

Lecture

Information Retrieval

for Music and Motion

Meinard Müller

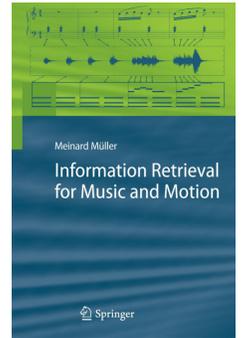
Summer Term 2008



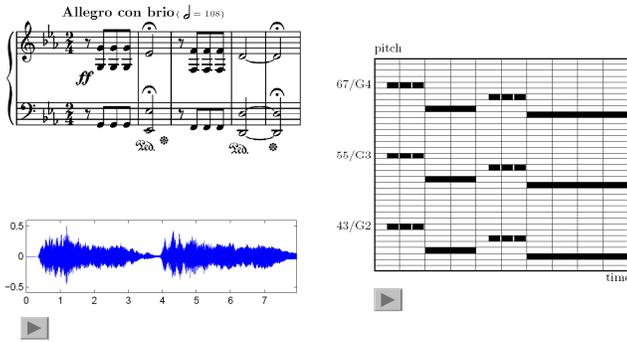
Organization

- Thu. 16-18
- MPI, Room 024
- Website
- Layout
- Examinations

Textbook



Music Data



Music Data

Various interpretations – Beethoven's Fifth

Bernstein	
Karajan	
Scherbakov (piano)	
MIDI (piano)	

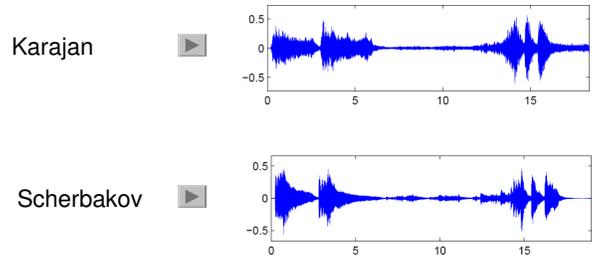
General Goals

- Automated organization of complex and inhomogeneous music collections
- Generation of annotations and cross-links
- Tools and methods for multimodal search, navigation and interaction

Music Information Retrieval (MIR)

