

# Index

- Absolute difference 85
- Absolute maximum amplitude 222
- Acceleration ( $\Delta\Delta$ ) 214
- Accuracy 113
- Acoustic analysis 104, 105
- Acoustic class model 260
- Acoustic class 250
- Acoustic feature vectors 237, 244
- Acoustic frames 105
- Acoustic measurements 141
- Acoustic model 106, 143, 150
- Acoustic observation 104
- Acoustic score 107, 138
- “Acoustic segmentation” 110
- Acoustic similarities 135, 156
- Acoustic vectors 213
- Acoustical segment 146
- AcousticScore 163
- Activation function 71
- AFF 37
- AFF descriptors 37
- AFF plot 38
- Affix 116
- Agglomerative clustering 237
- AH 36
- AH features 36, 37
- All-pole 235
- Amplitude 40
- Amplitude normalization 212
- Amplitude spectrum 42, 43, 80
- AnalogClip* 227
- Analysis 2
- Analysis windows 42
- Annotation tool 50
- Application scenario 4
- Arithmetic mean 31, 32
- Artificial intelligence 70
- ASB 77
- ASC 25, 28, 36, 46, 49, 75
- ASE descriptor 37
- ASE extraction procedure 75
- ASE feature 59
- ASF coefficients 32
- ASF vectors 32
- ASP 77
- ASP features 248, 258
- ASR 11, 124, 161
- ASR transcription machine 137
- ASS 29, 36
- Attack 7, 40, 41, 171
- Attack, Decay, Sustain, Release (ADSR) 39, 40
- Attack phase 40
- Attack portion 40
- Attack time feature 41
- Attack volume 40
- Attributes 9, 18, 19, 20, 22
- Audio 3
- Audio analysis 2, 59, 65
- Audio and video retrieval 2
- Audio attribute 164
- Audio broadcast 4
- Audio class 77, 258
- Audio classification 50, 66, 71, 74
- Audio classifier 32
- Audio content 259
- Audio content analysis 1, 2
- Audio content description 2
- Audio description tools 6
- Audio event detection 259
- Audio events 50
- Audio feature extraction 1, 74
- Audio feature space 72
- Audio features 13, 259
- Audio fingerprinting 8, 11, 207
- Audio fundamental frequency (AFF) 13, 36
- Audio fundamental frequency (AFF) descriptor 33, 36
- Audio harmonicity (AH) 13, 33

- Audio harmonicity (AH) descriptor 33
- Audio indexing 84
- Audio-on-demand 4
- Audio power (AP) 24
- Audio power descriptor 5
- Audio segmentation 52, 127
- Audio signal 51
- Audio signature 6
- Audio signature description scheme 7, 8
- Audio spectrum basis (ASB) 13, 49, 73
- Audio spectrum centroid (ASC) 13, 27
- Audio spectrum envelope (ASE) 13, 24
- Audio spectrum envelope (ASE) descriptor 75
- Audio spectrum envelope (ASE) features 247
- Audio spectrum flatness (ASF) 13, 29
- Audio spectrum projection (ASP) 13, 49, 60, 73
- Audio spectrum projection (ASP) feature extraction 74
- Audio spectrum spread (ASS) 13, 29
- Audio tools 3
- Audio watermarking 208
- Audio waveform 23, 50
- Audio waveform (AWF) audio power (AP) 13
- Audio waveform (AWF) descriptor 5, 23
- AudioBPM* 192
- AudioSegment* 50, 220
- AudioSegment* description 14
- AudioSignalQuality* 220
- AudioSignature* 217
- AudioSpectrumFlatness* 217
- AudioTempo* 192
- Audiovisual 103
- Audiovisual content 2
- Audiovisual documents 8
- Auditory model-based approach 197
- Auditory scene analysis (ASA) 197
- Auditory system 52
- Autocorrelation 33, 34, 223
- Autocorrelation function 35, 37, 43
- Automatic classification 79
- Automatic discrimination 91
- Automatic expansion 148
- Automatic expansion of queries 138
- Automatic extraction and description 2
- Automatic indexing 74
- Automatic sound identification 8
- Automatic speech recognition (ASR) system 8, 103, 265
- Automatic transcription 98
- Automatic word annotation 165
- Automatic word-to-phone transcription 144
- Average document transcription length 146
- Average linkage between groups 240
- Average linkage within groups 241
- Average power 16
- Average spectrum spread 47
- Average square 24
- Average zero crossing rate 215
- Back propagation (BP) algorithm 71
- Back vowel 141
- Background noise 209, 264
- BackgroundNoiseLevel* 221
- Balance* 221
- Band edge indexes 31
- Band periodicity 91
- Band representative vectors 213
- Bandwidth 214, 221
- Bark-scaled bands 213
- Basic descriptors 5, 13
- Basic signal parameters 13
- Basic spectral audio descriptors 6
- Basic spectral descriptors 6, 13, 24
- Basis decomposition 77
- Basis decomposition algorithm 75, 77
- Basis functions 75, 77
- Basis matrix  $C_E$  88
- Basis projection 75
- Basis vectors 88
- Baum–Welch 69, 78
- Baum–Welch algorithm 107, 245, 260
- Baum–Welch re-estimation training patterns 78
- Bayes' rule 106
- Bayesian information criterion (BIC) 234
- Beat* 202
- Beats per minute (BPM) 190, 215
- Best alignment 107
- Best-match 79
- Best-matching search* 130
- Best performance 89
- Best scoring 111
- Bhattacharyya distance 239
- BIC 244
- $\Delta BIC$  245
- Bigram phone language model 150
- Bigrams 110, 149, 152
- BiM 9
- BiM coders 10
- Bin 49
- Binary-coded representation 3
- Binary format 9
- Binary query–document similarity measure 132
- Binary weighting 152, 153
- Binary-valued vectors 131
- Blackboard architectures 197
- Block 121
- BlockRepeating* 227
- BOMP<sup>2</sup> dictionary 144
- Boolean matching searching 130, 132
- Bottom-up hierarchical clustering 240
- Boundary 73
- Boundary curve 73

