

Index

- 8PSK, 227
- acceleration features, 15, 223
- ACELP coders, 97
- acoustic environment, 66
 - model, 74
- acoustic model, 20
- adaptive DPCM, 87
- additive noise, 67
- additive white Gaussian noise, 48, 67
- AFE front end, 197
- AMDF function, 93
- AMR coders, 99
- analysis-by-synthesis, 93
- Aurora
 - error detection algorithm, 222
 - mitigation algorithm, 223
- Aurora mitigation algorithm, 143
- Aurora working group, 3, 112
- AWGN channel, 48, 51, 227
- babble noise, 70
- backtracking, 26
- backward probability, 24
- bad frame indicator, 138
- bad frame noise, 157
- base station, 43
 - controller, 44
 - system, 44
- base transceiver system, 44
- Baum–Welch algorithm, 26
- BCH codes, 232
- bit allocation
 - for DSR, 114, 119
 - in transform coders, 91
- bit error mask, 50
- bit error rate, 49
- bitstream-based NSR, 105
- blind equalization, 211
- block codes, 134, 229
- block interleaving, 235
- Bluetooth, 42
- BPSK, 49, 227
- bursty channel, 52
- carrier-to-interference ratio, 49
- CDF matching, 190
- CDMA-2000, 42
- cellular structure, 43
- CELP coders, 96
- cepstral
 - coefficients, 14
 - distance, 14
- cepstrum, 13
- channel coding, 131
- channel noise, 72
- class index, 218
- classification-oriented quantization, 118
- CMN, 182
 - real-time, 183
 - segmental, 184
- CMVN, 187
- cocktail transformation, 104
- coding degradation, 103

- colored noise, 67
- conditional loss probability, 61
- constellation, 227
- context-independent phones, 30
- continuous HMM, 28
- continuous speech recognition, 9, 30
- continuously variable slope delta modulation, 87
- convolutional codes, 135, 232
- convolutional interleavers, 235
- CRC codes, 232
- critical bands, 13
- CS-ACELP coder, 99
- CTN, 189
- cyclic codes, 231

- DAMPS, 42
- DCT-based coding for DSR, 121, 123
- deletion, 9
- delta cepstrum, 15
- delta modulation, 87
- differential pulse code modulation, 86
- discrete cosine transform, 91
- discrete HMM, 29
- distance measure, 16
- distributed speech recognition, 4, 85, 112
- Doppler shift, 47
- DTW, 19
- dynamic features, 15
- dynamic time warping, 19

- EFR coder, 97
- EGPRS, 45
- embedded speech recognition, 1
- EPH
 - suppression rule, 173
- erasure channel, 134
- error concealment, 132, 141
- error correction, 135
- error detection, 133
- error mitigation, 132
- error rate, 9
- estimation, 145
- ETSI, 3
- EVRC coder, 98

- exponential feature weighting, 36, 154
- extended Gilbert channel, 64
- extension features, 198, 218
- extrapolation, 145

- fading, 46, 48
- FE front end, 197
- feature enhancement, 109
- feature extraction, 207
- Feature normalization, 182
 - nonlinear, 189
- feature pair, 217
- Federal Standard
 - 1016 (CELP), 96
 - LPC10e, 92
 - MELP, 92
- filterbank, 10
- forward error correction, 132
- forward MMSE estimation, 147
- forward probability, 24
- forward-backward algorithm, 24
- forward-backward MMSE estimation, 147
- FR coder, 95
- frame, 10
 - pair format, 220
- frame packet stream, 220
- frame pair, 217
- Frame-dropping, 174

- gain factorization, 204
- gaussian mixture models, 124
- Gilbert channel, 62
- Gilbert-Elliot channel, 52, 62
- GMM-based encoder, 124
- GMSK, 49, 227
- GPRS, 45
- grammar, 9
- GSM, 42, 43
 - error patterns, 50
 - traffic channels, 45

- Hamming codes, 230
- Hamming window, 11
- hangover, 176, 202
- hard decision, 53, 133, 148, 228

- HEQ, 191
 - on-line, 195
 - quantile-based, 194
- hidden Markov models, 7
- higher-layer-oriented channel models, 53
- HMM
 - definition, 22
- HTK, 7

- IEEE 802.11, 42
- iLBC coder, 99
- IMT-2000, 42
- initial state probability, 22
- insertion, 9
- interframe VQ, 118
- interleaving, 132, 139, 234
 - spread, 235
- interpolation, 142
- IP protocol, 54, 56
- IS-136, 42
- IS-54, 42
- IS-641 coder, 98
- IS-95, 42
- isolated word recognition, 9
- ITU-T
 - G.721, 88
 - G.722, 90
 - G.723, 88
 - G.726, 88
 - G.727, 88
 - G.728, 96
 - G.729, 99
 - G.729A, 99
 - G.729B, 99
 - G.729D, 104
 - G.729E, 104
 - G723.1, 99

- k-means algorithm, 17
- Karhunen–Loewe transform, 90
- KLT-based coding for DSR, 124

- language, 9
- language model, 20
- LBG algorithm, 17
- LBND detector, 215

- left-to-right model, 29
- LFCC coefficients, 14
- liftering, 14
- line spectrum pairs, 97, 99, 108, 110, 111, 121, 226
- linear frequency cepstrum coefficients, 14
- linear interpolation, 142
- linear prediction, 10
- LMS algorithm, 110, 211
- log-area ratios, 95, 106, 226
- log-energy, 15, 208, 223
- long-term predictor, 95
- LPC, 11
 - analysis, 11
 - cepstral coefficients, 14
 - order, 12
 - spectrum, 11
 - speech production model, 11, 91
- LPC coefficients, 11
- LPCC coefficients, 14

