

Entwicklung einer Methode zur Identifikation dreidimensionaler Blickbewegungen in realer und virtueller Umgebung

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von Dipl.-Inf. Sascha Weber
geboren am 05.04.1978 in Erlabrunn

eingereicht im Mai 2016, verteidigt am 07.07.2016

Die Dissertation wurde in der Zeit von März/2011 bis Mai/2016
im Institut für Arbeits-, Organisations- und Sozialpsychologie angefertigt.

Gutachter: Prof. Dr. rer. nat. habil. Sebastian Pannasch
Jun.-Prof. Dr. Stefan Scherbaum

Diese Promotion wurde durch die Europäische Union und den Freistaat Sachsen aus
Mitteln des Europäischen Sozialfonds (ESF) SAB Nr.: 100080340 gefördert.



Danksagung

Ich möchte allen Beteiligten auf das Herzlichste danken, die mich während meiner Forschung und der Anfertigung dieser Arbeit in den zurückliegenden Jahren unterstützt haben.

Mein größter Dank gilt meiner Frau Natalie und meinen Töchtern Melina und Melody, die mich mit sehr viel Verständnis, Geduld und vor allem Liebe immer wieder neu motivieren konnten. Ohne sie hätte ich mich der Forschung nicht mit dem notwendigen Enthusiasmus widmen können. Auch meinen Vater möchte ich herzlich danken, durch dessen Expertise als Ingenieur die technische Umsetzung der Untersuchungsumgebung erst möglich wurde.

Ebenso dankbar bin ich Prof. Boris M. Velichkovsky und Prof. Sebastian Pannasch für das entgegengebrauchte Vertrauen, die wissenschaftlich wertvollen Hinweise und letztendlich die hervorragende Betreuung dieser Arbeit. Weiterhin bedanke ich mich bei Jun.-Prof. Stefan Scherbaum für die Bereitschaft der Begutachtung dieser Arbeit.

Ein herzliches Dankeschön an dieser Stelle auch an die Kolleginnen und Kollegen der Professur für Ingenieurpsychologie und angewandte Kognitionsforschung der TU Dresden für die kollegiale Unterstützung, die vielen fachlichen Diskussionen und damit verbunden auch die Inspiration neue Wege zu gehen. Insbesondere danke ich Dr. Jens Helmert und Dipl. Psych. Rebekka Schubert für die freundschaftliche Zusammenarbeit aber auch konstruktiven Anmerkungen, die mich meine Arbeit aus anderen Blickwinkeln betrachten ließen.

Ein weiteres Dankeschön möchte ich dem Lehrstuhl für Mediengestaltung der TU Dresden Prof. Rainer Groh, Dr.-Ing. Ingmar S. Franke, Dr.-Ing. Jan Wojdziak und Dr.-Ing. Martin Zavesky für die interdisziplinäre Kooperation und Hilfe bei der Umsetzung der stereoskopischen Visualisierung widmen. Dr.-Ing. Dietrich Kammer möchte ich hierbei besonders hervorheben, der in vielen gemeinsamen Entwicklungsschritten nie die Motivation verlor und durch die Umsetzung der Schnittstelle zwischen Visualisierungs- und Messmethode entscheidend zum Erfolg beitrug.

Den Mitarbeitern der Interactive Minds Dresden GmbH sowie des Partnerunternehmens LC Technologies, Inc. (USA) möchte ich für die organisatorische sowie technologische Mithilfe recht herzlich danken. Dixon Cleveland gilt ein besonderes Dankeschön, der mich bei der Einarbeitung in das komplexe Thema der Blickbewegungsmessung und Fixationserkennung unterstützte und mit mir die neuen Ansätze dieser Arbeit in zahlreichen Gesprächen diskutierte.

Inhaltsverzeichnis

1	Einleitung.....	1
2	Physiologische Grundlagen und psychologische Aspekte	3
2.1	Blickbewegungen	3
2.2	Blickbewegungen und kognitive Prozesse	5
3	Räumliche visuelle Wahrnehmung und Stereoskopie.....	9
3.1	Tiefenkriterien	9
3.2	Stereoskopische Visualisierung.....	14
3.2.1	Polarisationsfilter	14
3.2.2	Shutterverfahren	15
3.2.3	Stereobildgenerierung.....	15
3.3	Wahrnehmungsprobleme in stereoskopischer Umgebung.....	17
3.3.1	Visuelle Überforderung durch erzwungene Vergenz	18
3.3.2	Perspektivische Verzerrung.....	19
4	Blickbewegungsmessung	21
4.1	Parameter des menschlichen Auges	22
4.2	PCCR-Methode.....	22
4.3	Algorithmus zur Blickortbestimmung	24
4.4	3D-Verfahren	26
4.4.1	Eyetracking in realer Umgebung.....	26
4.4.2	Stereoprojektion und Eyetracking.....	29
4.5	Technologische Herausforderungen.....	35
5	Zielsetzung und Konzeption	39
6	Studie I: Stereoskopische Visualisierung und Eyetracking	43
6.1	Fragestellungen und Hypothesen.....	44
6.2	Methode	45
6.3	Ergebnisse.....	49
6.4	Diskussion.....	52
7	Bestimmung dreidimensionaler Blickpositionen.....	55
7.1	Anforderungsprofil eines universellen Algorithmus.....	55
7.2	Ein vektorbasierter Ansatz.....	56

7.2.1	Mathematisches Modell.....	56
7.2.2	Eyetracking-Parameter	58
7.2.3	Koordinatensysteme	59
7.2.4	Gaze3D.dll	60
7.2.5	Performance-Test.....	62
7.2.6	Charakteristik der Berechnungsmethode.....	63
8	Studie II: Eyetracking in realer Welt	65
8.1	Fragestellung und Hypothesen.....	65
8.2	Methode.....	66
8.3	Ergebnisse	71
8.4	Diskussion	75
9	Studie III: Eyetracking in einer VR-Umgebung	77
9.1	Fragestellungen und Hypothesen	77
9.2	Methode.....	78
9.3	Ergebnisse	84
9.4	Diskussion.....	86
10	Bestimmung dreidimensionaler Blickbewegungen	89
10.1	Zweidimensionale Methoden	89
10.2	Anforderungen an eine räumliche Blickbewegungserkennung	95
10.3	Ein ellipsoider 3D-Fixationserkennungsalgorithmus	97
10.3.1	Relation zwischen Fixationstiefe und Messgenauigkeit.....	98
10.3.2	Mathematisches Grundmodell.....	98
10.3.3	Fixation3D.dll.....	101
10.4	Berechnung und Analyse dreidimensionaler Fixationen	103
10.4.1	Zeitlicher und örtlicher Schwellwert	103
10.4.2	3D-Fixationsparameter	103
10.5	Fazit.....	107
11	Zusammenfassung und Perspektiven.....	109
	Literaturverzeichnis	121
	Abbildungsverzeichnis	129
	Tabellenverzeichnis	133
	Quellcodeverzeichnis.....	135

1 Einleitung

Das Verständnis über die Blickbewegungen und kognitive Prozesse ist nicht nur für die Kognitionsforschung von großem Interesse. Auch in alltäglichen Bereichen des Lebens stellt sich die Frage, wie wir unsere Umwelt in unterschiedlichen Situationen visuell wahrnehmen. Entsprechende Untersuchungen können in realen Szenarien und aufgrund neuer innovativer 3D-Verfahren in Umgebungen der virtuellen Realität (VR) durchgeführt werden.

Für eine stereoskopische Visualisierung virtueller Welten gibt es verschiedene Technologien wie Shutterbrille, Anaglyph- oder Polarisationsfilter. Diese ermöglichen die Darbietung getrennter Bilder für linkes bzw. rechtes Auge und erzeugen durch unterschiedliche Perspektiven einen Tiefeneindruck, welcher die dargestellte Szene realitätsnah erscheinen lässt. Derartige Verfahren erfreuen sich in 3D-Kinos oder im Entertainmentbereich (Nguyen et al., 2005) großer Beliebtheit und werden darüber hinaus auch in verschiedene Forschungs- und Anwendungsbereiche integriert. Durch die Simulation von Situationen mit hinreichender Vergleichbarkeit zur Realität lassen sich unter anderem sicherheitsrelevante Aspekte zur Analyse von Fehlverhalten und Unfallvermeidung gezielt untersuchen. Die planmäßige und systematische Erzeugung von Gefahrensituationen, beispielsweise im Straßenverkehr, ist in der Realität weder möglich noch ethisch vertretbar, lässt sich jedoch in VR Umgebungen realitätsnah umsetzen. Mit dreidimensionalen Visualisierungen eröffnen sich neue Perspektiven der Wahrnehmungsanalyse, insbesondere bei Fragen zur Sicherheit und Unfallvorbeugung.

Bei der Erforschung von Wahrnehmungs- und Aufmerksamkeitsprozessen wird unter anderem die Methode der Blickbewegungsmessung (*Eyetracking*) angewandt, da das Sehen für uns Menschen die wichtigste Sinnesmodalität darstellt. Grundvoraussetzung für visuelle Wahrnehmung sind die Bewegungen der Augen (vgl. Hayhoe & Ballard, 2005), speziell vor dem Hintergrund, dass die Ausrichtung der visuellen Aufmerksamkeit und die der Augen korrespondieren (Henderson, 2007). Im Straßenverkehr ist es oftmals eine Frage von Sekundenbruchteilen, um relevante Aspekte der Fahrumgebung zu sehen und angemessen darauf reagieren zu können oder bei autonomer Fahrunterstützung in gefährlichen Situationen die Kontrolle über das Fahrzeug zurückzugewinnen (Braunagel, Stolzmann, Kasneci & Rosenstiel, 2015; Gold, Dambock, Lorenz & Bengler, 2013). Ebenso ist es für einen Chirurgen während der Diagnose und der Operation von entscheidender Wichtigkeit, visuelle Informationen wahrzunehmen und in Entscheidungen zu berücksichtigen (Burgert et al., 2007; Schulz et al., 2014). Weniger zeitkritisch, aber von dennoch zunehmender Bedeutung, ist

die Anwendung der Blickbewegungsmessung im Bereich der Mensch-Maschine-Interaktion, insbesondere wenn es darum geht, Schnittstellen durch die Analyse des Benutzerverhaltens zu evaluieren (Weber, Joos & Velichkovsky, 2004).

Die Erfassung von Blickbewegungen ist mittels verschiedener Methoden durchführbar (vgl. Eggert, 2007). Videobasierte Eyetracking-Systeme zeichnen Blickorte zeitlich hochauflösend mit bis zu 2.000 Hz und räumlich präzise mit weniger als $0,5^\circ$ Abweichung auf. Dabei erfolgt die Messung relativ zu einer ebenen Fläche, wie beispielsweise einem Monitor, einer Projektionsleinwand oder einem Szenevideo aus Sicht des Probanden. Aus den einzelnen Blickorten werden anhand von Identifikationsalgorithmen Blickbewegungen in Bezug auf die Messebene bestimmt und anschließend den Objekten der Szene zugeordnet. Allerdings sind Blickorte sowie Blickbewegungen bislang nur innerhalb dieser zweidimensionalen Ebene bestimmbar, obwohl natürliches Blickverhalten vorwiegend dreidimensional stattfindet. Eine Erweiterung der Blickbewegungsmessung auf den dreidimensionalen Raum würde der Erforschung visueller Wahrnehmung völlig neue Möglichkeiten eröffnen, jedoch sind herkömmliche Systeme für eine dreidimensionale Blickortmessung nicht konzipiert. Folglich ist auch wenig bekannt, wie sich Eyetracker in eine reale 3D-Messung oder stereoskopisch projizierte VR-Umgebung integrieren lassen und inwieweit anhand der verfügbaren Eyetracking-Daten der Tiefenparameter des Blickortes bestimmt werden kann. Schlussendlich gibt es bislang keine Algorithmen zur Identifikation räumlicher Blickbewegungen aus dreidimensionalen Blickorten.

Die vorliegende Arbeit stellt eine Methode vor, mit der dreidimensionale Blickorte sowohl in einer realen als auch in einer stereoskopisch projizierten VR-Umgebung anhand binokularer Eyetracking Technologien bestimmt werden können. Ein neu entwickelter Algorithmus identifiziert aus den berechneten 3D-Blickorten räumliche Blickbewegungen mit zeitlichen und örtlichen Parametern. Dadurch wird die räumliche Blickbewegungsanalyse möglich. Menschliches Blickverhalten kann mit dieser Methode zukünftig dreidimensional erforscht werden.

2 Physiologische Grundlagen und psychologische Aspekte

Die Erforschung der visuellen Wahrnehmung durch unsere Augen im Zusammenhang mit Okulomotorik und kognitiven Prozessen spielt in der wissenschaftlichen Psychologie von Beginn an eine tragende Rolle. Insbesondere Korrelationen zwischen Blickbewegungen und visueller Aufmerksamkeit wurden vielfältig untersucht und in der Literatur in Form von theorie- und hypothesenleitenden Metaphern umschrieben. Beispielsweise wird das Zentrum visueller Aufmerksamkeit mit einer Lupe (Eriksen & Yeh, 1985) oder dem Gaußschen Gradienten (Downing & Pinker, 1985) verglichen, wobei im Zentrum immer eine bessere Verarbeitung visueller Informationen stattfindet als im Randbereich.

Eine aus Sicht der Blickbewegungsforschung naheliegende Analogie zur Beschreibung von Zusammenhängen zwischen dem aktuellen Blickort und visueller Aufmerksamkeit ist die *Attentional Spotlight Metaphor*. Sie erklärt den Prozess anhand eines Scheinwerfers oder einer Taschenlampe, deren Lichtstrahl Objekte beleuchtet und diese erst dadurch erkannt werden können (Posner, Snyder & Davidson, 1980). Cavanagh (2004) erweiterte diesen Vergleich und vermutet, dass visuelle Aufmerksamkeit auch an mehreren unterschiedlichen Orten gleichzeitig vorhanden sein kann, wobei die Anzahl parallel wahrnehmbarer Bereiche von deren Größe abhängt. Er geht vom kleinsten Detail, welches ein Mensch sehen kann, bis hin zum gesamten visuellen Feld aus und konnte in seinen Experimenten fünf unterschiedliche Wahrnehmungsfelder nachweisen (Intriligator & Cavanagh, 2001). Dabei korrelierten Sehschärfe und Detailgrad der Objekte mit deren Erkennbarkeit. Fixiert man mit seinen Augen ein Fixationskreuz und wird gleichzeitig im peripheren Abstand ein einzelner Buchstabe dargeboten, ist dieser eindeutig identifizierbar. Wird derselbe Buchstabe in eine Buchstabenkombination bzw. in ein Wort eingebettet, ist eine eindeutige Identifikation nicht mehr gegeben. Die Ergebnisse zeigten, dass es zwar möglich ist Aufmerksamkeit in unterschiedlichen Bereichen des Sichtfeldes auszurichten, jedoch um visuelle Informationen detailliert aufnehmen zu können, müssen die Augen und die visuelle Aufmerksamkeit stets auf dasselbe Objekt gezielt ausgerichtet sein.

2.1 Blickbewegungen

Bei der visuellen Wahrnehmung werden die auf der Netzhaut (*Retina*) befindlichen Fotorezeptoren durch elektromagnetische Schwingungen des sichtbaren Lichtspektrums gereizt. Dabei finden in den Außensegmenten der Rezeptoren biochemische Reaktionen statt,

wobei elektrische Signale erzeugt und diese über den Sehnerv weitergeleitet werden (Dowling & Boycott, 1966). Es gibt zwei formentypische Arten von Rezeptoren, *Stäbchen* und *Zapfen* (Goldstein, Ritter & Herbst, 2002). Beide Zellentypen sind lichtempfindlich, jedoch ist das Stäbchensystem für kürzere Wellenlängen sensibler als das Zapfensystem. Stäbchen sind vorwiegend für das Sehen in der Dämmerung und bei Nacht von Bedeutung (*skotopisches Sehen*). Durch sie können keine Farben unterschieden werden, da im Gegensatz zu den Zapfen alle Stäbchen dasselbe Empfindlichkeitsspektrum aufweisen. Das Zapfensystem hingegen ist für Wellenlängen im oberen Bereich des elektromagnetischen Spektrums sensibler. Sie ermöglichen die Unterscheidung differenter Wellenlängenbereiche und somit die Farbwahrnehmung (*photopisches Sehen*).

Ein weiterer Aspekt ist die ungleichmäßige Verteilung der ca. 120 Millionen Stäbchen und 5 Millionen Zapfen auf der Retina. Erstere sind größtenteils in der Peripherie und letztere hauptsächlich in der Netzhautmitte angeordnet. Das Zentrum mit der größten Dichte an Zapfen wird als *Fovea Centralis* bezeichnet und bildet den Bereich des schärfsten Sehens. Hier werden Farben und Konturen am deutlichsten wahrgenommen. Das gesamte binokulare Blickfeld schließt einen Bereich von etwa 208° horizontal bzw. 120° vertikal ein (Wade & Swanston, 2013), jedoch ist scharfes Sehen auf einen Bereich von etwa 2° um den fixierten Punkt aufgrund des optischen Apparates und der Rezeptorverteilung eingegrenzt. Weiter entfernte Objekte werden wegen der abnehmenden Zapfendichte zunehmend unschärfer und farbloser wahrgenommen. Dementsprechend definiert sich der foveale ($< 2^\circ$), parafoveale ($2^\circ - 5^\circ$) und periphere ($> 5^\circ$) Sehbereich (Parkin, 2014).

Um seine Umwelt visuell wahrnehmen bzw. neue Informationen aufnehmen und verarbeiten zu können, muss der Mensch seine Augen ständig neu ausrichten. Dabei werden für die Ausrichtung und Fokussierung verschiedene Bewegungen ausgeführt. Die Neuausrichtung des Auges bzw. zielführende Blickbewegung wird als *Sakkade* bezeichnet und positioniert die Fovea von einem Ort der visuellen Informationsaufnahme zum nächsten. Damit wir bei einem Blickortwechsel unsere Umwelt klar und stabil wahrnehmen, wird die visuelle Informationsverarbeitung während einer Sakkade unterdrückt (Kowler, 2011). Im Gegensatz dazu erfolgt die Aufnahme visueller Informationen innerhalb einer stabilisierenden Blickbewegung bzw. *Fixation*. Dabei befindet sich der Blickort in Bezug auf einen Wahrnehmungsgegenstand im relativen Stillstand und die Rezeptorinformationen werden von der Retina zu den kortikalen Regionen des visuellen Systems weitergeleitet. Die visuelle Informationsaufnahme aus der Umwelt kann daher primär als Funktion der Fixationen betrachtet werden. Neben Sakkaden und Fixationen gibt es weitere Arten von Blickbewegungen wie beispielsweise Blinzeln, Mikrobewegungen, Folgebewegungen in Bezug

auf sich bewegende Objekte oder Nystagmen. Da diese nicht im Fokus dieser Arbeit stehen, wird für einen umfassenden Überblick auf die Arbeiten von Joos, Rötting und Velichkovsky (2003) sowie Holmqvist et al. (2011) verwiesen.

In vielen alltäglichen Situationen, wie beispielsweise beim Lesen eines Textes oder beim Betrachten eines Bildes, können die okulomotorischen Aktivitäten als Zusammenspiel von Sakkaden und Fixationen betrachtet werden (Velichkovsky et al., 2013). Sie bilden die Grundlage für unsere visuelle Wahrnehmung und geben nicht nur Aufschluss über die lokale Ausrichtung der Informationsaufnahme. Nach der *levels of processing*–Theorie (Craik & Lockhart, 1972) wird ein Zusammenhang zwischen der Dauer einer Fixation und der Verarbeitungstiefe angenommen (Velichkovsky, Dornhoefer, Pannasch & Unema, 2001). Die Analyse menschlicher Blickbewegungen zählt daher zu den objektiven Methoden der Erforschung visueller Wahrnehmung und kognitiver Prozesse.

2.2 Blickbewegungen und kognitive Prozesse

Bereits vor mehr als 80 Jahren führte Guy Thomas Buswell Blickbewegungsmessungen durch. Er gehörte zu den ersten Forschern, die sowohl die Blickrichtung bzw. den Blickort als auch die Dauer einzelner Fixationen innerhalb einer Szene bestimmen und analysieren konnten (Buswell, 1935). Dafür entwickelte er einen Apparat, mit dem er Licht von der Cornea auf einen lichtempfindlichen Film reflektierte. Über die Zwischenschaltung eines Ventilators wurde anhand der Blattrotation in einer Frequenz von 30 Hz die Zeit quantisiert. Der aufgezeichnete Lichtstrahl erschien im Film als Punktlinie, wobei jeder Punkt einer Zeitspanne von $1/30$ s entsprach. Damit ließ sich die Dauer der Fixation anhand der abgezählten Punkte ermitteln. Buswell präsentierte Versuchspersonen Gemälde und zeigte, dass Fixationen nicht zufällig in einer Szene angeordnet waren. Stattdessen neigten die Betrachter dazu Fixationen in informativen Bildregionen zu gruppieren. Buswell konnte nicht nur durchschnittliche Fixationsdauern und Sakkadenamplituden sondern auch die Variabilität innerhalb der Messungen abschätzen. Er schlussfolgerte, dass es einen Zusammenhang zwischen Blickbewegungen und visueller Aufmerksamkeit gab: „*Eye movements are unconscious adjustments to the demands of attention during a visual experience.*“

Nach den Anfängen bildeten sich durch die technologischen Weiterentwicklungen der Blickbewegungsmessung und dem steigenden Interesse an der Erforschung kognitiver Wahrnehmungsprozesse verschiedene Untersuchungsschwerpunkte heraus. Beispielsweise waren Blickbewegungen für McConkie und Rayner (1975) in der Leseforschung von besonderem Interesse. Sie entwickelten eine blickkontingente Methode, um Zusammenhänge

zwischen Fixationen und Wahrnehmungsprozessen während des Lesens zu analysieren. In einem Experiment variierten sie die dargebotenen Texte peripher mittel unterschiedlichen Mustern in Abhängigkeit vom aktuellen Fixationsort und beobachteten dabei Veränderungen im Leseverhalten. Die Forscher wiesen nach, dass die Probanden Informationen über Wortlängen-Muster von 12 bis 15 Buchstabenpositionen in Leserichtung ausgehend vom aktuellen Fixationsort erfassen konnten. Dass diese perzeptuelle Spanne asymmetrisch und entgegengesetzt der Leserichtung lediglich 3-4 Buchstabenpositionen lang ist, zeigten Rayner, Well und Pollatsek (1980). Spätere Blickbewegungsstudien in der Leseforschung setzten sich unter anderem mit okulomotorischen Prozessen (Rayner, Sereno & Raney, 1996), der Steuerung von Blickbewegungen (Rayner & Reingold, 2015) oder verschiedenen Lesestrategien (Liu, 2014; Rau, Moll, Snowling & Landerl, 2015) auseinander.

Beim Lesen ist die Ausrichtung der Aufmerksamkeit durch die vorgegebene Leserichtung, Zeilenumbrüche oder Anordnung von Textblöcken vorhersehbar. Anders verhält es sich bei der Wahrnehmung komplexer Szenen wie z.B. auf Gemälden, Bildern, Zeichnungen oder Fotos, denn hierbei können Objekte richtungsunabhängig betrachtet werden. Einen Überblick relevanter Studienergebnisse zur Szenenwahrnehmung geben Henderson und Hollingworth (1999). Die Autoren konstatieren, dass Fixationsorte nicht zufällig in einer Szene verteilt sondern in visuell oder semantisch informativen Regionen gruppiert sind. Die ersten Fixationen beim Anblick einer neuen Szene werden durch visuelle Auffälligkeiten und globale semantische Charakteristiken (z.B. das Szenekonzept) jedoch nicht durch semantische Merkmale lokaler Regionen gesteuert. Erst mit fortschreitender Betrachtungsdauer erfolgt die Fixierung sowie semantische Analyse lokaler Bereiche. Eisenberg und Zacks (2016) konnten unterschiedliche Blickmuster ebenso in dynamischen Szenen nachweisen. Sie zeigten ihren Probanden Videosequenzen mit alltäglichen Handlungsabläufen und fanden dabei sowohl exploratives als auch fokales Blickverhalten.

In der Leseforschung aber auch in den Untersuchungen zur Szenenwahrnehmung wird den Probanden vorwiegend in Laborexperimenten artifizielles Stimulusmaterial auf einer zweidimensionalen Präsentationsfläche wie dem Computermonitor oder einer Leinwand dargeboten. Weitaus anspruchsvoller ist die Blickbewegungsmessung in natürlichen Kontexten, denn hierbei werden Objekte der realen Welt in unterschiedlichen Tiefenebenen betrachtet und manipuliert. Tatler (2001) ließ Probanden in einer Küche Tee zubereiten und schaltete während dieser Tätigkeit zu unvorhersehbaren Zeitpunkten das Licht aus. Die Versuchspersonen sollten nun beschreiben, was sie als letztes gesehen hatten. Während des Experimentes wurden Blickbewegungen gemessen und eine am Kopf angebrachte Videokamera zeichnete zeitgleich die Szene aus Sicht des Probanden auf. Anhand der

erhobenen Eyetracking-Daten konnte der aktuelle Blickort im 2D-Szenevideo in Form eines Punktes dargestellt und in einem weiteren Analyseschritt manuell die berechneten Fixationen den Objekten der Szene zugeordnet werden. Tatler wies nach, dass die Teilnehmer leichten Zugriff zu Informationen in Bezug auf das Blickortzentrum hatten, allerdings die Angabe zum zuletzt gesehenen Objekt vom Zeitpunkt des Lichtausschaltens relativ zur aktuellen Fixation abhing.

In weiteren Studien mit natürlichen Kontexten wurde beispielsweise das Blickverhalten beim Autofahren mit Assistenzsystemen (Barakat et al., 2015; Gold et al., 2013; Vicente et al., 2015) oder Fahren bei Nacht (Winter, Fotios & Voelker, 2016), in Einkaufssituationen (Otterbring, Wästlund & Gustafsson, 2016) oder während sportlicher Aktivitäten (Hayhoe, McKinney, Chajka & Pelz, 2012; Land, 2009) untersucht. Auch in diesen Experimenten kamen mobile Eyetracker für die Blickbewegungsmessung mit einer zusätzlichen Kopfkamera zum Einsatz. Die ermittelten Blickorte waren später im Szenenvideo als farbige Markierung sichtbar und mussten den Objekten im Videobild zugewiesen werden. Eine dreidimensionale Blickortbestimmung und Objektzuordnung fand dabei nicht statt.

Durch die technologische Entwicklung der letzten 15 Jahre wurden die Methoden zur Blickbewegungsmessung in Kombination mit computergestützten Visualisierungsverfahren auch in Simulationsumgebungen integrierbar. Hierbei konnten natürliche Handlungsabläufe in eine definierbare Laborumgebung übertragen und beispielsweise kritische Situationen gezielt untersucht werden, wie Gefahren im Straßenverkehr (Borowsky, Shinar & Oron-Gilad, 2010) oder die Fahrzeugübernahme nach autonomen Fahren (Braunagel et al., 2015), in der Unfallforschung im Luftverkehr (Sarter, Mumaw & Wickens, 2007), in Simulationen für Handlungsabläufe im Kontrollraum einer Fabrikanlage (Sharma, Bhavsar, Srinivasan & Srinivasan, 2016) oder im Operationssaal (Di Stasi et al., 2016; He & Zheng, 2016; Schulz et al., 2014). Aber auch in diesen Beispielen wurden die Blickorte und Blickbewegungen zweidimensional in Bezug auf eine Projektionsfläche oder im Szenevideo des Betrachters bestimmt.

Es gibt viele weitere Anwendungsfelder, in der die Messung von Blickbewegungen eine Rolle spielt, wie in der Wirtschaft für Usability-Untersuchungen (Ricard & Garrison, 2015), bei der Bewertung von Produktdesigns (Li, Huang & Christianson, 2016) oder E-Learning Applikationen (Kabugo, Muyinda, Masagazi, Mugagga & Mulumba, 2016), in der Medizin für diagnostische und therapeutische Zwecke (Curie et al., 2016; Giuliani & Marchetti, 2015; Kumar, Dutta, Das & Lahiri, 2016), aber auch bei blickgesteuerten Benutzungsschnittstellen wie den Kommunikationshilfsmitteln für Locked-in-Syndrom Patienten (Calvo et al., 2008).

Die angeführten Beispiele zeigen, welche Bedeutung der Messung von Blickbewegungen in Forschung, Wirtschaft und Medizin zukommt, allerdings wurden Fixationen und Sakkaden weitestgehend in einer Ebene wie dem Computerbildschirm, einer Projektionswand oder dem aufgezeichneten Szene-Video aus Sicht des Probanden analysiert. Die dritte Dimension und damit verbundene Tiefeninformationen blieben bislang unberücksichtigt, obwohl menschliches Blickverhalten überwiegend dreidimensional erfolgt, z.B. für die Orientierung bzw. Einschätzen der eigenen Entfernung zu Objekten der Umwelt (Goldstein et al., 2002), der Wahrnehmung relativer Größeninformationen (Palmer, 1999) oder der Kontrolle der eigenen Bewegung und Bewegungen der Objekte in unserer Umgebung (Kandel, Schwartz & Jessell, 2000; Wade & Swanston, 2013). Aufgrund der Einschränkung auf zweidimensionale Messmethoden sind räumliche Analysen bis jetzt nicht möglich. Eine dreidimensionale Blickbewegungsmessung würde die Möglichkeit eröffnen, bislang in zwei Dimensionen untersuchte Blickmuster auch räumlich zu bestimmen und damit grundlegende Zusammenhänge zwischen Blickbewegungen, visueller Wahrnehmung und kognitiven Prozessen naturalistischer zu erforschen.

3 Räumliche visuelle Wahrnehmung und Stereoskopie

Damit der Mensch in einer dreidimensionalen Welt präzise und zielführend handeln kann, hat das visuelle System komplexe Mechanismen ausgebildet, um aus der wahrgenommenen Umgebung Informationen über die räumliche Tiefe bzw. die Entfernung zu Objekten extrahieren zu können.

3.1 Tiefenkriterien

Die dem visuellen System dafür zur Verfügung stehenden Informationen werden auch als *Tiefenkriterien* bezeichnet und lassen sich wie folgt gliedern:

- **Okulomotorische Informationen:** entstehen durch die Auswertung der Konvergenzstellung und Akkommodation der Augen
- **Monokulare Informationen:** ergeben sich aus den strukturellen Informationen eines unbewegten Bildes (z.B. Perspektive, Verdeckung, Schatten)
- **Bewegungsinduzierte Informationen:** erzeugt durch die Bewegung des Beobachters selbst oder durch die Bewegung von Objekten in dessen Umwelt
- **Stereoskopische Informationen:** basieren auf Interpretation von Bildunterschieden zwischen linkem und rechtem Auge

Okulomotorische Informationen

Um Objekte an unterschiedlichen Positionen der realen Umgebung scharf auf der Netzhaut abzubilden, müssen sich die Augen mit Hilfe der mit ihnen verbundenen Muskeln zum Objekt hin ausrichten und dabei die Brechkraft der Linse durch eine Formänderung zum Fixieren in unterschiedlichen Entfernungen anpassen (Fincham & Walton, 1957). Bei einer Fokussierung unterschiedlich entfernter Objekte beschreibt die *Vergenz* zusammenfassend das Einwärtsdrehen (*Konvergenz*) und das Auswärtsdrehen (*Divergenz*) der Augen. Nach Palmer (1999) nähert sich der Konvergenzwinkel ab einer Entfernung von 1,8 m - 2,4 m gegen 0° an und entspricht damit der Fokussierung eines Objektes in unendlicher Entfernung. Da die Winkelunterschiede bei weiter entfernten Objekten kaum messbar sind, die Konvergenz zur Tiefeninformationsgewinnung auf den Nahbereich eingegrenzt.

Als *Akkommodation* bezeichnet man den Prozess, bei dem die Ziliarmuskeln des Auges den optischen Fokus der Linse durch eine Formveränderung kontrollieren. Je nach Entfernung wird die Linse dabei unterschiedlich gekrümmt. Ziel der Akkommodation ist die Abbildung

eines scharfen Bildes auf der Netzhaut. Sobald die Unschärfe einen Schwellwert übersteigt, wird ein neuronales Signal zu den Ziliarmuskeln gesandt (Kotulak & Schor, 1986). Daraufhin findet eine Korrekturreaktion statt, die je nach Situation ein Zusammenziehen bzw. Entspannen der Ziliarmuskeln bewirkt. Bis zum vollständigen Abschluss dieser Fehlerkorrektur vergehen bis zu 500 ms.

Eine Akkommodationsanpassung allein bildet zwar für den Moment ein scharfes Bild auf der Netzhaut ab, die Ziliarmuskeln würden sich jedoch nach Abklingen des neuronalen Signals sofort entspannen und das Bild wieder unscharf erscheinen lassen. Dieses Problem wird durch den *Tonischen Controller* gelöst, der durch ständiges Überprüfen des neuronalen Akkommodationssignals (aktueller Fokus) mit einem Referenzsignal (optimaler Fokus) vergleicht. Bei Unterschieden über die Toleranzgrenzen hinaus werden entsprechende Fehlerkorrektursignale erzeugt und ermöglichen somit eine Aufrechterhaltung des Spannungszustandes der Muskulatur (Ebenholtz, 2001). Die Formveränderung der Linse ist physiologisch eingeschränkt. Goldstein et al. (2002) gehen davon aus, dass sich die Ziliarmuskeln des Auges ab einer Entfernung des betrachteten Objektes von ca. 6 Metern vollkommen entspannen. Innerhalb dieses Bereiches ist die Akkommodation als Tiefenkriterium hilfreich.

Nach Lambooi, Fortuin, Heynderickx und Ijsselsteijn (2009) sind Akkommodation und Vergenz zwei parallele, aneinander gekoppelte Regelsysteme (Abb. 3-1). Während die Akkommodation die retinale Sehschärfe steuert, regelt die Vergenz die retinale Disparität der beiden Augenbilder. Unter natürlichen Betrachtungsbedingungen interagieren beide Systeme, um ein klares Bild der Szene binokular wahrnehmen zu können. Dabei geht das Akkommodationssignal in den Regelkreis der Vergenz ein und umgekehrt. Beide Systeme tolerieren Abweichungen zum Referenzsignal der optimalen Abbildungsschärfe bzw. Disparität, um einen stabilen Zustand der Regelsysteme aufrecht zu erhalten und Änderungen der Akkommodation und Vergenz zu vermeiden. Die Ausgangssignale für Akkommodation sowie Vergenz auf der rechten Seite gehen rückgekoppelt als Eingangssignal in den jeweiligen Regelkreis ein und werden mit einem Referenzsignal der optimalen Abbildungsschärfe bzw. Disparität verglichen. Eine retinale Unschärfe innerhalb des *depth of focus* bewirkt noch keine Akkommodationsanpassung und variiert typischerweise zwischen 0.2 - 0.5 Dioptrien. Ähnlich wird ein Toleranzbereich innerhalb der *panum's fusional area* für die Vergenz angenommen. Um ein Objekt mit den Augen fixieren zu können, müssen beide optische Achsen daraufhin ausgerichtet sein. Minimale Abweichungen führen zu retinaler Disparität, liegen allerdings im Zentrum der Fovea bis 0.1° , bei 6° Exzentrizität bis 0.33° und ab 12° Exzentrizität bis 0.66° innerhalb des Toleranzbereichs, ohne eine Änderung der Vergenz auszulösen. Kommt es zu

Abweichungen über die Toleranzgrenzen hinaus, werden Korrektursignale durch die entsprechenden Controller erzeugt und zum aktuellen Entfernungssignal (*proximity*) addiert. Das Ergebnissignal der Akkommodation wird einerseits an das Regelsystem der Disparität weitergeleitet und andererseits mit dem Signal der tonischen Komponente kombiniert, welche eine verlangsamte Adaption an sich ändernde Betrachtungssituationen bewirkt. Hieraus ergibt sich das Ausgangssignal der Akkommodation, das rückgekoppelt als Vergleichssignal zur optimal scharfen Retinaabbildung eingeht. Parallel wird das Ergebnissignal der Vergenz an das Regelsystem der Akkommodation übertragen und ergibt in Kombination mit dem Signal der tonischen Komponente das Vergenz-Ausgangssignal, welches ebenfalls als Vergleichssignal zur optimalen Disparität dem Kreislauf neu zugeführt wird.

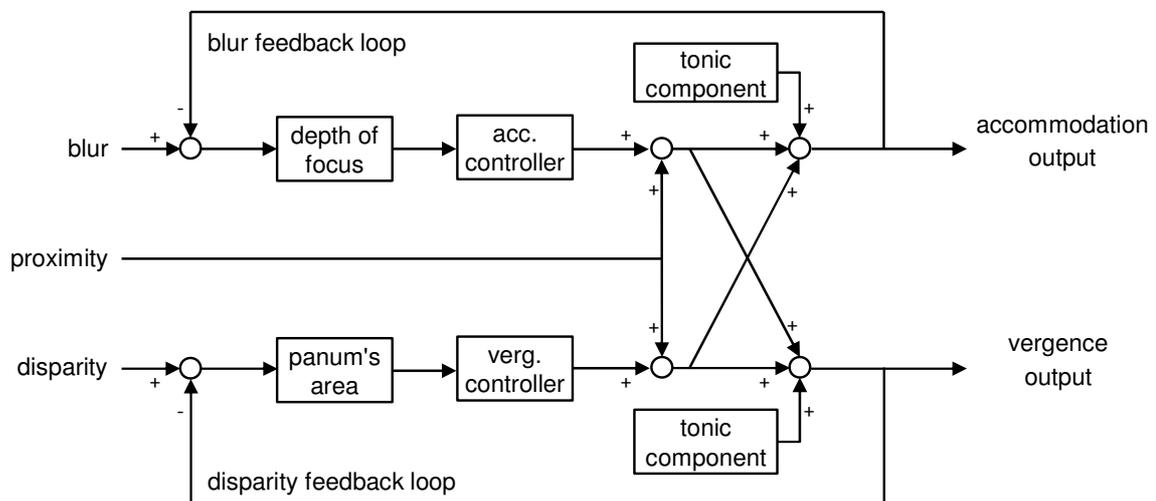


Abb. 3-1: Akkommodation-Vergenz-Modell nach Lambooj et al. (2009): Akkommodation und Vergenz modelliert als zwei parallele Feedback-Steuerysteme mit adaptiver Interaktion.

Monokulare Informationen

Tiefeninformationen, die auch wirksam sind, wenn man sie mit nur einem Auge betrachtet, nennt man *monokulare Tiefeninformationen*. Sie ergeben sich aus den räumlichen Strukturen der Objekte. Beispiele für derartige strukturelle Regelmäßigkeiten sind:

- **Verdeckung von Objekten:** Falls Objekt A einen Teil von Objekt B verdeckt, wird A als vor B liegend wahrgenommen.
- **Relative Höhe im Gesichtsfeld:** Höhere Objekte unterhalb der Horizontlinie und Objekte über dem Horizont, die im Gesichtsfeld weiter unten erscheinen, werden als weiter entfernt gesehen.
- **Relative Größe im Gesichtsfeld:** Bei physikalisch gleich großen Objekten nimmt das nähere einen größeren Teil im Gesichtsfeld ein.

- **Atmosphärische Perspektive:** Entferntere Objekte werden weniger scharf gesehen, da sich zwischen Beobachter und Objekt mehr Luft und damit mehr Staub, Wassertröpfchen, etc. befinden.
- **Lineare Perspektive:** Parallel verlaufende Linien konvergieren in der Ferne. Je größer die Entfernung, desto stärker nähern sie sich einander an, bis sie sich im Fluchtpunkt vereinigen.

Bewegungsinduzierte Informationen

Bewegungsinduzierte Tiefeninformationen entstehen durch die eigene Bewegung. Verändert man seine Position parallel zu Objekten der Umgebung, gleiten nähere schneller vorbei als entferntere. Objekte am Horizont scheinen sich nur wenig oder gar nicht zu bewegen. Dieser Geschwindigkeitsunterschied wird als *Bewegungsparallaxe* bezeichnet. Eine weitere bewegungsinduzierte Information entsteht bei sich teilweise verdeckenden Flächen in unterschiedlicher Entfernung, sobald man sich auf sie zu oder von ihnen weg bewegt. Die hintere Fläche wird dabei von der vorderen zu- bzw. aufgedeckt. Bewegungsinduzierte Tiefeninformation wirken insbesondere an Kanten (Kaplan, 1969).

Stereoskopisches Sehen und Querdisparation

Anders als bei okulomotorischen, monokularen oder bewegungsinduzierten Informationen benötigt man für die Tiefeninformationen des stereoskopischen Sehens beide Augen. Aufgrund des horizontalen Abstandes werden zwei verschiedene Bilder aus unterschiedlicher Perspektive auf die jeweilige Netzhaut projiziert. Die entstehende retinale Abbildungsdifferenz wird als *Querdisparation* definiert und ist von der Vergenzstellung der Augen abhängig (Goldstein et al., 2002). Um die genauen Unterschiede in den Netzhautbildern beschreiben zu können, muss man die Funktionsweise der *korrespondierenden Netzhautpunkte* näher betrachten. Jeder Punkt auf der Netzhaut des einen Auges entspricht exakt einem Punkt auf der Netzhaut des anderen. Diese korrespondierenden Netzhautpunkte sind mit denselben Ortswerten im visuellen Cortex verbunden. Die Gesamtheit der Punkte, die bei einer festen Augenstellung auf korrespondierende Stellen der Netzhaut abgebildet werden, beschreiben den in Abb. 3-2 A dargestellten *Horoopter*. Der Horoopter ist eine gedachte Kreislinie, die durch den Fixationsort und die optischen Brennpunkte beider Augen führt. Punkte, die nicht auf dem Horoopter liegen, werden auch als *nichtkorrespondierende Netzhautpunkte* bezeichnet.

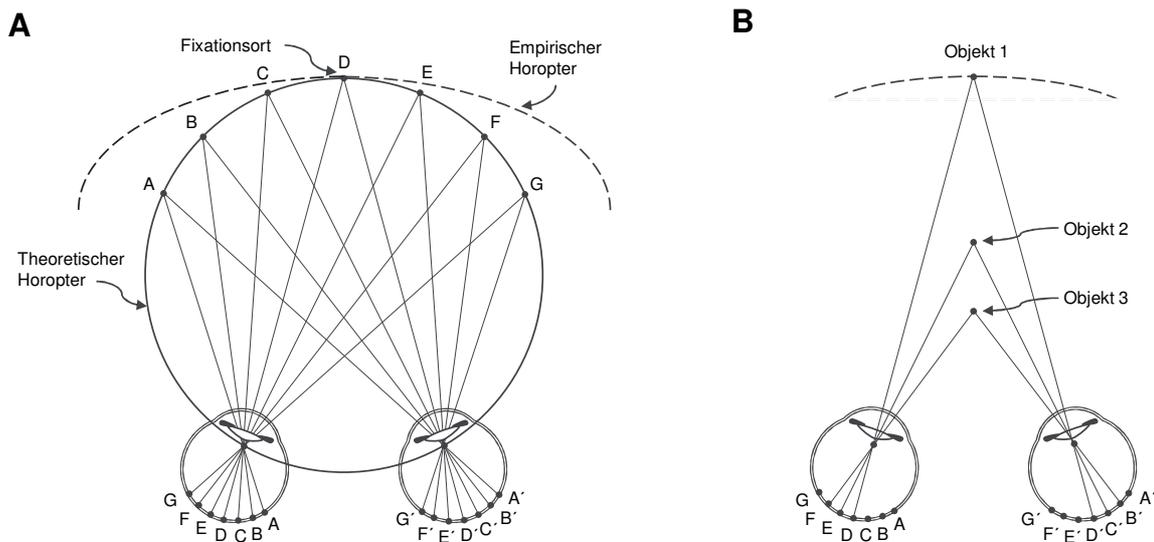


Abb. 3-2: Der Horopter (A). Die Objektpunkte A-G stellen die korrespondierenden Netzhautpunkte in beiden Augen dar. Der theoretische Horopter ist kreisförmig und auch als Vieth-Müller-Kreis bekannt. Der empirische Horopter (gestrichelte Linie) liegt hinter dem theoretischen Horopter (nach Palmer, 1999). Bei Objekten, die nicht auf dem Horopter liegen (B), fallen deren Abbildungen auf nichtkorrespondierenden Netzhautpunkte (nach Posner, 1977).

Wie Abb. 3-2 B zeigt, fällt Objekt 2 auf die nichtkorrespondierenden Punkte E und C'. Der korrespondierende Punkt zu E befindet sich bei E', weit entfernt von C'. Der zwischen E' und C' liegende *Querdisparationswinkel* variiert in Abhängigkeit der Objektentfernung zum Horopter. Da der Querdisparationswinkel von Objekt 3 größer als der von Objekt 2 ist, muss Objekt 3 weiter vom Horopter entfernt sein bzw. sich näher am Betrachter befinden als Objekt 2. Liegen Objekte vor dem Horopter, befinden sich ihre Abbildungen in Richtung äußerer Randbereich der Netzhaut. Die resultierende Querdisparation wird hierbei auch *gekreuzte Querdisparation* genannt. Liegen andererseits die Objekte hinter dem Horopter und dementsprechend deren Abbildungen in den nasalen Randbereichen der Netzhaut, spricht man von *ungekreuzter Querdisparation*.

Zusammenfassend lässt sich festhalten, dass räumliche Wahrnehmung auf unterschiedlichen Informationen basiert, und die dafür notwendigen Blickbewegungen, abgesehen von den monokularen Tiefeninformationen, binokular in verschiedenen Tiefenebenen erfolgen. Auch bei der Darbietung einer virtuellen dreidimensionalen Umgebung anhand stereoskopischer Visualisierungstechnologien müssen die binokularen Tiefeninformationen dem Betrachter zur Verfügung stehen.

3.2 Stereoskopische Visualisierung

Ähnlich wie bei der realen visuellen Wahrnehmung wird während einer stereoskopischen Visualisierung dem linken und rechten Auge ein verschiedenes, jeweils durch die Szene und dem Blickwinkel spezifiziertes Bild dargeboten. Wiedergabemethoden für die Visualisierung sind u.a. Stereoblick, örtliche Bildtrennung (Stereoskop, Head-Mounted Display), Anaglyphen, Polarisation, Shutterverfahren, Interferenzfilter, Autostereoskopie und Pulfrich-Effekt. Durch die verschiedenen Technologien und Verfahrensweisen in den Wiedergabemethoden entstehen Vor- und Nachteile für unterschiedliche Anwendungsszenarien in Bezug auf Präsentationsmedium, Zuschaueranzahl, Farbtreue, Geisterbilder, Tiefeneindruck und Notwendigkeit einer 3D-Brille. Während sich für Printmedien die Anaglyphentechnologie eignet, sind Polarisations-, Shutter- und Interferenzverfahren für den Einsatz am Computerbildschirm oder an Projektionsleinwänden praxisrelevant. Im Nachfolgenden werden die in den durchgeführten Studien eingesetzten Polarisations- und Shutter-Verfahren näher vorgestellt.

3.2.1 Polarisationsfilter

Bei einer stereoskopischen Bildtrennung durch Polarisationsfilter geht man von der physikalischen Eigenschaft des Lichtes aus, sich ähnlich den Wellen elektromagnetischer Schwingungen in alle Richtungen auszubreiten (Pietsch, 1957). Lässt man die Lichtstrahlen durch ein vorgeschaltetes Gitter nur in einer Schwingungsebene passieren, so werden alle Lichtstrahlen, deren Schwingungen quer zu den Gitterstäben verlaufen, zurückgehalten (Abb. 3-3 A). Mit Polarisationsfiltern können die zwischen zwei Polen schwingenden Lichtstrahlen je nach Anordnung der in den Filtern vorhandenen Gitterstäbe ausgerichtet bzw. polarisiert werden. Für die stereoskopische Visualisierung projiziert man die Bilder für linkes und rechtes Auge durch zwei unterschiedlich ausgerichtete Polarisationsfilter gleichzeitig auf dieselbe Fläche. Hier entsteht infolge der parallaktischen Unterschiede beider Augenbilder zunächst ein Doppelbild. Dieses wird durch zwei in ein Brillengestell gefasste Polarisationsfilter betrachtet, welche das Licht in denselben Schwingungsebenen wie die der Projektion vorgeschalteten Polfilter passieren lassen. Dadurch werden von jedem Auge nur noch die Lichtstrahlen des ihm bestimmten Bildes gesehen und die Lichtstrahlen des jeweils anderen Bildes blockiert. Während des Tragens der 3D-Brille wird aufgrund der Lichtfilterung die Helligkeit der projizierten Augenbilder verringert und die Szene vom Betrachter dunkler wahrgenommen als beim Betrachten ohne 3D-Brille.

