


Ringvorlesung
Perspektiven der Informatik
Wintersemester 2010/2011

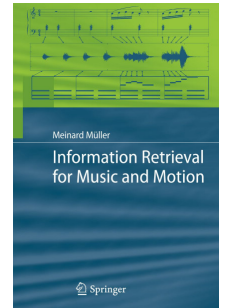
Meinard Müller
Universität des Saarlandes und MPI Informatik
meinard@mpi-inf.mpg.de

Music Signal Processing

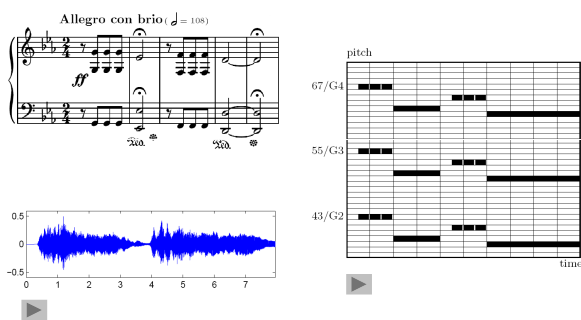


Priv.-Doz. Dr. Meinard Müller

- 2007 Habilitation, Bonn
- 2007 MPI Informatik, Saarland
- Cluster of Excellence

- 5 PhD Students
(2 Cluster, 3 DFG)



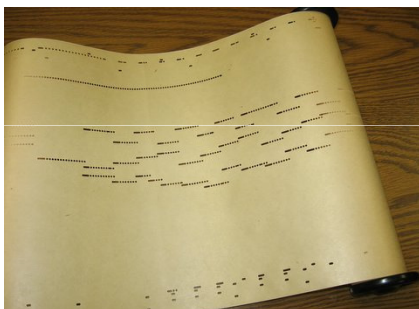
Music Data



Music Information Retrieval (MIR)

- Detection of semantic relations, e.g., harmonic, rhythmic, or motivic similarity
- Extraction of musical entities such as note events, instrumentation, or musical form
- Tools and methods for multimodal search, navigation, and interaction

Piano Roll Representation



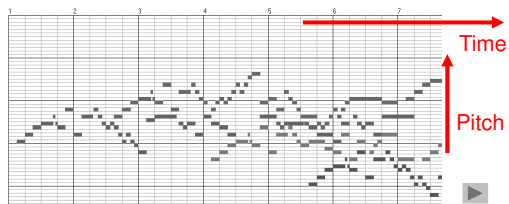
Piano Roll Representation

Player Piano (1900)



Piano Roll Representation (MIDI)

J.S. Bach, C-Major Fuge
(Well Tempered Piano, BWV 846)

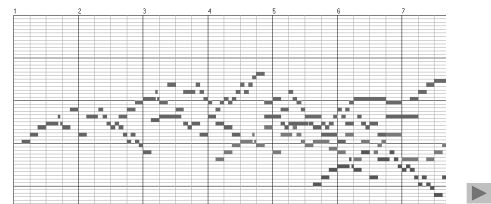


Piano Roll Representation (MIDI)

Query:



Goal: Find all occurrences of the query



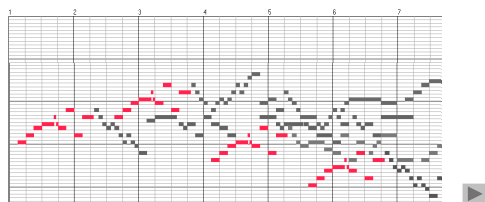
Piano Roll Representation (MIDI)

Query:



Goal: Find all occurrences of the query

Matches:



Audio Data

Various interpretations – Beethoven's Fifth

Bernstein



Karajan



Scherbakov (piano)



MIDI (piano)



