

**Fachbereich 3
Mathematik und Informatik**

Masterarbeit

**Evaluation von multisensorischen Antizipationshilfen
für Performance Animation in Virtual Reality**

**Evaluation of multisensoric anticipation guides
for performance animation in Virtual Reality**

Autor: Stefan Witt
switt@uni-bremen.de

Betreuer: Thomas Münder

Gutachter: Prof. Dr. Rainer Malaka
Dr. René Weller

Abgabedatum: 02.10.2019

I Abstract

Thematik: In der heutigen Zeit spielen Antizipationshilfen in unserer alltäglichen Umgebung eine große Rolle, ohne dass diese dabei bewusst wahrgenommen werden müssen. Auch in der Virtual Reality (VR) werden sie bereits zur Animation zeitkritischer Szenarien verwendet, wobei sie sich verschiedener Sinnesmodalitäten bedienen können.

Fragestellung: Es wurde untersucht, ob visuelle Antizipationshilfen, durch das Hinzufügen von synchron rhythmischer Stimulation der taktilen und auditiven Sinnesmodalität an Mehrwert für die Animation zeitkritischer Handlungsabläufe in der VR gewinnen.

Methodik: 20 Teilnehmende hatten zur Aufgabe, jeweils in zwei Szenarien mit mehreren zeitkritischen Handlungsabläufen einen Charakter zu animieren. Dabei durchliefen sie nacheinander zwei VR-Variationen, die jeweils visuelle oder multisensorische (visuell, taktil & auditiv) Antizipationshilfen enthielten. Im Anschluss gaben die Teilnehmenden Bewertungen zur Güte und Nutzung der Antizipationshilfen ab. Mittels NASA-TLX wurde die persönliche Belastung nach Abschluss beider Versuchsvarianten, deren Reihenfolge über die Teilnehmenden pseudo-randomisiert wurde, abgefragt, um Reihenfolgeeffekt und somit Belastungsunterschiede zwischen den Varianten zu vergleichen.

Ergebnisse: Die Studienteilnehmenden nahmen die Antizipationshilfen beider Versuchsvarianten (rein visuell vs. multisensorisch) als hilfreich wahr, wobei ein signifikanter Mehrwert der Güte der multisensorischen Antizipationshilfen beschrieben wurde. Dabei stützten sich die Teilnehmenden zur Aufgabenbewältigung in der multisensorischen Versuchsvariante teils auf unterschiedliche Kombinationsmöglichkeiten der dargebotenen Reizmodalitäten. Die empfundene Belastung nach Durchführung beider Versuchsvarianten zeigte über alle Teilnehmenden keine Beeinflussung durch deren Reihenfolge und deutete somit auch keine höhere Belastung durch die multisensorische Variante an.

Diskussion: Die Ergebnisse deuten darauf hin, dass sowohl visuelle als auch multisensorische Antizipationshilfen bei der Animation zeitkritischer Szenarien unterstützend und nicht überfordernd oder ablenkend wirken. Multisensorische Hilfen vermögen den Animierenden dabei noch effektiver zu unterstützen als rein visuelle Hilfen. Unterschiedliche Kombinationen sensorischer Antizipationshilfen konnten dabei nicht im Detail in ihrer Wirkung und ihrem Aufgabennutzen untersucht werden.

Signifikanz: Antizipationshilfen können im Bereich der VR und der Performance Animation einen entscheidenden Beitrag zu einer Steigerung der effizienten Animation eines komplexen Szenarios leisten, um gerade auch mit geringem personellem Aufwand beispielsweise in ein bestehendes Animationsszenario zusätzliche Charaktere mit zeitkritischen Handlungsereignissen einzufügen, und spielen daher eine wichtige Rolle im wachsenden Feld der VR.

II Inhaltsverzeichnis

I	Abstract	I
II	Inhaltsverzeichnis	II
III	Abbildungsverzeichnis	IV
IV	Tabellenverzeichnis	V
1	Einleitung	1
2	Beschreibung des aktuellen Technikstands	2
2.1	Virtual Reality	2
2.1.1	Projektions-basierte Virtual Reality	3
2.1.2	Virtual-Reality-Headsets	5
2.1.3	Nebenwirkungen von VR auf den menschlichen Körper	7
2.1.4	Anwendungsgebiete der VR	8
2.2	Performance Animation und Motion Capture	9
2.2.1	Optische Systeme	10
2.2.2	Nicht-optische Systeme	11
2.2.3	Inverse Kinematik	12
2.2.4	Motion Capture in Virtual Reality	13
2.3	Allgemeine Zeitantizipation	16
2.3.1	Antizipation in der Animation	17
2.3.2	Antizipation mit Multisensorik	19
3	Fragestellung	21
3.1	Motivation	21
3.2	Zielsetzung	22
4	Design und Implementierung	22
4.1	Quellcode-Basis	23
4.1.1	Das Plugin Final IK	24
4.1.2	Bereits vorhandene Funktionen	26
4.1.3	Neue Funktionen	28
4.2	Antizipationshilfen	30
4.2.1	Design der visuellen Antizipationshilfen	30
4.2.2	Design der multisensorische Antizipation	32
4.2.3	Steuerung	33
5	Studie	34
5.1	Fragestellung	34
5.2	Methodik	34
5.2.1	Stichprobe	34
5.2.2	Design	35
5.2.3	Versuchsaufbau	37

5.2.4	Durchführung	37
5.2.5	Auswertung der Daten	38
5.3	Ergebnisse	40
5.3.1	Fragebogen 1	40
5.3.2	Fragebogen 2	42
5.3.3	NASA-TLX	44
5.4	Diskussion	45
5.4.1	Diskussion der Studienergebnisse	45
5.4.2	Limitationen	47
5.4.3	Ausblick	48
6	Fazit	50
7	Quellenverzeichnis	51

III Abbildungsverzeichnis

Abb. 1	Vereinfachte Darstellung des „Virtual Continuum“ (Jerald, 2016, S. 30)	3
Abb. 2	Beispiele für (a)Augmented Virtuality und (b)Augmented Reality . . .	4
Abb. 3	Beispielhafter skizzierter Aufbau eines „Cave“ (Unigine, 2019)	4
Abb. 4	Direkte und inverse Kinematik. (Wikipedia, 2019)	13
Abb. 5	Mögliche Hardwarekonfigurationen der Software Orion (IKinema, 2019a)	14
Abb. 6	Antizipation im Spiel Pole Position (PolePosition, 1982)	16
Abb. 7	Guitar Hero 5 (GuitarHero, 2019b)	17
Abb. 8	Antizipation einer schnellen Bewegung (Lasseter, 1987)	18
Abb. 9	Die vier Antizipationshilfen der Bachelorarbeit „Evaluation von Antizipationshilfen im Kontext von Echtzeitsteuerung für Computerspiele und Performance Animation“ (Witt, 2016)	19
Abb. 10	HTC Vive (HTC, 2019a)	23
Abb. 11	Funktionsweise Tracking HTC Vive (HAWK, 2019)	23
Abb. 12	Zeitachse Animationen	24
Abb. 13	An HTC Vive Controller angepasste Handpositionen	25
Abb. 14	Konfiguration der Komponente VR IK des Unity-Plugins Final IK . .	26
Abb. 15	Teleportation oberhalb des vordefinierten Bodens	27
Abb. 16	Hauptmenü des Prototyps	29
Abb. 17	Erstes Szenario: Ein virtueller Charakter läuft ein paar Schritte nach vorne und macht ein „High Five“	31
Abb. 18	Zweites Szenario: Verschiedene Aktionen in unterschiedlichen Zeitabständen hintereinander	36
Abb. 19	Darstellung der Antwortverteilungen für die drei Fragen des Fragebogens 1 (1-3A) sowie Mittelwerte mit Standardabweichung der von 1-7 skalierten Antwortmöglichkeiten und des Gesamtwertes über Fragebogen 1 mit Andeutung signifikanter Modalitätsunterschiede zwischen den Versuchsbedingungen rein visuell und multisensorisch (* $p < 0,05$ & ** $p < 0,01$; $N=20$).	41
Abb. 20	Darstellung der relativen Antworthäufigkeiten [in %] der jeweiligen Antwortkategorien der Fragen 4 (A) und 5 (B) aus Fragebogen 2 für alle Teilnehmenden ($N=20$).	44
Abb. 21	Darstellung von Mittelwerten mit Standardabweichungen der Gesamtwerte des NASA-TLX Belastungstests für die Versuchsmodalitäten rein visuell und multisensorisch ($N=20$).	45

IV Tabellenverzeichnis

Tab. 1	Mittelwerte (M) und Standardabweichungen (SD) Antwortmöglichkeiten der Fragen 1-3 des 1. Fragebogens für die beiden Versuchsbedingungen visuell und multisensorisch	40
Tab. 2	Relative Antworthäufigkeiten [in %] der jeweiligen Antwortkategorien der Fragen aus Fragebogen 2 aller Teilnehmenden (N=20).	43

1 Einleitung

Performance Animation ist heute ein weit verbreitetes Forschungsgebiet, findet in der Produktion von fast allen modernen Filmen und Spielen Anwendung und entwickelt sich rasant weiter. Was mit einfachem Motion Capture begann, welches früher noch benutzt wurde, um menschliche oder tierische Bewegungen zu analysieren, spielt sich heute schon in einer eigenen virtuellen Realität ab. Spätestens seit der Vorstellung der Oculus Rift im Juni 2012 ist Virtual Reality (folgend nur noch als VR bezeichnet) nicht nur ein theoretischer Begriff, der an mangelnder Rechenleistung und Speicherproblemen scheitert, sondern ein eigenes immer wachsendes Forschungsgebiet, welches sich heute bereits in viele Bereiche integriert hat. So auch in das Gebiet der Performance Animation.

Durch die Kombination von VR mit Motion Capture-Systemen gibt es eine Vielzahl von neuen Herausforderungen, die es zu bewältigen gilt. Eine dieser Herausforderungen ist das richtige Timing. Beim Animieren ist es oft sehr entscheidend, wann genau der Akteur eine bestimmte Bewegung ausführt, um die Authentizität der Animation zu gewährleisten. Es ist folglich essentiell für eine Vielzahl von Animationen.

Das Problem des Timings einer Handlung trifft mit den Grenzen der Animationsmöglichkeiten in der VR zusammen. Daraus können Probleme in der grundlegenden Interaktion mit virtuellen Charakteren resultieren, beispielsweise wenn das Szenario ein einfaches Anrempeln oder Schubsen eines anderen Charakters verlangt. Das Ausbleiben taktile Reize in dieser VR-Situation erfordert eine exakte zeitliche Abstimmung von sowohl Position als auch Bewegungsabläufen, um in einer realistisch wirkenden Animation zu resultieren. Ein solch komplexes Animationsszenario erfordert daher ein hohes Maß an Vertrautheit mit der zu animierenden Situation und gegebenenfalls eine anschließende Nachbearbeitung. In der Performance Animation, in der Animationen direkt in Echtzeit aufgenommen werden, wird unter anderem das Konzept der Antizipation eingesetzt, um ein effizientes Animieren zu ermöglichen und Schritte der Nachbearbeitung zu minimieren. Nutzern wird es hierbei ermöglicht, durch Einsatz verschiedener Mechanismen wie Counter oder Zeitleisten einen Zeitpunkt zu antizipieren, sodass dieser vorzugsweise direkt getroffen werden kann. Folglich spielen Antizipationshilfen, die uns auf ein eintretendes Ereignis vorbereiten, eine entscheidende Rolle für zeitkritische Ereignisse in der VR.

Antizipationshilfen spielen bereits in sehr vielen Bereichen des menschlichen Alltags eine ausschlaggebende Rolle, wie beispielsweise als Ampel im Straßenverkehr, als Countdown in Computerspielen, oder als Sanduhr für das Zähneputzen von Kindern. Zum großen Teil werden diese zeitlichen Hilfen als selbstverständlich wahrgenommen und nicht tiefer gehend betrachtet. Dies spiegelt sich in einer nur geringen Zahl wissenschaftlicher Literatur

wider, die sich explizit mit der Thematik von Antizipationshilfen befasst.

Antizipationshilfen können sich verschiedener sensorischer Kanäle bedienen, um die Handlung auf ein zeitlich exaktes Ziel auszurichten. Zwei der vorangegangenen Alltagsbeispiele bedienen sich der visuellen Wahrnehmung. Im Gegensatz dazu können in Computerspielen eingesetzte Countdowns sowohl visuelle, als auch auditive oder taktile Reize ansprechen. In der VR strömen auf den Interakteur vor allem komplexe visuelle und auditive Reize ein. Es ist anzunehmen, dass dabei eine geeignete Wahl der Stimulation als Hilfestellung angepasst an Situation und persönliche Wahrnehmung des Menschen besonders effektiv zu wirken vermag.

Dieser Arbeit wurde die Frage zugrunde gelegt, inwiefern visuelle Antizipationshilfen allein ausreichend zur Animation einer zeitkritischen Aufgabe in der VR ist. Des Weiteren wurde untersucht, ob zusätzliche auditive und taktile Reize die Antizipation weiter stützen können oder für den Animierenden als hinderlich oder störend empfunden werden. Die Arbeit dient damit der Verbesserung von Effizienz-steigernden Antizipationshilfen in der VR.

2 Beschreibung des aktuellen Technikstands

2.1 Virtual Reality

Der Begriff „Virtual Reality“ ist heute ein gängiger Fachbegriff, der in dem Buch „Understanding Virtual“ (Sherman and Craig, 2018, S. 6) auf Grundlage des „New Universal Unabridged Dictionary von Webster, 1989, passend erklärt wird. „Virtual“ wird dabei als in der Essenz oder der Wirkung, aber nicht in der Wirklichkeit existent, beschrieben. „Reality“ wird beschrieben als der Zustand oder die Qualität echt zu sein. Etwas, das unabhängig von Vorstellungen, die darüber vorhanden sind, existiert. Etwas, das eine reale oder tatsächliche greifbare Sache darstellt und sich von dem unterscheidet, das nur offensichtlich ist. Mit anderen Worten: Ein Ort, der existiert und den wir erfahren können. Die VR schafft einen Ort, der real erlebbar wird, ohne dabei als realer Ort zu existieren und kann somit als Schnittstelle beider Konstrukte verstanden werden.

VR ist eine komplett vom Computer generierte virtuelle Umgebung, die nur aus virtuellen, also physisch nicht existenten, Objekten besteht. Als technische Herausforderung gilt es hierbei diese Umgebung als möglichst immersiv zu gestalten. Immersiv oder auch die Immersion, beschreibt den Effekt, die virtuelle Umgebung als möglichst real und im Gegenzug das Empfinden, in der realen Welt zu sein, als möglichst gering wahrzunehmen.

Umgangssprachlich wird in diesem Zusammenhang auch vom „Eintauchen“ in die virtuelle Umgebung gesprochen.

Auf der Abstraktionsebene gibt es neben der Virtual Reality mit ihrer komplett virtuellen Umgebung, auch noch die „Mixed Reality“ (dt. vermischte / gemischte Realität), in der Teile der realen Umgebung mit der virtuellen Umgebung interaktiv eingesetzt werden. Insgesamt wird vom „Virtual Continuum“ gesprochen, welches erstmals 1994 von Paul Milgram und Fumio Kishino (Milgram and Kishino, 1994) eingeführt wurde. Eine vereinfachte Darstellung des Virtual Continuum ist in Abbildung 1 zu sehen.

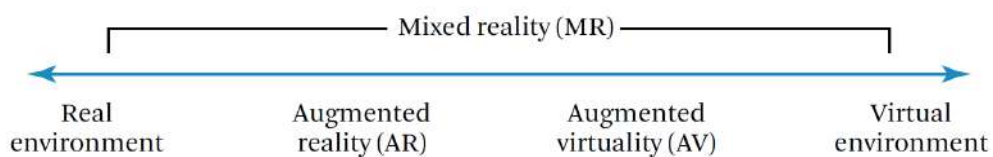


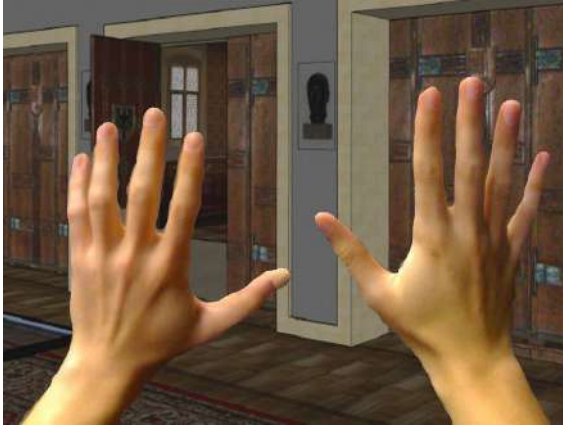
Abbildung 1: Vereinfachte Darstellung des „Virtual Continuum“ (Jerald, 2016, S. 30)

Das Kontinuum der Mixed Reality reicht dabei von der Realität bis zur Virtualität und beinhaltet dabei vor allem zwei große Konzepte: „Augmented Virtuality“ (dt. erweiterte Virtualität), in der real existierende Objekte in eine virtuelle Umgebung übertragen werden und die „Augmented Reality“ (dt. erweiterte Realität), in der virtuelle Objekte in die echte Welt übertragen werden (Abbildung 2). Gerade Augmented Reality hat spätestens durch das Spiel „Pokemon Go“ im Jahre 2016 eine starke Bekanntheit erlangt.

2.1.1 Projektions-basierte Virtual Reality

Die Erzeugung bzw. Simulation einer VR kann mittels verschiedener Techniken erreicht werden. Eine Möglichkeit ist die Projektions-basierte VR (engl. projection-based VR) - eine virtuelle Realität, deren Grundlage Projektionen, zumeist Rückprojektionen, darstellen, die sich abhängig von der Position des Nutzers verschieben. Rückprojektionen sind Projektionen, die von hinter der Projektionsfläche in Richtung Betrachter projiziert werden.

Prägend für diese Art von VR war das „Cave Automatic Virtual Environment“ (Cruz-Neira et al., 1993). Der Nutzer stand dabei in einem kubischen Raum, in dem drei Wände und der Boden Rückprojektionsflächen waren. Die virtuelle Umgebung, die darauf projiziert wurde, änderte sich je nach Position und Richtung des Kopfes des Nutzers. Dadurch entstand der gewünschte immersive Eindruck, Teil der virtuellen Umgebung zu sein. Durch



(a) Virtual body in an augmented virtuality scenario (Bruder et al., 2009)



(b) Bild des Augment Reality Spiels Pokemon Go (PokemonGo, 2019)

Abbildung 2: Beispiele für (a)Augmented Virtuality und (b)Augmented Reality

eine 3D-Shutterbrille war es außerdem möglich, 3D-Bilder zu erzeugen. (evltube, 2007) Der Begriff „Cave“ wird bis heute noch für Projektions-basierte VR-Systeme in Räumen ähnlich dem originalen Cave benutzt. Teilweise werden dabei bis zu sechs Rückprojektoren pro Raum benutzt, um eine komplett umschließende Umgebung zu schaffen.

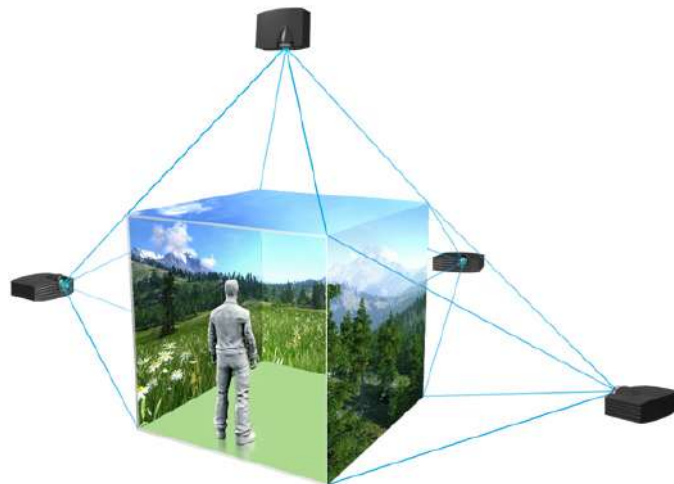


Abbildung 3: Beispielhafter skizzierter Aufbau eines „Cave“ (Unigine, 2019)

Zusätzlich zu der Möglichkeit des Stereoskopisches Sehens bietet diese Art von VR weitere Vorteile. Anders als bei anderen Techniken ist der eigene Körper und auch die direkte physische Umgebung sichtbar. Dadurch bleibt auch bei abstrakten Szenen das eigene Körpergefühl erhalten. Außerdem können andere Personen am VR-Erlebnis des Nutzers teilhaben, was gerade für Präsentationen oder Vorführungen einen enorm hohen Nutzen bietet. Nachteile Projektions-basierter VR-Systeme sind der große Platzbedarf und

die hohen Kosten. Möchte man Faktoren wie Leistung, Präsenz, Engagement, Benutzerfreundlichkeit, Vertrautheit und hohen Leistungsoutput in solchen System erreichen, muss die Displayqualität und die Präzision der Steuerung hoch sein (McMahan et al., 2012). Dies verursacht wiederum erhöhte Kosten.

