

Lehrstuhl für Mensch-Maschine-Kommunikation  
Technische Universität München

## **Multimodales, kontextadaptives Informationsmanagement im Automobil**

Markus Ablaßmeier

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. Norbert Hanik  
Prüfer der Dissertation: 1. Univ.-Prof. Dr.-Ing. habil. Gerhard Rigoll  
2. Univ.-Prof. Dr. rer.nat. Heiner Bubb, i.R.

Die Dissertation wurde am 06.10.2008 bei der Technischen Universität München eingereicht  
und durch die Fakultät für Elektrotechnik und Informationstechnik am 24.02.2009 angenommen.



# Inhaltsverzeichnis

<b>Zusammenfassung</b>	<b>xiii</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Problemstellung . . . . .	2
1.3 Zielsetzung und Aufgabenstellung . . . . .	2
<b>2 Einführung und Grundlagen</b>	<b>5</b>
2.1 Mensch-Maschine Interaktion . . . . .	5
2.1.1 Passungsmodell der Mensch-Maschine Interaktion . . . . .	6
2.1.2 Drei-Ebenen-Modell nach Rasmussen . . . . .	7
2.2 Fahrzeugdomäne . . . . .	7
2.2.1 Drei-Ebenen-Modell der Fahrzeugführung . . . . .	8
2.2.2 Fahraufgaben im Automobil nach Geiser . . . . .	9
2.2.3 Fahrerinformationssysteme . . . . .	10
2.2.4 Fahrerassistenzsysteme . . . . .	10
2.2.5 Fahrerzustandsüberwachung . . . . .	11
2.2.6 Entwicklungstrends bei Fahrerinformationssystemen . . . . .	12
2.2.7 Zukünftige Fahrerbedarfe . . . . .	15
2.2.8 Einflussgrößen und Kontext im Fahrzeug . . . . .	16
2.3 Kognitionswissenschaftliche Grundlagen . . . . .	18
2.3.1 Aufmerksamkeit . . . . .	19
2.3.2 Verhaltensmodelle . . . . .	19
2.3.3 Situation Awareness . . . . .	21
2.3.4 Workload . . . . .	21
2.3.5 Mentale Modelle . . . . .	21
2.4 Usability – Normen, Gestaltungshinweise und Richtlinien . . . . .	21
2.4.1 Definition Usability . . . . .	22
2.4.2 Normen . . . . .	22
2.4.3 Guidelines . . . . .	23
2.4.4 Styleguides . . . . .	24
2.4.5 Gestaltgesetze . . . . .	24
2.4.6 Bedienfreude - Joy of Use . . . . .	24
2.4.7 Empfehlungen und Standards für FIS . . . . .	25
2.5 Bewertungsmethoden . . . . .	25
2.5.1 Fahrsimulator . . . . .	25

2.5.2	Peripheral Detection Task . . . . .	27
2.5.3	Semantisches Differential . . . . .	28
2.5.4	Blickverhaltensmessung . . . . .	29
2.6	Zusammenfassung . . . . .	29
<b>3</b>	<b>Multimodales, kontextadaptives Dialogmanagement</b>	<b>31</b>
3.1	Dialogsysteme, Dialogführung . . . . .	31
3.1.1	Dialogformen . . . . .	31
3.1.2	Menüdialoge . . . . .	33
3.1.3	Sprachdialoge . . . . .	34
3.2	Multimodale Interaktion . . . . .	38
3.2.1	Vorteile multimodaler Kommunikation . . . . .	39
3.2.2	Ebenen multimodaler Fusion . . . . .	40
3.2.3	Wichtige Integrations- bzw. Fusionsmethoden . . . . .	41
3.2.4	Multimodalität im Automobil . . . . .	42
3.3	Kontextadaptive Dialogsysteme . . . . .	43
3.3.1	Einordnung adaptiver Benutzerschnittstellen . . . . .	43
3.3.2	Aktionsvorhersage und Antizipation . . . . .	44
3.3.3	Adaptivität im Automobilkontext . . . . .	46
3.4	Dialogmodellierung . . . . .	47
3.4.1	Zustandsautomaten . . . . .	47
3.4.2	Graphische Repräsentation . . . . .	48
3.4.3	Formale Beschreibung . . . . .	50
3.5	Multimodales, kontextadaptives Dialogframework . . . . .	50
3.5.1	Etablierte Architekturmodelle . . . . .	50
3.5.2	Multimodale, kontextadaptive Dialogarchitektur . . . . .	51
3.5.3	Datenmanagement-System . . . . .	54
<b>4</b>	<b>Multimodale, datenübergreifende Fahrzeug-Suchmaschine</b>	<b>57</b>
4.1	Suche in verschiedenen Applikationen . . . . .	57
4.1.1	Datei- und Datensuche . . . . .	57
4.1.2	Internetsuche . . . . .	58
4.1.3	Suche auf mobilen Endgeräten . . . . .	59
4.1.4	Suche durch Exploration . . . . .	59
4.2	Anforderungen der Fahrzeug-Suchmaschine . . . . .	61
4.2.1	Datenbasen . . . . .	61
4.2.2	Suchanfrage . . . . .	61
4.2.3	Ergebnisaufbereitung . . . . .	62
4.2.4	Multimodale Bedienstrategien . . . . .	63
4.3	Mathematische Methoden . . . . .	64
4.3.1	Exakte Suche . . . . .	64
4.3.2	Invertierte Indizes . . . . .	67
4.3.3	Unschärfe Suche . . . . .	69
4.3.4	Einsatz der Suchverfahren im Fahrzeug . . . . .	73
4.3.5	Rankingverfahren . . . . .	74
4.4	Alphanumerische Eingabestrategien . . . . .	76
4.4.1	Taktile Eingabegeräte mit haptischem Feedback . . . . .	76

---

4.4.2	Touch-Eingabegeräte . . . . .	77
4.4.3	Sprache . . . . .	78
4.5	Anzeige-/Bedienkonzepte für die Fahrzeug-Suchmaschine . . . . .	78
4.5.1	Suchinterface CARhoo! . . . . .	78
4.5.2	Suchinterface IMSENIO . . . . .	79
4.5.3	Suchinterface CARpedia . . . . .	80
4.6	Evaluierung der Suchinterfaces . . . . .	83
4.6.1	Untersuchung Suchinterface CARhoo! . . . . .	83
4.6.2	Untersuchung Suchinterface CARpedia . . . . .	84
4.6.3	Untersuchung alphanumerischer Eingabeverfahren . . . . .	87
<b>5</b>	<b>Objektorientiertes, vernetztes Datenbrowsing</b>	<b>91</b>
5.1	Verwendete Methoden . . . . .	91
5.1.1	Objektorientierte Struktur . . . . .	91
5.1.2	Verknüpfung durch Tagging . . . . .	93
5.1.3	Filtern von Objektlisten . . . . .	94
5.2	Graphisches Interface . . . . .	94
5.2.1	Statusmenü . . . . .	94
5.2.2	Objektorientierte Bedienung über Domänen . . . . .	96
5.3	Dialogformen, Verknüpfungsmethoden und Filtermechanismen . . . . .	98
5.3.1	Objektorientierter Funktionszugriff in den Domänen . . . . .	99
5.3.2	Filtermechanismen . . . . .	99
5.4	Multimodale Bedienung . . . . .	102
5.5	Zusammenfassung . . . . .	104
<b>6</b>	<b>Kontextsensitive Informationsagenten</b>	<b>105</b>
6.1	Aufbau der Informationsagenten . . . . .	105
6.1.1	Grundlagen von Agentensystemen . . . . .	107
6.1.2	Systemarchitektur der Informationsagenten . . . . .	109
6.2	Mathematische Grundlagen und Methoden . . . . .	110
6.2.1	Einführung in Bayes'sche Netze . . . . .	111
6.2.2	Einführung in Neuronale Netze . . . . .	115
6.3	Konzeption der Informationsagenten . . . . .	119
6.3.1	Tankagent . . . . .	119
6.3.2	Kontaktagent . . . . .	124
6.3.3	Terminagent . . . . .	126
6.3.4	Restaurant-Recommender . . . . .	127
6.3.5	Weitere Informationsagenten . . . . .	130
6.4	Bewertung und Diskussion . . . . .	131
<b>7</b>	<b>Informationsvisualisierung und multimodale Bedienstrategien</b>	<b>133</b>
7.1	Darstellungsmethoden . . . . .	133
7.1.1	Metaphern . . . . .	134
7.1.2	Direkte Manipulation . . . . .	134
7.1.3	Zoomable User Interface . . . . .	134
7.1.4	Animationen . . . . .	136
7.2	Zoomable User Interfaces für das Fahrzeug . . . . .	136

7.2.1	Zoomable User Interfaces mittels Annäherungstouchscreen . . . . .	137
7.2.2	Zoomable User Interface mittels Dreh-/Drücksteller . . . . .	140
7.3	Verwendung von 3D-Metaphern im Fahrzeug . . . . .	141
7.3.1	Skalierbares 3D-Filmrollenkonzept und Verknüpfungsnetzwerk . . . . .	142
7.3.2	3D-Zylindermetapher als Menütiefe . . . . .	143
7.3.3	3D-Rollkartei und Karteiflip . . . . .	143
7.4	Bewertung . . . . .	146
7.4.1	Versuchsdurchführung . . . . .	146
7.4.2	Subjektive Ergebnisse . . . . .	146
7.5	Zusammenfassung . . . . .	148
<b>8</b>	<b>Großflächiges Head-Up Display</b>	<b>149</b>
8.1	Grundlagen zum Head-Up Display . . . . .	149
8.1.1	Funktionsweise und technische Auslegung . . . . .	150
8.1.2	Vor- und Nachteile des HUD . . . . .	154
8.2	Innovatives, multimodales Anzeige-/Bedienkonzept im HUD . . . . .	155
8.2.1	Gesamtkonzept HUD_MAX für ein großes schwebendes HUD . . . . .	155
8.2.2	HUD Menüinteraktion MATRIX_MMI . . . . .	157
8.2.3	Multimodale Bedienung . . . . .	158
8.3	Integriertes HUD-Kombi-Konzept . . . . .	162
8.3.1	Prioritätenorientiertes Informationssplitting . . . . .	162
8.3.2	Fahrerzentrierter Kombi-Fokuswechsel . . . . .	163
8.4	Kontaktanalogen Head-Up Display . . . . .	163
8.4.1	Augmented Reality . . . . .	163
8.4.2	Kontaktanaloge Konzeptentwürfe . . . . .	165
8.4.3	Sitzkistenaufbau eines quasi-kontaktanalogen HUDs . . . . .	168
8.4.4	Gaze- und Eyetracking . . . . .	171
8.4.5	Anforderungen quasi-kontaktanaloger Anzeigen im Realfahrzeug . . . . .	176
8.5	HUD Untersuchungen . . . . .	177
8.5.1	Vergleich Serien-HUD mit HDD (Feldstudie I) . . . . .	177
8.5.2	Untersuchung von HUD-Inhalten (Feldstudie II) . . . . .	179
8.5.3	Vergleich Menüinteraktion HUD vs. CID (Simulatorstudie Ia) . . . . .	180
8.5.4	Untersuchung der Anzeigemenge im HUD (Simulatorstudie Ib) . . . . .	184
8.5.5	Vergleich Matrix- vs. Listendarstellung (Simulatorstudie Ic) . . . . .	186
8.5.6	Untersuchung der multimodalen Bedienung im HUD (Simulatorstudie II) . . . . .	189
<b>9</b>	<b>Resumé und Ausblick</b>	<b>191</b>
9.1	Methoden zur Informationsverarbeitung . . . . .	191
9.2	Dialogstrategien . . . . .	192
9.3	Multimodale Anzeige und Bedienung . . . . .	194
9.4	Ausblick . . . . .	196
	<b>Abkürzungsverzeichnis</b>	<b>197</b>
	<b>Index</b>	<b>201</b>
	<b>Literaturverzeichnis</b>	<b>205</b>

# Abbildungsverzeichnis

1.1	Informationsfluss im Fahrzeug . . . . .	3
1.2	Informationsmanagement auf drei Ebenen . . . . .	4
2.1	Passungsmodell der Mensch-Maschine Interaktion . . . . .	6
2.2	Drei-Ebenen-Modell nach Rasmussen [Ras83] . . . . .	7
2.3	Überblick Drei-Ebenen-Modell der Fahrzeugführung [Bub08a] . . . . .	8
2.4	Fahraufgaben nach Geiser [Gei90] . . . . .	9
2.5	Fahrerinformations- und Fahrerassistenzsysteme . . . . .	11
2.6	Das BMW iDrive der 5er Serie [BMW08b] . . . . .	13
2.7	Das Audi MMI des A8 [Aud08] . . . . .	13
2.8	Das Mercedes COMAND der S-Klasse [Dai08] . . . . .	14
2.9	Das Lexus LS Cockpit [Lex08] . . . . .	15
2.10	Überblick Fahrtchronologie . . . . .	17
2.11	Fahrtszenarienmatrix nach Bubb [BRT <sup>+</sup> 04] . . . . .	18
2.12	Beispielhafter Verlauf eines Szenarios (Geschäftsfahrt) . . . . .	18
2.13	Aufteilung in Ressourcenpools der Multiple Resource Theory [Wic84] . . . . .	20
2.14	POC-Kurve für zwei simultan bearbeitete Aufgaben A und B . . . . .	20
2.15	Grundsätze der Dialoggestaltung, Zusammenfassung ISO9241 Teil 10 [Nie02] . . . . .	23
2.16	Überblick Laborräume des Lehrstuhls für Mensch-Maschine-Kommunikation . . . . .	26
2.17	Schematischer Laboraufbau des Fahrersimulators . . . . .	26
2.18	Fahrersimulator-Labor am Lehrstuhl für Mensch-Maschine-Kommunikation . . . . .	27
2.19	Sichtsimulationen für den Fahrersimulator . . . . .	27
2.20	Beispiel für ein Adjektiv-Paar aus dem Semantischen Differential . . . . .	28
2.21	Das Janus Blickverhaltensmessungssystem [Sch03b] . . . . .	29
3.1	Menüdialoge . . . . .	33
3.2	Graphische Repräsentation von Frame-Based Dialogen . . . . .	36
3.3	Verarbeitungskapazitäten der Sinne [SR08] . . . . .	38
3.4	Klassifikation multimodaler Systeme nach [NC93] . . . . .	39
3.5	Klassifizierung von Benutzerschnittstellen hinsichtlich der Adaptivität [Loe89] . . . . .	43
3.6	Elemente eines Zustandsautomaten . . . . .	47
3.7	Erweiterung der Darstellungsform durch Statecharts . . . . .	49
3.8	Überblick Seeheim-Modell . . . . .	51
3.9	Architektur eines multimodalen Dialogmanagements . . . . .	52
3.10	Frame-Based Dialogmanagement mit Kontextbedingungen . . . . .	53
3.11	Aufbaudiagramm des Datenmanagement-Systems . . . . .	55

4.1	Datei- und Datensuche . . . . .	58
4.2	Internetaktionshaus ebay . . . . .	59
4.3	Suche Google mobile . . . . .	60
4.4	FaThumb Interface [KRR <sup>+</sup> 06] . . . . .	60
4.5	Beispiel für den naiven Suchalgorithmus . . . . .	65
4.6	Beispiel für die fehlgeschlagene Suche mit einem abgewandelten naiven Suchalgorithmus . . . . .	65
4.7	Die Idee des Knuth-Morris-Pratt Algorithmus . . . . .	66
4.8	Invertierter Index auf Ebene der Einträge . . . . .	68
4.9	Invertierter Index auf Ebene der Wörter . . . . .	68
4.10	Berechnung der Levenshtein Distanz mittels dynamischer Programmierung . . . . .	71
4.11	Berechnung der Levenshtein Distanz mittels dynamischer Programmierung für eine Suche im Text . . . . .	71
4.12	Berechnungsvorschrift für den Soundex Code . . . . .	72
4.13	Taktile alphanumerische Eingabestrategien . . . . .	76
4.14	Texteingabe über Speller . . . . .	76
4.15	CARhoo! . . . . .	78
4.16	IMSENIO . . . . .	80
4.17	CARpedia . . . . .	81
4.18	CARsearch . . . . .	82
4.19	Verwendung von Suche und Exploration über alle Aufgaben . . . . .	85
4.20	Ergebnisse des Attradiff-Fragebogens von CARpedia . . . . .	85
4.21	Verwendung der Features der Suchmaschine . . . . .	86
4.22	Mittlere Bediendauern für die Aufgaben . . . . .	86
4.23	Mittlere Spurabweichung für die Aufgaben . . . . .	87
4.24	Mittlere Spurabweichung der einzelnen Bedienelemente . . . . .	88
4.25	Mittlere Aufgabendauer bei Verwendung der einzelnen Bedienelemente . . . . .	88
4.26	Befragung zur Modalitätenpräferenz . . . . .	89
5.1	Überblick Statusmenü und Domänen . . . . .	95
5.2	Statusmenü Sprungfunktionen . . . . .	100
5.3	explizite Filterung . . . . .	100
5.4	Implizite Filterung . . . . .	102
5.5	Überblick über die Verknüpfungen der Domänen und des Statusmenüs . . . . .	103
6.1	Bündelung der Dialogschritte durch die Informationsagenten . . . . .	106
6.2	Überblick Informationsagenten . . . . .	106
6.3	Beispiel eines Tankagentendialogs . . . . .	107
6.4	Agentensystem als Erweiterung eines Dialogsystems . . . . .	109
6.5	Einfachstes Bayes'sches Netz . . . . .	112
6.6	Bayes'sches Netz mit drei Variablen (Knoten) und deren Abhängigkeiten (Kanten) . . . . .	112
6.7	Darstellung eines Neurons . . . . .	116
6.8	Binäre Schwellenwertfunktion . . . . .	117
6.9	Semilineare Ausgabefunktion . . . . .	117
6.10	n-schichtiges feedforward-Netz mit verdeckten Schichten . . . . .	118
6.11	Modellierung des Tankagenten: Dialogmanager und Informationsmanager . . . . .	120
6.12	Exponentialfunktion zur Einteilung des Tanks in festgelegte Bereiche . . . . .	122

6.13	Bayes'sches Netz des Kontaktagenten . . . . .	126
6.14	Abgewandelte Netz-Struktur für einen dreidimensionalen Eingangsvektor . . . .	129
7.1	Zoomable User Interface . . . . .	135
7.2	Aufnahmen mit einem Fisheye-Objektiv . . . . .	136
7.3	Fisheye Techniken . . . . .	137
7.4	Aufteilung des Sensorraums . . . . .	138
7.5	PreTouch-System . . . . .	139
7.6	MatrixKey . . . . .	140
7.7	Flussdiagramm für multimodale Schiebermanipulation . . . . .	140
7.8	Flussdiagramm für einen <i>Point-and-Talk</i> -Vorgang . . . . .	140
7.9	Zoomable User Interface mittels Dreh-Drücksteller . . . . .	141
7.10	Zoomable User Interface mittels Dreh-Drücksteller (Menüpanning) . . . . .	142
7.11	Gekrümmte Listendarstellung mit Zooming . . . . .	142
7.12	Taktile alphanumerische Eingabeverfahren . . . . .	142
7.13	Überblick skalierbares 3D-Filmrollenkonzept . . . . .	144
7.14	3D-Zylindermetapher als Menütiefe . . . . .	144
7.15	3D-Rollkartei und Karteiflip . . . . .	145
7.16	Mittelwerte der Attradiff-Auswertung . . . . .	148
8.1	Das HUD der BMW 5er Reihe [BMW08b] . . . . .	150
8.2	Die Funktionsweise des HUDs . . . . .	151
8.3	Prinzip des HUDs im BMW 5er . . . . .	151
8.4	Korrektur von Doppelbildern bei der Windschutzscheibe . . . . .	152
8.5	Querschnitt durch das menschliche Auge . . . . .	152
8.6	Überblick HUD_MAX-Konzept . . . . .	156
8.7	Beispieldarstellung des Lösungskonzeptes HUD_MAX . . . . .	157
8.8	Matrix Menü . . . . .	158
8.9	Überblick Matrix- und Listenstruktur . . . . .	159
8.10	Bedienelement am Lenkrad: 5-Wege-Navipad . . . . .	159
8.11	Verarbeitungsstatus . . . . .	159
8.12	Sprachfeedback in der Statusleiste . . . . .	160
8.13	Tankagent . . . . .	161
8.14	Restaurantagent . . . . .	161
8.15	Schematischer Überblick des integrierten HUD-Kombi-Konzeptes . . . . .	162
8.16	Kontaktanaloge Darstellung eines Navigationspfeils . . . . .	164
8.17	Zusammenhang Realität-Virtualität [Azu97] . . . . .	164
8.18	Kontaktanaloge Darstellungsarten . . . . .	166
8.19	Hervorhebung von POIs im HUD . . . . .	167
8.20	Überblick Kompasskonzept . . . . .	168
8.21	Schematischer Aufbau der verwendeten Sitzkiste . . . . .	169
8.22	Integration des remote Gazetrackers im Simulator . . . . .	171
8.23	Systemaufbau Hennessey [HNL06] . . . . .	174
8.24	Geometrie und Hilfskoordinatensystem [HNL06] . . . . .	175
8.25	Detektion des Pupillenmittelpunktes [HNL06] . . . . .	175
8.26	Herausforderungen HUD Entwicklung . . . . .	176
8.27	Häufigkeitsverteilung der Blickabwendungsdauern aller Probanden . . . . .	178

8.28	Mittlere Blickabwendungsdauern im HUD in Abhängigkeit des Alters . . . . .	178
8.29	Mittlere Bearbeitungsdauer in Abhängigkeit vom Aufgabentyp . . . . .	181
8.30	Mittlere Spurabweichung in Abhängigkeit des Aufgabentyps . . . . .	181
8.31	Sinnvoller Anzeigeort für Menübedienung . . . . .	182
8.32	Einschätzung der Menübedienung im Anzeigort als gefährlich . . . . .	183
8.33	Die zwei verschiedenen Anzeigenmengen im HUD . . . . .	184
8.34	Mittlere Anzahl nicht erkannter PDT-Punkte für die jeweiligen Positionen . . .	185
8.35	Mittlere Reaktionszeit auf die PDT-Punkte in Abhängigkeit der Anzeigemenge	185
8.36	Mittlere Spurabweichung in Abhängigkeit der Anzeigemenge und Alter . . . . .	186
8.37	Mittlere Auswahldauer aus Matrix und Liste . . . . .	187
8.38	Mittlere Anzahl der benötigten Schritte für Matrix und Liste . . . . .	188
8.39	Bevorzugte Darstellungsweise . . . . .	188
9.1	Überblick Dialogstrategien . . . . .	193
9.2	CAVE des Lehrstuhls für Mensch-Maschine-Kommunikation . . . . .	196

# Tabellenverzeichnis

3.1	Übergangstabelle . . . . .	48
3.2	XML-Dialogformulierung . . . . .	54
4.1	Datenbasen . . . . .	61
4.2	next-Tabelle für die Suchfolge “mmk” . . . . .	66
4.3	Codierung von Buchstaben nach der “Kölner Phonetik” . . . . .	73
4.4	Eingabefehler . . . . .	89
4.5	Bewertung des semantischen Differentials (Mittelwerte) . . . . .	89
5.1	Vordefinierte Filter . . . . .	101
6.1	Bedingte Wahrscheinlichkeiten $P(G W)$ für $G$ unter Voraussetzung $W$ und $P(S W)$ für $S$ unter Voraussetzung $W$ . . . . .	113
6.2	Verbundwahrscheinlichkeiten $P(G, W)$ und $P(S, W)$ . . . . .	113
6.3	Auftrittswahrscheinlichkeiten $P(S)$ und $P(G)$ durch Marginalisierung . . . . .	113
6.4	Rückschlußwahrscheinlichkeiten $P(W S)$ für $W$ unter Voraussetzung $S$ und $P(W G)$ für $W$ unter Voraussetzung $G$ . . . . .	114
6.5	Bedingte Wahrscheinlichkeiten $P(G W)$ für $G$ unter Voraussetzung $W$ und $P(S W)$ für $S$ unter Voraussetzung $W$ . . . . .	114
6.6	Häufigkeit der Wahl eines Rangs der Kriterien (aus Voruntersuchung) . . . . .	123
6.7	Häufigkeit der jeweils gewählten <i>Marke</i> (aus Voruntersuchung) . . . . .	123
6.8	Wahrscheinlichkeitstabelle für den Knoten <i>Entscheidung nach Tankinhalt</i> . . . . .	123
6.9	Standardabweichung und Normierung für die Gewichtung der Kriterien . . . . .	123
6.10	CPT für den Knoten <i>Wahrscheinlichste Tankstelle</i> . . . . .	124
6.11	Aktualisierte Wahrscheinlichkeitstabelle für den Knoten <i>EntschNachTankinhalt</i> . . . . .	125
7.1	Bewertung der ZUI-Effekte (semantisches Differential) . . . . .	146
7.2	Bewertung der Touch-Bedienung . . . . .	147
7.3	Bewertung der multimodalen Bedienung . . . . .	147
7.4	Bewertung des Menüs (semantisches Differential) . . . . .	147
8.1	Sichtbereiche des menschlichen Auges . . . . .	153



# Zusammenfassung

Die vorliegende Arbeit liefert neuartige Lösungsansätze zum Umgang mit großen Datenmengen im Fahrzeug. Hierzu werden effiziente Dialogstrategien für den Fahrer identifiziert, geeignete multimodale Interfaces entwickelt und die erforderlichen Methoden zur Informationsaufbereitung aufgezeigt. Um ein ganzheitliches Informationsmanagement im Automobil zu schaffen, dient als Grundstock für die konzipierten Interfaces ein multimodales, kontextadaptives Dialogframework. Damit ein schneller Informationszugriff für den Fahrer möglich wird, werden in dieser Arbeit drei innovative Ansätze vorgestellt: eine datenübergreifende Fahrzeug-Suchmaschine, ein objektorientiertes, vernetztes Datenbrowsing und der Einsatz von „intelligenten“ Informationsagenten. Die datenübergreifende Fahrzeug-Suchmaschine unterstützt den Fahrer bei der effektiven Informationsauffindung und bietet einen Schnellzugriff auf sämtliche Daten und Funktionen des Menüs. Hierzu werden mehrere multimodale Suchinterfaces konzipiert und in Nutzerstudien untersucht. Beim vernetzten Datenbrowsing-Ansatz erfolgt die Integration verschiedenster Informationsquellen (z. B. mobiler Endgeräte) über eine datenorientierte Grundstruktur. Mittels impliziter und expliziter Filterkonzepte lässt sich die relevante Datenmenge schnell und effektiv eingrenzen. Die Vernetzung der Daten ermöglicht es, die Bedienungsaufgaben des Fahrers durch schnelle, kontextsensitive Sprünge nutzer- und aufgabengerecht abzubilden. Dabei können die nötigen Funktionen durch den objektorientierten Ansatz kontextabhängig verwendet werden. Anhand neuartiger Visualisierungsstrategien, wie 3D- und Zoomable User Interfaces, wird die intuitive Darstellung umfangreicher Informationen im Fahrzeug auf Basis des Datenbrowsers aufgezeigt. Die „intelligenten“ Informationsagenten vereinfachen gerade ständig wiederkehrende und komplexe Handlungsabläufe für den Fahrer. Hier erfährt der Nutzer eine aufgabenorientierte und schrittweise Führung mit Hilfe von Wizard-Dialogen. Dazu werden die erforderlichen Informationen und Daten durch statistische Methoden situations- und nutzergerecht aufbereitet. Anhand eines Tank-, Kontakt- und Terminagenten, sowie eines Restaurant-Recommendere wird die Funktionsweise der Informationsagenten exemplarisch veranschaulicht. Schließlich werden auf Grundlage eines großflächigen Head-Up Displays ein prioritätenorientiertes, multimodales Anzeige- und Bedienkonzept vorgestellt und wesentliche Aspekte der Konzepte in Nutzerstudien evaluiert. Darüber hinaus werden die Möglichkeiten eines quasi-kontaktanalogen Head-Up Displays anhand eines prototypischen Sitzkistenaufbaus aufgezeigt.



# Kapitel 1

## Einleitung

### 1.1 Motivation

Heutzutage ist das Automobil weit mehr als ein reines Beförderungsmittel. Elektronische Entwicklungen aus unterschiedlichsten Lebensbereichen haben seit mehreren Jahrzehnten Einzug in das Fahrzeug gehalten. Das ursprüngliche Röhrenradio wurde längst durch umfangreiche Entertainmentapplikationen wie CD-, DVD-, MP3-Player und TV ersetzt. Auch Navigationssysteme und Autotelefone sind heute keine großen Besonderheiten mehr. Auf den Komfort und die Unterstützung dieser Geräte möchte der Fahrer nicht mehr verzichten. Diese Systeme wollen gerade auch während der Fahrt angemessen bedient werden können.

Das Informations- und Kommunikationszeitalter definiert die Mobilität in einem neuen und erweiterten Ausmaß. Die Etablierung von Kommunikationstechnologien, wie z.B. EDGE, UMTS und W-LAN im Fahrzeug, schafft dort völlig neue Möglichkeiten. Bei manchen Fahrzeugherstellern hat die Integration von Onlinediensten bereits in den letzten Jahren begonnen [BMW08a]. Auskunftsdienste, Reiseführer, E-mail und sogar Internet können so dem Fahrer bereitgestellt werden. Die Etablierung dieser neuen Technologien im Automobil führt jedoch auch weiter zu stark steigenden Informationsmengen für den Fahrer und dies wiederum stellt sehr hohe Ansprüche an die Benutzerschnittstelle.

Möglich wird ein solch moderner Fahrerarbeitsplatz durch die wachsende CPU-Leistung im Fahrzeug und Fortschritte in der Informationsverarbeitung. So können die anfallenden Informationen schnell verarbeitet und angemessen aufbereitet werden. Innovative Anzeigetechnologien und Bedienelemente sowie die Weiterentwicklung von Erkennertechnologien, wie z. B. die Spracherkennung, und multimodaler Systeme, eröffnen neue Wege in der Kommunikation zwischen Fahrer und Fahrzeug.

Bereits vor Jahren haben die Automobilhersteller damit begonnen, zur Bewältigung des immer komplexer werdenden Informations- und Interaktionsumfangs, neue Wege zu beschreiten. Viele Funktionen wurden zusammengefasst und gewohnte Hardkeys<sup>1</sup> durch zentrale multifunktionale Bedienelemente ersetzt. Informationen werden nun meist gebündelt auf einem zentralen Mitteldisplay integriert. Die Dialogführung erfolgt dabei weitgehend über hierarchische Menüstrukturen, die Bedienung über ein multifunktionales Bedienelement. Durch die stark wachsenden Informations- und Datenmengen der sogenannten *Fahrerinformationssysteme* kommen diese Dialogstrategien jedoch mehr und mehr an ihre Grenzen.

---

<sup>1</sup>Tasten mit fest vorgegebener Funktion

## 1.2 Problemstellung

Als zentrale Herausforderung stellt sich nun diese Vielzahl neuer Informationen und multimedialer Applikationen auch während der Fahrt sinnvoll nutzbar und bedienbar zu machen. Durch das vielfältige Informationsangebot ist der Fahrer mit den gegenwärtigen Fahrerinformationssystemen jedoch bereits schon heute überfordert und so sorgen diese Systeme – auch seitens der Medien [Wüs05, Rau04] – immer wieder für Skepsis und Kritik. Die aktuelle Funktionsgruppierung in streng hierarchische Menüstrukturen wird dabei den wachsenden Anforderungen nur unzureichend gerecht [ANR04].

Des Weiteren stellt auch die steigende Verbreitung *mobiler Endgeräte* die Automobilhersteller vor neue Herausforderungen. Aktuelle Trends der Unterhaltungselektronik sind neben multifunktionalen Mobiltelefonen (z. B. Apple iPhone), auch PDAs und portable MP3-Player (z. B. Apple iPod), die der Nutzer unterwegs überall mit sich tragen kann. Mitten im Trend von umfangreichen Kontaktdatenbanken und tragbaren Navigationssystemen muss hier das Fahrerinformationssystem seinen Platz rechtfertigen. Wenn der Benutzer sein Adressbuch und seinen elektronischen Reiseführer sowohl in der Tasche als auch im Auto verwenden kann – was für eine Funktion erfüllt da noch das Fahrerinformationssystem?

Die Anforderungen des Fahrers an ein solches System erwachsen aus den besonderen Umständen der Fahrsituation, und diese werden von den tragbaren Geräten auf keinen Fall erfüllt. Die Aufgabe des Fahrerinformationssystems ist es daher, die umfangreichen Daten und Funktionen, die im Auto genutzt werden wollen, auch fahrergerecht zu präsentieren und bedienbar zu gestalten. Dies umfasst zum Einen die Möglichkeit, dass auf die Daten und Funktionen der Geräte einfach zugegriffen werden kann, zum Anderen die sinnvolle Einbettung in den Kontext der Fahrsituation. Dabei darf das Fahrzeuginterface den Fahrer jedoch keinesfalls bei der Fahrzeugführung beeinträchtigen.

## 1.3 Zielsetzung und Aufgabenstellung

Wesentliches Ziel dieser Arbeit ist die Erarbeitung von geeigneten Lösungsstrategien und Konzepten zur Bewältigung der steigenden Informations- und Datenmengen verschiedenster Art und Herkunft im Automobil.

Aus der wachsenden Informationsmenge resultieren insgesamt die Fragestellungen *Was*, *Wann*, *Wo* und *Wie* im Auto sinnvoll angezeigt werden sollte. Anhand der Abbildung 1.1 kann der schematische Informationsfluss erläutert werden. Zunächst bedarf es der Identifikation relevanter und erforderlicher Informationen entsprechend der Fahrsituation. Danach sollte ein passender Informationszeitpunkt (*Timing*) gewählt werden, der eine Überlastung des Fahrers vermeidet. Hier sollte auf adäquate Kommunikationskanäle (*Modalität*) und Anzeigorte zurückgegriffen werden, die weitgehend mit der Fahrzeugführung vereinbar sind. Schließlich muss entschieden werden in welcher Art und Weise die Informationen aufbereitet und verständlich präsentiert werden. Die anfallenden Informationen, Daten und auch Funktionen sollen dem Fahrer durch ein geeignetes Informationsmanagement kontextgerecht zugänglich gemacht werden. Informationsmanagement kann dabei wie folgt definiert werden:

„Information management (IM) is the collection and management of information from one or more sources and the distribution of that information to one or more audiences.(...) Management means the organization of and control over the structure, processing and delivery of information.“ [Wik08]

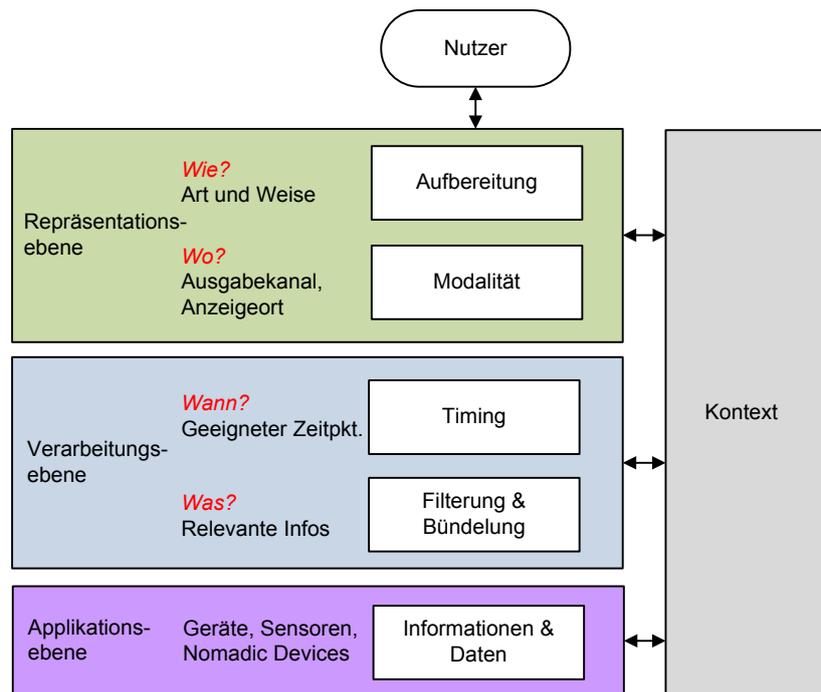


Abbildung 1.1: Informationsfluss im Fahrzeug

Um ein übergreifendes Informationsmanagement im Fahrzeug zu erreichen, werden in dieser Arbeit geeignete Methoden zur Informationsverarbeitung und -filterung aufgegriffen, passende Dialogstrategien für den Fahrer identifiziert und multimodale Anzeige- und Bedienkonzepte erarbeitet. Anschließend werden die verwendeten Ansätze in zahlreichen Usability-Studien einer umfassenden Bewertung unterzogen.

Zunächst liefert das Kapitel 2 die notwendigen Grundlagen zum exakten Verständnis der Fahrer-Fahrzeug Schnittstelle und zeigt Methoden und Kriterien zur Bewertung dieser Systeme auf. Des Weiteren wird in Kapitel 3 ein multimodales, kontextadaptives Dialogframework eingeführt, das die erforderliche Informationsverarbeitung und prototypische Realisierung der neuen Dialogansätze ermöglicht.

Eine zukünftige Herausforderung stellt das Auffinden und Suchen relevanter Informationen im Fahrzeug dar. Trends zum Umgang mit großen Datenmengen und Methoden zur Auslegung der Interfaces liefern dabei PC- und Internet-Anwendungen. Analog zu einer Internet-Suchmaschine werden daher in Kapitel 4 multimodale Interfaces für eine datenübergreifende Suchmaschine im Fahrzeug entwickelt. Eine besondere Problemstellung von Suchmaschinen stellt dabei die alphanumerische Eingabemethode dar. Auch Algorithmen und Methoden zur „intelligenten“ Verarbeitung der Suchanfragen spielen eine wesentliche Rolle. Im Fahrzeug können beispielsweise sehr einfache Eingabefehler auftreten, die durch Methoden zur Verarbeitung unscharfer Eingaben möglichst abgefangen werden sollten. Zur Aufbereitung des Suchergebnisses müssen geeignete Methoden zur Darstellung gewählt werden.

In Kapitel 5 wird ein erweiterter Ansatz vorgestellt, der die spezifischen Anforderungen geräteübergreifender Bedienungsaufgaben in bestimmten Fahrsituationen aufgreift. Dabei gilt es zunächst eine einfache Grundstruktur des Interfaces zu schaffen, welche die Vielzahl an Daten und Funktionen mobiler Endgeräte zusammenführt und dem Fahrer einen explorativen Zugriff

ermöglicht. Es stellt sich dabei die Frage, wie eine Reihe zusammenhängender BediENAufgaben unter Berücksichtigung des Kontextes einfach durchgeführt werden können. Leistungsfähige Filterstrategien, welche die großen Datenmengen effektiv reduzieren und eingrenzen, sind hier erforderlich.

Ähnlich einem „intelligenten Beifahrer“ wäre es wünschenswert, dass technische Assistenten den Fahrer bei seinen Aufgaben weitgehend autonom unterstützen. Die Verwendung von „intelligenten“ Informationsagenten sollen den Fahrer gerade bei häufig wiederkehrenden Handlungsabläufen entlasten und dabei relevante Informationen vorfiltern. Dazu soll eine effiziente Dialogführung den Fahrer aufgabengerecht leiten. Hierzu werden in Kapitel 6 am Beispiel eines Tank-, Kontakt- und Terminagenten und eines Restaurant-Recommenders Lösungsansätze entwickelt. Dabei bedarf es der Auswahl geeigneter mathematischer Verarbeitungsmethoden, welche die erforderlichen Daten kontextgerecht für den Fahrer filtern und aufbereiten.

Neue Methoden der Informationsvisualisierung bieten vielversprechende Ansätze, die neben einer effizienten Informationspräsentation auch einen großen Anteil an Bedienfreude mit sich bringen (siehe dazu Kapitel 7). Auch im Fahrzeug können diese Methoden helfen, umfangreiche Datenmengen einfach und effizient zu erfassen. Daher werden zur Repräsentation von Daten- und Menüstrukturen auf Grundlage verschiedener Visualisierungsstrategien neuartige Fahrzeuginterfaces vorgestellt. Eine wesentliche Herausforderung stellt dabei auch die Identifikation geeigneter, multimodaler Bedienstrategien dar.

Durch neue Anzeigetechnologien im Fahrzeug wie dem Head-Up Display (HUD) eröffnen sich völlig neue Möglichkeiten. In Kapitel 8 werden auf Basis eines großflächigen HUDs innovative Anzeige- und Bedienkonzepte erarbeitet. Mit dem Anzeigekonzept soll eine situations- und nutzergerechte Informationspräsentation ermöglicht werden. Durch multimodale Bedienung können die entwickelten Informationsagenten in das Konzept integriert werden. Darüber hinaus sollen die Möglichkeiten und Anforderungen eines quasi-kontaktanalogen HUDs anhand eines prototypischen Sitzkistenaufbaus aufgezeigt werden. Auftretende Fragestellungen der erarbeiteten Lösungskonzepte werden in diversen Studien bewertet.

Die beschriebenen Aufgabenstellungen liefern Beiträge für ein ganzheitliches Informationsmanagement im Fahrzeug. Diese lassen sich entsprechend Abbildung 1.2 drei Ebenen zuordnen.

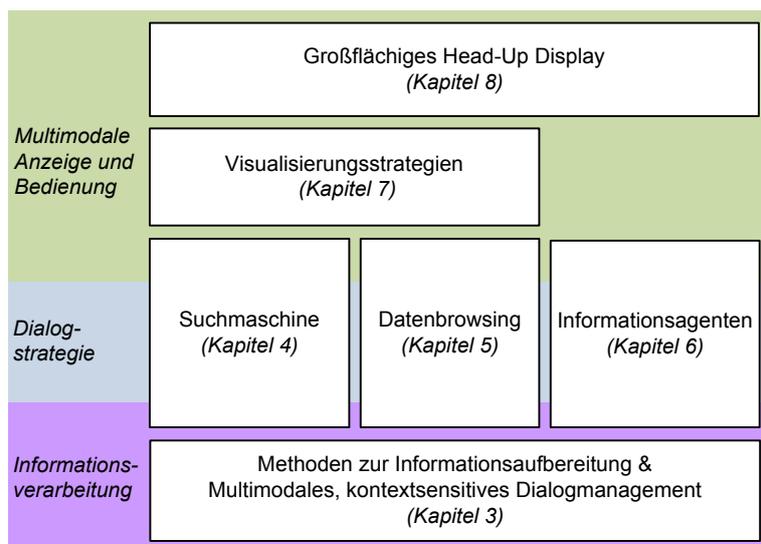


Abbildung 1.2: Informationsmanagement auf drei Ebenen

## Kapitel 2

# Einführung und Grundlagen

Für ein genaues Verständnis der Fahrer-Fahrzeug Schnittstelle gibt dieses Kapitel zunächst eine Einführung in das Forschungsgebiet der Mensch-Maschine Interaktion. Die hier erläuterten Grundlagen stellen eine wesentliche Voraussetzung für die Entwicklung geeigneter Ansätze zur fahrgerechten Bereitstellung von Informationen dar. Dazu werden wichtige Handlungsmodelle aufgezeigt und die verschiedenen Aufgaben des Fahrers erläutert. Als Weiteres erfolgt die Beschreibung gegenwärtiger Entwicklungstrends bei Anzeige- und Bedienkonzepten im Fahrzeug und der dabei verwendeten Dialogstrategien. Durch die Analyse von Fahrerbedarfen, Einflussgrößen und Fahrerszenarien werden die situativen Anforderungen des Fahrers ersichtlich. Kognitionswissenschaftliche Grundlagen schaffen ein erweitertes Verständnis für die Verarbeitungsprozesse des Fahrers, die wesentlich für die sichere Auslegung der Nutzerschnittstellen im Fahrzeug sind. Der Abschnitt Usability liefert Leitlinien zur Bewertung von Fahrerinformationssystemen. Vorgaben für die Gestaltung geben hier Normen, Guidelines, Richtlinien und auch Gestaltgesetze. Abschließend werden Bewertungsmethoden erläutert, mit denen die nachfolgenden Lösungsansätze eingehend untersucht werden können.

### 2.1 Mensch-Maschine Interaktion

Das Forschungsgebiet der Mensch-Maschine Interaktion (MMI)<sup>1</sup> beschäftigt sich mit der nutzergerechten Gestaltung von interaktiven Systemen und der Entwicklung geeigneter Benutzerschnittstellen, den sogenannten Mensch-Maschine Schnittstellen. Lenkrad, Gaspedal und Bremse sind zum Beispiel die Schnittstellen zur Steuerung des Fahrzeuges auf der Straße. Das Fachgebiet Human-Computer Interaction (HCI) kann nach [HBC<sup>+</sup>96] folgendermaßen definiert werden:

„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.“ [HBC<sup>+</sup>96]

In dem äußerst interdisziplinären Forschungsfeld fließen Erkenntnisse u. a. aus der Informationstechnik, der Informatik, der Psychologie, der Arbeitswissenschaft, der Kognitionswissenschaft, der Ergonomie, der Soziologie und des Designs mit ein. Durch die Verwendung geeigneter Ein- und Ausgabekanäle soll die Interaktion möglichst natürlich und intuitiv für den Nutzer gestaltet werden (siehe Abschnitt 3.2). Dabei wirken in einem Mensch-Maschine

---

<sup>1</sup>engl. *Human-Machine Interface* (HMI)

System (MMS) der Mensch und ein technisches System zusammen, um ein bestimmtes Ziel zu erreichen bzw. eine Aufgabe zu lösen [Gei90] (siehe dazu das Passungsmodell im Abschnitt 2.1.1).

Zur Modellierung eines MMS muss sowohl die technische als auch die menschliche Komponente abgebildet werden. Abhängig von den Entwicklungs- und Untersuchungszielen wird dabei auf geeignete Modelle zurückgegriffen, die entsprechende Aspekte menschlichen Verhaltens und dessen Fähigkeiten abbilden. Beispielsweise wird auf Basis des Drei-Ebenen-Modells nach Rasmussen im Abschnitt 2.1.2 die Anpassungsfähigkeit des Menschen an kognitiv unterschiedlich anspruchsvolle Aufgaben beschrieben.

### 2.1.1 Passungsmodell der Mensch-Maschine Interaktion

Bei der Interaktion zwischen Mensch und Maschine gilt es, den *Benutzer*, das *System* und die zu leistende *Aufgabe* in Einklang zu bringen [Gei90]. Wie in Abbildung 2.1 zu sehen ist, stehen diese drei Aspekte durch die Aufgabenbewältigung, die Benutzungsfreundlichkeit und die Funktionalität des Systems untereinander in Beziehung. Oberstes Ziel ist es, eine optimale Passung zwischen diesen drei Aspekten zu erreichen. Dieser Zusammenhang wird daher auch als *Passungsmodell* der Mensch-Maschine Interaktion bezeichnet (nach Frese und Brodbeck [FB89]).

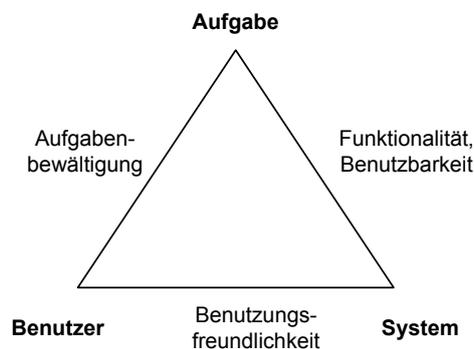


Abbildung 2.1: Passungsmodell der Mensch-Maschine Interaktion

Bereits in einer frühen Entwicklungsphase müssen die Aspekte der Gebrauchstauglichkeit und der Bedienungsfreundlichkeit berücksichtigt werden. Während des Entwicklungsprozesses werden daher Usability-Untersuchungen (mehr dazu in den Abschnitten 2.4 und 2.5) auf Basis von Prototypen durchgeführt, und die Erkenntnisse fließen in die Weiterentwicklung mit ein. Gemäß dem *Spiralmodell* ist die Entwicklung ein iterativer Prozess und sieht eine zyklische Wiederholung der Phasen des *Wasserfallmodells* vor [Ba198]. Die einzelnen Phasen sind Analyse, Design, Implementierung, Test und Wartung.

Bei der Gestaltung eines MMS ist vor allem die Kenntnis der Nutzergruppe sehr wichtig. Das Design der Benutzungsoberfläche muss abhängig von den Grundfertigkeiten und dem Kenntnisstand der Zielgruppe ausgeführt werden. Weiterhin muss den zu bewältigenden Aufgaben Rechnung getragen werden.

### 2.1.2 Drei-Ebenen-Modell nach Rasmussen

Die Verhaltensweisen und Handlungen des Menschen im MMS werden nach Jens Rasmussen (1983) in drei Ebenen eingeteilt (siehe Abbildung 2.2) [Ras83]:

- **Wissensbasiertes Verhalten (*knowledge-based behavior*):**  
bewusst gesteuerte Verhaltensweisen, die auf analytischen Prozeduren beruhen
- **Regelbasiertes Verhalten (*rule-based behavior*):**  
Anwendung gelernter Regeln, wie z. B. Wenn-Dann-Regeln
- **Fähigkeitsbasiertes Verhalten (*skill-based behavior*):**  
hochgeübte, automatisierte Verhaltensweisen, Routinehandlungen

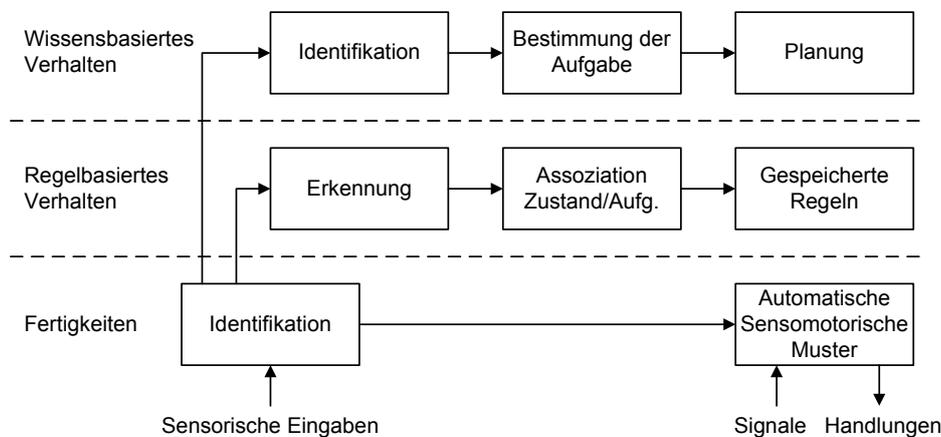


Abbildung 2.2: Drei-Ebenen-Modell nach Rasmussen [Ras83]

Auf der untersten Handlungsebene werden Kontrolltätigkeiten und stark automatisierte Gewohnheitshandlungen ausgeführt. Das *Fähigkeitsbasierte Verhalten* erfordert dabei keinerlei Steuerung oder willentliche Aufmerksamkeit des Menschen. Dagegen ist *Regelbasiertes Verhalten* gekennzeichnet durch Handlungsschemata in vertrauten Situationen und die Anwendung erlernter *Wenn-Dann-Regeln*. Eingehende Informationen (Ereignisse oder Zustände) werden dabei qualitativ reduziert und als Zeichen wahrgenommen. Die meisten Aufgaben von MMS stellen Regelbasiertes Verhalten dar. Auf der obersten Ebene erfolgt *Wissensbasiertes Verhalten*, und dies wird beim Fehlen geeigneter Regeln sowie in unbekanntem Situationen erforderlich. Informationen über das technische System und die Umgebung werden als Symbole für komplexe strukturelle und funktionelle Zusammenhänge wahrgenommen. Durch die Anwendung mentaler Wissensmodelle (siehe Abschnitt 2.3.5) werden Entscheidungen bezüglich der durchzuführenden Aufgabe getroffen. Dieses Modell dient als Grundlage für viele weitere Theorien und kann auch auf die Fahrzeugführung übertragen werden.

## 2.2 Fahrzeugdomäne

Allgemeine Modelle der MMS lassen sich auf die Fahrzeugdomäne transferieren und dort anwenden. Dieser Abschnitt geht zunächst auf vorherrschende Modelle der Fahrzeugführung ein. Anschließend werden Fahrerinformations- und Fahrerassistenzsysteme charakterisiert und

aktuelle Entwicklungstrends der Automobilhersteller aufgezeigt. Die Analyse und Erläuterung von Fahrerbedarfen und Einflussgrößen liefern Anhaltspunkte zu den situativen Anforderungen und Aufgaben des Fahrers.

### 2.2.1 Drei-Ebenen-Modell der Fahrzeugführung

Das Führen eines Fahrzeugs ist eine komplexe Angelegenheit, und hierbei laufen mehrere Prozesse parallel ab. Ein weit verbreitetes Modell dazu ist das Drei-Ebenen-Modell der Fahrzeugführung, das eine Unterteilung der Fahraufgabe analog zum Modell nach Rasmussen in drei Ebenen vornimmt [Ber70, Mic85, Erk93].

Diese sind (siehe auch Abbildung 2.3):

1. **Navigationsebene**
2. **Führungsebene**
3. **Stabilisierungsebene**

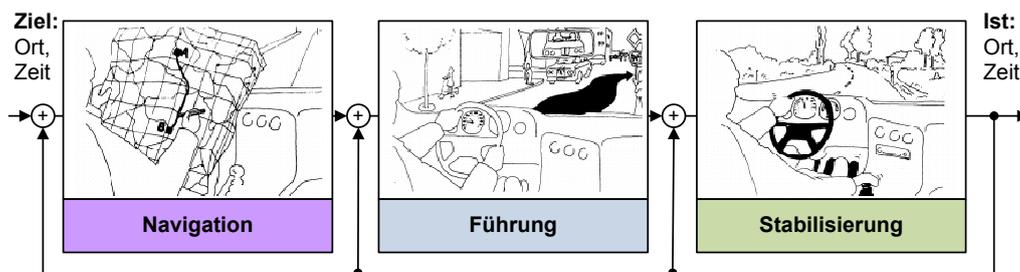


Abbildung 2.3: Überblick Drei-Ebenen-Modell der Fahrzeugführung [Bub08a]

Auf der *Navigationsebene* erfolgt die Festlegung der gewünschten Route durch den Fahrer. Die Wahl der Strecke erfordert die Berücksichtigung verschiedener Aspekte. Beispielsweise ist die zu erwartende Dauer der Fahrt abhängig von der Tageszeit. Ggf. müssen Zwischenziele (z. B. Anfahrt einer Tankstelle) eingeplant werden oder auch Gefahren einer Strecke (z. B. Straßenglätte) berücksichtigt werden. Auch kann die Umfahrung eines Staus erforderlich werden. Dabei spielt ebenfalls der Zweck der Fahrt eine weitere Rolle bei der Planung der Strecke.

Auf der *Führungsebene* wird die gewünschte und erforderliche Fahrtroute umgesetzt. Dabei folgt der Fahrer dem Verlauf der Straße und muss entsprechend weiterer Verkehrsteilnehmer reagieren bzw. ausweichen. Teilaufgaben wie dem Wechsel einer Spur, das Überholen, das Abbiegen an Kreuzungen und die Berücksichtigung von Verkehrszeichen gehören zur Führungsebene.

Die *Stabilisierungsebene* entspricht der untersten Ebene des Modells. Diese teilt sich in Längs- und Querverführung auf. Bei der *Längsführung* kontrolliert der Fahrer seine Geschwindigkeit und den Abstand zum Vordermann. Die seitlichen Abstände des Fahrzeugs zu Fahrspur (dies entspricht der *Querverführung*) werden durch Gas geben, Bremsen, Lenken, Kuppeln und Schalten geregelt.

Insgesamt stellen die drei Ebenen unterschiedliche *Anforderungen* an den Fahrer. Zunächst erfolgt das bewusste Festlegen einer bestimmten Route auf der Navigationsebene. Dies beansprucht ein hohes Maß an mentalen Ressourcen. Auch Handlungen auf der Führungsebene

werden weitgehend bewusst ausgeführt. Dabei müssen Verkehrssituationen und Verkehrsregeln wahrgenommen und interpretiert werden und erfordern entsprechende Reaktionen und Handlungen. Hier ist die Inanspruchnahme kognitiver Ressourcen jedoch geringer als bei Tätigkeiten der Navigationsebene. Dagegen werden auf der Stabilisierungsebene hoch automatisierte Handlungen ausgeführt. Sie sind dem Fahrer oftmals nicht bewusst und erfordern daher nur sehr geringe mentale Kapazitäten.

Unbewusste Prozesse erfolgen schnell und ohne Anstrengung. Auch mehrere Aufgaben können hier gleichzeitig ausgeführt werden. Dagegen laufen bewusste Prozesse (wie z. B. die Orientierung in einer fremden Stadt) seriell ab. Sie erfordern die Aufmerksamkeit des Fahrers und sind daher langsam und arbeitsaufwendig. Für diese Prozesse steht nur eine begrenzte Kapazität zur Verfügung. [ZDZ97, Grü05]

### 2.2.2 Fahraufgaben im Automobil nach Geiser

Inhaltlich lassen sich die Aufgaben des Fahrers nach Geiser in drei (Teil-)Aufgaben untergliedern (siehe Abbildung 2.4) [Gei90]:

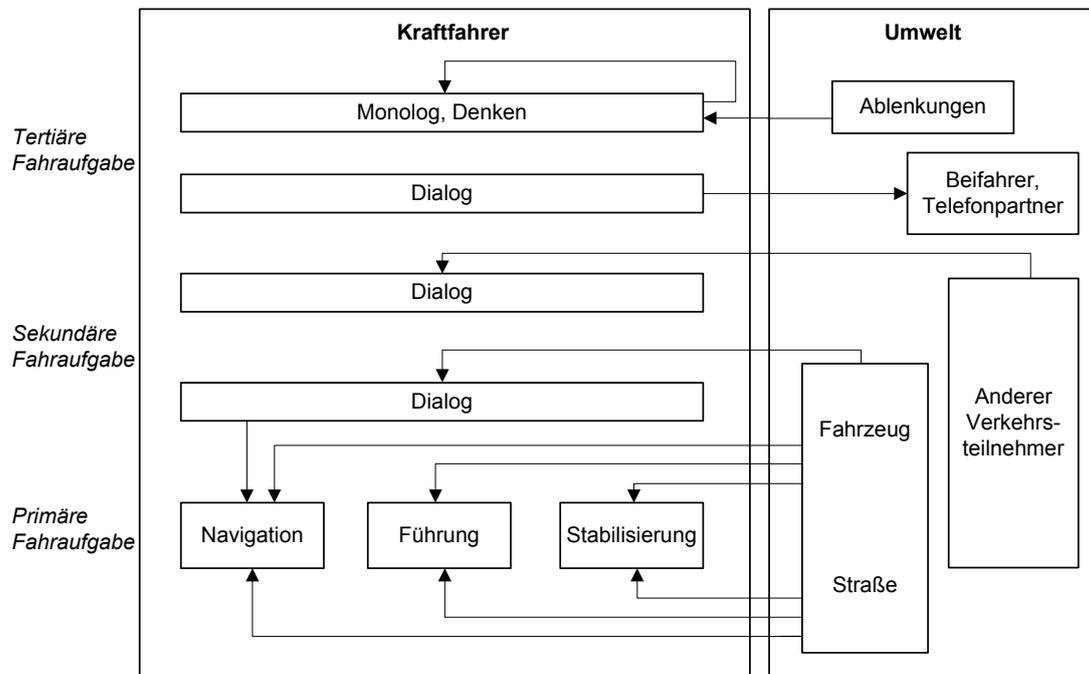


Abbildung 2.4: Fahraufgaben nach Geiser [Gei90]

**Primäre Fahraufgabe:** Die primäre Fahraufgabe entspricht dem eigentlichen Fahrprozess, wie dem Halten des Fahrzeuges auf der Fahrbahn und der Stabilisierung (Längs- und Querregelung) (siehe Abschnitt 2.2.1).

**Sekundäre Fahraufgabe:** Daneben sind vom Fahrer noch weitere Aufgaben zu erledigen. Sie fallen zwar im Rahmen der Fahraufgabe verkehrs- bzw. umweltbedingt an, dienen jedoch nicht dem eigentlichen Halten des Fahrzeugs auf der Straße. Sie werden deshalb den sekundären Aufgaben zugerechnet. Dazu zählen sämtliche Stellteile die in indirekter oder direkter Abhängigkeit von der Fahraufgabe zu bedienen sind, wie z. B. die Betätigung des Blinkers,

des Scheibenwischers, des Lichtschalters oder der Hupe. Auch die Betätigung von Fahrerassistenzsystemen und die Ausführung weiterer Kontrollaufgaben, wie das Ablesen der Drehzahl und der Öltemperatur, gehören dazu.

**Tertiäre Fahraufgabe:** Die tertiären Aufgaben stehen nicht mit der Fahraufgabe in Verbindung, sondern dienen lediglich der Zufriedenstellung des Komfort-, Unterhaltungs- und Informationsbedürfnisses. Dazu zählen beispielsweise die Betätigung der Heizung und Klimaanlage, des Radios und des Telefons. Auch weitere in Zukunft vermehrt hinzukommende Applikationen aus dem Kommunikations- und Internetbereich, sowie der Büro- und Haustechnik können der tertiären Fahraufgabe zugeordnet werden. Ein Großteil der tertiären Aufgaben kann über Fahrerinformationssysteme bedient werden, die nachfolgend beschrieben werden.

In der Literatur wird oftmals auch von *Hauptaufgabe* bzw. *Primäraufgabe* und *Nebenaufgabe* bzw. *Sekundäraufgabe* gesprochen. In der Hauptaufgabe werden dabei die erläuterte primäre und sekundäre Fahraufgabe zusammengefasst. Als Nebenaufgabe wird dabei die tertiäre Fahraufgabe bezeichnet.

### 2.2.3 Fahrerinformationssysteme

Die Fahrerinformationssysteme (FIS)<sup>2</sup> sind technische Systeme, die im Fahrzeug umfangreiche Informations-, Kommunikations-, Unterhaltungs- und Komfortfunktionen bereitstellen [Nie02]. Diese integrieren Informationen der tertiären Fahraufgabe, können allerdings auch zu einem gewissen Grad weitere unkritische Information und Funktionen aus dem primären und sekundären Bereich beinhalten. Daher werden oftmals auch Teile der Navigationsaufgabe zu den FIS hinzugerechnet (z. B. die Bedienung der Zieleingabe). Die Interaktion mit dem FIS erfolgt heute über ein integriertes Anzeige- und Bedienkonzept. Dieses System fasst beispielsweise Radio-, CD-, Telefon-, Navigations- und Klimafunktionen in einem Interface zusammen. Die Funktionen sind meist über eine überwiegend hierarchische Menüstruktur zugänglich. Dabei erfolgt die Anzeige über ein zentrales Display, die Bedienung über ein multifunktionales Bedienelement, mit dem über eine möglichst konsistente Bedienlogik durch das Menü navigiert werden kann. Es gibt verschiedene technische Ausprägungen dieser FIS. Im Abschnitt 2.2.6 werden die gegenwärtigen Entwicklungstrends und Grundzüge aktueller FIS verschiedener Automobilhersteller<sup>3</sup> vorgestellt.

Bei der Auslegung von FIS muss darauf geachtet werden, dass diese Systeme keine bedeutende Beanspruchung darstellen und der erforderliche Bedienaufwand in einem vertretbaren Verhältnis zur primären Fahraufgabe bleibt. Deshalb liegt das wesentliche Augenmerk dieser Arbeit an einer effizienten und fahrergerichten Gestaltung der Dialoge und des Interfaces.

### 2.2.4 Fahrerassistenzsysteme

Unter Fahrerassistenzsystemen (FAS)<sup>4</sup> versteht man elektronische Zusatzeinrichtungen in Kraftfahrzeugen, welche den Fahrer bei der Fahrzeugführung unterstützen und entlasten. Bei diesen Systemen erfolgt meist eine Assistenz in der primären Fahraufgabe.

Ziel der FAS ist es in erster Linie Unfälle zu vermeiden, die aufgrund von Fahrfehlern oder unerwarteten Ereignissen im Straßenverkehr verursacht werden. Sie helfen dem Fahrer das

<sup>2</sup>engl. *In-Vehicle Information Systems* (IVIS) or *Driver Information Systems* (DIS)

<sup>3</sup>auch OEM, engl. *Original Equipment Manufacturer* genannt

<sup>4</sup>engl. *Advanced Driver Assistance Systems* (ADAS) or *Driver Assistance Systems* (DAS)

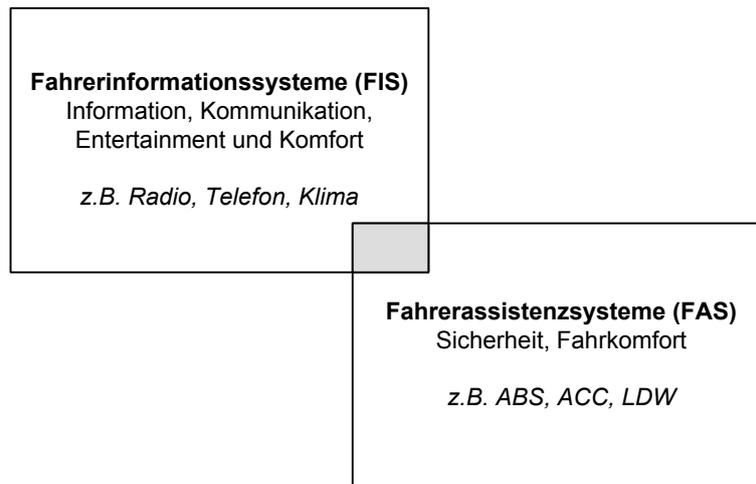


Abbildung 2.5: Fahrerinformations- und Fahrerassistenzsysteme

Fahrzeug unter Kontrolle zu halten, indem sie aktiv in die Längs- oder Querregelung eingreifen bzw. durch Warnungen oder Systemanpassungen, wie z. B. Druckaufbau im Bremssystem, die Reaktionszeit des Fahrers verkürzen. Darüber hinaus tragen diese Systeme auch dazu bei, den Fahrkomfort zu erhöhen.

Aktuelle FAS sind z. B. Geschwindigkeitsregelanlage (GRA) oder Tempomat, Adaptive-Cruise-Control (ACC), Collision-Warning (CW), Park-Distance-Control (PDC), Park-Assistent (PA), Night-Vision (NV) und Adaptives Kurvenlicht. Die FAS zählen zu den *aktiven* Sicherheitssystemen, da sie in direkter Weise dazu beitragen können, Unfälle zu vermeiden. Dagegen können *passive* Sicherheitssysteme wie Airbag und Gurtstraffer lediglich die Unfallfolgen mildern.

Die Assistenzsysteme lassen sich gemäß der Art und Weise des Eingriffs in *informierend*, *beratend*, *kommandierend* und *intervenierend* untergliedern. Informierende und beratende FAS werden zum Teil auch den FIS zugeordnet. Die Überschneidung von FIS und FAS wird in Abbildung 2.5 verdeutlicht.

Assistenzsysteme können auch für weitere Aufgaben des Fahrers, die neben der primären auch die sekundäre und tertiäre Fahraufgaben mit einschließen, realisiert werden. Im Kapitel 6 werden erweiterte Assistenzsysteme in Form von Informationsagenten umgesetzt, welche durch kontextgerechte Datenfilterung und effiziente Dialoge den Fahrer entlasten sollen.

### 2.2.5 Fahrerzustandsüberwachung

Eine Besonderheit stellen Systeme zur Fahrerzustandsüberwachung (engl. *driver monitoring*) dar, die sich aktuell noch überwiegend im Stadium der Forschung befinden. Zum Beispiel können diese Systeme den Wachsamkeitszustand des Fahrers beobachten und bei einer bedenklichen Absenkung eine Warnung ausgeben oder in kritischen Fällen direkt in die Fahrzeugführung eingreifen. Im Gegensatz zu den oben angeführten Systemen nutzen sie nicht den Vergleich des Ist-Fahrzustandes des Fahrzeugs mit einem aus der jeweiligen Situation technisch abgeleiteten Soll, sondern den Vergleich des beobachteten Fahrerverhaltens (z. B. Lid-schlagfrequenz) mit einem definierten Sollverhalten (z. B. Solllidschlagfrequenz) des Fahrers. Sie sind überwiegend den primären Aufgaben des Autofahrers zuzuordnen, da die Warnung oder der Eingriff auf dieser Ebene stattfindet. In erster Linie geht es bei diesen Systemen

darum, Unzulänglichkeiten heutiger FAS zu beheben und Absichten des Fahrers besser zu erkennen. So kann ein Spurwechselassistent beispielsweise eine Spurverlassenswarnung unterdrücken, falls der Fahrer willentlich die Spur wechselt und dies zuvor mit einem Spiegel- und Schulterblick abgesichert hat [TMRK07a]. Aber auch bei tertiären Systemen gewinnt die Fahrerzustandsüberwachung an Bedeutung. Bei der Bedienung von FIS kann auf Emotionen oder die Ablenkung des Fahrers Rücksicht genommen und so Fehler vermieden bzw. bei kritischem Verhalten auch Warnungen ausgegeben werden. Auch die Informationsmenge könnte entsprechend der Belastung des Fahrers richtig dosiert werden und Meldungen zu geeigneten Zeitpunkten erscheinen. Diese Systeme erfordern natürlich die Detektion des Fahrerzustandes, beispielsweise über Blickverfolgungssysteme (siehe dazu die Konzeption eines Gazetrackers in Abschnitt 8.4.4), Emotionserkennung, Headtracking, aber auch Systeme zur physiologischen Zustandsbestimmung wie z. B. die Bestimmung der Herzfrequenz, usw.

### 2.2.6 Entwicklungstrends bei Fahrerinformationssystemen

In heutigen FIS werden die Funktionen meist auf einem zentral positionierten Display, dem *Central Information Display* (CID), angezeigt und über ein zentrales Bedienelement gesteuert. Durch ein einheitliches Anzeige- und Bedienkonzept (ABK) soll die Kontrolle der Vielzahl an Funktionen und Informationen ermöglicht werden. Vor allem in Limousinen der Ober- und Mittelklasse haben solche Systeme Einzug gehalten. Die ABKs verschiedener Hersteller unterscheiden sich vor allem in Anzeigeort, Bedieneinheit und -logik. Dabei zeigen sich aktuell zwei Entwicklungstrends.

Die Hersteller Audi, BMW und Daimler setzen auf die *Trennung von Anzeige und Bedienung* im Bereich der Mittelkonsole [Aud08, BMW08b, Dai08]. Die Bedienung erfolgt hier über ein *zentrales Bedienelement* (ZBE oder auch *Controller* genannt), meist in Form eines Dreh-/Drückstellers, der an einem ergonomisch gut zu erreichenden Ort im unteren Bereich der Mittelkonsole nahe dem Gangwahlhebel platziert ist. Der Bildschirm des CID befindet sich dagegen am oberen Rand der Mittelkonsole, damit der Blick nicht zu weit nach unten abgewendet werden muss. Teilweise sind wichtige Funktionen zusätzlich als *Hardkey* verfügbar. In Abgrenzung dazu werden oftmals auch *Softkeys* verwendet, deren Funktion dynamisch angepasst wird.

Bei Modellen von Toyota, Lexus und neuerdings auch VW und Ford hingegen wird die *Anzeige und Bedienung in Form eines Touchscreens direkt an einem Ort integriert* [Toy08, Lex08, Vol08, For08]. Bei beiden Ansätzen sind die umfangreichen Funktionen in einer weitgehend *hierarchischen Menüstruktur* untergebracht. Nachfolgend werden aktuell verfügbare FIS erläutert.

**BMW iDrive:** Das BMW iDrive wurde erstmals im Jahre 2001 in der 7er Reihe vorgestellt [BMW08b]. In der darauf folgenden 5er Reihe wurde das FIS weiter überarbeitet. Das *i* steht dabei für *intelligent, integriert, informativ, innovativ, intuitiv* und *interaktiv*. Das iDrive integriert die Steuerung des Entertainment- (z. B. Radio, CD-, DVD-Player und TV), Kommunikations- (wie Telefon), sowie des Navigations- und Teile des Klimasystems. Des Weiteren sind verschiedene Systemeinstellungen möglich. Die grafische Darstellung erfolgt zentral auf einem 8 Zoll großen Farbdisplay in der Mitte der *i-Tafel*<sup>5</sup>. Als zentrales Bedienelement wird ein Dreh-/Drücksteller verwendet, welcher neben Drehen und Drücken außerdem ein Schieben ermöglicht. Hinter dem Controller ist ein Menü-Knopf vorhanden, über den aus

<sup>5</sup>Abkürzung für Instrumententafel, Armaturen



(a) Fahrzeug-Cockpit



(b) iDrive-Controller

Abbildung 2.6: Das BMW iDrive der 5er Serie [BMW08b]

beliebigen Untermenüs in das Hauptmenü zurückgesprungen werden kann. Daneben gibt es eine *Push-To-Talk*-Taste (PTT-Taste) zur Aktivierung der optionalen Sprachbedienung. Die Menüstruktur ist hierarchisch und überwiegend geräte- und funktionsorientiert gegliedert. Aus dem Hauptmenü gelangt man durch Schieben in die Untermenüs Kommunikation, Navigation, Entertainment, Klima und durch Drücken des Controllers zu den Einstellungen. Anschließend ist es möglich, durch Drehen des Controllers in den Untermenüs und Listen zu navigieren und durch anschließendes Drücken eine Auswahl zu treffen. Durch Schieben des Controllers kann durch das sogenannte *Pebble-Hopping* zwischen Bildschirmsegmenten und Menüebenen gewechselt werden. Texteingaben erfolgen durch den sogenannten *Speller*, dessen Funktionsweise im Abschnitt 4.4 beschrieben wird.

**Audi MMI:** Das Audi MMI (Abkürzung für *MultiMedia Interface*) kommt mit vollem Funktionsumfang im Modell A8 zum Einsatz [Aud08]. Das System besteht ebenfalls aus einer räumlich getrennten Bedien- und Anzeigeeinheit. Als Anzeigeelement wird dabei ein ausklappbarer 7 Zoll Farbbildschirm verwendet. Die Bedienung erfolgt über die sogenannte *Quattro Logic*. Als zentrales Steuerelement dient ein ergonomisch günstig platzierter Dreh-/Drücksteller in der Mittelkonsole, der im Gegensatz zum iDrive-Controller jedoch nicht geschoben werden kann. Über diesen Dreh-/Drücksteller kann ebenfalls durch Menüs navigiert und durch Lis-



(a) Fahrzeug-Cockpit



(b) MMI-Controller

Abbildung 2.7: Das Audi MMI des A8 [Aud08]



(a) Fahrzeug-Cockpit

(b) COMAND-Controller

Abbildung 2.8: Das Mercedes COMAND der S-Klasse [Dai08]

ten gescrollt werden. Durch das Drücken des Controllers wird die Auswahl bestätigt. Um den Drehknopf sind vier Softkeys platziert, die in den verschiedenen Menükontexten variable Funktionen annehmen können und die Navigation in Untermenüs ermöglichen. Unterhalb befinden sich noch eine Return-Taste und Tasten für Skiplen nach Links und Rechts. Darüber hinaus sind zwei Hardkey-Leisten angeordnet, die den Einsprung in die Hauptmenüs Navigation, Radio, CD/TV, Info, Setup, Adressbuch, Car und Telefon ermöglichen. Texteingaben erfolgen ähnlich dem iDrive über einen Speller. Die Buchstaben sind hier allerdings kreisförmig angeordnet.

**Mercedes COMAND:** Seit Herbst 2005 baut Mercedes eine Weiterentwicklung des COMAND (*COckpit MANagement and Data System*) in die Fahrzeuge der S-Klasse ein und folgt somit auch dem aktuellen Trend zu integrierten FIS [Dai08]. Es besteht aus einer separaten Bedieneinheit an der Mittelarmlehne, sowie einer 8 Zoll großen Anzeigeeinheit. Das Hauptbedienelement ist ein Dreh-/Drücksteller, der sich auch schieben lässt, was jedoch mit einem leichten Kippen des Knopfes einhergeht. Am oberen Bildschirmrand befindet sich eine Steuerleiste mit den Untermenüepunkten. Durch Drehen kann man sich innerhalb der Menüs und Listen bewegen, über horizontales Schieben kann man zwischen den parallelen Menüebenen springen, während durch vertikales Schieben zwischen den Menüebenen gesprungen werden kann. Vor dem Dreh-/Drücksteller sind vier Hardkeys platziert, über die alternative, direkte Einsprünge in die Menüebenen Radio, CD und DVD, Telefon und Navigation möglich sind. Außerdem gibt es noch eine frei konfigurierbare Favoritentaste, sowie eine Returntaste. Teile der Klimabedienung, wie Luftverteilung und Standheizung sind in das COMAND integriert, der Rest ist durch direkte Knöpfe ausgelagert. Als Weiteres besitzt das System noch eine separate Telefontastatur.

**Lexus Multiinformationsdisplay:** Das Lexus FIS verfügt über ein 8 Zoll großes Touchscreen und daher erfolgt die Anzeige und Bedienung – anders als bei den vorherigen Systemen – am selben Ort (siehe Abbildung 2.9) [Lex08]. Aufgrund der ergonomischen Erreichbarkeit ist das Display etwas niedriger positioniert. Links und rechts des Displays sind jeweils vier Hardkeys untergebracht, mit denen man direkt in die jeweiligen Menüs springen kann. In den Untermenüs erfolgt die Bedienung direkt über Buttons auf dem Touchscreen. Eine alphanumerische Eingabe ist über ein *On-Screen-Keyboard*<sup>6</sup> (OSK) mittels Touchscreen möglich, das

<sup>6</sup>graphische Darstellung einer QWERTY-Tastatur am Bildschirm



Abbildung 2.9: Das Lexus LS Cockpit [Lex08]

bei Bedarf eingeblendet wird. Ein kleiner Teil der Klimafunktionen ist in das Multifunktionsdisplay integriert.

### Kombiinstrument

Als Kombiinstrument, kurz Kombi genannt, wird die Ansammlung von Instrumentenanzeigen hinter dem Lenkrad bezeichnet. Es setzt sich meist aus Tachometer, Drehzahlmesser, Temperaturanzeige, Tankanzeige und Warnlampen zusammen, also Anzeigen für die primäre und sekundäre Fahraufgabe. Die Kombis verschiedener Hersteller und Fahrzeugtypen unterscheiden sich einerseits in der Anordnung und Größe der Anzeigen, andererseits in der Informationsdichte und Technologie. Gründe für die Anordnung sind auf Design und technische Randbedingungen zurückzuführen und orientieren sich an der Unterstützung verschiedener Aspekte der Fahraufgabe. Wurden die Informationen in den vergangenen Jahren hauptsächlich durch analoge Anzeigeelemente und Warn-/Hinweislämpchen realisiert, so halten heute in zunehmendem Maße digitale Anzeigen Einzug in das Kombi (siehe z. B. [Aud08] und [Dai08]). Graphikfähige Displays, die Teile oder auch den gesamten Bereich des Kombis ausfüllen, werden zukünftig an Bedeutung gewinnen und eröffnen neue Darstellungsmöglichkeiten. Diese Entwicklung führt zunehmend auch dazu, dass immer mehr Informationen und Funktionen der tertiären Fahraufgabe, die bisher im CID angesiedelt waren, im Kombi platziert werden. Zur Platzierung der unterschiedlichen Informationen an sinnvollen Anzeigeeorten im Fahrzeug liefert das Kapitel 8 neue Lösungsansätze.

### 2.2.7 Zukünftige Fahrerbedarfe

Zukunftsstudien wie Expertenbefragungen geben Anhaltspunkte für die Weiterentwicklung von FIS und FAS [BRT<sup>+</sup>04, Zöl04]. Assistenzsysteme für die primäre Fahraufgabe werden um Situationsbewusstsein und -adaption erweitert und weiterentwickelt werden. Vor allem Funktionen rund um das Heading Control und die Car-2-Car Vernetzung werden eine zentrale Rolle spielen. Diese Systeme sollen dem Fahrer Entscheidungshilfen bieten und in gefährlichen Situationen korrigierend eingreifen.

Im tertiären Aufgabenbereich erwarten die Prognosen die übergreifende Vernetzung und Integration von Informations- und Kommunikationsdiensten und -medien. Diese system- und dienstübergreifende Integration dient zunächst vorrangig komplexen Infotainmentanwendun-

gen im Sinne eines ganzheitlichen Mission-Supports. Fahrer und Beifahrer werden durch die Bereitstellung von Zusatzinformationen mit hochaktuellen Reiseinfos versorgt (z. B. elektronischer Reiseführer). Darüber hinaus werden Wartungs- und Instandhaltungsinformationen Teil einer vernetzten Kommunikationsstruktur. Hinsichtlich der Bereitstellung fahrerrelevanter Informationen erwarten die Experten detailliertere und zuverlässigere Informationen über die Routen-, Verkehrs- und Umweltbedingungen. Hierbei werden zunehmend verschiedene Kontextgrößen mit einfließen.

### 2.2.8 Einflussgrößen und Kontext im Fahrzeug

Auf das Fahrzeug und den Fahrer wirken verschiedene Einflussgrößen und Kontextgrößen, die hier erläutert und klassifiziert werden. Für eine nutzer- und situationsgerechte Auslegung der Fahrzeuginterfaces ist die Berücksichtigung dieser Größen sehr wichtig. Diese lassen sich zunächst abstrakt in System-, Umwelt- und Nutzerkontext untergliedern.

#### System-, Umwelt- und Nutzerkontext

Beim Führen eines Fahrzeugs sind verschiedene Faktoren zu berücksichtigen, die auf das MMS einwirken. Diese Größen haben unterschiedlich ausgeprägten Einfluss auf die jeweiligen Komponenten und werden deshalb – gemäß der Relation – System-, Umwelt- und Nutzerkontext zugeordnet. Bei diesen Kontextgrößen interessieren nur die zum aktuellen Zeitpunkt auftretenden Werte.

**Systemkontext:** Hier finden sämtliche, das technische System betreffenden Größen, Berücksichtigung. Beispielsweise ist dies die Verfügbarkeit der Systemkomponenten, die aktuell angesprochen werden können und zur Verfügung stehen. In einem FIS kann dies z. B. die Auslastung des Fahrzeugcomputers oder der aktuelle System- bzw. Dialogzustand sein.

**Umweltkontext:** Im Umweltkontext sind diejenigen Faktoren zu finden, die von außen auf Nutzer und Fahrzeug einwirken. Dazu zählen z. B. Verkehrsdichte, Straßenbeschaffenheit, Witterungsverhältnisse oder auch das Verhalten anderer Verkehrsteilnehmer, aktuelle Positionsdaten, usw. Oftmals wird hier auch von Situationskontext gesprochen.

**Nutzerkontext:** Zum Nutzerkontext gehören sämtliche Faktoren, die etwas über den Zustand des Nutzers aussagen. Dazu zählen der emotionaler Zustand des Fahrers wie Stress, aber auch dessen Verhalten und Präferenzen. Ein spezielle Herausforderung ist stellt hier die Erfassung dieser Größen dar.

#### Gliederung in acht Bereiche

Insgesamt lassen sich die Einflussgrößen auch in acht fahrzeugspezifische Bereiche einteilen [BRT<sup>+</sup>04]:

- **Auftrag:**  
Motiv, Charakter der Fahrmission
- **Fahrermerkmale:**  
physische Merkmale, soziodemographische Merkmale, Persönlichkeit
- **Soziale Situation:**  
soziale Rahmensituation, soziale Gruppe

- **Fahrzeugmerkmale:**  
Fahrzeugtyp, Abmessungen/Gewicht, Technische Daten/Ausstattung, Verschleiß/Verbrauch, Transport, Fahrdynamik
- **Routenplanung:**  
Routenplanungsdaten, saisonale Besonderheiten, Zeitdruck
- **Fahrsituation:**  
Geographie, Meteorologie, Route, Kommunikationsnetz
- **Verkehrssituation:**  
Konstellation, Fahrzeugumgebung
- **Fahrtchronologie**

**Zeitliche Untergliederung – Fahrtchronologie:** Die Kategorie der Fahrtchronologie setzt sich lediglich aus einer Dimension zusammen. Sie beschreibt den momentanen Standpunkt innerhalb der zeitlichen Abfolge einer Fahrt. Eine Unterteilung in sechs Phasen erschien in diesem Zusammenhang angemessen, um eine eindeutige Abgrenzung und Zuordnung zu bewahren. Diese sind *vor der Fahrt  $t_v$* , *beim Antritt der Fahrt  $t_b$* , *während der Fahrt  $t_w$* , *kurz vor Fahrtende  $t_e$* , *nach der Fahrt  $t_n$*  und *zwischen den Fahrten  $t_z$* . Die Ergebnisse der Bedarfsanalyse (siehe Abschnitt 2.2.7) haben gezeigt, dass die Anforderungen an die Fahrzeugsysteme sehr eng mit der zeitlichen Situation korrelieren.



Abbildung 2.10: Überblick Fahrtchronologie

**Untergliederung nach Auftrag:** Die Kategorie Auftrag definiert allgemein den Sinn und Zweck einer Fahrt (siehe 3x3-Fahrtszenarienmatrix in Abbildung 2.11). Die dort vorgenommenen Unterteilungen des Motivs sind dienstlich, privat und aus Spaß. Als Weiteres hängt sie von der Form der Beteiligten (Einzelperson, Personengruppe und mit Objekttransport) ab. Die Matrix ermöglicht die einfache Identifikation von bedeutenden Szenarien und die Zuordnung von Situationen.

### Fahrtszenarien

Die Lösungskonzepte dieser Arbeit wurden auf Basis folgender Fahrtszenarien entwickelt. Dazu wurden diese vier repräsentativen Szenarien ausgewählt:

- *Alltagsfahrt zur Arbeit* (siehe Abbildung 2.12)
- *Geschäftsfahrt zum Flughafen*

Beteiligte Auftrag	Einzelperson	Personen- gruppe	Objekt
Dienstlich			
Privat			
Aus Spaß			

Abbildung 2.11: Fahrscenarienmatrix nach Bubb [BRT<sup>+</sup>04]

- *Ausflugsfahrt mit Shopping*
- *Urlaubsfahrt mit der Familie*

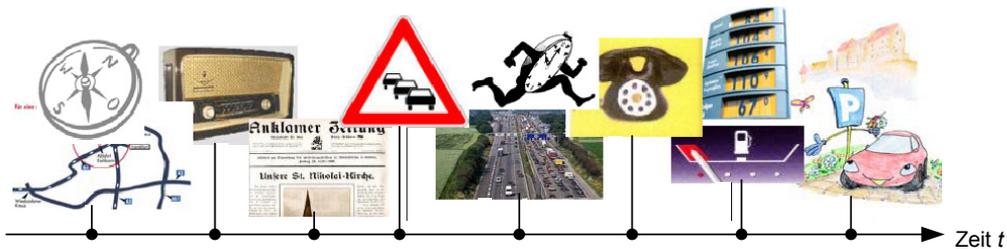


Abbildung 2.12: Beispielhafter Verlauf eines Szenarios (Geschäftsfahrt)

Auf Basis dieser Szenarien lassen sich die situativen Anforderungen und die jeweiligen Informationsbedarfe des Fahrers ableiten. Sie liefern Anhaltspunkte zur Bündelung von Aufgabensequenzen und der erforderlichen Information im jeweiligen Kontext. Zur Entwicklung der Ansätze des Datenbrowsings in Kapitel 5 und der Informationsagenten in Kapitel 6 wird auf diese Szenarientechnik zurückgegriffen.

## 2.3 Kognitionswissenschaftliche Grundlagen

Für das Verständnis der Verarbeitungs- und Denkprozesse des Fahrers werden in diesem Abschnitt kognitionswissenschaftliche Grundlagen erläutert. Diese Grundlagen müssen bei der Auslegung der Interfaces, wie z. B. der Wahl eines geeigneten Informationskanals (siehe Abschnitt 3.2), unbedingt berücksichtigt werden, um eine Vereinbarkeit mit der Fahraufgabe zu gewährleisten. Zunächst werden daher der Begriff Aufmerksamkeit definiert und etablierte Verhaltensmodelle beschrieben. Anschließend wird auf die Aspekte Situation Awareness und Workload, sowie mentale Modelle eingegangen.

### 2.3.1 Aufmerksamkeit

Aufmerksamkeit entspricht der Zuweisung von Ressourcen des Gehirns zur Wahrnehmung und kognitiven Verarbeitung [Sty97, Bro58]. Da die Verarbeitungskapazität beschränkt ist, werden relevante Reize aktiv ausgewählt und irrelevante Reize ausgeblendet. Insgesamt wird zwischen selektiver, geteilter und fokussierter Aufmerksamkeit unterschieden. Werden Reize auf Kosten anderer Informationen ausgewählt und verarbeitet, wird das als *selektive Aufmerksamkeit* bezeichnet. *Geteilte Aufmerksamkeit* bedeutet, dass die Aufmerksamkeit zwischen mehreren Informationen gesplittet wird. Bei der *fokussierten Aufmerksamkeit* wird der Fokus auf ganz bestimmte Informationen gelenkt. Diese verschiedenen Formen spielen während des Fahrens eine große Rolle. Gerade bei Sekundäraufgaben ist der Aspekt der geteilten Aufmerksamkeit von zentraler Bedeutung.

### 2.3.2 Verhaltensmodelle

Verhaltensmodelle bilden wesentliche Funktionen des Menschen nach und bieten eine gute Grundlage für das Verständnis der Aufmerksamkeit und Belastbarkeit des Fahrers.

**Ein-Kanal-Theorie:** Die Ein-Kanal-Theorie von Broadbent (1958) war das erste entwickelte Ressourcenmodell [Bro58]. Simultan eintreffende Informationen können nur auf der Rezeptorebene gleichzeitig verarbeitet werden. Mit Hilfe eines Filters werden die Informationen anschließend selektiert und entsprechend weitergeleitet. Jedoch lassen sich damit viele parallele Prozesse und Aufgaben nicht erklären.

**Working Memory:** Das Modell des Working Memory besteht aus einer zentralen Einheit, die als eine kontrollierende Einheit fungiert [BH74]. Für einen kurzen Zeitraum ist sie für die momentane Informationsverarbeitung und Speicherung von Informationen zuständig und beeinflusst zwei untere Ebenen. Auditorische und verbale Informationen werden direkt vom *Phonological Loop* verarbeitet. Das *Visual Spatial Working Memory* ist dafür zuständig, die visuellen und räumlichen Eigenschaften von Objekten, wie z. B. Umriss und Orientierung, zu erfassen. Diese Theorie erklärt unser Verständnis und die Interaktion mit unserer Umwelt. Das Working Memory hilft uns, neue Fähigkeiten zu erlernen und alte abzurufen, um Probleme zu lösen und Ziele zu erreichen.

**Multiple Resource Theory:** Wickens untersuchte die Ressourcenzuteilung bei parallelen Aufgaben. Aus seinen Experimenten entwickelte er ein qualitatives Modell, das die Informationsverarbeitung in drei Informationsübertragungsstufen unterteilt [Wic84]. Die Einteilung erfolgt in *Wahrnehmung* (Encoding), *zentrale Informationsverarbeitung* (Central Processing) und *Antwortausführung* (Responding) (siehe Abbildung 2.13).

Die Seite der Informationsaufnahme wird in die Aufnahmemodalitäten, *visuell* und *auditiv*, und den Verarbeitungs-codes, *räumlich* und *verbal*, unterteilt. Die Antwort- bzw. Ausführungsmodalitäten sind in manuelle und spracherzeugende Codes gegliedert. Die begrenzte Verfügbarkeit der Ressourcen wird durch den Würfel repräsentiert. Wickens geht davon aus, dass Prozesse in mehreren Zellen unabhängig voneinander parallel ausgeführt werden können. Aufgaben, die nicht die gleichen Ressourcenpools in Anspruch nehmen, können ohne Einbußen in der Leistung parallel verarbeitet werden. Die Fähigkeit zur gleichzeitigen Bearbeitung von Aufgaben ist unterschiedlich stark ausgeprägt. Je stärker die Interferenz zweier Aufgaben ist, umso mehr gemeinsame Ressourcen werden von ihnen beansprucht. Beispielsweise können Spazieren gehen und Unterhalten gleichzeitig ausgeführt werden, während die Fahrzeugführung

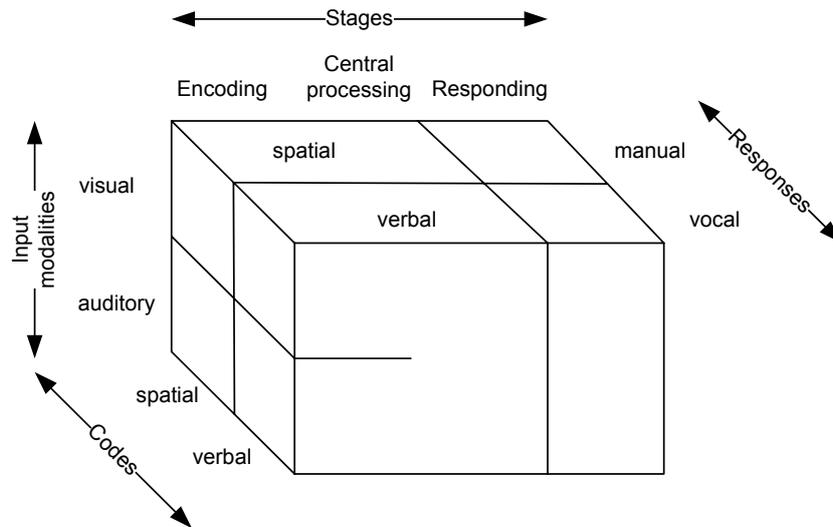


Abbildung 2.13: Aufteilung in Ressourcenpools der Multiple Resource Theory [Wic84]

und die Bedienung des Radios zu Leistungseinbußen führt, da sie auf dieselben Ressourcen zurückgreifen.

