



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Ausarbeitung Anwendungen 2

Olaf Potratz

Physikbasierte Interaktion im virtuellen Raum

Olaf Potratz

Thema der Ausarbeitung Anwendungen 2

Physikbasierte Interaktion im virtuellen Raum

Stichworte

Mensch-Maschine-Interaktion, Motion Tracking, Gestenerkennung, Physik-Engine, Kinect

Kurzzusammenfassung

In der Ausarbeitung soll ein kurzer Überblick über verwandte Arbeiten gegeben werden, die sich mit Interaktion mittels Physik im virtuellen Raum befassen. Zu diesem Zweck werden exemplarisch drei Arbeiten ausgesucht. Diese Arbeiten sollen näher betrachtet und ihre Vor- und Nachteile im Hinblick auf die eigene bevorstehende Masterarbeit aufgezeigt werden.

Olaf Potratz

Title of the paper

Physics-based interaction in virtual spa

Keywords

Human-Computer-Interface, Motion Tracking, Gesture Recognition, Physics-Engine, kinect

Abstract

In the paper, a brief overview of related work is given, dealing with interaction using physics in the virtual space. To this end, examples of three works will be selected. This work should be considered in more detail and their advantages and disadvantages in terms of their own impending thesis are presented.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Aufbau der Arbeit	1
1.2	Motivation	1
2	Verwandte Arbeiten	2
2.1	A User Interface Framework for Multimodal VR Interactions	2
2.1.1	Aufbau	3
2.1.2	Relevanz & Bewertung	3
2.2	Vision-based 3D Finger Interactions for Mixed Reality Games with Physics Simulation	4
2.2.1	Aufbau	4
2.2.2	Ergebnisse	6
2.2.3	Relevanz & Bewertung	6
2.3	HoloDesk: Direct 3D Interactions with a Situated SeeThrough Display	7
2.3.1	Aufbau	7
2.3.2	Ergebnisse	9
2.3.3	Relevanz & Bewertung	9
3	Zusammenfassung und Ausblick	10
3.1	Zusammenfassung	10
3.2	Ausblick	10
	Literaturverzeichnis	11
	Abbildungsverzeichnis	12

Kapitel 1

Einleitung

1.1 Aufbau der Arbeit

Die Arbeit ist in drei Kapitel aufgeteilt. Kapitel 1 diese Gliederung und die Motivation für diese Arbeit.

Im 2. Kapitel werden die drei näher betrachteten Arbeiten aufgeführt. Jede Arbeit wird vorgestellt, ihre Ziele und Eigenarten aufgeführt. Zum Ende einer jeden Vorstellung eine Bewertung der Arbeit im Blick auf die Relevanz und evtl. Verwendbarkeit der in den Arbeiten aufgeführten Techniken und Ideen gegeben.

In Kapitel 3 enthält eine abschließende Zusammenfassung und darüber hinaus wird ein Ausblick auf die weitere Arbeit gegeben.

1.2 Motivation

Diese Arbeit schließt an zwei bereits geschriebenen Arbeiten an. Die ursprüngliche Motivation stammt aus der Bachelorarbeit [Potratz (2011)]. Hier galt es, eine Gestenerkennung unter Einbeziehung einer Physik-Engine zu entwickeln. Im Umfeld der Veranstaltung „Anwendungen 1“ richtet sich der Fokus auf Medienfassaden. Mittlerweile ist das Thema etwas genereller gefasst, es geht um die Entwicklung eines Frameworks, das mithilfe einer Physik-Engine Interaktionen und Gestenerkennung bereitstellt. Die Aussage ist hierbei nicht mehr zwingend auf eine Medienfassade beschränkt, sondern auch andere Ausgabemedien sollen betrachtet werden. Auch bei den Eingabegeräten soll das Framework möglichst offen gehalten werden, und man sich nicht von einem bestimmten Typ Sensor abhängig machen. Die folgenden Arbeiten sollen genau hier ansetzen und verschiedene Verfahren aufzeigen und als mögliche Entwicklungshilfen dienen.

Kapitel 2

Verwandte Arbeiten

2.1 A User Interface Framework for Multimodal VR Interactions

[Latoschik (2005)]Die erste Arbeit, die untersucht wurde, ist „A User Interface Framework for Multimodal VR Interactions“ von Marc Erich Latoschik aus dem Jahre 2005. Ziel der Arbeit war eine multimodale Interaktion mit einer virtuellen Umgebung. Die Interaktion sollte mittels Gestik, Sprache und vorhandenen Kontextinformationen realisiert werden. Darüberhinaus wird das Framework noch durch eine KI unterstützt.



