

Verwendung von Eye-Tracking zur Steuerung von Bedienelementen in Virtual Reality

Studienarbeit

des Studiengangs -todo-

an der Dualen Hochschule Baden-Württemberg Stuttgart

von

Clemens Mollik, Jörn Herbstritt

-todo-

Bearbeitungszeitraum
Matrikelnummer, Kurs
Ausbildungsfirma
Betreuer

Oktober 2019 - Juni 2020
3394182, 7008632, STG-TINF17ITA
Robert Bosch GmbH, -todo-
-todo-

Abstract

TODO: deutscher Abstract....

Abstract

TODO: english abstract....

Inhaltsverzeichnis

Abkürzungsverzeichnis	V
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VII
Formelgrößenverzeichnis	VIII
Formelverzeichnis	IX
Listings	X
1 Einleitung	2
1.1 Auswertung	2
1.1.1 Tescht	2
2 Grundlagen	3
2.1 Virtual Reality	3
2.2 VR-Headset	3
2.2.1 Head-Mounted Display	4
2.2.2 Controller	4
2.2.3 Lighthouse Tracking System	4
2.3 Eyetracking	6
2.4 Unity	7
2.4.1 steamVR	7
2.4.2 hmd_eyes	7
2.5 Fitts' Law	7
3 Stand der Technik	8
3.1 Fitts' Gesetz	8
3.1.1 Fitts' Gesetz und Eye-Tracking	9
3.1.2 Fitts' Gesetz und Virtual Reality	10
3.1.3 Fitts' Gesetz mit Eye-Tracking in Virtual Reality	10
3.2 Interaction techniques using head gaze for virtual reality	10
3.2.1 Einordnung der Relevanz	11

3.3	Electrooculogram-based virtual reality game control using blink detection and gaze calibration	12
3.3.1	Einordnung der Relevanz	12
3.4	Eye-Gaze-Controlled Telepresence Robots for People with Motor Disabilities	13
3.4.1	Einordnung der Relevanz	13
3.5	Behavior Analysis of Indoor Escape Route-Finding Based on Head-Mounted VR and Eye Tracking	13
3.5.1	Einordnung der Relevanz	13
3.6	VR-HMD Eye Tracker in Active Visual Field Testing	13
3.6.1	Einordnung der Relevanz	13
3.7	Eye-gaze-triggered Visual Cues to Restore Attention in Educational VR . .	13
3.7.1	Einordnung der Relevanz	13
4	Architektur der Testumgebungen	14
5	Durchführung der Probandenversuche	15
5.1	Versuchsaufbau	15
5.2	Fragebogen	15
5.3	Versuche	15
6	Diskussion der Ergebnisse	16
7	Fazit	17
7.1	Zusammenfassung	17
7.2	kritische Würdigung	17
7.3	Ausblick	17
Anhang		A
	Literatur	A
Glossar		B

Abkürzungsverzeichnis

BSP	Board Support Package
HMD	Head-Mounted Display
VR	Virtual Reality

Abbildungsverzeichnis

2.1	HTC Vive HMD	4
2.2	Controller der HTC Vive	5
2.3	Lighthouse Basis Station	5
2.4	Pupil Core Headset	6
3.1	Resultat der Befragung	11

Tabellenverzeichnis

Formelgrößenverzeichnis

a	rad	Bedeutung von a
b	rad	Bedeutung von b
λ	rad	Bedeutung von lambda
ϕ	rad	Bedeutung von phi

Formelverzeichnis

Listings

Ideen für Studie

Vergleich Controller mit Eye-Tracker. Bsp.: Laserpointer

Zielscheibe (Pfeil & Bogen)

Allgemein Zuverlässigkeit / Genauigkeit Eye-Tracking-Systems

Anschauen der Kalibrierung

Auswertung

Bogen anfertigen für Bewertung der Versuche

10 Sachen auswählen (Zeit messen insgesamt (Controller & Augen))

Tescht

1 Einleitung

Vergleich Controller mit Eye-Tracker. Bsp.: Laserpointer

Zielscheibe (Pfeil & Bogen)

Allgemein Zuverlässigkeit / Genauigkeit Eye-Tracking-Systems

Anschauen der Kalibrierung

1.1 Auswertung

Bogen anfertigen für Bewertung der Versuche

10 Sachen auswählen (Zeit messen insgesamt (Controller & Augen))

1.1.1 Tescht

2 Grundlagen

In dem folgenden Kapitel werden die Grundlagen dieser Arbeit beschrieben. Zuerst wird eine Definition für Virtual Reality festgelegt. Danach werden die Headsets für Virtual Reality und für das Eye-Tracking vorgestellt. Am Ende der Grundlagen folgt eine Einführung in die Laufzeit- und Entwicklungsumgebung Unity.

