
Lecture
Music Processing

Beethoven, Bach, and Billions of Bytes

Meinard Müller

International Audio Laboratories Erlangen
meinard.mueller@audiolabs-erlangen.de

Prof. Dr. Meinard Müller



- PhD and Habilitation
Bonn University
- 2007 Senior Researcher
Multimedia Information Retrieval and Music
Processing
Max-Planck Institut für Informatik, Saarland
- 2012 Professor
Semantic Audio Processing
University of Erlangen-Nuremberg

Prof. Dr. Meinard Müller

Teaching assistants and current PhD students



Jonathan Driedger



Thomas Prätzlich



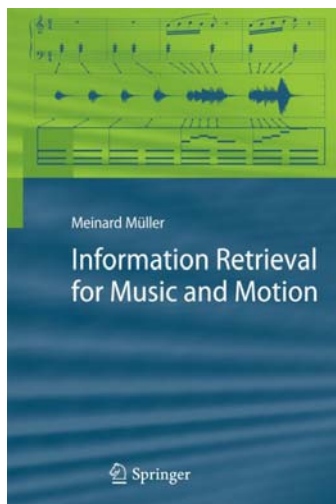
Nanzhu Jiang



Prof. Dr. Meinard Müller

Textbook:

Information Retrieval for Music and Motion

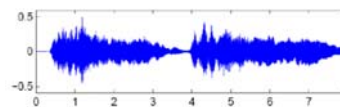


Music Representations

Sheet Music (Image)



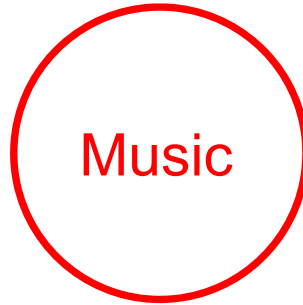
CD / MP3 (Audio)



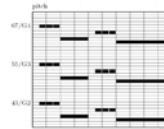
MusicXML (Text)

```
<note>
  <pitch>
    <step>E</step>
    <alter>-1</alter>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>half</type>
</note>
```

Dance / Motion (Mocap)



MIDI



Singing / Voice (Audio)



Music Film (Video)



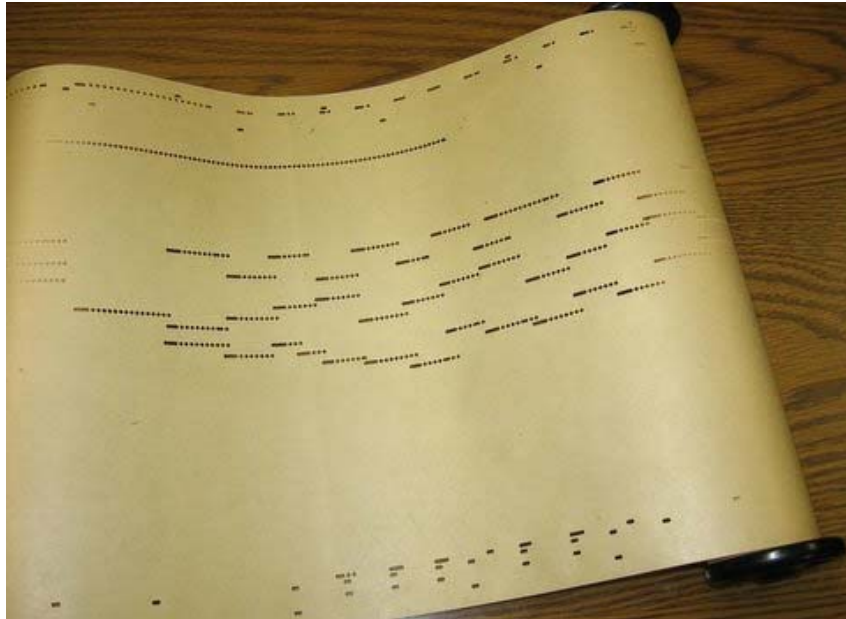
Music Literature (Text)



Research Goals

- Music Information Retrieval (MIR) → **ISMIR**
- Analysis of music signals (harmonic, melodic, rhythmic, motivic aspects)
- Design of musically relevant audio features
- Tools for multimodal search and interaction

Piano Roll Representation



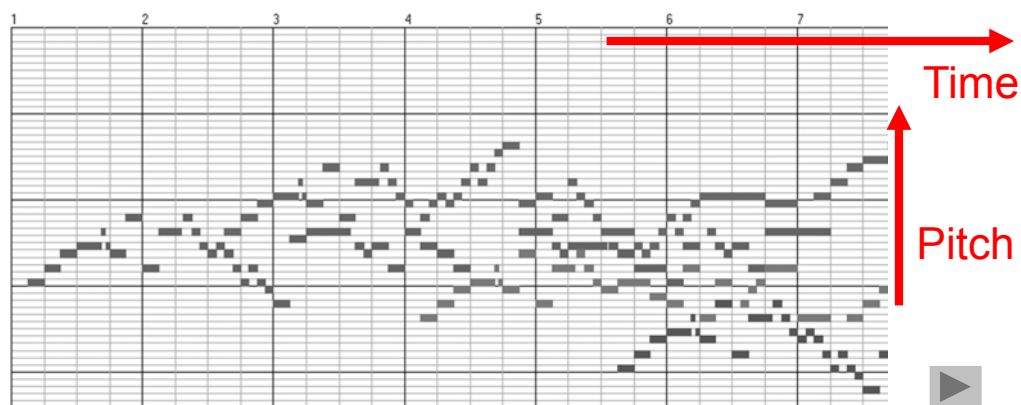
Player Piano (1900)



Piano Roll Representation (MIDI)

J.S. Bach, C-Major Fuge

(Well Tempered Piano, BWV 846)

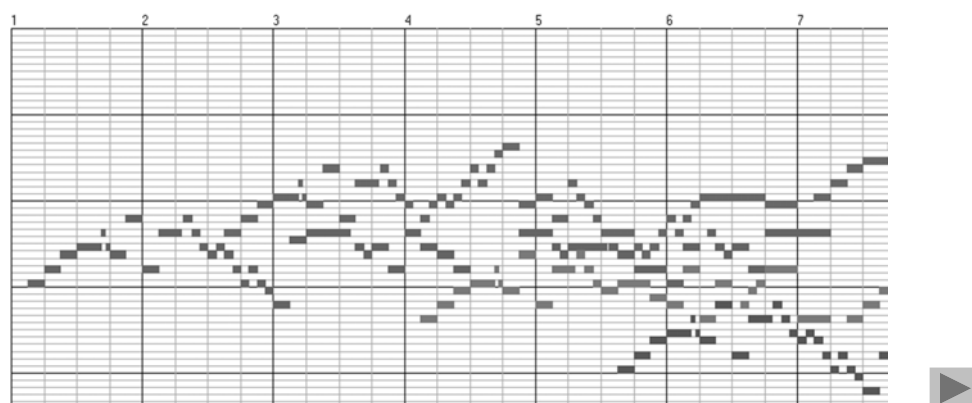


Piano Roll Representation (MIDI)

Query:

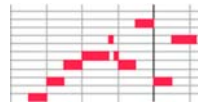


Goal: Find all occurrences of the query



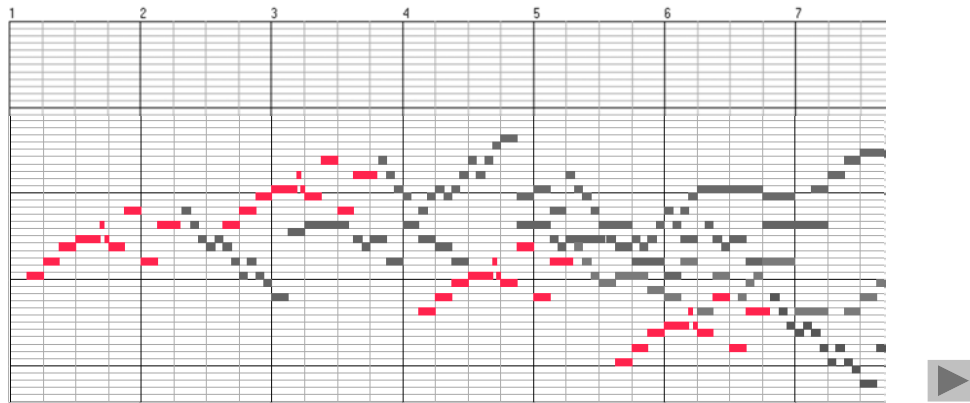
Piano Roll Representation (MIDI)

Query:



Goal: Find all occurrences of the query

Matches:



