Segmentation and Retrieval Techniques for Music and Motion Data

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

Digital Sports Colloquium, FAU
05.10.2016

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

Meinard Müller

- 2007 Habilitation
  Bonn University

- 2007 – 2012
  Senior Researcher
  Saarland University & MPI Informatik

- Since 2012
  Professor: Semantic Audio Processing
  Erlangen-Nürnberg University

Group Members

- Frank Zalkow
- Thomas Prätzlich
- Stefan Balke
- Christian Dittmar
- Patricio López-Serrano
- Christof Weiß

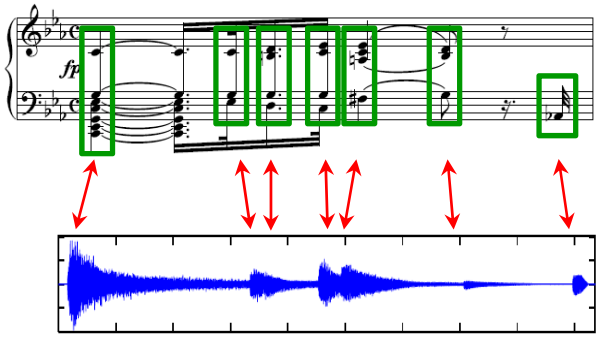




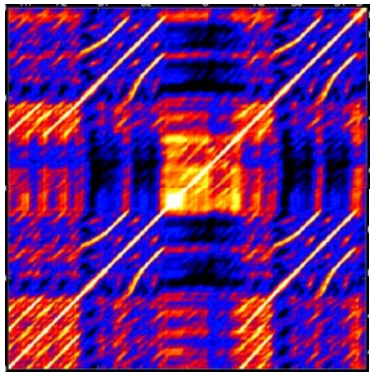


Music Processing

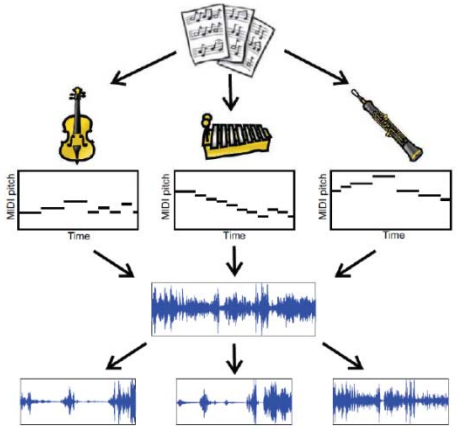
Music Synchronization



Structure Analysis

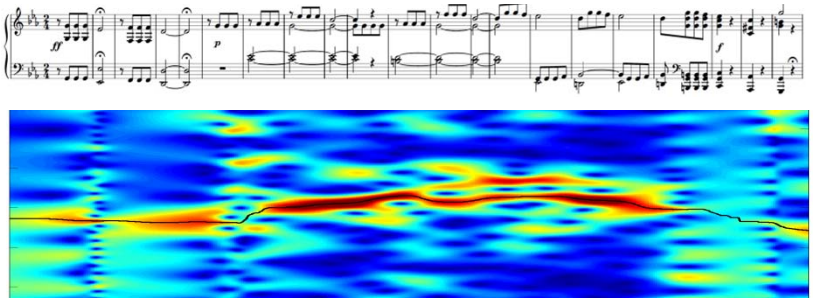


Score-Informed Source Separation

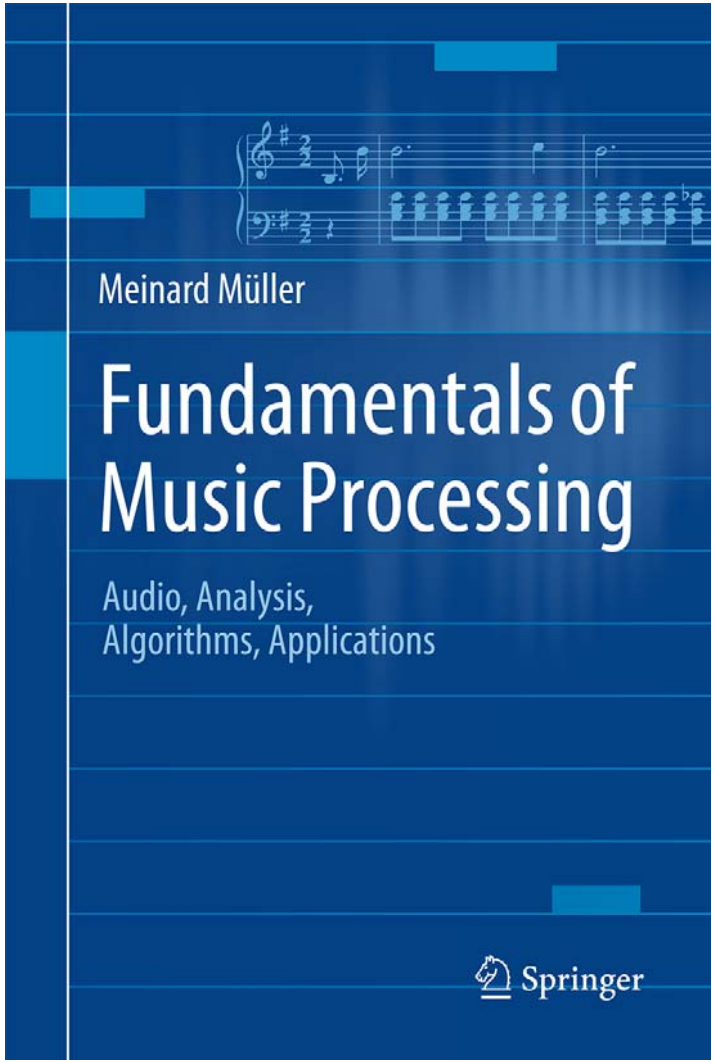


User Interfaces

Tempo Estimation and Beat Tracking



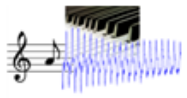
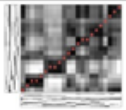
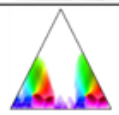
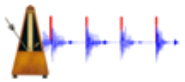
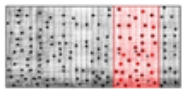
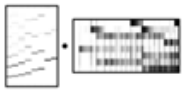
Book: Fundamentals of Music Processing



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

Accompanying website:
www.music-processing.de

Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Represenations
2		Fourier Analysis of Signals
3		Music Synchronization
4		Music Structure Analysis
5		Chord Recognition
6		Tempo and Beat Tracking
7		Content-Based Audio Retrieval
8		Musically Informed Audio Decomposition

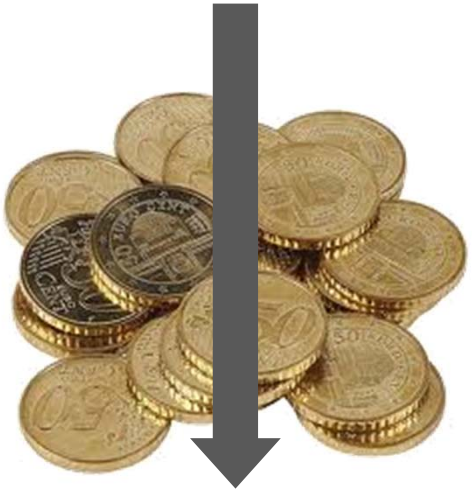
Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

International Audio Laboratories Erlangen

Fraunhofer IIS

mp3



FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

AUDIO LABS

AudioLabs – FAU

- Prof. Dr. Jürgen Herre
  Audio Coding
- Prof. Dr. Bernd Edler
  Audio Signal Analysis
- Prof. Dr. Meinard Müller
  Semantic Audio Processing
- Prof. Dr. Emanuël Habets
  Spatial Audio Signal Processing
- Prof. Dr. Tom Bäckström
  Speech Processing
- Dr. Stefan Turowski
  Coordinator AudioLabs-FAU





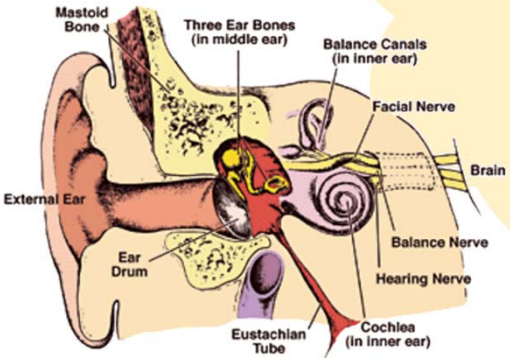

International Audio Laboratories Erlangen

International Audio Laboratories Erlangen

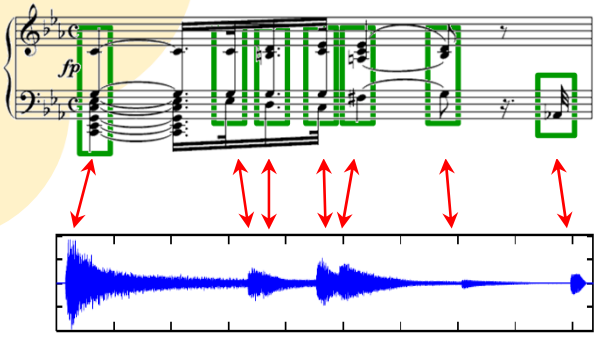
Audio Coding

3D Audio

Audio



Psychoacoustics

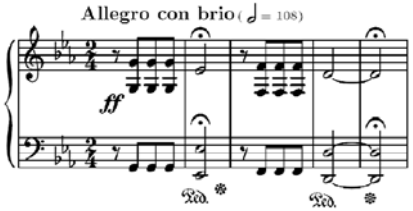


Music Processing

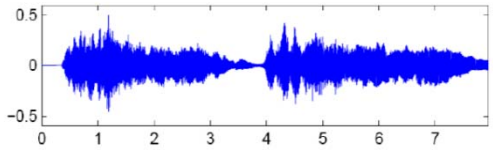
Music

Music Processing

Sheet Music (Image)



CD / MP3 (Audio)



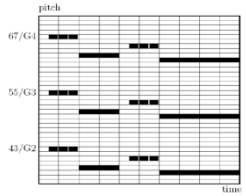
MusicXML (Text)

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

Dance / Motion (Mocap)

Music

MIDI



Singing / Voice (Audio)

Music Film (Video)

Music Literature (Text)

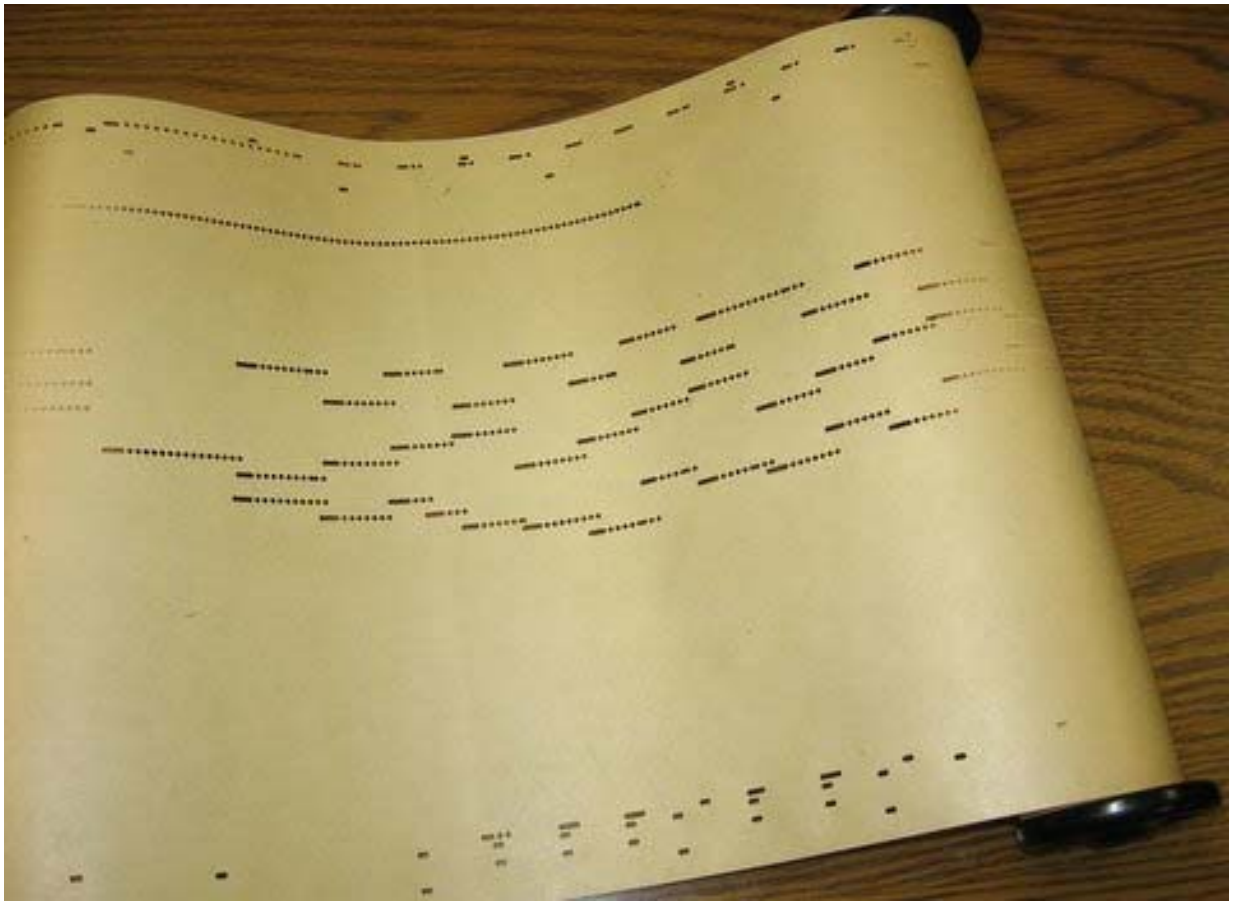
Research Goals

- 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

Overview

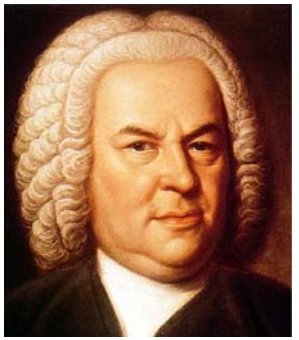
- Introduction
- Music Retrieval
- Motion Retrieval
- Music Structure Analysis
- Beat Tracking

Piano Roll Representation



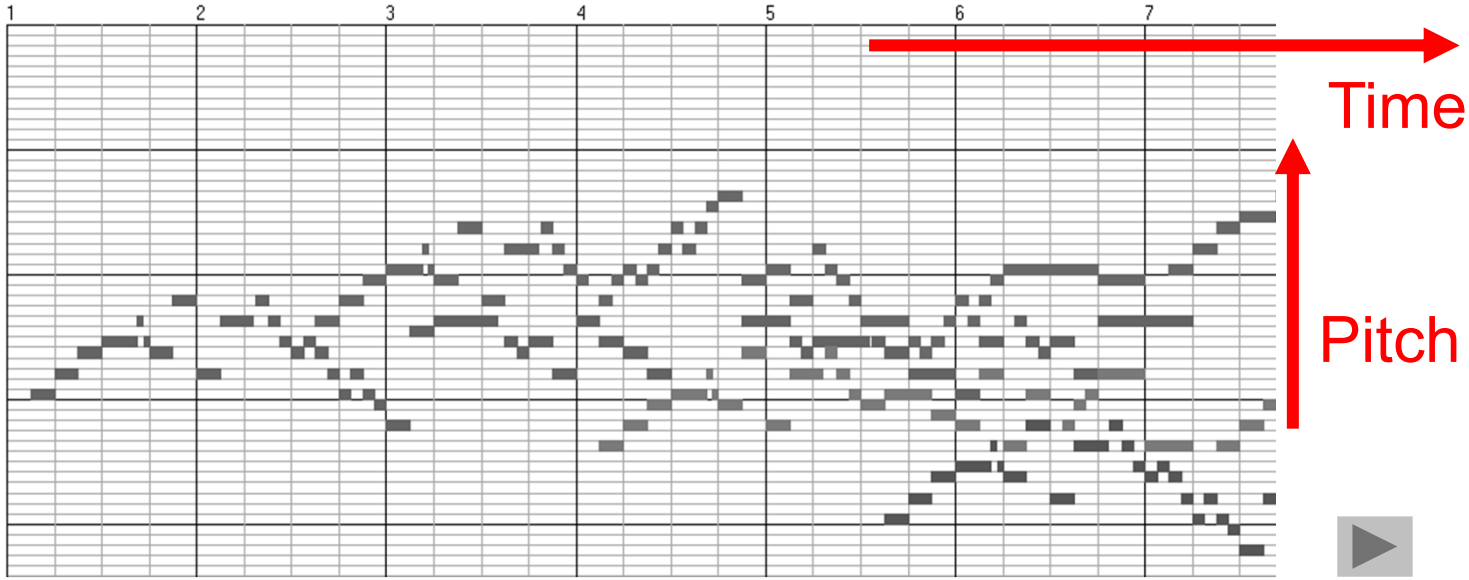
Player Piano (1900)

Piano Roll Representation (MIDI)

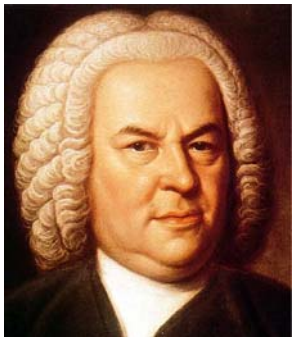


J.S. Bach, C-Major Fuge

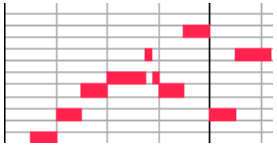
(Well Tempered Piano, BWV 846)