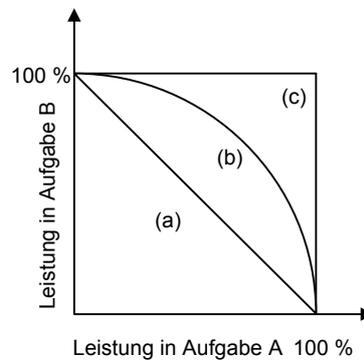


Abbildung 2.14: POC-Kurve für zwei simultan bearbeitete Aufgaben A und B

Die Abhängigkeiten zweier Aufgaben bei gleichzeitiger Bearbeitung (*Timesharing*) können in Form von *Performance Operation Characteristics* (kurz POC-Kurven) graphisch repräsentiert werden [LN81]. In Abbildung 2.14 ist dies für zwei Aufgaben dargestellt, die für sich alleine zu 100% erfüllbar sind. Greifen die Aufgaben auf abhängige Ressourcen zu, ergeben sich die Arbeitskurven (a) und (b). Hier führt der Kapazitäteneinsatz der einen Aufgabe zu einer Verringerung der Leistungsfähigkeit bei der Anderen. Während in der Kurve (a) kein Timesharing von Ressourcen möglich ist, ist dies in Kurve (b) zum Teil möglich. Herrscht dagegen ein optimales Timesharing, d. h. in beiden Aufgaben kann die volle Leistung erbracht werden, so ergibt sich Kurve (c).

### 2.3.3 Situation Awareness

Das Situationsbewusstsein (engl. *situation awareness*) beschreibt den Zustand oder den Prozess des Fahrers, sich seiner Umgebung zutreffend bewusst zu sein. Um die Ablenkung des Fahrers zu bewerten, ist dieser Aspekt von besonderem Interesse [End88, Wic96]. Insgesamt werden drei Ebenen unterschieden [BPK06]: *Wahrnehmungsebene*, *Verstehensebene* und die *Generierung von Annahmen*. Auf jeder dieser Ebenen kann Fehlverhalten entstehen. Ein häufiger Fehler bei kognitiver Ablenkung ist *looked-but-did-not-see*, d. h. der Fahrer blickt zwar zum relevanten Objekt, hat aber offenbar die Bedeutung des Objekts nicht erkannt [BPK06].

### 2.3.4 Workload

Unter einer Arbeitsbelastung (engl. *workload*) versteht man die Gesamtheit aller Einflüsse, die in einem Arbeitssystem auf den Menschen einwirken [Sch93]. Die Arbeitsbelastung summiert sämtliche Teilbelastungen und besteht aus wahrnehmbaren und nicht wahrnehmbaren Faktoren. Qualitativ erfassbare Teilbelastungen bezeichnet man als Belastungsfaktoren während quantifizierbare Teilbelastungen als Belastungsgrößen bezeichnet werden.

Obwohl die Arbeitsbelastung für alle Menschen bei gleicher Tätigkeit gleich ist, ist die Wirkung auf den Einzelnen sehr unterschiedlich und wird als Arbeitsbeanspruchung bezeichnet. Bei hoher Belastung sollte der Fahrer vor jeder unnötigen Ablenkung geschützt werden. Daher werden heute „intelligente“ Fahrzeugsysteme entwickelt, die bei großer Arbeitsbelastung, etwa wenn der Fahrer in eine scharfe Kurve einbiegt oder zum Überholen ansetzt, Telefonate oder Informationen zurückhalten können. Diese Systeme werden als *Workload-Manager* bezeichnet [TMRK07b].

### 2.3.5 Mentale Modelle

Ein mentales Modell (engl. *mental model*) entsteht bei der Interaktion mit einem System und wird fortlaufend durch die Erfahrung eines Nutzers erweitert. Mentale Modelle helfen dem Anwender dabei, Entscheidungen zu treffen oder Situationen zu interpretieren [Nor02]. Dieser Aspekt gewinnt mit zunehmender Komplexität technischer Systeme verstärkt an Bedeutung [Ras05]. So kann der Nutzer technische Systeme antizipieren, Strategien entwickeln und Erklärungen für Zusammenhänge finden [Cha92].

## 2.4 Usability – Normen, Gestaltungshinweise und Richtlinien

Der Fokus technischer Entwicklungen lag anfangs meist auf der Realisierung der geforderten Funktionalität. Hierbei genügte das Vorhandensein eines bestimmten Features oder einer Funktion zur Vermarktung. Doch heutzutage werden erhöhte Anforderungen hinsichtlich einer einfachen und intuitiven Gebrauchstauglichkeit gestellt. Dieser Umstand ist vor allem darin begründet, dass Kunden die Produkte, mit denen sie nicht zufrieden sind oder mit denen sie nicht zurecht kommen, zunehmend unberücksichtigt lassen [JJ95]. Im Fahrzeugbereich geht es jedoch im Besonderen um die Sicherheit. Die hier erläuterten Aspekte wie Usability, Normen, Gestaltungshinweise und Richtlinien liefern Leitlinien zur Gestaltung der Benutzerschnittstellen.

### 2.4.1 Definition Usability

Usability (zu deut. Gebrauchstauglichkeit) kann wie folgt definiert werden:

„Gebrauchstauglichkeit ist das Ausmaß, in dem ein Produkt durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um bestimmte Ziele effektiv, effizient und zufriedenstellend zu erreichen.“ [ISO98] (siehe dazu auch Abschnitt 2.4.2)

Die drei Ziele *Effektivität*, *Effizienz* und *Zufriedenheit* orientieren sich hauptsächlich am Anwender und kennzeichnen insgesamt die Nutzungsqualität einer Anwendung [Dah06]. Diese hängt also davon ab, wie schnell und zielstrebig ein Benutzer sein Ziel erreichen kann, und wie wohl er sich dabei fühlt. Nielsen hat in seinem Buch „Usability Engineering“ ein Grundwerk über die Gestaltung und Bewertung der Gebrauchsfähigkeit von technischen Geräten geschaffen [Nie93]. Usability setzt sich laut Nielsen aus mehreren Komponenten zusammen und ist eng mit messbaren Charakteristika eines interaktiven Anwendungssystems verbunden. Er bezeichnet diese Komponenten als *Usability Faktoren* und zeigt auf, wie vielseitig das Verständnis des Begriffs Usability ist. [Nie93]

### 2.4.2 Normen

Normen sind formale, von Standardisierungsbehörden veröffentlichte Dokumente, die durch einen Prozess aus Konsens und formalen Abstimmungen entwickelt wurden. Zur Festlegung von Normen existieren zahlreiche Verbände und Organisationen. Zu den wichtigsten gehören die *International Organization for Standardization* (ISO), das *Deutsche Institut für Normung e. V.* (DIN) sowie das *American National Standards Institute* (ANSI). Die folgenden ISO Normen sollen bei der software-ergonomischen Gestaltung und der Umsetzung von grundlegenden Eigenschaften helfen [Dah06]. Eine der bedeutendsten Normen für die Software-Ergonomie ist die *ISO 9241* mit ihren vielen Teilnormen. Daneben gibt es Normen zur Produktqualität, ISO 12119 und ISO 9216, und zur benutzerorientierten Entwicklung die ISO 13407. Für die allgemeine Gestaltung von Bildschirmarbeitsplätzen, worunter auch die Fahrerinformationssysteme fallen, existiert die ISO Norm 9241 mit den Titel „Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten“.

#### Anforderungen an die Gebrauchstauglichkeit (ISO 9241 – Teil 11)

Die oben genannte Definition von Usability ist in Teil 11 der ISO Norm 9241 eingebunden. Die drei Usability-Faktoren Effektivität, Effizienz und Zufriedenheit sind dabei wie folgt definiert [ISO98]:

- *Effektivität* bezeichnet die Genauigkeit und Vollständigkeit, mit welcher Benutzer ein bestimmtes Ziel erreichen.
- *Effizienz* ist das Verhältnis des vom Nutzer eingesetzten Aufwands, mit welcher er ein bestimmtes Ziel erreicht.
- *Zufriedenheit* entspricht der Freiheit von Beeinträchtigung und der positiven Einstellung gegenüber der Benutzung des Produkts.

Diese Kriterien legen einen allgemeinen Bewertungsrahmen zur Evaluierung der Gebrauchsfähigkeit von interaktiven Geräten fest.

### Grundsätze der Dialoggestaltung (ISO 9241 – Teil 10)

Richtlinien zur Dialoggestaltung sind in der ISO Norm 9241 in Teil 10 niedergelegt [ISO06]. Eine Überblick mit kurzer Erläuterung dieser Gestaltungsrichtlinien ist in Abbildung 2.15 zu sehen.

ISO 9241-10	Erwartungskonformität	Konsistente Verwendung von Begriffen und Bedienweisen
	Selbstbeschreibungsfähigkeit	Unmittelbare Verständlichkeit der Bedienschritte
	Aufgabenangemessenheit	Effektive und effiziente Aufgabenerledigung
	Lernförderlichkeit	Geringer Lernaufwand und einprägsame Bedienstrategien
	Individualisierbarkeit	Anpassung an Fähigkeiten und Vorlieben der Nutzer
	Fehlerrobustheit	Unempfindlich gegenüber Bedienfehlern
	Steuerbarkeit	Kontrolle des Dialogablaufs liegt jederzeit beim Nutzer

Abbildung 2.15: Grundsätze der Dialoggestaltung, Zusammenfassung ISO9241 Teil 10 [Nie02]

### 2.4.3 Guidelines

Guidelines sind Empfehlungen aufgrund bewährter Konzepte, die sich auf die Glaubwürdigkeit und Autorität der Autoren verlassen. Ben Shneiderman von der Universität Maryland formulierte einige Regeln, die generell für die Gestaltung von Bedienoberflächen gelten [Shn98]. Nachfolgend werden diese „Eight Golden Rules for Dialogue Design“ vorgestellt [Her94]:

1. *Versuche Konsistenz zu erreichen*  
Konsistente Befehlsabläufe in ähnlichen Situationen; Identische Terminologie in Menüs
2. *Ermögliche regelmäßigen Nutzern die Verwendung von Shortcuts*  
Reduzierung der Interaktionsschritte und Erhöhung des Tempos
3. *Biete informatives Feedback*  
Informatives Feedback für jede Handlung
4. *Entwerfe abgeschlossene Dialoge*  
In Gruppen organisierte Handlungssequenzen
5. *Biete einfache Fehlervermeidung und -behandlung*  
Vermeidung von Fehlern; Einfache Behebung, falls Fehler auftritt
6. *Erlaube einfache Umkehr von Aktionen*  
Handlungen sollten rückgängig gemacht werden können

7. *Unterstütze benutzergesteuerte Dialoge*

Der Benutzer sollte das Gefühl der Kontrolle haben

8. *Reduziere die Belastung des Kurzzeitgedächtnisses*

Begrenze die Zahl der Einträge und halte den Bildschirm einfach

#### 2.4.4 Styleguides

Styleguides sind ein Satz an Empfehlungen von Softwareherstellern oder Entwicklungsorganisationen, welche die Konsistenz des Designs verstärken und die Umsetzung bewährter Konzepte in einem Designprozess unterstützen. Beispiele dafür sind der Windows User Experience Guide, die Apple Human Interface Guidelines und das OSF/Motif Style Guide. [Dah06]

#### 2.4.5 Gestaltgesetze

In diesem Abschnitt werden Fragestellungen der Wahrnehmung und Räumlichkeit diskutiert. Die wichtigsten Gestaltgesetze sind [Gol02, Met75]:

- *Gesetz der Nähe:*  
Objekte die räumlich (oder zeitlich) benachbart sind, werden als zusammengehörend erlebt.
- *Gesetz der Gleichheit und Ähnlichkeit:*  
Ähnliche bzw. gleichartige Objekte werden eher als zusammengehörend und dadurch als eine Gruppe wahrgenommen.
- *Gesetz der guten Fortsetzung:*  
Harmonische Anordnung von Objekten führt zur Wahrnehmung als ein Objekt oder eine Linie.
- *Gesetz der Geschlossenheit:*  
Eine unvollständige Figur wird in der Wahrnehmung ergänzt.
- *Gesetz der Symmetrie:*  
Zwischenräume zwischen symmetrischen Konturen werden eher als Objekt wahrgenommen, als die zwischen nicht-symmetrischen Konturen.
- *Texturgradient:*  
Änderungen von Texturen unterstützen die Tiefenwahrnehmung.
- *Virtuelle Konturen:*  
Entstehung von Objekten durch zarte Andeutung.

#### 2.4.6 Bedienfreude - Joy of Use

*Joy of Use* kann als die Freude an der Bedienung umschrieben werden. [Hat00] widmet sich der Bedienfreude für Softwareanwendungen. Der Joy of Use ergibt sich durch das ausgezeichnete Funktionieren und die optisch und ästhetisch ansprechende Gestaltung. Wichtig sind jedoch auch persönlichkeitspsychologische Aspekte der Person, wie etwa die Neugier des Nutzers, aber auch dessen Geschlecht und Expertise. Der Kontext, in dem das System eingesetzt wird, darf auch nicht außer Acht gelassen werden. Beispielsweise macht Stress das Erleben von Joy of Use schwierig. [Hat00, HBB01]

Bei vorhandener Bedienfreude ist die Effizienz, die Motivation und die Zufriedenheit der Nutzer im Umgang mit dem Produkt erhöht. Daher stellt Joy of Use eine gewichtige Größe bei der Gestaltung technischer Systeme dar. Im Abschnitt 2.5.3 wird mit dem Semantischen Differential eine Methode vorgestellt, welche die Bewertung dieser Aspekte ermöglicht.

### 2.4.7 Empfehlungen und Standards für FIS

Für FIS sind neben zahlreichen Guidelines, die sich mit allgemeinen Darstellungs- und Bedienkonzepten beschäftigen, folgende Empfehlungen und Standards veröffentlicht (ein detaillierter Überblick findet sich in [SGK<sup>+</sup>04, Nie02]):

- Prinzipien zur Gestaltung von Fahrerinformations- und -assistenzsystemen nach ISO/DIS 15005 [ISO03a] und 15008 [ISO03b].
- Die Empfehlungen der *Alliance of Automobile Manufacturers* [All03] und der Europäischen Kommission fassen wesentliche Sicherheitsaspekte zusammen, die bei der Mensch-Maschine-Schnittstelle im Fahrzeug zu beachten sind
  - Die Empfehlung “Safe and Efficient In-Vehicle Information and Communication Systems” der *Europäischen Kommission* [Eur99]
  - Der Europäische Grundsatzkatalog für die Gestaltung der Mensch-Maschine-Schnittstelle für bordeigene Kommunikations- und Informationssysteme (2006) [Eur06].
- Die Empfehlungen der *Society of Automotive Engineers* J2364 [Soc04] und J2365 [Soc02].
- Die Guidelines des *US-Department of Defence* [Dep99] und des *US-Department of Transportation* [Dep98, GLPS95].
- “Guideline for In-Vehicle Display Systems“ der *Japan Automobile Manufactures Association* [Jap04].

## 2.5 Bewertungsmethoden

Die in dieser Arbeit entwickelten Dialogstrategien und Anzeige- und Bedienkonzepte zum Umgang mit großen Datenmengen müssen hinsichtlich der Eignung für den Fahrer eingehend untersucht und bewertet werden. Dazu wurden etablierte Bewertungsmethoden ausgewählt, welche die Evaluierung von Usability-Kriterien zulassen. Die verwendeten Methoden werden nun in diesem Abschnitt erläutert.

### 2.5.1 Fahr Simulator

Eine sehr bewährte Untersuchungsmethode zur Eignung von Konzeptansätzen im Fahrzeug ist die Durchführung von Studien im Fahr Simulator. Im Rahmen dieser Arbeit wurden dazu zwei neue integrierte Laborräume am Lehrstuhl für Mensch-Maschine-Kommunikation geplant und aufgebaut. Diese bestehen aus einem Fahr Simulator-Labor (siehe Abbildungen 2.18) und einem angegliederten Virtual Reality Labor (mit einer 3-Seiten CAVE, siehe auch Abbildung 9.2). Der schematische Aufbau der beiden Laborräume ist in Abbildung 2.16 zu entnehmen. Wesentliche Bestandteile des Fahr Simulators sind die Sitzkiste und die Projektion einer Sichtsimulation. Ein Teil der Untersuchungen wurde noch im ehemaligen Fahr Simulatorlabor des Lehrstuhls

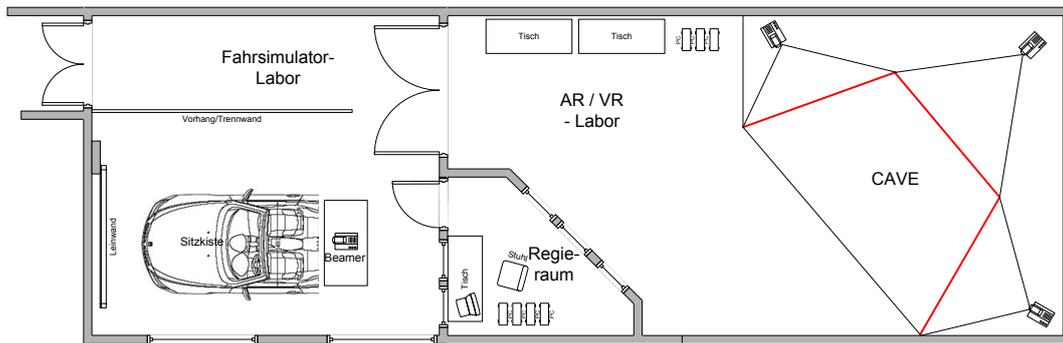


Abbildung 2.16: Überblick Laborräume des Lehrstuhls für Mensch-Maschine-Kommunikation

durchgeführt (bezüglich des Laboraufbaus siehe [Neu01]). Nachfolgend werden eingesetzte Sichtsimulationen aufgeführt.

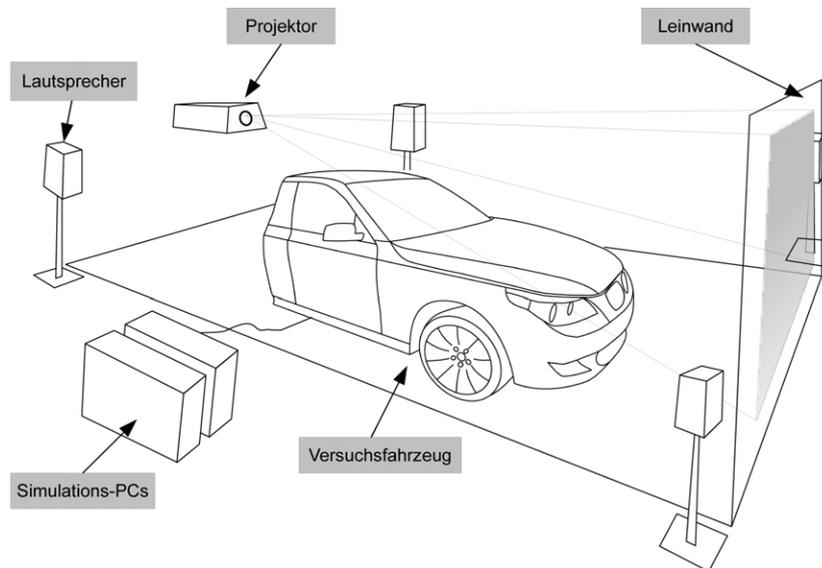


Abbildung 2.17: Schematischer Laboraufbau des Fahr Simulators

**Sichtsimulation Lane Change Task:** Für den Großteil der Untersuchungen wird der Lane Change Task (LCT) als Fahraufgabe verwendet. Dieser wird von der Daimler AG, Research and Technology für wissenschaftliche Untersuchungen zur Verfügung gestellt. Diese Untersuchungsmethode wurde entwickelt, um die Ablenkung des Fahrers, die durch Aufgaben im Fahrzeug verursacht wird, zu messen. [Dai04, ISO02, Lan06]

Die Fahraufgabe besteht aus einer dreispurigen geraden Strecke, wie sie in der Abbildung 2.19(a) zu sehen ist. Am Straßenrand werden auf beiden Seiten Schilder in festen Abständen eingeblendet, die erst in einer Entfernung von 40 m sichtbar werden. Die Schilder geben vor, welche der drei Spuren im nächsten Fahrabschnitt zu befahren ist. Die Höchstgeschwindigkeit von 60 km/h wird durch vollständiges Durchdrücken des Gaspedals erreicht. Die Aufgabe der Versuchspersonen besteht darin, die Maximalgeschwindigkeit zu halten und die Spurwechsel möglichst unmittelbar und exakt durchzuführen. Die mittlere Spurabweichung  $x_{a,i}$  wird anhand



(a) Sitzkiste hinten



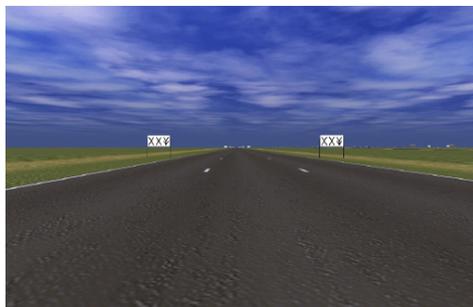
(b) Sitzkiste vorne

Abbildung 2.18: Fahrersimulator-Labor am Lehrstuhl für Mensch-Maschine-Kommunikation

folgender Formel berechnet:

$$\bar{x}_{at} = \frac{1}{T} \sum x_{a,i} \Delta t_i; \text{ Einheit [m] (Wegx; Zeitt)} \quad (2.1)$$

**Sichtsimulation CAVE\_SIM:** Zur Realisierung individueller Streckenabschnitte und Szenarien wurde eine eigene Sichtsimulation entwickelt [PARR07]. Diese basiert auf der sehr leistungsfähigen Unreal-Graphik-Engine. Ein Eindruck der entwickelten Sichtsimulation kann in Abbildung 2.19(b) gewonnen werden. Sie ermöglicht auch die stereoskopische Darstellung in der entstandenen CAVE des VR-Labors (siehe Abbildung 9.2).



(a) Lane Change Task



(b) CAVE\_SIM

Abbildung 2.19: Sichtsimulationen für den Fahrersimulator

### 2.5.2 Peripheral Detection Task

Bei der sogenannten Peripheral Detection Task (PDT) handelt es sich um eine Erkennungsaufgabe. Hierbei werden auf einer Horizontalen im peripheren Sichtfeld fünf Punkte an den Positionen von  $\pm 23^\circ$ ,  $\pm 11,5^\circ$  und  $0^\circ$  eingeblendet. Die Punkte erscheinen in einem Winkel von  $0,9^\circ$ . Diese Punkte werden für ca. 2 s eingeblendet und sollen in dieser Zeit erkannt und über eine Taste am Lenkrad mit der linken Hand bestätigt werden. Nach den ca. 2 s verschwinden die Punkte wieder. Dies erfordert eine schnelle Entdeckung und Reaktion seitens der Testperson. Aus den Reaktionszeiten der Testperson bei den einzelnen Punkten lässt sich folglich eine Aussage über die momentane Belastung treffen. Zudem ist es zum Teil möglich

bei übersehenen Punkten eine Verengung des Blickfeldes nachzuweisen [JOKG05, MW00]. Da in der Literatur meist auf PDT-Versuche bei realen Fahrversuchen verwiesen wird, weicht die Implementierung für den Fahrsimulator davon etwas ab. Damit bei der Durchführung keine Gewöhnungseffekte eintreten und die Bestätigung der Punkte auch nur erfolgt, wenn wirklich ein Punkt angezeigt wird, wurde die Zeit zwischen dem Erscheinen aufeinanderfolgender Punkte variabel umgesetzt. Die Zeit variiert zwischen 4 s und 6 s, ähnlich wie in [JOKG05] vorgeschlagen.

### 2.5.3 Semantisches Differential

Zur Bewertung des Joy of Use bzw. „Erleben und Gefallen“ der entwickelten Benutzerschnittstellen durch Versuchspersonen wird ein semantisches Differential zur fundierten und differenzierten Messung verwendet. Dieses umfangreiche semantische Differential wurde im Rahmen des TUMMIC-Projektes entwickelt [AAB<sup>+</sup>06]. Dabei handelt es sich um 59 bipolare siebenstufige Skalen, die aus gegensätzlichen Adjektiv-Paaren bestehen. Ein Beispiel zeigt Abbildung 2.20.

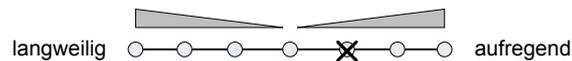


Abbildung 2.20: Beispiel für ein Adjektiv-Paar aus dem Semantischen Differential

Bei der Bewertung eines Systems kreuzt die Versuchsperson einen bestimmten Kreis an, der sich umso näher an einem Pol der Skala befindet, je mehr dieser Begriff auf die Einstellung des Befragten zutrifft. Wichtig ist dabei, dass ein Begriff nicht notwendigerweise wörtlich auf den zu bewertenden Gegenstand zutreffen muss, sondern dass es sich in erster Linie um die konnotative (übertragene) Bedeutung des Begriffs handelt.

### AttrakDiff

AttrakDiff ist ein Online Tool, mit dessen Hilfe ein System durch die Nutzer beurteilt werden kann. Über dieses semantische Differential lässt sich ermitteln, wie attraktiv das getestete Produkt im Hinblick auf Bedienbarkeit und Aussehen empfunden wird, und ob Optimierungsbedarf besteht [Att08, HBK03]. Der standardisierte AttrakDiff Fragebogen besteht aus vielen Wortpaaren, mit deren Hilfe die Beurteilung vorgenommen wird. Die Wortpaare stellen jeweils extreme Gegensätze dar, zwischen denen eine Abstufung möglich ist. Dieses Verfahren hat gegenüber der direkten Befragung den Vorteil, dass die Ergebnisse weniger von der Fragestellung beeinflusst werden. Bei AttrakDiff werden vier unterschiedliche Dimensionen bewertet [HBK03]:

- *Pragmatische Qualität* (PQ) beschreibt die Benutzbarkeit eines Produktes, und verdeutlicht, wie gut der Nutzer seine Ziele mit Hilfe des Systems erreichen kann, oder auch wie kontrollierbar dies ist.
- *Hedonische Qualität - Stimulation* (HQ-S) bildet ab, inwieweit ein Produkt die persönliche Weiterentwicklung unterstützen kann, indem es neuartige, interessante und anregende Funktionalitäten, Inhalte, Interaktions- und Präsentationsstile bietet, z. B. wie innovativ ein System ist.

- *Hedonische Qualität - Identität* (HQ-I) beschreibt, inwieweit ein Produkt seinem Nutzer ermöglicht, sich mit dem Produkt zu identifizieren oder auch wie wertvoll es für ihn ist.
- *Attraktivität* (ATT) beschreibt eine globale Bewertung des Produkts auf der Basis der wahrgenommenen Qualität, z. B. wie sympathisch das Produkt empfunden wird.

#### 2.5.4 Blickverhaltensmessung

Auch die Blickverhaltensmessung ist ein sehr bewährtes Mittel zur Untersuchung von Interfaces im Fahrzeug. In den Untersuchungen kamen das JANUS-Blickverhaltensmessungssystem des Lehrstuhls für Ergonomie [Sch03b] (siehe Abbildung 2.21) und das im Rahmen dieser Arbeit entwickelte remote Gazetracking-System aus Abschnitt 8.4.4 zum Einsatz.



Abbildung 2.21: Das Janus Blickverhaltensmessungssystem [Sch03b]

## 2.6 Zusammenfassung

Dieses Kapitel hat das grundlegende Verständnis der Fahrer-Fahrzeug Schnittstellen erläutert. Hierzu wurden Modelle, die sich bei der Fahrzeugführung anwenden lassen, aufgezeigt. Die anfallenden Aufgaben des Fahrers lassen sich unterschiedlichen Ebenen zuordnen. Die Bedienung von Fahrerinformationssystemen erfordert die bewusste Aufmerksamkeit durch den Fahrer. Dabei liegt das größte Augenmerk darauf, dass bei der Interaktion mit den Fahrerinformationssystemen die Fahrzeugführung keinesfalls beeinträchtigt werden darf. Aus der Betrachtung der kognitionswissenschaftlichen Grundlagen wurde deutlich, dass die Ressourcen der menschlichen Informationsverarbeitung zwar insgesamt begrenzt sind, jedoch eine Aufteilung auf verschiedene Ressourcenpools nicht unmittelbar zu Leistungseinbußen führt. Um eine Überlastung des Fahrers zu vermeiden, müssen Wege gefunden werden, welche die Grenzen der kognitiven Belastbarkeit nicht überschreiten und dabei die beanspruchten Ressourcen nicht mit den Aufgaben der Fahrzeugführung in Konkurrenz treten. Durch den Einbezug von Kontexteinflüssen können die vielseitigen Bedarfe und Aufgaben des Fahrers vereinfacht und dadurch die Belastung reduziert werden. Zur Vermeidung der Beeinflussung bei der Fahrzeugführung müssen geeignete Kommunikationskanäle ausgewählt werden. Der Überblick zu aktuellen Anzeige-/Bedienkonzepten im Fahrzeug hat aufgezeigt, dass der Trend zu integrierten Fahrerinformationssystemen klar ersichtlich ist. Die relativ starren Strukturen dieser Systeme machen die Nutzung eines vielfältigen Daten- und Informationsangebotes nur sehr bedingt möglich. Des Weiteren hat dieses Kapitel Leitlinien zur Auslegung und Bewertung von Nutzerschnittstellen aufgezeigt. Die folgenden Kapiteln liefern nun einzelne Lösungsbausteine zur Schaffung eines umfassenden Informationsmanagements im Fahrzeug.



## Kapitel 3

# Multimodales, kontextadaptives Dialogmanagement

Dieses Kapitel liefert zunächst die erforderlichen Grundlagen zu Dialogsystemen sowie einen Überblick etablierter Dialogformen. Der Aufbau von Menü- und Sprachdialogen sind dabei für das Interface im Fahrzeug von wesentlicher Bedeutung. Zur Wahl geeigneter Ein- und Ausgabekanäle für den Fahrer werden wissenschaftliche Erkenntnisse und Methoden multimodaler Interfaces aufgezeigt. Um die kontextgerechte Auslegung der Fahrerinformationssysteme zu ermöglichen, wird der Einsatz adaptiver Nutzerschnittstellen diskutiert.

Darüber hinaus werden etablierte Methoden zur Dialogmodellierung erläutert. Zur Schaffung eines umfassenden Informationsmanagements liefert ein multimodales, kontextadaptives Dialogframework den Grundstock zur Realisierung der identifizierten Dialogstrategien. Um gerade dem Trend großer Datenmengen im Fahrzeug Rechnung zu tragen, wird das Dialogframework außerdem um ein Datenmanagement-System erweitert.

### 3.1 Dialogsysteme, Dialogführung

Im alltäglichen Sprachgebrauch wird mit *Dialog* (altgriech. *dialégomai* – sich unterhalten) meistens die Konversation zwischen zwei oder mehreren Personen verstanden. Der Dialog im technischen Sinn ist der wechselseitige Austausch von Informationen zur Erfüllung einer Aufgabe zwischen Mensch und Maschine. Ein *Dialogsystem* ist dabei ein Computersystem, das den Nutzer mittels eines geeigneten Interfaces schrittweise zur Erfüllung seiner Aufgabe führt. Gegenwärtig erfolgt die Dialogführung meist über eine graphische Nutzerschnittstelle (engl. Graphical User Interface (GUI)) mittels taktiler Eingabegeräte. Entsprechend den Anforderungen und Fähigkeiten des Nutzers sollten geeignete Ein- und Ausgabekanäle verwendet und kombiniert werden (siehe dazu Abschnitt 3.2). Gerade im Fahrzeugbereich gewinnen daher die Sprache und weitere Modalitätenkombinationen zur Entlastung des Fahrers zunehmend an Bedeutung.

#### 3.1.1 Dialogformen

Dialogformen, auch Dialogarten oder -strategien genannt, bieten dem Nutzer verschiedene Möglichkeiten mit technischen System zu kommunizieren. Zunächst werden die allgemein geläufigen Formen des Dialogs beschrieben [Gei90, Shn98, ISO06, SR08].

**Frage-Antwort-Dialog:** Bei dieser Form stellt das System Fragen, auf die der Nutzer antwortet.

**Formular-Dialog:** Ähnlich den Formularen aus dem alltäglichen Bereich, wird hier meist ein Formular mit Text und Eingabefeldern elektronisch dargestellt. Die Bearbeitungsreihenfolge spielt hier keine Rolle.

**Menü-Dialog:** Bei einem Menü-Dialog werden dem Benutzer die Optionen in Form einer Liste angeboten, aus der er auswählen kann. Meist erfolgt die hierarchische Untergliederung in mehrere Untermenüs. Auf die Eigenschaften von Menüs wird im Abschnitt 3.1.2 speziell eingegangen.

**Kommando-Dialog:** Hier formuliert der Nutzer seine Eingaben in einer definierten Sprache, die durch Grammatik und Vokabular definiert ist.

**Natürlich-sprachlicher Dialog:** Dies ist der Informationsaustausch in einer Sprache aus dem Bereich der zwischenmenschlichen Kommunikation. Auf Sprachdialoge wird im Abschnitt 3.1.3 eingegangen.

**Direkte Manipulation:** Die direkte Manipulation basiert auf einer graphischen Repräsentation mittels optischer Anzeige. Durch ein Eingabegerät, meist Zeigegerät, kann der Nutzer Objekte anwählen und Aktionen auslösen.

**Multimedia-Dialog:** Hierunter versteht man interaktive digitale Videosysteme und deren Einsatz in Verbindung mit anderen Medien wie Sprachein- und Sprachausgabe als Multimedia-Dialogsysteme.

Die vorgestellten Dialogformen unterscheiden sich aufgrund der benutzten Eingabegeräte und der erforderlichen Eigenschaften der Ausgabegeräte.

### Interaktionsstile

Für die Ausführung von Dialogen gibt es verschiedene Interaktionsstile, die bestimmen in welcher Reihenfolge die Interaktionsschritte zu Dialogschritten zusammengestellt werden. Hierzu lassen sich u. a. folgende Arten unterscheiden:

- *Funktionsorientierte Interaktion:*  
Hier erfolgt zunächst die Auswahl der Operation (Funktion) und danach die Anwendung auf den Operanden (Objekt) (z. B. in einem Malprogramm: Pinsel auswählen, dann malen).
- *Objektorientierte Interaktion:*  
Bei dieser Art wird erst der Operand (Objekt) ausgewählt und anschließend die Operation (Funktion) (z. B. in einem Graphikprogramm: Kreis auswählen und dann Farbe zuweisen). Dieser objektorientierte Interaktionsstil wird in den GUIs der Kapiteln 4 und 5 aufgegriffen.

Des Weiteren zeigen Erkenntnisse der Software-Ergonomie, dass es wichtig ist, zwischen *sequentiellen* (hier ist die einzugebende Information sachlich festgelegt) und *parallelen bzw. simultanen* Aufgaben (keine sachlich bedingte Reihenfolge) zu unterscheiden [Ras04, Bub08b]. Bei der Gestaltung eines nutzergerechten Dialogverlaufs ist es wichtig, die entsprechenden Anforderungen der jeweiligen Bedienungsaufgaben zu berücksichtigen.

### 3.1.2 Menüdialoge

Gerade in der Unterhaltungselektronik, aber auch im Fahrzeugbereich halten menübasierte Dialogsysteme Einzug. Bisher waren diese Systeme meist rein textbasiert, doch heute sind die Dialoge oftmals in einer aufwendigen graphischen Nutzerschnittstelle eingebettet. Diese Menüsysteme sind meist hierarchisch aufgebaut, so dass man ausgehend von der Hauptkategorie durch die Unterkategorien navigieren muss, um zu den gewünschten Funktionen zu gelangen [Nor91]. Dabei können folgende Arten von Menüs unterschieden werden:

**Einzelnes Menü(-fenster):** Dies ist die einfachste Form. Hierbei wird dem Anwender eine Auswahlliste zur Verfügung gestellt.

**Sequentielle, lineare Menüs:** Dabei handelt es sich um eine Aneinanderreihung von einzelnen Menüs, welche *sequentiell* erscheinen, und somit die Bearbeitungsreihenfolge vorgeben. Führen sie den Nutzer durch einen Dialog so spricht man auch von sogenannten *Wizard-Systemen*.

**Simultane Menüs:** Bei dieser Form von Menü wird, im Gegensatz zum sequentiellen Menü, die Bearbeitungsreihenfolge nicht vorgegeben, sondern es werden alle möglichen Dialoge auf einer Seite präsentiert, welche dann in *beliebiger Reihenfolge* durchgeführt werden können.

**Hierarchische Menüs:** Der Grundgedanke ist hier, dass der Anwender seine Auswahl schrittweise mit Hilfe von Kategorien und Unterkategorien verfeinert. Ein hierarchisches Menü kann auch als Baumstruktur bezeichnet werden, wobei die Struktur entweder *symmetrisch* oder, wie in den meisten praktischen Anwendungen, *unsymmetrisch* ist. Ausgehend von den Wurzelknoten, können über die Eltern- und Kindknoten die Blätter erreicht werden (siehe auch Abbildung 3.1(a)).

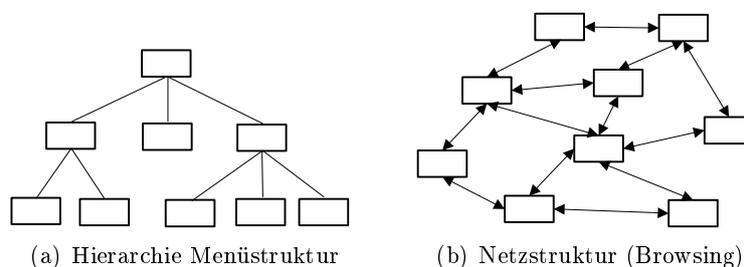


Abbildung 3.1: Menüdialoge

Bei hierarchischen Menüs ist die Breite und die Tiefe entscheidend. Als Menübreite  $b$  wird die Anzahl der auswählbaren Optionen in einer Menüebene bezeichnet. Die Menütiefe  $t$  ist die Anzahl der Menüebenen, um an einen Endpunkt bzw. an ein Blatt zu gelangen. Bei unsymmetrischen Menüstrukturen variieren diese Werte natürlich. Über folgende Funktionen lassen sich die Auswahlzeit

$$T = k + c \cdot \log b \quad (3.1)$$

wobei  $k$  und  $c$  empirisch ermittelte Konstanten sind und die Menütiefe

$$D = \log bN \quad (3.2)$$

berechnen [Shn98]. Als grober Richtwert wird nach [Nor91] eine Spanne von drei bis zwölf Auswahloptionen empfohlen: Nach [Mil81] ist auch die *7+/-2 Regel* bekannt, also fünf bis neun

*Chunks*<sup>1</sup> können im Kurzzeitgedächtnis gespeichert werden [May92]. Für eine optimale Menütiefe werden zwei bis drei Ebenen vorgeschlagen. In vielen Studien wurde gezeigt, dass breite Menüs bei Desktop Anwendungen tiefen Menüs vorzuziehen sind [GST06]. Dies wird damit begründet, dass der Anwender in tiefen Menüs Orientierungsprobleme bekommen kann, und dass eine falsche Entscheidung zu Beginn des Dialogs erst später bemerkt wird und so Rücksprünge nötig sind. Im Gegenzug nehmen bei tiefen Menüs zwar die Such- und Reaktionszeiten je Menüebene ab, über alle Ebenen hinweg führt dies doch zu etwa ähnlichen Bedienzeiten wie bei breiten Menüs. Die Aussage, dass breite Menüs grundsätzlich besser sind, lässt sich aber nicht generell auf FIS übertragen. Es wurde gezeigt [TRK03], dass in einer „Dual Task“ Situation, also während der Fahrt, tiefe Menüs hinsichtlich Bediengeschwindigkeit und visueller Beanspruchung besser abschneiden. Auch auf mobilen Geräten werden, wegen der begrenzten Bildschirmgröße, tiefe Menüs von den Nutzern bevorzugt [GST06].

**Netzförmige Menüstrukturen:** Dies sind Erweiterungen der hierarchischen Baumstruktur, bei denen Sprünge zwischen verschiedenen Menüpunkten, welche nicht den selben Elternknoten haben, möglich sind (siehe Abbildung 3.1(b)). Somit hat der Nutzer die Möglichkeit auf verschiedenen Wegen zum gewünschten Ziel (sprich Endknoten) zu gelangen, was die Flexibilität und die Bediengeschwindigkeit erhöht. Es besteht jedoch auch die Gefahr, dass der Anwender Orientierungsprobleme bekommt, und sich in der Fülle der Querverweise verirrt [May92]. Diese Dialogformen werden auch als *Hypertext*-Systeme bezeichnet und stellen die Grundlage des Internet dar. Im Folgenden wird daher auch von *Browsing* gesprochen, da der Nutzer die Möglichkeit hat, das Umfeld auf verschiedenen Wegen zu explorieren. Dieser Browsing-Ansatz dient v. a. als Grundlage des entwickelten Interfaces im Kapitel 5. In [ANR04] wurde gezeigt, dass netzförmige Dialogformen zu Effizienzsteigerungen bei der Bedienung im Fahrzeug führen können. Dabei werden kontextadaptiver, personalisierter und paralleler Informationszugriff unterschieden.

Die Vorteile von menübasierten Interfaces können laut [May92] in einer geringen kognitiven Beanspruchung (recognition vs. recall), wenigen Eingaben zum Treffen einer Auswahl, einfachen Fehlerbehandlung und Selbstbeschreibungsfähigkeit der Menüs liegen. Dagegen können jedoch auch abhängig von der Auslegung auch folgende Nachteile auftreten: Bei großer Anzahl von Funktionen sind die Menüs ineffizient, unflexibel in der Struktur und erfordern großen Platzbedarf bei vielen Optionen am Bildschirm.

### Navigations- und Orientierungshilfen

Systemkomponenten, die den Benutzer bei der Menü-Navigation unterstützen, um zu einer bestimmten Information zu gelangen, nennt man Navigationshilfen. Orientierungshilfen geben dem Benutzer Unterstützung, wo er sich gerade im Menü befindet. Beispiele hierfür sind Verweise und Landmarks, History Lists, Breadcrumbs bzw. Footprints, Bookmarks, Backtracking und Index. Weitere Infos dazu finden sich in [Abl03].

#### 3.1.3 Sprachdialoge

Zur Entlastung des visuellen Kanals nimmt die Bedeutung von Sprachdialogen im Fahrzeug stark zu. Allgemein sind *Command-and-Control* und *Frage-Antwort* Systeme am häufigsten verbreitet. Bei Command-and-Control Systemen setzt sich die Befehlssyntax oftmals wie folgt zusammen:

<sup>1</sup>Chunks sind einzelne Informationseinheiten wie z. B. wie einzelne Wörter oder auch Menüoptionen

<Befehl> (+ <Parameter>) bzw.  
<Objekt> (+ <Action> + <Parameter>)

Bei Frage-Antwort Systemen hingegen kann der Nutzer neben einzelnen Wörter und Phrasen auch teilweise mit vollständigen Sätzen antworten. Ein mögliche Rückmeldung des Systems wird mit *Prompting* umschrieben. Systeme die sich an die Form Mensch-Mensch-Dialoge annähern, werden als *natürlich-sprachliche* Systeme bezeichnet. Die technischen Anforderungen an natürlich-sprachliche Systeme sind jedoch sehr hoch.

Eine weitere Möglichkeit Dialoge zu klassifizieren ist die Dialoginitiative:

- *Systeminitiiert*:  
das System behält während des gesamten Dialogs die Initiative
- *Nutzerinitiiert*:  
der Nutzer startet den Dialog
- *Wechselseitig initiiert (mixed-initiativ)*:  
die Initiative wechselt während des Dialogs zwischen System und Nutzer

Abhängig von der Anwendung und dem Dialogsystem ergeben sich hier verschiedene Einschränkungen. Ein ideales Dialogsystem bietet wechselseitig initiierte Dialoge an, um bei der Interaktion größtmöglichen Freiraum bereitzustellen.

Auf Basis der Dialogführung können sprachbasierte Dialogsysteme in verschiedene Gruppen eingeteilt werden. [McT04] klassifiziert diese Systeme als *Finite-State-Based*, *Frame-Based* oder *Agent-Based* Systeme.

**Finite-State-Based:** In *Finite-State-Based* (oder auch Graph-Based) Systemen wird der Nutzer durch eine vorher exakt festgelegte Abfolge von Zuständen geleitet. In jedem Zustand wird eine weitere Information abgefragt. Die Anzahl an Informationen, die an das System übergeben werden, ist also gleich der Anzahl der Dialogschritte. Abhängig vom aktuellen Zustand des Systems und der Eingabe eines Benutzers geht das System entsprechend eines hinterlegten, meist starren Ablaufdiagramms oder eines Zustandsautomaten in einen neuen Zustand über. Das System ist meist sehr restriktiv und lässt in Abhängigkeit des Zustands nur bestimmte Eingaben zu. Oftmals ist im Allgemeinen keine Möglichkeit gegeben, eine Abfrage zu überspringen, wenn diese beispielsweise nicht beantwortet werden kann. Finite-State-Based Systeme eignen sich vor allem zur Darstellung und Lösung kleinerer Probleme. Wenig verzweigte Menüstrukturen und eine geringe Anzahl an Auswahlmöglichkeiten sind hier dienlich. Andernfalls werden solche Systeme sehr schnell überaus komplex. Im Vorhinein muss dann nicht nur festgelegt werden, welche Informationen abgefragt werden sollen, sondern auch die Reihenfolge der Abfrage beachtet werden, um eine konsistente und in sich schlüssige Dialogführung zu gewährleisten.

**Frame-Based oder Slot-Filling:** Anders als bei Finite-State-Based Systemen können bei *Frame-Based* Systemen pro Dialogschritt auch mehr als nur eine Information übergeben werden. Es folgen zwei denkbare Dialoge (entnommen aus [McT04]):

System1: What is your destination?

User1: London.

System1: What day do you want to travel?

User1: Friday.  
 System1: What time do you want to depart?  
 User1: At 10 am.  
 System1: I have the following connection ...  
  
 System2: What is your destination?  
 User2: London on Friday around 10 in the morning.  
 System2: I have the following connection ...

Während der Dialog zwischen *User1* und *System1* in seiner Ausprägung eher an ein Finite-State-Based System erinnert, zeigt das zweite Beispiel, dass Frame-Based Anwendungen auch mehrere Informationen auf einmal verarbeiten können – selbst wenn nach diesen nicht explizit gefragt ist. Das auch als Slot-Filling bezeichnete System füllt Slots mit vom Benutzer bereitgestellten Informationen. (siehe Abbildung 3.2) Je nachdem, welche Slots mit Daten versorgt wurden und welche noch nicht ausgefüllt sind, bietet das System dem Nutzer mögliche Informationen aus einer Datenbank oder fordert vom Nutzer noch benötigte Daten an.

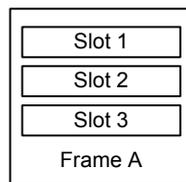


Abbildung 3.2: Graphische Repräsentation von Frame-Based Dialogen

Ein Vorteil ist die Möglichkeit der wesentlich flexibleren Gestaltung von Anfragen und Ausgaben. Der menschliche Dialogpartner ist nicht in einem starren Ablaufplan „gefangen“, sondern kann nach eigenem Ermessen mehr oder weniger Daten in einem einzelnen Bedienschritt übermitteln (siehe auch [AR07]). Besonders sprachbasierte Anwendungen profitieren von der Möglichkeit, dass bei einer falschen Interpretation der Spracheingabe seitens des Systems oder wenn der Nutzer eine Angabe korrigieren möchte, nicht der gesamte Dialog von Neuem durchlaufen werden muss:

System3: Do you want to travel to Sweden  
 on Friday around 10 in the morning?  
 User3: No. I want to travel to London.  
 System3: Do you want to travel to London  
 on Friday around 10 in the morning?...

Der Aufwand zur Erstellung von Frame-Based Systemen ist aufgrund der flexibleren Dialogführung gegenüber zustandsbasierten Varianten erheblich höher. Im Falle von natürlich-sprachlichen Dialogen müssen jedoch die relevanten Informationen erst extrahiert und interpretiert werden. Zur Gestaltung effizienter und gebündelter Dialoge findet dieser Ansatz Anwendung bei den Informationsagenten in Kapitel 6.

**Agent-Based:** Wird ein zu führender Mensch-Maschine-Dialog auf einem wesentlich komplexeren Hintergrund durchgeführt als in den beiden vorangegangenen Abschnitten, so müssen intelligentere Verfahren für das Dialogmanagement eingesetzt werden. Es existieren verschiedene Arten von *Agent-Based* Systemen, abhängig vom Grad der gewünschten Intelligenz und vielen weiteren Aspekten. Intelligente Agenten repräsentieren einen komplexen Themenbereich und werden deshalb in Kapitel 6 ausführlich diskutiert. Kurz zusammengefasst können in einem solchen (Dialog-)System mehrere Agenten nebeneinander existieren; jeder mit einem eigenen Aufgabenbereich und zu verfolgenden Zielen. Eine Besonderheit liegt nun darin, dass die Agenten in Abhängigkeit des Kontextes handeln, d. h. sie können mit Hilfe aktuell vorhandener Daten die Intentionen des Nutzers interpretieren. Stellt der Dialogmanager eine Anfrage, kann der Nutzer, statt diese zu beantworten, z. B. Hilfe anfordern und ein Hilfe-Agent erschließt aus dem Kontext, welche Hilfe für den User am nützlichsten ist. Die relevanten Kontextgrößen sind in sogenannten Wissensquellen hinterlegt und werden nachfolgend in Abschnitt 3.1.3 erläutert (siehe dazu auch Abschnitt 2.2.8). Eingesetzt werden solche Systeme besonders bei komplexen Problemstellungen und bei Dialogen, deren Verlauf antizipiert werden muss. Dem enormen technischen Aufwand steht die erhebliche Freiheit des Dialogs gegenüber.

### Wissensquellen

Nach [McT04] benötigen Dialogsysteme Kenntnisse aus unterschiedlichen Wissensquellen. Der Umfang und die Ausprägung der Wissensquellen sind vom Einsatz und der Umsetzung des Dialogmanagers abhängig. Im Folgenden werden mögliche Wissensquellen aufgelistet:

- **Dialog History:**  
Dies ist eine Aufzeichnung des bisherigen Dialogverlaufs mit allen getätigten Ein- und Ausgaben.
- **Task Record:**  
Dabei handelt es sich um eine Repräsentation der zu sammelnden Informationen. Es wird gespeichert, welche Daten bereits vorhanden sind und welche noch benötigt werden. Dadurch ist es möglich, nur einzelne Daten zu ändern, anstatt alle relevanten Informationen wiederholen zu müssen.
- **World Knowledge Model (aus engl. Weltwissen):**  
In diesem Modell ist allgemeines Hintergrundwissen über menschliche Auffassungen zu finden, das den Dialogmanager befähigt, Angaben eines Nutzers richtig zu verstehen. So sollte das System „am Mittag“ korrekterweise als „um 12:00 Uhr“ interpretieren können oder wissen, dass der 03. Oktober in Deutschland ein Feiertag ist.
- **Domain Model:**  
Besonderes Wissen bezüglich der Domäne, in welcher das System zum Einsatz kommt, ist hier untergebracht.
- **Generic Model of Conversational Competence:**  
Allgemeines Wissen über den Ablauf und den Gehalt einer Konversation gehören zu diesem Modell. Dazu gehört auch die Bereitstellung einer bestimmten Information als Folge einer Anfrage oder die Angabe eines Grundes, weshalb dazu keine Auskunft gegeben werden kann.

- **User Model (aus engl. Nutzermodell):**

Dieses Modell enthält zum Einen Wissen und Informationen über den Nutzer, die kurzfristig gesehen weitgehend konstant bleiben, wie etwa Präferenzen, Alter oder Geschlecht. Aber auch die sich im Dialogverlauf verändernden Informationen über die Ziele, Annahmen und Absichten (goals, beliefs, intentions) des Nutzers gehören dazu (siehe auch Abschnitt 3.1.3).

### 3.2 Multimodale Interaktion

*Multimodalität* bezeichnet in dieser Disziplin die Interaktionsform zwischen Menschen und Computern, bei denen mehrere Modalitäten, also Kommunikationskanäle, verwendet werden. Diese teilen sich generell auf die fünf menschlichen Sinne – *sehen, hören, fühlen, schmecken* und *riechen* – auf, wobei nicht jeder Kanal genutzt werden bzw. bidirektional ausgelegt sein muss [Sei02]. Abbildung 3.3 zeigt die allgemein bekannten fünf menschlichen Sinne mit ihren theoretisch möglichen Aufnahmekapazitäten. Für die Interaktion eignen sich vor allem der *visuelle, auditive, haptische* und *taktile* Kanal. Nach [Ovi99] wird ein multimodales System folgendermaßen definiert:

„Multimodal systems process combined natural input modes – such as speech, pen, touch, manual gestures, gaze and head and body movements – in a coordinated manner with multimedia system output.“ [Ovi99]

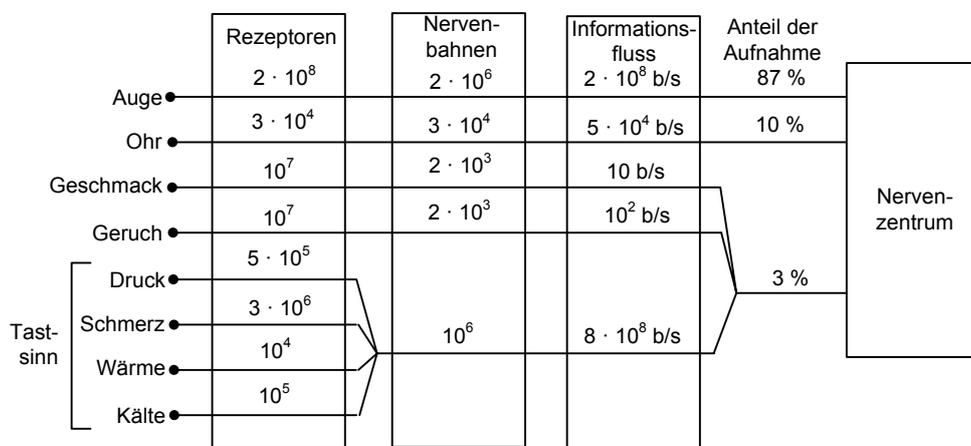


Abbildung 3.3: Verarbeitungskapazitäten der Sinne [SR08]

Als Eingabe können also beispielsweise Sprache, Gesten, Touchscreens, aber auch Tastatur und Maus zum Einsatz kommen und als Ausgabe werden neben grafischen Displays häufig auch Sprachausgabe, Signaltöne (z. B. Earcons<sup>2</sup>) oder haptisches Feedback genutzt. Generelles Ziel ist es, die Interaktion zwischen Mensch und Maschine möglichst den Gewohnheiten zwischenmenschlicher Kommunikation anzunähern und Informationen auf die verschiedenen Sinnesmodalitäten so aufzuteilen, dass die menschliche Informationsverarbeitungskapazität optimal ausgenutzt werden kann [Ovi02].

<sup>2</sup>dieser Neologismus ist an den Begriff Icon angelehnt und bezeichnet ein kurzes akustisches Signal, das einfach interpretiert werden kann

Grundlage des Ansatzes ist das bereits erläuterte Modell des Arbeitsgedächtnisses (siehe Abschnitt 2.3.2), welches funktionell trennbare Komponenten postuliert. Sollte nun eine Komponente überfüllt sein, so könnten theoretisch andere – noch nicht ausgelastete – Komponenten genutzt werden, um parallel mehr Informationen zu verarbeiten. Durch die gleichzeitige Inanspruchnahme verschiedener Komponenten des Arbeitsgedächtnisses kann somit die Verarbeitungskapazität optimiert werden. Dafür ist es aber notwendig, die entsprechenden Informationen über eine andere Sinnesmodalität zu vermitteln. Die multimodale Interaktion sieht ein großes Potenzial für neue Mensch-Maschine-Schnittstellen, die sich an die Arbeitsbelastung des Benutzers anpassen können. Neben der schon erwähnten bestmöglichen Kapazitätsausnutzung lässt sich durch multimodale Interaktion auch die Reaktionsfähigkeit von Benutzern erheblich erhöhen und eine umfassendere Verarbeitungstiefe der Informationen erreichen.

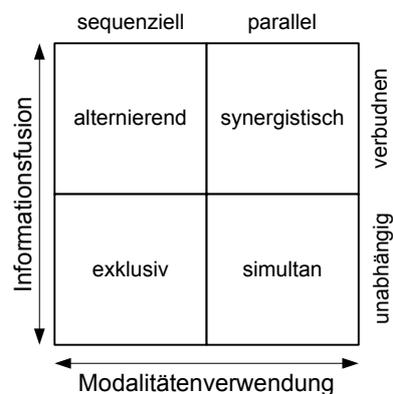


Abbildung 3.4: Klassifikation multimodaler Systeme nach [NC93]

Als Voraussetzung für die Entwicklung von multimodalen Mensch-Maschine Schnittstellen ist die computergestützte Erfassung der natürlichen menschlichen Ausdrucksmittel zu nennen. Hierzu werden Erkennertechnologien für Sprache, Gesten, Mimik, Blickrichtung und Hautleitfähigkeit benötigt. Die Idee, mehrere Kommunikationskanäle zu kombinieren, hat Bolt bereits 1980 formuliert und umgesetzt [Bol80]. Seine mit „Put-That-There“ titulierte Anwendung erlaubt die Manipulation von Objekten auf einer Leinwandprojektion durch verbundene Sprach- und Zeigegesten-Befehle. Dazu wird eine Software zur automatischen Spracherkennung und ein elektromagnetisches Ortungsgerät zur Erfassung der Zeigerichtung benutzt. Abhängig von der Kombination der Kanäle lassen sich multimodale Systeme in *exklusive*, *konkurrierende*, *alternierende* und *synergistische* Multimodalität klassifizieren (siehe Abbildung 3.4) [NC93, HMS96, Alt04, Nie02].

### 3.2.1 Vorteile multimodaler Kommunikation

Multimodale Schnittstellen sind weitgehend durch das Ziel, Interaktionen *transparenter*, *flexibler*, *effektiver*, *effizienter* und *robuster* zu gestalten, inspiriert [MS99, Ovi99]. Der flexible Einsatz von verschiedenen Eingabemodalitäten ist eine wichtige Designfrage. Dazu gehört insbesondere die Wahl der geeigneten Modalität für die jeweiligen Arten von Informationen, die Nutzung von kombinierten Eingabemodi, oder die alternative Nutzung zwischen den Modi. Eingabemodalitäten können abhängig von Kontext und der Aufgabe durch den Benutzer gewählt werden. Insbesondere für komplexe Aufgaben und Umgebungen erlauben es multimodale

Systeme dem Benutzer effektiver zu interagieren. Da es große individuelle Unterschiede in den Fähigkeiten und Präferenzen der Nutzer gibt, ist es von wesentlicher Bedeutung für die Auswahl und die Kontrolle die unterschiedlichsten Benutzergruppen zu unterstützen [OCL04]. Aus diesem Grund wird erwartet, dass multimodale Interfaces einfacher zu lernen und zu verwenden sind. Die sich ständig ändernden Anforderungen mobiler Anwendungen, wie im Fahrzeugbereich, ermöglichen es dem Anwender, die Modalitäten problemlos zu wechseln. Viele Studien zeigen, dass multimodale Interfaces die Nutzererwartungen besser zufriedenstellen. Der größte Vorteil ist vermutlich der Effizienzgewinn, der aus der menschlichen Fähigkeit, Eingaben parallel zu verarbeiten, resultiert. Die multimodale Interface-Gestaltung ermöglicht zudem eine überlegene Handhabung von Fehlern, diese zu vermeiden und zu korrigieren. Folglich können MMS in einer robusten und stabilen Art und Weise funktionieren [Rea88, Rea90, MALR04]. Ein zukünftiges Ziel ist die Interpretation von kontinuierlichen Eingaben aus visuellen, auditorischen und taktilen Kanälen zur Unterstützung einer intelligenten Anpassung an Nutzer, Aufgabe und Umwelt.

### 3.2.2 Ebenen multimodaler Fusion

In diesem Abschnitt wird auf die unterschiedlichen Methoden multimodaler Fusion eingegangen. Die multimodale Fusion lässt sich in *frühe*, *späte* und *hybride Fusion* untergliedern [OCW<sup>+</sup>00, WOC99].

**Feature Ebene bzw. Frühe Fusion:** Die frühe Fusion stellt die niedrigste Fusionsebene dar. Dabei werden zwei Arten zur Fusion der Eingangssignale unterschieden [Sei02, SAM<sup>+</sup>06]. Die am häufigsten angewandte Methode extrahiert zunächst die Merkmale der einzelnen Signale. Diese werden anschließend fusioniert und an den Entscheider weitergeleitet. Dadurch werden die unimodalen Merkmale getrennt extrahiert und zu einer gemeinsamen Repräsentation gebündelt. [ISM06] Die Merkmalsfusion der Modalitäten beeinflussen dabei den gesamten Erkennungsprozess [OCW<sup>+</sup>00]. Eine weitere Möglichkeit stellt die direkte Fusion der jeweiligen Sensorrohdaten dar. Die Merkmale des Signals werden zuerst fusioniert und erst danach extrahiert und an den Entscheider weitergegeben.

Die frühe Fusion basiert oftmals auf Hidden Markov Models (HMM) oder temporalen Neuronal Networks (NN) [OCW<sup>+</sup>00, Alt04]. Die Merkmalsfusion wird in der Regel bei eng verbundenen und zeitsynchronen Modalitäten, wie z. B. Sprache und Lippenbewegungen, verwendet. Sie ist jedoch nicht für Modalitäten mit unterschiedlicher zeitlicher Skalierung oder wesentlich abweichendem Informationsinhalt geeignet. Ein Beispiel hierfür ist die Kombination von Sprach- und Stifteingabe. Beide Modalitäten liefern unterschiedliche, komplementäre Daten und können erst später fusioniert werden.

**Interpretationsebene bzw. späte (semantische) Fusion:** Bei der späten Fusion findet die Klassifikation für jede einzelne Modalität getrennt statt. Durch Auswertung der einzelnen Ergebnisse der Klassifikatoren wird die Kombination bzw. Fusion der Modalitäten erreicht. So kann dieses Konzept als eine Kombination von Klassifikatoren, mit dem Ziel der Verbesserung des Gesamterkennungsergebnisses, gesehen werden. Bei der späten Fusion werden die Eingangssignale auf der semantischen Ebene integriert [Sei02, Alt04]. Ein solches System besteht aus mehreren unimodalen Erkennern, die sich, z. B. durch Post-Processing von Merkmalen auch gegenseitig beeinflussen können. Jedes System besitzt zudem eine eigene Vorverarbeitung, Merkmalsextraktion und Entscheidungsebene [SAM<sup>+</sup>06]. Eine übergeordnete Komponente vereint die einzelnen Interpretationen anschließend zu einer einzigen gesamten

Interpretation.

Diese semantische Fusion wird meist bei Modalitäten angewandt, welche sich in den zeitlichen Skalierungscharakteristiken ihrer Merkmale unterscheiden. Diese Kombination der Daten auf rein semantischer Ebene bewirkt eine hohe Flexibilität und Wiederverwendbarkeit der einzelnen Komponenten, da diese nicht a-priori aufeinander abgestimmt werden müssen. So können die Modalitäten einzeln trainiert und später ohne erneutes Training integriert werden. Ein Nachteil der späten Fusion ist der potentielle Verlust von Korrelation im fusionierten Merkmalsraum [SWS05].

**Hybride Mehrebenenfusion:** In [Vo98] wird eine weitere Art der Fusion diskutiert. Bei dieser Art der Fusion wird den verschiedenen Modalitäten erlaubt, miteinander auf dem Decodier-Level zu interagieren, so dass vorhergehende Ergebnisse das Sprachmodell beeinflussen und die Erkennungsleistung gesteigert wird. [Vo98] nennt hier als Beispiel ein System, bei dem Eingaben per Sprach- und Stifterkennung möglich sind. Dort kann man zum Beispiel, wenn man weiß, was der Benutzer sagt, die erwarteten Eingaben per Stift-Gesten eingrenzen und so die Zuverlässigkeit des Systems stärken.

### 3.2.3 Wichtige Integrations- bzw. Fusionsmethoden

Nachfolgend werden wichtige Integrations- bzw. Fusionsmethoden aufgeführt. [SAM<sup>+</sup>06]

**Temporale Integration:** Die temporale Integration stellt meist den ersten Schritt eines kombinierten Integrationsprozesses dar. Mit ihr kann bestimmt werden, ob ein Signal alleine oder im Zusammenhang mit anderen Signalen interpretiert werden sollte. Sie kann anschließend z. B. mit statistischen, regelbasierten oder auch semantischen Ansätzen kombiniert werden. Man unterscheidet zwei Arten von temporaler Integration: Die *mikrotemporale* Integration, bei der Informationen aus den Erkennern parallel oder pseudo-parallel verarbeitet werden, und die *makrotemporale* Integration, bei der die Informationen seriell, also nacheinander, verarbeitet werden. Hierbei überlappen die Informationen nicht, gehören aber zum gleichen Zeitfenster. [SAM<sup>+</sup>06, CMBM03]

**Statistische Integration:** Bei der statistischen Integration produziert jeder Erkenner eine N-Best-Liste, welche jedem erkannten Ergebnis eine Wahrscheinlichkeit zuordnet. Der Integrator berechnet nun zu jeder möglichen Kombination aus den verschiedenen N-Best-Listen eine Wahrscheinlichkeit. Die Auswahl einer gültigen Kombination kann durch semantische Auswertung oder mit Hilfe einer Datenbank, welche alle gültigen Kombinationen enthält, erfolgen. [SAM<sup>+</sup>06, CMBM03]

**Regelbasierte Integration:** Ein weiterer Fusionstyp ist die regelbasierte Integration. Dabei werden die möglichen Kombinationen der verschiedenen Eingänge durch ein Regelwerk oder Look-Up-Tabellen angegeben. Die regelbasierte Fusion ist der temporalen Integration sehr ähnlich, jedoch nicht strikt an Timing-Bedingungen gebunden. Ein solches System lässt sich sehr gut an gegebene Rahmenbedingungen anpassen, ist jedoch später nur schwer erweiterbar. Die Komplexität des Systems steigt mit der Anzahl der verwendeten Regeln. [SAM<sup>+</sup>06, Alt04]

**Semantische Integration:** Die semantische Integration arbeitet, ähnlich der statistischen Integration, mit N-Best-Listen als zugeführte Daten aus den Eingabe-Erkennern. Diese werden anhand Ihrer Wahrscheinlichkeit sortiert und dem semantischen Integrator zugeführt, welcher dann entscheidet, ob er aus den verschiedenen Kombinationen der eingehenden N-Best-Listen einen gültigen Befehl erzeugen kann. Aufgrund dieses einfachen Aufbaus sind semantische

Integratoren leicht zu erweitern, jedoch nimmt die Komplexität bei steigender Anzahl von Systemeingängen stark zu. Um dies zu verhindern, kann man dem System noch einen temporalen oder einen statistischen Integrator hinzufügen, welcher Eingaben behandelt, die vom semantischen Integrator nicht kombiniert werden konnten. [SAM<sup>+</sup>06, Alt04, CMBM03]

**Unification-Based Integration:** Allgemein ist Unification-Based Integration eine Methode, bedeutungstragende Datenfragmente aus verschiedenen Eingabemodalitäten in eine gemeinsame Bedeutung zu integrieren [Ovi03]. Die Vereinigung ist eine Operation, bei der die Konsistenz zweier Teilinformationen geprüft und bei positivem Befund zu einem einzelnen Ergebnis kombiniert wird. Diese Methode ist beispielsweise ideal, um die Kompatibilität einer Gesteneingabe und einer Spracheingabe zu prüfen und gegebenenfalls zu kombinieren, um dem System eine Interpretation zu ermöglichen [JCM<sup>+</sup>97, OCW<sup>+</sup>00, SAM<sup>+</sup>06]. Die Vereinigung von Merkmalsstrukturen ermöglicht die Kombination von komplementären und redundanten, jedoch nicht von widersprüchlichen Eingaben.

**Hybride Verarbeitung:** Um die Vorteile der jeweiligen Fusionsmethoden optimal auszunutzen, sind verschiedene hybride Verfahren entwickelt worden. So resultiert ein System, welches statistische Verarbeitungstechniken mit einem Unification-Based Verfahren kombiniert, in einem hybriden symbolisch-statistischen System. Die wichtigsten Techniken sind Associative Mapping und Members-Teams-Committee (MTC); Sie entwickeln und optimieren die Zuordnungsstruktur zwischen multimodalen Befehlen und ihre jeweiligen Bestandteile, und die Art und Weise der Verbindung von posterioren Wahrscheinlichkeiten mit statistischen Ansätzen. [OCW<sup>+</sup>00, Alt04]

### 3.2.4 Multimodalität im Automobil

Die Fahrzeugführung erfordert eine möglichst ununterbrochene visuelle Aufmerksamkeit und beansprucht die Hände und Füße zur Steuerung des Fahrzeugs. Gleichzeitig sind während des Fahrbetriebs oftmals weitere Aufgaben aus dem sekundären und tertiären Bereichen zu erfüllen, wie etwa die Bedienung eines Radios oder der Klimaanlage. In der Regel müssen auch diese mit den Händen und mit kurzen Kontrollblicken bedient werden. Das kann zu Konflikten führen, weil die Leistungsfähigkeit einzelner Ressourcen begrenzt ist [ALK01]. Aus dieser Betrachtung wird klar, dass eine effizientere Bearbeitung multipler Aufgaben nur durch die Steigerung der Timesharingfähigkeit oder durch die Verwendung unabhängiger Ressourcen ermöglicht wird.

Dies bestätigt auch die Theorie der multiplen Ressourcen von Wickens [Wic84], der zufolge der Mensch verschiedene unabhängige Ressourcen besitzt, mit denen er diese Effizienzsteigerung erreichen kann (siehe Abschnitt 2.3.2). Für den Informationsaustausch unterscheidet Wickens zwischen sensorischen und effektorischen Ressourcen, die wiederum nach Modalitäten aufgeteilt sind. Demnach besteht eine intermodale Unabhängigkeit, durch die simultanes Sprechen und Gestikulieren oder auch Hören und Sehen ohne Leistungseinbußen möglich ist [VT03]. Ebenso ist eine gewisse Unabhängigkeit zwischen der sensorischen und der effektorischen Verarbeitung zu beobachten. So wird z. B. die Sehleistung nicht durch zeitgleiches Sprechen vermindert. Dies gilt vor allem dann, wenn keine mentalen Ressourcen gemeinsam genutzt werden. Es ist demnach theoretisch möglich, die Arbeitslast durch ein multimodales Eingabesystem zu senken und dadurch die Bedienungseffizienz zu steigern [Abl07]. Mit der multimodalen Bedienung im Fahrzeug befassen sich zahlreiche Untersuchungen wie [Ben01, Alt04, Gei03, SD01, Nie02]. Sie haben gezeigt, dass der Einsatz von multimodalen

Benutzerschnittstellen zu geringeren Beanspruchungen, besserer Leistung bei komplexen Informationen, geringerer Fehlerrate und höherer Reaktionsfähigkeit als bei unimodalen Benutzerschnittstellen führen. In dieser Arbeit wird daher für die entwickelten Anzeige- und Bedienkonzepte *alternierende* und *synergistische* Multimodalität in Form von später Fusion eingesetzt (siehe Abbildung 3.4). Vor allem Sprache in Kombination mit visueller Nutzerführung scheint eine viel versprechende Lösungsmöglichkeit für die sichere Bedienung im Fahrzeug zu sein. Bei der alternierenden Verwendung der Modalitäten sollte zu jedem Zeitpunkt ein problemloser Wechsel zwischen den Modalitäten gewährt werden.

### 3.3 Kontextadaptive Dialogsysteme

Eine besondere Herausforderung an das Design von grafischen Benutzerschnittstellen im Fahrzeug stellt die Diversität der möglichen Benutzergruppen dar. Einerseits können die potentiellen Nutzer der Schnittstellen gemäß der Erfahrung mit der Benutzerschnittstelle, beziehungsweise der Versiertheit mit Schnittstellen in Anfänger/Erstnutzer, Sachkundige/sporadischen Nutzer und Experten/Vielnutzer eingeteilt werden [SP04]. Darüber hinaus unterscheiden sich die Benutzer hinsichtlich ihrer Interessen und der von ihnen bevorzugt ausgeführten Funktionen. Manche Fahrer lassen sich bei jeder Fahrt vom Navigationssystem leiten, andere nutzen diese Funktion nur selten, etwa bei Urlaubsreisen. Bestimmte Gruppen wünschen sich die persönliche Musiksammlung im Fahrzeug und verwalten eine Vielzahl von MP3 Dateien, während andere Nutzer wiederum auf den Lieblingsradiosender setzen. Benutzerschnittstellen im Auto müssen universell für alle Benutzergruppen einsetzbar sein. Bei einer zunehmenden Anzahl an Funktionen führt dies zu zunehmend unübersichtlichen Menüstrukturen. Abhilfe versprechen hier adaptive Benutzerschnittstellen [LF98], die sich basierend auf den Interaktionen des Nutzers automatisch an dessen Bedürfnisse anpassen. Dieser Abschnitt gibt eine Einführung in adaptive Benutzerschnittstellen, sowie eine Übersicht publizierter Anwendungen im Fahrzeugumfeld.

#### 3.3.1 Einordnung adaptiver Benutzerschnittstellen

Adaption bezeichnet generell die Fähigkeit eines Objektes sich an externe Einflussfaktoren anzupassen. Abbildung 3.5 klassifiziert Benutzerschnittstellen hinsichtlich dieser Fähigkeit nach Löffler [Loe89].

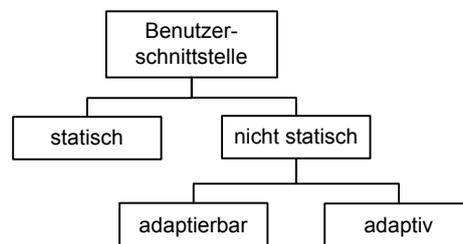


Abbildung 3.5: Klassifizierung von Benutzerschnittstellen hinsichtlich der Adaptivität [Loe89]

Bei statischen Benutzerschnittstellen sind Art und Weise der Interaktion, sowie die Funktionalität und das Verhalten der Schnittstelle festgelegt und unveränderlich. Adaptive Schnittstellen haben die Fähigkeiten, sich in einem bestimmten Rahmen automatisch an Nutzerverhalten

und Kontextinformationen anzupassen. Durch diese automatische Anpassung des Systems ohne Einwirkung des Nutzers grenzen sich adaptive von adaptierbaren Benutzerschnittstellen ab. Ist ein System adaptierbar, so können Einstellungen des Systems durch externe Eingriffe, beispielsweise durch den Nutzer selbst, angepasst werden [PR05].

Die Zielstellung von adaptiven Benutzerschnittstellen ist eine flexible Anpassung an Benutzer, Aufgabe und Kontext zur Optimierung des Informationsaustausches zwischen Benutzer und System [Loe89, DMKSH93].

### Modellbasierte Anpassung von Benutzerschnittstellen

Die Grundlage für die Adaptivität von Benutzerschnittstellen bilden Modelle für die Einflussfaktoren der Anpassung. Von besonderer Bedeutung für adaptive Benutzerschnittstellen sind Benutzer-, Kontext- und Aufgabenmodelle:

- **Benutzermodell:**

Das Benutzermodell enthält die gesamten Informationen, die ein System über den Nutzer besitzt. In der Regel startet das System mit einem Modell mit vorgegebenen Werten oder erfragt die initialen Werte explizit bei der ersten Nutzung. Während der Interaktion des Nutzers mit dem System werden bestimmte Aspekte, wie beispielsweise genutzte Funktionen oder Suchbegriffe, vom System im Benutzermodell gespeichert. Klassische Elemente eines Benutzermodells sind daher [Kul00]:

- Vorlieben, Interessen oder Ziele eines Nutzers
- Kenntnisse des Nutzers
- Interaktionshistorie (genutzte Funktionen, erledigte Aufgaben, Anfragen bei der Hilfefunktion, usw.)

- **Kontextmodell:**

In einem Kontextmodell wird der augenblickliche Systemkontext festgehalten. Ein Kontextmodell gewinnt insbesondere bei Anwendungen auf mobilen Geräten oder auch in Fahrzeugen an Bedeutung. Im Kontextmodell können Informationen wie die Geschwindigkeit oder der Ort gespeichert werden.

- **Aufgabenmodell:**

Aufgabenmodelle repräsentieren mögliche Folgen von Benutzeraktionen. Der Abgleich dieser mit dem Nutzerverhalten erlaubt es dem adaptiven System, Hypothesen über die Ziele des Nutzers oder mögliche Probleme aufzustellen.

### 3.3.2 Aktionsvorhersage und Antizipation

Ausgehend von den Modellen des Nutzers, des Kontextes und möglicher Aufgaben können Annahmen über die augenblickliche Aufgabe und die Ziele eines Nutzers abgeleitet werden. Ausgehend von diesem Wissen kann sich eine Benutzerschnittstelle anpassen und den Nutzer bei seiner Aufgabe optimal unterstützen. Beispielsweise werden an dieser Stelle Aktionsvorhersage-Algorithmen (AVA) vorgestellt, die auf Basis einer Nutzungsstatistik potentielle Folgebefehle bestimmen.

Kunzer et al. stellen AVAs basierend auf vergangenen Handlungssequenzen vor. Beispielhaft wird kurz die Idee des IPAM (Incremental Probabilistic Action Modeling) Algorithmus nach Davison et al. [DH98] vorgestellt.

Der IPAM Algorithmus wurde für die Vorhersage von Kommandobefehlen für eine Unix-Kommandozeile entwickelt und basiert auf einem einfachen Markov Modell: Es wird lediglich die Information des vorhergehenden Befehls für die Vorhersage des nächsten Kommandos genutzt. Für jeden Befehl wird eine Wahrscheinlichkeitsverteilung für Folgebefehle verwaltet. Bei der Eingabe eines Befehls wird die Wahrscheinlichkeitsverteilung des vorangehenden Befehls angepasst, wobei kürzlich aufgetretene Befehle höher gewichtet werden als länger zurück liegende. Für die Vorhersage des nächsten Befehls wird der Befehl mit der höchsten Wahrscheinlichkeit in der Wahrscheinlichkeitsverteilung für das Folgekommando des zuletzt ausgeführten Befehls gewählt. Weitere Ansätze sind PPM (Prediction by Partial Matching) und seine Erweiterungen, welche ebenfalls auf Markovmodellen basieren sowie der von Künzer et al. vorgestellte KO-Algorithmus, welcher mehrere bekannte Handlungsabfolgen zur Vorhersage der nächsten Aktion berücksichtigt.

### Arten von Adaptivität

Eine Benutzerschnittstelle kann auf verschiedenen Betrachtungsebenen adaptiv gestaltet werden. Beispielsweise ist denkbar, dass sich die Visualisierung in Abhängigkeit von der Tageszeit oder die Menüstruktur in Abhängigkeit der gefahrenen Geschwindigkeit ändert. Dieterich et al. [DMKSH93] schlägt folgende Liste von adaptierbaren Bestandteilen einer Benutzerschnittstelle vor:

- **Generische Funktionen:**
  - Fehlerkorrektur
  - Aktive Hilfe
- **Interaktionsebenen:**
  - Form der Eingabe von Information durch den Nutzer
  - Form der Systemausgabe von Informationen
  - Zugriff auf bestimmte Funktionen
  - Vereinfachung von bestimmten Aufgaben (durch benutzergenerierte Makros)

Im Rahmen der vorliegenden Arbeit wurden mögliche Arten von Adaptivität grafischer Benutzerschnittstellen im Automobilkontext konkretisiert und die folgenden möglichen Eigenschaften die adaptiv gestaltet werden können identifiziert:

- **Visualisierung:**

Bei einer adaptiven Visualisierung wird die Darstellung von Menükomponenten im User Interface abhängig von Kontextinformation gestaltet. Dies umfasst sowohl die Anordnung von Elementen des Menüs, als auch Farben und Formen von Menükomponenten. Ein Beispiel für adaptive Visualisierung ist die Anpassung der Farbgebung an die Helligkeit oder Tageszeit.
- **Bedienlogik:**

Wird die Bedienlogik adaptiv gestaltet, so betrifft dies die Navigation durch ein User Interface. Denkbare Variablen sind hier die Menüstruktur oder Shortcuts, die in Abhängigkeit vom Nutzerverhalten oder anderen Kontextinformationen eingefügt werden können.

- **Funktionsumfang:**

Die Menge der angebotenen Funktionen kann in Abhängigkeit vom Systemkontext variieren. Denkbar ist hier, dass gewisse Funktionen, die hohe Aufmerksamkeit des Nutzers erfordern, bei hoher Geschwindigkeit und Abwesenheit eines Beifahrers nicht angeboten werden.

- **Datenumfang:**

Bei den dargestellten Daten kann ein User Interface in zwei Dimensionen adaptiv gestaltet werden:

- Auswahl der Daten
- Sortierung der Daten

Mit Auswahl ist hier gemeint, dass bestimmte Daten gar nicht angezeigt werden. Der Fall, dass beispielsweise die fünf am häufigsten genutzten Daten angezeigt werden sollen, fällt unter den Punkt Sortierung.

### 3.3.3 Adaptivität im Automobilkontext

Wie zu Beginn erwähnt, ist der Einsatz von adaptiven Benutzerschnittstellen im Automobilkontext sehr interessant, da eine große Bandbreite an möglichen Benutzern mit verschiedenen Interessen und Fertigkeiten angesprochen wird. Der folgende Abschnitt gibt eine kurze Übersicht über implementierte adaptive Benutzerschnittstellen für Automobile bei Daimler Chrysler Research [RFT00]:

- *Adaptive Route Advisor:*

Der Adaptive Route Advisor [RFL99] gibt dem Fahrer Routenvorschläge basierend auf seinem bisherigen Fahrverhalten und seinen Präferenzen. Der Fahrer bekommt für ein angegebenes Fahrziel einen Vorschlag für verschiedene Routen zur Auswahl. Die einzelnen Routen werden über bestimmte Parameter wie Gesamtzeit, Anzahl von Kurven oder Autobahnkilometer bewertet. Die bisherige Auswahl des Fahrers sowie bisher bekannte Strecken werden bei den Vorschlägen des Systems berücksichtigt. Demnach passt das System die zur Auswahl stehenden Daten und deren Sortierung an den Nutzer an.

- *Adaptive News Reader:*

Der Adaptive News Reader [BP99] bezieht Nachrichten von einem Internetserver und liest diese dem Fahrer vor. Über einen Touchscreen und ein Sprachinterface hat der Nutzer die Möglichkeit, dem System Feedback über sein Interesse an den Daten zu geben. Anhand dieses Feedbacks werden in einem Nutzermodell Hypothesen über die Vorlieben eines Fahrers verwaltet und auf Basis dieses Modells die Auswahl der vorzulesenden Nachrichten getroffen.

### Studien zum Einsatz adaptiver Benutzerschnittstellen

In der Literatur findet sich eine Vielzahl von Studien zur Erforschung der Effizienz von adaptiven Benutzerschnittstellen [GCTW06, MS89, SS94, MBB02, FM04, GCH<sup>+</sup>05, GW85]. Die Studien beschäftigen sich durchwegs mit der Effizienz bei der Auswahl von bestimmten Funktionen aus Menüstrukturen in grafischen Benutzerschnittstellen für Desktop-Anwendungen. In [GCTW06] wird eine kurze Übersicht über den Inhalt der einzelnen Studien gegeben.

Als Zusammenfassung der genannten Studien bleibt festzuhalten: haben sich die der Untersuchung zugrunde liegenden Menüstrukturen sehr stark geändert oder war die Änderung vom Nutzer nicht nachvollziehbar, so hat sich die Adaptivität meist negativ auf die Effizienz bei der Bedienung ausgewirkt. Eine durchwegs positive Bewertung haben adaptive Menüstrukturen erhalten, die oft genutzte Funktionen dupliziert an einer exponierten Stelle im Menü angeboten haben.

Die Erkenntnis, dass eine zu starke, oder unerwartete Änderung der Benutzerschnittstelle den Nutzer von seiner eigentlichen Aufgabe ablenkt, deckt sich mit der Theorie des *stages of action* Modells von Norman [Nor02]. Angewandt auf Benutzerschnittstellen wird in diesem Modell die Diskrepanz zwischen der Informationsrepräsentation der Benutzerschnittstelle und des mentalen Modells des Nutzers als Kluft der Auswertung (*gulf of evaluation*) bezeichnet. Wird diese Kluft zu groß, muss sich der Nutzer neu orientieren und vermehrte Aufmerksamkeit auf die Interpretierung der Ausgaben der Benutzerschnittstelle aufwenden, was ihn von der effizienten Bearbeitung seiner eigentlichen Aufgabe abhält.

## 3.4 Dialogmodellierung

Zur Modellierung von Dialogen gibt es verschiedene etablierte Methoden, die in dieser Arbeit Anwendung finden. Sehr bewährt hat sich hierbei die Verwendung von *Zustandsautomaten* (auch *endliche Automaten* genannt). Dazu existieren verschiedene graphische Repräsentationsarten und formale Beschreibungsformen, die in diesem Abschnitt näher erläutert werden.

### 3.4.1 Zustandsautomaten

Ein Zustandsautomat besteht aus *Zuständen*  $s_i$  und *Transitionen*  $t_{ij}$ , die einen Quell- und einen Zielzustand verbinden [HMU01].



Abbildung 3.6: Elemente eines Zustandsautomaten

**Zustände:** In einem Zustand wird eine bestimmte unveränderliche Bedingung modelliert. Will man eine explizite Bedingung formulieren, kann man sie als Einschränkung dem Zustand zuordnen (s. u.). Dabei können dem Zustand drei Verhaltensspezifikationen zugeordnet werden:

- *Eintrittsverhalten:*  
dieses Verhalten wird ausgeführt, wenn der Zustandsautomat in den Zustand eintritt
- *Aktivitätsverhalten:*  
während sich der Zustandsautomat im Zustand befindet, wird diese Bedingung ausgeführt
- *Ausgangsverhalten:*  
dieses Verhalten tritt beim Verlassen des Zustandsautomaten auf

Ein Zustand wird meist als Kreis dargestellt (siehe Abbildung 3.6(a)).

**Transitionen:** Ein Quell- und ein Zielknoten wird über eine Transition verbunden (siehe Abbildung 3.6(b)). Die Transition beschreibt ein Verhalten, das beim Durchlaufen ausgeführt wird. Ein Wächterausdruck kann die Transition schützen. Die Transition wird nur dann durchlaufen, wenn der Wächterausdruck wahr ist.

Zur Modellierung können komplexe Zustandsautomaten durch weitere Elemente und Methoden erweitert werden (siehe Abschnitt 3.4.2).

### Mathematisches Modell

Formal lässt sich ein Zustandsautomat  $Z$  als 5-Tupel formulieren [SR08]:

$$Z = \{S, X, T, s_0, F\} \quad (3.3)$$

wobei

- $S$  eine endliche nicht leere Menge von Zuständen  $s_i$ ,
- $X$  das Eingabealphabet aller  $x$  (eine endliche nicht leere Menge von Symbolen),
- $T$  die Menge der Zustandsübergangsfunktionen für die Zustände in  $S$ ,
- $s_0$  der Anfangszustand und ein Element aus  $S$ ,
- $F$  die Menge von Endzuständen und eine (möglicherweise leere) Untermenge von  $S$

ist.

Ist die Ausgabefunktion eine Funktion von Zustand und Eingabealphabet, dann handelt es sich um ein Mealy-Modell. Falls die Ausgabefunktion nur vom Zustand abhängt, dann ist es ein Moore-Automat.

Einfache Zustandsautomaten lassen sich in Form von Übergangstabellen (auch Look-Up-Table genannt) formulieren:

MOMENTANER ZUSTAND/ BEDINGUNG	Zustand $s_1$	Zustand $s_2$	Zustand $s_3$
Bedingung $T_1$	...	...	...
Bedingung $T_2$	...	<i>Zustand <math>s_3</math></i>	...
Bedingung $T_3$	...	...	...

Tabelle 3.1: Übergangstabelle

### 3.4.2 Graphische Repräsentation

Neben der Darstellung von Zuständen und Transitionen (siehe Abbildung 3.6) gibt es weitere etablierte graphische Repräsentationen.

#### Statecharts

David Harel hat die so genannten *Statecharts* als Darstellungsform für Endlichen Automaten eingeführt [Har84, Har87]. Die Notation ermöglicht die exakte Spezifikation von zustandsbasierten Systemen. Dabei hat David Harel u. a. folgende Notationselemente eingeführt, um die Komplexität großer Systeme handhabbar zu machen (siehe Abbildung 3.7):

- *Entry-, Exit-, Throughout-Aktionen:*  
sie geben an, welche Aktionen beim Eintreten, Verlassen bzw. während des Aufenthalts in einem Zustand ausgeführt werden.
- *Condition-Konnektor:*  
abhängig von einer Bedingung wird ein Zustandsübergang (Transition) in verschiedene Zielzustände überführt. Als graphisches Symbol für den Condition-Konnektor wird ein eingekreistes *C* verwendet.
- *Hierarchische Komposition:*  
hier befindet sich in einem Oberzustand durch hierarchische Schachtelung ein weiterer vollständiger Unterzustandsautomat (die Darstellung erfolgt als Rahmen, siehe Abbildung 3.7). Die Einführung von Unterzuständen führt zu einer übersichtlicheren Darstellung mit weniger Kanten.
- *History-Konnektor:*  
beim Wiedereintritt in einen Unterzustandsautomaten wird der zuletzt aktive Unterzustand eingenommen. Ein eingekreistes *H* beschreibt den History-Konnektor.
- *Parallele Komposition:*  
durch die Verwendung von *AND* und *OR*-Kompositionen können ein gleichzeitiges oder abwechselndes Schalten der Automaten vorhergesehen werden.
- *Temporale Logik:*  
dadurch können beispielsweise Timeouts in Transition angegeben werden.

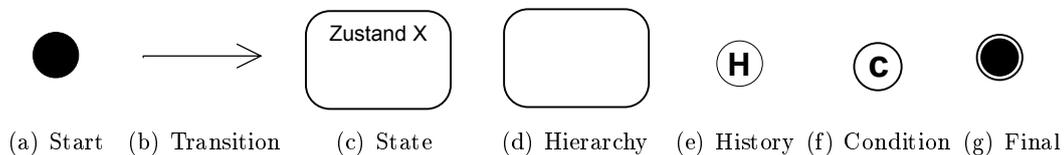


Abbildung 3.7: Erweiterung der Darstellungsform durch Statecharts

Die Darstellungsform der Statecharts ist heute Teil der *Unified Modeling Language*<sup>3</sup> (UML). UML wurde von der Object Management Group (OMG) entwickelt und ist eine standardisierte Sprache zur Beschreibung von Softwaremodellen. UML definiert dabei graphische Notationen für viele Begriffe und legt mögliche Beziehungen zwischen diesen Begriffen fest.

### Ablaufdiagramm

Eine Methode zur graphischen Darstellung von logischen Programmabläufen ist das *Fluss- oder Ablaufdiagramm*. Diese Darstellung findet Anwendung, um die Funktionsweise von Algorithmen oder Computerprogrammen zu veranschaulichen. Darüber hinaus eignet es sich zur Darstellung von verschiedenen Prozessen und Tätigkeiten. Auch die Aufgaben des Fahrers lassen sich als Ablaufdiagramme darstellen. So können gedankliche Abläufe und erforderliche Bedienschritte auf Basis der Zeit aufgeschlüsselt und analysiert werden.

<sup>3</sup>engl. vereinheitlichte Modellierungssprache

### 3.4.3 Formale Beschreibung

Zur formalen Beschreibung von Zustandsautomaten und Dialogen eignen sich folgende Metasprachen.

#### Extensible Markup Language – XML

Die *Extensible Markup Language* (XML) ist eine Metasprache zur textbasierten Darstellung hierarchisch strukturierter Daten. Sie wird zum Austausch von Daten zwischen verschiedenen Computersystemen verwendet. XML eignet sich auch sehr gut um Dialogabläufe zu formulieren und wird daher zur Dialogbeschreibung in dieser Arbeit verwendet. Folgendes Beispiel zeigt eine mögliche Notation eines simplen Dialogs:

```
<State Name='MainMenu'>
  <Transition Name='RadioOn' Destination='RadioMenu'>
    <Intention='RadioOn' Confidence='0.65'>
      Radio einschalten
    <\Intention>
  <\Transition>
<\State>
```

#### Backus-Naur-Form

Zur Darstellung kontextfreier Grammatiken (engl. Context Free Grammar (CFG)) findet die *Backus-Naur-Form* (BNF) Anwendung. Diese kompakte formale Metasprache wird hier z. B. zur Beschreibung von Grammatiken des Spracherkenners verwendet:

```
<Radio> ::= <Radio hören> | <Radio einschalten> | <Radio an>
```

## 3.5 Multimodales, kontextadaptives Dialogframework

In diesem Abschnitt wird ein Dialogframework beschrieben, das die Verarbeitung multimodaler Dialoge und den Einbezug von Kontextkriterien ermöglicht. Das Framework dient als Basis der in den nachfolgenden Kapiteln realisierten Prototypen. Unter einem *Framework* (aus engl. für Rahmenstruktur, Fachwerk) versteht man ein Programmiergerüst, welches in der Softwaretechnik insbesondere im Rahmen der objektorientierten Softwareentwicklung sowie bei komponentenbasierten Entwicklungsansätzen verwendet wird. Ein Framework ist selbst noch kein fertiges Programm, sondern stellt den Rahmen, innerhalb dessen der Programmierer eine Anwendung erstellt, zur Verfügung.