Music Synchronization: Audio-Audio

Beethoven's Fifth

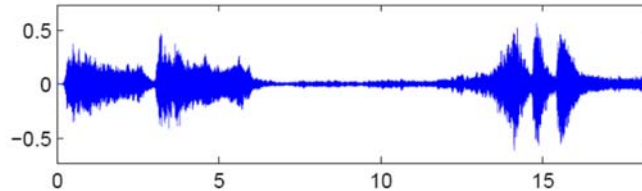


Music Synchronization: Audio-Audio

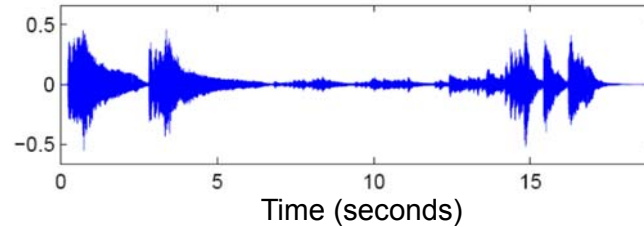
Beethoven's Fifth



Orchester
(Karajan)



Piano
(Scherbakov)

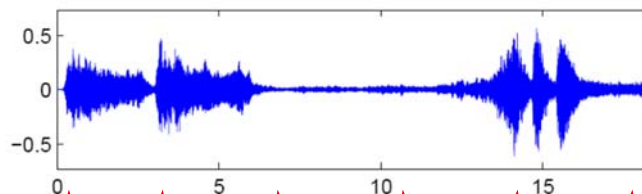


Music Synchronization: Audio-Audio

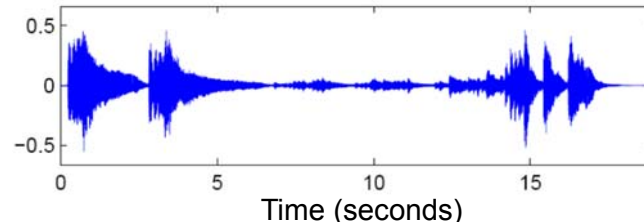
Beethoven's Fifth



Orchester
(Karajan)



Piano
(Scherbakov)



Application: Interpretation Switcher



Music Synchronization: Audio-Audio

Two main steps:

1.) Audio features

- Robust but discriminative
- Chroma features
- Robust to variations in instrumentation, timbre, dynamics
- Correlate to harmonic progression

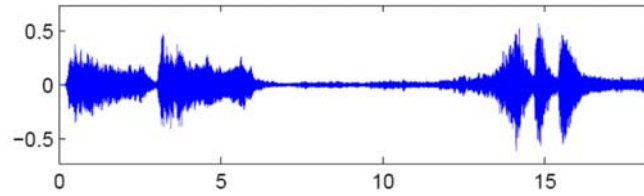
2.) Alignment procedure

- Deals with local and global tempo variations
- Needs to be efficient

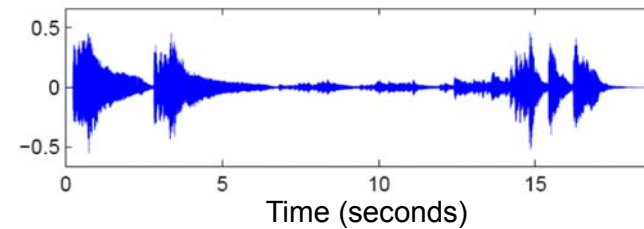
Music Synchronization: Audio-Audio

Beethoven's Fifth

Orchester
(Karajan)



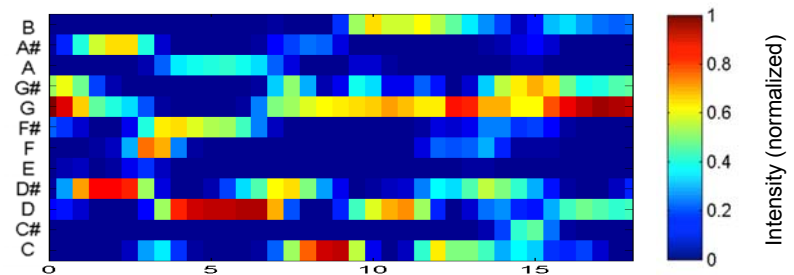
Piano
(Scherbakov)



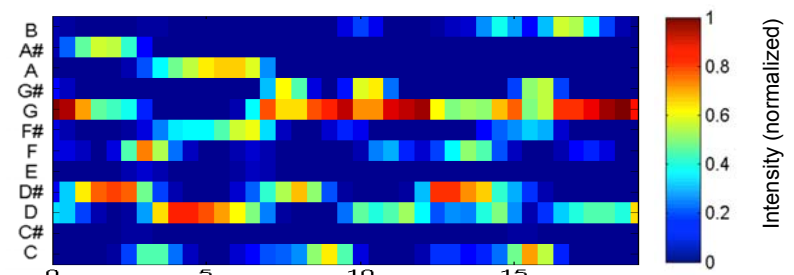
Music Synchronization: Audio-Audio

Beethoven's Fifth

Orchester
(Karajan)



Piano
(Scherbakov)



Time (seconds)

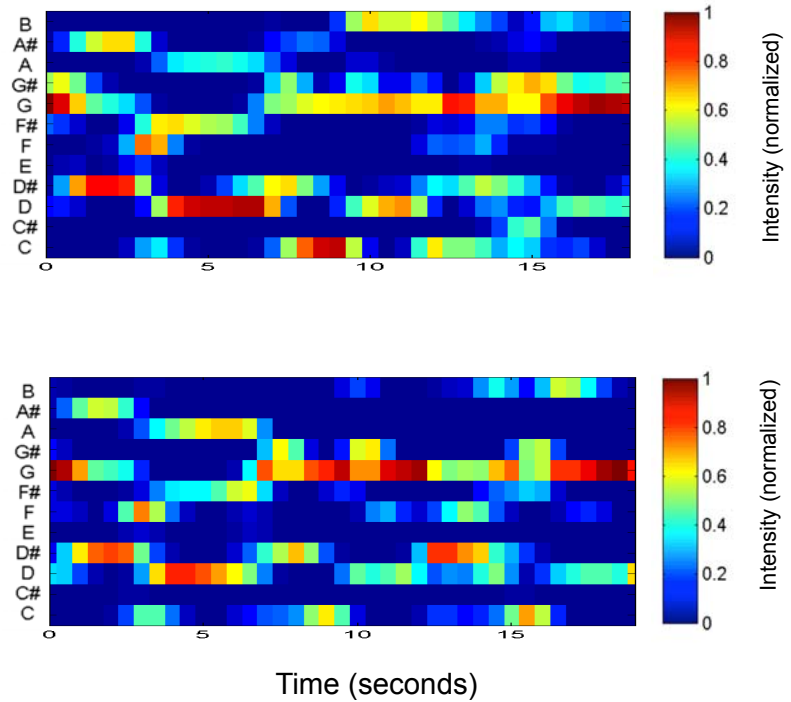
Music Synchronization: Audio-Audio

Beethoven's Fifth

Orchester
(Karajan)



Piano
(Scherbakov)



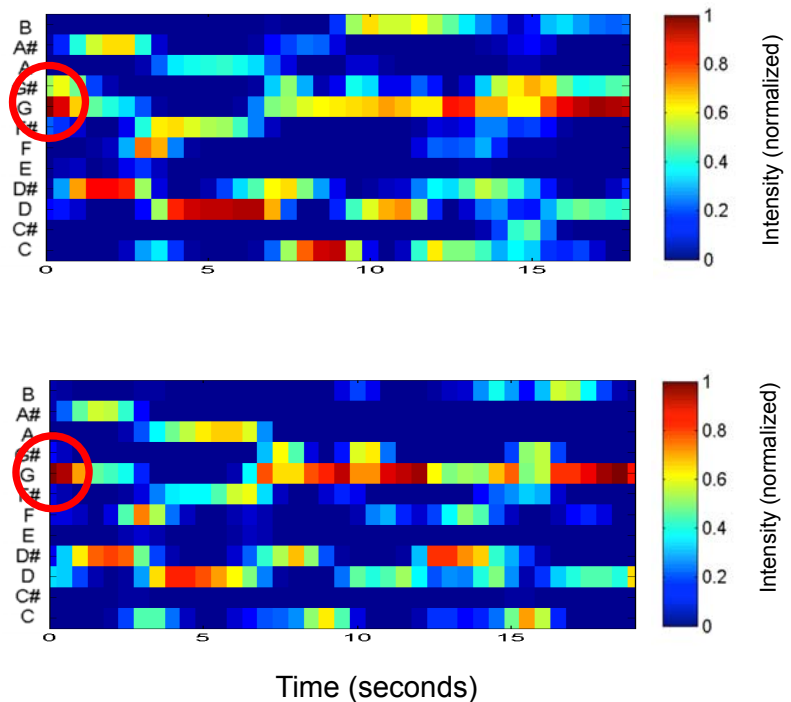
Music Synchronization: Audio-Audio

Beethoven's Fifth

Orchester
(Karajan)



