freiheitsgrad beeinflusst wird. Die Extremposition *integrierte Freiheitsgrade* ist im Gegensatz dazu dadurch gekennzeichnet, dass alle zur Verfügung stehenden Parameter bzw. Freiheitsgrade gleichzeitig beeinflusst werden können. Diese Gestaltungsdimension lässt sich als Dimensionalität des Ausgabevektors verstehen.

Die integrierte Interaktion mit Freiheitsgraden ist dadurch gekennzeichnet, dass es zwangsläufig eine Beeinflussung zwischen den Freiheitsgraden gibt, diese sind also von einander abhängig. Soll gezielt nur einer der Parameter der Operation manipuliert werden, müssen alle Eingabeparameter in einer Weise verändert werden, so dass die unvermeidliche Änderung an den übrigen Ausgabeparametern neutralisiert wird. Der Grad der tatsächlichen Beeinflussung hängt damit von der Feinmotorik und dem visuellen und räumlichen Vorstellungsvermögen bzw. dem Verständnis und der Übung der Benutzer ab. Diese Gestaltungsdimension kann als ein Maß dafür aufgefasst werden, wie *leicht* oder *schwer* es dem Benutzer gemacht wird, die Freiheitsgrade zu integrieren bzw. zu separieren.

Im Folgenden soll dies an einem Beispiel aus dem Bereich der 3D-Modellierung verdeutlicht werden. Betrachtet man die Translation bzw. die Rotation von Objekten im dreidimensionalen Raum, so verfügt die Translation grundsätzlich über drei Freiheitsgrade entsprechend den drei Raumachsen und selbiges gilt für die Rotation. Eine separierte Manipulation würde bedeuten, dass die Eingabedaten – also die Fingerberührungen – lediglich wenige dieser Freiheitsgrade gleichzeitig manipulieren. Im Extremfall könnte dies auf einen Freiheitsgrad, etwa die Translation entlang der X-Achse beschränkt sein. Bzgl. einer integrierten Manipulation wären die gleichzeitige Translation entlang der X- und der Y-Achse oder die Kombination mit einer Rotation um eine der Raumachsen denkbar. Die Kombination aus Translation und Rotation wäre auch ein Beispiel für eine Integration von Freiheitsgraden zwischen verschiedenen Operationen, während die Kombination mehrerer Freiheitsgrade der Translation sich lediglich auf die Integration von Freiheitsgraden innerhalb derselben Operation bezieht.

Um den Charakter der oben angesprochenen Beeinflussung der Parameter genauer zu erläutern, sei angenommen, in einem Interface bestünde die

Möglichkeit, mit zwei Fingern gleichzeitig ein Objekt in der X-Y-Ebene zu verschieben und um die Z-Achse zu rotieren. Die Translation werde bestimmt durch die relative Bewegung des geometrischen Schwerpunktes der beiden Finger und die Rotation durch die Summe der Winkel, ermittelt durch die Differenz der Geraden durch beide Finger und den geometrischen Schwerpunkt mit der X-Achse. Hier wäre eine völlig integrierte Kontrolle aller angesprochenen Freiheitsgrade mit einer Bewegung durch den Benutzer möglich. Andererseits kann der Benutzer, indem er die relative Positionierung beider Finger zueinander möglichst konstant hält und beide zusammen bewegt, auch nur eine Translation ausführen oder sogar – wenn er die Bewegung exakt horizontal oder vertikal ausführt – die Translation auf eine Achse beschränken. Allerdings ist diese Trennung durch die Präzision der Eingabedaten und vor allem durch die eigene Feinmotorik bzw. motorische und kognitive Leistungs- und Konzentrationsfähigkeit des Benutzers beschränkt. Durch Einsatz von Filterung oder „toten Zonen“ könnte das System dem Benutzer eine Trennung erleichtern, würde aber gleichzeitig das integrierte Ansprechverhalten verschlechtern, d.h. beide Positionen markieren tatsächlich gegensätzliche Endpunkte auf einem kontinuierlichen Spektrum.

Integrierte Manipulation ist in diesem Zusammenhang nicht grundsätzlich auf räumliche Dimensionen beschränkt, sondern kann auf das Zusammenspiel nahezu beliebiger Parameter angewendet werden. Beispiele wären Raum und Zeit oder andere grafische Objekteigenschaften wie Farbe oder Transparenz. Bzgl. Raum und Zeit stellt beispielsweise die Animation ein interessantes Anwendungsfeld dar.

Grundsätzlich stellt sich die Frage, wann es sinnvoll ist, die Manipulation von verschiedenen Freiheitsgraden zu integrieren bzw. zu separieren. Dies kann nach Durchsicht der Literatur im Allgemeinen nicht eindeutig beantwortet werden (vgl. Hancock u. a. (2007); Moscovich und Hughes (2008); Nacenta u. a. (2009)). Nach Jacob u. a. (1994) sollte die *Perceptual Structure* beachtet werden, d.h. wichtig ist die Wahrnehmung des Benutzers. Nimmt der Benutzer verschiedene elementare Operationen als einen *Task* wahr oder jeweils als getrennte Aufgabe, so sollte diese Struktur im Design des Interfaces beachtet werden. Eine ähnliche Herangehensweise findet sich auch bei Bowman u. a. (2004). Parameter, die vom Benutzer als zu einer Aufgabe gehörend wahr-

genommen werden, sollten meist auch zusammen manipulierbar sein. Hinckley und Wigdor (2012) sprechen in diesem Zusammenhang von verschiedenen Möglichkeiten des *Chunking*. So kann eine Manipulationstechnik eher die *Low-level*-Struktur bedienen, d.h. elementare Operationen werden getrennt durchgeführt oder diese kann auf der *High-level*-Ebene arbeiten, d.h. elementare Operationen werden integriert durchgeführt.

Im Fall dieser Designdimension ist eine Zuordnung zu der Frage Emulation vs. Paradigmenwechsel nicht eindeutig möglich, da auch bei etablierten Eingabegeräten und Softwaresystemen unterschiedliche Strategien verfolgt und abgebildet werden. Allerdings ist das Potenzial für eine Integration vieler Freiheitsgrade bei MT größer, da auf der Eingabeseite mehr Freiheitsgrade vorhanden sind. Es existieren Hinweise, dass MT das *Phrasing* (Hinckley und Wigdor, 2012) von elementaren Operationen durch die Nutzung der propriozeptiven Wahrnehmung begünstigt.

Sei \vec{o} der resultierende Ausgabevektor der Abbildungen (siehe oben) und d die Anzahl der räumlichen und sonstigen Freiheitsgrade einer Operation bzw. eines Objektes, so ist die Extremposition integrierte Freiheitsgrade durch

$$\dim \vec{o} = d \quad (4.6)$$

und die Extremposition der separierten Freiheitsgrade durch

$$\dim \vec{o} = 1 \quad (4.7)$$

beschrieben.

4.3.4 Direkte vs. indirekte Interaktion

Die Begriffe direkte bzw. indirekte Interaktion sowie absolute und relative Interaktion, Gesten oder Eingaben werden in der Literatur häufig unpräzise, teils synonym und teils widersprüchlich verwendet und sollen daher zunächst klarer gefasst werden. Diese Arbeit folgt dabei der Anschauung von Zhai und Milgram (1998).

Direkte Interaktion bedeutet eine starke Korrespondenz von Eingabedaten und manipulierten Freiheitsgraden bzw. Objekten. Bei der indirekten Interaktion ist dagegen der Abstraktionsgrad zwischen Eingabedaten und den manipulierten Freiheitsgraden bzw. Objekten größer (vgl. Abbildung 4.6). Wie

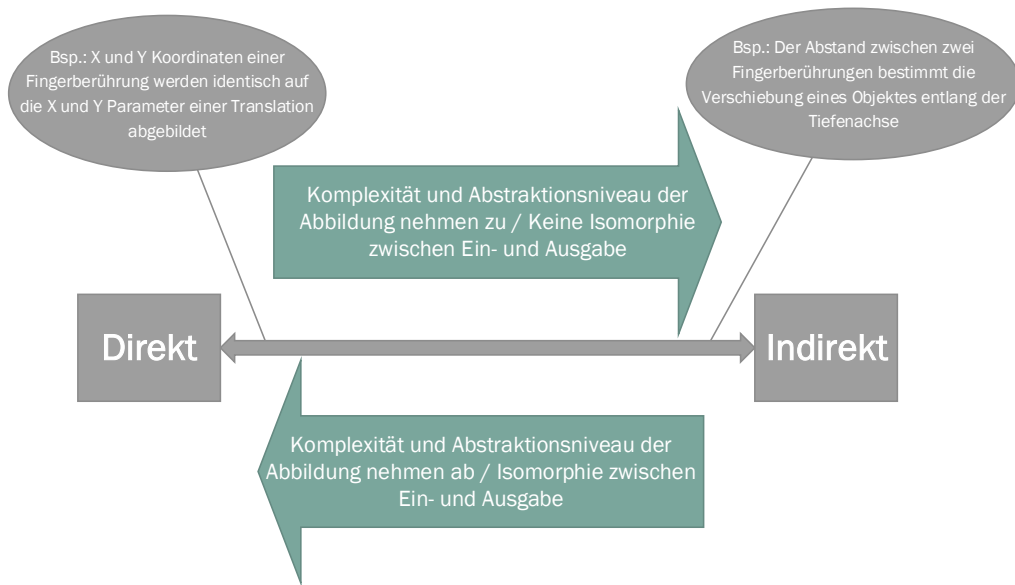


Abbildung 4.6: Schematische Darstellung zu der Gestaltungsdimension *direkte vs. indirekte Interaktion*

bei den anderen vorgestellten Designdimensionen liegen direkte und indirekte Interaktion auf einem Kontinuum (vgl. Zhai und Milgram (1998)).

Die Begriffe absolute und relative Interaktion beschreiben Eigenschaften der verwendeten Eingabegeräte. Ein absolutes Eingabegerät ist in der Lage, Positionen zu erfassen und auf diese Weise jede Eingabe in einem fixen, absoluten Koordinatensystem zu verorten (Hinckley u. a., 1994). Dagegen erfassen relative Eingabegeräte Veränderungen in der Position bzw. Bewegungen. Absolute Eingabegeräte sind in der Lage, relative Veränderungen aus den gemessenen Positionen abzuleiten und können daher sowohl in einem absoluten als auch relativen Modus betrieben werden. Relative Eingabegeräte sind dagegen auf den relativen Modus beschränkt. MT-Displays oder auch Touchpads erfassen Positionen und können daher sowohl absolute als auch relative Daten liefern, während die Maus lediglich relative Daten liefern kann. Nach Zhai und Milgram (1998) sind absolute Eingaben tendenziell als direkt und relative Eingaben tendenziell als indirekt zu klassifizieren.

Der Extremwert *direkte Interaktion* bedeutet eine isomorphe Abbildung zwischen Eingabe und Ausgabe. Nach Zhai und Milgram wäre das „direkteste“ denkbare Interface vom Benutzer gar nicht mehr wahrnehmbar. Dieser hätte

einfach das Gefühl unmittelbar das Zielobjekt zu manipulieren. Der Extremwert *indirekte Interaktion* ist dagegen durch eine Abstraktion, welche sich auch in der räumlichen Entfernung äußern kann, zwischen Eingabe und Ausgabe charakterisiert. Dies bedeutet, dass die Interaktion nach Zhai und Milgram eher dem Einsatz von Werkzeugen ähnelt oder einen symbolischen Charakter annimmt.

Zum Verständnis hilfreich ist die Einordnung von Sturman (1992) bzw. Moscovich (2007). Sturman unterteilt zunächst die möglichen Aktionen der Hand in die Kategorien „continuous features“ und „discrete features“. Kontinuierliche Eigenschaften oder Daten sind alle kontinuierlichen Quantitäten, die von der Hand oder den Fingern abgeleitet werden können, insbesondere Positionen, Bewegungen, Beschleunigungen, Winkel und Kräfte. Diskrete Eigenschaften sind hauptsächlich Posen oder bestimmte Handhaltungen, in denen ganze Bereiche der kontinuierlichen Quantitäten zusammengefasst werden. Auf der anderen Seite lassen sich nach Sturman die Interpretationen des Systems bzw. die Abbildungen der Eingaben auf Aktionen in die Kategorien „direct“, „mapped“ und „symbolic“ einordnen (vgl. Sturman (1992)). *Direct* entspricht der isomorphen Abbildung der Eingabedaten, ähnlich wie bei Zhai und Milgram. *Mapped* bedeutet die Abstraktion bzw. Abbildung auf virtuelle Werkzeuge oder auch Schnittstellenelemente wie Slider oder Buttons (vgl. Tools bei Zhai und Milgram (1998)). *Symbolic* kann als die Abbildung auf Elemente einer (formalen) Sprache interpretiert werden, also auf *Tokens* bzw. Symbole mit einer bestimmten Semantik.

Direkte Interaktion auf MT-Bildschirmen bedeutet, dass sich der Finger genau auf dem zu manipulierenden Objekt befindet und dass Wege und Distanzen von Bewegungen direkt abgebildet werden. Der „control space“ und der „display space“ korrespondieren (Zhai und Milgram, 1998) und Fingerberührungen werden in diesem absolut verortet. Im Gegensatz dazu ist die indirekte Interaktion durch ein variableres Verhältnis der Distanz zwischen Fingerberührungen und manipuliertem Objekt gekennzeichnet. Erstens ist ein deutlich wahrnehmbarer Abstand vorhanden und zweitens muss dieser Abstand nicht konstant sein, sondern kann kontext-abhängig variieren. Beispielsweise kann in bestimmten Anwendungen ein nicht-linearer Beschleunigungsfaktor zwischen Eingabe und Ausgabe hilfreich sein. Ähnliche Techniken stehen unter den gän-

gigen Betriebssystemen für die Maus zur Verfügung (vgl. Bowman u. a. (2004)).

Auch das Abstraktionsniveau unterscheidet sich insgesamt, so werden bei der direkten Interaktion im Wesentlichen gleichförmige Daten aufeinander abgebildet, d.h. Positionen auf Positionen bzw. Bewegungen auf Bewegungen. Dagegen wird bei der indirekten Interaktion stärker abstrahiert. Die Bewegungen der Hand bzw. der Finger können auf symbolische Gesten abgebildet werden, welche dann bestimmte Aktionen im System auslösen (Sturman, 1992) oder generell komplexeren mathematischen Berechnungen unterliegen (vgl. Hinckley u. a. (1994)).

Bei Zhai und Milgram (1998) findet sich eine Taxonomie, die das *Isomorphismus-Werkzeug-* bzw. *Direct-Indirect-Kontinuum* darstellt. MT-Bildschirme werden hier der direkten Interaktion zugeordnet. Touchpads sind etwas „indirekter“, da immer ein Offset zwischen „control space“ und „display space“ vorliegt.

Zur Bewertung der beiden Extrempositionen ist zu sagen, dass zunächst die direkte Interaktion als einer der wesentlichen Vorteile von interaktiven Oberflächen angesehen wird (Hinckley und Wigdor, 2012; Bowman u. a., 2004; Wigdor und Wixon, 2011). Insbesondere für 2D-Anwendungen, etwa Bild- oder Videobrowser, liegt hier eine isomorphe Abbildung vor und damit ein für den Benutzer schon auf Basis seiner „Alltagsintuition“ einfach verständliches mentales Modell (Sturman, 1992; Wigdor und Wixon, 2011). Tatsächlich hängt dieser Effekt natürlich vom Gesamtdesign des Interfaces und insbesondere von den anderen oben aufgeführten Gestaltungsdimensionen ab. Außerdem können psychologische Effekte auftreten, welche die „natürliche“ Interaktion behindern. Einige Benutzer erwarten von einem Computerinterface scheinbar kein „natürliches“ Verhalten im Sinne der Anlehnung an die physikalische Welt (North u. a., 2009; Terrenghi u. a., 2007; Epps u. a., 2006). Ein weiterer Vorteil der direkten, isomorphen Abbildung ist die Einbeziehung des kinesthetischen Sinns (Bowman u. a., 2004; Sturman, 1992; Hinckley und Wigdor, 2012; Balakrishnan und Hinckley, 1999).

Ein Nachteil der direkten Abbildung ist die Begrenzung der Interaktion durch körperliche Einschränkungen wie Armlänge, Schnelligkeit oder auch Feinmotorik des Benutzers sowie die Einschränkung des Interaktionsdesigns. Auch in der realen Welt setzen Menschen aus gutem Grund für viele Aktivitäten

Werkzeuge ein (Sturman, 1992; Hinckley u. a., 1994; Bowman u. a., 2004). Bowman u. a. (2004) weisen darauf hin, dass die Mehrzahl der Interaktionstechniken zu den nicht-isomorphen, also indirekten Techniken gehört.

Bezogen auf die Diskussion Emulation vs. Paradigmenwechsel ist festzustellen, dass viele etablierte Eingabegeräte wie die Maus oder auch passive Grafiktablets technisch bedingt lediglich eine indirekte Interaktion ermöglichen. Erst MT-Bildschirme bzw. aktive Grafiktablets mit Display erlauben eine vollständig direkte Interaktion wie oben beschrieben (vgl. Kapitel 3).

Mathematisch kann diese Gestaltungsdimension durch die Komplexität der Abbildung ausgedrückt werden. Dies wird der Übersicht halber stellvertretend am Beispiel der zweidimensionalen Translation eines Objektes beschrieben. Sei $\vec{o} = (o_1 \dots o_n)$ der resultierende Ausgabevektor und $\vec{p} = (p_1 \ p_2) \in \mathfrak{R}^2$ die Position des Objektes.

Der Extremwert der direkten Interaktion lässt sich in diesem Fall z.B. als

$$p_1 = o_1 \quad p_2 = o_2 \quad (4.8)$$

beschreiben.

Die einfachste (sinnvolle) Abbildung ist also die Identität. Eine komplexere Abbildung könnte eine konstante Verschiebung

$$p_1 = o_1 + 100 \quad p_2 = o_2 + 100 \quad (4.9)$$

oder lineare Faktoren

$$p_1 = 0,5 * o_1 + 100 \quad p_2 = 0,5 * o_2 + 100 \quad (4.10)$$

einschließen.

Der Extremwert der indirekten Interaktion ist dabei theoretisch nur begrenzt durch die Anforderung der Berechenbarkeit. In der praktischen Umsetzung spielen auch andere Anforderungen wie Rechenzeit und Speicheranforderungen eine wichtige Rolle.

4.3.5 Mehrhändige Interaktionsmuster

In der realen Welt setzen Menschen für die Bewältigung vieler Aufgaben grundsätzlich beide Hände ein. Dies ist vielfach auch bei Tätigkeiten der Fall, die auf

den ersten Blick wie einhändige Tätigkeiten wirken, z.B. das Schreiben (Guiard, 1987). Die Einbeziehung beider Hände in die Interaktion und die Beachtung der speziellen Rollen von starker und schwacher Hand spielen deshalb auch für die Gestaltung von Benutzungsschnittstellen eine wichtige Rolle (Buxton und Myers, 1986).

Insbesondere empirische Studien von Guiard (1987) werden häufig als Grundlage für die Gestaltung von zwei- bzw. beidhändigen Interfaces herangezogen (Bowman u. a., 2004; Balakrishnan und Kurtenbach, 1999; Goble u. a., 1995; Hinckley u. a., 1994; Cutler u. a., 1997). Guiard belegt, dass Menschen für verschiedene Manipulationsaufgaben in der realen Welt die starke und schwache Hand in spezifischer Weise koordinieren.

Guiard unterscheidet zwischen einhändiger, beidhändig symmetrischer und beidhändig asymmetrischer Interaktion (Guiard, 1987). Außerdem kann die beidhändige Interaktion entweder synchron oder asynchron erfolgen. Besonderes Augenmerk legt Guiard in seinen Untersuchungen auf die beidhändig asymmetrische Interaktion und stellt hier drei grundlegende Prinzipien fest, welche er als „kinematic chain“ bezeichnet (Guiard, 1987):

1. Die schwache Hand passt den räumlichen Bezugsrahmen für die Interaktionen der starken Hand dynamisch an.
2. Die starke Hand führt motorisch feine, präzise Bewegungen aus, die schwache Hand dagegen grobe Manipulationen.
3. Im zeitlichen Ablauf initiiert die schwache Hand die Interaktion vor der starken Hand.

Als typisches Beispiel für die Effektivität von asymmetrisch beidhändiger Interaktion führt Guiard das Schreiben an. Beidhändiges Schreiben (die schwache Hand führt das Papier) ist in seinen Untersuchungen im Vergleich zu einhändigem Schreiben um ca. 20% schneller (Guiard, 1987).

Nach Leganchuk u. a. (1998) besteht der Vorteil beidhändiger Interaktion nicht unbedingt nur in einem Geschwindigkeitsvorteil durch den gleichzeitigen Einsatz beider Hände, sondern es können darüber hinaus auch qualitative kognitive Vorteile resultieren. Die genaue Natur dieser Vorteile können Leganchuk u. a. (1998) aus ihren Experimenten zwar nicht bestimmen, sie vermuten aber,

dass die Interaktion mit beiden Händen dem Benutzer ein besseres mentales Modell vermittelt.

Die Einteilung erklärt und beschreibt zwar die wesentlichen Grundlagen der beidhändigen Interaktion. Für das Interfacedesign ist aber eine weitere Differenzierung hilfreich, insbesondere um technische Anforderungen an das System zu verdeutlichen. Im Gegensatz zu den anderen Designdimensionen stellt diese Dimension kein Kontinuum dar, sondern lässt sich in vier Kategorien (vgl. Abbildung 4.7) gliedern:

- einhändige Interaktion
- mehrhändige Interaktion (synchron/asynchron)
- symmetrisch beidhändige Interaktion (synchron/asynchron)
- asymmetrisch beidhändige Interaktion (synchron/asynchron)

Um zu diskutieren, was die einzelnen Kategorien bedeuten, ist es wichtig, zunächst genauer zu definieren, auf welches *Interfacelement* sich der Begriff *Interaktion* hier genau bezieht. Hierzu wird eine aufgabenbasierte Sicht eingenommen, d.h. Interaktion bezieht sich auf Aktionen, die der Benutzer zur Ausführung bzw. Lösung einer bestimmten Aufgabe ausführen muss.

Aufgaben können dabei unterschiedliche Granularität besitzen, z.B. eine andere Perspektive mit der Kamera einstellen oder auch die Durchführung einer komplexen Modellierungsoperation, die verschiedene Kameramanipulationen und Modellierungsoperationen einschließt. Auf jeder dieser Granularitätsstufen kann ein Interface in verschiedene der vier genannten Kategorien dieser Gestaltungsdimension fallen (vgl. Hinckley und Wigdor (2012); Bowman u. a. (2004)).

So kann eine Kameramanipulation wie das Zoomen beidhändig durchgeführt werden (symmetrisch), z.B. zur Erhöhung der Präzision, oder aber diese Operation für sich genommen ist nur einhändig durchführbar, dafür aber im Zusammenspiel mit anderen Modellierungsoperationen beidhändig auszuführen (asymmetrisch). Die Kameramanipulation kann zeitlich jeweils im Wechsel mit einer Modellierungsoperation stattfinden (asynchron) oder während der Benutzer mit der einen Hand die Kamera bewegt, manipuliert er gleichzeitig das Objekt (synchron). Eine typische symmetrisch synchrone Geste aus dem

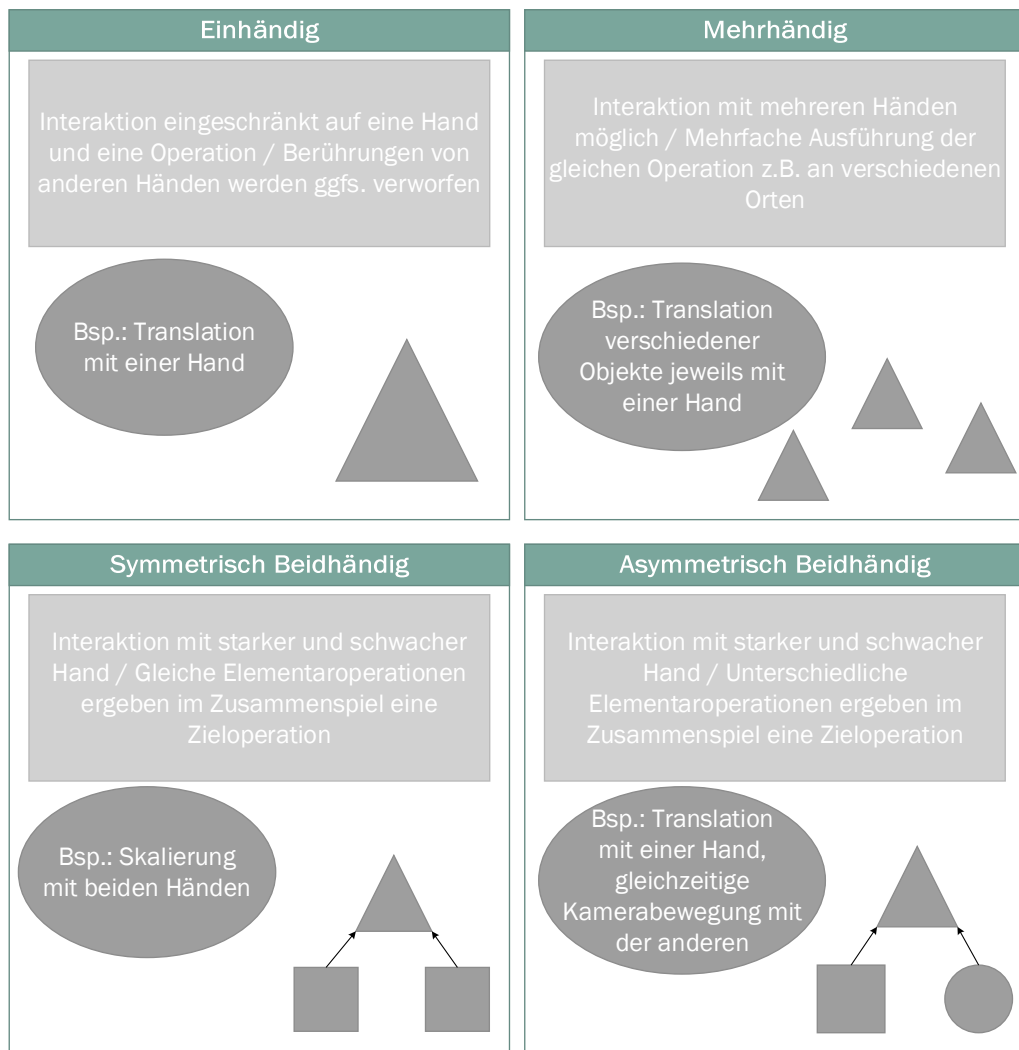


Abbildung 4.7: Schematische Darstellung zu der Gestaltungsdimension *mehrhändige Interaktionsmuster*

MT-Umfeld ist die Pinch-Geste. Ein symmetrisch asynchrones Beispiel wäre das Panning mit Hilfe beider Hände. Hier könnte der Benutzer, um das Panning schneller auszuführen, jeweils abwechselnd mit beiden Händen ein Objekt verschieben, ähnlich als würde er eine virtuelle Leiter hinaufsteigen. Beidhändiges Schreiben auf einem Touch-Keyboard (oder einer normalen Tastatur) würde ebenfalls in diese Kategorie fallen.

Mit einhändiger Interaktion entsprechend der ersten oben genannten Kategorie ist gemeint, dass das System tatsächlich während dieser Operation nur einhändig bedient werden kann. Dies impliziert auch, dass nur eine elementare Operation zeitlich durchgeführt werden kann. Alle weiteren Fingerberührungen werden ggf. ignoriert. Dies stellt eine relativ starke Einschränkung dar, kann aber unter Umständen aus Stabilitätsgründen oder aus anderen externen Gründen notwendig sein. Neben technischen Beschränkungen des Eingabegeräts, könnte es in bestimmten Arbeitsumfeldern erwünscht oder notwendig sein, dass ein Benutzer stets nur eine Hand für das MT-Interface verwendet.

Besteht in einem System die Möglichkeit, für eine Operation die Finger beider Hände einzusetzen oder verschiedene einhändige Operationen gleichzeitig durchzuführen, so fällt dies zunächst in die Kategorie *mehrhändige Interaktion*. Diese ist gegenüber *beidhändiger Interaktion* zu unterscheiden, da hier vom Designer bzw. dem Interface keine besondere Rollenzuweisung der Hände oder eine explizite Unterstützung der beidhändigen Interaktion geleistet wird. Gemäß der oben angesprochenen aufgabenorientierten Sicht, werden hier unter Umständen völlig unabhängige Operationen gleichzeitig durchgeführt, die nicht der Erreichung eines gemeinsamen Zieles bzw. der Erfüllung einer gemeinsamen Aufgabe dienen. Es besteht keine gewollte Zuordnung (entweder technisch erzwungen oder zumindest durch das Interfacedesign suggeriert) zwischen Operationen und starker bzw. schwacher Hand. Die Interaktion von mehreren Benutzern, die an verschiedenen Stellen der Oberfläche die gleiche Operation ausführen, fällt ebenfalls in diese Kategorie.

Von beidhändiger Interaktion kann im Gegensatz zu einer mehrhändigen Interaktion nur dann gesprochen werden, wenn beide Hände zusammen an einer Elementaroperation beteiligt sind (symmetrisch) oder an verschiedenen Elementaroperationen, welche sich aber eindeutig in der Ausführung einer übergeordneten Operation unmittelbar ergänzen (asymmetrisch). Voraussetzung

ist die Intention des Designers, da ein effizientes Zusammenspiel der Hände sich im Allgemeinen Fall nicht automatisch ergibt, sondern durch das Interface explizit unterstützt oder zumindest suggeriert werden sollte. Bei einem Mehrbenutzersystem würde eine beidhändige Interaktion bedeuten, dass jeder Benutzer jeweils seine starke und schwache Hand im Zusammenspiel einsetzen kann.

Die Ergebnisse von Guiard sind wie bereits erwähnt Grundlage für das Design der beidhändigen Interaktion in einer Reihe von Systemen und konnten dort vielfach erfolgreich eingesetzt werden (Leganchuk u. a., 1998; Hinckley u. a., 1998b; Balakrishnan und Hinckley, 1999; Balakrishnan und Kurtenbach, 1999; Cutler u. a., 1997). Einige empirische Versuche deuten jedoch darauf hin, dass die Ergebnisse von Guiard nicht in jedem Fall auf interaktive Oberflächen übertragen werden können (Terrenghi u. a., 2007; North u. a., 2009; Epps u. a., 2006; Walther-Franks u. a., 2011b; Herrlich u. a., 2011a). So scheint unter anderem die große Vertrautheit von vielen Benutzern mit etablierten Eingabegeräten bzw. Interfaceparadigmen, welche nicht auf eine beidhändige Interaktion ausgelegt sind, dazu zu führen, dass viele Benutzer dies von *Computern* einfach nicht erwarten und daher im Rahmen von Versuchen mit meist relativ kurzer Dauer kaum beidhändiges Interaktionsverhalten zeigen. Andererseits belegt die Zunahme von beidhändiger Interaktion gegen Ende einiger Versuche (Balakrishnan und Kurtenbach, 1999; Walther-Franks u. a., 2011b) und die beobachtete stärkere Ausprägung von beidhändiger Interaktion bei einigen Experten (Buxton und Myers, 1986), dass die wenig zu beobachtende beidhändige Interaktion kein prinzipielles Problem der verwendeten Eingabegeräte darstellt. Beidhändige Interaktionstechniken sollten demnach eher für Systeme eingesetzt werden, bei denen eine längere Lernphase von den Nutzern akzeptiert wird.

Neben den in den folgenden Kapiteln 5 und 6 präsentierten Ansätzen finden sich in der Literatur einige Beispiele für beidhändige Interaktionstechniken mit Bezug zur 3D-Modellierung. Ein Beispiel, allerdings aus dem Bereich der 3D-Interfaces, stellt die Responsive Workbench dar (Cutler u. a., 1997; Frohlich u. a., 2000), welche die beidhändige Manipulation von Objekten mittels zweier Datenhandschuhe erlaubt und explizit einhändige, symmetrisch beidhändige und asymmetrisch beidhändige Operationen entsprechend der oben

vorgestellten Klassifikation unterstützt. Balakrishnan und Kurtenbach (1999) untersuchen die beidhändige Kamerakontrolle bzw. Objektmanipulation mittels zweier Mäuse und Buxton und Myers (1986) erforschen die beidhändige Ausführung von elementaren Operationen wie Selektion und Navigation mittels speziell konstruierter Hardware.

Im Zusammenspiel mit den oben diskutierten Gestaltungsdimensionen wirkt eine mehrhändige bzw. beidhändige Interaktion wie ein Filter bzw. eine weitere Meta-Abbildung, welche die Zuordnung von spezifischen Teilmengen der Eingabedaten zu bestimmten Operationen bzw. Abbildungen im Sinne der vorherigen Gestaltungsdimensionen ermöglicht. Diese partitioniert also den Eingabevektor \vec{i} und beeinflusst den Kontext der Anwendung bzw. die Auswahl der Mappings.

4.4 Feedback

Wie bei anderen Eingabeschnittstellen auch, spielt Feedback für den Benutzer bei interaktiven Oberflächen eine wesentliche Rolle (Wigdor und Wixon, 2011). Nach Bowman u. a. (2004) sind grundsätzlich drei Arten von Feedback zu unterscheiden: das reaktive Feedback, das instrumentelle Feedback und das operative Feedback.

