References


396 References


404 References


406 References


408 References


References


References


426 References


Index

Accent, 107, 112, 374
Acoustic features, see Feature extraction
Acoustic model, see Source model
Adaptive modelling, see Source model adaptation
Adaptive oscillator, see Oscillator
Additive synthesis, see Sinusoidal model
Aerophone, 167, 231
Amplitude envelope, 176
Amplitude evolution model, 208
damped amplitudes, 208, 218
exponential decay model, 207, 219
Amplitude modulation
as clue for component fusion, 306
tremolo, see Tremolo
Amplitude prior, 212
Analysis by synthesis, 140
Analysis frame, 23, 66, 68
Analytic mode of listening, 6
Annotation of music, 17, 355
Arrhythmic, 105
Articulation, 9, 193
in singing, 364
Artificial neural network, see Neural network
Attack time, 177
Attack-point, 134
Audio coding, 3, 70, 88, 91
Auditory cortex, 12
left/right asymmetry, 12
Auditory filters, 235
bandwidth, 236
centre frequency, 236
frequency response, 237
impulse response, 237
Auditory model, 6, 14, 234, 372
of Cooke, 246
of Ellis, 245
of Meddis and Hewitt, 241
toolboxes, 238
Auditory nerve, 234, 235, 237, 238, 242
Auditory scene analysis, 7, 246, 247, 299, 306
Auditory system, 65, 234, 300
auditory cortex, 234
central auditory processing, 235, 241, 242, 265
cochlea, see Cochlea
peripheral hearing, 14, 234
Autocorrelation, 241, 243, 245
autocorrelation function, 233
calculation via FFT, 256
enhanced, 288
generalized ACF, 252, 256
of a frequency-domain signal, 233
summary ACF, see Summary autocorrelation
use for beat tracking, 112
use for F0 estimation, 232, 369
Automatic music transcription, see Transcription
Automatic threshold selection, 350
Autoregressive model, 39, 175, 211
Autoregressive noise, 211
Average magnitude difference function, 251
Bandwise energy features, 135
Bar line, 10, 106
Bar line estimation, 124, 159, 341
Basilar membrane, 234, 235
Basis function, 269
amplitude, see Gain function
frequency domain, 143, 146, 149, 273
time domain, 271
Bass line, 13, 321, 329
Bass line transcription, 330
Bassoon, 167
Bayes’s rule, 38, 40
Bayesian
classification, 62
estimation, 40
model, 203
network, 313
Beat, 10, 105, 134, 329, 341
Beat induction
by humans, see Metre perception
computational, see Beat tracking
Beat tracking, 6, 101, 341
Beating, 239, 256
Bell, 167
Binning, 154
Blackboard architecture, 246, 247, 312
Blind source separation, see Source separation
Bottom-up processing, 16, 307, 312, 342, 352
Box-Cox power transform, 179
Brain damage, 11
Brain imaging, see Neuroimaging
Brass instrument, 167
Bugle, 167
Cancellation filter, 196, 250
Canonical discriminant analysis, 180
Celesta, 167
Cello, 167, 209, 340
Cent, 332
Central auditory processing, see Auditory system
Central limit theorem, 274
Cepstral coefficients, 25, 26
MFCC, see Mel-frequency cepstral coefficients
Cepstrum, 25, 253
Cepstrum pitch detection, 253, 264, 371
Channel selection
in auditory model, 244, 249, 251
Choral music, 109, 125
Chord, 9, 13
chord change detection, 117, 328, 341
notation, 3
recognition, 319
root detection, 246, 254, 263
Chordophone, 167, 231
Chorus measure, 345, 353
Chorus sections, 329
estimation, 342
Chroma and height, 346
Chroma vector, 345, 347
Clarinet, 96, 166, 167, 176, 177, 206, 207, 319
Classical music, 7, 112, 125, 128, 340, 355
Classification, 52, 137, 139, 163, 184
combining, 193
ensemble, 184
flat, 184
generalization, 169, 189
genre, 13, 112, 356
hierarchical, 184, 192
unsupervised, 53, 133, 164
Clustering, 133, 151, 159, 164, 193, 286, 307, 311, 340, 352, 354
Cochlea, 164, 234
cochlear filters, see Auditory filters
computational models, 234
Cocktail party problem, 300
Comb filter, 115, 243, 257
Combining classifiers, 193
Common musical notation, 3, 4
expression symbols, 389
Complete family, 72
Complete transcription, 3
Compression
in auditory model, 238, 239, 248
νth-law compression, 255
Computational auditory scene analysis, see Auditory scene analysis
Concurrent vowel identification, 249
Conditional probability density function, 31
Conjugate prior, 47
Contextual information, 198
in auditory scene analysis, 313
in instrument classification, 166
in singing, 385
short-term acoustic context, 154
tonal context, 320, 385
Continuous-length evaluation measure, 125
Contrabass, 167, 198
Convolution, 22, 25, 289
time-frequency, 25
Correlation-based feature selection, 182
Correlogram, 227, 242, 245, 339
Covariance matrix, 29, 30
empirical, 53, 61
Cramér–Rao bound, 34
Creative Commons initiative, 170
Critical band, 236, 274
Critical-band scale, 236
Cross-validation, 170