### 3.5.1 Etablierte Architekturmodelle

Zunächst werden bekannte Architekturmodelle wie das *Seeheim-Modell* und das *MVC-Modell* erläutert.

#### Seeheim-Modell

Ein interaktives System lässt sich nach dem Seeheim-Modell als logisches Architekturmodell, wie in Abbildung 3.8 dargestellt, in folgende Komponenten unterteilen [Pfa83, Nie02]:

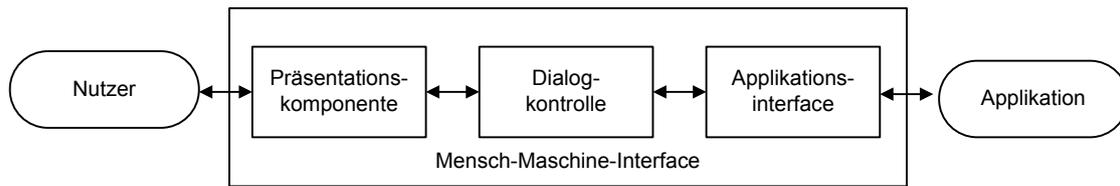


Abbildung 3.8: Überblick Seeheim-Modell

- Die *Präsentationskomponente* dient der Ein- und Ausgabe von Informationen. Ihre Aufgabe besteht in der Verwaltung des Bildschirms und der physikalischen Eingabegeräte, der textuellen und grafischen Ausgabe sowie der Realisierung verschiedener Interaktionstechniken [FWC84]. Sie enthält die dem Benutzer sichtbaren und zugänglichen Komponenten einer Benutzungsschnittstelle.
- Die *Dialogkontrolle* definiert und überwacht die korrekte Abfolge von Dialogschritten, die durch Interaktionstechniken der Präsentationskomponente realisiert werden. Sie steuert dadurch das dynamische Verhalten einer Benutzungsschnittstelle. Zu ihren Aufgaben gehört u. a. die Überprüfung der syntaktischen Korrektheit ganzer Folgen von Benutzereingaben, Fehlerbehandlungen und die Umwandlung von Ausgabeoperationen der Anwendungsschnittstelle in mehrere Aussagetoken, die an die Präsentationskomponente weitergeleitet werden [Göt94].
- Die *Anwendungsschnittstelle* beschreibt die Systemfunktionalität aus der Sicht der Benutzungsschnittstelle. Sie definiert die für die Benutzungsschnittstelle zugreifbaren Datenstrukturen und Funktionen der Anwendung sowie die Struktur der Informationen, die zwischen Benutzungsschnittstelle und Anwendung ausgetauscht werden.

### MVC-Modell

Das aus der Smalltalk-Welt stammende *Model-View-Controller* Modell (MVC) bezeichnet eine Architektur, welche die Objekte eines Softwaresystems in drei Bereiche einteilt [Kra88]:

- das *Model* enthält die Objekte des Anwendungskerns
- die *Views* präsentieren die Objekte
- der *Controller* wird verwendet, um Daten einzugeben und um das Model zu Handlungen zu veranlassen.

Durch dieses Modell kann ein flexibles Programmdesign erreicht werden, das u. a. eine spätere Änderung oder Erweiterung erleichtert und die Wiederverwendbarkeit einzelner Komponenten ermöglicht.

### 3.5.2 Multimodale, kontextadaptive Dialogarchitektur

Ein möglicher Aufbau einer multimodalen Architektur wurde bereits in [HMS96] vorgestellt. Mit dem Aufbau multimodaler Systeme im Fahrzeug beschäftigen sich u. a. auch [ALK01] und [MALR04]. In dieser Arbeit wird die in Abbildung 3.9 dargestellte Architektur verwendet. Nachfolgend werden die einzelnen Komponenten der verwendeten Architektur beschrieben.

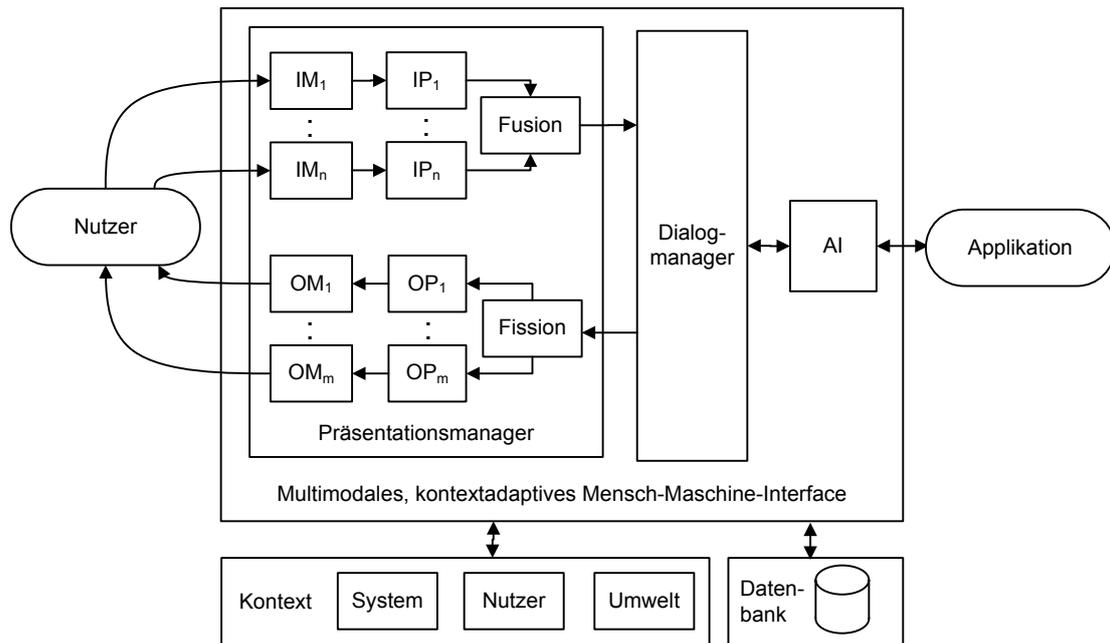


Abbildung 3.9: Architektur eines multimodalen Dialogmanagements

**Input Modality ( $IM_x$ ):** Diese Module dienen zur Verarbeitung der Eingaben. In dieser Arbeit sind dies taktile Bedienelemente wie z. B. Dreh-/Drücksteller oder Touchpad und ein Spracherkenner.

- *Spracherkenner bzw. Automatic Speech Recognition (ASR):*

Als Spracherkenner wurde der Scansoft Vocon 3200 verwendet. Dem Erkenner werden Vokabular in phonembasierter Beschreibung und die definierte Grammatik in BNF Form hinterlegt. Unterschiedliche Grammatiken können zur Laufzeit über den Dialogmanager dynamisch geladen werden. Der Erkenner wird über eine Push-To-Talk-Taste (PTT) oder ein Schlüsselwort initiiert. Die Analyse von gesprochenen Eingaben führt zu Hypothesen linguistischer Einheiten wie Phoneme oder Wörter. Die Hypothesen werden dann in sogenannten N-Best-Lists ausgegeben. Dazu wird meist ein Konfidenzmaß ausgegeben, das die Wahrscheinlichkeit der Hypothesen widerspiegelt.

<Hypothesis='Zielführung starten' Confidence='0.56' NR='1'> <\Hyperthesis>

**Input Processing ( $IP_x$ ):** Aufbereitung der sensorischen Eingabedaten.

- *Sprachinterpretation (engl. Natural Language Understanding (NLU)):*

Hier erfolgt die Interpretation der Hypothesen des Spracherkenners zu möglichen Intentionen. Die Interpretation von natürlich-sprachlichen Äußerungen erfolgt mit dem in [Mül03] vorgestellten Ansatz, der auf der Verarbeitung mittels Bayes'scher Netze beruht. Daraus werden dann mögliche Intentionen generiert.

<Intention='StartNavigation' Confidence='0.76' NR='1'> <\Intention>

Außerdem kann ein Garbage-Modell definiert werden, um nicht relevante Eingaben zu filtern.

**Multimodale Fusion:** Diese Komponente integriert die Ergebnisse der Eingabekanäle. Hierzu gibt es verschiedene Fusionsstrategien, die in Abschnitt 3.2 vorgestellt wurden. In dieser Architektur wurde Late Semantic Fusion (siehe Abschnitt 3.2.2) verwendet. Zur Verarbeitung der Intentionen erfolgt eine temporäre Integration anhand eines in XML definierten Regelwerks bzw. von Look-Up Tables.

**Multimodaler, kontextadaptiver Dialogmanager (engl. *Dialog Manager (DM)*):** Der DM ist die zentrale Komponente des Dialogsystems, das als Vermittler zwischen Nutzer und System fungiert und für die Dialogführung zuständig ist. Die Formulierung der Dialoge basiert auf den in Abschnitt 3.4 vorgestellten Zustandsautomaten bzw. Statecharts. Der Dialogmanager ermöglicht die Realisierung von Frame-Based Dialogen (siehe Abschnitt 3.1.3), die in den Frames beliebige Dialogreihenfolgen zulassen. Ein entsprechender Dialogverlauf ist in Abbildung 3.10 graphisch dargestellt.

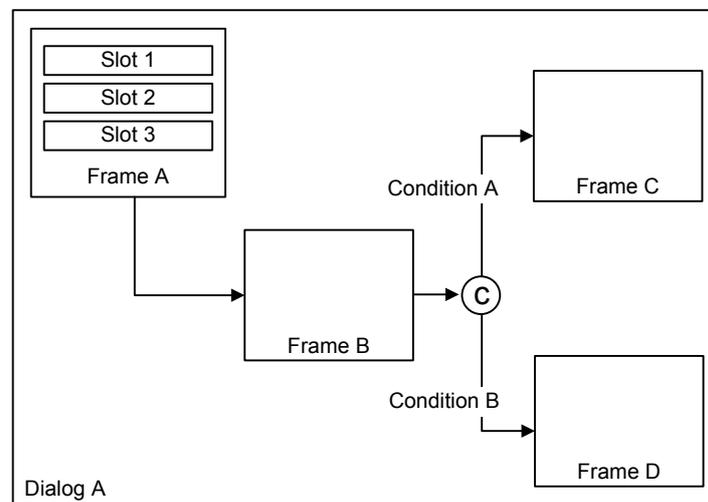


Abbildung 3.10: Frame-Based Dialogmanagement mit Kontextbedingungen

Darüber hinaus können adaptive Dialogverläufe durch *bedingte Transitionen* (z. B. durch Kontextbedingungen), *Verbosität* und *Erfahrungslevel* formuliert werden. Außerdem kann auch die gewünschte Modalität definiert werden, über die Rückmeldung erfolgen soll. Durch Koppelung mit Bedingungen kann die geeignete Modalität entsprechend von Kontextkriterien (z. B. Geschwindigkeit) gewählt werden. Zur XML-Formulierung der multimodalen, kontextadaptiven Dialoge siehe Tabelle 3.2:

**Data Base (DB):** Datenbank zur Speicherung von Daten und Informationen. Die vorgestellte Architektur kann durch das in Abschnitt 3.5.3 vorgestellte Datenmanagement-System erweitert werden.

**Applikation Interface (AI):** Dies ist die Schnittstelle des Dialogsystems zur Anwendung. Die Kommunikation erfolgt über Socket-Kommunikation (s. u.).

**Multimodale Fission:** Darunter versteht man die Aufteilung der Ausgabe auf verschiedene Modalitäten. Es kann z. B. die Ausgabe auf verschiedene Displays im Fahrzeug wie HUD, CID und Kombi sowie auch die Sprachausgabe erfolgen.

Dialogelement	Beispiel
Zustand $S$	<code>&lt;State Name='navigation'&gt; &lt;\State&gt;</code>
Transition $T$	<code>&lt;Transition Name='startNavigation'&gt; &lt;\Transition&gt;</code>
Frame	<code>&lt;Frame Name='address'&gt; &lt;\Frame&gt;</code>
Slot	<code>&lt;Slot Name='streetname'&gt; &lt;\Slot&gt;</code>
Bedingung	<code>&lt;Condition='criticalFuel'&gt; &lt;\Condition&gt;</code>
Modalität	<code>&lt;Modality='TTS'&gt; Route berechnet &lt;\Modality&gt;</code>
Verbosität	<code>&lt;Verbosity='Ort'&gt;&lt;\Verbosity&gt;</code>
Erfahrungslevel	<code>&lt;Level='novice'&gt;</code>
Dynamische Grammatik	<code>&lt;Grammar='streetnames'&gt; &lt;\Grammar&gt;</code>
Datenbank-Anfrage	<code>&lt;Data='Munich' Get='streetname'&gt; &lt;\Data&gt;</code>

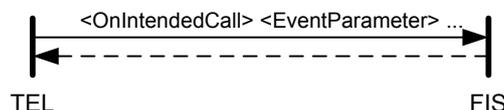
Tabelle 3.2: XML-Dialogformulierung

**Output Processing ( $OP_x$ ):** Hier wird das Ausgabeergebnis aufbereitet. Im Falle der Sprachverarbeitung wird über die Natural Language Generation (NLG) eine Rückmeldung in Form eines natürlich-sprachlichen Satzes formuliert.

**Output Modality ( $OM_x$ ):** Diese Komponente ermöglicht die Ausgabe des Ergebnisses über eine gewählte Modalität wie einem Display oder die Sprachausgabe bzw. Sprachsynthese (auch Text-To-Speech (TTS) genannt). Sie dient zur Erzeugung der akustischen Sprachausgabe.

**Netzwerkkommunikation:** Bei komplexen Architekturen ist es sinnvoll auf netzwerkfähige Lösungen zurückzugreifen. Daher werden einzelne Komponenten der multimodalen Dialogarchitektur durch Module in dem Netzwerk gekapselt. Ähnlich zu [MALR04] werden in dieser Arbeit die Module über eine Server-/Client-Struktur angebinden. Hierzu werden TCP/IP- und UDP-Socketverbindungen zum Nachrichtenaustausch mittels Context Free Grammar (CFG) im Netzwerk verwendet. Der Informationsaustausch findet statt, indem über den Kommunikationskanal Nachrichten über ASCII-Strings im XML-Format versendet werden. Der Vorteil dieser XML-Dateien liegt in der Darstellung von strukturierten Daten. Die Meldungen enthalten einen Adressblock und einen Informationsblock. Im Adressblock stehen der Empfänger und der Absender der Nachricht:

```
Empfänger | Absender | Informationsblock
TEL       FIS       <OnIntendedCall> <EventParameter> ...
```



### 3.5.3 Datenmanagement-System

Motiviert durch Konzepte, die nach komplexeren Datenmodellen verlangen, und die Wiederverwendbarkeit von Komponenten in verschiedenen Prototypen, wird nun ein Datenmanagement-System vorgestellt. Es erweitert die zuvor vorgestellte multimodale, kontextadaptive Dialogarchitektur.

### Architektur

Die Daten mit der Benutzerschnittstelle werden über *Simple Object Access Protocol* (SOAP) basierte *Web Services* ausgetauscht. Abbildung 3.11 zeigt die Systemarchitektur als Aufbaudiagramm in der Modellierungssprache FMC (Fundamental Modeling Concepts, siehe [KGT06]).

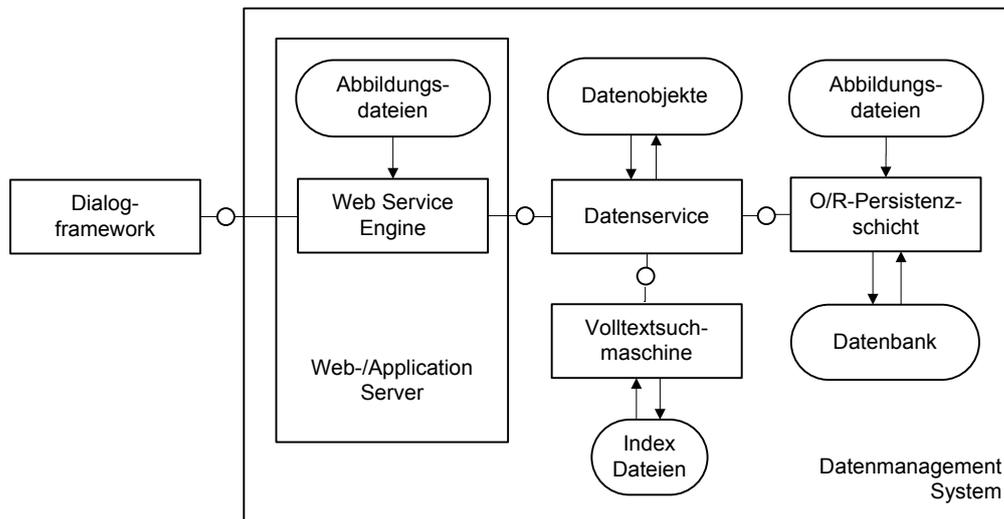


Abbildung 3.11: Aufbaudiagramm des Datenmanagement-Systems

Der Nutzer interagiert lediglich mit der grafischen Benutzerschnittstelle. Diese nutzt für die Datenhaltung das Datenmanagement-System und kommuniziert mit diesem über die SOAP Schnittstelle. Die Kommunikation zwischen Benutzerschnittstelle und Datenmanagement-System erfolgt nach dem Prinzip Request-Response. Das bedeutet, dass die Benutzerschnittstelle Anfragen an das Datenmanagement-System stellt, dieses die Anfragen bearbeitet und anschließend eine Antwort zur Benutzerschnittstelle sendet. Eine Initiierung der Kommunikation durch das Datenmanagement-System ist nicht vorgesehen.

Das Datenmanagement-System regelt den Zugriff auf die Datenbank (MySQL). Es besteht aus den Komponenten für die Bereitstellung des Web Services: Webserver (Apache Tomcat) sowie Web Service Engine (Apache Axis) für den Zugriff auf den Datenservice via SOAP. Die Web Service Engine ist hierbei für die Serialisierung und Deserialisierung der SOAP Nachrichten und für die Aufrufe der entsprechenden Java Methoden zuständig. Die Zuordnung von SOAP Nachrichten auf Java Methoden und Informationen für die Serialisierung der Datenobjekte werden in Abbildungsdateien abgelegt. Der eigentliche Datenservice stellt die Zugriffsmethoden auf die Daten bereit. Da die Daten in der Datenbank relational gespeichert werden, die Daten über den Web Service aber als Datenobjekte im Sinne eines objekt-orientierten Entwurfs übergeben werden, beinhaltet die Architektur eine objektrelationale Persistenzschicht (hibernate), welche die Abbildung von Objekten auf relationale Datenstrukturen vornimmt und den Zugriff auf die Datenbank regelt (*O/R-Persistenzschicht*). Die Abbildung von Objekten auf die Datenbank wird in Abbildungsdateien abgelegt. Des Weiteren interagiert der Datenservice mit einer Volltextsuchmaschine (Apache Lucene), welcher die Funktionalität der Volltextsuche kapselt und hierfür einen Index verwaltet.



## Kapitel 4

# Multimodale, datenübergreifende Fahrzeug-Suchmaschine

Ziel dieses Kapitels ist es, eine leistungsfähige Zugriffsstrategie für das stark steigende Datenaufkommen im Fahrzeug zu schaffen. Von einer ehemals überschaubaren Anzahl gespeicherter Radiosender ist der Fahrer heute mit einer enormen Anzahl an Musik- und Senderlisten, Telefon- und Adressbüchern und vielen weiteren Daten wie z. B. Point-Of-Interests (POI) konfrontiert. Allerdings mangelt es an Methoden zur fahrergerichten Nutzung dieser unzähligen Daten und Informationen.

Im Büro- und am Heimarbeitsplatz ist der PC-Anwender schon seit langem einer Unmenge an Daten und Informationen über rasant wachsende Speicherkapazitäten und globale Rechnernetze, wie Internet oder Firmennetzwerk, ausgesetzt. Klassische Ansätze der rein hierarchischen Strukturierung von Dokumenten in Ordner- und Menüstrukturen sind auch hier an ihrer Grenzen geraten. Ein zunehmend bewährtes Mittel ist die Verwendung von Suchmaschinen, ohne die das Internet von heute undenkbar wäre.

Um auch im Fahrzeug auf die großen Datenmengen zugreifen zu können, werden nun die Anforderungen einer datenübergreifenden Fahrzeug-Suchmaschine erarbeitet, dazu geeignete Interfaces konzipiert und anschließend bewertet.

### 4.1 Suche in verschiedenen Applikationen

Nachfolgend wird zunächst die Funktionsweise von Suchmaschinen in verschiedenen Anwendungsdomänen beschrieben. Eine Suchmaschine ist ein Programm zur Recherche von Dokumenten, die in einem Computer oder einem Computernetzwerk gespeichert werden. Sie hat ihren Ursprung aus dem Bereich des Information Retrievals (IR). Bei unzähligen PC-Anwendungen, wie der Dateiverwaltung, Datenbank-Applikationen und v. a. zur Auffindung von Websites im Internet, zeigen sich deutlich die Stärken der Suche.

#### 4.1.1 Datei- und Datensuche

Die Datei- und Ordnersuche ist ein bewährtes Mittel, um die Festplatte oder weitere Speichermedien eines PCs auf das Vorkommen von bestimmten Strings zu durchsuchen. Beispielsweise lässt sich bei Windows XP zunächst die gewünschte Kategorie wie Bilder, Musik, Videos, Dokumente oder Dateien festlegen. Nach der Eingabe des Textstrings kann man eine Eingren-

zung des Suchortes oder des Änderungsdatums über „erweiterte Suchfunktionen“ vornehmen. Jetzt kann die Suche gestartet werden und die Ergebnisse werden in Abhängigkeit des Suchfortschritts nach und nach aufgelistet.

Bei Windows Vista wurde die Suche stark überarbeitet, sodass Suchergebnisse nun sofort angezeigt werden, im Gegensatz zur Windows XP Suche, bei der mehrere Minuten bis zur Anzeige von Suchergebnissen vergehen können. Dies ist möglich, da die Vista-Suche auf einer erweiterten und verbesserten Version des Indizierungsdienstes basiert. Außerdem ist es möglich die Suche durch das nachträgliche Hinzufügen von Filtern stetig zu verfeinern. Zum schnellen Auffinden und Start von Anwendungen ist direkt im Startmenü ein Suchfeld integriert, das ebenfalls das Auffinden von Dokumenten und E-Mails ermöglicht (siehe Abbildung 4.1(a)).

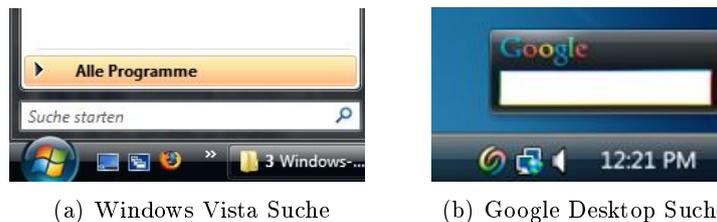


Abbildung 4.1: Datei- und Datensuche

Spezifische Suchinterfaces stellen PC-Applikationen wie z. B. Microsoft Outlook zur Verwaltung von E-Mails und iTunes zur Musikwiedergabe bereit. Eine übergreifende Suche bietet *Google Desktop Search* an (siehe Abbildung 4.1(b)). Neben der Internetsuche indiziert es lokale Dateiinhalte und auch zusätzlich relevante Programmdateien wie E-Mails, aufgerufene Webseiten, usw. auf dem Rechner. Über eine zentrale Eingabeleiste kann die datenübergreifende Suche initiiert werden. Eine Treffervorschau erscheint unmittelbar bei der Eingabe von Buchstaben, durch Bestätigung mit Enter wird die gesamte Trefferliste sichtbar.

#### 4.1.2 Internetsuche

Zum Auffinden von Websites und Netzinhalten haben Suchmaschinen die größte Bedeutung gewonnen. 2006 verwendeten in Deutschland ca. 86 % der Internetbenutzer Suchmaschinen. Mittlererweile ist die Firma Google wertvollstes Medienunternehmen der Welt geworden. Insgesamt lassen sich folgende Arten von Internet-Suchmaschinen unterscheiden:

**Allgemeine Suchmaschinen:** Sie durchsuchen automatisch Adressbereiche im Internet und erstellen daraus einen durchsuchbaren Index. Diese Systeme werden auch *Spider*, *Crawler* oder *Robot* genannt. Die gefundenen Dokumente werden auf Inhalt und enthaltene Links analysiert. Die Trefferliste einer Suchanfrage wird nach Relevanz sortiert ausgegeben. Hierbei haben vor allem *Ranking-Algorithmen* wie PageRank stark an Bedeutung gewonnen (siehe hierzu Abschnitt 4.3.5). Neben Google zählen u. a. Altavista, Lycos und Fireball zu den bekanntesten Vertretern dieser spiderbasierten Art.

**Katalogbasierte Suchmaschinen:** Dies sind hierarchisch gegliederte Linksammlungen, die nach definierten Kategorien untergliedert sind. Die einzelnen Kategorien werden redaktionell erstellt. Neben einer Suchanfrage können hier die Inhalte auch über die entsprechenden Kategorien durchforstet werden. Bekanntester Vertreter der katalogbasierten Suchmaschinen ist Yahoo!.



Abbildung 4.2: Internetaktionshaus ebay

**Meta Suchmaschinen:** Sie verfügen über keinen eigenen Datenbestand, sondern nutzen den Datenbestand von angeschlossenen Suchmaschinen. Über die Nutzung von Meta-Suchmaschinen ist es möglich, mehrere Anfragen parallel zu stellen. Vertreter hierfür sind z. B. Apollo7 und Alphasuche.

**Spezielle Suchmaschinen:** Insgesamt gibt es sehr viele lokale Suchmaschinen, die länder- bzw. regionalspezifisch ausgelegt sind. Als Weiteres existieren Fachdatenbanken, Nachschlagewerke, Telefonbücher, Wörterbücher, Lexika, Karten und Bibliothekskataloge. Auch das größte Internetauktionshaus *ebay* lebt von einer maßgeschneiderten Suchfunktion und geeigneter Ergebnisaufbereitung. Hierbei kommen *hybride Suchsysteme* zum Einsatz. Die Suche lässt sich stets durch weitere Stichwörter als auch durch vorgeschlagene Kategorien weiter einschränken. Auch bei *Amazon*, der größte Online Buchhändler, können mittlerweile Artikel jeglicher Art erworben werden. Die Suche ist ebenfalls hybrid aufgebaut und ermöglicht neben der direkte Eingabe von Schlagwörtern auch die hierarchische Navigation in Kategorien. Eine Besonderheit stellt die Empfehlung von Artikeln dar, die auf einer Analyse des Kaufverhaltens basieren (siehe auch Abschnitt 6.3.4). Dies ist ein bewährtes Mittel, um den Kunden bei der Produktsuche weiter zu unterstützen.

### 4.1.3 Suche auf mobilen Endgeräten

Die großen Internet-Suchdienste nutzen den Trend mobiler Endgeräte mit Datendiensten. Die Einschränkungen dieser Geräte, wie kleine Displays und Bedienelemente mit niedrigen Übertragungsraten, müssen hier berücksichtigt werden. Suchmaschinen-Anbieter wie Google, Yahoo und MSN bieten daher spezifische Interfaces an, welche die Inhalte deutlich reduzieren. Google hat seinen Suchdienst für mobile Geräte im Jahre 2005 gestartet. Um mit dem Microbrowser zu kommunizieren, wird ein XHTML-Standard verwendet (siehe Abbildung 4.3).

### 4.1.4 Suche durch Exploration

Ein anderer Weg als die Texteingabe zur Informationsauffindung wird in [KRR<sup>+</sup>06] beschrieben, wo das facettenbasierte Suchsystem FaThumb vorgestellt wird. FaThumb ist konzipiert zur Suche und Exploration großer Datenmengen wie POIs. Diese Methode kann auch als Browsing (siehe Abschnitt 3.1.2 und Kapitel 5) aufgefasst werden. Die Grundidee dieses Sys-

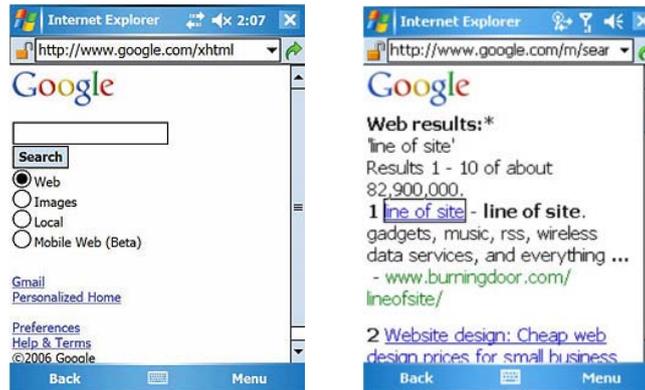
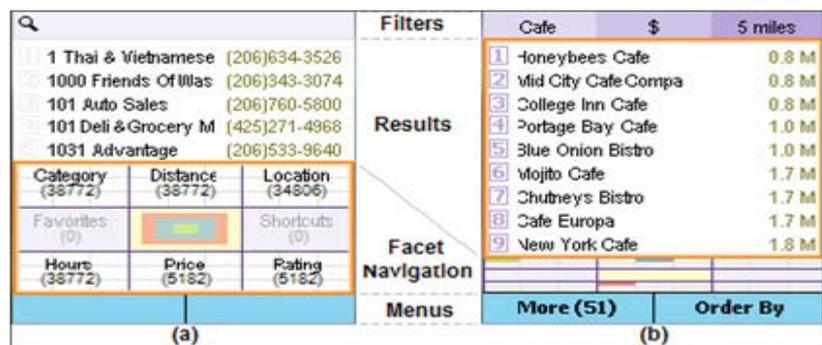


Abbildung 4.3: Suche Google mobile

tems ist es, den Suchraum sukzessiv immer weiter einzuzugrenzen, um zum gesuchten Objekt zu gelangen. Funktionsweise und Interface werden im Folgenden kurz vorgestellt.

**Facettenbasierte Suche FaThumb:** Das Interface von FaThumb ist in vier verschiedene Zonen aufgeteilt (siehe Abbildung 4.4). In der obersten Zeile werden die momentan aktiven Filter angezeigt, darunter die Ergebnisse der Suche. Über das Facettennavigationsmenü in Form einer 3x3 Matrix werden Filter zur Verfeinerung der Suche angezeigt. Die Filter können über den räumlich korrespondierenden Ziffernblock des Mobiltelefons angewählt werden. Am unteren Bildschirmrand werden die Belegungen der zwei Softkeys angezeigt. Mit dem linken Softkey kann der Nutzer dann Filter dauerhaft auswählen, mit dem Rechten kann er wieder ins Hauptmenü zurück navigieren, um dort weitere Filter aus einer anderen Hauptkategorie zu wählen.

Abbildung 4.4: FaThumb Interface [KRR<sup>+</sup>06]

Die Grundfilter sind Kategorie, Entfernung, Ort, Öffnungszeiten, Preis und Bewertung. So muss der Nutzer, um beispielsweise ein günstiges chinesisches Restaurant im Umkreis von 5 km zu finden, dreimal vom Hauptmenü ausgehend durch die jeweiligen Unterkategorien navigieren, um den gewünschten Filter auszuwählen. Hat man die gewünschten Filter eingestellt, so kann man die neun dargestellten Ergebnisse direkt mit den neun Zifferntasten anwählen, oder mit einem Textfilter weiter verfeinern. Die Ergebnisse der Evaluierung zeigten, dass das Browsing über Filter zwar keine bedeutende Zeitersparnis gegenüber reiner Textsuche bringt, dafür aber

vom User deutlich besser im Bereich Joy of Use bewertet wurde.

## 4.2 Anforderungen der Fahrzeug-Suchmaschine

Zunächst werden in diesem Abschnitt die Anforderungen und die Grundidee einer Suchmaschine im Fahrzeug erläutert. Die nachfolgend konzipierten Suchinterfaces berücksichtigen die speziellen Anforderungen im Fahrzeug und bieten zukunftsweisende Lösungen zum Umgang mit großen Datenmengen an. [AMG<sup>+</sup>05]

### 4.2.1 Datenbasen

Daten aus unterschiedlichsten Bereichen und Quellen halten Einzug in das Fahrzeug und wollen dort vom Fahrer genutzt werden. Diese sind:

DATENART	BEISPIELE
<i>Point-Of-Interests</i> (POIs)	Restaurants, Tankstellen, Supermärkte, Hotels, WCs, Sehenswürdigkeiten
<i>Adressen</i>	PC-Adressbücher, Adressbücher des Mobilfunktelefons, Adressbücher des Fahrzeugmenüs (z. B. Navigation, Telefon), Telefonbücher
<i>Musik</i>	Radiosender, Radiostreams, CDs, MP3s
<i>PC-/Office-Anwendungen</i>	E-mails, Kalender, Dokumente
<i>Informations- und Netzdienste</i>	Internet, podcasts, Staumeldungen, Nachrichten, Wetter, Aktienkurse, Sportmeldungen
<i>Funktionen der Fahrzeugmenüs</i>	Entertainment, Navigation, Kommunikation, Klima, Einstellungen

Tabelle 4.1: Datenbasen

Eine Unterscheidung zwischen *lokalen* und *globalen* Daten ist möglich. Lokale Daten liegen direkt im Fahrzeug, während globale Daten auf mobilen Endgeräten oder auf entfernten Datenservern bzw. im Internet liegen können. Zum Abrufen globaler Daten bedarf es funkbasierter Datendienste im Fahrzeug wie beispielsweise GPRS, EDGE oder UMTS. Mobile Endgeräte können sowohl kabelgebunden als auch kabellos über Infrarot, Bluetooth oder WLAN angebunden werden. Dabei soll dem Fahrer ein einfacher Zugriff auf diese vielschichtigen Daten ermöglicht werden.

### 4.2.2 Suchanfrage

Des Weiteren werden die Anforderungen an das Formulieren einer Suchanfrage erläutert. Das Formulieren erfolgt über alphanumerische Eingabestrategien. Geeignete Eingabegeräte und -methoden werden gesondert in Abschnitt 4.4 genauer erläutert. Der Suchstring wird in eine überall sichtbare, zentrale Suchleiste eingegeben. Dieser kann sich aus einem oder mehreren Teilen zusammensetzen. Folgende Methoden vereinfachen und erleichtern die Suche:

- *Personalisierte Suche*:  
z. B. letzte Anfragen, gespeicherte Anfragen (Favoritenlisten)

- *Automatische Ergänzungsvorschläge:*  
z. B. „Search as you type“
- *Unschärfe Eingaben:*  
z. B. Suchverfahren wie Levenshtein, Soundex
- *Semantische Interpretation:*  
z. B. Algorithmen wie z. B. TVSM (siehe Abschnitt 4.3.5)
- *Filter-Keywords:*  
z. B. Kategorien, Schlagwörter, örtliche Einschränkung (z. B. Entfernung von 5 km)

Gerade die Suchmaschine im Fahrzeug sollte mit einer gewissen Robustheit bei dem Formulieren der Anfrage aufwarten und auch unscharfe Eingaben interpretieren. Da der Fahrer von seiner primären Aufgabe, dem Steuern des Fahrzeuges, so wenig wie möglich abgelenkt werden darf, müssen eventuelle Schreibfehler durch eine unscharfe Suche abgedeckt werden.

### 4.2.3 Ergebnisaufbereitung

Von großer Bedeutung für den Fahrer ist die angemessene Präsentation und Aufbereitung des Suchergebnisses. Hierzu werden die erforderlichen Darstellungs-, Sortierungs- und Filtermethoden sowie die Berücksichtigung von Kontexteinflüssen erläutert.

Nachfolgend wird lediglich die Ausgabe des Suchergebnisses in Form von vertikalen Listen betrachtet.<sup>1</sup> Überschreitet das Suchergebnis den verfügbaren Anzeigeraum am Bildschirm, ist ein Aufsplitten des Ergebnisses auf mehrere Seiten bzw. durch Scrollen erforderlich. Dazu kommen folgende Möglichkeiten in Betracht:

- *Objektorientierte Ergebnislisten* mit zugeordneten Funktionen
- *Clustering* in Kategorien (z. B. über Karteireiter und Trefferanzeige)
- *Sortierung und Ranking* nach Kontext
- *Filterung* der Daten

Nun werden Methoden erläutert, die das Auffinden in den Ergebnislisten vereinfachen und vernetzte Funktionen anbieten.

**Objektorientierte Ergebnislisten mit vernetzten Funktionen:** Die Ausgabe des Suchergebnisses erfolgt in Form von Objektlisten, wie Adressen, Personen, Musiktiteln, usw. Diese Objekte können weitere Erläuterungen beinhalten und mit ihnen sind zugehörige Funktionen verknüpft bzw. vernetzt. Diese Funktionen sind beispielsweise bei einer Person das Tätigen eines Anrufes, das Hinzufügen als Navigationsziel, das Schreiben einer SMS, usw. Dies ist natürlich abhängig von den hinterlegten Daten und Funktionen der angeschlossenen Geräte.

**Sortierungsmethoden:** Zur Sortierung des Ergebnisses eignen sich verschiedene Methoden. Die Ausgabe ist in folgenden Sortierungen denkbar:

- *Alphabetisch A-Z*
- *Zuletzt Gewählt*

---

<sup>1</sup>Daneben existieren graphische Visualisierungsstrategien, die im Kapitel 7 detaillierter betrachtet werden.

- *Häufigkeit*
- *Entfernung*
- *Ranking* (siehe dazu Abschnitt 4.3.5)

**Filtermethoden:** Die Filterung der Daten stellt eine sehr bedeutende Unterstützung für die Suche dar. Folgende Filtermethoden bieten sich dabei an:

- *Schlüsselwörter:*  
z. B. Ortsangaben, Kategorienamen, unscharfe Einschränkung,...
- *Kategorien:*  
z. B. Funktions- und Datengruppen,...
- *Individualisierung:*  
z. B. selbst definierte Filterprofile,...
- *Kontext:*  
z. B. örtliche Eingrenzung wie im Umfeld von 10 km,...

Die Filterung lässt sich in Abhängigkeit des Ausführungszeitpunktes unterscheiden:

- *vorgelagert*
- *nachgelagert*

Die Kombination von Suchen und Filtern wird hier als *Search and Filter* bezeichnet.

### Kontextberücksichtigung

Gerade bei dem Ranking und bei der Darstellung des Suchergebnisses ist es sinnvoll Kontextkriterien mit einzubeziehen. Dabei sind nach Abschnitt 2.2.8 folgende Kontextgrößen relevant:

- *Nutzerkontext:*  
z. B. Vorlieben, Auswahlhäufigkeit, zuletzt gewählt, Belastung des Fahrers,...
- *Systemkontext:*  
z. B. Eingehender Anruf,...
- *Umwelt- und Fahrzeugkontext:*  
z. B. Aktueller Standort, Entfernung, Ziel, Termine,...

#### 4.2.4 Multimodale Bedienstrategien

Die Interaktion mit einer Suchmaschine im Fahrzeug erfordert die geeignete Kombination von Ein- und Ausgabegeräten (siehe Abschnitt 3.2). Zur visuellen Ausgabe wird ein grafikfähiges Display verwendet. Die hier vorgestellten Anzeige- und Bedienkonzepte beschränken sich auf ein maximal 10 Zoll großes Farbdisplay in der Mittelkonsole, dem CID. Weiterer Ausgabekanal ist die Sprachausgabe z. B. für ein Feedback oder zur Dialogführung. Als Weiteres könnten Rückmeldungen über Signaltöne oder haptisches Feedback gegeben werden.

Die Bedienungsaufgabe teilt sich in die Formulierung der Suchanfrage, Navigation in der Ergebnisliste, Sortier- und Filterfunktionen und das Ausführen objektverknüpfter Funktionen auf. Zur Formulierung der Suchanfrage sind alphanumerische Eingabestrategien erforderlich. In

Abschnitt 4.4 werden Eingabestrategien mittels Dreh-/Drücksteller, Touchscreen, Touchpad, Tastatur, Scrollrad und Sprache vorgestellt. Das Starten der Anfrage muss, falls nicht automatisch über einen Timer erfolgt, vom Nutzer aktiviert werden (z.B. über eine Taste). Zur Navigation in den Ergebnislisten, Auswahl von Sortier- und Filterfunktionen und der Nutzung von objektverknüpften Funktionen eignen sich Dreh-/Drücksteller, Touchscreen, Hardkeys und Sprache. Für die Suchmaschine wird auf alternierende multimodale Interaktion zurückgegriffen.

### 4.3 Mathematische Methoden

Dieser Abschnitt gibt eine Einführung zum Thema Volltextsuche. Dabei werden verbreitete und wichtige Algorithmen dargestellt und auf weiterführende Literatur verwiesen. Am Ende wird die Einsetzbarkeit der verschiedenen Suchverfahren im Auto diskutiert.

#### 4.3.1 Exakte Suche

Die Aufgabe der exakten Suche einer Zeichenfolge in einem Text besteht daraus, ein oder alle Vorkommen der Zeichenfolge in einem Text zu finden, wobei sowohl die Zeichenfolge, als auch der Text aus Zeichen eines bestimmten Alphabets stammen. Im Folgenden werden verschiedene Algorithmen kurz vorgestellt. Für eine ausführliche Behandlung der vorgestellten Algorithmen wird auf die Zusammenstellung von Baeza-Yates [BY89] verwiesen.

Für die Vorstellung der Suchalgorithmen wird von einer Länge des Textes  $n$  und einer Länge der zu suchenden Zeichenfolge  $m$  ausgegangen.

#### Naiver Suchalgorithmus

Der nahe liegendste Ansatz in einem Text eine Zeichenfolge zu finden besteht darin, den Text von vorne nach hinten durchzugehen und bei jedem der  $n$  Zeichen im Text zu prüfen, ob dies und die folgenden  $m - 1$  Zeichen mit der Suchfolge übereinstimmen. Also zuerst wird überprüft, ob die ersten  $m$  Zeichen im Text der zu suchenden Zeichenfolge entsprechen. Ist dies nicht der Fall, werden vom zweiten Zeichen an die Zeichen mit der Suchfolge verglichen, bis am Ende der gesamte Text durchlaufen ist.

Im Idealfall kann nach dem ersten oder zweiten Zeichen entschieden werden, ob eine Zeichenfolge gefunden wurde oder nicht. Daher ist die Laufzeit im Durchschnitt  $O(m + n)$ . Im schlimmsten Fall, wenn beispielsweise die Zeichenfolge "mmk" im Text "mmmmmmmmmmmk" gesucht wird, beträgt die Laufzeit  $O(mn)$ .

#### Knuth-Morris-Pratt-Algorithmus

Verbesserungen des naiven Suchalgorithmus basieren darauf, dass sie entweder den Text oder die zu suchende Zeichenfolge vor der eigentlichen Suche analysieren oder aufbereiten. Der Knuth-Morris-Pratt (KMP) Algorithmus nutzt die Kenntnis über den Aufbau der zu suchenden Folge, um Rücksprünge im Text nach fehlgeschlagenen Vergleichen möglichst klein zu halten [KMP77].

Das Beispiel in Abbildung 4.5 soll die Motivation für den KPM Algorithmus verdeutlichen. Dabei soll die Folge  $F$ ="tum" im Text  $T$ ="tututum" mit dem naiven Suchalgorithmus gesucht

	0	1	2	3	4	5	6
T:	t	u	t	u	t	u	m

i=0:	t	u	X				
i=1:	X						
i=2:		t	u	X			
i=3:			X				
i=4:				t	u	m	

Abbildung 4.5: Beispiel für den naiven Suchalgorithmus

	0	1	2	3	4	5
T:	t	u	m	m	m	k

i=0:	X					
i=1:	X					
i=2:		m	m	X		
i=3:				m	X	
i=4:					m	

Abbildung 4.6: Beispiel für die fehlgeschlagene Suche mit einem abgewandelten naiven Suchalgorithmus

werden. Ein “X” deutet an, dass der Vergleich fehlgeschlagen ist. Im Beispiel ist unschwer zu erkennen, dass nach der ersten Iteration ( $i = 0$ ) mit der Information, dass  $T[1] = 'u'$  und dass  $F[0] = 't'$  ist in der zweiten Iteration der Vergleich von  $T[1]$  und  $F[0]$  nicht mehr notwendig ist. Mit dem Wissen über den Aufbau von  $F$  hätte man gleich zum Vergleich von  $T[2]$  und  $F[0]$  übergehen, also zur Textstelle  $T[2]$  springen können.

Es existieren aber Fälle, in denen man nicht einfach zur Stelle des fehlgeschlagenen Vergleichs springen kann, um den Vergleich erneut mit  $F[0]$  zu beginnen. Hierzu noch ein Beispiel: Es soll die Folge  $F = \text{“mmk”}$  in dem Text  $T = \text{“Tummmk”}$  gesucht werden. Diesmal mit dem abgewandelten naiven Algorithmus, der nach fehlgeschlagenem Vergleich die Suche an dieser Stelle mit einem Vergleich des Textbuchstabens mit dem ersten Zeichen der Suchfolge fortsetzt. Die einzelnen Schritte sind in Abbildung 4.6 dargestellt. In diesem Fall würde der abgewandelte naive Suchalgorithmus die Suchfolge nicht finden. Dies passiert immer dann, wenn die Anfangsteilfolge der Suchfolge, die schon verglichen wurde bevor der fehlerhafte Vergleich verursacht wird, eine weitere Anfangsteilfolge der Suchfolge enthält. Auf das Beispiel bezogen ist die Anfangsteilfolge die schon verglichen wurde “mm” und diese enthält erneut eine Anfangsteilfolge, nämlich “m”. In dem Fall, müsste der Algorithmus an der Stelle des fehlgeschlagenen Vergleichs  $T[x]$ , nicht  $T[x]$  mit  $F[0]$ , sondern mit einem  $F[y]$  vergleichen, wobei  $y$  die Länge der Anfangsteilfolge der Suchfolge ist, welche in der Anfangsteilfolge der Suchfolge ist, die schon vor dem fehlgeschlagenen Vergleich verglichen wurde. Die Anfangsteilfolge müsste also um die Länge der Teilfolge verschoben werden. In Abbildung 4.7 wird dies nochmals am Beispiel verdeutlicht. Im vierten Schritt wird die für den Vergleich herangezogene Zeichenfolge um eins nach rechts verschoben. Daher wird nicht mit dem ersten, sondern dem zweiten “m” verglichen und schließlich die Zeichenfolge gefunden.

Diese Idee wird von dem KMP Algorithmus aufgegriffen, indem vor der eigentlichen Suche durch Analyse der Suchfolge die Information für diese Schiebevorgänge berechnet und in einer Tabelle `next` festgehalten werden. Sei  $j$  die Position der Suchfolge an deren Stelle ein fehlgeschlagener Vergleich stattgefunden hat, dann gibt `next[j]` die Position der Suchfolge wieder, die beim nächsten Vergleich untersucht wird. Die Suchfolge wird also um  $j - \text{next}[j]$  Stellen relativ zum Text nach rechts geschoben. Tabelle 4.2 zeigt die `next`-Tabelle die sich für die Suchfolge "mmk" ergibt. Für die genaue Berechnungsvorschrift für die Tabelle sei auf [KMP77] verwiesen.

Die Suche findet basierend auf der Information aus der Tabelle nach der Berechnungsvorschrift

j	0	1	2
next[j]	0	1	1

Tabelle 4.2: `next`-Tabelle für die Suchfolge "mmk"

wie in Algorithm 1 dargestellt statt. Der KMP-Algorithmus hat eine Laufzeit von  $O(m + n)$ ,

---

#### Algorithm 1 Knuth Morris Pratt Algorithmus

---

```

1: procedure KMP SUCHE(TextT, SuchfolgeF)
2:   Beginne am Anfang von T
3:   while F noch nicht gefunden and T noch nicht am Ende do
4:     while Zeichen aus F  $\neq$  aktuelles Zeichen aus T do
5:       Schiebe F um next[j] weiter
6:     end while
7:     Gehe zu nächstem Zeichen in T
8:   end while
9: end procedure

```

---

wobei die anfängliche Analyse der zu suchenden Zeichenfolge linear von deren Länge, also  $O(m)$ , abhängt. Die Dauer der eigentlichen Suche ist daher unabhängig von der Länge der Suchfolge.

Eine exakte Formulierung des Algorithmus, sowie einige Optimierungen sind in [KMP77] zu finden.

```

      0  1  2  3  4  5
T:   t  u  m  m  m  k

i=0: X
i=1:   X
i=2:     m  m  X
i=3:       m  m  k

```

Abbildung 4.7: Die Idee des Knuth-Morris-Pratt Algorithmus

### Boyer-Moore Algorithmus

Ein weiterer prominenter Algorithmus zur exakten Suche in einem Text ist der Algorithmus von Boyer und Moore. Er ist insbesondere für natürlich-sprachliche Alphabete effizient, allerdings recht kompliziert. Im Folgenden soll die Idee kurz skizziert werden; für eine detaillierte Beschreibung sei auf die Originalschrift [BM77] verwiesen.

Eine grundlegende Eigenschaft des Boyer-Moore (BM) Algorithmus ist die Tatsache, dass der Vergleich am Ende der Zeichenfolge beginnt. Er begründet sich in der Beobachtung, dass in diesem mehr nützliche Informationen abgeleitet werden können als ein Vergleich von links her und der Algorithmus so mit weniger Vergleichen auskommt.

Die Prinzipien des BM Algorithmus sollen im Folgenden informell erläutert werden.

Findet ein Vergleich mit einem Zeichen  $c$  aus dem Text  $T$  statt, welches nicht in der Suchfolge  $F$  vorkommt, so kann die Suchfolge bis an die Stelle nach dem Vorkommen von  $c$  verschoben werden, ohne einen weiteren Vergleich anzustellen. Hierzu ein Beispiel:

```
Suchfolge  $F$ :      mmk
Text  $T$ :          .....Sonntag ist ein schöner Tag
```

Im Beispiel werden 'k' aus  $F$  und 'n' aus  $T$  verglichen. 'n' kommt in  $F$  nicht vor, daher kann mit dem Vergleich  $|F|$  Stellen später fortgesetzt werden ( $|F|$  ist die Länge von  $F$ ):

```
Suchfolge  $F$ :      mmk
Text  $T$ :          .....Sonntag ist ein schöner Tag
```

Verallgemeinert: Angenommen wir vergleichen das letzte Zeichen der Suchfolge  $F$  mit dem Zeichen der aktuellen Stelle in  $T$ ,  $c$ , und der Vergleich schlägt fehl. Ist  $c$  nicht in  $F$  enthalten, dann können wir, wie oben im Beispiel, den Vergleich  $|F|$  Stellen später fortsetzen ohne weitere Vergleiche durchzuführen. Ist  $c$  in  $F$  enthalten und ist das sich am weitesten rechts befindende Vorkommnis von  $c$  in  $F$   $\delta$  Zeichen vom rechten Ende von  $F$  entfernt, dann kann die Suchfolge  $|F|$  Zeichen nach rechts verschoben werden, ohne weitere Vergleiche anzustellen. Der nächste Vergleich findet wieder am Ende von  $F$  statt. Der Grund hierfür ist, dass wenn  $F$  um weniger als  $\delta$  Zeichen nach links geschoben werden würde, dann müsste es noch ein Vorkommnis von  $c$  in  $F$  weiter rechts, als dem am weitesten rechts befindende Zeichen geben, um die Suchfolge bei den dazwischenliegenden Vergleichen zu finden.

Im BM-Algorithmus werden noch zwei weitere Prozeduren für den Fall definiert, dass die letzten  $x$  Vergleiche erfolgreich waren und jetzt ein fehlgeschlagener Vergleich vorliegt.

Der BM Algorithmus weist eine Zeitkomplexität von  $O(m + n)$  auf. Er arbeitet immer dann am besten, wenn viele Zeichen vorgefunden werden, die nicht in der Suchfolge enthalten sind. Dies ist insbesondere dann der Fall, wenn das Alphabet viele Zeichen enthält und die Suchfolge sehr kurz ist. Für diese Fälle ist die praktische Laufzeit oft sehr viel kürzer.

#### 4.3.2 Invertierte Indizes

Die beiden oben genannten Algorithmen, der Knuth-Morris-Pratt und der Boyer-Moore Algorithmus sind klassische Suchalgorithmen in einem Text. Auf Datenbanken angewandt, können die Algorithmen auf jedes Textfeld angewandt werden. Beide analysieren die Suchfolge vorab, um den eigentlich Suchvorgang effizienter zu gestalten. Ebenso kann auch der Text vorab

analysiert werden, um anschließend Operationen schneller zu gestalten. In diesem Zusammenhang spricht man oftmals von einer Indexstruktur, die ausgehend vom Text aufgebaut wird. Indexstrukturen sind sehr verbreitet im Bereich von Datenbanken, da sehr häufige Anfragen an Datenbanken erwartet werden und der Aufwand für das initiale Erstellen des Index daher vertretbar ist. Eine gängige Datenstruktur hierfür sind *Suffix trees*, welche in linearer Zeit aus dem Text erstellt werden können [Ukk95, Far96]. Je nach Implementierung können dann Operationen, wie Suche oder Einfügen in konstanter Zeit, durchgeführt werden. Das Problem bei diesen Indexstrukturen ist allerdings, dass der Speicherplatz der Indexstruktur oftmals größer ist als der benötigte Speicherplatz für den reinen Text. Weniger Speicherbedarf benötigen invertierte Indizes. Invertierte Indizes werden unter anderem für die Volltextsuche in Datenbanken wie MySQL oder Volltextsuchmaschinen wie Apache Lucene, als auch in bekannten Websuchmaschinen [BP98] eingesetzt. Eine umfangreiche Erklärung mit aktuellen Beiträgen zu invertierten Indizes sowie ein umfangreiches, kommentiertes Literaturverzeichnis mit den wichtigsten Veröffentlichungen in diesem Bereich wird in [ZM06] gegeben. An dieser Stelle soll lediglich die grundlegende Idee von invertierten Indizes anhand eines Beispiels vermittelt werden.

Invertierte Indizes speichern die Vorkommnisse von Wörtern in Dokumenten oder Datenbankeinträgen (*records*) ab. Angenommen es sind 3 Einträge in der Datenbank:  $E_0$ ="TUM ist spitze",  $E_1$ ="MMK an der TUM",  $E_2$ ="MMK ist interessant", dann ergeben sich für den Index auf Ebene der Einträge (oft als *inverted file index* bezeichnet) die Werte wie in Abbildung 4.8 dargestellt. Der Index auf Wortebene enthält noch zusätzlich die Position des Wortes im Ein-

```

''TUM'':           {0, 1}
''ist'':           {0, 2}
''spitze'':       {0}
''MMK'':           {1, 2}
''an'':            {1}
''der'':           {1}
''interessant'':  {2}

```

Abbildung 4.8: Invertierter Index auf Ebene der Einträge

trag, wie in Abbildung 4.9 dargestellt. Wird nun eine Suchanfrage gestellt, muss lediglich im

```

''TUM'':           {(0, 0), (1, 3)}
''ist'':           {(0, 1), (2, 1)}
''spitze'':       {(0, 2)}
''MMK'':           {(1,0), (2, 0)}
''an'':            {(1, 1)}
''der'':           {(1, 2)}
''interessant'':  {(2, 2)}

```

Abbildung 4.9: Invertierter Index auf Ebene der Wörter

invertierten Index das zu suchende Wort gefunden werden, um anschließend alle Vorkommnisse in der Datenbank zu kennen. Die Suche des Suchwortes im Index kann beispielsweise mittels binärer Suche in einer Laufzeit von  $O(\log N)$  realisiert werden.

### 4.3.3 Unscharfe Suche

Bei der unscharfen Suche werden in einem Text Zeichenfolgen gesucht, die mit einer zu suchenden Folge zu einem bestimmten Grad übereinstimmen. Im Folgenden werden zwei verschiedene Formen der unscharfen Suche vorgestellt. Im ersten Fall wird der Grad der Übereinstimmung zweier Zeichenfolgen über eine Distanzfunktion berechnet. Im Anschluss werden Verfahren definiert, welche auf der Aussprache von Wörtern beruhen – hierzu werden Phoneme definiert. Es werden dann als Ergebnisse bei einer Suche nach einer bestimmten Zeichenfolge Wörter gefunden, die in der Aussprache ähnlich klingen, auch wenn sie ganz verschieden buchstabiert werden.

#### Distanzfunktion

Bei der unscharfen Suche, die auf Distanzfunktionen basiert, wird der Grad der Übereinstimmung meist durch die Anzahl der erlaubten Fehler beim Vergleich definiert. Formal kann das Problem wie folgt definiert werden [Nav01]:

Sei  $\Sigma$  ein endliches Alphabet mit  $\sigma$  Zeichen,  $T \in \Sigma^*$  ein Text der Länge  $n$ ,  $F \in \Sigma^*$  eine Zeichenfolge der Länge  $m$  und  $k \in \mathbb{R}$  die Anzahl der maximal erlaubten Fehler und  $d : \Sigma^* \times \Sigma^* \rightarrow \mathbb{R}$  eine Distanzfunktion, dann ist das Problem der unscharfen Suche wie folgt definiert:

Gegeben sei  $T$ ,  $F$ ,  $k$  und  $d()$ ; finde die Menge aller Positionen  $j$  im Text, so dass gilt:  $\exists i : d(P, T_{i..j}) \leq k$ .

#### Verschiedene Distanzfunktionen

Der Wert des Unterschieds zweier Zeichenfolgen berechnet sich bei den gängigen Distanzfunktionen durch die minimale Anzahl der Anwendungen von gewissen Operationen die benötigt werden um eine Zeichenfolge in die andere zu überführen. Die prominentesten Distanzfunktionen sind:

- **Levenshtein Distanz:**

Die kleinste Anzahl von Operationen Einfügen, Löschen und Austausch von Zeichen, um eine Zeichenfolge in die andere zu überführen. Jede Operation kostet 1.

Beispiel:  $d_{Levenshtein}(\text{“Stefan”}, \text{“Steffen”}) = 2$

- **Hamming Abstand:**

Bei dem Hamming Abstand ist nur der Austausch erlaubt (Kosten 1). Es können daher nur Zeichenfolgen ineinander überführt werden, die die gleiche Länge besitzen. Zeichenfolgen ungleicher Länge haben einen Hammingabstand von  $\infty$ .

Beispiel:  $d_{Hamming}(\text{“Meier”}, \text{“Mayer”}) = 2$

- **Episodendistanz:**

Bei der Episodendistanz ist nur das Einfügen erlaubt (Kosten 1). Daher ist die Distanz zweier Zeichenfolgen  $a$  und  $b$  nicht symmetrisch. Es gilt entweder  $d(a, b) = |b| - |a|$  oder  $d(a, b) = \infty$ .

Beispiel:  $d_{Episode}(\text{“aaabb”}, \text{“aaabbba”}) = 2$

aber  $d_{Episode}(\text{“aaabbba”}, \text{“aaabb”}) = \infty$

- **Längste gemeinsame Teilfolge (Longest Common Substring – LCS):**

Es sind nur die Operationen Einfügen und Löschen erlaubt, jeweils mit Kosten 1. Die Distanz ist die Anzahl Zeichen für die es keinen Partner in der anderen Zeichenfolge

gibt. Hierbei wird die Länge der längsten gemeinsamen Teilfolge berechnet. Diese Distanzfunktion wird insbesondere auch beim Vergleich von DNA Ketten eingesetzt.

Beispiel:  $d_{LCS}(\text{“Steffen”}, \text{“Stefan”}) = 3$

- **Tippfehlerdistanz:**

Diese Distanzfunktion entspricht der Levenshtein Distanz, erweitert um die Operation Vertauschen zweier Zeichen (Kosten jeweils 1).

Beispiel:  $d_{Tippfehler}(\text{“Hans”}, \text{“Hnas”}) = 1$

Neben diesen grundlegenden Distanzfunktionen existieren noch viele weitere Funktionen. Denkbar, und sinnvoll für bestimmte Einsatzgebiete, ist auch die Gewichtung von Operationen. Eine ausführlichere Diskussion von Distanzfunktionen kann in [Nav01] gefunden werden. Für die unscharfe Suche in natürlich-sprachigen Texten ist die Levenshtein Distanz am weitesten verbreitet.

### Algorithmen zur Berechnung der unscharfen Suche

Die bekanntesten Berechnungsverfahren sind dynamische Programmierung, Algorithmen die den Bitparallelismus von Computern ausnutzen, Nichtdeterministische Automaten und Filteralgorithmen. Dynamische Programmierung und Nichtdeterministische Automaten haben wegen des hohen Zeit- und Speicherbedarf keine wirkliche praktische Relevanz. Ein auf dynamischer Programmierung basierender Algorithmus wird jedoch aufgrund seiner guten Verständlichkeit kurz vorgestellt. Für die Behandlung von Automaten und auf Bitparallelismus basierenden Algorithmen sei an dieser Stelle auf [Nav01] verwiesen. Als Distanzfunktion wird für die Vorstellung des Algorithmus die Levenshtein Distanz angenommen.

**Dynamische Programmierung:** Es wird ein einfacher Algorithmus vorgestellt, welcher die Levenshteindistanz zwischen zwei Zeichenfolgen  $F_1$  und  $F_2$  berechnet. Dieser Algorithmus wird anschließend so abgeändert, dass er auf die Suche einer Zeichenfolge  $F$  in einem Text  $T$  bei  $k$  erlaubten Fehlern angewendet werden kann.

Der Algorithmus basiert auf dynamischer Programmierung. Hierzu wird eine Matrix  $D_{0..|F_1|, 0..|F_2|}$  betrachtet. Im Eintrag  $D_{i,j}$  ist die minimale Anzahl an Einfüge-, Lösch- und Austauschoperationen eingetragen, die benötigt wird um  $F_{1..i}$  zu  $F_{2..j}$  zu überführen. Die Einträge berechnen sich nach der Vorschrift:

$$\begin{aligned} D_{i,0} &= i \\ D_{0,j} &= j \\ D_{i,j} &= \begin{cases} D_{i-1,j-1} & \text{für } F_{1_i} = F_{2_j} \\ 1 + \min(D_{i-1,j}, D_{i,j-1}, D_{i-1,j-1}) & \text{sonst} \end{cases} \end{aligned}$$

Der Algorithmus muss nun dafür sorgen, dass vor der Berechnung eines Eintrags der Matrix, der Eintrag links von dem aktuell zu berechnenden, links-oben und oben schon berechnet sind, da diese Werte für die Berechnung des Wertes des aktuellen Feldes herangezogen werden können. Dies kann leicht durch eine reihenweise Berechnung, jeweils von links-nach rechts, oder eine spaltenweise Vorgehensweise von oben nach unten realisiert werden.

Die Matrix für den oben vorgestellten Vergleich  $F_1 = \text{“Stefan”}$  und  $F_2 = \text{“Steffen”}$  ist in Abbildung 4.10 dargestellt. Um nun eine Zeichenfolge in einem Text  $T$  zu suchen, wird  $F_2$  gleich  $T$  gesetzt und alle  $D_{0,j} = 0$ . Letzteres bewirkt, dass jedes Zeichen aus  $T$  der Beginn der zu suchenden Zeichenfolge sein kann und daher mit der Distanz 0 startet. Für die Suche der Folge

		S	t	e	f	f	e	n
	0	1	2	3	4	5	6	7
S	1	0	1	2	3	4	5	6
t	2	1	0	1	2	3	4	5
e	3	2	1	0	1	2	3	4
f	4	3	2	1	0	1	2	3
a	5	4	3	2	1	1	2	3
n	6	5	4	3	2	2	2	2

Abbildung 4.10: Berechnung der Levenshtein Distanz mittels dynamischer Programmierung

$F_1 = \text{“Stefan”}$  in dem Text  $T = \text{“Steffen”}$  ergibt sich die Matrix aus Abbildung 4.11.

Sind  $k = 2$  Fehler erlaubt, so findet der Algorithmus 4 Übereinstimmungen der Zeichenfolge

		S	t	e	f	f	e	n
	0	0	0	0	0	0	0	0
S	1	0	1	1	1	1	1	1
t	2	1	0	1	2	2	2	2
e	3	2	1	0	1	2	2	3
f	4	3	2	1	0	1	2	3
a	5	4	3	2	1	1	2	3
n	6	5	4	3	2	2	2	2

Abbildung 4.11: Berechnung der Levenshtein Distanz mittels dynamischer Programmierung für eine Suche im Text

“Stefan” im Text “Steffen”, nämlich “Stef”, “Steff”, “Steffe” und “Steffen”.

### Phonetische Suche

Eine weitere Klasse von Algorithmen zur unscharfen Suche sind phonetische Algorithmen. Ziel ist das Auffinden von Zeichenfolgen in einem Text, welche gleich oder ähnlich wie die zu suchende Zeichenfolge ausgesprochen werden. Denkbare Beispiele hierfür sind in deutscher Sprache “Schmidt” und “Schmitt”, oder “Meier” und “Mair”. Der bekannteste aller phonetischen Algorithmen ist Soundex. Er wurde schon im 19. Jahrhundert unter dem Namen *Miracode* verwendet und 1918 patentiert [HD80]. Einsatzgebiet des Algorithmus war ein Index für Karteikarten mit Namen, die mittels der Soundex Codes nach der Aussprache geordnet wurden, um die Suche zu erleichtern. Auch heute sind Soundex und seine Weiterentwicklungen noch im Einsatz: Datenbanksysteme wie Mysql und Postgresql bieten Funktionen für unscharfe Suche mittels Soundex an. Im Folgenden wird das Prinzip von soundex vorgestellt und auf Weiterentwicklungen hingewiesen.

### Soundex

Bei Soundex werden Wörtern in einen vierstelligen Soundex Code überführt [HD80]. Die Berechnungsgrundlage für die Umwandlung ist in Abbildung 4.12 dargestellt. Aus einem Wort wird der Code berechnet, indem man alles außer den ersten Buchstaben gemäß der obigen

A E I O U H W Y	0
B F P V	1
C G J K Q S X Z	2
D T	3
L	4
M N	5
R	6

Abbildung 4.12: Berechnungsvorschrift für den Soundex Code

Tabelle durch die entsprechende Zahl ersetzt, alle direkten Wiederholungen von Zahlen im Code, sowie alle Nullen löscht und die ersten vier Zeichen des Codes zurück gibt.  
Es folgen einige Beispiele:

DIXON	D25
DICKSON	D25
RODGERS	R326
ROGERS	R262
KNIGHT	K523
NIGHT	N23
SCHMIDT	S253
SCHMITT	S253
MEIER	M6
MAIR	M6
CATHERINE	C365
COTRONEO	C365

Anhand der Beispiele erkennt man, dass viele Namen mit gleicher Aussprache auf denselben Code abgebildet werden. Allerdings gilt das nicht für alle wie das Beispiel “Rodgers” und “Rogers” zeigt. Ein prinzipielles Problem ist, dass Wörter mit verschiedenen Anfangsbuchstaben immer auf verschiedene Codes abgebildet werden (“KNIGHT” und “NIGHT”). Des Weiteren werden auch Zeichenfolgen die nicht zur selben Aussprache führen auf den identischen Soundex Code abgebildet. Insbesondere die Tatsache, dass Soundex für die englische Aussprache entwickelt wurde, führt in anderen Sprachen zu ungewünschten Abbildungen von Wörtern auf Soundex Codes, wie folgendes Beispiel zeigt, welches “Britney Spears” einer “bewährten Superzicke” gleichstellt:

BRITNEY SPEARS	B635 S162
BEWÄHRTEN SUPERZICKE	B635 S162

Es gibt viele Weiterentwicklungen von Soundex, die teilweise die oben genannten Probleme kompensieren. Die „Kölner Phonetik“ (von Postel, 1969), dargestellt in 4.3 passt die Codierung von Buchstaben auf Zahlen an die deutsche Sprache an [Wil05]. Mit Anlaut ist in der folgenden Tabelle der erste Buchstabe eines Wortes gemeint.

In [ZD96] und [Wil05] werden verschiedene phonetische Algorithmen wie Metaphon oder Phonix vorgestellt und deren Zweckmäßigkeit diskutiert. Erwähnenswert ist ein Ansatz bei Phonix, welcher die Endsilben trunkiert, um beispielsweise eine Ähnlichkeit von Wörtern mit Einzahl- und Pluralendungen zu erkennen.

#### 4.3.4 Diskussion der verschiedenen Suchverfahren für den Einsatz im Fahrzeug

Der Einsatz von Suchalgorithmen im Automobil ist geprägt von der Tatsache, dass die Interaktion mit der grafischen Benutzeroberfläche neben der eigentlichen Fahraufgabe ausgeführt wird und daher nicht die volle Aufmerksamkeit des Nutzers erwartet werden kann. Dies lässt eine höhere Fehleranfälligkeit bei Eingaben erwarten. Des Weiteren soll dem Nutzer bei der Präsentation der Suchergebnisse eine manuelle Beurteilung der Suchergebnisse weitgehend erspart bleiben um nicht zu viel Aufmerksamkeit zu erzeugen. Daher wird einerseits eine hohe Korrektheit der Suchergebnisse gefordert; andererseits darf jedoch die Vollständigkeit der Suchergebnisse nicht darunter leiden, da das Fehlen von Suchergebnissen, von deren Existenz der Nutzer überzeugt ist, den Nutzer irritiert und von der primären Fahraufgabe ablenkt.

Unterschieden werden muss die Form der Eingabe des Nutzers. Eine Eingabe des Nutzers mittels Sprache legt den Einsatz eines phonetischen Suchalgorithmus nahe. Erfolgt die Eingabe durch Texteingabe, ist ein Fehler durch Verschreiben am wahrscheinlichsten. Diese Fehler können am besten durch Algorithmen gefunden werden, die auf Distanzfunktionen basieren. Bei phonetischen Algorithmen ändert sich mit Vertauschen eines Buchstaben meist auch der Code und es können keine Ähnlichkeiten mehr festgestellt werden.

Unabhängig von der Eingabeform ist ein stufenweiser Ansatz zu empfehlen. Denkbar ist im ersten Schritt eine exakte Suche. Liefert diese Suche keine Ergebnisse zurück, so wird eine unscharfe Suche gestartet. Diese kann mehrere Iterationen durchlaufen, wobei jeweils die geforderte Ähnlichkeit gelockert werden kann. Auch die Kombination verschiedener Algorithmen ist denkbar. Beispielsweise kann bei Texteingabe erst eine Suche basierend auf Levenshtein

Zeichen	Kontext	Symbol
A,E,I,J,Y,O,U	im Anlaut	0
H		-
B,P		1
D,T	nicht vor C,S,Z	2
F,PH,V,W		3
G,K,Q		4
C	im Anlaut, vor A,H,K,L,O,Q,R,U,X ansonsten, vor A,O,U,H,K,X,Q	4
X	wenn nicht nach C,K,Q	48
L		5
M,N		6
R		7
S,Z		8
C	im Anlaut, nicht vor A,H,K,L,O,Q,R,U,X folgt ansonsten, nicht vor A,O,U,H,K,X,Q nach S, Z D,T vor S,C, Z X nach C,K,Q	

Tabelle 4.3: Codierung von Buchstaben nach der "Kölner Phonetik"

Distanz in mehreren Iteration mit größerem erlaubten Abstand durchgeführt werden und anschließend noch eine phonetische Suche. Denkbar ist auch, dass erneute Suchiterationen vom Nutzer angestoßen werden können, insofern sein erwartetes Ergebnis noch nicht in der Liste der Resultate auftaucht.

Hinsichtlich Laufzeit und Speicherbedarf der Algorithmen ist der Einsatz von Indizes zu prüfen. Bei Distanzfunktionen muss die Distanz bei jeder Suchanfrage berechnet werden. Hier könnten beispielsweise Inverted Files zur Verkürzung der Laufzeit beitragen. Indizes führen allerdings zu einem erhöhten Speicherbedarf, da sie zusätzlich zu den eigentlichen Daten persistent gespeichert werden müssen. Aktuell ist Speicher noch als knappe Ressource in Fahrzeugen anzusehen – dies dürfte sich aber in der näheren Zukunft mit dem Verbau von Festplatten als weniger kritisch darstellen.

### 4.3.5 Rankingverfahren

In diesem Abschnitt wird nun auf verschiedene Rankingverfahren eingegangen. Das Ranking (Rangliste, Rangordnung) ist die Sortierung von relevanten Objekten entsprechend einer Bewertung. Betrachtet man hier die Internet-Suchmaschinen, so verwenden diese verschiedene Ranking-Algorithmen. Bei der Abfrage von Suchmaschinen werden die Ergebnisse in einer vom System vorgegebenen Rangordnung dargestellt: Die ersten Treffer sollten dabei die größte Relevanz für den Suchbegriff aufweisen. Eines der Kriterien ist die Analyse des Inhalts einer Seite (On-Page-Faktor). Eines weiteres Kriterium ist die Zahl der Verweise anderer Webseiten auf eine bestimmte Seite (OFF-Page-Faktor). Je mehr Links im Internet auf die betreffende Webseiten verweisen, desto besser wird sie eingestuft. Den Indikator dafür nennt man Linkpopularität. Die Kunst aktueller Ranking-Verfahren besteht in einer geeigneten Kombination verschiedenster Verfahren.

#### PageRank

Der PageRank-Algorithmus ist wohl das bekannteste Verfahren zur Bewertung von Internet-Seiten. Der Algorithmus wurde Ende der 90er von Larry Page und Sergey Brin an der Stanford University entwickelt und wird modifiziert bei Google eingesetzt [PBMW98]. Der PageRank-Wert ist ein Off-Page-Faktor und spiegelt nur die Verlinkungsstruktur auf eine Seite wieder. Dabei besitzt jede Seite ein Gewicht, das um so größer ist, je mehr Seiten auf diese verweisen. Das PageRank-Gewicht  $PR_i$  der auf  $i$  verlinkte Seiten  $j$  berechnet sich über folgende Gleichung:

$$PR_i = \frac{1-d}{N} + d \sum_{\forall j \in \{(j,i)\}} \frac{PR_j}{C_j} \quad (4.1)$$

( $N$  ist die Gesamtanzahl der Seiten und  $d$  ein Dämpfungsfaktor zwischen 0 und 1,  $C_j$  ist die Anzahl von Links auf Seite  $j$ )

#### Hubs und Authorities

Hubs und Authorities ist eine Ranking-Methode der Graphentheorie zur Erkennung von thematischen Zusammenhängen. Diese Methode wurde 1999 unter dem Namen *Hypertext-Induced Topic Selection* (HITS) bekannt und wurde von Jon Kleinberg entwickelt.

Ein Hub bezeichnet eine Seite, die auf andere Seiten verlinkt, beispielsweise populäre Linksammlungen. Mit Authority meint man eine Seite, die von anderen (Hubs) verlinkt ist. Der

HITS-Algorithmus nutzt das Gleichgewicht zwischen Hubs und Authorities im Web-Graphen. Jeder Seite  $i = 1 \dots n$  aus einer Grundmenge von Seiten wird ein Authority-Gewicht  $a_i$  und ein Hub-Gewicht  $h_i$  zugeordnet und können folgendermaßen berechnet werden:

$$h_i = \delta \sum_{j=1}^n A_{ij} a_j \quad (4.2)$$

$$a_i = \lambda \sum_{k=1}^n A_{ik}^T h_k \quad (4.3)$$

(A ist die Verlinkungsmatrix)

### Vektor Space Model

Das Vektorraum-Retrieval (engl. Vector Space Model (VSM)) ist ein Verfahren des Information Retrievals (IR), das Dokumente und Suchanfragen (Queries) als Punkte in einem hochdimensionalen, metrischen Vektorraum abbildet. Zum Retrieval wird die Distanz zwischen dem Queryvektor und dem Dokumentvektor genutzt.

**Topic-based Vector Space Model (TVSM):** Das *Topic-based Vector Space Model* (TVSM) stellt eine Erweiterung des Vector Space Model dar und wurde erstmals von [BK03] präsentiert. Im Unterschied zum VSM müssen die Termvektoren jedoch nicht orthogonal zueinander sein. Das TVSM ist deshalb in der Lage linguistische Phänomene besser abzubilden. In dieses Modell können Stop-Word-Lists, Stemming und Thesaurus vollständig integriert werden. Die Grundvoraussetzung des TVSM ist die Existenz eines  $d$  dimensionalen Raums  $R$ , der nur positive Werte besitzt:

$$R \in \mathfrak{R}_{\geq 0}^d \text{ mit } d \in N_{>0} \quad (4.4)$$

Jede Dimension von  $R$  repräsentiert ein grundlegendes Thema (Topic) und nimmt an, dass jedes Thema voneinander unabhängig ist. In der Fahrzeugsuchmaschine könnten z. B. Restaurants, Tankstellen, usw. einzelne Topics darstellen. Ein Term  $i \in \{1, \dots, n\}$  wird durch einen Term-Vektor  $\vec{t}_i$  repräsentiert und hat ein Term-Gewicht zwischen 0 und 1. Das Term-Gewicht zeigt die Relevanz eines Terms in der Relation zu einem Topic im Allgemeinen. Die Ähnlichkeit zwischen zwei Dokumenten  $k$  und  $l$  ist das Skalarprodukt der Dokumentvektoren. Eine detaillierte Beschreibung des TVSM kann in [BK03] gefunden werden. Dieser Algorithmus findet Anwendung in dem Suchinterface IMSENIO (siehe Abschnitt 4.5.2)

### Diskussion

Die Integration der Internet-Ranking-Verfahren in die Fahrzeug-Suchmaschinen stellt eine gewisse Schwierigkeit dar. Die Daten liegen meist strukturiert vor und dadurch können nicht so einfache Beziehungen der Daten untereinander hergestellt werden. Hier bieten sich evtl. eher einfache Ranking-Algorithmen an, wie z. B. die Sortierung nach Auswahlhäufigkeit, Auswahlhistorie oder in Abhängigkeit von Nutzerbewertungen. Ein weiteres Problem stellt möglicherweise auch die geringe Transparenz von Ranking-Algorithmen dar, was zu Verwirrungen beim Fahrer führen kann.

## 4.4 Alphanumerische Eingabestrategien

Alphanumerische Eingabeverfahren stellen für den Einsatz von Suchmaschinen eine entscheidende Rolle dar. Daher werden in diesem Abschnitt verschiedene Eingabemethoden erläutert. Dabei ist die geringe Ablenkung, die minimale Blickabwendung, die Robustheit und Verständlichkeit dieser Strategien von großer Bedeutung, um den Fahrer möglichst wenig von seiner Hauptaufgabe, dem Fahren, abzulenken.

### 4.4.1 Taktile Eingabegeräte mit haptischem Feedback

Nun werden verschiedene taktile Eingabegeräte für die Alphanumerik im Fahrzeug vorgestellt.

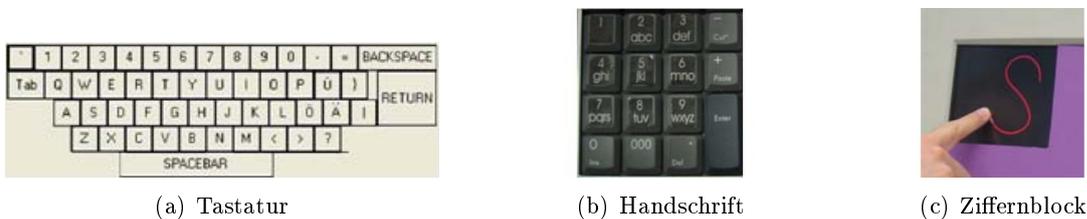


Abbildung 4.13: Taktile alphanumerische Eingabestrategien

**Dreh-/Drücksteller (DDS):** Eine sehr häufig verwendete Technik aktueller FIS ist die Eingabe über einen Dreh-/Drücksteller. Durch Drehen des Dreh-/Drückstellers wandert man durch die meist alphabetisch angeordneten Buchstaben und Zeichen. Abbildung 4.14 zeigt beispielhaft das Verfahren des aktuellen BMW 5er iDrives. Durch Drehen des Bedienelements wandert ein Cursor durch das Alphabet. Drücken man auf das Bedienelement, wird der aktuell markierte Buchstabe übernommen. Zum Löschen eines oder mehrerer Zeichen kann das Bedienelement nach rechts geschoben werden, wo dann der „Backspace“-Button durch Drücken ausgeführt werden kann. Zum Fortführen der Zeicheneingabe gelangt man über das Linksschieben wieder zurück in das Alphabet. Ist beispielsweise eine Ortsnamenliste hinterlegt, werden die nicht möglichen Buchstaben nach jeder Eingabe ausgegraut dargestellt und werden bei der Auswahl übersprungen. So können Fehleingaben vermieden und die Auswahl teilweise beschleunigt werden. Auch der Wechsel in die Liste ist hier durch Schieben des Bedienelements jederzeit möglich. Vorteil dieses Verfahrens ist die sehr gute Unterbrechbarkeit der Aufgabe und relativ geringe Blickzuwendungen. Jedoch werden oftmals die sehr langen Bediendauern als großer Kritikpunkt aufgeführt



Abbildung 4.14: Texteingabe über Speller

**Tastatur bzw. Ziffernblock:** Das etablierteste Eingabeverfahren im Computerbereich ist die Tastatur. Hierbei gibt es verschiedene Layouts. Das am weitesten verbreitetste Layout ist QWERTY, das aus noch aus dem Zeitalter der Schreibmaschinen stammt. Für den Fahrzeugbereich ist die Verwendung von Tastaturen aufgrund des hohen Platzbedarfs und der Bedienbarkeit nicht geeignet.

Eine reduzierte Form einer Tastatur ist der Nummern- bzw. Ziffernblock. Dieser wäre leichter in der Mittelkonsole integrierbar und bei Erstellung einer gut ertastbaren Profioberfläche zum Teil auch blindbedienbar. Das Eingabeverfahren von Ziffern wäre hier analog zu den Mobiltelefonen denkbar. Jeder Zahl bzw. jeder Taste sind mehrere Buchstaben zugeordnet. Durch mehrmaliges nacheinander folgendes Drücken einer Taste erfolgt die sequenzielle Auswahl des möglichen Buchstabens. Durch einmaliges Drücken der 2 schreibt man beispielsweise ein *A*, durch zweimaliges Drücken ein *B* usw. Aber auch das sogenannte *T9-Verfahren*<sup>2</sup> ist hier denkbar. Hierbei wird die jeweilige Taste lediglich einmal gedrückt. Bei der Eingabe mehrerer Buchstaben erfolgt anhand eines Lexikons der Vorschlag eines Begriffs.

#### 4.4.2 Touch-Eingabegeräte

Darüber hinaus ist die Verwendung berührungsempfindlicher Oberflächen zur alphanumerischen Eingabe möglich. Technische Ausführungen wären ein Touchpad oder auch ein Touchscreen. Bei einem Touchscreen erfolgt die Anzeige direkt überlagert mit der berührungsempfindlichen Oberfläche.

**Touchpad:** Bei einem Touchpad erweist sich besonders die Handschrifterkennung als eine zukunftsweisende Eingabestrategie alphanumerischer Zeichen im Fahrzeug. Dieses Verfahren wird oft auf PDAs oder PNAs<sup>3</sup> verwendet. Die Abmessungen eines solchen Gerätes sind gering und können daher sehr einfach z. B. in die Mittelkonsole integriert werden. Wie bei der Handschrift werden die gewünschten Buchstaben nacheinander über einem Finger auf das Touchpad geschrieben. Der Erkenner identifiziert über die angebundene Mustererkennung das gewünschte Zeichen und überträgt es an das FIS. Dabei kann es jedoch sehr einfach zu Erkennerfehlern kommen, die aber über ein geeignetes Fehlermanagement wie Feedback und einfache Korrektur kompensiert werden können. Großer Vorteil dieser Strategie ist bei entsprechender Auslegung wie Feedback über den haptischen und akustischen Kanal eine mögliche Blindbedienbarkeit.

**Touchscreen:** Bei der Verwendung eines Touchscreens hat man eine sehr flexible Möglichkeit zur Eingabe von alphanumerischen Eingaben. Durch die Koppelung mit einer graphischen Anzeige lassen sich hier verschiedenste Layouts realisieren.

Ein sehr bewährte Methode ist die Darstellung einer virtuellen Tastatur, sogenanntes On-Screen-Keyboards (OSK). Auch hier hat man die Möglichkeit der Handschrifterkennung mit visuellen Feedback. Der Touchscreen sollte so positioniert sein, dass er einen Kompromiss zwischen ergonomischer Erreichbarkeit mit der Hand und geringer Blickabwendung darstellt. Um die genaue Positionierung des Fingers im Fahrzeug zu erleichtern, ist hier eine Handgelenkstütze sinnvoll. Bei einem Touchscreen sind viele weitere Visualisierungsstrategien denkbar. Siehe dazu Abschnitt 7.2.1

---

<sup>2</sup>Abkürzung für Text on 9 keys (T9)

<sup>3</sup>Portable Navigation Assistant, ein mobiles Navigationssystem

### 4.4.3 Sprache

Auch Sprache eignet sich zur Eingabe von Text. Für die Nutzerakzeptanz ist jedoch ein großes Vokabular und eine hohe Robustheit zwingend erforderlich. Aktuelle Spracherkennung unterstützen diese hohen Anforderungen bisher nur sehr bedingt. Auch Filterung von Hintergrundgeräuschen ist im Fahrzeug von großer Bedeutung [SWM<sup>+</sup>08]. Als Alternative wäre eine Eingabe durch Buchstabieren möglich. Hier können die Buchstaben einfach nacheinander gesprochen, oder durch das Buchstabier-Alphabet eindeutiger gesprochen werden. Nachteil des Buchstabierens ist der hohe Zeitaufwand und die Fehlerrate.

## 4.5 Anzeige-/Bedienkonzepte für die Fahrzeug-Suchmaschine

In diesem Abschnitt werden nun mehrere multimodale Suchmaschinen-Interfaces für das Fahrzeug konzipiert. Dort fließen die vorgestellten Anforderungen und Methoden ein, bevor die Interfaces dann im nachfolgenden Abschnitt hinsichtlich ihrer Eignung im Fahrzeug untersucht werden.

### 4.5.1 Suchinterface CARhoo!

Das Look&Feel<sup>4</sup> der Fahrzeug-Suchmaschine CARhoo! ist nahe an das BMW iDrive Design angelehnt. Mit dem CARhoo! lassen sich neben umfangreichen Daten, z. B. MP3s, Kontakte oder POIs, auch klassischen Menüinhalte durchsuchen und auffinden. Die hybride Struktur ermöglicht neben der lokalen Suche auch eine Schnittstelle zu Internet-Suchmaschinen via Webservices. Des Weiteren können Hilfestellungen durchsucht und bereitgestellt werden. Die Bedienung des Suchinterfaces erfolgt multimodal, d. h. der Fahrer kann zwischen den folgenden Eingabemöglichkeiten wählen:

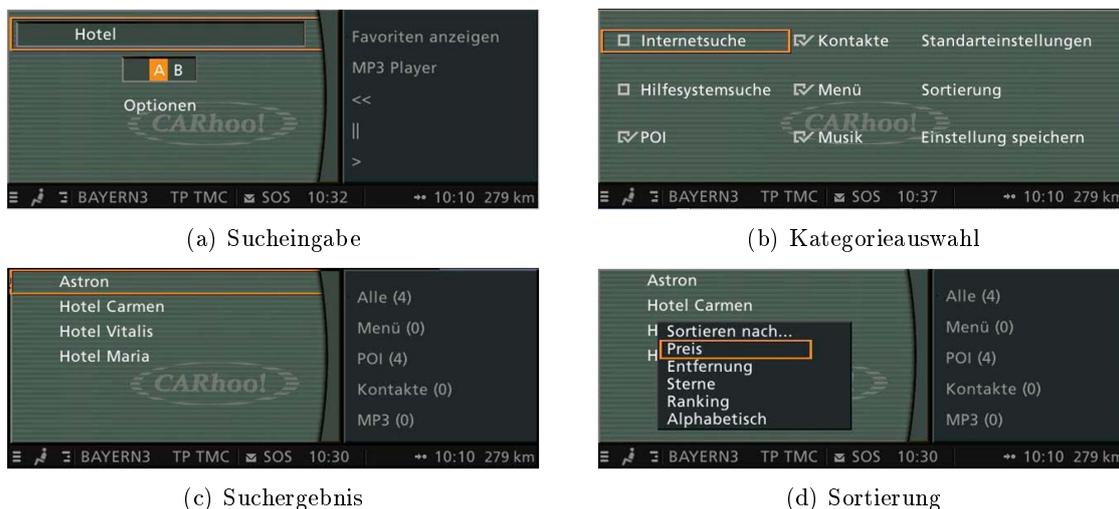


Abbildung 4.15: CARhoo!

- *Dreh-/Drücksteller* (DDS):

Die alphanumerische Eingabe wird analog zum iDrive-Speller realisiert. Die Funktions-

<sup>4</sup>zu deut. Aussehen und Handhabung, „Anfühlen“, Anmutung

weise ist identisch zu der Texteingabe von Navigationszielen. Durch Drehung des Controllers wird Buchstabe für Buchstabe aus der horizontalen, alphabetischen Liste ausgewählt und durch Drücken übernommen. Ebenfalls können alle weiteren Funktionen wie das Navigieren in den Menüs über den Dreh-/Drücksteller bedient werden.

- *Texteingabe mittels T9 (NUM)*:  
Neben der Texteingabe über den Dreh-/Drücksteller ist auch die Eingabe über einen Ziffernblock möglich. Diese Eingabemöglichkeit ist von Mobiltelefonen bekannt und bietet geübten Nutzern eine schnelle Eingabe von Text. Des Weiteren ermöglichen die Ziffern eine direkte Anwahl von Einträgen der Ergebnisliste.
- *Spracherkenner*:  
Mit dem Spracherkenner ist sowohl die Eingabe von Suchbegriffen als auch die Menübedienung möglich. Häufig auftretende Suchbegriffe wie auch Kommandos zur Interfacesteuerung sind im Vokabular und der Grammatik des Spracherkenners hinterlegt.

Durch Verwendung der vorgestellten Eingabegeräte können nun Suchbegriffe in die Suchzeile eingegeben werden (siehe Abbildung 4.15). Vor dem Ausführen der Suchanfrage können über das Feld „Optionen“ die Kategorien (Internet, Hilfe, POI, Kontakte, Musik oder Menü) eingeschränkt werden. Nach dem Starten der Suchanfrage erfolgt die Ausgabe des Ergebnisses in einer vertikalen Trefferliste (siehe Abbildung 4.15(c)). Einen schnellen Überblick über die Treffer der einzelnen Kategorien findet der Nutzer im rechten Teil des Splitscreens. Als Nächstes kann die Trefferliste abhängig vom Kontext z. B. nach Entfernung, Preis, Sterne, Ranking sortiert werden. Hat der Fahrer einen Eintrag selektiert, so erscheint ein Detailfenster mit weiteren Informationen. Durch die Anwahl eines Eintrags aus dem Detailfenster wird der Fahrer direkt zu der verknüpften Fahrzeugfunktion weitergeleitet. Wählt der Fahrer z. B. eine Adresse aus, so kann diese zur Berechnung der Route direkt in das Navigationsmenü übernommen werden. Bei MP3s z. B. wechselt das Interface zum Musik-Player. Auch können sich in der Trefferliste Funktionen aus dem Menü finden. Nach Anwahl erfolgt ein unmittelbarer Einsprung in das verknüpfte Menü.

#### 4.5.2 Suchinterface IMSENIO

Analog zum CARhoo! erfolgt bei IMSENIO ebenfalls die Einbindung von Datenbanken und Menüfunktionen sowie die Einbindung der Bedienungsanleitung. Das System IMSENIO unterstützt neben dem Dreh-/Drücksteller und Sprache, auch die Bedienung über einen Touchscreen. Abhängig von der Aufgabe und der Situation sollen hier dem Fahrer geeignete Modalitäten zur Verfügung gestellt werden:

- Während der Fahrt sollten die Eingaben über den *Spracherkenner* getätigt werden. Als Besonderheit unterstützt dieses System natürlich-sprachliche Suchanfragen (siehe Abbildung 4.16). Hierfür wurde eine geeignete Grammatik mit Vokabular erstellt.
- Bei der Benutzung im Stand oder bei Eingaben von Wörtern oder Begriffen, welche nicht im Sprachwortschatz enthalten sind, kommt hier eine On-Screen-Keyboard (OSK) zum Einsatz. Die Buchstaben sind hier für eine schnelle Orientierung in alphabetischer Reihenfolge angeordnet, anstatt des QWERTY-Layouts.
- Ein Dreh-/Drücksteller (DDS), der die Navigation im Menü, wie z. B. das Scrollen und Auswahl in einer Liste ermöglicht, wurde integriert.