- “Brightness” 7, 171
- BroadcastReady* 221
- Browsing 1, 4
- Brute-force 217
- Candidate peaks 45
- Canonical phonetic transcription 117
- Cardinalities 9
- Category 4, 60
- CDs MP3 1
- Centre frequencies 258
- Centre of gravity 27
- Centring 62
- Centroid 6
- Cepstral coefficient 52, 53, 54, 55
- Cepstrum 52
- Cepstrum coefficients (MFCCs) 100
- Cepstrum mean normalization (CMN) 214
- $C_E W$  88
- ChannelNo* 227
- Classes 59
- Classical expectation and maximization (EM) algorithm 96
- Classification 1, 2, 5, 51, 52, 61, 67, 74
- Classification methods 66
- Classification scheme 77, 78
- Classification systems 60
- Classifier model 61
- Classifiers 60, 61, 66, 73, 74
- Click 227
- ClickSegment* 227
- Cluster purity 245
- Clustering methods 141
- CMU dictionary<sup>1</sup> 144
- Coarse-level audio classification 232
- Codebook 216
- Coding artefacts 1
- Coefficient grouping procedure 31
- Coherence 172
- Collections of objects 4
- Comb filter 35
- Comment 227
- Compactness 210
- Comparison 74, 200
- Complete linkage 240
- Components 75, 116
- Compressed binary form 6
- Compression 9, 213
- Computation of retrieval scores 164
- Computational considerations 79
- Computational simplicity 210
- Concatenation 109
- Confidence 50, 209
- Confidence measure 37, 50, 123
- Conflation 137
- Confusion-based expansion 154
- Confusion probabilities 157
- Confusion statistics 118
- ConfusionCount* 117, 118
- ConfusionInfo* 115, 117, 121, 164
- ConfusionInfoRef* 121
- Conjugate gradient algorithm 71
- Consistency 99, 100
- Content analysis 11
- Content-based description 2
- Content-based retrieval 2
- Content descriptions 3
- Content identifier 8
- Content-integrity 210
- Continuous HMMs 69
- Continuous spoken requests 138, 144
- Convolution product 52
- Coordinate matching (CM) 201
- Coordination level* 132
- Cornell SMART system 136
- Correlated information 87
- Corresponding phone pairs 158
- Cosine retrieval function 152, 153
- Cosine similarity measure 132, 136
- Count distinct measure 201
- Covariance matrix 21, 63, 67, 69
- Covariance 21, 22
- Creation 3
- Critical-band analysis 213
- Cross-entropy estimation 216
- Cross-validation 242
- CrossChannelCorrelation* 221
- Cutting 241
- Cyclic Redundancy Checking (CRC) 208
- Data-adaptive techniques 198
- Data descriptor 1
- Data fusion 162
- Data merging 162
- Data mining framework 259
- Data variables 63
- Database 209
- DcOffset* 63, 221
- DCT 79
- DDL 9
- Decay 40
- Decay phase 40
- Decoders 10
- Decoding 74, 104
- Decomposition 64
- Decomposition algorithms 79
- Decorrelated MFCC features 79
- Decorrelation 213
- Decorrelation of features 77
- DefaultLattice 121
- DefaultSpeakerInfoRef* 122
- DefaultSpeakerRef* 122
- Degree of reliability 163
- Del* 118
- Deletion 118, 157, 201
- Deletion errors 113, 117, 147
- Delta coefficients* 54

- Dendrogram 241
- Denominator  $P$  106
- Derivative ( $\Delta$ ) 214
- Description 1, 2, 3, 11
- Description definition language (DDL) 3, 6, 9
- Description of a melody 201
- Description schemes (DSs) 3, 6
- DescriptionMetadata 115, 121
- Descriptors 1, 9, 32
- Detection approach 260
- Detection of goal events in 260
- Detection of target sub-strings 159
- DetectionProcess 227
- DFT 16, 49, 51, 213
- Diagonal covariance matrices 67
- Difference measure 201
- Digital analysis 1
- Digital audio 1
- Digital audio and video 2
- Digital audio waveforms 1
- DigitalClip 227
- DigitalZero* 227
- Dimensions 74, 94
- Dimension reduction 60, 80, 88, 248
- Dimensionality reduction methods 137
- Dimensionality reduction 61, 62, 65, 74, 75, 77, 88, 210
- Direct approach 92
- Direct measure* 202
- Discrete cosine transform (DCT) 53, 60, 213
- Discrete Fourier transform (DFT) 15
- Discrete HMM 159
- Discrimination capability 94
- Discrimination power 151, 210
- Discriminative objective functions 70
- Distance measures 176
- Distributional assumptions 69
- Divergence 66
- Divergence shape distance 239
- Document 124
- Document description 124
- Document representation 124
- Document selection in database 164
- Document transcription 159
- Document transcription, length of 146
- Domain-specific data 129
- Dot plot 202
- Doubled pitch errors* 36
- DP 155, 157
- DP matrix 158
- DropOut* 227
- DS 9
- DSD 244
- DSs 6, 9
- DVDs 1
- Dynamic features 97
- Dynamic programming 107, 201
- Dynamic programming algorithm 79
- Dynamic programming (DP) principle 147
- Edit distance 155, 158
- Efficiency 209
- Eigenvalue decomposition 63
- Elements 9
- EMIM 165
- Emphasize 81, 105
- Energy envelope 41, 75
- Envelope (*Env*) 6, 39
- Environmental sounds 52, 75
- Equal acoustic distance 156
- Ergodic 89, 249
- Ergodic HMM 250
- Ergodic HMM segmentation 264
- Ergodic topology 77
- Error back propagation (EBP) algorithm 71
- Error rates 111, 113, 129
- ErrorClass* 227
- ErrorEventList* 221
- Estimated power spectrum 49
- Euclidean distance  $\delta(a, b)$  99, 216
- Event candidate detection 260
- Excitation signal 52
- Expansion 153
- Expansion of the document representation 134
- Expectation maximization algorithm 67
- Expected mutual information measure (EMIM) 137
- Exploitation of multiple recognition outputs 162
- Exponential pseudo norm* (EPM), 216
- Extraction of sub-word indexing terms 162
- Extraction process 2
- Euclidean distance 85
- Face detection 71
- Factors 63
- False alarms 111
- Fast Fourier transform (FFT) 16, 213
- FastICA 64, 84, 100, 249, 258
- FastICA algorithm 88
- Feature dimensionality reduction 61
- Feature extraction 60, 73, 74, 75, 93, 212, 260
- Feature extraction method 75
- Feature matrix 61, 62
- Feature space 71, 72
- Feature vectors 1, 32, 52, 60, 61, 66, 77, 244
- Features 75
- Feedforward nets 70
- Feedforward NNs 71
- FFT 42, 53
- FFT bins 43, 44
- Filler models 111
- Film/video indexing 92