Da Feedback im Einklang mit den eingesetzten Interaktionstechniken stehen muss, erscheint eine herausgelöste Diskussion des Feedbacks nicht sinnvoll. Feedback wird daher nicht als eigene Dimension betrachtet, sondern als zugehöriges Gestaltungsmerkmal der Interaktionstechniken gemäß der bereits aufgeführten Designdimensionen. Der Designprozess sollte sich am Interaktionsziel orientieren und dieses im Kontext der Anforderungen und der Hardwarefähigkeiten zunächst entsprechend der oben vorgestellten Designdimensionen abbilden. Feedback sollte jeweils verstärkend bzw. unterstützend in die Gestaltung der Techniken einbezogen werden.

Für interaktive Oberflächen und speziell MT ist die Visualisierung der Fingerberührungen besonders wichtig, da dem Benutzer kaum anderweitige Kontrollmöglichkeiten über den Zustand des Systems zur Verfügung stehen (Hinkley und Wigdor, 2012).

Indirekte Interaktion kann durch ein visuelles Feedback unterstützt werden,

welches die Verbindung zwischen Berührungen und dem aktuellen Zielobjekt betont bzw. für den Benutzer verständlicher macht (Wigdor u. a., 2009). Bei der direkten Interaktion wirkt darüber hinaus das propriozeptive Feedback durch die Korrespondenz von Eingabe- und Ausgaberaum (Hinckley und Wigdor, 2012; Balakrishnan und Hinckley, 1999). Bei der direkten Interaktion ist eine hohe Aktualisierungsrate des visuellen Feedbacks bzw. der Auswirkungen der Interaktion wichtig, da eine fühlbare Latenz zwischen Eingaben durch den Nutzer und Darstellung der Auswirkungen die diskutierten Vorteile der direkten Interaktion beeinträchtigen kann. Ein Moduswechsel erfordert in der Regel auch eine entsprechende Repräsentation, damit der Benutzer weiß, in welchem Modus er sich gerade befindet. In gewissen Grenzen ist eine gegenseitige Substitution verschiedener Feedback-Kanäle bzw. -Modalitäten möglich (Bowman u. a., 2004; Balakrishnan und Hinckley, 1999).

Kapitel 5

Objektmanipulation in 3D-Szenen

Die Manipulation von 3D-Objekten ist eine Kernaufgabe jeder 3D-Modellierungsumgebung. Damit ein Objekt überhaupt manipuliert werden kann, muss dieses zuerst selektiert werden. Die Selektion steht also am Anfang jeder Manipulationsoperation und soll daher auch als erstes betrachtet werden.

Dieses Kapitel gliedert sich in zwei wesentliche Teile. Der erste Teil diskutiert eine Untersuchung zu proxybasierter Selektion auf Multi-touch (MT)-Bildschirmen. Die dargestellte Untersuchung wurde unabhängig von dem 3D-Modellierungskontext durchgeführt. Wie eingangs bereits erwähnt, stellt die Selektion allerdings eine Grundvoraussetzung für die Modellierung dar und die Ergebnisse sind somit auch für diesen speziellen Anwendungsfall relevant und übertragbar. Der zweite Teil dieses Kapitels beschäftigt sich besonders mit beidhändigen Mehrfingertechniken zur Objektmanipulation.

5.1 Selektion

Die korrekte, effiziente und präzise Selektion von Objekten ist eine grundsätzliche Anforderung von graphischen 2D- und 3D-Anwendungen. Insbesondere in Anwendung, welche mit vielen verdeckten oder sich bewegenden Objekten umgehen müssen, kann die Selektion dieser Objekte eine schwierige Aufgabe sein. Der Bereich der 2D- und 3D-Modellierung und -Animation ist hierfür ein gutes Beispiel. In typischen Szenen ist eine Vielzahl verschiedener Objekte vorhanden, welche sich insbesondere im Fall der häufig eingesetzten perspektivischen Projektion gegenseitig verdecken.

In diesem Abschnitt wird eine Selektionstechnik auf Basis von Proxyobjekten zur schnellen und präzisen Selektion von verdeckten und sich bewegenden Objekten vorgestellt. Die präsentierte Technik skaliert prinzipiell über verschiedene Eingabegeräte und ist sowohl für die Maus als auch für Single- und Multi-touch geeignet. Im Folgenden wird der Gestaltungsraum der Proxyobjekte und deren Eigenschaften beschrieben. Eine Teilmenge dieser Eigenschaften – *color similarity* und *motion similarity* – wurde im Rahmen einer Nutzerstudie näher untersucht, deren Ergebnisse ebenfalls präsentiert und diskutiert werden.

5.1.1 Stand der Technik

Es existieren eine Reihe von verwandten Arbeiten im Bereich der Selektionstechniken für verdeckte Objekte. Die TapTap-Technik (Roudaut u. a., 2008) und ähnliche *Magic-Lense*-Techniken (Bier u. a., 1993) verwenden (lokale) Vergrößerung, um Verdeckungen aufzulösen bzw. zu reduzieren. Die Handle-Flags-Technik (Grossman u. a., 2009) basiert auf einer verkleinerten Darstellung einzelner Teile eines Schriftzugs, um diese innerhalb des gesamten Schriftzugs auszuwählen bzw. Operationen auf diesen auszuführen. Die Splatter-Technik (Ramos u. a., 2006) separiert verdeckte Objekte temporär voneinander, um mit diesen zu interagieren. Andere Techniken schränken die Freiheitsgrade der Interaktion gezielt ein, um die Selektion der Objekte zu vereinfachen (Bier, 1990).

Die Interaktion mit sich bewegenden Objekten wurde bisher weniger intensiv untersucht. Bei der Target-Lock-Technik (Gunn u. a., 2009) bleibt jeweils das letzte Objekt selektiert, welches der Cursor berührt hat. Schröder-Kroll u. a. (2008) präsentieren verschiedene Techniken für die Selektion sich bewegnender Objekte, von denen die Snapping-Pointer-Technik der im Rahmen dieser Untersuchung verwendeten Referenztechnik am ähnlichsten ist. Gleiches gilt für die Bubble-Cursor-Technik (Grossman und Balakrishnan, 2005).

Verschiedene Gestaltgesetze (Goldstein, 2009, S. 104ff) sind relevant für die Gestaltung der Proxyobjekte, beispielsweise *Verbundenheit* (connectedness), *Übereinstimmung* (similarity) und *gleiches Schicksal* (common fate).

5.1.2 Proxybasierte Selektion

Die grundlegende Idee aller proxybasierten Selektionstechniken ist es, ein Stellvertreterobjekt einzuführen, welches mit dem zu selektierenden Objekt verbunden ist. Indem der Benutzer mit dem Stellvertreterobjekt interagiert, d.h. dieses selektiert, interagiert er mit dem Zielobjekt. Die Komplexität der Gesten ist dabei nicht grundsätzlich beschränkt. Ein Vorteil der Proxyobjekte ist, dass der Interfacedesigner diese im Gegensatz zu den Zielobjekten frei gestalten kann und diese somit für bestimmte Aufgaben bzw. Interaktionstechniken optimieren kann. Insbesondere im Bereich der 3D-Modellierung können die Zielobjekte nicht einfach angepasst werden, da deren Gestaltung ja gerade im Fokus der Anwendung steht. Ein potenzieller Nachteil von Proxyobjekten ist die Einführung einer zusätzlichen Abstraktionsebene (vgl. Kapitel 4). Daher ist es wichtig, dass die Repräsentation der Stellvertreterobjekte es dem Benutzer ermöglicht, die Verbindung zwischen dem Stellvertreter- und dem Zielobjekt schnell und korrekt zu erkennen.

Im Fokus der Entwicklung und Untersuchung steht der Einsatz auf MT-Displays. Dennoch eignet sich die Technik grundsätzlich auch für die Benutzung mit der Maus. Bisher wurden die vorgestellten Techniken lediglich anhand einer 2D-Anwendung untersucht (siehe nächster Abschnitt). Eine Generalisierung auf den 3D-Fall ist aber leicht möglich, in dem die Proxyobjekte immer auf der *near plane*, d.h. über allen anderen Objekten, angeordnet werden.

Ob proxybasierte Selektionstechniken nachweisbar einen Vorteil gegenüber Alternativtechniken bieten und wie sich verschiedene Gestaltungsmöglichkeiten der Proxyobjekte auswirken, wird mit Hilfe einer im nächsten Abschnitt diskutierten Nutzerstudie untersucht. Folgende Thesen sollen im Rahmen des Experiments überprüft werden:

1. Proxybasierte Techniken ermöglichen eine schnellere und präzisere Selektion mit einer geringeren Fehlerrate.
2. Proxybasierte Techniken sind besonders hilfreich für die Selektion von verdeckten und sich bewegenden Objekten.
3. Proxybasierte Techniken können Bewegung als markantes Merkmal für die Selektion dynamischer Objekte einsetzen.

Ansatz und Implementierung

Proxyobjekte erscheinen auf dem Bildschirm für Objekte innerhalb eines festgelegten Radius um eine Fingerberührung (bzw. den Mauscursor). Proxyobjekte sind nur sichtbar, wenn der Benutzer den Bildschirm berührt, um den sichtbaren Bereich nicht zusätzlich zu reduzieren. Der Radius hängt prinzipiell von der Bildschirm- und Eingabeauflösung ab und muss empirisch im Vorfeld für den verwendeten Bildschirm ermittelt werden. Proxyobjekte bleiben sichtbar, solange der Finger den Bildschirm berührt (oder bis die Maustaste losgelassen wird), auch wenn dieser anschließend bewegt wird. Dies ermöglicht eine vollständig einhändige und optional eine beidhändige Kontrolle.

Der prinzipielle Ablauf der Interaktion ist für alle Proxytechniken gleich und sieht im Detail wie folgt aus:

1. Der Benutzer berührt (klickt) in die Nähe des zu selektierenden Objektes.
2. Falls sich nur ein Objekt innerhalb des festgelegten Radius befindet, wird dieses direkt selektiert (die Selektion ist abgeschlossen). Falls sich mehrere Objekte innerhalb des festgelegten Radius befinden, wird für jedes Zielobjekt ein Proxyobjekt erzeugt.
3. Der Benutzer kann den Finger (oder den Mauspointer) auf ein Proxyobjekt bewegen und den Finger heben (die Maustaste loslassen), um das verknüpfte Zielobjekt zu selektieren. In einer Multi-touch Umgebung kann der Benutzer auch mehrere Objekte gleichzeitig durch den Einsatz weiterer Finger selektieren.
4. Falls der Benutzer den Finger anhebt (die Maustaste loslässt), ohne ein Objekt zu selektieren, werden die Proxyobjekte entfernt.

Durch die festgelegte Radiusbegrenzung können in einer MT-Umgebung mehrere Benutzer gleichzeitig in verschiedenen Bereichen des Bildschirms Objekte selektieren (mehrhändige Interaktion; vgl. Kapitel 4), ohne sich gegenseitig zu beeinflussen.

Gestaltungsmöglichkeiten

Da die Gestaltung der Stellvertreterobjekte grundsätzlich nur durch die Kreativität des Designers beschränkt ist, ist der theoretische Gestaltungsraum sehr

groß. Deshalb werden für die folgende Untersuchung einige Optionen im Vorfeld durch Expertenrunden ausgeschlossen, um eine handhabbare Anzahl an Proxyeigenschaften mit dem größten Potenzial zu identifizieren. In zusätzlichen informellen Vortests wird die Anzahl der in der Nutzerstudie zu untersuchenden Eigenschaften weiter reduziert. Um das Potenzial der Gestaltungsmöglichkeiten aufzuzeigen, schließt die folgende Darstellung verschiedener Gestaltungsmöglichkeiten daher auch Eigenschaften ein, die nicht Teil des folgenden Experiments sind.

Unterschiedlichste Gestaltungseigenschaften können potenziell dazu beitragen, die Verbindung zwischen dem Proxyobjekt und dem referenzierten Objekt für den Benutzer klarer erkennbar zu machen, um so eine schnelle und präzise Selektion zu ermöglichen.

Radius-Schwellwert (Threshold Radius) für die Proxyerzeugung: Für den Algorithmus können verschiedene Schwellwerte verwendet werden, um die Menge der Objekte zu bestimmen, für die Proxyobjekte generiert werden. Verschiedene Einstellungen beeinflussen im Wesentlichen die Anzahl der Proxyobjekte auf dem Bildschirm. Im Verlauf des Experiments wird diese Eigenschaft nach einmaliger Festlegung nicht verändert.

Verbindende grafische Elemente zwischen Proxy und Zielobjekt: In allen implementierten Techniken wird als Standardelement eine verbindende Linie zwischen den Proxyobjekten und den referenzierten Objekten angezeigt. Dies ist eine etablierte Technik und dient als Baseline, um festzustellen, ob andere Eigenschaften zusätzliche Vorteile bieten.

Farbübereinstimmung zwischen Proxyobjekt und Zielobjekt: Werden Proxyobjekt und Zielobjekt in den gleichen oder ähnlichen Farben angezeigt, kann dies ein starker visueller Hinweis auf eine Verbindung sein. Farbe ist daher eine der Eigenschaften, die im Fokus des Experiments stehen.

Bewegungsübereinstimmung zwischen Proxy- und Zielobjekt: Dies bedeutet, dass Proxyobjekte, welche sich bewegende Zielobjekte referenzieren, die Bewegung des Zielobjektes übernehmen, allerdings mit einer deutlich reduzierten und festgelegten Geschwindigkeit. Diese Eigenschaft ist Teil der Untersuchung im Rahmen des Experimentes.

Gerichtete Vibration der Proxyobjekte: Diese Eigenschaft ist verwandt mit der Bewegungsübereinstimmung. Anstatt einer konstanten, linearen Bewegung,

bewegt sich das Proxyobjekt hierbei jedoch kontinuierlich vor und zurück, entlang der Bewegungsrichtung des referenzierten Objektes. Diese Eigenschaft wird auf Grund negativer Resultate aus den informellen Vortests (kein erkennbarer Einfluss auf die Selektion) nicht für das Experiment in Betracht gezogen.

Kollisionserkennung und Physik für Proxyobjekte: Kollisionserkennung und Physik sind eine Möglichkeit, um sicherzustellen, dass sich Proxyobjekte niemals überlappen. Eine andere Möglichkeit (siehe unten) sind spezielle räumliche Anordnungen. Proxyobjekte schieben sich gegenseitig aus dem Weg. Diese Option wird während aller Untersuchungen aktiviert.

Räumliche Anordnung der Proxyobjekte: Die räumliche Anordnung kann einen entscheidenden Einfluss bzgl. Geschwindigkeit und Präzision der Selektion haben. Es sind verschiedene Strategien denkbar, z.B. könnte die räumliche Anordnung auf die typische Fingerhaltung des Benutzers optimiert werden. Um zunächst Bewegung und Farbe fokussiert untersuchen zu können, wird für das Experiment stets eine einfache kreisförmige Anordnung der Proxyobjekte (vgl. Abbildung 5.1) gewählt.

Relative Größe der Proxyobjekte: Die Größe der Proxyobjekte kann proportional zu der Größe des referenzierten Objektes (im 2D Fall) oder auch zur Entfernung des Objektes (im 3D Fall) angepasst werden. Wichtig ist eine Normalisierung, um eine minimale bzw. maximale Größe zu garantieren, so dass eine Auswahl mit dem Finger möglich ist. Außerdem besteht nach *Fitts' Law* ein Zusammenhang zwischen Objektgröße, Präzision und Geschwindigkeit der Selektion. Es ist daher schwierig, die Größen der Proxyobjekte – insbesondere für dynamische Zielobjekte – zu normalisieren und kann zu verwirrenden Effekten führen. Auf eine weitere Untersuchung dieser Eigenschaft wird daher zunächst verzichtet und die Größe der Proxyobjekte auf einen konstanten Wert eingestellt.

Die Form der Proxyobjekte: Um zunächst ein besseres Verständnis von den übrigen Eigenschaften zu gewinnen, wird die Form der Proxyobjekte als komplexe Gestaltungsdimension ebenfalls ausgeschlossen und die Untersuchung auf kreisförmige Proxyobjekte beschränkt.

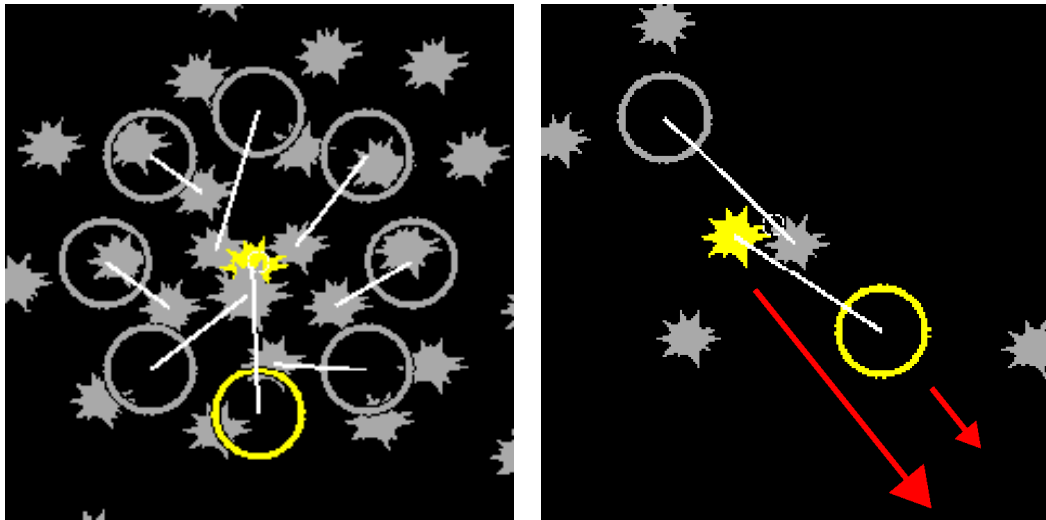


Abbildung 5.1: Screenshots der proxybasierten Techniken für verdeckte (links) und dynamische Objekte (rechts) – Die Bewegungsrichtung der Proxyobjekte ist durch rote Pfeile angezeigt. In den dargestellten Beispielen ist *Farbübereinstimmung* und *Bewegungsübereinstimmung* aktiviert (Abbildung aus Herrlich u. a. (2011b)).

5.1.3 Experiment

Im Fokus des Experiments stand die Gegenüberstellung von verschiedenen proxybasierten Selektionstechniken. Zusätzlich wurde *Nearest-target*-Selektion¹ ohne Proxyobjekte als Kontrolltechnik in die Untersuchung mit einbezogen.

Als unabhängige Variablen wurden wie im vorherigen Abschnitt beschrieben Farbübereinstimmung und Bewegungsübereinstimmung untersucht, d.h. vier Technikvarianten (keine Farbe/keine Bewegung, Farbe/keine Bewegung, keine Farbe/Bewegung, Farbe/Bewegung) mit Proxyobjekten und zusätzlich die Referenztechnik als Kontrollbedingung ganz ohne Proxyobjekte. Es wurde ein intraindividuelles Design (abhängige Gruppen) verwendet und eine Randomisierung mit der *Latin-Square*-Methode durchgeführt, um die Anzahl der Testfälle zu reduzieren.

Die Aufgabe der Teilnehmer bestand jeweils darin, ein farblich hervorgehobenes Zielobjekt zwischen einer Anzahl anderer Objekte zu selektieren. Dabei wurden drei verschiedene Szenarien getestet, welche jeweils vier verschiedene

¹Hierbei wird jeweils das Objekt innerhalb eines festgelegten Radius ausgewählt, dass sich in der geringsten Entfernung zu der Berührung bzw. dem Mauspointer befindet.

Positionen des Zielobjektes beinhalteten. In Szenario 1 (*statisches Szenario*) hat sich weder das Zielobjekt bewegt, noch wurde es durch andere Objekte verdeckt. In Szenario 2 (*verdecktes Szenario*) wurde das Zielobjekt stets zu $\frac{2}{3}$ seiner Grundfläche durch andere Objekte verdeckt (vgl. Abbildung 5.1). In Szenario 3 (*dynamisches Szenario*) bewegte sich das Zielobjekt auf einer festgelegten Bahn von einer zur anderen Seite des Bildschirms (vgl. Abbildung 5.1). Jedes Szenario wurde vier Mal wiederholt, es wurden also insgesamt 240 Einzelversuche pro Proband durchgeführt (5 Techniken x 4 Wiederholungen x 3 Szenarien x 4 Positionen). Alle Tests wurden auf einem 3M M2256PW MT-Bildschirm durchgeführt. Für jeden Einzelversuch hatte ein Proband maximal 16 Sekunden Zeit, nach deren Ablauf der Einzelversuch als nicht erfolgreich gewertet wurde. Alle innerhalb dieser Zeitspanne falsch selektierten Objekte wurden erfasst.

5.1.4 Ergebnisse

Insgesamt nahmen acht Probanden (5 m, 3 w) an dem Experiment teil. Eine einfaktorielle Varianzanalyse der Selektionszeiten zeigte signifikante Unterschiede zwischen den Wiederholungen auf ($F_{3;21} = 14,835, p < 0,0005$). Paarweise Post-hoc-Tests ergaben jeweils einen signifikanten Unterschied zwischen der ersten Wiederholung und den Wiederholungen zwei bis vier (Sidak-korrigierte p-Werte nach SPSS Post-hoc-Tests für wiederholte Messungen für die Wiederholungskombinationen 1+2, 1+3 und 1+4: $p = 0,016, p = 0,021, p = 0,007$). Dagegen zeigten sich jedoch keine signifikanten Unterschiede zwischen den Wiederholungen zwei bis vier. Für alle folgenden statistischen Auswertungen wurde auf Basis dieser Ergebnisse die erste Wiederholung nicht betrachtet. Die Wiederholungen zwei bis vier wurden durch arithmetrische Mittelwertbildung aggregiert.

Im Folgenden wurde eine zweifaktorielle Varianzanalyse der Selektionszeiten für die unabhängigen Variablen Szenario und Technik durchgeführt (vgl. Abbildung 5.2). Ein Test nach Mauchly ergab, dass die Voraussetzung der Sphärizität in allen Fällen verletzt wurde. Daher wurden jeweils die korrigierten Werte nach Greenhouse-Geisser verwendet. Diese zeigten signifikante Unterschiede zwischen den Szenarien ($F_{1,076;7,535} = 33,136, p < 0,0005$; korr. nach Greenhouse-Geisser) und zwischen den Selektionstechniken ($F_{1,791;12,539} =$

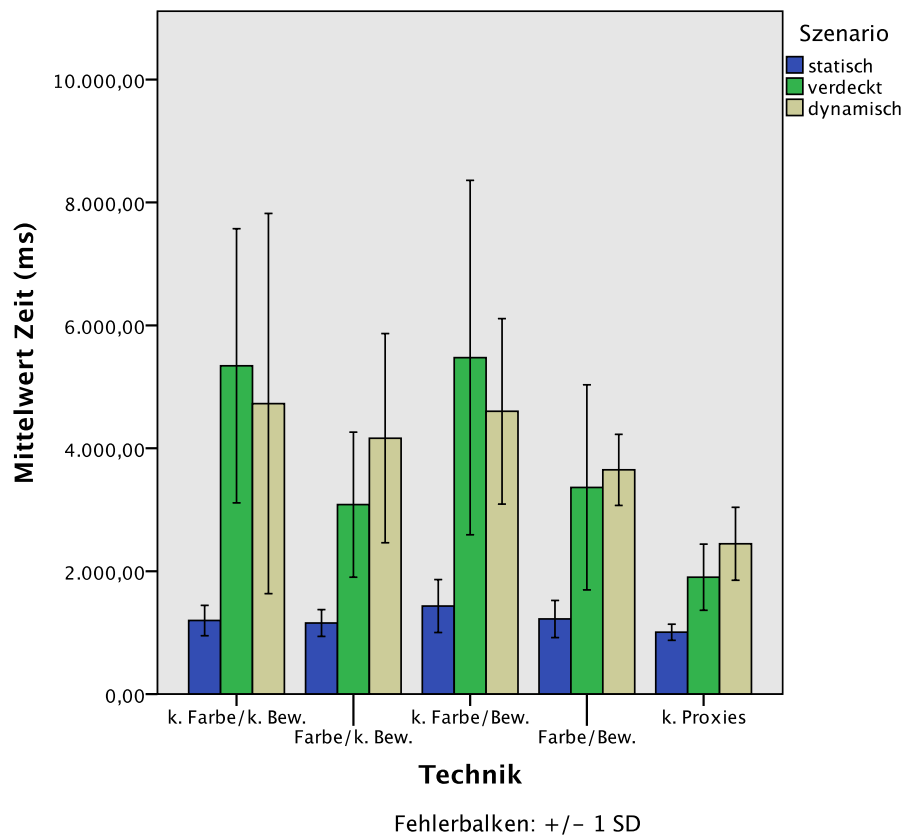


Abbildung 5.2: Mittelwerte der Selektionszeiten je Technik und Szenario

11,792, $p = 0,002$; korr. nach Greenhouse-Geisser) sowie eine signifikante Interaktion zwischen Szenario und Technik ($F_{1,442;10,094} = 4,547, p = 0,048$; korr. nach Greenhouse-Geisser).

Paarweise Post-hoc-Tests ergaben für das statische Szenario signifikant kürzere Selektionszeiten als für das verdeckte Szenario (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,004$) und ebenfalls signifikant kürzere Selektionszeiten für das verdeckte Szenario im Vergleich zum dynamischen Szenario (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests; $p < 0,0005$).

Für die Variable Technik ergaben paarweise Post-hoc-Tests signifikant längere Selektionszeiten für die Selektionstechnik *keine Farbe/Bewegung* im Vergleich zur Referenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,013$) sowie signifikant längere Selektionszeiten für die Selektionstechnik *Farbe/Bewegung* im Vergleich zur Referenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,01$). Auf Grund der eingesetzten Sidak-Korrektur ist zu erwarten, dass die Wahrscheinlichkeit von Typ II Fehlern ansteigt (resultierend aus einer zu konservativen Überkorrektur der p-Werte). Daher werden im Folgenden ebenfalls die Ergebnisse aufgeführt, welche das angestrebte Signifikanzniveau von 5% um weniger als 2% überschreiten. Diese sollten allerdings nur als vorsichtiger Trend gewertet werden. Die Technik *Farbe/keine Bewegung* war im Mittel schneller als die Technik *keine Farbe/Bewegung* (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,052$), allerdings langsamer als die Referenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,065$). Außerdem war die Technik *keine Farbe/Bewegung* langsamer als die Technik *Farbe/Bewegung* (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,052$).

Für die Interaktion zwischen den Variablen Szenario und Technik ergaben paarweise Post-hoc-Tests, dass für das verdeckte Szenario die Technik *keine Farbe/keine Bewegung* signifikant langsamer war als die Techniken *Farbe/keine Bewegung* (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,015$), *Farbe/Bewegung* (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,009$) und als die Referenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,016$). Außerdem war für dieses Szenario die Technik *Farbe/keine Bewegung* ebenfalls signifikant langsamer als die Re-

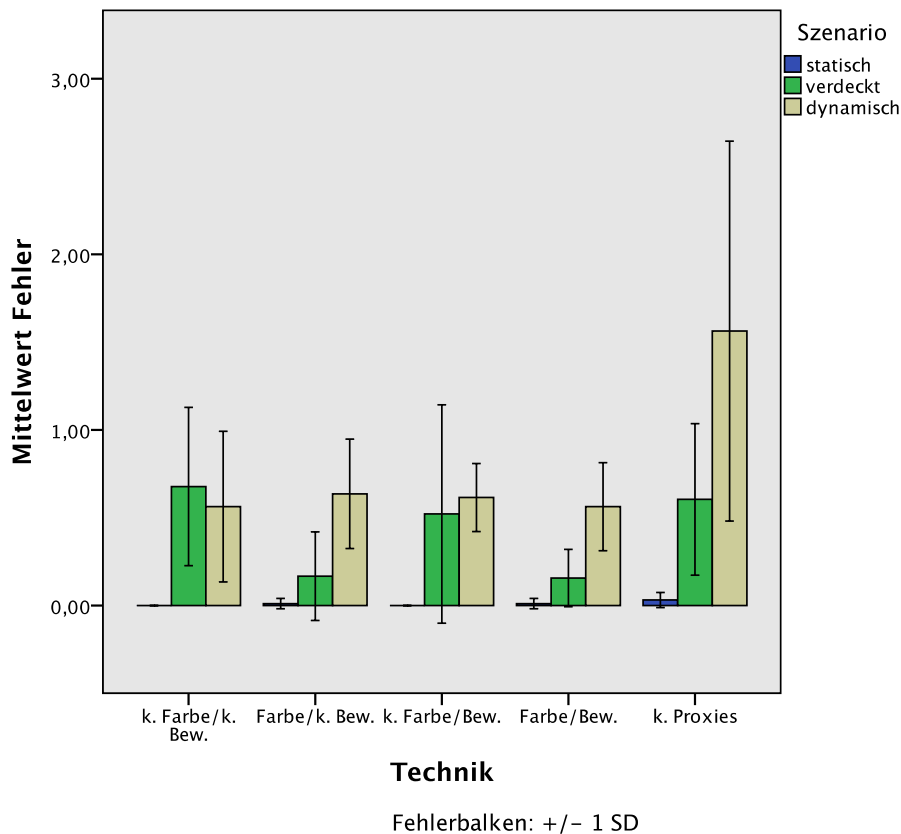


Abbildung 5.3: Mittelwerte der Fehleranzahlen je Technik und Szenario

ferenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,044$). Für das dynamische Szenario war außerdem die Technik *Farbe/Bewegung* signifikant langsamer als die Referenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,01$).

Für die abhängige Variable *Anzahl der Selektionsfehler* wurde ebenfalls eine zweifaktorielle Varianzanalyse nach den Variablen Szenario und Technik durchgeführt (vgl. Abbildung 5.3). Diese ergab jeweils signifikante Unterschiede für Szenario ($F_{2;14} = 21,384, p < 0,0005$; Mauchly n.s.) und Technik ($F_{1;543;10,8} = 12,464, p = 0,002$; korr. nach Greenhouse-Geisser) sowie eine signifikante Interaktion zwischen diesen Variablen ($F_{1;548;10,833} = 6,322, p = 0,02$; korr. nach Greenhouse-Geisser).

Paarweise Post-hoc-Tests zeigten eine signifikant geringere Fehleranzahl für das statische Szenario im Vergleich zum verdeckten (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,033$) und zum dynamischen Szenario (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,004$). Außerdem zeigte

sich eine signifikant geringere Fehleranzahl für das verdeckte Szenario im Vergleich zum dynamischen Szenario (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,005$).

Die Anzahl der Fehler für die Techniken *Farbe/keine Bewegung* und *keine Farbe/Bewegung* war jeweils signifikant geringer als für die Referenztechnik ohne Proxies (Sidak-korrigierte p-Werte nach SPSS Post-hoc-Tests: $p = 0,031, p = 0,04$).

Bzgl. der Interaktion zwischen den beiden Variablen zeigte sich, dass im verdeckten Szenario die Fehleranzahl für die Technik *keine Farbe/keine Bewegung* signifikant höher lag als für die Techniken *Farbe/keine Bewegung* (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,018$) und *Farbe/Bewegung* (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,034$). Außerdem war für dieses Szenario die Fehleranzahl für die Technik *Farbe/keine Bewegung* signifikant geringer als für die Referenztechnik ohne Proxies (Sidak-korrigierter p-Wert nach SPSS Post-hoc-Tests: $p = 0,018$).

Gepaarte t-Tests für die gemittelten Werte aller Proxytechniken gegenüber der Referenztechnik ohne Proxies ergaben, dass die Proxytechniken (vgl. Abbildung 5.4) signifikant langsamer abschnitten ($t_7 = 4,65, p = 0,002$), die Anzahl der Selektionsfehler (vgl. Abbildung 5.5) im Vergleich jedoch signifikant geringer war ($t_7 = -4,118, p = 0,004$).

Qualitative Beobachtungen zeigten, dass fünf Teilnehmer gelegentlich dadurch verwirrt wurden, dass sie ihre Finger zu früh wieder von der Oberfläche des Bildschirms nahmen und dementsprechend die Proxyobjekte sofort verschwanden. Im dynamischen Szenario versuchten vier Teilnehmer mehrfach das Zielobjekt zu „fangen“, wobei sie die Proxyobjekte vollständig ignorierten. Dagegen verwendeten alle Teilnehmer die Proxyobjekte erfolgreich für verdeckte Zielobjekte.

