

Inhaltsverzeichnis

	Einführung	1
1	Grundsätzliches zur Sprache	3
1.1	Sprache als Kommunikationsmittel	5
1.1.1	Lautsprachliche Kommunikation.....	5
1.1.2	Geschriebene vs. gesprochene Sprache	6
1.2	Die Beschreibung von Sprache.....	8
1.2.1	Die linguistischen Ebenen	8
1.2.2	Die phonetische Schrift.....	10
1.2.3	Die akustisch-phonetische Ebene	11
1.2.4	Die Prosodie der Sprache	11
1.3	Die menschliche Sprachproduktion.....	12
1.3.1	Übersicht über den Sprechapparat	12
1.3.2	Die Funktion des Sprechapparates	12
1.3.3	Die Artikulation der Sprachlaute	14
1.4	Das menschliche Gehör	15
1.4.1	Wahrnehmung der Schallintensität.....	16
1.4.2	Periodizität und Tonhöhe	17
1.4.3	Die Phasenwahrnehmung	18
1.4.4	Der Verdeckungseffekt.....	18
1.4.5	Wahrnehmung der Sprachlaute	19
1.5	Verarbeitung natürlicher Sprache.....	21
2	Übersicht über die Sprachverarbeitung	23
2.1	Was in einem Sprachsignal steckt	25
2.2	Teilgebiete der Sprachverarbeitung.....	26
2.3	Sprachsynthese.....	27
2.4	Spracherkennung.....	28
2.5	Sprecheridentifikation	29
2.6	Sprachidentifikation	29
2.7	Sprechertransformation	30
2.8	Sprachsignalcodierung	30
2.8.1	Signalformcodierung	32
2.8.2	Modellierung von Sprachsignalen.....	35
3	Darstellung und Eigenschaften des Sprachsignals . . .	37
3.1	Digitalisieren von Sprachsignalen	39
3.1.1	Bandbegrenzungsfilter	39
3.1.2	Zeit- und Amplitudendiskretisierung	41
3.1.3	Rekonstruktionsfilter	41

3.2	Darstellung digitaler Sprachsignale im Zeitbereich	43
3.3	Darstellung im Frequenzbereich	45
3.4	Kombinierte Zeit-Frequenz-Bereichs-Darstellungen	48
3.5	Darstellung der Phase eines Sprachsignals	50
3.6	Sprachmerkmale und ihre Darstellung	52
3.6.1	Grundfrequenz	52
3.6.2	Formanten	53
3.6.3	Dauer der Laute	54
3.6.4	Intensität der Laute	55
4	Analyse des Sprachsignals	57
4.1	Kurzzeitanalyse	59
4.2	Schätzung des Kurzzeitspektrums	60
4.2.1	Diskrete Fouriertransformation	60
4.2.2	Eigenschaften der DFT	64
4.2.3	Fensterfunktionen	64
4.2.4	Die Frequenzauflösung der DFT	66
4.2.5	Zeitabhängige Fouriertransformation	69
4.3	Schätzung des Leistungsdichtespektrums	71
4.4	Autokorrelation	73
4.4.1	Definition der Autokorrelationsfunktion	73
4.4.2	Eigenschaften der Autokorrelationsfunktion	73
4.4.3	Kurzzeit-Autokorrelation	74
4.5	Lineare Prädiktion	77
4.5.1	Herleitung der LPC-Analyse	77
4.5.2	Sprachmodellierung mittels linearer Prädiktion	80
4.5.3	Interpretation der linearen Prädiktion	86
4.6	Homomorphe Analyse	89
4.6.1	Das verallgemeinerte Superpositionsprinzip	89
4.6.2	Homomorphe Systeme	89
4.6.3	Das DFT-Cepstrum	90
4.6.4	Cepstrale Glättung	92
4.6.5	Das Mel-Cepstrum	94
4.6.6	Das Delta-Cepstrum	96
4.6.7	Mittelwertfreie Cepstren	97
4.6.8	Cepstrale Distanz	99
4.7	Vektorquantisierung	100
4.7.1	Realisation der Vektorquantisierung	101
4.7.2	Generieren eines Codebuches	103
5	Hidden-Markov-Modelle	107
5.1	Struktur und Parameter eines HMM	109