Eingesetzt wird diese Technik vermehrt in Trainingssimulationen (Doer et al., 2001) oder Konstruktionsmodellierung (Seth et al., 2011a). Eine Firma, die sich beispielsweise auf diesen Sektor spezialisiert hat, ist „WorldViz“ (WorldWiz, 2019).

2.1.2 Virtual-Reality-Headsets

Die weitaus günstigere und heute weit verbreitete Technik um eine VR zu simulieren, ist das Virtual-Reality-Headset. Solch ein Headset ist ein visuell abgeschottetes System, welches bei Nutzung keinerlei Sicht auf die reale Umgebung bietet. Es verbindet ein sogenanntes Head-Mounted Display (kurz HMD) mit mehreren Sensoren zur Erfassung von Position, Rotation und Blickrichtung des Kopfes wie u.a. Gyroskop und Beschleunigungssensor. Dadurch wird das vom Träger zu sehende Bild je nach Position und Blickrichtung angepasst und es entsteht der gewünschte immersive Effekt, Teil der virtuellen Umgebung zu sein. Einige Modelle von VR-Headsets unterstützen zusätzlich Augentracking mit dem einzelne Fokussierungen der Augen erkannt werden können.

Durch den hohen Berechnungsaufwand müssen VR-Headsets zumeist im Vergleich zu einigen anderen HMD-basierten Brillen an einen Computer oder eine Spielekonsole angeschlossen werden. Die bekanntesten Beispiele für VR-Headsets sind Oculus Rift, Oculus Quest, HTC Vive und die PlayStation VR. Die Oculus Quest ist eines der ersten VR-Headsets und in der Aufzählung das einzige, welches ohne Anschluss an einen Computer oder an eine Spielekonsole auskommt. In ihr übernimmt ein moderner Snapdragon-Prozessor die Verarbeitung, der auch in aktuellen Smartphones zum Einsatz kommt. Daher ist auch mit einem hardware-technisch leistungsstark ausgestatteten Smartphones die Erzeugung einer VR möglich. Dazu kann entweder eine visuell abschottende Halterung selbst gebaut, oder das Google Cardboard oder die Samsung Gear VR erworben werden. Das Google Cardboard ist eine günstige Halterung aus Pappe, welche zum Einsetzen eines beliebigen VR-fähigen Smartphones dient. Die Samsung Gear VR ist ein Headset mit einer Halterung, die nur aktuelle Samsung Handys wie das Galaxy S9 oder das Galaxy S10 unterstützt. Allerdings besitzt es noch einige Funktionstasten, die eine Steuerung innerhalb der VR erlauben.

Aufgrund der vielseitigen Ausstattung moderne Smartphones mit Sensoren wie Gyroskop und Beschleunigungssensor wird der Einsatz zur Simulation einer VR ermöglicht. Allerdings ist die Leistung der Smartphones zumeist bereits stark vom Betriebssystem und

verschiedenen Anwendungen eingeschränkt, so dass oft die flüssige Ausgabe einer VR bis dato nicht möglich ist.

Die wichtigste Komponente der eigenständigen VR-Headsets stellt das HMD dar. Grundlegend ist das HMD ein am Kopf befestigtes Gerät, welches vor einem Auge (monokular) oder vor beiden Augen (binokular) einen kleinen Bildschirm besitzt. Dabei existieren verschiedene Arten von HMDs, die je nach Zielnutzung sehr unterschiedlich sein können. Darunter fallen unter anderem Augmented-Reality-Brillen, die für den Träger durch teiltransparente Bildschirme virtuelle Objekte in die reale Welt projizieren. Diese Brillen sind im Gegensatz zu den visuell abgeschotteten VR-Headsets eher offene System, damit der Bezug zur physischen Welt nicht verloren geht. Beispiele für AR-Brillen sind Smartglasses wie die Google Glass (Google, 2019) oder die Microsoft HoloLens (Microsoft, 2019a).

Eine andere Bauweise hingegen unterliegen Videobrillen oder auch FPV-Brillen, die mit extra Sichtklappen die reale Welt komplett ausblenden. Videobrillen sind HMDs, die mit Kopfhörern oder Ohrhörern kombiniert Videomaterial wie Filme oder Fernsehprogramm abspielen können. First-Person-View-Brillen (kurz FPV-Brillen) sind spezielle Videobrillen, die zusätzlich über ein Empfängermodul verfügen, um die Live-Bildübertragungen eines Quadrocopters anzuzeigen.

Das erste HMD, bekannt unter dem Namen „The Sword of Damocles“ (dt. Das Damoklesschwert), wurde bereits im Jahr 1968 unter der Leitung von Ivan Sutherland gebaut (Sutherland, 1968). Es wird heute weitgehend als das erste immersive VR-Headset angesehen und stellte die Grundlage für heutige HMDs und damit auch VR-Headsets. Das Sword of Damocles war ein HMD, welches durch ein an der Decke befestigtes System zusätzlich den Kopf des Trägers tracken und somit die Simulation einer virtuellen Umgebung schaffen konnte. Es existierten allerdings mehrere Limitationen. Zum einen war das HMD so schwer, dass es nicht einfach tragbar war. Ein weiterer Grund für die Halterungssystem, welches an der Decke befestigt war. Zum anderen war der Sensor, der für das Tracking des Kopfes zuständig war, stark in seiner (Winkel-)Detektion begrenzt. Eine Neigung des Kopfes nach unten oder oben mit einem Winkel über 40 Grad wurde nicht erkannt. Außerdem hatte der Träger im Durchmesser nur 6 Fuß ($\sim 1,8288\text{m}$), um sich zu bewegen.

Einen weiteren wichtigen Meilenstein heutiger VR-Technik stellte die National Aeronautics and Space Administration (bekannt unter der Abkürzung „NASA“) mit der „Virtual Interactive Environment Workstation“ (kurz VIEW) im Jahre 1985 (Fisher et al., 1988).

Die von diesem System erzeugte VR ermöglichte eine Steuerung durch einen Datenhandschuh und Spracherkennung. Der Datenhandschuh wurde dabei wie ein normaler Handschuh angezogen, war allerdings auch in der VR sichtbar ist. Die Position von Arm, Hand und Finger wurde dabei erfasst und direkt an den Computer übertragen. Die relative Position zum Kopf wurde dann in der VR berücksichtigt und machte eine Steuerung möglich.

2.1.3 Nebenwirkungen von VR auf den menschlichen Körper

Neben den Vorteilen der Kosteneffizienz, Mobilität, Portabilität und Benutzerfreundlichkeit existieren bei der VR auch starke nachteilige Wirkung auf den menschlichen Organismus.. Gerade die Realitätsnähe, die durch den hohen Grad an Immersion erreicht wird, verursacht einige Nebenwirkungen. Folgende Nebenwirkung können bei Benutzung unter anderem auftreten: Starker Schwindel, Krampfanfälle, Augen- oder Muskelzucken, Ohnmacht, Desorientierung, Übelkeit, Benommenheit, Augenschmerzen, gestörtes Sehvermögen, gestörtes Gleichgewicht, gestörte Hand-Augen-Koordination.

Zusammengefasst werden viele dieser Symptome mit dem Begriff der „Motion Sickness“, teilweise auch unter „Simulator Sickness“ oder „Virtual Reality Sickness“ (dt. VR-Krankheit) bekannt. Motion Sickness ist wissenschaftlich nachgewiesen und stellt einen sehr großen Nachteil bei der Nutzung von VR dar (Munafa et al., 2016). Bekannte VR-Headsets wie Oculus Rift und HTC Vive warnen in ihren Gesundheit- und Sicherheitshinweisen ebenfalls vor diesen Symptomen bei Benutzung und empfehlen einen sofortigen Abbruch bei Auftreten (Facebook, 2019a; HTC, 2019b).

Weitere Folgen, die mit der Benutzung von VR einhergehen, sind psychologischer Natur. Stressige Szenarien innerhalb der VR, wie z.B. das Lehnen über einen Abgrund oder eine Achterbahnfahrt, erhöhen durch den hohen Grad der Immersion und dem Echtheitsgefühl dieser Szenarien die Herzfrequenz und die elektrodermale Aktivität. Des Weiteren treten der sogenannte Proteus Effekt sowie weitere Veränderungen im Verhalten auf, die sich auch nach der Nutzung von VR feststellen lassen. Der Proteus Effekt bedeutet, dass Nutzer von VR sich je nach Aussehen ihres Charakters unterschiedlich verhalten und zwar genau so, wie sie denken, dass andere es erwarten (Madary and Metzinger, 2016). Bei großen Charakteren wird also häufig ein aggressiveres Verhalten an den Tag gelegt, als bei kleinen. Andere verhaltensveränderte Nachwirkungen sind erhöhtes altruistisches Verhalten nach einer „Superman-artigen“ Flugsimulation innerhalb einer Stadt oder auch verringerte rassistische Vorurteile nachdem in VR z.B. ein schwarzhäutiger Charakter gesteuert wurde (Madary and Metzinger, 2016). Bestimmte Szenarien können Einfluss auf das Essverhalten haben. Opfer einer virtuellen Vergewaltung können psychische Traumata erleiden (Madary and Metzinger, 2016). Das zeugt von einer sehr hohen und bedenklichen

manipulativen Auswirkung die VR auf Nutzende haben kann.

2.1.4 Anwendungsgebiete der VR

Durch die immer fortschreitendere Entwicklung von Technik eröffneten sich stetig mehr Anwendungsgebiete für Virtual Reality. Heute ist VR ein wichtiger Bestandteil der Forschung und hat sich auch in der Wirtschaft fest integriert. Eine Vielzahl an Studien konnte bereits den Mehrwert von VR belegen. Gängige Einsatzbereiche sind beispielsweise Ausbildung (Cardenas et al., 2017), Medizin (Maples-Keller et al., 2017), Risikoprävention (Grabowski and Jankowski, 2015; Hee Cho et al., 2014), Produktdesign (Seth et al., 2011b), Animationsbearbeitung (Deering, 1995; NVRMIND, 2019; Facebook, 2019b) und Spiele- und Filmentwicklung (CloudheadGames, 2016).

Gerade im Bereich der Medizin gibt es kontinuierliche Fortschritte, die sich neben Ausbildung bei Operationen (Seymour et al., 2002), Alternativen durch VR gestützte Behandlungsmöglichkeiten widmen, wie beispielsweise bei der Behandlung von Angstzuständen und anderen psychologischen Störungen (Maples-Keller et al., 2017) sowie bei der Behandlung von Schlaganfallpatienten (Laver et al., 2017).

Die VR ermöglicht es ebenfalls, andere Forschungsgebiete in der Umsetzung experimenteller Paradigmen zu unterstützen. So findet sie zunehmend Einsatz in neuen Forschungsfeldern wie den Sozialen Neurowissenschaften, in denen soziales Verhalten und dessen neuronale Korrelate untersucht werden. Komplexe soziale Situationen lassen sich gezielt manipulieren durch den Einsatz von virtuell simulierten Charakteren. Durch die hohe Immanenz und geschickte Studiendesignwahl, die dank VR hohe Kontrollmöglichkeiten bietet, kann sich die Versuchsperson der experimentelle Manipulation nicht entziehen. Gerade in den Neurowissenschaften bedarf es gezielter systematische und unabhängiger Testveränderungen, um die neurobiologischen Grundlagen sozialen Verhaltens zu verstehen (Parsons et al., 2017).

Durch den mittlerweile hohen Verbreitungsgrad, auch in privaten Haushalten, werden heute ganze Spiele speziell für VR konzipiert. Auch bei der Produktion von Spielen selbst unterstützt VR z.B. für das Animierens von Charakteren (CloudheadGames, 2016). Meist handelt es sich dabei um Performance Animation, die in Kombination mit VR zum Einsatz kommt.

2.2 Performance Animation und Motion Capture

Bei der Performance Animation steht der Performer bzw. die Performance im Mittelpunkt. Es existiert ein Akteur, der Bewegungen ausführt, die dann live oder im Nachhinein auf virtuelle Menschen, Tiere oder Objekte übertragen werden. Bei der Filmproduktion aktueller Filme findet sich Performance Animation mittels Motion Capture wieder, um z.B. Dinosaurier- oder Affenbewegungen schnell aufnehmen zu können, um später die entsprechenden Grafiken auf die virtuellen Charakter zu platzieren (Steiger, 2019).

Motion Capture (kurz mo-cap oder mocap) bezeichnet den Prozess der Bewegungserfassung und -aufzeichnung des Körpers sowie die Übertragung an ein Endgerät, in der Regel einen Computer. Die erfassten Bewegungsdaten können dann digital weiter verarbeitet werden. Es wird größtenteils dazu benutzt, um Bewegungen direkt auf virtuelle Charaktere oder Objekte zu übertragen. Bei einer Live-Übertragung mittels Mocap spricht man von Real-Time Motion Capture. Während sich Motion Capture nur auf die Körperbewegungen im Ganzen bezieht, werden heute genauere Techniken eingesetzt, die zeitgleich dazu Finger- und auch Gesichtsbewegungen erfassen können. Ist Mimik inkludiert, so spricht man von „Performance Capture“ (Okun et al., 2010).

Das Erstellen von animationsbasierten Szenarien kann mitunter komplexe Prozesse nach sich ziehen. Während bei Film- oder Spieleentwicklung ganze Teams die Animationen nachbearbeiten, existieren bereits im Amateurbereich und in der Forschung viele Tools und Programme, die ein alleiniges Arbeiten ermöglichen. Ein Beispiel hierfür ist die „Animation Loop Station“, die durch Sprachbefehle die Aufnahme aktivieren und stoppen kann (Walther-Franks et al., 2012). Werden bei solchen Programmen komplexere Szenen mit mehreren Charakteren animiert, wird mit mehreren Aufnahmen, sogenannten Layern gearbeitet. Ein Layer beinhaltet genau eine Aufnahme eines Charakters. Wird diese Technik in der Performance Animation eingesetzt, spricht man auch von „Layered Performance Animation“ (Walther-Franks et al., 2012).

Motion Capture kann durch viele verschiedene Techniken realisiert werden. Man unterscheidet im Allgemeinen zwischen den weit verbreiteten optischen Systemen und alternativen Systemen.

2.2.1 Optische Systeme

Optische Mocap-Systeme zeichnen sich dadurch aus, dass Erfassung und Aufzeichnung von Bewegungen durch mindestens zwei hochauflösende Kameras erfolgt. Ziel ist es dabei, die Position und Rotation der einzelnen Körperteile zu bestimmen (Moeslund et al., 2006).

Um dies zu erreichen werden meist sogenannte „Marker“ benutzt, die am ganzen Körper befestigt werden und entweder an den Kameras angebrachtes Licht reflektieren (passive Marker) oder selbst ein Signal emittieren (aktive Marker). Erfassen nun mehrere Kameras einen Marker, kann die dreidimensionale Position dessen durch Triangulation errechnet werden. Dies geschieht zu jedem Zeitpunkt bzw. für jedes Bild der Aufnahme und für jeden angebrachten Marker, so dass zu jedem Zeitpunkt die Daten der drei Achsen vorliegen. Die Marker sind je nach Anforderung so positioniert, dass jedes gewünschte Gelenk und jedes gewünschte Körperteil abgebildet wird, um die Animation eines authentischen Bewegungsablaufs zu gewährleisten.

Die Unterscheidung der einzelnen Marker kann durch spezielle Algorithmen erreicht werden. Bei aktiven Markern besteht die Möglichkeit, verschiedene Farbspektren, Blinkmuster oder Lichtintensitäten zu wählen.

Statt Marker zu verwenden, sind auch farbige Kleidungsstücke möglich, die für unterschiedliche Teile des Körpers unterschiedliche Farben besitzen. Die Farbwerte werden dann für jedes Bild über einen Algorithmus bestimmt und geben so Aufschluss über Lage und Position des entsprechenden Körperteils (Wang et al., 2011).

Parallel zu den optischen Systemen mit Markern existieren auch optische Systeme ohne Marker oder spezielle Kleidung. Solche Verfahren können Menschen von ihrer Umgebung anhand spezieller Bildverarbeitungsalgorithmen trennen und erfassen (Deutscher et al., 2000). Für diesen Ansatz können beliebige handelsübliche Handkameras verwendet werden (Berger et al., 2011). Es gibt allerdings auch speziell für Bewegungserfassung ausgelegte Kamerasysteme, wie die Microsoft Kinect (Microsoft, 2019b), die Infrarotsensor, Infrarotkamera und RGB-Kamera in einem System vereinen und durch die Kombination der daraus erzeugten Bilder, Menschen und Objekte klar erkennen können (Biswas and Basu, 2011); (Berger et al., 2011).

Marker-basierte Systeme sind bei richtigem Einsatz extrem genau, teilweise auf 200 Mikrometer (Aurand et al., 2017). Allerdings können die Marker verrutschen oder falsch angebracht werden. Ein weiteres Fehlerpotential ist, dass Marker für Kameras verdeckt sein können, was dazu führt, dass diese Kameras für die Triangulation des Markers entfal-

len. Es werden folglich mitunter sehr viele Kameras benötigt. Des Weiteren beanspruchen diese System sehr viel Platz und sind durch die spezielle Hardware insgesamt sehr teuer. Markerlose Systeme sind um einiges günstiger, oft portable Systeme und vermeiden die oben genannten Fehlerpotentiale der Systeme mit Markern. Die Erfassung erfordert insgesamt auch deutlich weniger Aufwand. Nachteilig ist, dass sie nicht sehr genau und weniger robust gegenüber Fehlern sind. Zudem müssen für einige bestimmte Systeme, wie die Kinect, die Darsteller dem System zugewandt sein, damit die Animationen erkannt werden (Han et al., 2013).

2.2.2 Nicht-optische Systeme

Neben optischen Systemen für Motion Capture existiert eine große Zahl alternativer Lösungen, die beispielsweise mit Hilfe von Licht, Magnetfeldern, Strom, Ultraschall oder Trägheit realisiert werden. Es kann zum Beispiel anhand von Licht gemessen werden, wie stark ein Glasfaserkabel geknickt und damit wie stark ein Gelenk angewinkelt ist. In einem Magnetfeld kann die Stärke von drei senkrecht aufeinander liegenden Magnetfeldern an einem bestimmten Punkt gemessen und die Position trianguliert werden (Yabukami et al., 2000). Ein Ultraschall-Impuls kann dazu benutzt werden, um die Position von Körperteilen zu ermitteln, indem die Zeit, die der Impuls von einem Sender zu mehreren Empfängern benötigt, in Abhängigkeit zueinander gemessen wird (Okun et al., 2010, S. 338).

Je nach eingesetzter Technik existieren verschiedene Vor- und Nachteile. Lösungen mit Magnetfeldern beispielsweise durchdringen den menschlichen Körper und können nicht wie Marker verdeckt werden, sind dafür aber sehr störanfällig gegenüber metallischen oder elektrischen Gegenständen. Mechanische Systeme, die ähnlich eines Exoskeletts getragen werden, nehmen Bewegungen in Echtzeit wahr und sind sehr robust gegenüber Störungen von außen, müssen aber oft kalibriert werden und können weder die absolute Position und die Blickrichtung bestimmen, noch den Bezug zum Boden herstellen. Im Gegensatz zu optischen Systemen vereinen sie aber alle den Vorteil, sehr kostengünstig zu sein.