Abbildung 2.1: Erfolgreich erkannte Hebegeste (Quelle: [Latoschik (2005)])

2.1.1 Aufbau

Die Kernkomponente des Frameworks ist in Abbildung 2.2 aufgeführt. Die ermittelten Sensordaten werden mittels Multicast an das Framework übergeben. Hier wird dann abhängig von der vorhandenen Wissensbasis, der zugrunde liegenden Umgebung und weiteren Informationen auf eine mögliche Geste geschlossen. Unter anderem kann es von Bedeutung sein, in welchem Kontext eine Geste produziert wird, und wer diese tätigt. Auch der Scene-Graph, das Objekt das alle in der virtuellen Welt befindlichen Objekte verwaltet, hat Einfluss auf die Entscheidung, wie sich das Gesamtsystem verhält.

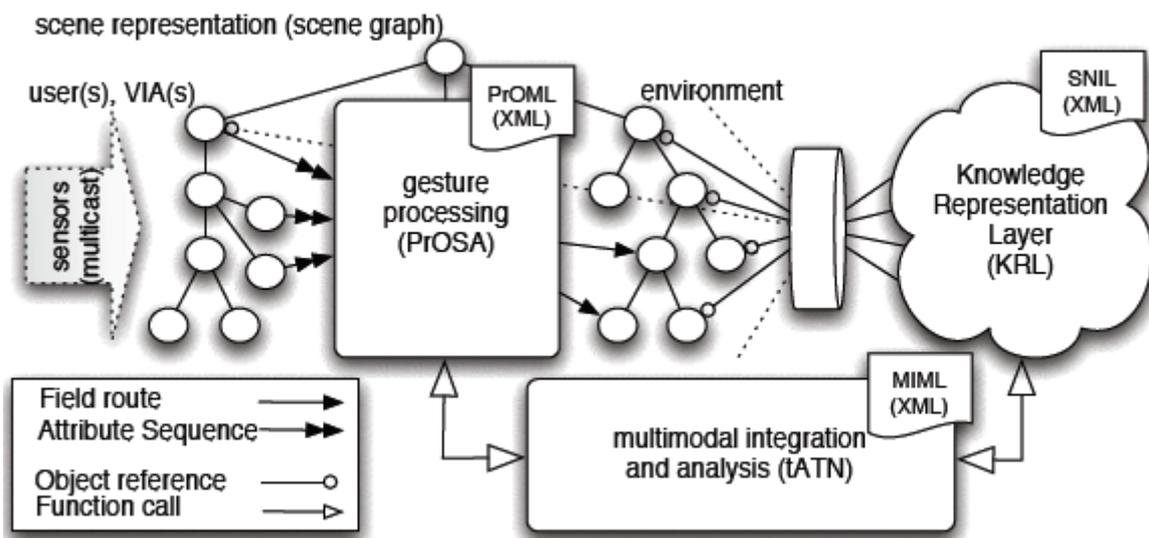


Abbildung 2.2: Architektur des Frameworks (Quelle: [Latoschik (2005)])

2.1.2 Relevanz & Bewertung

Schaut man sich die gestellten Aufgaben in dieser Arbeit an, stellt man fest, dass sich diese Aufgaben zum Teil sehr mit denen ähneln, die selbst bearbeitet werden sollen. Es handelt sich hier auch um ein Framework zur Interaktion in einem virtuellen Raum.

Die Arbeit stammt aus dem Jahr 2005 und ist daher von den verwendeten Softwarekomponenten nicht mehr auf der Höhe der Zeit. Außerdem handelt es sich hier um ein System, was ausschließlich für Indoorzwecke gedacht ist. Abschließend kommt noch hinzu, dass der komplexe Aufbau des Frameworks es schwer macht, hier Systemteile zu ermitteln, die man selbst verwenden könnte. Auch das Fehlen einer Physik-Engine und die somit gänzlich andere Art und Weise der Lösung der Gestenerkennung stellt eine Verwendbarkeit infrage.