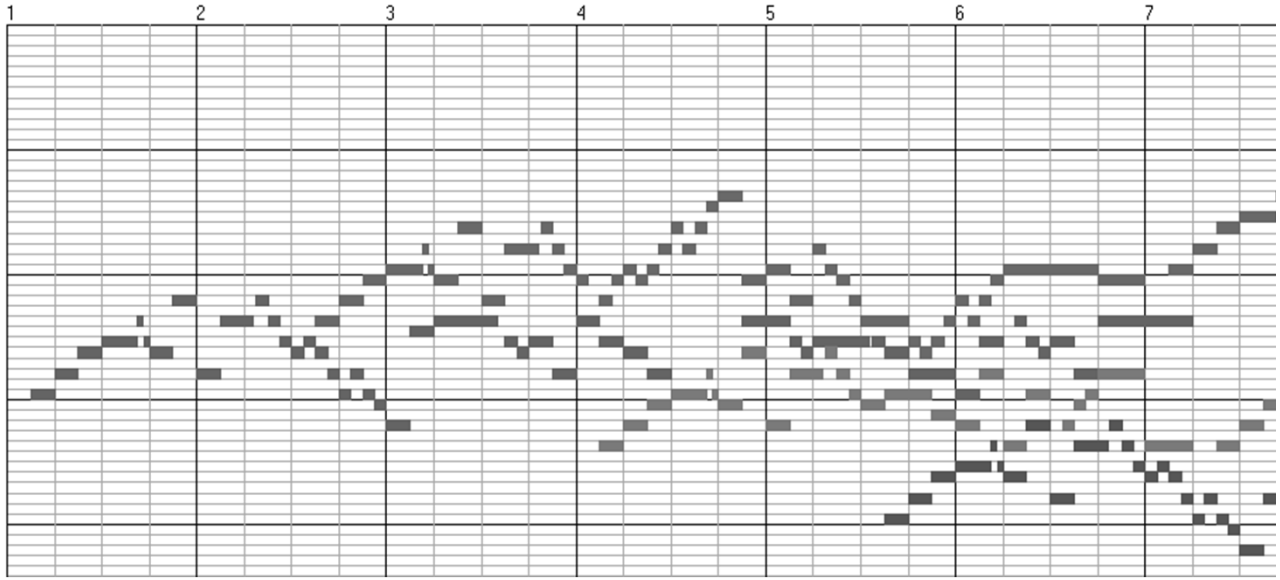
Piano Roll Representation (MIDI)



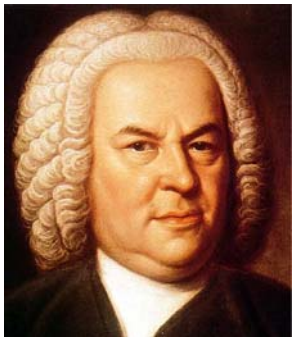
Query:



Goal: Find all occurrences of the query



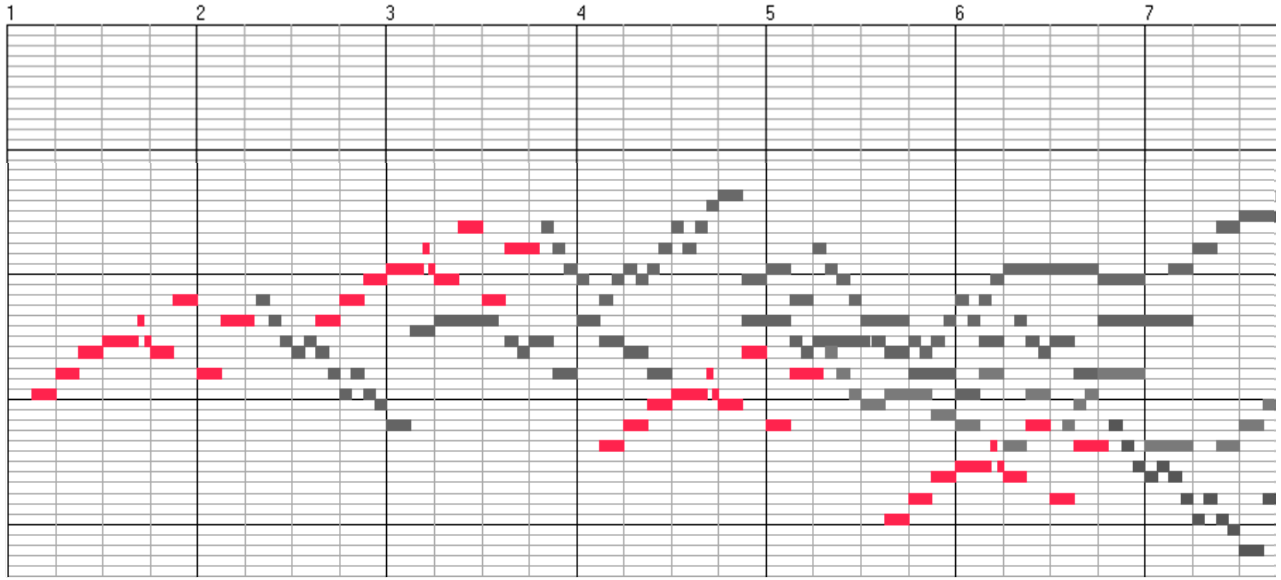
Piano Roll Representation (MIDI)



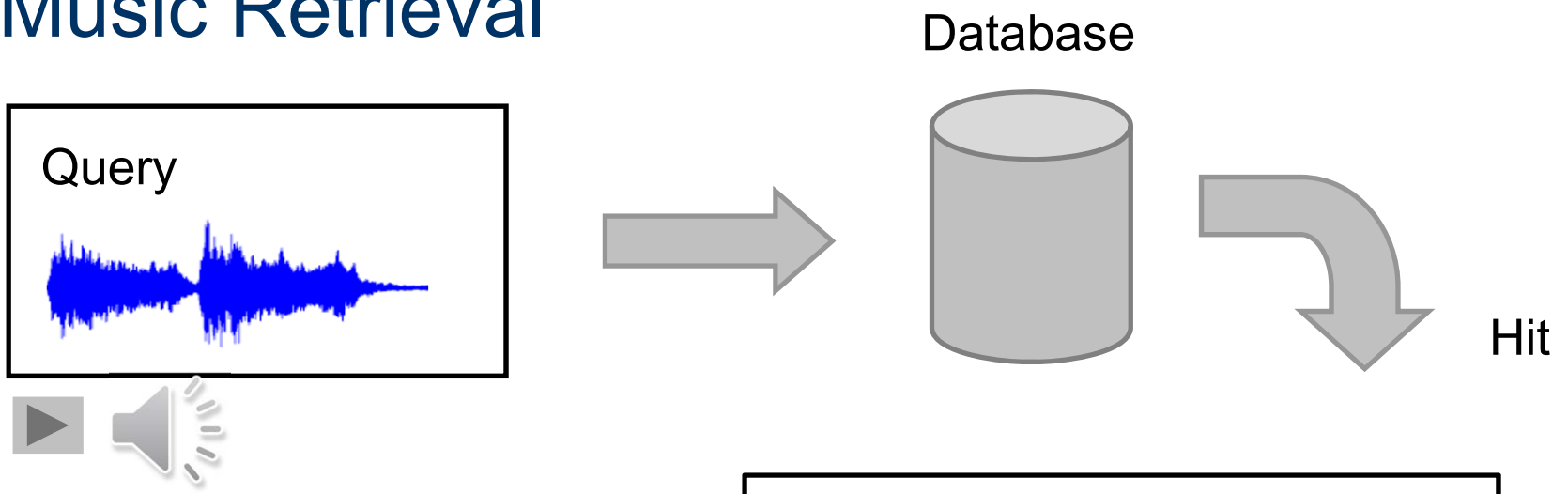
Query:

Goal: Find all occurrences of the query

Matches:



Music Retrieval



Audio-ID

Bernstein (1962)
Beethoven, Symphony No. 5

Version-ID

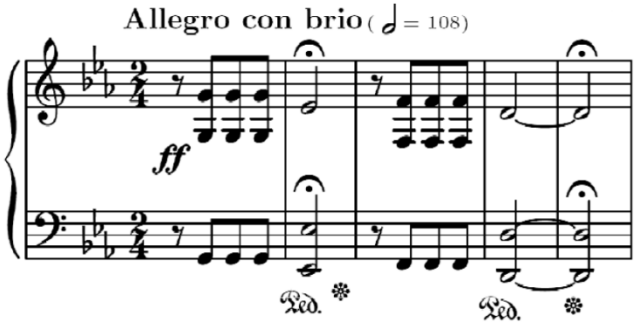
Beethoven, Symphony No. 5:
- Bernstein (1962)
- Karajan (1982)
- Gould (1992)

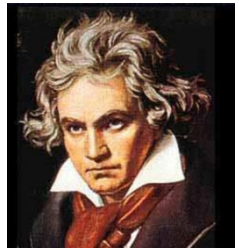
Kategorie-ID

- Beethoven, Symphony No. 9
- Beethoven, Symphony No. 3
- Haydn Symphony No. 94

Music Synchronization: Audio-Audio

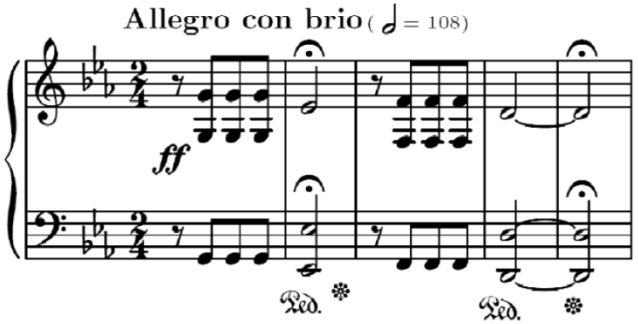
Beethoven’s Fifth

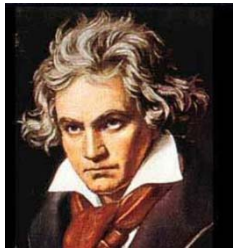




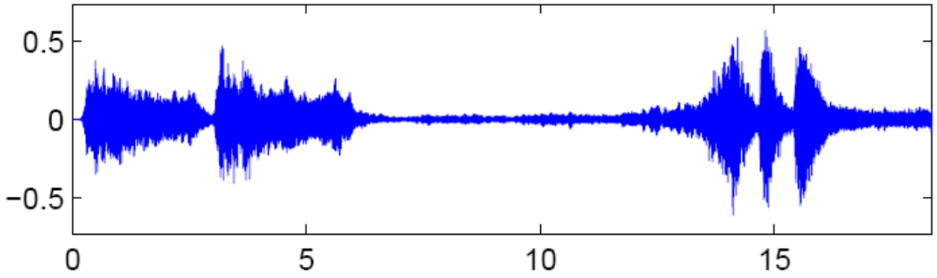
Music Synchronization: Audio-Audio

Beethoven's Fifth

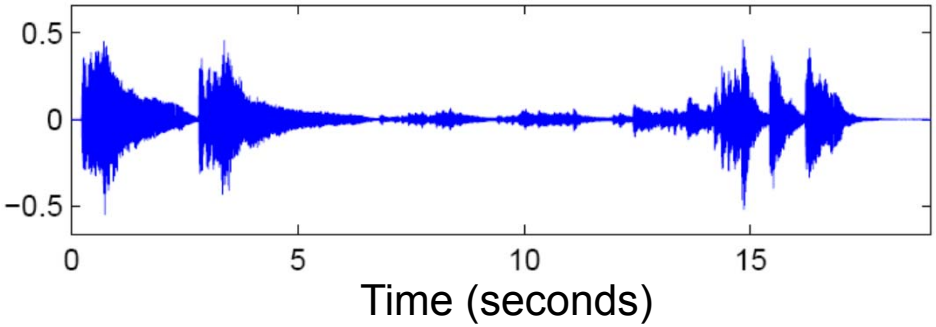




Orchester (Karajan)

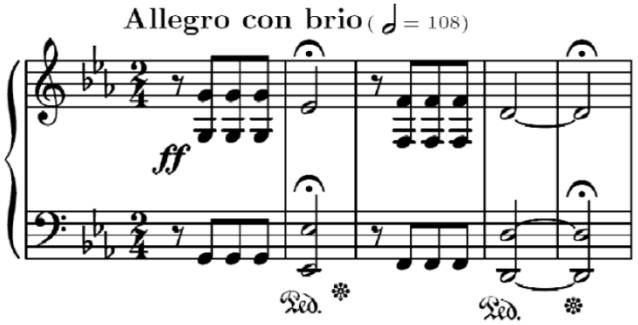


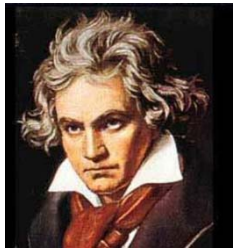
Piano (Scherbakov)



Music Synchronization: Audio-Audio

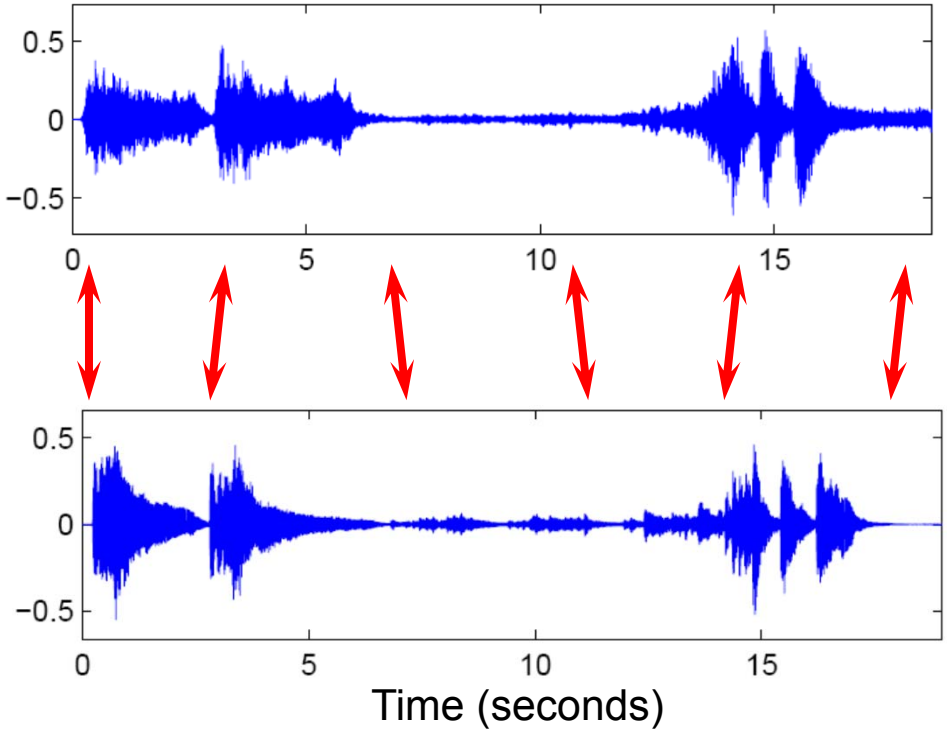
Beethoven's Fifth



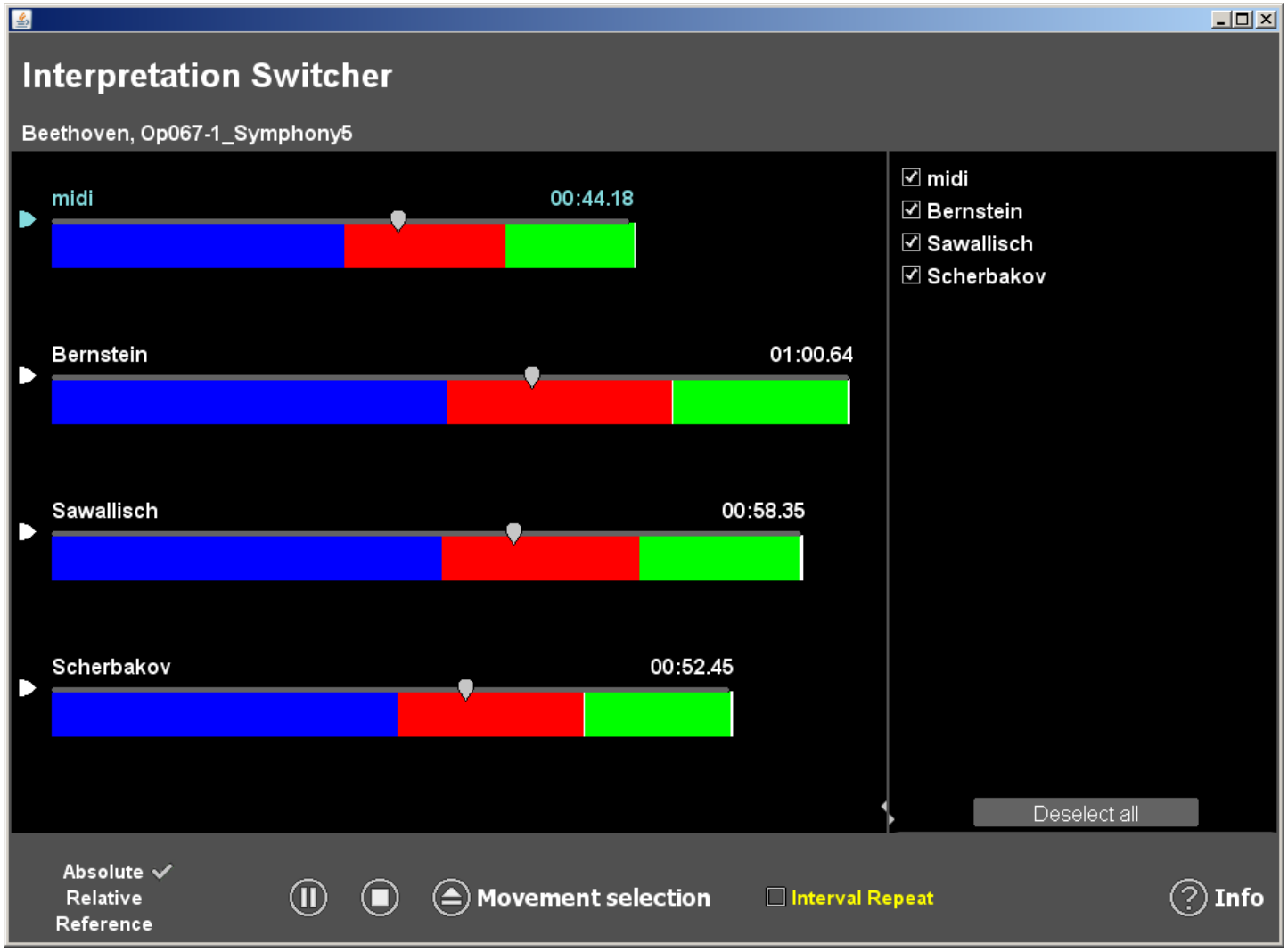


Orchester (Karajan)

