

Inhaltsverzeichnis

Vorwort	xi
1 Einführung	1
1.1 Automatisches Erkennen und Verstehen von Sprache	1
1.2 Spracherkennungssysteme	3
1.3 Das Spracherkennungs- und Dialogsystem EVAR	7
1.4 Die Komponenten des Erkennungs- und Dialogsystems EVAR	9
1.5 Charakterisierung der Worterkennung im Spracherkennungssystem	14
1.6 Konzept der Arbeit	17
2 Ein flexibles Experimentiersystem zur Worterkennung	20
2.1 Anforderungen an ein Experimentiersystem zur Worterkennung	20
2.2 Aufbau des Experimentiersystems	21
2.3 Beispielkonfigurationen für verschiedene Aufgabenstellungen der Worterkennung	24
2.4 Beurteilungskriterien für Vergleichsverfahren	27
3 Erzeugen von Aussprachemodellen	32
3.1 Variation der Aussprache in kontinuierlich gesprochener deutscher Sprache	32
3.2 Definition von Regelsätzen für Aussprachevarianten	36
3.3 Repräsentation und Konstruktion von Aussprache-Referenzmodellen	39
3.4 Automatisches Erzeugen von Aussprachemodellen anhand eines Lexikons	43
3.5 Interaktives Erzeugen von Aussprachemodellen	45
3.6 Einbettung der lexikalischen Einheit in das Experimentiersystem	46
4 Hidden-Markov-Modelle	49
4.1 Einführung in die theoretischen Grundlagen	49
4.2 Das Vergleichsverfahren	53
4.3 Schätzung der Modellparameter	55
4.4 Berechnungsvorschriften für kantenorientierte Hidden-Markov-Modelle	58
4.5 Ersetzung der Kanten der Aussprachemodelle durch elementare Markov-Modelle	59

5	Klassifikation vorgegebener Zeitbereiche mit Hidden-Markov-Modellen anhand kleiner Modellinventare	64
5.1	Motivation der Aufgabenstellung	64
5.2	Konkretisierung der Parameterschätzung und des Vergleichsverfahrens für HMM mit kontinuierlichen Ausgabeverteilungen	67
5.3	Versuchsaufbau und Ergebnisse	71
6	Explizite Segmentierung in Wortuntereinheiten mit Hidden-Markov-Modellen	79
6.1	Struktur des Segmentierers	79
6.2	Bestimmung einer Initialsegmentierung und Suche der optimalen Zerlegung	82
6.3	Versuchsergebnisse	85
7	Positionieren von Aussprachemodellen mit Hidden-Markov-Modellen (großer Wortschatz)	93
7.1	Vergleichsverfahren und Parameterschätzung für Aussprachemodelle auf der Basis von Wortuntereinheiten	93
7.2	Positionieren von Aussprachemodellen durch 'Vertikale Summation'	103
7.3	Einflußgrößen zur Erzeugung und Normierung von Hypothesen	111
8	Verifizieren von Hypothesen mit Hidden-Markov-Modellen	121
8.1	Verschiedene Gründe für die Verifikation von Hypothesen	121
8.2	Verifizieren von Hypothesen auf Laut- und Lautkomponentenebene	122
8.3	Verifizieren von Worthypothesen bei Verwendung von Aussprachevarianten	127
9	Verifizieren von Hypothesen mit dynamischer Zeitverzerrung (DTW)	131
9.1	Funktionsprinzip	131
9.2	Abstandsmaße zur Berechnung der Ähnlichkeit	136
9.3	Globale Wegbeschränkung, Überführungsregeln und Bewertungsnormalisierung	139
10	Eine Suchstrategie zur kontextfreien Analyse von Wortkettenhypothesen unter Verwendung von Hidden-Markov-Modellen	143
10.1	Der A*-Algorithmus zur Suche von Wortketten	143
10.2	Konkretisierung des Suchverfahrens zur kontextfreien Analyse von Wortketten	146
10.3	Versuchsergebnisse	154

11 Erwartungsgesteuerte Analyse von Wortketten	165
11.1 Einbettung der Worterkennung in eine erwartungsgesteuerte Analyseumgebung	165
11.2 Untersuchungen zur erwartungsgesteuerten syntaktischen Analyse von Konstituentenhypothesen	167
12 Zusammenfassung	171
13 Literatur	177
14 Anhang	186
14.1 Liste der erkennbaren Lautkomponenten des Akustik-Phonetik-Moduls in EVAR	186
14.2 Liste der vereinbarten Lautklassen des Akustik-Phonetik-Moduls in EVAR	187
14.3 Regelsatz für Aussprachevarianten	189
14.4 Liste der Wortarten, semantischen Klassen und pragmatischen Konzepte	191
14.5 Modellinventar zur Unterscheidung von 14 Lautoberklassen	194
14.6 Lauterkennungsdaten für unterschiedliche Modellinventare	195
14.7 Klassifikationsergebnisse aufgeschlüsselt nach 36 Lautklassen	196
14.8 Regelsatz zur Transkription von Lauten nach Lautkomponenten	197
14.9 Stichprobe für die Verifikationsverfahren mit dynamischer Zeitnormierung	197
14.10 Sprachumfang der ATN-Grammatiken	198
14.11 Stichprobe für die Konstituenten- und Satzanalyse	200
14.12 Erkennungsraten bei einer lokalen Beschränkung des Suchraums	200