2.2 Vision-based 3D Finger Interactions for Mixed Reality Games with Physics Simulation

Die Arbeit „Vision-based 3D Finger Interactions for Mixed Reality Games with Physics Simulation“ [Song u. a. (2008)] umfasst die Umsetzung von zwei Spielen. Es handelt sich hier um „Finger Fishing“ und „Jenga“. Da „Jenge“ etwas anspruchsvoller ist, und sämtliche Funktionen die für „Finger Fishing“ benötigt werden, auch in „Jenge“ enthalten sind, wird sich weiterhin nur mit dem „Jenga“ befasst. Das Spiel selbst, für die es nicht kennen, ist recht einfach, man muss Klötzchen aus dem Turm ziehen, bis dieser zusammenstürzt. Abbildung 2.3 zeigt „Jenga“, vom Start, über die Interaktion, bis zum Einsturz des Turms.

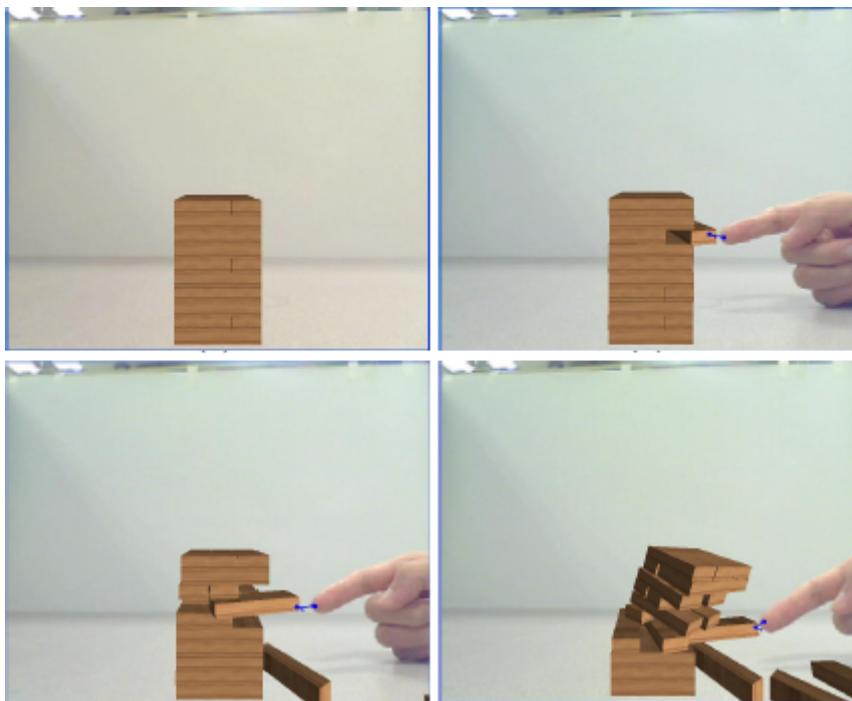


Abbildung 2.3: Erfolgreich erkannte Hebegeste (Quelle: [Song u. a. (2008)])

2.2.1 Aufbau

Als Eingabegerät dient eine Kamera mit zwei Objektiven wie in Abbildung 2.4 zu erkennen ist. Die zwei Objektive dienen einer Tiefenwahrnehmung des Systems. Es handelt sich bei dieser Arbeit um ein Trackingverfahren auf kurze Distanz. Es wird die Hand in relativ kurzen Abstand vor diese Kamera gehalten. Mittels der Suche nach einer möglichst lange Kette von aneinanderhängenden Pixeln wird dann die Position des Fingers und vor allem der Fingerspitze aus dem aufgenommenen Bild herausgearbeitet. Die Abbildung 2.5 veranschaulicht

diesen Sachverhalt noch einmal.



Abbildung 2.4: Erfolgreich erkannte Hebegeste (Quelle: [Song u. a. (2008)])

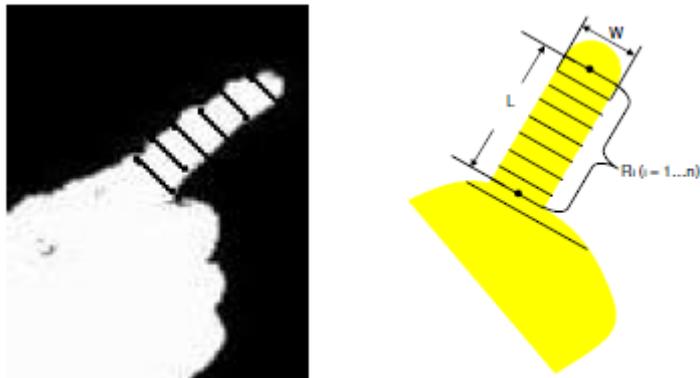


Abbildung 2.5: Links: Bild einer aufgenommenen Hand
Rechts: schematische Darstellung der Hand (Quelle: [Song u. a. (2008)])