2.1 Virtual Reality

Für den Begriff Virtual Reality (VR) oder Virtuelle Realität existiert keine einheitliche Definition. Für Singh und Singh ist VR eine wissenschaftliche Technologie für die Mensch-Maschine-Interaktion zum Verstehen und Simulieren einer realen oder virtuellen Umgebung. der Mensch-Maschine-Interaktion zum Verstehen und Simulieren der realen oder virtuellen Umgebung in einem 3D-Raum, die der realen Umgebung sehr ähnlich ist und die Natur in einem bestimmten Bereich erlebt, indem sie auftaucht und physische Präsenz in dieser Umgebung zeigt, indem Hochleistungscomputer zusammen mit sensorischer Ausrüstung wie Handschuhen, Kopfhörern usw. verwendet werden.“[SS17, vgl.]

2.2 VR-Headset

Im Rahmen dieser Arbeit wird das VR-Headset HTC Vive verwendet. Das Headset wurde gemeinsam von den Firmen HTC und Valve entwickelt. Das VR-Headset lässt sich in die folgenden 3 Teile unterteilen.

[CKK19]



Abbildung 2.1: HTC Vive HMD [Corc]

2.2.1 Head-Mounted Display

Head-Mounted Display (HMD)

TODO: Abbildung von Headset

2.2.2 Controller

TODO: Abbildung von Controller

2.2.3 Lighthouse Tracking System

TODO: Abbildung von Basisstation



Abbildung 2.2: Controller der HTC Vive [Corb]



Abbildung 2.3: Lighthouse Basis Station [Cora]

2.3 Eyetracking

Eyetracking ist eine Technologie, die erkennt, in welche Richtung eine Person ihren Blick richtet. Hierfür werden beim Eyetracking Blick sowie Augenbewegungen erfasst. Die hauptsächlichen Parameter, die durch das Eyetracking erfasst werden, sind Fixationen (von dem Benutzer fixierter Punkt oder Objekt), Sakkaden (schnelle Augenbewegungen bei der Erfassung eines neuen Fixpunktes) und Regression (Rücksprung zu vorherigen Fixpunkte oder Objekte).