Piano
(Scherbakov)



Time (seconds)

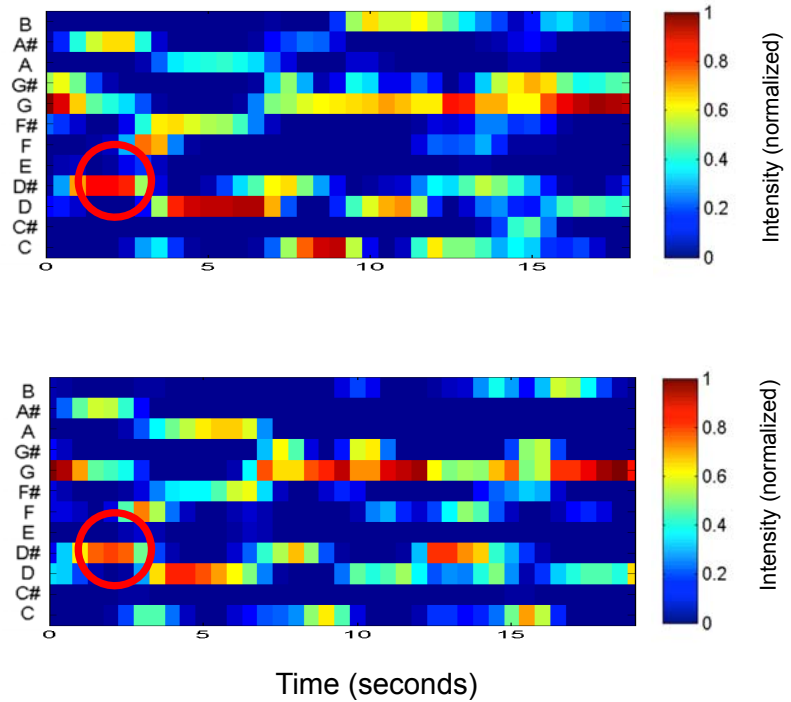
Music Synchronization: Audio-Audio

Beethoven's Fifth

Orchester
(Karajan)



Piano
(Scherbakov)



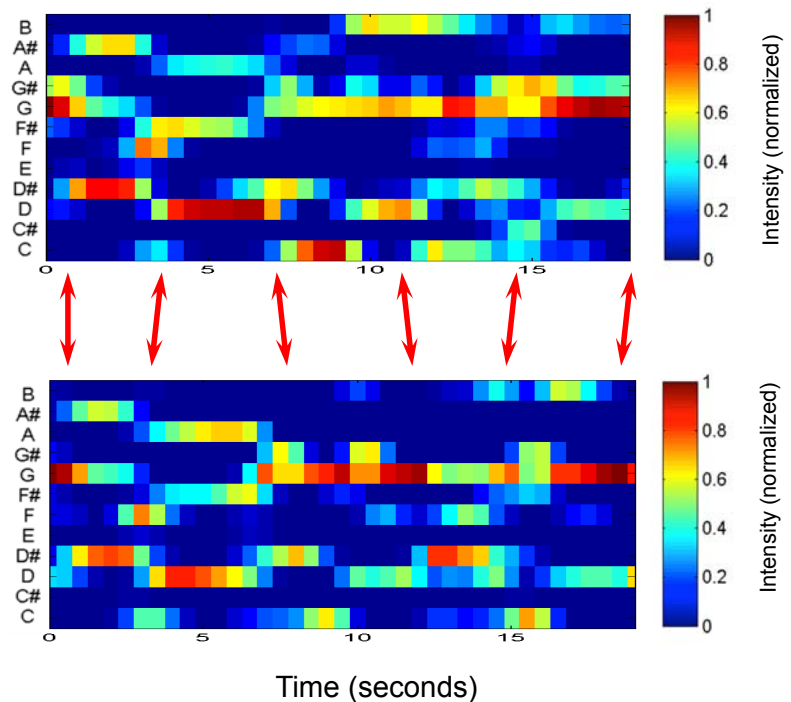
Music Synchronization: Audio-Audio

Beethoven's Fifth

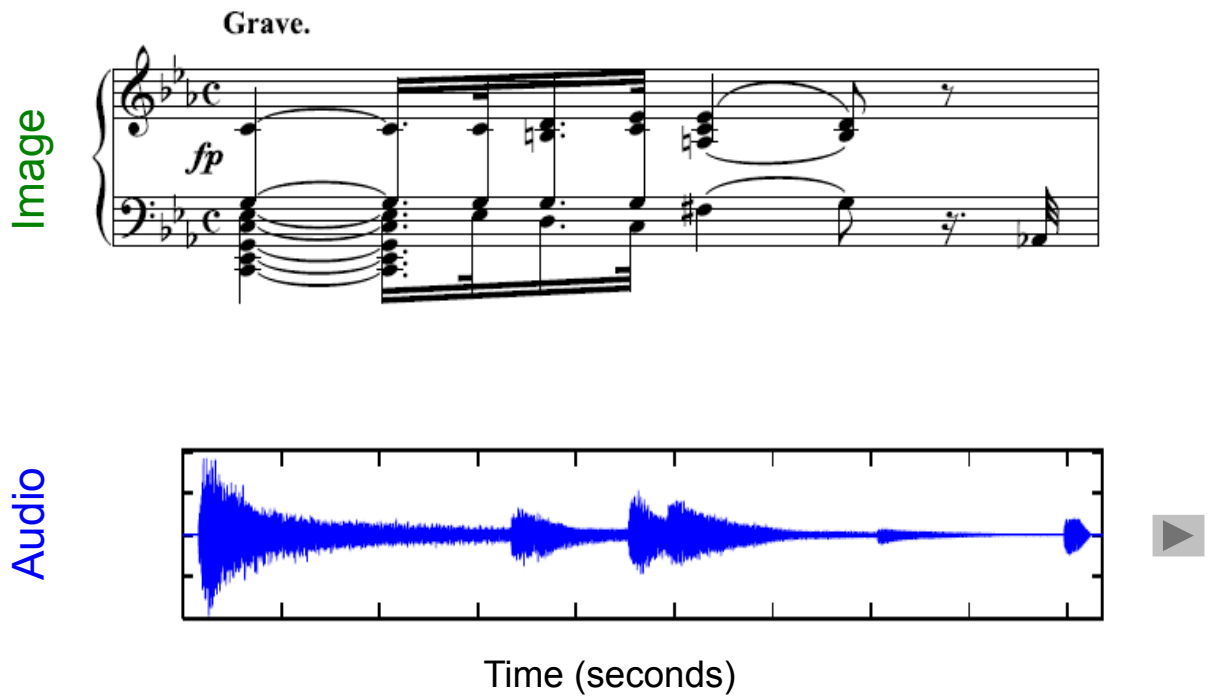
Orchester
(Karajan)



