

TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Mensch-Maschine-Kommunikation

Blickbasierte Mensch-Maschine Interaktion im Automobil

Tony Matthias Poitschke

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr.-Ing. Ulf Schlichtmann
Prüfer der Dissertation: 1. Univ.-Prof. Dr.-Ing. habil. Gerhard Rigoll
2. Univ.-Prof. Dr. rer. nat. habil. Thomas Huckle

Die Dissertation wurde am 19.04.2011 bei der Technischen Universität München eingereicht
und durch die Fakultät für Elektrotechnik und Informationstechnik am 22.09.2011 angenom-
men.

Kurzfassung

Ziel der vorliegenden Arbeit ist die Konzeption neuartiger, intuitiver Benutzerschnittstellen zur Reduktion der Fahrerablenkung im Kraftfahrzeug. Dazu stellt die Arbeit unterschiedliche Verfahren zur Erkennung der Fahrerablenkung vor: (a) mittels einer Überwachung der Interaktion, (b) mittels einer Analyse von Fahrzeugparametern sowie (c) der Erfassung der Blickrichtung des Fahrers. Aufbauend auf einer umfassenden Analyse dieser unterschiedlichen Parameter stellt die Arbeit verschiedene Konzepte zur Anpassung des Informationsflusses an den Verlauf der visuellen Aufmerksamkeit des Fahrers vor und gelangt im Rahmen verschiedener Nutzerstudien zu einer eingehenden Bewertung der erarbeiteten Anzeige-Bedienkonzepte. Abschließend gibt die Arbeit einen Ausblick auf das Nutzungspotential der vorgestellten Ansätze zur Erfassung der Blickrichtung im Kontext anderer Anwendungsdomänen.

Inhaltsverzeichnis

Kurzfassung	3
Abbildungsverzeichnis	viii
Tabellenverzeichnis	ix
1 Einleitung und Motivation	1
2 Grundlagen und Thematische Einführung	5
2.1 Mensch-Maschine Interaktion – Begriffsdefinition	5
2.2 Verhaltensmodelle in der Mensch-Maschine Interaktion	6
2.2.1 Wickens Multiple Resource Theory	7
2.2.2 Strategic Task Management	8
2.2.3 SRK-Modell nach Rasmussen	8
2.3 Domänenspezifische Grundlagen	9
2.3.1 Fahraufgabe	10
2.3.2 Einfluss der Fahraufgabe auf die mentale Belastung	12
2.3.3 Fahrerarbeitsplatz	12
2.4 Usability-Untersuchung von Fahrer-Fahrzeug Schnittstellen	15
2.4.1 Verwendete Bewertungsmethoden	16
2.5 Blickbasierte Benutzerschnittstellen	21
2.5.1 Blickbedienung	21
2.5.2 Attentive User Interfaces	22
2.5.3 Gestaltungshinweise für blickbasierte Benutzerschnittstellen	23
3 Kamera-basierte Fahrerüberwachung	27
3.1 Grundbegriffe und geschichtliche Entwicklung	28
3.1.1 Anwendungsfelder	28
3.1.2 Physiologie des menschlichen Auges	28
3.1.3 Augenbewegungen	31
3.1.4 Geschichtliche Entwicklung des Eye Tracking	32

3.1.5	Videookulografie	36
3.1.6	Fazit	40
3.2	Referenzsystem EyeSeeCam	40
3.2.1	Head-mounted Eye Tracking System <i>EyeSeeCam</i>	40
3.2.2	Transformation der Blickdaten in ein Weltkoordinatensystem	41
3.2.3	Display Manager Jeannie	43
3.2.4	Evaluierung des Gesamtsystems	47
3.3	Remote Eye Tracking System <i>eyeDrive</i>	48
3.3.1	Rekonstruktion der Blickgeraden in 3D	50
3.3.2	Extraktion der relevanten Parameter zur Blickrichtungsbestimmung	56
3.3.3	Hardwareaufbau	66
3.3.4	Darstellung der Blickdaten als HeatMap	70
3.3.5	Systemevaluierung	72
4	Sensor-basierte Fahrerüberwachung	77
4.1	Grundlagen	77
4.1.1	Longitudinale Führungsparameter	78
4.1.2	Laterale Führungsparameter	79
4.2	Fuzzy Logik	80
4.2.1	Fuzzy Set Theorie	80
4.2.2	Fuzzy-Inferenzsysteme	82
4.3	Das Assistenzsystem AWARE	85
4.3.1	Modellierung des Fahrverhaltens	87
4.3.2	Abnormalitätsbestimmung	89
4.4	Systembewertung	92
5	Attentive User Interfaces	95
5.1	<i>ABK2020</i> – ein Basisframework für Anzeige-Bedienkonzepte	96
5.2	Örtliches Scheduling von Dialogen	97
5.2.1	Örtliches Scheduling von fahrzeuginitiierten Dialogen	98
5.2.2	Iterative Anpassung der Tiefe von Dialogen	99
5.2.3	Versuchsdurchführung und -ergebnisse	101
5.3	Aufmerksamkeitsbeeinflussung mit Warnung und Locking	104
5.3.1	Versuchsergebnisse	106
5.3.2	Fazit	111
5.4	Zeitliches Scheduling von Dialogen	111
5.4.1	Versuchsergebnisse	113
5.4.2	Fazit	116
5.5	Fahreraufmerksamkeitslenkung durch graphische und akustische Anima- tionen	116
5.5.1	Theoretische Grundlagen	116
5.5.2	Konzept und Implementierung	120
5.5.3	Versuchsdurchführung	122
5.5.4	Versuchsergebnisse	123
5.5.5	Fazit	126

5.6	Intentionsbasierte Parametrierung der Mensch-Maschine Schnittstelle . . .	127
5.6.1	Theoretische Grundlagen	128
5.6.2	Intentionsbasierte Planerkennung	130
5.6.3	Versuchsergebnisse: Methodenbewertung	136
5.6.4	Versuchsergebnisse: Intentionsbasierte Parametrierung des HMI . . .	141
6	Blickbasierte Interaktion im Kraftfahrzeug	145
6.1	Konzept	145
6.1.1	Hauptmenü	147
6.1.2	Applikationen	148
6.1.3	Blickbasierte Alphanumerikeingabe	150
6.2	Versuchsergebnisse	152
6.2.1	Versuchsergebnisse – Alphanumerik	152
6.2.2	Versuchsergebnisse – Menübedienung	155
7	Integration in weitere Anwendungsgebiete	161
7.1	Blickbasierte Interaktion in Fabrikumgebungen	161
7.1.1	Evaluierungsergebnisse	163
7.1.2	Fazit	164
7.2	Blickbasierte Interaktion in Desktop Szenarien	164
7.2.1	GazeSpeed – Blickunterstützte Maus	165
7.2.2	GazeAnywhere – Blickgesteuerte Shortcuts	168
7.2.3	GazeRocket – Blickgesteuerter Programmstart	170
7.3	Multi-User Gaze Tracking für Robotiksznarien	172
7.3.1	Verwendete Kameraorientierungssysteme und Hardwareaufbau . . .	173
7.3.2	Gesichtsdetektion	175
7.3.3	Ausrichtung der Blickkameras	175
7.3.4	Systemevaluierung	178
8	Zusammenfassung und Ausblick	181
	Literaturverzeichnis	185
A	Anhang	201
A.1	Ausführliche Versuchsergebnisse zur Intentionserkennung	201

Abbildungsverzeichnis

2.1	Passungsmodell der Mensch-Maschine Interaktion.	6
2.2	Ressourcenpools von Wickens Multiple Resource Theory.	7
2.3	SRK-Modell nach Rasmussen.	9
2.4	Platzierung der Bedienelemente im Fahrzeugcockpit.	10
2.5	Die drei Ebenen der primären Fahraufgabe.	11
2.6	Anzeige-Bedienkonzepte – Trennung von Anzeige und Bedienung.	14
2.7	Anzeige-Bedienkonzepte – Kombination von Anzeige und Bedienung.	15
2.8	Beispiele unterschiedlicher Sichtsimulationen.	16
2.9	Der MMK Fahr Simulator.	17
2.10	Schematischer Versuchsaufbau Aufbau beim PDT.	18
2.11	Skala zur Erfassung subjektiv erlebter Anstrengung.	19
2.12	Organisationsstrukturen graphischer Benutzerschnittstellen.	26
3.1	Das menschliche Auges.	29
3.2	Schematische Darstellung des menschlichen Auges.	29
3.3	Schärfebereiche des menschlichen Auges.	30
3.4	Mathematisches Augenmodell.	31
3.5	Prinzip der Hebelnystagmographie.	33
3.6	Schematische Darstellung der Purkinje-Reflexionen.	34
3.7	Elektronystagmographie.	35
3.8	Limbus Tracking.	35
3.9	Apparaturen des Search Coil Verfahrens.	35
3.10	Videookulographie.	36
3.11	Head Supported Eyetracker SR Research EyeLink 1000/2000.	38
3.12	Remote Eye Tracking Systeme des Herstellers Tobii.	39
3.13	Head-mounted Eye Tracker EyeSeeCam.	40
3.14	Head-mounted Eye Tracker mit Erweiterungen zum Head Tracking.	41
3.15	Transformationskette.	42
3.16	Veruschsaufbau mit Eye Tracker basierend auf der EyeSeeCam.	44
3.17	Aufgezeichnete Fixationsdaten am kalibrierten Monitor.	48
3.18	Kamerabilder des remote Eye Tracking Systems <i>eyeDrive</i>	49

3.19	Geometrische Überlegungen zur Bestimmung des Korneazentrums. . . .	51
3.20	Geometrische Überlegungen zur Abschätzung des Pupillenmittelpunktes. . . .	52
3.21	Benutzerkalibrierung.	56
3.22	Differenz aus on- und off-axis beleuchtetem Eingangsbild.	57
3.23	Supersampling mit Faktor 8.	58
3.24	Basisbilder für den Kontrast basierten Klassifikator.	59
3.25	Inneres und äußeres maskiertes Bild.	60
3.26	Vergößerung der Bounding Boxen bei nicht gefundenen Pupillen.	61
3.27	Verhinderung von Überschneidungen der beiden Bounding Boxen.	61
3.28	Hilfsbild, Auge im Kamerabild und Ergebnis des Kantendetektors.	63
3.29	Ergebnis der Maskierung des Suchbereiches.	64
3.30	Kantenbilder mit unterschiedlicher zugrunde liegender Bildschärfe.	66
3.31	Hardwareaufbau des remote Eye Tracking Systems <i>eyeDrive</i>	67
3.32	Kamera Point Grey FireFly MV.	68
3.33	Flüssiglinse Arctic 314 und deren Funktionsweise.	69
3.34	Flüssiglinse integriert in den Objektivhalter.	70
3.35	Blickverläufe von zwei Probanden.	71
3.36	Vergleich der Iterationsgeschwindigkeiten.	73
3.37	Aufgezeichnete Rohdaten des Gaze Trackers.	75
3.38	Versuchsaufbau auf der Sitzkiste.	76
4.1	Scharfe Menge (links) und Fuzzymenge (rechts).	81
4.2	Struktur eines Fuzzy-Inferenzsystem nach Mamdani.	84
4.3	Modell der Abstandsregelung.	88
4.4	Modell des Beschleunigungsverhaltens.	89
4.5	Partitionierung der Eingangsvariablen.	90
4.6	Partitionierungsparameter.	91
4.7	Partitionierung der Ausgabevariablen.	91
4.8	Rohdaten der Abnormalitätsbewertung I.	93
4.9	Rohdaten der Abnormalitätsbewertung II.	94
5.1	Displayaufbau mit dreigeteiltem FPK, HUD und CID Touchscreen.	96
5.2	Exemplarische graphische Oberflächen für das CID.	97
5.3	GUI des FPK.	98
5.4	GUI des HUD.	99
5.5	Beispiel einer Systemmeldung im HUD.	100
5.6	Abbildung des Abnormalitätswertes auf den Helligkeitswert des CID.	105
5.7	Abbildung des Fahrabnormalitätswertes auf den variablen Warnrahmen.	106
5.8	Mittlere Spurabweichung der Probanden.	107
5.9	Objektiv gemessene Beanspruchung.	108
5.10	Ergebnisse SEA und Sekundäraufgabendaten.	109
5.11	(Soft-)Locking, allgemeine Bewertung des Systems.	110
5.12	Portfoliodarstellungen der getesteten Warnkonzepte.	111
5.13	Allgemeine Bewertung des Systems.	112
5.14	Mittlere Spurabweichung der Probanden.	114

5.15	Objektiv gemessene Beanspruchung.	115
5.16	Ergebnis SEA sowie Anteil falscher Antworten auf Systemmeldungen. . .	115
5.17	Das Gesichtsfeld des Menschen.	117
5.18	Cues sowie Gesetz der Ähnlichkeit und Gesetz der Nähe.	118
5.19	Optische Täuschung und Color Key.	119
5.20	Graphische Oberfläche des prototypischen Interaktionssystems.	120
5.21	Ortung einer Schallquelle.	121
5.22	Nachbefragungsergebnisse zum Warnton.	124
5.23	Ergebnisse AttrakDiff.	125
5.24	Nachbefragungsergebnisse zur Aufmerksamkeitslenkung.	126
5.25	Nachbefragungsergebnisse zum Warnton mit Führungspunkt.	127
5.26	Grundlegende Architektur der Intentionsbasierten Planerkennung.	131
5.27	Allgemeine Struktur eines Intentionsmodells.	133
5.28	Durchschnittswerte der Handdaten aller Versuchspersonen VT I.	138
5.29	Durchschnittswerte der Handdaten aller Versuchspersonen VT II.	139
5.30	Allgemeine Bewertung des Systems.	142
6.1	Exemplarische Darstellung der graphischen Oberflächen.	146
6.2	Graphische Oberflächen des Hauptmenüs auf dem CID.	147
6.3	Graphische Oberfläche Tempomat/ACC.	149
6.4	Graphische Oberfläche Navigationsmenü.	150
6.5	Graphische Oberflächen der blickbasierten Alphanumerikeingaben. . . .	151
6.6	Versuchsaufbau.	152
6.7	Übersicht der mittleren Eingabezeiten mit und ohne Fahraufgabe.	153
6.8	Mittlere Spurabweichung und PDT.	153
6.9	Umfrageergebnisse zu den Konzepten.	154
6.10	Portfoliodarstellung des Systems.	156
6.11	Bewertung vom Menü und Schätzung eigener Fahrleistung.	157
6.12	Bedienzeiten bei Blick- (blau) und Touchinteraktion (rot).	158
6.13	Ergebnisse der PDT.	158
6.14	Mittlere Spurabweichung.	159
7.1	Verwendeter Laboraufbau eines manuellen Montageplatzes.	162
7.2	Graphische Oberfläche und das zu fertigende Werkstück.	163
7.3	GazeSpeed, typischer und korrigierter Mausweg.	166
7.4	Übersicht der Ergebnisse des NASA TLX Fragebogens für alle Konzepte. .	167
7.5	GazeAnywhere Versuchsaufbau.	168
7.6	Screenshot der GazeRocket Benutzeroberfläche.	171
7.7	Demonstrationsplattformen.	173
7.8	Autofokus zur Analyse von Nutzern in variierenden Distanzen.	174
7.9	Kameraorientierungssysteme.	174
7.10	Gesamtaufbau <i>EyeSeeBot</i>	176
7.11	Verwendete Koordinatensysteme im <i>EyeSeeBot</i> Gesamtsystem.	177
7.12	Erweiterter Arbeitsbereich des <i>EyeSeeBot</i> Aufbaus.	179

Abbildungsverzeichnis

A.1	Versuchsteil I, Handdaten aller Versuchspersonen.	202
A.2	Versuchsteil II, Handdaten aller Versuchspersonen.	203

Tabellenverzeichnis

3.1	Struktur der Daten, die über den TCP-Socket übertragen werden.	43
3.2	Vergleich der benötigten Anzahl an Iterationen.	73
3.3	Berechnungsdauer in μs der einzelnen Bildvorverarbeitungsschritte. . . .	74
3.4	Berechnungsdauer in μs der einzelnen Bildverarbeitungsschritte.	74
4.1	Abnormalitätsbestimmung, linguistische Ausdrücke.	90
5.1	Abbildung der Abnormalität auf Helligkeits- bzw. Deckkraftwerte.	106
5.2	Beispielsätze für das Scheduling-Testverfahren	113
5.3	Kriterien zur Festlegung der verwendeten Datenbasis.	135
5.4	Gültigkeit und Erkennungsrate.	137
5.5	Auswertung der Versuchsdaten nach verwendeter Datenbasis.	137

Einleitung und Motivation

Ausgangslage und Problemstellung

Das moderne Automobil dient heutzutage nicht mehr nur der Befriedigung individueller Mobilitätsbedürfnisse, sondern entwickelt sich immer stärker zu einem multimodalen Infotainment- und Assistenzsystem, das vielschichtige Komfortbedürfnisse des Fahrers bedient. So sind bereits die meisten Modellreihen verschiedener Hersteller mit einer Vielzahl an Infotainmentfunktionen – wie z.B. Satellitenradio, CD-/DVD-Player, TV-Funktion, Navigationssystem und voll integrierte Mobiltelefone – ausgestattet. Zudem ist es heute keine Seltenheit, dass Fahrzeuge mittels drahtloser Datenverbindungen (z.B. EDGE oder UMTS) Zugang zum Internet besitzen, um dem Fahrer aktuellste Informationen (z.B. Podcasts, Newsfeeds, Hotel- und Restaurantführer, Wetter) und synchronisierte Officefunktionen (z.B. Email, Kalender) bereit zu stellen. Die stetig voranschreitende Etablierung neuartiger Kommunikationstechnologien, wie z.B. auf IEEE 802.11p basierende Car2X Kommunikation, ermöglicht eine Vielzahl weiterer Funktionen und Meldungen in zukünftigen Fahrzeuggenerationen. Aufgrund dieser stetig zunehmenden Menge verfügbarer Informationen und Funktionen sind sie in aktuellen Mittel- und Oberklassefahrzeugen meist in einer hierarchischen Menüstruktur organisiert: Über verschiedenste integrierte Anzeige-Bedienkonzepte lassen sich Funktionen, wie beispielsweise Navigation, Bordcomputer und das gesamte Entertainmentsystem bedienen.

Hinsichtlich dieser mit steigenden Komfortansprüchen einhergehenden Technisierung darf jedoch nicht außer Acht gelassen werden, dass neue Systeme im Fahrzeug zu bedienen sind und auf eintreffende Meldungen zu reagieren ist: Ein eingehender Telefonanruf will beantwortet werden, eine erhaltene Einladung zu einem Treffen muss in den Kalender geladen werden, das Navigationsgerät erhält neue Daten zur Routenführung, deren Verarbeitung Entscheidungen vom Fahrer erfordern, und vieles mehr. All diese Zusatzaufgaben stellen eine Quelle der Fahrerablenkung dar, die die Aufmerksamkeit des Fahrers hinsichtlich seiner primären Fahraufgaben – Lenken, Spurhaltung, Geschwindigkeit, etc. – beeinträchtigen. Um dies zu veranschaulichen, sei an dieser Stelle auf Unfallstatistiken verwiesen, die verdeutlichen, dass eine Vielzahl an Unfällen durch Unaufmerksamkeit und Ablenkung verursacht werden: Nach internationalen Studien sind ca. 25% der Unfälle auf Fahrerablenkung oder -unaufmerksamkeit [199] und ca. 20% der schwerwiegenden Unfälle auf Übermüdung [169, 204] zurückzuführen.

Deshalb gilt es, den Fortschritt im Bereich der Informationsverarbeitung dank wachsender Rechenleistung und neuer Technologien zu nutzen, um sowohl die Bedienung einzelner Systeme, als auch die Aufnahme und Verarbeitung verfügbarer Informationen durch den Fahrer intuitiver zu gestalten. Eine übergreifende Koordination der Systeme und Funktionen in Kombination mit intuitiven und wenig ablenkenden Bedienkonzepten sind essentielle Schritte, um die auf den Fahrer einwirkende Informationsflut zu bändigen. So ist es nach den Ausführungen von *Muigg* unverzichtbar,

„zukünftig eine Instanz im Fahrzeug zu installieren, welche dem Fahrer trotz aller Vernetzungen, Systemen und Technologien die Möglichkeit bietet, in Ruhe die Kurve zu kriegen“ [166].

Es sind also Überlegungen anzustellen, wie einer drohenden Gefahr der Fahrerüberlastung durch den wachsenden Funktionsumfang zukünftiger Kraftfahrzeuge bei gleichzeitig komplexer werdenden Verkehrsbedingungen entgegengewirkt werden kann.

Vorgehensweise und Aufbau der Arbeit

Aufbauend auf der Erfassung des mentalen Auslastungszustands eines Menschen im komplexen Szenario des Steuerns eines Kraftfahrzeugs – dabei ist für die vorliegende Arbeit vor allem die Erfassung der Komponente *Ablenkung* von besonderer Bedeutung – kann eine adaptive, an den aktuellen Zustand des Fahrers angepasste Koordination der Fahrzeugsysteme erfolgen. Somit wird der Fahrer durch die zusätzlichen Informationen in der jeweiligen Fahrsituation nicht überfordert und die Fahrsicherheit im Idealfall zu keinem Zeitpunkt beeinträchtigt. Die der Arbeit zugrunde liegende Herausforderung war es, zunächst geeignete Daten zu erfassen, die einen effektiven und zuverlässigen Rückschluss auf die Aufmerksamkeitsverteilung des Nutzers im Kraftfahrzeug erlauben, und auf deren Grundlage Konzepte für die Realisierung einer intuitiven Fahrer-Fahrzeug Interaktion zu entwickeln.

Die einfachste Möglichkeit zur Abschätzung des Aufmerksamkeitsfokus stellt die Überwachung der Interaktion dar. Zur Registrierung von Bedienvorgängen und der damit verbundenen Aufmerksamkeitsverlagerung von der Straße hin zum Fahrzeuginnenraum wird in Kapitel 5 auf die Integration eines Touchscreens mit Annäherungssensorik eingegangen, welcher es erlaubt, die Annäherung der Hand zu erfassen, bevor die eigentliche Bedienung stattfindet.

Eine solche Aufmerksamkeitsverlagerung hin zum Fahrzeuginnenraum manifestiert sich zudem in (zum Teil minimalen) Fahrfehlern. Deshalb stellt Kapitel 4 ein erarbeitetes Verfahren zur Überwachung des Fahrerverhaltens auf Basis von Sensordaten vor, um auf diesem Wege Rückschlüsse auf den Aufmerksamkeitsfokus des Nutzers zu ziehen.

Die effektivste Möglichkeit zur Überwachung der Nutzeraufmerksamkeit stellt die Erfassung und Auswertung der Blickrichtung dar. Da diese sehr stark mit der (visuellen) Aufmerksamkeit korreliert, resümiert Kapitel 3 umfangreiche Forschungsarbeiten im Bereich der Blickrichtungserfassung. Diese brachten neben mehreren prototypischen Eye Tracking Systemen schließlich den integrierten remote Eye Tracker *eyeDrive*, eine kombinierte Hard- und Softwarelösung, hervor. Dieses System erwies sich aufgrund seiner

Eigenschaften hinsichtlich Datenrate und Genauigkeit als optimaler Ausgangspunkt zur Umsetzung blickbasierter Konzepte für eine intuitive Fahrer-Fahrzeug Interaktion.

Die realisierte Sensorik erlaubt eine umfassende Analyse der Aufmerksamkeit des Nutzers im Fahrzeug. Auf Basis der gewonnenen Daten erfolgt in Kapitel 5 die Erarbeitung und Bewertung geeigneter Anzeige- und Bedienkonzepte, die den Verlauf der Aufmerksamkeit einbeziehen und dazu nutzen, die Fahrer-Fahrzeug Interaktion intuitiver und weniger ablenkend zu gestalten. In diesem Zusammenhang präsentiert die Arbeit Lösungsvorschläge zur Warnung vor einem drohenden Aufmerksamkeitsverlust, zur effektiven Lenkung der Aufmerksamkeit sowie zur Antizipation von Bedienwünschen.

Zudem präsentiert Kapitel 6 ein multimodales, blickbasiertes Bedienkonzept für Mensch-Maschine Schnittstellen im Fahrzeug. Dieses nutzt die in Kapitel 3 vorgestellte Blickrichtungserkennung in Kombination mit Touch als neuartige Eingabemodalität.

Abschließend gibt Kapitel 7 einen Ausblick auf das Nutzungspotential der vorgestellten Ansätze zur Erfassung der Blickrichtung im Kontext anderer Anwendungsdomänen.

Als Einstieg in die Thematik erarbeitet das folgende Kapitel 2 zunächst die zum Verständnis der Arbeit relevanten Grundlagen und gibt eine thematische Einführung in die Bereiche der Mensch-Maschine-Kommunikation (mit Fokus auf Grundlagen der Anwendungsdomäne Kraftfahrzeug), der Bewertung von Mensch-Maschine Schnittstellen sowie in den Forschungsbereich blickbasierter Benutzerschnittstellen.

2

Grundlagen und Thematische Einführung

Das folgende Kapitel 2 stellt zunächst das Forschungsgebiet der Mensch-Maschine Kommunikation (MMK) vor und vermittelt das nötige Basiswissen zum Verständnis der Begriffe, die in den darauf folgenden Kapiteln Verwendung finden. Aufbauend auf domänenspezifischen, wissenschaftlichen Erkenntnissen erarbeitet das Kapitel ein tieferes Verständnis der Mensch-Maschine Schnittstelle im Automobil. Dazu werden unterschiedliche, psychologisch motivierte Handlungsmodelle vorgestellt und die Aufgaben des Fahrers in Teilaspekten der Fahraufgabe erläutert. Weiterhin schaffen kognitionswissenschaftliche Erkenntnisse ein detailliertes Verständnis für die kognitiven Verarbeitungsprozesse des Fahrers. Der Abschnitt zum Thema der Usability liefert Leitlinien zur Bewertung von Fahrerinformationssystemen. Abschließend werden Methoden zur Bewertung der entwickelten Lösungsansätze sowie Grundlagen der blickbasierten MMK vorgestellt.

2.1 Mensch-Maschine Interaktion – Begriffsdefinition

Bedingt durch den kontinuierlichen Fortschritt im Bereich der Informationstechnik und der damit einhergehenden, zunehmenden Verbreitung von interaktiven, technischen Systemen (Maschinen) im beruflichen und privaten Alltag wird der intuitiven Auslegung der Kommunikation zwischen Maschinen und dem Menschen (als Benutzer) eine besondere Bedeutung beigemessen. Aus dieser Entwicklung hat sich der eigenständige, interdisziplinäre Forschungsbereich der Mensch-Maschine Interaktion¹ (MMI) etabliert, der die Analyse, Konzeption und Evaluierung entsprechender Mensch-Maschine Schnittstellen zum Ziel hat. Das Fachgebiet der HCI ist nach der *ACM Special Interest Group on Computer-Human Interaction* folgendermaßen definiert

„Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.“ [107]

¹Ebenfalls häufig als Mensch-Maschine Kommunikation (MMK) oder um die Bedeutung des Computers herauszustellen als Mensch-Computer-Interaktion (MCI) bezeichnet. Im Englischen wird dafür der Begriff Human-Computer Interaction (HCI) verwendet.

2. Grundlagen und Thematische Einführung

und bündelt neben klassischen Ingenieurs- und Arbeitswissenschaften, der Informatik und der (System- und Software-)Ergonomie ebenfalls Erkenntnisse aus den Forschungsfeldern der Kognitionswissenschaft, der Psychologie, der Medizin und des Designs.

Das Zusammenwirken von technischem System und Benutzer, um ein bestimmtes Ziel zu erreichen bzw. eine Aufgabe zu lösen, wird dabei als Mensch-Maschine System (MMS) bezeichnet [91]. Unter dem Begriff Maschine versteht man dabei allgemein ein technisches System, wie beispielsweise ein Computer-System, eine Fabrikanlage, oder – wie in der vorliegenden Arbeit von Hauptinteresse – ein Kraftfahrzeug. In einem MMS

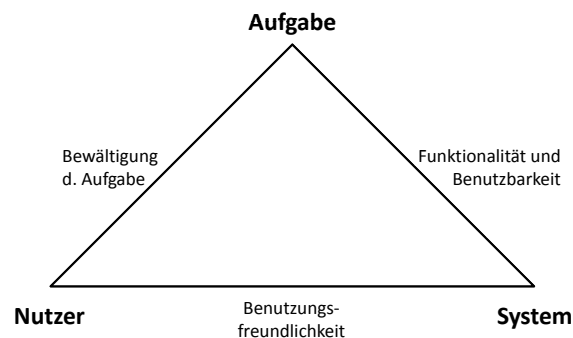


Abbildung 2.1: Passungsmodell der Mensch-Maschine Interaktion zur Beschreibung des Zusammenhangs zwischen Benutzer, System und der zu leistenden Aufgabe.

gilt es, den *Benutzer*, das *System* und die zu leistende *Aufgabe* in Einklang zu bringen. Diese Aspekte stehen durch die Aufgabenbewältigung, die Benutzungsfreundlichkeit und die Funktionalität des Systems untereinander in Beziehung. Oberstes Ziel ist es, eine optimale Passung zwischen den Aspekten zu erreichen,² weshalb bereits in frühen Entwicklungsphasen die Usability (vgl. Kapitel 2.4) berücksichtigt werden muss.

Zur Beschreibung menschlicher Verhaltensweisen in einem MMS existieren in der Literatur diverse Ansätze. Die folgenden Kapitel stellen exemplarisch das *SRK-Modell* nach *Rasmussen* [200] sowie die *MRT* nach *Wickens* [250] vor.

2.2 Verhaltensmodelle in der Mensch-Maschine Interaktion

In der Literatur existiert eine Vielzahl von Modellen zur Beschreibung der Rolle bzw. des Verhaltens eines Menschen im MMS. Diese Modelle bilden zugrunde liegende Eigenschaften des Menschen nach und dienen als gute Basis zum Verständnis unterschiedlicher Handlungsmuster des Menschen im MMS. Die Bandbreite existierender Modelle reicht dabei von der einfachen *Ein-Kanal-Theorie* [46] und der Theorie des *Working Memory* [20] hin zu komplexeren Theorien, wie der *Multiple Resource Theory* [250], dem *Strategic Task Management* [143] oder dem *SRK-Modell* [200]. Die folgenden Abschnitte beschreiben die letzten drei Modelle und für weitere Einzelheiten zu Verhaltensmodellen in der Mensch-Maschine Interaktion sei z.B. auf die ausführlichen Zusammenfassungen in

²Nach [86] auch als auch als Passungsmodell der Mensch-Maschine Interaktion bezeichnet

den kürzlich erschienenen Dissertationen im Automobilen Umfeld von *Ablaßmeier* [7], *Broy* [51], *Lange* [137], *Milicic* [160] und *Tönnis* [230] verwiesen.

2.2.1 Wickens Multiple Resource Theory

Die Kernaussage von Wickens *Multiple Resource Theory* (MRT) ist, dass für verschiedene Modalitäten (visuell, auditiv, taktil, olfaktorisch) eigene Ressourcen zur Verfügung stehen und deshalb manche Dual-Task-Aufgaben ohne gegenseitige Beeinflussung, andere wiederum nur unter gegenseitiger Beeinflussung, durchgeführt werden können [250]. Die MRT teilt die Informationsverarbeitung in drei Dimensionen ein: (1) Verarbeitungsstufe (stage) – (2) Wahrnehmungsmodalität (modality) – (3) Verarbeitungstyp (code). Abbildung 2.2 stellt diese Dimensionen graphisch dar. Informationen durchlaufen die drei Stufen Wahrnehmung (perception/encoding), zentrale Verarbeitung (central processing) und Reaktion (responding). Die wichtigsten Modalitäten sind visuell und auditiv. Der Verarbeitungstyp der Informationen kann räumlich (spatial) oder verbal (verbal) sein.³

Benötigen zwei Aufgaben gleichzeitig dieselben Ressourcen (z.B. verbale Verarbeitung), sinkt aufgrund der gegenseitigen Störung die Leistung bei beiden Aufgaben. Überlappen sich bei der Bearbeitung der Aufgaben die Ressourcen nicht, können beide Aufgaben ohne Beeinträchtigung durchgeführt werden. So kann man z.B. beim Sehen

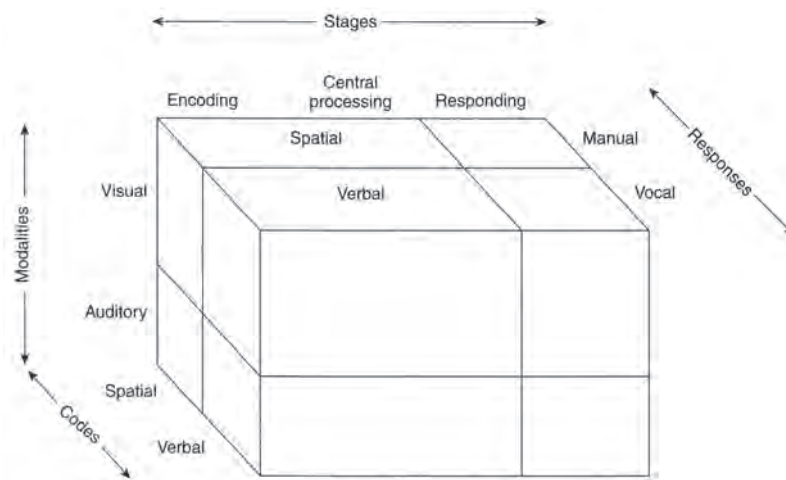


Abbildung 2.2: Ressourcenpools von Wickens Multiple Resource Theory nach [250].

eines Spielfilms ohne Probleme die Bildbewegungen verarbeiten (räumliche Verarbeitung) während man den Ton interpretiert (verbale Verarbeitung). Dagegen ist es nicht möglich, beim Lesen eines Buches gleichzeitig ein Hörbuch zu hören (beides verbale Verarbeitung).

Die Autoren von [142] weisen zusätzlich auf die verschiedenen Ressourcen von Wahrnehmung/Verarbeitung und Reaktion hin. So ist es möglich auf einen Stimulus

³Der Einfachheit halber wurden die weniger wichtigen Modalitäten (taktil und olfaktorisch) sowie ein Verarbeitungstyp (unterbewusst) in Abbildung 2.2 weggelassen.

zu reagieren und gleichzeitig einen anderen Stimulus wahrzunehmen, aber nicht für beide Stimuli gleichzeitig eine Reaktion vorzubereiten/durchzuführen. Dieser von *Pashler* gefundene Zusammenhang wurde von *Strayer* und *Johnston* in einem Experiment bestätigt. Sie zeigten, dass das Folgen eines Gesprächs mit dem Beifahrer (verbale Reaktion auf auditiven verbalen Stimulus) während dem Autofahren zu signifikant schlechterer Fahrleistung führte, als das Hören eines Hörbuches (ebenfalls auditiver verbaler Stimulus, jedoch keine Reaktion notwendig).

Die MRT ist ein leistungsfähiges Modell, um Situationen zu interpretieren oder vorherzusagen, in denen kognitive Aufgaben die Kapazitäten des Fahrers überfordern [143], sie erklärt jedoch nicht alle Effekte im Kontext des Mental Workload: So ist z.B. die elastische Kapazität nicht in das Modell integriert. Eine Theorie, die sich mit diesem Effekt auseinandersetzt, ist die *Malleable Attentional Resource Theory* (MART) von *Young* und *Stanton* [143]. Die MART ist besonders in Bereichen von sehr niedrigen Aufgabenanforderungen hilfreich, wogegen die MRT eher in Bereichen hoher Anforderungen Anwendung findet. Weitere Faktoren, welche die Leistung und auch den Mental Workload bei einer Aufgabe stark beeinflussen, jedoch nicht in diesen Ressourcen-Theorien berücksichtigt werden, sind:

- Aufgabenpriorisierung
- Aufmerksamkeitsabbruch
- Aufgabenvorherbestimmung
- Aufgabenscheduling und -switching statt Parallelbearbeitung

Die Theorie des *Strategic Task Management* setzt sich mit diesen Punkten auseinander.

2.2.2 Strategic Task Management

Lee et.al. liefern eine Zusammenfassung zum Konzept des *Strategic Task Management* [143], weshalb an dieser Stelle lediglich die wichtigsten Punkte zusammengefasst werden. Bearbeitet ein Nutzer parallel mehrere Aufgaben, geschieht deren Bearbeitung oft nicht gleichzeitig. Die Bearbeitungszyklen der einzelnen Aufgaben unterbrechen sich vielmehr gegenseitig. Der sog. *Zeigarnik-Effekt* besagt, dass Subjekte einen sehr starken Drang haben, die begonnene (Teil-)Aufgabe zu beenden. Die Unterbrechung einer (Teil-)Aufgabe durch eine andere (Teil-)Aufgabe führt zu einer fortwährenden Belastung des menschlichen Arbeitsspeichers durch die nicht beendete (Teil-)Aufgabe.

Das Wechseln zwischen Aufgaben (task switching) führt zu einer schlechteren Einzelleistung bei beiden Aufgaben. Die zusätzliche Belastung (switching cost) kann durch Vorbereitung auf die Aufgabe oder durch Vorhersehbarkeit der Aufgabe verringert (jedoch nicht eliminiert) werden. Die Ebene, auf der die jeweilige Aufgabe vom Akteur bearbeitet wird, hat ebenfalls Einfluss auf den Mental Workload. Das *SRK-Modell* nach *Rasmussen* veranschaulicht diesen Zusammenhang.

2.2.3 SRK-Modell nach Rasmussen

Das SRK-Modell – oder auch Drei-Ebenen-Modell – nach *Rasmussen* ist ein gebräuchlicher Ansatz zur Beschreibung von menschlichem Verhalten oder Fähigkeiten in einem

MMS. SRK steht für *skill-* (fertigungs-), *rule-* (regel-) und *knowledge-based* (wissensbasiert) und bezeichnet die drei Ebenen, auf denen Informationen verarbeitet werden können [200]. Abbildung 2.3 zeigt den zugrunde liegenden Informationsverarbeitungsprozess.

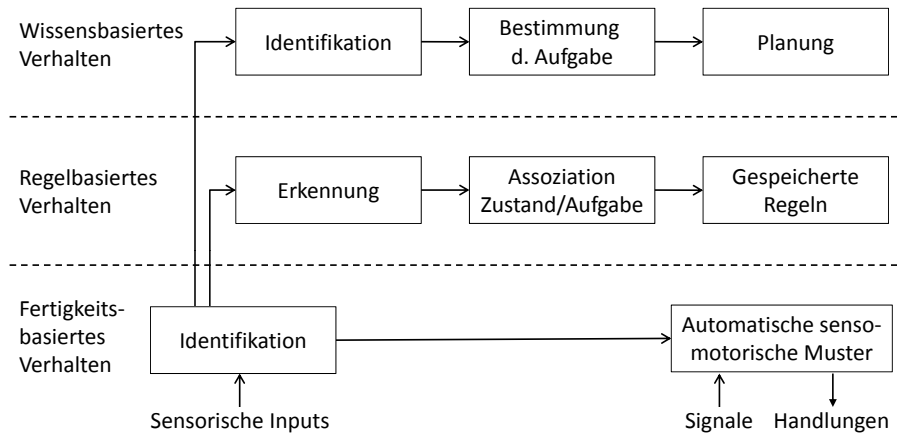


Abbildung 2.3: SRK-Modell nach Rasmussen [200].

Dieser ist durch die einzelnen Handlungsebenen „wissensbasiertes“ und „regelbasiertes Verhalten“ sowie „Fertigkeiten“ hierarchisch strukturiert, wobei die einzelnen Handlungsebenen sich hauptsächlich im Zeitbedarf zur Informationsverarbeitung unterscheiden.

Fertigkeitsbasiertes Verhalten stellt die unterste Handlungsebene dar und findet statt, wenn die Person sehr erfahren im Umgang mit dem Aufgabentyp ist. Hierbei führen Signale in erlernten, schnell und unbewusst ablaufenden Reiz-Reaktions-Automatismen zu Handlungen. Regelbasiertes Verhalten kommt zum Einsatz, wenn die Person den Aufgabentyp zwar kennt, aber noch nicht sehr erfahren ist. Das Verhalten zeichnet sich durch das Erkennen von bekannten Zuständen oder Situationen aus. Für bekannte Zustände wurde im Rahmen von früheren Gelegenheiten (z.B. durch aktives oder mentales Training) ein Set von Regeln angelegt. Aus diesem Repertoire wird bei Bedarf das effizienteste Verhaltensmuster abgerufen. Wissensbasierte Verarbeitung kommt zum Einsatz, wenn das Subjekt mit der Aufgabe gar nicht vertraut ist, und somit auch noch keine Regeln – aufbauend auf Erfahrung mit der Aufgabe – bilden konnte. Auf der Grundlage von vorhandenem und beschaffbarem Wissen werden vor der Umsetzung der vielversprechendsten Variante unterschiedliche Handlungsalternativen entwickelt und verifiziert [45, 143].

2.3 Domänenspezifische Grundlagen

Das folgende Kapitel transferiert diese allgemeinen Grundlagen der MMS in die Fahrzeugdomäne. Dazu werden die Rolle, das Verhalten und die Aufgaben des Nutzers in seiner Funktion als Kraftfahrzeugführer und Verkehrsteilnehmer genauer definiert. Für umfassendere Ausführungen zu den einzelnen Teilaspekten sei wiederum auf die Arbeiten von *Ablaßmeier* [7], *Broy* [51], *Lange* [137], *Milicic* [160] und *Tönnis* [230] verwiesen.

2.3.1 Fahraufgabe

Das sichere Führen eines Fahrzeugs ist eine hochkomplexe Aufgabe: Der Fahrer muss eine Vielzahl an Sinneseindrücken verschiedenster Modalitäten aufnehmen, verarbeiten und entsprechend darauf reagieren. Die Aufnahme der Informationen aus der Umgebung erfolgt dabei im Wesentlichen über den visuellen Sinneskanal.⁴ Weiterhin beeinflussen akustische (z.B. Motorgeräusch, Hupen), kinästhetische (z.B. Längs- und Querschleunigung) und haptische (z.B. Kraftübertragung vom Fahrwerk ans Lenkrad) Reize den Fahrer. Die Bewältigung aller beim Fahren anfallenden Aufgaben erfordert also ein hohes Maß an Konzentration.

Ausgehend von den Erkenntnissen von *Bernotat* [34], *Rasmussen* [200] und *Donges* [64] erarbeiteten verschiedene Arbeiten (z.B. [52, 78, 90, 91, 158]) unterschiedliche Taxonomien, deren Gemeinsamkeit in einer Unterteilung der Fahreraufgaben in drei Bereiche unterschiedlicher Priorität und Komplexität besteht. Vor allem im englischsprachigen Bereich wird zudem eine Unterteilung in Primär- und Nebenaufgaben (Primary und Secondary Tasks) verwendet. Ähnlich dem Drei-Ebenen-Modell von *Rasmussen* (vgl.



Abbildung 2.4: Platzierung der Bedienelemente im Fahrzeugcockpit entsprechend ihrer Zugehörigkeit zu primärer (blau), sekundärer (grün) und tertiärer (rot) Fahraufgabe [230].

Kapitel 2.2.3) lassen sich nach [34, 52, 90, 78, 91, 158] die Aufgaben des Fahrers in die drei Bereiche *primäre*, *sekundäre* und *tertiäre Fahraufgabe* mit absteigender Priorität unterteilen. Abbildung 2.4 zeigt nach [231] die Bedienelemente im Fahrzeug-Cockpit farblich hervorgehoben entsprechend ihrer Zugehörigkeit zu diesen Bereichen.

2.3.1.1 Primäre Fahraufgabe

Die primäre Aufgabe des Fahrers ist die sichere Steuerung des Fahrzeugs durch den Verkehr. Dazu zählt es u.a., das Fahrzeug auf der richtigen Fahrspur zu halten und

⁴Nach [209] werden ca. 86% der Informationen aus der Umwelt über den visuellen Kanal aufgenommen.

Berührung oder Gefährdung anderer Objekte zu vermeiden. Nach [90] kann man diese Aufgabe in die drei Ebenen *Stabilisierung*, *Führung* und *Navigation* untergliedern. Auf der untersten Ebene erfolgt die Stabilisierung des Fahrzeugs aufgeteilt in Längs- und Querverführung, durch Lenken, Gasgeben und Bremsen wird das Fahrzeug auf der gewählten Fahrspur und im gewünschten Abstand zu anderen Objekten gehalten. Die mittlere Ebene stellt die Führungsaufgabe dar und beinhaltet die Wahl der Geschwindigkeit und des Weges (z.B. die Wahl der Fahrspur, die Berücksichtigung von Verkehrszeichen und Fremdverkehrsteilnehmern oder den Entschluss zum Überholen). Die oberste Ebene,



Abbildung 2.5: Die drei Ebenen der primären Fahraufgabe nach [52].

die Navigationsaufgabe, dient der Wahl des Weges vom Start zum Ziel abhängig von örtlichen, zeitlichen oder auch meteorologischen Bedingungen, da z.B. die gewählte Strecke und die erwartete Fahrtdauer abhängig von der Tageszeit und damit verbundenem Verkehrsaufkommen ist. Diese Aufgaben des Fahrers sind als essentiell zu erachten, um das Fahrziel zu erreichen und ihnen wird während der Fahrt die höchste Priorität eingeräumt.

2.3.1.2 Sekundäre Fahraufgabe

Sekundäre Fahraufgaben ergeben sich aus den Verkehrsregeln und den Verkehrs- sowie Umweltbedingungen, haben jedoch nicht direkt damit zu tun, das Fahrzeug auf der Straße zu halten. Sie umfassen reaktive und aktive Aufgaben [201]: Zu den reaktiven Aufgaben zählen beispielsweise Abblenden, Betätigen des Scheibenwischers, Kuppeln und Schalten oder auch das Bedienen des Tempomats, aktive Aufgaben dienen eher dem Zweck der Kommunikation mit anderen Verkehrsteilnehmern (z.B. Blinken oder Hupen).

2.3.1.3 Tertiäre Fahraufgabe

Tertiäre Fahraufgaben betreffen weniger das eigentliche Autofahren als vielmehr das Zufriedenstellen von Komfort-, Unterhaltungs- und Informationsbedürfnissen des Fahrers. Dazu zählen beispielsweise Klimaanlage, Sitzheizung, CD-Player, Autoradio und Telefon. Dabei können diese System jedoch indirekt Einfluss nehmen auf primäre und sekundäre Fahraufgaben: So kann z.B. einerseits durch Steigerung des Wohlbefindens des Fahrers bei hohen Außentemperaturen die Klimaanlage zu einem entspannteren und somit weniger unfallträchtigen Fahrstil führen, andererseits kann sich das Unfallrisiko durch Ablenkung bei komplizierter und wenig intuitiver Bedienung der Geräte stark erhöhen. Darauf ist bei der Entwicklung von Nutzerschnittstellen in Fahrzeugen besonders zu achten.

2.3.2 Einfluss der Fahraufgabe auf die mentale Belastung

Das Autofahren ist eine komplexe Aufgabe mit Prozessen auf mindestens drei Ebenen. Auf all diesen Ebenen kann der Workload zu hoch werden, was sich dann auch auf die Leistung in den anderen Ebenen und schlussendlich auf die Fahrsicherheit auswirkt (z.B. kann zu hohe Beanspruchung durch sehr viel Verkehr auf der Manövrierebene in ungewohnter Umgebung zum Übersehen von Verkehrsschildern führen) [248].

Die einzelnen Ebenen der Fahraufgabe stellen unterschiedliche Anforderungen an den Fahrer dar und laufen zumeist auf unterschiedlichen Handlungsebenen nach Rasmussens Modell (vgl. Kapitel 2.2.3) ab. Das Festlegen der Route auf der Navigationsebene erfolgt bewusst und erzeugt in hohem Maß mentale Last. Handlungen auf der Fahrzeugführungsebene erfolgen weitgehend bewusst, erfordern jedoch in der Regel ein geringeres Maß an mentaler Anstrengung. Handlungen auf der Stabilisierungsebene werden hingegen hoch automatisiert durchgeführt und erfordern geringere mentale Kapazitäten. Die Bearbeitung unbewusster Prozesse erfolgt schnell und ohne Anstrengung, weshalb mehrere Aufgaben parallel bearbeitet werden können. Bewusste Prozesse laufen hingegen seriell ab und sie beanspruchen die Aufmerksamkeit in hohem Maße. Ihre Bearbeitung gestaltet sich langsam und aufwendig, weshalb für diese Aufgaben nur begrenzte Kapazitäten verfügbar sind [95, 259].

Zusätzlich wirken Infotainmentsysteme⁵ mit zunehmender Frequenz auf den Fahrer ein und zwingen ihn, neben der stark visuell-motorischen Fahraufgabe zusätzlich visuelle, auditive und taktile Informationen zu verarbeiten und teilweise (mit einer Reaktion) zu beantworten. Es wird geschätzt, dass diese Aufgaben einen Großteil der Verkehrsunfälle verursachen. [50] bezeichnet dies unter Berufung auf zwei Studien sogar als Hauptunfallursache. Auch *Verwey* ist sich sicher, dass in 30–50% der Unfälle mentale Ablenkung eine Rolle spielt [242]. Nach [223] sinkt zwar die Zahl der Todesopfer jährlich, die Anzahl der Unfälle insgesamt steigt jedoch von Jahr zu Jahr kontinuierlich an. Da die Anzahl sekundärer Aufgaben mit der stetig steigenden Anzahl an Systemen im Fahrzeug ebenso zunimmt, ist es in Zukunft unabdingbar, die mentale Beanspruchung des Fahrers zu beachten: Wenn er schon nicht durch die Systeme entlastet wird, sollte er wenigstens nicht durch zusätzliche Informationen überfordert werden. Dies könnte durch sog. *adaptive Informationssysteme*, abhängig vom Mental Workload, geschehen.

2.3.3 Fahrerarbeitsplatz

Dieses Kapitel betrachtet die Mensch-Maschine Schnittstelle im Fahrzeug näher. Dazu werden zunächst die Begriffe Fahrerassistenz- und Fahrerinformationssystem gegeneinander abgegrenzt und anschließend verschiedene aktuelle Ansätze und Anzeige-Bedienkonzepte einzelner Fahrzeughersteller vorgestellt.

2.3.3.1 Fahrerinformation

Unter dem Begriff *Fahrerinformationssystem* (FIS) versteht man

⁵Gebräuchliches Kunstwort, gebildet aus Fahrerinformationssystem und Entertainmentsystem.

„[...] eine Software-Plattform mit Applikationen [...], welche in einem Fahrzeug Informationen, beispielsweise über den Verkehr, Fahrtrouten, Nachrichten, Kommunikationsmöglichkeiten, wie Mobilfunk oder Internet, und Unterhaltung, wie Radio, CD, MP3, mit einer integrierten Steuerung zur Verfügung [stellt].“ [149]

Im Englischen werden diese häufig als *In-Vehicle Information Systems* (IVIS) oder *Driver Information Systems* (DIS) bezeichnet. Eine reduzierte Version eines FIS stellt der sog. *Bordcomputer* dar, der die Informationen im Kombi-Instrument, auf einem separaten Bildschirm in der Mittelkonsole oder auf dem Armaturenbrett anzeigt. Zu den anzeigbaren Informationen eines Bordcomputers gehören z.B. der Durchschnitts-/Momentanverbrauch, die verbleibende Reichweite, die Durchschnittsgeschwindigkeit, die Außentemperatur sowie Datum und Uhrzeit. In modernen Fahrzeugen stellen FIS weiterhin eine Vielzahl an Komfort- und Kommunikationsfunktionen (z.B. Radio, CD, Telefon, Internet, etc.) der tertiären Fahraufgabe bereit. Zumeist sind diese in einem integrierten Anzeige-Bedienkonzept zusammengefasst.

2.3.3.2 Fahrerassistenz

Als *Fahrerassistenzsystem* (FAS) bezeichnet man elektronische Systeme in Fahrzeugen, die den Fahrer bei der Bewältigung der verschiedenen Fahraufgaben (zumeist Inhalte der primären Fahraufgabe) unterstützen. Im Englischen werden diese als (*Advanced*) *Driver Assistance Systems* (A)DAS bezeichnet. Im Vordergrund stehen meist Sicherheitsaspekte oder auch die Steigerung des Fahrkomforts. Die Möglichkeiten der Unterstützung reichen dabei von Warnungen vor kritischen Situationen bis zu selbständigen Regeleinriffen. FAS unterstützen den Fahrer dabei, Anforderungen der primären und sekundären Fahraufgabe zu erfüllen, greifen (teil-)autonom in Antrieb, Steuerung und Signalisierungseinrichtungen des Kraftfahrzeuges ein und warnen vor kritischen Situationen.

FAS können dem jeweiligen Teilaspekt der Fahraufgabe zugeordnet werden, den sie unterstützen. Die bekannten Assistenzsysteme Antiblockiersystem (ABS) und elektronisches Stabilitätsprogramm (ESP) arbeiten auf der Stabilisierungsebene der primären Fahraufgabe. Der Führungsaufgabe kann man den Tempomat – in modernster Ausführung mit kombinierter automatischer Geschwindigkeits- und Abstandsregelung als Adaptive Cruise Control (ACC) bezeichnet – oder den Spurhalteassistenten (Lane Departure Warning, LDW) zuordnen. Assistenzsysteme für die Navigationsaufgabe in Form von Navigationssystemen existieren ebenfalls bereits seit vielen Jahren. Beispiele für Assistenzsysteme der sekundären Fahraufgaben sind das automatische Schaltgetriebe, der Fernlichtassistent (FLA) oder der automatische Scheibenwischer. Schließlich existiert eine Vielzahl an Systemen, die tertiäre Aufgaben unterstützen: Dazu zählt u.a. die Möglichkeit individuelle Sitzpositionen unterschiedlicher Fahrer abzuspeichern und diese beim Zurückkehren ins Fahrzeug selbständig wiederherzustellen. Einen heutigen Trend im Bereich der FAS stellen intelligente Fahrzeugsysteme dar, die Informationen priorisieren und diese bei großer Arbeitsbelastung filtern und verzögern können. So werden z.B. wenn der Fahrer zum Überholen ansetzt, Telefonate oder Informationen von diesen als Workload-Manager bezeichneten Systemen zurückgehalten [166].

2.3.3.3 Aktuelle Anzeige-Bedienkonzepte

Heutige FIS bündeln die Vielzahl verfügbarer Informationen zumeist in einem zentralen Display (CID)⁶. Hier ermöglicht ein einheitliches Anzeige-Bedienkonzept (ABK) den konsistenten Zugriff auf Informationen und die Kontrolle (sämtlicher) Fahrzeugfunktionen, indem sie Bedienvorgänge kaskadieren und in hierarchischen Menüsystemen mit (möglichst) wenigen Tasten zusammenfassen, um die Aufmerksamkeitsverlagerung während der Bedienung möglichst gering zu halten. Ein Vorteil dieses Ansatzes ist es, dass mit wenigen gut bekannten Tasten der gesamte Funktionsumfang bedient werden kann. Als Nachteil stellt sich die zunehmende Komplexität der Menüführung und der lange Bedienverlauf heraus [55], der benötigt wird, um die gewünschte Funktion auszuführen, wodurch solche Systeme für den Fahrer schwerer zu überblicken und zu bedienen werden [97].

Die Systeme verschiedener Hersteller unterscheiden sich hauptsächlich in den Bereichen Anzeigeort und Bedienstrategie. Bei Automobilherstellern und Anbietern von Nachrüstsystemen zeichnen sich aktuell zwei grundlegende Philosophien ab: *Trennung von Anzeige und Bedienung* (vgl. Abbildung 2.6) und *Kombination von Anzeige und Bedienung an einem Ort* (vgl. Abbildung 2.7) Die Hersteller Audi, BMW und Mercedes-Benz setzen



(a) Cockpit Audi A8.



(b) Cockpit BMW 5er.



(c) Mercedes-Benz S-Klasse.



(d) Bedienelement Audi A8.



(e) Bedienelement BMW 5er.



(f) Bedienelement S-Klasse.

Abbildung 2.6: Aktuelle Anzeige-Bedienkonzepte unterschiedlicher Automobilhersteller mit strikter Trennung von Anzeige und Bedienung.

auf eine strikte Trennung von Anzeige und Bedienung. Die Anzeige erfolgt auf dem CID im mittleren, oberen Bereich des Armaturenbrettes (vgl. Abbildung 2.6). Somit muss der Fahrer seinen Blick bei der Bedienung nicht zu weit nach unten richten. Die Bedie-

⁶Auch im deutschen oft gebrauchte Abkürzung des englischen Begriffs *Central Information Display*.

2.4. Usability-Untersuchung von Fahrer-Fahrzeug Schnittstellen

nung erfolgt über ein zentrales Bedienelement⁷, das in der Mittelkonsole gut erreichbar positioniert ist. Grundfunktionen sind über sog. Hard Keys⁸ zugänglich.

Die Hersteller Toyota/Lexus, Ford, Volkswagen, etc. setzen auf Anzeige und Bedienung an einem gemeinsamen Ort unter Verwendung eines Touchscreens (vgl. Abbildung 2.7). Dies findet sich auch bei aktuellen Nachrüstsystemen (z.B. Blaupunkt, JVC, Pioneer,



(a) Ford.



(b) Volkswagen.



(c) Lexus.



(d) Toyota Prius.



(e) Nachrüstsystem Alpine.



(f) Nachrüstsystem Zenec.

Abbildung 2.7: Aktuelle Anzeige-Bedienkonzepte unterschiedlicher Automobilhersteller mit Kombination von Anzeige und Bedienung an einem gemeinsamen Ort.

Zenec, etc.). Einzig der Hersteller Alpine bietet zu seinen Touch-basierten Nachrüstgeräten ein optionales zentrales Bedienelement in Form eines Dreh-/Drückstellers an.

2.4 Usability-Untersuchung von Fahrer-Fahrzeug Schnittstellen

Der englische Begriff „Usability“ kann mit den Begriffen der „Gebrauchstauglichkeit“ oder „Brauchbarkeit“ übersetzt werden, was jedoch die Begriffsdefinition nicht exakt wiedergibt. Im Deutschen hat sich deshalb ebenfalls der Begriff „Usability“ durchgesetzt [215], der 1990 folgendermaßen definiert wurde:

„Usability is a concept comprising the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment.“ [6]

In einer Usability-Studie werden dem Probanden unbekannte Systeme vorgestellt. Er kann sich kurz mit dem neuen System vertraut machen und soll anschließend anhand vordefinierter und reproduzierbarer Abläufe Aufgaben und Bedienabläufe – sog. Use-cases – vollziehen, um eine qualifizierte Aussage über die Gebrauchstauglichkeit des

⁷Zumeist als multifunktionaler Dreh-/Drücksteller ausgeführt.

⁸Knöpfe mit fest vergebener Funktion, z.B. Einsprünge in die Menüs Navigations, Klimasteuerung, etc.

2. Grundlagen und Thematische Einführung

Systems zu geben. Die Studie wird grundsätzlich von einem Versuchsleiter geführt, den die Probanden während des Ablaufes bei Fragen und Problemen kontaktieren können. Nach dem Versuch müssen die Probanden einen Fragebogen ausfüllen, um ihre subjektive Bewertung zum Versuchsablauf festzuhalten. Die objektive Bewertung erfolgt durch den Simulationsablauf während der Fahraufgabe und der relevanten Fahrdaten, welche in einer Protokolldatei hinterlegt sind. Mittels dieser Daten kann eine Analyse erfolgen, inwieweit dem Probanden der Umgang mit dem System gelang bzw. er in seiner primären Fahraufgabe beeinträchtigt war. Dabei ist herauszustellen, dass die empfundene Attraktivität eines Produktes bei Usability-Untersuchungen mehr und mehr in den Vordergrund rückt. Viele Verfahren beschränken sich im Wesentlichen auf die Bedienbarkeit (pragmatische Qualität). Aus diesem Grund führt [101] den Begriff der User Experience (UX) ein, der eine deutliche Erweiterung und Veränderung der klassischen Usability darstellt. UX betont besonders die ganzheitliche Sicht auf die subjektiv wahrgenommene Produktqualität, wobei nicht nur die sog. pragmatische, sondern auch die hedonische Qualität eines Produkts von Interesse ist.

2.4.1 Verwendete Bewertungsmethoden

Zur eingehenden Bewertung der im Rahmen der Arbeit entwickelten Anzeige-Bedienkonzepte hinsichtlich ihrer Eignung zum Einsatz im Fahrzeug wurden etablierte Werkzeuge ausgewählt, die im Folgenden beschrieben werden.

2.4.1.1 Fahrsimulator

Ein unverzichtbares Werkzeug zur Evaluierung und Bewertung von neuartigen Anzeige-Bedienkonzepten im Konzeptstadium stellt die Durchführung von Studien im Fahrsimulator dar. Um den Bediensituationen im realen Umfeld gerecht zu werden, ist eine „Nebenaufgabe“ zur Bedienung des MMI unerlässlich. Der tatsächlichen Aufgabe des Fahrens im späteren Fahrzeug soll der Versuchsaufbau idealerweise möglichst nahe kommen. Die Fahraufgabe induziert dabei eine Workload, die möglichst über alle Versuchspersonen und Aufgaben konstant gehalten werden soll. Die folgenden Sichtsimulationen wurden im Rahmen der Arbeit als primäre Fahraufgabe eingesetzt.

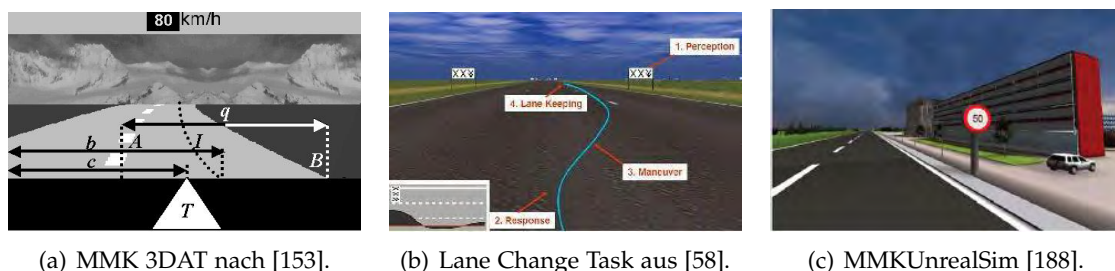


Abbildung 2.8: Beispiele unterschiedlicher Sichtsimulationen.

Abbildung 2.8(a) zeigt die Fahrersicht des Fahrsimulators *MMK 3DAT*. Dieser stellt eine simple Regelaufgabe dar, bei der der Proband versuchen muss, das Fahrzeug immer

möglichst in der Mitte des eigenen Fahrschlauchs zu halten. Als Maß für die Güte der Erfüllung der primären Fahraufgabe wird aus den aufgezeichneten Fahrzeugdaten (Geschwindigkeit und Spurhaltung) im Anschluss an die Versuchsfahrten der *Error Score* für die einzelnen Teilaufgaben und Versuchspersonen berechnet. Dieser ist ein unabhängiges numerisches Validierungsmaß, um die Fahrleistung der einzelnen Versuchspersonen objektiv beurteilen zu können. Zur Berechnung des Error Score sei auf [153] verwiesen.

Abbildung 2.8(b) zeigt den Simulator *Lane Change Task*, der von der Daimler AG - Research & Technology für wissenschaftliche Untersuchungen zur Verfügung gestellt wird [58]. Die Aufgabe des Probanden besteht im Fahren auf einer dreispurigen, geraden Straße mit fest vorgegebener Höchstgeschwindigkeit, welche im Rahmen der beschriebenen Fahrversuche auf 60 km/h festgesetzt und bei durchgedrücktem Gaspedal erreicht wurde. An beiden Seiten des Straßenrands sind in festen Abständen Schilder platziert, die dem Probanden die zu befahrene Spur anzeigen. Die Aufgabe der Probanden besteht im Halten der Maximalgeschwindigkeit bei gleichzeitiger möglichst unmittelbarer und exakter Durchführung der Spurwechselaufforderungen. Als Maß für die Güte dient die nach [58] berechnete mittlere Spurabweichung. Diese Fahrsimulation stellt jedoch lediglich ein Szenario zur Verfügung (dreispurige Straße mit Spurwechselanweisungen) und ist ein in sich geschlossenes Softwarepaket – Kontextparameter können nicht in das zu testende Konzept integriert werden (z.B. Geschwindigkeit, Motordrehzahl, etc. zur Anzeige im Kombiinstrument).



Abbildung 2.9: Der MMK Fahrsimulator. Die Abbildung zeigt die stereoskopische Projektion der Fahrszene sowie der Cockpitenelemente in der Vierseiten-Cave des Lehrstuhls. Die Darstellung ist dabei an die Kopfposition des Probanden angepasst.

Aus diesen Gründen wurde im Rahmen der Arbeit die Plattform *MMKUnrealSim* konzipiert (vgl. Abbildung 2.8(c) und [188]), wobei im Besonderen auf eine modulare Softwarestruktur, eine sehr gute Visualisierungsqualität sowie freie Konfiguration der

Szenarien geachtet wurde. Weiterhin sollte die Plattform die Verwendung von Standardhardware (z.B. Spielelenkrad), eine beliebige Skalierbarkeit der Laborumgebung (z.B. Desktopumgebung, einfache Sitzkiste, Mehrseiten Setup mit großem Field-of-View) sowie einen beliebig erweiterbaren Funktionsumfang (z.B. Integration von Fahrerassistenzfunktion wie ABS, etc.) zulassen. Auf Basis der Vorüberlegungen hat man sich für die Verwendung der Unreal Engine 2.5 des Computerspiels UT2004 [77] entschieden, die folgende Vorteile vereint: beliebige Modifikation von Spielinhalten, einen leistungsfähigen Editor zur Erstellung eigener Szenarien sowie eine sehr gute Grafikengine. Weiterhin sind aufgrund der hohen Popularität der Engine [76] eine Vielzahl an Erweiterungen aus der großen Entwicklergemeinschaft [37] verfügbar. Zu Details der Umsetzung der einzelnen Komponenten sei u.a. auf die am Lehrstuhl entstandenen studentischen Arbeiten [71, 88, 126, 170, 207] verwiesen.

2.4.1.2 Peripheral Detection Task (PDT)

Als besonders verlässliche Methode zur Messung des Mental Workload mittels Sekundäraufgaben hat sich in diversen Untersuchungen der sog. *Peripheral Detection Task* (PDT) herausgestellt [32, 150, 178]. Im Rahmen dieser Detektionsaufgabe muss der Proband Stimuli durch peripheres Sehen erkennen und per Knopfdruck bestätigen. Als Maße dienen die Reaktionszeit und das Verhältnis nicht detektierter Punkte zur Gesamtzahl der Stimuli. Der PDT ist vor allem zur Erfassung kurzfristiger, hoher Änderungen der

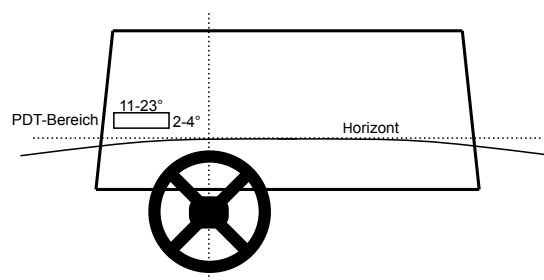


Abbildung 2.10: Schematischer Versuchsaufbau Aufbau beim PDT.

mentalen Beanspruchung geeignet, die die meisten herkömmlichen Maße nicht erfassen können. Weiterhin bildet der PDT den kognitiven Tunneleffekt ab [150]. Dabei verkleinert sich das funktionale visuelle Feld mit größerer kognitiver Belastung. [150] und [178] beschreiben zwar leicht unterschiedliche Konfigurationen des PDT. Im Wesentlichen gleichen sie sich aber in folgenden Parametern: Im peripheren Blickfeld werden Stimuli (kleine rote Punkte) für jeweils 1 Sekunde eingeblendet, deren zeitlicher Abstand zufällig (gleichverteilt) zwischen 3 und 6 Sekunden variiert. Die Stimuli werden in einem Bereich von 11° bis 23° links der vertikalen Achse, die durch die Augenmitte des Probanden läuft, und im Bereich von 2° bis 4° oberhalb der horizontalen Achse, die durch den Horizont gebildet wird, präsentiert. Eine Reaktion (durch Drücken eines Knopfes) ist nur zwischen 200 ms und 2 Sekunden nach Präsentation des Stimulus gültig (sonst wird sie nicht ausgewertet).

2.4.1.3 Skala zur Erfassung subjektiv erlebter Anstrengung (SEA)

Die RSME-Skala (Rating Scale Mental Effort) ist eine subjektive Selbstbewertungsskala, die von *Zijlstra* 1983 entwickelt und von *Eilers* 1986 ins Deutsche zur SEA-Skala übertragen wurde. Dabei erfolgte – aufgrund der abweichenden Bedeutungen der Begriffe im Deutschen – eine Umplatzierung der Ankerpunkte, um die psychometrische Validität beizubehalten [74]. Die SEA-Skala ist eine kontinuierliche Skala von 110 mm Länge, auf der alle 10 mm gekennzeichnet sind und entlang derer sich sog. Ankerpunkte befinden (vgl. Abbildung 2.11). Probanden können auf dieser Skala ihre subjektiv erlebte mentale Beanspruchung bewerten, indem sie auf der Skalenlinie einen entsprechenden Wert ankreuzen, wobei die Ankerpunkte zur Orientierung dienen. Im Idealfall erfolgt die Bewertung der subjektiven Anstrengung durch den Probanden direkt nach Erfüllung der Aufgabe.

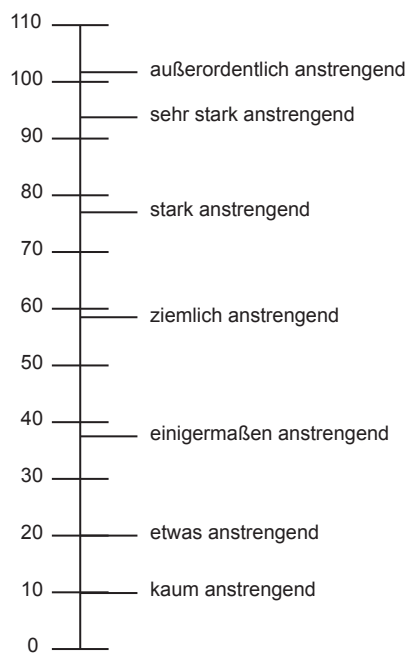


Abbildung 2.11: Skala zur Erfassung subjektiv erlebter Anstrengung.

Die SEA-Skala wird zwar in vielen Bereichen zur Erfassung der mentalen Beanspruchung eingesetzt, eine Überprüfung der Validität der SEA-Skala zeigte jedoch, dass die Skala zwar grundsätzlich – aber nicht unter allen möglichen Randbedingungen – zwischen mentaler Anstrengung differenzieren kann und einfacher zu verwenden ist als vergleichbare mehrdimensionale Skalen (z.B. Nasa TLX). Im GEM-Projekt (*Generic Evaluation Methodology for integrated driver support applications*) wurde festgestellt, dass die RSME/SEA-Skala (neben Sekundäraufgaben-Maßen) sehr gut zur Messung von (visuellen und Mental) Workload-Spitzen geeignet ist [109].

2.4.1.4 Polaritätsprofil (AttrakDiff)

Für den Erfolg eines neuen Systems ist neben der reinen Funktionalität auch die Akzeptanz bzw. Attraktivität von Bedeutung. Zur Erfassung dieser Aspekte wurde der *AttrakDiff*-Fragebogen [4] eingesetzt, der mittels eines Polaritätsprofils (semantisches Differential) die Wahrnehmung des Systems durch die Probanden erfasst. Es sei jedoch darauf hingewiesen, dass einige Versuche im Wesentlichen einen (funktionalen) Prototypen untersuchten, weshalb die Ergebnisse eher als Tendenz und nicht als Endergebnis zu interpretieren sind. Es war für die Probanden teils sehr schwierig, die oft sehr emotional ausgerichteten Bewertungsmöglichkeiten mit dem (eher provisorischen) Konzept zu assoziieren. Der Fragebogen erfasst folgende Dimensionen des Konzepts: Pragmatische Qualität (PQ), hedonische Qualität – Stimulation (HQ-S), hedonische Qualität – Identität (HQ-I) und Attraktivität (ATT). Die pragmatische Qualität beschreibt die Benutzbarkeit eines Produktes und verdeutlicht, wie gut der Nutzer seine Ziele mit Hilfe des Produkts erreichen kann. Die hedonische Qualität – Stimulation bildet ab, inwieweit das Produkt das Bedürfnis des Menschen nach Weiterentwicklung unterstützen kann, indem es neuartige, interessante und anregende Funktionalitäten, Inhalte, Interaktions- und Präsentationsstile bietet. Die hedonische Qualität – Identität beschreibt, inwieweit ein Produkt seinem Nutzer ermöglicht, sich mit dem Produkt zu identifizieren. Die Attraktivität schlussendlich beinhaltet eine globale Bewertung des Produkts auf Basis der wahrgenommenen Qualität. Die pragmatische und hedonische Qualität sind dabei unabhängig voneinander und tragen in etwa gleich stark zum Attraktivitätsurteil bei.

2.4.1.5 System Usability Scale (SUS)

Neben Attraktivität und Akzeptanz des Systems ist die Bedienbarkeit (usability) des Systems von Bedeutung. Zur Erfassung der Bedienbarkeit des getesteten Systems wurde die sog. System Usability Scale (SUS) verwendet, die eine einfache Zehnpunkt-Skala im Stile einer Likert-Skala darstellt, und zum Erfassen eines globalen, von den Benutzern subjektiv erlebten Eindruckes der Bedienbarkeit dient [48]. Im Rahmen der Arbeit wurde die ins Deutsche übersetzte Version des SUS Fragebogens eingesetzt.

2.4.1.6 Subsidiary Task Paradigma

Zum Testen der (adaptiven) Systeme muss das Subsidiary Task Paradigma verwendet werden, da die Primäraufgabe (das Autofahren) aus sicherheitsrelevanten Aspekten priorisiert werden muss. Die Änderung der mentalen Beanspruchung lässt sich beim Subsidiary Task Paradigma aus den Sekundäraufgabendaten ablesen. Da die Versuchspersonen jedoch – trotz strikter Anweisung – dieses Paradigma nicht immer genau befolgen, wirken sich die Sekundäraufgaben beim Autofahren auch auf die Primärleistung aus, weshalb die Änderung der mentalen Beanspruchung aus den Leistungsdaten der Primär- und Sekundäraufgabe interpretiert werden muss. Die strenge Einhaltung des Subsidiary Task Paradigma ist also nicht ganz gegeben. Eine Abgrenzung zum Loading Task Paradigma ist jedoch durchaus noch vorhanden.

Zur Auswertung der Konzepte erfolgt die Aufzeichnung folgender Daten: Fahrleistungsdaten, objektive mentale Beanspruchung, subjektiv erlebte mentale Beanspruchung

und Sekundäraufgabendaten. Die Fahrleistungsdaten (Spurabweichung, etc.) werden von der Fahr Simulator-Software automatisch erfasst. Zur objektiven Messung des Mental Workload wurde der PDT verwendet (vgl. Kapitel 2.4.1.2). Die Messung der subjektiv erlebten mentalen Beanspruchung wurde mittels SEA-Skala (vgl. Kapitel 2.4.1.3) vorgenommen. Die Sekundäraufgabendaten werden entweder vom Versuchsleiter manuell protokolliert oder direkt in einer Logdatei der Menüsoftware aufgezeichnet.

2.5 Blickbasierte Benutzerschnittstellen

Der Bereich der blickbasierten Mensch-Maschine-Kommunikation ist ein relativ junges, stetig wachsendes Forschungsfeld. In Europa beschäftigen sich die Partner des COGAIN Konsortiums mit der Erforschung blickbedienter Schnittstellen [1]. Dabei liegt der Fokus der Forschungsarbeiten darauf, motorisch eingeschränkten Nutzern den Zugang zu Computersystemen zu erleichtern. Ein weiterer Schwerpunkt der internationalen Forschungsarbeiten in diesem Umfeld stellen blickbasierte Schnittstellen im Desktop Bereich dar. In diesem Umfeld existieren hauptsächlich zwei Forschungsschwerpunkte: blickbasierte Mensch-Maschine Interaktion (z.B. Eye Typing und einfache Comand & Control Anwendungen) und Attentive User Interfaces (AUI). *Sibert* und *Jacob* fassen das Potential blickbasierter Benutzerschnittstellen folgendermaßen treffend zusammen:

„When eye gaze interaction is working well, the system can feel as though it is anticipating the user’s commands, almost as if it were reading the user’s mind.“ [218]

Aus dem Forschungsfeld der Mensch-Maschine-Kommunikation im Kraftfahrzeug sind bisher keine Veröffentlichungen zur blickbasierten MMK bekannt.

2.5.1 Blickbedienung

Blick als Interaktionsmodalität ist ein interessanter Ansatz im Forschungsbereich der Mensch-Maschine Interaktion. Zunächst war diese Bedienart vor allem ein vielversprechender Ansatz, um motorisch eingeschränkten Personen einen komfortablen Zugang zu Computersystemen zu ermöglichen. Dabei dient die Blickrichtung als Musersatz zur Interaktion [59, 141]. Weiterhin existiert eine Vielzahl an Gründen, weshalb Blickbedienung als erweiterte Interaktionsmöglichkeit in Betracht zu ziehen ist [241]:

1. Werden die Hände anderweitig eingesetzt bzw. stehen diese aufgrund motorischer Einschränkungen nicht für die Interaktion zur Verfügung, so stellt der Blick einen zusätzlichen leistungsfähigen Interaktionskanal dar.
2. Die Muskulatur zur Ausrichtung der Augen ermöglicht die schnellsten Bewegungen eines menschlichen Körperteils (vgl. Kapitel 3.1.3). Dies impliziert, dass Blickbedienung eine der schnellsten Eingabemöglichkeiten darstellt.
3. Untersuchungen von *Land* haben gezeigt, dass Menschen dazu tendieren, Zielobjekte zunächst mit den Augen zu selektieren, bevor sie manuelle Aktionen ausführen [135, 136]. Durch umfassende Analyse der Augenbewegungen lassen sich somit zukünftige Aktionen des Nutzers vorhersagen.

4. Menschen können tausende Augenbewegungen vollziehen, ohne davon merkbar zu ermüden. Der Einsatz von Blickbedienung vermindert zudem die Notwendigkeit von sich wiederholenden manuellen Handlungen und mindert somit das Risiko einer Verletzung durch wiederholte Belastung.
5. Der Umgang mit dem Blick muss nicht erlernt werden, vielmehr nutzt der Mensch ihn schon über das ganze Leben hinweg in einer intuitiven Weise. Bei der Interaktion in einer Gruppe richten sich z.B. gesprochene Worte an die angeschaute Person.

Der Einsatz von Blickbedienung unterliegt jedoch gewissen Einschränkungen. Das Auge ist in erster Linie ein Rezeptor. Evolutionsbedingt dient das Auge dem Menschen zur Wahrnehmung seiner Umgebung und er ist es deshalb nicht gewohnt, das Auge als Aktor zur direkten Manipulation/Steuerung einzusetzen. Zu den größten Einschränkungen bzw. Problemen zählt allerdings die Ungenauigkeit bei der Blickrichtungsbestimmung, die u.a. im physiologischen Aufbau des Auges begründet ist: Der Bereich des schärfsten Sehens (Fovea Centralis) des menschlichen Auges beträgt etwa zwei Grad. Deshalb unterliegt die (Re-)Fixationsgenauigkeit des Menschen und somit auch die Bestimmung des Blickpunktes einer Ungenauigkeit (vgl. Kapitel 3.1). Weiterhin gilt zu beachten, dass das Auge zu keinem Zeitpunkt still steht (vgl. Kapitel 3.1.3). Um ein scharfes Gesamtbild zu erhalten führt das Auge kleine Zuckbewegungen durch, um das Abbild der Umgebung auf unterschiedliche Rezeptorbereiche abzubilden. Im Vergleich zur „langsamen und bedachten“ Interaktion mit Computermäusen oder anderen manuellen Eingabegeräten, bewegt sich der Blick sehr schnell über die Anzeige. Dies lässt die Annahme zu, dass sich Blickbedienung nicht zur Steuerung kontinuierlicher Mausbewegungen eignet, sondern eher zur Auswahl diskreter Objekte (z.B. Buttons). Auch wenn der Benutzer davon ausgeht, dass er ein Objekt fixiert, bewegt sich das Auge kontinuierlich unbewusst um das Objekt herum. Dies fordert eine Dialoggestaltung, die solche Fixationsbewegungen kompensieren kann (z.B. durch ausreichend dimensionierte Schaltfläche).

Ein weiteres bekanntes Problem im Bereich der Blickbedienung ist das sog. *Midas Touch Problem*. Dieses von *Jacob* titulierte Problem tritt auf, wenn das Auge zusätzlich zu sensorischen Zwecken zur Bedienung verwendet wird [117]. Im Normalfall geht ein Benutzer davon aus, Objekte betrachten zu können, ohne direkt eine Aktion auszulösen. Die Anwendung darf also nicht jedes Mal reagieren, wenn sich das Ziel der Blickrichtung ändert, sondern nur in „angemessenen“ Situationen und zur „richtigen“ Zeit. Wenn der Anwender ein Objekt lediglich betrachtet, sollten versehentliche Aktionen bzw. Selektionen vermieden werden. Die Selektion eines Objektes erfolgt nur dann, wenn der Nutzer dies auch wirklich beabsichtigt. Im Idealfall reagiert das Interface also nur dann, wenn der Nutzer es will; andernfalls sollte die Möglichkeit bestehen, den Blick ohne weitere Reaktionen zu bewegen [117]. Menschliche Partner können dies sehr gut erkennen, für Computersysteme stellt es jedoch eine große Herausforderung dar. Somit stellt sich die Frage, wann ein mit den Augen fixiertes Objekt als selektiert zu betrachten ist [175]. Kapitel 2.5.3 stellt exemplarisch Ansätze zur Minimierung dieser Probleme vor.

2.5.2 Attentive User Interfaces

Adaptive Interfaces sind eine vielversprechende Lösung, um einer wachsenden Systemkomplexität zu begegnen. Sie sind derart gestaltet, dass sie einerseits den individuellen

Bedürfnissen der Nutzer gerecht werden, aber auch die Einflüsse der jeweiligen Umgebung und Situation berücksichtigen [62]. Anwendungen, die die visuelle Aufmerksamkeit eines Benutzers auswerten und sich daran anpassen, sind in der Literatur als Attentive User Interfaces (AUI) oder Attention Aware Systems (AAS) bekannt. Sie zeichnen sich dadurch aus, dass sie das Benutzerverhalten überwachen, Ziele/Interessen des Nutzers modellieren und seine Bedarfe vorhersagen. Auf Basis dieser Eigenschaften interagieren sie mit dem Nutzer und stellen ihm Informationen bereit, indem sie anhand der Analyse der visuellen Aufmerksamkeit auf die Intention des Anwenders schließen [114, 146]. Übertragen auf den Einsatz im Kfz bedeutet dies, dass AUI den Zustand des Fahrers (z.B. Aufmerksamkeit), dessen Aktivitäten (z.B. Lenk- oder Bremsmanöver) und externe Einflüsse (z.B. Verkehrsaufkommen) berücksichtigen müssen. In diesem Zusammenhang gewinnt der Begriff *Holistic Sensing* mehr und mehr an Bedeutung [235].

Attentive User Interfaces priorisieren Informationen dynamisch, so dass die Informationsverarbeitung des Nutzers und des Systems optimal auf die jeweiligen Aufgaben ausgelegt sind [239]. Sie greifen auf eine Vielzahl fortgeschrittener Eingabe-, Interpretations- und Ausgabetechniken zurück, um z.B. zu erkennen, mit welcher Aufgabe ein Nutzer beschäftigt ist [240]. [239] unterscheidet grob zwischen zwei verschiedenen Arten von AUIs: explizite und implizite. Implizite AUI schließen ohne „sensorische“ Messung auf die Aufmerksamkeit des Nutzers. Der Nutzer kann dem System jedoch seinen Aufmerksamkeitsstatus mitteilen (z.B. durch spezielles Anordnen der verwendeten Fenster). Explizite AUI verwenden zur Aufmerksamkeitsbestimmung dedizierte Sensoren (z.B. ein Eye Tracking System). Im Rahmen der Arbeit sollen explizite AUI betrachtet werden.

In Kombination mit Eye Trackern existiert eine Vielzahl potentieller Anwendungsfelder für AUI, wobei es nicht immer nötig ist, die exakte Blickrichtung zu kennen, da auch simple Sensoren (z.B. *EyeContact*) erkennen können, ob die Aufmerksamkeit des Nutzers auf sie gerichtet ist, ohne eine exakte Rekonstruktion durchzuführen. Ein Beispiel dafür stellt *Kembels iLights* dar [125]. Eine genauere Bestimmung des Aufmerksamkeitsfokus ist durch explizite Messung der Blickrichtung möglich. *iDict* nutzt dieses Vorgehen zur Unterstützung des Nutzers beim Lesen fremdsprachlicher elektronischer Texte [113]. Dazu erfolgt eine kontinuierliche Analyse der Blickrichtung während des Lesens. Weicht diese vom typischen Verhalten ab, schließt das System darauf zurück, dass der Leser Probleme beim Verständnis eines Wortes hat. Das System bietet dem Leser dann eine Hilfestellung an, indem es eine Übersetzung in der Muttersprache – direkt über dem problematischen Wort platziert – zur Verfügung stellt. Genügt diese Übersetzung nicht zum Verständnis des Textes, kann er sich durch Fixation des rechten Bildschirmbereichs den gesamten Wörterbucheintrag anzeigen lassen. Neben *iDict* existiert eine Vielzahl weiterer interessanter Ansätze im vergleichsweise jungen Forschungsfeld der AUIs [114].

2.5.3 Gestaltungshinweise für blickbasierte Benutzerschnittstellen

Wie in den vorangegangenen beiden Kapiteln bereits erwähnt, müssen bei der Umsetzung blickbasierter Benutzerschnittstellen gewisse Rahmenbedingungen zur Kompensation konzeptbedingter Nachteile beachtet werden. Um das *Midas Touch* Problem zu umgehen, bieten sich multimodale Interaktionsstile an. Darunter versteht man die Verwendung verschiedener Interaktionsstile, die eine möglichst natürliche und intuitive Bedienung

ermöglichen [209]. Bezogen auf das Midas Touch Problem bedeutet dies, dass eine Aktion erst durch eine zusätzliche Tastatur- oder Mauseingabe, ein akustisches Kommando oder eine Geste aktiviert wird [220]. Ist die Blickrichtung die einzig verfügbare Modalität, sollte eine *dwell-time* eingesetzt werden. Dies bewirkt, dass ein Objekt zunächst für eine gewisse Zeit fixiert werden muss, um eine Aktion auszulösen. Dadurch ist sichergestellt, dass die Anwendung Befehle nur dann aktiviert, wenn es der Anwender auch wirklich beabsichtigt und er die Schaltfläche länger fixiert. Es hat sich allerdings als schwierig erwiesen, die Dauer der Zeitschwelle angemessen zu wählen [237]: Wird sie zu kurz gewählt, so sind hohe Fehlerraten die Folge, wird eine zu große Zeitschwelle gewählt, so vermindert sich der Geschwindigkeitsvorteil des Eye Tracking Systems und reduziert die Akzeptanz beim Benutzer. In der Literatur konnten bisher keine verwertbaren Ergebnisse zur Wahl der *dwell-time* im Fahrzeugkontext gefunden werden. Jedoch kann man davon ausgehen, dass für dieses sicherheitskritischere Anwendungsszenario der Wahl dieses Parameters eine höhere Priorität beizumessen ist als z.B. bei Desktopanwendungen, für die sich eine *dwell-time* im Bereich von 300 bis 500 ms als optimal erwiesen hat.

Im Fahrzeugkontext stellt die Verwendung einer zusätzlichen Taste zur Bestätigung gewünschter Eingaben einen vielversprechenden Lösungsansatz dar. Dies ähnelt der Bedienung mit einer Computermaus – Zeigen auf das gewünschte Objekt und Bestätigen mit einer Taste – und ist deswegen leicht und schnell erlernbar. Dabei scheint eine ergonomische Platzierung dieser Taste am Lenkrad am naheliegendsten, um eine schnelle, sichere und bequeme Bedienung zu ermöglichen. Diese Taste ist leicht mit dem Daumen bedienbar und der Fahrer muss zu keiner Zeit die Hand vom Lenkrad nehmen oder die Taste im Cockpitbereich ausfindig machen. Zudem konnte *Vertegaal* zeigen, dass diese Kombination zu geringstmöglichen Fehlerraten führt, was sich u.a. in kürzeren Bedienzeiten widerspiegelt [241].

Zur Kompensation von Ungenauigkeiten existieren unterschiedliche Lösungsansätze. Durch eine Lupe [161] oder eine Fischaugenlinse [18] wird der Bereich um den Blickpunkt vergrößert dargestellt. Die Vergrößerung findet nur bei einer detektierten Fixation statt. Innerhalb des vergrößerten Bereichs kann eine genauere Selektion durchgeführt werden. Das Konzept *Star Gazer* setzt auf eine kontinuierliche Vergrößerung des Bereichs um den Blickpunkt bis dieser bildschirmfüllend dargestellt ist [98]. Ist eine parametrierbare Vergrößerung erreicht, wird das betrachtete Objekt selektiert und die Anwendung hinsichtlich des Zoomfaktors wieder in den ursprünglichen Zustand versetzt.

Für blickbediente Menüs ergeben sich verschiedene Anforderungen an die Gestaltung der graphischen Benutzeroberfläche, wobei in erhöhtem Maße Anforderungen an die Auslegung der *Schaltflächen* eines Menüs gestellt werden. Zur einfachen Fixation benötigt das Gehirn einen Reiz, weshalb die Schaltfläche ein Element enthalten sollte, das die Aufmerksamkeit auf sich zieht. Studien bestätigten, dass z.B. ein komplexeres Icon eine enorme Hilfe bei der Fokussierung der Schaltfläche darstellt. Auf diese Weise „kreist“ das Auge um das Icon, um die Details wahrnehmen zu können. Gleichzeitig darf die Gestaltung der Schaltflächen im vorliegenden Anwendungsfall keine zu große Ablenkung des Fahrers vom Fahren hervorrufen. Die Anforderungen an die *Größe einer Schaltfläche* sind in erster Linie von der möglichen Auflösung und Genauigkeit des verwendeten Gaze Trackers abhängig.