Music Synchronization: Audio-Audio

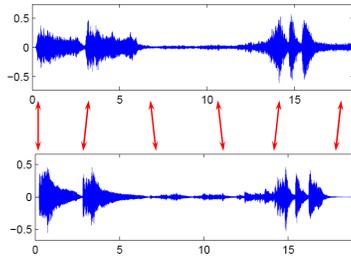
Beethoven's Fifth



Music Synchronization: Audio-Audio

Beethoven's Fifth

Karajan



Scherbakov



Synchronization: Karajan → Scherbakov

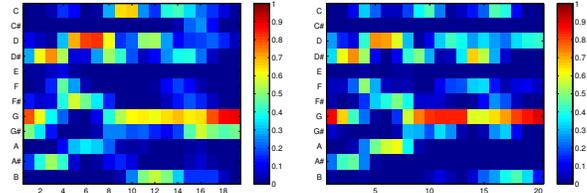
Music Synchronization: Audio-Audio

Feature extraction: chroma features

Karajan

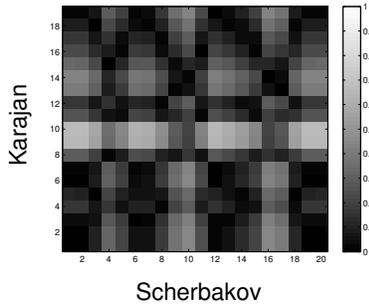


Scherbakov



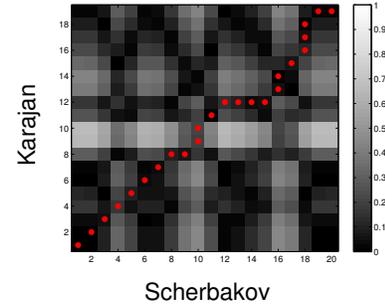
Music Synchronization: Audio-Audio

Cost matrix

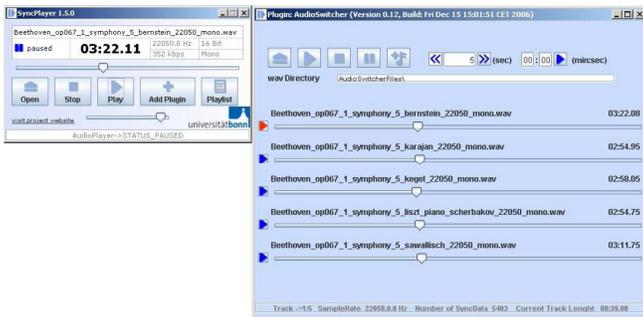


Music Synchronization: Audio-Audio

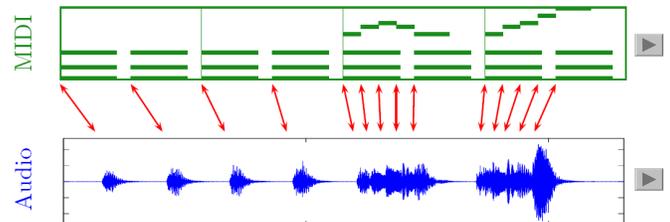
Cost-minimizing warping path



System: SyncPlayer/AudioSwitcher



Music Synchronization: MIDI-Audio



Music Synchronization: MIDI-Audio

MIDI = metadata

Automated annotation

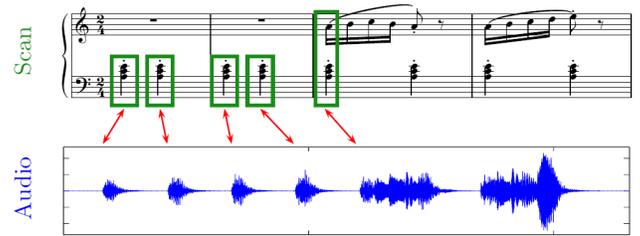
Audio recording

Sonification of annotations



13

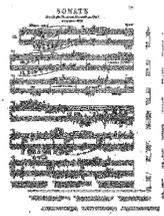
Music Synchronization: Scan-Audio



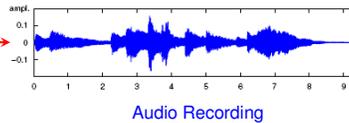
14

Music Synchronization: Scan-Audio

Scanned Sheet Music



Correspondence

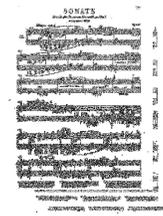


Audio Recording

15

Music Synchronization: Scan-Audio

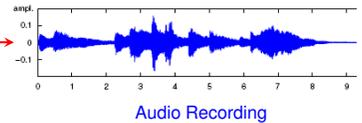
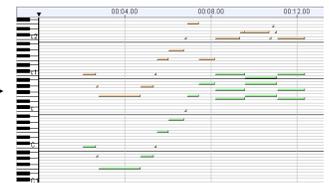
Scanned Sheet Music



Correspondence

Symbolic Note Events

OMR

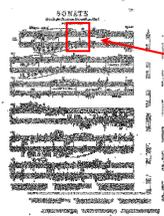


Audio Recording

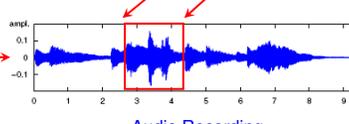
16

Music Synchronization: Scan-Audio

Scanned Sheet Music



Correspondence

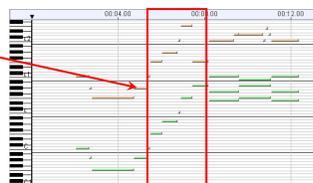


Audio Recording

17

Symbolic Note Events

OMR



Music Synchronization: Scan-Audio

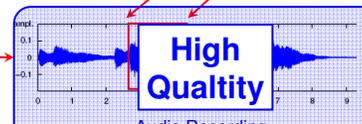
Scanned Sheet Music



Correspondence

Symbolic Note Events

OMR



Audio Recording

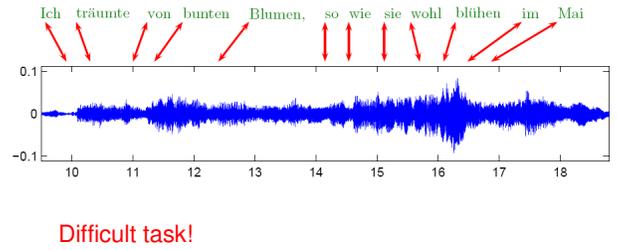
18

System: SyncPlayer/SheetMusic



19

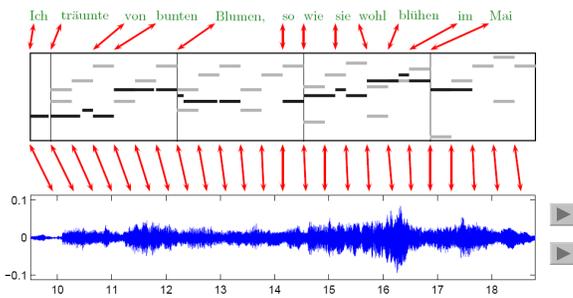
Music Synchronization: Lyrics-Audio



20

Music Synchronization: Lyrics-Audio

Lyrics-Audio → Lyrics-MIDI + MIDI-Audio



21

Music Synchronization

- Turetsky/Ellis (ISMIR 2003)
- Soulez/Rodet/Schwarz (ISMIR 2003)
- Arifi/Clausen/Kurth/Müller (ISMIR 2003)
- Hu/Dannenberg/Tzanetakis (WASPAA 2003)
- Müller/Kurth/Röder (ISMIR 2004)
- Raphael (ISMIR 2004)
- Dixon/Widmer (ISMIR 2005)
- Müller/Mattes/Kurth (ISMIR 2006)
- Dannenberg/Raphael (Special Issue ACM 2006)
- Fujihara/Goto (ICASSP 2008)
- Wang/Iskandar/New/Shenoy (IEEE T-ASLP 2008)