- Filter 52
- Filter banks 79, 258
- Filtering 1
- Final retrieval score 145
- Fingerprint 1, 7, 59, 60
- Fingerprint modelling 214
- First 19, 21
- First-order HMMs 69
- First-order Markov assumption 69
- Five-state left–right models 77
- Flat spectrum shape 32
- Flatness 6, 32
- Flatness features 30
- Flexible modelling 109
- F-measure  $F$  250
- Folk songs 177
- Formants 45
- Forward and backward HMM 250
- Foundation layer 13
- Fourier transform 52, 79, 105
- Frame-based analysis 42
- Frame index domain 39, 42
- Frame-level centroids *LHSC* 46
- Frame-level spectral variation values *LHSV* 48
- Frame-level spread values *LHSS* 48
- Frequency band coefficients *ASF (b)* 32
- Frequency bands 25, 32
- Frequency bins 44, 76
- Frequency resolution 30, 76
- Frequency-warped 83
- Front-end 211
- Functions 70
- Fundamental frequency  $f_0$  2, 6, 36, 37, 42, 43, 45, 91
- Fundamental periods 34, 35, 37, 42
- Fuzzy matching search 155
- Fuzzy sub-lexical matching 155
  
- Gamma distribution 69
- Gaussian density 67
- Gaussian distribution 69
- Gaussian mixture models (GMMs) 60, 66
- Gaussian radial basis functions 73
- Gender of speaker 93
- Generalized likelihood ratio (GLR) 239
- Genomics 201
- Genres 59
- Geometric mean 32
- Global redundancies 215
- Global retrieval framework 162
- Global values 42
- GLR 244
- GMMs 236, 245
- Goal detection 259
- Goal event segment selection 260
- Gram–Schmidt decorrelation 65
- Gram–Schmidt orthogonalization 64
  
- Granularity 209
- Grouped coefficients 32
- Grouping procedure 31
- GSM encoding/decoding 212
  
- Haar transform 213
- Hamming distance 216
- Hamming window 42
- Hamming window function 81, 247
- Hand-held speech recognition devices 130
- Handwriting recognition 68
- Hanning 53
- Hanning windowing 79
- Harmonic 88
- Harmonic peak detection algorithm 42
- Harmonic peaks 42, 43, 44, 45, 47, 48
- Harmonic power 33
- Harmonic ratio (HR) 2, 33, 91
- Harmonic signal 35
- Harmonic sounds 36, 87, 172, 177
- Harmonic spectra 42
- Harmonic spectral centroid (HSC) 13, 38, 45
- Harmonic spectral deviation (HSD) 13, 38, 47
- Harmonic spectral spread (HSS) 13, 38, 47
- Harmonic spectral variation (HSV) 13, 38, 48
- Harmonic structure 32, 44, 45, 48
- HarmonicInstrumentTimbre* 174
- Harmonicity 6, 32, 35, 214
- Harmonics 45
- HarmonicSpectralCentroid* 173, 174
- HarmonicSpectralDeviation* 173, 175
- HarmonicSpectralSpread* 173, 175
- HarmonicSpectralVariation* 174, 175
- Harmony 172
- Heterogeneous databases of spoken content metadata 163
- Hidden control neural network (HCNN) 70
- Hidden (intermediate) layers 70
- Hidden layer 71
- Hidden Markov models (HMMs) 60, 66, 68, 216
- Hidden units 71
- HiEdge* 24, 25, 30, 218
- Hierarchical agglomerative method 248
- Hierarchical approach 93
- Hierarchical approach with hints 92
- Hierarchical approach without hints 92
- Hierarchical clustering algorithm 141
- Hierarchies 8
- High-dimensional representation 49
- High-dimensional spectrum 88
- High zero crossing rate 91
- HighEdge* 81
- Higher-level information 68
- Highest layer–sound samples 95
- Highest maximum likelihood 93
- Highest maximum likelihood score 84

- High-level description 113
- High-level descriptors 4
- High-level tool 8
- High-pass 105
- Highlights extraction 259
- HiLimit* 37
- Hint experiments 93
- Histogram 69
- HMM 73, 143, 245
- HMM parameters 78
- HMM topologies 77, 89, 250
- HMM training module 77
- Homomorphic transformation 52
- HopSize* 15, 42
- Horizontal dimension  $E$  88
- HR 34, 35, 36
- HR coefficient 33
- HR descriptor 37
- HR values 34
- HSC 45, 46, 47, 49
- HSD 47
- HSS 48
- Human auditory system 213
- Human frequency perception 53
- Human perception 54
- Human voice 52
- Hybrid segmentation 243, 246, 252, 254
- Hybrid speaker-based segmentation system 244
- Hyperbolic tangent 71
- Hyperplanes 71, 72
- Hypotheses 112
  
- ICA 65, 75, 77
- ICA algorithms 64, 83
- Identification 209
- Image recognition 70
- Impulse response 52
- In-vocabulary words 128
- Inconsistency 100
- Independence assumptions 69
- Independent component analysis (ICA) 49, 60, 61, 63, 100
- Independent components 63
- Independent sources 64
- Independent sub-lexical indexing system 162
- Index 103
- Index and retrieve spoken documents 163
- Indexing 61, 125
- Indexing and browsing systems 264
- Indexing and retrieval process 140
- Indexing and retrieval system 125
- Indexing description tools 6
- Indexing features 128
- Indexing lexicon 266
- Indexing of audiovisual data 166
- Indexing recognizer 159
- Indexing symbols 141
  
- Indexing term set 138
- Indexing term similarities 132
- Indexing term statistics 131
- Indexing terms space 130
- indexing terms 128
- Indexing units 266
- Indexing vocabulary 129
- INFOMAX 64
- Information content maximization 88
- Information loss 77
- Information maximization 248
- Information retrieval (IR) 123
- Information theory 155
- Informedia project 162
- Inharmonic sounds 172
- Initial indexing term 136
- Initial state 68
- Initial state probabilities 68
- Input layer 70
- Input space 72, 73
- Input vector 71
- Ins* 118
- Insertion 117, 118, 147, 157, 201
- Insertion errors 113
- Instantaneous bandwidth* 29
- Instantaneous power 24
- Instantaneous values 42
- Instrument Timbre* 172
- Instrument timbre description 11
- Integer confusion counts 147
- Intensity–loudness 235
- Interference 212
- Intermediate level 165
- Internal hidden variables 85
- Invariance to distortions 210
- Inverse document frequency (IDF) 131, 136, 146
- Inverse normalized edit distance* 156
- IpaNumber 117
- IpaSymbol 117
- IR 146
- IR text pre-processing 136
- ISO Motion Picture Experts Group 2
- Isolated handwritten digit recognition 71
- Isolated query terms 138
- Isolated sound recognition 78
- IsOriginalMono* 221
- Iterations 90
- Iterative re-estimation 78
- Iterative relevance feedback* 125
  