Ein aussagekräftiges *Feedback* zum Systemzustand ist vor allem bei berührungslosen Interaktionskonzepten unabdingbar. Dabei gilt es, die aktuell selektierte bzw. fixierte Schaltfläche vor den restlichen hervorzuheben. Dazu hat sich die Verwendung eines Cursors (zur Visualisierung des rekonstruierten Blickpunktes) als ablenkend und somit unvorteilhaft herausgestellt, als vorteilhaft hingegen die Hervorhebung der Schaltfläche durch einen Farbwechsel. Dabei wird die Schaltfläche, solange der Blickpunkt deren Konturen nicht verlässt, in einer anderen Farbe dargestellt, wodurch beim Fahrer der Eindruck entsteht, dass ihn das System „versteht“, und nicht, dass er den Cursor mit gezielten Augenbewegungen in eine bestimmte Richtungen bewegen muss.

Animierte Informationsdarstellungen können bei blickbedienten Menüs einen negativen Einfluss haben, da Bewegungen die Aufmerksamkeit und damit den Blick des Fahrers auf sich ziehen. Erlaubt man beispielsweise während einer Eingabe Animationen am Rand des Displays werden diese die Fixation der relevanten Schaltflächen stören, indem sie den Blick auf sich ziehen, was zur Unterbrechung der Eingabe führt [80].

Vor allem bei der Erstellung von Interaktionskonzepten zur Nutzung im Kraftfahrzeug gilt es, der *Positionierung* der Displays, der Schaltflächen auf den Displays sowie der Kamera besondere Beachtung zu schenken. Die Kamera sollte dabei den Fahrer immer möglichst frontal erfassen und wichtige Schaltflächen möglichst zentral auf dem Display angeordnet sein. Je kleiner der Winkel zwischen Kamera und Nutzer bzw. zwischen Blickgerade und Display ausfällt, desto fehleranfälliger ist zumeist das Gaze Tracking. Deshalb sollten Schaltflächen am Rand von (größeren) Displays möglichst vermieden und andernfalls am Rand größer als in der Mitte dargestellt werden.

Zur Bündelung von Funktionen in Menüs existieren unterschiedliche Organisationsstrukturen. Das lineare Menü (Pull-Down-Menü, vgl. Abbildung 2.12(a)) ist eine der bekanntesten Formen hierarchischer Menüstrukturen und optimal für Mauseingabekonzepte geeignet. Es besteht aus einem Hauptmenü (oben horizontal) und mehreren hierarchisch aufgebauten Untermenüs. Aufgrund der Struktur und der Darstellung auf dem Monitor sind die Schaltflächen klein und daher nur schwer bedienbar [220]. Außerdem nimmt ihre Größe mit steigender Anzahl der Elemente und Untermenüs linear ab. Weiterhin sind kleine bzw. fehlende Abstände zwischen einzelnen Schaltflächen problematisch. Deshalb verlangt eine fehlerfreie Menünavigation nach exakten Bewegungen. Die Fehlerquote bei der gezielten Auswahl eines Untermenüs liegt bereits bei Mauseingabe sehr hoch, bei Blickbedienung kann man aufgrund des sprunghaften Charakters der Augenbewegungen von signifikant höheren Fehlerquoten ausgehen. Die Unterbrechbarkeit der Eingabe stellt ein Problem dar, da der erreichte Zwischenzustand verloren geht, sobald der Nutzer seinen Blick vom Display abwendet. Aufgrund der mangelnden Unterbrechbarkeit sowie der fehleranfälligen Navigation durch die aufklappenden Hierarchieebenen verstärken Pull-Down-Menüs den Tunnelblick des Anwenders [128]. Ein weiterer Nachteil von Pull-Down-Menüs ist es, dass sie für die Darstellung von Menüpunkten in Textform und nicht für die Darstellung intuitiver Symboliken optimiert sind. Somit bleibt festzuhalten, dass blickbediente Pull-Down-Menüs weniger geeignet sind für den Einsatz im Kraftfahrzeug.

Kreis-Menüs (Pie Menüs, vgl. Abbildung 2.12(b)) sind hinsichtlich Schaltflächengröße und Abstand zwischen den einzelnen Schaltflächen eine gute Alternative zu Pull-Down-Menüs. *Urbina* konnte die Eignung von Pie Menüs für blickbediente Texteingabe anhand

2. Grundlagen und Thematische Einführung

von sehr guten Ergebnissen in mehreren Studien zeigen (z.B. [237]). Die Schaltflächen sind beginnend in der Mitte mit zum Rand hin größer werdenden Schaltflächen über das gesamte Display verteilt. Untersuchungen zeigten, dass dieses Layout für blickbasierte Anwendungen von Vorteil ist, da sich Kalibrierungsfehler auf die Ränder des Bildschirms stärker auswirken als auf den Bildschirmmittelpunkt [30]. Pie Menüs bieten jedoch keinen adäquaten Lösungsansatz zu den Problemen der Unterbrechbarkeit sowie zur Kompensation des „sprunghaften“ Charakters der Augenbewegungen. Die Integration zusätzlicher intuitiver Symboliken gestaltet sich aufgrund des Platzbedarfes schwierig.

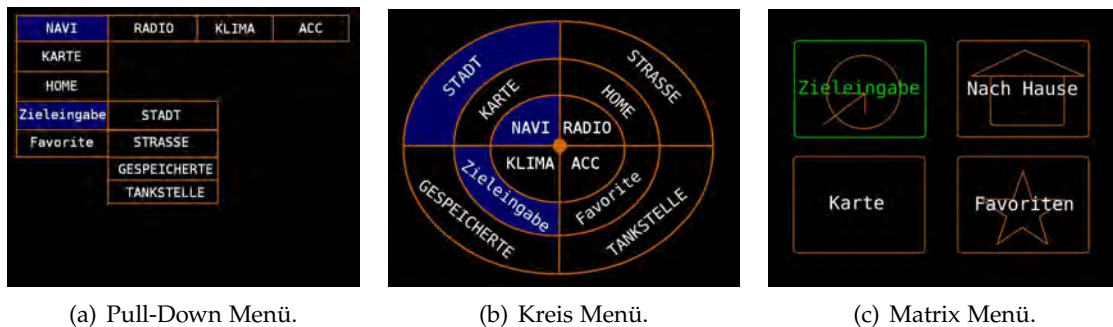


Abbildung 2.12: Verschiedene Organisationsstrukturen für graphische Benutzerschnittstellen.

Eine bestmögliche Nutzung der verfügbaren Anzeigefläche ermöglicht die Darstellung in Form einer Matrix (vgl. Abbildung 2.12(c)), welche größtmögliche Schaltflächen und Abstände zwischen den einzelnen Schaltflächen erlaubt, was eine benutzerfreundliche und intuitive Interaktion ermöglicht. Dabei ist anzumerken, dass die Übersichtlichkeit des Menüs leidet, sobald (abhängig von der Displaygröße) eine gewisse Anzahl gleichzeitig dargestellter Elemente überschritten wird. Menüstrukturen werden durch Ausblenden der oberen Hierarchieebene und Einblenden des Untermenüs angezeigt. Somit steht für Untermenüs wiederum die gesamte Displayfläche zur Verfügung, was ebenfalls das Problem der Unterbrechbarkeit der Interaktion löst. Wendet der Fahrer seinen Blick auf die Straße, so verbleibt das Menü im angezeigten Zustand und der Nutzer kann seine Eingabe jederzeit in diesem Zustand fortsetzen. Zusätzlich eignen sich Matrix Menüs für eine symbolische Informationsrepräsentation. Die Matrix Darstellung stellt somit den besten Ausgangspunkt für die Konzeption blickbasierter Bedienkonzepte im Automobil dar.

Kamera-basierte Fahrerüberwachung

Kapitel 3 beinhaltet eine thematische Einführung in das Forschungsgebiet der Blickrichtungsbestimmung und vermittelt das Basiswissen zum Verständnis der Begriffe, die in den darauf folgenden Kapiteln Verwendung finden.

Das Auge ist das bedeutendsten Sinnesorgan des Menschen und über den visuellen Sinneskanal werden ca. 87% der Informationen aufgenommen [209]. Somit sind Augenbewegungen ein essentieller Bestandteil des menschlichen Explorationsverhaltens. Die Augen werden so ausgerichtet, dass der Mensch einen gewissen Ausschnitt des Gesichtsfeldes in hoher Auflösung wahrnimmt und somit in der Lage ist, detailliert zu betrachten, was sich im zentralen Sichtbereich befindet. Dabei wird oftmals die gesamte (örtliche) Aufmerksamkeit auf ein bestimmtes Objekt bzw. eine bestimmte Region gerichtet.

Zudem kann man davon ausgehen, dass man den Verlauf der Aufmerksamkeit analysieren kann, wenn man die Augenbewegungen eines Menschen aufzeichnet und auswertet. Das Wissen zur Verteilung der (bewussten) Aufmerksamkeit wird verwendet, um höherwertige Informationen zu kognitiven Prozessen abzuleiten, welche anderweitig schwer bzw. überhaupt nicht zugänglich wären. Somit können aufgrund der Kenntnis über den Verlauf Blickes eines Menschen zukünftige Handlungen vorhergesagt werden.

Um diese Vorgänge zu veranschaulichen, sei auf die Untersuchungen von Land verwiesen, die am Beispiel des Verlaufs des Blicks beim Händewaschen oder Teekochen zeigen, dass die Aufmerksamkeit bzw. der Blick eines Menschen bei alltäglichen Handlungen der eigentlichen Aktion um teilweise mehrere Sekunden vorausseilt (engl. look-ahead fixations) [135, 136]. Der Mensch verschafft sich zunächst einen Überblick der Umgebung, lokalisiert zur Ausführung der geplanten Aktion nötige Gegenstände, plant seine Wege durch die Umgebung und führt (geplante) Handlungen kontrolliert durch.

Zur automatischen Bestimmung des Blickzieles oder Blickpunktes (engl. point-of-gaze, POG) werden Eye Tracking Systeme eingesetzt. Die Echtzeitfähigkeit moderner Eye Tracker erweitert deren Einsatzgebiet von rein diagnostischen Zwecken (offline Auswertung der aufgezeichneten Daten) hin zur Kontrolle von (Computer-) Systemen.

Im folgenden Kapitel 3.1 werden Grundbegriffe zum Thema Eye Tracking, dessen geschichtliche Entwicklung sowie unterschiedliche Ansätze zur Registrierung der Blickrichtung vorgestellt. Aufbauend auf diesem Überblick sowie den physiologischen Eigenschaften des menschlichen Auges, wird der entwickelte Ansatz zur berührungsl-

sen Schätzung der Blickrichtung sowie das entworfene Framework beschrieben und im Weiteren Ergebnisse zur Akkuratheit sowie Performance des Gesamtsystems diskutiert.

3.1 Grundbegriffe und geschichtliche Entwicklung

Das folgende Kapitel gibt einen Überblick unterschiedlicher Einsatzgebiete von Eye Tracking Systemen, gefolgt von physiologischen und mathematischen Grundlagen zur Modellierung des menschlichen Auges. Anschließend folgt ein Abriss der geschichtlichen Entwicklung der Blickrichtungsregistrierung von anfänglich simplen mechanischen Apparaturen hin zu modernen, berührungslosen Eye Tracking Systemen. Ausgehend vom Vergleich und der Bewertung der verschiedenen Methoden befasst sich der letzte Abschnitt dieses Kapitels mit der Videookulographie, die besonders erfolgversprechend und für die Mensch-Maschine Interaktion von herausragender Bedeutung ist.

3.1.1 Anwendungsfelder

Systeme zur Blickregistrierung werden heute in unterschiedlichen Bereichen genutzt. Hierbei werden zwei wesentliche Bereiche unterschieden:

- Analytische Einsatzgebiete: der Eye Tracker zeichnet lediglich Augenbewegungen auf bzw. analysiert diese (z.B. im medizinischen Umfeld)
- Mensch-Maschine Interaktion: Eye Tracker werden eingesetzt, um mit Hilfe von Blickbewegungen einen Computer – bzw. eine Maschine – zu bedienen

Erste Verfahren zur Bestimmung der Blickrichtung stammen aus dem medizinischen Umfeld, wo sie z.B. im Bereich der Schlafforschung oder der Gleichgewichtsuntersuchung zu diagnostischen Zwecken eingesetzt werden. Dabei handelt es sich eher um Eye als um Gaze Tracker¹, da eher die qualitative Analyse der Augenbewegung von Interesse ist.

Im Bereich der Werbung und des Marketings werden Eye Tracker explizit zur Bestimmung der Blickrichtung eingesetzt. Hier steht die Frage nach dem anvisierten Objekt im Vordergrund. Typische Einsatzgebiete sind Productplacement und Usability Studien. Die aufgezeichneten Blickverteilungen lassen Rückschlüsse auf die Bedienbarkeit von Websites, auf die optimale Platzierung von Produkten im Supermarkt oder von Werbebannern zu.

Weiterhin werden Gaze Tracker bereits erfolgreich im Bereich der Mensch-Maschine Interaktion eingesetzt. Zur Eingabe wurden Gaze Tracker hauptsächlich für Menschen mit motorischen Einschränkungen eingesetzt. Mit spezieller Software zur Texteingabe [99, 147, 236] oder zur Steuerung der Maus [138] sind sogar nahezu vollständig gelähmte Personen in der Lage, einen Computer zu bedienen und z.B. Bücher zu „schreiben“ [141].

3.1.2 Physiologie des menschlichen Auges

Abbildung 3.1 zeigt zwei Aufnahmen des Auges, bei dem es sich um ein sog. „Linsenauge“ handelt [214]. Das linke Bild zeigt die äußeren Bestandteile des Auges – die weiße Sklera

¹Beide Bezeichnungen werden heute teilweise synonym verwendet, auch wenn im eigentlichen Sinne ein Eye Tracker lediglich Daten bezüglich der Augenbewegung zur Verfügung stellt, wohingegen ein Gaze Tracker explizite Blickrichtungsdaten (Blickpunkt auf einem Monitor bzw. in 3D) berechnet.

3.1. Grundbegriffe und geschichtliche Entwicklung

bzw. Lederhaut, die farbige Iris bzw. Regenbogenhaut und die schwarze Pupille. Ein weiterer impliziter Bestandteil ist der Übergang von Sklera zu Iris: der Limbus. Im rechten Bild ist eine Auswölbung über der Iris zu erkennen – die Kornea bzw. die Hornhaut (vgl. Abbildung 3.1 bzw. Abbildung 3.2).

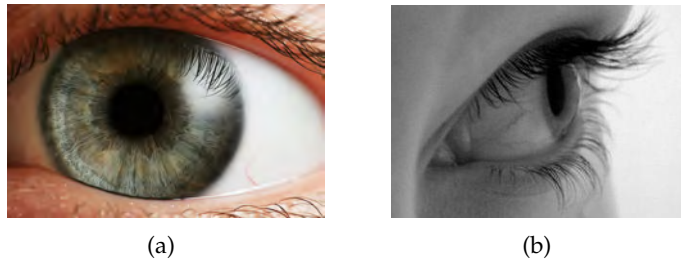


Abbildung 3.1: Abbildungen des menschlichen Auges, entnommen aus [252].

Folgende physiologische Parameter des menschlichen Auges sind für die anschließende mathematische Modellierung des Auges (siehe Kapitel 3.1.2.1) von essentieller Bedeutung: Der Augapfel (Bulbus Oculi) ist ein kugelförmiger Körper. Sein Durchmesser beträgt im Populationsmittel ca. 24 mm [94]. Der Durchmesser der Hornhaut beträgt frontal betrachtet 11 bis 13 mm und ist horizontal 1 mm größer als vertikal. Im Zentrum (Durchmesser ca. 4 mm) kann die Krümmung der Hornhaut annähernd als sphärisch und einheitlich betrachtet werden, was einem Öffnungswinkel von $\pm 12,5^\circ$ entspricht (Annahme: Krümmungsradius von 7,7 mm). Im Randbereich ändert sich der Radius der Hornhaut auf 8 mm, bis sie in die Lederhaut mit 13mm Radius übergeht [132].

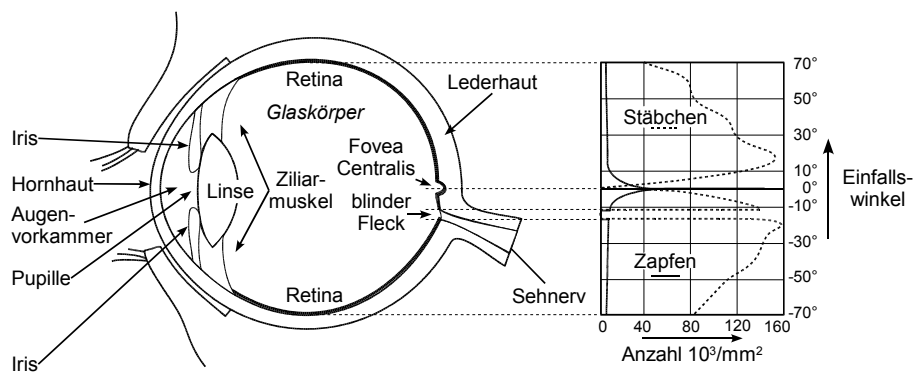


Abbildung 3.2: Schematische Darstellung des menschlichen Auges. Links: Querschnitt durch das menschliche Auge. Rechts: Verteilung der Rezeptoren (Zapfen und Stäbchen) auf der Retina in Abhängigkeit des Einfallswinkels des Lichts. Entnommen aus [209].

Abbildung 3.2 zeigt die drei Schichten, in die der Aufbau des Auges unterteilt ist: die äußere Augenhaut schützt das Auge und besteht aus Lederhaut und Hornhaut; die Lederhaut besteht aus weißem, festem Bindegewebe und geht in die stark gewölbte, durchsichtige Hornhaut über. Die zweite Schicht besteht aus Aderhaut, Iris und Ziliarkörper; das durch die Gefäße der Aderhaut fließende Blut versorgt die Netzhaut mit Nährstoffen und Sauerstoff. Die innerste Augenhaut ist die Netzhaut (Retina), auf der der

3. Kamera-basierte Fahrerüberwachung

Gelbe Fleck und der Blinde Fleck zu finden sind. Im Blinden Fleck ist der Sehnerv mit der Netzhaut verbunden und es sind keine Sehzellen vorhanden. In der Mitte der Retina befindet sich die Macula, deren Zentrum die Zone des schärfsten Sehens darstellt (Fovea Centralis) und als Gelber Fleck bezeichnet wird. Im Inneren des Auges befinden sich weiterhin der Glaskörper, die Linse und die beiden Augenkammern, die zusammen mit der Hornhaut für die Brechung des Lichts verantwortlich sind. Die Linse wird über die Zonulafasern gehalten und bei Bedarf gekrümmt, damit auf der Netzhaut ein scharfes Abbild des betrachteten Objektes entsteht.

Das von einem Objekt reflektierte Licht passiert dabei zunächst die Hornhaut und die mit Kammerwasser gefüllte vordere Augenkammer. Dabei werden die Lichtstrahlen gebrochen ($Brechungsindex_{Kammerwasser} \gg Brechnugsindex_{Luft}$). Diese Brechung kann mit einer Brechung an einer sphärischen Oberfläche verglichen werden. Die Grenzschicht Kornea-Kammerwasser wirkt als Sammellinse². Den Hauptteil der abbildenden Funktion übernimmt also die gekrümmte Kornea, die Linse hinter der Iris trägt dann jeweils die exakte Brechkraftanpassung zur genauen Fokussierung des Bildes auf der Netzhaut bei.

Nach der vorderen Augenkammer passiert das Licht die Iris, die vereinfacht als kreisförmige Aperturblende angesehen werden kann. Deren Öffnung ist variabel und wird als Pupille bezeichnet. Sie reguliert die Bestrahlungsstärke der Netzhautoberfläche. Unmittelbar hinter der Pupillenöffnung ist die in ihrem Krümmungsradius variable Linse an den Zonulafasern aufgehängt. Je nach Krümmungsradius der Linse, welcher durch Kontraktion des Ziliarmuskels verändert wird, bricht sich der Lichtstrahl an der Linsenoberfläche unterschiedlich stark und trifft anschließend auf die Netzhaut (Retina).

Auf der Retina entsteht ein bezüglich der optischen Achse gespiegeltes Abbild des fokussierten Objekts. Hier befinden sich die für das Sehen nötigen Rezeptoren. Zweihunderttausend der insgesamt 7 Millionen Zapfen, die für das Farbsehen (photopisches Sehen) zuständig sind, befinden sich im sog. Gelben Fleck, dessen Fläche weniger als einen Quadratmillimeter beträgt. Der Gelbe Fleck erfasst einen Blickwinkel von ca. 2° und stellt das Zentrum des scharfen Sehens (Fovea Centralis) dar, wobei schärfstes Sehen in einem Winkel von ca. 1° um den fixierten Ort möglich ist. An den Bereich der Fovea Centralis grenzen die Parafovea (bis ca. 10° Blickwinkel) und die Peripherie [21]. Abbildung 3.3 zeigt den Schärfebereich des Auges. Im äußeren Bereich erfolgt lediglich eine Helligkeitswahrnehmung, mit zunehmender Verengung des Sehfeldes nehmen die Farbempfindlichkeit und örtliche Genauigkeit zu.

Durch Lichtreize tauchen die Zapfen in die Pigmentzellen, welche Licht mit einer gewissen Wellenlänge absorbieren, ein und leiten elektrische Impulse über die Ganglienzellen weiter. Die Maxima der einzelnen Pigmente liegen für die Zäpfchen bei 430 nm (blau), 535 nm (grün) und 590 nm (rot). Weitere 130 Millionen Sinneszellen (Stäbchen), sind für das Dämmerungssehen (skotopisches Sehen) verantwortlich. Durch ihre Verteilung auf der Netzhaut ergibt sich ein kegelförmiges Blickfeld von ca. 100° [206]. Da



Abbildung 3.3: Schärfebereiche des menschlichen Auges nach [67].

²Verglichen mit der variablen Augenlinse ist die Brechkraft der Kornea sogar sehr viel größer (nach [164] etwa 40:19).

diese Rezeptoren weniger dicht als die Zapfen im Gelben Fleck angeordnet sind, kann man beim Dämmerungssehen nur ein farbloses, unscharfes Bild wahrnehmen. Über den Sehnerv werden schließlich die von der Retina registrierten optischen Reize zur weiteren Verarbeitung an das Gehirn abgeleitet.

3.1.2.1 Mathematisches Augenmodell

Ausgehend von diesen physiologisch motivierten Betrachtungen soll im Folgenden ein mathematisches Modell zur Beschreibung des menschlichen Auges vorgestellt werden, das zur Rekonstruktion der Blickrichtung mittels eines Eye Trackers Verwendung findet. Abbildung 3.4 zeigt ein vereinfachtes Rechenmodell des menschlichen Auges, das das

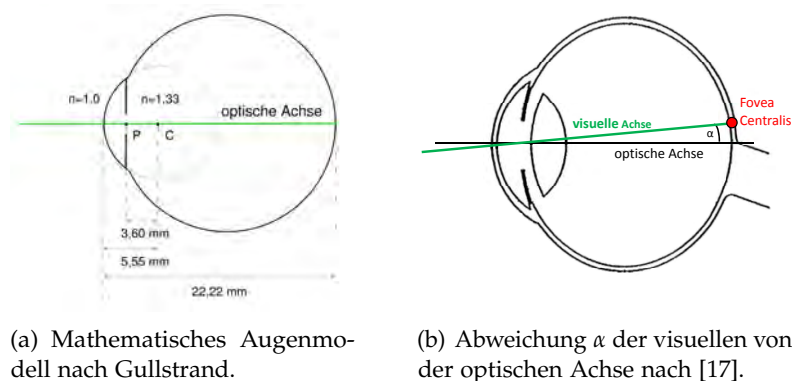


Abbildung 3.4: Mathematisches Augenmodell.

Auge als zwei Kugeln (Augapfel und Kornea) beschreibt. Mathematische Augenmodelle basieren, wenn auch in unterschiedlichen Graden der Komplexität, häufig auf dem schematischen Modell von *Gullstrand* [94]. Dieses Modell liefert als initiale Modellparameter die Populationsmittel der benötigten Augenparameter (vgl. Abbildung 3.4(a) sowie Kapitel 3.1.2). Abweichungen von diesem Populationsmittel werden im Rahmen einer einmaligen Nutzerkalibrierung (siehe Kapitel 3.3.1.4) bestimmt und kompensiert.

Im rudimentären Augenmodell ist weiterhin die optische Achse – als Verbindung zwischen Korneazentrum C und Pupillenzentrum P – eingezeichnet. Die optische Achse entspricht jedoch nicht der eigentlichen bzw. wahrgenommenen Blickrichtung (vgl. Abbildung 3.4(b)), da die wahrgenommene Blickrichtung ausgehend von der Fovea Centralis durch das Pupillenzentrum und schließlich bis zum tatsächlich betrachteten Objekt im Raum verläuft. Die Position der Fovea Centralis unterliegt (starken) interpersonellen Schwankungen und kann bis zu 5° von der optischen Achse abweichen [94]. Diese Abweichung ist für einen Nutzer fest und kann mit der in Kapitel 3.3.1.4 vorgestellten (einmal pro Nutzer durchzuführenden) Kalibrierung bestimmt und kompensiert werden.

3.1.3 Augenbewegungen

Zur scharfen Wahrnehmung eines kontinuierlichen Abbildes der Umgebung sind Augenbewegungen notwendig. Entgegen dem Eindruck, der beispielsweise beim Lesen eines Buches entsteht, sind diese nicht kontinuierlich, sondern in *Fixationen* und *Sakkaden*

unterteilt. Während den Fixationen wird ein Objekt/Punkt fixiert und es erfolgt eine Informationsaufnahme. Da man nur innerhalb des Blickwinkels von 2° scharf sieht (vgl. Kapitel 3.1.2), muss das Auge bewegt werden. Während diesen (sehr schnellen) Sakkaden werden keinerlei Informationen aufgenommen. Weiterhin ist die Aufnahme von Informationen etwa 30-40 ms vor sowie bis zu 120 ms nach dem Start einer Sakkade weitestgehend eingeschränkt. Diese sprunghaften Augenbewegungen ermöglichen einen sehr schnellen Blickrichtungswechsel mit Geschwindigkeiten bis $1000^\circ/\text{s}$. Die Bewegungen werden von sechs quergestreiften Muskeln bewerkstelligt, die das Auge präzise und schnell in einer kreisförmigen Region von 100° in alle Richtungen bewegen können [253]. Die Ausführung von Sakkaden erfolgt sowohl bewusst (z.B. als Reaktion auf Ereignisse im peripheren Gesichtsfeld), als auch unbewusst (z.B. Exploration der Umwelt) [159].

Zur Informationsaufnahme ist das Auge relativ zum betrachteten Objekt in Ruhe zu halten. Dauert dieser Zustand länger als 100 ms an, spricht man von einer Fixation und es ist möglich, Bildinformationen aufzunehmen [124, 255]. In der Regel liegt die Fixationsdauer bei ca. 250 ms [253]. Bei Fixationszeiten von unter 200 ms spricht man von einer Suchfixation und ab 300 ms von einer Verarbeitungsfixation [144]. Weiterhin existieren Folgebewegungen: Beim Betrachten eines bewegten Objektes verfolgt der Blick weiterhin sein Ziel, wobei sich die Blickrichtung relativ zum betrachteten Objekt nicht ändert. Während der Folgebewegungen (smooth pursuit), welche bei einem durchschnittlichen Menschen eine maximale Winkelgeschwindigkeit von $40^\circ/\text{s}$ erreichen [41], können – wie bei der Fixation – Informationen aufgenommen werden. Diese Folgebewegungen werden von Sakkaden unterstützt.

Drift, Tremor (Mikro-Nystagmus) und Mikrosakkaden sind weitere Augenbewegungen mit einer wesentlich geringeren Amplitude, und werden als Mikrobewegungen bezeichnet. Drift und Tremor haben die Aufgabe, das Auge so zu bewegen, dass die Netzhaut um den Bereich von einigen Sehzellen verschoben wird. Bei dieser Verschiebung erhalten die Nervenzellen unterschiedliche Lichtreize, wodurch einer Rezeptorer müdung aufgrund eines konstanten Bildes vorgebeugt wird. Den Mikrosakkaden wird die Aufgabe zugeschrieben, durch die zuvor genannten Bewegungen entstandene Verschiebungen zu korrigieren und die stabile Abbildung auf der Netzhaut zu ermöglichen [27].

3.1.4 Geschichtliche Entwicklung des Eye Tracking

Die Geschichte der Augenbewegungsmessung geht auf den Anfang des 18. Jahrhunderts und die Experimente von *Porterfield* im Jahre 1737 zurück [197]. Eine chronologische Aufzählung präsentiert ausgewählte Meilensteine von besonderer Relevanz für die spätere Forschung im Bereich der Blickrichtungserfassung. Zudem erfolgt eine Diskussion der Vor- und Nachteile einzelner Verfahren. Für eine ausführliche Beschreibung der geschichtlichen Entwicklung sowie der vorgestellten Verfahren sei auf [210] verwiesen.

Sakkaden - 1879 L.E. Javal

L.E. Javal erbrachte 1879 den ersten physikalischen Nachweis für das Auftreten von Sakkaden beim Lesen [120]. Zu diesem Zweck befestigte er ein kleines Glöckchen über eine Schnur am Auge des Probanden. Durch die ruckartigen und schnellen Augenbewegungen klingelte das Glöckchen als Indikator für eine Sakkade.

Hebelnystagmographie - 1891 Ahrens

Ahrens konnte mit der Hebelnystagmographie 1891 erste quantitative Messungen der Augenbewegungen durchführen [13]. Dazu befestigte er einen kleinen Hebel am Auge und konnte damit die Augenbewegungen mittels eines Kymographen auf eine Rußschicht übertragen (vgl. Abbildung 3.5).

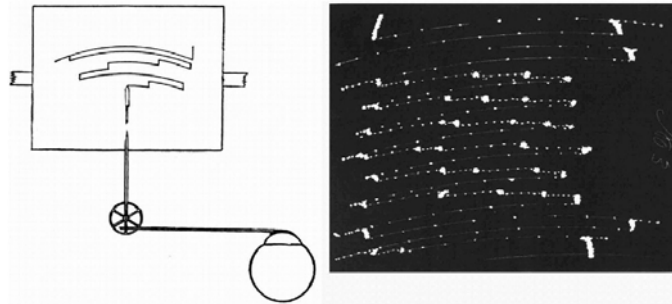


Abbildung 3.5: Prinzip der Hebelnystagmographie. Links: träge Hebelmechanik, die über einen Saugnapf am Auge befestigt ist. Rechts: Aufzeichnung von Augenbewegungen auf dem Kymographen. Entnommen aus [212].

Fotografische Registrierung von Blickbewegungen - 1901 Dodge und Cline

Dodge und Cline entwickelten 1901 die erste fotografische Methode zur Registrierung von Blickbewegungen. Sie zeichneten die Reflexion eines hellen vertikalen Lichtstreifens auf der Hornhaut auf [63]. Dieses Verfahren stellt einen frühen Vorgänger moderner Double-Purkinje Image (DPI) Tracker dar [57].

DPI-Tracker schließen ohne direkten Kontakt zum Auge auf die Blickrichtung. Dazu werten sie Reflexionen an der Hornhaut und der Linse aus. Sie benötigen ein komplexeres Augenmodell als alle anderen Verfahren, was zu präzisen Ergebnissen führt. Abbildung 3.6 zeigt alle Purkinje-Bilder, wobei zur Blickrichtungsbestimmung lediglich das erste und das vierte Bild/Reflexion Verwendung finden. Auflösung sowie die Genauigkeit dieses Verfahrens liegen im Bereich von 0.02° bei einer Abtastrate von 400Hz. Dies ist ausreichend, um die physiologische Bandbreite der Augenbewegungen zu erfassen. DPI-Tracker sind aufgrund des komplexen Aufbaus sehr teuer. Bei Sakkaden zeigt sich ein dynamisches Überschwingen: Führt das Auge eine schnelle Bewegung (Sakkade) aus, so bewegt sich zuerst der Glaskörper und damit die Hornhaut. Erst kurze Zeit später bewegt sich die Augenlinse mit. Dies liegt daran, dass die Linse nicht fest mit der Hornhaut verbunden ist, sondern „federnd eingebettet“, was dazu führt, dass sich die Linse bei einer Sakkade schneller bewegt als „das Auge an sich“ und ebenso nachschwingt, wenn das Auge schon wieder eine feste Position eingenommen hat.

Elektronystagmographie - 1922 Schott

Im Jahre 1922 maß Schott erstmals die Potentiale rund um das menschliche Auge [216]. Er ging dabei allerdings von der fehlerhaften Annahme aus, dass die gemessenen Spannungen von den Aktivitäten der Augenmuskeln herrühren. Diese Annahme wurde 1936 von Mowrer verworfen. Er zeigte auf, dass die unterschiedlichen Spannungen durch die Veränderungen der räumlichen Orientierung des Auges herrühren [165]. Das menschliche

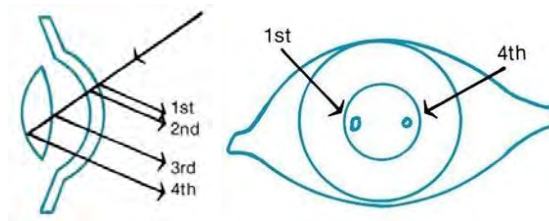


Abbildung 3.6: Schematische Darstellung der Purkinje-Reflexionen, entnommen aus [210].

Auge bildet zwischen Retina (Netzhaut) und der Kornea (Hornhaut) einen Dipol, der sich mit dem Auge mitdreht. Dadurch treten während der Messung Spannungen rund um das Auge auf, welche über Elektroden ausgewertet werden (vgl. Abbildung 3.7). Abhängig von der Ausrichtung des Auges kann man an den Elektroden unterschiedliche Spannungsdifferenzen im μV -Bereich messen. 1939 konnte *Jung* dieses Verfahren auf eine zweidimensionale Messung der Augenbewegungen erweitern [122].

Die ENG ist heute klinische Routine und wird beispielsweise im Bereich der Schlaf-forschung eingesetzt, da eine Messung auch bei geschlossenen Augen möglich ist. Nachteilig an diesem Verfahren ist, dass keine Torsionsmessung möglich ist. Weiterhin entstehen Artefakte bzw. Messfehler durch Aktivitäten von peripheren Gesichtsmuskeln. Spricht der Proband während der Untersuchung, runzelt die Stirn oder zieht die Augenbrauen hoch, hat das verfälschende Auswirkungen auf die gemessenen Spannungen. Deshalb müssen Probanden möglichst ruhig sitzen, was ein intuitives reagieren auf präsentierte Stimuli erschwert, wenn nicht ganz verhindert. Weiterhin verursacht die Dunkel-/Helladaption der lichtempfindlichen Rezeptoren auf der Retina ein Driften der Potentiale und erschwert damit die Auswertung der Messdaten. Der Messbereich beträgt typischerweise $\pm 40^\circ$ horizontal und $\pm 20^\circ$ vertikal, die Auflösung 0.5° und die Genauigkeit $1.5^\circ - 2^\circ$.

Messung des Augenzitterns - 1928 Ohm

Ohm nutzte zur Messung der Augenbewegungen kleine Spiegel, welche er seinen Probanden auf die Augen aufbrachte. Ein Schirm zeichnet das von den Spiegeln reflektierte Licht auf, was die Bestimmung der Blickrichtung in horizontaler und vertikaler Richtung ebenso möglich wie die Bestimmung der torsionellen Drehung [174].

Infrared Reflection Devices - 1951 Torok

Infrared Reflection Devices (IRD) sind eine nicht-invasive Methode zur Registrierung von Augenbewegungen über die Position des Übergangs zwischen Sklera und Iris (Limbus) anhand der Intensität von Infrarotlicht, das vom Auge reflektiert wird [234]. Dazu wird das Auge mit infrarotem Licht beleuchtet. Zusätzlich werden Photodioden vor dem Auge platziert, welche die Intensität des reflektierten Lichts messen. Der Abstand zwischen Auge und Photodiode beträgt dabei ca. 1 cm. Sklera und Iris haben unterschiedliche Reflektionseigenschaften, wodurch eine Bestimmung der Blickrichtung nach erfolgter Kalibrierung durchgeführt werden kann.

Abtastrate (1kHz) und Auflösung ($\leq 0.05^\circ$) sind bei diesem Verfahren sehr gut. Allerdings kann kein numerischer Wert für die Genauigkeit angegeben werden, da bereits eine minimale Relativverschiebung zwischen Messapparatur (Diode) und Auge



Abbildung 3.7: Proband mit Messelektroden für das Elektronystagmographie Verfahren, entnommen aus [210].

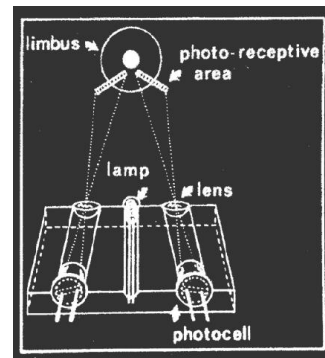


Abbildung 3.8: Schematische Darstellung eines Messaufbaus für das Limbus Tracking. Entnommen aus [210].

von 1mm in einem Fehler von bis zu 5° für die Blickrichtung resultieren kann. Das erfordert eine mehrmals zu wiederholende Kontrolle der Kalibrierung und erschwert die Beherrschbarkeit des Verfahrens. Weiterhin verursachen die Augenlider nicht zu vernachlässigende Artefakte bzw. erschweren eine Auswertung der Messdaten, wenn sie in den Empfindlichkeitsbereich der Diode gelangen.

Search-Coil Methode - 1963 Robinson

Die Search-Coil Methode wurde 1963 von *Robinson* vorgestellt [203] und ist ein höchst invasives Verfahren zur Blickbewegungsregistrierung. Es wird eine Anwendungsdauer von maximal einer halben Stunde empfohlen. Auflösung und Genauigkeit sind bei diesem Verfahren jedoch so gut, dass es als Referenzverfahren im klinischen Umfeld gilt. Zur



(a) Verwendete Kontaktlinse mit elektromagnetischer Spule.



(b) Käfig mit 3-dimensionalen Spulensystem.

Abbildung 3.9: Apparaturen des Search Coil Verfahrens, entnommen aus [203].

Messung wird dem Probanden eine Kontaktlinse mit ein oder zwei elektromagnetischen Spulen eingesetzt (vgl. Abbildung 3.9(a)). Der Kopf des Probanden befindet sich während der Messung innerhalb eines externen Magnetfeldes (vgl. Abbildung 3.9(b)). Dieses induziert in den einzelnen Spulen der Kontaktlinse Spannungen, die zur Blickanalyse ausgewertet werden. Dazu führt ein kleiner Draht von der Kontaktlinse zum eigentlichen Messgerät. Problematisch ist hierbei, dass dieser Draht ebenfalls leitend ist und dadurch eine Offsetspannung induziert.

3. Kamera-basierte Fahrerüberwachung

Zwei Spulen in der Kontaktlinse und drei externe Magnetfelder ermöglichen eine objektive Kalibration. Diese Vorteile machen das Search-Coil Verfahren zur Referenzmethode im Bereich der Medizin. Weiterhin bietet dieses Verfahren sehr gute Eigenschaften hinsichtlich Auflösung (horizontal & vertikal $\leq 0.01^\circ$, torsionell $\leq 0.03^\circ$), Genauigkeit (horizontal 0.458° , vertikal 0.948° , torsionell: 1.628°) und Abtastrate (1kHz). Nachteilig wirkt sich der hohe Grad der Invasivität aus. Es kann zur Austrocknung des Auges und Deformationen der Hornhaut kommen, was zu einer Beeinträchtigung der Sehkraft von bis zu drei Dioptrien führt. Das Gewicht der Kontaktlinse führt weiterhin zu einer Trägheitszunahme (Sakkaden bis zu 5% langsamer und bis zu 8% länger).

Videookulographie in Echtzeit - 1974 Merchant

Merchant stellte 1974 seinen Ansatz zur berührungslosen videobasierten Registrierung der Blickrichtung in Echtzeit vor.

„The remote oculometer is a *new* instrument for the remote measurement of eye direction and pupil diameter.“ [155]

Die Videookulographie stellt den modernsten – und für den Bereich der Mensch-Maschine Interaktion vielversprechendsten – Vertreter der Gaze Tracker dar. Für die VOG existiert eine Vielzahl von unterschiedlichsten Messaufbauten, Berechnungsstrategien und Algorithmen, die eine Bestimmung der Blickrichtung genau und robust ausführen. Kapitel 3.1.5 gibt exemplarisch einen Überblick verschiedener Ansätze der Videookulographie.

3.1.5 Videookulografie

Die Videookulografie (VOG) stellt, verglichen mit beispielsweise der Elektronystagmografie oder dem Search Coil Verfahren, die Methode mit der geringsten Invasivität dar [70]. Deshalb handelt es sich dabei – v.a. hinsichtlich der Mensch-Maschine Interaktion – um das Verfahren mit dem größten Zukunftspotential.



(a) Head-mounted Eye Tracker.



(b) Head-supported Eye Tracker.



(c) Remote Eye Tracker.

Abbildung 3.10: Vertreter der Videookulographie, Unterteilung hinsichtlich der Art des verwendeten Messaufbaus.

Zur Bestimmung der Blickrichtung nutzt die VOG moderne Videotechnik. Sie ermöglicht als einziges Verfahren eine berührungslose Augenbewegungsmessung. Zudem werden die hohen Anforderungen für (medizinisch) verwertbare Untersuchungen hinsichtlich Genauigkeit, Robustheit und Schnelligkeit gewährleistet. Zur Registrierung

von Augenbewegungen werden im Kamerabild verschiedene Merkmale verfolgt und deren Position ausgewertet. Über unterschiedliche Berechnungsstrategien können dann beispielsweise Mapping-Funktionen aufgestellt werden, welche einen Zusammenhang zwischen der 2D-Position der Merkmale im Kamerabild und der Position des Fixationspunktes auf dem Monitor bzw. der Blickrichtung herstellen. Weiterhin existieren Ansätze, die die Position des Auges in 3D und somit auch die Blickrichtung in 3D rekonstruieren.

Die unterschiedlichen Verfahren lassen sich hinsichtlich der verwendeten Merkmale (z.B. Purkinje-Bilder, Pupille, Limbus, gesamtes Auge), der Art des verwendeten Messaufbaus (z.B. head-mounted, head-supported, remote; vgl. Abbildung 3.10) oder der verwendeten Hardware (binokular bzw. monokular) gruppieren. Weiterhin kann die Gruppierung nach den verwendeten Algorithmen erfolgen, hierbei unterscheidet man zwischen *impliziten* und *expliziten* Verfahren [258].

Implizite Verfahren setzen zur Berechnung der Blickrichtung eine benutzerspezifische Kalibrierung voraus. Dies liegt darin begründet, dass keine exakte Augenposition berechnet wird. Vielmehr nutzen sie die unterschiedlichen geometrischen Verhältnisse der Bildmerkmale (z.B. Stellung der Pupille im Videobild bei Fixation eines Referenzpunktes), um den Blickpunkt auf einer Referenz- beziehungsweise Kalibrierebene zu bestimmen. Diese Verfahren haben den Nachteil, dass die Kalibrierung nur für eine Tiefenebene Gültigkeit hat und dadurch bei veränderten Positionen versagt. Dies wirkt sich auch auf die Verwendung von head-mounted Geräten aus, denn diese können mangels permanenter Fixierung am Kopf des Probanden verrutschen und dadurch ungültige Ergebnisse liefern. Trotz unterschiedlicher Ansätze zur Kompensation des Verrutschens leiden die impliziten Verfahren unter diesem Nachteil. *Explizite* Verfahren rekonstruieren anhand unterschiedlicher Bildmerkmale die Augenparameter (z.B. Position, Blickrichtung, etc.) in 3D und können dadurch im Gegensatz zu den impliziten Verfahren auch bei unterschiedlichen Positionen – Kopf relativ zur Kamera bzw. Proband relativ zum Stimulus – gültige Ergebnisse liefern. Ein weiterer Vorteil der expliziten Verfahren ist es, auch ohne Kalibrierung eine auswertbare Blickrichtung berechnen zu können. Zwar entspricht diese technische Blickrichtung (optische Achse) nicht der subjektiv wahrgenommenen Blickrichtung (visuelle Achse), doch ermöglicht sie eine erste Einschätzung, die je nach physiologischen Eigenschaften des Benutzers mehr oder weniger gut (im Mittel 5°) an die finale Wahrnehmung des Probanden heranreicht. Typischerweise erreichen Eye Tracker eine Genauigkeit von $0,4^\circ$ bis 1° [60]. Verglichen mit der menschlichen Refixationsgenauigkeit von ca. 1° [212], stellt dies einen ausreichend guten Wert dar.

Entsprechend obiger Unterscheidung nach verwendetem Messaufbau werden im Folgenden head-mounted, head-supported und remote Eye Tracking Systeme diskutiert. Prinzipiell nutzen alle Messaufbauten sowohl implizite als auch explizite Algorithmen, wobei head-mounted und head-supported Eye Tracker typischerweise implizite Strategien verwenden, während remote Eye Tracker auf beide Verfahren zurückgreifen.

3.1.5.1 Head-mounted Eye Tracker

Head-mounted Geräte werden direkt am Kopf des Probanden befestigt. Die Kamera ist über eine mechanische Vorrichtung, z.B. einen Helm, mit dem Kopf verbunden. Dadurch ändert sich die relative Lage von Sensor- und kopffestem Augenkoordinatensystem nicht.

3. Kamera-basierte Fahrerüberwachung

Der Nutzer ist zwar in seiner Bewegungsfreiheit nicht eingeschränkt, muss aber durch das Tragen dieser Einrichtung unter Umständen auf Komfort verzichten.

Aufgrund der sehr geringen Entfernung zwischen Sensor und Auge ermöglicht diese Methode bereits mit günstiger Kamerahardware genaue Messungen, wobei jedoch sichergestellt werden muss, dass das System nicht am Kopf verrutscht, was zu Verfälschung der Messwerte führen würde. Durch eine geeignete Algorithmik kann eine solche Verschiebung erkannt und (teilweise) kompensiert werden [211]. Head-mounted Eyetracker nutzen typischerweise implizite Berechnungsstrategien. Ein favorisiertes Einsatzgebiet für diese Systemklasse sind Studien im medizinischen Umfeld. Hierbei werden okulomotorische Daten (z.B. Sakkadengeschwindigkeit und -rate) als Reaktion auf entsprechende Stimuli ausgewertet und für diagnostische Zwecke (z.B. Untersuchung der okulomotorischen Dynamik) eingesetzt. Weiterhin werden sie für mobile Anwendungen, wie beispielsweise Marktforschungsstudien (z.B. wie gestaltet sich der Blickpfad eines Probanden beim Gang durch einen Supermarkt) oder Usability Studien (z.B. Analyse des Blickpfades eines Probanden im Fahrsimulator) verwendet. Die aufgezeichnete Blickrichtung wird typischerweise dem Bild einer – ebenfalls am Kopfgestell montierten – Blickfeldkamera überlagert. Für diese Art der Datenaufzeichnung und -analyse ist eine 2D Erfassung der Blickrichtung ausreichend. Abbildung 3.10(a) zeigt exemplarisch den kommerziellen head-mounted Eye Tracker der Firma SMI, mit den folgenden Systemeigenschaften: Abtastrate 200Hz, Auflösung 0.1° , Genauigkeit im Mittel 0.5° - 1.0° .

3.1.5.2 Head-supported Eye Tracker

Genauere Ergebnisse als head-mounted Geräte liefern sog. head-supported Geräte. Bei diesen Geräten positioniert der Proband seinen Kopf innerhalb einer Messapparatur. Der Kopf des Probanden wird entweder durch eine Stirn- oder Kinnstütze oder durch eine Bissvorrichtung (bite-bar) fixiert. Hierbei sind in der Regel alle Koordinatensysteme fest und der Proband in seiner Bewegungsfreiheit völlig eingeschränkt. Head-supported

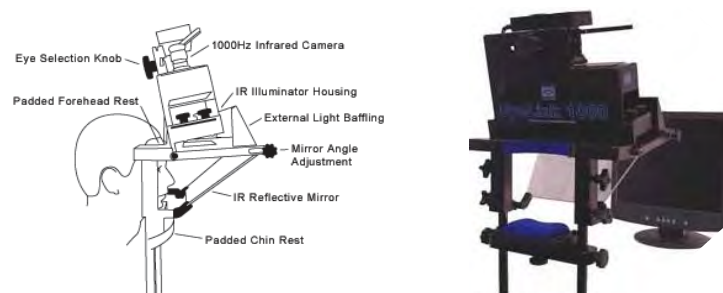


Abbildung 3.11: Head Supported Eyetracker SR Research EyeLink 1000/2000 [221].

Systeme weisen zwar den geringsten Komfort, aber auch die höchste Messgenauigkeit auf. Head-supported Systeme verwenden zur Blickrichtungsbestimmung ähnliche Berechnungsstrategien wie head-mounted Geräte und erreichen typischerweise Auflösungen $\leq 0.02^\circ$, Genauigkeiten um 0.15° (typisch: 0.25 bis 0.5°) und einen Messbereich von 45° vertikal und 55° horizontal bei Abtastraten bis zu 2kHz.

3.1.5.3 Remote Eye Tracker

Den maximal möglichen Komfort ermöglichen sog. remote Eye Tracking Systeme. Der Benutzer kann seinen Kopf „völlig frei“ im Sichtfeld einer kontaktlos angebrachten Kamera bewegen. Dies bedeutet, dass keine Einschränkungen bezüglich der relativen Lage von Sensor- und Augenkoordinatensystem gesetzt werden. Für viele Anwendungen, in denen Komfort von Bedeutung ist, stellt dies die einzig mögliche Variante dar.

Bedingt durch die kontaktlose und komfortable Messung entstehen einige Nachteile: Die größere Entfernung führt zu einer geringeren örtlichen Auflösung der benötigten Merkmale auf dem Sensor. Je nach gewünschter Messgenauigkeit und Bewegungsfreiheit müssen die Pixeldichte, Sensorfläche und Brennweite der verwendeten Optik abgestimmt werden. Generell steigt auch der Aufwand in der Algorithmik. So muss beispielsweise eine Blickrichtungsbestimmung nun auch unter Kopfbewegungen relativ zur Kamera sicher gestellt werden. Die Bildverarbeitung muss also in der Lage sein, Bewegungen des Kopfes von Drehbewegungen des Auges (Torsion) zu unterscheiden.

Erwähnenswert sind weiterhin Systeme, die zwar als Remote Eye Tracker verkauft werden, aber auf technische Hilfen wie Marker oder Stirnbänder angewiesen sind. Das EyeLink Remote der Firma SR Research nutzt einen, an der Stirn des Probanden platzierten kontrastreichen Marker mit konzentrischen Kreisen, um auch während Lidschlägen den Kopf und dessen Abstand zur Kamera zu verfolgen.

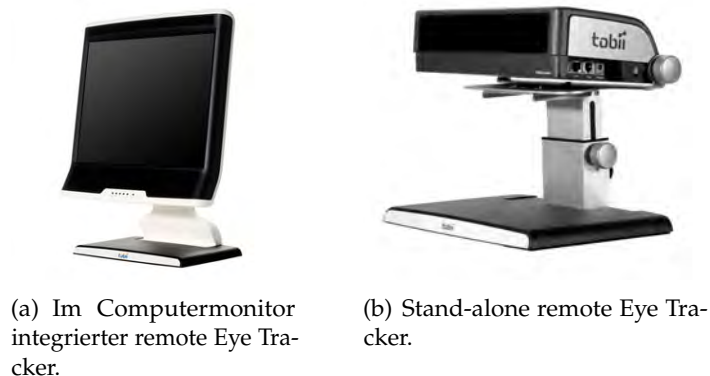


Abbildung 3.12: Remote Eye Tracking Systeme des Herstellers Tobii [232].

Remote Eye Tracker haben jedoch weniger gute Eigenschaften hinsichtlich Abtastrate, Auflösung und Genauigkeit. Für den Bereich der Mensch-Maschine Interaktion reichen diese Werte jedoch aus. Lediglich bei medizinischen Untersuchungen des okulomotorischen Apparats wären die gezeigten Eigenschaften nicht ausreichend. Remote Systeme erreichen typischerweise Abtastraten bis 120 Hz, Auflösungen von 0.1° und eine Genauigkeit von ca. 0.5° bei einem Abstand zwischen Kamera und Proband von 60cm bis 80cm.

3.1.6 Fazit

Das Search Coil Verfahren gilt – vor allem aufgrund seiner hohen Genauigkeit – als Referenzmethode für die Augenbewegungsmessung im klinischen Umfeld. Ebenso ist die Elektronystagmographie klinische Routine, da dieses Verfahren für Schlaforschung auch bei geschlossenen Augen einsetzbar ist. Im Bereich der Mensch-Maschine-Kommunikation stellt die Videookulographie die zukunftsweisendste Alternative dar. Sie liefert nicht nur gute Ergebnisse, sondern sorgt auch für die größte Akzeptanz beim Nutzer, was an der – vor allem bei remote Systemen – nicht-invasiven Arbeitsweise liegt.

3.2 Referenzsystem EyeSeeCam

Zur Durchführung initialer Versuche wurde ein Eye Tracker auf Basis des head-mounted Tracking System *EyeSeeCam* – entwickelt am Center for Sensorimotor Research der Ludwig-Maximilians-Universität München [3] – aufgebaut. Dieses System berechnet Blickdaten in einem fixen Kopfkoordinatensystem (gaze-in-head data). Zur Nutzung im automobilen Umfeld wurde das System so erweitert, dass es Daten in einem 3D Fahrzeugkoordinatensystem (gaze-in-space data) berechnen kann (vgl. [194, 195]). Dazu wurde das System um ein Stereokamerasystem zur Lagebestimmung des kopfgebundenen Aufbaus sowie um eine entsprechende Datentransformation erweitert. Im Folgenden werden das Tracking System *EyeSeeCam* sowie die entsprechenden Erweiterungen zur Nutzung im Fahrzeug vorgestellt.

3.2.1 Head-mounted Eye Tracking System *EyeSeeCam*

Die *EyeSeeCam* (vgl. Abbildung 3.13) ist ein mobiler, kopfgebundener Eye Tracker [61]. Das System bestimmt Blickdaten in einem Kopfkoordinatensystem mit einer Genauigkeit von 0.5° und einer Auflösung von 0.01° bei Datenraten bis zu 600Hz.

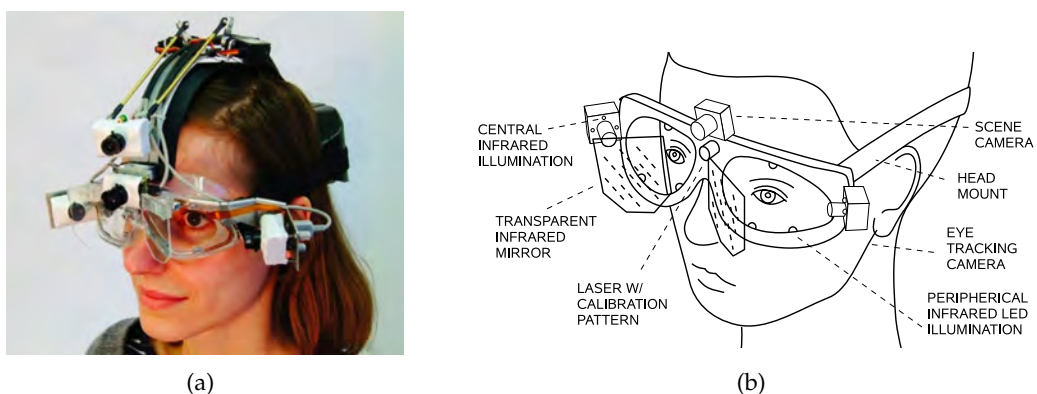


Abbildung 3.13: Head-mounted Eye Tracker *EyeSeeCam*. (a) Head-mounted Eye Tracker mit blickgesteuerter Kamera. (b) Schematische Darstellung des Eye Trackers. Entnommen aus [3]

Als Basis des Systems (head-mount) dient eine leichte, bequem zu tragende und eng am Kopf anliegende Schwimmbrille. An dieser Brille sind alle mechanischen Teile des Aufbaus fixiert. Die Passform der Brille verhindert ein Verrutschen des Aufbaus am Kopf (slippage) weitestgehend. Zur Beleuchtung der Szene (Augenregion, vgl. Abbildung 3.15(a)) sind am unteren Rand der Brille infrarot (IR) LEDs angebracht. Am seitlichen Rand des Aufbaus angebrachte Kameras (Point Grey Research Firefly IEEE1394) zeichnen die zur Blickbestimmung benötigten Videodaten auf. In der Mitte des oberen Randes befindet sich eine Laserdiode zur Projektion des für die Benutzerkalibrierung benötigten 9-Punkt Fixationsmusters (Abstand der Punkte jeweils 8.5° horizontal und vertikal). Das Gesamtgewicht des Aufbaus beträgt ca. 170g. Da die in Abbildung 3.13 zusätzlich dargestellte Aktorik der Blickkamera für die vorliegende Arbeit keine Relevanz hat, soll auf diesen Teil des Aufbaus nicht weiter eingegangen werden.

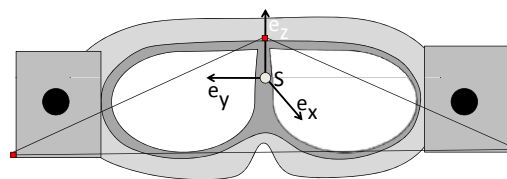
Die Blickrichtungsrekonstruktion basiert auf der Pupillenposition im Videobild (vgl. Abbildung 3.15(a)). Das Pupillentracking sowie die Methoden zur Blickrichtungsbestimmung sind hinsichtlich ihrer Performance ausgewählt und optimiert, um ebenfalls auf mobilen Laptop Computern mit reduzierter Rechenleistung Echtzeitanforderungen zu erfüllen. Zur genauen Funktionsweise (z.B. Pupillendetektion, Blickrichtungsrekonstruktion, Benutzerkalibrierung, etc.) sei an dieser Stelle auf [39, 61] verwiesen. Die berechneten Blickdaten werden mit Hilfe der im folgenden Kapitel 3.2.2 vorgestellten Transformation vom Kopfkoordinatensystem ins Weltkoordinatensystem überführt.

3.2.2 Transformation der Blickdaten in ein Weltkoordinatensystem

Zur Transformation der Blickdaten aus dem Kopf- bzw. Brillenkoordinatensystem in ein 3D Weltkoordinatensystem muss zunächst die Lage (Position und Orientierung) des Kopfgestells im Weltkoordinatensystem bekannt sein. Dazu wurden drei zusätzliche IR LEDs (p_1, p_2, p_3) am Eye Tracker montiert, eine LED in der Mitte des oberen Randes des Kopfgestells und zwei weitere seitlich an den Kamerahalterungen (vgl. rote Punkte in Abbildung 3.14(a)). Der Ursprung des lokalen Brillenkoordinatensystems befindet sich



(a) EyeSeeCam mit LEDs (rote Punkte).



(b) Resultierendes Brillenkoordinatensystem.

Abbildung 3.14: Head-mounted Eye Tracker mit Erweiterungen zum Head Tracking.

im Schwerpunkt s des Dreiecks, welches durch die drei LEDs aufgespannt wird (vgl. Abbildung 3.14(b)). Dementsprechend ergeben sich die x , y und z -Achse des lokalen Koordinatensystems folgendermaßen:

$$\mathbf{e}_x = \mathbf{v}_2 \times \mathbf{v}_3 = (\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_3 - \mathbf{p}_1); \quad \mathbf{e}_y = -\mathbf{v}_2 = \mathbf{p}_1 - \mathbf{p}_3; \quad \mathbf{e}_z = \mathbf{e}_x \times \mathbf{e}_y \quad (3.1)$$

3. Kamera-basierte Fahrerüberwachung

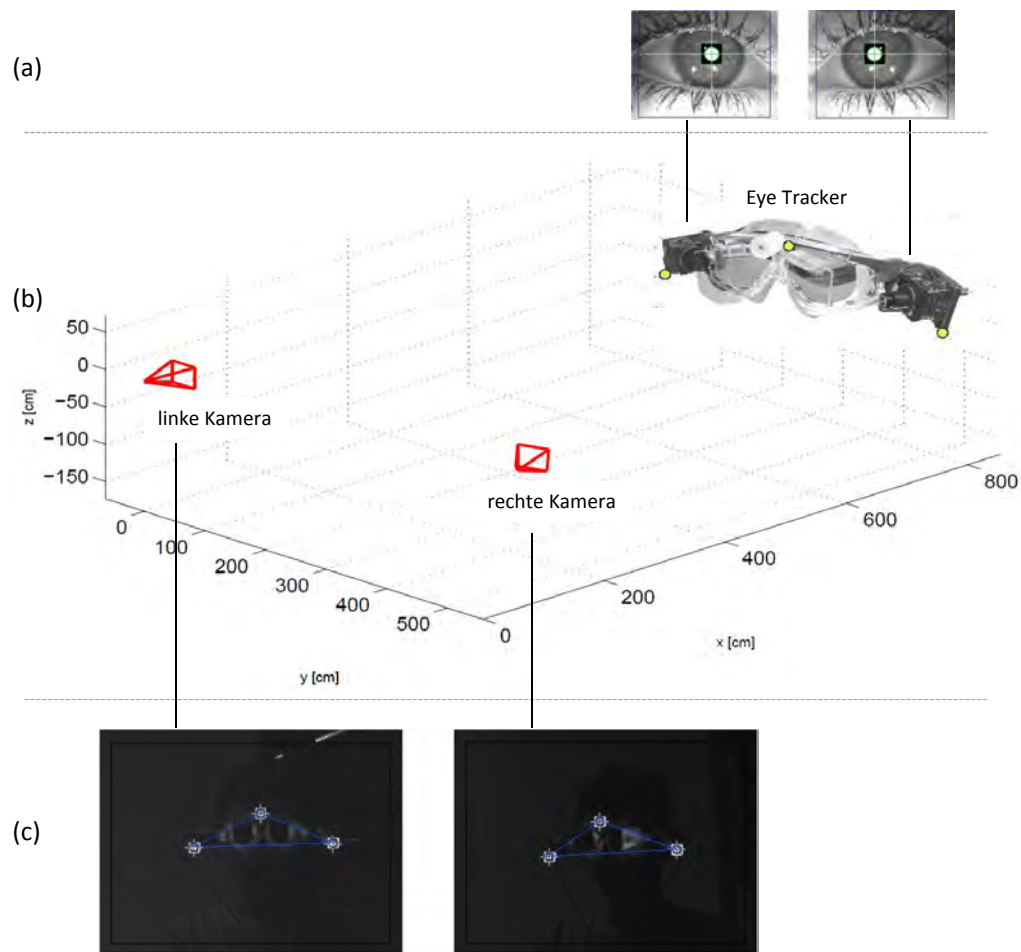


Abbildung 3.15: (a) Videobilder der linken und rechten Eye Tracking Kamera mit erkannter Pupillenposition. (b) Kopf Tracker und Weltkoordinatensystem COS_W , mit Ursprung in der rechten Kamera. Die Position des Eye Trackers im Weltkoordinatensystem wird anhand von drei zusätzlichen LEDs am head-mount bestimmt. (c) Kamerabilder des Stereo-Kamera-Systems zur Lagebestimmung des Eye Trackers. In den Videobildern ist die Lage der LED-Blobs sowie die Lage der aufgespannten Ebene in blau markiert.

Die LED Positionen werden mit Hilfe eines Stereo-Kamera-Systems bestimmt, welches vor dem Benutzer (z.B. auf einem Computer Monitor) platziert ist [194]. Der Ursprung des Weltkoordinatensystems COS_W liegt dabei in der rechten Kamera. Dessen X-Achse wird durch die optische Achse der rechten Kamera beschrieben, die Y-Achse steht senkrecht auf der X-Achse und zeigt in Richtung der linken Kamera. Die Z-Achse steht senkrecht auf der X-Y-Ebene, so dass ein Rechtssystem entsteht. Das Stereo-Kamera-System nutzt die gleichen Kameras wie der Eye Tracker, die synchron zu den Kameras des head-mounted Eye Trackers betrieben werden und ebenfalls mit Infrarot-Filtern ausgestattet sind. Folglich existieren im Videobild (im Idealfall) lediglich Blobs³ der LEDs. Zur Differenzierung zwischen den eigentlichen LEDs und ihren Reflexionen (z.B. auf der

³Abkürzung für binary large object oder basic large object aus dem Englischen.

Haut des Benutzers oder auf dem Eye Tracker, vgl. Abbildung 3.15(c)) werden weiterhin die Parameter Form, Größe und Intensität der einzelnen Blobs ausgewertet. Aus den Zentren der übrigen Blobs im Bild kann dann die Position der LEDs berechnet werden.

Im nächsten Schritt erfolgt die Berechnung der 3D Koordinaten der LEDs über die extrinsischen Parameter des kalibrierten Stereo-Kamera-Systems, um kontinuierlich die relative Translation T und Rotation R des Eye Trackers zu bestimmen. Aus diesen Werten wird die nötige Transformationsmatrix bestimmt, um die Blickgeraden aus dem Brillenkoordinatensystem in das verwendete Weltkoordinatensystem zu überführen.

Datenblock	Inhalt
[0-8]	X, Y, Z -Koordinaten der drei LEDs im Welt-KS, beginnend mit LED1
[9-11]	Normierter Gaze-Vektor im Brillen-KS
[12-14]	Normierter Gaze-Vektor im Welt-KS
[15]	Error-LED: Statusinformation, welche angibt, ob alle drei LEDs erkannt werden können. Kann eine der drei LEDs nicht erkannt werden, dann wird ein Zähler solange inkrementiert, bis alle LEDs wieder korrekt detektiert werden können.
[16]	ErrorGaze: Statusinformation, welche angibt, ob beide Augen erkannt werden können. Kann ein Auge nicht detektiert werden, dann liefert dieses Feld den Wert 1, ansonsten den Wert 0.
[17-19]	X, Y, Z -Koordinaten des Ursprungs des Brillen-KS
[20]	Time-Stamp

Tabelle 3.1: Struktur der Daten, die über den TCP-Socket übertragen werden.

Über einen TCP-Socket stehen die berechneten Informationen zur Position und zur Blickrichtung des Benutzers sowie Statusinformationen, welche Aufschluss über die Verwendbarkeit der Daten geben, anderen Anwendungen zur Verfügung. Eine detaillierte Beschreibung der Daten, die als Float-Array übermittelt werden, findet sich in Tabelle 3.1.

3.2.2.1 Systemperformance

Die Genauigkeit bzw. die Auflösung des head-mounted Eye Trackers ohne zusätzliche Transformation beträgt nach [3] 0.5° bzw. 0.01° . Im Rahmen einer initialen Untersuchung wurden die entwickelten Ansätze zur Lagebestimmung untersucht: Das System erreichte dabei mit einer Genauigkeit von 0.007mm bzw. 0.0032° für Position bzw. Orientierung sehr vielversprechende Werte. Die Genauigkeit des erarbeiteten Gesamtsystems bei der Bestimmung des Blickpunktes auf einem (kalibrierten) Monitor (vgl. Kapitel 3.2.4) kann mit 0.61° angegeben werden. Die Latenz des Gesamtsystems (inkl. Eye Tracking, Lagebestimmung, Transformation der Daten) konnte mit 10ms bestimmt werden [194].

3.2.3 Display Manager Jeannie

Versuchsaufbauten benötigen oftmals nicht nur Kenntnis der aktuellen Blickgraden des Nutzers im Raum, sondern vielmehr den exakten Blickpunkt des Nutzers, beispielsweise



(a) Gesamter Versuchsaufbau mit Sitzkiste. Oberhalb des Armaturenbretts ist das Stereo-Kamera-System zu erkennen.



(b) Versuchsperson mit Eye Tracking Brille aus Sicht der rechten Kamera.

Abbildung 3.16: Versuchsaufbau mit Eye Tracker basierend auf der EyeSeeCam.

auf einem Monitor (area-of-interest, AOI). Aus diesem Grund wurde eine Software zur weiteren Verarbeitung der vom Eye Tracker bestimmten Blickrichtungsinformationen konzipiert, welche die Verwaltung und Kalibrierung von einzelnen AOIs ermöglicht.

Die Aufgaben des AOI Handlers lassen sich grundsätzlich in zwei Bereiche gliedern: (1) Kalibrierung der 3D Position eines Monitors im Weltkoordinatensystem lediglich durch Verwendung von Blickinformationen, (2) Bestimmung des exakten Blickpunktes auf einer zuvor kalibrierten AOI und Zuordnung zu unterschiedlichen AOIs [23, 24]. Dazu wurde eine Kalibrieroutine entworfen, bei der der Benutzer nacheinander bestimmte Punkte auf dem Bildschirm fixiert (vgl. Abbildung 3.17). Aus den dabei gewonnenen Blickrichtungsinformationen wird dann auf eine Beschreibung der Bildebene im Weltkoordinatensystem geschlossen. Im Folgenden werden die einzelnen Schritte der Kalibrierung vorgestellt. Für weitere Details zur Implementierung sei auf [180] verwiesen.

Ziel der Kalibrierung ist es, auf Basis von aufgezeichneten Blickvektoren eine mathematische Beschreibung der Bildebene zu finden. Zur Minimierung von Messungenauigkeiten werden bei der Kalibrierung die Weltkoordinaten von neun Kalibrierpunkten der Bildebene ermittelt und daraus mittels Hauptkomponentenanalyse⁴ eine Ebene approximiert. Diese ermöglicht die Approximation einer Geraden oder (Hyper-)Ebene durch beliebig viele Datenpunkte aus höherdimensionalen Räumen. Das Ergebnis stellt die beste Lösung bezüglich der kleinsten Fehlerquadrate dar. Liegen die Datenpunkte als Vektoren der Form $\mathbf{a}_i = [x_i \ y_i \ z_i]^T$ vor, wird für M Datenpunkte mittels PCA eine Lösung gefunden, die $\sum_{i=1}^M d_i^2 = \min$ erfüllt. Dabei beschreibt d_i den Abstand eines Datenpunktes \mathbf{a}_i von der gefundenen Gerade. Bei der PCA wird ein neues Koordinatensystem gefunden, dessen Ursprung im Mittelwertvektor $\bar{\mathbf{a}}$ liegt und dessen Koordinatenachsen durch die Eigenvektoren \mathbf{v}_1 und \mathbf{v}_2 der Kovarianzmatrix Φ des mittelwertfreien Ensemble

⁴Engl.: *principal components analysis*, PCA. Eine detaillierte Beschreibung der PCA findet sich in [183].

beschrieben ist. Die approximierte Gerade verläuft dann entlang des ersten Eigenvektors der Kovarianzmatrix, d.h. entlang der stärksten Änderung der Datenpunkte.

Fixiert ein Anwender einen bestimmten Punkt, so liefert das Eye Tracking System die Blickrichtungsinformation in Form einer Sichtgeraden:

$$g: \mathbf{x} = \mathbf{p} + s\mathbf{u}, \quad s \in \mathbb{R} \quad (3.2)$$

Um bei der Aufzeichnung der Blickgeraden Messungenauigkeiten weitestgehend auszuschließen, werden mehrere Messwerte aufgenommen. Aus dieser Schar von Sichtgeraden wird anschließend eine Stellvertretende mittels PCA approximiert. Zur Ermittlung der Weltkoordinaten eines bestimmten Kalibrierungspunktes (KP) muss dieser aus unterschiedlichen Positionen betrachtet werden. Dies bedeutet, dass der Anwender – soweit möglich – lediglich seine Augen zur Fixierung der Kalibrierungspunkte bewegen soll. Danach wechselt der Benutzer die Kopfposition und es werden ein weiteres Mal die Blickrichtungsinformationen zu den neuen Kalibrierungspunkten aufgenommen. Über den Schnittpunkt der approximierten Sichtgeraden eines KP (aus unterschiedlicher Betrachtungsposition) können abschließend dessen Weltkoordinaten bestimmt werden.

3.2.3.1 Kalibrierung einer AOI

Zur Kalibrierung einer AOI werden für jeden Kalibrierungspunkt und jede Kopfposition 20 Sichtgeraden (vgl. Gleichung (3.2)) mit $N = 3$ Merkmalen (X , Y und Z -Koordinaten) aufgenommen. Somit liegen insgesamt $M = 2 \cdot 20 = 40$ Datenpunkte vor,

$$\mathbf{a}_1 = \mathbf{p}_1, \mathbf{a}_2 = \mathbf{p}_1 + \mathbf{u}_1, \dots, \mathbf{a}_{M-1} = \mathbf{p}_{M/2}, \mathbf{a}_M = \mathbf{p}_{M/2} + \mathbf{u}_{M/2} \quad (3.3)$$

aus welchen zunächst der Mittelwertvektor und das mittelwertfreie Ensemble

$$\mathbf{\Psi} = [(\mathbf{a}_1 - \bar{\mathbf{a}}) \quad \dots \quad (\mathbf{a}_M - \bar{\mathbf{a}})] \quad \text{mit } \bar{\mathbf{a}} = \frac{1}{M} \sum_{i=1}^M \mathbf{a}_i, \quad \mathbf{\Psi} \in \mathbb{R}^{N \times M} \quad (3.4)$$

der Datenpunkte berechnet werden. Über

$$\mathbf{\Phi} = \frac{1}{M} \mathbf{\Psi} \mathbf{\Psi}^T, \quad \mathbf{\Phi} \in \mathbb{R}^{N \times N} \quad (3.5)$$

erhält man die Kovarianzmatrix, deren Eigenvektoren \mathbf{v}_i anschließend bestimmt werden. Letztlich ergibt sich die approximierte Sichtgerade zu

$$\hat{g}: \mathbf{x} = \bar{\mathbf{a}} + s\mathbf{v}_1 \quad (3.6)$$

Aus den zu einem KP gehörenden approximierten Sichtgeraden \hat{g}_1 und \hat{g}_2 lassen sich durch Schneiden die Weltkoordinaten eines KP bestimmen. Da die Blickgeraden jedoch im Allgemeinen windschief zueinander sind, müssen auf beiden Geraden die Punkte \mathbf{x}_1 und \mathbf{x}_2 mit minimalem Abstand zueinander gefunden werden. Als „Schnittpunkt“ ergibt sich der Mittelpunkt der Verbindungsgeraden $\overline{\mathbf{x}_1\mathbf{x}_2}$ zwischen diesen beiden Punkten,

$$\min_{s,t} \{ |(\mathbf{p} + s\mathbf{u}) - (\mathbf{q} + t\mathbf{v})| \} \quad (3.7)$$

was sich über die Pseudoinverse \mathbf{A}^+ der Matrix \mathbf{A} lösen lässt.

$$\mathbf{x} = \mathbf{A}^+ \mathbf{b} \quad \text{mit } \mathbf{A} = [\mathbf{u} \quad -\mathbf{v}], \quad \mathbf{b} = \mathbf{q} - \mathbf{p}, \quad \mathbf{x} = \begin{pmatrix} s \\ t \end{pmatrix} \quad (3.8)$$

Aus s und t sowie den resultierenden Geradengleichungen \hat{g}_1 und \hat{g}_2 erhält man schließlich die Weltkoordinaten des Kalibrierungspunktes KP_i

$$\mathbf{p}_{KP,i} = \frac{1}{2} (\mathbf{x}_1 + \mathbf{x}_2) \quad (3.9)$$

woraus im nächsten Schritt die Bildebene bestimmt werden kann. Im \mathbb{R}^3 ist eine Ebene eindeutig über den Stütz- \mathbf{p} und zwei Richtungsvektoren \mathbf{u} und \mathbf{v} beschrieben.

$$E : \mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}, \quad s, t \in \mathbb{R} \quad (3.10)$$

Aus den KP_i wird mittels PCA eine Ebene bestimmt, die den quadratischen Abstand zwischen den einzelnen $\mathbf{p}_{KP,i}$ und der approximierten Ebene minimiert. Dazu werden zunächst das mittelwertfreie Ensemble (vgl. Gleichung (3.4)) und die Kovarianzmatrix (vgl. Gleichung (3.5)) bestimmt.

$$\tilde{\Psi} = \Psi^T \mathbf{V} \quad \text{mit } \mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3] \quad (3.11)$$

Der Stützvektor der Ebene ergibt sich durch den Mittelwertvektor $\bar{\mathbf{a}}$ und die Richtungsvektoren aus den beiden ersten Eigenvektoren \mathbf{v}_1 und \mathbf{v}_2 der Kovarianzmatrix.

$$\hat{E} : \mathbf{x} = \bar{\mathbf{a}} + s\mathbf{v}_1 + t\mathbf{v}_2 \quad (3.12)$$

Somit liegen die (im PCA Raum dargestellten) Kalibrierungspunkte in der approximierten Ebene. Nach einer Rücktransformation können die Kalibrierungspunkte durch eine Linearkombination der ersten beiden Eigenvektoren beschrieben

$$\hat{\Psi} = [\hat{\psi}_1 \quad \hat{\psi}_2 \quad \dots \quad \hat{\psi}_9] = [\mathbf{v}_1 \quad \mathbf{v}_2] \tilde{\Psi}^T \quad (3.13)$$

und in der neuen Datenmatrix $\hat{\mathbf{A}}$

$$\hat{\mathbf{A}} = [(\bar{\mathbf{a}} + \hat{\psi}_1) \quad (\bar{\mathbf{a}} + \hat{\psi}_2) \quad \dots \quad (\bar{\mathbf{a}} + \hat{\psi}_9)] \quad (3.14)$$

zusammengefasst werden. Aus diesen Kalibrierungspunkten ergibt sich die Beschreibung der Bildebene. Als Stützpunkt dient der erste Kalibrierungspunkt und die Richtungsvektoren ergeben sich aus der Differenz aus drittem und ersten bzw. aus siebtem und ersten Kalibrierungspunkt:

$$\hat{E} : \mathbf{x} = \hat{\mathbf{p}}_{KP,1} + s \cdot (\hat{\mathbf{p}}_{KP,3} - \hat{\mathbf{p}}_{KP,1}) + t \cdot (\hat{\mathbf{p}}_{KP,7} - \hat{\mathbf{p}}_{KP,1}) \quad (3.15)$$

$$\text{mit } \hat{\mathbf{p}}_{KP,1} = \begin{pmatrix} \hat{a}_{11} \\ \hat{a}_{21} \\ \hat{a}_{31} \end{pmatrix}, \quad \hat{\mathbf{p}}_{KP,3} = \begin{pmatrix} \hat{a}_{13} \\ \hat{a}_{23} \\ \hat{a}_{33} \end{pmatrix}, \quad \hat{\mathbf{p}}_{KP,7} = \begin{pmatrix} \hat{a}_{17} \\ \hat{a}_{27} \\ \hat{a}_{37} \end{pmatrix}$$

3.2.3.2 Bestimmung des Blickpunktes

Aus der beschriebenen Kalibrierung ergibt sich eine mathematische Beschreibung der Monitorebene im Weltkoordinatensystem. Während des Betriebs liefert der Eye Tracker Blickrichtungsinformationen in Form einer Blickrichtungsgeraden (ebenfalls in Weltkoordinaten). Zur Bestimmung des Blickpunktes wird die aktuelle Sichtgerade g mit der Bildschirmenebene E geschnitten und eine Beschreibung in Pixelkoordinaten gefunden. Sind die Bildschirmenebene und die Sichtgerade durch die Gleichungen

$$E: \mathbf{x} = \mathbf{p} + s \cdot \mathbf{u} + t \cdot \mathbf{v}$$

$$g: \mathbf{x} = \mathbf{q} + r \cdot \mathbf{w}$$

gegeben, so erhält man den Schnittpunkt in Weltkoordinaten durch simples Gleichsetzen von E und g . Anhand der bekannten Monitorabmessungen und -auflösung wird dieser Schnittpunkt in das entsprechende Pixelkoordinatensystem überführt.

3.2.3.3 Verwaltung von unterschiedlichen AOIs

Die beschriebene Kalibrierung kann abhängig vom jeweiligen Versuchsaufbau für beliebig viele Bildschirmenebenen bzw. AOIs durchgeführt werden. Zur Erfüllung der zweiten Anforderung an das System, werden die einzelnen Datensätze vom System gespeichert und verwaltet. Eine beliebige Client-Anwendung kann unter Angabe der entsprechenden ID den Gaze Point in Pixelkoordinaten zu einem bestimmten Monitor im Verbund über einen TCP-Socket abfragen. Somit ist sichergestellt, dass die einzelnen Anwendungen nur dann Blickrichtungsinformationen erhalten, wenn der Nutzer auch tatsächlich den entsprechenden Monitor bzw. die entsprechende AOI betrachtet.

3.2.4 Evaluierung des Gesamtsystems

Die Genauigkeit der vorgestellten Kalibrierungsmethode wurde im Rahmen eines Probandenversuchs untersucht. An den Versuchen, in denen das Gesamtsystem bestehend aus dem Eye Tracker EyeSeeCam sowie dem vorgestellten System zur Kalibrierung von AOIs und Verarbeitung von Blickdaten zum Einsatz kam, nahmen zehn VPn teil.

Die Teilnehmer mussten fünf verschiedene AOIs an bekannten und vorher vermessenen Positionen kalibrieren. Der Abstand zwischen Benutzer und AOI betrug 65cm. Im Rahmen der Versuche erreichte das System, gemittelt über alle Teilnehmer und AOI Positionen, eine Kalibrierungsgenauigkeit von 0.61° [23, 24]. Zur Bestimmung der Genauigkeit der Rekonstruktion des Blickpunktes auf einem zuvor kalibrierten Monitor, mussten die Probanden neun Punkte auf dem Monitor fixieren. Diese Referenzpunkte waren dabei analog zum verwendeten Kalibrieremuster (3×3 Matrix, vgl. Abbildung 3.17) angeordnet. Der Abstand zwischen Nutzer und AOI betrug ebenfalls 65cm. Während der Messung konnten die Probanden Kopfdrehungen um $\pm 37^\circ$ um die horizontale sowie um $\pm 40^\circ$ um die vertikale Achse durchführen. Die Genauigkeit für die Rekonstruktion des Blickpunktes über alle Probanden und alle Fixationspunkte gemittelt betrug $0,458^\circ$ bei einer Auflösung von $0,056^\circ$ (vgl. Abbildung 3.17) [23, 24].

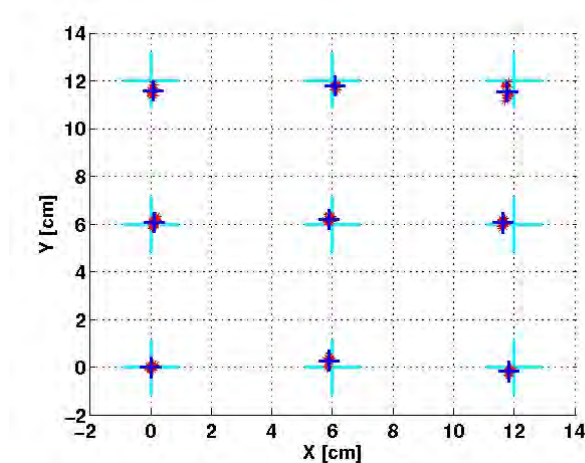


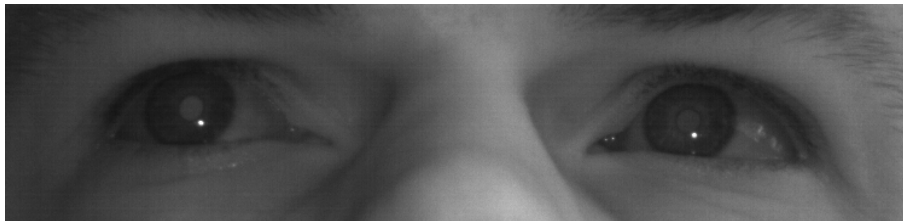
Abbildung 3.17: Aufgezeichnete Fixationsdaten am zuvor kalibrierten Monitor. Die hellblauen Kreuze markieren die angezeigten Referenz- bzw. Targetpunkte. Die roten Punktwolken markieren alle aufgezeichneten Blickpunkte und die dunkelblauen Kreuze markieren die Mittelwerte der berechneten Schnitt- bzw. Blickpunkte.

3.3 Remote Eye Tracking System *eyeDrive*

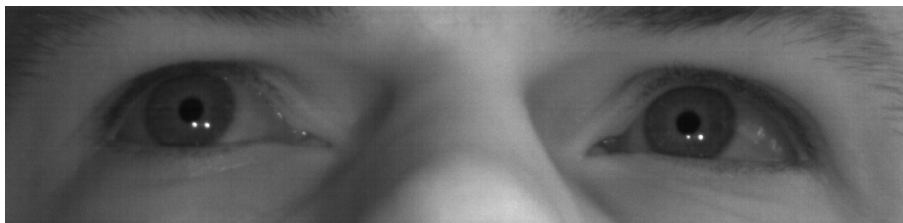
eyeDrive ist ein Ein-Kamera System zur berührungslosen Erfassung der Blickrichtung eines Probanden in 3D. Dazu bestimmt das System relevante Parameter im Videobild und berechnet daraus eine Blickgrade in 3D. Hierfür muss eine Verbindung zwischen dem 2D Bild- und dem 3D Weltkoordinatensystem hergestellt werden. In der Literatur existieren dazu unterschiedliche Verfahren, die sich in der Anzahl der Kameras, der verwendeten Optik und dem Aufbau der Beleuchtung unterscheiden:

Ohno [176] verwendet ein Autofokus Objektiv dessen Fokusweite und damit auch die Entfernung des Auges sich per Software abfragen lässt. Ebenfalls mit einer Kamera aber mit zwei LEDs ausgestattet arbeitet das System von *Hennessey* [106]. Über die geometrischen Beziehungen zwischen den Sensorkoordinaten der Korneareflexionen und der bekannten Position der LEDs im Raum kann die Position des Korneamittelpunktes im Raum berechnet werden. Das Ein-Kamera System von *Wang* [249] kommt ganz ohne LEDs aus, indem es den Irisrand als Merkmal benutzt. Im Gegensatz zum Rand der Pupille lässt sich der Rand der Iris als Kreis mit konstantem Durchmesser beschreiben. Über eine perspektivische Projektion dieses Kreises auf den im Bild detektierten, elliptischen Irisrand wird eine Beziehung zwischen den beiden Koordinatensystemen hergestellt. Schließlich sind noch diverse Systeme zu nennen, die mit Stereo-Kameras arbeiten, z.B. [36, 129, 228]. Sie berechnen die Weltkoordinaten von Augenmerkmalen wie dem Pupillenmittelpunkt oder -rand über Triangulation oder homographische Abbildung. Ist mithilfe einer dieser Techniken eine Verbindung zu Weltkoordinaten hergestellt, lassen sich unter Verwendung eines dreidimensionalen Augenmodells selbst im Videobild nicht ersichtliche Punkte in Weltkoordinaten rekonstruieren (z.B. der Korneamittelpunkt). Der Blick lässt sich dann beispielsweise als Normale auf den Pupillenrand bestimmen [249].

eyeDrive nutzt zur Bestimmung der Blickrichtung in 3D einen Hardwareaufbau bestehend aus einer Kamera und insgesamt drei aktiven infraroten Beleuchtungseinheiten (vgl. Kapitel 3.3.3). Eine dieser Beleuchtungseinheiten ist nahe der Kameraachse (konzentrisch um die optische Kameraachse, on-axis Beleuchtung) und zwei der Beleuchtungseinheiten sind fern der optischen Achse (off-axis Beleuchtung) angeordnet. Die on-axis Beleuchtung erzeugt den sog. „Rote-Augen-Effekt“ und dient vor allem der vereinfachten Merkmalsextraktion (Differenzbildtechnik, vgl. Kapitel 3.3.2) bzw. der Verstärkung/Hervorhebung natürlicher Merkmale (Pupille, vgl. Abbildung 3.18(a)), welche zur Blickrichtungsbestimmung erforderlich sind. Die off-axis Beleuchtungen dienen der Berechnung des Korneazentrums, indem sie künstliche Merkmale – in Form von Reflexionen sog. Glints auf dem Auge (vgl. Abbildung 3.18(b)) – erzeugen (vgl. Kapitel 3.3.1). Die Beleuchtungen arbeiten im nicht sichtbaren infraroten Bereich, weshalb der Nutzer sie nicht als störend empfindet. Bei der Systemspezifikation ist jedoch darauf zu achten, die Grenzwerte hinsichtlich der Strahlenexposition des menschlichen Auges mit IR Licht nicht zu überschreiten.



(a) On-Axis beleuchtetes Bild (bright pupil).



(b) Off-Axis beleuchtetes Bild (dark pupil).

Abbildung 3.18: Kamerabilder des remote Eye Tracking Systems *eyeDrive*.

Aus den (so erzeugten) Merkmalen im Bild kann man zudem die Lage von Augenparametern – z.B. Pupillenmittelpunkt oder -rand – in Weltkoordinaten berechnen. Bei der Blickrichtungsrekonstruktion über den Pupillenmittelpunkt ist weiterhin die Brechung des Lichtes an der Korneaoberfläche zu berücksichtigen, um nicht den Mittelpunkt einer virtuellen Pupille zu bestimmen. Dazu verwenden verschiedene Ansätze ein 3D Augenmodell [36, 106, 176], das die geometrischen und optischen Eigenschaften des Auges approximiert (vgl. Abbildung 3.4 und Kapitel 3.1.2). Dieses bricht die Lichtstrahlen zusätzlich gemäß der unterschiedlichen Brechindizes von Luft n_L und Kammerwasser n_K . Unter Verwendung der Daten aus dem Augenmodell und mindestens einer LED, deren Lichtstrahl an der Korneaoberfläche reflektiert und auf die Reflexionspunkte abgebildet wird, lässt sich der Mittelpunkt der Korneakugel K in 3D Koordinaten bestimmen.