3.2.2 Shutterverfahren

Bei der stereoskopischen Visualisierung mittels Shutterverfahren werden die Bilder für linkes und rechtes Auge aufeinanderfolgend projiziert. Die Bildtrennung erfolgt durch synchrones Öffnen bzw. Schließen der in eine Brille integrierten Blenden. Diese Blenden schalten mithilfe piezoelektrischer Elemente, mechanischer Bauteile oder Flüssigkeitskristalle zwischen durchlässig und undurchlässig um. Während die Bildinformationen für das linke Auge angezeigt werden, ist die zuständige Blende auf durchlässig geschaltet (Abb. 3-3 B). Im nächsten Schritt schließt die linke Blende, das rechte Augenbild wird dargeboten und die rechte Blende öffnet sich. Da die beiden Bilder abwechselnd dem jeweiligen Auge präsentiert werden, halbiert sich die Bildwiederholfrequenz. Für eine Darbietung von 25 Bildern pro Sekunde und Auge müssen dementsprechend 50 Bilder in einer Sekunde projiziert werden. Aufgrund der alternierenden Augenabdeckung verringert sich auch bei diesem Verfahren die einfallende Lichtmenge bzw. Helligkeit der wahrgenommenen Szene. Die Synchronisation zwischen Shutterbrille und Projektion erfolgt kabelgebunden oder kabellos über Funk- bzw. Infrarotlichtsignale.

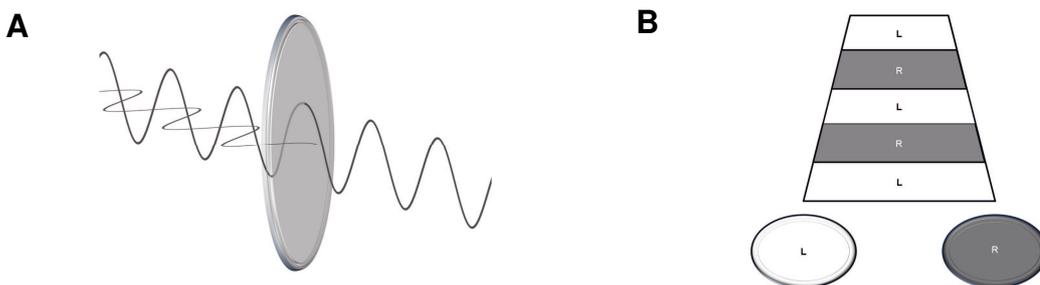


Abb. 3-3: Verfahren zur stereoskopischen Visualisierung: Schematische Darstellung der Wirkungsweise eines Polarisationsfilters (A) und zeitliche Synchronisation der abwechselnd sichtbaren Bilder für linkes und rechtes Auge im Shutterverfahren (B).

3.2.3 Stereobildgenerierung

Für die Berechnung eines Stereobildpaares der virtuellen Umgebung wird die Szene anhand zweier virtueller Kameras aus zwei unterschiedlichen Positionen heraus erstellt und dadurch die für das stereoskopische Sehen notwendige Querdisparation erzeugt. Dabei entsprechen die Kamerapositionen und -ausrichtungen denen für linkes und rechtes Auge. Die horizontale Verschiebung der virtuellen Kameras wird auch als *Stereobasis* und der horizontale Versatz eines Punktes in den beiden Bildern für linkes und rechtes Auge als *Bildschirmparallaxe* bezeichnet. Die Methoden zur Berechnung eines Stereobildpaares sind in der Literatur vielfach beschrieben (Jones, Lee, Holliman & Ezra, 2001; Kim & Sohn, 2010; Masaoka et al., 2006; Smith & Collar, 2012; Woods, Docherty & Koch, 1993), wobei jedoch immer der Augenabstand als Maß für die Stereobasis eingesetzt wurde. Untersuchungen von Dodgson

(2004) belegen, dass der Augenabstand der überwiegenden Mehrheit erwachsener Menschen zwischen 50 – 75 mm variiert. Bezieht man auch die Extrema der interokularen Distanzwerte ein, liegt die Variation aller erfassten Personen zwischen 45 – 80 mm. Der Durchschnittswert wird von Dodgson ähnlich vieler weiterer Literaturangaben mit 63 mm angegeben.

Hodges und Davis (1993) untersuchten die Zusammenhänge zwischen retinaler Disparität, Fixationsort, binokularer visueller Ausrichtung und Bildschirmparallaxe. Dabei verwendeten sie ein vereinfachtes stereogeometrisches Modell, bei dem linkes und rechtes Auge in einer Ebene parallel zur Projektionsfläche angeordnet sind. Bei virtuell dargestellten Objekten hinter der Projektionsfläche entsteht eine positiv (Abb. 3-4 A) und vor der Projektionsfläche eine negativ ausgerichtete Bildschirmparallaxe (Abb. 3-4 B).

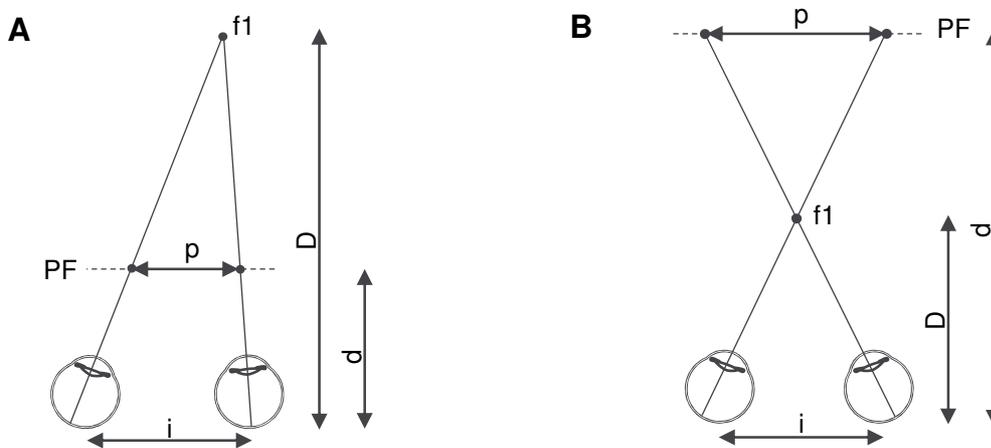


Abb. 3-4: Bildschirmparallaxe (nach Hodges & Davis, 1993): Der Fixationspunkt $f1$ befindet sich (A) hinter der Projektionsfläche PF bei ungekreuzter Parallaxe und (B) vor der Projektionsfläche bei gekreuzter Parallaxe.

Nach den Berechnungen der Autoren ergibt sich ein Bildschirmparallaxenwert p für einen Punkt $f1$, sofern dieser auf eine ebene Fläche im Abstand d bezüglich der Augenebene projiziert wird. Der modellierte Augenabstand wird mit i und die Entfernung zwischen $f1$ und der Augenebene mit D angegeben.

$$p = \frac{i(D - d)}{D} \quad (3.1)$$

Setzt man die Stereobasis mit dem individuellen Pupillenabstand gleich, erhält man bei einer angenommenen Pupillendistanz von 63 mm, einer Entfernung zur Projektionsfläche von 650 mm und einem Abstand zwischen Augen bzw. virtuellem Objekt von 300 mm eine Bildschirmparallaxe von -73,5 mm. Betrachten Versuchspersonen mit Pupillenabständen zwischen 50 - 70 mm (Dodgson, 2004) dasselbe virtuelle Objekt, ergeben sich variable Wahrnehmungsdistanzen von 263 - 328 mm (vgl. Abb. 3-5).

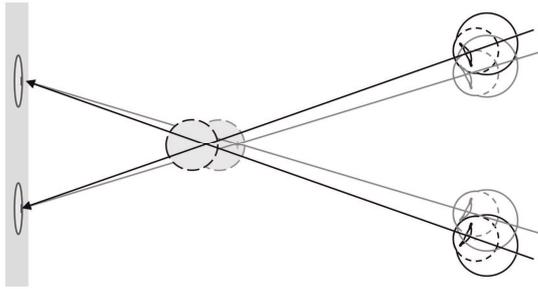


Abb. 3-5: Einfluss des Augenabstands auf die wahrgenommene Entfernung eines Objektes in einer stereoskopisch dargebotenen virtuellen Szene. Das gleiche Objekt erscheint mit größerem Augenabstand entfernter als bei einem kleineren Augenabstand.

Hierbei wird deutlich, dass Personen mit unterschiedlichem Augenabstand dasselbe visualisierte Bild unterschiedlich wahrnehmen. Diesen Zusammenhang haben auch Renner, Velichkovsky, Helmert und Stelzer (2013) nachgewiesen. In einer Untersuchung mit stereoskopisch projizierter VR-Umgebung konnten sie zeigen, dass die Bilderzeugung für linkes und rechtes Auge mit der Stereobasis des individuellen Pupillenabstands eine genauere Tiefenwahrnehmung bewirkt als eine davon abweichende. Versuchspersonen sollten in einem Experiment innerhalb einer fünfseitigen CAVE (Cave Automatic Virtual Environment) die Distanz eines virtuell dargebotenen Kegels mit dem real vorhandenen Kegelständer abgleichen. In mehreren Durchläufen wurde dabei die Stereobasis variiert. Renner et al. stellten fest, dass sich mit zunehmender Verfälschung der Stereobasis von der individuellen Pupillendistanz die Abweichung zwischen eingestelltem virtuellen und realen Objekt vergrößerte. Die Positionierung des virtuellen Kegels über den realen Ständer war mit angepasster Stereobasis am genauesten.

3.3 Wahrnehmungsprobleme in stereoskopischer Umgebung

Beim natürlichen Sehen wird die Umgebung anhand sakkadischer Blickbewegungen abgetastet bzw. der Blickort an einzelnen Orten stabilisiert, um Objekte fixieren und wahrzunehmen zu können. Hierbei sind Akkommodation und Vergenz an jeder beliebigen Blickposition im Raum auf denselben Ort ausgerichtet (Abb. 3-6 A). Bei stereoskopisch projizierten Szenen können technologisch bedingt Akkommodation und Vergenz in verschiedenen Tiefenebenen liegen (Abb. 3-6 B) und dabei zu Wahrnehmungsproblemen führen. Außerdem kommt es in einer VR-Umgebung zu perspektivischen Verzerrungen sobald der aktuelle Blickort nicht mit dem Fluchtpunkt der Szene übereinstimmt.

3.3.1 Visuelle Überforderung durch erzwungene Vergenz

Bei einer stereoskopischen Projektion müssen Akkommodation und Vergenz entkoppelt werden, da die Akkommodation stets auf die Projektionsfläche erfolgt, während sich die Vergenz in Abhängigkeit der wahrgenommenen Tiefe ändert (Yano, Ide, Mitsuhashi & Thwaites, 2002). In der Literatur wird diese Diskrepanz auch als *Erzwungene Vergenz* oder *Vergenz-Akkommodations-Konflikt* beschrieben (Emoto, Niida & Okano, 2005). Je näher sich das projizierte Objekt zum Betrachter hin und von der Projektionsebene weg befindet, umso schwieriger ist die Entkopplung. Der Vergenz-Akkommodations-Konflikt wird als Grund für *visuelle Überforderung* (Tauer, 2010) bis hin zu Schwindelgefühlen angenommen. Eine zunehmende Disparität bei konstanter Akkommodation verstärkt den Konflikt und kann nach Beobachtungen von Lambooij et al. (2009) zwei Reaktionen bewirken:

- 1- Die Vergenzstellung der Augen bleibt aufrechterhalten, jedoch driftet die Akkommodation in Richtung der durch die Vergenz empfundenen Tiefenebene ab. Das virtuelle Objekt weiterhin dreidimensional jedoch unscharf wahrgenommen.
- 2- Die Akkommodation in Projektionsebene bleibt aufrechterhalten, allerdings ist es nun nicht mehr möglich beide Augenbilder zu einem dreidimensionalen Objekt zu vereinen. Statt einem 3D-Objekt werden nun die beiden Einzelbilder für linkes und rechtes Auge getrennt als Doppelbild innerhalb der Projektionsebene gesehen.

Um den Vergenz-Akkommodations-Konflikt zu vermeiden, entwickelten Hoffman, Girshick, Akeley und Banks (2008) ein Volumendisplay, in dem anhand verschiedener Spiegel visuelle Informationen in unterschiedlichen Tiefenebenen sowohl mit korrektem Vergenzwinkel als auch in entsprechender Akkommodationsdistanz präsentiert werden können. Jedoch begrenzt die Anzahl an Spiegeln die möglichen Akkommodationsebenen. Des Weiteren erstellten u.a. Bando, Iijima und Yano (2012) einen Richtlinienkatalog zur Darbietung stereoskopischer Informationen, um der visuellen Überforderung entgegen zu wirken.

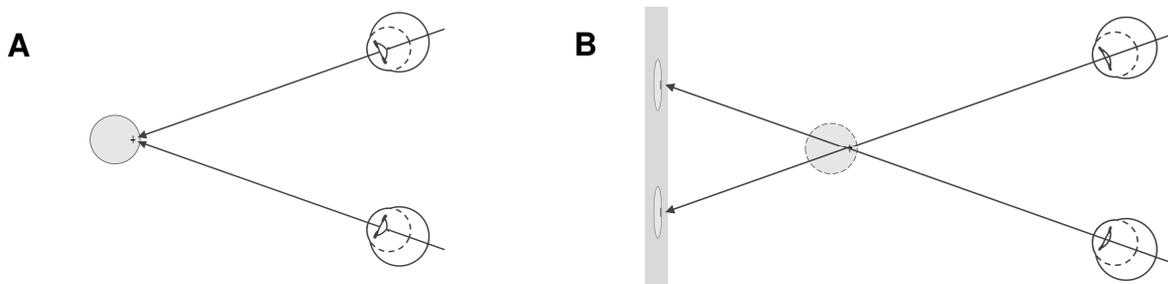


Abb. 3-6: Akkommodation und Vergenz in realer und stereoskopisch projizierter VR-Umgebung. Reale Umgebung (A): Akkommodation und Vergenz sind auf dieselbe Tiefe ausgerichtet. Virtuelle Umgebung (B): Akkommodation und Vergenz sind in unterschiedliche Tiefenebenen ausgerichtet. Hierbei kommt es zu einem Vergenz-Akkommodations-Konflikt.

3.3.2 Perspektivische Verzerrung

In realer Umgebung richten sich die Augen auf zu fixierende Objekte der Umgebung ständig neu aus. Damit verändern sich mit den Fixationsorten auch der Fluchtpunkt der Szene und damit die egozentrische Perspektive. In einer virtuellen 3D-Umgebung wird jedoch der Fluchtpunkt fest vorgegeben und meist in das Zentrum der Projektion gelegt sowie alle Objekte der Szene daran angepasst. Dies führt zu geometrischen Verzerrungen der nicht im Fluchtpunkt liegenden Objekte und damit auch zu einer unrealistischen Wahrnehmung (Franke et al., 2008). Je weiter ein Objekt vom Fluchtpunkt entfernt dargestellt wird, desto stärker wirken sich diese Verzerrungen aus (Abb. 3-7). Für eine vergleichbare Darstellung zur Realität müssen in stereoskopisch projizierten Umgebungen perspektivische Verzerrungen vermieden bzw. Korrekturmechanismen angewandt werden.

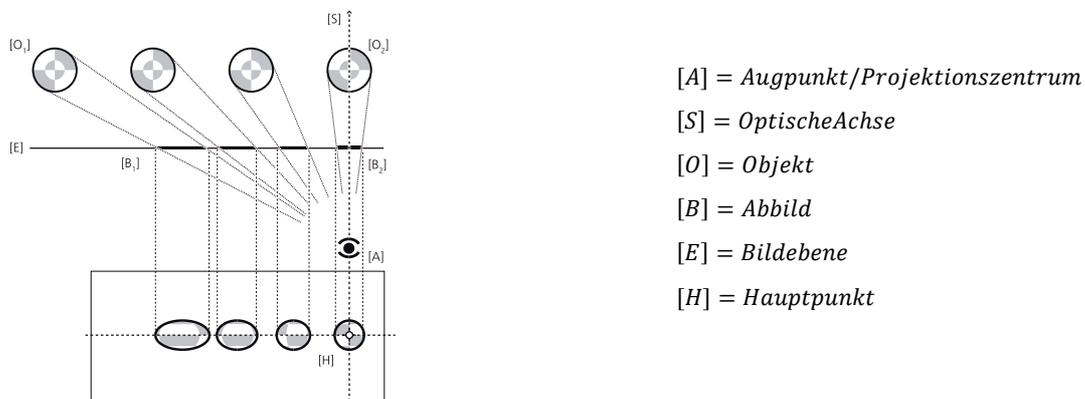


Abb. 3-7: Projektion von Objekten auf eine Bildebene (aus Zavesky, 2012). Das Bild wird nach Abbildungsprinzip der computergrafischen Kamera generiert, wobei alle Objekte gleich behandelt werden. Dabei entstehen projektionsbedingte Veränderungen der Proportionen. Diese wirken sich besonders auf am Rand positionierte Objekte aus und werden durch Verringerung des Abstandes zwischen Modell und Bildebene verstärkt.

4 Blickbewegungsmessung

Zur Messung von Blickbewegungen gibt es in der Literatur vielfältig beschriebene Methoden, (vgl. Joos et al., 2003). Die subjektive Erfassung von Augenbewegungen durch die Erzeugung retinaler Nachbilder oder die direkte Beobachtung der Augen sind dabei einfach durchzuführende Verfahren, lassen jedoch nur wenig Aussagen über örtliche und zeitliche Parameter zu. Sie finden überwiegend in der Gleichgewichtsforschung und in der Medizin bei der Überprüfung von Reflexen Anwendung. Wesentlich genauer aber auch technisch aufwendiger sind Methoden, bei denen elektrische Spannungen gemessen, digitalisiert und ausgewertet werden. Beispielsweise erfolgt bei einem *Elektrookulogramm* die Spannungsmessung mithilfe von Oberflächen Elektroden nahe am Auge, wobei durch horizontale und vertikale Augenbewegungen unterschiedliche Spannungswerte entstehen. Im Gegensatz dazu wird beim *Searsh Coil* Verfahren eine Kontaktlinse direkt auf dem Auge positioniert. Auf der Linse sind mehrere Spulen befestigt, die durch ein magnetisches Wechselfeld induzierte Spannungen messen. Je nach Ausrichtung des Auges und damit Lage der Kontaktlinse im magnetischen Feld verändern sich auch die induzierten Spannungswerte. Allerdings erfordern diese Methoden einen hohen Vorbereitungsaufwand und werden bei längeren Messungen als unangenehm empfunden.

Moderne Blickbewegungsmessungen arbeiten videobasiert. *Eyetracking*-Kameras zeichnen ein (monokular) oder beide (binokular) Augen auf. Dabei sind die Kameras bei *Head-Mounted Eyetrackern* direkt am Kopf angebracht oder befinden sich bei *Remote Eyetrackern* vor dem Probanden. Anhand der Augenbilder können die charakteristischen Merkmale durch rechnergestützte Bildverarbeitungsprozesse bestimmt und daraus die Blickrichtung sowie der Blickort abgeleitet werden. Videobasiertes Eyetracking ist mit vergleichsweise wenig Vorbereitungsaufwand verbunden und kann bei Remote Systemen auch über einen längeren Zeitraum durchgeführt werden.

Dieses Kapitel beschreibt die für eine videobasierte Blickbewegungsmessung erforderlichen Parameter des menschlichen Auges, ein Verfahren zur Berechnung der Blickrichtung sowie einen Algorithmus zur Blickortbestimmung. Anschließend werden aus der Forschung bekannte Versuche einer dreidimensionalen Blickbewegungsmessung vorgestellt, dabei aufgezeigte Probleme diskutiert und die technologischen Herausforderungen an eine räumliche Methode abgeleitet.

4.1 Parameter des menschlichen Auges

Um bei einer Blickbewegungsmessung den optischen Strahlengang der Blickachse bestimmen zu können, ist ein geometrisches Modell des Auges erforderlich. Abb. 4-1 zeigt die wichtigsten Schlüsselcharakteristiken. Die *visuelle Achse* des Auges verläuft als gedachte Linie vom Zentrum der *Fovea* durch den Mittelpunkt der *Cornea-Kugel*, einer durch die Krümmung der *Cornea* modellierten Kugel. Die *optische Achse* hingegen wird als Symmetrieachse des Auges definiert. Da der Ort der Fovea zur optischen Achse nach außen hin verschoben liegt, ist die visuelle Achse gegenüber der optischen nasal ausgerichtet. Der Winkel zwischen beiden Achsen wird in der Ophthalmologie als *Winkel Alpha* bezeichnet und in der Literatur mit bis zu $5,2^\circ$ horizontal bzw. bis zu $1,4^\circ$ vertikal angegeben (Holladay, 2007). Der Abstand zwischen Pupillenzentrum und visueller Achse im Augenvideobild ist durch den *Winkel Kappa* definiert. Kappa beträgt ca. $2,6^\circ$ horizontal bzw. $0,6^\circ$ vertikal. Die Oberfläche der Cornea ist annähernd sphärisch und etwas abflachend in Richtung des äußeren Randes gekrümmt. Dabei liegt der typische Krümmungsradius zwischen $7,7 \pm 2,0$ mm (LC Technologies, 2002). Die Cornea-Kugel ist kleiner als der Augapfel und ragt aus dem Augapfel um ca. 1,5 mm heraus.

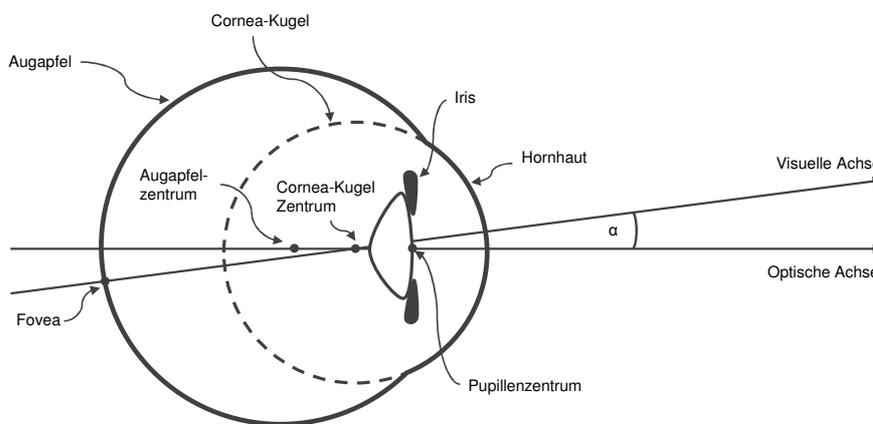


Abb. 4-1: Schematische Darstellung (Draufsicht) des rechten Auges mit charakteristischen Merkmalen zur Messung der Blickrichtung (nach LC Technologies, 2002).

4.2 PCCR-Methode

Bei der *Pupil Center Corneal Reflection (PCCR)-Methode* werden das Zentrum der Pupille und ein optischer Reflexionspunkt an der Cornea-Oberfläche im Videokamerabild bestimmt (Mason, 1969; Young & Sheena, 1975). Zur Detektion der Pupille im Videokamerabild gibt es zwei unterschiedliche Herangehensweisen. Wird das Auge mit einer Lichtquelle beleuchtet, die sich nicht in der optischen Kameraachse befindet, erscheint die Pupille im Kamerabild dunkel

(„*dark pupil*“), siehe Abb. 4-2 A. Dementsprechend durchsuchen Bildverarbeitungsprozesse das Videokamerabild nach einer dunklen Ellipse. Hierbei kann es zu Komplikationen kommen, sobald mehrere dunkle Bereiche z.B. durch Makeup, Hautverfärbungen oder Brillengestelle im Kamerabild vorhanden sind oder die Messung in dunkler Umgebung stattfindet (Holmqvist et al., 2011).

Die zweite Methode basiert auf dem Prinzip der Infrarotlichtmessung. Beleuchtet man das Auge mit einer sich nah oder direkt in der optischen Kameraachse befindlichen Infrarotlichtquelle, wird dieses durch die Netzhaut reflektiert und tritt durch die Pupille wieder aus. Im Videobild einer infrarotlichtsensitiven Kamera erscheint die Pupille hell („*bright pupil*“) und grenzt sich deutlich von der dunkleren Iris ab, siehe Abb. 4-2 B. Da bei dieser Methode nach einer im Videokamerabild weiß dargestellten Pupille gesucht wird, können Messungen unter sehr grellen Lichtverhältnissen zu Fehlerdetektionen führen. Morimoto, Koons, Amir und Flickner (2000) setzten beide Verfahren alternierend ein. Dabei wurden zwei Infrarotlicht-Ringe bestehend aus jeweils 8 LED mit dem Takt der Kamerabildrate zeitlich synchronisiert. Der innere LED-Ring befand sich in der Nähe der optischen Kameraachse, der zweite entfernt davon. Dadurch wurden die Pupillen in einem Bild hell und im nächsten dunkel abgebildet. Anhand der Differenz dieser beiden Bilder ließ sich die Pupille detektieren.



Abb. 4-2: Videokamerabilder nach dem *dark-pupil* Effekt (A) und dem *bright-pupil* Effekt (B). Unterhalb der Pupille ist der Cornea-Reflex als punktförmige Reflexion der Infrarotlichtquelle an der Cornea-Oberfläche zu erkennen.

Zur Erzeugung eines optischen Reflexionspunktes an der Cornea-Oberfläche beleuchtet man das Auge mit einer punktförmigen Infrarotlichtquelle. Dabei wird ein Teil des Lichtes an der Oberfläche der Cornea reflektiert und im Videokamerabild als heller Glanzpunkt bzw. *Glint* sichtbar. Die der PCCR Methode zugrunde liegende Theorie basiert auf der Annahme, dass die Richtung des Blickorts in Abhängigkeit zum Vektor zwischen Cornea-Reflex und Zentrum der Pupille im Videokamerabild steht. Dieser Vektor wird als *Glint-Pupil Vektor* bezeichnet (Abb. 4-3). Blickt eine Person direkt in die Kamera, liegen Cornea-Reflex und Mittelpunkt der Pupille sehr dicht beieinander. Auch wenn sich die Infrarot-LED direkt im Zentrum der Kameralinse befindet, überlagern sich die beiden Punkte nicht, denn die Fovea liegt typischerweise auf der visuellen statt optischen Achse des Auges. Verschiebt die Person nun den eigenen Blickort durch Rotieren des Auges von der Kamera ausgehend beispielsweise

nach oben, wandert auch das Zentrum der Pupille in vertikale Richtung und entfernt sich zunehmend vom Cornea-Reflex. Die horizontalen und vertikalen Anteile des Blickwinkels im Verhältnis zur Achse, die durch das Zentrum der Kameralinse und Cornea-Kugel geht, werden anhand des Vektorabstands zwischen cornealen Reflexpunkt und Pupillenmittelpunkt im Kamerabild berechnet.

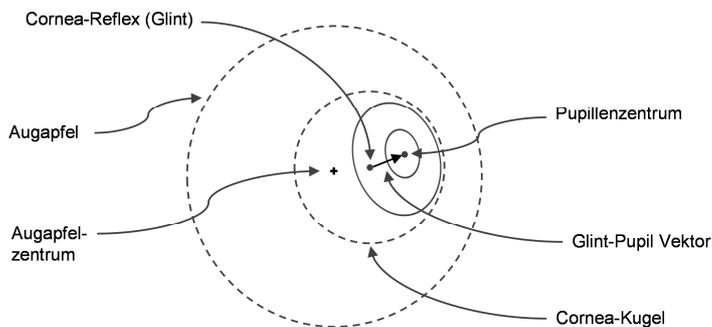


Abb. 4-3: Parameter der PCCR-Methode mit dem Glint-Pupil Vektor zur Berechnung der Blickrichtung.

Der Glint-Pupil Vektor hängt nicht nur von der Blickrichtung sondern auch den physiologischen Eigenschaften des Auges wie beispielsweise der Hornhautkrümmung ab. Deshalb muss das Eyetracking-System vor Beginn der Blickbewegungsmessung auf die jeweils zu messende Person kalibriert werden. Während der Kalibrierung wird ein mit den Augen zu fixierendes Objekt auf einer Projektionsfläche (Monitor, Leinwand, usw.) an definierten Positionen dargeboten und die die zugehörigen Parameter des Glint-Pupil Vektors erfasst.

4.3 Algorithmus zur Blickortbestimmung

Die exakte Berechnung der Blickposition erfolgt nach den optimierten Algorithmen der Eyetracker-Hersteller. In den folgenden Ausführungen wird die Berechnung des Blickvektors am Beispiel des *Eyegaze Analysis Systems* (LC Technologies, 2002) erläutert.

Lineares Modell

Der Koordinatenursprung $(0,0)$ des Videokamerabildes befindet sich in der linken oberen Ecke. Die Pixel-Koordinaten werden mit i bzw. j bezeichnet, wobei der horizontale Pixelindex i nach rechts hin und der vertikale Pixelindex j nach unten hin positiv bewertet ist. Die Parameter di und dj des Glint-Pupil Vektors sind variabel und repräsentieren die Ausrichtung des Auges.

$$di = i_{pupil-cent} - i_{glint} \quad (4.1)$$

$$dj = j_{pupil-cent} - j_{glint} \quad (4.2)$$

Als eine erste Annäherung, kann der Blickort als lineares Gleichungspaar modelliert werden:

$$x_{gaze} = A_0 + A_i * di \quad (mm) \quad (4.3)$$

$$y_{gaze} = B_0 + B_j * dj \quad (mm) \quad (4.4)$$

A und B sind Konstanten, die von den geometrischen Größen der jeweiligen Kamera wie Linse, Sensor oder Abstand zwischen Sichtfeld und Kamera abhängen. Die Berechnung des Glint-Pupil Vektors (di, dj) bezieht sich auf die Ausrichtung der optischen Achse des Auges. Zur Bestimmung des Blickvektors, muss die optische Achse um den Winkel κ gedreht werden. Die Konstanten A_0 und B_0 berücksichtigen den Winkelversatz, setzen jedoch feste x- und y-Offsets zwischen der Augenebene und der x-y-Ebene der Kamera voraus. So lange sich das Auge in einem fest definierten Abstand zur Blickebene befindet, kann der Winkelversatz als annähernd konstant betrachtet werden. Die linearen Ausdrücke A_i und B_j in den oberen Gleichungen repräsentieren die zeitabhängigen Rotationen des Auges. Für Blickwinkel innerhalb von 20° , ausgehend von der optischen Kameraachse, verhalten sich Veränderungen des Blickorts annähernd linear zum Glint-Pupil Vektor (di, dj) .

Erweiterung des linearen Modells

Beim *Eyegaze Analysis System* werden die Blickbewegungen in Bezug auf den visuellen Inhalt eines Computermonitors bestimmt. Dazu befinden sich die Eyetracking-Kameras am unteren Bildschirmrand. Die im vorherigen Abschnitt erläuterten linearen Gleichungen zur Ermittlung des Blickorts sind für diesen Eyetracker unzureichend, da die Z-Achse des Monitors zur Z-Achse der Kameras geneigt ist. Außerdem wird für eine Genauigkeitsverbesserung berücksichtigt, dass die Oberfläche der Cornea des Auges zu den Rändern hin abflacht. Um diese Nichtlinearitäten in die Berechnungen einzubeziehen, wurden die Gleichungen in den Verarbeitungsprozessen des *Eyegaze Analysis Systems* wie folgt erweitert:

$$x_{gaze} = A_0 + A_i * di + A_j * dj + A_{ij} * di * dj \quad (mm) \quad (4.5)$$

$$y_{gaze} = B_0 + B_j * dj + B_i * di + B_{jj} * dj^2 \quad (mm) \quad (4.6)$$

Die Terme $(A_j * dj)$ in x_{gaze} und $(B_i * di)$ in y_{gaze} berücksichtigen den Winkelversatz zwischen den horizontalen Achsen von Kamera und Monitor. Beachtet man diesen Versatz nicht, entstehen selbst bei rein horizontalen Blickbewegungen auch vertikale Veränderungen des berechneten Blickortes. Die Ausdrücke $(A_{ij} * di * dj)$ in x_{gaze} und $(B_{jj} * dj^2)$ in y_{gaze} wurden ergänzt, um di und dj bezüglich der aktuellen Blickorthöhe zu skalieren. Dadurch soll der unterschiedliche Einfluss von di auf x_{gaze} und dj auf y_{gaze} in Abhängigkeit der vertikalen Komponente des Blickvektors (dj) in die Berechnung einbezogen werden. Innerhalb des Bereichs eines Computermonitors gleichen diese beiden Ausdrücke die meisten Nichtlinearitäten in Bezug auf die Neigung des Monitors aus. Der Term $(B_{jj} * dj^2)$ in y_{gaze} berücksichtigt außerdem das Abflachen der Cornea-Oberfläche zu den Rändern hin, wie es für das menschliche Auge typisch ist. Die Bestimmung der Koeffizienten A_0, A_i, A_j, A_{ij} bzw. B_0, B_i, B_j, B_{ij} erfolgt während der Kalibrierung. Anhand des in der Mathematik als *Kleinste-Quadrate-Schätzung* bezeichneten Verfahrens werden die Werte derart ermittelt, dass diese optimal die gemessenen Glint-Pupil Vektoren auf die dargebotenen Kalibrierpunkte bzw. bekannten Blickorte abbilden.

4.4 3D-Verfahren

Für eine dreidimensionale Analyse der visuellen Wahrnehmung wird der Blickort bzw. die Blickbewegung räumlich erfasst und den Objekten einer dreidimensionalen Szene zugeordnet. In realer Umgebung ist dies komplexer als in virtueller, da sowohl die Position des sich frei bewegenden Betrachters als auch die geometrische Ausrichtung aller in der Umgebung vorhandenen Objekte für eine korrekte Blickortzuweisung bekannt sein müssen.

4.4.1 Eyetracking in realer Umgebung

Rötting, Göbel und Springer (1999) kombinierten ein Head-Mounted Eyetracking-System mit einem Headtracker, welcher Bewegungen in 6 Freiheitsgraden messen konnte. Zusätzlich wurde am Kopf des Probanden eine nach vorn ausgerichtete Szenekamera angebracht, um dessen Blickfeld aufzuzeichnen. Die Forscher versuchten Blickbewegungen in natürlicher Umgebung zu erfassen und automatisch den Objekten zuzuordnen. Dabei sollte sich der Proband in der realen Welt frei bewegen können. Die Blickortzuweisung zu den Objekten erfolgte nach der Datenaufzeichnungsphase in einem mehrstufigen Auswertungsprozess. Als erstes wurden die Objektkonturen im Szenekameravideo aus mindestens zwei verschiedenen

Kameraperspektiven identifiziert, daraus die Tiefeninformation extrahiert und ein geometrisches Modell der Objekte im Raum erstellt. Die aufgezeichneten Blickbewegungsdaten bezogen sich zunächst auf das Koordinatensystem des sich mitbewegenden Eyetrackers, konnten jedoch anhand der Headtracking-Daten in Weltkoordinaten transformiert werden. Durch das Headtracking System waren Position und Ausrichtung der Szenekamera bekannt und die Voraussetzungen für die Zuordnung der Blickpositionen zu den Objekten im Szenevideo bzw. räumlichen Modell gegeben. In einem weiteren Schritt erfolgte eine Fixationserkennung, wobei sich die Definition einer Fixation in Röttings Setting auf das „relative Stillstehen des Auges innerhalb einer Zeitperiode“ bezieht. Dafür wird sowohl ein Beschleunigungs- als auch Zeitkriterium vom Anwender festgelegt. Das Beschleunigungskriterium ist erfüllt, solange sich das Auge nicht schneller als der maximale Beschleunigungswert bewegt und das Zeitkriterium, sobald das Beschleunigungskriterium für eine minimale Dauer zutrifft. Nach welcher Berechnung der dreidimensionale Blickort dem realen Objekten zugeordnet bzw. der Tiefenparameter ermittelt wurde, geht aus Röttings Beschreibung nicht hervor. Fest steht jedoch, dass für die Anwendung dieser Methode eine statische Szene vorausgesetzt wird. Bei einer dynamischen Veränderung der Objektposition, wie beispielsweise im Straßenverkehr, wäre sowohl die Blickortzuweisung als auch Fixationserkennung nicht mehr möglich.

Einen ähnlichen Ansatz verfolgten Johnson, Liu, Thomas und Spencer (2007). Sie koppelten einen Head-Mounted Eyetracker mit einem dreidimensionalen Motion-Tracking-System und entwickelten einen Kalibrier-Algorithmus, um Blickbewegungen bei uneingeschränkter Kopfbewegungsfreiheit auf einer zweidimensionalen Präsentationsoberfläche zu messen. In einem Experiment trugen die Probanden eine modifizierte Baseball-Kappe, an der sowohl ein monokularer Eyetracker als auch ein optisch-elektronisches Bewegungsanalysesystem angebracht waren. In einem zweistufigen Kalibrierprozess wurde zunächst der Eyetracker auf die Augen und anschließend die Augen- in Bezug zu den Kopfbewegungen kalibriert. Danach betrachteten die Versuchspersonen projiziertes Stimulusmaterial auf einer Tischplatte. Anhand der Bezugsdaten zwischen Eyetracker, Bewegungsanalysesystem und Projektionsfläche konnten die Blickbewegungen aus dem zweidimensionalen Augenvideo mit den Kopfbewegungen abgeglichen und der Bezug zu den auf dem Tisch abgebildeten visuellen Inhalten hergestellt werden. Allerdings wurde dabei der Schnittpunkt zwischen Blickachse und Präsentationsoberfläche ermittelt. Eine räumliche Blickort- oder Blickbewegungsbestimmung in unterschiedlichen Tiefenebenen ist mit dieser Methode nicht durchführbar.

In Experimenten, in denen Positionsänderungen des Betrachters oder der Umgebungsobjekte nicht erfasst werden können und die Voraussetzungen für die Erstellung eines geometrischen

Modells nicht gegeben sind, kommen *Mobile Eyetracker* zum Einsatz (Hayhoe & Ballard, 2005; Land, 2006). Diese sind ähnlich dem vorhergehenden Setting mit einer Szenekamera jedoch ohne Headtracker ausgestattet. Während der Aufzeichnung entstehen ein (monokulares) oder zwei (binokulare) Augenvideos und ein zusätzliches Video der Umgebung aus Sicht des Betrachters. Der aus den Augenvideobildern ermittelte Blickort wird auf die 2D-Koordinaten der Szenekamera abgebildet und im Szenevideo beispielsweise als farbiger Punkt dargestellt. Die Zuordnung des Blickortes zu den Objekten muss vom Auswerter manuell durchgeführt werden (Pfeiffer, 2012). Zwar gibt es bereits Ansätze für automatische Objekterkennungen in Videos (Nascimento & Marques, 2006), allerdings kommt es insbesondere bei sich bewegenden oder mehreren sich gegenseitig verdeckenden Objekten zu Mehrdeutigkeiten und Fehlerkennungen, so dass diese in der Praxis noch keine Anwendung finden.

Ein weiteres Problem dieser Aufzeichnungsmethode ist der entstehende *Parallaxeneffekt* (Evans, Jacobs, Tarduno & Pelz, 2012; Mansouryar, Steil, Sugano & Bulling, 2016). Dadurch, dass sich die Szenekamera räumlich versetzt zu den Augen befindet, unterscheidet sich das Szenevideo von der Wahrnehmung des Betrachters. Eine Fixationserkennung nach Röttings Methode ist in diesem Verfahren nicht eindeutig, da sich trotz Beibehaltung des Fixationsortes bei Kopfbewegungen der Blickort im Koordinatensystem des Eyetrackers ändert. Ohne die zusätzlichen Bezugsdaten eines Headtrackers und Transformation der Positionsdaten in das Weltkoordinatensystem ist eine Überprüfung des Beschleunigungs- und Zeitkriteriums für eine Fixationserkennung nicht möglich.

Dieses Problem versuchten Pfeiffer, Renner und Pfeiffer-Leßmann (2016) zu lösen, indem sie zwischen den Objekten der realen Szene Marker mit unterschiedlichen Mustern platzierten und diese im Szenevideo des mobilen Eyetrackers identifizierten. Pfeiffer et al. konnten die Marker im Szenevideo in ihrer Position sowie Lage bestimmen und dadurch auf die Positionierung der am Kopf angebrachten Szenekamera rückschließen. Aus der Kopfposition sowie den Blickdaten eines mobilen Eyetracking Systems berechneten die Forscher *Blickstrahlen* im Raum. In einem Experiment wurden diese Blickstrahlen durch eine Kollisionserkennung den realen Objekten der Szene zugeordnet. Dafür wurden alle Marker- und Objekt-Positionen in einem geometrischen Modell abgebildet. Pfeiffer et al. berichteten die typischen Probleme der Kollisionserkennung. Bei nahe aneinander liegenden Objekten war eine Zuordnung nicht eindeutig möglich. Außerdem war eine Kollision zwingend notwendig, um ein Objekt als mit den Augen „fixiert“ zu definieren. Eine dreidimensionale Blickortberechnung oder räumliche Blickbewegungserkennung fand dabei nicht statt.

4.4.2 Stereoprojektion und Eyetracking

Das Problem der Objektpositionsbestimmung gibt es in virtuellen 3D-Umgebungen nicht. Die Geometrie der Szene und der darin enthaltenen Objekte ist selbst bei dynamischen Veränderungen zu jeder Zeit bekannt. Allerdings muss für eine virtuelle 3D-Projektion die stereoskopische Visualisierung mit einer Blickbewegungsmessung kombiniert werden. Erste Versuche, Blickbewegungen in einer stereoskopisch dargestellten VR-Umgebung zu messen, sind seit der Jahrtausendwende bekannt. Je nach Kombination aus Eyetracking-System und Stereoprojektion wurden von den Autoren speziell auf das Setting ausgerichtete Konzepte der dreidimensionalen Blickortbestimmung und Objektzuweisung entwickelt und angewandt. Die für diese Arbeit relevanten Ansätze werden im Folgenden vorgestellt.

Head-Mounted Display

Duchowski et al. (2000) integrierten eine binokulare Blickbewegungsmessung in ein *Head-Mounted Display* (HMD). Darin wurden anhand zweier Projektionsflächen die Bilder für linkes und rechtes Auge örtlich getrennt projiziert, wodurch ein räumlicher Eindruck entstand. Gleichzeitig konnte ein Eyetracker den Blickort der Augen auf dem jeweiligen Display bestimmen. In einem Versuch wurde Probanden der Innenraum eines Flugzeugladeraums zur Durchführung eines visuellen Inspektionstrainings virtuell dargeboten. Die Visualisierung der Laderauminnenwand erfolgte anhand einer realitätsnahen Oberflächentextur mit künstlich generierten Defekten, die während der Inspektion von den Versuchspersonen erkannt werden sollten. Um sich in der VR-Umgebung frei bewegen zu können, erfasste ein Headtracking-System sowohl die Position als auch Ausrichtung des Kopfes im dreidimensionalen Raum. Damit war es bei Bewegungen des Probanden möglich, einerseits die Projektion der VR und der darin wahrgenommene Szene neu auszurichten und andererseits die Blickposition der in der VR-Umgebung dargestellten Objekte zuzuordnen. Für die Bestimmung des ‚*Point of Regard*‘ (POR) fand zunächst eine Berechnung des dreidimensionalen Blickortes statt.

Bei dem verwendeten stereogeometrischen Modell (Abb. 4-4) gingen Duchowski et al. davon aus, dass beide Blickachsen in einer horizontalen Ebene liegen und ermittelten deren Schnittpunkt. Die Augenpositionen im HMD definierten jeweils den Blickachsenursprung. Für eine Richtungsbestimmung benötigte man einen zweiten Bezugspunkt auf jeder Blickachse und bezog die Ausgabewerte der integrierten Blickbewegungsmessung ein. Der Eyetracker berechnete die Blickorte der Augen auf dem entsprechenden Display, der zweidimensionalen Messebene rechtwinklig zum Auge, und gab diese in Pixelkoordinaten aus. Nach einer Koordinatentransformation in die VR-Umgebung konnte die Lage der Blickachsen und deren Schnittpunkt ermittelt werden.

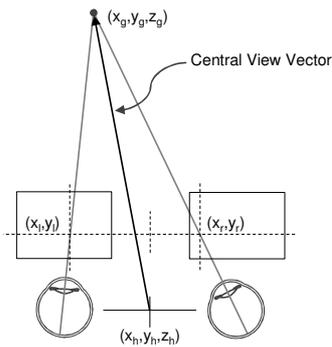


Abb. 4-4: Stereogeometrisches Modell zur dreidimensionalen Blickortbestimmung in einem HMD nach Duchowski et al. (2000). Die Blickachsen sind durch die Positionen der Augen bzw. den gemessenen Blickorten auf den Displays definiert und schneiden sich in derselben horizontalen Ebene. Der Blickachsenschnittpunkt gibt den dreidimensionalen Blickort an.

Duchowski et al. stellten allerdings fest, dass die berechneten 3D-Blickorte nicht immer die Textur der Laderauminnenwand trafen. Um den POR einem Texturfehler zuordnen zu können, musste jedoch eine Ortsbeziehung zwischen Blickort und Textur hergestellt werden. Daraufhin führten sie in einem weiteren Schritt eine Kollisionserkennung durch. Dabei wurde ein Strahl in die Szene gelegt und geprüft, mit welchen Objekten dieser kollidierte. Die Wissenschaftler positionierten den Ursprung ihres Kollisionsstrahls in die dreidimensionale Mitte beider Augen und nutzten zur Richtungsbestimmung den vorher errechneten 3D-Blickachsenschnittpunkt. Durch die Erkennung einer Kollision dieses *Central View Vectors* mit den Polygonen der Textur konnte nun überprüft werden, ob ein Defekt gesehen wurde oder nicht.

Wheatstone Stereoskop

Duchowski, Pelfrey, House und Wang (2011) kombinierten ein *Wheatstone Stereoskop* zur örtlichen Bildtrennung mit einer Blickbewegungsmessung. Ziel war es hierbei die Tiefeninformationen des Blickortes während der Projektion virtueller 3D-Objekte zu bestimmen. Bei dieser Technologie wurden über ein Spiegelsystem die Bilder zweier Monitore getrennt auf linkes und rechtes Auge projiziert. Mit einer Auflösung von jeweils 3.840 x 2.400 px, einer Abmessung von 480 x 300 mm und einem Betrachtungsabstand von 860 mm zu den Augen wurde ein visueller Winkel von 30" pro Pixel erreicht. Dieser Bereich war nach Duchowskis Aussagen mit der fovealen Rezeptorgröße vergleichbar und ausreichend, um die vom menschlichen Auge kleinste wahrnehmbare Ortsfrequenz bzw. das feinste erkennbare Muster darzustellen. Unterhalb jedes Monitors befand sich je eine Eyetracking-Kamera, welche ebenfalls über Spiegel das entsprechende Auge monokular aufzeichnete. Ein proprietärer Algorithmus des Herstellers gab neben den xy -Koordinaten auch die Tiefeninformation des Blickortes und den durch die Blickachsen einschließenden

Vergenzwinkel aus. Details zur Berechnung dieser Parameter wurden nicht veröffentlicht. In ihren Untersuchungen verglichen Duchowski et al. die z-Koordinate des Blickortes beim Betrachten einer auf 4 unterschiedliche Arten dargestellten Szene: ohne Stereo- oder Bewegungsparallaxe, ohne Stereo- jedoch mit Bewegungsparallaxe, mit Stereo- aber ohne Bewegungsparallaxe und schließlich mit Stereo- und Bewegungsparallaxe. Für die Erstellung der Stimuli mit Stereoparallaxe wurde eine interokulare Distanz von 63 mm angenommen. In den Ergebnissen konnte ein signifikanter Effekt der Tiefeninformation in beiden stereoskopischen Fällen jedoch in keinem der monoskopischen Darstellungen nachgewiesen werden. Duchowski et al. bemerkten allerdings ein starkes Rauschen in den Vergenzdaten und setzten daraufhin einen *Butterworth-Filter* zur Glättung der Daten ein. Ein Grund für diese Schwankungen wurde von den Autoren nicht angegeben.

Anaglyph-Verfahren

Essig, Pomplun und Ritter (2006) projizierten in ihren Untersuchungen stereoskopische Bildinhalte nach dem *Anaglyph-Verfahren*. Dabei bestand die Gesamtdarstellung aus zwei überlagerten Einzelbildern für linkes und rechtes Auge in den Komplementärfarben. Die Bildtrennung erfolgte anhand zweier in eine Brille eingefasster Farbfilter. Bei der Wahrnehmung durch die Anaglyph-Brille sperrte der Farbfilter die für das andere Auge bestimmten Informationen und ließ gleichzeitig das eigene Bild passieren. Essig et al. entwickelten zur Blickortbestimmung eine eigene Kalibrierungsprozedur in Verbindung mit einem künstlichen neuronalen Netz, der *Parametrized Self-Organizing Map* (PSOM). Die Motivation für die Verwendung eines künstlichen neuronalen Netzes stammte aus vorangegangenen Studien von Pomplun, Velichkovsky und Ritter (1994), in denen eine ähnliche Methode zur Reduzierung der Ungenauigkeit in zweidimensionalen Eyetracking-Messungen entwickelt wurde.