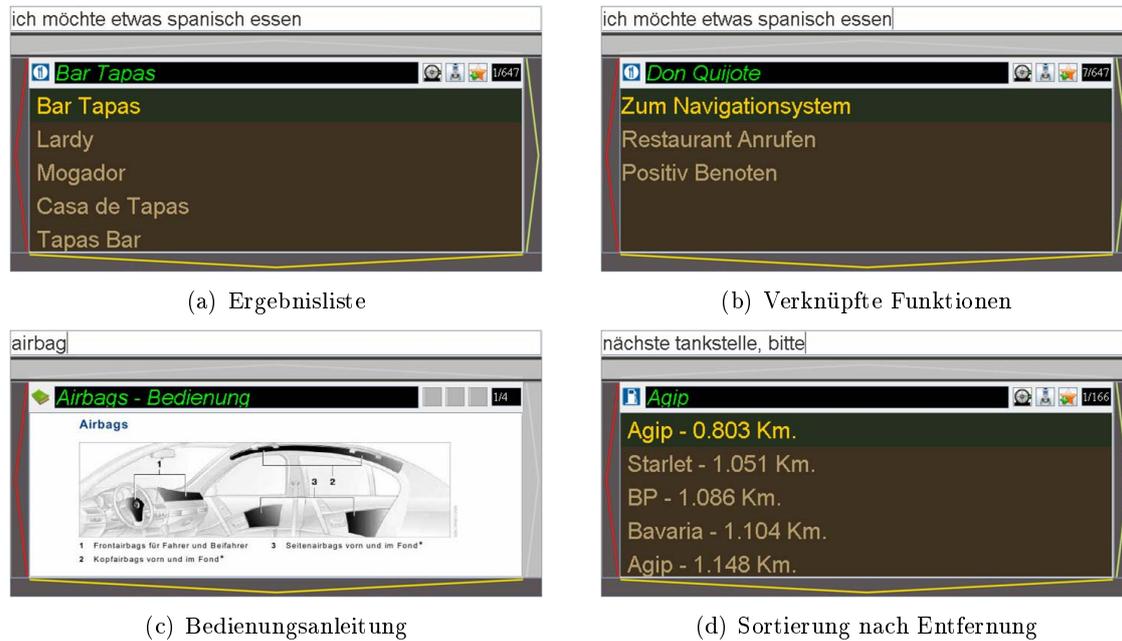


Abbildung 4.16: IMSENIO

Das Interface ist, wie man in Abbildung 4.16 sehen kann, sehr schlicht aufgebaut. Direkt unterhalb der Eingabezeile wird der angewählte Treffer aus der Liste nochmals redundant hervorgehoben und in Abhängigkeit des Kontextes mit drei Direktfunktionen versehen. Bei einem Adressbucheintrag werden beispielsweise die Funktionen „Nummer anrufen, Daten ins Navigationssystem übernehmen und zu Favoriten hinzufügen“ angeboten. Darunter wird dann die Trefferliste angezeigt. Im Normalfall sind dies fünf Einträge, bei höheren Geschwindigkeiten wird die Anzahl auf zwei reduziert, um den Fahrer nicht abzulenken.

Für die Steuerung des Interfaces benutzt man entweder die an den Rändern des Bildschirms angebrachten Touchbuttons oder den Dreh-/Drücksteller. Die oberen und unteren angebrachten Pfeilbuttons am Bildschirmrand ermöglichen ein Listenscrollen. Der rechte Pfeilbutton ermöglicht ein Eintauchen in Unterebenen wie dem Detailfenster, der linke Pfeilbutton fungiert als Zurücktaste. Im Detailfenster werden auch die mit dem Eintrag verknüpften Funktionen angezeigt.

Natürlich-sprachliche oder taktile Eingaben des Fahrers werden zunächst vorverarbeitet. Über ein Garbage-Modell, das Füllwörter entfernt, und die Verwendung von Stemming (Reduktion von Wörter auf die Grundform) wird die Informationsmenge zunächst reduziert und anschließend über den *TVSM-Algorithmus* (siehe Abschnitt 4.3.5) interpretiert. Hierbei werden auch unscharfe Eingaben durch eine Kombination aus Levenshtein- und Soundex-Algorithmus berücksichtigt.

### 4.5.3 Suchinterface CARpedia

Dieses Interface wurde entwickelt, um neben der reinen Suche auch weitere Möglichkeiten bei der Suche aufzuzeigen. Die sinnvolle Kombination der Suche aus Filter- und Sortierfunktionen mit kategoriebasiertem Zugriff zeigen hier neue Wege auf. Daraus ergibt sich ein neuer Zugriff auf die Daten durch *Exploration* (siehe Abschnitt 4.1.4). Die Vorteile dieser beiden Dialogstra-

tegien sollen in diesem Suchinterface vereint werden. Es ermöglicht sowohl auf verschiedene Daten, wie Adressbuch, POIs und MP3s zuzugreifen, als auch auf die Fahrzeug-Funktionen und Einträge des Menüs. Eine unscharfe Suche wird über die Algorithmen Levenshtein und Soundex realisiert.

CARpedia wurde in der objektorientierten Programmiersprache Java 5 entwickelt. Die Einbindung der Daten erfolgt über MySQL. Es wurde jeweils eine Datenbank für das Adressbuch (95 Einträge), für MP3s (261 Einträge), für Fahrzeug-Funktionen (157 Einträge), für Tankstellen (299 Einträge), für Unterkünfte (11 Einträge), für Einkaufsmöglichkeiten (20 Einträge) und für Restaurants (278 Einträge) eingebunden.

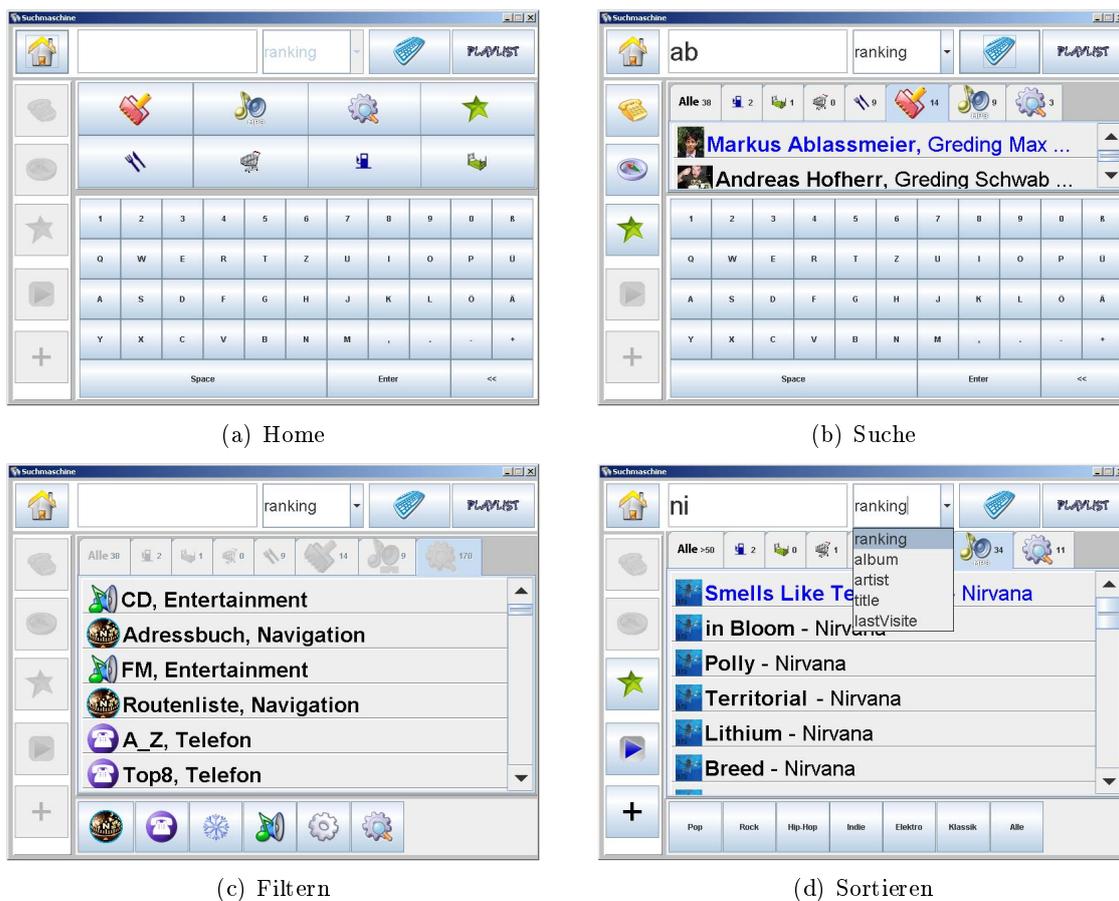
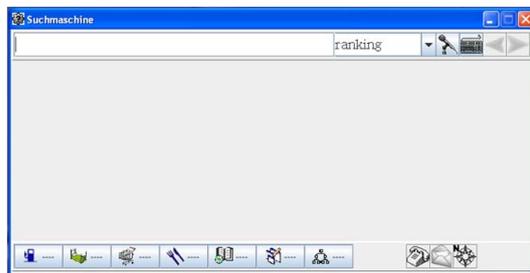


Abbildung 4.17: CARpedia

Das Interface von CARpedia wurde zur Bedienung über einen Touchscreens (10 Zoll) in der Mittelkonsole entwickelt. Bereits am Startbildschirm können die beiden Zugriffsstrategien *Suche* und *Exploration* verwendet werden (siehe Abbildung 4.17). Die oberste Zeile ermöglicht die Eingabe von Suchbegriffen. Neben dem Home-Button (oben links) folgt das Texteingabefenster, danach die Sortierfunktion über Drop-Down-Liste (als Default ist Ranking gewählt) und ein Button zum Ein- und Ausblenden des Keyboards. In der nächsten Zeile wird der explorative Zugriff über Kategorie-Buttons bereitgestellt. Dies sind Adressbuch, MP3s, Favoriten, Restaurants, Geschäfte, Tankstellen, Unterkünfte und Fahrzeug-/Menüfunktionen. Durch Auswahl eines Kategorie-Buttons springt man direkt in die entsprechende Kategorieliste. Danach

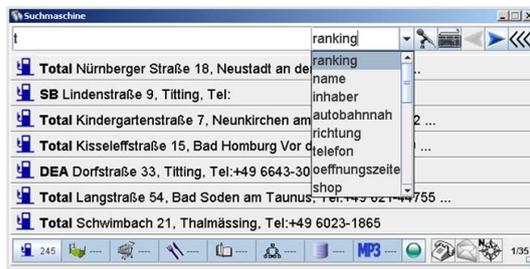
erfolgt die Anzeige der einzelnen Kategorien als Karteireiter und darunter die Trefferliste mit der Anzahl der Treffer. Im Startbildschirm befindet sich im unteren Bildschirmbereich ein On-Screen-Keyboard (OSK, Layout ist QWERTY) zur Eingabe der Suchbegriffe. Wird das OSK ausgeblendet so wird die Trefferanzahl von zwei (verkürzte Trefferanzeige) auf sechs (volle Trefferanzeige) Treffer vergrößert. Am linken Rand finden sich die dem Listenobjekt zugeordneten Funktionen (Aktionsbuttons), die je nach Verfügbarkeit der Daten als aktiv erscheinen. Die Eingabe eines Suchbegriffs gestaltet sich folgendermaßen: Hierzu wird über das OSK der Suchstring oder Teile davon eingegeben. Die Suche wird nach der Betätigung des Return-Buttons oder automatisch nach einem Timeout gestartet. Die Anzeige der Treffer erfolgt zunächst in der verkürzten Variante. Außerdem werden die Karteireiter eingeblendet. Die Anzeige aller Treffer erfolgt über den Karteireiter „Alle“. Die Art des Eintrags wird über ein Icon oder ein Bild des Treffers zusätzlich visuell repräsentiert. Ein Scrollen der Liste ist über die rechte Scrollleiste möglich. Bei Bedarf kann nun der Suchstring weiter ergänzt werden.



(a) Home



(b) Ergebnisliste



(c) Filtern



(d) Selektion

Abbildung 4.18: CARsearch

Um den Nutzer nun beim Auffinden des gewünschten Ergebnisses zu unterstützen, stehen auch *Filter-* und *Sortierfunktionen* zur Verfügung. Die Sortierung kann über das Drop-Down Menü neben dem Texteingabefeld verändert werden. In Abhängigkeit von der Kategorie stehen verschiedene Sortiermethoden zur Verfügung (z. B. Ranking, alphabetisch, zuletzt gewählt, Entfernung, usw.). Die Einschränkung des Ergebnisses über die Filterung erfolgt mehrstufig. Zuerst kann die gesamte Treffermenge durch Anwahl eines Karteireiters nach Kategorie eingeschränkt werden. Des Weiteren kann auch die Filterung über Buttons am unteren Rand vorgenommen werden. Handelt es sich um MP3s, so kann z. B. Einschränkungen über Genre wie Pop, Rock, Hip-Hop, Indie, usw. vorgenommen werden. Wurde eine Filterung ausgeführt können bei Bedarf weitere (Sub-)Filterungen getätigt werden und es erscheinen hier neue Filter-Buttons. Durch diese Kategorie- und Filtermethoden kann der Fahrer alternativ einen explorativen Zugriff wählen, neben der Einschränkung durch weitere Suchstrings. Wurde

der gewünschte Eintrag aufgefunden, können aktive Funktionen der linken Leiste ausgeführt werden oder weitere Details angefordert werden. Über den explorativen Zugriff können auch sämtliche Fahrzeug-Funktionen sortiert und gefiltert werden. In Abbildung 4.18 ist eine platzoptimierte Variante dargestellt.

## 4.6 Evaluierung der Suchinterfaces

In diesem Abschnitt soll geprüft werden inwieweit sich die erläuterten Suchstrategien zum Umgang mit großen Datenmengen im Fahrzeug eignen und in welcher Form die Interfaces gestaltet sein müssen. Hierzu werden die zuvor erläuterten Fahrzeug-Interfaces herangezogen und in mehreren Nutzerstudien untersucht.

### 4.6.1 Untersuchung Suchinterface CARhoo!

Zunächst wurde das Interface CARhoo! (siehe Abschnitt 4.5.1) in einer Usability-Studie bewertet. Ziel der Untersuchung war es dabei, die generelle Akzeptanz von Suchmaschinen im Fahrzeug zu evaluieren. Des Weiteren soll festgestellt werden, welche Anforderungen bei der Bedienung einer Fahrzeug-Suchmaschine entstehen. Eigenen sich die verwendeten Eingabemodalitäten? Wie soll die Ausgabe der Ergebnisse gestaltet sein? Dazu wird das Suchmaschinen-Interface mit den klassischen hierarchischen FIS verglichen.

#### Versuchsaufbau und -durchführung

Zur Untersuchung des CARhoo!-Konzeptes wurde der im Abschnitt 2.5.1 vorgestellte Fahr Simulator verwendet. Dabei diente die FERMUS-Sichtsimation als Fahraufgabe für die Versuchspersonen, die aus einer einfachen Spurhalteaufgabe besteht. Dieser Test wurde im Rahmen des Projekts FERMUS entwickelt [FER08]. In [MASL02] wird die genaue Funktionsweise der Fahraufgabe beschrieben. An dieser Studie nahmen insgesamt 14 Versuchspersonen im Alter von 19 bis 48 Jahren teil (Durchschnittsalter war 28 Jahre). In der Sitzkiste war ein 10 Zoll Bildschirm in der Mittelkonsole angebracht. Die taktilen Eingabemodalitäten, Dreh-/Drücksteller und Ziffernblock, waren in der Mittelkonsole nahe des Gangwahlhebels integriert. Die Eingabe von Suchbegriffen konnte auch über Sprachkommandos getätigt werden.

Die Testpersonen hatten in mehreren Durchläufen Aufgaben zur Zieleingabe (A1: z. B. Auswahl eines Orts- und Straßennamens), Eingabe von Telefonnummern (A2: z. B. aus einer Telefonliste) und Auswahl von Fahrzeugfunktionen (A3: z. B. Änderung der Klangeinstellungen) auszuführen. Dabei mussten die Versuchspersonen die Aufgaben sowohl mit der CARhoo!-Suchmaschine als auch mit dem Referenzsystem durchzuführen. Als Referenz diente ein hierarchisch gegliedertes FIS. Nach der Durchführung der Aufgaben im Simulator wurden die Versuchspersonen mittels Fragebogen hinsichtlich Handhabung, Akzeptanz, Nutzen und Ablenkung vom Versuchsleiter interviewt.

#### Ergebnisse

Betrachtet man die Bediendauern, so zeigt sich, dass durch die Verwendung der Suchmaschine wesentliche Zeitersparnisse erreicht werden konnten. Die Aufgabe A1 (Zieleingabe) konnte mit einer durchschnittlichen Zeitersparnis von 45,2% gelöst werden. Bei der Aufgabe A2 (Eingabe von Telefonnummern) wurde eine durchschnittliche Zeitersparnis von 8,6% im Vergleich

zum Menü erreicht. Die Auswahl von Fahrzeugfunktionen zeigte unterschiedliche Tendenzen: fanden sich die Fahrzeugfunktionen sehr tief in der Menüstruktur verborgen, so konnten deutliche Zeitersparnisse durch die Verwendung der Suchmaschine erreicht werden. Hinsichtlich der mittleren Spurabweichung traten bei der Verwendung beider Systeme keine großen Auffälligkeiten auf.

Bei der subjektiven Befragung zeigte die Suchmaschine eine hohe Zustimmung und Zufriedenheit bei den Versuchspersonen: Der Aussage „Diese Funktion ist sinnvoll“, stimmten 12 von 14 Personen (entspricht 85,7 %) zu, und sie erreichte eine durchschnittliche Bewertung von 1,14 (mit 1 - ich stimme zu; 4 - ich stimme nicht zu).

Bei der Befragung zu den unterschiedlichen Eingabemodalitäten erzielte die Spracheingabe die beste Bewertung mit einem Durchschnittswert von 1,36. Eine geringere Akzeptanz erzielte der Ziffernblock mit 2,79 gefolgt vom Dreh-/Drücksteller mit 3,71 (1 - sehr gut; 6 - unbefriedigend). Für das Scrollen und Auswählen in der Trefferliste bevorzugten jedoch die meisten Versuchspersonen den Dreh-/Drücksteller. Alle Versuchspersonen waren sich einig, dass die Sprache für die Eingabe von Suchbegriffen besonders komfortabel ist. 9 von 14 konnten sich auch die Verwendung eines Ziffernblocks vorstellen. Darüber hinaus äußerten 5 Personen den Wunsch lieber einen Touchscreen oder ein Touchpad verwenden zu wollen.

Die POI-Suchergebnisse sollten dabei in folgender Sortierung ausgegeben werden: 7 von 14 wünschten sich ein allgemeines Ranking, das neben Kontextgrößen auch persönliche Präferenzen mit einbeziehen darf; 2 konnten sich auch den Einbezug des Suchverhaltens anderer Nutzer in das Ranking vorstellen. 4 von 14 wollten die POI-Ergebnisse nach Zielentfernung sortiert und 3 der Versuchspersonen präferierten eine alphabetische Reihenfolge. Die Sortierung nach Entfernungszeit wurde des Weiteren von 4 Versuchspersonen vorgeschlagen.

### Zusammenfassung

Insgesamt erreichte der Suchmaschinen-Ansatz eine hohe subjektive Akzeptanz durch die Testpersonen (siehe auch [AMG<sup>+</sup>05, APR06]). Die Vorteile der Suchmaschine lagen vor allem in der Zeitersparnis bei der Auswahl aus großen Daten. Bei der Bedienung waren keine größeren Unregelmäßigkeiten in der Fahrleistung erkennbar. Zur Eingabe der Suchbegriffe präferierten die Versuchspersonen die Sprache, die Auswahl in der Trefferliste wollten sie jedoch lieber über den Dreh-/Drücksteller vornehmen. Der Ziffernblock fand dabei bedingten Zuspruch. Als Weiteres wünschten die Versuchspersonen verschiedene Sortiermethoden. Neben einem Ranking wünschten die Versuchspersonen die Ausgabe alphabetisch und nach Entfernung.

#### 4.6.2 Untersuchung Suchinterface CARpedia

Ziel dieser Untersuchung ist die Ermittlung der Nutzerakzeptanz bei der Bedienung von CARpedia (siehe Abschnitt 4.5.3). Dabei stellt sich als zentrale Fragestellung inwiefern die Verwendung der umfangreichen Sortierungs- und Filtermethoden auf Nutzerakzeptanz stoßen.

#### Versuchsaufbau und -durchführung

Die Untersuchung wurde im Fahr Simulator durchgeführt und als Sichtsimulation wurde der LCT aus Abschnitt 2.5.1 verwendet. Neben dem Think Aloud (Lautes Denken der Nutzer bei der Aufgabenbearbeitung) und einem Beobachtungsprotokoll wurde ein eigener Fragebogen erstellt und der AttrakDiff-Test durchgeführt. Die Untersuchung des CARpedias unterteilte sich in folgende drei Aufgabenteile: Zunächst erfolgte ein Kennenlernen des CARpedia-Interfaces

mit einem Think-Aloud-Protokoll. Des Weiteren waren 5 Aufgaben erst ohne Fahraufgabe und anschließend mit Fahraufgabe durchzuführen. Als Referenz diente ein menübasiertes FIS. Neben dem Nutzerverhalten wurden die Bedienzeiten und die Spurhaltegüte erfasst. An der Studie nahmen insgesamt 16 Testpersonen (davon 13 männlich, 3 weiblich) teil, davon 13 Studenten. Das Durchschnittsalter war 25,3 Jahre und reichte von 22 bis 28 Jahre.

## Ergebnisse

Nachfolgend werden die Ergebnisse der Untersuchung des CARpedia-Interfaces aufgeführt.

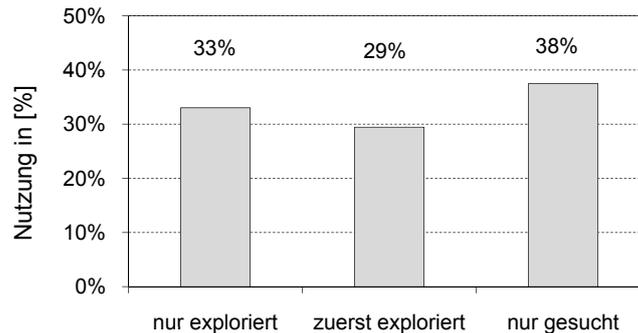


Abbildung 4.19: Verwendung von Suche und Exploration über alle Aufgaben

**Subjektive Ergebnisse:** Subjektiv wurde CARpedia sehr positiv bewertet. 14 der 16 Versuchspersonen sahen das Suchinterface für den Einsatz im Fahrzeug als gut geeignet und 15 Testpersonen stufte das entwickelte Interface als intuitiv ein. Eine große Akzeptanz herrschte bei den integrierten Filterfunktionen. Die Sortierfunktionen wurden von der überwiegenden Anzahl der Probanden ebenfalls akzeptiert.

Betrachtet man die Ergebnisse des Attrakdiff-Fragebogens (siehe Abbildung 4.20), so lagen die Stärken des CARpedias vor allem in einer höheren pragmatischen Qualität und gesteigerter Attraktivität.

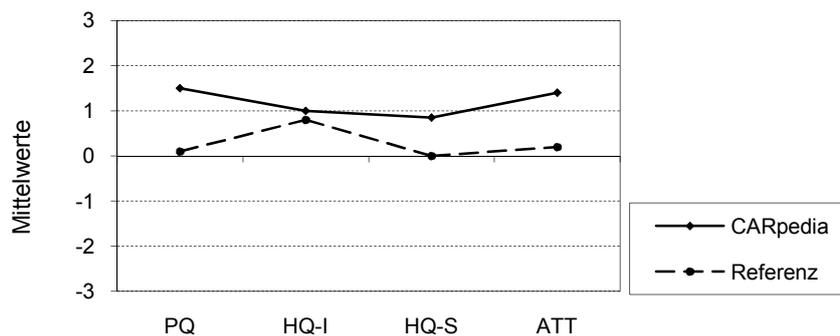


Abbildung 4.20: Ergebnisse des Attrakdiff-Fragebogens von CARpedia

**Objektive Ergebnisse:** Die Verwendung der zwei Zugriffsmethoden Exploration und Sucheingabe erfolgte in drei Mustern: Gemittelt wurden 38% der Aufgaben ausschließlich mit der Sucheingabe gelöst (siehe Abbildung 4.19). Bei 29% der Aufgaben wurde von anfänglicher

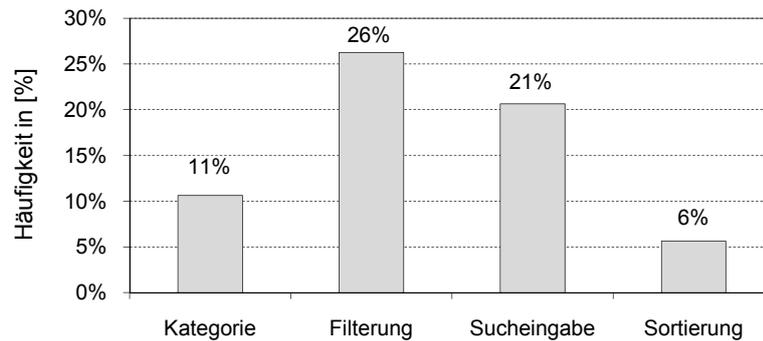


Abbildung 4.21: Verwendung der Features der Suchmaschine

Exploration auf die Sucheingabe gewechselt. 33 % der Aufgaben begannen die Versuchspersonen mit der Suche und schränkten anschließend das Ergebnis über Filterung ein.

Die mittlere Anzahl der eingegebenen Buchstaben bei der Sucheingabe war 4,5. Insgesamt wurden die Filterbuttons am Häufigsten (26 %) verwendet, gefolgt von den Tasten des OSK zur alphanumerischen Eingabe (21 %), den Kategorie-Reitern (11 %) und der Sortierungsfunktion (6 %) (siehe Abbildung 4.21). Die häufige Verwendung der Filterbuttons deutet auf die hohe Akzeptanz der nachgelagerten Filterung hin.

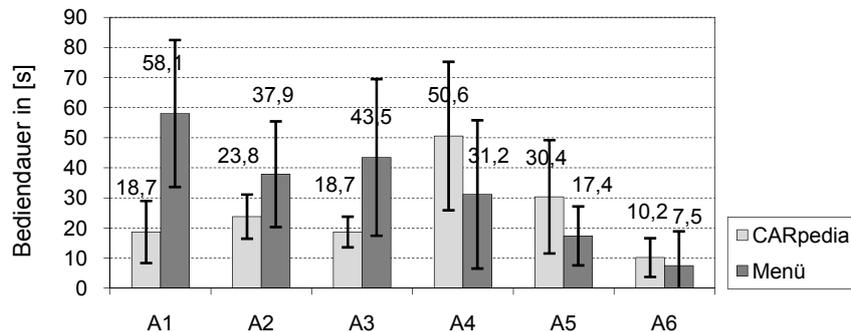


Abbildung 4.22: Mittlere Bediendauern für die Aufgaben

Bei den Bediendauern zeigten sich Zeitersparnisse bei CARpedia, aber auch das menübasierte Referenzsystem lag bei manchen Aufgabentypen vorne (siehe Abbildung 4.22). Insgesamt waren hier die Zeitunterschiede signifikant (Aufgaben 1 bis 3 hoch signifikant). Abhängig vom Aufgabentyp, wie bei der Datensuche und bei versteckten Fahrzeug-Funktionen, zeigte die CARpedia-Suchmaschine ihre Stärken. Auch bei der mittleren Spurhaltegröße ergaben sich leicht wechselnde Tendenzen, jedoch keine signifikanten Werte (siehe Abbildung 4.23).

### Zusammenfassung

Insgesamt wurde die CARpedia-Suchmaschine sehr positiv bewertet. Das System erwies sich entsprechend den AttrakDiff-Ergebnissen als sehr pragmatisch, sprich handlungsorientiert. Die Bedienzeiten konnten bei manchen Aufgaben deutlich gesenkt werden. Die Suchfunktion wurde für sämtliche Daten und Funktionen gewünscht und genutzt. Am Häufigsten wurde die nachgelagerte Filterung verwendet. Auch die Verwendung der Kategorie-Buttons und der Sor-

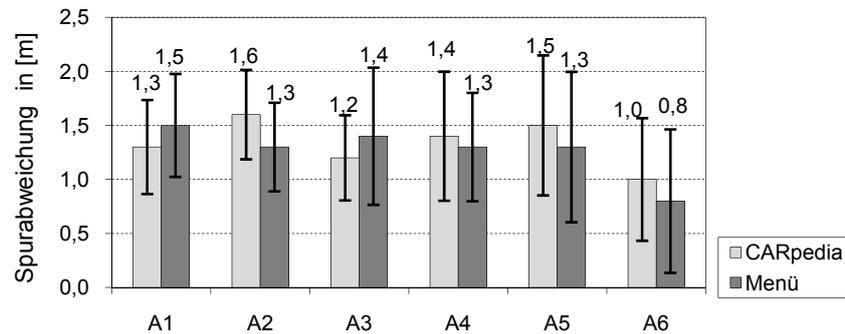


Abbildung 4.23: Mittlere Spurabweichung für die Aufgaben

tierfunktionen fand großen Zuspruch. Daraus lässt sich folgern, dass gerade der explorative Zugriff große Akzeptanz bei den Versuchspersonen bekam. Der explorative Ansatz wird daher im Kapitel 5 in Form des objektorientierten, vernetzten Datenbrowsers aufgegriffen und weiterentwickelt.

### 4.6.3 Untersuchung alphanumerischer Eingabeverfahren

Damit der Fahrer einen Nutzen aus der wachsenden Datenmengen im Fahrzeug ziehen kann, ist eine einfache und robuste Eingabemethode für Alphanumerik bei der Suche erforderlich. In dieser Untersuchung wurden Einflüsse der verschiedenen Interaktionsmethoden auf die resultierenden Bediendauern und Fahrleistungen analysiert. Zusätzlich wurden die subjektiven Aspekte hinsichtlich des Eingabeverfahrens ausgewertet.

#### Versuchsaufbau und -durchführung

Folgende vier Eingabegeräte wurden untersucht: ein Dreh-/Drücksteller (DDS), eine Touchpad mit Handschrifterkennung (HWR), ein numerischer Ziffernblock mit T9-Texteingabe (NUM) und ein Touchscreen mit Bildschirmtastatur (On-Screen-Keyboard (OSK), QWERTY-Layout). Die 15 Teilnehmer (Durchschnittsalter 38 Jahre), deren Alter von 22 bis 50 reichte, hatten als Nebenaufgabe mehrere Adresseingaben während der Fahrt (Ort, Straße, Hausnummer) im Simulator zu tätigen. Als Sichtsimulation wurde der Lane Change Task (LCT) im ehemaligen Fahrsimulator des Lehrstuhls für Mensch-Maschine-Kommunikation (siehe Abschnitt 2.5.1) verwendet. Zusätzlich zu Videodaten wurden die Spurabweichungen und die Bearbeitungsdauern während des Tests aufgezeichnet. Die Analyse subjektiver Kriterien erfolgte auf Basis eines strukturierten Fragebogens mit Verwendung eines semantischen Differentials.

#### Objektive Ergebnisse

Nun werden die Ergebnisse der Spurabweichungen und Bearbeitungsdauern präsentiert.

**Spurabweichung:** Die Fahrleistung wurde durch die Bestimmung der mittleren Spurabweichung ermittelt. Die Abbildung 4.24 fasst die Resultate zusammen und stellt die mittlere Spurabweichung beim Durchführen der kompletten Aufgabensequenzen dar. Bei der Eingabe über das NUM ergibt sich die größte Spurabweichung und somit die schlechteste Fahrleistung. Dies kann unterschiedlichen Aspekten zugeschrieben werden. Eine wichtige Tatsache

ist, dass der Benutzer zur Mittelkonsole nach unten blicken musste, um das Tastenfeld zu bedienen, was die Spurabweichung negativ beeinflusste. Das OSK und das HWR schnitten diesbezüglich besser ab und zeigten mittlere Spurabweichungen von 1,45 m bzw. 1,46 m. Das HWR konnte teilweise blind bedient werden, was die Blickabwendung stark verringerte. Lediglich Kontrollblicke zum CID waren notwendig, anstelle von konstantem Sichtkontakt mit dem Eingabegerät. Der DDS war hier im Mittelfeld.

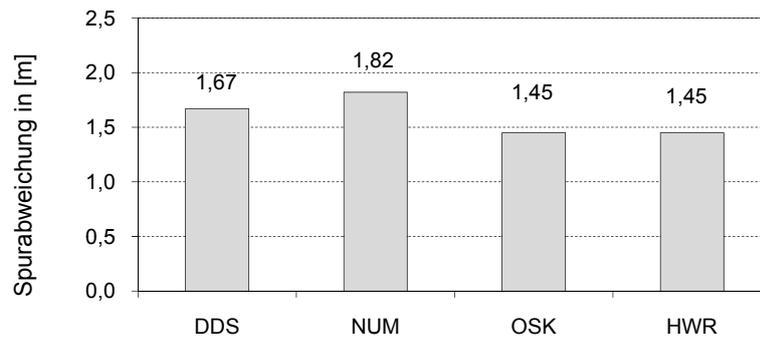


Abbildung 4.24: Mittlere Spurabweichung der einzelnen Bedienelemente

**Bearbeitungsdauer:** Ein weiterer Aspekt, der während des Tests betrachtet wurde, war die tatsächliche Zeit, die jeder Benutzer für die Eingabe aufwendete. Die Resultate können der Abbildung 4.25 entnommen werden. Die Bearbeitungsdauer (auch *Total Task Time*, TTT genannt) ist definiert als die Zeit, die der Benutzer für das Erledigen der gesamten Aufgabe benötigt. Hier gibt es große Abweichungen zwischen den Systemen. Das OSK zeigte sich als schnelle Eingabemethode, da nur ein Tastendruck pro Buchstabe notwendig ist.

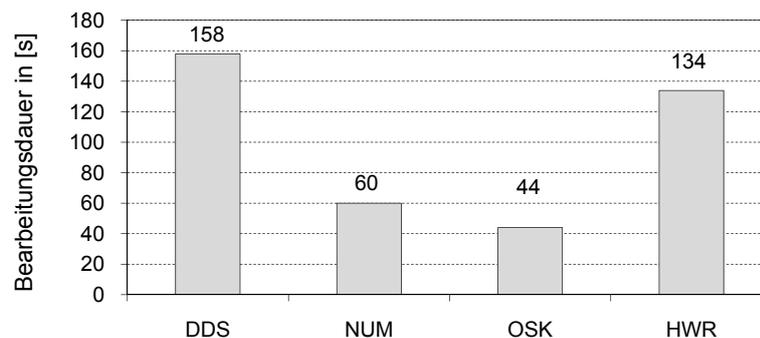


Abbildung 4.25: Mittlere Aufgabendauer bei Verwendung der einzelnen Bedienelemente

Dagegen hängt die Aufgabendauer mit dem NUM stark von der Vorerfahrung ab, die der Benutzer mit dem System hatte. Während einige Benutzer große Schwierigkeiten hatten, konnten die geübten Nutzer den Tastaturblock sehr schnell bedienen. Der Unterschied basiert hauptsächlich auf der Erfahrung mit dem vorbestimmten Texteingabesystem und dem Wissen über das Tastatur-Layout. Während die meisten Benutzer normalerweise ihre Mobiltelefone mit dem Daumen benutzen, musste das NUM z. B. mit dem Zeigefinger bedient werden, da die Einrichtung in die Mittelkonsole örtlich festgelegt war.

EINGABEGERÄT	DDS	NUM	OSK	HWR
<i>Fehlerrate</i>	4,9 %	6,7 %	4,9 %	23,5 %

Tabelle 4.4: Eingabefehler

Währenddessen war die Handhabung des HWR um einiges zeitaufwendiger, dies muss jedoch relativiert betrachtet werden. Zwar erforderte das Schreiben des Buchstabens auf das Touchpad eine gewisse Zeit, jedoch stellte vor allem die Erkennerleistung des Handschrifterkenners einen wesentlichen Störfaktor dar (siehe Tabelle 4.4). Auch der Gebrauch des Spellers erfordert eine nicht unwesentliche Zeitmenge. Hier muss der DDS zunächst zum gewünschten Buchstaben gedreht werden und anschließend zu Selektion gedrückt werden.

### Subjektive Ergebnisse

Nachfolgend werden die subjektiven Ergebnisse erörtert.

ADJEKTIV-PAARE (Skala 1..7)	DDS	NUM	OSK	HWR
<i>sicher – riskant</i>	3,8	3,7	3,7	2,7
<i>entspannend – anstrengend</i>	5,1	4,3	3,3	3,6
<i>einfach – schwierig</i>	3,9	2,9	2,4	3,0

Tabelle 4.5: Bewertung des semantischen Differentials (Mittelwerte)

**Befragung zu den Modalitäten:** Einzelne Ergebnisse des semantischen Differentials können in Tabelle 4.5 betrachtet werden. Hinsichtlich der Sicherheit stimmten die meisten Benutzer darin überein, dass alphanumerische Eingaben beim Fahren durchaus riskant sein können. Als sicherstes Bedienelement wurde das HWR betrachtet, hauptsächlich weil es kaum Blickabwendung erforderte und weitgehend blind bedient werden konnte. Der Rest der Bedienelemente wurde als riskant eingestuft, da alle Blickzuwendung zu den eingegebenen Daten und teilweise zum Bedienelement erforderten. Betrachtet man die Beanspruchung, wurde keines der Bedienelemente als entspannend eingestuft, jedoch erhielten das OSK und das HWR hier die beste Einstufung. Der DDS und NUM wurden als die anstrengendsten Eingabemethoden bewertet. Bezüglich einer einfachen Bedienbarkeit schnitt das OSK am Besten ab.

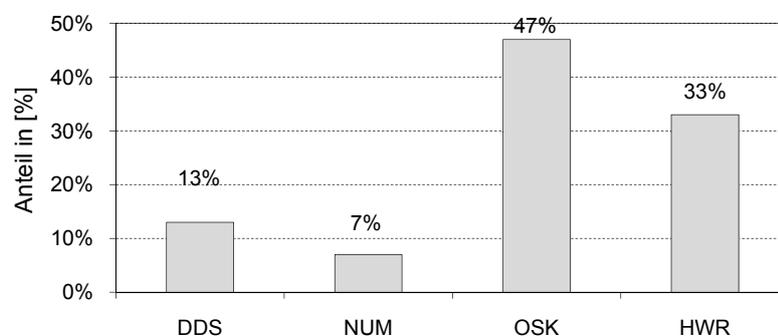


Abbildung 4.26: Befragung zur Modalitätenpräferenz

**Allgemeine Präferenz:** Die Frage „Welche Eingabemethode bevorzugen Sie im Auto?“ wurde wie folgt beantwortet (siehe Abbildung 4.26): Die subjektiven Daten betreffend bevorzugten 47 % der Benutzer das OSK als Eingabegerät während der Fahrt, gefolgt vom HWR mit 33 % und dem DDS mit 13 %. Lediglich 7 % wünschten das NUM.

### Schlussfolgerungen

Aus den Ergebnissen zur Untersuchung der Alphanumerik können folgende Schlussfolgerungen in Form von Empfehlungen ausgesprochen werden (siehe auch [AR06]):

- **Unterschiede in Zeitdauer und Fahrperformance bei alphanumerischen Eingaben:**

Hinsichtlich der Bearbeitungsdauer hatte das On-Screen-Keyboard (OSK) am Besten abgeschnitten. Dateneingaben mit Handschrifterkennung (HWR) und dem Dreh-/Drücksteller (DDS) waren dagegen relativ zeitaufwendig. Die großen Bearbeitungsdauern bei dem HWR waren jedoch überwiegend auf die schlechte Erkennerrate zurückzuführen. Hinsichtlich der Spurabweichung schnitten HWR und OSK am Besten ab, gerade das HWR ermöglichte den Versuchspersonen eine weitgehend blinde Bedienung. Das NUM erforderte für ungeübte Nutzer längere Blickabwendungen, was die Spurhaltung hier negativ beeinflusste. Die Versuchspersonen bewerteten das HWR als die sicherste Eingabemethode. Das On-Screen-Keyboard (OSK) wurde von den Versuchspersonen als einfachstes Eingabeelement betrachtet. Bei allen Eingabemethoden forderten die Nutzer unbedingt akustisches Feedback.

- **Zwei parallele Eingabegeräte für Alphanumerik gewünscht:**

Als alphanumerische Eingaben wünschten sich ein Großteil der Versuchspersonen zwei parallel verfügbare Eingabegeräte, die sie abhängig von der Situation wählen können. In einem stehenden Auto bevorzugten die Benutzer den Gebrauch des OSK, da es ein einfaches, schnelles und zuverlässiges Eingabegerät ist. Während der tatsächlichen Fahrt präferierten die Personen den Gebrauch des Touchpads mit Handschrifterkennung oder als Fallback den Dreh-/Drücksteller.

Zusammenfassend kann auch gemäß [Ras04, Bub08b] gefolgert werden, dass für die Alphanumerik simultane Eingabeverfahren wie das OSK und das HWR im Fahrzeug bevorzugt werden. Sie schnitten hinsichtlich Ablenkungswirkung und Bearbeitungsdauern größtenteils besser ab und boten hohe Akzeptanz bei den Nutzern. Dagegen zeigten sequenzielle Eingabemethoden wie dem DDS und auch das NUM wesentlich geringere Nutzerakzeptanzen und zum Teil längere Bearbeitungsdauern.

## Kapitel 5

# Objektorientiertes, vernetztes Datenbrowsing

In diesem Kapitel wird eine Dialogstrategie erarbeitet, die den Fahrer bei der Ausführung mehrerer zusammenhängender Teilaufgaben unterstützen soll. Im Besonderen soll hier einer möglichst nahtlose Integration mobiler Endgeräte im Fahrzeug Rechnung getragen werden. Dabei muss gewährleistet werden, dass der Fahrer auf die vorliegenden Daten einfach und fahrgerecht zugreifen kann. Mit gegenwärtigen FIS sind zusammenhängende Aufgaben bisher meist nur über viele komplexe Teilschritte durchführbar, die außerdem oftmals sogar den Wechsel zwischen verschiedenen Geräten erfordern. Auf Grundlage des vorgestellten Suchmaschinen-Ansatzes aus Kapitel 4 wird nun ein erweiterter Ansatz entwickelt, der dem Fahrer gerade bei umfassenden Handlungssequenzen einen größeren Grad an Nutzerführung bereitstellt. In diesem erweiterten Ansatz wird zur nahtlosen und fahrgerechten Integration der benötigten Daten und Funktionen auf einen Browsing-Ansatz zurückgegriffen. Das Interface soll dabei auf die Bedürfnisse des Fahrers in verschiedenen Situationen ausgelegt werden.

### 5.1 Verwendete Methoden

Dieser Ansatz basiert auf einer sehr flachen datenorientierten Grundstruktur, die dem Fahrer eine gute Orientierung und Führung bietet. Die Gliederung erfolgt anhand der relevanten Daten in die Domänen Kontakte, Kalender, Route, Unterhaltung, Ereignisse und Einstellungen. Sie bieten dem Fahrer eine datenorientierte Zugangsart und ermöglichen ihm, die einzelnen Domänen zu erforschen und zu explorieren. Zur Datenexploration, im Folgenden – analog zum Internetbrowsing – auch Datenbrowsing, kurz Browsing genannt, stehen dem Fahrer verschiedene Methoden zur Verfügung. Der Einsprung in die Domänen erfolgt über ein kontextadaptives Statusmenü, das dem Fahrer einen Überblick über situativ relevante Informationen, Daten und Funktionen gibt. Innerhalb der Domänen erfolgt die Bedienung primär über Filterfunktionen, die sogenannte explizite Filterung. Bewegt sich der Nutzer zwischen den Domänen erfolgt die Berücksichtigung von Systemkontext und somit eine implizite Filterung.

#### 5.1.1 Objektorientierte Struktur

In diesem Ansatz wird auf eine objektorientierte Struktur zurückgegriffen. Beim Zugriff auf Informationen existieren zwei Vorgehensweisen: einerseits kann zuerst die Aktion oder Funk-

tion gewählt werden und anschließend das Objekt, auf das man sie anwenden möchte. Jedoch gestaltet sich der umgekehrte Weg, beginnend mit der Auswahl des Objektes und danach die mögliche Aktion als wesentlich einfacher und intuitiver. Bei einem Textdokument, das formatiert werden soll, ist dies die übliche Vorgehensweise: man markiert zuerst den Text (Objekt) und anschließend wählt man dann eine gewünschte Formatierung (Aktion). Über ein Kontextmenü werden dann hier die für das Objekt relevanten Funktionen aufgeführt. Die Objekt-Aktions-Konstruktion bietet daher folgende Vorteile:

- *Fehlerreduktion:*  
Sobald bei dem Aktion-Objekt-Stil eine Aktion ausgewählt ist, hat dies Auswirkungen auf das nächste Objekt. Nach einer Unterbrechung kann dies zu einer großen Irritation führen. Bei dem Objekt-Aktions-Stil werden die Befehle sofort ausgeführt, solange diese sich noch im Zentrum der Aufmerksamkeit befinden.
- *Bediengeschwindigkeit:*  
Bei Objekt-Aktion-Konstruktionen bleibt die Aufmerksamkeit zuerst beim Objekt und lediglich ein kleiner Wechsel ist zum Ausführen der Aktion notwendig. Dagegen muss der Nutzer bei der Aktion-Objekt-Konstruktion erst die Aufmerksamkeit weg vom Inhalt (Objekt) wenden, danach die Aktion wählen und anschließend wieder zurückspringen, um die gewünschte Änderung vorzufinden.
- *Einfache Anwendung und Umkehrbarkeit:*  
In der Aktion-Objekt-Konstruktion muss eine Abbruchfunktion zur Verfügung stehen, wenn die ausgewählte Aktion doch nicht ausgeführt werden soll. Bei der Objekt-Aktion-Konstruktion lässt sich die Markierung jederzeit wieder ändern, es ist hierfür keine Abbruchfunktion notwendig.

Generell kann man feststellen, dass ein gelungenes Interface der Reihenfolge menschlicher Gedankenschritte (siehe auch Abschnitt 2.3.5) angepasst werden muss. Das Zentrum der Aufmerksamkeit wird sonst zu großen Sprüngen gezwungen, für die möglicherweise gerade keine Ressourcen zur Verfügung stehen. So wie bei dem oben beschriebenen Beispiel der Textformatierung, wo die Aktionsreihenfolge den menschlichen Gedankenschritten entspricht, sollte auch die Aktionsreihenfolge in Infotainmentsystemen den Gedankenschritten entsprechen.

Will der Nutzer beispielsweise zu einem bestimmten Zielort geführt werden, ist es deshalb sinnvoll, ihn zu allererst nach diesem Zielort zu fragen, da dieser im Zentrum der Aufmerksamkeit des Benutzers steht. Es wäre falsch, ihn in diesem Moment mit den Routenkriterien zu konfrontieren, da er dann seine Aufmerksamkeit neu ausrichten müsste. Solche Gedankensprünge können außerdem Verwirrung und Ärger auslösen, wenn der Benutzer den Zusammenhang zwischen seiner eigentlichen Intention und der gestellten Frage nicht nachvollziehen kann. Wird die Frage dann auch noch als Behinderung empfunden und als nicht zielführend eingeschätzt, schenkt der Benutzer ihr entsprechend wenig Aufmerksamkeit, was wiederum zu Missverständnissen und erfolgloser Bedienung führt.

Es ist daher sinnvoll, die Bedienschritte vom Hauptfokus einer Aufgabe zu ihren Randbereichen hin aufzubauen, wobei sicherlich bei manchen Aufgaben Kompromisse eingegangen werden müssen: Nicht immer lässt sich das Zentrum einer Aufgabe so leicht bestimmen wie bei der Navigationseingabe. Jedes noch so einfach und übersichtlich gestaltete Menü wird den Nutzer hin und wieder zu Sprüngen zwingen.

Hinzu kommt, dass der Benutzer – falls er das Menü nicht auswendig und genau kennt – seine Aufmerksamkeit auf das Navigieren im Systemmenü richten muss, dabei jedoch die

eigentlich Aufgabe nicht aus den Augen verlieren darf, die fast zwangsläufig in den Hintergrund gedrängt wird. Die Aufgabe des MMI ist in diesem Fall, die Menünavigation so an die Aufgaben anzupassen, dass sie sich weitestgehend mit den natürlichen Lösungswegen decken. Daher wurde in diesem Browsing-Ansatz auf eine objektorientierte Nutzerführung zurückgegriffen.

### 5.1.2 Verknüpfung durch Tagging

Der Begriff *Tagging* (zu deut. Verknüpfung, Etikett oder Schlagwort) wird oftmals für unterschiedliche Zwecke verwendet. Tags werden unter anderem in Markup-Sprachen, wie zum Beispiel XML, verwendet, um Daten zu klassifizieren. Im Allgemeinen werden sogenannte Tags jedoch dafür eingesetzt, Daten mit zusätzlichen Informationen zu versehen. Bei Musikdaten werden in sogenannten ID3-Tags Interpret, Name des Musikstücks u. ä. abgelegt. Dadurch, dass einem Element ein Tag angefügt ist, kann es auch einer bestimmten Gruppe zugeordnet werden.

Auf manchen Webseiten findet man sogenannte Tag-Clouds (Wortwolken). Diese Tag-Clouds sind Wortlisten (aus Schlagworten), die meist alphabetisch und zweidimensional angeordnet sind. Jedes Wort (Tag) stellt einen Link dar. Der Schriftgrad repräsentiert die unterschiedliche Gewichtung der Tags. Je größer ein Tag dargestellt ist, desto größer ist auch seine Bedeutung. Diese kann etwa von der Häufigkeit der Aufrufe abhängen. Klickt man auf ein Tag, erscheinen alle Daten, die mit diesem Tag assoziiert sind, also alle Daten, die in dieser Gruppe enthalten sind. Auch eine explizite Suche nach bestimmten Tags ist meist möglich. In diesem Zusammenhang handelt es sich bei den Tags um das Ergebnis einer „gemeinschaftlichen Indexierung“. Dabei werden Tags durch Indexierer, die in einer öffentlichen Gruppe zusammenkommen (z. B. Nutzer einer Internetseite), ohne festgelegte Regeln an bestimmte Objekte vergeben. Die so entstehenden Tag-Sammlungen heißen *Folksonomien* (Neuschöpfung aus *Folk* und *Taxonomie*).

In [Mat04] werden die Vor- und Nachteile der Folksonomien genauer beleuchtet. Sie besitzen eine chaotische Natur, oft gibt es mehrere Tags, die das gleiche beschreiben. Aber auch in der chaotischen Art und Weise scheint ein gewisser Vorteil zu liegen. Solche von Nutzern angelegten Metadaten<sup>1</sup> können sehr schnell wachsen. In derartigen Gebilden gibt es sehr viele Wege an ein bestimmtes Ziel bzw. zu den gesuchten Daten zu gelangen. Zudem orientieren sich diese Schlagwortsammlungen am aktuell verwendeten Sprachgebrauch. Einige Internetseiten bieten als Service sogenannte „*Social Bookmarks*“ an. Der Nutzer macht hier seine Internet-Lesezeichen öffentlich und kann diese mit Tags versehen. Somit wird das Internet durch gemeinschaftliches Indexieren erschlossen. Ein Beispiel für einen solchen Dienst ist in [Mis08] zu finden.

Ein großer Vorteil bei der Verwendung von Tags liegt darin, dass unterschiedlichste Datentypen miteinander verknüpft werden können, weshalb diese Technik besonders bei heterogenen Datenstrukturen Anwendung findet. Daten jeder Art können mit jedem Tag assoziiert werden und jedem Element können mehrere Tags zugewiesen werden. Durch Tags können leistungsstarke Filter-, Such- und Sortierfunktionen umgesetzt werden.

Durch die Verwendung von Tags können große Datenmengen besser beherrscht werden. Tags können auch als Gedächtnisstütze dienen. Ist man über einen bestimmten Pfad zu einem Objekt gekommen, prägt sich dieser. Ist einem beispielsweise der Name eines Kontakts entfallen, kann man diesen über assoziierte Tags auffinden.

---

<sup>1</sup>Daten, die Informationen über andere Daten beinhalten

Insgesamt kann man noch zwischen internen und externen Tags unterscheiden. Interne Tags sind selbst angelegte Tags, z. B. für eigene Kontakte oder Musikstücke. Externe kommen von öffentlichen Systemen, wie etwa Musikdatenbanken, und ergeben sich z. B. aus den oben beschriebenen Folksonomien.

### 5.1.3 Filtern von Objektlisten

Durch ein gemeinsames Datenmodell können Querverbindungen über Tags zwischen verschiedenen Datenquellen geschaffen werden. So lassen sich relevante Daten sehr schnell eingrenzen und auffinden, wenn sie gezielt gefiltert werden. Zum Einen entstehen dadurch sehr individuelle Filter, zum Anderen ist es möglich, gebräuchliche Filter im Voraus zu definieren.

Eine Filterung mittels vordefinierter Datenfilter, wie beispielsweise „Geschäftskontakte“, kann die Datenmenge gezielt – aber transparent für den Benutzer – einschränken. Eine Einzelauswahl kommt dann entweder durch Suche auf der gefilterten Liste oder explizite Auswahl durch den Benutzer zustande. Die aktuelle Auswahl eines Objektes (oder mehrerer) in einer Liste legt den aktuellen Kontext des Systems fest. Auch die Historie an vormals ausgewählten Objekten kann den Kontext weiter bestimmen, sofern sie nicht durch eine längere Pause oder einen Kontextwechsel zurückgesetzt wird. Nach einem Wechsel in eine andere Unterkategorie wird dieser Kontext als Input für den nächsten Filter verwendet und nur Daten angezeigt, die mit diesem Ausgangsobjekt in Bezug stehen.

Das Konzept sieht zwei Möglichkeiten der Filterung vor: die explizite, vom Benutzer über Bedienelemente oder Spracheingaben ausgelöste Filterung der aktuell angezeigten Liste, die dadurch um einen bestimmten Faktor reduziert wird. Die implizite Filterung einer Liste aufgrund von Kontextinformationen, die vor allem aus der Auswahl und Filterung der aktuellen Liste hervorgehen. Die Ergebnisliste ist in der Regel einer anderen Domäne zugeordnet als die Liste, der die Kontextinformationen entnommen werden.

## 5.2 Graphisches Interface

Unter Berücksichtigung der Vorüberlegungen wird nun das entwickelte Lösungskonzept vorgestellt. Das entstandene graphische Interface zeigt als Ausgangsbildschirm das *Statusmenü* bzw. die Hauptansicht und gliedert sich in die *Domänen* Kontakte, Termine, Route, Unterhaltung und Ereignisse. Die multimodale Bedienung des GUIs erfolgt sowohl über einen Touchscreen in der Mittelkonsole des Fahrzeuges und außerdem mittels Sprache (weitere Einzelheiten dazu im Abschnitt 3.2). Das sehr übersichtlich gestaltete Interface ermöglicht dem Fahrer den objektorientierten Schnelzugriff auf Datenobjekte und Funktionen des Fahrzeuges und der angeschlossenen mobilen Endgeräte. In den folgenden Unterkapiteln erfolgt eine Beschreibung der einzelnen Screens.

### 5.2.1 Statusmenü

Die Statusmenü-Ansicht (siehe Abbildung 5.1(a)) ist üblicherweise der normale Zustand des Displays. Sie gliedert sich in fünf Teilbereiche und bietet dem Fahrer einen perfekten Überblick über relevante Informationen. Durch die starke Reduzierung wird der Fahrer nicht mit unnötigen Informationen überfordert.

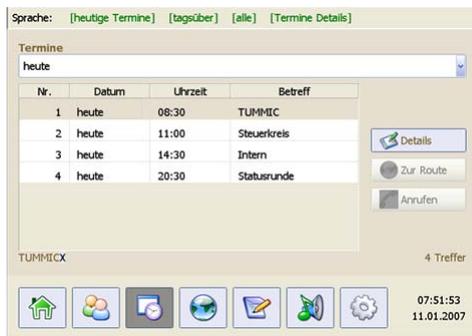
Aufgrund der Bedeutung der Systemnachrichten wird ihnen mit der Domäne Ereignisse der größte Anzeigebereich im Statusmenü eingeräumt. Diese Systemnachrichten stellen üblicher-



(a) Statusmenü



(b) Domäne Kontakte



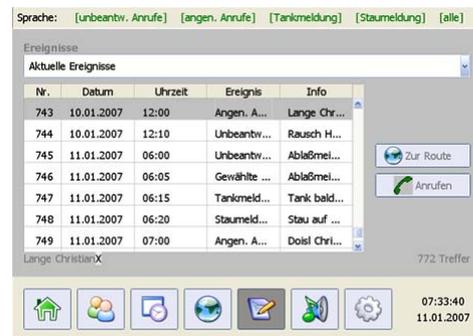
(c) Termine



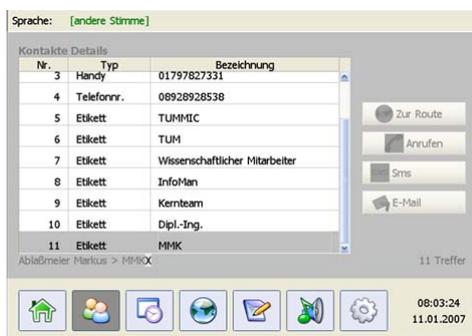
(d) Route



(e) Unterhaltung



(f) Ereignisse



(g) Kontakte Details



(h) Termine Details

Abbildung 5.1: Überblick Statusmenü und Domänen

weise einen Kontextbezug zu den übrigen Domänen Termine, Route, Kontakte und Unterhaltung dar. Beispielsweise bezieht sich eine Staumeldung auf die geplante Route bzw. ein entgangener Anruf auf einen Eintrag der Kontakte. Allerdings gibt es auch Ereignisse, die einen übergreifenden Bezug aufweisen, wie z.B. Kontrollmeldungen des Fahrzeuges, Fälligkeit des Kundendienstes oder spezielle Ereignisse mobiler Endgeräte. Aus Platzgründen findet eine Beschränkung auf sieben Ereignisse im Statusmenü statt. Dies erfordert eine automatisierte Vorfilterung relevanter Ereignisse (beispielsweise nach Wichtigkeit, Bedeutung und Zeitdauer). Die Kriterien für die Filterung müssen für den Fahrer transparent und nachvollziehbar sein.

Bestimmte Ereignisse können auch automatisiert weitere Aktionen auslösen (siehe auch Kapitel 6). Eine Staumeldung beispielsweise triggert eine automatische Anpassung der Routenplanung zur Umfahrung des Staus. Allerdings kann der Fahrer dem jederzeit auch widersprechen und auf diese Ereignisse selbst reagieren. Die Meldung „Stau wird umfahren“ weist den Fahrer dezent auf die geplante Stauumfahrung im Statusmenü hin, ein Widerspruch ist durch Reaktion über den Touchscreen oder ein Sprachkommando jederzeit möglich. Ebenfalls kann hier eine Tankwarnung zur Einplanung eines Tankstopps führen (siehe dazu Abschnitt 6.3.1).

Auf der linken Seite des Statusmenüs befinden sich zwei weitere Bereiche: der obere Bereich gibt einen Überblick über die nächsten drei anstehenden Termine mit zugehörigen Uhrzeiten und der untere Bereich die nächsten anzufahrenden Ziele der Domäne Route. Die anstehenden Termine werden automatisiert in die Routenliste nach Synchronisation mit einem mobilen Kalender übernommen. Beim Einstieg in das Auto wird der Fahrer über die erfolgte Datenübernahme in die Routenliste per Sprachausgabe informiert. Das zunächst angefahrte Ziel wird durch einen kleinen Pfeil vor dem Objekt markiert. Auf Wunsch werden die häufigsten, zuletzt gewählten oder vordefinierten Ziele, wie beispielsweise die Heimatadresse („nach hause“), angeboten.

Auf der rechten Seite des Statusmenüs finden sich ebenfalls zwei Bereiche: im Bereich Kontakte werden die zuletzt bzw. häufig gewählten Personen angezeigt. Der Bereich Unterhaltung zeigt die zur Verfügung stehenden Musikquellen und deren Status an. In Abbildung 5.1(e) ist im Radio der Sender „Bayern 3“ gewählt bzw. der MP3-Player spielt momentan den Song „Dieser Weg“. Die Quelle CD ist ausgegraut, da kein Medium eingelegt ist. Der momentane Zustand ist „Aus“, was der aktivierte Radiobutton zeigt. Als besonderes Feature ermöglichen die Radiobuttons direkte Funktionen im Statusmenü. Durch Anwahl einer Musikquelle wird diese aktiviert und abgespielt. Dabei erfolgt kein Wechsel in die verknüpfte Domäne Unterhaltung, sondern die Anzeige des Statusmenüs bleibt erhalten. So erhält der Autofahrer einen Schnellzugriff auf häufige Funktionen bereits im Statusmenü.

Anders verhält es sich, wenn der Fahrer die großen Icon-Buttons im Statusmenü anwählt. Diese zeigen auf die zugehörigen Domänen und es erfolgt der Sprung auf den Screen der „Domäne“ Termine, Route, Ereignisse, Kontakte oder Unterhaltung. Wählt der Fahrer direkt ein Objekt unterhalb der Buttons an, erfolgt ebenfalls ein Wechsel in die zugehörige Domäne, allerdings bleibt hier der Kontext erhalten, sprich im resultierenden Domänen-Screen ist das ausgewählte Objekt bereits vorselektiert. Aus Platzgründen werden die Bereiche Termine, Route, Kontakte und Unterhaltung auf jeweils drei Objekte begrenzt.

### 5.2.2 Objektorientierte Bedienung über Domänen

Im Folgenden wird nun die objektorientierte, durchgängige Aufteilung der Domänen-Screens erläutert. Die Domänen gliedern sich in *Kontakte*, *Termine*, *Route*, *Unterhaltung* und *Ereig-*

nisse. Zur Unterstützung der Orientierung des Nutzers besitzt jede Domäne einen identischen Screen-Aufbau: Drop-Down Filterliste (oben), Objektliste (mitte), Funktionsbuttons (rechts), Kontextpfad (unten), Domänenwechsel (Fusszeile). Zur Erhöhung der Orientierung und Einprägsamkeit sind die unterschiedlichen Domänen farblich kodiert.

Hinweis: Die Domäne „Einstellungen“ ist im Folgenden nicht weiter ausgearbeitet.

**Domäne Kontakte:** In dieser Domäne werden sämtliche Kontakteinträge, unabhängig von der Datenquelle, in einer gemeinsamen Listendarstellung präsentiert. Insgesamt sind in der Liste maximal sieben Einträge sichtbar. Ist die Treffermenge (die rechts unten angezeigt wird) größer, kann die Liste über einen beweglichen Balken und Pfeile am Textfeldrand gescrollt werden. Die angezeigten Objekte werden mittels einer inkrementierenden Zahl durchnummeriert.

Im Normalfall werden entweder alle Treffer oder durch Berücksichtigung des Kontextes nur eine Submenge angezeigt. Im Kopfbereich des Screens wird der aktuell ausgewählte Filter angezeigt. Durch die Wahl geeigneter Filter lässt sich die Treffermenge über eine Drop-Down Liste jederzeit direkt einschränken. Hier finden sich vordefinierte Filter wie „Geschäftskontakte“, „private Kontakte“, „In der Umgebung“, „Zuletzt gewählt“, usw. Über das Feld „Suchen“ kann die Treffermenge durch einzelne Buchstaben oder Wörter zusätzlich eingeschränkt werden. Auf der rechten Seite des Domänen-Screens können Funktionen, die den Objekten zugeordnet sind, kontextabhängig über Buttons ausgeführt werden, wie z. B. Auslösen eines *Anrufs*, einer *SMS* oder *Zur Route*. Sind Daten zum jeweilige Objekt vorhanden, werden die Funktionsbuttons auf aktiv gesetzt, sonst ausgegraut. Über den Details-Button werden weitere Einzelheiten zum ausgewählten Objekt angefordert. Hier erfolgt ein Wechsel zum Kontaktdetails-Screen. In der Fußzeile befindet sich eine Buttonleiste mit Icons, über die sehr schnell zwischen den einzelnen Domänen gewechselt bzw. ins Statusmenü zurückgesprungen werden kann.

**Domäne Termine:** Hier werden alle erfassten Termine aufgelistet und stehen dem Fahrer so zur Verfügung. Diese können beispielsweise aus einem PDA stammen oder durch ein PC Kalenderprogramm übermittelt worden sein. Durch eine automatische Zielübernahme der Kalendereinträge in die Routenliste findet, wie bereits angesprochen, eine Entlastung des Fahrers statt, dem er auch widersprechen kann. Erfolgt eine Übernahme in die Routenliste, z. B. nach Anschluss und Synchronisation mit einem mobilen Endgerät, wird dies nach Erfolg dem Fahrer mitgeteilt. Über den Drop-Down Filter können die angezeigten Termine sehr einfach eingeschränkt werden. Standardmäßig werden hier nur die heutigen Termine angezeigt. Möglich ist hier allerdings auch die Einschränkung auf die nächsten Stunden.

Durch eine intelligente Verknüpfung der Termindaten mit verschiedensten Objekten des Systems kann der Fahrer beispielsweise sehr einfach auf die Teilnehmerlisten bzw. Zielorte der Termine zugreifen. Über den Button „Details“ werden dem Fahrer beispielsweise nach einem Wechsel in das Detailmenü weitere Einzelheiten zum Termin angezeigt. Über den Button „Zur Route“ kann ein Termin zur Routenliste hinzugefügt werden.

**Domäne Route:** Die Routenliste enthält chronologisch die nächsten anzufahrenden Ziele. Die Liste ist üblicherweise mittels Filter auf die heutigen Ziele eingeschränkt. Die Ziele können automatisiert aus dem Kalender stammen. Auch Ereignisse können Einfluss auf die Routenliste nehmen. So kann ein notwendiger Tankstopp (automatisiert oder manuell) eingeplant werden oder eine Stauumfahrung mit einberechnet werden.

Über die mit den Routenobjekten verknüpften Funktionen kann ein Ziel vorzeitig angefahren werden, die Liste umsortiert werden oder auch ein Ziel entfernt werden. Falls nicht im

Adressbuch vorhanden, können bei Bedarf auch manuell neue Ziele hinzugefügt werden. Ist ein Ziel im Adressbuch vorhanden, erfolgt die Auswahl generell über die Domäne Kontakte. Nach Selektion eines Routenobjektes kann ein Sprung unter Berücksichtigung des Kontextes in die Domänen Kontakte oder Termine erfolgen. Unter Kontakte werden dann beispielsweise die dem Routenobjekt zugeordneten Personen angezeigt. So können z. B. Verspätungen über die verknüpften Personen gemeldet werden.

**Domäne Unterhaltung:** In der Domäne Unterhaltung werden sämtliche zur Verfügung stehende Musikobjekte verschiedener Quellen aufgelistet. Diese können z. B. Radiosender, MP3-Dateien oder CD-Titel sein, die auch von mobilen Endgeräten stammen können. Über den Filter werden die einzelnen Medien bzw. Quellen vorgefiltert. Spezielle Filter ermöglichen die Anzeige und Auswahl von häufigen, zuletzt gewählten Musikobjekten oder einer vordefinierten Playliste. Den Musikobjekten stehen Funktionen wie spielen, anhalten, löschen und zur Playliste (hinzufügen) zur Verfügung. Denkbar ist auch hier die Verknüpfung der Musikobjekte mit verschiedenen Kontaktobjekten. So können individuelle Playlists und Musikpräferenzen verschiedenster Personen zugeordnet und abgespielt werden. Dies entlastet den Fahrer bei der Wahl des Musikprogramms.

**Domäne Ereignisse:** Den Systemnachrichten mit der Domäne „Ereignisse“ wird ein eigener Bereich im MMI eingeräumt. Dort sind Nachrichten aufgelistet, die zwar inhaltlich meist einer der anderen Domänen „Kontakte, Termine, Route, Unterhaltung“ zugeordnet werden können, jedoch immer vom System oder von außen und nicht aber vom Benutzer ausgelöst wurden. Die Filterung der Ereignisse ist auch hier über eine Drop-Down Liste möglich. So können beispielsweise alle „unbeantworteten Anrufe“ oder vorliegende „Staumeldungen“ angezeigt werden.

Eine Reaktion auf die Ereignisse ist zum Einen wieder durch Buttons auf der rechten Seite möglich. Sie bieten verknüpfte Funktionen zum jeweiligen Ereignisobjekt an. Ereignisse können so zu einem beliebigen Zeitpunkt bearbeitet werden. Zum Anderen stehen durch Selektion eines Ereignisobjektes und anschließendem Wechsel in die beabsichtigte Domäne weitere Funktionen zur Verfügung.

**Detailmenüs:** Die Detailmenüs ermöglichen es dem Fahrer, weitere Einzelheiten der zugeordneten Objekten bzw. Daten zu erhalten. Diese sind im Falle der Domäne „Kontakte“ Adressdaten, Rufnummern und zugeordnete Etiketten. Durch Auswahl eines Etiketts kann der Kontext verändert werden. Bei einem erneutem Wechsel, beispielsweise in die Domäne Kontakte, werden die angezeigten Objekte in Abhängigkeit des ausgewählten Etiketts gefiltert. Im Falle der Domäne „Termine“ werden in „Terminedetails“ die Art des Termins und das zugehöriges Datum angezeigt.

### 5.3 Dialogformen, Verknüpfungsmethoden und Filtermechanismen

Dieser Abschnitt greift die zentralen Überlegungen des Grundlagenkapitels auf und stellt die wesentlichen Dialogformen der graphischen Benutzeroberfläche vor. Im zentralen Mittelpunkt stehen hier zum Einen die Zugriffsmethoden auf die Objekte und deren Funktionen. Außerdem wird der Wechsel zwischen verschiedenen Kontexten, hier vor allem Domänen und Statusmenü, erläutert. Zum Anderen spielen in diesem Konzept Filtermethoden eine zentrale Bedeutung die hier nochmals zusammenfassend erläutert werden.

### 5.3.1 Objektorientierter Funktionszugriff in den Domänen

Nach Selektion eines Objektes in einer beliebigen Domäne stehen dem Fahrer dazu weitere Funktionen zur Verfügung. Die kontextabhängigen den Objekten zugeordneten Funktionen lassen sich einfach auslösen, wie beispielsweise einen Anruf ausführen, eine SMS senden oder zur Route hinzufügen. Für das jeweilige Objekt werden Funktionen bei vorhanden Daten auf aktiv gesetzt, nicht vorhandene Funktionen ausgegraut. In der Fußzeile befindet sich eine Buttonleiste über die man sehr schnell zwischen den einzelnen Domänen wechseln kann bzw. ins Statusmenü zurückkehren kann.

#### Schnellzugriff aus dem Statusmenü

Das Interface bietet dem Fahrer zwei Arten von objektorientiertem Schnellzugriff aus dem Statusmenü: eine Sprungfunktion und eine Direktfunktion.

**Sprungfunktion:** Über das Stautsmenü erhalten die Nutzer einen Schnellzugriff zu den wichtigsten Objekten. Beispielsweise kann bei einer für die Route relevanten Staumeldung entweder eine automatische Umfahrung erfolgen oder ein bevorstehender Stau wird einfach nur im Bereich Ereignisse angezeigt. Nun kann der Fahrer explizit die Umfahrung durch die Anwahl des Ereignisobjektes einleiten, aufheben oder ändern. Damit springt der Nutzer in die Domäne Ereignisse, wo das ausgewählte Objekt in der Liste bereits vorselektiert ist. Zugeordnete Funktionsbuttons wie umfahren oder löschen können hier ausgeführt werden. Alternativ kann der Benutzer über den expliziten Filtermechanismus in eine gewünschte Domäne wechseln, um weitere verknüpfte Objekte mit deren zugeordneten Funktionen auszuführen. Für weitere Ereignisse im Statusmenü, wie z. B. dem Tankassistenten, bietet sich ebenfalls die Sprungfunktion an.

**Direktfunktion:** Als weiteres ermöglichen die Quellen-Objekte Radio, MP3, CD im Statusmenü der Kategorie Unterhaltung direkte Funktionen, wie beispielsweise Wechsel der Quelle ohne expliziten Sprung in die Domäne Unterhaltung.

### 5.3.2 Filtermechanismen

Die verschiedenen Filtermechanismen stützen sich auf die Methode Etikettierung bzw. Labeln der Daten. Dadurch ist eine Filterung langer Listen über zwei Methoden möglich: einerseits eine explizite Filterung direkt durch den Fahrer und andererseits eine implizite Filterung durch den jeweiligen Systemkontext.

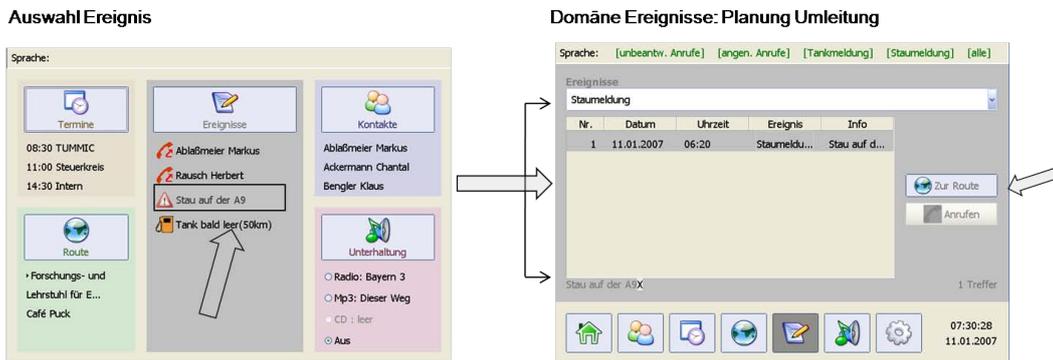
#### Explizite Filterung

Als explizite Filterung bieten sich neben der Suche, vordefinierte und individuelle Filter an.

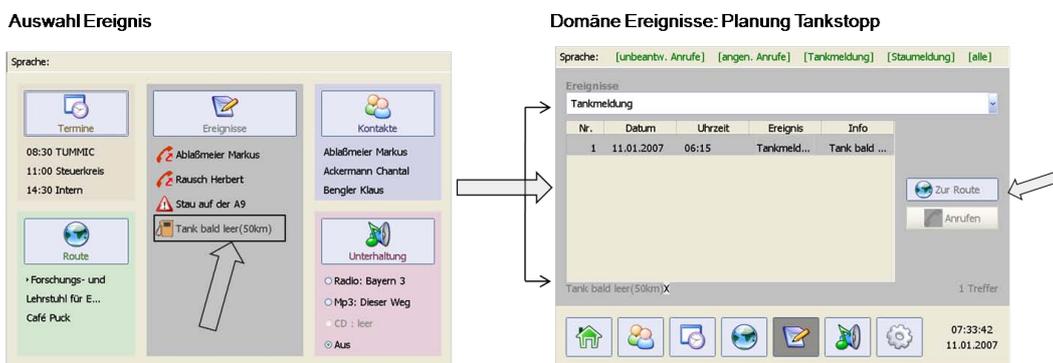
**Vordefinierte Filter:** Die explizite Filterung wird durch den Fahrer ausgeführt. Sie erfolgt durch Auswahl eines vordefinierten Filters aus dem Drop-Down Menü über der Liste.

In Abhängigkeit von den Anforderungen in den jeweiligen Domänen wurden häufig benötigte Filter vordefiniert und stehen dem Fahrer über die Drop-Down Liste zur Verfügung (siehe Tabelle 5.1).

Insgesamt kann man die Filter in die Klassen diskrete, wie z. B. „private Kontakte“ und unscharfe Filter, wie „in der Umgebung“, untergliedern. Default-Filter für die Domänen sind alle Anzeigen bzw. Kontext-Filter im Falle von impliziter Filterung. Ebenfalls ist ein Sprung



(a) Beispiel Staumeldung



(b) Beispiel Tankmeldung

Abbildung 5.2: Statusmenü Sprungfunktionen

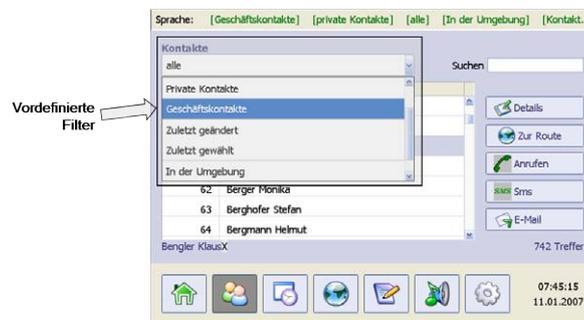


Abbildung 5.3: explizite Filterung

zum Anfangsbuchstaben möglich. Als weitere unscharfe Filter wären Begriffe wie „mehr“ oder „weniger“ denkbar, welche die Trefferliste um den Datenkontext erweitern bzw. reduzieren lassen.

**Suchleiste:** Über eine Suchleiste können einzelne Buchstaben und Wörter eingegeben werden, worauf die Treffermenge auf Übereinstimmungen mit der Eingabe reduziert wird. Dies erfolgt durch Sprache bzw. über taktile alphanumerische Eingabeverfahren, wie ein Onscreen-Keybord, das hier allerdings nicht genauer ausgearbeitet wurde (siehe dazu auch [AAB<sup>+</sup>06]).

DOMÄNE	FILTER
Kontakte	Geschäftskontakte, private Kontakte, in der Umgebung, zuletzt gewählt
Termine	heutige, tagsüber, in den nächsten Stunden
Route	heute
Unterhaltung	Radio, MP3, CD, Playliste
Ereignisse	Unbeantwortete Anrufe, Angenommene Anrufe, Tankmeldung, Staumeldung, Hinzugefügt

Tabelle 5.1: Vordefinierte Filter

**Individuelle Filterung:** Ebenfalls denkbar wäre die Erstellung eigener Filter, die in der Liste speicherbar bzw. selbst definierbar werden können. Als denkbare Lösungsmöglichkeit kann in dem Demonstrator über die Option Etikett eine zweite Liste mit sämtlichen definierten Etiketten parallel darstellen, die über Checkboxen kombiniert werden können. Auch eine Definition von eigenen Filtern über die Suchleiste wäre denkbar.

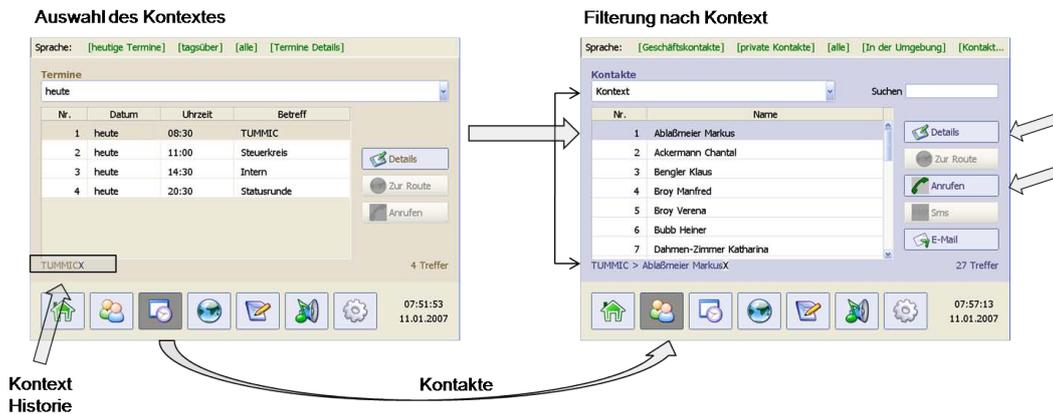
### Implizite Filterung

Die implizite Filterung erfolgt durch den Systemkontext. Der Fahrer hat die Möglichkeit einfach und schnell über die Statusleiste zwischen verschiedenen Domänen zu wechseln. Falls ein Objekt in der Liste selektiert ist und anschließend ein Wechsel zu einer anderen Domäne durchgeführt wird, werden nur Einträge, die mit dem Ausgangsobjekt in Beziehung stehen, angezeigt. Dadurch bleibt der aktuelle Kontext bestehen. Um den Überblick über den Kontext zu bewahren, wird unterhalb des Listenfeldes die Kontexthistorie angezeigt. Als Filter erscheint die Auswahl Kontext. Der Kontext kann über Auswahl eines expliziten Filters oder nach zwei Sprüngen zurückgesetzt werden. Denkbar wäre auch, dass der Kontextpfad durch Anwahl gelöscht bzw. nach Bedarf reduziert werden kann. Sinnvoll, jedoch nicht im Demonstrator realisiert, wäre die Darstellung der aktuellen Kontexthistorie alternativ direkt im ausgewählten Filter.

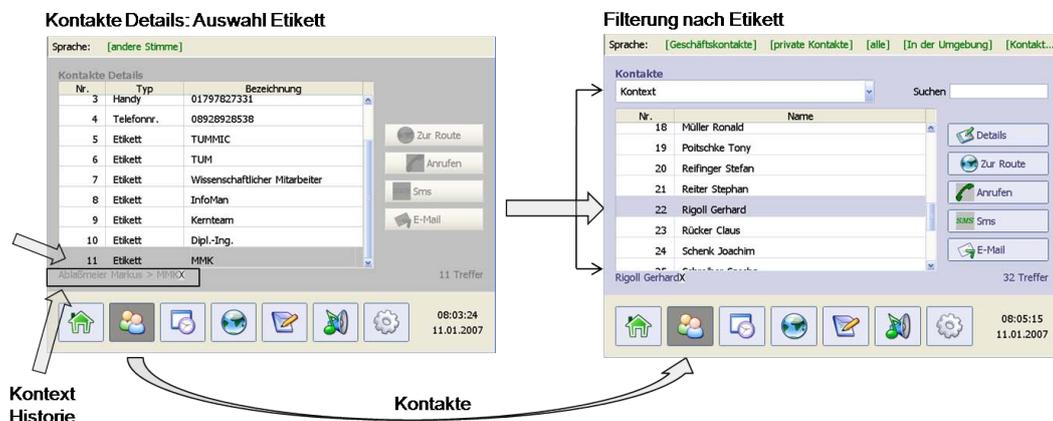
Im ersten Beispiel wird die implizite Filterung bei einer möglichen Verspätungsmeldung dargestellt. Hierzu wählt der Fahrer den relevanten Termin über die Domäne Termine oder Route. Durch Selektion eines Termins und anschließendem Wechsel über die Statusleiste zu der Domäne Kontakte werden die dem Termin zugeordneten Teilnehmer in der Liste angezeigt. Als Filter erscheint hier Kontext und der Kontextpfad zeigt den Titel des Termins. Nach Selektion einer Person können nun z. B. ein Anruf ausgelöst oder SMS versendet werden.

Erreicht der Fahrer aus der angezeigten Personengruppe niemanden, kann er sich weitere Personen über folgende verknüpfte implizite Filterung anzeigen lassen. Er wechselt in die Detailansicht einer Person, wählt hier ein passendes Etikett oder sonstiges Feld und bekommt nach erneuter Anwahl der Domäne Kontakte sämtliche Personen, die zu diesem Etikett zugeordnet sind, angezeigt. Alternativ wäre ja auch eine explizite unscharfe Filterung über „mehr“ oder „weniger“ denkbar.

In Abbildung 5.5 wird ein Überblick über die möglichen Verknüpfungen bzw. expliziten Filtermöglichkeiten dargestellt.



(a) Beispiel Termine



(b) Beispiel Kontakte Details

Abbildung 5.4: Implizite Filterung

## 5.4 Multimodale Bedienung

Wesentlicher Schwerpunkt des Konzeptes liegt auf einer multimodalen Bedienung. Hierbei ergänzen sich Sprache und taktile Interaktion in einer idealen Weise. Um eine möglichst direkte Interaktion zu ermöglichen, wurde ein Touchscreen gewählt. Darüber können Objekte, Funktionen und Filter gewählt werden. Listen können über Pfeile und bewegliche Balken gescrollt werden. Die Spracheingabe erfolgt mittels Verben, Substantiven und Phrasen. Das visuelle Interface und der Wechsel von system- und nutzerinitiierten Dialogen bietet eine gute Menüführung und kontextabhängige Einschränkung der Erkennungsgammatik. Besonders großen Anklang findet hier auch die multimodale Interaktionsstrategie *Point-and-Talk*. Durch einfache Objektselektion über den Touchscreen kann dabei eine gewünschte Funktion über Sprachbefehl ausgeführt werden.

**Spracheingabe:** Um eine fehlerrobuste Spracherkennung für den Fahrer zu ermöglichen, wurde auf die Einbindung eines kommerziellen Spracherkenners zurückgegriffen. Hierbei handelt es sich um den Spracherkennung Vocon ASR 3200 der Firma Scansoft. Dieser basiert auf einer C++ Entwicklungsumgebung, die mittels einer TCP/IP-Socketverbindung mit dem Java-

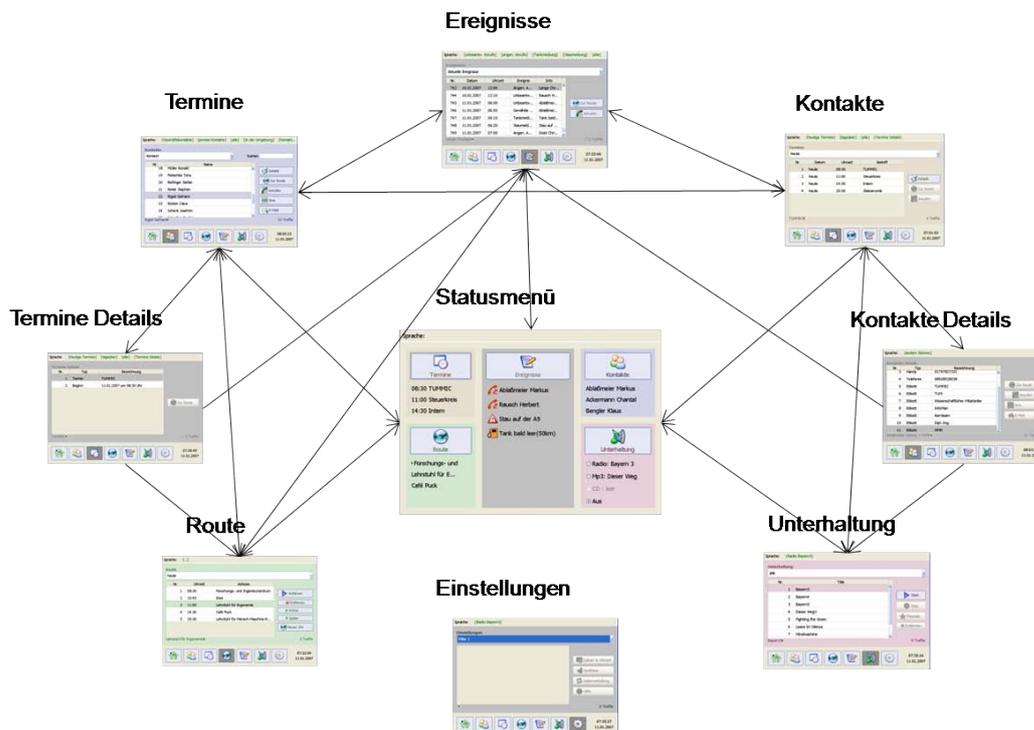


Abbildung 5.5: Überblick über die Verknüpfungen der Domänen und des Statusmenüs

Demonstrator verknüpft wird. Dem Erkenner wird die spezifizierte, dynamische Grammatik über eine bnf-Datei zur Verfügung gestellt. Hier werden sämtliche mögliche Satzkonstruktionen in der bnf-Form vordefiniert. Durch drücken der PTT-Taste (PTT: Push-To-Talk) wird der Erkenner aktiv geschaltet. Nach einem Timeout und einer definierten Schwelle erfolgt die Übermittlung des Ergebnisses an den Demonstrator. Sämtliche Begriffe, die am Bildschirm zu sehen sind, können auch gesprochen werden („Speak what you see“). Darüber hinaus werden wichtige Sprachbefehle für die jeweilige Domäne in der Kopfzeile des Interfaces dargestellt.

**Touchscreen:** Das Konzept wurde für eine multimodale Bedienung ausgelegt. Neben Sprache lassen sich sämtliche Funktionen direkt über das graphische Display mittels Touchscreen bedienen. Hierzu liefert ein 10 Zoll Monitor über einen seriellen Anschluss die Push-Events an die Software. Der Touchscreen befindet sich in der Mittelkonsole knapp unterhalb der Windschutzscheibe, jedoch noch im ergonomischen Greifraum des Fahrers. Eine Ablagefläche der Handfläche am unteren Rand des Bildschirms ermöglicht ein entspanntes, zielgenaues Bedienen und Positionieren. Auf eine Trennung von Anzeige und Bedienung wurde hier bewusst verzichtet. Ein Touchscreen wurde vor allem deshalb gewählt, da das Bedienprinzip eine sehr steile Lernkurve besitzt und mit wenig Vorwissen sehr einfach und direkt bedient werden kann. Dies kommt vor allem der Strategie des objektorientierten Datenzugriffs im Besonderen zu Gute. Die Einzelnen Objekte können so in einen sehr breiten Funktionskontext gesetzt werden, was so einen großen Spielraum für den Fahrer bietet.

## 5.5 Zusammenfassung

Das entwickelte Interface verwendet keine komplizierten hierarchischen Strukturen. In dem Datenbrowsing-Ansatz werden die vorliegenden Daten über ein gemeinsames Datenmodell fünf Domänen zugeordnet. Durch den objektorientierten Ansatz werden natürliche und logische Dialogformen und mentale Modelle unterstützt. Erst nach Anwahl des Objektes in der Domäne kann der Fahrer entscheiden, welche Funktion er ausführen möchte. Dadurch wird auf eine Menüstruktur im herkömmlichen Sinne komplett verzichtet, sondern die Funktionen in den Datenkontext einfach und sinnvoll eingebettet. Bisherige hierarchische Menüstrukturen unterstützen natürliche Denkprozesse nicht und führen den Nutzer in Sackgassen.

Der zentrale Statusbildschirm gibt dem Fahrer stets einen Überblick über die gesamten Daten und bietet einen Schnellzugriff auf diese. Um in den Domänen effizient und effektiv an das Ziel zu kommen wurde ein sehr wirksames Filterverfahren entwickelt. Durch implizite und explizite Filterung gelangt der Fahrer schnell an die gewünschten Daten und Funktionen.

Die Nutzung von Systemkontext und Etikettierung der Daten ermöglicht die Durchführung von mehreren Handlungssequenzen in sehr wenigen Schritten. So ist ein äußerst schneller Wechsel und Sprung zwischen den Domänen sehr einfach möglich, ohne dass der Nutzer die Orientierung verliert und der plausible Kontext verloren geht. Die multimodale Bedienung über Sprache und Touchscreen ergänzen sich nahezu nahtlos. Es wäre zu erwarten, dass der Fahrer das Interface mit einer steilen Lernkurve sehr einfach bedienen kann.

## Kapitel 6

# Kontextsensitive Informationsagenten

Um dem Fahrer bei häufig wiederkehrenden Handlungsabläufen größtmögliche Unterstützung zu bieten, wird nun eine weitere Dialogstrategie zum Umgang mit großen Datenmengen entwickelt. Nachdem sich bereits die *Assistenz bei der primären Fahraufgabe* durch FAS (siehe Abschnitt 2.2.4) etabliert hat, ist auch eine *Assistenz bei FIS bezüglich der sekundären und tertiären Fahraufgabe* denkbar. Ähnlich eines intelligenten Beifahrers unterstützen diese Assistenten (auch sogenannte *Sekretärinnenfunktionen*) den Fahrer bei seinen Aufgaben. In dem hier entwickelten Ansatz kommen „intelligente“ *Agentensysteme* zum Einsatz, die durch Einbezug von Kontextinformationen geeignete Methoden zur Dialogführung und zur Informationsfilterung verwenden. Diese Systeme werden in dieser Arbeit als *Informationsagenten* bezeichnet.

Die Ziele der Informationsagenten sind (siehe auch [LK03]):

- die Vereinfachung der direkten und spontanen Nutzung wichtiger Systemfunktionen,
- die Optimierung und Bündelung von Arbeitsschritten, und damit die Verringerung der subjektiv empfundenen Komplexität,
- die Entlastung des Fahrers durch Teilautomatisierung,
- die Filterung des Informations- und Datenaufkommens,
- die Unterstützung beim Erlernen der Systemfunktionen und der Bedienung.

Die folgenden Abschnitte erläutern den prinzipiellen Aufbau dieser Informationsagenten, die erforderlichen mathematischen Methoden und die entwickelten Lösungskonzepte für den Einsatz im Automobil.

### 6.1 Aufbau der Informationsagenten

Die Grundidee der Informationsagenten ist zunächst die sehr einfach gehaltene Dialogführung für eine Folge verknüpfter Handlungsabläufe des Fahrers. Um die erforderlichen Bedienschritte sinnvoll zu bündeln und den Nutzer in wenigen Schritten und ohne Umwege an sein Ziel zu führen, kommen zur Dialogführung sogenannte *Wizards*<sup>1</sup> zum Einsatz (siehe Abbildung 6.1 und Abschnitt 3.1). Damit sich der Dialog sinnvoll an die jeweiligen Bedürfnisse der Fahrer

---

<sup>1</sup>hier wird der Nutzer über ein Sequenz von Dialogschritten bzw. -fenstern geführt

anpasst, wird dieser zum Teil *situations- und nutzeradaptiv* gestaltet und relevante Kontextinformationen mit einbezogen (siehe auch Abschnitt 3.3). Eine wesentliche Eigenschaft der Agenten ist es, dass sie sowohl nutzer- als auch systeminitiiert aufgerufen werden können. Um die Dialoge möglichst effizient zu gestalten, werden diese multimodal ausgelegt (siehe auch [HHH06, Ben01]). Neben visueller und sprachbasierter Nutzerführung über die Wizard-Dialoge kann der Fahrer notwendige Eingaben und Entscheidungen über Sprache und taktile Eingabegeräte tätigen. Gerade Frame-Based Dialoge (siehe Abschnitt 3.1.3) ermöglichen dabei die gebündelte Eingabe mehrerer Intentionen in einem oder wenigen Bedienschritten. Die Realisierung und Integration der multimodal gestalteten Informationsagenten wird auf Basis des HUD\_MAX-Konzepts in Abschnitt 8.2 aufgezeigt.