Die optische Achse des Auges berechnet sich als Verbindungsgerade zwischen den Mittelpunkten von Pupille und Korneakugel. Die visuelle Achse, die Verbindungsgerade von fixiertem Objekt und Fovea Centralis, weicht von der optischen Achse um ca. 5° ab. Um den daraus resultierenden Fehler zu minimieren wird im Rahmen einer nutzerspezifischen Kalibrierung diese Abweichung für jeden Probanden bestimmt.

Die folgenden Kapitel beschreiben das im Rahmen der Arbeit realisierte System *eyeDrive* zur berührungslosen Rekonstruktion der Blickrichtung in 3D. Im Einzelnen stellt das Kapitel die mathematischen Strategien zur Berechnung einer 3D Blickgeraden auf Basis lediglich einer Kamera sowie einer speziellen Beleuchtung (vgl. Kapitel 3.3.1), die zugrunde liegende Parameterextraktion (vgl. Kapitel 3.3.2), die verwendete Hardware (vgl. Kapitel 3.3.3) sowie Evaluierungsergebnisse des Gesamtsystems (vgl. Kapitel 3.3.5) vor. Die Qualität der Blickrichtungserkennung ist dabei sowohl von der Parameterextraktion, als auch von der Akkuratheit der Vermessung des Hardwareaufbaus abhängig.⁵

3.3.1 Rekonstruktion der Blickgeraden in 3D

Das folgende Kapitel beschreibt in Anlehnung an [187] die einzelnen Berechnungsschritte, die notwendig sind, um aus den in Kapitel 3.3.2 extrahierten Bildparametern die Blickrichtung des Probanden in 3D zu rekonstruieren.

3.3.1.1 Berechnung des Korneazentrums

Zur Rekonstruktion der Blickrichtung gilt es zunächst das Korneazentrum zu bestimmen. Dieser Parameter ist die Grundlage für alle weiteren Berechnungen, da sich die optische Achse des Auges als Gerade durch Pupillen- und Korneazentrum ergibt (vgl. Kapitel 3.1.2). Abbildung 3.19 zeigt die geometrischen Überlegungen zur Bestimmung des Korneazentrums. Zwecks besserer Anschaulichkeit illustriert die Abbildung diese lediglich für eine der beiden externen Beleuchtungseinheiten bzw. für eine der beiden Reflexionen/Glints dieser Beleuchtungseinheiten auf dem Auge. Tatsächlich wird eine weitere Beleuchtungseinheit und damit eine weitere parametrische Kurve $c(\lambda)$ benötigt, um die Lage des Korneazentrums in 3D bestimmen zu können. Mathematisch basiert der Ansatz auf den folgenden zwei Formeln 3.16 und 3.17 (vgl. Abbildung 3.19):

$$\mathbf{s}(\lambda) = \mathbf{g} + \lambda \cdot \mathbf{v} \quad (3.16)$$

\mathbf{g} beschreibt den Vektor vom Ursprung $\mathbf{0}$ zum Abbild des Pupillenrandes auf dem Sensor und \mathbf{v} den negativen Einheitsvektor von \mathbf{g} . Für das Korneazentrum $\mathbf{c}(\lambda)$ gilt

$$\mathbf{c}(\lambda) = \mathbf{s}(\lambda) + r_c \cdot \mathbf{I}(\lambda). \quad (3.17)$$

r_c entspricht dem Kornearadius und $\mathbf{I}(\lambda)$ der normierten Winkelhalbierenden zwischen $\mathbf{i}(\lambda)$ und \mathbf{v} , mit $\mathbf{I}(\lambda) = \frac{\mathbf{v} + \mathbf{i}(\lambda)}{\|\mathbf{v} + \mathbf{i}(\lambda)\|}$. Mit Gleichung 3.17 ergibt sich für $\mathbf{c}(\lambda)$

$$\mathbf{c}(\lambda) = \mathbf{s}(\lambda) + r_c \cdot \frac{\mathbf{v} + \mathbf{i}(\lambda)}{\|\mathbf{v} + \mathbf{i}(\lambda)\|} \quad (3.18)$$

⁵Hierbei ist vor allem die relative Lage von Kamera, off-axis Beleuchtungseinheiten sowie dem zur Kalibrierung verwendeten Monitors zueinander von besonderem Interesse.

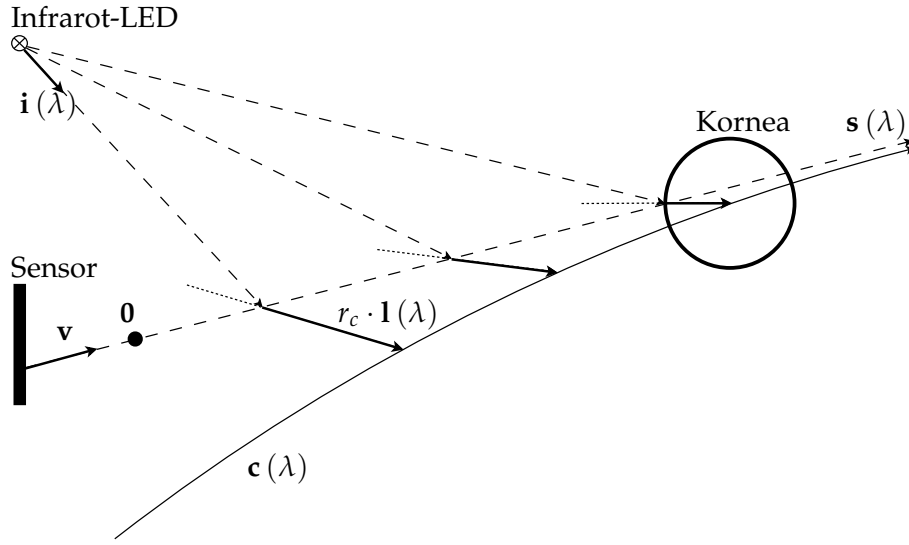


Abbildung 3.19: Geometrische Überlegungen zur Bestimmung des Korneazentrums.

$\mathbf{i}(\lambda)$ kann man in Abhängigkeit der Geraden $\mathbf{s}(\lambda)$ und der Position der Infrarot-LED \mathbf{L} angeben. Mit $\mathbf{i}(\lambda) = \frac{\mathbf{s}(\lambda) - \mathbf{L}}{\|\mathbf{s}(\lambda) - \mathbf{L}\|}$ ergibt sich für $\mathbf{c}(\lambda)$

$$\mathbf{c}(\lambda) = \mathbf{s}(\lambda) + r_c \cdot \frac{\mathbf{v} + \frac{\mathbf{s}(\lambda) - \mathbf{L}}{\|\mathbf{s}(\lambda) - \mathbf{L}\|}}{\|\mathbf{v} + \frac{\mathbf{s}(\lambda) - \mathbf{L}}{\|\mathbf{s}(\lambda) - \mathbf{L}\|}\|} \quad (3.19)$$

und durch Einsetzen von Gleichung 3.16 für $\mathbf{s}(\lambda)$ ergibt sich

$$\mathbf{c}(\lambda) = \mathbf{g} + \lambda \cdot \mathbf{v} + r_c \cdot \frac{\mathbf{v} + \frac{\mathbf{g} + \lambda \cdot \mathbf{v} - \mathbf{L}}{\|\mathbf{g} + \lambda \cdot \mathbf{v} - \mathbf{L}\|}}{\|\mathbf{v} + \frac{\mathbf{g} + \lambda \cdot \mathbf{v} - \mathbf{L}}{\|\mathbf{g} + \lambda \cdot \mathbf{v} - \mathbf{L}\|}\|} \quad (3.20)$$

Gleichung 3.20 kann als nichtlineares, unterbestimmtes Gleichungssystem mit vier Unbekannten ($\mathbf{c}(\lambda) = [c_x \ c_y \ c_z]^T, \lambda$) und drei Gleichungen aufgefasst werden. Wird Gleichung 3.20 nicht nur für eine der beiden infrarot LEDs $\mathbf{c}(\lambda_1)$, sondern für beide aufgestellt, erhält man durch Subtraktion dieser beiden parametrischen Kurven⁶

$$\mathbf{c}(\lambda_1) - \mathbf{c}(\lambda_2) = \mathbf{0} \quad (3.21)$$

und damit ein nichtlineares, überbestimmtes Gleichungssystem der Form

$$\mathbf{g}_1 + \lambda_1 \cdot \mathbf{v}_1 + r_c \cdot \frac{\mathbf{v}_1 + \frac{\mathbf{g}_1 + \lambda_1 \cdot \mathbf{v}_1 - \mathbf{L}_1}{\|\mathbf{g}_1 + \lambda_1 \cdot \mathbf{v}_1 - \mathbf{L}_1\|}}{\|\mathbf{v}_1 + \frac{\mathbf{g}_1 + \lambda_1 \cdot \mathbf{v}_1 - \mathbf{L}_1}{\|\mathbf{g}_1 + \lambda_1 \cdot \mathbf{v}_1 - \mathbf{L}_1\|}\|} - \mathbf{g}_2 + \lambda_2 \cdot \mathbf{v}_2 + r_c \cdot \frac{\mathbf{v}_2 + \frac{\mathbf{g}_2 + \lambda_2 \cdot \mathbf{v}_2 - \mathbf{L}_2}{\|\mathbf{g}_2 + \lambda_2 \cdot \mathbf{v}_2 - \mathbf{L}_2\|}}{\|\mathbf{v}_2 + \frac{\mathbf{g}_2 + \lambda_2 \cdot \mathbf{v}_2 - \mathbf{L}_2}{\|\mathbf{g}_2 + \lambda_2 \cdot \mathbf{v}_2 - \mathbf{L}_2\|}\|} = \mathbf{0} \quad (3.22)$$

⁶Dies entspricht der Annahme, dass beide Kurven exakt durch ein und den selben Punkt – den Korneamittelpunkt – verlaufen.

In Gleichung 3.21 bzw. 3.22 sind nunmehr lediglich $\lambda_{1,2}$ unbekannt. Aufgrund der Komplexität von Gleichung 3.22 ist eine analytische Lösung nicht möglich, weshalb zur numerischen Lösungsbestimmung ein Gradientenverfahren eingesetzt wird, welches die auf Basis von Gleichung 3.22 entstehende Energiefunktion bzw. das entsprechende Kostenfunktional minimiert. Zur numerischen Lösung wird das konjugierte Newton-Verfahren eingesetzt [47]. Zu weiteren Einzelheiten der Implementierung sei an dieser Stelle auf [69] verwiesen. Die Lösung des Kostenfunktionals ergibt numerische Werte für $\lambda_{1,2}$, die nach Einsetzen in Gleichung 3.17 die Position des Korneazentrums ergeben.

3.3.1.2 Berechnung des Pupillenzentrums

Zur Bestimmung der optischen Achse muss weiterhin das Pupillenzentrum (bzw. der Pupillenmittelpunkt) bekannt sein. Dieses wird anhand der in Abbildung 3.20 dargestellten geometrischen Zusammenhänge bestimmt. Primär wird dabei folgende Geradenglei-

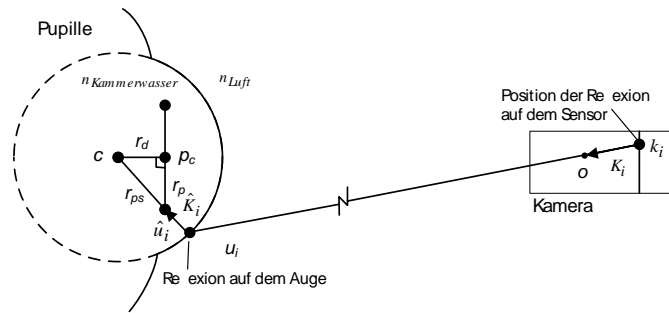


Abbildung 3.20: Geometrische Überlegungen zur Abschätzung des Pupillenmittelpunktes nach [106].

chung im dreidimensionalen Raum

$$\mathbf{u}_i = \mathbf{k}_i + s \cdot \mathbf{K}_i \quad (3.23)$$

mit \mathbf{k}_i als Vektor vom Brennpunkt zum Aufpunkt (Abbildung der Reflexion auf dem Sensor) und \mathbf{K}_i als normierten Richtungsvektor vom Aufpunkt durch den Brennpunkt der Kamera (was dem negativen Einheitsvektor des Aufpunktvektors entspricht) verwendet.

$$\mathbf{K}_i = \frac{-\mathbf{k}_i}{\|\mathbf{k}_i\|} \quad (3.24)$$

\mathbf{u}_i ergibt sich als Schnittpunkt dieser Geraden mit der Korneaoberfläche. Es wird also eine weitere Gleichung benötigt, um auf die vier Unbekannten $\mathbf{u} = [u_{ix} \ u_{iy} \ u_{iz}]^T$ und s schließen zu können. Hierfür wird die Bedingung genutzt, dass sich die Gerade aus Formel 3.23 und die Kornea schneiden müssen (vgl. Abbildung 3.20). Formal lässt sich das wie folgt ausdrücken:

$$(u_{ix} - c_x)^2 + (u_{iy} - c_y)^2 + (u_{iz} - c_z)^2 = r^2 \quad (3.25)$$

Setzt man die Gleichung(en) 3.23 in Gleichung 3.25 ein, so erhält man eine quadratische Gleichung. Zuerst werden die Quadrate in Gleichung 3.25 aufgelöst

$$\begin{aligned} (u_{ix} - c_x)^2 + \dots - r^2 &= 0 \\ u_{ix}^2 - 2u_{ix}c_x + c_x^2 + \dots - r^2 &= 0 \end{aligned} \quad (3.26)$$

und Einsetzen von Gleichung 3.23 in Gleichung 3.26 ergibt

$$(k_i + sK_i)^2 - 2(k_i + sK_i)c_x + c_x^2 + \dots - r^2 = 0,$$

Durch Ausmultiplizieren und Sortieren erhält man daraus

$$K_i s^2 + (2k_i K_i - 2K_i c_x) s + c_x + \dots - r^2 = 0, \quad (3.27)$$

was eine quadratische Gleichung ergibt, die mit Hilfe der Mitternachtsformel gelöst wird. Von den zwei Lösungen kann die Größere verworfen werden, da diese den zweiten fiktiven Schnittpunkt der Geraden aus Gleichung 3.24 und der „Rückseite der Korneakugel“ darstellt.

Der Faktor s liefert – eingesetzt in Gleichung 3.23 – den Schnittpunkt der Geraden mit der Kornea. Um den gesuchten Punkt $\hat{\mathbf{u}}_i$ auf der Pupille zu bestimmen, muss man den Richtungsvektor \mathbf{K}_i mit Hilfe des Snelliusschen Brechungsgesetzes

$$\frac{n_1}{n_2} = \frac{\sin \theta_1}{\sin \theta_2} \quad (3.28)$$

rotieren und eine weitere Geradengleichung aufstellen. Hierzu existieren verschiedene Realisierungs- und Optimierungsmöglichkeiten [102]. Diese Gerade besitzt als Aufpunkt den zuvor berechneten Schnittpunkt \mathbf{u}_i auf der Korneaoberfläche und als Richtungsvektor den gebrochenen Richtungsvektor $\hat{\mathbf{K}}_i$.

$$\hat{\mathbf{u}}_i = \mathbf{u}_i + w\hat{\mathbf{K}}_i \quad (3.29)$$

Diese Gleichung wird analog zu obiger Vorgehensweise (Berechnung von \mathbf{u}_i) gelöst. Zur Bestimmung der Pupillenrandpunkte gilt (vgl. Abbildung 3.20):

$$\|\hat{\mathbf{u}}_i - \mathbf{c}\| = r_{ps} \quad \text{wobei} \quad r_{ps} = \sqrt{r_d^2 + r_p^2} \quad (3.30)$$

mit r_d als Abstand zwischen Kornea- und Pupillenzentrum. Sein numerischer Wert wurde von Gullstrand im Populationsmittel mit 6mm berechnet [70, 94]. r_p (Pupillenzentrum) kann unter Verwendung des aktuellen Kamerabildes und des Strahlensatzes bestimmt werden. Hierbei gilt $r_p/d = r_{p\text{Sensor}}/fl$, woraus für r_p folgt

$$r_p = d \cdot \frac{r_{p\text{Sensor}}}{fl} \quad (3.31)$$

Zur Bestimmung von r_p aus Gleichung 3.31 muss d bekannt sein. Hierzu wird vom Korneazentrum \mathbf{c} der Abstand zwischen Pupillen- und Korneazentrum subtrahiert. Man erhält also $d = |\mathbf{c}| - r_d$. Für Punkte, die sich weder am Rand noch genau in der Mitte der Pupille befinden, lässt sich kein exakter Wert angeben. Es bleiben also zur robusteren und genaueren Bestimmung des Pupillenzentrums – durch Mittelung mehrerer Punkte – nur Randpunkte oder der Mittelpunkt. Dies reicht jedoch aus, um beispielsweise mit drei Punktpaaren (Punkte eines Paares sollten auf dem Pupillenrand gegenüber liegen) eine robuste Abschätzung des Pupillenmittelpunktes durchzuführen [106].

3.3.1.3 Berechnung der optischen Achse

Die optische Achse **OA** berechnet sich als Differenzvektor zwischen zwei Punkten im dreidimensionalen Raum

$$\mathbf{OA} = \mathbf{PC} - \mathbf{CC}, \quad (3.32)$$

mit **PC** als Pupillenzentrum und **CC** als Cornea- bzw. Korneazentrum. Abbildung 3.4(b) veranschaulicht diesen Zusammenhang, weiterhin ist die visuelle Achse **VA** eingezeichnet, welche die eigentlich wahrgenommene Blickrichtung darstellt. Zur Bestimmung der **VA** aus der **OA** wird eine Benutzerkalibrierung benötigt, welche das folgende Kapitel 3.3.1.4 vorstellt. Hier sei angemerkt, dass sich die **VA** als Vektor ausgehend von der Fovea Centralis (Punkt des schärfsten Sehens) Richtung Pupillenzentrum **PC** definieren lässt. Der Unterschied zwischen beiden Achsen beträgt im Bevölkerungsmittel 5° [75, 94].

3.3.1.4 Benutzerkalibrierung

Zur Kompensation des interpersonellen Unterschiedes zwischen optischer **OA** und visueller Achse **VA** ist eine benutzerspezifische Kalibrierung notwendig. Dazu existieren in der Literatur prinzipiell zwei unterschiedliche Strategien. Beide schreiben eine einmalige Kalibrierung pro Benutzer vor, unterscheiden sich vor allem in der Allgemeingültigkeit:

In den Arbeiten von *Hennessey* [106] und *Ohno* [176] fixiert der Benutzer während der Kalibrierung vorgegebene Fixationspunkte (z.B. die Eckpunkte des Bildschirms). Gleichzeitig ermittelt das System die Abweichung der berechneten von den tatsächlichen Punkten. Im Betrieb wird für jeden berechneten Fixationspunkt durch Interpolation (der Abweichungen während der Kalibrierung) ein Korrekturvektor ermittelt und addiert.

Beymer [36] und *Zhu* [258] ermitteln nicht die auf einer bestimmten Zielfläche resultierende Abweichung zwischen vorgegebenen und berechneten Punkten. Sie bestimmen die für den Benutzer charakteristische Abweichung der Achsen um einen gemeinsamen Drehpunkt im Auge. Im Betrieb wird dann die berechnete optische Achse um diese benutzerspezifischen Abweichungen gedreht und somit in die visuelle Achse überführt.

Die erste, quasi extrinsische Methode kann den systematischen Fehler für jeden Zielpunkt/jede Zielebene nur beheben, wenn der Benutzer während der Kalibrierung genau diesen Punkt oder Punkte auf genau dieser Ebene fixiert. Deshalb erweist sich diese Strategie als nicht allgemeingültig. Letztere Methode löst das Problem intrinsisch, indem die physiologisch bedingte Abweichung im Auge ermittelt und dann zur Korrektur herangezogen wird, unabhängig davon welche Punkte im Raum fixiert werden.

Das vorgestellte System nutzt zur Nutzerkalibrierung einen intrinsischen Ansatz unter Verwendung von neun Kalibrierpunkten (vgl. Abbildung 3.21). Aus der Annahme, dass beide Achsen im selben Punkt – dem Korneazentrum (vgl. Abbildung 3.4(b)) – entspringen, lässt sich folgende Gleichung ableiten

$$\mathbf{VA} = \mathbf{R} \cdot \mathbf{OA}, \quad (3.33)$$

R ist die Rotationsmatrix, welche die **OA** in die **VA** überführt. Dementsprechend setzt sich das Gleichungssystem aus drei Gleichungen und neun Unbekannten (den Einträgen

der Rotationsmatrix) zusammen

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (3.34)$$

Für Gleichung 3.33 ergeben sich folgende Gleichungen

$$r_{11} \cdot \mathbf{OA}_x + r_{12} \cdot \mathbf{OA}_y + r_{13} \cdot \mathbf{OA}_z = \mathbf{VA}_x \quad (3.35)$$

$$r_{21} \cdot \mathbf{OA}_x + r_{22} \cdot \mathbf{OA}_y + r_{23} \cdot \mathbf{OA}_z = \mathbf{VA}_y \quad (3.36)$$

$$r_{31} \cdot \mathbf{OA}_x + r_{32} \cdot \mathbf{OA}_y + r_{33} \cdot \mathbf{OA}_z = \mathbf{VA}_z, \quad (3.37)$$

die sich als Gleichungssystem $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ ausdrücken lassen. Dabei sind die Matrix \mathbf{A} als

$$\mathbf{A} = \begin{bmatrix} \mathbf{OA}_x & \mathbf{OA}_y & \mathbf{OA}_z & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{OA}_x & \mathbf{OA}_y & \mathbf{OA}_z & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{OA}_x & \mathbf{OA}_y & \mathbf{OA}_z \end{bmatrix}$$

und die Vektoren \mathbf{b} und \mathbf{x} als

$$\mathbf{b} = \begin{bmatrix} \mathbf{VA}_x \\ \mathbf{VA}_y \\ \mathbf{VA}_z \end{bmatrix} \quad \text{und} \quad \mathbf{x} = [r_{11} \ r_{12} \ r_{13} \ r_{21} \ r_{22} \ r_{23} \ r_{31} \ r_{32} \ r_{33}]^T$$

definiert. Über mehrere Referenzpunkte (mindestens jedoch drei) kann das unterbestimmte Gleichungssystem in ein überbestimmtes Gleichungssystem überführt werden. Dabei wird \mathbf{A} von 3×9 auf $3n \times 9$ ($n \hat{=}$ Anzahl der Referenzpunkte) „aufgebläht“,

$$\mathbf{A} = \begin{bmatrix} \mathbf{OA}_{1,x} & \mathbf{OA}_{1,y} & \mathbf{OA}_{1,z} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{OA}_{1,x} & \mathbf{OA}_{1,y} & \mathbf{OA}_{1,z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{OA}_{1,x} & \mathbf{OA}_{1,y} & \mathbf{OA}_{1,z} \\ \mathbf{OA}_{2,x} & \mathbf{OA}_{2,y} & \mathbf{OA}_{2,z} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{OA}_{2,x} & \mathbf{OA}_{2,y} & \mathbf{OA}_{2,z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{OA}_{2,x} & \mathbf{OA}_{2,y} & \mathbf{OA}_{2,z} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{OA}_{n,x} & \mathbf{OA}_{n,y} & \mathbf{OA}_{n,z} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{OA}_{n,x} & \mathbf{OA}_{n,y} & \mathbf{OA}_{n,z} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{OA}_{n,x} & \mathbf{OA}_{n,y} & \mathbf{OA}_{n,z} \end{bmatrix}$$

während sich der Vektor \mathbf{b}

$$\mathbf{b} = [\mathbf{VA}_{1,x} \ \mathbf{VA}_{1,y} \ \mathbf{VA}_{1,z} \ \mathbf{VA}_{2,x} \ \mathbf{VA}_{2,y} \ \mathbf{VA}_{2,z} \ \dots \ \mathbf{VA}_{n,x} \ \mathbf{VA}_{n,y} \ \mathbf{VA}_{n,z}]^T$$

vergrößert und \mathbf{x} unverändert bleibt. Dies ergibt ein überbestimmtes Gleichungssystem, welches über $\mathbf{x} = \mathbf{A}^+ \cdot \mathbf{b}$ mit

$$\mathbf{A}^+ = \mathbf{V} \cdot \mathbf{S}^{\text{inv}} \cdot \mathbf{U}^T, \quad (3.38)$$

gelöst wird⁷. \mathbf{S}^{inv} wird folgendermaßen gebildet (mit σ_i als Singulärwerte der Matrix \mathbf{A}):

$$\mathbf{S}^{\text{inv}}_{i,j} = \begin{cases} \frac{1}{\sigma_i}, & \text{für } i = j \leq r \\ 0, & \text{sonst} \end{cases}$$

Nach Bestimmung der Pseudoinversen können die Werte von \mathbf{x} und damit die Rotationsmatrix \mathbf{R} bestimmt werden, welche jede berechnete optische Achse in die entsprechende visuelle Achse überführt (vgl. Abbildung 3.21).

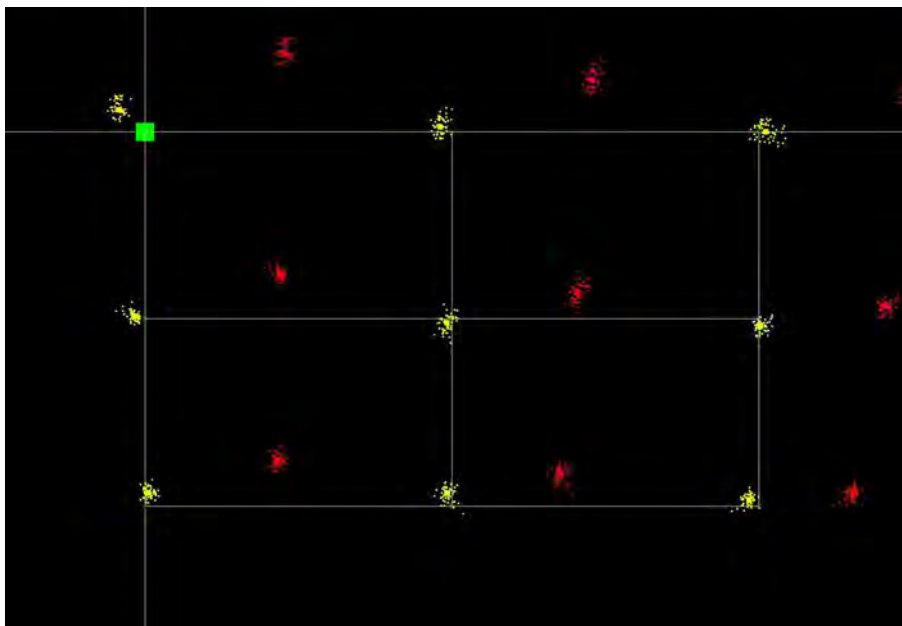


Abbildung 3.21: Darstellung am Monitor für die Benutzerkalibrierung: 9-Punkte Fixationsgitter; rot: Punktswirbel der Fixationen vor der Kalibrierung; gelb: Punktswirbel der Fixationen nach der Kalibrierung.

3.3.2 Extraktion der relevanten Parameter zur Blickrichtungsbestimmung

Die Lokalisation der Augen in den Videobildern basiert auf dem 1993 von *Ebisawa* [73] vorgestellten Ansatz unter Verwendung des *Rote-Augen-Effekts*. Dieser nutzt den Unterschied zwischen nahe (on-axis) und fern (off-axis) der optischen Achse beleuchteten Bildern zur Detektion der Pupillen. Ausgehend von einem solchen Bilderpaar (on-axis/off-axis beleuchtet) läuft die Bildverarbeitung zur Parameterextraktion grundsätzlich folgendermaßen ab:

- *Gaußfilterung* beider Quellbilder
- Bildung des *Differenzbildes* aus den gefilterten Bildern

⁷Für die Bestimmung der inversen Matrix wird die Berechnung der Pseudoinversen mit Hilfe der Singulärwertzerlegung durchgeführt, da die Zeilenvektoren von \mathbf{A} nicht linear unabhängig sind, was zu einer Nichtinvertierbarkeit der Matrix führt. Hierbei wird die Matrix \mathbf{A} in drei Matrizen zerlegt ($\mathbf{A} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T$).

- *Binarisierung* des Differenzbildes mit gegebenem Schwellwert
- Extraktion der *Pupillenandidaten*
- *Klassifizierung* der Pupillenandidaten
- *Pupillentracking*
- Extraktion der *Glints* aus dem off-axis (dark-pupil) Bild
- Erkennung des *Pupillenrandes* (Ellipse fitting)

Die initiale Gaußfilterung dient zur Glättung der Eingangsbilder und soll bei der späteren Kantenfilterung (vgl. Kapitel 3.3.2.5) verhindern, dass unerwünschte Kanten durch Rauschen im Bild detektiert werden. Der Gaußfilter (linearer Glättungsfilter/Tiefpassfilter) verringert Bildrauschen im Differenzbild [118] und verbessert damit die Stabilität und Geschwindigkeit der Kantenerkennung. Zu weiteren Details zur Implementierung der Gaußfilterung in *eyeDrive* sei auf [181] verwiesen.

Anschließend wird das Differenzbild aus zwei aufeinander folgenden Bildern (Subtraktion des off-axis vom zugehörigen on-axis Bild) berechnet (vgl. Abbildung 3.22). Das

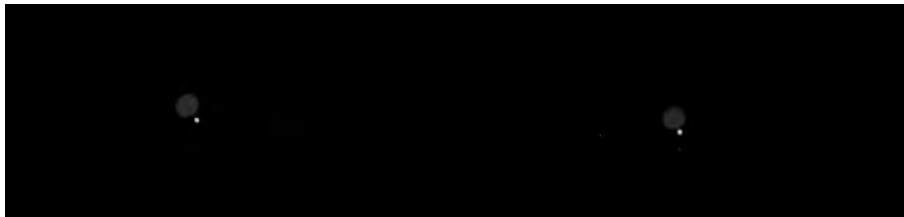


Abbildung 3.22: Differenz aus on-axis (Abbildung 3.18(a)) und off-axis (Abbildung 3.18(a)) beleuchtetem Eingangsbild.

Differenzbild enthält im Idealfall lediglich die hellen Pupillen. Praktisch sind jedoch ebenfalls Beleuchtungsartefakte im Differenzbild enthalten. Diese treten vor allem dann verstärkt auf, wenn sich die Kegel von on-axis und off-axis Beleuchtung nicht vollständig überlappen oder das Licht der aktiven LED Beleuchtung beispielsweise an der Brille des Nutzers reflektiert wird. Diese Artefakte können die Detektion der Pupillen erheblich erschweren und sollen deshalb durch die in Kapitel 3.3.2.3 vorgestellte Klassifikation der Pupillenandidaten weitestgehend eliminiert werden. Die weiteren Schritte werden in den folgenden Abschnitten näher erläutert.

3.3.2.1 Bibliotheken

Das Gesamtsystem ist in C/C++ implementiert und ist größtenteils über Funktionen der Bibliotheken *IPP* (Integrated Performance Primitives) und *OpenCV* der Firma Intel realisiert. *OpenCV* gilt als Standardbibliothek für die Computer Vision und *IPP* bietet eine Vielzahl weiterer Funktionen aus der Bildverarbeitung, welche auf Assembler Ebene speziell auf die Multimedia Register MMX und SSE der Intel Prozessoren zugeschnitten sind und daher im PC Bereich die größtmögliche Effizienz garantieren. Die verschiedenen

Threads, die Benutzerinteraktion und die Darstellung des Live Videos basieren auf Klassen und Funktionen der Bibliothek *Qt*. Für die Darstellung des Videos wird über *Qt* die Low Level Grafik Bibliothek *OpenGL* eingebunden. Die Kontrolle der Kamera erfolgt über die API des Kameraherstellers.

3.3.2.2 Extraktion der Pupillenkandidaten

Die Pupillen sind als helle, zusammenhängende Bereiche in einem sonst nahezu schwarzen (Differenz-)Bild enthalten. Deshalb bietet sich zur Detektion ein *Connected Components* oder *Blob Detection* Algorithmus an. Dazu wird zunächst das Differenzbild binarisiert (einstellbarer aber fester Schwellwert). Anschließend werden zusammenhängende weiße Bereiche – sog. „Blobs“ – erfasst und regelbasiert überprüft, ob es sich um eine Pupille handeln könnte. Zur Unterscheidung zwischen Pupillen und sonstigen Blobs werden einfache Regeln angewendet. Neben der Pupille lassen sich Brillenreflexionen als wahrscheinlichste Blobs nennen, welche jedoch wesentlich heller sind, als die durch den Redeye-Effekt ausgeleuchtete Pupille. Weiterhin kommen Blobs mit sehr vielen oder wenigen Pixeln und mit einem ungünstigen Breite/Höhe-Verhältnis nicht als Pupillenkandidaten in Frage und werden deshalb vor der Klassifizierung verworfen.

Zur Beschleunigung der Suche im Binärbild wird ein Supersampling Bild berechnet. Der Faktor 8^8 hat sich experimentell als niedrigster Faktor erwiesen, bei dem eine Erkennung von sehr kleinen Pupillen noch zuverlässig erfolgt. Dabei verringert sich der Aufwand zur Überprüfung auf $1/64$. Die Grauwerte im Supersampling Bild entsprechen

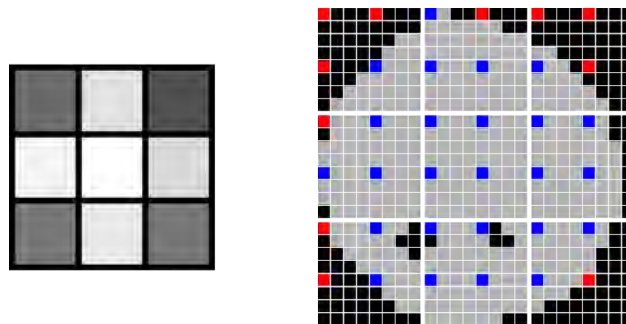


Abbildung 3.23: Supersampling mit Faktor 8. Links: 3x3 Pixel im Supersampling Bild. Rechts: Entsprechend 24x24 Pixel im originalen Bild. Startpunkte für die genauere Untersuchung markiert in rot für Pixel außerhalb des Blobs bzw. blau für Pixel innerhalb des Blobs.

den Mittelwerten der korrespondierenden Bereiche im Originalbild. Überschreitet der Grauwert eine Schwelle, erfolgt eine genauere Überprüfung. Die Schwelle ist aufgrund der Mindestgröße von Pupillenkandidaten auf 80 festgelegt, was bei Farbwerten von 0 und 255 im Binärbild ca. $1/3$ entspricht. Sofern also im 8x8 Pixel Ausschnitt mehr als ein Drittel der Pixel weiß sind, erfolgt eine genauere Überprüfung beginnend bei vier Startpunkten im 8x8 Pixel Ausschnitt (farblich markiert in Abbildung 3.23). Enthält ein Ausschnitt nur den Rand einer Pupille, wird die Schwelle erst im benachbarten Ausschnitt mit dem größeren Teil der Pupille überschritten.

⁸1x1 Pixel im Supersampling Bild entspricht 8x8 Pixel im Originalbild.

Das Binärbild enthält lediglich die Grauwerte 0 und 255. Um ein Pixel während der Suche als *besucht* zu markieren, wird zu Beginn das erste Bit des Farbwertes jedes Pixels gelöscht (Farbwerte 0 bis 127). Zur Blobdetektion kommt die Funktion `FloodFill` – mit der Farbe 255 und den jeweils genauer zu prüfenden vier Pixeln als Startpunkten – zum Einsatz. Der Algorithmus verwirft Pixel mit dem Farbwert 0, da sie keine nützliche Information enthalten. Pixel mit dem Farbwert 255 sind bereits Teil eines gefundenen Blobs (Farbwert während der Erkennung gesetzt). Der Farbwert 127 repräsentiert einen (weißen) Pixel, der Teil eines noch nicht geprüften Blobs ist und dient als Startpunkt. Da die verwendete Funktion Informationen über den Verlauf in einer Struktur speichert, können Anzahl der Pixel sowie Breite/Höhe eines Blobs direkt übernommen werden. Die X-Koordinate des Zentrums bestimmt sich aus der Summe der X-Koordinaten in der Struktur und der halben Breite; Entsprechendes gilt für die Y-Koordinate. Die gefundenen Blobs werden zur weiteren Verarbeitung – Beurteilung, ob es sich dabei um die gesuchte(n) Pupille(n) handelt – an die im folgenden Kapitel beschriebene Klassifikation übergeben. Dieses Vorgehen ist nicht zur Erkennung der Glints in den Pupillen geeignet (vgl. Kapitel 3.3.2.6), da diese zu klein sind und durch das beschriebene Suchraster fallen.

3.3.2.3 Klassifizierung der Pupillenkandidaten

Zur Beurteilung, ob es sich bei einem gefundenen Blob tatsächlich um eine Pupille handelt, kommt ein Kontrast-basierter Klassifikator zum Einsatz, der auf dem konstanten Kontrastverhältnis zwischen Pupille und Iris beruht. [127] untersuchte verschiedene Verfahren des maschinellen Lernens (Support Vector Machine (SVM), Neuronale Netze (NN), Principal Components Analysis (PCA)) hinsichtlich ihrer Tauglichkeit für die vorliegende Problemstellung. Im Rahmen dieser Untersuchung zeigte die Klassifikation auf Basis von SVM die besten Klassifikationsleistungen, jedoch ist dieser Ansatz zu wenig rechenzeiteffizient, um hohe Datenraten im laufenden Betrieb sicherzustellen.

Der Klassifikator bestimmt den Kontrast zwischen Pupille und Iris in einem 40x40 Pixel großen Ausschnitt um jeden gefundenen Blob im dunklen Kamerabild (Abbildung 3.24, links) und im binärisierten Differenzbild (Abbildung 3.24, rechts). Hierbei wird zunächst im Binärbild die Pixelzahl innerhalb des Blobs bestimmt und angenommen, dass sich im Bildausschnitt um den Blob wenig bis keine weiteren Artefakte befinden. Da

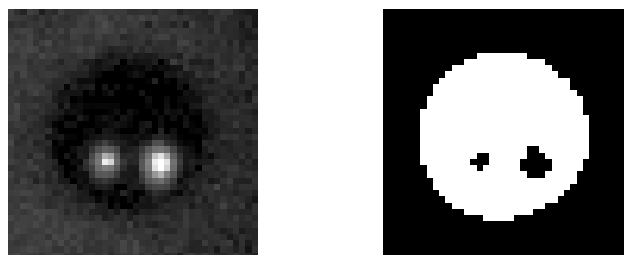


Abbildung 3.24: Basisbilder für den Kontrast basierten Klassifikator.

die Farbwerte entweder 0 oder 255 betragen, ist die Zahl der Pixel innerhalb des Blobs $1/255$ der Summe der Grauwerte. Die Zahl der Pixel außerhalb des Blobs berechnet sich aus der Differenz der vorhandenen ($40 * 40 = 1600$) und der Anzahl der inneren Pixel.

Danach erfolgt die Maskierung des Ausschnitt aus dem dunklen Kamerabild mit Hilfe des Binärbildes (vgl. Abbildung 3.25 links). Aus der Summe der verbleibenden Pixel und der berechneten Anzahl weißer Pixel wird der innere Mittelwert berechnet. Nach Invertierung des Binärbildes berechnet sich der äußere Mittelwert analog (siehe

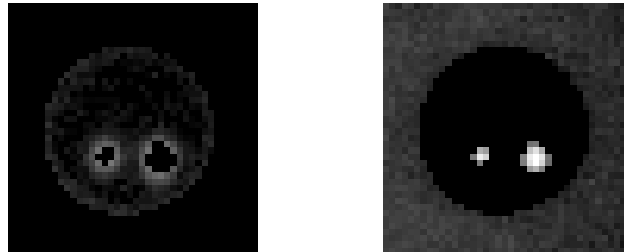


Abbildung 3.25: Inneres und äußeres maskiertes Bild.

Abbildung 3.25 rechts) ohne Beachtung der hellen Glints. Abschließend wird über Formel 3.39 der Michelson-Kontrast als Ergebnis der Klassifizierung berechnet,

$$\frac{\text{äußerer Mittelwert} - \text{innerer Mittelwert}}{\text{äußerer Mittelwert} + \text{innerer Mittelwert}} \quad (3.39)$$

wobei anzumerken ist, dass wegen der nahezu schwarzen Pupille im dunklen Bild der äußere Mittelwert als Maximum zu betrachten ist. Die Erkennungsleistung ist (experimentell bestimmt) der des SVM-basierten Klassifikator aus [127] ebenbürtig.

3.3.2.4 Pupillentracking

Augenbewegungen und damit die Position der Augen im Videobild unterliegen zufälligen Mustern. Die reinen Augenbewegungen sind zusätzlich von Kopfbewegungen überlagert. Zur Verfolgung der Pupillenposition wurde deshalb ein regelbasierter Algorithmus entwickelt, der die aktuellen Kandidaten den früheren Pupillenpositionen zuordnet. Der Algorithmus basiert auf folgenden (empirisch bestimmten) Einschränkungen:

- maximaler/minimaler Augenabstand: 200-460pix
- maximale Augenbewegung: 30pix/Bild

Jeder Proband hat grundsätzlich einen konstanten Augenabstand, der jedoch (im Videobild) im Rahmen der angegebenen Toleranz variieren kann. Der Augenabstand im Videobild variiert dann relativ stark, wenn der Nutzer den Kopf relativ zur Kamera verdreht. Die Begrenzung erlaubt es, ungültige Blobs außerhalb der (im Folgenden beschriebenen) Bounding Box auszuschließen. Zur Überprüfung der zweiten Randbedingung wird in jedem Frame die Augengeschwindigkeit berechnet. Mit Berücksichtigung von obigem Maximalwert ist sichergestellt, dass die Augen immer scharf abgebildet werden.⁹ Somit können alle Blobs außerhalb einer Kreisfläche mit 30pix Radius um die letzte Pupillenposition von der Bestimmung der aktuellen ausgeschlossen werden.

⁹Bei schnelleren Augenbewegungen ist bei der gewählten Bildrate die Bewegungsunschärfe zu hoch und die Blob Kandidaten können im Differenzbild nicht mehr robust gefunden werden.

Zur Bestimmung der aktuellen Position dienen die im Differenzbild detektierten Blobs. Erfüllen mehrere Blobs obige Bedingungen, wird der Blob ausgewählt, der vom Klassifikator (vgl. Kapitel 3.3.2.3) die höchste Wahrscheinlichkeit erhielt und am besten mit der letzten Augenposition korreliert. Der regelbasierte Algorithmus liefert so gute Ergebnisse (Genauigkeit und Rechenzeit), da lediglich wenige Kandidaten vorliegen.

Weiterhin muss für die gefundenen Pupillen bekannt sein, ob es sich um eine linke oder eine rechte Pupille handelt. Dieses Labeling der Pupillen ist vor allem dann schwierig, wenn lediglich eine Pupille gefunden wurde. Dafür kommen unterschiedliche Bounding Boxen zum Einsatz. Für jede Pupille existieren je zwei Bounding Boxen, deren Größe sich kontinuierlich entsprechend der Augenbewegungen anpasst. Gleichzeitig dienen diese Bounding Boxen zur Umsetzung obiger Tracking Regeln. Ein Blob wird erst dann als Pupille betrachtet, wenn er sich innerhalb der Bounding Boxen befindet. Die Kennzeichnung linkes oder rechtes Auge ist möglich, sobald erstmals beide Augen gefunden und die entsprechenden Labels zugewiesen wurden.

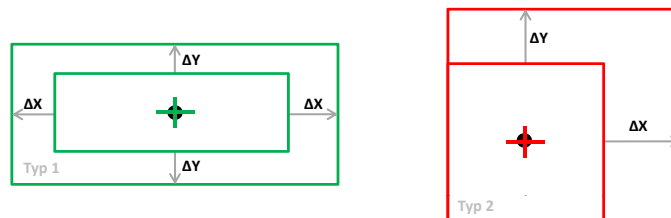


Abbildung 3.26: Vergrößerung der Bounding Boxen bei nicht gefundenen Pupillen.

Als Bounding Boxen dienen drei Rechtecke (linke Pupille, rechte Pupille, unbekannte Pupille), die entsprechend der aktuellen Augenbewegungen kontinuierlich ihre Breite und Höhe ändern. Jede Bounding Box ist dabei immer so klein wie möglich, da sich deren Größe direkt auf den Bildbereich auswirkt, in dem die Bild(vor)verarbeitung durchgeführt wird (vgl. Kapitel 3.3.2.7). Die Anpassung der Größe der Bounding Box erfolgt abhängig davon, wie schnell bzw. weit sich die Pupillenkandidaten in den letzten Frames bewegt haben und wie lange eine Pupille nicht zu detektieren war. Diese Faktoren müssen sehr sorgfältig festgelegt sein, um den besten Kompromiss zwischen Größe der Bounding Box und Tracking Qualität zu finden.

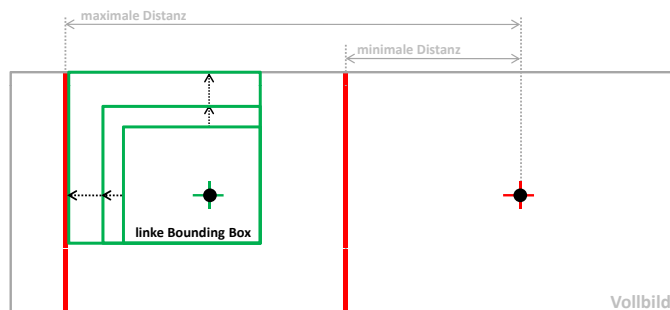


Abbildung 3.27: Verhinderung von Überschneidungen der beiden Bounding Boxen.

Für den Fall, dass lediglich eine Pupille detektierbar ist, jedoch Daten zu ihrer Bewegung aus den letzten Frames vorhanden sind, wird die Bounding Box der zweiten Pupille in ihrer Breite und Höhe dementsprechend in Richtung und Größe um δX bzw. δY vergrößert (vgl. Abbildung 3.26, Typ 2). Andernfalls vergrößern sich Höhe und Breite der Bounding Box in beide Richtungen (vgl. Abbildung 3.26, Typ 1). Um Überschneidungen beider Bounding Boxen zu vermeiden, kommt der minimale und maximale Abstand beider Pupillen (erste Regel) in horizontaler Richtung zum Einsatz, wenn mindestens ein Auge gefunden ist (vgl. Abbildung 3.27). Abhängig davon, ob und wie viele Pupillen im vorherigen Bilderpaar detektiert sind, lassen sich folgende Fälle unterscheiden:

- **Keine gefunden:** Es werden die Pupillenkandidaten ausgewählt, die von der Entfernung zwischen ihnen und von ihrem Klassifikationsgewicht her am ehesten einem Augenpaar entsprechen.
- **Links gefunden, rechts nicht:** Für die linke Pupille wird der Kandidat mit dem geringsten Abstand zur alten Position gewählt. Wurde dieser gefunden, wird nach einem weiteren Pupillenkandidaten gesucht, der vom Abstand und dem Klassifikationsgewicht am ehesten der rechten Pupille entsprechen könnte.
- **Rechts gefunden, links nicht:** Entsprechend zu links gefunden, rechts nicht.
- **Beide gefunden:** Es werden die Pupillenkandidaten mit dem minimalsten Abstand zur vorherigen Position gewählt.

Ist eine Pupille gefunden, erfolgt eine Beschränkung des Suchbereichs für das nächste Bilderpaar auf einen Bereich um die Position. Wird die Pupille nicht mehr gefunden, wächst der Bereich von einem Bild zum nächsten bis zum gesamten Bildausschnitt an.

3.3.2.5 Pupillenranderkennung

Die Pupillenranderkennung ermittelt die Ellipse, die den Pupillenrand am besten beschreibt. Die Ellipsenerkennung dient dazu, den Mittelpunkt und Randpunkte der Pupille anhand der Randlinie genauer zu bestimmen. Die Berechnung des Gaze-Vektors (vgl. Kapitel 3.3.1.2) benötigt diese Parameter. Dazu ermittelt ein Kantendetektor im Bereich um die gefundene Pupille enthaltene Kanten. Kantenpunkte, die zu weit vom Pupillenzentrum entfernt sind, werden eliminiert. Die Ermittlung der Kantenpunkte basiert auf dem 1986 vorgestellten *Canny Edge Detector* [53], der bis heute zu den Standardkantendetektoren in der industriellen Bildverarbeitung zählt. Er vereint folgende Eigenschaften:

1. **Gute Detektion:** Der Algorithmus markiert so viele tatsächliche Kanten wie möglich im Bild.
2. **Gute räumliche Lokalisierung:** Markierte Kanten liegen so nah wie möglich an den tatsächlichen Kanten.
3. **Eine einzige Antwort auf eine Kante:** Eine tatsächliche Kante wird nur ein einziges Mal markiert und falls möglich werden keine falschen Kanten durch Rauschen erzeugt.

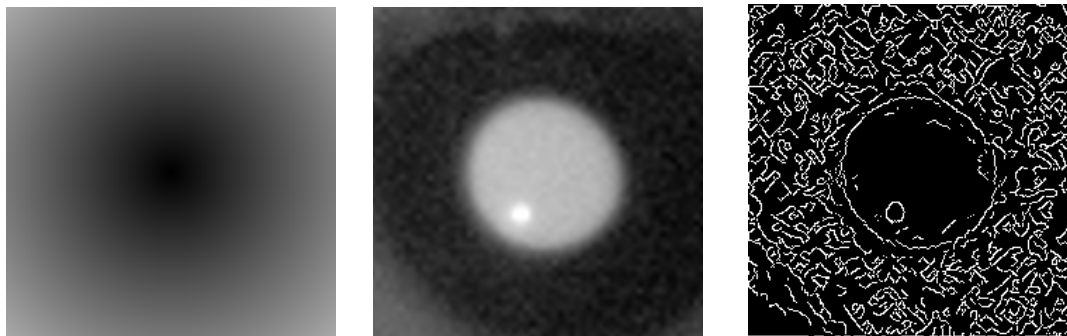
Zur Bestimmung der Kantenpunkte wird das Bild zunächst mit einem Gauß-Filter geglättet, um Verfälschungen zu unterdrücken. Anschließend werden im geglätteten Bild Betrag und Richtung des Gradienten berechnet.

$$M[i, j] = \sqrt{G_x[i, j]^2 + G_y[i, j]^2}, \quad (3.40)$$

$$\theta[i, j] = \begin{cases} \arctan\left(\frac{G_y[i, j]}{G_x[i, j]}\right), & \text{falls } G_x[i, j] \neq 0, \\ 0, & \text{falls } G_x[i, j] = 0 \text{ und } G_y[i, j] = 0, \\ \frac{\pi}{2}, & \text{falls } G_x[i, j] = 0 \text{ und } G_y[i, j] \neq 0 \end{cases}, \quad (3.41)$$

Zur Bestimmung des Kantenverlaufs werden danach im Gradientenbild lokale Maxima $M[i, j]$, sog. Bergrücken bestimmt. Dieser Prozess der Verdünnung wird Nicht-Maximum Unterdrückung genannt und liefert als Ergebnis eine Kante von der Breite eines Pixels entlang eines Bergrückens. Um die Anzahl falsch detektierter Kanten zu minimieren sowie einzelne Kantenpunkte zu Konturen zu verbinden, wird abschließend ein doppeltes Schwellwertverfahren angewandt. Da sich dabei die Wahl der Schwellwerte deutlich auf die Qualität der detektierten Kanten auswirkt, verwendet *eyeDrive* ein adaptives Verfahren zur Schwellwertbestimmung [181].

Ausgangspunkt der Kantenbestimmung ist das in Abbildung 3.28(a) gezeigte Hilfsbild, dessen Grauwerte die Entfernungen zum Mittelpunkt des Bildes repräsentieren. Da die Verarbeitung auf einem Bildausschnitt stattfindet, der von 60x60 Pixel auf 240x240 Pixel vergrößert wurde¹⁰, ist der Mittelpunkt des Hilfsbildes bei Position (120,120). Abbildung 3.28 zeigt neben einem Ausschnitt eines hellen Kamerabildes das entsprechende Ergebnis des Canny Filters für die weitere Analyse. Da die Größe der Pupille bereits



(a) Hilfsbild für Ellipseerkennung.

(b) On-axis beleuchtetes Kamerabild.

(c) Ergebnis des Canny Filters.

Abbildung 3.28: Hilfsbild, Auge im Kamerabild und Ergebnis des Kantendetektors.

aus der Blob Detektion (vgl. Kapitel 3.3.2.2) bekannt ist, kann eine minimale und eine maximale Entfernung des Pupillenrandes abgeschätzt werden. Aufgrund der Tatsache, dass die Farbwerte im Hilfsbild die Entfernung zum Mittelpunkt angeben, reduziert sich die Auswahl des Zwischenraumes auf einen Funktionsaufruf mit einem unteren

¹⁰Ein 60*60 Pixel Ausschnitt um die detektierte Pupille wird ausgeschnitten und mit einem Faktor von 4 kubisch interpoliert. Das Zentrum der Pupille liegt dabei in der Mitte des ausgeschnittenen Bildbereiches.

und oberen Schwellwert. Dabei dienen die minimale und die maximale Entfernung als Farbgrenzwerte. Alle Farbwerte darunter und darüber werden auf 0 gesetzt und man erhält einen Ring aus dem Hilfsbild (vgl. Abbildung 3.29(a)).

Die Maske dient (mittels einer logischen Und-Verknüpfung auf das Canny Bild) einer Eliminierung der meisten Artefakte (vgl. Abbildung 3.29(c)). Die restlichen Randpunkt-

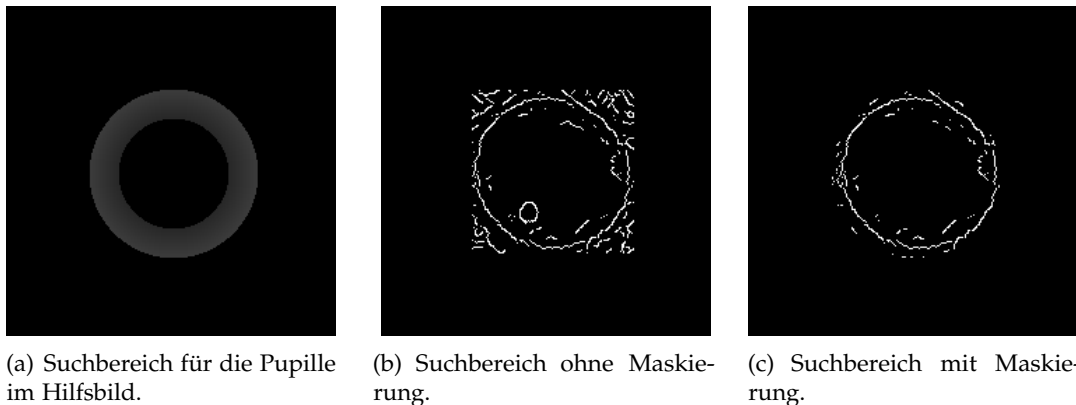


Abbildung 3.29: Ergebnis der Maskierung des Suchbereiches.

kandidaten werden gesammelt und jeweils vier Pixel darauf überprüft, ob sie schwarz sind. Trifft dies zu, können die Pixel übersprungen werden; andernfalls erfolgt eine einzelne Überprüfung. Ohne Maskierung kann das System Artefakte in den Ecken des Suchbereichs nicht eliminieren (vgl. Abbildung 3.29(b)). Auf alle übrigen, relevanten Kantenpunkte erfolgt mit Hilfe des Algorithmus aus Open CV ein Ellipse Fit, der die Ellipse bestimmt, die den kleinsten quadratischen Fehler zu allen Punkten liefert. Der gesuchte Pupillenmittelpunkt wird als Mittelpunkt der gefitteten Ellipse bestimmt.

3.3.2.6 Glintextraktion

Die Glintextraktion erfolgt unter der Voraussetzung, dass das Tracking der Pupille erfolgreich war, für jede Pupille getrennt. Der Blob Detektor zur Glintextraktion basiert auf *Vandevennes* Scanline-Floodfill-Algorithmus [238]. Als Berechnungsgrundlage dient das dunkle Kamerabild (siehe Abbildung 3.18(b)). Ein 60x60 Pixel großer Ausschnitt um die gefundene Pupille herum wird kopiert und binarisiert. Für jedes Pixel im Binärbild ermittelt das System, ob dieses die Farbe weiß hat und noch nicht überprüft ist. Trifft beides zu, wird die Position auf dem Stack gespeichert.

Solange der Stack nicht leer ist, sucht der Algorithmus ausgehend von der obersten Position das untere Ende der Scanline, indem er solange die y-Koordinate erhöht, bis entweder das untere Ende des Bildes (bzw. der ROI) erreicht oder das nächste Pixel nicht mehr weiß ist. Von dieser Position aus erfolgt die Überprüfung der linken und rechten Nachbarspalte. Sind unverarbeitete weiße Pixel vorhanden, wird die Pixelposition auf dem Stack gespeichert. Über zwei Flags ist sichergestellt, dass die Nachbarspalte nur einmal auf dem Stack gespeichert ist. Für jedes gültige Pixel innerhalb eines Blobs erhöht sich die Gesamtanzahl der gespeicherten Pixel. Daraus lässt sich später das Zentrum

berechnen. Die Grenzen des Blobs ergeben sich aus den minimalen bzw. maximalen X/Y-Koordinaten der überprüften Pixel. Da der Suchbereich und die Glintblobs (verglichen mit den Pupillenblobs) sehr klein sind, bietet dieses Vorgehen einen enormen zeitlichen Vorteil gegenüber der Verwendung des in Kapitel 3.3.2.2 vorgestellten Blob Detektors.

3.3.2.7 Region of Interest (ROI)

Basierend auf den gefundenen Pupillenpositionen erfolgt die Berechnung der Bildverarbeitungs- und Detektionsalgorithmen auf unterschiedlichen Regions of Interest (ROI) der aufgezeichneten Videodaten. Dies hat den Vorteil, dass die Menge an zu verarbeitenden Daten auch bei hohen Bildraten möglichst gering ist. Ist die Position der Pupille aus dem letzten Bild bekannt, kann man die Bildverarbeitung und die Suche nach den Pupillenkandidaten auf einen kleinen Bereich eingrenzen. Das Potential dieser Vorgehensweise besteht darin, dass die Bild(vor-)verarbeitung mit Gaussfilter, binarisiertem Differenzbild und Suche der Pupillenkandidaten nicht mehr im gesamten Bild stattfindet. Das Vollbild wird nur dann untersucht, wenn keine Pupillen gefunden wurden oder das Tracking mehrere Bilder hintereinander nicht erfolgreich war.

3.3.2.8 Autofokus

Das vorgestellte System verwendet eine nicht-invasive remote Videokamera, weshalb natürliche Kopfbewegungen ein Problem darstellen. Diese führen zu Veränderungen des Abstandes zwischen Kamera und Objekt, weshalb eine Defokussierung des Bildes bei einem Fixfokus-Setup unvermeidlich ist, sobald es dem Nutzer erlaubt sein soll, sich „frei“ vor der Kamera zu bewegen. Eine bestmögliche Bildqualität und -schärfe ist jedoch erforderlich, um rauschfreie Blickdaten berechnen zu können. Daher wird eine möglichst latenzfreie Autofokus-Funktionalität benötigt, um die Bildschärfe zu verbessern.

Zur Bestimmung der Bildschärfe kommt ein Ansatz basierend auf Kantenbildern zum Einsatz [190]. Die zugrunde liegende Idee ist folgende: Ein Bild mit vielen detektierbaren Kanten ist schärfer, als ein Bild mit wenig Kanten. Zur Detektion wird der Canny Detektor angewendet (vgl. Kapitel 3.3.2.5) und in der ROI um das Auge die Anzahl der Kantenpixel berechnet¹¹. Für eine rein numerische Maximumsuche wird häufig ein „Verfahren des steilsten Anstiegs“ verwendet, wobei der aktuelle Schärfeindex mit dem vorhergehenden verglichen wird und eine entsprechende Änderung der Linsenkonstellation erfolgt. Dieses Vorgehen wiederholt sich so lange, bis sich eine Verschlechterung der Fokussierung (d.h. weniger Kantenpixel) einstellt. Im Anschluss wird die letzte Änderung der Linsenkonstellation revidiert.

Bevor der Autofokus starten kann, durchläuft er einmal den gesamten Schärfebereich, um das Minimum der erreichbaren Schärfe zu bestimmen. Überschreitet der aktuelle Schärfewert im Betrieb das 1,5fache dieses Minimums nicht, wird eine große Schrittweite zur schnellen Fokussierung verwendet. Überschreitet der aktuelle Schärfeindex das dreifache des Minimums, setzt der Algorithmus die Schrittweite herab, um das globale (Schärfe-)Maximum nicht zu überspringen. Sobald sich der neu berechnete Schärfewert dreimal in Folge verschlechtert, erfolgt eine Invertierung der Fokussierungsrichtung.

¹¹Als Maß wird hierbei der absolute Gradient im Kantenbild berechnet: $\bar{x} = \sum_{i=2}^m \sum_{j=2}^n |(x_{i,j} - x_{i+1,j})|$.

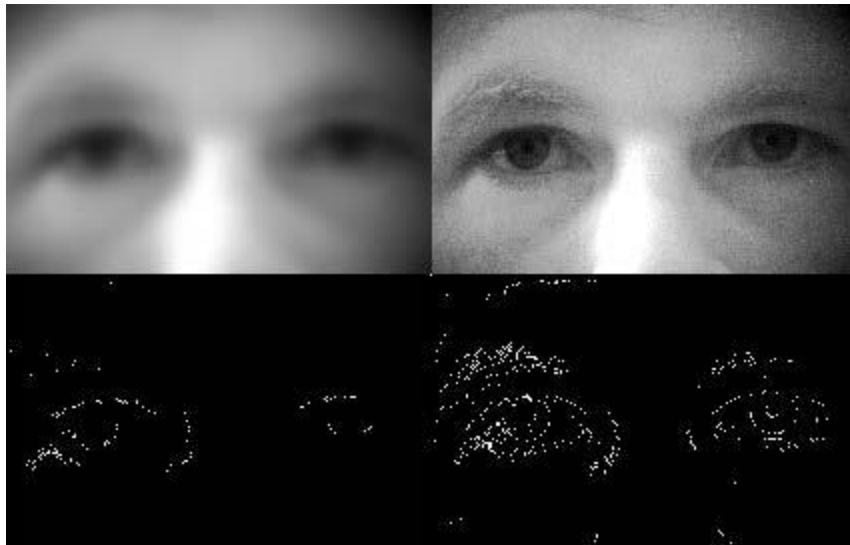


Abbildung 3.30: Kantenbilder mit unterschiedlicher zugrunde liegender Bildschärfe.

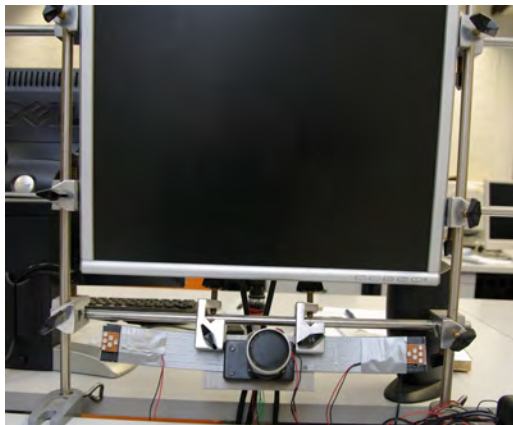
Dieses Vorgehen unterdrückt Rauschen und stellt sicher, dass die Fokusebene stabil um das Maximum pendelt. Die Schwankungen können bei kleinster Schrittweite vernachlässigt werden, da sich das Objekt dennoch im (Tiefen-)Schärfebereich befindet. Das Autofokus System ist fundamentaler Bestandteil des in Kapitel 7.3 vorgestellten Systems zur Blickrichtungsrekonstruktion mehrerer Nutzer in Robotikszszenarien (Abbildung 7.8).

3.3.3 Hardwareaufbau

Die folgenden Abschnitte stellen den konzipierten Hardwareaufbau vor. Wie bereits in den vorangegangenen Kapiteln 3.3.1 und 3.3.2 beschrieben, verwendet das System zwei unterschiedlich beleuchtete Bildtypen zur Bestimmung der benötigten Merkmale. Die Bilder werden alternierend nah und fern der optischen Achse der Kamera mit infrarotem Licht beleuchtet (vgl. Abbildungen 3.18 und 3.31). Die Steuerung der Beleuchtung verarbeitet das Triggersignal der Kamera und schaltet synchron zur Aufnahme jedes Bildes die Beleuchtung abwechselnd zwischen on- und off-axis LEDs hin und her. Dazu sendet die Kamera beim Start des Bildauslesens ein Triggersignal, das die Steuerung nutzt, um die einzelnen Beleuchtungseinheiten alternierend zu aktivieren. Die folgenden Abschnitte beschreiben die einzelnen Komponenten des finalen, integrierten Hardwareaufbaus. Die Umsetzung des Aufbaus erfolgte im Rahmen von [26].

3.3.3.1 Peripherie

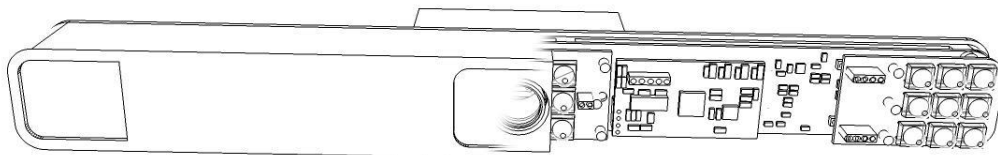
Zur Steuerung und Parametrierung des Gesamtsystems wird der Mikrokontroller *MSP430* der Firma Texas Instruments eingesetzt. Diese Controllerfamilie zeigt unter anderem die Vorteile einer geringen Stromaufnahme, vielseitiger Timerbaugruppen sowie eine geringe Anzahl zusätzlicher Bauelemente [38, 224]. Als Steuerzentrale dient das Derivat *MSP430FG439*, das alle für die Schaltung wichtigen Funktionen beinhaltet:



(a) Initialer Labortestlaufbau mit externer Beleuchtungssteuerung.



(b) Aufbau mit externer Beleuchtungssteuerung.



(c) CAD-Explosionszeichnung (Hauptplatine, Beleuchtungseinheiten, Kamera) des finalen komplett integrierten Hardwareaufbaus.



(d) Finaler komplett integrierter Hardwareaufbaus in mobilem Testsetup mit Laptop Computer.

Abbildung 3.31: Hardwareaufbau des remote Eye Tracking Systems *eyeDrive*.

- zwei unabhängige Timerbausteine mit jeweils drei Capture/Compare Registern
- 12 A/D-Wandler-Eingänge
- möglicher Anschluss zweier unterschiedlicher Quarzoszillatoren (32 kHz, 8MHz)
- UART-Schnittstelle
- JTAG-Programmierschnittstelle sowie BSL (Boot-Strap-Loader)

3. Kamera-basierte Fahrerüberwachung

- Zwei Interrupt-Ports
- 20 IO-Pins

Zur Kommunikation des Mikrocontrollers mit dem USB-Port des PCs ist weiterhin ein parametrierbares Interface (FT232R) integriert. Zu den wichtigsten Einstellmöglichkeiten zählen die Anforderung des benötigten Stroms (bis 500 mA) vom USB-HOST, das Einstellen der Stromversorgung als BUS-Powered und die Vergabe von Seriennummern.

3.3.3.2 Kamera

Als Kamera kommt die *FireFly MV USB* der Firma Point Grey [186] zum Einsatz (siehe Abbildung 3.32). Die Kamera verfügt über einen CMOS Sensor (Micron Mt9v022) mit



Abbildung 3.32: Kamera Point Grey FireFly MV.

752*480 Pixeln. Bei maximaler Auflösung erreicht die Kamera eine Bildfrequenz von 60 Hz. Zudem unterstützt sie die Region-of-Interest (ROI) Funktion. Dadurch ist es möglich, lediglich die Augenpartie auszulesen, was beispielsweise bei einer Auflösung von 752x240 Pixel eine Bildfrequenz von bis zu 110 Hz ermöglicht. Für höhere Bild-/Datenraten muss der vertikale Auslesebereich weiter eingeschränkt werden. Der Anschluss an den PC erfolgt über USB 2.0. Weiterhin besitzt die Kamera fünf GPIO Pins die mit Pegeln von 3.3V sowie 5V arbeiten und folgende Funktionalität zur Verfügung stellen: externe Stromversorgung der Kamera, Spannungsversorgung für weitere Anschlüsse, 2 Eingang/Ausgang Pins, Masse (GND). Über diese Pins wird der für die Beleuchtungssteuerung benötigte Kameraktakt ausgelesen.

3.3.3.3 Beleuchtung und Filter

Damit das System weitestgehend unabhängig von der Umgebungsbeleuchtung arbeiten kann und der Nutzer gleichzeitig nicht durch die Beleuchtung beeinträchtigt wird, arbeitet das System mit einer infraroten Szenenbeleuchtung. Als Lichtquelle nutzt das System die LEDs *FL850-03-08* der Firma Marubeni [151]. Diese emittieren einen schmalbandigen Wellenlängenbereich mit einem Maximum bei 850 nm und einem Abstrahlwinkel von ca. +/- 15°. Die Stromversorgung der LEDs erfolgt mittels dem Konstantstromtreiber-IC *LTC3217* der Firma LINEAR. Der Strom ist zur Helligkeitsregelung der LEDs mittels PWM (Puls-Weiten-Modulation) zwischen 0% und 100% regelbar.

Zur Verbesserung der Bildqualität ist eine möglichst hohe Leuchtintensität nötig. Da bei IR Beleuchtung kein Pupillenlichtreflex stattfindet, muss sichergestellt sein, dass die dissipierte Lichtleistung die gültigen Grenzwerte unterschreitet, um eine dauerhafte Schädigung der Netzhaut zu vermeiden. Bei der Bestrahlung mit infrarotem Licht

ergeben sich laut BGV B9 [35] sowie IEC TR 60825-9 [79] insgesamt drei Gefahrenpotentiale: Linsentrübung (Gefährdung durch Bestrahlung) sowie thermische Gefährdung der Netzhaut (Gefährdung durch Strahlungsdichte) und des Auges (Gefährdung durch Bestrahlungsstärke). Berechnungen zur Auslegung der Leuchtstärke zeigen, dass das System den Grenzwert zur Strahlenexposition unterschreitet [28].

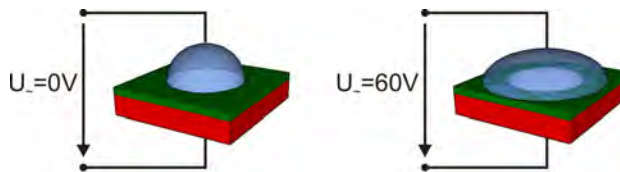
Zur Unterdrückung äußerer Störeinflüsse durch Umgebungslicht mit einer Wellenlänge kleiner als 850 nm kommen Infrarotfilter des Typs *Solaris S306* zum Einsatz. Die Filter absorbieren ca. 4% des Lichts bei 850 nm. Da das IR-Licht der LEDs zwei mal gefiltert wird, entsteht ein Verlust von 7,84%. Die verwendeten Filter besitzen einen Bandpasscharakter. Licht mit einer Wellenlänge bis 2200 nm kann den Filter passieren wobei der Bildsensor lediglich bis zu einer Wellenlänge von 950nm sensitiv ist.

3.3.3.4 Variable Anpassung der Fokusweite

Zur Realisierung der in Kapitel 3.3.2.8 vorgestellten Autofokusfunktionalität kommt die Flüssiglense *Arctic 314* der Firma Varioptic zum Einsatz (vgl. Abbildung 3.33(a)). Das Funktionsprinzip einer Flüssiglense wird in Abbildung 3.33(b) gezeigt. Zum Betreiben der Linse ist zusätzlich das *Liquid Lense Driver Board 60 LL3* (hier in PWM-Ausführung) nötig. Dies wandelt das vom Mikrocontroller gesendete PWM-Steuersignal in eine äquivalente Wechselspannung bis 60 V. Die Flüssiglense besteht aus zwei Flüssigkeiten gleicher Dichte



(a) Flüssiglense Arctic 314.



(b) Funktionsprinzip einer Flüssiglense.

Abbildung 3.33: Flüssiglense Arctic 314 und deren Funktionsweise.

[33, 87]. Im spannungsfreien Zustand verharrt die Linse im Ruhezustand und weist die größte Wölbung bzw. Brechkraft auf. Wird eine Spannung angelegt, ändert sich die Oberflächenspannung der Flüssigkeiten und die Linse beginnt mit der Verformung, was sich in einer Verringerung der Brechkraft auswirkt. Flüssiglinsen wurden primär für Handykameras entwickelt, um eine automatische Fokussierung auf kleinstem Raum zu ermöglichen. Außerdem bieten sie – im Vergleich zu motorbetriebenen, handelsüblichen Fokussiersystemen – folgende Vorteile: geringer Energieverbrauch (im Vergleich zu diversen Motoren, wie beispielsweise Ultraschallmotoren von professionellen Objektiven für Spiegelreflexkameras), schnelle Ansprech- und Reaktionszeiten, lange Lebenszeit aufgrund des Fehlens mechanisch beweglicher Teile und geringer Bauraum (Höhe: 2,5 mm; Durchmesser: 10,5 mm; Linsendurchmesser: 3 mm).

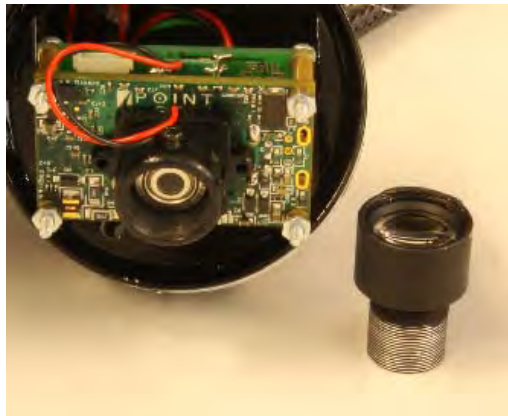


Abbildung 3.34: Flüssiglinse integriert in den Objektivhalter des Hardwareaufbaus aus Abbildung 3.31(b). Hinter der Kamera ist die Linsensteuerung auf einer Huckepackplatine montiert.

3.3.3.5 Gehäuse

Sämtliche (hardwaretechnisch realisierten) Funktionen des Eye Trackers sind in einem Gehäuse untergebracht, welches über nur ein USB Kabel (für die Steuerung, Stromversorgung und Übertragung der Bilddaten) mit dem Aufzeichnungsrechner verbunden wird. Um ein Höchstmaß an Flexibilität und Mobilität zu erreichen, sind die Abmessungen des Versuchsaufbaus möglichst gering gehalten. Die Größe ist vor allem durch den benötigten Abstand der beiden off-axis LED-Gruppen von einander vorgegeben (bei geringerem Abstand wären die beiden Glints im Videobild äußerst schlecht separierbar). Den kompletten Hardwareaufbau zeigt Abbildung 3.31(d).

Neben dem einfachen Anschluss über ein USB Kabel liegt ein weiterer Vorteil des integrierten Aufbaus in den definierten Abmessungen. Die relative Lage der Komponenten zueinander kann sich nicht mehr ändern. Das vereinfacht die Inbetriebnahme eines neuen Versuchssetups deutlich, da die komplizierte Vermessung entfällt. Es müssen lediglich die einzelnen Anzeigeflächen über den AOI Manager (vgl. Kapitel 3.2.3) kalibriert werden, um ein neues Setup für Probandenversuche einzusetzen.

3.3.4 Darstellung der Blickdaten als HeatMap

Zur Visualisierung von Blickdaten für eine Auswertung von Nutzerstudien existieren unterschiedliche ein- und zweidimensionale Darstellungsweisen. Dabei stellen eindimensionale Graphen als einfachste Form der Visualisierung lediglich die X- und Y-Koordinaten der Blickpunkte als zeitlichen Verlauf in Form eines Plots dar. Beispiele für zweidimensionale Darstellungsweisen sind Scanpaths, FixationMaps oder HeatMaps, die die aufgezeichneten Daten dem ursprünglichen Stimulus überlagert visualisieren.

Scanpaths stellen dabei lediglich die Augenbewegungen als Rohdaten in einem X/Y Diagramm mit dem Testbild als Hintergrund dar. FixationMaps sind im Wesentlichen eine zweidimensionale Aufzeichnung der Positionen aller Fixationen, die analysiert werden sollen. Dabei ist das Testbild zu Beginn einer Auswertung nicht sichtbar, d.h. die Fixation

Map ist dunkel. Nach und nach werden die Stellen, die am intensivsten betrachtet wurden, transparenter und das Testbild sichtbar. Somit sind, sobald die FixationMap vollständig aufgebaut ist, lediglich die Stellen des Stimulus für eine weitere Auswertung sichtbar, die der Proband auch wirklich betrachtet hat.

Eine weitere zweidimensionale Visualisierungsmöglichkeit sind HeatMaps [42, 56]. Diese leiten sich von FixationMaps ab. Die Intensität bzw. Dauer einer Fixation stellen HeatMaps mittels Farbwerten dar. Die verwendeten Farben sind meist schwarz/grau, blau/grün, gelb und rot, wobei kältere Farben kurze und wärmere Farben längere Blickdauern darstellen. Die Darstellung der Farbverläufe erfolgt dabei mit entsprechender Transparenz, um das Testbild im Hintergrund noch genau analysieren zu können, wie die Aufmerksamkeitsverteilung mit einzelnen Elementen des Stimulus korreliert. Ein Vorteil ist, dass die Farbskala größer ist, als die durch Transparenzwerte erzeugte Skala.

Das Erstellen der Heatmaps erfolgt in zwei Schritten: Im ersten Schritt erstellt man eine Intensitätsmaske, die den aufgezeichneten Fixationsdauern in den Blickdaten entspricht. Im zweiten Schritt erfolgt die Abbildung der Intensitätsstufen auf Farbwerte. Dabei werden Pixel mit hoher Intensität, also mit einer langen Fixationsdauer, auf wärmere Farben und dementsprechend Pixel mit niedriger Intensität auf kältere Farben abgebildet. Der Versuchsleiter hat bei der Auswertung zudem die Möglichkeit einzelne Parameter der Darstellung anzupassen: Abbildung der Intensitäten auf den Farbverlauf, bzw. die Fixationszeit, die zu einer entsprechenden Intensität gehört; das „Einzugsgebiet“ einer Fixation bzw. den Radius, den das Auge um ein gewisses Blickziel kreisen darf; etc.



Abbildung 3.35: Unterschiedliche Blickverläufe von zwei Probanden beim Betrachten des Bildschirminhaltes.

Zu Beginn eines Versuchsdurchlaufes startet die Aufzeichnung der Blickdaten. Dazu speichert die Anwendung jeden Blickpunkt in einer Struktur und erstellt bei Änderung des Bildschirminhaltes einen neuen Screenshot mit entsprechendem Timestamp. Im Anschluss an den Versuch wird daraus die Heatmap erstellt. Hierzu ruft eine Schleife jedes Bild auf, durchläuft die Liste der Blickpunkte (in Pixelkoordinaten abgespeichert) und zeichnet an den korrespondierenden Stellen ein Punkt mit Farbverlauf (abhängig

von der einstellbaren Verweildauer auf einem Blickpunkt) ein. Die einzelnen Fixationen sind zusätzlich durch eine aus den ScanPaths bekannte Verbindungslinie untereinander verbunden. Ist die Liste der Blickpunkte zu einem Screenshot durchlaufen, ist das Erstellen der Heatmap abgeschlossen und der zugehörige Screenshot wird mit der Heatmap überlagert. Die einzelnen überlagerten Bilder können zudem als Videodatei exportiert werden. Dazu kann man die Framerate des Videos einstellen, die letztlich die Granularität des HeatMap-Videos bestimmt. Abhängig davon müssen die einzelnen Frames des Videos schrittweise berechnet und mit den zugehörigen Blickdaten überlagert werden. Dadurch kann man im Anschluss an den Versuch den Blickverlauf des Probanden anhand einer zeitlichen Sequenz untersuchen und den Versuchsverlauf im Anschluss offline mit integrierten Funktionen zur Zeitmessung auswerten.

3.3.5 Systemevaluierung

Das folgende Kapitel präsentiert Ergebnisse der Systemevaluierung hinsichtlich des Konvergenzverhaltens der Rekonstruktionalgorithmen, der Laufzeit der Bildverarbeitungsalgorithmen sowie der Genauigkeit der Blickrichtungsrekonstruktion.

3.3.5.1 Konvergenzverhalten

Vor der eigentlichen Umsetzung des Gesamtsystems wurde das Konvergenzverhalten der in Kapitel 3.3.1 vorgestellten Algorithmen zur Rekonstruktion der Blickgeraden auf synthetischen Daten untersucht. Hiefür wurden die einzelnen Verarbeitungsschritte in Maple simuliert. Der von Maple exportierte Code dient gleichzeitig als Grundlage für die spätere Implementierung der Algorithmen in C++. Zur Evaluierung wurden händisch annotierte Daten (Lage von Glints, Pipillenrand und -mittelpunkt im Videobild) aus einem aufgezeichneten Videostream verwendet.

Von besonderem Interesse war dabei die Entscheidung für ein möglichst performantes Gradientenverfahren zur Optimierung des Kostenfunktional aus Gleichung 3.22. Als Abbruchbedingung der numerischen Lösungsverfahren dient der Unterschied ϵ zwischen zwei aufeinanderfolgenden Iterationsschritten. Tabelle 3.2 zeigt stellvertretend die benötigten Iterationsschritte sowie die zeitliche Dauer der Berechnung des Alpha- und des konjugierten Newton-Verfahrens. Hier ist eine klare Dominanz des konjugierten Newton-Verfahrens bezüglich einer möglichst geringen Anzahl an Iterationen bzw. einer möglichst kleinen Berechnungsdauer zu erkennen. Weiterhin zeigt sich eine beschränkte Genauigkeit des Alpha-Verfahrens. Bei Werten für $\epsilon < 1 \cdot 10^{-5}$ ist sehr deutlich zu sehen, dass die obere Schranke für die Anzahl der Iterationen (100000) stets ausgereizt wird. Dies entspricht einem Nichterreichen der vorgegebenen Genauigkeit für ϵ .

Der qualitative Vergleich der Iterationsgeschwindigkeiten der Gradientenverfahren (vgl. Abbildung 3.36) zeigt ebenfalls, dass das Newton-Verfahren die deutlich effektivere Methode darstellt. Die Abbildungen zeigen die Iterationsschritte durch Linien verbunden, wobei die Koordinatenachsen die Werte der Variablen widerspiegeln. Beispielfhaft sind der erste Iterationsschritt (rot), der letzte Iterationsschritt (grün) und der Iterationsschritt, der nach einem Zehntel der insgesamt ausgeführten Iterationen berechnet wurde (orange) eingezeichnet. Das konjugierte Newton-Verfahren ist wiederum das schnellere Verfah-

$\epsilon <$	Alpha-Verfahren		Newton-Verfahren	
	Iterationen	Berechnungsdauer	Iterationen	Berechnungsdauer
$1 \cdot 10^{-1}$	1365	1,539	470	0,716
$1 \cdot 10^{-2}$	3568	4,182	517	0,789
$1 \cdot 10^{-3}$	4648	4,912	562	0,899
$1 \cdot 10^{-4}$	7063	9,257	607	0,921
$1 \cdot 10^{-5}$	100000	110,776	651	0,988
$1 \cdot 10^{-6}$	100000	104,675	696	1,051
$1 \cdot 10^{-7}$	100000	108,232	741	1,129
$1 \cdot 10^{-8}$	100000	104,819	786	1,194

Tabelle 3.2: Vergleich der benötigten Anzahl an Iterationen sowie der benötigten zeitlichen Dauer (in ms) für Alpha- und konjugiertes Newton-Verfahren.

ren: Das Alpha-Verfahren zeigt eine deutliche „Zick-Zack“-Bewegung, wohingegen das konjugierte Newton-Verfahren zielgerichtet und ohne Umwege in den Zielpunkt führt.

3.3.5.2 Bildverarbeitung

Die folgenden Berechnungszeiten wurden auf einem aktuellen Standard Laptop PC¹² bestimmt. Zur Zeitmessung der einzelnen Bildverarbeitungs-komponenten wurde dieselbe Testsequenz (unter Laborbedingungen aufgezeichnete Livebilder) verwendet.

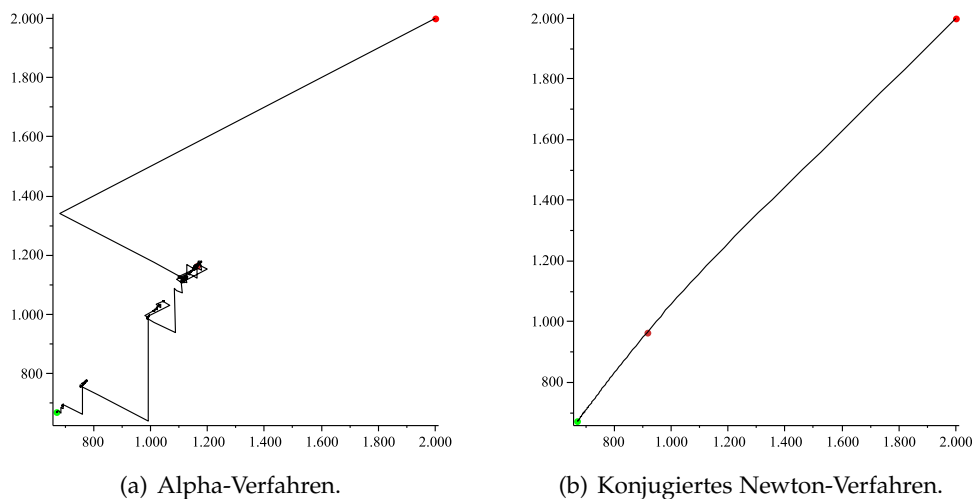


Abbildung 3.36: Vergleich der Iterationsgeschwindigkeiten von Alpha- und konjugiertem Newton-Verfahren.

¹²Prozessortyp: Intel Core 2 Duo T7700; Maximale Taktfrequenz: 2,4 GHz; Arbeitsspeicher: 3,00 GB; Grafikkarte: ATI Mobility Radeon; IPP Optimierung: Intel Core 2 Duo

Die benötigte Zeit der Bildvorverarbeitung (Berechnung von Gauß-, Differenz- und Binärbild) hängt stark von der Güte des Pupillentrackings und damit der Größe der betrachteten Region of Interest (ROI) ab (vgl. Kapitel 3.3.2.4 und 3.3.2.7), weshalb die Werte für die Berechnung auf dem Vollbild (752x240 Pixel) sowie einer ROI Größe von 160x80 Pixel (pro Auge) angegeben werden (siehe Tabelle 3.3).

Größe der ROI	752x240	2x160x80
Gaußbilder	547	121
Differenzbild	45	16
Binärbild	157	48

Tabelle 3.3: Berechnungsdauer in μs der einzelnen Bildvorverarbeitungsschritte.

Die vollständige Dauer der Bildvorverarbeitung setzt sich aus je einmal Gauß-Filter für das helle und das dunkle Kamerabild sowie je einmal die Dauer zur Berechnung von Differenz- und Binärbild zusammen. Somit ergibt sich für das Vollbild eine Dauer von 1,3 ms bzw. eine typische Dauer von 0,3ms (ROI: 2x160x80 Pixel) für die Bildvorverarbeitung.

Die Blob Detektion (vgl. Kapitel 3.3.2.2) benötigte im „durchschnittlichen Fall“ eine Berechnungsdauer von 32 μs . Für alle gefundenen Pupillenkandidaten, deren Größe einer Pupille entspricht, muss für das Tracking der Pupille eine Klassifikation erfolgen. Der Kontrast-basierte Klassifikator benötigte eine durchschnittliche Berechnungszeit von 8 μs pro Pupillenkandidat. Als Vergleich erfolgte ein Test des in [127] vorgestellten SVM basierten Klassifikators auf denselben Bilddaten. Für diesen ergab sich eine durchschnittliche Berechnungszeit von 1279 μs pro Pupillenkandidat. Die Ellipsenerkennung dient schließlich dazu, den Pupillenrand zu ermitteln (vgl. Kapitel 3.3.2.5). Die durchschnittliche Berechnungszeit für das Ellipsefitting auf dem Teststream beträgt 1521 μs pro Pupille.

	Berechnungsdauer
Blob Detektion	32
Klassifikation	8/Pupillenkandidat
Ellipsefitting	1521/Pupille

Tabelle 3.4: Berechnungsdauer in μs der einzelnen Bildverarbeitungsschritte.

3.3.5.3 Genauigkeit

Zur Überprüfung der Genauigkeit der Blickrichtungsrekonstruktion erfolgte eine Probandenstudie unter Laborbedingungen [190]. Zur Aufzeichnung der Bilddaten wurde die Kamera Point Grey FireFly MV (vgl. Kapitel 3.3.3.2) in Kombination mit einem 35 mm Objektiv sowie der Autofokuseinheit (vgl. Kapitel 3.3.2.8 und 3.3.3.4) verwendet. Die Probanden mussten – unter Beibehaltung einer möglichst fixen Kopfposition – aus unterschiedlichen Distanzen neun Referenzpunkte (Abstand horizontal: 10cm, Abstand vertikal: 8cm) auf einem Computermonitor fixieren. Zu jedem Referenzpunkt und jeder Entfernung (1.30 m, 1.00 m, 0.80 m und 0.60 m) wurden die rekonstruierten Blickpunkte (50 Messwerte pro Referenzpunkt und Entfernung) aufgezeichnet und analysiert. Als Maß

für die Rekonstruktionsgenauigkeit dient der Median der Distanzen zwischen Referenzpunkt und den berechneten bzw. geschätzten Blickpunkten auf dem Computermonitor. Für die Genauigkeit ergaben sich Werte von 0.40° , 0.38° , 0.47° und 0.53° (im Mittel: 0.45°) in Distanzen von 0.60 m, 0.80 m, 1.00 m und 1.30 m (vgl. Abbildungen 3.37(a) bis 3.37(d)).

