

Inhaltsverzeichnis

1	Einleitung	1
2	Stand der Forschung	5
2.1	Akustische Szenenanalyse	5
2.2	<i>Middleware</i> und ambiante Intelligenz	8
2.3	Ambiente Kommunikation	9
3	Wissenschaftliche Ziele	11
3.1	Akustische Szenenanalyse	11
3.2	<i>Middleware</i> und ambiante Intelligenz	12
3.3	Ambiente Kommunikation	13
4	Akustische Szenenanalyse	15
4.1	Merkmalsextraktion	15
4.1.1	Störgeräuschunterdrückung	15
4.1.2	<i>Mel-Frequency Cepstral Coefficients</i>	16
4.1.3	<i>Maximum Autocorrelation Value</i>	17
4.2	Akustische Positionsschätzung	18
4.2.1	<i>Generalized Cross Correlation with Phase Transformation</i>	18
4.2.2	Akustische Strahlformung	19
4.2.3	Lokalisation mittels verteilter Mikrophongruppen	20
4.3	Segmentierung und Sprecheridentifikation	26
4.3.1	Sequentielle Sprecherwechseldetektion und Identifikation	27
4.3.2	Gemeinsame Sprecherwechseldetektion und Identifikation	33
4.3.3	Experimentelle Ergebnisse	39
4.4	Audio-visuelle Sprecherprotokollierung	48
4.4.1	System zur Gesichtsidentifikation	48
4.4.2	Gesichtsdetektion	48
4.4.3	Gesichtsidentifikation	51
4.4.4	Kamerasteuerung und Systemintegration	53
4.4.5	Integration der visuellen Information	55
4.4.6	Experimentelle Ergebnisse	56
5	Akustische Ereignisdetektion	61
5.1	Datenbasis Ereignisdetektion	61
5.2	Experimente zur Modellierung	62
5.2.1	Modellierung mit Gauß'schen Mischungsverteilungen	62

5.2.2	Modellierung mit universellen Hintergrundmodellen	64
5.3	Diskriminative Lernverfahren	66
5.3.1	<i>MMI</i> -Parameterschätzung	67
5.3.2	Experimentelle Ergebnisse	71
5.4	Quellenauswahl und Fusion	73
5.4.1	Ansätze zur Fusion von Modellbewertungen	74
5.4.2	Experimentelle Ergebnisse	76
6	<i>Middleware</i> und ambiente Intelligenz	79
6.1	Semantisches Netz	79
6.1.1	Ontologien	80
6.1.2	Kontextinformation	80
6.1.3	Abfragesprache für Kontextinformationen	81
6.1.4	Verzeichnisdienst	82
6.2	<i>Webservice</i>	83
6.3	Amigo Architektur	84
6.3.1	Plattform	84
6.3.2	<i>Amigo Middleware</i>	85
6.3.3	Intelligente Dienste	86
6.4	Kontextmanagement	87
6.4.1	Schnittstellendefinition und Kommunikation	87
6.4.2	Kontextbewusste Applikationen	88
6.4.3	Akustische Szenenanalyse als Kontextquelle	89
7	Ambiente Kommunikation	91
7.1	Systemarchitektur und <i>Middleware</i> -Integration	91
7.2	Signalverarbeitung	94
7.2.1	Begrenzer	95
7.2.2	Sprachaktivitätsdetektion	95
7.2.3	Echounterdrückung	96
7.3	Echtzeitkommunikation	99
7.3.1	Lokalisation von Nutzern	100
7.3.2	Sitzungsverwaltung	100
7.3.3	Datenaustausch	101
7.4	Kontextbasierte Steuerung	103
7.4.1	<i>Follow-Me</i> -Fähigkeiten	103
7.4.2	<i>SAInt</i> als Kontextquelle	104
7.4.3	Schutz der Privatsphäre	105
7.5	Visuelle Kommunikation	106
7.5.1	Systemintegration	106
7.5.2	Kommunikationsbeispiel	107
7.5.3	<i>Follow-Me</i> -Fähigkeiten	108
7.6	Demonstration	108
8	Zusammenfassung	109

A Anhang	113
A.1 Herleitung ΔBIC	113
A.2 Herleitung <i>MMI</i> -Parameterschätzung	115
A.3 Experimentelle Ergebnisse der Ereignisdetektion	120
A.4 <i>ML</i> - und <i>MMI</i> -Parameterschätzung	121
Abkürzungsverzeichnis	123
Formelzeichen	127
Abbildungsverzeichnis	131
Tabellenverzeichnis	132
Literaturverzeichnis	135
Eigene Publikationen	147