- MAP
 - decision rule, 20
- MAP estimation, 150
- MAPLR, 35
- marginalization, 38, 152
- Markov channel models, 54, 62, 63, 65
- Markov process, 20
- Max-Log-MAP algorithm, 234
- maximum a posteriori linear regression, 35
- MCE estimation, 27
- mean opinion score, 86
- media-independent FEC, 132, 229
- media-specific FEC, 132, 140
- mel-frequency
 - cepstral coefficients, 14
 - filterbank, 13, 203
- mel-IDCT, 205
- MFCC coefficients, 14, 208
- missing data techniques, 37, 152
- mixture, 28
- ML decision rule, 228
- ML estimation, 26
- MLLR adaptation, 33, 104
- MMI estimation, 27

- MMSE estimation, 146
 - for erasure channels, 149
 - for wireless channels, 147
- MMSE-LSA estimation, 110
- mobile station, 44
- model adaptation, 32
- modulation, 227
- Modulation frequency, 185
- MPEG-4, 96
- multiframe format, 220
- multipath effect, 46
- multiple access interference, 47
- multiple description coding, 109
- multipulse coders, 94
- multistage vector quantization, 116
- MUSICAM algorithm, 90

- nearest neighbor rule, 16
- network speech recognition, 85, 100
- network subsystem, 44
- nonlinear interpolation, 142
- nonstationary noise, 68

- observation probability, 22
- offset compensation, 206
- OQPSK, 227
- out-of-vocabulary detection, 156

- packet loss, 58
 - Bernoulli model, 61
 - four-state Markov model, 64
 - higher-order Markov models, 65
 - rate, 61
 - three-state Markov model, 63
 - two-state Markov model, 62
- packet switching, 54
- packet trace, 61
- path loss, 46
- pattern matching, 18
- payload, 54
 - format, 58, 221
- PDC, 42
- perceptual linear prediction, 120
- physical-layer-oriented channel models, 51
- pink noise, 67

- pitch, 198
 - smoothing, 223
 - tracking, 223
- pitch estimation, 92, 214
- pitch quantization, 218
- pre-emphasis, 10
- probability of symbol error, 49
- product codes, 113
- pseudo-cepstrum, 111
- pulse code modulation, 86

- QCELP coder, 97
- QMF filters, 88, 208
- QPSK, 49, 227
- quefrency, 14

- RASTA, 187
- Rayleigh fading, 47, 48
 - channel, 51
- RCELP coder, 98
- recognition unit, 8
- recovery techniques, 131
- Reed–Solomon codes, 137, 232
- reflection coefficients, 225
- regression class trees, 33
- regular pulse excitation, 95
- remote speech recognition, 2
- reverberation, 72
- Ricean fading, 48
- root-cepstrum, 15, 103
- round trip time, 60
- router, 59
- RPE-LTP coder, 95
- RTCP protocol, 58
- RTP protocol, 56, 57

- sampling frequency extension, 208
- scalability, 126
- scalar quantization for DSR, 113
- semicontinuous HMM, 29
- sender-driven techniques, 131
- server feature processing, 223
- short-term predictor, 92
- signal space, 227
- signal vector, 227
- soft bit, 147

- soft data, 38, 153
- soft decision, 53, 133, 147, 228, 234
- SOVA algorithm, 234
- speaker variability, 32
- speaker-dependent ASR, 8
- speaker-independent ASR, 8
- spectral comb analysis, 93, 216
- spectral subtraction
 - magnitude, 168, 202, 204
 - musical noise, 169
 - nonlinear, 172
 - power, 167
- speech coding, 86
- speech-enabled services, 5
- split vector quantization, 114, 217
- stationary noise, 68
- subband coding, 88
- substitution, 9
- subword units, 8

- tandem free operation, 103
- tandeming, 102
- TCP protocol, 55, 57
- TCP/IP protocol suite, 55
- transform coding, 90
- transition probability, 22
- TRAU, 45
- tree-structured VQ, 17, 117
- triphones, 9, 30
- TSVQ, 17

- UDP protocol, 56, 57
- UMTS, 42
- unconditional loss probability, 61
- unequal error protection, 136
 - for lossy packet channels, 137
 - for wireless channels, 136

- VAD, 174
 - full-band, 175
 - LTSD, 179
 - LTSE, 179
 - MO-LRT, 180
 - noise estimation, 177
 - statistical, 177
 - using long-term information, 178
- VAD flag, 198
- variability, 7
- vector quantization, 16
 - for DSR, 113
- velocity features, 15, 223
- Viterbi algorithm, 25
- Viterbi decoding
 - with missing data, 37
 - with soft data, 38
- vocoder, 91
- voice activity detection, 212
- voicing class, 198, 217
- VSELP coder, 97

- waveform processing, 206
- weighted Viterbi algorithm, 36, 153
- weighting filter, 95
- Wiener filter, 161, 201
 - FIR, 162
 - frequency domain, 164
 - noise reduction, 165
- wireless transmission, 48
- word accuracy, 9
- word error rate, 9

- XAFE front end, 197
- XFE front end, 197

- zonal sampling, 91