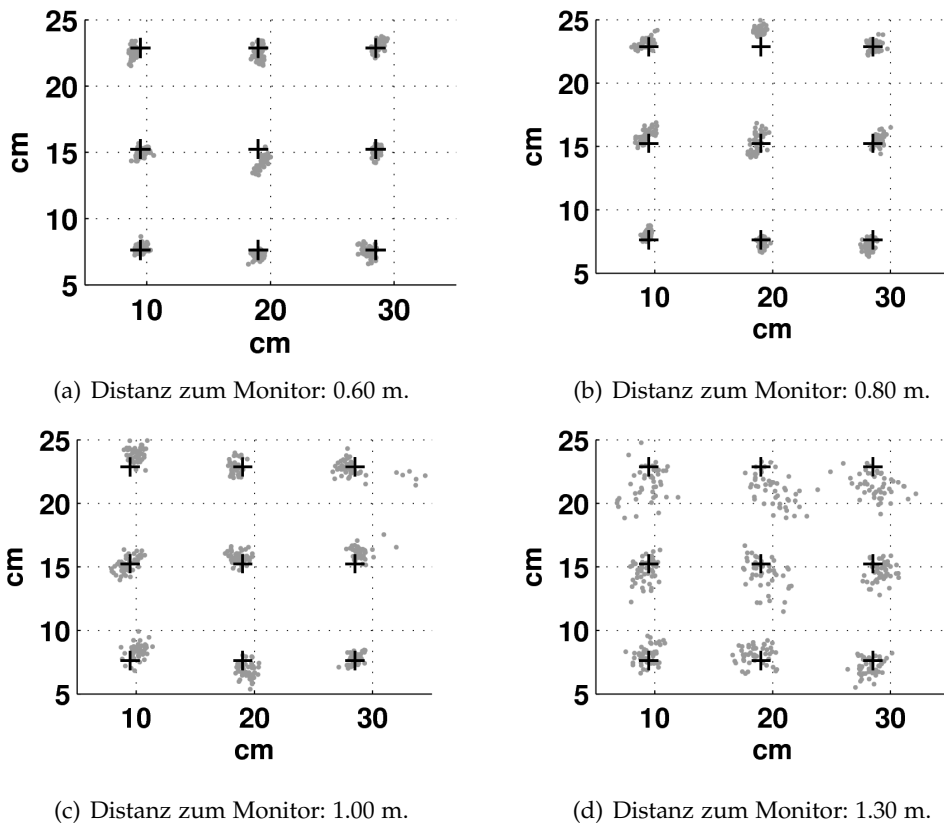


Abbildung 3.37: Aufgezeichnete Rohdaten (berechnete Blickpunkte) des Gaze Trackers in den unterschiedlichen Distanzen zum Monitor. Die schwarzen Kreuzen zeigen die Positionen der Referenzpunkte. Die grauen Punktwolken entsprechen den aufgezeichneten bzw. berechneten Blickpunkten auf dem Monitor.

Die schlechtere Auflösung bzw. das größere Rauschen (Standardabweichung der einzelnen berechneten Blickpunkte der einzelnen Fixationen) sowie die geringere Rekonstruktionsgenauigkeit in größeren Distanzen werden durch unterschiedliche Faktoren hervorgerufen: Aus der größeren Entfernung zur Kamera und den Beleuchtungseinheiten rühren eine geringere Größe und ein geringerer Abstand der Reflexionen (Glints) im Videobild, weshalb sich deren exakte Position schwerer bestimmen lässt. Weiterhin können bei zu großer Distanz zum Tracker die beiden Glints im Videobild mehr und mehr zu einem Blob „verschmelzen“. Diese Faktoren führen zu einem schlechteren Konvergenzverhalten der Rekonstruktionsalgorithmen. Zusätzlich wirkt sich die wachsende Distanz drastisch auf die Bildqualität aus. Als letzter Punkt sei die geringere örtliche Auflösung des Auges im Videobild angeführt, welche sich in größeren Distanzen negativ auf die Ergebnisse der Parameterextraktion auswirkt. Diese Probleme lassen sich jedoch

3. Kamera-basierte Fahrerüberwachung

mit einem auf die Arbeitsdistanz bzw. Gegenstandsweite angepassten Versuchsaufbau minimieren. Dazu sind der Abstand der Beleuchtungseinheiten, die Leuchtstärke der LEDs und die Brennweite des Objektivs (Abhilfe kann hier auch ein Zoom-Objektiv schaffen) entsprechend zu variieren.

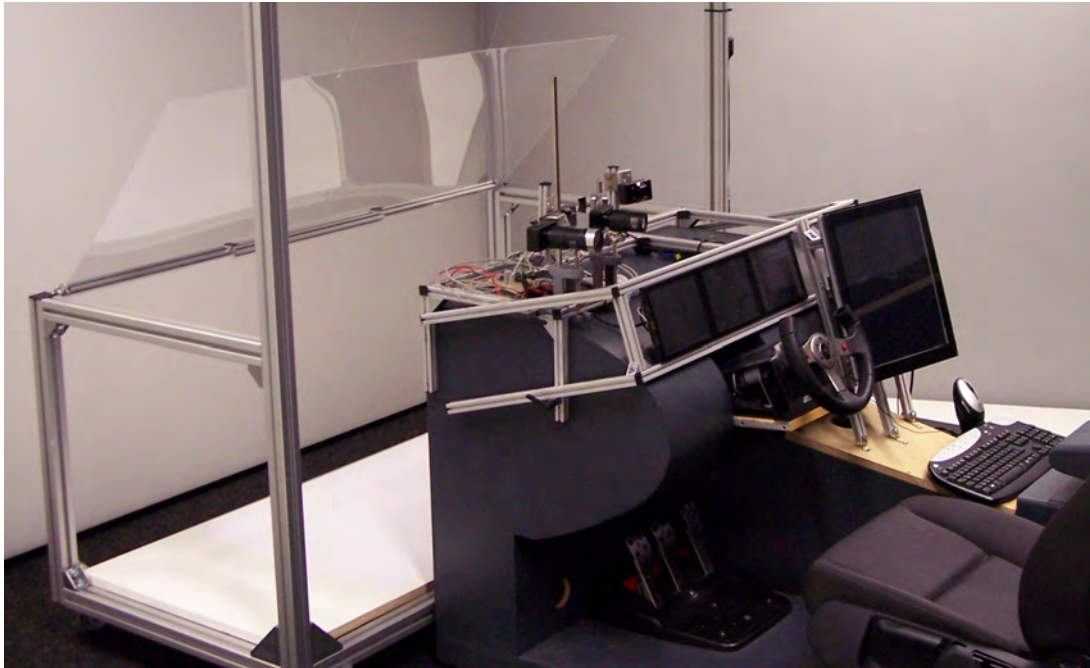


Abbildung 3.38: Versuchsaufbau auf der Sitzkiste. Das Bild zeigt die Sitzkiste mit Kameras zur Blickrichtungserfassung sowie den Displays zur Visualisierung von Anzeige Konzepten.

4

Sensor-basierte Fahrerüberwachung

Zur Bestimmung von Fahrerzustandsparametern wurden in Kapitel 3 kamerabasierte Ansätze vorgestellt, die physiologische Messgrößen bestimmen, um Rückschlüsse auf den Fahrerzustand zu ermöglichen. Kapitel 4 diskutiert Möglichkeiten zur Schätzung des Fahrerzustands auf Basis von Fahrzeugsensoren und stellt Methoden zur praktischen Anwendung vor. Dabei geht es im Wesentlichen darum, durch die Feststellung von Abnormalitäten im Regelverhalten des Fahrers auf Unaufmerksamkeit zu schließen.

Fahrfehler werden hauptsächlich durch Aufmerksamkeitsverluste hervorgerufen (vgl. Kapitel 2), welche sich in einem abnormalen Regelverhalten manifestieren. Diese Abnormalitäten können durch die Auswertung von Fahrzeugdaten identifiziert werden. Das folgende Kapitel stellt Ansätze vor, die es ermöglichen aus Fahrzeugdaten – z.B. Verlauf von Pedalstellung oder Lenkwinkel, Pendelbewegungen in der Spur (lateral oder longitudinal), Beschleunigungen, etc. – Indikatoren des momentanen Fahrerzustandes abzuleiten. Fahrzeugdaten sind jederzeit über den Fahrzeug CAN Bus¹ verfügbar. Daher stellt diese Art der indirekten Beobachtung einen vielversprechenden und robusten Ansatz zur Bewertung des Fahrers bzw. seines Fahrverhaltens dar.

Das folgende Kapitel ist in vier Teile gegliedert: Kapitel 4.1 stellt Parameter vor, die Rückschlüsse auf die Aufmerksamkeitsverteilung und den aktuellen Belastungszustand des Fahrers erlauben. Kapitel 4.2 führt mit der Fuzzy Set Theorie ein Modell ein, das es ermöglicht, unpräzise Informationen (aus den zuvor erörterten Parametern) der Umwelt in Rahmen rechnergestützter Systeme effizient zu verarbeiten. Darauf aufbauend stellt Kapitel 4.3 das erarbeitete Konzept AWARE vor [191], welches auf der Fuzzy Logik basiert, und anhand dessen konkret die Nutzung von Fahrzeugdaten zur Modellierung und Analyse des Zusammenspiels von Umwelt und Fahrer erörtert wird. Anschließend findet in Kapitel 4.4 eine Bewertung des zuvor vorgestellten Fahrerassistenzsystems anhand der Kriterien Abnormalitätsbewertung und Prädikation von Fahrmanövern statt.

4.1 Grundlagen

Um während der Fahrt auf den aktuellen Belastungszustand oder auf die Unaufmerksamkeit zurück zu schließen, können neben den bereits in Kapitel 3 vorgestellten physio-

¹Controller Area Network: asynchrones, serielles Bussystem (Feldbus) in heutigen Kraftfahrzeugen.

logischen Daten sog. Leistungsdaten (engl.: performance metrics) Verwendung finden. Diese beziehen sich direkt auf Daten zur Güte der Erfüllung einzelner Komponenten der Fahraufgabe (z.B. Güte der Spur- bzw. Geschwindigkeitshaltung, Abstandsregelung zum Vorderfahrzeug, etc). Dazu wird die Auswirkung unterschiedlicher Fahrerzustände (fit, müde, abgelenkt, etc.) auf die Fahrleistung untersucht, indem die Variation unterschiedlicher fahrdynamischer Parameter wie Geschwindigkeit, Beschleunigung, laterale Position (Position in der Fahrspur) oder Lenkwinkel protokolliert und analysiert wird. Experimentelle Studien untersuchen dazu den Leistungsunterschied zwischen Fahrten mit/ohne Stress- oder Ablenkungsfaktoren, um die Beeinträchtigung der Fahrleistung durch diese zu beurteilen. Zusätzlich zu fahrdynamischen Parametern erfolgt die Analyse von Antwortzeiten² oder die Anzahl der Fahrfehler im jeweiligen Zeitfenster. Zur Entwicklung von Assistenzsystemen werden diese Metriken hauptsächlich verwendet, um zusätzliche Ablenkung, Unaufmerksamkeit, Zerstreuung und allgemeine Belastung des Fahrers zu untersuchen. Hierbei erfolgt hauptsächlich eine Unterscheidung zwischen longitudinalen und lateralen Führungsparametern.

4.1.1 Longitudinale Führungsparameter

Longitudinale Führungsparameter erfassen alle Variablen bezüglich der Längsführung des Fahrzeugs. Hierzu gehören z.B. Geschwindigkeit, Beschleunigung, Abstand zum Vorderfahrzeug oder Gaspedal- bzw. Bremspedalstellung. Die Relevanz dieser Parameter hinsichtlich der Diagnostizierbarkeit des Fahrerzustandes wurde bereits in zahlreichen Studien – meist in Verbindung mit Ablenkung bzw. Unaufmerksamkeit – untersucht.

Eine häufige Annahme ist, dass der Fahrer versucht, die Belastung durch die Hauptaufgabe³ zu reduzieren, während er mit anderen Aufgaben beschäftigt ist. Dies schlägt sich vor allem in der Geschwindigkeitsregulierung nieder. *Zylstra* et al. stellten in ihren Experimenten fest, dass Testpersonen im Allgemeinen die Geschwindigkeit während Nebenaufgaben – insbesondere während der Eingabe von Navigationszielen – signifikant herabsetzten [260]. Weiterhin zeigten etwa 80 % der Probanden während Nebentätigkeiten Schwächen im Bereich der feinen Geschwindigkeitskorrekturen. Waren mehrere Aufgaben zu bearbeiten zeigte sich, dass die Probanden nicht alle Aufgaben (z.B. Nebenaufgabe, Lenken und Geschwindigkeitsregelung) kontinuierlich bearbeiten konnten, was dazu führte, dass sie zeitweise die Längsregelung (Geschwindigkeitssanpassung) vernachlässigten. Darüber hinaus ziehen die Autoren den Schluss, dass Parameter der Längsführung (vor allem bei geraden Straßen mit leichtem bis mäßigem Verkehr) geeignetere Indikatoren zur Bestimmung von Unaufmerksamkeit sind als Querführungsparameter.

Auch *Jamson & Merat* konnten dieses Phänomen bestätigen. Beim Experiment aus [119] war es die Aufgabe der Teilnehmer zufällig auftretende Klänge in einer Audiosequenz zu zählen sowie zu beurteilen ob ein gesuchter Pfeil in einer Gruppe von Pfeilen (dargestellt auf einem LCD Display) vorhanden war. Beide auditiven und visuellen Aufgaben wurden auf verschiedenen Schwierigkeitsstufen durchgeführt. Diese Nebenaufgaben verursachten nicht nur eine Reduktion der gefahrenen Geschwindigkeit, sondern ebenfalls der TTC

²Z.B. in Laborumgebungen Reaktion auf spezielle Stimuli, vgl. Kapitel 2.4.1.2.

³Primäre Fahraufgabe, vgl. Kapitel 2.3.1.1.

(Time to Collision⁴), wenn das vorausfahrende Fahrzeug unerwartet abbremst. Diese Charakteristik zeigte sich vor allem während visuellen Nebenaufgaben (Identifikation des gesuchten Pfeils) und stärker mit zunehmender Komplexität der Aufgabe. Die Tests erfolgten sowohl auf geraden als auch auf kurvigen Strecken. [16, 49] beobachteten ähnliche Ergebnisse in Bezug auf die Ressourcen Allokation von Autofahrern.

Zusätzlich zeigte sich, dass eine gesteigerte Variation von Geschwindigkeit und Distanz zum Vorderfahrzeug bereits vorhanden ist, bevor kompensatorisches Verhalten auftrat [68]. Diese Variationen lassen frühzeitig auf ein höheres Risiko für Unaufmerksamkeit bzw. höhere mentale Last des Fahrers schließen. Die Nutzung von Mobiltelefonen während kognitiver Aufgaben beeinflusst das Stopverhalten dahingehend signifikant, dass die Fahrer vermehrt stärker bremsen müssen bzw. die Distanz zu Haltelinien und Kreuzungspunkten zu Beginn der Bremsung deutlich kürzer ausfällt [84]. Die Ergebnisse der unterschiedlichen Studien in Bezug auf die Korrelation zwischen Längsführungsverhalten und (mentaler) Belastung lassen sich wie folgt zusammenfassen:

- Zusätzliche mentale Belastung reduziert vor allem die Fähigkeiten des Fahrers zur Interaktion mit der Umgebung, Verkehr, etc.
- Zunehmende Komplexität der Aufgaben bzw. steigende kognitive Anforderungen bedingen eine reduzierte Geschwindigkeitskontrolle.
- Typische Indikatoren sind erhöhte Geschwindigkeitsschwankungen, erhöhte Variation des Abstands zum Vorderfahrzeug und stärkere Bremsmanöver.
- Diese Merkmale sind besonders ausgeprägt während komplexer (Neben-)Aufgaben mit hohem Informationsgehalt und langer Dauer (z.B. Zieleingabe).
- Besonders Geschwindigkeitsverringern dient als kompensatorisches Verhalten.

4.1.2 Laterale Führungsparameter

Laterale Führungsparameter fassen alle Variablen zusammen, welche die Querführung des Fahrzeugs beschreiben. Hierzu gehören beispielsweise der Lenkwinkel, die Lenkfrequenz, die Querbeschleunigung und die laterale Position in der Spur bzw. die Abweichung von der Mitte des Fahrschlauches. [244] zeigt beispielsweise, dass zusätzliche Belastung durch (visuelle) Nebenaufgaben eine Erhöhung der Lenkfrequenz verursacht, während die Geschwindigkeit keine Beeinflussung erfuhr. Der entscheidende Punkt ist die Veränderung der Verteilung der Lenkbewegungen im Frequenzspektrum. Die Zahl der Lenkbewegungen in höheren Frequenzbereichen stieg im Rahmen der Versuche deutlich an. Somit ist der Anteil hochfrequenter Lenkbewegungen als objektive Metrik für den Grad der (mentalen) Belastung interpretierbar [145]. Je größer der Prozentsatz hochfrequenter Lenkradbewegungen ausfällt, desto größer ist die Belastung des Fahrers. Darauf basierend führt [168] die sog. Lenkentropie zur Quantifizierung der Bemühungen des Fahrers ein, einen „seitlichen Sicherheitsabstand“ bzw. das Fahrzeug in der Spur zu

⁴Engl. Time To Collision: Zeit bis zum Aufprall; typischerweise berechnet aus dem Abstand und der relativen Geschwindigkeit zu dem voraus fahrenden Fahrzeug.

halten. [40] verbessert die Empfindlichkeit bezüglich der Beurteilung von Veränderungen im Lenkverhalten, welche mit Änderungen der Arbeitsbelastung korrelieren.

[96] präsentiert Versuchsergebnisse zum Einfluss des Telefonierens auf die Leistungsfähigkeit des Fahrers. Abhängig von der Nutzung einer Freisprechanlage ergaben sich Veränderungen in der Häufigkeit des Spurverlassens, wobei sich ohne Nutzung einer Freisprechanlage eine deutlich höhere Zahl solcher Events zeigte. [233] belegt ebenfalls eine signifikante Änderung der Querführungsgüte bei Nutzung eines Telefons. Die Autoren registrierten unabhängig von der Nutzung einer Freisprechanlage eine erhöhte Standardabweichung der lateralen Spurablage. Basierend auf der statistischen Analyse in [205], stellen die Standardabweichungen der lateralen Spurablage und des Lenkwinkels die zwei wichtigsten Indikatoren für die Erkennung von Fahrerablenkung dar. Wesentliche Erkenntnisse bezüglich der Korrelation zwischen Querführungsgüte und zusätzlicher Belastung durch Nebenaufgaben lassen sich folgendermaßen zusammenfassen:

- Zusätzliche Belastungen während der Fahrt (Nebenaufgaben) können Schwankungen im Lenkverhalten hervorrufen.
- Die Lenkfrequenz stellt eine geeignete Metrik zur Erkennung von Unaufmerksamkeit bzw. Stress dar.
- Die Standardabweichung der lateralen Spurablage ist eine weitere geeignete Metrik zur Erkennung von Unaufmerksamkeit bzw. Stress.

4.2 Fuzzy Logik

Im täglichen Leben bewerten wir Eigenschaften oder äußere Umstände typischerweise mit Ausdrücken wie *ungefähr*, *viel*, *ein wenig* oder *fast*. Z.B. beschreibt man ein Fahrzeug, das mit einer Geschwindigkeit von 140 km/h auf der Autobahn fährt, als *schnell*. Nach der klassischen Mengenlehre ist ein Fahrzeug mit einer Geschwindigkeit von 135 km/h als langsam zu bezeichnen, wohingegen „menschliche“ Logik das Fahrzeug immer noch als schnell bezeichnet. Zur effizienten Verarbeitung von unscharfen, vagen oder unpräzisen Informationen in rechnergestützten Systemen wurde – angelehnt an das menschliche Naturell – die Theorie der *Fuzzy Logik* entwickelt. Sie verknüpft menschliches Wissen mit unscharfen oder verschwommenen Eigenschaften. Während die klassische Mengenlehre üblicherweise zwischen zwei Werten scharf unterscheidet (entweder ist ein Element Teil einer Menge oder nicht), befasst sich die Fuzzy Set Theorie mit der unscharfen Mengenlehre und erweitert die Boolesche Logik um Werte zwischen *wahr* und *falsch*.

4.2.1 Fuzzy Set Theorie

Die heute bekannte Fuzzy Theorie wurde 1965 von *Zadeh* vorgestellt [256]. In dem Artikel „Fuzzy Sets“ beschreibt er die mathematischen Grundlagen der Fuzzy Set Theorie zur Verarbeitung von unscharfen Informationen. Die folgenden Abschnitte stellen zum besseren Verständnis kurz die Grundzüge der Theorie vor. Für weitere Einzelheiten sei auf die ausführlichen Beschreibungen der Grundlagen in [172, 256] verwiesen.

In der traditionellen „zweiwertigen“ Logik ist ein Element entweder in einer Menge enthalten oder nicht. Wenn ein Mitglied einer „gewöhnlichen“ Menge (crispe Menge;

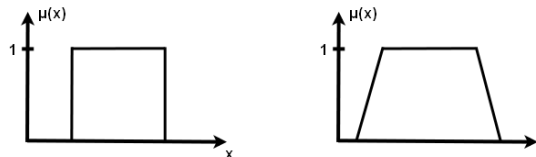


Abbildung 4.1: Scharfe Menge (links) und Fuzzymenge (rechts).

engl.: crisp set) immer auf die Werte *wahr* oder *falsch* ($\{0, 1\}$) abgebildet wird, erlauben Fuzzymengen (engl.: fuzzy sets) eine „anteilige“ Mitgliedschaft der einzelnen Elemente $[0, 1]$. Eine Fuzzymenge \tilde{A} ist mathematisch folgendermaßen beschrieben:

$$\tilde{A} := \{(x, \mu_{\tilde{A}}(x)) | x \in X\} \quad (4.1)$$

wobei x ein Element aus der Fuzzymenge X ist und $\mu_{\tilde{A}}(x)$ die Zugehörigkeitsfunktion (engl.: membership function) darstellt. Die Zugehörigkeitsfunktion $\mu_{\tilde{A}}(x)$ bildet also jedes Element x der Fuzzymenge X des Universums U auf das Einheitsintervall ab:

$$\mu_{\tilde{A}} : U \rightarrow [0, 1] \quad (4.2)$$

Der Unterschied zu gewöhnlichen Mengen ist demnach, dass nicht nur festgelegt wird, ob x in der Menge enthalten ist, sondern vielmehr zu welchem Grad es in der Menge liegt (*graduelle* Mitgliedschaft). Verfügt die membership function dabei lediglich über zwei diskrete Ausgangswerte, handelt es sich bei \tilde{A} um ein crisp set. Eine scharfe Menge kann als Spezialfall der Fuzzy Theorie betrachtet werden:

Crispe Menge: $\mu_A : X \rightarrow \{0, 1\}$

Fuzzymenge: $\mu_{\tilde{A}} : X \rightarrow [0, 1]$

Zur Bestimmung der Zugehörigkeitsfunktion sind Erfahrungen, Beobachtungen oder empirische Studien des Anwendungsszenarios notwendig. Vor allem bei neuen Anwendungsgebieten, zu denen noch kein statistisches Vorwissen existiert, sind Wissen und Gespür von Expertengruppen ausschlaggebend. Heuristisches Expertenwissen lässt sich im Sinne der Zugehörigkeitsfunktionen durch linguistische Terme und Variablen abbilden. Die Steuerung eines Tempomaten könnte z.B. folgendermaßen ausfallen:

- die aktuelle *Ist-Geschwindigkeit* sowie die gesetzte *Soll-Geschwindigkeit* des Egofahrzeuges als Eingang
- die *Steuerung* der Geschwindigkeitsregelanlage als Ausgang

Somit lautet eine mögliche Regel der Fuzzy-Steuerung wie folgt:

WENN <Ist-Geschwindigkeit> KLEINER ALS <Soll-Geschwindigkeit>
DANN Steuerung beschleunigen bis <Soll-Geschwindigkeit> erreicht.

Fuzzy-Steuerungssysteme kombinieren typischerweise mehrere Eingänge, die durch die logischen Operatoren UND und ODER verknüpft sind. UND ist typischerweise mit dem Minimum-Operator $\min\{.,.\}$ und ODER mit dem Maximum-Operator $\max\{.,.\}$ realisiert. Zur einfachen Berechenbarkeit und Anpassbarkeit werden üblicherweise Funktionen mit einer einfachen Form und wenigen, an das Modellierungsproblem angepassten Parametern als Zugehörigkeitsfunktion verwendet.

Darauf aufbauend verallgemeinern weitere Definitionen und Operationen der klassischen Mengenlehre die Fuzzy Theorie:

- *Gleichheit*: zwei Fuzzymengen \tilde{A} und \tilde{B} sind dann und nur dann gleich, wenn $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$:

$$\tilde{A} = \tilde{B} \iff \mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x) \quad (4.3)$$

- *Leere Menge*: eine Fuzzymenge \tilde{A} wird dann und nur dann als leere Menge bezeichnet, wenn $\mu_{\tilde{A}}(x) = 0$:

$$\tilde{A} = \emptyset \iff \mu_{\tilde{A}}(x) = 0 \quad (4.4)$$

- *Teilmenge*: eine Fuzzymenge \tilde{A} ist dann und nur dann eine Untermenge einer Fuzzymenge \tilde{B} , wenn $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$ gilt:

$$\tilde{A} \subseteq \tilde{B} \iff \mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x) \quad (4.5)$$

- *Negierung*: wenn \tilde{A} eine Fuzzymenge aus dem Universum X ist, wird die Negierung $\neg\tilde{A}$ folgendermaßen definiert:

$$\neg\tilde{A} = \{(x, (1 - \mu_{\tilde{A}}(x))) | x \in X\} \quad (4.6)$$

- *Vereinigungsmenge*: die Vereinigung der Fuzzymengen \tilde{A} und \tilde{B} mit der Zugehörigkeitsfunktion $\mu_{\tilde{A} \cup \tilde{B}}$ der Vereinigung $\tilde{A} \cup \tilde{B}$ ist definiert als:

$$\mu_{\tilde{A} \cup \tilde{B}}(x) = \mu_{\tilde{A}}(x) \oplus \mu_{\tilde{B}}(x) \quad (4.7)$$

- *Schnittmenge*: der Schnitt der Fuzzymengen \tilde{A} und \tilde{B} mit der Zugehörigkeitsfunktion $\mu_{\tilde{A} \cap \tilde{B}}$ des Schnittes $\tilde{A} \cap \tilde{B}$ ist definiert als:

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \mu_{\tilde{A}}(x) \otimes \mu_{\tilde{B}}(x) \quad (4.8)$$

4.2.2 Fuzzy-Inferenzsysteme

Fuzzy-Inferenzsysteme kombinieren die Eingänge eines Systems mit den sprachlich formulierten, unscharfen Regeln um eine Folgerung ziehen zu können. Im Bereich der klassischen Regelungstechnik wird die Regelung eines technischen Systems typischerweise in Form von Differentialgleichungen dargestellt. Auch wenn eine Spezifikation in Form von Differentialgleichungen nahe liegt, lässt sich nicht jedes Modellierungsproblem – mit angemessenem Aufwand – auf ein geeignetes mathematisches Modell zurückführen, da die Bestimmung der entsprechenden Differentialgleichungen einen nicht zu

vernachlässigenden Aufwand bedeutet. Zusätzlich erfordert die „klassische Art und Weise“ der Steuerungsauslegung ein hohes Maß an physikalischem und mathematischem Expertenwissen. Zudem ist die Kontrolle/Regelung eines komplexen Gesamtsystems auch ohne vorherige detaillierte Kenntnis der zugrunde liegenden physikalischen und/oder mathematischen Hintergründe möglich. Beispielsweise kann ein Mensch ohne ein Grundlagenwissen zu Differentialgleichungen das Radfahren meistern.

Eine alternative Methode zur Beherrschung dieses Problems stellt z.B. die Modellierung des Verhaltens eines Menschen (Experten) dar, der dieses Gesamtsystem (Fahrrad/Mensch/Straße/Physik) kontrollieren kann. Die Modellierung des Expertenwissens erfolgt dabei in Form von linguistischen Regeln, die aus einer *Prämisse* (Beschreibung einer bestimmten Situation inkl. einer vagen Spezifikation von Messgrößen) und einer entsprechenden Schlussfolgerung (angemessener Stellwert für diese Situation) bestehen. Im nächsten Schritt lassen sich die linguistischen Regeln mittels Fuzzy Sets mathematisch modellieren, um das entsprechende Wissen durch den Computer berechenbar auszuwerten. Dieser Ansatz wird als Fuzzy-Inferenz bezeichnet. Fuzzy-Controller (Regelungen) sind heutzutage in vielen technischen Geräten (z.B. Waschmaschinen [12]) anzutreffen und somit die wichtigste – auch kommerziell erfolgreiche – Anwendung der Fuzzy Theorie. Fuzzy-Inferenzsysteme bestehen im Wesentlichen aus vier Komponenten:

- *Fuzzyfikator*: Der Fuzzyfizierer berechnet aus den reellen Eingangsparametern Fuzzymengen. Weiterhin wird er für zusätzliche einfache Transformationen der Eingangswerte herangezogen (einfache Abbildung auf adäquate Wertebereiche).
- *Regelbasis*: Die Regelbasis besteht aus zwei Teilen. Der erste Teil besteht aus der Gesamtheit aller linguistischen Regeln. Weiterhin enthält die Datenbank alle Informationen über Ein- und Ausgabeparameter sowie den Zusammenhang zwischen den linguistischen Regeln und den Fuzzy Sets.
- *Entscheidungslogik*: Die Entscheidungslogik ist die Verarbeitungseinheit. Auf Grundlage der Regelbasis werden hier aus den gegebenen Eingangsgrößen die entsprechenden Schlussfolgerungen gezogen.
- *Defuzzyfikator*: Der Defuzzyfikator bestimmt aus den Informationen der Entscheidungslogik einen scharfen (crisp) Ausgabewert.

Die wichtigsten Typen von Fuzzy-Inferenzsystemen sind die Systeme nach *Mamdani* und *Takagi-Sugeno*. Die Mamdani Methode zur Auslegung eines Fuzzy-Controllers wurde 1974 beim Versuch, eine Dampfmaschine und einen Boiler mit einem durch erfahrene Bediener aufgestelltem linguistischen Regelsatz zu steuern, entworfen [148].

Die Formulierung des Expertenwissens als linguistische Regeln ist der erste Schritt bei der Umsetzung nach Mamdani. Alle sprachlichen Begriffe, die in den Regeln auftreten können, müssen definiert und eine angemessene sprachliche Bezeichnung für jedes Wertepaar bestimmt werden. Obwohl der gleiche sprachliche Ausdruck für zwei verschiedene Input-Parameter verwendet werden kann, kann die Interpretation anders sein. Zur mathematischen Beschreibung erhält jeder Ausdruck eine geeignete Fuzzy-Menge gemäß dem Expertenwissen. Die sprachlichen Begriffe repräsentieren somit lediglich Namen für die zugehörigen Fuzzy Sets. Auf diese Art und Weise sind alle Mengen des Eingabe-

4. Sensor-basierte Fahrerüberwachung

(X_1, \dots, X_n) und des Ausgabebereiches (Y_1, \dots, Y_m) mittels Fuzzy Sets partitioniert. Die Anzahl der verwendeten Sets p_i, q_j für Ein- und Ausgabebereich kann dabei unterschiedlich sein.

$$\tilde{A}_{i,1}, \dots, \tilde{A}_{i,p_i} \in F(X_i) \quad (i = 1, \dots, n) \quad (4.9)$$

$$\tilde{B}_{j,1}, \dots, \tilde{B}_{j,q_j} \in F(Y_j) \quad (j = 1, \dots, m) \quad (4.10)$$

$F(X_i)$ und $F(Y_j)$ beschreiben die Menge aller Fuzzy Sets über X_i bzw. Y_j . Die Partitionierungen und die linguistischen Regeln formen die Regelbasis. Zur deren Vervollständigung muss weiterhin das Regelwerk – der eigentliche Kern des Systems – definiert werden. Dazu wird jede Regel aus einem vorangehenden Teil (Beschreibung der Situation anhand der Eingabeparameter) und einer entsprechenden Konsequenz (Reaktion) formuliert:

$$\text{WENN } x_1 = \tilde{A}_{1,*} \text{ UND } \dots \text{ UND } x_n = \tilde{A}_{n,*} \text{ DANN } y_1 = \tilde{B}_{1,*} \text{ UND } \dots \text{ UND } y_m = \tilde{B}_{m,*}$$

Unabhängig von der Dimension des Ausgabebereiches m ist es ausreichend, Regeln zu formulieren, deren Ergebnisse auf eine Dimension des Ausgabebereiches verweisen, wohingegen Schlussfolgerungen, die auf mehrere Dimensionen verweisen

$$\text{WENN } x_1 = \tilde{A}_{1,*} \text{ UND } \dots \text{ UND } x_n = \tilde{A}_{n,*} \text{ DANN } y_1 = \tilde{B}_{1,*} \text{ UND } y_2 = \tilde{B}_{2,*}$$

relativ einfach durch mehrere Regeln ausgedrückt werden können,

$$\text{WENN } x_1 = \tilde{A}_{1,*} \text{ UND } \dots \text{ UND } x_n = \tilde{A}_{n,*} \text{ DANN } y_1 = \tilde{B}_{1,*}$$

$$\text{WENN } x_1 = \tilde{A}_{1,*} \text{ UND } \dots \text{ UND } x_n = \tilde{A}_{n,*} \text{ DANN } y_2 = \tilde{B}_{2,*}$$

ohne dass das Ergebnis verfälscht wird [172]. Somit lassen sich die einzelnen Verarbeitungsschritte eines Fuzzy-Inferenzsystems nach Mamdani folgendermaßen beschreiben:

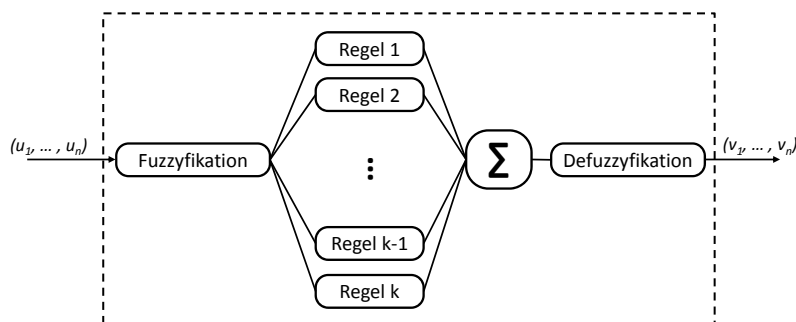


Abbildung 4.2: Struktur eines Fuzzy-Inferenzsystem nach Mamdani [148].

1. *Fuzzyfikation*: Anhand der Zugehörigkeitsfunktionen werden die exakten, scharfen Eingangswerte gewichtet und auf diese Weise der Zugehörigkeitsgrad jeder einzelnen Eingangsvariablen bestimmt.
2. *Aggregation*: Die Aggregation bestimmt den Wahrheitsgehalt einer gesamten Regel. Besteht eine Regel aus mehreren Eingängen, sind diese entsprechend den linguistischen Termen kombiniert und zu einem Wahrheitsgehalt zusammengeführt.

3. *Implikation*: Der in der Aggregation bestimmte Wahrheitsgehalt wird auf die Konsequenz des Regelwerks übertragen.
4. *Akkumulation*: Erfüllen mehrere Regeln die ermittelten Wahrheitswerte, so werden diese in der Akkumulation mit geeigneten Operatoren weiter verarbeitet.
5. *Defuzzifikation*: Die Defuzzifikation führt das Ergebnis der Akkumulation auf einen scharfen Ausgangswert zurück.

Der Fuzzy-Controller nach Takagi-Sugeno [226, 227] ist eine Modifizierung der Methode von Mamdani. Der prinzipielle Unterschied besteht in der Aufstellung der Regeln sowie der Berechnung der entsprechenden Schlussfolgerungen. Im Gegensatz zu den Fuzzy Mengen eines Mamdani Reglers, liegen die Zugehörigkeitsfunktionen des Ausgangs als Linearkombinationen der Eingangsgrößen vor (Mittelwert aus den gewichteten Ergebnissen sowie den Erfüllungsgraden der einzelnen Regeln). Als Ergebnis liegt eine scharfe Stellgröße vor, was eine Defuzzifizierung unnötig macht. Das Regelverhalten eines Takagi-Sugeno Reglers wird dazu mit Regeln der Form

$$R_l : \text{WENN } x_1 = \tilde{A}_{1,*}^l \text{ UND } \dots \text{ UND } x_n = \tilde{A}_{n,*}^l \text{ DANN } y_j = f_l(x_1, \dots, x_n)$$

spezifiziert, wobei f_l ($l = 1, \dots, k$) einer Transformation von $X_1 \times \dots \times X_n$ nach Y , x_i den scharfen Eingangswerten und $\tilde{A}_{n,*}^l$ dem Wert einer linguistischen Variablen entspricht. Aus einem Eingangstupel $(u_1, \dots, u_n) \in X_1 \times \dots \times X_n$ wird mithilfe der Entscheidungslogik der Wert von $f_l(u_1, \dots, u_n)$ sowie der Erfüllungsgrad α_l berechnet. Die finalen Ausgabe- v_1, \dots, v_m ergeben sich folgendermaßen:

$$v_j = \frac{\sum_{R_l \in RL(j)} \alpha_l \cdot f_l(u_1, \dots, u_n)}{\sum_{R_l \in RL(j)} \alpha_l} \quad (4.11)$$

wobei $RL(j)$ die Menge aller Regeln beschreibt, deren Konklusionen der Dimension des Ausgabebereiches Y_j entsprechen. Ein Vorteil der Takagi-Sugeno Methode liegt in der weniger komplexen Bestimmung der scharfen Ausgänge. Die komplexe Berechnung der Ausgabe Fuzzy Mengen sowie die Defuzzifizierung entfallen. Ein Nachteil liegt in der deutlich komplexeren Gestaltung des Regelwerks, da für die Konklusionen reelle Parameter oder lineare Funktionen statt linguistischer Formulierungen Verwendung finden, was die Spezifikation erschwert [229].

4.3 Das Assistenzsystem AWARE

Das System AWARE – Agent-based Warning and Risk Estimation – dient dazu, den Fahrer als eine Art „virtueller“ Beifahrer frühzeitig vor kritischen Situationen zu warnen [191]. Dazu macht es sich die Fähigkeiten eines menschlichen Co-Piloten zu Nutzen, gefährliche Verkehrssituationen frühzeitig zu antizipieren, wenn sie den entsprechenden Fahrer kennen bzw. dessen fahrerische Fähigkeiten einschätzen können. Die Grundidee ist die Nutzung von Fahrzeugdaten, um das Zusammenspiel zwischen Umwelt und Fahrer bestmöglich zu modellieren und zu analysieren. Das System ist mit Hilfe eines dynamischen Fahrermodells in der Lage, Abnormalitäten im Fahrverhalten zu detektieren:

Alle relevanten Informationen zum aktuellen Fahrverhalten (laterales sowie longitudinales Führungsverhalten) und Verkehrsbedingungen werden kontinuierlich protokolliert. Sie bilden die (Verlaufs-) Datenbank, auf deren Basis das System den Fahrer bzw. seine typische Fahrweise „erlernt“. Bei erheblichen Diskrepanzen zwischen dem tatsächlich vorliegenden und dem bisher „gelernten“ (Fahr-)Verhalten kann auf sicherheitskritische Beeinträchtigungen des Fahrers zurück geschlossen werden. Um den Anforderungen an ein solches Assistenzsystem gerecht zu werden, sind verlässliche Datenquellen zur Informationsakquise ebenso elementar, wie die Wahl einer entsprechenden Metrik zur Bestimmung der relevanten Fahrerparameter. Die Schätzung des Fahrerzustandes soll dabei auf Basis von Fahrzeugdaten erfolgen, da diese ohne weitere Beeinträchtigung des Fahrers (keine zusätzlichen Messgeräte) erhoben werden können und gleichzeitig Rückschlüsse auf das Ausmaß einer Aufmerksamkeitsstörung erlauben.

Basis des gesamten Systems ist die Software-Agenten Plattform *JADE* (Java Agent Development Framework) [31]. Die Verarbeitung aller erfassten Informationen erfolgt durch verschiedene Agenten, die autonome, proaktive, reaktive, soziale und kooperative Handlungsfähigkeiten besitzen. *JADE* stellt dazu gewisse Basisstrukturen zur Umsetzung von Agenten sowie Verhaltensweisen (behaviours), Kommunikationsfunktionen, Management-Tools und eine entsprechende Runtime Umgebung zur Verfügung.

Das Gesamtsystem ist in mehrere funktionale Blöcke, repräsentiert durch unterschiedliche (Software-)Agenten, gegliedert [72, 191]. Insgesamt existieren sieben Agenten mit unterschiedlichen Verhaltensweisen und Aufgabengebieten. Fünf dieser Agenten sind für die Verarbeitung und Analyse der Rohdaten hinsichtlich folgender Parameter verantwortlich:

- **Fahrverhalten:** Der Agent sammelt und überwacht kontinuierlich alle relevanten Daten zur Modellierung des aktuellen Fahrverhaltens. Signifikante Unterschiede zur gelernten Fahrweise stellen Indikatoren für Fahrerablenkung dar und werden an den verantwortlichen Agenten weitergeleitet.
- **Fahrerabsicht:** Dieser Agent schätzt bevorstehende Fahrmanöver anhand des Fahrtverlaufes (Datenbank) bei früheren Manövern ab.
- **Fahrertyp:** Hauptaufgabe dieses Agenten ist es, den Fahrertyp bzw. das Fahrverhalten auf Basis des Folge- sowie des Auflaufverhaltens auf vorausfahrende Fahrzeuge in drei Stufen einzuschätzen bzw. zu klassifizieren (vorsichtig, normal, aggressiv).
- **Longitudinal-Strategie:** Auf Basis des Fahrertypus, der aktuellen Fahrdaten und der aktuellen Verkehrsbedingungen schlägt der zuständige Agent geeignete Längsführungsmaßnahmen vor, um die vorliegende Situation zu bewältigen.
- **Lateral-Strategie:** Auf Basis der Querführungsdaten analysiert dieser Agent die für einen Fahrer „typische“ laterale Spurabweichung und kann somit Manöver vorschlagen, die dem Fahrer beim Spurhalten unterstützen.

Zusätzlich steuert und überwacht der Chief Executive Agent (CEA) alle Agenten. Er ist die letzte Instanz des gesamten Systems: Alle von den informationsverarbeitenden Agenten bereitgestellten Informationen werden an den CEA übermittelt. Er entscheidet

daraufhin, ob dem Fahrer ein Hinweis gegeben wird oder nicht, indem er die vorgeschlagenen Strategien mit dem aktuellen Fahrverhalten abgleicht. Abhängig von der CEA-Bewertung wird die empfohlene Strategie direkt angezeigt, entsprechend angepasst oder verworfen. Weiterhin erfüllt der CEA zusätzlich die Aufgaben der Fehlerbehandlung für jeden einzelnen Agenten sowie für das Gesamtsystem. Der GUI Agent kontrolliert und stellt die Benutzeroberfläche dar und fungiert als Management und Konfigurations-Tool für das Gesamtsystem. Die GUI stellt alle relevanten Informationen über den Fahrer sowie die einzelnen Agenten dar. Der Agent ermöglicht es darüber hinaus, spezifische Systemanpassungen zur Laufzeit vorzunehmen. Neben den Funktionalitäten, die von den Software-Agenten zur Verfügung gestellt werden, ist die Datenbank eine weitere wesentliche Komponente des Systems, in der alle Rohdaten sowie die Berechnungen der Agenten gespeichert sind. Alle Daten dienen als Grundlage für die automatische Parametrisierung und Anpassung der verwendeten Algorithmen und Modelle.

4.3.1 Modellierung des Fahrverhaltens

Der Zusammenhang zwischen einer bestimmten Verkehrssituation und einer entsprechenden Handlung seitens des Fahrers ist ein essentieller Anhaltspunkt, um einschätzen zu können, ob der Fahrer aktuell von seiner typischen Fahrweise abweicht. Der Fahrer leitet Aktionen oder Fahrmanöver ein, um auf Basis seiner Wahrnehmung der aktuellen Situation eine gewisse Absicht umzusetzen. Grundsätzlich wird die aktuelle Situation in Form von vagen Aussagen wie beispielsweise „der Abstand zum Vorderfahrzeug ist zu gering“ oder „meine Geschwindigkeit ist zu hoch“ klassifiziert. Auf Basis von Erfahrungen bzw. erlernten Kenntnissen werden nötige Aktionen festgelegt, z.B. „kommt das Vorderfahrzeug zu schnell näher, sollte ich abbremesen“. Die daraus resultierende Reaktion „Vollbremsung“ erfolgt ebenfalls in einer vagen Art und Weise. Die Fuzzy-Inferenzsysteme der Agenten repräsentieren diese unterbewussten Prozesse des Fahrers.

Die Agenten werten vor allem Längsführungsparameter während Folgefahrten und Spurhaltephasen aus, um Fahrerablenkung detektieren zu können. Diesem Vorgehen liegt – wie ausführlich in [123] untersucht – der Gedanke zugrunde, dass ein Fahrer dazu neigt, zusätzlichen Belastungen entgegen zu wirken, statt diese beispielsweise durch einen Spurwechsel weiter zu erhöhen (siehe Kapitel 4.1.1). Die Modellierung des Längsführungsverhaltens ist in zwei separate Blöcke unterteilt, die jeweils durch ein an die Wahrnehmung des Fahrers angepasstes, adaptives Fuzzy-Inferenzsystem repräsentiert werden. Beide Fuzzy-Inferenzsysteme besitzen die gleiche Struktur und die Regeln der beiden Regelbasen sind folgendermaßen formuliert:

$$R_l : \text{WENN } x_1 = \tilde{F}_1^l \text{ UND } \dots \text{ UND } x_n = \tilde{F}_n^l \text{ DANN } y = \tilde{G}^l \quad (l = 1, \dots, k)$$

wobei k und n die Zahl der Regeln bzw. der Eingangsgrößen beschreiben. Die Fuzzy-Menge \tilde{F}_i^l der Eingangsvariablen x_i wird durch folgende Gauß'sche Zugehörigkeitsfunktion beschrieben:

$$\mu_{\tilde{F}_i^l}(x_i) = \exp \left[- \left(\frac{x_i - c_i^l}{\sigma_i^l} \right)^2 \right] \quad (4.12)$$

4. Sensor-basierte Fahrerüberwachung

mit σ_i^l als Standardabweichung und c_i^l als Mittelwert. Zur Vereinfachung der Optimierung werden Singleton Zugehörigkeitsfunktionen für die Ausgabe verwendet. Der Ausgabe- wert v wird nach dem *Center of Singletons* Verfahren (COS/COGS) bestimmt:

$$v = \frac{\sum_{l=1}^k \left[\left(\prod_{i=1}^n \mu_{\tilde{F}_i^l}(x_i) \right) \cdot s_{G^l} \right]}{\sum_{l=1}^k \left(\prod_{i=1}^n \mu_{\tilde{F}_i^l}(x_i) \right)} \quad (4.13)$$

mit s_{G^l} als Position des Singleton G^l . Die Ausgabe-Zugehörigkeitsfunktionen werden zur Laufzeit über ein Gradientenabstiegsverfahren mit einer kleinen Lernrate η optimiert,

$$\Delta s_{G^l} = -\eta \frac{\partial E^2}{\partial s_{G^l}} \quad (4.14)$$

wobei $E = [v_{act} - v]$ den Modellierungsfehler⁵ beschreibt. Die Optimierung bricht ab, wenn eine bestimmte Grenze E unterschritten ist. Die Fuzzy-Inferenzsysteme modellieren die Abstandsregelung und das Beschleunigungsverhalten des Fahrers auf folgende Weise [171, 191]:

Abstandsregelung Primäres Ziel einer Folgefahrt ist es, einen „angemessenen“ Abstand zum Vorderfahrzeug zu halten. Die Festlegung dieses Abstandes erfolgt in der Regel

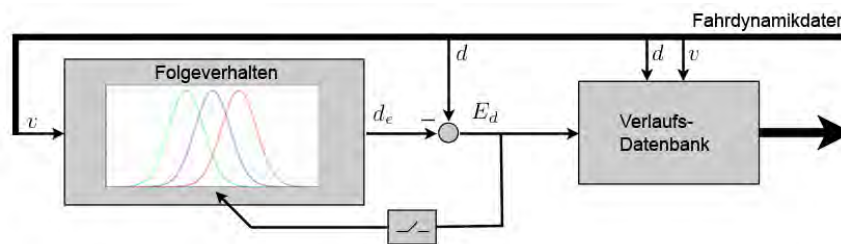


Abbildung 4.3: Modell der Abstandsregelung.

anhand der aktuellen Geschwindigkeit und des subjektiven Verständnisses von „angemessenem“ Abstand. Das erste Fuzzy-Inferenzsystem greift diese Annahme auf, indem es anhand der aktuellen Geschwindigkeit v des Ego-Fahrzeuges den Wunschabstand d_e schätzt. Der Agent bestimmt den Schätzfehler durch den Vergleich des geschätzten Wertes mit dem tatsächlich gewählten Abstand d . Der Schätzfehler $E_d = d - d_e$ repräsentiert die Differenz zur typischen Abstandsregelung. Sofern das aktuelle Fahrverhalten als stabil betrachtet werden kann, verwendet der Agent den Fehler e_d zusätzlich zur Kalibrierung der Ausgangs-Zugehörigkeitsfunktionen. Dafür überwacht der Agent die eingehenden Geschwindigkeits- und Abstandsdaten.

Beschleunigungsverhalten Während Folgefahrten steuert der Fahrer sein Fahrzeug in der Regel entsprechend seiner Wahrnehmung der Distanz zum vorausfahrenden Fahrzeug d , seiner eigenen Geschwindigkeit v und der relativen Geschwindigkeit $v_{rel} = v_{hw} - v$.

⁵Berechnet aus dem Vergleich von Modellergebnis v und dem tatsächlichen Wert v_{act} .

Abhängig von der wahrgenommenen Veränderung dieser Werte, versucht der Fahrer die Situation durch Beschleunigen oder Abbremsen seines Fahrzeugs in einem stabilen Zustand zu halten. Ohne äußere Einflüsse, z.B. kein oder sehr weit entfernt vorausfahrendes Fahrzeug, hält der Fahrer typischerweise eine selbst gewählte Geschwindigkeit v_{des} . Das zweite Fuzzy-System modelliert dieses Verhalten, indem es die Beschleunigung a_e abhängig von d , v , v_{rel} und v_{des} ⁶ schätzt. Der Schätzfehler $E_a = a - a_e$ dient der Kalibrierung des Modells. Da viele menschliche Prozesse durch Unsicherheit bzw. Rauschen geprägt sind,

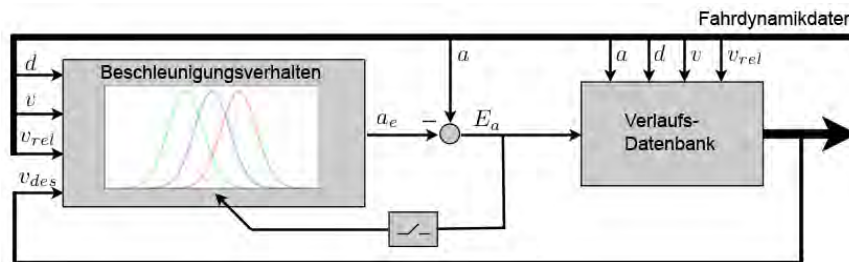


Abbildung 4.4: Modell des Beschleunigungsverhaltens.

fallen Aktionen eines Fahrers nicht konstant und eindeutig reproduzierbar aus, sondern variieren zu einem gewissen Grad: Z.B. ist es unmöglich für einen Fahrer, ständig einen exakten Abstand oder eine exakte Geschwindigkeit zu halten. Derartige Abweichungen stellen eine wesentliche Komponente des Führungsverhaltens eines Fahrers dar, weshalb diese Schätzfehler kontinuierlich in der Datenbank gespeichert werden.

4.3.2 Abnormalitätsbestimmung

Anhand der im vorherigen Abschnitt beschriebenen Modellierungen stehen dem Agenten-System zwei wesentliche Fahrerparameter zur Verfügung. Durch die Verwendung von kleinen Lernraten (in den folgend vorgestellten Experimente Werte ≤ 0.5) ist der Agent in der Lage, das Führungsverhalten eines Fahrers während eines gewissen Zeitfensters zu „erlernen“. Aus dem Vergleich der Mittelwerte und den tatsächlichen Messwerten, kann das System den Grad der Abweichung von der typischen Fahrweise bestimmen. Da ein menschlicher Fahrer keinen hoch präzisen Sensor/Aktor repräsentiert, ist immer eine Abweichung vorhanden. Diese Schwankungen charakterisieren die Fahrweise eines Fahrers und ihr Ausmaß spiegelt die Fähigkeit eines Fahrers wider, eine Konstantfahrt zu halten. Die Leistungsfähigkeit unterliegt Veränderungen durch physische und psychische Einflüsse (vgl. Kapitel 4.1.1). Besonders im Falle von (zusätzlichen) Nebenaufgaben, zeigen sich Ablenkungseffekte vor allem in erhöhten Schwankungen der Geschwindigkeit, des Abstands und der Beschleunigung. Daher erlauben Veränderungen in den regelmäßigen Schwankungen des Fahrverhaltens Rückschlüsse auf einen veränderten Fahrerzustand.

Durch Akkumulation aller Schätzfehler in einem bestimmten Zeitfenster, in dem Konstantfahrt Bedingungen vorliegen,⁷ modelliert das System die typischen Abweichungen

⁶ v_{des} ist dabei ein Durchschnittswert, der auf empirischen Daten aus früheren Zeiträumen der Fahrt basiert.

⁷Da der Fahrer z.B. ungestört einem Vorderfahrzeug folgen kann.

4. Sensor-basierte Fahrerüberwachung

im Verhalten eines Fahrers. Die Fuzzy-Inferenzsysteme generieren aus diesen Daten eine entsprechende Häufigkeitsverteilung. Um die Höhe der Abweichung, und somit die entsprechende Höhe des Leistungseinbruchs, zu bestimmen, verfügt der verantwortliche Agent über ein zusätzliches Fuzzy-Inferenzsystem, das in regelmäßigen Abständen nach der Häufigkeitsverteilung feinjustiert wird. Das Fuzzy-System bewertet das Führungsverhalten durch Vergleich der Ist-Abweichung mit den Attributen der Häufigkeitsverteilung in einem Bereich zwischen „keine Auffälligkeiten“ und „extremer Anomalie“.

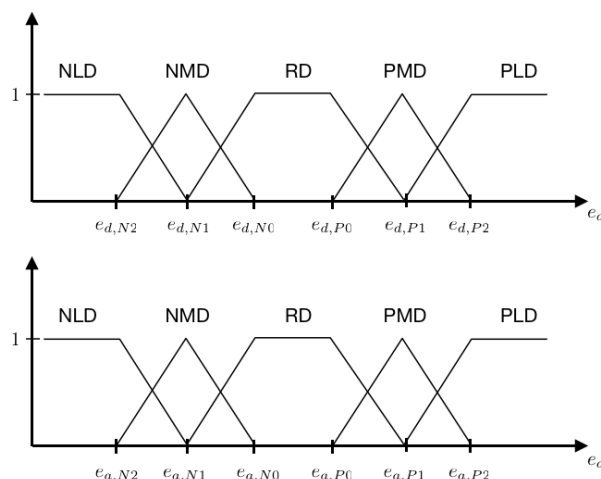


Abbildung 4.5: Abnormalitätsbestimmung, Partitionierung der Eingangsvariablen (vgl. Tabelle 4.1).

Als Eingangsvariablen ergeben sich die skalierten Fehler aus der Schätzung von Abstand ($e_d = \frac{d-d_e}{d_e}$) und Beschleunigung ($e_a = \frac{a-a_e}{a_e}$). Die Partitionierung der Eingangsvariablen ist in Abbildung 4.5 dargestellt, wobei sich die Partitionierungsparameter aus den Fehlerhäufigkeitsverteilungen (vgl. Abbildung 4.6) ergeben. Der Wert $e_{d,p0}$ stellt

Linguistische Ausdrücke, Eingabe	Linguistische Ausdrücke, Ausgabe
große positive Abweichung (NLD)	keine Abnormalität (NA)
mittlere positive Abweichung (NMD)	kleine Abnormalität (SA)
normale Abweichung (RD)	mittlere Abnormalität (MA)
mittlere negative Abweichung (PMD)	große Abnormalität (LA)
große negative Abweichung (PLD)	extreme Abnormalität (EA)

Tabelle 4.1: Abnormalitätsbestimmung, linguistische Ausdrücke.

die positive Grenze der typischen Abweichung in der Abstandsregelung eines Fahrers dar und ist durch das Intervall $[0, e_{d,p0}]$ in der Fehlerverteilung (vgl. Abbildung 4.6) definiert, in dem sich 90% der positiven kleinsten Werte befinden, sofern die Grenzen $0.05 \leq e_{d,p0} \leq 0.20$ nicht überschritten sind. Andernfalls finden die entsprechenden Grenzwerte Verwendung. Die positive Grenze $e_{d,p2} = 1,0$ entspricht dem doppelten

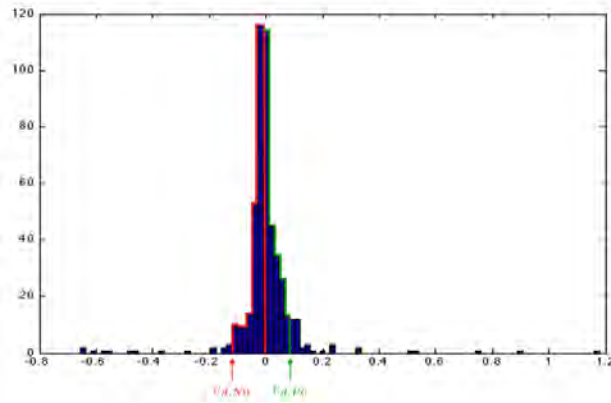


Abbildung 4.6: Abnormalitätsbestimmung, Partitionierungsparameter aus den Fehlerhäufigkeitsverteilungen.

typischerweise gewählten Abstand zum Vorderfahrzeug und $E_{d,P1}$ entspricht der Mitte zwischen $e_{d,P0}$ und $e_{d,P2}$. Im Gegensatz dazu entspricht $e_{d,N0}$ der negativen Grenze der typischen Regelabweichung und wird in ähnlicher Weise wie $e_{d,P0}$ jedoch als negative Hälfte der Verteilung mit $-0.15 \leq e_{d,N0} \leq -0.05$ als Grenzen definiert. Die negative Grenze beschreibt den kritischen Abstand, bei dem eine sofortige Vollbremsung zur Vermeidung einer Kollision nötig ist und ist folgendermaßen definiert

$$e_{d,N2} = \frac{1}{d_{e,avg}} \left(\frac{v_{avg}^2}{2 \cdot a_{max-}} - d_{e,avg} \right)$$

mit v_{avg} bzw. $d_{e,avg}$ als Schätzwerte der durchschnittlichen Geschwindigkeit bzw. Entfernung im betrachteten Zeitraum und a_{max-} als der maximalen Bremsleistung des Fahrzeuges. $e_{d,N1}$ ist der Mittelpunkt zwischen $e_{d,N0}$ und $e_{d,N2}$. Der Bereich des typischen Beschleunigungsverhaltens ist durch $e_{a,N0}$ und $e_{a,P0}$ definiert, die analog zu $e_{d,P0}$ jedoch ohne Begrenzung bestimmt werden. $e_{a,N1}$ und $e_{a,P1}$ entsprechen der maximalen Brems- bzw. Beschleunigungsleistung. $e_{a,N2}$ ist der Mittelpunkt zwischen $e_{a,N1}$ und dem Fehlerwert zur maximalen Bremsleistung a_{max-} . Ferner ist der Wert $e_{a,P2}$ der Mittelpunkt zwischen $e_{a,P1}$ und dem Fehlerwert zur maximalen Beschleunigung a_{max+} . Die feste Partitionierung der Ausgabevariablen ist in Abbildung 4.7 dargestellt.

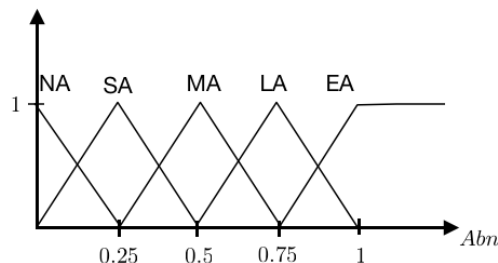


Abbildung 4.7: Abnormalitätsbestimmung, Partitionierung der Ausgabevariablen.

4.4 Systembewertung

AWARE wurde zunächst im Rahmen einer initialen Simulatorstudie hinsichtlich der Funktionalitäten zur Erkennung von Abnormalität bzw. der Manöverprädiktion untersucht. Dazu wurde im Fahrsimulator ein Rundkurs bestehend aus Autobahnabschnitten (vornehmlich gerade Streckensegmente mit zwei Fahrspuren pro Richtung) sowie innerstädtischen Bereichen (Geradeaus- sowie Kreuzungssegmente mit einer Fahrspur pro Richtung) zusammengestellt. Über den gesamten Streckenverlauf sind Fremdverkehrsfahrzeuge verteilt, die mit weitgehend konstanter Geschwindigkeit auf der Idealspur fahren und vorbestimmte Fahrmanöver durchführen, um gewünschte Verkehrssituationen zu induzieren. Die Geschwindigkeit wurde dabei auf einen relativ geringen Wert festgesetzt, um Überholvorgänge seitens der Versuchspersonen zu provozieren.

Zu Beginn waren die Testpersonen angewiesen, einem Fremdverkehrsfahrzeug mit selbst gewähltem Abstand und Geschwindigkeit auf der Autobahn zu folgen. Zu vorbestimmten Zeit- bzw. Wegpunkten leitete das vorausfahrende Fahrzeug abrupte Geschwindigkeitsänderungen ein, um das typische Folgeverhalten (z.B. Reaktion auf Abstandsvariationen) des Probanden zu erfassen. Im Verlauf des Versuches wurden solche Szenarien ebenfalls während der Bearbeitung von Nebenaufgaben (z.B. Navigationseingabe) sowie in Bereichen von „freier Fahrt“ ohne Anweisungen eingeleitet und aufgezeichnet. Dieser Ablauf wiederholte sich anschließend in städtischen Bereichen, wobei der Fahrer zusätzliche Instruktionen für Abbiegevorgänge bearbeiten musste.

Im Rahmen der Versuche zeigte sich, dass die verwendeten Modellierungsansätze sehr gut dazu geeignet sind, das typische Längsführungsverhalten der Testpersonen zu approximieren. Ebenfalls konnte das System zur Abnormalitätsbewertung in den meisten Fällen bestimmen, ob eine Änderung im Fahrverhalten der Probanden vorlag. Hohe Werte von langer Dauer traten dabei während der Bearbeitung von Nebenaufgaben auf. Dies deckt sich ebenfalls mit der subjektiven Einschätzung der Probanden. Abbildungen 4.8 und 4.9 zeigen exemplarisch zwei Ausgabesequenzen (Rohdaten) des entsprechenden Agenten. Der obere Teil entspricht der Distanz zwischen Ego-Fahrzeug und Fremdverkehr, der mittlere dem Verlauf der Beschleunigung (blau: tatsächlicher Verlauf; grün: Schätzwert) und der untere dem abgeleiteten Trend der Abnormalität.

Die gezeigten Sequenzen umfassen drei Teilbereiche: Der erste Bereich entspricht einer Folgefahrt mit vom Probanden selbst gewähltem Abstand zum Vorderfahrzeug. In diesem Intervall konnte der Agent die typische Fahrweise des Probanden „lernen“. Im zweiten Teil bearbeitete der Proband parallel zur Folgefahrt eine Nebenaufgabe (Eingabe eines Navigationszieles, rote Markierung im unteren Teil der Abbildung). Im letzten Abschnitt konnte der Proband frei fahren.

Besonders auffällig sind die deutlich höheren und dichten Abnormalitätswerte während der Bearbeitung der Nebenaufgabe, die durch eine erhöhte Variation im Fahrverhalten im Vergleich zum vorangegangenen Abschnitt der Fahrt verursacht wurden. Dieses Phänomen zeigte sich in 70% aller gültigen Versuchsdurchläufe. Die meisten Probanden konnten ihren typischen Fahrstil aufgrund der zusätzlichen Interaktion mit dem Navigationssystem nicht aufrechterhalten. Die Systembewertungen des letzten Fahrabschnittes liegen darin begründet, dass fast alle Probanden deutlich aggressiver fuhren und alle Fremdverkehrsfahrzeuge überholten. Dieses Verhalten ist am „sägezahnförmigen“

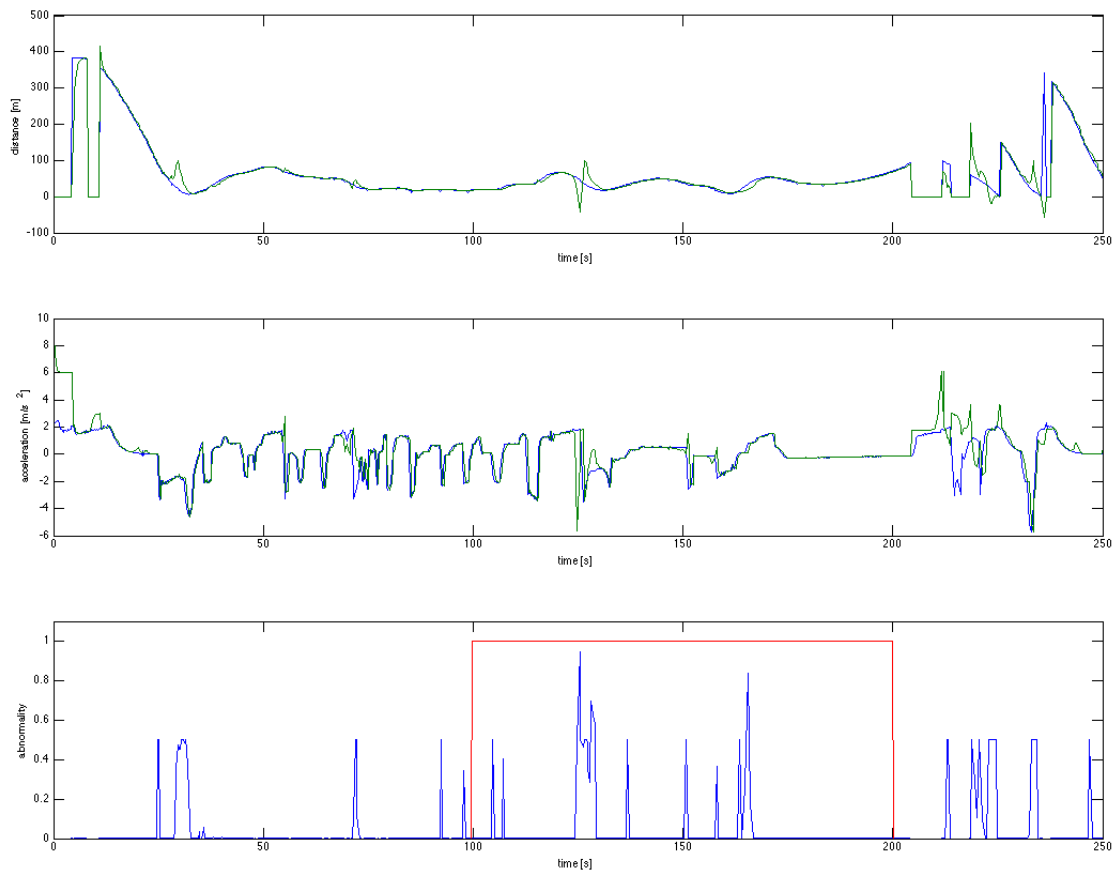


Abbildung 4.8: Beispielsequenz 1: Rohdaten der Abnormalitätsbewertung (blau: tatsächlicher Verlauf; grün: Schätzwert). Oben: Abstand zum Vorderfahrzeug. Mitte: Beschleunigung des Egofahrzeuges. Unten: abgeleiteter Abnormalitätswert.

Verlauf der Abstandskurven zum Ende hin ersichtlich, welcher auch die entsprechenden Abnormalitätswerte erklärt. Diese Werte werden jedoch im Normalfall durch den nächsten Verarbeitungsschritt ebenso herausgefiltert, wie (kurzzeitige) Abnormalitätswerte, die durch ein den vordefinierten Sensorerfassungsbereich betretendes oder verlassenes Fremdverkehrsfahrzeug hervorgerufen sind.

4. Sensor-basierte Fahrerüberwachung

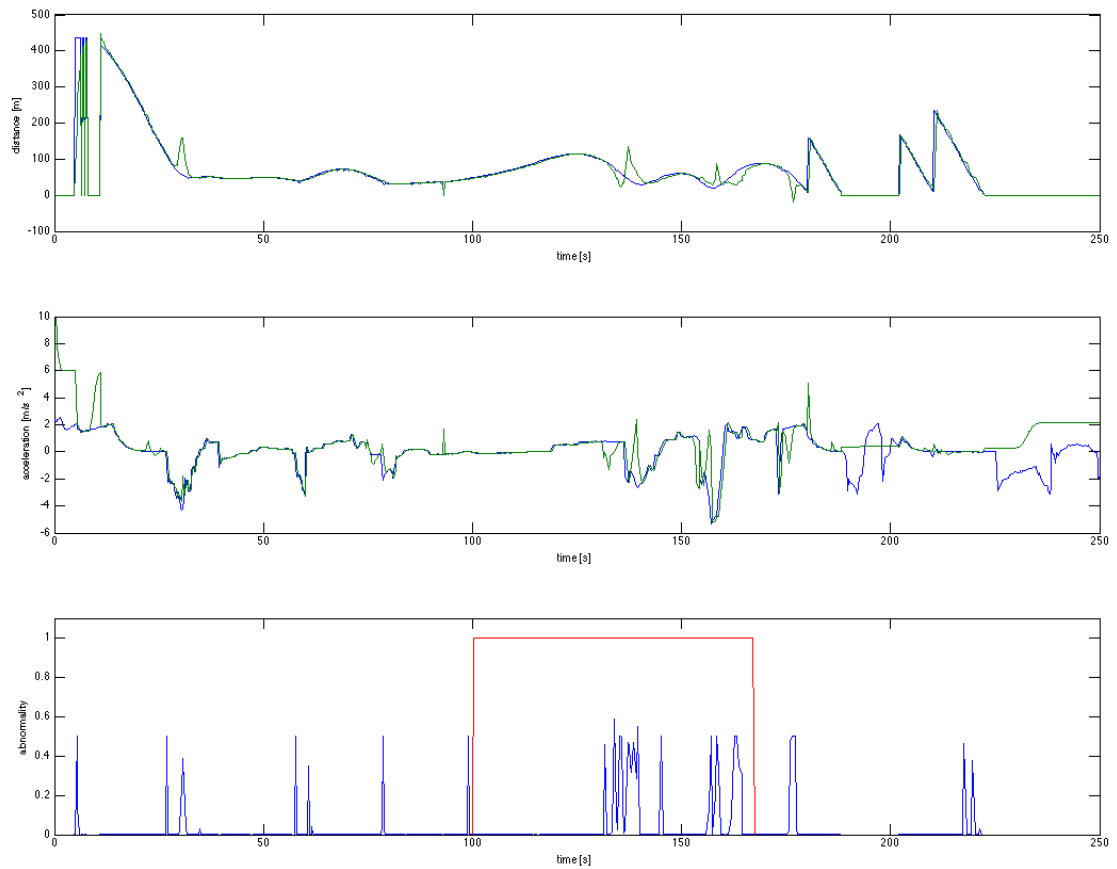


Abbildung 4.9: Beispielsequenz 2: Rohdaten der Abnormalitätsbewertung (blau: tatsächlicher Verlauf; grün: Schätzwert). Oben: Abstand zum Vorderfahrzeug. Mitte: Beschleunigung des Egofahrzeuges. Unten: abgeleiteter Abnormalitätswert.

Attentive User Interfaces

Das folgende Kapitel 5 stellt, aufbauend auf den Erkenntnissen aus den Kapiteln 3, 4 und 2.5, exemplarisch Konzepte zur adaptiven Anpassung des Informationsflusses im Mensch-Maschine System Fahrer-Fahrzeug an den aktuellen Fahrerzustand vor.

Unfallstatistiken zufolge spielt Unaufmerksamkeit in einem zunehmend höheren Prozentsatz aller Unfälle eine zunehmende Rolle. Die amerikanischen National Highway Traffic Safety Administration stellte z.B. in einer Untersuchung fest, dass ca. 25 % aller Unfälle Unaufmerksamkeit oder Ablenkung des Fahrers zugrunde liegt [199]. Eine österreichische Studie identifiziert bei rund 10 % der Unfälle mit tödlichem Ausgang ein Defizit der Aufmerksamkeit als Ursache [156]. Das Statistische Bundesamt schlüsselt zwar nur nach Unfallursachen (z.B. zu hohe Geschwindigkeit) auf, man kann jedoch davon ausgehen, dass diese Zahlen in Deutschland nicht signifikant davon abweichen.

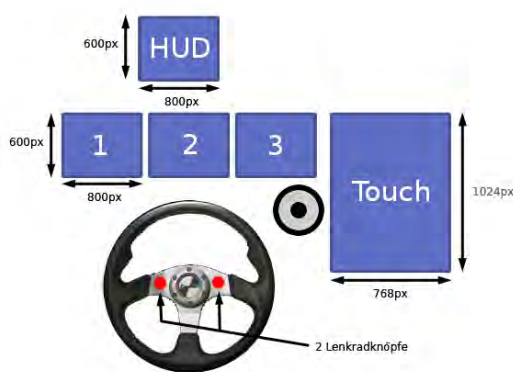
Das Highway Safety Research Center hat sich eingehender mit Fahrerablenkung beschäftigt. *The Role of Driver Distraction in Traffic Crashes* [225] wertet Daten zur genauen Art der Ablenkung genauer aus. Als Ergebnis lässt sich festhalten, dass in 8,3 % der Unfälle der Fahrer zum Zeitpunkt des Unfalls abgelenkt war, was jedoch eine eher konservative Schätzung darstellt, da allein in ca. 1/3 der Fälle der Fahrerzustand als 'nicht bekannt' eingestuft worden ist. Von der Zahl der als abgelenkt bekannten Fälle entfallen 11,9 % auf die Bedienung eines Radios/CD Players, 2,8 % auf die Bedienung der Klimaanlage und 1,5 % auf die Benutzung eines Mobiltelefons. Zu beachten ist hierbei allerdings, dass diese Statistik auf Zahlen aus den Jahren 1995 bis 1999 basiert. Da seitdem die Anzahl der Mobiltelefone und die Verwendung von Navigationsgeräten gestiegen sind, sind diese Zahlen eher als steigend denn als stagnierend zu betrachten.

Der Anstieg ist vor allem damit zu begründen, dass die Ablenkungsgefahr mit zunehmender Verbreitung von Navigationssystemen sehr hoch ist. Im Gegensatz zur Bedienung eines Radios oder der Eingabe einer Telefonnummer, benötigt die Eingabe eines Navigationszieles mehr Zeit [225] und die (visuelle) Aufmerksamkeit ist somit für längere Zeit nicht auf der Straße sondern im Auto. Aufgrund der stetig wachsenden Informationsflut in modernen Kraftfahrzeugen gilt es, neue Methoden zu entwickeln, die den Fahrer dabei unterstützen, diese zusätzlichen Aufgaben möglichst schnell und effizient zu bewältigen.

Ziel dieses Kapitels ist es daher, Konzepte vorzustellen, die durch eine aufmerksamkeitsbasierte Informationsausgabe dem Fahrer Informationen in angemessener Weise zugänglich machen, um die Ablenkung von der eigentlichen Fahraufgabe zu minimieren.

5.1 ABK2020 – ein Basisframework für Anzeige-Bedienkonzepte im Automobil

Das Framework *ABK2020* wurde im Rahmen der Umsetzung einzelner Anzeige-Bedienkonzepte entworfen und dient als Grundlage für die im Folgenden vorgestellten Konzepte. *ABK2020* stellt alle grundlegenden Funktionen einheitlich zur Verfügung: die Darstellung von Menüs, die Kommunikation einzelner Komponenten untereinander, Schnittstellen zur Fahrsimulation und der restlichen Laborinfrastruktur, den Eye Tracker und weiteren Eingabegeräte (z.B. Multifunktionslenkrad und Touchscreen). Das System greift den aktuellen Trend hin zu großflächigen Multi-Display Anzeige-Bedienkonzepten (vgl. Mercedes-Benz Comand) im Fahrzeug auf. Deshalb wurde bei der Konzeption gezielt die Verwendung großflächiger Multi-Display Anzeigesysteme adressiert: ein Head-Up Display (im Folgenden als *HUD* bezeichnet), ein zentrales Informationsdisplay mit Touchinteraktion (im Folgenden als *CID* bezeichnet) sowie ein großflächiges, freiprogrammierbares Kombiinstrument (im Folgenden als *FPK* bezeichnet).



(a) Schematische Darstellung.



(b) Verwendete Sitzkiste.

Abbildung 5.1: Displayaufbau mit dreigeteiltem FPK, HUD und CID Touchscreen.

Zur Anzeige der Menüstrukturen bietet die Sitzkiste mehrere Displays (vgl. Abbildung 5.1). Als CID dient ein hochkant montiertes 15" 4:3 TFT Display mit einer Auflösung von 1024x768 pix. Zusätzlich ist das Display mit einem kapazitiven Sensoraufsatz¹ ausgestattet, der das Display zum Touchscreen erweitert. Die Detektion der Handposition (in 3D) erfolgt mittels der in [140] vorgestellten Ansätze und Algorithmen. Das FPK ist aus drei separat ansprechbaren Displayteilen (7" 4:3 TFT Displays mit einer Auflösung von 800x600 pix) aufgebaut. Abhängig vom getesteten Prototypen kommt als HUD entweder ein großflächiger Aufbau zum Einsatz, welcher das Bild wie im Fahrzeug auf die Frontscheibe projiziert. Zwei Beamer (mit einer Auflösung von je 1280x1024 pix) projizieren dabei das virtuelle Bild auf eine reflektierende Fläche, von dort wird es reflektiert und als virtuelles Bild in der Frontscheibe dargestellt. Das Bild erstreckt

¹Aufsatz ZYTOUCH der Firma zytronic. Nach Angabe von zytronic liefert dieser Aufsatz Daten mit einer Abweichung von weniger als +/- 1,5 % von der tatsächlichen Position der Hand [261]. Das entspricht bei der Auflösung von 768x1024 Pixel einer Abweichung von 11 bis 15 Pixel.

sich über den Bereich der gesamten Frontscheibe. Dieser Aufbau ist prädestiniert für Versuche mit kontaktanalogen HUDs [189]. Da sich dieses realitätsnahe HUD Bild schlecht kalibrieren lässt² (vgl. Kapitel 3.2.3), kommt für Versuche mit Blickbedienung ein auf der Sitzkiste montiertes TFT Display zum Einsatz (vgl. Abbildung 5.1(a)).

Die graphischen Oberflächen der Displays sind als einzelne Applikation in Adobe Flash umgesetzt. Zur Kommunikation zwischen den verteilt auf den Displays dargestellten Menüteilen ist eine auf dem TCP/IP Protokoll basierende Kommunikationsstruktur über Netzwerk integriert. Dabei dient das CID als Server, der die Verbindungen zu den einzelnen Displays verwaltet. Weiterhin verwaltet diese Applikation die gesamte (Anzeige-)Logik im System und verteilt – abhängig vom zu erstellenden Systemprototypen – die Anzeigehalte auf die entsprechenden Displays. Zusätzlich verwaltet die CID Anwendung die Daten der externen Eingabegeräte vom Multifunktionslenkrad bis hin zum Eye Tracker und stellt diese allen im Verbund genutzten Anzeigeelementen zur Verfügung (z.B. Menübedienung auf dem FPK über das Multifunktionslenkrad).

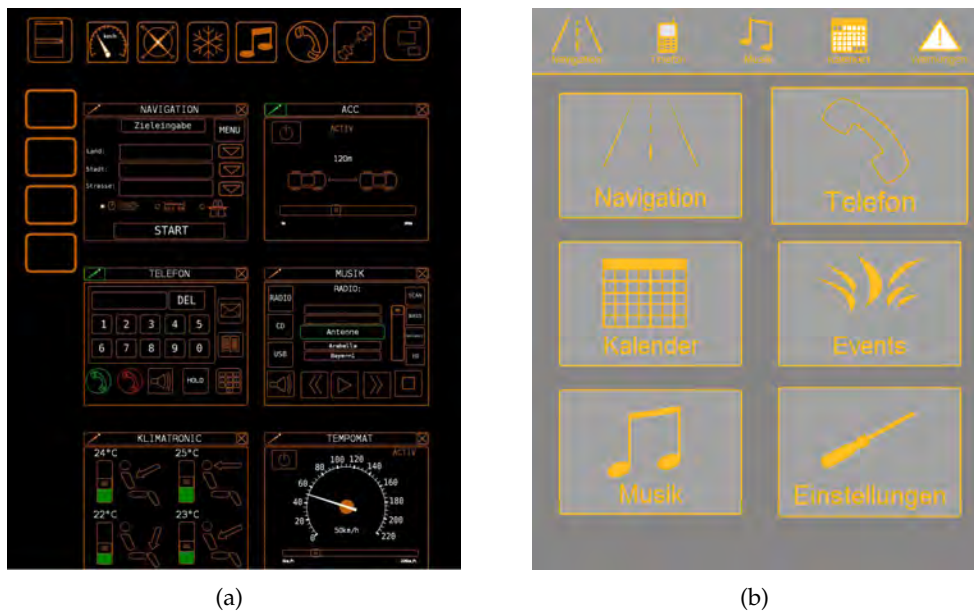


Abbildung 5.2: Exemplarische graphische Oberflächen für das CID, die mittels ABK2020 realisiert wurden.

5.2 Örtliches Scheduling und iterative Anpassung der Tiefe von fahrzeuginitiierten Dialogen

Ziel des im Folgenden präsentierten adaptiven Ansatz ist es, durch eine aufmerksamkeitsbasierte Informationsausgabe, dem Fahrer die bereits angesprochene Informationsflut in modernen Kraftfahrzeugen in angemessener Weise zugänglich zu machen und die

²Zur Verwendung der in Kapitel 3 vorgestellten Blickbedienung müssen die exakten 3D Koordinaten der Displayfläche bekannt sein. Die exakten Koordinaten der virtuellen Anzeigefläche dieses Aufbaus lassen sich jedoch praktisch nicht vermessen.

Ablenkung von der (primären) Fahraufgabe zu minimieren. Die Grundidee hierbei ist es, Informationen dort anzuzeigen, wo ohnehin schon die Aufmerksamkeit des Fahrers liegt (siehe Kapitel 5.2.1), und somit eine schnellere Aufnahme und Verarbeitung dieser Informationen zu ermöglichen. Dazu werden drei unterschiedliche Konzepte zur adaptiven Dialoggestaltung (Menge der präsentierten Informationen pro Dialogschritt, siehe Kapitel 5.2.2) vorgestellt und miteinander verglichen.

5.2.1 Örtliches Scheduling von fahrzeuginitiierten Dialogen

Das Anzeigekonzept basiert auf der in Kapitel 5.1 vorgestellten Grundstruktur. Ebenfalls kommen die beschriebenen verteilten großflächigen Anzeigorte (großflächiges FPK, HUD, CID) zum Einsatz. Das Hauptmenü wird auf dem CID (vgl. Abbildung 5.2(b)) angezeigt und enthält im Grundzustand sechs Buttons zum Direkteinsprung in die Menüs für Navigation, Telefon, Kalender, Events, Musik und Einstellungen. Zusätzlich wird die Statusleiste mit Symbolen verwendet, die das Eintreffen von Systemmeldungen symbolisieren. Die Anzeigorte FPK und HUD sollen ebenfalls eine Bedienung des Systems (per

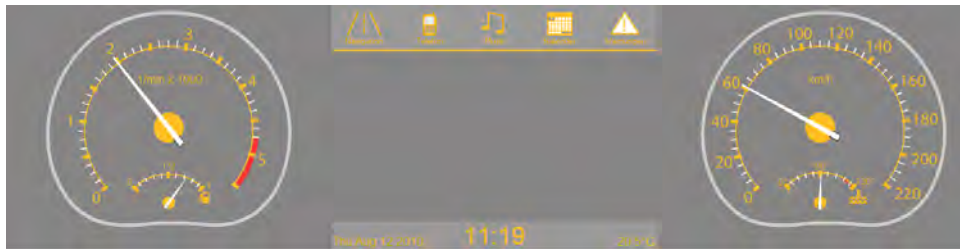


Abbildung 5.3: GUI des FPK mit Geschwindigkeitsanzeige, Drehzahlanzeige, sowie Uhrzeit, Datum und Temperaturanzeige. Oben in der Mitte, befinden sich die Symbole für die einzelnen Systemmeldungen.

Blickbedienung) ermöglichen. Die Darstellung orientiert sich am CID, der bedienbare Funktionsumfang ist jedoch signifikant geringer als im Hauptmenü. Das FPK (vgl. Abbildung 5.3) enthält eine Geschwindigkeitsanzeige, eine Drehzahlanzeige und eine Anzeige für die eintreffenden Systemmeldungen. Das HUD (vgl. Abbildung 5.4) zeigt typische Inhalte wie Navigationspfeil, Geschwindigkeit, Tempolimits, etc. an. Außerdem ist die Darstellung eintreffender Systemmeldungen vorgesehen. Für die Blickbedienung wird ein Target an den berechneten Pixelkoordinaten des Blickpunktes angezeigt. Um diesen nicht zu aufdringlich und ablenkend wirken zu lassen, erfolgt die Darstellung leicht transparent. Da Multidisplaysysteme eine Reihe neuer Eingabe- und Ausgabemöglichkeiten ermöglichen, wurden zwei verschiedene Informationsausgabestrategien konzipiert:

1. **Nicht-adaptive Informationsausgabe:** Dies stellt die konventionelle Methode dar, Meldungen bzw. Informationen an fixen Anzeigorten im Fahrzeug darzustellen. Dabei erscheinen Meldungen immer am selben Ort.
2. **Adaptive Informationsausgabe:** Diese Methode nutzt die Möglichkeit, Informationen dynamisch an verschiedenen Anzeigorten (FPK, HUD und CID) anzuzeigen. Systemmeldungen erscheinen in der Nähe des aktuellen Aufmerksamkeitsfokus des

Fahrers. Konzentriert sich der Fahrer z.B. auf die Straße, dient das HUD zur Ausgabe der Meldungen. Bedient der Fahrer das Hauptmenü des Infotainmentsystems im CID (z.B. Eingabe eines neuen Navigationszieles), werden die Systemmeldungen in das CID umgeleitet. Liegt die Aufmerksamkeit des Fahrers zum Zeitpunkt einer Systemmeldung auf dem Kombiinstrument, so werden diese dort angezeigt.



Abbildung 5.4: GUI des HUD mit Geschwindigkeitsanzeige, Navigationspfeil und Radiofrequenzanzeige.

Zu diesen Verfahren ist anzumerken, dass der Fahrer Meldungen bei nicht-adaptiver Informationsausgabe als ablenkender empfinden kann, da er die Aufmerksamkeit für längere Zeit von der Straße abwenden muss. Weiterhin kann man davon ausgehen, dass der Fahrer Meldungen bei adaptiver Informationsausgabe schneller wahrnimmt, als bei nicht-adaptiver Informationsausgabe. Zudem sollte die Bearbeitungsdauer einer Meldung geringer ausfallen, da er seine Aufmerksamkeit nicht zusätzlich verlagern muss.

Die adaptive Möglichkeit der Informationsausgabe bietet großes Potential und einen immensen Spielraum für Optimierungen: Es ist z.B. möglich, Systemmeldungen der Aufmerksamkeit des Fahrers nachzuführen. Verschiebt der Fahrer seine Aufmerksamkeit von der Straße ins CID, kann auch die Meldung vom HUD ins CID verschoben werden. Dazu erfolgt eine Aufteilung des Fahrerarbeitsplatzes in unterschiedliche räumliche Zonen, die jeweils einem Anzeigeort/Display entsprechen. Um auf die Aufmerksamkeit des Fahrers schließen zu können, muss seine Blickrichtung bekannt sein. Dazu wird das in Kapitel 3.3 vorgestellte Eye Tracking System eingesetzt. Aus den berechneten Blickdaten leitet das System zu jedem Zeitpunkt ab, welches Display sich in nächster Nähe zum aktuellen Aufmerksamkeitsfokus befindet. Die Zone, in der sich die Straße befindet, ist folglich eindeutig dem HUD zugewiesen. Der Bereich der Mittelkonsole und die Beifahrerseite sind dem CID und der Bereich des Lenkrads und der Armaturen dem FPK zugeordnet.