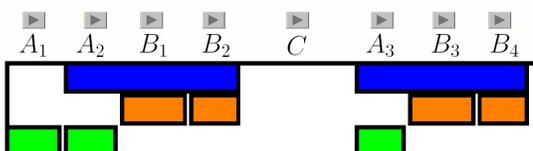
22

Audio Structure Analysis

Given: CD recording

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

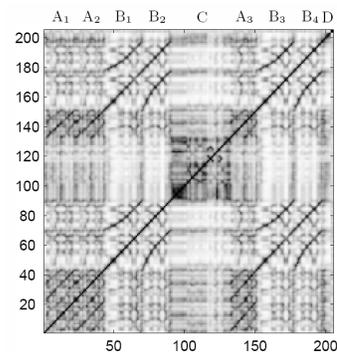
Example: Brahms Hungarian Dance No. 5 (Ormandy)



23

Audio Structure Analysis

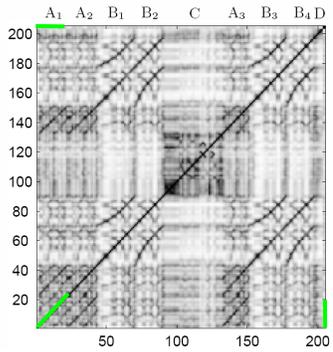
Self-similarity matrix



24

Audio Structure Analysis

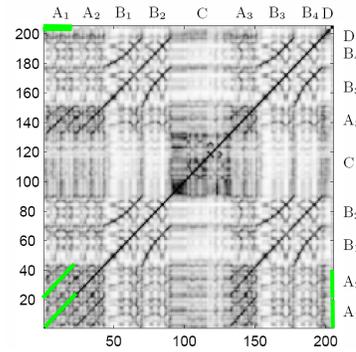
Self-similarity matrix



25

Audio Structure Analysis

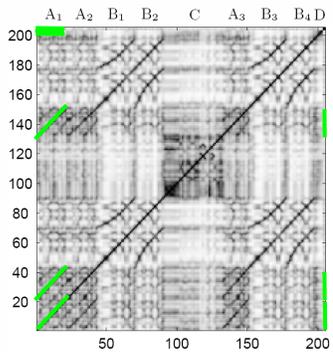
Self-similarity matrix



26

Audio Structure Analysis

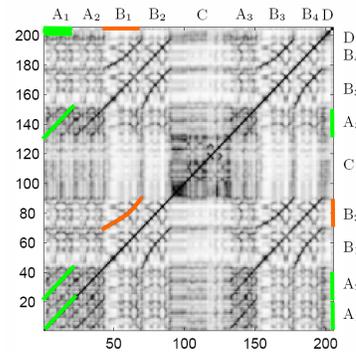
Self-similarity matrix



27

Audio Structure Analysis

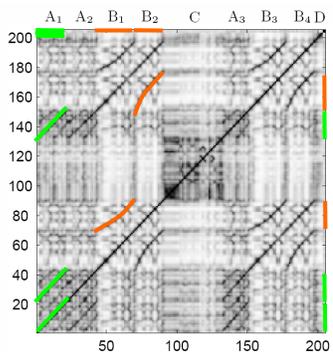
Self-similarity matrix



28

Audio Structure Analysis

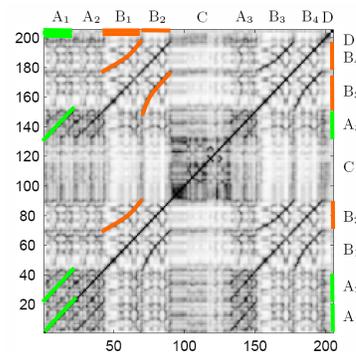
Self-similarity matrix



29

Audio Structure Analysis

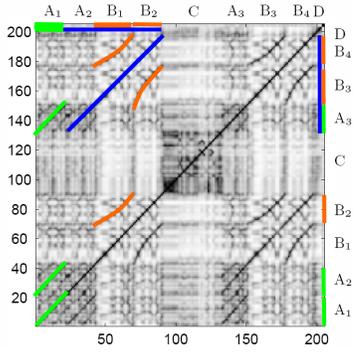
Self-similarity matrix



30

Audio Structure Analysis

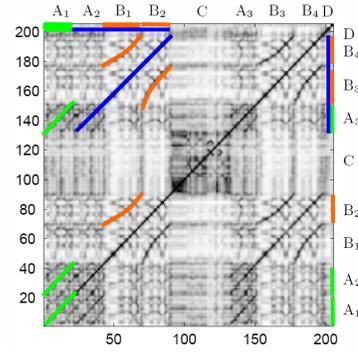
Self-similarity matrix



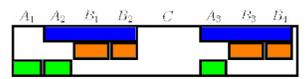
31

Audio Structure Analysis

Self-similarity matrix



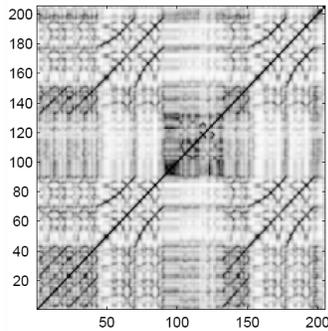
Similarity cluster



32

Audio Structure Analysis

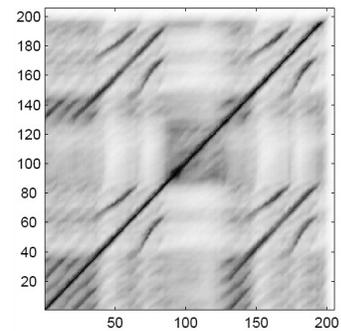
Path extraction



33

Audio Structure Analysis