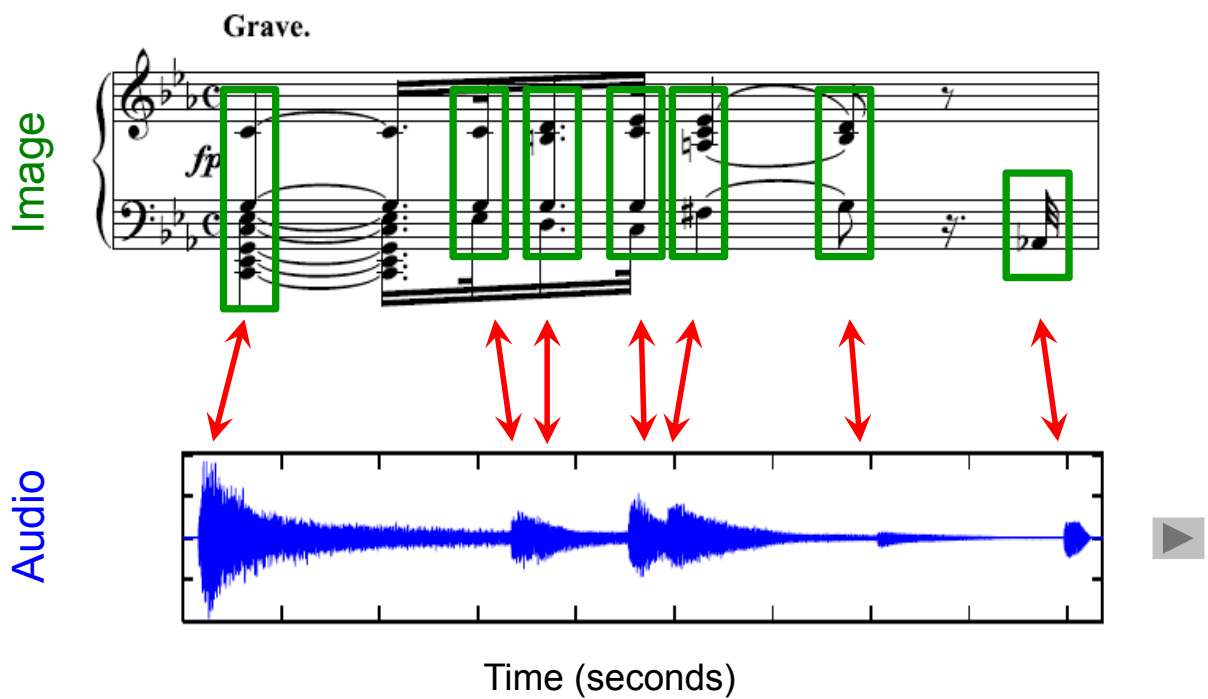
Piano
(Scherbakov)



