

Inhaltsverzeichnis

1	Einführung	1
1.1	Motivation	1
1.2	Grundlagen der Sprachwissenschaft	3
1.2.1	Sprachproduktion	4
1.2.2	Sprachperzeption	6
1.2.3	Kategorisierung	8
1.3	Stand der Technik	11
1.4	Beitrag dieser Arbeit zur Forschung	13
2	Parametrisierung	15
2.1	Diskretisierung	15
2.2	Spektrale Analyse	16
2.3	Lineare Vorhersage	19
2.4	Perzeptionsmodelle	22
2.5	Cepstrale Analyse	25
2.6	Berücksichtigung von zeitlichem Kontext	28
3	Klassifikation	32
3.1	Entscheidungstheoretischer Ansatz	33
3.2	Klassifikation bei vorgegebener Form der Wahrscheinlichkeitsdichte	34
3.2.1	Klassenweise normalverteilte Merkmale	35
3.2.2	Schätzung der Parameter	36
3.3	Klassifikation mit direkter Schätzung der a-posteriori Wahrscheinlichkeiten	39
3.4	Klassifikation mit neuronalen Netzwerken	42
3.4.1	Zusammenhang mit parametrischen Klassifikatoren	43
3.4.2	Vorwärtsgerichtete Netze	44
3.4.3	Zeitverzögernde und rückgekoppelte Netze	46
3.4.4	Phonotopische Karten	49
3.5	Integrierte Ansätze mit Hidden-Markov-Modellen	52
3.5.1	Hidden-Markov-Modelle	53
3.5.2	Semi-kontinuierliche Hidden-Markov-Modelle	56
3.5.3	Kombination von HMM und Neuronalen Netzen	58

4	Merkmalsstransformationen	60
4.1	Lineare Transformationen	60
4.2	Nichtlineare Transformationen	63
5	Kompensation von Schwankungen	68
5.1	Modifikationen auf der Ebene der Merkmalsberechnung	69
5.1.1	Normierungen	71
5.1.2	Adaption	79
5.2	Modifikationen auf der Ebene der Klassifikation	82
5.2.1	Robuster Ansatz	82
5.2.2	Adaption	85
6	Stichproben	90
6.1	Beispiele für derzeit verfügbare Stichproben	91
6.2	Die ERBA-Stichprobe	94
6.2.1	Entwurf und Generierung des Korpus	94
6.2.2	Festlegung der Größe der Stichprobe	100
6.2.3	Aufnahme der Stichprobe	101
6.2.4	Automatische Etikettierung	102
6.3	Beschreibung der in dieser Arbeit verwendeten Stichproben	103
7	Experimente und Ergebnisse	105
7.1	Kriterien zur Leistungsbewertung	105
7.2	Parametrisierung	109
7.2.1	Untersuchungen zu statischen Merkmalen	110
7.2.2	Berücksichtigung von Kontext	118
7.3	Klassifikation	122
7.3.1	Vokalklassifikation mit Neuronalen Netzen	122
7.3.2	Vektorquantisierung mit Kohonen-Netzen	128
7.3.3	Zusammenfassung	135
7.4	Merkmalsstransformationen	135
7.4.1	Vergleich verschiedener linearer Transformationen	135
7.4.2	Untersuchungen zur Linearen Diskriminanzanalyse (LDA)	138
7.4.3	Einsatz von MLPs zur Diskriminanzanalyse (NLDA)	139
7.4.4	Zusammenfassung	142
7.5	Kompensation von Schwankungen auf der Ebene der Merkmalsberechnung	142
7.5.1	Vergleich der Normierungsverfahren zur Reduktion der Sprecher- abhängigkeit	142
7.5.2	Einsatz der Normierungsverfahren zum Ausgleich von Kanal- schwankungen	146
7.5.3	Zusammenfassung	149
7.6	Kompensation von Schwankungen auf der Ebene der Klassifikation	149

7.6.1	Sprechergruppen	149
7.6.2	Codebuchadaption	151
7.6.3	Vergleich von Codebuchadaption und dynamischer Cepstraler Sub- traktion	152
7.6.4	Zusammenfassung	154
7.7	Beschleunigung der Verfahren	154
7.7.1	Parametrisierung	155
7.7.2	Klassifikation	156
7.7.3	Zusammenfassung	158
7.8	Ergebnis	159
8	Zusammenfassung	161
A	Verwendete Lautinventare	167
A.1	Die Phoneme des Deutschen	167
A.2	Liste der vereinbarten Lautkomponentenklassen der ERBA-Stichprobe . . .	169
A.3	Lautklassen der TIMIT-Stichprobe	171
B	Beschreibung der Merkmalberechnung	173
B.1	Mel-Cepstrum	173
B.2	Mel-Cepstrum mit DACS	176
C	Die ERBA-Stichprobe	178
C.1	Liste aller verwendeten Schablonen	178
C.2	Vollständige Liste der Regeln	179
C.3	Informationen zur Stichprobe	187
	Literaturverzeichnis	188