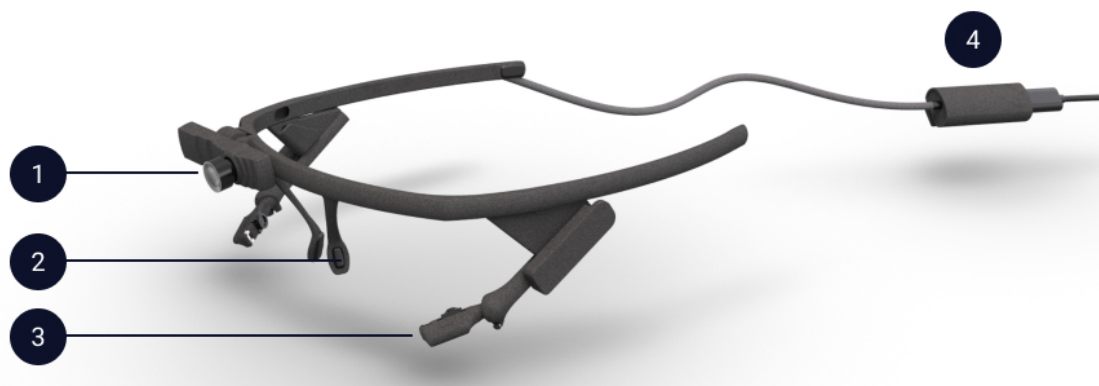


Abbildung 2.4: Pupil Core Headset

TODO: Quelle aus Pupil Labs Dokumentation -> Core -> Hardware

Im Rahmen dieser Arbeit wird ein Eyetracker von Pupil Labs verwendet. Sie entwickeln seit 2014 die Plattform Pupil Core, die aus einer Open-Source-Suite sowie dem Eyetracker selber besteht. In Abbildung 2.4 ist das tragbare Eyetracker Headset Pupil Core zu sehen, welches wie eine Brille getragen wird. An dem Headset ist eine Blickfeldkamera (Nummer 1) angebracht, welche das Blickfeld des Benutzers aufnimmt. Mithilfe der Augenkameras (Nummer 3) lässt sich das komplette Auge erfassen. Die Augenkameras lassen sich individuell auf die Augen einstellen. Ein USB-C Kabel (Nummer 4) dient als Stromversorgung sowie für den Austausch der Videodaten der Kameras. Wird der Eyetracker mit dem Computer verbunden, dann lässt sich mit der Pupil Core Software das Eyetracking starten. Mithilfe der Software lassen sich die Eyetracking-Daten aus den Videostreams auslesen, auswerten und über eine Netzwerk Schnittstelle zur Verfügung stellen. Zudem kann in das Umgebungsvideo der Blickfeldkamera der Punkt angezeigt

werden, auf den der Benutzer seinen Blick fixiert.

Da in dieser Arbeit das Eyetracking innerhalb einer VR-Umgebung untersucht werden soll, wird ein speziell für die HTC Vive entwickelter Eyetracker von Pupil Labs verwendet. Dies ist ein Add-on von Pupil Labs, welches in das HTC Vive Headset eingebaut wird und während dem Tragen des Headsets verwendet werden kann.

TODO: Bild von Add-on hinzufügen

[KPB14]

2.4 Unity

2.4.1 steamVR

2.4.2 hmd_eyes

2.5 Fitts' Law

3 Stand der Technik

In diesem Kapitel werden bisherige Forschungsarbeiten mit ähnlichen Ansätzen vorgestellt und mit dieser Arbeit in den Kontext gestellt. Es wurden bereits mehrere alternative Steuermöglichkeiten zu den herkömmlichen VR-Controllern erforscht und erprobt.

3.1 Fitts' Gesetz

Fitts' Gesetz (im englischen Fitts's Law) ist eine vom US-Amerikanischen Psychologe Paul Morris Fitts im Jahre 1954 veröffentlichte Metrik, welche die Schwierigkeit der Auswahl eines Zielobjekts numerisch beschreibt. Die zwei Parameter zur Bestimmung des "Index of Difficulty" (Index der Schwierigkeit) sind zum einen die Distanz D zu dem auszuwählenden Objekt und zum anderen die Breite dieses Objekts W . Die Berechnung erfolgt dann nach folgender Formel:

$$ID = \log_2 \left(\frac{2D}{W} \right) \quad (3.1)$$

Eine Erhöhung der Distanz hat eine Erhöhung der Schwierigkeit zur Folge, wohingegen eine Erhöhung der Breite das Senken des Schwierigkeitsindex ID zur Folge hat. Ein niedriger Wert für ID bedeutet, dass das Objekt leicht zu treffen ist, ein hoher Wert bedeutet, dass das Objekt schwer zu treffen ist. Basierend auf dieser Berechnung ist es möglich, die Bewegungsdauer (Movement Time) MT zu berechnen. Die Berechnung von MT ist mit folgender Formel möglich:

$$MT = a + b * ID = a + b * \log_2 \left(\frac{2D}{W} \right) \quad (3.2)$$

Die Konstanten a und b sind Zeitkonstanten, die von dem verwendeten Eingabegerät abhängig sind. Dabei beschreibt a die Zeit, bis eine Bewegung beginnt und b beschreibt kann als Beschleunigungskonstante gesehen werden. Daher wird mit einer erhöhten Distanz die Zeit, bis das Objekt ausgewählt wird ebenfalls höher.

Fitts' Gesetz wird vor allem in der Erstellung von Benutzeroberflächen (User Interfaces, kurz UIs) eingesetzt, um Schaltflächen für Nutzereingaben möglichst so zu gestalten, dass sie schnell und einfach getroffen werden können. Durch die begrenzte Fläche auf einem Computerbildschirm bietet sich hier ebenfalls die Möglichkeit, an den Rändern des Bildschirms wichtige Bedienelemente zu platzieren, um eine schnellere Bedienung des Programms vornehmen zu können. Im Betriebssystem Windows 10 ist zum Beispiel der Startknopf in der unteren linken Ecke platziert. Dadurch hat dieser Knopf eine "unendliche" Höhe und Tiefe, da der Nutzer die Maus unendlich weit nach unten links bewegen kann und immer den Startknopf treffen wird. Dadurch wird die Zeit, bis ein Nutzer den Knopf getroffen hat und aktivieren kann deutlich reduziert.