5.2.2 Iterative Anpassung der Tiefe von Dialogen

Ein weiteres zentrales Anliegen war es, den Einfluss unterschiedlichen Informationsgehalts pro Dialogschritt zu untersuchen. Grundgedanke dabei ist es, die Reaktionen des Fahrers auf Systemausgaben in die Dialoggestaltung einfließen zu lassen: Das System signalisiert dem Fahrer im ersten Schritt den Bedarf einer Reaktion bzw. einer Eingabe seinerseits. Kann der Fahrer in der entsprechenden Situation Ressourcen zur Bearbeitung

der zusätzlichen Aufgabe aufwenden, setzt er sich mit der Systemausgabe auseinander bzw. schenkt dieser Aufmerksamkeit. Ist dies im aufgezeichneten Blickverhalten ersichtlich, wird die zunächst einfach gehaltene Systemausgabe mit weiteren Informationen näher spezifiziert und der Fahrer erhält weitere Informationen, in welcher Art und Weise sein Eingriff nötig ist. Dazu erfolgt – anhand von einfachen Dialogen, die entweder mit Zustimmung (z.B. *Ja*) oder Ablehnung (z.B. *Nein*) zu beantworten sind – die Untersuchung drei verschiedener Dialogabläufe:

- **Dialogablauf 1:** Bei Eintreffen einer Systemmeldung fängt zunächst das Icon des jeweiligen Themenbereichs an zu blinken (vgl. Abbildung 5.5(a)), um den Fahrer auf die neue Meldung aufmerksam zu machen. Durch Anschauen des Icons bzw. durch Anklicken des blinkenden Icons auf dem Touchscreen öffnet sich das kleine Betrefffenster. Dies beinhaltet erste Informationen über den tatsächlichen Inhalt der Meldung (vgl. Abbildung 5.5(b)). Bei Klick auf das Betrefffenster öffnet sich der komplette Dialog, der alle Informationen der Meldung enthält (vgl. Abbildung 5.5(c)). Zudem enthält dieser zwei Buttons, zum Zustimmung oder Ablehnen einer Aktion im Zusammenhang mit der Meldung. Der Dialog kann auch ohne vorzeitige Entscheidung über das X-Symbol unterbrochen werden.

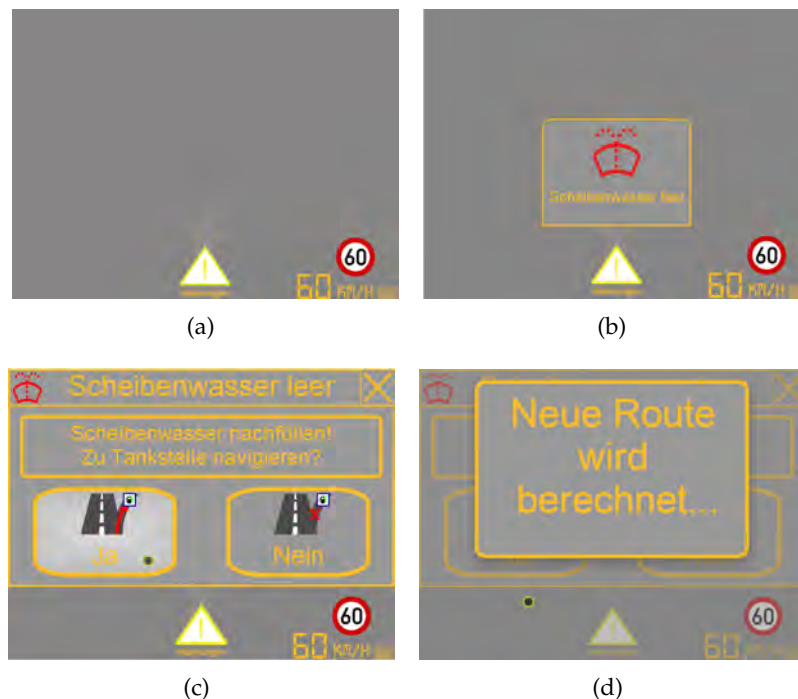


Abbildung 5.5: Beispiel einer Systemmeldung im HUD.

- **Dialogablauf 2:** Er verhält sich wie Dialogablauf 1, allerdings mit der Ausnahme, dass sich zu Beginn das kleine Betrefffenster (vgl. Abbildung 5.5(b)) sofort öffnet. Durch Anklicken öffnet sich der komplette Dialog. Um den gesamten Dialogablauf unterbrechbar zu machen, schließt sich das Betrefffenster nach einer bestimmten

Zeit wieder und es bleibt nur das blinkende Icon zurück. Im weiteren Verlauf ist der Dialog wie bei Dialogablauf 1 zu bearbeiten.

- **Dialogablauf 3:** Bei Eintreffen einer Systemmeldung öffnet sich ganz am Anfang schon der komplette Dialog mit den beiden Buttons zur Bearbeitung (vgl. Abbildung 5.5(c)). Alle Informationen werden hierbei in einem Schritt angezeigt. Durch Anklicken des X-Symbols kann man den Dialog unterbrechen und dieser reduziert sich auf das blinkende Icon.

Gegen eine Anzeige des kompletten Dialoges von Beginn an spricht die Überlegung, dass der Fahrer den kompletten Dialog – verglichen mit dem blinkenden Icon – als zu aufdringlich und ablenkend empfindet. Weiterhin erweisen sich Dialogablauf (DA) 2 und 3 als störend bei Eingabetätigkeiten, da sie ein Teil der Interaktionsfläche verdecken. Zusätzlich können Reaktionen auf DA 3 schneller erfolgen, da Meldungen allein aufgrund der Anzeigegröße sehr schnell wahrnehmbar sind. DA 1 erfordert viele Eingaben und somit eine längere Bearbeitungsdauer, wohingegen DA 3 lediglich eine Eingabe benötigt.

Abbildung 5.5 zeigt exemplarisch einen Dialogablauf für die Warnung bei zu geringem Wischwasserstand: Zunächst ist lediglich ein blinkendes Icon dargestellt, im nächsten Schritt das Betrefffenster mit reduziertem Informationsgehalt und im dritten Schritt erscheint der komplette Dialog mit Zustimmungs- und Ablehnungsbutton. Bei getätigter Eingabe erfolgt eine Rückmeldung und der Dialog gilt als beendet.

5.2.3 Versuchsdurchführung und -ergebnisse

Die Evaluierung der Konzepte erfolgte im Rahmen eines Versuchs mit 15 Probanden (13 männlich, Durchschnittsalter 23,6 Jahre). Lediglich eine Versuchsperson (VP) hatte im Vorfeld Erfahrungen mit einem Eye Tracker gesammelt. Ein Drittel der VPn hatte den Begriff Infotainmentsystem noch nie gehört. 40% hatten Erfahrung im Umgang mit Infotainmentsystemen gesammelt oder waren sogar im Besitz eines solchen. Der Großteil der VPn war technisch sehr interessiert und sicher im Umgang mit Technik.

Zu Beginn erhielten die VPn eine Einführung zum System und den bevorstehenden Aufgaben. Danach konnten sich die VPn an die Fahrsimulation (LCT) gewöhnen. Anschließend begann die erste Testfahrt. Hierbei mussten die VPn geforderte Zieleingaben in das Navigationssystem tätigen. Zusätzlich waren Systemmeldungen zu bearbeiten. Darauf folgte eine Testfahrt mit geänderter Informationsausgabestrategie. Um Einflüsse durch Lerneffekte auszuschließen, erfolgte eine Permutation der Reihenfolge bei den einzelnen VPn. Im Anschluss wurde eine Nachbefragung durchgeführt.

Die Probanden gaben an, Systemmeldungen bei adaptiver Informationsausgabe schneller zu bemerken, als bei einer Ausgabe nur im CID (statistisch signifikant). Es zeigte sich kein signifikanter Unterschied der gemessenen Reaktionszeiten. Die VPn erklärten, dass sich Systemmeldungen in der Nähe des Blickfeldes schneller beenden lassen, als eine Meldung im CID. Der Signifikanztest ergab sowohl für die Ergebnisse des Fragebogens als auch für die gemessenen Bearbeitungsdauern keine Ergebnisse.

Die VPn empfanden eine höhere Ablenkung durch Systemmeldungen bei nicht-adaptiver Informationsausgabe (Ausgabe aller Systemmeldungen im CID) als bei Erscheinen der Meldungen in der Nähe der aktuellen Blickrichtung. Diese Aussage ist statistisch

signifikant ($p < 0.0001$). Die VPn empfanden das Umschalten einer Systemmeldung in der Nähe ihres Blickfeldes nicht als aufdringlich – eine mittlere Ablehnung von 4,14 auf einer 5-fach Likert-Skala. Die VPn gaben an, dass die Abarbeitung einer Systemmeldung im CID ihre Fahrleistung negativ beeinflusste – im Durchschnitt eine starke Zustimmung von 1,8 auf einer 5-fach Likert-Skala. Insgesamt würden die VPn allerdings eine adaptive Informationsausgabe vor einer statischen bevorzugen – durchschnittlicher Wert von 2,07 auf einer 5-fach Likert-Skala.

Die Benutzbarkeit (erhoben über den SUS Fragebogen) der nicht-adaptiven Informationsausgabe wurde mit 84,5 besser bewertet (adaptiv 76,25), wobei der etwas schlechtere Wert der adaptiven Variante teilweise auf eine gestörte Funktionsfähigkeit des Eye Tracking Systems zurückzuführen ist³. Man kann erkennen, dass VPn, bei denen das Eye Tracking Systems zuverlässig und genau funktionierte, die adaptive Informationsausgabe besser bewerteten als die nicht-adaptive (nicht statistisch signifikant).

Die nicht-adaptive Informationsausgabe wurde im Durchschnitt mit 105,33 („ziemlich anstrengend“) und die adaptive mit durchschnittlich 89,29 als weniger anstrengend empfunden. Die subjektiv empfundene Beanspruchung (SEA Fragebogen) deckt sich somit auch mit der Aussage, dass Dialoge bei nicht-adaptiver Verteilungsstrategie ablenkender sind und eine adaptive Informationsausgabe vor einer statischen bevorzugt wird.

Hinsichtlich der pragmatischen Qualität (erfasst über den AttrakDiff Fragebogen) schneidet die adaptive Informationsausgabe (Produkt B) schlechter ab, als die nicht-adaptive Informationsausgabe (Produkt A). Dieser Unterschied ist nicht statistisch signifikant. Die hedonische Qualität der adaptiven Informationsausgabe schneidet (statistisch signifikant) besser ab. Das Konfidenzintervall der pragmatischen Qualität ist für die nicht-adaptive Informationsausgabe kleiner als das der adaptiven Informationsausgabe. Die VPn sind sich folglich bei der Beurteilung von Produkt A einiger als bei Produkt B. Das Konfidenzintervall der hedonischen Qualität ist bei Produkt B kleiner. Bei der hedonischen Qualität sind sich die VPn demnach bei Produkt B einiger. Die Konfidenzintervalle erstrecken sich über mehrere Charakterbereiche, weshalb eine eindeutige Charakterzuordnung der beiden Produkte nicht möglich ist. Das große Konfidenzintervall (pragmatische Qualität) der adaptiven Strategie lässt sich auf eine geringe Stichprobengröße oder sehr unterschiedliche Beurteilungen des Produktes zurückführen. Lediglich wenige VPn bevorzugten eindeutig die nicht-adaptive Variante. Die Auswertung der Fahrdaten ergab keine signifikanten Unterschiede zwischen den beiden Strategien zur Informationsausgabe. Es ergaben sich lediglich Tendenzen, dass die adaptive Informationsausgabe weniger ablenkend ist.

Die VPn gaben an, dass sie das zunächst nur blinkende Icon am wenigsten irritiert hat. Der komplett erscheinende Dialog wurde als irritierender empfunden (nicht signifikant, $p < 0.052$). Es konnte bestätigt werden ($p < 0.001$ im Vergleich zu DA 1 und $p < 0.0003$ im Vergleich zu DA 2), dass DA 3 subjektiv am aufdringlichsten ist. Die Aufdringlichkeit des blinkenden Icons und des sofort erscheinenden Betreffs ist (ungefähr) gleich. DA 3

³Probanden, bei denen der Tracker schlechter funktionierte, bewerteten die adaptive Variante signifikant schlechter, als die nicht-adaptive. An dieser Stelle sei auf die getrennte Analyse der Expertengruppe und die entsprechenden Ergebnisse in Kapitel 5.2.3.1 hingewiesen. Da eine Analyse der Expertengruppe durchgeführt wurde, sei an dieser Stelle auf eine getrennte Gegenüberstellung der Ergebnisse der VP verzichtet, bei denen der Tracker zufriedenstellend arbeitete.

rief bei den VPn die größte Ablenkung hervor, wobei die Ablenkungen durch das Icon und den Betreff etwa gleich empfunden wurden (statistisch signifikant).

DA 2 und 3 konnten (subjektiv) am schnellsten auf die Meldungen aufmerksam machen. Der Unterschied zwischen diesen beiden ist statistisch nicht signifikant. Der Unterschied zu DA 1 ist signifikant ($p < 0.008$ bzw. 0.003). Dies bestätigt die Annahme, dass die VPn diese Meldungen schneller wahrnehmen. Die Reaktionszeiten der einzelnen Abläufe zeigen für beide Ausgabestrategien keine signifikanten Unterschiede.

Die Bearbeitungsdauer hängt vor allem von der Anzahl Bedienschritte ab. Die Abarbeitung von DA 1 benötigte durchschnittlich 16,8s, bei DA 2 11,5s und bei DA 3 nur 4,9s. Deshalb sind hinsichtlich der Bearbeitungszeit Dialogabläufe mit höherem Informationsgehalt pro Schritt vorzuziehen. Subjektiv empfanden die VPn DA 3 bei anderen Eingabetätigkeiten als störender als das blinkende Icon (statistisch signifikant). Der Unterschied zwischen DA 2 und 3 ist nicht statistisch signifikant. Ein zeitliches Scheduling (vgl. Kapitel 5.4) kann das Problem der Störung während Eingabetätigkeiten vermeiden.

DA 2 erhielt mit einem SUS-Score von 83,5 die besten Bewertungen. Damit ist die Benutzbarkeit als sehr gut einzuschätzen. DA 1 bzw. DA 3 erhielten mit 76,16 bzw. 79,83 etwas schlechtere Bewertungen (kein signifikanter Unterschied). Das signifikant bessere Abschneiden von DA 2 lässt sich damit begründen, dass dieser die Vorteile von DA 3 (z.B. kürzere Bearbeitungsdauer) mit den Vorteilen von DA 1 (geringere Ablenkung) kombiniert. DA 2 wird mit einem durchschnittlichen SEA Wert von 80,66 als am wenigsten beanspruchend bewertet (DA 1: 97, DA 3: 86). Dies ist wiederum darauf zurückzuführen, dass DA 2 einen Kompromiss zwischen den Konzepten darstellt. Die Auswertung der Fahrdaten ergab keine signifikanten Unterschiede zwischen den drei Dialogabläufen.

Im Schnitt stimmten die VPn zu, dass ein blinkendes Icon ausreicht (2,06 auf 5-fach Likert-Skala). Die Aussage „Ich finde es gut, dass der Informationsgehalt mit der Dialogtiefe zunimmt“ erreichte mittlere Zustimmung (2,03) und die Aussage „Ich finde es gut, dass alle Informationen gleichzeitig angezeigt werden“ eher neutrale Bewertungen (2,86).

5.2.3.1 Ergebnisse der Expertengruppe

Der Versuch wurde mit einer „Expertengruppe“ wiederholt. Als Experten dienten Probanden, bei denen es während dem ersten Versuch keine Probleme mit dem Tracking System gab. Die Experten hatten zusätzlich die Möglichkeit, sich besser mit den Konzepten vertraut zu machen. Im Folgenden werden die Ergebnisse des Expertentests vorgestellt.

Sowohl die Bearbeitungszeiten als auch die mittlere Spurabweichung waren bei adaptiver Informationsausgabe besser, als bei nicht-adaptiver Informationsausgabe. Die Probanden erreichten eine mittlere Spurabweichung von 1,14 Meter bei der nicht-adaptiven Testfahrt und 1,06 Meter beim adaptiven Testfall. Dies entspricht einer Verbesserung von ungefähr 7%. Die Bearbeitungsdauer bei DA 1 verbesserte sich um 13%, bei DA 2 um 27,4% und die DA 3 sogar um 31,95%. Dabei sei erwähnt, dass die mittlere Spurhaltung und die Bearbeitungszeiten alle besser sind, als obige Durchschnittswerte.

Die Auswertung der SEA-Skala ergab einen durchschnittlichen Wert von 70 für das blinkende Icon und einen Wert von 61,25 für das sofort erscheinende Betrefffenster. Die subjektive Beanspruchung, durch den komplett erscheinenden Dialog, ist mit 57,5 als am

geringsten einzuschätzen. Die nicht-adaptive Verteilungsstrategie wurde im Durchschnitt mit 90 als eher „anstrengend empfunden“, wogegen die VPn die adaptive Verteilungsstrategie mit durchschnittlich 77,5 als „einigermaßen anstrengend“ empfanden.

Die Expertengruppe bewertete die Usability des adaptiven Systems mit einem SUS-Score von 91,25 besser, als das nicht-adaptive System (SUS-Score 86,5). Der SUS-Score des nicht-adaptiven Systems weicht nur wenig vom Wert der Gesamtstichprobe ab. Der SUS-Score des adaptiven Systems ist allerdings sehr viel höher/besser, als bei der Gesamtstichprobe. Dies bestätigt die Annahme, dass die Funktionsfähigkeit des Eye Trackers einen großen Anteil an der Beurteilung des Systems hat. Es sei jedoch angemerkt, dass diese Expertengruppe (von lediglich vier Personen) kein Rückschluss auf die Allgemeinheit erlaubt. Vielmehr zeigen die Ergebnisse den Einfluss des Eye Tracking auf die Ergebnisse sowie das Potential zur Verbesserung der Fahrleistung durch die Konzepte.

5.3 Aufmerksamkeitsbeeinflussung mit Warnung und Locking

Zur Vermeidung einer zu hohen mentalen Beanspruchung und den damit verbundenen Effekten (wie kognitive Gefangenheit oder veränderte Blickmuster) während der Interaktion mit Informationssystemen in ungünstigen Situationen ist es denkbar den Fahrer entweder durch eine Warnung oder gar durch das Abschalten von Funktionen wieder auf die Fahraufgabe zu lenken. Das gezielte Ausblenden von Infotainmentfunktionen abhängig von der Verkehrssituation fasst man dabei unter dem Begriff *Locking* zusammen.

Fahraufgabe und Sekundäraufgaben unterbrechen sich gegenseitig. Deshalb sprechen sich *Donmez et al.* nach ihren Studien zum Unterbrechungsmanagement für die Möglichkeit der Warnung oder des Lockings aus [65]. Sie weisen darauf hin, dass z.B. ein unkoordiniertes Sperren von Funktionen den Mental Workload erhöhen kann, wenn der Fahrer umso mehr versucht die Sekundäraufgabe fortzusetzen. Der Zeitpunkt einer Unterbrechung muss also sehr gut gewählt werden, auch um den Leistungsabfall beim Zurückkehren zur Zweitaufgabe klein zu halten.⁴ Auch *Iqbal* betont die Wichtigkeit des Unterbrechungszeitpunktes und zitiert Arbeiten, aus denen hervorgeht, dass Anwender bei unkoordinierter Unterbrechung beim Fortsetzen der Sekundäraufgabe mehr Fehler machen, mehr Zeit benötigen, größeren Ärger sowie Frustration erfahren [115]. Die Schwierigkeiten beim Sperren sprechen eher für das Warnen vor Sekundäraufgaben.

Eine Warnung kann visuell, auditiv oder taktil erfolgen. [66] zeigt, dass visuelle Warnungen jedoch nicht für alle Altersgruppen gleich gut geeignet sind. So wurden diese von mittelalten Probanden besser verarbeitet, als von älteren Teilnehmern. Auditive Hinweise stellten sich bei dem Experiment teilweise als störend dar, sollten jedoch für sehr wichtige Warnungen in Gefahrensituationen in Betracht gezogen werden. Taktile Warnungen bewirken die kleinste zusätzliche mentale Beanspruchung und sind daher ebenfalls geeignet, den Fahrer wieder auf das Fahrgeschehen zu lenken.

Um den Fahrer vor kognitiver Ablenkung oder Gefangenheit bei fahrerinitiierten Menüinteraktionen zu schützen, wurden die Warnkonzepte (*Soft-)Locking* und *Warnrahmen* getestet. Die Konzepte basieren auf einer adaptiven Ausblendung des Menüs im CID bzw. auf der Einblendung eines variablen Warnrahmens um das Menü und

⁴Diese zusätzliche Belastung wird in der Literatur als „switching cost“ bezeichnet (vgl. Kapitel 2.2.2).

sollen den Fahrer im Gefahrenfall wieder auf das Fahrgeschehen lenken. Der Gefahrenfall kann in der Realität ein zu hoher Workload Wert (physiologisch gemessen), ein plötzlich auftauchendes Hindernis (über Radar- oder Kamerasysteme erkannt) oder eine schwierige Fahrsituation (aus geografischen Informationen/Straßenkarte bekannt) sein. Das CID-Menü mit den entsprechenden Warnungen ist in den Abbildung 5.6 und 5.7 zu sehen. Die Anzeige des Warnrahmens sowie das Funktionslocking basieren auf den Fahrabnormalitätswerten des AWARE Systems aus Kapitel 4. Dies berechnet stetig den Abnormalitätswert, den die GUI weiterverarbeitet. Der Wertebereich für Abnormalität liegt dabei zwischen Null ($\hat{=}$ normales Fahren) und Eins ($\hat{=}$ sehr starke Abnormalität).

(Soft-)Locking Um den Nutzer auf einen drohenden Konzentrationsverlust hinzuweisen, ohne die erwähnten Nachteile des Lockings in Kauf nehmen zu müssen, wurde das Softlocking Konzept entworfen. Dieses basiert auf den zu jedem Zeitpunkt berech-

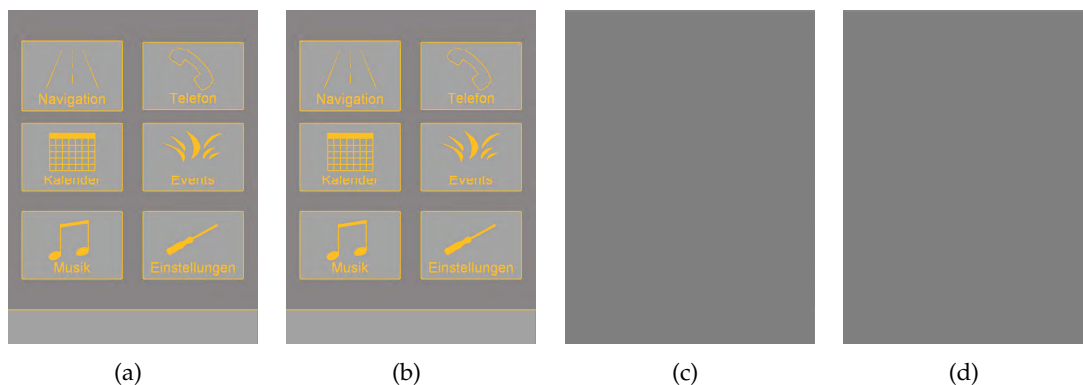


Abbildung 5.6: Softlockingfunktionalität: Abbildung des Abnormalitätswertes auf den Helligkeitswert des Mitteldisplays.

neten Fahrabnormalitätswerten, die dem Fahrer während tertiären (Bedien-)Aufgaben direkt als Visualisierung/Maß für seine Ablenkung vom Fahrgeschehen dienen. Dabei erfolgt keine gänzliche Unterbrechung der Bedienbarkeit des Interaktionssystems, sondern lediglich eine (leichte) Ausblendung. Die Umsetzung des hier verwendeten Ansatzes der Helligkeitsänderung entspricht eher einem „weichen“ Locking von Funktionen, da die Bedienfunktionalität jederzeit erhalten bleibt (der Fahrer wird nicht bevormundet). Zusätzlich signalisiert ihm die Helligkeitsänderung des Displays⁵, dass sein momentanes Fahrverhalten Abnormalität zeigt. Dazu erfolgt eine direkte Abbildung des Abnormalitätswertes ($0 \leq abnorm \leq 1$) auf die entsprechende Helligkeits- bzw. Deckkraftwerte.

Warnrahmen Die Umsetzung des Warnrahmens verfolgt einen ähnlichen Ansatz, jedoch ohne Abdunklung des gesamten Displays abhängig von der Ablenkung bzw. der Fahrabnormalität. Der Rand des Displays zeigt lediglich einen (kleinen) variablen Warnrahmen an. Dazu passen sich sowohl Farbe als auch Deckkraft des Warnrahmens an den

⁵Vgl. ausgegraute Schaltflächen bzw. aktuell nicht verfügbare Funktionen in graphischen Oberflächen von Desktop PC Software.

Fahrabnormalität <i>abnorm</i>	$0.1 < \textit{abnorm} < 0.2$	< 0.4	< 0.6	< 0.8	≥ 0.8
Deckkraft <i>d</i>	20%	40%	60%	80%	100%

Tabelle 5.1: Abbildung der Abnormalität auf Helligkeits- bzw. Deckkraftwerte.

Abnormalitätswert an, um bei geringer Ablenkung lediglich einen Hinweis zu geben bzw. bei großer Ablenkung eine Akutwarnung realisieren zu können.

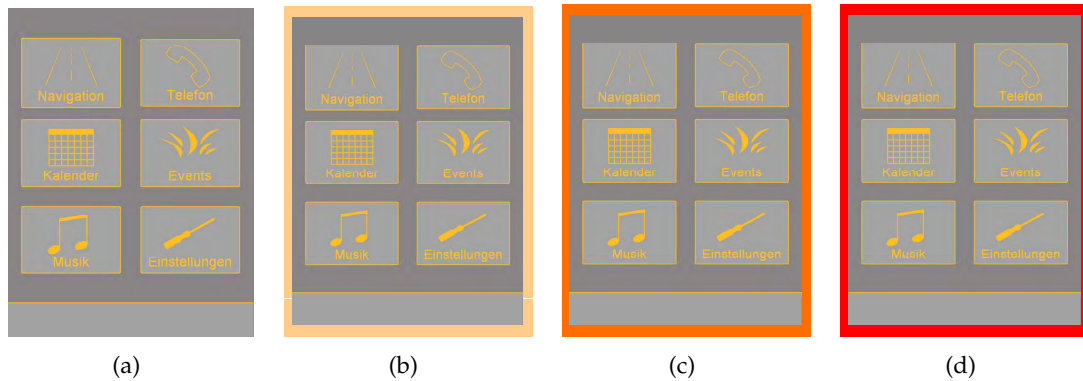


Abbildung 5.7: Abbildung des Fahrabnormalitätswertes auf den variablen Warnrahmen.

5.3.1 Versuchsergebnisse

Zur Bewertung dieser Warnkonzepte hinsichtlich ihrer Wirkung und Usability wurden die Warn- bzw. Hinweiskonzepte (vgl. Kapitel 5.3) in Laborfahrversuchen untersucht. An der Evaluierung der Warnfunktionalitäten nahmen 20 Probanden (14 männlich) im Durchschnittsalter von 28,2 Jahren teil. Die Teilnehmer bewerteten sich selbst mehrheitlich als erfahren im Straßenverkehr. Der Großteil der Probanden war technisch interessiert, offen für Innovationen und sicher im Umgang mit Technik. Diese Aussage bestätigte sich auch während der Versuchsfahrten, da es keinerlei Schwierigkeiten oder Unklarheiten in der Bedienung der Versuchsanordnungen gegeben hat. Auch mit Menüsystemen im Fahrzeug waren die Probanden bereits relativ gut vertraut. Aufgrund des jungen Durchschnittsalters, der durchschnittlichen Fahrerfahrung und der Technikrobustheit der VPn ging man davon aus, dass sich ein sinnvolles Konzept in einer Verbesserung der Fahrleistung und der mentalen Beanspruchung widerspiegelt. Die Aussage über die Fahrrentlastung seitens des Automobils erhielt im Voraus tendenziell Zustimmung, wobei sich dennoch Vorbehalte auftun: Die meisten Probanden sprachen sich strikt gegen Locking Ansätze aus. Trotzdem war in der Stichprobe der Wunsch nach Entlastung vorhanden, wenn es das Autofahren einfacher und sicherer gestaltet. Für diesen Versuch wurden oben beschriebene Systemverhalten anhand der Infotainmentmenüs der Telefon- und Navigationsfunktion untersucht. Dazu mussten die Probanden bei einer Kolonnenfahrt mit Fremdverkehr verschiedene Eingabeaufgaben auf dem Touchscreen durchführen:

- **Telefon:** Die Probanden wurden gebeten, zwei Namen in der langen und scrollbaren Telefonbuchliste zu suchen und durch Antippen einen Anrufaufbau zu initiieren.
- **Navigation:** Um einen erhöhten Workload zu induzieren und langfristige Interaktionen mit dem System zu gewährleisten, mussten die Probanden die längeren Städtenamen *Flensburg* und *Berchtesgaden* in die Zieleingabemaske des Navigationsmenüs eingeben.

Auch wenn dieses Verhalten im realen Straßenverkehr nicht erwünscht ist – da es Ablenkung stark fördert – war für diesen Versuch genau dieser Effekt erstrebenswert, da dies Fahrabnormalität provoziert und die Warnstrategien gut zur Geltung kommen.

Fahrleistung Die Auswertung der Fahrleistung erfolgte anhand der Spurabweichung. Abbildung 5.8 stellt jeweils die mittlere Spurabweichungen der Fahrten mit Warnrahmen und Softlocking gegenüber. Die Spurabweichung der Baseline-Erhebung ist links eingezeichnet und liegt mit 0,49m unter dem geforderten Wert von 1m [184]. Es zeigt

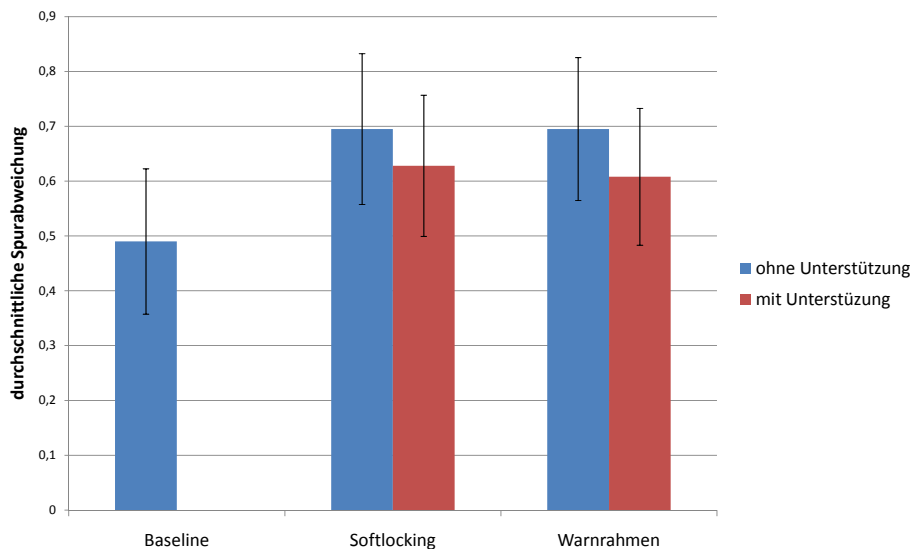


Abbildung 5.8: Mittlere Spurabweichung der Probanden ohne (blau) und mit aktiviertem Warnsystem (rot). Die linke Säule zeigt die mittlere Spurabweichung der Baseline-Erhebung.

sich, dass eine Tendenz zur Verbesserung der Fahrleistung durch die Warnkonzepte besteht, die durchschnittliche Verbesserung der Spurabweichung ist allerdings nicht statistisch signifikant. Es sei jedoch auf den großen Streubereich der Daten hingewiesen, der im Wesentlichen von einem Ausreißer (durch die älteste VP) verursacht wurde. Vergleicht man die Daten ohne diesen Extremwert, so ergibt sich eine statistisch signifikante Verbesserung um 8,7 cm bzw. 12,5% ($p < 0,006$).

Mentale Beanspruchung (PDT und SEA) Die mentale Beanspruchung der Probanden wurde objektiv über die Peripheral Detection Task (PDT, vgl. Kapitel 2.4.1.2) sowie subjektiv über die SEA-Skala (subjektiv erlebte Beanspruchung, vgl. Kapitel 2.4.1.3) gemessen. Die Warnungen hatten hinsichtlich der Reaktionszeiten auf die PDT-Punkte jeweils eine

tendenzielle Verbesserung zur Folge (nicht statistisch signifikant, vgl. Abbildung 5.9(a)). Es ist erkennbar, dass die Reaktionszeit generell die mentale Beanspruchung abbildet, der Einfluss des Warnsystems sich darin jedoch nicht äußerte.

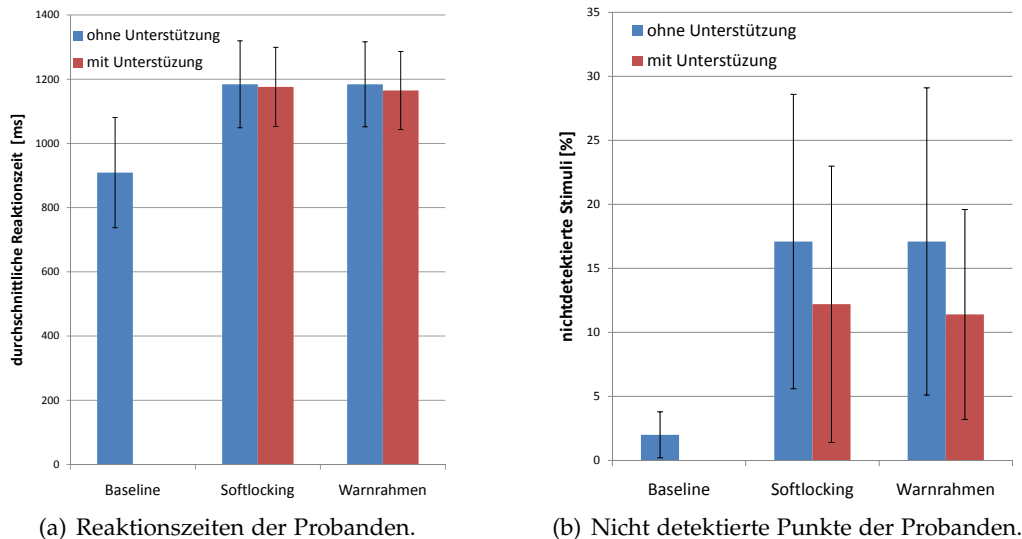


Abbildung 5.9: Auswertung der mittels PDT objektiv gemessenen Beanspruchung der Probanden.

Abbildung 5.9(b) stellt das Verhältnis nicht detektierter Punkte zur Gesamtzahl der Stimuli dar. Im Gegensatz zu den Reaktionszeiten zeigt sich ein positiver Einfluss des adaptiven Warnsystems: Mit Warnung wurden 5,7% weniger Punkte übersehen. Laut den t-Tests sind diese Ergebnisse zwar nicht statistisch signifikant, aber es ist ein eindeutig positiver Trend zu erkennen, der sich mit den Ergebnissen der Fahrleistungen deckt.

Die subjektiv erlebte mentale Beanspruchung der Probanden (SEA-Skala) bei den verschiedenen Versuchsfahrten stimmt gut mit der objektiv gemessenen Beanspruchung (nicht detektierte Punkte) überein (vgl. Abbildung 5.10(a)). Es zeigt sich eine deutliche Tendenz zur Verbesserung um durchschnittlich 6,5% (Softlocking: 5,8%; Warnrahmen: 7,2%) durch das System, die jedoch nicht statistisch signifikant ausfällt ($p = 0,058$).

Die kleinere Verbesserung der subjektiv erlebten mentalen Beanspruchung beim Warnsystem könnte in einem Akzeptanzproblem mancher Probanden begründet sein. Sie gaben – trotz objektiv besserer Leistungsdaten bei Fahr-, PDT- und Sekundäraufgabe – eine höhere subjektiv erlebte mentale Beanspruchung an. Als Gründe dafür konnten die zusätzlich zu verarbeitenden Stimuli (Deckkraftänderung bzw. Warnrahmen) sowie die als unangenehm empfundene Bevormundung durch die Warnungen identifiziert werden.

Sekundäraufgabendaten Abbildung 5.10(b) zeigt die durchschnittliche Anzahl eingegebener Fahrtziele und die durchschnittliche Anzahl der Fehler (falsch ausgewählte Buchstaben) für die Fahrten mit und ohne Unterstützung durch das Warnsystem. Die Anzahl der eingegebenen Fahrtziele hat sich mit Warnsystem durchschnittlich um 7,2% (Softlocking: -5,2%; Warnrahmen: -9,3%) verringert (nicht statistisch signifikant). Nach dem Subsidiary Task Paradigma würde die schlechtere Leistung bei der Sekundärauf-

5.3. Aufmerksamkeitsbeeinflussung mit Warnung und Locking

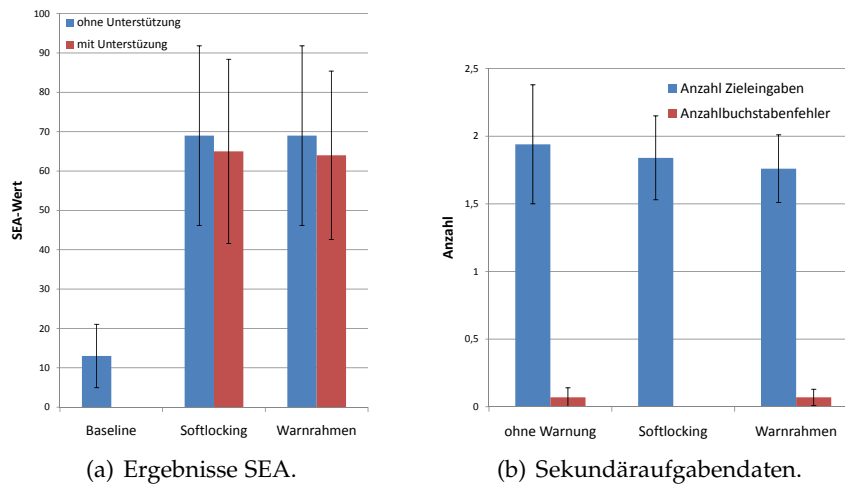


Abbildung 5.10: (a) Subjektiv erlebte mentale Beanspruchung der Probanden. (b) Anzahl eingegebener Fahrtziele und Buchstabenfehler der Probanden.

gabe auf eine Verringerung der zur Verfügung stehenden Ressourcen und damit auf eine gestiegene mentale Beanspruchung schließen lassen. Es muss in Erinnerung gerufen werden, dass die Probanden durch das Warnsystem gezwungen waren, die Menüeingabe öfter und länger zu vernachlässigen. Das bessere Abschneiden des Softlocking Ansatzes ist – laut Aussagen der Probanden – darauf zurückzuführen, dass sie bei leichter Fahrabnormalität und nur leichtem Abdunkeln des Displays zunächst weiter bedient haben, bevor sie im weiteren Verlauf des Abdunkelns die Eingabe unterbrochen haben. Dies spiegelt sich ebenfalls in der etwas schlechteren Fahrleistung wieder (vgl. Abbildung 5.8).

Die Daten lassen sich dahingehend interpretieren, dass das System die für die Menüeingabe aufgewendeten Ressourcen gezielt verringerte, um einer Vernachlässigung der Fahraufgabe vorzubeugen. Hier sei auf die Aktivierungsschwelle⁶ hingewiesen, die einerseits die Fahrsicherheit bestimmt (z.B. je früher die Warnung, desto besser) und andererseits Einfluss auf die Akzeptanz hat (z.B. je öfter, desto schlechter).

Subjektive Ergebnisse Die subjektiven Ergebnisse zur Nutzerakzeptanz der Warnstrategien wurden sowohl über Standardfragebögen (SUS und AttrakDiff) als auch über allgemeine Thesen (über Likert-Skalen) erhoben. Abbildung 5.11 zeigt die Zusammenfassung der allgemeinen Thesen zum (Soft-)Locking Konzept. Grundsätzlich bewerteten die Versuchsteilnehmer den Hinweis auf Fahrabnormalität bzw. drohende zu hohe mentale Last als sinnvoll (beide Ansätze). Bei der Umsetzung mittels der Deckkraftänderung waren die Probanden eher uneinig, was damit zusammenhängt, dass nicht alle diese Art der Visualisierung als intuitiv empfanden. Beim SUS (vgl. Kapitel 2.4.1.5) erreichte der Visualisierungsansatz 59,83 Punkte (von 100 möglichen), wobei sich die VPn nicht einig in der Bewertung waren (Standardabweichung 20,25). Da der Ansatz lediglich in

⁶Zum Problem der Wahl der Aktivierungsschwelle von Funktionen (z.B. Warnungen), sowie der Diskrepanz zwischen hoher Sicherheit und geringer Fehlalarmraten sei an dieser Stelle auf *Hoedemaekers* Bericht zu adaptiven Systemen im Kraftfahrzeug [109] verwiesen.

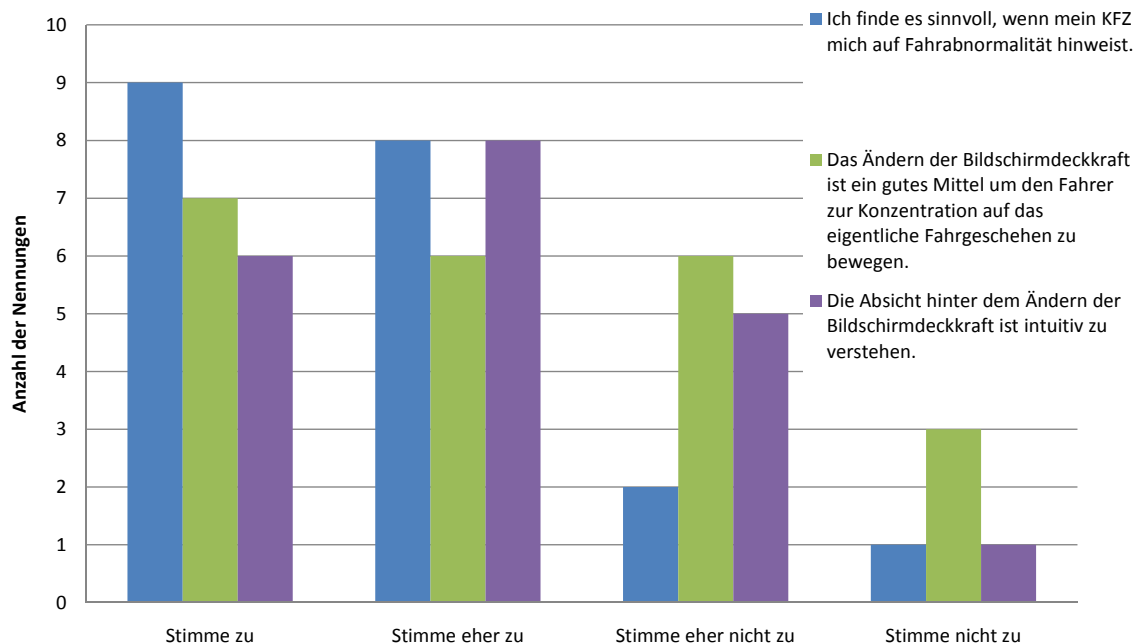


Abbildung 5.11: (Soft-)Locking, allgemeine Bewertung des Systems hinsichtlich wie sinnvoll der Hinweis auf Fahrabnormalität ist (blau), ob die Änderung der Deckkraft ein geeignetes Mittel ist um sich wieder auf das Fahren zu konzentrieren (grün) sowie ob die Änderung der Deckkraft intuitiv ist (lila).

einer prototypischen Umsetzung präsentiert wurde, kann davon ausgegangen werden, dass die Bewertungen in einem realistischeren Kontext deutlich bessere Werte erreichen. Das unterstreicht auch das Ergebnis des AttrakDiff Fragebogens (vgl. Kapitel 2.4.1.4): „Die Benutzeroberfläche des Prototypen wurde als neutral eingestuft. Diese Zuordnung der pragmatischen Qualität ist eindeutig. Der Nutzer wird durch den Prototypen zwar unterstützt, allerdings erreicht er lediglich mittlere Werte. Es besteht Verbesserungspotenzial hinsichtlich der Bedienbarkeit. Die hedonische Qualität erreichte eine eindeutige Charakterzuordnung. Der Nutzer ist durch das Produkt angeregt, allerdings erreicht die hedonische Qualität lediglich mittlere Werte. Es besteht somit noch Verbesserungspotenzial hinsichtlich hedonischer Aspekte.“

Abbildung 5.13 zeigt die Zusammenfassung der allgemeinen Thesen zum Warnrahmen. Die Umsetzung der Warnung mittels eines Warnrahmens bewerteten die Probanden etwas besser. Sie empfanden diesen Ansatz als intuitiver und besser geeignet, um die Aufmerksamkeit wieder auf das Fahrgeschehen zu lenken. Die Auswertung ergab (im Mittel) einen SUS Wert von 74,4. Die Bedienbarkeit erreicht also – trotz der prototypischen Integration – bereits zufriedenstellende Werte. Dies zeigt sich weiterhin im Ergebnis des AttrakDiff Fragebogens: „Die Benutzeroberfläche des Konzepts wird als „neutral“ eingestuft. Diese Zuordnung ist für die pragmatische Qualität nicht eindeutig, da das Konfidenzintervall über den Charakterbereich hinausgeht. Der Nutzer wird durch das Produkt unterstützt. Es besteht Verbesserungspotenzial hinsichtlich der Bedienbarkeit.“

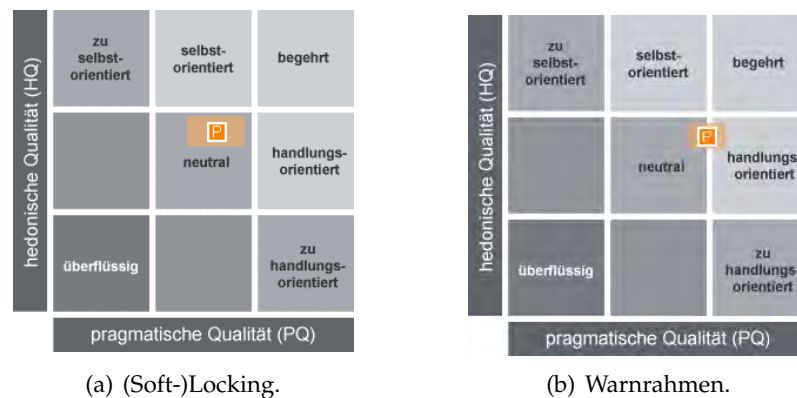


Abbildung 5.12: Vergleich der Portfoliodarstellungen der getesteten Warnkonzepte *Soft-locking* und *Warnrahmen*.

Für die hedonische Qualität trifft die Charakterzuordnung eindeutig zu. Der Nutzer wird durch das Produkt angeregt, die hedonische Qualität erreicht mittlere Werte. Es besteht Verbesserungspotential hinsichtlich hedonischer Aspekte.“ Ein Verbesserungspotential in der Bedienbarkeit ist aufgrund des reduzierten Interfaces gegeben. Trotzdem ist es bemerkenswert, dass sich die Nutzer durch das System unterstützt fühlten. Das kleine Konfidenzrechteck entspricht einer großen Einigkeit der Benutzer in ihren Beurteilungen. Persönliche Einstellungsmöglichkeiten (z.B. für die Art der Warnung, etc.) würden in einer seriennahen Umsetzung zur Verbesserung der Bewertung des Systems beitragen.

5.3.2 Fazit

Alle aufgezeichneten Daten zeigen positive Tendenzen. Die Ergebnisse aus Kapitel 5.3.1 zeigen, dass es durch ein einfaches Warnsystem gelingt, die Probanden vor zu hoher kognitiver Ablenkung bei Menüeingaben zu schützen, was sich in einer besseren Fahrleistung zeigt. Weiterhin ist festzuhalten, dass der Warnrahmen geringfügig bessere Ergebnisse (v.a. geringere mentale Anstrengung und bessere Fahrperformance) erreicht.

Mehrere Probanden gaben im Rahmen der Nachbefragung an, dass eine Mischung aus beiden Konzepten die Akzeptanz weiterhin steigern könnte. Sie schlugen vor, dass in „weniger kritischen Situationen“ über das Ausblenden des Monitors ein Hinweis gegeben werden könnte. In „kritischen Situationen“ erfolgt dann eine Akutwarnung über den Warnrahmen. Weiterhin sehen die VPn den Vorteil, dass der Beifahrer diese Warnungen weniger intensiv wahrnimmt. Dies steigert die Akzeptanz weiterhin, da der Fahrer den Hinweis, sein Fahrverhalten anzupassen, deutlich subtiler erhält.

5.4 Zeitliches Scheduling von fahrzeuginitiierten Dialogen

Die Informationspräsentation und die damit verbundene kognitive Ablenkung kann zu einem ungünstigen Zeitpunkt zur Überforderung der Verarbeitungskapazität des Fahrers führen. Untersuchungen dazu haben gezeigt, dass Probanden ihre Aufmerksamkeit

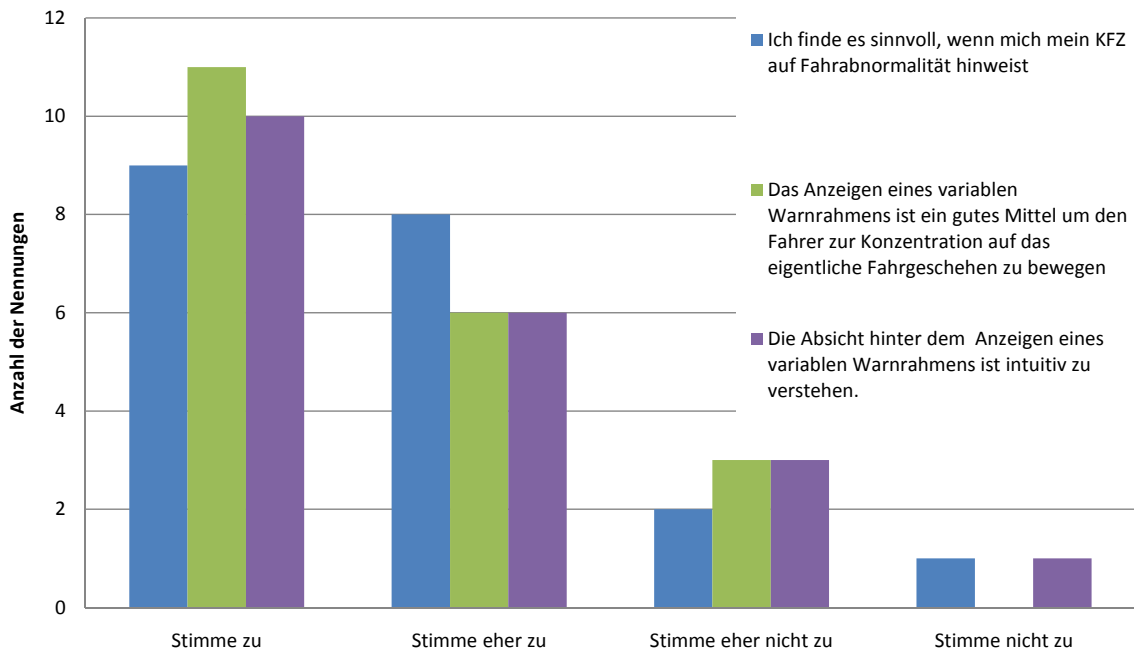


Abbildung 5.13: Allgemeine Bewertung des Systems hinsichtlich wie sinnvoll der Hinweis auf Fahrabnormalität ist (blau), ob die Anzeige eines variablen Warnrahmens ein geeignetes Mittel ist um sich wieder auf das Fahren zu konzentrieren (grün) sowie ob die Anzeige eines Warnrahmens intuitiv ist (lila).

immer in Richtung Fahrzeuginnenraum richten, sobald eine (akustische) Warnung erfolgt [166]. Ein Ansatz zur Reduktion der Ablenkung ist es also, Informationen zu Zeitpunkten, in denen der Fahrer mental sehr beansprucht ist, zeitlich zu verzögern oder gar zu blockieren. Dieses Vorgehen wurde bereits 1989 von *Smiley* [219] und später in diversen anderen Arbeiten vorgeschlagen [65, 152, 242, 243]. So könnte eine empfangene SMS erst angezeigt werden, wenn die mentale Beanspruchung des Fahrers dies zulässt.

Hierzu müssen sämtliche Meldungen einer Klasse zugeordnet sein, die einer gewissen Priorität entspricht. Informationen mit niedriger Priorität (z.B. SMS) würden dann bei hohem Workload (temporär) blockiert, während das System Informationen hoher Priorität (z.B. Abbiegehinweis) sofort ausgibt. Ebenso können Informationen, die gleichzeitig präsentiert würden, durch die Prioritäteneinteilung hintereinander dargestellt werden⁷. Obwohl der Mensch nach Wickens MRT in der Lage ist, gewisse Informationen (verschiedener Ressourcenbeanspruchung) parallel zu verarbeiten, kann dieses Vorgehen negative Effekte von gleichzeitigen auditiven und taktilen Informationen (z.B. Navigationsansage und Vibrieren des Lenkrads aufgrund einer Spurverlassenswarnung) verhindern [109]. Neben dem Scheduling von Informationen gibt *Verwey* die Empfehlung, Informationen in verschiedenen – jedoch geeigneten – Modalitäten auszugeben [243]. Dem Ändern des

⁷Dabei sollte wiederum ein gewisser Abstand eingehalten werden. [109] gibt z.B. 3 Sekunden als optimalen Abstand zwischen zwei Audioinformationen an.

Formats einer Information stehen die Autoren allerdings sehr kritisch gegenüber, da dies eher zu einer zusätzlichen mentalen Beanspruchung führt [157].

Ausgehend von diesen Erkenntnissen sowie dem in Kapitel 4 vorgestellten System AWARE wurde ein Informations-Scheduling entworfen und getestet, dass Informationen zu Zeitpunkten hoher mentaler Beanspruchung/ Abnormalität blockiert, bis der Mental Workload wieder sinkt. Ohne Informations-Scheduling konnten die Informationen jederzeit auf den Fahrer einwirken, also auch zu Zeitpunkten hoher mentaler Beanspruchung. Die Versuchssteuerung schickt dazu in zeitlich vorgegebenen Abständen Ausgaben auf das CID. Zur Simulation von Informationen, die eine gewisse kognitive Verarbeitung

Satz	richtige Antwort
Der Apfel fällt vom Baum.	ja
Das Wetter ist schwer.	nein

Tabelle 5.2: Beispielsätze für das Scheduling-Testverfahren

und eine aktive Antwort benötigen,⁸ kam das in [66] vorgestellte Verfahren zum Einsatz, dass dem Probanden ein Satz entweder auditiv oder visuell präsentiert. Dieser muss den Satz – abhängig davon, ob der Satz semantisch richtig oder falsch ist – mit „ja“ oder „nein“ quittieren (durch Bestätigen der Schaltfläche auf dem Touchscreen). Die Antworten der Probanden werden aufgezeichnet und ausgewertet (Zeitpunkt der Fragestellung, Antwortzeiten, richtig oder falsch geantwortet). Eine falsche Antwort lässt auf eine beeinträchtigte Perzeption oder Verarbeitung rückschließen.

Beim Versuch wurde das Informations-Scheduling jeweils mit auditiven und visuellen Informationen getestet, wobei mit eingeschaltetem System die Informationen blockiert wurden, wenn der Fahrabnormalitätswert (vgl. Kapitel 4) über einer – im Rahmen von Vorversuchen gewonnenen – Schwelle von $abnorm \geq 0.35$ lag. Dabei verzögerte das System alle Ausgaben unabhängig ihrer Priorisierung. Für eine detaillierte Betrachtung zur unterschiedlichen Priorisierung von Systemmeldungen sei auf [166] verwiesen. Zur Entscheidung, ob eine Meldung ausgegeben wird oder nicht, verwendet der Autor Fahrzeugdaten, wie z.B. Blinkerstatus, Bremsdruck und Geschwindigkeit. Insgesamt mussten die Probanden 17 Sätze pro Fahrt verarbeiten und beantworten. Zur Bewertung positiver oder negativer Effekte dieses Systems erfolgte ein Vergleich der Fahrten mit und ohne Informations-Scheduling, jeweils untereinander und zusätzlich mit einer Baseline Fahrt.

5.4.1 Versuchsergebnisse

Die Bewertung der vorgestellten Schedulingkonzepte hinsichtlich ihrer Effektivität und Usability erfolgte im Rahmen der Laborfahrversuche aus Kapitel 5.3. An der Evaluierung nahmen ebenfalls die 20 in Kapitel 5.3 näher beschriebenen Probanden teil. Der Versuch untersuchte das beschriebene Systemverhalten anhand der Telefon- und Navigationsfunktion. Dazu mussten die Probanden während einer Kolonnenfahrt mit Fremdverkehr verschiedene Navigationsziele und Telefonnummern über den Touchscreen eingeben.

⁸Stimuli, die eine Antwort benötigen, führen zu höherem Mental Workload (vgl. Kapitel 2.2.1). Die Verarbeitung solcher Ausgaben ist ähnlich der Verarbeitung einer Navigationsansage oder eines Telefongesprächs.

Fahrleistung Die Fahrleistung wurde anhand der Spurabweichung ausgewertet. Die Werte der Baseline-Erhebung stimmen mit den in Kapitel 5.3 präsentierten Werten überein. Abbildung 5.14 zeigt die mittleren Spurabweichungen der Fahrten (Baseline, akustische

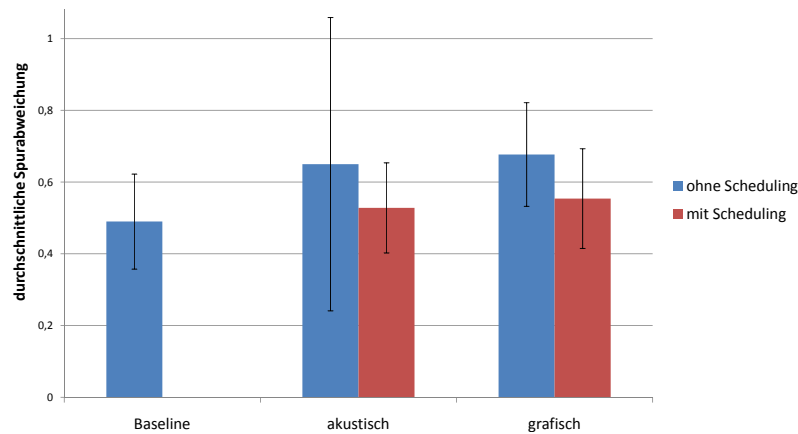


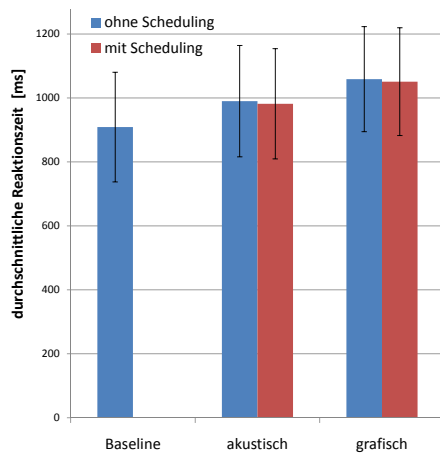
Abbildung 5.14: Mittlere Spurabweichung der Probanden ohne (blau) und mit aktiviertem Scheduling (rot). Die linke Säule zeigt die mittlere Spurabweichung der Baseline-Erhebung.

und visuelle Systemausgaben). Es zeigt sich eine Tendenz zur Verbesserung der Fahrleistung durch das Scheduling. Die durchschnittliche Verbesserung der Spurabweichung um 12,2 cm beim Scheduling von akustischen Ausgaben ist nicht statistisch signifikant.

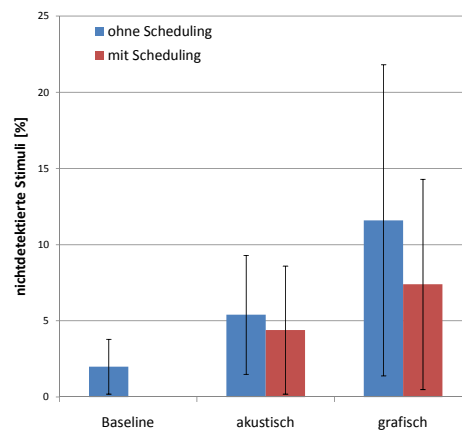
Es sei jedoch auf den großen Streubereich der Daten hingewiesen, den im Wesentlichen abermals die älteste VP verursachte. Vergleicht man die Daten ohne diesen Extremwert, so ergibt sich eine statistisch signifikante Verbesserung um durchschnittlich 5,2 cm ($p < 0,004$). Das Scheduling graphischer Ausgaben führte zu einer durchschnittlichen Verbesserung der Spurabweichung von 12,3 cm (statistisch signifikant, $p < 0,001$). Eine adaptive Unterstützung des Fahrers äußert sich also in einer Verbesserung der Fahrleistung um 8% (akustisch) bzw. 18,2% (graphisch). Interessant ist an dieser Stelle weiterhin, dass sich ein Einfluss auf die Fahrleistung bei der zeitlichen Verzögerung von akustischen Systemmeldungen ergab, da laut Wickens MRT dies kaum der Fall sein dürfte. Dieser gegenteilige Effekt ist in der Literatur bereits bekannt.

Mentale Beanspruchung (PDT und SEA) Die mentale Beanspruchung der Probanden wurde über die PDT (objektiv) sowie über die SEA-Skala (subjektiv) erfasst. Abbildung 5.15(a) zeigt die Reaktionszeiten der PDT, wobei jeweils lediglich eine kleine Tendenz der Verbesserung (nicht statistisch signifikant) zu erkennen ist. Abbildung 5.15(b) zeigt das Verhältnis nicht detektierter Punkte zur Gesamtzahl der Stimuli. Im Gegensatz zu den Reaktionszeiten ist hier ein positiver Einfluss zu identifizieren. Mit Unterstützung durch das zeitliche Scheduling wurden 1% (akustisch) bzw. 3,7% (graphisch) weniger Punkte verpasst. Auch hier zeigt sich ein eindeutig positiver Trend (nicht statistisch signifikant). Die subjektiv erlebte mentale Beanspruchung (vgl. Abbildung 5.16(a)) der Probanden bei den verschiedenen Versuchsfahrten korreliert gut mit der objektiv gemessenen Beanspruchung (nicht detektierte Punkte). Auch hier zeigt sich im Falle der Audio-Kondition kaum ein Unterschied, während bei den visuellen Aufgabentypen eine

5.4. Zeitliches Scheduling von Dialogen



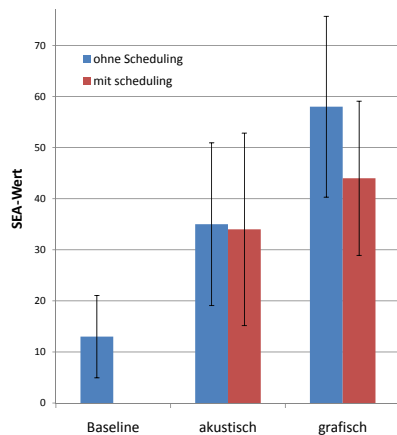
(a) Reaktionszeiten der Probanden.



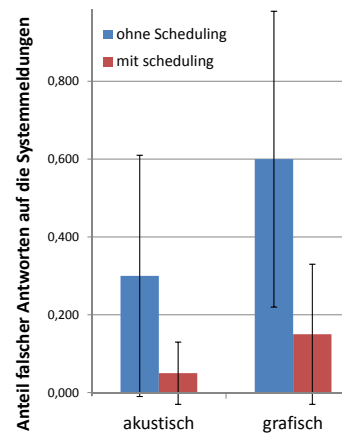
(b) Nicht detektierte Punkte der Probanden.

Abbildung 5.15: Auswertung der mittels PDT objektiv gemessenen Beanspruchung der Probanden ohne (blau) und mit aktiviertem Scheduling (rot).

deutliche Tendenz zur Verbesserung durch das System zu verzeichnen ist. Die Verbesserung um 22,8% für das Scheduling visueller Informationen ist statistisch hoch signifikant ($p < 0,001$).



(a) Subjektiv erlebte mentale Beanspruchung der Probanden.



(b) Anteil falscher Antworten auf die Systemmeldungen der Probanden.

Abbildung 5.16: Ergebnis SEA sowie Anteil falscher Antworten auf Systemmeldungen ohne (blau) und mit aktiviertem Scheduling (rot).

Es zeigt sich, dass die Reaktionszeit die allgemeine Beanspruchung einer Aufgabe (vgl. auch Ergebnisse zur mentalen Beanspruchung aus Kapitel 5.3) und die Anzahl nicht detektierter Punkte Workload-Spitzen (mentale Überforderung) abbilden. Damit lassen sich die Daten dahingehend interpretieren, dass die Konzepte die grundsätzliche mentale Beanspruchung von Aufgaben (z.B. Informationen während dem Autofahren verarbeiten) nicht verringern, jedoch den Fahrer vor mentaler Überlastung schützen können.

Sekundäraufgabendaten Abbildung 5.16(b) zeigt die durchschnittliche Anzahl falscher Antworten der VPn (jeweils mit/ohne Informations-Scheduling). Die Baseline Anteile (ohne Autofahrt) falscher Antworten betragen für akustische Meldungen 0,05 und für visuelle Meldungen 0,15. Für beide Kategorien zeigt sich eine deutliche Verbesserung durch das Scheduling. Die Probanden erreichten exakt die jeweiligen Baseline Werte: für akustische Meldungen konnte eine Verbesserung von 83,3% (nicht statistisch signifikant) und für graphische eine Verbesserung von 75% (statistisch signifikant, $p < 0,016$) erreicht werden. Interpretiert man diese Daten im Sinne des Subsidiary Task Paradigmas, so hatten die VPn mehr Ressourcen für Sekundäraufgaben verfügbar und Situationen mentaler Überforderung (keine Verarbeitung der Informationen möglich) wurden verringert.

Subjektive Ergebnisse Hinsichtlich Attraktivität, Akzeptanz und Bedienbarkeit erreichte das System zufriedenstellende Ergebnisse. Die Auswertung ergab einen mittleren SUS Score von 74,2, womit die Bedienbarkeit des Systems zufriedenstellend ist. Die Nutzer fühlten sich vom Prototypen unterstützt: Das System wurde nach dem AttrakDiff Fragebogen als „neutral“ eingestuft. Die pragmatische Qualität des Konzeptes wurde nicht eindeutig bewertet, das Konfidenzintervall geht über den Charakterbereich hinaus. Das Produkt unterstützt den Nutzer und regt ihn an, allerdings besteht Verbesserungspotential hinsichtlich „Bedienbarkeit“ und hedonischer Qualität. Das kleine Konfidenzrechteck im Portfolio entspricht einer großen Übereinstimmung der Beurteilungen.

5.4.2 Fazit

Die Auswertung der Versuchsdaten zeigt durchweg positive Ergebnisse, die den Erwartungen entsprachen. Es zeigte sich, dass eine adaptiv angepasste Informationspräsentation zu einer besseren Fahrleistung sowie einer besseren Informationsverarbeitung führt. Damit wurde das vielversprechende Potential adaptiver Fahrer-Fahrzeug Schnittstellen bestätigt.

5.5 Fahreraufmerksamkeitslenkung durch graphische und akustische Animationen

Aufgrund der eingangs gewonnenen Erkenntnisse ergibt sich das Ziel, die Aufmerksamkeit des Benutzers bei drohendem Aufmerksamkeitsverlust (v.a. in Gefahrensituationen) schnell und sicher dorthin zu lenken, wo sie am dringendsten benötigt wird. Deshalb sollen im Folgenden effektive Ansätze vorgestellt und bewertet werden, um die Aufmerksamkeit des Fahrers sowohl mit optischen als auch akustischen Mitteln gezielt zu lenken, um eine schnellere und bessere Reaktion zu ermöglichen.

5.5.1 Theoretische Grundlagen

Aufbauend auf den im Folgenden hergeleiteten Grundlagen zu den Themen Aufmerksamkeit und Aufmerksamkeitslenkung stellt Kapitel 5.5 im weiteren Verlauf Konzepte zur Fahreraufmerksamkeitslenkung durch graphische und akustische Animationen vor.

5.5.1.1 Aufmerksamkeit

Einen umfassenden Überblick zum Thema Aufmerksamkeit geben Müller und Krummenacher [163]. Sie definieren als eine wesentliche Funktion der Aufmerksamkeit die Selektion von bestimmten Inhalten mit dem Ziel, bestimmte Informationen dem Bewusstsein bzw. der Steuerung von Denken und Handeln zugänglich zu machen. Als Beispiel führen sie dabei das „Cocktailparty-Phänomen“ auf: Auch wenn in der Umgebung viele einzelne Gespräche geführt werden, ist es möglich, einem bestimmten zu folgen. Hört man währenddessen den eigenen Namen, kann man die Aufmerksamkeit verlagern und für eine Weile einem anderen Gespräch folgen. Auch wenn der Mensch durchaus alle Reize aufnimmt, erfolgt gleichzeitig eine Filterung nach Relevanz, was zu einer bewussten Aufnahme lediglich eines geringen Teils verfügbarer Informationen führt. Diese Selektion existiert sowohl für die auditive als auch für die visuelle Aufmerksamkeit. Da die visuelle Aufmerksamkeit sehr eng mit der aktuellen Blickrichtung korreliert ist [179], dient im Folgenden zur Bestimmung des Aufmerksamkeitsfokusses das in Kapitel 3.2 vorgestellte Eye Tracking System.

5.5.1.2 Graphische Möglichkeiten der Aufmerksamkeitslenkung

Um aufzuzeigen, inwieweit es möglich ist, die Aufmerksamkeit eines Menschen gezielt zu lenken, fasst dieser Abschnitt die relevanten Grundlagen der Wahrnehmung zusammen.

Das Gesichtsfeld besteht aus dem zentralen und dem peripheren Blickfeld und erstreckt sich über ca. 180° in horizontaler Richtung. Alle Information im zentralen Blickfeld werden scharf abgebildet und bewusst wahrgenommen. Damit ist es möglich, Bewegungen wahrzunehmen und auf potentielle Gefahren schnell zu reagieren, ohne stetig das komplette Gesichtsfeld detailliert zu verarbeiten. Farbe und Sehschärfe sind bei der Detektion von Bewegung von geringerer Bedeutung und im peripheren Sichtfeld nur bedingt verfügbar [131].

Auf der Unterteilung des Blickfeldes baut der urzeitlich verankerte *Orientierungsreflex* auf. Um sicherzustellen, dass von einer (unerwarteten) Bewegung im peripheren Blickfeld keine Gefahr ausgeht, blicken wir reflexartig in die Richtung der Bewegung [92]. Was sich dabei bewegt ist primär irrelevant, und wird bei Zuwendung zum bewegten Objekt ohnehin festgestellt. Wichtig ist hier eine möglichst schnelle Reaktion.

Weitere Hinweise auf den Unterschied zwischen peripherem und zentralem Blickfeld finden sich bei sog. *peripheren Cues*. Müller und Rabbitt untersuchten diesen Zusammen-

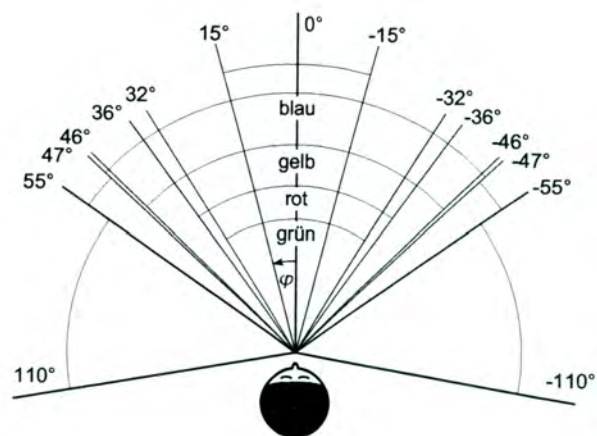


Abbildung 5.17: Das horizontale Gesichtsfeld des Menschen erstreckt sich von -55° bis +55°. Durch Drehung des Kopfes lässt sich das Blickfeld zusätzlich ausdehnen. Entnommen aus [208].

hang [167]: Probanden sollten dabei die Position peripher angezeigter Punkte möglichst schnell angeben. Der Hinweis auf eine mögliche Position des nächsten Punktes erfolgte dabei entweder durch einen Pfeil im zentralen Blickfeld (zentraler Cue) oder durch einen Blitz im peripheren Sichtfeld (peripherer Cue, siehe Abbildung 5.18(a)), wobei der angezeigte Ort nicht zwangsläufig mit dem Ort des folgenden Punktes (ungültiger Cue) korrelierte. Sie fanden heraus, dass auf zentrale Cues eine bewusste Verlagerung der visuellen Aufmerksamkeit erfolgt. Bei peripheren Cues dagegen läuft dieser Vorgang unbewußt und folglich auch schneller ab, was dem Orientierungsreflex entspricht.

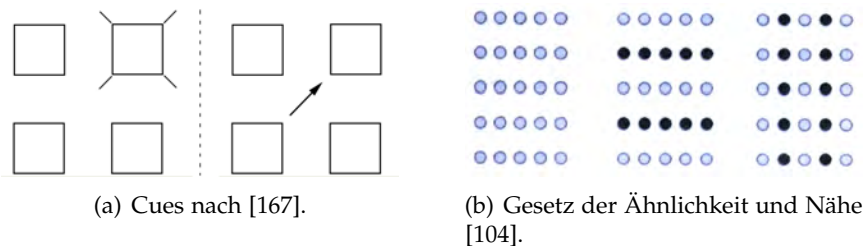


Abbildung 5.18: (a) In [167] verwendete Cues, links: peripherer Cue durch blinken; rechts: zentraler Cue durch Pfeil. (b) Das Gesetz der Ähnlichkeit (links) wirkt stärker als das Gesetz der Nähe (rechts). In der Mitte verstärken sie sich gegenseitig. Entnommen aus [104].

Bei der Bewegung unserer Augen von einem Punkt zu einem anderen kann zwischen *attentional saccades* und *attentional pursuit* unterschieden werden (vgl. Kapitel 3.1.3), also danach ob der Blick von einem Objekt zum nächsten springt oder einem bewegten Objekt folgt. *Horowitz* stellte dabei fest, dass sich die Aufmerksamkeit deutlich schneller verlagert, wenn sie sich mit einem bewegten Objekt mitbewegen kann. Dabei kann man durchaus eine Verbesserung um den Faktor 2 erreichen [111]. Dies lässt sich darauf zurückführen, dass bei „attentional saccades“ erst ein neuer Ankerpunkt zum Fixieren und Fokussieren selektiert werden muss. Beim „attentional pursuit“ hingegen kann der Aufmerksamkeitsfokus mit dem Objekt mitwandern.

Aus diesen Ergebnissen lässt sich folgern, dass ein sich direkt zum Ziel bewegendes Objekt besser geeignet ist, als ein peripherer Cue, um die Aufmerksamkeit gezielt auf einen Punkt zu lenken. Damit reduziert sich die zur Verlagerung der Aufmerksamkeit benötigte Zeit und ein nahe am zentralen Blickfeld auftauchendes Objekt ist weitaus schwerer zu übersehen, als ein blinkendes Objekt (weit) im peripheren Sichtfeld. Ein solches Objekt ist dann allerdings wieder als zentraler Cue zu verstehen, bei dem die Reaktion nicht automatisch sondern bewusst erfolgt. Da das Geschehen auf der Straße immer Vorrang vor auf einem Monitor angezeigten Informationen hat, kann man im Falle eines zentralen Cue verhindern, dass er die Aufmerksamkeit auf den Monitor lenkt, obwohl sie dringender zur Erfüllung der primären Fahraufgabe nötig ist.

Ein zentraler Aspekt der visuellen Wahrnehmung des Menschen ist es, dass sie nicht auf einzelnen Bildpunkten basiert, wie sie ein Computer abspeichert, sondern auf Formen. Dies zeigt sich beispielsweise in diversen *optischen Täuschungen*, bei denen wir Muster anders wahrnehmen als sie objektiv sind (vgl. Abbildung 5.19(a)). Aus der Wahrnehmung lassen sich *Gestaltgesetze* ableiten, denen man spezielle Wirkung zuschreiben kann. Ein Beispiel dafür ist das *Gesetz der Nähe*. Danach werden Objekte, die sich in direkter Nähe

befinden, automatisch als zusammengehörig wahrgenommen (vgl. Abbildung 5.18(b)). Gestaltgesetze können sich gegenseitig verstärken oder abschwächen. Stärker als das Gesetz der Nähe ist beispielsweise das *Gesetz der Ähnlichkeit*. Danach nimmt man Objekte mit gleicher Farbe/Form als Einheit wahr. Verwendung findet dies unter anderem in Diagrammen, um den Zusammenhang zwischen mehreren Datenreihen auszudrücken. Ist eine Datenreihe durch rote und die andere durch blaue Punkte repräsentiert, erkennt man sofort deren Zusammengehörigkeit. Dies setzt das Gesetz der Nähe außer Kraft.

Gegensätzlich ist die Wirkung des *Ausnahmegesetzes*, welches besagt, dass die Aufmerksamkeit unwillkürlich auf Objekte gelenkt wird, die anders sind. Je deutlicher sich diese von der Umgebung unterscheiden, umso besser nehmen wir sie wahr [131]. Eine zusätzliche Erhöhung des Kontrastes zwischen Farbe und schwarz/weiß gehaltener Umgebung führt zu einer stärkeren Wirkung dieses Gesetzes (vgl. Abbildung 5.19(b)).

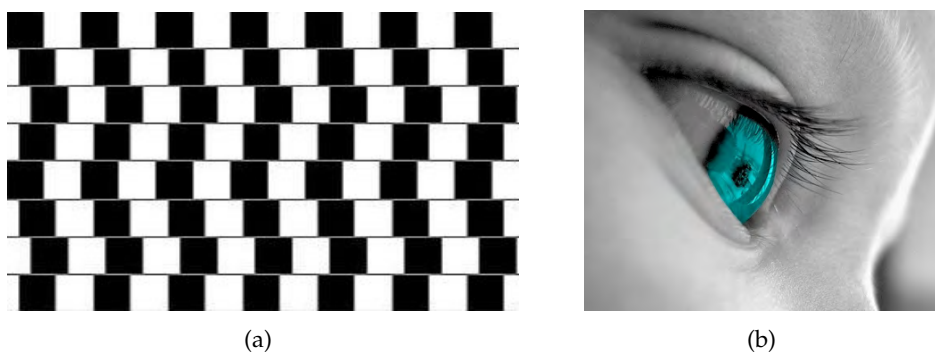


Abbildung 5.19: (a) Optische Täuschung, obwohl die horizontalen Linien parallel sind wirken sie schief. Entnommen aus [104]. (b) ColorKey.

5.5.1.3 Akustische Möglichkeiten der Aufmerksamkeitslenkung

Sobald eine graphische Aufmerksamkeitslenkung nicht möglich ist, können – wie z.B. heute bereits bei zahlreichen wichtigen Warnmeldungen im Automobil – akustische Reize zum Einsatz kommen. Um den Blick nicht fälschlicherweise auf die Warnmeldungsanzeige zu lenken, müssen sich diese von bekannten (statischen) Warntönen unterscheiden.

Der Vorteil der akustischen Ausgabe ist, dass der Fahrer sie wahrnimmt, egal wohin er blickt. Im Rahmen von Laborversuchen kann man zwar davon ausgehen, dass sich der Blick des Probanden nur zwischen Straße und FIS bewegt, im alltäglichen Fahrbetrieb ist es aber nicht ungewöhnlich, dass der Blick auch im Innenraum oder auf der Landschaft, anstatt auf der Straße liegt. Da eine optische Warnmeldung in solchen Situationen wenig sinnvoll ist, stellt der Warnton in diesem Fall eine zielführende Methode dar.

Zur signifikanten Differenzierung von anderen Warnmeldungen ist auf das Potential von bewegten Warntönen hinzuweisen. Ein solcher Ton sollte so konzeptioniert sein, dass er den Blick des Fahrers möglichst intuitiv vom aktuellen Fokus der Aufmerksamkeit in die Region lenken kann, die seine Aufmerksamkeit aktuell am dringendsten benötigt. Der Ton soll dabei so im Raum platziert sein, dass er sich vom aktuellen Aufmerksamkeitsfokus zur gewünschten Region bewegt. Bedient der Fahrer beispielsweise das

Menüsystem während das Fahrzeug eine drohende Gefahrensituation erkennt, wird ein Ton eingespielt, der von der Mittelkonsole hin zur Straße wandert.

Zur Realisierung solcher Warntöne kommt die am Lehrstuhl entwickelte *Virtuelle Akustik* zum Einsatz, die es ermöglicht, dem Probanden über einen Kopfhörer Schalle – in einem nahezu beliebigen Winkel horizontal um ihn herum platziert – darzubieten [247]. Das System nutzt die Fähigkeit des Menschen, Schalle über Laufzeitunterschiede räumlich zu lokalisieren. Zur virtuellen Emulierung dieser Ortung bedient man sich der *Head Related Transfer Function* (HRTF). Die HRTF beschreibt – abhängig von Umgebung (Wände, Möbel etc.) und personenabhängigen Parametern (Ohrmuschel, Nase etc.) – die Übertragung eines an einem bestimmten Ort platzierten Schalls zu den Ohren des Menschen [81].

5.5.2 Konzept und Implementierung

Die folgenden Abschnitte beschreiben die Umsetzung der Konzepte zur Aufmerksamkeitslenkung (AL). Um den Probanden eine möglichst realistische Aufgabe zu stellen und ihre Aufmerksamkeit gezielt zu lenken, erfolgte die prototypische Integration der Konzepte in ein Fahrerinformationssystem (vgl. Kapitel 5.1). Die Oberfläche gliedert

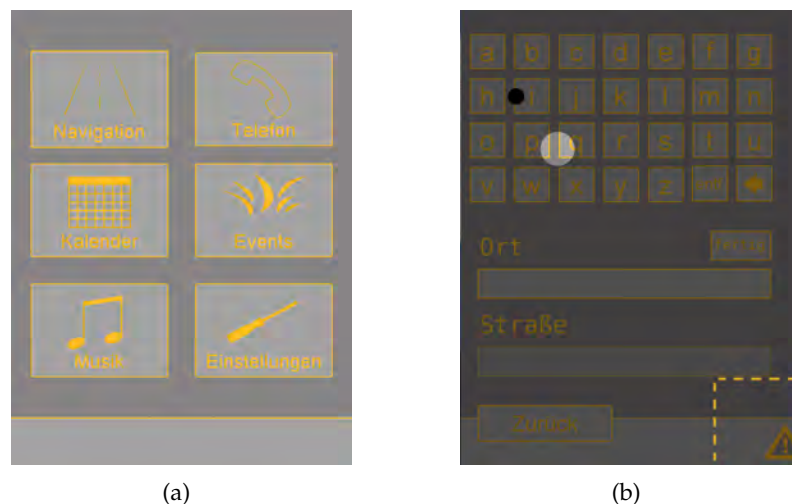


Abbildung 5.20: Graphische Oberfläche des prototypischen Interaktionssystems. (a) zeigt das Hauptmenü. (b) zeigt das Konzept der optischen Aufmerksamkeitslenkung mit Führungspunkt. Der kleine schwarze Punkt repräsentiert dabei die aktuelle Blickrichtung, die orangene Linie begrenzt das Zielfeld. Beide sind normalerweise nicht zu sehen.

sich in eine Statusleiste am unteren Rand und ein großflächiges Bedienfeld im restlichen Bereich (siehe Abbildung 5.20(a)). Die Statusleiste dient zur Anzeige von Warnhinweisen. Für die durchgeführten Versuche sind nur das Navigationssystem und das Telefon von Relevanz, da diese das größte Ablenkungspotential darstellen [225].

5.5.2.1 graphische Aufmerksamkeitslenkung

Es ergeben sich verschiedene Möglichkeiten die Aufmerksamkeit auf dem Bildschirm selbst zu lenken. Nach dem Orientierungsreflex zieht ein Blinken (=Bewegung) im peripheren Blickfeld genügend Aufmerksamkeit auf sich. Ein blinkendes Icon symbolisiert die Systemmeldung (vgl. Abbildung 5.20(b) rechts unten). Beim Blick auf das Warnschild erscheint die komplette Mitteilung, die durch einfaches Antippen quittiert werden kann.

In Phasen hoher Konzentration oder falls die Aufmerksamkeit auf der Straße und nicht auf dem Monitor liegt, kann der Fahrer das Blinken übersehen. Deshalb existiert der sog. Führungspunkt (FP). Zusätzlich zum blinkenden Icon dunkelt sich der Bildschirm mit Ausnahme eines kleinen Bereiches/Punktes ab. Dieser Punkt zieht die Aufmerksamkeit auf sich (vgl. Ausnahmegesetz [104]) und lenkt sobald er sich bewegt den Blick des Fahrers zum Zielbereich. Die Bewegung erfolgt entlang der Verbindungslinie von aktueller Blickposition zum Ziel, wobei der Punkt dem Blick vorausseilt. Dafür hat sich während Vorversuchen eine Entfernung von 150 Pixel als sinnvoll erwiesen.

Dieses Vorgehen ist vergleichbar mit einem Staubkorn, das sich auf dem Auge im Blickfeld befindet. Diesem folgt das Auge unbewusst und blickt dabei z.B. kontinuierlich immer weiter nach oben. Da die Bewegung des FP abhängig von der Blickrichtung erfolgt, ist sie schwerer zu übersehen oder zu ignorieren. Blickt der Nutzer zwischenzeitlich wieder auf die Straße, ist die Bewegung bei der erneuten Zuwendung zum FIS nicht abgeschlossen, sondern wird automatisch von der aktuellen Blickposition aus weitergeführt. Dadurch kann der Nutzer die Bewegung des Punktes nicht verpassen. Der Punkt wird ausgeblendet, sobald der Blick des Nutzers das Zielgebiet erreicht hat.

5.5.2.2 Akustische Aufmerksamkeitslenkung

Um den Schall an einem festen Ort – unabhängig von der Kopfposition der VP – zu platzieren, ist es nötig die Kopfposition des Probanden möglichst exakt zu registrieren. Dazu wird ein Infrarot-Tracking System sowie ein entsprechendes – auf dem Kopfhörer angebrachtes – Target benutzt. Somit ist es möglich, den Ton z.B auf der Straße zu platzieren, oder von der aktuellen Blickrichtung zur Straße wandern zu lassen.

Die verwendete Virtuelle Akustik kann Schalle nur in einer horizontalen Ebene um den Nutzer positionieren. Deshalb sind zur Simulation der Elevation des Warntones



Abbildung 5.21: Ein schmalbandiger 8kHz-Ton wird unabhängig vom Ort der Schallquelle als von oben kommend lokalisiert. Entnommen aus [81].

Anteile bei 8kHz hinzugefügt, die der Nutzer als von oben kommend wahrnimmt (vgl.

Abbildung 5.21). Da sich ein weißes Rauschen nicht als Warnton eignet (ungewohnt) und ein einzelner 8kHz Sinuston als unangenehm empfunden wird, müssen weitere Komponenten im Warnton vorhanden sein. Als Grund-Warnton dienen drei Sinustöne⁹ unterlegt mit einem breitbandigen Rauschen von 0,1 bis 8kHz sowie einem schmalbandigen Rauschen mit höherem Pegel um 8kHz. Damit ist der Ton gut zu lokalisieren und zieht die Aufmerksamkeit auch mit sich nach oben, ohne dabei unangenehm zu sein.¹⁰ Ähnlich dem Führungspunkt existiert ein Ton, der zwischen der aktuellen Blickrichtung und der zu erreichenden Zielblickrichtung – z.B. der Straße – positioniert ist und sich abhängig vom Blick bewegt. Somit soll der Ton die Aufmerksamkeit des Nutzers mit seiner Bewegung mit sich ziehen. Dagegen wurde ein fest in Richtung der Straße angesiedelter Ton evaluiert, der die Aufmerksamkeit eventuell schneller – weil direkter – lenkt.

5.5.2.3 Kombinierte Aufmerksamkeitslenkung

Zusätzlich ist die Kombination von graphischer und akustischer Aufmerksamkeitslenkung (AL) möglich. Hierbei wird der Führungspunkt genutzt, um die Aufmerksamkeit des Nutzers während der Menübedienung auf die Straße zu lenken. Die Verlagerung der Aufmerksamkeit nach vorn wird durch den Punkt gestartet und vom Ton weiter geführt, sobald der Blick am Randbereich des Displays angelangt ist.

5.5.3 Versuchsdurchführung

Zur Bewertung der Effektivität und Usability der Konzepte wurden sie im Rahmen einer Laborfahrtstudie untersucht. Dazu erfolgte eine Umsetzung auf der in Kapitel 5.1 vorgestellten Sitzkiste. Zur Darbietung der akustischen Hinweistöne tragen die VPn Kopfhörer, deren Position und Lage ein Tracking System registrierte. An den Versuchen nahmen 15 Probanden (14 männlich) mit einem Durchschnittsalter von 23,6 Jahren teil. Drei gaben an, noch nie ein Navigationssystem benutzt zu haben, sieben haben schon ein Navigationssystem mit Touchscreen benutzt. Das Interesse an der Technik war eher hoch angesiedelt und alle VPn bekundeten ihr Interesse an technischen Neuheiten.

Zur Bewertung der Fahrleistung dienten die Spurdaten des LCT (vgl. Kapitel 2.4.1.1) und zur Bewertung der Ablenkung von der primären Fahraufgabe die PDT (vgl. Kapitel 2.4.1.2). Zudem erfolgte eine Protokollierung der Reaktionszeiten auf Warnmeldungen und Warntöne, sowie der entsprechenden Blickpositionen zur Auswertung bei welcher Blickrichtung die VPn wie schnell reagieren (Blick auf Straße/FIS). Zusätzlich zu diesen objektiven Daten mussten die Probanden nach jedem Teilversuch einen Fragebogen ausfüllen (AttrakDiff, SEA und SUS). Nach den Versuchen erfolgte eine Abschlussbefragung.

Bei den Versuchen waren die VPn instruiert, dass während den Fahrten Warnmeldungen (akustisch/visuell) auftauchen. Dazu erhielten sie zusätzlich die Erklärung, dass vor einzelnen Spurwechseln Warntöne darauf hinweisen, dass sie die Aufmerksamkeit auf die Straße richten sollten. Da der LCT in sich geschlossen ist, war es nicht möglich, die Warntöne automatisiert auf die Folge der Schilder abzustimmen, weshalb sie der Versuchsleiter möglichst gleichverteilt (mit/ohne Führungspunkt) triggern musste. Als

⁹Jeweils im Abstand von einer Oktave bei 500Hz, 1kHz und 2kHz.