Piano (Scherbakov)

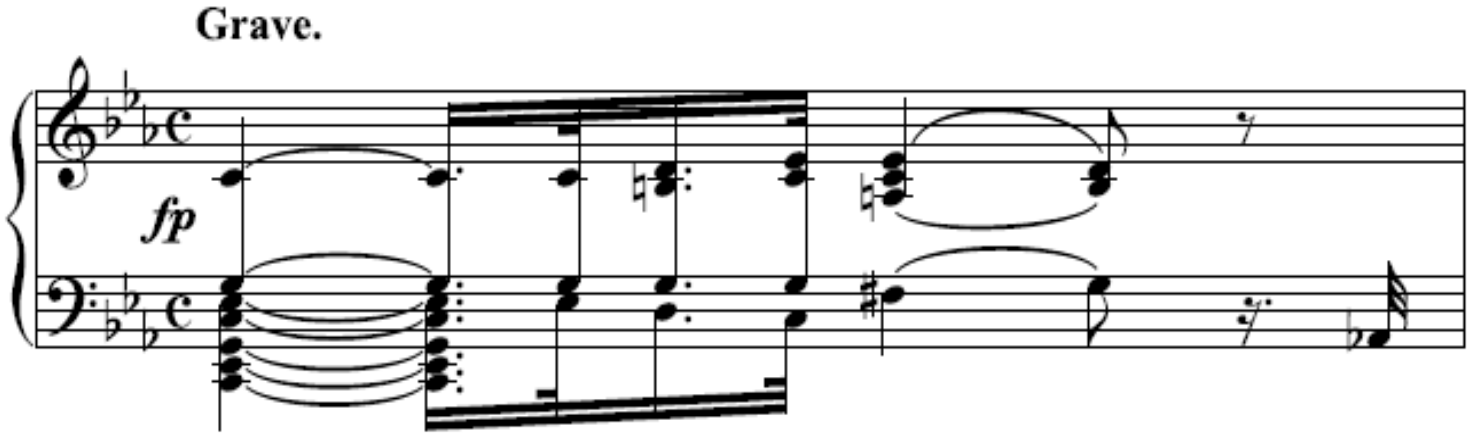


Application: Interpretation Switcher

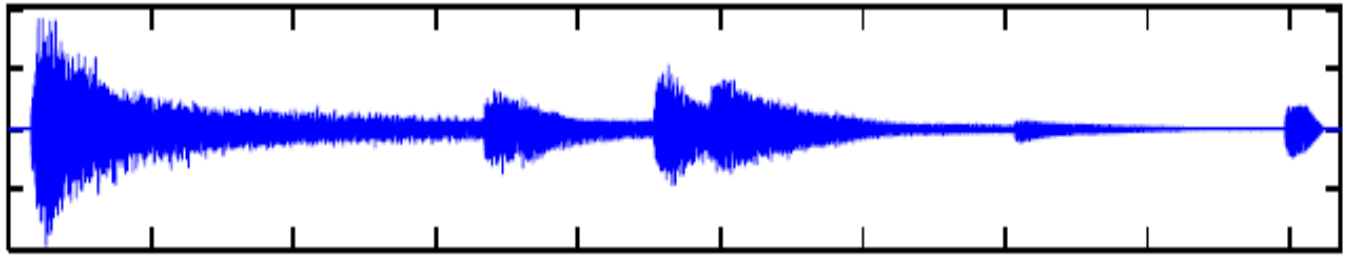


Music Synchronization: Image-Audio

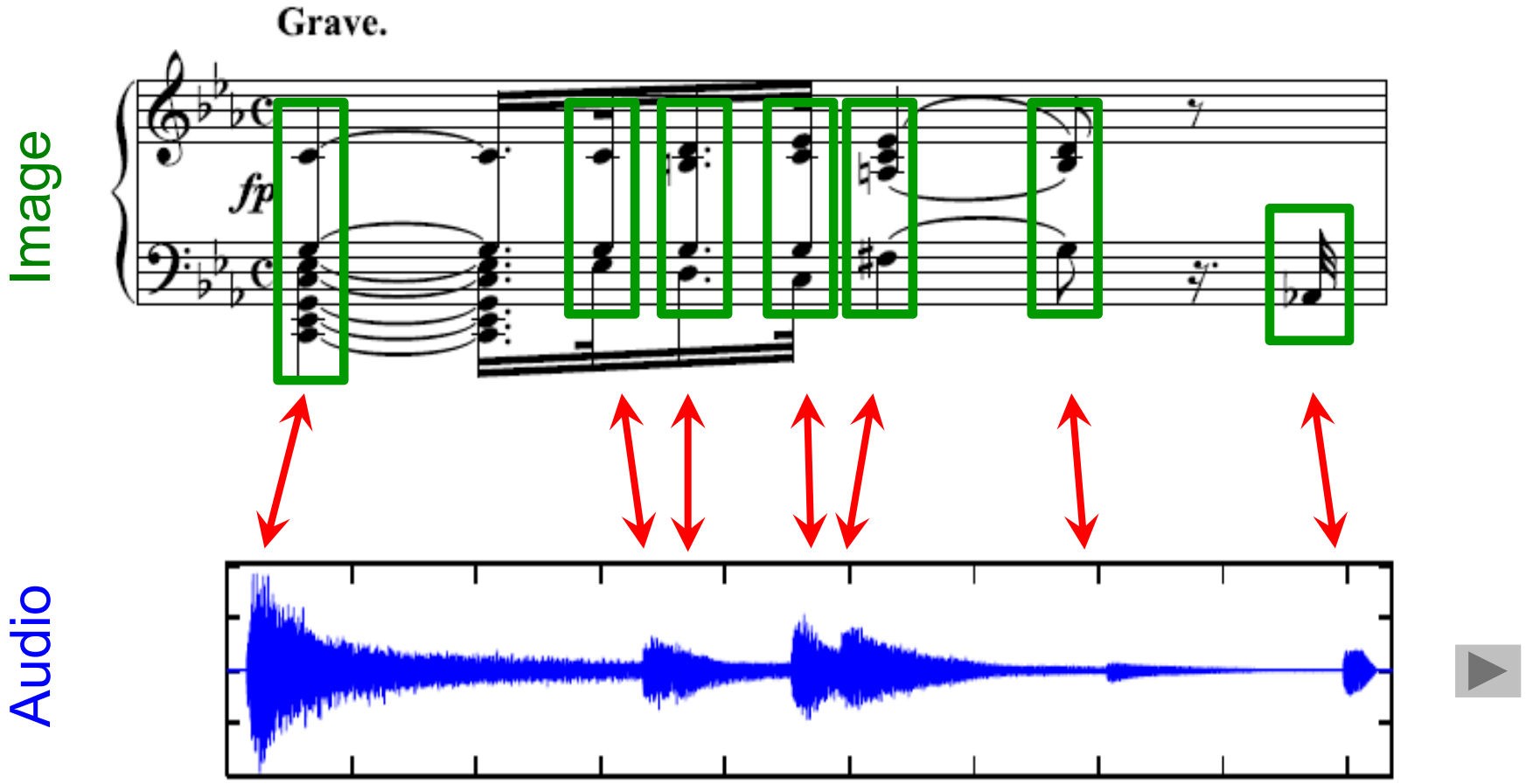
Image



Audio

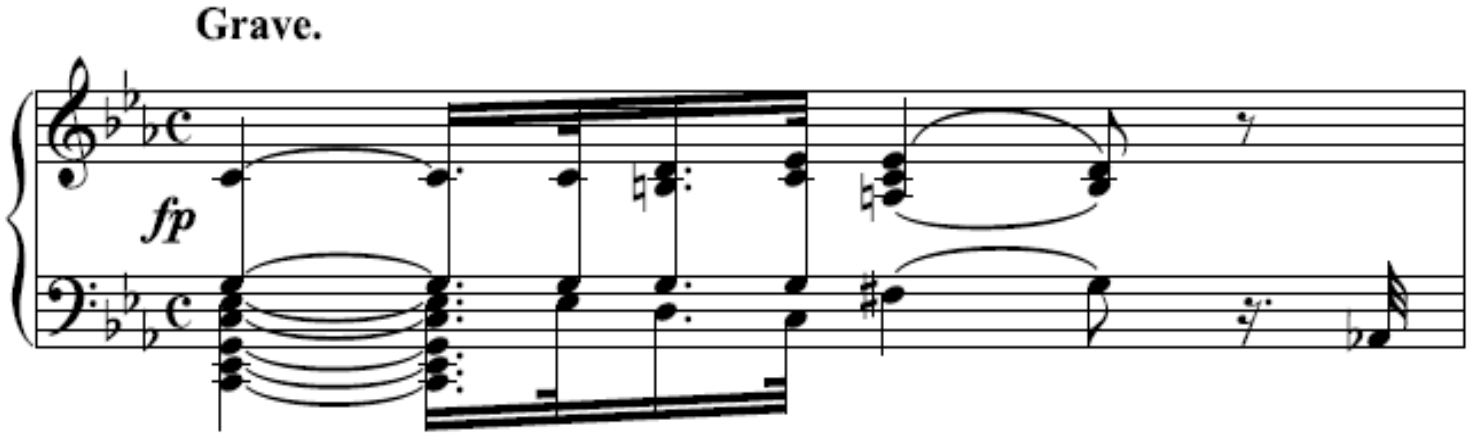


Music Synchronization: Image-Audio

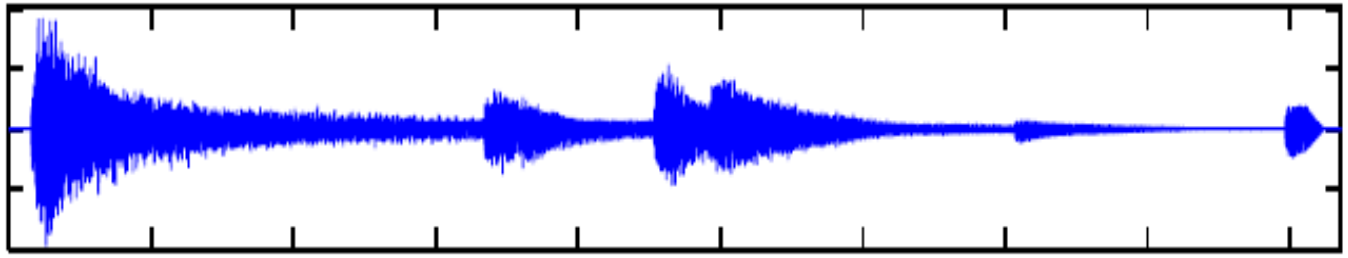


How to make the data comparable?

Image



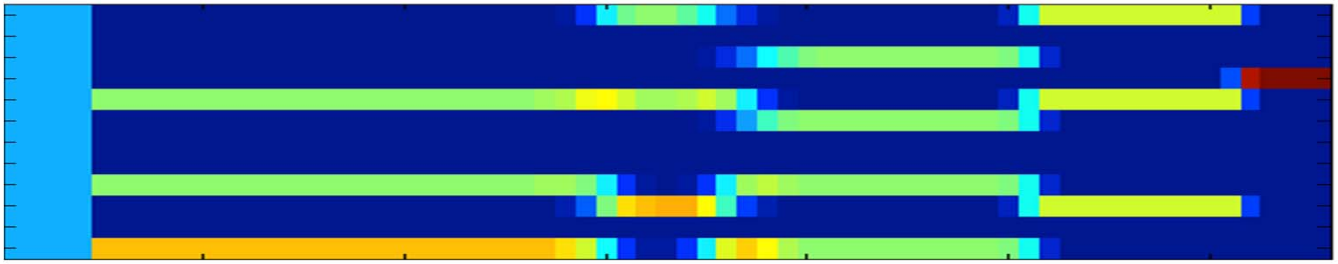
Audio



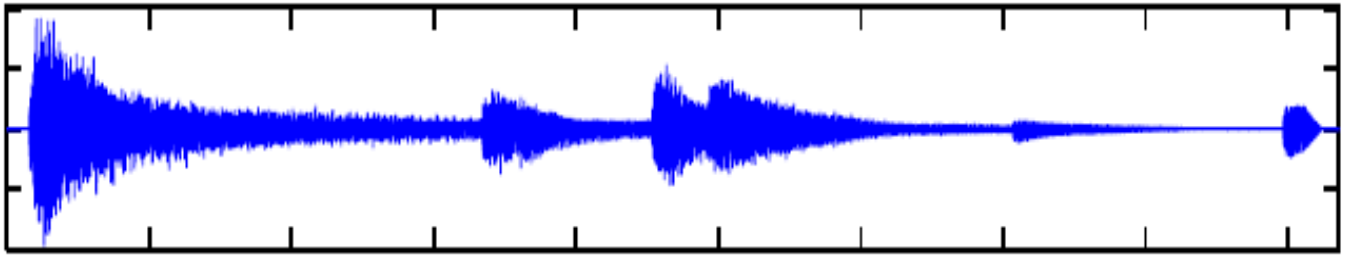
How to make the data comparable?

Image Processing: Optical Music Recognition

Image



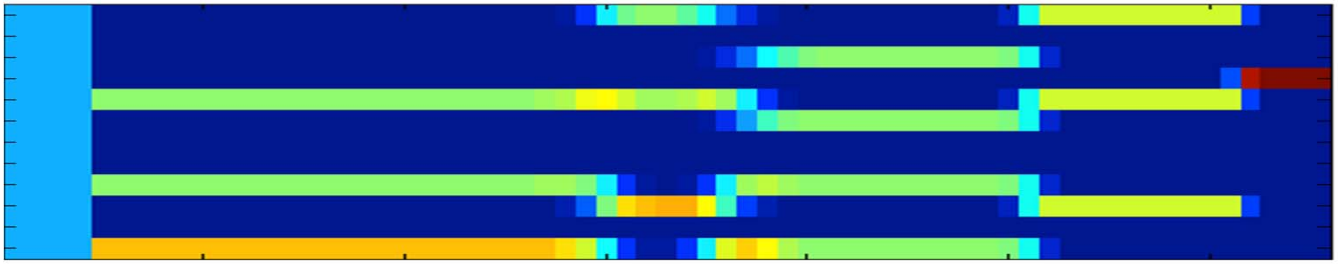
Audio



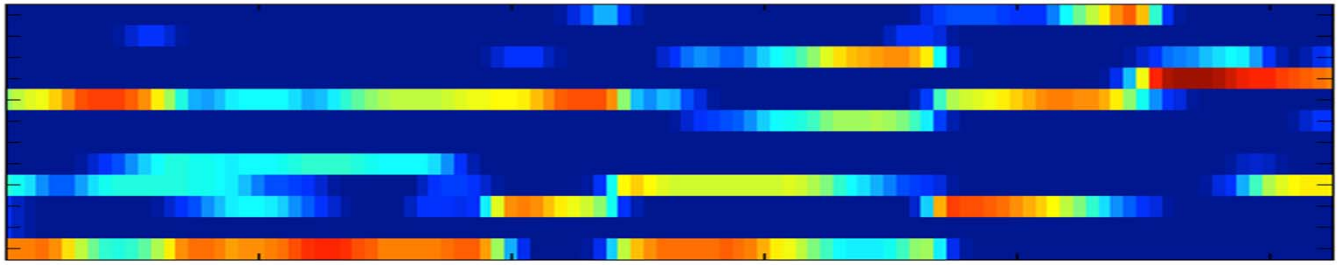
How to make the data comparable?