Physikalische Interaktion

Ist der Finger erkannt, wird die Position der Fingerspitze in die Physik-Engine übertragen und durch einen Aktor repräsentiert. Die Haltung des Daumens ist entscheidend, wenn der Zeigefinger einen virtuellen Stein berührt. Ist der Daumen gehoben passiert nichts. Ist der Daumen hingegen gesenkt, bzw. nicht weiter sichtbar, weil er von der Hand selbst verdeckt wird, sorgt eine Berührung der Fingerspitze dazu, dass zwischen Repräsentation des Zeigefingers und dem berührten Physikobjekt ein Joint aufgespannt wird. Eine nähere Erklärung über die Funktionsweise eines Joint kann hier entnommen werden [Potratz (2011)].

2.2.2 Ergebnisse

Editieren Der Aufbau funktionierte und eine Interaktion mit den virtuellen Bauklötzchen mittels Zeigefinger war möglich.

Es wurde eine Benutzerstudie mit insgesamt 57 Personen durchgeführt. Diese Studie kam zu dem Schluss, dass wenn man Maus/Tastatur Eingaben, der Eingabe mittels Interaktion mit den Fingern gegenüberstellt, für Spiele, das Agieren mittels Finger mehr Spaß macht, aber die Probanden länger brauchen das Spiel zu spielen.

2.2.3 Relevanz & Bewertung

Es gibt in dieser Arbeit durchaus relevante Punkte, die es hervorzuheben gilt. Es handelt sich hier um ein sehr einfaches, passives Kamerasystem, und ermöglicht trotzdem eine zuverlässige Variante den Aktor in der Realität zu finden. Als möglicher Ideengeber für ein eigenes für einen größeren Suchraum konzipiertes System.

Auch sind die Eigenschaften der Physik-Engine nicht allzu weit von denen entfernt, die man selbst benötigt.

Wie in den Ergebnissen zu dieser Arbeit erwähnt umfasst diese Arbeit eine Benutzer-Studie. Ohne im Detail auf deren genaue Inhalte einzugehen, ist allein die Vorgehensweise zum Erstellen einer solchen Studie interessant. Hier können zu einem späteren Zeitpunkt hilfreiche Anhaltspunkte entnommen werden, wenn man selbst eine ähnliche Benutzer-Studie durchführen möchte.

Diese Arbeit ist als Ideengeber sehr wertvoll, auch wenn die eigentliche Umsetzung aufgrund der für kurze Distanz und das primitive Interaktionsverfahren mittels eines Joint durch eine Initialgeste nicht weiter verwendbar.

2.3 HoloDesk: Direct 3D Interactions with a Situated See-Through Display

Als letztes soll die Arbeit von Otmar Hilliges u.a. [Hilliges (2012)] mit dem Titel „HoloDesk: Direct 3D Interactions with a Situated See-Through Display“ vorgestellt werden. Es handelt sich hier um einen Aufbau mit durchsichtigen Display und einem Kinect Sensor. Wie in Abbildung 2.6 zu sehen, sieht der Benutzer das Objekt genau an der Position, wo er mit ihm, im virtuellen Physik-Raum interagiert.

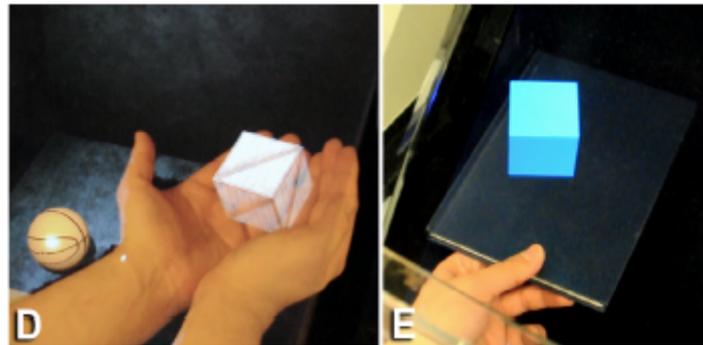


Abbildung 2.6: Interaktion mit einem virtuellen Würfel (Quelle: [Hilliges (2012)])