5.1.5 Diskussion

Die Auswertung des Experiments zeigt, dass die Proxytechniken grundsätzlich erfolgreich auch durch ungeübte Teilnehmer angewendet werden konnten. Allerdings muss die These bzgl. des Geschwindigkeitsvorteils im Vergleich zu der Referenztechnik revidiert werden. Dagegen konnte bestätigt werden, dass Proxyobjekte die Fehlerrate bei der Selektion senken. Dies könnte insbeson-

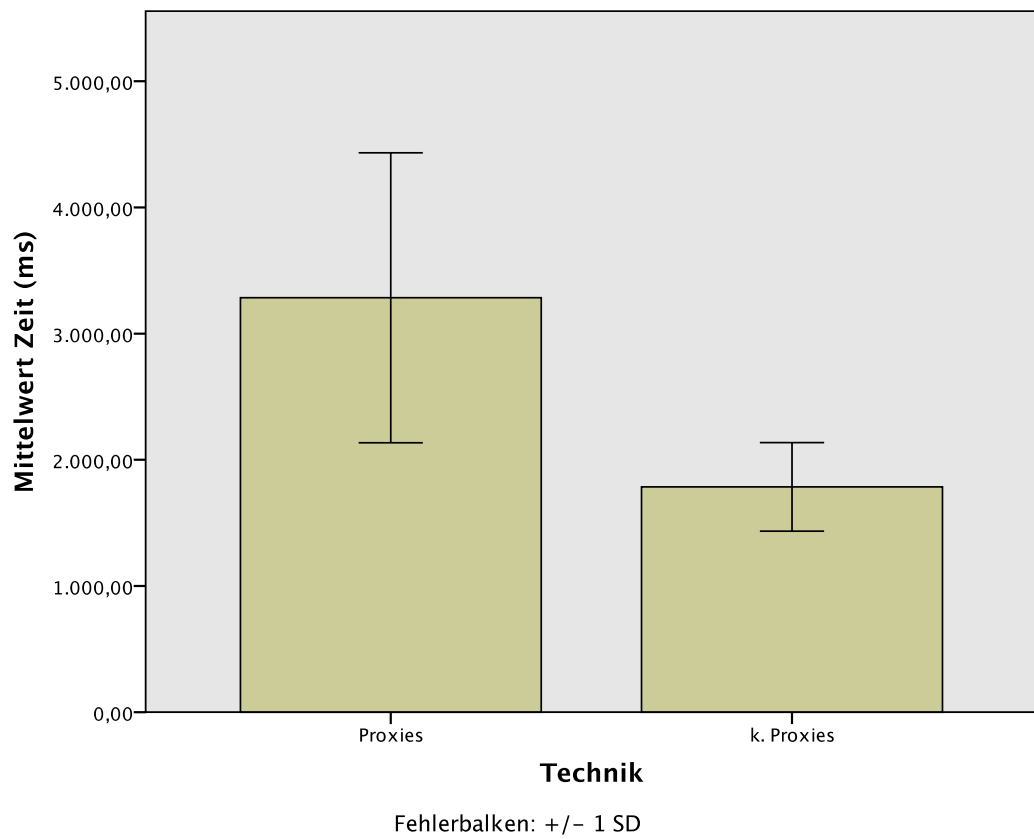


Abbildung 5.4: Kombiniertes Mittelwert der Selektionszeiten aller Proxytechniken im Vergleich zur Referenztechnik ohne Proxies

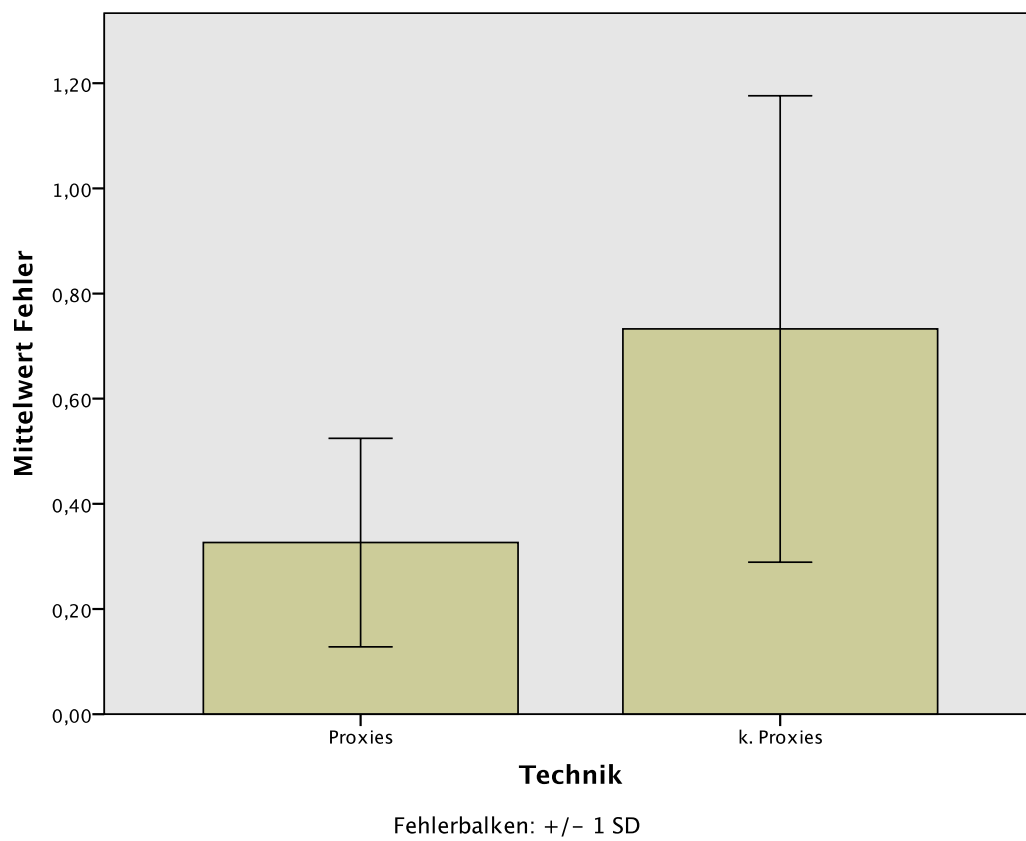


Abbildung 5.5: Kombiniertes Mittelwert der Fehleranzahlen aller Proxytechniken im Vergleich zur Referenztechnik ohne Proxies

Gestaltungsdimension	Einordnung	Erklärung
Multi-point vs. Multi-touch	Multi-point	Berührungen werden als Punkte verarbeitet
Statischer vs. dynamischer Kontext	eher dynamisch	kein expliziter Moduswechsel notwendig; Distanz als einfache Heuristik
Integrierte vs. Separierte Freiheitsgrade	integriert	nur absolute 2D Positionierung möglich
Direkte vs. indirekte Interaktion	eher indirekt	Interaktion über Proxyobjekte
Mehrhändige Interaktionsmuster	mehrhändig	keine bevorzugte Rollenverteilung der Hände

Tabelle 5.1: Einordnung der proxybasierten Selektionstechniken gemäß den Gestaltungsdimensionen aus Kapitel 4

dere für 3D-Modellierungsanwendungen ein großer Vorteil sein. Auch die Fehlerrate für die Selektion sich bewegender Objekte war für die proxybasierte Selektion niedriger als für die Vergleichstechnik. Allerdings scheint der visuelle Reiz durch die sich bewegenden Zielobjekte so stark zu sein, dass dieser Faktor einige der Benutzer ablenkt. Eventuell erfordert die Anwendung dieser Technik daher eine längere Lernphase. In den durchgeführten Experimenten zeigte sich *Farbe* als der wichtigste visuelle Anreiz, jedoch hatte auch *Bewegung* grundsätzlich einen positiven Effekt und könnte zusätzlich zu *Farbe* oder in Anwendungen eingesetzt werden, in denen *Farbe* nicht zur Verfügung steht. In realen Anwendungen könnte eine Kombination aus Proxytechniken und *Nearest-Target*-Selektion die besten Ergebnisse liefern.

5.1.6 Fazit

In diesem Abschnitt wurde eine proxybasierte Technik zur Selektion in 2D- und 3D-Umgebungen vorgestellt. Verschiedene Eigenschaften und Gestaltungsmöglichkeiten wurden vorgeschlagen und aus dieser Menge wurden *Farbübereinstimmung* und *Bewegungsübereinstimmung* im Rahmen eines Experiments im Vergleich zu einer Referenztechnik näher untersucht. Es zeigte sich, dass in der durchgeführten Untersuchung die Proxytechniken die Fehlerrate der Selektion signifikant senken konnten. Die Referenztechnik scheint dagegen die schnellere Selektion zu ermöglichen. Eine Kombination beider Ansätze könnte demnach in realen Applikationen hilfreich sein, um anwendungsspezifisch einen Kompro-

miss zwischen Geschwindigkeit und Fehlerrate zu finden. Tabelle 5.1 fasst die Einordnung der vorgestellten Techniken gemäß der Gestaltungsdimensionen nach Kapitel 4 zusammen.

5.2 Manipulation in sechs Freiheitsgraden

Im Bereich der graphischen 2D-Anwendungen haben sich einige MT-Interaktionstechniken etabliert, beispielsweise für das Betrachten und Sortieren von Bildern oder Videos. Diese sind meist charakterisiert durch eine nahtlose Integration verschiedener Operationen, z.B. Translation, Rotation und Zoom. Für das Feld der 3D-Anwendungen gibt es dagegen bisher kaum etablierte Interaktionstechniken, obwohl MT-Eingaben potenziell zusätzliche Freiheitsgrade (DOF) bieten, die für solche Anwendungen genutzt werden könnten. Grundsätzlich ist die Interaktion mit einer 2D-Projektion eines 3D-Raums komplexer als die reine 2D-Interaktion, dennoch ist es das Ziel der in diesem Abschnitt präsentierten Interaktionstechniken, einige der erfolgreichen Eigenschaften der 2D-Interaktionstechniken auf den 3D-Fall zu übertragen.

In diesem Abschnitt werden zwei MT-Interaktionstechniken für die integrierte Manipulation von sechs Freiheitsgraden (6DOF) vorgestellt. Grundidee ist die Abbildung von affinen 2D-Transformationen auf den 3D-Raum und die Integration mit etablierten Kontrolltechniken wie der *Turntable*-Metapher, die in einigen 3D-Modellierungswerkzeugen zum Einsatz kommt. Beide Varianten sind effizient zu berechnen, können flexibel mit einer oder beiden Händen verwendet werden und erlauben sowohl die integrierte als auch die separierte Ausführung von Translation und Rotation im 3D-Raum (vgl. Kapitel 4).

Die *PieRotate*-Technik erlaubt eine sehr spezifische Separation der einzelnen Rotations- und Translationsfreiheitsgrade, während die *Turn&Roll*-Technik auf die integrierte Manipulation und eine intuitive Bedienung optimiert ist (vgl. Abb. 5.6). Eine durchgeführte Nutzerstudie zum Vergleich beider Techniken belegt, dass die integrierte Manipulation mittels *Turn&Roll* der separierten Manipulation mittels *PieRotate* im Rahmen des Experiments klar überlegen war. Dennoch finden sich auch Argumente zu Gunsten des Ansatzes von *PieRotate*. Die Analyse der qualitativen Beobachtungen vermittelt weitere Einsichten in Manipulationsstrategien für 3D-Objekte auf großen MT-Bildschirmen.

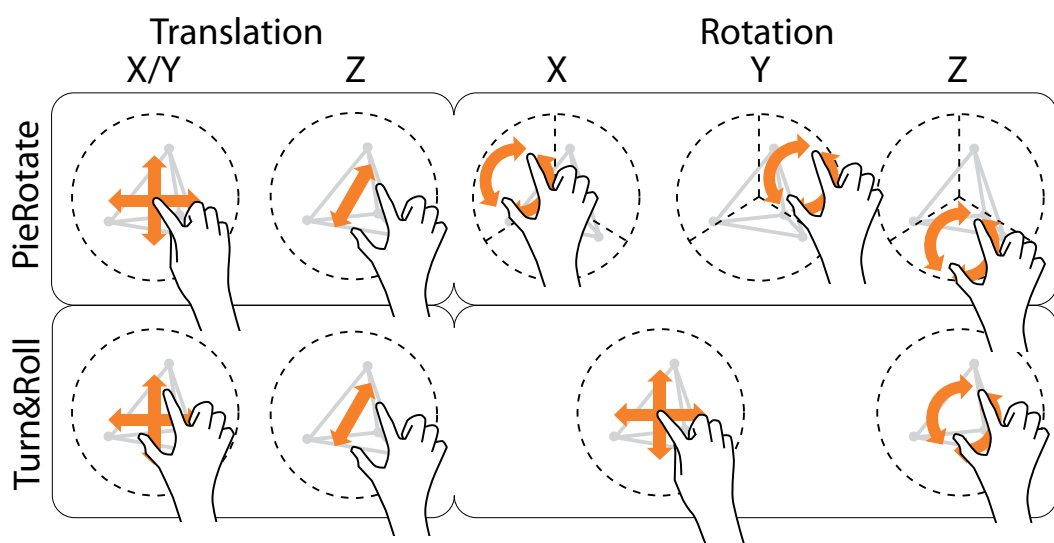


Abbildung 5.6: Gesten für die 3D-Translation und -Rotation mit den vorgestellten Techniken. Alle Zweifingergesten können auch mit mehr als zwei Fingern ausgeführt werden. Es wird nur die einhändige Bedienung gezeigt, die Bedienung mit beiden Händen ist aber ebenfalls möglich (Abbildung aus Herrlich u. a. (2011a)).

5.2.1 Stand der Technik

2D-MT-Manipulationstechniken wurden bereits intensiv erforscht (vgl. auch Kapitel 4). Kruger u. a. (2005) integrieren 2D-Rotation und -Translation in einem physikalisch-inspirierten Ansatz basierend auf virtueller Reibung in Abhängigkeit der relativen Positionen der Fingerberührung und des Objektzentrums. In einer Nutzerstudie können sie zeigen, dass ihr *RNT*-Ansatz der separierten Manipulation über Handles an den Objektkanten überlegen ist. Moscovich und Hughes (2006) stellen Mehrfingertechniken vor, welche 2D-Rotation, -Translation und -Skalierung durch die Berechnung der approximierten affinen Transformationen aus den Fingerberührungen integrieren. *PieRotate* bzw. *Turn&Roll* verwenden das gleiche Approximationsverfahren wie Moscovich und Hughes (2006), bilden die resultierenden Transformationen dann allerdings unterschiedlich in den 3D-Raum ab. Für eine grundlegende Diskussion einfacher 2D-Translations- und -Rotationstechniken wird auf Hancock u. a. (2006) verwiesen.

Hancock u. a. (2007) beschreiben verschiedene Mehrfingertechniken, welche Krugers *RNT* auf den 3D-Raum erweitern und theoretisch die Manipulation von fünf bzw. sechs Freiheitsgraden erlauben. Die vorgestellten Techniken sind jedoch in der Tiefendimension stark eingeschränkt. Sie basieren auf dedizierten Bereichen und einer festgelegten Abbildung und Reihenfolge der ersten drei Fingerberührungen. Zudem scheint ihr Ansatz am geeignetsten für eine beidhändige Kontrolle. In einer neueren Arbeit erweitern Hancock u. a. (2009a) diese Techniken und schlagen ebenfalls die *Pinch-Geste* zur Kontrolle der Translation entlang der Tiefenachse vor. Die Abbildung und Reihenfolge der Fingerberührungen sind jedoch auch bei diesen Techniken immer noch fest vorgegeben und auf maximal drei Finger beschränkt. Reisman u. a. (2009) verwenden ein nicht-lineares Energieoptimierungsverfahren basierend auf der Methode der kleinsten Quadrate, um 3D-Transformationen direkt aus den Screen-Space-Koordinaten der Fingerberührungen abzuleiten. Der vorgestellte Ansatz scheint sehr leistungsfähig, der Rechenaufwand ist allerdings vergleichsweise groß und einige intrinsische Ambiguitäten des Verfahrens können lediglich durch empirisch ermittelte, zusätzliche Konstanten ausgeglichen werden. Martinet u. a. (2010) studieren Techniken, welche einen zusätzlichen Finger für die Translation in drei Freiheitsgraden verwenden, untersuchen al-

lerdings keine Rotationsoperationen.

Aus dem Bereich der physikalisch-basierten Interaktion stellen Wilson u. a. (2008) und Wilson (2009) Verfahren vor, um Berührungseingaben in eine Physiks simulation einzubinden und das *Greifen* von 3D-Objekten auf Basis einfacher Mechanik zu simulieren. Beide Verfahren sind in der beschriebenen Form aber im Wesentlichen auf eine 2D-Interaktion beschränkt, da Operationen entlang der Tiefenachse kaum kontrolliert ausgeführt werden können.

Die Frage, ob eine Integration von Rotation und Translation im Allgemeinen vorteilhaft ist oder nicht, wurde in der Literatur bereits intensiv diskutiert, allerdings bisher nicht abschließend entschieden. Es scheint hier eine starke Abhängigkeit von dem konkreten Anwendungskontext bzw. der Aufgabe vorzuliegen, wie Jacob u. a. (1994) anmerken. Auch Hancock u. a. (2007) diskutieren die widersprüchlichen Forschungsergebnisse in diesem Bereich. So weisen Ergebnisse von Moscovich und Hughes (2008) darauf hin, dass die einhändige integrierte Kontrolle von Position und Orientierung möglich und hilfreich ist. Auf der anderen Seite stellen Nacenta u. a. (2009) verschiedene Techniken für die Separierung von Rotations- und Translationsoperationen vor.

5.2.2 Interaktionstechniken

Existierende 3D-Manipulationstrategien auf MT-Bildschirmen können grob in drei Ansätze unterteilt werden: Handles/Widgets, affine Transformationen und Mehrfingergesten (vgl. Kapitel 4).

Handles (oder *Gizmos*), wie sie als Werkzeuge in etablierter CAD oder Animationssoftware häufig eingesetzt werden, sind ohne weitreichende Anpassungen für MT-Bildschirme ungeeignet, da diese für eine sehr präzise *Single-Pointer*-Kontrolle ausgelegt sind. MT-Eingaben sind dagegen durch eine im Vergleich zur Maus unpräzisere Kontrolle und mehr Freiheitsgrade, sowie *Multi-Pointer*-Interaktion gekennzeichnet.

Die direkte Berechnung von *affinen 3D-Abbildungen/-Transformationen* ist sehr rechenintensiv, die Bedienung teilweise kontraintuitiv und die Lösungsverfahren bedürfen der Einführung von empirisch zu ermittelnden zusätzlichen Konstanten, um Ambiguitäten zu verhindern und die Verfahren zu stabilisieren (vgl. Reisman u. a. (2009)).

Mehrfingergesten können die zusätzlichen Freiheitsgrade von MT auf unter-

schiedliche Weise abbilden: Die Anzahl der Finger kann verwendet werden, um modal zwischen verschiedenen Manipulationszuständen oder Werkzeugen zu wechseln (Matejka u. a., 2009). Dieses Vorgehen erlaubt allerdings per Definition keine integrierte Manipulation. Eine feste Zuordnung bestimmter Freiheitsgrade auf einzelne Finger, z.B. anhand der Reihenfolge der Berührungen (Hancock u. a., 2007) macht eine gleichzeitige Kontrolle verschiedener Freiheitsgrade schwierig und der Benutzer muss sich die feste Zuordnung merken. Bei mehr als drei Freiheitsgraden wird die Bedienung hier zunehmend komplex.

Der hier vorgeschlagene neue Ansatz besteht darin, affine 2D-Transformationen auf geeignete Weise in den 3D-Raum abzubilden, um so 3D-Manipulationen durchzuführen. Dieser Ansatz bietet drei klare Vorteile gegenüber den existierenden Techniken:

- *Geringer Rechenaufwand und stabile Berechnung:* Affine 2D-Transformationen sind mit wesentlich geringerem Aufwand zu berechnen als 3D-Transformationen. Die Berechnung ist robust und flexibel in Bezug auf die Anzahl der Fingerberührungen und beruht nicht auf einer festgelegten Reihenfolge/Rolle der Finger. Je mehr Finger eingesetzt werden, umso stabiler und genauer wird die Kontrolle, da diese auf alle Finger aufgeteilt wird.
- *Einhändige oder beidhändige Kontrolle:* Die vorgeschlagenen Techniken (siehe nächster Abschnitt) können mit einer Hand bedient werden, um die andere z.B. für Kameramanipulationen freizuhalten. Sie funktionieren aber auch mit beiden Händen und erlauben so eine präzisere Kontrolle.
- *Gleichzeitige und unabhängige Kontrolle:* Die Techniken erlauben sowohl eine gleichzeitige als auch eine unabhängige Ausführung verschiedener Operationen, d.h. entweder Rotation oder Translation oder beides gleichzeitig auf eine konsistente Art und Weise. Hierbei liegt die Annahme zu Grunde, dass die gleichzeitige Ausführung effizienter, in bestimmten Fällen aber auch die unabhängige Kontrolle wünschenswert ist (Nacenta u. a., 2009).

Z-Pinch

Die Nutzung der Pinch-Geste zur Kontrolle der Z- oder Tiefenachse basiert auf existierenden Ansätzen, welche einen zusätzlichen Finger verwenden, um die Änderung in der Tiefendimension zu bestimmen (Hancock u. a., 2007; Martinet u. a., 2010). Konzeptuell kann man dies als äquivalent zur Verwendung der Pinch-Geste für das Skalieren von Objekten in 2D-Anwendungen betrachten. Die Änderungen des Abstandes zwischen den Fingern wird auf die Tiefe abgebildet: Eine Verkleinerung des Abstandes vergrößert die Entfernung zwischen Objekt und Betrachter in Richtung der Tiefenachse, eine Vergrößerung des Abstandes bewegt das Objekt auf der Tiefenachse näher an den Betrachter. Typischerweise kann diese Geste einhändig, z.B. mit Zeigefinger und Daumen ausgeführt werden. Es können aber auch andere Finger oder mehr als zwei Finger derselben oder der anderen Hand eingesetzt werden.

Die folgenden Techniken – *Turn&Roll* bzw. *PieRotate* (vgl. Abb. 5.6) – sind jeweils zusammengesetzt aus einer gemeinsamen Translationstechnik – *Z-Pinch* – und zwei verschiedenen Rotationstechniken. Die Kombination von *Z-Pinch* mit *Turn&Roll* ist für eine einfachere Integration von Rotation und Translation entworfen, während die Kombination mit *PieRotate* stärker auf eine separierte Rotation und Translation ausgerichtet ist. Die Kontrollen werden aktiviert, wenn Fingerberührungen innerhalb des *Objektkontrollbereiches* auftreten. Dies ist die (nicht sichtbare) Projektion der *Bounding Sphere* eines Objektes.

Turn&Roll

Die *Turn&Roll* Technik kombiniert die *Turntable* (Plattenteller)-Metapher, welche häufig in 3D-Modellierungswerkzeugen zum Einsatz kommt, mit einer Rotation um die Z-Achse (*roll*). Mit der horizontalen Bewegung eines Fingers kann das Objekt um die Y-Achse in Bildschirmkoordinaten rotiert werden. Die vertikale Bewegung rotiert das Objekt um die X-Achse. Mit zwei oder mehr Fingern kann eine *Twist*-Geste ausgeführt werden, um das Objekt um die Z-Achse zu rotieren und gleichzeitig eine Translation entlang der X-Y-Achsen und der Z-Achse (*Z-Pinch*) durchgeführt werden (vgl. Abb. 5.6). Dafür werden die Fingerkontakte zunächst als 2D-Eingaben interpretiert und die resultierenden 2D Transformationen berechnet und anschließend auf die entsprechen-

den Achsen abgebildet. Dadurch ermöglicht Turn&Roll zusammen mit Z-Pinch für zwei oder mehr Finger die gleichzeitige Kontrolle von vier Freiheitsgraden (3D-Position und Z-Rotation) und einen nahtlosen Wechsel zur Kontrolle der übrigen zwei Freiheitsgrade (X-Y-Rotation) mit einem Finger.

PieRotate

PieRotate unterteilt den Objektkontrollbereich in verschiedene Regionen, welche es dem Benutzer erlauben, eine Rotation um jeweils eine spezifische Achse auszuführen. Eine *Twist*-Geste rotiert das Objekt um die X-, Y-, oder Z-Achse, je nachdem in welcher Region diese ausgeführt wird. Der kreisförmige Objektkontrollbereich wird dazu in drei „Tortenstück“-artige Bereiche aufgeteilt. Gleichzeitig zur Rotation kann jeweils eine Translation entlang der X-Y-Achsen und der Z-Achse (Z-Pinch) ausgeführt werden (vgl. Abb. 5.6). Wie bei Turn&Roll werden die aus den Fingerkontakten berechneten 2D-Transformationen anschließend auf die entsprechenden Achsen abgebildet. Die Translationsoperation ist unabhängig von der Aufteilung des Kontrollbereiches. PieRotate zusammen mit Z-Pinch erlaubt daher die gleichzeitige Kontrolle von vier Freiheitsgraden (3D Position und eine Rotationsachse). Die Rotationsachse wird ausgewählt durch die Region, in der die *Twist*-Geste begonnen wird.

5.2.3 Experiment

Ziele

Beide Techniken basieren auf dem gleichen Prinzip der Abbildung von 2D-Transformationen auf die 3D-Manipulation. Die Abbildung auf die Translationsoperation ist sehr ähnlich, die Techniken unterscheiden sich allerdings grundlegend bezüglich der Rotationsmetapher und wie diese mit der Translation kombiniert wird. Daher wurde eine Nutzerstudie durchgeführt, um die Unterschiede beider Techniken bezüglich Performanz und Nutzerzufriedenheit zu untersuchen. Ein weiteres Ziel der Studie bestand darin, Strategien und Interaktionsmuster der Testpersonen bzgl. der gleichzeitigen bzw. unabhängigen Ausführung von Translations- und Rotationsoperationen sowie einhändiger und beidhändiger Interaktion zu beobachten.

Versuchsaufbau

Die Probanden führten die Aufgaben im Stehen an einem *impressx xdesk* MT-Tisch aus. Der *xdesk* verwendet die DI-Technik. Der Tisch wies eine Höhe von 90 cm und eine Bildschirmdiagonale von 52 Zoll auf. Zur Darstellung wurde ein Projektor mit einer Auflösung von 1280×800 Pixeln verwendet. Für das Tracking kamen zwei Kameras mit einer Auflösung von jeweils 640×480 Pixeln zum Einsatz, welche mit überlappenden Sichtbereichen angeordnet waren und die Oberfläche mit einer virtuellen Kameraauflösung von ca. 640×900 Pixeln erfassten. Die proprietäre Tracking-Software kommunizierte mittels des TUIO-Protokolls mit einer speziell für den Versuch entwickelten Applikation, welche die Techniken und die Testumgebung bereitstellte.

Aufgabe

Für den Versuch wurde die gleiche 3D-Docking-Aufgabe verwendet, die in früheren Untersuchungen von 3D-Manipulationstechniken zum Einsatz kam (Hancock u. a., 2007, 2009b). Ein Tetraeder mit farblich eindeutig gekennzeichneten Ecken und Kanten musste in einem transparenten Zieltetraeder platziert werden (vgl. Abbildung 5.7). Das Zieltetraeder färbte sich gelb, wenn die Entfernungen zwischen den korrespondierenden Ecken unter einem festgelegten Schwellwert lagen und der Einzelversuch galt als erfolgreich abgeschlossen. Für den Versuch lag der Schwellwert bei 10% der Kantenlänge. Wurde ein Einzelversuch erfolgreich abgeschlossen, begann der nächste mit einem Zielobjekt an anderer Stelle und mit einer anderen Orientierung. Als visuelle Hilfen dienten die perspektivische Darstellung, Verdeckung und Schatten auf der Bodenebene.

Design

Es wurde ein intraindividueller (abhängige Gruppen) Versuchsaufbau mit wiederholten Messungen gewählt. Die unabhängigen Variablen waren Technik (Turn&Roll und PieRotate), (Ziel-)Position (oben-links-vorne, unten-links-hinten, oben-rechts-hinten, unten-rechts-vorne in einer kubischen Anordnung um das Zielobjekt) und Orientierung (in 180° Rotationsschritten um die X-, Y- und Z-Achse). Die Reihenfolge der Techniken wurde zwischen den Probanden

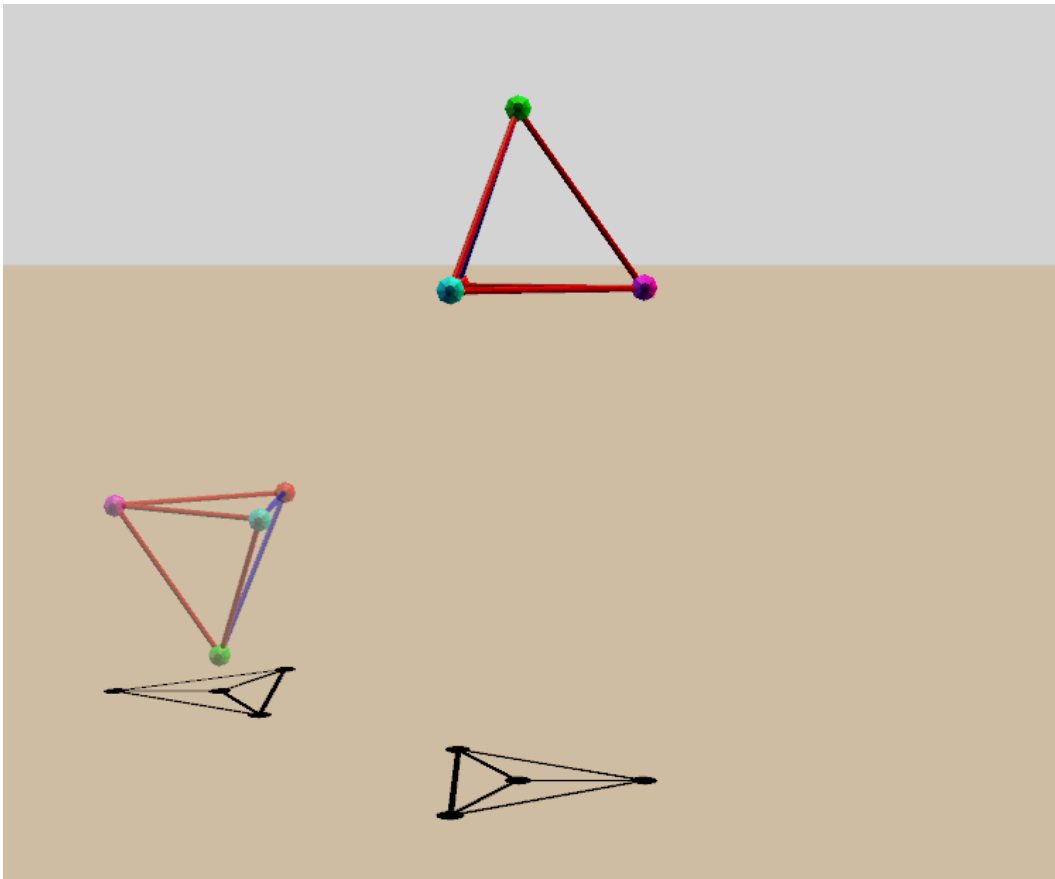


Abbildung 5.7: Ausschnitt aus einem Bildschirmfoto der Versuchsumgebung – Das Tetraeder im Vordergrund muss unter Anwendung der jeweiligen Technik mit dem (leicht transparenten) Zieltetraeder möglichst genau in Deckung gebracht werden. Schatten auf der Bodenebene dienen als visuelle Hilfen für die korrekte Platzierung in der Tiefendimension.

randomisiert.

Jeder Proband führte zwei Wiederholungen der gleichen (zufälligen) Reihenfolge von Positionen/Orientierungen durch, d.h. insgesamt 48 Einzelversuche pro Versuchsteilnehmer. Zu Beginn jedes Technikblocks, wurde die jeweilige Technik erklärt und die Teilnehmer hatten Zeit, um sich mit der Technik anhand einer Beispielaufgabe vertraut zu machen. Dabei wurde besonders darauf geachtet, keine Hinweise auf bestimmte Interaktionsstrategien, z.B. einhändige oder beidhändige Kontrolle, zu geben.

5.2.4 Ergebnisse

Teilnehmer

Insgesamt nahmen acht Probanden (5 m, 3 w) an der Studie teil. Alle waren Rechtshänder, der überwiegende Teil Studenten der Informatik oder verwandter Fächer. Zwei Probanden hatten begrenzte Erfahrung mit 3D-Anwendungen, drei besaßen fortgeschrittene Kenntnisse in diesem Feld und weitere drei sehr fortgeschrittene Kenntnisse. Fünf hatten oberflächliche Erfahrungen mit Touch-Eingaben (meist aus dem mobilen Bereich), drei hatten in dieser Hinsicht bereits umfangreiche Erfahrungen. Es wurden bewusst Teilnehmer mit unterschiedlichen Vorkenntnissen zugelassen, um den Einfluss der Vorerfahrung mit zu beobachten.