Image Processing: Optical Music Recognition

Image



Audio

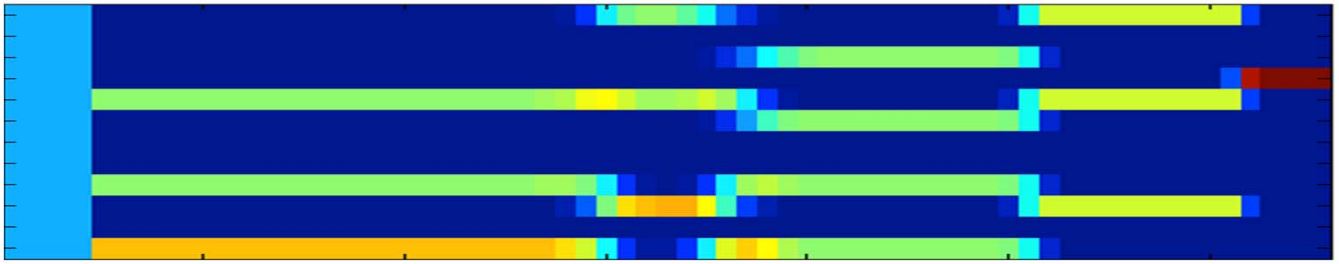


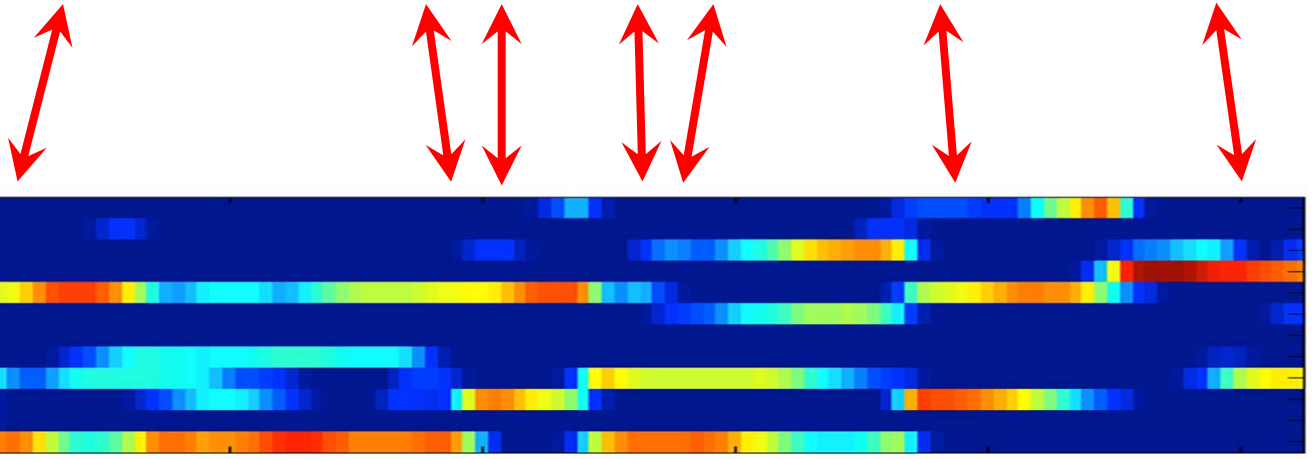
Audio Processing: Fourier Analyse

How to make the data comparable?

Image Processing: Optical Music Recognition

Image

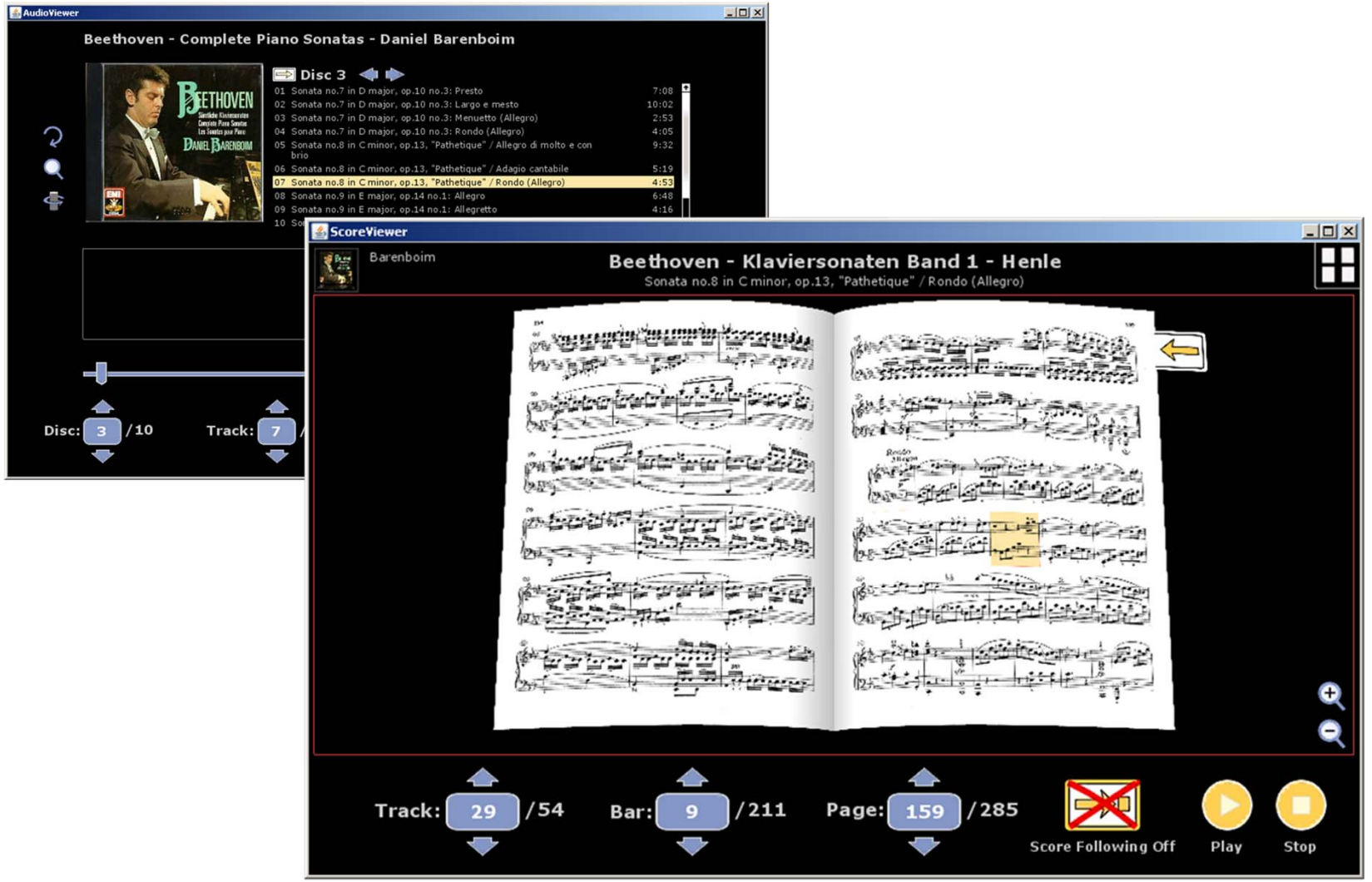




Audio

Audio Processing: Fourier Analyse

Application: Score Viewer



Audio Matching

Database: Audio collection containing:

- Several recordings of the same piece of music
- Different interpretations by various musicians
- Arrangements in different instrumentations

Goal: Given a short query audio fragment, find all corresponding audio fragments of similar musical content.

Notes:

- Instance of fragment-based retrieval
- Medium specificity
- A single document may contain several hits
- Cross-modal retrieval also feasible

Audio Matching

Two main ingredients:

1.) Audio features

- Robust but discriminating
- Chroma-based features
- Correlate to harmonic progression
- Robust to variations in dynamics, timbre, articulation, local tempo

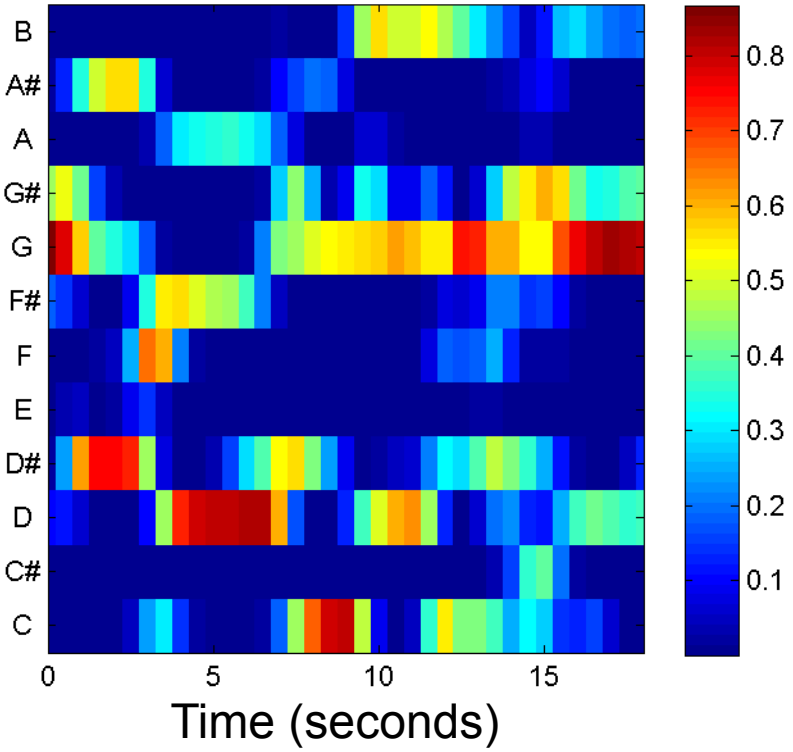
2.) Matching procedure

- Efficient
- Robust to local and global tempo variations
- Scalable using index structure

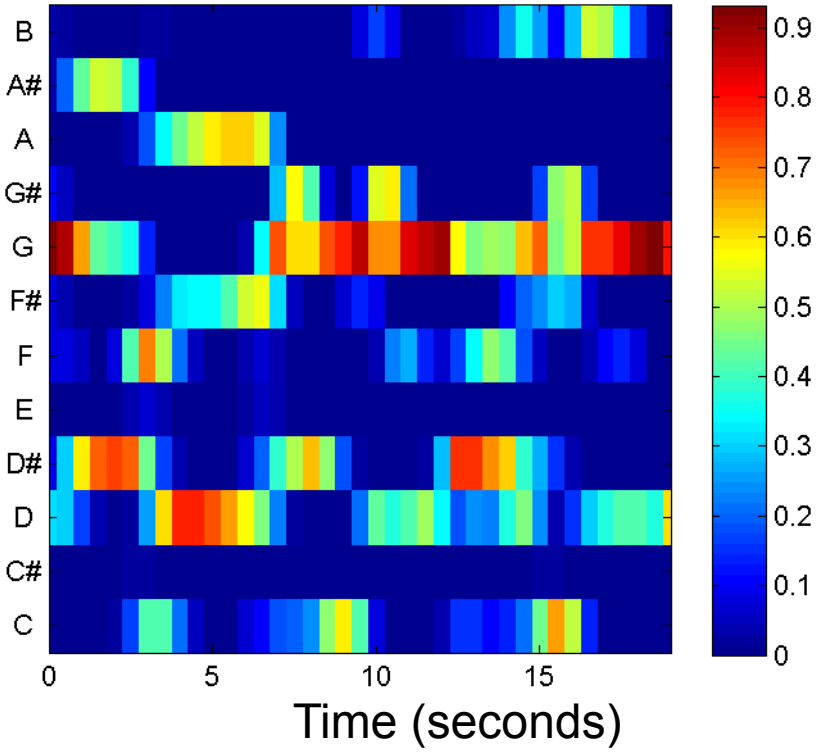
Audio Features

Example: Beethoven's Fifth
Chroma representation (normalized, 2 Hz)