5.1.1	Zustandsübergangswahrscheinlichkeiten.....	110
5.1.2	Beobachtungswahrscheinlichkeiten.....	112
5.2	Die grundlegenden HMM-Probleme	114
5.3	Trellis-Diagramm	115
5.4	Grundlegende Algorithmen für DDHMM	117
5.4.1	Evaluationsproblem.....	117
5.4.2	Forward-Algorithmus für DDHMM.....	118
5.4.3	Decodierungsproblem	120
5.4.4	Viterbi-Algorithmus für DDHMM	120
5.4.5	Schätzproblem.....	122
5.4.6	Backward-Algorithmus für DDHMM	122
5.4.7	Baum-Welch-Algorithmus für DDHMM.....	123
5.4.8	Viterbi-Training für DDHMM	126
5.4.9	Initial-DDHMM	128
5.5	Grundlegende Algorithmen für CDHMM.....	129
5.5.1	Forward-Algorithmus für CDHMM.....	129
5.5.2	Viterbi-Algorithmus für CDHMM	129
5.5.3	Backward-Algorithmus für CDHMM.....	129
5.5.4	Baum-Welch-Algorithmus für CDHMM.....	130
5.5.5	Viterbi-Training für CDHMM	131
5.5.6	Initial-CDHMM	132
5.5.7	Mixture Splitting.....	133
5.6	Training mit mehreren Beobachtungssequenzen.....	134
5.7	Underflow bei HMM	135
6	Darstellung und Anwendung linguistischen Wissens..	137
6.1	Formale Sprachen und Grammatiken	139
6.2	Die Sprachhierarchie nach Chomsky	140
6.2.1	Reguläre Sprachen (Typ 3)	143
6.2.2	Kontextfreie Sprachen (Typ 2).....	145
6.2.3	Kontextsensitive Sprachen (Typ 1).....	146
6.2.4	Allgemeine Sprachen (Typ 0)	150
6.2.5	Das Wortproblem.....	150
6.3	Die Wortanalyse	152
6.3.1	Wortanalyse für Typ-3-Grammatiken	152
6.3.2	Wortanalyse für Typ-2-Grammatiken	160
6.3.3	Wortanalyse für Typ-1- und Typ-0-Grammatiken	167
6.4	Formalisierung natürlicher Sprachen	167
6.5	Der DCG-Formalismus.....	172
6.5.1	Definition und Eigenschaften von DCG	172
6.5.2	Unifikation	175

6.5.3	DCG-Ableitungen	175
6.5.4	DCG-Ableitungsbaum.....	177
6.5.5	DCG und Chart-Parsing	178
6.6	Two-Level-Regeln und Transduktoren	178
6.6.1	Einführung	178
6.6.2	Two-Level-Regeln	179
6.6.3	Transduktoren	182
7	Einführung in die Sprachsynthese	191
7.1	Überblick über die Geschichte der Sprachsynthese.....	193
7.2	Aufgabe der Sprachsynthese.....	194
7.3	Zusammenhang zwischen Lautsprache und Schrift	195
7.4	Teile der Sprachsynthese	196
7.4.1	Die Transkription	197
7.4.2	Die phonoakustische Stufe.....	199
7.5	Lautinventar für die Sprachsynthese	200
7.5.1	Linguistische Grundlagen.....	200
7.5.2	Festlegen der Lautdifferenzierung	201
8	Sprachsynthese: Transkription	203
8.1	Linguistische Grundlagen für die Transkription	205
8.1.1	Ermitteln der Lautfolge.....	206
8.1.2	Ermitteln der Prosodie	211
8.2	Automatische Transkription	214
8.2.1	Der "direkte" Ansatz der Transkription	214
8.2.2	Der linguistische Ansatz der Transkription	214
8.3	Automatische morphosyntaktische Analyse.....	216
8.3.1	Morphologische Analyse mit DCG.....	217
8.3.2	Generierung der phonetischen Umschrift in einer DCG ..	218
8.3.3	Hauptprobleme der morphosyntaktischen Analyse	219
8.4	Realisation einer Transkriptionsstufe.....	223
8.4.1	DCG in SVOX.....	224
8.4.2	Morphologische Analyse in SVOX.....	225
8.4.3	Syntaxanalyse in SVOX	225
8.4.4	Analyse unbekannter Wortstämme.....	227
8.4.5	Phonetische Umschrift ganzer Wörter.....	228
8.4.6	Akzentuierung	230
8.4.7	Phrasierung	233
8.4.8	Generierung der phonologischen Darstellung.....	234
8.4.9	Weiterverarbeitung der phonologischen Darstellung	235

9	Sprachsynthese: Phonoakustische Stufe	237
9.1	Verfahren für die Sprachsignalproduktion	239
9.1.1	Der artikulatorische Ansatz	240
9.1.2	Der Signalmodellierungsansatz	241
9.1.3	Der Verkettungsansatz	245
9.2	Sprachsynthese nach dem Verkettungsansatz	247
9.2.1	Wahl der Grundelemente	248
9.2.2	Ausschneiden von Grundelementen	249
9.2.3	Veränderung der prosodischen Grössen	252
9.2.4	Signalveränderung mittels LPC-Analyse-Synthese	253
9.2.5	Signalveränderung mittels Fourier-Analyse-Synthese	254
9.2.6	Signalveränderung mittels PSOLA	260
9.3	Steuerung der Prosodie	262
9.3.1	Dauersteuerung	264
9.3.2	Grundfrequenzsteuerung	273
9.3.3	Intensitätssteuerung	284
9.3.4	Umsetzung der prosodischen Grössen auf die Laute	286
9.3.5	Prosodische Veränderung der Grundelemente	286
10	Polyglotte Sprachsynthese	289
10.1	Motivation und Einführung	291
10.1.1	Aufgabe der polyglotten Sprachsynthese	291
10.1.2	Gemischtsprachige Texte	292
10.2	Konzept der polyglotten Sprachsynthese	294
10.3	Transkription gemischtsprachiger Texte	297
10.3.1	Morphologische Analyse gemischtsprachiger Wörter	297
10.3.2	Syntaktische Analyse gemischtsprachiger Sätze	303
10.3.3	Probleme der polyglotten morphosyntaktischen Analyse	307
10.3.4	Phonologische Transformationen	309
10.3.5	Akzentuierung und Phrasierung	314
10.3.6	Rekapitulation der polyglotten Transkription	316
10.4	Polyglotte phonoakustische Stufe	317
10.4.1	Polyglotte Prosodiesteuerung	318
10.4.2	Polyglotte Sprachsignalproduktion	321
11	Einführung in die Spracherkennung	323
11.1	Zur Geschichte der Spracherkennung	325
11.2	Ansätze zur Spracherkennung	327
11.3	Probleme der Spracherkennung	328
11.4	Anwendungen	331
11.5	Einteilung der Spracherkennungssysteme	332
11.6	Evaluation der Erkennungsleistung	334

