

Contents

Preface	xiii
Acknowledgements	xv
1 Introduction	1
2 Coding Strategies and Standards	5
2.1 Introduction	5
2.2 Speech Coding Techniques	6
2.2.1 <i>Parametric Coders</i>	7
2.2.2 <i>Waveform-approximating Coders</i>	8
2.2.3 <i>Hybrid Coding of Speech</i>	8
2.3 Algorithm Objectives and Requirements	9
2.3.1 <i>Quality and Capacity</i>	9
2.3.2 <i>Coding Delay</i>	10
2.3.3 <i>Channel and Background Noise Robustness</i>	10
2.3.4 <i>Complexity and Cost</i>	11
2.3.5 <i>Tandem Connection and Transcoding</i>	11
2.3.6 <i>Voiceband Data Handling</i>	11
2.4 Standard Speech Coders	12
2.4.1 <i>ITU-T Speech Coding Standard</i>	12
2.4.2 <i>European Digital Cellular Telephony Standards</i>	13
2.4.3 <i>North American Digital Cellular Telephony Standards</i>	14
2.4.4 <i>Secure Communication Telephony</i>	14
2.4.5 <i>Satellite Telephony</i>	15
2.4.6 <i>Selection of a Speech Coder</i>	15
2.5 Summary	18
Bibliography	18
3 Sampling and Quantization	23
3.1 Introduction	23

3.2	Sampling	23
3.3	Scalar Quantization	26
3.3.1	<i>Quantization Error</i>	27
3.3.2	<i>Uniform Quantizer</i>	28
3.3.3	<i>Optimum Quantizer</i>	29
3.3.4	<i>Logarithmic Quantizer</i>	32
3.3.5	<i>Adaptive Quantizer</i>	33
3.3.6	<i>Differential Quantizer</i>	36
3.4	Vector Quantization	39
3.4.1	<i>Distortion Measures</i>	42
3.4.2	<i>Codebook Design</i>	43
3.4.3	<i>Codebook Types</i>	44
3.4.4	<i>Training, Testing and Codebook Robustness</i>	52
3.5	Summary	54
	Bibliography	54
4	Speech Signal Analysis and Modelling	57
4.1	Introduction	57
4.2	Short-Time Spectral Analysis	57
4.2.1	<i>Role of Windows</i>	58
4.3	Linear Predictive Modelling of Speech Signals	65
4.3.1	<i>Source Filter Model of Speech Production</i>	65
4.3.2	<i>Solutions to LPC Analysis</i>	67
4.3.3	<i>Practical Implementation of the LPC Analysis</i>	74
4.4	Pitch Prediction	77
4.4.1	<i>Periodicity in Speech Signals</i>	77
4.4.2	<i>Pitch Predictor (Filter) Formulation</i>	78
4.5	Summary	84
	Bibliography	84
5	Efficient LPC Quantization Methods	87
5.1	Introduction	87
5.2	Alternative Representation of LPC	87
5.3	LPC to LSF Transformation	90
5.3.1	<i>Complex Root Method</i>	95
5.3.2	<i>Real Root Method</i>	95
5.3.3	<i>Ratio Filter Method</i>	98
5.3.4	<i>Chebyshev Series Method</i>	100
5.3.5	<i>Adaptive Sequential LMS Method</i>	100
5.4	LSF to LPC Transformation	101