Bei Experimenten mit Fitts' Gesetz werden sowohl die Breite W und die Distanz D variiert, jedoch sollten die beiden Parameter nicht gleichzeitig variiert werden, um eine genauere Aussagekraft zu erlangen.

Da Fitts das Ursprungsexperiment ausschließlich in einer Dimension durchgeführt hat, war es ausreichend, ausschließlich die Breite des Objekts senkrecht zur Bewegungsrichtung als Breite W zu nutzen. Da aber ein Computerbildschirm zweidimensional ist, muss der Parameter angepasst werden, um die zweite Dimension ebenfalls in W darzustellen. Ansätze hierfür sind beispielsweise die Fläche des Objekts als W zu nutzen, oder die Breite mit der Höhe zu addieren. Da diese Ansätze aber nicht die Bewegungsrichtung beachten, wird in den meisten Fällen die Länge des Objekts in Bewegungsrichtung verwendet. Deshalb werden in den meisten Untersuchungen zu dem Thema Kreise verwendet, da hier W unabhängig von der Bewegungsrichtung konstant ist.

3.1.1 Fitts' Gesetz und Eye-Tracking

Bei Eye-Tracking Steuerung wird weiterhin diskutiert, ob, und wenn dann wie, Fitts' Gesetz weiterhin einsetzbar ist. Eine weite Bewegung mit dem Auge zeichnet sich, im Gegensatz zu einer weiten Bewegung mit der Maus, nicht über die zurückgelegte Distanz, sondern über den Winkel zwischen zwei Objekten aus. Unter Berücksichtigung dieses Aspekts ist es dann möglich, Fitts' Gesetz anzuwenden, allerdings können hier Ungenauigkeiten auftreten, die z.B. bei der Steuerung mit einer Maus nicht auftreten können. Dies kann zum Beispiel durch Bewegungen des gesamten Kopfes und nicht nur der Augen geschehen.

3.1.2 Fitts' Gesetz und Virtual Reality

Bei der Erstellung einer Nutzeroberfläche in der virtuellen Realität ist Fitts' Gesetz grundsätzlich weiterhin anwendbar. Allerdings ist hier die Distanz zwischen zwei Punkten durch die komplette Immersion des Nutzers in der virtuellen Welt besser durch die Verwendung von Winkeln, als der Verwendung von Pixeln als Maßeinheit anzugeben. Durch eine festgesetzte Position des Nutzers ist der Winkel zwischen zwei Objekten immer der selbe und somit die Bewegung eines Zeigers in Form eines Laserpointers oder des Blicks immer mit der gleichen Bewegungslänge verbunden. Dies gilt auch für die Breite eines Objektes. Allerdings können sich diese Werte durch unbeabsichtigt/intuitiv stattfindende Kopfbewegungen ebenfalls wieder verändern. Dies muss bei Experimenten in der virtuellen Realität stets berücksichtigt werden.

3.1.3 Fitts' Gesetz mit Eye-Tracking in Virtual Reality

3.2 Interaction techniques using head gaze for virtual reality

Diese im Mai 2016 veröffentlichte Arbeit hat sich zum Ziel gesetzt, die Position des Kopfes zu nutzen, um die ungefähre Blickrichtung des Nutzers zu erkennen und somit natürlichere Interaktionen in der virtuellen Realität zu ermöglichen. Die Motivation hinter der Arbeit war es, die doch sehr unnatürliche Steuerung über Controller zu ersetzen, um so eine bessere Immersion in die virtuelle Welt zu gewährleisten. Die Wahl, die Kopfbewegung als Steuerung zu nutzen, ist gefallen, da diese Bewegung von allen natürlich vorkommenden interaktiven Bewegungen eines Nutzers (wie z.B. Handbewegungen, Augenbewegungen, Fuß/Beinbewegungen und Kopfbewegungen) am einfachsten zu erfassen und in VR übertragbar ist. Da ein Großteil der VR-Headsets auf dem Erfassen des Kopfes in einem Raum basiert (TODO vgl Kpaitel von Jörn wie er das erklärt), ist die Erfassung dieser Daten meist ohne größere Umstände möglich. Um ein Aussagekräftiges Ergebnis zu erlangen, wurden mehrere Versuchsreihen und anschließende Befragungen der Probanden durchgeführt. Die Autoren der Arbeit haben dafür einige Spiele (teils bereits in VR existierend, teils Portierungen von Smartphone Spielen in VR) so angepasst, dass Kopfbewegungen zur Steuerung eingelesen werden können. Als Eingabegeräte wurden ein Leap Motion Controller (ein Hand-Tracking Gerät), ein herkömmlicher Xbox Controller und das von den Autoren "Head Gaze"benannte Kopf-Tracking miteinander verglichen. Als VR-Headset wurde ein Oculus DK2 verwendet. Nach der ersten Versuchsreihe in dem Spiel SSlash the