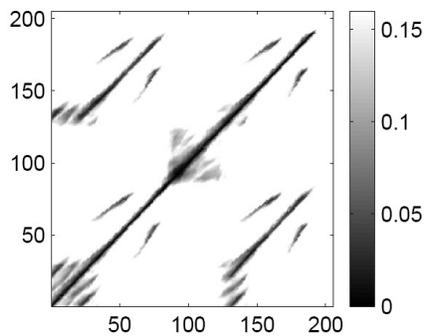
Path extraction



34

Audio Structure Analysis

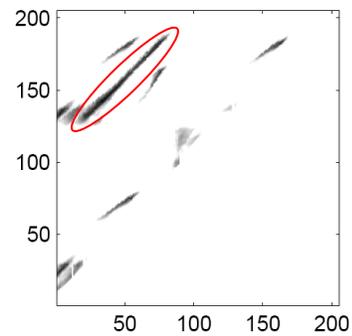
Path extraction



35

Audio Structure Analysis

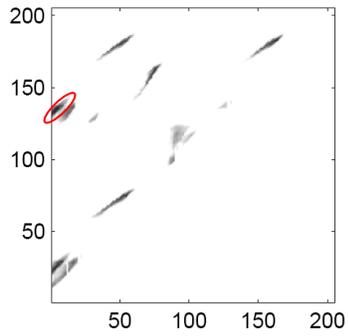
Path extraction



36

Audio Structure Analysis

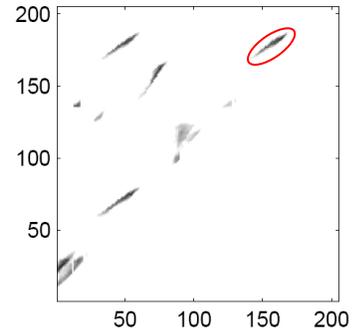
Path extraction



37

Audio Structure Analysis

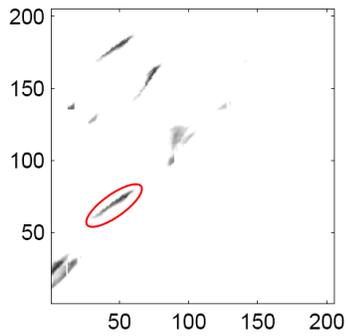
Path extraction



38

Audio Structure Analysis

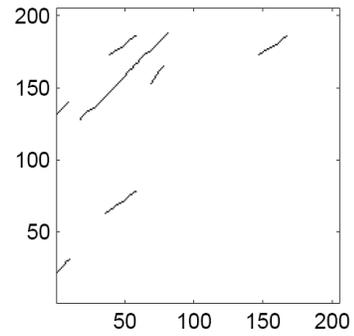
Path extraction



39

Audio Structure Analysis

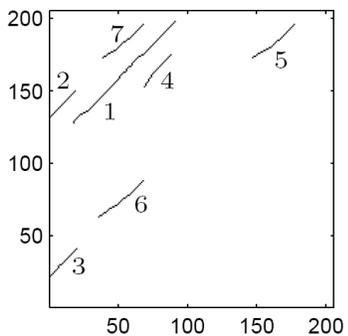
Path extraction



40

Audio Structure Analysis

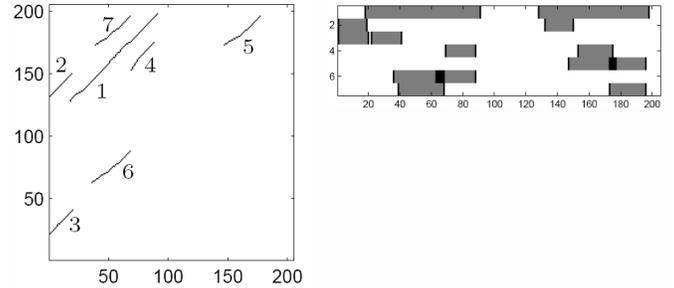
Path extraction



41

Audio Structure Analysis

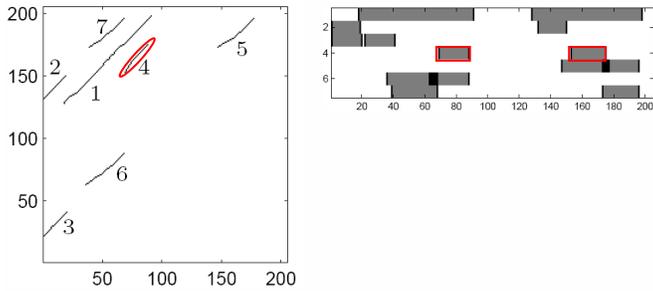
Global structure



42

Audio Structure Analysis

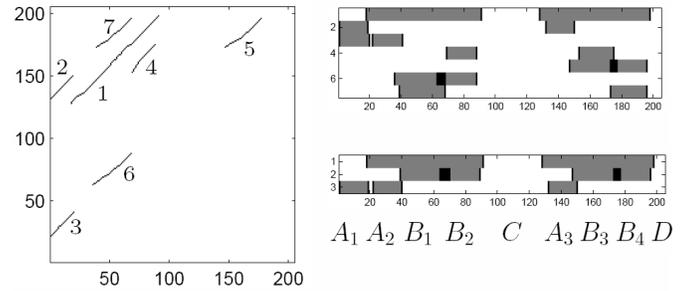
Global structure



43

Audio Structure Analysis

Global structure



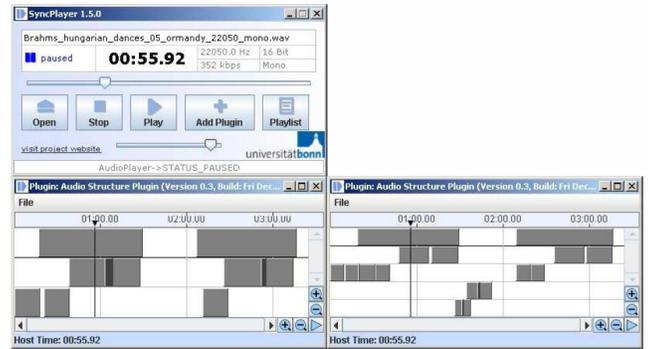
44

Audio Structure Analysis

- Dannenberg/Hu (ISMIR 2002)
- Peeters/Burthe/Rodet (ISMIR 2002)
- Cooper/Foote (ISMIR 2002)
- Goto (ICASSP 2003)
- Chai/Vercoe (ACM Multimedia 2003)
- Lu/Wang/Zhang (ACM Multimedia 2004)
- Bartsch/Wakefield (IEEE Trans. Multimedia 2005)
- Goto (IEEE Trans. Audio 2006)
- Müller/Kurth (EURASIP 2007)

45

System: SyncPlayer/AudioStructure