Damped amplitudes, see Amplitude evolution model
Damping factor, 208
Dance music, 102, 125, 160
ballroom dance, 117
Data representation, see Mid-level data representation
Data-driven processing, 16
Databases, 16
of musical instrument sounds, 169, 260, 355
RWC Music Database, 355
DCT-IV transform, 75
decibel, 8, 294
Decision tree, 138, 186
Decorrelation, 275
nonlinear, 275
Delta features, 175
Delta-MFCC, see Mel-frequency cepstral coefficients
Description of music signal, 328
Deterministic approach to signal modelling, 67
Diatonic scale, 271, 293
Dictation, see Human transcription
Dictionary of waveforms, see Waveform dictionary
Dilation, 71
Directed acyclic graph, 120, 313
Discrete cosine transform, 75
Discrete Fourier transform, see Fourier transform, 273
Discrete wavelet transform, see Wavelet transform
Discriminant analysis, 61, 185
Dissonance, 9, 287
Distribution, see Probability density function
Divergence, 282
Double bass, see Contrabass
Drums, see Percussion
Dual frame, 78
Duplex theory of pitch perception, 241
Duration
of a note, 362
perceived duration, 8
Dyadic grid, 74
Dynamic model
finite, discrete state space, 63
Dynamic range
of hearing, 238, 273
of inner hair cells, 238
Ear training, 12
Eardrum, 234
Eigenvalue, 54, 180
Eigenvector, 54
Electric guitar, see Guitar
Elementary waveform, 70
EM algorithm, see Expectation maximization
Empirical average, 29
Empirical probability density function, 82
Empirical risk, 56
English horn, 167
Enhanced summary ACF, see Summary autocorrelation function
Ensemble of classifiers, 184
Equal-tempered scale, 9, 224
Equivalent rectangular bandwidth, 236
Estimation,
bias, 33, 41
covariance, 33
unbiased estimator, 34
see also Fundamental frequency estimation, 21
Estimation theory, 33
Euclidean distance, 282
Evaluation methodology, 170
Expectation, 29
Expectation maximization, 35, 221, 336
Exponential decay, see Amplitude evolution model
Expression, 193
  estimation and encoding, 389
FastICA algorithm, 275
Feature extraction, 66, 135, 171, 247, 270, 287, 307, 344, 345, 368
delta, 175
  sliding window, 155
Feature scale transformation, 178
Feature selection, 137, 181
Feature transformation, 191
Fisher ratio, 61
Flat classifier, 184
Flute, 205, 206, 209, 319, 323, 340
Folk music, 125, 323
Formant, 253, 294, 365
Fourier transform, 21, 22
  continuous, 22
  discrete, 22
  inverse, 22
Fractional-delay filter, 250
Frame, see Analysis frame
Frame (family of vectors), 77
dual, 78
  frame inversion, 78
  frame representation, 78
Gabor frame, 78
  hybrid system, 81
  multiple Gabor frames, 79
  overcomplete system, 78
  quilted frame, 81
French horn, 166, 167
Frequency evolution model, 209
Frequency grid, 219, 224
Frequency modulation
  as clue for component fusion, 306
  vibrato, see Vibrato
Frequency proposal distribution, 216
Frequency warping, 254
Frobenius norm, 280
Front end, 244
Full-wave $\nu$th-law compression, 255
Fundamental frequency estimation
  basic principles, 232
  multiple FO estimation, 203, 229, 248, 287
  of melody and bass lines, 330
  single FO estimation, 232, 369
  typical errors, 233, 253
Fundamental frequency pdf, 332
Fundamental frequency, term definition, 8
g-prior, 212
Gabor frame, 78
Gabor function, 78
Gabor representation, 24, 209
Gabor waveforms, 71
Gain function, 143, 269
Gamma prior, 211
Gammatone filter, 237, 241, 247, 255, 306
  efficient implementation, 238
Gaussian distribution, 28
  Gaussianity of variable, 274
  generalized, 286
  truncated, 213
Gaussian mixture model, 35, 37, 139, 191, 385, 386
Gaussian noise, 279, 282
Generalization in classification, 169, 189
Generalized autocorrelation, see
  Autocorrelation
Generative model, 62, 112, 122, 204, 277, 324
Generative signal, speech and music, 16
Genre classification, 13, 112, 356
Gestalt psychology, 300
Glissando, 389
  in singing, 367
Glockenspiel, 232
Greedy algorithm, 84, 89, 225, 354
Grid, temporal, see Temporal grid
Grouping, see Rhythmic grouping
Guitar, 79, 166, 167, 177, 227
  electric, 198, 200, 209
  onset detection, 108
  sound separation, 270, 290
  transcription, 301
Hair cell, see Inner hair cell, Outer hair cell
Half-wave rectification, 239, 256, 258
Harmonic matching pursuit, 92
Harmonic partial, 107, 231, 291
Harmonic pattern matching, 233
Harmonic selection, 257
Harmonic sound, 231
Harmonic trajectory, 225
Harmonicity, as clue for component fusion, 306
Harmonics and Individual Lines plus Noise, 69
Harmony, 9
Harp, 167
Harpsichord, 167
transcription, 282
Heisenberg-Gabor inequality, 24
Hidden Markov model, 63, 121, 122, 139, 191, 286, 380
Hierarchical beat structure, see Metre
Hierarchical classifier, 184, 192
Higher-order statistics, 275, 287
History of automatic music transcription, 6, 301
Human auditory system, see Auditory system
Human transcription, 5, 12
Human-computer interaction, 341
Human-computer interaction, 5, 361
Hybrid representation, 67
Hybrid system, see Frame (family of vectors)
Hyperparameter, 47, 62