Bevor das eigentliche Experiment begann, wurde eine blickkontingente Kalibrierung durchgeführt. Dabei fixierte die Versuchsperson eine an unterschiedlichen Bildschirmpositionen dargebotene Markierung. Das künstliche neuronale Netz wurde mit den dabei erhobenen Eyetracking-Daten parametrisiert und glich den Messfehler der Blickposition aus. In empirischen 2D-Untersuchungen konnten Pomplun et al. den Fehler in der Blickortberechnung durch die PSOM auf weniger als 40% des Initialwertes reduzieren. Für Untersuchungen im dreidimensionalen Bereich erweiterten Essig et al. diesen Ansatz. Sie entwickelten eine 3D-Kalibrierung, um ihre Version der PSOM entsprechend für räumliche Messungen anzupassen. Dazu zeigten sie den Versuchspersonen zu fixierende virtuelle Markierungen in drei verschiedenen Zeilen, Spalten und Ebenen des Raumes (3 x 3 x 3),

erhoben die Blickdaten und bestimmten daraus die PSOM-Parameter. Während des anschließenden Experimentes wurde die Blickbewegungsmessung mit einem binokularen Head-Mounted Eyetracker bei einem konstanten Abstand von 500 mm zwischen Versuchsperson und Bildschirm durchgeführt. Für die Erstellung der Stereobilder nahmen Essig et al. einen durchschnittlichen Augenabstand von 66 mm an. Ausgewertet wurde der Messfehler als euklidische Distanz zwischen gemessener und realer Position bezogen auf die Stimulus-Tiefenebene. Der Fehler durch die PSOM berechneter 3D Blickpositionen betrug im Gegensatz zur geometrischen Lösung nur noch 45%.

Daugherty, Duchowski und House (2010) nutzten ebenfalls die stereoskopische Bildtrennung durch Anaglyphen und kombinierten diese mit einem Remote Eyetracker. Das Videokamerasystem zum Aufzeichnen der Augenbilder befand sich dabei unterhalb des Bildschirms und wurde auf die Augen des Probanden ausgerichtet. Die Forscher versuchten festzustellen, inwieweit ein Remote Eyetracker zur okularen Vergenzmessung einsetzbar ist und ob es Unterschiede zwischen monoskopischer und stereoskopischer Stimuluspräsentation auf einem Standarddisplay gibt. Daugherty, et al. verwendeten ähnlich Duchowski et al. (2000) ein vereinfachtes geometrisches Modell, bei dem die Augen gleich weit von der Mess- bzw. Monitorebene entfernt sind und sich die Blickachsen horizontal schneiden. Die Bestimmung der Blickorttiefe zu einem Punkt P erfolgte über die Winkelbeziehung zwischen den Positionen der Augen und der zweidimensionalen Blickorte $(x_l, y_l), (x_r, y_r)$ in Projektions- bzw. Messebene. Auch hier wurden ausgehend von den Augenzentren und der binokularen Eyetracking-Daten die Blickachsen bzw. deren Schnittpunkt bestimmt. Daugherty et al. nahmen einen Augenabstand von $a = 63 \text{ mm}$ und eine Entfernung zwischen Augen und Monitor von $D = 500 \text{ mm}$ an. Die Distanz zwischen Monitor und virtuellem 3D-Objekt wird als Δd bezeichnet. Als Maß der Vergenz berechnet sich die Disparität η aus:

$$\eta = \frac{-a\Delta d}{D^2 + D\Delta d} \quad (4.7)$$

In einem Experiment bot man den Versuchspersonen ein Video dar, in dem sich ein zu fixierendes Objekt in Form einer Scheibe in drei unterschiedlichen Tiefenebenen an die jeweiligen 4 Eckpositionen bewegte. Die Tiefenebenen befanden sich vor, innerhalb und hinter der Darstellungsfläche. Entfernungsangaben gingen aus der Veröffentlichung allerdings nicht hervor. Dieses Video wurde zum einen monoskopisch und zum anderen stereoskopisch visualisiert. Die Ergebnisse zeigten einen hochsignifikanten Unterschied bzgl. der okularen Vergenz beim Betrachten des stereoskopischen Inhaltes im Gegensatz zum monoskopisch

dargebotenen. Ein Effekt konnte auch für die Vergenz-Antwort in den verschiedenen Tiefenebenen nachgewiesen werden, wobei ein Unterschied lediglich zwischen *vorderer* und *mittlerer* bzw. *vorderer* und *hinterer* Ebene bestand. Keine signifikanten Unterschiede gab es zwischen *mittlerer* und *hinterer* Ebene. Weshalb Daughtery et al. keine Vergenzunterschiede zwischen den beiden Ebene finden konnten wurde von den Autoren nicht erklärt, jedoch sind die Vergenzwinkel mit zunehmender Fixationsdistanz geringer und damit auch schwieriger zu unterscheiden.

Shutter-Technologie

Das Shutterverfahren ermöglicht eine bessere Farbwahrnehmung als das Anaglyph-Verfahren und wurde von Pfeiffer, Donner, Latoschik und Wachsmuth (2007) mit zwei Head-Mounted Eyetrackern verschiedener Hersteller gekoppelt. Allerdings gelang es den Forschern nicht durch die Shutterbrille hindurch zu messen. Daraufhin positionierten sie die Kameras der Eyetracker, deren absolute Position zu den Augen auch bei Kopfbewegungen unverändert blieb, unterhalb der Brille. Pfeiffer et al. setzten ähnlich Essig et al. (2006) das PSOM-Verfahren zur blickortunterstützten Objektauswahl (*Picking*) ein. In einem Experiment sollten Probanden dargebotene 3D-Objekte auf einem Computerbildschirm mit den Augen fixieren. Zunächst verglichen Pfeiffer et al. die errechneten Tiefeninformationen (Z-Koordinate) aus der PSOM mit denen der geometrischen Verfahren und konnten eine höhere Genauigkeit der Daten aus dem künstlichen neuronalen Netzwerk nachweisen. Dennoch wichen alle Ergebnisse vom Sollwert signifikant ab.

In einem weiteren Schritt setzten Pfeiffer et al. die Objektselektion als qualitatives Kriterium ein. Analysiert wurde, ob durch eine 3D-Fixationsbestimmung mehr korrekte Zuordnungen möglich sind als durch die Bestimmung der euklidischen Distanz aus den 2D-Daten. Bezogen auf die gesamte Szene führte die Berücksichtigung der Fixationstiefe entgegen der Hypothese zu weniger korrekten Selektionen (42%) als der 2D-Ansatz (75%). Lediglich in den Bereichen, in denen sich Objekte teilweise verdeckten, bewirkte die Beachtung der Tiefeninformation eine Steigerung der Selektionsleistung (42% korrekte Zuweisungen statt 30%). Ein Vergleich der durch die PSOM errechneten xy-Koordinaten mit den vom Eyetracker ausgegebenen Daten zeigte, dass die berechneten Werte stärker streuten. Gründe hierfür sahen Pfeiffer et al. in ungenauen 3D-Kalibrierungen durch das Projektionsverfahren (Geisterbilder) und einer unzureichend auflösbaren Blickwinkeldifferenz von 0.9° zwischen Fixationen der vordersten und hintersten Tiefenebene. Die Autoren wiesen darauf hin, dass das PSOM-Verfahren aufgrund der Ungenauigkeiten für eine Blickortzuweisung nicht generell zu empfehlen ist.

Polarisationsfilter

Auch Maggia, Guyader und Guérin-Dugué (2013) versuchten ihre Blickbewegungsmessung mit dem Shutterverfahren zu kombinieren, verwendeten dafür allerdings einen Remote Eyetracker. Hierbei stellten sie Probleme bei der Synchronisation zwischen 3D-Brille und Monitor fest (3.2.2). Die Infrarotlichtmessung des Remote Eyetrackers interferierte mit dem Synchronisationssignal der Shutterbrille und verhinderte eine korrekte Bildtrennung. Als Konsequenz wechselten sie zum Polarisationsfilter-Verfahren und untersuchten in einem Experiment die Disparitäten in unterschiedlichen Tiefenebenen. Ein Ziel dabei war es, die Tiefeninformation des 3D-Blickortes anhand der zweidimensionalen Ausgabeparameter des binokularen Remote Eyetrackers zu bestimmen. Dazu wurde die Disparität als horizontale Differenz zwischen den x- Koordinaten beider Blickorte in der zweidimensionalen Messebene des Eyetracking-Systems berechnet. In einer Kalibrierprozedur präsentierten die Forscher ihren Versuchspersonen zu fixierende Zielobjekte an unterschiedlichen Positionen im Raum: in Bildschirmenebene (ohne Disparität), hinter dem Bildschirm (ungekreuzte Disparität) sowie davor (gekreuzte Disparität). Anhand der zweidimensionalen Eyetracking-Daten wurden die individuellen Disparitäten berechnet und als Ausgangsparameter in Regressionsmodelle (linear und nichtlinear) eingesetzt. Für die Autoren waren lediglich die Abweichungen der Regressionsmodelle zu den realen Disparitäten von Interesse. Ergebnisse zu berechneten 3D-Blickorten wurden nicht berichtet.

Autostereoskopie

Die Bildtrennung für linkes und rechtes Auge ohne zusätzliche Brille ist mit einem autostereoskopischen Display möglich. Bei dieser Technologie werden beide Augenbilder gleichzeitig auf demselben Bildschirm dargestellt und mittels Parallaxbarrieren, einem Raster bestehend aus optischen Linsen oder Prismen, in verschiedene Richtungen zum jeweiligen Auge hin abgelenkt. Dabei ist eine Verschachtelung der beiden Bilder über die Displayfläche notwendig, die allerdings eine Halbierung der Horizontalauflösung pro Auge zur Folge hat. Sollen zusätzliche Perspektiven angeboten werden, erhöht sich die Zahl der Parallaxbarrieren und die Auflösung wird nochmals reduziert.

Kwon et al. (2006) kombinierten ein autostereoskopisches Display mit einem monokularen Remote Eyetracker. Die Kamera mit zwei Infrarotlichtquellen befand sich unterhalb des 3D-Bildschirms und zeichnete frontal das Gesicht bzw. die Augen des Probanden auf. Dabei erzeugten die beiden Infrarotlichtquellen zwei Reflexionspunkte in jedem Auge. Anhand der Winkelbeziehung zwischen Pupillenmittelpunkt und Cornea-Reflexionen ließ sich mit der PCCR-Methode auf die Blickposition in Monitorebene schließen. Um jedoch dreidimensional

wahrgenommene Objekte vor bzw. hinter dem Bildschirm zuordnen zu können, benötigten die Forscher zusätzlich die Tiefeninformation des Blickortes. Dazu wendeten Kwon et al. das Modell der *Pupil Center Distance* (PCD) an. In Abhängigkeit der Vergenzbewegung ändert sich durch die Augenrotation auch der Abstand zwischen den Pupillenzentren in Abhängigkeit der fixierten Tiefenebene. Da die physiologischen Augenparameter und damit auch die PCD zwischen verschiedenen Personen beim Betrachten eines Objektes gleicher Entfernung unterschiedlich sein können, wurde zur Erfassung der individuellen Kenngrößen vor Beginn der Datenerhebung eine Kalibrierung durchgeführt. Nachteil dieses Settings ist die Notwendigkeit der Kopffixierung, denn bei Kopfbewegungen variieren die Kenngrößen für die PCCR- sowie die PCD-Berechnung im Videokamerabild auch ohne Änderung des Blickortes oder Vergenzwinkels. Kwon et al. nutzten das Verfahren, um dreidimensionale Interaktionsmöglichkeiten zu schaffen. Dabei war es weniger bedeutsam den exakten 3D-Blickort zu ermitteln sondern vielmehr Interaktionsobjekte wie beispielsweise einen virtuellen Dartpfeil zu selektieren, welcher sich in vorgegebenen Sektoren (3x4) in 5 unterschiedlichen Tiefenebenen befand.

Die aufgeführten Beispiele verdeutlichen, dass es für eine Kombination stereoskopischer Visualisierungsverfahren mit einer Blickbewegungsmessung unterschiedliche Möglichkeiten in Abhängigkeit von der Zielsetzung gab, jedoch in Hinblick auf die stereoskopische Visualisierung, den eingesetzten Eyetracking-Systemen und der resultierenden Daten verschiedene Vor- und Nachteile aufzeigten.

4.5 Technologische Herausforderungen

Um zukünftig visuelle Wahrnehmungsprozesse räumlich analysieren zu können, müssen Blickorte dreidimensional erfasst, daraus Blickbewegungen berechnet und diese den wahrgenommenen Objekten der Umgebung zugeordnet werden. Eine korrekte Blickortzuweisung setzt dabei Kenntnisse über die Geometrie der Szene voraus. Hierfür bedarf es eines einheitlichen Koordinatensystems, in dem der Betrachter, alle vorhandenen Objekte sowie Blickorte und Blickbewegungen in Bezug gesetzt werden können. In der realen Welt ist es erforderlich, sämtliche Positionen auszumessen und konstant zu halten oder über Trackingsysteme sowie Videobildanalysen zu verfolgen. Bei computergenerierten Visualisierungen ist dies weniger komplex, denn hierbei ist die Geometrie der Szene und der darin enthaltenen Objekte selbst bei dynamischen Veränderungen zu jeder Zeit bekannt.

Die Herausforderung in virtuellen 3D-Umgebungen besteht in der Kombination der stereoskopischen Visualisierung mit einer Blickbewegungsmessung. Bei Verwendung eines autostereoskopischen Displays kann zwar auf das Tragen einer zusätzlichen 3D-Brille verzichtet und die Blickbewegungen wie im zweidimensionalen Setting ohne weitere Einschränkungen gemessen werden, dennoch erscheint dieses Bildtrennungsverfahren aufgrund der beschriebenen Auflösungsprobleme als ungeeignet. In einer stereoskopischen Visualisierung durch Anaglyphen ist die Bildtrennung über nur einen Kanal von Vorteil, jedoch wirken sich die Filter negativ auf die Farbwiedergabe aus. Bildinhalte sind in den Filterfarben nicht binokular kodierbar und dadurch nur für ein Auge sichtbar. Für eine farbtreuere Darstellung eignen sich Polarisations- und Shutterverfahren. Allerdings werden zur stereoskopischen Bildtrennung Anteile des sichtbaren Lichtes gefiltert. Dementsprechend muss geprüft werden, ob Eyetracking durch die für beide Verfahren notwendigen 3D-Brillen hindurch möglich ist und die Kombination einen Einfluss auf die Messgenauigkeit hat.

Der individuelle Augenabstand hat einen erheblichen Einfluss auf die Stereobasis und damit auf die wahrgenommene Tiefe dargestellter 3D-Objekte. In den 3D-Blickbewegungsstudien mit stereoskopischer Projektion wurde jedoch der Augenabstand für alle Versuchspersonen gleichermaßen mit beispielsweise 63 mm (Daugherty et al., 2010; Duchowski et al., 2011) oder 65 mm (Essig et al., 2006) angenommen. Damit in einer virtuellen Umgebung ein Objekt von unterschiedlichen Betrachtern an derselben Position wahrgenommen wird, muss bei der Berechnung der Augenbilder auch die individuelle Pupillendistanz einbezogen werden.

Blickbewegungsmessungen in realer Umgebung werden überwiegend mit Mobile Eyetrackern und in Labor-Experimenten mit Remote Eyetrackern durchgeführt. Dabei hat jedes System ein eigenes Koordinatensystem und gibt die Blickparameter in Bezug auf eine zweidimensionale Messebene aus. So wird beispielsweise der Blickort bei einem Mobile Eyetracker auf dem Szenevideo und bei einem Remote Eyetracker auf dem Computerbildschirm berechnet. Außerdem sind die Ausgabeparameter über verschiedene Eyetracking Systeme hinweg nicht standardisiert. Es gibt Eyetracking-Systeme, die lediglich die Blickorte für linkes und rechtes Auge innerhalb der Messebene angeben. Andere wiederum unterstützen zusätzlich die Ausgabe von Blickvektoren oder Positionsangaben der Augen während der Messung. Eine einheitliche Berechnung räumlicher Blickparameter setzt eine Methode voraus, die möglichst unabhängig vom verwendeten Eyetracking System mit verschiedenen Ausgabeparametern anwendbar ist.

Die Berechnungsmethoden zur Bestimmung dreidimensionaler Blickorte wurden in den berichteten Studien speziell auf die eingesetzte Stereoprojektion oder Blickbewegungsmessung ausgerichtet. Um die Ergebnisse aus verschiedenen Experimenten zukünftig vergleichen zu können, ist universeller Algorithmus zur Berechnung dreidimensionaler Blickorte notwendig. Außerdem ist bislang keine Berechnungsvorschrift zur Bestimmung räumlicher Blickbewegungen bekannt und muss entwickelt werden.

5 Zielsetzung und Konzeption

In dieser Arbeit wird eine Methode zur Bestimmung dreidimensionaler Blickorte und Blickbewegungen in realer sowie virtueller Umgebung entwickelt. Dazu werden anhand verfügbarer Eyetracking-Parameter dreidimensionale Blickorte berechnet und daraus Blickbewegungen abgeleitet.

Kombination von Eyetracking und stereoskopischer Visualisierungen

Im ersten Teil dieser Forschungsarbeit wird geprüft, inwieweit Eyetracking in einer stereoskopisch projizierten Umgebung durchführbar ist und welche ausgegebenen Kenngrößen für eine dreidimensionale Blickortberechnung relevant sind. Hierbei kommen Remote Eyetracker zum Einsatz, damit der Nutzer nicht durch am Kopf angebrachte Hardware in dessen Interaktion beeinflusst wird. Für die Kombination der Blickbewegungsmessung mit einer stereoskopischen Projektionstechnologie wird aufgrund der hohen örtlichen Auflösungen und einer weitestgehend unverfälschten Farbwiedergabe zur Bildtrennung das Polarisationsfilter- und Shutter-Verfahren getestet. Die Studie I in Kapitel 6 soll Aufschlüsse darüber geben, ob eine Blickbewegungsmessung durch die für die stereoskopische Bildtrennung notwendigen 3D-Brillen hindurch möglich ist und ob die Kombination einen Einfluss auf die erhobenen Eyetracking-Daten insbesondere auf deren Genauigkeit hat. Weiterhin wird untersucht, inwieweit das bei der Blickbewegungsmessung emittierte Infrarotlicht mit den Synchronisationssignalen der Projektionstechnik interferiert und eine korrekte Bildtrennung verhindert. Um die Eyetracking-Parameter den visuellen Informationen korrekt zuordnen zu können, werden beide Verfahren zeitlich gekoppelt. Für die Umsetzung sind die Konzeption und Implementierung entsprechender Softwareschnittstellen, Synchronisationssignale sowie Übertragungsprotokolle zwischen Eyetracking-System und stereoskopischer Visualisierung erforderlich.

Algorithmus zur 3D-Blickortbestimmung

In Kapitel 7 wird aus den bisherigen Erkenntnissen ein Algorithmus entwickelt, welcher dreidimensionale Blickorte anhand der Ausgabewerte zur Verfügung stehender Eyetracking-Systeme berechnet. Diese Parameter müssen unter Betrachtung der im ersten Teil gewonnenen Erkenntnisse sowohl bei Untersuchungen in realer Umgebung als auch in Kombination mit einer stereoskopischen Visualisierung messbar sein. Die Bestimmung des Tiefenparameters ist dabei von besonderem Interesse, da sich diese in der Vergangenheit mit zunehmender Fixationsdistanz als problematisch und ungenau erwies. Eine Optimierung soll

erreicht werden, indem der Algorithmus individuelle physiologische Kenngrößen wie die 3D-Positionen für linkes bzw. rechtes Auge und Ausrichtung der Blickachsen im Raum berücksichtigt. Je genauer die Berechnung der 3D-Blickposition erfolgt, desto präziser können später die räumlichen Blickbewegungen abgeleitet werden.

Dreidimensionale Blickortmessung in realer und virtueller Umgebung

In zwei weiteren Studien werden Eyetracking-Parameter beim Betrachten sowohl realer (Kapitel 8) als auch stereoskopisch projizierter (Kapitel 9) 3D-Stimuli anhand eines binokularen Remote Eyetrackers erhoben und die Abweichungen der durch den Algorithmus ausgegebenen 3D-Blickorte zu den Stimuluspositionen bestimmt. Für eine gleiche Tiefenwahrnehmung der virtuellen Szene wird dabei die individuelle Stereobasis gemessen und in die Generierung der stereoskopischen Bilder einbezogen.

Algorithmus zur Erkennung dreidimensionaler Blickbewegungen

Um aus den berechneten 3D-Blickorten räumliche Blickbewegungen ableiten zu können, ist zu prüfen, welche zweidimensionalen Methoden bekannt sind und inwieweit sich diese auf den dreidimensionalen Bereich erweitern lassen. Aus den Erkenntnissen wird in Kapitel 10 ein Algorithmus zur Identifikation von 3D-Fixationen konzipiert und implementiert, der im Vergleich zu existierenden Lösungen zusätzlich den Tiefenparameter berücksichtigt. Anschließend werden die in Studie II und III berechneten 3D-Blickorte dem Algorithmus übergeben und daraus 3D-Fixationen mit unterschiedlichen Erkennungsparametern bestimmt. Die resultierenden Fixationsparameter sollen Aufschluss über die Genauigkeit der erhobenen Blickbewegungen geben und zeigen, dass der Algorithmus sowohl in realer als auch stereoskopisch projizierter Umgebung anwendbar ist.

Allgemeine Angaben zur Datenanalyse

Die innerhalb dieser Arbeit durchgeführten Experimente zielen überwiegend darauf ab, die Genauigkeit der erhobenen Blickparameter zu analysieren. Dementsprechend ist das generelle Vorgehen der statistischen Auswertung in den einzelnen Studien ähnlich. Für die Genauigkeitsbestimmung wird jede einzelne Parameterabweichung als zweidimensionale oder dreidimensionale Distanz zur Stimulusposition berechnet. Die Überprüfung der Werte auf Normalverteilung erfolgt mit einem Komolgorov-Smirnov-Test. Bei asymptotischer Signifikanz wird der Median als Lageparameter gewählt, da sich dieser unempfindlicher gegenüber Extremwerten verhält und zur Beschreibung der Datenmenge besser eignet. Hierbei wird die Abweichungsdifferenz für jedes Sample einzeln berechnet und anschließend der Median über die Abweichungen für jede Versuchsperson in der jeweiligen Bedingung

bestimmt. In den mehrfaktoriellen ANOVAS mit Messwiederholung werden die post-hoc Analysen mittels Bonferroni-korrigierten Einzelvergleichen (t-Tests) gerechnet, um genauere Aussagen zu Richtung und Art der erzielten Haupt- und Interaktionseffekte zu erhalten. Als Maß für die Effektstärke ist EtaQuadrat (Levine & Hullett, 2002) und bei Verletzung der Sphärizitäts-Annahmen zusätzlich der Greenhouse-Geisser-Korrekturfaktor Epsilon angegeben. Die statistische Datenanalyse wird mit IBM Statistics SPSS (17, 19, 21), sowie Mathworks Matlab (7.11.0, R2014a) durchgeführt.

6 Studie I: Stereoskopische Visualisierung und Eyetracking

Ziel der im Folgenden dargestellten Studie war die Integration stereoskopischer 3D-Visualisierungstechniken anhand von Polarisationsfilter- und Shutterverfahren in eine Blickbewegungsmessung mit Remote Eyetrackern. Es wurde untersucht, ob die Verwendung von 3D-Brillen die Aufzeichnung einschränkt bzw. die Genauigkeit der aufgezeichneten Eyetracking-Parameter beeinflusst. Die sich aus der Kombination beider Technologien zwingend ergebende Frage lag darin, ob das für die Blickbewegungsmessung notwendige Infrarotlicht von den verwendeten stereoskopischen Visualisierungstechniken gefiltert wird. In diesem Fall wäre eine Blickortbestimmung in Kombination mit stereoskopischen Visualisierungsverfahren nur teilweise bzw. gar nicht möglich. Anhand der technischen Dokumentationen und Nachfragen bei den Herstellern konnten hierzu keine eindeutigen Aussagen getroffen werden.

In einem Vorversuch sollte zunächst geprüft werden, ob eine Kalibrierung und damit auch Messung von Blickdaten unter Verwendung stereoskopischer Visualisierungen prinzipiell gewährleistet ist. Sowohl beim Tragen der Polarisationsfilterbrille (Abb. 6-1 A) als auch der Shutterbrille (Abb. 6-1 B) startete der Kalibrierprozess problemlos, jedoch kam es im Setting mit Shutterbrille zu Darstellungsproblemen. Die Übertragung des Synchronisationssignals zum Öffnen und Schließen der jeweiligen Blende für linkes und rechtes Auge erfolgte durch ein Infrarotlichtsignal zwischen dem per USB-Schnittstelle direkt mit dem Computer verbundenen Sender und dem in den linken Brillenbügel integrierten Empfänger. Bei gleichzeitiger Aktivierung von Blickbewegungsmessung und stereoskopischer Visualisierung wurde das infrarote Licht des Eyetrackers ebenfalls vom Empfänger der Shutterbrille erfasst und bewirkte ein Öffnen beider Blenden unabhängig vom dargebotenen Augenbild. Zur Lösung dieses Problems wurde der Empfänger im Brillenbügel mit einer Abdeckung versehen, die den Empfang von Signalen nur aus seitlicher Richtung zuließ. Dementsprechend musste der Sender neben dem Kopf der Versuchsperson positioniert werden. Da der Strahlengang des eingesetzten Infrarotlichts von den Eyetracking-Systemen ausgehend frontal auf das Gesicht des Probanden verlief, konnte dieses durch die Abdeckung vom Empfänger der Shutterbrille nicht mehr erfasst werden. Damit wurde die korrekte Synchronisation der Blenden mit den dargebotenen Augenbildern wieder hergestellt.

In drei Durchgängen konnten die eingesetzten Remote Eyetracker beim Tragen einer Polarisationsfilter- bzw. Shutterbrille erfolgreich kalibriert und damit eine vollständige

Filterung mit resultierendem Datenausfall ausgeschlossen werden. Allerdings ist eine erfolgreiche Kalibrierung auch bei diskontinuierlicher Blockierung möglich. Ungültige Messwerte entstehen ebenso bei Blinzelbewegungen und werden vom Eyetracking-System toleriert. Deshalb sollte die anschließende Untersuchung zeigen, ob eine teilweise Beeinträchtigung der Blickbewegungsmessung vorliegt und inwieweit die Genauigkeit der erhobenen Daten durch die stereoskopische Visualisierung beeinflusst wird.

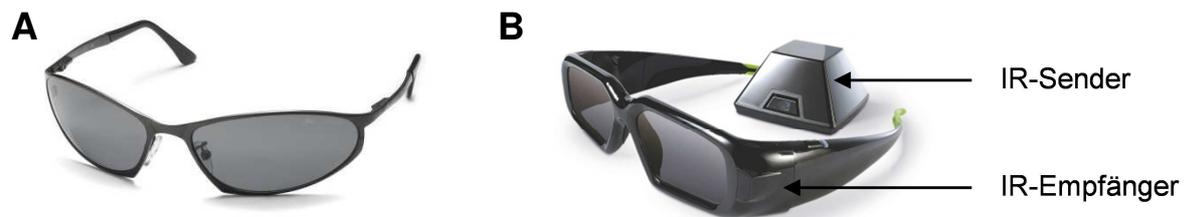


Abb. 6-1: Technologien zur stereoskopischen Bildtrennung. (A) Polfilterbrille (vgl. Bild: <http://www.my-bolles.com>) und (B) NVIDIA 3D VISION® Shutter Glasses (Bild: <http://www.nvidia.com>) mit Sender und Empfänger zur Synchronisation der Blenden für linkes und rechtes Auge in Abhängigkeit des dargebotenen Bildschirminhaltes.

6.1 Fragestellungen und Hypothesen

Bei einer Blockierung des Infrarotlichtes durch Polarisationsfilter- oder Shutterbrille wären die Voraussetzungen der PCCR-Methode nicht mehr gegeben und es würden ungültige Sampledaten entstehen. Da ein vollständiger Datenausfall aufgrund der erfolgreichen Kalibrierungen im Vorversuch ausgeschlossen werden konnte, blieb die Frage nach einer diskontinuierlichen Filterung. Falls Polarisationsfilter- oder Shutterbrille die PCCR-Methode teilweise verhindern, gäbe es in Messreihen mit 3D-Brillen weniger gültige Sampledaten.

Ob die stereoskopische Visualisierungstechnologie einen Einfluss auf die Genauigkeit der Blickbewegungsmessung hat, sollte einerseits durch die Qualitätsmerkmale der Hersteller-Kalibrierung als Distanz der ermittelten Blickorte zu den angezeigten Kalibrierpunkten und andererseits durch die Abweichungen der Sampledatenpositionen und Fixationsorte in Bezug auf dargebotene 2D-Stimuli getestet werden. Generell sind geringere Abweichungen bzgl. der Fixationsorte als die der Samplepositionen zu erwarten, da Fixationen mittels entsprechender Erkennungsalgorithmen aus aggregierten Sampledaten berechnet werden. Vereinzelt, stark abweichende Werte gehen dabei weniger gewichtet in die Berechnung ein. Da die Genauigkeit der Messwerte aber auch von der Blickortposition auf dem Bildschirm abhängig sein kann (Holmqvist et al., 2011; Zhang & Hornof, 2011), sollten die Abweichungen nicht nur in den einzelnen Bedingungen der stereoskopischen Visualisierung sondern auch für unterschiedliche Bereiche des Computerbildschirms betrachtet werden. Eine weitere Frage ergab sich aus den

in Polarisationsfilter- und Shutterbrillen eingesetzten Filtermechanismen. Falls einfallendes, sichtbares Licht und damit visuelle Informationen blockiert werden, müsste sich dieser Effekt anhand vergrößerter Pupillenwerte in den Blickbewegungsmessungen mit zusätzlicher 3D-Brille widerspiegeln.

6.2 Methode

In einem Experiment erfolgte die Integration stereoskopischer Visualisierungstechnologien in eine Blickbewegungsmessung. Dabei wurden in mehreren Durchgängen Stereoprojektion, Eyetracking-System sowie die Position der angezeigten Stimuli variiert.

Stichprobe

An der Studie nahmen 12 Versuchspersonen (8 weiblich und 4 männlich) im Alter zwischen 19 und 45 Jahren teil. Das Durchschnittsalter lag bei 26 Jahren. Alle Versuchsteilnehmer hatten normales oder durch eine Brille bzw. Kontaktlinsen korrigiertes Sehvermögen.

Apparatus und Aufbau

Für die stereoskopische Visualisierung kamen das Polarisations- und das Shutterverfahren zum Einsatz, wobei die jeweiligen Filtertechnologien aktiv waren, die Stimuli jedoch zweidimensional in Bildschirmenebene präsentiert wurden. Bei ersterem hatten die verwendeten Polarisationsfilter eine Ausrichtung von 45° für das linke bzw. -45° für das rechte Auge und wurden in ein Brillengestell zwischen Auge des Probanden und Eyetracking-System geschaltet. Die technische Umsetzung des Shutterverfahrens erfolgte mit der 3D VISION® Technologie des Herstellers NVIDIA bei einer Frequenz von 60 Hz pro Auge bzw. 120 Hz alternierend. Auf einem 19" LCD-Monitor mit einer Auflösung von 1.280×1.024 px und einer Bildwiederholrate von 60 Hz wurden die nichtstereoskopischen Stimuli dargeboten. Der Abstand zwischen Versuchsperson und Monitor (Auge-Monitor-Distanz) betrug 70 cm. Die Blickdaten wurden mit zwei verschiedenen videobasierten Remote Eyetrackern erhoben, dem *Eyegaze Analysis System* (EAS) sowie dem *EyeLink™ 1000 System* (ELS). In Tab. 6-1 sind die technischen Details beider Eyetracker dargestellt.

Tab. 6-1: Technische Details der Remote Eyetracker für die Blickbewegungsmessung.

	Eyegaze Analysis System (EAS)	EyeLink™ 1000 System (ELS)
Örtliche Auflösung	0.4°	0.5°
Zeitliche Auflösung	120 Hz (2 * 60 Hz alternierend)	500 Hz
Methode	bright-pupil, binokular	dark-pupil, monokular
Hersteller	LC Technologies, Inc. (USA)	SR-Research, Ltd.(CAN)

Die spätere Auswertung der Ergebnisse erfolgte für beide Eyetracking-Systeme getrennt voneinander, da sich das Messprinzip bzw. die Bestimmung der Pupille im Kamerabild unterschied (*bright pupil* vs. *dark pupil*) und ein Effekt auf die Kombination mit stereoskopischen Visualisierungstechnologien nicht ausgeschlossen werden konnte.

Stimulusmaterial

Den teilnehmenden Versuchspersonen wurde ein Fixationsobjekt in Form eines Kreisrings (siehe Abb. 6-2 A) auf einem Computerbildschirm dargestellt. Jedes Fixationsobjekt hatte einen äußeren Durchmesser von 0.5° sowie einen Innendurchmesser von 0.1° und wurde als schwarzes (RGB: 0, 0, 0) 2D-Objekt vor einem grauen (RGB: 204, 204, 204) Hintergrund präsentiert. Das Fixationsobjekt wechselte in einem 1.000 ms - Rhythmus die Position, welche randomisiert an 13 verschiedenen Koordinaten gleichmäßig über den Bildschirm verteilt war (Abb. 6-2 B).

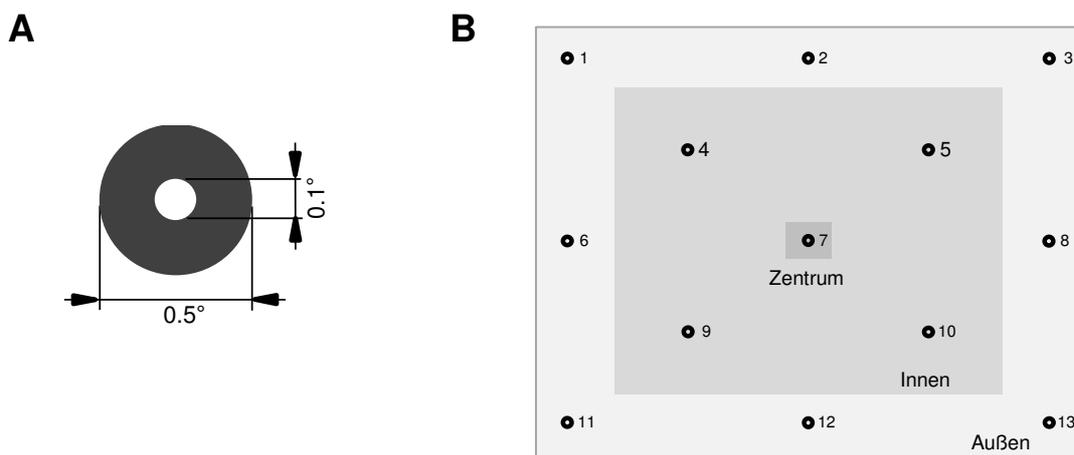


Abb. 6-2: Fixationsobjekt in Form eines Kreisrings (A) und die mögliche Positionierung auf dem Bildschirm (B).

Versuchsablauf

Das Experiment bestand aus 6 Durchgängen. Variiert wurden dabei die Eyetracking-Systeme: EAS und ELS sowie die Visualisierungsbedingungen: ohne stereoskopische Technologie (Ohne3D), unter Verwendung einer Polarisationsfilterbrille (Polfilter) sowie durch Zwischenschaltung einer Shutterbrille (Shutter). Jede Versuchsperson wurde erst mit einem der Eyetracker in den verschiedenen Visualisierungsbedingungen getestet und anschließend mit dem anderen Eyetracking System. Die Reihenfolge der Blickbewegungsmessung sowie der Visualisierungsbedingungen waren ausbalanciert. Abb. 6-3 zeigt beispielhaft Kamerabilder der beiden Eyetracking-Systeme in den unterschiedlichen Visualisierungsbedingungen.

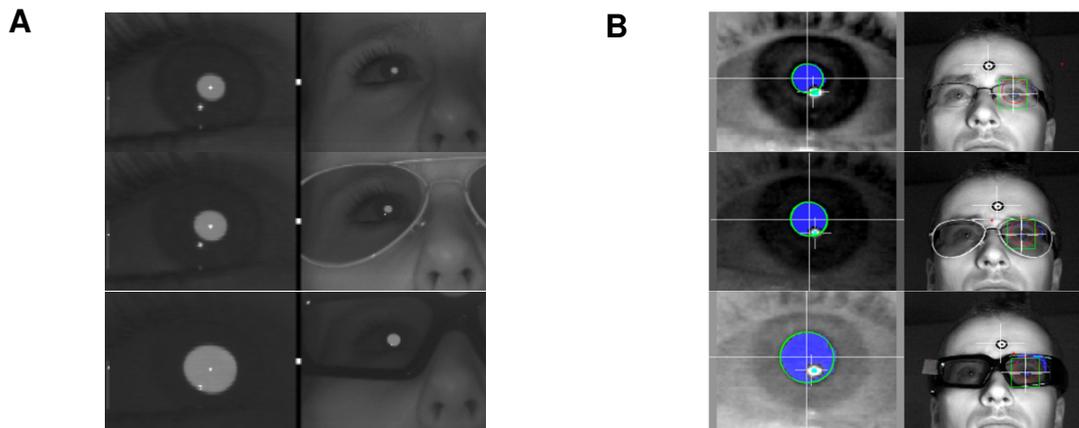


Abb. 6-3: Videokamerabilder des Eyegaze Analysis Systems (A) und des EyeLink™ 1000 Systems (B) in den verschiedenen Bedingungen (von oben nach unten): ohne 3D-Brille, mit Polarisationsfilterbrille, mit Shutterbrille.

Vor jedem Durchgang fand die vom Hersteller vorgegebene Kalibrierung des jeweiligen Eyetrackers auf die zu messende Person statt, um die individuellen physiologischen Parameter zu bestimmen. Während des Kalibrierprozesses wurde auf dem Computerbildschirm ein Kalibrierpunkt dargeboten, der mit den Augen möglichst genau fixiert werden sollte. Dieser erschien nacheinander an neun verschiedenen Positionen. Beim ELS fand nach der Kalibrierung zusätzlich eine Validierung statt, wobei in ähnlicher Weise ein Validierungspunkt an verschiedenen Monitorpositionen mit den Augen fixiert werden sollte. Sofern die durchschnittliche Abweichung zwischen Validierungspunkt und gemessenem Blickort $< 1^\circ$ und der maximale Fehler $< 1.5^\circ$ betrug, wurde die Messgenauigkeit für das Experiment akzeptiert. Beim EAS erfolgte die Validierung anhand der bereits erhobenen Messdaten aus der Kalibrierung. Diese wurden zusammen mit den Kalibrierparametern der Blickortberechnung übergeben und dadurch die zu den Kalibrierpunkten zugehörigen Blickpositionen bestimmt. Das EAS bestätigte eine erfolgreiche Kalibrierung bei einer mittleren Abweichung $< 0.3^\circ$.

Bei hinreichender Genauigkeit der Kalibrierergebnisse wurde das Experiment fortgesetzt, andernfalls die Kalibrierung wiederholt. Jeder Durchgang bestand aus einem Probedurchgang plus zehn weiterer Blöcke. Im Probedurchgang sowie jedem Block im Experiment erschien das Fixationsobjekt randomisiert an allen 13 Positionen für jeweils 1.000 ms. Die Probanden wurden instruiert, stets das Zentrum des Fixationsobjektes mit den Augen zu fixieren. Nach jedem Block konnte die Versuchsperson eine Erholungspause selbstgewählter Dauer einlegen, um Müdigkeitserscheinungen entgegenzuwirken bzw. die Konzentration aufrechtzuerhalten.

Datenaufbereitung

Zunächst mussten die Blickdaten bezogen auf das jeweilige Fixationsobjekt gefiltert werden, wobei die sich die Stimulus-Übergänge als einziges Filterkriterium als unzureichend herausstellten. Während bereits der neue Zielreiz erschien, verweilte das Auge noch für kurze Zeit an der alten Position, um die Sakkade zum neuen Zielreiz programmieren und auslösen zu können. Das entstehende Zeitfenster bzw. die *Sakkadenlatenz* ist in der Literatur mit ca. 200 ms angegeben (Trottier & Pratt, 2005). Erst nach dem Blickortwechsel war die Fovea auf die veränderte visuelle Information neu ausgerichtet. Deshalb wurden zunächst alle Fixationen, die innerhalb der jeweiligen Stimulusdarbietung begannen, diesem Fixationsobjekt zugeordnet.

Allerdings ist nach einem Stimuluswechsel nicht zwingend die erste Fixation relevant. Beim Erscheinen des zu fixierenden Objektes springt das Auge mittels einer Sakkade in die Nähe des Zielreizes und versucht diesen so genau wie möglich zu treffen. Gelingt dies nicht sofort, muss die Blickposition durch eine Korrektursakkade angepasst werden (siehe Abb. 6-4). Die vorhergehende Sakkade wird dann als *Under-* bzw. *Overshoot* bezeichnet, je nachdem ob sie zu kurz oder zu lang war (McLaughlin, 1967). Dementsprechend wurde ein weiterer Filteralgorithmus angewandt, um die dem Stimulus zugehörige Fixation mit dem geringsten Abstand zum Fixationsobjekt zu selektieren.

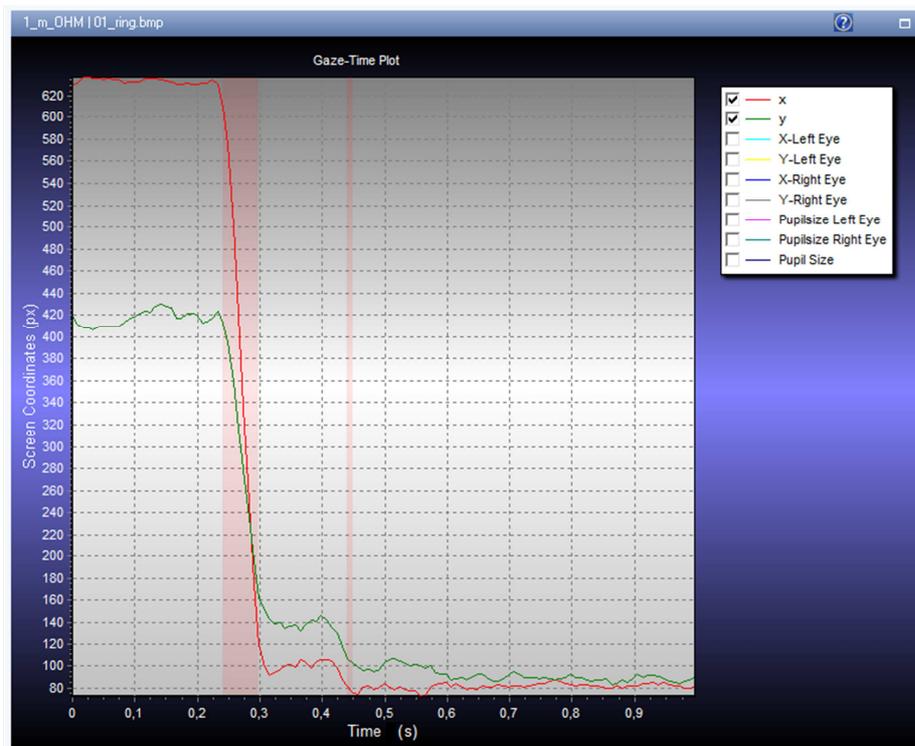


Abb. 6-4: Gaze Time Plot (NYAN 2.0^{XT}): Darstellung der Blickortkoordinaten (x = rot, y = grün) in Abhängigkeit der Zeit. Der Stimuluswechsel erfolgte zum Zeitpunkt $t = 0$ und die erste Sakkade zum neuen Blickort nach 0,22 s. Anschließend fand eine Korrektursakkade nach 0,43 s statt. Für die Auswertung bedeutsam ist die zweite Fixation beginnend ab 0,45 s.

6.3 Ergebnisse

Um ausschließlich valide Messungen in die Berechnungen einfließen zu lassen, sind ungültige Samples während der Fixationsobjektdarbietung aus der Datenmenge entfernt worden. Als ungültig wurden diejenigen Samples deklariert, denen kein valider x- und y-Wert des Blickorts zugeordnet werden konnte (Tab. 6-2). Die übrigen 0,3 – 1,5% ungültige Samples sind nicht zwingend durch Messfehler sondern auch anhand stattgefundener Blinzelbewegungen (*Blinks*) erklärbar. In einer einfaktoriellen (Ohne3D, Polfilter, Shutter) ANOVA mit Messwiederholung waren keine signifikanten Unterschiede bezüglich der Datenausfälle zwischen den jeweiligen Bedingungen für EAS und ELS, mit jeweils $F < 1$, nachweisbar.

Tab. 6-2: Gültige Samples mit validem x- bzw. y-Wert.

	Ohne3D	Polfilter	Shutter
EAS	98,59%	99,15%	99,24%
ELS	99,70%	99,52%	99,56%

Kalibrierergebnisse

Die Kalibrierergebnisse als mittlere Abweichung zwischen den angezeigten Kalibrierpunkten und berechneten Blickorten wurden mittels einer einfaktoriellen (Bedingung: Ohne3D, Polfilter, Shutter) ANOVA mit Messwiederholung ausgewertet. Signifikante Unterschiede in den Kalibrierwerten durch die Verwendung von 3D-Brillen konnten weder für das EAS, $F(2,22) = 1.989$, $p = .161$ (Ohne3D = 0.20° [SE = 0.02], Polfilter = 0.23° [SE = 0.019], Shutter = 0.24° [SE = 0.15]), noch für das ELS, $F(2,22) = 2.261$, $p = .128$ (Ohne3D = 0.35° [SE = 0.03], Polfilter = 0.46° [SE = 0.04], Shutter = 0.41° [SE = 0.03]), nachgewiesen werden.

Blickorte

Die Abweichungsunterschiede zu den Fixationsobjekten wurden einerseits in Bezug auf die drei verschiedenen Messbedingungen und andererseits hinsichtlich der Positionierung auf dem Bildschirm ausgewertet. Dazu gab es eine Unterteilung der 13 Fixationsobjekte in 3 vom Zentrum unterschiedlich weit entfernte Bereiche (Abb. 6-2 B): Zentrum, innerer Bereich (Innen) und äußerer Bereich (Außen). Für die Sampledaten der beiden Messsysteme wurde jeweils eine 3 (Bedingung: Ohne3D, Polfilter, Shutter) x 3 (Bereich: Zentrum, Innen, Außen) faktorielle ANOVA mit Messwiederholung berechnet. Für das EAS konnte ein höchst signifikanter Haupteffekt für den Faktor Bedingung, $F(2,22) = 11.13$, $p < .001$, $\eta^2 = .313$, sowie Signifikanz für den Faktor Bereich, $F(2,22) = 6.242$, $p = .023$, $\varepsilon = .594$, $\eta^2 = .076$,

aufgezeigt werden. Die Wechselwirkung zwischen Bedingung und Bereich erwies sich als nicht signifikant, $F < 1$. Eine post-hoc Analyse mittels Bonferroni korrigierter gepaarter t-Tests bestätigte, dass die Abweichungen zur Originalposition ohne stereoskopische Hilfsmittel am geringsten waren. Bei Verwendung einer Polfilter- bzw. Shutterbrille nahmen die Abweichungen zu (Abb. 6-5 A), allerdings war ein hoch signifikanter Unterschied ausschließlich zwischen der Bedingung Ohne3D und Shutter, $p < .004$, vorhanden. Bezüglich der verschiedenen Bereiche divergierten die gemessenen von den originalen Positionen im Zentrum des Monitors am wenigsten. Mit Entfernungszunahme ausgehend von der Monitormitte wurden auch die Abweichungen größer, wobei sich diese lediglich zwischen dem zentralen und äußeren Bereich höchst signifikant unterschieden, $p < .001$.

Unter Verwendung stereoskopischer Hilfsmittel nahmen die Blickortabweichungen auch beim ELS zu (Abb. 6-5 B), signifikante Unterschiede waren allerdings nicht nachweisbar, $F < 1$. Wie beim EAS fanden sich in den verschiedenen Monitorbereichen höchst signifikante Abweichungsdifferenzen, $F(2,22) = 25.116$, $p < .001$, $\eta^2 = .346$. Die Messungen im Zentrum des Monitors waren auch hier am genauesten und die Abweichungen verstärken sich nach außen hin. Signifikante Unterschiede gab es zwischen den Bereichen Zentrum – Innen, $p = .012$, hohe Signifikanz zwischen Innen – Außen, $p = .003$, und höchst signifikante Differenzen für Zentrum – Außen, $p < .001$. Eine Wechselwirkung zwischen stereoskopischer Visualisierung und Monitorbereich bestand nicht, $F < 1$.

