

Inhaltsverzeichnis

1	Einleitung	1
2	Sprechen und Hören	5
2.1	Lautsprachliche Kommunikationskette	5
2.2	Die menschlichen Sprechorgane	8
2.2.1	Anregung des Sprechtrakts	10
2.2.2	Formanten und Vokalartikulation	13
2.2.3	Röhrenmodell des Vokaltrakts	20
2.2.4	Lineares Modell der Spracherzeugung	26
2.3	Aufbau und Funktionsweise des menschlichen Gehörs	28
2.3.1	Grundsätzlicher Aufbau des Gehörorgans	28
2.3.2	Hörfäche	31
2.3.3	Mithörschwelle, Frequenzgruppe, Lautheit, Maskierung	33
2.3.4	Differentielle Wahrnehmbarkeitsschwellen	38
2.4	Aspekte der Phonetik und Linguistik	39
2.4.1	Linguistische Einheiten: Phonem und Phon	39
2.4.2	Die verschiedenen Informationsebenen in einem sprachverarbeitenden System	42
2.4.3	Lautlehre: Vokale und Konsonanten	44
2.4.4	Internationales Lautschriftalphabet; Transkription	50
2.4.5	Distinktive Merkmale	52
2.4.6	Lautübergänge und Koartikulation	54
2.4.7	Phonotaktik	57
2.4.8	Prosodie	58
3	Signale, Spektren und Systeme	61
3.1	Signalbeschreibungen	61
3.2	Spektraltransformationen	62
3.3	Systembeschreibung im Zeit- und Spektralbereich	66
3.4	Logarithmisches Spektrum und Cepstrum	68

4	Spektralanalyse	71
4.1	Kurzzeitspektren	71
4.1.1	Filterung einer Spektralkomponente	71
4.1.2	Taktreduktion	73
4.1.3	Messung eines Gesamtspektrums	75
4.2	Unmittelbare Transformations-Berechnung	75
4.2.1	Effiziente DFT-Berechnung (FFT)	76
4.2.2	DFT als Filterbank	81
4.2.3	Taktreduktion	85
4.2.4	Verallgemeinerte Spektraltransformationen	85
4.2.5	Verallgemeinerte DFT (GDFT)	86
4.2.6	Diskrete Cosinus-Transformation (DCT, GDCT)	87
4.2.7	Karhunen-Loève-Transformation (KLT)	91
4.3	Filterbänke	94
4.3.1	Filterrealisierungen	94
4.3.2	Baumstrukturen	102
4.3.3	Polyphasen-Filterbänke	107
4.3.4	Modulierte Filterbänke	110
4.3.5	Polyphasen-Filterbänke mit ungleichmäßiger Auflösung	111
5	Spektralsynthese, Analyse-Synthese-Systeme	117
5.1	Begriffsklärung	117
5.2	Synthese- und Analyse-Synthese-Filterbänke	119
5.3	QMF-Bänke	120
5.3.1	Gleichphasige und versetzte Taktreduktion	123
5.3.2	Zweikanal – QMF – System	123
5.3.3	Filterentwurf	124
5.4	Polyphasen-Filterbänke	126
5.4.1	Struktur und Varianten	126
5.4.2	Taktreduktion	127
5.4.3	Aliaskompensation	129
5.4.4	Filterentwurf	131

6	Statistische Analyse	137
6.1	Benötigte Begriffe	137
6.1.1	Verteilung, Verteilungsdichte, Stationarität	137
6.1.2	Erwartungswerte, Momente, Korrelationen	139
6.1.3	Unkorreliertheit, Orthogonalität, Dekorrelation	142
6.1.4	Statistische Unabhängigkeit	142
6.1.5	Korrelations- und Kovarianzmatrizen	143
6.1.6	Spektren	144
6.1.7	Lineare Filterung von Zufallssignalen	145
6.1.8	Beispiele	147
6.2	Messung statistischer Kenngrößen	152
6.2.1	Verteilungsdichte, Histogramm	152
6.2.2	Zeitmittelwerte, Ergodizität, Schätzung	155
6.2.3	Zeitliche Momente, zeitliche Korrelation	157
6.2.4	Zeitliche Korrelations- und Kovarianzmatrizen	161
6.2.5	Kurzzeitspektren	162
7	Lineare Prädiktion	165
7.1	Zugrundeliegendes Modell und Kurzzeitprädiktion	165
7.2	Optimale Prädiktorkoeffizienten bei Stationarität	171
7.3	Adaptive Einstellung des linearen Prädiktors	174
7.3.1	Blockorientierte Adaption	174
7.3.2	Sequentielle Adaption	184
7.4	Langzeitprädiktion	188
8	Grundperiode, Grundfrequenz, Anregungsart	195
8.1	Grundfrequenzbestimmung: Übersicht	196
8.1.1	Definitionen des Parameters Sprachgrundfrequenz	197
8.1.2	Grobunterteilung der GFB-Algorithmen	200
8.2	GFB nach dem Prinzip der Kurzzeitanalyse	201
8.2.1	Überblick	201
8.2.2	Beispiel: GFB mit Hilfe doppelter Spektraltransformation und nichtlinearer Verzerrung im Frequenzbereich	204
8.2.3	GFB mit aktiver Modellierung; weitere Entwicklungen	207
8.3	Grundperiodenbestimmung im Zeitbereich	209
8.3.1	Analyse der Zeitstruktur	210
8.3.2	Bestimmung der ersten Teilschwingung	211
8.3.3	Mehrkanalalgorithmen	213

