



# Fundamentals of Music Processing

## Preface

Meinard Müller

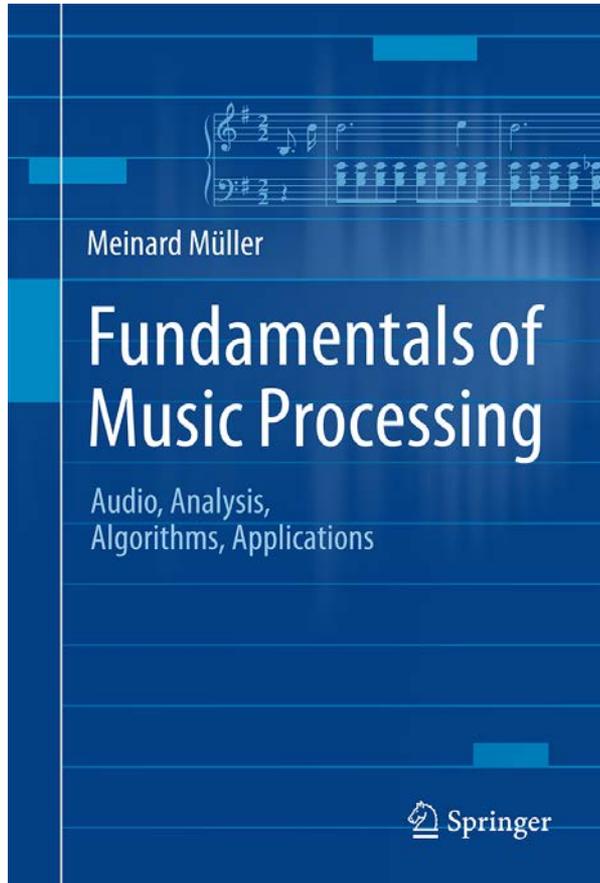
International Audio Laboratories Erlangen

[www.music-processing.de](http://www.music-processing.de)

 Springer

**AUDIO**  
**LABS**

# Book: Fundamentals of Music Processing



Meinard Müller  
Fundamentals of Music Processing  
Audio, Analysis, Algorithms, Applications  
483 p., 249 illus., 30 illus. in color, hardcover  
ISBN: 978-3-319-21944-8  
Springer, 2015

Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

# Preface

This textbook provides both profound technological knowledge and a comprehensive treatment of essential topics in music processing and music information retrieval. Including numerous examples, figures, and exercises, this book is suited for students, lecturers, and researchers working in audio engineering, computer science, multimedia, and musicology.

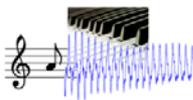
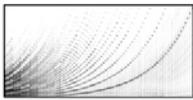
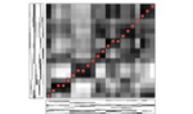
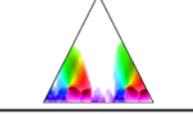
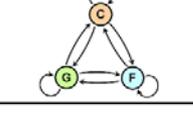
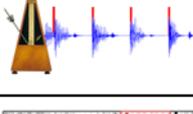
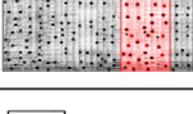
In the following, we give an overview of the book's content. The preface of the book contains further information on the overall structure of the book, the interconnections between the various topics and techniques, as well as suggestions on how this book may be used as a basis for different courses.

# Preface

This textbook consists of eight chapters. The first two chapters cover fundamental material on music representations and the Fourier transform—concepts that are required throughout the book. These two chapters make the book self-contained to a great extent. In the subsequent chapters, concrete music processing tasks serve as starting points for our investigations. Each of these chapters is organized in a similar fashion. A chapter starts with a general description of the music processing scenario at hand and integrates the topic into a wider context. Motivated by the scenario at hand, each chapter discusses important techniques and algorithms that are generally applicable to a wide range of analysis, classification, and retrieval problems. All these techniques are treated in a mathematically rigorous way. At the same time, the techniques are immediately applied to a concrete music processing task. By mixing theory and practice, the book's goal is to convey both profound technological knowledge as well as a solid understanding of music processing applications. Each of the chapters ends with a section that includes links to the research literature, hints for further reading, a list of references, and exercises. We now give an overview of the individual chapters and the main topics.

