

Contents

1	Music Representations	1
1.1	Sheet Music Representations	2
1.1.1	Musical Notes and Pitches	3
1.1.2	Western Music Notation	5
1.2	Symbolic Representations	10
1.2.1	Piano-Roll Representations	11
1.2.2	MIDI Representations	13
1.2.3	Score Representations	15
1.2.4	Optical Music Recognition	17
1.3	Audio Representation	18
1.3.1	Waves and Waveforms	19
1.3.2	Frequency and Pitch	21
1.3.3	Dynamics, Intensity, and Loudness	24
1.3.4	Timbre	26
1.4	Further Notes	30
	References	34
	Exercises	36
2	Fourier Analysis of Signals	39
2.1	The Fourier Transform in a Nutshell	40
2.1.1	Fourier Transform for Analog Signals	42
2.1.2	Examples	48
2.1.3	Discrete Fourier Transform	49
2.1.4	Short-Time Fourier Transform	53
2.2	Signals and Signal Spaces	57
2.2.1	Analog Signals	58
2.2.2	Digital Signals	60
2.2.3	Signal Spaces	63
2.3	Fourier Transform	69
2.3.1	Fourier Transform for Periodic CT-Signals	69
2.3.2	Complex Formulation of the Fourier Transform	71

2.3.3	Fourier Transform for CT-Signals	77
2.3.4	Fourier Transform for DT-Signals	82
2.4	Discrete Fourier Transform (DFT)	86
2.4.1	Signals of Finite Length	86
2.4.2	Definition of the DFT	88
2.4.3	Fast Fourier Transform (FFT)	89
2.4.4	Interpretation of the DFT	92
2.5	Short-Time Fourier Transform (STFT)	93
2.5.1	Definition of the STFT	94
2.5.2	Spectrogram Representation	98
2.5.3	Discrete Version of the STFT	102
2.6	Further Notes	105
	References	109
	Exercises	110
3	Music Synchronization	115
3.1	Audio Features	117
3.1.1	Log-Frequency Spectrogram	118
3.1.2	Chroma Features	123
3.2	Dynamic Time Warping	131
3.2.1	Basic Approach	131
3.2.2	DTW Variants	141
3.3	Applications	146
3.3.1	Multimodal Music Navigation	146
3.3.2	Tempo Curves	151
3.4	Further Notes	154
3.4.1	Audio Features	154
3.4.2	Dynamic Time Warping	156
3.4.3	Music Synchronization	156
3.4.4	Applications	158
	References	160
	Exercises	164
4	Music Structure Analysis	167
4.1	General Principles	169
4.1.1	Segmentation and Structure Analysis	170
4.1.2	Musical Structure	172
4.1.3	Musical Dimensions	175
4.2	Self-Similarity Matrices	178
4.2.1	Basic Definitions and Properties	178
4.2.2	Enhancement Strategies	184
4.3	Audio Thumbnailing	195
4.3.1	Fitness Measure	195
4.3.2	Scape Plot Representation	202
4.3.3	Discussion of Properties	203

4.4	Novelty-Based Segmentation	207
4.4.1	Novelty Detection	208
4.4.2	Structure Features	212
4.5	Evaluation	215
4.5.1	Precision, Recall, F-Measure	216
4.5.2	Structure Annotations	217
4.5.3	Labeling Evaluation	218
4.5.4	Boundary Evaluation	220
4.5.5	Thumbnail Evaluation	221
4.6	Further Notes	224
4.6.1	Self-Similarity Matrices	225
4.6.2	Audio Thumbnailing	226
4.6.3	Segmentation Approaches	227
4.6.4	Evaluation and Sources	228
	References	230
	Exercises	234
5	Chord Recognition	237
5.1	Basic Theory of Harmony	239
5.1.1	Intervals	239
5.1.2	Chords and Scales	243
5.2	Template-Based Chord Recognition	253
5.2.1	Basic Approach	254
5.2.2	Evaluation	257
5.2.3	Ambiguities in Chord Recognition	260
5.2.4	Enhancement Strategies	266
5.3	HMM-Based Chord Recognition	273
5.3.1	Markov Chains and Transition Probabilities	273
5.3.2	Hidden Markov Models	276
5.3.3	Evaluation and Model Specification	279
5.3.4	Application to Chord Recognition	287
5.4	Further Notes	293
	References	297
	Exercises	300
6	Tempo and Beat Tracking	303
6.1	Onset Detection	305
6.1.1	Energy-Based Novelty	306
6.1.2	Spectral-Based Novelty	309
6.1.3	Phase-Based Novelty	313
6.1.4	Complex-Domain Novelty	315
6.2	Tempo Analysis	316
6.2.1	Tempogram Representations	317
6.2.2	Fourier Tempogram	319
6.2.3	Autocorrelation Tempogram	321

6.2.4	Cyclic Tempogram	325
6.3	Beat and Pulse Tracking	328
6.3.1	Predominant Local Pulse	329
6.3.2	Beat Tracking by Dynamic Programming	333
6.3.3	Adaptive Windowing	338
6.4	Further Notes	341
6.4.1	Onset Detection	343
6.4.2	Tempo Analysis and Beat Tracking	343
6.4.3	Applications	345
	References	347
	Exercises	351
7	Content-Based Audio Retrieval	355
7.1	Audio Identification	357
7.1.1	General Requirements	358
7.1.2	Audio Fingerprints Based on Spectral Peaks	360
7.1.3	Indexing, Retrieval, Inverted Lists	364
7.1.4	Index-Based Audio Identification	367
7.2	Audio Matching	371
7.2.1	General Requirements and Feature Design	371
7.2.2	Diagonal Matching	376
7.2.3	DTW-Based Matching	379
7.3	Version Identification	384
7.3.1	Versions in Music	385
7.3.2	Identification Procedure	389
7.3.3	Evaluation Measures	394
7.4	Further Notes	399
7.4.1	Audio Identification	400
7.4.2	Audio Matching	402
7.4.3	Version Identification	404
7.4.4	Alignment Scenarios	405
7.4.5	Category-Based Retrieval	407
	References	408
	Exercises	411
8	Musically Informed Audio Decomposition	415
8.1	Harmonic–Percussive Separation	417
8.1.1	Horizontal–Vertical Spectrogram Decomposition	420
8.1.2	Signal Reconstruction	425
8.1.3	Applications	429
8.2	Melody Extraction	431
8.2.1	Instantaneous Frequency Estimation	434
8.2.2	Saliency Representation	439
8.2.3	Informed Fundamental Frequency Tracking	444
8.3	NMF-Based Audio Decomposition	450

- 8.3.1 Nonnegative Matrix Factorization 452
- 8.3.2 Spectrogram Factorization 459
- 8.3.3 Audio Decomposition 464
- 8.4 Further Notes 468
 - 8.4.1 Harmonic–Percussive Separation 469
 - 8.4.2 Melody Extraction 470
 - 8.4.3 NMF-Based Audio Decomposition 472
- References 474
- Exercises 479

- Index** 481



<http://www.springer.com/978-3-319-21944-8>

Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications

Müller, M.

2015, XXIX, 487 p. 249 illus., 30 illus. in color.,

Hardcover

ISBN: 978-3-319-21944-8