Memory Requirements

1 Bit	=	1: on 0: off
1 Byte	=	8 Bits
1 Kilobyte (KB)	=	1 Thousand Bytes
1 Megabyte (MB)	=	1 Million Bytes
1 Gigabyte (GB)	=	1 Billion Bytes
1 Terabyte (TB)	=	1000 Billion Bytes

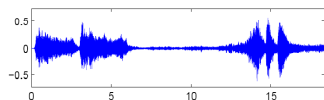
Memory Requirements

12.000 MIDI files	<	350 MB
One audio CD	≈	650 MB
Two audio CDs	>	1 Billion Bytes
1000 audio CDs	≈	Billions of Bytes

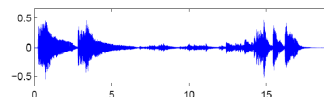
Music Synchronization: Audio-Audio

Beethoven's Fifth

Karajan



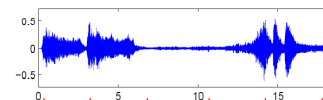
Scherbakov



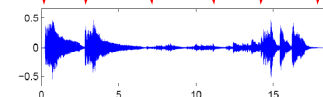
Music Synchronization: Audio-Audio

Beethoven's Fifth

Karajan



Scherbakov



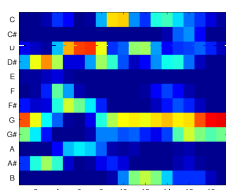
Synchronization: Karajan → Scherbakov



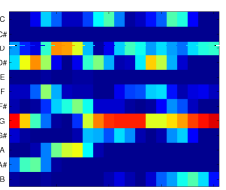
Music Synchronization: Audio-Audio

Feature extraction: chroma features

Karajan