Fixationen

Den folgenden Berechnungen lagen die Fixationsdaten aus den Erkennungsalgorithmen der Hersteller zugrunde. Ziel war es, die Abweichungen zwischen den ausgegebenen Fixationsorten und den Positionen der Fixationsobjekte in den verschiedenen Visualisierungsbedingungen zu überprüfen. Dazu wurde jeweils eine 3 (Bedingung: Ohne3D, Polfilter, Shutter) x 3 (Bereich: Zentrum, Innen, Außen) faktorielle ANOVA mit Messwiederholung gerechnet. Für das EAS konnten ähnlich den Sampledaten höchst signifikante Effekte in Bezug auf Abweichungen resultierend durch die stereoskopischen Visualisierungsbedingungen, $F(2,22) = 11.115$, $p < .001$, $\eta^2 = .371$, und signifikante Ungenauigkeiten für unterschiedliche Monitorbereiche, $F(2,22) = 7.613$, $p = .013$, $\varepsilon = .607$, $\eta^2 = .107$, nachgewiesen werden. Die Abweichungen zur Originalposition waren ohne Verwendung stereoskopischer Verfahren am geringsten. Durch das Tragen einer Polfilterbrille bzw. einer Shutterbrille nahmen diese zu. Hoch signifikante Unterschiede wurden zwischen Ohne3D und Shutter, $p = .006$, gefunden. Im Zentrum des Monitors wichen die gemessenen von den originalen Werten am wenigsten ab. Mit Zunahme der Entfernung von der

Monitormitte verstärkten sich die Abweichungen. Höchste Signifikanz lag zwischen dem inneren und äußeren Bereich, $p < .001$, vor. Eine Wechselwirkung zwischen stereoskopischer Visualisierung und einzelnen Monitorbereichen bestand nicht, $F < 1$.

Beim ELS gab es keine signifikanten Genauigkeitsunterschiede durch die Verwendung stereoskopischer Hilfsmittel, $F(2,22) = 1.171$, $p = .329$. Die Abweichungen nahmen jedoch auch hier in den Messungen mit 3D-Brille zu. In den verschiedenen Bereichen wurden höchst signifikante Abweichungsdifferenzen, $F(2,22) = 20.109$, $p < .001$, $\eta^2 = .434$, nachgewiesen. Die Messgenauigkeit war im Zentrum am größten und nahm wie beim ELS nach außen hin ab. Eine Wechselwirkung zwischen stereoskopischer Visualisierung und den Monitorbereichen bestand ebenfalls nicht, $F < 1$.

Pupillengrößen

Für die folgenden Aussagen bleibt zu beachten, dass sich die Angaben zur Pupillengröße lediglich auf errechnete Daten im Videokamerabild beziehen und somit keine Aussage auf das physiologische Maß zulassen. Der EAS-Hersteller berechnet den Pupillenradius (mm) im Videokamerabild (LC Technologies, 2002), während SR Research in den ELS-Daten die Pixelanzahl der erkannten Pupillenfläche im Videobild ausgibt (SR Research, 2009). Betrachtet wurden die Pupillengrößenangaben aus den Sampledaten in der jeweils relevanten Fixation. Da Pupillengrößen zwischen den Menschen variieren, wurden die erhobenen Daten mit einer Baseline korrigiert. Die Bestimmung der individuellen Baseline erfolgte anhand der gemessenen Pupillengrößen bei Fixierung der im Zentrum dargebotenen Fixationsobjekte ohne stereoskopische Visualisierungstechnologien.

Es wurde jeweils eine 3 (Bedingung: Ohne3D, Polfilter, Shutter) x 3 (Bereich: Zentrum, Innen, Außen) ANOVA mit Messwiederholung in Bezug auf die Pupillenparameter gerechnet. Für das EAS konnte ein höchst signifikanter Effekt resultierend durch die unterschiedlichen stereoskopischen Visualisierungsbedingungen, $F(2,22) = 33.759$, $p < .001$, $\varepsilon = .595$, $\eta^2 = .753$, gezeigt werden. Die ausgegebenen Messwerte waren ohne stereoskopische Visualisierungstechnologien am kleinsten und nahmen unter Verwendung einer Polfilter- bzw. einer Shutterbrille zu (Abb. 6-5 C). Hoch signifikante Unterschiede gab es zwischen den Bedingungen Ohne3D - Polfilter, $p = .002$, sowie Polfilter – Shutter, $p = .023$, und höchst signifikante Differenzen für Ohne3D – Shutter, $p < .001$. Die Pupillenwerte in den verschiedenen Monitorbereichen unterschieden sich ebenfalls höchst signifikant voneinander, $F(2,22) = 12.451$, $p < .001$, $\eta^2 = .001$, und waren im Zentrum kleiner als in den Randbereichen. Signifikanz konnte zwischen Zentrum und äußerer Bereich, $p = .007$, sowie innerer und äußerer Bereich, $p = .011$, festgestellt werden. Eine Wechselwirkung zwischen

stereoskopischer Visualisierung und Monitorbereich bestand nicht, $F < 1$. In den ELS-Daten gab es ebenfalls einen höchst signifikanten Effekt durch die Verwendung stereoskopischer Visualisierungstechnologien, $F(2,22) = 39.691$, $p < .001$, $\varepsilon = .686$, $\eta^2 = .782$. Mit zusätzlicher 3D-Brille vergrößerten sich die Pupillenwerte (Abb. 6-5 D). Dabei zeigten sich hoch signifikante Unterschiede zwischen allen drei Bedingungen, $p \leq .004$. Für den Faktor Bereich wurde kein Effekt gefunden, $F < 1$.

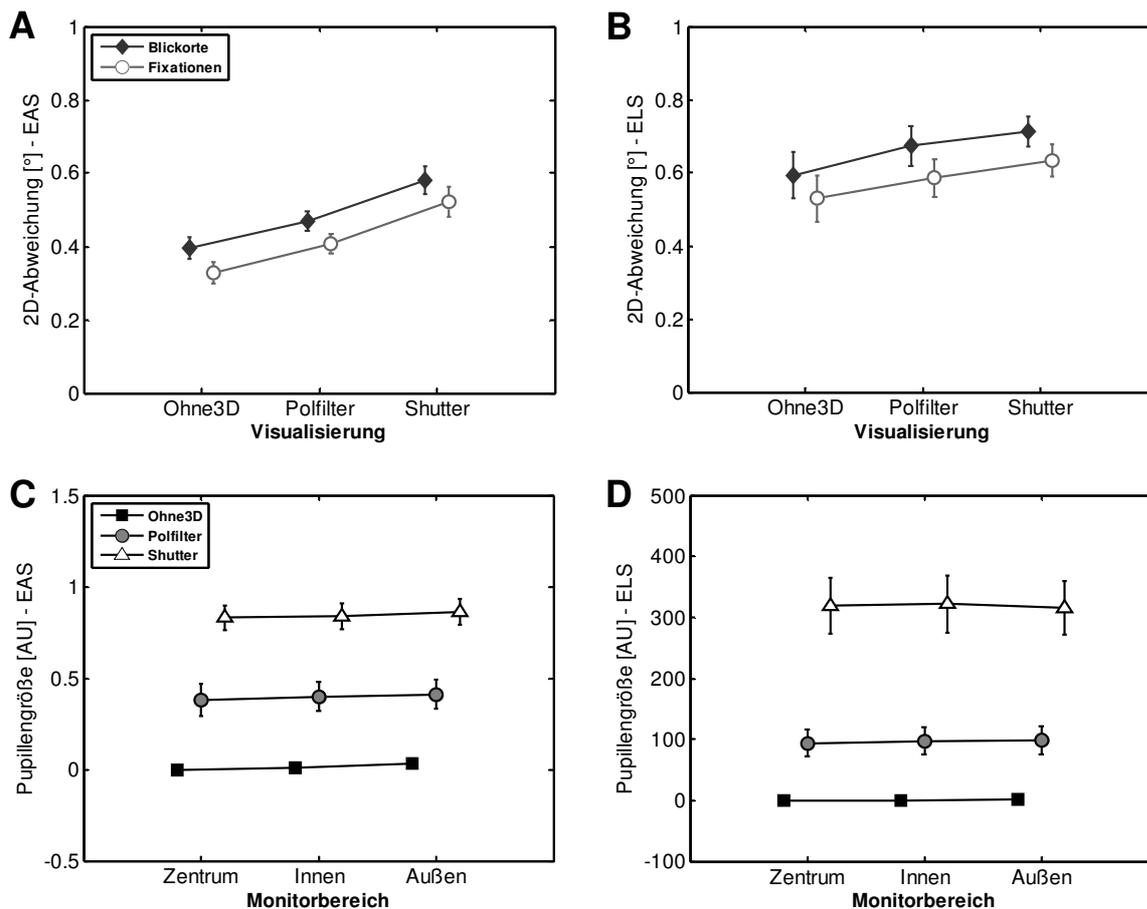


Abb. 6-5: Zweidimensionale Abweichung der Blickorte- (A) und Fixationen (B) in den verschiedenen Visualisierungsbedingungen. Die gemessenen Pupillenwerte (Arbitrary Units) in den Videokamerabildern vergrößerten sich bei Einsatz stereoskopischer Visualisierungstechnologien im Vergleich zur Baseline sowohl in der EAS-Messung (C) als auch in der ELS-Messung (D). Fehlerbalken kennzeichnen das 95%-CI.

6.4 Diskussion

Innerhalb dieser Studie sollte geprüft werden, ob eine Blickbewegungsmessung in Kombination mit einer stereoskopischen Projektion möglich ist oder aufgrund der notwendigen 3D-Brillen teilweise oder sogar vollständig verhindert wird. Signifikante Unterschiede bezüglich der Anzahl gültiger und ungültiger Sampledaten in den Messungen mit und ohne 3D-Brille waren jedoch nicht nachweisbar. Dementsprechend konnten Blickbewegungen durch die Brillen hindurch gemessen werden, allerdings ließen diese

Ergebnisse noch keinerlei Aussage über einen Einfluss der stereoskopischen Visualisierungstechnologie auf die Genauigkeit der Messwerte zu. Deshalb wurden die Kalibrierwerte beider Eyetracker in den unterschiedlichen Bedingungen mit und ohne 3D-Brillen verglichen. Die Ergebnisse zeigten, dass eine Kalibrierung der verwendeten Systeme auch mit stereoskopischen Visualisierungsverfahren möglich war. Zwar nahmen die Kalibrierabweichungen beim Tragen der Polfilter- bzw. Shutterbrille zu, die Unterschiede waren jedoch nicht signifikant. In einem weiteren Schritt wurden die Blickortabweichungen zum Fixationsobjekt berechnet. Die Abweichungen in den Bedingungen mit 3D-Brille waren für beide Eyetracker größer als ohne, erreichten jedoch nur beim EAS das Signifikanzniveau.

Die Distanzen der Fixationsorte zu den Fixationsobjekten verhielten sich ähnlich. Ohne zusätzliche 3D-Brillen waren die Abweichungen am kleinsten und stiegen bei Verwendung einer Polarisationsfilter- bzw. Shutterbrille an. Signifikante Unterschiede wurden allerdings nur für das EAS nachgewiesen. Die Fixationsortabweichungen waren geringer als die der einzelnen Blickorte. Der Grund hierfür lag in den Berechnungsmethoden. Während für die Bestimmung der Blickortabweichung jede Differenz zwischen Blickort und Zielreiz eingeht, werden durch die Fixationserkennungsalgorithmen mehrere Samples zu einer Fixation zusammengefasst (Komogortsev, Gobert, Jayarathna, Koh & Gowda, 2010). Hierbei finden Gewichtungen in Richtung Fixationszentrum statt, so dass ‚Ausreißer‘ weniger Einfluss auf das Gesamtergebnis haben. Die Abweichungen der Fixationsorte in den Messungen ohne stereoskopische Visualisierungstechnologie stimmten mit den Herstellerangaben zur örtlichen Genauigkeit der Eyetracking-Systeme überein und deuteten auf eine korrekte Aufzeichnung der Blickdaten hin.

Für beide Eyetracker waren unterschiedliche Blickort- und Fixationsortabweichungen in verschiedenen Monitorbereichen nachweisbar. Tendenziell war die Messung im Zentrum des Bildschirms genauer als in den Randbereichen, auch ohne stereoskopische Visualisierung. Weshalb die stereoskopischen Visualisierungsverfahren einen Einfluss auf die Messgenauigkeit in Bezug auf Sampledaten bzw. Fixationen aber nicht auf die Kalibrierung des Eyetrackers hatten, kann nicht abschließend geklärt werden. Dafür sind die erhobenen Daten nicht aussagekräftig und es wäre die Durchführung einer weiteren Studie in Hinblick auf die Kalibrierergebnisse notwendig.

Die Auswertung der Pupillengröße ergab, dass diese in den Bedingungen mit 3D-Brille zunahm. Die Pupillenerweiterung lässt sich durch den geringeren Lichteinfall erklären, welcher aus der Filterung nicht ausgerichteter Lichtstrahlen beim Polarisationsfilter- bzw. einer Halbierung der einfallenden Lichtmenge beim Shutterverfahren resultiert. Um zu prüfen,

inwieweit dabei visuell wahrnehmbare Informationen verloren gehen, müsste eine entsprechende Untersuchung durchgeführt werden, bei der beispielsweise die Wahrnehmbarkeit von Reizen systematisch manipuliert wird. Signifikant verschiedene Pupillengrößen in Bezug auf unterschiedliche Monitorbereiche waren nur für das EAS nachweisbar, allerdings können diese nach Herstellerangaben beim ELS ebenfalls blickortabhängig variieren. Erklärt wird dies durch schwankende Abstände der Augen zu den Kameras bei Rotationsbewegungen, optischen Verzerrungen der Pupille im Videokamerabild und kameraspezifischen Faktoren (SR Research, 2009). Der Hersteller empfiehlt bei derartigen Untersuchungen die Augen nicht zu bewegen und die Stimuli an derselben bzw. ausbalanciert an vergleichbaren Positionen zu präsentieren.

Zusammenfassend lässt sich konstatieren, dass zweidimensionale Blickbewegungsmessungen auch in Kombination mit stereoskopischen Visualisierungen mit den getesteten Technologien durchführbar sind. Beachtet werden sollte die Verringerung der Messgenauigkeit und Zunahme der Pupillengröße unter Verwendung von Polarisationsfilter- oder Shutterverfahren. Um Blickbewegungen in einer realen oder virtuellen Umgebung räumlich bestimmen zu können, müssen im nächsten Schritt die dafür notwendigen Blickorte dreidimensional berechnet werden.

7 Bestimmung dreidimensionaler Blickpositionen

Die in der Forschung für eine Blickbewegungsmessung eingesetzten Eyetracking-Systeme sind hauptsächlich für zweidimensionale Projektionsflächen wie Bildschirm, Leinwand oder Printmedien konzipiert. Eine Positionsbestimmung der Blickorte und der daraus ableitbaren Blickbewegungen im dreidimensionalen Raum erfordert zusätzlich die Berechnung des Tiefenparameters. Erst dadurch kann natürliches Blickverhalten im Raum analysiert werden. Dieses Kapitel beschreibt die Entwicklung eines Algorithmus zur Berechnung dreidimensionaler Blickorte anhand binokularer Eyetracking-Parameter sowohl in realer als auch virtueller Umgebung.

7.1 Anforderungsprofil eines universellen Algorithmus

Für die Berechnung der 3D-Blickposition gibt es bereits verschiedene Ansätze wie beispielsweise stereogeometrische Verfahren (Duchowski et al., 2000), Kollisionserkennung bzw. *Intersection Points* (Barabas et al., 2004; Duchowski et al., 2002), die Berechnung der Vergenz in einem Head-Mounted-Display (HMD) (Medlin, 2003), neuronale Netze bzw. *Parametrized Self-Organizing Maps* (Essig et al., 2006; Pfeiffer et al., 2007) oder trianguläre Winkelberechnungen (Daugherty et al., 2010; Duchowski et al., 2011). Jedoch sind alle bekannten Berechnungsmethoden zur Bestimmung dreidimensionaler Blickorte speziell auf die verwendete Kombination von Stereoprojektion und Blickbewegungsmessung ausgerichtet. Um einen universell einsetzbaren Algorithmus zur dreidimensionalen Blickortberechnung entwickeln zu können, wird zunächst ein Anforderungsprofil aus vorangegangenen Erkenntnissen abgeleitet.

Anforderung 1: Der Algorithmus erstellt ein räumliches Modell des Versuchsaufbaus und bildet das Blickverhalten des Betrachters realitätsnah ab.

Anforderung 2: Der Algorithmus eignet sich für Messungen sowohl in realer als auch virtueller Umgebung und ist dabei unabhängig von der verwendeten Visualisierungstechnologie (z.B. HMD, 3D-Monitor, Leinwand).

Anforderung 3: Der Algorithmus ist mit verschiedenen Blickbewegungsmessungen einsetzbar (z.B. Head-Mounted, Remote, dark-pupil, bright-pupil).

- Anforderung 4:* Horizontale und vertikale Kopfbewegungen werden vom Algorithmus ebenso berücksichtigt wie in die Tiefe, Drehungen und Neigungen.
- Anforderung 5:* Individuelle physiologische Parameter wie Augenabstand oder Augensymmetrie gehen in die Berechnungen ein.
- Anforderung 6:* Die Bestimmung des dreidimensionalen Blickortes muss in einem Zeitfenster erfolgen, welches die Verwendung der Daten in Echtzeit ermöglicht.

7.2 Ein vektorbasierter Ansatz

Aus *Anforderung 1* ergibt sich, dass ein Modell des optischen Systems der Augen abstrahiert und in den Algorithmus integriert werden muss. Je präziser das Modell die Realität abbildet, desto genauer lassen sich 3D-Blickorte bestimmen. Während des Sehvorgangs wird von Objekten reflektiertes Licht durch die Rezeptoren auf der Netzhaut des Auges wahrgenommen. Da dieses geradlinig verläuft, wurde der dreidimensionale Blickort oftmals als Schnittpunkt der Blickachsen betrachtet. Allerdings schneiden sich in der Realität diese nur in wenigen Fällen, sie liegen überwiegend windschief zueinander. Dementsprechend wird der dreidimensionale Blickort in dem Punkt ermittelt, der den beiden Blickachsen am nächsten liegt. Zur Berechnung werden die die Blickachsen als *Blickvektoren* definiert. Ein Punkt auf dem Blickvektor liegt im Auge bzw. auf der Fovea. Für die Bestimmung der Vektorrichtung ist die Kenntnis über einen zweiten Punkt auf der jeweiligen Blickachse erforderlich. Hierfür eignen sich die Ausgabeparameter binokularer EyeTracker, da derartige Systeme die beiden Blickorte für linkes und rechtes Auge in Bezug auf eine zweidimensionale Messebene getrennt ausgeben. In einem weiteren Schritt ist derjenige Vektor zu berechnen, der den geringsten Abstand zwischen den beiden Blickvektoren hat. Der Mittelpunkt dieses Verbindungsvektors mit geringstem Abstand wird als dreidimensionale Blickposition angenommen.

7.2.1 Mathematisches Modell

Die Blickachsen können jeweils als Gerade g_L und g_R durch zwei Punkte beschrieben werden. Die Ausgangsparameter bilden die beiden Augpunkte \vec{A}_L und \vec{A}_R sowie die jeweiligen Blickorte \vec{B}_L und \vec{B}_R in Messebene. In der analytischen Geometrie lassen sich die Blickachsen als Blickvektoren \vec{X}_L und \vec{X}_R mit den Ortsvektoren \vec{A}_L und \vec{A}_R , den Richtungsvektoren \vec{B}_L und \vec{B}_R und den Parametern t_L und t_R in folgender Form darstellen:

$$g_L: \quad \vec{X}_L = \vec{A}_L + t_L(\vec{B}_L - \vec{A}_L) = \vec{A}_L + t_L \vec{AB}_L \quad \text{mit } t_L \in \mathbb{R} \quad (7.1)$$

$$g_R: \quad \vec{X}_R = \vec{A}_R + t_R(\vec{B}_R - \vec{A}_R) = \vec{A}_R + t_R \vec{AB}_R \quad \text{mit } t_R \in \mathbb{R}$$

Als erstes wird die Lagebeziehung der Blickachsen ermittelt, um die entsprechenden mathematischen Verfahren für die einzelnen Fälle korrekt zu bestimmen. Die beiden Blickachsen können in drei Lagebeziehungen zu einander stehen:

Fall 1: Blickachsen liegen parallel zueinander

Fall 2: Blickachsen schneiden sich

Fall 3: Blickachsen liegen windschief zueinander

1. Schritt: Prüfen auf Kollinearität der Richtungsvektoren. Hierzu wird getestet, ob der Richtungsvektor der einen Blickachse sich durch den Richtungsvektor der zweiten ausdrücken lässt. Falls dem so ist, liegen die Blickachsen parallel zu einander.

$$t_L(\vec{B}_L - \vec{A}_L) = t_R(\vec{B}_R - \vec{A}_R) \quad (7.2)$$

2. Schritt: Fällt die Prüfung auf Kollinearität negativ aus, wird getestet, ob die Blickachsen einen gemeinsamen Schnittpunkt besitzen. Dazu werden durch Gleichsetzungsverfahren die Parameter t_L und t_R ermittelt und diese in die Ausgangsgleichungen (7.1) eingesetzt:

$$\vec{A}_L + t_L \vec{AB}_L = \vec{A}_R + t_R \vec{AB}_R \quad (7.3)$$

Ist das lineare Gleichungssystem (7.3) eindeutig lösbar, existiert genau ein gemeinsamer Schnittpunkt S zur Lösung t_L, t_R :

$$S = \vec{A}_L + t_L(\vec{B}_L - \vec{A}_L) \quad \text{bzw.} \quad (7.4)$$

$$S = \vec{A}_R + t_R(\vec{B}_R - \vec{A}_R)$$

3. Schritt: Kann (7.3) nicht gelöst werden, liegt kein gemeinsamer Schnittpunkt vor, d.h. die Geraden bzw. Blickachsen liegen windschief zueinander. In diesem Falle soll die Position gefunden werden, die beiden Geraden am nächsten liegt. Dazu wird das gemeinsame Lot beider Geraden berechnet, auf dessen Hälfte sich der gesuchte Punkt befindet.

Annahme: Auf der Blickachse des linken Auges g_L befindet sich ein Punkt P und auf der Blickachse des rechten Auges g_R ein Punkt Q . \overrightarrow{PQ} ist der geringste Abstand zwischen den beiden Achsen. Er wird genau dann angenommen, wenn:

$$\overrightarrow{PQ} \perp g_L \Rightarrow \overrightarrow{PQ}(\overrightarrow{B_L} - \overrightarrow{A_L}) = 0 \quad \text{und} \quad (7.5)$$

$$\overrightarrow{PQ} \perp g_R \Rightarrow \overrightarrow{PQ}(\overrightarrow{B_R} - \overrightarrow{A_R}) = 0$$

Somit steht \overrightarrow{PQ} senkrecht auf beiden Blickachsen. \overrightarrow{PQ} ergibt sich aus $\vec{P} - \vec{Q}$ und mit (7.1):

$$\overrightarrow{PQ} = (\overrightarrow{A_L} + t_L \overrightarrow{AB_L}) - (\overrightarrow{A_R} + t_R \overrightarrow{AB_R}) \quad (7.6)$$

Aus (7.5) und (7.6) folgt:

$$\left((\overrightarrow{A_L} + t_L \overrightarrow{AB_L}) - (\overrightarrow{A_R} + t_R \overrightarrow{AB_R}) \right) * (\overrightarrow{B_L} - \overrightarrow{A_L}) = 0 \quad (7.7)$$

und

$$\left((\overrightarrow{A_L} + t_L \overrightarrow{AB_L}) - (\overrightarrow{A_R} + t_R \overrightarrow{AB_R}) \right) * (\overrightarrow{B_R} - \overrightarrow{A_R}) = 0 \quad (7.8)$$

Mittels (7.7) und (7.8) lassen sich t_L und t_R bestimmen, so dass nach Einsetzen in (7.1) die Punkte P und Q berechnet werden können.

Der Halbierungspunkt H_{PQ} ergibt sich aus:

$$H_{PQ} = \frac{1}{2}(P + Q) \quad (7.9)$$

H_{PQ} entspricht dem Punkt, der beiden Blickachsen am nächsten liegt und definiert den Blickort im dreidimensionalen Raum.

7.2.2 Eyetracking-Parameter

Nach dem vorgestellten vektorbasierten Ansatz sind jeweils zwei pro Blickachse identifizierbare, dreidimensionale Punkte notwendig, um die Blickvektoren für linkes und rechtes Auge abzuleiten. Hierfür können nur Eyetracking-Systeme eingesetzt werden, die

Informationen für beide Augen getrennt ausgeben. Jedoch sind die Ausgabewerte herstellerspezifisch, eine standardisierte Schnittstelle für alle Modelle ist derzeit nicht gegeben. Deshalb müssen die für die Bildung der Blickvektoren benötigten Parameter in Abhängigkeit der vom Eyetracking-System zur Verfügung gestellten Daten ausgewählt werden. Ein binokulares Eyetracking-System gibt zumindest die getrennten Blickortkoordinaten für linkes und rechtes Auge bezogen auf eine zweidimensionale Bezugsebene aus, beispielsweise die Bildschirmkoordinaten eines Monitors. Diese Positionsangaben können auch dreidimensional in das Weltkoordinatensystem transformiert werden. Damit lässt sich der erste Punkt einer Blickachse bestimmen. Zusätzlich benötigt man einen zweiten Punkt, um den Blickvektor bilden zu können.

Fixe Augenposition (FixEye)

Sollten keine weiteren Parameter verfügbar sein, können die Positionen für linkes und rechtes Auge manuell ausgemessen und während des Experimentes konstant gehalten werden. Hierbei ist die Verwendung einer Kinnstütze sinnvoll.

Variable Augenposition (VarEye)

Gibt ein Eyetracking-System weitere blickachsenrelevanter Positionsangaben aus, sind diese unmittelbar für die Berechnung der Blickvektoren nutzbar. Hierbei handelt es sich beispielsweise um Informationen zur Lokalisation der Pupille, Linse oder dem Augapfelzentrum. Stehen diese Daten zeitlich synchron zu den Blickkoordinaten zur Verfügung, kann der Proband während der Messung den Kopf frei bewegen.

Blickvektoren (GazeVec)

Unterstützt das Eyetracking-System die Ausgabe der beiden Blickvektoren, werden diese direkt dem Algorithmus zur Berechnung der 3D Blickposition übergeben. Auch Kombinationen aus Orts- und Richtungsvektor sind möglich. Ausgehend von einem Bezugspunkt auf der Blickachse wie beispielsweise des Cornea-Zentrums und einem Einheitsvektor als Richtungsangabe ist die Berechnung des Blickvektors möglich. Wie bei *VarEye* kann bei zeitlich synchroner Parameterausgabe der Kopf frei bewegt werden.

7.2.3 Koordinatensysteme

In einer Blickbewegungsmessung treffen unterschiedliche Koordinatensysteme aufeinander, die im Setting zusammengebracht werden müssen. So befindet sich der Proband im Weltkoordinatensystem und betrachtet den Monitor aus seiner Sicht symmetrisch in einem Abstand von ca. 65 cm. Der Monitor bzw. der darauf abgebildete visuelle Inhalt wird in

Pixelkoordinaten angegeben wobei der Koordinatenursprung in der linken oberen Monitorecke definiert ist. Auch das Eyetracking-System hat ein eigenes herstellerspezifisches Koordinatensystem, in dem z.B. die Positionsdaten des Pupillenzentrums berechnet werden. Unterscheidungsmerkmale der beteiligten Koordinatensysteme sind die Dimension (2D, 3D), Maßeinheiten (px, mm) und Koordinatenursprung. Um das geometrische Modell aus Kapitel 7.2.1 anwenden zu können, müssen die zur Verfügung stehenden Parameter in ein gemeinsames, dreidimensionales Koordinatensystem transformiert werden. Maßeinheit und Koordinatenursprung sind dabei frei wählbar.

7.2.4 Gaze3D.dll

Der Algorithmus aus Kapitel 7.2.1 wurde in eine Funktionsbibliothek bzw. *Dynamic Link Library* (DLL) implementiert und kann dadurch in verschiedenen Programmierumgebungen geladen werden. Innerhalb dieser Bibliothek sind die mathematischen Berechnungen in Funktionen hinterlegt. Je nach Art und Umfang der zur Verfügung stehenden Eyetracking-Parameter wird die geeignete Berechnungsmethode ausgewählt. Zur Berechnung des dreidimensionalen Blickortes bei vorgegebenen Augenpositionen (*FixEye*, *VarEye*) wird die Funktion *CalculateGaze3DFromEyePos* mit der in Listing 7-1 angeführten Syntax aufgerufen. Übergeben werden hierbei die räumlichen Augenpositionen sowie die binokularen Blickortkoordinaten mit dem z-Wert der Eyetracking-Messebene. Voraussetzung für eine korrekte Berechnung ist die Transformation aller Positionsangaben in ein gemeinsames 3D-Koordinatensystem.

```
function CalculateGaze3DFromEyePos( x_EyePos_Left,
                                   y_EyePos_Left,
                                   z_EyePos_Left,

                                   x_GazePos_Left,
                                   y_GazePos_Left,
                                   z_GazePos_Left,

                                   x_EyePos_Right,
                                   y_EyePos_Right,
                                   z_EyePos_Right,

                                   x_GazePos_Right,
                                   y_GazePos_Right,
                                   z_GazePos_Right: Int32;

  out  x_Gaze3D,
       y_Gaze3D,
       z_Gaze3D: Int32): Int32; stdcall;
```

Listing 7-1: Funktionsaufruf zur Berechnung der 3D-Blickortposition aus gegebenen Augenpositionen und binokularen Blickortkoordinaten. Die Rückgabewerte beinhalten den dreidimensionalen Blickort.

Folgendes Beispiel soll die Anwendung des Algorithmus veranschaulichen: Der gemeinsame Koordinatenursprung wird in die Mitte des Bildschirms gelegt, wobei x-Koordinaten nach

rechts, y-Koordinaten nach oben und z-Koordinaten zum Betrachter hin positiv ausgerichtet sind. Alle Maße sind in Millimeter angegeben. Die Versuchsperson hat eine symmetrische Pupillendistanz von 66 mm. Sie wird im Versuchsaufbau so positioniert, dass beide Augen 700 mm vom Monitor entfernt sind und den gleichen horizontalen Abstand zum Ursprung haben. Die Augenhöhe liegt 100 mm oberhalb der Bildschirmmitte. Der binokulare Eyetracker gibt folgende 2D-Koordinaten (mm) aus: Blickort des linken Auges (60,50), Blickort des rechten Auges (-60,50). Da sich der Koordinatenursprung auch innerhalb der Eyetracking-Messebene befindet, sind die z-Werte in beiden Fällen 0. Aus diesem Setting ergeben sich folgende Übergabeparameter für den Algorithmus: Position des linken Auges (-33,100,700); Blickort des linken Auges (60,50,0); Position des rechten Auges (33,100,700); Blickort des rechten Auges (-60,50,0). Als Rückgabewert erhält man die Koordinaten des dreidimensionalen Blickortes (0,82,496).

Unterstützt das Eyetracking-System die Ausgabe beider Blickvektoren, kann die in Listing 7-2 aufgeführte Funktion *CalculateGaze3DFromGazeVec* verwendet werden. Auch hier wird die berechnete 3D-Blickortposition zurückgegeben.

```
function CalculateGaze3DFromGazeVec(  x_EyePos_Left,
                                     y_EyePos_Left,
                                     z_EyePos_Left,

                                     x_GazeVec_Left,
                                     y_GazeVec_Left,
                                     z_GazeVec_Left,

                                     x_EyePos_Right,
                                     y_EyePos_Right,
                                     z_EyePos_Right,

                                     x_GazeVec_Right,
                                     y_GazeVec_Right,
                                     z_GazeVec_Right: Int32;

    out  x_Gaze3D,
         y_Gaze3D,
         z_Gaze3D: Int32): Int32; stdcall;
```

Listing 7-2: Funktionsaufruf zur Berechnung der 3D Blickortposition aus gegebenen Blickvektoren. Es wird der dreidimensionale Blickort zurückgegeben.

Der Vergenzwinkel ergibt sich aus dem Winkel zwischen den beiden visuellen Achsen. Fixiert eine Person einen Punkt in unendlicher Entfernung, stehen die Augen parallel und der Vergenzwinkel ist Null. Der Winkel wird umso größer, je mehr die Augen konvergieren. Bei symmetrischer Konvergenz ist der horizontale Vergenzwinkel φ abhängig von der interokularen Distanz (Augenabstand) a und dem Abstand zum Fixationsort D ausgehend vom Mittelpunkt zwischen den Augen (Daugherty et al., 2010):

$$\tan\left(\frac{\varphi}{2}\right) = \frac{a}{2D} \quad (7.10)$$

Dadurch wirken sich Entfernungsänderungen bzgl. des zu fixierenden Objektes D im Nahbereich stärker aus. Nach dem vektorbasierten Ansatz lässt sich der Winkel zwischen den beiden Blickvektoren \vec{a} und \vec{b} anhand des Skalarproduktes berechnen:

$$\cos \varphi = \frac{\vec{a} * \vec{b}}{|\vec{a}| * |\vec{b}|} \quad (7.11)$$

Die beschriebenen Rechenoperationen sind in der Funktion *CalculateVergenceAngle* umgesetzt und können mit dem im Listing 7-3 beschriebenen Aufruf ausgeführt werden.

```
function CalculateVergenceAngle( x_GazeVec_Left,
                               y_GazeVec_Left,
                               z_GazeVec_Left,

                               x_GazeVec_Right,
                               y_GazeVec_Right,
                               z_GazeVec_Right: Int32;

                               out VergenceAngle: Double): Int32; stdcall;
```

Listing 7-3: Funktionsaufruf zur Berechnung des Winkels zwischen den Blickvektoren für linkes und rechtes Auge.

7.2.5 Performance-Test

In einem Leistungstest wurde geprüft, wie viel Zeit die einzelnen Funktionen zur Berechnung der dreidimensionalen Blickortkoordinaten benötigen. Es kam ein Desktop-PC mit folgender Konfiguration zum Einsatz:

CPU	Intel Core™ 2 Duo P8600, 2 x 2.4 GHz
RAM	2 x 2 GB DDRII, 800
HDD	500 GB, 7.200
Grafik	NVidia GeForce 500 Series 1 GB DDRIII

Während des Tests wurde jede Funktion aufgerufen und die Zeit vom Aufruf bis zur Rückgabe der Ergebniswerte gemessen. Dieser Vorgang wurde 1 Million Mal wiederholt und anschließend der Durchschnittswert der Berechnungsdauer ermittelt, um durch die Computerhardware bedingte Performanceschwankungen zu kompensieren. Die durchschnittliche Berechnungsdauer der Funktionen lag im Performancetest unter 10^{-3} ms (siehe Tab. 7-1). Verwendet man zur stereoskopischen Visualisierung beispielsweise einen Monitor mit 120 Hz, ergibt sich für jedes dargestellte Bild ein Zeitfenster von 8,3 ms. Damit sind die Funktionen zwischen zwei Bildanzeigen vollständig durchführbar.

Tab. 7-1: Ergebnisse des Performancetests zur Überprüfung der implementierten Funktionen

Funktion	Zeit
CalculateGaze3DFromEyePos	$1,09 * 10^{-3}$ ms
CalculateGaze3DFromGazeVec	$8,89 * 10^{-4}$ ms
CalculateVergenceAngle	$2,34 * 10^{-4}$ ms

7.2.6 Charakteristik der Berechnungsmethode

Das Blickverhalten wird identisch abgebildet (*Anforderung 1*). Die Berechnung benötigt keinerlei Informationen über die eingesetzte Visualisierungstechnologie (*Anforderung 2*). Der Algorithmus ist von der verwendeten Eyetracking-Technologie unabhängig, sofern die Daten binokular ausgegeben werden (*Anforderung 3*). Da die Augenpositionen für linkes und rechtes Auge getrennt in die Berechnungen einfließen, wird implizit auch Augenabstand und Symmetrie berücksichtigt (*Anforderung 4*). Über die dreidimensionalen Augenpositionen werden ebenfalls die Kopfposition bzw. sämtliche Bewegungen in allen Freiheitsgraden erfasst (*Anforderung 5*). Die Funktionen sind echtzeitfähig und die Ergebnisse können in die Generierung des nächsten Frames einbezogen werden (*Anforderung 6*). Damit erfüllt der vektorbasierte Ansatz alle Anforderungen aus Kapitel 7.1.

8 Studie II: Eyetracking in realer Welt

Innerhalb dieser Studie sollten dreidimensionale Blickorte anhand des unter Kapitel 7.2 vorgestellten vektorbasierten Ansatzes bestimmt werden. Dazu wurden Probanden in einem Experiment reale 3D-Objekte dargeboten und gleichzeitig Blickdaten mit einem binokularen Remote Eyetracker erhoben. Anschließend erfolgte die Berechnung der dreidimensionalen Blickorte mit manuell ausgemessener Augenposition sowie den vom Eyetracking-System ausgegebenen Informationen zur Cornea-Position und den Blickvektoren. Je nach Verfügbarkeit der Eyetracking-Parameter wurden für *FixEye*, *VarEye* und *GazeVec* die zugehörigen Algorithmen eingesetzt.

8.1 Fragestellung und Hypothesen

Das Hauptaugenmerk der Analysen bezog sich auf die Genauigkeit der durch die drei Algorithmen berechneten 3D-Blickorte. Hierbei sollten die Abweichung bezüglich der Positionen im Messraum und über den Verlauf der Zeit untersucht werden. Anhand der Abweichungswerte lässt sich beispielsweise in späteren Untersuchungen die notwendige Minimalgröße der Szenenobjekte ableiten, um eine korrekte Zuordnung zwischen der Blickposition und dem fokussierten Objekt vornehmen zu können. In einem weiteren Schritt wurde der Tiefenparameter als gemittelte Blickortentfernung ohne Beachtung der horizontalen und vertikalen Komponente berechnet, um die Genauigkeit der Fokussierdistanz zu analysieren.

Es ist anzunehmen, dass die 3D-Blickortabweichung in Abhängigkeit der Fokussierebene variiert, da mit zunehmender Entfernung die Blickachsen annähernd parallel verlaufen und sich geringe Schwankungen in den Vektoren verstärkt auf die berechnete Blickposition auswirken. Hierfür wurden die Vergenzwinkel in den Tiefenebenen untersucht.

Bei Messungen, in denen ein Eyetracking-System keine Informationen bezüglich der Augenparameter oder Blickvektoren zur Verfügung stellt, kann der Versuchsleiter die interokulare Pupillendistanz manuell bestimmen. Da sich die Genauigkeit dieser Messung auf die Berechnung des Blickortes auswirkt, wurden die vom Versuchsleiter erhobenen Pupillenabstände mit den vom Eyetracker ausgegebenen Augenpositionen verglichen.

8.2 Methode

In einem Experiment wurden Blickdaten beim visuellen Fixieren realer Objekte erhoben. In mehreren Durchgängen erfolgte die Darbietung der Fixationsobjekte an unterschiedlichen Positionen im Raum. Für die Blickbewegungsmessung wurde ein binokulares Remote-Eyetracking-System eingesetzt, dessen Ausgabewerte die Berechnung der dreidimensionalen Blickortkoordinaten nach *FixEye*, *VarEye* und *GazeVec* unterstützt.

Stichprobe

An der Studie nahmen 12 Versuchspersonen (5 weiblich) im Alter zwischen 26 und 47 Jahren teil. Das Durchschnittsalter lag bei 33 Jahren. Alle Versuchsteilnehmer hatten normales oder durch eine Brille bzw. Kontaktlinsen korrigiertes Sehvermögen.

Apparatus und Aufbau

Die Blickbewegungsmessung wurde mit einem *EyeFollower 2.1 System* (EFS) des Herstellers LC Technologies, Inc. durchgeführt. Der für die Kalibrierung des Eyetrackers auf eine Versuchsperson notwendige Monitor hatte eine Abmessung von 534 x 300 mm bei einer Bildschirmdiagonale von 24", eine Auflösung von 1.920 x 1.080 px und eine Bildwiederholfrequenz von 60 Hz. Die Projektionsfläche des Monitors stellte gleichzeitig die Bezugsebene für ausgegebene 2D-Blickpositionen dar. Die Besonderheit des EFS liegt in der Kompensation von Kopfbewegungen. Um durch hochauflösende Eyetracking-Kameras eine örtliche Genauigkeit von 0.4° über einen durch die Kameras erfassbaren Messbereich hinaus aufrechtzuerhalten, werden die Augenkameras in Echtzeit den Kopfbewegungen des Probanden nachgeführt. Dazu sind im Eyetracking-System zwei weitere Kameras integriert, deren Videobilder in Kombination mit einer Gesichtserkennungssoftware die Berechnung der dreidimensionalen Position des Kopfes bzw. der Augen ermöglichen. Bei Kopfbewegungen über definierte Toleranzgrenzen hinaus werden die Augenkameras neu ausgerichtet. Das Eyetracking-Interface unterstützt sowohl die Ausgabe der Blickortkoordinaten als auch Informationen bzgl. 3D-Position der Cornea-Zentren sowie der Blickvektoren für linkes und rechtes Auge.

Der Abstand vom Kalibrierbildschirm des EFS zur Cornea-Oberfläche des Probanden betrug 700 mm und wurde mittels einer Kinnstütze konstant gehalten. Zwischen Proband und Eyetracking-System positionierte der Versuchsleiter während des Experimentes die Fixationsobjekte an unterschiedliche Positionen im Messraum (5 Tiefenebenen x 3 Spalten x 3 Zeilen). Die erste Fixationsebene (E1) befand sich 100 mm ausgehend von der

Kalibrierebene bzw. 600 mm vom Probanden entfernt. In 100 mm Abständen folgten weitere vier Ebenen, so dass bis zu einer Nahdistanz von $z = 200$ mm (E5) gemessen werden konnte (siehe Abb. 8-1).

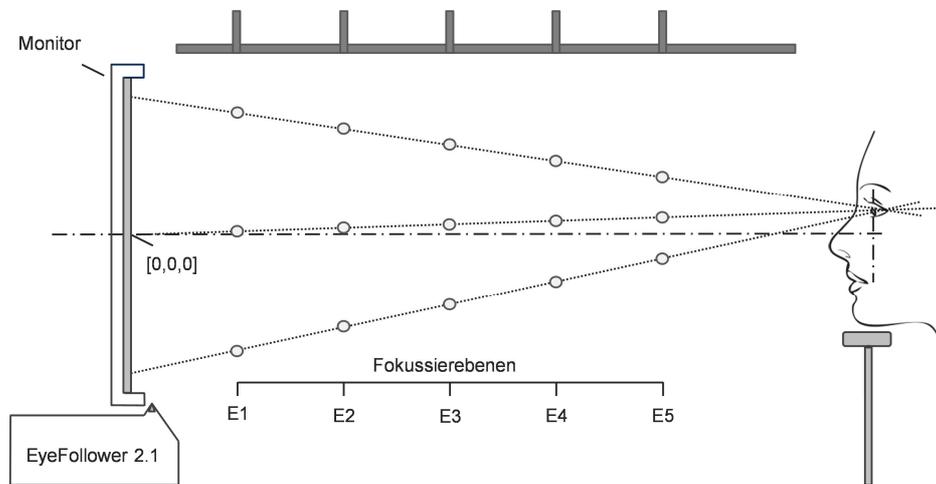


Abb. 8-1: Versuchsaufbau mit Eyetracking-System, Kalibrierbildschirm, Proband und Fixationsobjekten in den fünf Fokussierebenen.

Während der Messung durfte sich nur ein Fixationsobjekt im Messraum befinden, da sich andernfalls hintereinander liegende gegenseitig verdecken. Aus diesem Grund wurde die in Abb. 8-2 dargestellte mechanische Konstruktion konzipiert und hergestellt, um ein präzises Positionieren der Fixationsobjekte in den Messraum hinein und auch wieder hinaus zu ermöglichen. Zur Überprüfung der unter Kapitel 7.2.2 beschriebenen Ausgangsbedingungen musste für *FixEye* eine feste Augenposition vorgegeben werden. Zur Stabilisierung des Kopfes kam deshalb die bereits erwähnte Kinnstütze zum Einsatz, um den möglichen Einfluss von Kopfbewegungen auf die Messgenauigkeit auch in den Bedingungen *VarEye* und *GazeVec* auszuschließen. Für eine gemeinsame Berechnungsgrundlage wurde ein 3D-Weltkoordinatensystem wie folgt definiert: Das Koordinatensystem umfasst drei Dimensionen (x, y, z), Maßeinheit ist Millimeter, der Ursprung wurde in die Mitte der Bezugsebene des Eyetracking-Systems bzw. der Kalibrierfläche gelegt.

Stimulusmaterial

Während des Experimentes sollten Probanden die Fixationsobjekte an 45 verschiedenen Positionen im dreidimensionalen Raum fokussieren: 5 Tiefenebenen (E) mit je 3 Spalten (S) x 3 Zeilen (Z) pro Ebene. Damit konnte der dreidimensionale Messraum (Frustum) zwischen den Augen des Probanden und der Bezugsebene des Eyetracking-Systems vollständig erfasst werden. Die Fixationsobjekte bestanden aus Kugeln mit einem Durchmesser von 4 mm und wurden zum besseren Erkennen bzw. Fixieren mit einem Marker versehen (Abb. 8-2 C).

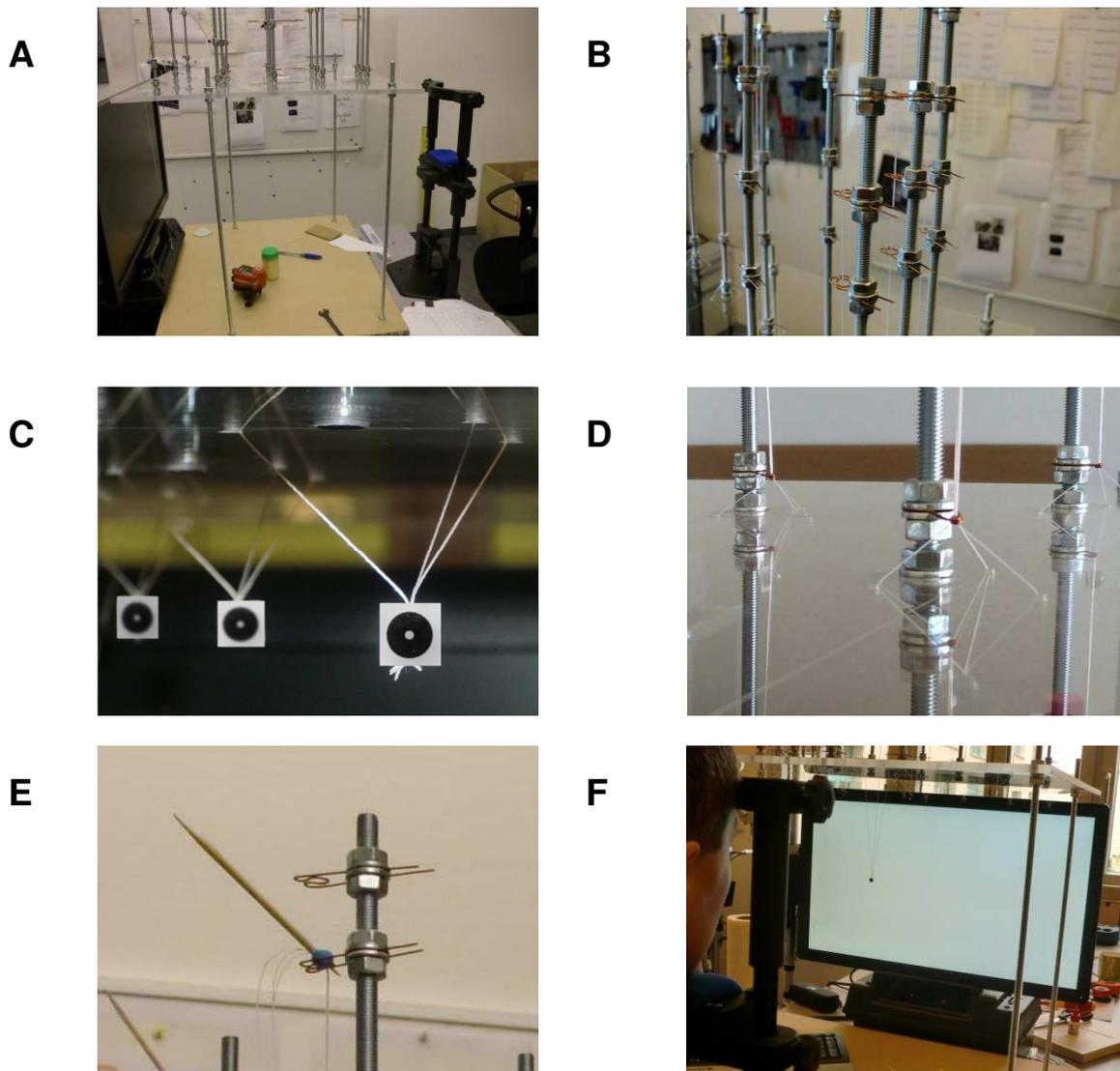


Abb. 8-2: Experiment-Aufbau: Eine Acrylscheibe als Grundplatte für die Fixationsobjektaufhängungen ist zwischen Monitor und Augen der Versuchsperson oberhalb des Messbereiches horizontal angeordnet (A). Jede Fixationsobjektaufhängung besteht aus einer Gewindestange mit in drei unterschiedlichen Höhen befestigte Einhängeösen (B). Die Fixationsobjekte sind unterhalb der Acrylglasplatte an je 3 Fäden aufgehängt (C), um ein Pendeln während der Messung zu verhindern. Die 3 Fäden werden durch die Platte hindurch zur jeweiligen Aufhängevorrichtung geführt und durch eine Führungsöse (D) gebündelt. Die Fäden halten am oberen Ende durch eine Kunststoffperle zusammen, welche in die Ösen der Aufhängevorrichtung nach Bedarf eingehängt werden kann (E). Ein an die Kunststoffperle angebrachtes Holzstäbchen erleichtert die Bedienung während der Positionswechsel. Durch die unterschiedlichen Höhen der Einhängeösen wird das am unteren Ende hängende Fixationsobjekt in die exakte Position gebracht. Für eine gleichmäßige Helligkeit und einen definierten Kontrast ist im Hintergrund der Kalibrierbildschirm mit einem Grauton ausgeleuchtet (F). Mit den 15 Fixationsobjektaufhängungen und jeweils drei Einhängeösen sind 45 Messpositionen gegeben.