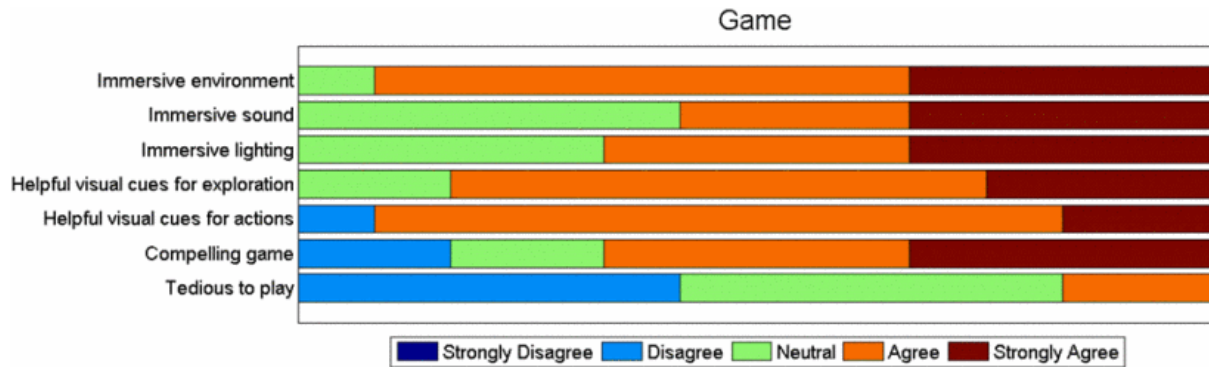


Abbildung 3.1: Resultat der Befragung

Fruit VR" hat sich Head Gaze als die intuitivste Eingabemethode herausgestellt. Daher wurden in den nachfolgenden beiden Versuchen mit den Spielen "Virtual Reality Experience" (kurz VREx) und "Dungeon VR" nur noch Head Gaze verwendet. Der Grund hierfür war zum Einen die Unzuverlässigkeit des Leap Motion Controllers, zum Anderen war der Xbox Controller nicht intuitiv genug für Bewegungen in VR. Da die Kernfrage die Immersion der neuen Steuerung war, wurden die Teilnehmer der Versuchsreihe nach ihrem Immersionsbefindnis befragt.

Wie in den oben stehenden Ergebnissen zu sehen ist, empfand der Großteil der Nutzer diese Steuerung als immersiv und nur wenige Nutzer haben sich von der Steuerung gestört gefühlt. Als weiteres Ziel der Forscher wird von den Forschern eine Umsetzung in Smartphone VR-Anwendungen erwähnt.

3.2.1 Einordnung der Relevanz

Die beschriebene Arbeit ist für den in dieser Arbeit vorgesehenen Versuchsaufbau von großer Relevanz, da es zu einem gewissen Grad eine Vorstufe darstellt. Der Begriff Head Gaze wird von den Forschern bewusst verwendet, da mit der Kopfbewegung auch der Blick (engl. Gaze) angenommen werden kann. Somit sind der Versuchsaufbau und die damit einhergehenden Ergebnisse sehr gut mit den in dieser Arbeit entstandenen Ergebnissen vergleichbar. (TODO Vergleich)