i.i.d., 32
Idiophone, 131, 145, 167
Ill-posed problem, 38
Importance probability density function, 49
Importance sampling, 49, 223
Independence, see Statistical independence
Independent component analysis, 109, 143, 274, 305
multidimensional, 276, 277
non-negative, 278
spatiotemporal, 277
Independent subspace analysis, 144, 276, 285
sub-band, 145
inertia ratio maximization using feature space projection, 183
Information gain, 182
Information retrieval, see Music information retrieval
Inharmonicity, 175, 206, 222
in string instruments, 231
inharmonicity factor, 232
model for piano, 206, 222, 232
Inner ear, 234
Inner hair cell, 234, 235
models, 238, 241, 246, 247, 255
Inner lines in music, 6, 13
Inner product, 71
Inner product space, 71
Instance-based classification, 138, 185
k-NN, see k-nearest neighbours
Instantaneous frequency, 246, 332
Instrument
sample databases, see Databases sounds, see Musical sounds
Instrument classification, 7, 65, 163
in humans, 164
percussion, see Percussion sound recognition
Instrument families, 167
Integration of information, 16, 154, 313, 383
Inter-onset interval, 8, 115
Intermediate data representation, see Mid-level data representation
Internal model, 15
Interval, 12, 321, 340, 341, 362
Invariant feature extraction, 276, 277
Inverse Fourier transform, see Fourier transform
Inverse gamma distribution, 47
Iterative F0 estimation and cancellation, 196, 250, 254, 259, 263