46

Music Information Retrieval

- Music Synchronization
- Audio Structure Analysis
- Audio Matching
- ...

47

Multimodal Computing and Interaction

Music

48

Multimodal Computing and Interaction

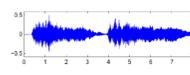
Sheet Music (Image)



MIDI



CD / MP3 (Audio)



Music

49

Multimodal Computing and Interaction

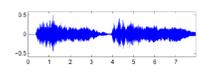
Sheet Music (Image)



MIDI



CD / MP3 (Audio)



MusicXML (Text)

```
<note>
  <pitch>
    <step>B</step>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>quarter</type>
</note>
```

Music

Singing / Voice (Audio)



Music Literature (Text)



50

Multimodal Computing and Interaction

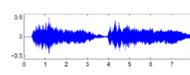
Sheet Music (Image)



MIDI



CD / MP3 (Audio)



MusicXML (Text)

```
<note>
  <pitch>
    <step>B</step>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>quarter</type>
</note>
```

Music

Singing / Voice (Audio)



Music Literature (Text)



Music Film (Video)



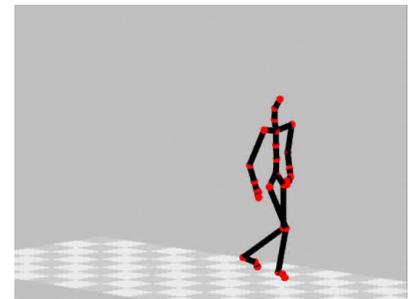
Dance / Motion (Mocap)



51

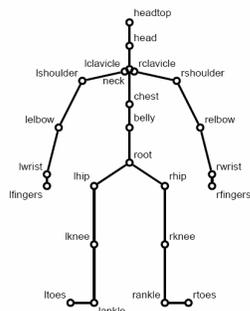
Motion Capture Data

- Digital 3D representations of motions
- Computer animation
- Sports
- Gait analysis



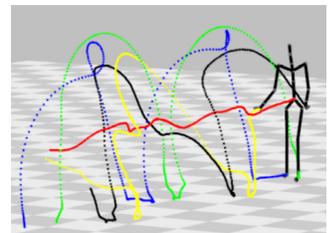
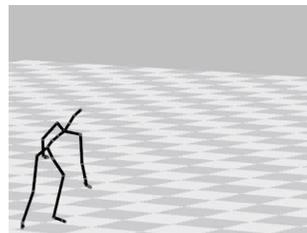
52

