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

Player Piano (1900)
Piano Roll Representation (MIDI)

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

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

- 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: MIDI-Audio

MIDI = meta data

Automated annotation
Audio recording

Sonification of annotations

Music Synchronization: Scan-Audio

Scanned Sheet Music
Symbolic Note Events

OMR

Correspondence
Audio Recording

Music Synchronization: Scan-Audio

Scanned Sheet Music
Symbolic Note Events

OMR

Correspondence
Audio Recording
Music Synchronization: Scan-Audio

High Quality

Symbolic Note Events

„Dirty“ but hidden

High Quality

Audio Recording

System: SyncPlayer/SheetMusic

Music Synchronization: Lyrics-Audio

Difficult!

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

Music Processing

<table>
<thead>
<tr>
<th>Coarse Level</th>
<th>Fine Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>What do different versions have in common?</td>
<td>What are the characteristics of a specific version?</td>
</tr>
</tbody>
</table>

Music Processing

<table>
<thead>
<tr>
<th>Coarse Level</th>
<th>Fine Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>What do different versions have in common?</td>
<td>What are the characteristics of a specific version?</td>
</tr>
<tr>
<td>What makes up a piece of music?</td>
<td>What makes music come alive?</td>
</tr>
</tbody>
</table>
Music Processing

<table>
<thead>
<tr>
<th>Coarse Level</th>
<th>Fine Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>What do different versions have in common?</td>
<td>What are the characteristics of a specific version?</td>
</tr>
<tr>
<td>What makes up a piece of music?</td>
<td>What makes music come alive?</td>
</tr>
<tr>
<td>Identify despite of differences</td>
<td>Identify the differences</td>
</tr>
</tbody>
</table>

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")

Strategy: Compute score-audio synchronization and derive tempo curve
Performance Analysis

Performance:

Tempo curve:

Score (reference):

What can be done if no reference is available?

Music Processing

<table>
<thead>
<tr>
<th>Relative</th>
<th>Absolute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given: Several versions</td>
<td>Given: One version</td>
</tr>
</tbody>
</table>

Comparison of extracted parameters

Extraction errors have often no consequence on final result

Extraction errors immediately become evident

Example tasks:
- Music Synchronization
- Genre Classification
- Tempo Estimation

Music Processing

<table>
<thead>
<tr>
<th>Relative</th>
<th>Absolute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given: Several versions</td>
<td>Given: One version</td>
</tr>
</tbody>
</table>

Comparison of extracted parameters

Extraction errors have often no consequence on final result

Extraction errors immediately become evident

Example tasks:
- Music Synchronization
- Genre Classification
- Music Transcription
- Tempo Estimation
Tempo Estimation

Measure

<table>
<thead>
<tr>
<th>Measure</th>
<th>Measure</th>
</tr>
</thead>
</table>
| \( \begin{array}{c}
    \text{Tempo Estimation} \\
    \text{Tactus (beat)} \\
    \text{Local tempo deviations} \\
    \text{Sparse information} \\
    \text{(e., only note onsets available)} \\
    \text{Vague information} \\
    \text{(e., extracted note onsets corrupt)}
\end{array} \) |
| \( \begin{array}{c}
    \text{Tempo Estimation} \\
    \text{Tatum (temporal atom)} \\
    \text{Which temporal level?} \\
    \text{Local tempo deviations} \\
    \text{Sparse information} \\
    \text{(e., only note onsets available)} \\
    \text{Vague information} \\
    \text{(e., extracted note onsets corrupt)}
\end{array} \) |
| \( \begin{array}{c}
    \text{Tempo Estimation} \\
    \text{Performance}
\end{array} \) |