- JADE 64
- Jazz 177
- Jitter* 227
  
- Karhunen–Loève (KL) 213
- Kernel 72
- Kernel-based SVM 71

- Kernel functions 73
- Kernel mapping 72
- Key 178, 181
- Keyword lexicon 255
- Keyword spotting 111, 138, 155
- Keywords 111, 258
- Kohonen self-organizing maps (SOMs) 70
- Kullback–Leibler (KL) distance 237
  
- Lag 34, 35
- Lagrange functional 72
- Lagrange multipliers 72
- Language model (LM), 107
- Large-vocabulary ASR 163
- Last 19, 21
- LAT 40, 41
- Latent variables 63
- Lattice expansion 154
- Lattice with an upper limit 121
- Lattices 112, 114
- $L$ -best likelihood scores 244
- LD 155
- Learning phase 73
- Left–right 89
- Left–right HMM 78, 250
- Left–right HMM classifiers 96
- Left–right models 106
- Legacy audio content 8
- Levenberg–Marquadt algorithm 71
- Levenshtein distance 155
- $LHSC_i$  45
- $LHSD$  47
- Likelihood 67, 69
- Likelihood distance 245
- Likelihood maximization 69
- Likelihood Selection block 245
- Likelihood values 68
- Linear and second-order curvature 97
- Linear bands 81
- Linear combination function 149
- Linear correlations 63
- Linear discriminant 70
- Linear prediction cepstrum coefficients (LPCCs) 52, 105
- Linear prediction coefficients (LPCs) 52, 235
- Linear spectral pairs (LSPs) 235
- Linear transformation 87, 213
- LinguisticUnit 165
- LLDs 32, 37, 52
- Local alignment 201
- Local HSC  $LHSC_i$  47
- Local spectra 47
- LoEdge 24, 25, 30, 218
- Log attack time (LAT) 13, 38, 40
- Log-energy 53, 54
- Log-frequency bands 30
- Log-frequency power spectrum 24, 27, 75
- Log-frequency spectrum 29, 37
  
- Log-scale octave bands 80
- Log-spectral features 73
- Logarithmic bands 81
- Logarithmic frequency descriptions 24
- Logarithmic maximum power spectrum (LMPS) 226
- LogAttackTime 173, 174, 176
- LoLimit 37
- Longest common subsequence 202
- Loudness 49, 52, 214
- Low-complexity description 75
- Low-dimensional feature 79
- Low-dimensional representation 49, 75, 88
- Low short-time energy ratio 91
- LowEdge 81
- Low-level audio descriptors 11
- Low-level audio features 14, 66
- Low-level descriptors (LLDs) 13, 50, 217
- Low-level features 3
- Low-level MPEG-7 descriptors 5
- Low-level signal-based analyses 8
- LVCSR 128, 136, 144, 161
- $L_w$  42
  
- Magnitude 225
- Magnitude spectrum 53
- Mahalanobis distance 238
- Manhattan distance 216
- MAP 153
- Matrix diagonal 118
- Max 19, 21, 22
- Maximal statistical independence 83
- Maximum 19
- Maximum accumulative probability 79
- Maximum coefficients 21
- Maximum fundamental period 33, 34
- Maximum likelihood 236
- Maximum likelihood criteria 118, 147
- Maximum likelihood score 79, 89, 96, 244
- Maximum samples 23
- Maximum squared distance MSD 21
- MaxRange 23
- MaxSqDist 21, 22
- Mean 19, 20, 21, 22
- Mean average precision (mAP), 151
- Mean of original vectors 21
- Mean power 222
- Mean sample value 19
- Mean square error 71
- Measure of acoustic similarity 146, 158
- Measure of similarity 126
- MediaTime 121
- Mel 52
- Mel-filter bank 53
- Mel-frequency 81, 100
- Mel-frequency cepstrum coefficients (MFCCs) 50, 52, 213
- Mel-frequency scale 52
- Mel scale 53

- Melody 177
- Melody contour representation 8
- Melody contours 8, 200
- Melody description 6
- Melody description tools 8
- Melody DS* 171, 177
- Melody sequence 8
- MelodyContour* 178, 182
- MelodyContour DS* 183, 185
- MelodySequence* 178
- Mel-scaling 80
- Mel-spaced filter bank 53
- Mercer condition 72
- Merging 244
- Message Digest 5 (MD5) 208
- Metadata 2, 8, 103
- Meter* 178
- Metric 216
- Metric-based model-based or hybrid segmentation 233
- Metric-based segment 251
- Metric-based segmentation 237, 252
- MFCC +  $\Delta$  +  $\Delta\Delta$  97
- MFCC features 246, 252
- MFCC vector 54, 55
- MFCCs 54, 55, 60, 61, 79, 254
- Mid-level statistical models 8
- Min 19, 21, 22
- MinDuration* 50
- MinDurationRef* 50
- Minimum 19, 23
- Minimum block power 223
- Minimum coefficients 21
- Minimum fundamental frequency 33
- MinRange* 23
- Missed words 111
- MissingBlock* 227
- Mixtures of Gaussian densities 83
- MLP function 73
- Model 61, 67
- Model-based classifiers 66
- Model-based resegmentation 243, 246
- Model building and model-level clustering 245
- Model-cluster 244
- Model-level clustering 243, 245
- Model of highlights 260
- Model selection 242
- Model-selection-based method 234
- Modified HR 35
- Modified power spectrum 28, 29
- Modulation 214
- Module 265
- Monograms 149
- Monophonic melody transcription 194
- Morpheme 116
- MP3 208
- MPEG-7 2, 3, 37, 49, 50, 217
- MPEG-7 ASP 79, 88, 246
- MPEG-7 audio 50
- MPEG-7 audio classification tool 77
- MPEG-7 audio group 75
- MPEG-7 audio LLD 66
- MPEG-7 audio spectrum projection (ASP) 100, 236
- MPEG-7 basis function database 77
- MPEG-7 descriptor 84, 256
- MPEG-7 feature extraction 75
- MPEG-7 features 262
- MPEG-7 foundation layer 13
- MPEG-7 LLDs 39, 91
- MPEG-7 low-level audio descriptors 5, 77
- MPEG-7 *MelodyContour* 203
- MPEG-7 "Multimedia Content Description Interface" 2
- MPEG-7 NASE 86
- MPEG-7 *Silence* 50
- MPEG-7 *Silence* descriptor 50
- MPEG-7 sound recognition classifier 74
- MPEG-7 sound recognition framework 73
- MPEG-7 Spoken Content description 104
- MPEG-7 *SpokenContent* 113, 163
- MPEG-7 *SpokenContent* high-level tool 266
- MPEG-7 *SpokenContent* tool 265
- MPEG-7 standard 3, 34, 37, 40, 42, 49, 52, 60, 61, 73
- MPEG-7 standardized features 254
- MPEG-7 timbre description 38
- Multi-dimensional 100
- Multi-dimensional data 73
- Multi-phone units 141
- Multi-hypotheses lattices 157, 164
- Multi-hypothesis 154
- Multi-layer feedforward perceptron 71
- Multi-layer perceptron (MLP) 70
- Multi-level descriptions 166
- "Multi-pitch estimation" 197
- Multigrams 142
- Multimedia applications 91
- Multimedia content 3
- Multimedia documents 8
- Multimedia mining indexer 231
- Multimedia mining server 231
- Multimedia mining system (MMS) 231
- "Multiple fundamental frequency estimation" (MFFE) 197
- Multiple Index Sources 161
- Multiplexing 3, 10
- Multiplicative divergence update 66
- Music 1, 45, 50, 52, 59, 171
- Music description tools 11, 171
- Music genre classification 51
- Music indexing 98
- Music information retrieval (MIR) 193
- Music sounds 75
- Musical instrument classification 85
- Musical instrument timbre tool 7

