Contents

1 Introduction to Music Similarity and Retrieval 1
- 1.1 Music Information Retrieval 2
- 1.2 MIR from an Information Retrieval Perspective 3
  - 1.2.1 Retrieval Tasks and Applications in MIR 4
  - 1.2.2 Browsing Interfaces in MIR 6
  - 1.2.3 Recommendation Tasks and Applications in MIR 9
  - 1.2.4 MIR Beyond Retrieval, Browsing, and Recommendation 10
- 1.3 Music Similarity 12
  - 1.3.1 Computational Factors of Music Similarity 13
  - 1.3.2 Music Features 14
- 1.4 Contents of this Book 18
- 1.5 Evaluation of Music Similarity Algorithms 19
  - 1.5.1 Evaluation Using Prelabeled Data 20
  - 1.5.2 Evaluation Using Human Judgments 24
  - 1.5.3 Evaluation Using Listening Histories 26
  - 1.5.4 Music Collection and Evaluation in this Book 26
- 1.6 Further Reading 30

Part I Content-Based MIR

2 Basic Methods of Audio Signal Processing 33
- 2.1 Categorization of Acoustic Music Features 33
- 2.2 Simplified Scheme of a Music Content Feature Extractor 36
  - 2.2.1 Analog–Digital Conversion 37
  - 2.2.2 Framing and Windowing 39
  - 2.2.3 Fourier Transform 40
- 2.3 Common Low-Level Features 41
  - 2.3.1 Time Domain Features 43
  - 2.3.2 Frequency Domain Features 47
- 2.4 Summary 49
- 2.5 Further Reading 50

Bibliografische Informationen
http://d-nb.info/1083021893

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

3 Audio Feature Extraction for Similarity Measurement 51
3.1 Psychoacoustic Processing 51
3.1.1 Physical Measurement of Sound Intensity 52
3.1.2 Perceptual Measurement of Loudness 52
3.1.3 Perception of Frequency 53
3.2 Frame-Level Features and Similarity 55
3.2.1 Mel Frequency Cepstral Coefficients 55
3.2.2 Statistical Summarization of Feature Vectors 58
3.2.3 Vector Quantization 58
3.2.4 Gaussian Mixture Models 60
3.2.5 Single Gaussian Model 65
3.3 Block-Level Features and Similarity 67
3.3.1 Fluctuation Pattern 68
3.3.2 Logarithmic Fluctuation Pattern 70
3.3.3 Spectral Pattern 71
3.3.4 Correlation Pattern 74
3.3.5 Similarity in the Block-Level Feature Framework 77
3.4 Hubness and Distance Space Normalization 78
3.5 Summary 81
3.6 Further Reading 83

4 Semantic Labeling of Music 85
4.1 Genre Classification 86
4.2 Auto-tagging 90
4.2.1 Differences to Classification 90
4.2.2 Auto-Tagging Techniques 91
4.3 Mood Detection and Emotion Recognition 94
4.3.1 Models to Describe Human Emotion 94
4.3.2 Emotion Recognition Techniques 99
4.4 Summary 102
4.5 Further Reading 103

Part II Music Context-Based MIR

5 Contextual Music Meta-data: Comparison and Sources 107
5.1 Web-Based Music Information Retrieval 108
5.1.1 The Web as Source for Music Features 108
5.1.2 Comparison with Content-Based Methods 110
5.1.3 Applications Using Web Data 111
5.2 Data Formats for Web-Based MIR 112
5.3 Tags and Annotations 114
5.3.1 Expert Annotations 115
5.3.2 Collaborative Tagging 115
5.3.3 Games with a Purpose 117

5.4 Web Texts 119
5.4.1 Web Pages Related to Music 120
5.4.2 Biographies, Product Reviews, and Audio Blogs 126
5.4.3 Microblogs 127
5.5 Lyrics 128
5.5.1 Analysis of Lyrics on the Web 128
5.5.2 Retrieval and Correction 129
5.6 Summary 131
5.7 Further Reading 132

6 Contextual Music Similarity, Indexing, and Retrieval 133
6.1 Text-Based Features and Similarity Measures 133
6.1.1 Vector Space Model 134
6.1.2 Latent Semantic Indexing 139
6.1.3 Applications of Latent Factor Approaches 142
6.2 Text-Based Indexing and Retrieval 144
6.2.1 Pseudo Document Indexing 145
6.2.2 Document-Centered Rank-Based Scoring 146
6.2.3 Auto-Tag Indexing 147
6.3 Similarity Based on Co-occurrences 148
6.4 Combination with Audio Content Information 151
6.4.1 Combined Similarity Measures 151
6.4.2 Contextual Filtering 153
6.4.3 Combined Tag Prediction 154
6.5 Stylistic Analysis and Similarity 155
6.6 Summary 156
6.7 Further Reading 157

Part III User-Centric MIR

7 Listener-Centered Data Sources and Aspects: Traces of Music Interaction 161
7.1 Definition and Comparison of Listener-Centered Features 161
7.2 Personal Collections and Peer-to-Peer Network Folders 163
7.3 Listening Histories and Playlists 164
7.4 User Ratings 169
7.5 Modeling User Context 170
7.5.1 Sensor Data for Modeling User Context 170
7.5.2 Social Networks and User Connections 173
7.6 Factors of User Intentions 174
7.7 Summary 176
7.8 Further Reading 177

8 Collaborative Music Similarity and Recommendation 179
8.1 Similarity Estimation via Co-occurrence 180
8.2 Graph-Based and Distance-Based Similarity 182

8.3 Exploiting Latent Context from Listening Sessions 186
8.3.1 Latent Dirichlet Allocation 186
8.3.2 Case Study: Artist Clustering from Listening Events 187
8.3.3 Music Recommendation 190
8.4 Learning from Explicit and Implicit User Feedback 191
8.4.1 Memory-Based Collaborative Filtering 192
8.4.2 Model-Based Collaborative Filtering 194
8.5 Multimodal Combination 198
8.5.1 Hybrid Recommender Systems 198
8.5.2 Unified Metric Learning 205
8.6 Summary 208
8.7 Further Reading 210

Part IV Current and Future Applications of MIR

9 Applications 215
9.1 Music Information Systems 215
9.1.1 Band Members and Their Roles 216
9.1.2 Artist's or Band's Country of Origin 216
9.1.3 Album Cover Artwork 217
9.1.4 Data Representation 218
9.2 User Interfaces to Music Collections 218
9.2.1 Map-Based Interfaces 218
9.2.2 Other Intelligent Interfaces 230
9.3 Automatic Playlist Generation 233
9.4 Music Popularity Estimation 239
9.4.1 Popularity Estimation from Contextual Data Sources 240
9.4.2 Comparison of Data Sources 244
9.5 Summary 245

10 Grand Challenges and Outlook 247
10.1 Major Challenges 247
10.1.1 Methodological Challenges 248
10.1.2 Data-Related Challenges 249
10.1.3 User-Centric Challenges 250
10.1.4 General Challenges 251
10.2 Future Directions 252

A Description of the Toy Music Data Set 255
A.1 Electronic Music 255
A.2 Classical Music 257
A.3 Heavy Metal 259
A.4 Rap 261
A.5 Pop 263

References 266

Index 293