Zeiten

Insgesamt schlossen die Probanden 57% der Einzelaufgaben mit PieRotate und 85% der Einzelaufgaben mit Turn&Roll erfolgreich ab. Im zweiten Durchlauf konnten für die PieRotate-Technik im Mittel etwas niedrigere Zeiten beobachtet werden, dieser Lerneffekt erwies sich allerdings nicht als statistisch signifikant. Ein t-Test für gepaarte Stichproben ($t_{97} = 3,166$, $p = 0,002$) bestätigte, dass für Turn&Roll im Mittel signifikant niedrigere Zeiten als für PieRotate gemessen werden konnten (vgl. Abb. 5.8). Eine Varianzanalyse für wiederholte Messungen ergab außerdem signifikante Einflüsse für Position ($F_{3,30} = 3,971$, $p = 0,017$) und eine signifikante Interaktion zwischen Position und Technik ($F_{3,24} = 3,815$, $p = 0,023$). Ein t-Test für gepaarte Stichproben ergab für die PieRotate-Technik, dass die Zeiten für Positionen in der oberen Hälfte der An-

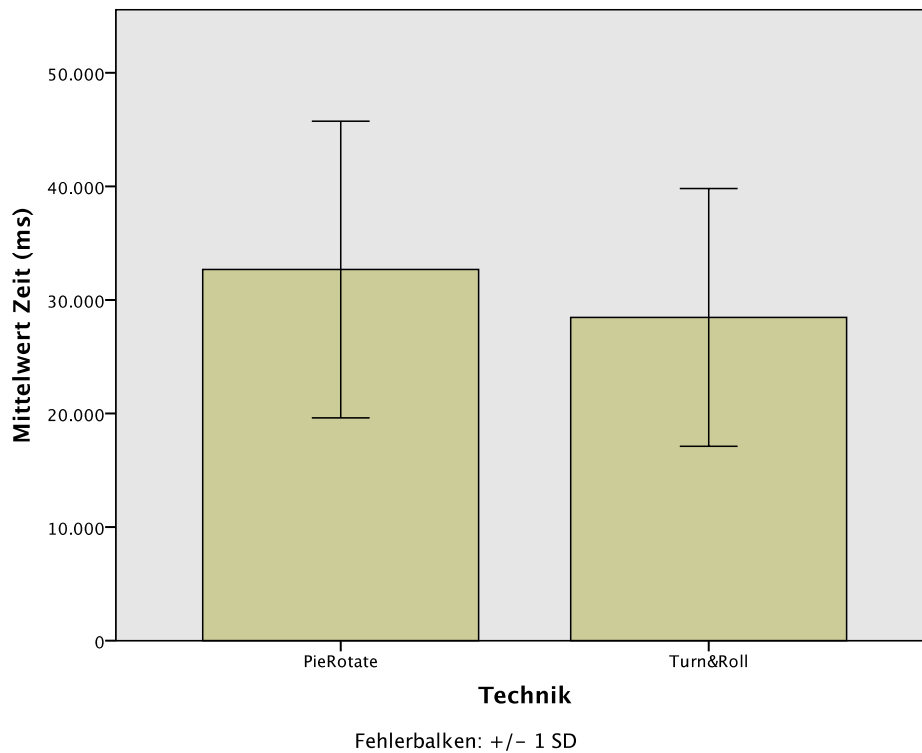


Abbildung 5.8: Mittelwerte der Abschlusszeiten der Aufgaben für beide Techniken

zeige signifikant höher waren als für Positionen in der unteren Anzeigehälfte ($t_{25} = 3,403$, $p = 0,002$) (vgl. Abbildung 5.9).

Manipulationsstrategien

Eine generelle Strategie der Teilnehmer bestand darin, zunächst die Orientierung des manipulierten Objektes grob an das Zielobjekt anzupassen, es dann über das Zielobjekt zu bewegen und schließlich feinere Anpassungen der Orientierung und der Position vorzunehmen.

Fünf Teilnehmer führten Rotationen und Translationen nur separiert aus, während drei Teilnehmer zu einem wesentlichen Teil auch die Möglichkeit zur integrierten Ausführung nutzten. Es konnte keine besondere Präferenz bezüglich der verwendeten Hand/Finger festgestellt werden. Drei Probanden führten Zweifingergesten beidhändig mit den Zeigefingern aus, drei mit Daumen und Zeigefinger oder Mittelfinger und Zeigefinger einer Hand. Bei zwei Teilnehmern war die beidhändige Interaktion beschränkt auf bestimmte Situationen.

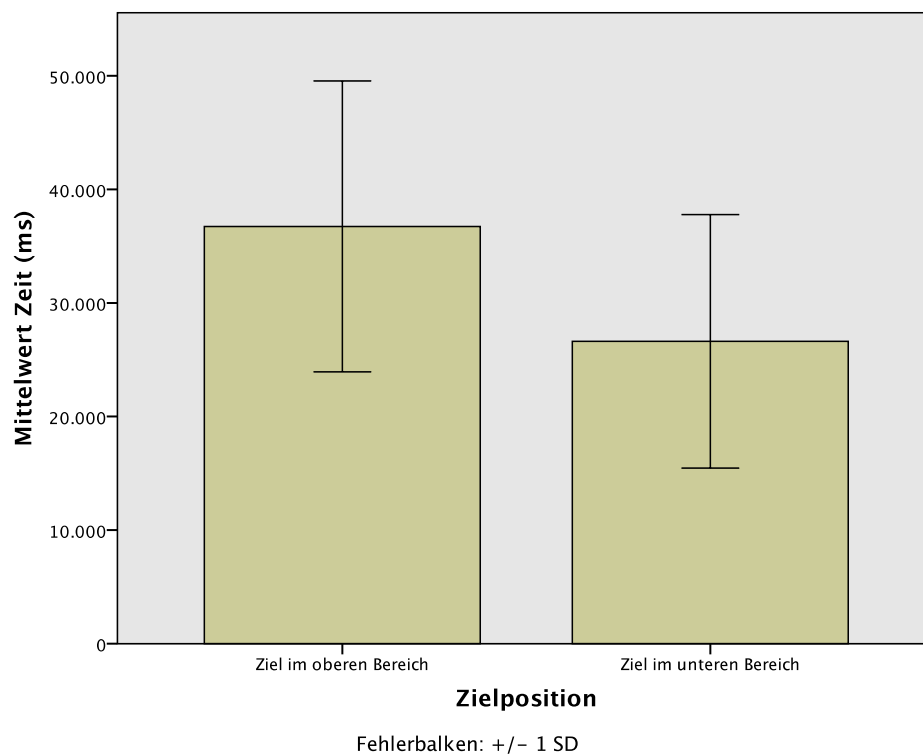


Abbildung 5.9: Mittelwerte der Abschlusszeiten der Aufgaben für unterschiedliche Zielpositionen für PieRotate

So nahm einer dieser Teilnehmer ausschließlich für Z-Pinch den Zeigefinger der anderen Hand zur Hilfe, während der andere Teilnehmer die zweite Hand einsetzte, um die Präzision² der Rotation oder Translation zu bestimmen.

Die meisten Probanden verwendeten die Schatten der Tetraeder auf der Bodenebene als wesentliche Hilfe für die Z-Translation. Oftmals war deutlich erkennbar, dass die Probanden lediglich auf die Schatten achteten und die Objekte gar nicht anschauten.

Im Allgemeinen bereitete die Verwendung der verschiedenen Kontrollzonen bei der PieRotate-Technik keine Probleme. Einige Probanden wechselten die Hand zwischen den oberen linken und rechten Bereichen. Ein Teilnehmer verstand die Kontrolltechnik nicht und konnte keine Einzelaufgabe erfolgreich abschließen. Nur ein Proband kombinierte die Ausführung von Rotation und

²Die Anzahl der eingesetzten Finger beeinflusst den Anteil der einzelnen Finger an der ausgeführten Operation und damit, wie groß der *Ausschlag* im Verhältnis zu der ausgeführten Bewegung ist.

Translation mit PieRotate.

Turn&Roll ermutigte drei Teilnehmer zur integrierten Ausführung von Translation und Rotation. Bei zweien dieser Teilnehmer konnte sogar die gleichzeitige Kontrolle von vier Freiheitsgraden (3D-Position und Z-Rotation) beobachtet werden. Zwei Teilnehmer verwendeten häufig einen schnellen Wechsel zwischen einem und zwei Fingern für die Rotation in drei Freiheitsgraden.

Qualitatives Feedback

Die qualitative Befragung ergab, dass alle Teilnehmer insgesamt die Turn&Roll-Technik bevorzugten. Als Hauptgrund wurde genannt, dass die Bedienung einfacher und intuitiver sei. Es wurde auch angegeben, dass es leichter sei, Fehler zu korrigieren und ein Teilnehmer lobte explizit den flüssigen Wechsel zwischen Rotation und Translation. Die Teilnehmer gaben an, dass es ihnen bei der PieRotate-Technik schwer gefallen sei, die Abbildung der Kontrollbereiche auf die Rotationsachsen im Gedächtnis zu behalten. Dagegen beurteilten viele Teilnehmer PieRotate als die präzisere der beiden Techniken. Einige waren der Meinung, dass sie zwar länger bräuchten, um PieRotate zu erlernen, diese am Ende aber die leistungsfähigere der beiden Techniken sei.

5.2.5 Diskussion

Die Abschlussquoten und -zeiten der Einzelaufgaben sowie das subjektive Feedback sprechen eindeutig für die Turn&Roll-Technik. Diese wurde als intuitiver und leichter zu erlernen bewertet. In einigen Fällen ermutigte diese Technik sogar zum schnellen Wechsel zwischen Ein- und Zweifingergesten. Dies konnte für PieRotate nicht beobachtet werden.

Dennoch hat auch PieRotate einige Vorteile. Diese Technik ist schwieriger zu erlernen, könnte allerdings für Experten eine präzisere Kontrolle ermöglichen. Hier stellt sich die Frage, ob *Walk-up-and-use* (Olsen, 2007) ein sinnvolles Kriterium für Experimente darstellt, welche komplexe Aufgaben wie 3D-Manipulation beinhalten, die immer eine gewisse Übung/Expertise voraussetzen. PieRotate würde sehr wahrscheinlich von visuellem Feedback bzgl. der Kontrollbereiche profitieren. Dies würde den *Widget*-Charakter der Technik betonen. Es wären weitere Untersuchungen notwendig, um eine endgültige

Gestaltungsdimension	Einordnung	Erklärung
Multi-point vs. Multi-touch	Multi-point	Berührungen werden als Punkte verarbeitet
Statischer vs. dynamischer Kontext	eher statisch	PieRotate: expliziter Moduswechsel über den Ort der Interaktion; Turn&Roll: expliziter Moduswechsel über die Anzahl der Finger
Integrierte vs. Separierte Freiheitsgrade	unterschiedlich	je nach Modus werden mehr oder weniger Freiheitsgrade gleichzeitig beeinflusst
Direkte vs. indirekte Interaktion	eher indirekt	indirekte Abbildung der Freiheitsgrade der Rotation, besonders im Fall vom PieRotate
Mehrhändige Interaktionsmuster	mehrhändig	keine bevorzugte Rollenverteilung der Hände

Tabelle 5.2: Einordnung der Techniken PieRotate und Turn&Roll gemäß den Gestaltungsdimensionen aus Kapitel 4

Aussage über den Vergleich einer solchen Technik mit Turn&Roll treffen zu können.

Sequenzielle (separierte) Manipulationsstrategien haben im Versuch klar dominiert, die parallele (integrierte) Ausführung von Operationen konnte aber ebenfalls mehrfach beobachtet werden. Insbesondere bei Turn&Roll hat dies möglicherweise zu den niedrigeren Abschlusszeiten beigetragen.

Insgesamt favorisieren die im Versuch erzielten Ergebnisse die Integration von Rotation und Translation. Zusätzliche Optimierungen könnten allerdings hilfreich sein, beispielsweise eine Filterung der Eingaben (vgl. Nacenta u. a. (2009)). Das allgemeine Problem der Dissonanz zwischen 3D-Raum und 2D-Kontrolle betrifft auch die vorgestellten Techniken in gewissem Maße. Die Bereitstellung zusätzlicher visueller Anhaltspunkte könnte hier Abhilfe schaffen.

5.2.6 Fazit

In diesem Abschnitt wurden zwei Techniken für die 3D-Manipulation von Objekten auf MT-Bildschirmen vorgestellt. Beide Techniken basieren auf der Abbildung von 2D-Transformation in den 3D-Raum. Dieser Ansatz macht die Verfahren robust und effizient berechenbar und erlaubt eine flexible Kontrolle

mit einer oder beiden Händen ohne eine feste Reihenfolge der Fingerberührungen vorzugeben. Die 3D-Manipulation von Objekten ist eine grundlegende Voraussetzung für die 3D-Modellierung auf interaktiven Oberflächen. Tabelle 5.2 fasst die Einordnung der vorgestellten Techniken gemäß der Gestaltungsdimensionen nach Kapitel 4 zusammen.

Kapitel 6

Modellierungsansätze

In etablierten Softwaretools werden viele unterschiedliche Ansätze und Techniken zur Modellierung von 3D-Objekten eingesetzt. Einige Techniken eignen sich besonders gut oder schlecht für bestimmte Klassen von Modellen, diese sind z.B. besser für technische oder künstliche Objekte als für „organische“ Objekte einsetzbar oder anders herum. Auch Präferenz und Vorkenntnisse des Benutzers können bei der Auswahl der Technik eine Rolle spielen. Im professionellen Bereich ist der Prozess der 3D-Modellierung zudem durch eine stark ausgeprägte Arbeitsteilung und Spezialisierung gekennzeichnet. Die einzelnen Arbeitsschritte unterscheiden sich dementsprechend ebenfalls durch die eingesetzten Werkzeuge.

Es stellt sich daher die Frage, welche Modellierungsansätze von Multi-touch (MT)-Eingaben und Techniken profitieren können. In diesem Kapitel werden zwei Ansätze und das dazugehörige Interaktionsdesign vorgestellt und diskutiert. Ein wichtiger Fokus ist dabei in beiden Fällen die beidhändige Interaktion zur Kamerakontrolle und Objektmanipulation.

Zunächst wird ein Ansatz auf Basis von impliziten Flächen (*Metaballs*) vorgestellt. Dann folgt die Diskussion von virtuellem oder digitalem Sculpting mit MT.

6.1 Implizite Flächen

Für das Feld der Computergrafik und der Forschung im Bereich Mensch-Computer-Interaktion ist die leichte und effektive Erstellung von 3D-Inhalten

ein wichtiger Anwendungsfall. Mit der fortschreitenden Durchdringung unseres Alltags durch digitale Medien ist die Nachfrage nach einfach zu benutzenden Werkzeugen für diesen Zweck über die letzte Dekade bedeutend gestiegen. 3D-Modelle werden heute in fast allen Arbeitsbereichen von der Industrie über die Forschung bis zur Unterhaltungsbranche eingesetzt. Vielfach stellt die Erstellung von 3D-Inhalten jedoch einen Engpass im Entwicklungsprozess von interessanten interaktiven grafischen Anwendungen dar. Die etablierten Softwarepakete im Bereich der Modellierung, z.B. Autodesk Maya oder 3ds Max, sind schwierig zu lernen und zu beherrschen. Oft ist viel Hintergrundwissen und ein tiefes Verständnis der zu Grunde liegenden mathematischen Modelle und Algorithmen notwendig, um mit diesen Werkzeugen umzugehen. Zusätzlich bereiten die Navigation und Orientierung innerhalb der virtuellen 3D-Umgebung vielen Benutzern Schwierigkeiten.

MT bietet im Vergleich zu etablierten Eingabegeräten wie Tastatur und Maus zusätzliche Freiheitsgrade, z.B. durch den Einsatz unterschiedlicher Fingergesten. Zusätzlich wird die beidhändige Interaktion ermöglicht, welche die Effizienz des Interfaces weiter erhöhen kann. Besonders für die 3D-Modellierung ist beidhändige Interaktion eine Möglichkeit mit großem Potenzial. Die gleichzeitige Kontrolle von Kamera und Objekten könnte zu einer besseren Wahrnehmung von räumlichen Anordnungen im virtuellen 3D-Raum beitragen, da diese schnelle kleine Änderungen der Perspektive ermöglicht. Die lineare direkte Abbildung zwischen den Bewegungen des Benutzers auf dem MT-Bildschirm und den Objekten in der virtuellen Szene verbessert potenziell ebenfalls das Gefühl für die Proportionen und räumliche Anordnungen durch die verbesserte Nutzung des „Muskelgedächtnisses“ (vgl. Kapitel4).

Im Allgemeinen haben Menschen viel Erfahrung damit, Objekte in der realen Welt zu bewegen, auszurichten und anzudocken. In der realen, physischen Welt „greift man“ einfach ein Objekt und bewegt es an die gewünschte Stelle – 3D-Modellierung auf MT-Eingabegeräten sollte diese Möglichkeit ebenfalls bieten. Die Funktionalität des Interfaces sollte jedoch nicht übermäßig beschränkt werden, sofern diese im Einklang mit dem genannten Grundprinzip steht. Daher sollte die Schnittstelle auch erweiterte Funktionalität bieten, die nicht direkt an die reale Welt angelehnt ist, aber gut für MT funktioniert. Die im nächsten Abschnitt beschriebenen relativen Abbildungen stehen nicht

in direkter Beziehung zu Abläufen in der realen Welt sondern erweitern die absoluten Abbildungen, welche aus der realen Welt bekannt sind (vgl. *direkte vs. indirekte Interaktion* aus Kapitel 4).

Eine Voraussetzung, um 3D-Modellierung auf MT-Interfaces zugänglich zu machen, sind geeignete Modellierungsmetaphern, welche gut zu der Art der Interaktion passen. Physikalisch-inspirierte Metaphern, z.B. angelehnt an das Arbeiten mit Ton oder Knete, sind gute Kandidaten hierfür. In diesem Abschnitt wird die Nutzung von impliziten Flächen, speziell von *Metaballs* nach Blinn (1982), als Metapher für die 3D-Modellierung auf MT-Displays näher untersucht. Eine weitere Voraussetzung sind dazu passende Gesten und Abbildungen der Nutzereingaben auf Operationen im virtuellen 2D- bzw. 3D-Raum. Daher wird in diesem Abschnitt ebenfalls die Kombination mit einfachen einhändigen und beidhändigen Mehrfingergeräten untersucht.

MT-Interaktion ist nicht auf einen einzigen Fokuspunkt für die Eingabe beschränkt. Diese Eigenschaft kann genutzt werden, um den Anteil der sequentiellen Ausführung von Aufgaben zu reduzieren. In etablierten 3D-Modellierungswerkzeugen muss der Benutzer oft sich wiederholende und zeitraubende Serien von Operationen durchführen, z.B. das Objekt ein kleines Stück bewegen, die Kamera ausrichten, das Objekt etwas weiter bewegen, die Kamera wieder ausrichten usw. Der vorgestellte Ansatz erlaubt es dem Benutzer, beide Operationen gleichzeitig oder zumindest in schnellerem Wechsel durch den Einsatz beider Hände auszuführen. Die parallele¹ Natur von MT-Interaktion erfordert außerdem ein geeignetes Anwendungsframework und eine andere Behandlung von Eingabeevents und der Abstimmung zwischen verschiedenen Operationen, als dies bei Maus und Tastatur der Fall ist. Im folgenden Abschnitt werden notwendige Modifikationen an der Eventverarbeitung etablierter Modellierungswerkzeuge am Beispiel des Open Source Tools Blender² vorgestellt.

Dieser Abschnitt trägt in verschiedener Hinsicht zum Feld der Interaktionsforschung bzw. der Computergrafik bei. Existierende allgemeine Ansätze für die beidhändige Interaktion, speziell Arbeiten von Balakrishnan und Kurtinbach (1999) und das Framework von Guiard (1987), werden auf MT-Interfaces mit dem Anwendungsfall 3D-Modellierung übertragen und für diesen speziel-

¹Gemeint ist die gleichzeitige Interaktion an verschiedenen Stellen des Bildschirms.

²<http://blender.org>

len Fall evaluiert. Es werden Abbildungen für die einhändige und beidhändige Objekt- und Kameramanipulation entwickelt und Metaball-Modellierung als geeignete Metapher untersucht. Außerdem werden die Detektion und Integration von Mehrfingergesten in eine existierende komplexe Anwendung diskutiert. Abschließend werden die präsentierten Werkzeuge in einer ersten Nutzerstudie getestet und bewertet.

6.1.1 Stand der Technik

Etablierte 3D-Bearbeitungssoftware wird hauptsächlich einhändig bedient. Die dominante Hand (DH) kontrolliert die Maus oder einen Stift und die schwache Hand (NDH) kommt eher selten zum Einsatz, um etwa mittels Tastatur zu einem anderen Werkzeug zu wechseln. Für viele Aufgaben in der realen Welt setzen Menschen dagegen beide Hände in asymmetrischer Aufgabenteilung ein (vgl. Kapitel 4; Guiard (1987)). Motiviert durch Studien über die Vorteile beidhändiger Interaktion, verwenden Zeleznik u. a. (1997) zwei Mauszeiger, um Operationen in einer 3D-Desktopapplikation auszuführen und berichten, dass Techniken auf Basis physikalischer Intuition besser in ihren Untersuchungen abschneiden. Die Responsive Workbench von Cutler u. a. (1997) ist ein immersives Virtual Reality (VR)-Tabletop-System und erlaubt die beidhändige, physikbasierte 3D-Interaktion mit einer stereoskopischen Darstellung. Ein Nachteil dieses Systems ist jedoch der hohe Grad der Instrumentarisierung des Benutzers. Dieser muss einen speziellen Head Tracker und eine Stereobrille, ein spezielles 6DOF Eingabegerät sowie VR-Handschuhe verwenden. Balakrishnan und Kurtenbach (1999) untersuchen die beidhändige asymmetrische Interaktion (mit zwei Mäusen) für die Kontrolle der Kamera mit der NDH und Objektmanipulation mit der DH. Ihre Studie ergibt, dass die beidhändige Kontrolle für 3D-Selektionsaufgaben 20% schneller als die sequenzielle einhändige Kontrolle ist. Forlines u. a. (2007) vergleichen direkte Touch-Eingabe mit indirekter Mausinteraktion und stellen fest, dass Touch, auf Grund der kognitiven Anstrengung, gleichzeitig zwei Zeiger indirekt kontrollieren zu müssen, besser für die beidhändige Interaktion geeignet ist. Untersuchungen von Terrenghi u. a. (2007) bzgl. der Unterschiede zwischen beidhändiger Interaktion mit realen und virtuellen Objekten ergeben, dass in der virtuellen Umgebung die einhändige Interaktion oftmals überwiegt.

Eine Reihe von Arbeiten hat sich mit verschiedenen Ansätzen beschäftigt, um die Freiheitsgrade von MT-Input auf interaktiven Oberflächen effektiv für Softwareanwendungen zu nutzen. Wu u. a. (2006) schlagen allgemeine Gestaltungsprinzipien für Gesten auf interaktiven Oberflächen vor: Gestenerkennung (registration), Lockerung der Kriterien der Erkennung (relaxation) und Wiederverwendung von Gesten (reuse). Die RoomPlanner-Anwendung von Wu und Balakrishnan (2003) verwendet einhändige und beidhändige Gesten zur 2D-Objektmanipulation, z.B. wird ein Finger zur Selektion und Translation und zwei Finger zum Skalieren und Rotieren von Objekten eingesetzt. Benko u. a. (2006) diskutieren verschiedene beidhändige Techniken zur pixelgenauen Selektion, beispielsweise die Verwendung jeweils eines Fingers jeder Hand und die Nutzung des Mittelpunktes zwischen diesen beiden Fingern als Referenzpunkt für die Selektion. Mit geeigneten Approximationsverfahren ermitteln Moscovich und Hughes (2006) affine 2D-Transformationen aus den Fingerkontakten und untersuchen verschiedene Abbildungen der Mehrfingereingaben für die 2D-Objektmanipulation. Eine weitere Möglichkeit ist die physik-basierte Objektmanipulation. Touch-Eingaben werden auf geeignete Weise in eine Physikengine eingebunden (Wilson u. a., 2008), um das „Greifen“ von Objekten mittels virtueller Kräfte zu simulieren (Wilson, 2009).

Viele Arbeiten versuchen die Vorteile der direkten Interaktion und der absoluten Abbildung zu nutzen, die sich aus der Übereinstimmung von Eingabe- und Ausgaberaum auf interaktiven Oberflächen ergibt. Dies ist jedoch nicht immer wünschenswert. Begrenzte Reichweite (ein zu kleiner Eingaberaum), begrenzte Präzision (Eingaberaum zu grob) oder die Verdeckung von Objekten können Gründe sein, die indirekte Abbildungen notwendig und interessant machen (Forlines u. a., 2006; Moscovich und Hughes, 2008).

Im Bereich der 3D-MT-Interaktionstechniken existieren im Vergleich zu den 2D-Techniken bisher weniger Forschungsarbeiten. Hancock u. a. (2007) erweitern ihre integrierten 2D-Rotations- und -Translationskontrollen auf die dritte Dimension allerdings mit stark begrenzten Interaktionsmöglichkeiten bzgl. der Tiefendimension. Dagegen präsentieren Reisman u. a. (2009) vollwertige 3D-Transformationen, deren Berechnung allerdings numerisch relativ aufwändig und nicht unproblematisch ist (vgl. Kapitel 5). Bisher wurde keiner der Ansätze in einer funktionierenden 3D-Modellierungsumgebung getestet. In zwei

weiteren Arbeiten untersuchen Hancock u.a. die 2D- und 3D-Interaktion auf interaktiven Oberflächen im Hinblick auf Vor- und Nachteile von *Tangible Interaction* gegenüber MT für die Kamerakontrolle (Hancock u. a., 2009b) und bzgl. des Einflusses der verwendeten Projektion auf die Wahrnehmung von räumlichen Anordnungen von Objekten (Hancock u. a., 2009c). Sie berichten, dass die Benutzer in ihren Untersuchungen *Tangibles* für die Kamerakontrolle und Touch für die Objektmanipulation präferierten.

Für weitere grundlegende Arbeiten zu Interaktionstechniken für interaktive Oberflächen wird der Leser auf Kapitel 4 verwiesen.

6.1.2 Metaball-Metapher

Aktuelle 3D-Modellierungssoftwarepakete stellen verschiedene Werkzeuge basierend auf verschiedenen Oberflächenrepräsentationen zur Verfügung, aus denen der Benutzer wählen kann (vgl. Kapitel 2). Unter diesen Verfahren erscheinen implizite Oberflächen – speziell *Metaballs* oder *Blobbies* (Blinn, 1982) – aus den folgenden Gründen als geeignet für die Modellierung auf einem MT-Bildschirm:

- Die Modellierungsmetapher ist verglichen mit den Werkzeugen³ für polygonale Netze (Meshes) sehr einfach zu verstehen.
- Die grundlegende Modellierung benötigt fast ausschließlich die Translationsoperation.
- Modellierung mit Metaballs weist Ähnlichkeiten zu der Arbeit mit Ton oder Knete auf und passt daher zu dem physikalisch-inspirierten Interaktionscharakter von MT.
- Metaballs können nach Abschluss der Modellierung leicht in eine polygonale Darstellung überführt werden.

Ein Metaball ist durch seine Position und einen positiven oder negativen numerischen Wert definiert (manchmal „Steifheit“ genannt). Eine implizite

³Werkzeuge zum Entfernen, Hinzufügen, Aufspalten usw. von Polygonen oder Polygonringen, welche in allen etablierten Softwaretools zur Modellierung mit Meshes zur Verfügung stehen, setzen Kenntnisse über Topologie und Aufbau des Meshes für den Einsatz der Werkzeuge voraus.

Funktion definiert die Abnahme (oder Zunahme) der Steifheit in Abhängigkeit von der Position bzw. dem Ursprung des Metaballs. Für jeden Punkt im 3D Raum kann der Einfluss jedes Metaballs in der Szene addiert und so jedem Punkt der resultierende Wert zugewiesen werden. Auf diese Weise definiert die Gesamtheit der Metaballs in einer Szene ein Skalarfeld. Dieses Skalarfeld wird ausgewertet und mit einem festgelegten Schwellwert verglichen, um eine Isofläche zu extrahieren. Mit verschiedenen Methoden kann die resultierende Isofläche schließlich gerendert werden. Ein häufiger Ansatz ist, diese zu polygonisieren (Bloomenthal, 1988). Dabei wird die Fläche in Echtzeit in ein polygonales Mesh konvertiert, welches dann von der Standard Renderingpipeline verarbeitet werden kann. Blender stellt diese Funktionalität bereits zur Verfügung und verwendet eine leicht modifizierte Variante der Octree-basierten Polygonisierungstechnik von Bloomenthal (Bloomenthal, 1994).

Obwohl die Theorie hinter der Metaballmetapher durchaus komplex ist, fällt die Anwendung auch unerfahrenen Benutzern leicht. Im Vergleich zu polygonalen Meshes benötigt Metaball-Modellierung nur eine minimale Menge an Operationen, um das „sinnvolle“ Modellieren zu ermöglichen. Ist eine Auswahl von Metaballs mit verschiedenen Eigenschaften gegeben (ähnlich einer „Malerpalette“), werden nur Operationen zur Vervielfältigung und Translation benötigt, um einfache Modelle zu kreieren. Metaballs haben nicht alle Eigenschaften von echten Virtual Clay-Ansätzen, aber die Art wie diese scheinbar miteinander „verschmelzen“ weist einige Ähnlichkeiten mit Ton oder Knete auf und unterstützt die Art der physikalisch-inspirierten Interaktion. Schließlich können Metaballs leicht in polygonale Meshes konvertiert werden und ermöglichen so die Anwendung aller Standardwerkzeuge für die Modellierung mit Meshes. Die Abbildungen 6.1 und 6.2 geben einen Eindruck von der Modellierung mit Metaballs mit dem vorgestellten System.

6.1.3 Interaktionsdesign und Gesten

Die grundlegendsten Operationen in einer 3D-Modellierungsanwendung sind die Selektion von Objekten, die Ausführung von Transformationen – wie Translation, Rotation und Skalierung – sowie die Kontrolle der virtuellen Kamera – Schwenken, Rotieren und Zoomen. Zusätzlich ist es erforderlich, Objekte zu erstellen, zu löschen, zu gruppieren usw.

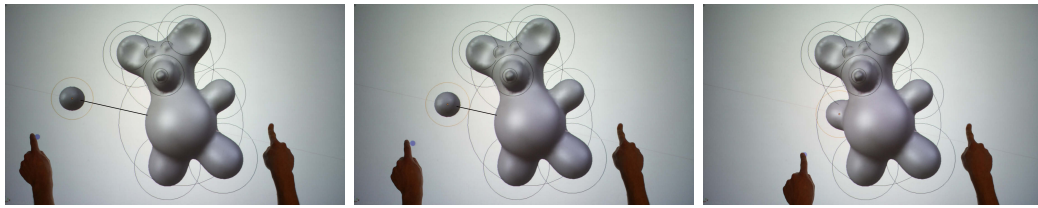


Abbildung 6.1: Indirekte beidhändige Objektmanipulation mit *auto-constraints* (die Constraint-Achse wurde für die Abbildung farblich hervorgehoben)

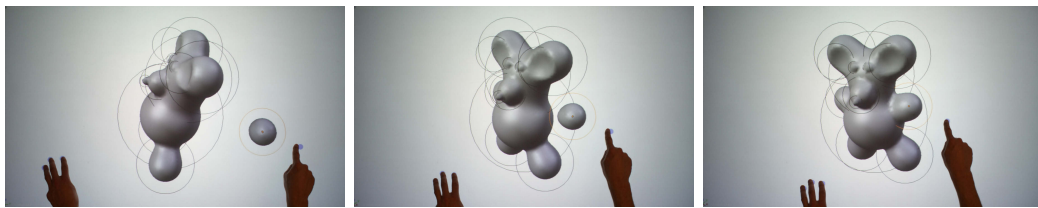


Abbildung 6.2: Indirekte gleichzeitige beidhändige Kamerakontrolle und Objektmanipulation mit *object pinning*

Im Fokus der folgenden Untersuchung stehen die grundlegenden Operationen zur Manipulation von Objekten und zur Navigation in der 3D-Szene. Diese werden mit spezifischen MT-Techniken erweitert. Um die Machbarkeit dieser erweiterten MT-Interfacekonzepte für den speziellen Anwendungsfall der 3D-Modellierung zu untersuchen, ist es aber notwendig, dass die Testumgebung die weiteren notwendigen Werkzeuge bereitstellt (Ansatz der iterativen Weiterentwicklung im Gegensatz zu einem vollständigen Paradigmenwechsel – vgl. Kapitel 4). Die übrigen vorhandenen Werkzeuge müssen daher – zumindest zu einem gewissen Grad – mit MT benutzbar gemacht werden. Es wird daher ein Ansatz gewählt (Mehrfingertechniken), der diese beiden Anforderungen in einem gemeinsamen Konzept zu erfüllen verspricht. Die hierfür notwendigen Modifikationen an der Eventverarbeitung der Softwareumgebung (hier: Blender), um eine Emulation von Maus- und Tastatureingaben mittels MT zu ermöglichen, werden in Abschnitt 6.1.4 beschrieben.

In den folgenden Abschnitten werden die MT-Techniken, insbesondere der Bezug zu mehrhändigen Interaktionsmustern, im Detail vorgestellt.