Immer beliebter sind ganze Motion Capture-Anzüge, die dazu benutzt werden, um den kompletten Körper in Echtzeit zu animieren. Eingesetzt wird bei diesen Anzügen ein Netzwerk von sogenannten „Smart Inertial Measurement Units“ (kurz: IMUs), welche an den entscheidenden Stellen innerhalb des Anzugs untergebracht werden. Meist besitzen Anzüge 19 - 20 solcher Trägheitssensoren. Da eine IMU unter anderem einen Beschleunigungs-, einen Gyroskop-Sensor und ein Magnetometer enthält, können Winkelgeschwindigkeit und Geschwindigkeit im Raum, abhängig von der Startposition, zu jeder

Zeit bestimmt werden (Brigante et al., 2011). Bei bekannter Startposition kann auch die absolute Position bestimmt werden. Vorweg muss bei diesen Anzügen eine Kalibrierung durchgeführt werden. Vorteile bei dieser Technik, die teilweise auch als Inertial Motion Capture bezeichnet wird (Xsens, 2019a), sind unter anderem die leichte Tragfähigkeit und die damit verbundene Bewegungsfreiheit. Sie können in engen Räumen oder auch außerhalb von Gebäuden eingesetzt werden (Rokoko, 2018). Des Weiteren sind sie kabellos und sind sehr robust gegenüber äußeren Einflüssen. Nachteilig ist allerdings, dass Messungenauigkeiten oder kleinere Messfehler sich bei einer Ableitung über Zeit fortführen und vor allen Dingen vergrößern können. Um dieser Problematik entgegen zu wirken, müssen die Sensoren bzw. der Anzug in regelmäßigen Abständen neu kalibriert werden.

Beispiele für Motion Capture Anzüge sind der „Xsens MVN“ (Xsens, 2019b), der „Smartsuit Pro“ (Rokoko, 2019) oder der „Percepton Neuron“ (Noitom, 2019).

2.2.3 Inverse Kinematik

Zur Unterstützung von bestimmten Motion Capture-Techniken kann mit inverser Kinematik gearbeitet werden. Inverse Kinematik errechnet anhand der Positionen und Orientierung von beispielsweise Händen und Füßen die Winkel der Gelenke zwischen diesen und dem Körper und richtet sie danach aus. Berücksichtigt werden dabei die Freiheitsgrade der Gelenke, so dass unmögliche Stellungen vorab ausgeschlossen werden können. Wird beispielsweise eine Hand in eine Pose gebracht, so positioniert inverse Kinematik Handgelenk und Ellenbogen, aber auch die Schulter, um so eine authentische Ausrichtung des Charakters zu gewährleisten.

Inverse Kinematik kann in der Performance Animation dazu benutzt werden, um Gelenke, wie Ellenbogen- oder Kniegelenk, die nicht von Sensoren abgedeckt sind, authentisch zu animieren (IKinema, 2019b).

Die direkte Kinematik oder auch Vorwärtskinematik stellt, wie in Abbildung 4 zusehen, das Gegenstück zur inversen Kinematik dar. Diese errechnet anhand der Gelenkwinkel die Position von zum Beispiel Händen oder Füßen.

Eingesetzt wird inverse Kinematik neben der Produktion von Animationen auch in der Robotik, um so Endstücke, die an Gelenken hängen, richtig auszurichten (Petrescu and Petrescu, 2016).

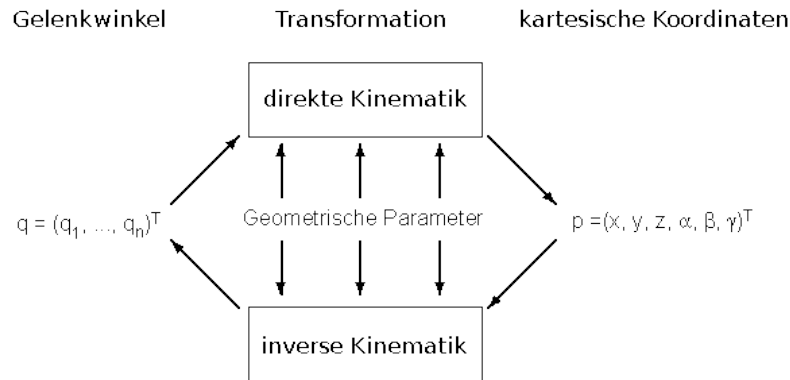


Abbildung 4: Direkte und inverse Kinematik. (Wikipedia, 2019)

2.2.4 Motion Capture in Virtual Reality

Beispiele für die Kombination von Motion Capture und Virtual Reality finden sich heutzutage sehr viele. So existieren Systeme, die Teile des Körpers in VR animieren oder den ganzen Körper mit Ausnahme von Finger, Zehen oder dem Gesicht.

Ein Beispiel für Teilanimation ist Dexmo. Dexmo ist ein leichtgewichtiges Exoskelett für die Hand, welches Hand- und Fingerbewegung in Echtzeit erkennt und in die VR überträgt. Durch die Positionierung oberhalb des Handrückens ist es möglich, dass Nutzer Gegenstände oder auch Controller von VR-Systemen greifen, ohne dabei behindert zu werden (Gu et al., 2016).

Ein zweites Beispiel ist die Software Orion von IKinema (IKinema, 2019b). Ein System, welches auf die Animation vom ganzen Körper ausgelegt ist und mit verschiedener Hardwarekonfiguration gekauft werden kann. Grundlage der Hardwareausstattung stellt dabei das VR-System „HTC Vive“ dar. Auswählbare Hardwarekomponenten sind in Abbildung 5 zu sehen.

Wird die Konfiguration mit weniger Hardware ausgewählt, so werden die Gelenke und Körperteile, die nicht von Sensoren erfasst werden, über inverse Kinematik animiert. (IKinema, 2019b). Anders als bei Dexmo halten die Nutzer dauerhaft die beiden Controller der Vive in Händen, was die Möglichkeiten der Hände einschränkt. Des Weiteren müssen sie die verschiedenen Sensoren, wie in Abbildung 5 zu sehen, tragen, welche verrutschen und zu Bewegungseinschränkungen führen können. Die zwei großen Vorteile des Systems sind die Präzision, da optisches Tracking benutzt wird und der geringe Kostenfaktor. Außerdem wird eine Schnittstelle zu den bekannten Spieleengines „Unity“ und „Unreal Engine“ bereitgestellt.

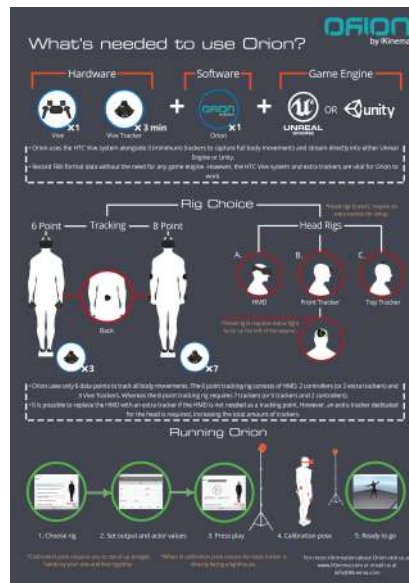


Abbildung 5: Mögliche Hardwarekonfigurationen der Software Orion (IKinema, 2019a)

Die Firma Cloudhead Games nutzte für ihre Produktion des Spiels „The Gallery“ eine Kombination des Motion Capture-Anzugs „Perception Neuron“ der Firma Noitom und der HTC Vive, um einen Charakter von einem Schauspieler animieren zu lassen. Der Schauspieler durfte eine bereits fertige Szene innerhalb der VR betreten und dort den Platz des Charakters einnehmen. Dadurch konnte er die Stimmung und tatsächliche Umgebung der Szene wahrnehmen und sich daran anpassen, um eine noch authentischere Animation zu erzeugen. Durch diese kostengünstige Variante kann auch ein kleines Spielestudio, wie Cloudhead Games, Motion Capture durchführen, ohne ein teures und platzintensives optisches System zu benutzen (CloudheadGames, 2016).

Auch zu anderen Motion Capture-Anzügen existieren bereits Lösungen zur Kombination mit VR. Felix Unger entwickelte in seiner Masterarbeit „Performance Animation in VR“ an der Universität Bremen (Unger, 2018) ein Programm, welches den „Smartsuit Pro“ von Rokoko (Rokoko, 2019) mit der HTC Vive kombinierte. Der Vorteil dieses Programms liegt darin, dass eine Person genügt, um dieses zu bedienen, da die Produktionssteuerung über die Controller der Vive durchgeführt werden kann. Mit Produktionssteuerung ist die Steuerung rund um die Produktion der Animation gemeint, wie beispielsweise das Starten und Stoppen einer Animationsaufnahme. Nachteilig ist die Störanfälligkeit gegenüber Metall oder Strahlung von elektronischen Geräten in der direkten Umgebung.

Die Firma Xsens vertreibt ein Produkt, welches ihren eigenen Motion Capture-Anzug

„Xsens Motion Capture“ mit der Oculus Rift verknüpft, um so in Unity Animationen zu erzeugen. Weitere Firmen, die Motion Capture mit Virtual Reality kombinieren, sind beispielsweise Optitrack (OptiTrack, 2019) oder Motion Workshop (MotionWorkshop, 2019).

Über Inverse Kinematik lässt sich Motion Capture auch ohne externes MoCap-System in VR-Systemen durchführen. Spezielle Plugins wie „Final IK“ für Unity animieren einen Charakter nur anhand der Position und Ausrichtung des VR-Headsets. Die Animation aller Körperteile, die nicht von Sensoren abgedeckt sind, kann von diesen Plugins automatisch durchgeführt werden. Dabei können verschiedene Konfigurationen vorgenommen werden wie Geschwindigkeit oder Elastizität der Körperteile.

So ist es möglich, lediglich mit einem VR-Headset einen virtuellen Charakter in die Knie oder auf Zehenspitzen gehen zu lassen, sich zu drehen oder sich nach vorne oder nach hinten zu bewegen. Sind ebenfalls Controller inkludiert, können auch die Arme animiert werden und nur das Laufen geschieht automatisch. Sind zusätzlich Fußsensoren vorhanden, kann auch der Unterkörper komplett animiert werden. Diese Technik stößt je nach Anzahl an Sensoren an ihre Grenzen, denn beispielsweise wird das Beugen des Oberkörpers in eine Richtung ohne Fußsensoren als ein Schritt interpretiert.

Konkrete Anwendungsgebiete sind neben der Spieleproduktion z.B. Trainingsprogramme, wie das VR Dance Training System von Chan et al (Chan et al., 2011). In diesem System können mit Hilfe eines virtuellen Lehrers, der Tanzschritte vortanzt und verschiedenen Arten von Feedback, die ein Nutzer bekommt, Tanzschritte selbst erlernt werden.

Die Performance Animation innerhalb der VR, in der das Interagieren mit Gegenständen, Charakteren und der virtuellen Umgebung an sich ein alltäglicher Prozess ist, werden die Grenzen der virtuellen Realität schnell spürbar. Einfach Ereignisse, wie ein Händeschütteln mit einer anderen Person, die in der realen Welt mehrmals täglich nebenher geschehen, sind mit einem virtuellen Charakter eine große Herausforderung. Anders als bei Bewegungsabläufen, die einen virtuellen Charakter nur kurz oder einmalig berühren, muss beim Händeschütteln die Hand des virtuellen Charakters zur richtigen Zeit an einem bestimmten Ort ergriffen und in der gleichen Geschwindigkeit geschüttelt werden, damit die Bewegung real erscheint. Minimale Abweichungen in der Bewegung stören die authentische Wirkung. Ohne eine intensive Nachbearbeitung scheint dieses unscheinbare Szenario in der Echtzeitanimierung unmöglich. Bereits die Synchronisation des zeitlichen Ablaufs zu perfektionieren und vorweg korrekt zu antizipieren, ist ohne technische Hilfe unvorstellbar. An diesem Punkt werden daher verschiedene Konzepte eingesetzt, um zumindest

Teillösungen zu dieser Problematik zu schaffen. Eines dieser Konzepte ist die Antizipation.

2.3 Allgemeine Zeitantizipation

Das Konzept der Antizipation von Ereignissen ist allgegenwärtig. Ampeln im Straßenverkehr beispielsweise signalisieren uns, wann wir anhalten oder fahren bzw. gehen dürfen. Mittlerweile werden sogar Ampeln eingesetzt, die durch einen angezeigten Countdown darüber informieren, wann die Ampel wieder auf grün umschalten wird. Anzeigetafeln des öffentlichen Verkehrs zeigen uns an, wann Verkehrsmittel an Haltestellen abfahren. Computer geben uns vor, wann ein Programm installiert oder aktualisiert, wann eine Datei heruntergeladen oder verschoben oder wann eine Suche abgeschlossen ist. Und in fast jedem Fahrstuhl ertönt ein Signal, kurz bevor sich die Tür öffnet. Antizipationshilfen sind so alltäglich geworden, dass sie als solche nicht mehr richtig wahrgenommen werden.

Schon im ersten Rennspiel „Pole Position“ aus dem Jahr 1982 wurde die Startampel mit Ton hinterlegt und diente so als einer der ersten multisensorische Hilfe zum Antizipieren eines Ereignis innerhalb eines Computerspiels (Abbildung 6). Mittlerweile existieren komplette Spieleserien, die auf dem Antizipieren von Ereignissen basieren. So zum Beispiel die Spiele „Guitar Hero“ (GuitarHero, 2019a) oder „Rock Band“ (RockBand, 2019), in denen die Spieler bestimmte Tasten eines Controllers zur genau richtigen Zeit betätigen müssen, um den richtigen Ton zu treffen und Punkte zu sammeln (siehe Abbildung 7).



Abbildung 6: Antizipation im Spiel Pole Position (PolePosition, 1982)

Im Gegensatz zu den zuvor genannten Antizipationshilfen, die aktiv und bewusst Aufmerksamkeit fordern, um zu einem bestimmten zeitkritischen Ereignis eine Reaktion aus-



Abbildung 7: Guitar Hero 5 (GuitarHero, 2019b)

zulösen, existieren Antizipationshilfen, die den gegenteiligen Effekt bezwecken und den Nutzer vom Ereignis ablenken und zurück zu seinem eigentlichen Arbeitsfluss lenken sollen. So zum Beispiel im Bereich der Softwareentwicklung. Bei der Entwicklung von Programmen mit Zeitaspekten, wo Nutzer auf bestimmte Ereignisse warten, sollen diese Antizipationshilfen den Fokus auf das Ereignis zerstreuen. Zeitangaben werden dafür nur ungefähr angegeben, um den Grad der Informationen von Beginn an zu reduzieren. Die Wahrnehmung wird getäuscht, in dem die vermutete Restdauer kleiner ausfällt als die tatsächliche (Seow, 2008).

2.3.1 Antizipation in der Animation

Eine weitere Art von Antizipation entspringt dem Bereich der Zeichentrickfilm- und Animationsfilmindustrie, in der Zuschauer auf schnelle oder unerwartete Bewegungen von Film- oder Comic-Figuren vorbereitet werden.

In diesem Bereich wurde der Begriff der Antizipation schon sehr früh geprägt und später von Personen wie Lasseter, einem sehr bekannten Produzent von Computeranimationsfilmen, sogar als eine von mehreren fundamentalen Prinzipien traditioneller Animation betitelt (Lasseter, 1987). Durch diese Art von Antizipation werden Bewegungen animierter Figuren nicht zu steif oder zu unnatürlich oder wirken zu abrupt. Wichtig ist dabei laut Lasseter, das richtige Gleichgewicht zwischen Aktionsgeschwindigkeit und Länge der vorweg gegriffenen Antizipation. Dadurch sollen die Zuschauer entscheidende Aktionen nicht verpassen und ihre Aufmerksamkeit zur richtigen Zeit auf die richtige Stelle fokus-

sieren. Bei zu schnellen Aktionen muss die Antizipation entweder sehr lang ausfallen oder die Aktion muss verlangsamt werden, damit die Zuschauer sie nicht verpassen. Ein Beispiel einer solchen Antizipation ist in Abbildung 8 zu sehen.



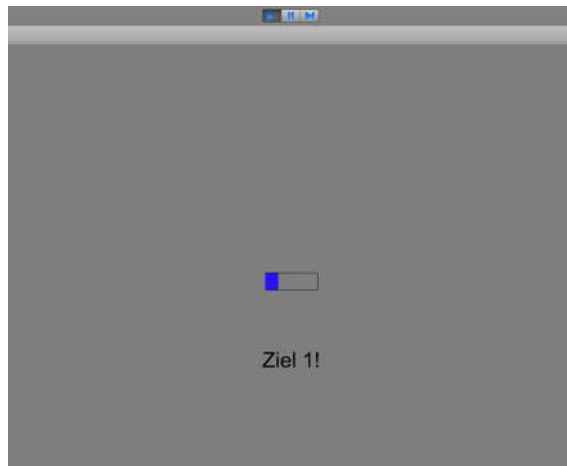
Abbildung 8: Antizipation einer schnellen Bewegung (Lasseter, 1987)

In der Performance Animation, in der die Animation in Echtzeit aufgenommen wird, dienen Antizipationskonzepte hingegen der Vorbereitung auf bestimmte Ereignisse, auf die der Animierende reagieren muss. Als Antizipationshilfen dienen dabei meist Zeitleisten, die den Anfangs- und Endzeitpunkt sowie den aktuellen Zeitpunkt darstellen. Der Abspielpunkt der bereits vorhandenen Animation kann dann so verschoben werden, dass ein besseres Timing möglich ist, so wie z.B. im Multitouch-System von Walter-Franks et al. (Walther-Franks et al., 2011).

In meiner eigenen Bachelorarbeit „Evaluation von Antizipationshilfen im Kontext von Echtzeitsteuerung für Computerspiele und Performance Animation“ an der Universität Bremen (Witt, 2016) habe ich vier verschiedene Arten von visuellen Antizipationshilfen entwickelt, die ich in einem simplen Computerspiel auf ihre Tauglichkeit evaluiert habe. Gegenübergestellt wurden dabei zwei rhythmische und zwei kontinuierliche Lösungen (Abbildung 9). Die rhythmischen Antizipationshilfen starteten jeweils drei Sekunden vor einem eintreffenden Ereignis. Dieser Zeitpunkt war für jedes Ereignis identisch, beispielsweise drei Sekunden vor einem Ereignis. Die kontinuierlichen Antizipationshilfen antizipierten dauerhaft das nächste Ereignis und waren dabei in ihrer Geschwindigkeit abhängig von der zeitlichen Entfernung des nächsten Punktes. Die zwei rhythmischen Antizipationshilfen waren ein klassischer Countdown mit Zahlen und ein Countdown mit Farben, welcher innerhalb eines Kreises von grün über gelb zu rot wechselte. Die kontinuierlichen Antizipationshilfen waren eine sich blau füllende Progressbar und ein rot gefüllter Kreis, der durch abnehmende Transparenz das nächste Ereignis antizipiert.

Es zeigte sich, dass beide Arten von Antizipationshilfen, kontinuierlich und rhythmisch,

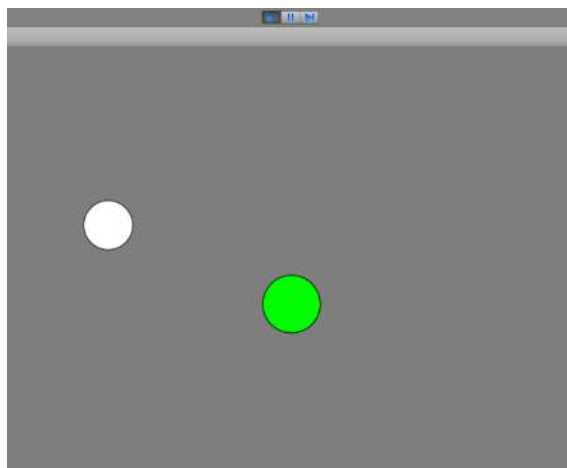
geeignet sind, um präzise das nächste Ereignis vorherzusagen. Aufgrund ihrer salienten Natur, auch außerhalb des Fokusbereiches, wurde die Progressbar gewählt, die sich in abgewandelter Form im Prototyp dieser Arbeit wiederfindet.



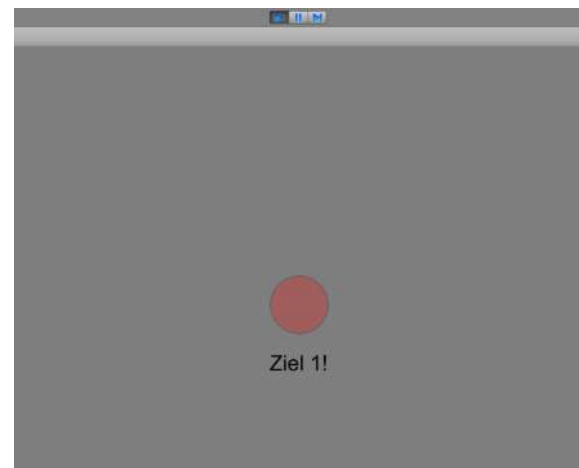
(a) Zeitleiste



(b) Countdown



(c) Ampel mit zunehmender Transparenz zwischen den Übergängen



(d) LED mit abnehmender Transparenz

Abbildung 9: Die vier Antizipationshilfen der Bachelorarbeit „Evaluation von Antizipationshilfen im Kontext von Echtzeitsteuerung für Computerspiele und Performance Animation“ (Witt, 2016)

2.3.2 Antizipation mit Multisensorik

Durch den zunehmenden Einfluss von Echtzeitanimation in der VR werden zeitliche Aspekte automatisch wichtiger. Denn wie zuvor beschrieben, ist das richtige Timing in vielen Animations-Szenarien unabdingbar. Um dies zu erreichen, werden die Nutzer der meisten Systeme mit bestimmten Hilfen unterstützt, um die Ereignisse zu antizipieren.