¹⁰Siehe Evaluierungsergebnisse: mit 93% Zustimmung wird der Warnton nicht als unangenehm bewertet.

Beispiel wurden ihnen die akustischen Warnungen (bewegt oder fest) vor den eigentlichen Versuchsfahrten präsentiert. Währenddessen konnten sich die VPn die Lautstärke angenehm einstellen. Während der Fahrt erfolgte die Steuerung von Zeitpunkt und Art der Warnmeldungen durch den Versuchsleiter. Dabei wurde auf eine Gleichverteilung der Meldungen geachtet, und der Zeitpunkt zufällig gewählt. Nach 10 Sekunden ohne Bestätigung durch den Probanden galt eine Meldung als nicht erkannt.

Zu Beginn erfolgte eine Referenzfahrt ohne Belastung (nur LCT und PDT) zur Bestimmung des Baseline Werts. Danach mussten die Probanden eine Fahrt mit Zieleingaben absolvieren. Getestet wurden dann die Anzeigen der Warnmeldungen auf dem FIS sowie der bewegte/feste Warnton. Zur Vermeidung von Einflüssen durch Lerneffekte erfolgte eine Permutation der Reihenfolge. Nach erfolgter Zieleingabe galt die (einzelne) Fahrt als beendet, da eine Weiterfahrt ohne Belastung das Ergebnis verfälscht hätte. Die einzugebenden Orte und Straßen waren so gewählt, dass die VPn möglichst lange beschäftigt waren.

5.5.4 Versuchsergebnisse

Die folgenden Abschnitte fassen die Versuchsergebnisse kurz zusammen. Für eine vollständige Übersicht der Versuchsergebnisse sei an dieser Stelle auf [85, 192] verwiesen.

Referenzfahrten Während der Baselinefahrten (ohne Bedienung) wurden alle PDT Punkte mit einer durchschnittlichen Reaktionszeit von 766ms erkannt. Die mittlere Spurabweichung des LCT betrug 0,51m. Bei den Referenzfahrten mit Zieleingabe ohne Warnung wurden bei einer mittleren Spurabweichung von 1,03m im Schnitt lediglich 81 Prozent der Punkte mit einer mittleren Reaktionszeit von 1043ms (durchschnittlich 277ms bzw. 36% langsamer) erkannt. Diese Ergebnisse sind statistisch signifikant.

Fahrten mit Warnton Während der Versuchsfahrten konnten lediglich 4 der 15 VPn den bewegten und den festen (auf der Straße liegenden) Ton voneinander unterscheiden, wohingegen bei einem Test ohne Fahr- und Sekundäraufgabe alle VPn die Bewegung des Tones registrierten. Der Unterschied ist vermutlich durch die erhöhte Belastung während der Versuche (Zieleingabe, PDT, LCT) zu erklären. Eine Person gab konkret an, die Bewegung nicht wahrgenommen zu haben, weil sie nicht darauf geachtet hat. Die Personen, die einen Unterschied bemerkten, empfanden die Aufgabe mit dem festen Ton als anstrengender (105 zu 90 auf der SEA-Skala), was zu einer größeren Spurabweichung (1,03m zu 0,95m) führte (nicht statistisch signifikant). Bei der Nachbefragung tendierten zwei VPn zum bewegten Ton, die beiden anderen gaben an, dass für sie kein Unterschied bestünde. Es ist also lediglich eine Tendenz hin zum bewegten Warnton vorhanden.

Verglichen mit der Fahrt ohne Warntöne konnte eine signifikante Verbesserung der Spurhaltung erreicht werden. Sie betrug beim festen Ton nur noch 0,83m (Faktor 1,6 verglichen mit der Referenzfahrt). Die Daten der Fahrten mit bewegtem Ton waren etwas schlechter (nicht statistisch signifikant). Die zusätzliche Spurabweichung ist auf schlechtes Spurhalten während der Eingabe zurückzuführen, übersehene Schilder sowie sehr späte Spurwechsel konnten die Warntöne allerdings nicht reduzieren.

Die PDT Reaktionszeit konnten die Warntöne nicht signifikant verändern (1064ms zu 1043ms). Dies war auch nicht anzunehmen, da die Warntöne zwar mit den Schildern, nicht

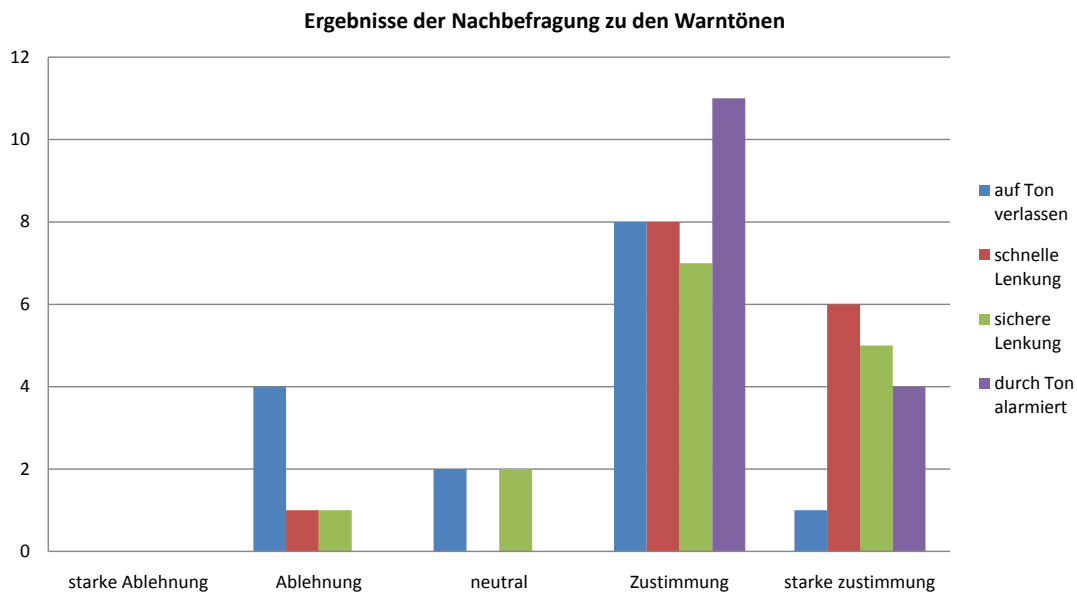


Abbildung 5.22: Nachbefragungsergebnisse zum Warnton jeweils als Anzahl VPn, die dem entsprechende Wert zustimmen, auf folgende Statements: Ich habe mich sehr auf den Warnton verlassen (blau). Der Ton führte meine Aufmerksamkeit schnell wieder auf die Straße (rot). Der Ton führte meine Aufmerksamkeit sicher wieder auf die Straße (grün). Durch den Ton bin ich sofort alarmiert (lila).

aber mit der PDT synchronisiert waren. Die Detektionsrate hat sich von 80,92% auf 70,06% verschlechtert. Dies lässt sich vermutlich dadurch begründen, dass sich die Probanden auf den Warnton verlassen und Sicherungsblicke seltener durchgeführt haben. Die VPn gaben dies auch in der Nachbefragung an (vgl. Abbildung 5.22, blaue Säulen). Dieser Effekt ist nicht zu vernachlässigen, jedoch sei darauf hingewiesen, dass während einer realen Fahrt Warntöne weniger häufig auftreten. Somit ist ein Gewöhnungseffekt weniger stark ausgeprägt. Die VPn gaben weiterhin an, dass der Warnton ihre Aufmerksamkeit schnell und sicher auf die Straße lenkt und sie sofort alarmiert sind (vgl. Abbildung 5.22).

Fahrt mit Aufmerksamkeitslenkung auf dem Monitor Bei der AL am Monitor erfolgten für die Konzepte mit/ohne Führungspunkt keine separaten Fahrten, da anzunehmen war, dass die Ergebnisse nicht signifikant voneinander abweichen. Die PDT Reaktionszeit liegt etwas höher als bei den Fahrten ohne Warnungen (1109ms zu 1043ms). Der Unterschied ist zwar signifikant, aber klein. Deutlicher zeigt sich die Belastung in den PDT Erkennungsraten von 81% zu 66%. Dies ist vermutlich darauf zurückzuführen, dass der Blick öfter auf dem FIS liegt, da Warnmeldungen erkannt und bestätigt werden mussten. Aus gleichem Grund ist die Spurabweichung beim LCT mit 1,18m etwas höher als bei den Fahrten mit Zieleingabe ohne Warnungen (nicht statistisch signifikant). Man kann annehmen, dass die Warnmeldungen nur eine geringe zusätzliche Belastung darstellen. Ein ähnliches Ergebnis liefert auch die subjektiv erlebte Anstrengung: 97 bei Fahrten mit Führungspunkt, 98 ohne Punkt und 89 mit Warntönen (entspricht „einigermaßen bis ziemlich anstrengend“).

5.5. Fahreraufmerksamkeitslenkung durch graphische und akustische Animationen

Die Auswertung der Reaktionszeiten auf Warnungen lieferte keine aufschlussreichen Ergebnisse, da die Bearbeitung der Primäraufgabe (LCT) stets mit höherer Priorität erfolgte. Im Schnitt galten 37% der blinkenden Warnmeldungen als abgebrochen, da sie nach 10 Sekunden nicht erkannt waren. Dies liegt zum Teil in der Platzierung der Warnungen begründet. Obwohl die VPn zu Beginn entsprechende Instruktionen erhielten, kam es des Öfteren zu Verdeckungen der Warnmeldungen durch den Arm der Probanden. Zusätzlich waren die VPn oft so stark mit der Primäraufgabe beschäftigt, dass sie das Blinken nicht registrierten. Dieser Umstand lässt sich auch durch eine andere Platzierung der Warnmeldung nicht ausschließen. Die AL mit Führungspunkt führte immer zum erwünschten Ziel und schnitt in allen Dimensionen des semantischen Differentials (vgl. Kapitel 2.4.1.4) am besten ab (vgl. Abbildung 5.23(a)). In den Dimensionen HQ-I und

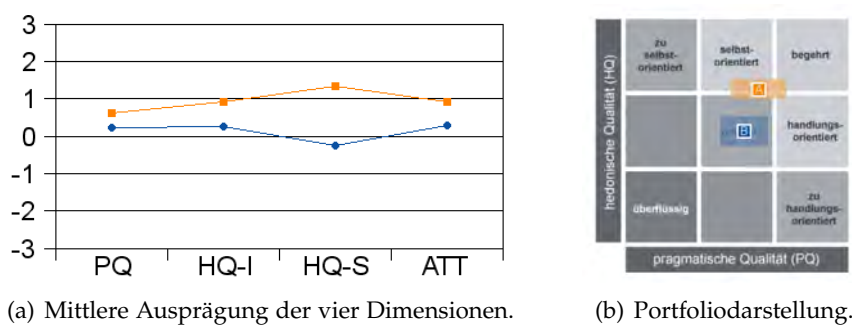


Abbildung 5.23: Ergebnisse AttrakDiff. In orange ist dabei die Aufmerksamkeitslenkung mit Führungspunkt dargestellt und in blau die ohne. (a) Mittlere Ausprägung (Ordinate) der vier Dimensionen. (b) Portfolio mit der durchschnittlichen Ausprägung der Dimensionen PQ und HQ und den jeweiligen Konfidenz-Rechtecken.

HQ-S ist der Unterschied statistisch signifikant. Beim Profil der Wortpaare schneidet die AL-FP besser ab. Besonders in der Dimension HQ-S sind die Unterschiede deutlich: der FP wird als besonders innovativ, originell, kreativ und neuartig bewertet. Die Bewertung als „mittelmäßig voraussagbar“, liegt in der Nutzung des Auges als Akteur begründet. Die Bewertung als praktisch, handhabbar und einfach lässt vermuten, dass sich die Augen bei häufiger Nutzung gewöhnen. Beim SUS schnitt die Lenkung mit FP ebenfalls besser ab als die Variante ohne FP (75 vs. 64 von 100 Punkten, statistisch signifikant).

Mehr als die Hälfte der VPn gab an, dass ihre Reaktion schneller erfolgte, wenn sie dem FP mit den Augen folgen konnten (siehe Abbildung 5.24, blaue Datenreihe). Rund die Hälfte der VPn fühlte sich durch den FP eingeschränkt/bevormundet (siehe Abbildung 5.24, rote Datenreihe). Man kann jedoch davon ausgehen, dass das Gefühl der Einschränkung während realen Fahrten deutlich zurückgeht, da Warnmeldungen unter realen Bedingungen weniger häufig und nicht stets während der Interaktion auftreten. Etwa ein Drittel der Probanden gab an, dass ein blinkendes Icon ausreicht (siehe Abbildung 5.24, grüne Datenreihe). Diese Bewertung gaben jedoch genau die Personen mit den meisten verpassten Warnmeldungen beim blinkenden Icon (40 bis 50 Prozent) ab. Diese Fehleinschätzung kann zustande kommen, da ihnen nicht bewusst war, ob und wie viele Warnmeldungen sie verpasst haben. Fast alle Probanden gaben an, dass ihr Blick durch den FP immer sicher auf dem Monitor landet (siehe Abbildung 5.24, lila

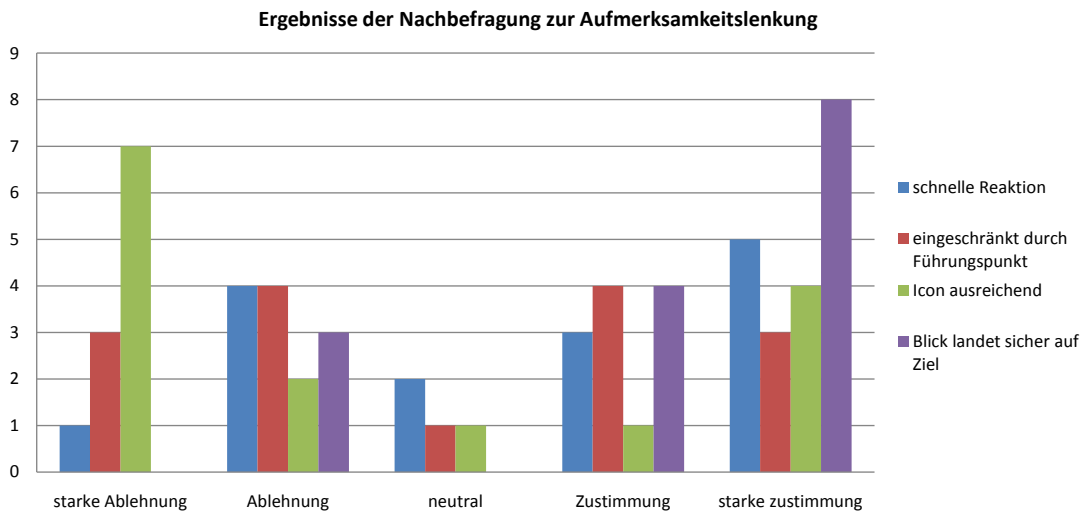


Abbildung 5.24: Nachbefragungsergebnisse zur Aufmerksamkeitslenkung, jeweils als Anzahl VPn, die dem entsprechende Wert zustimmen, auf folgende Statements: Meine Reaktion erfolgt schneller, wenn ich dem Führungspunkt mit den Augen folgen kann (blau). Ich fühle mich durch den Führungspunkt eingeschränkt (rot). Es reicht aus, wenn ein Icon blinkt (grün). Durch den Führungspunkt landet mein Blick auf jeden Fall auf dem Icon/Zielgebiet (lila).

Datenreihe), 12 Personen (80 Prozent) bewerteten den FP als sicherer und 10 Personen (67 Prozent) als angenehmer.

Fahrt mit Warnton und Führungspunkt Den Führungspunkt mit dem Warnton zu kombinieren scheint wenig sinnvoll: Er stärkt weniger die Orientierung nach vorn, sondern schwächt sie vielmehr ab. 10 VPn gaben an, dass der FP der Orientierung nach vorne nicht zuträglich sei, nur drei empfanden eine Verbesserung. Eine VP vermutete hinter dem FP eine Warnmeldung. Eine optische Aufmerksamkeitslenkung zusätzlich zum Warnton ist somit in dieser Ausführung nicht sinnvoll.

5.5.5 Fazit

Die Kombination von FP und Ton scheint wenig sinnvoll, da der FP den Blick eher von der Straße weg als auf den Monitor lenkte. Statt einer Lenkung der Aufmerksamkeit in die richtige Richtung wirkte dieses Konzept eher gegenteilig.

Die einzelnen Methoden sind jedoch sinnvoll und zielführend. Die AL auf dem Monitor stellte zwar eine geringe zusätzliche Ablenkung dar, die Ergebnisse zeigen aber eine durchweg positive Tendenz für die AL-FP. Etwa die Hälfte der Probanden fühlte sich zwar durch den Punkt eingeschränkt, die Vorteile sind jedoch deutlich in der Detektionsrate der Warnmeldungen zu erkennen. Während die Nutzer über ein Drittel der lediglich blinkenden Warnmeldungen verpassten, führte der FP den Blick immer sicher zum Ziel. Diese Methode sollte folglich nur dann eingesetzt werden, wenn eine möglichst zuverlässige Reaktion unabdinglich ist.

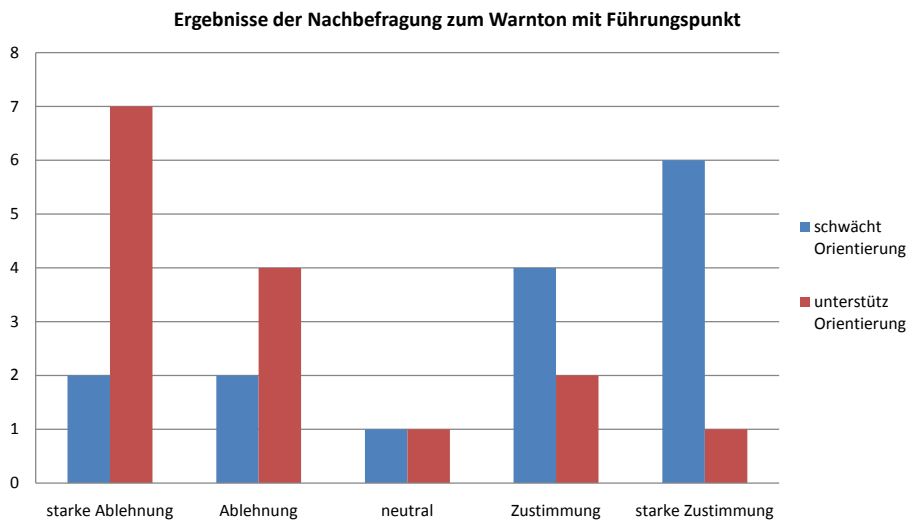


Abbildung 5.25: Nachbefragungsergebnisse zum Warnton mit Führungspunkt, jeweils als Anzahl VPn, die dem entsprechende Wert zustimmen, auf folgende Statements: Der Führungspunkt schwächt meine Orientierung nach vorn (blau). Der Führungspunkt unterstützt meine Orientierung nach vorn (rot).

Die AL mittels Warnton zeigte eine deutliche Verbesserung in der primären Fahraufgabe, die Spurabweichung reduzierte sich um rund ein Drittel durch Vermeidung von zu spätem Spurwechsel und verpassten Schildern. Alle VPn gaben an, durch den Ton sofort alarmiert zu sein. Vermutlich aufgrund der hohen Beanspruchung beim Versuch konnten die meisten VPn keinen Unterschied zwischen dem bewegten und dem festen Ton ausmachen. Die Integration von speziellen Warntönen zur schnellen und sicheren AL ist somit als sinnvoll und sehr effizient anzunehmen.

5.6 Intentionsbasierte Parametrierung der Mensch-Maschine Schnittstelle

Die Aktivierung von Infotainmentfunktionen erfolgt erst bei Bedarf. Sie erfordern somit einen aktiven Eingriff des Fahrers und gleichzeitig einen Teil seiner Aufmerksamkeit. Dies führt zu einem (teilweisen) Aufmerksamkeitsverlust für seine eigentliche Aufgabe und einem erhöhten Unfallrisiko [177]. Um dieses Problem zu entschärfen, sollen in den nächsten Abschnitten effektive Ansätze vorgestellt und bewertet werden, um die (Bedien-)Intentionen des Fahrers während der Bedienung (jedoch vor Erreichen des eigentlichen Bedienziesels) zu bestimmen und zur Parametrierung des Interaktionssystems zu verwenden. Dies ist eine Voraussetzung für eine effiziente Fahrer-Fahrzeug Interaktion, da man durch eine korrekte Interpretation der Benutzerintention von einer adäquaten Systemreaktion auf Benutzeraktionen ausgehen und eine intuitivere und weniger ablenkende Bedienung bewerkstelligen kann. Die Prädiktion von Benutzerintentionen basiert auf einer Analyse von Augen- und Handbewegungen [185].

5.6.1 Theoretische Grundlagen

Die folgenden Abschnitte erarbeiten zunächst theoretische Grundlagen zu Intentionen sowie zur Intentionserkennung. Aufbauend darauf werden die umgesetzten und untersuchten Konzepte zur Intensionsprädiktion und Ergebnisse einer Laborstudie vorgestellt.

5.6.1.1 Definition von Intentionen

Im allgemeinen Sprachgebrauch ist eine Intention als eine Absicht, ein Plan oder ein Vorhaben definiert. In der Psychologie beschreibt der Begriff der Intention den Prozess von einer intern getroffenen Instruktion — sowohl spezifische Handlungsabsichten als auch allgemeine Handlungskonzepte — hin zu einer konkreten Handlungsumsetzung und hängt somit mit der willentlichen Handlungssteuerung zusammen.

„Much of our understanding of ourselves and others is rooted in a commonsense psychological framework, one that sees intention as central. Within this framework we use intention to characterise both people’s actions and their minds. [...] These characterizations provide a basis for our everyday attempts to predict what others will do, explain what they have done and coordinate our projects with theirs.“ Nach [44], entnommen aus [105].

Kuhl [133] unterscheidet vier Makrosysteme der Handlungssteuerung. Eines davon, das Intensionsgedächtnis, ist mit dem Aufmerksamkeitssystem verknüpft und verstärkt absichtsrelevante Objekte. Die Intention und mit ihr der Weg zur Umsetzung eines Handlungszieles ist in zwei Komponenten zu unterteilen: die *Zielintention* und die *Durchführungsentention* [93]. Die Bildung einer Zielintention legt den angestrebten Zielzustand fest. Ab diesem Punkt werden verschiedene volitionale Mechanismen zur Unterstützung der Zielrealisierung relevant und die Durchführungsentention gewinnt an Bedeutung. Sie legt fest, *wann, wo und wie* zielrealisierende Handlungen auszuführen sind.

Dementsprechend ist laut Kuhl das wichtigste Hilfssystem, mit dem das Intensionsgedächtnis kooperiert, das analytische Denken: Das Intensionsgedächtnis enthält die unmittelbar zur Ausführung anstehenden Teile von Handlungsplänen, die die Prozesse des analytischen Denkens je nach ihrer Komplexität und Neuartigkeit produzieren [133]. Dabei wird die Aufrechterhaltung von Absichten von einer spezifischen Form der Aufmerksamkeit unterstützt. Sie verstärkt jene Signale der Umwelt, die eine hinreichende Passung zu den im Intensionsgedächtnis repräsentierten Abbildern von geplanten Handlungen aufweisen und zu Objekten, die für die Ausführung oder für das Erkennen einer Ausführungsgelegenheit relevant sind. Konkret bedeutet dies, dass die mentale Repräsentation der in der Durchführungsentention spezifizierten günstigen Handlungsgelegenheit eine erhöhte Aktivierung erfährt. Folglich ist diese Gelegenheit im Gedächtnis leichter zugänglich und wird bei ihrem Auftreten schneller entdeckt.

5.6.1.2 Motivations- und Handlungstheorie

Klassische Motivationstheorien erklären die Zielwahl anhand von Erwartungswert-Wert Ansätzen, welche die Erwartungen bei der Durchführung von Handlungen sowie deren Ergebnisse multiplikativ mit dem Wert, den der Handelnde den Auswirkungen der Handlung auf sich und seine Umwelt zumisst, verknüpft. Dabei wird die sog. Motivationsstärke gebildet [14, 82, 108].

Das Hauptinteresse psychologischer Motivationstheorien bestand bis in die 1980er Jahre in der Erklärung der Zielwahl. Dabei wurden Fragen der Zielrealisierung sowie zu den entsprechenden psychologischen Prozessen, die diese steuern und moderieren, kaum oder gar nicht thematisiert. *Heckhausen* charakterisiert diese Auslassung als "Handlungsloch" [103]. Mit dem *Rubikonmodell der Handlungsphasen* stellt er eine theoretische Konzeption vor, welche diese Lücke schließen soll. Dieses Modell teilt die zur Erreichung eines Ziels notwendigen Aktivitätssequenzen in vier unterschiedliche Phasen ein:

1. In der *prädeziSIONalen Phase* wird aus unterschiedlichen Wünschen einer ausgewählt und in eine verbindliche Intention umgewandelt.
2. In der *präaktionalen Phase* wird die Ausführung von Handlungen geplant.
3. In der *aktionalen Phase* werden die Handlungen ausgeführt.
4. In der (abschließenden) *postaktionalen Phase* wird die Handlung (Ausführung und Erfolg) rückblickend bewertet.

Gollwitzer ergänzt diese vier Phasen weiterhin um das Konzept der *Bewusstseinslagen* [93]. Damit beschreibt er eine Einstellung des kognitiven Systems, die dabei hilft, in der jeweiligen Handlungsphase anfallende Aufgaben möglichst effektiv zu bewältigen. Um eine möglichst objektiv realisierbare Intention zu bilden, wird die abwägende Bewusstseinslage beispielsweise mit der prädeziSIONalen Handlungsphase verknüpft. Weiterhin verarbeitet die planende Bewusstseinslage mögliche, rein subjektive Anreize und Erwartungen. Zusätzlich entwickelte sich unter anderem die *Theorie des geplanten Verhaltens*. Im Mittelpunkt dieser Theorie steht der Zusammenhang zwischen der Intentionstärke und dem Verhalten. Die Intentionstärke setzt sich dabei aus drei Faktoren zusammen [14]:

1. die Einstellung zum Verhalten (bezeichnet die subjektive Wahrscheinlichkeit mit der das Verhalten bestimmte, gewollte Konsequenzen hat und welche Bedeutung diese Konsequenzen haben),
2. die subjektive Norm (ergibt sich aus Überzeugungen bezüglich der normativen Erwartungen und der Motivation sich normgerecht zu verhalten) und
3. die wahrgenommene Verhaltenskontrolle (fasst Faktoren zusammen, die die Ausführung eines Verhaltens erleichtern oder erschweren).

5.6.1.3 Intentionserkennung

Heinze definiert den Vorgang der Intentionserkennung als

„[...] the process of becoming aware of the intention of another agent, and more technically, as the problem of inferring an agent's intention through its actions and their effects in the environment[...]“, Entnommen aus [105].

Bewusstwerdung oder Inferenz der Intention eines Gegenübers durch Beobachtung und Auswertung seiner Aktionen sowie deren Effekte auf die Umwelt. Aufgrund der

breitgefächerten Einsatzgebiete ist die Intentionserkennung in den letzten Jahren für unterschiedliche Wissenschaftsdisziplinen von großem Interesse geworden: [89] beschreibt z.B. ein Planerkennungssystem zur Unterstützung älterer Menschen und [173] stellt ein System zur Intentionserkennung im Fahrzeug vor, das zur effektiven Spurhalteassistenten intentionale von nicht-intentionalen Spurwechselmanövern unterscheiden kann.

Zur Bestimmung von Intentionen auf Basis von Handlungstheorien muss eine mess- und interpretierbare Grundlage vorhanden sein, da diese lediglich beschreiben, dass es verschiedene Intentionen gibt und dass zielführende Handlungen geplant, durchgeführt und kontrolliert werden. Diese Handlungen werden von einem charakteristischen Blickverhalten begleitet. Land stellte z.B. in seinen Studien fest¹¹, dass den meisten Handlungen ein Blick (*look-ahead fixation*) vorausseilt, um zu manipulierende Gegenstände sowie Hindernisse in der Umgebung zu lokalisieren und zu identifizieren. Auf Basis dieser Blicke erfolgt die kognitive Bestimmung der Manipulation. Anschließend Handlungen werden von Blicken begleitet (*guiding fixation*). Sie dienen dem Ausweichen von Hindernissen und einem korrekten Abschluss der Handlung. Ist das Ziel erreicht, können abschließende Kontrollblicke (*look-back fixation*) erfolgen. Bei komplexeren Handlungsabläufen erfolgt eine Zerlegung in Einzelschritte, die demselben Schema folgen.

Eine alleinige Analyse von Blickmustern kann nicht alle Handlungen hinreichend genau beschreiben. Befindet sich der Gegenstand beispielsweise bereits in oder in der Nähe der Hand (z.B. Knopf am Lenkrad), kann die Fixation auf diesen ganz (oder teilweise) entfallen, da die bereits gewonnene räumliche Vorstellung ausreicht, um die Handlung durchzuführen. Dies gilt insbesondere bei einstudierten Handlungen¹² [135]. Daraus ergeben sich für eine Intentionserkennung folgende Basiskomponenten:

- eine Datenbank mit allen Intentionen sowie den zugehörigen Handlungsabläufen
- ein System zum Aufzeichnen und Auswerten von Blicken bzw. Blickpunkten
- ein System, das die Handpositionen und -bewegungen auswertet
- komplexe Regelstrukturen, welche die Auswertungen zusammenfassen und mit der Intensionsdatenbank vergleichen, um daraus die eigentliche Intention zu bestimmen

Dies entspricht den Grundzügen *Hofmanns* intentionsbasierten Ansatzes zur Interpretation von Benutzeraktionen und der Planerkennung [110].

5.6.2 Intentionsbasierte Planerkennung

Hofmann stellt den intentionsbasierten Ansatz zur Interpretation von Benutzeraktionen vor, der eine applikations- und modalitätenunabhängige Systemarchitektur nutzt [110]. Im Folgenden wird der Ansatz erläutert und an die vorliegende Problemstellung angepasst.

Abbildung 5.26 skizziert die grundlegende Systemarchitektur der intentionsbasierten Interpretation von Benutzeraktionen. Die *Intentionsbibliothek I* (vgl. Kapitel 5.6.1.3) enthält das Wissen über potenzielle Benutzerziele sowie den sog. *Intentionshypothesenraum*.

¹¹[135] und [136] untersuchen z.B. alltägliche Vorgänge wie das Händewaschen oder das Teekochen.

¹²Nach Rasmussen fertigungs-basierte Aufgaben (vgl. Kapitel 2.2.3), z.B. das der Kuppeln oder das Schalten.

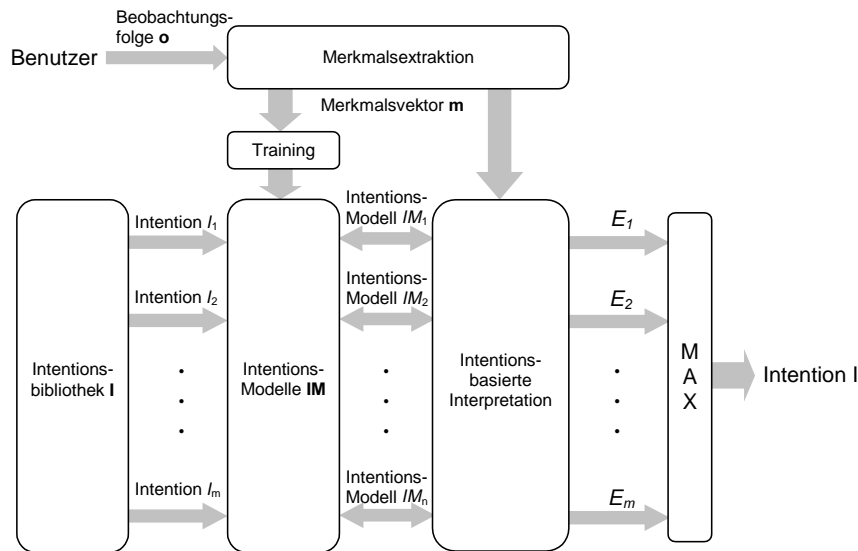


Abbildung 5.26: Grundlegende Systemarchitektur der intensionsbasierten Interpretation von Benutzeraktionen. Entnommen aus [110].

Dieser wird durch einzelne *Intensionsmodelle* **IM** repräsentiert. Sie stellen das Bindeglied zwischen den Intensionshypothesen und den aus der *Beobachtungsfolge* o hervorgehenden Merkmalen dar und enthalten das gesamte Wissen über den Verstehensprozess. o wird durch einzelne Benutzeraktionen erzeugt und mittels einer Merkmalsextraktion auf die entscheidenden informationstragenden Elemente/Merkmale reduziert. Der *Merkmalsvektor* m kann zum Training der Intensionsmodelle oder zur Interpretation von o dienen. Während der Interpretationsphase wird m auf die **IM** abgebildet, um die Nutzereingaben zu rekonstruieren. Durch Interpretation der rekonstruierten Eingabe kann auf die eigentliche Absicht des Nutzers geschlossen und entsprechend darauf reagiert werden. Dabei beschreibt das *Evaluierungsergebnis* E die Güte der Repräsentation der einzelnen Intensionshypothesen durch die Merkmale. Als Ergebnis des Interpretationsprozesses erhält man schließlich die Intention mit dem maximalen Evaluierungsmaß.

5.6.2.1 Beobachtungsfolge und Merkmalsextraktion

Der Fahrer artikuliert seine Absichten mittels Aktionen, die als Beobachtungsfolge vorliegen. Diese wird im laufenden Betrieb analysiert und dient der Rekonstruktion der Nutzeraktion auf Basis der Beobachtungsfolge. Aus der rekonstruierten Nutzeraktion wird im Rahmen der Intensionsdecodierung die zugrunde liegende Intention bestimmt.

Dabei können einzelne Beobachtungen eintreten oder nicht. Weiterhin können Beobachtungen von anderen Beobachtungen und Zeitverläufen abhängig sein. Deshalb sollten möglichst alle Beobachtungen, die sich aus der Umgebung, der Situation und den Intentionen eines Fahrers ergeben können, sowie deren zeitliches Auftreten festgehalten und betrachtet werden. Die Beobachtungen selbst ergeben sich dabei vor allem aus Aktionen des Fahrers, der Umwelt und der Situation. Daraus kann man eine Liste erstellen, in der für jeden Zustand einer Beobachtung eine Variable hinterlegt ist. Fahrerbezogene

Beobachtungen geben z.B. wieder, wohin der Fahrer sieht (Straße, Cockpit, Beifahrer, etc.) und welche Aktionen (Blinken, etc.) er ausführt. Umwelt- bzw. situationsbezogene Beobachtungen können z.B. ein Fremdverkehrsteilnehmer (vorhanden, nicht vorhanden, auf der eigenen Fahrspur, etc.) oder das Betätigen des Blinkers (Blinker gesetzt/nicht gesetzt) sein. Die Gesamtheit dieser Beobachtungen wird als Beobachtungsfolge \mathbf{o} bezeichnet und stellt eine enorme Fülle an situativen Informationen zur Verarbeitung und Analyse bereit. Da dies jedoch zu viele und teilweise irrelevante Informationen beinhaltet, kann und muss \mathbf{o} über eine geeignete Form der *Merkmalsselektion* eingeschränkt sowie anschließend entsprechende Merkmale in der *Merkmalsextraktion* herausgefiltert werden.

Die Merkmalsextraktion schließt zunächst alle Beobachtungen aus, die keinen Beitrag zur Bewertung der aktuellen Situation liefern (z.B. wird es auf Überlandstraßen keine plötzlich einscherenden Verkehrsteilnehmer geben). Weiterhin schließt dieser Vorgang beispielsweise Beobachtungen aus, deren Zustand sich nicht ändert. Der zweite Schritt ist die Einschränkung der Beobachtungsfolge auf Basis einer Referenzsituation, die beispielsweise folgende Parameter enthält

- Grundeinstellungen der Fahrzeugsysteme (z.B. Voreinstellungen Radio, etc.)
- Bestimmte Fahrsituationen (5. Gang, Tempo 130 auf der Autobahn)
- Bestimmtes Fahrerverhalten (Blick auf Fahrbahn, beide Hände am Lenkrad)

Auf Basis dieser Referenzsituationen werden alle Teile von \mathbf{o} verworfen, die dazu führen, die Referenzsituation wiederherzustellen. Weiterhin muss man \mathbf{o} auf Korrekturverhalten überprüfen: revidiert der Nutzer z.B. Aktionen zeitnah, kann man davon ausgehen, dass diese nicht der Intention des Fahrers entsprachen. Diese Zyklen können (mit wenigen Ausnahmen) aus \mathbf{o} gestrichen werden. Daraus resultiert der Merkmalsvektor \mathbf{m} .

5.6.2.2 Bildung der Intensionsmodelle

Intensionsmodelle **IM** verknüpfen die Intentionen mit zugehörigen Beobachtungen. Dazu sind alle Intentionen in der Intensionsbibliothek **I** hinterlegt. [110] geht davon aus, dass "der Benutzer in der Regel nicht mehrere Ziele gleichzeitig verfolgt" und daher "genau ein Eintrag der Intensionsbibliothek der aktuellen Absicht des Benutzers" entspricht.

Diese Absicht wird durch Einzelaktionen bzw. komplexere Aktionssequenzen erreicht. Abhängig von der Komplexität des betrachteten Systems ist jede Systemfunktion als eigene Intention modelliert oder mehrere Systemfunktionen zu einem Funktionsblock zusammengefasst (z.B. Funktionsblock „Tastatur“ zur Intention Texteingabe). Die Zuordnung von korrespondierenden Beobachtungen zu einzelnen Intentionen erfolgt über syntaktisch-semantische Bezüge (Intensionsmodelle **IM**) [177, 182].

Grundlage der Intensionsmodelle ist das Bayestheorem (siehe Formel 5.1), das sich aufgrund des umgekehrten probabilistischen Ansatzes am besten eignet, das nötige Wissen für die Interpretationsprozesse bereit zu halten. Der probabilistische Ansatz besagt, dass alle Aussagen zu einer bestimmten Sicherheit getroffen werden. Die Umkehrung stellt das Bayestheorem dar und erlaubt das *Schließen unter Unsicherheit*, was gleichbedeutend ist mit der Fähigkeit, unvollständige Informationen zu verarbeiten. Dazu müssen jedoch

weitere Voraussetzungen erfüllt sein: Die Ergebnismenge der Variablen (der Intentionen und der Beobachtungen) muss endlich sein und die a-priori Wahrscheinlichkeiten sowie die bedingten Wahrscheinlichkeiten der Variablen müssen bekannt sein [29, 110, 121].

Die Darstellung erfolgt in sog. *Bayes'schen Netzen*, die aus einem Graphen und bedingten Wahrscheinlichkeitsverteilungen bestehen. Der Graph enthält Knoten, die Zufallsvariablen repräsentieren, sowie gerichtete Kanten zwischen den Knoten, die (direkte) kausale Zusammenhänge symbolisieren. Für einen ausführlichen Überblick zu Bayes'schen Netzen sei an dieser Stelle unter anderem auf [15, 121, 198] verwiesen. Im vorliegenden Kontext modellieren sie hauptsächlich die Beziehungen zwischen extrahierten Merkmalen und den Intentionshypothesen. Abbildung 5.27 skizziert diesen Zusammenhang exemplarisch anhand der allgemeinen Struktur eines Intensionsmodells nach [110], welches die Struktur hierarchisch in folgende drei Ebenen aufteilt:

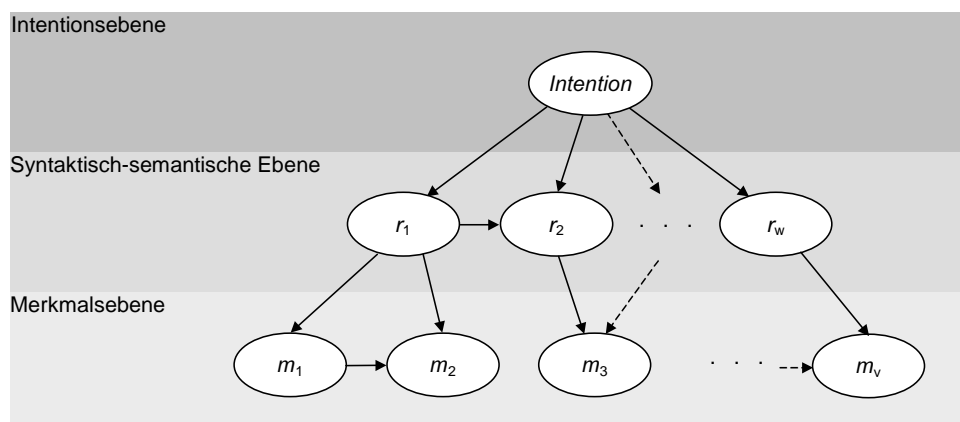


Abbildung 5.27: Allgemeine Struktur eines Intensionsmodells. Entnommen aus [110].

1. Die *Intentionsebene* repräsentiert alle (berücksichtigten) Intentionen und besteht entweder aus Knoten (Intentionsknoten) mit jeweils einem Zustand für jede Intention oder einem Knoten mit einem Boole'schen Zustandsraum für alle Intentionen. Diese sog. Wurzelknoten ermöglichen die wahrheitsbasierten Aussagen.
2. Die *syntaktisch-semantische Ebene* modelliert die statistischen Zusammenhänge zur Umsetzung der Klassifikation und verbindet die Intentionsebene mit der Merkmalsebene. Hier werden Beobachtungen durch probabilistisches Schließen verarbeitet und bezüglich der mit dem IM korrespondierenden Intention evaluiert.
3. Die *Merkmalsebene* besteht aus den sog. Merkmalsknoten, die Beobachtungen bestimmter Merkmale repräsentieren. Diese Ebene stellt die einzige Schnittstelle eines Intensionsmodells zur Merkmalsextraktion dar.

5.6.2.3 Interpretationsprozesse

Im Verlauf der Interpretation wird aus dem Abgleich des Merkmalsvektors mit den Intensionsmodellen unter Bildung eines Evaluierungsmaßes die wahrscheinlichste Intention

bestimmt. Dazu kann z.B. eine Strichliste geführt werden, bei der für jede Intention das Auftreten einer zugehörigen Beobachtung des Merkmalsvektors festgehalten wird. Die Häufigkeit des Auftretens stellt das Evaluierungsergebnis der jeweiligen Intention dar. Die Intention mit dem größten Evaluierungsergebnis wird als wahrscheinlichste Fahrerintention interpretiert. Weiterhin erfolgen Wahrscheinlichkeitsbetrachtungen über den Satz von Bayes, der aus der Umkehrung der bedingten Wahrscheinlichkeit Rückschlüsse von der Beobachtung auf die Intention erlaubt.

$$P(I|o) = \frac{P(o|I) \cdot P(I)}{P(o)} \quad (5.1)$$

Alle Intentionen erhalten eine Grundwahrscheinlichkeit $P(I)$, die mit einem Bewertungsfaktor multipliziert wird, der die Wahrscheinlichkeit festsetzt, dass eine Beobachtung zu einer bestimmten Intention gehört $P(o|I)$. Anschließend werden die bewerteten Wahrscheinlichkeiten durch die Grundwahrscheinlichkeit der Beobachtung $P(o)$ dividiert. Die wahrscheinlichste Intention ergibt sich aus dem Maximum aller Intentionen. Dieses Verfahren ermöglicht die Analyse und Bewertung komplexer Sachverhalte und hat den Vorteil, nicht nur lineare und statische Ereignisse erfassen zu können, sondern erlaubt auch das Erkennen von Präferenzen. Der Bewertungsfaktor bringt Vorkenntnisse ein, die sich z.B. aus der aktuellen Fahrsituation oder einer Benutzermodellierung ergeben.

5.6.2.4 Verwendete Intentionsmodelle

Da die Intentionsmodelle, die sich bei der Betrachtung eines Fahrers im Verkehr ergeben, sehr umfassend sind und sich die vorliegende Arbeit speziell mit der Bedienung von Infotainmentsystemen beschäftigt, sollen im Folgenden lediglich für diesen Bereich relevante (Bedien-)Intentionen betrachtet werden. Einflüsse aus der Umwelt und der Fahrsituation bleiben dabei zunächst unberücksichtigt, weshalb sich relevante Intentionen direkt aus dem Aufbau des Infotainmentmenüs ableiten lassen:

1. Fahrt zu einem einzugebendem Ziel
2. Fahrt zu einem bekannten Ziel
3. Fahrt nach Hause
4. Fahrt zur Arbeit
5. Karte anzeigen
6. Staumeldungen abfragen
7. Telefonnummer eingeben und anrufen
8. Telefonnummer im Adressbuch suchen und anrufen
9. Radiosender einstellen
10. CD/DVD wechseln
11. Fernsehsender einstellen
12. Internetverbindung herstellen
13. Klimaeinstellungen ändern
14. Fahrzeugeinstellungen ändern
15. Berichtigung von Eingaben
16. Rücksprünge im Menü

Diese Intentionen ($I_1 \dots I_{16}$) erfassen alle Funktionstasten (vgl. Abbildung 5.2(b)), die zum Merkmalsvektor beitragen. *Rücksprünge im Menü* (I_{16}) erfasst alle Funktionen ohne Zuordnung zu anderen Intentionen. Aufgrund der eindeutigen Zuordnung der Funktionstasten und dem Aufbau des Menüs gestaltet sich die Bildung der Intentionsmodelle einfach, denn mit Ausnahme von I_5 , I_{15} und I_{16} folgen die Aktionen des Benutzers immer der Menüführung. Vom Hauptmenü aus wird immer ein Untermenü aufgerufen, welches bereits die Intention als solche ausführt oder durch Auswahl eines zweiten Untermenüs die beabsichtigte Intention erreicht. I_5 stellt einen Sonderfall dar, da die Kartenanzeige auf zwei Wegen erreichbar ist (über das Hauptmenü oder das Untermenü Navigation). I_{15} und I_{16} sind losgelöst zu betrachten, da sie zur Erfassung von Fehleingaben und Rücksprüngen dienen.

5.6.2.5 Interpretationsprozess während der Evaluierung

Der Interpretationsprozess entspricht einer zeitgesteuerten Schleife [110], die alle Berechnungen von der Extraktion von \mathbf{m} bis hin zur Bestimmung des Maximums von \mathbf{E} ausführt. Die Schleife muss zeitgesteuert sein, da das Auslesen und Aktualisieren von Hand- und Blickdaten zu bestimmten Zeitpunkten erfolgen muss. Der Eye Tracker (vgl. Kapitel 3.3) liefert Daten mit einer Rate von 30 Hz. Der Touchscreen liefert laut Hersteller bei einer 4 mm dicken Sensorscheibe (vgl. Kapitel 5.1) alle 13 ms (bei dickeren Gläsern bis zu alle 20 ms) neue Datenpakete. Zudem ist eine Verarbeitungszeit nötig, weshalb alle 10 ms Daten ausgelesen werden sollten um stets mit aktuellen Werten zu rechnen.

Der Interpretationsprozess beginnt mit der Bestimmung der relevanten Daten. Dazu liest das System die aktuellen Hand- und Blickdaten aus und ermittelt danach, ob sich die Hand in der Nähe oder auf dem Touchscreen (TS) befindet und entscheidet anhand von Tabelle 5.3 nach dem Abstand der Hand (Z) auf welcher Datenbasis eine Beobachtung erfolgt. Blickdaten sind im Gegensatz zu Handdaten per Definition immer

Handdaten	Datenbasis
Hand abseits oder $Z > 35$	Blickdaten
Hand unbewegt und $Z > 20$	Blickdaten
Hand bewegt und $Z > 20$	Handdaten
Hand am TS und $Z < 20$	Handdaten

Tabelle 5.3: Kriterien zur Festlegung der verwendeten Datenbasis.

verfügbar. Liegt der Blick nicht auf dem Touchscreen oder wird der Blick nicht richtig erfasst, kann das System keine aktuelle Handlung bestimmen. In diesem Fall wird *keine Beobachtung* festgelegt. Ändert sich der Blickpunkt von einem Datum zum nächsten nur minimal bzw. ist kein Blickpunkt bestimmbar, verwendet der Algorithmus die zuletzt gültigen Blickdaten. Dies eliminiert kurzzeitige Signalverluste (z.B. Blinzeln) und minimale Schwankungen, die sich aus *look-ahead fixations* oder minimalen Kopfbewegungen ergeben.

Sobald die aktuelle Handlung des Fahrers bestimmt ist, wird diese an das bisherige \mathbf{o} angehängt und auf den letzten Rücksprung ins Hauptmenü überprüft. \mathbf{m} setzt sich

aus allen Beobachtungen zusammen, die zeitlich nach dem Rücksprung folgen. Somit müssen lediglich Rücksprünge eliminiert und mit den **IM** verglichen werden.

Abhängig vom aktuellen Menüzustand wird eine Vorauswahl bzgl. der **IM** getroffen. Intentionen ohne Rücksprünge im Menü werden als wahrscheinlicher angesehen und entsprechend mit hoher Wahrscheinlichkeit bewertet. Die restlichen Intentionen erhalten eine niedrigere Wahrscheinlichkeit. Dies sind die Wahrscheinlichkeiten der letzten Beobachtung von **m** unter Voraussetzung des jeweiligen **IM** $P(o|I_x)$. Als nächstes werden die Wahrscheinlichkeiten mit den Grundwahrscheinlichkeiten der **IM** $P(I_x)$ multipliziert und durch die Wahrscheinlichkeit der Beobachtung $P(o)$ dividiert. Damit stehen die Evaluierungsmaße fest und man kann das Maximum bestimmen.

Eine weitere Möglichkeit zur Festlegung von $P(o|I_x)$ besteht in der Berechnung von Wahrscheinlichkeitspfaden in den einzelnen Intentionsmodellen. Dazu erhält $P(o|I_x)$ für jede Intention den Wert der Wahrscheinlichkeit des Rücksprungs vom aktuellen Menüpunkt zum Hauptmenü. Anschließend wird dieser Wert dem Pfad des jeweiligen **IM** folgend mit den Wahrscheinlichkeiten der Knotenpunkte multipliziert. Da ein Rücksprung im Menü weniger wahrscheinlich ist als ein Vorwärtssprung, entsteht bei Intentionen ohne Rücksprung ein Fehler. Zum Ausgleich wird die Wahrscheinlichkeit dieser Intentionen mit einem Faktor multipliziert, der umgekehrt proportional zur Wahrscheinlichkeit der doppelt durchlaufenen Knotenpunkte ist.

Ein alternativer Interpretationsprozess ist die Bildung einer Häufigkeitstabelle. Dabei wird aus der Häufigkeit des Auftretens von Beobachtungen eines Intentionsmodells im Merkmalsvektor das Evaluierungsmaß bestimmt. Diesen Interpretationsprozess kann man um eine Gewichtung der festgestellten Beobachtungen erweitern, um den gerichteten Kanten der **IM** eine höhere Bedeutung zukommen zu lassen.

5.6.3 Versuchsergebnisse: Methodenbewertung

Der zweiteilige Versuch dient der Bewertung der Intentionserkennung. Im ersten Teil konnten die Probanden das Interaktionssystem ohne Vorgaben kennenlernen. Die aufgezeichneten (und später annotierten) Handdaten der Probanden dienen zur (folgenden offline) Optimierung der Prädiktion der Nutzerintentionen. Der zweite Teil beinhaltete begleitend zum Versuch ein kurzes Szenario, nach dem sich die Probanden bei der Bedienung richten sollten. Während dieses Versuchs erfolgte eine Aufzeichnung und Analyse von sowohl Hand- als auch Blickdaten. Dies diente der Verifizierung, ob das System alle Intentionen richtig erkennt und wie lange vor dem Betätigen einer Funktionstaste ersichtlich ist, dass diese gedrückt werden soll.

Tabelle 5.4 gibt einen Überblick zur Gültigkeit der Daten¹³ sowie die Erkennungsraten der Fahrerintention mit den Verfahren Häufigkeitstabelle (HT) und intentionsbasierte Planerkennung (IP). Die Gültigkeit der Daten ergibt sich aus der Erkennung von Hand- und Blickdaten sowie der Bildung des richtigen Merkmalsvektors. Ebenfalls verzeichnet ist die Häufigkeit der Handeingaben auf dem Touchscreen. Die Werte hinter dem Schrägstrich geben an, wie viele Daten auf den Blick zurückzuführen sind und wie oft

¹³Versuchsteil II konnte von VP5 und 6 nicht ausgewertet werden, da der Eyetracker die Pupillen nicht ausreichend gut über die gesamte Versuchsdauer erfassen konnte, was durch zu dominante Reflexionen auf Brillengläsern hervorgerufen wurde.

5.6. Intentionbasierte Parametrierung der Mensch-Maschine Schnittstelle

Blicke die Eingaben begleiteten. Da sich die Erkennungsraten der Interpretationsprozesse um bis zu 10% unterscheiden, erfolgte zur Bewertung der Interpretationsprozesse unter unterschiedlichen Bedingungen eine getrennte Betrachtung von Hand- und Blickdaten bzw. der kombinierten Daten.

Versuchsperson	1		2		3	
Versuchsteil (VT)	I	II	I	II	I	II
Gültigkeit d. Daten	86,2%	96,5%	81,4%	90,0%	94,0%	97,1%
Anzahl Daten	65	141/93	113	211/126	116	211/136
mit HT erkannt	83,1%	96,5%	72,5%	90,0%	87,9%	97,1%
mit IP erkannt	80,0%	96,5%	68,1%	84,4%	95,3%	96,3%
Manipulationen	56	23/20	38	34/32	40	33/28

Versuchsperson	4		5		6	
Versuchsteil (VT)	I	II	I	II	I	II
Gültigkeit d. Daten	79,4%	78,8%	86,0%	keine	78,6%	keine
Anzahl Daten	126	156/39	171	Blick-	169	Blick-
mit HT erkannt	78,6%	78,2%	86,0%	daten	74,6%	daten
mit IP erkannt	74,6%	75,0%	86,0%	ver-	74,0%	ver-
Manipulationen	41	34/29	69	ffügbar	54	ffügbar

Tabelle 5.4: Gültigkeit und Erkennungsrate für beide Versuchsteile/alle Versuchspersonen.

	Gesamt	Hand	Blick	VT I	VT II
Daten gesamt	1479	1085	394	760	719
davon gültig	1291	938	353	637	654
Gültige in Prozent	87,3%	86,5%	89,6%	83,3%	91,0%
mit HT erkannt	1263	-	-	610	653
in Prozent	85,4%	-	-	80,3%	90,8%
mit IP erkannt	1228	-	-	594	634
in Prozent	83,0%	-	-	78,2%	88,2%
Index berichtigt, IP	+50	+40	+10	+30	+20
damit erkennt IP	1258	-	-	624	654
in Prozent	86,4%	-	-	82,1%	91,0%

Tabelle 5.5: Auswertung der Versuchsdaten nach verwendeter Datenbasis.

Tabelle 5.5 fasst das Ergebnis dieser Auswertung zusammen. Bei Rücksprüngen ändert sich der Menüindex nach der Berechnung aus dem Merkmalsvektor. Der Menüindex ist durch die Menüstruktur des Interaktionssystems festgelegt. Im vorliegenden Aufbau ergibt sich ein Offset, welches die Korrekturwerte in Tabelle 5.5 ausgleichen. Neben Fragebogen, Gültigkeit der Daten und den Erkennungsraten wurden die Zeiten zur Erkennung einer Intention ermittelt. Die Diagramme stellen die Ergebnisse nach Versuchsteil und Datenbasis als Mittelwert über alle VPn dar. Der Übersichtlichkeit halber sei an dieser Stelle auf die ausführliche Übersicht der Daten in Anhang A.1 verwiesen.

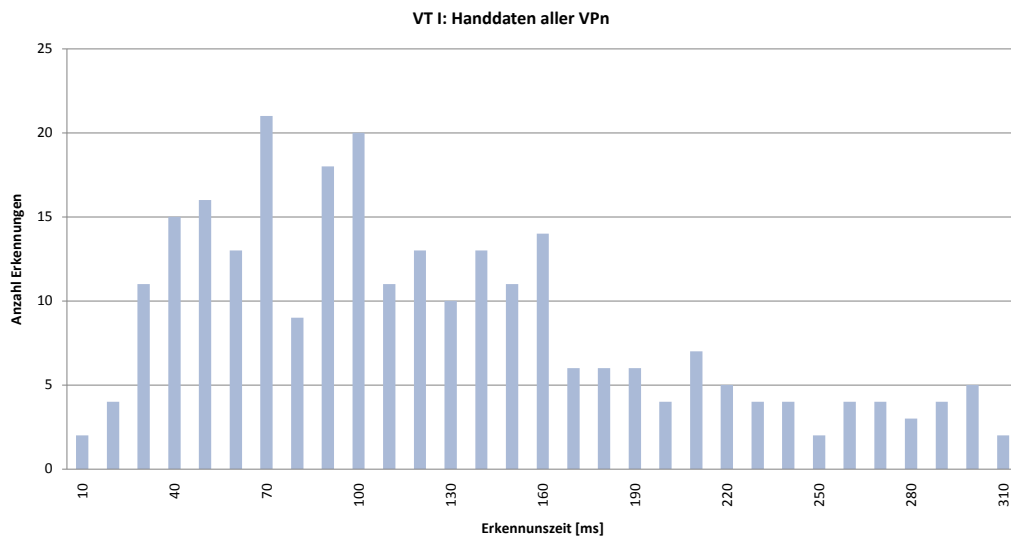


Abbildung 5.28: Versuchsteil I, Durchschnittswerte der Handdaten aller Versuchspersonen. Hier dargestellt der Erkennungsbereich von 10 ms bis 310 ms vor Durchführung der eigentlichen Aktion. Im restlichen Zeitbereich sind keine signifikanten Erkennungsanzahlen vorhanden.

5.6.3.1 Diskussion der Ergebnisse

Bei der Diskussion der Versuchsergebnisse sind verschiedene Aspekte zu beachten: Zum einen nahmen an VT I lediglich sechs VPn teil, zum anderen wurden 1479 Ereignisse (siehe Tabelle 5.4) verzeichnet, die eine gute Datenbasis zur Überprüfung der Fahrerintention bilden (davon 1085 auf Handdaten und 394 auf Blickdaten).

Erkennungsraten In der gesamten Datenmenge sind 1291 von 1479 Ereignissen als gültig verzeichnet (durchschnittliche Gültigkeit 87,3%) und es ergibt sich eine Gültigkeit von 86,5% für Hand- und 89,6% für Blickdaten. Eine getrennte Analyse von VT I und VT II ergibt Erkennungsleistungen von 83,8% (VT I, nur Handdaten) und 91,0% (VT II, kombiniert) und es ergibt sich eine Steigerung um ca. 7% bei kombinierter Erkennung.

Die Intentionserkennung lieferte in VT I eine durchschnittliche Trefferquote von 80,3% bei Verwendung der Häufigkeitstabelle (HT) und von 78,2% bei der intentionsbasierten Planerkennung (IP). In VT II schneidet die HT mit 90,8% ebenfalls besser ab als die IP mit 88,2%. Dies ist vor allem darin begründet, dass die einfache und flache Menühierarchie der HT sehr entgegenkommt. Ab einer bestimmten Länge des Merkmalsvektors lässt dieses Verfahren keine Fehlinterpretation mehr zu, was allerdings bei verfälschtem Merkmalsvektor unweigerlich zu Fehlinterpretationen führt.

Die IP führt die Beurteilung der Fahrerintention ausgehend von der letzten Beobachtung von \mathbf{m} durch, wodurch die Interpretation stets direkt von der aktuellen Handlung abhängig ist. Somit ergibt sich eine höhere Entscheidungssicherheit bei verfälschtem \mathbf{m} , da die Bestimmung des Menüindex bereits eine Vorentscheidung trifft. Das schlechtere Abschneiden liegt in der Art der Intensionsmodellierung und der Auswertung begründet. Die vorliegende Modellierung fasst in Modell I_{16} alle Funktionstasten zusammen, die

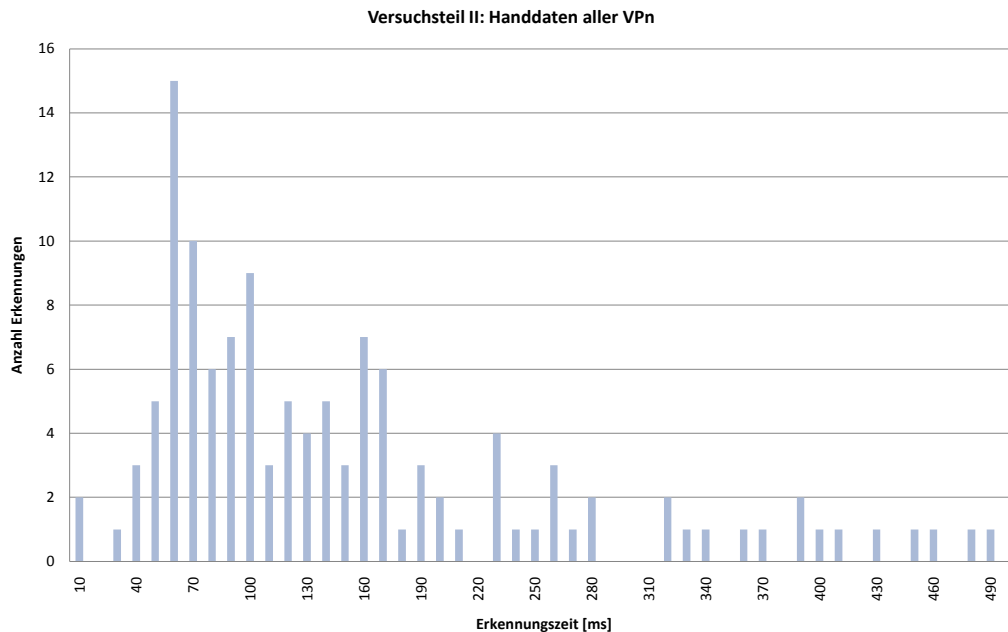


Abbildung 5.29: Versuchsteil II, Durchschnittswerte der Handdaten aller Versuchspersonen. Hier dargestellt der Erkennungsbereich von 10 ms bis 490 ms vor Durchführung der eigentlichen Aktion. Im restlichen Zeitbereich sind keine signifikanten Erkennungsanzahlen vorhanden.

Rücksprünge im Menü entsprechen. Im Versuch wurde allerdings der Menüindex zunächst direkt aus der Menüanwendung übernommen, anstatt ihn aus dem Merkmalsvektor zu berechnen, weshalb die Berechnung der Fahrerintention teilweise einen falschen Menüindex heranzog. Berechnet man den Menüindex aus der IP, zeigt sich, dass das System in allen Fällen die richtige Fahrerintention schätzt. Die Trefferquote steigt damit in VT I auf 82,1% und in VT II auf 91,0% und liegt in VT I über der Trefferquote der HT (80,3%). In VT II kann aus allen gültigen Daten die richtigen Fahrerintentionen bestimmt werden (siehe Tabelle 5.5). Dieses Ergebnis zeigt, dass der Ansatz der IP auf Basis von kombinierten Daten das geeignetere Verfahren darstellt.

Erkennungszeiten Abbildungen 5.28 und 5.29 zeigen die Erkennungszeiten. Für die Handdaten liegt in VT I die durchschnittliche Zeitspanne von Erfassung der Hand (z-Koordinate ≤ 35 mm) bis zur Eingabe bei 136,6 ms. In VT II liegt der Durchschnitt für die Handdaten bei 157,2 ms und für die Blickdaten bei 481,1 ms. Aufgrund der breiten Streuungen und Spitzen sind diese Werte für die Handdaten alleine nicht aussagekräftig. In VT I (vgl. Abbildungen 5.28 und A.1 in Anhang A.1) ergeben sich für

- VP1 deutliche Spitzen bei 40 ms und bei 100 ms,
- VP2 zwei deutliche Spitzen bei 70 ms und 140 ms auf, sowie einen leichten Anstieg zwischen 200 ms und 250 ms,
- VP3 Spitzen bei 30 ms und zwischen 90 ms und 100 ms zu finden,

- VP4 weist eine alles überragende Spitze bei 50 ms und eine weitere bei 130 ms auf,
- VP5 Spitzen zwischen 40 ms und 70 ms sowie bei 160 ms,
- VP6 Spitzen bei 70 ms und 160 ms.

Für jede VP existiert mindestens eine ausgeprägte und sehr steile Spitze und eine flachere Spitze. Zudem sind vereinzelte höhere Zeiten ersichtlich. Die Handdaten (vgl. Abbildungen A.1 und 5.29 in Anhang A.1) weisen ähnliche Spitzen bei jeder VP und vereinzelte höhere Zeiten auf: VP1 bei 60 ms und 150 ms, VP2 bei 160 ms und 260 ms, VP3 zwischen 60 ms und 80 ms sowie bei 130 ms, VP4 bei 60 ms und 170 ms. Außer bei VP2 sind diese Spitzen unwesentlich später angesiedelt, als in VT I.

Die Spitzen und vereinzelte höhere Zeiten sind in der Menüstruktur begründet. Neben den sechs Funktionstasten zur Menüanwahl existieren zwei Menüs mit alphanumerischen Eingaben (Telefonnummern- und Zieleingabe). Im Telefonmenü ergeben sich kurze Wege. Ist die Hand bereits am Touchscreen, ergeben sich kurze Erkennungszeiten – die erste Spitze. Weitere Spitzen resultieren aus der durchschnittlichen Bedienzeit beim Wechsel von Menüs und der Bedienung der Bildschirmtastatur (OSK). Vereinzelte höhere Zeiten ergeben sich aus den längeren Wegen beim OSK, wenn sich die Hand am Bildschirm entlang über verschiedene Tasten hinweg bis zur gewünschten Taste bewegt. Eine mittlere Erkennungszeit auf Handdaten lässt sich im vorliegenden Fall für bekannte, einstudierte Handlungen und wenn die Hand bereits in der Nähe des Touchscreens ist mit ca. 40 bis 70 ms angeben. Für normale Bedienvorgänge ergeben sich 130 bis 160 ms.

Von 394 erkannten Ereignissen mit Blickdaten sind 353 gültig. Von diesen drückte die VP bei 109 Ereignissen nach Betrachten einer Funktionstaste diese auch. In VT II existieren 15 Fälle, bei denen ein Drücken der Funktionstasten ohne vorherige Fixation erfolgt. Dies deckt sich mit den Studien von Land [136, 135]. Aus den 109 Ereignissen lässt sich der Zeitraum (Erkennungszeit) bestimmen, in dem einem Blick auf eine Funktionstaste die Annäherung der Hand und deren Betätigung folgen. In den Diagrammen sind Bereiche auszumachen, in denen sich die Anzahl an Erkennungen häuft. Allerdings sind diese weniger ausgeprägt.

Bei VP1 sind im vollen Bereich von 120 ms bis 890 ms gültige Blicke zu verzeichnen. Für VP2 liegen diese zwischen 230 ms und 830 ms. Bei etwa 400 ms ist eine schwach ausgeprägte Spitze zu erkennen. Für VP3 bzw. VP4 zeigen sich Häufungen bei 160 ms und zwischen 250 ms und 270 ms bzw. zwischen 130 ms und 290 ms. Aus diesen Daten lässt sich eine durchschnittliche Erkennungszeit von 481,2 ms berechnen. Die übrigen 244 Events entsprechen hauptsächlich (über 80%) Blicken, die der Orientierung oder dem Auffinden bestimmter Tasten dienen. Weiterhin existieren Blickverläufe, die das Lesen ganzer Schriftzüge widerspiegeln – vor allem bei den Text- und Auswahlfeldern im Adressbuch, Telefonbuch und der Senderübersicht. Sehr wenige Ereignisse sind dem Kontrollieren von Eingaben oder der Führung der Hand zuzurechnen.

5.6.3.2 Fazit

Die intentionsbasierte Planerkennung stellt einen geeigneten Interpretationsprozess zur Erkennung der Fahrerintention dar. Die Versuchsergebnisse zeigen, dass sich eine Erkennungsrate von 91,0% bei der Verwendung von Hand- und Blickdaten im verwendeten

Setup möglich ist. Dies entspricht einer richtigen Interpretation aller erkannten Ereignisse. Aufgrund der Einschränkung, dass alle Aktionen sequentiell durchzuführen sind, erkennt das System bei Überschneidungen von Aktionen nur die wahrscheinlichste Intention.

Die Erkennungszeit auf Basis von Handdaten liegt im Bereich von 130 ms bis 160 ms vor der eigentlichen Aktion. Befindet sich die Hand bereits am Touchscreen, sinkt die Erkennungszeit auf 40 ms bis 70 ms. Auf Basis von Blickdaten wurde eine durchschnittliche Erkennungszeit von etwa 480 ms ermittelt. Eine weitere Verbesserung der Erkennungsrate und der Erkennungszeiten ist durch die Verbesserung der Datenerfassung möglich. Das folgende Kapitel 5.6.4 stellt stellvertretend eine Anwendung vor, die die Ergebnisse der IP zur dynamischen Anpassung des Interaktionssystems nutzt.

5.6.4 Versuchsergebnisse: Intentionsbasierte Parametrierung des HMI

Die Evaluierung des vorgestellten Systems zur Erkennung von Nutzerintentionen erfolgte im Rahmen einer Studie als Teilkomponente. Das System wurde erweitert, in das Gesamtsystem integriert und anhand von zwei (Teil-)Aufgaben untersucht. Die Eignung von Nutzerintentionen zur Parametrierung des HMI konnte an einem exemplarischen Usecase belegt werden. Die VPn mussten dazu eine Zieleingabe (München, Arcisstrasse) als Aufgabe bearbeiten. Dabei waren sie angehalten, die Eingabe auf dem Touchdisplay (vgl. Abbildung 5.2) zu starten. Sobald das System den Bedienwunsch *Zieleingabe* erkennt und verifiziert, schlug das System dem Nutzer eine für diese Aufgabe passendere Eingabemodalität vor. Im vorliegenden Fall stellte das System dem Nutzer per TTS¹⁴-Engine die Rückfrage „Zu welchem Ziel möchten Sie navigieren?“, um dem Nutzer die Spracheingabe¹⁵ als adäquate Eingabemethode vorzuschlagen. Die VP sollte einen systemgetriebenen Wechsel bei der Eingabe eines Zielortes vollziehen. An den Versuchen nahmen 36 VPn (elf weiblich) im Alter von 35 bis 55 Jahren (durchschnittliches Alter 42 Jahren) und einer durchschnittlichen Fahrerfahrung von ca. 17300 Kilometer pro Jahr teil.

5.6.4.1 Versuchsergebnisse

Ziel der Tests war die Bewertung, ob die VPn einen systemgetriebenen Modalitätenwechsel erfassen und korrekt darauf reagieren *H1* und ob sie ein vorausschauendes Systemverhalten akzeptieren *H2*. Dazu erfolgte die Protokollierung des Verhaltens der VPn im jeweiligen Dialogschritt. Spezielles Augenmerk lag darauf, welche VPn die angebotene Modalität angenommen und beibehalten haben. Abschließend erfolgte eine Befragung der Probanden dahingehend, ob sie ein vorausschauendes Verhalten des Systems akzeptieren und wie verständlich das Angebot der zusätzlichen Modalität war.

Die Bewertung von *H1* stützt sich auf die Auswertung des Verhaltens im jeweiligen Dialogschritt. Im Rahmen der gestellten Aufgaben haben alle VPn den angebotenen Wechsel angenommen, weshalb keine numerische Aussage zur Signifikanz vorliegt. Aufgrund dieser Ergebnisse ist *H1* für die vorliegende Stichprobe als belegt zu werten.

Zur Bewertung von *H2* erfolgte eine Befragung, ob das Angebot der zusätzlichen Modalität verständlich war. Hierzu konnten die VPn auf einer Skala von 1 (sehr verständlich) bis 5 (nicht verständlich) eine Note vergeben. Weiterhin konnten sie bewerten,

¹⁴Text-to-Speech, engl. für Vorleseautomat. Eingesetzt zur Sprachsynthese.

¹⁵In diesem Fall als idealer Spracherkennung durch den Versuchsleiter realisiert.

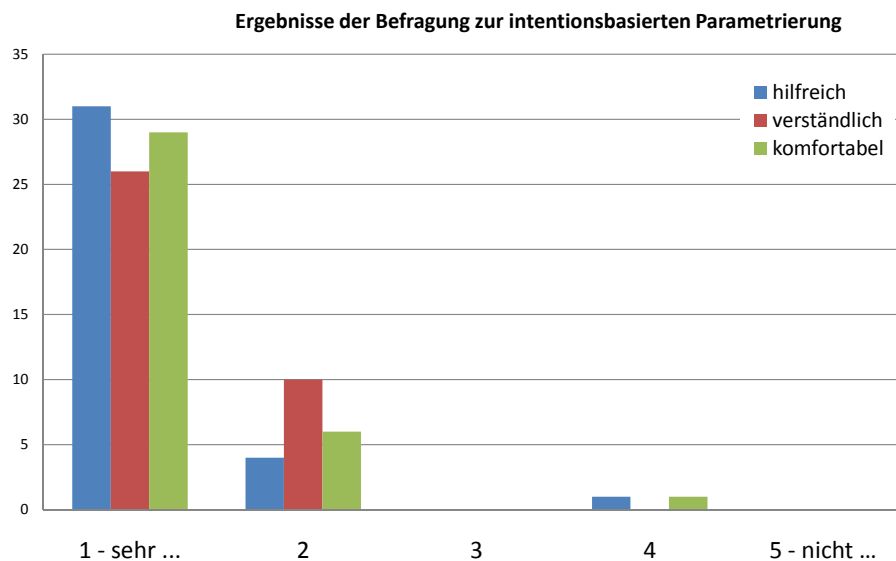


Abbildung 5.30: Bewertung des Systems hinsichtlich wie hilfreich (blau), verständlich (rot) und komfortabel (grün) das Angebot einer anderen Eingabemodalität auf die Probanden gewirkt hat.

wie hilfreich das Angebot der zusätzlichen Modalität war (Bewertungsskala: 1 $\hat{=}$ sehr hilfreich, 5 $\hat{=}$ nicht hilfreich) und wie komfortabel der Wechsel für diese Aufgabe war (Bewertungsskala: 1 $\hat{=}$ sehr komfortabel, 5 $\hat{=}$ nicht komfortabel).

Der angebotene Wechsel erhielt durchweg sehr positiv Bewertungen (vgl. Abbildung 5.30). 31 Probanden bewerten das Angebot einer geeigneteren Eingabemodalität als sehr hilfreich (Note 1) und vier Personen vergaben die Note 2. Ein T-Test ergab bei einem Signifikanzniveau von 95% ein Konfidenzintervall von 78%-100%. Das heißt, mit einer Sicherheit von 95% empfinden mindestens 78% der Vergleichsgruppe den Modalitätenwechsel als hilfreich bis sehr hilfreich. Den Komfort bewerten die Probanden durchschnittlich mit der Note 1,25 (29 VPn Note 1, 6 VPn Note 2 und 1 VP Note 4). Bei einem Signifikanzniveau von 95% ergab der T-Test ein Konfidenzintervall von 78%-100%. Das heißt, mit einer Sicherheit von 95% empfinden mindestens 78% der Vergleichsgruppe den Modalitätenwechsel als sehr komfortabel oder komfortabel. Weiterhin zeigt sich eine sehr gute Bewertung der Verständlichkeit des Modalitätenwechsels. Alle VPn bewerten die Verständlichkeit mit Note 1 (26 VPn) oder 2 (10 VPn). Somit ist H_2 für die vorliegenden Vergleichsprobe als belegt zu werten.

5.6.4.2 Fazit

Einen vom System initiierten und für den Probanden nachvollziehbaren Modalitätenwechsel nehmen die VPn sehr gut an. Erhält der Benutzer vom System das Angebot einer zusätzlichen Eingabeart, nimmt er dieses in der Regel an. Bei einem Wechsel zum Sprachdialog während der Zieleingabe wechseln in der vorliegenden Stichprobe alle VPn die Eingabeart. Der Nutzer wünscht und akzeptiert ein vorausschauendes Systemverhalten.

Weitere Möglichkeiten der Integration eines intentionsbasierten Ansatzes stellen direkte Anpassungen der graphischen Oberflächen an erkannte Bedienwünsche des Fahrers dar. Im Fall der Verwendung von Touch-Interaktion kann z.B. eine automatische Anpassung der Icon-Größe vorgenommen werden. Erkennt das System einen Bedienwunsch des Fahrers, können die entsprechenden Schaltflächen in der GUI vergrößert werden um ein einfaches und sicheres Treffen zu ermöglichen. Blickt der Fahrer öfter für „längere“ Zeit auf eine Bedienelementegruppe (z.B. Heizungssteuerung), kann das graphische Menü in den entsprechenden Menüpunkt (z.B. Klimatisierung) verfahren werden.

6

Blickbasierte Interaktion im Kraftfahrzeug

Die in Kapitel 3 vorgestellten Grundlagen und Systeme zur berührungslosen Bestimmung der Blickrichtung sollen im folgenden Kapitel auf die multimodale, blickbasierte Bedienung von Mensch-Maschine Schnittstellen im Kraftfahrzeug angewendet werden.

Ziel des vorzustellenden Konzeptes ist es, die Ablenkung des Fahrers während der Interaktion mittels intuitiver Menüzugriffe und kürzerer Bedienzeiten zu reduzieren. Das Konzept greift den Trend hin zu großflächigen Multi-Display Anzeige-Bedienkonzepten im Fahrzeug auf und erweitert ihn konsequent. Dazu kombiniert es Blick- und Touchinteraktion sowie aus der Desktop Domäne bekannte Interaktionsstile in einer intuitiven Art und Weise. Bei der Konzeption erfuhren folgende Aspekte besondere Beachtung:

- Vermeidung von Verdeckungen der Applikationen untereinander
- Gewährleistung des Zugriffs auf einzelne Applikationen und eine schnelle Rückkehr in den Ausgangszustand
- einfache Skalierbarkeit der einzelnen Applikation auf die gewünschte Größe
- konsistente Darstellung der Applikationen auf unterschiedlichen Displays und beim Verschieben auf ein anderes Display

Die folgenden Ausführungen stellen ein an diese Anforderungen angepasstes Anzeige-Bedienkonzept vor, das in seiner individuellen Gestaltung einen großen Beitrag dazu leistet, die Gefahr der Ablenkung des Fahrers während der Bedienung von Funktionen möglichst gering zu halten [193, 222].

6.1 Konzept

Grundlage des Interaktionssystems ist das in Kapitel 5.1 vorgestellte Basisframework *ABK2020*. Das Konzept nutzt eine applikations- bzw. fensterbasierten Organisation einzelner Menüteile, die deren separate Skalierung sowie freie Positionierung auf den Displays erlaubt. Diese Bedienphilosophie ist bekannt vom Umgang mit Desktopcomputern. Hier stellt die Fensterdarstellung seit langem die favorisierte graphische Informationsorganisation dar. Das Starten von Applikationen erfolgt über entsprechende Icons, wobei sich ein

6. Blickbasierte Interaktion im Kraftfahrzeug

neues Fenster öffnet. Durch Schließen des Fensters werden aktive Applikationen beendet. Einzelne Fenster können in beliebiger Reihenfolge geöffnet und geschlossen werden. Zur Reduktion der Darstellungskomplexität können Fenster jedoch im Gegensatz zu Heim-PCs nicht völlig frei angeordnet, sondern nur an vorgegebene Positionen bewegt und skaliert werden. Weiterhin wird verhindert, dass sich einzelne Fenster überlappen bzw. einander verdecken. Die Verwaltung der Fenster im FPK und HUD geschieht über das Hauptmenü im CID. Somit sollte der Umgang mit der graphischen Oberfläche als intuitiv und leicht bedienbar empfunden werden.

Zur Bedienung stehen dem Fahrer – abhängig von der Anzeigefläche – multimodale Interaktionsstile zur Verfügung: Applikationen auf dem CID können per Sprache, Touch- und Blickbedienung gesteuert werden, Applikationen auf dem HUD und dem FPK werden per Blick bedient. Die Blickbedienung beruht dabei auf einer Kombination aus Fixation zur Selektion der gewünschten Schaltfläche und der Bestätigung der Auswahl über eine Taste am Lenkrad, wodurch der Midas-Touch Effekt umgangen wird. Als Feedback für den Nutzer wird die fixierte Schaltfläche (bzw. die Schaltfläche unter dem Finger¹) durch einen Farbwechsel des Rahmens von den restlichen Schaltflächen hervorgehoben (vgl. grüne Schaltfläche in Abbildung 6.1(a)). Wird eine Schaltfläche ausgewählt, färbt sich die gesamte Schaltfläche als Rückmeldung ein, bevor die Applikation geöffnet bzw. die Funktion ausgeführt wird (vgl. grünes Telefonsymbol in Abbildung 6.1(b)).



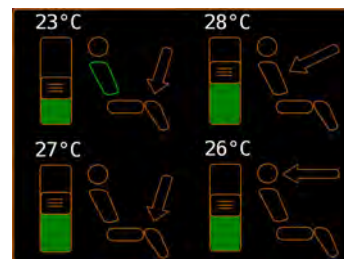
(a) Navigation Hauptmenü.



(b) Telefonmenü.



(c) Tempomat/ACC.



(d) Klimaautomatik.

Abbildung 6.1: Exemplarische Darstellung der graphischen Oberflächen (Navigation, Telefon, Abstandstempomat, Klimabedienung) des blickbasierten Interaktionssystems. Die Fenster können auf den einzelnen Anzeigeflächen dargestellt werden.

¹Die Annäherung des Fingers an eine Taste wird über den Touchscreen erfasst (vgl. Kapitel 5.6).

6.1.1 Hauptmenü

Das Menü im CID dient der Kontrolle des Multi-Display Systems und wird hauptsächlich per Touch bedient. Im CID werden die einzelnen Applikationen gestartet und auf die

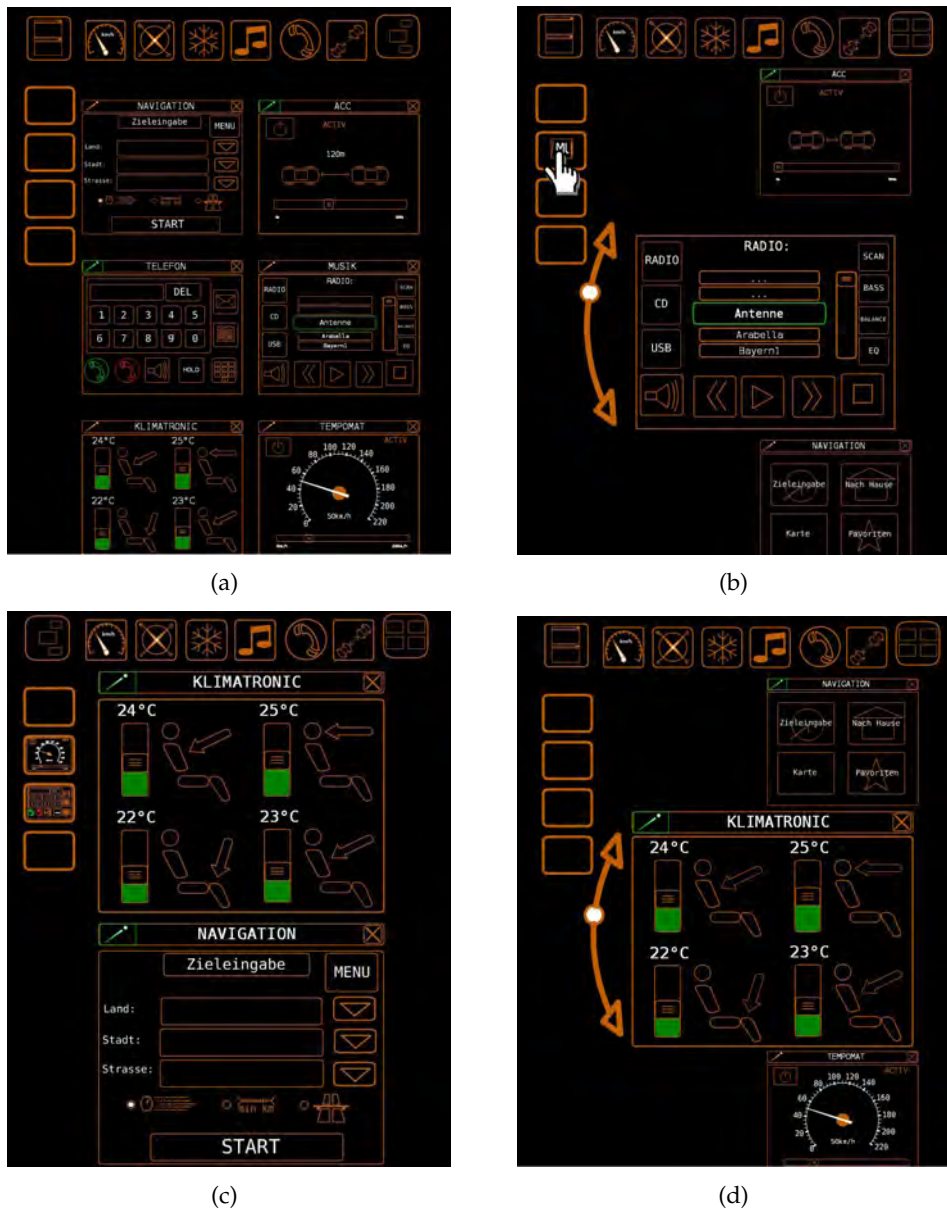


Abbildung 6.2: Graphische Oberflächen des Hauptmenüs auf dem CID.

verschiedenen Anzeigerteile im Cockpit verteilt. Wird eine Funktion nicht mehr benötigt, kann sie über das Menü im CID beendet werden. Weiterhin kontrolliert das CID die Kommunikation und Koordination zwischen allen Fenstern und Anzeigerteilen. Abbildung 6.2 zeigt den Aufbau des Hauptmenüs: (1) am oberen Rand befindet sich die Schnellstartleis-

te, (2) das Applikationsfenster und (3) am linken Rand ist der Konfigurationsbereich zur Verteilung der Fenster auf die Anzeigorte.

Die zwei äußersten Tasten der Schnellstartleiste dienen der Anpassung der Fensteranordnung im Applikationsbereich (vgl. Abbildung 6.2(a) und 6.2(b)), die restlichen zum Starten der einzelnen Applikationen per Direktzugriff. Ist das Fenster bereits geöffnet, aber aktuell minimiert oder in einem anderen Display geladen, wird es automatisch für die aktuelle Fensterkonstellation bedienbereit dargestellt und markiert.

Das Applikationsfenster stellt die aktiven Fenster dar, wobei die Fenster der einzelnen Applikationen (vgl. Abbildungen 6.1 und 6.5) in unterschiedlichen Konstellationen angeordnet sein können: alle aktiven Fenster in Matrixanordnung (Abbildung 6.2(a)), in Form einer vertikalen scrollbaren Liste (Abbildungen 6.2(b) und 6.2(d)) oder zwei festgepinnte Applikationsfenster (Abbildung 6.2(c)). Die Anordnung lässt sich über die äußeren Tasten der Schnellstartleiste wechseln und Reihenfolge bzw. Größe der Fenster lassen sich per Drag & Drop bzw. durch aufziehen der Fenster anpassen.

Die vier Rechtecke im dritten Bereich symbolisieren die Anzeigeflächen im HUD (oberstes Rechteck) sowie im FPK. Aus der Schnellstartleiste können Funktionen einzeln auf die Displays gezogen werden. Zieht der Fahrer ein Icon auf ein Rechteck und lässt es los, zoomt es in dieses hinein und taucht im entsprechenden Display auf. Auf diesem Wege erfolgt die freie Belegung der Anzeigeflächen mit Applikationen. Bei der Verteilung der Fenster auf die Anzeigeflächen wird als Erstes überprüft, ob der Zielbildschirm frei oder bereits durch eine andere Applikation belegt ist. Ist die Anzeigefläche verfügbar, wird die entsprechende Applikation geladen. Wird bereits eine andere Applikation dargestellt, muss diese auf ein freies Display verschoben werden. Dabei kann eine Verschiebung der Applikationen vom HUD ins FPK erfolgen, aber nicht umgekehrt. Ist keine freie Anzeigefläche verfügbar, erfolgt eine Verschiebung der Applikation ins CID. Gleichzeitig findet ein Statusabgleich der Applikation statt, um diese im selben Zustand auf der neuen Anzeigefläche zu starten und die alte Instanz zu löschen. Wird z.B. die Anzeige des ACC vom HUD ins FPK verschoben, müssen dieser die aktuellen Werte für Wunschgeschwindigkeit und Abstand zum Vordermann übergeben werden.

Zusätzlich kann per Drag & Drop die Position der Applikationen auf den Displays geändert oder getauscht werden. Verschiebt der Fahrer eine Applikation aus einem Rechteck und auf die freie Fläche im CID, wird sie vom aktuellen Display ins CID verschoben. Durch drücken der X-Schaltfläche oben rechts im Fenster wird die Applikation beendet. Somit kann der Fahrer als störend empfundene Anzeigen aus seinem Sichtfeld verlagern bzw. ausblenden. Die Synchronisation dieser Vorgänge wird vom CID durchgeführt. Außerdem verteilt das CID als Datenserver entsprechende Parameter im Gesamtsystem.

6.1.2 Applikationen

Applikationen stellen die bedienbaren Funktionalitäten der einzelnen Fahrerassistenz- und Infotainmentsysteme² bereit und sind aus verschiedenen rein anzeigenden sowie interaktiven Elementen aufgebaut. Die Applikationen können auf jeder Anzeigefläche

²Im Rahmen der vorliegenden Arbeit wurden Oberflächen für Tempomat/Adaptive Cruise Control, Navigation, Klimaautomatik, Musik und Telefon umgesetzt.

„gestartet“ und bedient werden, was den Vergleich der Interaktionsstile (konventionelle Bedienung per Touch vs. multimodale blickbasierte Bedienung) ermöglicht.

