

Otto-von-Guericke-Universität Magdeburg

Fakultät für Informatik



Diplomarbeit

Bestimmung fokussierter Objekte in medizinischen,
dreidimensionalen Visualisierungen mittels
2.5D-Eye-Tracking

Stefan Herold

Institut für Simulation und Graphik

Bestimmung fokussierter Objekte in medizinischen,
dreidimensionalen Visualisierungen mittels
2.5D-Eye-Tracking

Diplomarbeit

an der
Fakultät für Informatik
der Otto-von-Guericke-Universität Magdeburg

Verfasser:	STEFAN HEROLD
Geburtstag:	29. April 1984
Geburtsort:	Bad Frankenhausen
Studiengang:	Computervisualistik
Matrikelnummer:	16-81-36
1. Gutachter:	Prof. Dr.-Ing. BERNHARD PREIM,
2. Gutachter:	Jun.-Prof. Dr.-Ing. RAIMUND DACHSELT
Betreuer:	Prof. Dr.-Ing. BERNHARD PREIM, Dipl.-Ing. CHRISTIAN TIETJEN, Dipl.-Ing. MATHIAS NEUGEBAUER
Bearbeitungszeit:	23. Juni 2008 – 23. Dezember 2008

Selbstständigkeitserklärung

Hiermit versichere ich, Stefan Herold (Matrikelnummer 16-81-36), die vorliegende Arbeit allein und nur unter Verwendung der angegebenen Quellen angefertigt zu haben.

Für meine lieben Eltern und Natalia

Danksagung

In aller erster Linie möchte ich mich bei meinen Eltern bedanken, die es mir ermöglicht haben, das Studium der Computervisualistik aufzunehmen und erfolgreich abzuschließen. In den schwierigen Zeiten dieser fünf Jahre und vor allem im Zeitraum der Diplomarbeit haben sie mich immer wieder aufgemuntert und mir neue Motivation gegeben. Auch meine beiden Brüder unterstützten mich grundlegend, indem sie mir ihre vielfältigen Erfahrungen aus Forschung und Wirtschaft näher brachten. Dies befähigte mich nicht selten dazu die richtigen Argumentationen für mein Vorgehen während der Diplomarbeit zu finden.

Weiterhin bedanke ich mich bei all meinen Freunden und speziell bei Sarmad AL-Saiegh. Sie alle hatten immer ein offenes Ohr und waren zu jeder Zeit für eine entspannende Gesprächs- oder Kaffeerunde zwischen der Arbeit bereit. Ganz besonderer Dank gilt meiner lieben Freundin Natalia. Dich nach der Diplomarbeit endlich wieder in die Arme schließen zu können, unsere regelmäßigen Telefonate sowie Dein Vertrauen in mich gab mir den größten Ansporn dieses Projekt so zeitnah wie möglich abzuschließen.

Großer Dank gilt außerdem meinen Betreuern Prof. Dr.-Ing. Bernhard Preim, Dipl.-Ing. Christian Tietjen und Dipl.-Ing. Matthias Neugebauer für ihre konstruktiven Kritiken und Hilfestellungen. Weiterhin hatte Dipl.-Inf. Konrad Mühler stets Zeit, mir bei meinen Schwierigkeiten mit MEVISLAB unter die Arme zu greifen.

Außerdem hätte diese Arbeit ohne meine fleißigen Korrektoren Ines Gehrman, Monique Heimann und Claudia Schmidt niemals diesen orthographisch und grammatikalisch hohen Stand erreicht. Euch allen gilt deswegen ein ganz besonders großes Dankeschön.

Zum Schluss spreche ich den vielen Menschen ein Dank aus, die sich als Versuchsperson für meine Experimente zur Verfügung gestellt haben. Diese lustige Zeit hat mir gezeigt, dass ein Diplom neben der steinigen Arbeit auch sehr viel Spass machen kann.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Ziele und Aufgaben	2
1.2	Gliederung der Arbeit	2
2	Grundlagen	5
2.1	Funktionsweise des visuellen Systems	5
2.1.1	Aufbau des menschlichen Auges	6
2.1.2	Bedeutung okularer Bewegungen	7
2.1.3	Visuelle Aufmerksamkeit	9
2.1.4	Räumliche Wahrnehmung	10
2.1.5	Objektwahrnehmung	12
2.2	Eye-Tracking	15
2.2.1	Methoden	15
2.2.2	Kalibrierung	17
2.2.3	Metriken	17
2.3	Zusammenfassung	20
3	Verwandte Arbeiten	21
3.1	Datenaufnahme und -analyse	21
3.1.1	Erweiterte Verfahren zur Kalibrierung und Driftkorrektur	21
3.1.2	Auswertestrategien für Blickpunktdaten	23
3.1.3	Blickverhalten bei der Betrachtung dynamischer Stimuli	23
3.1.4	Fehldeutung von Eye-Tracking-Daten	24
3.1.5	Blickmuster von Radiologen bei Röntgenbildern	25
3.2	Weiterverarbeitung von Eye-Tracking-Daten	27
3.3	Objektbestimmung aus Bildschirmkoordinaten	29
3.3.1	Objektbasierter Ansatz	29
3.3.2	Bildbasierter Ansatz	30
3.4	Zusammenfassung	31
4	Konzept	35
4.1	Anforderungsanalyse	35
4.1.1	Merkmalsanalyse medizinischer Visualisierungen	35
4.1.2	Anforderungen an die Evaluierungsversuche	42
4.1.3	Schlussfolgerungen	45
4.2	Verfahren zur Bestimmung fokussierter Objekte	46
4.2.1	Schritte nach dem Experiment	47
4.2.2	Vorstellung objekt- und bildbasierter Metriken	48
4.2.3	Kombination aller Metriken	54
4.2.4	Datenfilterung	55
4.2.5	Zusammenfassung der Wahrscheinlichkeiten und globale Analyse	56
4.3	Zusammenfassung	58

5	Implementierung	59
5.1	Vorstellung der Entwicklungswerkzeuge	59
5.1.1	MeVisLab	59
5.1.2	Open Inventor	60
5.1.3	MATLAB	61
5.2	Das Eye-Tracking-System	61
5.2.1	Vorstellung von EyeLink I	61
5.2.2	Kopplung von EyeLink I und MeVisLab	64
5.3	Implementierungsdetails zur Objektbestimmung	65
5.3.1	Aufbau der Protokolldatei	65
5.3.2	Berechnung der Metriken	67
5.3.3	Datenauswertung in MATLAB	70
5.3.4	Diskussionen zur Laufzeit	72
5.4	Zusammenfassung	72
6	Evaluierung und Ergebnisse	73
6.1	Evaluierungsversuche und Hypothesen	73
6.1.1	Fixierungen in verdeckungsfreien Szenen	73
6.1.2	Semi-Transparente Verdeckung	75
6.1.3	Fixierungswechsel zwischen bewegten Objekten	77
6.2	Versuchsdurchführung	78
6.3	Ergebnisse der Evaluierungsversuche	79
6.4	Auswertung des Fragebogens	83
6.5	Schlussfolgerungen	84
7	Zusammenfassung und Ausblick	87
7.1	Weiterführende Entwicklungen	88
	Abkürzungen und Akronyme	III
	Literaturverzeichnis	V
A	Bedeutung der Gestaltprinzipien	XIII
B	Fragebogaufbau und -auswertung	XV
C	Experimentauswertung	XXI

Kapitel 1

Einleitung

Die Visualisierung dreidimensionaler Daten ist heute schon sehr weit fortgeschritten. Ein Beispiel hierfür sind Computertomographiedaten aus der Medizin. Sie lassen sich, beispielsweise durch geeignete Verarbeitungsschritte und Darstellungstechniken, in interaktive 3D-Visualisierungen überführen. Dadurch wird es möglich, sich räumlich durch Patientendaten zu bewegen sowie einzelne Objekte zu selektieren und Informationen zu diesen anzuzeigen.

Es existieren derzeit mehrere Visualisierungstechniken. Zwei der bekanntesten sind das direkte und das indirekte Volumenrendering. Durch letztere werden polygonale Oberflächenmodelle aus medizinischen Schichtdatensätzen generiert. Auf diese Weise entstandene Darstellungen werden in der vorliegenden Arbeit betrachtet. Zum Erzeugen spezieller Effekte gibt es außerdem Verfahren, wie semi-transparente Flächen für die Exploration verdeckter Objekte, Silhouetten zur Hervorhebung kritischer Strukturen, illustrative Darstellung um irrelevante Informationen zu Reduzieren oder wichtiger Objekte durch wahrnehmungsorientierte Farbgebung zu betonen.

Eine Vielzahl von Studien befasst sich mit der Effektivität und Effizienz medizinischer Visualisierungen. In BURGERT et al. [2007] wird ein direkter Vergleich zwischen Computertomographiedaten und 3D-Darstellungen hergestellt. Während der durchgeführten Experimente zeichnet ein Gerät zur Blickpunktregistrierung (Eye-Tracker) die teilweise unbewusst ausgeführten Augenbewegungen der teilnehmenden Chirurgen auf. Ziel dieser Untersuchung ist die Suche nach auffälligen Mustern in den einzelnen Blickpunkten.

Da 3D-Visualisierungen heute bereits zur Interventionsplanung in der Medizin eingesetzt werden, ist die Entwicklung von Evaluierungstechniken von großer Bedeutung. Dazu besteht die Möglichkeit der Bewertung auf der Basis eines Fragebogens, wobei subjektive Meinungen, zu verschiedenen Aspekten der Darstellungen, eingeholt werden. Einen objektiven Ansatz stellt die Aufzeichnung der Blickpunkte während des Experimentes mittels eines Eye-Trackers dar. Das hat den Vorteil der Messung und Analyse von Daten, auf die der Proband keinen direkten Einfluss hat. Dazu gehören z.B. unbewusste Augenbewegungen oder die Veränderung der Pupillengröße.

Mittels unterschiedlicher Techniken der Computergrafik sowie den 2D-Blickpunkten lassen sich fokussierte Objekte ermitteln. Dabei erzeugen jedoch gerade die in medizinischen Visualisierungen häufig eingesetzten Transparenzen Probleme. Die Fixierung einer Struktur ist schwierig zu identifizieren, wenn sich diese hinter einem oder mehreren transparenten Objekten befindet. Außerdem ist bislang wenig über den Einfluss der Präsentation einer animierten gegenüber einer interaktiv modifizierbaren Visualisierung als Stimulus bekannt. Mit Stimulus wird die, während des Experimentes gezeigte Darstellung, bezeichnet. Beispiele hierfür sind statische Bilder und animierte bzw. interaktive Visualisierungen.

1.1 Ziele und Aufgaben

Das Ziel dieser Diplomarbeit ist es, ein Verfahren zu entwickeln, das auf der Grundlage von bewegten, dreidimensionalen Darstellungen sowie Blickpunktdaten erstmals die Bestimmung von betrachteten Objekten ermöglicht. Diese Arbeit stellt damit eine Grundlage zur Evaluierung von medizinischen Visualisierungen dar. Zur Umsetzung dieses Ziels müssen verschiedene Aufgaben bewältigt werden.

Zunächst erfolgt die Identifikation und Beschreibung eines initialen Verfahrens zur Bestimmung fokussierter Objekte. Mit der Durchführung von Evaluierungsversuchen wird dieses Verfahren anschließend bewertet. In den Experimenten werden ausschließlich animierte Darstellungen verwendet. Der Betrachter kann sich dadurch vollständig auf die Visualisierung konzentrieren. Die Komplexität der dargebotenen Visualisierungen ist zunächst auf primitive Objekte beschränkt. Diese haben jedoch Gemeinsamkeiten mit Strukturen medizinischer Darstellungen, welche in dieser Arbeit hervorgehoben werden. Durch diese Vereinfachungen können erste Erkenntnisse gewonnen werden, welche die Basis für weitere Forschungen bilden.

Des Weiteren wird in der vorliegenden Diplomarbeit auf eines der wichtigsten Probleme eingegangen: Das Auflösen von Doppeldeutigkeiten in der Objektidentifikation, verursacht durch transparente Objekte.

Für die Realisierung dieser Aufgaben werden die Entwicklungsumgebung MEVISLAB, die Grafikbibliothek OPEN INVENTOR sowie das Analysewerkzeug MATLAB verwendet.

1.2 Gliederung der Arbeit

Kapitel 2 - Alle grundlegenden Kenntnisse zum Verständnis der Arbeit sowie der angewandten Verfahren werden in Kapitel zwei vorgestellt. Dies umfasst neben anatomischen und psychologischen Betrachtungen zum menschlichen, visuellen System ebenfalls Grundlagen zum Eye-Tracking. Der zweite Teil befasst sich anschließend mit der Gegenüberstellung alternativer Methoden zur Blickpunktregistrierung, der Kalibrierung eines Eye-Tracking-Systems sowie der Vorstellung von Möglichkeiten zur Auswertung aufgenommener Blickpunkte.

Kapitel 3 - Dieses Kapitel präsentiert und diskutiert die hier formulierte Problemstellung verwandte Arbeiten. Dabei werden allgemeine Erkenntnisse aus mehreren Studien, bezüglich des Eye-Trackings, zusammengetragen, welche die in diesem Gebiet auftretenden Probleme verdeutlichen und teilweise Hilfestellungen zur Lösung geben. Aber auch konkrete Techniken, die der Verbesserung des in Kapitel vier konzipierten Verfahrens dienen, werden hier herausgearbeitet, eingeschätzt und miteinander verglichen.

Kapitel 4 - Anhand einer Merkmalsbeschreibung medizinischer Datensätze erfolgt in diesem Kapitel die Anforderungsanalyse am Beispiel typischer Anwendungsszenarien. Neben den Anforderungen werden außerdem die für die Evaluierungsversuche geltenden Rahmenbedingungen vorgestellt. Den wesentlichen Teil bildet die Entwicklung eines Verfahrens zur Bestimmung fokussierter Objekte auf der Grundlage der Zuordnung von Fixierungswahrscheinlichkeiten.

Kapitel 5 - Zunächst werden MEVISLAB, OPEN INVENTOR, MATLAB und das verwendete Eye-Tracking-System EYELINK I vorgestellt. Es folgt eine Beschreibung, wie dieses System in die Entwicklungsplattform MEVISLAB integriert wurde. Der nachfolgende Teil bezieht sich

auf die Umsetzung der in Kapitel 4 vorgestellten Ansätze. Dazu gehören die Zuordnung der Blickpunkte zu Objekten, die Datenauswertung in MATLAB sowie eine Laufzeitbetrachtung.

Kapitel 6 - Das Kapitel beginnt mit der Erläuterung der Evaluierungsversuche. Nach der Beschreibung der Versuchsdurchführung erfolgt die Auswertung der von den Versuchspersonen auszufüllenden Fragebögen sowie der Ergebnisse aus den Versuchen. Der Fragebogen dient der Akquirierung subjektiver Meinungen der Versuchspersonen bezüglich der Experimentdurchführung sowie des Experimentdesigns. Dadurch wird die Qualitätssteigerung zukünftiger Studien ermöglicht. Die Ergebnisse werden am Beispiel repräsentativer Versuche diskutiert. Eine Zusammenfassung der wichtigsten Erkenntnisse aus den Versuchen schließt dieses Kapitel ab.

Kapitel 7 - Im letzten Kapitel wird das Wesentliche dieser Diplomarbeit zusammengefasst und anhand der in Abschnitt 1.1 definierten Ziele kritisch diskutiert. Die Erörterung offener Probleme sowie Vorschläge für zukünftige, themenbezogene Aufgabenstellungen schließen diese Arbeit ab.

Kapitel 2

Grundlagen

In dieser Diplomarbeit wird untersucht, inwiefern sich von einem Betrachter fokussierte Objekte in einer virtuellen, dreidimensionalen und bewegten Szene mit Hilfe eines Eye-Tracking-Systems (ETS) identifizieren lassen. Dazu sind Kenntnisse zur Anatomie und Funktionsweise des menschlichen visuellen Systems unerlässlich. In Abschnitt 2.1 werden der Aufbau des menschlichen Auges sowie verschiedene Augenbewegungen erläutert. Im weiteren Verlauf folgen Grundlagen psychologischer Natur. Die Bedeutung visueller Aufmerksamkeit sowie die Wahrnehmung räumlicher Tiefe spielen bei der Exploration von 3D-Umgebungen eine entscheidende Rolle. Da ein weiterer Schwerpunkt auf der Untersuchung von Blickmustern bei gegenseitiger Verdeckung von Objekten liegt, schließt der erste Grundlagenteil mit Betrachtungen zur Objektwahrnehmung ab.

Der zweite Teil erläutert Grundlagen zur Blickpunktbestimmung¹. Darunter wird die Registrierung und Auswertung von Punkten verstanden, die im visuellen Fokus einer Person liegen. Dabei werden 2D-Koordinaten auf einer Ebene (Bildschirm) erfasst. Um zu verdeutlichen auf welche Art und Weise diese Daten gewonnen werden können, erfolgt in Abschnitt 2.2.1 die Vorstellung mehrerer Verfahren zum Eye-Tracking. Eigenschaften, wie Genauigkeit oder Belastung für die Versuchsperson stehen dabei im Mittelpunkt. Die Kalibrierung des ETS repräsentiert einen bedeutenden Schritt vor jedem Experiment, da ohne sie keine korrekten Blickpunktdaten aufgenommen werden können. Daher wird in Abschnitt 2.2.2 der allgemeine Ablauf und die Notwendigkeit der Kalibrierung beschrieben. Das Grundlagenkapitel schließt mit der Vorstellung häufig verwendeter Eye-Tracking-Metriken ab.

2.1 Funktionsweise des visuellen Systems

Neben dem Aufbau des Auges wird in diesem Abschnitt die oft verwendete Einheit „Grad Sehwinkel“ erläutert. Darauf aufbauend erfolgt die Vorstellung typischer okularer Reflexe und Bewegungen. Dies umfasst beispielsweise Sakkaden und Fixierungen wie auch den Lidschlussreflex sowie die Veränderung der Pupillengröße. Des Weiteren wird der Begriff „Aufmerksamkeit“ und deren Verteilung auf einzelne Regionen des Sichtfeldes, erläutert. Daraus ergeben sich Hinweise, wie Blickpunkte „geplant“ werden. Die letzten beiden Abschnitte beziehen sich auf die menschliche Wahrnehmung. Zunächst erfolgt die Vorstellung von Tiefenhinweisen. Sie versetzen den Menschen in die Lage, die Entfernung eines Objektes zu bestimmen.

Ein weiteres, für diese Arbeit bedeutendes, psychologisches Fachgebiet ist die Objektwahrnehmung. Das menschliche Gehirn kann einerseits, ein Objekt als zusammenhängend erkennen,

¹ In der Fachliteratur und im weiteren Verlauf dieser Arbeit wird für den Begriff Blickpunktbestimmung auch das Synonym Eye-Tracking verwendet.

auch wenn es teilweise verdeckt ist. Andererseits werden Objekte gruppiert, die nur aufgrund ähnlicher Eigenschaften als zusammenhängend erscheinen. So genannte Gestaltprinzipien beschreiben diese Effekte. Da Verdeckungen sowie Gruppierungseffekte in dreidimensionalen Visualisierungen vermehrt auftreten, setzt sich Abschnitt 2.1.5 genauer mit diesen auseinander.

2.1.1 Aufbau des menschlichen Auges

Das Auge nimmt elektromagnetische Strahlung auf und wandelt diese mit Hilfe des Gehirns in Sinneseindrücke unserer Umwelt um. Die Wellenlänge des wahrnehmbaren Lichtes beträgt 380 bis 780 Nanometer. Oberhalb dieses Bereiches befindet sich u.a. die infrarote Strahlung. Sie ist für das menschliche Auge unsichtbar und wird deshalb bei einigen ETS's zur Ausleuchtung der Augenregion genutzt.

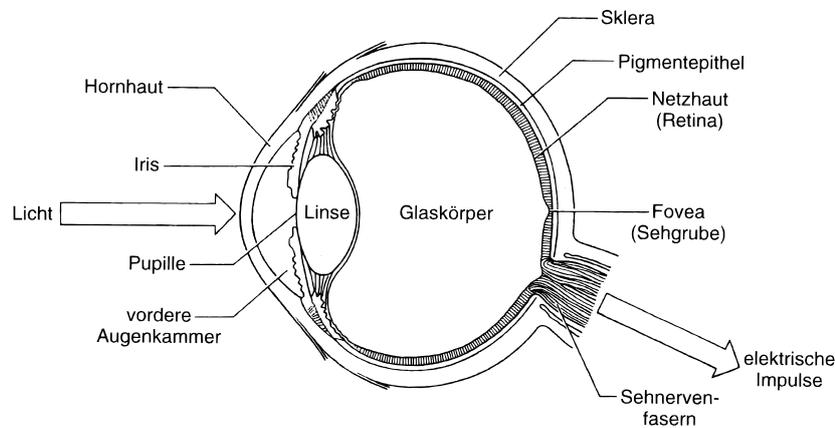


Abbildung 2.1: Bestandteile des menschlichen Auges [GOLDSTEIN, 1997]

Abbildung 2.1 zeigt den Querschnitt des menschlichen Auges. Einfallendes Licht durchdringt zunächst die Hornhaut. Die durch die Pupille gelangenden Wellen, treffen anschließend auf die Linse. Deren Brechkraft wird durch einen Vorgang namens Akkomodation reguliert, was zur korrekten Fokussierung des einfallenden Lichtes auf der Netzhaut (Retina) führt [GOLDSTEIN, 2008].

Die Retina besteht aus einem zentralen und peripheren Bereich. Der zentrale Bereich ist auch unter den Namen *Fovea* oder gelber Fleck bekannt. Die Wahrnehmung auf die Fovea abgebildeter Objekte erfolgt mit der größtmöglichen Sehschärfe, was an der Anordnung der Sinneszellen in diesem Bereich liegt. Die Sehschärfe nimmt mit zunehmendem Abstand zur Fovea ab. Der periphere Bereich ist laut DUCHOWSKI [2007] empfindlicher gegenüber Bewegungen. Peripherie und Fovea arbeiten des Weiteren auch eng im Bereich der Aufmerksamkeitsverteilung zusammen (vgl. Abschnitt 2.1.3).

Im Bereich des Eye-Tracking sowie im weiteren Verlauf dieser Diplomarbeit werden Objektgrößen in Form von Grad angegeben. Diese Einheit bezieht sich auf das Maß des „Schwinkels“. Er definiert das Ausmaß der Abbildung eines Objektes auf der Netzhaut und ist von dessen Größe und Entfernung abhängig. Der Schwinkel A , mit der Einheit $[1^\circ]$, nimmt mit zunehmender Größe S und abnehmender Distanz D eines Objektes zum Betrachter hin zu. Gleichung 2.1 drückt entsprechende Verhältnisse aus. Die Fovea beansprucht beispielsweise

eine Fläche von einem Grad. Ein fokussiertes Objekt ist damit immer noch scharf erkennbar, wenn die Fixierung maximal ein Grad neben dieses Objekt fällt.

$$A = 2 \arctan \frac{S}{2D} \quad (2.1)$$

Da beim Eye-Tracking meistens auf einen Bildschirm geschaut wird, ist S oft in der Einheit [1 *Pixel*] verfügbar. D , der Abstand des Auges zum Display, wird in diesem Fall in [1 *cm*] gemessen. Aus den Gründen ist es erforderlich, eine der beiden Einheiten in die jeweils andere zu überführen. Die Umrechnung von Pixel in Zentimeter ist abhängig von der Bildschirmauflösung und der Bildschirmdiagonale. Mit beiden Größen lässt sich die Umrechnung nach Gleichung 2.2 durchführen. Dabei bezeichnen S_{cm} und S_{pix} die Objektgröße in Zentimeter bzw. Pixel. Die horizontale und vertikale Bildschirmauflösung in Pixel wird durch res_x sowie res_y und die sichtbare Bildschirmdiagonale in Zentimeter wird durch $diag$ dargestellt.

$$S_{cm} = S_{pix} \frac{diag}{\sqrt{res_x^2 + res_y^2}} \quad (2.2)$$

2.1.2 Bedeutung okularer Bewegungen

Die Augen stehen fast nie still und führen ständig viele verschiedene Bewegungen bzw. Reflexe aus, die alle einem bestimmten Zweck dienen. Die für diese Diplomarbeit wichtigsten Bewegungen werden im Folgenden genannt und erläutert.

Sakkade - Zur Erfassung der Gesamtheit einer Szene ist der Mensch dazu gezwungen, Sakkaden auszuführen. Sie entstehen für gewöhnlich durch Reflexe, können aber auch freiwillig ausgelöst werden und dienen dem Verschieben der Aufmerksamkeit auf eine andere Region. Sakkaden sind außerdem sehr schnelle Augenbewegungen und haben eine Dauer von nur 10 bis 100 Millisekunden. Aufgrund dieser Geschwindigkeit ist der Mensch während einer Sakkade annähernd blind [TRIESCH et al., 2002]. Sie werden daher von vielen Forschern einfach ignoriert.

Fixierung - Sobald eine Sakkade abgeschlossen ist, folgt immer eine Fixierung. Viele Arbeiten bestätigen, dass diese den Wunsch des Betrachters repräsentieren, seine Aufmerksamkeit auf einer bestimmten Stelle zu halten [DUCHOWSKI, 2007; NODINE und KUNDEL, 1987]. Fixierungen haben eine minimale Dauer von 100 Millisekunden, da das Gehirn Zeit zur Verarbeitung des visuellen Inputs benötigt [RAYNER, 1998; POOLE et al., 2004]. Die Bezeichnung Fixierung kann den Eindruck vermitteln, dass der Augapfel während dieser Phase still steht. Jedoch beschrieb schon VON HELMHOLTZ [1867], dass dies nicht der Fall ist. Aus diesem Grund wird der Begriff Fixierung im weiteren Verlauf der Diplomarbeit auch als Augenbewegung bezeichnet.

Driftbewegung, Tremor, Mikrosakkade - Nach MARTINEZ-CONDE et al. [2004] können Fixierungen in drei Miniaturbewegungen unterteilt werden: Drift, Tremor und Mikrosakkaden. Von dem eigentlichen Fokuspunkt haben diese Miniaturbewegungen nach JACOB und KARN [2003] eine Abweichung von ca. zwei Grad. Dabei dienen Driftbewegungen zur Kompensation von Kopf- und Körperbewegungen. Die Fixierung bleibt dadurch auf das fokussierte Objekt gerichtet. Tremor bezeichnet Bewegungen relativ kleinen Ausmaßes, deren Erforschung sich bislang noch im Anfangsstadium befindet. Mikrosakkaden sind kleine, zufällige Sakkaden um die eigentliche Fixierung herum. Zusammen

mit den anderen beiden Miniaturbewegungen können Mikrosakkaden in computergestützten Analysen als Rauschen betrachtet werden, da sie reflexartig ausgelöst werden. Sie repräsentieren Ausgleichsbewegungen und damit nicht den Willen des Betrachters bestimmte Teile der Szene zu fokussieren.

(Langsame) Folgebewegung - Bei der Betrachtung bewegter Szenen am Computer stellt sich die Frage, ob hier ausschließlich sprunghafte Sakkaden und Fixierungen auftreten, wie es bei der Betrachtung statischer Bilder vorkommt. Bei der Betrachtung von Bildern ist das Auge im Allgemeinen nicht in der Lage, kontinuierliche Blickrichtungsänderungen auszuführen. Wird jedoch ein sich langsam bewegendes Objekt fixiert, kann es laut ANDREASSI [1995] verfolgt werden. Die Fixierung wird somit automatisch über die Bewegung hinweg aufrecht erhalten und die Informationsverarbeitung in diesem Zeitraum nicht unterbrochen. Der beschriebene Reflex wird nach CARPENTER [1988] durch ein nachgeschaltetes Kontrollsystem in der entsprechenden Hirnregion gesteuert. Er tritt nicht nur im realen Leben sondern auch bei der Betrachtung von z.B. animierten 3D-Visualisierungen am Bildschirm auf. Da bei der Veränderung der Kameraposition in einer virtuellen Darstellung ebenfalls Objektbewegungen verursacht werden, sind während dieser Zeit auch Folgebewegungen möglich. Das beschriebene Verhalten trägt den Namen (langsame) Folgebewegung und kann in speziellen Fällen auch freiwillig ausgelöst werden [KAUZLIS, 2005]. Da auch während Folgebewegungen das Interesse bestimmten Regionen oder Objekten gilt, werden diese nachfolgend ebenfalls als Fixierungen bezeichnet.

Vergenzbewegung - Eine der letzten, hier behandelten und auf die Augäpfel bezogenen Bewegungen heißen Vergenzbewegungen. Sie bezeichnen die Drehung der Augen zueinander oder voneinander weg. Wird z.B. ein Stift in Armlänge vor den Kopf gehalten und langsam auf die Nase zubewegt, so drehen sich beide Augen nach Innen. Bei einer Bewegung des Stiftes vom Kopf weg, drehen sie sich wieder nach außen. Dieser Prozess hält die Fixierung des betrachteten Objektes aufrecht. Solche Bewegungen können beim Eye-Tracking durch spezielle Formen der Stimuluspräsentation hervorgerufen werden. Mit Hilfe angepasster Kalibrierungsverfahren ist es dann möglich, dreidimensionale Blickpunkte zu messen. In DONNER [2007] wird genau dieser Ansatz verfolgt.

Pupille - In HOEKS und LEVELT [1993] sowie BEATTY [1982] wird erwähnt, dass der Pupillendurchmesser mit dem Konzentrationsgrad bzw. der Aufmerksamkeit eines Betrachters bei visuellen Suchaufgaben variiert. MORAD et al. [2000] haben Verfahren zur Quantifizierung der Eigenschaften Aufmerksamkeit und Müdigkeit mit Hilfe des Pupillendurchmessers entwickelt. Dies ist keine triviale Aufgabe, da sich durch äußere Einflüsse wie Lichteinfall oder Verdeckung der Pupille bei Lidschlägen die Pupillengröße nicht immer eindeutig bestimmen lässt. POMPLUN und SUNKARA [2003] entwickelten aus diesem Grund ein auf neuronalen Netzen basierendes Kalibrierungsverfahren, welches genau diese Einflüsse fast vollständig beseitigt. Somit wird die Pupillengröße, als objektives Maß, zur Auswertung von Eye-Tracking-Daten nutzbar.

Lidschlussreflex - Das Augenlid ist für die einwandfreie Funktion des Sehapparates von Bedeutung, da es durch die Abdeckung des Auges u.a. eine Schutzfunktion ausübt. Dadurch wird das Eindringen von Schmutzpartikeln verhindert. Sollten dennoch Fremdkörper ins Auge gelangen, so dient der Lidschlussreflex dazu, das Auge zu befeuchten und es von diesen zu befreien. Die Dauer des Reflexes hängt laut HÄKKÄNEN et al. [1999] vom Müdigkeitsgrad und laut PESHORI et al. [2001] vom Alter der Versuchsperson ab. Die meisten Eye-Tracker sind in der Lage, diesen Reflex aus den aufgenommenen Daten

zu erkennen und entsprechend zu markieren. Das ist notwendig, da die Daten während eines Lidschlusses aufgrund der Pupillenverdeckung hochgradig verrauscht sind. In SEN-SOMOTORIC INSTRUMENTS GMBH [1999] wird des Weiteren empfohlen, Fixierungen vor und nach einem detektierten Lidschluss genau zu untersuchen, da das Rauschen auch darauf Einfluss haben kann.

2.1.3 Visuelle Aufmerksamkeit

Aufmerksamkeit ist eine begrenzte kognitive Ressource, welche die Wahrnehmung ausgewählter Sinneseindrücke ermöglicht. Während sich der Mensch auf bestimmte Sinneseindrücke konzentriert, ignoriert er viele andere. Das ist notwendig, da er nicht alles, was um ihn herum geschieht, auf einmal verarbeiten kann. Aufmerksamkeit, bezogen auf die vom Auge wahrgenommenen Sinneseindrücke, wird im weiteren Verlauf als „visuelle Aufmerksamkeit“ bezeichnet. Als Beispiel hierfür können Werbeplakate angeführt werden. Da diese möglichst häufig gesehen werden sollen, nutzen werbende Firmen u.a. Portraits berühmter Personen, kontrastreiche Farben oder übergroße Darstellungen. All diesen „Aufmerksamkeitsmagneten“ ist gemein, dass sie unbewusst wahrgenommen werden und für viele Menschen interessant sind. Das führt häufig dazu, dass sich diese Personen die Werbung im Detail ansehen.

Um ein solches Werbeplakat komplett zu erfassen, ist die Ausführung von Fixierungen und Sakkaden notwendig. Beide sind für die visuelle Wahrnehmung wichtig, da sie eine detaillierte Analyse der vom Gehirn auserwählten Bereiche erlauben. Es gibt mehrere Ansätze, wie die visuelle Aufmerksamkeit auf verschiedene Regionen des Sichtfeldes verteilt wird.

Die *Feature Integration Theory*, vorgestellt in TREISMAN und GELADE [1980], ist eines der am meisten angewandten Modelle zur Beschreibung der Aufmerksamkeit. Nach dieser Theorie extrahiert das Gehirn in einem ersten Schritt der visuellen Verarbeitung primäre *Features* und speichert diese in einer *Feature-Map*. Der zweite Schritt besteht aus der Integration dieser Feature-Map in eine *Saliency-Map*, mit deren Hilfe die Aufmerksamkeit auf besonders hervorstechende Bereiche gelenkt wird. Bereits 1867 beschreibt Hermann von Helmholtz in VON HELMHOLTZ [1867] die Tendenz zur Verschiebung der Aufmerksamkeit auf markante Objekte. In TREISMAN und GELADE [1980] wird außerdem primitiven Merkmalen wie Farbe, Orientierung oder Intensität die Eigenschaft zugeschrieben, präattentiv wahrnehmbar zu sein. Diese werden parallel (schnell) erfasst. Die Erkennung von Kombinationen solcher Merkmale erfolgt dagegen seriell (langsam).

Aus der oben beschriebenen Theorie kann ein zyklisches Modell zur Aufmerksamkeitsverteilung abgeleitet werden, welches nach DUCHOWSKI [2007] wie folgt aufgebaut ist:

1. Die Erfassung des Stimulus (z.B. eines Bildes) erfolgt zunächst parallel durch peripheres Sehen. Einzelne Merkmale stechen hierbei heraus.
2. Eine Sakkade repositioniert die Augen auf eine in Schritt eins identifizierte und somit für den Betrachter interessante Region. Die Aufmerksamkeit wird während dieser Zeit deaktiviert.
3. Nachdem Schritt zwei abgeschlossen ist, wird die Aufmerksamkeit reaktiviert. Die Fovea ist nun auf einen neuen Bereich ausgerichtet. Dieser kann dadurch mit hoher Sehschärfe analysiert werden.

Dieses Modell bildet die Grundlage zur Erstellung erweiterter Modelle zur visuellen Aufmerksamkeit. Es ist jedoch nach DUCHOWSKI [2007] unvollständig, was anhand zweier Fragen deutlich gemacht werden soll:

1. Wenn allein Features des Stimulus für die Aufmerksamkeitskontrolle zuständig sind, wozu benötigt der Mensch dann freiwillige Augenbewegungen?
2. Liegt die Aufmerksamkeit des Betrachters immer auf den fixierten Regionen einer Szene?

Die erste Frage steht in direktem Zusammenhang mit den Experimenten in YARBUS [1967]. Verschiedenen Probanden wird je eine Suchaufgabe für ein und dasselbe Bild gestellt. Es hat sich herausgestellt, dass die Folge von Fixierungen, bei jedem Probanden verschieden und damit aufgabenabhängig ist. Jede Person macht also freiwillige Augenbewegungen auf der Basis ihrer jeweiligen Erwartungen und Erfahrungen bezüglich des Gesehenen. Frage zwei ist mit nein zu beantworten. Mehrere Studien, darunter POSNER et al. [1980], haben die Möglichkeit aufgezeigt, Aufmerksamkeit freiwillig auf den peripheren Bereich zu lenken. Dieser Prozess wird in der Fachliteratur mit „verdeckter Aufmerksamkeit“ bezeichnet. In der Fachliteratur wird zwar immer wieder darauf hingewiesen, dass diese Aufmerksamkeitsverschiebungen nicht messbar seien, jedoch untersuchen HAFED und CLARK [2002] sowie ENGBERT und KLIEGL [2003] Ansätze zu deren Detektion auf Basis von Mikrosakkaden. In dieser Diplomarbeit wird jedoch von der Annahme ausgegangen, dass die Aufmerksamkeit der Versuchsperson auf den jeweils fixierten Regionen liegt, so wie dies bei den meisten Eye-Tracking-Studien der Fall ist.

2.1.4 Räumliche Wahrnehmung

Alles was der Mensch sieht, wird als zweidimensionales Abbild auf die Retina projiziert. Durch diese Projektion des dreidimensionalen Raums auf eine zweidimensionale Fläche, gehen viele räumliche Informationen bezüglich wahrgenommener Objekte verloren. Dennoch ist es für die visuelle Wahrnehmung durch so genannte Tiefenhinweise möglich, das auf die Retina projizierte Bild in einen dreidimensionalen Eindruck unserer Umwelt umzuwandeln. WANGER et al. [1992] beschreiben diese Kriterien und weisen sie mittels Experimenten nach. Dieser Abschnitt befasst sich speziell mit der räumlichen Wahrnehmung von Bildern durch monokulare Tiefenhinweise.



Abbildung 2.2: Darstellung bildbezogener, monokularer Tiefenhinweise. In (a) werden *Verdeckung*, *relative Größe*, *perspektivische Konvergenz* und *Schatten* veranschaulicht. In (b) wirkt durch die *vertraute Größe* des Geldes das 1-Cent-Stück näher und die 2-Euro-Münze ferner als das 10-Cent-Stück.

Monokular bedeutet einäugig und bezieht sich auf Tiefenhinweise, welche mit einem Auge erkennbar sind. Daneben gibt es ebenfalls binokulare Tiefenhinweise. Zu den monokularen

gehören bildbezogene (vgl. Abbildung 2.2), bewegungsinduzierte und okulomotorische Hinweise. Die bildbezogenen können also mit Hilfe von nur einem Auge aus Bildern extrahiert werden. Sie werden im Folgenden erläutert.

Verdeckung - Jeder Mensch ohne entsprechende Sehstörungen nutzt Hilfen zur Entfernungsschätzung täglich. In Abbildung 2.2(a) kann durch die Verdeckung der hintereinander angeordneten Würfel erkannt werden, dass der rote vor dem grünen und der grüne vor dem blauen Würfel steht. Verdeckte Objekte nimmt der Mensch als weiter entfernt wahr. Besonders hervorzuheben ist hierbei die Verdeckung durch transparente Objekte. Ein Objekt wird als weiter entfernt wahrgenommen, wenn sich dessen tatsächliche Farbe mit der des transparenten Objektes vermischt und sich damit hinter diesem befindet.

Relative Größe und Texturgradient - Die relative Größe von Objekten steht ebenfalls im Zusammenhang mit deren Entfernung. Die Würfel in Abbildung 2.2(a) haben alle die gleichen Kantenlängen, wirken aber aufgrund ihrer Nähe oder Ferne zum Beobachter größer bzw. kleiner. An dieser Stelle ist auch der Texturgradient zu erwähnen, da dieser Tiefenhinweis im Zusammenhang mit der relativen Größe steht. Die Muster in Texturen scheinen, obwohl sie gleich groß sind, mit zunehmender Entfernung dichter angeordnet zu sein. Dies wird in Abbildung 2.2(a) an den rechten Würfelflächen deutlich.

Perspektivische Konvergenz - Ein Merkmal der perspektivischen Konvergenz ist, dass parallel in die Tiefe verlaufende Linien mit zunehmender Distanz als konvergierend wahrgenommen werden. Je weiter die Linien in die Tiefe reichen, desto stärker ist die Konvergenz. Dieser Effekt bewirkt außerdem, dass wahrgenommene Abstände und Winkel zwischen Objektkanten nicht mehr der Realität entsprechen. Da in medizinischen Visualisierungen die Notwendigkeit besteht, beispielsweise Abstände zwischen anatomischen Strukturen korrekt zu messen, wird dort die orthographische der perspektivischen Darstellung vorgezogen. Dies hat den Vorteil, dass gleichgroße Objekte, unabhängig von ihrer räumlichen Position, ebenfalls gleichgroß auf den Bildschirm projiziert werden. Die perspektivische Konvergenz geht dabei verloren, so dass sie nicht mehr als Tiefenhinweis genutzt werden kann.

Schatten - Schatten verstärken das dreidimensionale Erscheinungsbild. In WANGER et al. [1992] finden sich Beispiele dazu. Ein Mensch kann ohne Schatten in Abbildung 2.2(a) nicht sagen, ob der Zylinder auf der Ebene liegt oder in der Luft schwebt. Durch die zusätzliche Information wird jedoch sichtbar, dass letzterer Fall der Realität entspricht. In diesem Zusammenhang ist auch zu erwähnen, dass die räumliche Wahrnehmung von den Oberflächeneigenschaften eines Objektes abhängen. Fallen Licht bzw. Schatten auf zwei Objekte mit unterschiedlicher Oberflächenkrümmung, so wird die Krümmungsrichtung und damit die Objektstruktur erkennbar. Dies ist ebenfalls an dem Zylinder erkennbar.

Vertraute Größe - Auch die vertraute Größe von Objekten spielt eine wesentliche Rolle bei der Tiefenwahrnehmung. Liegen beispielsweise die Münzen aus Abbildung 2.2(b) vor dem Betrachter auf dem Tisch, so erzeugen sie bei diesem unterschiedliche Sehwinkel. Ihr wahrgenommener Durchmesser ist dadurch verschieden. Um die Münzen aus Abbildung 2.2(b) auf die dargestellte Weise wahrzunehmen, müssen alle den selben Sehwinkel einnehmen. Die Präsentation der einzelnen Münzen muss daher in verschiedenen Entfernungen zum Betrachter erfolgen. Die vertraute Größe steht laut LAW et al. [2004] außerdem in Zusammenhang mit der Erfahrung des Betrachters.

Bewegungsparallaxe - Die bislang beschriebenen Tiefenreize beziehen sich auf statische Bilder. Da die hier betrachteten Anwendungsfälle animiert sind, werden nun zwei bewegungsinduzierte Tiefenreize vorgestellt. Einerseits ist der Effekt der Bewegungsparallaxe bekannt. Die Bewegungsparallaxe ist schematisch in Abbildung 2.3 dargestellt. Es wird von einer Bewegung des Betrachters von Position A zu Position B ausgegangen. In diesem Zeitraum legt das retinale Abbild des gelben Kreises einen größeren Weg auf der Netzhaut zurück als das retinale Abbild des roten Kreises. Die verschiedenen langen Wege können dadurch mit der Entfernung der jeweiligen Objekte assoziiert werden.

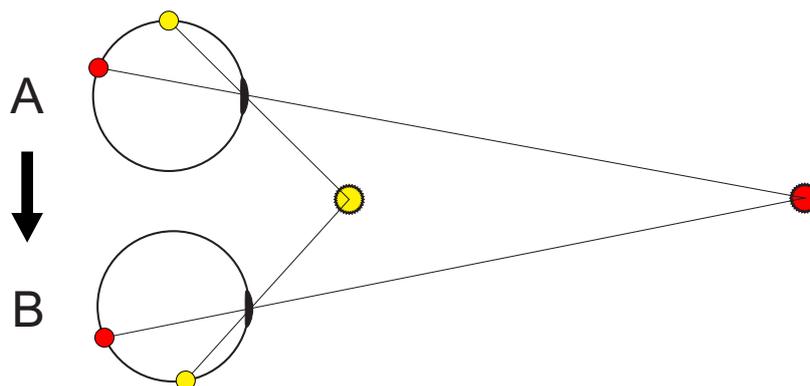


Abbildung 2.3: Schematische Darstellung der Bewegungsparallaxe bei der Betrachtung eines nahen und eines fernen Objektes nach [DONNER, 2007].

Die in diesem Abschnitt vorgestellten Tiefenreize beschreiben wahrgenommene Eigenschaften und Verhaltensweisen von Objekten der realen Welt. In den hier betrachteten Visualisierungen sind viele dieser Eigenschaften nachgebildet. So ist z.B. die Simulation von Schatten, Verdeckungen und Bewegung möglich. Auch bei einer Bewegung um die Szene herum, verhalten sich die oben beschriebenen Tiefenreize, wie es der Mensch aus dem täglichen Leben gewohnt ist. Das ist wichtig, um unvorhersehbare Blickmuster aufgrund ungewohnter, visueller Eindrücke zu vermeiden. Die realitätsfremde, orthographische Darstellung in medizinischen Visualisierungen bildet dabei eine Ausnahme. Dies ist jedoch kaum nachteilig zu sehen, da perspektivische Darstellungen den orthographischen ähneln und die menschliche Wahrnehmung diese Unterschiede kompensieren kann. Ärzte, die medizinische Visualisierungen häufig betrachten, gewöhnen sich an diese Darstellungsform, was zur Manifestation regelmäßiger Blickmuster führt. Des Weiteren ist hier festzuhalten, dass in dem Fall der Verdeckung durch semi-transparente Objekte eine Aufmerksamkeitsverschiebung vom Vorder- auf das Hintergrundobjekt bzw. umgekehrt möglich ist. In computergestützten Analysen kommt es dabei zu Doppeldeutigkeiten bei der Bestimmung des fokussierten Objektes, welche es zu umgehen gilt.

2.1.5 Objektwahrnehmung

Der vorliegende Abschnitt beschreibt die Grundlagen zur Objektwahrnehmung, einem zentralen Bestandteil dieser Arbeit. Psychologen erforschen bereits seit langem, wie das menschliche Gehirn die Bestandteile der betrachteten Welt zu einem Ganzen integriert. Dabei wurden auch Erkenntnisse zu Wahrnehmungseffekten gewonnen, welche den Fokus des Betrachters auf Bereiche lenken, die durch Gruppierung elementarer Teile der Szene entstehen.

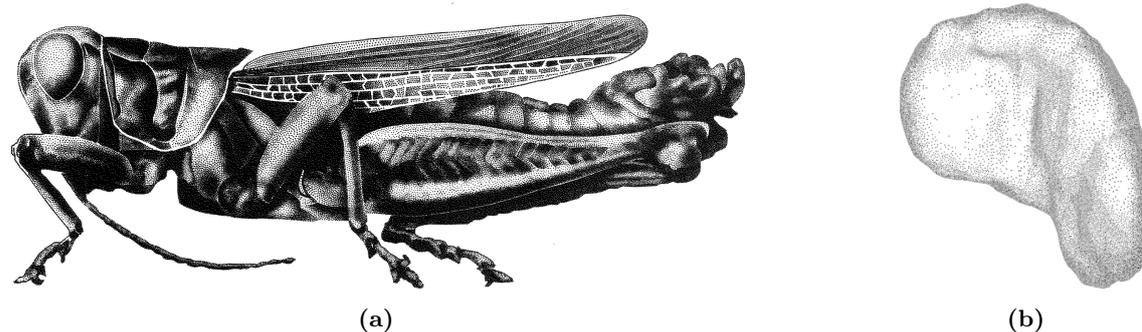


Abbildung 2.4: Abbildung (a) zeigt die Visualisierung eines Grashüpfers als Stipplingvariante [STROTHOTTE und SCHLECHTWEG, 2002]. In Abbildung (b) ist die menschliche Milz auf die selbe Weise dargestellt [BAER, 2005].

Die Forschungen zur Objektwahrnehmung können zurückgeführt werden auf den Ansatz des „Strukturalismus“. Er wurde von Psychologen im 19. Jahrhundert begründet und bezeichnet die Idee, dass Wahrnehmung durch die Verbindung elementarer Strukturen zustande kommt. Dieses Konzept wird durch Abbildung 2.4 verdeutlicht. Entstanden sind die dargestellten Bilder durch Stippling, ein aus der Kunst übernommenes Verfahren der Computergraphik, bei dem Bilder ausschließlich durch das Setzen von Punkten erzeugt werden. Genau betrachtet sind auf den Bildern daher nur diese Punkte (elementare Strukturen) zu sehen. Die Wahrnehmung erstellt daraus jedoch eine komplexere Form. Eine Person würde daher niemals auf die Idee kommen sich jeden einzelnen Punkt genauer zu betrachten, um Informationen über die eigentliche Darstellung zu beziehen. Stattdessen wird das Gesamtobjekt wahrgenommen, in welchem Objektmerkmale (wie Kanten) durch die spezielle Anordnung elementarer Strukturen erzeugt werden.

Die visuelle Wahrnehmung ist besonders sensibel für Kanten, da diese relativ kontrastreich sind und die Reizverarbeitung zusätzlich kontrastverstärkend wirkt. Auf dieser Basis werden in MARR [1982] Kantenfilter, auf dieser Basis, für den Einsatz in der Bildverarbeitung, vorgestellt. Die Grundlagen des Strukturalismus wurden zu Beginn des 20. Jahrhunderts durch die so genannte „Gestaltpsychologie“ wiederlegt. Sie war von diesem Zeitpunkt an das wichtigste Modell zur Beschreibung der menschlichen Wahrnehmung und wird durch GOLDSTEIN [2008] wie folgt definiert:

„Ein Ansatz in der Psychologie, der sich auf die Formulierung von Prinzipien der Wahrnehmungsorganisation konzentriert und dabei die Annahme zugrundelegt, dass „das Ganze Mehr ist, als die Summe ihrer Teile.“

Auch in der Gestaltpsychologie, wird von einer Gruppierung einzelner Objekte unter bestimmten Voraussetzungen ausgegangen. Es konnten so genannte „Gestaltprinzipien“ bzw. „Gestaltheuristiken“ identifiziert werden, die diesen Ansatz belegen. Tabelle A.1 stellt die Bedeutung der wichtigsten Gestaltprinzipien dar.

