

Index

- A posteriori probability, 180, 181, 192
- Accessibility, 3
- Acoustic mismatch, 9, 11
- Adaptation, 228
- Anti-HMM, 211
- Anti-model, 211
- Antisymmetric features, 79
- ASR, 8
- AT, 111, 118, 124
- ATM, 223
- Auditory
 - features, 135, 136
 - filters, 116
- Auditory transform, 111
 - dilation variable, 139
 - discrete-time, 123
 - fast inverse transform, 123
 - fast transform, 123
 - inverse, 120, 124
 - scale variable, 139
 - shift variable, 139
 - translation variable, 139
- Auditory-based
 - feature extraction, 135, 136, 138
 - transform, 111, 118
- Authentication, 1
 - biometric-based, 2, 3
 - human-human, 2
 - human-machine, 2
 - information-based, 2, 5, 6
 - machine-human, 2
 - machine-machine, 2
 - token-based, 2
- Automatic enrollment, 224
- Automatic speech recognition, 8
- Average individual equal-error rates, 174
- Background model, 153
- Background noise, 114
- Backpropagation, 44, 181, 192
- Band number, 139
- Bank application, 223
- Bark scale, 113, 116, 128, 131
- Basilar membrane, 116, 138
- Batch-mode process, 76
- Baye
 - decision rule, 69
- Bayes, 181
- Bayesian
 - decision theory, 68, 179
- Beam
 - search, 93
 - width, 93
- Benchmark, 38
- Central frequency, 139
- Centroids, 27
- Cepstral domain, 158
- Cepstral mean subtraction, 152, 158
- CFCC, 127, 140, 147
- Change-point
 - detection, 76, 95–97
 - problem, 95
 - state detection, 97
- Classification rules, 55
- Closed test, 9

- CMS, 152, 158, 226
- Cochlea, 116
- Cochlear
 - filter, 120, 139
 - filter bank, 138
 - filter cepstral coefficients, 140
- Cochlear filter cepstral coefficients, 127
- Code vector, 30
- Codewords, 27
- Cohort normalization, 172
- Combined system, 225
- Common thresholds, 228
- Communications, 75
- Computation noise, 114
- Conditional-likelihood functions, 182
- Confidence measure, 211
- Cryptographic protocols, 5
- Cubic lattice, 29
- Cubic root, 141

- Data space, 28
- Decision
 - boundary, 63
 - diagram, 75
 - parallel, 5
 - sequential, 5
- Decision procedure, 2
 - parallel, 2
 - sequential, 2
- Decoder
 - complexity, 93
 - complexity analysis, 103
 - detection-based , 93
 - forward-backward, 95
 - optimal path, 95
 - search space alignment, 94
 - search-space reduction, 93
 - subspace, 95
- Decoding algorithm, 94
- Decomposed frequency bands, 127
- Decryption, 2
- Denoising, 127
- DET, 4
- Detection, 93
- Dialogue procedure, 10
- Direct median method, 30
- Direct method, 10
- Discriminant analysis, 44
- Discriminant function, 44
- Discriminative
 - training, 192
 - training objectives, 179
- Distinctiveness, 3
- Distortion, 30
- Distribution estimation, 192

- Eardrum, 116
- Eavesdropping, 11
- EER, 71, 155, 156, 213
- Eigenvalue, 26, 50
- Eigenvector, 30
- EM algorithm, 62, 191
- Encoding error, 28
- Encryption, 2
 - key, 4
 - voice key, 4
- Endpoint detection, 75, 76, 86
 - beginning edge, 78
 - beginning point, 78
 - decision diagram, 82
 - ending edge, 78
 - ending point, 78
 - energy normalization, 83
 - filter, 78, 82
 - real-time, 81
 - rising edge, 78
- Energy
 - feature, 78
 - normalization, 75, 86
- Enrollment, 1, 224
 - session, 4, 7, 154, 157
- Equal loudness function, 130
- Equal-error rate, 155, 156, 213
- Equivalent rectangular bandwidth, 120
- ERB, 116, 120, 131, 139
- Error rate, 180
- Estimation
 - covariance matrices, 196
 - mean vectors, 197
 - mixture parameters, 198
- Evolution, 95
- Expectation-maximization algorithm, 191

- FA, 213
- False
 - acceptance, 71, 213
 - rejection, 71, 213

- Fast discriminative training, 191
- Fast Fourier transform, 111
- Fast training algorithm, 193
- FFT, 111
- Filter
 - bandwidth, 111
 - bank, 136
- Finite-state grammars , 87
- Fisher's LDA, 44
- Fixed
 - pass phrase, 8
 - phrase, 8
- Flops, 34
- Forced alignment, 152
- Forward auditory transform, 138
- Fourier transform, 111
- FR, 213
- Full search, 93