Die Hilfen zum Antizipieren von Ereignissen können dabei in sehr vielen verschiedenen Formen und basiert auf den menschlichen Wahrnehmungssinnen vorliegen. Die Basis kann dabei aufgrund der Eigenschaften der VR visuell, auditiv oder taktil sein oder auch eine Kombination einsetzen.

Ein Beispiel für eine Kombination mehrerer Sinne ist der sogenannte „Count Ticker“ von Wong (Wong, 2014). Dieser stellt ein erweitertes Metronom für Takte dar, welcher zwei verschiedenen Tönen für starke und schwache Takte in Kombination mit verschiedenen visuellen Zählern und Werten, z.B. für Taktart, Taktanzahl, Taktnummer, Geschwindigkeit, besitzt, die den Nutzer dabei unterstützen sollen, die nächsten Takte vorherzusehen. Ein weiterer Vorteil stellt ein Countdown dar, der vor Aufnahme start herunterläuft und dem Nutzer eine kurze Vorbereitungszeit bereitstellt.

Ein Vergleich der verschiedenen Sinnesmodalitäten oder ihrer Kombinationen bei der Verarbeitung aufgaben-relevanter Stimuli gestaltet sich aus diversen Gründen schwierig. Stimulationsreize können sich mitunter stark unterscheiden, z.B. aufgrund ihrer physikalischen Eigenschaften, ihrer individuell erlebten Intensität oder persönlicher Präferenzen. Nichtsdestotrotz versuchen immer wieder wissenschaftliche Untersuchungen einen Vergleich zwischen der Wahrnehmung und Aufgabenlösung basierend auf Stimulation durch unterschiedliche Sinnesmodalitäten zu untersuchen. Einige Studien deuten im Vergleich der Sinne vor allem auf die Vorteile auditiver Stimulation hin, z.B. im Hinblick auf Lernprozesse (Conway and Christiansen, 2005). Für auditive Stimulation scheinen die Fähigkeiten der Stimulusdiskrimination im Zeitbereich besonders fein abgestimmt zu sein (Sherrick and Cholewiak, 1986; Rubinstein and Gruenberg, 1971). Gerade in komplexen Aufgaben basierend auf zeitlich präziser Verarbeitung von Reizen deuten einige Studien auf die Überlegenheit des auditiven Sinnesmodalität hin. Visuelle Stimulation scheint dagegen bei Aufgaben zur räumlichen Wahrnehmung von Vorteil zu sein (Freides, 1974; Kubovy, 1988; Mahar et al., 1994; Penney, 1989; Saffran, 2002). Taktile Stimulation scheint wiederum für beide Aufgabentypen nicht auf ein gleiches Level der Aufgabenbewältigung zu führen (Mahar et al., 1994). Bei der Verarbeitung von Stimulationen verschiedener Sinne scheinen leicht unterschiedliche Verarbeitungsprozesse und/oder -strategien angewendet zu werden. Taktile sowie visuelle Stimulation ist dabei bei sequentiell Input besonders sensitiv für die Information zu Beginn der Sequenz. Dagegen scheint auditive Stimulation gegen Ende besonders bewusst wahrgenommen zu werden (Conway and Christiansen, 2005). Gerade die Kombination mehrerer Sinnesmodalitäten, die grundsätzlich den gleichen Informationsgehalt widerspiegeln sollen, vermag also unter Umständen die

Stärken der einzelnen Stimulationsarten zu vereinen. Gerade auch die variationsreichere Anregung neuronaler Netzwerke deutet auf einen Vorteil durch multisensorisches Lernen hin (Shams and Seitz, 2008). Übertragen auf die Verwendung von Antizipationshilfen in der VR, könnte eine multisensorische Stimulation den Nutzern ermöglichen, je nach Aufgabe und persönlicher Präferenz für eine Stimulusart jeweils spontan den Fokus zu wählen. So ist auch innerhalb der Antizipation eines Zeitpunktes ein Wechsel des persönlichen Aufmerksamkeitsfokus möglich, ohne dass dies bewusst als Wechsel wahrgenommen werden muss. Demgegenüber stehen die Bedenken, dass eine multimodale Stimulation durch Reizüberflutung auch zu einer Ablenkung weg von der Zielsetzung oder Überforderung und somit zu Stress führen kann.

3 Fragestellung

3.1 Motivation

Im Alltag werden Antizipationshilfen zahlreich eingesetzt. Gerade für zeitkritische Situationen erweisen sie sich als außerordentlich nützlich und sind in vielen Anwendungsbereichen unerlässlich. Gelänge es, das Konzept und die Wirkweise von Antizipation besser zu verstehen, so würde dies eine Entwicklung und Nutzung von Antizipationshilfen erleichtern. Wissenschaftlich fundierte Variationen von Antizipationshilfen in zukunftsorientierte Technik zu integrieren, könnte zu einer effizienzsteigernden und somit kostensenkenden Verbesserung dieser führen.

Die VR mit ihren vielfältigen Anwendungsgebieten stößt durch fehlende physische Präsenz in vielen Szenarien, wie beispielsweise ein einfaches Anrempeln eines virtuellen Charakters, an die Grenzen ihrer Möglichkeiten. Antizipationshilfen, wie Counter, Zeitleisten oder Taktsignale, werden bereits vermehrt eingesetzt, um viele zeitkritische Situationen so effizient wie möglich zu animieren. Der Vorteil solcher Antizipationshilfen liegt darin, dass sie oft intuitiv eine genauere zeitliche Einschätzung zu gewährleisten vermögen. Jedoch fehlen einerseits für komplexe Szenarien passende Hilfskonzepte und andererseits fehlen wissenschaftliche Studien, die sich mit eben dieser Thematik auseinandersetzen und verschiedene Ansätze und Möglichkeiten wissenschaftlich analysieren.

Gerade der Bereich Performance Animation stellt sehr hohe Ansprüche an den Faktor Zeit. Das Erstellen von Szenarien, die mehrere Charaktere beinhalten, bedarf oftmals einiger Zeit und Mühe zuzüglich Aufwand der Nachbearbeitung. Antizipationshilfen können dabei helfen, in der richtigen Kombination eingesetzt, die Effizienz des Arbeitsflusses zu verbessern, den Arbeitsfluss angenehmer zu gestalten und Schritte der Nachbearbeitung zu verringern.

In der hier vorliegenden Arbeit werden die Grenzen der VR und die Problematik von

komplexer Charakteranimation in konkreten Szenarien aufgegriffen und versucht, das Verständnis für die Wichtigkeit von Antizipationshilfen zu verstärken. Verschiedene Wahrnehmungssinne, die innerhalb der VR angesprochen werden können, werden genutzt, um eine Steigerung der Effizienz durch bestimmte Antizipationshilfen zu erreichen. Dadurch sollen oben genannte Probleme minimiert werden.

3.2 Zielsetzung

In der vorliegenden Arbeit wird geprüft, ob für eine Antizipation eines zeitkritischen Momentes in einer VR Szene zusätzliche sensorische Reize - auditiv und taktil - einen Mehrwert zu einer rein visuellen Stimulation bieten können oder der Einsatz paralleler sensorischer Kanäle mit ähnlichem Informationsgehalt zu einer Überforderung oder Ablenkung weg von der eigentlich zu lösenden Aufgabenstellung führt.

Dabei ist von Interesse, ob die Teilnehmenden sich zur Animation auf die Antizipationshilfen stützen und diese als hilfreich und somit nicht als störend empfinden. Es wird untersucht, welche Sinnesmodalitäten in der multisensorischen Variante wahrgenommen werden können und ob Teilnehmende sich gezielt einzelner oder bestimmter Teilkombinationen der Hilfen bedienen. Abschließend wird untersucht, ob die Durchführung beider Versuchsvarianten (rein visuell vs. multisensorisch) eine Belastung für die Teilnehmenden darstellt. Hierbei wird der Reihenfolgeeffekt der beiden Varianten genauer beleuchtet. Dies soll Aufschluss über den Grad der möglichen Überforderung geben.

4 Design und Implementierung

Für die Umsetzung des Prototyps wurde das VR-System Vive der Hersteller „High Tech Computer Corporation“ (kurz HTC) und „Valve Corporation“ verwendet. Wie in Abbildung 10 zu sehen, besteht das System aus einem VR-Headset, zwei Controllern und zwei Basisstationen. Die zwei Basisstationen sind dabei für Positions- und Rotationsbestimmung von Headset und Controllern zuständig, skizziert in Abbildung 11.

Entwickelt wurde in dieser Arbeit ausschließlich mit der Spiele-Engine Unity in der Version 2018.1.0f2, welches mit dem Plugin „SteamVR Plugin“ der Valve Cooperation (Version 1.2.3) eine einfache Einbindung und Nutzung der Vive innerhalb Unitys ermöglicht.

Des Weiteren wurde das Plugin „Final IK“ (Version 1.7) benutzt, auf welches im späteren Verlauf dieses Kapitels näher eingegangen wird.

Der gesamte Quellcode dieser Arbeit, welcher sich in Skripten für Unity manifestiert, wurde in C# geschrieben.



Abbildung 10: HTC Vive (HTC, 2019a)

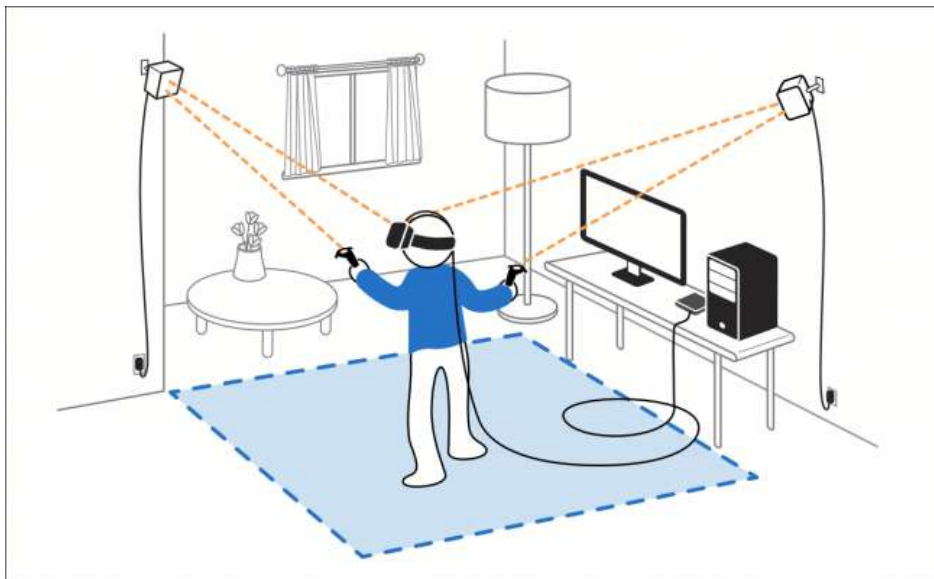


Abbildung 11: Funktionsweise Tracking HTC Vive (HAWK, 2019)

4.1 Quellcode-Basis

Als Grundlage zur Implementierung der Antizipationshilfen wurde der Prototyp 4.1.2 aus der Masterarbeit „Performance Animation in VR“ von Felix Ungar (Ungar, 2018) mit dem Ziel gewählt, die Kopplung eines Motion Capture-Anzug und der Vive zu verwenden, die in dieser Arbeit umgesetzt wurde. Innovativ war dabei der Einsatz des „Smartsuits Pro“ von Rokoko (Rokoko, 2019), einem Motion Capture-Anzug, der im optimalen Fall die Bewegungen des Trägers sehr präzise und in Echtzeit an das angeschlossene System bzw.

Unity übertragen konnte. Dadurch wurde es möglich, innerhalb der VR einen virtuellen menschlichen Charakter fast vollständig zu steuern und seine Bewegungen aufzunehmen. Lediglich Finger- und Zehenbewegungen oblagen der Grenzen des eingesetzten Motion Capture-Anzugs.

Das Aufnehmen der Animation geschah dabei nach dem Prinzip der Layered Performance Animation. Eine erfolgreiche Aufnahme wurde anhand einer Kopie des Charakters in einer Schleife wiedergegeben, so dass an jeder beliebigen Stelle dieser Schleife eine zweite Animation aufgenommen und so in mit der ersten in Zusammenhang gebracht werden konnte. Abbildung 12 stellt den zeitlichen Ablauf von zwei verschiedenen Szenen dar, um das Konzept näher zu erklären.

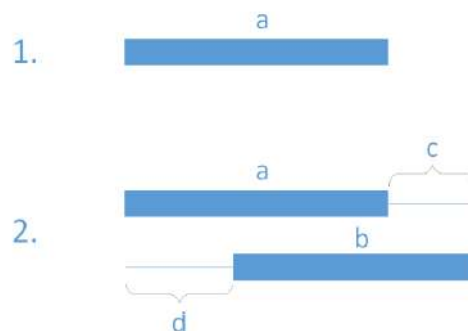


Abbildung 12: Zeitachsen zwei verschiedener Szenen. Beispiel 1 enthält nur eine Animation a, die in einer Schleife immer wieder abgespielt wird.

Beispiel 2 enthält zwei Animationen a und b. Die Aufnahme von b wurde innerhalb von a gestartet und geht über die Animation von a hinaus. Ist dabei das Ende von a erreicht, so stoppt die Animation a und verharrt an ihrem Ende bis die Aufnahme von b gestoppt wird, also der Zeitraum c. Dann wird die Szene zurückgesetzt und die Animation a startet erneut. Bis zu dem Zeitpunkt, an dem b startet, verharrt Animation b in der Startposition, also den gesamten Zeitraum d. Quelle: (Ungar, 2018, S. 17)

4.1.1 Das Plugin Final IK

Zur Vereinfachung der Studie, aufgrund mehrerer technischer Probleme durch neue Versionen und geänderte Lizenzmodelle seitens Rokoko wurde letztlich auf den „Smartsuit Pro“ verzichtet und stattdessen mit dem Unity-Plugin „Final IK“ gearbeitet. Dieses ist in der Lage, über Inverse Kinematic Arm- und Beinbewegungen automatisch zu animieren. Des Weiteren beinhaltet es bereits eine vorkonfigurierte Komponente „VR IK“, welches dem Charakter hinzugefügt werden muss.

Nach der grundlegenden Einrichtung der Komponente wurden Position und Rotation der beiden Hände des virtuellen Charakters manuell so verändert, dass sie, wie in Abbildung

13 zu sehen, an den Vive Controller angepasst sind. Die offene Handfläche wurde dabei bewusst gewählt, um die verschiedenen Aufgaben der Studie erfüllen zu können.

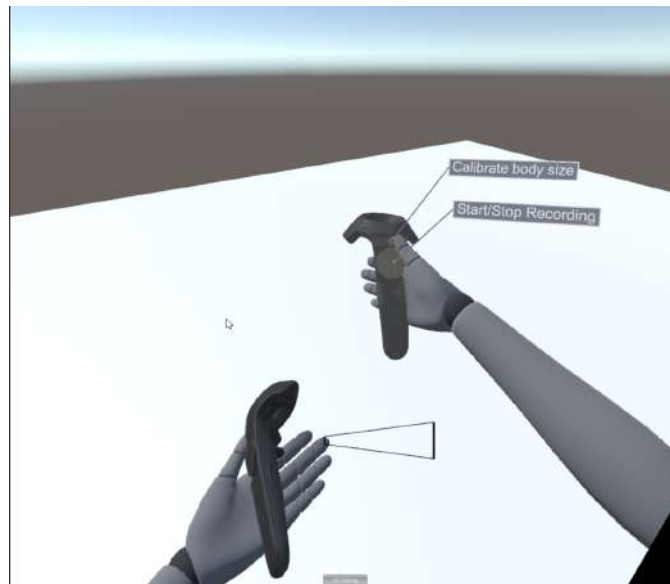


Abbildung 13: An HTC Vive Controller angepasste Handpositionen

Ähnlich der manuellen Handpositionierung wurde auch das VR-Headset auf die Position und Rotation der Augenpartie des virtuellen Charakters abgestimmt. Lediglich Fuß- und Beinbewegungen waren aufgrund fehlender Sensoren auf diese Weise nicht umzusetzen. Final IK stellt allerdings für solche Fälle Funktionen bereit, die diese Problematik zumindest teilweise umgehen. So hat die Komponente VR IK eine Funktion für automatisches Laufen, welches anhand der Drehung und Position der VR-Brille automatisch ein Gehen oder Drehen des Charakters animieren kann. Diese Funktion kann durch verschiedene Parameter konfiguriert und damit auf die eingesetzte virtuelle Figur angepasst werden. Wie in Abbildung 14 zu sehen, können unter anderem Schrittgröße, -schnelligkeit, die Schwerkraft, der Abstand der Füße bis zu bestimmten Grenzen gesteuert werden. Anhand der Höhe der Brille erkennt Final IK auch ein Stehen auf Zehenspitzen und ein in die Knie gehen. Ein Vor- oder Zurückbeugen des Oberkörpers wird allerdings nicht erkannt und resultiert in Schritten nach vorne bzw. hinten.

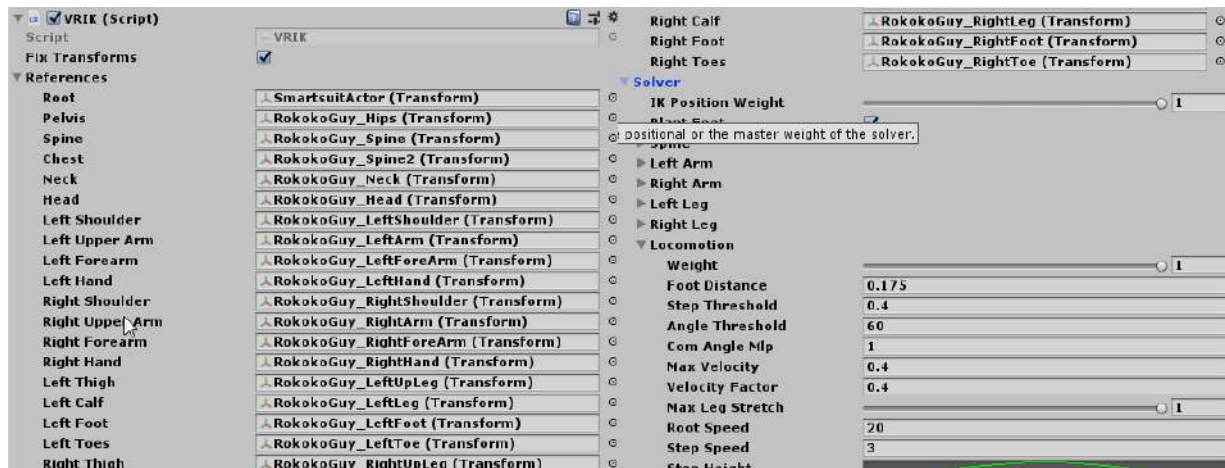


Abbildung 14: Konfiguration der Komponente VR IK des Unity-Plugins Final IK

4.1.2 Bereits vorhandene Funktionen

Trotz des Wechsels von Motion Capture-Anzug auf Inverse Kinematic wurden die meisten Grundfunktionen, die durch den Prototyp bereitgestellt wurden, beibehalten. Dazu zählen unter anderem die eigentliche Aufnahme von Animationen, die Teleportation innerhalb der VR und das Verschieben von bereits existierenden Aufnahmen. Diese werden im Folgenden kurz dargestellt.

Aufnahme:

Mit Druck auf die Taste „Start/Stop Recording“ des rechten Controllers kann eine Aufnahme gestartet werden. Eine gestartete Aufnahme wird anhand eines blinkenden roten Punktes und des Textes „Recording...“ signalisiert. Ab dem Zeitpunkt des Aufnahmestarts werden alle Bewegungen, die vom Nutzer gemacht werden, aufgenommen. Sobald die Aufnahme gestoppt wird, wird eine Kopie des virtuellen Charakters erstellt, die exakt dort startet, wo auch der Nutzer seine Aufnahme gestartet hat. Diese wird nun, je nachdem, ob es sich um die erste Aufnahme handelt, in einer Dauerschleife wiederholt oder passt sich, wie schon in Abbildung 12 beschrieben, der schon vorhandenen Szene an.