KAPADIA et al. [1995] führten psychologische Testexperimente durch und fanden heraus, dass bei der Präsentation von Reizmustern auf Basis dieser Prinzipien tatsächlich eine erhöhte neuronale Aktivität erfolgt.

Zwei der Gestaltprinzipien lassen sich an Abbildung 2.5 erläutern. Sie beziehen sich auf die Verdeckung der Lymphknoten durch den rechts dargestellten Halsmuskel. In Abbildung 2.5(a)

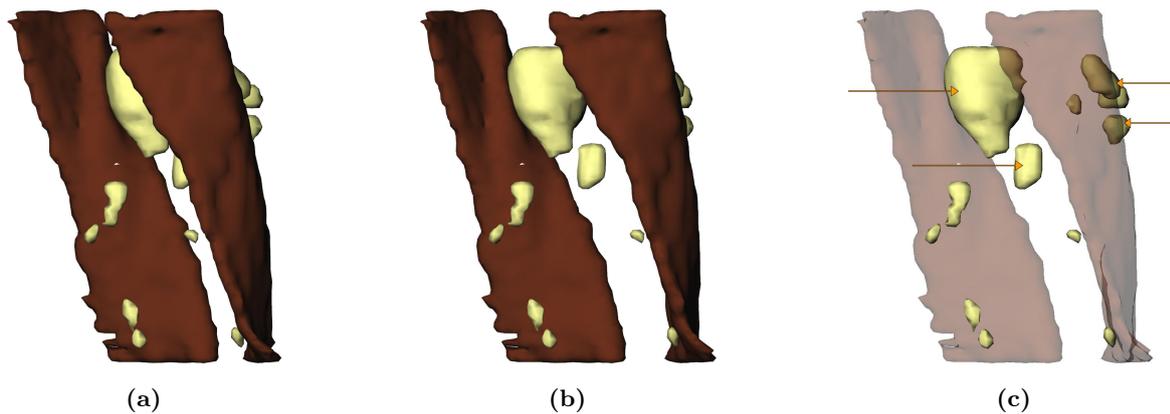


Abbildung 2.5: Veranschaulichung der Gestaltprinzipien der Fortsetzung und der Ähnlichkeit am Beispiel abnehmender Verdeckung ((a) und (b)) sowie Objektmerkmalen der Lymphknoten, die in (c) durch Pfeile markiert sind.

sieht es so aus, als ob der Muskel einen einzigen großen Lymphknoten verdeckt. Dies ist jedoch nicht der Fall, wie in Abbildung 2.5(c) zu sehen ist. Die einzelnen, in Abbildung 2.5(c) durch Pfeile markierten Lymphknoten, werden aufgrund der Verdeckung in Abbildung 2.5(a) als einheitliches Objekt erkannt. Dieser Effekt basiert auf den Gestaltprinzipien der Fortsetzung und der Ähnlichkeit. Das Prinzip der Fortsetzung besagt, dass Punkte als zusammenhängend wahrgenommen werden, wenn diese auf einer geraden bzw. leicht geschwungenen Linie liegen. Dies trifft für alle sichtbaren Punkte der Lymphknoten links und rechts des Halsmuskels zu. Durch das Prinzip der Ähnlichkeit wird dieser Effekt weiter verstärkt, da alle Lymphknoten der Szene die gleiche Farbe sowie eine ähnliche Form aufweisen. Die beschriebenen Hinweise vermitteln dem Gehirn, dass die in Abbildung 2.5(a) verdeckten Lymphknoten ein einziges Objekt repräsentieren müssen. Das visuelle System erzeugt in diesem Fall also ein falsches Bild der Realität. Aus diesem Grund werden die Gestaltprinzipien auch als Heuristiken bezeichnet, die meistens aber nicht immer zu korrekten Aussagen führen. Sie liefern jedoch schnelle Aussagen. Bei einer Veränderung des Betrachterstandpunktes werden die räumlichen Verhältnisse der in der Szene dargestellten Objekte schnell deutlich (Abbildung 2.5(b)).

Eine durch Herrmann von Helmholtz aufgestellte Theorie besagt, dass die menschliche Wahrnehmung teilweise durch Annahmen über die betrachtete Szene bestimmt ist. Hieraus kann geschlossen werden, dass erworbenes Wissen über die räumlichen Verhältnisse von Objekten in die weitere Erkundung der Szene mit einbezogen wird. Nachdem also die Lymphknoten aus Abbildung 2.5 einmal als nicht zusammenhängend gesehen wurden oder diese Tatsache bereits bekannt ist, wird auch nicht mehr vermutet, dass sie zu einem Objekt gehören könnten. Das Einbringen von Erfahrungen in den Erkennungsprozess erfolgt zudem unbewusst, was der Helmholtz'schen Theorie den Namen „Theorie der unbewussten Schlüsse“ verleiht.

In diesem Teil ist deutlich geworden, dass nicht immer ein tatsächlich präsenten Objekt wahrgenommen wird. Die Wahrnehmung erzeugt bei bestimmten Szenenkonstellationen den Eindruck neuer Objekte, die durch Gruppierung vorhandener Zustände kommen. In diesem Fall wird das wahrgenommene Objekt anstatt der tatsächlich existierenden Strukturen gesehen, was zu schwer interpretierbaren Blickmustern führen kann. Andererseits ist es einigen Personen aufgrund ihrer Erfahrung möglich, effizientere Blickmuster als unerfahrene Personen auszuführen. Beispielsweise ist ein erfahrener Arzt in der Lage, Lymphknoten schneller und

genauer einzuschätzen, als dies ein Assistenzarzt könnte. Diese Erkenntnisse sollten bei der Auswertung von Eye-Tracking-Daten unbedingt berücksichtigt werden.

2.2 Eye-Tracking

Eye-Tracking dient der Aufnahme von Blickpunkten einer Person, während diese etwas betrachtet oder eine visuelle Suchaufgabe löst. Ein Anwendungsbeispiel für Eye-Tracking ist die Usabilityanalyse. In entsprechenden Studien werden Geräte, Webseiten oder Software auf Benutzbarkeit bzw. Anwendbarkeit geprüft [RICHTER, 2008]. Eine Zusammenfassung mehrerer dieser Studien finden sich in JACOB und KARN [2003]. Die Blickpunktinformationen sind in vielen weiteren Bereichen anwendbar. Dazu gehören beispielsweise die Analyse des Blickverhaltens beim Lesen von Internetseiten, die Beurteilung der Effektivität von Werbung, die blickgesteuerte Interaktion mit Computerprogrammen [INFERENCE GROUP] oder die Bestimmung fokussierter Objekte, wie in diesem Fall.

In den folgenden Abschnitten sind die für diese Arbeit wichtigsten Grundlagen zum Eye-Tracking beschrieben. Zunächst werden Eigenschaften, wie das Fehlerverhalten verschiedener Verfahren zur Blickpunktbestimmung, miteinander verglichen. Der Genauigkeitsfehler moderner Geräte beträgt 0.5° - 1.0° . Im gegenseitigen Austausch mit Andrew T. Duchowski [DUCHOWSKI, 2007] hat sich bestätigt, dass der Genauigkeitsfehler als Radius eines Unsicherheitskreises um den tatsächlich betrachteten Punkt zu verstehen ist. In diesem Bereich können die ermittelten Blickkoordinaten gleichwahrscheinlich auftreten. Weitere Gespräche mit einem Mitarbeiter der Firma Tobii haben ergeben, dass zum Genauigkeitsfehler noch weitere Ungenauigkeiten hinzukommen. Dazu gehören auch die zufälligen Bewegungen des menschlichen Auges, welche zudem personenbezogen sind. Aufgrund der Addition aller Ungenauigkeiten ist es schwierig, die einzelnen Fehlerkomponenten aus den Daten zu bestimmen.

Bei der Betrachtung eines auf dem Bildschirm definierten Zielreizes gibt der Genauigkeitsfehler die mittlere Abweichung der vom ETS ermittelten Blickkoordinaten von der tatsächlichen Position des Zielreizes an. Soll der Fehler auf dem Bildschirm dargestellt werden muss dieser Wert mit der typischen Einheit [1°] in Pixel umgerechnet werden. Das erfolgt mittels der nach der Objektgröße S umgestellten Gleichung 2.1 und der Auflösung des Monitors. Beispielsweise entspricht, bei einer Auge-Display-Distanz von 75 cm , der Fehler von 1° einem Wert von $\approx 1.3\text{ cm}$. Bei einem Monitor mit einer Auflösung von 86 dpi ist dies ein Kreis mit 45 Pixeln Radius.

2.2.1 Methoden

Eye-Tracking-Methoden lassen sich nach YOUNG und SHEENA [1975] in Verfahren zur Messung der Augenpositionen relativ zum Kopf bzw. zur Messung absoluter Bildschirmkoordinaten unterteilen. Das „Search-Coil“-Verfahren wurde 1963 vorgestellt und ist nach MULLIGAN [1997] für Versuchspersonen eines der am belastendsten, da sie mit Spulen versehene Kontaktlinsen eingesetzt bekommen. Zudem werden sie von einem magnetischen Feld umgeben. Dadurch werden, unter Bewegung der Augen, Spannungsänderungen in den Spulen hervorgerufen. Positionsinformationen relativ zum Kopf können hierdurch, mit einer hohen Genauigkeit von ca. $1/60^\circ$, gemessen werden [DUCHOWSKI, 2007].

Des Weiteren gibt es das Verfahren der Elektrookulographie. Es wird vorwiegend von Psychophysiologen zur Messung von Augenbewegungen eingesetzt und hat nach ANDREASSI [1995]

eine Genauigkeit von 1.5° - 2.0° . Über Elektroden, welche um die Augen verteilt sind, werden elektrische Potentialänderungen gemessen. Da hierbei nichts in das Auge eingesetzt wird, ist dieser Ansatz weitaus weniger eindringlich als die Search-Coil-Methode. Die Messung erfolgt ebenfalls relativ zum Kopf.

Photo- bzw. Videookulographie bezeichnet die Zusammenfassung verschiedener Methoden zur Bestimmung der Veränderung von Merkmalen der Augen, während diese sich bewegen [DUCHOWSKI, 2007]. Da dies unter Verwendung von Videokameras erfolgt, belastet diese Methode den Probanden nur gering. Eine Ablenkung von der eigentlichen Aufgabe ist dadurch fast ausgeschlossen. Den Nachteil bildet jedoch die geometrische Verzerrung der Merkmale während sich die Augen bewegen. Bei einer zu großen Deformation können diese Merkmale von den Kameras nicht mehr erkannt werden. Diese Tatsache beeinflusst die Genauigkeit und Anwendbarkeit des Verfahrens negativ. Auch hier lassen sich keine absoluten Blickkoordinaten auf dem Bildschirm erfassen.



Abbildung 2.6: Abbildung (a) zeigt den kopfbefestigten Eye-Tracker EYELINK I der Firma SR Research. In Abbildung (b) ist das desktopplatzierte System MYTOBII P10 der Firma Tobii nach [TOBII EYE TRACKING & EYE CONTROL] dargestellt.

Photo- bzw. Videookulographie ähnelt einem heute oft angewandten Verfahren. In Abbildung 2.6 sind zwei moderne, videobasierte Systeme zur Blickpunktregistrierung dargestellt. Wie in der Photo- bzw. Videookulographie werden auch bei diesen Geräten Videokameras verwendet. Die Frequenz, mit der Bilder bzw. Blickpunkte aufgenommen und verarbeitet werden, beträgt 50 Hz bis 2000 Hz . Die ermittelten Informationen speichern solche Systeme als Samples² in einer zentralen Protokolldatei. Die entsprechenden Geräte sind entweder auf dem Kopf des Probanden befestigt (wenig eindringlich für die Versuchsperson, vgl. Abbildung 2.6(a)) oder befinden stationär auf dem Desktop (nicht eindringlich für die Versuchsperson, vgl. Abbildung 2.6(b)). Bei der stationären Variante sind die Kameras direkt im Bildschirm integriert. Die Versuchsperson hat damit keinerlei Kontakt mehr mit dem Eye-Tracker, was ein angenehmes Arbeiten ermöglicht. Die Genauigkeit solcher Systeme beträgt 0.5° - 1.0° [TOBII EYE TRACKING & EYE CONTROL; SENSO MOTORIC INSTRUMENTS GMBH, 1999].

²Ein Sample enthält die aus einer Videoaufnahme berechneten Daten. Ein Eye-Tracking-System mit einer zeitlichen Auflösung von 250 Hz , nimmt alle vier Millisekunden Bilder der Augen auf, errechnet daraus Blickkoordinaten sowie weitere Informationen und speichert die Daten als Sample in der Protokolldatei.

2.2.2 Kalibrierung

Die Güte der Daten aus dem Eye-Tracking-Protokoll hängt von vielen Faktoren ab. Der Proband muss beispielsweise ruhig sitzen und darf seinen Kopf während der Experimente kaum bewegen. Einige Tracking-Systeme sind in der Lage diese Bewegungen bis zu einem gewissen Grad auszugleichen. Eine der größten Fehlerquellen liegt jedoch in einem unsachgemäß ausgeführten Kalibrierungsprozess, welcher vor dem eigentlichen Experiment durchgeführt wird. Während dieser Zeit darf der Proband seinen Kopf nicht bewegen, da hier bereits kleinste Änderungen das Ergebnis verfälschen würden. Eine Kalibrierung wird durchgeführt, damit folgende Parameter bei der Blickpunktberechnung berücksichtigt werden können:

- Unterschiede im Aufbau des Auges einer jeden Person
- Relative Position der Kameras zum Kopf bei kopfbefestigten Systemen
- Relative Position des Kopfes zum Monitor

Die Aufgabe der Kalibrierung ist die möglichst präzise Abbildung verschiedener Augenpositionen auf Koordinaten des Bildschirms. Außerdem ist während dieser Phase die Bestimmung physiologischer und systembedingter Fixierungsprobleme möglich. Wenn zu erwarten ist, dass das Experiment fehlschlagen wird, kann es somit bereits vor dem eigentlichen Start abgebrochen werden.

Einige der beim Eye-Tracking auftretenden Abweichungen kann die Kalibrierung jedoch nicht beseitigen, sondern nur in gewissem Rahmen reduzieren. Dazu zählen die hardware-spezifischen Genauigkeitsfehler der einzelnen Eye-Tracking-Verfahren (vgl. Abschnitt 2.2.1), Driftbewegungen sowie die Ungenauigkeit des Auges an sich (vgl. Abschnitt 2.1.2). Driftbewegungen sind unbewusste Abweichungen von der initialen Körper- und Kopfhaltung und treten bei längeren Experimenten auf. Eine erfolgreiche Kalibrierung ist damit keine Garantie für korrekte Blickpunktdaten. Aufgrund einer mangelhaften Driftkompensation durch beispielsweise eine Kopfkamera am Eye-Tracker haben Driftbewegungen die fehlerhafte Berechnung der Blickkoordinaten zur Folge. Dieses Problem kann nur durch eine Driftkorrektur oder eine Rekalibrierung beseitigt werden. Bei einfachen Driftkorrekturen muss die Versuchsperson einen definierten Marker auf dem Bildschirm fixieren. Aus den dabei gewonnenen Blickpunkten wird die mittlere Abweichung zu den bekannten Koordinaten des tatsächlich präsentierten Punktes berechnet, welche anschließend zur Driftkompensation genutzt werden kann.

Aufgrund der vielfältigen, oft unvorhersehbaren Einflüsse auf die ermittelten Blickdaten, ist es nicht ausreichend, nur die zurückgelieferte Blickkoordinate zu untersuchen. Viel mehr ist die Analyse aller Punkte innerhalb einer gewissen Region um diese Blickkoordinate herum notwendig.

2.2.3 Metriken

Auf dem Gebiet des Eye-Tracking bezeichnen Metriken einerseits einfache statistische Größen wie Mittelwerte (mittlere Fixierungsdauer) oder Abweichungen (räumlicher Fixierungsabstand). Auf der anderen Seite repräsentieren sie aber auch komplexe Größen, wie die Richtung aufeinander folgender Fixierungen (Richtung des Scanpfades) oder die Fläche eines alle Fixierungspunkte umschließenden Polygons (konvexe Hülle). In Tabelle 2.1 sind häufig angewandte Metriken mit jeweils einer Beispielstudie zusammengefasst. Diese stammen aus der Arbeit von JACOB und KARN [2003], in welcher Metriken aus 21 Evaluierungsstudien

analysiert und detailliert erläutert werden. Die Tabelleneinträge 1–4 beziehen sich auf reine Fixierungen und deren Dauer. Die Metriken 5–7 beziehen sich auf räumlich und zeitlich nahe beieinander liegende Fixierungen³. Die Einträge 8–13 repräsentieren seltener verwendete Maße bezüglich der räumlichen Anordnung von Fixierungen. Ein wichtiger Aspekt ist, dass verschiedene ETS's, aufgrund heutzutage noch nicht standardisierter Algorithmen zur Fixierungserkennung, unterschiedliche Daten produzieren. Aus diesem Grund können die Metriken auf verschiedenen Systemen auch verschiedene Ergebnisse liefern. Sie sind daher immer im Kontext zum jeweils verwendeten System und vor allem zu den verwendeten Parametern zur Fixierungserkennung zu sehen.

Nr.	Metrik	Beispielstudien
1	Anzahl aller Fixierungen	[BACKS und WALRATH, 1992]
2	Anzahl aller Fixierungen pro AOI	
3	Fixierungsrate [Fixierungen/s]	[KOLERS et al., 1981]
4	Fixierungsdauer (Mittel, gesamt)	
5	Blickdauer [%] pro AOI	[SVENSSON et al., 1997]
6	Mittlere Blickdauer pro AOI	
7	Anzahl der Blicke pro AOI	[HENDRICKSON, 1989]
8	Räumliche Dichte aller Fixierungen	[COWEN, 2001]
9	Fixierungssequenz (Scanpfad)	[JOSEPHSON und HOLMES, 2002]
10	Scanpfadfläche (konvexe Hülle)	[KOTVAL und GOLDBERG, 1998]
11	Richtung des Scanpfades	[AALTONEN et al., 1998]
12	Länge des Scanpfades	[GOLDBERG et al., 2002]
13	Räumlicher Fixierungsabstand	

Tabelle 2.1: Ausgewählte Metriken mit den zugehörigen Usabilitystudien, in denen sie angewendet wurden (nach JACOB und KARN [2003]).

In der Tabelle 2.1 werden die Begriffe „Area of Interest“ (AOI) sowie „Scanpfad“ verwendet. Scanpfad ist ein Synonym für eine zeitlich geordnete Liste der Fixierungspositionen. In Abbildung 2.7(a) sind die farblich gekennzeichneten Scanpfade mehrerer Personen dargestellt. AOI's sind für die Versuchsperson unsichtbare Regionen in einem zweidimensionalen Bild. Diese können auf der Basis vorhandener Blickpunktdaten automatisch generiert oder wie in Abbildung 2.7(b) manuell um wichtige Bereiche gelegt werden. AOI's gelten für die Versuchsleiter als besonders interessant und werden im Vorfeld im Bild markiert. Durch sie wird es selbst bei der Präsentation von Bildern möglich, Objekte zu definieren, deren Fixierungen auf einfache Weise gemessen werden können. Bei dreidimensionalen Reizen können die Szenenobjekte an sich als „Objects of Interest“ betrachtet werden. Die Notwendigkeit einer Markierung besteht hierbei nicht, da es geeignete Techniken der Computergraphik zur Bestimmung der betrachteten Objekte gibt. Die folgenden Abschnitte beschreiben die wichtigsten Metriken aus Tabelle 2.1 mit ihrem tieferen, psychologischen Hintergrund.

³Diese Art der Fixierung wird in der englischen Fachliteratur als *dwell* oder *gaze* bezeichnet

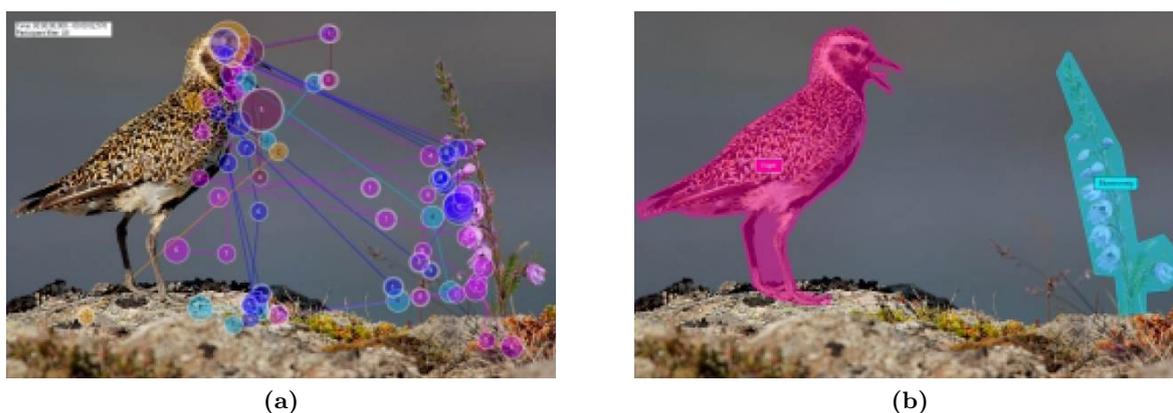


Abbildung 2.7: In Abbildung (a) sind die Scanfade mehrerer Personen dargestellt. Abbildung (b) zeigt die vor den Experimenten festgelegten Areas of Interest [RICHTER, 2008].

Die „Anzahl von Fixierungen“ in einem Experiment wird in elf der von JACOB und KARN [2003] untersuchten Studien zur Auswertung herangezogen und ist die in diesen Studien am häufigsten verwendete Metrik. Diese Größe ist immer im Verhältnis zur Experimentlänge zu sehen, da sie mit der Länge des Experimentes ebenfalls ansteigt. Laut KOTVAL und GOLDBERG [1998] sowie GOLDBERG und KOTVAL [1998] signalisieren viele kurze Fixierungen eine geringe Sucheffizienz des Betrachters, da der Fokus in geringen zeitlichen Abständen zwischen verschiedenen Elementen wechselt. Fixiert die Versuchsperson jedoch längere Zeit eine Position, so reduziert sich die Zahl aller Fixierungen und die mittlere Fixierungsdauer steigt an. Die „Dauer einer Fixierung“ wird von Forschern mit der Schwierigkeit in Verbindung gesetzt, Informationen aus dem Gesehenen zu extrahieren. Andererseits kann es sich, je nach Algorithmus zur Fixierungserkennung, bei langen Fixierungen auch um Folgebewegungen handeln, da beide Augenbewegungen ähnliche Charakteristiken aufweisen. Dieser Sachverhalt muss daher bei jeder Fixierung genau untersucht werden. Die „Fixierungsrate“ [$Fixierungen/s$] stellt, im Gegensatz zur Anzahl der Fixierungen, ein auf die Experimentlänge normiertes Maß dar. Die Normierung ermöglicht die Vergleichbarkeit der Fixierungsraten mehrerer Datenprotokolle.

Die bislang vorgestellten Maße sind Bausteine zum Entwurf komplexerer Metriken. Ein Beispiel dafür ist die „Anzahl aufeinanderfolgender Fixierungen pro AOI“. Dieses Maß besteht aus zwei Komponenten. Zum einen werden Fixierungen betrachtet, welche auf dieselbe AOI fallen. Dabei besteht die Annahme, dass die Versuchsperson sich zu diesen Zeiten auf das selbe Element konzentriert. Die Wichtigkeit jeder einzelnen Fixierung pro AOI wird zusätzlich gesteigert, wenn diese nacheinander erfolgen. Die Bedeutung von sukzessiven Fixierungen steigt daher, sobald sie die selbe AOI treffen. Darauf basierend werden die oben genannten Metriken auch auf diese speziellen Fixierungen angewandt. Es entstehen Maße wie „Gesamtdauer aufeinanderfolgender Fixierungen pro AOI“ oder „zeitlicher Anteil aufeinanderfolgender Fixierungen pro AOI“, bezogen auf die Experimentdauer“.

Die im letzten Abschnitt von Tabelle 2.1 genannten Metriken beziehen sich auf räumliche Zusammenhänge zwischen Fixierungen. Vor allem in der Analyse von Blickmustern, hervorgerufen durch die Betrachtung von Bewegtbildern, zeigt sich deren Stärke. Beispielsweise eignet sich die „Fixierungsdichte“ zur Bestimmung der räumlichen Konzentration aller Fixierungen. Andererseits bieten diese Metriken zusätzliche Informationen, wenn auf Basis der oben genannten Maße eine eindeutige Klassifizierung zu einzelnen AOI's nicht möglich ist.

Die vorgestellten Metriken werden in Eye-Tracking-Studien zur Analyse der Daten herangezogen. Dabei kann es vorkommen, dass sie an spezielle Anforderungen eines Anwendungsgebietes angepasst sowie mit klassischen Metriken kombiniert werden müssen. Erst dann ist das Ziehen aussagekräftiger Schlüsse aus den Experimenten möglich. In KOTVAL und GOLDBERG [1998] sowie GOLDBERG und KOTVAL [1998] werden viele weitere Maße detailliert beschrieben. Die Vielfalt an messbaren Daten sowie die Schwierigkeit der Analyse von komplexen Blickpunkt-daten, wird in vielen Arbeiten besonders hervorgehoben. Dazu gehören u.a. [SALVUCCI, 2000] sowie [SALVUCCI und ANDERSON, 2001].

2.3 Zusammenfassung

In diesem Kapitel wurden zunächst grundlegende Eigenschaften des visuellen Systems beschrieben. Dazu gehören neben dem Aufbau des menschlichen Auges vor allem Augenbewegungen, die während des Eye-Trackings gemessen werden. Hierbei geben die Charakteristiken der zu messenden Augenbewegungen Aufschluss darüber, welche bei einer Datenreduktion vernachlässigt werden können. Es wurde erläutert, dass ausschließlich Fixierungen das Interesse des Betrachters an einem fokussierten Objekt repräsentieren. Daher ist es naheliegend, dass diese bei der Blickmusteranalyse im Vordergrund stehen. Anhand der räumlichen Wahrnehmung sowie der Wahrnehmung von Objekten wurde verdeutlicht, dass Augenbewegungen sehr von der Erfahrung des Betrachters abhängen und in diesem Kontext auch ausgewertet werden müssen. Des Weiteren wurde aufgezeigt, dass es bei der Objektwahrnehmung zu Doppeldeutigkeiten kommen kann. Die daraus resultierenden Blickmuster sind möglicherweise schwer zu klassifizieren.

Im weiteren Verlauf wurden verschiedenste Techniken zur Messung der Blickpunkte eines Betrachters vorgestellt und miteinander verglichen. Der Schwerpunkt lag dabei auf den videobasierten Eye-Tracking-Geräten, die heute sehr häufig angewandt werden. Diese haben jedoch, im Vergleich zu traditionellen Verfahren, eine geringe Genauigkeit von nur 0.5° - 1.0° . Dies wird sich höchstwahrscheinlich nachteilig auf die genaue Bestimmung fokussierter Objekte auswirken, da einzelne Strukturen in den betrachteten Darstellungen nahe beieinander liegen. Im Weiteren Verlauf ist die Kalibrierung des ETS beschrieben. Dieser Prozess dient der Anpassung des Systems an die jeweilige Versuchsperson und die äußeren Gegebenheiten. Es wurde verdeutlicht, dass die Blickpunktbestimmung aufgrund der Vielfalt auftretender Messabweichungen mehr zur Bestimmung von Blickregionen als zur Bestimmung von Blickpunkten geeignet ist. Das Grundlagenkapitel schließt mit der Vorstellung von Metriken ab, die der automatischen Analyse eines Eye-Tracking-Protokolls dienen. Es hat sich gezeigt, dass immer mehrere Metriken zur Untersuchung der Blickpunkt-daten kombiniert werden müssen und dass spezielle Anwendungsgebiete auch die Entwicklung neuer Metriken erfordert.

Kapitel 3

Verwandte Arbeiten

Diese Arbeit beschäftigt sich speziell mit der Analyse von 2.5D-Blickpunktdaten. 2.5D bedeutet in diesem Zusammenhang, dass aufgenommene 2D-Blickpunkte mit zusätzlichen 3D-Informationen, zur Ausnutzung erweiterter Analysemöglichkeiten, kombiniert werden. Da die Bestimmung fokussierter Objekte in animierten bzw. interaktiven 3D-Darstellungen mittels Eye-Tracking eine hochgradig interdisziplinäre Aufgabe ist, sind dafür Kenntnisse aus verschiedensten Bereichen notwendig.

Intensive Literaturrecherchen in verschiedenen Stadien dieser Arbeit ergaben keine Ergebnisse zur direkten Lösung des gestellten Problems. Aus diesem Grund umfasst das vorliegende Kapitel verwandte Arbeiten zur Lösung kleinerer Teilprobleme. Zunächst werden Studien vorgestellt, die sich mit den Schwierigkeiten bei der Aufnahme von Blickpunktdaten beschäftigen. Anschließend folgen Diskussionen, wie mit diesen Daten, besonders in Bezug auf 3D-Szenen, weiter verfahren werden kann. Abschließend wird der Bezug zu dieser Arbeit mit möglichen Verfahren zur Zuordnung von 2D-Blickpunkten zu 3D-Objekten hergestellt.

3.1 Datenaufnahme und -analyse

Zunächst werden in diesem Abschnitt neuartige Verfahren zur Reduktion der beim Eye-Tracking auftretenden Fehler diskutiert, da dies eine wichtige Grundlage für die Eignung zur Weiterverarbeitung der Daten darstellt. Die vorliegende Arbeit beschäftigt sich mit Blickmustern bezüglich animierter 3D-Szenen, weswegen sich die Erörterung relevanter Studien anschließt. In diesem Teil werden auch Beispiele angeführt, bei denen die Blickmusteranalyse aufgrund unvorhergesehener Ereignisse im präsentierten Stimuli zu falschen Aussagen führen kann. Die Untersuchung von Augenbewegungen bei der Darbietung statischer 2D-Stimuli bildet für die Analyse von 2.5D-Blickpunktdaten eine wesentliche Grundlage, da sie historisch gesehen einen großen Anteil auf dem Gebiet des Eye-Tracking einnimmt. Aus diesem Grund werden abschließend Arbeiten vorgestellt, welche die Augenbewegungen von Radiologen bei der Suche anatomischer Anomalien in 2D-Schichtbildern untersuchen.

3.1.1 Erweiterte Verfahren zur Kalibrierung und Driftkorrektur

Die Kalibrierung stellt, wie in Abschnitt 2.2.2 beschrieben, eine der wichtigsten Phasen eines Eye-Tracking-Experimentes dar. Bei modernen, stationären Geräten erfolgt diese, aufgrund einer stark reduzierten Anzahl notwendiger Schritte, relativ zügig. Vor allem auf dem Kopf des Probanden befestigte Systeme haben in dieser Hinsicht jedoch große Nachteile. Da eine hohe Anzahl von die Kalibrierung beeinflussenden Parametern für jede Versuchsperson individuell eingestellt werden muss, hat dies Einfluss auf die Vergleichbarkeit der Daten zwischen

diesen Personen. Laut POMPLUN et al. [1994] implementieren viele Trackingalgorithmen die hochgradig nichtlineare Abbildung der Augenbewegungen auf Bildschirmkoordinaten nicht korrekt und produzieren dadurch einen Genauigkeitsfehler von bis zu drei Grad. Deshalb entwickelten sie ein Verfahren, welches in Echtzeit (online) oder als Nachverarbeitungsschritt (offline) den mittleren Fehler mittels einer *Parametrized Self-Organizing Map* (PSOM) auf weniger als 0.4° reduziert. PSOM's sind spezielle neuronale Netze mit folgenden Eigenschaften:

- Schnelles Lernen hochgradig nichtlinearer Funktionen
- Subpixelgenaue Ausgabe resultierender Bildschirmkoordinaten

Die Eingabeparameter der PSOM sind die unkalibrierten Blickkoordinaten des Eye-Trackers, wodurch sich das Verfahren als Ersatz für bestehende Kalibrierungsverfahren eignet. Da die meisten Eye-Tracking-Systeme (ETS's) solche Koordinaten erzeugen, macht dies das Verfahren vielfältig einsetzbar. In POMPLUN et al. [1994] wurde gezeigt, dass PSOM's die Berechnung reeller Blickkoordinaten auf der Basis einer sehr kleinen Trainingsmenge durchführen kann. Es genügt, als Trainingsmenge ein regelmäßiges Gitter mit nur 4×4 Kalibrierungsmarkern zu verwenden, welche vom Probanden nacheinander, in zufälliger Reihenfolge fixiert werden müssen. Die daraus gewonnenen Informationen reichen aus, um eine nichtlineare Korrektur der unkalibrierten Bildschirmkoordinaten zu erhalten. Weitere Details zu PSOM's finden sich in DONNER [2007], ESSIG et al. [2006] sowie ESSIG et al. [2004].

Mit Hilfe einer PSOM ist es möglich die Genauigkeit der registrierten Blickposition erheblich zu steigern und damit die Erkennung von Fixierungen eng beieinander liegender Objekte zu verbessern. Zudem können die Kalibrierungsinformationen für Onlineanwendungen genutzt oder auch zur Nachkorrektur von Blickpunktdaten gespeichert werden.

Neben der Kalibrierung stellt die Driftkorrektur einen weiteren Prozess zur Fehlerkompensation dar. Während der Betrachtung eines Stimuli ist es jedoch ungünstig, Methoden zur Driftkorrektur anzuwenden, welche den Betrachter plötzlich dazu auffordern, einen bestimmten Punkt zu fixieren. Die Ablenkung sowie die Beeinflussung des Blickverhaltens wären dabei zu groß.

HORNOF und HALVERSON [2002] haben ein Verfahren entwickelt, mit dem der durchschnittliche Fehler aller Versuchspersonen auf weniger als ein Grad reduziert werden konnte. Sie verwenden dazu den Zusammenhang zwischen der Position des Mauszeigers und dem Blickpunkt. Die Position des Mauszeigers wird als implizite Fixierungsposition genutzt, da verschiedene Studien eine Korrelation zwischen der Bewegung des Mauszeigers und des Blickpunktes nachgewiesen haben. Auch in der alltäglichen Bedienung eines graphischen Betriebssystems liegen Blickpunkt und Mauszeiger, während des Klicks auf ein Displayelement, häufig dicht beieinander. Aus diesem Grund stellt die Disparität zwischen registriertem Blickpunkt und Mausposition die jeweilige Blickpunktabweichung dar. Diese Disparität kann jedoch nur dann zur Driftkorrektur genutzt werden, wenn sich Blickpunkt und Mausposition zur selben Zeit am selben Ort befinden. Die Disparität kann in diesem Fall folgendermaßen weiterverarbeitet werden:

- Automatische Rekalibrierung bzw. Driftkorrektur bei zu großer Disparität
- Identifikation, Messung und Analyse der Blickpunktabweichungen
- Speichern aller Disparitäten und Reduktion der Blickpunktabweichung der gesamten Datenfolge nach dem Experiment

Durch eine automatische Rekalibrierung wird in den Ablauf des Experimentes direkt eingegriffen, um auftretende Abweichungen zu beseitigen. Dies kann sich störend auf die Konzentration des Probanden auswirken. Anwendung zwei und drei eignen sich dagegen gerade für eine Verarbeitung der Informationen nach dem Experiment. Die Disparität kann in der Protokolldatei vermerkt und zur Beseitigung eventueller Abweichungen nach dem Experiment genutzt werden. In den Versuchen von HORNOF und HALVERSON [2002] wurden die Probanden speziell dazu motiviert, die Zielobjekte zu fixieren, während sie auf diese klicken. Dies verleiht dem Verfahren Stabilität, ist für den praktischen Einsatz jedoch eher unerwünscht.

Der Nachteil dieser Methode zur Driftkorrektur ist die Notwendigkeit der Interaktion des Benutzers mit der Szene, was es ungeeignet für die Präsentation von Bildern oder Animationen erscheinen lässt. Für die Betrachtung interaktiver, medizinischer Visualisierungen ist das Verfahren jedoch gut einsetzbar, da keine expliziten Anweisungen, für die Driftkorrektur einen Bildschirmpunkt zu fixieren, gegeben werden müssen. Die Versuchsperson registriert diesen Vorgang also nicht, was die Ablenkung von der eigentlichen Aufgabe minimiert.

3.1.2 Auswertestrategien für Blickpunktdaten

Die Analyse von Eye-Tracking-Daten stellt aufgrund der hohen Interdisziplinarität eine besondere Herausforderung dar. Das größte Problem besteht in den nicht standardisierten Verfahren zur Erkennung von Fixierungen sowie zur Auswertung aufgezeichneter Daten [JACOB und KARN, 2003]. Diese Verfahren müssen jedoch an die jeweils im Experiment gestellten Aufgaben angepasst werden. Beispielsweise können Folgebewegungen nicht auf die selbe Art detektiert werden wie klassische Fixierungen auf statischen Stimuli. Aus diesen Gründen ist es schwierig, im Vorfeld Hypothesen bezüglich des Blickverhaltens der Versuchspersonen aufzustellen. Eine solche Strategie wird als *Top-Down-Ansatz* bezeichnet, da die Hypothesen im Vorfeld hergeleitet und durch die Experimentdaten überprüft werden. OLEG [2008] stellt eine Plattform zur Standardisierung von Aufnahme, Auswertung und Visualisierung der Blickpunktdaten vor. Sobald sich die in dieser Arbeit entwickelten Standards durchsetzen, sind vom jeweiligen Eye-Tracking-Verfahren sowie von den im Experiment gestellten Aufgaben unabhängige Analysen mit einem einzigen Werkzeug möglich.

Da dies bislang noch nicht geschehen ist, wird häufig der Weg des *Bottom-Up-Ansatzes* verfolgt. Dabei erfolgt die Erkenntnissgewinnung ohne das Aufstellen vorheriger Hypothesen ausschließlich auf der Basis aufgenommener Daten. In [RUSSO und LECLERC, 1994] betrachten die Probanden beispielsweise Videos. Zur Ergebnisanalyse werden die betrachteten Bildelemente kodiert und Muster auf Basis von deren Reihenfolge identifiziert. Im Allgemeinen erfolgt diese Identifikation durch die Betrachtung des Stimulusmaterials mit einer überlagerten Blickpunktvisualisierung. Da dies relativ zeitaufwändig ist, wird von JACOB und KARN [2003] die Verwendung der in Abschnitt 2.2.3 vorgestellten Metriken empfohlen. Dies sind abhängige Variablen, welche sich experimentell überwachen lassen.

3.1.3 Blickverhalten bei der Betrachtung dynamischer Stimuli

Eye-Tracking-Studien haben das Ziel, gemeinsame Muster aus den Blickdaten mehrerer Personen zu extrahieren. Diese Blickmuster verändern sich wesentlich mit der Art der präsentierten Stimuli. In TOSI et al. [1997] werden beispielsweise zehn Personen verschiedene Videos gezeigt und anschließend untersucht, ob Besonderheiten in den Blickdaten einzelner Personen aufgetreten sind. Es hat sich gezeigt, dass die Probanden häufig die Bildmitte fixieren und

nur während relativ statischer Szenen, wie z.B. eines Dialoges größere Sakkaden machen. Die Unterschiede in den Blickmustern einzelner Personen waren dagegen nur minimal. Interessant ist auch, dass eine Schwarz-Weiß-Darstellung der Videos keine Änderung an den Scanpfaden hervorrief. Ähnliche Experimente mit 20 Versuchspersonen und verschiedenen Videos wurden von [GOLDSTEIN et al., 2007] durchgeführt. Auch hier befanden sich die Fixierungen der Versuchspersonen relativ konstant an einer Position. Zusätzlich hat sich gezeigt, dass dieses Verhalten bei Frauen und jüngeren Menschen schwächer ausgeprägt ist, als bei Männern und älteren Personen.

Auch bei statischen Bildern befindet sich der Blick laut RENNINGER et al. [2007] häufig im Bereich des Bildschirmzentrums. Begründet wird dies damit, dass sich die Augen immer zu den Positionen bewegen, die eine Informationsmaximierung versprechen. Versuche haben gezeigt, dass Sakkaden das Auge zunächst auf die Mitte des zu untersuchenden Objektes positionieren. Von dort ausgehend erfolgt die Orientierung an Objektmerkmalen. Das heißt, dass die eigentlich fixierte Position auch knapp neben dem Objekt liegen kann. In TATLER [2007] wird aufgrund dieser Überlegungen vermutet, dass die Bildschirmmitte ebenfalls eine Region darstellt, die erstens einen guten Überblick über die Szene verschafft und zweitens einen idealen Ausgangspunkt zur weiteren Exploration darstellt. Häufige Fixierungen dieses Bereiches sind daher mit hoher Wahrscheinlichkeit zu erwarten.

Ein Anwendungsgebiet des Eye-Tracking ist die *Usability-Analyse*. Ziel von Usabilitystudien ist die Optimierung verschiedener Aspekte der Nutzbarkeit eines Systems. Dazu zählen u.a. Fehleranfälligkeit, Effizienz, Erlernbarkeit und Benutzerfreundlichkeit [NIELSEN, 1993]. Andere, klassische Eye-Tracking-Studien sind psychologischer Natur und befassen sich beispielsweise mit der Erforschung des menschlichen Leseverhaltens oder der Objektwahrnehmung. Beide Gebiete konzentrieren sich vorwiegend auf die Präsentation statischer Stimuli. Bezüglich des Blickverhaltens während der Betrachtung zeitlich veränderter Reize gibt es dagegen verhältnismäßig wenig Erkenntnisse [GOLDSTEIN et al., 2007; MUNN et al., 2008].

Algorithmen zur Erkennung von Fixierungen, welche häufig in den Softwarepaketen der ETS's integriert sind, gehen meistens von statischen Reizen aus und berücksichtigen nicht, dass bei der Betrachtung von Animationen auch Folgebewegungen auftreten. Kann der Eye-Tracker während einer Folgebewegung die Pupille, aufgrund von beispielsweise Lidschlägen nicht mehr detektieren, treten Datenverluste auf. Samples, welche zu *einer* Folgebewegung gehören, werden in zwei separate Fixierungen aufgeteilt und somit nicht korrekt klassifiziert.

MUNN et al. [2008] haben daher ein Verfahren entwickelt, welches in der Lage ist, mit den Datenverlusten umzugehen und korrekte Folgebewegungen zu erkennen. Die Grundlage für den vorgestellten Algorithmus bildet ein einfacher, auf der Bewegungsgeschwindigkeit des Auges basierender Algorithmus (*I-VT*, vgl. SALVUCCI und GOLDBERG [2000]). Der Kern des Verfahrens besteht darin, Datenverluste inmitten einer Fixierung teilweise zu ignorieren. Ausschließlich deren Dauer wird in die Berechnung der Gesamtfixierungsdauer einbezogen. Folgebewegungen werden damit nicht mehr durch Datenverluste unterbrochen, sondern als eine Fixierung behandelt.

3.1.4 Fehldeutung von Eye-Tracking-Daten

Bei der Betrachtung bestimmter Szenen (vgl. Abschnitt 2.1.5) treten auch Situationen auf, in denen die während des Versuchs aufgenommenen Blickpunkte keinem bzw. dem falschen

Objekt zugeordnet werden. Dies ist dann der Fall, wenn aufgrund der Szenencharakteristik Doppeldeutigkeiten an der aktuellen Blickposition bestehen. Solche Doppeldeutigkeiten können mehrere Ursachen haben, von denen hier zwei diskutiert werden.

Der Mensch kann über seinen peripheren Sichtbereich Bewegungen erkennen aber auch die Struktur und räumlichen Zusammenhänge mehrerer Objekte wahrnehmen. SOPHIAN und CROSBY [2000] haben durch Experimente gezeigt, dass es beim Vergleich von Proportionen zweier Objekte ausreicht, nur eines zu fixieren. Die Merkmale des anderen Objektes sind dagegen mit Hilfe der Peripherie erkennbar. Da dieses Blickverhalten also bereits beobachtet wurde, muss davon ausgegangen werden, dass es höchstwahrscheinlich auch beim Vergleich von Größenverhältnissen zwischen anatomischen Strukturen auftreten kann. In diesen Fällen würden die registrierten Blickpunkte keine Aussage darüber treffen, was die Versuchsperson tatsächlich interessiert und wahrgenommen hat.

Ein weiteres Problem, welches eng mit der in Abschnitt 2.1.5 vorgestellten Objektwahrnehmung verbunden ist, hat sich in den Versuchen von FEHD und SEIFFERT [2008] gezeigt. Dort wurden mehreren Probanden sich zufällig über den Bildschirm bewegende Ziel- sowie Ablenkungspunkte präsentiert. Wie angenommen hat sich gezeigt, dass Fixierungen häufiger auf die Ziel- als auf die Ablenkungsreize fallen. Viel interessanter jedoch ist, dass bei dreiecksähnlicher Konstellation der Punkte die Versuchspersonen dazu tendieren, die Dreiecksmitte anstatt die Zielreize zu fokussieren. Die Fixierungspositionen unterscheiden sich in dieser Zeit also von den Positionen der Objekte, welche von den Probanden verfolgt werden sollten. Es ist jedoch wahrscheinlich, dass die Zielobjekte aufgrund des peripheren Sehens im Fokus liegen. In CAVANAGH und ALVAREZ [2005] wird bestätigt, dass ein Mensch dazu in der Lage ist, seine Aufmerksamkeit auf ca. zwei bis sechs Objekte gleichzeitig zu richten. Damit ist es ihm auch ohne direkte Fixierung möglich, Objekteigenschaften wahrzunehmen und zu verarbeiten. Dies verdeutlicht, dass die von einem Eye-Tracker registrierten Blickpunkte nicht immer zu den eigentlich wahrgenommenen Objekten gehören müssen. Der Blick einer Person kann also von den elementaren Objekten abschweifen, sobald diese eine komplexe Struktur formen.

3.1.5 Blickmuster von Radiologen bei Röntgenbildern

Als Grundlage zur Blickmusteranalyse auf medizinisch gearteten 3D-Stimuli werden nun Arbeiten zur Wahrnehmung von 2D-Stimuli vorgestellt, welche durch medizinische Bildgebung entstanden sind. Dazu zählen Studien, in denen das Blickverhalten von Radiologen, bei der Analyse von Röntgenbildern, mittels Eye-Tracking aufgenommen und untersucht wird.

Nach NODINE und KUNDEL [1987] stellen die Suche bestimmter Bildregionen in Abbildung 3.1(a) sowie die Suche und Einschätzung krankhaften Gewebes in Röntgenbildern ähnliche Anforderungen an die visuelle Wahrnehmung. Diese Aussage wird in KRUPINSKI [2000] bestätigt. Beispiele von Röntgenbildern der menschlichen Lunge, mit überlagerter Blickpunktvisualisierung, sind in Abbildung 3.1(b) und Abbildung 3.1(c) dargestellt. Bei der visuellen Suche in Abbildung 3.1(a) sowie bei der Analyse eines Röntgenbildes müssen Vergleiche zwischen verschiedenen Bildregionen durchgeführt werden, um die betreffende Person bzw. das beschädigte Gewebe eindeutig identifizieren zu können. Ähnliche Effekte sind auch bei der Betrachtung dreidimensionaler Darstellungen denkbar, da beim Großteil der möglichen Kamerastellungen semi-transparente oder opake Verdeckungen der zu untersuchenden Strukturen auftreten. In der Radiologie wird nach NODINE und KUNDEL [1987] die Erkennung eines Tumors, aufgrund von Verdeckungen durch andere anatomische Strukturen, deutlich erschwert.

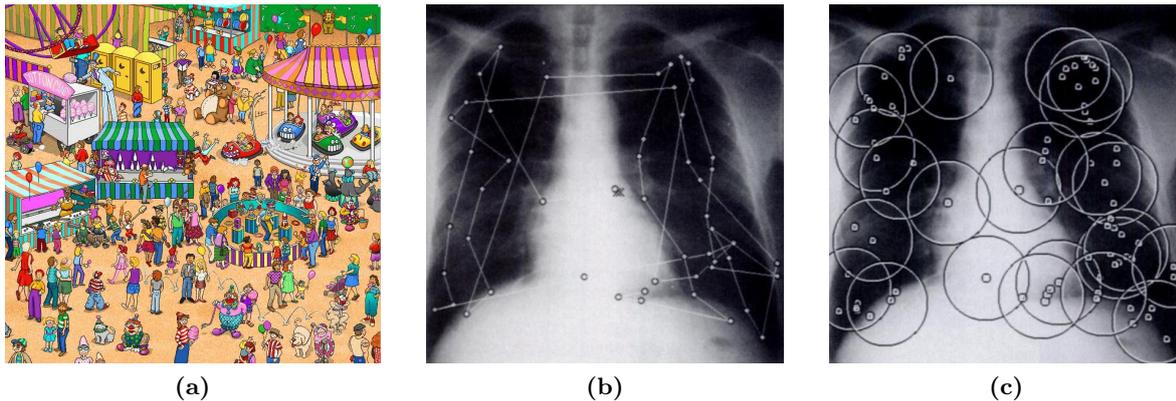


Abbildung 3.1: Abbildung (a) ist ein Beispiel für eine visuelle Suchaufgabe, in der es die Person *Waldo* zu finden gilt [ECE]. Dies ist ein Junge mit Brille, roter Mütze sowie einem rot-weiß gestreiftem Pullover. Diesbezüglich sind viele Ablenkungsreize im Bild vorhanden, was die Suchaufgabe erschwert. Abbildungen (b) und (c) zeigen Röntgenbilder mit überlagerten Blickpunkten [NODINE und KUNDEL, 1987]. In Abbildung (b) sind Fixierungen, verbunden durch einen Scanpfad dargestellt. Abbildung (c) dagegen visualisiert diese Fixierungen in Form von Kreisen. Jeder Kreis beschreibt eine Gruppe (*Cluster*) von räumlich nahe beieinander liegenden Fixierungen.