2.3.1 Aufbau

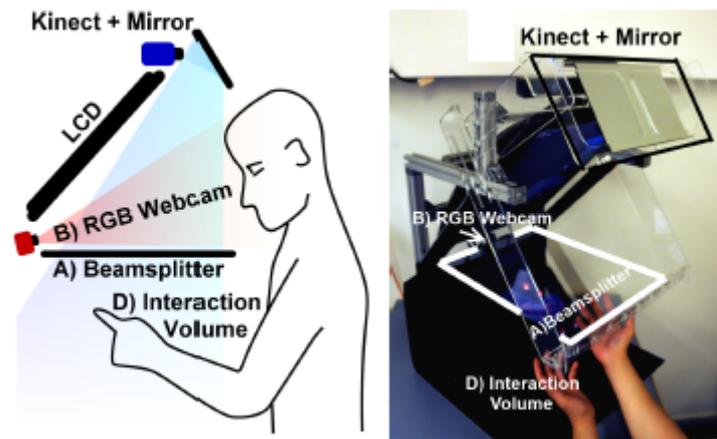


Abbildung 2.7: Links: Schematische Darstellung des Aufbaus
Rechts: Foto des fertigen Aufbaus (Quelle:[Hilliges (2012)])

Das Tracking erfolgt hier mittels eines Kinect-Sensors, der wie in Abbildung 2.7 zu sehen ist, mittels eines Spiegels so umgelenkt wird, das er die Hände des Probanden oder ein beliebiges anderes Objekt, das dort hineingehalten wird, von oben auszeichnen kann. Die Daten über das aufgezeichnete Objekt, im Fall von Abbildung 2.8 die Hände, werden weiterverarbeitet. Die Verarbeitung erfolgt in den drei Schritten von links nach rechts (siehe Abbildung 2.8). Am Ende dieser Verarbeitung gibt es ein, oder mehrere physikalische Konstruckte in der Physik-Engine. Diese Konstruckte bestehen aus einzlenen Physikobjekten, in diesem Fall, kleinen Kugeln. Diese kugeln werden durch Joints auf ihren Positionen gehalten.



Abbildung 2.8: Verarbeitungsschritte von der Realität in die Physik-Engine (Quelle: [Hilliges (2012)])

Physikalische Interaktion

Die Interaktion zwischen den realen Händen und dem Physikobjekt findet mittels der stellvertretenden Aktorobjekte statt, die die Hände repräsentieren. Wie in Abbildung 2.9 gut zu erkennen ist, bleiben die einzelnen Teile des Aktors ausserhalb des Würfels. In Teilbild D der Abbildung 2.9 sind gelbe Linien zu erkennen. Diese Linien repräsentieren die auftretenden Kräfte der Kugeln, auf den Würfel, da die Kugeln selbst auf die mit einem roten Punkt gekennzeichneten Positionen wollen. Mittels dieser Kräfte, die für eine Reibung sorgen, ist es möglich im Physik-Raum, den Würfel zu greifen, und zu bewegen, auch anzuheben.

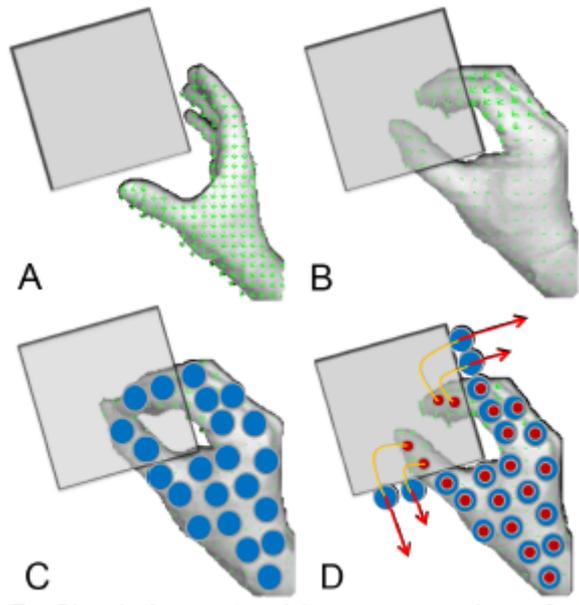


Abbildung 2.9: Interaktion mittels Physik (Quelle: [Hilliges (2012)])

2.3.2 Ergebnisse

Die Ergebnisse des Ausbaus waren vielversprechend, nicht zuletzt durch das durchsichtige Display war die Interaktion mit den Physikobjekten schon sehr realitätsnah für den Benutzer. Besonders das positionsgenaue Sehen der Objekte ist nochmal eine Steigerung, zu den herkömmlichen Visualisierungsverfahren, wie mit großen Bildschirmen, an einer nahegelegenen Wand.