- Gammatone
 - filter, 130
 - filter bank, 117, 142
 - function, 111, 117, 130
 - GFCC features, 142
 - MGFCC features, 142
 - transform, 117
- Gaussian
 - classifier, 54
 - correlated, 29
 - node, 54
 - uncorrelated, 29
- Gaussian mixture model, 23, 62, 153
- GD, 191
- Generalized minimum error rate, 179
- Generalized probabilistic descent, 67
- Generalized Rayleigh quotient problem, 51
- GMER, 179, 185, 191
- GMM, 23, 62, 153
- GPD, 67
- Gradient descent, 44, 191

- Hair cell, 116, 140
- Hearing system, 115
- Hidden Markov model, 66
 - left-to-rightl, 97
 - SD, 155
 - Search-space reduction, 100
- Hidden nodes, 43

- High density, 28
- Histogram, 29
- HMM, 66
 - left-to-right, 97
 - SD, 155
 - Search-space reduction, 100
- Holder norm, 182
- HSV, 166
- Hypercube, 29
- Hyperplane, 53

- Identification, 4
- Impostor, 9
- Impulse response
 - cochlear filters, 120
 - plot, 120
- Indirect method, 10
- Inner ear, 116
- Internet, 5
- Invariant, 29
- Inverse AT, 124
- Inverse auditory transform, 124
- Inverse transform, 30

- Joint probability distribution, 61

- K-means, 23, 28, 32
 - algorithm, 27
- Key information, 207
- Keyword spotting, 209

- Laplace
 - distribution, 37
 - source, 30, 37
- LBG, 27, 34
- LDA, 46, 51, 166
- Left-to-right HMM, 67, 152
- Likelihood
 - ratio, 71, 152
 - ratio test, 154
- Linear discriminant analysis, 51, 166
- Linear prediction cepstral coefficients, 136
- Linear separable, 168
- Linear transform, 157, 158, 160
 - fast estimation, 160
- LLR, 211
- Local optimum, 28
- Log likelihood

- ratio, 162, 211
- score, 70
- LPC, 152
- LPCC, 136
- MAP, 181
- Maximum a posteriori, 181
- Maximum allowable distortion, 29
- Maximum mutual information, 183
- MCE, 179, 188
- Mean-squared error, 35
- Measurability, 3
- Mel frequency cepstral coefficient, 136
- Mel frequency cepstral coefficients, 135, 136, 204
- Mel scale, 116
- MER, 179, 188
- MFCC, 135, 136, 144, 204
- MGFCC, 142, 144
- Middle ear, 116
 - transfer function, 116
- Minimum classification error, 179
- Minimum error rate, 179, 181
 - classification, 69
- Mismatch, 157, 158, 160
- ML, 179
- MMI, 179, 183, 188
- Modified GFCC, 142
- Modified radial basis function network, 58
- Mother's maiden name, 10
- Mother's maiden name, 5
- MRBF, 58
- MSE, 35–37
- Multidimensional, 36
- Multiple paths, 101
- Multispectral pattern recognition, 58
- Multivariate
 - Gaussian, 23
 - Gaussian distribution, 23
 - statistical analysis, 27
 - statistics, 23
- NDA, 165
- Neural networks, 44
- Neyman Pearson, 211
- Neymann Pearson
 - lemma, 210
 - test, 70
- NIST, 4
- Non-keyword rejection, 209
- Non-stationary
 - pattern recognition, 61
 - process, 61, 95
- Normalized confidence measures, 211
- Normalized discriminant analysis, 165, 166, 168
- Objective, 179
 - function, 179
 - generalized minimum error rate, 179, 185
- GMER, 179, 193
- HMM training, 182
- maximum mutual information, 179, 183
- maximum-likelihood, 179
- MCE, 179, 181
- MER, 179
- minimum classification error, 179, 181, 183
- minimum error rate, 179
- MMI, 179
- One-pass VQ, 23, 28
 - algorithm, 32
 - codebook design, 34
 - complexity analysis, 33
 - data pruning, 28
 - direct-median method, 30
 - principal component method, 30
 - pruning, 31
 - robustness, 38
 - updating, 32
- One-to-many, 4
- One-to-one, 4
- Open test, 9
- Open-set evaluation, 154
- Operator, 9
- Optimal Gaussian classifier, 53
- Outer ear, 116
 - transfer function, 116
- Outlier, 29, 39
- Oval window, 116
- Parametric model, 95
- Pass-phrase, 8
- Pattern recognition, 43
- PCA, 23, 25

