

1. Spracherzeugung und Wahrnehmung — Kommunikation zwischen Sprecher und Hörer	11
(S. Marcus)	
1.1 Das akustische Signal	11
1.2 Spracherzeugung	13
1.3 Frequenz-Analyse	18
1.4 Lautübergänge und phonetische Lauteinheiten	21
1.5 Unterschiede zwischen Sprechern	29
1.6 Zusammenfassung	30
1.7 Literaturverzeichnis	31
2. Mechanismen der Spracherzeugung und Sprachwahrnehmung	35
(S. Marcus)	
2.1 Das Zentralnervensystem	35
2.2 Spracherzeugung	37
2.3 Das Gehör	43
2.4 Empfindungsgrößen des Gehörs	48
2.5 Zusammenfassung	55
2.6 Literaturverzeichnis	55
3. Einführung in die digitale Signalverarbeitung	59
(R. Geppert)	
3.1 Einleitung	59
3.1.1 Vor- und Nachteile von DSP-Systemen	59
3.1.2 Grundelemente eines DSP-Systems	61
3.2 Binärzahlensystem und binäre Arithmetikelemente	63
3.2.1 Das Binärzahlensystem	63
3.2.2 Arithmetische Bausteine	64
3.2.3 Andere Bauelemente	66
3.3 Grundbegriffe der digitalen Signalverarbeitung	67
3.3.1 Der Systembegriff	67
3.3.2 Die Diskrete Fourier-Transformation	68
3.3.3 Das Abtast-Theorem	70
3.3.4 Die z-Transformation	72
3.3.5 Die inverse z-Transformation	75
3.3.6 Die Systemfunktion $H(z)$	76
3.4 Analog-Digital- und Digital-Analog-Umsetzung	77
3.4.1 D-A-Umsetzer	77
3.4.2 Verarbeitung von vorzeichenlosen und vorzeichenbehafteten Binärzahlen	79
3.4.3 A-D-Umsetzer	81
3.5 Digitale Filter	84
3.5.1 Nichtrekursive Systeme (FIR-Filter)	85

3.5.2 Rekursive Systeme (IIR-Filter)	90
3.5.3 Vergleich von FIR- und IIR-Filtern	95
3.6 Die schnelle Fourier-Transformation (FFT)	96
3.6.1 Die Basis-2-FFT	96
3.6.2 Einige Bemerkungen zum FFT-Algorithmus	99
3.6.3 Ein FFT-Fortran-Programm	100
3.6.4 Die inverse Diskrete Fourier-Transformation (IDFT)	102
3.6.5 Schlußbemerkung	103
3.7 Integrierte digitale Signalprozessoren	103
3.7.1 Intel i2920	105
3.7.2 AMI S2811	106
3.7.3 NEC μ PD 7720	108
3.7.4 Vergleichstabellen	111
3.8 Literaturverzeichnis	112
4. Methoden der Sprachanalyse	117
(H. Ney)	
4.1 Einführung	117
4.2 Analyse im Zeitbereich	118
4.2.1 Bestimmung der Energie	118
4.2.2 Nulldurchgänge	119
4.2.3 Autokorrelationsfunktion	121
4.3 Analyse im Frequenzbereich	126
4.3.1 Fourier-Transformation	126
4.3.2 Filterbank-Analyse	129
4.3.3 Spektrographische Darstellung	130
4.4 Homomorphe Analyse und Cepstrum	131
4.4.1 Prinzip der homomorphen Analyse	131
4.4.2 Cepstrum	133
4.5 LPC-Analyse	134
4.5.1 Prinzip der LPC-Analyse	134
4.5.2 Berechnungsverfahren	135
4.5.3 Interpretation im Frequenzbereich	137
4.5.4 Alternative Darstellungen der Prädiktor-Koeffizienten	141
4.6 Grundfrequenz-Analyse	142
4.6.1 Einführung	142
4.6.2 Feinstruktur des Spektrums	143
4.6.3 Fehlerfunktion und Kammfilter	144
4.6.4 Gesamttablauf der Grundfrequenz-Analyse	146
4.7 Zusammenfassung	148
4.8 Literaturverzeichnis	150