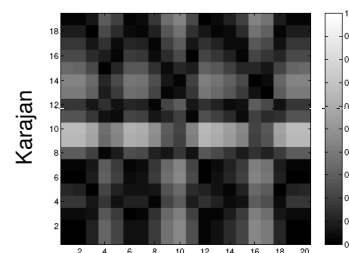


Scherbakov



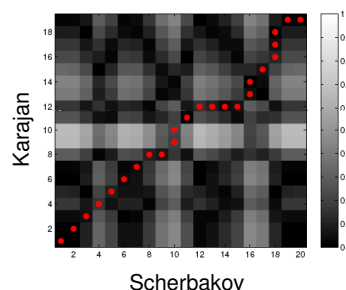
Music Synchronization: Audio-Audio

Cost matrix

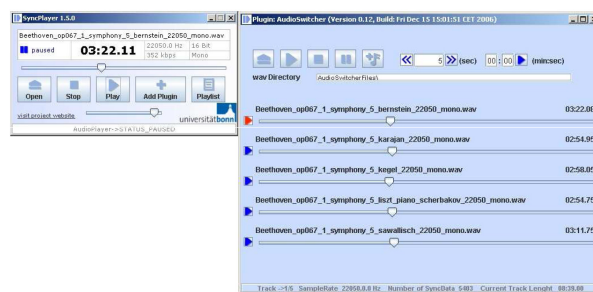


Music Synchronization: Audio-Audio

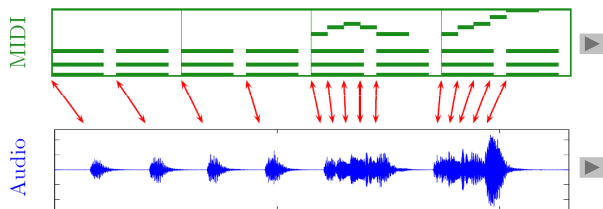
Cost-minimizing warping path



System: SyncPlayer/AudioSwitcher



Music Synchronization: MIDI-Audio



Music Synchronization: MIDI-Audio

MIDI = meta data

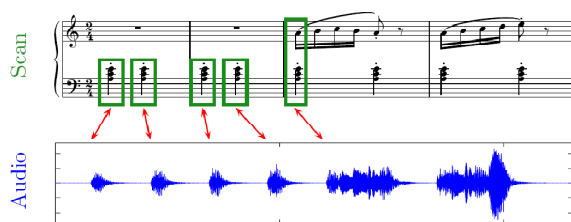
Automated annotation

Audio recording

Sonification of annotations



Music Synchronization: Scan-Audio

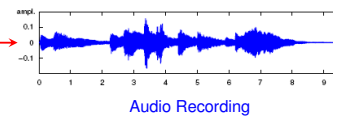


Music Synchronization: Scan-Audio

Scanned Sheet Music



Correspondence

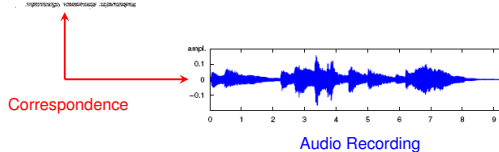


Music Synchronization: Scan-Audio

Scanned Sheet Music

Symbolic Note Events

OMR

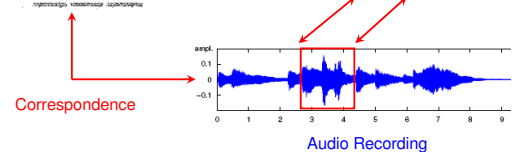


Music Synchronization: Scan-Audio