Solche Verdeckungen können oft durch eine intensive, visuelle Suche von der Wahrnehmung umgangen werden. Je gründlicher diese visuelle Suche ausgeführt wird, desto besser erfolgt ebenfalls die Erkennung typischer Merkmale des beschädigten Gewebes. Erfahrene Radiologen haben in dieser Hinsicht den Vorteil, betroffene Regionen schneller zu finden, da sie eine gute Vorstellung von diesen Merkmalen besitzen.

In einigen Fällen erfolgt die Erkennung der Läsionen jedoch nicht korrekt. Gesundes Gewebe wird fälschlicherweise als krank klassifiziert oder krankes Gewebe als gesund. Solche Fälle können anhand der Länge sowie der Häufigkeit ausgeführter Fixierungen, durch die Analyse von Blickpunktdaten registriert werden. KRUPINSKI [2000] teilt die Ursachen solcher Erkennungsfehler in drei Kategorien:

- Suchfehler
- Wahrnehmungsfehler
- Entscheidungsfehler

Suchfehler repräsentieren Läsionen, die aufgrund einer zu undeutlichen Darstellung im Bild, vom Menschen, nicht wahrgenommen und damit auch nicht fixiert werden. Wahrnehmungsfehler treten immer dann auf, wenn der Proband kurze bzw. wenige Fixierungen auf entsprechende Bereiche ausgeführt hat. Diese reichen jedoch nicht zur Erkennung der Merkmale des geschädigten Gewebes aus. Bei Entscheidungsfehlern wurde der entsprechende Bildbereich intensiv untersucht, aber trotzdem nicht korrekt klassifiziert.

Ein weiteres Phänomen wird in KRUPINSKI als *Satisfaction of Search* bezeichnet. Es beschreibt, dass die Entdeckung deutlich erkennbarer Läsionen mit der Vernachlässigung weniger gut erkennbarer Läsionen einhergeht. Dies könnte eine mögliche Quelle für Suchfehler repräsentieren, da unscheinbare Bereiche somit schneller übergangen werden. Dieses Verhalten wurde experimentell mit Radiogrammen verschiedener Anwendungsfälle nachgewiesen.

Die Beurteilung von Röntgenbildern erfolgt nach einem ähnlichem Schema, wie dem in Abschnitt 2.1.3 beschriebenen Modell zur Aufmerksamkeitsverteilung. Nachdem verdächtige Bereiche mittels peripherem Sehen identifiziert wurden, werden diese durch foveales Sehen weiter analysiert. Potentielle Läsionen werden dabei durch eine Reihe nahe beieinander liegender Fixierungen (Cluster) untersucht. In Abbildung 3.1(b) sind solche Fixierungen eines Röntgenbildes der menschlichen Lunge als helle Punkte dargestellt. Ein Scanpfad gibt die Reihenfolge der Blickpunkte wieder. Abbildung 3.1(c) stellt weiterhin räumlich benachbarte Fixierungen (Cluster) Form von Kreisen mit einem Radius von 2.5° dar. NODINE und KUNDEL [1987] sowie KRUPINSKI [2000] vermuten hinter jedem Cluster eine bewusste oder unbewusste Entscheidung, über den Infiltrationsgrad der entsprechenden Region. Längere Fixierungen deuten dabei auf die Tendenz des betrachtenden Arztes in Richtung Tumorbefall hin.

Die Länge und Häufigkeit von Fixierungen sind zudem stark von der Erfahrung des betrachtenden Arztes abhängig. Experimentelle Ergebnisse aus @KRUPINSKI belegen, dass Experten viel effizientere Suchstrategien verfolgen und betroffene Gebiete weitaus schneller finden als unerfahrene Radiologen. Des weiteren fixieren Experten seltener Regionen, die betroffenen Bereichen zwar ähnlich aber dennoch gesund sind. Auch außerhalb medizinischer Anwendungsgebiete wurde gezeigt, wie sehr die persönliche Erfahrung das menschliche Blickverhalten beeinflusst. Die experimentell gewonnenen Ergebnisse von AALTONEN et al. [1998] zeigen, dass Nutzer mit weniger als fünf Monaten Erfahrung in der Bedienung graphischer Benutzerschnittstellen (GUI) weitaus ineffizienter bei der Suche bestimmter Einträge aus Auswahllisten vorgehen, als Nutzer die schon mehr als sieben Jahre mit GUI's arbeiten.

3.2 Weiterverarbeitung von Eye-Tracking-Daten

Über die Jahre wurden im Bereich des Eye-Tracking eine Vielzahl von Techniken entwickelt. Diese werden u.a. in dessen Anwendungsmöglichkeiten deutlich. Der vorliegende Abschnitt beschreibt daher Techniken zur Interpretation und Weiterverarbeitung der Blickpunktdaten an einem der vorliegenden Problemstellung ähnlichen Beispiel.

In LU et al. [2006] werden Eye-Tracking-Daten zur Bestimmung fixierter Volumenelemente (*Voxel*) einer Volumenvisualisierung genutzt. Ziel dieser Arbeit ist die automatische Erzeugung von Parametern zum Rendern der Szene. Dazu gehören initiale Kamerapositionen oder auch Transferfunktionen zur Hervorhebung der ermittelten Voxel. Zunächst exploriert der Betrachter die Visualisierung mit gängigen Mitteln. Er ist dabei in der Lage, durch beispielsweise das Ausblenden von Objektteilen (*Clipping*), Einsicht ins Innere der Strukturen zu erlangen. Die Blickpunktaufzeichnung wird temporär eingestellt, sobald der Proband mit dem zur Verfügung stehenden Kontrollfenster zur Modifikation des Volumens interagiert, da diese Fixierungen nicht zur Generierung der Visualisierungsparameter genutzt werden sollen.

$$G(v) = \begin{cases} 60.0, & v \leq 0.825^\circ/s \\ 57.69 - 27.78 \log_{10}(v), & 118.3 \geq v > 0.825^\circ/s \\ 0.1, & v > 118.3^\circ/s \end{cases} \quad (3.1)$$

$$M(e) = \begin{cases} 1.0, & e \leq 5.79^\circ \\ 7.49/(0.3e + 1)^2, & e > 5.79^\circ \end{cases} \quad (3.2)$$

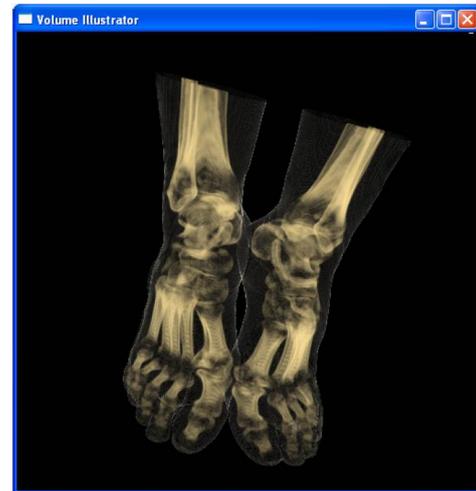
$$H(v, e) = G(v)M(e) \quad (3.3)$$

Ausgehend von den erfassten 2D-Blickpunkten auf dem Monitor erfolgt die Ermittlung fixierter Voxel im Raum. Um den Blickpunktdateien eine Wichtigkeit zuzuweisen, wird zunächst die von REDDY [2001] vorgestellte Gleichung 3.3 zur Approximation der menschlichen Sehschärfe (*visual acuity*) genutzt. Das Modell beschreibt mittels der Faktoren Kontrastsensitivität $G(v)$ und Sensitivitätsabfall $M(e)$, wie sich die Sehschärfe des Auges $H(v, e)$ von der Fovea bis in den peripheren Bereich der Netzhaut hin verändert. Die einzelnen Parameter sind in REDDY [2001] dokumentiert.

Zur Berechnung des 3D-Punktes, welcher von der Versuchsperson fixiert wurde, gehen die Autoren von einem gleichmäßig rotierenden Volumen aus. Die Blickpunkte ergeben bei orthographischer Projektion, zusammen mit dem Sehschärfemodell, zylinderförmige Teilvolumina. Da der Mensch jedoch gewisse Punkte im Raum fixiert und nicht an dem gesamten Teilvolumen interessiert ist, wird der 3D-Schnittpunkt zweier, konsekutiver Blickpunktstrahlen berechnet. Da diese Strahlen in der Regel windschief sind, ist mit dem Schnittpunkt die Position im Raum gemeint, an welcher der Abstand beider Strahlen minimal wird. Der Einfachheit halber wird diese Position jedoch im weiteren Verlauf als Schnittpunkt bezeichnet.



(a)



(b)

Abbildung 3.2: Volumenvisualisierungen der Füße auf Basis automatisch generierter Renderingparameter mittels Eye-Tracking-Daten [LU et al., 2006]. In Abbildung (a) lag der Fokus des Betrachters häufiger auf der Haut, während in Abbildung (b) eine höhere Aufmerksamkeitsverteilung für die Knochen gemessen wurde.

Das 2D-Sehschärfemodell muss anschließend dahingehend abgeändert werden, dass alle Voxel um den 3D-Schnittpunkt eine Wichtigkeit erhalten, welche mit zunehmendem Abstand von

diesem Punkt sinkt. Eben diese Wichtigkeiten dienen abschließend zur Berechnung geeigneter Darstellungsparameter zum Rendern der Visualisierungen aus den Volumendaten, wie sie in Abbildung 3.2 dargestellt sind. Diese Visualisierungen stellen die für den Betrachter relevanten Voxel so dar, dass entsprechende Strukturen besonders gut sichtbar sind.

3.3 Objektbestimmung aus Bildschirmkoordinaten

In diesem Abschnitt wird auf Techniken zur Bestimmung der sich an einem gegebenen Bildschirmpunkt befindlichen Objekte eingegangen. Der erste Teil konzentriert sich auf den Vergleich von Laufzeiten drei verschiedener objektbasierter Verfahren. Der Fokus des zweiten Teils liegt auf der Vorstellung bildbasierter Ansätze sowie der Berechnung zweier Metriken, die sich bei der Identifizierung fokussierter Objekte als nützlich erweisen könnten.

3.3.1 Objektbasierter Ansatz

Zu einem gegebenen Punkt auf der Bildschirmoberfläche kann das sich an dieser Stelle befindliche Objekt mit Hilfe eines objektbasierten Ansatzes ermittelt werden. Dazu muss der 2D-Punkt zunächst in die 3D-Szene projiziert werden. Durch eine solche Projektion entsteht ein Strahl von der Projektionsebene der Kamera bis zur hinteren Clippingebene. Diese beiden Ebenen werden mit der Kamera definiert und verhindern die Darstellung sehr naher bzw. sehr weit entfernter Objekte. Anschließend erfolgt ein Schnittest des Strahls mit der Szenengeometrie. Dies ermöglicht die Bestimmung exakter Punkte auf der Oberfläche geschnittener Objekte, weshalb das Verfahren im weiteren Verlauf auch als objektbasiert bezeichnet wird.

Eine solche Schnittpunktberechnung ist Teil von Selektionsverfahren der Computergrafik. Solche Verfahren dienen der Interaktion mit zwei- bzw. dreidimensionalen Darstellungen und im Speziellen der Selektion und anschließenden Hervorhebung eines Szenenobjektes. Zusätzlich zur Bestimmung der 3D-Schnittpunkte liefern ausgereifte Selektions-Algorithmen weitere Informationen wie z.B. das zugehörige Objekt oder das den Schnittpunkt umschließende Polygon. Hierdurch wird auch die Bestimmung getroffener Strukturen aus den vom Eye-Tracker aufgenommenen Blickpunkten möglich. Die hohe Anzahl durch Eye-Tracking aufgezeichneter Blickpunkte sowie die hohe Komplexität medizinischer Visualisierungen erfordern dabei eine besonders performante Schnittpunktberechnung.

STORK [2000] stellt einen Algorithmus für 3D-Eingabegeräte vor, der die Selektion und damit die Schnittberechnungen in Echtzeit durchführen kann. Zudem vergleichen die Autoren das Verfahren mit den in der Graphikbibliothek OPEN INVENTOR sowie im CAD-Programm „ACIS“ integrierten Algorithmen. Abbildung 3.3(a) zeigt, dass alle Verfahren außer OPEN INVENTOR nach dem ersten Durchlauf an Geschwindigkeit gewinnen. Die beiden anderen Algorithmen bauen während des ersten Durchlaufs spezielle, geometrische Datenstrukturen zur Beschleunigung sich anschließender Schnittpunktberechnungen auf. In Abbildung 3.3(b) ist deutlich zu erkennen, dass OPEN INVENTOR auf kleinen Modellen eine schnelle Schnittpunktberechnung ausführen kann. Die Laufzeit des in der Graphikbibliothek implementierten Algorithmus steigt jedoch mit der Szenenkomplexität stark an.

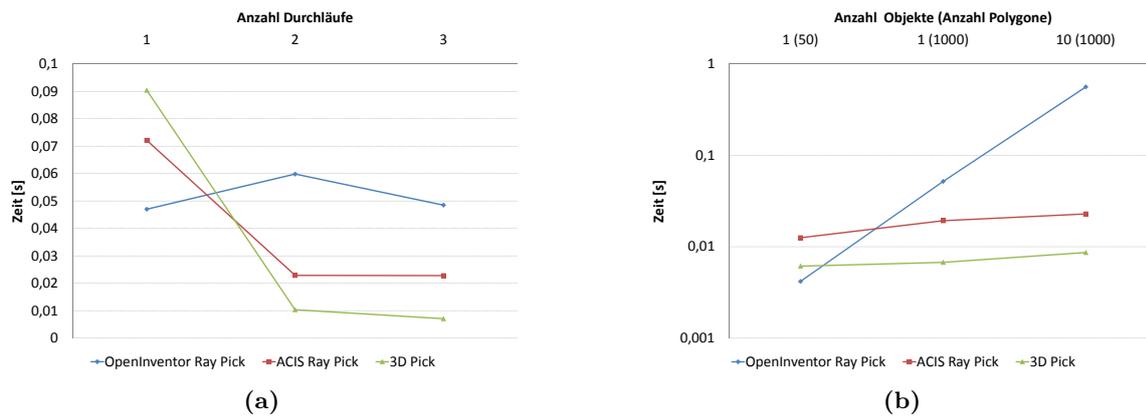


Abbildung 3.3: Vergleich der Laufzeiten von drei Selektionsalgorithmen. In Abbildung (a) werden Laufzeit und Anzahl hintereinander ausgeführter Selektionen gegenübergestellt. Abbildung (b) veranschaulicht dagegen die Laufzeit in Abhängigkeit der Szenenkomplexität nach [STORK, 2000].

3.3.2 Bildbasierter Ansatz

Die Bestimmung betrachteter Strukturen in einer Szene mit n Objekten kann nicht nur über die Schnittpunkte der Szenenobjekte mit den Blickpunktstrahlen erfolgen. Ein bildbasierter Ansatz besteht darin, nach Änderung der Kameraposition die Bildschirmprojektion aller Szenenobjekte auf Schnitt mit dem Blickpunkt zu untersuchen. In NEUGEBAUER [2006] wird ein Verfahren zur Sichtbarkeitsbestimmung beschrieben, welches in abgewandelter Form auch zur Bestimmung fokussierter Objekte eingesetzt werden kann. Teil dessen ist die Bestimmung der Metriken „Anzahl sichtbarer Pixel pro Objekt“ P_{si} ($i \in 1 \dots n$) und „Gesamtzahl aller Objektpixel bei verdeckungsfreier Projektion auf die Bildebene der Kamera“ P_{gi} , ($i \in 1 \dots n$). Da große Objekte häufiger im Sichtfeld des Betrachters liegen als kleine, diese aber eher als Kontextinformation dienen, können beide Maße dazu genutzt werden, großen Strukturen eine niedrigere Fixierungswahrscheinlichkeit zuzuordnen.

Die Bestimmung von P_{si} und P_{gi} erfolgt für jede Kameraposition erneut, da sich die Sichtverhältnisse mit der Änderung der Kameraparameter ebenfalls verändern. Im weiteren Verlauf wird der Algorithmus für eine gegebene Kameraposition erklärt. Nach der Positionierung der Kamera erfolgt das separate Rendern jedes der n Szenenobjekte und die Erzeugung von $n + 1$ z-Buffer-Bildern. Es wird ein z-Buffer-Bild für die gesamte Szene mit allen Objekten (Gesamtbuffer) sowie für jedes einzelne Objekt generiert. Ein z-Buffer ist mit der dazu gehörigen Visualisierung in Abbildung 3.4 dargestellt und gibt an, welche relative Tiefe die Szenenobjekte an jeder Pixelposition haben. Aus dem Gesamtbuffer ist zu erkennen, an welcher Bildposition sich Objekte befinden. Er ist notwendig, um leere Bildregionen nicht in allen n z-Buffern auf das Vorhandensein eines Objektes testen zu müssen. Die Ermittlung von P_{gi} kann bereits an dieser Stelle, durch das Auszählen aller Pixel der jeweiligen z-Buffer, die nicht der Hintergrundfarbe entsprechen, erfolgen.

Für die Ermittlung von P_{si} ist es erforderlich, die Werte aller z-Buffer für eine Pixelposition auszulesen und diese zu sortieren. Das Objekt, welches für die aktuell betrachtete Pixelposition den niedrigsten z-Wert besitzt, ist der Kamera am nächsten und somit an dieser Stelle sichtbar. Wird der Prozess für jede Pixelposition wiederholt, an der sich ein Objekt befindet

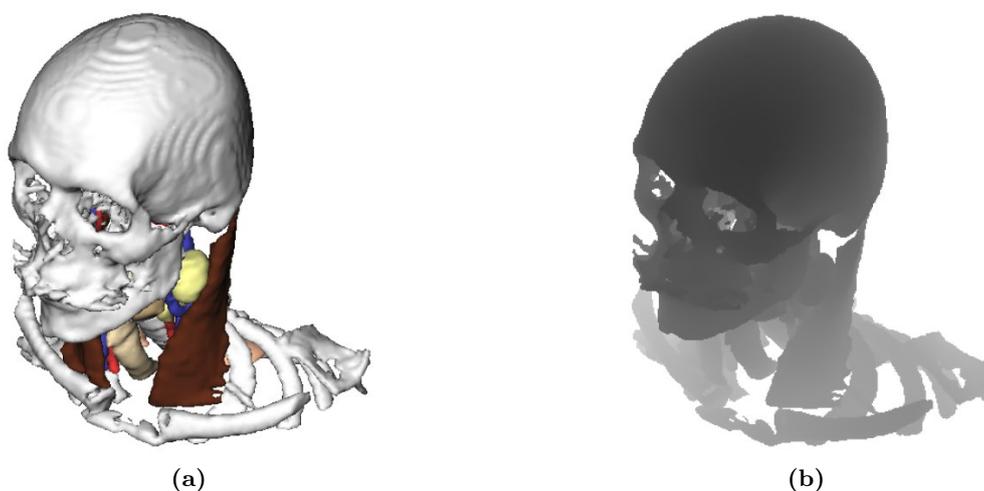


Abbildung 3.4: Darstellung eines Bildes, wie es der Nutzer sieht (a) und des dazugehörigen z-Buffer-Bildes (b) [NEUGEBAUER, 2006]. Es ist zu erkennen, dass der Kamera nahe Oberflächenregionen dunkler im z-Buffer dargestellt sind, während ferne Bereiche einen helleren Farbton besitzen.

und P_{si} inkrementiert, wenn das entsprechende Objekt den kleinsten z -Wert besitzt, so resultiert daraus die Anzahl sichtbarer Pixel für die aktuelle Kameraposition und jedes der n Objekte.

Es ist denkbar, ein ähnliches Verfahren für die Objektidentifizierung auf Bildebene zu verwenden. Dies hat den Vorteil, dass aufwändige Schnittpunktberechnungen zwischen 3D-Blickstrahlen und Objekten entfallen. Daraus resultiert jedoch auch der Nachteil, dass eben diese Schnittpunkte im dreidimensionalen Raum nicht mehr bei der Auswertung der Blickpunkte zur Verfügung stehen. Ein weiterer Nachteil resultiert zudem aus dem zusätzlichen Rendern der $n + 1$ Bilder nach jeder Kameraänderung, was sich besonders bei einem hybriden Verfahren (objekt- und bildbasiert) nachteilig auf die Laufzeit auswirkt.

3.4 Zusammenfassung

In diesem Kapitel wurden einerseits konkrete Techniken vorgestellt. Zum anderen lag der Fokus aber auch auf wichtigen Erkenntnissen bezüglich unbewusster Blickmusteränderungen. Diese wurden im Zusammenhang mit unvorhersehbaren Objektkonstellationen, Erkennungsfehlern sowie dem Erfahrungsgrad der Versuchsperson beschrieben. Unregelmäßigkeiten im Blickverhalten aufgrund von Erkennungsfehlern bzw. der Erfahrung sind erkennbar, da sich beide Faktoren genau bestimmen lassen. Daher sollte es einerseits möglich sein, die Güte einer Visualisierung auf Basis des Blickverhaltens zu bestimmen. Andererseits muss unbedingt die Erfahrung der Versuchspersonen in die Ergebnisanalyse einfließen, sobald gleichzeitig Spezialisten und medizinisch unversierte Personen an den Versuchen beteiligt sind.

In Tabelle 3.1 werden die wichtigsten, in diesem Kapitel vorgestellten Techniken, bezüglich ihrer Anwendbarkeit bei verschiedenen Präsentationsformen der dargebotenen Stimuli, verglichen. Einige dieser Techniken haben sich als nützlich zur Entwicklung eines Verfahrens für die Zuordnung von Fixierungen zu Szenenobjekten ergeben. Entsprechende Sachverhalte werden nachfolgend erläutert.

Technik	geeignet für Stimuli der Art		
	interaktiv	animiert	Bild
PSOM	✓	✓	✓
Mauscursor zur Driftkorrektur	✓		
Erkennung von Folgebewegungen	✓	✓	
Sehschärfemodell	✓	✓	✓
Blickstrahl-Kreuzungspunkte		✓	
Strahl-Objekt-Schnittpunkte	✓	✓	✓
Sichtbarkeitsbestimmung	✓	✓	✓

Tabelle 3.1: Vergleich herausgearbeiteter Techniken nach ihrer Anwendbarkeit bezüglich der Art des präsentierten Stimulus. Mit „Bild“ werden in dieser Arbeit statische Visualisierungen bezeichnet und keine Bitmaps im eigentlichen Sinne.

Es wurde erläutert, dass durch die Kalibrierung mittels einer PSOM Fixierungen dicht beieinander liegender Objekte besser voneinander unterschieden werden können. Aus diesem Grund bietet sich das Verfahren für die Kalibrierung der Eye-Tracking-Hardware vor der Präsentation einer medizinischen Visualisierung an. Weiterhin ist eine echtzeitfähige Korrektur der Daten möglich, um beispielsweise die Blickinformation zur Interaktion zu nutzen. Andererseits kann die Kalibrierungsinformation aber auch zur Korrektur der Bildschirmkoordinaten nach dem Experiment verwendet werden.

Zur Kompensation des Genauigkeitsfehlers und zur Interpretation der Blickpunkte kann auch ein Wahrnehmungsorientiertes Modell dienen. Es wurde ein solches Modell vorgestellt, das um den vom Eye-Tracker aufgenommenen Blickpunkt einen kreisförmigen Bereich annimmt. Die Wichtigkeit sinkt dabei mit größerem Abstand zum Blickpunkt. Dieses Modell stellt im Gegensatz zur Weiterverarbeitung einzelner Blickpunkte einen natürlichen, der Funktionsweise des Auges entsprechenden Ansatz dar. Der Genauigkeitsfehler eines ETS tritt gleichwahrscheinlich in einem gewissen Radius auf und das Sehschärfemodell stellt diesen Bereich mit abnehmender Wichtigkeit dar. Aus diesen Gründen muss vor der Anwendung des Modells zur Genauigkeitskompensation geprüft werden, inwiefern sich der Ansatz für diese Zwecke eignet.

Eine andere Art der Ungenauigkeit wird durch die sogenannte Driftkorrektur reduziert. Verfahren, die diese Aufgabe erledigen, sind stark vom jeweiligen Experiment abhängig. Die Driftkorrektur durch unbewusste Fixierungen des Mauszeigers ist zudem nur für interaktive Anwendungen geeignet. Nachteilig ist weiterhin, dass die Probanden in den zugrundeliegenden Experimenten speziell motiviert wurden, was das Blickverhalten beeinflusst. In realen Experimenten zur Evaluierung einer Visualisierung sollte das Blickverhalten jedoch natürlich sein, um valide Ergebnisse sicherzustellen.

In einem anderen Abschnitt dieses Kapitels wurden der Top-Down- sowie Bottom-Up-Ansatz zur Auswertung von Eye-Tracking-Protokollen vorgestellt. Während der Top-Down-Ansatz üblich für die meisten Experimente ist, wird speziell im Bereich des Eye-Tracking der Bottom-Up-Ansatz verwendet. Da eine gründliche Analyse von Blickpunktdaten jedoch relativ zeitaufwändig ist, wird die Verwendung von Metriken empfohlen [JACOB und KARN, 2003]. Mit deren Hilfe werden zudem Experimentergebnisse verschiedener Personen vergleichbar.

Bislang wurden relativ wenige Erkenntnisse bezüglich des Blickverhaltens bei dynamischen Stimuli gewonnen. Gerade bei der Erkennung von Folgebewegungen aus den Blickpunktdaten wirkt sich dies nachteilig aus. Aus diesem Grund ist die Integration eines Verfahrens empfehlenswert, welches auch bei der Betrachtung bewegter Reize auftretende Blickmuster erkennen kann.

Da die Analyse radiologischer Schichtbilder eine visuelle Suche darstellt, können daraus Konsequenzen für die Gestaltung geeigneter Evaluierungsversuche gezogen werden. Ist z.B. eine reale Halsvisualisierung zu komplex für die Evaluierung eines möglichen Verfahrens zur Objektidentifizierung können einfache Szenarien mit geeigneten Suchaufgaben entwickelt werden.

Des Weiteren ist auf Erkennungsfehler eingegangen worden. Abhängig von den Visualisierungsparametern bzw. der angewandten Techniken können ähnliche Effekte auch bei medizinischen Visualisierungen auftreten. Vor allem bei der Betrachtung einer statischen Darstellung sollte dieser Effekt besonders ausgeprägt sein. Anatomische Strukturen sind in diesem Fall aus nur einem Blickwinkel sichtbar und semi-transparent verdeckte Objekte können teilweise nicht korrekt erkannt werden. Solche Erkennungsfehler und damit verbundene Unregelmäßigkeiten im Blickverhalten sind durch genaue Analyse des betrachteten Stimulus sowie der dazugehörigen Eye-Tracking-Daten identifizierbar und lassen Rückschlüsse auf die Darstellungseigenschaften des Stimulusmaterials zu.

Mit LU et al. [2006] wurde eine Möglichkeit zur Weiterverarbeitung von Eye-Tracking-Daten vorgestellt. Die Bestimmung von Fixationspunkten im Raum, könnte besonders bei der Erkennung von Fixierungen transparent verdeckter Objekte vorteilhaft sein. Dadurch ist es möglich, von vornherein Objekte vor bzw. hinter dem fixierten Punkt auszuschließen. In Abschnitt 2.2 wurde dargestellt, dass selbst moderne ETS's einen relativ hohen Genauigkeitsfehler haben. Da die Bestimmung der 3D-Schnittpunkte dadurch beeinflusst wird, muss geprüft werden ob das Verfahren eine akzeptable Genauigkeit aufweist. Nachteilig ist ebenfalls, dass dieses Vorgehen nur für animierte Szenen geeignet ist. Ein weiterer Kritikpunkt ist die Ablenkung des Probanden aufgrund der notwendigen Modifikation des Volumens während der Versuche. Es kann nicht ausgeschlossen werden, dass eine solche Interaktion Einfluss auf das weitere Blickverhalten der Versuchsperson hat.

Das Clustering der Fixierungen wurde anhand von Röntgenbildern beschrieben. Von diesen Erkenntnissen ausgehend, kann vermutet werden, dass auch in 3D-Darstellungen Abnormitäten zunächst genau untersucht werden und Fixierungen, ähnlich zu Radiogrammen, häufiger in Clustern auftreten. Dies stellt eine erste Interpretationsmöglichkeit von Fixierungen bei 3D-Visualisierungen dar. Bildbasierte Verfahren reichen hierfür nicht aus, da sie keine Punkte im 3D-Raum bestimmen können. Zur Auswertung solcher Blickmuster ist daher ein Verfahren notwendig, welches Schnittpunkte zwischen Blickstrahlen und Objekten berechnet und somit potentiell betrachtete Objekte identifizieren kann. Algorithmen zur Schnittpunktberechnung, die immer die gesamte Geometrie untersuchen, haben jedoch eine hohe Rechenzeit, welche aber mittels verschiedener Verfahren reduzierbar ist.

Bildbasierte Verfahren sind zwar weniger gut geeignet zur Identifizierung von Objekten, ermöglichen jedoch die Ermittlung folgender Maße zur Charakterisierung von Größe sowie gegenseitigen Verdeckungen der Szenenobjekte:

- Gesamtanzahl der Pixel eines Objektes bei verdeckungsfreier Projektion
- Gesamtanzahl der sichtbaren Pixel eines Objektes
- Verdeckungsgrad eines Objektes

Kapitel 4

Konzept

In diesem Kapitel wird ein Verfahren zur automatischen Ermittlung fokussierter Objekte einer dreidimensionalen, medizinischen Visualisierung entwickelt. Der Schwerpunkt liegt auf der Untersuchung gleichmäßig animierter Szenen, da sich während der Betrachtung auftretende Folgebewegungen zur Beseitigung durch Semitransparenzen erzeugter Doppeldeutigkeiten eignen. Neben Animationen können Visualisierungen jedoch auch interaktiv exploriert oder als Bild präsentiert werden. Animationen stellen aus diesem Grund nur einen Teil des denkbaren Anwendungsspektrums dar. Für die zur Evaluierung des Verfahrens durchgeführten Versuche gelten weiterhin spezielle Regeln. Die Erläuterung der beschriebenen Aspekte erfolgt in der sich anschließenden, anwendungsorientierten Anforderungsanalyse. Danach wird ein an die Anforderungen angepasstes Konzept entwickelt, mit dessen Hilfe sich die Bestimmung fokussierter Strukturen, mittels 2.5D-Eye-Tracking, durchführen lässt.

4.1 Anforderungsanalyse

Es gibt eine Vielzahl von Gesichtspunkten, die bei der Konstruktion eines Verfahrens zur Identifikation fokussierter Objekte zu beachten sind. Das Einbeziehen von Blickpunktdaten bringt dabei verschiedene Schwierigkeiten mit sich. Aus diesem Grund werden nachfolgend die Anforderungen an ein solches Verfahren anhand konkreter Anwendungsbeispiele medizinischer Visualisierungen beschrieben. Die Erläuterung der Rahmenbedingungen für die Evaluierungsexperimente beendet diesen Abschnitt.

4.1.1 Merkmalsanalyse medizinischer Visualisierungen

Dieser Abschnitt erläutert die Besonderheiten von Visualisierungen unterschiedlicher Datensätze aus verschiedenen sowie medizinischen Anwendungsgebieten. Dabei werden folgende Gesichtspunkte untersucht, die ebenfalls im Rahmen des menschlichen Blickverhaltens eine Bedeutung haben:

- Räumliche Verhältnisse anatomischer Strukturen
- Komplexität der dargestellten Szene
- Materialeigenschaften
- Genauigkeitsbetrachtungen bezüglich des Eye-Tracking-Systems
- Präsentationsart der Visualisierung

Bezüglich der räumlichen Verhältnisse dargestellter Strukturen wird auf deren Größe, Form und Position eingegangen. Die genannten Objektmerkmale können vom Menschen über den peripheren Sichtbereich wahrgenommen und daher nach dem Aufmerksamkeitsmodell aus Abschnitt 2.1.3 im Gehirn als „interessante Regionen“ registriert werden. Sakkaden richten das Auge kurz darauf auf solche Regionen aus, um diese mit hoher Auflösung genauer analysieren zu können. Das Blickverhalten ist durch diese Objektmerkmale beeinflussbar. Des Weiteren ist gerade die Fixierung einer nahe bei anderen Objekten liegenden Struktur, aufgrund der in Abschnitt 2.2 beschriebenen Ungenauigkeiten, schwierig von der Fixierung der umliegenden Strukturen zu unterscheiden. Zudem steht der Anstieg der Komplexität medizinischer Datensätze, für objektbasierte Verfahren, in Verbindung mit einem Anstieg der Laufzeit.

Einen Schwerpunkt dieser Arbeit bildet die Auflösung von Doppeldeutigkeiten, verursacht durch semi-transparente Verdeckungen, weswegen die Materialeigenschaften der dargestellten Strukturen betrachtet werden müssen. Neben der Analyse von Transparenzen ist auch die Untersuchung anderer Merkmale, wie der Farbgebung wichtig, da Farbe eine Objekteigenschaft ist, die laut TREISMAN und GELADE [1980] präattentiv wahrgenommen werden kann. Sie hat somit, bei entsprechend auffällig gefärbten Objekten, Einfluss auf das Blickverhalten.

Nach der Erläuterung von Möglichkeiten zur Reduktion von Genauigkeitsabweichungen bei der Blickpunktaufnahme, werden Möglichkeiten zur Präsentation einer Visualisierung aufgezeigt. Es bieten sich hierbei die Möglichkeiten der animierten und der interaktiven Darstellung sowie der Präsentation als statisches Bild. Alle drei Formen sind bei der Präsentation medizinischer Visualisierungen denkbar. Sie wirken sich ebenfalls unterschiedlich auf das Blickverhalten aus und haben verschiedene Vor- und Nachteile bei der Objektidentifikation.

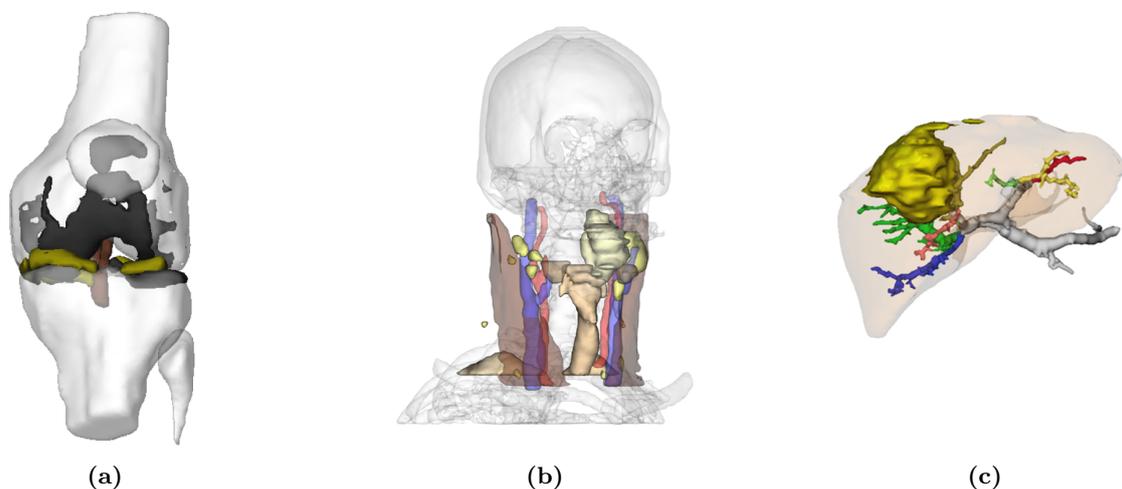


Abbildung 4.1: Vergleich dreier Visualisierungen von Datensätzen des Knies (a), der Hals-Kopf-Region (b) und der Leber (c) [NEUGEBAUER, 2006].

Räumliche Verhältnisse und Komplexität

Die untersuchten Datensätze stammen aus drei verschiedenen Anwendungsgebieten, um ein breites Spektrum an möglichen Objektkombinationen abzudecken. Die Wahl fiel auf Darstellungen des Kniegelenkes, der Hals-Kopf-Region und der Leber. In Abbildung 4.1 sind

alle Beispiele in dieser Reihenfolge abgebildet. Sie repräsentieren verschiedene Regionen des menschlichen Körpers und die jeweils enthaltenen Strukturen variieren bezüglich folgender Kriterien:

- Typ (Gefäßbäume, Gewebe, Organe, Knochen)
- Relative Größe (Schädel, Lymphknoten)
- Form (kompakt, elongiert, verästelt)

Weiterhin sind in den ausgewählten Visualisierungen teilweise stark umschlossene (innenliegende Teile des Kniegelenks), durchdringende (Tumore) und nahe beieinander liegende Strukturen (Lymphknoten und Gefäße) erkennbar. Außerdem sind die variierenden Transparenzwerte von Bedeutung, welche die Wichtigkeit der jeweiligen Szenenobjekte widerspiegeln [NEUGEBAUER, 2006]. Diese Tatsache kann hier jedoch nicht direkt genutzt werden, da semi-transparente Strukturen bis zu einem gewissen Transparenzgrad genauso fixiert werden können, wie opake Objekte. Bei einer Abstandseinschätzung, zwischen beispielsweise Tumor und Gefäß müssen Fixierungen beider Strukturen zunächst als gleichwertig betrachtet werden. Der Transparenzwert des Gefäßes hat in diesem Fall nichts mit der Wahrscheinlichkeit einer Fixierung dieser Struktur zu tun.

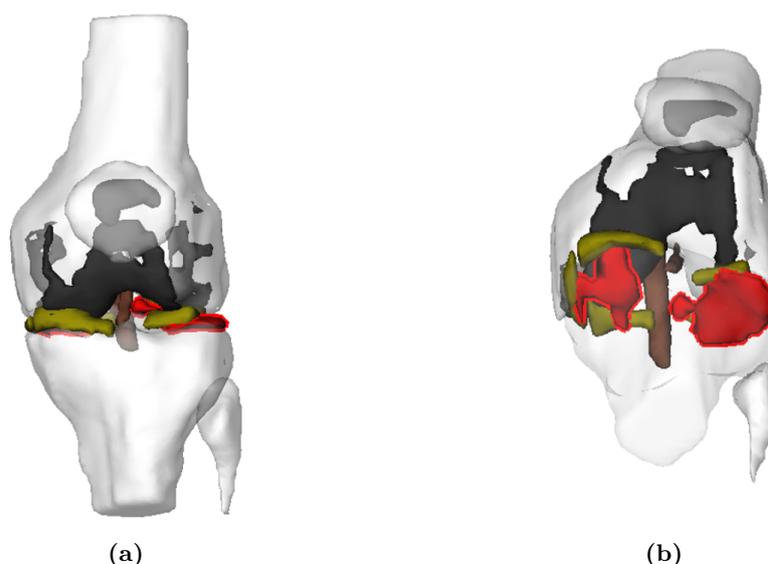


Abbildung 4.2: Darstellung des Knies aus der typischen Sicht des behandelnden Arztes (a) sowie von unten, aus Richtung des Fußes (b). Die betroffenen Gelenkknorpel sind rot eingefärbt und erst aus dem in (b) eingenommenen Blickwinkel deutlich analysierbar [NEUGEBAUER, 2006].

Ein typischer Anwendungsfall für den Einsatz der **Knievisualisierung** ist die Behandlung geschädigter Gelenknorpel. Wie in Abbildung 4.2 zu erkennen ist, liegen alle Bestandteile des Knies im Innern des Gelenks dicht beieinander und verdecken sich gegenseitig. Dies erschwert die Erkennung der Gelenknorpel in Abbildung 4.2(a) deutlich und bereitet einem Algorithmus zur Objektidentifikation ebenfalls Schwierigkeiten. Aufgrund der Ungenauigkeiten bei der Datenaufnahme wäre die Unterscheidung von Fixierungen der in Abbildung 4.2(a) rot markierten Bereiche durch die Betrachtung eines einzigen Zeitpunktes bzw. einer einzigen Fixierung aus dem Datenprotokoll kaum realisierbar. Durch die Drehung des Modells in die

in Abbildung 4.2(b) gezeigte Position, werden die betreffenden Regionen viel deutlicher und größer abgebildet. Dies bringt folgende Vorteile mit sich:

- Erleichterung der Planung durch den behandelnden Arzt
- Verbesserte Möglichkeiten zur Bestimmung des fokussierten Objektes, aufgrund der nun vorliegenden Größenverhältnisse

Das eigentliche Interesse an einer Struktur ist oft erst durch das Wechseln des Betrachtungswinkels deutlich aus den aufgenommenen Daten zu erkennen. Aus diesen Betrachtungen ergibt sich, neben lokalen Aussagen, die Forderung zeitlich gesehen globaler Aussagen über die Fokussierung eines Objektes.

Aus Abbildung 4.2 ist auch deutlich zu erkennen, dass die Zielstrukturen¹ bei einigen Kamerapositionen teilweise stark verdeckt sind. In diesen Fällen weisen entsprechende Objekte einen geringen Informationsgehalt auf und die Wahrscheinlichkeit, dass sie für den behandelnden Arzt als interessant erscheinen, sinkt. Andererseits kann argumentiert werden, dass der Arzt diesen Blickwinkel absichtlich gewählt hat, um beispielsweise Vergleiche zwischen zwei Strukturen durchführen zu können. Da ein solches Verhalten jedoch schwierig zu generalisieren ist, erscheint die erste Herangehensweise zunächst als sinnvoller. Der oben erwähnte Informationsgehalt ist ein Maß, welches z.B. in der Bildverarbeitung zur Datenkompression eingesetzt wird. Der durchschnittliche Informationsgehalt eines Signals ist auch unter dem Namen „Entropie“ bekannt. Die Information, welche zur Approximation der Entropie in medizinischen Visualisierungen herangezogen wird, ist bei der Sichtbarkeitsanalyse vom jeweiligen Anwendungsgebiet abhängig [NEUGEBAUER, 2006]. Dort genannte Möglichkeiten sind die projizierte Fläche einer Struktur oder aber Objektmerkmale wie die Oberflächenkrümmung.

Da an medizinische Visualisierungen hohe Anforderungen bezüglich der Genauigkeit von z.B. Abständen und Volumina gestellt werden, ist die Komplexität der zugrunde liegenden Datensätze in vielen Fällen auch entsprechend hoch. Durch die Ausnutzung von Verfahren zur Komplexitätsreduktion kann die Anzahl der Polygone jedoch, bei annähernd gleichbleibender Genauigkeit, reduziert werden. Alle Beispiele der Abbildung 4.1 bestehen aus polygonalen Modellen, welche durch Verfahren, wie *Marching Cubes*, entstanden sind. Das Modell der Leberdarstellung hat 360.520 und die Halsvisualisierung hat 139.536 Dreiecke. Gerade bei einem objektbasierten Verfahren ergibt sich daher ein relativ hoher, rechnerischer Aufwand.

Ein typisches Anwendungsbeispiel für die **Halsvisualisierung** ist die Planung einer Lymphknotenausräumung. Eine entsprechende Operation wird nach der Diagnose eines Tumorbefalls geplant. Der behandelnde Arzt muss dazu den Infiltrationsgrad gesunden Gewebes und vor allem die Nähe von Tumoren zu wichtigen Strukturen genau einschätzen. Kleine und verdeckte Objekte können dabei leicht übersehen werden.

Wie in Abbildung 4.3 zu erkennen ist, können Lymphknoten, im Vergleich zu anderen Objekten, relativ klein sein. Solche Strukturen sind leicht zu übersehen. Außerdem muss bei der Datenanalyse, aufgrund der Ungenauigkeit der Eye-Tracking-Systeme (ETS's), ein zusätzlicher Bereich um den registrierten Blickpunkt betrachtet werden. Der kleine Lymphknoten würde daher immer bei in dessen Nähe liegenden Fixierungen des benachbarten Halsmuskels von einem Algorithmus ebenfalls als „fokussiert“ erkannt werden. Auch hier stellt sich wieder die Forderung einer zeitlich globalen Betrachtung. Werden häufiger Bereiche um diesen

¹Mit Zielstrukturen werden für die Versuchsperson bzw. den Versuchsleiter interessante Objekte einer Visualisierung bezeichnet.

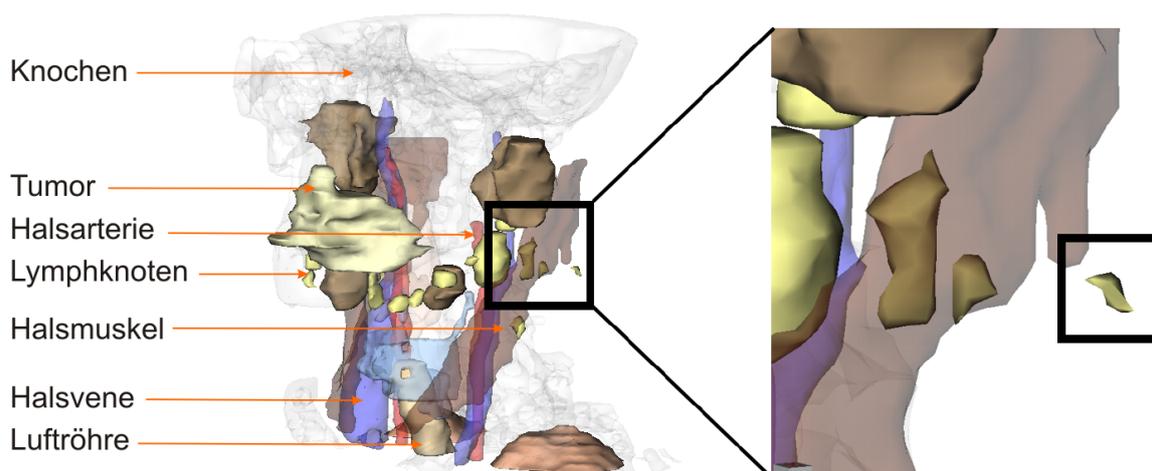


Abbildung 4.3: Größenvergleich zwischen einem Lymphknoten und den anderen anatomischen Strukturen nach [NEUGEBAUER, 2006].

Lymphknoten fokussiert, so kann das als Interesse des Betrachters an diesem Objekt aufgefasst werden. Zur Messung der beschriebenen Verdeckungen sowie der Größe der dargestellten Strukturen sind Informationen nötig, welche diese Eigenschaften adequat beschreiben. Hierfür eignen sich die in NEUGEBAUER [2006] vorgestellte Maße „Anzahl der Pixel eines Objektes bei verdeckungsfreier Projektion“ und „tatsächliche Anzahl der auf die Bildebene projizierten Pixel“. Außerdem ist ein Maß nötig, welches die räumlich gesehen um den Blickpunkt bestehende Sichtbarkeit fokussierter Objekte beschreibt.

Bei der obigen Erläuterung der Knievisualisierung sollte deutlich geworden sein, dass häufig eine gegenseitige Verdeckung anatomischer Strukturen eintritt. Dabei können Objekte durch Strukturen mit folgenden Materialeigenschaften verdeckt werden:

- opak (opake Verdeckung)
- semi-transparent (semi-transparente Verdeckung)

Bei der Verdeckung durch eine opake Struktur sind, abhängig von dessen Position, Form und Größe, Teile der bzw. die gesamte Zielstruktur nicht mehr sichtbar. Bei semi-transparenten Verdeckungen sind diese Faktoren zusätzlich abhängig vom Grad der Transparenz und der Anzahl semi-transparent verdeckender Objekte. Die Sichtbarkeit der Zielstruktur sinkt daher mit größer werdender Anzahl semi-transparent verdeckender Objekte.

Ein letztes Anwendungsbeispiel bezieht sich auf verästelte Strukturen, wie sie bei der **Lebervisualisierung** aus Abbildung 4.4 zu sehen sind. Eine typische Anwendung hierfür ist die Tumorentfernung. Das kranke Gewebe befindet sich dabei im Leberinnern und verdeckt teilweise die filigranen Gefäße. Bei der Entfernung des Tumors muss darauf geachtet werden, das Gefäßkonstrukt nicht zu verletzen, da die verschiedenen Leberregionen sonst nicht mehr mit Blut versorgt werden können und damit ernsthaft Schaden nehmen.