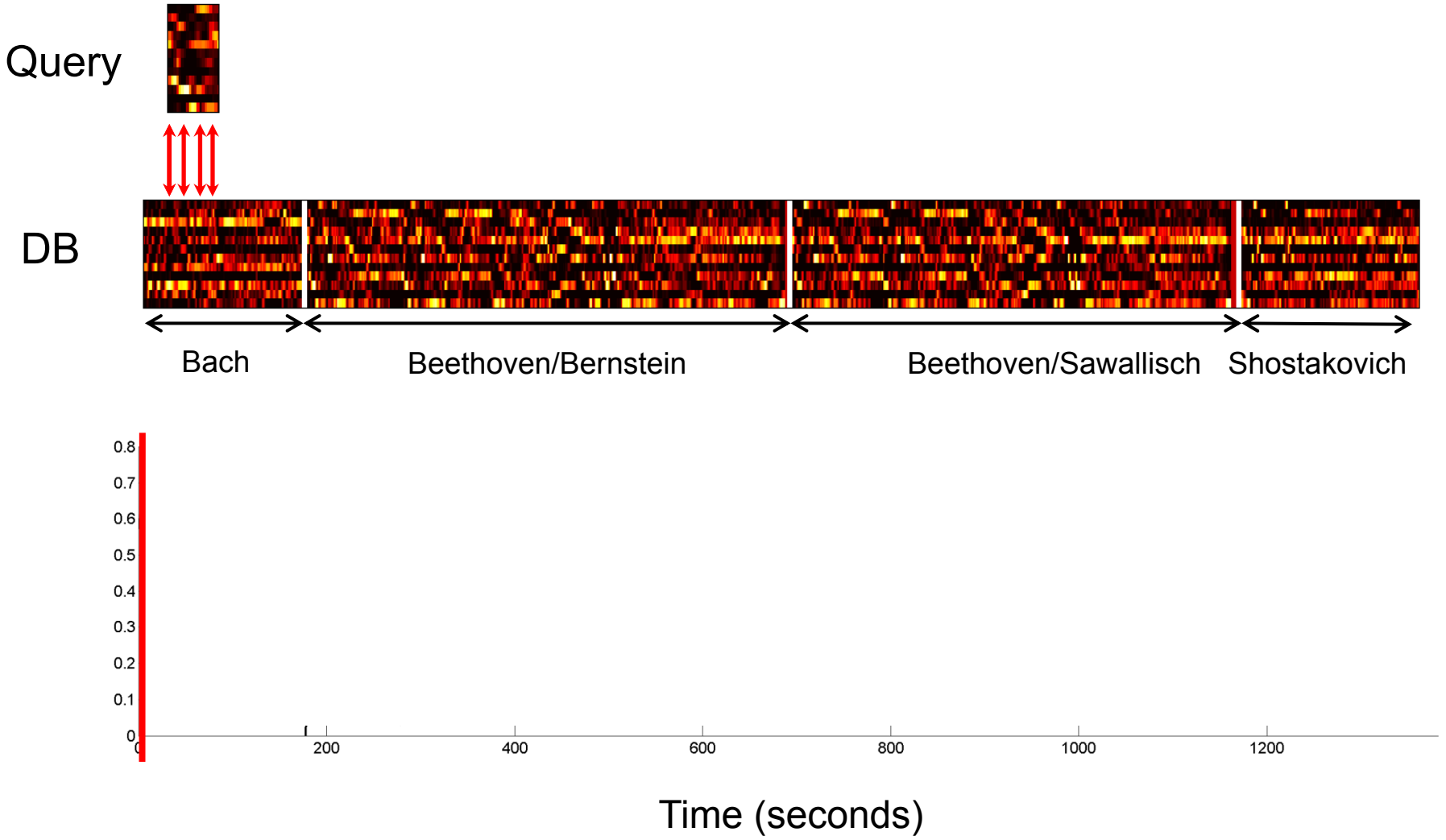
Karajan



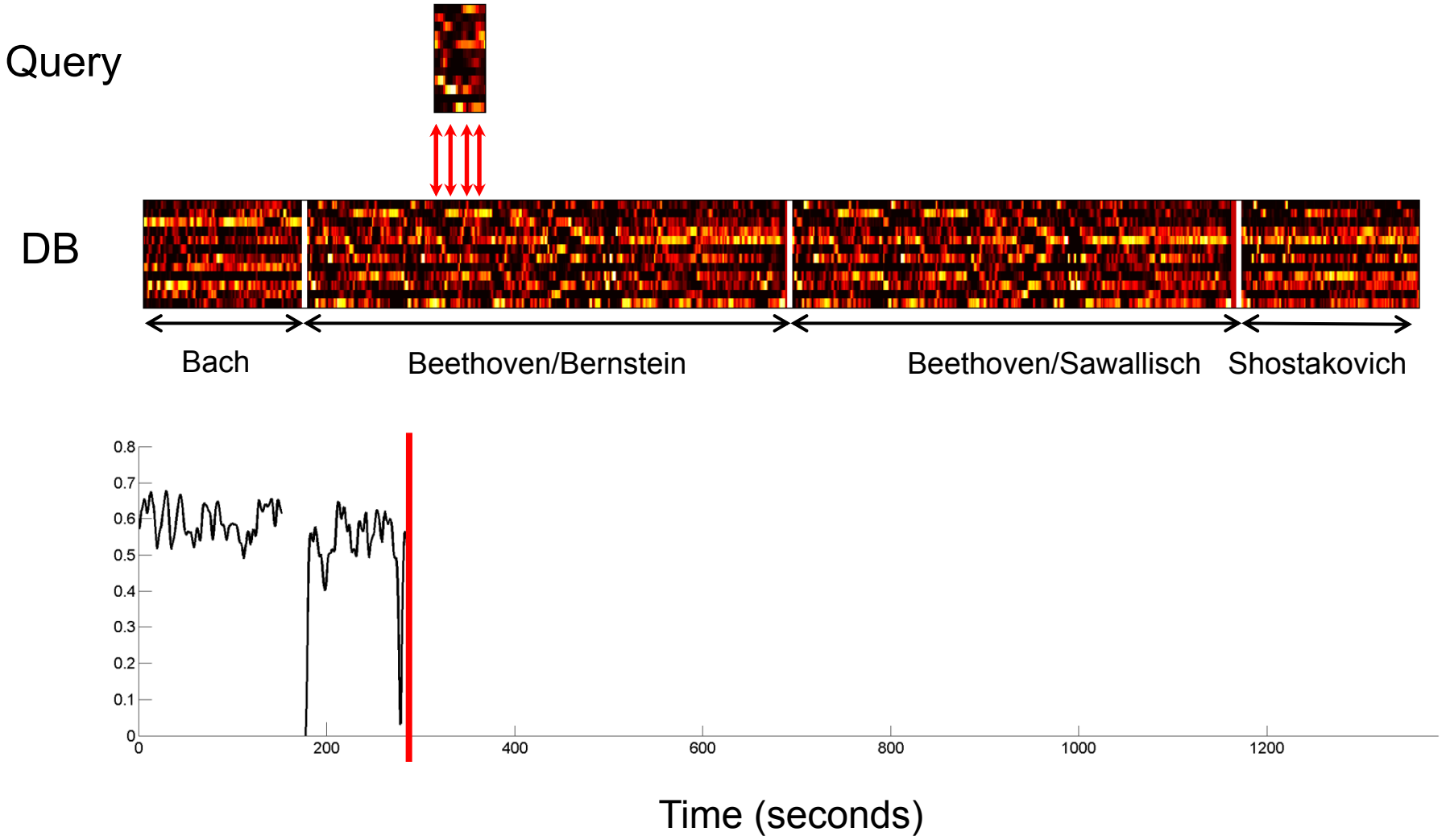
Scherbakov



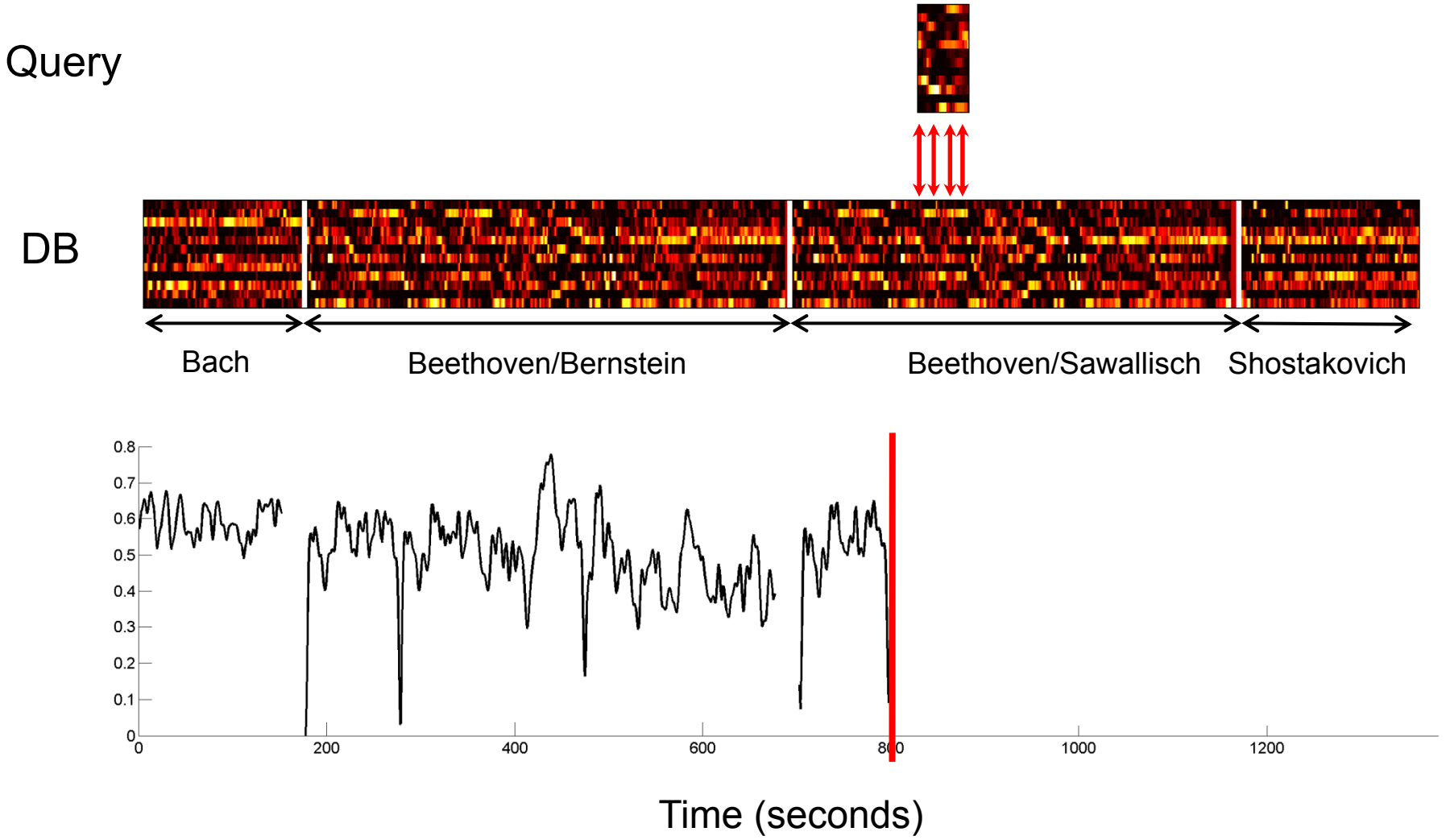
Matching Procedure



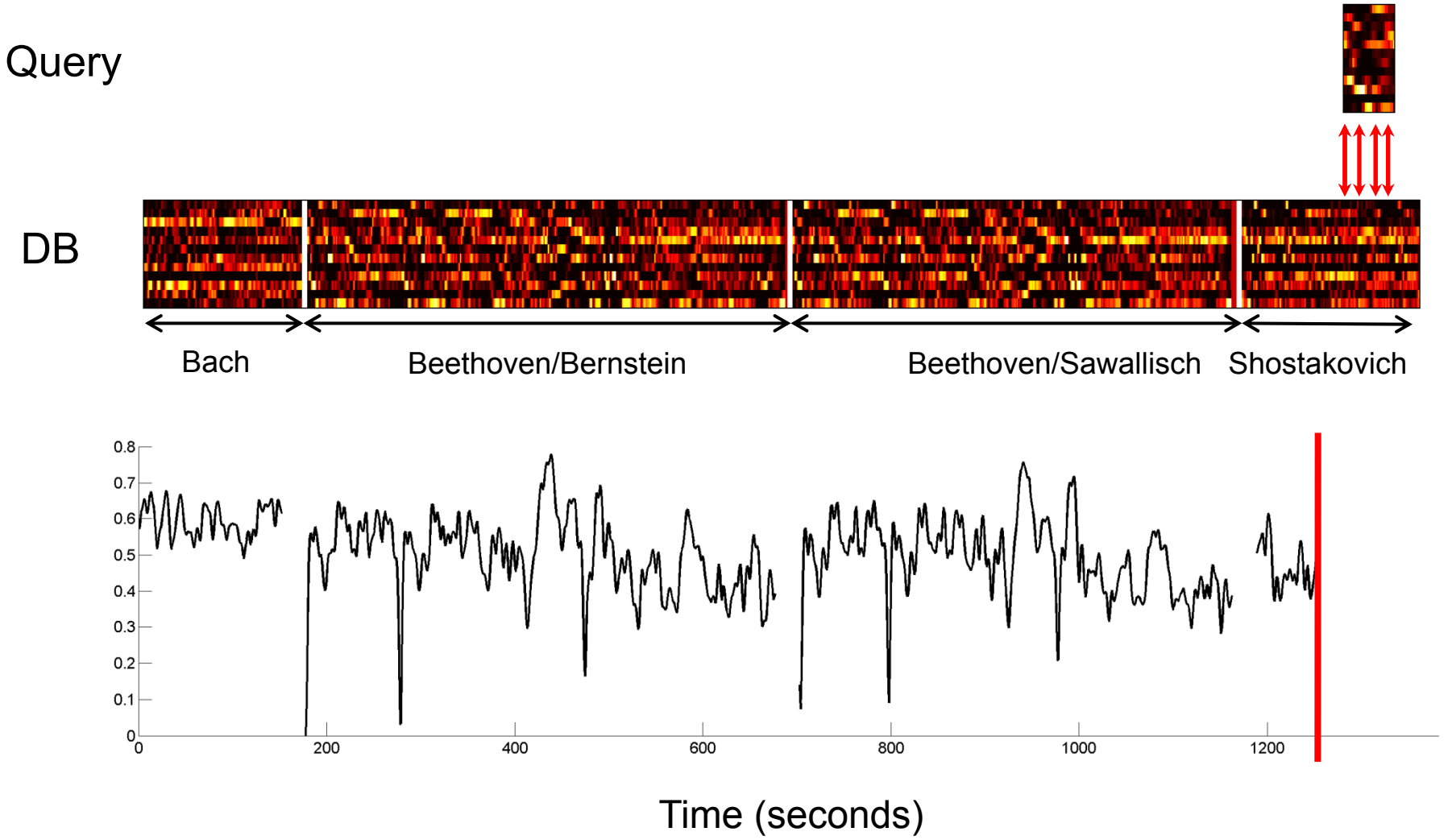
Matching Procedure



Matching Procedure



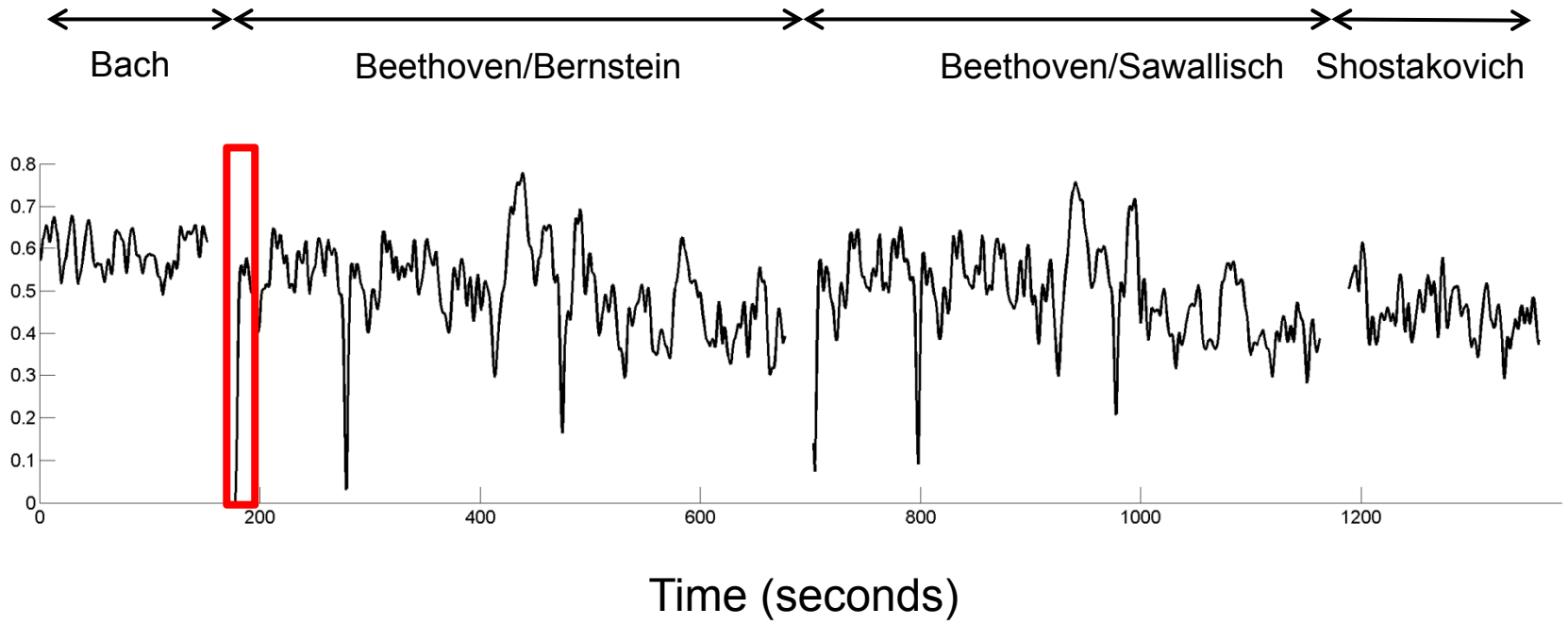
Matching Procedure



Matching Procedure

Matching curve

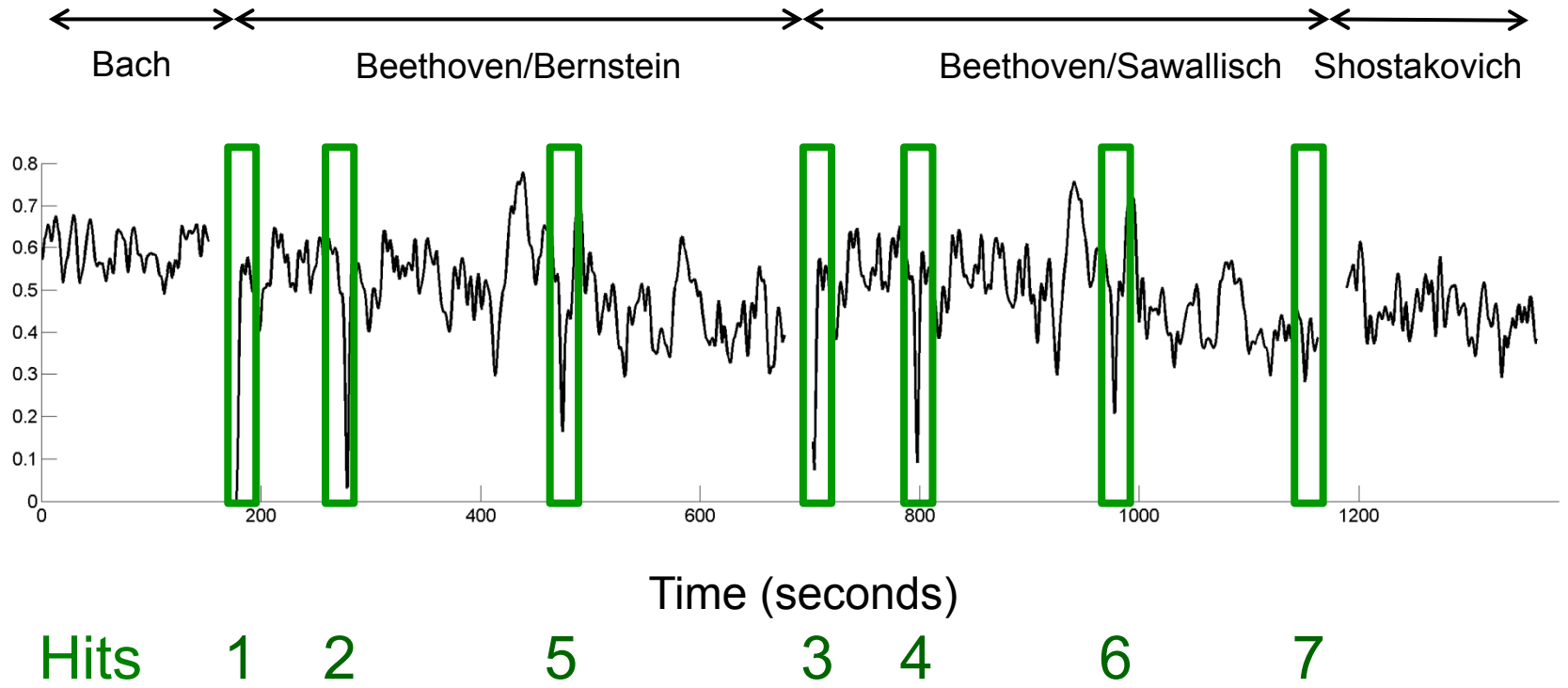
Query: Beethoven's Fifth / Bernstein (first 20 seconds)



Matching Procedure

Matching curve

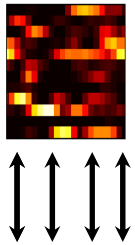
Query: Beethoven's Fifth / Bernstein (first 20 seconds)

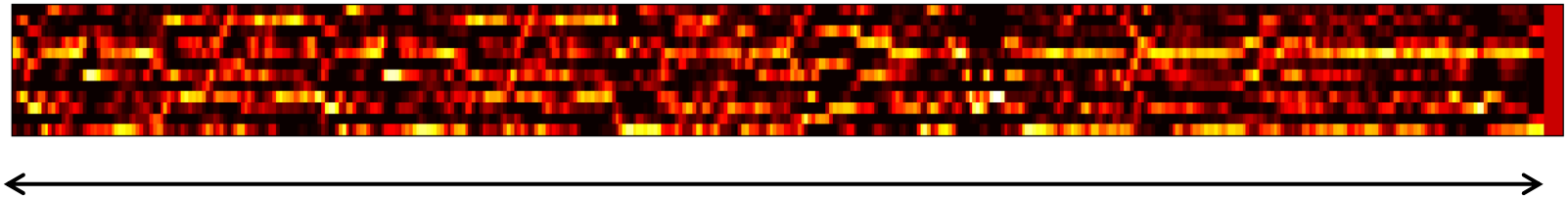


Matching Procedure

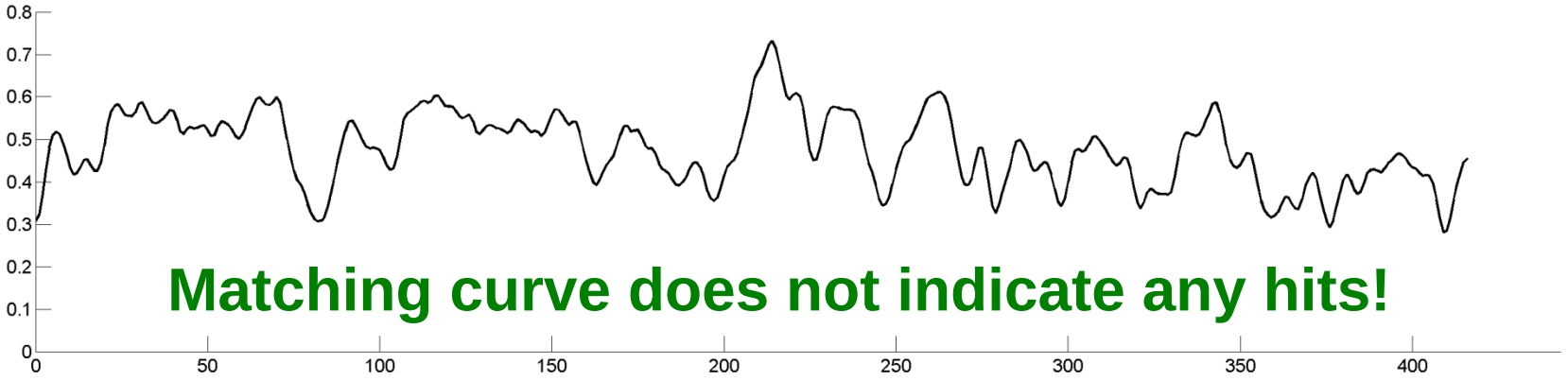
Problem: How to deal with tempo differences?