- Musical instrument timbres 6, 49
- Musical instruments 45, 171
- Musical signals 52, 171
- Musical sound 39
- Musical timbre 6
  
- NASE 75, 76, 88
- NASE features 79, 248
- NASE matrix 77, 94
- NASE vectors 75, 76
- NASE  $X(f, l)$  87
- Natural logarithmic operation 81
- Navigate 1
- N*-best 149, 151
- N*-best list 112
- N*-best transcriptions 112, 164
- Nearest neighbour 216
- Neural networks (NNs) 60, 66, 70
- Neurons 70, 71
- Newton's method 64, 65
- n*-gram confusion matrix 146
- n*-gram level 147
- n*-gram matching technique 137
- n*-gram models 109
- n*-gram similarities 157
- n*-gram techniques 200
- n*-grams 120, 141, 145, 152, 153, 161, 201
- NMF 94
- NMF method 1* 94
- NMF method 2* 94
- NN classifier 70
- Node 121, 122
- Noise 50
- Noise-like 29
- Noise-like sound 87
- Noise reduction method 261
- Non-coherent sounds 172
- Non-emitting states 89
- Non-harmonic sounds* 36
- Non-harmonicity tolerance 44
- Non-keyword 111
- Non-linear discrimination 72
- Non-linear function 70
- Non-linear technique 83, 88
- Non-matching terms 135, 146
- Non-MPEG-7 audio descriptors 91
- Non-negative components 66
- Non-negative factorization (NMF) 65
- Non-negative matrix factorization (NMF) 61, 101
- Non-negativity constraints 65
- Non-orthogonal directions 214
- Non-overlapping units 142
- Nonorthographic* 115
- Nonspeech* 116
- Normalization 86
  
- Normalized audio spectrum envelope (NASE) 75
- Normalized autocorrelation function 33
- Normalized histogram 85
- Normalized term frequency 149
- NoteArray 186
- Number of frequency bands 30
- Number of logarithmic bands 25
- Number of occurrences 131
- NumOfDimensions* 117
- NumOfElements* 18
- NumOfOriginalEntries* 115, 116
- Nyquist frequency 16, 30, 32
  
- Object recognition 71
- Objects and events 4
- Observation sequence 67
- Observation vectors 69, 105
- Octave 31
- Octave scale 36
- Off-line 210
- Okapi function 136
- Okapi similarity measure 146
- OM 263
- On-line audio classification 91
- Onsets 40
- OOV 128, 144, 161
- Open-vocabulary retrieval 129
- Operator* 221
- Optimal HMM state path 84
- Order of appearance 118
- Ordered principal component analysis (OPCA) 214
- Orthogonal directions 214
- Orthogonality 65
- Orthogonality constraint 65
- Orthogonalization 65
- Orthographic 115
- Other 116, 117
- Out-of-band* coefficients 26
- Out-of-band* energy 25
- Out of vocabulary 110, 128
- Output layer 70, 71
- Output vector 71
- Overlap 30, 105, 212
- Overlapping 79
- Overlapping sub-word units 142
- Overtraining 242
  
- Parallel processing units 70
- Parts-based representation 65
- Pass-band 53
- Pattern classification 68
- Pattern recognition framework 74
- Patterns 73
- PCA 63, 64, 65
- PCA algorithm 88
- PCA basis 77
- PCA/FastICA 88