Die Oberfläche des *Tempomat/ACC* dient zur Parametrierung der analogen Werte für Wunschgeschwindigkeit und gewünschtem Abstand zum Vorderfahrzeug. Die niedrige Komplexität der Aufgabe ermöglicht die Verwendung von großen Schaltflächen sowie redundanter Eingabearten. Abbildung 6.3 zeigt die graphischen Oberflächen. Bei Touch-



(a) Wunschgeschwindigkeit.

(b) Wunschabstand.

Abbildung 6.3: Graphische Oberfläche Tempomat/ACC.

Bedienung (nur auf dem CID möglich) werden Geschwindigkeit bzw. Abstand durch horizontales Ziehen des Balkens (oder durch das Tippen auf die Zahl auf der analogen Tachometerskala) gesetzt. Eine blickbasierte Manipulation eines solchen Parameters folgt dabei stets dem gleichen Muster:

1. Selektion der gewünschten Schaltfläche durch Fixation
2. Einstellen des gewünschten Wertes durch
 - (a) Drehen des Scrollbalkens am Lenkrad oder
 - (b) Fixation des gewünschten Wertes
3. Bestätigung der Einstellung durch Drücken des Scrollbalkens am Lenkrad

Die Applikation *Navigation* beinhaltet eine deutlich höhere Komplexität der Bedienung und deckt simple Einstellungen (z.B. Anwahl der Berechnung einer kurzen, schnellen oder benzinsparenden Route) bis hin zu komplexen alphanumerischen Eingaben (vgl. Kapitel 6.1.3) ab. Die Grundstruktur der Oberfläche besteht aus einer 2x2 Matrixstruktur (vgl. Abbildung 6.1(a)). Die eigentliche Zieleingabe kann über die Eingabe der detaillierten Zielinformationen (siehe dazu Kapitel 6.1.3) oder aber über die Auswahl eines Ziels aus den Favoriten erfolgen. Dazu existiert eine scrollbare Liste der meisteingegebenen Ziele (vgl. Abbildung 6.4(a)). Das Scrollen erfolgt analog zu obiger Manipulation der Geschwindigkeitseinstellung, welches per Touch über den Slider am Rand realisiert ist. Bei Blickinteraktion wird der Eingabefokus durch Fixieren auf die scrollbare Liste gelegt und durch Drehen des Scrollbalkens am Lenkrad wird durch die Liste navigiert. Dabei wandert die Liste jeweils in die entsprechende Richtung und das aktuell ausgewählte Element wird durch die hervorgehobene Markierung (in grün und größer dargestellt) visualisiert. Ein Druck auf den Scrollbalken/den Touchscreen wählt dann den markierten



Abbildung 6.4: Graphische Oberfläche Navigationsmenü.

Eintrag aus. Die Anwahl der weiteren in Abbildung 6.4 dargestellten Tasten geschieht in ähnlicher Art und Weise: Drücken der Tasten auf dem Touchscreen oder Fixation der Tasten und Drücken des Scrollrades am Lenkrad.

Abbildung 6.1(d) zeigt die Oberfläche der *Klimaautomatik*, welche in vier Segmente unterteilt ist und die Einstellung der Temperierung der einzelnen Zonen ermöglicht. Dabei kann man für jede Zone die Temperatur und die Luftausströmung (Kopf, Oberkörper, Beine) anpassen. Die Einstellung der Temperatur per Blick bzw. Touch erfolgt analog obiger Vorgehensweise, zur Anpassung der Luftausströmung kann die gewünschte Körperregion über Fixation und Scrollrad oder Touch selektiert werden. Die Applikationen *Entertainment* und *Kommunikation* sind analog aufgebaut.

6.1.3 Blickbasierte Alphanumerikeingabe

Zur Untersuchung des Potentials einer blickbasierten Alphanumerikeingabe im Kraftfahrzeug wurden die Konzepte *Tastaturfeld*, *Speller*, *MatrixKey* und *Kreisstruktur* entworfen und untersucht. Die Texteingabe ist ausschließlich für das HUD konzipiert, da die restlichen Displays für diese Aufgabe a-priori zu hohe Blickabwendung bedeuteten.

Das *Tastaturfeld* ordnet die Buchstaben in Form einer Bildschirmtastatur mit alphabetischer Reihenfolge an. In jeder Zeile befinden sich sieben Tastenfelder (Abbildung 6.5(a)), wodurch sich eine größtmögliche Buttonfläche von $1,2^\circ$ (Blickwinkel) ergibt. Fixiert der Nutzer einen Buchstaben, wird dieser (als Feedback) grün markiert und bei anschließendem Drücken der Bestätigungstaste am Lenkrad (während der Blick auf dem Buchstaben verweilt) in das Textfeld am oberen Rand eingetragen.

Die *Speller* Anordnung (vgl. Abbildung 6.5(b)) besteht aus sieben Buttons: Die unteren Tasten dienen dem Direktzugriff auf Leerzeichen und Löschen, die restlichen Tasten der Auswahl und Eingabe von Buchstaben (Größe $1,4^\circ$). Das mittlere Feld markiert den aktuell selektierten Buchstaben, der bei Betätigung der Lenkradtaste übernommen wird. Die restlichen vier Buttons dienen der Navigation durch die Buchstabenliste.³

Das Konzept *MatrixKey* (vgl. Abbildungen 6.5(c) und 6.5(d)) besteht aus vier quadratischen Feldern, die jeweils mit neun Buchstaben- und Zahlenfeldern belegt sind. Da im

³Der linke Button verschiebt die Buchstabenliste um zwei Positionen nach rechts ('A'→'Y', 'B'→'Z', etc.) und der Button halb links verschiebt die Liste um eins nach rechts ('A'→'Z', 'B'→'A', etc.). Die rechten Buttons verschieben die Liste in gleicher Weise nach links.

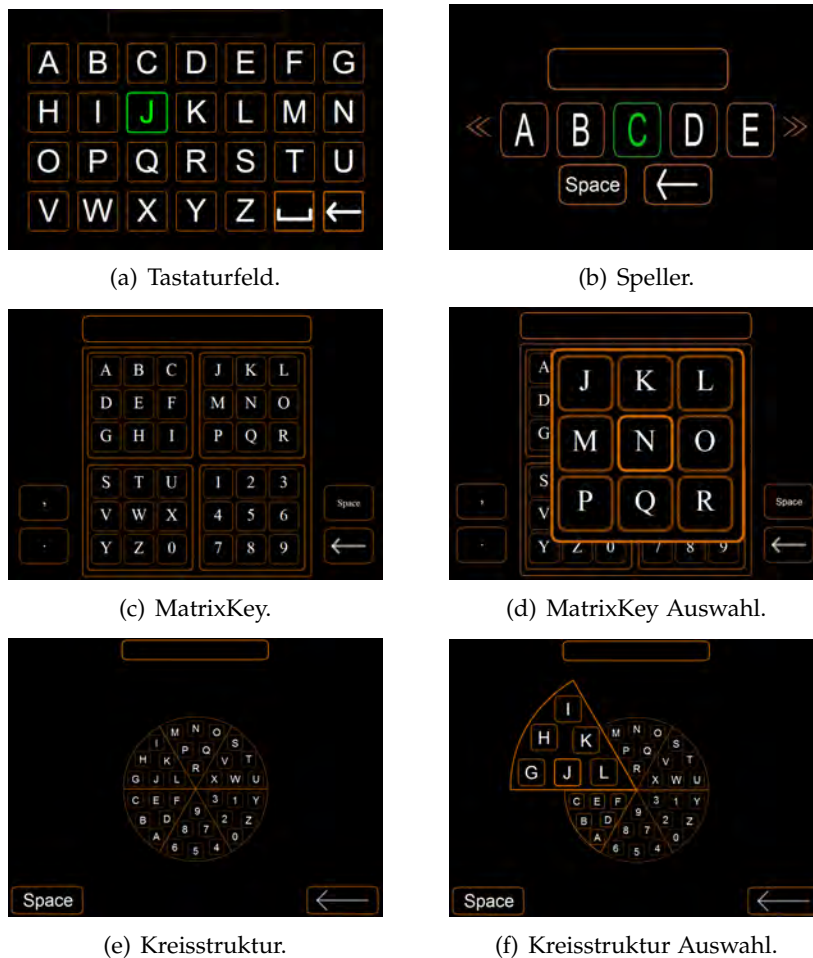


Abbildung 6.5: Graphische Oberflächen der blickbasierten Alphanumerikeingaben.

Fahrzeug nur eine begrenzte Displayfläche zur Verfügung steht, bedient sich das Konzept eines Zoomeffekts, was eine Minimierung der benötigten Fläche erlaubt. Verweilt der Blick auf einem der vier Felder (dwell-time: 200 ms), vergrößert sich dieses um den Faktor 1,8 (vgl. Abbildung 6.5(d), Buttongröße $1,6^\circ$). Verweilt der Blick auf einem Buchstaben im vergrößerten Feld, wird dieser zunächst farblich hervorgehoben und bei Bestätigung übernommen.

Die *Kreisstruktur* (vgl. Abbildungen 6.5(e) und 6.5(f)) benutzt ebenfalls die Zooming Strategie, um größtmögliche Tastengröße zu erreichen. Die Buchstabenfelder sind auf sechs „Kuchenstücken“⁴ angeordnet. Auch hier wurden als dwell-time für die Vergrößerung 200 ms verwendet. Die Größe eines Buttons im vergrößerten Zustand beträgt $0,8^\circ$. Aufgrund des schlechten Abschneidens im Rahmen einer Vorstudie (Buttongröße trotz Zooming zu gering) soll auf eine eingehende Evaluierung im Folgenden verzichtet werden.

⁴Im Bereich des Eye Typing sind in dieser Form angeordnete pEYE Menüs sehr beliebt, vgl. [112].

6.2 Versuchsergebnisse

Die vorgestellten Konzepte zur blickbasierten Menübedienung und Alphanumerikeingabe wurden im Rahmen von zwei Probandenversuchen auf ihre Verwendbarkeit im Fahrzeugkontext getestet. Zu Beginn der Versuche wurden die Probanden in das getestete System eingewiesen und konnten die Funktionalitäten testen.



Abbildung 6.6: Versuchsaufbau.

6.2.1 Versuchsergebnisse – Alphanumerik

Der Versuch zur blickbasierten Texteingabe bestand aus zwei Teilen: (1) Eingabe von Straßennamen im Stand, wodurch die Probanden den Umgang und die Bedienung der Konzepte erlernten. Die Zeiten dienen gleichfalls als Referenz. Hierbei mussten die Probanden folgende Straßennamen eingeben: LEOPOLDSTR, GEORGENSTR und EINSTEINSTR. (2) Eingabe von Städtenamen während der Fahrt (Lane Change Task): AUGSBURG, ROSENHEIM und DORTMUND. Während der gültigen Testfahrt fuhren die VPn zunächst einen Abschnitt ohne Texteingabe als Referenzfahrt. Zur Bewertung der Ablenkung wurden die Spurabweichung und zusätzlich die Ergebnisse der PDT ausgewertet. Um die Belastung der Probanden nicht zusätzlich zu verstärken, wurde auf die Korrektur von Rechtschreibfehlern während den Versuchen verzichtet. Die Texteingabe wurde in solchen Fällen mit dem nächsten Buchstaben fortgesetzt.

An den Versuchen nahmen 14 VPn (elf männlich) mit einem durchschnittlichen Alter von 25,6 Jahren teil. Acht Teilnehmer gaben eine Fahrleistung von weniger als 10.000km, vier weniger als 25.000km und zwei mehr als 25.000km im Jahr an. Der Begriff Eye Tracker war den meisten Teilnehmern bekannt. Fünf hatten bereits Erfahrungen im Umgang, während nur zwei noch nie davon gehört hatten. Ähnliches gilt für das HUD: Außer einer VP hatten alle davon gehört, wobei nur eine Person bereits ein

Fahrzeug mit HUD gefahren war. Alle Testpersonen hatten Erfahrung im Umgang mit tragbaren Navigationssystemen und zwei Erfahrung im Umgang mit integrierten Navigationssystemen, was darauf schließen lässt, dass die meisten bereits erste Erfahrung im Umgang mit Texteingabe während der Fahrt besaßen.

6.2.1.1 Objektive Daten

Abbildung 6.7 zeigt die mittleren Eingabezeiten mit und ohne Fahraufgabe. Ein Vergleich zeigt, dass das *Tastaturfeld* die kürzeste mittlere Eingabedauer erreicht, während *MatrixKey* (ohne Fahraufgabe) und *Speller* (mit Fahraufgabe) die längsten Eingabezeiten benötigten. Sowohl das *Tastaturfeld* als auch der *Speller* weisen einen Anstieg der Eingabezeiten mit Fahraufgabe auf, wohingegen *MatrixKey* einen Rückgang der mittleren Eingabedauer aufweist, was auf einen Lerneffekt schließen lässt.

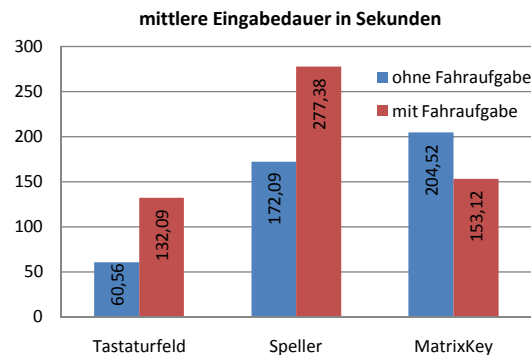
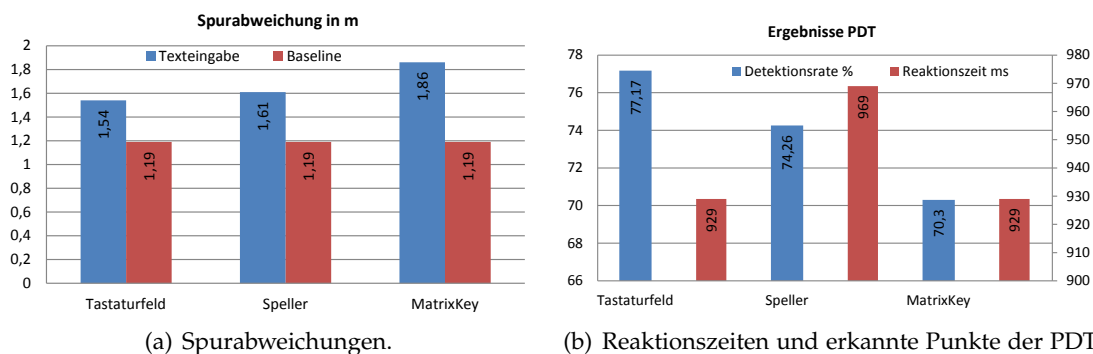


Abbildung 6.7: Übersicht der mittleren Eingabezeiten mit (rot) und ohne Fahraufgabe (blau).



(a) Spurabweichungen.

(b) Reaktionszeiten und erkannte Punkte der PDT.

Abbildung 6.8: Übersicht der mittleren Spurabweichungen sowie die Ergebnisse der PDT.

Abbildung 6.8 zeigt die Ergebnisse der einzelnen Texteingabestrategien hinsichtlich der Ablenkung von der Fahraufgabe: Ohne Alphanumerikeingaben erreichten die Probanden eine durchschnittliche Spurabweichung von 1,19 m und eine durchschnittliche Detektionsrate der PDT Punkte von 94,9%. Mit Texteingabe erhöhte sich die mittlere Spurabweichung auf Werte von bis zu 1,86 m beim Konzept *MatrixKey* (vgl. Abbildung

6.8(a)). Die Anzahl der bestätigten PDT-Punkte zeigt ebenfalls die geringer werdende Aufmerksamkeit. Der Speller erreichte als bestes Konzept eine Bestätigungsrate von 74,9% (vgl. Abbildung 6.8(b)). Mit Texteingabe ist die mittlere Reaktionszeit von 618,3ms auf minimal 929,2ms beim *MatrixKey* angestiegen (siehe Abbildung 6.8(b)).

6.2.1.2 Subjektive Daten

Zusätzlich wurden subjektive Daten erhoben, um den Eindruck der Probanden in Bezug auf Attraktivität, Erlernbarkeit, Benutzbarkeit, Anstrengung und den Einfluss auf den Straßenverkehr zu bewerten. Die Ergebnisse der freien Befragung sind in Abbildung 6.9 dargestellt. Generell zeigt sich, dass es den meisten Probanden leicht fiel, die Bedienung zu erlernen. Bei der mentalen Anstrengung unterscheiden sich die drei Konzepte jedoch stark von einander: Während hier das Urteil beim *Speller* eher neutral ausfiel, waren die Versuchsteilnehmer beim *Tastaturfeld* – im Gegensatz zum *MatrixKey* – nicht sehr angestrengt. Hinsichtlich der Wahrnehmung der Umgebung zeigt sich, dass alle Konzepte

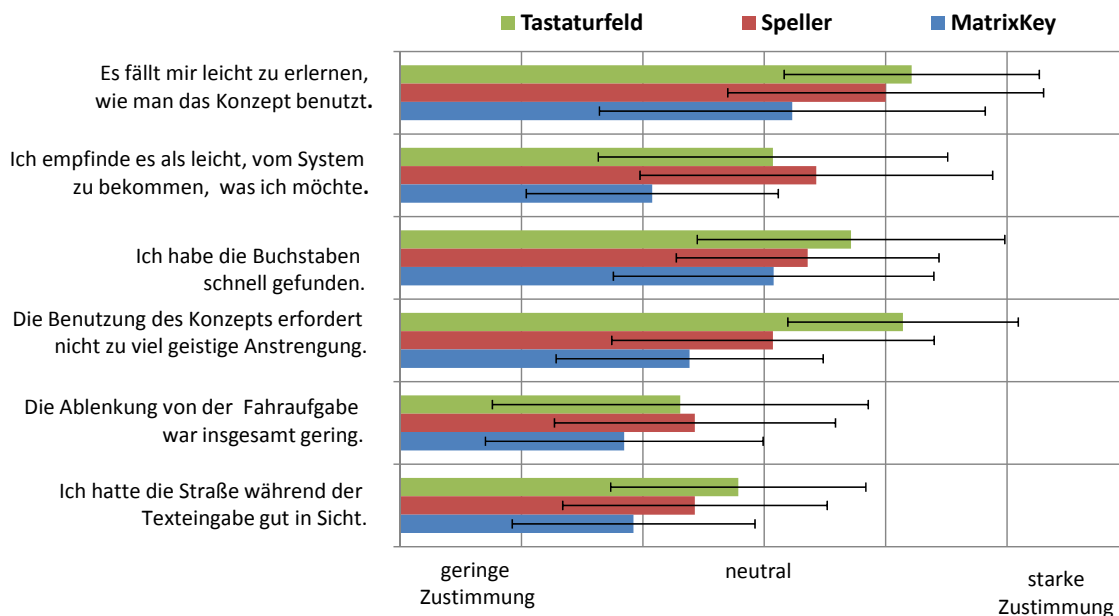


Abbildung 6.9: Umfrageergebnisse zu den Konzepten.

zu einer subjektiven Ablenkung von der Fahraufgabe führen. Die Probanden merkten an, dass es günstiger wäre, das Display höher zu positionieren, um die Straße während der Eingabe besser im Blick zu haben. Vergleicht man die Konzepte untereinander, wird das *Tastaturfeld* gefolgt vom *Speller* subjektiv am besten bewertet, *MatrixKey* hingegen hat bei allen Fragen am schlechtesten abgeschnitten.

Die Mittelwerte der subjektiv erlebten Beanspruchungshöhe (SEA) ergaben für das *Tastaturfeld* mit 118,5 eine geringere Anstrengung als für *Speller* (130,5) und *MatrixKey* (159). Es zeigt sich, dass unter den Teilnehmern eine große Streuung der subjektiven Beanspruchung herrschte. Anzumerken ist, dass es trotz der relativ hohen Mittelwerte zwei Probanden gab, deren Beanspruchung deutlich geringer war.

Zur allgemeinen Bewertung der Attraktivität einer blickbasierten Texteingabe im Fahrzeug wurde der AttrakDiff Fragebogen ausgewertet. Hierbei zeigt sich, dass die blickbasierte Texteingabe den Nutzer überdurchschnittlich anregt und motiviert. Das bewertete Produkt wird als eher selbstorientiert eingestuft. Die blickbasierte Texteingabe erreicht bei der pragmatischen Qualität mittlere Werte. Hinsichtlich Bedienbarkeit besteht also Verbesserungspotenzial. Das große Konfidenzintervall zeigt an, dass die Benutzer das Produkt sehr unterschiedlich bewertet haben. Zum Vergleich der vorgestellten Strategien zur Alphanumerikeingabe mit herkömmlichen Texteingabeverfahren sei auf die ausführlichen Untersuchungen in [7] hingewiesen.

6.2.1.3 Fazit

Die blickbasierte Texteingabe über das Tastaturfeld erzielte in der vorliegenden Untersuchung die besten Ergebnisse, jedoch benötigt dieses Konzept (dauerhaft) die größte Displayfläche. Hinsichtlich der benötigten Anzeigefläche stellt der *Speller* die bessere Alternative dar, sollte jedoch für den Einsatz im Fahrzeug um die Möglichkeit erweitert werden, die Navigation durch die Buchstabenliste ebenfalls mittels des Scrollrades zu steuern (vgl. Listennavigation in Kapitel 6.1.2), dadurch könnte viel Zeit gespart werden. Aufgrund der Versuchsergebnisse ist das Konzept *MatrixKey* als ungeeignet für den Einsatz im Fahrzeug einzustufen. Verglichen mit den Ergebnissen aus [7] bleibt festzuhalten, dass blickbasierte Verfahren nicht mit traditionellen Verfahren konkurrieren können.

6.2.2 Versuchsergebnisse – Menübedienung

Die vorgestellten Konzepte zur blickbasierten Interaktion wurden im Rahmen einer Probandenstudie mit 15 VPn (drei weiblich, durchschnittliches Alter 25,3 Jahre) untersucht. Alle VPn gaben ein hohes Interesse an technischen Neuigkeiten an, zwölf hatten bereits Erfahrungen im Umgang mit einem Touchscreen. Keine VP war bisher ein Fahrzeug mit HUD gefahren, jedoch hatten alle VPn erste Erfahrungen im Umgang mit Menüsystemen gesammelt. Drei Probanden hatten Erfahrung mit Gaze Tracking Systemen, zwei haben ein Gaze Tracking System mehr als einmal benutzt.

Im Rahmen der Studie mussten die Probanden zehn unterschiedliche Aufgaben (Einstellung Wunschabstand/Wunschgeschwindigkeit ACC, Auswahl Routenkriterien, Auswahl Telefonnummer, etc.) sowohl über blickbasierte als auch über klassische Touchinteraktion erfüllen. Der Versuchsablauf bestand aus vier Teilen: Referenzfahrt (nur LCT und PDT), zwei Versuchsteile mit Aufgaben und eine abschließende Referenzfahrt. Die folgenden Abschnitte stellen die Ergebnisse des Versuches vor.

6.2.2.1 Subjektive Daten

Abbildung 6.10 stellt die durchschnittliche Ausprägung der pragmatischen und hedonischen Qualität der getesteten Eingabemodalitäten als Konfidenzrechtecke (Blickbedienungs- bzw. Touchbedienungskonzept in blau bzw. rot) dar. Das Touchkonzept wurde als neutral eingestuft. Die pragmatische und hedonische Qualität des Blickkonzeptes erhielten deutlich bessere Bewertungen. Das Konzept wurde als eher begehrt und sehr

6. Blickbasierte Interaktion im Kraftfahrzeug

hedonisch bewertet. Die VPn können sich sehr gut mit dem System identifizieren und werden von ihm stimuliert/motiviert. Die kleinen Konfidenzrechtecke zeigen an, dass sich die VPn bei den Beurteilungen der Systeme einig waren. Während Touchbedienung auf die Nutzer mittelmäßig attraktiv wirkt, wird Blickbedienen als sehr attraktiv eingestuft. Die Unterschiede in der Bewertung sind durchweg statistisch signifikant (Signifikanzniveau 0.045).



Abbildung 6.10: Portfoliodarstellung des Systems per blickbasierter (blau) und per Touchinteraktion (rot).

Zur Bewertung der Attraktivität kam neben AttrakDiff ein freier Fragebogen zum Einsatz, welcher aus mehreren freien Fragen bestand, die auf 5-fach Likert Skalen (1 $\hat{=}$ trifft nicht zu; 5 $\hat{=}$ trifft voll zu) zu bewerten waren. Abbildung 6.11 zeigt die Ergebnisse der Befragung als BoxWhiskerPlot⁵. Die Probanden bewerteten die Bedienung des Systems per Blick als äußerst intuitiv, benutzerfreundlich und einfach. Der Aussage „Das Konzept Blickbedienung ist besser als eine konventionelle Steuerung“ stimmten die Testpersonen nicht einheitlich zu. Trotzdem gaben über 50% der Befragten an, das Konzept der Blickinteraktion gegenüber der Touchinteraktion zu bevorzugen. Weiterhin bewerteten die Probanden die Fähigkeit zur Überwachung des Verkehrsgeschehens während der Blickbedienung als signifikant besser, verglichen mit der Touchbedienung. Alle Probanden waren sich sicher, dass ihre Fahrleistungen bei Blickinteraktion deutlich besser waren als bei der Touchbedienung.

Die subjektiv erlebte Beanspruchung der Probanden bei der Bedienung des Systems per Blick und Touch wurde mittels der SEA-Skala erhoben. Touchbedienung wird mit großer Übereinstimmung der Probanden mit einem Median von 120 als „ziemlich anstrengend“ eingestuft, die Beanspruchung des Blickbedienens liegt mit einem Median von 70 zwischen „einigermaßen“ und „etwas anstrengend“ und somit wird die Bedienmodalität als weniger anstrengend empfunden, wobei sich die Testpersonen weniger einig waren.

⁵Dieser fasst verschiedene robuste Streuungs- und Lagemaße in einer Darstellung zusammen. Die blaue Fläche gibt den Bereich an, indem die mittleren 50% der Werte liegen. Die Länge der blauen Box entspricht dem Interquartilsabstand, der durch die Differenz des oberen und unteren Quartils bestimmt wird und ein Maß für die Streuung der Werte darstellt. Des Weiteren wird die Verteilung durch die Angabe des Medians (hier als roter Strich) in zwei Hälften geteilt. Die beiden Whisker geben die Lage der äußersten Ausreißer in der Verteilung an.

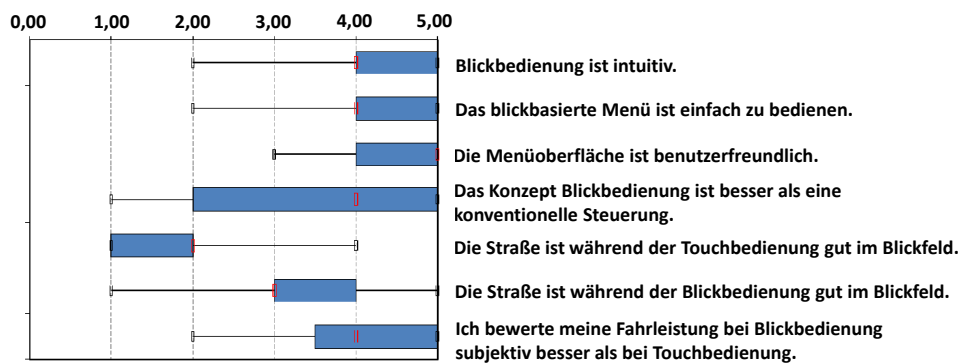


Abbildung 6.11: Bewertung vom Menü und Schätzung eigener Fahrleistung.

6.2.2.2 Objektive Daten

Neben subjektiven Daten wurden zur Beurteilung der Leistung der Probanden unterschiedliche objektive Daten erhoben. Bei der Auswertung der objektiven Daten zeigten sich keine signifikanten Unterschiede zwischen den Geschlechtern, Links- und Rechtshändern oder Personen mit/ohne Sehhilfe. Ausschlaggebend für die Unterschiede zwischen den Probanden war die Erfahrung mit einem Gaze Tracking System. Zwei Probanden, die bereits erste Erfahrungen mit einem Gaze Tracker gesammelt hatten, schnitten deutlich besser ab, als alle anderen. Deswegen wird ihre Leistung (in den Abbildungen jeweils mit einem roten Stern markiert) in der nachfolgenden Diskussion der Ergebnisse explizit dargestellt und diskutiert.

Bedienzeiten Die Komplexität einer Aufgabe hängt u.a. von der Größe des zu fokussierenden Objektes und der Anzahl benötigter Bedienschritte ab. Je kleiner die Schaltfläche ist und je mehr Bedienschritte notwendig sind, desto schwieriger ist eine Aufgabe per Blick einzustufen. Zum Vergleich von Aufgaben mit unterschiedlichen Schwierigkeitsgraden, soll dieser folgendermaßen abgeschätzt werden

$$\text{Schwierigkeitsgrad} = \frac{\text{Eingabeschritte} * \text{Fenstergröße}}{\text{Größe der kleinsten Schaltfläche}}$$

Der Schwierigkeitsgrad berechnet sich dabei aus dem Verhältnis zwischen Fenstergröße und der kleinsten Schaltfläche multipliziert mit den notwendigen Eingabeschritten. So wird z.B. die Auswahl von „Niederlanden“ in der Navigation unter Favoriten als Aufgabe mit Schwierigkeitsgrad 8 eingestuft. Für die Aufgabe sind zwei Eingabeschritte notwendig: Auswahl der Favoriten-Schaltfläche (genau 1/4 der Fenstergröße) und Scrollen der Favoritenliste bis „Niederlanden“. Dies ergibt: $2 \times 4 = 8$. Abbildung 6.12 zeigt die durchschnittlichen Bedienzeiten unterschiedlicher Aufgaben über den Schwierigkeitsgrad aufgetragen. Es zeigt sich eine deutliche Reduktion der Bedienzeiten bei Blickbedienung (durchschnittlich 31%). Je schwieriger eine Aufgabe eingestuft ist, desto kleiner wird der Unterschied zwischen den Modalitäten. Die Linien schneiden sich bei einem Schwierigkeitsgrad von ca. 10. Ab diesem Punkt ist die Eingabe per Touch vorzuziehen. Steigt der Schwierigkeitsgrad weiter an, empfinden die VPn die Eingabe per Blick als anstrengender und ablenkender.

6. Blickbasierte Interaktion im Kraftfahrzeug

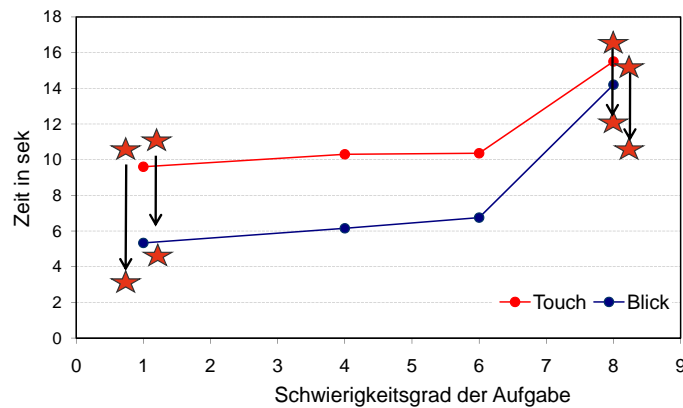


Abbildung 6.12: Bedienzeiten bei Blick- (blau) und Touchinteraktion (rot).

Hierbei sei auf die Zeiten der beiden im Umgang mit Blickbedienung versierten Probanden hingewiesen. Sie konnten ihre Bedienzeiten mit Blickbedienung im Schnitt um 38% verbessern, was auf einen großen Lerneffekt bei der Verwendung dieser zunächst sehr ungewohnten Eingabemodalität zurückzuführen ist. Die meisten Probanden ohne Erfahrung mit einem Gaze Tracker waren zu sehr auf die Fokussierung der Schaltflächen konzentriert, was weiterhin ihre Fahrleistung signifikant negativ beeinflusst hat.

PDT Interessante Ergebnisse zeigte der PDT (vgl. Abbildung 6.13). Es zeigt sich eine deutliche Verschlechterung der Reaktionszeit bei Blickbedienung, wobei die umgekehrte Tendenz bei geübten Probanden auffällig ist, was durch den Effekt des Starrens zu begründen ist: Widmet der ungeübte Fahrer seine Aufmerksamkeit völlig der Aufgabe, resultiert

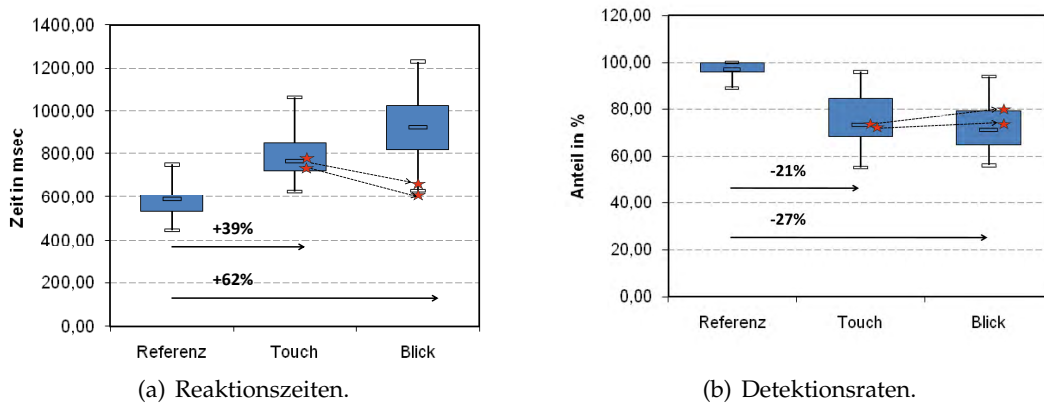


Abbildung 6.13: Ergebnisse der PDT.

daraus die Fokussierung auf das Eingabedisplay. Die Probanden sind fälschlicherweise der Überzeugung, dass sie die Straße im Blick haben (vgl. Abbildung 6.11) und vergessen für die Dauer der Eingabe die restlichen Aufgaben, was zu einem Tunneleffekt führt. Das Resultat waren nicht nur verspätete Reaktionen, sondern auch eine steigende Anzahl

nicht bestätigter Punkte (vgl. Abbildung 6.13(b)). Die PDT Ergebnisse der geübten VPn fallen signifikant besser aus (teilweise auf Niveau der Baseline Erhebung).

Fahrleistung Als eine weitere Messung der Ablenkung des Fahrers diente die mittlere Spurabweichung des LCT. Abbildung 6.14 zeigt diese für die Referenzfahrt sowie Fahrten mit Blick- und Touchinteraktion in Metern. Auch hier zeigt sich der Tunneleffekt. Geübte Fahrer konnten ihre Fahrleistung bei Menübedienung mit dem blickbasierten Konzept um 16% gegenüber der Touchbedienung verbessern, während die Aufgaben bei ungeübten Fahrern eine zusätzliche Belastung hervorriefen. Die in Abschnitt 2.2.1 beschriebene

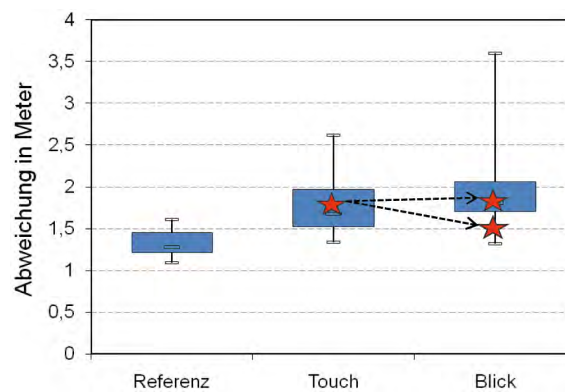


Abbildung 6.14: Mittlere Spurabweichung bei blickbasierter (blau) und bei Touchinteraktion (rot).

begrenzte Menge kognitiver Ressourcen zur Verarbeitung von Informationen und Reizen stößt hier leicht an ihre Grenzen, was die Allokation von Ressourcen zu einer schweren Aufgabe für den Fahrer macht, worunter in erster Konsequenz die selektive Aufmerksamkeit leidet. Wenn Prioritäten unter den Aufgaben falsch verteilt werden, werden in Folge dessen tertiäre Aufgaben gegenüber primären Aufgaben bevorzugt.

6.2.2.3 Fazit

Die Versuche haben gezeigt, dass die Kombination von Touch- und Blickbedienung als Eingabekonzept sehr gute Ergebnisse liefert und von den Probanden als benutzerfreundlich und leicht bedienbar eingestuft wird. Weiterhin zeigt sich, dass für geübte Nutzer eine signifikante Reduktion der Bedienzeiten und der Beanspruchungshöhe des Fahrers bei der Verwendung der Blickbedienung für einfachere Menüeingaben erreicht werden kann.

Integration in weitere Anwendungsgebiete

In den folgenden Abschnitten soll anhand exemplarischer Anwendungen das Potential der erarbeiteten Systeme zur Erfassung der Blickrichtung zum Einsatz in weiteren Anwendungsdomänen vorgestellt werden.

7.1 Blickbasierte Interaktion in Fabrikumgebungen

Langzeitziel des Forschungsclusters CoTeSys [2] ist die Schaffung der *Kognitiven Fabrik* als Fertigungskonzept der Zukunft im industriellen Umfeld. Dazu soll sich die heutige Zusammenarbeit von Menschen und Industrierobotern bzw. Fertigungseinrichtungen mehr und mehr in Richtung einer echten Kooperation entwickeln, wie sie zwischen menschlichen Partnern alltäglich ist. Um eine solche Kooperation auf Augenhöhe zu ermöglichen, müssen zukünftige Industrieroboter in menschlich dominierten Arbeitsbereichen unter Verwendung mehrerer Kommunikationskanäle integriert werden. Eine dafür ausgelegte intelligente Arbeitsumgebung muss folglich diverse Sensoren zur Überwachung der Umwelt, der menschlichen Person sowie der Fertigungsanlagen (z.B. Fließband, Industrieroboter, Material Versorgung, Werkzeuge, Arbeitstisch, etc.) integrieren.

Um eine kognitive und adaptive Interaktion zwischen Werker und Fertigungseinrichtungen zu ermöglichen, wurde die Modalität Blick im Kontext der Kognitiven Fabrik untersucht. Aus der Kenntnis über die aktuelle Blickrichtung bzw. deren zeitlichen Verlauf lassen sich z.B. die nächsten Arbeitsschritte vorhersagen (z.B. wird ein Werkzeug oder Werkstück länger betrachtet, kann davon ausgegangen werden, dass der Werker es in Kürze greifen wird). Ist der nächste Arbeitsschritt bekannt, kann der Materialfluss automatisch dahingehend angepasst werden, dass kein Material in zukünftigen Laufwegen der Arbeiter gelagert wird. Blickinformationen können in diesem Kontext weiterhin genutzt werden, um die HMIs von Terminals zur Kontrolle und Überwachung maschineller Abläufe intuitiver bedienbar zu gestalten und somit den Informationsfluss zwischen Fabrik, Umwelt und Benutzer angemessen zu organisieren.

Motiviert durch diese Forschungsinhalte wurde der im Folgenden präsentierte Prototyp für die blickbasierte Interaktion während manuellen Fertigungsszenarien in kognitiven Fabrikumgebungen der Zukunft umgesetzt. Abbildung 7.1 zeigt den prototypischen Laboraufbau eines manuellen Montageplatzes, einen Werker sowie das zu fertigende

Werkstück. Über die Displays werden dem Arbeiter z.B. dynamisch Arbeitsanweisungen oder Materiallisten angezeigt. Weiterhin ist das in Kapitel 3.2 vorstellte System zur

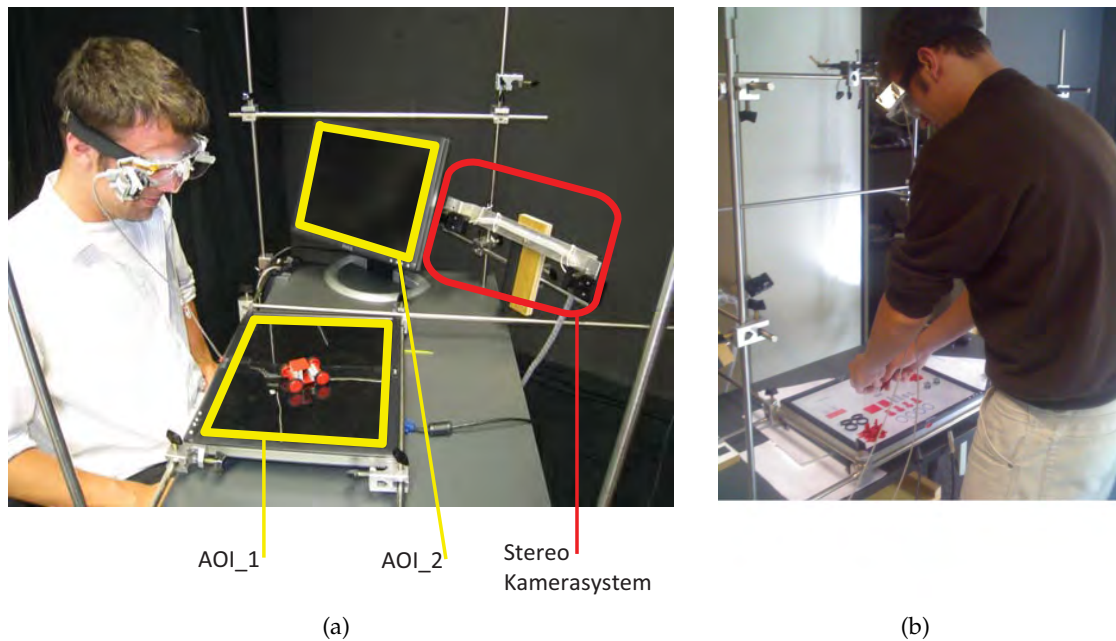


Abbildung 7.1: Verwendeter Laboraufbau eines manuellen Montageplatzes.

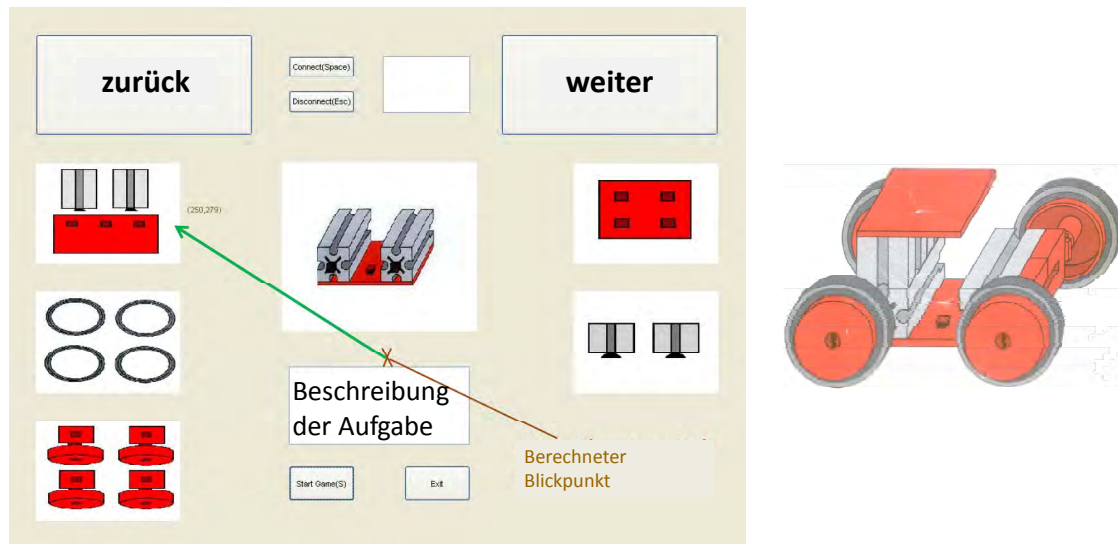
Blickrichtungserfassung bestehend aus kopfgebundenem Eye Tracker und Stereokamerasystem zu sehen. Auf dem liegend angeordneten Display (AOI_1 in Abbildung 7.1) wird dem Arbeiter die Benutzeroberfläche mit der Montageanleitung angezeigt.

Die GUI setzt sich aus reinen Anzeigeflächen sowie interaktiven Inhalten zusammen. Diese kann sowohl auf einem Monitor (wie in Abbildung 7.1 gezeigt) als auch über eine großflächige Beamer-Projektion auf dem Arbeitstisch angezeigt werden. Die Interaktion beruht auf einer Command & Control Funktionalität (z.B. Auswahl von Buttons) und einem Attentive User Interface (z.B. blickrichtungsabhängige Darstellung von Informationen). Zur Gestaltung eines reproduzierbaren Arbeitsablaufs mussten die VPn während der Evaluierung ein Spielzeugfahrzeug „fertigen“ (vgl. Abbildung 7.2(b)).

Der Arbeitsbereich stellt Informationen aus der Montageanleitung dar: z.B. benötigte Einzelteile sowie die Position, an der diese zu montieren sind (siehe Abbildung 7.2(a)). In den seitlichen Bereichen werden Teile bevorratet, die zur Fertigstellung des Arbeitsschrittes nötig sind und im oberen Bereich befinden sich die Schaltflächen zur Navigation durch die Montageanleitung. Sie werden per Blick und nach einer einstellbaren Verzögerungszeit (dwell-time) aktiviert. Stellt das System während der Fertigung Probleme fest, gibt es dem Werker automatisch Hilfestellungen. Kann er beispielsweise benötigte Teile nicht finden,¹ zeigt die Arbeitsfläche entsprechende Hinweise an (z.B. ein Pfeil zum nächsten benötigten Teil, siehe Abbildung 7.2(a)). Der Pfeil beginnt dabei direkt beim berechneten Blickpunkt und weist auf die benötigten Teile auf der Arbeitsfläche.

¹Verwirrung oder zu langes Suchen kann im Blickpfad erkannt werden.

Die Hilfestellung wird ausgelöst, wenn der Nutzer zu lange nach einem Teil suchen muss oder durch das Sprachkommando² *HILFE*. In Abhängigkeit der visuellen Aufmerk-



(a) Gui. Braun: aktueller Blickpunkt. Grün: adaptive Hilfestellung.

(b) Werkstück.

Abbildung 7.2: Grafische Oberfläche auf der Arbeitsfläche und das zu fertigende Werkstück.

samkeit werden weiterhin zusätzliche Informationen zum Arbeitsschritt angezeigt (z.B. Anzugsmomente für Schraubverbindungen). Die Visualisierungen beziehen sich also auf den jeweiligen Arbeitsschritt und werden in der Region dargestellt, auf die der Nutzer seine visuelle Aufmerksamkeit richtet. Dadurch kann der Arbeiter leicht auf zusätzliche Informationen zugreifen, ohne das Werkstück aus der Hand legen zu müssen (z.B. zum Umblättern der Montageanleitung).

7.1.1 Evaluierungsergebnisse

Das System wurde im Rahmen einer Nutzerstudie mit zehn Probanden (eine weiblich) im Alter von 23 bis 31 Jahren evaluiert [23]. Zur Bewertung wurden ein eigener Fragebogen sowie standardisierte Fragebögen (SUS und AttrakDiff) eingesetzt. Der Fragebogen bestand zum Teil aus freien Fragen sowie aus 5-fach Likert Skalen (1 = sehr gut/sehr große Zustimmung/etc., 5 = sehr schlecht/etc.).

Allgemeine Fragen Die Auswertung der allgemeinen Fragen zeigte, dass alle VPn die blickbasierte Interaktion als sinnvoll (durchschnittliche Bewertung 2,3; 1 $\hat{=}$ sehr sinnvoll), praktisch (durchschnittliche Bewertung 1,6; 1 $\hat{=}$ sehr praktisch) und nützlich (durchschnittlich 1,9; 1 $\hat{=}$ sehr nützlich) bewerten. Weiterhin gaben die Probanden an, dass sie subjektiv ein solches System schnell (durchschnittlich 1,7; 1 $\hat{=}$ sehr schnell) und vor allem intuitiv (durchschnittlich 2,1; 1 $\hat{=}$ sehr intuitiv) bedienen können.

²Zur Spracherkennung wurde der kommerzielle Spracherkennung VoCon 3000 der Firma Nuance eingesetzt.

Die meisten Probanden empfanden es als leicht, einen Menüpunkt/Button via Blick auszuwählen (durchschnittliche Bewertung 3,6; 1 $\hat{=}$ sehr leicht). Vor allem bewerteten sie die Visualisierung des aktuell berechneten Blickpunktes in der GUI als vorteilhaft (durchschnittliche Bewertung 2,1; 1 $\hat{=}$ sehr vorteilhaft), da sie so eine sehr gute Systemrückmeldung beim Anvisieren der Schaltflächen hatten. Dies wurde besonders von Anfängern im Umgang mit dem System als nützlich bewertet. Lediglich der Umstand, dass im vorgestellten System die Visualisierung von Rohdaten erfolgte, wurde als etwas verwirrend bewertet. Die Probanden wünschten sich, dass vor der Visualisierung eine Glättung (z.B. durch ein Tiefpassfilter) erfolgen sollte, um eine flüssigere Darstellung des Blickpunktes zu ermöglichen, da sich schnelle Augenbewegungen – z.B. Mikrosakkaden – negativ auf die Darstellung auswirken.

Die adaptive Unterstützung in Form des angezeigten Pfeils wurde als sehr hilfreich bewertet (durchschnittliche Bewertung 1,4; 1 $\hat{=}$ sehr hilfreich), da es eine sehr intuitive Art und Weise der Aufmerksamkeitslenkung des Benutzers auf die entsprechenden Teile (wörtlich von acht Teilnehmern angegeben) darstellt. Außerdem wurde die Beachtung des Blickverhaltens des Nutzers als sehr nützlich (durchschnittliche Bewertung 1,4; 1 $\hat{=}$ sehr nützlich) und hilfreich (durchschnittliche Bewertung 1,7; 1 $\hat{=}$ sehr hilfreich) bewertet. Für die vorliegende Untersuchung wurde dieses Systemverhalten teilweise direkt in das laufende System implementiert und teilweise vom Versuchsleiter übernommen (partieller Wizard-of-Oz Versuch).

System Usability Scale Das System erreichte einen durchschnittlichen SUS Wert von 74,4. Die meisten Teilnehmer bewerteten die Bedienung des System als leicht und eindeutig. Somit kann die Usability des Systems bereits als sehr gut angesehen werden.

AttrakDiff Die Bewertung der Attraktivität des Systems erfolgte mittels des AttrakDiff-Fragebogens. Die pragmatische Qualität des Systems wurde als sehr unterstützend bewertet, es gibt jedoch Raum für Verbesserungen hinsichtlich der Benutzerfreundlichkeit, was vor allem in der sehr einfach gestalteten graphischen Oberfläche des Prototypen begründet liegt. Die hedonische Qualität des Systems wurde von den Probanden nicht eindeutig bewertet: Die Nutzer waren angeregt, aber die Ausprägung der hedonischen Qualität erreicht bei dem Produkt lediglich mittlere Werte. Hier gibt es ebenfalls noch Raum für Verbesserungen. Dennoch ist bemerkenswert, dass sich die Probanden durch das System unterstützt fühlten.

7.1.2 Fazit

Die Untersuchung des prototypischen Interaktionskonzeptes zeigt, dass blickbasierte Interaktion eine adäquate Möglichkeit darstellen kann, um die Arbeit an manuellen Fertigungsplätzen zu vereinfachen. Die guten Ergebnisse des Prototyps können jedoch durch die Integration eines remote Gaze Trackers verbessert werden. Die Mehrheit gab an, sich beim Tragen der Tracking Brille nicht wohl und sehr überwacht gefühlt zu haben.

7.2 Blickbasierte Interaktion in Desktop Szenarien

Im Bereich der Desktop PCs wurde Gaze Tracking in der Vergangenheit hauptsächlich für Marktforschungs- bzw. Usabilitystudien eingesetzt. So wurde z.B. untersucht, wie

sich der Blick eines Probanden beim Verwenden einer Internetseite verhält, um Aussagen hinsichtlich des Nutzungsverhaltens bzw. der Benutzbarkeit zu treffen.

Augenbewegungen erlauben jedoch nicht nur Auskunft über die Interessen des Benutzers, sondern an ihnen können auch zukünftige Aktionen erkannt und Informationen über den Grad der Aufmerksamkeit oder Frustration identifiziert werden (z.B. durch langes Suchen). Blickpunkt sowie Blickverlauf beinhalten also wesentlich mehr Informationen. Dies lässt sich in Applikationen ausnutzen, um z.B. zukünftige Aktionen zu erkennen und so den Benutzer zu unterstützen. Die Verwendung des Blickpunkts als Input bietet des Weiteren folgende Vorteile:

- **Genauigkeit und Geschwindigkeit** Das Auge ist ein sehr schnelles Organ. Es kann sich mit Winkelgeschwindigkeiten von bis zu 600°/Sek sehr präzise bewegen.
- **Intuitivität** Die Verwendung des Auges ist intuitiv und muss nicht trainiert werden.
- **Indikator für Interesse** Der Blickpunkt des Menschen ist ein guter Indikator für dessen Interessen. Man kann davon ausgehen, dass bei einer Fixation auf einen Punkt durch den Benutzer, ein Interesse für Inhalte an dieser Position besteht.
- **Blick vor Bewegung** Der Blick eilt meist einer Handlung voraus. Das heißt z.B., dass der Blick zuerst an eine Position geht, bevor die Maus bewegt wird. Diese Eigenschaft nutzt die „blickunterstützte Maus“ (vgl. Kapitel 7.2.1).

Um diese Vorteile der Benutzung von Blick im Bereich der Interaktion mit Desktop PCs zu untersuchen, wurden exemplarisch die drei im Folgenden vorgestellten Prototypen entworfen, umgesetzt und im Rahmen einer Nutzerstudie untersucht [100]:

1. Die **blickunterstützte Maus** nutzt die Gaze Tracker Daten, um die Geschwindigkeit und die Bewegungsrichtung der Maus zu beeinflussen (vgl. Kapitel 7.2.1).
2. **Blickgesteuerte Shortcuts** erlauben die Definition von Zielen am Rand des Bildschirms, um beliebige Funktionen auszuführen und Programme zu steuern. Die einzige Voraussetzung ist, dass die Funktion einer Tastenkombination zugeordnet ist oder zugeordnet werden kann (vgl. Kapitel 7.2.2).
3. Der **Blickgesteuerter Programmstart** dient dem Starten von Programmen per Blick und verwendet Ziele, welche sich bei Annäherung vergrößern (vgl. Kapitel 7.2.3).

7.2.1 GazeSpeed – Blickunterstützte Maus

Ziel dieser Applikation ist es, das klassische Eingabegerät Maus mit Hilfe von Gaze Daten zu verbessern. Das Ersetzen der Maus hat in bisherigen Arbeiten nicht zu einer Verbesserung in Geschwindigkeit und Komfort geführt und war somit nicht angestrebt.

Zhai präsentierte in „Manual And Gaze Input Cascaded (MAGIC) Pointing“ [257] eine Methode, die darauf basiert, die Maus lediglich zur Feinpositionierung zu verwenden. Er verfolgt zwei unterschiedliche Ansätze: (1) der Mauscursor wird an den Blickpunkt versetzt, wenn ein neues Ziel mit den Augen fixiert wird. Dies kann zu Verwirrung

führen, da sich die Maus auch ohne aktives Zutun bewegt. (2) der Mauscursor wird versetzt, wenn zusätzlich zu einer Fixation eine entsprechende Mausbewegung erfolgt. In diesem Fall wird die Maus kurz vor den Blickpunkt gesetzt und der Cursor kann nun mit der Maus auf das Ziel bewegt werden.

Weiterhin sei auf die Ansätze [19, 134, 139] hingewiesen, die eine Vergrößerung der Region um den Blickpunkt verwenden, um eine Feinpositionierung in diesem Bereich zu ermöglichen. Diverse Ansätze setzen auf eine Kombination von Sprache, Maus und Blick zur Selektion von Icons (z.B. [162]).

GazeSpeed wurde entworfen, um folgende Anforderungen zu erfüllen: (1) Verbessern statt Ersetzen: Ziel ist es, die Maus mit Hilfe der Blickdaten zu verbessern und nicht diese zu ersetzen. (2) Unterbewusste Unterstützung: Der Benutzer soll möglichst nichts von der Unterstützung per Blick merken. Dazu werden die Mausgeschwindigkeit und die Mausrichtung mit Hilfe der Blickdaten folgendermaßen beeinflusst:

- Anpassung der Mausgeschwindigkeit in Abhängigkeit des Abstandes vom Mauszeiger zum Blickpunkt: Durch eine hohe Mausgeschwindigkeit bei großen Entfernungen lassen sich diese schneller überwinden. Kleine Objekte lassen sich genauer selektieren, da die Mausgeschwindigkeit am Blickpunkt deutlich herabgesetzt ist.
- Kompensation von Schlingerbewegungen der Maus (vgl. Abbildung 7.3(a)) durch Anpassung der Maustrajektorie direkt zum Blickpunkt (vgl. Abbildung 7.3(b)). Somit muss erstens weniger Weg zurück gelegt werden und zweitens wird das Ziel weniger häufig verfehlt. Dies bedeutet in Summe eine signifikante Zeitersparnis.



(a) Typischer Weg eines Mauszeigers. (b) Korrigierter Weg eines Mauszeigers.

Abbildung 7.3: GazeSpeed, typischer und korrigierter Mausweg bei der Annäherung an ein Ziel (z.B. Icon auf dem Desktop).

- Arretierung des Mauszeigers in einem fest definierten Bereich um den Blickpunkt zur Vermeidung des Überschwingens der Mausbewegung bei der Annäherung an das Ziel (graue Box in Abbildung 7.3(b)).

7.2.1.1 GazeSpeed – Evaluierungsergebnisse

Die Evaluierung der Applikation sollte zeigen, ob die Funktion einen Zeitgewinn ermöglicht, ob sich die Fehlclickrate verringert und auf welche Weise die Applikation auf den Anwender wirkt.

Die Gesamtuntersuchung (für alle in Kapitel 7.2 vorgestellten Konzepte) wurde mit 12 Personen unter Bürobedingungen durchgeführt. Zur Zeitmessung wurden zwei Vergleichsmessungen ohne Blickunterstützung durchgeführt: eine mit Mausgeschwindigkeit 10 (Standardeinstellung unter Windows) und eine zweite mit Mausgeschwindigkeit 17.

Mit diesen Einstellungen mussten die Probanden 20 zufällig erscheinende Kästchen (23x23 Pixel groß) in möglichst kurzer Zeit anklicken. Außerdem wurde die Anzahl der Klicks protokolliert, um die Anzahl der (Fehl-)Klicks aufzuzeichnen. Zusätzlich wurden die Fragebögen SUS, NASA TLX und AttrakDiff eingesetzt.

Das System erreichte beim SUS einen sehr guten Wert von 87,8. Die Einfachheit der Benutzung erhielt besonders gute Bewertungen. Die Konsistenz der Benutzbarkeit erreichte etwas schlechtere – aber mit 80 immer noch gute – Ergebnisse. Dies lag daran, dass Kästchen vereinzelt nicht sofort geklickt werden konnten, da der Gaze Tracker verrauschte Daten lieferte, wodurch es zu ungewollten Sprüngen des Cursors kam. Diese Erklärung stützt sich auf Aussagen der Probanden und Beobachtungen des Versuchsleiters.

AttrakDiff zeigt eine sehr hohe pragmatische Qualität (optimale Unterstützung des Benutzers) sowie Verbesserungspotenzial hinsichtlich der hedonischen Qualität. Die Identifikation mit dem Programm ist mittelmäßig, die Stimulation aber optimal. Dies lässt sich einerseits dadurch erklären, dass der Benutzer das System nur dann bemerkt, wenn er – zumindest gefühlt – schneller ist. Sonst wird das Programm nicht wahrgenommen, was die Gefahr birgt, dass sich der Benutzer nicht mit dem Konzept identifizieren kann. Die Attraktivität des Konzepts ist sehr hoch. Weiterhin gab ein Proband an, dass das Konzept besonders für CAD Anwendungen sehr sinnvoll sei, ein anderer, dass es sehr gut funktioniert hat, solange der Tracker robuste Daten lieferte. Mehrere Probanden äußerten, dass sie gefühlt schneller waren, auch wenn die Zeitmessungen dies widerlegten.

Die Ergebnisse des NASA TLX zeigen, dass der Faktor „Performance“ den höchsten Einfluss auf die Gesamtbelastung hat (vgl. Abbildung 7.4). Die Benutzer sind demnach am wenigsten damit zufrieden, wie sie die Aufgabe erfüllt haben. Mehrere Probanden waren

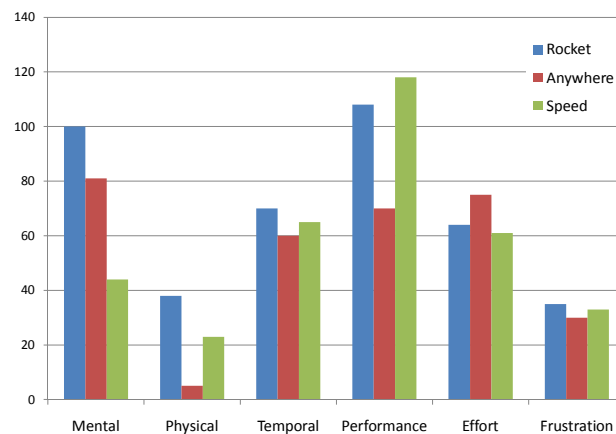


Abbildung 7.4: Übersicht der Ergebnisse des NASA TLX Fragebogens für alle Konzepte.

nach eigener Aussage gefühlt schneller, was jedoch nicht den objektiven Messwerten entspricht. Eine hohe Frustration ergab sich daraus nicht.

Die Bearbeitungszeiten zeigen, dass die blickunterstützte Maus langsamer als die Standardmaus ist. Zur Bestimmung der Signifikanz wurde ein *Two-sample Kolmogorov-Smirnov Test* durchgeführt. Dieser zeigt, dass der Unterschied zwischen der schnellen Maus (17) und der blickunterstützten Maus (Gaze) nicht signifikant ist ($p=0.7358$). Jedoch

ist die blickunterstützte Maus signifikant langsamer als die langsame Maus ($p=0.0121$). Der Vorteil der blickunterstützten Maus besteht nur bei einem gewissen Abstand zum Ziel. Ist der Abstand klein, so lässt sich kein positiver Effekt erzielen. Im Versuch wurde ein Standard Monitor eingesetzt, dessen Fläche jedoch – aufgrund von Ungenauigkeiten des Gaze Trackers am Rand – nicht vollständig genutzt werden konnte (verwendeter Bereich 1250×650 pix). Weiterhin musste der Bereich um das Ziel (keine Wirkung der Mausgeschwindigkeitsanpassung) relativ groß (250×250 px) gewählt und die minimale Geschwindigkeit erhöht werden, was den Effekt der blickunterstützten Maus minimiert. Trotz dieser Maßnahmen kam es zu Problemen bei der Auswahl von Punkten, welche durch Ungenauigkeiten (z.B. durch Kopfbewegungen) hervorgerufen wurden.

Abschließend lässt sich sagen, dass das Konzept äußerst positiv bewertet wurde. Auf einem größeren Bildschirm bzw. bei Multi-Monitor-Systemen sollten sich deutlichere Effekte erzielen lassen. Der Einsatz in Kombination mit Touchpads bei Laptops erscheint vielversprechend, der Einsatz in Kombination mit der Maus und einem Monitor ist jedoch nur bei sehr genauem und fehlerfreiem Tracking interessant.

7.2.2 GazeAnywhere – Blickgesteuerte Shortcuts

Ziel dieser Applikation ist es, unterschiedliche Anwendungen per Blick zu steuern. Fast jede Anwendung eines typischen Desktop PCs kann neben der Bedienung über Menüs auch per Tastenkombinationen/Shortcuts bedient werden (z.B. Strg-S zum Speichern).

GazeAnywhere ermöglicht den Zugriff auf Shortcuts per Blickbedienung. Aus Publikationen ist die Verwendung des Bildschirmrands als Target (*Off-Screen Target*) bekannt. Das Konzept ermöglicht es dem Benutzer, Shortcuts auf beliebige (frei definierbare) Bereiche am Bildschirmrand zu legen und diese durch Fixation und einen Tastendruck abzurufen. Abbildung 7.5 zeigt den Versuchsaufbau. Als Targets wurden PostIt Marker verwendet, um die Shortcut-Bereiche zu markieren. Abbildung 7.5 zeigt exemplarisch die Markierungen zur Verwendung mit Microsoft Word (linker Marker: linksbündiger Schriftsatz; mittlerer Marker: Blocksatz; rechter Marker: rechtsbündiger Schriftsatz).



Abbildung 7.5: GazeAnywhere Versuchsaufbau. Gaze Tracker und unterer Bildschirmrand mit angebrachten PostIts für frei programmierbare Shortcuts.

Die Steuerung von Applikationen über die Verknüpfung von Tastenkombinationen und Targets ist in der bisherigen Literatur nicht bekannt. *Kumar* schlägt vor, die Auswahl der zu drückenden Taste einer blickgesteuerten Maus über Off-Screen Targets (OST) zu lösen [134]. Weiterhin werden OST zur Navigation in Dokumenten verwendet. Hier werden Funktionen wie *Home*, *End*, *Page Up* und *Page Down* an den Rand des Bildschirms gelegt und per Blick und Verzögerungszeit (dwell-time) aktiviert. *Isokoski* verwendet OST

für die Eingabe von Morsecode über 3 Targets (Punkt, Strich und Ende) [116]. Auch hier erfolgt die Aktivierung über dwell-time. Da die Genauigkeit des Gaze Trackers zum Rand des Monitors hin abnimmt, müssen OST deutlich größer sein als Ziele im Zentrum des Bildschirms [30]. GazeAnywhere bietet dem Nutzer folgende Funktionalitäten:

- Aufzeichnen und Abspeichern von Tastenkombination, Applikation und Blickpunkt
- Erkennung der aktuell aktiven Applikation
- Ausführen der zugehörigen Shortcuts bei Aktivierung durch eine Taste
- Oberfläche zum Verwalten und Editieren der Shortcuts

Für jede Applikation existiert ein eigener Satz blickgesteuerter Shortcuts. Hierzu werden die auszuführende Tastenkombination sowie das entsprechende Target gespeichert. Weiterhin kann der Nutzer definieren, in welchen Anwendungen der Shortcut verfügbar ist. Im Betrieb wird ein Shortcut durch simples Drücken einer Taste (z.B. Leertaste) aufgerufen: Existiert zum aktuellen Blickpunkt ein passendes Target in der Shortcut-Bibliothek (bzw. das Target mit dem geringsten euklidischen Abstand), wird die gespeicherte Tastenkombination simuliert und an das entsprechende Zielfenster geleitet.

7.2.2.1 GazeAnywhere – Evaluierungsergebnisse

Die Untersuchung mit 12 Probanden sollte die zeitlichen Vorteile und die Verwendbarkeit des Systems bewerten. Abbildung 7.5 zeigt den verwendeten Versuchsaufbau.

Im Versuch sollten die VPn in Word zunächst ein neues Dokument erstellen und in der ersten Zeile „zentriert“ und „fett“ die Worte *Hallo Welt!* schreiben. In der nächsten Zeile sollen die gleichen Worte „linksbündig“ und „normal“ geschrieben und das Dokument abgespeichert werden. Der Versuchsleiter startet die Messung der Bearbeitungszeit, die automatisch beim Speichern des Dokuments endet (Standardordner und -dateinamen). Zunächst wurde die Untersuchung ohne Blickunterstützung durchgeführt. Es stand der VP dabei frei, ob sie Tastenkombinationen oder Icons verwendet. Anschließend wurde die Untersuchung mit blickgesteuerten Shortcuts durchgeführt.

Die SUS Bewertungen erreichten insgesamt ein sehr hohes Niveau (durchschnittlicher SUS Wert: 88,9). Das System überzeugt durch seine Einfachheit und klare Bedienung. Die Vertrautheit mit der Bedienung des Systems wird geringfügig schlechter bewertet. Mit ca. 80 entspricht dieser Wert aber dennoch einem Wert zwischen dem höchsten und dem zweithöchsten des Fragebogens und ist somit immer noch nahe am Optimum. Zu erklären ist dieser niedrigere Wert vermutlich dadurch, dass die VPn lediglich ca. eine Minute zur Gewöhnung an das System hatten. Durch Beobachtungen und Aussagen der VPn war zu erkennen, dass sie z.B. noch nicht auswendig wussten, an welcher Ecke des Monitors sich welche Funktion befindet. Auch mussten sie sich zuerst umgewöhnen, da sie nun eine bestimmte Stelle fixieren und eine Taste betätigen mussten.

Die Benutzer werden durch das Konzept optimal unterstützt (pragmatische Qualität). Die hedonische Qualität ist sehr hoch, die Dimension Identität ist jedoch geringer als die Stimulation ausgeprägt. Der Benutzer fühlt sich demnach nicht so stark an das Konzept gebunden, was auf die nicht vorhandene (graphische) Oberfläche zurückzuführen ist.

Die Anwendung ist zwar durch die Targets wahrnehmbar, ein direkter Kontakt besteht aber nur bei der Konfiguration. Dennoch ist das Konzept als sehr attraktiv einzustufen und stimuliert den Benutzer durch seine interessanten und anregenden Funktionen.

Abbildung 7.4 zeigt eine Übersicht der Bewertungen hinsichtlich der Arbeitsbelastung durch das System. Der Workload ist insgesamt sehr gering. Die Dimensionen „Mental“, „Effort“, „Perfomance“ und „Temporal“ tragen in absteigender Reihenfolge am meisten zum Gesamtworkload bei. Die VPn gaben an, dass es ungewohnt sei, die Augen zur Bedienung zu verwenden. Da die Benutzer lediglich ca. eine Minute zur Eingewöhnung hatten, ist die Umgewöhnungsphase während der Tests noch nicht abgeschlossen. Ein Proband äußerte weiterhin, dass das Wegschauen vom Dokument auf den Bildschirmrand mental anspruchsvoll (weil ungewohnt) sei. Das begründet auch den hohen „Effort“-Wert.

Die Mittelwerte der Bearbeitungszeiten mit/ohne Blickbedienung sind nahezu gleich. Die Steuerung per Blick ist im Schnitt lediglich 0,77 Sekunden langsamer als die herkömmliche Bedienung (25,0 vs. 24,23 Sekunden), wobei hier von den VPn sowohl Tastenkombinationen als auch die Maus verwendet wurden. Auch wurden Funktionen verwendet, die direkt über Symbole erreichbar sind. Der Unterschied ist nicht signifikant ($p=0.9913$). Somit ließ sich bereits ohne lange Eingewöhnung ein nahezu gleich gutes Ergebnis erzielen. Alle VPn erklärten, dass die Verwendung der Software schneller sei, sobald sie sich umgewöhnt haben. Positiv merkten sie weiterhin an, dass sie gar keine Maus mehr verwenden müssten und es für fast jegliche Software einsetzen könnten.

Zudem fiel die Unempfindlichkeit gegenüber Offsets sehr positiv auf. Bei einer schlechteren Kalibrierung müssen lediglich die Ziele größer sein. Je besser der Gaze Tracker bei einem Anwender funktioniert, desto kleiner kann dieser die Ziele gestalten und desto mehr Funktionen kann er verwenden.

7.2.3 GazeRocket – Blickgesteuerter Programmstart

Ein Programmstart häufig verwendeter Applikationen per Blick kann dem Benutzer Zeit sparen. In der Literatur existieren bereits Ansätze, die sich mit der Verwendung von vergrößernden Zielen (*expanding targets*) beschäftigen.

[154] verwendet Ziele, die ihre Größe in Abhängigkeit des Benutzerfokuses ändern. Die Treffsicherheit ist dabei nicht abhängig von der Ausgangsgröße, sondern ausschließlich von der Zielgröße, weshalb *Fitts' Law*³ beachtet werden sollte [83]. [25] stellt ebenfalls die Verwendung von vergrößerten Zielen als sehr sinnvoll dar und erkennt, dass hauptsächlich die Größe des Ziels die Performance beeinflusst. [19] setzt *expanding targets* in Form von Fisheye-Linsen zur Verbesserung der Blickbedienung ein, während [134] Blick u.a. zum Wechseln zwischen unterschiedlichen Applikationen einsetzt (*application switching*). Dazu wird durch Tastendruck ein 2D Raster von Applikationen angezeigt. Durch Blick auf die gewünschte Applikation und loslassen der Taste wird die gewünschte

³Fitts' Law sagt aus, dass sich bei Verkleinerung eines Ziels die zur Auswahl benötigte Zeit vergrößert. Dieser Zusammenhang wurde jedoch nur für statische Größen untersucht. Bislang wurde nicht untersucht, ob Fitts' Law gültig ist, wenn sich die Größe des Zieles erst ändert, nachdem der Benutzer begonnen hat die Maus zu bewegen. Es kann jedoch davon ausgegangen werden, dass sich vergrößernde Ziele ohne Performanceverlust einsetzen lassen.

Applikation ausgewählt. Ein Programm zum blickgesteuerten Programmstart – ähnlich des hier vorgestellten – ist jedoch in der Literatur bisher nicht bekannt.

Zum blickgesteuerten Starten wurde eine Startleiste entworfen. Da in diesem Bereich umfangreiche Applikationen existieren, wurde das Programm *RocketDock* [5] als Grundlage verwendet. RocketDock besteht aus einer Leiste, die sich bei Annäherung des



Abbildung 7.6: Screenshot der GazeRocket Benutzeroberfläche. Sobald sich der Blick über die Startleiste bewegt, werden die Inhalte um den aktuellen Blickpunkt (hier: der Nutzer fixiert das Firefox Icon) vergrößert dargestellt.

Cursors an den Bildschirmrand einblendet. Bewegt der Benutzer den Cursor über die Leiste, vergrößern sich die Icons bis zu einer einstellbaren Größe und die Applikation kann per Klick gestartet werden. Da die Software jedoch über keinerlei Schnittstellen verfügt, mussten die regulären Benutzereingaben über Maus/Tastatur emuliert werden.

Im Betrieb wird GazeRocket per Tastendruck aktiviert und die Startleiste (vgl. Abbildung 7.6) eingeblendet. Der Benutzer kann dann das gewünschte Icon fixieren und zum Starten des Programms erneut die Aktivierungstaste drücken. Während des Selektionsvorgangs werden Icons vergrößert dargestellt, sobald sich der Blick über sie hinweg bewegt, was dem Nutzer das Treffen des gewünschten Icons erleichtert.

7.2.3.1 GazeRocket – Evaluierungsergebnisse

Ziel der Evaluierung war es, die benötigte Zeit der unterschiedlichen Methoden des Programmstartens zu vergleichen und die Wirkung auf den Probanden zu untersuchen. Aufgabe der Probanden war das Starten von Programmen mittels der Methoden

- Start per Startmenü
- Start per RocketDock ohne Blick
- Start per RocketDock mit Blick

Verglichen mit den Ergebnissen aus den Kapiteln 7.2.1 und 7.2.2 erhielt dieses Konzept niedrigere SUS Bewertungen (Durchschnitt 78,4). Die meisten Werte liegen jedoch auch bei diesem Konzept im Bereich um 75%. Die weniger gute Bewertung ist laut den VPn darauf zurückzuführen, dass das System sehr wenige Möglichkeiten bietet. Die Nutzer sehen keinen Vorteil darin, sich auf die Benutzung des Blicks allein zum Programmstart umzustellen. Vermutlich ist eine solche Umstellung nur dann sinnvoll, wenn man weitere Aktionen per Blick ausführen kann und die Interaktionsmodalität optimal eingebettet ist.

Dies spiegelt sich ebenfalls in der pragmatischen und hedonischen Qualität wieder. Das Konzept erreicht mittlere Werte bei relativ großem Konfidenzintervall. Es besteht

Verbesserungspotential in der Bedienbarkeit. Die Stimulation ist optimal und das Konzept regt an, da es neuartig und interessant ist. Insgesamt ergibt sich eine hohe Attraktivität.

Abbildung 7.4 zeigt die Ergebnisse hinsichtlich der mentalen Anstrengung. Die Dimensionen „Performance“ und „Mental“ wirken sich besonders auf den Gesamtworkload aus. Nach Aussagen der Probanden kommt dies durch die Probleme beim Fixieren der Icons zustande, da es trotz der Vergrößerung per Fisheye oftmals kompliziert ist, das gewünschte Icon zu fixieren.

Beim Vergleich der Bearbeitungszeiten ergab sich, dass der Start per Startmenü signifikant länger dauert als der Start per GazeRocket oder RocketDock. Die im Schnitt schnellste Methode ist RocketDock ohne Blick, wobei der Unterschied zwischen GazeRocket und RocketDock nicht signifikant ist ($p=0.1862$). Der Unterschied zwischen RocketDock/GazeRocket und Startmenü ist signifikant. Weiterhin fällt auf, dass das 25% Quantil von GazeRocket unterhalb des 25% Quantil von RocketDock liegt. Das 75% Quantil von GazeRocket liegt ebenfalls relativ hoch, was bedeutet, dass die Benutzer stark variierende Zeiten hatten. Dies ist teilweise dadurch zu erklären, dass einige VPn nicht auf Anhieb das richtige Icon getroffen haben, sondern erst beim zweiten Versuch.

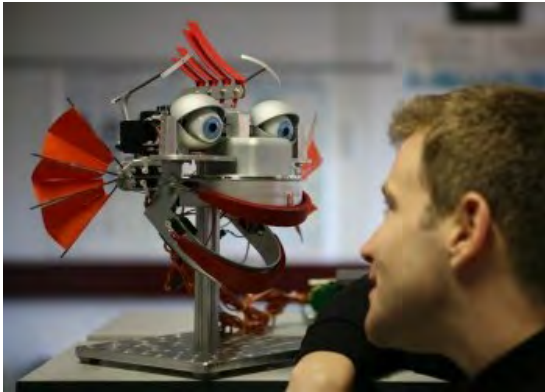
Abschließend kann festgestellt werden, dass das System alleinstehend für den Benutzer weniger interessant ist und vermutlich nur in Kombination mit anderen blickgesteuerten Anwendungen sein volles Potential erreicht.

7.3 Multi-User Gaze Tracking für Robotiksznarien

Einen weiteren Forschungsschwerpunkt des Exzellenzclusters CoTeSys (vgl. Kapitel 7.1) stellt die Interaktion zwischen Menschen und (humanoiden) Robotern dar [2]. Zur Ermöglichung einer kooperativen Mensch-Roboter Interaktion ist die Kenntnis über zukünftige Pläne des menschlichen Partners für die beteiligte Maschine von fundamentalem Interesse. Aus diesem Grund spielt die Blickrichtungserkennung auch im Robotik-Umfeld eine immer größer werdende Rolle. Ist die Blickrichtung des menschlichen Partners in einem Kooperationsszenario bekannt, kann der autonome humanoide Roboter ihm quasi die Wünsche im wahrsten Sinne des Wortes „von den Augen ablesen“, was im Folgenden an einem Beispiel erörtert werden soll: Fixiert bzw. betrachtet die Person einen Gegenstand in ihrer Umgebung und kennt der Roboter die Blickrichtung im 3D Raum, kann er die Blickgerade hinsichtlich greifbarer Objekte⁴ überprüfen. Kann der Roboter ein relevantes Objekt finden, bewegt er sich auf dieses zu, greift und überreicht es dem Menschen.

Zur Erforschung kooperativer Mensch-Roboter Interaktionsszenarien kommen im Rahmen von CoTeSys die Plattformen Eddie und Elias (vgl. Abbildung 7.7) zum Einsatz. Von besonderem Interesse für die vorliegende Arbeit sind dabei die Köpfe der Roboterplattformen. Angelehnt an das menschliche visuelle System, besitzen sie zwei „Augen“, welche – ausgestattet mit entsprechenden Kameras und Algorithmen zur Objektdetektion und -erkennung – zur Erfassung der Umwelt und Navigation durch diese eingesetzt werden. Analog zum menschlichen Verhalten während alltäglicher zwischenmenschlicher

⁴Zusätzlich zum benötigten Gaze Tracking System wird hier die Fähigkeit zur Objekterkennung und Lokalisation im Raum bzw. eine entsprechende Kenntnis über die Szene vorausgesetzt.



(a) Eddie.



(b) ELIAS.

Abbildung 7.7: CoTeSys Demonstrationsplattformen [2].

Interaktionen, war es das Ziel, die Roboteraugen mit der Funktionalität zur Rekonstruktion der Blickrichtung des Gegenübers auszustatten. In der Literatur existieren zwar bereits einige Ansätze, die ein aktives Kamerasystem zur Blickrichtungserkennung einsetzen (z.B. [36]). Diese sind jedoch im vorliegenden Kontext nicht einsetzbar, da sie zum einen für die Verwendung in stationären Desktopszenarien und zum anderen lediglich zur Erkennung der Blickrichtung einzelner Probanden ausgelegt sind.

Da im Rahmen von CoTeSys jedoch mobile Szenarien mit mehreren menschlichen Partnern und Robotern von besonderem Interesse sind, wurde das im Folgenden vorgestellte System *EyeSeeBot* entwickelt. Der Roboterkopf musste zur Integration der in Kapitel 3.3 vorgestellten Blickrichtungsrekonstruktion um folgende Komponenten erweitert werden:

- Integration eines Teleobjektives und des in Kapitel 3.3.3.4 vorgestellten Autofokus, um Nutzer in variierenden Distanzen analysieren zu können (vgl. Abbildung 7.8).
- Integration der in [245] vorgestellten aktiven, hochdynamischen Kameraorientierungssysteme, die eine nahezu latenzfreie, äußerst schnelle und vor allem hochgenaue Ausrichtung der Kameras ermöglichen. Dadurch können die Kameras sakkadenartig auf die Augenregion der einzelnen Nutzer ausgerichtet werden.
- Integration einer Weitwinkel Szenenkamera zur Erfassung von Personen. Die Daten werden genutzt, um die Blickkameras auf die jeweilige Augenregion auszurichten.

7.3.1 Verwendete Kameraorientierungssysteme und Hardwareaufbau

Das verwendete Kameraorientierungssystem besitzt zwei Freiheitsgrade und weist dynamische Eigenschaften ähnlich dem menschlichen Auge auf. Für weitere Details inklusive dem inversen Kinematik-Modell zur korrekten Positionierung sei auf [245] verwiesen.

Abbildung 7.9 zeigt die beiden Rotationsachsen sowie die Antriebsmechanik. Ein Kardangelen ermöglicht Schwenk- und Neigebewegungen der Kamera, wobei als Antriebe zwei Piezo-Aktuatoren dienen, die jeweils eine Schubstange steuern (Auslenkungen Θ_1 ,



Abbildung 7.8: Einsatz des Autofokus zur Analyse von Nutzern in variierenden Distanzen.

Θ_2), die die Bewegung auf die Kamera überträgt. Die Antriebsmechanik stellt ein paralleles kinematisches System dar, was sich für eine mechanische Orientierung von Kameras aufgrund der Dynamik besonders gut eignet, da die Motoren nicht mit bewegt werden müssen und sich deren Trägheit nicht auf die Trägheit der anderen Teile addiert. Eine auf zwei voneinander entkoppelten PID-Regelkreisen basierende Regelungsarchitektur ermöglicht in Kombination mit der verwendeten Mechanik Rotationsbewegungen der Kamera mit einer Winkelgeschwindigkeit von mehr als 1000 Grad/s. Ausgehend von der Ruhelage⁵ ermöglicht diese Mechanik Neigungen von ± 30 Grad (um y -Achse) und Schwenkbewegungen von ± 40 Grad (um x -Achse) der Kamera. Das Orientierungssystem

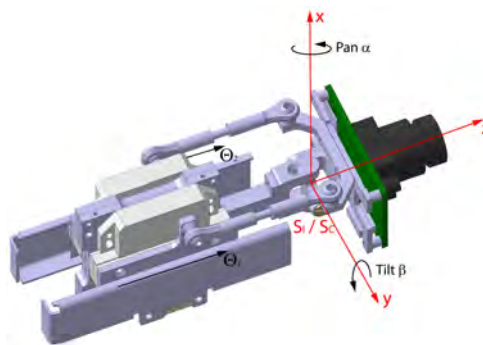


Abbildung 7.9: CAD Modell des Kameraorientierungssystems mit Freiheitsgraden [245].

tem führt lediglich Kamera und on-axis Beleuchtung nach. Auf eine Nachführung der off-axis Beleuchtungseinheiten wurde verzichtet, um die Trägheit des Systems möglichst gering zu halten. Zur Berechnung der Blickgeraden müssen alle 3D Geometriedaten des verwendeten Aufbaus (Position der Kamera sowie der einzelnen Beleuchtungsgruppen) zusätzlich zu den extrahierten Bildmerkmalen bekannt sein. Deshalb muss eine erneute Berechnung der geometrischen Daten des Gesamtaufbaus nach jeder Bewegung der Kamera aus den Motordaten erfolgen (vgl. Kapitel 7.3.3).

⁵ S_I koinzident mit S_C wie in Abb. 7.9

7.3.2 Gesichtsdetektion

Zur Detektion relevanter Nutzer wurde der Aufbau um eine unbewegliche Kamera ergänzt (mittlere Kamera in Abbildung 7.10), die mit einem Weitwinkelobjektiv bestückt ist und zur Aufzeichnung eines Gesamtüberblicks der Szene (entsprechend dem Gesichtsfeld eines Menschen) dient. Das Videobild (vgl. Abbildung 7.10) dient zur Detektion von Personen mittels eines Gesichtsdetektors. Dafür wurde der in OpenCV implementierte Algorithmus nach *Viola und Jones* [246] verwendet, der auf einem mit frontalen Gesichtern trainierten Klassifikator basiert. Als Ergebnis des Detektors erhält man die Positionsdaten (x - und y -Pixelkoordinaten) und die Größe im Videobild detektierter Gesichter. Diese Ergebnisse werden an ein simples, regel-basiertes Trackingmodul übergeben. Dadurch können mehrere Gesichter im Bild unterschieden werden, um ein alternierendes Scannen der einzelnen Augenregionen zu realisieren. Hierbei werden für Kopfbewegungen maximale Geschwindigkeiten angenommen. Weiterhin werden gegenseitige Verdeckungen von Gesichtern und ein folgendes Wiederauftauchen der Gesichter (merge & split) aufgrund der Regeln ausgeschlossen. Dieser erste Prototyp geht von einer festen Entfernung der Gesichter aus. Zukünftige Erweiterungen sollten daher einen 3D Ansatz (z.B. Stereo-Kamerasystem) für die Gesichtsdetektion adressieren. Ist lediglich eine Person in der Szene vorhanden, folgt das System den Kopfbewegungen dieser. Wird mehr als eine Person detektiert, folgt die Blickkamera alternierend den jeweiligen Kopfbewegungen der einzelnen Personen, um ihre Blickrichtung zu analysieren.

7.3.3 Ausrichtung der Blickkameras

Basierend auf den Positionsdaten der Gesichter erfolgt im Betrieb die Ausrichtung der Roboteraugen auf die Augenregion der Probanden. Im Videobild können dann die benötigten Parameter zur Rekonstruktion der Blickgeraden berechnet werden. Aus den Positionen der Gesichter wird zunächst die 3D Position der Köpfe berechnet. Basierend auf diesen Koordinaten werden über trigonometrische Beziehungen die notwendigen Auslenkwinkel $winkel_x$ und $winkel_y$ um die x - und y -Achse (siehe Abb. 7.9) berechnet. Anschließend können aus dem inversen Kinematik-Modell [245] die (ganzzahligen) Schrittweiten $step_x$ und $step_y$ und die Servo-Kommandos zur Ansteuerung der Motoren bestimmt werden.

Zur Anpassung der Geometriedaten nach einer Bewegung der Blickkamera wurde initial die Lage der einzelnen Koordinatensysteme zueinander sowie die Positionen der off-axis Beleuchtungen (vgl. Kapitel 3.3) im Koordinatensystem S_i vermessen. Zur Rekonstruktion der Blickrichtung müssen die Koordinaten der LEDs im Koordinatensystem S_{cam} bekannt sein.

Dabei handelt es sich um eine verkettete Transformation: Zuerst werden die Koordinaten über cT_i vom initialen Koordinatensystem nach S_c und anschließend mittels ${}^{cam}T_i$ nach S_{cam} überführt. Für cT_i gilt (mit α und β als Auslenkwinkel im Bogenmaß):

$${}^cT_i = \begin{bmatrix} \cos(\beta) & \sin(\alpha) \cdot \sin(\beta) & -\cos(\alpha) \cdot \sin(\beta) & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) & 0 \\ \sin(\beta) & -\sin(\alpha) \cdot \cos(\beta) & \cos(\alpha) \cdot \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7.1)$$

7. Integration in weitere Anwendungsgebiete

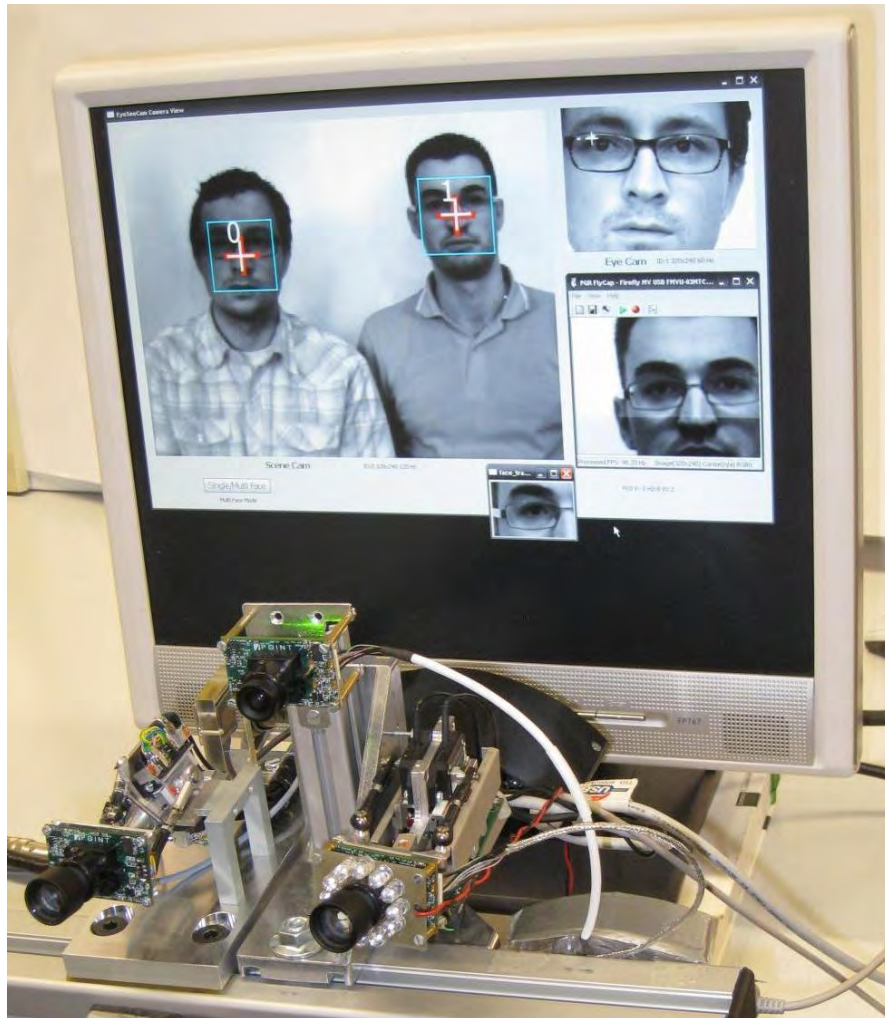


Abbildung 7.10: Gesamtaufbau *EyeSeeBot*. Unten im Bild die beiden Beweglichen Roboter-Augen/Kameras sowie die Szenenkamera. Auf dem Bildschirm ist die graphische Benutzeroberfläche des Systems mit dem Bild der Szenenkamera (großes Videofenster mit markierten Positionen der gefundenen Gesichter sowie deren IDs) sowie den einzelnen Bildern der ausgerichteten Blickkameras (kleine Videofenster) zu erkennen.