Teleportation:

Dem Nutzer ist es möglich, innerhalb der weißen Teleportationsplattform zu teleportieren (Abbildung 15). Dazu muss am linken Controller die Taste „Teleport“ gedrückt und gehalten werden. Es erscheint eine Animation, die je nach Ziel anzeigt, ob und wohin teleportiert werden kann. Dies hat den Sinn, das auch in kleineren Räumen, der Platz innerhalb der VR voll ausgenutzt werden kann.

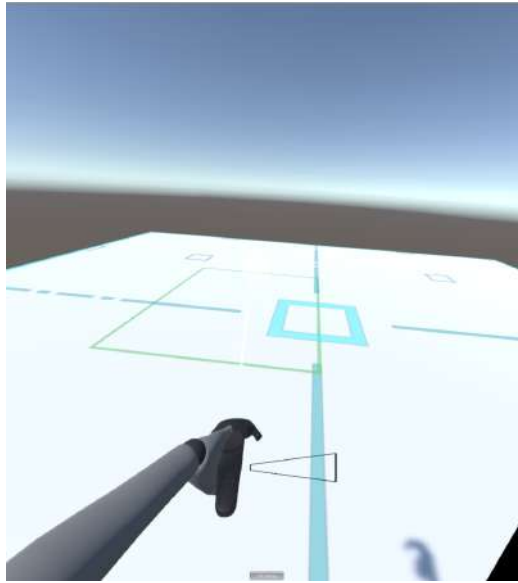


Abbildung 15: Teleportation oberhalb des vordefinierten Bodens

Verschieben von Animationen:

Bereits aufgenommene Animationen können verschoben werden. Dadurch ist es möglich, die Position und auch Ausrichtung zu verändern. Dazu muss der Nutzer so nah an eine Animation herantreten, dass er sie mit dem linken oder rechten Controller erreichen kann. Sobald der Nutzer mit dem Controller die Animation berührt, färbt sich seine Umrandung gelb ein. Betätigt der Nutzer in diesem Zustand die Taste „Move Body“ des entsprechenden Controllers und hält diese Taste gedrückt, so hat er den Charakter „gegriffen“. Bis zum Loslassen der Taste ist die Animation frei verschiebbar.

Kalibrierung:

Da die Kalibrierungsfunktion des Anzuges nicht mehr benötigt wurde, wurde sie durch eine eigene ersetzt. Diese skaliert den virtuellen Charakter je nach Körpergröße des Nutzers. Dabei wird die Position des VR-Headsets im Bezug zur errechneten Körpermitte betrachtet.

Der Nutzer kann durch Betätigen der Taste „Calibrate body size“ am rechten Controller die Kalibrierungsfunktion aufrufen und die Größe des virtuellen Charakters ändert sich.

4.1.3 Neue Funktionen

Alle anderen Funktionen, die den Smartsuit Pro betreffen wurden entfernt. Hinzu kamen neue Funktionalitäten, die für den weiteren Verlauf wichtig waren. Erst im Anschluss daran folgte das Einbauen der eigentlichen Antizipationshilfen.

Zusätzlich zu den eben vorgestellten Funktionen wurden folgende Funktionalitäten hinzugefügt:

- Einführung einer Zeitleiste
- Einführung eines Hauptmenü
- Speichern einer Szene
- Laden einer Szene
- Pausieren einer Szene / Pause Modus
- Neue Szene erstellen

Zeitleiste:

Bei der bisherigen Programmierung gab es keinen Hinweis darauf, wann eine Szene endete bzw. wieder erneut startete. Ist also mindestens eine Animation aufgenommen worden und wird in einer Schleife wiedergegeben, konnte der Start dieser Schleife bislang nicht ersichtlich wahrgenommen werden. Generell war es ohne Kenntnis der Szene nicht möglich, die genaue zeitliche Position in dieser zu bestimmen. Daher wurde eine beschriftete Zeitleiste eingeführt, die einerseits die Szenenlänge in Sekunden angibt und andererseits mit einem laufenden Balken die genaue zeitliche Position innerhalb der Szene kenntlich macht.

Hauptmenü:

Durch das Betätigen der Taste „Main Menu“ des rechten Controllers erscheint ein Hauptmenü. Die Szene wird dabei pausiert. Dort kann nun in den Pausemodus gewechselt werden, eine Szene gespeichert, geladen oder eine neue Szene erstellt werden. Die Auswahl geschieht über Betätigen des Touchpads des rechten Controllers, sofern der Menüpunkt anvisiert wurde (Abbildung 16).

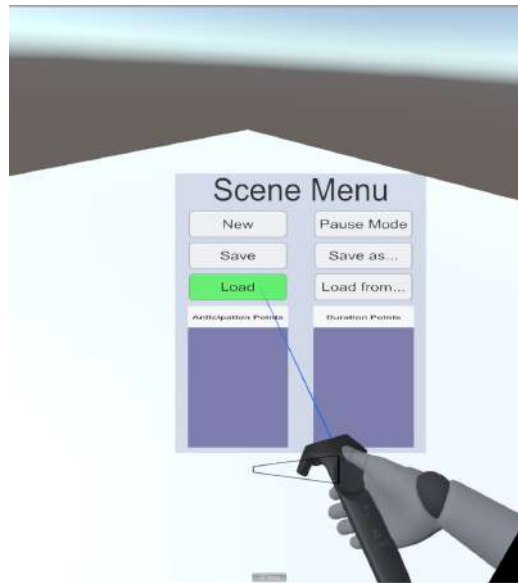


Abbildung 16: Hauptmenü des Prototyps

Speichern einer Szene:

Wählt der Nutzer im Hauptmenü „Save“, so wird die aktuelle Szene mit allen Aufnahmen sowie Antizipations- und Interaktionspunkten schnell gespeichert. Es wird also eine Datei mit dem Namen „Scene.gd“ innerhalb des Unity-Projektordners angelegt. Wird zu einem späteren Zeitpunkt erneut der Punkt Speichern ausgewählt, wird diese Datei überschrieben. Sobald die Szene gespeichert wurde, schließt sich automatisch das Hauptmenü und die Szene wird fortgeführt. Der Vorteil dieser Speicherfunktion ist, dass die VR zum Abspeichern so nicht verlassen werden muss.

Wählt der Nutzer im Hauptmenü „Save as...“ so öffnet sich ein Dialog auf dem auszuführenden Rechner, also nicht in der VR, in dem der Nutzer nun einen Pfad auswählen sowie einen Namen bestimmen kann.

Laden einer Szene:

Wählt der Nutzer im Hauptmenü „Load“ aus, so wird die zuletzt per „Save“ gespeicherte Szene geladen, sofern eine vorhanden ist. Sobald das Laden abgeschlossen ist, schließt sich das Hauptmenü automatisch und die geladene Szene, inklusive aller Animationen und Antizipationspunkte, wird abgespielt. Die zuvor bearbeitete Szene ist dann nicht mehr verfügbar.

Wählt der Nutzer „Load from...“ öffnet sich ein Dialog auf dem auszuführenden Rechner, in dem der Nutzer nun eine Datei auswählen kann, die dann versucht wird, zu laden. Ist dies möglich, schließt sich das Hauptmenü automatisch und startet die in der Datei gespeicherte Szene. Die zuvor bearbeitete Szene ist auch hier nicht mehr verfügbar.

Pausieren einer Szene / Pausemodus

Der Nutzer hat die Möglichkeit in einen Pausemodus zu wechseln. Dafür muss im Hauptmenü der Punkt „Pause Mode“ ausgewählt werden. Das Hauptmenü wird anschließend geschlossen, die Szene bleibt dabei pausiert. Nun hat der Nutzer die Möglichkeit, sich frei in der Szene zu bewegen, zu teleportieren und auch Animationen zu verschieben. Lediglich das Starten einer neuen Aufnahme ist nicht möglich. Der Nutzer kann den Pause-Modus verlassen, in dem er das Hauptmenü öffnet und wieder schließt. Daraufhin wird die Szene an der Stelle fortgesetzt, an der sie angehalten wurde.

Neue Szene erstellen:

Durch die Auswahl „New Scene“ im Hauptmenü wird die aktuelle Szene gelöscht und eine neue, leere Szene wird aufgerufen. Bereits aufgenommene Animationen und auch jedwede Änderung werden dadurch verworfen.

4.2 Antizipationshilfen

Durch das Hinzufügen der eben vorgestellten Funktionen wurde die Grundlage geschaffen, um eine funktionierende Szenenbearbeitung innerhalb der VR durchzuführen. Auf Basis dessen wurden die Antizipationshilfen integriert.

4.2.1 Design der visuellen Antizipationshilfen

Als visuelle Antizipationshilfe wurde die „Progressbar“ (Abbildung 17) gewählt. Diese stellt, wie in meiner Bachelorarbeit nachgewiesen werden konnte, ein gutes visuelles Konzept zur Antizipation dar und weist zusätzlich durch ihre kontinuierliche Bewegung eine große Salienz auf. Das Grundprinzip der Progressbar ist, dass ein Rechteck am Startpunkt komplett leer und am Endpunkt komplett farblich gefüllt ist. Vom Start- zum Endzeitpunkt füllt sich die Progressbar in einer kontinuierlichen Geschwindigkeit abhängig von der zeitlichen Entfernung des Punktes. Ist ein Punkt also zeitlich sehr nah am Startpunkt, füllt sie sich schneller im Vergleich zu einem Punkt, zeitlich sehr weit entfernt ist. Der Endzustand spiegelt dabei den zu antizipierenden Zeitpunkt wider.

Sind mehrere Antizipationspunkte vorhanden, so sind sie abhängig voneinander. Da es nur eine Progressbar gibt, die dauerhaft sichtbar ist, wird immer der nächste zu antizipierende Zeitpunkt dargestellt. Es wird also kontinuierlich das nächste Ereignis, sofern vorhanden, antizipiert. Die Position wurde dabei am unteren Sichtfeldrand gewählt, um

so wenig Sichtfeld wie möglich zu verdecken.

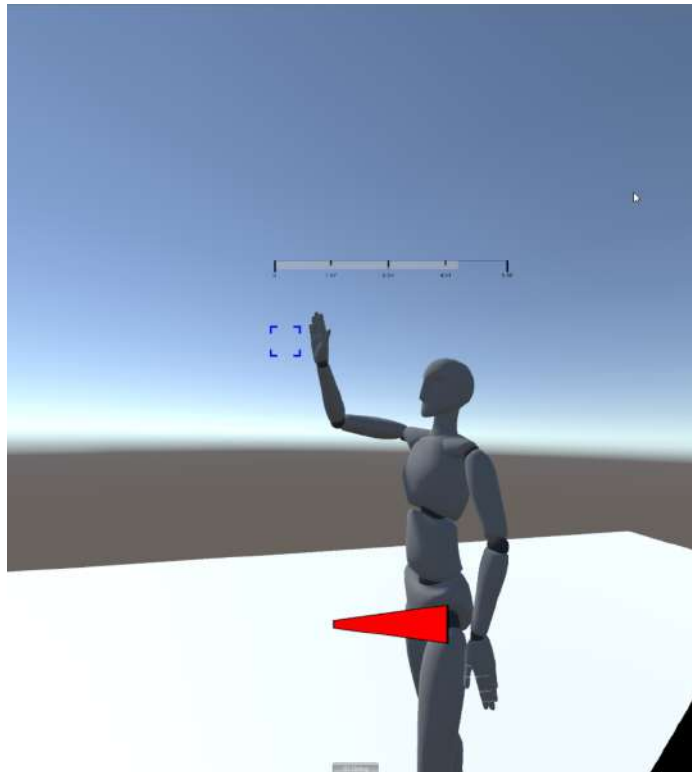


Abbildung 17: Erstes Szenario: Ein virtueller Charakter läuft ein paar Schritte nach vorne und macht ein „High Five“

In einem Experteninterview mit einer Mitarbeiterin von „Arx Animation“, einem Animationsstudio für High-End Charakteranimation (ArxAnima, 2019), wurde das Fazit gezogen, dass die Progressbar und das Erreichen des Endzustands einer erhöhten Salienz bedarf. Um dies zu erreichen, wurde einerseits die rechteckige Form in eine Trichterform umgewandelt, so dass der Endzustand der Progressbar die breiteste Stelle darstellt. Darüber hinaus wurde die Salienz weiter verstärkt, indem eine Interpolation der Füllfarbe von grün als Startfarbe zu rot als Endfarbe eingefügt wurde.

Des Weiteren wurde in dem Interview das Fazit gezogen, dass eine örtliche Markierung im VR-Raum eine essentielle Ergänzung wäre, da zumeist Ereignisse, die antizipiert werden, an einen bestimmten Ort gekoppelt sind. Daher wurde eine Positionsmarkierung in Form eines blauen Quadrates eingearbeitet.

Um auch auf Ereignisse außerhalb des Sichtfelds aufmerksam zu machen, wurde zusätzlich ein kleiner zweidimensionaler Pfeil implementiert, der nur sichtbar ist, sofern sich die Positionsmarkierung nicht im Sichtfeld des Nutzers befindet. Der kleine Pfeil zeigt die Rich-

tung an, in der das Ereignis stattfinden wird.

Nach Umsetzung der Antizipationskonzepte wurde als letzte Ergänzung eine weitere zeitliche Variable mit in den Prototyp aufgenommen: Die Interaktionsdauer mit Charakteren bzw. bereits vorhandenen Animationen. Durch einen Kreis, welcher sich mit der Zeit füllt, soll vermittelt werden, wie lange mit einem bestimmten Charakter interagiert werden soll. Anders als bei der Progressbar verschwindet dieser Kreis nach seinem Endzustand, sodass der Nutzer weiß, dass der Fokus zu diesem Zeitpunkt nicht mehr auf diesem Charakter liegt. Dadurch sollen komplizierte Szenen mit mehreren zu animierenden Handlungen übersichtlicher gestaltet werden. Diese zeitlichen Kreisdiagramme existieren parallel zu den Antizipationspunkten, sind aber vom Grundprinzip her gleich. Sie füllen sich mit einer konstanten Geschwindigkeit abhängig vom zeitlichen Abstand des nächsten Punktes. Es existiert immer nur ein Kreis zur gleichen Zeit und erst sobald dieser abgeschlossen ist, erscheint der nächste. Während also die Progressbar den nächsten Handlungszeitpunkt verdeutlicht, dient das Kreisdiagramm der Verdeutlichung, wie lange mit einem Charakter interagiert werden soll. Es ist also möglich, dass in einer Szene mehrere Handlungen und somit Handlungszeitpunkte in einem Kreisdiagramm zeitlich zusammengefasst sind.

4.2.2 Design der multisensorische Antizipation

Bei den multisensorischen Antizipationshilfen wurde eine synchrone Stimulation gewählt. Taktile und auditive Signale sollten also zeitgleich und im selben Rhythmus zu spüren bzw. zu hören sein. Dies diente der vereinfachten Wahrnehmung der multisensorischen Stimulation ohne Reizüberflutung durch deren starke Varianz. So wurde eine Vereinfachung des Vergleiches zwischen visueller und multisensorischer Antizipation bewirkt.

Bei den taktilen Signalen bietet die Vive nur die Möglichkeit, die beiden Controller vibrieren zu lassen. Dabei kann zwar die Länge der Vibration, jedoch nicht die Vibrationsstärke eingestellt werden. Daher wurde durch Eigenversuche zunächst die kleinste spürbare Pause zwischen zwei taktilen Signalen ermittelt. Es wurde festgestellt, dass eine Pause von unter 50 ms nicht spürbar war. Anhand dieses Wertes wurde ein Algorithmus geschrieben, der eine Signalreihe erzeugte, die drei Sekunden vor Eintreten des zu antizipierenden Events startete. Diese Reihe gab mehrere Signale mit Pausen unterschiedlicher Länge zwischen den Signalen aus. Die Pausen wurden dabei zunehmend kürzer und die Signale kamen in immer kleiner werdenden Abständen bis hin zum Endzeitpunkt der Antizipation. Die letzte Pause hat exakt die Länge von 50 ms und war damit noch als solche gerade noch wahrnehmbar.

Basierend auf diesem Algorithmus wurden synchron zu den taktilen Signalen auditive Si-

gnale ausgegeben. Dabei wurden drei verschiedene Töne ausgewählt, die jeweils drei Mal hintereinander ertönten. Um die Salienz zu erhöhen und damit dem Nutzer beim Antizipieren zu helfen, wurde die Reihenfolge so gewählt, dass die Töne zum Endzeitpunkt der Antizipation hin immer höher und schriller wurden.

4.2.3 Steuerung

Für das Setzen und Entfernen der zuvor beschrieben zeitlichen Punkte wurden jeweils Funktionen zur Verfügung gestellt. Die Voraussetzung für das Einfügen eines zeitlichen Punktes ist, dass bereits eine Animation, also eine Aufnahme, existiert, die in einer Schleife wiedergegeben wird.

Setzen von Punkten:

Durch Betätigen der Taste „Set Point“ kann der Nutzer einen Antizipationspunkt in Form eines blauen Quadrates setzen. Der aktuelle Zeitpunkt bei Betätigung der Taste ist dabei der Zeitpunkt, der beim nächsten Schleifendurchlauf antizipiert wird. Sind bereits mehrere Punkte gesetzt, so wird dieser Punkt abhängig von den anderen zeitlich dazwischen eingefügt. Das Setzen eines Punktes wird durch ein kurz aufleuchtendes Textfeld mit der exakten Zeit des Punktes signalisiert.

Berührt der linke Controller beim Betätigen der Taste eine Animation, ist also gelb umrandet, wird ein Interaktionspunkt in Form eines Kreisdiagramms auf dem ausgewählten Charakter gesetzt. Dieser Punkt wird beim nächsten Schleifendurchlauf berücksichtigt. Sind auch hier bereits mehrere Punkte gesetzt, wird er passend eingefügt. Auch das Setzen dieses Punktes wird durch einen kurzen Text verdeutlicht.

Entfernen von Punkten:

Betätigt der Nutzer die Taste „Remove Point“ wird der Antizipationspunkt entfernt, der gerade antizipiert wird und es wird unmittelbar der nächste Punkt antizipiert. Wird beim Betätigen der Taste mit dem linken Controller eine Animation berührt und dort existiert ein Interaktionspunkt, der gerade antizipiert wird, wird dieser entfernt und es erscheint unmittelbar der nächste Kreis für den nächsten Interaktionspunkt. Auf das erfolgreiche Entfernen eines Punktes wird ebenfalls mit einem kurzen Textfeld hingewiesen.

Da zwei Punkte durch falsches oder doppeltes Setzen unmittelbar hintereinander sein

können und es für den Nutzer mit der oberen Löschkfunktion schwierig wäre, den hinteren der beiden Punkte zu entfernen, wurde eine weitere Möglichkeit geschaffen, Punkte zu löschen. Dazu kann das Hauptmenü geöffnet werden. Dort sind die gesetzten Punkte in zwei Tabellen aufgezeigt. Wird ein Punkt mit dem Laser anvisiert und das Touchpad des rechten Controllers betätigt, wird dieser Punkt entfernt. Das Hauptmenü bleibt danach geöffnet, falls der Nutzer weitere Punkte löschen möchte.

Das Umschalten zwischen den Versuchsvarianten (visuell vs. multisensorisch) muss in Unity direkt erfolgen. Dafür wurde im Hinblick auf die Studie keine eigene Funktion bereitgestellt. Damit sollte vermieden werden, dass Studienteilnehmenden bewusst oder unbewusst die Variante verändern können.

5 Studie

5.1 Fragestellung

Der Arbeit liegt die Frage zugrunde, ob eine visuelle Antizipationshilfe durch das Hinzufügen von synchroner rhythmischer Beanspruchung von taktilen und auditiven Sinnen eine bessere zusätzliche Hilfestellung darbietet, ohne dabei den Animierenden Reizstimulationen auszusetzen, die für ihn ungenutzt oder störend sind. Des Weiteren soll festgestellt werden, ob eine Überforderung durch die Beanspruchung der drei Sinne vorliegt.

5.2 Methodik

5.2.1 Stichprobe

An der Studie nahmen insgesamt 20 Personen im Alter zwischen 20 und 32 Jahren teil. Das Durchschnittsalter betrug rund 26,00 Jahre ($SD = 4,27$). Von den 20 Teilnehmenden waren 65 % weiblich und 35 % männlich. 40 % hatten bereits Erfahrungen mit VR gesammelt, davon wiederum 75 % mehr als ein Mal. 40 % der Teilnehmenden arbeiteten in einem Beruf oder studierten in einem Studiengang, in der IT eine tragende Rolle spielt. 90 % der Teilnehmenden waren Akademiker, die entweder studieren oder bereits einen Abschluss der Universität erlangt haben.