8.4	Korrektur und Glättung von GF-Verläufen	214
8.5	Stimmbandschwingung, Glottisverschußzeitpunkt	215
8.5.1	Rekonstruktion der Stimmbandschwingung	216
8.5.2	Bestimmung des Glottisverschußzeitpunktes	217
8.6	Bestimmung der Anregungsart	219
8.6.1	Schwellwertanalyse mit wenigen Parametern	220
8.6.2	Simultane Bestimmung von Anregungsart und Grundfrequenz	222
8.6.3	ABA mit Hilfe der Mustererkennung	225
8.7	Evaluierung und Robustheit	227
9	Quantisierung und Codierung	233
9.1	Klassifikation und Kriterien	233
9.1.1	Klassifikation der Algorithmen zur Sprachcodierung	234
9.1.2	Kriterien zur Beurteilung	236
9.1.3	Quantisierung und Codierung	238
9.2	Gleichmäßige Quantisierung	239
9.3	Quantisierung mit Kompondierung	247
9.4	Optimalquantisierung	256
9.5	Adaptive Quantisierung	257
9.6	Vektorquantisierung	262
9.6.1	Prinzip	262
9.6.2	Das Komplexitätsproblem	265
9.6.3	Lattice-Quantisierung	266
9.6.4	Entwurf von optimalen Vektor-Codebüchern	267
9.6.5	Gain-Shape-Vektorquantisierung	270
10	Codierung im Zeitbereich	271
10.1	Modellgestützte prädiktive Codierung	271
10.2	Differentielle Signalform-Codierung	273
10.2.1	Grundstrukturen	273
10.2.2	Quantisierung des Restsignals	278
10.2.3	ADPCM: Adaptive Differenz-Puls-Code-Modulation	288
10.3	Parametrische Codierung	290
10.3.1	Vocoder-Strukturen	290
10.3.2	LPC-Vocoder	293
10.3.3	Quantisierung der Prädiktorkoeffizienten	295
10.4	Hybrid-Codierung	301
10.4.1	Gemeinsame Grundlage der Codec-Konzepte	301
10.4.2	Restsignal-Codierung: RELP	311
10.4.3	Analyse-durch-Synthese: CELP	319
10.4.4	Analyse-durch-Synthese: MPE, RPE	327
10.5	Codec-Verbesserung durch adaptive Nachfilterung	331

11 Codierung im Frequenzbereich	337
11.1 Hintergrund	337
11.2 Transformationscodierung (TC)	340
11.2.1 Prinzip	340
11.2.2 Fehlervarianz	341
11.2.3 Mögliche Transformationen	342
11.2.4 Optimale Bitzuteilung	343
11.2.5 Minimale Fehlerleistung, Störspektrum	346
11.2.6 Transformationsgewinn, Wortlängenreduktion	346
11.2.7 Optimale und praktikable Transformationen	348
11.2.8 Adaptive Transformationscodierung (ATC)	351
11.2.9 Realisierung	353
11.3 Teilbandcodierung (SBC)	355
11.3.1 Prinzip	355
11.3.2 Bandbreiten und Bitzuteilung	356
11.3.3 Adaption	357
11.3.4 Teilband-Differenzcodierung	358
11.3.5 Vielkanal-SBC	359
11.3.6 „Gehörrichtige“ Quantisierung	361
11.3.7 Codierung mit Polyphasen-Filterbänken	363
11.3.8 Realisierungen	364
11.4 Sinusmodellierung und Harmonische Codierung	365
11.4.1 Prinzip der Sinusmodellierung	365
11.4.2 Prinzip der Harmonischen Codierung	368
11.4.3 Probleme und Lösungsansätze	368
11.4.4 Realisierung, Aufwand, Qualität	372
11.4.5 Multiband-Codierung (MBE)	373
12 Geräuschreduktion	377
12.1 Begriffsklärung und Motivation	377
12.2 Ansätze	379
12.3 Einkanaliges Optimalfilter (Wienerfilter)	380
12.3.1 Ansatz und Zeitbereichslösung	380
12.3.2 Frequenzbereichslösung	382
12.3.3 Realisierung, Adaption	384
12.4 Spektrale Subtraktion	387
12.4.1 Ansatz und Zusammenhang mit dem Optimalfilter	387
12.4.2 Realisierung, Adaption	390