Music Synchronization: Image-Audio

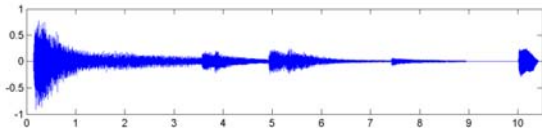


Music Synchronization: Image-Audio



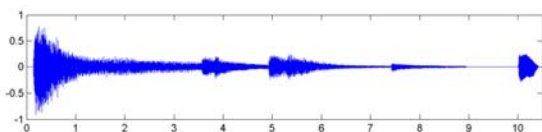
Music Synchronization: Image-Audio

Convert into common mid-level feature representation

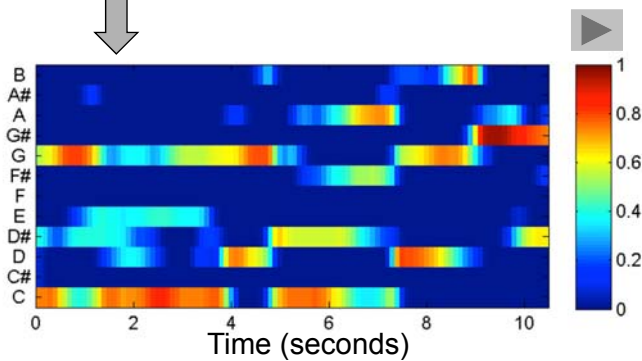


Music Synchronization: Image-Audio

Convert into common mid-level feature representation



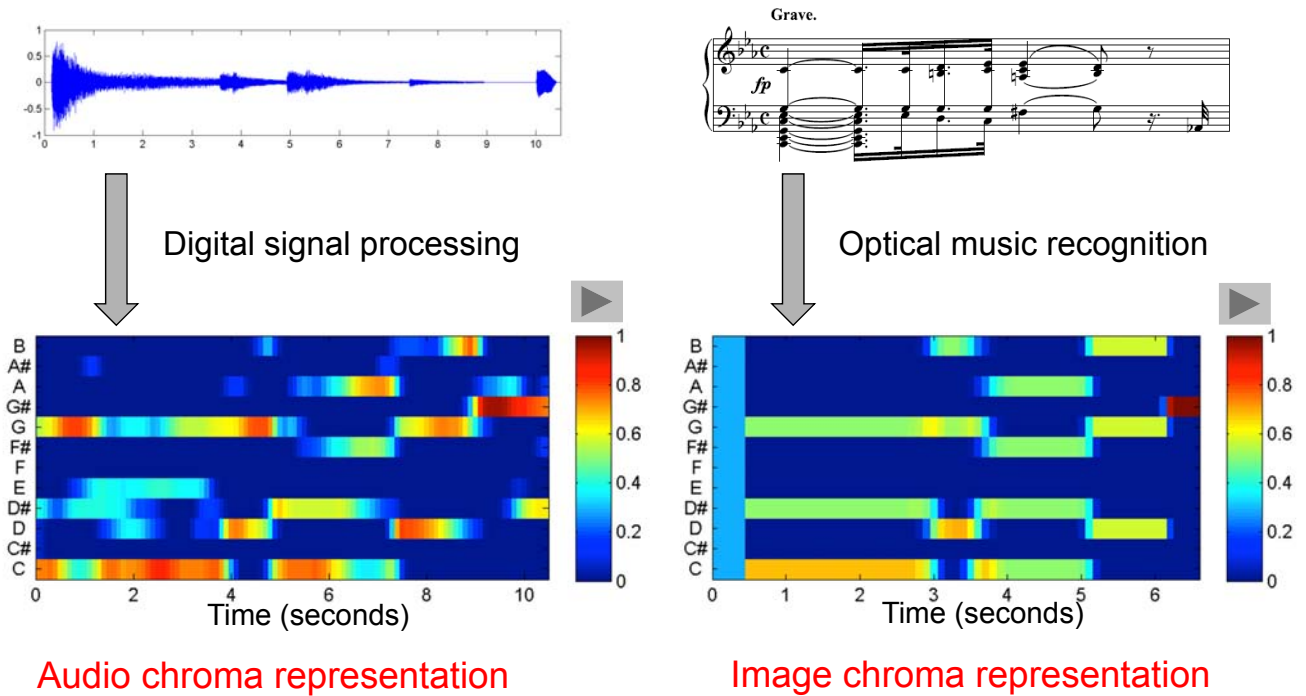
Digital signal processing



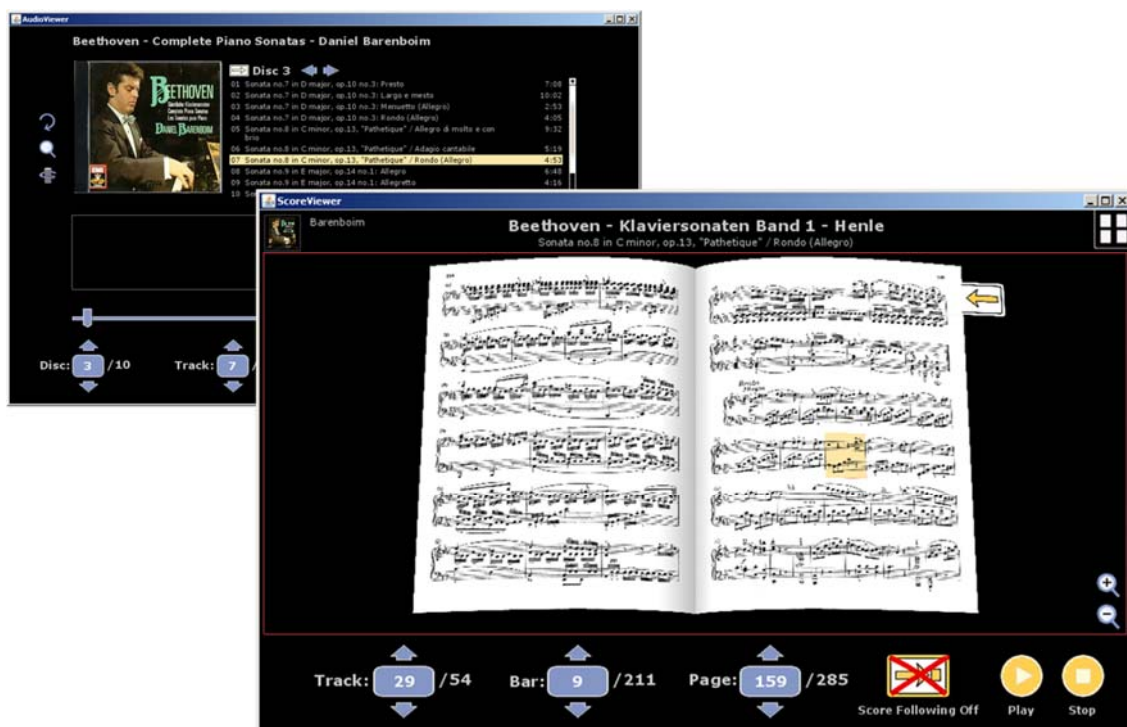
Audio chroma representation

Music Synchronization: Image-Audio

Convert into common mid-level feature representation



Application: Score Viewer

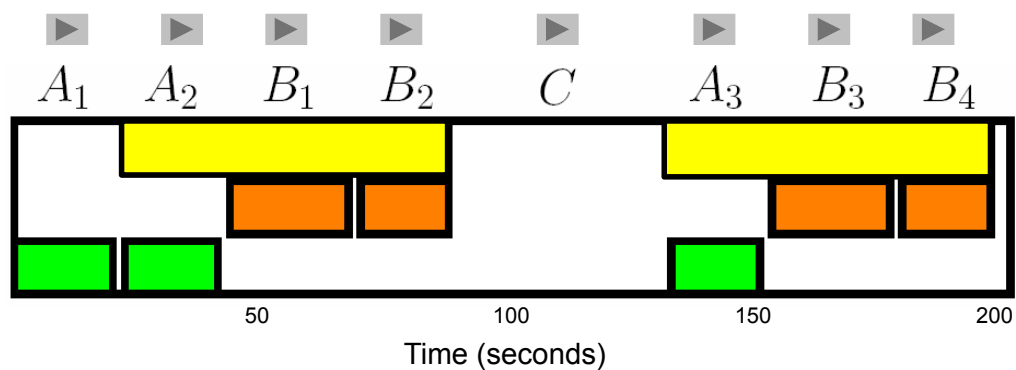


Audio Structure Analysis

Given: CD recording

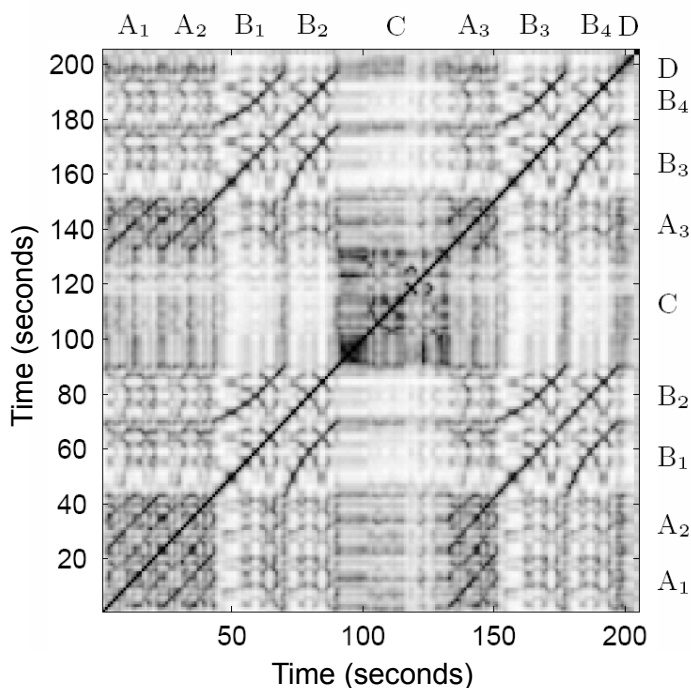
Goal: Automatic extraction of the **repetitive structure**
(or of the **musical form**)

Example: Brahms Hungarian Dance No. 5 (Ormandy)