5. Sprachausgabe	155
(K. Sickert)	
5.1 Sprachausgabe aus analogen Speichern	155
5.2 Sprachausgabe aus digitalen Speichern	156
5.2.1 Signalformcodierung	158
5.2.1.1 Direkte Quantisierung	158
5.2.1.2 Differentielle Quantisierung	161
5.2.2 Mozer-Codierung	169
5.2.3 Quellencodierung — »Source Coding«	175
5.2.3.1 Formantensynthese	178
5.2.3.2 LPC-Synthese	197
5.3 Synthese nach Regeln — Vollsynthese	205
5.3.1 Synthese im Zeitbereich	206
5.3.2 Synthese aus Parametern	208
5.4 Anwendungen von Sprachausgabesystemen	215
5.4.1 Systeme mit einer Sprachausgabeeinheit	216
5.4.2 Auskunft- und Informationssysteme	217
5.5 Literaturverzeichnis	219
6. Spracheingabe	223
(R. Geppert, M.H. Kuhn, H. Ney)	
6.1 Verfahren der Mustererkennung	225
6.1.1 Struktur eines Erkennungssystems	226
6.1.2 Merkmalsextraktion	228
6.1.2.1 Statistische Merkmale	230
6.1.2.2 Merkmale durch Segmentierung	231
6.1.2.3 Zeitabhängige Merkmale	232
6.1.3 Vorverarbeitungsschritte	232
6.1.3.1 Normierung	232
6.1.3.2 Glättung der Zeitverläufe	233
6.1.3.3 Ortskurvensegmentierung	233
6.1.4 Nichtlineare Zeitanpassung	235
6.1.4.1 Vorgegebener Anfangs- und Endpunkt	235
6.1.4.2 Unbekannter Anfangs- und Endpunkt (»Word Spotting«)	242
6.1.5 Klassifikation	244
6.1.5.1 Statistische Entscheidungstheorie	245
6.1.5.2 Normalverteilung	247
6.1.5.3 Diskriminantenfunktion	251
6.1.5.4 Geometrische Klassifikatoren	252
6.1.5.5 Nachlernende Klassifikatoren	254
6.1.5.6 Unüberwachte Klassifikation oder »Clustering«	255
6.1.5.7 Bestimmung der Fehlerrate	256

6.2 Sprecherkennung	258
6.2.1 Sprechertypische Parameter, Merkmale und Muster	263
6.2.1.1 Statistische Merkmale	267
6.2.1.2 Dynamische Merkmale	272
6.2.2 Sprecherverifikation in geschlossenen Systemen	276
6.2.2.1 Randbedingungen für die Systembewertung	279
6.2.2.2 Beschreibung verschiedener Systemrealisierungen	281
6.2.3 Benutzerverhalten und Bedienungsaspekte	292
6.2.4 Sprecherverifikation über Telefon	297
6.2.4.1 Randbedingungen bei Telefonübertragung	298
6.2.4.2 Überblick über verschiedene Verfahren	298
6.2.4.3 Vergleichende Untersuchung verschiedener Erkennungsverfahren ..	299
6.3. Spracherkennung	313
6.3.1 Anwendungsgebiete und wichtige Systemeigenschaften	315
6.3.2 Einzelworterkennung (EWE)	323
6.3.2.1 Sprachanalyse	324
6.3.2.2 Detektion von Wortanfang und -ende	326
6.3.2.3 Vorverarbeitung und Normierung	332
6.3.2.4 Nichtlineare Zeitanpassung	341
6.3.2.5 Referenzerstellung und Mustererkennung	344
6.3.2.6 Vorteile einer Strukturierung des Dialogs	350
6.3.2.7 Experimentelle Ergebnisse	352
6.3.3 Wortkettenerkennung	358
6.3.3.1 Erkennung von Wortketten ohne Syntax	359
6.3.3.2 Erkennung von Wortketten mit Syntax	367
6.3.4 Erkennung und Interpretation kontinuierlicher Sprache	372
6.3.5 Kommerziell verfügbare Systeme	390
6.4 Literaturverzeichnis	405
7. Kombinierte Systeme mit Spracheingabe und Sprachausgabe	421
(K. Sickert)	
7.1 Literaturverzeichnis	426
8. Sachwortverzeichnis	429
9. Produktinformation	435