JADE algorithm, 275
Jazz music, 105, 125, 127, 321, 355
Joint estimation of multiple F0s, 214, 251, 252, 263
Joint probability density function, 31
Junction tree algorithm, 316
k-nearest neighbours, 60, 185, 191
Kalimba, 167
Kalman filter, 51, 52, 117, 121, 221–223, 226, 324
Kernel, 57
  Gaussian, 57
  positive definite, 57
Kullback–Leibler divergence, 282
  symmetric, 287
Kullback–Leibler information, 335
Kurtosis, 275
  spectral, 136
Language model, 15
Laplace approximation, 225
Laplacian distribution, 280
Latent variable, 35
Law of large numbers, 30
Lazy learning, 185
Least-squares, 285
Leave-one-out cross-validation, 170
Lebesgue measure, 28
Legato, 9
  in singing, 367
Level adaptation in auditory model, 239, 246
Level compression, see Compression
Likelihood, 210, 219, 332
Likelihood function, 32
  degenerate, 33, 38
  penalized, 38, 41, 57
Linear discriminant analysis, 186, 191
Linear interpolation, 253
Linear prediction, 39
Linear programming, 286
Local cosine basis, 73
Localized source model, 142
Locally harmonic sources, 69
Log-Gaussian distribution, 224
Loss function, 55
Loudness, 8, 172
  of instrument sounds, 319
  of melody vs. accompaniment, 340
  of singing, 366
Mallet percussion instrument, 232
Marginal MMSE, 214
Marginal probability density function, 31
Marimba, 167, 232
Markov chain, see N-gram model
Markov chain Monte Carlo, 43, 117, 121, 215
Markov tree, 90
Masking, 236
Matching pursuit, 84
Matrix diagonalization, 54
Matrix factorization, see Non-negative matrix factorization
Maximum a posteriori estimation, 40, 121, 214, 279, 331
Maximum likelihood estimation, 33, 35
Mbira, 167
McNemar’s test, 171
Mean, 29, 30
  empirical, 54, 61
Mean square error of an estimator, 34
Measure
  musical measure, see Bar line
  musical measure estimation, see Bar line estimation
Mechanical-to-neural transduction, 238
Mel frequency cepstral coefficients, 26
Mel frequency scale, 26
Mel-frequency cepstral coefficients, 63, 135, 174
delta-MFCC, 175
mel-frequency cepstral coefficients, 270
Mel-frequency scale, 173
Mel-scale filterbank, 26
mel-scale filterbank, 173
Melodic phrase, see Phrase
Melody, 9, 12, 13, 329
  perceptual coherence, 15
  segregation of melodic lines, 247
  transcription, see Predominant F0 estimation
Membranophone, 131, 145, 147, 167
Memory for music, 11, 13
Message passing, 221
Metadata, see Annotation
Metre, 10, 105, 312, 329, 341
Metre analysis, 101, 134, 341, 388
Metre perception, 10, 11, 102
MFCC, see Mel frequency cepstral coefficients
Mid-level data representation, 12, 13, 65, 244, 248, 251, 256, 264
desirable qualities, 14
hybrid, see Hybrid representation
Middle ear, 234
MIDI, 3, 4, 9, 101, 102
Minimum mean square error estimation, 41, 214
Missing feature theory, 197
Mixing matrix, 270
Mixture-of-experts approach, 193
Model adaptation, source, see Source model adaptation
Model selection, 39
Model, signal, see Signal model
Modified discrete cosine transform, 75
Mixing matrix, 270
Molecular matching pursuit, 94
Monte Carlo, 41
Monte Carlo, 41
MPEG-7, 136, 167, 172, 270, 354
Multi-class classification, 55
Multi-layer perceptron, 187, 191
Multidimensional scaling, 168, 179
Multiple F0 (non)stationary model, see Signal model
Multiple F0 estimation, see Fundamental frequency estimation
Multiplicative update rule, 283
Multiresolution analysis, 74
Multiwavelet, 77
Music cognition, 11
impaired cognition, 11
Music information retrieval, 5, 102, 170, 327, 356, 363
Music listening station, 357
Music map, 357
Music perception, 5, 103, 327
Music scene analysis, 299
Music scene description, 327, 328
Music structure, see Structure
Music structure analysis, see Structure analysis
Music thumbnail, 342
Music transcription, see Transcription
Music-playback interface, 357
Music-synchronized computer graphics, 356
Musical context, see Contextual information
Musical instrument, see Instrument
Musical instrument classification, see Instrument classification
Musical key, 9, 385
c change, 344
estimation, 386
Musical metre, see Metre
Musical scale, see Scale
Musical sounds
percussive, 107, 131
pitched, 107, 167, 231
Musicological modelling, 15, 153-155
melodic continuity, 331
musical key, 385
of periodic patterns, 157
of rhythmic patterns, 159
short-term context modelling, 154
with N-grams, see N-gram model
Mutual information, 274