Motion Capture Data



53

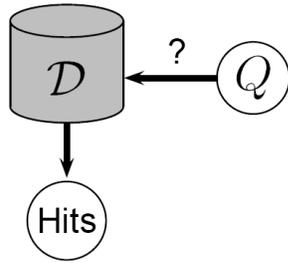
Motion Capture Data



54

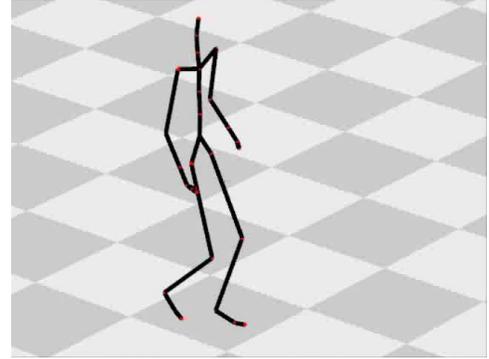
Motion Retrieval

- \mathcal{D} = MoCap database
- Q = query motion clip
- **Goal:** find all motion clips in \mathcal{D} similar to Q



55

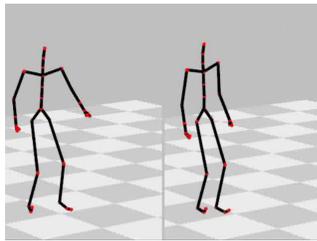
Motion Retrieval



56

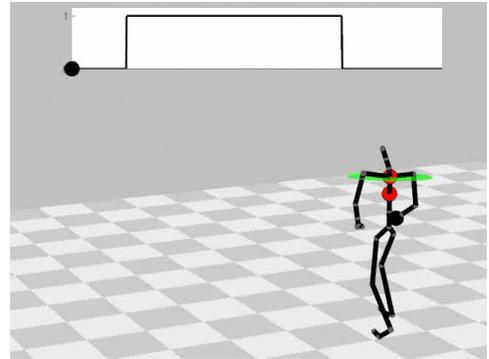
Notion of Similarity

- **Numerical** similarity vs. **logical** similarity
- Logically related motions may exhibit significant **spatio-temporal** variations



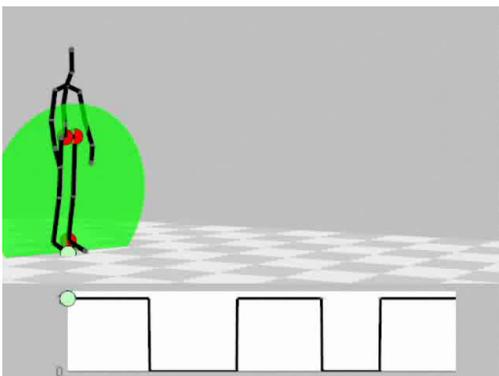
57

Relational Features



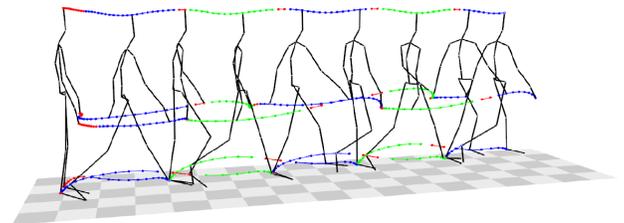
58

Relational Features



59

Relational Features

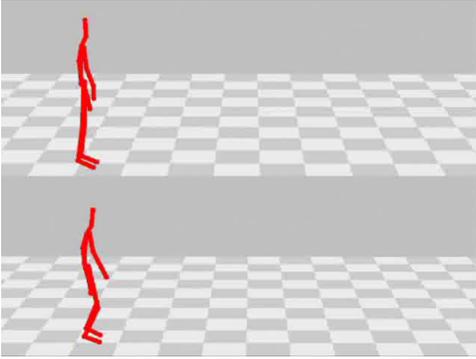


Induced feature sequence:

$$\left(\binom{1}{1}, \binom{0}{1}, \binom{1}{1}, \binom{1}{0}, \binom{1}{1}, \binom{0}{1}, \binom{1}{1}, \binom{1}{0}, \binom{1}{1}, \binom{0}{1} \right)$$