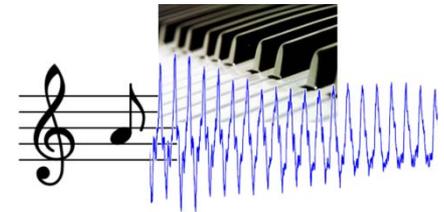
# Preface

## Overview

Chapter	Music Processing Scenario	Notions, Techniques & Algorithms
1 	<b>Music Representations</b>	Music notation, MIDI, audio signal, waveform, pitch, loudness, timbre
2 	<b>Fourier Analysis of Signals</b>	Discrete/analog signal, sinusoid, exponential, Fourier transform, Fourier representation, DFT, FFT, STFT
3 	<b>Music Synchronization</b>	Chroma feature, dynamic programming, dynamic time warping (DTW), alignment, user interface
4 	<b>Music Structure Analysis</b>	Similarity matrix, repetition, thumbnail, homogeneity, novelty, evaluation, precision, recall, F-measure, visualization, scape plot
5 	<b>Chord Recognition</b>	Harmony, music theory, chords, scales, templates, hidden Markov model (HMM), evaluation
6 	<b>Tempo and Beat Tracking</b>	Onset, novelty, tempo, tempogram, beat, periodicity, Fourier analysis, autocorrelation
7 	<b>Content-Based Audio Retrieval</b>	Identification, fingerprint, indexing, inverted list, matching, version, cover song
8 	<b>Musically Informed Audio Decomposition</b>	Harmonic/percussive component, signal reconstruction, instantaneous frequency, fundamental frequency (F0), trajectory, nonnegative matrix factorization (NMF)

# Chapter 1: Music Representations

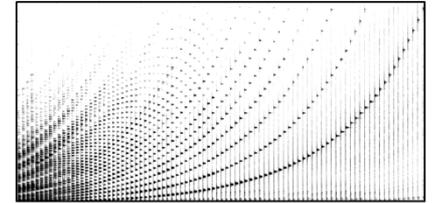
- 1.1 Sheet Music Representations
- 1.2 Symbolic Representations
- 1.3 Audio Representation
- 1.4 Further Notes



Musical information can be represented in many different ways. In Chapter 1, we consider three widely used music representations: sheet music, symbolic, and audio representations. This first chapter also introduces basic terminology that is used throughout the book. In particular, we discuss musical and acoustic properties of audio signals including aspects such as frequency, pitch, dynamics, and timbre.

# Chapter 2: Fourier Analysis of Signals

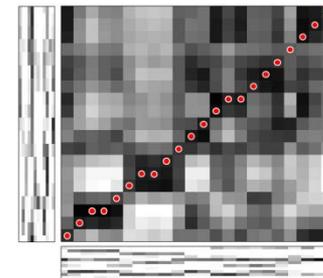
- 2.1 The Fourier Transform in a Nutshell
- 2.2 Signals and Signal Spaces
- 2.3 Fourier Transform
- 2.4 Discrete Fourier Transform (DFT)
- 2.5 Short-Time Fourier Transform (STFT)
- 2.6 Further Notes



Important technical terminology is covered in Chapter 2. In particular, we approach the Fourier transform—which is perhaps the most fundamental tool in signal processing—from various perspectives. For the reader who is more interested in the musical aspects of the book, Section 2.1 provides a summary of the most important facts on the Fourier transform. In particular, the notion of a spectrogram, which yields a time–frequency representation of an audio signal, is introduced. The remainder of the chapter treats the Fourier transform in greater mathematical depth and also includes the fast Fourier transform (FFT)—an algorithm of great beauty and high practical relevance.

# Chapter 3: Music Synchronization

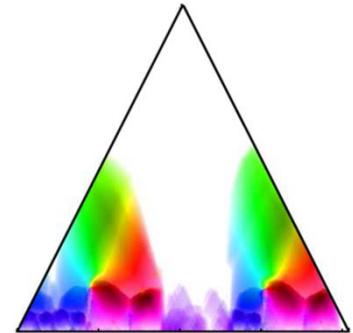
- 3.1 Audio Features
- 3.2 Dynamic Time Warping
- 3.3 Applications
- 3.4 Further Notes



As a first music processing task, we study in Chapter 3 the problem of music synchronization. The objective is to temporally align compatible representations of the same piece of music. Considering this scenario, we explain the need for musically informed audio features. In particular, we introduce the concept of chroma-based music features, which capture properties that are related to harmony and melody. Furthermore, we study an alignment technique known as dynamic time warping (DTW), a concept that is applicable for the analysis of general time series. For its efficient computation, we discuss an algorithm based on dynamic programming—a widely used method for solving a complex problem by breaking it down into a collection of simpler subproblems.