12.4.3	Reststörungen: Musical Tones	393
12.4.4	Variation der „Subtraktionsregel“	396
12.5	Verwendung verallgemeinerter Spektraldarstellungen	398
12.6	Eigenwert-/Eigenvektor-orientierte Geräuschreduktion	398
12.7	Geräuschkompensation	404
12.7.1	Ansatz und Zeitbereichslösung	404
12.7.2	Frequenzbereichsüberlegungen	406
12.7.3	Adaptionsmöglichkeiten	407
12.8	Zweikanal-Geräuschreduktion	409
12.8.1	Vorüberlegungen	409
12.8.2	Kohärenzfunktion und Kompensationsgewinn	411
12.8.3	Zweikanalige Geräuschreduktion im Zeitbereich	416
12.8.4	Zweikanalige Geräuschreduktion im Frequenzbereich	417
12.9	Mehrkanal-Geräuschreduktion	419
12.10	Beamforming	422
12.10.1	Mikrofonarrays mit superdirektiven Richteigenschaften	424
13	Kompensation akustischer Echos	429
13.1	Aufgabenstellung und Lösungsansatz	429
13.2	Objektive Beurteilungskriterien	434
13.3	Adaptionsalgorithmus: LMS, NLMS	436
13.3.1	Zusatzmaßnahmen zur Verbesserung der Echodämpfung	451
13.4	Frequenzbereichsverfahren und Blockverarbeitung	457
14	Sprachsynthese	465
14.1	Sprachsynthese und akustische Mensch-Maschine-Kommunikation	465
14.2	Synthese auf segmentaler Ebene – Verkettung	468
14.2.1	Regeln versus natürlichsprachliche Daten; Koartikulation	468
14.2.2	Segmentale Einheiten und Elemente	470
14.3	Akustische Synthese	474
14.3.1	Parametrische Synthese	474
14.3.2	Synthese durch Signalmanipulation im Zeitbereich	477
14.3.3	Inventarerstellung. Auf dem Weg zum „Personal Synthesizer“	483
14.3.4	Verkettung in Systemen mit natürlichsprachlichen Bausteinen	484
14.3.5	Direkte Sprachsynthese vom Sprachkorpus	485
14.4	Zur Frage der Prosodie	487
14.4.1	Dauersteuerung	488
14.4.2	Intonationssteuerung	489
14.5	Einige ausgewählte Anwendungen	491
14.5.1	Einsatz der Sprachsynthese im Behindertenbereich	492
14.5.2	Multilinguale Systeme	493
14.5.3	Inhaltsgesteuerte Sprachsynthese (<i>Concept to Speech</i>)	494

15 Sprachsignal-Qualität	499
15.1 Problematik	499
15.2 Auditive Qualitätsbestimmung	500
15.2.1 Beurteilungsansätze	500
15.2.2 Mean-Opinion Score (MOS)	503
15.2.3 Anker-Beurteilungen, MNRU	503
15.2.4 Attributbewertungen	504
15.2.5 Faktoren- und Hauptkomponentenanalyse	504
15.3 Instrumentelle Qualitätsbestimmung	507
15.3.1 Problematik der ACR-Nachbildung	507
15.3.2 Zwei Vergleichsbasen	507
15.3.3 Bekannte Maße	509
15.3.4 Neuere psychoakustisch motivierte Ansätze	516
15.4 Evaluierung der Qualität von Sprachsynthesystemen	528
15.4.1 Evaluierung der Verständlichkeit und Verstehbarkeit	528
15.4.2 Bewertung der Natürlichkeit und zugehöriger Attribute	531
15.4.3 Beispiele von Qualitätsauswertungen	531
15.5 Schlußbemerkungen	536
 Anhang: Codec-Standards	 537
A.1 ITU-T/G.726: Adaptive Differential Pulse-Code Modulation	538
A.2 ITU-T/G.728: Low-Delay CELP Speech Coder	539
A.3 ITU-T/G.729: Conjugate-Structure Algebraic CELP-Codec	542
A.4 ITU-T/G.722: 7 kHz Audio Coding within 64 kbit/s	545
A.5 ETSI-GSM 06.10: Full-Rate Speech Transcoding	546
A.6 ETSI-GSM 06.20: Half-Rate Speech Transcoding	548
A.7 ETSI-GSM 06.60: Enhanced Full-Rate Speech Transcoding	550
A.8 INMARSAT: Improved Multi-Band Excitation Codec (IMBE)	552
A.9 ISO-MPEG1 Audio Codierung	554
 Literaturverzeichnis	 557
 Sachwortverzeichnis	 581