N-fold cross-validation, 170
N-gram model, 155
for chord sequences, 319
of melody, 321, 387
of percussion sequences, 155, 156
periodic N-gram, 157, 158
Neural firing probability, 237, 238
Neural impulse, 234, 238, 242
Neural network, 187, 191
MLP, 187, 191
time-delay neural network, 247
Neural spike, see Neural impulse
Neuroimaging, 11
Neuropsychology of music cognition, 11
Noise, 32
autoregressive, 211
Gaussian, 279, 282
Poisson, 282
white, 32, 39
Noise robustness in F0 estimation, 251, 253, 264
Noisy sum-of-sines model, see Signal model
Non-negative matrix deconvolution, 293
Non-negative matrix factorization, 148, 282
Norm, 71
Normal distribution, see Gaussian distribution
Notation, see Common musical notation
Note, 3, 362
  MIDI note number, 362
Note birth move, 216
Note death move, 216
Note labelling in singing transcription, 362, 378
Note model, see Source model
Note segmentation, see Temporal segmentation
Note update move, 216
Nuisance parameters, 220
Nyquist frequency, 23
Oboe, 167, 209
Observation density, see Likelihood
Octave equivalence, 9
Odd-to-even ratio, 175
Offset (a)synchrony, 306
Onset, 101, 260, 262
  and beat, 106
  of percussive sounds, 108
  of pitched sounds, 108
Onset (a)synchrony, 306
Onset detection, 102, 107, 134, 195, 287
Organ of Corti, 234
Orthogonal matching pursuit, 85
Orthonormal basis, 72
Oscillator, 113
  adaptive, 113, 247
  comb filter, 115, 244, 257
  oscillator net, 247
Outer ear, 234
Outer hair cell, 234
Overcomplete system, see Frame (family of vectors)
Overlap-add, 288
Overtone partial, see Harmonic partial
Parametric model of signal, see Signal model
Parseval formula, 73
Partial de-tuning, 206, 306
Partial transcription, 3
Particle filter, 50, 117, 121
  Rao-Blackwellized, 222
Peak selection in autocorrelation function, 251
Penalized likelihood, see Likelihood function
Perception
  of metre, see Metre perception
  of music, see Music perception
  of pitch, see Pitch perception
Perceptual attributes of sounds, 8
Perceptual categorization, 9
Perceptual sound vs. physical sound, 302
Percussion notation, 3
Percussion sound recognition, 133, 137, 174, 184
  clustering and labelling, 159
Percussion transcription, 6, 7, 131, 329, 342
  pattern recognition-based, 133
  separation-based, 142
Periodicity
  in the frequency domain, 233
  in the time domain, 232
Peripheral hearing, see Auditory system
Phase generation, 288
Phenomenal accent, see Accent
Phonation, 364
  frequency, 365
  types, 365
Phoneme, 16
Phrase, 10
Phrasing, 193
Physical sound vs. perceptual sound, 302
Piano, 24, 167, 203, 206, 207, 209, 319, 323
  identification in music, 198
  inharmonicity, see Inharmonicity
  keyboard, 9
  onset detection, 108
  transcription, 245, 247
Piano roll, 4, 220
Pitch, 8, 107
  of noise signals, 229
  perception, see Pitch perception
  tonal encoding, see Tonal encoding
  zoo of pitch effects, 229
<table>
<thead>
<tr>
<th>Term</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch class</td>
<td>9, 345, 385</td>
</tr>
<tr>
<td>Pitch label of a note</td>
<td>362</td>
</tr>
<tr>
<td>Pitch perception</td>
<td>229, 242</td>
</tr>
<tr>
<td>Pitch perception model</td>
<td>234, 241</td>
</tr>
<tr>
<td>autocorrelation model</td>
<td>241, 242</td>
</tr>
<tr>
<td>duplex theory</td>
<td>241</td>
</tr>
<tr>
<td>shortcomings</td>
<td>248, 262</td>
</tr>
<tr>
<td>unitary model</td>
<td>242, 252</td>
</tr>
<tr>
<td>Pitched musical sounds</td>
<td>107, 167, 231</td>
</tr>
<tr>
<td>Pizzicato</td>
<td>192</td>
</tr>
<tr>
<td>Poisson noise</td>
<td>282</td>
</tr>
<tr>
<td>Poisson prior</td>
<td>214</td>
</tr>
<tr>
<td>Polyphonic signal</td>
<td>5</td>
</tr>
<tr>
<td>Popular music</td>
<td>10, 125, 128, 131, 137, 151, 329, 330, 342, 344, 345, 353, 355</td>
</tr>
<tr>
<td>Predominant F0 estimation</td>
<td>198, 260, 330</td>
</tr>
<tr>
<td>PreFEst method</td>
<td>330</td>
</tr>
<tr>
<td>Principal component analysis</td>
<td>54, 137, 139, 144, 151, 152, 155, 179, 181, 275</td>
</tr>
<tr>
<td>Prior</td>
<td>40</td>
</tr>
<tr>
<td>conjugate</td>
<td>47</td>
</tr>
<tr>
<td>sequential</td>
<td>217, 220, 221, 225</td>
</tr>
<tr>
<td>Prior distribution</td>
<td>40</td>
</tr>
<tr>
<td>Prior selection</td>
<td>211</td>
</tr>
<tr>
<td>Prior subspace analysis</td>
<td>146, 150</td>
</tr>
<tr>
<td>input generated priors</td>
<td>151</td>
</tr>
<tr>
<td>non-negative</td>
<td>148</td>
</tr>
<tr>
<td>Probabilistic model</td>
<td>31</td>
</tr>
<tr>
<td>Probability density function</td>
<td>28</td>
</tr>
<tr>
<td>conditional</td>
<td>31</td>
</tr>
<tr>
<td>empirical</td>
<td>82</td>
</tr>
<tr>
<td>Gaussian</td>
<td>see Gaussian distribution function</td>
</tr>
<tr>
<td>importance</td>
<td>49</td>
</tr>
<tr>
<td>inverse gamma</td>
<td>47</td>
</tr>
<tr>
<td>joint</td>
<td>31</td>
</tr>
<tr>
<td>Laplacian</td>
<td>280</td>
</tr>
<tr>
<td>marginal</td>
<td>31</td>
</tr>
<tr>
<td>normal</td>
<td>see Gaussian distribution function</td>
</tr>
<tr>
<td>prior</td>
<td>40</td>
</tr>
<tr>
<td>proposal</td>
<td>44</td>
</tr>
<tr>
<td>uniform</td>
<td>28</td>
</tr>
<tr>
<td>Projected steepest descent algorithm</td>
<td>281</td>
</tr>
<tr>
<td>Proposal distribution</td>
<td>44</td>
</tr>
<tr>
<td>Pseudoinverse</td>
<td>147</td>
</tr>
<tr>
<td>Psychoacoustics</td>
<td>8, 173, 235, 308</td>
</tr>
<tr>
<td>Pulse, see Beat; Metre</td>
<td></td>
</tr>
<tr>
<td>Pulse Code Modulation (PCM)</td>
<td>65</td>
</tr>
<tr>
<td>Quantization</td>
<td>9</td>
</tr>
<tr>
<td>of F0 values</td>
<td>213, 362</td>
</tr>
<tr>