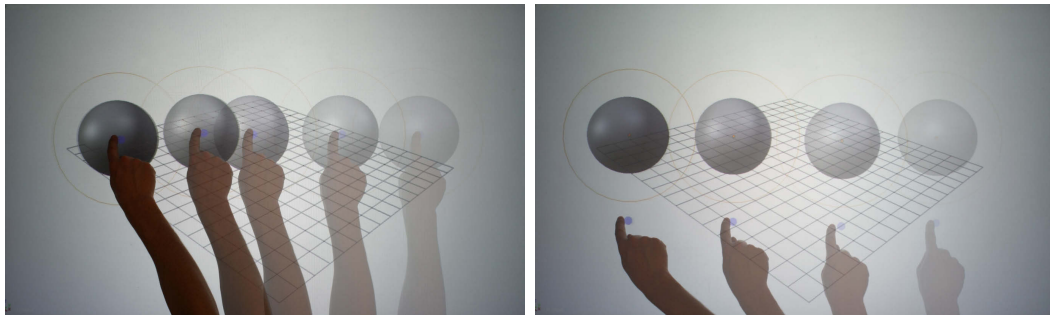


Abbildung 6.3: Links: Direkte Selektion und Manipulation; Rechts: Indirekte Manipulation eines selektierten Objektes

Einhändige Objektmanipulation

Damit die schwache Hand (NDH) für die Kontrolle der Kamera (oder eines zusätzlichen Objektes) zur Verfügung steht, muss der Benutzer Kernaufgaben wie die Selektion oder Translation mit einer Hand ausführen können. Außerdem gehören die Selektion und Manipulation zu den am häufigsten ausgeführten Operationen bei der 3D-Modellierung. Das Interaktionsdesign sieht daher die einfachste Geste für diese Aufgaben vor: das *tapping* bzw. *dragging* mit einem Finger.

Eine direkte Translation kann ausgeführt werden, indem ein Objekt berührt und der Finger, ohne abzusetzen, bewegt wird (vgl. Abb. 6.3, links). Alternativ kann ein Objekt durch eine kurze Berührung (tap) zunächst selektiert werden. Ist ein Objekt selektiert, kann durch das Bewegen eines Fingers an beliebiger Stelle auf der Oberfläche eine indirekte Translation ausgeführt werden (vgl. Abb. 6.3, rechts). Dabei wird das Objekt relativ zu der Fingerberührung bewegt. Dies ist besonders hilfreich für eine präzise Kontrolle sehr kleiner Objekte oder falls sich viele weitere Objekte in der Nähe des Zielobjektes befinden. Die Selektion eines beliebigen anderen Objektes deselektiert das aktuelle Objekt. Eine explizite Deselektion von Objekten ist nicht vorgesehen. Implementiert ist die Selektion von Objekten mittels *Ray Casting* an der Stelle der Fingerberührung. Tabelle 6.1 fasst die Gesten für einhändige Objektmanipulation nochmals zusammen.

Geste	Operation
„Tap“ (ein Finger)	Selektion
„Touch + move“ (ein Finger)	Selektion + direkte Translation
„Move“ (ein Finger)	Indirekte Translation

Tabelle 6.1: Einhändige Objektselektion und -manipulation

Einhändige Kamerakontrolle

Da der Sichtbereich der Kamera bei der 3D-Modellierung häufig verändert werden muss, sollten die hierfür verwendeten Gesten so einfach und robust zu erkennen sein wie möglich. Zum Zeitpunkt der in diesem Abschnitt durchgeführten Experimente war eine Handunterscheidung nicht möglich (vgl. Kapitel 3). Aus diesem Grund müssen die verwendeten Gesten eindeutig sein, so dass die Zuordnung auch bei beidhändiger Bedienung ohne Probleme funktioniert. Insbesondere darf es keinen Konflikt mit den für die Objektmanipulation verwendeten Gesten geben, welche in der Regel mit der starken Hand ausgeführt werden. Um die direkte Interaktion zu unterstützen, soll auf spezielle Bereiche und zusätzliche visuelle Elemente verzichtet werden.

Für die Kamerakontrolle erscheinen daher Mehrfingergersten als gut geeignet. Eine Unterscheidung der auszuführenden Operationen findet dabei über die Anzahl der detektierten Fingerberührungen statt. Um die gleichzeitige Ausführung mehrerer Gesten/Operationen und die Interaktion mit beiden Händen zu ermöglichen, werden nur die Berührungen zusammen gezählt, die innerhalb einer begrenzten Zeitspanne und innerhalb eines festgelegten Radius erkannt werden. Als Referenzpunkt für die Interaktion wird der geometrische Mittelpunkt aller zu der Geste gehörenden Fingerberührungen verwendet. Weitere Details zu dem verwendeten Algorithmus finden sich in Abschnitt 6.1.4.

Es existiert keine kanonische optimale Abbildung zwischen Mehrfingergersten und Kamerakontrollfunktionen. Es finden sich Argumente für verschiedene Abbildungen und individuelle Präferenzen sind zu erwarten (Wobbrock u. a., 2009). In einer Produktionsumgebung sollten diese Einstellungen daher konfigurierbar sein.

Ein wichtiges Maß könnte die Häufigkeit der Verwendung bestimmter Funktionen darstellen und die häufiger verwendeten Funktionen könnten dement-

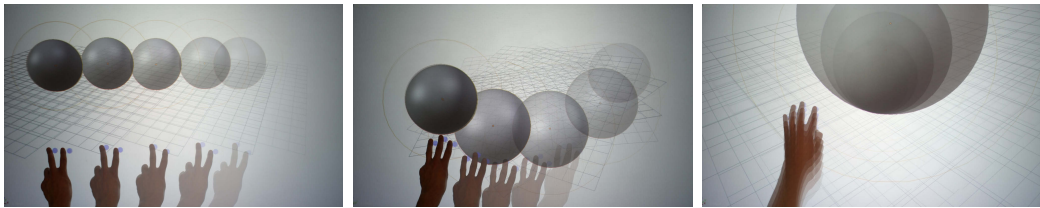


Abbildung 6.4: Links: Kamerashwenk (zwei Finger) Mitte: Kamerarotation (drei Finger) Rechts: Kamerazoom (vier Finger)

sprechend auf Gesten abgebildet werden, die einfacher auszuführen sind, d.h. z.B. weniger Finger benötigen (dieser Ansatz wurde in den Untersuchungen zu Virtual Sculpting in Abschnitt 6.2 verfolgt). Im Hinblick auf die Lernkurve des Systems darf jedoch auch das Vorwissen der Benutzer nicht außer Acht gelassen werden. Auf einigen Plattformen existieren z.B. etablierte Mehrfingergesten für das Scrollen von Fenstern. Da beispielsweise das Scrollen von Fenstern dem Kamerashwenk sehr ähnlich ist, könnte es sinnvoll sein, die gleiche Geste zu verwenden. Die einfachen Mehrfingergesten könnten potenziell auch mit komplexeren Gesten kombiniert werden, z.B. könnte die relative Position der Finger zur Bestimmung der Rotationsachse oder anderer Eigenschaften verwendet werden. Entsprechend müsste die verwendete Geste genug Freiheitsgrade für die auszuführende Funktionalität bieten.

Die Erfahrung hat gezeigt, dass das Zoomen der Kamera im Vergleich zu den anderen Operationen bei der Arbeit mit Metaballs selten verwendet wird. Daher wird diese Funktion auf die Vier-Finger-Geste abgebildet. Die Kamera kann gezoomt werden, indem vier Finger vertikal hoch oder runter bewegt werden (vgl. Abb. 6.4, rechts). Die Zoom-Funktionalität ist implementiert als sog. *dolly move*, d.h. die virtuelle Kamera wird bewegt, die Parameter der Abbildung (die virtuelle Linse) bleiben unverändert. Zwei Finger werden für den Schwenk (vgl. Abb. 6.4, links) und drei Finger für die Kamerarotation verwendet (vgl. Abb. 6.4, mitte). Für die Rotation wird die *Turntable*-Metapher verwendet, d.h. die XY-Ebene kann um den Ursprung rotiert und gekippt werden. Tabelle 6.2 fasst die verwendeten Gesten nochmals zusammen.

Geste	Operation
„Move“ (zwei Finger)	Kameraschwenk („pan view plane“)
„Move“ (drei Finger)	Kamerarotation („turntable metaphor“)
„Move“ (vier Finger)	Kamerazoom („dolly move“)

Tabelle 6.2: Gesten für die einhändige Kamerakontrolle

Beidhändige Kamerakontrolle und Objektmanipulation

Die zuvor beschriebenen einhändigen Gesten für die Objektmanipulation bzw. Kamerakontrolle sind im Hinblick auf die gleichzeitige beidhändige Nutzung gestaltet. Basierend auf dem Framework von Guiard (1987) und der Arbeit von Balakrishnan und Kurtenbach (1999) erlaubt dies dem Benutzer, die dominante Hand (DH) zur Objektmanipulation und die nicht-dominante Hand (NDH) zur Kontrolle der Kamera einzusetzen.

Eines der Grundprinzipien von asymmetrischer beidhändiger Interaktion ist nach Guiard, dass die NDH für die Aktionen/Bewegungen der DH einen Referenzrahmen bereitstellt (vgl. Kapitel 4). Das entwickelte Interaktionsdesign baut auf diesem Prinzip auf, indem die NDH den Sichtbereich und damit den räumlichen Bezugsrahmen für die Objektmanipulation mit der DH kontrolliert.

Die Bewegung der Kamera vermittelt einen wichtigen visuellen Hinweis für die Tiefenwahrnehmung, den sog. „kinetic depth effect“. Die gleichzeitige Kontrolle der Kamera kann diesen Effekt unterstützen, da diese kontinuierliche kleine Änderungen des Sichtbereiches bei gleichzeitiger Bewegung der Objekte ermöglicht (vgl. Balakrishnan und Kurtenbach (1999)).

Die beidhändige Interaktion auf einem MT-Display birgt einige zusätzliche Herausforderungen verglichen mit dem Einsatz relativer Eingabegeräte, etwa von zwei Mäusen wie bei Balakrishnan und Kurtenbach (1999). Eine Anpassung des Sichtbereiches verändert im Allgemeinen auch die Bildschirmposition der dargestellten Objekte. Auf Grund der absoluten Abbildung der Fingerberührungen würde dies bedeuten, dass ein gleichzeitig manipuliertes Objekt „unter den Fingern“ des Benutzers „weggezogen“ wird – ein verwirrender Effekt.

Um diesen Effekt zu beseitigen, wird eine Technik implementiert, die im Kontext dieser Arbeit als *object pinning* bezeichnet wird (vgl. Abb. 6.5). Diese funktioniert wie folgt: die Position eines Objektes in Bildschirmkoordinaten

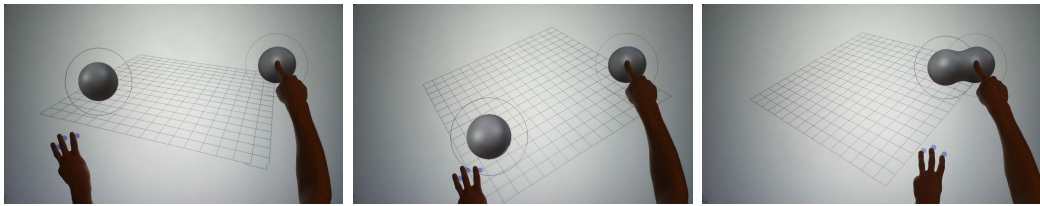


Abbildung 6.5: Direkte beidhändige Kamerakontrolle und Objektmanipulation (*object pinning*)

vor und nach einer Änderung des Sichtbereiches wird in Weltkoordinaten zurückgerechnet. Mit Hilfe dieser Informationen wird die Position des Objektes in Weltkoordinaten so angepasst, dass sich die neue Bildschirmposition nach Anpassung des Sichtbereiches in konstantem Abstand zu der jeweiligen Fingerberührung befindet. „Object pinning“ funktioniert sowohl für direkte als auch für indirekte Manipulation. Obwohl mathematisch korrekt, würde die so beschriebene Anpassung der Position zu sprunghaften Veränderungen des Abstandes zwischen Objekt und virtueller Kamera führen. Daher wird als weitere Bedingung eingeführt, dass dieser Abstand konstant bleiben muss.

Object pinning beseitigt nicht nur die Nebeneffekte der gleichzeitigen Kamerakontrolle, sondern erlaubt auch das Stückweise „herantasten“ an die Zielposition beim Platzieren von Objekten (vgl. Abb. 6.2). Diese Technik ermöglicht auch die indirekte Ausführung von Transformationen. Durch das „pinning“ eines bestimmten Objektes kann dieses indirekt durch das Verändern der Kameraperspektive transformiert werden.

Object pinning funktioniert für das Schwenken und Rotieren der Kamera. In Kombination mit dem Kameraschwenk kann es z.B. verwendet werden, um ein spezifisches Objekt „festzuhalten“ und die übrigen Objekte relativ zu diesem Objekt zu verschieben.

Beidhändige Objektmanipulation

Obwohl die beidhändige Kontrolle der Kamera wie oben dargelegt die Wahrnehmung der räumlichen Anordnungen in einer virtuellen 3D-Szene verbessern kann, ist es in bestimmten Situationen dennoch schwierig, Objekte präzise zu platzieren. Insbesondere können Objekte in der 2D-Darstellung nahe beieinander erscheinen, obwohl diese im 3D-Raum weit auseinander liegen.

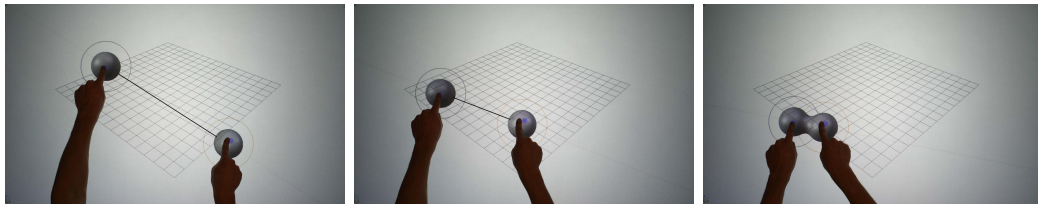


Abbildung 6.6: Direkte beidhändige Objektmanipulation mit *auto-constraints* (für die Abbildung farblich zusätzlich hervorgehoben)

Um die Benutzer bei der präzisen Platzierung von Objekten zu unterstützen, wird eine Technik implementiert, die im Rahmen dieser Arbeit als *auto-constraints* bezeichnet wird. Auto-constraints machen Gebrauch von beidhändig symmetrischer Interaktion (beide Hände werden zur Objektmanipulation eingesetzt). Die Technik funktioniert folgendermaßen: Der Benutzer selektiert ein Objekt zur Manipulation. Dieses Objekt dient als „Anker“ und kann frei bewegt werden. Selektiert der Benutzer gleichzeitig ein weiteres Objekt mit der anderen Hand, so ist die Bewegung des zweiten Objektes automatisch auf die Achse zwischen den beiden Objekten beschränkt (vgl. Abb. 6.1 und Abb. 6.6). Auf diese Weise kann das zweite Objekt präzise auf das erste Objekt zu bewegt werden. Außerdem kann diese Technik eingesetzt werden, um symmetrische Objektteile (eine häufige Operation bei der Modellierung) zueinander korrekt auszurichten.

In der derzeitigen Version wird immer die Verbindungsachse zwischen den Mittelpunkten beider Objekte für die Einschränkung der Bewegung verwendet. Andere Möglichkeiten wären denkbar. Beispielsweise könnte die Ebene, welche beide Objekte enthält und möglichst parallel zur Projektionsebene liegt eine gute Wahl sein. Eine weitere Möglichkeit wäre eine flexibel gewichtete Beschränkung der Bewegung bzgl. der verbindenden Achse, um eine Bewegung in diese Richtung zu erleichtern, aber nicht zu erzwingen.

6.1.4 System und Softwareintegration

Die Open Source 3D-Modellierungssoftware Blender dient als Basis für die Entwicklung und Untersuchung der beschriebenen Techniken. Dies ermöglicht es, auf einen vollständigen Satz an Werkzeugen zurückzugreifen, macht allerdings auch Modifikationen des internen Eventsystems von Blender notwendig.

Um existierende Werkzeuge mit MT benutzbar zu machen und die erweiterten Techniken zu implementieren, wird eine Emulation von Mausevents für die existierenden Werkzeuge integriert und zusätzlich als Grundlage für die neuen Techniken die Möglichkeit geschaffen, direkt auf MT-Events zuzugreifen. Zusätzliche Erweiterungen ermöglichen die parallele Ausführung von Operationen an verschiedenen Stellen und in verschiedenen Bereichen der Anwendung.

Gestenerkennung

Die Erkennung („gesture registration“) der Mehrfingergersten basiert auf zeitlichen und räumlichen Schwellwerten, um die Fingerberührungen zu gruppieren. Wird eine neue Berührung innerhalb eines begrenzten Bereiches um den geometrischen Mittelpunkt einer existierenden Gruppe von Berührungen detektiert, wird diese der Gruppe hinzugefügt und der Mittelpunkt der Gruppe entsprechend aktualisiert. Der Typ der Gruppe entspricht der Anzahl der Berührungen und kann zur Abbildung auf verschiedene Funktionen verwendet werden. Bis zu vier Berührungen können zur einer Gruppe zusammengefasst werden. Hat eine Gruppe für mindestens 120 ms bestand, wird diese als ein virtueller Touch per Event an die Anwendung weitergeleitet, wobei die Anwendung nach Bedarf aus dem virtuellen Touch auch die einzelnen Berührungen extrahieren kann. Jede Gruppe/jeder virtuelle Touch besitzt eine eindeutige ID, welche über die gesamte Lebensdauer der Gruppe erhalten bleibt. Jedes Event mit Bezug zu dieser Gruppe ist ebenfalls mit der gleichen ID markiert. Wird nach der erstmaligen Erkennung ein Finger „verloren“ oder kommt hinzu, ändert dies die Geste nicht mehr. Die Erkennungskriterien werden also gelockert („gesture relaxation“), um diese robuster gegenüber Unterbrechungen in der Erkennung oder unbeabsichtigtem Fingerauf-/absetzen des Benutzers zu machen. Die Gruppe bleibt erhalten bis der letzte Finger entfernt wird. Einzelne Berührungen werden ebenfalls mit einem Zeitschwellwert (80 ms) gefiltert bevor diese einer Gruppe hinzugefügt werden, um die Stabilität der Erkennung weiter zu erhöhen.

Eingabeverarbeitung für mehrere Cursor

Beidhändige MT-Interaktion setzt voraus, dass der Benutzer in unterschiedlichen Bereichen der graphischen Oberfläche gleichzeitig interagieren und dort

verschiedene Operationen ausführen kann. Daher sind Modifikationen an der Eingabeeventverarbeitung von Blender notwendig, um die vorhandenen modalen Dialoge und Operationen MT-fähig zu machen. Jede Eingabe wird auf einen *Operator* abgebildet. Operatoren sind in Blender verantwortlich für die Ausführung von Aktionen auf der Szene. Insbesondere die Kamerakontrolle und die Ausführung von Transformationen sind in Blender als modale Operatoren ausgelegt, welche jeweils exklusiv ausgeführt werden. Um dieses Verhalten zu ändern und die parallele Ausführung dieser Operatoren zu ermöglichen, wird die oben beschriebene Event-ID verwendet, um die Operatoren zu markieren. Auf diese Weise werden jedem Operator nur die zugehörigen Events zur Verarbeitung zugeordnet und diese können andere Operatoren nicht beeinflussen.

6.1.5 Experiment

Ziel der Untersuchung war, die grundsätzliche Herangehensweise der Probanden bei der Benutzung des Systems zu studieren. Dabei lag das Augenmerk auf der prinzipiellen Benutzbarkeit der Gesten und Funktionen und auf den eventuell zu beobachtenden Interaktionsmustern. Hier interessierte insbesondere beidhändige Interaktion und die Rollenaufteilung der Hände. Ziel war es außerdem, einen Eindruck davon zu bekommen, ob das System den Benutzern intuitiv zugänglich erscheint.

Versuchsaufbau und -ablauf

Für alle im Folgenden dargestellten Untersuchungen wurde als Hardwareplattform ein auf der Diffuse Illumination (DI) (vgl. Kapitel 3) basierender MT-Tisch eingesetzt.

Nach einer kurzen mündlichen Einführung folgte eine 5-minütige moderierte Aufwärmphase, um die Teilnehmer mit den grundlegenden Funktionen und der Modellierungsmetapher vertraut zu machen. Es wurde dabei darauf geachtet, den Probanden die zur Verfügung stehenden Funktionen lediglich verbal und möglichst neutral in ihrer Funktionsweise zu erklären, um die Beeinflussung durch die Demonstration bestimmter Interaktionsmuster möglichst zu vermeiden.

Das graphische Interface zeigte im Wesentlichen eine 3D-Ansicht der Szene

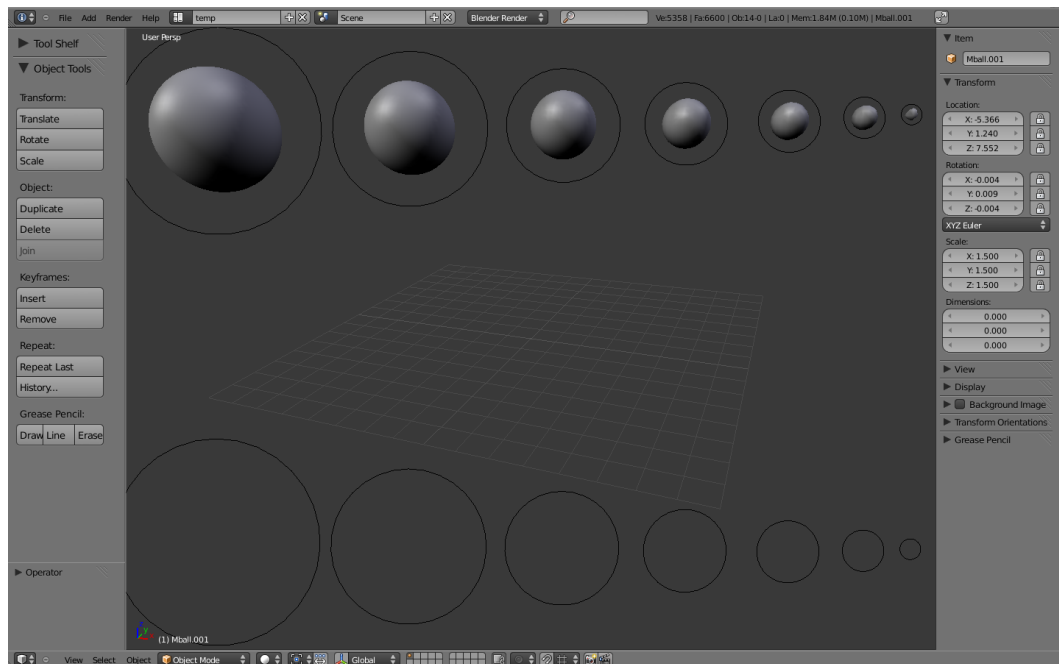


Abbildung 6.7: Die für den Versuch verwendete Modellierungsumgebung: Im oberen und unteren Bereich der 3D-Ansicht ist die Palette der positiven bzw. negativen Metaballs, die den Teilnehmern zur Verfügung stand, zu erkennen.

(vgl. Abb. 6.7). Diese enthielt eine Palette von positiven und negativen Metaballs verschiedener Größen, die als „Schablonen“ verwendet werden konnten. Dazu mussten die vorhandenen Metaballs über einen Button dupliziert und dann mittels der beschriebenen MT-Techniken bewegt und ausgerichtet werden. Soweit möglich, wurden alle nicht benötigten Standard-Interface-Elemente abgeschaltet oder verkleinert. Die einzige Funktionalität, welche im Rahmen des Versuches benötigt wurde und nicht per Touch zu bedienen war, bestand in der besagten Verwendung eines Buttons zum Duplizieren der Schablonen-Metaballs.

Zum Ende der Aufwärmphase wurde den Probanden ein einfaches Beispielmodell gezeigt (vgl. Abb. 6.1 und Abb. 6.2), um den Teilnehmern eine ungefähre Vorstellung von den Möglichkeiten des Systems zu vermitteln. Den Teilnehmern war jedoch freigestellt, ein beliebiges eigenes Modell in der nachfolgenden 15-minütigen freien Modellierungsphase zu kreieren und dabei die Werkzeuge nach eigenem Ermessen einzusetzen. Lediglich bei größeren Problemen stand ein Versuchsleiter für Fragen zur Verfügung.

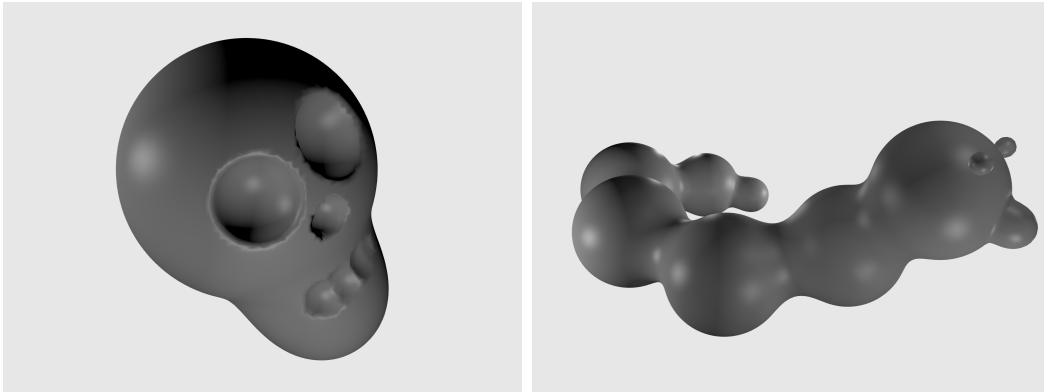


Abbildung 6.8: Beispiele für die Ergebnisse, die von den Teilnehmern während des Versuchs erstellt wurden

Nach Abschluss der Modellierungsphase wurden die Probanden gebeten, einen QUESI-Fragebogen (Hurtienne und Naumann, 2010) auszufüllen, um subjektives Feedback bzgl. der Verständlichkeit des Systems zu gewinnen. Der QUESI-Fragebogen misst die wahrgenommene „Intuitivität“ auf fünf verschiedenen Subskalen: mentale Belastung, Zielerreichung, Lernaufwand, Vertrautheit und Fehlerrate.

6.1.6 Ergebnisse

Insgesamt nahmen sechs Probanden (4 m, 2 w) an dem Versuch teil. Die Teilnehmer waren ausschließlich Rechtshänder mit einem Durchschnittsalter von 27,5 Jahren (SD 2,81). Alle Teilnehmer kamen aus einem Informatik bzw. digitale Medien nahen Umfeld.

Alle Teilnehmer verstanden die Modellierungsmetapher und die grundlegenden Techniken zur Objektmanipulation und Kamerakontrolle sehr schnell und erstellten erfolgreich einfache Modelle mit dem Versuchssystem (vgl. Abb. 6.8).

Keiner der Probanden hatte das System zuvor verwendet und die Hälfte der Teilnehmer gab an, keine oder nur sehr begrenzte Erfahrung mit 3D-Modellierungssoftware im Allgemeinen zu haben. Es konnte beobachtet werden, dass die Teilnehmer die Kamerakontrolle zwar technisch ohne Probleme verwendeten, es einigen aber schwer viel, die richtige Perspektive zu wählen bzw. sich in der 3D-Szene zu orientieren. Einmal musste der Versuchsleiter eingreifen und einem Probanden helfen, die gesuchten Objekte mit der Kamera

wieder zu finden.

Ähnliche Beobachtungen konnten bzgl. der Orientierung und Anordnung der Metaballs gemacht werden. Den unerfahrenen Teilnehmern fiel es sehr viel schwerer, die räumliche Anordnung der Objekte zu begreifen.

Die Selektion verdeckter und kleiner Objekte stellte für viele Teilnehmer ein Problem dar, obwohl die meisten Teilnehmer dieses Problem schließlich durch eine entsprechende Rotation der Kamera lösten.

Nach einer kurzen Erinnerung durch den Versuchsleiter, setzten die Probanden die indirekte Translation erfolgreich für die Manipulation sehr kleiner Objekte oder in Situationen mit vielen anderen Objekten in der Nähe ein.

Im Einklang mit den Ergebnissen von Terrenghi u. a. (2007) konnte festgestellt werden, dass die meisten Teilnehmer zunächst kaum von der Möglichkeit der beidhändigen Interaktion Gebrauch machten. Lediglich ein Proband verwendete die beidhändige Kamerakontrolle (NDH) und Objektmanipulation (DH) von Beginn an und mit der erwarteten Rollenverteilung der Hände. Ansonsten dominierte bei den Probanden weitgehend die einhändige Interaktion und ein häufiger Wechsel zwischen Kamerakontrolle und Objektmanipulation. Erst gegen Ende des Versuches konnte häufiger der Einsatz beider Hände beobachtet werden. Ein Proband verwendete erfolgreich die „auto-constraints“-Technik zur Ausrichtung der Objekte.

Insgesamt waren die Kommentare der Versuchsteilnehmer sehr positiv. Viele gaben an, dass die Benutzung des Systems trotz einiger Probleme „großen Spaß“ bei der Benutzung gemacht habe.

Die Mittelwerte der QUESI-Subskalen (vgl. Tabelle 6.3) zeigen, dass das Testsystem am besten in der Bewertung des Lernaufwands abschnitt. Im Hinblick auf die kurze Lernphase sind die Resultate auch für die übrigen Subskalen und die Gesamtbewertung zufriedenstellend, so erreicht das (experimentelle) System im Schnitt bei allen Kriterien eine neutrale bis positive Bewertung mit einem Score größer als drei.

6.1.7 Diskussion

Insgesamt zeigen die Ergebnisse, dass die einhändigen Gesten und die Techniken zur Kamerakontrolle und Objektmanipulation funktionieren. Die Teilnehmer haben diese schnell und ohne größere Probleme verstanden und konn-

Erfahrung (Modellierung)	Mentale Belastung	Zielerreichung	Lernaufwand	Vertrautheit	Fehlerrate	Gesamtbewertung
nein	2,3	2,7	2,7	2,7	1,5	2,4
ja	4,7	4	4,7	4,7	4	4,4
nein	3	2	3	3	3	2,8
ja, sehr begrenzt	3,3	3	4,3	3,3	4	3,6
ja	4,7	4,7	4,3	4,3	4,5	4,5
nein	3,3	2,7	4	3,7	3,5	3,4
Mittel	3,6	3,2	3,8	3,6	3,4	3,5

Tabelle 6.3: QUESI Ergebnisse pro Teilnehmer – Skala von 1,0 (niedrig) bis 5.0 (hoch)

ten sie nach einer kurzen mündlichen Einführung erfolgreich anwenden. Die Techniken und Möglichkeiten zur beidhändigen Interaktion wurden ebenfalls verstanden und erfolgreich eingesetzt, allerdings deutlich seltener und erst nach einer längeren Eingewöhnungsphase. Eine Erklärung dafür ist die umfangreiche Erfahrung mit den etablierten Eingabegeräten Tastatur und Maus der Probanden und die daraus resultierende Gewöhnung an das herkömmliche Single-Point-Eingabeschema sein. Dies deckt sich mit den Ergebnissen anderer Studien (Terrenghi u. a., 2007; North u. a., 2009). Gegen Ende der jeweiligen Versuchsdurchläufe zeichnete sich bei einigen Teilnehmern jedoch ein Trend hinsichtlich zunehmender beidhändiger Interaktion ab (vgl. auch Kapitel 4). Hier wären allerdings weitere Versuche notwendig, um dies genauer zu ergründen.

Das Selektieren und Manipulieren der Objekte bereitete den Teilnehmern wenig Probleme, lediglich die Interaktion mit sehr kleinen oder verdeckten Objekten führte häufig zu Schwierigkeiten. Ein Teil der Probanden konnte dies durch eine entsprechende Kamerarotation ausgleichen. Zusätzlich könnten hier erweiterte Selektionstechniken, beispielsweise die vorgestellten proxybasierten Selektionstechniken aus Kapitel 5, die Benutzer bei der Selektion unterstützen.

Gestaltungsdimension	Einordnung	Erklärung
Multi-point vs. Multi-touch	Multi-point	Berührungen werden als Punkte verarbeitet
Statischer vs. dynamischer Kontext	eher statisch	expliziter Moduswechsel über die Anzahl der Finger
Integrierte vs. Separierte Freiheitsgrade	eher separiert	je nach Technik werden maximal zwei Freiheitsgrade beeinflusst
Direkte vs. indirekte Interaktion	eher direkt	absolute Positionierung bzw. konstanter Abstand
Mehrhändige Interaktionsmuster	beidhändig	intendierte Rollenverteilung der Hände

Tabelle 6.4: Einordnung der Metaball-Interaktionstechniken gemäß den Gestaltungsdimensionen aus Kapitel 4

Die Beobachtungen während der Studie zeigten auch, dass einige Teilnehmer grundsätzliche Schwierigkeiten damit hatten, von der 2D-Darstellung der Szene auf die räumlichen Anordnungen der Objekte im 3D-Raum zu schließen. Trotz der erwarteten Vorteile durch die Kamerakontrolle bzgl. „kinetic depth effect“ in dieser Hinsicht, stellt die Orientierung in einem solchen Szenario viele Benutzer vor große Probleme. Eventuell könnten spezifische Einschränkungen der möglichen Kameraorientierung solchen Benutzern helfen, den Überblick nicht zu verlieren.