Es wurde nicht unterschieden zwischen Links- oder Rechtshändern. Personen mit Behinderungen, starken Sehschwächen oder Beeinträchtigungen durch z.B. Rot-Grün-Schwäche wurden vorab von der Studie ausgeschlossen.

5.2.2 Design

Die Studie wurde so designt, dass die Teilnehmenden Aufgaben in verschiedenen Varianten absolvieren mussten - einer Variante mit rein visuellen Antizipationshilfen und einer Variante mit visuellen zuzüglich weiterer multisensorischer Antizipationshilfen. Innerhalb der Varianten mussten die Teilnehmenden Aufgaben unterschiedlicher Komplexität bewältigen. Dadurch wurde gewährleistet, dass die Studie einfache, übersichtliche Szenen, aber auch schwere, unübersichtliche Szenarien abdeckt.

Aufgabe der Teilnehmenden war es, jeweils eine einzelne Animation hinzuzufügen, die alle Teile der in dem Szenario vorgesehenen Bewegungen enthält. Die Inhalte des Szenarios sowie die vorgesehenen Bewegungen, welche die Aufnahme beinhalten soll, wurden dem Teilnehmenden vorweg erklärt.

Ursprünglich war es vorgesehen, dass die Teilnehmenden zur Vorbereitung auf ihre Aufnahme alle bereitgestellten Funktionen nutzen durften. Dabei sollten sie nach Belieben entscheiden, wann und wo sie Punkte für Antizipation oder Interaktionsdauer setzen. In einer Vorstudie, die aus zwei Teilnehmendenn bestand, kristallisierte sich jedoch deutlich heraus, dass die Steuerung des Prototypen erhöhte Aufmerksamkeit bedarf, sodass die Fokussierung auf die Antizipationshilfen nicht entsprechend Beachtung fand. Damit die Fragestellung wieder in den Fokus rücken konnte, wurde der Prototyp in seinen Funktionen für die Studie auf die Aufnahme von Animationen beschränkt.

Infolgedessen wurden die vorgegebenen Szenen mit Antizipations- und Interaktionspunkten bestückt und die Teilnehmenden vor jeder Szene instruiert, wie sie diese zu deuten haben.

Die zwei Szenarien wurden den Teilnehmenden beginnend mit dem einfachen Szenario unmittelbar hintereinander gestellt. Dies geschah zunächst in einer Variante und dann, nach dem Ausfüllen eines kurzen Fragebogens in der anderen Variante. Jeweils bei männlichen und weiblichen Teilnehmenden wechselte mit jeder Person die Variante, mit der gestartet wurde. Mit dieser Gleichverteilung innerhalb der Geschlechter sollte der Einfluss von eventuell vorliegenden Kohärenzen zwischen Geschlecht und Technik-Affinität minimiert werden.

Das erste und einfache Szenario beinhaltete nur eine Animation und einen Antizipationspunkt. Die Animation machte ein paar Schritte nach vorne, hob dabei den rechten Arm zu einem sogenannten „High Five“ und führte die dazugehörige Bewegung aus. Der Antizipationspunkt war örtlich genau auf die Position der Handfläche orientiert und zeitlich genau dann gesetzt, wenn die Hand der Animation wieder zurückgezogen wird (Abbildung 17). Aufgabe der Teilnehmenden war es nun, eine Animation aufzunehmen, die in die angebotene Hand einschlägt. Einzige Bedingung war, dass die Position und die zeitli-

che Abstimmung dabei beachtet werden.

Das zweite Szenario bestand aus drei Animationen und vier Antizipationspunkten, bei denen vier verschiedene Bewegungen ausgeführt werden mussten (Abbildung 18). Zunächst kam eine Animation auf den Teilnehmenden zu, mit der nur ein Gespräch anhand von Gestik animiert werden sollte. Anschließend musste sich der Teilnehmenden nach links zur zweiten Animation drehen, die in zwei Hände einschlagen soll, d.h. der Teilnehmenden musste seine beiden Hände mit der Handfläche nach oben an die von der Antizipationshilfe vorgegebene Position halten. Nach einem kurzen „Gespräch“ mit der zweiten Animation sollte der Teilnehmenden wiederum in die Hände des Charakters einschlagen, die sie nun mit den Handflächen nach oben hält. Nachfolgend musste sich der Teilnehmenden um 180° drehen, wo ihn dann ein „High Five“ des dritten Charakters erwartete. Kurze Zeit später musste diese Animation letztlich geschubst werden.

Die Komplexität der Szene mit ihren unterschiedlichen Bewegungen und ihrer langen Szenenlänge wurde bewusst gewählt, damit die Teilnehmenden ein grobes Verständnis für komplexe Animationserstellung erhalten und die Antizipationsvarianten auch unter diesem Gesichtspunkt bewerten können.

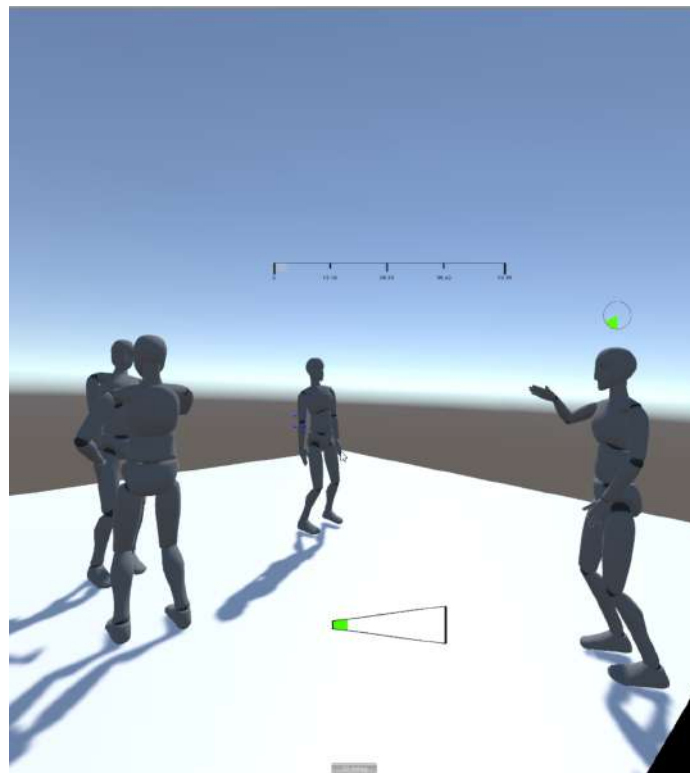


Abbildung 18: Zweites Szenario: Verschiedene Aktionen in unterschiedlichen Zeitabständen hintereinander

Nach jeder Variante wurde den Teilnehmenden der Fragebogen 1 (Anhang A & B) vorgelegt, um eine direkte Bewertung der Antizipationshilfen nach Durchführung des Szenarios abzufragen mit den drei Bewertungskategorien: Unterstützung, zeitkritische Vorbereitung und zeitkritische Reaktion. Nachdem beide Animationsvarianten und erneut der Fragebogen 1 abgeschlossen waren, sollten die Teilnehmenden den demografischen Fragebogen (Anhang C) und anschließend den Fragebogen 2 (Anhang D) ausfüllen. Der erste demografische Fragebogen erfragte über die demografischen Fragen hinaus auch die Technik-Affinität sowie die Häufigkeit der VR-Benutzung. Der zweite Fragebogen bestand aus elf Fragen, der feine Unterschiede zwischen den zwei Versuchsvarianten aufdecken sollte. Der abschließende NASA-Task Load Index (NASA, 2019, folgend nur noch NASA-TLX genannt) wurde gewählt, um zu überprüfen, ob es durch die zusätzliche multisensorische Stimulation zu einer Mehrbelastung oder eine Überforderung im Gegensatz zu der rein visuellen Variante vorliegt.

5.2.3 Versuchsaufbau

Der Aufbau wurde so gewählt, dass die Studie zeitlich möglichst effizient ablaufen konnte. Daher wurde die Einführungspräsentation und die späteren Fragebögen nicht an dem Gerät durchgeführt, an dem die Vive angeschlossen war. Dies hatte den Vorteil, dass die Teilnehmenden den Fragebogen 1 ausfüllen konnten, während in Unity die Versuchsvariante verändert werden konnte. Dafür war ein Neustarten des Prototypen erforderlich. Außerdem konnten die erstellten Lösungen der Teilnehmenden abgespeichert werden, ohne dass eine Wartezeit entstand. Die Studie wurde in einem separaten Raum abgehalten, in dem kein öffentlicher Zugang möglich war, um eine möglichst ungestörte Testung zu garantieren. Der Versuchsaufbau war für jeden Teilnehmenden identisch.

5.2.4 Durchführung

Der folgende Ablauf der Studie war für jeden Teilnehmenden gleich:

1. Einführung
2. Einverständniserklärung
3. Kurze freies „Vertraut machen“ mit der VR
4. Einführungsszenario
5. Studie
 - a) Beide Szenarien: Erst leicht, dann schwere Szene

- b) Fragebogen 1
 - c) Gleiche Szenarien in anderer Versuchsvariante
 - d) Fragebogen 1
 - e) Demografischer Fragebogen
 - f) Fragebogen 2
 - g) NASA-TLX
6. Verabschiedung

Zunächst wurden die Teilnehmenden begrüßt und mit einer Einführungspräsentation in die Studie eingeleitet. In dieser Präsentation wurde der Ablauf der Studie vorgestellt, auf Risiken hingewiesen und dass die Studie zu jeder Zeit abgebrochen werden kann. Des Weiteren wurde den Teilnehmenden eine grundlegende Einführung in den Prototypen gegeben. In ihr wurde das Prinzip der Bewegungsaufnahme und ihrer Folgen sowie aller visuellen Elemente von Zeitleiste bis hin zur Antizipationshilfe näher gebracht.

Im Anschluss mussten die Teilnehmenden zunächst die Einverständniserklärung unterschreiben. Danach durften sie das Vive-Headset aufsetzen und damit die VR betreten. Hier fand, nachdem der Prototyp gestartet wurde, zunächst ein kurzes Zurechtfinden statt. Die Teilnehmenden durften sich frei bewegen und Animationen aufnehmen. Sobald die Teilnehmenden sich bereit fühlten, spätestens jedoch nach fünf Minuten, wurde ein Einführungsszenario gestartet, welches kurz erklärt wurde. Dabei handelte es sich um eine ungefähr sechs Sekunden lange Aufnahme eines still stehenden Charakters, der nach drei Sekunden an der linken Schulter berührt werden sollte. Der genaue Zeitpunkt wurde dabei von der visuellen Antizipation unterstützt.

Nach dem Absolvieren dieser Einführungsaufgabe begann die eigentliche Studie nach dem oben angegebenen Ablauf. Erst wurden den Teilnehmenden die zwei Aufgaben gestellt, dann mussten sie die VR verlassen, um Fragebogen 1 zu beantworten. Im Anschluss, zurück in der VR, mussten die Teilnehmenden die Aufgaben in der anderen Variante lösen und danach wieder Fragebogen 1 beantworten. Letztlich sollten sie den demografischen Fragebogen, Fragebogen 2 und den NASA-TLX ausfüllen.

5.2.5 Auswertung der Daten

Alle 20 erhobenen Datensätze konnten vollständig in die Analyse eingeschlossen werden. Für jeden Teilnehmenden lagen vollständig ausgefüllte und valide Fragebögen vor. Die

Datenauswertung erfolgte mit Hilfe der Statistik-Software „IBM SPSS“ der Version 25 (IBM, 2019). Die grafische Aufbereitung der Datenauswertung (Abbildung 19) wurde mit dem Programm „Prism“ in der Version 5.03 (GraphPad, 2019) durchgeführt. Aufgrund des kleinen Datensatzes (N=20) wurde keine Normalverteilung aller Antworten angenommen.

Zur Überprüfung der Fragestellung, ob multisensorische gegenüber rein visuellen Antizipationshilfen in der VR einen Zugewinn bieten, wurde als einseitiger t-Test für abhängige Stichproben der Wilcoxon-Test durchgeführt. Hierfür wurde die ordinalskalierte Antwort-Skala der Fragen 1-3 aus Fragebogen 1 auf eine 7-stufige Likert-Skala von 1 (trifft gar nicht zu) bis 7 (trifft völlig zu) übertragen. Die drei Fragen zielen mit unterschiedlichen Aspekten auf die Güte der Hilfestellung durch die jeweilige Variante visuell oder multisensorisch ab, wobei ihnen eine gemeinsame Antwort-Skala zugrunde liegt. Daher wurden zur Analyse die Fragen als Summe zu einem Gesamtwert zusammengefasst.

Eine anschließende Post-hoc-Analyse mittels Wilcoxon-Test prüft den Einfluss der einzelnen Fragen auf den Unterschied der Gesamtwerte zwischen den Bedingungen.

Darüber hinaus wurden Mittelwerte mit Standardabweichung und die absoluten Häufigkeiten der einzelnen Antwortmöglichkeiten ermittelt.

Antwortmöglichkeiten aus Fragebogen 1 mit einem Wert größer 4 deuten eine positive Bewertung der jeweiligen Frage an.

Frei formulierte Antworten des Fragebogens 2 wurden zunächst in normalskalierte Antwortkategorien übertragen. Somit setzte dieser sich größtenteils aus Fragen mit dichotomen Antwortmöglichkeiten (Fragen 1, 2, 2a, 3, 3a und 6-9) zusammen. Frage 4 und 5 lag respektiv eine vier- und fünfteilige Antwortskala zugrunde. Alle Antworten wurden einer deskriptiven Beschreibung unterzogen.

Zum Ende der Versuchsdurchführung wurde der NASA-TLX als normierter Test zur empfundenen Belastung des Experiments durchgeführt. Der NASA-TLX enthält mehrere Subskalen, die durch unterschiedliche Gewichtung, in einen Wert zur Gesamtbeanspruchung eingehen. Die Teilnehmenden können in zwei Untergruppen aufgeteilt werden anhand der Reihenfolge, in welcher sie die beiden Versuchsbedingungen durchlaufen haben. Jeweils zehn Teilnehmenden starteten mit den rein visuellen Antizipationshilfen und wurden danach mit der multisensorischen Variante konfrontiert und vice versa. Die zuletzt durchgeführte Variante kann hierbei als präsenderer Einfluss auf die momentan empfundene Belastung angenommen werden. Der Gruppenvergleich der Gesamtbelastung wurde mittels nicht-parametrischem Mann-Whitney-U-Test für unabhängige Stichproben durchgeführt.

5.3 Ergebnisse

5.3.1 Fragebogen 1

Der Gesamtwert aus Frage 1-3 spiegelt das Gütekriterium der Hilfestellung der beiden Bedingungen visuell und multisensorisch wieder. Die Verteilung der Antwortmöglichkeiten ist durch die jeweiligen absoluten Häufigkeiten der Fragen 1-3 in Abbildung 19 1-3A dargestellt.

Für Frage 1 und 2 wählten bei der visuellen Variante 80 % der Teilnehmenden einen Wert größer 4, bei der multisensorischen Variante 95 %. Für Frage 3 wählten 55 % der Teilnehmenden bei der visuellen Methode einen Wert größer 4, 100 % wurden bei der multisensorischen Methode gewählt (Abbildung 19 1-3A).

Abbildung 19 stellt jeweils Mittelwert und Standardabweichung der Antworten aller drei Fragen (1-3B) sowie der Gesamtsumme (4) dar. Die Mittelwerte und Standardabweichungen können Tabelle 1 entnommen werden, wobei die Werte für die multisensorische Variante systematisch höher ausfallen.

Die Gesamtbewertung der Hilfestellung durch multisensorische Stimulation erweist sich als signifikant größer im Gegensatz zur visuellen Variante ($Z = -2,91$; $p < 0.01$). Die post hoc Analyse zeigt, dass dieser Unterschied durch signifikant höhere Bewertungen aller drei einzelnen Fragen zustande kommt ($p < 0,05$; vgl. Anhang E).

Variante	Frage 1		Frage 2		Frage 3		Summe	
	M	SD	M	SD	M	SD	M	SD
visuell	5,60	1,31	5,50	1,32	5,20	1,47	16,30	3,74
multisensorisch	6,30	0,98	6,35	0,81	6,10	0,85	18,75	2,15

Tabelle 1: Mittelwerte (M) und Standardabweichungen (SD) Antwortmöglichkeiten der Fragen 1-3 des 1. Fragebogens für die beiden Versuchsbedingungen visuell und multisensorisch

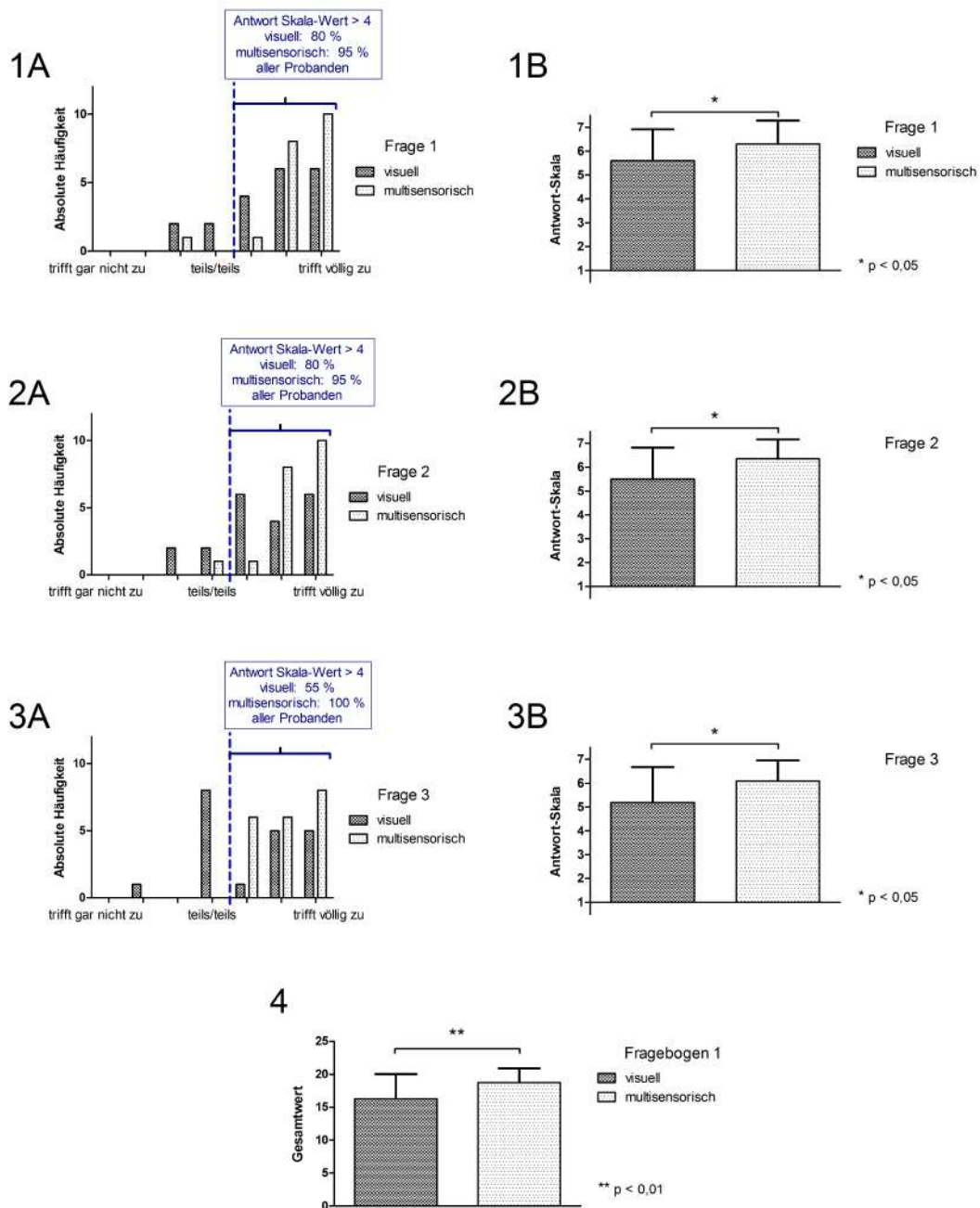


Abbildung 19: Darstellung der Antwortverteilungen für die drei Fragen des Fragebogen 1 (1-3A) sowie Mittelwerte mit Standardabweichung der von 1-7 skalierten Antwortmöglichkeiten und des Gesamtwertes über Fragebogen 1 mit Andeutung signifikanter Modalitätsunterschiede zwischen den Versuchsbedingungen rein visuell und multisensorisch (* p < 0,05 & ** p < 0,01; N=20).