<td>of onset times</td>
<td>101, 106, 118, 389</td>
</tr>
<tr>
<td>of signal sample values</td>
<td>65</td>
</tr>
<tr>
<td>Query by humming</td>
<td>356, 361</td>
</tr>
<tr>
<td>Quilted frame, see Frame (family of vectors)</td>
<td></td>
</tr>
<tr>
<td>Random variable</td>
<td>28</td>
</tr>
<tr>
<td>expectation, see Expectation</td>
<td></td>
</tr>
<tr>
<td>i.i.d., see i.i.d.</td>
<td></td>
</tr>
<tr>
<td>independent, see Statistical independence</td>
<td></td>
</tr>
<tr>
<td>Random walk</td>
<td>208</td>
</tr>
<tr>
<td>Rayleigh quotient</td>
<td>61</td>
</tr>
<tr>
<td>Recognition</td>
<td></td>
</tr>
<tr>
<td>of percussion sounds, see Percussion</td>
<td></td>
</tr>
<tr>
<td>sound recognition</td>
<td></td>
</tr>
<tr>
<td>of pitched sounds, see Instrument</td>
<td></td>
</tr>
<tr>
<td>classification</td>
<td></td>
</tr>
<tr>
<td>Rectification, see Half-wave rectification</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>78</td>
</tr>
<tr>
<td>Reed instrument</td>
<td>167</td>
</tr>
<tr>
<td>RefraID method</td>
<td>345</td>
</tr>
<tr>
<td>Register (in singing)</td>
<td>365</td>
</tr>
<tr>
<td>Regular grid</td>
<td>75</td>
</tr>
<tr>
<td>Regularized risk</td>
<td>57, 58</td>
</tr>
<tr>
<td>Repeated sections, see Structure</td>
<td></td>
</tr>
<tr>
<td>Reproducing kernel Hilbert space</td>
<td>57</td>
</tr>
<tr>
<td>Resolvability</td>
<td>258</td>
</tr>
<tr>
<td>Rest (in music)</td>
<td>362</td>
</tr>
<tr>
<td>Rhythm, 10, 105</td>
<td></td>
</tr>
<tr>
<td>analysis, see Metre analysis</td>
<td></td>
</tr>
<tr>
<td>Rhythmic grouping</td>
<td>10, 105</td>
</tr>
<tr>
<td>Rhythmic pattern modelling</td>
<td>159</td>
</tr>
<tr>
<td>Risk, 55</td>
<td></td>
</tr>
<tr>
<td>empirical</td>
<td>56</td>
</tr>
<tr>
<td>regularized</td>
<td>57, 58</td>
</tr>
<tr>
<td>Rock music</td>
<td>105, 112, 125, 131, 151</td>
</tr>
<tr>
<td>Root-mean-square level</td>
<td>172</td>
</tr>
<tr>
<td>Rotation matrix</td>
<td>218</td>
</tr>
<tr>
<td>Royalty-free music</td>
<td>355</td>
</tr>
<tr>
<td>RWC Music Database</td>
<td>355</td>
</tr>
<tr>
<td>Salience</td>
<td></td>
</tr>
<tr>
<td>of an onset</td>
<td>107</td>
</tr>
</tbody>
</table>
of F0 candidate, 257
Sampling, 65
Saxophone, 167, 198, 207, 256, 259
Scale, 10, 385
equal-tempered, 9, 224
Scale tone, 10, 385
Scale transformation, 178
Scaling function, 74
Scene analysis, 299, see also Auditory
scene analysis
Schema-based segregation, 312
Score, see Common musical notation
Scoretime, 105
Segmentation, see Temporal segmentation
Semitone, 291, 292, 332, 362
Sequential dependency, see N-gram
model
Sequential Monte Carlo, 117
Sequential prior, 217, 220, 221, 225
Shannon theorem, 23
Short-time Fourier transform, 23
Signal model, 66, 204
multiple F0 non-stationary, 209
noisy sum-of-sines, 204
parametric, 67, 70
single F0 non-stationary, 207
single F0 stationary, 204
sinusoidal, see Sinusoidal model
sum-of-sines model, 204
weighted mixture of tone models, 332
Signal space, 71
Significance map, 88
Significance tree, 89
Similarity matrix, 349
Simplex algorithm, 286
Singer’s formant, 366
Singing
acoustic characteristics, 364
production of singing sounds, 364
singing out of tune, 367
Singing transcription
applications, 361
expression encoding, 389
problem formulation, 361
segment-and-label approach, 377
statistical approach, 379
Single F0 (non)stationary model, see
Signal model
Singular value decomposition, 277
Sinusoidal model, 14, 68, 225, 246
other variants, see Signal model
single sinusoid, 31
Sinusoids + transients + noise model, 70
Sliding window method, 155
SmartMusicKIOSK, 357
Sound production mechanism
in musical instruments, 167
in singing, 364
Sound source separation, see Source
separation
Source model, 15
percussion sounds, 149
pitched sounds, 319
statistical note model, 379
tone model, 319, 333
Source model adaptation, 140, 142, 197,
333
in time domain, 140
in time-frequency domain, 141
Source separation, 15, 65, 142, 143, 195,
249, 267, 304, 305, 327
Sparse coding, 143, 278, 285
non-negative, 153, 281, 285
Sparse expansion, 82
Sparse representation, 67
Spatial information, 267, 306
Spectral features
spectral centroid, 136
spectral flatness, 173
spectral flux, 174
spectral irregularity, 174
spectral kurtosis, 136
spectral rolloff, 174
spectral shape, 135
spectral skewness, 136
spectral spread, 136
Spectral model synthesis, 69
Spectral organization, see Auditory
scene analysis
Spectral smoothness, 207, 240
Spectral whitening, 239, 248, 252, 255,
263
Spectrogram, 23, 24, 65
Spectrogram factorization, 143, 268
Spectrum envelope, 207
Speech recognition, 15, 16, 63, 156, 307, 358
speech separation, 249
speech signals, 16
Squared difference function, 250
Staccato, 9
Statistical independence, 31, 274
Statistical significance tests, 170
Stiffness of vibrating strings, 232
Stochastic signal components, 69
Strobed temporal integration, 244
Structure (of a musical work), 10, 329
Structure analysis, 10
by humans, 13
computational, 342
Structured approximation, 87
Structured audio coding, 3
Student–Fisher t-test, 170
Style detection, see Genre classification
Sub-band coding, 74
Sub-beat structure, 103, 106, 117
Subglottic pressure, 364
Sum-of-sines model, 204
Summary autocorrelation function, 242, 245, 248
enhanced summary ACF, 253
Super-beat structure, 103, 106
Supervised classification, 55, 164
Support vector machine, 57, 138, 191
output moderating, 154
Swing, 105
Symbolic representation, 9, 102, 107, 302
Synchrony strand, 247
Synthesis of separated sources, 288