Karajan is much
faster then Bernstein!





Beethoven/Karajan



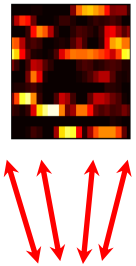
Matching curve does not indicate any hits!

Time (seconds)

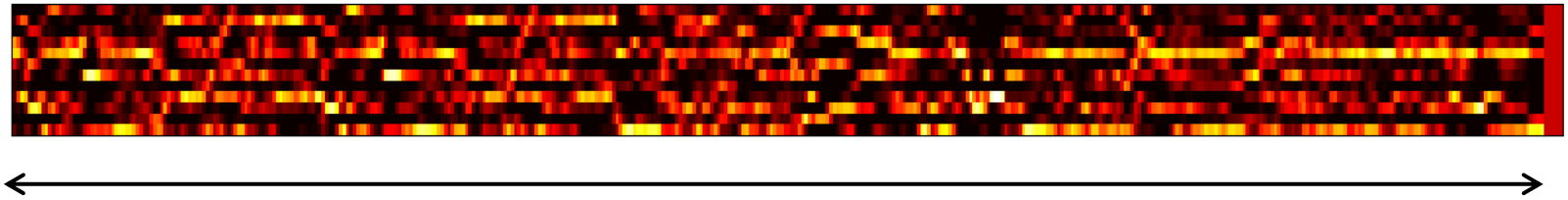
Matching Procedure

1. Strategy: Usage of local warping

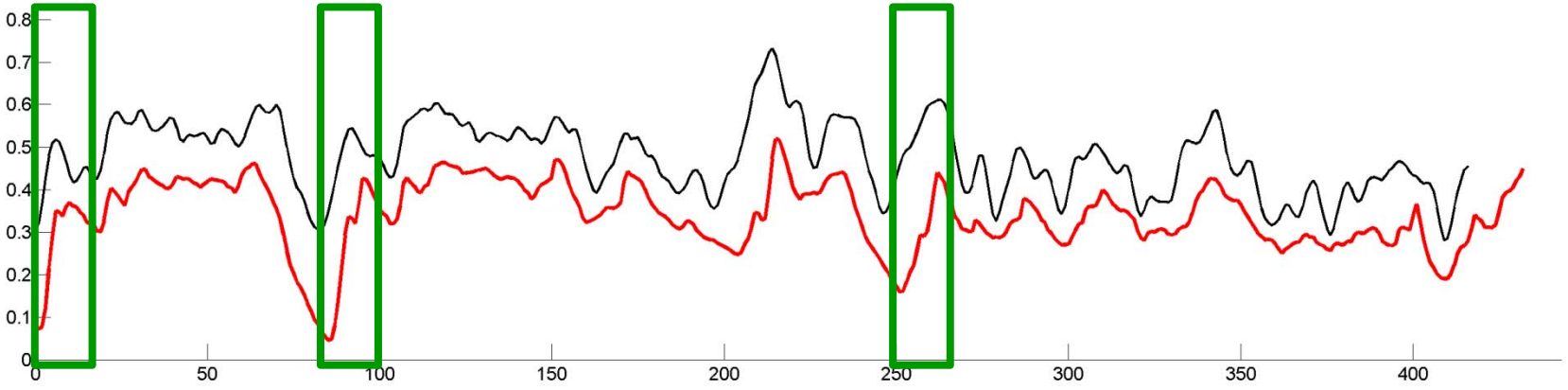
Karajan is much
faster then Bernstein!



Warping strategies are computationally expensive and hard for indexing.



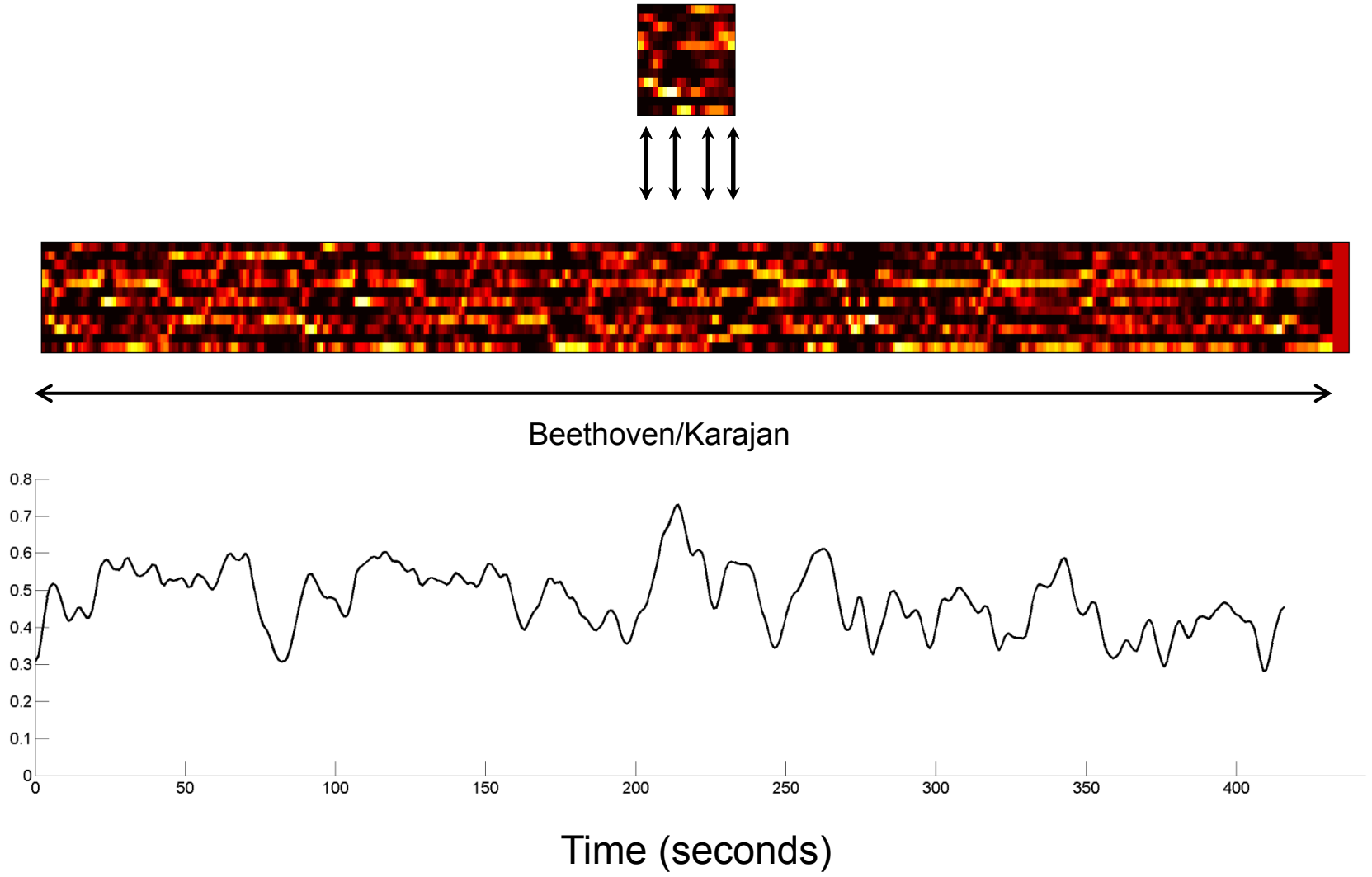
Beethoven/Karajan



Time (seconds)