6.1.8 Fazit

In diesem Abschnitt wurden einhändige und beidhändige Interaktionstechniken für die 3D-Modellierung auf MT-Bildschirmen vorgestellt. Die Ergebnisse eines durchgeführten ersten Nutzertests zeigen, dass die Techniken verständlich sind und auch von unerfahrenen Benutzern nach nur kurzer Lernphase erfolgreich zur Erstellung einfacher Modelle angewendet werden können. Die Wahl von impliziten Flächen bzw. Metaballs als Modellierungsmetapher hat sich ebenfalls als geeignet erwiesen. In der durchgeführten Untersuchung dominierte klar die einhändige Interaktion. Dies macht deutlich, dass eine längere Eingewöhnungsphase notwendig ist, um die Benutzer mit den beidhändigen Interaktionstechniken vertraut zu machen. Viele Benutzer scheinen im Umgang mit interaktiven Systemen stark von etablierten Eingabegeräten wie der Maus geprägt zu sein. Eventuell könnte für eine zukünftige Studie auch die explizi-

te Anleitung zum Ausprobieren der beidhändigen Interaktionstechniken hier einen Unterschied ausmachen.

Der begrenzte Umfang der Studie bedingt, dass alle Ergebnisse grundsätzlich nur als Trend oder Hinweis und nicht als Beleg gedeutet werden dürfen. Weitere Untersuchungen sind notwendig, um diese Trends kritisch zu analysieren. Dennoch ist es bereits ein positives Ergebnis, dass alle Benutzer mit dem Testsystem in der Lage waren, innerhalb von nur 15 Minuten ein einfaches Modell zu erstellen.

Zusätzliche Hilfen wie die bereits in Kapitel 3 untersuchte Handunterscheidung oder die in Kapitel 4 vorgestellten proxybasierten Selektionstechniken könnten einige der festgestellten Probleme vermeiden oder reduzieren. Auch sollten alternative Modellierungsmetaphern (siehe nächster Abschnitt) untersucht werden. Tabelle 6.4 fasst die Einordnung der vorgestellten Techniken gemäß der Gestaltungsdimensionen nach Kapitel 4 zusammen.

6.2 Virtuelles Sculpting

Bei der 3D-Modellierung sollte das bearbeitete Objekt im Fokus der Aufmerksamkeit stehen, nicht das Interface oder das mathematische Modell dahinter. Aktuelle 3D-Softwarepakete stellen jedoch häufig die mathematische Repräsentation, z.B. polygonale Meshes, in den Vordergrund. Dies gilt ebenfalls für die Werkzeuge, welche z.T. sogar numerische Eingaben durch den Benutzer für einige Operationen voraussetzen. In den letzten Jahren konnte sich das virtuelle oder digitale „Sculpting“ neben den weit verbreiteten Mesh-orientierten Werkzeugen etablieren. Virtuelles Sculpting (VS) abstrahiert von dem mathematischen Modell und stellt dem Anwender eine Auswahl an Werkzeugen zur Verfügung (sog. „brushes“ oder „Pinself“), welche zu einem gewissen Grad die realen Werkzeuge von Bildhauern und anderen traditionellen Künstlern nachbilden. In den meisten VS Systemen existieren beispielsweise Brushes, um „Material“ auf die Oberfläche aufzutragen, diese zu glätten, abzufachen oder zu falten (vgl. Abb. 6.9). VS wird in etablierten Softwarepaketen mit der Maus oder einem Stift-Tablet in Kombination mit der Tastatur bedient, wobei die letztere Kombination eindeutig bevorzugt wird. Obwohl VS den Künstler zu einem gewissen Grad davon befreit, über die darunter liegenden mathemati-

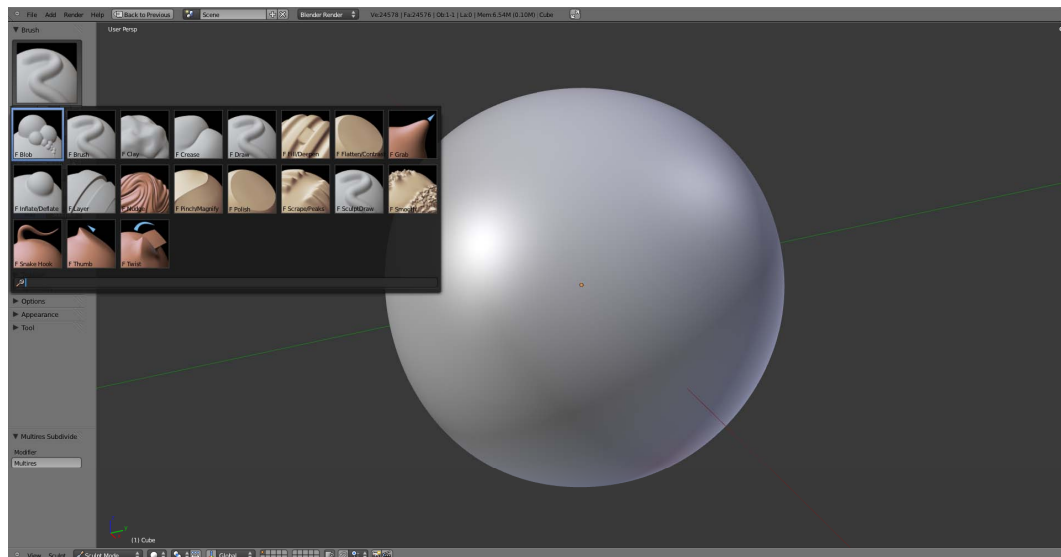


Abbildung 6.9: Screenshot zur Illustration von virtuellem Sculpting (Abbildung aus Herrlich u. a. (2012a)) – Abgebildet ist die Modellierungsumgebung, wie sie dem Benutzer präsentiert wird. Das Zielobjekt befindet sich in der Mitte des Bildschirms; auf der linken Seite kann der Benutzer verschiedene Werkzeuge („Brushes“) auswählen. Modelliert wird durch das (wiederholte) Anwenden der aktiven Brush an den gewünschten Stellen des Zielobjektes.

schen Modelle nachzudenken, muss er sich dabei immer noch sehr stark auf die Eigenheiten und Bedienung der verwendeten Eingabegeräte konzentrieren. VS ist damit immer noch weit entfernt von der Selbstverständlichkeit mit der wir physische Werkzeuge einsetzen.

Zum Teil resultiert dies aus der fehlenden physischen Repräsentation und dem nicht vorhandenen taktilen Feedback. Zum Teil kann dies aber auch dem Arbeitsablauf zugeschrieben werden, der aus der Verwendung der etablierten Eingabegeräte resultiert. Dies gilt besonders im Hinblick auf die beidhändige Interaktion.

Die Vorteile beidhändiger Interaktion wurden bereits in einer Reihe von Studien demonstriert (Leganchuk u. a., 1998; Hinckley u. a., 1998a,b; Gribnau und Hennessey, 1998; Balakrishnan und Hinckley, 1999; Buxton und Myers, 1986; Guiard, 1987; Balakrishnan und Kurtenbach, 1999) und es wurden Hinweise gefunden, dass beidhändige Interaktion der einhändigen Interaktion nicht nur bzgl. Effizienz und Geschwindigkeit sondern auch im Hinblick auf das Auf-

gabenverständnis überlegen sein kann. Die gleichen Studien zeigten allerdings auch, dass diese Vorteile von der konkreten Aufgabe und dem verwendeten Eingabegerät abhängig sind. Daher ist es wichtig, spezifische Kombinationen von Aufgaben und Eingabegeräten auch innerhalb spezifischer Anwendungskontexte zu studieren und nicht nur für verallgemeinerte, abstrahierte Aufgaben. In diesem Abschnitt wird ein Interfacedesign für beidhändiges VS auf interaktiven Oberflächen präsentiert und die Ergebnisse einer Fallstudie diskutiert, welche einhändige und beidhändige Interaktionsmuster für VS auf interaktiven Oberflächen mit dem etablierten Pen-Tablet vergleicht. Die Ergebnisse der Studie weisen daraufhin, dass interaktive Oberflächen tatsächlich erfolgreich für VS genutzt werden können und eine stärkere Affordanz für beidhändige Interaktion als etablierte Eingabegeräte in diesem Anwendungskontext besitzen. Das Modell der „kinematic chain“ von Guiard (1987) für die asymmetrische Aufgabenteilung der Hände bei der beidhändigen Interaktion funktioniert auch für VS bezogen auf die Modellierung und die Kontrolle der Kamera mit MT-Eingaben, allerdings nicht so deutlich wie erwartet. Es konnte außerdem beobachtet werden, dass die Anwendung beidhändiger Interaktion substanzielles Training erfordert und dass der Pen/Stylus besser für die Anwendung der Brushes geeignet ist. Dies ist ein deutlicher Hinweis auf mögliche Vorteile einer zukünftigen Kombination aus Pen- und MT-Interaktion für diesen Anwendungsbereich.

6.2.1 Stand der Technik

In der Literatur werden verschiedene Ansätze für das VS vorgeschlagen. Frühe Systeme erlauben es dem Benutzer einfache polygonale Formen auf den Bildschirm zu malen, welche dann automatisch in 3D-Objekte extrudiert und mit anderen Objekten verschnitten werden, um ein finales Objekt zu erzeugen (Parent, 1977). Andere Arbeiten verbessern das mathematische Modell mit Hilfe verschiedener voxel-basierter Repräsentationen (Galyean und Hughes, 1991; Wang und Kaufman, 1995) oder Modellen auf Basis impliziter Funktionen (Raviv und Elber, 2000; Ferley u. a., 2001). Letztere nutzen verschiedene adaptive Sampling-Verfahren zur Extraktion von Isoflächen für das Rendering.

Viele Systeme beinhalten typische Werkzeuge, um Material hinzuzufügen, zu entfernen oder die Oberfläche zu glätten. Einige Systeme arbeiten direkt

auf polygonalen Meshes (Bill und Lodha, 1994), während andere Constructive Solid Geometry (CSG) einsetzen (Mizuno u. a., 1998). Parviainen u. a. (2004) untersuchen geometrische „Hilfsformen“, z.B. Boxen, Ebenen oder Graden, um die Wahrnehmung und das Verständnis des Benutzers in der 3D-Welt beim VS zu verbessern.

Viele Interfacedesigns und Forschungsarbeiten im Bereich der beidhändigen Interaktion basieren auf den Untersuchungen zu beidhändigen Interaktionsmustern von Guiard (1987). Dieser berichtet, dass auch bei Tätigkeiten, die oberflächlich betrachtet als einhändige Aufgaben eingeordnet werden könnten, die nicht-dominante Hand (NDH) die Aktionen der dominanten Hand (DH) in spezifischer Weise unterstützt und so insgesamt die Leistung verbessert. Er spricht von einer „kinematic chain“.

Da verschiedene Eingabegeräte unterschiedliche Möglichkeiten bzgl. beidhändiger Interaktion aufweisen, existieren Ansätze, diese zu formalisieren und Eingabegeräte entsprechend ihres Interaktionsmodells zu gruppieren (Hinckley u. a., 1998a). Andere Studien bzgl. der potenziellen Vorteile beidhändiger Interaktion zeigen, dass die Benutzer auf verschiedenen Ebenen von beidhändiger Interaktion profitieren können. Je nach dem spezifischen Szenario, kann beidhändige Interaktion zu Geschwindigkeits- und Präzisionsvorteilen beitragen oder auch das mentale Modell des Benutzers von der Aufgabe verbessern (Buxton und Myers, 1986; Leganchuk u. a., 1998; Balakrishnan und Hinckley, 1999).

Balakrishnan und Kurtenbach (1999) nutzen zwei Mäuse, um die Vorteile einer Aufteilung von Manipulation und Kamerakontrolle zwischen der DH und der NDH zu explorieren. Frisch u. a. (2009) untersuchen die Kombination aus MT- und Stift-Eingabe für das Editieren von Diagrammen. Im Gegensatz zu unseren Ergebnissen können sie keine klare Präferenz der Benutzer zu Gunsten der Stift-Eingabe für bestimmte Operationen feststellen und die Möglichkeit der kombinierten MT- und Stift-Interaktion wird von ihren Teilnehmern nur äußerst selten genutzt.

6.2.2 Interaktionsdesign und Gesten

Bei der Gestaltung des Interaktionsdesigns wird ein nutzerzentrierter Ansatz verfolgt. Als Datenbasis dient eine Korpus von Videotutorials für virtuelles Sculpting, welche ausgewertet wird, um die wichtigsten Werkzeuge zu iden-

tifizieren und deren typische Anwendung im Modellierungsprozess besser zu verstehen.

Hauptziele der Videoanalyse sind somit: Die Identifizierung einer grundlegenden Auswahl an Werkzeugen für das virtuelle Sculpting (vorausgesetzt eine solche Auswahl existiert) und die Beobachtung typischer Verwendungsmuster.

Der Korpus von Videotutorials aus dem Web wird nach den folgenden Kriterien ausgewählt: Ein klarer Fokus auf virtuelles Sculpting, d.h. keine „gemischten“ Tutorials mit anderen 3D-Modellierungstechniken. Die Abdeckung des vollständigen Modellierungsprozesses von Anfang bis Ende. Zusätzlich werden, soweit möglich, die Expertise und Qualität der Tutorials in die Auswahl mit einbezogen. Hierfür werden der allgemeine Bekanntheitsgrad und, sofern vorhanden, öffentliche Nutzerbewertungen der Tutorials zu Grunde gelegt.

Insgesamt wurden 12 Videotutorials (ca. 5h:30m) analysiert. Dazu wurden alle Operationen mit Bezug zum Sculpting (inklusive der Menüs für das Ändern von Eigenschaften und Werkzeugen) ausgezählt. Die Kontrolle des Fensters und allgemeine Anwendungsfunktionen wurden dagegen nicht einbezogen. Die Auswertung der Tutorials ergab, dass die grundlegende Auswahl an Operationen für das virtuelle Sculpting aus sieben Aktionen bestand: Anwendung des aktuellen Werkzeugs/Bedienung des Menüs (38,27%), Rotation der Sicht (21,94%), Zoomen der Sicht (12,33%), Anwendung des Glättungswerkzeuges (9,89%), Verschieben der Sicht (8,22%), Anpassung der Größe des Werkzeuges (2,44%) und „Undo“ (1,2%). Alle weiteren Aktionen machten für sich genommen jeweils weniger als 1,2% aller beobachteten Aktionen aus. Die Anwendung des aktuellen Werkzeuges und die Verwendung des Menüs wurden zusammen gezählt, da beide Operationen in gleicher Weise (Aufsetzen des Stiftes auf die Tablet Oberfläche) ausgeführt wurden, während die anderen Operationen den Einsatz zusätzlicher Tasten erforderlich machten.

Mehrfingergesten (Matejka u. a., 2009; Walther-Franks u. a., 2011b) sind eine effektive Möglichkeit für die Abbildung von allgemeinen Operationen auf Multi-touch-Gesten, sofern die Anzahl der Operationen klein genug ist (vgl. Kapitel 4). Außerdem ermöglichen diese verschiedene Formen der beidhändigen Interaktion.

Neben der Anzahl der Finger werden weitere Parameter in das Gestendesign mit einbezogen, um die Anzahl der möglichen Abbildungen zu erweitern. So

	<p>Ein Finger wird auf die häufigsten Operationen abgebildet, d.h. die Anwendung des aktuellen Werkzeuges und die Selektion aus dem Menü.</p>
	<p>Zwei Finger auf dem Modellierungsobjekt werden auf das Glättungswerkzeug abgebildet. Zwei Finger außerhalb des Objektes dienen zum Verschieben der Sicht.</p>
	<p>Drei Finger dienen zum Rotieren der Sicht oder zum Ändern der Werkzeuggröße. Bilden die Finger einen Winkel von ungefähr 90 Grad, so wird die Werkzeuggröße geändert, andernfalls die Sicht rotiert.</p>
	<p>Vier Finger dienen dem Zoomen der Sicht.</p>
	<p>Fünf Finger führen die „Undo“ Operation aus (kurzer „tap“).</p>

Tabelle 6.5: Mehrfingergesten für das Sculpting (Tabelle aus Herrlich u. a. (2012a))

hängt die ausgeführte Aktion auch davon ab, ob eine Geste auf dem Modellierungsobjekt oder daneben ausgeführt wird und auch die relative Positionierung der Finger (die Handpose) kann eine Rolle spielen. Basierend auf den Ergebnissen der Videoanalyse werden die in Tabelle 6.5 dargestellten Gesten für die grundlegenden Operationen definiert.

6.2.3 Experiment

Um das Interfacedesign und beidhändige Benutzungsmuster für virtuelles Sculpting näher zu untersuchen, wurde eine Nutzerstudie durchgeführt. Die spezifischen Ziele dieses Experiments waren wie folgt: Die Demonstration der Mach-

barkeit, d.h. die grundsätzliche Eignung von MT-Eingaben (speziell die oben angeführten Mehrfingergeräten) für das virtuelle Sculpting. Die Verifikation von Guiards Modell zur beidhändigen Interaktion, angewendet auf virtuelles Sculpting und eventuell die Identifikation von alternativen beidhändigen und einhändigen Interaktionsmustern. Der Vergleich von Interaktionsmustern und -problemen zwischen MT und der etablierten Eingabegerätekombination aus Stift-Tablet und Tastatur.

Viele existierende Studien im Bereich der MT-Interfaces zielen auf *Walk-up-and-use*-Anwendungen, d.h. auf einen möglichst „intuitiven“ Zugang (im Sinne von „ohne spezielle Vorbereitung/Lernphase“) für unerfahrene Benutzer. 3D-Modellierung ist jedoch ein sehr komplexes Anwendungsfeld. Vorerfahrungen aus anderen Untersuchungen (vgl. vorheriger Abschnitt) legen nahe, dass eine gewisse Lernphase daher stets erforderlich ist. Des Weiteren weisen mehrere Studien darauf hin, dass die beidhändige Interaktion auf MT Bildschirmen eine Expertentechnik ist, welche nicht ad hoc von unerfahrenen Benutzern ausgeführt wird (North u. a., 2009; Terrenghi u. a., 2007; Walther-Franks u. a., 2011b). Daher war es nicht Ziel dieser Studie, ein „natürliches“ Interface für unerfahrene Benutzer zu demonstrieren, sondern den Verlauf der Interaktion im Rahmen einer längeren Lernphase zu dokumentieren. Daher wurde die Untersuchung als Fallstudie mit wenigen Probanden, aber einer längeren Sitzungsdauer pro Proband angelegt. In jeder Sitzung konnte ein Proband das System zusammen mit einem Versuchsleiter ausführlich explorieren.

Um einen Vergleich zwischen MT und Stift/Tastatur herstellen zu können, haben die Probanden beide Eingabegerät getestet. Die Einstellungen für Stift/Tastatur basierten auf den Standardeinstellungen der 3D-Modellierungsanwendung Blender, reduziert auf die gleichen grundlegenden Operationen, die sich aus der oben beschriebenen Videoanalyse ergaben (vgl. Tabelle 6.6).

Aufbau und Ablauf

Jeder Proband testete beide Interfacevarianten (MT vs. Stift/Tastatur) ausführlich (ca. 1,5 Stunden). Auf Grund des erheblichen Zeitaufwandes wurde den Probanden eine Aufwandsentschädigung von 15 EUR gezahlt. Die Teilnehmer wurden aufgefordert jederzeit eine Pause zu machen, falls sie sich erschöpft fühlten.

Aktuelles Werkzeug / Menüselektion	Stiftspitze
Rotieren der Sicht	Stifttaste + Stiftspitze
Zoomen der Sicht	Strg + Stifttaste + Stiftspitze
Glätten	Stiftende („Radierer“)
Verschieben der Sicht	Umschalttaste + Stifttaste + Stiftspitze
Größenänderung des Werkzeugs	F-Taste + Stiftspitze
„Undo“	Strg + Z-Taste

Tabelle 6.6: Interfaceeinstellungen für Stift-Tablet und Tastatur.

Links: Die auszuführende Operation; Rechts: Stift-/Tastatur-Aktion, um die jeweilige Operation auszuführen (Tabelle aus Herrlich u. a. (2012a))



Abbildung 6.10: Foto des experimentellen Aufbaus (anonymisiert)

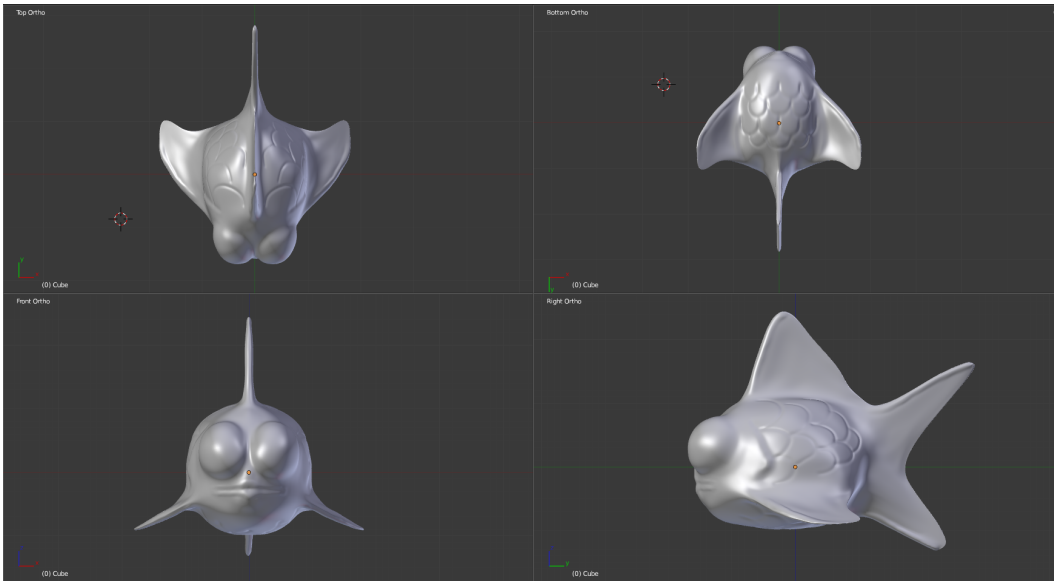


Abbildung 6.11: Verschiedene Ansichten des Modells, welches den Teilnehmern für das Experiment als Ziel vorgegeben wurde (Abbildung aus Herrlich u. a. (2012a)).

Als Modellierungsumgebung wurde eine modifizierte Version der Open Source Software Blender eingesetzt, welche mit den oben beschriebenen MT- bzw. Stiftgesten gesteuert werden konnte. Als MT Eingabegerät kam ein 22" Monitor vom Typ 3M M2256PW zum Einsatz (vgl. Abbildung 6.10), welcher auf die native Auflösung von 1680x1050 Pixeln eingestellt war. Derselbe Monitor wurde auch als nicht-interaktiver Bildschirm mit abgeschalteter MT-Funktionalität in der Stift-Tablet Kondition verwendet. Als Stift-Tablet wurde ein Wacom Intuos A3 wide Tablet zusammen mit einer drahtlosen deutschsprachigen Tastatur verwendet. Der eingesetzte Stift besaß zwei zusätzliche Tasten und das hintere Ende („Radiergummi“) konnte ebenfalls mit Funktionen belegt werden (vgl. Tabelle 6.6). Um eine Verwechslung der beiden Knöpfe auszuschließen, wurden diese für den Test auf die gleiche Funktionalität eingestellt.

Der vollständige Ablauf des Experimentes bestand für jede der zwei Konditionen aus den folgenden Schritten: Die Teilnehmer wurden begrüßt, der grundsätzliche Ablauf des Experiments kurz erklärt und demographische Daten abgefragt. Daraufhin wurde den Probanden zunächst durch ein kurzes Erklärungsvideo (ca. 3 Minuten) die jeweilige Steuerung und der grundlegende Arbeitsprozess beim virtuellen Sculpting erklärt.

Anschließend an das Video hatten die Teilnehmer die Möglichkeit, sich selbst mit den Kontrollen vertraut zu machen und Verständnisfragen zu stellen. Wenn die Teilnehmer angaben, sich mit der Benutzung sicher genug zu fühlen, wurde ihnen das Zielmodell (vgl. Abbildung 6.11) anhand verschiedener Bilder gezeigt und sie wurden gebeten, dieses so genau wie möglich mit den jeweiligen Kontrollen nach zu modellieren.

Das Zielmodell stammte aus einem der analysierten Videotutorials für Anfänger und wurde gewählt, da es den vollständigen Satz an grundlegenden Operationen ausnutzt. Den Probanden wurde eine ausgedruckte einseitige Übersicht der Gesten bzw. Operationen während des Experiments zur Verfügung gestellt. Die Modellierungsphase war zeitlich nicht begrenzt, sondern wurde als abgeschlossen betrachtet, wenn die Teilnehmer angaben, fertig zu sein.

Während des Versuchs wurden die Teilnehmer angehalten, die Position und Neigung des Bildschirms und die Position der anderen Eingabegeräten nach eigenem Empfinden komfortabel einzustellen. Sie wurden außerdem ermutigt, sowohl eine sitzende als auch eine stehende Position bei der Interaktion auszuprobieren.

Abschließend wurde die subjektive Einschätzung der Teilnehmer abgefragt. Dazu wurden die Probanden gebeten, den „System Usability Scale“ (SUS) Fragebogen auszufüllen. Darauf folgte ein semi-strukturiertes Interview. Außerdem wurden Video- und Audioaufnahmen jeder Sitzung für die weitere Analyse gesichert. Die Reihenfolge der Konditionen wurde abgewechselt und die Zuordnung der Teilnehmer erfolgte zufällig.

6.2.4 Ergebnisse

Vier Probanden (2 m, 2 w) nahmen an dem Experiment teil. Das Durchschnittsalter lag bei 20,75 Jahren (SD 2,86) und alle Teilnehmer waren Rechtshänder. Die Teilnehmer hatten unterschiedliche Vorerfahrungen bzgl. MT-/Stift-Tablet-Eingabe und 3D-Modellierung.

Teilnehmer 1 besaß Erfahrung mit MT auf mobilen Geräten, hatte aber keine Erfahrung mit 3D-Modellierung, MT auf dem Desktop oder Stift-Tablets. Der Proband konnte während der Experiments grundsätzlich alle MT-Gesten erfolgreich ausführen. Zu Beginn der Sitzung verwendete der Teilnehmer fast ausschließlich die dominante Hand, später konnte jedoch auch häufiger die Ver-

wendung der nicht dominanten Hand beobachtet werden. Diese wurde dabei hauptsächlich für die asynchrone Kontrolle der Kamera eingesetzt. Synchrone beidhändige Kamerakontrolle konnte nur einmal beobachtet werden. Obwohl die Illustrationen auf der einseitigen Anleitung (vgl. Tabelle 6.5) nicht die Verwendung des Daumens zeigten, führte der Teilnehmer dennoch viele Gesten unter Einsatz des Daumens aus. Mehrmals konnte beobachtet werden, dass der Proband versuchte die Pinch-Geste für das Zoomen einzusetzen, obwohl diese nicht von unserem System unterstützt wird. Nach ungefähr der Hälfte der Zeit wechselte der Proband für die MT Interaktion von einer sitzenden in eine stehende Position. Während der Interaktion mit dem Stift-Tablet konnte ausschließlich die einhändige Kontrolle beobachtet werden. Auf Nachfrage erklärte der Proband das Stift-Interface vorzuziehen.

Teilnehmer 2 verfügte laut Selbsteinschätzung über umfangreiche Erfahrung mit MT auf mobilen Geräten, hatte jedoch keine Vorerfahrung mit Desktop-MT, 3D-Modellierung oder Stift-Tablets. Bei der Verwendung des Stift-Tablets konnte zunächst lediglich die einhändige Interaktion beobachtet werden. Gegen Ende der Sitzung verwendete der Proband gelegentlich die Tastatur mit der nicht dominanten Hand. Der Proband versuchte häufig den Stift wie eine Maus einzusetzen und diesen z.B. mehrmals anzuheben und wieder aufzusetzen, um den Bildausschnitt zu verschieben. Auf Grund des absoluten Tracking des Stiftes funktionierte dies nicht und der Proband wirkte verwirrt. Der Teilnehmer verwendete für die MT-Interaktion hauptsächlich die dominante Hand und zog häufig die Anleitung zu Rate. Während des Versuchs wechselte der Teilnehmer von einer sitzenden in eine stehende Position. Nach ungefähr der Hälfte Zeit der MT-Bedingung konnte beobachtet werden, dass der Proband begann, beide Hände knapp über der Oberfläche zu halten und die nicht-dominante Hand für Kameramanipulationen sowie die dominante Hand für die Objektmanipulation einzusetzen. Zum Ende der Sitzung wechselte der Proband wiederum das Interaktionsmuster und setzte beide Hände sowohl für die Kamerakontrolle als auch für die Objektmanipulation ein. Dabei schien hauptsächlich die Entfernung zum gewünschten Manipulationsbereich die Wahl der Hand zu beeinflussen. Der Proband gab an, dass MT „ein besseres Gefühl für das was man tut“ vermittele.

Teilnehmer 3 verfügte über begrenzte Vorerfahrung mit MT auf mobilen

Geräten, jedoch über keinerlei Erfahrung mit 3D-Modellierung, Desktop-MT oder Stift-Tablets. Der Proband interagiere in beiden Versuchsbedingungen fast ausschließlich einhändig mit der dominanten Hand. Der Teilnehmer konnte im Versuch grundsätzlich alle MT-Gesten erfolgreich ausführen. Der Proband interagiere ausschließlich aus einer sitzenden Position, veränderte diese jedoch häufig. Es konnten ähnliche Probleme wie bei Teilnehmer 2 bzgl. der Stift-Interaktion beobachtet werden. Auch dieser Teilnehmer versuchte, den Stift wie eine Maus zu verwenden, was auf Grund des absoluten Trackings zu Problemen führte. Auf Nachfrage schätzte der Teilnehmer die Interaktion mit dem Stift als „etwas weniger anstrengend“ im Vergleich zu MT ein.

Teilnehmer 4 verfügte über Vorerfahrung mit Desktop-MT, Stift-Tablets und 3D-Modellierung. Für den Fall der Stift-Interaktion konnte ausschließlich einhändige Interaktion mit der dominanten Hand beobachtet werden. Bei der Interaktion mit MT verwendete der Proband von Anfang an beide Hände. Dabei wurde die nicht dominante Hand ausschließlich für die Kontrolle der Kamera und die dominante Hand für alle übrigen Funktionen eingesetzt. Der Teilnehmer interagiere ausschließlich aus einer sitzenden Position. Der Proband lobte ausdrücklich die Möglichkeit der beidhändige Interaktion mit dem MT-Interface, schätzte die Stift-Interaktion aber als „etwas weniger anstrengend“ ein.

Die Ergebnisse des SUS Fragebogens ergaben einen Wert von 24,0 (SD 5,24) für MT und 32,25 (SD 1,92) für Stift-Tablet. Die Aktionsfrequenzen waren in beiden Bedingungen vergleichbar und lagen bei ungefähr 15 Aktionen pro Minute. Bzgl. beidhändiger Interaktionsmuster lassen sich die Beobachtungen wie folgt zusammenfassen: Nach einer Lernphase von ca. 30-40 Minuten verwendeten alle Teilnehmer gelegentlich beide Hände für die MT-Interaktion. Ein Teilnehmer mit erheblicher Vorerfahrung zeigte von Beginn an beidhändige Interaktion. Im Falle der beidhändigen Interaktion dominierte klar ein asymmetrisches, asynchrones Interaktionsmuster. Dabei wurde die nicht dominante Hand für die Kamerakontrolle und die dominante Hand für die Objektmanipulation eingesetzt. Im Fall der Stift-Interaktion konnte nur in einem Fall eine Form der beidhändigen Interaktion beobachtet werden, bei der auf der Tastatur zusätzliche Funktionstasten bedient wurden.

6.2.5 Diskussion

Bzgl. der ursprünglichen Ziele der Untersuchung ist festzustellen, dass die grundsätzliche Anwendbarkeit von MT und Prinzipien der beidhändigen Interaktion auf das virtuelle Sculpting erfolgreich demonstriert werden konnte. Die Ergebnisse früherer Studien (Terrenghi u. a., 2007; North u. a., 2009) bzgl. des Einflusses bzw. der Prägung durch die Interaktion mit der Maus auf das mentale Modell der Benutzer konnten ebenfalls bestätigt werden.

Auf Grund des Fallstudienansatzes und einer längeren Sitzungsdauer je Teilnehmer konnte ein klarer Anstieg der beidhändigen Interaktion über den individuellen Versuchsverlauf beobachtet werden. Das Modell bzw. die Ergebnisse von Guiard (1987) konnten im Wesentlichen für die Kamerakontrolle und die Objektmanipulation bestätigt werden, welche beinahe ausschließlich asymmetrisch mit nicht-dominanter bzw. dominanter Hand ausgeführt wurden. Es muss jedoch festgehalten werden, dass von diesem Modell nur begrenzt Heuristiken beispielweise zur Detektion der Hände abgeleitet werden können, da in Einzelfällen auch das Wechseln der Hände und symmetrische beidhändige Interaktion beobachtet werden konnten.

Bzgl. des Vergleichs zwischen Stift und MT-Interaktionsmustern muss zunächst beachtet werden, dass die Stift-Eingabegeräte etablierter und ausgereifter sind. Darin könnte ein Grund für das bessere Abschneiden im Versuch liegen. Die Teilnehmer bevorzugten den Stift insbesondere für die Manipulation des Objektes, machten in Kombination mit der Tastatur jedoch kaum von den Möglichkeiten der beidhändigen Interaktion Gebrauch.