- PCM format 258  
 PCM 208  
 Peak candidate 44  
 "Peakiness" 44  
 Peaks 45  
 Perceived loudness 81  
 Perceptual 79  
 Perceptual features 11, 38, 49  
 Perceptual features of musical sounds 171  
 Perceptual grouping of frequency  
   partials 197  
 Perceptual identification 208  
 Perceptual information 61  
 Perceptual linear prediction (PLP) 235  
 Perceptual sharpness 28  
 Perceptually motivated mel-frequency  
   scaling 79  
 Percussive sounds 40, 172, 177  
*PercussiveInstrumentTimbre* 176  
 Periodic signals 34  
 Periodicity 37  
*Person* 119  
 PhonDat 150  
 Phone-based retrieval system 119, 144  
 Phone-based VSM 165  
 Phone confusion counts 150  
 Phone confusion matrix (PCM) 118, 146  
 Phone confusion statistics 164  
 Phone error count matrix 146  
 Phone error rate 140, 151, 154  
 Phone HMMs and bigram language  
   model 266  
 Phone lattices 149, 150, 154  
 Phone-level similarities 147  
 Phone lexicon 150  
 Phone loop 110, 111  
 Phone-only recognizer 150  
 Phone pair 146  
 Phone-phoneme mapping 140  
 Phone recognition systems 129, 266  
 Phone recognizer 151  
 Phone sequences 149  
 Phone strings 158  
*PhoneIndex* 120  
*PhoneLexicon* 114, 115, 116, 118, 123, 165  
*PhoneLexiconref* 121  
*PhoneLinks* 122, 165  
 Phoneme error rate 140  
 Phoneme recognizers 68, 129  
 Phonemes 129, 140, 145, 160  
 Phones 110, 129, 140, 145, 152, 157, 223  
 Phonetic-based approach 161  
 Phonetic classes 141  
 Phonetic error confusion matrix 148  
 Phonetic lattices 141  
 Phonetic level 144  
 Phonetic name matching 201  
*Phonetic n-grams* 141  
 Phonetic recognition 110  
 Phonetic retrieval 151  
 Phonetic similarity 135  
 Phonetic similarity measures 148  
 Phonetic string matching methods 161  
 Phonetic sub-string matching 162  
 Phonetic transcription errors 161  
 Phonetic transcriptions 141, 142, 149,  
   150, 157  
*PhoneticAlphabet* 115, 116  
 Phrase 116  
 Pitch contour 38  
 Pitch curve 37  
 Pitch estimation 33  
 Pitch estimation techniques 36  
 Pitch frequency 42  
*Pitch halving* 36  
 Pitch, loudness 38  
 Pitch period 36  
 Pitch shifting 209  
 Pitches 36, 37, 42, 43, 45, 52, 53  
 Pitching 215  
 Polynomial 73  
 Polyphonic melody transcription 196  
 Polyphonic music transcription 65  
*Pop* 227  
 Porter's algorithm 137  
 Post-processing 212  
 Posteriori probability 67  
 Power 91  
 Power coefficient grouping 32  
 Power coefficients 25, 31, 32  
 Power spectra 32, 35, 49  
 Power spectrum 16, 31, 33, 79  
 Power spectrum coefficients 25  
 Pre-emphasis 105, 212  
 Pre-processing 212  
 Precision 151, 250  
*Precision-Recall* curve 151  
 Principal component analysis (PCA) 60, 61,  
   100, 62, 214  
 Principal components 62  
 Probabilistic retrieval models 126  
 Probabilistic string matching 157  
 Probabilities of confusion 158  
 Probabilities of correct recognition 158  
 Probabilities of transition  $a_{ij}$  106  
 Probability 69  
 Probability densities 66  
 Probability density functions 68, 106, 107  
 Probability of confusing phone  $r$  118  
 Probability ranking principle 126  
 Production 3  
 Professional sounds archiving 92  
 Prominent pitch 215  
 Pronunciation dictionaries 144, 161  
 Prosodic feature 37  
 Provenance 121  
 Provenance attribute 165

- Pruning 241
- PSM algorithm structures 158
- PSM algorithm with restrictions on deletions and insertions 159
- PSM procedure 158
- Psycho-acoustical feature 52
  
- QBH 193
- Quarter-octave frequency 218
- Quasi-Newton algorithm 71
- Query 125
- Query-by-humming 8, 11, 193
- Query expansion 139
- Query string will 155
- Quorum* matching function 132
  
- Radio broadcasts 1
- Random 19, 21
- Ranked list of documents 125
- RASTA-PLP 235
- Ratio 18, 20, 22
- Raw* 18, 20
- Raw* vector 19
- Real-world 103
- REC 117
- Recall* 151, 250
- Recognition 1, 2, 67
- Recognition and indexing description tools 8
- Recognition errors 148
- Recognition improvement 95
- Recognition rates 88
- Recognition vocabulary 128
- Rectangular filters 81
- Reduced dimension  $E$  88
- Reduced log-energy spectrum 54
- Reduced SVD basis 77
- Reduced term space 136
- Reduction of redundancy 213
- REF 117
- Reference 245
- Register 209
- RelAtive SpecTrAl technique* 235
- RelativeDelay* 221
- Release 40
- Release phase 40
- Relevance 227
- Relevance assessment 139
- Relevance feedback 139
- Relevance feedback techniques 138
- Relevance score 125, 150, 159
- Relevancy 131
- Reliability 209
- Reliability of recognition output 163
- Representation 115
- Request 125
- Rescaling 22
- Resegmentation 244
- Retrieval 98
- Retrieval effectiveness 151
- Retrieval efficiency 139
- Retrieval functions 152
- Retrieval model 126
- Retrieval performance 138, 148, 153
- Retrieval scores 145, 149
- Retrieval status value (RSV) 125, 130
- Retrieval system 84, 266
- Retrieve 103
- Retrieving 4
- Rhythm 215
- “Richness” 7, 171
- RMS 39, 47, 52, 75
- RMS energy envelope 81
- RMS-norm gain values 77
- Robust segmentation 246
- Robust sub-word SDR 149
- Robustness 209
- Rocchio algorithm 139
- Rock music 177
- Rolloff frequencies 51
- Root-mean-square deviation (RMS) 29
- RSV 134, 135, 144, 148, 155, 159
- Rule-based phone transcription system 144
- Rule-based text-to-phone algorithms 144, 161
  
- Saddle point 72
- SAMPA phonetic alphabet 119
- SAMPA 117, 140
- Sample value 19
- Sample weights 20
- SampleHold* 227
- Sampling frequency 51
- Sampling period 23
- Sampling rate 14
- Scalability 209
- ScalableSeries* 17
- ScalableSeries* description 14, 17, 18
- Scale 178, 179
- Scale ratio 18
- Scaled samples 17, 18, 20
- Scaling 17, 18, 19, 22
- Scaling ratio 218
- Schema 9
- Schema language 9
- Schema of an ASR system 105
- SDR 123, 160
- SDR matching 145
- SDR system 126, 127
- Search engines 1, 4, 59
- Search, identify, filter and browse 2
- Searching 92
- Second central moment 29
- Second-order moments 62
- Second-order statistics 63
- Segment-based 110
- Segment clustering 237