Tactus, 10, 105
Tatum, 106, 134
analysis, 116, 124, 134, 159
Taxonomy, 164
of musical instruments, 167
Template matching, 139
Tempo, 6, 101, 105
estimation, 103, 118, 354, 374
variation, 105, 106, 388
Tempogram, 121
Temporal centroid, 136, 178
Temporal grid, 134
Temporal segmentation, 12, 134
of instrument sounds, 177
of singing notes, 362, 373, 377
Threshold of hearing, 273
Thumbnail, 342
Tick, see Tatum
Timbre, 8, 231
acoustic correlates, 168
acoustic features, see Feature extraction
F0 dependency, 189
of singing sounds, 365
perceived similarity, 168
use for auditory organization, 306
Timbre space, 168
Time quantization, see Quantization
Time–frequency molecule, 87
Time–frequency atom, 70
Time–frequency covariance, 24
Time–frequency jigsaw puzzle, 85
Time–frequency lattice, 209
Time–frequency representation, 13, 23, 25, 65
Time-lag triangle, 349
Time-persistence, 88
Timing deviations, 105, 375
Timpani, 167
Token-passing algorithm, 380
Tonal encoding of pitch, 10, 11
Tonal music, 10, 304, 330, 385
Tone model, see Source model
Tonic note, 10
Top-down processing, 16, 312, 342, 352
Transcription
by humans, 5, 12
complete vs. partial, 3
designing transcription system, 11
state of the art, 7
subtopics, 5, 11
trends and approaches, 6, 301
Transient, 70
Transientness index, 89
Translation, 71
Tremolo, 178, 389
in singing, 367
Tristimulus, 176
Trombone, 166, 167
Trumpet, 167, 176, 198, 209, 245, 254, 319
Tuba, 167
Index