Aufzeichnungssoftware

Zur Durchführung der Studie war die Konzeption und Entwicklung eine Software notwendig, um durch das Experiment zu leiten, die Blickbewegungsmessung zu steuern, Eyetracking-Parameter abzufragen, 3D-Blickorte zu berechnen und sämtliche Daten zeitlich synchronisiert zu speichern. Dazu wurde die Steuerung des Eyetracking-Systems (Initialisierung, Start, Stopp, Kalibrierfunktion) implementiert und die Schnittstelle zur Abfrage der Ausgabeparameter angebunden. Vor Beginn des Experiments waren Versuchspersonen-Code, Speicherort

aufzuzeichnender Daten, individueller Augenabstand und Symmetrie sowie die Kalibrierbildschirmparameter manuell in die Benutzeroberfläche einzugeben (Abb. 8-3). Nach erfolgreicher Kalibrierung und Darbietung der Instruktionsanweisung startete die Blickbewegungsmessung. Die zu präsentierenden Stimuli wurden dem Versuchsleiter zur Positionierung auf einem separaten Monitor angezeigt.

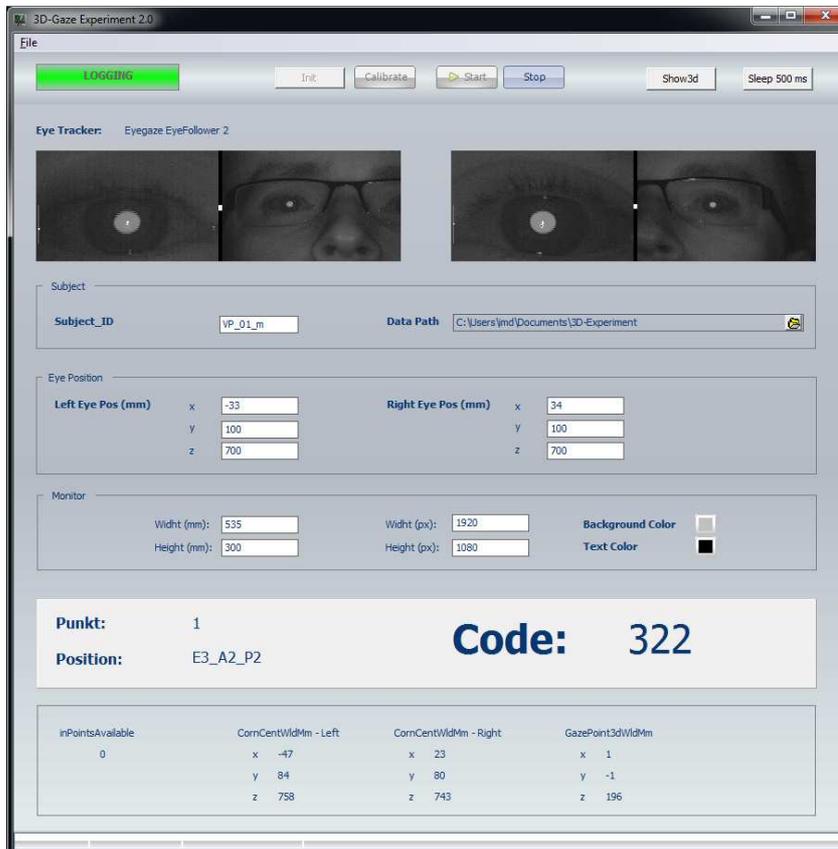


Abb. 8-3: Grafische Benutzeroberfläche der Software „3D-Gaze Experiment 2.0“ zur Berechnung und Aufzeichnung dreidimensionaler Blickorte.

Während der Aufzeichnungsphase las das Programm die aktuellen Blickdaten aus einem Ringpuffer des Eyetracking-Systems aus und übergab diese zusammen mit den 3D-Augenpositionen durch entsprechende Funktionsaufrufe an die Gaze3D.dll. In der Bedingung *FixEye* wurden die manuell erhobenen physiologischen Daten und in *VarEye* die vom Eyetracker ausgegebenen Augenpositionen an *CalculateGaze3DFromEyePos* übergeben. Für *GazeVec* gingen die vom EFS ausgegebenen Blickvektoren in die Berechnungen von *CalculateGaze3DFromGazeVec* direkt ein (Tab. 8-1). Die ermittelten 3D-Blickorte wurden zeitlich synchronisiert zu den angezeigten Stimuli und Tastatureingaben zur späteren Auswertung gespeichert.

Tab. 8-1: Berechnungsgrundlage für die drei unterschiedlichen Blickortbestimmungen

	Beschreibung	Eyetracking-Parameter (EFS)	Funktionsaufruf
<i>FixEyePos</i>	Die Augenpositionen werden manuell ausgemessen und bleiben während der Datenerhebung unverändert.	2D Blickortkoordinaten in Monitorebene (Pixel)	<i>CalculateGaze3D</i> <i>FromEyePos</i>
<i>VarEyePos</i>	Das Eyetracking-System gibt während der Messung Daten über die Position beider Augen aus.	3D-Position der Cornea (mm) 2D Blickortkoordinaten in Monitorebene (Pixel)	<i>CalculateGaze3D</i> <i>FromEyePos</i>
<i>GazeVec</i>	Das Eyetracking-System gibt die Blickvektoren für linkes und rechtes Auge getrennt aus.	3D-Position der Cornea (mm) Blickrichtungsvektoren für linkes und rechtes Auge	<i>CalculateGaze3D</i> <i>FromGazeVec</i>

Versuchsablauf

Vor Beginn des Experimentes beantworteten alle Probanden einen Fragebogen zu demographischen Daten, Sehvermögen, Eyetracking-Erfahrung und Motivation. Anschließend wurde vom Versuchsleiter die Augensymmetrie mit einer Augenabstandsschablone ausgemessen und mit dem Versuchspersonencode in die Software eingetragen. Nach Instruktion des Probanden bzgl. des Ablaufs sowie der Aufgaben im Experiment nahm dieser innerhalb der Versuchsumgebung Platz. Nun richtete der Versuchsleiter die Kinnstütze so aus, dass die Augenhöhe 50 mm unterhalb des oberen Monitorrandes lag und startete die Kalibrierprozedur. Dabei kam die 2D-Kalibrierung des Herstellers zum Einsatz, wobei wie schon in Studie I ein Kalibrierpunkt an 9 verschiedenen Positionen auf dem Bildschirm mit den Augen fixiert werden sollte. Bei einer Kalibrierabweichung größer 0.3° wurde die Kalibrierung wiederholt andernfalls mit dem Experiment fortgefahren. Anschließend wurde ein Fixationskreuz auf dem Bildschirm dargeboten und dabei die dreidimensionalen Augenpositionen der Versuchsperson durch das Eyetracking-System erhoben.

Nach einem Testdurchlauf sowie der Klärung noch offener Fragen wurde erneut kalibriert und der Experimentdurchgang gestartet. Dazu positionierte der Versuchsleiter das erste Fixationsobjekt an einen durch die Software vorgegebenen Ort innerhalb des Messraums. Die Aufgabe des Probanden bestand darin, den jeweiligen Stimulus so exakt wie möglich mit den Augen zu fixieren und dies durch Betätigung der Leertaste zu bestätigen. Ab diesem Zeitpunkt sollte die aktuelle Fixation so lange aufrechterhalten werden, bis nach 1.000 ms ein akustisches Signal erklang. Nun konnte die Versuchsperson einen beliebig anderen Blickort zum Entspannen der Augen wählen, während der Versuchsleiter das Fixationsobjekt aus dem Messraum entfernte und im nächsten Schritt ein neues positionierte. Dieser Vorgang wurde für alle 45 Messpunkte wiederholt.

Datenaufbereitung

Bevor mit der Auswertung der Daten begonnen werden konnte, wurden die relevanten Segmente aus dem Datenstrom selektiert, in denen eine Versuchsperson das Fixationsobjekt nach eigenem Empfinden fixierte. Dies betraf alle Samples zwischen dem Drücken der Leertaste und dem Ertönen des akustischen Signals. Anschließend wurden alle ungültigen Werte entfernt, die auf Blinks oder einen Blickort außerhalb des Messbereichs zurückzuführen waren (6,15%). Während der Filterung fiel auf, dass es sowohl bei Messungen in Ebene 1 – Spalte 1 – Zeile 1 als auch in Ebene 1 – Spalte 3 – Zeile 1 häufiger zu ungültigen Messungen kam als in anderen Positionen. Dies ist allein durch Blinks nicht erklärbar. Bei Betrachtung der Video-Augenbilder wurde ersichtlich, dass an den Grenzen des Eyetracking-Messbereichs der Winkel zwischen Eyetracking-Kameras und Blickort ausgehend vom Auge des Betrachters so groß werden konnte, dass sich der Cornea-Reflexpunkt auf die Sklera verschob und dabei Mehrfachreflexionen auftraten (Abb. 8-4 B). Dementsprechend gab es an diesen Positionen weniger gültige Datensamples.

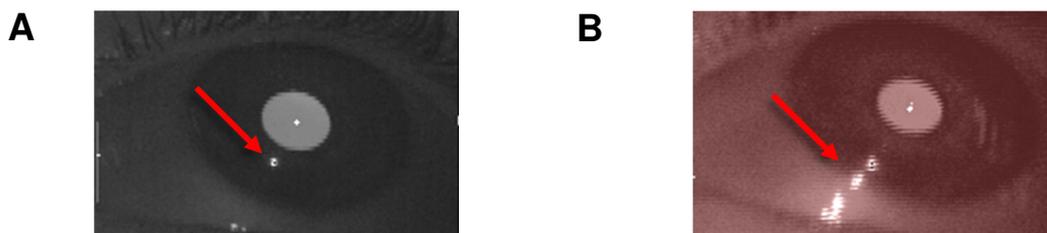


Abb. 8-4: Kamerabilder des Eyetracking-Systems: (A) Eindeutig identifizierbarer Cornea-Reflexpunkt. (B) Wird der Winkel zwischen dem Blickvektor und optischer Achse der Kamera zu groß bzw. liegt der Blickort oberhalb des Messbereiches, verschiebt sich der Cornea-Reflexpunkt auf die Sklera und es kommt zu Mehrfachreflexionen. Eine eindeutige Lokalisation des Cornea-Reflexpunktes ist hierbei nicht möglich.

8.3 Ergebnisse

3D Blickortkoordinaten

In einer einfaktoriellen (Algorithmus) ANOVA mit Messwiederholung konnten höchst signifikante Unterschiede in den 3D-Abweichungen nachgewiesen werden, $F(2,22) = 37.839$, $p < .001$, $\varepsilon = .503$, $\eta^2 = .775$. Die dreidimensionalen Abweichungen der Sampledaten waren im Algorithmus *Gaze3dFromFixEyePos* mit 16 mm am geringsten und nahmen in *Gaze3dFromVarEyePos* mit 23 mm bzw. *Gaze3dFromGazeVec* mit 22 mm zu (Abb. 8-5 A). Obwohl die Abweichungswerte für *Gaze3dFromVarEyePos* bzw. *Gaze3dFromGazeVec* nahe beieinander lagen ergab die statistische Auswertung signifikante Unterschiede. Eine Erklärung für diesen Effekt wäre eine systematische Abweichung zwischen den Algorithmen. Zur genaueren Betrachtung wurden deshalb die 3D-Abweichungen jeder einzelnen

Versuchsperson in beiden Bedingungen von einander subtrahiert. Die Differenzwerte ($Gaze3dFromVarEyePos - Gaze3dFromGazeVec$) zeigten eine durchgehend positive Ausprägung und bestätigen die Annahme der systematischen Abweichung.

Genauigkeit der Blickorte in den unterschiedlichen Tiefenebenen

In einer 3 (Algorithmus) x 5 (Tiefenebene) faktoriellen ANOVA mit Messwiederholung konnten höchst signifikante Unterschiede für Algorithmus, $F(2,22) = 48.701$, $p < .001$, $\eta^2 = .543$, Tiefenebene, $F(4,44) = 33.326$, $p < .001$, $\eta^2 = .091$, sowie die Interaktion Algorithmus x Tiefenebene, $F(8,88) = 15.247$, $p < .001$, $\eta^2 = .124$, nachgewiesen werden. Wechselwirkungen zwischen verschiedenen Algorithmen in unterschiedlichen Ebenen waren vorhanden, jedoch nicht systematisch. Wie man Abb. 8-5 B entnehmen kann, sind die Abweichungen im Nahbereich mit 8 – 16 mm am geringsten. Mit zunehmendem Abstand zwischen Fixationsobjekt und Betrachter verringerte sich die Genauigkeit der dreidimensionalen Blickortpositionen auf 35 – 36 mm.

Genauigkeit der Blickorte innerhalb des Messzeitraumes

Für die Überprüfung, ob sich die Genauigkeit in Abhängigkeit der Dauer der Fixierung ändert, wurde der 1.000 ms Messzeitraum in 10 Zeitfenster a 100 ms unterteilt. Die durchgeführte 3 (Algorithmus) x 10 (Zeitfenster) ANOVA mit Messwiederholung zeigte keine signifikanten Effekte bzgl. Algorithmus, $F(2,22) = 1.532$, $p = .240$, und Zeitfenster, $F(9,99) = 1.018$, $p = .431$. Auch Unterteilungen in 5 mal 200 ms bzw. 2 mal 500 ms ergaben keine Unterschiede.

Genauigkeit der Blickorte über den Verlauf des Experimentes

Um zu prüfen, ob sich die Blickortgenauigkeit im Verlauf des Experimentes ändert, wurden die 45 Trials in 3 Blöcke mit jeweils 9 zusammengefassten Stimuluspositionen unterteilt. Eine 3 (Algorithmus) x 3 (Blöcke) ANOVA mit Messwiederholung ergab keine signifikanten Unterschiede für Algorithmus, $F < 1$, jedoch für den Faktor Blöcke, $F(2,22) = 36.913$, $p < .001$, $\varepsilon = .505$, $\eta^2 = .282$. Die Abweichungen im ersten Drittel des Experimentes waren mit 17 mm am geringsten, nahmen allerdings im zweiten und dritten Drittel auf 23 mm zu. Ob dieses Ergebnis auf ein Nachlassen der Fixationsleistung hindeutet oder der zunehmende Abstand zur Kalibrierung eine Verschlechterung der Messung bewirkt, kann abschließend nicht geklärt werden. Hierfür wäre eine erneute Kalibrierung vor jedem der 3 Blöcke erforderlich gewesen.

Tiefenparameter

Die vorangegangenen Tests bezogen sich auf die gesamte dreidimensionale Abweichung des ermittelten 3D-Blickortes und beinhaltet die Summe der Einzelabweichungen für die x-, y- bzw. z-Komponente. Allerdings werden in Algorithmen zur Identifikation von beispielsweise 2D-Fixationen mehrere Blickpositionen zu einer Bewegung zusammengefasst (Holmqvist et al., 2011). Dabei können sich entgegengesetzte horizontale und vertikale Abweichungen mathematisch kompensieren. Inwieweit sich diese Eigenschaft auch auf die räumliche Tiefe übertragen lässt, sollte in den folgenden Berechnungen überprüft werden. Hierfür sind die z-Werte der 3D-Blickorte als Tiefenparameter gemittelt und in verschiedenen Bedingungen überprüft worden.

Eine einfaktorielle (Algorithmus) ANOVA mit Messwiederholung zeigte höchstsignifikante Unterschiede in den Z-Abweichungen zwischen allen drei Algorithmen, $F(2,22) = 157.902$, $p < .001$, $\varepsilon = .504$, $\eta^2 = .935$. Die Abweichung war im Algorithmus *Gaze3dFromFixEyePos* mit 2 mm am geringsten und stieg in *Gaze3dFromVarEyePos* mit 15 mm bzw. *Gaze3dFromGazeVec* mit 14 mm an (Abb. 8-5 C). Da die Abweichungsunterschiede zwischen *Gaze3dFromVarEyePos* und *Gaze3dFromGazeVec* gering jedoch signifikant waren, wurden die Einzelabweichungen jeder Versuchsperson analysiert. Hier gab es eine systematisch positive Ausprägung über alle Probanden hinweg und erklärt die Signifikanz.

Analog zum Blickort wurde die Genauigkeit des Tiefenparameters in den verschiedenen Ebenen betrachtet. In einer 3 (Algorithmus) x 5 (Tiefenebene) faktoriellen ANOVA mit Messwiederholung konnten keine signifikanten Z-Abweichungsdifferenzen für Tiefenebene, $F(4,44) < 1$, und wiederum höchst signifikante Effekte für Algorithmus, $F(2,22) = 101.105$, $p < .001$, $\varepsilon = .505$, $\eta^2 = .21$, gefunden werden. Unterschiede gab es zwischen allen drei Methoden. Je nach Algorithmus lagen zwischen berechneter Tiefeninformation und tatsächlicher Stimulusposition in den Ebenen Abweichungen zwischen 0 – 17 mm vor. In Abb. 8-5 D ist erkennbar, dass die berechneten Tiefenparameter in allen drei Algorithmen positiv ausgerichtet sind und damit näher am Betrachter liegen als die Stimulus-Ebenen. Die Abweichungen waren im Algorithmus *Gaze3dFromFixEyePos* (0 – 6 mm) geringer als in *Gaze3dFromVarEyePos* (9 – 17 mm) und *Gaze3dFromGazeVec* (9 – 16 mm).

Vergenzwinkel

In einem weiteren Schritt wurden die Vergenzwinkel in den einzelnen Ebenen berechnet. Die einfaktorielle (Tiefenebene) ANOVA mit Messwiederholung ergab höchstsignifikante Unterschiede, $F(4,44) = 3777.653$, $p < .001$, $\varepsilon = .354$, $\eta_p^2 = .997$. Abb. 8-5 E veranschaulicht die nichtlineare Abnahme des Vergenzwinkels in Abhängigkeit der Betrachtungsentfernung.

Augenabstände

Die Augenabstände wurden zu Beginn des Experimentes einerseits manuell ermittelt und andererseits anhand der Cornea-Positionen vom Eyetracker automatisch berechnet (Abb. 8-5 F). Der t-Test für verbundene Stichproben ergab einen hoch signifikanten Unterschied, $t(11) = 4.709$, $p = .001$. Der manuell erhobene mittlere Abstand lag bei 66,3 mm und die in den Eyetracking-Daten ausgegebene mittlere Entfernung beider Cornea-Kugeln bei 63,8 mm.

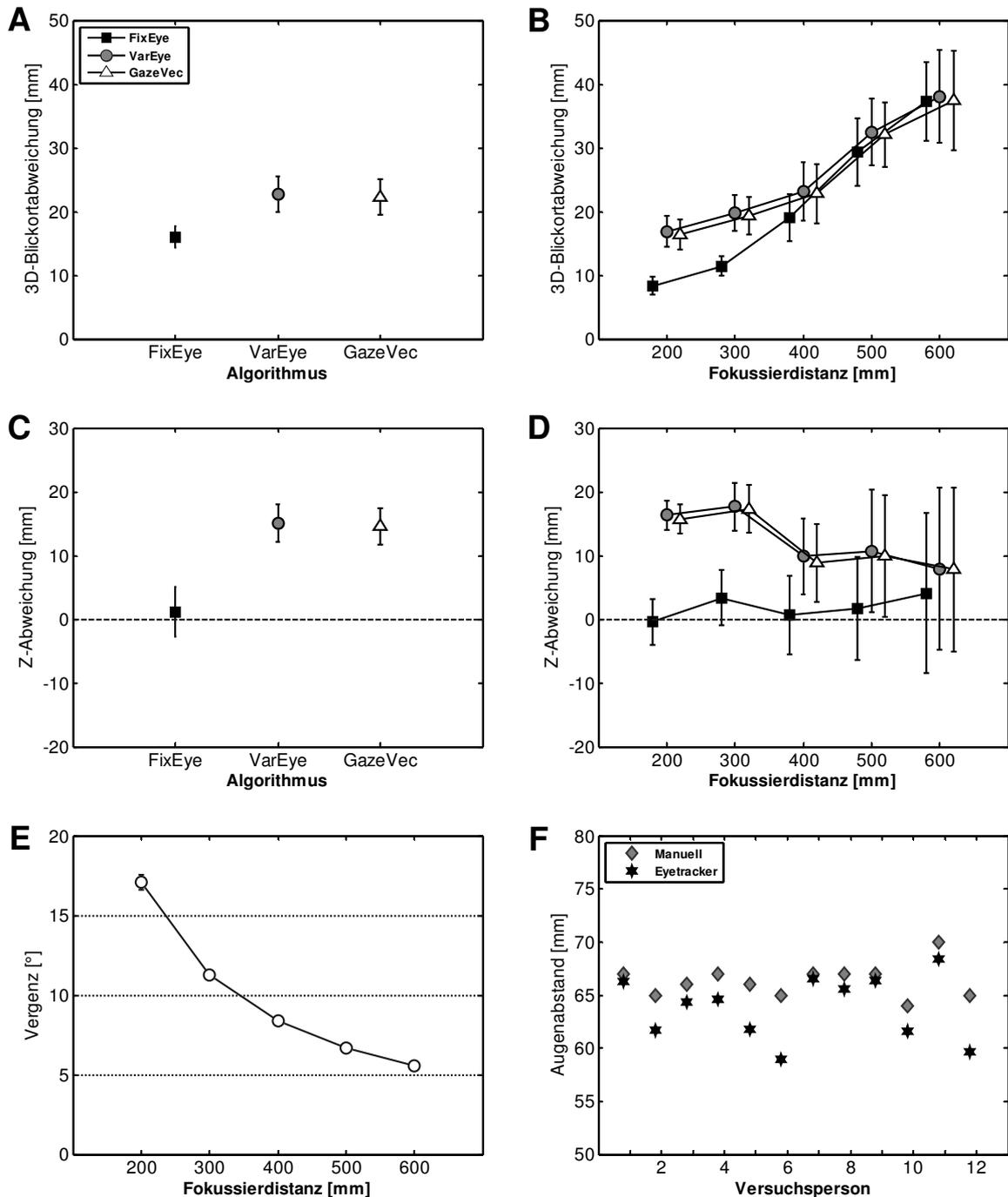


Abb. 8-5: 3D-Blickortabweichung in den unterschiedlichen Algorithmen (A) und Tiefenebenen (B). Genauigkeit des Tiefenparameters in Bezug auf Algorithmus (C) und Tiefenebene (D), Vergenzwinkel (E) sowie die manuell vom Versuchsleiter und automatisch vom Eyetracker ermittelten Augenabstände (F). Fehlerbalken kennzeichnen das 95%-CI.

8.4 Diskussion

Die Abweichungen zwischen den Fixationsobjekten und aus den Blickdaten ermittelten Blickortpositionen betragen dreidimensional 16–23 mm. Es wurden signifikante Unterschiede zwischen den getesteten Algorithmen sowie den verschiedenen Tiefenebenen nachgewiesen. Dabei bestätigte sich die Annahme, dass mit zunehmender Fokussierentfernung die Abweichungen ansteigen. Die separate Auswertung des Tiefenparameters ergab eine Genauigkeit von 2–15 mm. Auch hier zeigten sich signifikante Unterschiede zwischen den Algorithmen, allerdings nicht zwischen den Ebenen. Die Varianz in den 3D-Blickort- sowie die Tiefenparameterabweichungen stieg mit zunehmender Fokussierdistanz an. Dies ist auf den geringer werdenden Vergenzwinkel zurückzuführen. Je kleiner der Vergenzwinkel, umso größer sind die Auswirkungen von Messwertschwankungen auf die berechnete Blickposition. Im Abstand von 200 mm wurde ein durchschnittlicher Vergenzwinkel von 17° gemessen. Dieser nahm mit zunehmender Distanz des zu fixierenden Objektes ab und erreichte bei einer Fokussierdistanz von 600 mm Werte zwischen $5\text{--}6^\circ$. Die vom Versuchsleiter mit einer Augenabstandsschablone gemessenen Abstände wichen von den durch das Eyetracking-System ausgegebenen Kenngrößen signifikant ab. Welche Werte näher an der Realität liegen erklärt die Studie nicht. Auch konnte kein Zusammenhang zwischen den ermittelten Augenabständen (Manuell/Eyetracker) und der Abweichung in Bezug auf 3D-Blickposition bzw. Tiefenparameter hergestellt werden.

Der Algorithmus *Gaze3dFromFixEyePos* wies geringere Abweichungen bezüglich Blickort und Tiefenparameter auf als *Gaze3dFromVarEyePos* bzw. *Gaze3dFromGazeVec* auf. Hier stellte sich die Frage, wie genau die vom Eyetracker ausgegebenen Augenparameter (Cornea-Zentren und Blickvektoren) tatsächlich waren. Zu geringe Augenabstände würden die Verschiebung der Fokussierebene zum Betrachter hin und die Unterschiede zwischen manuell und automatisch ermittelter Augensymmetrie erklären. Die Ergebnisse wurden deshalb mit dem Hersteller diskutiert. Nach Überprüfung der Software- und Hardwarekonfiguration stellte sich heraus, dass die Ausrichtung der Kameraachsen des Eyetrackers außerhalb der Toleranzgrenzen von 0.2° lag. LC Technologies, Inc. nahm daraufhin Veränderungen an der Treibersoftware vor und kalibrierte die Kameras auf das Koordinatensystem des Eyetrackers neu.

Der vektorbasierte Ansatz hat sich in diesem Experiment bewährt und ergab insbesondere im nahen Fokussierbereich dreidimensionale Abweichungen von weniger als 20 mm. Er eignet sich damit zur Berechnung dreidimensionaler Blickorte in realer Umgebung.

9 Studie III: Eyetracking in einer VR-Umgebung

Die vorangegangene Studie wurde innerhalb einer virtuellen Umgebung nachgebildet und auf ähnliche Weise überprüft. In einem Experiment sollten die Probanden virtuelle Fixationsobjekte mit den Augen fixieren. Für die stereoskopische Visualisierung wurde das in Kapitel 6 getestete Shutterverfahren eingesetzt. Insbesondere musste die Diskrepanz zwischen Akkommodation der Augen in Monitorebene und Fokussierung bzw. Vergenz auf ein virtuelles Objekt in einer davon abweichenden Tiefe berücksichtigt werden. Für die Erstellung der virtuellen Szene wurden die dreidimensionalen Augenpositionen gemessen und in die Berechnung der stereoskopischen Einzelbilder einbezogen.

9.1 Fragestellungen und Hypothesen

Wichtigstes Qualitätskriterium war wiederum die 3D-Blickortabweichung in den drei verschiedenen Algorithmen. Entsprechend wurde die Genauigkeit der Blickpositionen im Raum sowie über den zeitlichen Verlauf hinweg analysiert. In der stereoskopisch projizierten Umgebung war dabei von besonderem Interesse, ob sich Versuchspersonen während des Experimentes an den Akkommodation-Vergenz-Konflikt gewöhnen und dadurch die Fixationsleistung verbessert wird oder andernfalls die Augen ermüden und sich damit die Fixationsleistung verschlechtert.

In den Berechnungen des dreidimensionalen Blickortes wurde bislang davon ausgegangen, dass die beiden Blickachsen in Nähe des Blickortes den geringsten Abstand zu einander haben. Allerdings zeigen bildschirmbasierte Untersuchungen aus der Leseforschung, dass die Überkreuzung der Blickachsen beim Lesen eines Textes auf einem Computermonitor sowohl vor der Bildschirmenebene (Nuthmann & Kliegl, 2009) als auch dahinter (Liversedge, Rayner, White, Findlay & McSorley, 2006) erfolgen kann. Eine mögliche Erklärung für diese Inkonsistenz erklären Huckauf, Watrin, Yuras und Koepsel (2013) anhand des eingesetzten Stimulusmaterials. Nuthmann und Kliegl präsentierten dunklen Text auf hellem Hintergrund während Liversedge et al. helle Buchstaben auf dunklem Hintergrund darboten. Um einen möglichen Einfluss der Stimulushelligkeit auf die Fokussiertiefe und damit der 3D-Blickposition zu erkennen, wurden den Versuchspersonen während des Experiments sowohl helle Fixationsobjekte auf dunklem Hintergrund als auch dunkle Fixationsobjekte auf hellem Hintergrund dargeboten und auf Abweichungsunterschiede hin untersucht.

9.2 Methode

In einem Experiment wurden Blickbewegungen beim Betrachten virtueller Fixationsobjekte in einer stereoskopisch projizierten VR-Umgebung gemessen. In mehreren Durchgängen erfolgte die Darbietung mit dem Shutter-Verfahren an unterschiedlichen Positionen im Raum. Für die Blickbewegungsmessung wurde ein binokulares Remote-Eyetracking-System eingesetzt, dessen Ausgabewerte die Berechnung der 3D-Blickortkoordinaten nach den vorgestellten Möglichkeiten *FixEye*, *VarEye* und *GazeVec* unterstützt.

Stichprobe

An der Studie nahmen 24 Versuchspersonen (12 weiblich) im Alter zwischen 25 und 49 Jahren teil. Das Durchschnittsalter lag bei 32 Jahren. Alle Versuchsteilnehmer gaben ein normales (55%) oder durch eine Brille bzw. Kontaktlinsen korrigiertes Sehvermögen an. 77% der Probanden hatten bereits stereoskopisch projizierte Filme gesehen. Ein Proband bestand den *Vision Assessment Corporation Fly Stereo Acuity Test with Lea Symbols® P/N 1000* zur Stereopsis nicht (Disparitäts-Level < 32") und musste von der Untersuchung ausgeschlossen werden.

Apparatus und Aufbau

Für die dreidimensionale Darstellung virtueller Objekte wurde ein 3D-Monitor mit einer Shutterbrille kombiniert. Das eingesetzte Bildschirmmodell hatte eine Abmessung von 510 x 290 mm bei einer Bildschirmdiagonale von 23", eine Auflösung von 1.920 x 1.080 Pixel und eine Bildwiederholfrequenz von 120 Hz. Die auf dem 3D-Monitor dargebotenen Bildinhalte wurde mit einem 3DVISION™ System des Herstellers NVIDIA® synchronisiert. Durch das Shutterverfahren halbierte sich diese auf 60 darstellbarer Bilder pro Sekunde. Für die Blickbewegungsmessung kamen wiederum das EFS sowie eine Kinnstütze zur Vermeidung von Kopfbewegungen zum Einsatz. Die Anordnung des Eyetrackers, 3D-Monitors sowie Positionierung der Versuchsperson und Kinnstütze erfolgte zunächst nach den Vorgaben des vorangegangenen Experimentes. Ein wesentlicher Unterschied zur vorhergehenden Studie bestand in einer aktualisierten Treibersoftware sowie Neujustierung der Eyetracking-Kameras, mit dem Ziel, eine höhere Genauigkeit der vom Eyetracking-System ausgegebenen Augendaten zu erreichen.

Während des Experimentes sollten den Probanden ähnlich dem vorangegangenen Experiment dreidimensionale virtuelle Fixationsobjekte an 45 verschiedenen Positionen in einer stereoskopisch visualisierten VR-Umgebung dargeboten werden: 5 Tiefenebenen (E), 3

Spalten (S) x 3 Zeilen (Z) pro Ebene. Durch die stereoskopische Visualisierung war jedoch ein kleinerer Messbereich als in realer Umgebung verfügbar, da die beiden Bilder für linkes und rechtes Auge bei der Darstellung auf der Bildschirmfläche einen horizontalen Versatz in Abhängigkeit der Wahrnehmungstiefe erfordern. Die Fixationsobjekte mit der weitesten Entfernung zur Projektionsfläche weisen gleichzeitig die größte Disparität auf und grenzen den 3D-Darstellungsbereich ein. Unter Einbehaltung des Monitorseitenverhältnisses von 16:9 ergab sich eine 3 stufige Aufteilung von $0/7/14^\circ$ vertikal und $-12/0/12^\circ$ horizontal. Damit wurde der dreidimensionale Messraum (Frustum) zwischen den Augen des Probanden, der Bezugsebene des Eyetracking-Systems sowie den darstellbaren virtuellen Objekten der Szene auf dem 3D-Monitor vollständig erfasst.

Stimulusmaterial

Die Fixationsobjekte in Form virtueller Kugeln hatten in Monitorebene einen Durchmesser von 7 mm und waren für eine bessere Fokussierung mit einer Markierung versehen. Um einen möglichen Einfluss der Stimulushelligkeit auf die Fokussiertiefe und damit der 3D-Blickposition zu untersuchen, sind den Versuchspersonen sowohl helle Fixationsobjekte auf dunklem Hintergrund als auch dunkle Fixationsobjekte auf hellem Hintergrund dargeboten worden (Abb. 9-1). Allerdings beeinflusste die Szenenbeleuchtung die tatsächlichen Farbwerte der Fixationsobjekte in Abhängigkeit der Entfernung zwischen Objekt und Lichtquelle. So variierte beispielsweise der Farbwert für die Kugeloberfläche in der hellen Szene zwischen 55/55/57 – 88/88/90.

Tab. 9-1: RGB-Werte des Stimulusmaterials.

	Helle Szene	Dunkle Szene
Hintergrund	230 / 230 / 230	26 / 26 / 26
Kugel	88 / 88 / 90	167 / 167 / 165
Markierung	0 / 0 / 0	38 / 37 / 36

Um für alle Versuchspersonen die gleiche Tiefenwahrnehmung zu erzeugen, ging die individuelle Pupillendistanz in die Stereobildgenerierung ein. Da in diesem Experiment der aktuelle Blickort zu jederzeit bekannt war, wurde für eine möglichst realitätsnahe Darstellung der Fluchtpunkt angepasst bzw. die Perspektive in der Szene korrigiert.

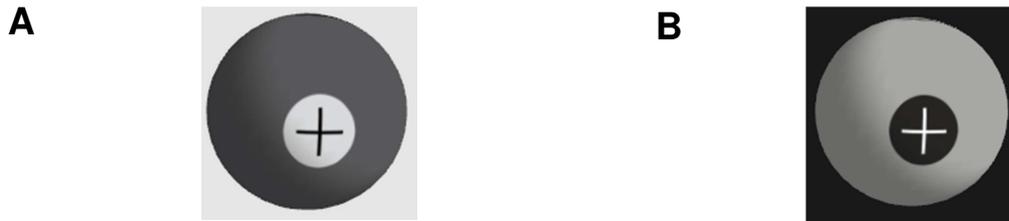


Abb. 9-1: Rechtes Augenbild eines virtuellen Fixationsobjektes im Zentrum der 4. Tiefenebene mit dunkler Kugel auf hellem Hintergrund (A) sowie heller Kugel auf dunklem Hintergrund (B).

Experimentsoftware

Zur Durchführung des Experimentes wurde ein Computerprogramm mit folgenden Funktionalitäten konzipiert und implementiert:

- Instruktionspräsentation
- Steuerung des Eyetracking-Systems (Init, Start, Stop, Calibrate)
- Abruf der Eyetracking-Parameter aus dem Rohdatenstrom bzw. Ringpuffer
- Berechnung der randomisierten Stimulusposition
- Rendering der 3D Szene mit Perspektivkorrektur
- Funktionsaufrufe der Gaze3D.dll zur Berechnung der 3D Blickorte
- Synchronisierung und Speicherung der Daten

In Kooperation mit der *AG Technische Visualistik der Professur Mediengestaltung, Institut für Software- und Multimediatechnik, Fakultät Informatik der Technischen Universität Dresden* wurde eine Schnittstelle für die am Lehrstuhl entwickelte Visualisierungssoftware BiLL (Bildsprache Live Lab) konzipiert, mit der die virtuelle 3D Umgebung in Echtzeit erzeugt werden konnte. Auch perspektivische Korrekturen nahm BiLL automatisch vor. Zusätzlich erfolgte die Anbindung an das Eyetracking-System sowie die Ansteuerung der Shutter-Technologie.

Vorversuche

In vier Vorversuchen wurde das Experiment mit jeweils 5 Probanden durchgeführt.

Vorversuch 1

Zunächst erfolgte die stereoskopische Visualisierung der dreidimensionalen virtuellen Fixationsobjekte, ähnlich dem vorangegangenen Experiment, an 45 unterschiedlichen Positionen über den gesamten Darstellungsraum verteilt.

Problem: Das Eyetracking-System lieferte nur wenige gültige Datensamples, da sowohl Pupille als auch Cornea-Reflexion unzureichend erkannt wurden.

Die Ursache für dieses Problem ist auf eine unzureichende Helligkeit in den Videokamerabildern des EFS aufgrund eines zu geringen Infrarotlichtanteils zurückzuführen.

Demnach blockierte die Shutterbrille ein Teil des Infrarotlichts und die Voraussetzungen für eine Messung nach der PCCR-Methode waren nicht mehr gegeben. Um diesen Effekt entgegenzuwirken, musste entweder die Menge des infraroten Lichts z.B. durch zusätzliche Infrarotlichtquellen erhöht oder der Abstand zur Quelle verringert werden. Da eine Modifikation des Eyetracking-Systems technisch bedingt nicht möglich war, wurde die Distanz zwischen Monitor bzw. Eyetracker und Augenposition auf 600 mm verkürzt und damit die notwendige Intensität an infrarotem Licht gewährleistet. Durch die Anzahl unterschiedlicher Tiefenebenen in je 100 mm Abstand ergab sich nun jedoch eine Nahdistanz von 100 mm für die 5. Ebene.

Vorversuch 2

Problem: Die Versuchspersonen konnten die Fixationsobjekte in der 5. Ebene nicht mehr fokussieren. Sie berichteten von nur noch zweidimensionaler Wahrnehmung der einzelnen Halbbilder in Monitorebene. Offenbar wurde hierbei die Diskrepanz zwischen Akkommodation und Vergenz derart groß, dass eine dreidimensionale Wahrnehmung mit der eingesetzten stereoskopischen Visualisierungstechnologie in einer Nahdistanz von 100 mm (Abstand zwischen Fixationsobjekt und Auge) nicht mehr gegeben war.

Als Konsequenz dieser Beobachtungen wurden die Fixationsobjekte der 5. Tiefenebene aus dem Setting entfernt. Dementsprechend änderte sich der ursprüngliche Versuchsaufbau wie folgt: Das EFS mit 3D-Bildschirm wurde in einer Entfernung von ca. 600 mm ausgehend von der Cornea-Oberfläche positioniert (Abstand: Auge-Projektionsfläche). Die erste Fixationsebene (E1) befand sich 100 mm ausgehend von der Kalibrierebene bzw. 500 mm vom Probanden entfernt. In 100 mm Abständen folgten drei weitere Ebenen, so dass bis zu einer Nahdistanz von $z = 200$ mm (E4) gemessen werden konnte (siehe Abb. 9-2).

Vorversuch 3

Problem: Es fiel den getesteten Probanden schwer, sprunghafte Positionsänderungen der Fixationsobjekte insbesondere bei Tiefenänderungen von Fern auf Nah zu verfolgen bzw. die beiden Halbbilder mit nun größerer Disparität zu einem dreidimensionalen Bild zusammenzuführen.

Für eine langsamere Anpassung an die Disparität wurde eine dynamische Bewegung der Fixationsobjekte mit einer Geschwindigkeit von 10 cm/s aus der Monitorebene in Richtung der Endposition implementiert. Durch die geführte Vergenz konnte die dreidimensionale Wahrnehmung nach Angaben der Probanden stabiler aufrechterhalten werden.

Vorversuch 4

In einem weiteren Test absolvierten wiederum 5 Versuchspersonen das Experiment. Diesmal verlief die stereoskopische Visualisierung erfolgreich bzw. es gab keinerlei negative Rückmeldung von den Probanden die Objekte nicht dreidimensional wahrnehmen zu können.

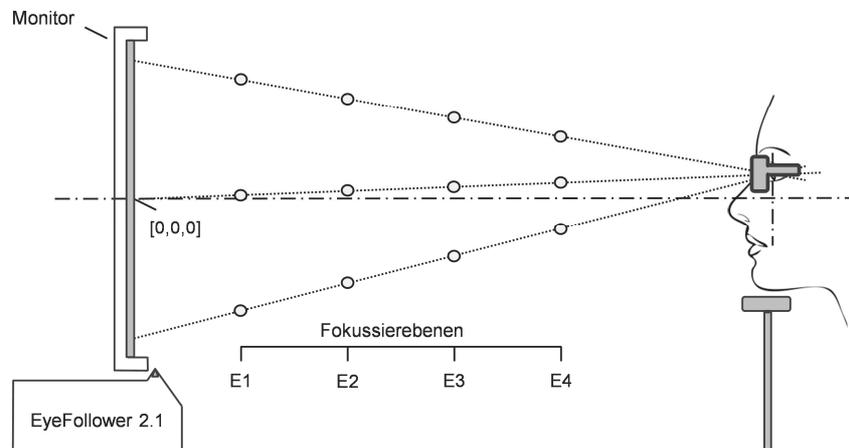


Abb. 9-2: Versuchsaufbau mit Eyetracking-System, 3D-Bildschirm, virtuelle Fixationsobjekte in den 4 Tiefenebenen, Proband mit Shutterbrille und Kinnstütze.

Versuchsablauf

Vor Beginn des Experimentes beantworteten die Probanden einen Fragebogen zu demographischen Daten, Sehvermögen, Erfahrungen mit 3D-Filmen sowie Experimenten zur Blickbewegungsmessung und Motivation. Anschließend wurde anhand eines *Pupillometers* (Essilor Instruments) der Augenabstand bzw. die Augensymmetrie gemessen und der *Vision Assessment Corporation Fly Stereo Acuity Test with Lea Symbols[®] P/N 1000* zur Überprüfung der Stereopsis durchgeführt. Bei einer Erkennung stereoskopisch dargebotener Kreise mit einer Disparität kleiner 32" galt der Test als bestanden. Damit sollte sichergestellt werden, dass die Versuchspersonen in der Lage sind das stereoskopische Bildmaterial dreidimensional wahrzunehmen.

Nach Instruktion des Probanden nahm dieser innerhalb der Versuchsumgebung Platz und die Kinnstütze wurde derart justiert, dass die Augenhöhe 50 mm unterhalb des oberen Monitorrandes lag. Das Experiment bestand aus einem Probedurchlauf und zwei weiteren Durchgängen, wobei sich letztere ausschließlich im verwendeten Stimulusmaterial unterschieden. In einem Durchgang wurden dunkle Fixationsobjekte auf hellem Hintergrund und im anderen Durchgang helle Fixationsobjekte auf dunklem Hintergrund dargeboten. Die Reihenfolge war ausbalanciert. Nach Positionierung der Versuchsperson fand die Kalibrierung des Eyetrackers statt. Je nach Stimulusvariante des Durchgangs wurde auch die Kalibrierung auf einem hellen (RGB: 255/255/255) bzw. dunklen (RGB: 0/0/0) Hintergrund

durchgeführt, um einer möglichen Pupillenmittelpunktdrift entgegenzuwirken. Bei einem Kalibrierwert unterhalb 0.3° wurde mit dem Experiment fortgefahren, andernfalls die Kalibrierung wiederholt.

Nach Darbietung der Instruktionsanleitung und Bestätigung durch Drücken der Leertaste sollte der Proband ein zentrales Fixationsobjekt in Monitorebene mit den Augen fokussieren. Während dieser Phase sind die 3D-Augenpositionen erhoben und die Durchschnittswerte für linke sowie rechte Augenposition der Prozedur zum Rendern der 3D-Szene übergeben worden, um die individuelle Stereobasis zur Bildgenerierung zu verwenden. An dieser Stelle wurde bewusst auf die manuell erhobenen Daten verzichtet, um später auch bei freier Kopfbewegung die Positionsdaten vom Eyetracker nutzen zu können. Anschließend folgte der Probedurchgang, in dem nacheinander die zentralen Fixationsobjekte in jeder Ebene fixiert werden sollten. Bereits hier stellte sich heraus, ob eine Versuchsperson die virtuelle Kugel in allen 4 Ebenen dreidimensional wahrnehmen und fixieren konnte.

Daraufhin startete der erste Durchgang, in dem nacheinander 36 Fixationsobjekte in randomisierter Reihenfolge im virtuellen 3D-Raum dargeboten wurden. Dabei bewegten sich die Fixationsobjekte ausgehend von der Monitorebene solange auf den Betrachter zu, bis die vorher errechnete 3D-Position erreicht war. Jetzt sollte die Versuchsperson möglichst genau die Markierung auf der Kugeloberfläche mit den Augen fokussieren, die Leertaste drücken und die Fixation bis zum Ertönen eines akustischen Signals (nach 1.000 ms) aufrecht erhalten. Gelang es einer Versuchsperson nicht, die beiden unterschiedlichen Bilder für linkes und rechtes Auge zu einem dreidimensionalen Gesamtbild zu vereinen, musste die F-Taste (F=Fehler) statt der Leertaste gedrückt werden, um Fehlwahrnehmungen im Datenstrom zu kennzeichnen und diese Daten in den späteren Analysen zu filtern. Dieser Vorgang wurde für alle 36 Messpunkte wiederholt. Nach Beendigung des ersten Durchgangs konnte sich die Versuchsperson in einer Pause mit selbstgewählter Dauer erholen. Danach startete die Kalibrierung für den zweiten Durchgang mit angepasstem Hintergrund. Der Ablauf des zweiten Durchgangs entsprach dem des ersten.

Datenaufbereitung

In die Berechnung gingen die Daten von 23 Versuchspersonen ein. Bevor jedoch mit der Auswertung begonnen werden konnte, wurden diejenigen Segmente aus dem Datenstrom gefiltert, in denen die Versuchsperson das Fixationsobjekt nach eigenem Empfinden fixiert hatte. Dies betraf die Daten zwischen dem Betätigen der Leertaste und dem Ertönen des akustischen Signals. Bei 23 Probanden mit je 2 Durchgängen und 36 Stimuluspositionen ergab sich bei 120Hz Aufzeichnungsrate eine Anzahl von 198.720 möglichen Messwerten. Jedoch

wurden nicht alle Fixationsobjekte dreidimensional wahrgenommen und die durch Drücken der Taste ‚F‘ angegebenen Fehlwahrnehmungen entfernt. In einem weiteren Schritt sind alle für die Auswertung der Genauigkeit bedeutungslosen Werte wie Blinks oder fehlende Binokularwerte gefiltert worden. Letztere traten auf, falls die Blickdaten von nur einem Auge vorhanden waren, beispielsweise aufgrund von Messproblemen bei Reflexionen. Durch die Filterung reduzierte sich die Datenmenge um 8,75%.

9.3 Ergebnisse

3D-Blickortkoordinaten

Anders als in der vorangegangenen Untersuchung waren die 3D-Abweichungen in allen Algorithmen mit 17 mm annähernd gleich (Abb. 9-3 A), signifikante Unterschiede gab es nicht, $F < 1$.

Genauigkeit der Blickorte in den unterschiedlichen Tiefenebenen

In einer 3 (Algorithmus) x 4 (Tiefenebene) faktoriellen ANOVA mit Messwiederholung konnten keine Effekte für Algorithmus, $F(2,44) = 1.758$, $p = .198$, $\varepsilon = .575$, höchst signifikante Abweichungsdifferenzen für Tiefenebene, $F(3,66) = 78.257$, $p < .001$, $\varepsilon = .716$, $\eta^2 = .761$, sowie Signifikanz für die Interaktion Algorithmus x Tiefenebene, $F(2,54) = 2.889$, $p = .011$, $\eta^2 = .001$, nachgewiesen werden. Die Abweichungen waren im Nahbereich bei 200 mm mit 11 mm am geringsten und nahmen mit steigender Entfernung auf die weiteste Fokussierebene bei 500 mm auf 30 mm zu (Abb. 9-3 B).

Genauigkeit der Blickorte innerhalb des Messzeitraumes

Für die Überprüfung, ob sich die Genauigkeit in Abhängigkeit der Dauer der Fixierung ändert, wurde der 1.000 ms Messzeitraum in Zeitfenster a 100 ms, 200 ms und 500 ms unterteilt. Die durchgeführten Varianzanalysen mit Messwiederholung zeigten jedoch keine signifikanten Effekte bzgl. Algorithmus und Zeitfenster, $F < 1$. Außerdem stellte sich die Frage, wie hoch die Abweichungen vor dem Tastendruck waren, insbesondere wenn man davon ausgeht, dass zwischen der Entscheidung die Leertaste zu drücken und der motorischen Umsetzung weitere Zeit vergeht. Um die Abweichungen vergleichen zu können, wurde die Zeitspanne von 500 ms jeweils vor und nach dem Leertastendruck in je 5 Zeitfenster a 100 ms aufgeteilt. Die einfaktorielle (Zeitfenster) ANOVA mit Messwiederholung ergab keine signifikanten Unterschiede der 3D Abweichungen zwischen den Zeiträumen vor und nach dem Betätigen der Leertaste, $F < 1$.

