

# INHALT

Vorwort .....	11
<b>1 Lautsprachliche Kommunikation .....</b>	<b>15</b>
1.1 Die Bedeutung gesprochener Sprache für die Kommunikation zwischen Mensch und Maschine .....	15
1.2 Gedanken und Sprache .....	16
1.3 Das Verhältnis von geschriebener und gesprochener Sprache .....	16
1.4 Phonetik und Phonologie .....	17
1.5 Das akustische Signal .....	17
1.6 Phoneme, Phone und Allophone .....	18
1.7 Vokale und Konsonanten .....	20
1.8 Phoneme und Orthographie .....	21
1.9 Prosodische Merkmale .....	21
1.10 Sprache, Akzent und Dialekt .....	23
1.11 Die Ergänzung des akustischen Signals .....	24
1.12 Die Komplexität der Sprachverarbeitung .....	26
Zusammenfassung .....	26
Übungen .....	28
<b>2 Mechanismen und Modelle der Sprachproduktion .....</b>	<b>29</b>
2.1 Einführung .....	29
2.2 Schallquellen .....	30
2.3 Das Resonanzsystem .....	35
2.4 Interaktion zwischen der Funktion von Kehlkopf und Vokaltrakt ...	41
2.5 Schallabstrahlung .....	43
2.6 Zeitsignale und Spektrogramme .....	45
2.7 Sprachproduktionsmodelle .....	49
2.7.1 Anregungsmodelle .....	51

2.7.2	Vokaltraktmodelle .....	53
	Zusammenfassung .....	59
	Übungen .....	61
<b>3</b>	<b>Mechanismen und Modelle des menschlichen Gehörs .....</b>	<b>63</b>
3.1	Einleitung .....	63
3.2	Physiologie des Außen- und Mittelohres .....	63
3.3	Der Aufbau der Cochlea .....	65
3.4	Neuronale Verarbeitung .....	67
3.5	Psycho-physikalische Messungen .....	69
3.6	Analyse von einfachen und komplexen Signalen .....	72
3.7	Modelle des Gehörs .....	73
3.7.1	Mechanische Filterung .....	73
3.7.2	Modelle der neuronalen Übertragung .....	74
3.7.3	Neuronale Verarbeitung auf höherer Ebene .....	75
	Zusammenfassung .....	76
	Übungen .....	77
<b>4</b>	<b>Digitale Codierung gesprochener Sprache .....</b>	<b>79</b>
4.1	Einführung .....	79
4.2	Einfache Signalformcodierung .....	81
4.2.1	Pulsmodulation .....	81
4.2.2	Delta-Modulation .....	84
4.3	Analyse/Synthese-Systeme (Vocoder) .....	86
4.3.1	Kanalvocoder .....	87
4.3.2	LPC-Vocoder .....	88
4.3.3	Formantvocoder .....	90
4.3.4	Effiziente Codierung der Parameter .....	92
4.4	Zwischensysteme .....	92
4.4.1	Teilbandcodierung .....	93
4.4.2	Adaptive Transformationscodierung .....	94
4.4.3	Auf linearer Prädiktion basierende Methoden .....	94
4.5	Die Wahl des Codierers .....	99
	Zusammenfassung .....	99
	Übungen .....	100
<b>5</b>	<b>Sprachausgabe mit gespeicherter menschlicher Sprache .....</b>	<b>102</b>
5.1	Einführung .....	102
5.2	Die Verknüpfung von Sprachsignalen .....	102
5.3	Die Verknüpfung von vocoder-generierten Wörtern .....	106

---

5.4	Die Verknüpfung von Einheiten unterhalb der Wortebene .....	107
5.5	Hardware-Anforderungen .....	110
	Zusammenfassung .....	111
	Übungen .....	112
<b>6</b>	<b>Regel-basierte Sprachsynthese .....</b>	<b>113</b>
6.1	Einführung .....	113
6.2	Akustisch-phonetische Regeln .....	115
6.2.1	Regeln für Formantsynthetisatoren .....	116
6.2.2	Tabellengesteuerte phonetische Regeln .....	117
6.2.3	Optimierung phonetischer Regeln .....	124
6.2.4	Regeln für verschiedene Sprechertypen .....	126
6.3	Intensitätsregeln .....	127
6.4	Dauerregeln .....	127
6.5	Grundfrequenzregeln .....	129
6.6	Regeln zur Textanalyse .....	131
6.6.1	Graphem-Phonem-Konversion .....	131
6.6.2	Betonungsbestimmung .....	134
6.7	Vorverarbeitung von Zahlwörtern, Abkürzungen, etc. ....	134
6.8	Hardware-Implementation .....	135
6.9	Fähigkeiten der gegenwärtigen Regelsysteme .....	137
	Zusammenfassung .....	137
	Übungen .....	138
<b>7</b>	<b>Spracherkennung durch Mustervergleich ganzer Wörter .....</b>	<b>140</b>
7.1	Allgemeine Prinzipien .....	140
7.2	Abstandsmaße .....	141
7.2.1	Filterbankanalyse .....	143
7.2.2	Pegelnormierung .....	144
7.2.3	Andere einfache, auf dem Spektrum basierende Abstandsmaße. ....	147
7.2.4	Analyse durch lineare Prädiktion .....	148
7.2.5	Analyse auf der Basis von Modellen auditorischer Perzeption ..	149
7.3	Ende-Erkennung bei isolierten Wörtern .....	150
7.4	Erlaubte Variationen im Zeitablauf .....	151
7.5	Dynamische Programmierung für die Zeitanpassung .....	151
7.6	Verfeinerte Anwendungen der DP beim Vergleich isolierter Wörter ...	154
7.7	Begrenzung des Vergleichs .....	157
7.8	Zulässige Fehler am Wortende .....	159
7.9	Analyse mit variabler Frame-Rate .....	160
7.10	Erweiterung der dynamischen Programmierung auf Wortketten .....	160
7.11	Erkennung kontinuierlicher Sprache .....	164
7.12	Syntaktische Beschränkungen .....	166