5.4.1	<i>Direct Expansion Method</i>	101
5.4.2	<i>LPC Synthesis Filter Method</i>	102
5.5	Properties of LSFs	103
5.6	LSF Quantization	105
5.6.1	<i>Distortion Measures</i>	106
5.6.2	<i>Spectral Distortion</i>	106
5.6.3	<i>Average Spectral Distortion and Outliers</i>	107
5.6.4	<i>MSE Weighting Techniques</i>	107
5.7	Codebook Structures	110
5.7.1	<i>Split Vector Quantization</i>	111
5.7.2	<i>Multi-Stage Vector Quantization</i>	113
5.7.3	<i>Search strategies for MSVQ</i>	114
5.7.4	<i>MSVQ Codebook Training</i>	116
5.8	MSVQ Performance Analysis	117
5.8.1	<i>Codebook Structures</i>	117
5.8.2	<i>Search Techniques</i>	117
5.8.3	<i>Perceptual Weighting Techniques</i>	119
5.9	Inter-frame Correlation	121
5.9.1	<i>LSF Prediction</i>	122
5.9.2	<i>Prediction Order</i>	124
5.9.3	<i>Prediction Factor Estimation</i>	125
5.9.4	<i>Performance Evaluation of MA Prediction</i>	126
5.9.5	<i>Joint Quantization of LSFs</i>	128
5.9.6	<i>Use of MA Prediction in Joint Quantization</i>	129
5.10	Improved LSF Estimation Through Anti-Aliasing Filtering	130
5.10.1	<i>LSF Extraction</i>	131
5.10.2	<i>Advantages of Low-pass Filtering in Moving Average Prediction</i>	135
5.11	Summary	146
	Bibliography	146
6	Pitch Estimation and Voiced–Unvoiced Classification of Speech	149
6.1	Introduction	149
6.2	Pitch Estimation Methods	150
6.2.1	<i>Time-Domain PDAs</i>	151
6.2.2	<i>Frequency-Domain PDAs</i>	155
6.2.3	<i>Time- and Frequency-Domain PDAs</i>	158
6.2.4	<i>Pre- and Post-processing Techniques</i>	166
6.3	Voiced–Unvoiced Classification	178
6.3.1	<i>Hard-Decision Voicing</i>	178
6.3.2	<i>Soft-Decision Voicing</i>	189

6.4 Summary	196
Bibliography	197
7 Analysis by Synthesis LPC Coding	199
7.1 Introduction	199
7.2 Generalized AbS Coding	200
7.2.1 <i>Time-Varying Filters</i>	202
7.2.2 <i>Perceptually-based Minimization Procedure</i>	203
7.2.3 <i>Excitation Signal</i>	206
7.2.4 <i>Determination of Optimum Excitation Sequence</i>	208
7.2.5 <i>Characteristics of AbS-LPC Schemes</i>	212
7.3 Code-Excited Linear Predictive Coding	219
7.3.1 <i>LPC Prediction</i>	221
7.3.2 <i>Pitch Prediction</i>	222
7.3.3 <i>Multi-Pulse Excitation</i>	230
7.3.4 <i>Codebook Excitation</i>	238
7.3.5 <i>Joint LTP and Codebook Excitation Computation</i>	252
7.3.6 <i>CELP with Post-Filtering</i>	255
7.4 Summary	258
Bibliography	258
8 Harmonic Speech Coding	261
8.1 Introduction	261
8.2 Sinusoidal Analysis and Synthesis	262
8.3 Parameter Estimation	263
8.3.1 <i>Voicing Determination</i>	264
8.3.2 <i>Harmonic Amplitude Estimation</i>	266
8.4 Common Harmonic Coders	268
8.4.1 <i>Sinusoidal Transform Coding</i>	268
8.4.2 <i>Improved Multi-Band Excitation, INMARSAT-M Version</i>	270
8.4.3 <i>Split-Band Linear Predictive Coding</i>	271
8.5 Summary	275
Bibliography	275
9 Multimode Speech Coding	277
9.1 Introduction	277
9.2 Design Challenges of a Hybrid Coder	280
9.2.1 <i>Reliable Speech Classification</i>	281
9.2.2 <i>Phase Synchronization</i>	281
9.3 Summary of Hybrid Coders	281
9.3.1 <i>Prototype Waveform Interpolation Coder</i>	282