11.6.1	Wortfehlerrate	335
11.6.2	Algorithmus zur Bestimmung der Wortfehlerrate	335
11.7	Merkmalsextraktion	336
11.7.1	Mel Frequency Cepstral Coefficients (MFCC)	338
11.7.2	Geglättete Mel-Spektren von Vokalen und Frikativen	339
11.7.3	Plosivlaute im geglätteten Mel-Spektrogramm	340
11.7.4	Spektrale Variabilität	342
11.7.5	Rekonstruktion des Signals	342
11.8	Aufnehmen einer Äusserung	344
11.8.1	Anwendungsszenarien	344
11.8.2	Anfangs- und Endpunktdetektion	346
12	Spracherkennung mit Mustervergleich	351
12.1	Das Prinzip des Sprachmustervergleichs	353
12.2	Zeitliche Anpassung zweier Sprachmuster	356
12.3	Randbedingungen für die Warping-Kurve	357
12.4	Der DTW-Algorithmus	360
12.5	Spracherkennung mittels DTW	363
12.5.1	Generieren von Referenzmustern	363
12.5.2	Einsatzmöglichkeiten und Grenzen	365
13	Statistische Spracherkennung	367
13.1	Informationstheoretische Sicht	369
13.2	Spracherkennung gemäss MAP-Regel	369
13.3	Modellierung von Merkmalssequenzen	371
13.3.1	Variabilität von Merkmalssequenzen	369
13.2n68.8 g9 g -246638 0 60.9 l b.7 s1 le 0.9 1s2eme9 n -6.2 z -1.2 -71429 n -1.		

13.6.2	Erzeugen von Grundelementmodellen	397
13.7	Modelle für verschiedene akustische Ereignisse	403
13.7.1	Modelle für Pausen	403
13.7.2	Modelle für Geräusche	404
13.7.3	Modell für beliebige Sprachsignale	404
13.8	Spracherkennung mit Laut-HMM	405
13.8.1	Erkennung einzeln gesprochener Wörter	405
13.8.2	Erkennung kontinuierlicher Sprache	405
13.8.3	Reduktion des Rechenaufwands (Pruning)	406
13.9	Stärken und Schwächen von HMM	407
14	Sprachmodellierung	409
14.1	Zum Begriff der Sprachmodellierung	411
14.2	Statistische Sprachmodellierung	412
14.2.1	Sprachmodellierung bei der Einzelworterkennung	413
14.2.2	Sprachmodellierung für Wortfolgen	415
14.2.3	Das allgemeine statistische Sprachmodell	415
14.2.4	N-Gram-Sprachmodelle	417
14.2.5	Schätzen der Parameter von N-Gram-Sprachmodellen	420
14.2.6	Kategorielle N-Gram-Sprachmodelle	423
14.2.7	Anwendung von N-Gram-Sprachmodellen	424
14.2.8	Bewertung von Sprachmodellen	425
14.2.9		

B.1.1	Regeln der Wahrscheinlichkeitsrechnung	459
B.1.2	Wahrscheinlichkeitsverteilungen	461
B.2	z-Transformation.....	467
B.3	Neuronale Netze: Mehrschicht-Perzeptron	469
B.3.1	Das Neuronenmodell	469
B.3.2	Das Mehrschicht-Perzeptron	470
B.3.3	Anwendungen von Mehrschicht-Perzeptronen	471
B.3.4	Training eines Mehrschicht-Perzeptrons.....	472
B.3.5	Hinweise zum Einsatz von neuronalen Netzen	474
B.3.6	Komplexe neuronale Netze	476
	Notationen	479
	Glossar	483
	Literaturverzeichnis	497
	Index	501

Sprachverarbeitung

Grundlagen und Methoden der Sprachsynthese und
Spracherkennung

Pfister, B.; Kaufmann, T.

2017, XIV, 507 S. 175 Abb. Mit Online-Extras., Softcover

ISBN: 978-3-662-52837-2