# Chapter 4: Music Structure Analysis

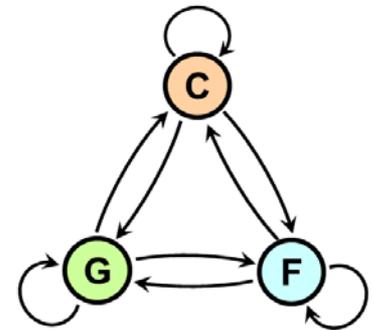
- 4.1 General Principles
- 4.2 Self-Similarity Matrices
- 4.3 Audio Thumbnailing
- 4.4 Novelty-Based Segmentation
- 4.5 Evaluation
- 4.6 Further Notes



In Chapter 4, we address a central and well-researched area within MIR known as music structure analysis. Given a music recording, the objective is to identify important structural elements and to temporally segment the recording according to these elements. Within this scenario, we discuss fundamental segmentation principles based on repetitions, homogeneity, and novelty—principles that also apply to other types of multimedia beyond music. As an important technical tool, we study in detail the concept of self-similarity matrices and discuss their structural properties. Finally, we briefly touch the topic of evaluation, introducing the notions of precision, recall, and F-measure.

# Chapter 5: Chord Recognition

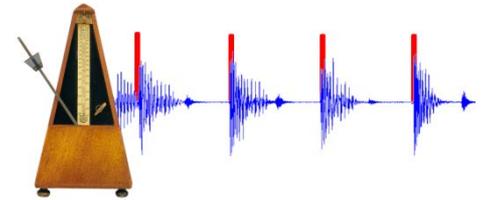
- 5.1 Basic Theory of Harmony
- 5.2 Template-Based Chord Recognition
- 5.3 HMM-Based Chord Recognition
- 5.4 Further Notes



In Chapter 5, we consider the problem of analyzing harmonic properties of a piece of music by determining a descriptive progression of chords from a given audio recording. We take this opportunity to first discuss some basic theory of harmony including concepts such as intervals, chords, and scales. Then, motivated by the automated chord recognition scenario, we introduce template-based matching procedures and hidden Markov models—a concept of central importance for the analysis of temporal patterns in time-dependent data streams including speech, gestures, and music.

# Chapter 6: Tempo and Beat Tracking

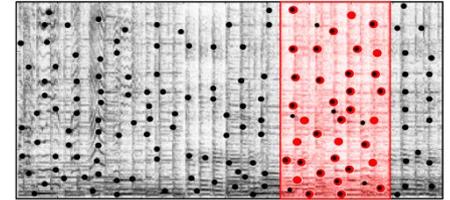
- 6.1 Onset Detection
- 6.2 Tempo Analysis
- 6.3 Beat and Pulse Tracking
- 6.4 Further Notes



Tempo and beat are further fundamental properties of music. In Chapter 6, we introduce the basic ideas on how to extract tempo-related information from audio recordings. In this scenario, a first challenge is to locate note onset information—a task that requires methods for detecting changes in energy and spectral content. To derive tempo and beat information, note onset candidates are then analyzed with regard to quasiperiodic patterns. This leads us to the study of general methods for local periodicity analysis of time series.

# Chapter 7: Content-Based Audio Retrieval

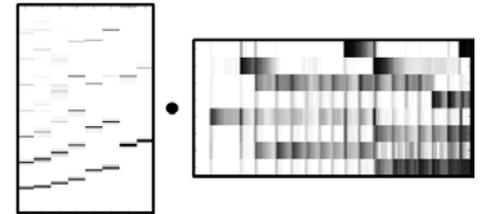
- 7.1 Audio Identification
- 7.2 Audio Matching
- 7.3 Version Identification
- 7.4 Further Notes



One important topic in information retrieval is concerned with the development of search engines that enable users to explore music collections in a flexible and intuitive way. In Chapter 7, we discuss audio retrieval strategies that follow the query-by-example paradigm: given an audio query, the task is to retrieve all documents that are somehow similar or related to the query. Starting with audio identification, a technique used in many commercial applications such as Shazam, we study various retrieval strategies to handle different degrees of similarity. Furthermore, considering efficiency issues, we discuss fundamental indexing techniques based on inverted lists—a concept originally used in text retrieval.