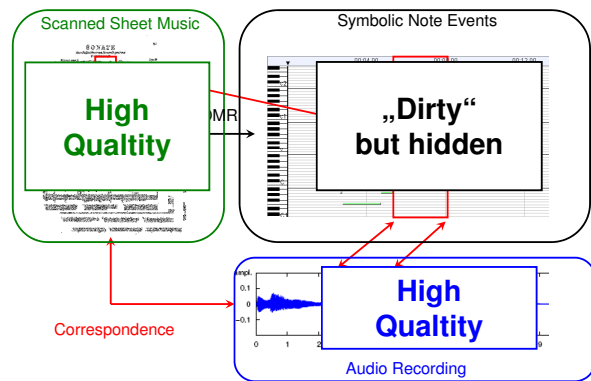
Scanned Sheet Music

Symbolic Note Events

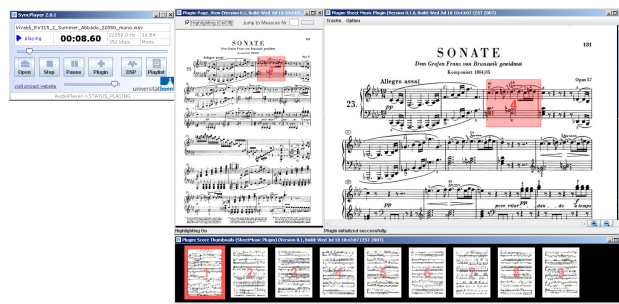
OMR



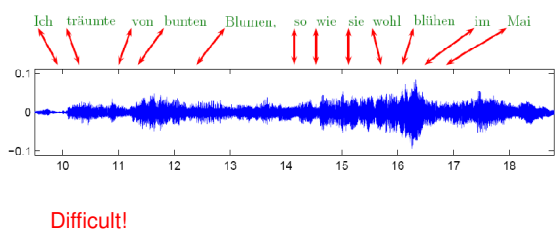
Music Synchronization: Scan-Audio



System: SyncPlayer/SheetMusic

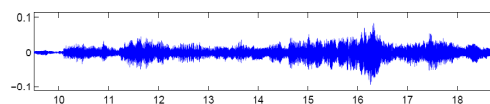


Music Synchronization: Lyrics-Audio

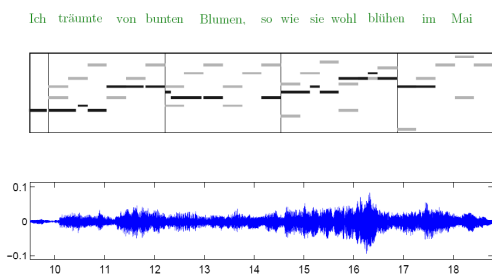


Music Synchronization: Lyrics-Audio

Ich träumte von bunten Blumen, so wie sie wohl blühen im Mai

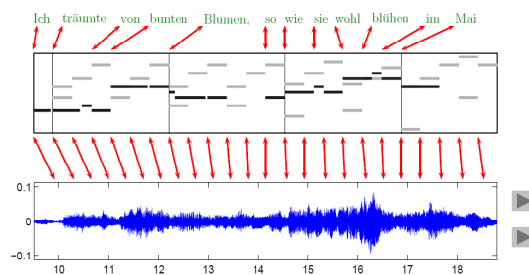


Music Synchronization: Lyrics-Audio



Music Synchronization: Lyrics-Audio

Lyrics-Audio → Lyrics-MIDI + MIDI-Audio

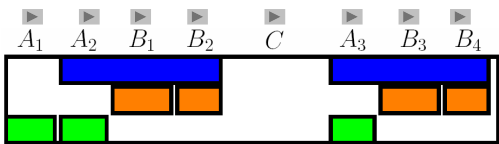


Audio Structure Analysis

Given: CD recording

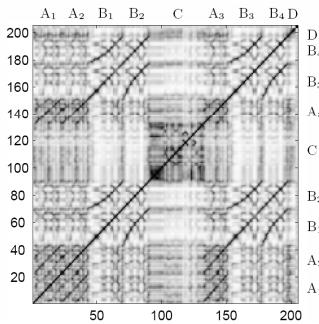
Goal: Automatic extraction of the repetitive structure (or of the musical form)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