Genauigkeit der Blickorte über den zeitlichen Verlauf

Es sollte geprüft werden, ob sich die Genauigkeit im zeitlichen Verlauf verändert. Eine Verschlechterung könnte auf Müdigkeitserscheinungen der Augen bzw. Abnahme der Fixationsleistung und eine Verbesserung auf einen Lern- oder Gewöhnungseffekt bzgl. der stereoskopischen Visualisierung hindeuten. Dazu wurde jeder Durchgang in 3 Intervalle a 12 Fixationsobjekte unterteilt. Die gerechneten Varianzanalysen mit Messwiederholung ergaben keine Abweichungsdifferenzen für Durchgang sowie Intervall, jeweils $F < 1$.

Einfluss der Stimulushelligkeit auf die Genauigkeit

Das Experiment bestand aus zwei Durchgängen, wobei in einem der Durchgänge dunkle Fixationsobjekte auf hellem Hintergrund und im anderen helle Fixationsobjekte auf dunklem Hintergrund dargeboten wurden. Um zu prüfen, welchen Einfluss die Helligkeit der Stimuli auf die Genauigkeit der Blickortmessung hat, wurde eine einfaktorielle (Stimulushelligkeit) ANOVA mit Messwiederholung gerechnet. Es konnten keine signifikanten Unterschiede nachgewiesen werden, $F < 1$.

Tiefenparameter

Die einfaktorielle (Algorithmus) ANOVA mit Messwiederholung ergab höchstsignifikante Effekte in den Z-Abweichungen, $F(2,44) = 948.022$, $p < .001$, $\varepsilon = .606$, $\eta^2 = .977$, wobei die Unterschiede zwischen den Algorithmen *Gaze3dFromFixEyePos* und *Gaze3dFromGazeVec* sowie *Gaze3dFromVarEyePos* und *Gaze3dFromGazeVec* nachgewiesen wurden. Die Z-Abweichungen betragen sowohl bei *Gaze3dFromFixEyePos* als auch *Gaze3dFromVarEyePos* ca. 5 mm und bei *Gaze3dFromGazeVec* hingegen -1 mm (Abb. 9-3 C). Für eine genauere Betrachtung der Z-Abweichung in den Tiefenebenen wurde eine 3 (Algorithmus) x 5 (Tiefenebene) faktorielle ANOVA mit Messwiederholung gerechnet. Die Auswertung ergab ähnlich den realen Bedingungen keine signifikanten Unterschiede für Tiefenebene, $F(3,66) = 1.836$, $p = .182$, $\varepsilon = .501$, jedoch für Algorithmus, $F(2,44) = 1224.771$, $p < .001$, $\varepsilon = .674$, $\eta^2 = .101$. Auch hier unterschieden sich alle drei Algorithmen voneinander wobei die Effekte zwischen *Gaze3dFromFixEyePos* zu *Gaze3dFromVarEyePos* hoch signifikant, $p = .004$, und zwischen *Gaze3dFromFixEyePos* zu *Gaze3dFromGazeVec* und *Gaze3dFromVarEyePos* zu *Gaze3dFromGazeVec* höchst signifikant, $p < .001$, waren. Die Abweichungen des Tiefenparameters mit *Gaze3dFromGazeVec* fallen mit -2 mm geringer aus als die der anderen beiden Algorithmen mit ca. 4 mm (Abb. 9-3 D).

Vergenzwinkel

Es wurden die ermittelten Vergenzwinkel in den einzelnen Ebenen betrachtet. Eine einfaktorielle (Tiefenebene) ANOVA mit Messwiederholung ergab höchstsignifikante Unterschiede, $F(3,66) = 5054.353$, $p < .001$, $\varepsilon = .457$, $\eta^2 = .996$. Ähnlich den vorherigen Ergebnissen nahm der Vergenzwinkel mit steigender Entfernung zwischen Augen und Fokussierebene ab (Abb. 9-3 E).

Augenabstände

Innerhalb der Studie wurden die Augenpositionen einerseits vor Beginn der Blickbewegungsmessung mittels eines Pupillometers manuell ausgemessen und andererseits während des Experimentes anhand der 3D Positionen der Cornea-Zentren vom Eyetracker erfasst. Der t-Test für verbundene Stichproben in Bezug auf die manuell und vom der Blickbewegungsmessung automatisch ermittelten Augenabstände ergab einen höchst signifikanten Unterschied, $t(22) = -6.966$, $p < .001$ (2-Seitig), jedoch unterschieden sich beide Durchschnittswerte nur gering voneinander. Der mit dem Pupillometer erhobene Abstand lag im Mittelwert bei 66,1 mm und die in den Eyetracking-Daten ausgegebene mittlere Distanz beider Cornea-Kugeln 67,5 mm (Abb. 9-3 F).

9.4 Diskussion

Im Gegensatz zur vorangegangenen Untersuchung unterschieden sich im virtuellen Setting die dreidimensionalen Blickortabweichungen der Algorithmen nicht voneinander und betragen nur noch 16-17 mm. Mit der angepassten Treibersoftware sowie der erneuten Kamera-Kalibrierung wurden vom Eyetracker präzisere Messwerte für die Blickachsenparameter ausgegeben und somit eine höhere Genauigkeit der berechneten Blickorte erreicht. Auch im virtuellen Setting hing die Genauigkeit von der Fokussierdistanz ab. Je weiter weg sich der dreidimensionale Blickort befand, umso größer waren die Abweichungen. Der gemittelte Tiefenparameter lag mit einer Genauigkeit von -1-5 mm sehr nahe an der Fokussierebene. Signifikante Unterschiede gab es zwischen den Algorithmen und nicht zwischen den Ebenen. Wie in der vorangegangenen Studie vergrößerte sich die Varianz mit zunehmender Fokussierdistanz. Zu beachten bleibt, dass die Datenerhebung in unterschiedlichen Messbereichen stattfand. Der Abstand zwischen Versuchsperson und Eyetracker betrug 700 mm im realen und 600 mm im virtuellen Setting. Dadurch sind die Ergebnisse nicht direkt vergleichbar.

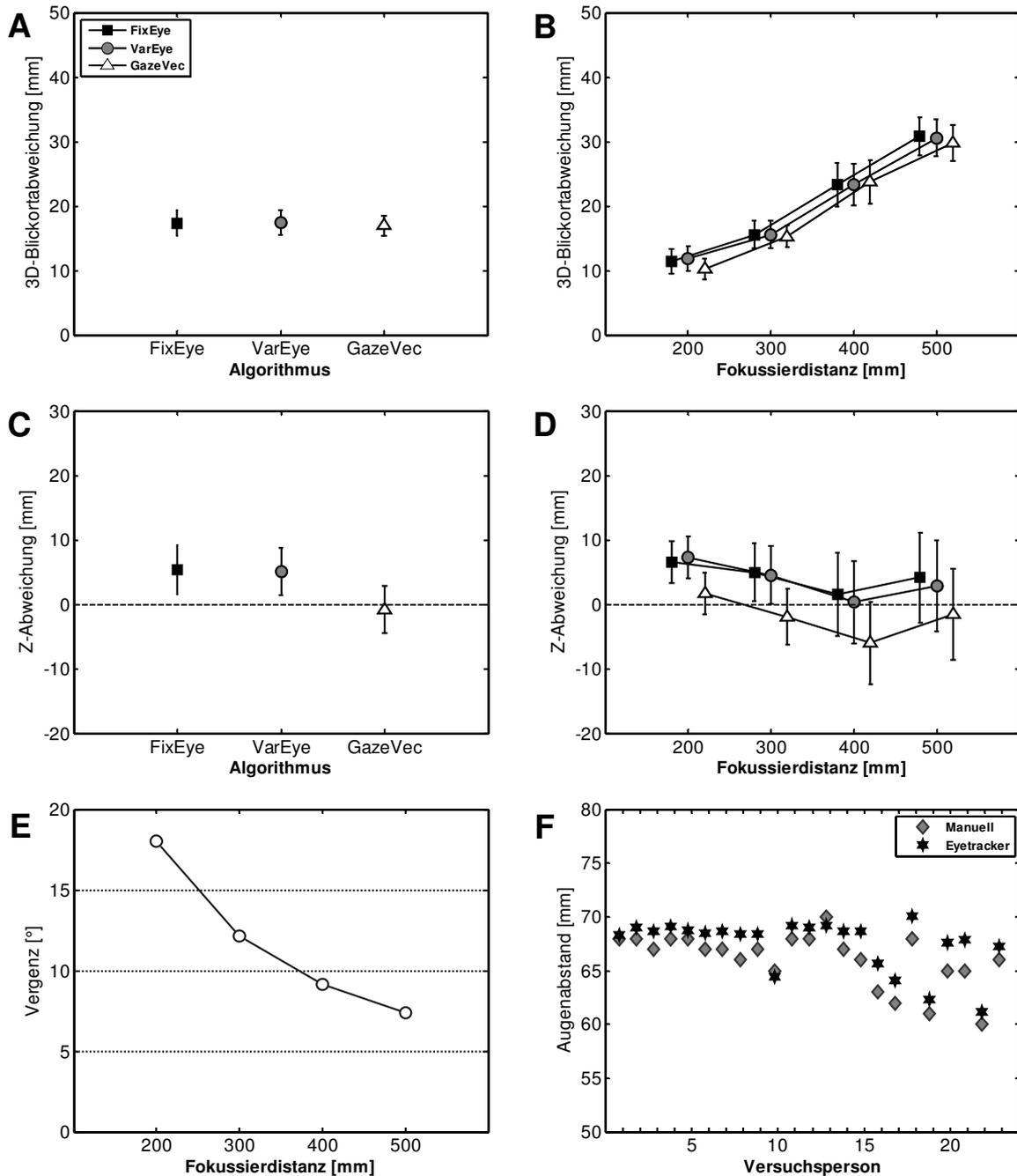


Abb. 9-3: Dreidimensionale Blickortabweichung in den unterschiedlichen Algorithmen (A) sowie Abweichungen der Blickorte in Bezug auf Tiefenebene (B). Genauigkeit des Tiefenparameters in den unterschiedlichen Algorithmen (C) bzw. Tiefenebenen (D), Vergenzwinkel (E) sowie die manuell vom Versuchsleiter bzw. automatisch vom Eyetracker ermittelten Augenabstände (F). Fehlerbalken kennzeichnen das 95%-CI.

Die Vergenzwinkel im realen und virtuellen Setting ähnelten sich in den jeweiligen Tiefenebenen. Demnach wurde im virtuellen Experiment der Vergenzwinkel trotz Akkommodation in Projektionsebene dem dreidimensional wahrgenommenen Objekt angepasst. Auch hier verringerte sich der Winkel mit zunehmender Fixationsdistanz. Da das Einwärtsdrehen der Blickachsen eine direkte Auswirkung auf die berechnete 3D-Blickposition hat, können sich Messwertschwankungen bei kleinerem Vergenzwinkel verstärkt auswirken und erklären die zunehmenden Abweichungen sowie die erhöhte Varianz.

Innerhalb dieser Studie konnte gezeigt werden, dass sich die Berechnung der dreidimensionalen Blickposition nach dem vektorbasierten Ansatz auch in einer stereoskopisch projizierten VR-Umgebung eignet. Die Ergebnisse aus den Studien II und III bilden die Basis für die Entwicklung eines neuen, 3D-Fixationserkennungsalgorithmus und dessen Schwellwertparameter.

10 Bestimmung dreidimensionaler Blickbewegungen

Um Fixationen und Sakkaden räumlich identifizieren zu können, müssen aus 3D-Blickorten räumliche Blickbewegungen abgeleitet werden. In diesem Kapitel sollen im ersten Teil Algorithmen zur Identifikation zweidimensionaler Blickbewegungen aus 2D-Blickorten erklärt und zusätzlich geprüft werden, inwieweit eine Erweiterung auf die dritte Dimension möglich ist. Anschließend wird ein ellipsoider Ansatz für die räumliche Fixationserkennung vorgestellt. Im dritten Teil werden die aus Studie II und III ermittelten 3D-Blickorte der ellipsoiden Fixationserkennung übergeben und die daraus berechneten Fixationsparameter analysiert.

10.1 Zweidimensionale Methoden

In der Literatur sind Algorithmen zur Identifikation von Blickbewegungen auf eine zweidimensionale Bezugsebene mehrfach beschrieben (Holmqvist et al., 2011; Komogortsev et al., 2010; Nyström & Holmqvist, 2010; Salvucci & Goldberg, 2000). Die Identifikation von Fixationen bzw. Sakkaden in Eyetracking-Protokollen ist essentieller Bestandteil der Blickbewegungsanalyse und dient als Basis für Auswertungen höherer Ebene. Typischerweise werden Parameter wie *Fixationsort* und *Fixationsdauer*, *Sakkadenbeschleunigung* und *Sakkadenamplitude* aber auch verschiedene transitionsbasierte Kenngrößen zwischen Fixationen oder definierten Bereichen ausgewertet. Die Analyse von Fixationen und Sakkaden erfordert die Identifikation dieser Blickbewegungsmuster aus den Sampledaten bzw. Blickorten des Eyetracking-Protokolls.

Salvucci und Goldberg (2000) beschrieben erstmalig eine Taxonomie von Fixationserkennungsalgorithmen, welche durch die unterschiedliche Auswertung örtlicher und zeitlicher Informationen in Eyetracking-Protokollen klassifiziert sind. Die örtlichen Klassifikationskriterien beziehen sich auf Geschwindigkeits-, Dispersion- und Bereichs-Informationen, während Dauer und lokale Adaption als zeitliche Kriterien die Fixationserkennungsalgorithmen in verschiedene Klassen unterteilen. Geschwindigkeitsbasierte Algorithmen analysieren bei konstanter Abtastfrequenz die örtliche Veränderung des Blickortes, der sich während der Fixation langsam bis gar nicht und während der Sakkade sehr schnell bewegt. Dispersionsbasierte Algorithmen hingegen berechnen die Abstände zwischen den einzelnen Blickorten unter der Annahme, dass diese innerhalb einer Fixation dichter aneinander liegen als während einer Sakkade. Bereichsbasierte Algorithmen filtern Messwerte

innerhalb vorgegebener Regionen (*Areas of Interest – AOI*), welche die visuell relevanten Ziele repräsentieren. Diese akkumulierten Blickorte mit entsprechender Verweildauer werden zu *Makro-Fixationen* zusammengefasst. Das erste zeitliche Kriterium ist die Fixationsdauer. Es wird angenommen, dass Fixationen selten unter 100 ms dafür häufig zwischen 200 – 400 ms andauern. Die lokale Adaption definiert das zweite zeitliche Kriterium und ermöglicht die Interpretation eines gegebenen Datenpunktes durch die Auslegung der zeitlichen Nachbarpunkte. Dies dient beispielsweise dazu, individuelle Unterschiede zwischen Personen mit trägen und denen mit agilen Augenbewegungen zu kompensieren. Um die Algorithmen-Klassen mit den unterschiedlichen Kriterien der Taxonomie vergleichbar zu machen, formalisierten Salvucci und Goldberg fünf Beispielalgorithmen, welche die grundlegenden Ideen jeder Klasse repräsentieren (siehe Tab. 10.1).

Tab. 10.1: Taxonomie der Fixationserkennungsalgorithmen (nach Salvucci & Goldberg, 2000).

Kriterien		Algorithmen				
		I-VT	I-HMM	I-DT	I-MST	I-AOI
Örtlich	Geschwindigkeit	•	•			
	Dispersion			•	•	
	Bereich					•
Zeitlich	Dauer			•		•
	Lokale Adaption		•	•	•	

Geschwindigkeitsbasierte Identifikation

Die geschwindigkeitsbasierte Fixationserkennung (*Velocity-threshold fixation identification – I-VT*) ordnet Messpunkte den Fixationen und Sakkaden anhand der Punkt-zu-Punkt Geschwindigkeiten zu. Nach Salvucci und Goldberg (2000) lassen sich diese in zwei Geschwindigkeitsprofile gliedern: geringe Geschwindigkeiten für Fixationen (z.B. $< 100^\circ/\text{s}$) und größere für Sakkaden (z.B. $> 300^\circ/\text{s}$). Jeder Geschwindigkeitswert wird aus der Distanz zwischen dem vorangegangenen und aktuellen Messpunkt und der dazwischen liegenden Zeit berechnet. Falls die Geschwindigkeit geringer als ein vorgegebener Schwellwert ist, wird der aktuelle Punkt einer Fixation, andernfalls einer Sakkade zugeordnet. Anschließend werden aufeinanderfolgende Fixationspunkte zu Fixationen zusammengefasst und die sakkadischen Messpunkte verworfen. Jede Fixation wird durch den Fixationsort als Zentrum der Punktmasse (x, y) , Startzeitpunkt t und Fixationsdauer d definiert.

Die Fixationserkennung nach dem *Hidden Markov Modell* (I-HMM) erweitert die I-VT Methode und nutzt wahrscheinlichkeitstheoretische Analysen zur Identifikation von Fixationen und Sakkaden. Hierbei wird kein fester Schwellwert vorgegeben, sondern anhand des Eyetracking-Protokolls und der daraus ableitbaren Geschwindigkeitsverteilungen für Fixationen und Sakkaden bestimmt. I-HMM verwendet ein Zwei-Zustands-HMM, in dem die Geschwindigkeitsverteilungen für Fixationen- und Sakkadenpunkte als zwei unterschiedliche Zustände repräsentiert sind. Zunächst werden alle Punkt-zu-Punkt Geschwindigkeiten im Protokoll berechnet. Dabei sollten sich zwei Normalverteilungen unterscheiden lassen: eine über einen geringeren Geschwindigkeitswert für Fixationen und eine über die höhere Sakkadengeschwindigkeit. Zusätzlich wird die Übergangswahrscheinlichkeit zwischen den Zuständen einbezogen. So ist es wahrscheinlicher, dass auf eine Fixation eine Sakkade und anschließend wiederum eine Fixation folgt. Ein Blickbewegungsmuster mit zwei aufeinanderfolgenden Sakkaden oder Fixationen hingegen ist wenig wahrscheinlich. Anhand der probabilistischen Vorgehensweise des I-HMM können die Geschwindigkeitswerte zwischen Messpunkten für Fixationen und Sakkaden individuell unterschieden werden und ermöglichen, im Gegensatz zum I-VT mit fest eingestelltem Schwellwert, eine lokale Adaption. Die identifizierten, aufeinanderfolgenden Fixationspunkte werden auch bei dieser Methode zu Fixationen aggregiert (x, y, t, d) .

Dispersionsbasierte Identifikation

Im Gegensatz zur geschwindigkeitsbasierten Erkennungsalgorithmen I-VT und I-HMM, geht die dispersionsbasierte Identifikation (*Dispersion-Threshold Identifikation – IDT*) davon aus, dass Punkt-zu-Punkt Abstände innerhalb einer Fixation geringer sind als während einer Sakkade. IDT erkennt Fixationen als Gruppen aufeinanderfolgender Punkte innerhalb eines räumlichen Gebietes mit einer maximalen Streuung. Weil Fixationen typischerweise eine Minimaldauer von 100 ms aufweisen, berücksichtigen dispersionsbasierte Algorithmen häufig einen Schwellwert für die minimale Fixationsdauer. Der Algorithmus initialisiert ein Zeitfenster über die ersten Messpunkte in Abhängigkeit der minimalen Fixationsdauer. Falls die Streuung der im Zeitfenster liegenden Messpunkte innerhalb des Dispersions-Schwellwertes liegt, werden alle folgenden Messpunkte, die diese Grenze ebenfalls einhalten, zum Zeitfenster hinzugefügt. Sobald das erste Sample diese Bedingung nicht mehr erfüllt, fasst der Algorithmus alle zuvor gesammelten Fixationspunkte des Zeitfensters zu einer Fixation zusammen und der Vorgang beginnt von vorn. Streuen die Messpunkte innerhalb des initialisierten Zeitfensters zu stark, wird der erste Wert entfernt, ein neuer hinzugefügt und die Dispersions-Bedingung geprüft. IDT benötigt zwei Parameter: den Schwellwert für die maximale Streuung bzw. Abstand des

aktuellen Fixationszentrums zum neuen Messwert und die minimale Fixationsdauer bzw. Anzahl an Samples zur Identifikation. Nach Salvucci und Goldberg (2000) eignen sich hierfür $0.5 - 1^\circ$ des Blickwinkels für den Dispersions-Schwellwert, sofern der Abstand des Auges zum Monitor bekannt ist, sowie ein Schwellwert für die minimale Fixationsdauer zwischen 100 – 200 ms. Die Ausgabeparameter entsprechen denen von I-VT und I-HMM.

Eine weitere Identifikationsmethode basiert auf *Minimum Spanning Trees (I-MST)*. Dabei werden alle Messpunkte nach *Prim's Algorithm* durch Linien miteinander verbunden (Camerini, Galbiati & Maffioli, 1988). Die Besonderheit der daraus resultierenden Baumstruktur liegt darin, dass die Gesamtlänge aller Baumsegmente minimiert ist. Von jedem Messpunkt aus gehen mindestens eine oder auch mehrere Verbindungen ab und es bilden sich Cluster. Messpunkte die an den Randbereichen eines Clusters liegen weisen weniger Verbindungen zu Nachbarpunkten auf und werden als Sakkadenpunkte identifiziert. Der Schwellwert für die minimale Anzahl notwendiger Verbindungen definiert das Filterkriterium. Die Kantenlängen der verbleibenden Messpunkte werden statistisch ausgewertet. Der berechnete Mittelwert sowie die Standardabweichung bilden weitere adaptive Filterkriterien, um aus den Clustern Sakkadenpunkte zu entfernen (Goldberg & Schryver, 1995). Alle jetzt noch vorhandenen Messpunkte eines Clusters werden zu einer Fixation zusammengefasst.

Bereichsbasierte Identifikation

Die vier vorangegangenen Identifikationsmethoden bestimmen Fixationen an beliebigen Orten des visuellen Feldes. Im Gegensatz dazu sucht die *Area of Interest Identifikation (I-AOI)* nach Fixationen in einem definierten Zielbereich. Diese Regionen repräsentieren Einheiten visueller Informationen und sind durch die zu analysierenden Zielobjekte definiert. Zu Beginn ordnet I-AOI Datenpunkte den definierten Zielbereichen zu. Dazu werden die innerhalb einer AOI gemessenen Punkte ortsabhängig dem entsprechenden Bereich als zugehörig markiert und alle außerhalb liegenden als Sakkaden-Messpunkte definiert. Anschließend fasst der Algorithmus alle aufeinanderfolgenden Messpunkte innerhalb einer AOI zu Fixationsgruppen zusammen und filtert diejenigen, die den Schwellwert für die minimale Dauer nicht überschreiten. Dadurch sollen Messpunkte durchstreifender Sakkaden gefiltert werden. Allerdings sind Sakkaden zwischen Fixationen innerhalb einer AOI nicht identifizierbar. Es handelt sich bei I-AOI also nicht um einen Fixationserkennungsalgorithmus, sondern vielmehr als ein Konzept zur Bestimmung von Fixationsgruppen höherer Ordnung, die visuellen Zielen bzw. Bereichen zugeordnet sind.

Evaluation und Vergleich

Salvucci und Goldberg (2000) analysierten die verschiedenen Identifikationsalgorithmen anhand eines Eyetracking-Protokolls in Bezug auf: Genauigkeit, Geschwindigkeit, Robustheit, Schwierigkeitsgrad der Implementierung sowie die Anzahl benötigter Parameter. Dabei stellten sie fest, dass geschwindigkeitsbasierte Algorithmen nach I-VT zwar unkompliziert umzusetzen, effizient und in Echtzeit lauffähig sind, es jedoch bei Geschwindigkeiten in Nähe des Schwellwertes aufgrund von Messwertschwankungen des Eyetrackers zu Problemen kam. Hierbei wurden Fixationen bestehend aus ein oder zwei Messpunkten erkannt. Um diesen Effekt zu vermeiden, ist die Festlegung einer Minimaldauer für die Aufrechterhaltung der Sakkaden-Mindestgeschwindigkeit von beispielsweise 10 ms sinnvoll (Sen & Megaw, 1984). I-HMM verwendet statt eines definierten Geschwindigkeitsschwellwertes ein probabilistisches Modell, um Messpunkte für Fixationen bzw. Sakkaden unterscheiden zu können und löst damit auch das Problem der Messwertschwankungen in den Eyetracking-Daten. Darüber hinaus kann das Zwei-Zustands-HMM in weitere kognitive Modelle integriert werden, um die Zuverlässigkeit der Blickbewegungserkennung zu steigern (Salvucci & Anderson, 1998). Die Implementierung dieser Methode ist nach Angaben der Autoren komplex, benötigt allerdings keine festzulegenden Schwellwerte.

Bei den dispersionsbasierten Modellen berechnet I-DT nahezu identische Ergebnisse wie I-HMM. Durch den zusätzlichen Schwellwert für eine minimale Fixationsdauer wurde auch hier das Zuordnungsproblem schwankender Eyetracking-Daten im Grenzbereich gelöst. Beachtet werden muss jedoch, dass sich die Schwellwerte für den örtlichen Toleranzbereich sowie der minimalen Fixationsdauer wechselseitig beeinflussen. Ein geringer Dispersions-Schwellwert in Kombination mit einer großen Mindestfixationsdauer führte zu weniger erkannten Fixationen als umgekehrt. Deshalb erfordert I-DT die größte Sorgfalt bei der Schwellwertfestlegung unter Berücksichtigung der zugrundeliegenden Dispersionsberechnung (Blignaut, 2009). Typische Werte sind ca. 0.5° für die Dispersionsstoleranz und 80 – 150 ms für die minimale Fixationsdauer (Nyström & Holmqvist, 2010).

Anders als die bisherigen Modelle fasst I-MST Fixationspunkte zu Clustern zusammen und ignoriert dabei die sequentielle Information im Eyetracking-Protokoll. Einen direkten Vergleich nahmen Salvucci und Goldberg (2000) deshalb nicht vor. Ähnlich verhält es sich bei I-AOI. Hier werden aufeinanderfolgende Messpunkte innerhalb eines Bereichs zu einer Fixation zusammengefasst und eingeschlossene Sakkadenpunkte als fixationszugehörig betrachtet. Sowohl I-MST als auch I-AOI eignen sich für Analysen höherer Ordnung, allerdings weniger für eine Blickbewegungserkennung. Nach Einschätzung der Autoren sind I-HM und I-DT für eine genaue und robuste Fixationserkennung aus den Messpunkten unter

Einbehaltung der sequentiellen Reihenfolge am besten geeignet. I-MST bietet ebenfalls eine robuste Identifikation, diese ist allerdings langsamer als alle anderen Algorithmen. I-VT kann wesentlich einfacher implementiert werden und arbeitet sehr schnell, funktioniert jedoch in den Schwellwertgrenzbereichen unzuverlässig.

Erweiterte Ansätze

Die bisherigen Identifikationsmethoden wurden erweitert, um einerseits die bekannten Probleme bei Schwankungen in den Messdaten mit neuen Ansätzen lösen und andererseits Blickbewegungen durch weiterentwickelte Technologien mit höheren Abstraten effizienter und genauer bestimmen zu können. Smeets und Hooge (2003) griffen die Problematik der Ungenauigkeiten bei geschwindigkeitsbasierter Sakkadenerkennung in Nähe des Schwellwertbereiches bei I-VT auf und präsentierten einen Ansatz, in dem sie zusätzlich eine Konturenabschätzung zur Bestimmung des Sakkaden-Peaks durchführten. Anschließend prüft der Algorithmus, ob sich dieser im Zentrum der durch den geschwindigkeitsbasierten Ansatz erkannten Sakkade befindet. Falls dem so ist, wird die Sakkade anerkannt oder andernfalls verworfen. Geschwindigkeitsbasierte Algorithmen können auch mit Beschleunigungskriterien kombiniert werden, um Anfang und Ende einer Sakkade zu detektieren. Beispielsweise verwendet die EyeLink™-Software Parameter für Geschwindigkeits-, Beschleunigungs- und Bewegungsschwellwerte von $30^\circ/\text{s}$, $8.000^\circ/\text{s}^2$ und $0,15^\circ$ (SR Research, 2009).

Einen adaptiven Algorithmus zur Erkennung von Fixationen, Sakkaden und *Glissaden* entwickelten Nyström und Holmqvist (2010). Als Glissaden werden Blickbewegungen am Ende einer Sakkade beschrieben, wobei die Augenbewegung nicht abrupt stoppt, sondern das Auge für einen Moment ‚zittert‘, bevor es für die folgende Fixation wieder relativ stillsteht. Dieser Effekt wird nicht in jeder Sakkade nachgewiesen, kann allerdings die erkannte Blickbewegung je nach Zuordnung zur Sakkade oder Fixation beeinflussen und sollte nach Auffassung der Autoren gefiltert werden. Sie erweiterten den Algorithmus von Smeets und Hooge (2003) mit verschiedenen Filterfunktionen, um einerseits das Rauschen (Messwertschwankungen des Eyetrackers) zu entfernen und andererseits neben Fixationen und Sakkaden zusätzlich Glissaden aus den Blickdaten bestimmen zu können.

Dass verschiedene Algorithmen für die Erkennung von Blickbewegungen aus Eyetracking-Protokollen nicht nur unterschiedliche Performances erreichen, sondern darüber hinaus auch mit denselben Daten unterschiedliche Ergebnisse berechnen, berichten Komogortsev et al. (2010). Letzteres Problem kann in vergleichbaren Studien mit derselben Methode zu widersprüchlichen Schlussfolgerungen führen. Beispielsweise wiesen Crevits, Hanse, Tummers und Van Maele (2000) in ihren Eyetracking-Daten nach, dass Personen mit leichtem Schädel-

Hirn-Trauma in der Lage sind, normale antisakkadische Blickbewegungen auszuführen. Daraus schlussfolgerten sie, dass diese Art von Untersuchungen keine diagnostische Relevanz besitzt. Das Gegenteil zeigten Heitger et al. (2009) in ihren Daten, obwohl die Art der Methode sowie die Messtechnik vergleichbar waren. Allerdings wurden in den beiden Studien unterschiedliche Methoden zur Identifikation der Blickbewegungen eingesetzt.

Komogortsev et al. (2010) evaluierten in ihrer eigenen Studie fünf Algorithmen zur Identifikation von Blickbewegungen in Hinblick auf okulomotorisches Fixations- bzw. Sakkadenverhalten und stellten dabei fest, dass die Ergebnisse der Algorithmen aus demselben Eyetracking-Protokoll selbst bei einer einfachen Stimulusdarbietung und gleichen Schwellwerten stark variierten. Um dies zu vermeiden, schlugen sie eine standardisierte Identifikation von Fixationen und Sakkaden vor. Dazu entwickelten die Forscher ein Bewertungssystem, in dem die Anzahl an Fixationen, durchschnittliche Fixationsdauer, Anzahl an Sakkaden, durchschnittliche Sakkadendauer sowie weitere qualitative und quantitative Metriken zwischen Algorithmen verglichen werden können. Demnach müssten die Schwellwerte der einzelnen Algorithmen derart angepasst werden, dass nach dem Bewertungssystem alle Algorithmen vergleichbare Resultate wie Anzahl an Fixationen, durchschnittliche Fixationsdauer, usw. erzielen. Jedoch ist ein automatisches Berechnen der Schwellwerte schwierig umsetzbar, da die Parameter des Bewertungssystems auch stimulus- und aufgabenabhängig sind.

10.2 Anforderungen an eine räumliche Blickbewegungserkennung

Über die Erkennung räumlicher Blickbewegungen ist aufgrund der zweidimensionalen Messtechnik nur wenig bekannt. Duchowski et al. (2002) kombinierten beispielsweise den geschwindigkeitsbasierten Ansatz mit einer Kollisionserkennung in einem HMD. Dabei verwendeten sie die 2D-Berechnung der Blickortgeschwindigkeit innerhalb der Eyetracker-Messebene zur Identifikation der zeitlichen Fixationsparameter und die Blickachsenkollision mit der virtuellen Szene zur Positionsbestimmung. Für eine szenenunabhängige Erkennung dreidimensionaler Blickbewegungen eignet sich diese Methode nicht. Sowohl Essig et al. (2006) als auch Pfeiffer et al. (2007) entkoppelten in einer VR-Umgebung den räumlichen Blickort von der Szene. Sie setzten zur Bestimmung der dreidimensionalen Fixationsorte die von den Eyetracking-Systemen ausgegebenen 2D-Fixationskoordinaten in ihre PSOM ein. Da die Fixationserkennung von den Herstellern auf die zweidimensionale Messebene ausgelegt ist, blieb der Tiefenparameter für die Identifikation unberücksichtigt.

Aus den bereits bekannten zweidimensionalen Methoden geht hervor, dass entweder Fixationen erkannt und dazwischen Sakkaden angenommen werden oder umgekehrt. Für eine Sakkadenerkennung ist eine hohe Dichte an Messpunkten erforderlich, da diese in ihrer Dauer wesentlich kürzer auftreten. Um auch Eyetracking-Systeme mit geringer Frequenz für die räumliche Blickbewegungserkennung einsetzen zu können, ist die Fixationserkennung einer Sakkadenerkennung vorzuziehen. Für eine geschwindigkeitsbasierte Fixationserkennung müssen allerdings sämtliche Punkt-zu-Punkt Geschwindigkeiten berechnet und den Blickbewegungen zugeordnet werden. Bei räumlichen Messpunkten ist nicht nur die horizontale und vertikale Komponente sondern auch der Tiefenparameter für die Blickbewegungsgeschwindigkeit einzubeziehen.

Hinzu kommt die Erkenntnis aus Studie II und III, dass mit zunehmender Entfernung des Blickortes vom Betrachter der Tiefenparameter stärker variiert. Die Punkt-zu-Punkt Geschwindigkeit kann demnach allein durch größere Schwankungen des Tiefenparameters zu Fehlinterpretationen im Erkennungsalgorithmus führen. Geschwindigkeitsbasierte Ansätze nach I-VT und I-HMM sind hierdurch ausgeschlossen. Eine Identifikation nach I-MST ist theoretisch auch räumlich möglich und kann die Schwankungen des Tiefenparameters mit entsprechenden Schwellwerten berücksichtigen, jedoch geht die sequentielle Beziehung und damit die zeitliche Aufeinanderfolge der Blickorte bei diesem Algorithmus verloren. Bei I-AOI ist für die Fixationszuordnung die Festlegung objektbezogener Bereiche bzw. Räume notwendig und damit eine Szenenunabhängigkeit nicht gegeben. Sowohl I-MST als auch I-AOI eignen sich für Analysen höherer Ordnung, weniger für die Identifikation von Blickbewegungen.

Bessere Voraussetzungen ergeben sich für die abstandsbaasierte 3D-Fixationserkennung nach I-DT. Abstände können nicht nur zweidimensional sondern auch räumlich berechnet und nach dem Ansatz des 2D-Algorithmus innerhalb einer Toleranzgrenze einer 3D-Fixation zugeordnet werden. Geometrisch ergibt sich für eine homogene Grenze in alle Dimensionen ein kugelförmiger Toleranzbereich. Die Mindestanzahl notwendiger Samples innerhalb der Grenze zur Fixationserkennung bestimmt auch hier die minimale Fixationsdauer. Allerdings müssen die Schwellwerte für Abstand und minimale Samples sinnvoll gewählt werden. Aus den bisherigen Erkenntnissen lassen sich folgende Anforderungen ableiten:

- Anforderung 1:* Blickbewegungen werden unabhängig einer Objektkollision erkannt.
- Anforderung 2:* Für die Identifikation wird der Tiefenparameter dreidimensionaler Blickorte einbezogen.

- Anforderung 3:* Der Algorithmus erkennt räumliche Fixationen. Sofern keine Blinks stattfinden, werden zwischen den Fixationen liegende Sakkaden angenommen.
- Anforderung 4:* Tiefenschwankungen des gemessenen Blickortes werden in Abhängigkeit der Wahrnehmungsentfernung berücksichtigt.
- Anforderung 5:* Die 3D-Fixationserkennung erfolgt anhand der Abstandsanalyse zwischen 3D-Blickorten.

10.3 Ein ellipsoider 3D-Fixationserkennungsalgorithmus

Um einen Überblick über die räumliche Verteilung der berechneten Sampledaten zu erhalten, wurde eine OpenGL-Software zur Visualisierung des Experimentsettings sowie der berechneten Blickpositionen entwickelt (Abb. 10-1). Betrachtet man die in Studie III bestimmten 3D-Blickorte im Profil, stellt man fest, dass die durch Schwankungen des Tiefenparameters bedingte Abstandsberechnung anhand einer Kugel nachteilig ist. Wählt man den Toleranzbereich für die örtliche Auflösung derart groß, dass alle ermittelten Blickorte innerhalb einer Kugel erfasst werden, erweitert sich gleichzeitig auch die horizontale bzw. vertikale Fixationsausdehnung (Abb. 10-1 A). Die räumliche Auflösung identifizierbarer Fixationen wird dadurch eingeschränkt. Wählt man hingegen den Bereich nach den x-y-Toleranzen des zweidimensionalen Algorithmus, schließt man dadurch der Fixation zugehörige Messpunkte aus (Abb. 10-1 B). Eine Kugel beschreibt die Sampledatenverteilung demnach nur unzureichend.

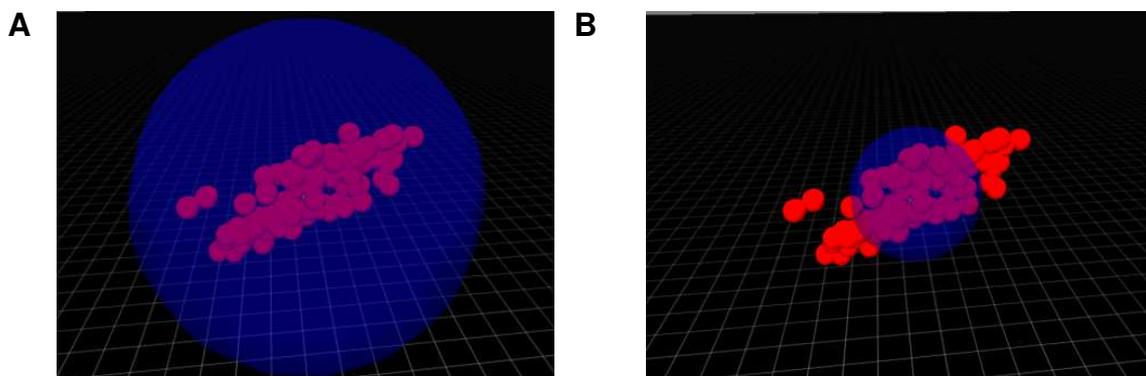


Abb. 10-1: Dreidimensionale Fixationsbestimmung durch Aggregation räumlicher Blickorte (rot) innerhalb eines kugelförmigen Toleranzbereichs (blau). Bei großem Schwellwert wird der Fixation zu viel Raum zugeordnet und damit die räumliche Auflösung erkennbarer Fixationen verringert (A). Wird dieser zu klein gewählt, bleiben der Fixation zugehörige Samples unberücksichtigt (B).

10.3.1 Relation zwischen Fixationstiefe und Messgenauigkeit

Dass die Genauigkeit der gemessenen Blickorte in Abhängigkeit von der Entfernung zwischen Auge und Fixationsobjekt variiert, wurde bereits in Studie II und Studie III nachgewiesen. Je weiter der Fixationsort vom Betrachter entfernt liegt, umso größer sind die Abweichungen in den ermittelten Blickpositionen. Eine mögliche Ursache liegt in den Messwertschwankungen des Eyetrackers. Die räumliche Auflösung des Eyetrackers wird in Winkelgrad angegeben und hängt von der Genauigkeit des ermittelten Blickvektors ab. Aufgrund der Geometrie wirken sich dabei horizontale und vertikale Disparitäten der Blickorte für linkes und rechtes Auge in Messebene verstärkt auf die berechnete Blickorttiefe aus. Wie man der Abb. 10-2 entnehmen kann, ergibt sich für die Blickortbestimmung aufgrund Überschneidungen der Blickvektorschwankungen ein annähernd ellipsenförmiger Toleranzbereich. Je weiter sich der Blickort vom Betrachter weg befindet, desto größer die horizontale und vertikale Ausdehnung. Für eine dreidimensionale, abstandsbezogene Fixationserkennung wird im Folgenden der Toleranzbereich anhand eines Ellipsoids konstruiert.

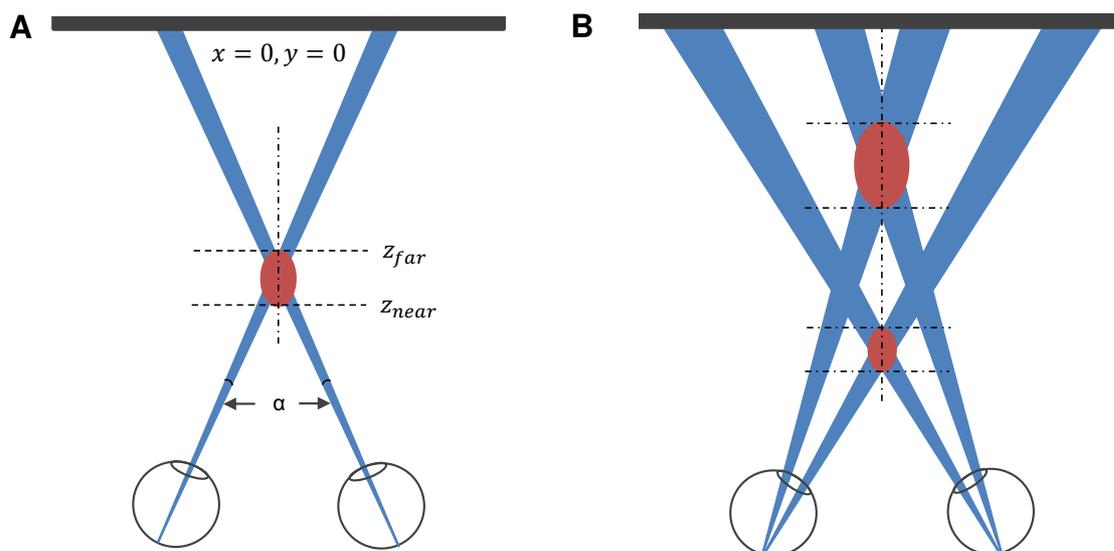
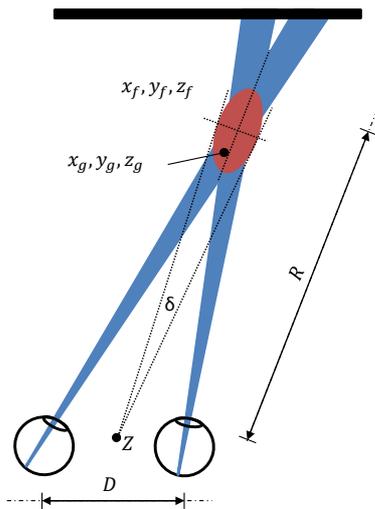


Abb. 10-2: Schematische Darstellung der Messgenauigkeit eines binokularen Eyetrackers mit der Messgenauigkeit α (blau) für die beiden Blickvektoren und dem sich daraus ergebenden elliptischen Schwankungsbereich (rot) für die Bestimmung des Blickortes (A). Die Ellipsengröße variiert in Abhängigkeit der Entfernung zwischen Betrachter und Fixationsort (B).

10.3.2 Mathematisches Grundmodell

Für die Fixationserkennung muss geprüft werden, ob der dreidimensionale Blickort innerhalb des ellipsoiden Toleranzbereiches liegt oder nicht. Falls ja, wird er der aktuellen Fixation zugeordnet und falls nicht, die aktuelle Fixation für abgeschlossen erklärt. Die Größe und dreidimensionale Lage des Ellipsoids wird vom aktuellen Fixationszentrum in Relation zu den Augen des Betrachters bestimmt (siehe Abb. 10-3).



D = Augenabstand

Z = $3D$ – Position des Zyklopauges

R = Entfernung zwischen Zyklopaugenauge und
aktuellem Fixationsort

φ = Gierwinkel (Rotation um die y – Achse)

θ = Nickwinkel (Rotation um die x – Achse)

δ = Schwellwert der Blickortabweichung

Beispiel: $\delta = 1^\circ = \frac{1}{57.3} \text{ rad}$

x_g, y_g, z_g = Blickort

x_f, y_f, z_f = Aktueller Fixationsort

Abb. 10-3: Ellipsoider Toleranzbereich der Fixationserkennung.

Die Definition des ellipsoiden Toleranzbereiches sowie die Erkennung einer dreidimensionalen Fixation erfolgt nach folgendem Berechnungsschema:

1. Im ersten Schritt wird der räumliche Mittelpunkt zwischen beiden Augen als *Zyklopaugenauge* bestimmt:

$$x_c = (x_{le} + x_{re})/2; y_c = (y_{le} - y_{re})/2; z_c = (z_{le} - z_{re})/2 \quad (10.1)$$

2. Anschließend erfolgt die Entfernungsberechnung zwischen Zyklopaugenauge und dem aktuellen Fixationsort.

$$R = \sqrt{(x_c - x_f)^2 + (y_c - y_f)^2 + (z_c - z_f)^2} \quad (10.2)$$

3. Die Parameter des Ellipsoiden werden in Abhängigkeit der Entfernung zur aktuellen Fixation und dem Schwellwert der zulässigen Blickortabweichung definiert.

$$r_x = r_y = R * \delta, r_z = r_x * \frac{R}{D} \quad (10.3)$$

4. Der Gier- und Nickwinkel zwischen zykloter Augenposition und aktuellem Fixationsort ergeben sich folgendermaßen:

$$\psi = +\tan^{-1} * \left(\frac{x_f - x_c}{z_f - z_c} \right) \quad (10.4)$$

ψ ist positiv ausgerichtet, sobald der Blickort ausgehend von der Vertikalebene des Zyklopenauges nach links ausgerichtet ist, bzw. negativ in entgegengesetzter Richtung.

$$\theta = -\tan^{-1} * \left(\frac{y_f - y_c}{z_f - z_c} \right) \quad (10.5)$$

θ ist positiv ausgerichtet, sobald der Blickort ausgehend von der Horizontalebene des Zyklopenauges nach oben ausgerichtet ist, bzw. negativ in entgegengesetzter Richtung.

5. Berechnung der Abstände zwischen neuem Blickort und aktueller Fixation.

$$\Delta x = x_g - x_f ; \Delta y = y_g - y_f ; \Delta z = z_g - z_f \quad (10.6)$$

6. Jetzt müssen die die Blickort-Abweichungen entsprechend dem durch die aktuelle Fixation vorgegebenen ellipsoiden Toleranzbereich rotiert werden.

$$\Delta x' \cong \Delta x * \cos \psi + \Delta z * \sin \psi \quad (10.7)$$

$$\Delta y' \cong \Delta y * \cos \theta + \Delta z * \sin \theta \quad (10.8)$$

$$\Delta z' \cong \Delta z * \cos \theta * \cos \psi - \Delta x * \sin \psi - \Delta y * \sin \theta \quad (10.9)$$

7. Schlussendlich wird geprüft, ob der neue Blickort sich innerhalb oder außerhalb des ellipsoiden Toleranzbereiches befindet.

$$\text{if } \left[\left(\frac{\Delta x'}{r_{x'}} \right)^2 + \left(\frac{\Delta y'}{r_{y'}} \right)^2 + \left(\frac{\Delta z'}{r_{z'}} \right)^2 \right] < 1 \quad (10.10)$$

then

*der neue Blickort befindet sich innerhalb
des ellipsoiden Toleranzbereiches*

else

*der neue Blickort befindet sich außerhalb
des ellipsoiden Toleranzbereiches*

8. Befindet sich der neue Blickort innerhalb des ellipsoiden Toleranzbereiches, wird dieser der Fixation hinzugefügt und die Fixationsparameter (3D-Position des Zentrums, Fixationsdauer) neu bestimmt.

9. Mit dem nächsten Datensample beginnt die Überprüfung der Fixationshypothese von vorn (Schritt 1).

Nach Anwendung des ellipsoiden statt kugelförmigen Toleranzbereiches können die erkannten Fixationen durch die zugrundeliegenden Blickorte genauer beschrieben werden (Abb. 10-4).

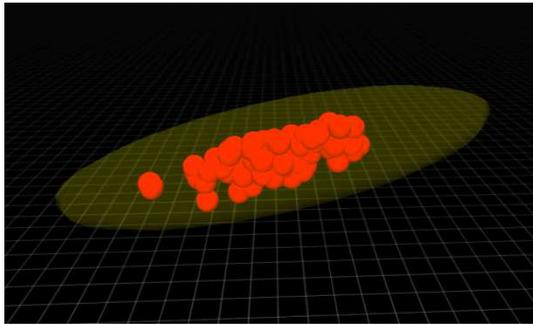


Abb. 10-4: Dreidimensionale Fixationsbestimmung durch Aggregation räumlicher Blickorte (rot) innerhalb eines ellipsoiden Toleranzbereiches (gelb).