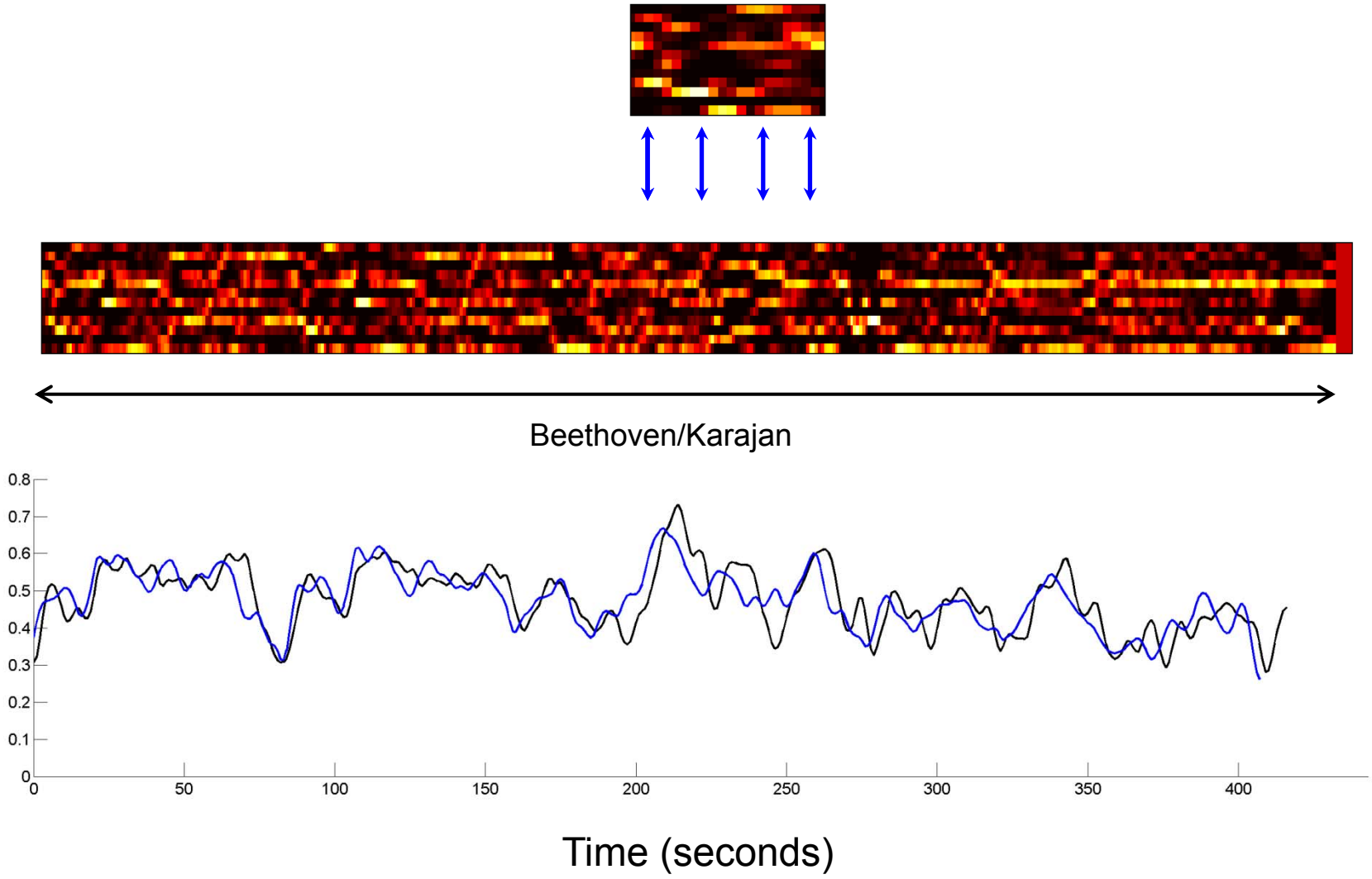
Matching Procedure

2. Strategy: Usage of multiple scaling



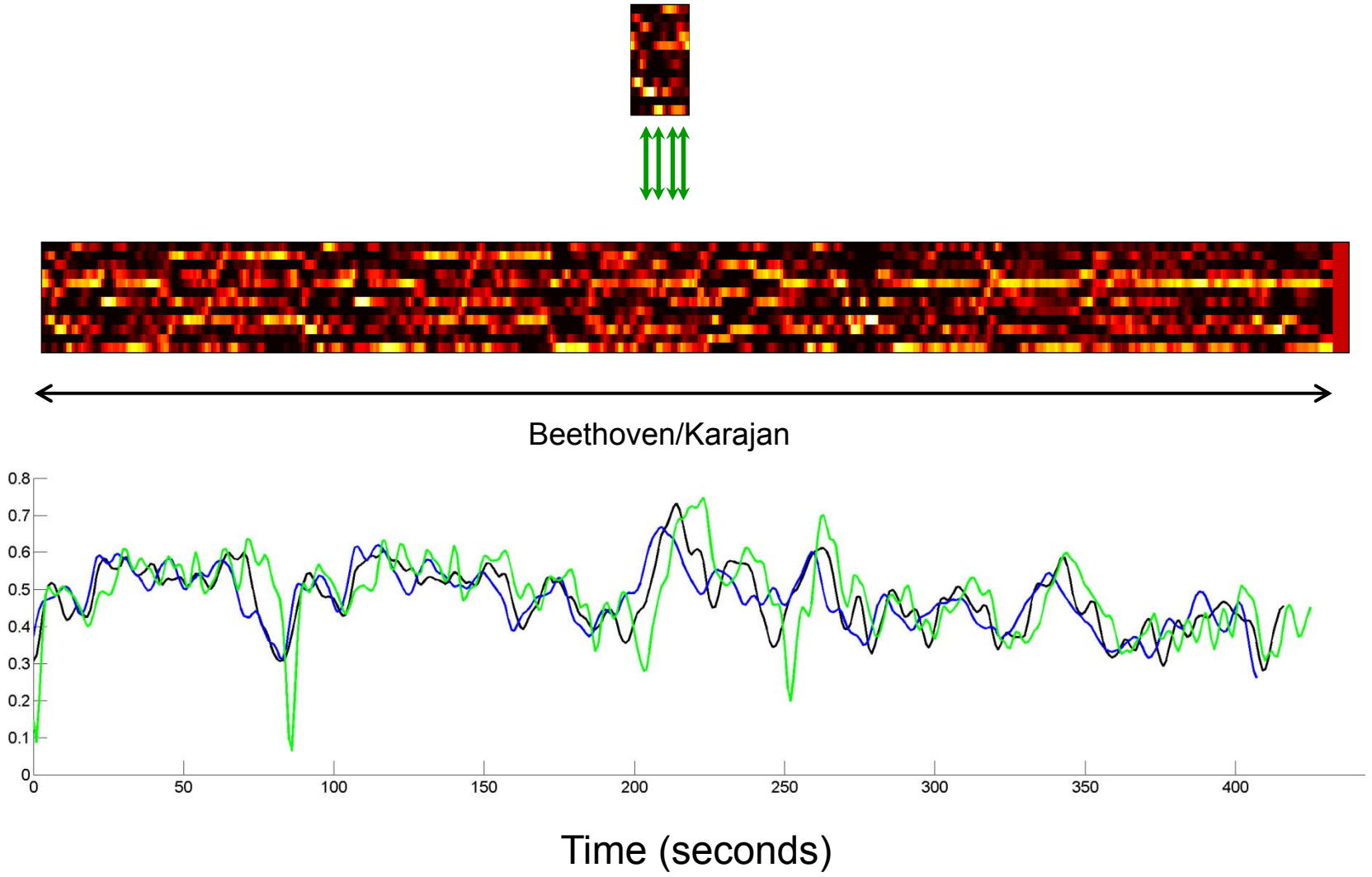
Matching Procedure

2. Strategy: Usage of multiple scaling



Matching Procedure

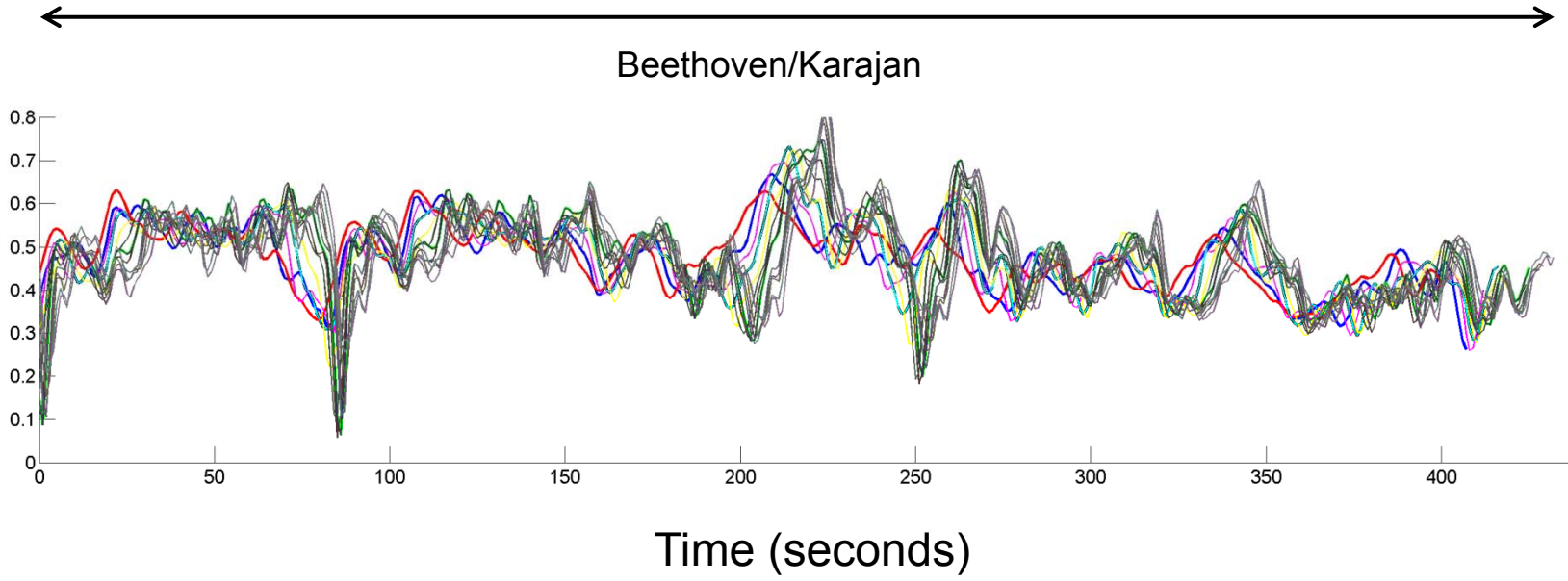
2. Strategy: Usage of multiple scaling



Matching Procedure

2. Strategy: Usage of multiple scaling

Query resampling simulates tempo changes

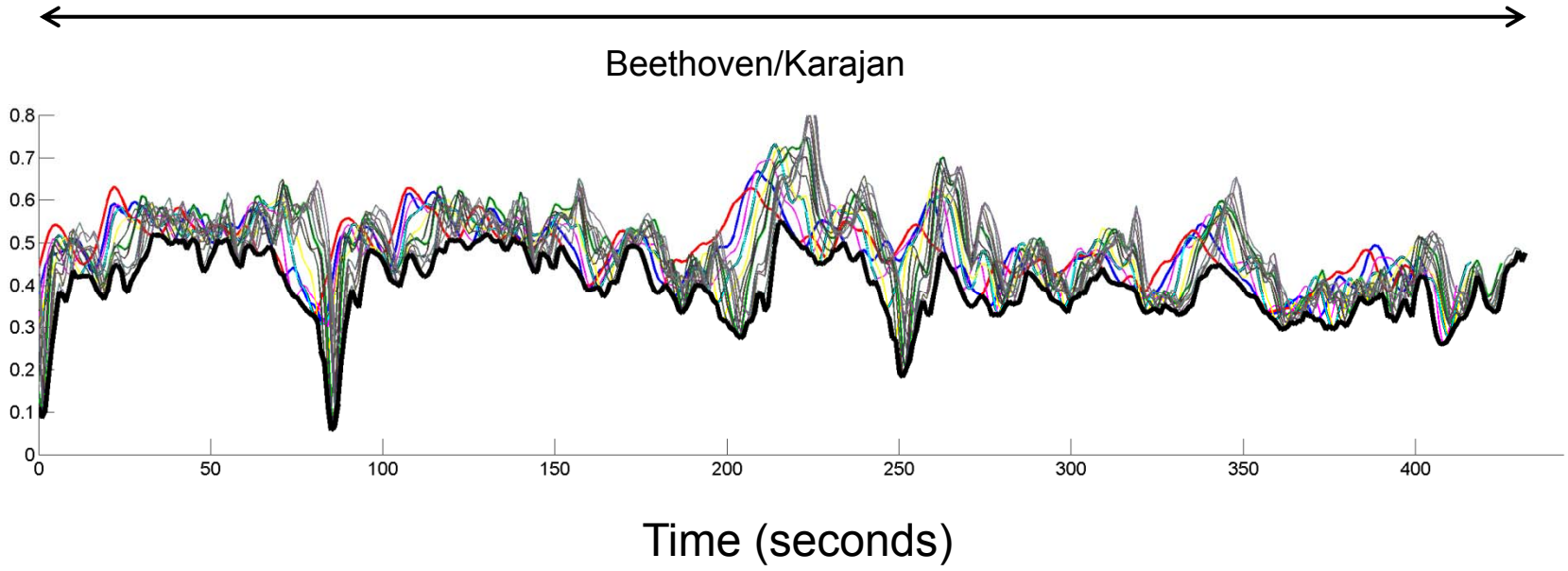


Matching Procedure

2. Strategy: Usage of multiple scaling

Query resampling simulates tempo changes

Minimize over all curves



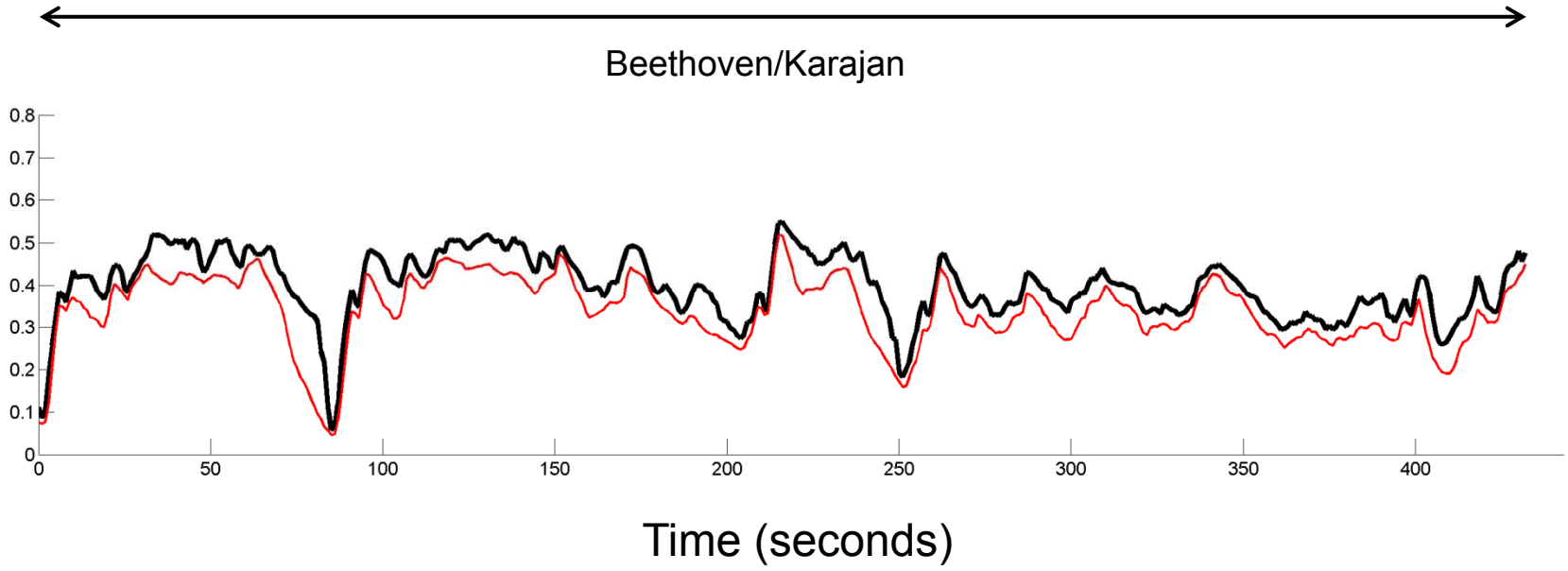
Matching Procedure

2. Strategy: Usage of multiple scaling

Query resampling simulates tempo changes

Minimize over all curves

Resulting curve is similar warping curve



Audio Matching for Feature Analysis

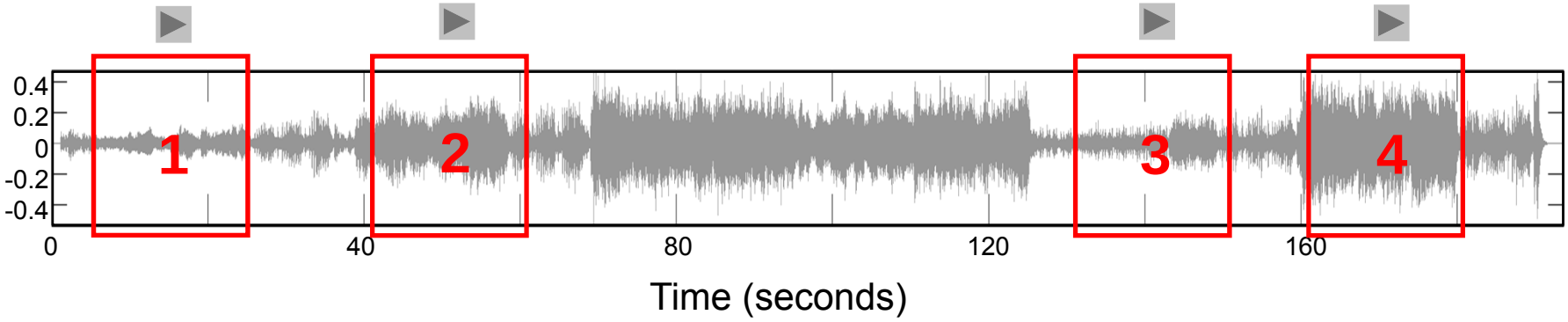
Idea:

Use “Audio Matching” for analyzing and understanding audio & feature properties:

- Relative comparison
- Compact
- Intuitive
- Quantitative evaluation

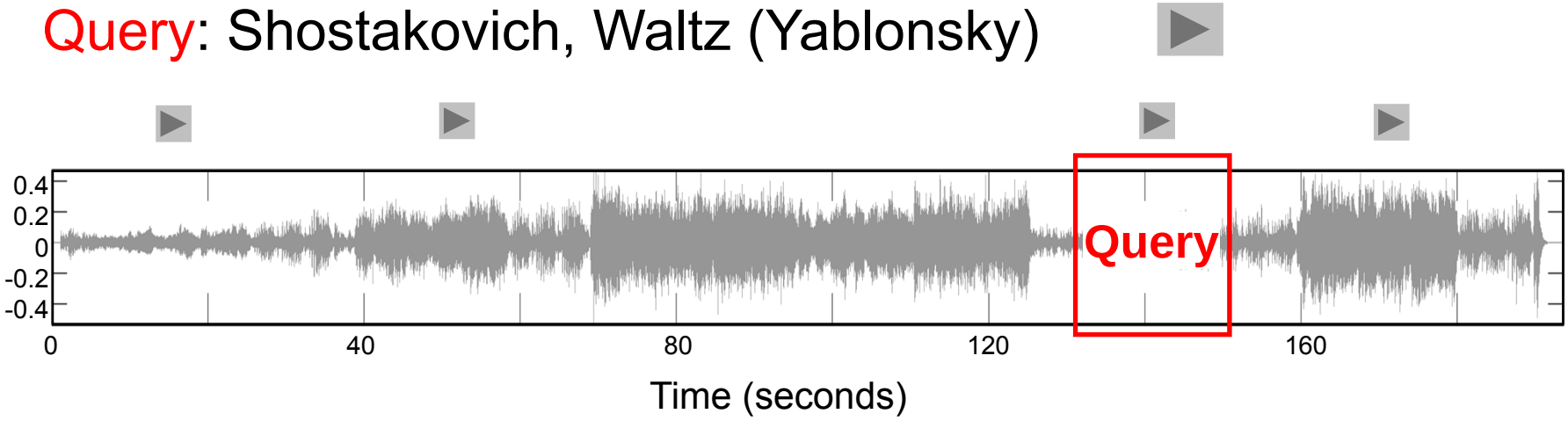
Audio Matching for Feature Analysis

Example: Shostakovich, Waltz (Yablonsky)



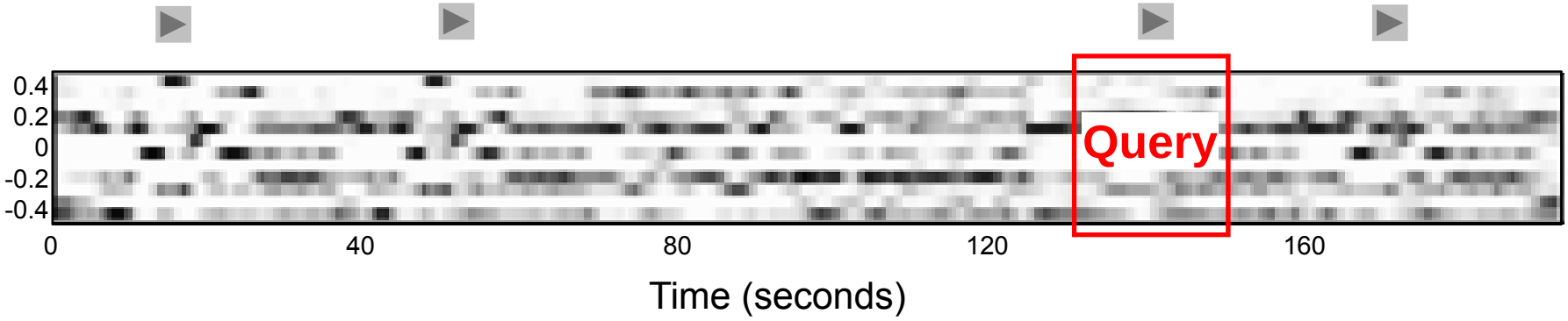
Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



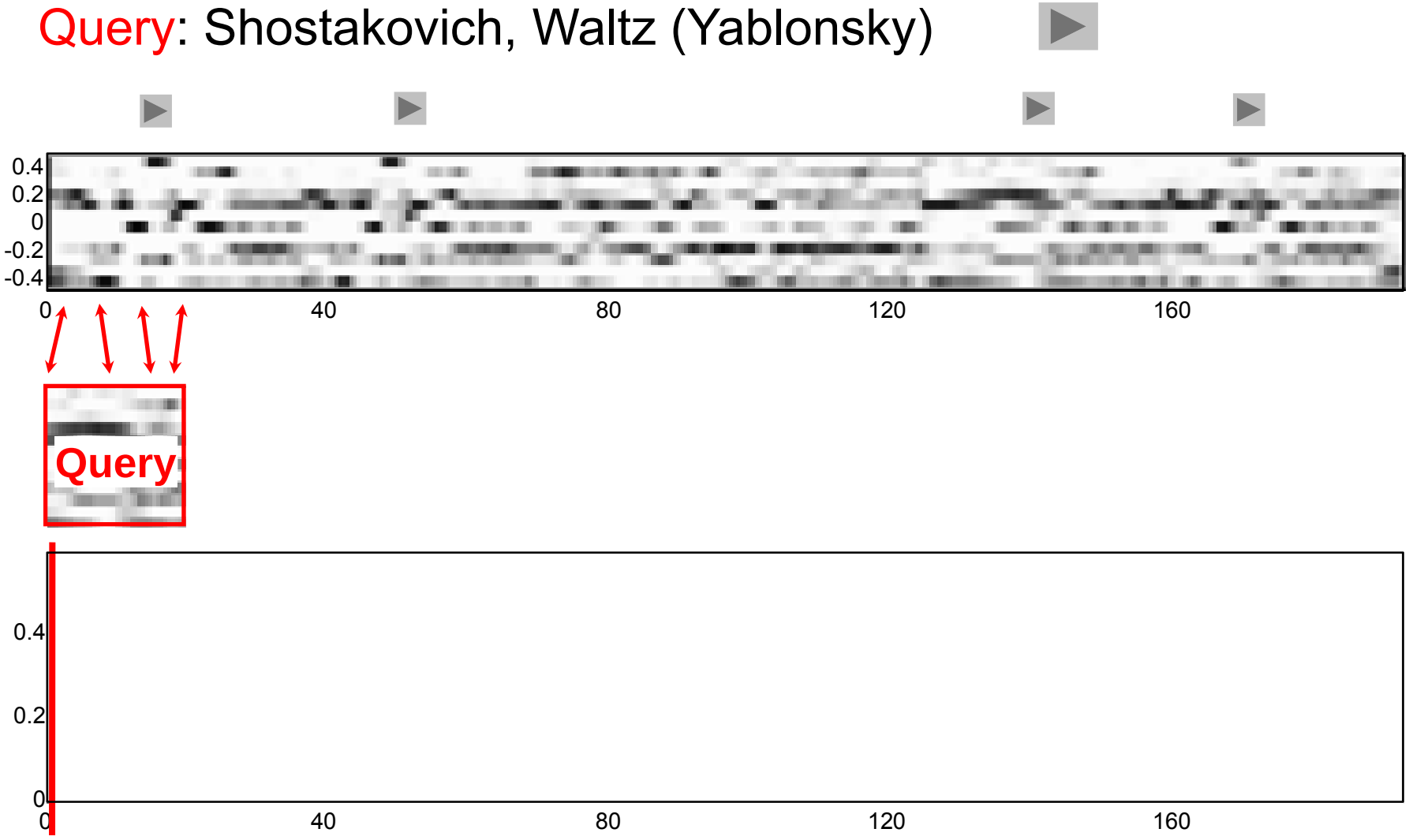
Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