Daraus erschließt sich, dass eine Kombination von Stift-Eingaben für die Manipulation und MT-Eingaben für die Kamerakontrolle viel Potenzial besitzt. In einer ersten Iteration des Testsystems wurde das kontinuierliche Wechseln zwischen verschiedenen Operationen und Werkzeugen durch das Auf- und Absetzen von Fingern „on-the-fly“ in Betracht gezogen. Im Versuch zeigte sich jedoch, dass die Teilnehmer die Gesten klar getrennt einleiteten und zwischen verschiedenen Gesten die Hände vollständig von der Oberfläche nahmen.

Das relativ schlechte Abschneiden in den SUS-Scores ist sehr wahrscheinlich bedingt durch die komplexe Aufgabe und den Fokus des Experiments auf die Erhebung von qualitativen Daten.

Gestaltungsdimension	Einordnung	Erklärung
Multi-point vs. Multi-touch	Multi-point	Berührungen werden als Punkte verarbeitet
Statischer vs. dynamischer Kontext	eher statisch	expliziter Moduswechsel über die Anzahl der Finger bzw. den Ort der Interaktion
Integrierte vs. Separierte Freiheitsgrade	eher separiert	je nach Technik werden maximal zwei Freiheitsgrade beeinflusst
Direkte vs. indirekte Interaktion	eher indirekt	Einsatz von abstrakten Werkzeugen
Mehrhändige Interaktionsmuster	beidhändig	intendierte Rollenverteilung der Hände

Tabelle 6.7: Einordnung der Virtual Sculpting-Interaktionstechniken gemäß den Gestaltungsdimensionen aus Kapitel 4

6.2.6 Fazit

In diesem Abschnitt wurde ein Ansatz zur Nutzung von MT für das virtuelle Sculpting mit besonderem Fokus auf die beidhändige Interaktion präsentiert. Die Ergebnisse einer ersten Fallstudie zeigen, dass interaktive Oberflächen erfolgreich für das virtuelle Sculpting eingesetzt werden können. Es zeigte sich außerdem, dass die von Guiard beschriebenen Prinzipien der Arbeitsteilung von nicht dominanter und dominanter Hand auf bestimmte Aktionen beim virtuellen Sculpting – Kamerakontrolle und Objektmanipulation – anwendbar sind, dass die Benutzung der beidhändigen Interaktion allerdings eine substanzielle Lernphase voraussetzt und daher weniger für *Walk-up-and-use*-Szenarien geeignet ist. Der Vergleich von MT- und Stift-Interaktionsmustern ergab, dass MT verständlicher sein kann als die Stift-Interaktion und dass die Interaktion mit beiden Eingabegeräten stark von der Vorerfahrung der Benutzer mit der Maus beeinflusst wird. Der Stift wurde insbesondere für die Objektmanipulation von den Benutzern präferiert. Dagegen wurden die Möglichkeiten der MT-Kameramanipulation von einigen Benutzern positiv erwähnt. Dies lässt den Schluß zu, dass VS von einer Kombination aus Stift- und MT-Eingabe profitieren könnte. Die präsentierten Ergebnisse sind auf Grund der Natur und des Umfangs der beschriebenen Studien nicht als abschließend zu betrachten, bestärken jedoch weitere Untersuchungen in diesem Bereich. Tabelle 6.7 fasst die Einordnung der vorgestellten Techniken gemäß der Gestaltungsdimensionen nach Kapitel 4 zusammen.

Kapitel 7

Fazit

In diesem Abschnitt wird eine abschließende Zusammenfassung und kritische Bewertung der erzielten Ergebnisse vorgenommen.

Die wissenschaftliche Motivation der vorliegenden Arbeit bestand in der übergeordneten Fragestellung, ob die 3D-Modellierung als Anwendungsfall von den (potenziell) erweiterten Möglichkeiten der Multi-touch (MT)-Interaktion im Vergleich zu den etablierten Eingabegeräten Maus und Tastatur profitieren kann. Ein besonderer Fokus lag dabei auf beidhändigen Interaktionstechniken. Aus dieser Fragestellung wurden weitere Detailfragen abgeleitet und untersucht:

- Was ist der aktuelle Stand der Technik im Bereich der Hardware? Welche Möglichkeiten bietet diese prinzipiell? Können oder müssen auf dieser Ebene die Interaktionsmöglichkeiten erweitert werden?
- Welche Interaktionstechniken und Modellierungsansätze/-metaphern sind mit MT-Eingaben einsetzbar bzw. profitieren davon?
- Wie ist die Nutzererfahrung? Gibt es Hinweise auf relevante Unterschiede zwischen Experten und unerfahrenen Benutzern?
- Ist eine schrittweise Erweiterung der Techniken bzw. der Softwaretools durch MT-Eingaben möglich und sinnvoll oder ist ein vollständiger Paradigmenwechsel notwendig?

Insgesamt war die gewählte Vorgehensweise durch die Beschäftigung mit dem Themenfeld entlang der gesamten Verarbeitungskette von der Hardware

bis hin zu einer komplexen Modellierungsumgebung gekennzeichnet.

7.1 Diskussion der eigenen Beiträge

Zu Beginn dieser Arbeit (vgl. Kapitel 3) erfolgte zunächst eine ausführliche Einführung in den Stand der Technik im Bereich der MT-Hardware und eine Diskussion der unterschiedlichen Erkennungs- bzw. Verarbeitungsverfahren. Dabei wurden die Stärken und Schwächen der verschiedenen Ansätze herausgestellt.

Ein Ergebnis dieser Diskussion bestand darin, dass bei MT-Hardware, welche die Erkennung nach dem Prinzip der Diffuse Illumination (DI) realisiert, die Beleuchtungskonstruktion einen kritischen Punkt darstellt. Im Rahmen dieser Arbeit wurde dazu ein Vorschlag für ein solides Beleuchtungskonzept für MT-Tische auf DI-Basis vorgestellt. Es wurde auch festgestellt, dass das Potenzial von DI noch nicht voll ausgeschöpft wird. Besonders gilt dies in Bezug auf die Interaktion im Nahbereich über der Oberfläche und die Zuordnung von Fingerberührungen zu einer bestimmten Hand. Ein weiterer Beitrag dieser Arbeit bestand daher in der Entwicklung eines Ansatzes für die Interaktion im „Nah-3D“-Bereich über der Oberfläche („Hovering“). Ein erster Proof-of-Concept-Testlauf verlief positiv. Außerdem wurde ein Ansatz zur Handunterscheidung auf interaktiven Oberflächen vorgestellt. Hier verlief ein Proof-of-Concept-Testlauf ebenfalls erfolgreich.

Einen Hauptbeitrag dieser Arbeit stellte auch die Diskussion und Ordnung der Gestaltungsmöglichkeiten in verschiedenen Dimensionen dar (vgl. Kapitel 4). Diese Dimensionen werden durch ihre Extrempositionen charakterisiert, um die Vielfalt der Möglichkeiten aufzuzeigen und diesen eine Ordnung zu geben, die als Leitlinie für das Design von zukünftigen und die Einschätzung von derzeitigen Interfaces dienen kann.

In der zweiten Hälfte der Arbeit (Kapitel 5) wurden empirische Fallstudien zu grundlegenden Techniken für die Selektion und Manipulation von Objekten auf interaktiven Oberflächen vorgestellt. Selektion und Manipulation sind fundamentale Operationen jeder 3D-Modellierungsanwendung. Die Ergebnisse der Untersuchungen zeigten ein differenziertes Bild. So konnten die Techniken grundsätzlich erfolgreich von den Probanden eingesetzt werden, es zeigte

sich jedoch auch, dass speziell die Manipulation von 3D-Objekten auf einer 2D-Projektion eine inhärent schwierige Aufgabe ist.

In Kapitel 6 wurden schließlich Modellierungstechniken-/metaphern näher auf ihre Eignung und Vor- bzw. Nachteile für die Benutzung auf interaktiven Oberflächen betrachtet. Insbesondere die beidhändige Interaktion stand hier bei beiden Techniken im Fokus der Untersuchung. Hier zeigte sich wieder, dass die Techniken einerseits verständlich und benutzbar sind, aber andererseits, insbesondere die Verwendung der beidhändigen Techniken auch eine signifikante Lernphase und Eingewöhnungszeit erfordert bzw. mit sich bringt.

7.2 Ergebnis bzgl. der wissenschaftlichen Fragestellung

Zusammenfassend im Hinblick auf die Fragestellung dieser Arbeit kann gesagt werden, dass gezeigt werden konnte, dass MT-Eingabegeräte flexibel einsetzbar sind und Potential im Bereich der 3D-Modellierung aufweisen.

Aus der Diskussion des Gestaltungsraums ist erkennbar, dass es dem Interfacedesigner freisteht, zwischen einer „sanften“ Erweiterung mittels MT-Eingaben und eher „radikalen“ neuen Eingabetechniken zu wählen.

Außerdem zeigten die empirischen Untersuchungen, dass insbesondere beidhändige Techniken erfolgreich eingesetzt werden können und den Modellierungsprozess parallelisieren. Sie ergaben jedoch auch, dass *Walk-up-and-use* nicht das einzige Kriterium für MT ist und auch MT-Techniken gelernt werden müssen und dass „Natürlichkeit“ in diesem Zusammenhang ein relativer, schwer zu fassender Begriff ist.

7.3 Offene Fragen und Ausblick

Wie im letzten Abschnitt bereits erwähnt, konnten die Techniken zur Hover-Detektion und zur Handunterscheidung im Rahmen dieser Arbeit bisher nur als erster Proof-of-Concept umgesetzt werden. Allerdings zeigte sich bereits in diesem Stadium ein großes Potenzial. Daher wäre eine Verbesserung und intensive empirische Untersuchung dieser Ansätze in zukünftigen Arbeiten sinnvoll.

Außerdem wurden diese Techniken bisher nicht im Rahmen einer komplexen Modellierungsanwendung erprobt.

Die durchgeführten Experimente im empirischen Teil der Arbeit haben teilweise den Charakter von Vorstudien. Die Ergebnisse sollten daher mit größeren Benutzergruppen validiert werden. Eine explizite Gegenüberstellung von Anfängern und Experten und eine Begleitung über einen längeren Zeitraum könnten außerdem weitere wertvolle Erkenntnisse liefern. Beispielsweise könnten Tagebuchstudien eingesetzt werden. Die Ergebnisse haben überwiegend qualitativen Charakter. Im Rahmen der durchgeführten Experimente konnten z.B. keine quantitativen Daten zur Lernkurve der einzelnen Techniken erhoben werden. Eine umfangreichere Studie als sie im Rahmen dieser Arbeit bisher durchgeführt werden konnte, könnte stärkere Hinweise auf quantitative Zusammenhänge liefern, beispielsweise Fehlerraten, Präzision, Geschwindigkeiten und ähnliches bei der Modellierung.

Die Kombination von Stift- und 3D-Interaktion und der Einsatz von Tangibles könnten weitere interessante Interaktionstechniken ermöglichen und wurden in dieser Arbeit nicht betrachtet.

Die in Kapitel 5 diskutierten Selektions- und Manipulationstechniken wurden im Rahmen dieser Arbeit lediglich im Rahmen von herausgelösten Testanwendungen untersucht, aber noch nicht in eine komplexere Modellierungsumgebung integriert.

Aktuelle Trends wie gekrümmte Displays und taktiles Feedback für Multi-touch eröffnen weitere interessante Möglichkeiten für zukünftige Untersuchungen.

Außerdem existieren noch weitere Modellierungsansätze, die auf ihr Potenzial für den Einsatz auf MT-Eingabegeräten untersucht werden sollten. Beispielsweise physikbasierte oder physik-inspirierte Ansätze auf Basis von Echtzeit-Physikengines oder verwandten Technologien.

Literaturverzeichnis

- [Akenine-Moller u. a. 2008] AKENINE-MOLLER, Tomas ; HAINES, Eric ; HOFFMAN, Naty: *Real-Time Rendering, Third Edition*. 3. AK Peters, Juli 2008. – URL <http://www.worldcat.org/isbn/1568814240>. – ISBN 1568814240
- [Aras u. a. 2011] ARAS, Hidir ; WALTHER-FRANKS, Benjamin ; HERRLICH, Marc ; RODACKER, Patrick ; MALAKA, Rainer: ElasticSteer – Navigating Large 3D Information Spaces via Touch or Mouse. In: DICKMANN, Lutz (Hrsg.) ; VOLKMANN, Gerald (Hrsg.) ; MALAKA, Rainer (Hrsg.) ; BOLL, Susanne (Hrsg.) ; KRÜGER, Antonio (Hrsg.) ; OLIVIER, Patrick (Hrsg.): *Smart Graphics* Bd. 6815. Berlin, Heidelberg : Springer Berlin / Heidelberg, 2011, Kap. 14, S. 138–141. – URL http://dx.doi.org/10.1007/978-3-642-22571-0_14. – ISBN 978-3-642-22570-3
- [Arun u. a. 1987] ARUN, K. S. ; HUANG, T. S. ; BLOSTEIN, S. D.: Least-Squares Fitting of Two 3-D Point Sets. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-9* (1987), Nr. 5, S. 698–700. – URL <http://dx.doi.org/10.1109/tpami.1987.4767965>
- [Aust 2010] AUST, Markus: *Rechts-Links-Handunterscheidung für Multitouchinteraktion auf der Basis von Diffuse-Illumination*. Bremen, Germany, University of Bremen, Diplomarbeit, Dezember 2010
- [Balakrishnan und Hinckley 1999] BALAKRISHNAN, Ravin ; HINCKLEY, Ken: The role of kinesthetic reference frames in two-handed input performance. In: *Proceedings of the 12th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 1999 (UIST '99), S. 171–178. – URL <http://dx.doi.org/10.1145/320719.322599>. – ISBN 1-58113-075-9

- [Balakrishnan und Kurtenbach 1999] BALAKRISHNAN, Ravin ; KURTENBACH, Gordon: Exploring bimanual camera control and object manipulation in 3D graphics interfaces. In: *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 1999, S. 56–62. – URL <http://dx.doi.org/10.1145/302979.302991>. – ISBN 0-201-48559-1
- [Barr 1984] BARR, Alan H.: Global and local deformations of solid primitives. In: *SIGGRAPH Comput. Graph.* 18 (1984), Januar, Nr. 3, S. 21–30. – URL <http://dx.doi.org/10.1145/964965.808573>. – ISSN 0097-8930
- [Benko u. a. 2006] BENKO, Hrvoje ; WILSON, Andrew D. ; BAUDISCH, Patrick: Precise selection techniques for multi-touch screens. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. New York, NY, USA : ACM, 2006 (CHI '06), S. 1263–1272. – URL <http://dx.doi.org/10.1145/1124772.1124963>. – ISBN 1-59593-372-7
- [Bier 1990] BIER, Eric A.: Snap-dragging in three dimensions. In: *Proceedings of the 1990 symposium on Interactive 3D graphics*. New York, NY, USA : ACM, 1990 (I3D '90), S. 193–204. – URL <http://dx.doi.org/10.1145/91385.91446>. – ISBN 0-89791-351-5
- [Bier u. a. 1993] BIER, Eric A. ; STONE, Maureen C. ; PIER, Ken ; BUXTON, William ; DEROSE, Tony D.: Toolglass and magic lenses: the see-through interface. In: *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*. New York, NY, USA : ACM, 1993 (SIGGRAPH '93), S. 73–80. – URL <http://dx.doi.org/10.1145/166117.166126>. – ISBN 0-89791-601-8
- [Bill und Lodha 1994] BILL, James R. ; LODHA, Suresh K.: Computer Sculpting of Polygonal Models Using Virtual Tools / University of California at Santa Cruz. Santa Cruz, CA, USA : University of California at Santa Cruz, 1994. – Forschungsbericht. – URL <http://portal.acm.org/citation.cfm?id=902839>
- [Blinn 1982] BLINN, James F.: A Generalization of Algebraic Surface Drawing. In: *ACM Trans. Graph.* 1 (1982), Juli, Nr. 3, S. 235–256. – URL <http://dx.doi.org/10.1145/357306.357310>. – ISSN 0730-0301

- [Bloomenthal 1988] BLOOMENTHAL, J.: Polygonization of implicit surfaces. In: *Comput. Aided Geom. Des.* 5 (1988), November, Nr. 4, S. 341–355. – URL [http://dx.doi.org/10.1016/0167-8396\(88\)90013-1](http://dx.doi.org/10.1016/0167-8396(88)90013-1). – ISSN 0167-8396
- [Bloomenthal 1994] BLOOMENTHAL, Jules: An Implicit Surface Polygonizer. In: *Graphics Gems IV*. Morgan Kaufmann, 1994, S. 324–349. – URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.24.3429>
- [Bloomenthal und Wyvill 1990] BLOOMENTHAL, Jules ; WYVILL, Brian: Interactive techniques for implicit modeling. In: *SIGGRAPH Comput. Graph.* 24 (1990), Februar, Nr. 2, S. 109–116. – URL <http://dx.doi.org/10.1145/91394.91427>. – ISSN 0097-8930
- [Bowman u. a. 2004] BOWMAN, Doug A. ; KRUIJFF, Ernst ; LAVIOLA, Joseph J. ; POUYREV, Ivan: *3D User Interfaces: Theory and Practice*. 1. Addison-Wesley Professional, August 2004. – URL <http://www.worldcat.org/isbn/0201758679>. – ISBN 0201758679
- [Buxton 2007] BUXTON, Bill: *Multi-Touch Systems that I Have Known and Loved*. Website. Januar 2007. – URL <http://www.billbuxton.com/multitouch0verview.html>
- [Buxton und Myers 1986] BUXTON, W. ; MYERS, B.: A study in two-handed input. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 1986 (CHI '86), S. 321–326. – URL <http://dx.doi.org/10.1145/22627.22390>. – ISBN 0-89791-180-6
- [Cao u. a. 2008] CAO, Xiang ; WILSON, Andrew D. ; BALAKRISHNAN, Ravin ; HINCKLEY, Ken ; HUDSON, Scott E.: ShapeTouch: Leveraging contact shape on interactive surfaces. In: *2008 IEEE International Workshop on Horizontal Interactive Human Computer Systems (TABLETOP)*, IEEE, Oktober 2008, S. 129–136. – URL <http://dx.doi.org/10.1109/tabletop.2008.4660195>. – ISBN 978-1-4244-2897-7
- [Cutler u. a. 1997] CUTLER, Lawrence D. ; FRÖHLICH, Bernd ; HANRAHAN, Pat: Two-handed direct manipulation on the responsive workbench. In:

- SI3D '97: Proceedings of the 1997 symposium on Interactive 3D graphics*. New York, NY, USA : ACM, 1997, S. 107–114. – URL <http://dx.doi.org/10.1145/253284.253315>. – ISBN 0-89791-884-3
- [Dang u. a. 2009] DANG, Chi T. ; STRAUB, Martin ; ANDRÉ, Elisabeth: Hand Distinction for Multi-Touch Tabletop Interaction. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, ACM, 2009
- [Dewaele und Cani 2004] DEWAELE, G. ; CANI, M.: Interactive global and local deformations for virtual clay. In: *Graphical Models* 66 (2004), November, Nr. 6, S. 352–369. – URL <http://dx.doi.org/10.1016/j.gmod.2004.06.008>. – ISSN 15240703
- [Dietz und Leigh 2001] DIETZ, Paul ; LEIGH, Darren: DiamondTouch: a multi-user touch technology. In: *Proceedings of the 14th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2001 (UIST '01), S. 219–226. – URL <http://dx.doi.org/10.1145/502348.502389>. – ISBN 1-58113-438-X
- [Epps u. a. 2006] EPPS, Julien ; LICHMAN, Serge ; WU, Mike: A study of hand shape use in tabletop gesture interaction. In: *CHI '06 extended abstracts on Human factors in computing systems*. New York, NY, USA : ACM, 2006 (CHI '06), S. 748–753. – URL <http://dx.doi.org/10.1145/1125451.1125601>. – ISBN 1-59593-298-4
- [Erleben u. a. 2005] ERLEBEN, Kenny ; SPORRING, Jon ; HENRIKSEN, Knud ; DOHLMAN, Kenrik: *Physics-based Animation (Graphics Series)*. Rockland, MA, USA : Charles River Media, Inc., 2005. – URL <http://portal.acm.org/citation.cfm?id=1051390>. – ISBN 1584503807
- [Esenther und Ryall 2006] ESENTHER, Alan ; RYALL, Kathy: Fluid DT-Mouse: better mouse support for touch-based interactions. In: *AVI '06: Proceedings of the working conference on Advanced visual interfaces*. New York, NY, USA : ACM, 2006, S. 112–115. – URL <http://dx.doi.org/10.1145/1133265.1133289>. – ISBN 1-59593-353-0
- [Everitt u. a. 2005] EVERITT, Katherine ; SHEN, Chia ; RYALL, Kathy ; FORLINES, Clifton: Modal spaces: spatial multiplexing to mediate direct-touch

- input on large displays. In: *CHI '05 extended abstracts on Human factors in computing systems*. New York, NY, USA : ACM, 2005 (CHI EA '05), S. 1359–1362. – URL <http://dx.doi.org/10.1145/1056808.1056916>. – ISBN 1-59593-002-7
- [Ferley u. a. 2001] FERLEY, Eric ; CANI, Marie-Paule ; GASCUEL, Jean-Dominique: Resolution Adaptive Volume Sculpting. In: *Graphical Models* 63 (2001), November, Nr. 6, S. 459–478. – URL <http://dx.doi.org/10.1006/gmod.2001.0558>. – ISSN 15240703
- [Foley u. a. 1995] FOLEY, James D. ; DAM, Andries van ; FEINER, Steven K. ; HUGHES, John F.: *Computer Graphics: Principles and Practice in C (2nd Edition)*. 2. Addison-Wesley Professional, August 1995. – URL <http://www.worldcat.org/isbn/0201848406>. – ISBN 0201848406
- [Forlines u. a. 2006] FORLINES, Clifton ; VOGEL, Daniel ; BALAKRISHNAN, Ravin: HybridPointing: fluid switching between absolute and relative pointing with a direct input device. In: *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2006, S. 211–220. – URL <http://dx.doi.org/10.1145/1166253.1166286>. – ISBN 1-59593-313-1
- [Forlines u. a. 2007] FORLINES, Clifton ; WIGDOR, Daniel ; SHEN, Chia ; BALAKRISHNAN, Ravin: Direct-touch vs. mouse input for tabletop displays. In: *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2007, S. 647–656. – URL <http://dx.doi.org/10.1145/1240624.1240726>. – ISBN 978-1-59593-593-9
- [Freeman u. a. 2009] FREEMAN, Dustin ; BENKO, Hrvoje ; MORRIS, Meredith R. ; WIGDOR, Daniel: ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. New York, NY, USA : ACM, 2009 (ITS '09), S. 165–172. – URL <http://dx.doi.org/10.1145/1731903.1731935>. – ISBN 978-1-60558-733-2
- [Frisch u. a. 2009] FRISCH, Mathias ; HEYDEKORN, Jens ; DACHSELT, Raimund: Investigating Multi-Touch and Pen Gestures for Diagram Editing on

- Interactive Surfaces. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, 2009
- [Frohlich u. a. 2000] FROHLICH, B. ; TRAMBEREND, H. ; BEERS, A. ; AGRAWALA, M. ; BARAFF, D.: Physically-based manipulation on the Responsive Workbench. In: *IEEE Virtual Reality 2000* Bd. 0. Los Alamitos, CA, USA : IEEE Comput. Soc, 2000, S. 5–11. – URL <http://dx.doi.org/10.1109/vr.2000.840357>. – ISBN 0-7695-0478-7
- [Galyean und Hughes 1991] GALYEAN, Tinsley A. ; HUGHES, John F.: Sculpting: an interactive volumetric modeling technique. In: *SIGGRAPH Comput. Graph.* 25 (1991), Juli, Nr. 4, S. 267–274. – URL <http://dx.doi.org/10.1145/127719.122747>. – ISSN 0097-8930
- [Gibson und Mirtich 1997] GIBSON, Sarah F. F. ; MIRTICH, Brian: A Survey of Deformable Modeling in Computer Graphics / Mitsubishi Electric Research Laboratories. Mitsubishi Electric Research Laboratories, 1997 (TR-97). – Forschungsbericht. – URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.6.4155>
- [Goble u. a. 1995] GOBLE, J. C. ; HINCKLEY, K. ; PAUSCH, R. ; SNELL, J. W. ; KASSELL, N. F.: Two-handed spatial interface tools for neurosurgical planning. In: *Computer* 28 (1995), Juli, Nr. 7, S. 20–26. – URL <http://dx.doi.org/10.1109/2.391037>. – ISSN 0018-9162
- [Goldstein 2009] GOLDSTEIN, E. B.: *Sensation and Perception (with Virtual Lab Manual CD-ROM)*. 8. Wadsworth Publishing, Februar 2009. – URL <http://www.worldcat.org/isbn/9780495601494>. – ISBN 0495601497
- [Gribnau und Hennessey 1998] GRIBNAU, Maarten W. ; HENNESSEY, James M.: Comparing single- and two-handed 3D input for a 3D object assembly task. In: *CHI 98 conference summary on Human factors in computing systems*. New York, NY, USA : ACM, 1998 (CHI '98), S. 233–234. – URL <http://dx.doi.org/10.1145/286498.286720>. – ISBN 1-58113-028-7
- [Grossman und Wigdor 2007] GROSSMAN, T. ; WIGDOR, D.: Going Deeper: a Taxonomy of 3D on the Tabletop. In: *2007 IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, IEEE, Oktober 2007,

S. 137–144. – URL <http://dx.doi.org/10.1109/tabletop.2007.18>. – ISBN 0-7695-3013-3

[Grossman und Balakrishnan 2005] GROSSMAN, Tovi ; BALAKRISHNAN, Ravin: The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2005 (CHI '05), S. 281–290. – URL <http://dx.doi.org/10.1145/1054972.1055012>. – ISBN 1-58113-998-5

[Grossman u. a. 2009] GROSSMAN, Tovi ; BAUDISCH, Patrick ; HINCKLEY, Ken: Handle Flags: efficient and flexible selections for inking applications. In: *GI '09: Proceedings of Graphics Interface 2009*. Toronto, Ont., Canada, Canada : Canadian Information Processing Society, 2009, S. 167–174. – URL <http://portal.acm.org/citation.cfm?id=1555880.1555918>. – ISBN 978-1-56881-470-4

[Grossman u. a. 2004] GROSSMAN, Tovi ; WIGDOR, Daniel ; BALAKRISHNAN, Ravin: Multi-finger gestural interaction with 3d volumetric displays. In: *Proceedings of the 17th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2004 (UIST '04), S. 61–70. – URL <http://dx.doi.org/10.1145/1029632.1029644>. – ISBN 1-58113-957-8

[Guiard 1987] GUIARD, Yves: Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. In: *Journal of Motor Behaviour* 19 (1987), S. 486–517. – URL <http://cogprints.org/625/>

[Gunn u. a. 2009] GUNN, Tyler J. ; IRANI, Pourang ; ANDERSON, John: An evaluation of techniques for selecting moving targets. In: *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*. New York, NY, USA : ACM, 2009 (CHI EA '09), S. 3329–3334. – URL <http://dx.doi.org/10.1145/1520340.1520481>. – ISBN 978-1-60558-247-4

[Hahne u. a. 2009] HAHNE, Uwe ; SCHILD, Jonas ; ELSTNER, Stefan ; ALEXA, Marc: Multi-touch focus+context sketch-based interaction. In: *SBIM '09*:

-
- Proceedings of the 6th Eurographics Symposium on Sketch-Based Interfaces and Modeling*. New York, NY, USA : ACM, 2009, S. 77–83. – URL <http://dx.doi.org/10.1145/1572741.1572755>. – ISBN 978-1-60558-602-1
- [Han 2005] HAN, Jefferson Y.: Low-cost multi-touch sensing through frustrated total internal reflection. In: *Proceedings of the 18th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2005 (UIST '05), S. 115–118. – URL <http://dx.doi.org/10.1145/1095034.1095054>. – ISBN 1-59593-271-2
- [Hancock u. a. 2006] HANCOCK, M. ; CARPENDALE, S. ; VERNIER, F. ; WIGDOR, D. ; SHEN, Chia: Rotation and Translation Mechanisms for Tabletop Interaction. In: *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)*, IEEE, Januar 2006, S. 79–88. – URL <http://dx.doi.org/10.1109/tabletop.2006.26>. – ISBN 0-7695-2494-X
- [Hancock u. a. 2007] HANCOCK, Mark ; CARPENDALE, Sheelagh ; COCKBURN, Andy: Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2007 (CHI '07), S. 1147–1156. – URL <http://dx.doi.org/10.1145/1240624.1240798>. – ISBN 978-1-59593-593-9
- [Hancock u. a. 2009a] HANCOCK, Mark ; CATE, Thomas T. ; CARPENDALE, Sheelagh: Sticky Tools: Full 6DOF Force-Based Interaction for Multi-Touch Tables. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, 2009
- [Hancock u. a. 2009b] HANCOCK, Mark ; HILLIGES, Otmar ; COLLINS, Christopher ; BAUR, Dominikus ; CARPENDALE, Sheelagh: Exploring Tangible and Direct Touch Interfaces for Manipulating 2D and 3D Information on a Digital Table. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, 2009
- [Hancock u. a. 2009c] HANCOCK, Mark ; NACENTA, Miguel ; GUTWIN, Carl ; CARPENDALE, Sheelagh: The Effects of Changing Projection Geometry on the Interpretation of 3D Orientation on Tabletops. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, 2009

- [Herrlich u. a. 2012a] HERRLICH, Marc ; BRAUN, Andreas ; MALAKA, Rainer: Towards Bimanual Control for Virtual Sculpting. In: *Mensch und Computer 2012*, Oldenbourg Verlag, September 2012
- [Herrlich u. a. 2008] HERRLICH, Marc ; KRAUSE, Markus ; SCHWARTEN, Lasse ; TEICHERT, Jens ; WALTHER-FRANKS, Benjamin: Multitouch Interface Metaphors for 3D Modeling. In: *IEEE Tabletops and Interactive Surfaces 2008*. Los Alamitos, CA, USA : IEEE Computer Society, 2008
- [Herrlich u. a. 2011a] HERRLICH, Marc ; WALTHER-FRANKS, Benjamin ; MALAKA, Rainer: Integrated Rotation and Translation for 3D Manipulation on Multi-Touch Interactive Surfaces. In: *Proceedings of Smart Graphics 2011*, Springer, Juli 2011
- [Herrlich u. a. 2012b] HERRLICH, Marc ; WALTHER-FRANKS, Benjamin ; MALAKA, Rainer: Daten zum Anfassen: Be-greifen mit Interaktiven Bildschirmen. In: ROBBEN, Bernard (Hrsg.) ; SCHELHOWE, Heidi (Hrsg.): *Be-greifbare Interaktionen: Der allgegenwärtige Computer: Touchscreens, Wearables, Tangibles und Ubiquitous Computing*. Transcript Verlag, März 2012. – ISBN 978-3-8376-2005-4
- [Herrlich u. a. 2011b] HERRLICH, Marc ; WALTHER-FRANKS, Benjamin ; SCHRÖDER-KROLL, Roland ; HOLTHUSEN, Jan ; MALAKA, Rainer: Proxy-Based Selection for Occluded and Dynamic Objects. In: *Proceedings of Smart Graphics 2011*, Springer, Juli 2011
- [Herrlich u. a. 2011c] HERRLICH, Marc ; WALTHER-FRANKS, Benjamin ; WEIDNER, Daniel ; MALAKA, Rainer: Designing for Social Interaction in Collaborative Games on Large Multi-Touch Displays. In: *Proceedings of Mensch und Computer 2011*, Oldenbourg, September 2011
- [Hilliges u. a. 2009] HILLIGES, Otmar ; IZADI, Shahram ; WILSON, Andrew D. ; HODGES, Steve ; MENDOZA, Armando G. ; BUTZ, Andreas: Interactions in the air: adding further depth to interactive tabletops. In: *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2009, S. 139–148. – URL <http://dx.doi.org/10.1145/1622176.1622203>. – ISBN 978-1-60558-745-5

- [Hinckley u. a. 1998a] HINCKLEY, Ken ; CZERWINSKI, Mary ; SINCLAIR, Mike: Interaction and modeling techniques for desktop two-handed input. In: *Proceedings of the 11th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 1998 (UIST '98), S. 49–58. – URL <http://dx.doi.org/10.1145/288392.288572>. – ISBN 1-58113-034-1
- [Hinckley u. a. 1994] HINCKLEY, Ken ; PAUSCH, Randy ; GOBLE, John C. ; KASSELL, Neal F.: A survey of design issues in spatial input. In: *Proceedings of the 7th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 1994 (UIST '94), S. 213–222. – URL <http://dx.doi.org/10.1145/192426.192501>. – ISBN 0-89791-657-3
- [Hinckley u. a. 1998b] HINCKLEY, Ken ; PAUSCH, Randy ; PROFFITT, Dennis ; KASSELL, Neal F.: Two-handed virtual manipulation. In: *ACM Trans. Comput.-Hum. Interact.* 5 (1998), September, Nr. 3, S. 260–302. – URL <http://dx.doi.org/10.1145/292834.292849>. – ISSN 1073-0516
- [Hinckley und Sinclair 1999] HINCKLEY, Ken ; SINCLAIR, Mike: Touch-sensing input devices. In: *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*. New York, NY, USA : ACM, 1999 (CHI '99), S. 223–230. – URL <http://dx.doi.org/10.1145/302979.303045>. – ISBN 0-201-48559-1
- [Hinckley und Wigdor 2012] HINCKLEY, Ken ; WIGDOR, Daniel: Input Technologies and Techniques. In: JACKO, Julie A. (Hrsg.): *The Human-computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Third. CRC Press Inc, Mai 2012, Kap. 9. – ISBN 978-1439829431
- [Hodges u. a. 2007] HODGES, Steve ; IZADI, Shahram ; BUTLER, Alex ; RRUSTEMI, Alban ; BUXTON, Bill: ThinSight: versatile multi-touch sensing for thin form-factor displays. In: *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2007, S. 259–268. – URL <http://dx.doi.org/10.1145/1294211.1294258>. – ISBN 978-1-59593-679-2
- [Hsu u. a. 1992] HSU, William M. ; HUGHES, John F. ; KAUFMAN, Henry: Direct manipulation of free-form deformations. In: *Proceedings of the 19th*