10.3.3 Fixation3D.dll

Der Algorithmus zur Identifikation räumlicher Fixationen wurde in die Funktionsbibliothek *Fixation3D.dll* implementiert und ist damit in verschiedene Programmierumgebungen integrierbar. Die Berechnung der dreidimensionalen Fixationsparameter erfolgt dabei nach dem vorgestellten ellipsoiden Fixationsmodell und baut auf die zweidimensionale Logik der Fixationserkennung des *Eyegaze Analysis Systems* (LC Technologies, 2014) auf. Zunächst wird der Algorithmus initialisiert und dabei die Schwellwertparameter für die minimale Anzahl an Datensamples zur Bestimmung einer gültigen 3D-Fixation übergeben. Aus der Aufzeichnungsfrequenz des Eyetrackers und der Anzahl minimaler Samples leitet sich die minimale Dauer einer 3D-Fixation ab. Bei einer Aufzeichnungsrate von beispielsweise 500 Hz und einer Initialisierung mit 50 Messpunkten ergibt sich eine minimale Fixationsdauer von 100 ms. Bei erfolgreicher Initialisierung gibt die Funktion zur Bestätigung den Wert 1 zurück, andernfalls 0. Anschließend werden dem Algorithmus die Parameter für jedes Datensample übergeben. Als Rückgabewert erhält man den Wert 0 (Fixation ist nicht bestätigt), den Wert 1 (Fixation ist bestätigt aber noch nicht abgeschlossen) oder den Wert 2 (Fixation ist abgeschlossen). Zusätzlich gibt der Algorithmus die Parameter der aktuellen Fixation sowie der vorangegangenen Sakkade aus. Die Funktionsaufrufe erfolgen mit der in Listing 10-1 angeführten Syntax. Für eine genauere Beschreibung der einzelnen Parameter und Rechenoperationen wird an dieser Stelle auf den Quellcode der Funktionen verwiesen.

```

function Init3DFixation(iMinimumSamples: Integer): Integer; stdcall;

function Calculate3DFixation( bGazeVectorFound: Integer;
                             fXLeftEye, fYLeftEye, fZLeftEye,
                             fXRightEye, fYRightEye, fZRightEye,
                             fXGaze, fYGaze, fZGaze: Single;
                             fAccuracyAngleRad: Single;
                             iMinimumSamples: Integer;

                             out pbGazepointFoundDelayed: Integer;
                             out pfXGazeDelayed, pfYGazeDelayed, pfZGazeDelayed: Single;
                             out pfXFixDelayed, pfYFixDelayed, pfZFixDelayed: Single;
                             out pfXEllipsoidRDelayed, pfYEllipsoidRDelayed,
                             pfZEllipsoidRDelayed: Single;
                             out pfEllipsoidYawDelayed, pfEllipsoidPitchDelayed: Single;
                             out piSacDurationDelayed,
                             piFixDurationDelayed: Integer):Integer; stdcall;

```

Listing 10-1: Funktionsaufrufe zur Initialisierung der ellipsoiden Fixationserkennung sowie zur Bestimmung dreidimensionaler Fixationen. Die Rückgabewerte beinhalten die Parameter der aktuellen 3D-Fixation sowie der vorangegangenen Sakkade.

Zur Implementierung der Berechnungsvorschrift wurde eine Software entwickelt, mit der die in Studie II und III erhobenen Daten importiert, 3D-Blickorte sowie 3D-Fixationen berechnet und innerhalb einer OpenGL-Umgebung (Abb. 10-5) visualisiert werden können.

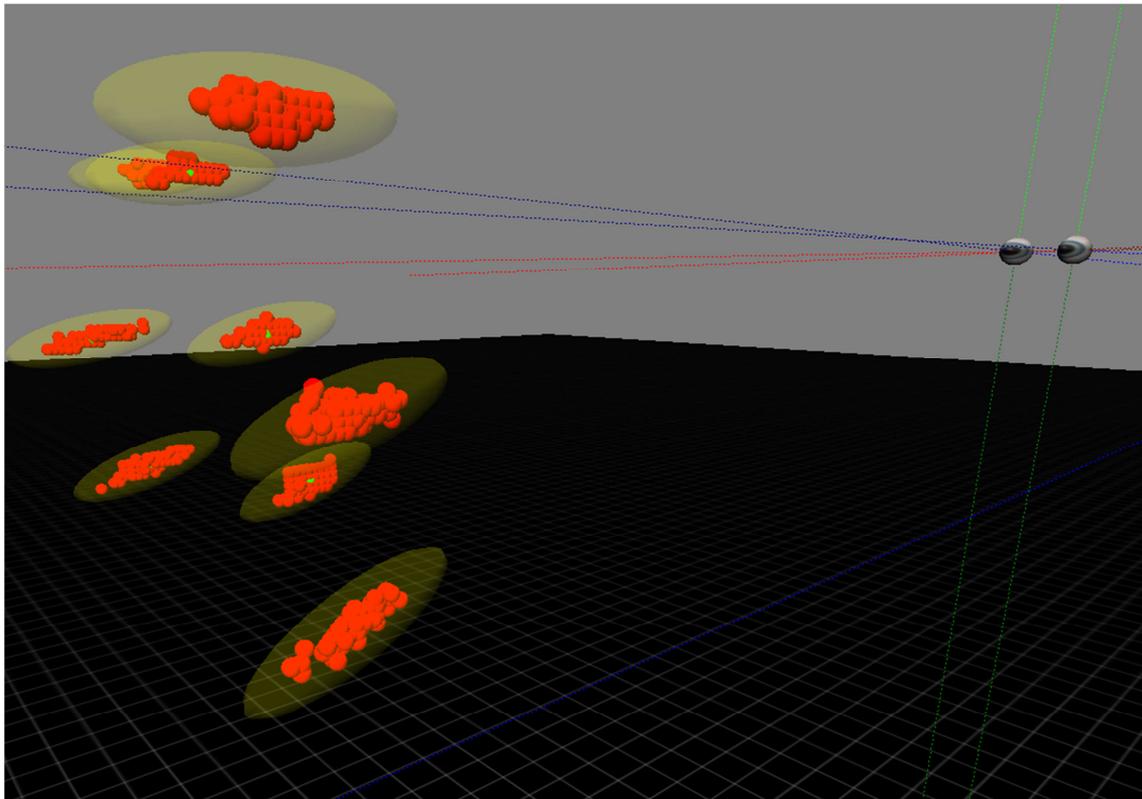


Abb. 10-5: Visuelle Darstellung der berechneten 3D-Fixationen durch Aggregation räumlicher Blickorte (rot) innerhalb eines ellipsoiden Toleranzbereiches (gelb) mit dem Fixationszentrum (grün).

10.4 Berechnung und Analyse dreidimensionaler Fixationen

Die mit den Algorithmen *FixEye*, *VarEye* sowie *GazeVec* ermittelten 3D-Blickorte aus Studie II und III sollen der ellipsoiden Fixationserkennung übergeben und die daraus berechneten Fixationsparameter analysiert werden. Da es sich um einen Ansatz nach I-DT handelt, müssen für die Fixationsbestimmung zuvor die Schwellwerte für die minimale Sampleanzahl einer Fixation sowie der örtlichen Genauigkeit der Blickortmessung festgelegt werden.

10.4.1 Zeitlicher und örtlicher Schwellwert

Nach Salvucci und Goldberg (2000) eignen für den zeitlichen Parameter minimale Fixationsdauern zwischen 100 – 200 ms sowie $0.5 - 1^\circ$ für den 2D-Dispersions-Schwellwert. Der Softwareentwickler des Eyetracking-Systems empfiehlt eine minimale Fixationsdauer von 100 ms sowie einen 2D-Dispersions-Schwellwert von 0.7° (Interactive Minds, 2011). Unter Berücksichtigung der angeführten Annahmen wurde die minimale Fixationsdauer auf 100 ms festgelegt. Der Dispersions-Schwellwert bezieht sich in beiden Fällen auf die maximal zulässige Abweichung eines gemessenen Blickortes vom aktuellen Fixationszentrum innerhalb der zweidimensionalen Messebene. Betrachtet man die Fixation als symmetrische Kreisform, entspricht der Dispersions-Schwellwert dem Radius des Kreises bzw. wird die Fixation mit einem Durchmesser zwischen $1 - 2^\circ$ definiert.

Bei der ellipsoiden Fixationserkennung ist dreidimensional keine Symmetrie gegeben. Hier muss der Dispersions-Schwellwert als ellipsoide Fixationshülle betrachtet werden. Die Dimension der Hülle wird durch den örtlichen Schwellwertparameter δ definiert, der horizontalen Ausdehnung des Ellipsoiden im Zentrum. Die Tiefe berechnet sich aus Augenabstand und Entfernung der Fixation. Um sowohl die Vorgaben von Salvucci und Goldberg als auch die des Herstellers einzubeziehen, wurden für die ellipsoide Fixationserkennung Schwellwerte zwischen 1° und 2° untersucht. Diese Auswahl stimmt mit den in Kapitel 2 vorgestellten Erkenntnissen überein, dass scharfes Sehen auf einen Bereich von etwa 2° um den fixierten Punkt aufgrund des optischen Apparates und der Rezeptorverteilung eingegrenzt ist (Parkin, 2014).

10.4.2 3D-Fixationsparameter

Für die Auswertung der dreidimensionalen Fixationsparameter waren insbesondere die Genauigkeit des Fixationszentrums sowie die berechnete Fixationsdauer von Interesse. Dazu

wurden alle präsentierten Stimuli, welche die Versuchspersonen als „fixiert“ markierten, den detektierten gegenübergestellt. Da Fixationen durch Blinks oder ungültige Messungen unterbrochen werden und zu falschen Ergebnissen führen, sind diese vor den weiteren Analysen ausgeschlossen worden. Dadurch reduzierte sich die Datenbasis für Studie II auf 414 und Studie III auf 1.154 Trials. Allerdings gab der Fixationserkennungsalgorithmus sowohl im realen als auch virtuellen Kontext durchweg mehr detektierte Fixationen aus. Abb. 10-6 A und B verdeutlichen, dass die Anzahl erkannter Fixationen in realer und virtueller Umgebung mit einer Erweiterung des örtlichen Toleranzbereiches abnahm, jedoch die Anzahl der Trials nicht erreichte. Dementsprechend wurden innerhalb eines Trials mehrere Fixationen detektiert.

Fixationsdauern

Als logische Konsequenz mehrfach geteilter Fixationen ist mit kürzeren Fixationsdauern zu rechnen. Dies bestätigte sich in beiden Experimenten. Im realen Setting konnten in einer 3 (Algorithmus) x 11 (Toleranzbereich) faktoriellen ANOVA mit Messwiederholung höchst signifikante Unterschiede in den Fixationsdauern für Algorithmus, $F(2,22) = 14.922$, $p < .001$, $\eta^2 = .015$, sowie Toleranzbereich, $F(10,110) = 40.618$, $p < .001$, $\varepsilon = .239$, $\eta^2 = .728$, nachgewiesen werden. Die Fixationsdauern unterschieden sich auf Datenbasis der Algorithmen *GazeVec* zu *FixEyePos* und *GazeVec* zu *VarEyePos* voneinander, wobei Fixationen mit den Sampledaten von *GazeVec* durchschnittlich längere Dauern (725 ms) aufwiesen als *FixEyePos* (681 ms) und *VarEyePos* (658 ms). Über alle drei Algorithmen hinweg berechnete die Fixationserkennung für den kleinsten Toleranzbereich die kürzesten Fixationsdauern, welche mit zunehmenden Schwellwert größer wurden (Abb. 10-6 C).

Ähnliche Ergebnisse zeigten sich im virtuellen Experiment. Höchste signifikante Unterschiede waren für Algorithmus, $F(2,44) = 79.774$, $p < .001$, $\eta^2 = .062$, und Toleranzbereich, $F(10,220) = 70.910$, $p < .001$, $\varepsilon = .193$, $\eta^2 = .602$, nachweisbar. Auch hier wurden auf Datenbasis des Algorithmus *GazeVec* im Durchschnitt mit 910 ms längere Fixationen erkannt als bei *VarEyePos* mit 815 ms bzw. *FixEyePos* mit 785 ms. Ebenso stiegen die Fixationsdauern mit zunehmendem örtlichem Toleranzbereich an (Abb. 10-6 D). Daraus ließ sich ableiten, dass die ellipsoide Fixationserkennung in beiden Experimenten auf Basis der *GazeVec*-Daten mehr Samples einer Fixation zuordnete als bei *FixEyePos* und *VarEyePos*.

Genauigkeit der Fixationszentren

Eine 3 (Algorithmus) x 11 (Toleranzbereich) faktorielle ANOVA mit Messwiederholung ergab signifikante Abweichungsunterschiede der Fixationszentren für den zugrundeliegenden Algorithmus der 3D-Sampledaten aus Studie II, $F(2,22) = 20.535$, $p = .001$, $\varepsilon = .548$, $\eta^2 = .59$, allerdings nicht für den Toleranzbereich, $F < 1$. Die Werte unterschieden sich zwischen

FixEye und *VarEye* sowie zwischen *FixEye* und *GazeVec*. Die geringsten Zentren-Deviations wiesen Fixationen aus *FixEye* mit durchschnittlich 17 mm auf. Größere Ungenauigkeiten ergaben die beiden andern Sampledaten-Algorithmen mit jeweils 22 mm. Diese Effekte sind anhand der in Studie II festgestellten Ungenauigkeiten der 3D-Sampledaten aufgrund der Parametrisierung des Eyetracking-Systems erklärbar. Die Abweichungen blieben innerhalb der Datenbasis über die getesteten Toleranzgrenzen annähernd gleich (Abb. 10-6 E).

In Studie III gab es hingegen keine Unterschiede in den Abweichungen bezüglich des Sampledatenalgorithmus, $F(2,44) = 1.498$, $p = .235$, $\varepsilon = .596$, aber höchst signifikante Differenzen für den Toleranzbereich, $F(10,220) = 8.591$, $\varepsilon = .173$, $p < .001$, $\eta^2 = .084$. Die Fixationszentren wichen in allen drei Algorithmen durchschnittlich mit 17 mm ab. Eine Verbesserung der Genauigkeit um ca. 5 mm in *VarEye* und *GazeVec* deutet auf eine Auswirkung des Treiberupdates sowie Neujustierung des Eyetrackers im Vergleich zum vorangegangenen Experiment hin. Über die verschiedenen Toleranzgrenzen verringerten sich die Abweichungen von 18 mm bei 1° auf 16 mm bei 2° örtlichem Schwellwert (Abb. 10-6 F).

Fixationsdauer und 3D-Genauigkeit in Abhängigkeit der Fokussierebene

Für die Analyse der Fixationsparameter in unterschiedlichen Tiefenebenen, erfolgte die ellipsoide Fixationserkennung auf Basis der 3D-Sampledaten des *FixEye*-Algorithmus bei einer minimalen Fixationsdauer von 100 ms und einer maximal zulässigen örtlichen Toleranz von 2°. Es wurden die 3D-Sampledaten des *FixEye*-Algorithmus gewählt, da die Fixationsparameter in beiden Experimenten ähnliche Werte aufwiesen.

In einer einfaktoriellen (Tiefenebene) ANOVA mit Messwiederholung zeigten sich höchst signifikante Unterschiede der Fixationsdauern sowohl im realen, $F(4,40) = 14.235$, $p < .001$, $\varepsilon = .669$, $\eta^2 = .587$, als auch virtuellen Setting, $F(3,66) = 20.963$, $p < .001$, $\varepsilon = .703$, $\eta^2 = .488$. Dabei hatten Fixationen, die näher am Betrachter lagen, eine größere Dauer als weiter entfernte (Abb. 10-6 G). Offensichtlich wurden im Nahbereich mehr Samples einer Fixation zugeordnet. Ebenfalls höchst signifikante Unterschiede über die Tiefenebene waren für die 3D-Abweichungen der Fixationszentren in Studie II, $F(4,40) = 46.89$, $p < .001$, $\varepsilon = .521$, $\eta^2 = .824$, sowie Studie III, $F(3,66) = 33.571$, $p < .001$, $\varepsilon = .780$, $\eta^2 = .604$, nachweisbar. Näher liegende Fixationen hatten geringere dreidimensionale Abweichungen als Fixationen in größerer Entfernung (Abb. 10-6 H).

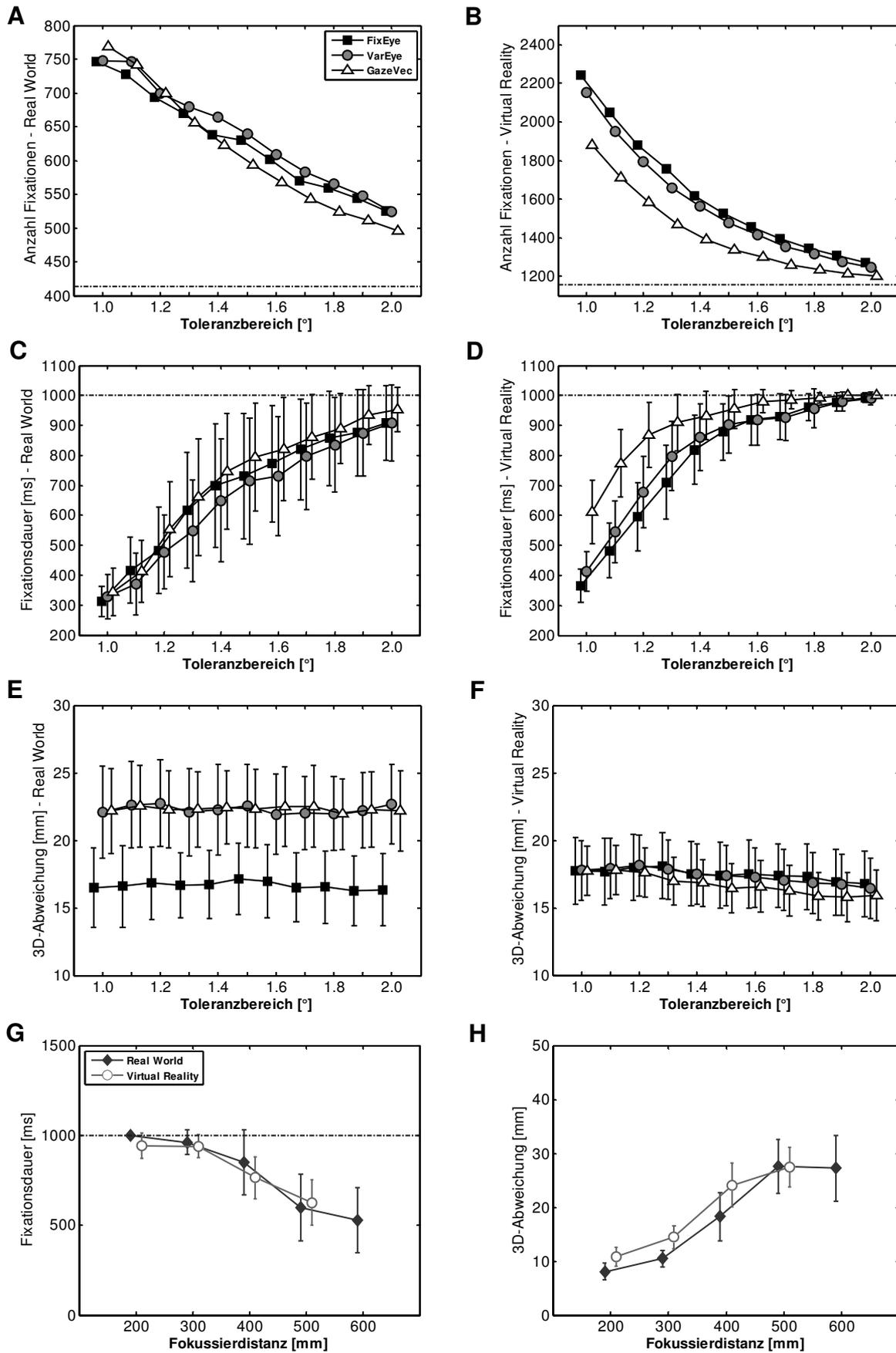


Abb. 10-6: Berechnete Fixationsparameter aus Studie II und III bei einer min. Fixationsdauer von 100 ms in unterschiedlichen Schwellwertbereichen: Anzahl erkannter Fixationen (A) und (B), Abhängigkeit der Fixationsdauer (C) und (D) sowie 3D-Abweichung der Fixationszentren (E) und (F). Berechnete Fixationsparameter in Abhängigkeit der Fokussierebene in Bezug auf Fixationsdauer (G) und 3D-Genauigkeit (H). Fehlerbalken kennzeichnen das 95%-CI.

10.5 Fazit

Mit der entwickelten Methode sind Blickbewegungen unabhängig einer Kollision mit den Objekten einer Szene bestimmbar (*Anforderung 1*). Der Tiefenparameter des Blickortes wird in die Berechnung einbezogen (*Anforderung 2*). Der Algorithmus identifiziert dreidimensionale Fixationen (*Anforderung 3*) und nimmt, sofern kein Blink stattfindet, dazwischenliegende Sakkaden an. Tiefenschwankungen des 3D-Blickortes werden in Abhängigkeit der Entfernung zu den Augen berücksichtigt (*Anforderung 4*) und Abstände zwischen dreidimensionalen Blickorten räumlich berechnet sowie innerhalb zeitlicher und örtlicher Toleranzgrenzen einer Fixation zugeordnet (*Anforderung 5*). Damit erfüllt die ellipsoide Fixationserkennung alle Anforderungen aus Kapitel 10.2.

Für eine Analyse relevanter Fixationsparameter wurde die ellipsoide Fixationserkennung mit einer minimalen Fixationsdauer von 100 ms in örtlichen Toleranzgrenzen zwischen $1 - 2^\circ$ durchgeführt. Die Datenbasis bestand aus den 3D-Samples, die in den beiden Studien II und III erhoben wurden. Sowohl für das Experiment in realer als auch virtueller Umgebung nahm die Anzahl erkannter Fixationen mit Vergrößerung des Toleranzbereiches ab, gleichzeitig stiegen die Fixationsdauern an. Hierbei wurden mehr Samples einer Fixation zugeordnet. Die 3D-Abweichungen der Fixationszentren lagen im realen Experiment mit *FixEye* und im virtuellen Experiment mit allen drei Sampledatenalgorithmen bei durchschnittlich 17 mm. Mehr Abweichungen gab es im realen Setting bei *VarEye* und *GazeVec* mit ca. 22 mm aufgrund ungenauer Eyetracking-Daten. Würde man die Studie II mit der verbesserten Treibersoftware erneut durchführen, wäre mit geringeren dreidimensionalen Blickort- und schlussendlich auch Fixationsabweichungen zu rechnen.

Die Distanz zwischen Augen und Fixationsort hatte einen signifikanten Einfluss auf die ermittelten Fixationsparameter. In geringer Entfernung wurden weniger Fixationen von höherer Dauer berechnet. Größere Messwertschwankungen in den 3D-Sampledaten führten zu Fixationsteilungen, da hierbei der örtliche Toleranzbereich häufiger überschritten wurde. Ob diese auf Messwertschwankungen des Eyetracking-Systems zurückzuführen sind oder die Versuchspersonen eine dreidimensionale Fixation über die Dauer von 1.000 ms nicht aufrechterhalten konnten, war nicht abschließend erklärbar. Auch die Genauigkeit der Fixationen hing von der betrachteten Tiefenebene ab. Die getesteten Fixationszentren auf Basis der *FixEye*-Sampledaten hatten im Nahbereich 3D-Abweichungen von 8–12 mm welche mit Entfernungszunahme auf ca. 27 mm anstiegen.

Sowohl die Auswahl der Schwellwertgrenzen als auch die Entfernung zu den betrachteten Stimuli haben einen Einfluss auf die dreidimensionale Fixationserkennung und müssen im Setting berücksichtigt werden. Wählt man beispielsweise den örtlichen Schwellwert sehr klein, werden Fixationen häufiger geteilt als in größeren Toleranzbereichen. Dafür lassen sich aber Objekte in geringen Abständen positionieren und einer Fixation zuordnen. Allerdings ist hierbei mit kürzeren Fixationsdauern zu rechnen und genau das Gegenteil ist der Fall, wenn man die örtliche Schwellwertgrenze großräumiger wählt.

Mit der entwickelten Methode konnten dreidimensionale Blickbewegungen anhand des ellipsoiden Fixationserkennungsalgorithmus auf Datenbasis der berechneten 3D-Blickorte aus den Studien II und III bestimmt werden. Insbesondere mit der verbesserten Eyetracking-Technologie wurden 3D-Abweichungen von weniger als 20 mm erreicht. Die Methode eignet sich zur Bestimmung räumlicher Blickbewegungen sowohl in realer als auch virtueller Umgebung.

11 Zusammenfassung und Perspektiven

Im theoretischen Teil dieser Arbeit wurden Zusammenhänge zwischen visueller Wahrnehmung und den damit verbundenen menschlichen Blickbewegungen beschrieben. Insbesondere die Aufeinanderfolge von Fixationen und Sakkaden repräsentieren das menschliche Blickverhalten und geben nicht nur Aufschluss über die lokale Ausrichtung der Informationsaufnahme, sondern lassen auch Rückschlüsse auf die Verarbeitungstiefe zu. In zahlreichen Untersuchungen wurden dazu Blickbewegungen in verschiedenen Kontexten gemessen und Rückschlüsse auf kognitive Prozesse gezogen. Allerdings blieben in bisherigen Untersuchungen, aufgrund der auf zwei Dimensionen eingegrenzten Blickortmessung, die dritte Dimension und damit verbundene Tiefeninformationen weitestgehend unberücksichtigt, obwohl menschliches Blickverhalten bei räumlicher Wahrnehmung vorwiegend dreidimensional stattfindet.

Das dritte Kapitel beschreibt die räumlich Wahrnehmung und die dafür notwendigen Tiefenkriterien. Anders als bei okulomotorischen, monokularen oder bewegungsinduzierten Informationen benötigt der Mensch für die Tiefeninformationen des stereoskopischen Sehens gleichzeitig beide Augen. Die durch retinale Abbildungsunterschiede erzeugte Querdissipation ermöglicht nicht nur in der realen Umgebung sondern auch in stereoskopischen Visualisierungen eine unterschiedliche Entfernungswahrnehmung. Für eine gleiche Tiefenwahrnehmung über verschiedene Personen hinweg muss bei der Stereobildgenerierung der individuelle Augenabstand einbezogen werden. Des Weiteren wurde die praktische Umsetzung der Querdissipation anhand von Polarisationsfilter- und Shutterverfahren zur Generierung virtueller 3D-Umgebungen detailliert beschrieben. Bei stereoskopisch projizierten Szenen gibt es technologische Besonderheiten zu beachten. Abweichungen zwischen Vergenz- und Akkommodationsebene sowie perspektivische Verzerrungen können zu visueller Überforderung und Wahrnehmungsproblemen führen.

Die Grundlagen der videobasierten Blickbewegungsmessung wurden im vierten Kapitel erklärt. Wichtigste Kenngrößen sind die Pupille bzw. deren Mittelpunkt und die Position des durch eine Infrarotlichtquelle entstehenden Reflexpunktes auf der Cornea. Anhand der PCCR-Methode lässt sich aus diesen beiden Parametern anhand vektorbasierter Algorithmen die Blickrichtung bestimmen. Jedoch beziehen sich bisherige Blickortmessungen vorwiegend auf zweidimensionale Ebenen. Um visuelle Wahrnehmung räumlich analysieren zu können, ist es erforderlich Blickorte dreidimensional zu berechnen, daraus Blickbewegungen abzuleiten und diese den wahrgenommenen Objekten der Umgebung zuzuordnen. In realer Umgebung

gestaltet sich die Umsetzung komplexer als in virtueller, denn dazu muss sowohl die Position des sich frei bewegenden Betrachters als auch die geometrische Ausrichtung aller in der Umgebung vorhandenen Objekte für eine korrekte Blickortzuweisung bekannt sein.

In virtuellen 3D-Umgebungen sind die Informationen über die Geometrie der Szene und der darin enthaltenen Objekte selbst bei dynamischen Veränderungen zu jeder Zeit verfügbar. Allerdings ist es notwendig, die stereoskopische Visualisierungstechnologie mit einer Blickbewegungsmessung zu kombinieren. Je nach Kombination aus Blickbewegungsmessung und Stereoprojektion wurden bisher speziell auf das Setting ausgerichtete Konzepte der räumlichen Blickortbestimmung und Objektzuweisung entwickelt und angewandt. Dabei stellten sich in Bezug auf die stereoskopische Visualisierung, den eingesetzten Eyetracking-Systemen und den resultierenden Daten verschiedene Vor- und Nachteile heraus. Eine einheitliche Methode zur Bestimmung dreidimensionaler Blickorte und Blickbewegungen gibt es bislang nicht. Aus den Erkenntnissen vergangener Forschung und den vorhandenen technischen Möglichkeiten der Blickbewegungsmessung wurde das Ziel dieser Arbeit abgeleitet. Es sollte eine Methode zur Bestimmung dreidimensionaler Blickorte und Blickbewegungen in realer als auch virtueller Umgebung zur Analyse der räumlichen visuellen Wahrnehmung entwickelt werden.

Studie I: Stereoskopische Visualisierung und Eyetracking

Zunächst musste geprüft werden, inwieweit Eyetracking in virtueller Umgebung unter Verwendung stereoskopischer Visualisierungstechnologien durchführbar ist und welche ausgegebenen Kenngrößen für eine 3D-Blickortberechnung in Betracht kommen. Die Untersuchungen fanden mit binokularen Remote Eyetrackern statt, um einerseits die Nutzer nicht durch am Kopf angebrachte Hardware zu beeinflussen und andererseits die räumliche Komponente des Blickortes aus der Blickachsenausrichtung bzw. Vergenzstellung beider Augen ableiten zu können. Für die Kombination der Blickbewegungsmessung mit einer Stereoprojektion wurden aufgrund der hohen örtlichen Auflösungen und einer weitestgehend unverfälschten Farbwiedergabe zur Bildtrennung das Polarisationsfilter- sowie das Shutterverfahren getestet. Die durchgeführte Studie I sollte Aufschluss darüber geben, ob eine Blickbewegungsmessung durch die für die stereoskopische Bildtrennung notwendigen 3D-Brillen hindurch möglich ist und ob die Kombination einen Einfluss auf die erhobenen Eyetracking-Daten insbesondere auf deren Genauigkeit hat. In einem Experiment wurden in mehreren Durchgängen Stereoprojektion, Eyetracking-System sowie die Position der angezeigten Stimuli variiert. Die Stimulusdarbietung erfolgte auf einem 2D-Monitor. In beiden Visualisierungsverfahren verliefen die Kalibrierrouniten erfolgreich. Obwohl die von den

Eyetracking-Systemen ausgegebenen Kalibrierabweichungen beim zusätzlichen Tragen einer 3D-Brille zunahmen, konnten keine signifikanten Unterschiede auch im Vergleich zur Bedingung ohne Brille nachgewiesen werden.

Unter Einsatz des Shutter-Verfahrens kam es jedoch zu Darstellungsproblemen. Die Übertragung des Synchronisationssignals zum Öffnen und Schließen der jeweiligen Blende für linkes und rechtes Auge erfolgte durch ein Infrarotlichtsignal zwischen dem per USB-Schnittstelle direkt mit dem Computer verbundenen Sender und dem in den linken Brillenbügel integrierten Empfänger. Bei einer gleichzeitigen Aktivierung von Blickbewegungsmessung und stereoskopischer Visualisierung wurde das infrarote Licht des Eyetrackers ebenfalls vom Empfänger der Shutterbrille erfasst und bewirkte ein Öffnen beider Blenden unabhängig vom dargebotenen Augenbild. Zur Lösung dieses Problems wurde der Empfänger im Brillenbügel mit einer Abdeckung versehen, die den Empfang von Signalen nur aus seitlicher Richtung zuließ.

Die Analyse der Sampledaten ergab keine signifikanten Unterschiede im Datenausfall durch die Verwendung der Polarisationsfilter- oder Shutter-Technologie, jedoch nahmen die 2D-Blickortabweichungen beim Tragen einer 3D-Brille zu und erreichten beim EAS das Signifikanzniveau. Diese Ergebnisse bestätigten sich auch in den Abweichungen der 2D-Fixationszentren, wobei diese ca. 0.06° - 0.08° näher an den Stimuluspositionen lagen als die 2D-Blickorte. Für beide Eyetracker konnten Genauigkeitsunterschiede in verschiedenen Monitorbereichen auch unabhängig der Visualisierungsverfahren nachgewiesen werden. Im Zentrum sind sowohl Blickort- als auch Fixationsabweichungen geringer als in den Randbereichen. Durch das Tragen einer 3D-Brille stieg die Pupillengröße höchst signifikant an. Begründen lässt sich die Pupillenerweiterung durch den geringeren Lichteinfall, welcher aus der Filterung nicht ausgerichteter Lichtstrahlen beim Polarisationsfilter- bzw. einer Halbierung der einfallenden Lichtmenge beim Shutterverfahren resultiert.

Die Ergebnisse zeigten, dass zweidimensionale Blickbewegungsmessungen in Kombination mit stereoskopischen Visualisierungen anhand der getesteten Technologien durchführbar sind. Um Blickbewegungen in einer realen oder virtuellen Umgebung dreidimensional bestimmen zu können, müssen die dafür notwendigen Blickorte räumlich berechnet werden.

Algorithmus zur 3D-Blickortbestimmung

Zur Berechnung dreidimensionaler Blickorte wurde das Anforderungsprofil an einen universellen Algorithmus erstellt und anschließend mit einem vektorbasierten Ansatz umgesetzt. Besonderheit hierbei ist die Berechnung der Blickvektoren anhand der Augen- bzw. Foveaposition und binokularen Eyetracking-Daten. Da sich Blickvektoren geometrisch selten schneiden sondern überwiegend windschief zueinander liegen, wird derjenige Vektor berechnet, der den geringsten Abstand zwischen den beiden Blickvektoren einnimmt. Der Mittelpunkt dieses Verbindungsvektors wird als dreidimensionale Blickposition angenommen. Um diese Berechnungsmethode über verschiedene Programmiersprachen hinweg zur Verfügung zu stellen, erfolgte die Implementierung in die Funktionsbibliothek *Gaze3D.dll*. Wie genau dreidimensionale Blickorte anhand dieses Algorithmus berechnet werden können, sollte in zwei Studien sowohl in realer als auch stereoskopisch projizierter VR-Umgebung überprüft werden.

Studie II: Eyetracking in realer Umgebung

Ziel dieser Studie war die Berechnung dreidimensionaler Blickorte anhand des vektorbasierten Ansatzes. Dazu wurden Probanden in einem Experiment reale 3D-Fixation dargeboten und gleichzeitig Blickdaten mit einem Remote Eyetracker erhoben. Die Blickortbestimmung erfolgte in Abhängigkeit unterschiedlicher Eyetracking-Parameter. Zur Durchführung der Studie wurde eine Experimentalsoftware konzipiert sowie implementiert, um durch das Experiment zu leiten, die Blickbewegungsmessung zu steuern, Eyetracking-Parameter abzufragen, 3D-Blickorte zu berechnen und sämtliche Daten zeitlich synchronisiert zu speichern.

Das Hauptaugenmerk der Analysen bezog sich auf die Genauigkeit der ermittelten 3D-Blickorte bzw. inwiefern diese von den Stimuluspositionen abwichen. Die Resultate ergaben höchst signifikante Unterschiede in den 3D-Abweichungen bzgl. der verwendeten Datenbasis und des dafür vorgesehenen Algorithmus. Die dreidimensionalen Abweichungen der Sampledaten waren in den Berechnungen mit manuell ausgemessener Augenposition am geringsten (16 mm) und nahmen entgegen der Hypothese in den Bedingungen mit den vom Eyetracker zur Verfügung gestellten Messwerten von Cornea-Zentren (23 mm) bzw. den Blickvektoren (22 mm) zu. Es wurden signifikante Unterschiede zwischen den getesteten Algorithmen sowie den verschiedenen Tiefenebenen nachgewiesen, wobei die Abweichungen mit zunehmender Fokussierentfernung anstiegen. Dies wird auf den geringer werdenden Vergenzwinkel zurückgeführt. Je kleiner der Vergenzwinkel, umso stärker wirken sich Messwertschwankungen auf die berechnete 3D-Blickposition aus.

Die Erkenntnisse dieser Studie wurden mit dem Hersteller des Eyetracking-Systems diskutiert, insbesondere die Abweichungen in den Berechnungen mit denen vom Eyetracker ausgegebenen Angaben zu Cornea-Zentren und Blickvektoren. LC Technologies, Inc. nahm daraufhin Veränderungen der Treibersoftware vor und kalibrierte die Kameras auf das Koordinatensystem des Eyetrackers neu. Der vektorbasierte Ansatz hatte sich in diesem Experiment bewährt und ergab insbesondere im nahen Fokussierbereich dreidimensionale Abweichungen von weniger als 20 mm.

Studie III: Eyetracking in einer stereoskopischen VR-Umgebung

Die vorhergehende Studie wurde innerhalb einer virtuellen Umgebung nachgebildet. Für die stereoskopische Visualisierung kam das in Studie I getestete Shutterverfahren zum Einsatz. In einem Experiment sollten die Probanden virtuell dargebotene Fixationsobjekte mit den Augen fixieren. Um die virtuelle Szene realitätsnah abzubilden, wurden die individuellen Augenpositionen ausgemessen und gingen in die Berechnung der stereoskopischen Einzelbilder für rechtes und linkes Auge ein. Bereits in den Vorversuchen zeigten sich verschiedene Effekte bei der Kombination von Eyetracking und Shutterverfahren. Als erstes stellte sich heraus, dass die 3D-Shutterbrille einen Teil des Infrarotlichtes filterte und damit die Voraussetzungen für eine Messung nach der PCCR-Methode nicht mehr gegeben waren. Zur Lösung dieses Problems wurde der Abstand zwischen Eyetracker und Augen verkürzt und damit die Intensität des infraroten Lichts erhöht. Des Weiteren fiel es den getesteten Probanden schwer, sprunghafte Positionsänderungen der Fixationsobjekte insbesondere bei Tiefenänderungen von Fern auf Nah zu verfolgen bzw. die beiden Halbbilder mit jetzt größerer Disparität zu einem dreidimensionalen Bild zusammenzuführen. Für eine langsamere Anpassung an die Disparität wurde eine dynamische Bewegung der Fixationsobjekte aus der Monitorebene in Richtung der Endposition implementiert.

Im Gegensatz zur vorangegangenen Studie unterschieden sich im virtuellen Setting die Blickortabweichungen der Algorithmen nicht voneinander und betrug dreidimensional nur noch 16-17 mm. Die Nachjustierung des Eyetracking-Systems führte zu präziseren Messwerten der Blickachsenparameter und somit zu einer höheren Genauigkeit der berechneten 3D-Blickorte. Auch im virtuellen Setting hingen die Abweichungen des 3D-Blickortes und Tiefenparameters von der Fokussierdistanz ab. Je geringer die Entfernung, desto höher die Genauigkeit. Ebenso gab es einen Zusammenhang zwischen Fokussierebene und Vergenz. In der stereoskopisch projizierten VR-Umgebung wurde der Vergenzwinkel trotz Akkommodation innerhalb der Projektionsebene dem dreidimensional wahrgenommenen Objekt angepasst. Er verringerte sich gleichermaßen mit zunehmender

Fokussierdistanz. Damit konnte gezeigt werden, dass die Berechnung der dreidimensionalen Blickposition nach dem vektorbasierten Ansatz auch in einer stereoskopisch projizierten VR-Umgebung möglich ist. Die Ergebnisse aus beiden Studien bildeten die Basis für die Entwicklung eines neuen, dreidimensionalen Fixationserkennungsalgorithmus.

Bestimmung dreidimensionaler Blickbewegungen

Zu Beginn dieses Kapitels wurden die aus der Literatur bekannten Methoden zur Bestimmung zweidimensionaler Blickbewegungen erklärt und geprüft, inwieweit sich diese für eine räumliche Erkennung erweitern lassen. Naheliegender wäre es, bei einer abstandsbezogenen Fixationserkennung eine Kugel statt eines Kreises als Fixationshülle zu wählen. Allerdings wirken sich horizontale und vertikale Disparitäten der Blickorte für linkes und rechtes Auge innerhalb der Messebene verstärkt auf die berechnete Blickorttiefe aus, so dass die Verteilung der dreidimensionalen Blickorte nicht durch eine Kugel sondern vielmehr eine Ellipse beschrieben wird. Daraus ableitend wurde ein ellipsoider Fixationserkennungsalgorithmus entwickelt. Eine Besonderheit des ellipsoiden Ansatzes liegt in der Geometrie. Im Gegensatz zur Kugel ist ein Ellipsoid nicht radialsymmetrisch und muss durch Anpassung des Gier- und Nickwinkels stets auf die Blickvektoren des Probanden angepasst werden. Eine Veränderung der örtlichen Fixationsparameter durch die Zuordnung weiterer 3D-Blickpositionen erfordert eine Neuausrichtung des Ellipsoiden und damit auch Ortskenntnis über die Position beider Augen. Weiterhin muss durch den Anstieg dreidimensionaler Blickortabweichung mit zunehmender Fokussierdistanz die Dimensionierung der Fixationshülle in Abhängigkeit der Entfernung variiert werden. Weiter entfernte Fixationen benötigen aufgrund des geringeren Vergenzwinkels und der damit verbundenen Zunahme an Messwertschwankungen einen größeren Toleranzbereich als nahe liegende. Diese und auch alle weiteren definierten Anforderungen wurden in die ellipsoide Fixationserkennung integriert und der Algorithmus die Funktionsbibliothek Fixation3D.dll implementiert.

Für die räumliche, dispersionsbasierte Blickbewegungserkennung sind sowohl ein zeitlicher als auch örtlicher Parameter für die Bestimmung einer Fixation erforderlich. Da es bislang noch keinerlei Erkenntnisse im dreidimensionalen Bereich gab, wurden die in Studie II und III ermittelten 3D-Blickorte der ellipsoiden Fixationserkennung übergeben und die daraus berechneten Fixationsparameter analysiert. Dabei ist eine minimale Fixationsdauer von 100 ms festgelegt und die örtliche Toleranzgrenze zwischen $1 - 2^\circ$ variiert worden. In beiden Experimenten nahm die Anzahl erkannter Fixationen mit Vergrößerung der Fixationshülle ab, gleichzeitig stiegen die Fixationsdauern an. Dies lässt sich darauf zurückzuführen, dass bei größerem Volumen auch mehr Samples einer Fixation zugeordnet wurden.

Im realen Experiment mit manuell bestimmter Augenposition lagen die 3D-Abweichungen der Fixationszentren bei durchschnittlich 17 mm. Die größeren Abweichungen der beiden anderen Algorithmen von jeweils 22 mm sind ähnlich der Resultate in Studie II durch die ungenauen Eyetracking-Daten erklärbar. Dies bestätigte sich in den 3D-Fixationen aus den Blickorten der Studie III. Hier waren die Abweichungen unabhängig der Algorithmen mit ca. 17 mm annähernd gleich. Die Abweichungen der 3D-Blickorte und 3D-Fixationen im realen und virtuellen Setting lassen eine ähnliche Genauigkeit vermuten, allerdings war der Messbereich im Experiment mit stereoskopischer Visualisierung eingeschränkt. Berechnet man die Abweichungen im realen Setting für den Messbereich des virtuellen, ergeben sich geringere Abweichungen sowohl für die 3D-Blickorte als auch 3D-Fixationen (Abb. 11-1 A und B). Eine mögliche Ursache liegt in den größeren Abweichungen durch das Tragen einer 3D-Brille, wie sie die bereits in Studie I nachgewiesen wurden. Die berechneten Vergenzwinkel wiesen im virtuellen Setting geringfügig größere Werte auf als im realen Experiment (Abb. 11-1 C), beide liegen aber nahe der mathematisch korrekten Lösung.

Wie in den vorhergehenden Analysen hatte die Fokussierdistanz einen signifikanten Einfluss auf die ermittelten Fixationsparameter. In kürzerer Entfernung wurden weniger Fixationen von höherer Dauer berechnet und umgekehrt. Auch die 3D-Abweichungen der Zentren waren im Nahbereich geringer als in größerer Entfernung. Dabei hatte der örtliche Schwellwert in den getesteten Grenzen kaum einen Einfluss auf die dreidimensionale Fixationsabweichung dafür umso mehr auf die Anzahl erkannter 3D-Fixationen bzw. deren Dauern. Ein erweiterter Toleranzbereich führt zwar zu längeren Fixationsdauern, allerdings nimmt die Fixationshülle dabei einen größeren Raum ein. Der örtliche Parameter muss sorgfältig gewählt werden, da nahe beieinanderliegende Fixationen fälschlicherweise zu einer einzigen Fixation zusammengefasst werden könnten.

Blickbewegungen zwischen Fixationen ohne Unterbrechungen durch Blinks können wie in den zweidimensionalen Methoden als Sakkaden angenommen werden. Aus der zeitlichen Abfolge von 3D-Fixationen und 3D-Sakkaden ergibt sich ein dreidimensionaler Blickpfad. Dadurch eröffnet sich die Möglichkeit, bislang in zwei Dimensionen untersuchte Blickverläufe nunmehr räumlich zu bestimmen und grundlegende Zusammenhänge zwischen Blickbewegungen und kognitiven Prozessen zukünftig dreidimensional sowohl in einer realen als auch virtuellen Umgebung zu analysieren.

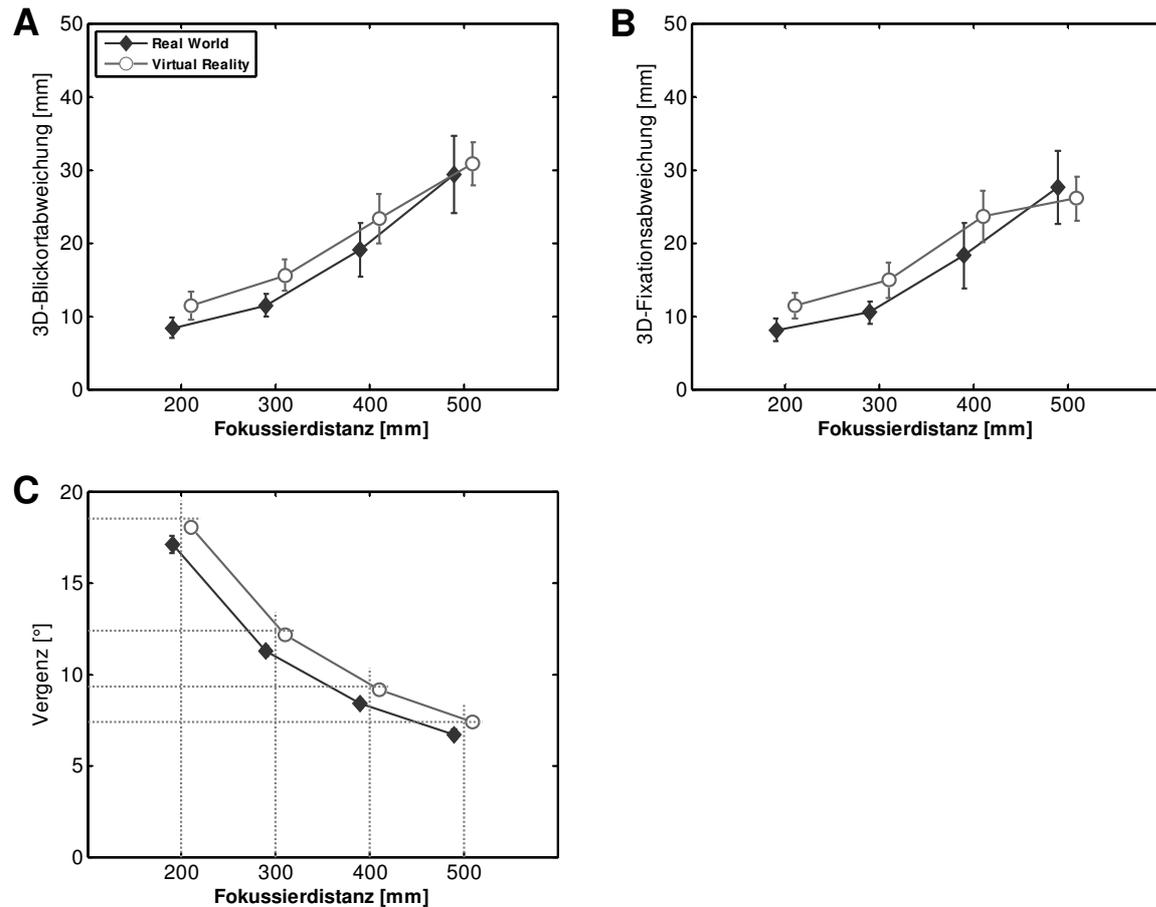


Abb. 11-1: Gegenüberstellung der Ergebnisse aus dem realen und virtuellen Experiment in gleichen Fokussierentfernungen. Abweichungen der dreidimensionalen Blickorte (A) und (B) auf Datenbasis des FixEye-Algorithmus sowie die berechneten Vergenzwinkel (C) im Vergleich zu den mathematisch korrekten Werten (gestrichelte Linien) bei einer interokularen Pupillendistanz von 65 mm. Fehlerbalken kennzeichnen das 95%-CI.