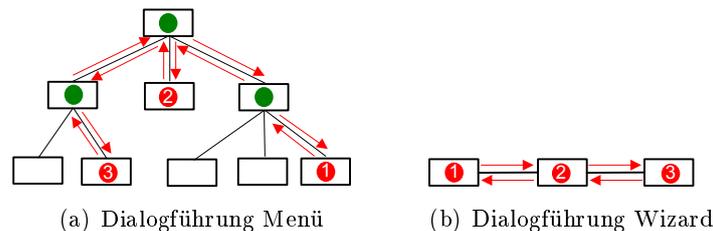


Abbildung 6.1: Bündelung der Dialogschritte durch die Informationsagenten

Der Schwerpunkt dieses Kapitels ist die Identifikation von Methoden und Ansätze für die *kontextsensitive Informationsaufbereitung*. Eine wesentliche Aufgabe der Informationsagenten besteht in der geeigneten (Vor-)Filterung und Bereitstellung relevanter Informationen für den Fahrer (siehe auch [Lug01]). Statistische Methoden bieten hierbei gute Möglichkeiten, um aus früheren Aktionen der Nutzer Muster zu identifizieren und dadurch auf Intentionen des Fahrers zu schließen. Mit der adaptiven Informationsbereitstellung im Fahrzeug haben sich bereits die Arbeiten [HBL01, HL01, GT00, RFL99] beschäftigt.

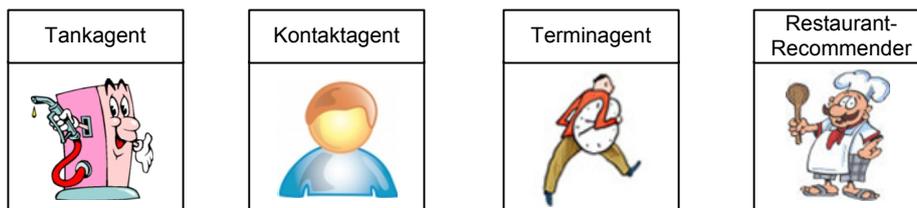


Abbildung 6.2: Überblick Informationsagenten

In dieser Arbeit werden Lösungskonzepte für folgende Informationsagenten erarbeitet (siehe Abbildung 6.2):

- **Tankagent:**

Der Schwerpunkt des Tankagenten liegt in der Unterstützung und Optimierung der Tankaufgabe des Fahrers [APRR07]. In Abbildung 6.3 ist ein Beispieldialog eines Tankagenten aufgeführt. Der Tankagent lernt aus dem Verhalten des Fahrers und filtert entsprechend den Tankgewohnheiten die in Frage kommenden Tankstellen vor. Dabei kommen Bayes'sche Netze zur Anwendung, mit denen mittels bedingter Wahrscheinlichkeiten Schlüsse aus Beobachtungen gezogen werden können.

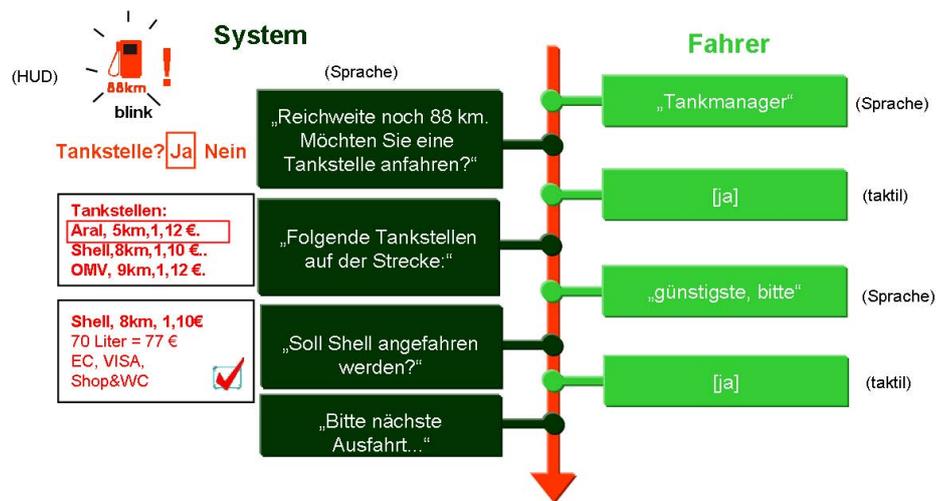


Abbildung 6.3: Beispiel eines Tankagentendialogs

- **Kontaktagent:**

Bei dem Kontaktagenten erfolgt die situative Unterstützung des Fahrers bei der Auswahl des geeigneten Kommunikationskanals (z. B. Büronummer, Privatnummer, SMS, usw.). Durch die Analyse der verwendeten Kommunikationsmöglichkeiten wird der geeignete Kanal vorausgewählt. Zur Analyse des Verhaltens kommen ebenfalls Bayes'sche Netze zur Anwendung.

- **Terminagent:**

Er verwaltet die anstehenden Termine und überprüft, ob diese eingehalten werden können. Kommt es voraussichtlich zu Verspätungen informiert er den Fahrer und bietet die Information der Terminteilnehmer an.

- **Restaurant-Recommender:**

Recommender Systeme verhelfen Benutzern bei zunehmender Anzahl an Daten die gewünschten zu finden (siehe auch [GT00, APRR07]). Sie verwenden die Kenntnis über das Nutzerverhalten, um sie durch die Datenvielfalt zu leiten. Internetanbieter wie beispielsweise eBay, Amazon und CDNOW setzten solche Systeme ein. Um dem Fahrer die Auswahl von POIs im Navigationssystem zu erleichtern, wird am Beispiel des Restaurant-Recommenders ein Konzept erarbeitet, das unter Berücksichtigung von kontextuellen Informationen und der Analyse des Benutzungsverhaltens eine Vorbewertung vornimmt. Zur Datenanalyse werden Neuronale Netze verwendet.

### 6.1.1 Grundlagen von Agentensystemen

In diesem Abschnitt werden zunächst die wesentlichen Grundlagen zum Verständnis von Agentensystemen erläutert. In der Literatur gibt es unterschiedliche Definitionen für „intelligente“ Softwareagenten. Ein *Agent* (lat. Handelnder) als elementarer Begriff kann wie folgt beschrieben werden:

„Eine Person oder Sache, die in der Lage oder ermächtigt ist, im Auftrag Dritter zu handeln.“ [CH98]

Ein Softwareagent ist dabei:

„Eine Softwareentität, die vom Anwender delegierte Aufgaben autonom erfüllt.“ [CH98]

Eine Schwierigkeit liefert die Abgrenzung „intelligenter Agenten“ zu „normalen Softwareprogrammen“. Je nach Definition und Betrachtungsweise kann beim Einsatz von lernfähigen Agenten auch von *künstlicher Intelligenz* (KI) gesprochen werden.

### Eigenschaften

Um entsprechend der Definition im Auftrag eines Nutzers handeln zu können, müssen Agenten mit bestimmten Fähigkeiten versehen werden. Die wesentlichen Eigenschaften, die im Folgenden genannt werden, müssen dabei nicht zwingend vorhanden sein – andererseits kann ein Agent mehrere dieser Eigenschaften auch gleichzeitig in sich vereinen. Nach [BZW98] sind dies:

- Autonomie
- Mobilität
- Kommunikation/Kooperation
- Soziales Verhalten
- Reaktivität
- Deliberatives Verhalten

Unter einem *autonomen* Agenten versteht man, dass dieser Aufgaben ohne den Eingriff oder eine explizite Anweisung des Menschen ausführt. Durch diese Selbständigkeit entlastet er den Anwender. Autonomie ist die bedeutendste Eigenschaft von Softwareagenten (vgl. Definition nach [CH98]). Besonders für Internet-Anwendungen ist die *Mobilität* eine wichtige Eigenschaft. Die Agenten verbreiten sich hierbei selbständig von Host zu Host und erledigen lokal ihre vorbestimmte Aufgabe. Der Herkunftsrechner muss dabei nicht ständig mit dem Netz verbunden sein. Als Beispiele sind hier Cookies zu nennen – aber auch Computerviren können als mobile Agenten aufgefasst werden. Ein weiterer wichtiger Aspekt ist die Interaktion mit der Umwelt. So muss der Agent entsprechende Werkzeuge besitzen, um mit dem Nutzer oder anderen Agenten *kommunizieren* zu können. Mit der Fähigkeit zur *Kooperation* können Agenten gemeinsam eine Aufgabe lösen, indem jeder Agent eine Teilaufgabe bearbeitet. Die Befähigung zu *sozialem Verhalten* besagt, dass Agenten gewisse Regeln und Normen einzuhalten haben. Dabei dient das menschliche Verhalten als Vorbild. Die letzten beiden Begriffe *Reaktivität* und *deliberatives Verhalten* beinhalten neben den Eigenschaften vor allem auch den Aufbau und die Funktionsweise von Softwareagenten. Agentensysteme werden oftmals auch als *bot* (abgeleitet von robot) bezeichnet.

**Reaktive und deliberative Agenten:** Zwei Architekturen gilt es zu unterscheiden. Einerseits existiert die Gruppe der sogenannten reaktiven Agenten und als Gegenstück die Agenten mit deliberativem Verhalten. Allerdings werden die Vorteile dieser beiden Architekturen auch in einer Mischform vereint. Diese Arten von Softwareagenten werden als hybride Agenten bezeichnet. *Reaktive Agenten* sind nach dem *Stimulus-Response-Prinzip* aufgebaut und geben eine eher einfache Agentenstruktur wider. Entsprechend vorgegebener Verhaltensregeln reagiert der Agent auf Zustände oder Umwelteinflüsse. Reaktive Agenten sind zwar weniger

komplex aufgebaut als deliberative Agenten, dafür aber leichter an neue Herausforderungen anzupassen. Ihre Intelligenz rührt aus dem Zusammenspiel mehrerer Agenten her und wird daher *verteilte KI* genannt. Dem internen Zustand *deliberativer Agenten* liegt eine BDI-Struktur mit Überzeugungen (*beliefs*), Wünschen (*desires*) und Zielen (*intentions*) zugrunde. Sie besitzen ein Modell ihrer Umwelt als Wissensbasis und ein Lern- sowie ein Kommunikationsmodul entsprechend der klassischen KI.

### 6.1.2 Systemarchitektur der Informationsagenten

Die Architektur des Informationsagenten basiert auf dem in Abschnitt 3.5 vorgestellten multimodalen, kontextadaptiven Dialogframework. Zur Realisierung der Anforderungen des Informationsagenten wird die Architektur um die Komponente des *Informationsmanagers* erweitert (siehe Abbildung 6.4). In dieser Komponente erfolgt die Vorverarbeitung der anliegenden Informationen entsprechend des Kontextes.

Die erläuterten Informationsagenten erfordern auch die automatische *Adaption* der *Präsentationskomponente* und der *Dialogkomponente* an die aktuelle Situation, Aufgabe und den Benutzer (siehe dazu auch [ALK01, LK03]). Die Adaption in diesen Komponenten ermöglicht die Minimierung der notwendigen Interaktionsschritte zur Erfüllung der Aufgabe und somit der Bedienzeit. Durch die Berücksichtigung des Wissensstands und der Präferenzen des Benutzers werden mentale Überforderungen bei der Bedienung reduziert. Das Zusammenspiel dieser drei Komponenten kann letztlich zu einer kürzeren und effizienteren Interaktion mit FIS und damit zur Reduktion des Ablenkungspotentials führen.

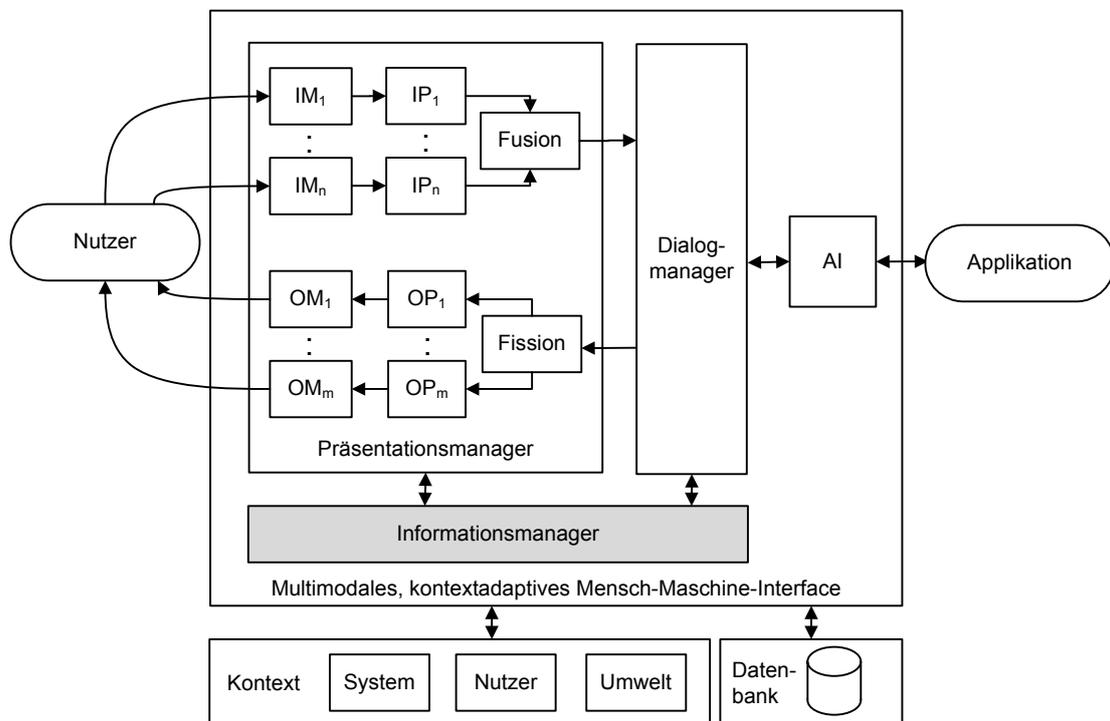


Abbildung 6.4: Agentensystem als Erweiterung eines Dialogsystems

Nachfolgend werden die Aufgaben der Komponenten erläutert:

**Informationsmanager:** Diese Komponente legt fest, *Welche* Informationen (*bzw. Was*) und *Wann* diese dargestellt werden. Nur die in der jeweiligen Situation nützlichen Informationen sollen angeboten werden. Dagegen sind die in einer bestimmten Situation irrelevanten Informationen herauszufiltern, da unerwünschte Hilfestellungen leicht zum Ärgernis werden. Die Aufgaben des Informationsmanagers sind im Einzelnen:

- *Filtern* irrelevanter Daten
- *Priorisierung* von Anzeigen
- *Timing* von Nachrichten (Zeitpunkt, Dauer, Häufigkeit)
- *Guiding & Tutoring* (Erläuterung der Funktion wie Help-Funktionen; Hinweise auf wenig genutzte Funktionen, wie z. B. „Wussten Sie schon?“; Erläuterung der Bedienung wie Hinweise auf alternative Eingabemöglichkeiten und erweiterte Beschreibung von Menüpunkten)

Darüber hinaus hat der Informationsmanager Einfluss auf die Komponenten:

**Dialogmanager:** Dabei soll der Dialogmanager den Fahrer mittels angepasster Dialoge führen. Die *Wizard-Dialoge* ermöglichen eine schrittweise Nutzerführung bei den Aufgaben und stellen eine Erweiterung der übrigen Dialoge dar. Die Initiierung und Steuerung dieser Dialoge erfolgt über den Informationsmanager oder direkt auf Wunsch des Fahrers.

**Präsentationsmanager:** Gerade auch das *Wie* und *Wo* lässt sich an den Kontext des Fahrers anpassen. Hierzu kommt der Präsentationsmanager zum Einsatz. So kann beispielsweise die Anordnung, Schriftgröße und Modalität an den aktuellen Kontext angepasst werden. Auf die Funktion des Präsentationsmanager wird in diesem Kapitel nicht explizit eingegangen.

Der Informationsmanager kann u. a. folgende Systeme umfassen:

- *Planerkenner:*  
Dieses System versucht die Intention der Nutzer zu erkennen und antizipieren.
- *Konfliktmanager bzw. Prioritätenmanager:*  
Treffen mehrere Informationen gleichzeitig aufeinander so kommt es zu einem Konflikt. Hierzu müssen Konzepte erarbeitet werden, die einen solchen Konflikt auflösen. Diese lassen sich zum Einen durch geschicktes Timing und Entzerrung und zum Anderen durch geschickte Präsentation auflösen. Ein Konfliktmanager soll dabei zwischen verschiedenen Agenten vermitteln.

## 6.2 Mathematische Grundlagen und Methoden

In diesem Abschnitt werden zunächst die mathematischen Grundlagen der Informationsagenten erläutert. Zunächst ist hier der Einsatz regelbasierter Ansätze möglich. Bei besonders vielen Informationen, Daten und Einflussgrößen kommen diese Regelwerke jedoch schnell an ihre Grenzen. Aus diesem Grund werden für die Informationsmanager wahrscheinlichkeitbasierte Ansätze verwendet. Dabei kommen Bayes'sche Netze und Neuronale Netze zur Anwendung, die im Folgenden erläutert werden.

### 6.2.1 Einführung in Bayes'sche Netze

Der Alltag ist vielfach durch voneinander abhängige Ereignisse geprägt. Zum Beispiel hat das Auftreten eines bestimmten Sachverhaltes Einfluss auf eine (oder mehrere) Entscheidung(en), die wir treffen müssen. Anders als in der Boole'schen Logik, bei der eine getroffene Annahme entweder *wahr* (*ja* bzw. 1) oder *falsch* (*nein* bzw. 0) ist, existieren in der Realität nicht nur klar abgrenzbare Begebenheiten. Vielmehr begegnen uns unbestimmte (unscharfe) Ereignisse. Anstatt zu sagen, "es sind ein paar Wolken am Himmel, also wird es regnen", ist die Aussage "der Himmel ist zu 90 % bedeckt, also wird es mit 60 %iger Wahrscheinlichkeit regnen" präziser. Bayes'sche Netze bieten eine relativ gute Möglichkeit, Ereignisse und die resultierenden Reaktionen zu modellieren und Wahrscheinlichkeiten für deren (Nicht-)Auftreten zu berechnen.

#### Bayes-Theorem

Um die Wahrscheinlichkeit für das Auftreten bzw. Nichtauftreten eines Ereignisses unter einer (oder mehrerer) Voraussetzungen berechnen zu können, wird eine mathematische Vorschrift benötigt. Die den Bayes'schen Netzen zugrunde liegende Formel liefert das Theorem des Mathematikers und Pfarrers Thomas Bayes, der in England des 18. Jahrhunderts gelebt hat. Das Bayes-Theorem, dem das Bayes'sche Netz seinen Namen verdankt, gibt eine Berechnungsvorschrift für *bedingte Wahrscheinlichkeiten* an. Für zwei gegebene Ereignisse  $A$  und  $B$  – wobei hier  $A$  die Ursache und  $B$  die Wirkung repräsentiert – lautet das Bayes-Theorem [Jen01].

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} \quad (6.1)$$

Dieses lässt sich leicht aus dem grundlegenden Gesetz der bedingten Wahrscheinlichkeiten zur Berechnung der Verbundwahrscheinlichkeiten herleiten. Aus

$$P(A|B) \cdot P(B) = P(A, B) \quad (6.2)$$

und

$$P(B|A) \cdot P(A) = P(A, B) \quad (6.3)$$

erhält man durch Gleichsetzen

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A). \quad (6.4)$$

Durch Umformen ergibt sich aus Gleichung 6.1 das Bayes-Theorem. Hierbei besitzt die Variable  $A$  endlich viele *Zustände*  $a_1, \dots, a_n$  ( $a_i \neq a_n; i, j = 1, \dots, n$ ) – analog sind  $b_1, \dots, b_n$  endlich viele *Zustände* für die Variable  $B$ . Die allgemeine Wahrscheinlichkeit  $P(A)$  für das Auftreten des Ereignisses  $A$  wird als a-priori-Wahrscheinlichkeit bezeichnet. Mit  $P(B|A)$  erhält man die Wahrscheinlichkeit für das Auftreten von  $B$  unter der Voraussetzung, dass  $A$  eintritt. Für die Summe der Wahrscheinlichkeiten aller möglichen Zustände von  $A$  gilt

$$\sum_{i=1}^n P(a_i) = 1 \quad (6.5)$$

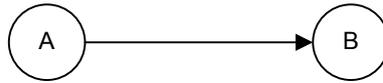


Abbildung 6.5: Einfachstes Bayes'sches Netz

**Graphische Darstellung:** In Abbildung 6.5 ist das einfachste Bayes'sche Netz mit einer Ursache  $A$  und einer Wirkung  $B$  aufgezeigt. Ursache und Wirkung werden als Knoten bezeichnet; die Verbindung dazwischen als Kante. Letztere repräsentiert eine bedingte Abhängigkeit und gibt in Form eines Pfeils die Richtung von der Ursache zur Wirkung an. Ein Bayes'sches Netz ist ein gerichteter, azyklischer Graph (*engl. directed acyclic graph*, kurz DAG).

In Abbildung 6.6 ist Knoten  $W$  eindeutig als Ursache und Knoten  $S$  und  $G$  sind unverkennbar als Wirkung definiert.

### Bayes'sche Netze in der Anwendung

Ein Bayes'sches Netz besteht aus

- einer Menge an Variablen
- einem graphischen Gerüst, das die Abhängigkeiten der Variablen zeigt
- aus Tabellen mit den bedingten Wahrscheinlichkeitsverteilungen für jede Variable.

Anhand eines Beispiels soll gezeigt werden, wie mit Bayes'schen Netzen aus Beobachtungen Wahrscheinlichkeiten für das Auftreten nachfolgender Ereignisse modelliert und berechnet werden können. Zunächst werden alle Variablen bestimmt, die für das zu erstellende System relevant sind. Im Beispiel (Abbildung 6.6) sind dies die Knoten *schönes Wetter* ( $W$ ), *Sonnencreme* ( $S$ ) und *glücklich* ( $G$ ). Mit Wissen über die kausalen Zusammenhänge werden die Abhängigkeiten zwischen den Variablen festgelegt (Kanten im Graphen). Die Ursache ist schönes Wetter mit der Folge, dass eine Person glücklich ist und der Auswirkung, ob diese Person Sonnencreme verwendet. Die Grenzzustände in allen drei Knoten lauten: *trifft zu* und *trifft nicht zu*.

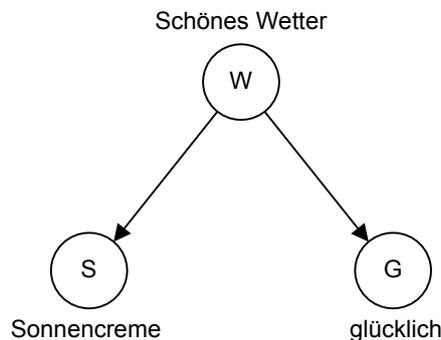


Abbildung 6.6: Bayes'sches Netz mit drei Variablen (Knoten) und deren Abhängigkeiten (Kanten)

**Initialisierung:** Bis jetzt kann man aus dem Netz die beteiligten Variablen auslesen und deren kausale Abhängigkeiten erkennen. Nun müssen in das so entstandene Bayes'sche Netz noch die Werte für die einzelnen Zustände jeder Variablen eingebracht werden, d. h. das Netz wird initialisiert. Die allgemeine Auftrittswahrscheinlichkeit für schönes Wetter sei  $P(W = 1) = 0,55$ . Nach Gleichung 6.5 ist die Wahrscheinlichkeit für den komplementären Fall (das Wetter ist nicht schön) damit gegeben zu  $P(W = 0) = 0,45$ . Herrscht schönes Wetter, ist die Person, die diesem Modell zugrunde liegt, zu 80 % glücklich und verwendet mit 70 %iger Wahrscheinlichkeit Sonnencreme. Die Person ist allerdings auch zu 30 % glücklich, wenn das Wetter nicht schön ist. In mathematischer Notation:  $P(G = 1|W = 0) = 0,3$ .

In Tabelle 6.1 sind alle für die Initialisierung benötigten Wahrscheinlichkeiten eingetragen.

$P(G W)$	$W = 1$	$W = 0$	$P(S W)$	$W = 1$	$W = 0$
$G = 1$	0,8	0,3	$S = 1$	0,7	0,1
$G = 0$	0,2	0,7	$S = 0$	0,3	0,9

Tabelle 6.1: Bedingte Wahrscheinlichkeiten  $P(G|W)$  für  $G$  unter Voraussetzung  $W$  und  $P(S|W)$  für  $S$  unter Voraussetzung  $W$

Aus Gleichung 6.2 bzw. 6.3 lassen sich nun die Verbundwahrscheinlichkeiten berechnen (siehe Tabelle 6.2).

$P(G, W)$	$W = 1$	$W = 0$	$P(S, W)$	$W = 1$	$W = 0$
$G = 1$	0,44	0,135	$S = 1$	0,385	0,045
$G = 0$	0,11	0,315	$S = 0$	0,165	0,405

Tabelle 6.2: Verbundwahrscheinlichkeiten  $P(G, W)$  und  $P(S, W)$

**Marginalisierung:** Um die bedingten Wahrscheinlichkeiten nach dem Bayes-Theorem berechnen zu können, werden noch die Auftrittswahrscheinlichkeiten der Zustände  $b_1, \dots, b_n$  des Kindknotens  $B$  benötigt. Allgemein erhält man  $P(B)$  durch Marginalisierung aus

$$P(b_i) = \sum_{j=1}^m P(a_j, b_i) \quad (6.6)$$

Für das Beispiel müssen also  $P(S)$  und  $P(G)$  berechnet werden. Die Ergebnisse dieser einfachen Summationen sind in Tabelle 6.3 für  $P(S)$  und für  $P(G)$  dargestellt.

$P(S, W)$	$W = 1$	$W = 0$	$P(S)$	$P(G, W)$	$W = 1$	$W = 0$	$P(G)$
$S = 1$	0,385	0,045	0,43	$S = 1$	0,44	0,135	0,575
$S = 0$	0,165	0,405	0,57	$S = 0$	0,11	0,315	0,425

Tabelle 6.3: Auftrittswahrscheinlichkeiten  $P(S)$  und  $P(G)$  durch Marginalisierung

Mit diesen Ergebnissen und dem Bayes-Theorem können nun die Rückschlusswahrscheinlichkeiten – also von der Wirkung zur Ursache – berechnet werden. Dazu werden die Verbundwahr-

scheinlichkeiten  $P(A,B)$  durch die Auftrittswahrscheinlichkeiten der Zustände in den Kindknoten dividiert. Mit den Ergebnissen aus Tabelle 6.4 wurden alle Parameter des Bayes'schen Netzes berechnet. Auch wenn die Algorithmen relativ leicht nachzuvollziehen sind, ist der

$P(W S)$	$W = 1$	$W = 0$	$P(W G)$	$W = 1$	$W = 0$
$S = 1$	0,895	0,105	$G = 1$	0,765	0,235
$S = 0$	0,289	0,711	$G = 0$	0,259	0,741

Tabelle 6.4: Rückschlußwahrscheinlichkeiten  $P(W|S)$  für  $W$  unter Voraussetzung  $S$  und  $P(W|G)$  für  $W$  unter Voraussetzung  $G$

Notationsaufwand schon bei kleinen Netzen nicht unerheblich. Mit entsprechenden Software-Tools kann dieser Aufwand aber erheblich verringert werden.

**Evidenz:** Interessant wird das Arbeiten mit Bayes'schen Netzen, wenn man den Zustand einer oder mehrerer Variablen beobachtet und daraus Wahrscheinlichkeiten für das Auftreten eines Zustandes in einer anderen Variable kalkulieren möchte. Wie eingangs erwähnt, liegt der Vorteil der hier behandelten Netze in der Möglichkeit, Aussagen der Art „Ich bin mir nicht sicher, aber mit 80 %-iger Wahrscheinlichkeit befindet sich  $A$  in Zustand  $a1$ .“ zu berücksichtigen und damit unsicheres Wissen einzubringen. Andererseits ist es aber auch möglich, die Intensität eines Zustandes mit einzubeziehen. Die Variable Tankinhalt kann entsprechend eines zu drei Vierteln gefüllten Tanks mit  $P(\text{„Tankinhalt“} = \text{„Voll“}) = 0,75$  bemessen werden. Wird eine konkrete Beobachtung gemacht – z. B. „es regnet“ – so wird dem Zustand einer entsprechenden Variable volle Evidenz zugewiesen, also  $P(\text{„Wetter“} = \text{„Regen“}) = 1$ . Erlangt ein Knoten (bzw. eine Variable) Evidenz, so werden die Parameter der Kindknoten nach den in diesem Abschnitt dargestellten Schemata Initialisierung und Marginalisierung berechnet. Existieren zum evidenten Knoten Elternknoten, so erhält man deren Parameter durch Berechnung der Verbundwahrscheinlichkeit aus  $P(\text{Elternknoten}|\text{evidenter Kindknoten})$  und  $P(\text{evidenter Kindknoten})$ . Gemäß dem Beispiel wird angenommen, jemand ist überglücklich, also  $P(G = 1) = 1$ . Aus Gleichung 6.2 erhält man Tabelle 6.5 und daraus durch Marginalisierung nach Gleichung 6.6  $P(W = 1) = 0,765$  und  $P(W = 0) = 0,235$ .

$P(G W)$	$W = 1$	$W = 0$	$P(S W)$	$W = 1$	$W = 0$
$G = 1$	0,765	0,235	$S = 1$	0,536	0,023
$G = 0$	0	0	$S = 0$	0,230	0,211

Tabelle 6.5: Bedingte Wahrscheinlichkeiten  $P(G|W)$  für  $G$  unter Voraussetzung  $W$  und  $P(S|W)$  für  $S$  unter Voraussetzung  $W$

Schließlich hat eine Veränderung eines Elternknotens wiederum Auswirkungen auf seine anderen Kindknoten. Um die Parameter für die Variable Sonnencreme zu berechnen, benötigt man zuerst die Verbundwahrscheinlichkeit  $P(S,W)$  aus  $P(W)$  und  $P(S|W)$  (siehe Tabelle 6.5). Durch Marginalisierung folgt daraus die Auftrittswahrscheinlichkeit  $P(S)$  mit  $P(S = 1) = 0,559$  und  $P(S = 0) = 0,441$ .

Man sieht also, dass die Beobachtung, dass sich jemand sehr glücklich fühlt, auch die Wahrscheinlichkeit vergrößert, dass Sonnencreme benutzt wird. Da sich die Wahrscheinlichkeit ver-

größert, dass die modellierte Person glücklich ist, wenn das Wetter schön ist, kann man aus der Stimmungslage dieser Person auf das Wetter schließen, welches wiederum die Wahrscheinlichkeit der Benutzung von Sonnencreme beeinflusst. Die Veränderung einer Variablen hat also nicht nur Auswirkungen auf direkt benachbarte Kind- oder Elternknoten.

Da auch die Möglichkeit gegeben ist, dass mehrere Knoten zur gleichen Zeit Evidenz erlangen, ist zu beachten, dass bei jeder neu eingebrachten Evidenz die Parameter der Knoten im gesamten Netz neu kalkuliert werden müssen. Aufgrund der Kommutativität der verwendeten Rechenoperationen (Addition und Multiplikation) ist die Reihenfolge, in der man die Evidenzen einbringt, allerdings unerheblich für die Bearbeitung.

### Training

Erste Anwendungen erfuhren Bayes'sche Netze in medizinischen Diagnosesystemen. Aus Art und Intensität gewisser Symptome kann so auf mögliche Krankheiten rückgeschlossen werden. Dabei wurden Erfahrungen in die Netze eingebracht, welche Anzeichen bei physischen oder psychischen Erkrankungen festzustellen waren. Das Netz wurde sozusagen vor seinem Einsatz trainiert. Das Netz kann – muss aber nicht – auch während seiner Verwendung durch Anpassung der beobachteten Zustandsvariablen verändert bzw. verbessert werden. Es kann also online trainiert werden. Ein sehr gutes Beispiel hierfür bieten die Spam-Filter in E-Mail-Programmen, wie etwa Mozilla Thunderbird. Durch sogenannte bayes'sche Filter lernt das Programm vom Nutzer, welche E-Mails als Spam einzuordnen sind. "Der Filter schafft bereits nach kurzem Training mit zirka 30 E-Mails erstaunlich hohe Trefferquoten – auch wenn für die produktive Nutzung mindestens ein paar hundert Mails beider Kategorien empfohlen werden."

### Logikschaltung

Wie weiter oben erwähnt, ist einer der Hauptgründe für die Verwendung eines Bayes'schen Netzes der, dass man aus gegebenen, beobachteten Ereignissen die Wahrscheinlichkeit für das Auftreten eines weiteren Ereignisses berechnen und angeben kann. Daneben ist aber auch der regelbasierte Ansatz darstellbar, indem in den Wahrscheinlichkeitstabellen für die Zustände nur noch die Werte 0 und 1 zugelassen werden. Die Regel kann dann wie folgt umgesetzt werden:

WENN im Elternknoten Zustand  $E1$  gleich 0 UND Zustand  $E2$  gleich 1,  
DANN im Kindknoten Zustand  $K1$  gleich 1 UND Zustand  $K2$  gleich 1.

Weiterhin lassen sich so sehr einfach Logikschaltungen aufbauen, die jedoch über eine rein binäre Darstellung hinausgehen. So können in ein und demselben Bayes'schen Netz binäre Informationen mit wahrscheinlichkeitsbasierten Ereignissen gekoppelt und bearbeitet werden, anstatt dass mehrere eigene Systeme erstellt werden müssen. Im Kontaktagent in Abschnitt 6.3.2 kommen sowohl der wahrscheinlichkeits- als auch der logikbasierte Ansatz zur Anwendung.

### 6.2.2 Einführung in Neuronale Netze

Das Gebiet der Neuronalen Netze ist sehr umfangreich und breit gefächert. Aus diesem Grund wird in diesem Kapitel nur ein kurzer Überblick gegeben und grob auf die *Probabilisti-*

schen Neuronalen Netze (PNN) (siehe dazu [Rig94]) eingegangen, auf die der Restaurant-Recommendier zurückgreift.

### Grundlagen

Eine weitere Methode der Klassifikation stellen künstliche neuronale Netze dar. Mit ihnen wird versucht, biologische Netze als informationsverarbeitende Systeme nachzubilden. In den meisten Anwendungsgebieten dieser Netze ist herkömmliche Informatik mit Schleifen und Abfragen zu aufwendig, deshalb werden künstliche Neuronale Netze programmiert und trainiert. Eine vereinfachte Darstellung eines *künstlichen Neurons* ist in Abbildung 6.7 dargestellt. Ein künstliches Neuron ist ein Tupel

$$(\vec{x}, \vec{w}, f_a, f_o, o), \quad (6.7)$$

das aus einem Eingabevektor

$$\vec{x} := (x_1, \dots, x_n), \quad (6.8)$$

einem Gewichtungsvektor

$$\vec{w} := (w_1, \dots, w_n), \quad (6.9)$$

einer Aktivierungsfunktion

$$f_a : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \quad (6.10)$$

und einer Ausgabefunktion

$$f_o : \mathbb{R} \rightarrow \mathbb{R} \quad (6.11)$$

besteht. Das Ausgabesignal wird durch

$$f_o(f_a\{\vec{x}, \vec{w}\}) = o \quad (6.12)$$

gebildet und kann an weitere Neuronen weitergeleitet werden [Lip04]. Die häufigste Aktivie-

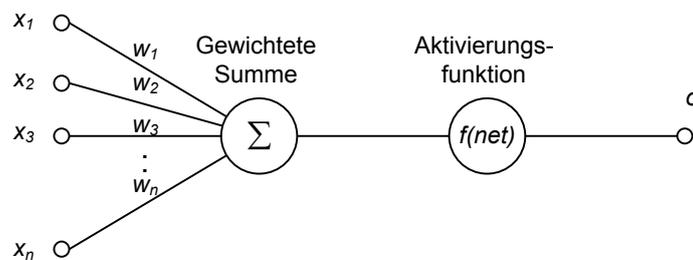


Abbildung 6.7: Darstellung eines Neurons

rungsfunktion ist die gewichtete Summe:

$$f_a\{\vec{x}, \vec{w}\} = \sum_{i=1}^n x_i w_i \quad (6.13)$$

Der Wertebereich der Ausgabe ist abhängig von der Ausgabefunktion und kann einen kontinuierlichen – alle reellen Zahlen oder Intervall – oder einen diskreten Wertebereich aufweisen. Dies wird durch unterschiedliche Ausgabefunktionen realisiert. In Abbildung 6.8 ist eine Schwellenwertfunktion mit binärem Ausgabebereich dargestellt und darunter die entsprechende Ausgabefunktion.

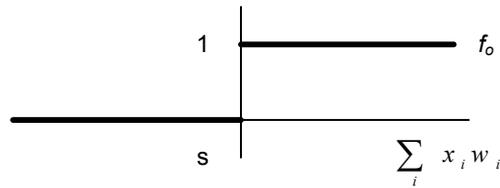


Abbildung 6.8: Binäre Schwellenwertfunktion

$$f_o \left( \sum_{i=1}^n x_i w_i \right) = \begin{cases} 1 & \text{falls } \sum_{i=1}^n x_i w_i \geq S \\ 0 & \text{sonst} \end{cases} \quad (6.14)$$

In Abbildung 6.9 ist eine semilineare Ausgabefunktion mit Sättigung realisiert.

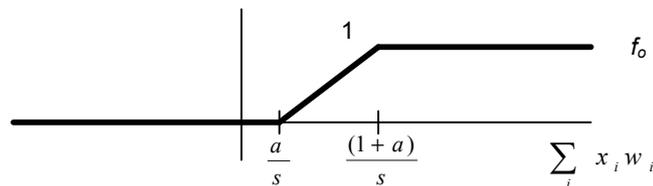


Abbildung 6.9: Semilineare Ausgabefunktion

$$f_o \left( \sum_{i=1}^n x_i w_i \right) = \begin{cases} 1 & \text{falls } \sum_{i=1}^n x_i w_i \geq \frac{1+a}{s} \\ s \left( \sum_{i=1}^n x_i w_i \right) - a & \text{falls } \frac{a}{s} < \sum_{i=1}^n x_i w_i < \frac{1+a}{s} \\ 0 & \text{sonst} \end{cases} \quad (6.15)$$

Weitere kontinuierliche Funktionen können unter anderem in Form von Sinus bis Sättigung, Tangens Hyperbolicus oder als logistische Funktion realisiert werden. Der Verbund mehrerer Neuronen bildet ein Neuronales Netz. Sie bestehen im Allgemeinen aus folgenden Komponenten [Zel94]:

- *Neuronen* mit Aktivierungszustand, Aktivierungsfunktion und Ausgabefunktion (vgl. oben)
- *Verbindungsnetzwerk der Zellen:*  
Ein neuronales Netz kann als gerichteter, gewichteter Graph angesehen werden. Die Kanten stellen die gewichteten Verbindungen ( $w_{ij}$  – von Neuron  $i$  nach Neuron  $j$ ) zwischen den Neuronen dar. Die Matrix  $W$  aller Verbindungen heißt Gewichtsmatrix.
- *Propagierungsfunktion:*  
Sie gibt an, wie sich die Netzeingabe eines Neurons aus den Ausgaben der anderen Neuronen und den Verbindungsgewichten berechnet. Dabei handelt es sich um die gewichtete Summe der Ausgaben der Vorgängerzellen.
- *Lernregel:*  
Die Lernregel ist ein Algorithmus, gemäß dem das neuronale Netz für eine vorgegebene Eingabe eine gewünschte Ausgabe zu produzieren lernt

Viele in der Praxis verwendete Neuronale Netze haben Verbindungen in nur eine Richtung, also ohne Rückkopplung. Die Verbindungen dieser sogenannten feedforward-Netze (Abbildung 6.10) gehen von der Eingabeschicht zur Ausgabeschicht. Angeordnete Neuronen zwischen diesen Schichten dienen der Informationsverarbeitung innerhalb des neuronalen Netzes und werden als verdeckte Schichten (Hidden Layers) bezeichnet, weil sie für einen außenstehenden Betrachter nicht zu sehen sind. Die Bezeichnung 2-, 3- oder  $n$ -schichtig geht mit der Anzahl an Layer mit trainierbaren Verbindungen einher. Ein 3-schichtiges Netz besteht folglich aus einem Input Layer, zwei Hidden Layer und einem Output Layer. Netze mit Rückkopplung, sogenannte recurrente Netze, weisen hingegen auch Verbindungen auf, die in Richtung der Eingabeschicht „rückwärtsgerichtet“ sind.

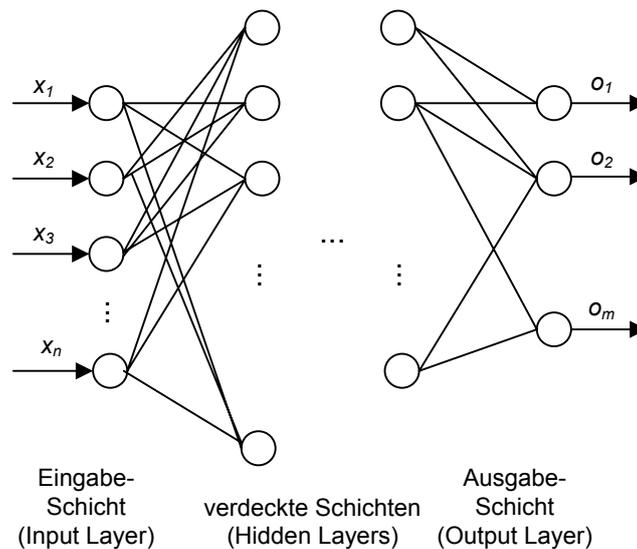


Abbildung 6.10:  $n$ -schichtiges feedforward-Netz mit verdeckten Schichten

Die Lernregel ist die wichtigste Komponente von neuronalen Netzen. Sie ermöglicht es, dass ein Netz weitgehend selbstständig eine Aufgabe anhand von Beispielen lernt. Prinzipiell kann der Lernprozess aus folgenden Lernmechanismen bestehen, wobei Punkt 3 die mit Abstand am Häufigsten verwendete ist:

1. Entwicklung neuer Verbindungen
2. Löschen existierender Verbindungen
3. Modifikation der Stärke  $w_{ij}$  von Verbindungen
4. Modifikation des Schwellenwertes von Neuronen
5. Modifikation der Aktivierungs-, Propagierungs- oder Ausgabefunktion
6. Entwicklung neuer Zellen
7. Löschen von Zellen

Eine weitere Unterscheidung besteht in der Art des Lernparadigmas, wobei sich im Wesentlichen drei Typen unterscheiden lassen:

- **Überwachtes Lernen (supervised learning):**  
Ein externer Lehrer gibt dem System zu jeder Eingabe die korrekte Ausgabe. Über die Minimierung der Abweichung wird das System angelernt. Voraussetzung ist die Existenz von Trainingsdaten aus Paaren von Ein- und Ausgaben
- **Bestärkendes Lernen (reinforcement learning):**  
Der Unterschied zum supervised learning besteht darin, dass dem System nur eine Angabe zur Richtigkeit der Ausgabe gemacht wird, ohne Zielwerte anzugeben.
- **Unüberwachtes Lernen (unsupervised learning):**  
Bei unüberwachten Lernmethoden existieren keine Angaben, ob und bis zu welchem Grad falsch klassifiziert wurde. Der Algorithmus versucht selbstständig ähnliche Eingabevektoren zu identifizieren und auf Gruppen ähnlicher oder benachbarter Neuronen abzubilden.

### Probabilistische Neuronale Netze

Ein entscheidender Nachteil eines Großteils der neuronalen Netze ist, dass sie eine sehr große Anzahl von Trainingsdaten benötigen. Dies ist begründet in der Art des Lernens. Sind zu wenig Datensätze vorhanden, kann es passieren, dass das Netz diese auswendig lernt (*Overfitting*) und falsche Ergebnisse für neue errechnet. Bedenkt man, dass der Benutzer eines Fahrzeugs nicht bei jeder Gelegenheit z. B. ein Restaurant aufsucht, vielleicht zwei bis drei mal in der Woche, dann würde das Anlernen des Netzes einen sehr langen Zeitraum benötigen. Dieses sogenannte Cold-Start-Problem muss unbedingt verhindert werden, um schon nach den ersten Restaurantbesuchen ein Verhaltensmuster feststellen zu können. Außerdem erreichen viele Netze nur eine Klassifizierung nach ja oder nein, richtig oder falsch, oder durch das Auffinden von Gruppenzugehörigkeiten. Eine derartige Auswertung ist im Automobil kaum brauchbar. Werden keine Restaurants oder nur eins in der aktuellen Umgebung des Benutzers gefunden, das seinen Präferenzen entspricht, hat er nicht die Möglichkeit andere auszuwählen, da keine weiteren oder gar keine als Alternative erkannt werden. Aus diesem Grund ist es sinnvoll, die Eigenschaften der Restaurants im aktuellen geografischen Umfeld mit den bereits besuchten zu vergleichen und dementsprechend für jedes Objekt eine Ähnlichkeit zu berechnen. So kann eine Sortierung einer dargestellten Ergebnisliste erreicht werden, die an oberster Stelle diejenigen mit der größten Ähnlichkeit zu denen in der Vergangenheit ausgewählten enthält. Diese Möglichkeit bieten Probabilistische Neuronale Netze (PNN). Dabei handelt es sich um einen Klassifikator nach dem bayes'schen Prinzip. Ein entsprechender Klassifikator berechnet aus den Mustervektoren für jede Klasse eine Wahrscheinlichkeitsverteilung der Musterklassen. Weitere Einzelheiten zu PNNs finden sich in [Rig94].

## 6.3 Konzeption der Informationsagenten

Nachfolgend wird nun die Konzeption der realisierten Informationsagenten beschrieben. Dies sind der Tank-, Kontakt-, Terminagent und Restaurant-Recommendier.

### 6.3.1 Tankagent

An dieser Stelle wird die Funktionsweise des *Tankagenten* vorgestellt und dessen intelligente, benutzeradaptive Verarbeitungsmethoden beschreiben (siehe auch [APRR07]).

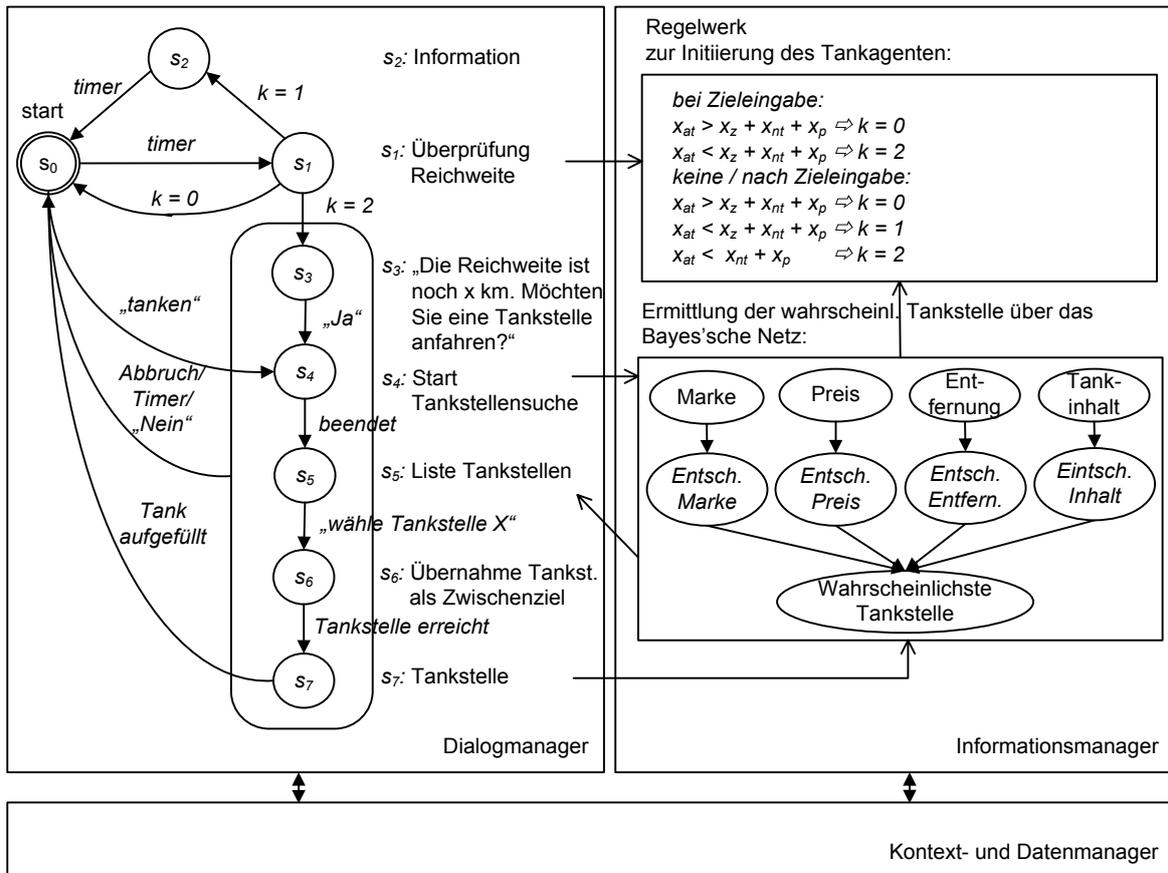


Abbildung 6.11: Modellierung des Tankagenten: Dialogmanager und Informationsmanager

## Aufgaben

Wie der Name vermuten lässt, unterstützt der Tankagent den Fahrer bei allen Aspekten, die mit dem Vorgang *Tanken* in Verbindung stehen. Neben der Gestaltung des adaptiven Dialogs (siehe Abschnitt 3) liegt hier der Hauptaugenmerk auf der geeigneten Datenvorfilterung. Der Tankagent kann sowohl explizit vom Benutzer aufgerufen werden als auch selbst die Initiative ergreifen. Es gilt gemäß dem Informationsfluss (siehe Abbildung 1.1), die Fragestellungen nach dem *Was bzw. Welche* (bevorzugte Tankstelle) und dem *Wann* (Timing) unter Berücksichtigung von Nutzerpräferenzen zu beantworten. Der Agent muss also entscheiden, wann der Nutzer benachrichtigt werden soll und welche Tankstellen in Frage kommen.

## Ein- und Ausgabedaten

Um seine Aufgabe erfüllen zu können, müssen dem Tankagenten Informationen über die *Entfernungen* zum aktuellen Zwischen- bzw. Endziel bereitgestellt werden. Neben dem *maximal möglichen Tankinhalt* wird auch ein *mittlerer Verbrauchswert* benötigt. Nach einer entsprechenden Anfrage an das Datenmanagement-System (siehe Abschnitt 3.5.3) bekommt der Tankagent eine Vorauswahl an Tankstellen übermittelt. Darin enthalten sind Informationen der Tankstellen bezüglich *Entfernung*, *Luftlinienentfernung*, *Tankstellenmarke*, *Spritpreis*, *Spritart*, *Adresse bzw. Ort* und nicht zuletzt über den *Umweg*, den das Anfahren der Tankstelle

verursachen würde. Als Nächstes bewertet der Tankagent, mit welcher Wahrscheinlichkeit der Fahrer eine bestimmte Tankstelle im gegebenen Kontext wählen würde.

### Funktionsweise

Der Tankagent überwacht fortwährend den aktuellen *Tankinhalt*  $v_{at}$  und gleicht die daraus berechnete *Restreichweite*  $x_{at}$  mit der Entfernung zu vorhandenen (Zwischen-)Zielen  $x_z$  ab. Hierzu wird ein einfaches Regelwerk verwendet (siehe Abbildung 6.11). Auch die Entfernung einer nächsten Tankstelle am Ziel  $x_{nt}$  unter Berücksichtigung eines Sicherheitspuffers muss  $x_p$  berücksichtigt werden. Reicht der Tankinhalt  $v_{at}$  nicht, um ein vorgegebenes Ziel incl. nächster Tankstelle und Puffer zu erreichen, so teilt er dies dem Fahrer mit und fragt diesen, ob nach Tankstellen gesucht werden sollte. Abhängig von den Tankvorlieben des Fahrers wird der Tankagentendialog entweder sofort beim Einsteigen (z. B. über den Naviagenten) initiiert oder beim Erreichen einer Restspritmenge  $v_s$  bzw. Restreichweite  $x_s$ , die der Agent als wahrscheinlichste Schwelle, bei der er tanken möchte, interpretiert. Diese Schwelle ist abhängig vom Kontext; es können die Straßenart (z. B. Autobahn, Nebenstraße, usw.), die Entfernung der Tankstelle und weitere Faktoren in die Entscheidung mit einfließen.

Wurde die Suche nach Tankstellen veranlasst, so stellt der Tankagent eine Anfrage an das Datenmanagement-System und erhält daraufhin eine Vorauswahl mit den zehn nächsten Tankstellen übermittelt. Mit Hilfe des Bayes'schen Netzes (siehe Abschnitt 6.3.1) werden nun diese Tankstellen bewertet und die Daten anschließend als Auswahl angeboten. Wählt der Fahrer eine Tankstelle aus, so interpretiert der Tankagent das Nutzerverhalten und aktualisiert daraufhin das Bayes'sche Netz nach Verarbeitung mit einem Lernfaktor  $L$ . Über die Zeit passt sich der Tankagent dem Nutzerverhalten immer besser an und kann die Fahrerintentionen zunehmend besser antizipieren.

**Tankinhaltsbereiche:** Durch eine Voruntersuchung wurde die Annahme bestätigt, dass der Fahrer mit zunehmender Wahrscheinlichkeit tankt, je weniger Sprit noch verfügbar ist. Aus diesem Grund wird der Tankinhalt in mehrere diskrete Bereiche eingeteilt, die um so kleiner werden, je geringer der Restinhalt im Tank ist. Dies lässt sich mit Hilfe folgender Exponentialfunktion

$$f(x) = \exp(-5x) \quad (6.16)$$

und dem empirisch ermittelten Faktor 5 im Intervall  $[0; 1]$  darstellen.

Dieses Intervall wird in zehn gleich große Bereiche eingeteilt. Zunächst multipliziert man dazu  $1/9$  mit den ganzen Zahlen im Wertebereich  $\{0; 1; 2; \dots; 9\}$ . Setzt man die Ergebnisse in Gleichung 6.16 ein, erhält man so die obersten zehn Bereichsgrenzen. Multipliziert man diese noch mit dem maximal möglichen Tankinhalt (hier als 65 Liter) und setzt man die unterste, elfte Bereichsgrenze gleich 0, so erhält man die Werte aus Abbildung 6.12.

### Einsatz des Bayes'schen Netzes

Aus einer Voruntersuchung hat sich ergeben, dass die vier Kriterien *Marke*  $m_{nt}$ , *Entfernung*  $x_{nt}$ , *Preis*  $p_{nt}$  und *Restspritmenge des Puffers*  $v_p$  bei der Auswahl der nächstgelegenen Tankstelle eine Rolle spielen. Unter *Restspritmenge des Puffers*  $v_p$  ist die Spritmenge zu verstehen, die der Fahrer beim Erreichen der Tankstelle noch im Tank hat. Diese variiert entsprechend den Fahrerpräferenzen. Die vier Eingangskriterien werden in dem Bayes'schen Netz für den

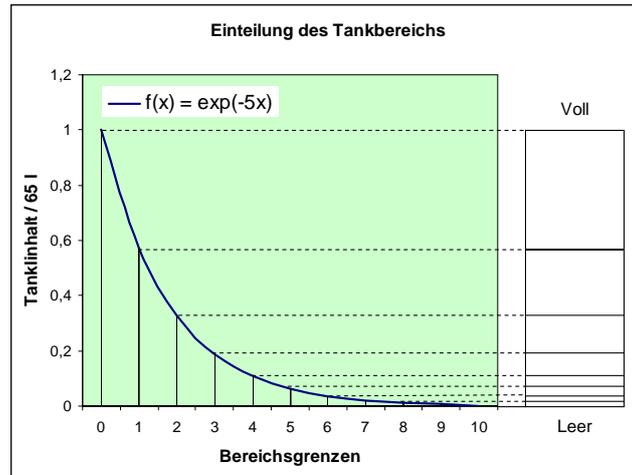


Abbildung 6.12: Exponentialfunktion zur Einteilung des Tanks in festgelegte Bereiche

Tankagenten (siehe Abbildung 6.11) verarbeitet:

$$B\vec{N}_n = \begin{pmatrix} \text{Marke } m_{nt} \\ \text{Entfernung } x_{nt} \\ \text{Preis } p_{nt} \\ \text{Restspritmenge } v_p \end{pmatrix} \quad (6.17)$$

Zunächst werden alle kontinuierlichen Werte in diskrete Bereiche überführt. Entfernung, Preis und Restspritmenge werden in zehn Bereiche, Marke in sechs Ränge eingeteilt. Berücksichtigt wurden die Marken Aral, BFT, Esso, Jet, OMV und Shell.

In der zweiten Ebene (Knoten *Entscheidung nach Kriterium*) werden die diskreten Bereiche der darüber liegenden Kriterien bezüglich den Nutzerpräferenzen bewertet. Im Knoten *Wahrscheinlichste Tankstelle* werden die vier Kriterien entsprechend zusammengeführt und schließlich die Gesamt-Wahrscheinlichkeit berechnet, mit welcher der Fahrer die jeweilige Tankstelle wählen würde.

Zunächst werden die Kriterien der angefragten Tankstellenliste untereinander verglichen und in die entsprechenden Bereiche und Ränge eingeteilt. Ist eine Tankstelle beispielsweise die am weitesten entfernte, entspricht dies dem Entfernungsrang 10, ist sie die drittgünstigste, erhält sie Preisrang 3 usw. Daraufhin werden im Bayes'schen Netz entsprechend der vorangegangenen Einteilung Evidenzen in den Kriterien vergeben. Die Werte für die Wahrscheinlichkeitsverteilungen wurden auf Grundlage der Voruntersuchung berechnet. In Tabelle 6.6 und Tabelle 6.7 ist ersichtlich, welcher Rang von den Versuchspersonen wie oft „gewählt“ wurde. Aus den relativen Häufigkeiten werden die bedingten Wahrscheinlichkeiten CPTs gebildet. Beispielhaft ist diese für den Tankinhalt in Tabelle 6.8 angegeben. Um die Intention des Nutzers berechnen zu können, bedarf es einer mathematischen Umsetzung seiner Präferenzen. Hierzu wird die Standardabweichung der relativen Häufigkeiten herangezogen, da diese um so größer ist, je öfter der Nutzer einen bestimmten Rang wählt, d. h. je wichtiger für ihn ein bestimmtes Kriterium ist. So ist die Standardabweichung für den Preis wesentlich höher als für die Marke.

Nach einer Normierung können die Standardabweichungen – beides dargestellt in Tabelle 6.9 – für die Wahrscheinlichkeitsverteilungen im Knoten *Wahrscheinlichste Tankstelle* (siehe Tabelle

Rang	Preis	Entfernung	Tankinhalt
1	128	18	7
2	27	13	4
3	21	10	10
4	10	12	11
5	4	12	23
6	5	12	21
7	4	16	36
8	1	26	51
9	0	41	26
10	0	40	11

Tabelle 6.6: Häufigkeit der Wahl eines Rangs der Kriterien (aus Voruntersuchung)

Marke	Aral	BFT	Esso	Jet	OMV	Shell
Anzahl	41	30	24	34	34	37

Tabelle 6.7: Häufigkeit der jeweils gewählten *Marke* (aus Voruntersuchung)

$P(N B)$ Tankinhaltsbereich B	Nutzerintention N	
	ja	nein
01	0,035	0,965
02	0,020	0,980
03	0,050	0,950
04	0,055	0,945
05	0,115	0,885
06	0,105	0,895
07	0,185	0,815
08	0,250	0,750
09	0,130	0,870
10	0,055	0,945

Tabelle 6.8: Wahrscheinlichkeitstabelle für den Knoten *Entscheidung nach Tankinhalt*

6.10) verwendet werden. Dazu werden die normierten Werte entsprechend den auftretenden Zuständen addiert.

Kriterium	Entfernung	Preis	Marke	Tankinhalt
Standartabw.	0,055	0,185	0,027	0,069
normiert	0,165	0,550	0,079	0,206

Tabelle 6.9: Standardabweichung und Normierung für die Gewichtung der Kriterien

Um sich besser an die Intentionen des Nutzers anpassen zu können, wird ein Lernalgorithmus benötigt. Diese relativ einfache Rechenvorschrift ist mit einem *Lernfaktor*  $L$  versehen, der je

<i>Entscheidung nach Kriterium</i>				<i>CPT</i>	
Entfernung	Marke	Preis	Tankinhalt	ja	nein
ja	ja	ja	ja	1	0
ja	ja	ja	nein	0,79	0,21
ja	ja	nein	ja	0,45	0,55
ja	ja	nein	nein	0,24	0,76
ja	nein	ja	ja	0,92	0,08
ja	nein	ja	nein	0,71	0,29
ja	nein	nein	ja	0,37	0,63
ja	nein	nein	nein	0,17	0,83
ja	ja	ja	ja	0,83	0,17
ja	ja	ja	nein	0,63	0,37
ja	ja	nein	ja	0,29	0,71
ja	ja	nein	nein	0,08	0,92
ja	nein	ja	ja	0,76	0,24
ja	nein	ja	nein	0,55	0,45
ja	nein	nein	ja	0,21	0,79
ja	nein	nein	nein	0	1

Tabelle 6.10: CPT für den Knoten *Wahrscheinlichste Tankstelle*

nach gewünschter Adaptionsgeschwindigkeit aus dem Bereich  $\{0,01; 0,02; c; 0,10\}$  (langsame > schnelle Anpassung) gewählt werden kann. Für jedes Kriterium wird der neue Wert für den Rang, den der Nutzer gewählt hat, mit folgender Vorschrift bestimmt:

$$Rang_{neu} = \frac{Rang_{alt} + L}{1 + L}. \quad (6.18)$$

Für die Berechnung der neuen Werte für die verbleibenden, nicht vom Nutzer gewählten Ränge, findet nachfolgende Formel Verwendung:

$$Rang_{neu} = \frac{Rang_{alt}}{1 + L} \quad (6.19)$$

Hat der User beispielsweise die Tankstelle gewählt, bei deren Erreichen die verbleibende Spritmenge in den Bereich 9 eingeordnet wird, so ändert sich die CPT für den Knoten *Entscheidung nach Entfernung* aus Tabelle 6.8 gemäß Tabelle 6.11.

Danach werden von Neuem die Standardabweichungen wie oben geschildert kalkuliert und anschließend normiert. Das Bayes'sche Netz ist somit aktualisiert.

### 6.3.2 Kontaktagent

Ähnlich wie der Tankagent hat der Kontaktagent die Aufgabe, aus dem Verhalten des Nutzers zu lernen. Von besonderer Wichtigkeit ist die Verarbeitung unscharfer Eingaben, um damit eine Vorhersage der Intentionen des Fahrers zu ermöglichen. Eine Spracheingabe, wie beispielsweise „Ruf mal bei Person X an!“ kalkuliert daraufhin die Wahrscheinlichkeiten für einen Anruf auf das Handy, die Privat- oder die Büronummer.

$P(N B)$ Tankinhaltbereich B	Nutzerintention N	
	ja	nein
01	0,032	0,968
02	0,018	0,981
03	0,045	0,955
04	0,050	0,950
05	0,105	0,895
06	0,095	0,905
07	0,168	0,831
08	0,227	0,773
09	0,209	0,791
10	0,050	0,950

Tabelle 6.11: Aktualisierte Wahrscheinlichkeitstabelle für den Knoten EntschNachTankinhalt

### Daten-Input und Output

Wird der Kontaktagent initiiert, so müssen ihm Informationen bezüglich des aktuellen Wochentages und der aktuellen Uhrzeit bereitgestellt werden. Des Weiteren muss ihm mitgeteilt werden, ob zu dem zu kontaktierenden Kommunikationspartner ein Termin hinterlegt ist. Zusätzlich sind der Terminstatus und die Terminart von Bedeutung. Der Status hängt davon ab, ob der Termin pünktlich oder verspätet erreicht werden wird oder ob er bereits verpasst wurde. Bei der Terminart wird zwischen den beiden Optionen geschäftlich und privat unterschieden. Zuletzt erhält der Kontaktagent noch eine Rückmeldung über das vom Nutzer gewählte Medium, um das Bayes'sche Netz, das hier zum Einsatz kommt, zu aktualisieren (siehe folgender Abschnitt). Der Kontaktagent liefert die Wahrscheinlichkeiten für die Verwendung fünf möglicher Kommunikationsmedien wie die Privatnummer, Büronummer, Mobiltelefon, SMS und E-Mail.

### Interne Verarbeitung durch Bayes'sches Netz

Zur internen Berechnung der Wahrscheinlichkeiten kommt das Bayes'sche Netz aus Abbildung 6.13 zum Einsatz. Entsprechend der Attribute, die für die Kriterien *Wochentag*, *Uhrzeit*, *Terminart* und *Terminstatus* übermittelt wurden, werden in diesem Netz Evidenzen vergeben. Für ein Kriterium muss nicht zwingend eine Eigenschaft zugewiesen werden; es wurde dann sozusagen keine Beobachtung für diesen Zustand gemacht. Will der User mit einem Gesprächspartner in Kontakt treten, muss er eine Anweisung geben. Entweder er entscheidet sich explizit für eines der fünf Medien – Handy (Mobiltelefon), private Festnetznummer (privat), Festnetznummer im Büro, Kurznachricht auf das Handy (SMS) und E-Mail – oder er äußert nur einen unklaren Befehl. Im letzteren Fall gibt es die Option, dass er irgendeine der fünf Methoden (*Kontakt\_allgemein*) zur Kontaktaufnahme wünscht oder dass er zwar ein Telefonat führen möchte, sich aber noch nicht auf ein Medium festgelegt hat (*Anruf\_allgemein*). Es ergäbe sich noch die Möglichkeit, an dieser Stelle auch eine Anweisung *schriftliche\_Benachrichtigung* einzuführen. Da in einer Voruntersuchung jedoch kaum die Möglichkeit zum Schreiben einer E-Mail genutzt wurde und die SMS wie oben beschrieben direkt ausgewählt werden kann, findet dieser zusätzliche Befehl keine Verwendung. Die Uhrzeit wurde in sechs Bereiche unterteilt, um zum Einen der Tatsache gerecht zu werden, dass die Wahl eines Mediums davon

abhängt, ob ein Anruf während der Arbeitszeit, der Mittagszeit oder nach Feierabend, usw. getätigt werden soll. Zum Anderen sollen durch diese Klassifikation die möglichen Knoten für die Uhrzeiten im Bayes'schen Netz eingegrenzt werden.

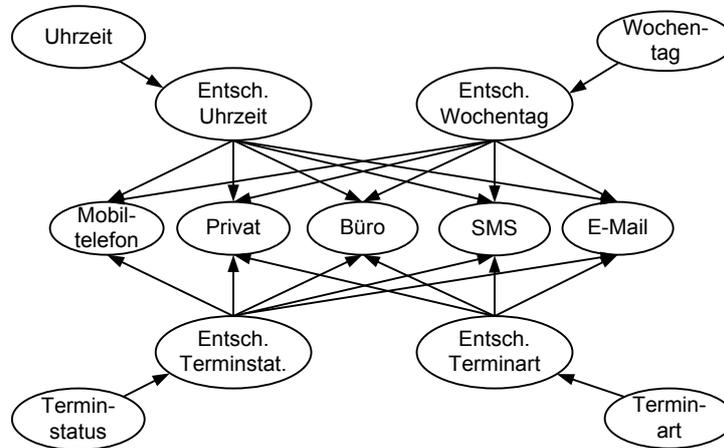


Abbildung 6.13: Bayes'sches Netz des Kontaktagenten

Für die Berechnung der relativen Häufigkeiten werden jeweils die Werte der Medien, die zum selben Attribut gehören, berechnet. Das bedeutet, dass beispielsweise jeweils die Häufigkeiten aller Medien nur für Dienstag oder nur für private Termine kalkuliert werden. Daraus ergeben sich die Tabellen mit den Wahrscheinlichkeitsverteilungen.

Die Knoten *Mobiltelefon*, *Privat*, *Büro*, *SMS* und *E-Mail* bestehen jeweils aus einer CPT mit 112 Zeilen. Die Illustration an dieser Stelle würde allerdings den Rahmen sprengen. Zur Gewichtung gehen die vier Kriterien mit 0,25 ein und werden den Zuständen entsprechend addiert. Wählt der Nutzer ein konkretes Medium, wird dies durch die Zuteilung von 1,0 für das in der Anweisung genannte Medium und von 0,0 für die anderen vier Medien berücksichtigt. Die Aktualisierung erfolgt nach demselben Lernalgorithmus, wie für das Bayes'sche Netz beim Tankagenten.

### 6.3.3 Terminagent

Der Terminagent verwaltet die Einträge im Terminplaner und im Adressbuch. Er prüft, ob anstehende Termine eingehalten werden können und informiert den Nutzer, falls dieser einen Termin voraussichtlich verpassen wird oder der Termin bereits verpasst wurde. Spricht der Nutzer den Befehl „Ruf mal an!“ aus, so überprüft der Assistent, ob sich aus dem Kontext das Ziel des Anrufs ergibt. Bei einer kurz zuvor empfangenen SMS könnte der Fahrer den Absender der Kurzmitteilung kontaktieren wollen. Falls der Nutzer gerade im Terminplaner oder Adressbuch blättert, möchte er wohl die zum aktuell ausgewählten Eintrag gehörige Person anrufen. Ist aus dem Kontext keine erfolgreiche Schlussfolgerung möglich, erfolgt eine explizite Nachfrage beim Nutzer und der Terminagent meldet die Daten dem Kontaktagenten. In seiner Funktion als Schnittstelle präsentiert er das vom Kontaktagent berechnete wahrscheinlichste Medium dem Nutzer und leitet das von diesem letztlich gewählte Medium an den genannten Agenten weiter.

### Datenempfang und Datenversand

Aus der Datenbank holt sich der Agent alle Informationen über gespeicherte Termine. Neben Datum und Uhrzeit sind dabei vor allem die Art des Termins (privat oder geschäftlich) interessant. Zudem werden die voraussichtliche Ankunftszeit und die aktuelle Uhrzeit benötigt. Um eine eventuelle Verzögerung einkalkulieren zu können, muss dem Terminagenten zusätzlich mitgeteilt werden, ob ein Tankstopp geplant ist. Vom Kontaktagenten erhält er zudem Infos über die Wahrscheinlichkeiten der verschiedenen zur Verfügung stehenden Medien, um diese an den Nutzer weiterleiten zu können. In der Realität wäre noch die Versendung vorgefertigter SMS bzw. E-Mail-Nachrichten denkbar, die eine Verspätung o.ä. ankündigen. Die vom Kontaktagenten für dessen Berechnungen benötigten Daten (Terminart, -status, Wochentag, Uhrzeit) erhält dieser vom Terminagent. An diesen Agenten sendet er für dessen Aktualisierung auch die Informationen zum vom User ausgewählten Medium zurück.

### Interne Abläufe

Wird das System gestartet, gleicht der Terminagent vorhandene Termine mit der aktuellen Uhrzeit ab und informiert den Nutzer über verpasste Termine. Wird dem Naviagenten vom Nutzer ein Ziel mitgeteilt, für das ein Termin existiert, so gleicht der Terminagent die voraussichtliche Ankunftszeit mit dem Zeitpunkt des Termins ab und informiert den Fahrer, falls dieser sich verspäten wird oder den Termin verpasst hat. Verpasst bedeutet dabei, dass die aktuelle Uhrzeit größer ist als der Zeitpunkt des Termins. Der Agent vermutet eine Verspätung, wenn seine Berechnungen nach der Formel

$$T_{\text{Zeitspanne}} = t_{\text{Termin}} - (t_{\text{akt.Uhrzeit}} + T_{\text{Fahrtdauer}}) \quad (6.20)$$

ergeben, dass zwischen voraussichtlicher Ankunftszeit und des Termins weniger als 15 Minuten bleiben. Ist diese Zeitspanne größer, nimmt der Agent eine pünktliche Ankunft an. Wird vom Nutzer ein Tankstopp eingeplant, berücksichtigt der Terminagent dies und rechnet die für diesen Stopp benötigte Zeit mit ein. Dabei greift er auf Erfahrungswerte zurück, die ihm für die Dauer eines solchen Vorgangs vorliegen. Durch die gegebenen Einschränkungen des Simulators ist die Zeit für einen Tankaufenthalt nicht dynamisch. Allerdings kann der statische Wert für einen Stopp (Tankstellenan- und -abfahrt, Tankvorgang) manuell zur Laufzeit des Systems geändert werden. Als Defaultwert ist ein Zeitraum von 20 Minuten eingestellt. Liegt einer der beiden Zustände verspätet oder verpasst vor, so wird eine Anfrage an den Fahrer gestellt, ob und wie der Termin zu benachrichtigen ist. Möchte der User den Kontakt über z. B. eine Verspätung informieren, sendet der Terminagent die nötigen Informationen an den Kontaktagenten und erhält daraufhin die Wahrscheinlichkeiten, mit der der Nutzer die verschiedenen Medien wählen würde. Anschließend wird der Fahrer darüber informiert, dass eine Verbindung mittels der wahrscheinlichsten Methode initiiert wird, woraufhin dieser fünf Sekunden Zeit hat, den Agenten zu überstimmen und ein anderes Medium zu wählen. Genauso wird auch verfahren, wenn der Benutzer einen Eintrag aus dem Adressbuch kontaktieren möchte. Als letztes wird die Art des gewählten Mediums dem Kontaktagenten zu dessen Aktualisierung mitgeteilt.

### 6.3.4 Restaurant-Recommender

Hier wird das entwickelte Restaurant-Recommender-System vorgestellt. Zunächst wertet eine Suchmaschine die in der aktuellen Umgebung befindlichen Restaurants nach realer Entfer-

nung und dem zurückzulegenden Umweg bei Bekanntheit einer Route aus. Das Recommender-System sortiert diese entsprechend der in der Vergangenheit getätigten Auswahl. Dabei greift das System auf eine Datenbasis mit Restaurants zurück, die mit Eigenschaften hinterlegt sind. Anhand der Analyse der Eigenschaften der bereits ausgewählten Restaurants wird ein Probabilistisches Neuronales Netz angelernt.

### Relevante Restaurants

Die Identifikation relevanter Daten für Restaurants gestaltet sich ungemein schwieriger. Aufgrund der Tatsache, dass es kaum größere Restaurant-Ketten gibt und auch keinen standardisierten Internetauftritt, ist die Information über diese teilweise unvollständig oder einfach nicht vorhanden. Aus dem Namen des Restaurants lässt sich allein die Art der Küche ableiten. Zahlreiche Restaurant-Führer und Internetportale versuchen diesem Defizit entgegenzuwirken. Viele Internetportale vertrauen darauf, dass Besucher Restaurants eintragen und eine kurze Bewertung abgeben, wie zum Beispiel *restaurantfuehrer24*<sup>2</sup>. Allerdings variieren die gespeicherten Daten, da nur subjektive Auffälligkeiten registriert werden und somit oft objektiv vergleichbare fehlen. Restaurant-Führer, die in Buchform zu erwerben sind, sind auch über das Internet zu erreichen. Ihre Vorteile liegen in einer vollständigen und vereinheitlichten Informationswiedergabe und in Bewertungen, die von einem festgelegten Team durchgeführt werden. Beliebte Anbieter sind z. B. ViaMichelin, Aral Schlemmer-Atlas, Varta-Führer, Der Feinschmecker Guide oder Gault Millau. Sie bilden eine objektive Grundlage für Details über Restaurants, die in eine kontextspezifische Auswertung einfließen können. ViaMichelin z. B. setzt auf eine kurze Präsentation des Objekts mit Adresse, Öffnungszeiten, Preise, Art der Küche, Spezialitäten und einem kurzen Kommentar.

**Kontextsensitive Auswertung der Restauranteigenschaften:** Nachdem eine regionale Vorauswahl an verfügbaren Restaurants durchgeführt worden ist, können diese jetzt kontextsensitiv ausgewertet werden. Die kontextsensitive Betrachtung soll Benutzerpräferenzen berücksichtigen und situationsabhängig eine Sortierung der Ergebnisliste durchführen. Dies wurde zu Anfang der Arbeit für die automotive Umgebung postuliert, um Bedienzeiten zu reduzieren und dementsprechend eine geringere Ablenkung von der eigentlichen Fahraufgabe zu erreichen. Eine Darstellung der Restaurants, geordnet nach Entfernung, Preiskategorie oder Alphabet, kann diese Bedingung auch bei wenigen Ergebnissen kaum erfüllen. Aus diesem Grund sollen persönliche Vorzüge und situative Zusammenhänge mit in die Aufbereitung der Darstellung einfließen.

### Einsatz des Probabilistischen Neuronalen Netzes

Nun wird die Umsetzung des Probabilistischen Neuronalen Netzes beschrieben (siehe auch Abschnitt 6.2.2 und [Rig94]). Die restaurant-spezifischen Eigenschaften, die in der Datenbank gespeichert sind, lassen sich in drei Kategorien unterteilen, die separat betrachtet werden müssen. Dazu zählen neben der Preiskategorie, die einen anderen Wertebereich aufweist (1 bis 5), auch die Art der Restaurants (*feinschmecker*, *weinlokal*, *biergarten*, *cafe*, *bistro*) und die Küche, die auf *international*, *deutsch*, *regional*, *italienisch*, *asiatisch*, *griechisch*,  *mexikanisch*, *vegetarisch* und *fastfood* beschränkt ist. Zusätzlich zu diesen Daten erhält der Restaurant-Recommender aus der regional eingeschränkten Suche der Suchmaschine die reale Entfernung vom aktuellen

---

<sup>2</sup>[www.restaurantfuehrer24.de](http://www.restaurantfuehrer24.de)

Standort. Diese Aufteilung muss auch in die Bewertung der Restaurants einfließen. Variieren zum Beispiel die Entfernungen der in der Vergangenheit ausgewählten Restaurants stark, so muss dieser Einfluss gering gehalten werden. Dies wird erreicht, indem das Netz in vier PNNs aufgeteilt wird, wobei jedes für die Wahrscheinlichkeitsberechnung einer anderen Eigenschaftskategorie zuständig ist. Die gemittelte Summe dieser Wahrscheinlichkeiten steht dann für die Ähnlichkeit einer Auswahl mit den in der History gespeicherten Restaurants. Wir setzen voraus, dass der Benutzer in der Vergangenheit 6 Restaurants ausgewählt hat, davon vier unterschiedliche einmal und zwei doppelt. Jedes dieser Restaurants stellt eine Klasse in einem PNN dar. Da allerdings zwei doppelt ausgewählt wurden, würde sich das Ergebnis dieser Berechnungen nicht unterscheiden und die Anzahl der Klassen kann auf 4 reduziert werden. Das bedeutet, dass wir vier Restaurants (die Beispielvektoren) und vier Klassen, jede Klasse ein Beispielvektor, haben. Die Beispielvektoren  $\vec{x}_i$  sind zugleich identisch mit den Gewichtungsvektoren  $\vec{w} = \vec{x}_i$ . Für jeden dieser Beispielvektoren ist ein Neuron zuständig. Jedes Neuron summiert seine gewichteten Eingänge und leitet die Summe an die Aktivierungsfunktion weiter. Der Bruch  $1/I_n$  und die Summe fallen raus, da die Anzahl  $I_n$  an Beispielvektoren für jede Klasse 1 ist. Das Ergebnis stellt eine Approximation an die Ähnlichkeit von Vektoren dar. Bei entsprechender Wahl von Sigma  $\sigma$  kann dieser Wert angepasst werden.

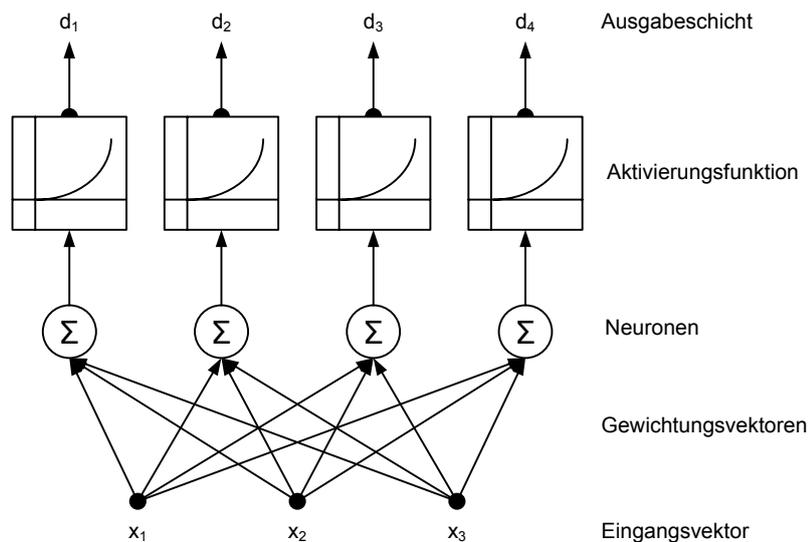


Abbildung 6.14: Abgewandelte Netz-Struktur für einen dreidimensionalen Eingangsvektor

In Abbildung 6.14 ist ebenfalls zu sehen, dass der Eingangsvektor aus drei Dimensionen besteht. Diese Reduktion ist aus Gründen der Übersichtlichkeit gemacht worden. Betrachtet man die Restaurantdaten, so erkennen wir vier Eigenschaftskategorien. Dies bedeutet, wie oben erklärt, dass die entsprechenden Ausgänge  $d_1, \dots, d_n$  dieser vier Probabilistischen Neuronalen Netze aufsummiert und durch die Anzahl der Eigenschaftskategorien  $Z$  dividiert werden müssen:

$$p_n(\vec{x}) = \frac{1}{Z} \sum_{j=1}^Z e^{-\frac{(\vec{x}-\vec{x}_j)^T \cdot (\vec{x}-\vec{x}_j)}{2\sigma^2}} \quad (6.21)$$

$p_n(\vec{x})$  ist dann ein Maß für die Ähnlichkeit eines neuen Restaurants mit den in der Vergangenheit ausgewählten, wobei  $n$  für die einzelnen Klassen steht. Folglich ergeben sich in diesem

Fall 4 Werte für  $p_1$  bis  $p_4$ . Das höchste Ergebnis gibt Aufschluss über die Zugehörigkeit zu einer Klasse. Diese Klasse bzw. das Restaurant ist dann als Referenz für die Ähnlichkeitsbewertung zu sehen. Ändert sich die Anzahl der Eigenschaftskategorien oder die Anzahl der Einträge, müssen dementsprechend die Anzahl der Verbundenen PNNs oder die Dimension des Eingangsvektors erhöht oder verringert werden.

**Training:** Das Lernverfahren wird mit der Festlegung der Anzahl an Klassen (entsprechend der Anzahl unterschiedlicher Beispielvektoren), der Herstellung der Verbindungen und deren Belegung mit den Gewichtungen in Form der Beispielvektoren durchgeführt. Außerdem muss eine Bestimmung der Unsicherheit  $\sigma$  der Beispielvektoren erfolgen. Dabei sollte berücksichtigt werden, dass ein Restaurant eventuell des öfteren besucht worden ist, da dies nicht in der Struktur des PNN eingeflossen ist. Über eine statistische Analyse der jeweiligen Eigenschaftskategorien der Restaurants lässt sich die Standardabweichung  $\sigma_x$  kalkulieren:

$$\sigma_x := \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (6.22)$$

$N$  ist dabei die Anzahl der Werte,  $x$  das arithmetische Mittel und  $x_i$  die Merkmalsausprägung am  $i$ -ten Element der Stichprobe. Das arithmetische Mittel ist definiert durch:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (6.23)$$

Diese Formeln gelten allerdings nur statistisch, d. h. bei einer Folge von Stichproben. Im Fall des PNN werden die Standardabweichungen der Vektoren über mittlere Abweichung der einzelnen Dimensionen berechnet. D. h. für jede Dimension  $J$  wird die Standardabweichung über alle Beispielvektoren berechnet, die Sigmas addiert und durch die Anzahl der Dimensionen geteilt.

$$\sigma = \frac{1}{J} \sum_{j=1}^J \sigma_j \quad (6.24)$$

**Verarbeitung:** Die Verarbeitung der Eingangsvektoren erfolgt nach der Lernphase des PNNs. Sind die Anzahl der Neuronen und die Verbindungen mit den Gewichtungen gesetzt und ist die Berechnung des Parameters  $\sigma$  abgeschlossen, können dem Netz Daten präsentiert werden. Dementsprechend werden die von der Suchmaschine gefundenen Restaurants schrittweise übergeben und für jedes Restaurant die Ähnlichkeitsbewertung der Klasse mit dem höchsten Ergebnis ausgewählt. Auf diese Weise erhält jedes Restaurant eine Bewertung beruhend auf den in der Vergangenheit ausgewählten und kann diese Information für die Aufbereitung der Darstellung verwenden.

### 6.3.5 Weitere Informationsagenten

Im Folgenden werden weitere mögliche Informationsagenten aufgezeigt.

- **Stauagent:**

Eng an den Naviganten gekoppelt, verarbeitet der Stauagent Informationen über Staus

und andere Verkehrsbehinderungen auf der geplanten Fahrtroute. Hat sich die Zielführung aufgrund einer Stauumfahrung geändert, werden diese Daten zusätzlich an den Tankagenten und den Terminagenten weitergeleitet. Auf die Umfahrroute wirken sich u. a. die Straßenbeschaffenheit und die Straßenart aus.

- **Ankunftsagent:**

Der Ankunftsagent sucht freie Parkmöglichkeiten am Zielort oder bietet Vorschläge an, dieses Ziel von außerhalb eines Ballungsraumes mit öffentlichen Verkehrsmitteln zu erreichen. Der Agent sucht dabei selbständig – unter Berücksichtigung der Vorlieben des Fahrers – die günstigste Verbindung (in Bezug auf Fahrpreis und -kosten). Eine weitere hilfreiche Funktion repräsentiert die Anzeige des verbleibenden Fußweges zur besseren Orientierung und die Übertragung dieser Informationen auf mobile Endgeräte (Handy, PDA, usw.). Gegebenenfalls wird mit Hilfe des Terminagenten eine Person am Fahrziel kontaktiert, um die Ankunft anzukündigen.

- **Klimaagent:**

Dieser Agent setzt unscharfe Eingaben des Nutzers unter Beachtung von Kontextgrößen um. Auf die Äußerung „Hier ist es kalt!“ werden z. B. erst die Fenster ganz oder teilweise geschlossen und eventuell wird die Klimaanlage höher gestellt. Blickt der Fahrer bei seiner Aussage zusätzlich in Richtung des Fußraums, interpretiert der Agent dies und erhöht nur die Temperatur für diesen Bereich.

- **Bahn-/Fluginfoagent:**

In den Aufgabenbereich dieses Agenten fallen die Bereitstellung von Informationen zu Flug- und Bahnverbindungen. Bei reservierten Reisen per Bahn oder Flugzeug teilt er dem Nutzer eventuelle Verspätungen mit, zeigt ihm Abflugterminal bzw. Gleis an und sucht mittels Naviagent eine Parkmöglichkeit in dessen Nähe. Auf Wunsch übernimmt er auch die Reservierung im Sinne des Nutzers.

## 6.4 Bewertung und Diskussion

Zur Bewertung der entstandenen Informationsagenten wurde auf einen virtuellen (Fahr-)Simulator und eine POI-Datenbank zurückgegriffen. Dadurch war es möglich, verschiedene Szenarien zu simulieren und das Verhalten der Nutzer zu studieren. Mit diesen Daten konnte die Funktionsweise der Informationsagenten ausgelegt und überprüft werden. Diese Untersuchung wurde für den Tankagenten, den Kontaktagenten und den Terminagenten durchgeführt. Der Tank- und Kontaktagent bewertet die Daten anhand des oben konzipierten Bayes'schen Netzes. Die für den Einsatz dieser Netze benötigten Trainingsdaten wurden im Rahmen einer Voruntersuchung gesammelt und aufbereitet. In dieser Untersuchung mussten zunächst die Testpersonen (8 Personen, davon 1 weiblich, Durchschnittsalter 26,0 Jahre) in 25 Usecases zeigen, welche Kriterien für sie bei den jeweiligen Szenarien von Bedeutung waren. Für die Kriterien, welche das Tanken betreffen, bleibt festzuhalten, dass allgemein eine Tankstelle mit dem günstigsten Preis präferiert wird und die Nutzer so spät wie möglich tanken möchten, also die Tankreichweite möglichst voll nutzen möchten. Betrachtet man die Marke, so sollte vor dem Einsatz eines Tankagenten eine mögliche „Lieblingsmarke“ erfragt werden. Dies ist besonders dann anzuraten, wenn der Fahrer Besitzer von Kundenkarten o. ä. ist.

In einer weiteren Befragung war es herauszufinden, unter welchen gegebenen Kontextgrößen die Testperson ein bestimmtes Medium zur Kontaktaufnahme im Fahrzeug bevorzugen. Kurz

zusammengefasst, wurde am Häufigsten das Mobiltelefon gewählt, während die E-Mail für den Fahrer aktuell kaum Bedeutung zu haben schien. Abhängig vom Kontext (z. B. werktags, am Wochenende, morgens, usw.) zeigten sich allerdings Unterschiede bei den gewählten Kontaktmedien.

Abschließend wurde untersucht, welchen Nutzen der Fahrer durch den Einsatz der Informationsagenten erlangt. Dazu wurde die Effizienz des Tank- und des Kontaktagenten objektiv bewertet, indem die Nutzeraktionen aus der Vorbewertung mit dem Verhalten der simulierten Informationsagenten verglichen wurden. Der Tankagent konnte nach einer Lernphase im Schnitt in 55 % der Fälle die Intention des Nutzers erkennen und berechnete die richtige Tankstelle als die Wahrscheinlichste. Wird dem Nutzer eine Vorauswahl an zehn Tankstellen in Listenform dargeboten, so befanden sich 81 % der von den Nutzern gewählten Tankstellen nach der Berechnung und Sortierung durch den Agenten auf Platz eins bis vier (also auf der ersten Seite der Liste); ohne Vorbereitung waren dies nur 44 %. Betrachtet man die Anzahl der dazu benötigten Auswahlsschritte (2,725), bedeutet dies eine Ersparnis von 48,1 % (im Median 53,7 %) im Vergleich zu einer Darstellung ohne Vorsortierung eines Agenten (5,25 Schritte). Beim Vergleich zwischen Nutzer und Agent, welches Medium in gegebenem Kontext gewählt wird, ist beim Kontaktagent eine Erfolgsquote von 57 % zu verzeichnen. In ungefähr weiteren 15 % der Fälle rangierte das vom Nutzer gewählte Medium nach Berechnungen des Agenten auf Platz zwei.

Als Ergebnis kann also festgehalten werden, dass Informationsagenten im Fahrzeug geeignete Methoden darstellen, um den Fahrer zu unterstützen. Besonders die Erkennung der Fahrertentionen durch intelligente Verfahren stellen dem Fahrer eine Art elektronische Sekretärin zur Seite. Agenten bieten durch ihre modulare Aufgabenteilung die Möglichkeit, die Art und den Umfang der Unterstützung an die Wünsche des Nutzers und die Erfordernisse der aktuellen Situation anzupassen. In diesem Kapitel wurde das Grundgerüst von Informationsagenten geschaffen und die Umsetzbarkeit überprüft. Um die potentiellen Vorteile einschätzen zu können, die sich aus dem Einsatz von intelligenten Agenten und der Integration Bayes'scher Ansätze ergeben, wurde eine objektive Bewertung vorgenommen.

Allerdings muss vor dem realen Einsatz in Automobilen die Akzeptanz seitens der Nutzer weiter überprüft werden. In Form von Langzeitstudien kann einerseits die Effizienz der Agenten verifiziert werden. Andererseits können so auch subjektive Eindrücke über Agenten erfragt werden; etwa, ob Agenten allgemein als sinnvoll erachtet werden und in welchen Bereichen diese eingesetzt werden sollen, dürfen oder auch nicht. Weitere Fragestellungen betreffen die Transparenz von Informationsagenten. Versteht der Nutzer, warum ein Agent gerade etwas macht, hat er es gar nicht bemerkt oder wird es als störend empfunden? Schlimmstenfalls führen die Agenten dann zur Verwirrung des Nutzers, bestenfalls werden sie als hilfreich eingeschätzt.

Neben der Durchführung weiterer Nutzerstudien empfiehlt sich auch die Erweiterung der Informationsagenten um kooperative Fähigkeiten der Systeme untereinander und dem Einbezug von Situationsbewusstsein. Besonderes Augenmerk kann dabei auf Konfliktmanagement und Timing gelegt werden. Das heißt, die Agenten sollten um Aspekte erweitert werden, die überprüfen, ob sicherheitskritische Situationen vorliegen. Als Beispiel hierfür kann herangezogen werden, ob ein Agent während eines Abbiegevorganges einen Telefonanruf durchstellen, nur visuell bzw. akustisch darauf hinweisen, oder den Anrufer direkt abweisen sollte. Ein multimodales Benutzerinterface für die Informationsagenten wird in Abschnitt 8.2.3 auf Basis des großflächigen Head-Up Displays vorgestellt.