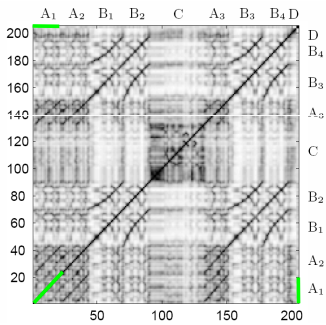
Audio Structure Analysis

Self-similarity matrix



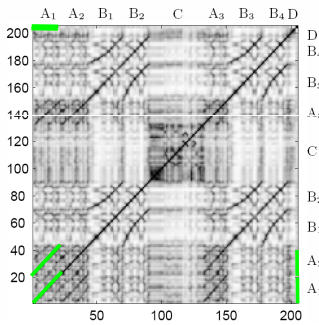
Audio Structure Analysis

Self-similarity matrix



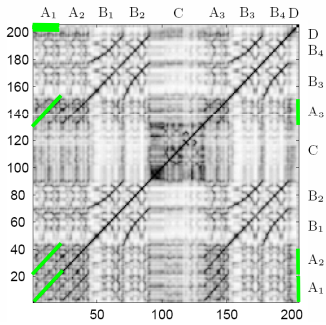
Audio Structure Analysis

Self-similarity matrix



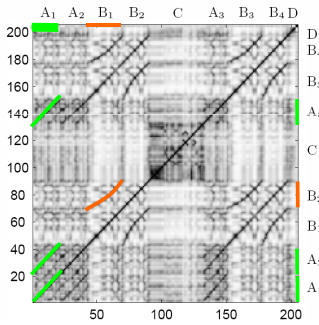
Audio Structure Analysis

Self-similarity matrix



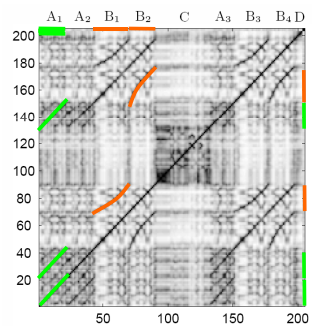
Audio Structure Analysis

Self-similarity matrix



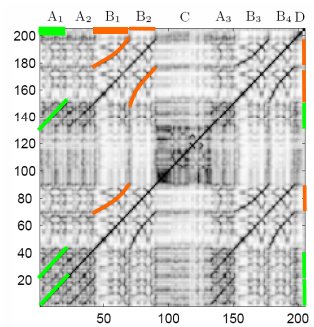
Audio Structure Analysis

Self-similarity matrix



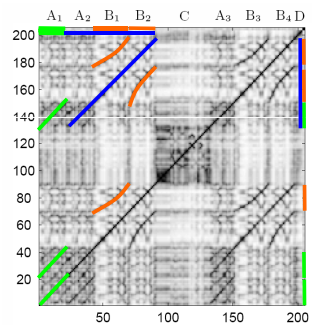
Audio Structure Analysis

Self-similarity matrix



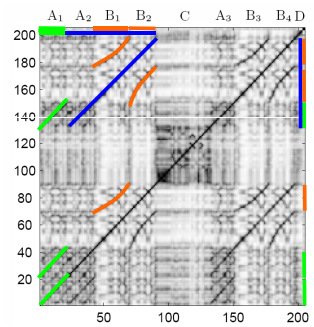
Audio Structure Analysis

Self-similarity matrix

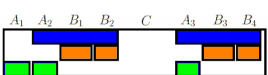


Audio Structure Analysis

Self-similarity matrix



Similarity cluster



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

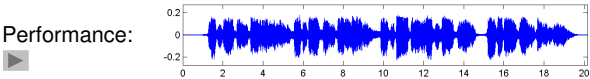
Music Processing

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

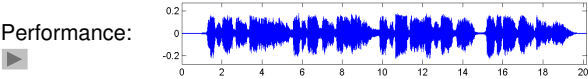
Performance Analysis

- 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´´)

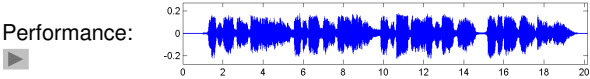
Performance Analysis



Performance Analysis



Performance Analysis

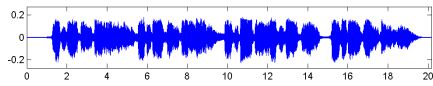


Strategy: Compute score-audio synchronization and derive tempo curve

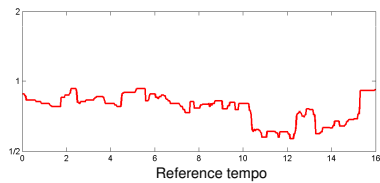


Performance Analysis

Performance:



Tempo curve:

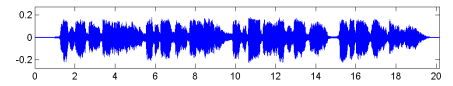


Score (reference):



Performance Analysis

Performance:



What can be done if no reference is available?

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

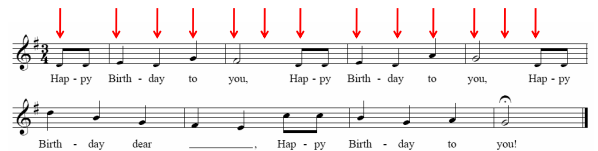
Tempo Estimation

Measure



Tempo Estimation

Tactus (beat)



Tempo Estimation

Tatum (temporal atom)