- Perceptual linear predictive, 135, 136
- Personal data profile, 11
- Personal identification number, 2
- Personal information, 11
- PFC, 44
- PFN, 43, 44
 - design procedure, 44
- Phoneme alignment, 227
- PIN, 2
- Pitch harmonics, 114
- PLP, 135, 136, 147
- Pooled thresholds, 228
- Posterior probability, 180, 181
- Pre-emphasize, 152
- Principal component, 25
 - analysis, 23, 25
 - discriminant analysis, 53
 - hidden node design, 52
 - method, 30
- Principal feature, 44
 - classification, 44
- Principal feature network, 43, 44
 - classified region, 46
 - construction procedure, 48
 - contribution array, 56
 - decision-tree implementation, 46
 - Fisher's node, 48, 51
 - Gaussian discriminant node, 49
 - hidden node, 46
 - hidden node design, 48
 - lossy simplification, 56
 - maximum SNR hidden node, 55
 - parallel implementation, 46, 48
 - PC Node, 53
 - principal component node, 48
 - sequential implementation, 48
 - simplification, 56
 - thresholds, 56
 - unclassified region, 46
- Pruning, 28, 44, 95

- Q-factor, 118, 120

- Radial basis network, 46
- Radius, 28
- RASTA, 147
- RASTA-PLP, 135, 136, 147
- RBF, 46
- Rectifier function, 120

- Recurrent neural network, 61
- Recursive, 97
- Registration, 1
- Repeatability, 3
- Resynthesis, 117
- Robust interval, 217
- Robust speaker identification, 135
- Robustness, 3
- Rotation, 158, 161
- RST, 161

- Scale, 158, 161
- SD, 7, 9, 154, 209
- SD target model, 228
- Search path, 95
- Searching, 93
- Segmental K-means, 39, 67
- Sequential
 - design algorithm, 30
 - detection, 95
 - detection scheme, 96
 - observations, 69
 - probability ratio test, 95
 - procedure, 44
 - process, 76
 - questions, 208
 - utterance verification, 212
- SI, 154, 209
- SID, 7
- Single Gaussian distribution, 24
- Single path, 101
- Single utterance
 - ASR, 209
 - verification, 209
- SNR, 44, 142
- Source histogram, 29
- Speaker authentication, 6
 - progressive integrated, 11
 - system, 223
- Speaker dependent, 7, 209
- Speaker identification, 7, 204
- Speaker independent, 209
- Speaker recognition, 6, 7
 - text constrained, 8
- Speaker verification, 7
 - background score, 154
 - connected digit, 166
 - contex-dependent, 152
 - enrollment phase, 152

- fixed phrase, 151, 152, 158, 226
- fixed-phrase, 165
- general phrase, 161
- global variance, 163
- hybrid, 174
- Hybrid system, 166
- language-dependent, 156
- language-independent, 151, 156
- pass-phrase, 151, 155
- randomly prompted, 165, 166
- robust, 157
- speaker dependent, 152
- speaker independent, 152
- stochastic matching, 161
- target model, 163
- target score, 154
- test phase, 152
- text dependent, 161
- text independent, 8
- text prompted, 8
- whole-phrase model, 152
- Speaker-dependent threshold, 217
- Speaker-independent models, 9
- Spectrogram, 113
 - AT, 120
 - auditory transform, 120
 - FFT, 120
- Spectrum, 129
- Speech activity detection, 76
- Speech detection, 76
- Speech recognition, 6
- Speech segmentation, 68
- Speech segments, 78
- Speech-sampling rate, 78
- Speed up, 34, 37
- SPRT, 95
- SR, 6, 7
- State change-point detection, 104
- State transition diagram, 82
- Stationary
 - pattern recognition, 61
 - process, 61
- Statistical properties, 3
 - non-stationary, 3
 - stationary, 3
- Statistical verification, 70
- Step-down procedure, 212
- Stochastic matching, 157
 - fast, 158
 - fast algorithm, 158
 - geometric interpretation, 158
 - linear transform, 158
 - rotation, 158
 - scale, 158
- Stochastic process, 61
- Subband signals, 138
- Subtest, 212
- Subword
 - hypothesis testing, 210
 - model, 210
 - sequence, 210
 - test, 211
- SV, 7
- Synthesized speech, 126
- System design, 223
- Target model, 154
- Test session, 7, 154
- Testing, 1
 - session, 4
- Text to speech, 10, 208
- Time-frequency
 - analyses, 117
 - auditory-based analysis, 137
 - transforms, 111
- Training procedure, 200
- Translation, 158, 161
- Traveling wave, 116
- TTS, 10, 208
- Uncorrelated Gaussian source, 34
- User authentication, 2
- Utterance verification, 71, 209
- Variable length window, 140
- Vector quantization, 27
- Verbal information verification, 6, 9, 207
- Verification, 1, 4
- Verification session, 158
- Viterbi, 67, 68
 - algorithm, 93
 - search, 94
- VIV, 6, 9, 207, 208, 223, 224
- Voice characteristics, 6
- VoIP, 75, 77
- VQ, 23, 27, 28
- Warped frequency, 128

Wavelet transform, 117

Window shifting, 128

Wireless communication, 5

Word error rates, 86

WT, 117