Tuning, 9
  absolute, 362
  drift of, 368
  in note labelling, 378
  singing out of tune, 367
Two-way mismatch, 288, 371

Ukulele, 167
Unbiased estimator, see Estimation
Uniform distribution, 28
Unitary model, see Pitch perception model
Unmixing matrix, 147, 275
Unsupervised
  classification, 53, 133, 164
  clustering, see Clustering learning, 7, 16, 267, 380, 382

Validation, 169, 170
Vibraphone, 232
Vibrating bar, 232
Vibrato, 178, 389
  in singing, 366
  rate and depth estimation, 389
Viola, 167
Violin, 9, 166, 167, 177, 207, 209, 231, 254, 259, 319, 323
  onset detection, 109
Viterbi algorithm, 63, 122, 225, 380
  Vocal organ, 364
  Voice source, 364
  Voicing, degree of, 373

Waveform dictionary, 67, 81
Waveform representation, 13, 72
Wavelet, 71
Wavelet basis, 73
Wavelet transform, 74
Weighted-mixture model, see Signal model
Well-tempered scale, see Equal-tempered scale
Western music, 3, 5, 9, 10, 231, 304
White noise, 32, 39
Whitening, see Spectral whitening
Wiener–Khintchine theorem, 252
Wigner–Ville representation, 25
Window function, 23, 75, 208
Windowing, 23
Woodwind instrument, 167
Written music, see Common musical notation
  Xylophone, 167, 232

Zero crossing rate, 136, 174, 182
Zero tree, 87