Um zu untersuchen, welche Objekte im einzelnen betrachtet wurden, gab es die Idee, ein bildbasiertes Verfahren einzusetzen. Mittels zweidimensionaler *Bounding-Boxes* sollte herausgefunden werden, in der 2D-Bounding-Box welcher Struktur sich die vom ETS aufgenommene Blickkoordinate befindet. Bounding-Boxes sind Hüllkörper zur Beschleunigung vieler Berechnungen auf komplexen Objekten, deren geometrische Komplexität relativ gering ist.

Sie ermöglichen zwar die Objektbestimmung mit einem erheblichen Geschwindigkeitsvorteil, Abbildung 4.4 zeigt jedoch, wie ungeeignet diese Form der Approximation unkompakter Strukturen ist. Ziel dieser Arbeit ist es, die Menge der sich an einem Blickpunkt befindlichen, potentiell fixierten Objekte durch eine Auswahl geeigneter Maße und Verfahren möglichst weit einzuschränken. Aufgrund der Möglichkeit, 3D-Szeneninformationen zu verwenden, bietet sich ein objektbasiertes Verfahren zur Bestimmung der an einem Blickpunkt liegenden Struktur(en) an. Ein Objekt kann damit über eine Schnittpunktberechnung zwischen den in die Szene projizierten Blickpunkten (Blickstrahlen) und dem polygonalen Modell des Datensatzes bestimmt werden. Durch den Vergleich mit anderen Algorithmen wurde in Abschnitt 3.3.1 gezeigt, dass sich die Rechenzeit der Strahlanfrage durch eine intelligente Schnittpunktberechnung reduzieren lässt.

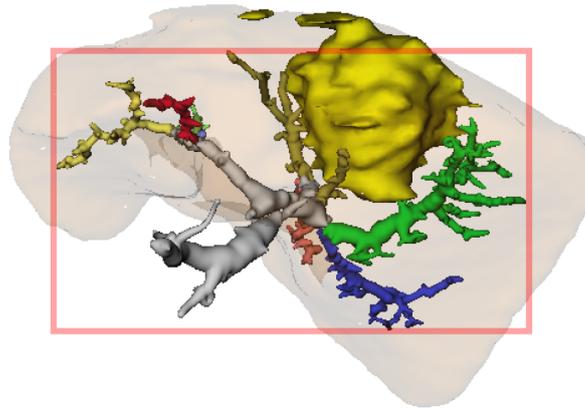


Abbildung 4.4: Darstellung der Leber. Der Gefäßbaum wird durch eine zweidimensionale Bounding-Box approximiert [NEUGEBAUER, 2006].

Ein letzter Aspekt, der an der Lebervisualisierung deutlich wird, ist die Umhüllung der Gefäße und des Tumors durch das Lebergewebe. Es liegt bei der Betrachtung aller Strukturen ebenfalls im Sichtfeld und ist gleichzeitig nur selten von wirklichem Interesse. Es dient vielmehr als Kontextinformation. Bei der Analyse der Blickpunktdaten muss daher eine Möglichkeit gefunden werden, die Fixierung des Lebergewebes auszuschließen, wenn tatsächlich eine der innenliegenden Strukturen betrachtet wurde. Ähnlich verhält es sich mit den Knochen (Schädel, Wirbelsäule, Schultern) aus der Halsvisualisierung. Auffällig bei beiden Strukturen ist ihre relative Größe zu anderen Objekten sowie deren hohe Transparenz. Eine Möglichkeit wäre, beide Objektmerkmale in die Bestimmung fokussierter Strukturen einfließen zu lassen. Hierbei sollte jedoch darauf geachtet werden, dass diese Merkmale nur zur Modifikation der Fixierungswahrscheinlichkeiten für die beschriebenen Strukturen verwendet werden, da dieser Schritt nicht allgemeingültig für alle Objekte anwendbar ist.

Materialeigenschaften

Neben den räumlichen Verhältnissen anatomischer Strukturen spielen ebenfalls Materialeigenschaften eine bedeutende Rolle für die Beurteilung des Blickverhaltens einer Person. Gerade an der Halsvisualisierung in Abbildung 4.3 ist die aus anatomischen Atlanten entnommene und mit Ärzten abgestimmte Farbgebung deutlich erkennbar.

Verschiedene Farben erregen die Aufmerksamkeit des Betrachters unterschiedlich stark und beeinflussen damit dessen Blickverhalten. Da das konzipierte Verfahren jedoch dazu genutzt

werden soll, um medizinische Visualisierungen bezüglich der Erfüllung ihres Visualisierungsziels zu evaluieren ist es für die Bestimmung fokussierter Objekte nicht zweckmäßig, Vorannahmen über die zu analysierende Farbinformation zu treffen. Aus diesem Grund muss der Einsatz dieser und ähnlicher Informationen genau bedacht werden.

Vor allem bei hochtransparenten und großen Objekten, wie dem Schädel oder dem Lebergewebe, sind Teile der Rückseite sichtbar. In diesen Fällen ist es wünschenswert, herauszufinden, welchen Teil des Objektes die Versuchsperson fixiert hat. Um dies zu erreichen, ist ein objektbasierter Ansatz notwendig, welcher die 3D-Position einer Fixierung ermitteln kann. Anhand dieser Position ist es dann möglich zu sagen, ob eher der vordere oder eher der hintere Bereich fokussiert wurde.

Genauigkeitsbetrachtungen und Art der Präsentation

Die nachfolgenden Erläuterungen beziehen sich auf das Problem der Trackerungenauigkeit (vgl. Abschnitt 2.2) sowie die Möglichkeiten zur Präsentation einer medizinischen Visualisierung. Der Genauigkeitsfehler wird selbst bei modernen ETS's mit bis zu einem Grad angegeben. Für die Bestimmung fokussierter Objekte in medizinischen Visualisierungen ist dies relativ hoch, da anatomische Strukturen einerseits eine geringe Größe haben können und andererseits teilweise eng beieinander liegen. Aus diesen Gründen muss versucht werden, den Genauigkeitsfehler möglichst weit zu reduzieren. Hierfür bieten sich die in Abschnitt 3.1.1 vorgestellten *Parametrized Self-Organizing Maps* an. Mit ihrer Hilfe kann der mittlere Genauigkeitsfehler auf weniger als 0.4° reduziert werden [POMPLUN et al., 1994].

Neben dem Genauigkeitsfehler sind Abweichungen durch Driftbewegungen ein Problem. Die Versuchsperson muss während des Experimentes sehr still sitzen und darf sich aus der vor der Kalibrierung eingenommenen Position nur minimal herausbewegen, da sonst Driftabweichungen verursacht werden. Eine Person, die noch keine Erfahrung im Eye-Tracking hat, muss gründlich in den Ablauf des Experimentes und die Rahmenbedingungen eingewiesen werden. Es ist einigen Menschen trotz größter Konzentration nicht möglich, Driftabweichungen zu vermeiden und entsprechend still zu sitzen. Deshalb ist es unbedingt nötig, diesen Fehler zu kompensieren.

Aus Tabelle 4.1 gehen verschiedene Vor- und Nachteile einer Visualisierung bei der Verwendung der unterschiedlichen Präsentationsarten hervor. Wichtig dabei ist, dass jede Art ihre Stärken bzw. Schwächen hat und somit alle drei Präsentationsmöglichkeiten für die Evaluierung einer Visualisierung zum Einsatz kommen können. Da bei interaktiver und animierter Darbietung Folgebewegungen auftreten können, ergibt sich die Forderung nach einer Methode zur korrekten Erkennung dieser Augenbewegungen. Der Algorithmus zur Objektidentifikation sollte außerdem für alle drei Präsentationsarten geeignet sein.

Durch die Bewegung dargestellter Objekte können weitere Möglichkeiten zur Objektidentifikation genutzt werden. Dazu gehören die über einen längeren Zeitraum andauernden Folgebewegungen. Diese treten auch bei interaktiven Visualisierungen auf, während der Benutzer die Kameraposition verändert. Da das Auge, während der Fixierung eines sich bewegenden Objektes bestrebt ist, diese Fixierung aufrecht zu erhalten, entsteht ein kontinuierlicher Pfad des Blickpunktes, der in gewissem Maße der Objektbewegung gleicht. In statischen Bildern und in Phasen, in denen eine interaktive Visualisierung nicht durch den Benutzer verändert wird, treten Folgebewegungen allerdings nicht auf.

Typ	Besonderheit
animiert	Die Definition besonders günstiger Animationspfade im Vorfeld ermöglicht die optimale Hervorhebung wichtiger Eigenschaften der Visualisierung. Die Wahl der Animationsparameter bestimmt dabei, wie lange und aus welchen Blickwinkeln die Zielstrukturen gesehen werden können. Die Möglichkeit zur freien Exploration der Szene bietet diese Form der Präsentation jedoch nicht.
interaktiv	Diese Form der Darbietung ermöglicht die größtmögliche Freiheit zur Exploration. Es kann jedoch vorkommen, dass Strukturen nicht korrekt wahrgenommen werden, da sie nicht aus dem optimalen Blickwinkel betrachtet wurden.
Bild	Ein statisches Bild bietet aufgrund des eingeschränkten Blickwinkels keine Möglichkeit zur freien Exploration. Aus dem selben Grund können Zielstrukturen leicht übersehen werden. Andererseits ist es möglich, ein Bild mit einer optimal positionierten Kamera zu erzeugen, so dass die Evaluierung ganz bestimmter Strukturen oder Visualisierungstechniken möglich wird.

Tabelle 4.1: Vergleich der Möglichkeiten zur Präsentation einer dreidimensionalen Visualisierung. Die entsprechenden Eigenschaften werden nach den Gesichtspunkten „Wahrnehmbarkeit der Zielstruktur(en)“ und „Möglichkeit zur freien Exploration“ gegenübergestellt. Daraus ergeben sich Möglichkeiten und Grenzen der Evaluierung einer solchen Visualisierung.

In dieser Studie werden Blickmusteranalysen auf möglichst einfachen Evaluierungsversuchen durchgeführt, um Fehlerquellen von vornherein möglichst auszuschließen. Beim Experimentdesign sollte daher weiterhin auf das Vermeiden jeglicher Interaktion durch die Versuchsperson geachtet werden. Da Bewegungen des Arms durch die Schulter auf angrenzende Körperregionen übertragen werden können, ist durch eine Interaktion mit der Maus oder anderen Eingabegeräten eine unbewusste Bewegung des ganzen Körpers und insbesondere des Kopfes möglich. Die Verwendung von interaktiven Experimenten könnte sich also negativ auf das Ergebnis der Blickpunktmessung auswirken. Aus diesem Grund wird es als sinnvoll angesehen, zunächst animierte Testszene zu entwerfen, welche die Versuchspersonen einfach betrachten müssen.

Hierfür ist die Wahl angemessener Animationsparameter notwendig. Wichtig ist dabei einerseits, dass die Animation nicht zu langsam ist, damit sich die Probanden nicht langweilen. Andererseits sollte sie nicht zu schnell ablaufen, um die Konzentration auf die eigentliche Aufgabe sicherzustellen. Da sich die Geschwindigkeit über das Experiment nicht ändert, wird die Position der virtuellen Kamera, während des Versuches, in konstanten Schritten und auf Basis einstellbarer Parameter verändert. Um außerdem die Auswertung nach dem Bottom-Up-Ansatz einfacher zu gestalten, ist weiterhin eine einfache Rotation einem komplexen Animationspfad vorzuziehen.

4.1.2 Anforderungen an die Evaluierungsversuche

Aus den eben erläuterten Problemen ergeben sich die Rahmenbedingungen für die Durchführung der Versuche zur Evaluierung des Verfahrens. Weitere bedeutende Hinweise stammen aus Erfahrungswerten mit ETS's. Diese Rahmenbedingungen beschreiben verschiedene Aspekte des Experimentdesigns und dienen damit der Minimierung ungewollter Einflüsse auf die Messergebnisse. Solche Einflüsse bilden beim Eye-Tracking z.B. eine ungünstige Wahl des Versuchsortes. Die physischen und psychischen Eigenschaften der teilnehmenden Versuchspersonen

sind ebenfalls von Bedeutung. Neben der Charakterisierung der Experimentausführung helfen die Rahmenbedingungen außerdem bei der Gestaltung reproduzierbarer Versuche. Daher werden folgend Festlegungen bezüglich der Auswahl von Versuchspersonen, der Experimentdurchführung sowie der dargebotenen Visualisierungen getroffen. Da für Studien dieser Art vorwiegend videobasierte Eye-Tracker verwendet werden, um die Belastung für die Versuchspersonen so gering wie möglich zu halten, wird folgend speziell kopf- und desktopplatzierte Geräte eingegangen (vgl. Abbildung 2.6).

Versuchspersonen

Insbesondere für die Versuchspersonen gelten besondere Regeln, da in den Experimenten teilweise unbewusste Reflexe des visuellen Systems gemessen werden und diese durch viele Einflüsse manipulierbar sind. Diese Regeln sind in Form der Tabelle 4.2 zusammenfassend dargestellt.

Einflussfaktor	Beschreibung
Kaffee, Alkohol, Medikamente	Diese Mittel wirken auf das Nervensystem und damit auf das Blickverhalten. Sie sollten daher vor dem Experiment nicht eingenommen werden.
Müdigkeit	Alle Versuchspersonen müssen ausgeschlafen sein, da sonst die Wahrnehmungs- und Fixierungsfähigkeit beeinträchtigt wird.
Experimentdetails	Vor dem Experiment werden außer der zu lösenden Aufgabe keine weiteren Details bekannt gegeben. Dadurch wird verhindert, dass die Probanden versuchen, bestimmte Vorgänge im Experiment vorherzusehen und die Daten sich dadurch verfälschen.
Sehstörungen	Es ist bekannt, dass Reflexionen durch Brillen und Kontaktlinsen Probleme beim Eye-Tracking bereiten. Aus diesem Grund erscheint es sinnvoll, bevorzugt Probanden mit gesunden Augen auszuwählen und dieses Problem anhand einer Testgruppe zu überprüfen.
Augenmakeup	Makeup auf den Augen erschwert die Einstellung des Schwellenwertes zur Pupillensegmentierung und damit die Kalibrierung. Es sollte daher vor einem Experiment darauf verzichtet werden.

Tabelle 4.2: Beschreibung der auf die Versuchspersonen wirkenden und das Experiment beeinflussenden Faktoren sowie sich daraus ergebende Rahmenbedingungen.

Experimentdurchführung

Das Handbuch des ETS, welches in dieser Arbeit verwendet wird, stellt Vorschläge zur Steigerung der Qualität der aufgenommenen Blickpunktdaten zur Verfügung. Es wird beispielsweise dazu geraten, helle Lichtquellen oder Fenster in der Nähe der Versuchsperson zu vermeiden, da beides Reflexionen auf Bildschirm und Hornhaut verursachen kann. Weiterhin sollte auf eine ausreichende Beleuchtung geachtet werden, da sich die Pupille des Menschen in dunkler Umgebung erweitert. Dadurch können Probleme bei der Kalibrierung auftreten. Des Weiteren

muss jegliche Ablenkung, hervorgerufen durch z.B. Umgebungsgeräusche oder andere Personen, auf ein Minimum reduziert werden. Die Sitzposition der Versuchsperson sollte zudem stabil sein.

Die auf dem Kopf befindlichen Geräte müssen an die jeweilige Versuchsperson angepasst werden. Dazu gehört die Befestigung mit Hilfe von Fixierungsbändern. Der Eye-Tracker muss fest auf dem Kopf sitzen, um Messfehler durch Verrutschen zu vermeiden. Dies ist auf längere Sicht unbequem und führt zu Kopfschmerzen. Daher ist es wichtig, einen Kompromiss zwischen gutem Sitz und Komfort zu finden. Eine weitere, hardwareseitige Anpassung besteht in der Ausrichtung der am Gerät befestigten Kameras. Dieser Prozess ist eng mit den softwareseitigen Einstellungen zur Pupillendetektion verbunden. Die Kameras müssen so justiert werden, dass sie die Pupille bei jeder Augenstellung möglichst gut aufnehmen können. Sie dürfen das Sichtfeld der Versuchsperson dabei allerdings nicht verdecken. Die Durchführung dieser Prozedur für beide Augen ist relativ zeitaufwändig. Hinzu kommt, dass die anschließende Kalibrierung mit beiden Kameras weitaus länger dauert, als mit nur einer. Da bei den meisten Menschen beide Augen die selbe Stelle fixieren, ist es somit zweckmäßig das Experiment mit nur einer Kamera durchzuführen, was als monokulares Eye-Tracking bezeichnet wird.

Eye-Tracking-Systeme geben häufig Auskunft über die Qualität der Kalibrierung. Es muss auf ein optimales Kalibrierungsergebnis geachtet werden. Ein Experiment darf nicht mit einem suboptimalen Ergebnis begonnen werden, da die Abweichung der Blickpunktdaten in diesem Fall zu groß wird. Die Versuchsperson muss während der Kalibrierungsphase ruhig sowie entspannt sitzen und darf auf keinen Fall ihren Kopf, sondern nur ihre Augen, bewegen. Stellt sich das Ergebnis der Kalibrierung als qualitativ ungenügend heraus, so muss dieser Schritt erneut durchgeführt werden.

Der Abstand des Betrachters vom Display spielt für den Eye-Tracker und die Versuchsperson eine wesentliche Rolle. Zu nahe am Bildschirm kann die Person nicht mehr das gesamte Display überblicken, ohne den Kopf zu bewegen. Ist der Abstand zu groß, steigt die Ungenauigkeit der Messung. Es wird eine nach @CDC ergonomische Entfernung von 75 *cm* zwischen Auge und Display festgelegt. Dies entspricht der doppelten Displaybreite eines 19-Zoll-Monitors und ist in SENSOMOTORIC INSTRUMENTS GMBH [1999] als optimale Entfernung angegeben.

Aktuell stehen noch keine fortgeschrittenen Auswertungsmöglichkeiten für die umfangreichen Experimentdaten zur Verfügung. Das Datenprotokoll für ein Experiment von 30 Sekunden mit einer Versuchsperson würde, bei einer Aufnahmerate von 250 *Hz*, 7500 Samples betragen. Um den Aufwand zur Auswertung zu begrenzen, ist die Anzahl der teilnehmenden Versuchspersonen, der im Rahmen dieser Arbeit durchgeführten Experimente, auf 14 beschränkt. Aus diesen Ergebnissen können zwar noch keine allgemeingültigen Daten erhoben werden, eine Gewinnung von Erkenntnissen bezüglich vorhandener Tendenzen ist jedoch möglich.

Dargebotene Visualisierungen

Die Experimente müssen so gestaltet werden, dass störende Einflüsse vermieden werden. Dies kann mit Hilfe geeigneter Szenarien erfolgen. Steht beispielsweise keine Driftkorrektur zur Verfügung, so müssen die Versuche in ihrer Länge begrenzt werden. In den Beispielexperimenten von EYELINK I wird die Driftkorrektur nach ca. 10 Sekunden durchgeführt. Dies stellt jedoch einen relativ kurzen Zeitraum für ein Experiment dar. Daher werden die hier durchgeführten Evaluierungsversuche auf maximal 15 Sekunden begrenzt. Sobald eine für

die Versuchsperson unauffällige Driftkorrektur zur Verfügung steht, ist es möglich längere Experimente durchzuführen.

In Abschnitt 3.1 wurde erläutert, dass die Untersuchung eines 2D-Schichtbildes auf krankhaft verändertes Gewebe einer Suchaufgabe ähnelt. Für fortgeschrittene Evaluierungsstudien muss die Präsentation des visuellen Stimulus ebenfalls mit einer Aufgabe kombiniert werden. Dadurch kann ein ähnliches Blickverhalten, wie das eines Arztes bei der Analyse medizinischer Visualisierungen, mit Personen ohne medizinische Fachkenntnisse nachgeahmt werden. Die Vorstellung der jeweiligen Aufgaben für die, im Rahmen dieser Arbeit, durchgeführten Versuche erfolgt bei der Erläuterung der Experimente in Abschnitt 6.1. Beispiele für einfache Aufgaben, die den Versuchspersonen zur Ausführung unbewusster Blickmuster veranlassen, sind:

1. Zählen Sie alle dargestellten Objekte!
2. Schätzen Sie die Abstände mehrerer Objekte zueinander ein!
3. Finden Sie heraus, welches der dargestellten Objekte am größten ist!
4. Identifizieren Sie das Objekt, welches für Sie am kompaktesten erscheint!

4.1.3 Schlussfolgerungen

In diesem Abschnitt wurden durch medizinische Anwendungsbeispiele die Anforderungen an ein Verfahren zur Identifikation fokussierter Objekte, mittels Eye-Tracking, verdeutlicht. Anhand verschiedener Kriterien wie Materialeigenschaften, Szenenkomplexität oder der Ungenauigkeit moderner ETS's wurden diese herausgearbeitet. Die wesentlichen Anforderungen werden im Folgenden beschrieben.

Globale Blickpunktanalyse - Es wurde gezeigt, dass Zielstrukturen aufgrund unterschiedlicher Blickwinkel und den damit zusammenhängenden globalen Verdeckungseigenschaften besser bzw. schlechter sichtbar sind. Da angenommen wird, dass sich das Blickverhalten zeitlich gesehen mit diesen Eigenschaften ändert, ist es nicht ausreichend eine Objektidentifizierung auf Basis einzelner Blickpunkte durchzuführen. Eine zeitlich globale Blickmusteranalyse ist deshalb erforderlich.

Erweiterte Metriken - Neben der globalen Blickmusteranalyse ist es nötig, Blickpunkte und damit Fixierungen einzeln zu betrachten, da letztere das Interesse des Betrachters an einer bestimmten Region bzw. Struktur repräsentieren. Aus diesen Grund muss die Objektidentifizierung zunächst auf den Daten basieren, die vom Eye-Tracker aufgenommen wurden. Die einzelne Bildschirmkoordinate reicht dabei aufgrund der Eye-Tracking-Ungenauigkeit nicht aus. Es ist erforderlich, einen Bereich um jeden Blickpunkt herum zu betrachten. Des Weiteren ist es notwendig, speziell bei der Verwendung von Animationen, die Anwendbarkeit von Folgebewegungen als Kriterium zur Ermittlung des fokussierten Objektes zu untersuchen. Neben den beschriebenen Metriken haben sich auch folgende Maße, mit denen sich Verdeckungseigenschaften und Größe von Strukturen beschreiben lassen, als bedeutend herausgestellt:

1. Anzahl sichtbarer Pixel eines Objektes
2. Anzahl sichtbarer Pixel bei verdeckungsfreier Projektion
3. räumlich lokale Sichtbarkeit

Als letzter Punkt ist die Entropie einer Struktur von Bedeutung. Sie sollte so berechnet werden, dass sie Auskunft über die aktuelle Sichtbarkeit erlaubt. Dazu eignet sich eine Kombination aus den Maßen zwei und drei der obigen Liste. Diese Informationen geben Aufschluss darüber, wie sehr eine Struktur maximal sichtbar sein kann bzw. gerade sichtbar ist.

Genauigkeit bei der Objektbestimmung - Bezüglich dieses Aspektes hat sich gezeigt, dass Bounding-Boxes eine zu grobe Repräsentation der dargestellten Strukturen repräsentieren. Um eine geeignete Genauigkeit zu erzielen, muss daher direkt mit dem 3D-Objekt gearbeitet werden (objektbasierter Ansatz). Um die Genauigkeit der Objektbestimmung weiter zu erhöhen, ist die Integration von Verfahren zur Reduktion des Genauigkeitsfehlers sowie der Abweichungen durch Driftbewegungen erforderlich.

Gesonderte Behandlung großer Objekte - Für diese Aufgabe erscheint die Ausnutzung von Transparenz und Größe entsprechender Strukturen geeignet. Dabei sollte jedoch darauf geachtet werden, dass Objekte nicht allgemeingültig, auf Basis dieser Eigenschaften eine Wichtigkeit zugewiesen bekommen. Die Anpassung entsprechender Metriken mit Hilfe von Wichtungsfaktoren, Schwellenwerten oder speziellen Funktionen stellt in diesem Falle eine mögliche Lösung dar.

Vereinfachte Evaluierungsversuche - Animationen haben sich in den obigen Betrachtungen als besonders geeignet zum einen für die Präsentation medizinischer Visualisierungen und zum anderen zur Identifikation fokussierter Objekte ergeben. Daher wird es als sinnvoll angesehen, die ersten Evaluierungsversuche als Animation zu gestalten.

4.2 Verfahren zur Bestimmung fokussierter Objekte

In diesem Abschnitt wird ein an den obigen Anforderungen orientiertes Verfahren entwickelt, welches die Bestimmung fokussierter Objekte in einem gewissen Rahmen erlaubt. Aufgrund der Anatomie des menschlichen, visuellen Systems, des beschriebenen Genauigkeitsfehlers und den in medizinischen Darstellungen auftretenden Transparenzen, liegen häufig mehrere Objekte im Fokus des Betrachters. Es wurde weiterhin beschrieben, dass sich der Mensch auf mehrere Objekte gleichzeitig konzentrieren kann, was mittels Eye-Tracking bislang nicht messbar ist. Daher kann das Ziel dieses Verfahrens nicht die Bestimmung eines einzigen, fixierten Objektes sein. Vielmehr soll eine möglichst weite Reduktion der Anzahl potentiell fixierter Objekte, zu einem bestimmten Zeitpunkt, ermöglicht werden. Zudem werden Fixierungswahrscheinlichkeiten für jedes Objekt und beliebige Zeitpunkte abrufbar sein. Diese Gedanken basieren auf den drei Schritten:

1. Entwicklung den Anforderungen entsprechender Metriken
2. Zusammenfassung der Wahrscheinlichkeiten innerhalb von Fixierungen
3. Zeitlich lokale sowie globale Analyse resultierender Wahrscheinlichkeiten

Die eingesetzten Metriken arbeiten lokal auf den vom ETS aufgenommenen Samples. Durch die Kombination aller Metriken wird anschließend entschieden, wie wahrscheinlich ein Objekt zu beliebigen Zeitpunkten bzw. für jede Fixierung fokussiert wurde. Diese Analyse kann letztendlich durch eine globale Betrachtung der ermittelten Wahrscheinlichkeiten unterstützt werden. All das wird im Rahmen des in den folgenden Abschnitten vorgestellten 3-Phasen-Modells erläutert.

4.2.1 Schritte nach dem Experiment

Nachdem das Eye-Tracking-Experiment mit dem Ziel einer sich anschließenden Datenanalyse (Offlineanalyse) beendet ist, liegen verschiedene Informationen in Form einer Protokolldatei vor. Einerseits können dies ausschließlich die Samples sein. Andererseits liefern verschiedene Hersteller Zusatzsoftware, welche die aufgenommenen Samples zu zeitlich größeren Ereignissen (Events) wie Fixierungen, Sakkaden oder Lidschlussreflexen zusammenfassen. Die entsprechenden Events können wahlweise, abhängig von der gestellten Aufgabe und den damit benötigten Informationen zur Analyse, ebenfalls in das Protokoll geschrieben werden. Im Falle der für diese Arbeit nötigen Analysen, ist weiterhin die Integration folgender Informationen notwendig:

- Kameraevents (Position, Orientierung, Projektionsart, etc. der virtuellen Kamera)
- Interaktionsevents (Maus- und Tastaturevents)

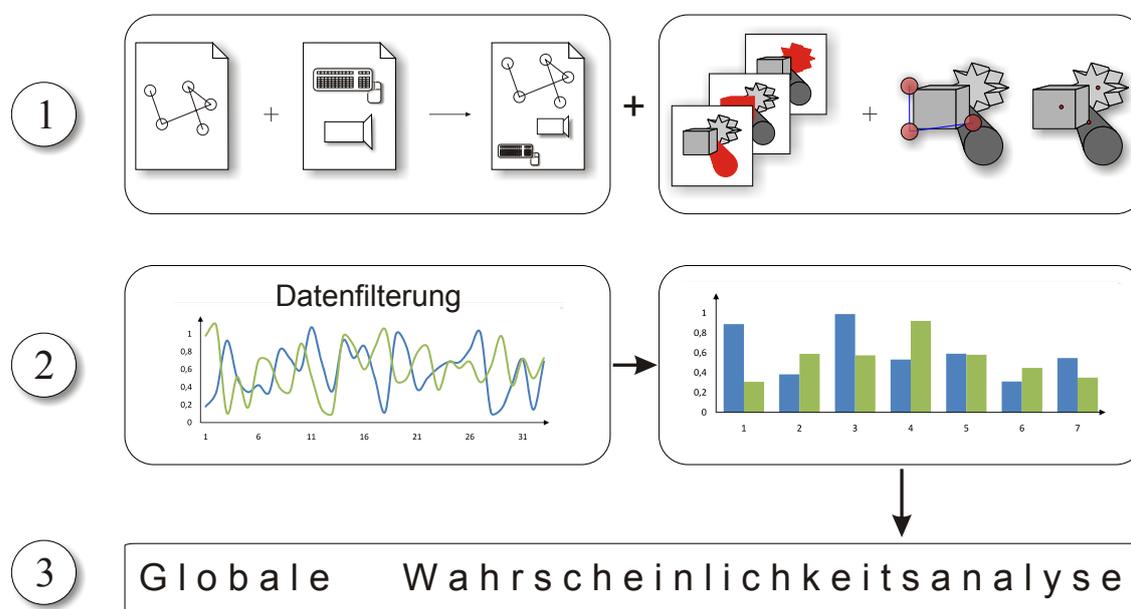


Abbildung 4.5: Darstellung des Gesamtprozesses zur Objektidentifizierung in seinen drei Schritten. Schritt 1: Vereinigung des Blickpunktprotokolls mit Kameraevents und Benutzerinteraktionsdaten (links) sowie Berechnung der bild- und objektbasierten Metriken (rechts) für jedes Sample (Ermittlung der Wahrscheinlichkeiten). Schritt 2: Datenfilterung und Zusammenfassung der gefilterten Wahrscheinlichkeiten innerhalb von Fixierungen. Schritt 3: Globale Analyse der Wahrscheinlichkeiten aus Phase zwei für jedes Objekt und jede Fixierung.

Abbildung 4.5 stellt schematisch das in drei vertikal angeordnete Stufen unterteilte 3-Phasen-Modell dar. Es dient zur formalen Beschreibung des Ablaufs der Objektidentifizierung. Die oberste Stufe verdeutlicht die ablaufenden Schritte direkt nach dem Experiment. Dazu zählen die Vereinigung der Eye-Tracking-Events mit den Kamera- und Interaktionsevents sowie die zeitlich entkoppelte Berechnung der Metriken. Die Aufzeichnung der Kameraevents findet statt, damit zu einem späteren Zeitpunkt das Experiment als Animation mit einer überlagerten Blickpunktvisualisierung wiedergegeben werden kann. Außerdem ist diese Information

zur Ausrichtung der virtuellen Kamera während der Analyse notwendig, so dass eine exakte Berechnung der Metriken möglich ist. Da die Speicherung von Interaktionsevents keinen großen Mehraufwand mit sich bringt und beispielsweise die Mausposition zur Driftkorrektur genutzt werden kann, bietet sich eine Integration in das Protokoll an. Diese Kamera-, Maus- und Tastaturevents müssen mit den Blickpunktdaten zeitlich synchronisiert werden, wenn beide Informationen auf verschiedenen Rechnern aufgenommen werden. Die vereinigte Datenmenge wird anschließend in einer Ascii-Datei abgelegt.

Bei der Durchführung mehrerer Experimente besteht das Ergebnis aus einer Protokolldatei pro Versuch, auf deren Grundlage die Objektidentifizierung durchgeführt wird. Die Experimentdurchführung sowie die Berechnung der Metriken bilden zwei zeitlich voneinander getrennte Schritte, da die Metrikenberechnung relativ zeitaufwändig ist. Durch diese zeitliche Entkopplung ist es möglich, mehrere Protokolldateien in einem Stapelverarbeitungsprozess, hintereinander abzuarbeiten. Im Folgenden werden die einzelnen Stufen des 3-Phasen-Modells kurz charakterisiert:

- 1. Stufe** - Die Berechnung der Metriken, welche im rechten Teil der ersten Stufe von Abbildung 4.5 dargestellt ist, dient der initialen Zuordnung einzelner Samples zu Objekten auf der Basis von Wahrscheinlichkeiten. Sie bildet zusammen mit der Vereinigung der verschiedenen Events die erste Stufe des 3-Phasen-Modells und findet bewusst vor der in Stufe zwei durchgeführten Datenfilterung statt, um die ermittelten Informationen zunächst für alle Samples verfügbar zu machen. Dies erleichtert die Anpassung des Konzeptes, falls im Rahmen dieser Arbeit gefilterte Blickpunkte zu einem späteren Zeitpunkt Verwendung finden sollten.
- 2. Stufe** - Die Datenfilterung leitet Stufe zwei ein, beseitigt unnötige Informationen aus der Protokolldatei und ermöglicht, aufgrund der Neustrukturierung aller Daten in Tabellenform, die Vereinfachung der Auswertung in Analyseprogrammen wie MATLAB oder SPSS. Im Anschluss an diese Filterung stehen für jedes erhaltene Sample und für jedes Objekt ein Wert zur Verfügung, der die Wahrscheinlichkeit der Fixierung eines Objektes zum entsprechenden Zeitpunkt angibt. Diese Wahrscheinlichkeiten werden nun innerhalb aller vom ETS detektierten Fixierungen zusammengefasst. Damit ergibt sich neben den Wahrscheinlichkeiten für jedes Sample auch eine Wahrscheinlichkeit für jedes Objekt über die Zeit einer jeden Fixierung.
- 3. Stufe** - Im Rahmen der dritten Stufe erfolgt die Vorstellung von Möglichkeiten zur globalen Analyse der in Phase zwei zusammengefassten Wahrscheinlichkeiten. Mit diesem Schritt berücksichtigt das entwickelte Verfahren sowohl die zeitlich lokale (Schritt eins und zwei) als auch die zeitlich globale Betrachtung der Blickpunktdaten, was sich in der Anforderungsanalyse als notwendig herausgestellt hat.

4.2.2 Vorstellung objekt- und bildbasierter Metriken

Der nachfolgende Teil beschreibt die im Rahmen dieser Arbeit entwickelten und eingesetzten Maße für die Zuordnung von Blickpunkten zu Objekten. Dieser Schritt ist in die erste Stufe des 3-Phasen-Modells einzuordnen. Jede der hier vorgestellten Metriken geht aus einer bestimmten Eigenschaft der Szenenobjekte oder des ETS hervor und wird zunächst für jedes Sample des Protokolls berechnet. Die Resultate stehen damit für alle möglichen Zeitpunkte, außer im Falle von Datenverlusten, zur Verfügung. Des Weiteren werden die Metriken

für jedes Objekt betrachtet und liefern daher gewisse Rückschlüsse über dessen Fixierungswahrscheinlichkeit. Das Ergebnis einer Metrik ist somit eine Wahrscheinlichkeit im Intervall $[0 \dots 1]$. Ihre Berechnung erfolgt getrennt für jedes Sample durch Division der ermittelten Werte einer Metrik, pro Objekt, durch deren Summe. Nach dieser Umrechnung gleicht die Summe der Werte einer Metrik für jedes Sample und über alle Objekte folglich dem Wert eins. Eine Metrik gibt damit für jedes Objekt an, wie wahrscheinlich dieses zu einem Zeitpunkt und auf Basis der jeweils zugrunde liegenden Eigenschaft fixiert wurde. Die Begriffe Sample und Zeitpunkt werden im weiteren Verlauf synonym verwendet, da die auswertbaren Zeitpunkte aus dem Blickpunktprotokoll durch Samples repräsentiert werden.

Die Umrechnung jeder Metrik trägt weiterhin zur einheitlichen Berechnung einer Gesamtwahrscheinlichkeit bei, welche alle Metriken in sich vereint. Diese Wahrscheinlichkeit wird durch die gewichtete Summe aller Metriken ermittelt. Damit eröffnet sich die Möglichkeit, mehrere Metriken in verschiedenem Maße in das Endergebnis einfließen zu lassen. Klassische Metriken, welche ausschließlich auf den registrierten Blickpunkten operieren, wurden in Abschnitt 2.2.3 erläutert. Die im Folgenden vorgestellten Maße beziehen dagegen ihre Informationen entweder direkt von den 3D-Objekten oder von deren Projektion auf die Bildebene der Kamera.

Objektbasierte Metriken: „Objekt innerhalb des Unsicherheitsbereiches“

Um zu erfahren, welche Objekte zu beliebigen Zeitpunkten im Fokus des Betrachters gelegen haben, müssen alle Samples einzeln analysiert werden. Aufgrund der sich bietenden Genauigkeit und der einfachen Möglichkeit zur Integration erweiterter Verfahren, wie der Analyse von 3D-Schnittpunkten zwischen Objekten und Blickstrahlen, wurde ein objektbasierter Ansatz als Grundlage für die initiale Zuordnung der Blickpunkte gewählt. Da zusätzlich die Ungenauigkeit eines ETS beachtet werden muss, ist die Untersuchung des Unsicherheitsbereiches um den eigentlichen Blickpunkt herum nötig. Dieser Bereich kann als Kreis mit dem Radius des entsprechenden Genauigkeitsfehlers approximiert werden. Er wird in Abbildung 4.6 anhand von drei potentiellen Blickpunkten schematisch dargestellt. Hierbei werden folgende zwei Aspekte deutlich:

1. Ein Verfahren, welches den Genauigkeitsfehler nicht beachtet, würde die Fixierung des Würfels nicht erkennen, da die Blickpunkte (Mittelpunkt des Unsicherheitskreises) neben diesen liegen. Aufgrund dieses Genauigkeitsfehlers muss der Unsicherheitsbereich ebenfalls analysiert werden.
2. Es kommt bei der Analyse des Unsicherheitskreises häufig vor, dass sich für einen Blickpunkt mehrere Objekte innerhalb dieses Bereiches befinden. Es muss versucht werden, diesen Effekt durch eine Verbesserung der Genauigkeit und den Einsatz weiterer Metriken zu minimieren.

Eine natürlichere Variante zum Unsicherheitskreis wäre die Anwendung des in Abschnitt 3.2 diskutierten, wahrnehmungsorientierten Modells. Hierbei wird nicht mehr der Kreis sondern der durch Gleichung 3.3 berechnete Bereich um den Blickpunkt analysiert. Die Wichtigkeit des den Blickpunkt umgebenden Bereiches sinkt dabei mit zunehmender Entfernung von diesem. Da der Mensch jedoch eine größere Region wahrnimmt, als den Unsicherheitsbereich, muss hierbei eine weitaus größere Umgebung des Blickpunktes untersucht werden. Aus diesem Grund wird das vereinfachte Modell eines Kreises mit gleichbleibender Wichtigkeit

vorgezogen. Für das objektbasierte Verfahren würde sonst ein zu großer Anstieg der Laufzeit resultieren.

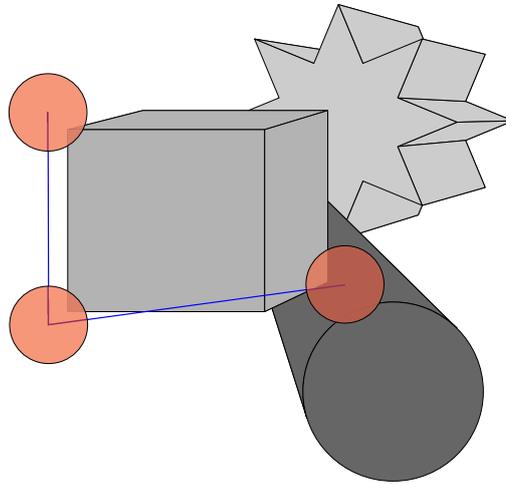


Abbildung 4.6: Schematische Darstellung der Idee hinter der objektbasierten Metrik „Objekt innerhalb des Unsicherheitsbereiches“.

Jeder der sich im Unsicherheitskreis befindlichen Bildschirmkoordinaten wird nun auf einen dreidimensionalen Strahl (Blickstrahl) projiziert. Anschließend erfolgt die Schnittpunktberechnung mit der Szenengeometrie, wobei Blickpunkte geschnittenen Objekten zugeordnet werden. Hieraus ergibt sich der hohe Rechenaufwand des objektbasierten Verfahrens, der allerdings durch den Einsatz von Hüllkörpern und der Ausnutzung räumlicher Zusammenhänge reduzierbar ist. Unter den in Abschnitt 2.2 beschriebenen Bedingungen, hat der Unsicherheitskreis einen relativ großen Radius von 45 Pixeln auf dem Bildschirm. Um zunächst sicherzustellen, dass auch kleine Objekte korrekt von dem Verfahren gefunden werden können, wird zunächst jeder der sich in dem Kreis befindlichen Pixel analysiert. Dies führt allerdings auch zu einem hohen Rechenaufwand. Bei der Wahl einer Samplingstrategie, die nicht jeden Punkt prüft, kann es jedoch passieren, dass kleine Objekte zwischen je zwei Blickpunkte des Unsicherheitsbereiches fallen und somit nicht registriert werden.

Bei der Schnittpunktberechnung muss sichergestellt werden, dass die gefundenen Objektpunkte und die entsprechenden Objekte einander zugeordnet werden können, um somit beispielsweise auf entsprechende Materialeigenschaften Zugriff zu haben. Dies erfolgt anhand einer eindeutigen Identifikation (ID), wie beispielsweise der Objektbezeichnung oder einer eindeutigen Nummer, welche in einem Vorverarbeitungsschritt zusammen mit den Transparenzwerten und den Bounding-Boxes eines jeden Objektes bestimmt wird. Letztere sind für eine weitere Metrik von Bedeutung.

Nachdem alle Schnittpunkte für einen Blickpunkt ermittelt sind, werden sie der Reihe nach, vom Standpunkt der virtuellen Kamera bis hin zur hinteren Clippingebene, traversiert. Anhand der ID sind auch dem jeweiligen Objekt entsprechende Transparenzwerte abrufbar. Während der Traversierung wird in einem Vektor markiert, welches Objekt vom Blickstrahl getroffen wurde. Zudem erfolgt die Berechnung der Sichtbarkeit auf Basis der Transparenzwerte, welche alle folgenden, auf dem Strahl liegenden Schnittpunkte maximal haben können. Diese

nimmt durch semi-transparent verdeckende Objekte, mit zunehmender Tiefe der Schnittpunkte, ab. Bei einer Sichtbarkeit von 0.05 terminiert das Verfahren, da der entsprechende Oberflächenpunkt in diesem Fall nicht mehr wahrnehmbar ist. Sobald diese Prozedur für jedes Sample und den dazugehörigen Unsicherheitsbereich abgeschlossen ist, liegt der Vektor vor, dessen Länge der Anzahl aller dargestellten Objekte entspricht. Jede Komponente des Vektors beschreibt für einen Blickpunkt, ob das entsprechende Objekt innerhalb oder außerhalb des Unsicherheitsbereiches gelegen hat. Dies erfolgt entsprechend durch den Eintrag 1 oder 0. Diese Metrik kann damit Strukturen ausschließen, die sich außerhalb des Unsicherheitsbereiches des aktuellen Blickpunktes befinden. Die Werte des Vektors werden nun, wie oben beschrieben, in eine Wahrscheinlichkeit umgerechnet. Zusammengefasst ergeben die Vektoren aller Samples und damit aller Blickpunkte eine Matrix, welche zunächst intern gespeichert wird.

Aufgrund semi-transparenter Verdeckungen und der notwendigen Analyse des Unsicherheitsbereiches, sind in dieser Matrix zu jedem Zeitpunkt jedoch mehrere Objekte, gleichwahrscheinlich als fixiert markiert. Um tatsächlich fokussierte Strukturen genauer identifizieren zu können, sind daher weitere Metriken nötig. Gerade bei Animationen bietet es sich an, langsame Folgebewegungen des Auges zu nutzen. Es wurde in Abschnitt 3.1.3 erläutert, dass diese Augenbewegungen von den meisten ETS's nicht korrekt erkannt werden. Das ist zum aktuellen Zeitpunkt jedoch kein Nachteil, da die Auswertung der Metriken zunächst unabhängig von jeglicher Klassifizierung der Samples zu Fixierungen und Sakkaden durchgeführt wird. Erst bei der Zusammenfassung der Wahrscheinlichkeiten kann sich ein Verfahren zur korrekten Detektion von Folgebewegungen als vorteilhaft erweisen.

Objektbasierte Metriken: „Abstand des Blickpunktes vom Objekt“

Die Idee einer zweiten, objektbasierten Metrik besteht in dem Vergleich der Blickkoordinate jedes einzelnen Samples mit der Position eines jeden Objektes. Hierbei wird die Tatsache ausgenutzt, dass das Auge dem bewegten und fixierten Objekt genau folgt. Dabei bleibt der Abstand des Blickpunktes zum Objekt über die gesamte Folgebewegung sehr klein und vor allem konstant, während sich der Abstand zu den anderen Objekten verändert. Die Blickpunktposition kann dabei ohne weiteren Rechenaufwand aus der Protokolldatei entnommen werden. Für die Berechnung eines Wertes, der die Objektposition widerspiegelt, sind dagegen weiterführende Überlegungen notwendig.

In Abschnitt 4.1.1 wurde erwähnt, dass medizinische Visualisierungen u.a. kleine, kompakte Strukturen wie Lymphknoten oder Mandeln enthalten. Solche Strukturen lassen sich relativ gut durch Bounding-Boxes approximieren. Vor allem das Zentrum der Bounding-Box ist dem Objektzentrum solcher Strukturen relativ nahe. Da der Fokus für das entwickelte Verfahren zunächst auf kleinen, kompakten Objekten liegt, kann die Position des auf die Bildebene projizierten Bounding-Box-Zentrums zur Annäherung der Objektposition genutzt werden. Die Bounding-Box jeder Struktur wird in einem Vorverarbeitungsschritt, einmalig bestimmt. Da sich während des Experimentes ausschließlich die Kamera um die Szene bewegt, bleibt jedes Objekt und damit auch die jeweilige Bounding-Box an ihrer Position im Weltkoordinatensystem. Die zweite Metrik beschreibt also damit für jedes Objekt den Abstand des Blickpunktes vom projizierten Bounding-Box-Zentrum.

Wie bei der ersten Metrik wird für jedes Sample der Abstand zu allen Objekten in Form von Vektoren gespeichert. Diese einzelnen Vektoren werden ebenfalls zu einer Matrix zusammengefasst. Aus dieser Matrix wird der größte Abstand bestimmt. Anschließend erfolgt die

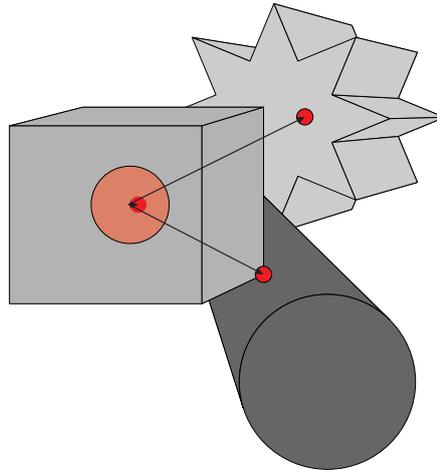


Abbildung 4.7: Schematische Darstellung der Idee hinter der objektbasierten Metrik „Abstand des Blickpunktes vom Objekt“. Die Abbildung stellt die Zentren der Bounding-Boxes für jedes Objekt als rote Punkte dar. Der rote, transparente Kreis verdeutlicht dabei den Blickpunkt mit Unsicherheitsbereich. Die Pfeile repräsentieren den Abstand dieses Punktes von den Bounding-Box-Zentren.

Subtraktion aller über die Experimentdauer ermittelten Abstände von diesem Wert. Um nun eine Wahrscheinlichkeit zu erhalten, müssen die Werte jeder Zeile der Matrix durch deren Summen dividiert werden. Das Ergebnis sind Wahrscheinlichkeiten, die bei großen Abständen klein und bei kleinen Abständen groß werden. Durch diese Metrik wird die Zuordnung von Fixierungen möglich, selbst wenn das entsprechende Objekt durch eine oder mehrere semi-transparente Strukturen verdeckt wird. Um diese Funktionalität auch für elongierte oder verzweigte Strukturen anwendbar zu machen, müssen jedoch alternative Möglichkeiten zur Beschreibung der Objektposition gefunden werden.

Objektbasierte Metriken: „Sichtbarkeit eines Objektes“

In den Anforderungen wurde die Bedeutung der Sichtbarkeit anatomischer Strukturen beschrieben. Je geringer diese aufgrund von opaken oder semi-transparenten Verdeckungen ist, desto unwahrscheinlicher ist es auch, dass entsprechende Objekte fixiert werden. Die Sichtbarkeit stellt daher eine dritte Metrik dar, welche im Folgenden mit dem mittleren, prozentualen Anteil der Objektfarbe innerhalb des Unsicherheitsbereiches, welche von einem Objekt sichtbar ist, definiert wird. Vordergrundobjekte haben daher eine Sichtbarkeit von eins. Die Sichtbarkeit aller hinter einem semi-transparenten Vordergrundobjekt liegenden Strukturen, ist dabei linear abhängig vom Transparenzwert der verdeckenden Struktur. Die Farben beider Objekte vermischen sich. Die Metrik wird nur im lokalen Umfeld eines jeden Samples ausgewertet. Wird dazu der zur Berechnung der Metrik „Objekt befindet sich im Unsicherheitsbereich“ genutzte Unsicherheitsbereich verwendet, kann die hier beschriebene Metrik im Rahmen der oben erläuterten Schnittpunktberechnung für jedes Sample und jedes Objekt mitberechnet werden.