- Segment-level clustering 244, 245
- Segmentation 59, 60, 69, 71, 232
- Segmenting 50, 61
- Semantic labels 8
- Semantic level 144
- Semantic similarity 137
- Semantic similarity measure 139
- Semantic term similarity 135
- Semantic term similarity measure 136
- Semantics 59
- Separated acoustic class models 264
- Sequence of phonetic units 144
- Sequence of symbols 201
- Sequences of phonetic units 141
- Series of scalar LLD features 19
- Series of scalars 18
- Series of vector LLD features 22
- Series of vectors 18
- SeriesOfScalar 18, 19, 37
- SeriesOfScalar* descriptor 18
- SeriesOfScalar* rescaling 19
- SeriesOfScalarBinary* 22
- SeriesOfVector* 20, 22
- SeriesOfVector* descriptor 20
- SeriesOfVectorBinary* 22
- SeriesOfVectorBinary* description 218
- Set of keywords 138
- Set of linguistic rules 141
- SF 52
- Sharpness* 49
- Shifting 215
- Short-term audio power spectrum 24
- Short-term spectral-based features 81
- Short-time Fourier transform (STFT) 75
- Shot-based multi-modal 259
- Sigmoid function 70
- Signal amplitude spectrum 51
- Signal discontinuities 53
- Signal envelope 23, 39, 42
- Signal-model-based probabilistic inference 198
- Signal parameter descriptors 6
- Signal quality description 11
- Silence 50, 59
- Silence descriptor 13
- Silence detection algorithms 50
- Silence-Header* 50
- Silence model 244
- Silence Segment 50
- Similarity 61
- Similarity matching 11
- Similarity matrix 146
- Similarity measure 155
- Similarity retrieval 61
- Similarity score 158, 202
- Similarity term function 133, 135
- Single linkage 240
- Single word queries 150, 151
- Singular value decomposition (SVD) 49, 60, 61, 213
- Slope threshold 44
- Slot detection* 160
- Smoothing 80
- SNF 263
- SNF algorithm is 262
- Sound 59
- Sound classes 50, 70, 75, 78
- Sound classification 6, 11, 59, 60, 61, 65, 70, 73, 77, 85, 92, 254
- Sound classification accuracy (%) 259
- Sound classification systems 61
- Sound classifiers 60
- Sound content 59
- Sound descriptors 256
- Sound events 59
- Sound feature space 78
- Sound ideas 92
- Sound models 67, 68, 74
- Sound properties 74
- Sound recognition 73, 74
- Sound recognition and indexing 6
- Sound recognition system 68, 74
- Sound recognizers 8
- Sound signals 1
- Sound similarity 11, 59, 59, 60, 61, 69
- Source coders 10
- Source components 77
- Source string 155
- Sources 63
- Speech recognition 68
- Speaker-based segmentation 85, 233
- Speaker change detection 233, 244, 244
- Speaker identification (likelihoods) 245
- Speaker model scores (likelihoods) 245
- Speaker models 244
- Speaker recognition 85, 96
- Speaker verification 68
- SpeakerInfo* 115, 119, 165
- SpeakerINfoRef* 122
- Speakers 59, 60
- Spectral and temporal centroids 7
- Spectral autocorrelation (SA) method 36
- Spectral basis descriptors 6
- Spectral basis functions 64, 73
- Spectral basis representations 13, 49
- Spectral centroid (SC) 13, 38, 48, 49
- Spectral centroid measure 46
- Spectral change 52
- Spectral coefficients 62, 76
- Spectral coverage 218
- Spectral descriptors 5, 42
- Spectral distributions 51
- Spectral domain 36, 66, 87
- Spectral energy 54
- Spectral energy distribution 2
- Spectral envelope 38, 47
- Spectral estimates 214

- Spectral features 59, 66
- Spectral flatness 32
- Spectral flatness coefficient 32
- Spectral flatness descriptor 7
- Spectral flatness features 32
- Spectral flatness measure (SFM) 213
- Spectral flux (SF) 51
- Spectral noise floor (SNF) 261
- Spectral peaks 44, 45
- Spectral power coefficients 32
- Spectral resolution 24, 81
- Spectral rolloff frequency 51
- Spectral shape 51
- Spectral timbral 42
- Spectral timbral descriptors 13, 38, 39, 42
- Spectral timbral features 42
- Spectral timbral LLDs 46
- Spectral variation 48
- SpectralCentroid* 174, 176
- Spectrogram 24, 27, 77
- Spectro-temporal autocorrelation (STA)
  - method 36
- Spectro-temporal autocorrelation function
  - (STA method) 43
- Spectrum 32, 33, 83
- Spectrum centroid 91
- Spectrum dimensionality 76
- Spectrum flux 91
- Spectrum projection features 77
- Speech 1, 50, 52, 59, 70, 75
- Speech endpoint detection 261
- Speech messages 124
- Speech/music 51
- Speech processing 66
- Speech recognition 52, 60, 66, 68, 70, 127, 255
- Speech signal 37
- Speech sounds 45
- Speech-to-text 109
- Speech transcripts 124
- Spoken annotation 4, 255
- Spoken annotation of databases 166
- Spoken content description 254
- Spoken content description tools 8
- Spoken content descriptors 256
- Spoken content information 123
- Spoken content 1, 6, 8, 103
- Spoken content metadata 124
- Spoken document retrieval (SDR) 11, 123, 166
- Spoken keywords 103
- Spoken language information 163
- Spoken retrieval system 126
- SpokenContent* description 163
- SpokenContent* descriptors 146
- SpokenContent* 8, 9, 11, 127, 130, 166
- SpokenContentHeader* 114, 117, 150, 165
- SpokenContent-Lattice* description 123
- SpokenContentLattice* 114, 121, 164
- SpokenContentLink* descriptor 123
- SpokenLanguage* 120
- Spread 6, 47
- STA 37
- STA pitch estimation 37
- Standardized tools 3
- State 69
- State duration 69
- State path histograms 85, 95
- State transition probabilities 68, 77
- State sequence 79
- States 68
- Statistical behaviour 88
- Statistical independence 64
- Statistical models 60, 68
- Statistical speaker model 245
- Statistically independent 69
- Statistically independent basis 77
- Statistics 62
- Steady-state portion 41
- Stem 116
- Stemming method 137
- STFT 76
- Still images 65
- Stochastic language models 129
- Stochastic similarity measure 158
- Stop words 136
- Storage 3
- Storage features 3
- Streaming 10
- String matching 155
- String matching techniques 159
- Sub* 118, 147
- Sub-lexical 111
- Sub-lexical indexing 145
- Sub-lexical indexing terms 140
- Sub-lexical term similarity measure 146
- Sub-lexical terms 161
- Sub-lexical transcriptions 161
- Sub-lexical units 129, 155
- Sub-string 155
- Sub-word indexing 145
- Sub-word indexing terms 129, 149
- Sub-word string matching 154
- Sub-word units 141
- Sub-word-based ASR system 162
- Sub-word-based retrieval 145
- Sub-word-based SDR 128, 154
- Subjective duration 38
- Subspace 65, 73
- Substitution 117, 118
- Substitution errors 113, 147
- Subword-based word spotting 155
- Suffix stripping* 137
- Sum 25
- Sum of frequencies (SF)* 201
- Sum of squared differences 85
- Summed variance coefficients 21