## Kapitel 7

# Informationsvisualisierung und multimodale Bedienstrategien

Die Informationsvisualisierung (engl. *Information Visualization*) ist ein noch recht junges Forschungsgebiet, das sich mit computergestützten Methoden zur graphischen Repräsentation großer Datenmengen beschäftigt. Im Informations- und Kommunikationszeitalter bedarf es graphischer Methoden und Metaphern, um die wachsenden Datenmengen in einer übersichtlichen und verständlichen Form zugänglich zu machen. Gerade auch im anspruchsvollen Arbeitsumfeld des Fahrzeugs, wie auch bei mobilen Endgeräten (z. B. Apple iPhone), gewinnen intuitive Visualisierungsstrategien zunehmend an Bedeutung und zeigen große Nutzerakzeptanz durch ein erhöhtes Maß an Bedienfreude (siehe Abschnitt 2.4.6) [Bro07]. Durch eine geeignete Darstellung können umfangreiche Informationen und komplexe Zusammenhänge über den visuellen Kanal schnell und effizient wahrgenommen werden. Gerade auch bei der Fahrzeugführung könnte dies einen essentiellen Mehrwert liefern [BAK06]. Um der Fragestellung nach der Art und Weise (Wie) der Informationspräsentation nachzugehen, werden in diesem Kapitel zunächst innovative Darstellungsmethoden beschrieben. Danach werden diese Darstellungsmethoden auf die Fahrerinformationssysteme übertragen und geeignete multimodale Anzeige- und Bedienkonzepte entworfen, die eine effiziente Darstellung umfangreicher Daten- und Menüstrukturen für den Fahrer ermöglichen.

### 7.1 Darstellungsmethoden

Aktuelle graphische Oberflächen basieren überwiegend auf dem *WIMP*-Konzept (*Windows, Icons, Menus, Pointing Device*). Alle bekannten Betriebssysteme wie Microsoft Windows, Mac OS oder Linux-Fenstermanager wie KDE verwenden die sogenannte Desktopmetapher. Dabei bestehen die Fenster aus verschiedenen Widgets, welche die Interaktion meist über ein Zeige-Instrument wie Computer-Maus ermöglichen. Diese Konzepte beruhen weitgehend auf 2D-Darstellungen. Neue innovative Methoden der Informationsvisualisierung beziehen auch die *dritte Dimension* (3D) mit ein. Kognitionswissenschaftliche Untersuchungen belegen, dass eine räumlich-orientierte Wahrnehmung von Informationen für Menschen wesentlich verständlicher ist [Ras05]. Zunächst werden die Begriffe Metaphern und direkte Manipulation erläutert. Danach werden Methoden und Effekte wie Zoomable User Interfaces und Fisheye-Effekte sowie Erkenntnisse zu graphischen Animationen vorgestellt.

### 7.1.1 Metaphern

Unter einer *Metapher* versteht man in diesem Zusammenhang ein Bild oder eine rhetorische Figur, welche mit der gemeinten Sache oder dem Objekt im übertragenen Sinne in Beziehung steht. Bekannte Metaphern im PC-Umfeld sind die bereits erwähnten Windows- und Desktopmetaphern. Des Weiteren werden Icons zur Repräsentation von Objekten verwendet. Insgesamt wird versucht, geläufige Gegenstände oder Interaktionsprinzipien aus dem Alltag in Form von graphischen Repräsentanten auf den Bildschirm zu übertragen. Die graphischen Metaphern können so die Erkennung eines Objektes oder die Bedienung eines Interfaces durch bekannte mentale Modelle (siehe Kapitel 2.3) plausibilisieren und beschleunigen. Dabei kann eine Metapher auf visueller, funktionaler oder struktureller Analogie beruhen. Beispielsweise können geläufige geometrische Körper, wie Quader oder Zylinder, die Struktur oder Funktionalität veranschaulichen.

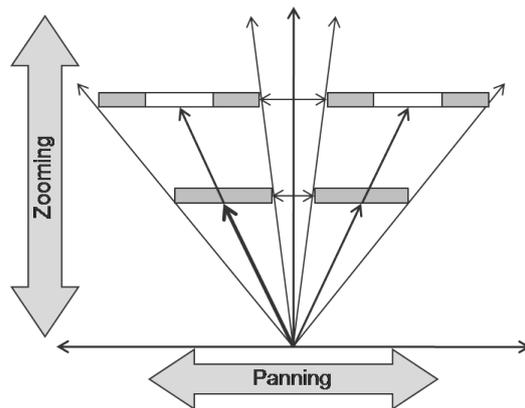
### 7.1.2 Direkte Manipulation

Shneiderman verdeutlicht das Prinzip der direkten Manipulation anhand des Autofahrens [Shn98]; dreht der Fahrer das Lenkrad nach links, so wirkt sich dies unmittelbar auf die Fahrtrichtung aus und das Fahrzeug fährt nach links. Eine geläufige Methode ist hier *Drag-and-Drop*. Als Beispiel sei hier das Löschen eines Dokumentes genannt. Durch Anwahl des Dokument-Icons und ziehen (z. B. mit Maus) über das Papierkorb-Icon verschwindet das Dokument vom Desktop und landet schließlich bei den gelöschten Objekten. Die Vorteile der Verwendung von direkter Manipulation liegen auf der Hand: die Vorgänge bzw. Handlungen sind leicht erlernbar und gut einprägsam. Des Weiteren muss der Anwender keine komplizierte Kommandosprache erlernen, um das System zu bedienen. Ein Nachteil ist, dass routinierte Handlungen stets aufs neue ausgeführt werden müssen, falls keine Möglichkeiten zu Shortcuts angeboten werden. Die charakteristischen Merkmale der direkten Manipulation sind laut [Shn98]:

- permanente, graphische Darstellung des zu manipulierenden Objekts,
- direkte, physische Einflussnahme auf die Gestaltung (über Maus, Joystick, usw.) anstelle ausformulierter Befehle in einer vorgegebenen Manipulationssprache,
- unmittelbar sichtbare, bei Bedarf auch umkehrbare Aktionen am Objekt,
- einfaches Erlernen der Funktionalität durch Benutzung.

### 7.1.3 Zoomable User Interface

Ein *Zoomable User Interface* (ZUI) ist eine Benutzeroberfläche, die es dem Nutzer ermöglicht, die gesamte Oberfläche oder einzelne Objekte zu skalieren bzw. zu *zoomen*. Durch heranzoomen kann der Bedienende bei Bedarf Bereiche am Bildschirm vergrößern bzw. durch herauszoomen einen umfassenderen Überblick gewinnen. Das im Jahre 1963 am Massachusetts Institute of Technology (MIT) entwickelte Sketchpad besaß die erste Zoom-Funktionalität. Neben dem rein *optischen Zoom* wird auch von *semantischen Zoom* gesprochen, wenn neben dem optischen Detailgrad auch der Informationsgehalt erhöht wird (siehe Abbildung 7.1(a)). Dies schafft einen guten Überblick und je nach Skalierungsgrad werden bei näherer Betrachtung zusätzliche Informationsdetails sichtbar. Kann die Arbeitsfläche über das Zooming hinaus auch horizontal sowie vertikal verschoben werden, spricht man von *Panning*.



(a) Überblick semantischer ZUI



(b) Photomesa Image Browser aus [Hum08]

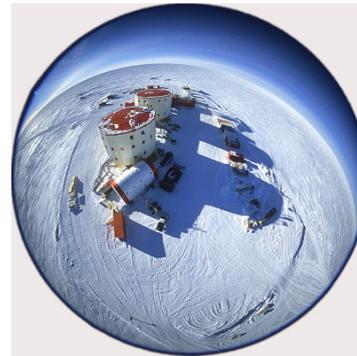
Abbildung 7.1: Zoomable User Interface

Mit der graphischen Repräsentation von Menüstrukturen über ZUIs beschäftigt sich u. a. das Human Computer Interaction Lab (HCIL) an der University of Maryland [Hum08]. Durch ZUIs wird die einfache Darstellung von hierarchischen Menüs erreicht. Hier unterstützt das geografische und räumliche Gedächtnis den Nutzer bei der Orientierung und Navigation im Menü. Ken Perlin und David Fox entwarfen 1993 an der New York University mit Pad das erste ZUI für Menüinterfaces. Mit „Piccolo“ steht heute eine Bibliothek für ZUIs in C++ und Java zur Verfügung [BGM04]. Das Interface besteht aus einer Ebene mit fest positionierten Elementen, auf die aus einer Vogelperspektive geblickt wird. Durch Zooming kann man sich nun den einzelnen Menüelementen nähern. Je weiter man in das Menü eintaucht, desto mehr Untermenüs werden angezeigt. Zudem können durch semantischen Zoom wichtige Menüeinträge schon im herausgezoomten Zustand angezeigt werden, weniger wichtige oder selten benötigte Daten und Funktionen werden erst sichtbar, wenn man weiter in das Menü hineinzoomt. Durch den semantischen Zoom können Daten bzw. Funktionen nach Bedarf gefiltert werden. Ein gut gestaltetes ZUI unterstützt dabei den Nutzer sehr gut bei der Menübedienung. Bei der Menünavigation können Hierarchiestufen übersprungen werden und sichtbare Objekte direkt angewählt werden.

[Kön06] stellt einige Evaluierungen, die zu ZUIs und ähnlichen Menüformen durchgeführt wurden, zusammen. Positive Aspekte von ZUIs sind die große Zufriedenheit der Nutzer und die hohe Bedieneffizienz. Ein weiterer Pluspunkt liegt laut [Kön06] in der guten Orientierung des Anwenders, da die räumliche Erinnerung wesentlich besser unterstützt wird und bestimmte Funktionen oder Daten mit festen örtlichen Positionen im Interface assoziiert werden können. Der Benutzer kann sich auf diese Weise leichter eine gedankliche „Karte“ des Interfaces bilden. In [GSB04] wird beschrieben, wie durch die Verwendung eines ZUI anstelle eines herkömmlichen, ordnerbasierten Systems eine Zeitersparnis von 30 % bei der Bearbeitung einer Aufgabe erzielt werden konnte. Am Human-Computer-Interaction Lab [Hum08] der University of Maryland wurden weitere ZUI Anwendungen entworfen, wie z. B. Photomesa, ein ZUI-Photobrowser (siehe Abbildung 7.1(b)).



(a) Panorama-Aufnahme



(b) Gebäude in der Mitte

Abbildung 7.2: Aufnahmen mit einem Fisheye-Objektiv

### Fisheye-Effekte

Der Fisheye-Effekt verfolgt einen ähnlichen Ansatz. Ein Fisheye-Objektiv zeichnet sich durch eine sehr kurze Brennweite und eine extreme Tonnenverzerrung aus. Daraus ergibt sich ein sehr großer Blickwinkel von etwa  $180^\circ$ . Die Perspektive, die daraus entsteht, ähnelt der eines Fisches auf die Wasseroberfläche. Mit einem Fisheye-Objektiv kann ein Fotograf umfassende bis zu  $360^\circ$ -Panorama-Aufnahmen machen, die allerdings verzerrt erscheinen (siehe 7.2(a)). Gegenstände, die sich dagegen in der Mitte der Linse befinden, werden relativ groß und kaum verzerrt dargestellt (siehe Abbildung 7.2(b)).

Diese Darstellung kann man auch bei der Darstellung von Menüs oder Daten in Benutzeroberflächen verwenden. In [Bed00] werden Fisheye-Effekte auf lineare Listen angewendet (siehe Abbildung 7.3(a)). Über die Mausposition kann der Fokus verändert werden. Auf diese Weise erhält man einen schnellen Überblick über die gesamte Liste. [Qui08] stellt einen Photobrowser vor, der sich dabei den Fisheye-Effekt zu Nutze macht (siehe Abbildung 7.3(b)).

#### 7.1.4 Animationen

Es gibt verschiedene Arten von Animation. Manche sind optisch ansprechend, aber liefern sonst keinen direkten Mehrwert. Andere verdeutlichen und unterstützen die gewählten mentalen Modelle. Die Animationen unterstützen den Nutzer, das mentale Modell des Interfaces aufzubauen. [BB99] belegt, dass angemessen konzipierte Animationen dem Benutzer helfen, sich in einer räumlichen, graphischen Oberfläche zurechtzufinden. Im Idealfall helfen die Animationen, Menüs schneller wiederzufinden und die Orientierung zu behalten. Einbußen in der Geschwindigkeit beim Erledigen von Aufgaben wurden bei [BB99] nicht gefunden. Bei der Verwendung von Animationen im Fahrzeug ist es wichtig, dass diese nicht unnötig die Aufmerksamkeit auf das Interface ziehen. In [Bro07] zeigten sich bei der Verwendung von Animationen keine negativen Einflüsse auf die Fahrzeugführung. Die Auslegung der Animations-Geschwindigkeit muss allerdings angemessen gewählt werden.

## 7.2 Zoomable User Interfaces für das Fahrzeug

Auf Basis der zuvor erläuterten Visualisierungsstrategien wurden für das in Kapitel 5 vorgestellte Datenbrowsing erweiterte Anzeige- und Bedienkonzepte für das Fahrzeug entworfen.

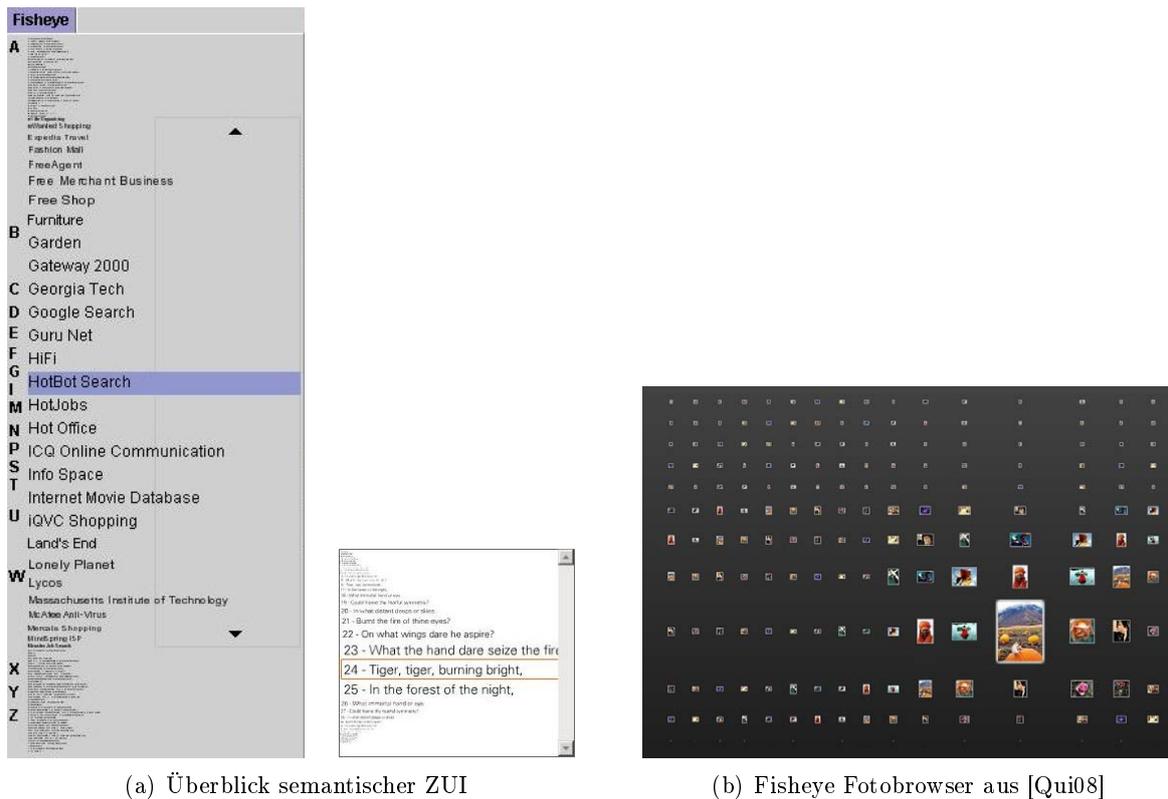


Abbildung 7.3: Fisheye Techniken

Im sogenannten PreTouch-System kommen hierbei ZUI- und Fisheye-Effekte zur Anwendung. Gerade bei den begrenzten Displaygrößen des Fahrzeugs könnten diese Effekte ihre Vorteile zeigen. Des Weiteren erhofft man sich, dass neben einer gesteigerten Bedienfreude, die Interfaces den Fahrer bei der Orientierung und intuitiven Bedienung unterstützen. Nachfolgend werden zwei prototypische Entwicklungen vorgestellt, die verschiedene multimodale Bedienstrategien bereitstellen.

### 7.2.1 Zoomable User Interfaces mittels Annäherungstouchscreen

In diesem Ansatz wird zunächst auf eine Touchinteraktion zurückgegriffen. Als Besonderheit verwendet das PreTouch-System einen Touchscreen mit Annäherungssensorik. Der Fingerabstand zum Bildschirm steuert dabei die Zoomstufe des Bildschirms und ermöglicht weitere Menüfunktionen. Durch Koppelung mit Sprache werden multimodale Bedienstrategien realisiert.

#### Technische Grundlagen

Als Hardware wird ein Rechner mit Flachbildschirm verwendet. Auf dem Bildschirm ist eine Touchsensorscheibe montiert. Zur Steuerung der Zoomstufen wird ein Eingabegerät benötigt, das zur kontinuierlichen Positionierung in der  $x$ - $y$ -Ebene des Bildschirms auch eine Tiefeninformation in  $z$ -Richtung zur Verfügung stellt. In diesem Prototypen wird ein kapazitiver Touchscreen verwendet, der mit Projected Capacitive Messtechnik arbeitet. Damit ist bis zu

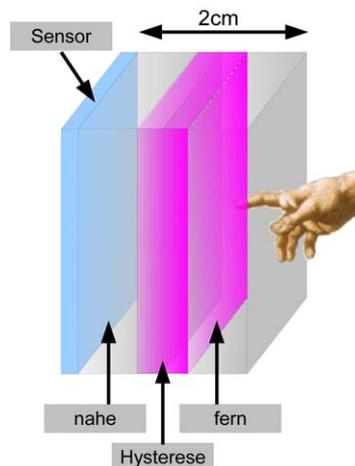


Abbildung 7.4: Aufteilung des Sensorraums

einem gewissen Grad ein markerloses Fingertracking möglich. Die Unterscheidung zwischen Übergangsstufen bei der Annäherung erfolgt mittels Hysterese (siehe Abbildung 7.4). Dadurch wird verhindert, dass an den Grenzbereichen ein „zittern“ entsteht. Das graphische Interface wird mit Adobe Flash realisiert und über MDM Zinc eingebettet, um die Touchpositionen auswerten zu können. Außerdem ist der Spracherkennungsscanner Scansoft Vocon 3200 über TCP/IP-Sockets integriert. Zur multimodalen Fusion wird eine Kombination aus temporal und Unification-Based Algorithmus verwendet. Das Datenmanagement-System aus Abschnitt 3.5.3 bietet die Nutzung von Web Services zur Datenbereitstellung und Aufbereitung.

## Bedienung

Im Grundzustand zeigt der Bildschirm in der obersten Ebene das Statusmenü mit den Domänen Navigation, Kontakte, Kalender, Events, Musik, Einstellungen. Findet keine Bedienung statt, werden lediglich die Domänen-Überschriften dargestellt. Bewegt der Nutzer jedoch seinen Finger in Richtung eines Domänenobjektes, wird das Feld vergrößert und durch semantischen Zoom werden weitere Informationen sichtbar (siehe Abbildung 7.5). Dabei werden mehrere Zoom-Stufen (semantischer Zoom) unterschieden: Zunächst vergrößert sich bei Annäherung (ab ca. 5 cm) die Fisheye-Liste (Stufe 1). Durch vertikales Bewegen in der y-Richtung kann man sich in der Liste bewegen. Sind weitere Einträge vorhanden und man erreicht den Fensterrand, dann beginnt die Liste zu scrollen. Überschreitet man einen definierten Bereich kurz vor der Touchscheibe (Stufe 2), werden zum selektierten Eintrag weitere Funktionen in Form von Icons sichtbar, die aktiviert werden können. In der Kontaktliste sind dies beispielsweise anrufen, editieren, usw. Die Bedienung von hierarchischen Menüs, die normalerweise mehrere Bedienschritte erfordern, kann über die Annäherung durch einen Bedienvorgang ausgeführt werden. Die Stufe 3 erreicht man über ein Icon (Pfeil-Symbol) in der rechten unteren Ecke, welche das Fenster bildschirmfüllend vergrößert. Hier steht dann eine größere Anzahl an Funktionen zur Verfügung. Ist eine Domäne maximiert, befinden sich alle anderen Domänen minimiert in der Taskleiste. Um wieder in den Ausgangszustand, das Statusmenü, zu springen befindet sich in der Taskleiste ein Homebutton. Außerdem können die Fenster über den Schließen-Button (X) am rechten oberen Rand wieder in die Ausgangsgröße zurückgeführt

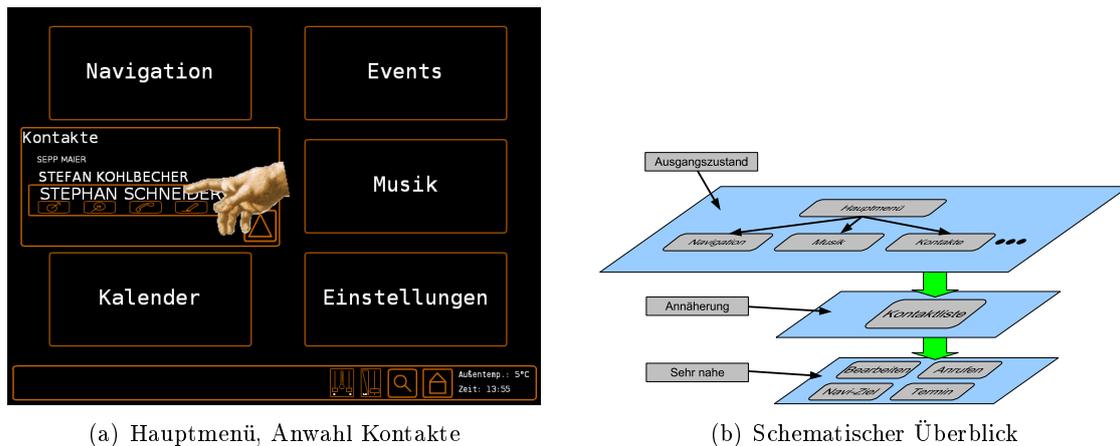


Abbildung 7.5: PreTouch-System

werden. Platzsparende Zoomable Buttons können sich durch Annäherung in FishEye-Listen oder Slider-Funktionen (z. B. Veränderung der Lautstärke) morphen. So kann schnell, ohne den Kontextwechsel zu verlassen, eine Auswahl getroffen werden. Durch Splashscreens, die in der Mitte des Bildschirms für kurze Zeit erscheinen, kann Feedback zu Interaktionen zurückgegeben werden.

Das Design wurde sehr schlicht gehalten, um den Fahrer nicht unnötig zu belasten. Lediglich die Farben schwarz für den Hintergrund, weiß für den Text und orange für die Menüränder, Schaltflächen und Piktogramme wurden verwendet. So wird die schnelle Orientierung und gute Lesbarkeit gefördert.

### Alphanumerische Eingaben

Zur alphanumerischen Eingabe wurden zwei Strategien entwickelt, die nun vorgestellt werden.

**Fisheye-Tastatur:** In diesem Ansatz wird eine graphische Tastatur auf dem Touchscreen dargestellt. Die Besonderheit dieser Methode ist die Verwendung eines Fisheye-Effektes (siehe Abbildung 7.6(a)). Wenn der Benutzer sich mit dem Finger dem Bildschirm nähert, erscheint ein Fisheye-Effekt auf der Tastatur, bei dem die Tasten, die unter dem Finger liegen, vergrößert dargestellt werden. Angrenzende Buchstaben wandern etwas zur Seite, um Überlappungen zu vermeiden. Diese Technik soll das Treffen der Tasten erleichtern. Insgesamt ist es so möglich eine Tastatur auf relativ kleinem Bildschirm darzustellen.

**MatrixKey-Touch:** Eine weitere Möglichkeit wurde ursprünglich für die Texteingabe mittels Nummernblock entwickelt (vgl. [IDZ04]). Die Buchstaben sind auf 9 Cluster aufgeteilt (siehe Abbildung 7.6(b)). Nach Anwahl eines Clusters können dann die dahinterliegenden Buchstaben ausgewählt werden. Nähert man sich mit dem Finger einem Cluster, detektiert dies die Annäherungssensorik des Touchscreens und das Cluster wird vergrößert. Nun kann bequem ein Buchstabe über das Berühren des Touchscreens gewählt werden.

### Multimodale Bedienstrategien

Ziel ist hier die Identifikation geeigneter multimodaler Bedienstrategien, die sich für diese Art der Darstellung besonders eignen. Die Bedienung des ZUIs erfolgt multimodal durch Kombina-

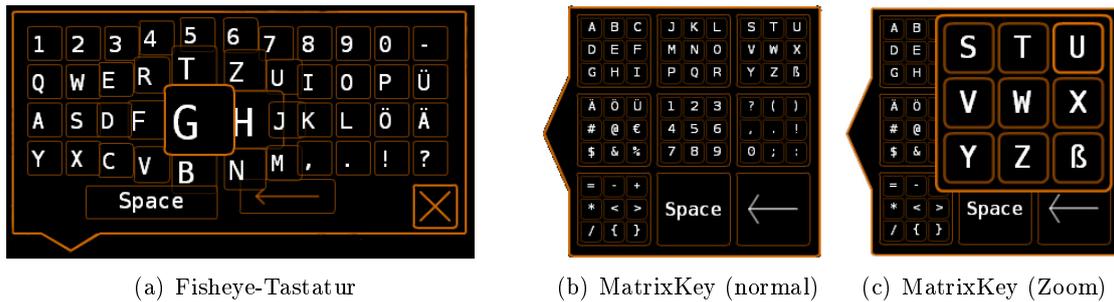


Abbildung 7.6: MatrixKey

tion der Modalitäten Annäherungstouchscreen und Spracheingabe. Dabei können die beiden Modalitäten zunächst vollkommen parallel und unabhängig voneinander eingesetzt werden. Als Beispiel für eine Schiebermanipulation siehe Abbildung 7.7.

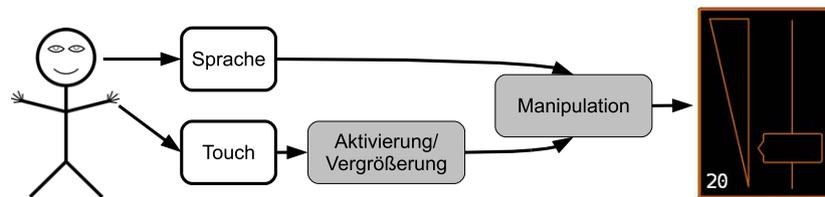
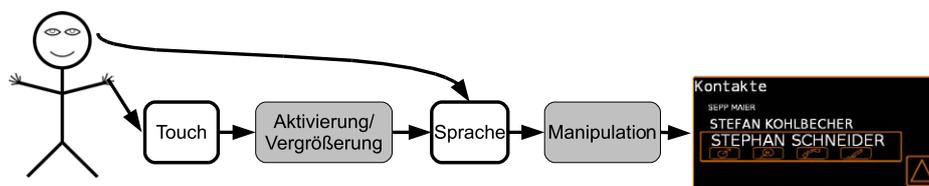


Abbildung 7.7: Flussdiagramm für multimodale Schiebermanipulation

Ein weitere Möglichkeit ergibt sich durch eine parallele multimodale Bedienung. Die Kombination von Zeigen auf ein Objekt und gleichzeitigem Sprechen einer Aktion zeigt sich hier als besonders geeignet (siehe Abbildung 7.8). Diese Interaktionsmethode wird als *Point-and-Talk* bezeichnet. Die Fusion der beiden Kanäle Sprache und Zeigegeste erfolgt auf Basis eines Unification-Based Algorithmus.

Abbildung 7.8: Flussdiagramm für einen *Point-and-Talk*-Vorgang

### 7.2.2 Zoomable User Interface mittels Dreh-/Drücksteller

Das oben vorgestellte ZUI zum Datenbrowsing lässt sich zum Teil auch auf die Bedienung mit taktilem Eingabegerät übertragen. Bei dem sogenannten CarZoom sind eine größere Anzahl an Freiheitsgraden über ein taktiles Eingabegerät erforderlich. Die Funktionalität wurde auf einen Dreh-/Drücksteller mit Schiebefunktion und weitere Softkeys übertragen. Durch Drehen

im Hauptmenü lassen sich einzelne Objekte anwählen und dazu erfolgt die Darstellung von weiteren Details (Stufe 1 bzw. 2) (siehe Abbildung 7.9).

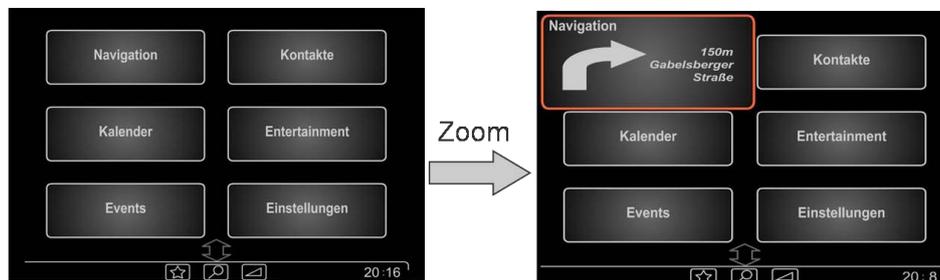


Abbildung 7.9: Zoomable User Interface mittels Dreh-Drücksteller

Durch Drücken des Dreh-/Drückstellers gelangt man in die jeweilige Domäne, in der anschließend weitere Informationen und Funktionen zur Verfügung stehen (Stufe 3). Hierbei erfolgt auch die Darstellung der direkt angrenzenden Domänen (siehe Abbildung 7.10). Analog zu der Anordnung im Statusmenü kann der Nutzer zwischen den angrenzenden Domänen durch Schieben wechseln (Menüpanning). Durch das mentale Modell des Matrixmusters kann sich der Nutzer gut orientieren und erlernt schnell die Anordnung der einzelnen Domänen. Durch diese kartenähnliche 2D-Navigation kann die Orientierung des Nutzers unterstützt werden. Eine spezielle Anforderung stellt die Bedienung von Listen über einen Dreh-/Drücksteller dar. Für eine analoge Repräsentation erfolgt die Listendarstellung in gekrümmter Form (siehe Abbildung 7.11). Der Blick des Fahrers kann so zielgerichtet in die Mitte des Bildschirms erfolgen. Durch die Fisheye-Darstellung werden unmittelbar angrenzende Einträge der aktuellen Listenposition besonders gut hervorgehoben.

### Alphanumerische Eingabe

**MatrixKey-Texteingabe:** Die MatrixKey-Texteingabe von CarZoom ist in Abbildung 7.12 zu sehen (siehe auch [IDZ04]). Die Tasten auf dem Nummernblock entsprechen jeweils einem der neun Felder in Abbildung 7.12. Drückt man eine dieser Tasten, vergrößert sich das entsprechende Feld. Wählt man z. B. die „1“, vergrößert sich das linke obere Feld (siehe Abbildung 7.12). Drückt man in diesem Zustand die Taste „1“ erneut, schreibt man ein „A“, drückt man die „2“, schreibt man ein „B“, usw. Bei entsprechender Übung könnte ein gewisses Maß an Blindbedingung erreicht werden.

**TwistKey-Texteingabe:** Bei CarZoom wird ein weiteres Texteingabeverfahren verwendet, welches das Drehen des Controllers mittels einer Drehmetapher auf dem Bildschirm visualisiert (siehe Abbildung 7.12(b)). Durch Drehen des Controllers rotiert die Scheibe, auf der die Buchstaben dargestellt sind. Das Auswahlsegment hebt die aktuell aktiven Buchstaben hervor. Drückt man in diesem Zustand den Controller einmal, wird der Buchstabe der innersten Scheibe gewählt, drückt man zweimal wird der Buchstabe auf der mittleren Scheibe gewählt.

## 7.3 Verwendung von 3D-Metaphern im Fahrzeug

In diesem Abschnitt werden weitere Darstellungskonzepte zum Umgang mit großen Datenmengen präsentiert, die vor allem auch auf die Verwendung von 3D-Metaphern zurückgreifen.

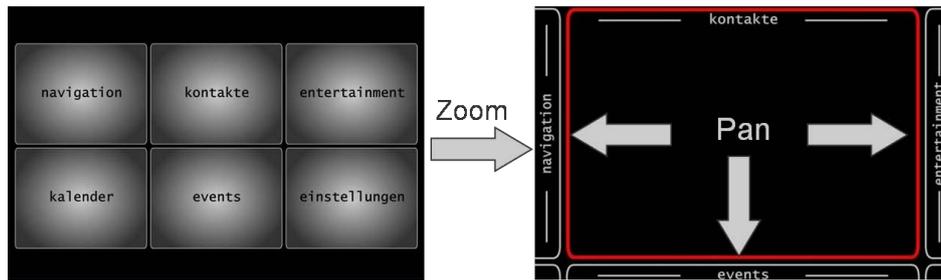


Abbildung 7.10: Zoomable User Interface mittels Dreh-Drücksteller (Menüpanning)

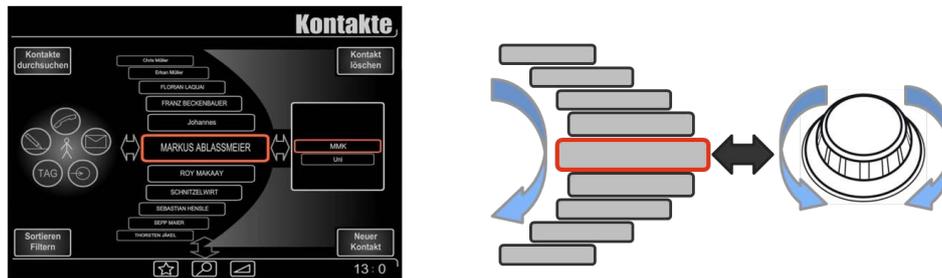


Abbildung 7.11: Gekrümmte Listendarstellung mit Zooming

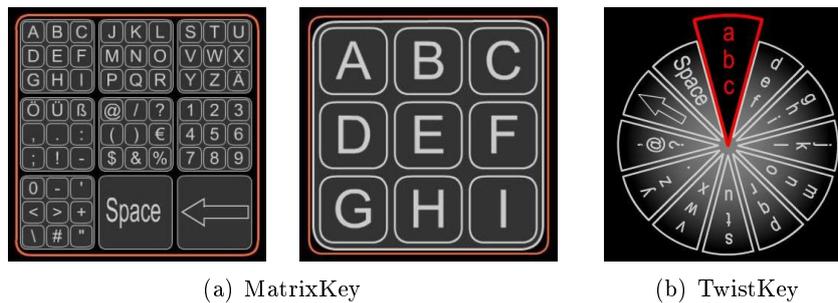


Abbildung 7.12: Taktile alphanumerische Eingabeverfahren

### 7.3.1 Skalierbares 3D-Filmrollenkonzept und Verknüpfungsnetzwerk

Dieses Konzept integriert verschiedene Visualisierungsmethoden zur erweiterten Bedienung des Browsingansatzes. Die Bedienung erfolgt multimodal durch einen Dreh-/Drücksteller mit Schiebefunktion, einen Touchscreen mit Annäherungssensorik und Sprache. Das Interface umfasst folgende Eigenschaften (siehe Abbildung 7.13):

- **Filmrollenmetapher:**  
Zwischen den einzelnen Domänen kann über eine Filmrollenmetapher navigiert werden. Durch Ziehen bzw. Streichen am linken Bildschirmrand kann die Rolle bewegt werden und es erfolgt der Wechsel zur angrenzenden Domäne.
- **Statusleiste:**  
Am unteren Rand befindet sich eine Statusleiste mit Buttons zur direkten Anwahl von Domänen. Diese können personalisiert werden.

- **Dynamische Segmentierung:**

Dieses Konzept ermöglicht eine dynamischen Bildschirmaufteilung. Die einzelnen Bildschirmbereiche können je nach Bedarf erweitert bzw. reduziert werden. Durch Anwahl eines reduzierten Bereiches erfolgt das animierte Ein- bzw. Ausblenden weiterer Einzelheiten über Pfeilbuttons. Der Bereich für die Listendarstellung wird dazu etwas reduziert. Die Grobaufteilung bleibt dabei bestehen und so behält der Nutzer hier den Überblick und die Orientierung. Zwischen diesen Bereichen kann durch Schieben über den Dreh-/Drücksteller gewechselt werden.

- **Schichtenkonzept:**

Popups öffnen sich vor der Filmrolle durch animierte Vergrößerung des Fensters vom Ursprungsbutton aus.

- **Netzwerkgraph:**

Eine Besonderheit stellt der Netzwerkgraph dar. Er zeigt Zusammenhänge zwischen verschiedensten Daten auf und ermöglicht das Browsen in den vernetzten Daten über die Domänen hinweg. Die Verknüpfung basiert auf der Verwendung von Tags. Es erfolgt eine Unterscheidung zwischen Datenknoten und Tag-Knoten. Nach Auswahl eines Elementes erfolgt die animierte Zentrierung des Graphes.

- **Gravitations-Listen:**

Die Fisheye-Listen können auch durch „Anschieben“ bzw. Ziehen des Fingers in Bewegung gesetzt werden, die in Abhängigkeit des Momentes dann wieder zum Stillstand kommen.

- **Overlay Elemente:**

Sie zeigen situativ wichtige Infos an. Es erfolgt die transparente Überblendung des Bildschirms. Die Inhalte der aktuellen Domäne bleiben jedoch weiterhin sichtbar.

### 7.3.2 3D-Zylindermetapher als Menütiefe

Dieses Visualisierungskonzept ermöglicht die graphische Darstellung von Menüstrukturen. Das sogenannte Donut-Konzept besteht aus 3D-Zylinderscheiben (siehe Abbildung 7.14). Die einzelnen Hierarchieebenen werden durch kreisförmige Menüscheiben repräsentiert. Die Auswahl auf der Kreisscheibe und die Menünavigation erfolgt über das sogenannte Donut-Bedienelement. Durch Drücken kann der fokussierte Eintrag ausgewählt werden und der Sprung in die nächste Hierarchiestufe wird ausgeführt. Über ein Skip-Scroll-Rad ist ein Sprung zwischen den Ebenen möglich. Ist die Anzeige einer großen Datenmenge erforderlich, wechselt die GUI durch Animation in eine Listendarstellung.

### 7.3.3 3D-Rollkartei und Karteiflip

Zur Repräsentation großer Datenmengen, wie Adressdaten, wurde hierbei unter Verwendung von 3D Metaphern ein Interaktionskonzept gestaltet. Als Hauptmetapher wurde eine Rollkartei, bestehend aus drehbaren Visitenkarten in Form von Quadern und Zylinderscheiben verwendet. Dieser Alltagsgegenstand wurde analog in die graphische Darstellung übertragen. Zur Realisierung wurde Java 3D verwendet. Diese Programmiersprache ermöglicht die graphische Darstellung von 3D Objekten.

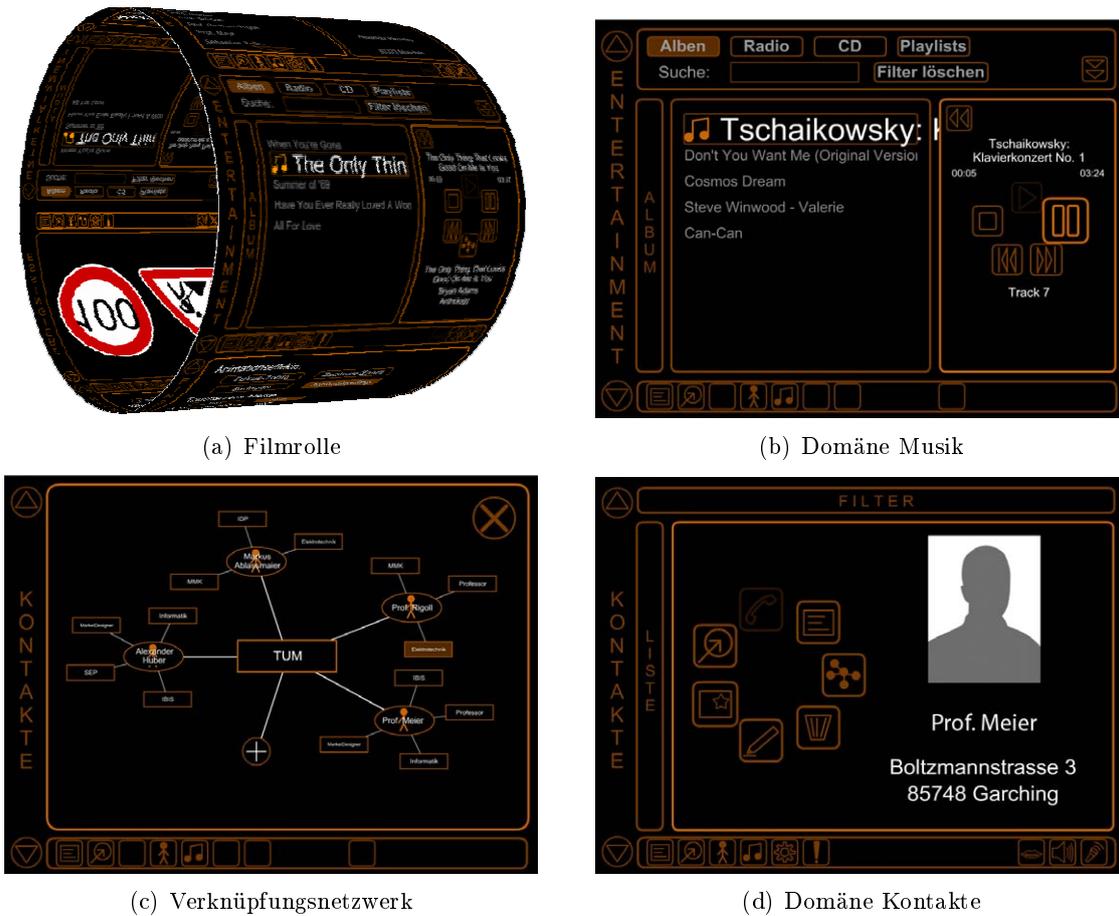


Abbildung 7.13: Überblick skalierbares 3D-Filmrollenkonzept

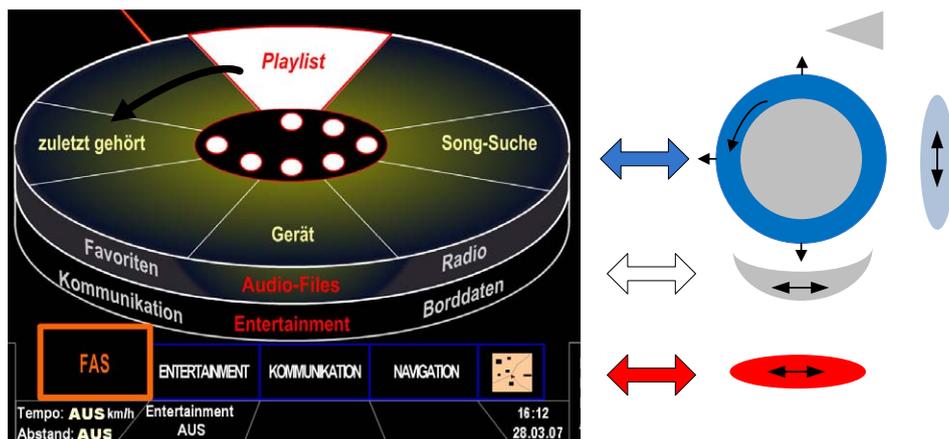


Abbildung 7.14: 3D-Zylindermetapher als Menütiefe

**Interaktionskonzept:** Zur Verwaltung von Adressdaten werden die Visitenkarten kreisförmig um eine drehbare Zylinderachse angeordnet (siehe Abbildung 7.15). Die Darstellung eignet sich besonders gut für die Bedienung über den Dreh-/Drücksteller. Der Nutzer kann damit die

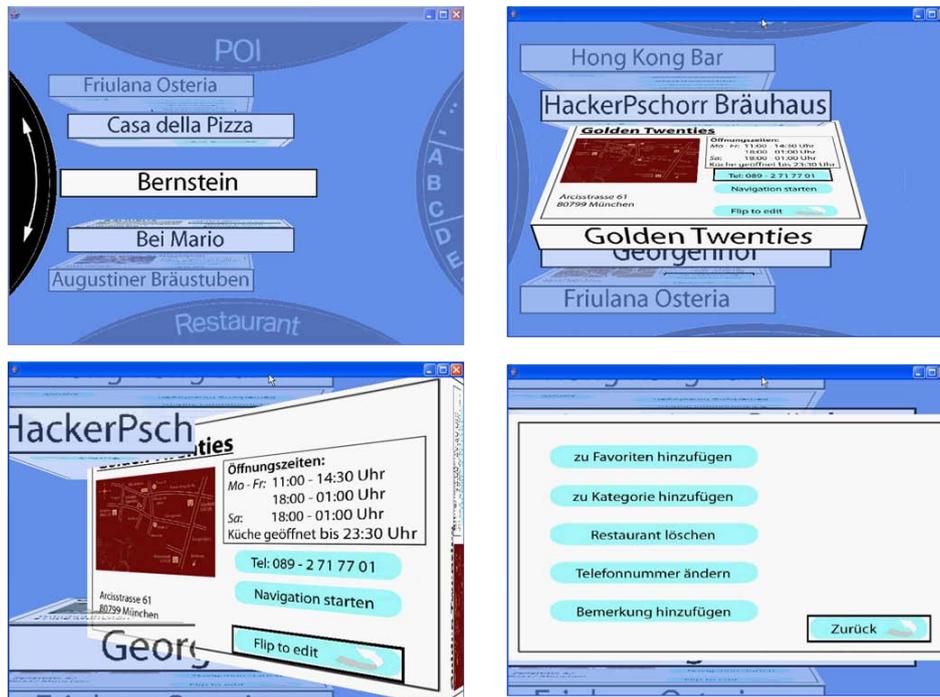


Abbildung 7.15: 3D-Rollkartei und Karteiflip

Zylinderachse der Rollkartei manipulieren bzw. in Rotation versetzen. In dem GUI wird die Rollkartei seitlich dargestellt. Um in den Adressdaten zu blättern, kann der Controller nach rechts und links gedreht werden. Der Fokus bzw. das selektierte Element liegt dabei stets in der Bildschirmmitte.

Durch Interaktion mit der Adresskartei können dem Nutzer weitere Details und Funktionen zugänglich gemacht werden. Die Manipulation wird mit Hilfe einer Animation veranschaulicht. Durch Drücken des Controllers dreht sich die Kartei im Uhrzeigersinn und anschließend erfolgt die Anzeige einer gewohnten Visitenkarte. Durch nochmaliges Drücken gelangt der Nutzer auf die Rückseite der Kartei, wo sich weitere Funktionen befinden, die durch Drehen selektiert werden können. Dieser Effekt wird als *Flipping* bezeichnet [BAK06]). Durch Schieben des Dreh-/Drückstellers in beliebige Richtung springt die Kartei in den Ausgangszustand.

Zur Datenfilterung gibt es verschiedene Zylinderscheiben am oberen, unteren und rechten Bildschirmrand. Durch Schieben in horizontaler bzw. vertikaler Richtung kann zwischen diesen Zylinderscheiben und der Rollkartei gesprungen werden. Die Zylinderscheibe am oberen Rand ermöglicht eine Grobfilterung durch Auswahl einer bestimmten Kategorie, die untere Scheibe bietet eine Feinfilterung der Daten abhängig von der Kategorie. Die Eingabe wird nach dem Drehen sofort ausgeführt, kann aber durch entgegengesetztes Drehen jederzeit rückgängig gemacht werden. Eine Spezialfunktion ist noch das Flipping der Zylinderscheibe. Auf der Rückseite finden sich weitere selten genutzte Funktionen, die durch Drücken des Controllers gedreht werden können. Am rechten Rand des Bildschirms gibt es die Möglichkeiten der alphanumerischen Eingabe. So können die Daten schnell gefiltert und durchsucht werden. Durch Flipping der Zylinderscheibe gelangt man von der Buchstabeneingabe zur Zahleneingabe.

## 7.4 Bewertung

Nachfolgend werden die Ergebnisse der Untersuchung von PreTouch (aus Abschnitt 7.2.1) präsentiert. Ziel war die Bewertung der verwendeten Visualisierungsstrategien hinsichtlich ihrer Eignung im Fahrzeug.

### 7.4.1 Versuchsdurchführung

Die Untersuchung fand im Fahrsimulator des Lehrstuhls für Mensch-Maschine-Kommunikation statt. Als Fahraufgabe wurde der Lane Change Task (LCT) verwendet (siehe Abschnitt 2.5.1). An der Studie nahmen 20 Probanden (davon 4 weiblich, Durchschnittsalter: 36,2 Jahre) teil. Nach einer Einführung hatten die Versuchspersonen zwei Aufgabensequenzen während der Fahrt mit dem Interface PreTouch durchzuführen. Die Aufgaben wurden ebenfalls auf einem Referenzsystem ohne Visualisierungseffekte und streng hierarchischer Menüstruktur durchgeführt. Das Referenzsystem wurde über einen Dreh-/Drücksteller bedient. Die erste Aufgabe bestand in der Auswahl und dem Bearbeiten eines Kontakts. Die zweite Aufgabe erforderte das Suchen eines Musiktitels im Unterhaltungsmenü. Dabei mussten die Versuchspersonen die vorgestellten multimodalen Eingabestrategien verwenden. Nach jeder Aufgabe wurde ein Fragebogen und am Ende der Untersuchung der Attradiff-Fragebogen ausgefüllt.

### 7.4.2 Subjektive Ergebnisse

Zunächst werden die Ergebnisse zu den ZUI-Effekten präsentiert (siehe Tabelle 7.1). Die verwendeten ZUI-Effekte bewertete die überwiegende Mehrheit der Probanden als hilfreich. Die Auslegung der Effekte war insgesamt angemessen. Des Weiteren wurde die schlichte und einfache Darstellung positiv bewertet. Ein Teil der Probanden wünschte noch weitere Piktogramme bzw. graphische Metaphern.

ADJEKTIV-PAARE	ANZAHL DER NENNUNGEN				
<i>hilfreich – störend</i>	5	12	2	1	0
<i>zu schnell – zu langsam</i>	0	6	9	5	0
<i>zu selten – zu oft</i>	0	1	16	2	1

Tabelle 7.1: Bewertung der ZUI-Effekte (semantisches Differential)

Die Bedienung des Interfaces PreTouch wurde wie folgt bewertet (siehe Tabelle 7.2): Die Kopplung der Bedienung über den Touchscreen mit der Annäherungssensorik wurde als eine direkte Eingabemethodik empfunden, die viel Spaß und Freude bereitet. Allerdings bemängelten die Probanden noch die Robustheit und mangelnde Präzision bei der Touch-Bedienung und zum Teil bei der Detektion der Annäherung.

Des Weiteren waren die multimodalen Bedienstrategien zu bewerten (siehe Tabelle 7.3). Die Versuchspersonen empfanden dabei die Spracheingabe als sehr hilfreich. Der Großteil der Probanden bevorzugte die multimodale Bedienung gegenüber einer reinen Touchbedienung. Sie waren weitgehend der Meinung, dass die multimodale Eingabe weniger Aufmerksamkeit beansprucht. Außerdem wünschten sich mehrere Personen auch noch akustisches Feedback zur Verbesserung der Blindbedienbarkeit.

EIGENSCHAFT (gut..schlecht)					
<i>Direktheit</i>	1	11	7	0	0
<i>Präzision</i>	0	1	4	14	1
<i>Annäherung</i>	1	6	6	6	1
<i>Robustheit</i>	1	3	12	3	0
<i>Spaß</i>	1	14	4	1	0

Tabelle 7.2: Bewertung der Touch-Bedienung

FRAGE (stimme zu..stimme nicht zu)					
Spracheingabe hilfreich	10	9	0	0	1
Multimodale Bedienung					
- wird bevorzugt	10	2	3	3	2
- erfordert weniger Aufmerksamkeit	10	5	5	0	0

Tabelle 7.3: Bewertung der multimodalen Bedienung

Der datenorientierte Menüaufbau wurde folgendermaßen bewertet (siehe Tabelle 7.4): Er erschien den Versuchspersonen als klar strukturiert und ermöglichte eine schnelle Navigation im Menü. Die Anzahl an Informationen war angemessen. Nach Ansicht der Probanden hatten sie durch das ZUI weniger das Gefühl sich zu „verirren“. Insgesamt wurde das Eintauchen in das Menü als intuitiv, sinnvoll und leicht bedienbar bewertet.

BEWERTUNG DES MENÜAUFBAUS (Anzahl Nennungen):					
<i>klar strukturiert – unübersichtlich</i>	13	6	0	1	0
<i>zu viele Optionen – zu wenige Optionen</i>	0	0	20	0	0
<i>schnelle Navigation – langsame Navigation</i>	6	11	1	1	0
BEWERTUNG DES EINTAUCHENS IN DAS MENÜ (Anzahl Nennungen):					
<i>schwer verständlich – intuitiv</i>	0	2	1	10	7
<i>unnützlich – sinnvoll</i>	0	3	2	8	7
<i>schwer bedienbar – leicht bedienbar</i>	0	0	2	12	6

Tabelle 7.4: Bewertung des Menüs (semantisches Differential)

Abschließend konnten sich 17 der 20 Probanden den Einsatz eines solchen Systems im Fahrzeug vorstellen. In Abbildung 7.16 ist das Endergebnis des AttrakDiff-Fragebogens zu sehen. Als Ergebnis bleibt festzuhalten, dass PreTouch in allen Bewertungskriterien durchwegs positiv bewertet wurde. Im Vergleich zu klassischen FIS erreicht man eine deutlich Verbesserung im Bereich Hedonische Qualität-Stimulation (HQ-S). Bezüglich diesem Aspekt schnitt PreTouch vor allem in den Punkten kreativ, innovativ, neuartig und originell sehr positiv ab.

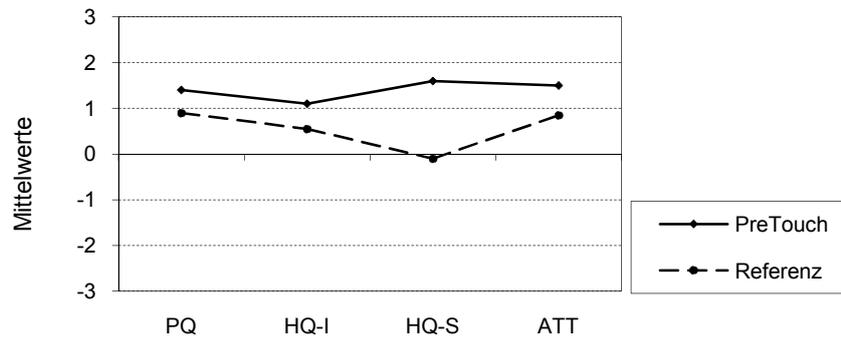


Abbildung 7.16: Mittelwerte der Attradiff-Auswertung

## 7.5 Zusammenfassung

Diese Kapitel hat Lösungsmöglichkeiten zur Anwendung von innovativen Visualisierungsstrategien aufgezeigt. Dabei kamen insbesondere ZUIs und Fisheye-Effekte zur Anwendung. Abschließend kann gefolgert werden, dass sich die graphische Gestaltung wie die Verwendung von Animationen und ZUIs sehr positiv auf die Akzeptanz und Bedienfreude niederschlägt. Gerade zur Bedienung umfangreicher Daten im Fahrzeug kann dies sehr hilfreich sein. Dabei wurde die datenorientierte Struktur des Browsing-Ansatzes als sehr positiv und übersichtlich bewertet. Die visuellen Effekte wecken hier die Neugier des Nutzers das Interface und die Daten zu explorieren. Es zeigte sich auch, dass die multimodale Bedienung über die Modalitäten Sprache und Touch ein gute Ergänzung darstellen und als weniger belastend empfunden wurde. Neben großer Akzeptanz der Sprachbedienung, wünschten die Nutzer des Weiteren Sprachfeedback bei der Bedienung. Keine Aussage konnte jedoch darüber getroffen werden, wie gut sich die Bedienung mit der Fahraufgabe vereinen lässt.

## Kapitel 8

# Großflächiges Head-Up Display

Moderne Anzeigetechnologien im Fahrzeug ermöglichen es, relevante Informationen für den Fahrer wesentlich komfortabler und intuitiver zu präsentieren. Das Head-Up Display (HUD) ist eine solche Technologie, die seit einigen Jahren in den Automobilbereich Einzug gehalten hat. Dieses Display blendet Anzeigehalte direkt in das Sichtfeld des Fahrers ein. Dadurch ergeben sich viele Vorteile – aber auch nicht zu vernachlässigende Nachteile. Unter der Prämisse eines großflächigen HUDs wird zunächst ein gesamtheitliches, prioritätenorientiertes Anzeige- und Bedienkonzept vorgestellt. Anschließend wird ein Lösungskonzept für ein sinnvolles Zusammenspiel zwischen dem HUD und einem frei programmierbaren Kombiinstrument aufgezeigt. Im letzten Abschnitt wird auf die Möglichkeiten eines kontaktanalogen HUDs eingegangen. Hierzu werden geeignete Anzeigekonzepte vorgestellt und ein prototypischer Aufbau eines quasi-kontaktanalogen HUDs beschrieben. Dazu wurde ein remote Gazetracking-System entwickelt, das für die technische Realisierung erforderlich ist. Verschiedene Fragestellungen, die ein solch großflächiges HUD aufwerfen, werden am Ende des Kapitels anhand von ausgewählten Untersuchungsergebnissen beantwortet.

### 8.1 Grundlagen zum Head-Up Display

Ein HUD ist ein visuelles Anzeigeelement, bei dem mittels Projektion Informationen in das Sichtfeld des Fahrers eingeblendet werden. Diese Technologie stammt ursprünglich aus dem militärischen Anwendungsbereich. HUDs wurden erstmals in den 70er Jahren in amerikanischen Kampfflugzeugen verbaut und später für tieffliegende Helikopter verwendet. Für die Militärpiloten spielt die Informationsbelastung eine entscheidende Rolle. Die Abwendung des Blickes zum Ablesen der Instrumente kann unter Umständen zu einer fatalen Ablenkung führen. Daher versorgt das HUD den Piloten mit Informationen über den Verlauf des Fluges, den Zustand des Flugzeuges und mögliche Zielobjekte. Hierbei hat der Pilot die Möglichkeit sich überwiegend auf die Informationen im HUD konzentrieren, da dort sämtliche flugrelevanten Informationen zur Verfügung gestellt werden. [Kau04]

Auf Basis dieser Entwicklungen wurde das HUD für eine Vielzahl weiterer Applikationen eingesetzt. Die erste Anwendung im Automobil erfolgte im Jahre 1998 durch die Firma General Motors in dem Modell „Cutlass Supreme“. Über fest positionierte Leuchtdioden (LEDs) wird das Bild auf die Windschutzscheibe gespiegelt. Das erste freiprogrammierbare und mehrfarbige HUD wurde im Jahr 2001 durch die Firma Siemens VDO Automotive vorgestellt. Daraufhin setzte der Automobilhersteller BMW dieses Display im Jahre 2003 serienmäßig in der 5er Rei-

he ein (siehe Abbildung 8.1). Bei diesem HUD bekommt der Fahrer auf einem etwa 18 cm breiten und 10 cm hohem Display folgende Anzeigen zu sehen: digitale Geschwindigkeitsanzeige, Navigationshinweise und Warnmeldungen (hier das Active-Cruise-Control (ACC)). [BMW06, Sch03a, Cor08]



Abbildung 8.1: Das HUD der BMW 5er Reihe [BMW08b]

Der englische Begriff *Head-Up* deutet daraufhin, dass der Kopf beim Ablesen nach oben gerichtet bleibt und damit nicht vom Straßengeschehen abgewendet werden muss. Analog dazu werden die Anzeigen unterhalb der Windschutzscheibe, wie CID und Kombi, auch als *Head-Down Displays* (HDD) bezeichnet. [Sch03a]

Grundsätzlich werden zwei Arten der Darstellung unterschieden: *kontaktanaloge* und *nichtkontaktanaloge* HUDs. Bei kontaktanalogen HUDs werden die angezeigten Symbole direkt in die Fahrszene integriert und verschmelzen so mit der Umwelt. Schon bei den Kampffjets wurden Zieleobjekte wie fremde Flugzeuge oder Bodenziele kontaktanalog über das HUD markiert. In Kraftfahrzeugen werden bisher noch keine kontaktanalogen HUDs eingesetzt, diese befinden sich jedoch im Forschungsstadium [Lan08, TLKB06]. Die kontaktanaloge Anzeige erfordert insgesamt einen sehr hohen technischen Aufwand. Auf das Prinzip eines quasi-kontaktanalogen HUDs im Fahrzeug, mögliche Konzepte und die Realisierung wird im Abschnitt 8.4 eingegangen.

### 8.1.1 Funktionsweise und technische Auslegung

Das Prinzip der HUDs basiert auf den Regeln der Strahlenoptik. Dabei wird eine Bildquelle  $b$  auf einer Glasfläche, dem Combiner  $c$ , projiziert und teilweise reflektiert. Gemäß dem Reflexionsgesetz gilt hier für den Einfallswinkel  $\alpha$  und dem Ausfalls- bzw. Reflexionswinkel  $\beta$  bezogen auf das Einfallslot  $l$ :

$$\alpha = \beta \quad (8.1)$$

Der reflektierte Anteil  $v$  wird für den Betrachter so als virtuelles Bild wahrgenommen. Hier ist der Abstand  $x$  zwischen dem virtuellen Bild  $v$  und dem Combiner  $c$  äquivalent zum Abstand  $y$  zwischen Bildquelle  $b$  und dem Combiner  $c$ . Das heißt, das virtuelle Bild wird im Abstand der Entfernung  $x$  der Bildquelle zur Glasfläche wahrgenommen und nicht in der Ebene der Glasfläche  $c$ . In der Abbildung 8.2 ist der prinzipielle Aufbau skizziert. Je größer der Strahlengang von der Bildquelle  $b$  zum Combiner  $c$  ist, desto weiter entfernt ist das wahrgenommene Bild

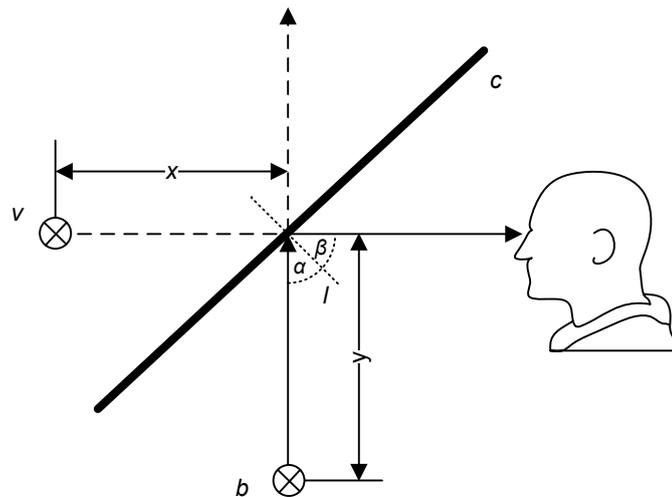


Abbildung 8.2: Die Funktionsweise des HUDs

vom Auge des Betrachters. Diese Entfernung  $x$  kann auch über mehrere Spiegel und weitere Optiken, wie Linsen, vergrößert werden.

Im HUD der BMW 5er Reihe übernimmt die Windschutzscheibe die Funktion des Combiners. Der Fahrer nimmt die Anzeige hier knapp über der Motorhaube schwebend in einer Entfernung von ca. 2,2 m wahr (siehe Abbildung 8.3). Technisch baut das HUD auf einer starken Lichtquelle mit 128 LEDs auf. Über ein TFT-Display mit 65.000 Bildpunkten werden die Informationen auf die Frontscheibe projiziert. Durch den Einsatz von vier Umlenkspiegeln wird die Länge des Strahlenganges vergrößert und der Bauraum entsprechend verkleinert. Verzerrungen, die durch die Biegung der Frontscheibe entstehen, werden ebenfalls kompensiert. Mit den teilweise beweglichen Spiegeln erfolgt die Anpassung des Sichtbereichs, der sogenannten *Eyebox*, an die individuelle Sitzposition des Fahrers. Über einen Helligkeitssensor wird die Leuchtstärke permanent kontrolliert und angepasst, so dass der Fahrer nachts nicht geblendet wird und er die Anzeigen auch bei starker Sonneneinstrahlung ablesen kann. Dabei wird ein Dimmverhältnis von 1000:1 verwendet. [Gei03, Kau04]

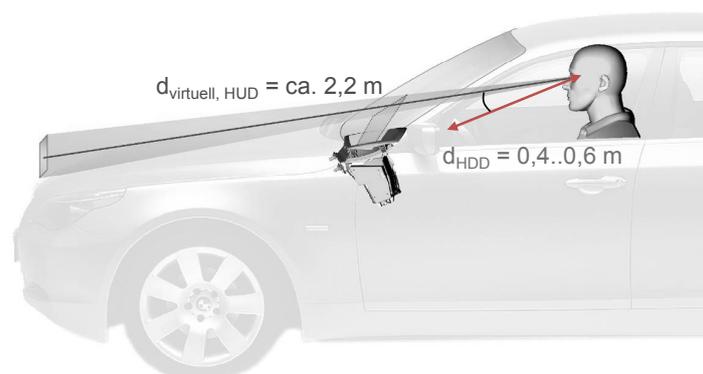


Abbildung 8.3: Prinzip des HUDs im BMW 5er

Die Frontscheibe im Fahrzeug setzt sich aus Stabilitäts- und Sicherheitsgründen aus zwei

Glasscheiben zusammen. Bei einer Projektion auf eine zweiteilige Scheibe führt dies jedoch auch zu zwei Reflexionen, die örtlich versetzt sind, wie in der Abbildung 8.4(a) zu sehen ist. Dieses Problem wird durch eine keilförmige Folie, die zwischen die beiden Scheiben gelegt wird, kompensiert. Dadurch werden die beiden Reflexionen optisch übereinander gelegt, wie in Abbildung 8.4(b) zu sehen ist. [Kau04, EET05]

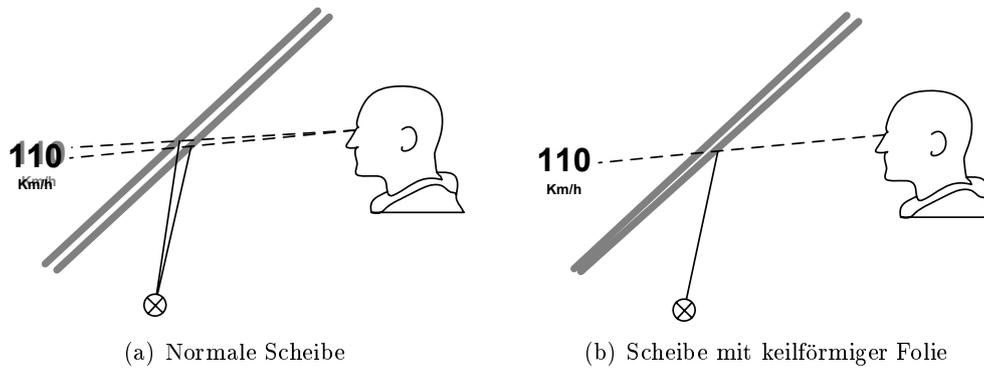


Abbildung 8.4: Korrektur von Doppelbildern bei der Windschutzscheibe

**Visuelle Wahrnehmung:** Die visuelle Sinneswahrnehmung ermöglicht es, komplexe Informationsinhalte aufzunehmen. Über diesen Kanal werden im Mittel 87% aller Informationen wahrgenommen. Damit stellt das Auge das wichtigste Sinnesorgan für die Informationsaufnahme dar. Der Sehbereich des Menschen erstreckt sich vertikal über einen Winkel von  $110^\circ$  und horizontal über einen Winkel von  $170^\circ$ . Dabei nimmt die Sehschärfe von der Mitte zum Rand hin ab. Dies liegt an der Dichteverteilung der Rezeptoren auf der Netzhaut, welche in der Mitte – dem sogenannten Gelben Fleck (fovea centralis) – am höchsten ist und zum Rand hin abnimmt. Das maximale Auflösungsvermögen des menschlichen Auges beträgt dabei etwa 0,5 Bogenminuten. Es werden insgesamt drei Sehbereiche unterschieden (siehe dazu Abbildung 8.5): *fovealer*, *parafovealer* und *peripherer* Sehbereich. [Kor78, Sch93, Sch03b]

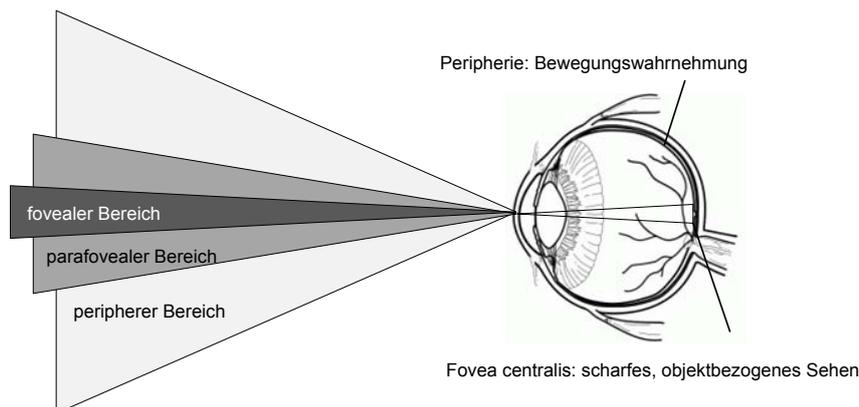


Abbildung 8.5: Querschnitt durch das menschliche Auge

In Tabelle 8.1 sind die relevanten Eigenschaften der einzelnen Sehbereiche zusammengestellt:

**Entfernung der virtuellen Bildebene:** Der empfohlene HUD-Fokus, d. h. die Entfernung der virtuellen HUD-Bildebene, für nicht-kontaktanaloge Anzeigen liegt bei ca. 2 bis 4 m [Kie00,

SICHTBEREICHE	BLICKWINKEL	WAHRNEHMUNG
foveal	2°	objektbezogenes scharfes Sehen
parafoveal	2° bis 10°	Farben, Bewegung, grobe Formen
peripher	größer 10°	Bewegung, Helligkeitsunterschiede

Tabelle 8.1: Sichtbereiche des menschlichen Auges

GS95, Oka96]. Bei einem größeren HUD-Fokus wäre die Diskrepanz zwischen HUD-Bild und der realen Fahrszene vor allem bei kurvigen Strecken zunehmend wahrnehmbar [Oka96].

Jedoch können ältere Autofahrer von dieser virtuellen Entfernung des HUDs profitieren, da die Akkomodationszeit zwischen HUD und Fahrszene kürzer ist, als zu den Anzeigen im Kombi oder CID. Bei Fahrern über 40 Jahren liegt die durchschnittliche Ruheakkommodation bei ungefähr 0,5 Dioptrien [SA90]. Entsprechend 1/Dioptrin wäre dies etwa eine Entfernung von ca. 2 m.

**Positionierung:** Der typische Bildausschnitt (engl. *Field-Of-View*; FOV) eines HUDs liegt bei ca. 5° vertikal und 5° horizontal [GS95]. Die HUD-Informationen sollten jedoch nicht im primären Sehfeld des Fahrers auf die Straße zentriert werden [WE92]. Gerade wenn HUD-Information vom Fahrer als Informationsquelle für sekundäre oder tertiäre Aufgaben genutzt werden. Verschiedene Studien zeigen auch, dass ein nicht-kontaktanaloges HUD zwischen 6° und 10° unterhalb der primären Sehachse des Fahrers positioniert werden sollte [Oka96, GS95]. Um eine Interferenz mit der primären Sehachse des Fahrers zu vermeiden, sollte die horizontale Anordnung um ca. 10° nach rechts verschoben werden.

**Grenzwerte der HUD Symbol-Leuchtdichten:** Allgemein gilt, dass Symbole vor schneeweißem Hintergrund ( $34000 \text{ cd/m}^2$ ) lesbar sein sollten. Als möglicher Richtwert gilt dabei das maximale Leuchtdichte-Kontrast-Verhältnis, das folgendermaßen definiert ist:

$$\frac{L_S + L_H}{L_H} = \frac{1.5}{1} \quad (8.2)$$

wobei  $L_S$  der Leuchtdichte der Symbolik und  $L_H$  der Leuchtdichte des Hintergrundes entspricht. Diese Forderung kann mit der aktuellen Technologie nicht erreicht werden und daher empfehlen [GS95] eine Reduzierung des Leuchtdichte-Kontrast-Verhältnisses auf 1.15:1. Dennoch erreichen heutige HUDs im Fahrzeug nur die maximale Display-Leuchtdichte von  $2842 \text{ cd/m}^2$ . Bei niedrigen Hintergrund-Leuchtdichten wird ein maximales Leuchtdichte-Kontrast-Verhältnis von 4:1 angegeben (siehe dazu [GS95]). Insgesamt muss dieses Verhältnis daher dynamisch an die Tages- und Nachtbedingungen angepasst werden. Aufgrund des sich ständig ändernden Hintergrundes sollte die Verwendung von Farben im HUD eingeschränkt werden. Dabei sollten vor allem Farben verwendet werden, die in der Umgebung selten vorkommen.

**Anzeigengröße:** Die minimale Anzeigengröße von Symbolen wird mit ca. 20 Bogenminuten empfohlen [WE92]. Für die Umrechnung der Symbolgrößen ergibt sich folgende Formel:

$$b = \frac{3438 \cdot h}{d} \quad (8.3)$$

wobei  $b$  die visuelle Größe eines Symbols in Winkelminuten,  $h$  die Symbolhöhe und  $d$  die Entfernung des Symbols zu den Augen des Fahrers ist.

### 8.1.2 Vor- und Nachteile des HUD

Im Folgenden wird auf die wesentlichen Vor- und Nachteile von HUDs im Fahrzeug eingegangen. Einen guten Überblick über diese Aspekte gibt [GS95]. Die wesentlichen Aspekte von HUDs im Flugzeug-Cockpit werden ausführlich in [WE92] diskutiert. Diese Erkenntnisse können jedoch nicht ohne Diskussion auf den Automobilbereich übertragen werden, da die Änderungen in Flugzeugen im Vergleich zum Fahrzeug weniger zeitkritisch sind [Tuf97].

#### HUD Vorteile

Nachfolgend werden die Vorteile des HUDs im Fahrzeug erläutert und diskutiert.

**Verbessertes Blickverhalten zur Verkehrsszene:** Ein großer Hauptvorteil beim Einsatz des HUDs ist das verbesserte Blickverhalten auf die Straße bzw. Verkehrsszene gegenüber Displays im Armaturenbrett [HW02, Kie98]. In der Literatur wird dieser Vorteil meist als *Improved-Forward-Visibility-Claim* oder *Increased-Eyes-on-the-road Time* bezeichnet. [Kie00] stellte fest, dass gerade bei älteren Fahrern durch die Anzeigen im HUD die Reaktionszeiten auf unerwartete Ereignisse auf der Straße signifikant verkürzt werden können.

**Schnellere Reaktion auf Fußgänger:** Des Weiteren stellte Kiefer fest, dass während dem Ablesen von Tachometerinformationen (HUD vs. HDD) eine schnellere Reaktion auf Fußgänger zu messen war [Kie00, Kie98].

**Zeitvorteil beim Ablesen von Informationen:** Vergleicht man das HUD mit dem HDD, dauert ein Blick auf das HUD etwa 0,25 bis 1 s kürzer. Dieser Zeitvorteil ist zurückzuführen auf [GS95]:

- den kürzeren Blickweg auf das HUD; der Blickwinkel zum HUD ist um bis zu 15° kleiner als zum HDD
- die geringere Akkomodation, Nah-/Fernfokussierung des menschlichen Auges; dies wirkt sich insbesondere bei älteren Menschen positiv aus
- den geringen hell/dunkel-Adaptionsaufwand im HUD

#### HUD Nachteile

Im Fahrzeug können jedoch auch folgende Nachteile von HUDs auftreten. Diese werden nun etwas ausführlicher diskutiert, da die Wirkung auf den Fahrer noch teilweise unklar ist.

**Cognitive Capture und Perceptual Tunneling:** Beim Einsatz des Head-Up Displays wird meist auf den *Cognitive Capture* Effekt hingewiesen [Tuf97, GS95]. Bei diesem Effekt verschiebt sich die Aufmerksamkeit des Fahrers weg von der Fahrsituation hin zum Head-Up Display, ohne dass dieser sich dessen bewusst ist. Die Folge kann sein, dass in einer kritischen Situation die Reaktion verspätet oder gar nicht erfolgt. Dieser Effekt kann den eigentlich beabsichtigten Vorteilen des Head-Up Displays entgegenwirken.

Die Studie von [HW02] hat insgesamt gezeigt, dass eine Zweitaufgabe im HUD und HDD zu einer erhöhten kognitiven Beanspruchung führt, jedoch die Reaktionszeit auf ein Hindernis bei der Durchführung der Aufgabe im HUD im Gegensatz zum HDD signifikant kürzer ist. In einer Studie des US Department of Transportation wurde als Weiteres die Tauglichkeit des HUD gegenüber einem HDD für Navigationszwecke ermittelt und dabei keine negativen Einflüsse des HUDs festgestellt [HG98]. Jedoch wurde auf eine mögliche Gefahr durch Cognitive Capture bei

großer Informationsmenge, neuen oder bewegten Symbolen hingewiesen. Dabei ist der Inhalt und die Struktur des HUDs von entscheidender Bedeutung [GS95].

Als Folge des Cognitive Captures könnte in diesem Zusammenhang auch das *Perceptual Tunneling* auftreten (ebenfalls Tunnelblick genannt). Hierbei verengt sich der Aufmerksamkeitsfokus auf die Anzeigen des Head-Up Displays. Objekte und Reize im peripheren Sichtfeld können dabei nur noch eingeschränkt wahrgenommen werden. Die Studie von [RN03] konnte über einen PDT-Test eine kognitive Beanspruchung und geändertes Blickverhalten bei Verwendung des HUDs gleichverteilt im gesamten Blickfeld feststellen, jedoch konnte hier kein Tunneling nachgewiesen werden.

**Entfernungsüberschätzungen:** Bei einer geringen virtuellen HUD-Bildebene kann es zu einer Entfernungsüberschätzung von Objekten (wie z. B. Fahrzeugen) kommen, die außerhalb der fokussierten Ebene liegen [Tuf97]. Ursache hierfür ist die Akkommodation des Auges auf die HUD-Symbolik. In reizreichen Umgebungen kann dieser Effekt jedoch weitgehend vernachlässigt werden, da die Augen des Fahrers in der Regel auf die Reize akkommodiert sind [WE92, SA90]. Dagegen könnte in reizarmen Situationen, wie dichtem Nebel oder bei Nachtfahrten mit wenig Verkehr, z. B. der Abstand zu vorausfahrenden Fahrzeugen überschätzt werden.

## 8.2 Innovatives, multimodales Anzeige-/Bedienkonzept im HUD

In diesem Abschnitt wird ein neuartiges, innovatives Anzeige und Bedienkonzept für ein großflächiges, schwebendes Head-Up Display vorgestellt. Ziel des Konzeptes ist es, die kognitive Belastung und die Blickabwendung des Fahrers weitgehend zu reduzieren und entsprechend den verschiedenen Aufgaben des Fahrers eine optimale Passung mit der Fahraufgabe zu erreichen. Dieses Konzept löst sich bewusst von konventionellen Informationsgliederungen im Fahrzeug, die zum Teil auf historisch gewachsenen und theoretischen Prinzipien beruhen.

### 8.2.1 Gesamtkonzept HUD\_MAX für ein großes schwebendes HUD

Das entwickelte Gesamtkonzept HUD\_MAX untergliedert sich in drei Anzeigenbereiche, dem *Priorität-I*-Bereich, den *Priorität-II*- und den *Priorität-III*-Bereichen. Diese erstrecken sich entlang des unteren Randes der Windschutzscheibe, der Blick des Fahrers auf die Straße wird jedoch nicht beeinträchtigt. In dem integrierten HUD\_MAX-Konzept verschmelzen Inhalte der primären, sekundären und tertiären Fahraufgabe. Stattdessen orientiert sich das HUD\_MAX-Konzept hauptsächlich an der Wichtigkeit und Relevanz der Informationen für den Fahrer zum jeweiligen Kontext. Studien zeigen, dass die Fahrer nicht strikt zwischen primären, sekundären und tertiären Aufgaben unterscheiden, sondern sich eher an Relevanz und Priorität von Informationen zur gegenwärtigen Situation orientieren. Beispielsweise könnte ein eintreffender Telefonanruf wichtiger sein, als ein nicht relevantes Verkehrsschild.

In Abbildung 8.6 wird das entwickelte HUD\_MAX-Konzept schematisch dargestellt. Die gesamte Darstellung des HUDs liegt unter dem Hauptsichtbereich des Fahrers auf die Straße. Die drei Bereiche gruppieren sich um das zentrale Sichtfeld des Fahrers. Analog zur Wichtigkeit des jeweiligen Anzeigenbereichs werden diese in drei unterschiedlich weit entfernten Blickabständen angeordnet. Wie ein solches Display in Aktion aussehen kann, ist exemplarisch in Abbildung 8.7 zu sehen. Nun werden die Inhalte der drei Anzeigenbereiche detailliert erläutert.

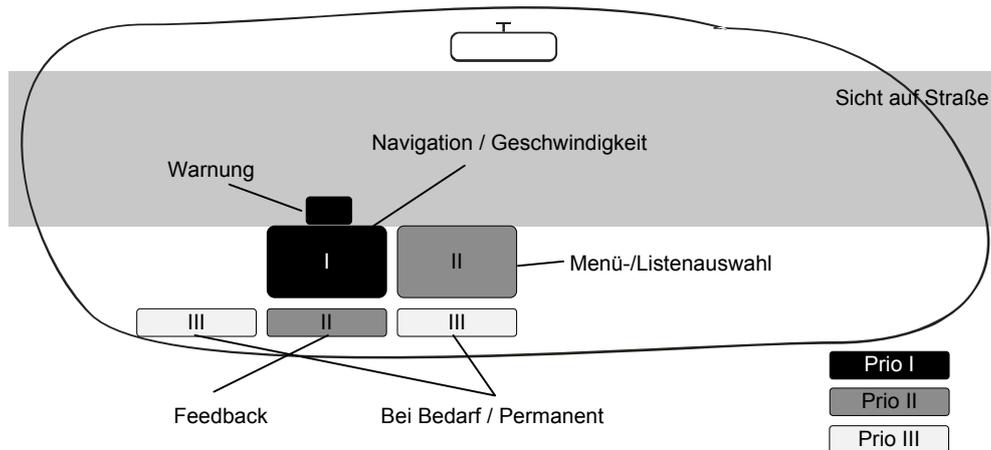


Abbildung 8.6: Überblick HUD\_MAX-Konzept

**Anzeigebereich Priorität I:** Dieser zentrale Priorität I Bereich wird überwiegend mit Informationen der primären und zum Teil sekundären Fahraufgabe belegt. Dies umfasst zum Einen permanente, für die Fahrzeugführung wichtige Informationen, wie z. B. die aktuelle Geschwindigkeit, aber auch Informationen die indirekt mit ihr zusammenhängen, wie die Drehzahl. Die Position und Größe des Anzeigebereichs orientiert sich an der von aktuellen Serien-HUDs wie vom BMW 5er. Als weiteres werden kritische Informationen der Fahrerassistenzsysteme situativ eingeblendet bzw. überblendet. Dies können Warnungen von Spurhaltesystemen oder ACC-Informationen (Active Cruise Control) sein (siehe Abbildung 8.7). Um kritische Informationen sofort wahrnehmen zu können, befindet sich dieser Anzeigebereich möglichst nahe am Blickfeld des Fahrers, beeinträchtigt aber den Blick auf die Straße nicht.

**Anzeigebereiche Priorität II:** Die beiden Priorität II Bereiche grenzen sowohl in horizontaler als auch in vertikaler Richtung an den Priorität I Bereich an. Diese Anzeigebereiche enthalten Informationen, welche temporär für den Fahrer von großer Bedeutung sind, jedoch nicht kritisch mit der Fahrzeugführung in Verbindung stehen. Die Positionen dieser Bereiche werden ebenfalls in Abbildung 8.6 dargestellt. Am unteren Rand, unterhalb des Priorität I Bereichs, findet man die *Feedbackzeile*. Die Idee der Feedbackzeile ist es, zum Einen den Fahrer auf systeminitiierte Meldungen aufmerksam zu machen, beispielsweise das Anzeigen der Nummer eines eintreffenden Telefonanrufs oder die Anzeige einer geringen Tankreichweite. Erfordert die Meldung eine Interaktion des Fahrers, dann wird der Dialog im Fahrerinteraktionsbereich (s. u.) fortgesetzt. Zum Anderen gibt dieser Bereich temporär Feedback bei direkter Interaktion mit verschiedenen Bedienelementen (wie Lenkstockhebel, Hardkeys oder Drehräder) an. Dabei wird der Status des aktivierten Bedienelements (ggf. auch berührsensitiv) rückgemeldet und ggf. weitere Informationen (z. B. Hilfe) dazu gegeben. Feedback wird beispielsweise beim Skippen eines Radiosenders, bei der Betätigung des Scheibenwischers oder beim Aktivieren der Beleuchtung gegeben. Die Informationsmenge wird jedoch weitgehend minimiert und daher auch nur situativ rückgemeldet.

Unmittelbar rechts neben dem Priorität I Bereich wird der *Fahrerinteraktionsbereich* definiert. Dieser Bereich ermöglicht Ja/Nein-Entscheidungen, die Auswahl aus Listen und die Bedienung eines reduzierten Menüs im HUD. Das genaue Menüinteraktionskonzept wird im Abschnitt 8.2.2 ausführlich beschrieben.

**Anzeigebereich Priorität III:** Der Priorität III Bereich besteht aus Informationen der sekundären und tertiären Fahraufgabe, wie Status- und Bordcomputerinformationen und wird entlang des unteren Rands der Windschutzscheibe angezeigt (links und rechts neben der Feedbackzeile, siehe Abbildung 8.6). Hier werden Informationen angezeigt, die sich entweder nur sehr langsam ändern oder Statusinformationen verschiedener Geräte beinhalten. Aufgrund der geringsten Wichtigkeit dieser Anzeigen werden diese am weitesten entfernt vom Fahrerblickfeld dargestellt. Die vom Fahrer angeforderten Informationen werden permanent oder temporär eingeblendet. Der Fahrer kann nach individuellen Anforderungen die angezeigten Informationen selbst festlegen. Beispielsweise kann dies die Tankreichweite, die Tages- oder Gesamtkilometeranzeige, der aktuelle Radiosender, das gewählte Ziel, die Ankunftszeit, die aktuelle Uhrzeit oder das Datum sein. Diese Informationen dienen in erster Linie zur Befriedigung des individuellen Informationsbedürfnisses.

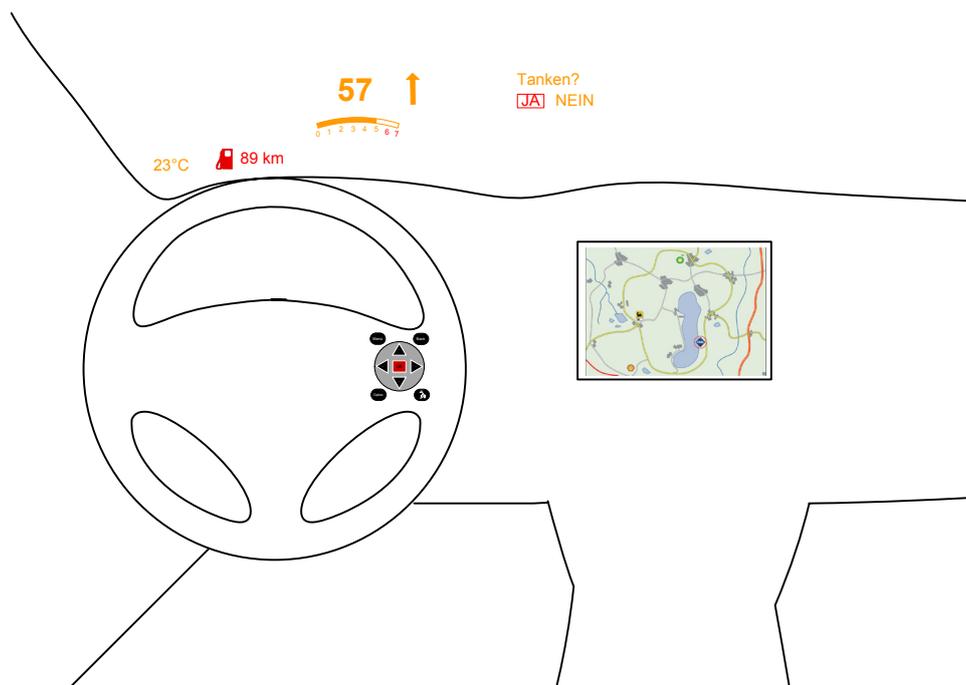


Abbildung 8.7: Beispieldarstellung des Lösungskonzeptes HUD\_MAX

### 8.2.2 HUD Menüinteraktion MATRIX\_MMI

Nun wird der Fahrerinteraktionsbereich des HUD\_MAX-Konzeptes erläutert. Wie bereits beschrieben, verwendet das HUD\_MAX-Konzept den Priorität II Bereich für die Interaktion. Dieser befindet sich direkt rechts neben dem Priorität I Bereich. Um eine Interferenz mit der primären Fahraufgabe zu vermeiden, ist der Menüinteraktionsbereich von der vertikalen Sehachse verschoben.

Für den Interaktionsbereich wurde das sogenannte *MatrixMMI*-Konzept entwickelt, das in erster Linie Basisfunktionen für den Fahrer zur Bedienung im HUD bereitstellt. Daher wurde im Speziellen darauf geachtet, dass das Menü sehr übersichtlich und einfach zu bedienen ist. Dabei wird das Menü nur bei der Bedienung bzw. nach Bedarf eingeblendet (durch Berührung des Bedienelementes am Lenkrad). In der Abbildung 8.8 sind drei Hierarchiestufen des Menüs

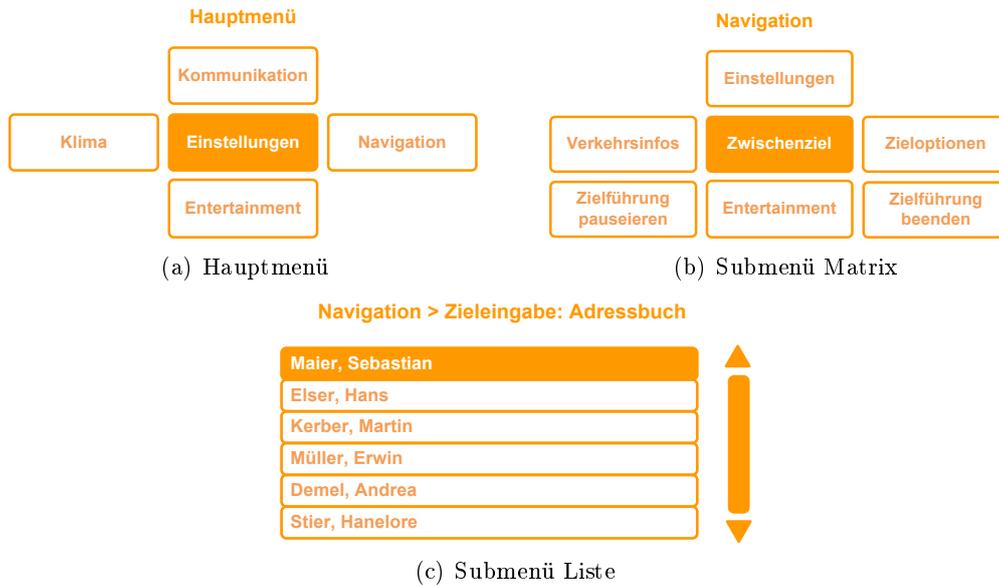


Abbildung 8.8: Matrix Menü

abgebildet. Auf den beiden obersten Hierarchiestufen verwendet das MatrixMMI eine 3x3 Matrix zur Präsentation von Kategorien und Funktionen. Die Darstellung erlaubt ein Maximum von neun Elementen. Um eine Informationsüberladung zu vermeiden, wird das Hauptmenü auf fünf Elemente reduziert. In tieferen Ebenen kann die Anzahl an Elementen auf maximal neun erhöht werden. Das Layout ändert sich in eine scrollbare Liste, sobald Daten angezeigt werden. Beispiele hierfür sind Adress-, Ziel- und Musiklisten.

Für die Bedienung des Menüs wird ein 5-Wege-Navipad am Lenkrad verwendet (siehe Abschnitt 8.2.3). Das Bedienelement enthält vier Navigationsknöpfe, um den Cursor nach oben bzw. unten und links bzw. rechts bewegen zu können. In der Mitte ist der Bestätigungsknopf positioniert und gut blind zu ertasten. Die beiden oberen Tasten sind mit den Funktionen Escape und Hauptmenü belegt.