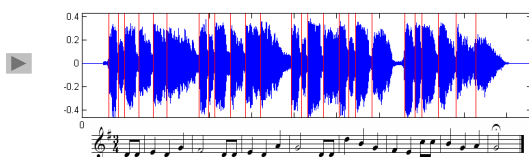


Tempo Estimation

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

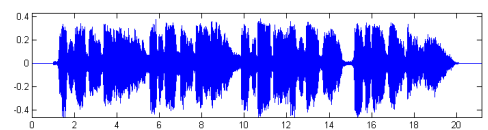
Tempo Estimation

Performance



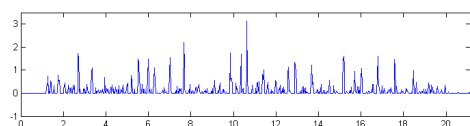
Tempo Estimation

Performance



Tempo Estimation

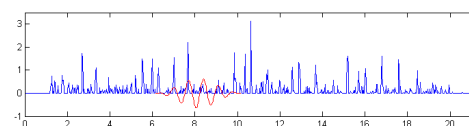
Novelty Curve



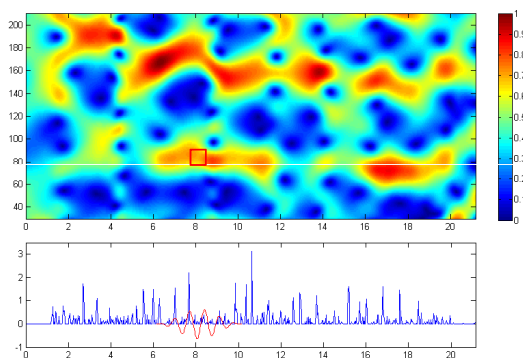
Tempo Estimation

Novelty Curve

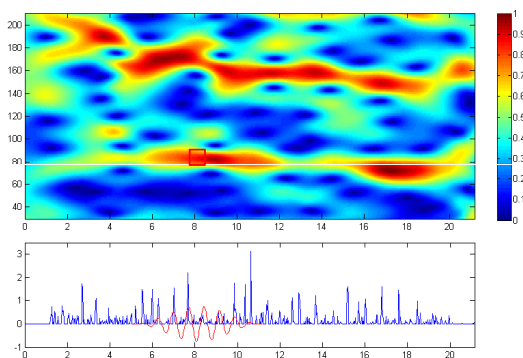
Periodicity Analysis



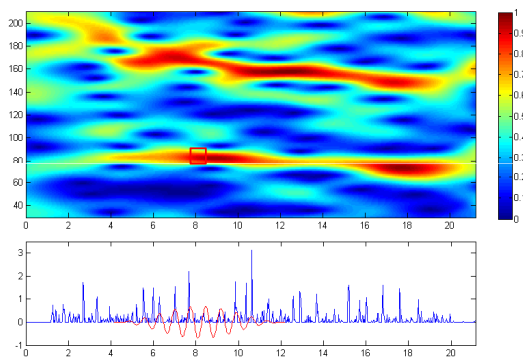
Tempo Estimation: Tempogram



Tempo Estimation: Tempogram



Tempo Estimation: Tempogram



Motivic Similarity

Var. 4: Vivace

A musical score for a piece titled "Var. 4: Vivace". The score is written for four staves, likely representing different instruments or voices. The music is in 2/4 time and features various rhythmic patterns, including eighth and sixteenth notes, and rests. The score is marked with dynamics such as *f* (forte) and *p* (piano). A play button icon is located below the score.

Motivic Similarity



Beethoven's Fifth (1st Mov.)



Motivic Similarity



Beethoven's Fifth (1st Mov.)



Beethoven's Fifth (3rd Mov.)



Motivic Similarity



Beethoven's Fifth (1st Mov.)



Beethoven's Fifth (3rd Mov.)



Beethoven's Appassionata



Multimodal Computing and Interaction

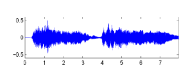
Sheet Music (Image)



MIDI



CD / MP3 (Audio)



Music

Multimodal Computing and Interaction

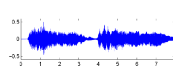
Sheet Music (Image)



MIDI



CD / MP3 (Audio)



MusicXML (Text)

```
<musicxml>
<staff>
<note>
<pitch>
<name>C4
<octave>4
</pitch>
<duration>4
</duration>
</note>
</staff>
</musicxml>
```

Music

Singing / Voice (Audio)



Music Literature (Text)



Music Film (Video)



Dance / Motion (Mocap)