# Chapter 8: Musically Informed Audio Decomposition

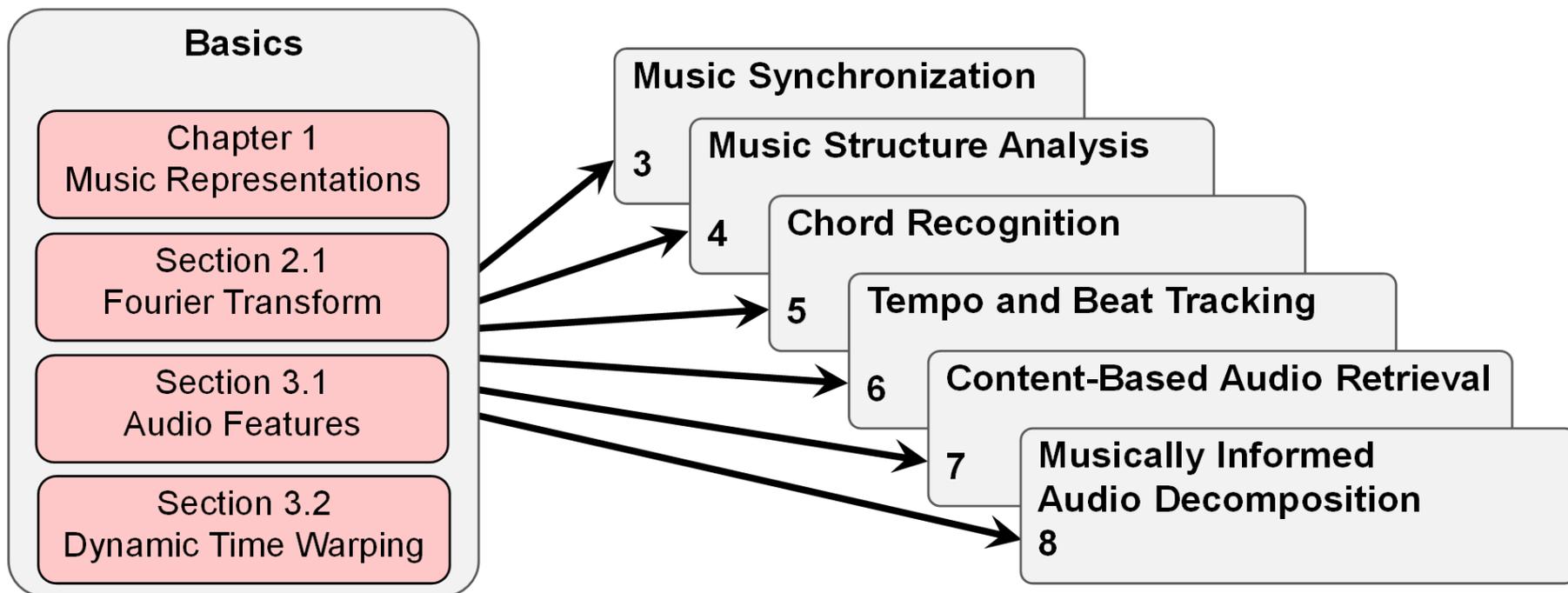
- 8.1 Harmonic-Percussive Separation
- 8.2 Melody Extraction
- 8.3 NMF-Based Audio Decomposition
- 8.4 Further Notes



In the final Chapter 8 on audio decomposition, we present a challenging research direction that is closely related to source separation. Within this wide research area, we consider three subproblems: harmonic–percussive separation, main melody extraction, and score-informed audio decomposition. Within these scenarios, we discuss a number of key techniques including instantaneous frequency estimation, fundamental frequency (F0) estimation, spectrogram inversion, and nonnegative matrix factorization (NMF). Furthermore, we encounter a number of acoustic and musical properties of audio recordings that have been introduced and discussed in previous chapters, which rounds off the book.

# Preface

View 1



# Preface

View 2

## Basics

Section 1.3  
Audio  
Representation

Section 2.1  
Fourier  
Transform

## Mathematical Theory

Section 2.2  
Signal Spaces

Section 2.3  
Fourier  
Transform

Section 2.4  
DFT, FFT

Section 2.5  
STFT

## Audio Features

Section 3.1.1  
Log-Frequency  
Spectrogram

Section 6.1.1  
Spectral-Based  
Novelty

Section 7.1.2  
Audio  
Fingerprints

## Phase Information

Section 6.1.2  
Phase-Based  
Novelty

Section 8.2.1  
Instantaneous  
Frequency  
Estimation

## Spectrogram Decomposition

Section 8.1.1  
Horizontal–  
Vertical  
Spectrogram  
Decomposition

Section 8.3.2  
NMF-based  
Spectrogram  
Factorization

# Preface

View 3

## Data Representations

Section 1.1  
Sheet Music Representation

Section 1.2  
Symbolic Representation

Section 1.3  
Audio Representation

Section 2.1  
Spectrogram Representation

Section 3.1  
Feature Representation

## Dynamic Programming

Section 3.2  
Dynamic Time Warping (DTW)

Section 7.2  
Audio Matching

Section 7.3  
Version Identification

Section 4.3  
Audio Thumbnailing

Section 6.3  
Beat and Pulse Tracking

## Further Algorithms

Section 2.4.3  
Fast Fourier Transform (FFT)

Section 5.3  
Hidden Markov Model (HMM)

Section 7.1.3  
Indexing, Retrieval, Inverted  
Lists

Section 8.3.1  
Nonnegative Matrix  
Factorization (NMF)