9.3.2	<i>Combined Harmonic and Waveform Coding at Low Bit-Rates</i>	282
9.3.3	<i>A 4 kb/s Hybrid MELP/CELP Coder</i>	283
9.3.4	<i>Limitations of Existing Hybrid Coders</i>	284
9.4	Synchronized Waveform-Matched Phase Model	285
9.4.1	<i>Extraction of the Pitch Pulse Location</i>	286
9.4.2	<i>Estimation of the Pitch Pulse Shape</i>	292
9.4.3	<i>Synthesis using Generalized Cubic Phase Interpolation</i>	297
9.5	Hybrid Encoder	298
9.5.1	<i>Synchronized Harmonic Excitation</i>	299
9.5.2	<i>Advantages and Disadvantages of SWPM</i>	301
9.5.3	<i>Offset Target Modification</i>	304
9.5.4	<i>Onset Harmonic Memory Initialization</i>	308
9.5.5	<i>White Noise Excitation</i>	309
9.6	Speech Classification	311
9.6.1	<i>Open-Loop Initial Classification</i>	312
9.6.2	<i>Closed-Loop Transition Detection</i>	315
9.6.3	<i>Plosive Detection</i>	318
9.7	Hybrid Decoder	319
9.8	Performance Evaluation	320
9.9	Quantization Issues of Hybrid Coder Parameters	322
9.9.1	<i>Introduction</i>	322
9.9.2	<i>Unvoiced Excitation Quantization</i>	323
9.9.3	<i>Harmonic Excitation Quantization</i>	323
9.9.4	<i>Quantization of ACELP Excitation at Transitions</i>	331
9.10	Variable Bit Rate Coding	331
9.10.1	<i>Transition Quantization with 4 kb/s ACELP</i>	332
9.10.2	<i>Transition Quantization with 6 kb/s ACELP</i>	332
9.10.3	<i>Transition Quantization with 8 kb/s ACELP</i>	333
9.10.4	<i>Comparison</i>	334
9.11	Acoustic Noise and Channel Error Performance	336
9.11.1	<i>Performance under Acoustic Noise</i>	337
9.11.2	<i>Performance under Channel Errors</i>	345
9.11.3	<i>Performance Improvement under Channel Errors</i>	349
9.12	Summary	350
	Bibliography	351
10	Voice Activity Detection	357
10.1	Introduction	357
10.2	Standard VAD Methods	360
10.2.1	<i>ITU-T G.729B/G.723.1A VAD</i>	361

10.2.2	<i>ETSI GSM-FR/HR/EFR VAD</i>	361
10.2.3	<i>ETSI AMR VAD</i>	362
10.2.4	<i>TIA/EIA IS-127/733 VAD</i>	363
10.2.5	<i>Performance Comparison of VADs</i>	364
10.3	Likelihood-Ratio-Based VAD	368
10.3.1	<i>Analysis and Improvement of the Likelihood Ratio Method</i>	370
10.3.2	<i>Noise Estimation Based on SLR</i>	373
10.3.3	<i>Comparison</i>	373
10.4	Summary	375
	Bibliography	375
11	Speech Enhancement	379
11.1	Introduction	379
11.2	Review of STSA-based Speech Enhancement	381
11.2.1	<i>Spectral Subtraction</i>	382
11.2.2	<i>Maximum-likelihood Spectral Amplitude Estimation</i>	384
11.2.3	<i>Wiener Filtering</i>	385
11.2.4	<i>MMSE Spectral Amplitude Estimation</i>	386
11.2.5	<i>Spectral Estimation Based on the Uncertainty of Speech Presence</i>	387
11.2.6	<i>Comparisons</i>	389
11.2.7	<i>Discussion</i>	392
11.3	Noise Adaptation	402
11.3.1	<i>Hard Decision-based Noise Adaptation</i>	402
11.3.2	<i>Soft Decision-based Noise Adaptation</i>	403
11.3.3	<i>Mixed Decision-based Noise Adaptation</i>	403
11.3.4	<i>Comparisons</i>	404
11.4	Echo Cancellation	406
11.4.1	<i>Digital Echo Canceller Set-up</i>	411
11.4.2	<i>Echo Cancellation Formulation</i>	413
11.4.3	<i>Improved Performance Echo Cancellation</i>	415
11.5	Summary	423
	Bibliography	426
	Index	429