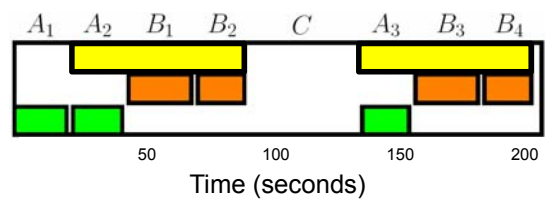


Basic Procedure

Self-similarity matrix

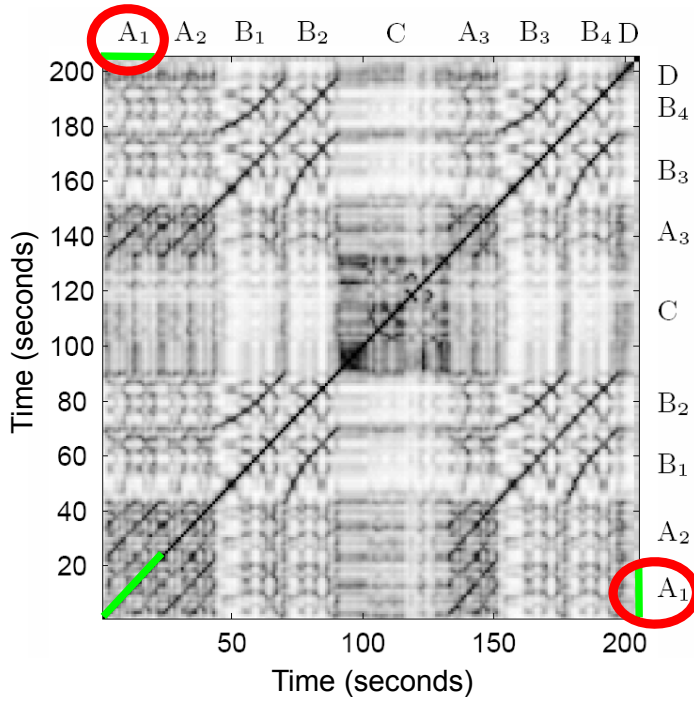


Similarity structure

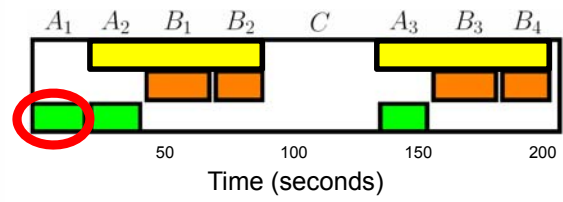


Basic Procedure

Self-similarity matrix

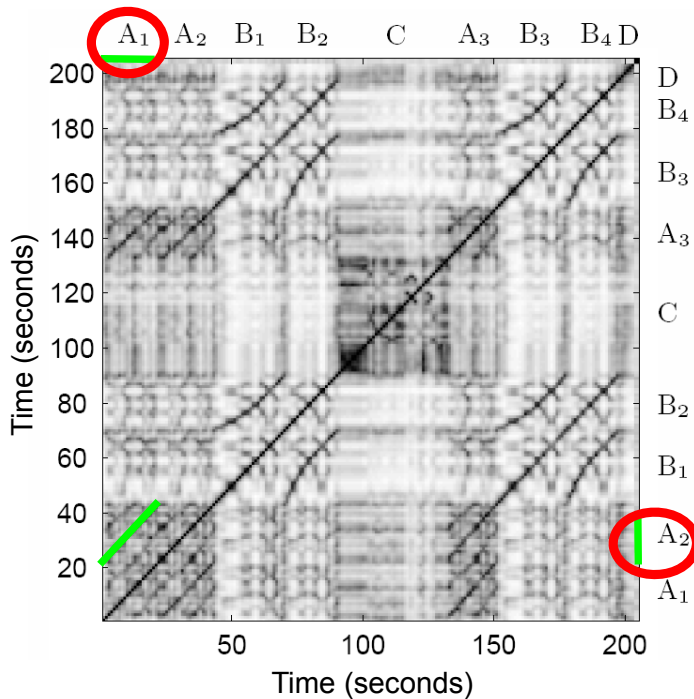


Similarity structure

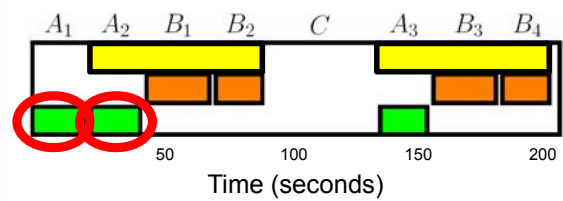


Basic Procedure

Self-similarity matrix

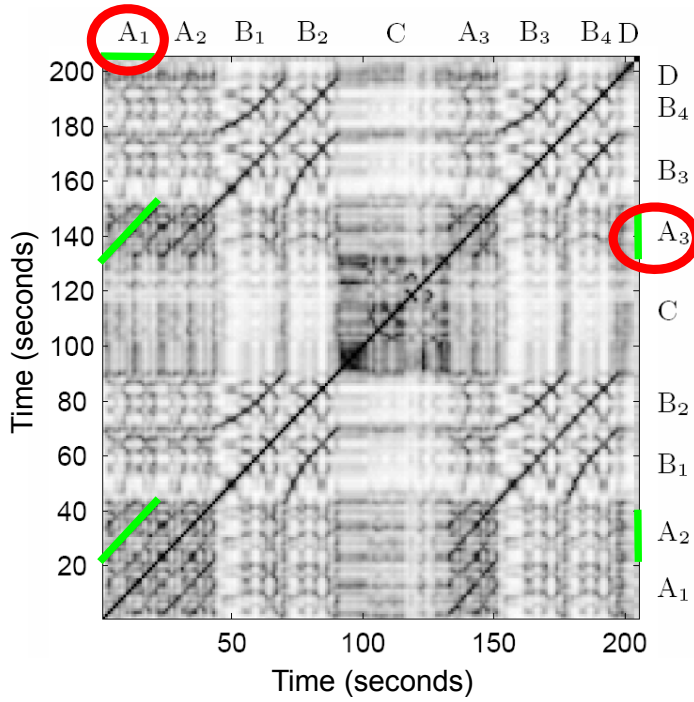


Similarity structure

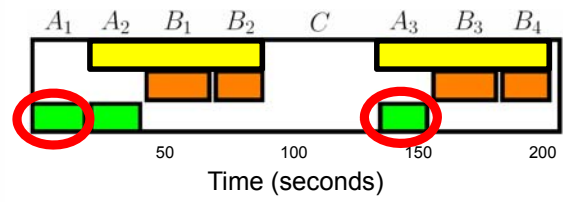


Basic Procedure

Self-similarity matrix

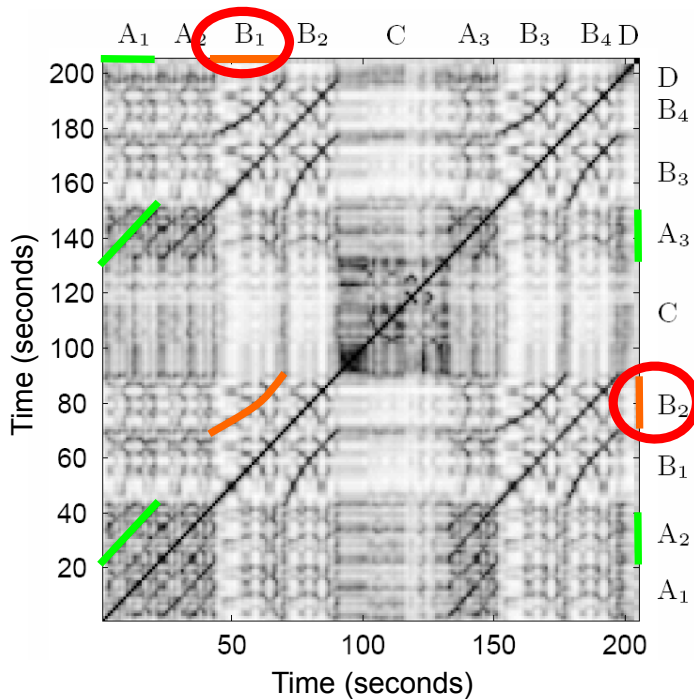


Similarity structure

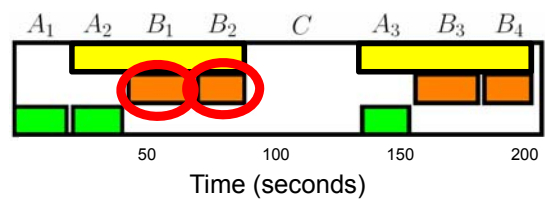


Basic Procedure

Self-similarity matrix

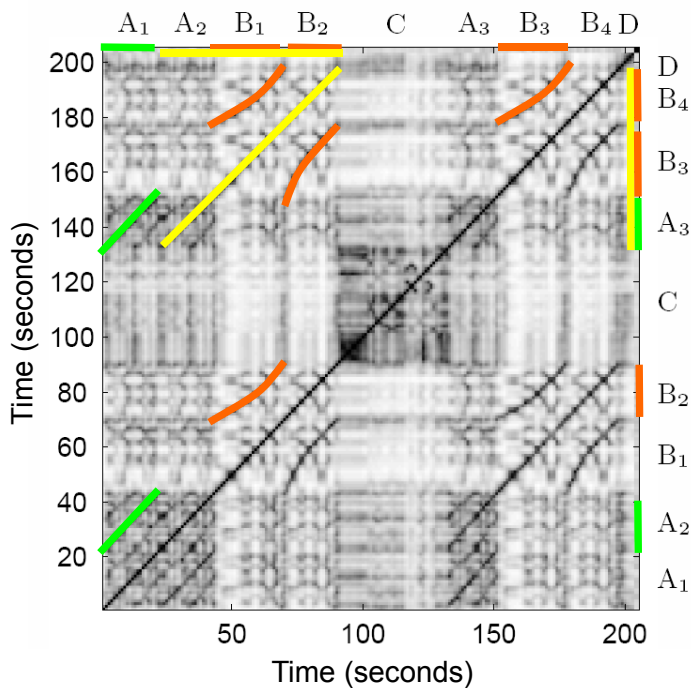


Similarity structure

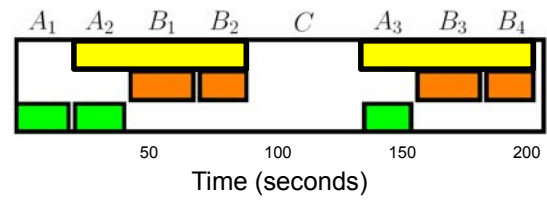


Basic Procedure

Self-similarity matrix



Similarity structure



Music Processing

Coarse Level	Fine Level
What do different versions have in common?	What are the characteristics of a specific version?

Music Processing

Coarse Level	Fine Level
What do different versions have in common?	What are the characteristics of a specific version?
What makes up a piece of music?	What makes music come alive?

Music Processing

Coarse Level	Fine Level
What do different versions have in common?	What are the characteristics of a specific version?
What makes up a piece of music?	What makes music come alive?
Identify despite of differences	Identify the differences

Music Processing

Coarse Level	Fine Level
What do different versions have in common?	What are the characteristics of a specific version?
What makes up a piece of music?	What makes music come alive?
Identify despite of differences	Identify the differences
Example tasks: Audio Matching Cover Song Identification	Example tasks: Tempo Estimation Performance Analysis

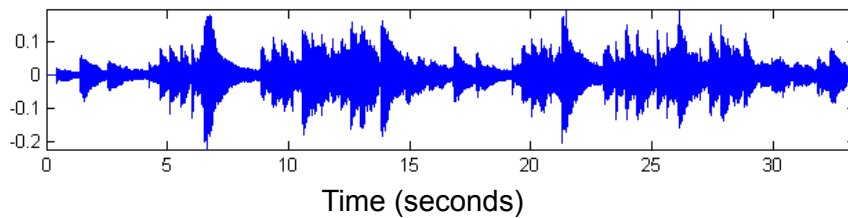
Performance Analysis

1. Capture nuances regarding tempo, dynamics, articulation, timbre, ...
2. Discover commonalities between different performances and derive general performance rules
3. Characterize the style of a specific musician (``Horowitz Factor``)

Performance Analysis: Tempo Curves

Schumann: Träumerei

Performance:



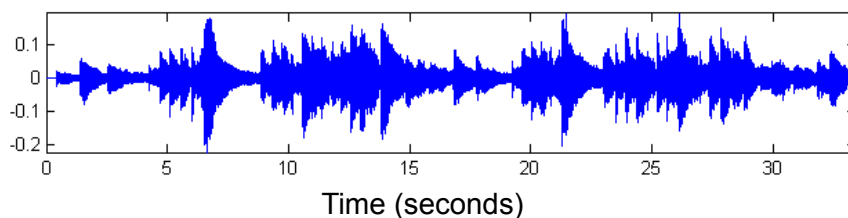
Performance Analysis: Tempo Curves

Schumann: Träumerei

Score (reference):



Performance:



Performance Analysis: Tempo Curves

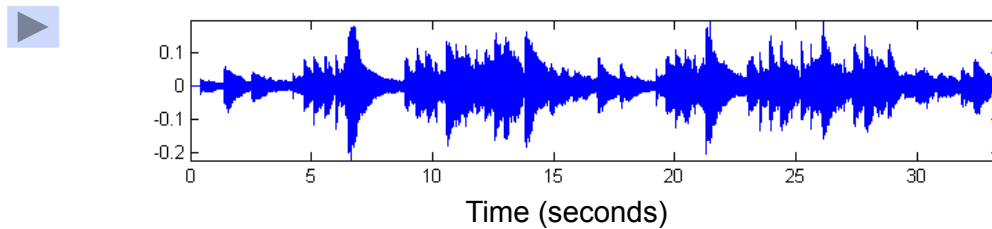
Schumann: Träumerei

Score (reference):



Strategy: Compute score-audio synchronization and derive tempo curve

Performance:



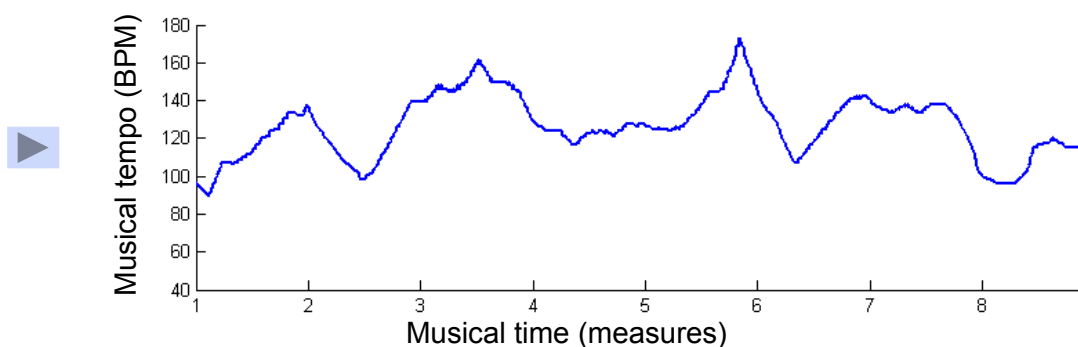
Performance Analysis: Tempo Curves

Schumann: Träumerei

Score (reference):



Tempo Curve:



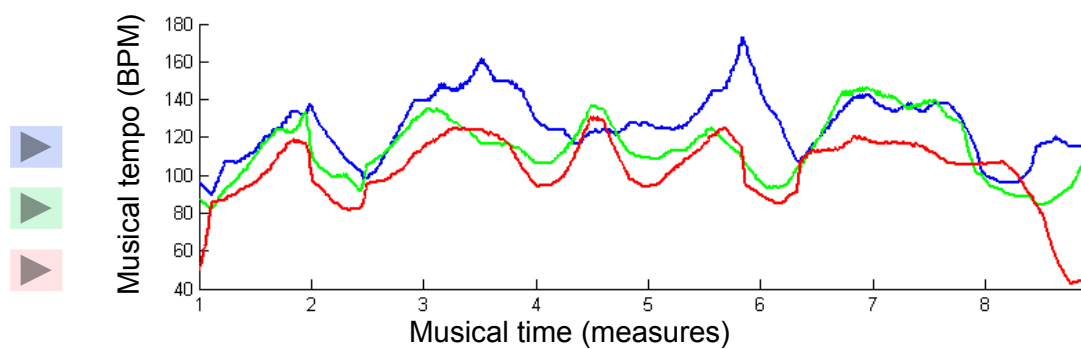
Performance Analysis: Tempo Curves

Schumann: Träumerei

Score (reference):



Tempo Curves:



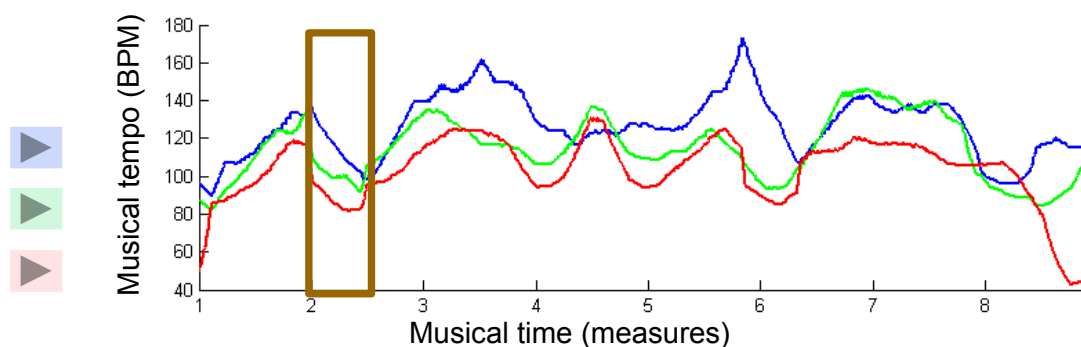
Performance Analysis: Tempo Curves

Schumann: Träumerei

Score (reference):



Tempo Curves:

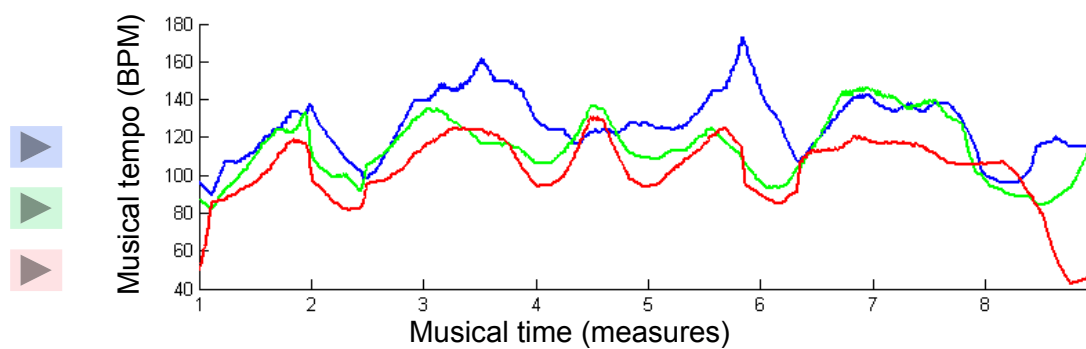


Performance Analysis

Schumann: Träumerei

What can be done if no reference is available?

Tempo Curves:



Music Processing

Relative	Absolute
Given: Several versions	Given: One version

Music Processing

Relative	Absolute
Given: Several versions	Given: One version
Comparison of extracted parameters	Direct interpretation of extracted parameters

Music Processing

Relative	Absolute
Given: Several versions	Given: One version
Comparison of extracted parameters	Direct interpretation of extracted parameters
Extraction errors have often no consequence on final result	Extraction errors immediately become evident

Music Processing

Relative	Absolute
Given: Several versions	Given: One version
Comparison of extracted parameters	Direct interpretation of extracted parameters
Extraction errors have often no consequence on final result	Extraction errors immediately become evident
Example tasks: Music Synchronization Genre Classification	Example tasks: Music Transcription Tempo Estimation

Tempo Estimation

Measure

The image shows two staves of musical notation for the song "Happy Birthday to you". The first staff is in 3/4 time and contains the lyrics: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py". The second staff continues with the lyrics: "Birth - day dear _____, Hap - py Birth - day to you!". Four red arrows point downwards to the first note of each of the four measures in the first staff, illustrating the concept of a measure in music.

Tempo Estimation

Tactus (beat)

Happy Birthday to you, Happy Birthday to you, Happy
Birthday dear _____, Happy Birthday to you!

The image shows a musical score for 'Happy Birthday to you' in 3/4 time. The melody is written on a treble clef staff with a key signature of one sharp (F#). The lyrics are: 'Happy Birthday to you, Happy Birthday to you, Happy Birthday dear _____, Happy Birthday to you!'. Red arrows point down to the first beat of each measure, indicating the Tactus (beat). There are 13 arrows in total, corresponding to the 13 measures of the piece.

Tempo Estimation

Tatum (temporal atom)

Happy Birthday to you, Happy Birthday to you, Happy
Birthday dear _____, Happy Birthday to you!

The image shows the same musical score for 'Happy Birthday to you' as above. However, instead of Tactus (beat) markers, red arrows point down to every individual note and rest in the melody, representing the Tatum (temporal atom). There are 26 arrows in total, corresponding to the 26 individual notes and rests in the piece.

Tempo Estimation and Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: ??? 

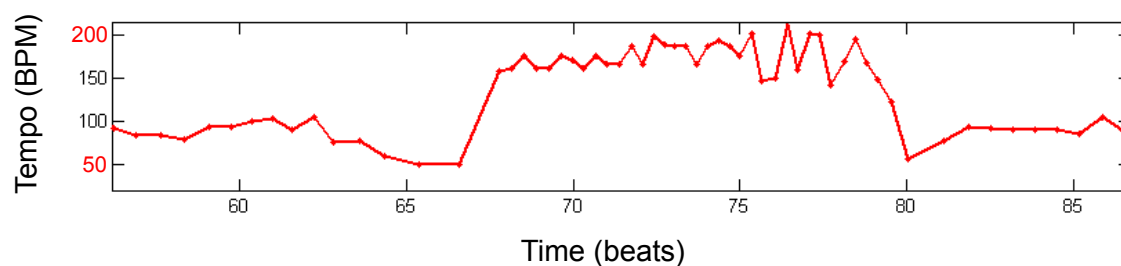
Tempo Estimation and Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: 50-200 BPM 

Tempo curve



Tempo Estimation

- Which temporal level?
- Local tempo deviations
- Sparse information
(e.g., only note onsets available)
- Vague information
(e.g., extracted note onsets corrupt)

Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3 

Mazurka.