5.3.2 Fragebogen 2

Die relativen Häufigkeiten der jeweiligen Antwortkategorien der Fragen aus Fragebogen 2 aller 20 Teilnehmenden können der Tabelle 2 entnommen werden.

Auf Frage 1 antworteten 95 % der Teilnehmenden, dass Ihnen die multisensorische Variante besser gefiel als die rein visuelle.

Alle Teilnehmenden haben die auditiven Signale wahrgenommen (Frage 2) und 95 % von ihnen fanden Sie hilfreich (Frage 2a).

In Frage 3 gaben 95 % der Teilnehmenden an, die taktilen Signale wahrgenommen zu haben und 65 % von ihnen empfingen sie als hilfreich (Frage 3a).

Die Frage, ob eines der beiden Signale wegen der anderen untergegangen bzw. nicht deutlich genug wahrgenommen werden konnten (Frage 4), wurde von 50 % der Teilnehmenden verneint. 40 % der Teilnehmenden sagten aus, dass das taktile Signal nicht deutlich genug war, 5 % der Teilnehmenden, dass das auditive Signal untergegangen ist und 5% der Teilnehmenden sagte aus, dass taktil und auditiv nicht deutlich genug wahrgenommen wurden. Abbildung 20 a stellt die Verteilung der Antwortmöglichkeiten für Frage 4 grafisch dar.

In Abbildung 20 b sind Antwortmöglichkeiten für Frage 5 grafisch skizziert:

20 % der Teilnehmenden sagten aus, dass sie sich auf die visuellen Signalart konzentriert haben, 50 % konzentrierten sich auf die auditiven Signale, 10 % der Teilnehmenden auf alle und 20 % der Teilnehmenden konzentrierten sich auf taktil und auditiv.

Ob die Kombination der Signalarten hilfreich war (Frage 6) bejahten 90 % der Teilnehmenden.

Alle Teilnehmenden empfanden die Abstimmung der Signale aufeinander gut (Frage 7).

70 % der Teilnehmenden verneinten die Frage, ob sie sich vorstellen könnten, die Aufgaben ohne visuelle Hilfen genauso gut zu bewältigen (Frage 8).

In Frage 9 gaben 100 % der Teilnehmenden an, dass die Aufgabe ohne jedwede Antizipationshilfen nicht genauso gut zu bewältigen gewesen wäre.

		Relative Antworthäufigkeit [in %]				
Frage	Antwort	Nein	Ja			
	1	5,00	95,00			
	2	0,00	100,00			
	2a	5,00	95,00			
	3	5,00	95,00			
	3a	35,00	65,00			
	6	10,00	90,00			
	7	0,00	100,00			
	8	70,00	30,00			
	9	100,00	0,00			
Frage	Antwort	Nein	Taktil	Auditiv	Beide	
	4	50,00	40,00	5,00	5,00	
Frage	Antwort	Visuell	Taktil	Auditiv	Alle	Visuell & auditiv
	5	20,00	0,00	50,00	10,00	20,00

Tabelle 2: Relative Antworthäufigkeiten [in %] der jeweiligen Antwortkategorien der Fragen aus Fragebogen 2 aller Teilnehmenden (N=20).

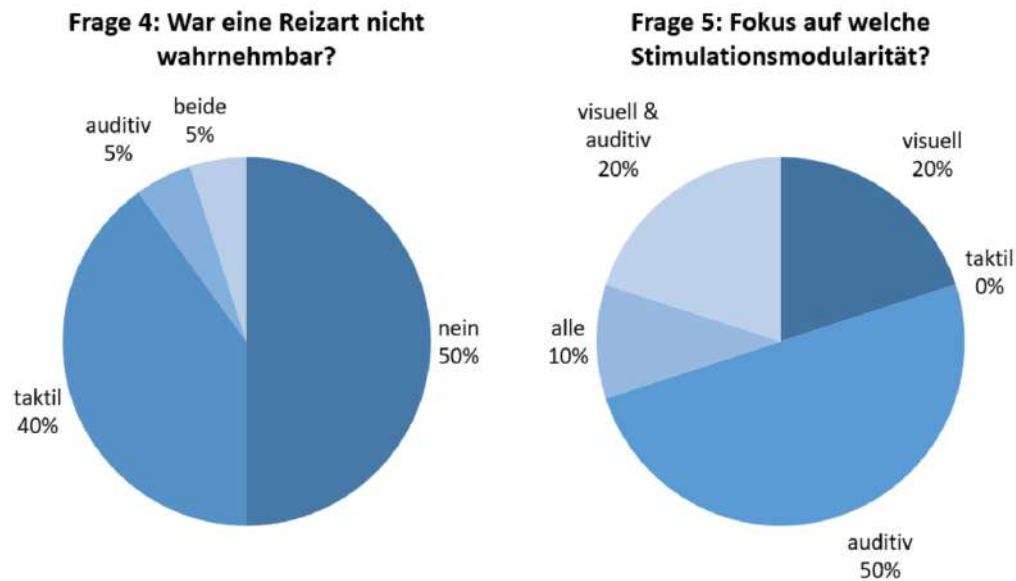


Abbildung 20: Darstellung der relativen Antworthäufigkeiten [in %] der jeweiligen Antwortkategorien der Fragen 4 (A) und 5 (B) aus Fragebogen 2 für alle Teilnehmenden (N=20).

5.3.3 NASA-TLX

Für Teilnehmenden, die mit der visuelle Variante abgeschlossen hatten (N = 10), betrug der durchschnittliche Gesamtwert der empfundenen Belastung im NASA-TLX 40,07 (SD = 16,27) und bei den Teilnehmenden, die die multisensorischen Variante zuletzt absolvierten (N = 10), 42,41 (SD = 20,42; vgl. Abbildung 21).

Es konnte kein Unterschied in der empfundenen Belastung basierend auf der Reihenfolge der Variantendurchführung aufgedeckt werden ($Z = -0,27$; $p = 0,79$; vgl. Abbildung 21 sowie Anhang E).

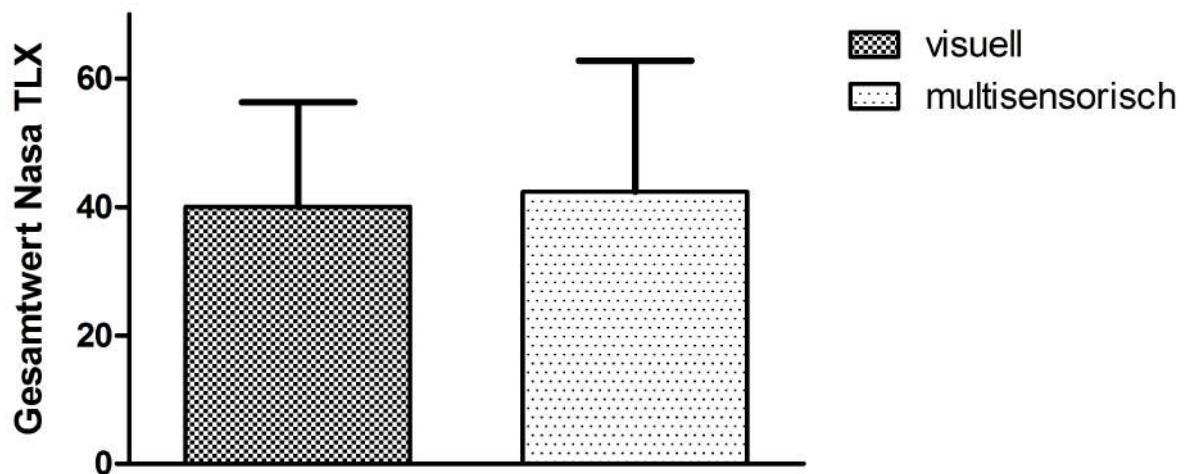


Abbildung 21: Darstellung von Mittelwerten mit Standardabweichungen der Gesamtwerte des NASA-TLX Belastungstests für die Versuchsmodalitäten rein visuell und multisensorisch (N=20).

5.4 Diskussion

5.4.1 Diskussion der Studienergebnisse

Der erste Fragebogen erfasste jeweils nach durchgeführter Versuchsvariante, wie die Antizipationshilfe für die hier gestellten Aufgabenstellungen empfunden wurde. 80 - 95 % der Antworten wurden auf den jeweiligen Likert-Skalen mit einem Wert von über 4 bewertet. Werte größer 4 deuten eine positive Bewertung an. Dies legt für Fragebogen 1 den Schluss nahe, dass die Antizipationshilfen in beiden Varianten für die Mehrzahl der Teilnehmenden generell eine nützliche Hilfestellung waren. Sowohl die visuellen als auch multisensorischen Reize wurden laut eigenen Angaben der Teilnehmenden wahrgenommen, überforderten im Szenario nicht und halfen, den zeitkritischen Moment zu animieren.

Die Gegenüberstellung der beiden Varianten stützt die Hypothese, dass die multisensorischen Antizipationshilfen einen Mehrwert darstellen. Dies geht aus dem Unterschied der Gesamtbewertung der Güte der Hilfen hervor und stützt sich auf einen signifikanten Unterschied in allen drei Bewertungskriterien: Unterstützung, zeitkritische Vorbereitung und zeitkritische Reaktion.

Fragebogen 2 geht auf den subjektiven Vergleich beider Versuchsvarianten im Anschluss an die Durchführung des gesamten Versuchsablaufes ein. Entscheidend ist, dass für alle

Teilnehmenden laut persönlicher Einschätzung die Aufgabe ohne jedwede Antizipationshilfen nicht genauso gut zu bewältigen gewesen wäre. Alle Teilnehmenden stützten sich folglich bei der Aufgabenbewältigung auf die dargebotenen Antizipationshilfen. Dabei bewerteten die meisten Teilnehmenden (95 %) die multisensorische Variante als besser.

In der multisensorischen Version wurden die auditiven Reize ausnahmslos von allen Teilnehmenden wahrgenommen, wobei fast alle Teilnehmenden diese dabei auch als hilfreich empfanden. Trotz vielfältiger Stimulationsreize ging der auditive Reiz für fast alle Teilnehmenden nicht unter in der Wahrnehmung und viele Teilnehmenden fokussierten gezielt darauf, um diesen Reiz zur Antizipationshilfe zu nutzen. Die gewählte Reizart und ihre Umsetzung zeigten sich für die in dieser Arbeit zu bewältigenden Aufgabenstellungen als deutliche und teils auch dominante Stimulation, die intuitiv als gute Hilfestellung genutzt werden konnte.

Die taktilen Reize in der multisensorischen Version wurden ebenfalls von fast allen Teilnehmenden wahrgenommen. Dabei gaben 65 % der Teilnehmenden an diese als nützlich empfunden zu haben. Keiner der Teilnehmenden stützte sich zur Aufgabenbewältigung allein auf diese Reizkategorie. Dies mag vor allem an der zu schwachen Intensität des taktilen Reizes gelegen haben, so dass in der multisensorischen Reizumgebung die anderen Reizarten als dominantere und verlässlichere Hilfestellung empfunden wurden. In der hier durchgeführten Studie mag daher die Vibrationsstärke nicht ausreichend deutlich genug ausgefallen sein. Taktile Reize mögen jedoch auch nicht die intuitiv erste Wahl zur Antizipation eines Ereignisses bei einer multimodalen Hilfestellung sein, da die taktile Stimulation durchaus von den meisten Teilnehmenden wahrgenommen wurde.

Die Ergebnisse verdeutlichen darüber hinaus, dass einige Teilnehmenden (20 %) sowohl auditive als auch taktile Stimulation in der Kombination zur Antizipation genutzt haben. Die jeweils synchrone und gleichförmige Taktung beider Stimulationen bestärkte vermutlich die Gesamtwahrnehmung der zum Ereigniszeitpunkt hin zunehmenden Stimulationsfrequenz. Dabei konnten sich jedoch 70 % der Teilnehmenden nicht vorstellen, die Aufgaben ohne visuelle nur auf Basis der auditiven und taktilen Hilfen genauso gut zu bewältigen. Einige Teilnehmenden (10 %) gaben an, dass sie sich gezielt zusätzlich auch auf die visuelle Hilfe stützten und nutzten somit bewusst das ganze Spektrum an dargebotenen Stimulationsreizen. Eine geeignete Kombination der Stimulationsmodalitäten vermag sich also zu Ergänzen und kann bewusst oder unbewusst als gemeinsame Hilfestellung genutzt werden ohne dabei den Teilnehmenden zu überfordern. Dies wird durch die grundsätzliche Aussage der Teilnehmenden bestärkt, dass für 90 % die Kombination der Signalarten hilfreich war und ausnahmslos alle die Abstimmung der Signale aufeinander als gut empfanden.

Die empfundene Belastung der Teilnehmenden direkt nach Durchführung der Szenarien mit visuellen Antizipationshilfen unterscheidet sich nicht von der Belastung nach der multisensorischen Variante. Dies deutet darauf hin, dass die multisensorische Variante mit zusätzlicher Stimulation keine Mehrbelastung und somit auch keine Überforderung für die Teilnehmenden darstellt.

5.4.2 Limitationen

Der Stichprobenumfang ist mit 20 Teilnehmenden vergleichsweise gering, sodass die Aussagekraft der Ergebnisse recht klein ausfällt. Aus der kleinen Stichprobe resultiert daher, dass kleine Unterschiede innerhalb des Variantenvergleichs unerkant bleiben könnten. Die Ergebnisse sind zudem nicht auf die Normalbevölkerung übertragbar, da 90 % der Teilnehmenden eine akademische Laufbahn absolvierten oder absolviert haben. Hinzu kommt, dass bereits 40 % der Teilnehmenden einen IT-intensiven Beruf ausübten oder ein IT-intensives Studium durchliefen.

Der NASA-TLX wurde einmalig am Ende der Studie durchgeführt, was in einer Aufspaltung in zwei Vergleichsgruppen resultierte. Dadurch verkleinerte sich die bereits geringe Vergleichsmenge von 20 je 10 Teilnehmende. Durch den Gruppenvergleich ist die Aussagekraft folglich weiter eingeschränkt. Zudem kann durch die einmalige Durchführung des Fragebogens am Ende der Studie kein interpersoneller - also ein Vergleich pro Person nach Durchführung der jeweiligen Versuchsvariante - sondern nur ein intrapersoneller Vergleich eines generellen Belastungsbefinden nach Durchführung der gesamten Studie erreicht werden.

Die Skalen der Fragen 1-3 des Fragebogens 1 waren mit sieben wählbaren Antwortmöglichkeiten stark beschränkt. So entstand ein Deckeleffekt der Antworten zum Rand der Skala. Teilnehmende, die bereits den höchsten Wert für eine Frage nach der ersten Versuchsvariante wählten, konnten bei der gleichen Frage für die zweite Versuchsvariante keinen höheren Wert mehr wählen. Ein Nachschauen oder Ändern der Ergebnisse aus Fragebogen 1 nach der ersten Versuchsvariante wurde nicht gestattet, um keine Antwortverzerrungen im Nachhinein zu riskieren.

Die Ergebnisse der Fragen des Fragebogens 2 konnte aufgrund der gewählten Fragestellungen nur deskriptiv erfolgen. Der manuellen Übertragung der textbasierten Antworten

in normalskalierte Antwortkategorien liegt ein Interpretationsspielraum zugrunde.

Der Vergleich in dieser Studie ist auf eine multisensorische Modalität beschränkt. Die passende Kombination der Sinnesmodalität für zahlreiche persönliche Vorliebemöglichkeiten zu bestimmen, gestaltet sich als schwierig. Ein Vergleich der verschiedenen Kombinationsmöglichkeiten blieb im Rahmen dieser Arbeit aus.

5.4.3 Ausblick

Zur Erforschung der Hilfestellung durch Antizipationshilfen in der VR-Umgebung erscheint eine Abdeckung aller Kombinationsmöglichkeiten von für den VR-Einsatz möglichen Wahrnehmungssinnen sinnvoll, um bestmögliche Kombinationen auf Szenarien und persönliche Präferenzen abzustimmen. So kann ein interpersoneller Vergleich aller Kombinationen durchgeführt werden und folglich zu jedem Teilnehmenden die bestmögliche Antizipationshilfe gefunden werden. Als nächster Schritt scheint dann eine Aufsplittung in zwei Gruppen plausibel: VR-Anfänger und VR-Experten. VR-Anfänger, die wenig bis keine Erfahrung mit VR besitzen, sind allein durch die Nutzung des VR-Systems stärker belastet, so dass dort geringere Reizflutung sinnvoller erscheint als bei VR-Experten. Daher würde die Möglichkeit, für VR-Experten den vollen Umfang des Prototyps, in dem sie u.a. Antizipationspunkte selbst bearbeiten können, freizuschalten, ein geringeres Risiko der Überforderung darstellen. Bei VR-Anfängern hingegen scheint eine starke Überforderung sehr wahrscheinlich. In der Vorstudie wurde dies bereits schnell ersichtlich.

Ein weiterer Ansatz durch die Kombination aller Sinnesmodalitäten könnte sein, zu unterschiedlichen Anwendungsszenarien jeweils die beste Kombination zu ermitteln, um so eine Gruppierung von Kombinationen zu Anwendungsszenarien zu erzeugen. Diese Gruppierung könnte wiederum genutzt werden, um einzelne Kombinationen innerhalb der Anwendungsszenarien weiter zu erforschen, in dem verschiedene Ausprägungen und Varianten dieser Kombination einander gegenübergestellt werden.

Ein weitere Betrachtung dieses Themengebietes erscheint sinnvoll.

Die Studie lies den Teilnehmenden die Möglichkeit, sich jedes Szenario unbegrenzt wiederholt anzuschauen. Dadurch hatte die Studie einen zu großen unterschiedlichen Lerneffekt auf die Teilnehmenden.

Für die Auswertung der Studie könnte zunächst der Deckeleffekt des Fragebogens 1 (Anhang A) entfernt werden. Dies könnte erreicht werden, in dem den Teilnehmenden erlaubt wird, ihren Fragebogen 1, den sie nach der ersten Variante ausgefüllt haben, am Ende

nachträglich zu ändern. Eine weitere Möglichkeit, den Deckeleffekt zu entfernen, stellt eine direkte Vergleichsfrage nach Abschluss der zweiten Variante dar. In jedem Fall wäre eine feinere Granularität mit einem größeren Skalenintervall angebracht.

Fragebogen 2 sollte für zukünftige Arbeiten eine normalskalierte Antwortmöglichkeit bieten, so dass eine bessere und genauere Auswertung stattfinden kann.

Der NASA-TLX sollte für beide Versuchsvarianten beantwortet werden, um einerseits einen interpersonellen Vergleich zu erreichen und um andererseits die Verringerung und Aufspaltung der Vergleichsmenge zu verhindern.

Für zukünftige Verwendung des Prototyps sollte zunächst die Feinjustierung der Funktion des automatischen Laufens des Plugins Final IK stattfinden, da dieser nicht authentisch war. Im Zuge dessen, sollte der Algorithmus der Kalibrierung der Körpergröße verbessert und in die Feinjustierung mit berücksichtigt werden, da diese bei mehrmaligen Ausführen Skalierungsprobleme verursachte.

Die Zeitleiste, welche die Szenendauer angibt, könnte so erweitert werden, dass ein Vor- und Zurückspulen auf ihr möglich ist. Die Option, die Zeitleiste an einen gewünschten beliebigen Ort zu verschieben, wäre ebenfalls eine sinnvolle Ergänzung.

Der positionsanzeigende quadratische Marker von Antizipationspunkten sollte von einer 2D- in eine 3D-Grafik umgewandelt werden, um auch die Sicht in die Tiefe zu gewährleisten. Eine Verschiebung des Blickwinkels auf den Marker resultierte in einer deutlichen Ungenauigkeit, da beim Setzen Punktes die z-Achse nicht wahrgenommen werden konnte.

Der Durationpoint für die Dauer einer Interaktion mit einem Charakter war zu devot, so dass er von den Studienteilnehmern wenig beachtet wurde.

Die Position der Progressbar sollte verändert werden, da einige Teilnehmende sie für zu tief positioniert empfanden. Eine Verschiebung in Richtung Fokusbereich erscheint ebenfalls sinnvoll. Die beste mögliche Lösung wäre eine vom Nutzenden anpassbare Version. Des Weiteren wäre eine dynamische Anpassung der Länge der Progressbar anhand der Entfernung zum nächsten Antizipationspunkt empfehlenswert, um eine noch deutlichere zeitliche Einschätzung zu gewährleisten.