Die Matrix cT_i entsteht aus der in [245] angegebenen Matrix iT_c ⁶ durch die Beziehung

$${}^cT_i = {}^iT_c^{-1}, \quad (7.2)$$

mit

$${}^iT_c = \begin{bmatrix} {}^iR_c & r \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (7.3)$$

⁶In [245] mit iR_c bezeichnet.

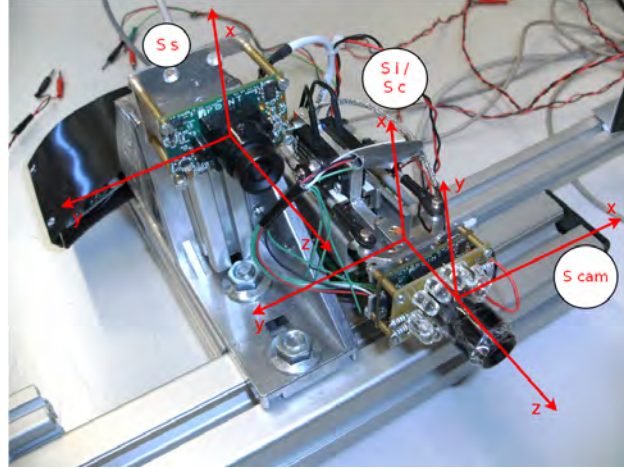


Abbildung 7.11: Verwendete Koordinatensysteme im *EyeSeeBot* Gesamtsystem. Der Ursprung von S_s liegt in der Szenenkamera mit der Distanz der Brennweite von dem Bildsensor entfernt. S_c ist hier koinzident mit S_i (Gaze Tracker Kamera nicht ausgelenkt), Ursprung beider Koordinatensysteme liegt jeweils im Kardangelenk. Der Ursprung von S_{cam} liegt in der Gaze Tracker Kamera, ebenfalls mit der Brennweiten-Distanz zum Bildsensor entfernt.

Bei einer reinen Rotation gilt:

$${}^i T_c^{-1} = {}^i T_c^T = \begin{bmatrix} {}^i R_c^T & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (7.4)$$

wobei Gleichung (7.4) lediglich für Sonderfälle homogener Transformationen gilt. Im vorliegenden Fall beschreibt T eine reine Rotation. ${}^i R_c$ setzt sich wie folgt zusammen:

$${}^i R_c = \text{pan}(\alpha) \cdot \text{tilt}(\beta) = R_x \cdot R_y \quad (7.5)$$

mit R_x und R_y als Rotationsmatrizen um die X- bzw. Y-Achse. Für ${}^{cam} T_c$ gilt:

$${}^{cam} T_c = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -(g+f) \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -41.9 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7.6)$$

${}^{cam} T_c$ setzt sich aus einer Rotation von S_c (um z-Achse von S_{cam} um $+90^\circ$) und einer Translation von $[0, 0, -(g+f)]^T$ zusammen. g bezeichnet den Abstand des Kardangelenks zur Sensorfläche und f die Brennweite der Gaze Tracker Kamera. g wurde vermessen (16 mm) und die Brennweite f (25,9 mm) wurde über die Kamerakalibrierung ermittelt. Die Position der LED-Gruppen im Koordinatensystem S_{cam} ergibt sich demnach als

$$\mathbf{l}_{cam}^{(l)} = {}^{cam} T_c \cdot {}^c T_i \cdot \mathbf{l}_i^{(l)} \quad (7.7)$$

$$\mathbf{l}_{cam}^{(r)} = {}^{cam} T_c \cdot {}^c T_i \cdot \mathbf{l}_i^{(r)} \quad (7.8)$$

Die (transformierten) Positionen der LED-Gruppen liegen danach im Koordinatensystem des Eye Trackers vor und können für die Rekonstruktion der Blickgeraden verwendet

werden. Sobald die Blickgerade bestimmt ist, muss diese für eine weitere Verwendung ins Weltkoordinatensystem S_s transformiert werden. Diese Transformation kann invers obiger Transformationskette sowie einer Transformation der Daten mittels sT_i von S_i nach S_s beschrieben werden. Es ergeben sich folgende Matrizen ${}^cT_{cam}$, iT_c und iT_c ⁷:

$${}^cT_{cam} = {}^{cam}T_c^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & g+f \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 41.9 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7.9)$$

$${}^iT_c = {}^cT_i^{-1} = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) & 0 \\ \sin(\alpha) \cdot \sin(\beta) & \cos(\alpha) & -\sin(\alpha) \cdot \cos(\beta) & 0 \\ -\cos(\alpha) \cdot \sin(\beta) & \sin(\alpha) & \cos(\alpha) \cdot \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7.10)$$

$${}^sT_i = \begin{bmatrix} 1 & 0 & 0 & -60 \\ 0 & 1 & 0 & -45 \\ 0 & 0 & 1 & 24 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7.11)$$

Mittels dieser Matrizen ergibt sich in S_s folgendes Ergebnis für Punkte der optischen Achse:

$$\mathbf{p}_s = {}^sT_i \cdot {}^iT_c \cdot {}^cT_{cam} \cdot \mathbf{p}_{cam} \quad (7.12)$$

Für weitere Einzelheiten zur echtzeitfähigen Implementierung der Transformationsketten sei auf [251] verwiesen.

7.3.4 Systemevaluierung

Die Untersuchung des Systems hatte zwei Ziele: (a) Evaluierung der Genauigkeit der Blickrichtungsrekonstruktion und (b) die Ermittlung des Arbeitsbereiches. Zur Messung der Genauigkeit wurden fünf (vermessene) Referenzpunkte im Raum verwendet (2 off-axis Beleuchtungen, Mitte der Szenenkamera, je 1 Referenzpunkt willkürlich links bzw. rechts des Aufbaus platziert). An der Evaluierung nahmen elf Versuchspersonen teil. Als Maß für die Genauigkeit der Blickrichtungsbestimmung wurde der Mittelwert der Distanzen zwischen Referenzpunkt und den geschätzten Blickpunkten für 200 Blickgeraden pro Referenzpunkt definiert. Es wurden 3 Messreihen mit festen aber unterschiedlichen Kopfpositionen (in der Mitte, am linken und am rechten Rand des Szenenkamerabildes) sowie eine Messreihe mit horizontalen und vertikalen Kopfbewegungen durchgeführt. Währenddessen mussten die Probanden sequentiell die fünf Referenzpunkte fixieren.

Genauigkeit Die mittlere Genauigkeit über alle VPn und Referenzpunkten fiel mit $3,28^\circ$ deutlich schlechter aus als in Kapitel 3.3.5. Dies liegt vor allem an der (für diesen Versuch fehlenden) Nutzerkalibrierung. Weiterhin mussten die Personen bei den beiden Referenzpunkten am Rand teilweise scharf nach links oder rechts blicken. Die Folge davon sind schwer oder teilweise nicht detektierbare Glints.

⁷Die Position des Kardangelenks relativ zum Koordinatensystem der Szenenkamera wurde vermessen und die Brennweite f (3,1 mm) der Szenenkamera wurde über die Kamerakalibrierung ermittelt.

Arbeitsbereich Das System ohne Kameranachführung bietet dem Nutzer ein Volumen (head-box) von 20x15x10 cm (BxHxT) für Kopfbewegungen. Zur Bestimmung der Vergrößerung des Arbeitsbereiches durch den vorgestellten Aufbau wurden vier Kopfpositionen frontal zur Kamera ermittelt: (a) zwei Kopfpositionen, bei denen die Augen dicht am linken sowie rechten Rand im Videobild der Gaze Tracker Kamera sichtbar waren, sowie (b) zwei Positionen, bei denen die Augen dicht am unteren sowie oberen Bildrand sichtbar waren. Dazu wurde jeweils der Kopf über die Szenenkamera verfolgt und die Blickkamera nachgeführt, bis das Kameraorientierungssystem am Ende seiner Beweglichkeit angelangte. Als Ergebnis lässt sich ein Bereich von 60x60x10 cm (BxHxT) angeben, in dem die Blickrichtung von Probanden erfasst werden kann [196].

In Kombination mit dem in Kapitel 3.3 vorgestellten Autofokus ergibt sich ein Arbeitsbereich von 60x60x70 cm (BxHxT). Dies ist ein Vielfaches des Arbeitsbereiches von kommerziell erhältlichen State-of-the-Art Eye Tracking Systemen. Somit kann das vorgestellte System *EyeSeeBot* alle geforderten Eigenschaften hinsichtlich Multi-User-Tauglichkeit und Arbeitsbereich bei gleichzeitigem mobilen Einsatz auf einer Roboterplattform erfüllen und ermöglicht erstmals die Blickrichtungserkennung mehrerer, beweglicher Nutzer und gleichzeitig einen – abhängig vom gewählten Kameraorientierungssystem – nahezu unbegrenzten Arbeitsraum der Blickrichtungserkennung [196].

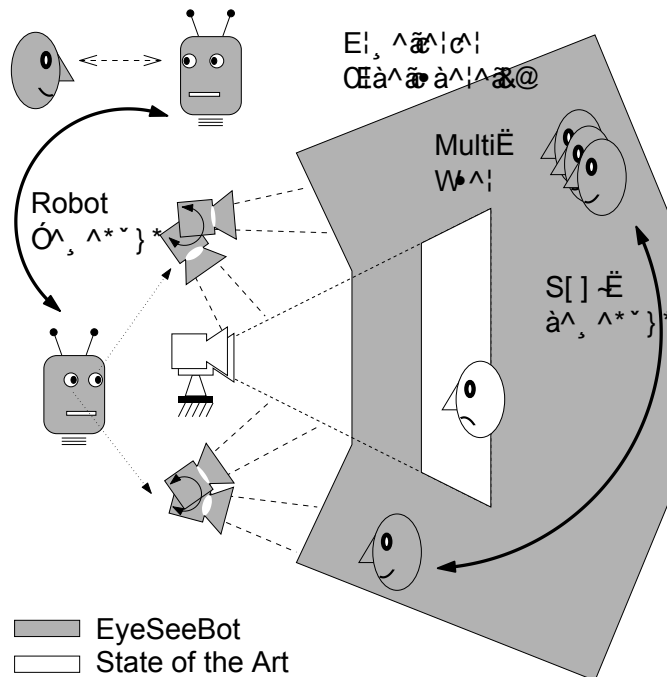


Abbildung 7.12: Erweiterter Arbeitsbereich des *EyeSeeBot* Aufbaus. Entnommen aus [196].

Zusammenfassung und Ausblick

Zusammenfassung

Das zentrale Anliegen der vorliegenden Arbeit bestand darin, umfassende Lösungsansätze zur Erfassung der Fahreraufmerksamkeit und einer darauf basierenden Anpassung des Informationsflusses im Kraftfahrzeug vorzustellen. Dazu lieferten die einzelnen Kapitel – aufbauend auf den in Kapitel 2 erarbeiteten domänenspezifischen Grundlagen – relevante Bausteine zur Umsetzung einer intuitiven, blickbasierten Mensch-Maschine Interaktion in zukünftigen Fahrzeuggenerationen.

Zur umfassenden Erfassung und Überwachung der Fahreraufmerksamkeit stellten Kapitel 3 und Kapitel 4 Ansätze zur Erkennung der Blickrichtung und einer Auswertung von Fahrzeugdaten vor. Im Rahmen umfangreicher Forschungsarbeiten im Bereich der Blickrichtungserkennung entstand mit dem in Kapitel 3 vorgestellten System *eyeDrive* eine kombinierte Hard- und Softwarelösung, die die berührungslose Analyse von Augenbewegungen ermöglicht. Mit einer Genauigkeit von bis zu 0.45° und Datenraten bis 100 Hz stellt das System einen optimalen Ausgangspunkt zur Realisierung blickbasierter Anzeige-Bedienkonzepte dar. Neben der entworfenen Strategie zur Rekonstruktion des Korneazentrums und der erstmaligen Integration eines nahezu latenzfreien Autofokus auf Basis einer Flüssiglense, grenzt sich *eyeDrive* weiterhin durch die integrierte, einfach in Betrieb zu nehmende Hard- und Software von anderen Systemen aus dem Forschungsbereich und kommerziell erhältlichen remote Eye Trackern ab.

Zusätzlich zu dieser Möglichkeit der expliziten Erfassung des Aufmerksamkeitsfokus mittels Eye Tracking stellt Kapitel 4 das System AWARE zur impliziten Erfassung von Ablenkung auf Basis von Fahrzeugdaten vor. Dieses System analysiert Fahrzeugparameter (z.B. Geschwindigkeit, Beschleunigung und Abstand zum Vorderfahrzeug), um Abnormalitäten im Regelverhalten des Fahrers festzustellen. Auf Basis dieser Analysen bestimmt das System den momentanen Abnormalitätsgrad des Fahrverhaltens und ermöglicht somit einen Rückschluss auf den Fahrerzustand, den die entworfenen Anzeigekonzepte zur Adaption der Mensch-Maschine Schnittstelle an den aktuellen Fahrerzustand nutzen.

Aufbauend auf diesen Systemen zur Erfassung der Aufmerksamkeit stellte Kapitel 5 exemplarisch Konzepte zur Anpassung der Mensch-Maschine Schnittstelle an den Nutzer vor. Diese adressieren das örtliche und zeitliche Scheduling von Systemaufgaben, die

Warnung vor drohendem Aufmerksamkeitsverlust, die effektive Lenkung der Fahreraufmerksamkeit sowie die frühzeitige, blickbasierte Erkennung von Bedienwünschen. Das in Kapitel 5.2 vorgestellte örtliche Scheduling hat das Ziel, Systemausgaben an die Aufmerksamkeit des Fahrers anzupassen und diese möglichst auf einer Anzeigefläche in der Nähe des Aufmerksamkeitsfokus darzustellen. Bereits nach einer kurzen Eingewöhnungsphase konnte das System signifikant bessere Ergebnisse hinsichtlich Akzeptanz, subjektiver Beanspruchung, Bedienbarkeit und vor allem Bedien-/Reaktionszeit – verglichen mit einer konventionellen Ausgabe – erreichen. Zudem zeigt Kapitel 5.3, dass es bereits durch ein relativ simpel gestaltetes System zur Warnung vor Aufmerksamkeitsverlust gelingt, die Probanden vor zu hoher kognitiver Ablenkung bei Menüeingaben zu schützen. In den Versuchen äußerte sich der Effekt des Systems einerseits in einer besseren Fahrleistung und andererseits in den sehr guten Akzeptanzwerten der untersuchten Konzepte. Das auf Basis der Fahrabnormalität umgesetzte System zum zeitlichen Scheduling von Systemausgaben untermauert in Kapitel 5.4 den positiven Einfluss adaptiver Systeme auf die Fahrperformance der Probanden. Eine adaptive Informationspräsentation zu Zeitpunkten geringer Fahrabnormalität resultierte während den Versuchen in einer besseren Informationsverarbeitung. Situationen mentaler Überforderung wurden verringert und durch die Informationsausgabe in entspannten Fahrsituationen standen den Probanden ausreichend mentale Ressourcen zur Bearbeitung von Sekundäraufgaben zur Verfügung, ohne die Güte der Primäraufgabe zu reduzieren. Aufbauend darauf präsentierte Kapitel 5.5 Ansätze zur Aufmerksamkeitslenkung im Fahrzeug. Die Methoden mittels blickrichtungsabhängigen Führungspunkten und Warntönen sind äußerst effektiv und zielführend, was sich in signifikant besseren Primär- und Sekundäraufgabendaten zeigte.

Neben diesen Warnkonzepten stellten Kapitel 5.6 und Kapitel 6 Konzepte zur blickbasierten Bedienung von Mensch-Maschine Schnittstellen vor. Dabei nutzt der Ansatz auf Basis der intentionsbasierten Planerkennung die Blickrichtung zur frühzeitigen Erkennung von Bedienintentionen. Der Interpretationsprozess ermöglicht bei der kombinierten Auswertung von Hand- und Blickdaten im verwendeten Setup eine korrekte Prädiktion von 91,0% der Nutzereingaben. Zudem zeigen Versuchsergebnisse zur intentionsbasierten Parametrierung der Mensch-Maschine Schnittstelle, dass Nutzer ein darauf aufbauendes, vorausschauendes Systemverhalten wünschen und akzeptieren. Kapitel 6 präsentierte ein Bedienkonzept, das den aktuellen Trend hin zu großflächigen Multi-Display Anzeige-Bedienkonzepten im Fahrzeug aufgreift, konsequent erweitert und als neuartige Bedienart die Modalität Blick integriert. Untersuchungen haben gezeigt, dass multimodale, blickbasierte Interaktionstechniken im Fahrzeug nach einer (kurzen) Gewöhnungsphase sehr gute Ergebnisse liefern. Die Probanden stufen diesen Interaktionsstil als benutzerfreundlich und leicht bedienbar ein und es zeigt sich, dass signifikant geringere Bedienzeiten und mentale Beanspruchung bei gleichzeitig sehr guten Akzeptanzwerten erreicht werden.

Abschließend gab Kapitel 7 einen Ausblick auf das Nutzungspotential der vorgestellten Ansätze zur Erfassung der Blickrichtung im Kontext anderer Anwendungsdomänen. Das Kapitel stellte die Integration blickbasierter Interaktionstechniken in einen manuellen Montageplatz einer kognitiven Fabrik der Zukunft, die Nutzung von Blickdaten zur einfachen und intuitiven Bedienung von Desktoparbeitsplätzen sowie das umgesetzte Multi-User Eye Tracking System EyeSeeBot für Robotiksznarien vor. Dieses System er-

weitert – im Vergleich zu State-of-the-Art Eye Tracking Systemen – den Arbeitsbereich in erheblichem Maße, wodurch erstmals die Blickrichtungserkennung mehrerer, beweglicher Nutzer und gleichzeitig ein nahezu unbegrenzter Erfassungsbereich möglich wird.

Ausblick

Neben dem bereits heute ersichtlichen Trend im Bereich der Forschung und Entwicklung hin zum kamerabasierten Driver Monitoring, zeigen die präsentierten Untersuchungsergebnisse durchweg das Potential einer Integration der Modalität Blick im kognitiven Automobil der Zukunft. Essentiell wichtige Punkte hinsichtlich einer hohen Nutzerakzeptanz sind dabei ein möglichst großer Erfassungsbereich sowie eine möglichst unkomplizierte Inbetriebnahme des Systems. Zur Steigerung des Erfassungsbereichs muss in Großserienfahrzeugen ein anderer, als der in Kapitel 7.3 vorgestellte, Weg eingeschlagen werden, da aktive Kameraorientierungssysteme aufgrund von Bauraumbeschränkungen und der mechanischen Gefährdung durch Vibrationen keine adäquate Option darstellen. Für dieses Einsatzszenario stellen einzig Kameras mit gesteigerter Sensorfläche zur Vergrößerung des Arbeitsbereiches eine Lösung dar. Hinsichtlich einer einfachen Inbetriebnahme eines videobasierten Systems zur Fahrerüberwachung besteht Forschungsbedarf im Bereich einer automatischen Nutzerkalibrierung, die ohne Fixation von Referenzpunkten die relevanten, nutzervarianten Augenparameter bestimmen kann. Dazu stellt beispielsweise [28] erste Ansätze einer neuartigen, kombinierten Hardware- und Nutzerkalibrierung vor. Eine Verbesserung der Genauigkeit und der – vor allem im Anwendungsszenario Fahrzeug wichtigen – Robustheit der Blickrichtungserfassung kann durch die Entwicklung neuer Augenmodelle, die die tatsächlichen geometrischen Gegebenheiten besser abbilden, zusätzlich vorangetrieben werden.

Die Auswertung der aufgezeichneten Versuchsdaten zeigt durchweg den positiven Einfluss der vorgestellten Anzeige-Bedienkonzepte auf die Fahrleistung und Bedienzeiten der Probanden im Fahrsimulator. Die Warnkonzepte konnten Situationen hoher mentaler Beanspruchung reduzieren, zudem hat sich als Eingabeart die Kombination von Touchbedienung und Eye Tracking als sehr erfolgreich und intuitiv erwiesen, jedoch ergibt sich für zukünftige Untersuchungen die Fragestellung: „Wie nehmen Nutzer dieses Systemverhalten im realen Fahrzeug an?“ Zur eingehenden Bewertung der erarbeiteten Ansätze sind deshalb weitere umfassende Feldstudien im realen Straßenverkehr zwingend erforderlich. Vor allem gilt es in Feldversuchen, den Einfluss von realen Verkehrssituationen auf die Fähigkeiten zur Blickbedienung sowie den Einfluss der Beleuchtungsbedingungen auf die erreichbare Genauigkeit der Blickrichtungsbestimmung zu untersuchen. Dabei gilt es das Potential der Blickbedienung auch im automobilen Umfeld ähnlich umfassend zu bestätigen, wie es Forschungsvorhaben im Desktopbereich bereits belegen konnten.

Da in den vorgestellten Versuchen Probanden oftmals explizit angaben, dass sie – allein für das im jeweiligen Versuch untersuchte Konzept – keinen Vorteil in der Integration der Modalität Blick sehen und eine Umstellung nur dann als sinnvoll erachten, wenn man „alle“ Aktionen per Blick ausführen kann und die Interaktionsmodalität optimal eingebettet ist, besteht weiterhin Forschungsbedarf hinsichtlich der Möglichkeiten einer durchgängig konsistenten Integration der Blickrichtung zur Bedienung der gesamten

8. Zusammenfassung und Ausblick

Mensch-Maschine Schnittstelle. Nur unter diesen Voraussetzungen kann diese Modalität ihr gesamtes Potential ausschöpfen und gleichzeitig eine hinreichend gute Akzeptanz seitens der Nutzer erreichen, die nötig ist, um den Aufwand zur Umsetzung eines feldtauglichen Systems zur Erfassung der Blickrichtung zu rechtfertigen.

Literaturverzeichnis

- [1] *COmmunication by GAze INteraction*. www.cogain.org, letzter Zugriff: 25.02.2011,
- [2] *Exzellenzcluster CoTeSys - COgnition for TEchnical SYStems*. www.cotesys.org, letzter Zugriff: 25.03.2011,
- [3] *EyeSeeCam HeadCam*. www.eyesecam.com, letzter Zugriff: 25.03.2011,
- [4] *Fragebogen AttrakDiff. Service der User Interface Design GmbH*. www.attrakdiff.de, letzter Zugriff: 25.02.2011,
- [5] *RocketDock Homepage*. <http://rocketdock.com/>,
- [6] *ISO Norm 9241-11*. 1990
- [7] ABLASSMEIER, M. : *Multimodales, kontextadaptives Informationsmanagement im Automobil*, Technische Universität München, Diss., 2009
- [8] ABLASSMEIER, M. ; MCGLAUN, G. ; GAST, J. ; POITSCHKE, T. ; RIGOLL, G. : A Robust, Context-adaptive and Multimodal Search Engine for Efficient Information Retrieval in Car Environments. In: SALVENDY, G. (Hrsg.): *Proceedings of 11th Intern. Conf. on Human-Computer Interaction, HCI 2005, Las Vegas, USA.*, Lawrence Erlbaum Associates, Inc., 2006, S. CD-ROM
- [9] ABLASSMEIER, M. ; POITSCHKE, T. ; REIFINGER, S. ; RIGOLL, G. : Context-Aware Information Agents for the Automotive Domain Using Bayesian Networks. In: SMITH, M. (Hrsg.) ; SALVENDY, G. (Hrsg.): *Proceedings of International Conference on Human-Computer Interaction HCI 2007, Beijing, China, Proc. Part I* Bd. LNCS 4557, Springer Verlag Heidelberg, 22-27 July 2007, S. 561–570
- [10] ABLASSMEIER, M. ; POITSCHKE, T. ; RIGOLL, G. : A new approach of a context-adaptive search agent for automotive environments. In: *Extended Abstracts on Human Factors in Computing Systems CHI 06, Montreal, Quebec, Canada*, ACM Press, New York, NY, 22-27 April 2006, S. 1613–1618
- [11] ABLASSMEIER, M. ; POITSCHKE, T. ; WALLHOFF, F. ; BENGLER, K.-J. ; RIGOLL, G. : Eye Gaze Studies comparing Head-Up and Head-Down Displays in Vehicles. In: *Proceedings of IEEE International Conference on Multimedia & Expo, ICME 2007, Beijing, China, 2007*, S. 2250–2252, CD ROM
- [12] AEG ELECTROLUX: *Product LAVAMAT 84950*. <http://www.aeg-electrolux.de/>, letzter Zugriff: 25.02.2010, June 2008
- [13] AHRENS, A. : *Die Bewegungen der Augen beim Schreiben*. Universität Rostock, 1891

- [14] AJZEN, I. : The theory of planned behavior. In: *Organizational Behavior and Human Decision Processes* 50 (1991), S. 179–211
- [15] AL-HAMES, M. A.: *Graphische Modelle in der Mustererkennung.*, Technische Universität München, Lehrstuhl für Mensch-Maschine-Kommunikation, Diss., 2008
- [16] ALM, H. ; NILSSON, L. : Changes in driver behaviour as a function of handsfree mobile phones - A simulator study. In: *Accident Analysis and Prevention Vol. 26* (1994), S. 441–451
- [17] AMIR, A. ; FLICKNER, M. : *System and method for eye gaze tracking using corneal image mapping.* 2001
- [18] ASHMORE, M. ; DUCHOWSKI, A. T. ; SHOEMAKER, G. : Efficient Eye Pointing with a FishEye Lens. In: *Proceedings of Graphics Interface (GI 2005)*, 2005
- [19] ASHMORE, M. ; DUCHOWSKI, A. ; SHOEMAKER, G. : Efficient eye pointing with a fisheye lens. In: *Proceedings of Graphics Interface*, 2005
- [20] BADDELEY, A. ; HITCH, G. : Working memory. In: *Recent advances in learning and motivation.* New York : Academic Press 8 (1974)
- [21] *Kapitel* Word recognition processes in foveal and parafoveal vision: the range of influence of lexical variables. In: BALOTA, D. A. ; RAYNER, K. : *Basic Processes in Reading: Visual word recognition.* Hillsdale, NJ, England: Lawrence Erlbaum Associates, S. 198–232
- [22] BANAIYAN, A. ; MAHDIANI, H. ; FAKHRAI, S. : Cost-Performance Co-Analysis in VLSI Implementation of Existing and New Defuzzification Methods. In: *Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, 2005
- [23] BARDINS, S. ; POITSCHKE, T. ; KOHLBECHER, S. : Gaze-based Interaction in Various Environments. In: *Proceedings of 1st ACM International Workshop on Vision Networks for Behaviour Analysis VNBA 2008 (in conj. with ACM Multimedia 2008), Vancouver, Canada*, 2008
- [24] BARDINS, S. ; POITSCHKE, T. ; KOHLBECHER, S. ; WALLHOFF, F. ; SCHNEIDER., E. : A Framework for Gaze-Driven Interaction in Arbitrary Environments. In: *Proceedings of 1st International Workshop on Cognition for Technical Systems, Cotesys*, Technische Universität München, 6–7 October 2008
- [25] BATES, R. ; ISTANCE, H. : Zooming interfaces!: enhancing the performance of eye controlled pointing devices. In: *Proceedings of the 5th international ACM SIGACCESS Conference on Assistive Technologies*, 2002
- [26] BAUER, B. : *Konzeption und Aufbau eines Gaze-Tracker-Beleuchtungssystems.* Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, April 2010
- [27] BAUMGARTNER, G. ; BORNSCHEIN, H. ; HANITZSCH, R. ; JUNG, R. ; KORNHUBER, H. H. ; RENTSCHLER, I. ; SCHOBER, H. ; THODEN, U. : *Sehen. Sinnesphysiologie III.* München, Wien, Baltimore: Urban & Schwarzenberg, 1978
- [28] BAY, E. : *Optimierung der Genauigkeit eines Remote Gazetrackers für den Desktop Bereich.* Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Juni 2009
- [29] BAYES, T. : An Essay towards solving a problem in the Doctrine of Chances. In: *Philosophical Transaction of the Royal Society of London* 53 (1763), S. 370–418

-
- [30] BEINHAUER, W. : A widget library for gaze-based interaction elements. In: *Proceedings of the ACM Symposium on Eye Tracking Research & Applications - ETRA 2006*, 2006
- [31] BELLIFEMINE, F. ; CAIRE, G. ; TRUCCO, T. ; RIMASSA, G. : *JADE programmer's guide*, June 2007
- [32] BELLOTTI, F. ; GLORIA, A. de ; MONTANARI, R. ; DOSIO, N. : COMUNICAR: designing a multimedia, context-aware human-machine interface for cars. In: *Cognition, Technology & Work* 7 (2005), Nr. 1, S. 36–45
- [33] BERGE, B. : Liquid lens technology: principle of electrowetting based lenses and applications to imaging. In: *Proceedings of 18th IEEE International Conference on Micro Electro Mechanical Systems, MEMS*, 2005
- [34] BERNOTAT, R. : Anthropotechnik in der Fahrzeugführung. In: *Ergonomics* 13 (1970), S. 353–377
- [35] BERUFGENOSSENSCHAFT DER FEINMECHANIK UND ELEKTROTECHNIK: *Expositionsgrenzwerte für künstliche optische Strahlung. Berufsgenossenschaftliche Informationen für Sicherheit und Gesundheit bei der Arbeit*. 2004
- [36] BEYMER, D. ; FLICKNER, M. : Eye gaze tracking using an active stereo head. In: *In Proceedings of Computer Vision and Pattern Recognition*, 2003, S. 451–458
- [37] BEYONDUNREAL, INC: *BeyondUnreal – Covering Unreal Tournament 2004, Unreal Tournament 2003 & Gears of War*. <http://www.beyondunreal.com/>, letzter Zugriff: 15.11.2010,
- [38] BIERL, L. : *Das große MSP430 Praxisbuch*. Franzis Verlag, 2004
- [39] BOENING, G. ; BARTL, K. ; DERA, T. ; BARDINS, S. ; SCHNEIDER, E. ; BRANDT, T. : Mobile Eye Tracking as a Basis for Real-Time Control of a Gaze Driven Head-Mounted Video Camera. In: *Proceedings of ACM 2006 Symposium on Eye tracking Research & Application, ETRA 2006, San Diego, California, USA, ACM Press, NY, 2006*, S. 56
- [40] BOER, E. ; RAKAUSKAS, M. ; WARD, N. ; GOODRICH, M. : Steering entropy revisited. In: *Proceedings of the 3rd International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (2005), S. 25–32
- [41] BOFF, K. R. ; LINCOLN, J. E.: *Engineering Data Compendium, Human Perception and Performance*. Harry G Armstrong Aerospace, 1988
- [42] BOJKO, A. : Informative or Misleading? Heatmaps Deconstructed. In: *Lecture Notes in Computer Science*. 5610 (2009), S. 30–39
- [43] BONNAY, F. ; LEMARCHAND, C. ; ZALILA, Z. ; COFFIN, F. : Evaluation of a Fuzzy System-based Automotive Copilot Dedicated to Lateral Guidance / Renault Research Division - Vehicle Dynamics Group. 2001. – Forschungsbericht
- [44] BRATMAN, M. E.: *Intention, Plans, and Practical Reason*. Center for the Study of Language and Information, 1999
- [45] BREUER, J. : *Ergonomische Beurteilung und Gestaltung der Sicherheit des Arbeitssystems Kraftfahrzeugführen.*, Universität Düsseldorf, Diss., 1996. – Düsseldorf: VDI-Verl., 1996. VII
- [46] BROADBENT, D. E.: *Perception and Communication*. Elsevier Science Ltd, 1958
- [47] BRONSTEIN, I. N. ; SEMENDJAJEW, K. A. ; MUSIOL, G. ; MUEHLIG, H. : *Taschenbuch der Mathematik*. Wissenschaftlicher Verlag Harri Deutsch GmbH, 2008
- [48] BROOKE, J. : SUS - A quick and dirty usability scale / Redhatch Consulting Ltd. 1986. – Forschungsbericht

- [49] BROOKHUIS, K. ; VRIES, G. D. ; WAARD, D. D.: The effects of mobile telephoning on driving performance. In: *Accident Analysis and Prevention Vol. 23* (1991), S. 309–316
- [50] BROOKHUIS, K. ; WAARD, D. de: Assessment of Drivers' Workload: Performance and Subjective and Physiological Indexes. In: HANCOCK, P. (Hrsg.) ; DESMOND, P. (Hrsg.): *Stress, Workload and Fatigue*. Lawrence Erlbaum Associates, 2001, Kapitel 2.5
- [51] BROY, V. : *Benutzerzentrierte, graphische Interaktionsmetaphern für Fahrerinformationssysteme.*, Technische Universität München, Diss., 2007
- [52] BUBB, H. : Fahrerassistenz – primär ein Beitrag zum Komfort oder für die Sicherheit? In: *Der Fahrer im 21. Jahrhundert : Anforderungen, Anwendungen, Aspekte für Mensch-Maschine-Systeme* Bd. 1768 VDI-Gesellschaft Fahrzeug- und Verkehrstechnik, VDI-Verlag, 2003 (VDI-Berichte), S. 25–44
- [53] CANNY, J. : A Computational Approach to Edge Detection. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (1986), S. 679–698
- [54] CHIANG, H.-H. ; MA, L.-S. ; PERN, J.-W. ; WU, B.-F. ; LEE, T.-T. : Longitudinal and Lateral Fuzzy Control Systems Design for Intelligent Vehicles / IEEE. 2006. – Forschungsbericht
- [55] CHRIST, S. ; BAUR, A. : Die Rolle des Fahrers bei der Entwicklung neuer Infotainment- und Fahrerassistenzsysteme im Automobil. In: *Der Fahrer im 21. Jahrhundert: Anforderungen, Anwendungen, Aspekte für Mensch-Maschine-Systeme*. Gesellschaft Fahrzeug- und Verkehrstechnik, Verein Deutscher Ingenieure (VDI), 2003 (Tagung Braunschweig), S. S. 257–268
- [56] CINAR, M. O.: Eye Tracking Method to Compare the Usability of University Web Sites: A Case Study. In: *Proceedings of the 1st International Conference on Human Centered Design: Held as Part of HCI International 2009*, 2009
- [57] CRANE, H. D. ; STEELE, C. M.: Generation-V dual-Purkinje-image eyetracker. In: *Applied Optics* 24 (1985), S. 527 – 537
- [58] DAIMLER AG - RESEARCH & TECHNOLOGY: *Lane Change Test 1.2 - User Guide*. Juli 2004
- [59] DALZEL-JOB, S. ; NICOL, C. ; OBERLANDER jon: Comparing behavioural and self-report measures of engagement with an embodied conversational agent: a first report on eye tracking in Second Life. In: *Proceedings of ACM 2008 Symposium on Eye tracking Research & Application, ETRA 2008, Savannah, Georgia, USA*. New York, NY, USA : ACM, 2008. – ISBN 978-1-59593-982-1, S. 83–85
- [60] DAUNYS G. ET. AL: Report on New Approaches to Eye Tracking / COGAIN Communication by Gaze Interaction. 2006. – Forschungsbericht
- [61] DERA, T. ; BOENING, G. ; BARDINS, S. ; SCHNEIDER, E. ; BRANDT, T. : Low-latency video tracking of horizontal, vertical, and torsional eye movements as a basis for 3DOF realtime motion control of a head-mounted camera. In: *Proceedings of the IEEE Conference on Systems, Man and Cybernetics (SMC2006), Taipei, Taiwan*, 2006
- [62] Kapitel State of the art in adaptive user interfaces. In: DIETERICH, H. ; MALINOWSKI, U. ; KÜHME, T. ; SCHNEIDER-HUFSCHMIDT, M. : *Adaptive user interfaces: Principles and practice*. North-Holland, Amsterdam, 1993, S. 13–48
- [63] DODGE, R. ; CLINE, T. S.: The angle velocity of eye movements. In: *Psychological Review* 8 (1901), S. 145–157

- [64] DONGES, E. : Das Prinzip Vorhersehbarkeit als Auslegungskonzept für Maßnahmen zur Aktiven Sicherheit Maßnahmen zur Aktiven Sicherheit. In: *Das Mensch-Maschine System im Verkehr, VDI-Berichte 948* (1992)
- [65] DONMEZ, B. ; BOYLE, L. ; LEE, J. ; McGEHEE, D. : SAVE-IT: SAfety VEhicles using adaptive Interface Technology (Task 4), A Literature Review of Distraction Mitigation Strategies / University of Iowa. 2004. – Forschungsbericht
- [66] DONMEZ, B. ; BOYLE, L. ; LEE, J. ; McGEHEE, D. : SAVE-IT: SAfety VEhicles using adaptive Interface Technology (Task 4), Experiments for Distraction Mitigation Strategies / University of Iowa. 2004. – Forschungsbericht
- [67] DORNHÖFER, S. M. ; PANNASCH, S. ; UNEMA, P. J.: Augenbewegungen und deren Registrierungsmethoden. / Applied Cognitive Research Unit, Institute of Psychology III, TU Dresden. – <http://rcswww.urz.tu-dresden.de/~cogsci/pdf/augenbewegungen.pdf>, letzter Zugriff: 15.10.2010
- [68] DRAGUTINOVIC, N. ; TWISK, D. : Use of mobile phones while driving - Effects on road safety / SWOV Institute for Road Safety Research, The Netherlands. 2006 (SWOV report R-2005-12). – Forschungsbericht
- [69] DREFERS, E. ; HACKNER, C. : *Implementierung und Visualisierung eines remote Gazetrackers*. Bachelorarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, September 2008
- [70] DUCHOWSKI, A. T.: *Eye Tracking Methodology: Theory & Practice*. Springer New York, 2007
- [71] DUMKE, D. : *Simulation und Evaluierung von verschiedenen Fahrphysikeinstellungen im Fahrsimulator*. Studienarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, November 2007
- [72] DUMKE, D. : *Intelligentes Fahrerassistenzsystem basierend auf einem Multiagentensystem zur Planung von Fahrstrategien*. Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Juni 2008
- [73] EBISAWA, Y. ; SATOH, S. : Effectiveness of pupil area detection technique using two light sources and image difference method. In: *Proceedings of 5th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1993, S. 1268–1269.
- [74] EILERS, K. ; NACHREINER, F. ; HÄNECKE, K. : Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. In: *Zeitschrift für Arbeitswissenschaft* 40 (1986), S. 215–224
- [75] EMSLEY, H. H.: *Visual Optics*. Hatton Press Limited, London, 1976
- [76] EPIC GAMES, INC.: *Make Something Unreal Contest*. <http://www.makesomethingunreal.com/>, letzter Zugriff: 15.11.2010,
- [77] EPIC GAMES, INC.: *Unreal Tournament 2004*. <http://www.unrealtournament2004.com/ut2004/>, letzter Zugriff: 15.11.2010,
- [78] ERKE, H. : Vorgaben an die Verkehrsplanung - Anforderungen an den Menschen aus der Sicht der ökologischen Psychologie. In: *Der Mensch im Straßenverkehr. Referate der Fünften Informationsmedizinischen Tage in Hamburg 1991. Band 6. Schriftenreihe der Hamburg-Mannheimer-Stiftung für Informationsmedizin*. Stuttgart: Ferdinand Enke Verlag. (1993)
- [79] FACHVERBAND FÜR STRAHLENSCHUTZ E.V: *Leitfaden Nichtionisierende Strahlung - Sichtbare und infrarote Strahlung*. 1998-2000

- [80] FALKMER, T. ; NILSSON, L. ; TOEMROS, J. : Detection and identification of information presented peripherally inside the car. Effects of driving task demands, stimulus position and direction of motion of the stimulus. VTI Rapport 461A 2000 / Linköping: Swedish National Road and transport Research Institute. 2000. – Forschungsbericht
- [81] FASTL, H. ; ZWICKER, E. : *Psychoacoustics: Facts and Models*. Springer, Berlin, 2010
- [82] FEATHER, N. T. ; NEWTON, J. W.: Values, expectations, and the prediction of social action: An expectancy-valence analysis. In: *Motivation and Emotion* 6 (1982), September, Nr. 3, S. 217–244
- [83] FITTS, P. M.: The information capacity of the human motor system in controlling the amplitude of movement. In: *Journal of Experimental Psychology*. 47 (1954), S. 381–391
- [84] FORSMAN, A. ; NILSSON, L. ; TÖRNOS, J. ; (Ö)STLUND, J. : Effects of cognitive and visual load in real and simulated driving / VTI Swedish National Road and Transport Research Institute. 2006 (VTI report 533A). – Forschungsbericht
- [85] FORSTNER, M. : *Nutzeraufmerksamkeitslenkung im Fahrzeug*. Bachelorarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Dezember 2008
- [86] FRESE, M. ; BRODBECK, F. C.: *Computer in Büro und Verwaltung : psychologisches Wissen für die Praxis*. Springer Verlag Berlin, 1989
- [87] GABAY, C. ; BERGE, B. ; DOVILLAIRE, G. ; BUCOURT, S. : Dynamic study of a Varioptic variable focal lens. In: *Proceedings of Current Developments in Lens Design and Optical Engineering III*, 2002
- [88] GAST, J. : *Generierung einer multifunktionalen Simulatorplattform zur Repräsentation und Evaluierung von visionären Interaktionskonzepten in der automotiven Domäne*. Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Januar 2006
- [89] GEIB, C. W.: Problems in Intent Recognition for Elder Care. In: *Proceedings of the AAAI-02 Workshop - Automation as Caregiver.*, 2002
- [90] GEISER, G. : Mensch-Maschine-Kommunikation im Kraftfahrzeug. In: *Automobiltechnische Zeitschrift ATZ-87* (1985), Nr. 2, S. 77–84
- [91] GEISER, G. : *Mensch-Maschine-Kommunikation*. Oldenbourg-Verlag München/Wien, 1990. – ISBN: 3486215051
- [92] GILULA, J. O.: Über Droh- und Orientierungsreflexe. In: *European Archives of Psychiatry and Clinical Neuroscience* 79 (1927), Dezember, Nr. 1, S. 407–417
- [93] GOLLWITZER, P. M. ; BARGH, J. A.: The Volitional Benefits of Planning. In: *The Psychology of Action: Linking Cognition and Motivation of Behavior*. (1996), S. 287–312
- [94] GOSS, D. A. ; WEST, R. W.: *Introduction to the Optics of the Eye*. 1. Auflage. Butterworth-Heinemann, 2002. – ISBN: 978-0750673464
- [95] GRÜNDL, M. : *Fehler und Fehlverhalten als Ursache von Verkehrsunfällen und Konsequenzen für das Unfallvermeidungspotenzial und die Gestaltung von Fahrerassistenzsystemen.*, Lehrstuhl für Experimentelle und Angewandte Psychologie, Universität Regensburg, Diss., 2005
- [96] HAIGNEY, D. ; TAYLOR, R. ; WESTERMAN, S. J.: Concurrent mobile (cellular) phone use and driving performance: task demand characteristics and compensatory processes. In: *Transport Research Part F: Traffic Psychology and Behaviour*, Vol. 3 (2000), S. 113–121

- [97] HAMBERGER, W. ; MAUTER, G. : Audi Multi Media Interface (MMI): Neue Spezifikationsmethoden zur interdisziplinären Bedienkonzeptentwicklung. In: *Der Fahrer im 21. Jahrhundert: Anforderungen, Anwendungen, Aspekte für Mensch-Maschine-Systeme*. Gesellschaft Fahrzeug- und Verkehrstechnik, Verein Deutscher Ingenieure (VDI), 2003 (Tagung Braunschweig), S. 217–233
- [98] HANSEN, D. W. ; SKOVSGAARD, H. H. ; HANSEN, J. P. ; MOLLENBACH, E. : Noise tolerant selection by gaze-controlled pan and zoom in 3D. In: *Proceedings of the ACM Symposium on Eye Tracking Research & Applications - ETRA 2008*, 2008
- [99] HANSEN, J. P. ; TØRNING, K. ; JOHANSEN, A. S. ; ITOH, K. ; AOKI, H. : Gaze typing compared with input by head and hand. In: *ETRA '04: Proceedings of the 2004 symposium on Eye tracking research & applications*. New York, NY, USA : ACM, 2004. – ISBN 1–58113–825–3, S. 131–138
- [100] HANSIS, B. : *Anwendung von GazeTracking im Desktopbereich*. Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, September 2009
- [101] HASSENZAHN, M. ; TRACTINSKY, N. : User Experience - a research agenda. In: *Behavior & Information Technology* 25 (2006), S. 91–97
- [102] *Kapitel Derivation of Refraction Formulas*. In: HECKBERT, P. : *Introduction to Ray Tracing*. Academic Press, London, 1989, S. 288–293
- [103] HECKHAUSEN, H. : *Motivation und Handeln*. 3. Auflage. Springer-Verlag Berlin und Heidelberg, 2009
- [104] HEINECKE, A. M.: *Mensch-Computer-Interaktion*. Hanser Fachbuchverlag, 2004
- [105] HEINZE, C. : *Modelling Intention Recognition for Intelligent Agent Systems*, Air Operations Division Systems Sciences Laboratory, Diss., 2004
- [106] HENNESSEY, C. ; NOUREDDIN, B. ; LAWRENCE, P. D.: A single camera eye-gaze tracking system with free head motion. In: *Proceedings of ACM 2006 Symposium on Eye tracking Research & Application*, 2006
- [107] HEWETT, T. ; BAECKER, R. ; CARD, S. ; CAREY, T. ; GASEN, J. ; MANTEL, M. ; PERLMAN, G. ; STRONG, G. ; VERPLANK, W. : *Curricula for Human-Computer Interaction*. <http://www.sigchi.org/cdg/index.html>. Version: 1996. – ISBN 0-89791-474-0
- [108] HIEMISCH, A. : *Der Begriff der Handlung in der Psychologie*. www.widerstreit-sachunterricht.de Ausgabe Nr. 12, http://www2.hu-berlin.de/wsu/ebeneIII/begriffe/hand_hie.pdf, letzter Zugriff: 25.02.2010, März 2009
- [109] HOEDEMAEKER, M. ; RIDDER, S. de ; JANSSEN, W. : Review of European Human Factors Research on Adaptive Interface Technologies for Automobiles / TNO Human Factors. 2002. – Forschungsbericht
- [110] HOFMANN, M. : *Intentionsbasierte, maschinelle Interpretation von Benutzeraktionen*., Technische Universität München, Lehrstuhl für Mensch-Maschine Kommunikation, Diss., 2003
- [111] HOROWITZ, T. ; HOLCOMBE, A. ; WOLFE, J. ; ARSENIO, H. ; DIMASE, J. : Attentional pursuit is faster than attentional saccade. In: *Journal of Vision* 4 (2004), S. 585–603
- [112] HUCKAUF, A. ; URBINA, M. : Gazing with pEYE: new concepts in eye typing. In: *Proceedings of the 4th ACM symposium on Applied perception in graphics and visualization APGV*, 2007
- [113] HYRSKYKARI, A. ; MAJARANTA, P. ; AALTONEN, A. ; RÄIHÄ, K.-J. : Design Issues of iDict: a gaze-assisted translation aid. In: *Proceedings of ACM 2000 Symposium on Eye tracking Research & Application, ETRA 2000*, 2000

- [114] HYRSKYKARI, A. ; MAJARANTA, P. ; RÄIHÄ, K.-J. : From Gaze Control to Attentive Interfaces. In: *Proceedings of International Conference on Human-Computer Interaction HCI 2005*, 2005
- [115] IQBAL, S. : MeWS-IT: A Mental Workload Based System for Interruption Timing. In: *Proceedings of the ACM Symposium on User Interface Software and Technology*, 2005
- [116] ISOKOSKI, P. : Text Input Methods for Eye Trackers Using Off-Screen Targets. In: *Proceedings of the ACM Symposium on Eye Tracking Research & Applications - ETRA 2000*, 2000, S. 15–22
- [117] JACOB, R. J. K.: The use of eye movements in human-computer interaction techniques: What you look at is what you get. In: *ACM Transactions on Information Systems (TOIS) - Special issue on computer—human interaction* 9 (1991), S. 152–169
- [118] JAIN, R. ; KASTURI, R. ; SCHUNCK, B. G.: *Machine Vision*. McGraw-Hill Science/Engineering/Math, 1995
- [119] JAMSON, A. H. ; MERAT, N. : Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. In: *Transport Research Part F: Traffic Psychology and Behaviour*, Vol. 8 (2005), S. 79–96
- [120] JAVAL, L. E.: Essai sur la physiologie de la lecture. In: *Annales d'Oculistique* 82 (1879), S. 242–253
- [121] JENSEN, F. V.: *An Introduction to Bayesian Networks*. Springer Verlag, 1997
- [122] JUNG, R. : Eine Elektrische Methode zur Mehrfachen Registrierung von Augenbewegungen und Nystagmus. In: *Journal of Molecular Medicine* 18 (1939), S. 21–24
- [123] KAMAL, M. ; KAWABE, T. ; MURATA, J. ; MUKAI, M. : Driver-adaptive assist system for avoiding abnormality in driving. In: *IEEE Transactions on Control Applications* (2007), S. 1247–1252
- [124] Kapitel Looking at looking: The amorphous fixation measure. In: KARSH, R. ; BREITENBACH, F. W.: *Eye movements and psychological functions: International Views*. Hillsdale, NJ: Erlbaum, 1983, S. 53–64
- [125] KEMBEL, J. A.: Reciprocal eye contact as an interaction technique. In: *Proceeding of CHI extended abstracts on Human factors in computing systems*, 2003
- [126] KHALIA, M. ; SCHWARZMANN, E. : *Analyse, Bereitstellung und Verarbeitung von Kontextdaten im Fahr Simulator*. Bachelorarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, M. Ablastmeier, November 2006
- [127] KIM, S. J.: *Robust Pupil Tracking for Optimization of a Gaze Tracking System*. Mastersthesis, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Dezember 2008
- [128] KOBAYASHI, M. ; IGARASHI, T. : Considering the direction of cursor movement for efficient traversal of cascading menus. In: *Proceedings of the 16th Annual ACM Symposium on User interface Software and Technology. UIST 2003*, 2003
- [129] KOHLBECHER, S. ; BARDINS, S. ; BARTL, K. ; SCHNEIDER, E. ; POITSCHKE, T. ; ABLASSMEIER, M. : Calibration-free eye tracking by reconstruction of the pupil ellipse in 3D space. In: *Proceedings of ACM 2008 Symposium on Eye tracking Research & Application, ETRA 2008, Savannah, Georgia, USA*, ACM Press, NY, March 2008, S. 135–138
- [130] KOHLBECHER, S. ; POITSCHKE, T. ; ABLASSMEIER, M. ; RIGOLL, G. ; BARDINS, S. ; SCHNEIDER, E. : Gaze vector detection by stereo reconstruction of the pupil contours. In: *Journal of Eye Movement Research* 1 (2007), S. 121. – Conference Abstracts of European Conference on Eye Movements, ECEM 2007, Potsdam, 19.-23.08.2007 ISSN 1995-8692.

- [131] KOMMDESIGN: *Kommdesign Homepage*. <http://www.kommdesign.de>, letzter Zugriff: 15.08.2010, 2010
- [132] KORN, A. : *Bildverarbeitung durch das visuelle System*. Springer-Verlag, Berlin, 1982
- [133] KUHLE, J. : *Motivation und Persönlichkeit: Interaktionen psychischer Systeme*. Hogrefe-Verlag, 2001
- [134] KUMAR, M. : *Gaze-Enhanced User Interface Design.*, Stanford University, Diss., 2007
- [135] LAND, M. F.: Eye movements and the control of actions in everyday life. In: *Progress in Retinal and Eye Research* 25 (2006), Mai, Nr. 3, S. 296–324
- [136] LAND, M. F. ; MENNIE, N. ; RUSTED, J. : The roles of vision and eye movements in the control of activities of daily living. In: *Perception*. 28 (1999), S. 1311–1328
- [137] LANGE, C. : *Wirkung von Fahrerassistenz auf der Führungsebene in Abhängigkeit der Modalität und des Automatisierungsgrades*. München, Technische Universität München, Dissertation, 2008
- [138] LANKFORD, C. : Effective eye-gaze input into Windows. In: *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*. New York, NY, USA : ACM, 2000. – ISBN 1-58113-280-8, S. 23–27
- [139] LANKFORD, C. : Effective eye-gaze input into Windows. In: *Proceedings of the ACM Symposium on Eye Tracking Research & Applications - ETRA 2000*, 2000, S. 23–27
- [140] LAQUAL, F. ; ABLASSMEIER, M. ; POITSCHKE, T. ; RIGOLL, G. : Using 3D Touch Interaction for a Multimodal Zoomable User Interface. In: *Proceedings of International Conference on Human-Computer Interaction HCI 2009, San Diego, USA, 2009*
- [141] LC TECHNOLOGIES, INC.: *Books written with the Eyegaze System*. <http://www.eyegaze.com/content/books-written-eyegaze-system>, letzter Zugriff: 15.11.2010,
- [142] LEE, J. ; REYES, M. ; MCGEHEE, D. : SAVE-IT: SAFETY VEHICLES using adaptive Interface Technology (Task 5), A Literature Review of Cognitive Distraction. / The University of Iowa. 2004. – Forschungsbericht
- [143] LEE, J. ; REYES, M. ; SMYSER, T. ; LIANG, Y. ; THORNBURG, K. : SAVE-IT: SAFETY VEHICLES using adaptive Interface Technology (Task 5), Final Report: Phase 1 / The University of Iowa. 2004. – Forschungsbericht
- [144] LEVEN, W. : *Blickverhalten von Konsumenten. Grundlagen, Messung und Anwendung in der Werbeforschung*. Physica-Verlag Heidelberg, 1991
- [145] MACDONALD, W. ; HOFFMANN, E. : Review of relationships between steering wheel reversal rate and driving task demand. In: *Human Factors Vol. 22* (1980), S. 733–739
- [146] MAGLIO, P. ; MATLOCK, T. ; CAMPBELL, C. ; ZHAI, S. ; SMITH, B. : Gaze and speech in attentive user interfaces. In: *Proceedings of the Third International Conference on Multimodal Interfaces*, 2000
- [147] MAJARANTA, P. ; RÄIHÄ, K.-J. : Twenty years of eye typing: systems and design issues. In: *Proceedings of the 2002 symposium on Eye tracking research & applications, ETRA*. New York, NY, USA : ACM, 2002. – ISBN 1-58113-467-3, S. 15–22
- [148] MAMDANI, E. : Application of fuzzy algorithms for control of simple dynamic plant. In: *Proc. IEE Vol. 121*, 1974

- [149] MANN, P. R. ; MAY, T. ; HORNBERG, B. ; WALOSSEK, E. : *Method for access control to resources of an application in a computer system*. <http://www.freepatentsonline.com/EP1482394A2.html>. Version: Dezember 2004
- [150] MARTENS, M. ; WINSUM, W. van: *Measuring distraction: the Peripheral Detection Task / TNO Human Factors*. 2000. – Forschungsbericht
- [151] MARUBENI CORPORATION: *Datasheet FL850-03-80*. www.marubeni.com, letzter Zugriff: 15.11.2010, 2009
- [152] MATTHEWS, G. ; DESMOND, P. : *Stress and Driving Performance: Implications for Design and Training*. In: HANCOCK, P. (Hrsg.) ; DESMOND, P. (Hrsg.): *Stress, Workload and Fatigue*. Lawrence Erlbaum Associates, 2001, Kapitel 1.8
- [153] MCGLAUN, G. ; ALTHOFF, F. ; SCHULLER, B. ; LANG, M. : *A New Technique for Adjusting Distraction Moments in Multitasking Non-field Usability Tests*. In: *Proceedings of the International Conference on Human Factors in Computing Systems, CHI2002, Minneapolis, MN, USA, 2002*
- [154] MCGUFFIN, M. ; BALAKRISHNAN, R. : *Acquisition of Expanding Targets*. In: *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI), 2002, S. 57–64*
- [155] MERCHANT, J. ; MORRISSETTE, R. ; PORTERFIELD, J. L.: *Remote measurement of eye direction allowing subject motion over one cubic foot of space*. In: *IEEE Transactions on Biomedical Engineering* 21 (1974), S. 309–317
- [156] MESSNER, U. ; OMANN, D. : *Verkehrsunfallstatistik 2007, Kuratorium für Verkehrssicherheit*. Mai 2008
- [157] MICHON, J. ; PIERSMA, E. ; SMILEY, A. ; VERWEY, W. ; WEBSTER, E. : *Design considerations*. In: MICHON, J. (Hrsg.): *Generic Intelligent Driver Support, A Comprehensive Report on GIDS*. Taylor & Francis, 1993, Kapitel 5
- [158] *Kapitel A critical view of driver behavior models: What do we know, what should we do?* In: MICHON, J. A.: *Human behavior and traffic safety*. New York: Plenum Press, 1985., 1985, S. 485–520
- [159] *Kapitel Blickbewegungsforschung. Einführung in die Physiologischen Grundlagen, Techniken und in die Problem- und Anwendungsbereiche*. In: MICKASCH, H. D. ; HAACK, J. : *Blickbewegung und Bildverarbeitung. Kognitionspsychologische Aspekte visueller Informationsverarbeitung*. Frankfurt am Main/New York: Peter Lang, 1986, S. 11–36
- [160] MILICIC, N. : *Sichere und ergonomische Nutzung von Head-Up Displays im Fahrzeug.*, Technische Universität München, Diss., 2010
- [161] MINIOTAS, D. ; ŠPAKOV, O. ; MACKENZIE, I. S.: *Eye gaze interaction with expanding targets*. In: *Proceeding of ACM CHI 2004 extended abstracts on Human factors in computing systems, 2004*
- [162] MINIOTAS, D. ; SPAKOV, O. ; TUGOY, I. ; MACKENZIE, S. I.: *Speech-Augmented Eye Gaze Interaction with Small Closely Spaced Targets*. In: *Proceedings of the ACM Symposium on Eye Tracking Research & Applications - ETRA 2006, 2006, S. 67–72*
- [163] *Kapitel Aufmerksamkeit*. In: MÜLLER, H. ; KRUMMENACHER, J. : *Lehrbuch Allgemeine Psychologie*. Spektrum Akademischer Verlag, Heidelberg, 2002, S. 120–132
- [164] MÜLLER, M. : *Vorlesung Auge*. Version: WS 2004/2005. <http://wwwuser.gwdg.de/~mmuelle7/Skripten/Auge1-WS04-05.pdf>, letzter Zugriff: 15.08.2010

- [165] MOWRER, O. R. ; RUCH, R. C. ; MILLER, N. E.: The corneoretinal potential difference as the basis of the galvanometric method of recording eye movements. In: *American Journal of Physiology* 114 (1936), S. 423
- [166] MUIGG, A. : *Implizites Workloadmanagement. Konzept einer zeitlich-situativen Informationsfilterung im Automobil.*, Technische Universität München, Diss., 2009
- [167] MULLER, H. J. ; RABBITT, P. M. A.: Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. In: *Journal of Experimental Psychology: Human Perception and Performance* 15 (1989), S. 315–330
- [168] NAKAYAMA, O. ; FUTAMI, T. ; NAKAMURA, T. ; BOER, E. : Development of a Steering Entropy Method for Evaluating Driver Workload. In: *Society of Automotive Engineers Technical Paper Series: 1999-01-0892* (1999)
- [169] NCSDR/NHTSA: Drowsy Driving and Automobile Crashes. / National Center on Sleep Disorders Research und National Highway Traffic Safety Administration. 1998. – Forschungsbericht
- [170] NGHIEM, S. L.: *Integration eines dynamischen Fremdverkehrs zur gezielten Induktion gewünschter Verkehrssituationen im Fahrsimulator.* Bachelorarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Oktober 2006
- [171] NGHIEM, S. L.: *Agent-based driver intention and abnormality modeling for intelligent drive assistance systems.* Master Thesis, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Juni 2008
- [172] NIENDIECK, S. : *Optimierung von Fuzzy-Controllern*, Universität Münster, Diss., 2003
- [173] OHASHI, K. ; YAMAGUCHI, T. ; TAMAI, I. : Humane Automotive System Using Driver Intention Recognition. In: *Proceedings of IEEE SICE 2004 Annual Conference* Bd. 2, 2004, S. 1164–1167
- [174] OHM, J. : *Zur Tätigkeit des Augenmuskelsenders.* Postberg: Bottrop. (1928)
- [175] OHNO, T. : Features of eye gaze interface for selection tasks. In: *Proceedings of the Third Asian Pacific Computer and Human Interaction.*, 1998
- [176] OHNO, T. ; MUKAWA, N. ; YOSHIKAWA, A. : FreeGaze: a gaze tracking system for everyday gaze interaction. In: *ETRA '02: Proceedings of the 2002 symposium on Eye tracking research & applications.* New York, NY, USA : ACM, 2002. – ISBN 1-58113-467-3, S. 125–132
- [177] OLSEN, S. : *Im Dialog mit dem Auto: Neueste Trends bei der Sprachsteuerung.* http://www.zdnet.de/arbeitsplatzrechner_peripherie_in_unternehmen_im_dialog_mit_dem_auto_neuste_trends_bei_der_sprachsteuerung_story-20000005-39141996-1.htm, letzter Zugriff: 25.02.2010, März 2006
- [178] OLSSON, S. ; BURNS, P. : Measuring Driver Visual Distraction with a Peripheral Detection Task / Department of Education & Psychology, Linköping University, Sweden und Volvo Technological Development Corporation, Gothenburg, Sweden. 2000. – Forschungsbericht
- [179] PASHLER, H. : *Attention.* Taylor & Francis Press, Philadelphia, 1988
- [180] PAVLIĆ, M. : *Blickbasierte Bildschirmlibrierung in 3D und Transformation des Gaze-Vektors in eine Bildebene.* Bachelorarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, September 2008
- [181] PAVLIĆ, M. : *Konzeption eines irisrandbasierten 3D Gazetracking-Systems.* Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Juni 2009

- [182] PEARL, J. : *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman Publ Inc, 1997
- [183] PEARSON, K. : On lines and planes of closest fit to systems of points in space. In: *Philosophical Magazine* 2 (1901), S. 559–572
- [184] PETERS, B. ; ÖSTLUND, J. ; THORSLUND, B. ; ENGSTRÖM, J. ; MARKKULA, G. ; KEINATH, A. ; HORST, D. ; JUCH, S. ; MATTES, S. ; FOEHL, U. : Evaluation and assessment methodology: Driver workload and distraction assessment methods and tools: Using performance indicators for driver state assessment / AIDE. 2004 (D 2.2.5). – Forschungsbericht
- [185] PFEUFER, B. : *Intentionsbasierte Planerkennung in der Bedienung von Mensch-Maschine-Schnittstellen im Fahrzeug*. Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Oktober 2009
- [186] POINT GREY RESEARCH, INC.: *Products: FireFly MV CMOS Camera*. <http://www.ptgrey.com/products/fireflymv>, letzter Zugriff: 15.11.2010,
- [187] POITSCHKE, T. ; BAY, E. ; DREFERS, E. ; LEINSLE, H. ; PAVLIC, M. ; RIGOLL, G. : *3D Gaze Tracking für die Mensch-Maschine-Interaktion*. Technischer Bericht, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, 2009
- [188] POITSCHKE, T. ; ABLASSMEIER, M. ; REIFINGER, S. ; RIGOLL, G. : A Multifunctional VR-Simulator Platform for the Evaluation of Automotive User Interfaces. In: JACKO, J. (Hrsg.): *Proceedings of International Conference on Human-Computer Interaction HCI 2007, Beijing, China, Proc. Part IV* Bd. LNCS 4553, Springer Verlag Heidelberg, 22-27 July 2007, S. 1120–1129
- [189] POITSCHKE, T. ; ABLASSMEIER, M. ; RIGOLL, G. ; BARDINS, S. ; KOHLBECHER, S. ; SCHNEIDER, E. : Contact-analog Information Representation in an Automotive Head-Up Display. In: *Proceedings of ACM 2008 Symposium on Eye tracking Research & Application, ETRA 2008, Savannah, Georgia, USA*, ACM Press, NY, March 2008, S. 119–122
- [190] POITSCHKE, T. ; BAY, E. ; LAQUAI, F. ; RIGOLL, G. ; BARDINS, S. ; BARTL, K. ; VOCKEROTH, J. ; SCHNEIDER, E. : Using Liquid-Lenses to extend the Operating Range of a Remote Gaze Tracking System. In: *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, SMC 2009*, 2009
- [191] POITSCHKE, T. ; LAQUAI, F. ; RIGOLL, G. : Agent-Based Driver Abnormality Estimation. In: *Proceedings of International Conference on Human-Computer Interaction HCI 2009, San Diego, USA*, 2009
- [192] POITSCHKE, T. ; LAQUAI, F. ; RIGOLL, G. : Guiding a Driver's Visual Attention using Graphical and Auditory Animations. In: *Proceedings of International Conference on Human-Computer Interaction HCI 2009, San Diego, USA*, 2009
- [193] POITSCHKE, T. ; LAQUAI, F. ; STAMBOLIEV, S. ; RIGOLL, G. : Gaze-based interaction on multiple displays in an automotive environment. In: *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, SMC 2011*, 2011
- [194] POITSCHKE, T. ; WALLHOFF, F. ; RIGOLL, G. ; BARDINS, S. ; SCHNEIDER, E. ; BRANDT, T. : Eyetracker-Based Driver State and Intention Recognition for Realtime Control and Configuration of Human-Machine Interfaces in Vehicles. / DFG, Cognition for Technical Systems - CoTeSys, Cluster of Excellence. 2007. – CoTeSys Progress Report for First Funding Period.
- [195] POITSCHKE, T. ; WALLHOFF, F. ; RIGOLL, G. ; BARDINS, S. ; SCHNEIDER, E. ; BRANDT, T. : Progress Report for CoTeSys Project (323), RealEYE - REmote And Low-latency EYEtracking. / DFG, Cognition for Technical Systems - CoTeSys, Cluster of Excellence. 2008. – CoTeSys Progress Report for Second Funding Period.

- [196] POITSCHKE, T. ; WALLHOFF, F. ; RIGOLL, G. ; BARDINS, S. ; SCHNEIDER, E. ; BRANDT, T. : Progress Report for CoTeSys Project (323), RealEYE - REmote And Low-latency EYEtracking. / DFG, Cognition for Technical Systems - CoTeSys, Cluster of Excellence. 2009. – CoTeSys Progress Report for Third Funding Period.
- [197] PORTERFIELD, W. : An essay concerning the motions of our eyes. PartI. Of their external motions. In: *Edinburgh medical essays and observations* 3 (1737), S. 160–263
- [198] RAMMELT, P. : *Probabilistische Netze*. Unveröffentlichte Arbeit, Technische Universität Berlin, 2005
- [199] RANNEY, T. A. ; MAZZAE, E. ; GARROTT, R. ; GOODMAN, M. J.: NHTSA Driver Distraction Research: Past, Present, and Future. / National Highway Traffic Safety Administration. 2000. – Forschungsbericht
- [200] RASMUSSEN, J. : Skills, rules, and knowledge: signals, signs, and symbols, and other distinctions in human performance models. In: *IEEE Transactions on Systems, Man and Cybernetics*. 13 (1983), Nr. 3, S. 257–266
- [201] RASSL, R. : *Ablenkungswirkung tertiärer Aufgaben im Pkw – Systemergonomische Analyse und Prognose*, Technische Universität München, Diss., 2004
- [202] REIFINGER, S. ; WALLHOFF, F. ; ABLASSMEIER, M. ; POITSCHKE, T. ; RIGOLL, G. : Static and Dynamic Hand-Gesture Recognition for Augmented Reality Applications. In: JACKO, J. (Hrsg.): *Proceedings of International Conference on Human-Computer Interaction HCI 2007, Beijing, China, Proc. Part III* Bd. LNCS 4552, Springer Verlag Heidelberg, 22-27 July 2007, S. 728–737
- [203] ROBINSON, D. A.: A method of measuring eye movement using a cleral search coil in a magnetic field. In: *IEEE Transactions on Biomedical Electronics* 10 (1963), S. 137–145
- [204] ROSPA - THE ROYAL SOCIETY FOR THE PREVENTION OF ACCIDENTS: *Driver Fatigue Policy Statements*. 2001
- [205] SANTANA-DIAZ, A. ; HERNANDEZ-GRESS, N. ; ESTEVE, D. ; JAMMES, B. : Discriminating sensors for driver's impairment detection. In: *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine & Biology* (2000)
- [206] SCHANDRY, R. : *Lehrbuch Psychophysiologie - Körperliche Interaktionen psychischen Geschehens*. Psychologie Verlags Union, München & Weinheim, 1989
- [207] SCHELLER, B. : *Sample-basierte Simulation von Fahrzeuginnengeräuschen und Untersuchung der Auswirkung derer auf den Aufmerksamkeitsfokus und das Fahrverhalten des Fahrers*. Diplomarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, November 2008
- [208] SCHENK, J. ; RIGOLL, G. : *Mensch-Maschine Kommunikation 1 - Skriptum zur Vorlesung*. Lehrstuhl für Mensch-Maschine Kommunikation, Technische Universität München, 2009
- [209] SCHENK, J. ; RIGOLL, G. : *Mensch-Maschine-Kommunikation*. 1. Auflage. Springer, 2010. – ISBN 978-3-642-05456-3
- [210] SCHNEIDER, E. : Möglichkeiten der Registrierung von Augenbewegungen. / Neurologische Poliklinik, Klinikum der Universität München Großhadern. 2006. – http://www.forbias.de/papers/dgkn/schneider_dgkn_2006.ppt, letzter Zugriff: 15.11.2010
- [211] SCHNEIDER, E. ; BARDINS, S. : *Mobile Eye Tracking as a Basis for Real-Time Control of a Gaze Driven Head- Mounted Video Camera.*, Lawrence Erlbaum Associates, Inc., 2006, S. CD-ROM

- [212] SCHNEIDER, E. ; EGGERT, T. : Methoden der Augenbewegungsmessung. / Neurologische Poliklinik, Klinikum der Universität München Großhadern. 2007. – <http://www.forbias.de/papers/dgkn06.pdf>, letzter Zugriff: 15.11.2010
- [213] SCHNEIDER, E. ; KOHLBECHER, S. ; VILLGRATTNER, T. ; BARTL, K. ; BARDINS, S. ; POITSCHKE, T. ; ULBRICH, H. ; BRANDT, T. : Vision System for Wearable and Robotic Uses. In: *Proceedings of IEEE 17th International Symposium on Robot and Human Interactive Communication, RO-MAN 2008, Munich, Germany*, 2008, S. 53–58
- [214] SCHNEIDER, M. : *Physiologie des Menschen*. 15. Auflage, Revidierter Neudruck. Springer-Verlag, Berlin, Heidelberg, New York, 1966
- [215] SCHNEIDER-HUFSCHMIED, M. : *Entwurf ergonomischer Benutzungsoberflächen*. Vorlesungsskript, Technische Universität München, 2010
- [216] SCHOTT, E. : Über die Registrierung des Nystagmus und anderen Augenbewegungen mittels des Seitengalvanometers. In: *Deutsches Archiv für Klinische Medizin* 140 (1922), S. 79–90
- [217] *Kapitel Speech Communication and Multimodal Interfaces*. In: SCHULLER, B. ; ABLASSMEIER, M. ; MÜLLER, R. ; REIFINGER, S. ; POITSCHKE, T. ; RIGOLL, G. : *Advanced Man-Machine Interaction*. Springer Verlag Berlin Heidelberg New York, ISBN: 3-540-30618-8, 2006, S. 141–190
- [218] SIBERT, L. ; JACOB, R. : Evaluation of eye gaze interaction. In: *Proceedings of Human Factors in Computing Systems Conference*, 2000, S. 281–288
- [219] SMILEY, A. : Mental Workload and information management / IEEE. 1989. – Forschungsbericht
- [220] SPAKOV, O. ; MONIOTAS, D. : Extending the limits for gaze pointing through the use of speech. In: *Information Technology and Control*. 34 (2005), S. 225–230
- [221] SR RESEARCH LTD., 5516 MAIN ST., OSGOODE ON CANADA K0A 2W0: <http://www.eyelinkinfo.com/>, letzter Zugriff: 15.11.2010,
- [222] STAMBOLIEV, S. : *Konzeption eines blickbedienten HMI für automobile Anwendungen*. Master Thesis, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, Juli 2009
- [223] STATISTISCHES BUNDESAMT DEUTSCHLAND: *Pressemitteilung Nr. 071 vom 25.02.2008*. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2008/02/PD08__071__46241,templateId=renderPrint.psml,
- [224] STURM, M. : *Microcontrollertechnik - Am Beispiel der MSP430-Familie*. Hanser Verlag, 2006
- [225] STUTTS, J. C. ; REINFURT, D. W. ; STAPLIN, L. ; RODGMAN, E. A.: The Role of Driver Disraction in Traffic Crashes. / AAA Foundation for Traffic Safety. 2001. – Forschungsbericht
- [226] SUGENO, M. : *Industrial Applications of Fuzzy Control*. New York, NY, USA : Elsevier Science Inc., 1985. – ISBN 0444878297
- [227] TAKAGI, T. ; SUGENO, M. : Fuzzy Identification of Systems and Its Applications to Modeling and Control. In: *IEEE Transactions on Systems, Man, and Cybernetics* 15 (1985), February, Nr. 1, S. 116–132
- [228] TALMI, K. ; LIU, J. : Eye and gaze tracking for visually controlled interactive stereoscopic displays. In: *Signal Processing: Image Communication* 14 (1999), S. 799–810

- [229] THE MATHWORKS, INC.: *Fuzzy Logic Toolbox 2*. Release 2008a. 3 Apple Hill Drive Natick, MA 01760-2098: The MathWorks, Inc., 3 2008
- [230] TÖNNIS, M. : *Towards Automotive Augmented Reality*. München, Technische Universität München, Dissertation, 2008
- [231] TÖNNIS, M. ; BROY, V. ; KLINKER, G. : *A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers*. Vorlesung „3D-Nutzerschnittstellen“, Chair for Computer Aided Medical Procedures & Augmented Reality, Technische Universität München, 24. Nov. 2006
- [232] TOBII TECHNOLOGY AB: *www.tobii.com*. , letzter Zugriff: 13.02.2010,
- [233] TORNROS, J. ; BOLLING, A. : Mobile phone use - Effects of handheld and hands-free phones on driving performance. In: *Accident Analysis and Prevention Vol. 37* (2005), S. 902–909
- [234] TOROK, N. ; GUILLEMIN, V. ; BARNOOTHY, J. M.: Photoelectric nystagmography. In: *Annals of Otology, Rhinology and Laryngology* 60 (1951), S. 917–926
- [235] TRIVEDI, M. M. ; CHENG, S. Y.: Holistic Sensing and Active Displays for Intelligent Driver Support Systems. In: *Computer* 40 (2007), May, Nr. 5, S. 60–68
- [236] TUISKU, O. ; MAJARANTA, P. ; ISOKOSKI, P. ; RÄIHÄ, K.-J. : Now Dasher! Dash away!: longitudinal study of fast text entry by Eye Gaze. In: *ETRA '08: Proceedings of the 2008 symposium on Eye tracking research & applications*. New York, NY, USA : ACM, 2008. – ISBN 978-1-59593-982-1, S. 19–26
- [237] URBINA, M. ; HUCKAUF, A. : Dwell time free eye typing approaches. In: *Conference on Communication by Gaze Interaction (GOGAIN) Leicester, UK, 2007*, S. 65–70
- [238] VANDEVENNE, L. : *Computer Graphics Tutorial – Flood Fill*. <http://www.student.kuleuven.be/m0216922/CG/floodfill.html>, letzter Zugriff: 18.11.2010,
- [239] VERTEGAAL, R. : Designing attentive interfaces. In: *Proceedings of the ACM Symposium on Eye Tracking Research & Applications - ETRA 2002, 2002*
- [240] VERTEGAAL, R. : Attentive User Interfaces. In: *Communications of the ACM* 46 (2003), S. 30–33
- [241] VERTEGAAL, R. : A Fitts' Law Comparison of Eye Tracking and Manual Input in the Selection of Visual Targets. In: *Proceeding of ICMI '08 Proceedings of the 10th international conference on Multimodal interfaces* , 2008
- [242] VERWEY, W. : How can we prevent overload of the driver? In: PARKES, A. (Hrsg.) ; FRANZEN, S. (Hrsg.): *Driving future vehicles*. Taylor & Francis, 1993, Kapitel 22
- [243] VERWEY, W. ; ALM, H. ; GROEGER, J. ; JANSSEN, W. ; KUIKEN, M. ; SCHRAAGEN, J. ; SCHUMANN, J. ; WINSUM, W. van ; WONTORRA, H. : GIDS functions. In: MICHON, J. (Hrsg.): *Generic Intelligent Driver Support, A Comprehensive Report on GIDS*. Taylor & Francis, 1993, Kapitel 7
- [244] VERWEY, W. ; VELTMAN, J. : Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. In: *Journal of experimental psychology: Applied*, 2 (1996), S. 270–285
- [245] VILLGRATTNER, T. : *Design and Control of Compact High Dynamic Camera Orientation Systems.*, Technische Universität München, Diss., 2010
- [246] VIOLA, P. ; JONES, M. : Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, 2001*, S. 511–518

- [247] VÖLK, F. ; KERBER, S. ; FASTL, H. ; REIFINGER, S. : Design und Realisierung von virtueller Akustik für ein Augmented-Reality-Labor. In: *Tagungsband Fortschritte der Akustik - DAGA 2007, Stuttgart* DEGA, 2007, S. 559–560
- [248] WAARD, D. de: *The measurement of drivers' Mental Workload.*, The Traffic Research Centre VSC, University of Groningen, Diss., 1996
- [249] WANG, J.-G. ; SUNG, E. : Gaze determination via images of irises. In: *Image and Vision Computing* 19 (2001), S. 891–911
- [250] WICKENS, C. D. ; HOLLANDS, J. G.: *Engineering Psychology and Human Performance.* Charles E. Merrill Publishing Co., Columbus, OH, 1984
- [251] WIESER, E. : *EyeSeeBot: Entwurf eines remote Eyetrackers für Robotik-Anwendungen.* Bachelorarbeit, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Betreuer: T. Poitschke, April 2010
- [252] WIKIPEDIA: *Auge.* <http://de.wikipedia.org/wiki/Auge>, Letzter Zugriff: 6.12.2008,
- [253] WITTLING, W. : *Einführung in die Psychologie der Wahrnehmung.* Hoffmann und Campe, Hamburg, 1976
- [254] YANG, J. ; ZHENG, N. : An Expert Fuzzy Controller for lateral Vehicle Control / Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, China. 2007. – Forschungsbericht
- [255] YOUNG, L. R. ; SHEEN, D. : Survey of eye movement recording methods. In: *Behaviour Research Methods and Instrumentations* 7 (1975), S. 397–429
- [256] ZADEH, L. A.: *Fuzzy Sets, Information and Control / University of California.* 1965. – Forschungsbericht
- [257] ZHAI, S. ; MORIMOTO, C. ; IHDE, S. : Manual and gaze input cascaded (MAGIC) pointing. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: the CHI is the limit.*, 1999, S. 246 – 253
- [258] ZHU, Z. ; JI, Q. : Novel Eye Gaze Tracking Techniques Under Natural Head Movement. In: *IEEE Transactions on Biomedical Engineering.* 54 (2007), Nr. 12, S. 2246 – 2260
- [259] ZIMMER, A. ; DAMEN-ZIMMER, K. : Situationsbezogene Sicherheitskenngrößen im Straßenverkehr. In: *Berichte der Bundesanstalt für Straßenwesen.* M78 (1997)
- [260] ZYLSTRA, B. ; TSIMHONI, O. ; GREEN, P. ; MAYER, K. : Driving Performance for Dialing, Radio Tuning, and Destination Entry while Driving Straight Roads / Ann Arbor, MI: The University of Michigan Transportation Research Institute. 2003 (Technical Report UMTRI-2003-35). – Forschungsbericht
- [261] ZYTRONIC DISPLAYS LTD: *Projected Capacitive Touch Sensor Technology.* www.zytronic.co.uk, letzter Zugriff: 13.10.2010,

A

Anhang

A.1 Ausführliche Versuchsergebnisse zur Intentionserkennung

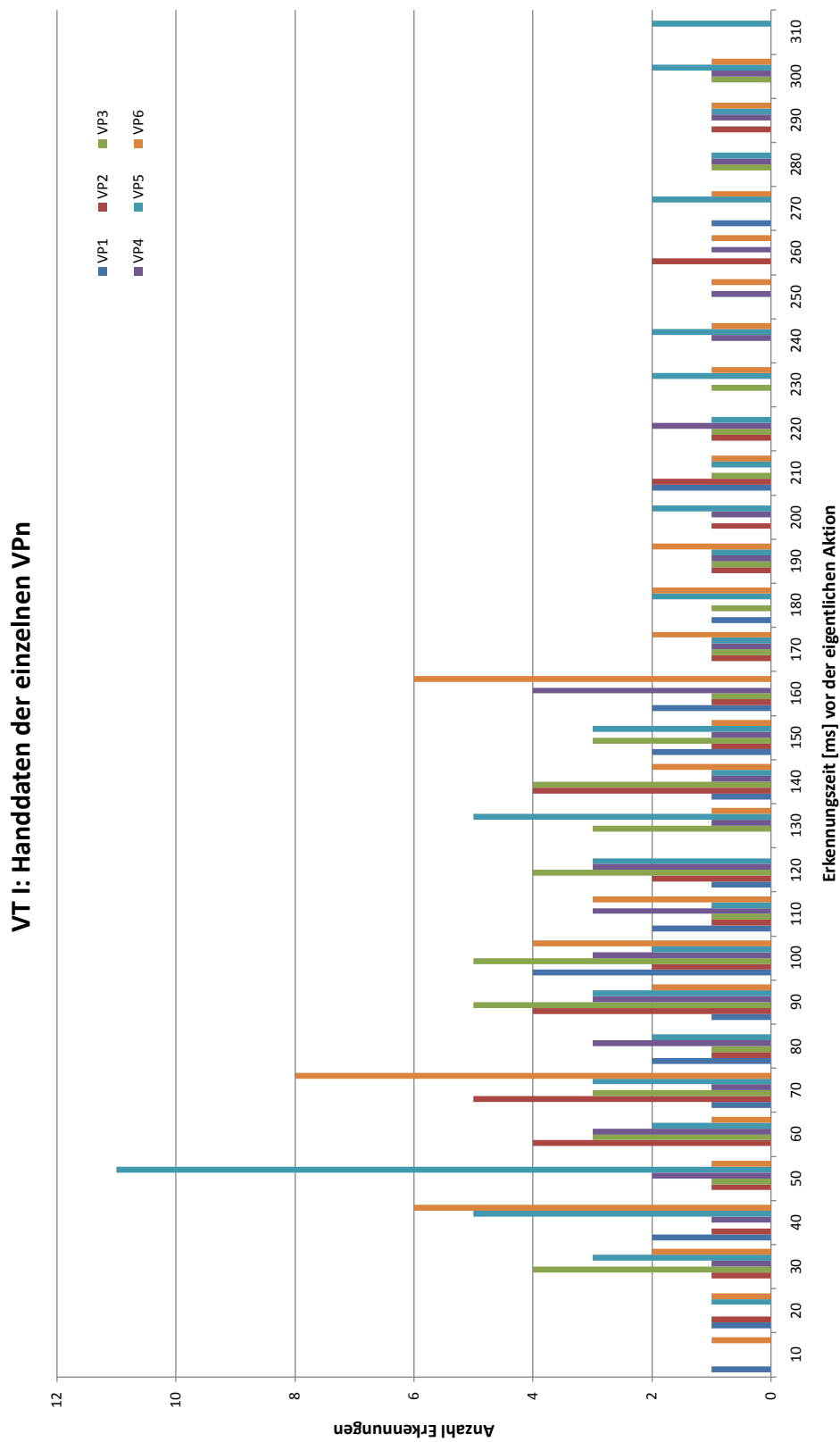


Abbildung A.1: Versuchsteil I, Handdaten aller Versuchspersonen. Hier dargestellt der Erkennungsbereich von 10 ms bis 310 ms vor Durchführung der eigentlichen Aktion. Im restlichen Zeitbereich sind keine signifikanten Erkennungsanzahlen vorhanden.

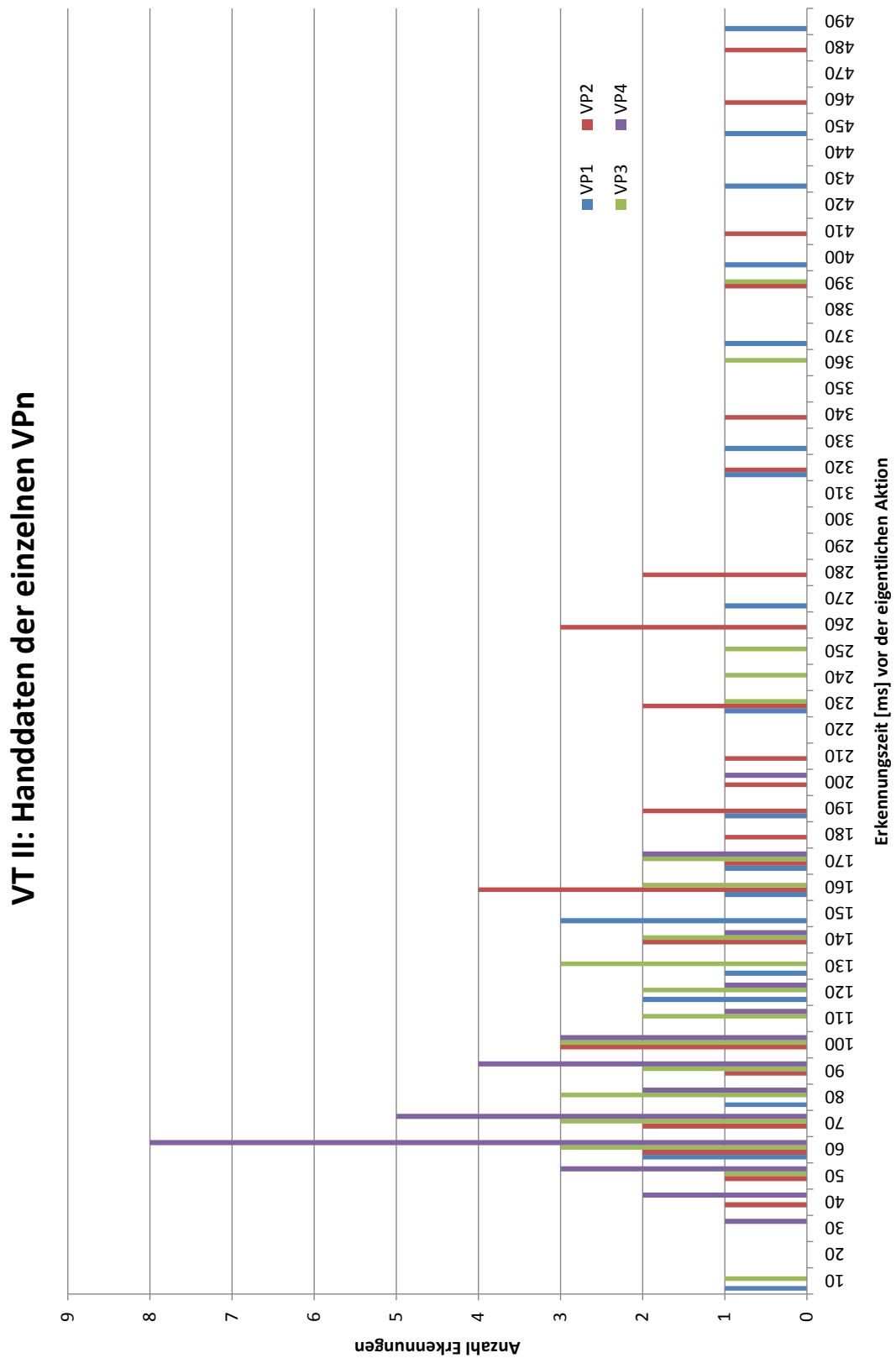


Abbildung A.2: Versuchsteil II, Handdaten aller Versuchspersonen. Hier dargestellt der Erkennungsbereich von 10 ms bis 490 ms vor Durchführung der eigentlichen Aktion. Im restlichen Zeitbereich sind keine signifikanten Erkennungsanzahlen vorhanden.