Der Sichtbarkeitswert wird zu Beginn der Strahlenverfolgung auf den Wert eins gesetzt. Alle Vordergrundobjekte erhalten daher diesen Wert. Trifft der Strahl eines Blickpunktes (Primärstrahl) oder ein Strahl aus dessen Unsicherheitsbereich (Sekundärstrahl) auf ein Objekt, wird der neue Sichtbarkeitswert an dieser Stelle intern gespeichert und mit der alten Sichtbarkeit multipliziert. Alle folgenden, sich auf dem Strahl befindlichen Schnittpunkte, können daher nur noch maximal diese neue Sichtbarkeit besitzen. Für jeden Primär- und Sekundärstrahl sowie jedes Objekt wird dabei die Opazität nur für den ersten, gefundenen Schnittpunkt ermittelt. Der Mittelwert aller berechneten Opazitäten für ein Objekt, über die Fläche des Unsicherheitsbereiches, stellt dabei das Ergebnis dar. Für Objekte, die sich nicht in diesem Bereich befinden, ist dieser Wert null. Damit ist es für diese Metrik unwahrscheinlich, dass ein solches Objekt, zum entsprechenden Zeitpunkt fixiert wurde. Auch für diese Metrik muss anschließend die Umrechnung in eine Wahrscheinlichkeit erfolgen, indem alle Werte eines Zeitpunktes durch deren Summe dividiert werden.

Bildbasierte Metriken

Einen wichtigen Beitrag zur Objektbestimmung liefern folgende, bildbasierte Metriken, welche anhand der schematischen Abbildung 4.8 erläutert werden:

1. Fläche des Objektes bei verdeckungsfreier Projektion (A_c)
2. Tatsächlich projizierte Fläche (A_v)

Für die Berechnung des ersten Wertes muss jedes Objekt einzeln gerendert werden (vgl. Abbildung 4.8(b)). Aus den entstehenden Bildern wird die Anzahl der Pixel ermittelt, deren Farbe nicht dem Hintergrund entspricht. Dabei wäre es günstig jedes Objekt in einem eindeutigen Farbton (Objekt-ID) zu rendern, um zu verhindern, dass die Hintergrundfarbe in der Objektfarbe vorkommt. Für die Bestimmung des zweiten Maßes werden zunächst alle Objekte mit ihrer eindeutigen ID in ein eigenes Bild, der Größe des Ursprungsbildes, gerendert. Anschließend wird für jede Pixelposition ermittelt, welchem Objekt die sich dort befindliche Farbe entspricht. Wichtig hierbei ist, dass semi-transparent verdeckte Objekte, abhängig von deren Transparenzwerten als sichtbar bzw. nicht sichtbar eingestuft werden sollten, da diese auch erkennbar sein können, wenn sie von einem bzw. mehreren transparenten Objekten verdeckt werden.

Die genaue Vorgehensweise zur Berechnung dieser Informationen ist in NEUGEBAUER [2006] dokumentiert. Aus ihnen kann, für jedes Objekt, der nach Gleichung 4.1 berechenbare Verdeckungsgrad V ermittelt werden. Dieser gibt nach der Umrechnung in eine Wahrscheinlichkeit ebenfalls Aufschluss darüber, wie wahrscheinlich ein Objekt zu einem Zeitpunkt fixiert wurde. Die Wahrscheinlichkeit ist dabei so zu bestimmen, dass sie bei einer hohen Verdeckung kleiner wird.

$$V = \frac{A_c}{A_v} \quad (4.1)$$

In den Anforderungen wurde die häufige Erkennung großer Objekte durch das hier beschriebene Verfahren erläutert. Beispiele hierfür waren der Schädel oder das Lebergewebe. Um eine häufige Erkennung solcher Strukturen zu vermeiden und damit die durch das Verfahren bestimmte Menge potentiell fixierter Objekte zu reduzieren, bietet sich eine Kombination aus

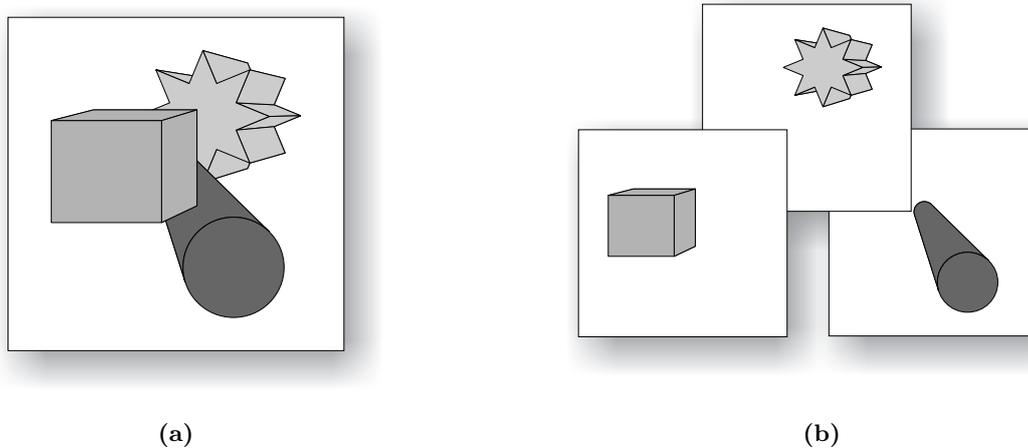


Abbildung 4.8: In Abbildung (a) ist eine fiktive Szene zu sehen. Abbildung (b) stellt die Objekte einzeln gerendert dar, so dass die Anzahl der Pixel bestimmt werden kann, die nicht der Hintergrundfarbe entsprechen.

deren Eigenschaften Größe und Transparenz als weitere Metrik an. Die Größe der Szenenobjekte wird durch das bildbasierte Maß A_c beschrieben und deren Transparenzwert kann in einem Vorverarbeitungsschritt ermittelt und gespeichert werden. Nach der Normierung der Objektgrößen über alle Samples auf einen Wert zwischen $[0 \dots 1]$ können die beiden Metriken mittels der Gleichung 4.2 zu einem neuen Maß M kombiniert werden. Hierfür werden Schwellenwerte ($S_{\text{Transparenz}}$ und $S_{\text{Größe}}$) verwendet, die zunächst relativ hoch angesetzt werden sollten (z.B. 0.95), so dass große und hochtransparente Strukturen die Wahrscheinlichkeit null zugewiesen bekommen. Anschließend erfolgt, äquivalent zu allen anderen Metriken, noch die Umwandlung in eine Wahrscheinlichkeit. Aus diesem Maß resultiert für große und hochtransparente Strukturen eine niedrigere Gesamtwahrscheinlichkeit nach der Kombination aller Metriken, welche in Abschnitt 4.2.3 beschrieben wird.

$$M = \begin{cases} 0, & \text{Größe} \geq S_{\text{Größe}} \ \&\& \ \text{Transparenz} \geq S_{\text{Transparenz}} \\ 1, & \text{Größe} < S_{\text{Größe}} \ || \ \text{Transparenz} < S_{\text{Transparenz}} \end{cases} \quad (4.2)$$

4.2.3 Kombination aller Metriken

Nachdem die Metriken für alle Samples des Datenprotokolls berechnet und in die Form einer Wahrscheinlichkeit überführt wurden, trifft jedes einzelne Maß eine Aussage drüber, ob ein Objekt zu einem Zeitpunkt potentiell vom Betrachter fixiert wurde. Einzeln haben diese Maße jedoch nur eine geringe Aussagekraft. Beispielsweise wird sich, aufgrund des Verdeckungsgrades noch kein Objekt eindeutig eingrenzen lassen, so dass eine Entscheidung über dessen Fixierung möglich ist. Durch die zusätzliche Anwendung weiterer und vor allem lokaler Metriken (Unsicherheitsbereich, Abstand des Blickpunktes vom Objekt, lokale Sichtbarkeit), kann die Menge potentiell fixierter Objekte weiter eingeschränkt werden. Aus diesen Gründen wurde es als zweckmäßig erachtet, die Einzelmetriken mit Gewichten zu versehen und sie zu einem Wahrscheinlichkeitsmaß zusammenzufassen. Die Wahrscheinlichkeit, dass ein Objekt o zu einem Zeitpunkt t fixiert wurde, ergibt sich somit aus der Summe von K normierten

und mit den Gewichten w_k versehenen Metriken m_{kot} , ($k = 1 \dots K$). Sie kann durch die Summenformel 4.3 ausgedrückt werden.

$$P_{ot} = \sum_{k=1}^K w_k * m_{kot} \quad \left(\sum_{k=1}^K w_k = 1 \right) \quad (4.3)$$

Jedes Maß trägt damit einen Teil zur Erhöhung der Gesamtwahrscheinlichkeit des tatsächlich fixierten Objektes bei. Strukturen, die objektiv gesehen unwahrscheinlich fixiert wurden, weil sie zu weit vom Blickpunkt entfernt oder zu stark verdeckt sind, erhalten dagegen eine niedrige Gesamtwahrscheinlichkeit. Dieses Verhalten kann weiterhin durch die Gewichte gesteuert werden. Jedes dieser Gewichte wurde zunächst auf den Wert $1/K$ gesetzt, was die Mittelwertbildung der Wahrscheinlichkeiten aller Metriken zur Folge hat. Dies wurde als sinnvoll eingeschätzt, da bislang noch keine Aussagen über den Einfluss unterschiedlicher Wichtungen auf das Endergebnis getroffen werden können. Im Anschluss an die Berechnung der Gesamtwahrscheinlichkeiten erfolgt die Datenfilterung.

4.2.4 Datenfilterung

Das Filtern eines Bild- oder Tonsignals dient der Entfernung oder der Generierung neuer Informationen durch Interpolation. Ein Fotofilter aus der Optik blockiert gewisse Wellenlängen des Lichtes, so dass andere Farbtöne hervorgehoben werden. In der Visualisierung umfasst der Filterungsprozess ebenfalls eine Datenreduktion aber auch die Durchführung von Fehlerkorrekturen und die Beseitigung von Ausreißern (Datenglättung). Alle Anwendungen haben die Gemeinsamkeit, die Daten auf weitere Verarbeitungsschritte vorzubereiten. Aus diesem Grund erfolgt die Filterung direkt nach der Datenaquisition.

Das Eye-Tracking-Protokoll wird ebenfalls gefiltert, nachdem alle Daten aufgenommen und die Metriken generiert wurden. Dieser Schritt gehört damit zur zweiten Stufe des 3-Phasen-Modells. Nach dem Filterungsprozess bleiben die Samples übrig, die zu einer Fixierung gehören und innerhalb des Bildschirms liegen. Diese Vorgehensweise wurde einerseits gewählt, da Fixierungen das Hauptinteresse des Betrachters widerspiegeln. Andererseits machen sie den Großteil des Datenprotokolls aus, da Sakkaden nur eine Dauer von wenigen Millisekunden haben. Im Einzelnen werden folgende Informationen gefiltert, die für die weitere Verarbeitung unnötig sind:

- Nachrichten des ETS
- Samples in denen das Auge nicht erkannt werden konnte (z.B. Lidschlag)
- Sakkaden
- Fixierungen außerhalb des Bildschirms

In Abschnitt 3.1.3 wurde erwähnt, dass das Auge nach einer Sakkade auf die Mitte des für den Betrachter interessanten Objektes bzw. der Objektregion positioniert wird. In der darauf folgenden Fixierung wird damit die gesamte Zeit diese Region untersucht. Anstatt alle Samples einer Fixierung zu analysieren, ergibt sich damit die Möglichkeit, nur die Zielkoordinate einer jeden Sakkade zu untersuchen. Für die Bestimmung der Schnittpunkte und die zusätzliche Betrachtung des Unsicherheitsbereiches würde dies zu einer Verringerung der Rechenzeit

führen. Da jedoch im Allgemeinen Fixierungen das Interesse des Betrachters widerspiegeln, wurden diese für die Berechnung der Metriken und zur Auswertung genutzt.

Während des Experimentes kann es vorkommen, dass der Proband zeitweise nicht auf den Bildschirm schaut. Gründe dafür können Langeweile oder fehlerhafte Messungen während des Lidschlussreflexes sein. Letzteres wird von vielen ETS's erkannt und als solches gekennzeichnet. Ersteres kann angenommen werden, wenn Fixierungen häufig außerhalb des Bildschirms liegen. Da einem Experiment jedoch meistens eine konkrete Aufgabe zugrunde liegt, wird hier davon ausgegangen, dass sich die Probanden auch auf diese konzentrieren und der Blick auf den Bildschirm gerichtet bleibt. Um Einflüsse auf die eigentlich wichtigen Daten auszuschließen, werden Blickpunkte außerhalb des Displays aus dem weiteren Verarbeitungsprozess exkludiert.

Im weiteren Verlauf soll kurz auf einen sich bietenden Vorzug des binokularen Eye-Trackings eingegangen werden (Verwendung beider Kameras). Wie oben beschrieben, umfasst die Filterung ebenfalls die Interpolation neuer Daten im Falle von fehlenden Informationen. Das ETS ist unter einigen Umständen nicht in der Lage, das Auge korrekt zu erfassen, was zu Datenverlusten führt. Für beide Augen gehen beim binokularen Tracking jedoch unterschiedlich viele Samples verloren. Aus diesem Grund kann fehlende Information durch die Daten des jeweils anderen Auges ausgeglichen werden. Dieses Vorgehen wird in SENSOMOTORIC INSTRUMENTS GMBH [1999] detailliert beschrieben. Da die binokulare Kalibrierung, aufgrund einer längeren Dauer, jedoch belastend für die Versuchsperson sein kann, muss überlegt werden, ob diese Möglichkeit des Datenausgleichs ein qualitativ schlechteres Gesamtergebnis rechtfertigt. Aufgrund von Konzentrationsschwierigkeiten nach einer langen Kalibrierung, vor allem mit auf dem Kopf befestigten Systemen, könnten solche Effekte auftreten. Soll die beschriebene Form der Datenrekonstruktion genutzt werden, so muss dies vor der Berechnung der Metriken erfolgen, da die interpolierten Koordinaten sonst nicht berücksichtigt werden.

4.2.5 Zusammenfassung der Wahrscheinlichkeiten und globale Analyse

Das Ergebnis aus Phase eins ist eine Wahrscheinlichkeit, die vom Zeitpunkt und vom Objekt abhängt. Dabei werden aufgrund der Datenfilterung nur die Zeitpunkte betrachtet, in denen eine Fixierung vorliegt. Dieses Ergebnis dient als Grundlage zur weiteren Verarbeitung der Daten in Phase zwei. Dort erfolgt, nach der Filterung, zunächst die Zusammenfassung der Wahrscheinlichkeiten aller Objekte über die Samples innerhalb jeder Fixierung durch Mittelwertbildung. Tabelle 4.3 verdeutlicht dieses Vorgehen an einem Beispiel mit drei Objekten und drei Fixierungen. Jede Fixierung besteht hier aus zwei Samples. Die Tabelleneinträge repräsentieren die durch Gleichung 4.3 ermittelten Gesamtwahrscheinlichkeiten aus den einzelnen Metriken.

Die Metriken sind so konzipiert, dass sie eine hohe Wahrscheinlichkeit liefern, wenn ein Objekt im Fokus liegt, bzw. gut sichtbar ist. Daher wird in diesen Fällen auch die Gesamtwahrscheinlichkeit hoch sein. Es wird nun weiter davon ausgegangen, dass eine Sakkade und damit eine neue Fixierung erfolgt, sobald sich eine Metrik drastisch ändert. Wird beispielsweise ein Objekt zunehmend verdeckt, kann es nicht mehr korrekt fixiert werden. Das bedeutet wiederum, dass während einer Fixierung die Metriken relativ stabil sind. Aus diesem Grund liegt der Zusammenfassung der Wahrscheinlichkeiten die Mittelwertbildung zu Grunde.

Das nach der Zusammenfassung resultierende Ergebnis ist die Wahrscheinlichkeit, dass Objekt o während der Fixierung f fokussiert wurde. Dies stellt bereits ein vorläufiges Ergebnis dar, woraus die Objekte mit den höchsten Wahrscheinlichkeiten während einer Fixierung ermittelt

Fixierung	P(Object 1)	P(Object 2)	P(Object 3)
1	0.574	0.071	0.355
1	0.621	0.059	0.320
2	0.252	0.725	0.023
2	0.245	0.678	0.077
3	0.069	0.613	0.319
3	0.059	0.621	0.320



Fixierung	P(Object 1)	P(Object 2)	P(Object 3)
1	0.598	0.065	0.338
2	0.436	0.392	0.172
3	0.248	0.702	0.050

Tabelle 4.3: Beispielhafte Darstellung der Zusammenfassung aller Samples der Fixierungen eins bis drei. Die obere Tabelle zeigt die zu den einzelnen Samples gehörenden Wahrscheinlichkeiten, nach Gleichung 4.3. In der unteren Tabelle sind diese für jede Fixierung zusammengefasst.

werden können. Es besteht nun weiterhin die Möglichkeit, im Rahmen der dritten Stufe, globale Betrachtungen der Fixierungswahrscheinlichkeiten durchzuführen. Hierfür eignet sich beispielsweise die Fixierungslänge, die das Interesse des Betrachters an der fixierten Region bzw. dem fixierten Objekt widerspiegelt. Daher ist es sinnvoll, langen Fixierungen mehr Bedeutung zuzuweisen als kürzeren. Weitere Informationen, welche im Rahmen einer globalen Analyse untersucht werden können sind:

1. Identifikation von Objekten, die in aufeinanderfolgenden Fixierungen eine hohe Wahrscheinlichkeit erhalten haben
2. Identifikation von Objekten, die im Großteil aller Fixierungen eine hohe Wahrscheinlichkeit erhalten haben
3. Im Falle von gleichen bzw. ähnlichen Wahrscheinlichkeiten: Entscheidung für das Objekt, welches über alle Fixierungen die höchste mittlere Wahrscheinlichkeit besitzt
4. Fixierungsdauer am wahrscheinlichsten fokussierter Objekte
5. Fixierungsdauer eines „Objects of Interest“ (vgl. Abschnitt 2.2.3)
6. Durchschnittliche Wahrscheinlichkeit eines Objektes über alle Fixierungen
7. Einen Vektor der für jede Fixierung die ID des am wahrscheinlichsten fixierten Objektes enthält
8. Einen Vektor der für ein Objekt die Position innerhalb des Wahrscheinlichkeitsranges aller Fixierungen angibt
9. Die Häufigkeit mit der für ein Objekt eine bestimmte Wahrscheinlichkeit erreicht wird

4.3 Zusammenfassung

In diesem Kapitel wurde der Gesamtprozess zur Verarbeitung eines Samples bis hin zur Fixierungswahrscheinlichkeit beschrieben. Die dazu notwendigen Schritte können mittels des entwickelten 3-Phasen-Modells beschrieben werden. Im ersten Teil des Kapitels erfolgte eine Charakterisierung der aus einem Eye-Tracking-Experiment resultierenden Daten, mit deren Hilfe im zweiten Schritt die Metriken aus Tabelle 4.4 berechnet werden können. Für alle Metriken wurden die notwendigen Schritte zur Umrechnung in eine Wahrscheinlichkeit erläutert, so dass alle Maße einfach zu einer Gesamtwahrscheinlichkeit kombiniert werden können. Aufgrund der Anpassung der Abstandsmetrik (vgl. Nummer zwei in Tabelle 4.4) an kleine, kompakte Strukturen einer medizinischen Visualisierung ist das Verfahren bislang ebenfalls nur zur Erkennung von Fixierungen entsprechender Strukturen geeignet.

Nummer	Bezeichnung der Metrik
1	Objekt befindet sich im Unsicherheitsbereich
2	Abstand von Objekt und Blickpunkt
3	Lokale Sichtbarkeit eines Objektes
4	Anzahl Pixel bei verdeckungsfreier Projektion
5	Tatsächliche Anzahl Pixel bei Darstellung aller Objekte
6	Verdeckungsgrad
7	Kombination aus Größe und Transparenz

Tabelle 4.4: Zusammenfassung der im Konzeptteil vorgestellten Metriken zur Objektbestimmung.

Der Forderung nach Beseitigung von Doppeldeutigkeiten, durch semi-transparente Verdeckungen, wurde mit der Analyse von Folgebewegungen, unter der Bedingung einer animierten Szene, begegnet. Da jedoch die Animation nicht die einzige Möglichkeit zur Präsentation einer Visualisierung darstellt, erfolgte ebenfalls die Beschreibung der Vorzüge von interaktiven Darstellungen und Bildern.

Da es bislang wenig Informationen bezüglich der Objektidentifizierung mit Hilfe von Eye-Tracking-Daten gibt und somit nur schwierig Hypothesen bezüglich auftretender Blickmuster ableitbar sind, wird in den Evaluierungsexperimenten ein kombinierter Ansatz aus Top-Down- und Bottom-Up-Analyse verwendet. Dieser erscheint sinnvoll, da die gefundenen Eye-Tracking-Studien, welche das Blickverhalten bei der Betrachtung animierter bzw. dreidimensionaler Stimuli untersuchen, sich auf recht primitive Szenarien beschränken und nicht das Blickverhalten bei der Betrachtung komplexer Visualisierungen untersuchen. Eine Ausnahme hierfür bildet die Analyse von Blickmustern in Computerspielen. Entsprechende Erkenntnisse stehen jedoch noch am Anfang ihrer Entwicklung und werden in den Arbeiten SENNERSTEN und LINDLEY [2008] sowie STELLMACH [2007] beschrieben.

Das vorgestellte Konzept enthält Lösungsansätze für verschiedene Problemstellungen, die sich bei der Objektidentifizierung ergeben. Es stellt jedoch noch keine finale Lösung dar. Aus diesem Grund wurde an mehreren Stellen dieses Kapitels bereits angedeutet, wie das hier vorgestellte Verfahren weiter verbessert und angepasst werden kann. Diese sowie weitere Ausblicke sind in Kapitel 7 detailliert erläutert.

Kapitel 5

Implementierung

Der Fokus dieses Kapitels liegt auf der Erläuterung der Umsetzung des Verfahrens zur Zuordnung von Blickpunkten zu konkreten Objekten der Szene. Dies erfolgt am Beispiel des Aufbaus einer typischen Protokolldatei, der Implementation zweier im Konzept vorgestellter Metriken sowie deren Auswertung. Neben der Vorstellung der dafür notwendigen Werkzeuge MEVISLAB, OPEN INVENTOR und MATLAB erfolgt außerdem die Beschreibung spezieller Eigenschaften des verwendeten Eye-Tracking-Systems (ETS) EYELINK I.

5.1 Vorstellung der Entwicklungswerkzeuge

Die Wahl der zentralen Entwicklungsumgebung fiel auf MEVISLAB [MEVISLAB], da diese zur Darstellung komplexer, medizinischer Visualisierungen geeignet ist und umfangreiche Möglichkeiten zur Verarbeitung der zugehörigen Datensätze bietet. Dabei übernimmt die integrierte Graphikbibliothek OPEN INVENTOR [WERNECKE, 1993, 1994] das Rendern der Modelle sowie die Verwaltung von Interaktionen. Zur Auswertung und zur graphischen Präsentation der Ergebnisse des Verfahrens erschien es sinnvoll, die Daten in die leistungsfähige Analyseapplikation MATLAB [MATLAB] zu exportieren.

5.1.1 MeVisLab

MEVISLAB ist eine Entwicklungsumgebung, gekoppelt mit einer graphischen Benutzerschnittstelle (GUI). Sie ist ausgelegt für die medizinische Bildverarbeitung sowie zur Datenvisualisierung. Außerdem eignet sich MEVISLAB zum schnellen Testen von Applikationskonzepten. Da Programme als Netzwerke von miteinander verknüpften Modulen repräsentiert werden, ist Anwendungsentwicklung intuitiv und schnell. MEVISLAB unterscheidet Makro-, *MeVis Image Processing Library*- (ML-) und Inventormodule. Des Weiteren ist die Entwicklung auf C++-, Netzwerk- und Skriptebene möglich.

Auf C++-Ebene können neue Inventor- und ML-Module durch Ableitung entsprechender Basisklassen generiert und mit Feldern versehen werden. Felder dienen der Steuerung modulin-terner Funktionen aus MEVISLAB heraus. Sie repräsentieren Datentypen bzw. -strukturen in Form von Klassen und ermöglichen eine ständige Überwachung entsprechender Werte. Dies wird durch deren Verbindung mit Sensoren realisiert. Felder gleicher Klassen können aber auch untereinander verbunden werden, um deren Werte synchron zu halten. Sie erlauben zudem die Definition von Ein- und Ausgängen, über welche der intermodulare Datenaustausch erfolgt.

Auf Netzwerkebene lassen sich Module, abhängig von deren Ein- und Ausgängen, zu Netzwerken verbinden. Dadurch wird festgelegt, welche Module gegenseitig Daten austauschen dürfen. Die Verbindung von Feldern aus MEVISLAB und OPEN INVENTOR auf C++-Ebene ist schwierig, da diese auf verschiedenen Klassen basieren, auch wenn sie den selben Datentyp kapseln. Auf Netzwerkebene ist dies jedoch einfach realisierbar, weil MEVISLAB die korrekte Synchronisierung der jeweiligen Werte verwaltet.

Die Skriptebene steht in enger Verbindung zu Makromodulen, welche Kontainer für in sich geschlossene Netzwerke repräsentieren. Sie werden direkt in MEVISLAB erzeugt und können ebenfalls Felder besitzen, welche auch hier zur Steuerung interner Prozesse dienen. Felder können allgemein über die GUI des entsprechenden Moduls verändert werden. Die Variablen und Funktionen aller Module sowie des internen Netzwerkes eines Makromoduls sind zudem über Skripte der Sprachen PYTHON oder JAVASCRIPT automatisiert steuerbar. Innerhalb von Makromodulen sind weitere Makromodule aber auch Inventor- und ML-Module denkbar. ML-Module dienen hauptsächlich der medizinischen Bildverarbeitung.

5.1.2 Open Inventor

Weiterhin gibt es Inventormodule, welche die Funktionen von OPEN INVENTOR kapseln. Sie sind auf C++-Ebene u.a. von den Basisklassen `SoMaterial`, `SoOrthographicCamera`, `SoTriangleStripSet` und `SoSeparator` ableitbar. Diese Module sind Teil eines Szenengraphen und beschreiben, entsprechend ihrer dortigen Position, wie die Szene dargestellt wird. Die Traversierung des Graphen erfolgt von einem beliebigen Knoten aus immer von oben nach unten und von links nach rechts. Beim Rendern wird der Graph von der Wurzel aus, in dieser Reihenfolge, durchlaufen.

Abhängig von der Basisklasse bieten Inventormodule verschiedene Funktionalität an. `SoShape`-Knoten definieren neue Geometrie und `SoMaterial`-Knoten speichern Materialeigenschaften, wie Transparenz. `SoSeparator`-Module, dienen der Gruppierung von Teilgraphen und dem Speichern sowie Wiederherstellen aktueller Zustände, wie Beleuchtungseinstellungen oder Transformationen, während der Traversierung. In einer solchen Gruppe befindliche Knoten sind nur lokal gültig und haben keinen Einfluss auf die Verarbeitung des restlichen Szenengraphen.

OPEN INVENTOR bietet besondere Klassen des Typs `SoAction` an. Sie traversieren den Graphen, führen Berechnungen durch und verwalten gleichzeitig Informationen bezüglich der Szene. Letztere werden in speziellen Methoden zur Verfügung gestellt. Actions lassen sich auf Gruppenknoten oder Pfade anwenden. Im ersten Fall wird der gruppierte Teilgraph traversiert. Bei der Anwendung auf einen Pfad werden nur die darin enthaltenen Knoten besucht, anstatt den Baum wie üblich zu durchlaufen.

Die Graphikbibliothek enthält außerdem ein Eventkonzept zur Registrierung von Interaktionen. Neben der Szenentransformation bzw. -exploration ist auch die weitere Verarbeitung dieser Benutzereingaben möglich. Beispielsweise gibt es Funktionen zur Projektion von Bildschirmkoordinaten auf 3D-Strahlen, was eine Grundvoraussetzung für die Schnittpunktberechnung mit der Szenengeometrie und die Objektbestimmung bildet. Zudem lässt sich in OPEN INVENTOR mit Vektoren und Matrizen rechnen. Dadurch wird die Entwicklung effizienterer Methoden zur Objektbestimmung möglich.

5.1.3 MATLAB

MATLAB ist ein auf Vektor- und Matrixoperationen spezialisiertes Programm zur Datenanalyse und enthält diesbezüglich viele integrierte Funktionen. Durch die Erweiterung um *Toolboxes* eröffnen sich Anwendungsmöglichkeiten u.a. in der Statistik, Bildverarbeitung oder Simulation. Dagegen ist jedoch auch die Basisversion vielseitig bezüglich der statistischen Analyse. Sie enthält eine Vielzahl an Möglichkeiten zur graphischen Datenaufbereitung. Komplexe Abläufe lassen sich zudem über Skripte automatisieren.

Als Analysetool kann MATLAB verschiedenste Datenformate lesen und schreiben. Tabellarisch strukturierte Textdateien sowie Binärdaten sind mittels einfacher Befehle importierbar. Eine andere Möglichkeit zum Datenimport stellt deren Bereitstellung als MATLAB-Skript dar, in welchem sie direkt in Form von Vektoren oder Matrizen gespeichert werden. Skripte können in anderen Skripten aufgerufen werden, was den Datenimport vereinfacht.

5.2 Das Eye-Tracking-System

Die für das hier verwendete ETS EYELINK I besonderen Kennwerte werden zunächst an einigen Beispielen aus Tabelle 5.1 vorgestellt. Anschließend folgt eine Beschreibung zum Aufbau des ETS sowie notwendiger Schritte zur Experimentvorbereitung. Den letzten Teil dieses Abschnitts bildet die Erläuterung der Integration von EYELINK I in MEVISLAB.

Bezeichnung	Wert
zeitliche Auflösung	250 Hz
Genauigkeit	0.5 – 1.0 °
Auge-Display-Distanz	40 – 240 cm
Kamera-Auge-Distanz	4 – 7 cm
Blickpunktauflösung	0.005 °
kleinste erkennbare Sakkade	0.3 °

Tabelle 5.1: Technische Eigenschaften des Eye-Tracking-Systems EYELINK I nach [SENSOMOTORIC INSTRUMENTS GMBH, 1999]

5.2.1 Vorstellung von EyeLink I

EYELINK I ist ein videobasiertes, auf dem Kopf befestigtes ETS. Seine technischen Eigenschaften (vgl. Tabelle 5.1) haben direkten Einfluss auf die Durchführung von Experimenten sowie auf Anwendungsmöglichkeiten des Systems. Aufgrund der hohen zeitlichen Auflösung und der damit verbundenen Erkennung kleiner Sakkaden [KARN, 2000] eignet es sich gut zur Registrierung von Folgebewegungen. Die Genauigkeit von EYELINK I ist für die zuverlässige Registrierung fokussierter Objekte in medizinischen Darstellungen jedoch relativ gering, da anatomische Strukturen teilweise relativ klein sind und nicht mit Sicherheit dem korrekten Blickpunkt zugeordnet werden können.

Während die Hochgeschwindigkeitskameras des ETS alle vier Millisekunden Bilder der Augen aufzeichnen, betrachtet die Versuchsperson den präsentierten Stimulus auf dem Monitor des Subject-PC's (SPC) (vgl. Abbildung 5.1). Die Signale werden zum Operator-PC (OPC) weitergeleitet, auf dem Schritte zur Pupillendetektion und zur Erkennung von Fixierungen

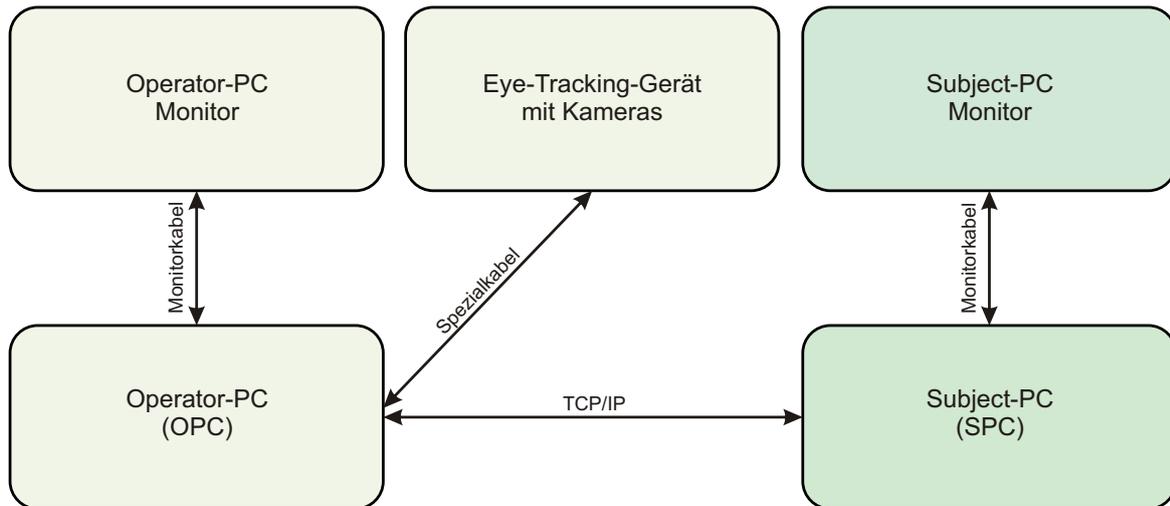


Abbildung 5.1: Schematischer Aufbau des Eye-Tracking-Systems EYELINK I.

stattfinden. Auf dem OPC werden anschließend alle relevanten Daten in eine Protokolldatei mit der Dateierweiterung EDF geschrieben. Wird diese nach dem Experiment zum SPC übertragen, kann die Objektidentifizierung zu einem beliebigen Zeitpunkt, auf diesem Computer, durchgeführt werden. Über das Interface des OPC lässt sich das Experiment zusätzlich überwachen und steuern.

Setup und Kalibrierung

Zunächst ist der Eye-Tracker mit Hilfe von Bändern am Kopf des Probanden zu befestigen. Zwei oder eine Kamera nehmen anschließend Bilder der Augen auf. Infrarot-LED's schaffen die nötige Ausleuchtung entsprechender Bereiche. Die LED's erzeugen außerdem sogenannte Reflexionen auf der Hornhaut. Zwischen Reflexions- und Pupillenzentrum werden Vektoren V_Z gebildet, die sich bei größeren Augenbewegungen verändern und bei kleineren Kopfbewegungen relativ konstant bleiben [DUCHOWSKI, 2007]. Sie dienen während der Kalibrierung und während des Experimentes der eindeutigen Erkennung aller Blickrichtungen.

An EYELINK I ist eine zusätzliche Kamera zur Registrierung der Kopfposition angebracht. Aus dieser Information lassen sich aus kopfrelativen Blickinformationen absolute Bildschirmkoordinaten berechnen [SENSOMOTORIC INSTRUMENTS GMBH, 1999]. Damit bietet das Verfahren, den Vorteil der sofortigen Verfügbarkeit von Bildschirmkoordinaten und erspart dem Anwendungsprogrammierer zusätzliche Arbeit.

Sobald der Eye-Tracker am Kopf des Probanden angebracht ist, wird die Kalibrierungsroutine gestartet. Hierbei ist es wichtig, dass die Kameras die Pupillen bei jeder Blickrichtung erkennen. Nun können Fokus der Kameras sowie Schwellwerte zur Pupillensegmentierung eingestellt werden. Die Kalibrierung erzeugt anschließend eine Abbildung von den Vektoren V_Z auf die Bildschirmkoordinaten des Displays. Bei EYELINK I erscheinen dazu zufällig Marker auf einem regelmäßigem Gitter der Größe 3×3 (Abbildung 5.2). Jeder Zielreiz muss ca. eine Sekunde fixiert werden, bis zufällig ein anderer Marker eingeblendet wird.

Der OPC berechnet während jeder Fixierung einen Vektor V_M von den bekannten Markerpositionen zu den aufgenommenen, unkalibrierten Fixierungspositionen. Für Punkte an denen

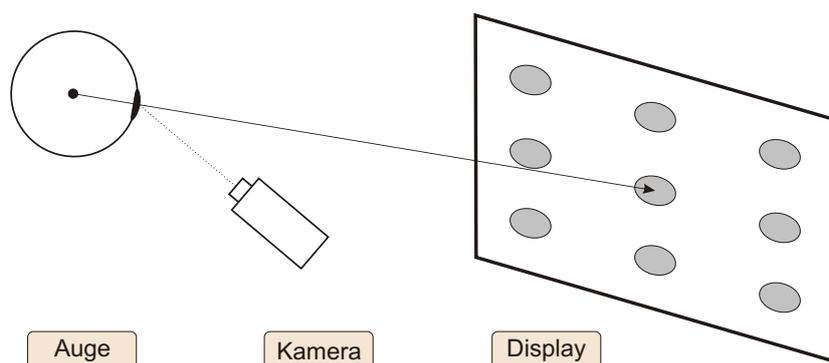


Abbildung 5.2: Darstellung der zweidimensionalen Kalibrierung nach [DONNER, 2007].

kein Marker zu sehen war, wird V_M durch Interpolation bestimmt. Da jeder Vektor einer bestimmten Blickrichtung entspricht, ist das System damit in der Lage, Bildschirmkoordinaten zu allen gemessenen Augenstellungen zu bestimmen.

Detektion von Fixierungen und Sakkaden

EYELINK I verwendet einen Algorithmus, der Sakkaden durch drei Schwellenwerte aus den Samples bestimmt. Dazu gehören der räumliche Abstand aufeinanderfolgender Blickpunkte (*Weg*) sowie die *Geschwindigkeit* und die *Beschleunigung* zu jedem Aufnahmezeitpunkt. Die Werte sind in Tabelle 5.2, für zwei Arten der Stimuluspräsentation dargestellt. Für die Klassifizierung eines Samples als Sakkade müssen alle drei Werte überschritten werden. Fixierungen ergeben sich dann aus jeweils allen Samples zwischen zwei Sakkaden.

Schwellenwert	animiert	statisch	Einheit
Weg	0	0.15	[°]
Geschwindigkeit	22	30	[°/s]
Beschleunigung	4000	8000	[°/s ²]

Tabelle 5.2: Schwellenwerte zur Detektion von Sakkaden bei animierten und statischen Stimuli [SENSOMOTORIC INSTRUMENTS GMBH, 1999].

Bei der Präsentation statischer Stimuli muss das ETS räumlich längere Sakkaden erkennen. Dazu gehören beispielsweise diejenigen, welche beim Lesen am Ende jeder Zeile auftreten. Geschwindigkeit und Beschleunigung der Blickbewegungen sind dabei höher, was zu größeren Schwellenwerten führt. Der „Weg“ bezeichnet den Abstand, den eine vom ETS aufgenommene Blickpunktposition von der mittleren Position der bisher betrachteten Samples haben muss.

Sind in einem Experiment Folgebewegungen von Bedeutung, so ist nach SENSOMOTORIC INSTRUMENTS GMBH [1999] die Anwendung der Schwellenwerte aus Spalte zwei der Tabelle 5.2 empfehlenswert. Mit diesen Parametern werden Sakkaden erkannt, die bis zu 0.3° klein sind. Das führt zur Detektion von mehr Sakkaden als mit den anderen Schwellenwerten. Eine Implikation dessen ist die verringerte Dauer sowie eine erhöhte Anzahl von Fixierungen.

5.2.2 Kopplung von EyeLink I und MeVisLab

Im Vorfeld dieser Arbeit wurde die Verwendung des ETS in der Entwicklungsplattform MEVISLAB ermöglicht. Ziel dabei war es, die wichtigsten Funktionen von EYELINK I aus MEVISLAB heraus steuern zu können. Da die Programmierschnittstelle von EYELINK I erlaubt, Konfigurationsnachrichten vom SPC über eine Netzwerkverbindung an den OPC zu senden, ist dieser Vorgang ohne weiteres möglich. Hierfür wurden die in Abbildung 5.3 dargestellten Netzwerke, in MEVISLAB, umgesetzt.

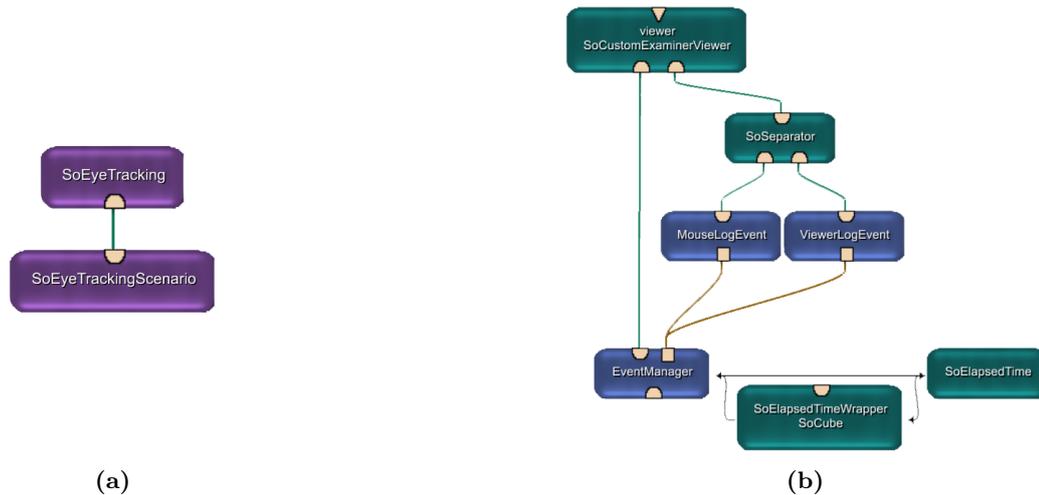


Abbildung 5.3: Die in MEVISLAB umgesetzten Netzwerke zur Ansteuerung von EYELINK I und zur Darstellung der Stimuli sowie der Blickpunktvisualisierungen. Abbildung (b) zeigt das interne Netzwerk des Markomoduls `SoEyeTracking` aus Abbildung (a).

Das Netzwerk der Abbildung 5.3(a) besteht aus den Makromodulen `SoEyeTracking` und `SoEyeTrackingScenario`. Die Module `MouseLogEvent` und `ViewerLogEvent` dienen der Registrierung von Maus- und Kameraevents. Sie sind mit dem zentralen `EventManager` verbunden. In diesem finden z.B. das Starten und Beenden eines Experimentes oder die Erzeugung der finalen Protokolldatei statt. Auf der Basis letzterer steuert das Modul das Playback von Experimenten und überlagert diese Animation zusätzlich mit der entsprechenden Blickpunktvisualisierung. Abbildung 5.4(a) zeigt einen Ausschnitt aus einer solchen Animation.

Auf der Scriptebene des `EventManagers` wurde die Funktionalität implementiert, mit den Tastenkombinationen ALT-F in den Vollbildmodus und mit ALT-N in den Fenstermodus zu wechseln. Dadurch kann der gesamte Bildschirm zur Darstellung des Stimulus genutzt werden. Durch ALT-F wird gleichzeitig das ETS in den Setupmodus versetzt und durch ALT-N wird das Experiment schließlich beendet. Über die GUI des `EventManagers` lässt sich außerdem ein Pfad einstellen, unter dem einerseits die Protokolldatei gespeichert wird und von wo andererseits die Daten für das Playback geladen werden.

Die Playbackfunktion ist im Rahmen des Bottom-Up-Ansatzes nützlich, um den Blickverlauf zu beobachten und den Genauigkeitsfehler abzuschätzen. Eine aussagekräftigere Visualisierung ist in Abbildung 5.4(b) zu sehen. Blickstrahlen und Kamerapfad sind hier Teil der Szene. Die Strahlen haben die Farbe schwarz wenn sie kein Objekt durchstoßen. Andernfalls ist die Strahlfarbe mittels der Temperaturfarbskala, nach dem Zeitstempel des jeweiligen Samples, kodiert [SCHUMANN und MÜLLER, 2000]. Diese Visualisierung kann exploriert werden und gibt Aufschluss über Fixierungen und das Ausmaß des Genauigkeitsfehlers.

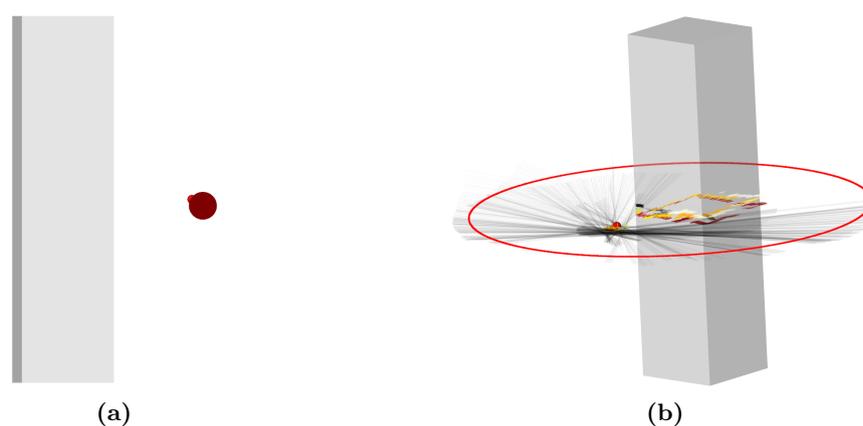


Abbildung 5.4: Darstellung von zwei Visualisierungsmöglichkeiten für Blickpunktdaten. Abbildung (a) ist ein Ausschnitt aus dem Playback eines Experimentes mit überlagerterem Blickpunkt. Abbildung (b) stellt die als Strahlen in die Szene projizierten Blickpunkte dar und visualisiert den den Kamerapfad als Folge roter Kreise.

Vor dem Starten der der Objektidentifikation muss die zum Protokoll passende Szene an den Eingang von `SoEyeTracking` gegangen werden, um korrekte Ergebnisse zu erhalten. Die Szene wird über eine Verbindung automatisch am Eingang des `EventManager`s bereitgestellt. Damit die Übersicht über alle Szenarien erhalten bleibt wurde das Modul `SoEyeTrackingScenario` entwickelt, welches alle für die Evaluierungsversuche verwendeten Szenen enthält. Diese werden über die GUI von `SoEyeTrackingScenario` aus einer Liste gewählt.

5.3 Implementierungsdetails zur Objektbestimmung

Neben der Beschreibung des Aufbaus der Protokolldatei erfolgen Betrachtungen zur Berechnung der Metriken, welche eng mit der Traversierung der Protokolldatei verbunden ist. Details bezüglich der Implementierung der zwei Metriken „Objekt im Unsicherheitsbereich“ sowie „Abstand von Objekt und Blickpunkt“ werden dabei diskutiert. Des Weiteren wird erklärt, wie die in MATLAB importierten Daten strukturiert sind und wie die einzelnen Schritte zur dortigen Datenauswertung umgesetzt wurden. Abschließend wird die Laufzeit des Verfahrens diskutiert.

5.3.1 Aufbau der Protokolldatei

Vor einem Experiment wird im `EventManager` ein Dateiname für jede anzulegende Datei generiert. Dieser besteht aus dem aktuellen Datum und hat die Form „20081131“ für ein am 31.11.2008 durchgeführtes Experiment. Dementsprechend müssen für verschiedene Versuche auch unterschiedliche Verzeichnisse gewählt werden. Die während des Experimentes erzeugten Dateien haben die Erweiterungen EDF, ASC, MRG und M. Die binäre Protokolldatei (EDF) wird nach dem Experiment zum SPC übertragen. Dort erfolgt die Dekodierung in eine neue Textdatei (ASC) mittels eines Programms aus dem Softwarepaket von EYELINK I. Die ASC-Datei besteht aus Samples, Events, Konfigurationseinträgen sowie zusätzlichen Nachrichten zur Kennzeichnung wichtiger Zeitpunkte, wie dem Präsentationsstart.