7.13	Trainingsphase eines Erkennungssystems .....	167
7.14	Sprecher-unabhängige Spracherkennung .....	168
7.15	Auswirkungen des Hintergrundgeräuschs .....	169
	Zusammenfassung .....	170
	Übungen .....	171
<b>8</b>	<b>Stochastische Modelle für die Worterkennung .....</b>	<b>173</b>
8.1	Berücksichtigung der Merkmalsvariabilität beim Mustervergleich ...	173
8.2	Einführung in die Hidden-Markov-Modelle .....	175
8.3	Berechnung der Wahrscheinlichkeiten in Hidden-Markov-Modellen ..	178
8.4	Der Viterbi-Algorithmus .....	181
8.5	Parameter-Abschätzung für Hidden-Markov-Modelle .....	183
	8.5.1 Vor- und Rückwärtswahrscheinlichkeiten .....	184
	8.5.2 Auswahl von Startwerten .....	187
8.6	Folgen von zu kleinem Trainingsmaterial (Teil 1) .....	188
8.7	Vektorquantisierung .....	189
8.8	Multivariate kontinuierliche Verteilungen .....	190
	8.8.1 Abschätzung von Normalverteilungen nach Baum-Welch .....	192
8.9	Modellierung der zeitlichen Dauer in Hidden-Markov-Modellen ...	193
8.10	Gebrauch von Normalverteilungen im Viterbi-Algorithmus .....	195
8.11	Praktische Probleme bei der Berechnung .....	197
8.12	Folgen von zu kleinem Trainingsmaterial (Teil 2) .....	199
8.13	Erweiterung von Hidden-Markov-Modellen auf Wortketten .....	200
	Zusammenfassung .....	201
	Übungen .....	202
<b>9</b>	<b>Spracherkennung bei sehr großem Vokabular .....</b>	<b>204</b>
9.1	Probleme bei der Erweiterung der Spracherkennungsmethoden für kleines Vokabular .....	204
9.2	Sprachtranskription und Sprachverständnis .....	206
9.3	Phonetische Merkmale .....	207
	9.3.1 Expertensysteme .....	209
	9.3.2 Erkennung phonetischer Merkmale mit statistischen Methoden.	209
9.4	Segmentgitter .....	211
9.5	Bottom-up und top-down Verfahren .....	211
9.6	Einsatz von Hidden-Markov-Modellen für große Vokabulare .....	213
9.7	Die IBM-Diktiermaschine .....	215
9.8	Berücksichtigung der Koartikulation .....	217
9.9	Regel-basierte Synthese als ein Modell der Spracherkennung .....	218
	Zusammenfassung .....	219
	Übungen .....	220

---

<b>10</b>	<b>Zukünftige Forschungsrichtungen für Sprachsynthese und Spracherkennung</b> .....	221
10.1	Einleitung .....	221
10.2	Sprachsynthese .....	222
10.3	Automatische Spracherkennung .....	224
10.4	Die Beziehung zwischen Synthese und Erkennung .....	227
10.5	Parallele verteilte Verarbeitung .....	228
10.5.1	Das menschliche Gehirn .....	228
10.5.2	Konnektionistische Modelle .....	230
10.5.3	Bedeutende Eigenschaften der PDP-Modelle .....	230
10.5.4	PDP und Sprachverarbeitung .....	233
	Zusammenfassung .....	234
	Übungen .....	235
<b>11</b>	<b>Weiterführende Literatur</b> .....	236
	<b>Literaturverzeichnis</b> .....	242
	<b>Lösungen der Übungsaufgaben</b> .....	246
	<b>Sachregister</b> .....	253