3.3 Electrooculogram-based virtual reality game control using blink detection and gaze calibration

Diese im September 2016 vorgestellte Arbeit beschreibt die Verwendung von einem Elektrookulogramm zur Bestimmung der Blickrichtung und zum Erkennen von Blinzeln. Diese Daten werden dann als Eingabe in einem VR-Spiel verwendet. Ein Elektrookulogramm (EOG) ist die Registrierung von Bewegungen der Muskeln, die zur Kontrolle der Augen dienen. Dabei werden an jeder Seite der Augenhöhle außerhalb Elektroden angebracht, mit welchen die Bewegung der Augen und Augenlider aufgezeichnet werden kann. Wie schon in Kapitel 3.2, lag auch hier das Ziel dabei, eine intuitivere und immersivere Eingabemethode zu entwickeln. Zur Überprüfung der Ergebnisse werden sowohl die Rate der falschen Eingaben als auch die subjektive Empfindung erfasst. Als Beispielapplikation wird das Spiel "VRrailSurfer" verwendet. Das Spielprinzip ist vergleichbar mit dem bekannten Smartphone Spiel Subway Surfers. Eine Bewegung der Augen nach links oder rechts hat eine Bewegung der Spielfigur in die entsprechende Richtung zur Folge. Zum Springen müssen die Augen nach oben bewegt werden. Die mit dieser Methode erzielte Genauigkeit der Steuerung lag nach einer Reihe von Versuchen mit 10 Probanden im Durchschnitt bei 78%, wohingegen die Blinzelerkennung mit 96% deutlich höher lag. Die Forscher haben allerdings in diesem Beispiel nur die Steuerung über den Blick im Spiel implementiert, sie gehen aber davon aus, dass eine Genauigkeit von rund 96% auch in einem Spiel mit Blinzelerkennung umsetzbar ist.

3.3.1 Einordnung der Relevanz

Das Paper ist für diese Arbeit auch relevant. Trotz der komplett anderen Eye-Tracking Methode sind einige Punkte auf diese Arbeit übertragbar. Der Ansatz, Augenbewegungen über ein EOG zu erfassen, ist allerdings eher suboptimal, da die Genauigkeit nicht so hoch ist, wie bei dem in dieser Arbeit verwendeten System, da mehr Komplikationen, wie etwa ein verrutschen der Elektroden, auftreten können. Der Fokus des Papers lag allerdings eher auf der Machbarkeit eines EOGs in Verbindung mit einem VR-Headset, weshalb kein Vergleich mit anderen Eingabemethoden stattgefunden hat.

3.4 Eye-Gaze-Controlled Telepresence Robots for People with Motor Disabilities

3.4.1 Einordnung der Relevanz

3.5 Behavior Analysis of Indoor Escape Route-Finding Based on Head-Mounted VR and Eye Tracking

3.5.1 Einordnung der Relevanz

3.6 VR-HMD Eye Tracker in Active Visual Field Testing

3.6.1 Einordnung der Relevanz

3.7 Eye-gaze-triggered Visual Cues to Restore Attention in Educational VR

3.7.1 Einordnung der Relevanz

4 Architektur der Testumgebungen

TODO:

5 Durchführung der Probandenversuche

TODO:

5.1 Versuchsaufbau

5.2 Fragebogen

5.3 Versuche

6 Diskussion der Ergebnisse

TODO:

7 Fazit

TODO:

7.1 Zusammenfassung

7.2 kritische Würdigung

7.3 Ausblick

Anhang

Literatur

- [CKK19] Viviane Clay, Peter König und Sabine König. „Eye tracking in virtual reality“. In: *Journal of Eye Movement Research* 12.1 (2019). DOI: [10.16910/jemr.12.1.3](https://doi.org/10.16910/jemr.12.1.3). URL: <https://bop.unibe.ch/JEMR/article/view/4332-Clay-final-sub>.
- [Cora] HTC Corporation. *Vive Hardware Base Stations*. URL: https://www.vive.com/media/filer_public/vive/product-overview/vive-hardware-base-stations.png.
- [Corb] HTC Corporation. *Vive Hardware Controllers 1*. URL: https://www.vive.com/media/filer_public/vive/product-overview/vive-hardware-controllers-1.png.
- [Corc] HTC Corporation. *Vive Hardware HMD 1*. URL: https://www.vive.com/media/filer_public/vive/product-overview/vive-hardware-hmd-1.png.
- [KPB14] Moritz Kassner, William Patera und Andreas Bulling. „Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-Based Interaction“. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. UbiComp '14 Adjunct. Seattle, Washington: Association for Computing Machinery, 2014, 1151–1160. ISBN: 9781450330473. DOI: [10.1145/2638728.2641695](https://doi.org/10.1145/2638728.2641695). URL: <https://doi.org/10.1145/2638728.2641695>.
- [SS17] N. Singh und S. Singh. „Virtual reality: A brief survey“. In: *2017 International Conference on Information Communication and Embedded Systems (ICICES)*. 2017, S. 1–6. DOI: [10.1109/ICICES.2017.8070720](https://doi.org/10.1109/ICICES.2017.8070720).

Glossar

Glossareintrag

Ein Glossar beschreibt verschiedenste Dinge in kurzen Worten.