Die entwickelten Algorithmen zur Bestimmung dreidimensionaler Blickorte sowie der ellipsoide Fixationserkennungsalgorithmus funktionieren unabhängig der eingesetzten Blickbewegungsmessung oder stereoskopischen Visualisierungstechnologie. Beispielsweise wurde in einem Experiment am *Lehrstuhl für Ingenieurpsychologie und angewandte Kognitionsforschung* den Probanden ein virtueller Tennisball auf einem 3D-Monitor mit Shutter-Technologie in unterschiedlichen Tiefenebenen stereoskopisch dargeboten. Die dabei erhobenen Eyetracking-Parameter mit einem EyeLink™ 1000 System sind den Algorithmen zur 3D-Blickortberechnung sowie 3D-Fixationserkennung übergeben und daraus räumliche Blickbewegungen berechnet worden. Die Visualisierung der Ergebnisse beschreibt deskriptiv die gleiche ellipsoide Anordnung der 3D-Blickorte wie in den Messungen mit dem EyeFollower™ System (Abb. 11-2). Für genauere Aussagen über die Unabhängigkeit der Algorithmen von der Blickbewegungsmessung müssen in weiteren Studien die Blickparameter verschiedener binokularer Eyetracking-Systeme der 3D-Blickortberechnung sowie Fixationserkennung übergeben und die Ergebnisse miteinander verglichen werden.

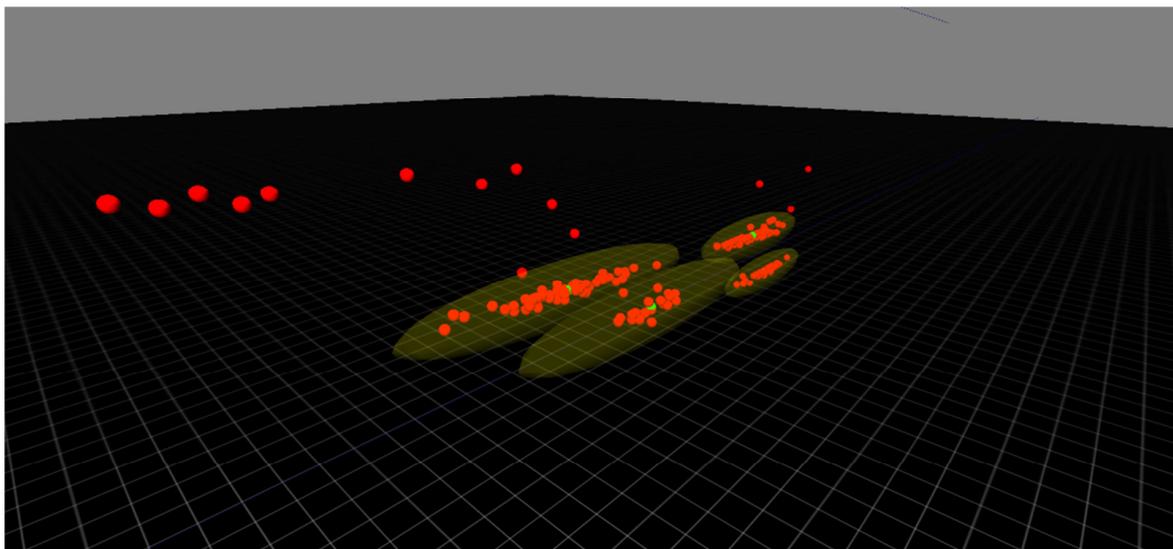


Abb. 11-2: Visualisierung der berechneten 3D-Blickorte (rot), erkannter 3D-Fixationen (gelb) sowie der Fixationszentren (grün) aus den Blickdaten des EyeLink™ 1000 Systems mit der ellipsoiden Fixationserkennung. Im abgebildeten Stadium wurden bereits vier Fixationen identifiziert. Eine weitere Fixation lässt sich anhand der Anordnung bereits berechneter Blickorte (links, oben) vermuten, allerdings ist die Mindestanzahl notwendiger Samples für die Bestätigung einer 3D-Fixation, in diesem Fall 25, noch nicht erreicht.

Die entwickelten Algorithmen beziehen Veränderungen in den Blickachsenparametern, wie beispielsweise die Positionsdaten der Augäpfel, in die Berechnung eines jeden Samples ein. Damit sind Kopfbewegungen theoretisch möglich, wurden jedoch innerhalb dieser Arbeit aufgrund des Experiment-Designs nicht zugelassen. Ob die Bewegung des Kopfes eine Auswirkung auf die Genauigkeit der bestimmten 3D-Blickparameter hat muss in weiteren Untersuchungen überprüft werden. Die Aufhebung der Restriktion einer festen Kopfposition würde Experimente mit Probanden ermöglichen, denen es schwer fällt über einen längeren Zeitraum stillzusitzen, wie beispielsweise Kinder oder Patienten mit motorischen Störungen.

Perspektiven

Mit der im Rahmen dieser Forschungsarbeit entwickelten Methode einer dreidimensionalen Blickortberechnung sowie Fixationserkennung sind okulomotorische Vorgänge, visuelle Wahrnehmungsaspekte und kognitive Prozesse räumlich analysierbar geworden. Blickmuster, wie Fixationen und Sakkaden können in unterschiedlichen Tiefenebenen bestimmt und mit Paradigmen aus dem zweidimensionalen Bereich verglichen werden.

Beispielsweise ließe sich untersuchen, inwiefern korrigierende Blickbewegungen zur Nachjustierung des Blickortes, wie sie in 2D-Experimenten bei Suchaufgaben (Mcpeck, Skavenski & Nakayama, 2000; Prablanc, Masse & Echallier, 1978) oder beim Lesen (vgl. Rayner, Juhasz & Pollatsek, 2005) nachgewiesen wurden, auch bei räumlicher Wahrnehmung vorhanden sind. Bei einem Blickortwechsel springt das Auge mittels einer Sakkade in die Nähe des Zielreizes und versucht diesen so genau wie möglich zu treffen. Gelingt dies nicht, muss

die Blickposition durch eine Korrektursakkade angepasst werden. Inwieweit dieser Mechanismus auch räumlich erfolgt, ist mit der dreidimensionalen Blickortanalyse erforschbar. Hierbei könnten Blickortwechsel zwischen Objekten in unterschiedlichen Tiefenebenen Aufschluss über Nachjustierungen bei einem Under- bzw. Overshoot geben. Ein weiterer Befund aus der Grundlagenforschung belegt, dass zweidimensionale Sakkaden weniger eine geradlinige sondern vielmehr eine bogenförmigen Bahn beschreiben (Nummenmaa, Hyönä & Calvo, 2009). Ob dieser gekrümmte Verlauf bei räumlichen Sakkaden vorhanden ist bzw. welche Bahn bei einem Blickortwechsel unterschiedlicher Tiefe beschrieben wird, ist mit der 3D-Blickerfassung analysierbar.

Nicht nur grundlagenorientierte sondern auch angewandte Forschungsaspekte werden auf die dritte Dimension erweiterbar. Während einer Szenenwahrnehmung oder natürlicher Handlungsabläufe lassen sich 3D-Blickbewegungen den Objekten der Umgebung zuordnen. Die vorhandene Tiefeninformation kann insbesondere bei sich teilweise verdeckenden Objekten in unterschiedlicher Tiefenebene eine genauere Zuweisung ermöglichen (Pfeiffer et al., 2007). Beispielsweise sind Blickortwechsel beim Autofahren zwischen Head-Up-Display und Fahrumgebung, in Einkaufssituationen zwischen verschiedenen Produkten oder während sportlicher Aktivitäten nunmehr dreidimensional analysierbar. Aber auch in virtuellen Umgebungen wird die 3D-Blickbewegungsmessung die Forschungsmöglichkeiten erweitern. In Kombination mit Stereoprojektionen lassen sich dreidimensionale Szenen kontrolliert erzeugen. Hierbei können Gefahrensituationen in Simulatoren projiziert und dabei die visuelle Aufmerksamkeitsverteilung untersucht werden. Durch eine realitätsnahe Visualisierung sind vergleichbarere Ergebnisse zur Wirklichkeit zu erwarten als mit zweidimensionalen Methoden.

In der Wirtschaft lassen sich mit der 3D-Blickanalyse Usability- und User Experience-Methoden ergänzen, um für Designstudien zu einem neuen Produktes nicht erst kostenintensive Prototypen herstellen zu müssen, sondern auf virtuelle 3D-Modelle zurückgreifen zu können (Rojas, Contero, Bartomeu & Guixeres, 2015; Yousefi, Mohammadpour, Karan & Asadi, 2015). Im medizinischen Bereich werden räumliche Blickbewegungsmessungen für Diagnosen und Therapien nutzbar. Beispielsweise könnte bisher zweidimensionales Stimulusmaterial wie Bilder oder Videos bei der Untersuchung von autistisch veranlagten Kindern (Chevallier et al., 2015; Fujioka et al., 2016; Guillon et al., 2016) auf die dritte Dimension erweitert werden und auch für die Steuerung von Mensch-Maschine-Schnittstellen, wie sie in einem „See-Through HMD“ (Benko, Ofek, Zheng & Wilson, 2015; Vogel et al., 2009) eingesetzt wird, kann das Wissen über die Blickorttiefe Aufschluss darüber geben, ob der Nutzer seine Umwelt wahrnimmt oder die semitransparenten Interaktionsflächen zur Blicksteuerung fokussiert.

Die entwickelte „Methode zur Identifikation dreidimensionaler Blickbewegungen in realer und virtueller Umgebung“ wird die Möglichkeiten psychologischer Forschungen, Anwendungen in der Wirtschaft und Medizin zukünftig erweitern. Um die Verwendung und eine Weiterentwicklung der Methode zu ermöglichen, werden die im Rahmen dieser Arbeit entwickelten Algorithmen, Softwareanwendungen sowie Dokumentationen in einem *Open Source Version Control System* der Öffentlichkeit zur Verfügung gestellt.

<https://github.com/sascha-weber/Gaze3DFix>

Literaturverzeichnis

- Barabas, J., Goldstein, R. B., Apfelbaum, H., Woods, R. L., Giorgi, R. G., & Peli, E. (2004). Tracking the line of primary gaze in a walking simulator: Modeling and calibration. *Behavior Research Methods, Instruments, & Computers*, 36(4), 757-770.
- Barakat, B., Crump, C., Cades, D., Rauschenberger, R., Schwark, J., Hildebrand, E., & Young, D. (2015). Eye tracking evaluation of driver visual behavior with a forward collision warning and mitigation system. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, pp. 1321-1325). Los Angeles, CA, USA: SAGE Publications.
- Benko, H., Ofek, E., Zheng, F., & Wilson, A. D. (2015). FoveAR: combining an optically see-through near-eye display with projector-based spatial augmented reality. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (pp. 129-135). Charlotte, NC, USA: ACM.
- Blignaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, 71(4), 881-895. doi: 10.3758/APP.71.4.881
- Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis & Prevention*, 42(4), 1240-1249. doi: 10.1016/j.aap.2010.02.001
- Braunagel, C., Stolzmann, W., Kasneci, E., & Rosenstiel, W. (2015). Driver-activity recognition in the context of conditionally autonomous driving. *Proceedings of the 18th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 1652-1657). Las Palmas, Spain: IEEE.
- Burgert, O., Örn, V., Velichkovsky, B. M., Gessat, M., Joos, M., Strauß, G., . . . Hertel, I. (2007). Evaluation of perception performance in neck dissection planning using eye tracking and attention landscapes. In Y. Jiorg & B. Sahiner (Eds.), *Medical Imaging 2007: Image Perception, Observer Performance, and Technology Assessment* (Vol. 6515, pp. 65150-65159). San Diego, CA, USA: SPIE.
- Buswell, G. T. (1935). How people look at pictures: A study of the psychology of perception in art. Chicago, IL, USA: University of Chicago Press.
- Calvo, A., Chiò, A., Castellina, E., Corno, F., Farinetti, L., Ghiglione, P., . . . Vignola, A. (2008). Eye tracking impact on quality-of-life of ALS patients. *Proceedings of the Conference on Computers Helping People with Special Needs* (Vol. 5101, pp. 70-77). Linz, Austria: Springer Berlin/Heidelberg.
- Camerini, P. M., Galbiati, G., & Maffioli, F. (1988). Algorithms for finding optimum trees: Description, use and evaluation. *Annals of Operations Research*, 13, 265-397.
- Cavanagh, P. (2004). Attention routines and the architecture of selection. In M. I. Posner (Ed.), *Cognitive Neuroscience of Attention* (pp. 14-28). New York, NY, USA: Guilford Press.
- Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. T. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research*, 8(5), 620-628. doi: 10.1002/aur.1479
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework of memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Crevits, L., Hanse, M., Tummers, P., & Van Maele, G. (2000). Antisaccades and remembered saccades in mild traumatic brain injury. *Journal of neurology*, 247(3), 179-182.

- Curie, A., Brun, A., Cheylus, A., Reboul, A., Nazir, T., Bussy, G., . . . des Portes, V. (2016). A novel analog reasoning paradigm: New insights in intellectually disabled patients. *Plos One*, 11(2). doi: 10.1371/journal.pone.0149717
- Daugherty, B. C., Duchowski, A. T., & House, D. H. (2010). Measuring vergence over stereoscopic video with a remote eye tracker. *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications (ETRA '10)* (pp. 97-100). New York, NY, USA: ACM.
- Di Stasi, L. L., Diaz-Piedra, C., Rieiro, H., Carrión, J. M. S., Berrido, M. M., Olivares, G., & Catena, A. (2016). Gaze entropy reflects surgical task load. *Surgical Endoscopy*, 1-10. doi: 10.1007/s00464-016-4851-8
- Dodgson, N. A. (2004). Variation and extrema of human interpupillary distance. *Proceedings of the Stereoscopic Displays and Virtual Reality Systems XI* (pp. 36-46). San Jose, CA, USA: SPIE.
- Dowling, J. E., & Boycott, B. B. (1966). Organization of the primate retina: Electron microscopy. *Proceedings of the Royal Society of London B: Biological Sciences* (Vol. 166, pp. 80-111).
- Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & S. Marin (Eds.), *Attention and Performance XI*. Hillsdale, NJ, USA: Lawrence Erlbaum.
- Duchowski, A. T., Medlin, E., Cournia, N., Murphy, H., Gramopadhye, A., Nair, S., . . . Melloy, B. (2002). 3-D eye movement analysis. *Behavior Research Methods*, 34(4), 573-591.
- Duchowski, A. T., Pelfrey, B., House, D. H., & Wang, R. (2011). Measuring gaze depth with an eye tracker during stereoscopic display. *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (pp. 15-22). Toulouse, France: ACM.
- Duchowski, A. T., Shivashankaraiah, V., Rawls, T., Gramopadhye, A. K., Melloy, B. J., & Kanki, B. (2000). Binocular eye tracking in virtual reality for inspection training. *Proceedings of the 2000 symposium on Eye tracking research & applications* (pp. 89-96). New York, NY, USA: ACM.
- Ebenholtz, S. M. (2001). *Oculomotor systems and perception*. Cambridge, UK: Cambridge University Press.
- Eggert, T. (2007). Eye Movement Recordings: Methods. In A. Straube & U. Büttner (Eds.), *Neuro-ophthalmology: Neuronal Control of Eye Movements* (Vol. 40, pp. 15-34). Basel, Switzerland: Karger.
- Eisenberg, M. L., & Zacks, J. M. (2016). Ambient and focal visual processing of naturalistic activity. *Journal of Vision*, 16(2), 5. doi: 10.1167/16.2.5
- Emoto, M., Niida, T., & Okano, F. (2005). Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *Journal of Display Technology*, 1(2), 328-340. doi: 10.1109/Jdt.2005.858938
- Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 11(5), 583-597.
- Essig, K., Pomplun, M., & Ritter, H. (2006). A neural network for 3D gaze recording with binocular eye trackers. *International Journal of Parallel, Emergent and Distributed Systems*, 21(2), 79-95.
- Evans, K. M., Jacobs, R. A., Tarduno, J. A., & Pelz, J. B. (2012). Collecting and analyzing eye-tracking data in outdoor environments. *Journal of Eye Movement Research*, 5(2), 6,1-19.

- Fincham, E. F., & Walton, J. (1957). The reciprocal actions of accommodation and convergence. *Journal of Physiology*, 137(3), 488-508.
- Franke, I. S., Pannasch, S., Helmert, J. R., Rieger, R., Groh, R., & Velichkovsky, B. M. (2008). Towards attention-centered interfaces: An aesthetic evaluation of perspective with eye tracking. *Acm Transactions on Multimedia Computing Communications and Applications*, 4(3), 1-13. doi: 10.1145/1386109.1386111
- Fujioka, T., Inohara, K., Okamoto, Y., Masuya, Y., Ishitobi, M., Saito, D. N., . . . Kosaka, H. (2016). Gazefinder as a clinical supplementary tool for discriminating between autism spectrum disorder and typical development in male adolescents and adults. *Mol Autism*, 7, 19. doi: 10.1186/s13229-016-0083-y
- Giuliani, F., & Marchetti, B. C. (2015). Is storytelling therapy useful for children with autism spectrum disorders and severe mental retardation? *Advanced Techniques in Biology & Medicine*, 04(01). doi: 10.4172/2379-1764.1000166
- Gold, C., Dambock, D., Lorenz, L., & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), 1938-1942. doi: 10.1177/1541931213571433
- Goldberg, J. H., & Schryver, J. C. (1995). Eye-gaze-contingent control of the computer interface: Methodology and example for zoom detection. *Behavior Research Methods, Instruments, & Computers*, 27(3), 338-350. doi: 10.3758/Bf03200428
- Goldstein, E. B., Ritter, M., & Herbst, G. (2002). *Wahrnehmungspsychologie* (Vol. 2). Heidelberg: Spektrum Akademischer Verlag.
- Guillon, Q., Rogé, B., Afzali, M. H., Baduel, S., Kruck, J., & Hadjikhani, N. (2016). Intact perception but abnormal orientation towards face-like objects in young children with ASD. *Scientific reports*, 6, 22119. doi: 10.1038/srep22119
- Hayhoe, M. M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, 9(4), 188-194. doi: 10.1016/j.tics.2005.02.009
- Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Experimental Brain Research*, 217(1), 125-136. doi: 10.1007/s00221-011-2979-2
- He, W., & Zheng, B. (2016). Collaborative performance in laparoscopic teams: behavioral evidences from simulation. *Surgical Endoscopy*, 1-6. doi: 10.1007/s00464-016-4794-0
- Heitger, M. H., Jones, R. D., Macleod, A., Snell, D. L., Frampton, C. M., & Anderson, T. J. (2009). Impaired eye movements in post-concussion syndrome indicate suboptimal brain function beyond the influence of depression, malingering or intellectual ability. *Brain*, 132(10), 2850-2870. doi: 10.1093/brain/awp181
- Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science*, 16(4), 219-222.
- Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243-271. doi: 10.1146/annurev.psych.50.1.243
- Hodges, L. F., & Davis, E. T. (1993). Geometric considerations for stereoscopic virtual environments. *Presence: Teleoperators and Virtual Environments*, 2, 34-43.
- Holladay, J. T. (2007). *Quality of vision: Essential optics for the cataract and refractive surgeon*. Thorofare, NJ, USA: Slack Incorporated.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, UK: Oxford

- University Press.
- Huckauf, A., Watrin, L., Yuras, G., & Koepsel, A. (2013). Brightness and contrast effects on binocular coordination. *Proceedings of the 55th Tagung experimentell arbeitender Psychologen*. Vienna, Austria: PABST.
- Interactive Minds. (2011). NYAN 2.0XT-Eye tracking data analysis suite [computer program] (Version 2.5.6). Dresden, Germany: Interactive Minds Dresden GmbH.
- Intriligator, J., & Cavanagh, P. (2001). The Spatial Resolution of Visual Attention. *Cognitive Psychology*, 43(3), 171-216. doi: 10.1006/cogp.2001.0755
- Johnson, J. S., Liu, L., Thomas, G., & Spencer, J. P. (2007). Calibration algorithm for eyetracking with unrestricted head movement. *Behavior Research Methods*, 39(1), 123-132.
- Jones, G. R., Lee, D., Holliman, N. S., & Ezra, D. (2001). Controlling perceived depth in stereoscopic images. *Proceedings of Photonics West 2001 - Electronic Imaging* (pp. 42-53). San Jose, USA: SPIE.
- Joos, M., Rötting, M., & Velichkovsky, B. M. (2003). Die Bewegungen des menschlichen Auges: Fakten, Methoden, innovative Anwendungen. In T. Herrmann, S. Deutsch & G. Rickheit (Eds.), *Handbuch der Psycholinguistik* (pp. 142-168). Berlin, Germany: De Greyter.
- Kabugo, D., Muyinda, P. B., Masagazi, F. M., Mugagga, A. M., & Mulumba, M. B. (2016). Tracking students' eye-movements when reading learning objects on mobile phones: A discourse analysis of luganda language teacher-trainees' reflective observations. *Journal of Learning for Development*, 3(1).
- Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of neural science* (Vol. 4). New York, USA: McGraw-Hill.
- Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics*, 6(4), 193-198.
- Kim, D., & Sohn, K. (2010). Depth adjustment for stereoscopic image using visual fatigue prediction and depth-based view synthesis. *Proceedings of the 2010 IEEE International Conference on Multimedia and Expo* (pp. 956-961). Singapore: IEEE.
- Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., & Gowda, S. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering*, 57(11), 2635-2645. doi: 10.1109/TBME.2010.2057429
- Kotulak, J. C., & Schor, C. M. (1986). The dissociability of accommodation from vergence in the dark. *Invest Ophthalmol Vis Sci*, 27(4), 544-551.
- Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, 51(13), 1457-1483. doi: 10.1016/j.visres.2010.12.014
- Kumar, D., Dutta, A., Das, A., & Lahiri, U. (2016). SmartEye: Developing a novel eye tracking system for quantitative assessment of oculomotor abnormalities. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*. doi: 10.1109/TNSRE.2016.2518222
- Kwon, Y. M., Jeon, K. W., Ki, J., Shahab, Q. M., Jo, S., & Kim, S. K. (2006). 3D gaze estimation and interaction to stereo display. *International Journal of Virtual Reality*, 5(3), 41-45.
- Lambooj, M., Fortuin, M., Heynderickx, I., & Ijsselstein, W. A. (2009). Visual discomfort and visual fatigue of stereoscopic displays: a review. *Journal of Imaging Science and Technology*,

- 53(3), 30201-30201. doi: 10.2352/J.ImagingSci.Technol.2009.53.3.030201
- Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, 25(3), 296-324. doi: 10.1016/j.preteyeres.2006.01.002
- Land, M. F. (2009). *Looking and acting: vision and eye movements in natural behaviour*. Oxford, UK: Oxford University Press.
- LC Technologies. (2002). *Eyegaze Edge Analysis System: User's manual*. Fairfax, USA: LC Technologies, Inc.
- LC Technologies. (2014). *Eyegaze Edge Analysis System: Programmer's manual*. Fairfax, USA: LC Technologies, Inc.
- Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, 28(4), 612-625. doi: 10.1093/Hcr/28.4.612
- Li, Q., Huang, Z., & Christianson, K. (2016). Visual attention toward tourism photographs with text: An eye-tracking study. *Tourism Management*, 54, 243-258. doi: 10.1016/j.tourman.2015.11.017
- Liu, P. (2014). Using eye tracking to understand learners' reading process through the concept-mapping learning strategy. *Computers & Education*, 78, 237-249. doi: 10.1016/j.compedu.2014.05.011
- Liversedge, S. P., Rayner, K., White, S. J., Findlay, J. M., & McSorley, E. (2006). Binocular coordination of the eyes during reading. *Current Biology*, 16(17), 1726-1729. doi: 10.1016/j.cub.2006.07.051
- Maggia, C., Guyader, N., & Guérin-Dugué, A. (2013). Using natural versus artificial stimuli to perform calibration for 3D gaze tracking. In B. E. Rogowitz, T. N. Pappas & H. d. Ridder (Eds.), *Human Vision and Electronic Imaging XVIII* (pp. 865113-865114): SPIE.
- Mansouryar, M., Steil, J., Sugano, Y., & Bulling, A. (2016). 3D gaze estimation from 2D pupil positions on monocular head-mounted eye trackers. *arXiv preprint arXiv:1601.02644*.
- Masaoka, K., Hanazato, A., Emoto, M., Yamanoue, H., Nojiri, Y., & Okano, F. (2006). Spatial distortion prediction system for stereoscopic images. *Journal of Electronic Imaging*, 15(1), 013002-013002. doi: 10.1117/1.2181178
- Mason, K. A. (1969). U.S. Patent Patent No. 3,462,604.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17(6), 578-586.
- McLaughlin, S. C. (1967). Parametric adjustment in saccadic eye movements. *Perception & Psychophysics*, 2(8), 359-362.
- Mcpeek, R. M., Skavenski, A. A., & Nakayama, K. (2000). Concurrent processing of saccades in visual search. *Vision Research*, 40(18), 2499-2516.
- Medlin, E. (2003). *Ocular vergence measurement in bi[n]ocular virtual reality*. (Master Theses), Graduate School of Clemson University, Clemson, SC, USA.
- Morimoto, C. H., Koons, D., Amir, A., & Flickner, M. (2000). Pupil detection and tracking using multiple light sources. *Image and Vision Computing*, 18(4), 331-335. doi: 10.1016/S0262-8856(99)00053-0
- Nascimento, J. C., & Marques, J. S. (2006). Performance evaluation of object detection algorithms for video surveillance. *Ieee Transactions on Multimedia*, 8(4), 761-774. doi: 10.1109/Tmm.2006.876287

- Nguyen, T. H., Qui, T. C., Xu, K., Cheok, A. D., Teo, S. L., Zhou, Z., . . . Kato, H. (2005). Real-time 3D human capture system for mixed-reality art and entertainment. *IEEE Trans Vis Comput Graph*, *11*(6), 706-721.
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2009). Emotional scene content drives the saccade generation system reflexively. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 305. doi: 10.1037/a0013626
- Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision*, *9*(5), 1-28. doi: 10.1167/9.5.31
- Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, *42*(1), 188-204. doi: 10.3758/BRM.42.1.188
- Otterbring, T., Wästlund, E., & Gustafsson, A. (2016). Eye-tracking customers' visual attention in the wild: Dynamic gaze behavior moderates the effect of store familiarity on navigational fluency. *Journal of Retailing and Consumer Services*, *28*, 165-170. doi: 10.1016/j.jretconser.2015.09.004
- Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA, USA: MIT Press.
- Parkin, A. J. (2014). *Essential cognitive psychology (classic edition)*. USA/CAN: Taylor and Francis.
- Pfeiffer, T. (2012). Measuring and visualizing attention in space with 3D attention volumes. *Proceedings of the Symposium on Eye Tracking Research & Applications* (pp. 29-36). Santa Barbara, CA, USA: ACM.
- Pfeiffer, T., Donner, M., Latoschik, M. E., & Wachsmuth, I. (2007). 3D fixations in real and virtual scenarios. *Journal of Virtual Reality and Broadcasting*, *5*(16).
- Pfeiffer, T., Renner, P., & Pfeiffer-Leßmann, N. (2016). EyeSee3D 2.0: Model-based real-time analysis of mobile eye-tracking in static and dynamic three-dimensional scenes. *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (pp. 189-196). Charleston, SC, USA: ACM.
- Pietsch, W. (1957). *Die Praxis der Stereo-Nahaufnahmen*. Halle (Saale): VEB Knapp.
- Pomplun, M., Velichkovsky, B. M., & Ritter, H. (1994). An artificial neural network for high precision eye movement tracking. In B. Nebel & L. Dreschler-Fischer (Eds.), *KI-94: Advances in Artificial Intelligence* (Vol. 861, pp. 63-69): Springer Berlin/Heidelberg.
- Posner, M. I. (1977). *Review of Human information processing: An introduction to psychology. 2nd ed* (Vol. 22). USA: American Psychological Association.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160-174.
- Prablanc, C., Masse, D., & Echallier, J. F. (1978). Error-correcting mechanisms in large saccades. *Vision Research*, *18*(5), 557-560.
- Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology*, *130*, 92-105. doi: 10.1016/j.jecp.2014.09.012
- Rayner, K., Juhasz, B. J., & Pollatsek, A. (2005). Eye movements during reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 79-97). Oxford, UK: Blackwell.
- Rayner, K., & Reingold, E. M. (2015). Evidence for direct cognitive control of fixation durations during reading. *Current Opinion in Behavioral Sciences*, *1*, 107-112. doi:

- 10.1016/j.cobeha.2014.10.008
- Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, 22(5), 1188.
- Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, 27(6), 537-544.
- Renner, R. S., Velichkovsky, B. M., Helmert, J. R., & Stelzer, R. H. (2013). Measuring interpupillary distance might not be enough. *Proceedings of the ACM Symposium on Applied Perception* (pp. 130-130). Dublin, Ireland: ACM.
- Ricard, S., & Garrison, K. (2015). Usability of video-based website designs. *CRIUS*, 3(1).
- Rojas, J. C., Contero, M., Bartomeu, N., & Guixeres, J. (2015). Using combined bipolar semantic scales and eye-tracking metrics to compare consumer perception of real and virtual bottles. *Packaging Technology and Science*, 28(12), 1047-1056. doi: 10.1002/pts.2178
- Rötting, M., Göbel, M., & Springer, J. (1999). Automatic object identification and analysis of eye movement recordings. *MMI-Interaktiv*, 1(2).
- Salvucci, D. D., & Anderson, J. R. (1998). Tracing eye movement protocols with cognitive process models. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 923-928). Hillsdale, NJ, USA: Lawrence Erlbaum.
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Symposium on Eye Tracking Research & Applications* (pp. 71-78). Palm Beach Gardens, FL, USA: ACM.
- Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3), 347-357. doi: 10.1518/001872007x196685
- Schulz, C. M., Schneider, E., Kohlbecher, S., Hapfelmeier, A., Heuser, F., Wagner, K. J., . . . Schneider, G. (2014). The influence of anaesthetists' experience on workload, performance and visual attention during simulated critical incidents. *Journal of clinical monitoring and computing*, 28(5), 475-480. doi: 10.1007/s10877-013-9443-8
- Sen, T., & Megaw, T. (1984). The effects of task variables and prolonged performance on saccadic eye movement parameters. *Advances in Psychology*, 22, 103-111.
- Sharma, C., Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2016). Eye gaze movement studies of control room operators: A novel approach to improve process safety. *Computers & Chemical Engineering*, 85, 43-57. doi: 10.1016/j.compchemeng.2015.09.012
- Smeets, J. B. J., & Hooge, I. T. (2003). Nature of variability in saccades. *Journal of Neurophysiology*, 90(1), 12-20. doi: 10.1152/jn.01075.2002
- Smith, M. D., & Collar, B. T. (2012). Perception of size and shape in stereoscopic 3D imagery. *Proceedings of Stereoscopic Displays and Applications XXIII* (Vol. 8288, pp. 82881O-82831): SPIE.
- SR Research. (2009). *EyeLink® 1000: User manual*. Mississauga, Ontario, Canada: SR Research Ltd.
- Tatler, B. W. (2001). Characterising the visual buffer: real-world evidence for overwriting early in each fixation. *Perception*, 30(8), 993-1006.
- Tauer, H. (2010). *Stereo 3D*. Berlin: Fachverlag Schiele & Schön.

- Trottier, L., & Pratt, J. (2005). Visual processing of targets can reduce saccadic latencies. *Vision Research*, 45(11), 1349-1354. doi: 10.1016/j.visres.2004.12.007
- Velichkovsky, B. M., Cornelissen, F., Geusebroek, J.-M., Graupner, S.-T., Hari, R., Marsman, J. B., . . . Pannasch, S. (2013). Measurement-related issues in the investigation of active vision. In B. Berglund, G. B. Rossi, M. C. Townsend & L. Pendrill (Eds.), *Measurement With Persons: Theory, Methods, and Implementation Areas* (pp. 281-300). London-New York: Taylor and Francis.
- Velichkovsky, B. M., Dornhoefer, S. M., Pannasch, S., & Unema, P. J. A. (2001). Visual fixations and level of attentional processing. In A. Duhowski (Ed.), *Proceedings of the Symposium on Eye Tracking Research & Applications* (pp. 79-85). Palm Beach Gardens, NY, USA: ACM Press.
- Vicente, F., Huang, Z., Xiong, X. F., De la Torre, F., Zhang, W., & Levi, D. (2015). Driver gaze tracking and eyes off the road detection system. *Ieee Transactions on Intelligent Transportation Systems*, 16(4), 2014-2027. doi: 10.1109/Itits.2015.2396031
- Vogel, U., Kreye, D., Richter, B., Bunk, G., Reckziegel, S., Herold, R., . . . Velichkovsky, B. M. (2009). Bi-directional OLED microdisplay for interactive see-through HMDs: Study toward integration of eye-tracking and informational facilities. *Journal of the Society for Information Display*, 17(3), 175-184. doi: 10.1889/Jsid17.3.175
- Wade, N., & Swanston, M. (2013). *Visual perception: An introduction, 3rd edition*. USA/CAN: Taylor and Francis.
- Weber, S., Joos, M., & Velichkovsky, B. M. (2004). Aufzeichnung und Analyse von Augenbewegungen beim Betrachten von Internetseiten. *Proceedings of the 46. Tagung experimentell arbeitender Psychologen*. Gießen, Germany: PABST.
- Winter, J., Fotios, S., & Voelker, S. (2016). Gaze direction when driving after dark on main and residential roads: Where is the dominant location? *Lighting Research and Technology*. doi:10.1177/1477153516632867
- Woods, A. J., Docherty, T., & Koch, R. (1993). Image distortions in stereoscopic video systems. *Proceedings of Symposium on Electronic Imaging: Science and Technology* (pp. 36-48): SPIE.
- Yano, S., Ide, S., Mitsuhashi, T., & Thwaites, H. (2002). A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. *Displays*, 23(4), 191-201. doi: 10.1016/S0141-9382(02)00038-0
- Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, 7(5), 397-429. doi: Doi 10.3758/Bf03201553
- Yousefi, M. V., Mohammadpour, A., Karan, E. P., & Asadi, S. (2015). Implementing eye tracking technology in the construction process. *Proceedings of 51st ASC Annual International Conference*. College Station, TX, USA: Texas A&M University.
- Zavesky, M. (2012). *Wahrnehmungsrealistische Projektion anthropomorpher Formen*. (Doctoral dissertation), Technische Universität Dresden, Germany.
- Zhang, Y., & Hornof, A. J. (2011). Mode-of-disparities error correction of eye-tracking data. *Behavior Research Methods*, 43(3), 834-842. doi: 10.3758/s13428-011-0073-0

Abbildungsverzeichnis

- Abb. 3-1: Akkommodation-Vergenz-Modell nach Lambooi et al. (2009): Akkommodation und Vergenz modelliert als zwei parallele Feedback-Steuersysteme mit adaptiver Interaktion.....11
- Abb. 3-2: Der Horopter (A). Die Objektpunkte A-G stellen die korrespondierenden Netzhautpunkte in beiden Augen dar. Der theoretische Horopter ist kreisförmig und auch als Vieth-Müller-Kreis bekannt. Der empirische Horopter (gestrichelte Linie) liegt hinter dem theoretischen Horopter (nach Palmer, 1999). Bei Objekten, die nicht auf dem Horopter liegen (B), fallen deren Abbildungen auf nichtkorrespondierenden Netzhautpunkte (nach Posner, 1977).....13
- Abb. 3-3: Verfahren zur stereoskopischen Visualisierung: Schematische Darstellung der Wirkungsweise eines Polarisationsfilters (A) und zeitliche Synchronisation der alternierend sichtbaren Bilder für linkes und rechtes Auge im Shutterverfahren (B).15
- Abb. 3-4: Bildschirmparallaxe (nach Hodges & Davis, 1993): Der Fixationspunkt f_1 befindet sich (A) hinter der Projektionsfläche PF bei ungekreuzter Parallaxe und (B) vor der Projektionsfläche bei gekreuzter Parallaxe.16
- Abb. 3-5: Einfluss des Augenabstands auf die wahrgenommene Entfernung eines Objektes in einer stereoskopisch dargebotenen virtuellen Szene. Das gleiche Objekt erscheint mit größerem Augenabstand entfernter als bei einem kleineren Augenabstand.17
- Abb. 3-6: Akkommodation und Vergenz in realer und stereoskopisch projizierter VR-Umgebung. Reale Umgebung (A): Akkommodation und Vergenz sind auf dieselbe Tiefe ausgerichtet. Virtuelle Umgebung (B): Akkommodation und Vergenz sind in unterschiedliche Tiefenebenen ausgerichtet. Hierbei kommt es zu einem Vergenz-Akkommodations-Konflikt.18
- Abb. 3-7: Projektion von Objekten auf eine Bildebene (aus Zavesky, 2012). Das Bild wird nach Abbildungsprinzip der computergrafischen Kamera generiert, wobei alle Objekte gleich behandelt werden. Dabei entstehen projektionsbedingte Veränderungen der Proportionen. Diese wirken sich besonders auf am Rand positionierte Objekte aus und werden durch Verringerung des Abstandes zwischen Modell und Bildebene verstärkt..19
- Abb. 4-1: Schematische Darstellung (Draufsicht) des rechten Auges mit charakteristischen Merkmalen zur Messung der Blickrichtung (nach LC Technologies, 2002).22
- Abb. 4-2: Videokamerabilder nach dem dark-pupil Effekt (A) und dem bright-pupil Effekt (B). Unterhalb der Pupille ist der Cornea-Reflex als punktförmige Reflexion der Infrarotlichtquelle an der Cornea-Oberfläche zu erkennen.....23
- Abb. 4-3: Parameter der PCCR-Methode mit dem Glint-Pupil Vektor zur Berechnung der Blickrichtung.....24
- Abb. 4-4: Stereogeometrisches Modell zur dreidimensionalen Blickortbestimmung in einem HMD nach Duchowski et al. (2000). Die Blickachsen sind durch die Positionen der Augen bzw. den gemessenen Blickorten auf den Displays definiert und schneiden sich in derselben horizontalen Ebene. Der Blickachsenschnittpunkt gibt den dreidimensionalen Blickort an.....30
- Abb. 6-1: Technologien zur stereoskopischen Bildtrennung. (A) Polfilterbrille (vgl. Bild: <http://www.my-bolles.com>) und (B) NVIDIA 3D VISION® Shutter Glasses (Bild: <http://www.nvidia.com>) mit Sender und Empfänger zur Synchronisation der Blenden für linkes und rechtes Auge in Abhängigkeit des dargebotenen Bildschirminhaltes.44

- Abb. 6-2: Fixationsobjekt in Form eines Kreistrings (A) und die mögliche Positionierung auf dem Bildschirm (B). 46
- Abb. 6-3: Videokamerabilder des Eyegaze Analysis Systems (A) und des EyeLink™ 1000 Systems (B) in den verschiedenen Bedingungen (von oben nach unten): ohne 3D-Brille, mit Polarisationsfilterbrille, mit Shutterbrille..... 47
- Abb. 6-4: Gaze Time Plot (NYAN 2.0^{XI}): Darstellung der Blickortkoordinaten (x = rot, y = grün) in Abhängigkeit der Zeit. Der Stimuluswechsel erfolgte zum Zeitpunkt $t = 0$ und die erste Sakkade zum neuen Blickort nach 0,22 s. Anschließend fand eine Korrektursakkade nach 0,43 s statt. Für die Auswertung bedeutsam ist die zweite Fixation beginnend ab 0,45 s. 48
- Abb. 6-5: Zweidimensionale Abweichung der Blickorte- (A) und Fixationen (B) in den verschiedenen Visualisierungsbedingungen. Die gemessenen Pupillenwerte (Arbitrary Units) in den Videokamerabildern vergrößerten sich bei Einsatz stereoskopischer Visualisierungstechnologien im Vergleich zur Baseline sowohl in der EAS-Messung (C) als auch in der ELS-Messung (D). Fehlerbalken kennzeichnen das 95%-CI. 52
- Abb. 8-1: Versuchsaufbau mit Eyetracking-System, Kalibrierbildschirm, Proband und Fixationsobjekten in den fünf Fokussierebenen. 67
- Abb. 8-2: Experiment-Aufbau: Eine Acrylscheibe als Grundplatte für die Fixationsobjektaufhängungen ist zwischen Monitor und Augen der Versuchsperson oberhalb des Messbereiches horizontal angeordnet (A). Jede Fixationsobjektaufhängung besteht aus einer Gewindestange mit in drei unterschiedlichen Höhen befestigte Einhängeösen (B). Die Fixationsobjekte sind unterhalb der Acrylglasplatte an je 3 Fäden aufgehängt (C), um ein Pendeln während der Messung zu verhindern. Die 3 Fäden werden durch die Platte hindurch zur jeweiligen Aufhängevorrichtung geführt und durch eine Führungsöse (D) gebündelt. Die Fäden halten am oberen Ende durch eine Kunststoffperle zusammen, welche in die Ösen der Aufhängevorrichtung nach Bedarf eingehängt werden kann (E). Ein an die Kunststoffperle angebrachtes Holzstäbchen erleichtert die Bedienung während der Positionswechsel. Durch die unterschiedlichen Höhen der Einhängeösen wird das am unteren Ende hängende Fixationsobjekt in die exakte Position gebracht. Für eine gleichmäßige Helligkeit und einen definierten Kontrast ist im Hintergrund der Kalibrierbildschirm mit einem Grauton ausgeleuchtet (F). Mit den 15 Fixationsobjektaufhängungen und jeweils drei Einhängeösen sind 45 Messpositionen gegeben. 68
- Abb. 8-3: Grafische Benutzeroberfläche der Software „3D-Gaze Experiment 2.0“ zur Berechnung und Aufzeichnung dreidimensionaler Blickorte. 69
- Abb. 8-4: Kamerabilder des Eyetracking-Systems: (A) Eindeutig identifizierbarer Cornea-Reflexpunkt. (B) Wird der Winkel zwischen dem Blickvektor und optischer Achse der Kamera zu groß bzw. liegt der Blickort oberhalb des Messbereiches, verschiebt sich der Cornea-Reflexpunkt auf die Sklera und es kommt zu Mehrfachreflexionen. Eine eindeutige Lokalisation des Cornea-Reflexpunktes ist hierbei nicht möglich. 71
- Abb. 8-5: 3D-Blickortabweichung in den unterschiedlichen Algorithmen (A) und Tiefenebenen (B). Genauigkeit des Tiefenparameters in Bezug auf Algorithmus (C) und Tiefenebene (D), Vergenzwinkel (E) sowie die manuell vom Versuchsleiter und automatisch vom Eyetracker ermittelten Augenabstände (F). Fehlerbalken kennzeichnen das 95%-CI. 74
- Abb. 9-1: Rechtes Augenbild eines virtuellen Fixationsobjektes im Zentrum der 4. Tiefenebene mit dunkler Kugel auf hellem Hintergrund (A) sowie heller Kugel auf dunklem Hintergrund (B). 80

- Abb. 9-2: Versuchsaufbau mit Eyetracking-System, 3D-Bildschirm, virtuelle Fixationsobjekte in den 4 Tiefenebenen, Proband mit Shutterbrille und Kinnstütze.....82
- Abb. 9-3: Dreidimensionale Blickortabweichung in den unterschiedlichen Algorithmen (A) sowie Abweichungen der Blickorte in Bezug auf Tiefenebene (B). Genauigkeit des Tiefenparameters in den unterschiedlichen Algorithmen (C) bzw. Tiefenebenen (D), Vergenzwinkel (E) sowie die manuell vom Versuchsleiter bzw. automatisch vom Eyetracker ermittelten Augenabstände (F). Fehlerbalken kennzeichnen das 95%-CI.....87
- Abb. 10-1: Dreidimensionale Fixationsbestimmung durch Aggregation räumlicher Blickorte (rot) innerhalb eines kugelförmigen Toleranzbereichs (blau). Bei großem Schwellwert wird der Fixation zu viel Raum zugeordnet und damit die räumliche Auflösung erkennbarer Fixationen verringert (A). Wird dieser zu klein gewählt, bleiben der Fixation zugehörige Samples unberücksichtigt (B).....97
- Abb. 10-2: Schematische Darstellung der Messgenauigkeit eines binokularen Eyetrackers mit der Messgenauigkeit α (blau) für die beiden Blickvektoren und dem sich daraus ergebenden elliptischen Schwankungsbereich (rot) für die Bestimmung des Blickortes (A). Die Ellipsengröße variiert in Abhängigkeit der Entfernung zwischen Betrachter und Fixationsort (B).....98
- Abb. 10-3: Ellipsoider Toleranzbereich der Fixationserkennung.....99
- Abb. 10-4: Dreidimensionale Fixationsbestimmung durch Aggregation räumlicher Blickorte (rot) innerhalb eines ellipsoiden Toleranzbereiches (gelb).....101
- Abb. 10-5: Visuelle Darstellung der berechneten 3D-Fixationen durch Aggregation räumlicher Blickorte (rot) innerhalb eines ellipsoiden Toleranzbereiches (gelb) mit dem Fixationszentrum (grün).....102
- Abb. 10-6: Berechnete Fixationsparameter aus Studie II und III bei einer min. Fixationsdauer von 100 ms in unterschiedlichen Schwellwertbereichen: Anzahl erkannter Fixationen (A) und (B), Abhängigkeit der Fixationsdauer (C) und (D) sowie 3D-Abweichung der Fixationszentren (E) und (F). Berechnete Fixationsparameter in Abhängigkeit der Fokussierebene in Bezug auf Fixationsdauer (G) und 3D-Genauigkeit (H). Fehlerbalken kennzeichnen das 95%-CI.106
- Abb. 11-1: Gegenüberstellung der Ergebnisse aus dem realen und virtuellen Experiment in gleichen Fokussierentfernungen. Abweichungen der dreidimensionalen Blickorte (A) und (B) auf Datenbasis des FixEye-Algorithmus sowie die berechneten Vergenzwinkel (C) im Vergleich zu den mathematisch korrekten Werten (gestrichelte Linien) bei einer interokularen Pupillendistanz von 65 mm. Fehlerbalken kennzeichnen das 95%-CI. ..116
- Abb. 11-2: Visualisierung der berechneten 3D-Blickorte (rot), erkannter 3D-Fixationen (gelb) sowie der Fixationszentren (grün) aus den Blickdaten des EyeLink™ 1000 Systems mit der ellipsoiden Fixationserkennung. Im abgebildeten Stadium wurden bereits vier Fixationen identifiziert. Eine weitere Fixation lässt sich anhand der Anordnung bereits berechneter Blickorte (links, oben) vermuten, allerdings ist die Mindestanzahl notwendiger Samples für die Bestätigung einer 3D-Fixation, in diesem Fall 25, noch nicht erreicht.117

Tabellenverzeichnis

Tab. 6-1: Technische Details der Remote Eyetracker für die Blickbewegungsmessung.....	45
Tab. 6-2: Gültige Samples mit validem x- bzw. y-Wert.....	49
Tab. 7-1: Ergebnisse des Performancetests zur Überprüfung der implementierten Funktionen	63
Tab. 8-1: Berechnungsgrundlage für die drei unterschiedlichen Blickortbestimmungen.....	70
Tab. 9-1: RGB-Werte des Stimulusmaterials.	79
Tab. 10.1: Taxonomie der Fixationserkennungsalgorithmen (nach Salvucci & Goldberg, 2000).	90

Quellcodeverzeichnis

Listing 7-1: Funktionsaufruf zur Berechnung der 3D-Blickortposition aus gegebenen Augenpositionen und binokularen Blickortkoordinaten. Die Rückgabewerte beinhalten den dreidimensionalen Blickort.....	60
Listing 7-2: Funktionsaufruf zur Berechnung der 3D Blickortposition aus gegebenen Blickvektoren. Es wird der dreidimensionale Blickort zurückgegeben.....	61
Listing 7-3: Funktionsaufruf zur Berechnung des Winkels zwischen den Blickvektoren für linkes und rechtes Auge.....	62
Listing 10-1: Funktionsaufrufe zur Initialisierung der ellipsoiden Fixationserkennung sowie zur Bestimmung dreidimensionaler Fixationen. Die Rückgabewerte beinhalten die Parameter der aktuellen 3D-Fixation sowie der vorangegangenen Sakkade.....	102

Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die vorliegende Arbeit wurde am Institut für Arbeits-, Organisations- und Sozialpsychologie, an der Professur für Ingenieurpsychologie und angewandte Kognitionsforschung der Technischen Universität Dresden unter der wissenschaftlichen Betreuung von Sen.-Prof. Dr. phil. habil. Boris M. Velichkovsky angefertigt.

Die Promotionsordnung der Fakultät Mathematik und Naturwissenschaften vom 23. Februar 2011 erkenne ich an.

Dresden, den

Sascha Weber