Jedes Sample besteht aus einem Zeitstempel, den registrierten Bildschirmkoordinaten und der Pupillengröße. Zeiten werden im Datenprotokoll generell in Millisekunden angegeben. Auf den elementaren Samples führt das ETS den Algorithmus zur Sakkadenerkennung aus und klassifiziert damit sie damit zu Events, wie Sakkaden, Fixierungen oder Lidschlägen. Diese werden ebenfalls in der Protokolldatei vermerkt. Ein Sample hat damit folgende Struktur:

```
ZEITSTEMPEL POSITIONX POSITIONY PUPILLENGRÖßE
```

Sakkaden sind durch die Schlüsselwörter **SSACC** und **ESACC** gekennzeichnet, welche am Anfang bzw. Ende der zu einer Sakkade klassifizierten Gruppe von Samples stehen. Während das **SSACC**-Event wenige Parameter besitzt, charakterisiert das **ESACC**-Event die Sakkade sehr genau. Dies umfasst u.a. einen Buchstabenwert (R/L) zur Identifizierung des aufgenommenen Auges, Start- und Endzeit der Sakkade, ihre Dauer sowie ihre Start- und Zielkoordinaten auf dem Bildschirm. Eine Sakkade hat im Protokoll damit den allgemeinen Aufbau:

```
SSACC AUGE STARTZEIT
```

```
:
```

```
:
```

```
ESACC AUGE STARTZEIT ENDZEIT DAUER STARTX STARTY ENDEX ENDEY WINKEL BESCHLEUNIGUNG
```

Fixierungsevents ähneln Sakkadenevents bis auf zwei Ausnahmen. Die begrenzenden Schlüsselwörter heißen **SFIX** sowie **EFIX** und die Charakterisierung des Events erfolgt durch die Mittelwerte der Koordinaten und Pupillengrößen aller eingeschlossenen Samples. Somit ergibt sich folgender Aufbau eines Fixierungsevents:

```
SFIX AUGE STARTZEIT
```

```
:
```

```
:
```

```
EFIX AUGE STARTZEIT ENDZEIT DAUER MITTELX MITTELY MITTELPUPILLENGRÖßE
```

Das letzte Event bilden Lidschläge, welche von **EYELINK I** erkannt werden. Entsprechende Samples werden hier mit den Schlüsselwörtern **SBLINK** und **EBLINK** eingeschlossen. In diesem Zeitraum konnte das System keine Daten erfassen. Zusätzlich wird dieser zeitliche Bereich sowie einige vorhergehende und nachfolgende Samples in **SSACC**- und **ESACC**-Events gehüllt, um das Rauschen vom Beginn des Blinzeln bis zum wieder komplett geöffneten Auge zu kennzeichnen. Es resultiert damit folgende Struktur in der Protokolldatei:

```
SSACC AUGE STARTZEIT
```

```
:
```

```
:
```

```
SBLINK AUGE STARTZEIT
```

```
:
```

```
:
```

```
EBLINK AUGE STARTZEIT ENDZEIT DAUER
```

```
:
```

```
:
```

```
ESACC AUGE STARTZEIT ENDZEIT DAUER STARTX STARTY ENDEX ENDEY WINKEL BESCHLEUNIGUNG
```

Nach der Erzeugung der ASC-Datei erfolgt im **EventManager** deren Vereinigung mit den Kameraevents aus **MEVISLAB** zu einer neuen Datei mit der Endung **MRG**. In dieser sind alle Informationen im Ascii-Format abgelegt. Die Aufnahme von Mausinteraktionen in das Datenprotokoll wurde deaktiviert da den Versuchspersonen ausschließlich Animationen gezeigt werden. Eine Reaktivierung ist jedoch einfach, da alle nötigen Module und entsprechender C++-Code bereits vorhanden ist.

Da Kameraevents auf dem SPC und die EDF-datei auf dem OPC erzeugt werden, müssen die Zeitstempel beider Datensätze synchronisiert werden. Dafür wird die Protokollnachricht GLOBALSTRT genutzt, welche je einen Zeitstempel von SPC und OPC enthält. Da die Zeit des SPC's, im Gegensatz zur Zeit des OPC's, von null beginnt, wurden mit Hilfe der Informationen aus GLOBALSTRT und der Gleichung 5.1 alle Zeiten der ASC-Datei an die Zeiten des SPC's angeglichen.

$$T_{ASC} = T_{ASC} - GLOBALSTRT_{TOPC} + GLOBALSTRT_{TSPC} \quad (5.1)$$

Kameraevents haben das Schlüsselwort PV und treten wie Samples zu einem Zeitpunkt auf, der durch den Zeitstempel gekennzeichnet wird. In der formalen Beschreibung folgen diesem 16 Werte einer Modelviewmatrix und weitere 16 Werte einer Projektionsmatrix. Anschließend wird die Größe des Viewports (Bildschirmauflösung) gespeichert, um später eine korrekte Projektion der 2D-Blickpunkte auf 3D-Blickstrahlen zu erhalten. Der wichtigste Teil besteht jedoch aus den sich anschließenden 12 Parametern einer orthographischen Kamera, welche in WERNECKE [1993] detailliert beschrieben sind. Der Aufbau eines Kameraevents ergibt sich damit folgendermaßen:

PV ZEITSTEMPEL MODELVIEWMATRIX PROJEKTIONSMATRIX VIEWPORT KAMERAPARAMETER

5.3.2 Berechnung der Metriken

Dieser Abschnitt beschreibt zunächst Aufbau und Funktionsweise des auf C++-Ebene im Modul `EventManager` integrierten und in Abbildung 5.5 dargestellten Netzwerkes. Im Anschluss folgen Betrachtungen zur Traversierung der MRG-Datei, welche alle notwendigen Informationen zur Berechnung der Metriken enthält. Traversierung und Metrikenberechnung sind eng miteinander verbunden, weshalb beide Aspekte simultan erläutert werden.

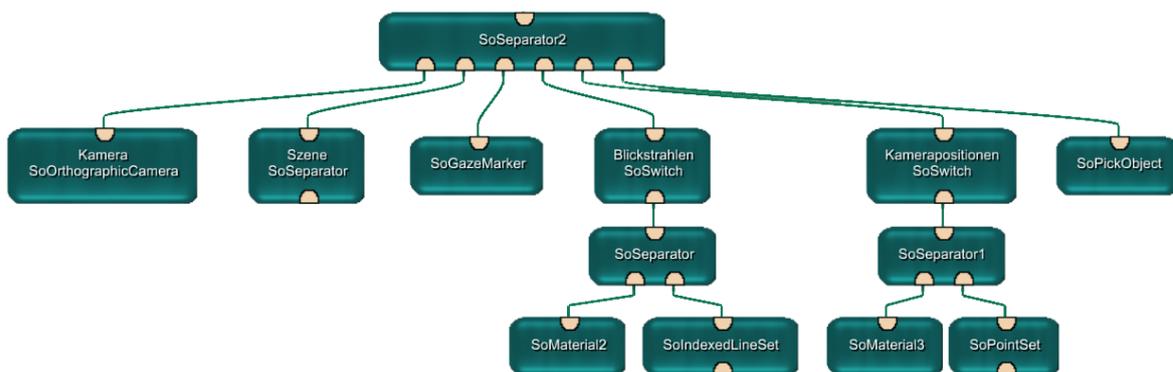


Abbildung 5.5: Darstellung des im `EventManager` integrierten Netzwerkes. Es enthält folgende Module: Eine orthographische Kamera, einen Knoten zur Gruppierung der Szene, drei Knoten bzw. Teilgraphen zur Blickpunktvisualisierung (`SoGazeMarker`, `Blickstrahlen`, `Kamerapositionen`) und dem Modul `SoPickObject`, in welchem die Traversierung der MRG-Datei sowie die Ermittlung der Metriken implementiert ist.

Die drei in Abbildung 5.5 dargestellten Visualisierungskomponenten (vgl. Abbildung 5.4) werden beim Rendern der Szene nur dann ausgewertet, wenn der Benutzer über die GUI des `EventManagers` die entsprechenden Befehle erteilt. `SoPickObject` nutzt die Inventormethode `rayPick()` zur Bestimmung der sich an einer Bildschirmkoordinate befindlichen Objekte.

Ursprünglich war diese Methode zur Auswahl von Objekten mit der Maus gedacht. Vor der nun folgenden Traversierung der Protokolldatei sind folgende Vorverarbeitungsschritte notwendig:

Objekte im Graphen identifizieren - Es erfolgt die Traversierung der Szene mittels einer `SoSearchAction`, um die Pfade zu allem Objektknoten zu ermitteln. Von jedem Pfad kann nun der jeweilige Objektknoten bezogen und dessen eindeutige ID, mittels der Inventormethode `getNodeId()`, ermittelt werden. Sie dient der Zuordnung korrekter Transparenzwerte und der Berechnung der Abstandsmetrik. Pfade zu den in medizinischen Datensätzen integrierten `SoLabel`-Knoten werden ebenfalls mit dieser Action bestimmt. Diese Bezeichnungen ermöglichen die Beschriftung der Diagramme in MATLAB.

Berechnung der Bounding-Boxes - Dieser Schritt wird mit Hilfe der `SoBoundingBoxAction` ausgeführt. Angewandt wird diese auf jeden der ermittelten Objektpfade. Dabei traversiert die Action den Szenengraphen bis zum Objekt und findet eventuelle Transformationen. Diese dienen der Ausrichtung der Bounding-Box am lokalen Koordinatensystem des Objektes, was zu einer besseren Approximation führt als die Ausrichtung an den Achsen des Weltkoordinatensystems. Diese speziellen Bounding-Boxes werden in OPEN INVENTOR durch die Klasse `SbXfBox3f` unterstützt.

Bestimmung der Transparenzwerte - In diesem Fall wird eine `SoCallbackAction` auf Objektpfade angewandt. Auf dem Weg zum Objekt traversiert sie entsprechende Materialknoten und speichert die darin enthaltenen Transparenzwerte. Das Auslesen dieser Werte erfolgt in einer speziellen Callbackfunktion, welche durch die Action aufgerufen wird, sobald das Objekt gefunden wurde.

Abschließend müssen zwei Schritte ausgeführt werden. Der erste besteht aus dem Suchen des ersten Kameraevents und dem Kopieren der dort gespeicherten Parameter in die Parameterfelder der in Abbildung 5.5 dargestellten Kamera. Diese ist nun genau, wie zu Beginn des Experimentes, ausgerichtet. Würde dieser Schritt entfallen, so könnten Blickpunkte nicht auf die korrekten Strahlen in der Szene projiziert werden. Der zweite Schritt besteht in der Initialisierung des MATLAB-Scriptes (M-Datei), welches am Ende alle nötigen Daten zur Auswertung beinhalten wird. Zunächst erfolgt jedoch nur das Schreiben von Headerinformationen und der Labels aller Objekte in Form eines Vektors von Zeichenketten mit dem Format „`ETLABELS = ['Label01'; 'Label02'; 'Label03']`“.

Traversierung der MRG-Datei

Die Traversierung der Protokolldatei erfolgt im Modul `SoPickObject` mittels einer `while`-Schleife, welche die Datei zeilenweise ausliest und nach den Schlüsselwörtern für Fixierungs- und Kameraevents sowie nach Samples sucht. Aufgrund der ausschließlichen Betrachtung von Fixierungs- und Kameraevents erfolgt mit diesem Schritt gleichzeitig die Datenfilterung. Die Verarbeitung eines Kameraevents umfasst zunächst die Positionierung der in Abbildung 5.5 gezeigten Kamera auf Basis der im Event gespeicherten Kameraparameter. Anschließend müssen die Zentren der Bounding-Boxes in Bildschirmkoordinaten neu bestimmt werden. Die Projektion der Hüllkörpermittelpunkte erfolgt mittels der Inventorfunktion `projectToScreen()`. Die dabei entstehenden 2D-Koordinaten werden in einer internen Datenstruktur gespeichert und bilden die Grundlage für die Metrik „Abstand des Blickpunktes vom Objekt“ (Spalte zwei in Tabelle 4.4). Das Setzen der Kameraparameter und Projektion müssen in

dieser Reihenfolge ausgeführt werden, da zur Projektion sonst die alten Kameraparameter verwendet werden.

```

1 for y = -rad to rad do
2   for x = -rad to rad do
3     if (x,y) innerhalb eines Kreises mit Radius rad then
4       PN(x) = (P(x) + x) / VP(x);
5       PN(y) = (P(y) + y) / VP(y);
6       if (PN(x), PN(y) ) innerhalb Viewport do
7         SoRayPickAction rpa( VP );
8         rpa.setNormalizedPoint( PN );
9         rpa.apply( root );

```

Algorithmus 5.1: Vereinfachte Darstellung der Methode `_computeSecondaryRays()` als Pseudocode. Zu dem durch ein Sample gegebenen Blickpunkt P werden nacheinander alle Koordinaten eines Kreises mit dem Pixelradius rad gerechnet. Die resultierenden Koordinaten dieses Unsicherheitskreises werden dann durch die Ausmaße des Viewports VP dividiert und damit auf das Intervall $[0 \dots 1]$ normiert. Wenn die resultierenden, normierten Koordinaten PN auf dem Bildschirm liegen, so erfolgt die Initialisierung einer `SoRayPickAction` mit dem jeweiligen Viewport und PN . Diese Action wird dann auf einen Gruppenknoten ($root$) angewandt, wodurch sie den darunter liegenden Teilgraphen traversiert.

Neben Kameraevents werden, in der MRG-Datei, Fixierungsevents (SFIX, EFIX) gesucht. Bei dem gefundenen Schlüsselwort SFIX wird intern eine Marke gesetzt, damit für die ab hier folgenden Fixierungssamples spezielle Verarbeitungsschritte ausgeführt werden. Für jedes dieser Samples werden Zeitstempel, X-, Y-Koordinate und Pupillengröße in einer internen Datenstruktur gespeichert. Anschließend erfolgt der Aufruf der Methode `_computeSecondaryRays()` (Algorithmus 5.1). Dort werden zunächst alle Bildschirmkoordinaten des Unsicherheitsbereiches für das aktuelle Sample berechnet und anschließend eine `SoRayPickAction` mit allen berechneten Koordinate initialisiert sowie auf den obersten, in Abbildung 5.5 dargestellten `SoSeparator`-Knoten angewandt.

Nach dem Anwenden der Action werden Kamera, Szene und `SoPickObject` in dieser Reihenfolge traversiert. Mittels der Kameraparameter kann die Action den übergebenen, normierten auf einen Strahl im Raum projizieren. Besucht die Action bei der Traversierung einen Geometrieknoten, erfolgt die Schnittpunktberechnung zugehöriger Polygone mit dem Blickstrahl. In der Methode `rayPick()` des Moduls `SoPickObject` werden schließlich die berechneten Schnittpunktkoordinaten für die Auswertung zur Verfügung gestellt.

In der in Algorithmus 5.2 dargestellten Methode `rayPick()`, erfolgt die Analyse der sortierten Schnittpunktliste L . Die in L enthaltenen Instanzen der Klasse `SoPickedPoint` verwalten Zeiger auf die zu den Schnittpunkten gehörenden Objekte. Zu diesen werden mittels der ID's entsprechende Transparenzwerte T aus einer internen Datenstruktur ermittelt, welche in einem Vorverarbeitungsschritt erzeugt wurde. Die Variable *attenuation* beschreibt im Algorithmus die Abschwächung der Sichtbarkeit entlang des Strahls. Wird entlang des Strahls eine Sichtbarkeit von 0.05 erreicht, werden nachfolgende Schnittpunkte nicht weiter analysiert, da zugehörige Objektregionen nicht wahrnehmbar sind.

Das Array *HIT* hat eine Komponente für jedes Objekt. Wird ein Objekt von einem Strahl geschnitten, so wird die entsprechende Komponente inkrementiert, wobei für jeden Primär- und Sekundärstrahl nur für den ersten gefundenen Schnittpunkt eine Inkrementierung erfolgt. Liegt der Unsicherheitsbereich also komplett auf einem Objekt, so ist der Wert der entsprechenden

```

1  if L.length() > 0 then
2    attenuation = 1;
3    for i=0 to L.length() do
4      obj = L[i].getObject();
5      if attenuation > 0.05 then
6        for idx = 0 to Anzahl Objekte do
7          if obj.getID() == ID[idx] then
8            if T[idx] < 1 then
9              HIT[idx]++;
10             attenuation *= T[idx];

```

Algorithmus 5.2: Vereinfachte Darstellung der Implementation der Methode `rayPick()` als Pseudocode. Objekte werden hier den vom ETS stammenden Blickpunkten zugeordnet. *L* stellt dabei die von der `SoRayPickAction` erzeugte Liste aller Objekt-Strahl-Schnittpunkte dar, welche nach Abstand von der Kamera sortiert ist.

Komponente gleich der Anzahl der Koordinaten des Unsicherheitsbereiches. Das Array *HIT* wird für jedes Sample in einer internen Datenstruktur zwischengespeichert und bildet die Grundlage für die Metrik „Objekt im Unsicherheitsbereich“ (Spalte eins in Tabelle 4.4).

1	2	3	4	5	6	7	8	9	10	11
1	116	543	674	1974	0	0	300	979	511	511
1	120	543	674	1974	0	0	300	979	511	511
2	452	443	593	1977	700	0	300	979	511	511
2	456	446	593	1977	582	0	300	979	511	511
3	580	946	537	1613	0	6361	300	979	511	511
3	584	949	542	1585	0	6361	300	979	511	511

Tabelle 5.3: Vereinfachte Darstellung der Datenmatrix, welche nach MATLAB exportiert wird. Jede der drei Fixierungen enthält in der Realität weitaus mehr als nur zwei Samples (Zeilen), die hier aber nicht dargestellt werden können. Die Spalten repräsentieren folgende Informationen: 1) Fixierungsnummer, 2) Zeitstempel, 3-4) X-/Y-Blickposition, 5) Pupillengröße, 6-7) Treffer durch Primär-/Sekundärstrahlen (hier für zwei Objekte), 8-11) X-/Y-Position der Bounding-Box-Zentren (hier für die Bounding-Boxes zweier Objekte).

Die beschriebenen Schritte werden für jedes Sample einer Fixierung, bis zum gefundenen EFIX-Event durchgeführt. Der Gesamtprozess wird für alle Fixierungen des Datenprotokolls, wiederholt. Am Ende der Traversierung, werden die in den internen Datenstrukturen zwischengespeicherten Informationen in tabellarischer Form zum MATLAB-Script exportiert. Eine vereinfachte Darstellung dieser Datenmatrix zeigt Tabelle 5.3. Diese enthält drei Fixierungen mit je zwei Samples sowie die Daten, deren Erzeugung in diesem Abschnitt beschrieben wurde. Diese Matrix wird als letztes in die bereits bestehende M-Datei geschrieben.

5.3.3 Datenauswertung in MATLAB

Nachdem die M-Datei in MATLAB geöffnet wurde, erfolgt die Berechnung der endgültigen Wahrscheinlichkeiten aus den importierten Daten sowie die graphische Darstellung der Ergebnisse mittels eines zweiten Skriptes. Bis alle, im Konzept beschriebenen Metriken umgesetzt wurden, erscheint es sinnvoll die Analyse in MATLAB anstatt in C++ durchzuführen. Nachdem sichergestellt wurde, dass das System mittels aller in Abschnitt 5.3.3 zusammengefassten

Metriken sowie einer globalen Analyse der Wahrscheinlichkeiten korrekte Ergebnisse liefert, ist es nur noch ein kleiner Schritt zur endgültigen Implementation der Metriken in C++.

Zunächst werden die Datenmatrix und der Labelvektor eingelesen. In der Teilmatrix der Metrik „Objekt im Unsicherheitsbereich“ (Spalte 6-7 in Tabelle 5.3) werden darauf folgend alle Werte größer null auf 1 gesetzt, was bedeutet, dass sich das entsprechende Objekt zum jeweiligen Zeitpunkt im Unsicherheitsradius befunden hat. Anschließend erfolgt die Konvertierung in eine Wahrscheinlichkeit (vgl. Abschnitt 4.2.2). Die Abstandsmetrik wird anschließend über den Blickpunkt (Spalten 3-4) und die Zentren der Bounding-Boxes (Spalten 8-11) berechnet und ebenfalls in eine Wahrscheinlichkeit konvertiert.

Das Resultat beider Metriken sind zwei Matrizen, deren Zeilenanzahl der Anzahl der Fixierungssamples aus der M-Datei entspricht. Die Spaltenzahl ist jeweils gleich der Anzahl der Szenenobjekte. Die Kombination beider Metriken und damit das Resultat aus Phase eins des 3-Phasen-Modells erfolgt durch komponentenweise Multiplikation dieser Matrizen mit dem Gewicht und deren anschließender Addition. Dies ist in MATLAB aufgrund integrierter Operatoren in wenigen Schritten möglich.

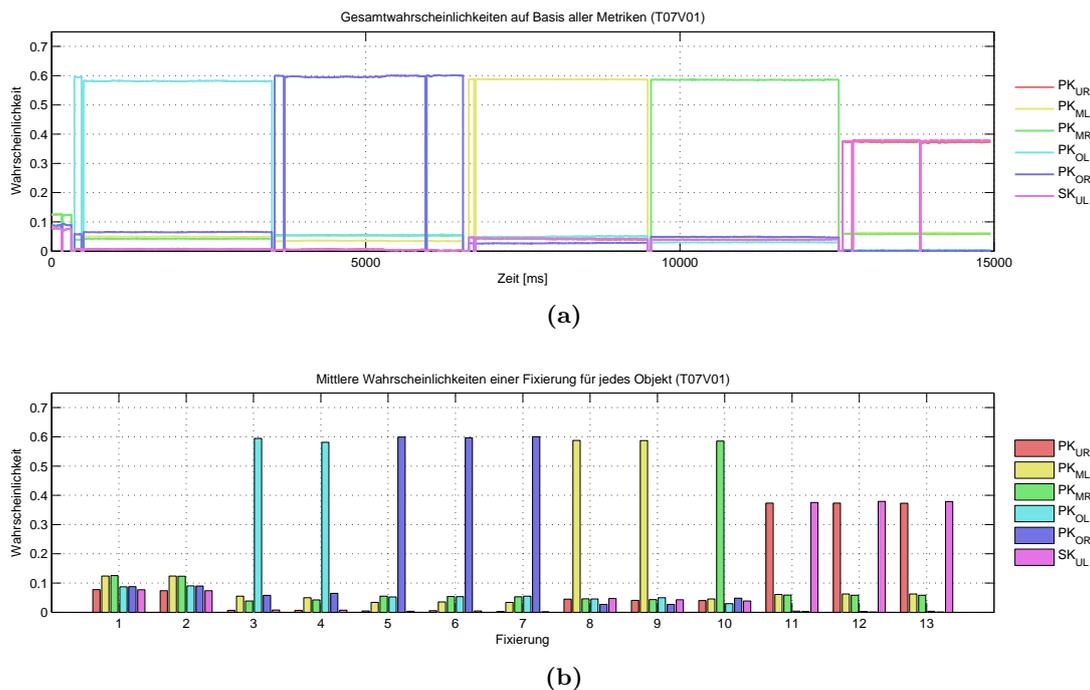


Abbildung 5.6: Veranschaulichung der Resultate zur Objektidentifizierung auf Basis zweier Metriken am Beispiel eines 15-Sekunden-Experimentes mit dem Szenario aus Abschnitt 6.1.2. Die korrekten Kugeln wurden fixiert, was der Algorithmus in diesem Fall auch bestätigt. Abbildung (a) zeigt das Ergebnis aus Phase eins und in Abbildung (b) aus Phase zwei des 3-Phasen-Modells dargestellt.

Das Resultat kann als Graph über jedes Sample einer Fixierung, wie in Abbildung 5.6(a) dargestellt werden. Das Endergebnis aus Phase zwei, also die aggregierten Wahrscheinlichkeiten für jede Fixierung, ist durch ein Balkendiagramm wie in Abbildung 5.6(b) anschaulich zu präsentieren. In diesem Beispiel arbeitet der Algorithmus gut und erkennt die Fixierungen der Kugeln korrekt.

5.3.4 Diskussionen zur Laufzeit

Die Laufzeit des Verfahrens hängt vor allem von der Schnittpunktberechnung der `SoRay-PickAction` und damit von der Komplexität der Szene ab. Um zunächst die Genauigkeit des Algorithmus nicht einzuschränken, wird für jedes Sample eine Kreisfläche von einem Grad Radius berechnet. Dies entspricht dem Genauigkeitsfehler von EYELINK I und bei einer Distanz des Auges zum Display von 75 cm und einer Monitorauflösung von 86 dpi einer Größe von 45 Pixeln. Für jedes Pixel innerhalb dieses Kreises erfolgt anschließend die Berechnung der Sekundärstrahlen. Dementsprechend werden für jedes der 3219 Samples aus allen 18 Fixierungen eines Experimentes, der Länge von 15 Sekunden, $\pi * 45^2 \approx 6361$ Strahlen berechnet. Das macht eine Gesamtanzahl von $3219 * 6361 = 20.476.059$ Strahlen, welche auf Schnitte mit der Geometrie getestet werden.

In Abschnitt 3.3.1 wurde der Aufwand bei der Schnittpunktberechnung verdeutlicht. Es wird davon ausgegangen, dass dieser Schritt einen Aufwand von $O(N)$ hat, wenn N die Anzahl aller Polygone ist. Einen weiteren Faktor stellt die Anzahl S der analysierten Fixierungssamples dar. Mit der Anzahl B , der Bildschirmkoordinaten innerhalb des Unsicherheitskreises, ergibt sich damit ein Gesamtaufwand von $O(S) * O(B) * O(N)$.

Aufgrund seiner Größe muss versucht werden, diesen Aufwand weiter zu verringern. Dazu können folgende Schritte unternommen werden:

- Verringerung analysierter Blickpunkte im Unsicherheitsbereich eines Samples durch spezielle Samplingstrategien
- Ausnutzung räumlicher Zusammenhänge, um nahe beieinander liegende Samples nicht analysieren zu müssen
- Implementation einer Hierarchie von Hüllkörpern, welche die Szenengeometrie einschließen und Schnittpunktberechnungen mit jedem Polygon verhindern

5.4 Zusammenfassung

Mit der Implementierung zweier Metriken wurden die Grundlagen zur Umsetzung der restlichen, konzipierten Maße vorgegeben. Am Beispiel der graphischen Auswertung der generierten Informationen konnte gezeigt werden, dass fixierte Objekte in einfachen Szenen korrekt bestimmt werden können. Teile des Konzeptes, wie die globale Analyse der resultierenden Wahrscheinlichkeiten, wurden nicht umgesetzt, da das Gesamtverfahren, bestehen aus zwei Metriken, noch keine optimalen Werte liefert.

Zusammenfassend kann gesagt werden, dass Ansätze des entwickelten Verfahrens erfolgreich umgesetzt wurden. Zusätzlich konnte gezeigt werden, dass dieses Vorgehen eine mögliche Lösung des in dieser Arbeit formulierten Problemes darstellt, da das Verfahren tatsächlich die fixierten Objekte durch höhere Wahrscheinlichkeiten kennzeichnet. Jedoch ist es notwendig, die Umsetzung um die im Konzept entwickelten Gesichtspunkte zu erweitern, um schließlich das gesamte Verfahren evaluieren zu können. Dazu gehört die Umsetzung vor allem der bildbasierten Metriken sowie die weitere Untersuchung von Möglichkeiten zur globalen Analyse der resultierenden Wahrscheinlichkeiten aus Phase zwei. Eine genaue Betrachtung der Möglichkeiten und Grenzen des Verfahrens erfolgt im nächsten Kapitel.

Kapitel 6

Evaluierung und Ergebnisse

Ziel der hier vorgestellten Evaluierung ist es zu zeigen, dass das Verfahren für alle Versuchspersonen ähnliche Ergebnisse. Zudem sollen Ausnahmefälle bzw. Einschränkungen hervorgehoben werden. Dazu erfolgt zunächst die Vorstellung der Experimente, die zur Erzeugung der für die Evaluierung nötigen Daten durchgeführt wurden. Im Anschluss findet die Charakterisierung der Population teilnehmender Probanden sowie eine Beschreibung des Versuchsablaufs und der verwendeten Computerhardware zur Berechnung der Metriken statt. Für die Evaluierung wurden mittels Fragebögen, Experimenten und deskriptiver Statistik quantitative Verfahren verwendet, welche die numerische Darstellung empirischer Sachverhalte erlauben [SNEDECOR und COCHRAN, 1989]. Nachdem die Evaluierungsergebnisse vorgestellt wurden folgt die Auswertung der ausgefüllten Fragebögen.

6.1 Evaluierungsversuche und Hypothesen

In diesem Abschnitt wird jedes Experiment genau charakterisiert. Das umfasst neben einer strukturellen Beschreibung der verwendeten Szenarien und der Erläuterung getroffener Annahmen auch die die Postulierung von Hypothesen, deren Gültigkeit in den Experimenten nachgewiesen werden soll. In diesen Experimenten ist das Stellen spezieller Suchaufgaben, wie sie in den Abschnitten 3.1.5 und 4.1.2 beschrieben wurden, noch nicht vorgesehen. Die Probanden erhielten zunächst genaue Anweisungen, wohin sie zu bestimmten Zeitpunkten schauen sollen. Somit können die erwarteten Blickmuster mit den vom Eye-Tracker registrierten Daten verglichen werden. Aus diesen Ergebnissen sind erste Vermutungen über das Verhalten des Verfahrens auf realen Szenen ableitbar.

6.1.1 Fixierungen in verdeckungsfreien Szenen

Dieser initiale Versuch demonstriert exemplarisch, dass in bestimmten Fällen die Klassifizierung von Fixierungen zu Objekten eindeutig, anhand einer einzigen Metrik, möglich ist. Die Objektidentifizierung erfolgt hier zunächst ausschließlich unter Verwendung des Blickpunktes mit dem dazugehörigen Unsicherheitsbereich. Es wird außerdem demonstriert, welche Probleme bei dicht beieinanderliegenden Objekten auftreten.

Hypothese zu Versuch 1: Die vom Eye-Tracker detektierten Fixierungen können eindeutig zu Objekten zugeordnet werden, wenn deren Abstand auf dem Bildschirm größer als die doppelte Genauigkeit des Eye -Trackers ist.

Es wird von der doppelten Genauigkeit ausgegangen, weil dieser Wert den Durchmesser statt nur den Radius des Unsicherheitskreises beschreibt (vgl. Abschnitt 2.2). Ist die Entfernung zweier Objekte O_1 und O_2 zu jedem Zeitpunkt größer als der Durchmesser des Unsicherheitskreises, kann theoretisch ausgeschlossen werden, dass die Fixierung von O_1 zusätzlich zu O_2 klassifiziert wird. Theoretisch deswegen, da es weiterhin die beschriebenen Driftabweichungen gibt, welche bei jedem Probanden unterschiedlich groß ausfallen können, da sie von der jeweiligen Körper- bzw. Kopfbewegung abhängen.

Aufgabe und Aufbau der Testszene

Das Testszenario, dargestellt in Abbildung 6.1(a), enthält insgesamt sechs graue, sich nicht bewegende Primärkugeln. Im Vollbildmodus auf einem Bildschirm mit einer Auflösung von 1280 mal 1024 Pixel haben diese einen gegenseitigen Abstand von ca. 112 Pixel. Die Materialeigenschaften wurden möglichst einfach gewählt, um Ablenkung aufgrund bevorzugter Farben oder Texturen zu vermeiden. Die roten Sekundärkugeln werden nacheinander, in der in Abbildung 6.1(b) gezeigten Reihenfolge, für jeweils drei Sekunden eingeblendet. Durch diesen Zeitrahmen entstehen ausreichend lange Fixierungen, die sich gut zur Datenanalyse eignen.

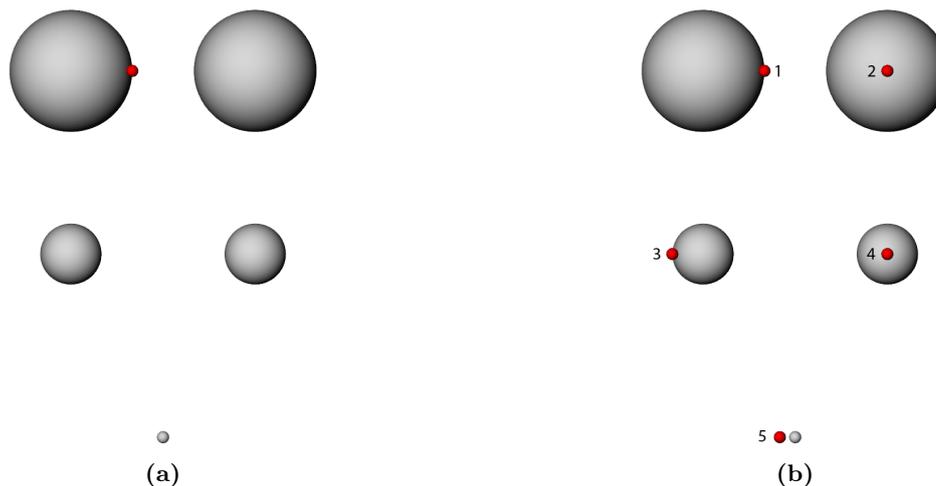


Abbildung 6.1: Darstellung des dargebotenen Stimulus für den ersten Evaluierungsversuch. Das in (a) gezeigte Bild sieht die Versuchsperson zu Beginn des Experimentes. Die roten Kugeln werden jeweils im Dreisekudentakt an den in (b) markierten Positionen eingeblendet, um den Betrachter an diese Stellen schauen zu lassen.

Der Benutzer hat während des Experimentes die Aufgabe, seinen Blick immer auf die roten Sekundärkugeln zu richten. Sie dienen dem Zweck der Fixierung von Kanten bzw. Zentren der Primärkugeln. Diese Positionen wurden gewählt, da sich der Mensch bei der Objektbetrachtung einerseits an Kanten orientiert und andererseits eine zentrierte Objektfokussierung simuliert werden soll (vgl. Abschnitt 3.1.3). Mit der Fixierung der oberen, linken Sekundärkugel in (vgl. Abbildung 6.1(b)) soll untersucht werden, ob die obere, rechte Primärkugel vom Algorithmus gleichzeitig als fokussiert erkannt wird und damit die Ungenauigkeit des Eye-Trackers größer ist als angenommen.

Die Größe der Primärkugeln nimmt von oben nach unten ab. Damit ist es möglich zu untersuchen, wie sich objektbasierte Metriken bei der Fixierung verschieden großer Objekte

verhalten. Die untersten beiden Kugeln liegen relativ eng beieinander. Diese Tatsache soll zeigen, ob die Identifikation der fokussierten roten Sekundärkugel vom Algorithmus tatsächlich erkannt wird oder ob häufig Fixierungen der rechten unteren Kugel ermittelt werden. Im letzteren Fall, müssen Wege gefunden werden, die Erkennung von Fixierungen zu verbessern, da die Erkennung dicht beieinander liegender Objekte, für den Einsatz des Verfahrens zur Evaluierung medizinischer Visualisierungen, von Bedeutung ist.

Bezeichnung	Kugeltyp und -position
PK_{OL}	Primärkugel oben, links
PK_{OR}	Primärkugel oben, rechts
PK_{ML}	Primärkugel mitte, links
PK_{MR}	Primärkugel mitte, rechts
SK_{UL}	Sekundärkugel unten, links
PK_{UR}	Primärkugel unten, rechts

Tabelle 6.1: Bezeichnungen der Szenenobjekte für den ersten Evaluierungsversuch.

Die Primärkugeln wurden mit Bezeichnungen entsprechend ihrer Position versehen. Tabelle 6.1 gibt die Namen der einzelnen Objekte wieder. Die Sekundärkugeln haben keine Namen, da sie vom Algorithmus nicht erkannt werden können. Eine Ausnahme bildet SK_{UL} , da ihre Fixierungen zur Unterscheidung von Fixierungen auf PK_{UR} benötigt werden. Die Bezeichnungen aus Tabelle 6.1 werden ebenfalls zur Kennzeichnung der Objekte in den Diagrammen zur Ergebnisauswertung verwendet.

Annahmen zu Versuch 1: Im Idealfall betrachtet die Versuchsperson solange eine Sekundärkugel, wie diese auf einer Stelle eingeblendet wird. Ihr Blick sollte sich verschieben, sobald diese Kugel repositioniert wird. Außerdem wird angenommen, dass die anderen Fehlerkomponenten, neben dem Genauigkeitsfehler, möglichst klein sind. Es kann versucht werden, dies durch eine hohe Konzentration der Versuchsperson während des Experimentes sicherzustellen. Daraus folgend sollte sich bei der Fixierung einer Primärkugel also zu keiner Zeit eine andere Primärkugel innerhalb des Ungenauigkeitsbereiches befinden.

6.1.2 Semi-Transparente Verdeckung

Visualisierungen medizinischer Datensätze enthalten häufig kleine Strukturen, wie Lymphknoten oder Mandeln. Diese sind zudem kompakt und liegen je nach Betrachterstandpunkt hinter Gefäßen, Muskeln oder anderen Strukturen, welche semi-transparent dargestellt sind. Somit stellt sich die Frage: „Wie können Fixierungen semi-transparent verdeckter, kleiner, kompakter Strukturen erkannt werden?“. Um diese Frage zu beantworten wurde ein Versuch konstruiert, mit dem auf Basis einer Animation untersucht werden soll, ob bei temporärer, semi-transparenter Verdeckung die Zuordnung von Fixierungen zu kleinen, kompakten Objekten möglich ist. Es wird dabei auf Schwierigkeiten eingegangen, die bei der Identifikation von Fixierungen großer, semi-transparenter Vordergrundobjekte auftreten können.

Hypothesen zu Versuch 2: Aus Objekt- und Blickpunktpositionen lassen sich bei Animation der Szene die jeweils fixierten Objekte auf Basis der Folgebewegungen ableiten. Außerdem ist mit diesem Vorgehen die Auflösung von Doppeldeutigkeiten bei der Fixierung semi-transparent verdeckter, kleiner, kompakter Objekte möglich.

Aufgabe und Aufbau der Testszene

Das Szenario, wie es in Abbildung 6.2 dargestellt ist, ist bewusst einfach gewählt. Es soll zunächst untersucht werden, ob aus der jeweiligen Position des Blickpunktes und der Position des Objektes Rückschlüsse auf die Fixierung eines kleinen kompakten Objektes gezogen werden können. Daher besteht die Szene aus einem grauen Quader, der zur semi-transparenten Verdeckung der roten Kugel dient. Beide Objekte rotieren mit einer Winkelgeschwindigkeit¹ von $36^\circ/s$ um eine vertikale Achse, welche parallel zur Y-Achse des dargestellten Koordinatensystems verläuft. In LU et al. [2006] wird diese Geschwindigkeit als idealer Wert angegeben. Es werden langsame Folgebewegungen provoziert und gleichzeitig verhindert, dass sich die Versuchsperson langweilt und ihren Blick damit vom zu fixierenden Objekt abwendet. Der Betrachter hat in diesem Versuch die Aufgabe, die rote Kugel zu fixieren.

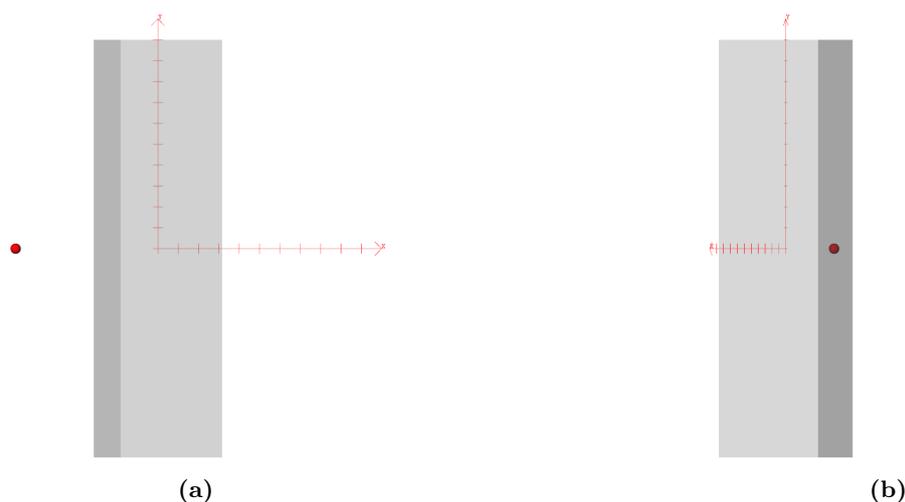


Abbildung 6.2: Darstellung des Szenarios für den zweiten Evaluierungsversuch. Die virtuelle Kamera dreht sich um die dargestellte Szene, was zur zeitweisen Verdeckung der roten Primärkugel durch den semi-transparenten Quader führt (b). Das Koordinatensystem ist während des Experimentes nicht zu sehen.

Annahmen zu Versuch 2: Auch bei Verdeckung fixiert der Betrachter das Zielobjekt und verschiebt seine Aufmerksamkeit nicht zum Quader. Dies würde die Versuchsergebnisse beeinträchtigen, da die Kugel dann eine niedrige Fixierungswahrscheinlichkeit bekommt. Es kann aber vorkommen, dass zu diesem Zeitpunkt beide Objekte im Fokus liegen. Die Fixierung des Zielobjektes sollte also die gesamte Zeit aufrecht erhalten werden, um die Eignung der Abstandsmetrik zu überprüfen.

¹ Winkelgeschwindigkeit ist die zeitliche Änderung des Drehwinkels um eine Rotationsachse. Sie ist unabhängig vom Abstand (Rotationsbahnradius) eines Objektes von dieser Achse.

6.1.3 Fixierungswechsel zwischen bewegten Objekten

Dieser Versuch soll zeigen, ob Fixierungswechsel zwischen Objekten mit unterschiedlicher Entfernung zueinander und durch die relative Bewegung des Blickpunktes zum projizierten Zentrum der Objektboundingbox, erkennbar ist. Solche Fixierungswechsel treten auf, wenn Abstände anatomischer Strukturen geschätzt werden müssen. Es wird des Weiteren Wert darauf gelegt, ob die Verfolgung eng beieinander liegender Objekte, die sich unterschiedlich bewegen, erkennbar ist. Messfehler können dabei so groß sein, dass Fixierungen eng beieinander liegender Objekte selbst mit Hilfe zusätzlicher Metriken nicht eindeutig zugeordnet werden können. Hierzu wird untersucht, ob Fixierungswechsel zwischen zwei dicht positionierten Objekten, auf der Basis des entwickelten Metrikensystems, unterscheidbar sind. Auch in diesem Experiment besteht wieder die Beschränkung auf kleine, kompakte Objekte, da die Bewegung des projizierten Zentrums in diesem Fall die Objektbewegung gut annähert und die verwendeten Kugeln kleinen, kompakten, anatomischen Strukturen ähneln.

Hypothesen zu Versuch 3: Fixierungswechsel zwischen Objekten mit unterschiedlich schnellen Bewegungen können mittels des entwickelten Metrikensystems unterschieden werden. Des Weiteren ist die Differenzierung von Wechselfixierungen zwischen den dicht zueinander positionierten Objekten K_3 und K_4 möglich.

Aufgabe und Aufbau der Testszene

Die Szene besteht aus vier Kugeln gleicher Größe, deren Bezeichnungen in Abbildung 6.3 dargestellt sind. Sie wurden so um die Y-Achse angeordnet, dass bei einer Rotation der Kamera um diese Achse sich jede Kugel mit einer anderen Tangentialgeschwindigkeit bewegt. Das resultiert aus den unterschiedlichen Abständen der Objekte zur Rotationsachse. Die relativen Abstände sind für jedes Objekt wie folgt festgelegt: $D_{K_1} = 10$, $D_{K_2} = 5$, $D_{K_3} = 1$, $D_{K_4} = 0$.

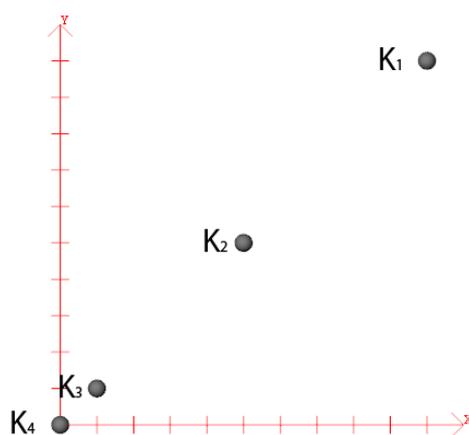


Abbildung 6.3: Darstellung des Szenarios für den dritten Evaluierungsversuch. Die Kugeln rotieren um die vertikale Y-Achse. Für jeweils drei Sekunden werden die Kugeln in einer bestimmten Reihenfolge rot eingefärbt. Die Abbildung zeigt den initialen Zustand der Szene und die in der Auswertung verwendeten Bezeichnungen. Das Koordinatensystem ist während des Experimentes nicht zu sehen.

Aus Abbildung 6.3 ist zu erkennen, dass die Abstände zwischen K_1 und K_2 bzw. K_2 und K_3 relativ groß sind. Das hat den Grund, damit sich kein anderes Objekt im Unsicherheitskreis des jeweils aufgenommenen Blickpunktes befindet, wenn eine Kugel gerade fixiert wird. Der Abstand zwischen K_3 und K_4 wurde dagegen klein gewählt, um zu untersuchen, inwiefern Fixierungen dicht positionierter Objekte bei Animation der Szene unterschieden werden können. Der Abstand ist in diesem Fall größer als bei den untersten beiden Kugeln in Experiment eins.

Des Weiteren halbieren sich die Abstände (Rotationsbahnradien) von K_1 , K_2 und K_4 jeweils zur Rotationsachse. Da zwischen Radius und Umfang der Rotationsbahn eines jeden Objektes ein linearer Zusammenhang besteht ($u = 2\pi r$), halbiert sich damit auch jeweils der zurückzulegende Weg für jedes Objekt, bei Rotation um die Y-Achse. Daraus und aus der konstanten Winkelgeschwindigkeit eines jeden Objektes folgt, dass Objekte mit kleinem Abstand zur Rotationsachse, auch eine kleinere Tangentialgeschwindigkeit haben. Solche Objekte bewegen sich damit langsam über den Bildschirm als diejenigen mit hoher Tangentialgeschwindigkeit. Die Kugel K_4 liegt direkt auf der Rotationsachse und dreht sich damit ausschließlich um sich selbst. Die verschiedenen Geschwindigkeiten sind in Bildschirmkoordinaten z.B. über die Änderung des projizierten Zentrums einer jeden Objektboundingbox messbar.

Der Betrachter hat die Aufgabe, sich jeweils auf die rot gefärbte Kugel zu konzentrieren. Die Objekte werden, wie beim letzten Experiment, im Dreisekudentakt und in folgender Reihenfolge eingefärbt: K_4 , K_3 , K_4 , K_2 , K_1 . Die erneute Fixierung von K_4 im dritten Schritt dient dazu, Fixierungswechsel zwischen nahe beieinander liegenden Objekten herbeizuführen, wie sie bei der Abstandseinschätzung zwischen zwei Objekten auftreten können.

Annahmen zu Versuch 3: Die Versuchspersonen konzentrieren sich tatsächlich die gesamten drei Sekunden auf das jeweils rot gefärbte Objekt. Dies ist für eine erste Evaluierung des Verfahrens notwendig, da sonst nicht untersucht werden kann, inwiefern das Metrikensystem korrekt arbeitet.

6.2 Versuchsdurchführung

Die Versuche hatten jeweils eine Dauer von 15 Sekunden und deren Auswertung wurde unter WINDOWS XP auf einem System mit einer INTEL PENTIUM 4 CPU (2.99 GHz) und einem GB Arbeitsspeicher durchgeführt. Als Grafikkarte ist eine NVIDIA GEFORCE 8800 GTX eingesetzt worden.

Die Versuche wurden für jede Person auf die gleiche Weise und ohne eine spezielle Motivation durchgeführt. Nachdem sich der Proband auf die Art und Weise auf den stabilen Stuhl gesetzt hat, dass eine Auge-Display-Distanz von 75 cm erreicht wird, hat der Versuchsleiter das Eye-Tracking-Gerät auf dem Kopf des Probanden befestigt. Dabei wurde ein Kompromiss zwischen Bequemlichkeit und stabilem Sitz gewählt, da bei zu starker Befestigung schnell Kopfschmerzen auftreten können. Nach dem Starten der Software wurden die jeweils verwendete Kamera optimal eingestellt, so dass sie einerseits das Sichtfeld der Versuchsperson nicht behinderte und sie andererseits die Pupille bei jeder Augenstellung erkennen konnte. Es wurde jeweils die Kamera gewählt, die das von der Versuchsperson als „schärfer sehend“ eingeschätzte Auge aufnimmt.

Da keine der Probanden an einer Eye-Tracking-Studie teilgenommen hatte, wurden vor Beginn der eigentlichen Aufgaben, auflockernde Experimente durchgeführt. Sie verdeutlichten

die Anwendungsmöglichkeiten des Eye-Tracking, nahmen den Versuchspersonen die Aufregung und trainierten das nötige Verhalten bei der Kalibrierung. Nachdem der Proband verstanden hatte, wie die Kalibrierung abläuft, wurden MEVISLAB und die Experimente gestartet. Nachdem vor jedem der drei Versuche die jeweilige Aufgabe bekannt gegeben wurde, konnte die Kalibrierung gestartet werden. Das System speichert zwar das Kalibrierungsergebnis, einige Versuchspersonen haben sich jedoch mitunter stark zwischen den Experimenten bewegt. Daher fand die Kalibrierung vor *jedem* Versuch statt. Wenn die aufgenommenen Blickpunkte zu ungenau waren, wurde das gesamte Experiment (inklusive Kalibrierung) wiederholt. Die entsprechende Überprüfung fand anhand der Visualisierung aus Abbildung 5.4(b) und der vom Eye-Tracking-System berichteten Kalibrierungsgüte statt.

Im Anschluss an die drei Experimente wurden die Probanden gebeten den in Abbildung B.1 abgebildeten Fragebogen auszufüllen. Bei Fragen stand der Experimentleiter jederzeit zur Verfügung. Um die Wahrscheinlichkeiten unterschiedlicher Versuchspersonen, für jedes Objekt miteinander vergleichen zu können, wurden diese in den gefilterten Zeitintervallen (Sakkaden, Lidschläge) auf den Wert null gesetzt. Dadurch hatten die Datenmatrizen aller Personen dieselbe Länge, was die Anwendung statistischer Maße ermöglichte. Daraus resultieren auch die kurzzeitigen „Einbrüche“ der Graphen, die in den Abbildungen C.2, C.3 und C.4 zu sehen sind.