- SUMMIT 110  
 Support vector machines (SVMs) 60, 66, 71  
 Support vectors 72, 73  
 Sustain 40, 172  
 Sustain level 40  
 Sustain phase 40  
 SVD 75  
 Syllable 111, 116, 129, 145  
 Syllable bigram 143  
 Syllable models 143  
 Syllable recognizer 143, 165  
 Symbolic representation 234  
 Synchronization 3, 209  
 Syntactic network 108
- TA method 36  
 Target string 115  
 Taxonomy methods 92  
 TC 41, 42  
 TC feature 42  
 Tempo 190  
*Tempo DS* 171  
 Temporal 5, 39, 66  
 Temporal and spectral domains 37  
 Temporal and spectral features 7  
 Temporal autocorrelation (TA) method 36  
 Temporal autocorrelation function (TA method) 43  
 Temporal centroid (TC) 13, 38, 41  
 Temporal components 3  
 Temporal correlation 211  
 Temporal envelope 38  
 Temporal properties 22  
 Temporal resolution 23, 218  
 Temporal scope 218  
 Temporal structure of musical sounds 171  
 Temporal structures 77  
 Temporal threshold 50  
 Temporal timbral 39  
 Temporal timbral descriptors 13, 38, 39  
*TemporalCentroid* 174, 176  
 Term frequencies 145  
*Term mismatch* problem 132  
*Term misrecognition* problem 133  
 Term similarity 133  
 Term weighting 149  
 Testing 91  
 Text 4  
 Text-based annotation 2  
 Text-based database management 2  
 Text-based IR 135  
 Text descriptions 2  
 Text normalization methods 137  
 Text pre-processing 144, 165  
 Text processing 145  
 Text requests 138, 143  
 Text retrieval 124, 136  
 Text retrieval methods 128
- Text-to-Phone 145, 266  
 Textual form 6  
 Theme 177  
 Thesauri 137  
 Three basis decomposition algorithms 97  
 Threshold 40, 111, 209  
 Timbral descriptors 38  
 Timbral sharpness 46  
 Timbral spectral descriptors 6  
 Timbral temporal descriptors 6  
 Timbre description 7  
*Timbre DS* 38, 171  
 Time average 41  
 Time-based centroid 41  
 Time-delay neural network (TDNN) 70  
 Time domain 23, 39, 42  
 Time domain descriptions 5  
 Time-frequency analysis 6  
 Time-frequency representation 213  
 Time-independent spectral peaks 86  
 Time Pitch Beat Matching I 203  
 Time variation 59  
 TimeStamp 227  
 TIMIT database<sup>1</sup> 91  
*Token* 116  
 Tonal 32  
*Tonal sounds* 29  
 Tone colour 171  
 Topology 249  
 Total recognition rates 96  
*TotalNumOfSamples* 18  
*TPBM I* 203  
 Trade-off 111, 210, 212  
 Trained sound class models 74  
 Training 61, 70, 74, 91  
 Training audio data 78  
 Training data 71, 73, 129  
 Training information 69  
 Training process 78  
 Transcription 8, 233  
 Transcription rules 161  
 Transformation 212  
 Transition probabilities 107, 157, 158  
 Transition schemes for different string matching algorithms 160  
 Transitional information 40  
 Transmission 3  
*TransmissionTechnology* 221  
 Trapezium-shaped 81  
 TREC 151  
 TREC SDR 130  
 TREC SDR experiments 151  
 Triangular filters 53, 79  
 Trigrams 110, 142, 150, 152
- Ukkonen 201  
 Ukkonen measure (UM) 201  
 Unified Modeling Language (UML) 9  
 Unvoiced speech 51

- Upper limit of harmonicity (ULH) 33,  
35, 36  
Usage 3  
*UsedTool* 221
- Variable-length phone sub-sequences 142  
Variable-length syllable units 142  
Variances 19, 20, 21, 88  
VarianceSummed 21, 22  
VCV 143  
Vector quantization (VQ) 220  
*VectorSize* 20  
Vector space model (VSM) 126, 130  
Vibrato 47, 48  
Violins 45  
Visual information processing 259  
Visual inspection 87  
Visual interpretation 87  
Viterbi algorithm 69, 79, 107, 216, 244  
Vocabulary 68  
Vocabulary-independent 151  
Voice 45, 60  
Voice conversion 66  
Voiced 51  
Voiced fricative, nasal 141  
Voiced speech 37  
Voiced/unvoiced speech 51  
Vowel classification 71  
Vowel–Consonant–Vowel method 143  
Vowels 45  
VSM 136, 139, 145, 149, 154  
VSM weighting schemes 152
- Walsh–Hadamard transform 213  
Wavetable music synthesis 41  
*Weight* 19, 20, 21, 22, 37  
Weight of term 131  
Weighting methods 131
- Weighting schemes 131  
Western classical music 177  
White noise 34  
Windowing function 15, 42, 53, 105  
Windows 246  
*Within-band* coefficient 26  
Within-band power coefficients 25  
Word 115, 116  
Word-based retrieval 162  
Word-based retrieval approaches 140  
Word-based SDR 128, 135, 136  
Word boundaries 144  
Word hypotheses 129  
Word-independent 111  
Word recognition 136  
Word similarity measure 136  
Word spoken in isolation 144  
Word spotting 138  
Word stemming 137, 144  
Word stopping 136  
Word stopping process 144  
Word substitution errors 161  
Word vocabulary 140  
WordIndex 120  
*WordLexicon* 114, 115, 116, 123, 165  
*Wordlexiconref* 121  
*WordLinks* 122, 165  
Words spoken 103  
Word-to-syllable 116
- XML 6  
 $Y = XC_E W$  88
- ZCR 214  
Zero crossing features 51  
Zero crossing rate (ZCR) 50  
Zero-lag 34  
Zero-mean variables 62