2.3.3 Relevanz & Bewertung

Editieren Die Arbeit ist top aktuell und zeigt gut die im Moment vorhandenen Möglichkeiten. Insbesondere das Physikmodell der Aktoren aus zusammengesetzten einzelnen Kugeln, und die darauf mögliche realitätsnahe Interaktion beim Greifen von Objekten ist sehr interessant. Der Versuchsaufbau selbst, der ausschließlich auf der Verwendung der Kinect fusst und nur für kurze Distanzen gedacht ist, bietet keine weiteren Punkte, die für das zu erstellende Framework hilfreich scheinen.

Kapitel 3

Zusammenfassung und Ausblick

3.1 Zusammenfassung

In dieser Arbeit wurden verschiedene Arbeiten untersucht, die sich alle mit der Interaktion in virtuellen Räumen befassen. Abschließend kann man sagen, dass alle diese Arbeiten eine Stütze für das weitere Vorgehen darstellen können. Allerdings wichen die Zielsetzungen der untersuchten Arbeiten doch soweit von den eigenen Zielen ab, dass keiner der Arbeiten als Leitfaden gesehen werden kann.

3.2 Ausblick

Parallel zu dieser Veranstaltung wurde in der Veranstaltung Projekt 1 bereits mit der Implementierung eines Frameworks begonnen. Viele der hier gewonnenen Erkenntnisse und Ideen werden in der Folgeveranstaltung Projekt 2 zur Anwendung kommen. Hier ist im Besonderen die Arbeit von Otmar Hilliges u.a. [Hilliges (2012)] zu nennen. Das dort zur Anwendung gekommene Aktor-Modell scheint sehr vielversprechend, da der eigene Aktor bislang ausschließlich aus einer nicht transformierbaren Kugel besteht. Somit würde ein Aktor, der aus mehreren kleinen Objekten besteht, sicherlich eine interessante Ergänzung darstellen.

Literaturverzeichnis

- [Hilliges 2012] HILLIGES, Otmar u.: *HoloDesk: direct 3d interactions with a situated see-through display*. 2012. – URL <http://doi.acm.org/10.1145/2208276.2208405>
- [Latoschik 2005] LATOSCHIK, Marc E.: *A User Interface Framework for Multimodal VR Interactions*. 2005. – URL <http://doi.acm.org/10.1145/1088463.1088479>
- [Potratz 2011] POTRATZ, Olaf: *Ein System zur physikbasierten Interpretation von Gesten im 3D-Raum*. 2011. – URL <http://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/bachelor/potratz.pdf>
- [Song u. a. 2008] SONG, Peng ; YU, Hang ; WINKLER, Stefan: *Vision-based 3D finger interactions for mixed reality games with physics simulation*. 2008. – URL <http://doi.acm.org/10.1145/1477862.1477871>

Abbildungsverzeichnis

2.1	Erfolgreich erkannte Hebegeste (Quelle: [Latoschik (2005)])	2
2.2	Architektur des Frameworks (Quelle: [Latoschik (2005)])	3
2.3	Erfolgreich erkannte Hebegeste (Quelle: [Song u. a. (2008)])	4
2.4	Erfolgreich erkannte Hebegeste (Quelle: [Song u. a. (2008)])	5
2.5	Links: Bild einer aufgenommenen Hand Rechts: schematische Darstellung der Hand (Quelle: [Song u. a. (2008)]) . .	5
2.6	Interaktion mit einem virtuellen Würfel (Quelle: [Hilliges (2012)])	7
2.7	Links: Schematische Darstellung des Aufbaus Rechts: Foto des fertigen Aufbaus (Quelle:[Hilliges (2012)])	7
2.8	Verarbeitungsschritte von der Realität in die Physik-Engine (Quelle: [Hilliges (2012)])	8
2.9	Interaktion mittels Physik (Quelle: [Hilliges (2012)])	9