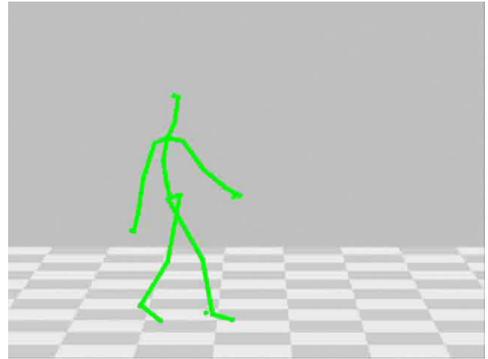
60

Relational Features



61

Motion Retrieval



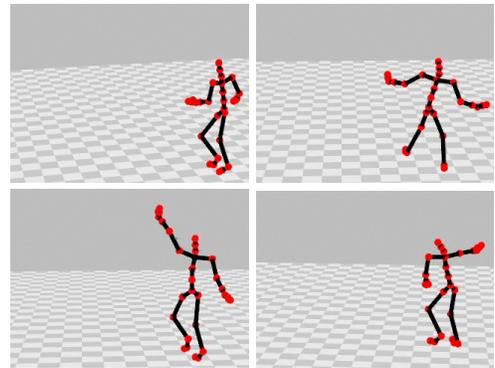
62

Relational Features

- Spatio-temporal invariance
- Indexing
- Efficient retrieval & preselection
- Problem: feature design & selection