Das aktuell fokussierte Feld wird orange hintermalt. Zur Selektion der Position muss der Bestätigungsknopf gedrückt werden. Dadurch springt das Menü zur nächsten Ebene bzw. es wird eine Funktion aktiviert. Mit dem Escape-Knopf springt man eine Ebene im Menü zurück und mit dem Hauptmenü-Knopf ist man in der Lage, zu einem beliebigen Zeitpunkt in die oberste Menüebene zu gelangen.

### 8.2.3 Multimodale Bedienung

Für die Bedienung des HUD\_MAX-Konzeptes wird eine multimodale Kombination aus 5-Wege-Navipad und Sprache verwendet. Beide Kanäle können parallel bedient werden. Erweiterte Funktionen werden darüber hinaus über einen Touchscreen in der Mittelkonsole bedient.

**5-Wege-Navipad:** Das 5-Wege-Navipad ist Teil des MatrixMMI-Konzeptes und befindet sich auf der rechten Seite des Lenkrades (siehe Abbildung 8.10). Der Fokus des Menü-Cursors wird über vier Richtungsknöpfe gesteuert. Die Auswahl erfolgt über den O.K.-Button bzw. Bestätigungsknopf im Zentrum des Bedienelementes. Um zum Ausgangspunkt zurückzukehren, existiert ein Hauptmenü-Knopf (oben links). Im Falle einer falschen Eingabe kann der Nutzer

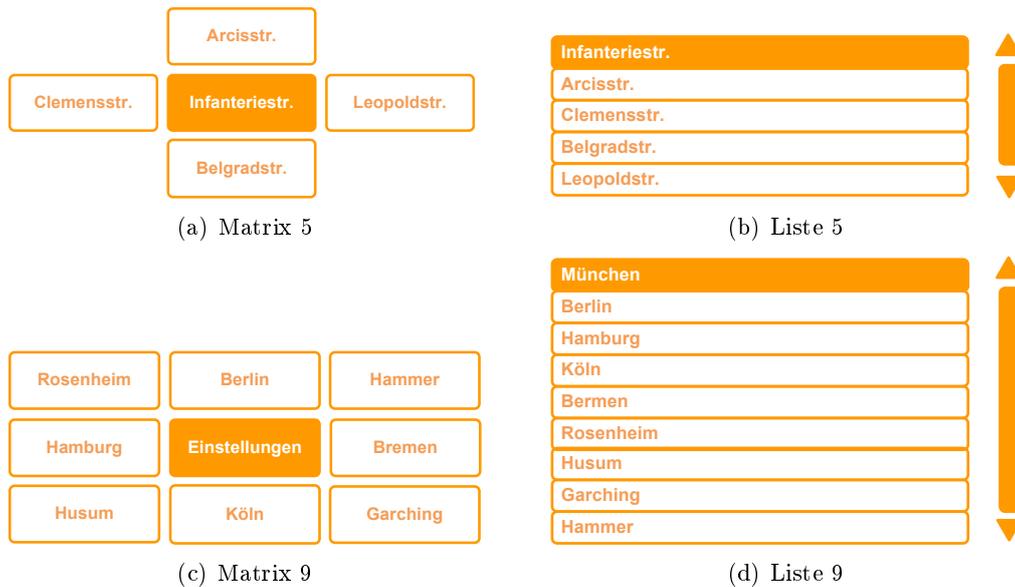


Abbildung 8.9: Überblick Matrix- und Listenstruktur

mit dem Esc-Button (oben rechts) zur vorhergehenden Menüebene zurückspringen und Fehler korrigieren.

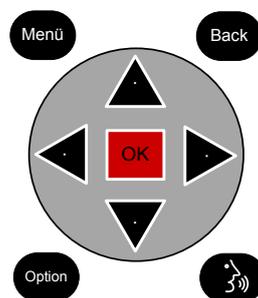


Abbildung 8.10: Bedienelement am Lenkrad: 5-Wege-Navipad

**Sprache:** Die taktile Eingabe wird sehr gut durch die Sprachbedienung ergänzt. Sprache kann die Belastung weiter reduzieren und eignet sich deshalb besonders gut für das HUD. Sämtliche Menüeinträge können über Sprachkommandos direkt ausgewählt werden (*speak-what-you-see*) und ermöglichen Shortcuts zu Untermenüs. Des weiteren ermöglichen Kommandos das Ein- und Ausblenden gewünschter Anzeigen des Priorität III Bereichs. Beispielsweise kann über den Befehl „Temperatur“ die Anzeige der Außen- und Innentemperatur erfolgen. Akustisches



Abbildung 8.11: Verarbeitungsstatus

Feedback wird insgesamt sparsam eingesetzt, um den Fahrer und Mitfahrer möglichst wenig akustisch zu belasten. Stattdessen erfolgt meist visuelles Feedback. So wird dem Fahrer selbst

überlassen, zu welchem Zeitpunkt er freie Ressourcen für die Informationsaufnahme hat. Der Zustand des Spracherkenners wird durch ein farbkodiertes Speechy-Symbol in der Feedback Leiste rückgemeldet (siehe Abbildung 8.11). Grün bedeutet, dass gesprochen werden kann, orange, dass gerade das Kommando verarbeitet wird und rot weist auf einen Fehler bei der Erkennung hin. Die erkannte Intention wird anschließend kurz in der Feedbackzeile eingeblendet. Mögliche Arten des Feedbacks werden in Abbildung 8.12 dargestellt. Müssen Entscheidungen getroffen werden, erfolgt die Auflistung der Optionen im Interaktionsbereich. Anschließend kann hier die Auswahl durch das Sprechen der gewünschten Option oder auch taktil über das 5-Wege-Navipad erfolgen. Darüber hinaus können auch natürlich-sprachliche Eingaben ermöglicht werden. Über den Frame-Based Ansatz können hier mehrere Intentionen in einem Dialogschritt in beliebiger Reihenfolge übermittelt werden (siehe Abschnitt 3.1.3). Dazu erfolgt die Intentionsinterpretation anhand des NLU aus dem Kapitel 3. Dieser natürlich-sprachliche Ansatz kommt bei den Informationsagenten zum Einsatz (siehe nächster Abschnitt).



Abbildung 8.12: Sprachfeedback in der Statusleiste

**Touchscreen im CID:** Zur Ergänzung der Anzeigen im HUD erfolgt eine Anzeige von erweiterten Informationen im CID. Das CID kann durch einen Touchscreen sehr direkt bedient werden (siehe 8.2.3). Im Normalzustand wird der aktuelle Kartenausschnitt dargestellt. Außerdem können zum aktuellen Dialogschritt ergänzende Informationen angezeigt werden. Beispielsweise können POIs zusätzlich zur Listendarstellung im HUD auch mit geographischen Labels oder Icons auf der Karte im CID visualisiert werden. Um eine Auswahl zu treffen, wählt der Fahrer direkt das gewünschte Objekt über den Touchscreen. Neben der Karte verfügt das CID über ein erweitertes Menü, das umfassender als das Basismenü im HUD ist. Es kann von Fahrer und Beifahrer direkt über das Touchscreen bedient werden. Auf die Funktionsweise des erweiterten Menüs wird hier nicht näher eingegangen.



### Informationsagenten im HUD

Die im Kapitel 6 vorgestellten *Informationsagenten* sind in das HUD\_MAX-Konzept integriert. Die Dialoge werden dabei knapp und einfach gehalten. Die Informationsagenten führen den Fahrer multimodal und schrittweise über einfache Dialoge im Menüinteraktionsbereich des HUDs und durch Frame-Based Sprachdialoge (siehe Abschnitt 3.1.3 und 8.2.3). Beispielsweise unterstützt der Tankagent den Fahrer mit Vorschlägen zu geeigneten Tankstellen (siehe Abbildung 8.13), der Restaurant-Recommender kann helfen, präferierte Restaurants schneller zu identifizieren (siehe Abbildung 8.14). Die Ausgabe möglicher POIs erfolgt als Liste im Interaktionsbereich. Diese können dann über Sprache oder das 5-Wege-Navipad angewählt werden.

- Tankagent



(a) Tankdialog



(b) Auswahl Tankstelle aus Liste

Abbildung 8.13: Tankagent

- Restaurant-Recommender



(a) Suche nach Essensmöglichkeiten



(b) Auswahl Restaurant

Abbildung 8.14: Restaurantagent

### 8.3 Integriertes HUD-Kombi-Konzept

In den kommerziell verfügbaren Konzepten dient das HUD als redundanter Darstellungsort bei dem einzelne Informationen des analogen Kombiinstruments dargestellt werden. Der technologische Trend zu einem freiprogrammierbaren, farbigen Graphikdisplay für das Kombi schafft neue Möglichkeiten. In diesem Abschnitt werden Ansätze diskutiert, die ein gutes Zusammenspiel zwischen HUD und Kombi ermöglichen. Die Vorteile der jeweiligen Anzeigertechnologien im Fahrzeug sollen optimalerweise zusammengeführt werden und so ein ganzheitliches, integriertes Konzept darstellen.

#### 8.3.1 Prioritätenorientiertes Informationssplitting

Über das sogenannte *Informationssplitting* ergänzen sich die Anzeigen im HUD und im Kombi in idealer Weise (siehe Abbildung 8.15). Situativ relevante und wichtige Infos für den Fahrer werden dabei im HUD dargestellt, erweiterte und detaillierte Infos finden hingegen im Kombi ihren definierten Anzeigort.

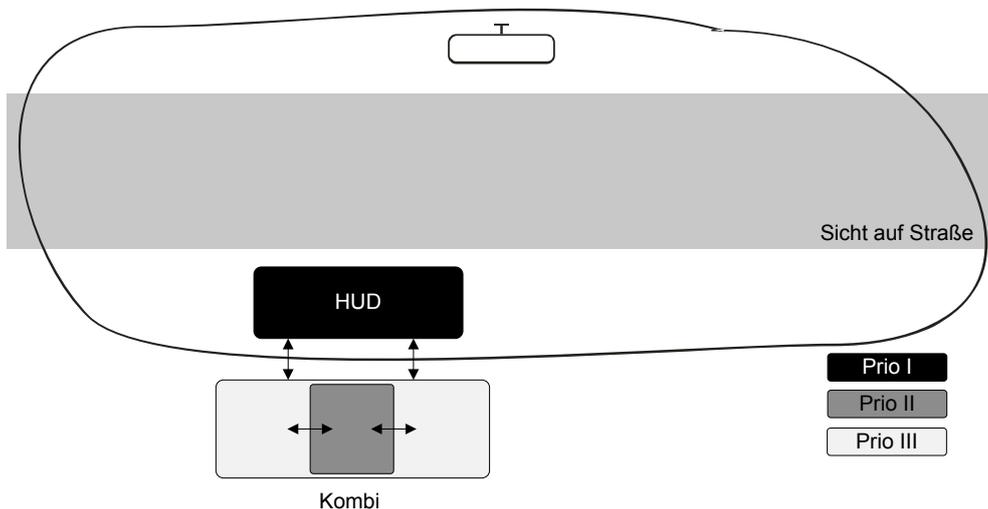


Abbildung 8.15: Schematischer Überblick des integrierten HUD-Kombi-Konzeptes

**Primäre Anzeigen:** Anstatt den Fahrer mit festen Informationsanzeigen im HUD permanent zu belasten, findet ein geregeltes Anzeigemanagement zwischen HUD und Kombi statt. Ziel des entwickelten Konzeptes ist es, dem Fahrer primäre Anzeigen möglichst schnell und intuitiv erfassbar zu präsentieren. Dazu wurde eine Fahrzeugsymbolik gewählt, in der die entsprechenden Informationen integriert dargestellt werden. Die Aktivierung der primären Anzeigen, sowie akute Warnungen werden im HUD kurzzeitig visualisiert und wandern dann zur detaillierten Darstellung in das Kombidisplay (Home-Position). Auch Navigationsanzeigen erfolgen in einem ergänzenden Zusammenspiel zwischen Kombi und HUD. Das Kombi stellt dabei in der Kartenansicht einen Umgebungsüberblick dar. Gelangt das Fahrzeug in den Nahbereich einer navigationsrelevanten Position, erfolgt ergänzend ein Navigationshinweis in Form eines Pfeils im HUD. Die situationsadaptive Navigationsausgabe im Kombi passt sich dabei an die Komplexität der Straßen- bzw. Navigationsssituation an, d. h. je komplexer die Navigation, desto detaillierter die Kartenanzeige.

**Menüdarstellung:** Neben der Bedienung des Menüs im Kombi, der Home-Position, ist auch die Bedienung über das HUD möglich. Ein Vorteil der Bedienung im HUD könnte u. a. die verringerte Blickabwendungszeit sein. Durch Menüsplitting stehen dem Fahrer folgende Darstellungsvarianten zur Verfügung:

- Reine Darstellung im Kombi-Instrument
- HUD-orientierte Bedienung: Anzeige von Menüebene und Menüpfade im HUD
- Kombi-orientierte Bedienung: Anzeige der aktuellen Menüebene im HUD, jedoch die Menüstruktur im Kombi

Ein Hardkey ermöglicht den Wechsel zwischen den verschiedenen Darstellungsprofilen. Ein wesentlicher Vorteil der HUD-orientierten Menübedienung ist die parallele fahrerzentrierte Darstellung der Navigationskarte im Kombidisplay.

### 8.3.2 Fahrerzentrierter Kombi-Fokuswechsel

Durch einen Button am Lenkradbedienelement kann ein Fokuswechsel im Kombi (zwischen den Anzeigenbereichen Prio II und Prio III) initiiert werden (siehe Abbildung 8.15). Hier kann zwischen FAS-Symbolik, der Karte und dem GUI-Bereich gewechselt werden. Die gewünschte Anzeige wandert anschließend von einer der äußeren Positionen (Prio III) in die mittlere, zentrale Position (Prio II) des Kombis. So kann der Fahrer die für ihn interessante Anzeige in den Vordergrund rücken.

## 8.4 Kontaktanaloges Head-Up Display

*Kontaktanalogie* beschreibt eine Informationsdarstellung, welche die Anzeigen perspektivisch korrekt in die Realität integriert [Bub76, TLKB06]. Im Idealfall soll die Darstellung für den Benutzer eine direkte Beziehung zur Umwelt verdeutlichen. Davon erhofft man sich eine Erhöhung der Sicherheit durch die geringere Ablenkung des Fahrers und eine Erhöhung des Fahrkomforts. Eine solche Informationsdarstellung ist im Bereich *Erweiterte Realität (Augmented Reality (AR))*, der im folgenden Abschnitt näher erläutert wird, angesiedelt. Kontaktanalogie bezeichnet damit nur die Art der Informationsdarstellung und nicht die Information an sich. Wichtige Anzeigen für ein kontaktanaloges HUD im Fahrzeug sind z. B. Navigationshinweise, Warnungen, Informationen zu POIs. Abbildung 8.16 zeigt ein Beispiel für die kontaktanaloge Darstellung eines Navigationshinweises.

In diesem Abschnitt werden zunächst Grundlagen zum Thema Augmented Reality erläutert. Anschließend werden mögliche Darstellungskonzepte für ein kontaktanaloges HUD vorgestellt und ein Ansatz zur prototypischen Realisierung eines quasi-kontaktanalogen HUDs präsentiert.

### 8.4.1 Augmented Reality

Zwischen der realen Welt (engl. *reality*) und der virtuellen Welt (engl. *virtual reality*) definiert Milgram einen Bereich, der als gemischte Realität (engl. *mixed reality*) bezeichnet wird [MK94, Azu97]. In der gemischten Realität verschwimmen die Grenzen zwischen virtuellen und realen Objekten. Dieser Bereich lässt sich noch weiter unterteilen und unterscheidet sich im Grad dessen, was der Benutzer noch von der realen Umgebung wahrnimmt. Unter der Erweiterten Realität versteht man die Anreicherung der realen Welt mit rechnergenerierten virtuellen



Abbildung 8.16: Kontaktanaloge Darstellung eines Navigationspfeils

Objekten, welche situationsgerecht im Kontext der zu betrachtenden Realität direkt in das Sichtfeld des Betrachters eingeblendet werden können.

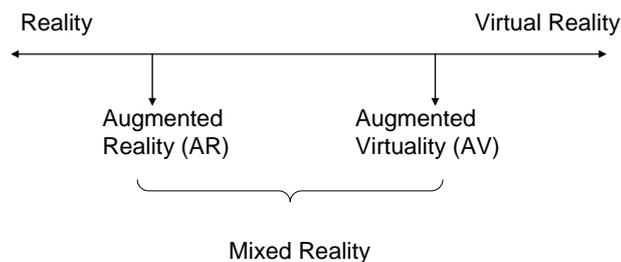


Abbildung 8.17: Zusammenhang Realität-Virtualität [Azu97]

Um die reale Welt mit virtuellen Informationen überlagern zu können, sind drei technische Komponenten nötig:

- **Trackingsystem:**  
Ein Trackingsystem erfasst die Position und Blickrichtung eines Benutzers und ggf. weiterer Objekte. Die Daten werden an den Visualisierungsrechner weitergeleitet.
- **Visualisierungsrechner:**  
Er funktioniert als Schnittstelle zwischen Trackingsystem und Display und berechnet die perspektivisch korrekte virtuelle Szene. Je nach Anwendung ist die Software des Visualisierungsrechners das eigentliche Herzstück einer AR Anwendung. Dabei werden hohe Anforderungen an die Echtzeitfähigkeit der Software gestellt.
- **Visuelle Ausgabereinheit:**  
Anschließend werden die berechneten virtuellen Szenen mit der Realität über eine Ausgabereinheit überlagert. Abhängig von der Anwendung kann es sich dabei um ein Head-Mounted Display (HMD) oder in unserem Fall um ein HUD handeln.

## 8.4.2 Kontaktanaloge Konzeptentwürfe

### Kontaktanaloge Darstellungen für die primäre Fahraufgabe

In diesem Abschnitt werden nun mögliche kontaktanaloge Konzeptideen vorgestellt (siehe auch [Abl07]). Zur Unterstützung der primären Fahraufgabe wurden folgende Konzeptvarianten erarbeitet. Detaillierte Untersuchungen von kontaktanalogen HUD-Konzepten finden sich in [Lan08, Bub76, TLKB06].

**Navigationsebene:** Die folgenden Konzepte dienen der Unterstützung des Fahrers bei der Navigation:

- **Pfeildarstellung:**

Äquivalent den Richtungshinweisen auf der Straße, können dem Fahrer auf der Straße liegende, virtuelle Pfeile zur Navigationsunterstützung eingeblendet werden (siehe Abbildung 8.18(a)).

- **Kieselsteindarstellung:**

Die Aneinanderreihung von Punkten auf der Straße markiert die zu wählende Route (siehe Abbildung 8.18(b)).

- **Taxiway:**

Bei dieser Darstellung wird die zu wählende Fahrstrecke durch farbige Hinterlegung hervorgehoben (siehe Abbildung 8.18(c)).

- **Straßenname:**

Die Einblendung der Straßennamen erfolgt bei kreuzenden Querstraßen (siehe Abbildung 8.18(g)). So kann der Fahrer bei der Orientierung in unbekanntem Städten unterstützt werden.

- **Verkehrsschilder:**

Wichtige Verkehrsschilder werden auf der Straße kontaktanalog eingeblendet (siehe Abbildung 8.18(d)). So können sicherheitsrelevante Schilder verbessert wahrgenommen werden.

- **Spurwechsel:**

Es erfolgt die Empfehlung eines erforderlichen Spurwechsels, so dass der Fahrer frühzeitig auf erforderliche Fahrmanöver aufmerksam gemacht wird (siehe Abbildung 8.18(f)).

**Längs- und Querregelung:** Zur Unterstützung der Stabilisierungsebene werden folgende Konzeptvarianten vorgeschlagen:

- **Fahrbahnmarkierung:**

Um die Fahrbahnmarkierungen bei schlechten Sichtverhältnissen hervorzuheben, werden diese mittels kontaktanaloger Begrenzungslinien überblendet (siehe Abbildung 8.18(e)).

- **Abstandsbalken:**

In Abhängigkeit der Entfernung zu einem vorausfahrenden Fahrzeug wird ein Abstandsbalken unterschiedlicher Farbe eingeblendet (siehe Abbildung 8.18(h)). Gelb symbolisiert einen ausreichenden Abstand, rot signalisiert einen Gefahrenbereich.



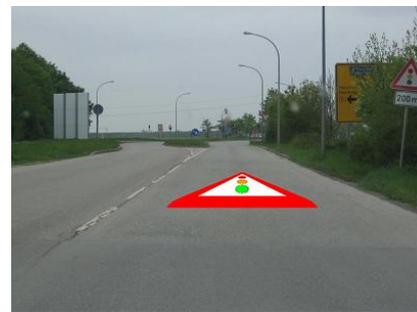
(a) Pfeildarstellung



(b) Kieselsteindarstellung



(c) Taxiwaydarstellung



(d) Verkehrsschilder



(e) Fahrbahnmarkierung



(f) Spurwechsel



(g) Straßenname



(h) Abstandsbalken

Abbildung 8.18: Kontaktanaloge Darstellungsarten

- **Gefahren:**

Schwächere Verkehrsteilnehmer, wie Fußgänger, Radfahrer oder Wildwechsel können durch kontaktanaloge Einrahmung bei Tag und Nacht hervorgehoben werden. Hierbei wäre besonders die Koppelung mit einem Nightvision System sinnvoll.

### Kontaktanaloge Anzeige von POIs

In Abbildung 8.19(a) werden relevante POIs direkt im Blickfeld des Fahrers hervorgehoben. Um den Fahrer beispielsweise auf eine Sehenswürdigkeit aufmerksam zu machen, kann diese durch farbliches Highlighten im HUD hervorgehoben werden. Ein zusätzliches Textfeld gibt weitere Detailinformationen. In einer Konzeptvariante werden bei einer Hotelsuche die sich im Blickfeld befindenden Hotels markiert. POIs, die sich nicht im Blickfeld befinden, werden durch ein Icon, das die Himmelsrichtung angibt, angedeutet. Eine weitere Darstellungsalternative von nicht sichtbaren POIs ist eine vertikale Linie mit einem Icon am oberen Rand der Windschutzscheibe.



(a) Touristenführer



(b) Hotelführer



(c) Hotel Labeling



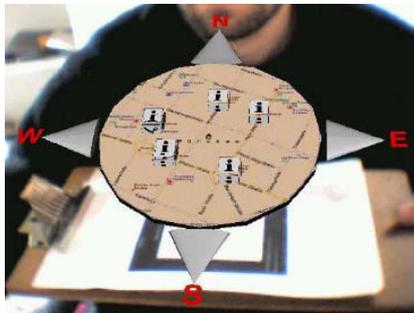
(d) Kompassdarstellung

Abbildung 8.19: Hervorhebung von POIs im HUD

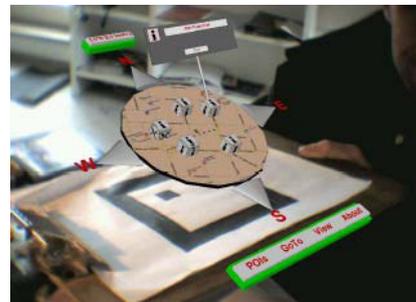
Oftmals möchte der Fahrer die relative Position der POIs zur aktuellen Position des Fahrzeuges wissen. In dem Szenario von Abbildung 8.19(d) zeigt das sogenannte Kompassmodell lediglich die relative Richtung über einen dreidimensionalen Pfeil im Blickfeld des Fahrers an. Dieser Pfeil zeigt stets in die Richtung des POIs. Die Größe des Pfeils dient als Indikator für die Entfernung.

**Kompasskonzept:** Das oben vorgestellte Kompassmodell wurde nun mit einer Kartendarstellung kombiniert. Die virtuelle Karte wird augmentiert im HUD auf der Motorhaube des Fahrzeuges präsentiert. Die Karte informiert den Fahrer über seine aktuelle Position und

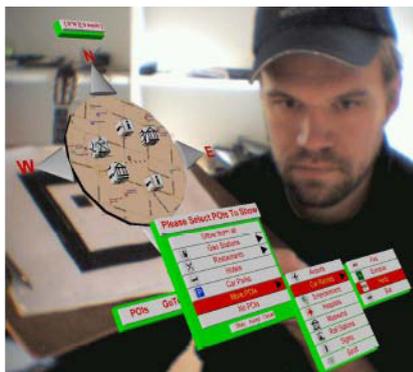
Fahrtrichtung. Die Daten werden über einen GPS-Receiver geliefert. Die 3D-Karte hat eine runde Form um den Eindruck des Kompasses darzustellen. Das Zentrum der Karte zeigt die gegenwärtige Position. Ein POI wird über einen Quader mit zugehöriger Textur symbolisiert. Interessante POIs werden dreidimensional in Form eines Quaders mit geeigneter Textur in der Karte symbolisiert. Auch eine Interaktion mit der Karte ist möglich und der Fahrer kann über ein Bedienelement, z. B. ein POI, auswählen oder weitere Informationen anfordern. Darüber hinaus verfügt die Anwendung auch über ein 3D-Menü, das unterhalb der Karte angezeigt wird. Über dieses Menü kann der Nutzer spezifische Einstellungen der Applikation vornehmen. Um die Darstellung der Karte anzupassen, gibt es ein Menü, das aus den Einträgen POIs (Auswählen der POI Kategorien), GoTo (Kartenzentrum definieren), View (Art der Kartendarstellung) und About (weitere Informationen) besteht. Über dieses 3D-Menü kann der Nutzer die POI Kategorie wählen. Die 24 Kategorien sind nach Wichtigkeit und Alphabet geordnet. Wenn ein Nutzer ein POI auf der virtuellen Karte direkt anwählt, erfolgt ein Pop-Up-Feld. Auf dem Feld wird der Name und die Kategorie dargestellt.



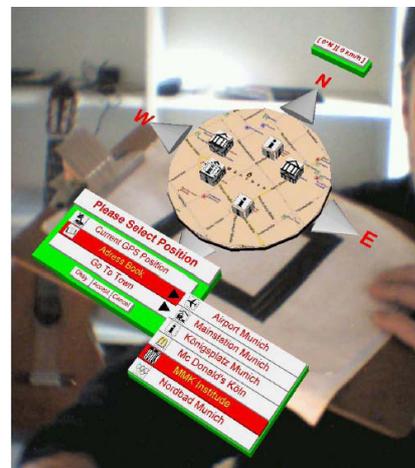
(a) Kartendarstellung



(b) Hauptmenü



(c) Untermenü POI



(d) Untermenü GoTo

Abbildung 8.20: Überblick Kompasskonzept

### 8.4.3 Sitzkistenaufbau eines quasi-kontaktanalogen HUDs

Üblicherweise wird ein HMD für die Anzeige von AR Inhalten verwendet, welches die Form eines Helmes oder einer Brille hat und vom Benutzer getragen wird. Für den Einsatz im Auto

ist dieses Verfahren jedoch nicht zweckmäßig, stattdessen wird hier die Windschutzscheibe als halb transparente Anzeigefläche für die Anzeige der virtuellen Informationen genutzt. Dazu wird ein prototypischer Aufbau eines quasi-kontaktanalogen HUDs in Form einer Sitzkiste realisiert (siehe Abbildung 8.21).

Die bewegliche, auf Rollen gelagerte Sitzkiste besteht aus einem Sitz, einem Steuerrad, einem Display und einer Combinerscheibe. Die Position und Orientierung des prototypischen Aufbaus wird über ein Infrarot-Trackingsystem getrackt. In dieser Arbeit wurde ein auf Infrarot basierendes optisches Tracking-System (ART) der Firma Advanced Realtime Tracking GmbH verwendet. Ein Monitor, der über einen Spiegel umgelenkt und in der Combinerscheibe reflektiert wird, zeigt den virtuellen Inhalt an. Um einen kontaktanalogen Eindruck zu erreichen, muss die Blickrichtung des Fahrers getrackt werden. Zur Detektion der Blickrichtung wurde ein Gazetracking-System aufgebaut (siehe hierzu Abschnitt 8.4.4).

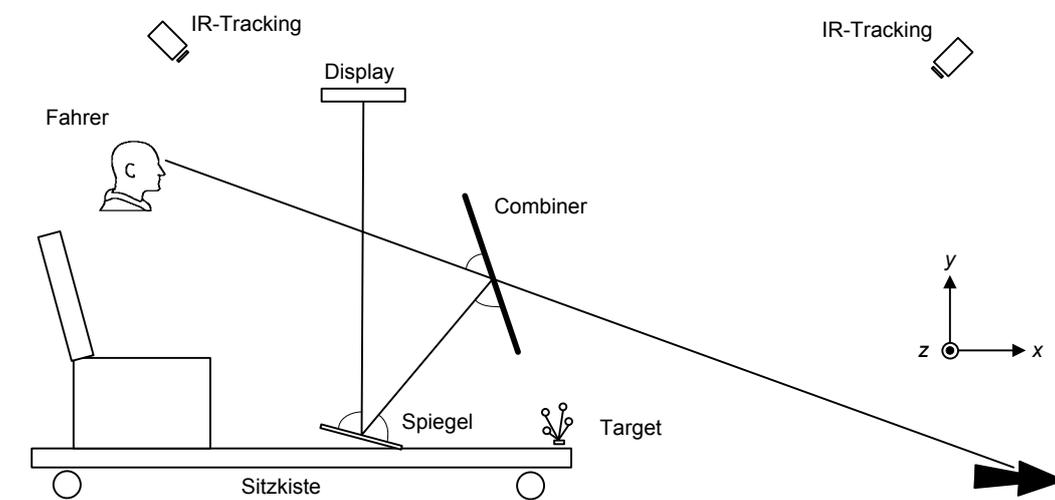


Abbildung 8.21: Schematischer Aufbau der verwendeten Sitzkiste

Bei einem kontaktanalogen HUD entfällt das Problem des HUD-Fokus, da kontaktanaloge Symbole in der Fahrscene integriert sind. Kontaktanaloge HUDs benötigen einen kollimierten Fokus (d. h. eine Ausrichtung auf das optisch Unendliche, in der Regel weiter als ca. 6 m), wodurch eine Akkommodation zwischen HUD-Informationen und Fahrscene nicht erfolgen muss (siehe [WE92]).

### Kalibrierung

In einer AR-Umgebung gibt es sowohl reale Objekte in der Umgebung des Nutzers als auch virtuelle Objekte, die nur in einem computerinternen Modell der Welt existieren. Um beim Benutzer das Gefühl hervorzurufen, er befände sich in einer komplett realen Welt, muss das interne Modell so exakt wie möglich mit der physischen Umgebung in Übereinstimmung gebracht werden. Dazu müssen die Parameter, die das interne mathematische Modell beschreiben, entsprechend so bestimmt werden, dass eine gute Übereinstimmung gegeben ist. Die Bestimmung dieser Parameter nennt man *Kalibrierung*.

Die in dieser Arbeit zu bestimmenden Parameter sind die Einträge der nachfolgend hergeleiteten *Projektionsmatrix*  $P$  des Modells der endlichen projektiven Kamera [PAR<sup>+</sup>08]. Dieses

Kameramodell besteht in der realen Welt aus dem Projektionszentrum und der in die Frontscheibe eingespiegelten Anzeigefläche (Bildebene). Dieses optische System wird im Folgenden als virtuelle Kamera bezeichnet. Die Kenntnis der Projektionsmatrix  $P$  ermöglicht es, Objekte aus der realen Welt, deren Positionen in einem definierten Weltkoordinatensystem bekannt sind, mit der virtuellen Darstellung des Objektes zu überlagern.

Die Projektionsmatrix  $P$  wird aus einer Anzahl gegebener Korrespondenzen zwischen 3D-Welt und dem 2D-Bild berechnet. Die einfachste Art solcher Korrespondenzen stellen 3D-Punkte und ihre dazugehörigen, durch die Zentralprojektion entstehenden 2D-Bildkoordinaten dar. Da bei dem Berechnungsverfahren normalerweise überbestimmte lineare Gleichungssysteme auftreten, kann  $P$  nicht exakt bestimmt werden, es handelt sich also lediglich um eine Näherung, die möglichst nahe an den echten Werten liegen soll. Dabei wird im Folgenden immer davon ausgegangen, dass die Abbildung von 3D nach 2D linear ist, also keine Linsenverzerrung auftritt. Nun wird ein Algorithmus vorgestellt, mit dem die gesuchte Kameramatrix  $P$  berechnet werden kann.

**Grundlegende Gleichungen:** Es wird davon ausgegangen, dass eine bestimmte Anzahl  $n$  an Punktkorrespondenzen  $\vec{X}_i \leftrightarrow \vec{x}_i$  zwischen 3D-Weltpunkten und 2D-Bildpunkten gegeben ist. Gesucht ist eine  $3 \times 4$  Kameramatrix  $P$ , so dass  $\vec{x}_i = P\vec{X}_i$  für  $1 \leq i \leq n$  gilt. Dabei ist zu bedenken, dass die Gleichung homogene Vektoren enthält und somit  $\vec{x}_i$  und  $P\vec{X}_i$  nicht identisch sind, sondern lediglich die gleiche Richtung haben. Sie können sich aber durch einen von 0 verschiedenen Skalierungsfaktor unterscheiden. Da zwei Vektoren, die dieselbe Richtung haben, im Kreuzprodukt den Nullvektor ergeben, kann die Gleichung auch folgendermaßen geschrieben werden:

$$\vec{x}_i \times P\vec{X}_i = \vec{0} \quad (8.4)$$

Diese Art der Darstellung ermöglicht es, eine einfache lineare Lösung für  $P$  herzuleiten. Wenn nun die  $j$ -te Zeile der Matrix  $P$  als  $\vec{P}^{jT}$  bezeichnet wird, ergibt sich für  $P\vec{X}_i$

$$P\vec{X}_i = \begin{pmatrix} \vec{P}^{1T} \vec{X}_i \\ \vec{P}^{2T} \vec{X}_i \\ \vec{P}^{3T} \vec{X}_i \end{pmatrix} \quad (8.5)$$

Mit  $x_i = (x_i, y_i, w_i)^T$ , kann das Kreuzprodukt aus 8.4 als

$$\vec{x}_i \times P\vec{X}_i = \begin{pmatrix} x_i \\ y_i \\ w_i \end{pmatrix} \times \begin{pmatrix} \vec{P}^{1T} \vec{X}_i \\ \vec{P}^{2T} \vec{X}_i \\ \vec{P}^{3T} \vec{X}_i \end{pmatrix} = \begin{pmatrix} y_i \vec{P}^{3T} \vec{X}_i - w_i \vec{P}^{2T} \vec{X}_i \\ w_i \vec{P}^{1T} \vec{X}_i - x_i \vec{P}^{3T} \vec{X}_i \\ x_i \vec{P}^{2T} \vec{X}_i - y_i \vec{P}^{1T} \vec{X}_i \end{pmatrix} \quad (8.6)$$

geschrieben werden. Da  $\vec{P}^{jT} \vec{X}_i = \vec{x}_i^T \vec{P}^j$  gilt, ergibt sich für jedes Punktepaar  $\vec{X}_i \leftrightarrow \vec{x}_i$  mit  $1 \leq i \leq n$ , ein Satz von drei Gleichungen in den Einträgen von  $P$ , der folgendermaßen dargestellt werden kann:

$$\begin{bmatrix} \vec{0}^T & -w_i \vec{X}_i^T & -y_i \vec{X}_i^T \\ w_i \vec{X}_i^T & \vec{0}^T & -x_i \vec{X}_i^T \\ y_i \vec{X}_i^T & x_i \vec{X}_i^T & \vec{0}^T \end{bmatrix} \begin{pmatrix} \vec{P}^1 \\ \vec{P}^2 \\ \vec{P}^3 \end{pmatrix} = \vec{0} \quad (8.7)$$

Da diese drei Gleichungen linear abhängig sind, kann das Gleichungssystem auch auf zwei linear unabhängige Zeilen reduziert werden:

$$\begin{bmatrix} \vec{0}^T & -w_i \vec{X}_i^T & -y_i \vec{X}_i^T \\ w_i \vec{X}_i^T & \vec{0}^T & -x_i \vec{X}_i^T \end{bmatrix} \begin{pmatrix} \vec{P}^1 \\ \vec{P}^2 \\ \vec{P}^3 \end{pmatrix} = \vec{0} \quad (8.8)$$

Abgekürzt geschrieben haben diese Gleichungen die Form  $A_i \vec{p} = \vec{0}$ , wobei  $A_i$  eine  $2 \times 12$  Matrix ist und  $\vec{p}$  der 12-Vektor mit den Einträgen aus  $P$ . Aus den gegebenen  $n$  Punktenpaaren erhält man durch „Stapeln“ der jeweiligen Gleichungen 8.8 eine  $2n \times 12$  Matrix  $A$ . Die Projektionsmatrix  $P$  wird dann durch Lösen des Gleichungssystems  $A_i \vec{p} = \vec{0}$  bestimmt, wobei die triviale Lösung  $\vec{p} = \vec{0}$  hier nicht von Interesse ist.

#### 8.4.4 Gaze- und Eyetracking

Im Rahmen dieser Arbeit wurde ein eigenes Gaze- und Eyetracking System entwickelt, das u. a. für den prototypischen Sitzkistenaufbau des kontaktanalogen HUDs zur Anwendung kommt. Darüber hinaus sind zahlreiche Anwendungen hinsichtlich Gaze Control, wie die Blickbedienung von Displays, denkbar. In Abbildung 8.22 ist die Anordnung des entwickelten remote Gazetrackers im Fahrzeug zu sehen. Zunächst wird eine Einführung in das Thema der Augen- und Blickdetektion gegeben und es werden verschiedene Methoden vorgestellt. Danach erfolgt die Beschreibung des verwendeten Gazetracking Ansatzes. Abschließend werden die Bildverarbeitungsmethoden zur Augendetektion beschrieben.

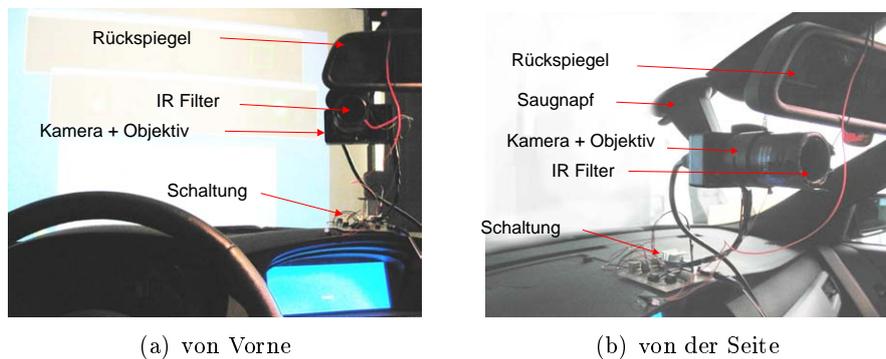


Abbildung 8.22: Integration des remote Gazetrackers im Simulator

#### Einführung

Eyetracking hat seine Wurzeln in der Augenbewegungsforschung, die sich bereits seit Ende des 19. Jahrhunderts mit der Messung der Bewegungen des menschlichen Auges und der damit verbundenen kognitionswissenschaftlichen Zusammenhänge beschäftigt [SE08]. Dazu wurden diverse Ansätze entwickelt, die sich in der Messgenauigkeit und vor allem der Beeinträchtigung des Benutzers unterscheiden lassen. Neben der Verwendung von Kontaktlinsen, in die kleine Spiegel oder Spulen verarbeitet wurden (Search Coil Methode) oder Elektroden, die neben den Augen angebracht werden (Elektrookulogramm (EOG)), haben sich insbesondere videobasierte Systeme durchgesetzt (Videookulogramm (VOG)). Eine genaue Zusammenstellung aller nennenswerten Methoden inklusive ihrer örtlichen und zeitlichen Genauigkeit, Kosten und Beeinträchtigung des Probanden, kann der Fachliteratur entnommen werden

[DPU08, JRV08, SE08]. Da videobasierte Eyetracking Systeme die geringste Beeinträchtigung vorweisen und sich wesentlich vielseitiger einsetzen lassen, wird im Rahmen dieser Arbeit ausschließlich diese Technologie weiter betrachtet. Des Weiteren wird in diesem Zusammenhang oft der verwandte Begriff Gazetracking genannt. Beide Begriffe sollen im Folgenden von einem systemtheoretischen Standpunkt her erklärt werden.

Ein *Eyetracker* lässt sich als ein videobasiertes System definieren, welches das Auge detektieren, verfolgen und Merkmale extrahieren kann. Diese Merkmale können sowohl in Sensor- als auch in Weltkoordinaten berechnet werden. Basierend auf diesen Merkmalen und eventuell noch weiteren benutzerspezifischen Daten berechnet ein Gazetracker den Blick des Benutzers. Hierbei werden 2 grundsätzlich verschiedene Typen von Gazetrackern unterschieden, 2D- und 3D-Gazetracker. Während 2D-Gazetracker den Blickpunkt (engl. Point of Gaze (POG)) auf einem Bildschirm ausgeben, liefern 3D-Systeme eine Blickgerade (engl. Line of Gaze (LOG)) in Weltkoordinaten.

Bei einem remote Eyetracker kann der Benutzer seinen Kopf völlig frei im Sichtfeld einer kontaktlos angebrachten Kamera bewegen. Dies bedeutet, dass keine Einschränkungen bezüglich der relativen Lage von Sensor- und Augenkoordinatensystemen gesetzt werden. Für viele Anwendungen in denen Komfort von Bedeutung ist, stellt dies die einzig mögliche Variante dar. Der maximal mögliche Komfort wird jedoch durch einige Nachteile erkauft. So führt die größere Entfernung zu einer geringeren örtlichen Auflösung. Je nach gewünschter Messgenauigkeit und Bewegungsfreiheit müssen die Pixeldichte, Sensorfläche und Brennweite der verwendeten Kamera bzw. Optik aufeinander abgestimmt werden. Generell steigt der Aufwand in der Algorithmen. Erwähnenswert sind noch Systeme, die zwar als remote Eyetracker verkauft werden, aber trotzdem auf technische Hilfen wie Marker oder Stirnbänder angewiesen sind. So wird beispielsweise beim Eyelink Remote der Firma SR Research [Eye08] ein kontrastreicher Marker mit konzentrischen Kreisen an die Stirn des Probanden platziert. Damit kann auch bei Lidschlag der Kopf verfolgt und dessen Abstand zur Kamera berechnet werden.

## Merkmale

Die Extraktion von Merkmalen ist eine der Hauptaufgaben eines Eyetracking Systems. Häufig handelt es sich dabei um Sensorkoordinaten von wichtigen Punkten wie dem Pupillenmittelpunkt oder einer LED-Reflexion auf der Hornhaut (Kornea) in der Einheit [Pixel]. Es können aber auch Merkmale in Weltkoordinaten (in [mm]) bestimmt werden. Hierfür muss jedoch eine Verbindung zwischen dem 2D-Bild und dem 3D-Weltkoordinatensystem hergestellt werden.

Dazu existieren diverse Techniken, welche sich in der Anzahl der Kameras und der verwendeten Optik unterscheiden: Ohno et al [OMY02] verwenden ein Autofokus-Objektiv, dessen Fokusweite und damit auch die Entfernung des Auges sich per Software abfragen lässt. Ebenfalls mit einer Kamera, aber mit 2 LEDs ausgestattet, arbeitet das System von Hennessey et al [HNL06]. 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 et al [WS01] kommt ganz ohne LEDs aus, in dem 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 kann somit eine Beziehung zwischen den beiden Koordinatensystemen hergestellt werden. Schließlich sind noch diverse Systeme [TL99, BF03] zu nennen, die mit *Stereo Kameras* arbeiten. Sie berechnen die Weltko-

ordinaten von Augenmerkmalen wie dem Pupillenmittelpunkt oder -rand über Triangulation oder homographische Abbildung. Ist mit Hilfe einer dieser Techniken eine Verbindung zu Weltkoordinaten hergestellt, so lassen sich unter Verwendung eines dreidimensionalen Augenmodells selbst im Videobild nicht ersichtliche Punkte in Weltkoordinaten lokalisieren, wie z. B. der Korneamittelpunkt im System von Hennessey et al.

### Blickdetektion

Die Aufgabe einer Blickrichtungserkennung besteht darin, den Blick des Probanden aus den extrahierten Bildmerkmalen und eventuell zusätzlichen, in einer Kalibrierung gewonnenen Daten, zu berechnen. Systeme der Blickrichtungserkennung, auch Gazetracker genannt, lassen sich grundsätzlich in 2D- und 3D-Gazetracker unterteilen. Während bei 2D-Gazetrackern lediglich Merkmale in Sensorkoordinaten berücksichtigt werden, stellt die Umrechnung der Merkmale von Sensor- in Weltkoordinaten einen wesentlichen Bestandteil eines 3D-Gazetrackers dar. Folglich können 2D-Gazetracker lediglich einen 2-dimensionalen Blickpunkt bestimmen, 3D-Gazetracker hingegen eine Blickgerade in Weltkoordinaten. Im Folgenden sollen die unterschiedlichen Ansätze erklärt und ihre Vor- und Nachteile untersucht werden. Im Gegensatz zu 2D-Systemen wird das Auge bei 3D-Gazetrackern explizit modelliert und der Blick in Weltkoordinaten berechnet. Dazu muss zunächst ein Zusammenhang zwischen dem 2D-Sensor und dem 3D-Weltkoordinatensystem hergestellt werden. Wie bereits erwähnt, kann dies z. B. durch ein Autofokus-Objektiv, eine Stereokamera, oder auch durch eine Kamera mit 2 LEDs realisiert werden. Ist dieser Zusammenhang hergestellt, kann die Lage von Merkmalen wie dem Pupillenmittelpunkt oder -rand auch in Weltkoordinaten ermittelt werden. Beim Rand von Iris oder Pupille lässt sich nun der Blick als Normalenvektor auf die Ebene, die durch die Randkurve aufgespannt wird, berechnen [WS01]. Soll der Pupillenmittelpunkt möglichst genau bestimmt werden, so muss auch die Brechung an der Korneaoberfläche berücksichtigt werden. Dazu wird bei diversen Ansätzen [HNL06, OMY02, BF03] ein 3D-Augenmodell verwendet, welches die geometrischen und optischen Eigenschaften des Auges approximiert. Um nicht den Mittelpunkt einer virtuellen Pupille zu erhalten, müssen hier die Lichtstrahlen gemäß der unterschiedlichen Brechindizes von Luft und Kammerwasser gebrochen werden. Unter Verwendung der Daten aus dem Augenmodell und mindestens einer LED, deren Lichtstrahl an der Korneaoberfläche reflektiert und auf die Reflexionspunkte im Bild abgebildet wird, lässt sich auch der Mittelpunkt der Korneakugel in 3D-Weltkoordinaten bestimmen.

### Gazetracking-Ansatz

Im Rahmen dieser Arbeit wird der Gazetracking-Ansatz nach Hennessey [HNL06] verwendet. Im folgenden Abschnitt wird der Ansatz vorgestellt und seine Vor- und Nachteile erläutert. Es werden eine Kamera und 2 LEDs verwendet, wobei die Kamera kalibriert und die Position der LEDs im 3D-Kamerakoordinatensystem bekannt sein muss. Die LEDs verursachen Reflexionen auf der Oberfläche der Kornea, welche auf den Sensor abgebildet werden. Abbildung 8.23 zeigt diesen Sachverhalt für ein Auge, welches sich vor dem Monitor befindet. Zur Rechten sind die beiden LEDs an einer bekannten Position angebracht. Die Visualisierung zeigt den Verlauf des einfallenden Lichtstrahls von der LED bis zur Kornea und des reflektierten Strahls von der Kornea bis zum Abbild der Reflexion auf dem Sensor, welcher sich mittig unterhalb des Monitors befindet.

Für jede LED und deren Reflexion lässt sich gemäß Abbildung 8.23 ein Hilfskoordinatensystem

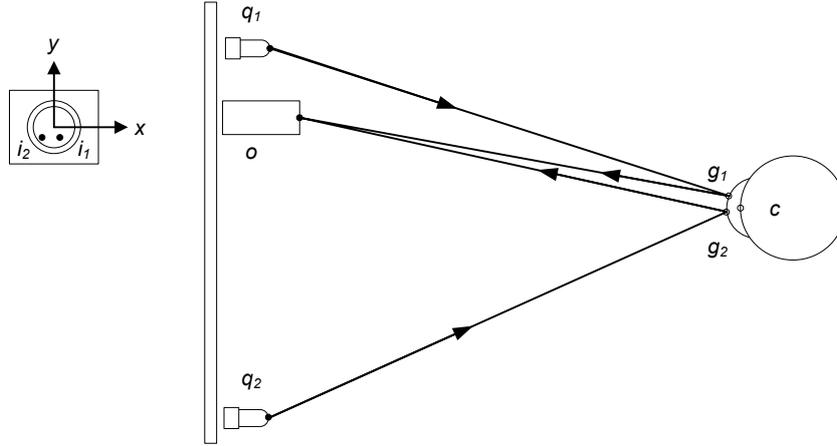


Abbildung 8.23: Systemaufbau Hennessey [HNL06]

aufstellen, wobei

- die  $X$ -Achse vom Ursprung (Zentrum der kalibrierten Kamera) zur LED zeigt
- die  $Z$ -Achse so gewählt wird, dass der Vektor vom Ursprung zum Abbild der Reflexion in der  $X - Z$  Ebene liegt
- die  $Y$ -Achse so gewählt wird, dass ein Rechtssystem entsteht

Im Hilfskoordinatensystem für die  $i$ -te LED lässt sich der Korneamittelpunkt  $(\hat{c}_{ix}, \hat{c}_{iy}, \hat{c}_{iz})$  wie folgt angeben:

$$\begin{bmatrix} \hat{c}_{iz} \\ \hat{c}_{iy} \\ \hat{c}_{ix} \end{bmatrix} = \begin{bmatrix} \hat{g}_{iz} - r \cdot \sin\left(\frac{\hat{\alpha}_i - \hat{\beta}_i}{2}\right) \\ 0 \\ \hat{g}_{iz} \cdot \tan(\hat{\alpha}_i) - \cos\left(\frac{\hat{\alpha}_i - \hat{\beta}_i}{2}\right) \end{bmatrix} \quad (8.9)$$

wobei die Winkel  $\hat{\alpha}_i$  und  $\hat{\beta}_i$  von der Unbekannten  $\hat{g}_{ix}$  und weiteren, bekannten geometrischen Größen abhängen:

$$\hat{\alpha}_i = \cos^{-1}\left(\frac{-\hat{I}_i \cdot \hat{Q}_i}{\|\hat{I}_i\| \cdot \|\hat{Q}_i\|}\right) \quad \hat{\beta}_i = \tan^{-1}\left(\frac{-\hat{g}_{iz} \cdot \tan(\hat{\alpha}_i)}{\hat{l}_i \cdot \hat{g}_{iz}}\right) \quad (8.10)$$

über eine Matrix  $R$  kann der Hilfs-Korneamittelpunkt  $\hat{c}_i$  zurück ins Weltkoordinatensystem transformiert werden

$$c_i = R_i^{-1} \hat{c}_i \quad (8.11)$$

und man erhält 3 Gleichungen für 4 Unbekannte  $(c_{ix}, c_{iy}, c_{iz}, \hat{g}_{ix})$ .

Die Matrix  $R$  muss hierbei für jedes Bild neu aufgestellt werden, da sie den Zusammenhang zwischen Hilfs- und Weltkoordinatensystem beschreibt und daher von der Position der LED-Reflexion auf dem Sensor abhängt. Führt man obige Überlegungen analog für die zweite LED durch, so ergeben sich 6 Gleichungen für 8 Unbekannte. Da jedoch der Korneamittelpunkt in Weltkoordinaten für beide LEDs identisch ist,

$$c_1 = c_2 \quad (8.12)$$



zwischen Pupillen- und Korneamittelpunkt  $r_d$ , so lässt sich  $r_{ps}$  berechnen:

$$r_{ps} = \sqrt{r_2^d + r_2^p} \quad (8.13)$$

Durch die Mittelung zweier gegenüberliegender Randpunkte lässt sich die Lage des Pupillenmittelpunktes abschätzen. Die Verwendung mehrerer Paare von gegenüberliegenden Randpunkten ermöglicht eine robustere Schätzung. Die optische Achse des Auges kann nun als Verbindungsgerade zwischen dem Pupillen- und Korneamittelpunkt angegeben werden. Der Fehler, der sich aus der Abweichung zur visuellen Achse ergibt, kann nach den beiden vorgestellten Methoden minimiert werden. Die resultierende Genauigkeit der Blickrichtungsdetektion liegt laut Autor bei  $0,4 - 0,9^\circ$ .

#### 8.4.5 Anforderungen quasi-kontaktanaloger Anzeigen im Realfahrzeug

Um den prototypischen Aufbau des quasi-kontaktanalogen HUDs in ein Realfahrzeug übertragen zu können, müssen noch folgende Aspekte berücksichtigt werden:

- **Positionierung und Visualisierung:**  
Zunächst wird auf die Auswirkungen des eingeschränkten Sichtbereichs eines HUDs im Fahrzeug und auf die zu erwartenden Messfehler durch die bedingte Positionsungenauigkeit eingegangen.

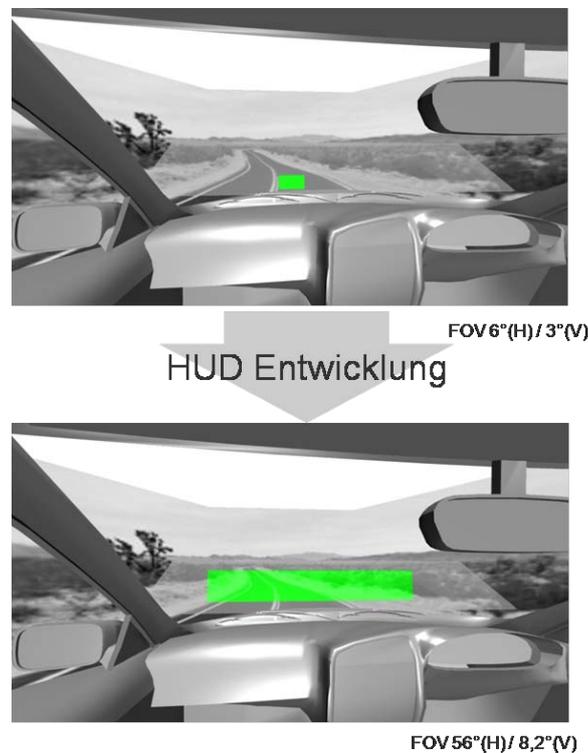


Abbildung 8.26: Herausforderungen HUD Entwicklung

- **Fahrzeugsensorik:**

Für das Tracking des Fahrzeuges wird ein Differential-GPS-System benötigt, das momentan die größtmögliche Genauigkeit aufweist. Über weitere Fahrzeugsensorik wie Gyrosensor, Kameras und Kartenmaterial kann die Genauigkeit des GPS-Systems erhöht werden.

Abschließend lassen sich folgende Aussagen zur Realisierung eines quasi-kontaktanalogen HUDs nennen:

Der prototypische Aufbau hat gezeigt, dass quasi-kontaktanaloge Darstellungen prinzipiell möglich sind. Dies stellt jedoch hohe Anforderungen an die Sensorik und deren Genauigkeit. Der Aufbau eines fahrbaren Prototyps ist vor allem dann sinnvoll, sobald großflächige HUDs technisch möglich und verfügbar sind.

Insgesamt liegen die Herausforderungen der weiteren Forschung an der Untersuchung und Entwicklung geeigneter Konzepte zur Informationsdarstellung, wozu in dieser Arbeit ein Beitrag geleistet wurde. Mit der intensiven Erforschung kontaktanaloger HUDs beschäftigen sich im Besonderen [Lan08, Bub76, TLKB06]. Wie erwähnt, müssen weitere Anstrengungen auf die Entwicklung von HUDs mit großem Field-Of-View gelegt werden, was im Hinblick auf den begrenzten Platz im Fahrzeug problematisch ist. Des Weiteren ist es notwendig, unerwünschte Überlagerungen, z. B. bei einem vorausfahrenden Fahrzeug, zu identifizieren und geeignet herauszufiltern.

## 8.5 HUD Untersuchungen

Nachfolgend werden die Ergebnisse verschiedener Untersuchungen zum Einsatz des HUDs im Fahrzeug präsentiert.

### 8.5.1 Vergleich Serien-HUD mit HDD (Feldstudie I)

In dieser Feldstudie wurde das BMW 5er Serien-HUD hinsichtlich Blickabwendung und Akzeptanz untersucht. Neben dem HUD war das Testfahrzeug mit konventionellem Kombidisplay und Mitteldisplay, auch Head-Down Display (HDD) genannt, ausgestattet. Das HUD zeigte neben der aktuellen Geschwindigkeit auch Navigationshinweise und ACC-Infos. Die gewählte Strecke erforderte eine Fahrzeit von ca. einer Stunde und hatte eine Gesamtlänge von 60 km. Um das HUD in verschiedenen Fahrsituationen zu untersuchen, beinhaltete die Strecke Abschnitte in Stadt, Überland und Autobahn. Die Streckenführung erfolgte über die Navigationshinweise des eingebauten Navigationssystems. Um weitere Blicke auf die Displays zu provozieren, wurden die Versuchspersonen damit beauftragt Geschwindigkeitsbeschränkungen laut vorzulesen und anschließend die Geschwindigkeitsanzeige mit dem vorgegebenen Display (HUD bzw. Kombidisplay) abzugleichen.

Zur Erfassung der Blickabwendungsdauern waren die Versuchspersonen mit dem JANUS Eye-Tracking System (siehe dazu Abschnitt 2.5.4) ausgestattet. Die Blickabwendungsdauer (engl. *Gaze Retention Period*, GRP) ermittelt sich aus der Dauer der beiden Blicksprünge  $BS$  und der Verweildauer  $VD$  im Display (auch sogenannte *Fixation*).

$$GPR = BS + VD; [s] \quad (8.14)$$

Am Ende der Fahrt wurden die Probanden zur Akzeptanz der Displays mit Hilfe eines Fragebogens interviewt. An der Studie nahmen insgesamt 18 Probanden (Durchschnittsalter:

43,9 Jahre, davon 3 weiblich) teil.

### Objektive Ergebnisse

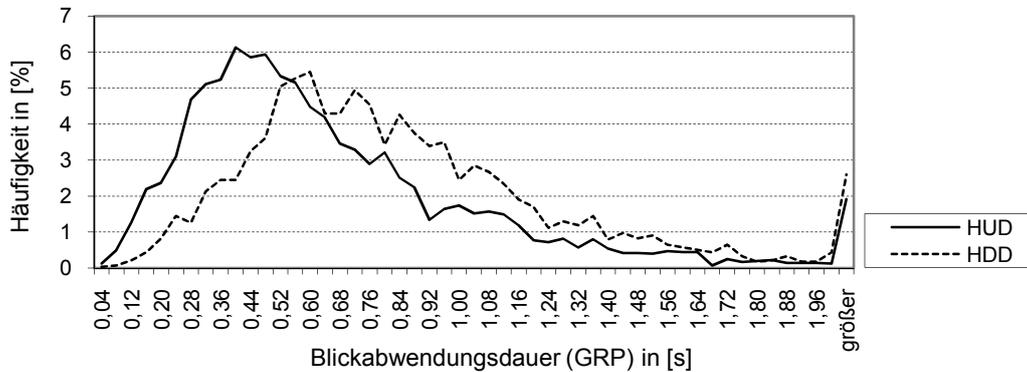


Abbildung 8.27: Häufigkeitsverteilung der Blickabwendungsdauern aller Probanden

Nach der Analyse der Blickabwendungsdauern (GRPs) über die gesamte Fahrstrecke und aller Probanden lieferte das HUD die beste Leistung (siehe dazu Abbildung 8.27<sup>1</sup>). Die Anzahl der Blicke auf das HUD war im Vergleich zum HDD größer, was überwiegend auf den Erstkontakt der Versuchspersonen mit der HUD-Technologie zurückzuführen ist. Im Mittel (Modus) war die Blickabwendungsdauer beim HUD um ca. 220 ms niedriger als beim HDD. Dabei war bei geringen Geschwindigkeiten wie im Stadtverkehr, die mittlere GRP bei der Verwendung des HUDs zwischen 20 % und 22 % geringer als bei der Nutzung des HDD. Bei hohen Geschwindigkeiten wie auf Landstraßen und Autobahnen, wurde die durchschnittliche GRP sogar um bis zu 25 % reduziert. Dies deutet u. a. darauf hin, dass das HUD ein hohes Potenzial für eine effiziente Erfassung von Informationen besonders bei schnellem Verkehr hat. Gerade für die Altersgruppe zwischen 51 und 60 Jahren zeigte sich auch die gute Eignung des HUDs, da hier der Prozess der Informationsaufnahme im Mittel (Modus) sogar um ca. 250 ms schneller war (siehe Abbildung 8.28).

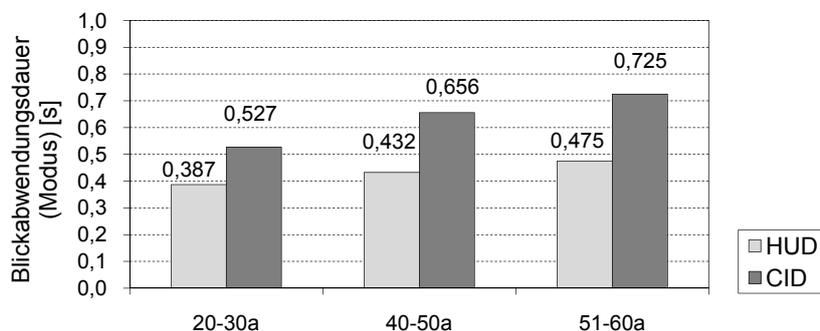


Abbildung 8.28: Mittlere Blickabwendungsdauern im HUD in Abhängigkeit des Alters

<sup>1</sup>Die Verzerrung des rechten Randes (ab 2 s) resultiert aus der Kumulierung von mehreren Bereichen.

### Subjektive Ergebnisse

85 % der Probanden waren der Meinung, dass sie das HUD während der Fahrt wünschen. Das HUD ist demnach eine wichtige Ergänzung zu den bestehenden Anzeigen und liefert unabhängig vom Alter eine gute Unterstützung. Dabei gaben 86 % an, dass die Geschwindigkeit und 90 %, dass die Navigation und die ACC-Symbolik unbedingt im HUD angezeigt werden sollten. Die Testpersonen begründeten die Bevorzugung des HUDs mit der reduzierten Blickabwendung (64 %), einem Komfortgewinn (44 %) und höherer Sicherheit (17 %). Dabei wollten mehr als 70 % der Probanden weitere Informationen und Funktionen im HUD. Neben den Individualisierungsmöglichkeiten wünschten die Fahrer eine Totenwinkelwarnung (86 %), Fahrzeugfehler-Hinweise (85 %), Informationen zur Stauumfahrung (83 %), Warnungen zu Geschwindigkeitsbegrenzungen (75 %), Anzeige der Tankreichweite (75 %), des Sicherheitsabstandes (55 %) und einer Kollisionswarnung (50 %). Darüber hinaus wurden Statusinformationen wie Außentemperatur, aktueller Radiosender, die Uhrzeit und die Anzeige der Gangwahl genannt. Nach Meinung der Versuchspersonen kann das HUD jedoch nicht die üblichen Displays des HDD (wie Kombi und CID) vollständig ersetzen. Insgesamt 50 % aller Versuchspersonen würden auf das Kombi und 40 % auf das CID verzichten.

### Schlussfolgerungen

Zusammenfassend kann aus der Untersuchung gefolgert werden (siehe auch [APW<sup>+</sup>07, AMR05]):

- Blickabwendungsdauern zum Ablesen von Informationen im HUD geringer als bei den HDDs; hohes Potential des HUD für effiziente Informationsaufnahme besonders bei hohen Geschwindigkeiten und älteren Personen.
- Häufigere Blickzuwendung v. a. bei Erstkontakt mit dem HUD.
- Sehr hohe subjektive Akzeptanz des HUDs.
- Weitere Informationen im HUD werden gewünscht.

#### 8.5.2 Untersuchung von HUD-Inhalten (Feldstudie II)

Um das HUD in Bezug auf Inhalte und Gestaltung zu untersuchen, wurde eine weitere Feldstudie konzipiert. In dieser Studie kam ein Laser HUD (monochrom) zum Einsatz, das das virtuelle Bild jedoch nur in ca. 1 m Entfernung zum Fahrer projizierte. Hierzu wurden verschiedene Layouts mit mehreren Symbolen entworfen. An diesem Test nahmen 14 Probanden teil. Die Studie erfolgte auf einem Testgelände. Dabei musste einem vorausfahrenden Fahrzeug gefolgt werden. Die Reaktionen der Probanden wurde über eine Videokamera aufgezeichnet. Ziel der Untersuchungen war es, die verschiedenen Layouts in Abhängigkeit des situativen Kontextes (z. B. Kurven im Gegensatz zu geraden Straßen) zu beurteilen.

### Ergebnisse

Im Hinblick auf die Informationsmenge wurde festgestellt, dass eine steigende Anzahl an Symbolen im HUD auch zu größeren Erfassungszeiten führte. Eine noch akzeptierte Informationsmenge der Versuchspersonen waren in etwa bis zu vier unbekannte Symbole. Für die kognitive Aufnahme der vier Symbole bzw. Junks (Informationseinheiten) benötigten die Probanden

im Durchschnitt insgesamt 4,9 s, in kurvenreichen Situationen sogar bis zu 5,8 s. Dies deutet darauf hin, dass die Informationsmenge im HUD in einem angemessenen Rahmen bleiben sollte.

### Schlussfolgerungen

Insgesamt kann gefolgert werden, dass nicht mehr als vier unbekannte Informationseinheiten auf einmal im HUD präsentiert werden sollten. Weitere Ergebnisse finden sich in [APW<sup>+</sup>07, AMR05].

#### 8.5.3 Vergleich Menüinteraktion HUD vs. CID (Simulatorstudie Ia)

Schwerpunkt dieser Untersuchung war die Bewertung und der Vergleich von Interaktionen unter Verwendung der Anzeigeorte HUD und CID. Hierzu wurde das entworfene HUD\_MAX-Konzept aus Abschnitt 8.2 herangezogen. Insgesamt wurden sechs Aufgaben entworfen, die drei unterschiedlichen Schwierigkeitsstufen zugeordnet werden konnten. Der Schwierigkeitsgrad wurde über den Quotienten aus Bearbeitungsdauer und Anzahl der Bedienschritte ermittelt. Dieser Quotient wurde als ein Maß zur Bewertung der kognitiven Beanspruchung herangezogen. Als Aufgabentypen handelte es sich um Regelaufgaben, einfache Listenauswahlen und Menübedienung von mehreren Hierarchieebenen. Bei den *Regelaufgaben* mussten die Versuchspersonen ins Klimamenu wechseln und einen spezifischen Temperaturwert einstellen. Des Weiteren musste im Entertainment-Menu eine neue Radiofrequenz gewählt werden. Bei den Aufgaben mit Listenauswahl wählten die Versuchspersonen Musiktitel aus einer MP3-Sammlung des Entertainment-Menus. Bei den schwierigen Aufgaben musste auch in der Liste gescrollt werden. Die Aufgaben mit der höchsten Belastung waren die Navigationsaufgaben, bei denen mehrere Hierarchieebenen durchlaufen werden mussten. Insgesamt nahmen 28 Versuchspersonen an der Untersuchung teil. Das Durchschnittsalter betrug 41,25 Jahre. Die Untersuchung fand im ehemaligen Fahrsimulator des Lehrstuhls für Mensch-Maschine-Kommunikation statt. Als Fahraufgabe diente der LCT-Test (siehe Abschnitt 2.5.1). Die HUD Inhalte wurden mit einem weiteren Beamer auf die Fahraufgabe projiziert (siehe Abbildung 8.33).

### Objektive Ergebnisse

Zunächst werden die gemessenen Ergebnisse erläutert.

**Bearbeitungsdauer:** Die Bearbeitungsdauer (auch Total Task Time, TTT), d. h. die Zeit vom Beginn der Bedienung, nach dem Hören der Anweisung, bis zur Lösung der Aufgabe, wurde aufgezeichnet und ausgewertet. Die mittlere Bearbeitungsdauer stellt die Summe der Bearbeitungsdauern der sechs Aufgaben gemittelt über alle 28 Versuchspersonen im HUD bzw. CID dar. Der Zeitvorteil betrug im Durchschnitt 10 s. Das Ergebnis ist mit  $p < 0,02$  signifikant und bestätigt, dass die Aufgaben im HUD schneller gelöst wurden.

Ein tendenzieller Zeitvorteil im HUD wurde in allen drei Aufgabenkategorien sichtbar (siehe Abbildung 8.29). Der größte Vorteil zeigte sich bei kognitiv schwierigen Aufgaben und betrug hier 4 s. Bei Aufgaben mittleren und schweren Schwierigkeitsgrades war ein signifikanter Unterschied vorhanden ( $p < 0,046$  und  $p < 0,011$ ). Auch ein Einfluss des Alters war erkennbar. Betrachtet man die jüngeren Testpersonen (unter 35 Jahre) zeigte sich nur bei schwierigen Aufgaben ein klarer mittlerer Zeitvorteil von ca. 4 s im HUD. Bei Testpersonen über 35 Jahre hingegen war ein Zeitvorteil über alle Aufgaben sichtbar. Am Stärksten zeigte sich dieser bei

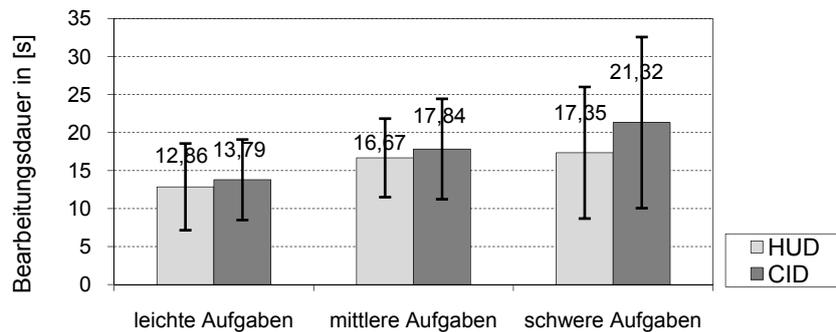


Abbildung 8.29: Mittlere Bearbeitungsdauer in Abhängigkeit vom Aufgabentyp

den Versuchspersonen über 50 Jahre, wo sich ein mittlerer Zeitvorteil von ca. 5 s bei schwierigen Aufgaben abzeichnete.

**Bedienschritte:** Jede Aufgabe konnte mit einer minimalen Anzahl an Bedienschritten gelöst werden. Diese wurde mit der tatsächlich benötigten Anzahl an Bedienschritten der Testpersonen verglichen. Bei der Mittelung der benötigten Bedienschritte zeigten sich tendenzielle Vorteile bei Verwendung des HUDs, diese sind jedoch nicht signifikant. Schlüsselte man die benötigten Schritte in Abhängigkeit von den Schwierigkeitsstufen auf, lassen sich bei schweren Aufgaben signifikante Unterschiede feststellen ( $p < 0,045$ ). Hier haben alle Altersgruppen im Vergleich zum HUD im Mittel 4 bis 5 Bedienschritte mehr im CID durchführen müssen.

**Spurabweichung:** Bei der Analyse der Leistung bei der Fahraufgabe wurde die mittlere Spurabweichung herangezogen. Hier zeigten sich leichte Unregelmäßigkeiten, jedoch konnte hier keine einheitliche Tendenz in Abhängigkeit des Anzeigeortes (siehe Abbildung 8.30) abgeleitet werden. Jedoch zeigte sich ein Einfluss der Übung und Vorerfahrung auf die Spurabweichung bei der Bedienung des CIDs. Erfolgte zuerst die Bedienung im CID, zeigte sich tendenziell eine Verschlechterung der mittleren Spurhaltung gegenüber der Bedienung erst im zweiten Durchlauf. Der Einfluss der Vorerfahrung auf die Spurabweichung zeigte sich aber nicht bei der Bedienung im HUD.

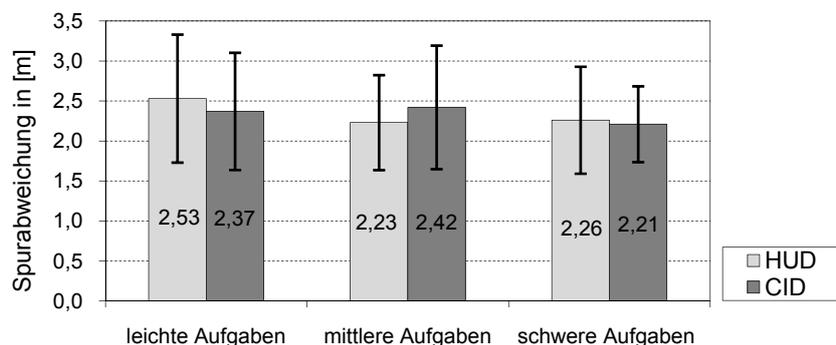


Abbildung 8.30: Mittlere Spurabweichung in Abhängigkeit des Aufgabentyps

**Blickverhalten:** Das Blickverhalten von sieben Versuchspersonen wurde mit dem remote Gazetracking (aus Abschnitt 8.4.4) aufgezeichnet. Dazu erfolgte die Erfassung und Auswertung

der Blickhäufigkeit und Blickdauern. Bei den kognitiv schweren Navigationsaufgaben schauten die Fahrer im HUD seltener auf die Anzeige als im CID. Bei den Aufgaben mit einer Listenauswahl war die Blickanzahl auf das HUD ebenfalls geringer als im CID. Bei den Regelaufgaben blicken die Versuchspersonen häufiger auf das HUD. Daraus könnte man schließen, dass das HUD für mittlere und schwierigere Aufgaben besser geeignet ist, dagegen das CID für leichte Regelaufgaben besser abschneidet. Diese Tendenz war auch bei den Bearbeitungsauern und der Anzahl der Bedienschritte feststellbar. Jedoch weist die durchschnittliche Blickdauer auf das HUD und CID eine andere Tendenz auf. Die durchschnittliche Blickdauer fällt im HUD fast bei allen Aufgaben länger aus. Die selteneren und längeren Blicke auf das HUD führen dazu, dass die Aufgaben in einer signifikant kürzeren Zeit erledigt werden. Am Stärksten zeigt sich die längere Blickdauer im HUD bei Listenauswahlen.

### Subjektive Ergebnisse:

Nun werden die subjektiven Einschätzungen der Probanden präsentiert.

**Anzeigeort:** Für das HUD als sinnvollen Anzeigeort entschieden sich vor der Untersuchung 67,9%; danach waren es sogar 92,4% (siehe Abbildung 8.31). Keine einzige Versuchsperson entschied sich danach gegen das HUD. Da nur drei der Versuchspersonen Vorerfahrung mit dem HUD hatten, kann man die Aussage bestätigen, dass eine große Akzeptanzsteigerung nach dem Erstbenutzer-Kontakt entsteht. Beim CID bleibt die Akzeptanz nach der Benutzung des Displays gleich 32,1%; allerdings stuften anschließend mehr Personen (42,9%) das CID als weniger sinnvollen Anzeigeort ein. Dies deutet darauf hin, dass sie, nachdem sie das HUD kennen gelernt haben, das CID nicht mehr so überzeugend fanden. Den Versuchspersonen wurden offene Fragen, was ihnen am jeweiligen Anzeigeort gefällt und was nicht, gestellt. Bei den expliziten Antworten gaben die Versuchspersonen am häufigsten die geringe Blickabwendung (67,9%) beim HUD an. Die zweithäufigste Antwort (14,3%) war der innovative Anzeigeort.

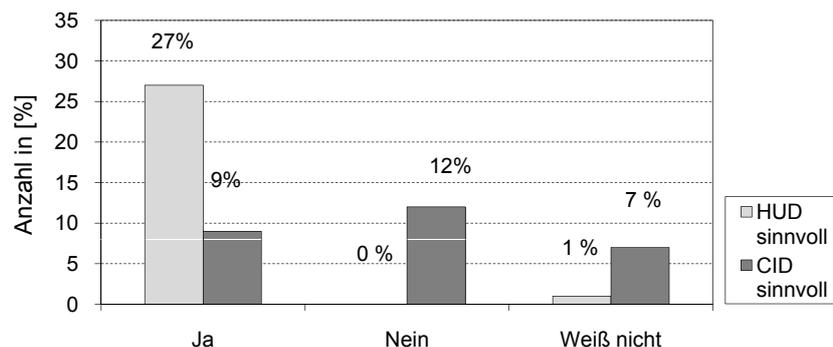


Abbildung 8.31: Sinnvoller Anzeigeort für Menübedienung

**Gefährlichkeit:** Die Einschätzung über die Gefahr der Menübedienung in den zwei Anzeigeorten wurde mit der Frage „Halten Sie eine Menübedienung im HUD/CID für gefährlich?“ ermittelt (siehe Abbildung 8.32). Das CID wird von deutlich mehr Personen (78,6%) als gefährlich eingestuft. Die Versuchspersonen scheinen wenig Gefahrenbewusstsein für das HUD zu haben und empfanden es mehrheitlich (53,6%) als nicht gefährlich, Aufgaben während der Fahrt zu erledigen. Nur 25% stuften die Bearbeitung von Aufgaben während der Fahrt als

gefährlich ein. Diese Daten deuten darauf hin, dass man bei der Vermarktung des Produktes HUD auch Aufklärung betreiben sollte. 64,4% der Fahrer stufte die Ablenkung durch das CID als stark ein. 35,7% von ihnen empfanden den Einfluss als leicht. Beim HUD zeichnet sich ein anderer Trend ab. Eine Versuchsperson (3,6%) fühlte sich durch das HUD sehr stark gestört. 14,3% kreuzten stark an. Die große Mehrheit (69,9%) stufte die Ablenkung als leicht ein und 10,7% stufte sie als gar nicht beeinflussend ein.

**Sichtbehinderung:** Es wurde die Frage gestellt, ob die Versuchspersonen die Anzeigen im großen HUD als eine Sichtbehinderung empfanden. Diese Frage beantworteten 85,7% der Fahrer mit „Nein“. Dies zeigt, dass sich der Großteil der Fahrer nicht durch das HUD gestört fühlte.

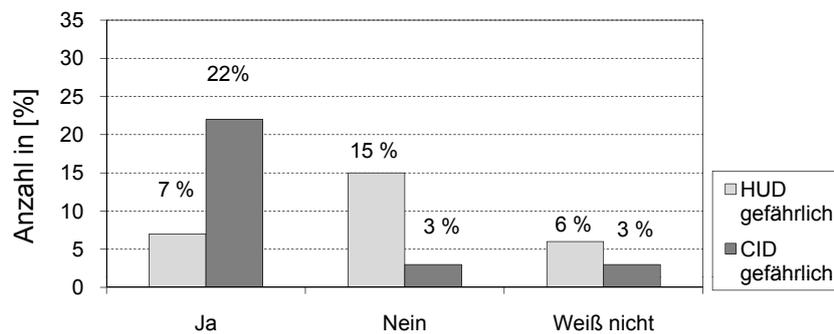


Abbildung 8.32: Einschätzung der Menübedienung im Anzeigort als gefährlich

### Schlussfolgerungen

Aus den Untersuchungen können folgende Schlussfolgerungen in Form von Empfehlungen gegeben werden (siehe auch [MAB08]):

- Das HUD bringt einen Zeitvorteil bei der Interaktion:**  
Das HUD brachte gegenüber dem CID einen signifikanten Zeitvorteil bei der Lösung der Interaktionsaufgaben. Vor allem bei kognitiv belastenden Aufgaben zeigte sich dieser Unterschied am deutlichsten. Beim CID spielte der Lerneffekt eine größere Rolle als beim HUD. Demnach scheint das HUD auch ohne Vorerfahrung intuitiver zu bedienen zu sein. Bei kognitiv schweren Aufgaben brauchten die Versuchspersonen im HUD signifikant weniger Bedienschritte bis zur Lösung der Aufgaben. Dies deutet auf einen geringeren Aufwand und eine geringere Anzahl von Fehlern hin.
- Das HUD wirkt sich nicht negativ auf die Spurhaltung aus:**  
In der Spurhaltung konnte kein signifikanter Unterschied zwischen den beiden Anzeigorten festgestellt werden.
- Die Blicke auf das HUD bei Bedienung sind seltener, aber länger:**  
Bei schweren Aufgaben schauten die Fahrer seltener, aber dafür länger auf das HUD. Dies führte zur schnelleren Bearbeitung in diesem Display.
- HUD gegenüber CID als Bedienort bevorzugt:**  
Subjektiv bevorzugten fast alle Versuchspersonen das HUD.

### 8.5.4 Untersuchung der Anzeigemenge im HUD (Simulatorstudie Ib)

In diesem Teil des Versuches lag der Schwerpunkt auf einer verträglichen Anzeigemenge im HUD. Es galt nachzuprüfen, wie sehr die Anzahl der eingeblendeten Anzeigen im HUD Einfluss auf die Performance beim Fahren haben.

Den Versuchspersonen wurden vier Screenshots mit großer Anzeigemenge (durchschnittlich 14 Symbolen bzw. Informationseinheiten) im HUD eingeblendet. Vier Screenshots dagegen mit mittlerer Anzeigemenge enthielten durchschnittlich 6,2 Symbole im HUD. In Abbildung 8.33 ist der Unterschied dargestellt. Um zu verhindern, dass die Fahrer das HUD ganz ignorieren, wurden sie immer wieder aufgefordert, Informationen abzulesen. Durch diese Maßnahme wurde die kognitive Beanspruchung erhöht. Jeder Screenshot wurde für ca. 40s eingeblendet. Die Reihenfolge der Screenshots mit mittlerer und großer Anzeigemenge wechselte sich ab. Neben der Spurhaltung wurde der PDT-Test (siehe Abschnitt 2.5.2) verwendet.



Abbildung 8.33: Die zwei verschiedenen Anzeigemengen im HUD

### Objektive Ergebnisse

Nachfolgend werden die objektiven Ergebnisse erläutert.

**Anzahl der erkannten PDT-Punkte:** Ohne Belastung wurden durchschnittlich 0,2 Punkte an der linken Position übersehen, was 3,5 % entspricht. Sonst wurden alle PDT-Punkte erkannt. Diese Daten zeigen auf, dass die Versuchspersonen ohne eine Zweitaufgabe nahezu 100 % aller Reize wahrnahmen. Bei den zwei unterschiedlichen Anzeigemengen „mittel“ und „voll“ konnte bei der Erkennungsrate der PDT-Punkte keine signifikante Unterscheidung getroffen werden. In beiden Fällen ist eine deutliche Zunahme der nicht erkannten Punkte im peripheren Sichtfeld erkennbar. Sie ist leicht höher bei der vollen Anzeigemenge. Diese Unterschiede sind aber nicht signifikant. Die geringe Anzahl der übersehenen Punkte deutet darauf hin, dass die Versuchspersonen kognitiv nicht stark genug gefordert wurden, um relevante Ereignisse zu übersehen. Die angezeigte Menge an Informationen hat keine erhöht messbare kognitive Beanspruchung verursachen können. Mit dem Alter nimmt die Anzahl der übersehenen Punkte deutlich zu.

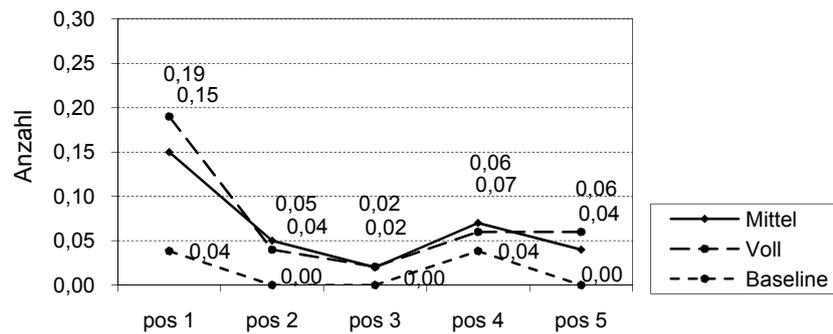


Abbildung 8.34: Mittlere Anzahl nicht erkannter PDT-Punkte für die jeweiligen Positionen

**Reaktionszeit auf die PDT-Punkte:** Aus der Reaktionszeit auf die PDT-Punkte kann keine Abhängigkeit von der angezeigten Menge der Informationen festgestellt werden (siehe Abbildung 8.35). Die Unterschiede sind nicht signifikant. Der Größte betrug 0,05 Sekunden an der Position 2. Im Vergleich zur Baseline war die Reaktion auf die Punkte bis zu 0,35 Sekunden langsamer. Mit dem Alter steigt auch die Reaktionszeit bei den einzelnen PDT-Punkten. Bei älteren Personen kann man auch einen Anstieg der Reaktionszeit in Abhängigkeit von der Anzeigemenge feststellen. Dieser Anstieg ist aber nicht signifikant.

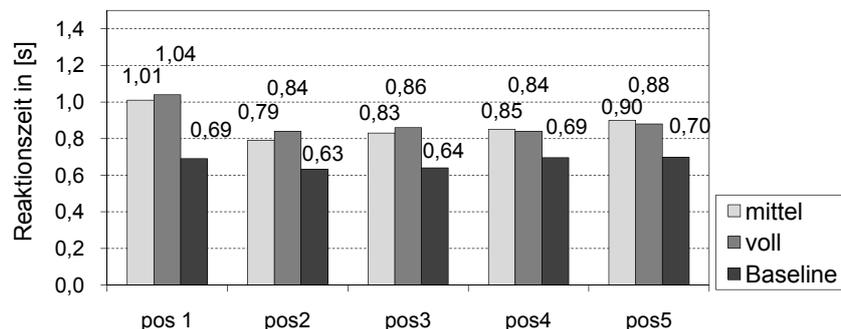


Abbildung 8.35: Mittlere Reaktionszeit auf die PDT-Punkte in Abhängigkeit der Anzeigemenge

**Spurabweichung:** Bei der Spurabweichung, konnte kein signifikanter Unterschied in Abhängigkeit von der angezeigten Informationsmenge festgestellt werden. Es zeigten sich auch keine signifikanten Unterschiede zur Baseline (siehe Abbildung 8.36). Die Aufschlüsselung nach den Altersgruppen weist auch keine signifikanten Unterschiede zwischen den Anzeigemengen auf.

### Schlussfolgerungen

Aufgrund der Ergebnisse können folgende Schlussfolgerungen gezogen werden (siehe auch [MAB08]):

- **Einfluss der Anzeigemenge im HUD mittels PDT, LCT und Blickdaten:**  
Zusammenfassend kann man sagen, dass die erhöhte angezeigte Informationsmenge keinen Einfluss auf die Anzahl übersehener PDT-Punkte, die Reaktionszeit und auch die

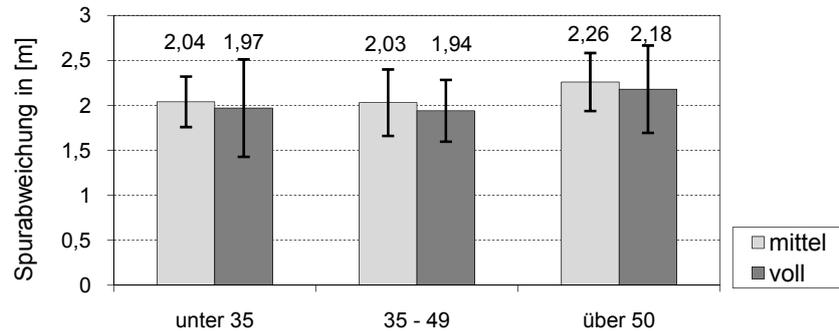


Abbildung 8.36: Mittlere Spurabweichung in Abhängigkeit der Anzeigemenge und Alter

Güte der Spurhaltung hatte. Die Anzahl der nicht erkannten Punkte nimmt mit dem Alter zu. Selbst aber bei über 50-jährigen konnte kein Ansatz zu „Perceptual Tunneling“ festgestellt werden. Damit können die Ergebnisse von Recarte und Nunes bestätigt werden. Die kognitive Beanspruchung zeigt sich in der Verschlechterung der Performance, aber es zeigt sich keine Beeinträchtigung der visuellen Wahrnehmung abhängig vom Sehwinkel. Bei der Reaktionszeit und Spurabweichung hatte auch das Alter keinen signifikanten Einfluss. Bei dieser kognitiven Belastung konnte kein Unterschied zwischen den eingeblendeten Informationsmengen gemessen werden.

### 8.5.5 Vergleich Matrix- vs. Listendarstellung (Simulatorstudie Ic)

In diesem Versuchsteil galt es, die Darstellungsweisen Matrix5 mit fünf Belegungselementen (M5) und Liste5 mit fünf belegten Feldern (L5), genauso wie die Matrix9 mit neun (M9) und Liste9 mit neun (L9) Belegungselementen zu vergleichen. Für jede Darstellung ist ein Beispiel in der Abbildung 8.9 dargestellt.

Die Versuchspersonen hatten als Hauptaufgabe die Mittelspur des LCTs zu halten und PDT-Punkte zu bestätigen. Als Zweitaufgabe mussten sie eine durch den Versuchsleiter vorgegebene Auswahl aus der dargestellten Matrix bzw. Liste treffen. Bei der M9 und L9 mussten sie 18 mal die ihnen vorgelesene Auswahl treffen, d. h. jedes Feld wurde zweimal bestätigt. Bei M5 und L5 wurde jedes Feld dreimal ausgewählt, was in der Gesamtzahl 15 Bedienvorgänge mit sich bringt. Bei der großen Anzahl der Bestätigungsvorgänge ist es möglich, Durchschnittswerte für die Dauer der Auswahl und die Anzahl der benötigten Bedienschritte, die charakteristisch für die jeweilige Darstellungsart sind, zu ermitteln. Diese charakteristischen Werte ermöglichen es, die Vor- und Nachteile der Darstellungsarten zu verdeutlichen. Um die Vergleichbarkeit zu gewährleisten, wurden die Einträge und die Platzierung bei Matrizen und Listen mit gleicher Anzahl an Belegungselementen gleich gewählt. Die PDT sollte die kognitive Belastung der Darstellungsarten vergleichend messen.

#### Objektive Ergebnisse

**Anzahl der erkannten PDT-Punkte:** Die Anzahl der nicht erkannten PDT-Punkte unterscheidet sich nicht signifikant zwischen den Darstellungsweisen bei gleicher Anzahl von belegten Feldern. Auch bei dieser Aufgabe war die kognitive Beanspruchung der Fahrer nicht groß genug, um eine größere Anzahl der PDT-Punkte zu übersehen.

**Reaktionszeit auf PDT-Punkte:** Die Reaktionszeiten auf die PDT-Punkte bei M5 und L5 sind jeweils zur Baseline signifikant. Untereinander ist keine Signifikanz wegen der großen Streuungen feststellbar. Die älteren Versuchspersonen erzielten bessere Ergebnisse bei den Listendarstellungen. Diese favorisierten die Liste auch in der subjektiven Bewertung. Die häufigste Begründung war, dass bei der Liste nur eine Auf- bzw. Abwärtsbewegung zur Auswahl nötig ist und sie dies leichter umsetzen können.

**Auswahldauer:** Die Auswahl aus einer Matrix erfolgt bei beiden Informationsmengen schneller (siehe Abbildung 8.37). Aber nur der Unterschied zwischen M9 und L9 ist signifikant ( $p < 0,038$ ).

Es zeigt sich jedoch, dass die durchschnittliche Schrittdauer bei den Listen jeweils signifikant kürzer ist als bei den Matrizen mit der gleichen Anzahl von belegten Feldern. Dies kann darauf hindeuten, dass die Auswahl aus der Matrix kognitiv belastender ist, denn die Personen haben mehr Zeit für jeden einzelnen Schritt gebraucht, obwohl die räumliche Aufteilung besser ist als bei der Liste. Die Auswahldauer steigt mit zunehmendem Alter.

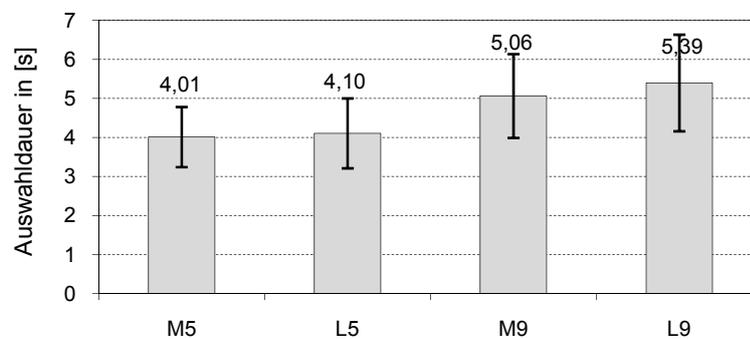


Abbildung 8.37: Mittlere Auswahldauer aus Matrix und Liste

**Anzahl der benötigten Schritte:** In Abbildung 8.38 ist die durchschnittliche Anzahl der Schritte, die bis zur Auswahl benötigt werden, angetragen. Sie ist bei den Matrizen signifikant geringer. Bei der Darstellung mit fünf Elementen ist der Unterschied 1,34 und mit neun Elementen sogar 2,99 Schritte. Diese Daten zeigen einen Vorteil für die Matrixdarstellung. Da es aber bei der Liste nur eine vertikale Bewegung der Blickrichtung und der Finger gibt, die kognitiv nicht so anspruchsvoll ist wie die horizontale und vertikale bei der Matrix, schlägt sich dieser Vorteil nicht so deutlich in der Dauer nieder. Vor allem bevorzugten die älteren Fahrer die Listendarstellung und empfanden sie als kognitiv leichter, obwohl deutlich mehr Bedienschritte bis zur Auswahl des gewünschten Feldes nötig sind.