Der Start der Wiederholung einer Szene wurde von den Teilnehmenden zu wenig wahrgenommen. Dort wäre es möglich, beim Zeitpunkt des Neustarts den Bildschirm aufleuchten zu lassen, so dass deutlich wahrgenommen werden kann, dass die Szene nun wieder neu beginnt.

Weiterhin sollte das Löschen von bereits vorhandenen Aufnahmen gewährleistet werden. Ohne diese Funktion ist der Bearbeitungsaufwand zu groß.

Abschließend könnten weitere Funktionen ergänzt werden, wie z.B. ein Springen zum

Start oder Ende der Szene oder um eine bestimmte Anzahl von Sekunden nach vorne oder hinten.

Dieser Prototyp besitzt viele ausbaufähige Funktionen, die in ihrer grundlegenden Funktionsweise auch auf andere Programme übertragbar sind.

6 Fazit

Es konnte in dieser Arbeit festgestellt werden, dass die gewählten Antizipationshilfen, unabhängig ihrer Versuchsvariante, bereits generell einen Mehrwert im Bezug auf die Bewältigung der Aufgabenstellung, eine zeitkritische Situation per Performance Motion zu animieren, boten. Hervorzuheben ist jedoch, dass die multisensorische Kombination von Stimulationsreizen hierbei tatsächlich genutzt und als hilfreich empfunden wurde. Es wurde nachgewiesen, dass durch die multisensorischen Antizipationshilfen keine Mehrbelastung oder Überforderung entstand, sondern, dass diese vielmehr einen signifikanten Mehrwert bieten. Die in dieser Arbeit gewählten Antizipationshilfen waren dabei gut auf die präsentierten Szenarien abgestimmt. Es kann somit zusammengefasst werden, dass sich multisensorische Antizipationshilfen eignen, um in ein bestehendes Animationsszenario zusätzliche Charaktere mit zeitkritischen Handlungsereignissen einzufügen.

Anzumerken ist, dass die Ergebnisse aufgrund der Beschränkung auf einen recht kleinen, gesellschaftlich spezifischen Stichprobenumfang auf einen größeren und breiteren Personenkreis ausgeweitet werden sollten, um die Allgemeingültigkeit der Ergebnisse zu gewährleisten.

Nichtsdestotrotz ist eine Tendenz erkennbar, die aufzeigt, dass Antizipationshilfen im Bereich der VR und der Performance Animation eine sehr große Hilfe darstellen können. Weiterführende Forschung sollte daher an diese Studie anknüpfen und die Kombinationsmöglichkeiten verschiedener sensorischer Stimulationsreize zur Antizipation in unterschiedlichen VR-Szenarien genauer untersuchen. Gerade die geeignete Kombination mehrerer Sinnesreize für Antizipationshilfen vermag einen entscheidenden Beitrag zu einer Steigerung der effizienten Animation eines komplexen Szenarios leisten.

7 Quellenverzeichnis

- ArxAnima (2019). Arx anima animation studio GmbH. <http://www.arxanima.com/>.
Aufgerufen am: 29.05.2019.
- Aurand, A. M., Dufour, J. S., and Marras, W. S. (2017). Accuracy map of an optical motion capture system with 42 or 21 cameras in a large measurement volume. *Journal of Biomechanics*, 58:237 – 240.
- Berger, K., Ruhl, K., Schroeder, Y., Bruemmer, C., Scholz, A., and Magnor, M. (2011). Markerless motion capture using multiple color-depth sensors. pages 317–324.
- Biswas, K. and Basu, S. (2011). Gesture recognition using microsoft kinect®. pages 100–103.
- Brigante, C., Abbate, N., Basile, A., Faulisi, A., and Sessa, S. (2011). Towards miniaturization of a mems-based wearable motion capture system. *Industrial Electronics, IEEE Transactions on*, 58:3234 – 3241.
- Bruder, G., Steinicke, F., Rothaus, K., and Hinrichs, K. (2009). Enhancing presence in head-mounted display environments by visual body feedback using head-mounted cameras. pages 43–50.
- Cardenas, I. S., Letdara, C. N., Selle, B., and Kim, J.-H. (2017). Immersify: Next generation of immersive pilot training. *2017 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 1203–1206.
- Chan, J., Leung, H., Tang, J., and Komura, T. (2011). A virtual reality dance training system using motion capture technology. *Learning Technologies, IEEE Transactions on*, 4:187 – 195.
- CloudheadGames (2016). Cloudhead Games: virtual reality performance capture with htc vive and perception neuron. <http://cloudheadgames.com/storyinvr/>. Aufgerufen am: 29.05.2019.
- Conway, C. and Christiansen, M. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of experimental psychology. Learning, memory, and cognition*, 31:24–39.
- Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: The design and implementation of the cave. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '93*, pages 135–142, New York, NY, USA. ACM.

- Deering, M. F. (1995). Holosketch: A virtual reality sketching/animation tool. *ACM Trans. Comput.-Hum. Interact.*, 2(3):220–238.
- Deutscher, J., Blake, A., and Reid, I. (2000). Articulated body motion capture by annealed particle filtering. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, volume 2, pages 126–133 vol.2.
- Doer, K.-U., Schiefel, J., and Kubbat, W. (2001). Virtual cockpit simulation for pilot training. page 8.
- evltube (2007). Cave® - a virtual reality theater - 1992. <https://www.youtube.com/watch?v=aKL0urEdtPU>. Aufgerufen am: 29.05.2019.
- Facebook (2019a). Oculus rift - gesundheits- und sicherheitshinweise. <https://securecdn.oculus.com/sr/gearvr-warning-german>. Aufgerufen am: 29.05.2019.
- Facebook (2019b). Quill. https://www.oculus.com/experiences/rift/1118609381580656/?locale=de_DE. Aufgerufen am: 29.05.2019.
- Fisher, S., Wenzel, E., Coler, C., and W. McGreevy, M. (1988). Virtual interface environment workstations. *Proceedings of the Human Factors Society Annual Meeting*, 32.
- Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. *Psychological bulletin*, 81:284–310.
- Google (2019). Glass. <https://www.google.com/glass/start/>. Aufgerufen am: 29.05.2019.
- Grabowski, A. and Jankowski, J. (2015). Virtual reality-based pilot training for underground coal miners. *Safety Science*, 72:310–314.
- GraphPad (2019). Prism. <https://www.graphpad.com/scientific-software/prism/>. Aufgerufen am: 29.09.2019.
- Gu, X., Zhang, Y., Sun, W., Bian, Y., Zhou, D., and Kristensson, P. (2016). Dexmo: An inexpensive and lightweight mechanical exoskeleton for motion capture and force feedback in vr. pages 1991–1995.
- GuitarHero (2019a). Guitar Hero. https://de.wikipedia.org/wiki/Guitar_Hero. Aufgerufen am: 29.05.2019.
- GuitarHero (2019b). Guitar hero 5. <http://www.gamrate.net/img/games/237/big4.jpg>. Aufgerufen am: 22.05.2019.

- Han, J., Shao, L., Xu, D., and Shotton, J. (2013). Enhanced computer vision with microsoft kinect sensor: A review. *IEEE transactions on cybernetics*, 43.
- HAWK, H. f. a. W. u. K. (2019). Funktionsweise htc - vive. <http://wikis.hawk-hhg.de/wikis/fields/cadsoftware/uploads/Main/OptischesTracking.jpg>. Aufgerufen am: 29.05.2019.
- Hee Cho, G., Hwangbo, G., and Soo Shin, H. (2014). The effects of virtual reality-based balance training on balance of the elderly. *Journal of physical therapy science*, 26:615–7.
- HTC (2019a). Htc vive - health and safety guide. https://brain-images-ssl.cdn.dixons.com/6/5/10144056/u_10144056.jpg. Aufgerufen am: 29.05.2019.
- HTC (2019b). Htc vive - health and safety guide. http://d14.htc.com/vive/safty_guide/91H02887-08M%20Rev.A.PDF. Aufgerufen am: 29.05.2019.
- IBM (2019). Spss statistics software. <https://www.ibm.com/de-de/analytics/spss-statistics-software>. Aufgerufen am: 29.09.2019.
- IKinema (2019a). Mögliche hardwarekonfigurationen der software orion. https://ikinema.com/docs/UserFiles/tiny_files/Orion_Leaflet_Digital_ENG.png. Aufgerufen am: 22.09.2019.
- IKinema (2019b). Orion. <https://www.youtube.com/watch?v=Khoer5DpQkE&t=33s>. Aufgerufen am: 29.05.2019.
- Jerald, J. (2016). *The VR Book: Human-Centered Design for Virtual Reality*. Association for Computing Machinery and Morgan & Claypool, New York, NY, USA.
- Kubovy, M. (1988). Should we resist the seductiveness of the space: Time:: Vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, 14:318–320.
- Lasseter, J. (1987). Principles of traditional animation applied to 3d computer animation. *SIGGRAPH Computer Graphics*, 21(4):S. 35–44.
- Laver, K., Lange, B., George, S., Deutsch, J., Saposnik, G., and Crotty, M. (2017). Virtual reality for stroke rehabilitation. *Cochrane Database of Systematic Reviews*, (11).
- Madary, M. and Metzinger, T. K. (2016). Real virtuality: A code of ethical conduct. recommendations for good scientific practice and the consumers of vr-technology. *Frontiers in Robotics and AI*, 3:3.
- Mahar, D., Mackenzie, B., and McNicol, D. (1994). Modality-specific differences in the

- processing of spatially, temporally, and spatiotemporally distributed information. *Perception*, 23(11):1369–1386. PMID: 7761246.
- Maples-Keller, J., Bunnell, B., Kim, S.-J., and Rothbaum, B. (2017). The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard Review of Psychiatry*, 25:103–113.
- McMahan, R. P., Bowman, D. A., Zielinski, D. J., and Brady, R. B. (2012). Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):626–633.
- Microsoft (2019a). Hololens 2. <https://www.microsoft.com/de-de/hololens>. Aufgerufen am: 29.05.2019.
- Microsoft (2019b). Kinect. <https://developer.microsoft.com/de-de/windows/kinect>. Aufgerufen am: 29.05.2019.
- Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Trans. Information Systems*, vol. E77-D, no. 12:1321–1329.
- Moeslund, T. B., Hilton, A., and Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.*, 104(2):90–126.
- MotionWorkshop (2019). Motion Workshop. <https://www.motionshadow.com/>. Aufgerufen am: 29.05.2019.
- Munafo, J., Diedrick, M., and Stoffregen, T. A. (2016). The virtual reality head-mounted display oculus rift induces motion sickness and is sexist in its effects. *Experimental Brain Research*, 235:889–901.
- NASA (2019). Task load index. https://www.keithv.com/software/nasatlx/nasatlx_german.html. Aufgerufen am: 29.09.2019.
- Noitom (2019). Percepton neuron. https://neuronmocap.com/products/perception_neuron. Aufgerufen am: 29.08.2019.
- NVRMIND (2019). Animvr. <https://nvrmind.io/>. Aufgerufen am: 29.05.2019.
- Okun, J., Zwerman, S., and Society, V. E. (2010). *The VES Handbook of Visual Effects: Industry Standard VFX Practices and Procedures*. Media Technology. Focal Press/Elsevier.
- OptiTrack (2019). OptiTrack. <https://optitrack.com/>. Aufgerufen am: 29.05.2019.

- Parsons, T., Gaggioli, A., and Riva, G. (2017). Virtual reality for research in social neuroscience. *Brain Sciences*, 7:1–21.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, 17(4):398–422.
- Petrescu, F. I. and Petrescu, R. (2016). Direct and inverse kinematics to the anthropomorphic robots. *ENGEVISTA*, 18:109–124.
- PokemonGo (2019). Pokemon go: Bild eines smartphones mit pokemon go. https://www.digitalartsonline.co.uk/cmsdata/features/3671619/ar_pokemon_go_istock_georgeclerk-niantic.jpg. Aufgerufen am: 29.09.2019.
- PolePosition (1982). Pole position: Start eines Rennens. <https://i.ytimg.com/vi/Rad4Dztj6qk/maxresdefault.jpg>. Aufgerufen am: 22.05.2019.
- RockBand (2019). Rock Band. https://de.wikipedia.org/wiki/Rock_Band. Aufgerufen am: 29.05.2019.
- Rokoko (2018). Siggraph 2018. <https://www.youtube.com/watch?v=cV5qvU7ILKU>. Aufgerufen am: 29.08.2019.
- Rokoko (2019). Smartsuit pro. <https://www.rokoko.com/en/products/smartsuit-pro>. Aufgerufen am: 29.08.2019.
- Rubinstein, L. and Gruenberg, E. M. (1971). Intramodal and crossmodal sensory transfer of visual and auditory temporal patterns. *Perception & Psychophysics*, 9(5):385–390.
- Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, 47(1):172 – 196.
- Seow, S. C. (2008). *Designing and Engineering Time: The Psychology of Time Perception in Software*. Addison-Wesley Professional, 1 edition.
- Seth, A., Vance, J., and Oliver, J. (2011a). Virtual reality for assembly methods prototyping: A review. *Virtual Reality*, 15:5–20.
- Seth, A., Vance, J. M., and Oliver, J. H. (2011b). Virtual reality for assembly methods prototyping: a review. *Virtual Reality*, 15(1):5–20.
- Seymour, N., Gallagher, A., A. Roman, S., K. O'Brien, M., K. Bansal, V., Andersen, D., and M. Satava, R. (2002). Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Annals of Surgery*, 236:458–464.

- Shams, L. and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11):411 – 417.
- Sherman, W. and Craig, A. (2018). *Understanding Virtual Reality—Interface, Application, and Design*.
- Sherrick, C. and Cholewiak, R. (1986). Handbook of perception and human performance. *Chapter Cutaneous Sensitivity*, pages 1–58.
- Steiger, N. (2019). Planet der affen: Die schauspieler hinter den affen. <http://www.blogbusters.ch/sidekicks/planet-der-affen>. Aufgerufen am: 29.05.2019.
- Sutherland, I. E. (1968). A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, AFIPS '68 (Fall, part I), pages 757–764, New York, NY, USA. ACM.
- Ungar, F. (2018). Performance animation in vr.
- Unigine (2019). Unigine: Beispielhafter aufbau eines cave. https://unigine.com/images/products/Engine/Output-Everywhere/07_output.jpg. Aufgerufen am: 29.09.2019.
- Walther-Franks, B., Biermann, F., Steenbergen, N., and Malaka, R. (2012). The Animation Loop Station: Near Real-time Animation Production. In *International Conference on Entertainment Computing (ICEC)*, pages 469–472, Berlin, Heidelberg. Springer.
- Walther-Franks, B., Herrlich, M., and Malaka, R. (2011). A Multi-Touch System for 3D Modelling and Animation. In *Proceedings of the International Symposium on Smart Graphics*, pages 48–59, Berlin, Heidelberg. Springer.
- Wang, R., Paris, S., and Popović, J. (2011). Practical color-based motion capture. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, pages 139–146, New York, NY, USA. ACM.
- Wikipedia (2019). Direkte und inverse kinematik. <https://upload.wikimedia.org/wikipedia/de/a/ae/Roboterkinematik.png>. Aufgerufen am: 22.09.2019.
- Witt, S. (2016). Evaluation von antizipationshilfen im kontext von echtzeitsteuerung für computerspiele und performance animation.
- Wong, S. T. (2014). Humancr-edit: Prototyping digital tools for dance creation. Technical report, Universität Bremen.
- WorldWiz (2019). Virtual reality projection system. <https://www.worldviz.com/virtual-reality-projection-system>. Aufgerufen am: 29.05.2019.

- Xsens (2019a). History of motion capture. <https://www.xsens.com/a-history-of-motion-capture>. Aufgerufen am: 29.08.2019.
- Xsens (2019b). Mvn animate. <https://www.xsens.com/products/mvn-animate>. Aufgerufen am: 29.08.2019.
- Yabukami, S., Kikuchi, H., Yamaguchi, M., Arai, K., Takahashi, K., Itagaki, A., and Wako, N. (2000). Motion capture system of magnetic markers using three-axial magnetic field sensor. *Magnetics, IEEE Transactions on*, 36:3646 – 3648.

Anhang A – Fragebogen 1

Variante 1

1. Die Antizipationshilfen haben mich sehr gut unterstützt, meine Aufgaben zu erfüllen.

trifft gar nicht zu

teils / teils

trifft völlig zu

2. Zeitkritische Ereignisse ließen sich durch die Hilfen sehr gut antizipieren.

trifft gar nicht zu

teils / teils

trifft völlig zu

3. Durch die Antizipationshilfen konnte ich genau zum richtigen Zeitpunkt reagieren

trifft gar nicht zu

teils / teils

trifft völlig zu

Anhang B – Fragebogen 1

Variante 2

1. Die Antizipationshilfen haben mich sehr gut unterstützt, meine Aufgaben zu erfüllen.

trifft gar nicht zu

teils / teils

trifft sehr gut zu

2. Zeitkritische Ereignisse ließen sich durch die Hilfen sehr gut antizipieren.

trifft gar nicht zu

teils / teils

trifft sehr gut zu

3. Durch die Antizipationshilfen konnte ich genau zum richtigen Zeitpunkt reagieren

trifft gar nicht zu

teils / teils

trifft sehr gut zu

Anhang C – Fragebogen 1

Demografischer Fragebogen

Alter:

Geschlecht:

Beruf / Studiengang:

Bereits VR benutzt:

Wenn ja, wie oft (1x, wenige Male, schon öfters, benutze ich regelmäßig):

Anhang D – Fragebogen 1

Multisensorisch vs. Visuell

1. Welche Variante fanden sie besser? Multisensorisch oder rein visuell?
 - a. Warum?

2. Haben Sie die auditiven Signale wahrgenommen?
 - a. Waren die auditiven Signale hilfreich?

3. Haben sie die taktilen Signale wahrgenommen?
 - a. Waren die taktilen Signale hilfreich?

4. Ist eines der beiden Signale wegen der anderen untergegangen bzw. konnten Sie es daher nicht deutlich genug wahrnehmen?

5. Auf welche Signalart haben Sie sich konzentriert, um die Aufgaben zu bewältigen? (auch visuell)

6. War die Kombination der Signalarten ihrer Meinung nach hilfreich? (auch visuell)

7. Finden Sie, dass die Signalarten gut aufeinander abgestimmt waren?

8. Könnten Sie sich vorstellen, dass Sie die Aufgabe ohne visuelle Hilfen genauso gut bewältigt hätten?

9. Könnten Sie sich vorstellen, dass Sie die Aufgabe ohne jedwede Antizipationshilfen genauso gut bewältigt hätten?

Anhang E – SPSS-Auswertung

Wilcoxon-Test

Statistik für Test^a

	1 multisensoris ch - 1 visuell	2 multisensoris ch - 2 visuell	3 multisensoris ch - 3 visuell	Summe MUL - Summe VIS
Z	-2,581 ^b	-2,799 ^b	-2,622 ^b	-2,907 ^b
Asymptotische Signifikanz (2-seitig)	,010	,005	,009	,004

a. Wilcoxon-Test

b. Basiert auf negativen Rängen.

Nichtparametrische Tests

Mann-Whitney-Test

Ränge

	ersterVersuch	N	Mittlerer Rang	Rangsumme
NASA_test_gesamtwert	VIS	10	10,15	101,50
	MUL	10	10,85	108,50
	Gesamt	20		

Statistik für Test^a

	NASA_test_g esamtwert
Mann-Whitney-U	46,500
Wilcoxon-W	101,500
Z	-,265
Asymptotische Signifikanz (2-seitig)	,791
Exakte Signifikanz [2* (1-seitige Sig.)]	,796 ^b

a. Gruppenvariable: ersterVersuch

b. Nicht für Bindungen korrigiert.

Anhang F - Einverständniserklärung

Einverständniserklärung

Name, Vorname: _____

Hiermit erkläre ich mich einverstanden, dass im Rahmen der Studie Daten über mich gesammelt und anonymisiert aufgezeichnet werden. Außerdem stimme ich zu, dass die Ergebnisse zu Studienzwecken weiter verwendet werden dürfen. Es wird gewährleistet, dass meine personenbezogenen Daten nicht an Dritte weitergegeben werden.

Datum, Unterschrift

Selbstständigkeitserklärung

Hiermit versichere ich, dass ich meine Masterarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Datum:

.....

(Unterschrift)