6.3 Ergebnisse der Evaluierungsversuche

Da die endgültigen Daten in MATLAB erzeugt wurden, erfolgte deren Auswertung ebenfalls dort. Hierzu wurden weitere Skripte zur Stapelverarbeitung mehrerer Protokolle und zur Visualisierung der Ergebnisse angelegt. Diese Ergebnisse setzen sich zusammen aus dem Resultat der ersten (Graphendiagramm) und zweiten (Balkendiagramm) Stufe des 3-Phasen-Modells zusammen. Aufgrund der hohen Anzahl der Abbildungen, konnten nicht alle in diese Arbeit integriert werden. Deshalb musste sich auf je ein qualitativ gutes und schlechtes Beispiel beschränkt werden. Es wurden jedoch die Gesamtwahrscheinlichkeiten für jedes Objekt über alle Personen gemittelt sowie entsprechende Standardabweichungen berechnet, um den Erfolg des Verfahrens im Mittel quantifizieren zu können. Tabelle 6.2 macht diesen Sachverhalt noch deutlicher. Zusätzlich wird auf Unterschiede zwischen zwei Gruppen „Brillen-/Kontaktlinsenträger“ und „normalsichtige Personen“ eingegangen. Die jeweiligen Diagramme können in Anhang C eingesehen werden.

Es wurde sich für die Aggregation der Wahrscheinlichkeiten pro Objekt über alle teilnehmenden Personen entschieden, da die entsprechenden Standardabweichungen klein sind. Diese Tatsache wird durch die, über alle Samples gemittelte, interpersonelle Standardabweichung der Gesamtwahrscheinlichkeiten in Abbildung C.1 dargestellt. Diese niedrigen Werte geben an, dass die Fixierungswahrscheinlichkeiten aller Personen nur wenig von deren Mittelwert abweichen. Das heißt, dass das Verfahren ähnliche Ergebnisse für jede Versuchsperson produziert und dass alle Probanden die Aufgaben im Schnitt auf die selbe Weise erfüllt haben.

Auswertung von Versuch I

Wie in Abbildung 6.4 zu sehen ist, lässt sich das fixierte Objekt zu jedem Zeitpunkt korrekt und auf Basis der Metrik eins (Unsicherheitsbereich) bestimmen, wenn die Objekte weiter als der Unsicherheitsbereich voneinander entfernt sind. Die Überschneidung der Graphen stammt

Versuch	Versuchsperson													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
I	(✓)	(✓)	(✓)	(✓)	(✓)	✓	✓	(✓)	✓	✓	✓		(✓)	✓
II	✓		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
III	✓	✓	✓	✓	(✓)	(✓)	✓	(✓)	(✓)	✓	(✓)	✓	(✓)	✓

Tabelle 6.2: Ergebnisvergleich nach Korrektheit zwischen allen Versuchen und Probanden. Diese Tabelle wurde auf Basis der innerhalb von Fixierungen aggregierten Wahrscheinlichkeiten erstellt. Ein Haken bedeutet, dass der Algorithmus das durch die jeweilige Aufgabe vorgegebene Blickmuster erkannt hat. Kein Haken bedeutet, dass in diesem Versuch bei einigen Fixierungen fokussierte Objekte eine niedrigere Wahrscheinlichkeit erhielten als andere. Bei einem in Klammern gesetzten Haken hatte der Betrachter Probleme, den konkreten Anweisungen zu folgen und das markierte Objekt zu fixieren. Beispielsweise traten einerseits viele Fixierungen von Sekundärobjekten auf (vorrangig Versuch III). Andererseits konnten Fixierungen eng beieinander liegender Objekte nicht eindeutig unterschieden werden (vorrangig Versuch I).

aus der Mittelung über alle Versuchspersonen, da jeder Proband, aufgrund unterschiedlicher Reaktionszeiten, die Sakkade zum nächsten Objekt, zu leicht verschiedenen Zeiten, initiiert hat. Da jedoch SK_{UL} und PK_{UR} innerhalb des Unsicherheitsbereiches liegen, wenn eine der beiden Kugeln betrachtet wird, ist die Wahrscheinlichkeit in Abbildung 6.4 für beide Objekte annähernd gleich. Die Kombination der beiden entwickelten Metriken hebt jedoch die Fixierungswahrscheinlichkeit für SK_{UL} leicht an (vgl. Abbildung C.5(a)). Da die Abstände beider Objekte jedoch sehr klein sind bringt in diesem Fall auch die Kombination dieser beiden Metriken kein zufriedenstellendes Ergebnis, was ebenfalls aus Tabelle 6.2 deutlich wird. Zusätzlich wurde bei keinem der Probanden eine Fixierung von PK_{OR} festgestellt, während SK_{OL} fokussiert wurde. Aus diesen Gründen kann die erste Hypothese akzeptiert werden.

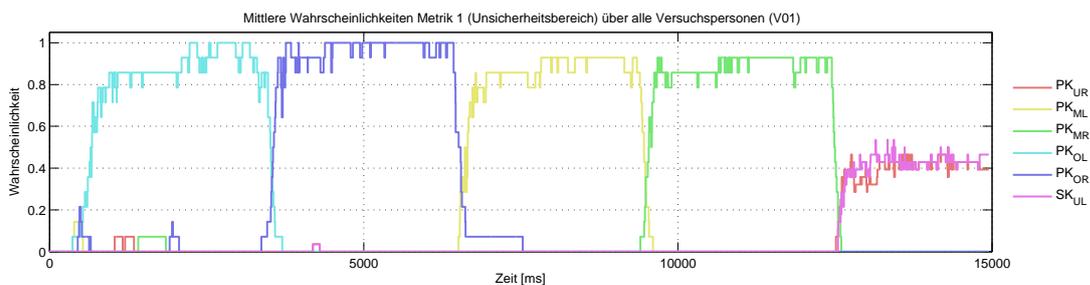


Abbildung 6.4: Wahrscheinlichkeit, resultierend aus dem Mittelwert der ersten Metrik (vgl. Tabelle 4.4) für jedes Objekt über alle Probanden.

In dem qualitativ guten Beispiel zum ersten Versuch (vgl. Abbildungen C.2(a) und C.2(b)) hat der Algorithmus alle, durch die gestellte Aufgabe vorgegebenen Blickmuster, korrekt erkannt und den Objekten entsprechend hohe Wahrscheinlichkeiten zugewiesen. Im qualitativ schlechten Beispiel (vgl. Abbildungen C.2(c) und C.2(d)) ist auffällig, dass die VP mehr Fixierungen machte und vor allem am Beginn des Experimentes sich zunächst die ganze Szene betrachtete anstatt die Aufgabe zu erfüllen. Aus diesem Grund entsteht auch hier ein unerwartetes Blickmuster. Dies zeigt deutlich, dass Blickmuster oft ungeahnte und schwer (automatisch) analysierbare Formen annehmen können.

Auswertung von Versuch II

Bei der Betrachtung von Abbildung C.5(b) fällt auf, dass die Wahrscheinlichkeiten der Fixierung von Kugel und Quader sich einander annähern, wenn sich beide Objekte gegenseitig verdecken, obwohl die ganze Zeit die Kugel fixiert wurde. Dieses Verhalten ist jedoch sinnvoll, da sich in diesem Zeitraum beide Objekte im Fokus befinden. Durch die Fixierung der Kugel über den gesamten Zeitraum des Experimentes und die Mittelung der berechneten Wahrscheinlichkeiten innerhalb jeder Fixierung, erzeugt das Verfahren in fast jedem Fall die erwartete Aussage (vgl. Tabelle 6.2 und Abbildung C.3). Außerdem wäre die Auflösung von Doppeldeutigkeiten verursacht durch Transparenzen ohne die zweite Metrik nicht möglich. Bei Verdeckung liegen beide Objekte im Unsicherheitsbereich und erhielten auf alleiniger Basis der ersten Metrik die selbe Wahrscheinlichkeit.

Aus Abbildung C.3(c) ist des Weiteren gut zu erkennen, dass die Versuchsperson, bei der semi-transparenten Verdeckung kurzzeitig von der roten Kugel abgewichen ist und den Quader betrachtet hat (Sprünge im Zeitbereich 3-4 Sekunden und ab 14 Sekunden). Dieses Verhalten hat sich ebenfalls auf die in Abbildung C.3(d) dargestellten Fixierungen ausgewirkt. Die Fixierungen 4, 14, 16, 17 zeigen daher, dass der Quader betrachtet wurde. Dies stellt im Grunde jedoch kein falsches Ergebnis dar.

Aus den gewonnenen Erkenntnissen kann geschlussfolgert werden, dass das Verfahren in der Lage ist, Fixierungen semi-transparent verdeckter Objekte zu erkennen. Dies ist jedoch nur möglich, da die Animation der Szene ausreichend lange Fixierungen (Folgebewegungen) ermöglicht und die Wahrscheinlichkeiten innerhalb dieser Fixierungen gemittelt werden. Daher können die für den zweiten Versuch aufgestellten Hypothesen angenommen werden.

Auswertung von Versuch III

K_3 und K_4 aus Experiment III sind weiter voneinander entfernt als SK_{UL} und PK_{UR} aus Versuch I. Dies wird auch besonders in Abbildung C.5(c) deutlich. Während der ersten drei Sekunden ist klar zu erkennen, dass K_4 fixiert wurde. Die sich anschließenden Fixierungswechsel zwischen den Objekten K_3 und K_4 kennzeichnet das Verfahren mit weitaus größeren Wahrscheinlichkeitsdifferenzen als die Fixierung von SK_{UL} und SK_{UR} bei Versuch I. Auch die sich anschließenden Fixierungen der Kugeln K_2 und K_3 sind deutlich hervorgehoben. Letzteres lässt darauf schließen, dass die Probanden allen sich bewegenden Objekten problemlos folgen können. Die Wahrscheinlichkeitsdifferenzen bei den Fixierungswechseln sind ebenfalls groß genug, um die Fixierungen eindeutig den korrekten Objekten zuordnen zu können. Diese Erkenntnisse unterstützen somit die für den dritten Versuch aufgestellten Hypothesen.

Tabelle 6.2 bestätigt, dass ein Großteil der Versuche erfolgreich ausgewertet werden konnte. Der andere Teil wurde als eingeschränkt erfolgreich klassifiziert, da die Versuchspersonen, vor allem zu Beginn des Experimentes, zunächst die gesamte Szene analysiert haben und sich nicht auf das jeweils markierte Objekt konzentrierten (bewusst oder unbewusst). Dieser Effekt ist im qualitativ schlechten Beispiel (Abbildung C.4(c)) deutlich zu erkennen und ruft im Vergleich zu Abbildung C.4(b) unerwartete Blickmuster, vor allem zu Beginn des in Abbildung C.4(d) ausgewerteten Experimentes, hervor. Da bereits bei solch einfachen Szenen diese Probleme auftreten, wird vermutet, dass in Experimenten mit realen Szenen noch größere Schwierigkeiten bei der Identifikation eindeutiger Blickmuster zu erwarten sind.

Brillen-/Kontaktlinsenträger gegenüber normalsichtigen Personen

Im Allgemeinen wird beschrieben, dass ETS's Schwierigkeiten bei der Blickpunktregistrierung haben, wenn die Versuchsperson eine Brille bzw. Kontaktlinsen trägt. Dieser Sachverhalt ist sehr interessant, da heutzutage ein Großteil der Bevölkerung eine Sehilfe benötigt. Es ist daher wünschenswert, dass dieser Personenkreis ebenfalls an Eye-Tracking-Experimenten teilnehmen kann. Um zu überprüfen, ob das Verfahren Schwierigkeiten mit der Blickpunktregistrierung bei Brillen-/ Kontaktlinsenträgern hat, wurden die 14 teilnehmenden Probanden in je eine Gruppe aus zehn normalsichtigen Personen und aus fünf Brillen-/ Kontaktlinsenträgern (Probanden fünf, acht, zehn, elf und vierzehn) geteilt. Die Gesamtwahrscheinlichkeiten für jedes Sample wurden objektweise für beide Gruppen über die Personen gemittelt. Nun wurden für korrespondierende Objekte die Standardabweichungen für jeden Zeitpunkt berechnet und anschließend über die Zeit gemittelt. In Abbildung 6.5 sind die resultierenden Werte zu sehen.

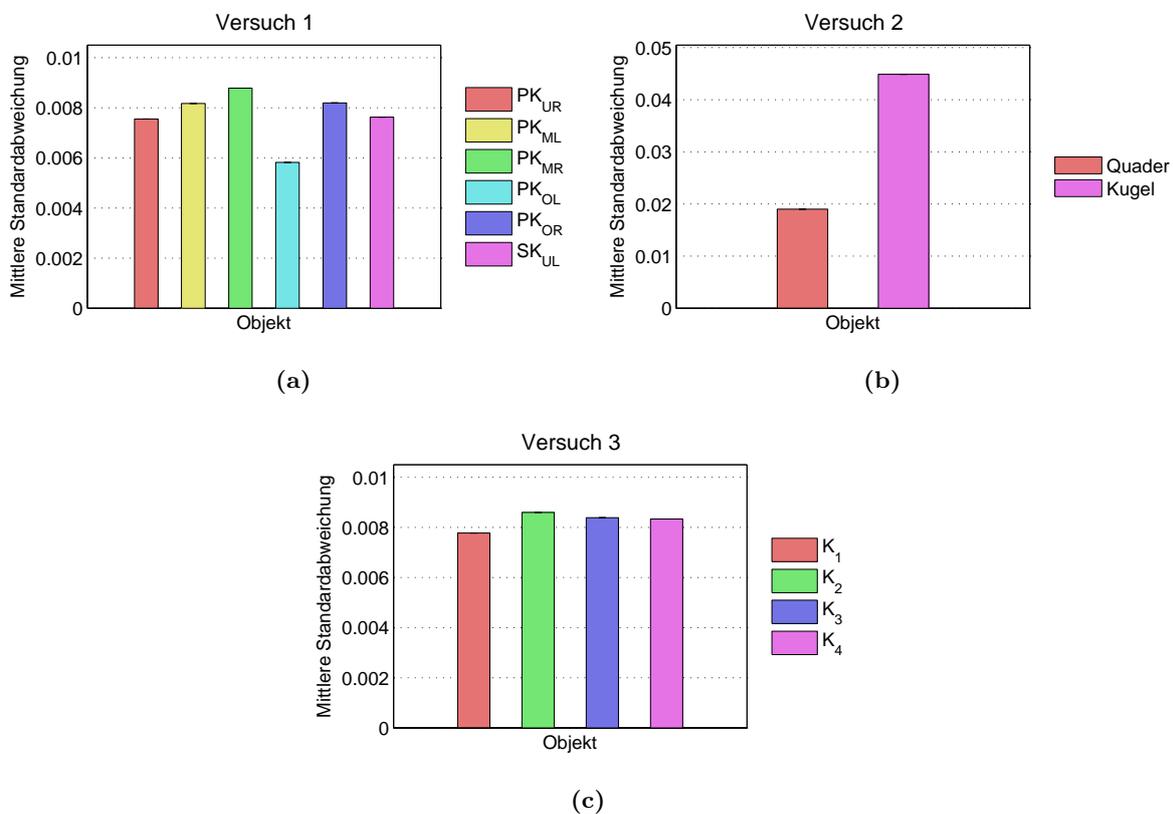


Abbildung 6.5: Die Abbildungen stellen für jedes Objekt des jeweiligen Versuches die mittlere Standardabweichung der gemittelten Gesamtwahrscheinlichkeiten (pro Objekt) zwischen den zwei Gruppen „Brillen-/ Kontaktlinsenträger“ und „normalsichtige Personen“ dar.

Diese Werte sind äußerst gering. Auch während der Experimente konnten keine nennenswerten Einschränkungen festgestellt werden. Die Kalibrierung lieferte ein optimales Ergebnis und Reflexionen durch Kontaktlinsen bzw. Brillen waren auf den Videobildern nicht zu sehen. Aus diesen Gründen kann davon ausgegangen werden, dass EYELINK I auch für Brillen-/ Kontaktlinsenträger akzeptable Blickpunkte liefert und das entwickelte Verfahren damit keiner Einschränkung unterliegt. Diese Vermutung kann darauf zurückgeführt werden, dass die

Kameras des ETS so dicht vor dem Auge positioniert werden können, dass tatsächlich nur das Auge und keine Bestandteile des Brillenrahmens aufgenommen werden. Die Kontaktlinsenträger verwendeten alle Tageslinsen, welche sich durch ihre geringe Dicke und hohe Transparenz auszeichnen. Tabelle 6.2 untermauert die hier postulierte Vermutung, da sich die Ergebnisse für die Gruppe der Brillen-/ Kontaktlinsenträger, außer für Person fünf, im Rahmen befinden.

6.4 Auswertung des Fragebogens

Insgesamt haben an den Experimenten 14 Personen im Alter von 20 bis 30 Jahren teilgenommen. Das Durchschnittsalter betrug 24 Jahre. Von den teilnehmenden Personen waren jeweils die Hälfte männlich und weiblich. Außer drei Personen hatte niemand vor dem Experiment Kaffee zu sich genommen. Bei ersteren konnten jedoch auch keine Defizite bei der Kalibrierung festgestellt werden. Außerdem haben alle Probanden angegeben, nicht unter dem Einfluss von Medikamenten zu stehen. Die Kalibrierung und die Experimente konnten zudem für alle Versuchspersonen erfolgreich durchgeführt werden. Alle Probanden wurden aus dem fakultätsnahen, studentischen Umfeld gewählt. Aus diesem Grund gaben die meisten an, sich gut bis sehr gut mit Computern auszukennen (vgl. Abbildung B.2(c)). Da Eye-Tracking-Studien an dieser Fakultät jedoch eher selten sind, hatten die teilnehmenden Personen auf diesem Gebiet auch wenig Erfahrung (vgl. Abbildung B.2(d)). Die meisten gaben an, „schon einmal etwas davon gehört“ zu haben. Durch diese neue Erfahrung fanden die meisten Probanden die Versuche interessant und waren auch dementsprechend hoch motiviert.

Es wurden zwei Personen mit Brille sowie drei Personen mit Kontaktlinsen getestet. In beiden Fällen hat es nicht den Anschein, dass sich diese Umstände nachteilig auf die Ergebnisse auswirken. Der Großteil der Probanden gab an, vor dem Experiment ausgeschlafen und entspannt zu sein. Somit konnte ein hoher Konzentrationsgrad angenommen werden. Für die Personen vier und elf, welche beim Müdigkeits- und Entspannungsgrad die Note schlecht bzw. mittelmäßig vergeben hatten, konnte trotzdem ein gutes Ergebnis erzielt werden (vgl. Tabelle 6.2, Abbildung B.2(a) und Abbildung B.2(b)).

Anhand der Ergebnisse aus den Abbildungen B.3(a), B.3(c), B.3(d) und B.3(f) kann geschlossen werden, dass das Experiment für den Großteil der Probanden angenehm war. Die einzige Ausnahme bildet das Eye-Tracking-Gerät, welches mit der Zeit als relativ schwer empfunden wurde (vgl. Abbildung B.3(b)). Außer bei einer Person (Proband eins) konnten sich alle Personen mittelmäßig bis sehr gut konzentrieren und wurden kaum während der Experimente abgelenkt. Der Großteil der Ablenkung stammte von sich bewegenden Personen im Raum oder im benachbarten Raum, der aufgrund der Trennung durch eine Glasscheibe im Sichtfeld der VP's lag. Die Kommentare der Versuchspersonen konnten eindeutig klassifiziert werden. Im Folgenden werden diese sowie die Häufigkeiten, mit denen sie geäußert wurden, dargestellt:

1. Kalibrierung ist unangenehm und häufig. [7]
2. Unbewusste Versuchung folgende Marker bei der Kalibrierung vorherzusehen. [4]
3. Eye-Tacking-Gerät war schwer und unbequem. [4]
4. Kalibrierung war zu schnell [1]

5. Besser nicht vorher sagen, wo die zu fixierenden Punkte im Experiment auftreten, da sonst versucht wird, diese vorherzusehen. [1]

Punkt eins kann vermieden werden, indem die Tatsache ausgenutzt wird, dass der Eye-Tracker das Kalibrierungsergebnis speichert. Die Versuchsperson darf sich dann jedoch, während des gesamten Experimentes, nicht bewegen. Punkt zwei führte bei vielen Personen dazu, dass die Kalibrierung mehrmals durchgeführt werden musste, wodurch die Häufigkeit bei Punkt eins zustande kommt. Dieser Aspekt kann durch eine qualitativ hochwertigere Erläuterung und das Training des Kalibrierungsprozesses vermieden werden. Die Unbequemlichkeit des Gerätes ist nur durch den Austausch des Eye-Trackers mit einer dektoplatzierten Variante bzw. einer kopfbefestigten Variante, wie dem EYELINK II-System, zu umgehen. Letzteres ist leichter als die hier verwendete Version. Die in Punkt vier kritisierte Geschwindigkeit der Kalibrierung lässt sich über die Programmierschnittstelle des ETS anpassen. Die letzte Aussage kam zu stande, da aufgrund hoher Genauigkeitsfehler das gesamte Experiment bei einigen Probanden mehrmals durchgeführt werden musste. Dies führte teilweise zum Drang den Experimentverlauf vorhersehen zu wollen und damit zu unerwarteten Blickmustern bei den Probanden.

6.5 Schlussfolgerungen

Eine Schwierigkeit bei der Auswertung stellte die unterschiedliche Anzahl erkannter Fixierungen pro Person dar. Da somit die Zahlenbasen der Balkendiagramme nicht einheitlich sind (vgl. Anhang C), wurden primär die nicht aggregierten Wahrscheinlichkeiten zu jedem Zeitpunkt, dargestellt als Graphen, für die Evaluierung verwendet. Zur automatischen Auswertung der Ergebnisse aus Stufe zwei des 3-Phasen-Modells muss daher ein geeignetes Verfahren entwickelt werden, welches interpersonelle Daten, auf dieser Ebene, vergleichbar macht.

In allen Versuchen hat sich zudem gezeigt, dass sich die Probanden mit den ersten Fixierungen, zu Beginn des Experimentes, einen Überblick über die Szene verschaffen. Da dieses Verhalten vom Probanden abhängt, müssen solche Fixierungen genau analysiert werden. Auch das Verwerfen dieser Daten wäre denkbar, wenn sie nicht benötigt werden. Allerdings könnten gerade die ersten Fixierungen wertvolle Informationen über besonders hervorstechende Objekte enthalten.

Eine Schwierigkeit die vor allem im ersten Versuch deutlich geworden ist, liegt in der korrekten Zuordnung von Fixierungen eng beieinander liegender Objekte, welche für die Evaluierung medizinischer Visualisierungen von großer Bedeutung ist. Hier ist es unbedingt notwendig, die Genauigkeit des ETS mittels der in Abschnitt 3.1.1 beschriebenen Verfahren zu verbessern. Da jedoch von 28.5 % der Versuchspersonen die Unbequemlichkeit des Gerätes durch Kommentare bestätigten, wäre die Anschaffung eines moderneren Eye-Tracking-Gerätes eine lohnenswerte Investition. Das EYELINK II-System ist beispielsweise 30 % leichter als EYELINK I und hat im Gegensatz zu diesem ETS einen maximalen Genauigkeitsfehler von nur 0.5°.

Abschließend wurde das Verfahren getrennt für die zwei Gruppen „Brillen-/ Kontaktlinsenträger“ und „normalsichtige Personen“ getestet. Da die erste Gruppe aus nur fünf Personen bestand, kann noch nicht von repräsentativen Ergebnissen gesprochen werden. Jedoch ergab sich in folgenden drei Bereichen, dass deren Ergebnisse qualitativ gleichwertig zu denen von normalsichtigen Menschen sind:

- Die Kalibrierung konnte mit einem optimalen Ergebnis durchgeführt werden
- Die Visualisierung zur Kontrolle der Genauigkeit des erzielten Ergebnisses (vgl. Abbildung 5.4), ließ keine Zweifel an deren Qualität aufkommen
- Die gemittelten Wahrscheinlichkeitsdaten beider Gruppen wichen nur minimal voneinander ab
- Die ermittelten Standardabweichungen waren sehr niedrig

Nachdem die gesamte Arbeit, in seinen Einzelheiten vorgestellt wurde, folgt deren Zusammenfassung im nun folgenden Kapitel.

Kapitel 7

Zusammenfassung und Ausblick

In dieser Arbeit erfolgte die Konzipierung eines Verfahrens zur automatischen Erkennung betrachteter Objekte einer 3D-Szene mittels Eye-Tracking. Es wurden wichtige Grundlagen bezüglich dieses Themengebietes anhand einer intensiven Literaturrecherche zusammengefasst. Einen zusätzlichen Gewinn repräsentieren die erstmals durch Umsetzung und Evaluierung des Verfahrens gewonnenen Erkenntnisse hinsichtlich der Güte der entstehenden Daten und damit der Eignung des Eye-Tracking-Systems EYELINK I für diese Aufgabe.

In einer umfangreichen Anforderungsanalyse wurden Charakteristiken von medizinischen Visualisierungen und von Eye-Tracking-Systemen herausgearbeitet. Untersucht wurden die räumlichen Verhältnisse anatomischer Strukturen, Szenenkomplexität, Materialeigenschaften, Genauigkeit des Eye-Trackers und die Präsentationsart der Visualisierung. Durch Berücksichtigung dieser Gesichtspunkte war es möglich ein Verfahren für die Zuordnung von Fixierungen zu Objekten auf Basis von Wahrscheinlichkeiten zu entwickeln sowie das Ziel „Beseitigung von Doppeldeutigkeiten verursacht durch Transparenz“ zu erreichen.

Dazu erfolgte zunächst die Entwicklung des 3-Phasen-Modells, das zur Beschreibung der für die Zuordnung von Fixierungen zu Objekten notwendigen Schritte dient. Es steigert die Übersichtlichkeit des Gesamtverfahrens und ermöglicht die dortige Einordnung aller Teilschritte.

Die Erkenntnisse der Anforderungsanalyse wurden sinnvoll mit angepassten, in Kapitel drei vorgestellten Techniken kombiniert, um somit die Ziele dieser Arbeit zu erfüllen. Den Hauptanteil der Arbeit umfasste die Entwicklung geeigneter Metriken auf Basis der in der Anforderungsanalyse herausgearbeiteten Kriterien. Zwei dieser Metriken wurden umgesetzt und es konnte durch die Evaluierung gezeigt werden, dass deren Kombination auch eine korrekte Zuordnung von Fixierungen zu semi-transparent verdeckten Objekten ermöglicht.

In MEVISLAB wurde die Generierung der für die Auswertung nötigen Daten erfolgreich in Form eines Netzwerkes umgesetzt. Aufgrund des hohen Zeitbedarfs, erfolgte die Vorbereitung weiterer Metriken lediglich konzeptuell. Da die endgültige Implementierung in C++ erst dann als sinnvoll erschien, wenn alle Metriken zur Verfügung stehen, wurde sich für die Datenauswertung in MATLAB entschieden. Dies ermöglicht, im Gegensatz zu C++, die einfache Modifikation sowie Präsentation großer Datenmengen auf experimenteller Ebene.

Das in MEVISLAB entwickelte Rahmenprogramm zur Ansteuerung von EYELINK I konnte zudem erfolgreich für eine Studie zur Applikationsevaluierung mittels Eye-Tracking-Daten eingesetzt werden. Dies zeigt, dass auch ein Anwendungsübergreifender Einsatz der im Rahmen dieser Diplomarbeit entwickelten Software möglich ist.

Diese Zusammenfassung sollte verdeutlicht haben, dass eine komplexe Entwicklung geschaffen und wichtige Erkenntnisse gewonnen wurden. Die zu Beginn der Arbeit definierten Ziele werden dadurch vollständig erfüllt.

7.1 Weiterführende Entwicklungen

Eine Anwendung des Verfahrens auf alle Objektformen, wie größere und vor allem elongierte bzw. verzweigte Strukturen (vgl. Abbildung 7.1(a)) ist zur Evaluation medizinischer Visualisierungen essentiell. Die entwickelte Methodik konnte aus Zeitgründen zunächst nur an kleine, kompakte Objekte adaptiert werden. Um den oben benannten Sachverhalt zu ermöglichen, müssen weitere Techniken in das bestehende Verfahren integriert und getestet werden. Beispielsweise wäre es denkbar, die Methodik aus Lu et al. [2006], zur Bestimmung häufig betrachteter Regionen an Oberflächenmodelle anzupassen. Ein Vorschlag wäre hier, um jeden Strahl-/Objektschnittpunkt eine Kugel mit dem Radius des Unsicherheitsbereiches zu legen und zu untersuchen, welche Knoten des zugrundeliegenden Polygonnetzes innerhalb der Kugel liegen und wie oft sich diese im Fokus befinden. Eine daran angepasste Datenstruktur müsste dazu die Knoten zusammen mit den Häufigkeiten speichern. Wird außerdem zu den Knotenpunkten das zugehörige Objekt gespeichert, ist es durch eine sich anschließende Analyse möglich dieses Verfahren in das bestehende Metrikensystem zu integrieren.

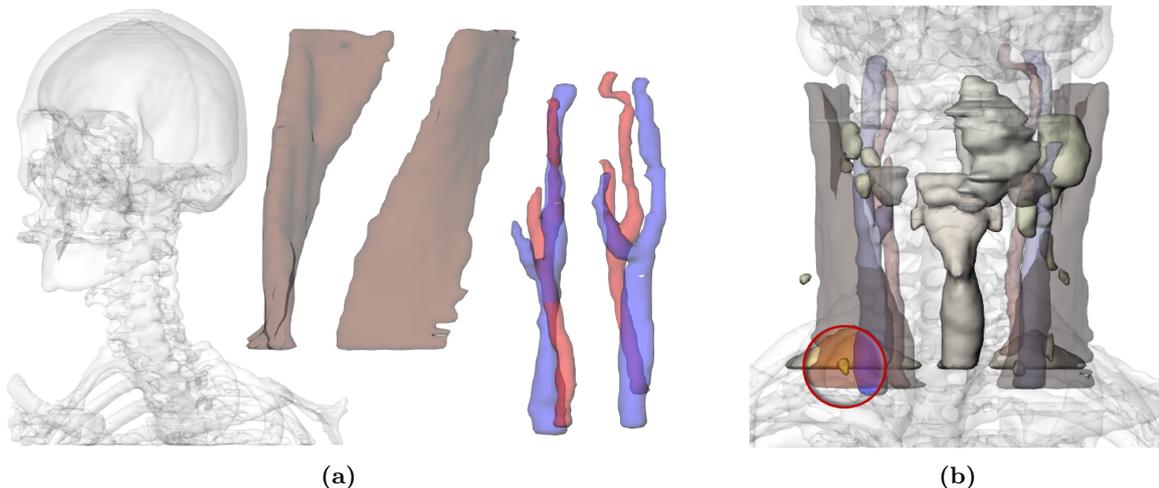


Abbildung 7.1: Abbildung (a) zeigt einen Vergleich der verschiedenen Strukturformen groß (Knochen), elongiert (Halsmuskel) und verzweigt (Arterien, Venen). In Abbildung (b) sind die Objekte hervorgehoben, welche im Falle der Fixierung des Lymphknotens durch die Objektidentifizierung erkannt werden.

Die Erkennung fixierter Teilbereiche einer Struktur, statt die Zuordnung von Fixierungen zu ganzen Objekten, ist ein großer Vorteil des oben beschriebenen Vorgehens. Mit diesem objektbasierten Ansatz kann zudem das in der Arbeit angesprochene Problem der Unterscheidung von Fixierungen der Vorder- und Rückseite eines Objektes gelöst werden. Diese Tatsache beruht jedoch auf der Annahme, dass die Szene interaktiv bzw. automatisch bewegt wird. Hieraus ergibt sich eine weitere Verbesserungsmöglichkeit. Bislang ist das Verfahren ungeeignet für die Anwendung auf statische Stimuli. Hierbei müssen an diese Reize angepasste Möglichkeiten zur Auflösung von Doppeldeutigkeiten durch semi-transparente Verdeckungen untersucht werden.

Zur Anwendung des Verfahrens auf statische Stimuli wird es als hilfreich angesehen, alle in Tabelle 4.4 aufgeführten Metriken zu implementieren. Nachdem dies geschehen ist, kann das in Kapitel 4 vorgestellte Gesamtsystem evaluiert werden. Vor allem die „Anzahl

tatsächlich projizierter Pixel“ [NEUGEBAUER, 2006] sollte ein nützliches Maß bezüglich statischer Stimuli darstellen. Wird beispielsweise häufig der in Abbildung 7.1(b) hervorgehobene Bereich fixiert, so ist anzunehmen, dass das Interesse des Betrachters dem Lymphknoten anstatt den diesen umgebenden Strukturen gilt. Auf Basis der neuen Metrik und der Häufigkeit der fixierten Region müssen also die umgebenden Objekte in ihrer Fixierungswahrscheinlichkeit abgeschwächt werden.

Fixierungen semi-transparenter Strukturen im Vordergrund müssen neben der korrekten Zuordnung von Fokussierungen kleiner Hintergrundobjekte ebenfalls berücksichtigt werden. Hierbei ist zunächst interessant, wie häufig und zu welchen Zeitpunkten diese über das Experiment betrachtet wurden. Diesbezüglich ist es zweckmäßig, die Eignung des oben vorgestellten, objektbasierten Ansatzes nach LU et al. [2006] analysieren.

Bei Anwendung dieses Ansatzes sollte untersucht werden, inwiefern er sich zur Repräsentation der Objektposition eignet, da der bislang verwendete, **projizierte Mittelpunkt der Bounding-Box** nur eine Annäherung dieses Wertes darstellt. Um die Abstandsmetrik jedoch weiter nutzen zu können, müsste jedes Objekt zunächst in Substrukturen bzw. Teilflächen unterteilt werden. Nach der Berechnung der Bounding-Box für jede Substruktur, kann der Abstand des Blickpunktes zu den projizierten Zentren der Hüllkörper berechnet und die mittels eines Schwellenwertes ermittelten, geringsten Distanzen in eine weitere Analyse einbezogen werden.

Die Durchführung der dritten Stufe des 3-Phasen-Modells und der damit verbundenen Auswertung der resultierenden Daten ist je nach Anwendungsgebiet möglich, nachdem die Wahrscheinlichkeiten für jedes Objekt und alle Fixierungen vorliegen (Ergebnisse der zweiten Stufe). Es wurden Informationen vorgestellt, die im Rahmen einer globalen Analyse generiert werden können.

Hierbei ist zu bemerken, dass diese Informationen große Ähnlichkeit mit den klassischen Metriken haben. Sie können durch das Anwenden der Erkenntnisse aus entsprechenden Eye-Tracking-Studien mit den Ergebnissen der ersten bzw. zweiten Stufe des 3-Phasen-Modells kombiniert werden. Diese globale Analyse ist jedoch erst dann sinnvoll durchführbar, wenn die Ergebnisse aus Stufe eins und zwei für alle Objekttypen korrekte Resultate liefern.

Zur Beschleunigung der Schnittpunktberechnung mit den polygonalen Geometriemodellen sollte untersucht werden, wie sich die lineare Abhängigkeit von der Anzahl der Polygone und der Anzahl der Blickstrahlen für ein Sample mittels Hüllkörper und der Ausnutzung räumlicher Zusammenhänge beschleunigen lässt. Mittels Hüllkörper ist es nicht mehr nötig alle Polygone auf Schnitt zu untersuchen. Eine Möglichkeit zur Ausnutzung räumlicher Kohärenzen stellt die Verwendung von Samplingstrategien bei der Berechnung der Sekundärstrahlen des Unsicherheitsbereiches dar. Somit ist es nicht mehr nötig, alle Blickstrahlen sondern nur ein gewisses Muster weniger Strahlen auf Schnitt zu untersuchen. Dieses Muster muss in derart konstruiert werden, dass auch kleine Objekte in jedem Fall erfasst werden können. Außerdem wäre es sinnvoll zeitlich nahe beieinander liegende Samples, z.B. innerhalb einer Fixierung (bei statischem Stimulus), zusammenzufassen.

Die Auslegung der Gewichte der Metriken stellt einen weiteren Punkt dar, den es zu ergründen gilt. Bislang wurde jedes Maß mit denselben Wertigkeiten belegt. Es muss jedoch untersucht werden, inwiefern sich diesbezügliche Änderungen auf die Gesamtwahrscheinlichkeit auswirken.

Auf die Genauigkeit des Eye-Tracking-Systems wurde an mehreren Stellen dieser Arbeit eingegangen. Speziell EYELINK I hat einen relativ ungünstigen Wert von maximal 1°.

Eine denkbare Lösung wäre, die in Abschnitt 3.1.1 diskutierten Kalibrierungsverfahren und Methoden zu Driftkorrektur umzusetzen und zu evaluieren. Die andere Möglichkeit besteht in der Anschaffung eines neuen Eye-Tracking-Systems. EYELINK II beispielsweise verfügt über eine höhere Genauigkeit und ermöglicht durch sein geringeres Gewicht (im Vergleich zu EYELINK I) eine niedrigere Belastung der Probanden.

Bei einem weitreichenden Blick in die Zukunft erscheint die Entwicklung **geeigneter Visualisierungsmöglichkeiten** als sinnvolle Ergänzung zum eigentlichen Verfahren. Neben der Verbesserung, der in Abbildung 5.4 vorgestellten Möglichkeiten, kann ebenfalls über die direkte Modifikation von Materialeigenschaften auf Basis der Blickpunktdaten nachgedacht werden.

Zur Zeit werden die Samples einer Fixierung genutzt, um die neu entwickelten Metriken zu Berechnen. Da Fixierungen jedoch relativ viele Samples umfassen würden sich Sakkaden als Alternative anbieten. In Abschnitt 3.1.3 wurde erläutert, dass diese den Blick über dem Objekt zentrieren, welches anschließend fixiert werden soll. Es sollte daher untersucht werden, inwiefern die **Sakkadenanalyse** bei der Auswertung rechenintensiver, objektbasierter Metriken zu validen Ergebnissen führt.

Die Bestimmung fokussierter Objekte in dreidimensionalen, medizinischen Visualisierungen mittels Eye-Tracking ist ein Gebiet, auf dem noch wesentlicher Forschungsbedarf nötig ist. In dieser Arbeit wurden mit der Entwicklung eines möglichen Verfahrens die Grundlagen für das weitere Vorgehen geschaffen. In diesem Abschnitt sind viele Möglichkeiten zur Fortführung der Themenstellung in verschiedenster Hinsicht aufgezeigt worden. Aus diesen Gründen kann behauptet werden, dass ein bedeutender Beitrag im Bereich der Mensch-Computer-Interaktion und der Evaluierung medizinischer Visualisierungen geleistet wurde.

Abkürzungen und Akronyme

AOI	Area of Interest
CAD	Computer Aided Design
CPU	Central Processing Unit
dpi	Dots Per Inch
ETS	Eye-Tracking-System
GB	Giga Byte
GHz	Giga Hertz
GUI	Graphical User Interface
ID	Identifikation
IP	Internet Protocol
LED	Light Emitting Diode
METK	Medical Exploration Toolkit
OPC	Operator-PC
PSOM	Parametrized Self-Organizing Map
SPC	Subject-PC
VP	Versuchsperson

Literaturverzeichnis

- [AALTONEN et al. 1998] AALTONEN, Antti ; HYRSKYKARI, Aulikki ; RÄIHÄ, Kari-Jouko: 101 Spots, or How do Users Read Menus? In: *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, 1998, S. 132–139
- [ANDREASSI 1995] ANDREASSI, John L.: *Psychophysiology: Human Behaviour and Physiological Response*. Lawrence Erlbaum Associates, 1995
- [BACKS und WALRATH 1992] BACKS, Richard W. ; WALRATH, Larry C.: Eye Movement and Pupillary Response Indices of Mental Workload During Visual Search of Symbolic Displays. In: *Applied Ergonomics* 23 (1992), Nr. 4, S. 243–254
- [BAER 2005] BAER, Alexandra: *Hardwaregestütztes Stippling von medizinischen Oberflächenmodellen*, Otto-von-Guericke Universität Magdeburg, Diplomarbeit, 2005
- [BEATTY 1982] BEATTY, Jackson: Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. In: *Psychological Bulletin* 91 (1982), Nr. 2, S. 276–92
- [BURGERT et al. 2007] BURGERT, Oliver ; ÖRN, Veronika ; JOOS, Markus ; STRAUSS, Gero ; TIETJEN, Christian ; PREIM, Bernhard ; HERTEL, Ilka: Evaluation of Perception Performance in Neck Dissection Planning using Eye-Tracking. In: JIANG, Yulei (Hrsg.) ; SAHINER, Berkman (Hrsg.): *Proceedings of SPIE Medical Imaging 2007: Image Perception, Observer Performance and Technology Assessment*, 2007
- [CARPENTER 1988] CARPENTER, Roger H. S.: *Movements of the Eyes*. Pion-Verlag, 1988
- [CAVANAGH und ALVAREZ 2005] CAVANAGH, Patrick ; ALVAREZ, George A.: Tracking multiple targets with multifocal attention. In: *Trends in Cognitive Sciences* 9 (2005), Nr. 7, S. 349–354
- [@CDC] CENTERS FOR DISEASE CONTROL AND PREVENTION: *Computer Workstation Ergonomics*. <http://www.cdc.gov/od/ohs/Ergonomics/compergo.htm>. – Link aufgerufen am 09.09.2008
- [COWEN 2001] COWEN, Laura: *An Eye Movement Analysis of Web-Page Usability*, Lancaster University, UK, Diplomarbeit, 2001
- [DONNER 2007] DONNER, Matthias: *Erfassung und Berechnung von Blickbewegungen auf dreidimensionalen Szenen*, Technische Fakultät der Universität Bielefeld, Diplomarbeit, 2007
- [DUCHOWSKI 2007] DUCHOWSKI, Andrew T.: *Eye Tracking Methodology: Theory and Practice*. Springer, 2007
- [@ECE] UNIVERSITY OF MAINE, DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING: *Visual Search - Where is waldo?* http://www.eece.maine.edu/research/gk12/Five_Senses/waldo.jpg. – Link aufgerufen am 20.10.2008

- [ENGBERT und KLIEGL 2003] ENGBERT, Ralf ; KLIEGL, Reinhold: Microsaccades Uncover the Orientation of Covert Attention. In: *Vision Research* 43 (2003), Nr. 9, S. 1035–1045
- [ESSIG et al. 2004] ESSIG, Kai ; POMPLUN, Marc ; RITTER, Helge: Application of a Novel Neural Approach to 3D Gaze Tracking: Vergence Eye-Movements in Autostereograms. In: FORBUS, Kenneth (Hrsg.) ; GENTNER, Dedre (Hrsg.) ; REIGER, Terry (Hrsg.): *Proceedings of the 26th Annual Meeting of the Cognitive Science Society, 2004*
- [ESSIG et al. 2006] ESSIG, Kai ; POMPLUN, Marc ; RITTER, Helge: A Neural Network for 3D Gaze Recording with Binocular Eye Trackers. In: *International Journal of Parallel, Emergent and Distributed Systems* 21 (2006), S. 79–95
- [FEHD und SEIFFERT 2008] FEHD, Hilda M. ; SEIFFERT, Adriane E.: Eye Movements During Multiple Object Tracking: Where do Participants Look. In: *Cognition* 108 (2008), July, Nr. 1, S. 201–209
- [GOLDBERG und KOTVAL 1998] GOLDBERG, Joseph H. ; KOTVAL, Xerxes P.: Eye Movement-Based Evaluation of the Computer Interface. In: *Advances in Occupational Ergonomics and Safety* (1998)
- [GOLDBERG et al. 2002] GOLDBERG, Joseph H. ; STIMSON, Mark J. ; LEWENSTEIN, Marion ; SCOTT, Neil ; WICHANSKY, Anna M.: Eye tracking in web search tasks: design implications. In: DUCHOWSKI, Andrew T. (Hrsg.) ; VERTEGAAL, Roel (Hrsg.) ; SENDERS, John W. (Hrsg.): *ETRA '02: Proceedings of the 2002 symposium on Eye tracking research & applications*, 2002, S. 51–58
- [GOLDSTEIN 1997] GOLDSTEIN, Eugen B. ; RITTER, Prof. Dr. M. (Hrsg.): *Wahrnehmungspsychologie: Eine Einführung*. Spektrum-Verlag, 1997
- [GOLDSTEIN 2008] GOLDSTEIN, Eugen B. ; IRTEL, Hans (Hrsg.): *Wahrnehmungspsychologie: Der Grundkurs*. Spektrum-Verlag, 2008
- [GOLDSTEIN et al. 2007] GOLDSTEIN, Robert B. ; WOODS, Russell L. ; PELI, Eli: Where People look when Watching Movies: Do all Viewers look at the Same Place? In: *Computers in Biology and Medicine* 37 (2007), Nr. 7, S. 957–964
- [HAFED und CLARK 2002] HAFED, Ziad M. ; CLARK, James J.: Microsaccades as an Overt Measure of Covert Attention Shifts. In: *Vision Research* 42 (2002), Nr. 22, S. 2533–2545
- [VON HELMHOLTZ 1867] HELMHOLTZ, Hermann von ; KÖNIG, Arthur (Hrsg.): *Handbuch der physiologischen Optik*. Voss, 1867
- [HENDRICKSON 1989] HENDRICKSON, J. J.: Performance, Preference, and Visual Scan Patterns on a Menu-Based System: Implications for Interface Design. In: *CHI '89: Proceedings of the SIGCHI conference on Human factors in computing systems SIGCHI (Veranst.)*, 1989, S. 217–222
- [HÄKKÄNEN et al. 1999] HÄKKÄNEN, Helinä ; SUMMALA, Heikki ; PARTINEN, Markku ; TIHONEN, Mikko ; SILVO, Jouni: Blink Duration as an Indicator of Driver Sleepiness in Professional Bus Drivers. In: *SLEEP* 22 (1999), Nr. 6, S. 798–802
- [HOEKS und LEVELT 1993] HOEKS, Bert ; LEVELT, Willem J. M.: Pupillary Dilation as a Measure of Attention: A Quantitative System Analysis. In: *Behavior Research Methods, Instruments, & Computers* 25 (1993), Nr. 1, S. 16–26

- [HORNOF und HALVERSON 2002] HORNOF, Anthony J. ; HALVERSON, Tim: Cleaning up systematic error in eye-tracking data by using required fixation locations. In: *Behavior Research Methods, Instruments, & Computers* 34 (2002), Nr. 4, S. 592–604
- [@INFERENCE GROUP] INFERENCE GROUP: *The Dasher Project*. <http://www.inference.phy.cam.ac.uk/dasher/>. – Link aufgerufen am 01.09.2008
- [JACOB und KARN 2003] JACOB, Robert J. K. ; KARN, Keith S.: Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises (Section Commentary). In: *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (2003), S. 573–605
- [JOSEPHSON und HOLMES 2002] JOSEPHSON, Sheree ; HOLMES, Michael E.: Visual Attention to Repeated Internet Images: Testing the Scanpath Theory on the World Wide Web. In: DUCHOWSKI, Andrew T. (Hrsg.) ; VERTEGAAL, Roel (Hrsg.) ; SENDERS, John W. (Hrsg.): *ETRA '02: Proceedings of the 2002 symposium on Eye tracking research & applications*, 2002, S. 43–49
- [KAPADIA et al. 1995] KAPADIA, Mitesh K. ; ITO, Minami ; GILBERT, Charles D. ; WESTHEIMER, Gerald: Improvement in Visual Sensitivity by Changes in Local Context: Parallel Studies in Human Observers and in V1 of Alert Monkeys. In: *Neuron* 15 (1995), S. 843–856
- [KARN 2000] KARN, Keith S.: „Saccade pickers“ vs. „fixation pickers“: the effect of eye tracking instrumentation on research. In: DUCHOWSKI, Andrew T. (Hrsg.): *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, 2000, S. 87–88
- [KAUZLIS 2005] KAUZLIS, Richard J.: The Control of Voluntary Eye Movements: New Perspectives. In: *The Neuroscientist* 11 (2005), Nr. 2, S. 124
- [KOLERS et al. 1981] KOLERS, Paul A. ; DUCHNICKY, Robert L. ; FERGUSON, Dennis C.: Eye Movement Measurement of Readability of CRT Displays. In: *Human Factors and Ergonomics Society* 23 (1981), Nr. 5, S. 517–27
- [KOTVAL und GOLDBERG 1998] KOTVAL, Xerxes P. ; GOLDBERG, Joseph H.: Eye Movements and Interface Components Grouping: An Evaluation Method. In: *Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* HFES (Veranst.), 1998, S. 486–490
- [@KRUPINSKI] KRUPINSKI, Elizabeth A.: *Eye-Position Recording Studies*. <http://www.radiology.arizona.edu/KRUPINSKI/epr.htm>. – Link aufgerufen am 20.10.2008
- [KRUPINSKI 2000] KRUPINSKI, Elizabeth A.: The Importance of Perception Research in Medical Imaging. In: *Radiat Med* 18 (2000), Nr. 6, S. 329–334
- [LAW et al. 2004] LAW, Benjamin ; ATKINS, M. S. ; KIRKPATRICK, Arthur E. ; LOMAX, Alan J.: Eye Gaze Patterns Differentiate Novice and Experts in a Virtual Laparoscopic Surgery Training Environment. In: DUCHOWSKI, Andrew T. (Hrsg.) ; VERTEGAAL, Roel (Hrsg.): *ETRA '04: Proceedings of the 2004 symposium on Eye tracking research & applications*, 2004, S. 41–48
- [LU et al. 2006] LU, Aidong ; MACIEJEWSKI, Ross ; EBERT, David S.: Volume Composition Using Eye Tracking Data. In: SANTOS, Beatriz S. (Hrsg.) ; ERTL, Thomas (Hrsg.) ; JOY, Kenneth I. (Hrsg.): *EuroVis06: Proceedings of Joint Eurographics - IEEE VGTC Symposium on Visualization*, 2006, S. 115–122