- annual conference on Computer graphics and interactive techniques*. New York, NY, USA : ACM, 1992 (SIGGRAPH '92), S. 177–184. – URL <http://dx.doi.org/10.1145/133994.134036>. – ISBN 0-89791-479-1
- [Hurtienne und Naumann 2010] HURTIENNE, J. ; NAUMANN, A.: QUESI – A questionnaire for measuring the subjective consequences of intuitive use. In: PORZEL, R. (Hrsg.) ; SEBANZ, N. (Hrsg.) ; SPITZER, M. (Hrsg.): *Interdisciplinary College 2010. Focus Theme: Play, Act and Learn* Fraunhofer Gesellschaft (Veranst.), 2010, S. 536+
- [Igarashi u. a. 1999] IGARASHI, Takeo ; MATSUOKA, Satoshi ; TANAKA, Hidehiko: Teddy: a sketching interface for 3D freeform design. In: *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1999 (SIGGRAPH '99), S. 409–416. – URL <http://dx.doi.org/10.1145/311535.311602>. – ISBN 0-201-48560-5
- [Izadi u. a. 2007] IZADI, Shahram ; HODGES, Steve ; BUTLER, Alex ; RRUSTEMI, Alban ; BUXTON, Bill: ThinSight: integrated optical multi-touch sensing through thin form-factor displays. In: *EDT '07: Proceedings of the 2007 workshop on Emerging displays technologies*. New York, NY, USA : ACM, 2007, S. 6+. – URL <http://dx.doi.org/10.1145/1278240.1278246>. – ISBN 978-1-59593-669-1
- [Jacob u. a. 1994] JACOB, Robert J. K. ; SIBERT, Linda E. ; MCFARLANE, Daniel C. ; MULLEN, M. P.: Integrality and separability of input devices. In: *ACM Trans. Comput.-Hum. Interact.* 1 (1994), März, Nr. 1, S. 3–26. – URL <http://dx.doi.org/10.1145/174630.174631>. – ISSN 1073-0516
- [Jones u. a. 2006] JONES, Mark W. ; BAERENTZEN, J. A. ; SRAMEK, Milos: 3D Distance Fields: A Survey of Techniques and Applications. In: *IEEE Transactions on Visualization and Computer Graphics* 12 (2006), Juli, Nr. 4, S. 581–599. – URL <http://dx.doi.org/10.1109/tvcg.2006.56>. – ISSN 1077-2626
- [Jordà u. a. 2005] JORDÀ, Sergi ; KALTENBRUNNER, Martin ; GEIGER, Günter ; BENCINA, Ross: The reacTable. In: *Proceedings of the International*

- Computer Music Conference (ICMC 2005)*, International Computer Music Association, 2005, S. 579–582. – URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.62.6306>
- [Joshi u. a. 2010] JOSHI, Amit ; ROBERTSON, Glen ; WUENSCHKE, Burkhard ; PLIMMER, Beryl: Bubbleworld Builder – 3D Modeling Using Twotouch And Sketch Interaction. In: *Proceedings of the 5th International Conference on Computer Graphics Theory and Applications (GRAPP 2010)*, 2010, S. 116–122
- [Kaltenbrunner und Bencina 2007] KALTENBRUNNER, Martin ; BENCINA, Ross: reacTIVision: a computer-vision framework for table-based tangible interaction. In: *Proceedings of the 1st international conference on Tangible and embedded interaction*. New York, NY, USA : ACM, 2007 (TEI '07), S. 69–74. – URL <http://dx.doi.org/10.1145/1226969.1226983>. – ISBN 978-1-59593-619-6
- [Kaltenbrunner u. a. 2005] KALTENBRUNNER, Martin ; BOVERMANN, Till ; BENCINA, Ross ; COSTANZA, Enrico: TUIO - A Protocol for Table Based Tangible User Interfaces. In: *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW 2005)*. Vannes, France : Springer, 2005
- [Karam und Schraefel 2005] KARAM, Maria ; SCHRAEFEL, M. C.: A taxonomy of gestures in human computer interactions / University of Southampton. Southampton, UK : University of Southampton, August 2005. – Forschungsbericht. – URL <http://eprints.soton.ac.uk/id/eprint/261149>
- [Kin u. a. 2009] KIN, Kenrick ; AGRAWALA, Maneesh ; DEROSE, Tony: Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation. In: *Proceedings of Graphics Interface 2009*. Toronto, Ont., Canada, Canada : Canadian Information Processing Society, 2009 (GI '09), S. 119–124. – URL <http://portal.acm.org/citation.cfm?id=1555910>. – ISBN 978-1-56881-470-4
- [Krause u. a. 2008] KRAUSE, Markus ; HERRLICH, Marc ; SCHWARTEN, Lasse ; TEICHERT, Jens ; WALTHER-FRANKS, Benjamin: Multitouch Motion

- Capturing. In: *IEEE Tabletops and Interactive Surfaces 2008*. Los Alamitos, CA, USA : IEEE Computer Society, 2008
- [Krueger u. a. 1985] KRUEGER, Myron W. ; GIONFRIDDO, Thomas ; HINRICHSEN, Katrin: VIDEOPLACE – An Artificial Reality. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 1985 (CHI '85), S. 35–40. – URL <http://dx.doi.org/10.1145/317456.317463>. – ISBN 0-89791-149-0
- [Kruger u. a. 2005] KRUGER, Russell ; CARPENDALE, Sheelagh ; SCOTT, Stacey D. ; TANG, Anthony: Fluid integration of rotation and translation. In: *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2005, S. 601–610. – URL <http://dx.doi.org/10.1145/1054972.1055055>. – ISBN 1-58113-998-5
- [Leganchuk u. a. 1998] LEGANCHUK, Andrea ; ZHAI, Shumin ; BUXTON, William: Manual and cognitive benefits of two-handed input: an experimental study. In: *ACM Trans. Comput.-Hum. Interact.* 5 (1998), Dezember, Nr. 4, S. 326–359. – URL <http://dx.doi.org/10.1145/300520.300522>. – ISSN 1073-0516
- [MacCracken und Joy 1996] MACCRACKEN, Ron ; JOY, Kenneth I.: Free-form deformations with lattices of arbitrary topology. In: *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. New York, NY, USA : ACM, 1996 (SIGGRAPH '96), S. 181–188. – URL <http://dx.doi.org/10.1145/237170.237247>. – ISBN 0-89791-746-4
- [Markosian u. a. 1999] MARKOSIAN, Lee ; COHEN, Jonathan M. ; CRULLI, Thomas ; HUGHES, John: Skin: a constructive approach to modeling free-form shapes. In: *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1999 (SIGGRAPH '99), S. 393–400. – URL <http://dx.doi.org/10.1145/311535.311595>. – ISBN 0-201-48560-5
- [Marquardt u. a. 2010] MARQUARDT, Nicolai ; KIEMER, Johannes ; GREENBERG, Saul: What caused that touch?: expressive interaction with a surface through fiduciary-tagged gloves. In: *ACM International Conference on Interactive Tabletops and Surfaces*. New York, NY, USA : ACM, 2010 (ITS

- '10), S. 139–142. – URL <http://dx.doi.org/10.1145/1936652.1936680>.
– ISBN 978-1-4503-0399-6
- [Martinet u. a. 2010] MARTINET, Anthony ; CASIEZ, Gery ; GRISONI, Laurent: The design and evaluation of 3D positioning techniques for multi-touch displays. In: *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, IEEE, März 2010, S. 115–118. – URL <http://dx.doi.org/10.1109/3dui.2010.5444709>. – ISBN 978-1-4244-6846-1
- [Matejka u. a. 2009] MATEJKA, Justin ; GROSSMAN, Tovi ; LO, Jessica ; FITZMAURICE, George: The design and evaluation of multi-finger mouse emulation techniques. In: *Proceedings of the 27th international conference on Human factors in computing systems*. New York, NY, USA : ACM, 2009 (CHI '09), S. 1073–1082. – URL <http://dx.doi.org/10.1145/1518701.1518865>. – ISBN 978-1-60558-246-7
- [Micire u. a. 2009] MICIRE, Mark ; DESAI, Munjal ; COURTEMANCHE, Amanda ; TSUI, Katherine M. ; YANCO, Holly A.: Analysis of Natural Gestures for Controlling Robot Teams on Multi-touch Tablet Surfaces. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, 2009
- [Milgram und Kishino 1994] MILGRAM, Paul ; KISHINO, Fumio: A Taxonomy of Mixed Reality Visual Displays. In: *IEICE Transactions on Information Systems* 77 (1994), S. 1321–1329. – URL http://vered.rose.utoronto.ca/people/paul_dir/IEICE94/ieice.html
- [Mizuno u. a. 1998] MIZUNO, Shinji ; OKADA, Minoru ; TORIWAKI, Jun-ichiro: Virtual sculpting and virtual woodcut printing. In: *The Visual Computer* 14 (1998), Juni, Nr. 2, S. 39–51. – URL <http://dx.doi.org/10.1007/s003710050122>. – ISSN 01782789
- [Moscovich 2007] MOSCOVICH, Tomer: *Principles and applications of multi-touch interaction*. Providence, RI, USA, Brown University, Dissertation, 2007. – URL <http://portal.acm.org/citation.cfm?id=1354812>
- [Moscovich 2009] MOSCOVICH, Tomer: Contact area interaction with sliding widgets. In: *UIST '09: Proceedings of the 22nd annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2009,

- S. 13–22. – URL <http://dx.doi.org/10.1145/1622176.1622181>. – ISBN 978-1-60558-745-5
- [Moscovich und Hughes 2006] MOSCOVICH, Tomer ; HUGHES, John F.: Multi-finger cursor techniques. In: *GI '06: Proceedings of Graphics Interface 2006*. Toronto, Ont., Canada, Canada : Canadian Information Processing Society, 2006, S. 1–7. – URL <http://portal.acm.org/citation.cfm?id=1143081>. – ISBN 1-56881-308-2
- [Moscovich und Hughes 2008] MOSCOVICH, Tomer ; HUGHES, John F.: Indirect mappings of multi-touch input using one and two hands. In: *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2008 (CHI '08), S. 1275–1284. – URL <http://dx.doi.org/10.1145/1357054.1357254>. – ISBN 978-1-60558-011-1
- [Nacenta u. a. 2009] NACENTA, Miguel A. ; BAUDISCH, Patrick ; BENKO, Hrvoje ; WILSON, Andy: Separability of spatial manipulations in multi-touch interfaces. In: *GI '09: Proceedings of Graphics Interface 2009*. Toronto, Ont., Canada, Canada : Canadian Information Processing Society, 2009, S. 175–182. – URL <http://portal.acm.org/citation.cfm?id=1555919>. – ISBN 978-1-56881-470-4
- [Nealen u. a. 2009] NEALEN, Andrew ; PETT, Justus ; ALEXA, Marc ; IGARASHI, Takeo: GridMesh: Fast and high quality 2D Mesh generation for interactive 3D shape modeling. In: *2009 IEEE International Conference on Shape Modeling and Applications (SMI)*, IEEE, Juni 2009, S. 155–162. – URL <http://dx.doi.org/10.1109/smi.2009.5170143>. – ISBN 978-1-4244-4069-6
- [Nealen 2007] NEALEN, Andrew V.: *Interfaces and Algorithms for the Creation, Modification, and Optimization of Surface Meshes*, TU Berlin, Dissertation, 2007
- [Norman 2002] NORMAN, Donald A.: *The design of everyday things*. Basic Books, September 2002. – URL <http://www.worldcat.org/isbn/0465067107>. – ISBN 0465067107

- [North u. a. 2009] NORTH, Chris ; DWYER, Tim ; LEE, Bongshin ; FISHER, Danyel ; ISENBERG, Petra ; ROBERTSON, George ; INKPEN, Kori: Understanding Multi-touch Manipulation for Surface Computing. In: HUTCHINSON, David (Hrsg.) ; KANADE, Takeo (Hrsg.) ; KITTLER, Josef (Hrsg.) ; KLEINBERG, Jon M. (Hrsg.) ; MATTERN, Friedemann (Hrsg.) ; MITCHELL, John C. (Hrsg.) ; NAOR, Moni (Hrsg.) ; NIERSTRASZ, Oscar (Hrsg.) ; PANDU RANGAN, C. (Hrsg.) ; STEFFEN, Bernhard (Hrsg.) ; SUDAN, Madhu (Hrsg.) ; TERZOPOULOS, Demetri (Hrsg.) ; TYGAR, Doug (Hrsg.) ; VARDI, Moshe Y. (Hrsg.) ; WEIKUM, Gerhard (Hrsg.) ; GROSS, Tom (Hrsg.) ; GULLIKSEN, Jan (Hrsg.) ; KOTZÉ, Paula (Hrsg.) ; OESTREICHER, Lars (Hrsg.) ; PALANQUE, Philippe (Hrsg.) ; PRATES, Raquel O. (Hrsg.) ; WINCKLER, Marco (Hrsg.): *Human-Computer Interaction – INTERACT 2009* Bd. 5727. Berlin, Heidelberg : Springer-Verlag, 2009, S. 236–249. – URL http://dx.doi.org/10.1007/978-3-642-03658-3_31. – ISBN 978-3-642-03657-6
- [Olsen 2007] OLSEN, Dan R.: Evaluating user interface systems research. In: *Proceedings of the 20th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2007 (UIST '07), S. 251–258. – URL <http://dx.doi.org/10.1145/1294211.1294256>. – ISBN 978-1-59593-679-0
- [Parent 1977] PARENT, Richard E.: A system for sculpting 3-D data. In: *Proceedings of the 4th annual conference on Computer graphics and interactive techniques*. New York, NY, USA : ACM, 1977 (SIGGRAPH '77), S. 138–147. – URL <http://dx.doi.org/10.1145/563858.563884>
- [Parviainen u. a. 2004] PARVIAINEN, Jyrki ; SAINIO, Nina ; RAISAMO, Roope: Perceiving Tools in 3D Sculpting Computer Human Interaction. In: MASOODIAN, Masood (Hrsg.) ; JONES, Steve (Hrsg.) ; ROGERS, Bill (Hrsg.): *Computer Human Interaction, 6th Asia Pacific Conference, APCHI 2004* Bd. 3101. Berlin, Heidelberg : Springer Berlin / Heidelberg, 2004, S. 328–337. – URL http://dx.doi.org/10.1007/978-3-540-27795-8_33. – ISBN 978-3-540-22312-2
- [Patten u. a. 2001] PATTEN, James ; ISHII, Hiroshi ; HINES, Jim ; PANGARO,

- Gian: Sensetable: a wireless object tracking platform for tangible user interfaces. In: *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2001, S. 253–260. – URL <http://dx.doi.org/10.1145/365024.365112>. – ISBN 1-58113-327-8
- [Ramos u. a. 2006] RAMOS, Gonzalo ; ROBERTSON, George ; CZERWINSKI, Mary ; TAN, Desney ; BAUDISCH, Patrick ; HINCKLEY, Ken ; AGRAWALA, Maneesh: Tumble! Splat! helping users access and manipulate occluded content in 2D drawings. In: *Proceedings of the working conference on Advanced visual interfaces*. New York, NY, USA : ACM, 2006 (AVI '06), S. 428–435. – URL <http://dx.doi.org/10.1145/1133265.1133351>. – ISBN 1-59593-353-0
- [Raviv und Elber 2000] RAVIV, A. ; ELBER, G.: Three-dimensional freeform sculpting via zero sets of scalar trivariate functions. In: *Computer-Aided Design* 32 (2000), August, Nr. 8-9, S. 513–526. – URL [http://dx.doi.org/10.1016/s0010-4485\(00\)00039-7](http://dx.doi.org/10.1016/s0010-4485(00)00039-7). – ISSN 00104485
- [Reisman u. a. 2009] REISMAN, Jason L. ; DAVIDSON, Philip L. ; HAN, Jefferson Y.: A screen-space formulation for 2D and 3D direct manipulation. In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2009 (UIST '09), S. 69–78. – URL <http://dx.doi.org/10.1145/1622176.1622190>. – ISBN 978-1-60558-745-5
- [Rekimoto 2002] REKIMOTO, Jun: SmartSkin: an infrastructure for freehand manipulation on interactive surfaces. In: *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2002, S. 113–120. – URL <http://dx.doi.org/10.1145/503376.503397>. – ISBN 1-58113-453-3
- [Rosenberg und Perlin 2009] ROSENBERG, Ilya ; PERLIN, Ken: The UnMousePad: an interpolating multi-touch force-sensing input pad. In: *SIGGRAPH '09: ACM SIGGRAPH 2009 papers*. New York, NY, USA : ACM, 2009, S. 1–9. – URL <http://dx.doi.org/10.1145/1576246.1531371>. – ISBN 978-1-60558-726-4

- [Roudaut u. a. 2008] ROUDAUT, Anne ; HUOT, Stéphane ; LECOLINET, Eric: TapTap and MagStick: improving one-handed target acquisition on small touch-screens. In: *Proceedings of the working conference on Advanced visual interfaces*. New York, NY, USA : ACM, 2008 (AVI '08), S. 146–153. – URL <http://dx.doi.org/10.1145/1385569.1385594>. – ISBN 978-1-60558-141-5
- [Roudaut u. a. 2009] ROUDAUT, Anne ; LECOLINET, Eric ; GUIARD, Yves: MicroRolls: expanding touch-screen input vocabulary by distinguishing rolls vs. slides of the thumb. In: *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*. New York, NY, USA : ACM, 2009, S. 927–936. – URL <http://dx.doi.org/10.1145/1518701.1518843>. – ISBN 978-1-60558-246-7
- [Rungjiratananon u. a. 2008] RUNGJIRATANANON, Witawat ; SZEGO, Zoltan ; KANAMORI, Yoshihiro ; NISHITA, Tomoyuki: Real-time Animation of Sand-Water Interaction. In: *Computer Graphics Forum* 27 (2008), Nr. 7, S. 1887–1893. – URL <http://dx.doi.org/10.1111/j.1467-8659.2008.01336.x>. – ISSN 1467-8659
- [Schmidt und Lanz 2003] SCHMIDT, Hans-Martin ; LANZ, Ulrich: *Chirurgische Anatomie der Hand*. 2., überarb. u. aktualis. A. Thieme Georg Verlag, Juli 2003. – URL <http://www.worldcat.org/isbn/3131306920>. – ISBN 3131306920
- [Schöning u. a. 2008] SCHÖNING, Johannes ; BRANDL, Peter ; DAIBER, Florian ; ECHTLER, Florian ; HILLIGES, Otmar ; HOOK, Jonathan ; LÖCHTEFELD, Markus ; MOTAMEDI, Nima ; MULLER, Laurence ; OLIVIER, Patrick ; ROTH, Tim ; ZADOW, Ulrich von: Multi-Touch Surfaces: A Technical Guide / University of Münster. 2008. – Forschungsbericht
- [Schröder-Kroll u. a. 2008] SCHRÖDER-KROLL, Roland ; BLOM, Kristopher ; BECKHAUS, Steffi: Interaction Techniques for Dynamic Virtual Environments. In: SCHUMANN, Marco (Hrsg.) ; KUHLEN, Torsten (Hrsg.): *Virtuelle und Erweiterte Realität, 5. Workshop der GI-Fachgruppe VR/AR*, Shaker Verlag, 2008

- [Sederberg und Parry 1986] SEDERBERG, Thomas W. ; PARRY, Scott R.: Free-form deformation of solid geometric models. In: *SIGGRAPH Comput. Graph.* 20 (1986), August, Nr. 4, S. 151–160. – URL <http://dx.doi.org/10.1145/15922.15903>. – ISSN 0097-8930
- [Shen 2006] SHEN, Chia: Multi-User Interface and Interactions on Direct-Touch Horizontal Surfaces: Collaborative Tabletop Research at MERL. In: *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)*, IEEE, 2006, S. 53–54. – URL <http://dx.doi.org/10.1109/tabletop.2006.22>. – ISBN 0-7695-2494-X
- [Shneiderman 1983] SHNEIDERMAN: Direct Manipulation: A Step Beyond Programming Languages. In: *Computer* 16 (1983), August, Nr. 8, S. 57–69. – URL <http://dx.doi.org/10.1109/mc.1983.1654471>. – ISSN 0018-9162
- [Shneiderman und Plaisant 2010] SHNEIDERMAN, Ben ; PLAISANT, Catherine: *Designing the User Interface: Strategies for Effective Human-computer Interaction.* 5. Addison-Wesley, April 2010. – URL <http://www.worldcat.org/isbn/0321601483>. – ISBN 9780321601483
- [Smith u. a. 2007] SMITH, J. D. ; GRAHAM, T. N. ; HOLMAN, David ; BORCHERS, Jan: Low-Cost Malleable Surfaces with Multi-Touch Pressure Sensitivity. In: *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*, IEEE, Oktober 2007, S. 205–208. – URL <http://dx.doi.org/10.1109/tabletop.2007.1>. – ISBN 0-7695-3013-3
- [Stoppe 2009] STOPPE, Jannis: *Bildbasierte Hover-Detektion für Multi-Touch Interfaces.* Bremen, Germany, University of Bremen, Diplomarbeit, Dezember 2009
- [Sturman 1992] STURMAN, David J.: *Whole-hand Input*, Massachusetts Institute of Technology, Dissertation, Februar 1992
- [Sutherland 1964] SUTHERLAND, Ivan E.: Sketch pad a man-machine graphical communication system. In: *Proceedings of the SHARE design automation workshop.* New York, NY, USA : ACM, 1964 (DAC '64). – URL <http://dx.doi.org/10.1145/800265.810742>

- [Teichert u. a. 2010] TEICHERT, Jens ; HERRLICH, Marc ; WALTHER-FRANKS, Benjamin ; SCHWARTEN, Lasse ; FEIGE, Sebastian ; KRAUSE, Markus ; MALAKA, Rainer: Advancing Large Interactive Surfaces for Use in the Real World. In: *Advances in Human-Computer Interaction 2010* (2010), Mai, S. 1–11. – URL <http://dx.doi.org/10.1155/2010/657937>. – ISSN 1687-5893
- [Terrenghi u. a. 2007] TERRENGHI, Lucia ; KIRK, David ; SELLEN, Abigail ; IZADI, Shahram: Affordances for manipulation of physical versus digital media on interactive surfaces. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2007 (CHI '07), S. 1157–1166. – URL <http://dx.doi.org/10.1145/1240624.1240799>. – ISBN 978-1-59593-593-9
- [Terzopoulos u. a. 1987] TERZOPOULOS, Demetri ; PLATT, John ; BARR, Alan ; FLEISCHER, Kurt: Elastically deformable models. In: *SIGGRAPH Comput. Graph.* 21 (1987), Juli, Nr. 4, S. 205–214. – URL <http://dx.doi.org/10.1145/37402.37427>. – ISBN 0897912276
- [Walther-Franks und Herrlich 2010] WALTHER-FRANKS, Benjamin ; HERRLICH, Marc: *Integrating Multi-Touch into Blender 2.5*. Talk at Blender Conference 2010. Oktober 2010
- [Walther-Franks u. a. 2011a] WALTHER-FRANKS, Benjamin ; HERRLICH, Marc ; AUST, Markus ; MALAKA, Rainer: Left and Right Hand Distinction for Multi-touch Displays. In: DICKMANN, Lutz (Hrsg.) ; VOLKMANN, Gerald (Hrsg.) ; MALAKA, Rainer (Hrsg.) ; BOLL, Susanne (Hrsg.) ; KRÜGER, Antonio (Hrsg.) ; OLIVIER, Patrick (Hrsg.): *Smart Graphics* Bd. 6815. Berlin, Heidelberg : Springer Berlin / Heidelberg, 2011, Kap. 17, S. 155–158. – URL http://dx.doi.org/10.1007/978-3-642-22571-0_17. – ISBN 978-3-642-22570-3
- [Walther-Franks u. a. 2010] WALTHER-FRANKS, Benjamin ; HERRLICH, Marc ; MALAKA, Rainer: Multi-Touch Interaction for 3D Modeling and Animation. In: *Interdisciplinary College 2010 (IK2010)*, März 2010
- [Walther-Franks u. a. 2011b] WALTHER-FRANKS, Benjamin ; HERRLICH, Marc ; MALAKA, Rainer: A Multi-touch System for 3D Modelling and

- Animation. In: DICKMANN, Lutz (Hrsg.) ; VOLKMANN, Gerald (Hrsg.) ; MALAKA, Rainer (Hrsg.) ; BOLL, Susanne (Hrsg.) ; KRÜGER, Antonio (Hrsg.) ; OLIVIER, Patrick (Hrsg.): *Smart Graphics* Bd. 6815. Berlin, Heidelberg : Springer Berlin / Heidelberg, 2011, Kap. 5, S. 48–59. – URL http://dx.doi.org/10.1007/978-3-642-22571-0_5. – ISBN 978-3-642-22570-3
- [Walther-Franks u. a. 2008] WALTHER-FRANKS, Benjamin ; SCHWARTEN, Lasse ; TEICHERT, Jens ; KRAUSE, Markus ; HERRLICH, Marc: User Detection for a Multi-touch Table via Proximity Sensors. In: *IEEE Tabletops and Interactive Surfaces 2008*. Los Alamitos, CA, USA : IEEE Computer Society, 2008
- [Wang u. a. 2009] WANG, Feng ; CAO, Xiang ; REN, Xiangshi ; IRANI, Pourang: Detecting and leveraging finger orientation for interaction with direct-touch surfaces. In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2009 (UIST '09), S. 23–32. – URL <http://dx.doi.org/10.1145/1622176.1622182>. – ISBN 978-1-60558-745-5
- [Wang und Kaufman 1995] WANG, Sidney W. ; KAUFMAN, Arie E.: Volume sculpting. In: *Proceedings of the 1995 symposium on Interactive 3D graphics*. New York, NY, USA : ACM, 1995 (I3D '95). – URL <http://dx.doi.org/10.1145/199404.199430>. – ISBN 0-89791-736-7
- [Westerman 1999] WESTERMAN, Wayne: *Hand Tracking, Finger Identification, and Chordic Manipulation on a Multi-Touch Surface*, University of Delaware, Dissertation, 1999
- [Wigdor u. a. 2009] WIGDOR, Daniel ; WILLIAMS, Sarah ; CRONIN, Michael ; LEVY, Robert ; WHITE, Katie ; MAZEEV, Maxim ; BENKO, Hrvoje: Ripples: utilizing per-contact visualizations to improve user interaction with touch displays. In: *Proceedings of the 22nd annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2009 (UIST '09), S. 3–12. – URL <http://dx.doi.org/10.1145/1622176.1622180>. – ISBN 978-1-60558-745-5

- [Wigdor und Wixon 2011] WIGDOR, Daniel ; WIXON, Dennis: *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. 1. Morgan Kaufmann, April 2011. – URL <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0123822319>. – ISBN 0123822319
- [Williams 1990] WILLIAMS, Lance: 3D paint. In: *SIGGRAPH Comput. Graph.* 24 (1990), Februar, Nr. 2, S. 225–233. – URL <http://dx.doi.org/10.1145/91394.91450>. – ISSN 0097-8930
- [Wilson 2004] WILSON, Andrew D.: TouchLight: an imaging touch screen and display for gesture-based interaction. In: *ICMI '04: Proceedings of the 6th international conference on Multimodal interfaces*. New York, NY, USA : ACM, 2004, S. 69–76. – URL <http://dx.doi.org/10.1145/1027933.1027946>. – ISBN 1-58113-995-0
- [Wilson 2005] WILSON, Andrew D.: PlayAnywhere: a compact interactive tabletop projection-vision system. In: *Proceedings of the 18th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2005 (UIST '05), S. 83–92. – URL <http://dx.doi.org/10.1145/1095034.1095047>. – ISBN 1-59593-271-2
- [Wilson 2009] WILSON, Andrew D.: Simulating Grasping Behavior on an Imaging Interactive Surface. In: *Proceedings of Interactive Tabletops and Surfaces 2009*, 2009
- [Wilson u. a. 2008] WILSON, Andrew D. ; IZADI, Shahram ; HILLIGES, Otmar ; MENDOZA, Armando G. ; KIRK, David: Bringing physics to the surface. In: *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2008, S. 67–76. – URL <http://dx.doi.org/10.1145/1449715.1449728>. – ISBN 978-1-59593-975-3
- [Wobbrock u. a. 2009] WOB BROCK, Jacob O. ; MORRIS, Meredith R. ; WILSON, Andrew D.: User-defined gestures for surface computing. In: *Proceedings of the 27th international conference on Human factors in computing systems*. New York, NY, USA : ACM, 2009 (CHI '09), S. 1083–1092. – URL

<http://dx.doi.org/10.1145/1518701.1518866>. – ISBN 978-1-60558-246-7

[Wu u. a. 2006] WU, M. ; SHEN, Chia ; RYALL, K. ; FORLINES, C. ; BALAKRISHNAN, R.: Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. In: *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP '06)* Bd. 0. Washington, DC, USA : IEEE, 2006, S. 185–192. – URL <http://dx.doi.org/10.1109/tabletop.2006.19>. – ISBN 0-7695-2494-X

[Wu und Balakrishnan 2003] WU, Mike ; BALAKRISHNAN, Ravin: Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In: *Proceedings of the 16th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2003 (UIST '03), S. 193–202. – URL <http://dx.doi.org/10.1145/964696.964718>. – ISBN 1-58113-636-6

[Wyvill u. a. 1986a] WYVILL, Brian ; MCPHEETERS, Craig ; WYVILL, Geoff: Animating soft objects. In: *The Visual Computer* 2 (1986), August, Nr. 4, S. 235–242. – URL <http://dx.doi.org/10.1007/bf01900347>. – ISSN 1432-8726

[Wyvill u. a. 1986b] WYVILL, G. ; MCPHEETERS, C. ; WYVILL, B.: Soft objects. In: *Proceedings of Computer Graphics Tokyo '86 on Advanced Computer Graphics*. New York, NY, USA : Springer-Verlag New York, Inc., 1986, S. 113–128. – URL <http://portal.acm.org/citation.cfm?id=23944.23952>. – ISBN 4-431-70011-0

[Zelevnik u. a. 1997] ZELEVNIK, Robert C. ; FORSBERG, Andrew S. ; STRAUSS, Paul S.: Two pointer input for 3D interaction. In: *SI3D '97: Proceedings of the 1997 symposium on Interactive 3D graphics*. New York, NY, USA : ACM, 1997. – URL <http://dx.doi.org/10.1145/253284.253316>. – ISBN 0-89791-884-3

[Zelevnik u. a. 1996] ZELEVNIK, Robert C. ; HERNDON, Kenneth P. ; HUGHES, John F.: SKETCH: an interface for sketching 3D scenes. In: *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer*

graphics and interactive techniques. New York, NY, USA : ACM, 1996, S. 163–170. – URL <http://dx.doi.org/10.1145/237170.237238>. – ISBN 0-89791-746-4

[Zhai und Milgram 1998] ZHAI, Shumin ; MILGRAM, Paul: Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In: *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM Press/Addison-Wesley Publishing Co., 1998, S. 320–327. – URL <http://dx.doi.org/10.1145/274644.274689>. – ISBN 0-201-30987-4

Erklärung zur Selbstständigkeit

Ich erkläre hiermit gemäß §6 Abs. 5 der Promotionsordnung des Fachbereichs 3 der Universität Bremen vom 14. März 2007, dass ich diese Arbeit ohne unerlaubte fremde Hilfe angefertigt habe. Ich habe dabei keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt. Die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen habe ich als solche kenntlich gemacht.

Ort, Datum

Unterschrift

Hinweis zu Überarbeitungen

Ich erkläre hiermit gemäß §11 Abs. 2 der Promotionsordnung des Fachbereichs 3 der Universität Bremen vom 14. März 2007, dass gegenüber der für die Zulassung zum Kolloquium eingereichten Version mit Zustimmung des Vorsitzenden der Prüfungskommission die folgenden Überarbeitungen für die Veröffentlichung vorgenommen worden sind:

- das Dokument und enthaltene Bilder wurden umformatiert/konvertiert, um Kompatibilität mit dem PDF/A-Standard herzustellen
- Widmung und Danksagung wurden hinzugefügt; die Reihenfolge der zusätzlichen Textteile (Erklärungen usw.) wurde angepasst/geändert
- die Zusammenfassungen (deutsch und englisch) wurden sprachlich überarbeitet
- das Fazit wurde sprachlich überarbeitet
- im Literaturverzeichnis wurden bei einzelnen Referenzen Angaben zum Verlag ergänzt

Ort, Datum

Unterschrift