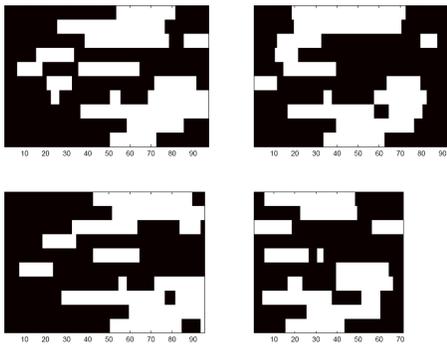
63

Motion Templates



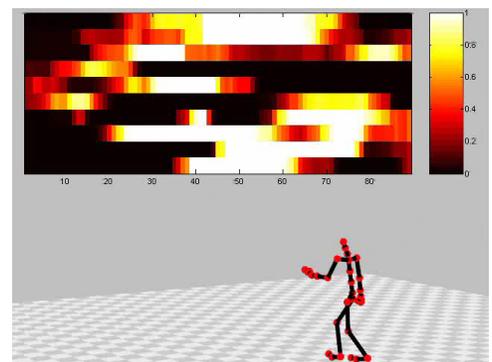
64

Motion Templates



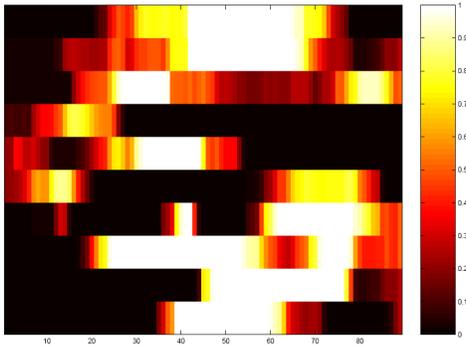
65

Motion Templates



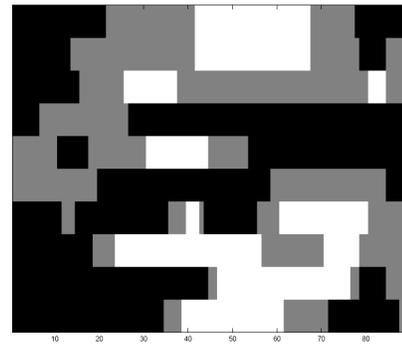
66

Motion Templates



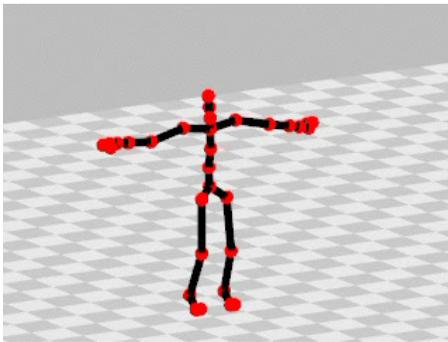
67

Motion Templates



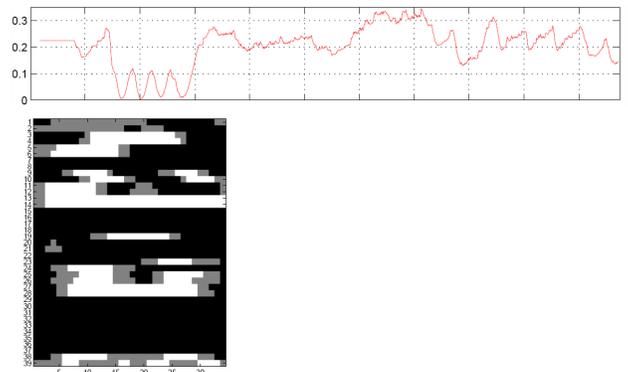
68

MT-based Motion Retrieval



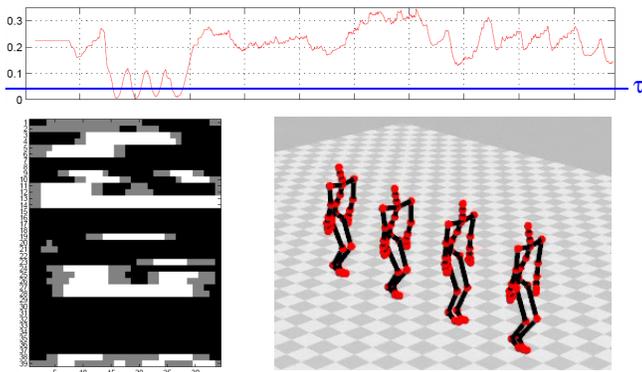
69

MT-based Motion Retrieval: Jumping Jack



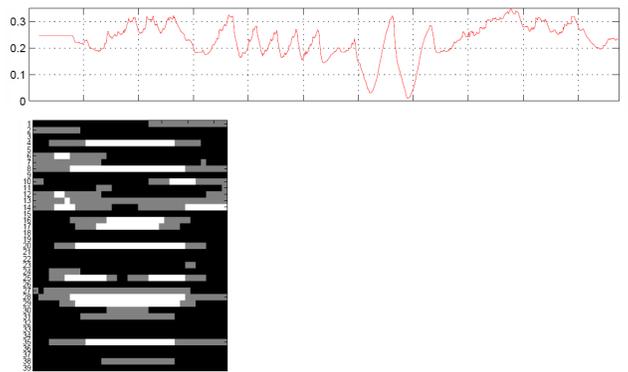
70

MT-based Motion Retrieval: Jumping Jack



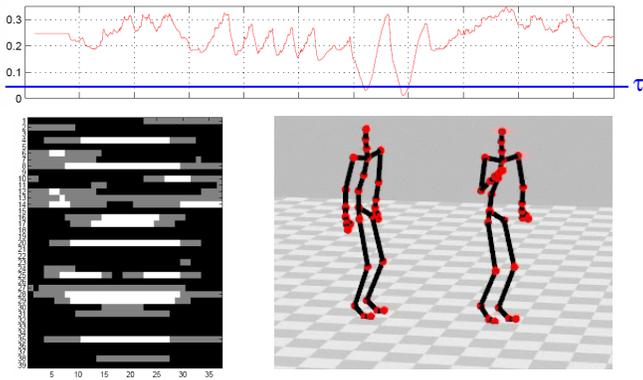
71

MT-based Motion Retrieval: Elbow-To-Knee



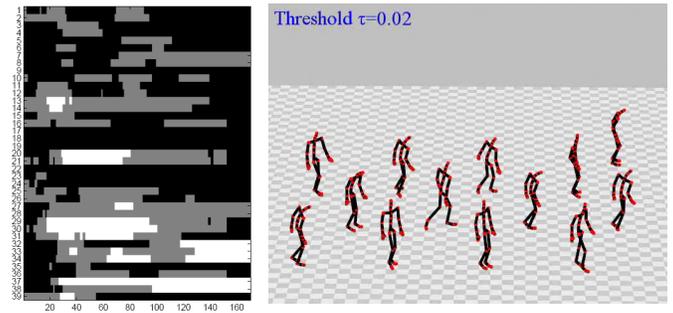
72

MT-based Motion Retrieval: Elbow-To-Knee



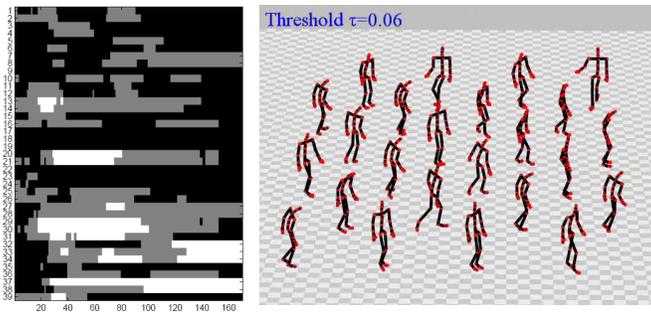
73

MT-based Motion Retrieval



74

MT-based Motion Retrieval



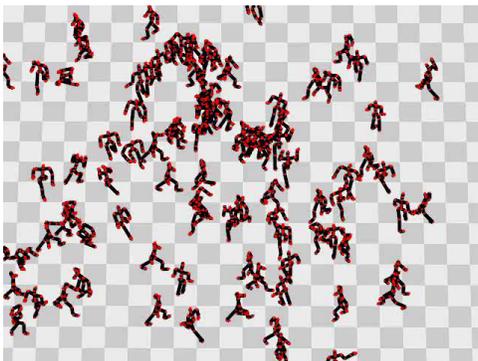
75

Conclusions

- Automated data organization
- Exploiting multimodality
- Synchronization
- Handling object deformations
- Indexing and efficiency

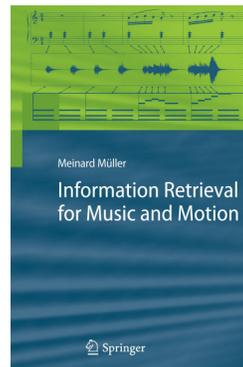
76

Conclusions



77

Book



Müller, Meinard
Information Retrieval for Music and Motion
2007, XVI, 318 pages
136 illus. 39 in Color, Hardcover
ISBN: 978-3-540-74047-6
www.springer.com/978-3-540-74047-6/
69,50 EUR

78