- [MARR 1982] MARR, David: *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York, NY, USA : W. H. Freeman, 1982
- [MARTINEZ-CONDE et al. 2004] MARTINEZ-CONDE, Susana ; MACKNIK, Stephen L. ; HUBBEL, David H.: The Role of Fixational Eye Movements in Visual Perception. In: *Nature Reviews Neuroscience* 5 (2004), Nr. 3, S. 229–240
- [@MATLAB] THE MATHWORKS, INC.: *MATLAB - The Language of Technical Computing*. <http://www.mathworks.com/products/matlab/>. – Link aufgerufen am 27.11.2008
- [@MEVISLAB] MEVIS RESEARCH GMBH: *MeVisLab - Medical Image Processing and Visualization*. <http://www.mevislab.de/>. – Link aufgerufen am 13.09.2008
- [MORAD et al. 2000] MORAD, Yair ; LEMBERG, Hadas ; YOFE, Nehemiah ; DAGAN, Yaron: Pupillography as an Objective Indicator of Fatigue. In: *Current Eye Research* 21 (2000), Nr. 1, S. 535–542
- [MULLIGAN 1997] MULLIGAN, Jeffrey B.: Image Processing for Improved Eye Tracking Accuracy. In: *Behavior Research Methods, Instruments, & Computers* 29 (1997), Nr. 1, S. 54–65
- [MUNN et al. 2008] MUNN, Susan M. ; STEFANO, Leanne ; PELZ, Jeff B.: Fixation-Identification in Dynamic Scenes: Comparing an Automated Algorithm to Manual Coding. In: *APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization SIGGRAPH (Veranst.)*, 2008, S. 33–42
- [NEUGEBAUER 2006] NEUGEBAUER, Mathias: *Entwicklung eines Verfahrens zur parametrisierbaren Kamerapositionierung in medizinischen Visualisierungen*, Otto-von-Guericke Universität Magdeburg, Diplomarbeit, 2006
- [NIELSEN 1993] NIELSEN, Jakob: *Usability Engineering*. San Diego , CA , USA : Academic Press, Inc., 1993
- [NODINE und KUNDEL 1987] NODINE, Calvin F. ; KUNDEL, Harold L.: Using Eye Movements to Study Visual Search and to Improve Tumor Detection. In: *Radiographics* 7 (1987), Nr. 6, S. 1241–50
- [OLEG 2008] OLEG, Špakov: *iComponent Device-Independent Platform for Analyzing Eye Movement Data and Developing Eye-Based Applications*, University of Tampere, Faculty of Information Sciences, Department of Computer Sciences, Dissertation, 2008
- [PESHORI et al. 2001] PESHORI, Kavita R. ; SCHICATANO, Edward J. ; GOPALASWAMY, Ramesh ; SAHAY, Eva ; EVINGER, Craig: Aging of the Trigeminal Blink System. In: *Experimental Brain Research* 136 (2001), S. 351–363
- [POMPLUN und SUNKARA 2003] POMPLUN, Marc ; SUNKARA, Sindhura: Pupil Dilation as an Indicator of Cognitive Workload in Human-Computer Interaction. In: STEPHANDIS, Constantine (Hrsg.): *HCI International 2003: Proceedings of the 10th International Conference on Human-Computer Interaction*, 2003
- [POMPLUN et al. 1994] POMPLUN, Marc ; VELICHKOVSKY, Boris ; RITTER, Helge: An Artificial Neural Network for High Precision Eye Movement Tracking. In: NEBEL, Bernhard (Hrsg.) ; DRESCHLER-FISCHER, Leonie S. (Hrsg.): *KI '94: Proceedings of the 18th Annual German Conference on Artificial Intelligence*, 1994, S. 63–69

- [POOLE et al. 2004] POOLE, Alex ; BALL, Linden J. ; PHILLIPS, Peter: In Search of Saliency: A Response-Time and Eye-Movement Analysis of Bookmark Recognition. In: FINCHER, Sally (Hrsg.) ; MARKOPOULOS, Panos (Hrsg.) ; MOORE, David (Hrsg.) ; RUDDLE, Roy A. (Hrsg.): *HCI 2004: Proceedings of the 18th British HCI Group Annual Conference: People and Computers XVIII: Design for Life*, 2004
- [POSNER et al. 1980] POSNER, Michael I. ; SNYDER, Charles R. ; DAVIDSON, Brian J.: Attention and the Detection of Signals. In: *Journal of Experimental Psychology* 109 (1980), Nr. 2, S. 160–174
- [RAYNER 1998] RAYNER, Keith: Eye Movements in Reading and Information Processing: 20 Years of Research. In: *Psychological Bulletin* 124 (1998), Nr. 3, S. 372–422
- [REDDY 2001] REDDY, Martin: Perceptually Optimized 3D Graphics. In: *IEEE Computer Graphics and Applications* 21 (2001), Nr. 5, S. 68–75
- [RENNINGER et al. 2007] RENNINGER, Laura W. ; VERGHESE, Preeti ; COUGHLAN, James: Where to Look Next? Eye Movements Reduce Local Uncertainty. In: *Journal of Vision* 7 (2007), Nr. 3, S. 1–17
- [RICHTER 2008] RICHTER, Stefan: *Bereitstellung eines Eyetracking-Systems für Lehre und Forschung*, Hochschule für Angewandte Wissenschaften Hamburg, Diplomarbeit, 2008
- [RUSSO und LECLERC 1994] RUSSO, J. E. ; LECLERC, France: An Eye-Fixation Analysis of Choice Processes for Consumer Nondurables. In: *Journal of Consumer Research* 21 (1994), S. 274–290
- [SALVUCCI 2000] SALVUCCI, Dario D.: An interactive model-based environment for eye-movement protocol analysis and visualization. In: DUCHOWSKI, Andrew T. (Hrsg.): *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, 2000, S. 57–63
- [SALVUCCI und ANDERSON 2001] SALVUCCI, Dario D. ; ANDERSON, John R.: Automated Eye-Movement Protocol Analysis. In: *Human-Computer Interaction* 16 (2001), Nr. 1, S. 39–86
- [SALVUCCI und GOLDBERG 2000] SALVUCCI, Dario D. ; GOLDBERG, Joseph H.: Identifying Fixations and Saccades in Eye-Tracking Protocols. In: DUCHOWSKI, Andrew T. (Hrsg.): *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, 2000
- [SCHUMANN und MÜLLER 2000] SCHUMANN, Heidrun ; MÜLLER, Wolfgang: *Visualisierung : Grundlagen und allgemeine Methoden*. Springer, 2000
- [SENNERSTEN und LINDLEY 2008] SENNERSTEN, Charlotte ; LINDLEY, Craig: Evaluation of Real-time Eye Gaze Logging by a 3D Game Engine. In: *Proceedings of the 12th IMEKO TC1 & TC7 Joint Symposium on Man Science & Measurement* IMEKO (Veranst.), 2008
- [SENSOMOTORIC INSTRUMENTS GMBH 1999] SENSOMOTORIC INSTRUMENTS GMBH: *EyeLink System User Documentation*. 1999
- [SNEDECOR und COCHRAN 1989] SNEDECOR, George W. ; COCHRAN, William G.: *Statistical Methods*. Iowa State Press, 1989

- [SOPHIAN und CROSBY 2000] SOPHIAN, Catherine ; CROSBY, Martha E.: What eye-movements tell us about ratios and spatial proportions. In: DUCHOWSKI, Andrew T. (Hrsg.): *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, 2000, S. 115
- [STELLMACH 2007] STELLMACH, Sophie: *A psychophysiological logging system for a digital game modification*, Otto-von-Guericke Universität Magdeburg, Diplomarbeit, 2007
- [STORK 2000] STORK, André: An Algorithm for Fast 3D Picking and Snapping using a 3D Input Device and 3D Cursor. In: *CAD Tools and Algorithms for Product Design [Dagstuhl Seminar, November 1998]*, Springer, 2000, S. 113–127
- [STROTHOTTE und SCHLECHTWEG 2002] STROTHOTTE, Thomas ; SCHLECHTWEG, Stefan: *Non-Photorealistic Computer Graphics: Modeling, Rendering, and Animation*. Morgan Kaufmann Publishers Inc., 2002
- [SVENSSON et al. 1997] SVENSSON, E. ; ANGELBORG-THANDERZ, M. ; SJOBERG, L. ; OLSSON, S.: Information Complexity-Mental Workload and Performance in Combat Aircraft. In: *Ergonomics* 40 (1997), Nr. 3, S. 362–380
- [TATLER 2007] TATLER, Benjamin W.: The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. In: *Journal of Vision* 7 (2007), Nr. 14, S. 1–17
- [@TOBII EYE TRACKING & EYE CONTROL] TOBII EYE TRACKING & EYE CONTROL: *Tobii Eye-Tracking Products*. http://www.tobii.com/corporate/products/product_overview.aspx. – Link aufgerufen am 25.08.2008
- [TOSI et al. 1997] TOSI, V. ; MECACCI, L. ; PASQUALI, E.: Scanning Eye Movements made when Viewing Film: Preliminary Observations. In: *International Journal of Neuroscience* 92 (1997), Nr. 1-2, S. 47–52
- [TREISMAN und GELADE 1980] TREISMAN, Anne M. ; GELADE, Garry: A feature-integration theory of attention. In: *ACM Computing Surveys Cognitive Psychology* 12 (1980), S. 97–136
- [TRIESCH et al. 2002] TRIESCH, Jochen ; SULLIVAN, Brian T. ; HAYHOE, Mary M. ; BALLARD, Dana H.: Saccade Contingent Updating in Virtual Reality. In: DUCHOWSKI, Andrew T. (Hrsg.) ; VERTEGAAL, Roel (Hrsg.) ; SENDERS, John W. (Hrsg.): *ETRA '02: Proceedings of the 2002 symposium on Eye tracking research & applications*, 2002, S. 95–102
- [WANGER et al. 1992] WANGER, Leonard R. ; FERWERDA, James A. ; GREENBERG, Donald P.: Perceiving spatial relationships in computer-generated images. In: *IEEE Computer Graphics and Applications* 12 (1992), Nr. 3, S. 44–58
- [WERNECKE 1993] WERNECKE, Josie: *The Inventor Mentor: Programming Object-Oriented 3d Graphics with Open Inventor, Release 3*. Boston, MA, USA : Addison-Wesley Longman Publishing Co., Inc., 1993
- [WERNECKE 1994] WERNECKE, Josie: *The Inventor Toolmaker: Extending Open Inventor, Release 2*. Boston, MA, USA : Addison-Wesley Longman Publishing Co., Inc., 1994
- [YARBUS 1967] YARBUS, Alfred L.: *Eye Movements and Vision*. Plenum Press, 1967

-
- [YOUNG und SHEENA 1975] YOUNG, Laurence R. ; SHEENA, David: Survey of Eye Movement Recording Methods. In: *Behavior Research Methods and Instrumentation* 7 (1975), Nr. 5, S. 397–429

Anhang A

Bedeutung der Gestaltprinzipien

Prinzip	Aussage
Der guten Gestalt	Jedes Reizmuster wird so gesehen, dass die resultierende Struktur so einfach wie möglich ist.
Der Ähnlichkeit	Ähnliche Dinge scheinen zu Gruppen geordnet.
Der Fortsetzung	Punkte, die als gerade oder sanft geschwungene Linien gesehen werden, wenn man sie verbindet, werden als zusammengehörig wahrgenommen.
Der Nähe	Dinge, die sich nahe beieinander befinden, erscheinen als zusammengehörig.
Des gemeinsamen Schicksals	Dinge, die sich in die gleiche Richtung bewegen, erscheinen als zusammengehörig.
Der Vertrautheit	Dinge bilden mit größerer Wahrscheinlichkeit Gruppen, wenn die Gruppen vertraut erscheinen oder etwas bedeuten.
Der gemeinsamen Region	Elemente, die innerhalb einer gemeinsamen Region liegen, werden zusammengruppiert.
Der Verbundenheit	Verbundene Elemente innerhalb einer Region mit gemeinsamen visuellen Charakteristiken wie Helligkeit, Farbe, Textur oder Bewegung werden als Einheit gesehen.

Tabelle A.1: Darstellung der Gestaltprinzipien und ihrer Bedeutung nach GOLDSTEIN [2008].

Anhang B

Fragebogenaufbau und -auswertung

Abbildung B.1 zeigt den Fragebogen, der zur Akquisition personen- und experimentbezogener Daten von den Versuchspersonen (VP) verwendet wurde. Er ist angelehnt an viele experimentbezogene Hinweise aus SENSOMOTORIC INSTRUMENTS GMBH [1999] und dient der Verbesserung zukünftiger Experimente mit EYELINK I. Dazu sind vor allem die subjektiven Kommentare hilfreich, welche in den blau hinterlegten Bereichen abgegeben werden konnten. Der Fragebogen ist in die folgenden zwei Fragenkategorien mit ihren zwei Unterkategorien geteilt:

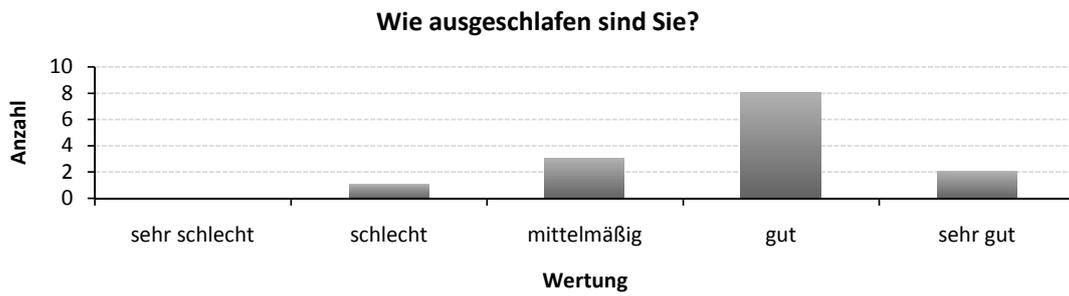
1. Personenbezogene Fragen
 - a) Beantwortung mittels Skala
 - b) Beantwortung ohne Skala
2. Experimentbezogene Fragen
 - a) Beantwortung mittels Skala
 - b) Beantwortung ohne Skala

Durch die Skala konnten die Fragen mit den Werten 1 (sehr schlecht) bis 5 (sehr gut) beantwortet werden. Dazu mussten die Probanden in die gelblichen Felder der Subkategorien 1a) und 2a) ein „x“ an der entsprechenden Skalenposition eintragen. Für die Kategorien 1b) und 2b) musste das „x“ an die jeweils dafür vorgesehenen Stellen für „ja“, „nein“, „männlich“, „weiblich“, etc. eingetragen werden. In die grünen Felder sollten die teilnehmenden Personen Zahlen eintragen. Die in Abbildung B.1 eingetragenen Werte entsprechen jeweils der Anzahl der VP's, welche ein Feld angekreuzt haben (gelbe Felder) bzw. dem Mittelwert der von allen VP's angegebenen Daten (grüne Felder). Die Skala ist am unteren Ende des Fragebogens erläutert.

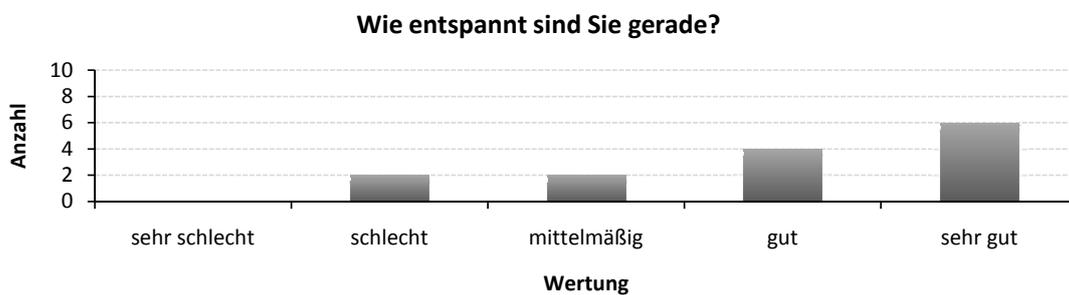
In Abbildung B.2 und Abbildung B.3 ist die Auswertung der Antworten auf die personenbezogenen bzw. experimentbezogenen Fragen dargestellt. Diese bezieht sich nur auf die Fragen, zu denen eine Bewertung, auf Basis der Skala, abgegeben werden konnte. Jedes Diagramm wertet eine Frage für alle VP's aus und gibt die Anzahl der Personen wieder, die sich für eine bestimmte Wertung entschieden haben. Insgesamt haben 14 Probanden teilgenommen.

Fragebogen "Eye-Tracking"		Bewertung					Datum	
		1	2	3	4	5	Uhrzeit	
Persönliche Fragen							14:30	
Wieviele Tassen koffeinhaltige Getränke haben Sie heute getrunken?		0,5						
Wie ausgeschlafen sind Sie?		0	1	3	8	2		
Wie entspannt sind Sie gerade?		0	2	2	4	6		
Wie kennen Sie sich mit Computern im Allgemeinen aus?		0	1	0	7	6		
Wie gut sind Ihre Eye-Tracking-Kenntnisse?		6	5	2	1	0		
Alter			24					
Geschlecht		m	7	w	7			
Nehmen Sie beruhigende Medikamente oder Augentropfen?		n	14	j	0			
Haben Sie eine Augenkrankheit?		n	8	j	6		Wenn ja, welche?	
Tragen Sie eine Sehhilfe?					Brille	2		
					Kontaktlinsen	3		
					Keine	9		
Fragen zu den Experimenten		1	2	3	4	5		
Wie bequem saßen Sie auf dem Stuhl?		0	0	1	9	4		
Wie bequem saß das Eye-Tracking-Gerät auf Ihrem Kopf?		1	2	4	6	1		
Wie gut konnten Sie, trotz des Eye-Trackers, den Bildschirm sehen?		0	0	1	4	9		
Wie gut entsprach der Bildschirmabstand Ihren Gewohnheiten?		0	0	0	8	6	Wie wäre es besser?	
Wie gut konnten Sie sich während des Experimentes konzentrieren?		1	0	4	8	1	Wodurch abgelenkt?	
Wie entspannt waren Sie während des Experimentes?		0	0	2	9	3	Warum nervös?	
Welches Auge wurde bei Ihnen aufgenommen?		l	11	r	3		Warum dieses?	
War das Experiment bei Ihnen erfolgreich?		n	0	j	14		Warum nicht?	
Allgemeine Bemerkungen / Hinweise zu den Experimenten:								
Skala:		1 (sehr schlecht) • 2 (schlecht) • 3 (mittel) • 4 (gut) • 5 (sehr gut)						

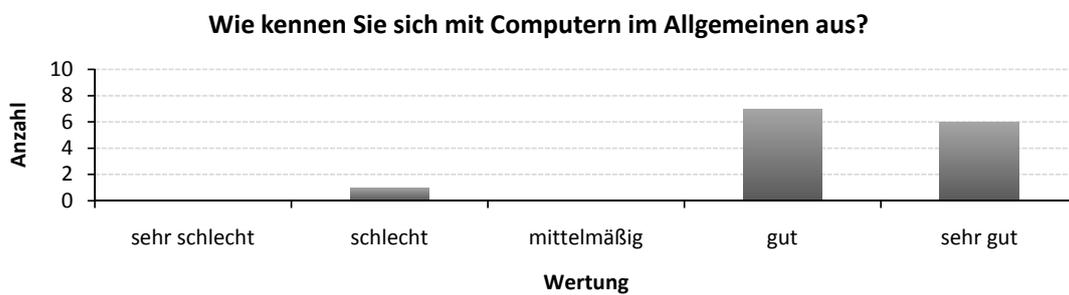
Abbildung B.1: Fragebogen zur Akquisition personen- und experimentbezogener Daten von den Versuchspersonen.



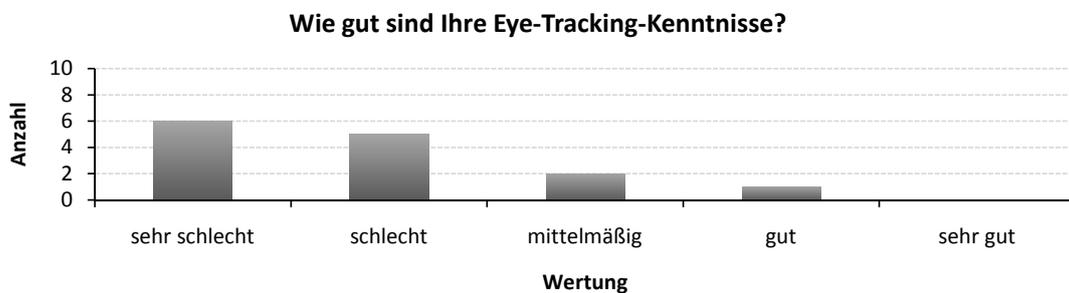
(a)



(b)

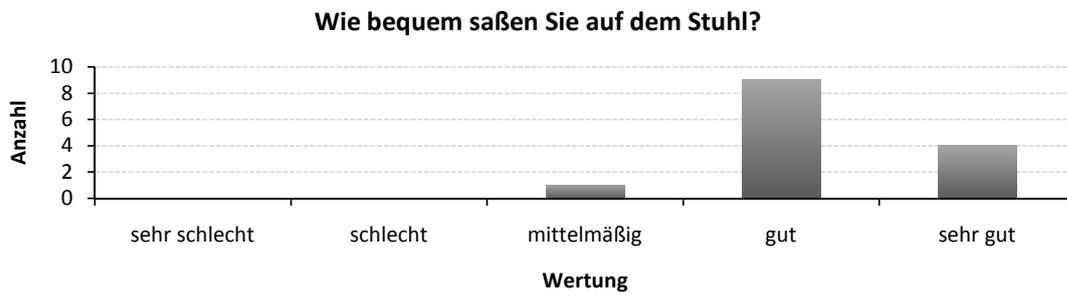


(c)

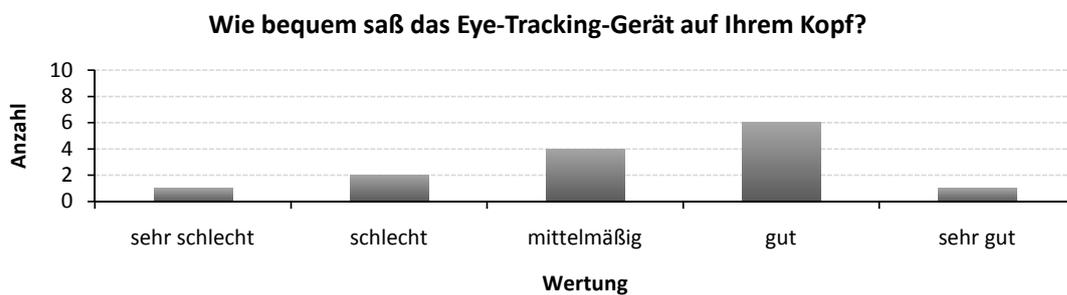


(d)

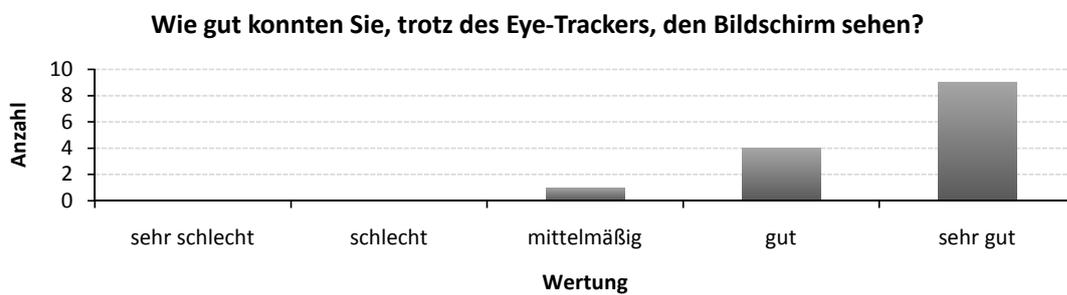
Abbildung B.2: Graphische Darstellung der erfassten Daten zum Abschnitt „Persönliche Fragen“ des Fragebogens.



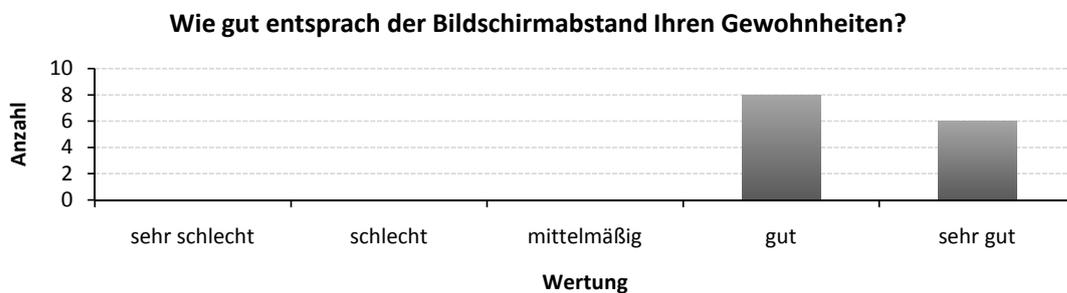
(a)



(b)



(c)

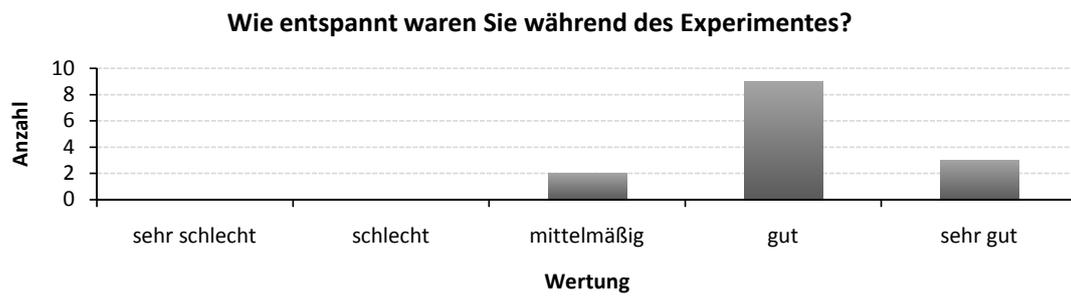


(d)

Abbildung B.3: Graphische Darstellung der erfassten Daten zum Abschnitt „Experimentbezogene Fragen“ des Fragebogens.



(e)



(f)

Abbildung B.3 (Fortsetzung): Graphische Darstellung der erfassten Daten zum Abschnitt „Experimentbezogene Fragen“ des Fragebogens.

Anhang C

Experimentauswertung

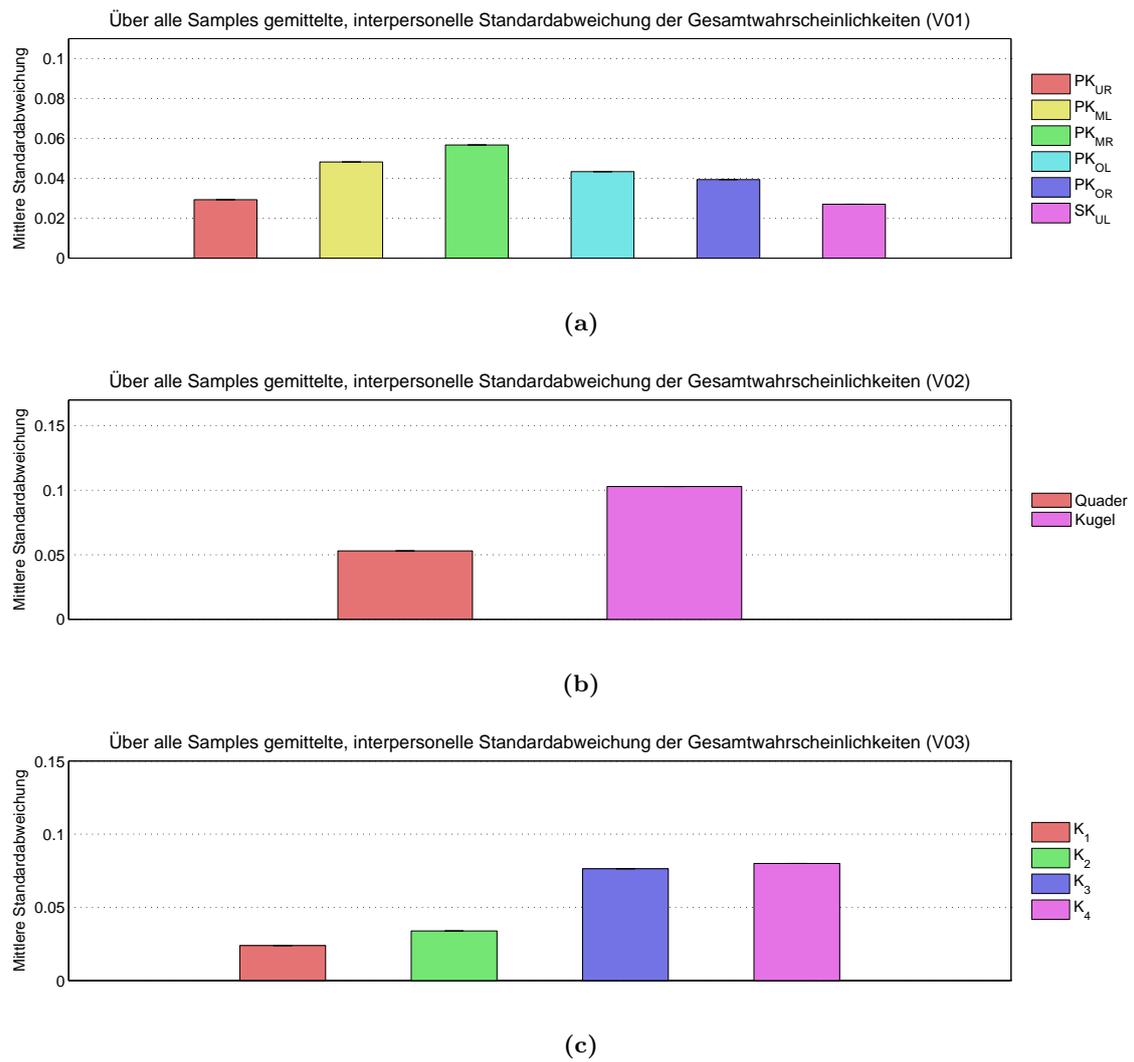


Abbildung C.1: Über alle Samples gemittelte, interpersonelle Standardabweichung der Gesamtwahrscheinlichkeiten. Abbildung (a) stellt die Standardabweichung für Versuch eins, Abbildung (b) für Versuch zwei und Abbildung (c) für Versuch drei dar.

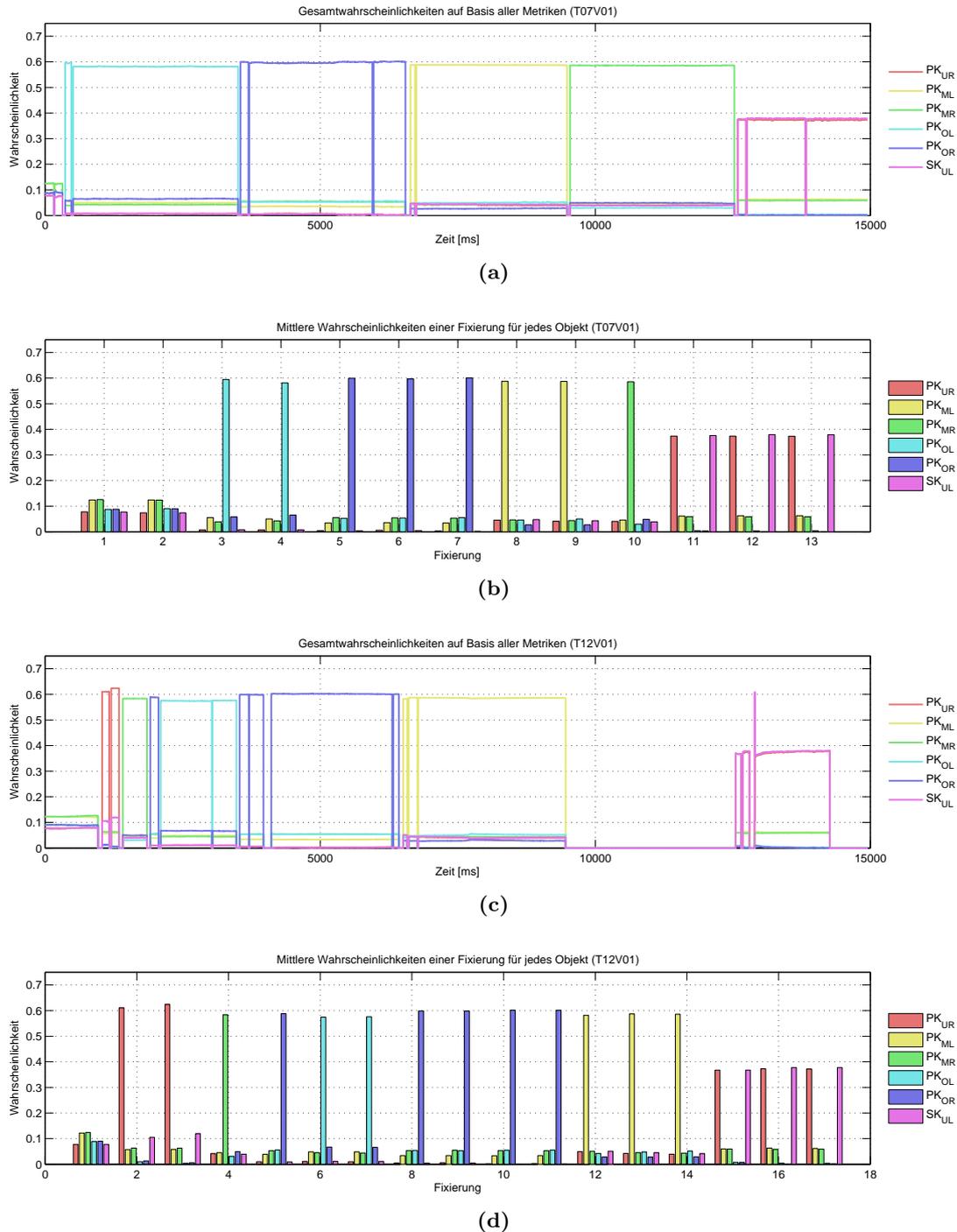


Abbildung C.2: Vergleich eines guten ((a), (b)) und eines schlechten ((c),(d)) Ergebnisses bezüglich der Datenqualität aus dem ersten Versuch. Abbildung (a) und (b) zeigen die ermittelten Fixierungswahrscheinlichkeiten von Proband sieben und die Abbildungen (c) und (d) entsprechen den Wahrscheinlichkeiten von Proband 12. Selbst im qualitativ guten Beispiel lassen sich Fixierungen von dicht beieinander liegenden Objekten nur schwer unterscheiden. Die Wahrscheinlichkeiten von SK_{UL} und PK_{UR} in Abbildung (b) sind sich ähnlich. Jedoch ist in Abbildung (b) die Wahrscheinlichkeit für SK_{UL} , wie erwartet, für die letzten Fixierungen am größten.

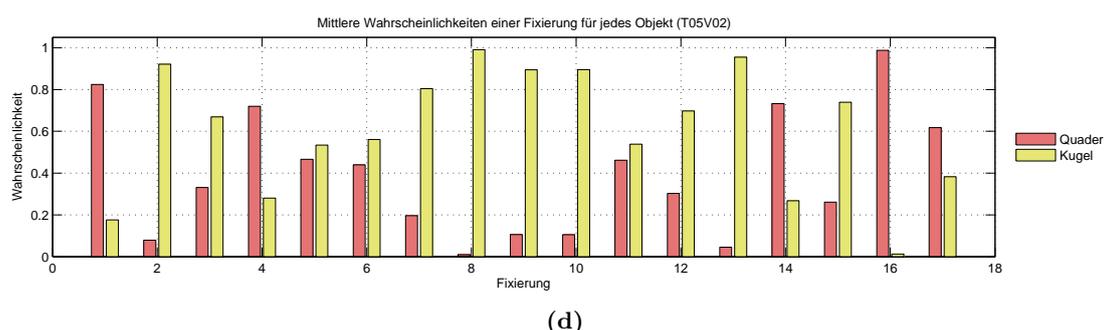
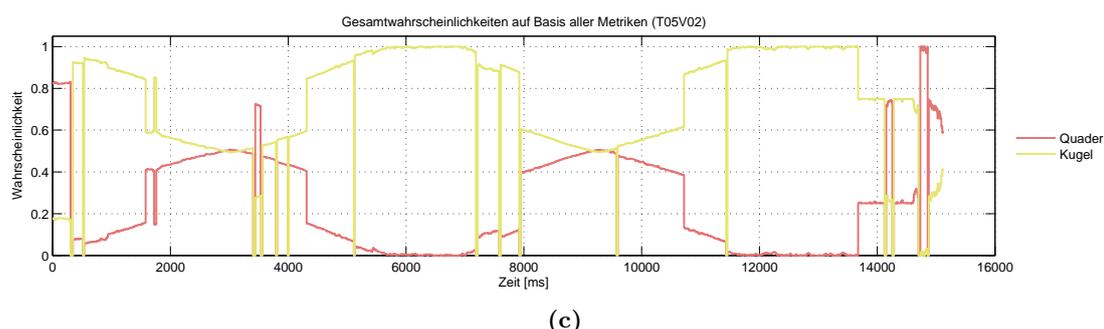
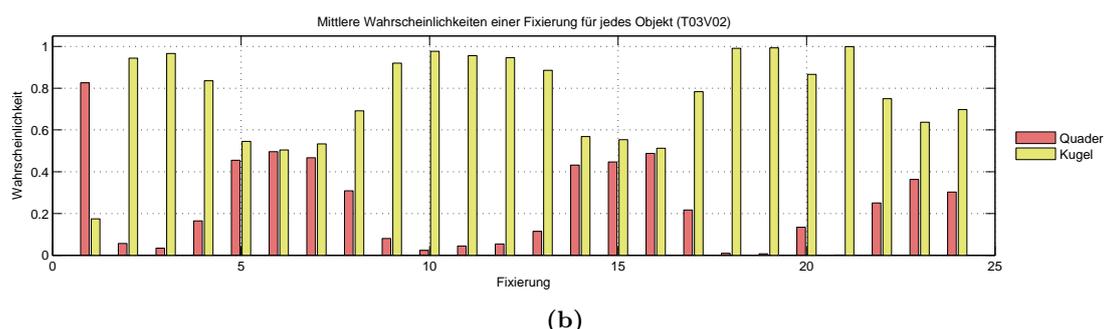
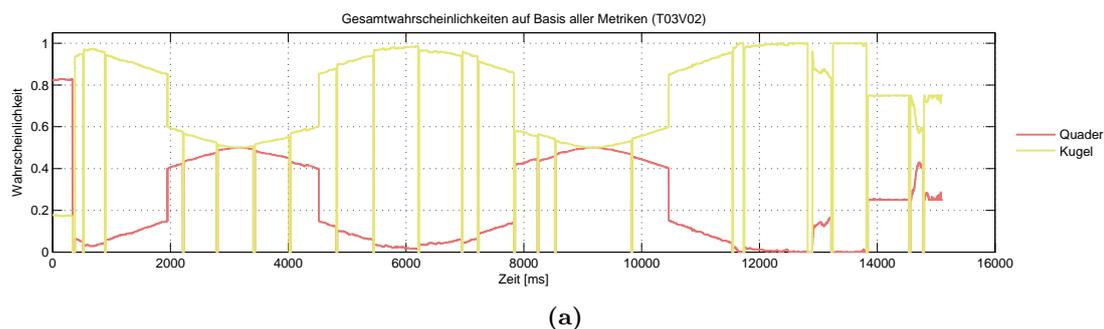


Abbildung C.3: Vergleich eines guten ((a) und (b)) und eines schlechten ((c) und (d)) Ergebnisses bezüglich der Datenqualität aus dem zweiten Versuch. Abbildung (a) und (b) zeigen die ermittelten Fixierungswahrscheinlichkeiten für Proband drei und die Abbildungen (c) und (d) entsprechen den Wahrscheinlichkeiten von Proband fünf. Im guten Beispiel hat der Algorithmus, wie erwartet, für jede Fixierung die Kugel als das am wahrscheinlichsten fixierte Objekt ermittelt. Dies ist auch für die Zeitpunkte der der Fall, wenn sich Kugel und Quader gegenseitig verdecken. Es ist jedoch zu erkennen, dass sich beide Wahrscheinlichkeiten für diesen Fall annähern.

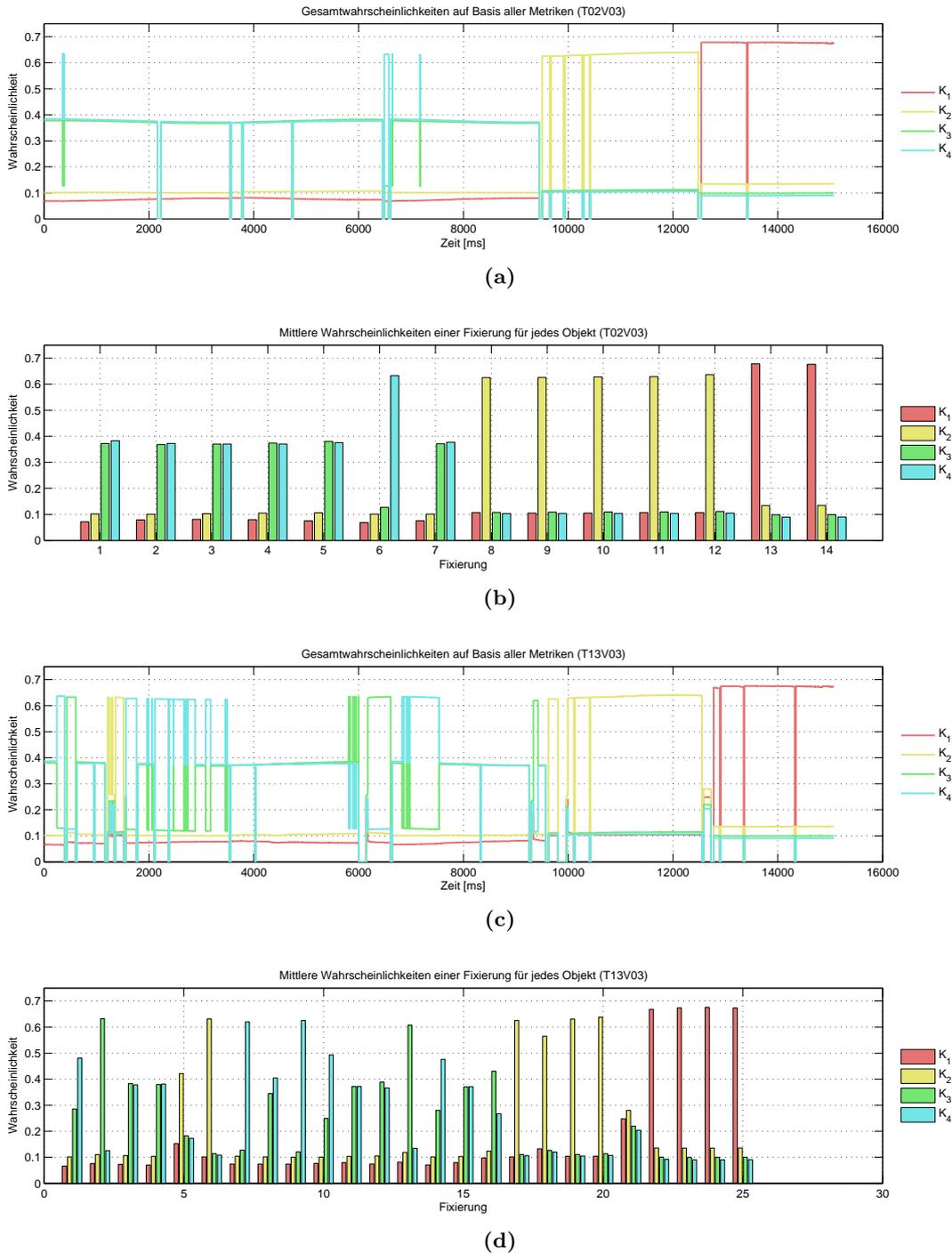


Abbildung C.4: Vergleich eines guten ((a) und (b)) und eines schlechten ((c) und (d)) Ergebnisses bezüglich der Datenqualität aus dem dritten Versuch. Abbildung (a) und (b) zeigen die ermittelten Fixierungswahrscheinlichkeiten für Proband eins und die Abbildungen (c) und (d) entsprechen den Wahrscheinlichkeiten von Proband 13. Im schlechten Beispiel wurden weitaus mehr Fixierungen ausgeführt. Die Konzentration lag entgegen der Annahmen nicht auf den Objekten die rot eingefärbt wurden, was sich negativ auf das Endergebnis auswirkt.

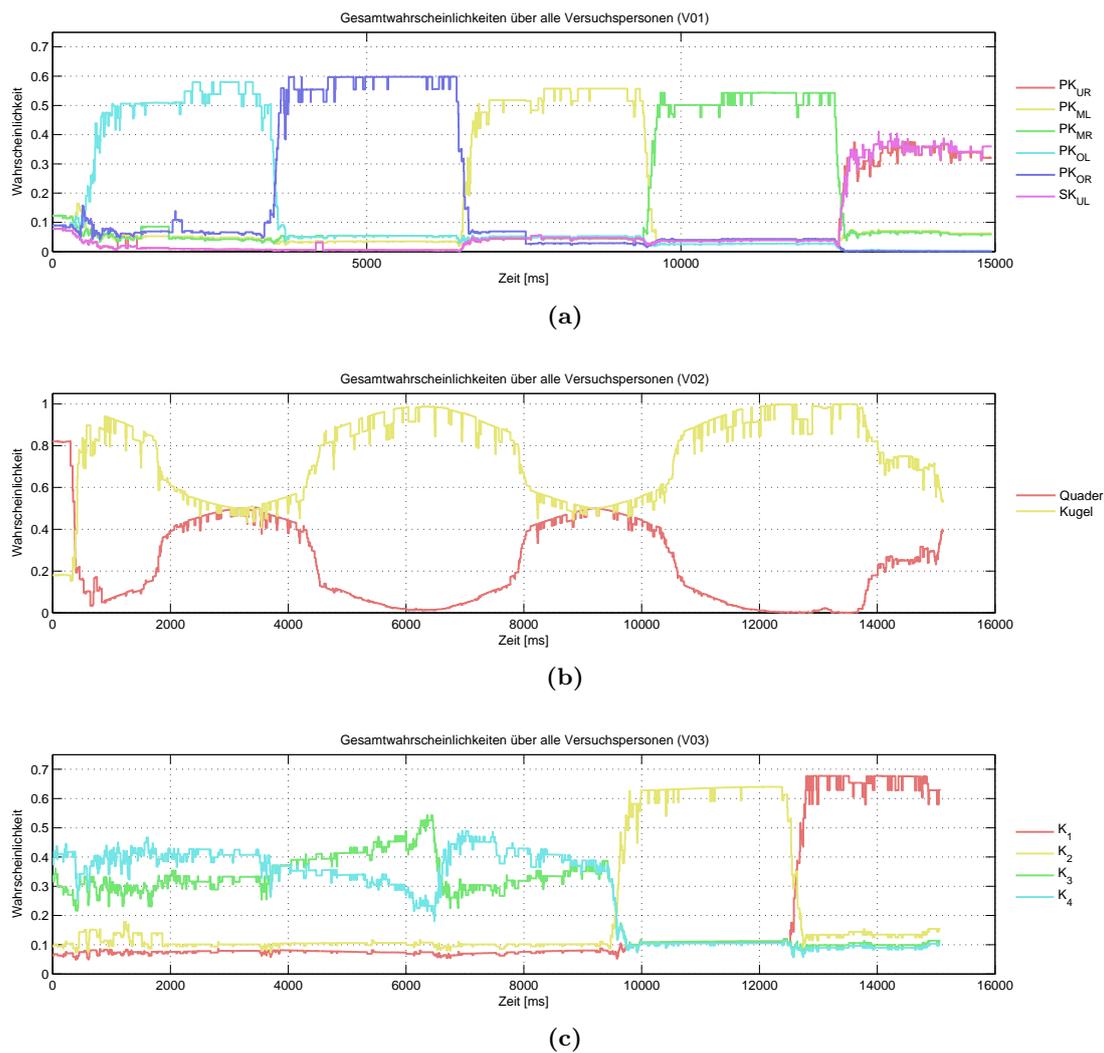


Abbildung C.5: Mittelwert der Gesamtwahrscheinlichkeiten pro Objekt über alle Probanden. Abbildung (a) stellt die Mittelwerte für Versuch eins, Abbildung (b) für Versuch zwei und Abbildung (c) für Versuch drei dar.