**Spurabweichung:** Die Werte für die Spurabweichung sind so gering, weil in diesem Versuchsabschnitt nur die Mittelspur gehalten werden musste. Aber auch in der Qualität der Spurhaltung kann man erkennen, dass die Matrix kognitiv belastender ist als die Liste. Die Unterschiede sind fast signifikant, da beim Vergleich von L5 und M5 der Wert  $p < 0,06$  und bei L9 und M9  $p < 0,07$  für die Signifikanz berechnet wurde. Die jungen Versuchspersonen kamen mit der Matrixdarstellung deutlich besser zurecht, als ältere Fahrer. Bei ihnen gibt es zwischen Matrix und Liste keinen signifikanten Unterschied in der Spurhaltegröße. Die älteren Fahrer erzielten bei M5 die schlechtesten Ergebnisse und bei der L5 waren sie genauso gut wie die Jüngeren.

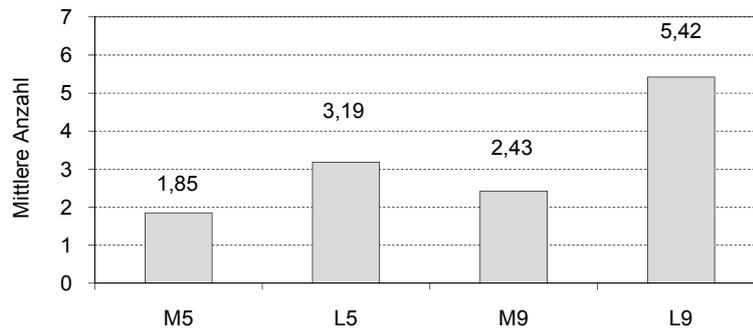


Abbildung 8.38: Mittlere Anzahl der benötigten Schritte für Matrix und Liste

### Subjektive Ergebnisse

**Bevorzugte Darstellungsweise:** Im Abschlussfragebogen wurde den Versuchspersonen die Möglichkeit gegeben, sich zu den Darstellungsarten zu äußern. Über die Hälfte der Fahrer bevorzugte die Matrixdarstellung; 35,7% entschieden sich für die Liste. Die vollständigen Ergebnisse sind in der Abbildung 8.39 zu sehen. Einige der Versuchspersonen äußerten die Anregung, für Kategorien die Matrixdarstellung zu verwenden und für viele Informationen in tieferen Hierarchiestufen Listen zu benutzen.

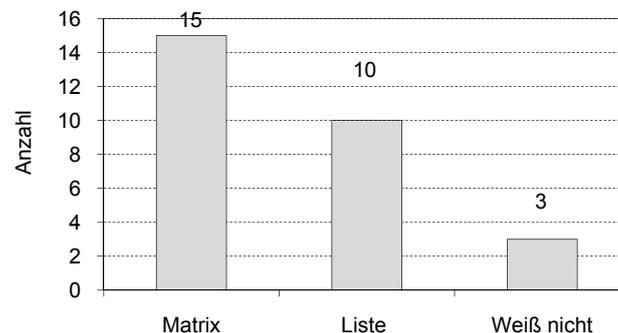


Abbildung 8.39: Bevorzugte Darstellungsweise

**Abhängigkeit der Anzahl der Einträge:** Am Besten wurde die Matrix mit fünf Elementen bewertet und erreichte eine Durchschnittsbewertung von 1,64. Darauf folgend wurde die Liste mit ebenfalls fünf Elementen mit einer 2,16 benotet. M9 und L9 bekamen jeweils eine drei. Dies zeigt deutlich, dass die Darstellungsweisen mit der kleineren Informationsmenge bevorzugt werden.

### Schlussfolgerungen

Zusammenfassend lassen sich folgende Empfehlungen aus den Ergebnissen ableiten:

- **Mehr Auswahlelemente führen zu längeren Bediendauern:**

Die Auswahldauer der Liste5 – unter Verwendung des HUDs – lag geringfügig über dem

Niveau der Matrix5. Die Darstellungsarten mit 9 Elementen benötigten jeweils mehr Zeit als die mit 5 Elementen.

- **Die Bediendauer der Matrixdarstellung ist kürzer als die der Listendarstellung:**

In beiden Fällen aber, also bei Darstellungen mit 5 oder 9 Elementen, schnitt jeweils die Matrixdarstellung etwas besser ab, was sich zudem bei Matrix9 gegenüber Liste9 als statistisch signifikant erwies. Auch hinsichtlich der Bedienschritte schnitt die Matrixdarstellung besser ab als die Listendarstellungen. Dies lag aber auch an der guten Abstimmung des Bedienelements (5-Wege-Navipad) mit der Darstellungsart.

- **Listendarstellung ist gegenüber Matrixdarstellung weniger ablenkend:**

Hinsichtlich der Spurabweichung zeigte sich jedoch, dass die Verwendung einer Listendarstellung weniger ablenkend als eine Matrixdarstellung ist. So ergibt sich, dass die Matrixdarstellung gegenüber einer Listendarstellung zu kürzeren Bediendauern führt und wesentlich effizienter hinsichtlich der Bedienschritte ist. Jedoch bewirkt die Matrixdarstellung auch eine größere Spurabweichung, was auf eine höhere kognitive Belastung schließen lässt.

- **Einflüsse der Darstellungsart auf PDT:**

Aus den Daten der PDT konnte man hinsichtlich der Reaktionszeiten und der Anzahl übersehener PDT-Punkte eine allgemeine Verschlechterung im Vergleich zu Fahrten ohne Menübedienung feststellen. Die Anzahl der übersehenen PDT-Punkte ist jedoch zwischen den Darstellungsweisen Matrix und Liste gering und nicht signifikant unterscheidbar. Auch in diesem Fall war die kognitive Beanspruchung der Fahrer möglicherweise nicht groß genug, um eine größere Anzahl der PDT-Punkte zu übersehen. Es zeigt sich lediglich eine leichte Tendenz in der Entdeckungsleistung von PDT-Punkten auf Höhe der HUD-Darstellung und an den Rändern sowie eine leichte Aufmerksamkeitsverschiebung des Fahrers. Die älteren Fahrer aber zeigten eine bessere Leistung bei der Listendarstellung. Bei der subjektiven Befragung präferierten die Fahrer eine Listendarstellung aufgrund erhöhter Übersichtlichkeit. Die Nutzer, welche die Liste bevorzugten, sahen die Vorteile in der Bedienung nur in einer Richtung, was weniger anstrengend ist.

- **Einfluss der Darstellungsart im HUD: Vergleich Matrix vs. Liste:**

Matrixdarstellung ist belastender, aber etwas schneller und übersichtlicher und daher für obere Hierarchiestufen geeignet. Die Liste spielt bei vielen Einträgen ihren Vorteil aus und ist weniger kognitiv belastend.

### 8.5.6 Untersuchung der multimodalen Bedienung im HUD (Simulatorstudie II)

Ziel dieser Untersuchung war es, die multimodale Bedienung des HUD\_MAX-Konzeptes im Fahrzeug zu untersuchen. Für die Evaluierung wurden 22 Versuchspersonen akquiriert, davon waren 15 männlich und 7 weiblich. Der Altersdurchschnitt lag bei 37,7 Jahren. Als Aufgaben hatten die Versuchspersonen die Zieleingabe und den Tankagenten zu bedienen. Die Wahl der Modalitäten sowohl die Sprache als auch das 5-Wege-Navipad wurde ihnen freigestellt. Es erfolgte eine Videoaufzeichnung und Befragung über einen Fragebogen.

## Ergebnisse

Die Akzeptanz der Informationsagenten war gut, sie wurden mit den Schulnoten 1,6 für den Tankagenten und 1,9 für den Restaurant-Recommender bewertet. Sprachliche und taktile Eingaben hielten sich bei ihrer Bedienung in der Waage. Vor allem die PTT-Taste lässt viel Raum für Verbesserungen. Nur 14 % der Teilnehmer wählten intuitiv die richtige Bedienstrategie. Allerdings erschweren die konkurrierenden Systemmodelle der Probanden eine einzelne optimale Lösungsfindung.

Die visuelle Rückmeldung des Spracherkenners durch das beschriebene Speechy-Symbol muss auffälliger platziert werden. Die Farbcodierung war verständlich.

Der Anteil von natürlichsprachlichen Eingaben war, über alle Aufgaben gemittelt, mit 21 % geringer als der von Kommandos. Dennoch ist der höhere technische Aufwand gerechtfertigt, der durch die Unterstützung natürlich-sprachlicher Eingaben entstand, da diese 21 % der Spracheingaben anderenfalls Fehler hervorgerufen hätten.

Sämtliche Personen, die zu Beginn die Zieladresse mittels Slot-Filling eingaben, wechselten zur Frame-Based Eingabe, nachdem sie diese kennen gelernt hatten. Der Wechsel von 50 % der Personen, die zu Beginn die direkte Eingabestrategie gewählt hatten, zur Eingabe mit Hilfe der Sprachmenüführung lässt sich vor allem durch Fehler des Spracherkenners aufgrund von Fehlbedienungen der PTT-Taste erklären. Die erhöhte Effektivität rechtfertigt zusammen mit dem in der Befragung geäußerten Wunsch, mehrere Informationen in einem Schritt einzugeben, die Verwendung eines Frame-Based Ansatzes (siehe Kapitel 3.1.3). Weitere Ergebnisse hierzu finden sich in [SAMR05].

# Kapitel 9

## Resumé und Ausblick

Die vorliegende Arbeit hat umfassende Lösungsansätze zur Nutzung großer Datenmengen im Fahrzeug aufgezeigt. Dabei liefern die einzelnen Kapitel Bausteine zur Schaffung eines ganzheitlichen Informationsmanagements auf drei verschiedenen Ebenen (siehe Abschnitt 1.3). Es wurden Methoden zur Informationsverarbeitung und effiziente Dialogstrategien für den Fahrer identifiziert, sowie geeignete multimodale Anzeige- und Bedienkonzepte entwickelt. Anschließend wurden diese in zahlreichen Usability-Studien hinsichtlich ihrer Eignung im Fahrzeug bewertet.

### 9.1 Methoden zur Informationsverarbeitung

Zunächst werden nun die gewonnenen Erkenntnisse zu den Methoden der Informationsverarbeitung im Automobil betrachtet.

Das aufgezeigte Dialogframework (siehe Kapitel 3) ermöglicht die Realisierung multimodaler und kontextadaptiver Dialogsysteme. Dabei gliedert sich das Dialogframework in die Komponenten Präsentations- und Dialogmanager sowie Applikationsinterface. Um gerade den Anforderungen großer Datenmengen gerecht zu werden, erfolgte die Erweiterung durch ein Datenmanagement-System, das über Webservices (SOAP) Dienste für Datenbank-Anfragen bereitstellt. Insgesamt schafft dieses Framework das Grundgerüst zur Realisierung der erarbeiteten Dialogstrategien für den Fahrer. Die Formulierung der Dialoge erfolgt über ein XML-Regelwerk, das neben der Steuerung des Dialogverlaufs auch Regeln zur Verarbeitung verschiedener Ein- und Ausgabekanäle und Kontextgrößen mit einbeziehen kann. Dabei lässt sich der Dialogverlauf über Zustandsautomaten und Statecharts anschaulich darstellen.

Zur Verarbeitung multimodaler Eingaben wurde der Präsentationsmanager u. a. durch die Komponenten Fusion und Fission erweitert. Für die Fusion verschiedener Kanäle wurde eine Reihe etablierter Integrationsmethoden präsentiert. Diese Arbeit verwendet die Late-Semantic Fusion, die weitgehend auf regelbasierten Fusionsstrategien sowie temporaler Integration in Kombination mit Unification-Based Ansätzen zurückgreift. Die gewählten Methoden bieten für den Fahrer eine transparente und verständliche Methodik zur Integration verschiedener Kommunikationskanäle.

Zur Verarbeitung von Suchanfragen der Fahrzeug-Suchmaschine wurden verschiedene Suchalgorithmen aufgezeigt (siehe Abschnitt 4). Sollen auch unscharfe Eingaben, wie Fehler in der Schreibweise und unvollständige Eingaben, mitberücksichtigt werden, kommen Distanzfunktionen und phonetische Verfahren zum Einsatz. Die Levenshtein-Distanz und der Soundex-

Algorithmen wurden als geeignet identifiziert. Allerdings muss dabei auch die Precision- vs. Recall-Problematik betrachtet werden, da große Treffermengen auch die Auffindung erschweren können. Zur sortierten Ausgabe der Treffer wurden Rankingmethoden, wie einfache Häufigkeitsalgorithmen und das Topic-Based Vector Space Model, verwendet. Bei der Realisierung des verknüpften Informationsbrowsers aus Kapitel 5 kommt das Tagging der Daten zur Anwendung. Dies stellt eine geeignete Methode dar, um vernetzte Daten mehreren Gruppen zuzuordnen. Dadurch können relevante Daten anschließend einfach über die Tags gefiltert werden.

Die Realisierung der Informationsagenten (siehe Kapitel 6) stellt spezielle Anforderungen an die Informationsverarbeitung. Zur Umsetzung der Wizard-Dialoge der Informationsagenten kommen Frame-Based Ansätze zur Anwendung (siehe Kapitel 3). Durch die Verwendung von Frame-Based Ansätzen können sowohl Dialogphrasen als auch natürlich-sprachliche Dialoge realisiert werden. Der Frame-Based Ansatz bietet dem Fahrer die Möglichkeit mehrere Intentionen gebündelt in einem Dialogschritt zu übermitteln. Die Dialogreihenfolge wird jedoch nicht zwingend vorgegeben. Zur Intentionsdetektion bei natürlich-sprachlichen Eingaben kommen neben Stemming-Verfahren Bayes'sche Netze zur Anwendung. Um die Anforderungen der Informationsagenten hinsichtlich einer „intelligenten“ Informationsfilterung zu bewerkstelligen, wurde das Dialogframework um die Komponente eines Informationsmanagers erweitert. Die Informationsmanager-Komponente übernimmt hauptsächlich die Filterung und das Timing von Informationen und Daten. Zur Realisierung der Datenfilterung kommen bei den Informationsagenten regelbasierte und statistische Ansätze, wie Bayes'sche und Neuronale Netze, zum Einsatz. Der Systemaufbau und die Funktionsweise wurden exemplarisch für einen Tank-, Kontakt- und Terminagenten und einen Restaurant-Recommender vorgestellt. Die theoretische Bewertung hat gezeigt, dass die Bayes'schen Netze gewünschte Tankstellen der Fahrer besser platzieren und dadurch Bedienschritte reduzieren können. In einer Voruntersuchung konnten die Marke, der Preis, die verbleibende Tankreichweite und die Entfernung der Tankstelle als die wichtigsten Kriterien identifiziert werden. Die Bayes'schen Netze bieten insgesamt den Vorteil, dass hier Vorwissen und regelbasiertes Wissen mit einfließen können und damit ein Training nicht erforderlich ist. Durch Berücksichtigung eines Lernfaktors kann das Bayes'sche Netz sich an Veränderungen des Tankverhaltens anpassen. Zur Empfehlung von Restaurants kamen Neuronale Netze zur Anwendung. Hier ist allerdings ein längeres Training des Neuronalen Netzes erforderlich, was einen gewissen Nachteil darstellt. Dabei wurden die Kriterien Preiskategorie, Küche, Art des Restaurants und Entfernung als Eingänge für das Probabilistische Neuronale Netz verwendet. Die Entscheidungen durch das neuronale Netz sind möglicherweise weniger transparent für den Nutzer, allerdings könnte dies durch den Empfehlungscharakter der Agenten hier als weniger störend empfunden werden.

## 9.2 Dialogstrategien

Als Nächstes werden nun die gewonnenen Erkenntnisse zu den effizienten Dialogstrategien zusammengetragen. Dabei wurde zunächst deutlich, dass die bestehenden, weitgehend hierarchischen Menüstrukturen den Anforderungen großer Datenmengen nur unzureichend gerecht werden. Sie erfordern viele Bedienschritte und können den Nutzer in Sackgassen führen.

Zur gezielten Auffindung von Daten und Informationen wurde zuerst eine datenübergreifende Fahrzeug-Suchmaschine identifiziert und entwickelt (siehe Kapitel 4). Dieses Konzept orientiert sich am Aufbau von Suchmaschinen aus dem PC- und Internet-Bereich. Die Fahrzeug-

Suchmaschine ist in der Lage auf sämtliche Daten des Fahrerinformationssystems inklusive Online-Systeme und mobile Endgeräte zuzugreifen. Die Bewertungen der Suchinterfaces im Fahrzeug haben gezeigt, dass diese Systeme akzeptiert und gewünscht werden. Sie stellen eine geeignete Methode für einen schnellen, zielgerichteten Datenzugriff dar und zeigen gerade bei großen Datenmengen und versteckten Funktionen im Menü ihre Stärken. Die Untersuchungen haben verdeutlicht, dass Methoden zur Sortierung, Kategorisierung und Filterung sehr wichtig sind. Diese Methoden bieten dem Fahrer Möglichkeiten zur Exploration der Treffer und zeigen eine alternative Zugriffsmethode auf, die einen großen Zuspruch fand.

Zur Bewältigung komplexer Aufgabensequenzen und motiviert durch die Ergebnisse zur Fahrzeug-Suchmaschine lieferte das Kapitel 5 einen Lösungsansatz, der ein vernetztes Datenbrowsing ermöglicht. Dabei erfolgt der Zugriff auf die Informationen unterschiedlicher Quellen (z. B. mobiler Endgeräte) über eine datenorientierte Grundstruktur. Die Grundstruktur ermöglicht eine verbesserte Orientierung für den Nutzer und gliedert sich in die Domänen Kontakte, Termine, Route, Unterhaltung und Ereignisse. Durch die objektorientierte Gliederung können Funktionen schnell und einfach ausgeführt werden. Das Statusmenü liefert einen informativen, fahrgerechten Überblick zu relevanten Inhalten der Domänen. Es bietet einen direkten Zugriff auf wichtige Funktionen und Einsprünge in die Domänen. Zur schnellen Datenauffindung wurden implizite und explizite Filterstrategien zur Verfügung gestellt. Gerade die Sprungfunktionen in Kombination mit den impliziten Filterstrategien ermöglichen durch Einbezug des Kontexts eine sehr gute fahrer- und aufgabenorientierte Bedienung. So können gerade komplexe Aufgabensequenzen für den Fahrer deutlich vereinfacht und gebündelt werden.

Im Hinblick auf eine umfassende Unterstützung des Fahrers wurden ähnlich einem „intelligenten“ Beifahrer die Informationsagenten eingeführt (siehe Abschnitt 6). Durch die Informationsagenten werden dem Fahrer komplexe und häufig wiederkehrende Handlungsabläufe erheblich vereinfacht. Die Nutzerführung erfolgt hier schrittweise über multimodale Wizard-Dialoge. Dabei ergänzen sich sprachbasierte Dialoge in Verbindung mit taktilem Eingabe und visueller Führung im Head-Up Display. Über die Frame-Based Sprachdialoge, welche die Übermittlung mehrerer Intentionen in einem Bedienschritt ermöglichen, lassen sich die Aufgabenschritte effizient und gebündelt ausführen.

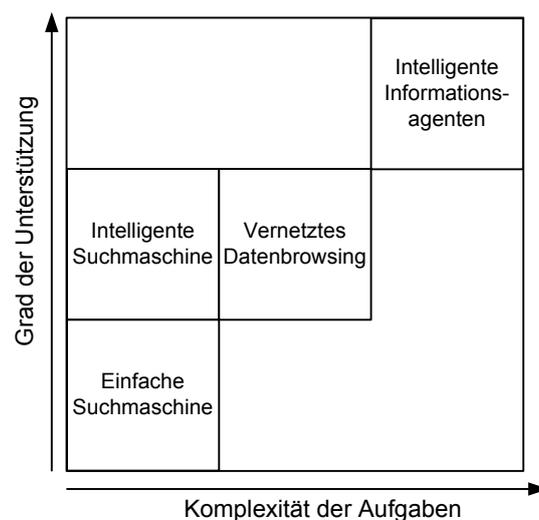


Abbildung 9.1: Überblick Dialogstrategien

Eine abschließende Betrachtung der präsentierten Dialogstrategien zur Nutzung großer Datenmengen im Fahrzeug macht deutlich, dass bei einer steigenden Komplexität der Aufgaben des Fahrers auch ein zunehmender Unterstützungsgrad bei der Dialogführung erforderlich wird (siehe Abbildung 9.1). Zur einfachen, schnellen Auffindung von Daten eignet sich die Fahrzeug-Suchmaschine. Die Verwendung von „intelligenten“ Suchalgorithmen kann das Auffinden von relevanten Treffern weiter verbessern. Filter- und Sortiermechanismen können dabei den Nutzer bei der Eingrenzung der Daten deutlich unterstützen. Werden allerdings vernetzte Aufgabensequenzen erforderlich, was aus den Szenarien und Fahrerbedarfen deutlich wird, so bedarf es einer verstärkten Nutzerführung. Der Datenbrowsing-Ansatz bietet durch die einfache datenorientierte Grundstruktur mit kontextabhängigen Verknüpfungen die einfache Durchführung umfangreicher Bediensequenzen. Zur Bewältigung sehr komplexer Bedienungsaufgaben ist die schrittweise, gebündelte Nutzerführung durch die vorgestellten Informationsagenten besonders hilfreich, da diese sehr effiziente Dialoge ermöglichen.

### 9.3 Multimodale Anzeige und Bedienung

Zunächst werden die Ergebnisse der *multimodalen Bedienung* betrachtet. Die Fahrzeug-Suchinterfaces stellen spezielle Anforderungen an die alphanumerische Eingabe der Suchbegriffe (siehe Kapitel 4). Große Akzeptanz fand hier die Spracheingabe, die jedoch Einschränkungen hinsichtlich Vokabulars mit sich bringt. Betrachtet man die taktilen Eingabeverfahren so fanden vor allem Touchtechnologien wie Handschrifterkennung und On-Screen-Keyboard eine große Zustimmung, die eine simultane Auswahl der Buchstaben ermöglichen. Für eine gute Blindbedienbarkeit wünschten die Versuchsteilnehmer unbedingt akustisches Feedback. Sequenzielle Eingabeverfahren wie Dreh-/Drücksteller und Ziffernblock fanden weniger Akzeptanz aufgrund der langen Aufgaben- und Bediendauern. Zur Auswahl aus den Trefferlisten wünschten die Nutzer weitgehend taktile Verfahren über Dreh-/Drücksteller oder 5-Wege-Navipad. Diese sequenziellen Verfahren ermöglichen eine gute Exploration von Treffermengen. Wurde ein Zielobjekt durch den Nutzer identifiziert, so wurden auch parallele multimodale Bedienstrategien identifiziert (siehe Kapitel 7). Durch Point-and-Talk kann das Objekt durch Zeigen selektiert und parallel über einen Sprachbefehl eine Aktion gestartet werden. Zur Bedienung der Informationsagenten über das Head-Up Display fanden hier die multimodale Bedienung über ein 5-Wege-Navipad am Lenkrad und Sprache große Akzeptanz bei den Nutzern (siehe Kapitel 8). Die Dialoge der Informationsagenten können entweder systeminitiiert über eine visuelle Anzeige im HUD in Kombination mit einer Sprachmeldung gestartet werden. Des Weiteren können die Informationsagenten über Spracheingabe durch den Nutzer bei Bedarf initiiert werden. Frame-Based Dialoge bieten dabei den Vorteil, dass mehrere Intentionen in einem Dialogschritt durch natürlich-sprachliche Äußerungen oder Phrasen übermittelt werden können. Die Nutzerführung erfolgt hier schrittweise visuell und akustisch über das HUD. Der Status des Spracherkenners und die erkannten Intentionen werden visuell über das HUD rückgemeldet. Die Auswahl aus mehreren Optionen kann multimodal wahlweise über Sprechen der Option oder Auswahl über das 5-Wege-Navipad ausgeführt werden. Abschließend lässt sich folgern, dass die Verwendung alternierender Multimodalität im Fahrzeug akzeptiert wird und weitere Effizienzsteigerungen ermöglicht. Wesentlich ist hier, dass der Nutzer, entsprechend verschiedener Aufgaben, geeignete Modalitäten bereitgestellt bekommt und der Modalitätenwechsel einfach und verständlich ausgeführt werden kann. Durch Multimodalität können komplexe Aufgaben und Handlungssequenzen weiter vereinfacht werden und die

Belastung bei der Bedienung für den Fahrer reduzieren.

Betrachtet man die *Darstellungskonzepte* der entwickelten Dialogstrategien, so haben sich folgende Erkenntnisse gezeigt. Zur Darstellung größerer Informations- und Datenmengen eignen sich die graphischen Displays innerhalb des Fahrzeugs, wie das zentrale Informationsdisplay (CID). Sowohl beim Suchmaschinen- als auch beim Browsing-Interface hat sich gezeigt, dass hierfür ein einfaches, übersichtliches Layout in wenigen Segmenten sinnvoll ist (siehe Abschnitt 4 und 5). Bei einer objektorientierten Darstellung bietet sich die Gliederung in die Bereiche Datenobjekte, Funktionen und Filter-/Sortiermethoden an. Zur einfachen und effizienten Aufnahme von umfangreichen Datenmengen bieten innovative Visualisierungsmethoden (siehe Kapitel 7) neue Möglichkeiten. 3D-Visualisierungsmethoden wie Zoomable User Interfaces schaffen gemeinsam mit innovativen Eingabetechnologien fortschrittliche Nutzerschnittstellen, die neben Effizienz auch einen wesentlichen Beitrag zur Bedienfreude leisten können. Das vorgestellte Datenbrowsing wurde hier unter Anwendung von Zoomable User Interfaces und Fisheye-Techniken erweitert. Zur Steuerung des semantischen Informationszooms wurde ein Touchscreen mit Annäherungssensorik verwendet. Neben einer Vergrößerung werden beim Eintauchen (Annäherung des Fingers) zunehmend weitere Details zum aktuellen Kontext sichtbar. Neben Fisheye-Listen, -Buttons und -Slider wurde auch eine Fisheye-Tastatur umgesetzt. Durch den Einsatz von bekannten 3D-Metaphern wurden weitere Konzepte vorgestellt, die große Datenmengen und Menüstrukturen anschaulich aufbereiten. Das Zoomable User Interface fand insgesamt eine große Akzeptanz bei den Testpersonen. Über die flache Grundstruktur des Browsing-Ansatzes konnten die Funktionen einfach ausgewählt werden. Durch die Verwendung der Zooming und Panning-Strategien hatten die Nutzer ein verbessertes Gefühl der Orientierung. Des Weiteren erreichten die ZUI-Effekte in Verbindung mit dem Annäherungs-Touchscreen eine deutlich gesteigerte Bedienfreude bei den Nutzern.

Durch die erweiterten Möglichkeiten eines großflächigen Head-Up Displays wird im Kapitel 8 ein fahrerorientiertes Anzeige- und Bedienkonzept vorgestellt, welches die Informationen in Abhängigkeit der Priorität für den Fahrer in drei verschiedenen Anzeigenbereiche im Head-Up Display gruppiert. Für die effiziente Dialogführung im HUD eignen sich ganz besonders die entwickelten Informationsagenten. Die Evaluierung wesentlicher Fragestellungen großflächiger Head-Up Displays hat folgende Erkenntnisse gebracht: Das Head-Up Display als neues Anzeigemedium wurde von den Probanden sehr akzeptiert und bot eine effiziente Informationsaufnahme. Es hat sich auch gezeigt, dass die Versuchspersonen im Vergleich zum herkömmlichen Serien-Head-Up Display weitere Informationen und Funktionen wünschen. Die Zahl der unbekanntenen Informationen sollte jedoch in einem sinnvollen Umfang bleiben. Bei der Verwendung des Head-Up Displays zur Menüinteraktion zeigten sich reduzierte Bearbeitungsdauern gegenüber dem zentralen Informationsdisplay. Große Unregelmäßigkeiten in der Spurhaltung konnten dabei nicht festgestellt werden. Allerdings waren die Blicke auf das Head-Up Display zum Teil länger, dafür weniger häufig als auf das zentrale Informationsdisplay. Insgesamt fand die Bedienung im HUD eine sehr große Nutzerakzeptanz. Dabei bewerteten die Versuchspersonen die Bedienung im HUD als weniger gefährlich, was allerdings auch ein Anzeichen eines mangelnden Gefahrenbewusstseins der Nutzer darstellen könnte. Bei der Bewältigung von Aufgaben bei zwei verschiedenen Informationsmengen im HUD war keine Verschlechterung der PDT-Erkennungsrate und keine Anzeichen eines Tunnelblicks zu erkennen. Der Vergleich der Darstellungsarten Matrix gegenüber einer Liste im Head-Up Display brachte geringe Zeitvorteile bei der Matrix, allerdings war ein etwas höherer kognitiver Aufwand wahrnehmbar. Während die jüngeren Probanden die Matrixdarstellung bevorzugten, wünschten die Versuchspersonen der Altersklasse 51-60 dagegen die einfachere Listendarstellung.

Des Weiteren wurden Ansätze für ein prioritätenorientiertes Zusammenspiel zwischen dem HUD und einem freiprogrammierbaren Kombi aufgezeigt. Schließlich wurde auf die technische Realisierung quasi-kontaktanaloger Anzeigen im Head-Up Display eingegangen. Neben möglichen Darstellungskonzepten wurde anschließend ein prototypischer Sitzkistenaufbau erläutert. Zur Detektion der Blickrichtung wurde ein eigener Remote-Gazetracker entwickelt.

## 9.4 Ausblick

Die eingesetzten Bewertungsmethoden dieser Arbeit lieferten erste Tendenzen zur Eignung der neuen Lösungskonzepte für das Fahrzeug. Hierbei zeigte sich eine große Akzeptanz und Effizienz der identifizierten Dialogstrategien und der entwickelten multimodalen Anzeige- und Bedienkonzepte zum Umgang mit den großen Datenmengen. Zur tiefgreifenden Bewertung der erarbeiteten Ansätze sind allerdings weitere umfassende Studien im Fahrzeug erforderlich.

Es bleibt insgesamt festzuhalten, dass im Zuge des Informations- und Kommunikationszeitalters ein steigender Informationsbedarf im Fahrzeug unumgänglich ist. Wie bereits eingangs erwähnt, werden die gegenwärtigen mobilen Endgeräte den aktuellen Anforderungen im Fahrzeug bisher keineswegs gerecht. Die Automobilhersteller müssen allerdings auf die technischen Innovationen der Consumer Produkte und die gesteigerten Bedürfnisse der Fahrer reagieren und sich den hohen Herausforderungen stellen. Diese Arbeit stellt dazu geeignete Ansätze bereit, welche die fahrgerechte Integration des gesteigerten Informationsumfangs ermöglicht. Um auch in Zukunft den hohen Anforderungen gerecht zu werden, muss sich die Automobilindustrie zunehmend den umfangreichen Softwareanforderungen stellen. Andernfalls wird es schwierig, die Wettbewerbsvorteile aktueller Fahrerinformationssysteme gegenüber den kurzen Entwicklungszyklen der Consumer Produkte zu verteidigen. Die Consumer Produkte werden durch Nachrüstlösungen versuchen, diesen Vorsprung zunehmend streitig zu machen.

Innovative Benutzerschnittstellen liefern daher weiterhin großes Forschungs- und Entwicklungspotential. Um die Entwicklungszyklen dieser komplexen Systeme zu verkürzen, kann die Etablierung virtueller Prototyping-Techniken einen weiteren Beitrag leisten. Im Rahmen dieser Arbeit wurde durch den Aufbau einer CAVE-Umgebung am Lehrstuhl für Mensch-Maschine-Kommunikation der Grundstein gelegt, um die schnelle Realisierung und Bewertung von multimodalen Ansätzen zu ermöglichen. Virtual Reality eröffnet dabei eine Methodik zur immersiven Bewertung von Interfaces bereits in einer frühen Entwicklungsphase.



Abbildung 9.2: CAVE des Lehrstuhls für Mensch-Maschine-Kommunikation

# Abkürzungsverzeichnis

<b>ABK</b>	Anzeige- und Bedienkonzept
<b>ABS</b>	Antiblockiersystem
<b>ACC</b>	Active Cruise Control
<b>AI</b>	Application Interface
<b>ANSI</b>	American National Standards Institute
<b>AR</b>	Augmented Reality
<b>ASCII</b>	American Standard Code for Information Interchange
<b>ASR</b>	Automatic Speech Recognition
<b>AVA</b>	Aktionsvorhersage Algorithmen
<b>BM</b>	Boyer-Moore
<b>BNF</b>	Backus-Naur-Form
<b>CAVE</b>	Cave Automatic Virtual Environment
<b>CD</b>	Compact Disc
<b>CFG</b>	Context Free Grammar
<b>CID</b>	Central Information Display
<b>CW</b>	Collision-Warning
<b>DB</b>	Data Base
<b>DDS</b>	Dreh-/Drücksteller
<b>DIN</b>	Deutsches Institut für Normung e.V.
<b>DIS</b>	Driver Information Systems
<b>DM</b>	Dialog Manager
<b>DSC</b>	Dynamic-Stabilitiy-Control
<b>DVD</b>	Digital Video Disc

<b>EDGE</b>	Enhanced Data Rates for GSM Evolution
<b>EOG</b>	Elektrookulogramm
<b>FIS</b>	Fahrerassistenzsysteme
<b>FIS</b>	Fahrerinformationssysteme
<b>FMC</b>	Fundamental Modeling Concept
<b>FOV</b>	Field of View
<b>GPS</b>	Global Positioning System
<b>GRA</b>	Geschwindigkeitsregelanlage
<b>GRP</b>	Gaze Retention Period
<b>GUI</b>	Graphical User Interface
<b>HCI</b>	Human-Computer Interaction
<b>HDD</b>	Head-Down Display
<b>HITS</b>	Hypertext-Induced Topic Selection
<b>HMD</b>	Head-Mounted Display
<b>HMI</b>	Human-Machine Interface
<b>HMM</b>	Hidden Markov Model
<b>HTML</b>	Hypertext Markup Language
<b>HUD</b>	Head-Up Display
<b>HWR</b>	Hand Writing Recognition
<b>ID3</b>	Identify an MP3
<b>IM</b>	Information Management
<b>IP</b>	Internet Protocol
<b>IPAM</b>	Incremental Probabilistic Action Modeling
<b>IR</b>	Information Retrieval
<b>ISO</b>	International Organization for Standardization
<b>IVIS</b>	In-Vehicle Information Systems
<b>KDE</b>	K Desktop Environment
<b>KI</b>	Künstliche Intelligenz
<b>KMP</b>	Knuth-Morris-Pratt

---

<b>LAN</b>	Local Area Network
<b>LCS</b>	Longest Common Substring
<b>LCT</b>	Lane Change Task
<b>LED</b>	Light Emitting Diode
<b>LOG</b>	Line of Gaze
<b>MMI</b>	Mensch-Maschine Interaktion
<b>MMS</b>	Mensch-Maschine System
<b>MP3</b>	MPEG-1 Audio Layer 3
<b>MPEG</b>	Moving Picture Experts Group
<b>MTC</b>	Members-Teams-Committee
<b>MVC</b>	Model-View-Controller Modell
<b>NLU</b>	Natural Language Understanding
<b>NN</b>	Neuronal Network
<b>NUM</b>	numerischer Ziffernblock
<b>NV</b>	Night Vision
<b>OMG</b>	Object Managment Group
<b>OSK</b>	On-Screen-Keyboard
<b>PA</b>	Park-Assistent
<b>PC</b>	Personal Computer
<b>PDA</b>	Personal Digital Assistant
<b>PDC</b>	Park-Distance-Control
<b>PDT</b>	Peripheral Detection Task
<b>PNA</b>	Portable Navigation Assistant
<b>PNN</b>	Probabilistic Neural Network
<b>POG</b>	Point of Gaze
<b>POI</b>	Point Of Interest
<b>PPM</b>	Prediction by Partial Matching
<b>PTT</b>	Push-To-Talk
<b>SMS</b>	Short Message Service

<b>SOAP</b>	Simple Object Access Protocol
<b>T9</b>	Text on 9 keys
<b>TCP/IP</b>	Transmission Control Protocol/Internet Protocol
<b>TFT</b>	Thin Film Transistor
<b>TSVM</b>	Topic-based Vector Space Model
<b>TTT</b>	Total Task Time
<b>UDP</b>	User Datagram Protocol
<b>UML</b>	Unified Modeling Language
<b>UMTS</b>	Universal Mobile Telecommunications System
<b>VOG</b>	Videookulogramm
<b>VSM</b>	Vector Space Model
<b>WIMP</b>	Windows, Icons, Menus and Pointer
<b>XHTML</b>	erweiterbares HTML
<b>XML</b>	Extensible Markup Language
<b>ZBE</b>	Zentrale Bedieneinheit
<b>ZUI</b>	Zoomable User Interface

# Index

- Ablaufdiagramm, 49
- Adaptivität, 46
- Agentensysteme, 107
- Aktionsvorhersage, 44
- Alphanumerische Eingabestrategien, 76
- Animationen, 136
- Architektur
  - multimodale, kontextadaptive, 51
- AttrakDiff, 28
- Audi MMI, 13
- Aufmerksamkeit, 19
- Augmented Reality, 163
  
- Backus-Naur-Form, 50
- Bayes'sche Netze, 111
- Bedienfreude, 24
- Bewertung
  - Informationsagenten, 131
  - Suchmaschine, 83
- Bewertungsmethoden, 25
- Blickdetektion, 173
- Blickverhaltensmessung, 29
- BMW iDrive, 12
- Browsing, 91
  
- Cognitive Capture, 154
  
- Datenbrowsing, 91
- Datenmanagement, 54
- Dialogführung, 31
- Dialogmanager, 110
- Dialogmodellierung, 47
- Dialogsysteme, 31
  - kontextadaptive, 43
- Dilogformen, 31
- Direkte Manipulation, 134
- Drei-Ebenen-Modell
  - der Fahrzeugführung, 8
  - nach Rasmussen, 7
  
- Ein-Kanal-Theorie, 19
- Einflussgrößen, 16
  
- Fahraufgabe
  - primäre, 9
  - sekundäre, 9
  - tertiäre, 10
- Fahraufgaben, 9
- Fahrerassistenzsysteme, 10
- Fahrerbedarfe, 15
- Fahrerinformationssysteme, 10
- Fahrerzustandsüberwachung, 11
- Fahringsimulator, 25
- Fahringszenarien, 17
- Fahrtchronologie, 17
- Fahrzeugdomäne, 7
- Fisheye-Effekte, 136
- Framework, 50
- Fusion
  - frühe, 40
  - hybride, 41
  - semantische, 40
  - späte, 40
- Fusionsmethoden, 41
  
- Gazetracking, 173
- Gestaltgesetze, 24
- Gestaltungshinweise, 21
- Guidelines, 23
  
- Head-Up Display, 149
  - Funktionsweise, 150
  - kontaktanaloges, 163
- HUD MAX Konzept, 155
  
- Informationsagenten, 105
  - Architektur, 109
  - Kontaktagent, 124
  - Konzeption, 119
  - Tankagent, 119

- Terminagent, 126
- Informationsmanager, 110
- Informationsvisualisierung, 133
- Integration
  - hybride, 42
  - regelbasierte, 41
  - semantische, 41
  - statistische, 41
  - temporale, 41
  - unification-based, 42
- Integrationsmethoden, 41
- Interaktionsstile
  - funktionsorientiert, 32
  - objektorientiert, 32
- Joy of Use, 24
- Kombiinstrument, 15
- Kontext, 16
- Lane Change Task, 26
- Lexus Multiinformationsdisplay, 14
- Menüdialoge, 33
- Menüstrukturen
  - hierarchische, 33
  - netzförmige, 34
- Mensch-Maschine Interaktion, 5
- Mentale Modelle, 21
- Mercedes COMAND, 14
- Metaphern, 134
- Modellierung
  - Dialoge, 47
- Multimodale Fusion, 40
- Multimodales Interface
  - Informationsagenten, 161
- Multimodalität, 38
  - im Automobil, 42
  - Datenbrowsing, 102
  - Vorteile, 39
- Multiple Resource Theory, 19
- MVC-Modell, 51
- Navigationshilfen, 34
- Neuronale Netze, 115
  - probabilistische, 119
- Normen, 22
- Nutzerkontext, 16
- Orientierungshilfen, 34
- Passungsmodell, 6
- Perceptual Tunneling, 154
- Peripheral Detection Task, 27
- POC-Kurven, 20
- Präsentationsmanager, 110
- Ranking, 74
- Rasmussen, 7
- Richtlinien, 21
  - FIS und FAS, 25
- Seeheim-Modell, 50
- Semantisches Differential, 28
- Sichtsimulation, 27
  - Lane Change Task, 26
- Situation Awareness, 21
- Sprachdialoge, 34
  - Agent-Based, 37
  - Finite-State-Based, 35
  - Frame-Based, 35
  - Slot-Filling, 35
- Statecharts, 48
- Styleguides, 24
- Suchalgorithmen, 64
- Suchinterface
  - CARhoo, 78
  - CARpedia, 80
  - IMSENIO, 79
- Suchmaschine, 57
- Systemkontext, 16
- Szenarien, 17
- Transitionen, 48
- Tunneling, 154
- Umweltkontext, 16
- Usability, 21, 22
- Verhaltensmodelle, 19
- Visualisierungsstrategien, 133
- Wahrnehmungspsychologie, 18
- Wissensquellen, 37
- Working Memory, 19
- Workload, 21
- XML, 50

Zoomable User Interface, 134

Zustände, 47

Zustandsautomaten, 47



# Literaturverzeichnis

- [AAB<sup>+</sup>06] ABLASSMEIER, M. ; ACKERMANN, Ch. ; BROY, V. ; GRÜNDL, M. ; LANGE, Ch. ; LIBOSSEK, M. ; WINTER, S: Mensch-Maschine-Interaktion im Kraftfahrzeug. Abschlussbericht Phase III. / TUMMIC. München, 2006. – Forschungsbericht
- [Abl03] ABLASSMEIER, M.: *Untersuchung netzförmiger Dialogstrukturen hinsichtlich ihrer Verwendbarkeit im Fahrzeug*. Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Diplomarbeit, 2003
- [Abl07] ABLASSMEIER, M.: Multimodale Anzeige- und Bedienkonzepte und Trends bei Head-Up Displays. In: *Kooperationsforum Fahrerassistenzsysteme, Konzepte - Sicherheit - Akzeptanz, Aschaffenburg*, Bayern Innovativ, Nürnberg, 2007. – 29.03.2007
- [ALK01] AKYOL, S. ; LIBUDA, L. ; KRAISS, K.-F. ; (HRSG.), Thomas Jürgensohn; Klaus-Peter T. (Hrsg.). *Kraftfahrzeugführung*. 2001
- [All03] ALLIANCE OF AUTOMOBILE MANUFACTURERS (AAM). *Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems*. Alliance of Automobile Manufacturers, Washington, D.C. 2003
- [Alt04] ALTHOFF, F.: *Ein generischer Ansatz zur Integration multimodaler Benutzereingaben*, Technische Universität München, Fakultät für Elektro- und Informationstechnik, Dissertation, 2004
- [AMG<sup>+</sup>05] ABLASSMEIER, M. ; MCGLAUN, G. ; GAST, J. ; POITSCHKE, T. ; RIGOLL, G.: A Robust, Context-adaptive and Multimodal Search Engine for Efficient Information Retrieval in Car Environment. In: SALVENDY, G. (Hrsg.): *Proc. of HCI 2005, 11th Intern. Conf. on Human-Computer Interaction, Las Vegas, USA*, Lawrence Erlbaum Associates, Inc., 2005. – 22.-27.10.2005, CD-ROM
- [AMR05] ABLASSMEIER, M. ; MCGLAUN, G. ; RIGOLL, G.: Evaluating the Potential of Head-Up Displays for a Multimodal Interaction Concept in the Automotive Environment. In: ZINN, Dale et a. (Hrsg.): *Proc. of WMSCI 2005, The 9th World Multi-Conference on Systemics, Cybernetics and Informatics, Orlando, USA* Bd. X, IIS, 2005. – 10.-13.10.2005, auch CD-ROM, S. 238–24
- [ANR04] ABLASSMEIER, M. ; NIEDERMAIER, B. ; RIGOLL, G.: A New Approach of Using Network Dialog Structures in Cars. In: CALLAOS, N. (Hrsg.) ; LESSO, W. (Hrsg.)

- ; SANCHEZ, B. (Hrsg.): *Proc. of 8th World Multiconference on Systemics, Cybernetics, and Informatics SCI 2004, Orlando, Florida, USA* Bd. XI, IIS Orlando, 2004. – 18.-21.07.2004, auch CD-ROM, S. 227–232
- [APR06] 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.-27.04.2006 2006, S. 1613–1618
- [APRR07] ABLASSMEIER, M. ; POITSCHKE, T. ; REIFINGER, S. ; RIGOLL, G.: Context-Aware Information Agents for the Automotive Domain Using Bayesian Networks. In: SMITH, M.J. (Hrsg.) ; SALVENDY, G. (Hrsg.): *2007, Beijing, China* Bd. LNCS 4557, Springer, Heidelberg, 2007. – July, 22-27, 2007, Proc. Part I, S. 561–570
- [APW+07] ABLASSMEIER, M. ; POITSCHKE, T. ; WALLHOFF, F. ; BENGLER, K. ; RIGOLL, G.: Eye Gaze Studies comparing Head-Up and Head-Down Displays in Vehicles. In: *2007, Beijing, China*, IEEE, 2007. – July 2-5, 2007, CD ROM, S. 2250–2252
- [AR06] ABLASSMEIER, M. ; RIGOLL, G.: A context-adaptive search engine concept and multimodal input strategies for automotive environments. In: *Proc. Workshop on Effective Multimodal Dialogue Interfaces, in conjunction with IUI 06 International Conference on Intelligent User Interfaces, Sydney, Australia*, ACM Press, 29.01.2006 2006, S. 4–10
- [AR07] ABLASSMEIER, M. ; RIGOLL, G. ; HÜSEYIN, Abut (Hrsg.): *Advances for In-Vehicle and Mobile Systems: Challenges for International Standards*. Springer Science-Business, 2007. – 223–235 S. – Chapter 19
- [Att08] ATTRAKDIFF, USER INTERFACE DESIGN GMBH: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.uidesign.de>
- [Aud08] AUDI AG: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.audi.de>
- [Azu97] AZUMA, R. T.: A Survey of Augmented Reality. In: *Presence: Teleoperators and Virtual Environments* 6 (1997), Nr. 4, S. 355–385
- [BAK06] BROY, V. ; ALTHOFF, F. ; KLINKER, G.: iFlip: A Metaphor for In-vehicle Information Systems. In: *AVI'06, May 23-26, 2006, Venezia Italy*, 2006
- [Bal98] BALZERT, H.: *Lehrbuch der Software-Technik. Software-Management, Software-Qualitätssicherung, Unternehmensmodellierung*. Berlin, Heidelberg : Addison-Wesley, 1998. – 129–133 S. – ISBN ISBN 3–8274–0065–1
- [BB99] BEDERSON, B. B. ; BOLTMAN, A.: Does Animation Help Users Build Mental Maps of Spatial Information? In: *INFOVIS '99: Proceedings of the 1999 IEEE Symposium on Information Visualization*. Washington, DC, USA : IEEE Computer Society, 1999. – ISBN 0–7695–0431–0, S. 28
- [Bed00] BEDERSON, B. B.: Fisheye menus. In: *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2000. – ISBN 1–58113–212–3, S. 217–225

- [Ben01] BENGLER, K.: Aspekte der multimodalen Bedienung und Anzeige im Automobil. In: *Kraftfahrzeugführung*. Jürgensohn, Thomas and Timpe, Klaus-Peter, 2001
- [Ber70] BERNOTAT, R.: Anthropotechnik in der Fahrzeugführung. In: *Ergonomics* 13 (1970), S. 353–377
- [BF03] BEYMER, D. ; FLICKNER, M.: *Eye Gaze Tracking Using an Active Stereo Head*. IBM Almaden Research Center, USA, 2003
- [BGM04] BEDERSON, B. B. ; GROSJEAN, J. ; MEYER, J.: Toolkit Design for Interactive Structured Graphics. In: *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING* (2004), S. 535–546
- [BH74] BADDELEY, A. ; HITCH, G.: Working memory. In: BOWER, G. (Hrsg.): *Recent advances in learning and motivation* Bd. 8. New York : Academic Press, 1974
- [BK03] BECKER, J. ; KUROPKA, D.: Topic-based Vector Space Model. In: GRAPHICS, IEEE C. (Hrsg.) ; APPLICATIONS (Hrsg.): *Proceedings of the 6th International Conference on Business Information Systems* Bd. 4(11). Colorado Springs, 2003, S. 7–12
- [BM77] BOYER, R. S. ; MOORE, J. S.: A fast string searching algorithm. In: *Commun. ACM* 20 (1977), Nr. 10, S. 762–772
- [BMW06] BMW MAGAZIN ONLINE: *Alles im Blick. Mit dem Head-Up Display fährt es sich entspannter*. 3/2006
- [BMW08a] BMW GROUP: BMW auf der CeBit 2008: Weltpremiere für uneingeschränkte Internet-Nutzung im Auto. (04.03.2008). – BMW Pressemitteilung
- [BMW08b] BMW GROUP: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.bmw.de>
- [Bol80] BOLT, R. A.: “Put-That-There”: Voice and Gesture at the Graphics Interface. In: *International Conference on Computer Graphics and Interactive Techniques*, 1980, S. 262–270
- [BP98] BRIN, S. ; PAGE, L.: The anatomy of a large-scale hypertextual Web search engine. In: *Computer Networks and ISDN Systems* 30 (1998), Nr. 1–7, S. 107–117
- [BP99] BILLSUS, D. ; PAZZANI, M. J.: A personal news agent that talks, learns and explains. In: *Proceedings of the Third International Conference on Autonomous Agents (Agents’99)*, 1999, S. 268–275
- [BPK06] BAUMANN, M. R. ; PETZOLD, T. ; KREMS, J.: Situation Awareness beim Autofahren als Verstehensprozess. In: *MMI-Interaktiv*. 2006 ( Nr. 11), S. 43–45
- [Bro58] BROADBENT, D. E.: *Perception and Communication*. Pergamon Press, 1958
- [Bro07] BROY, V. *Benutzerzentrierte, graphische Interaktionsmetaphern für Fahrerinformationssysteme*. 2007

- [BRT<sup>+</sup>04] BUBB, H. ; RIGOLL, G. ; TILLMANN, H. ; KLINKER, G. ; ABLASSMEIER, Draxler C. ; HAGEN, R. ; SANDOR, Ch.: Abschlussbericht zum Kooperationsvorhaben Mensch-Maschine-Interaktion im Kraftfahrzeug, 1. Phase, Grundlegende Vorgaben, Konzepte Visionen / TUMMIC. München, 2004. – Forschungsbericht
- [Bub76] BUBB, H.: *Untersuchung über die Anzeige des Bremsweges im Kraftfahrzeug*, Institut für Ergonomie der Technischen Universität München (Prof. Dr. H. Schmidtke, Dissertation, 1976
- [Bub08a] BUBB, H.: *Produktergonomie – Skriptum zur Vorlesung*. München : Lehrstuhl für Ergonomie, 2008
- [Bub08b] BUBB, H.: *Softwareergonomie – Skriptum zur Vorlesung*. München : Lehrstuhl für Ergonomie, 2008
- [BY89] BAEZA-YATES, R. A.: Algorithms for string searching. In: *SIGIR Forum* 23 (1989), Nr. 3-4, S. 34–58. – ISSN 0163–5840
- [BZW98] BRENNER, W. ; ZARNEKOW, R. ; WITTIG, H.: *Intelligente Software-Agenten – Grundlagen und Anwendungen*. Berlin : Springer-Verlag, 1998
- [CH98] CAGLAYAN, A. K. ; HARRISON, C. G.: *Intelligente Software-Agenten*. München, Wien : Carl Hanser Verlag, 1998
- [Cha92] CHARWAT, H. J.: *Lexikon der Mensch-Maschine-Kommunikation*. Oldenbourg, 1992. – ISBN 3–486–20904–3
- [CMBM03] CORRADINI, A. ; MEHTA, M. ; BERNSEN, N. ; MARTIN, J.-C.: Multimodal input fusion in human-computer interaction on the example of the on-going nice project. In: *Proceedings of the NATO-ASI conference on Data Fusion for Situation Monitoring, Incident Detection, Alert and Response Management*. Yerevan (Armenia), 2003
- [Cor08] CORPORATION, Continental: *Gute Fahrt: Wichtige Infos direkt im Sichtfeld: Das Head-up-Display zeigt fahrrelevante Informationen auf einen Blick*. Hannover, 03/2008
- [Dah06] DAHM, M.: *Grundlagen der Mensch-Maschine-Interaktion*. Pearson Studium, 2006
- [Dai04] Daimler AG - Research and Technology: *Lane Change Test 1.2 - User Guide*. Juli 2004
- [Dai08] DAIMLER AG: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.daimler.de>
- [Dep98] DEPARTMENT OF TRANSPORTATION: *Human Factors Design Guidelines* / U.S. Department of Transportation. Washington D.C., 1998. – Forschungsbericht
- [Dep99] DEPARTMENT OF DEFENCE: *Design Criteria Standard - Human Engineering* / U.S. Army Aviation and Missile Command. Redstone, 1999 ( MIL-STD-1472F). – Technischer Bericht

- [DH98] DAVISON, B. D. ; HIRSH, H.: Predicting Sequences of User Actions. In: *Predicting the Future: AI Approaches to Time-Series Problems*, 1998, S. 5–12
- [DMKSH93] DIETERICH, H. ; MALINOWSKI, U. ; KÜHME, T. ; SCHNEIDER-HUFSCHMIDT, M. *State of the Art in Adaptive User Interfaces*. 1993
- [DPU08] DORNHÖFER, S. M. ; PANNASCH, S. ; UNEMA, P. J. A.: *Augenbewegungen und deren Registrierungsmethoden*. Angewandte Kognitionsforschung, Institut für Psychologie, TU Dresden, 2008
- [EET05] EETIMES GERMANY: *Head-Up Displays im Auto: Augen bleiben auf der Straße*. Hannover, 17.02.2005
- [End88] ENDSLEY, M.R.: Situation Awareness Global Assessment Technique (SAGAT). In: *Proceedings of the National Aerospace and Electronics Conference (NAECON)* (1988), S. 789–795
- [Erk93] ERKE, H.: Vorgaben an die Verkehrsplanung. Anforderung an den Menschen aus Sicht der ökologischen Psychologie. In: LANG, E. (Hrsg.) ; ARNOLD, K. (Hrsg.): *Der Mensch im Straßenverkehr*. Stuttgart : Enke, 1993
- [Eur99] EUROPÄISCHE KOMMISSION. *Commission Recommendation of 21 December 1999 on Safe and Efficient In-Vehicle Information and Communication Systems: A European Statement of Principles on Human machine Interfaces*. Dokument Nr. C(1999) 4786. 1999
- [Eur06] EUROPÄISCHE KOMMISSION. *Empfehlung der Kommission: Neufassung des europäischen Grundsatzkatalogs für die Gestaltung der Mensch Maschine Schnittstelle für bordeigene Kommunikations- und Informationssysteme*. Brüssel. 2006
- [Eye08] EYELINK REMOTE - REMOTE EYETRACKER SR RESEARCH: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.eyelinkinfo.com>
- [Far96] FARACH, M.: Optimal Suffix Tree Construction with Large Alphabets. 1996 ( 96-48). – Forschungsbericht
- [FB89] FRESE, M. ; BRODBECK, F.: Computer in Büro und Verwaltung. In: *Psychologisches Wissen für die Praxis*. Heidelberg : Springer, 1989
- [FER08] FERMUS - MULTIMODALE FEHLERROBUSTE SPRACHDIALOGE: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.fermus.de/>
- [FM04] FINDLATER, L. ; MCGRENERE, J.: A comparison of static, adaptive, and adaptable menus. In: *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, 2004. – ISBN 1-58113-702-8, S. 89–96
- [For08] FORD AG: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.ford.de>
- [FWC84] FOLEY, J. D. ; WALLACE, V. L. ; CHAN, P.: The human factors of computer graphics interaction techniques., 1984, S. 13–48

- [GCH<sup>+</sup>05] GAJOS, K. ; CHRISTIANSON, D. B. ; HOFFMANN, R. ; SHAKED, T. ; HENNING, K. ; LONG, J. J. ; WELD, D. S.: Fast and Robust Interface Generation for Ubiquitous Applications. In: *UbiComp* Bd. 3660, 2005, S. 37–55
- [GCTW06] GAJOS, K. Z. ; CZERWINSKI, M. ; TAN, D. S. ; WELD, D. S.: Exploring the design space for adaptive graphical user interfaces. In: *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, 2006, S. 201–208
- [Gei90] GEISER, G. (Hrsg.): *Mensch-Maschine-Kommunikation*. Oldenbourg-Verlag, München, 1990
- [Gei03] GEIGER, Thomas: Head-Up Display: Alles im Blick. Hannover : Spiegel Online, 13. Mai 2003. – Forschungsbericht
- [GLPS95] GREEN, P. ; LEVISON, W. ; PAELKE, G. ; SERAFIN, C.: Preliminary Human Factors Design Guidelines for Driver Information Systems / Knowledge Systems Laboratory, Stanford University. 1995 ( KSL-93-04). – Technischer Bericht
- [Gol02] GOLDSTEIN, E. B.: *Wahrnehmungspsychologie*. Springer, 2002
- [Grü05] 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, Dissertation, 2005
- [GS95] GISH, K. W. ; STAPLIN, L.: Human Factors of Using Head Up Displays in Automobiles: A Review of the Literature. Washington : U.S. Department of Transportation, 1995
- [GSB04] GOOD, L. ; STEFIK, M. ; BEDERSON, B.: *A Comparison of Zoomable User Interfaces and Folders For Grouping Visual Objects*. Palo Alto Research Center / Human Computer Interaction Lab, University of Maryland, 2004
- [GST06] GEVEN, A. ; SEFELIN, R. ; TSCHELIGI, M.: Depth and Breadth away from the Desktop - the Optimal Information Hierarchy for Mobile Use. In: *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. Helsinki, Finnland, September 2006
- [Göt94] GÖTZE, R.: *Dialogmodellierung für multimediale Benutzerschnittstellen.*, Universität Oldenburg, Fachbereich Informatik, Diss., 1994
- [GT00] GOKER, M. ; THOMPSON, C.: The Adaptive Place Advisor: A Conversational Recommendation System. In: *Proceedings of the 8th German Workshop on Case Based Reasoning*. Lammerbuckel, Germany, March 2000
- [GW85] GREENBERG, S. ; WITTEN, I. H.: Adaptive Personalized Interfaces – A Question of Viability. In: *Behaviour and Information Technology* 4 (1985), S. 31–45
- [Har84] HAREL, D.: Statecharts: A visual approach to complex systems / Department of Applied Mathematics, The Weizmann Institute of Science. 1984 ( CS84-05). – Forschungsbericht

- [Har87] HAREL, D.: Statecharts: A Visual Formalism for Complex Systems. In: *Science of Computer Programming* 8 (1987), S. 231–274
- [Hat00] HATSCHER, M.: *Joy of use - Determinanten der Freude bei der Softwarenutzung*, Universität Osnabrück, Fachbereich Psychologie und Gesundheitswissenschaften, Fachgebiet Arbeits- und Organisationspsychologie, Diplomarbeit, 2000
- [HBB01] HASSENZAHL, M. ; BEU, A. ; BURMESTER, M.: Engineering Joy. In: *IEEE Software* (2001), Nr. 1, S. 70–76
- [HBC<sup>+</sup>96] HEWETT, T. (Hrsg.) ; BAECKER, R. (Hrsg.) ; CARD, S. (Hrsg.) ; CAREY, T. (Hrsg.) ; GASEN, J. (Hrsg.) ; MANTEI, M. (Hrsg.) ; PERLMAN, G. (Hrsg.) ; STRONG, G. (Hrsg.) ; VERPLANK, W. (Hrsg.): *Curricula for Human-Computer Interaction*. ACM Special Interest Group on Computer-Human Interaction, Curriculum Development Group, 1996
- [HBK03] HASSENZAHL, M. ; BURMESTER, M. ; KOLLER, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: *Mensch & Computer*, 2003
- [HBL01] HOFMANN, M. ; BENGLER, K. ; LANG, M.: Ein Assistenzsystem zur fahrer- und situationsadaptiven Prädiktion potentieller Zielorte für eine robuste Interaktion mit sprachgesteuerten Navigationssystemen. In: *Tagungsband „Elektronik im Fahrzeug“, Baden-Baden*, VDI-Bericht 1646, 2001
- [HD80] HALL, P. A. V. ; DOWLING, G. R.: Approximate String Matching. In: *ACM Comput. Surv.* 12 (1980), Nr. 4, S. 381–402. – ISSN 0360–0300
- [Her94] HERCZEG, M.: *Software-Ergonomie*. Addison-Wesley, 1994
- [HG98] HOOEY, B.L. ; GORE, B.F.: Advanced traveler information systems and commercial vehicle operations components of the intelligent transportation systems: head-up displays and driver attention for navigation information. (FHWA-RD-96-153) / US Department of Transportation Federal Highway Administration. 1998. – Forschungsbericht
- [HHH06] HOF, A. ; HAGEN, E. ; HUBER, A.: Adaptive Help for Speech Dialogue Systems Based on Learning and Forgetting of Speech Commands. In: *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue*. Sydney, Australia : Association for Computational Linguistics, July 2006, S. 1–8
- [HL01] HOFMANN, M. ; LANG, M.: User Appropriate Plan Recognition for Adaptive Interfaces. In: *HCI 2001, New Orleans, Louisiana, USA* Bd. 1, Lawrence Erlbaum Ass., NJ, 2001. – „Usability Evaluation and Interface Design“, S. 1130–1134
- [HMS96] HARTUNG, K. ; MÜNCH, S. ; SCHOMAKER, L.: MIAMI: Software Architecture, Deliverable Report 4. Report of ESPRIT III: Basic Research Project 8579, Multimodal Interface for Advanced Multimedia Interfaces (MIAMI). 1996. – Forschungsbericht
- [HMU01] HOPCROFT, J. E. ; MOTWANI, R. ; ULLMAN, J. D.: *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley, 2001

- [HNL06] HENNESSEY, C. ; NOUREDDIN, B. ; LAWRENCE, P.: *A Single Camera Eye-Gaze Tracking System with Free Head Motion*. University of British Columbia, 2006
- [Hum08] HUMAN-COMPUTER INTERACTION LAB, UNIVERSITY OF MARYLAND: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.cs.umd.edu/hcil/>
- [HW02] HORREY, W. J. ; WICKENS, C. D.: *Driving and Side Task Performance: The Effects of Display Clutter, Separation, and Modality*. Washington : University of Illinois, Aviation Human Factors Division, 2002, S. 502–520
- [IDZ04] INGMARSSON, M. ; DINKA, D. ; ZHAI, S.: TNT: a numeric keypad based text input method. In: *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2004. – ISBN 1–58113–702–8, S. 639–646
- [ISO98] ISO 9241-11. *Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten - Teil 11: Anforderungen an die Gebrauchstauglichkeit - Leitsätze*. 1998
- [ISO02] ISO 3888-1. *Passenger cars - Test track for a severe lane-change manoeuvre*. Deutsches Institut für Normung. Nov. 2002
- [ISO03a] ISO 15005. *Ergonomische Aspekte von Fahrer- und Assistenzsystemen - Anforderungen und Bewertungsmethoden der visuellen Informationsdarstellung im Fahrzeug*. Deutsches Institut für Normung e.V. 2003
- [ISO03b] ISO 15008. *Ergonomische Aspekte von Fahrerinformations- und Assistenzsystemen - Anforderungen und Bewertungsmethoden der visuellen Informationsdarstellung im Fahrzeug*. Deutsches Institut für Normung e.V. 2003
- [ISO06] ISO 9241-110. *Ergonomie der Mensch-System-Interaktion – Teil 110: Grundsätze der Dialoggestaltung (Ersatz für DIN EN ISO 9241-10:1996- 07)*. 8 2006
- [Jap04] JAPAN AUTOMOBILE MANUFACTURES ASSOCIATION (JAMA): *Guidelines for In-Vehicle Display Systems*. Tokyo, 2004 ( Version 3). – Forschungsbericht
- [JCM+97] JOHNSTON, M.I ; COHEN, P. R. ; MCGEE, D. ; OVIATT, S. L. ; PITTMAN, J. A. ; SMITH, I.: Unification-based multimodal integration. In: *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*. Morristown, NJ, USA : Association for Computational Linguistics, 1997, S. 281–288
- [Jen01] JENSEN, F. V.: *Bayesian Networks and Decision Graphs*. New York : Springer Verlag Inc., 2001
- [JJ95] JONES, T.O. ; JR., W.E. S.: *Why satisfied customers defect*. Harvard Business Review, 1995. – pp. 88–99 S
- [JOKG05] JAHN, G. ; OEHME, A. ; KREMS, J.F. ; GELAU, C.: Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. In: *Transportation Research Part F 8*, 2005, S. 255–275

- [JRV08] JOOS, M. ; RÖTTING, M. ; VELICHKOVSKY, B. M.: *Bewegungen des menschlichen Auges: Fakten, Methoden und innovative Anwendungen*. Angewandte Kognitionsforschung, Institut für Psychologie, TU Dresden, 2008
- [Kau04] KAUFMANN, J.: Head-Up-Display: Neue Technik für mehr Verkehrssicherheit. ZDNet, 20.10.2004. – Forschungsbericht
- [KGT06] KNOPFEL, A. ; GRONE, B. ; TABELING, P.: *Fundamental Modeling Concepts: Effective Communication of IT Systems*. John Wiley & Sons, 2006. – ISBN 047002710X
- [Kie98] KIEFER, R. J.: Defining the „HUD Benefit Time Window“. In: ET AL., A. G. G. (Hrsg.): *Visions in Vehicles - VII* Bd. Report No. AHFD-02-13/GM-02-2. Amsterdam : Elsevier, 1998, S. 133–142
- [Kie00] KIEFER, R. J.: Older Drivers' Pedestrian Detection Times Surrounding Head-Up versus Head-Down Speedometer Glances. In: ET AL., A. G. G. (Hrsg.): *Visions in Vehicles - VII* Bd. Report No. AHFD-02-13/GM-02-2. Amsterdam : Elsevier, 2000, S. 111–118
- [KMP77] KNUTH, D. E. ; MORRIS, J. H. ; PRATT, V. R.: Fast Pattern Matching in Strings. In: *SIAM Journal on Computing* 6 (1977), Nr. 2, S. 323–350
- [Kön06] KÖNIG, W. A.: *Referenzmodell und Machbarkeitsstudie für ein neues Zoomable User Interface Paradigma*. Universität Konstanz, FB Informatik und Informatikwissenschaft, Masterarbeit, 2006
- [Kor78] KORNUBER, H. H.: Blickmotorik. In: GAUER, Kramer K. Jung R. (Hrsg.): *Physiologie des Menschen. Band 13: Sehen (Sinnesphysiologie III)* Bd. 13, Urban & Schwarzenbeck , München, Wien, Baltimore, 1978, S. 357–426
- [Kra88] KRASNER, S. T.: A Cookbook for using the Model- View-Controller Interface Paradigm in Smalltalk-80. In: *Journal of Object Oriented Programming* 8/9 (1988), S. 26–49
- [KRR<sup>+</sup>06] KARLSON, A. K. ; ROBERTSON, G. ; ROBBINS, D. C. ; CZERWINSKI, M. ; SMITH, G.: FaThumb: A Facet-based Interface for Mobile Search. In: *ACM CHI 2006*. Montreal, Quebec, Canada, 2006
- [Kul00] KULES, B. *User Modeling for Adaptive and Adaptable Software Systems*. <http://www.otal.umd.edu/UUGuide/wmk/> (letzter Zugriff: 30.08.2007). 2000
- [Lan06] LANGE, M.; Bubb H.: Vergleichbarkeit von Usability Lab und Realversuch zur Bestimmung der ergonomischen Güte und der Ablenkungswirkung von Nebenaufgaben im Kfz. In: FÜR ARBEITSWISSENSCHAFT E.V. DORTMUND, Gesellschaft (Hrsg.): *Innovationen für Arbeit und Organisation. 52. Frühjahrskongress der Gesellschaft für Arbeitswissenschaft*. Stuttgart : GFA-Press, 2006, S. 692 – 697
- [Lan08] LANGE, C.: *Wirkung von Fahrerassistenz auf der Führungsebene in Abhängigkeit der Modalität und des Automatisierungsgrades*. München, Technische Universität München, Dissertation, 2008

- [Lex08] LEXUS: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.lexus.com>
- [LF98] LANGLEY, Pat ; FEHLING, Michael. *The Experimental Study of Adaptive User Interfaces*. 1998
- [Lip04] LIPPE, W.-M.: Einführung in SoftComputing (Skript zur Vorlesung) / Institut für Informatik, Westfälische Wilhelms Universität Münster. 2004. – Forschungsbericht
- [LK03] LIBUDA, L. ; KRAISS, K.-F.: Dialogassistenz im Kraftfahrzeug. In: *DGLR-Bericht Volume 2003-04*. Neubiberg, 2003, S. 255–270
- [LN81] LINDSAY, P. ; NORMAN, D.: *Einführung in die Psychologie*. Berlin : Springer, 1981
- [Loe89] LOEFFLER, L.: *Adaptierbare und adaptive Benutzerschnittstellen*. Institut für Werkzeugmaschinen und Betriebstechnik, Universität Karlsruhe, 1989
- [Lug01] LUGER, G. F.: *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Boston, MA, USA : Addison-Wesley Longman Publishing Co., Inc., 2001. – ISBN 0201648660
- [MAB08] MILICIC, N. ; ABLASSMEIER, M. ; BENGLER, K.: Das Head-Up Display im Fahrzeug - Potenzial zukünftiger Nutzung. In: *Produkt- und Produktions-Ergonomie - Aufgabe für Entwickler und Planer, Bericht zum 54. Arbeitswissenschaftlichen Kongress vom 9. - 11.4.2008 an der Technischen Universität München*, Gesellschaft für Arbeitswissenschaft GfA-Press, Dortmund, 2008. – ISBN 978-3-936804-06-5
- [MALR04] MCGLAUN, G. ; ALTHOFF, F. ; LANG, M. ; RIGOLL, G.: Development of a Generic Multimodal Framework for Handling Error Patterns during Human-Machine Interaction. In: *SCI 2004, 8th World Multi-Conference on Systems, Cybernetics, and Informatics, Orlando, FL, USA, 2004*
- [MASL02] MCGLAUN, G. ; ALTHOFF, F. ; SCHULLER, B. ; LANG, M.: A new technique for adjusting distraction moments in multi-tasking non-field usability tests. In: TERVEEN (Hrsg.) ; WIXON (Hrsg.) ; COMSTOCK (Hrsg.) ; SASSE (Hrsg.): *Proc. of Intern. Conf. on Human Factors in Computing Systems CHI 02, Minneapolis, USA,*, ACM SIGCHI, New York, 2002. – 20.-25.04.2002, S. 666–667
- [Mat04] MATHES, A. *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*. December 2004
- [May92] MAYHEW, D. J.: *Principles and Guidelines in Software User Interface Design*. Englewood Cliffs : Prentice Hall, 1992
- [MBB02] MCGRENERE, J. ; BAECKER, R. ; BOOTH, K. *An Evaluation of a Multiple Interface Design Solution for Bloated Software*. 2002
- [McT04] MCTEAR, M. F.: *Spoken Dialogue Technology: Toward the Conversational User Interface*. Springer Verlag, London, 2004. – ISBN 1–85233–672–2
- [Met75] METZGER, W.: *Gesetze des Sehens*. Waldemar Kramer, 1975

- [Mic85] MICHON, J. A.: A critical view of driver behavior models: what do we know, what should we do? In: EVANS, L. (Hrsg.) ; SCHWING, R. C. (Hrsg.): *Human behavior and traffic safety*. New York : Plenum Press, 1985
- [Mil81] MILLER, D. P.: The Depth/Breadth Tradeoff in Hierarchical Computer Menu. In: *HFES 25th Annual Meeting*, 1981
- [Mis08] MISTER WONG: *Homepage*. Letzter Zugriff: 20.03.2008. – [hwww.mister-wong.de](http://hwww.mister-wong.de)
- [MK94] MILGRAM, P. ; KISHINO, F.: A Taxonomy of Mixed Reality Visual Displays. In: *IEICE Trans. Information Systems* E77-D (1994), dec, Nr. 12, S. 1321–1329
- [Mül03] MÜLLER, R.: *Einsatz Bayes'scher Netze zur Interpretation von Benutzeraktionen*. Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Diplomarbeit, 2003
- [MS89] MITCHELL, J. ; SHNEIDERMAN, B.: Dynamic versus static menus: an exploratory comparison. In: *SIGCHI Bull.* 20 (1989), Nr. 4, S. 33–37. – ISSN 0736–6906
- [MS99] MAYBURY, M. T. ; STOCK, O.: Multimedia Communication, including Text. In: HOVY, E. (Hrsg.) ; IDE, N. (Hrsg.) ; FREDERKING, R. (Hrsg.) ; MARIANI, J. (Hrsg.) ; ZAMPOLLI, A. (Hrsg.): *Multilingual Information Management: Current Levels and Future Abilities*. A study commissioned by the US National Science Foundation and also delivered to European Commission Language Engineering Office and the US Defense Advanced Research Projects Agency, 1999
- [MW00] MARTENS, M. H. ; VAN WINSUM, W.: Measuring distraction: The Peripheral Detection Task. Netherlands : TNO Human Factors, 2000
- [Nav01] NAVARRO, G.: A guided tour to approximate string matching. In: *ACM Computing Surveys* 33 (2001), Nr. 1, S. 31–88
- [NC93] NIGAY, L. ; COUTAZ, J.: A design space for multimodal systems: Concurrent processing and Data fusion. In: PRESS, ACM (Hrsg.): *Proceedings of InterCHI'93*. Amsterdam : Ashlund, S. and Mullet, K. and Henderson, A. and Hollnagel, E. and White, T., 1993, S. 172–178
- [Neu01] NEUSS, R.: *Usability Engineering als Ansatz zum Multimodalen Mensch-Maschine-Dialog*. München, Technische Universität München, Dissertation, 2001
- [Nie93] NIELSEN, J.: *Usability Engineering*. Academic Press, Inc., 1993. – ISBN 0–12–518405–0
- [Nie02] NIEDERMAIER, F. B.: *Entwicklung und Bewertung eines Rapid-Prototyping Ansatzes zur multimodalen Mensch-Maschine-Interaktion im Fahrzeug*, Technische Universität München, Dissertation, 2002
- [Nor91] NORMAN, K. L.: *The Psychology of Menu Selection: Designing Cognitive Control at the Human Computer Interface*. Ablex Publishing Corporation, 1991
- [Nor02] NORMAN, D. A.: *The Design of Everyday Things*. Basic Books, September 2002. – ISBN 0465067107

- [OCL04] OVIATT, S. ; CAULSTON, R. ; LUNSFORD, R.: When Do We Interact Multimodally? Cognitive Load and Multimodal Communication Patterns. (2004)
- [OCW<sup>+</sup>00] OVIATT, S. ; COHEN, P. ; WU, L. ; VERGO, J. ; DUNCAN, L. ; SUHM, B. ; BERS, J. ; HOLZMAN, T. ; WINOGRAD, T. ; LANDAY, J. ; LARSON, J. ; FERRO, D.: Designing the User Interface for Multimodal Speech and Pen-based Gesture Applications: State-of-the-Art Systems and Future Research Directions. In: *Human Computer Interaction* (2000), Nr. 15(4), S. 263–322
- [Oka96] OKABAYASHI, S.: *Visual Optics of Head-up Displays (HUDs) in Automotive Applications*. Oldenbourg : Taylor & Francis, 1996, 1996. – ISBN 2884490337, 9782884490337
- [OMY02] OHNO, T. ; MUKAWA, N. ; YOSHIKAWA, A.: *A Gaze Tracking System for Everyday Gaze Interaction*. NTT Communication Science Laboratories, NTT Corporation, 2002
- [Ovi99] OVIATT, S.: Ten Myths of Multimodal Interaction. In: *Communications of the ACM* 42 11 (1999), S. 74–81
- [Ovi02] OVIATT, S. ; JACKO, J. (Hrsg.) ; SEARS, A. (Hrsg.): *Handbook of Human-Computer Interaction*. New Jersey : Lawrence Erlbaum, 2002
- [Ovi03] OVIATT, S.: Advances in Robust Multimodal Interface Design. In: *IEEE Comput. Graph. Appl.* 23 (2003), Nr. 5, S. 62–68. – ISSN 0272–1716
- [PAR<sup>+</sup>08] POITSCHKE, T. ; ABLASSMEIER, M. ; RIGOLL, G. ; BARDINS, S. ; KOHLBECHER, S. ; SCHNEIDER, E.: Contact-analog information representation in an automotive head-up display. In: *Proc. of the 2008 symposium on Eye tracking research & application ETRA 08, Savannah, Georgia, USA*, ACM Press, NY, 2008. – March 2008, S. 119–122
- [PARR07] POITSCHKE, T. ; ABLASSMEIER, M. ; REIFINGER, S. ; RIGOLL, G.: A Multifunctional VR-Simulator Platform for the Evaluation of Automotive User Interfaces. In: JACKO, J.A. (Hrsg.): *2007, Beijing, China* Bd. LNCS 4553, Springer, Heidelberg, 2007. – July, 22-27, 2007, Proc. Part IV, Human-Computer-Interaction, S. 1120–1129
- [PBMW98] PAGE, Lawrence ; BRIN, Sergey ; MOTWANI, Rajeev ; WINOGRAD, Terry: The PageRank Citation Ranking: Bringing Order to the Web / Stanford Digital Library Technologies Project. 1998. – Forschungsbericht
- [Pfa83] User Interface Management Systems, Proceedings. In: PFAFF, G. E. (Hrsg.): *Workshop on User Interface Management Systems (1. - 3.11.1983)*. Seeheim : Springer Verlag, 1983
- [PR05] PETKO, D. ; REUSSER, K.: Das Potenzial interaktiver Lernressourcen zur Förderung von Lernprozessen. (2005)
- [Qui08] QUIETLY SCHEMING: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.quietlyscheming.com>

- [Ras83] 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
- [Ras04] RASSL, R.: *Ablenkungswirkung tertiärer Aufgaben im Pkw - Systemergonomische Analyse und Prognose*, Technische Universität München, Diplomarbeit, 2004
- [Ras05] RASKIN, J.: *The humane interface*. Boston [u.a.] : Addison-Wesley, 2005. – ISBN 0–201–37937–6
- [Rau04] RAUNER, M.: Quasselnde Engel. In: *Die Zeit* 18 (22.04.2004)
- [Rea88] REASON, J.: Frameworks of human performance and error: a consumers guide. In: TAYLOR (Hrsg.) ; FRANCIS (Hrsg.): *Taks, errors and mental models*. London : I. P. Goodstein and H. Andersen and S. E. Olsen, 1988
- [Rea90] REASON, J.: *Human Error*. Cambridge University Press, 1990. – ISBN 0521314194
- [RFL99] ROGERS, S. ; FLECHTER, C.-N. ; LANGLEY, P.: An adaptive interactive agent for route advice. In: *Proceedings of the Third International Conference on Autonomous Agents (Agents'99)*, 1999, S. 198–205
- [RFT00] ROGERS, S. ; FIECHTER, C.-N. ; THOMPSON, C.: Adaptive User Interfaces for Automotive Environments. In: *Procs. IEEE Intelligent Vehicles Symposium 2000*. Detroit, USA, Oktober 2000, S. 662–667
- [Rig94] RIGOLL, G.: *Neuronale Netze: eine Einführung für Ingenieure*. Renningen-Malmsheim : expert-Verlag, 1994
- [RN03] RECARTE, M. A. ; NUNES, L. M.: Mental Workload while Driving: Effects on Visual Search, Discrimination, and Decision Making. In: *Journal of Experimental Psychology* Bd. 9. 2003, S. 119–137
- [SA90] SOJOURNER, R.J. ; ANTIN, J.F.: The effects of a simulated Head-up Display Speedometer on Perceptual Task Performance. In: FÜR ARBEITSWISSENSCHAFT E.V. DORTMUND, Gesellschaft (Hrsg.): *Human Factors*, Nr. 32, 1990
- [SAM+06] SCHULLER, B. ; ABLASSMEIER, M. ; MÜLLER, R. ; REIFINGER, S. ; POITSCHKE, T. ; RIGOLL, G.: Advanced Man-Machine Interaction. Springer Verlag Berlin Heidelberg New York, 2006. – ISBN: 3-540-30618-8, Kapitel Speech Communication and Multimodal Interfaces, S. 141–190
- [SAMR05] SCHEJA, Ch. ; ABLASSMEIER, M. ; MCGLAUN, G. ; RIGOLL, G.: Multimodal, Contextual Dialog Management for Natural Speech in Car Environments. In: *SDV Sprache und Datenverarbeitung, International Journal for Language Data Processing* 29 (2005), Nr. 2, S. 87–106. – Institut für Kommunikationswissenschaften, Bonn ISSN 0343-5202
- [Sch93] SCHMIDTKE, H.: *Ergonomie*. Hanser Verlag, München Wien, 1993
- [Sch03a] SCHMITT, S.: Augen geradeaus! heise online, 11/2003. – Forschungsbericht

- [Sch03b] SCHWEIGERT, M.: *Fahrerblickverhalten und Nebenaufgaben*, Technische Universität München, Dissertation, 2003
- [SD01] SCHATTENBERG, K. ; DEBUS, G.: Multimodale Anzeige- und Bedienkonzepte zur Steuerung technischer Systeme während der Fahrt im Kraftfahrzeug: Evaluationsbefunde zur Systemweiterentwicklung mit paralleler Sprachbedienung. In: *Kraftfahrzeugführung*. Jürgensohn, Thomas and Timpe, Klaus-Peter, 2001
- [SE08] SCHNEIDER, E. ; EGGERT, T.: *Methoden der Augenbewegungsmessung*. Neurologische Poliklinik, Klinikum der Universität München, 2008
- [Sei02] SEIFERT, K.: *Evaluation multimodaler Computer-Systeme in frühen Entwicklungsphasen*, Verkehrs- und Maschinensysteme (Fakultät V), TU Berlin, Dissertation, 2002
- [SGK<sup>+</sup>04] SCHINDHELM, R. ; GELAU, C. ; KEINATH, A. ; BENGLER, K. ; KUSSMANN, H. ; KOMPFER, P. ; AN M. MARTINETTO, C.P. C.: Report on the Review of the Available Guidelines and Standards. / AIDE - Adaptive Integrated Driver-Vehicle Interface. 2004 ( IST-1-507674-IP). – Forschungsbericht
- [Shn98] SHNEIDERMAN, B.: *Designing the user interface: Strategies for effective human-computer interaction*. 3rd ed. Addison-Wesley Publishing, 1998
- [Soc02] SOCIETY OF AUTOMOTIVE ENGINEERS (SAE): Recommended Practice Calculation of the Time to Complete In-Vehicle Navigation and Route Guidance Tasks / Society of Automotive Engineers. Warrendale, 2002 ( J2365). – Forschungsbericht
- [Soc04] SOCIETY OF AUTOMOTIVE ENGINEERS (SAE): Recommended Practice Navigation and Route Guidance Function Accessibility While Driving. Warrendale, 2004 ( J2364). – Forschungsbericht
- [SP04] SHNEIDERMAN, B. ; PLAISANT, C.: *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Pearson Addison Wesley, 2004. – ISBN 0321197860
- [SR08] SCHENK, J. ; RIGOLL, G.: *Mensch-Maschine-Kommunikation 1 – Skriptum zur Vorlesung*. München : Lehrstuhl für Mensch-Maschine-Kommunikation, 2008
- [SS94] SEARS, A. ; SHNEIDERMAN, B.: Split menus: effectively using selection frequency to organize menus. In: *ACM Transactions on Computer-Human Interaction* 1 (1994), Nr. 1, S. 27–51
- [Sty97] STYLES, E. A.: *The Psychology of Attention*. Hove : Psychology Press, 1997
- [SWM<sup>+</sup>08] SCHULLER, B. ; WÖLLMER, M. ; MOOSMAYR, T. ; RUSKE, G. ; RIGOLL, G.: Switching Linear Dynamic Models for Noise Robust In-Car Speech Recognition. In: RIGOLL, G. (Hrsg.): *Proc. 30th DAGM Symposium, Munich, Germany, June 2008* Bd. LNCS 5096, Springer, 2008, S. 244–253
- [SWS05] SNOEK, C. G. M. ; WORRING, M. ; SMEULDERS, A. W. M.: Early versus late fusion in semantic video analysis. In: *MULTIMEDIA '05: Proceedings of the 13th*

- annual ACM international conference on Multimedia*. New York, NY, USA : ACM, 2005. – ISBN 1-59593-044-2, S. 399–402
- [TL99] TALMI, K. ; LIU, J.: *Eye and Gaze Tracking for Visually Controlled Interactive Stereoscopic Displays*. 1999
- [TLKB06] TÖNNIS, M. ; LANGE, C. ; KLINKER, G. ; BUBB, H.: Transfer von Flugschlauchanzeigen in das Head-Up Display von Kraftfahrzeugen (Transfer of Flight Tunnel Visualizations into the Head-Up Display of Cars). In: *Proceedings of the VDI-Gemeinschaftstagung: Integrierte Sicherheit und Fahrerassistenzsysteme*, 2006
- [TMRK07a] TOTZKE, I. ; MÜHLBACHER, D. ; RAUCH, N. ; KRÜGER, H.-P.: The BMW SURF Project: A Contribution to the Research on Cognitive Vehicles. In: RÖTTING, M. (Hrsg.) ; WOZNY, G. (Hrsg.) ; KLOSTERMANN, A. (Hrsg.) ; HUSS, J. (Hrsg.): *Intelligent Vehicles Symposium, 2007 IEEE*. Istanbul : IEEE, 2007, S. 692 – 697
- [TMRK07b] TOTZKE, I. ; MÜHLBACHER, D. ; RAUCH, N. ; KRÜGER, H.-P.: Workload-Management im Fahrzeug: Braucht der Fahrer eine kontinuierliche Rückmeldung oder reicht eine kurze Vorinformation? In: RÖTTING, M. (Hrsg.) ; WOZNY, G. (Hrsg.) ; KLOSTERMANN, A. (Hrsg.) ; HUSS, J. (Hrsg.): *Prospektive Gestaltung von Mensch-Technik-Interaktion (VDI-Fortschritt-Bericht Reihe 22, Bd. 25)*, Düsseldorf: VDI-Verlag, 2007, S. 113–118
- [Toy08] TOYOTA: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.toyota.de>
- [TRK03] TOTZKE, I. ; RAUCH, N. ; KRÜGER, H.-P.: Kompetenzerwerb und Struktur von Menüsystemen im Fahrzeug: „Breiter ist besser?“. In: *VDI-Fortschritts-Berichte, Reihe 22 Band 16: Entwerfen uns gestalten*. VDI-Verlag, 2003
- [Tuf97] TUFANO, D. R.: Automotive HUDs: The overlooked safety issues. In: *Human Factors* Bd. 39, 1997, S. 303–311
- [Ukk95] UKKONEN, E.: On-Line Construction of Suffix Trees. In: *Algorithmica* 14 (1995), Nr. 3, S. 249–260
- [Vo98] VO, M. T.: *A framework and toolkit for the construction of multimodal learning interfaces*. Pittsburgh, PA, USA, Dissertation, 1998. – Chair-Alex Waibel
- [Vol08] VOLKSWAGEN AG: *Homepage*. Letzter Zugriff: 20.03.2008. – <http://www.volkswagen.de>
- [VT03] VOLLRATH, M. ; TOTZKE, I.: Möglichkeiten der Nutzung unterschiedlicher Ressourcen für die Fahrer-Fahrzeug-Interaktion. In: *Der Fahrer im 21. Jahrhundert*. VDI-Berichte Nr. 1768, 2003
- [WE92] WEINTRAUB, D. J. ; ENSING, M. J.: *Human factors issues in head-up display and design: the book of HUD : state-of-the-art report*. Crew System Ergonomics Information Analysis Center, 1992
- [Wic84] WICKENS, C. D.: *Engineering Psychology and Human Performance*. Columbus. Merrill, 1984

- [Wic96] WICKENS, C. D.: *Attention and Situation Awareness*, Univ. Illinois, Ph.d. thesis, 1996
- [Wik08] WIKIPEDIA, THE FREE ENCYCLOPEDIA: *Homepage (engl.)*. Letzter Zugriff: 20.03.2008. – <http://www.wikipedia.org>
- [Wil05] WILZ, M. *Aspekte der Kodierung phonetischer Ähnlichkeiten in deutschen Eigennamen*. 2005
- [WOC99] WU, L. ; OVIATT, S. ; COHEN, P.: Multimodal integration - A Statistical Review, 1999, S. 334–341
- [WS01] WANG, J.-G. ; SUNG, E.: *Gaze determination via images of irises*. Nanyang Technological University, Singapore, 2001
- [Wüs05] WÜST, C.: Irrfahrt durchs Untermenü. In: *Der Spiegel* Ausgabe 12 (2005), S. 151
- [ZD96] ZOBEL, J. ; DART, P. W.: Phonetic String Matching: Lessons from Information Retrieval. In: FREI, H.-P. (Hrsg.) ; HARMAN, D. (Hrsg.) ; SCHÄBLE, P. (Hrsg.) ; WILKINSON, R. (Hrsg.): *Proceedings of the 19th International Conference on Research and Development in Information Retrieval*. Zurich, Switzerland : ACM Press, 1996, S. 166–172
- [ZDZ97] ZIMMER, A. ; DAMEN-ZIMMER, K.: Situationsbezogene Sicherheitskenngrößen im Straßenverkehr. In: *Berichte der Bundesanstalt für Straßenwesen* Heft M 78 (1997)
- [Zel94] ZELL, A.: *Simulation Neuronaler Netze*. Oldenburg-Verlag, 1994
- [Zöl04] ZÖLLNER, R.: Was löst beim Fahrer den Wunsch aus, ein Fahrer-Assistenzsystem zu nutzen? Die Delphi-Expertenbefragung als Methode zur Bewertung der motivationalen Relevanz von Einflussgrößen. In: *VDI-Berichte 1864*, 2004, S. 49–61
- [ZM06] ZOBEL, J. ; MOFFAT, A.: Inverted files for text search engines. In: *ACM Comput. Surv.* 38 (2006), Nr. 2, S. 6. – ISSN 0360–0300

# Danksagung

Zunächst möchte ich mich ganz besonders bei Herrn Prof. Rigoll bedanken, der mir diese Arbeit und die Tätigkeit am Lehrstuhl ermöglicht hat. Herzlichen Dank für die vielen schönen Jahre am Lehrstuhl und das motivierende Arbeitsumfeld, welches mir äußerst viel Freiraum zur Selbstverwirklichung geboten hat. Sehr gefreut habe ich mich auch über die sofortige Bereiterklärung von Herrn Prof. Bubb, diese Arbeit als Zweitprüfer zu betreuen. Vielen Dank für die Unterstützung und die schöne Zeit der Zusammenarbeit im Projekt.

Als Weiteres gilt mein Dank Herrn Dr. Klaus Bengler von der BMW Forschung und Technik, der mir die Möglichkeit zur Mitarbeit im TUMMIC-Projekt eröffnet und mich in vielen fruchtbaren Diskussionen zu spannenden Forschungsfragen inspiriert hat. Gerade das interdisziplinäre Umfeld aus sechs Lehrstühlen war eine äußerst interessante und bereichernde Erfahrung für mich. Dabei möchte ich mich auch bei Herrn Johannes Fenk für die Koordination und die umfangreiche Unterstützung bedanken, die ganz wesentlich zum Erfolg des Projektes beigetragen hat. Mein großer Dank gilt auch den Doktoranden und Professoren des TUMMIC-Projektes für die konstruktive Zusammenarbeit und den lehrreichen Wissensaustausch. Außerdem möchte ich mich noch bei Herrn Dr. Bernhard Niedermaier bedanken, der während meiner Diplomarbeit meine Begeisterung für die Mensch-Maschine Interaktion im Automobil geweckt hat.

Bedanken möchte ich mich auch bei meinen Kollegen am Lehrstuhl, im Besonderen bei Herrn Tony Poitschke und Herrn Stefan Reifinger, mit denen ich gemeinsam viel bewegen konnte und die immer ein offenes Ohr für die zahlreichen Diskussionen hatten. Ein besonderer Dank gilt Herrn Gregor McGlaun, der mich für die Forschungsthemen am Lehrstuhl motiviert und mich bei sämtlichen Belangen unterstützt hat. Auch bei den Festangestellten des Lehrstuhls möchte ich mich für die Unterstützung und die guten Arbeitsbedingungen bedanken.

Schließlich möchte ich mich noch für die große Geduld und Ausdauer bei meiner Freundin Steffi, meinen Eltern und Geschwistern und meinen Freunden bedanken. Ohne Eure Unterstützung und Euer Verständnis wäre das Gelingen dieser Arbeit nicht möglich gewesen.

Vielen lieben Dank!

München, den 06.10.2008