F. CHOPIN. Op. 63, N^o 3.

Allegretto.

41. *p*



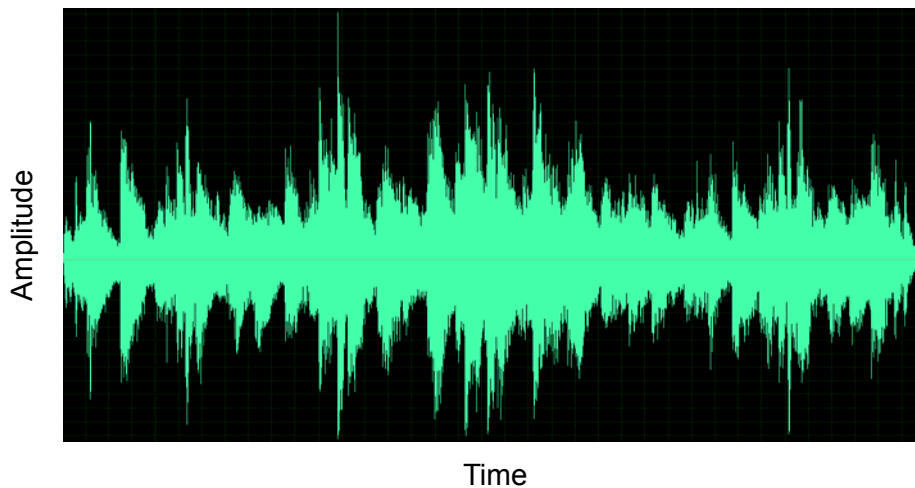
Rea * Rea * Rea * Rea * Rea *

Rea * Rea * Rea * Rea * Rea *

Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3 

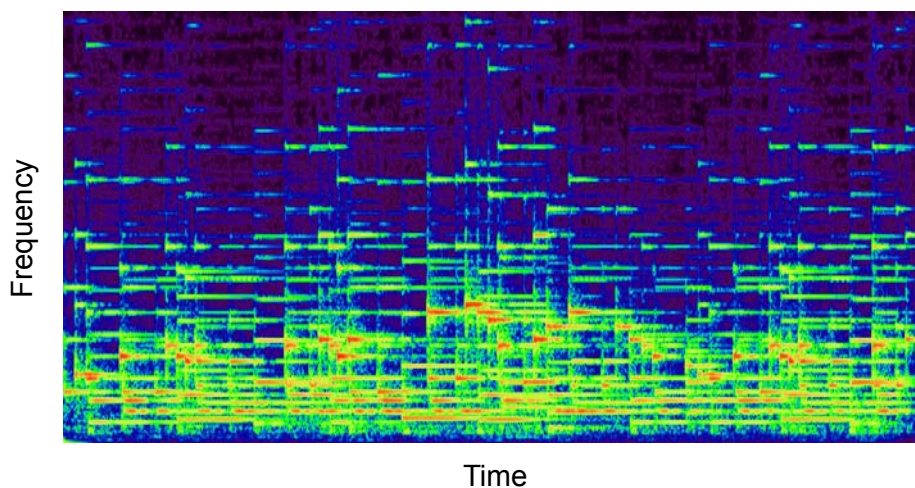
- Waveform



Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3 

- Waveform / Spectrogram




Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3 

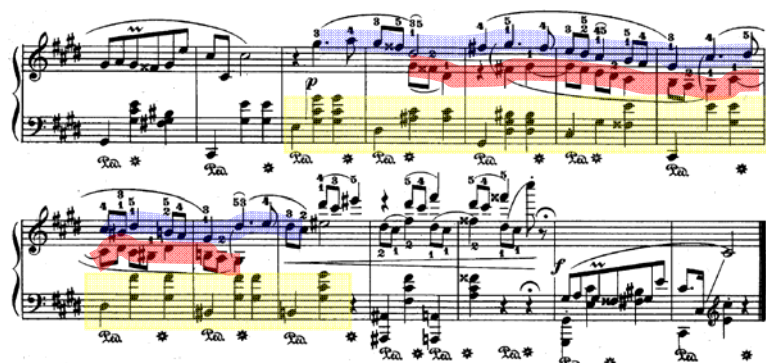
- Waveform / Spectrogram
- Performance
 - Tempo
 - Dynamics
 - Note deviations
 - Sustain pedal

Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3 




- Waveform / Spectrogram 

- Performance
 - Tempo
 - Dynamics
 - Note deviations
 - Sustain pedal



The image shows a musical score for Chopin's Mazurka Op. 63 No. 3, consisting of two systems of staves. The score is annotated with color-coded lines: blue for the main melody, red for an additional melody line, and yellow for the accompaniment. The annotations are applied to specific notes and chords across both systems. The score includes various musical notations such as notes, rests, and dynamic markings.

- Polyphony

	Main Melody
	Additional melody line
	Accompaniment

Motivic Similarity

Var. 4: Vivace

Musical score for Var. 4: Vivace, featuring four staves of music in bass clef with a key signature of two flats and a common time signature. The first staff begins with a forte (*f*) dynamic. The second and third staves continue the melodic and rhythmic patterns. The fourth staff includes dynamic markings of piano (*p*) and forte (*f*).



Motivic Similarity

Allegro con brio ($\text{♩} = 108$)

Musical score for Beethoven's Fifth (1st Mov.), featuring a grand staff with treble and bass clefs, a key signature of two flats, and a 2/4 time signature. The tempo is marked Allegro con brio ($\text{♩} = 108$). The score includes a forte (*ff*) dynamic marking and articulation marks (accents and asterisks) on the bass line.

Beethoven's Fifth (1st Mov.)



Motivic Similarity

Allegro con brio ($\text{♩} = 108$)

ff

rit. * rit. *

Beethoven's Fifth (1st Mov.)



Beethoven's Fifth (3rd Mov.)



Motivic Similarity

Allegro con brio ($\text{♩} = 108$)

ff

rit. * rit. *

Beethoven's Fifth (1st Mov.)



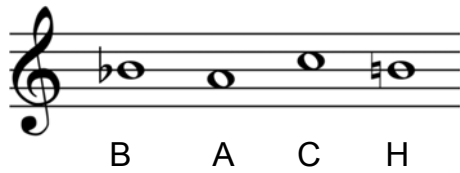
Beethoven's Fifth (3rd Mov.)



Beethoven's Appassionata



Motivic Similarity



S auf - - ge - rafft,
und nie - mand ach - - tet

A **b a c h** und nie - mand ach - - tet drauf

T und nie - mand ach - - tet

B auf - - ge - rafft,

Thanks

- Sebastian Ewert (Bonn University)
- Peter Grosche (Saarland University)
- Verena Konz (Saarland University)
- Christian Fremerey (Bonn University)
- David Damm (Bonn University)
- Verena Thomas (Bonn University)
- Michael Clausen (Bonn University)
- Frank Kurth (Fraunhofer-FKIE, Wachtberg)
- Jonathan Driedger (University of Erlangen-Nuremberg)
- Thomas Prätzlich (University of Erlangen-Nuremberg)
- Nanzhu Jiang (University of Erlangen-Nuremberg)

Selected Publications (Music Processing)

- M. Müller, P.W. Ellis, A. Klapuri, G. Richard (2011):
Signal Processing for Music Analysis.
IEEE Journal of Selected Topics in Signal Processing, Vol. 5, No. 6, pp. 1088-1110.
- P. Grosche and M. Müller (2011):
Extracting Predominant Local Pulse Information from Music Recordings.
IEEE Trans. on Audio, Speech & Language Processing, Vol. 19, No. 6, pp. 1688-1701.
- M. Müller, M. Clausen, V. Konz, S. Ewert, C. Fremerey (2010):
A Multimodal Way of Experiencing and Exploring Music.
Interdisciplinary Science Reviews (ISR), Vol. 35, No. 2.
- M. Müller and S. Ewert (2010):
Towards Timbre-Invariant Audio Features for Harmony-Based Music.
IEEE Trans. on Audio, Speech & Language Processing, Vol. 18, No. 3, pp. 649-662.
- F. Kurth, M. Müller (2008):
Efficient Index-Based Audio Matching.
IEEE Trans. Audio, Speech & Language Processing, Vol. 16, No. 2, 382-395.
- M. Müller (2007):
Information Retrieval for Music and Motion.
Monograph, Springer, 318 pages