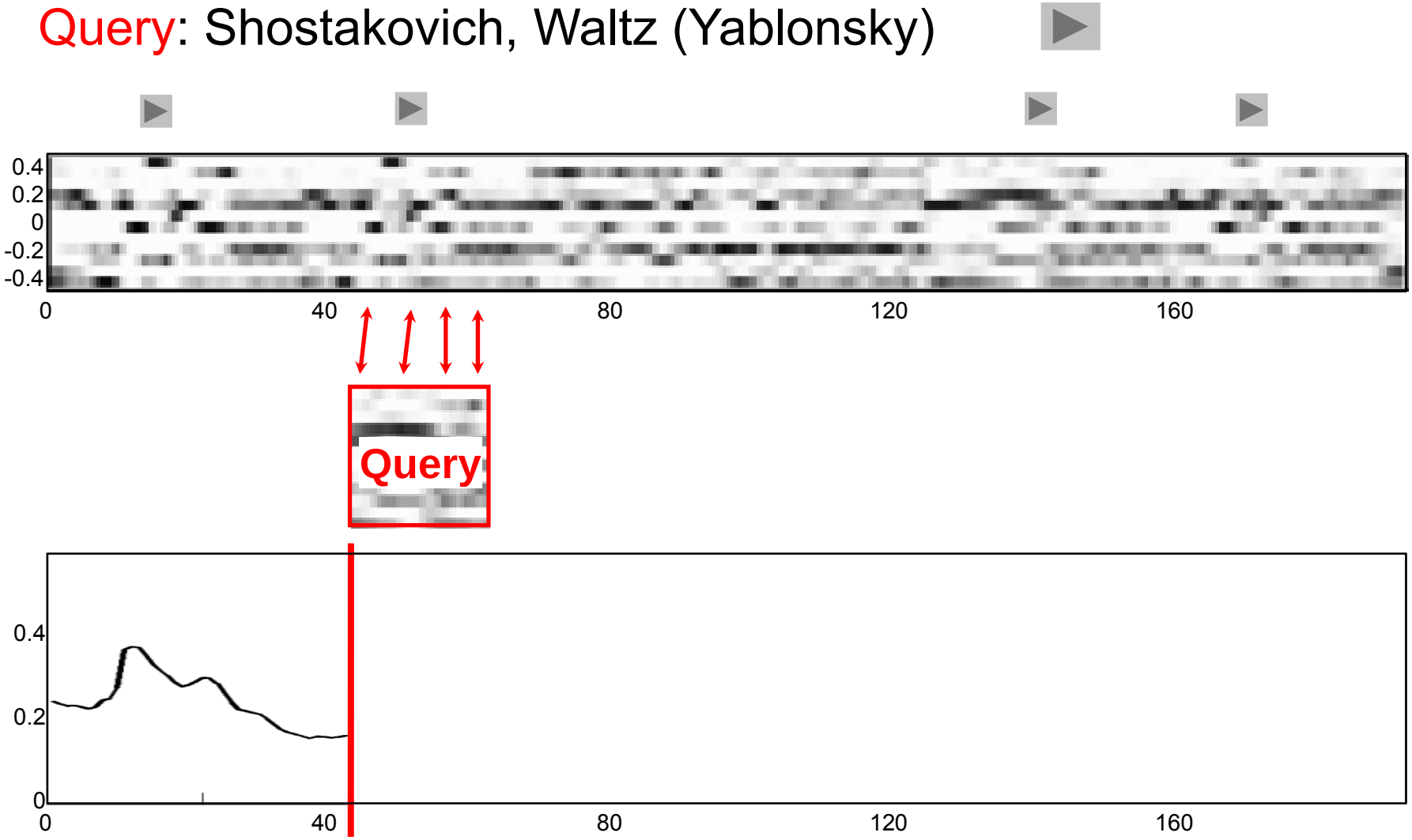
Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



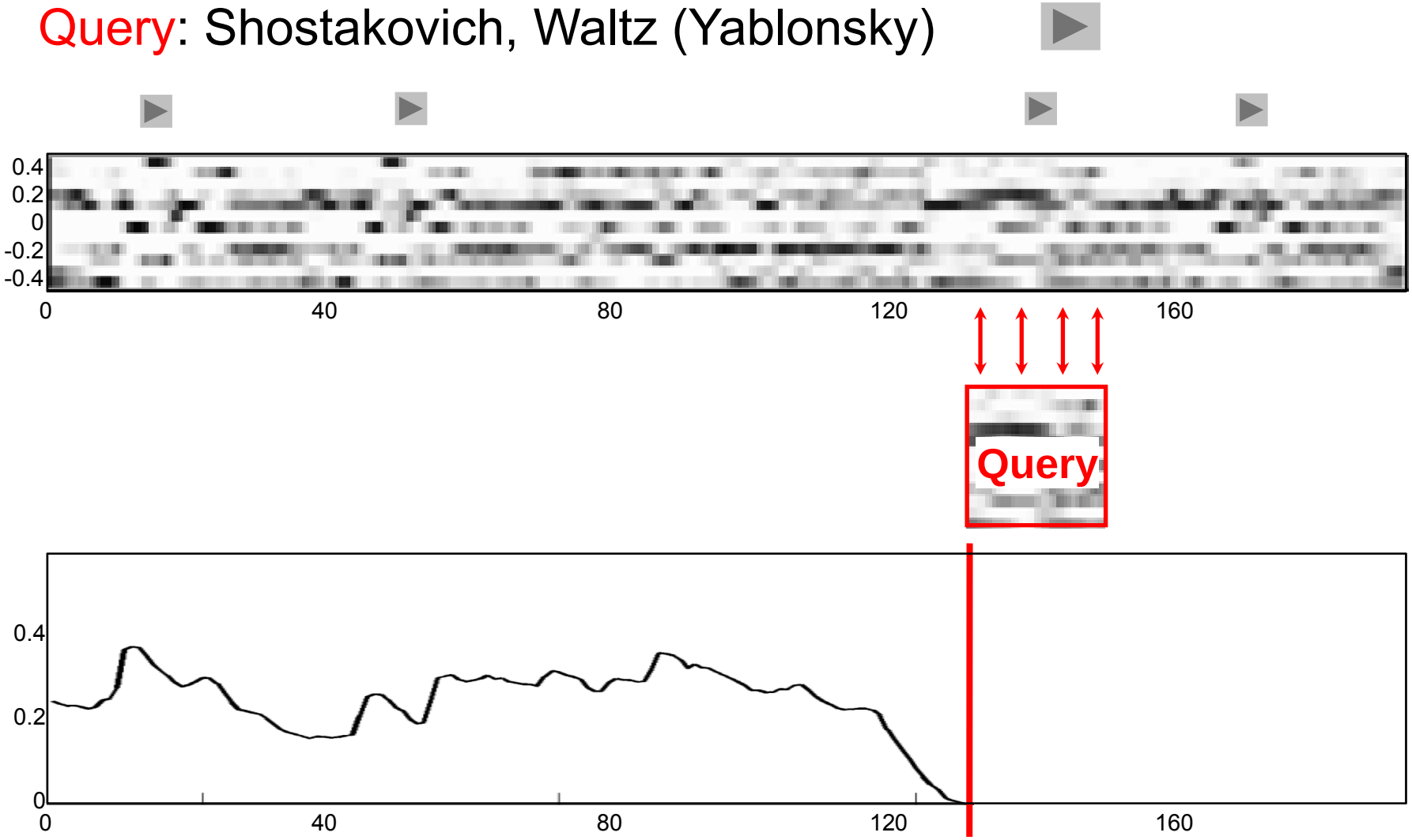
Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



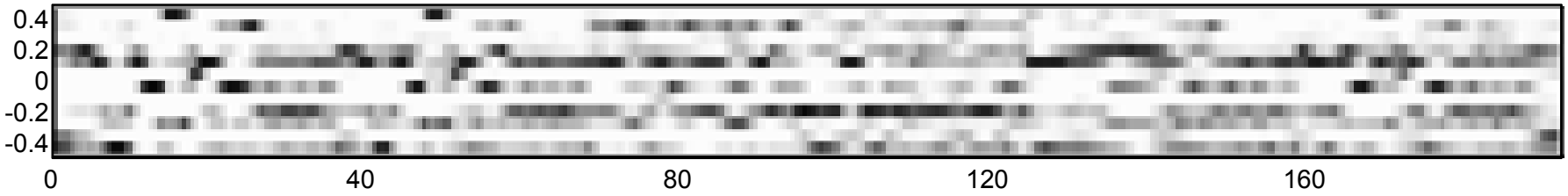
Audio Matching for Feature Analysis

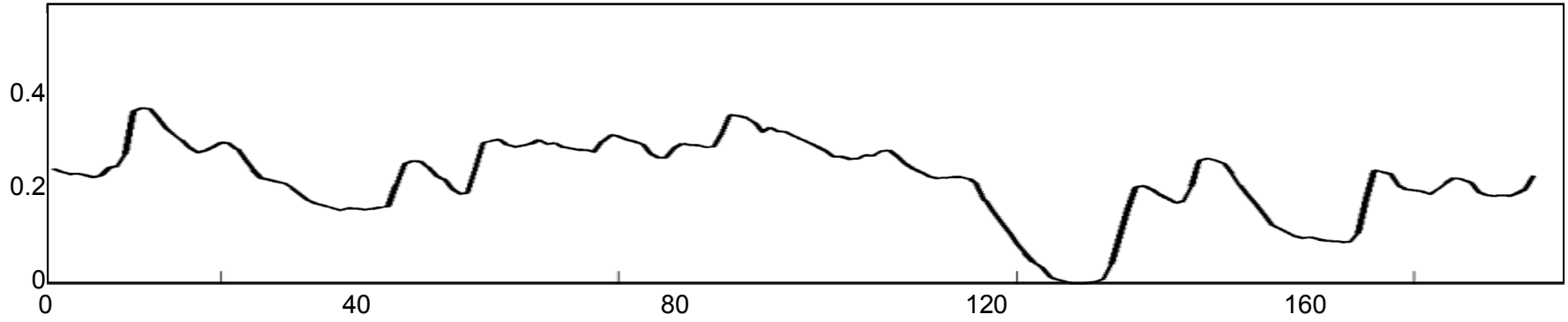
Query: Shostakovich, Waltz (Yablonsky)



Audio Matching for Feature Analysis

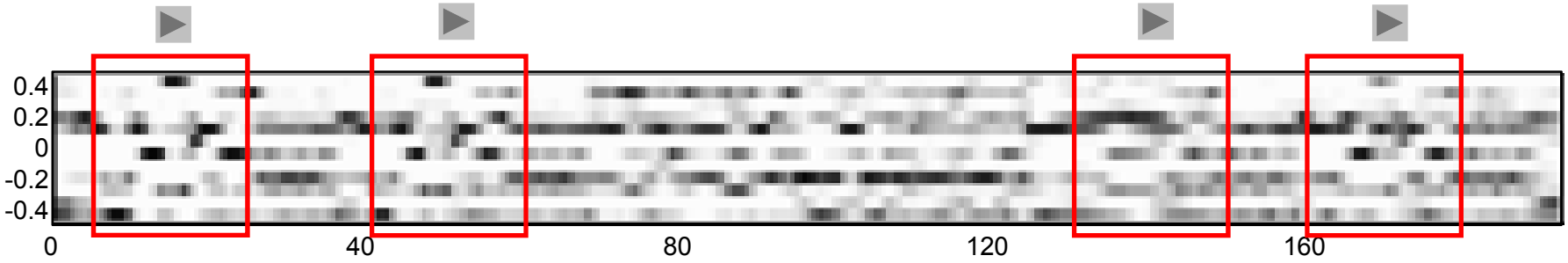
Query: Shostakovich, Waltz (Yablonsky)



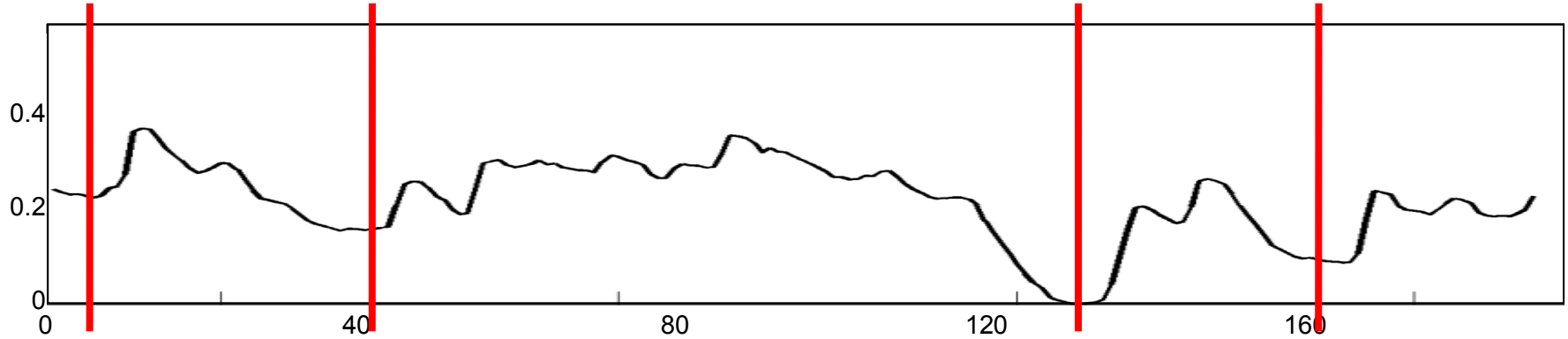


Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



Expected matching positions (should have local minima)

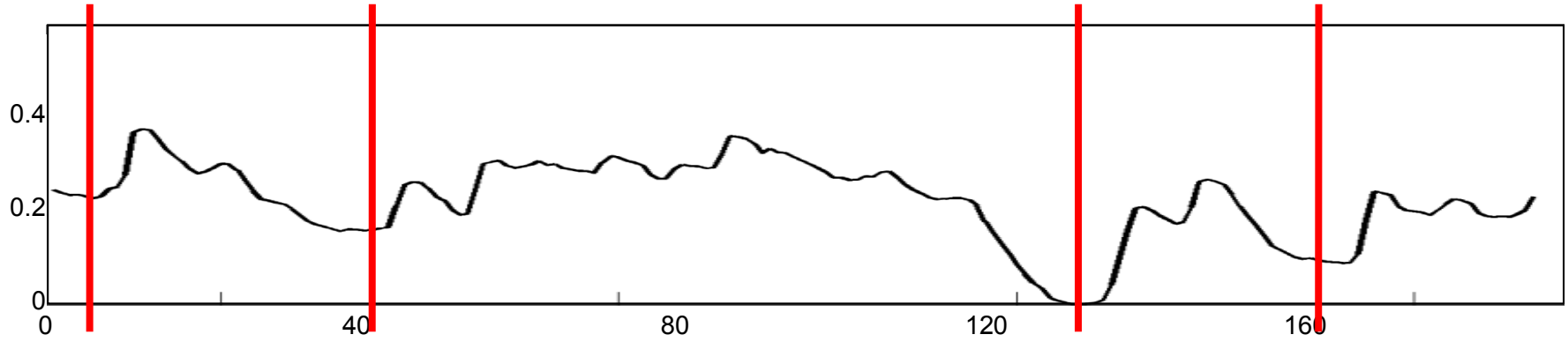


Audio Matching for Feature Analysis

Idea:

- Use matching curve for analyzing feature properties

Expected matching positions (should have local minima)

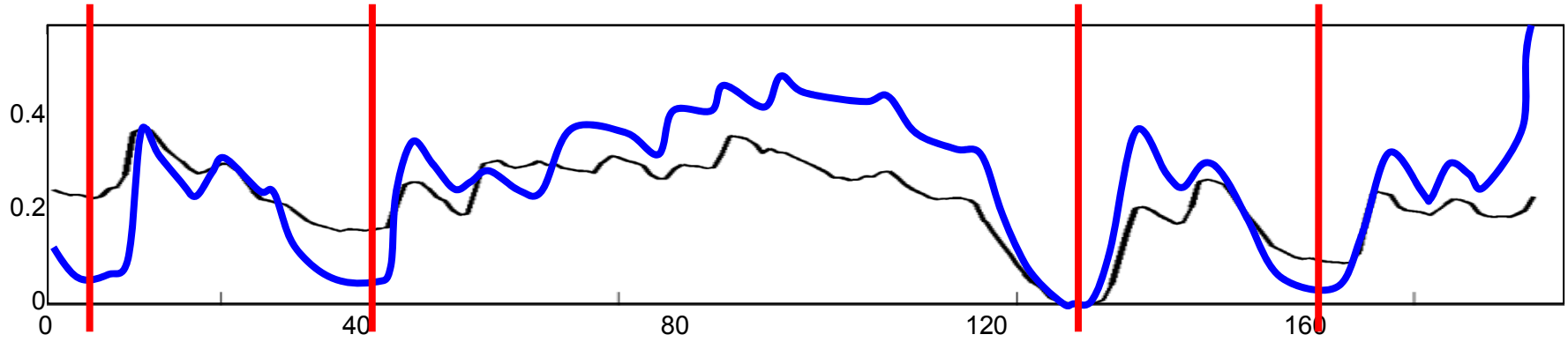


Audio Matching for Feature Analysis

Idea:

- Use matching curve for analyzing feature properties
- Example: Chroma feature of higher timbre invariance

Expected matching positions (should have local minima)

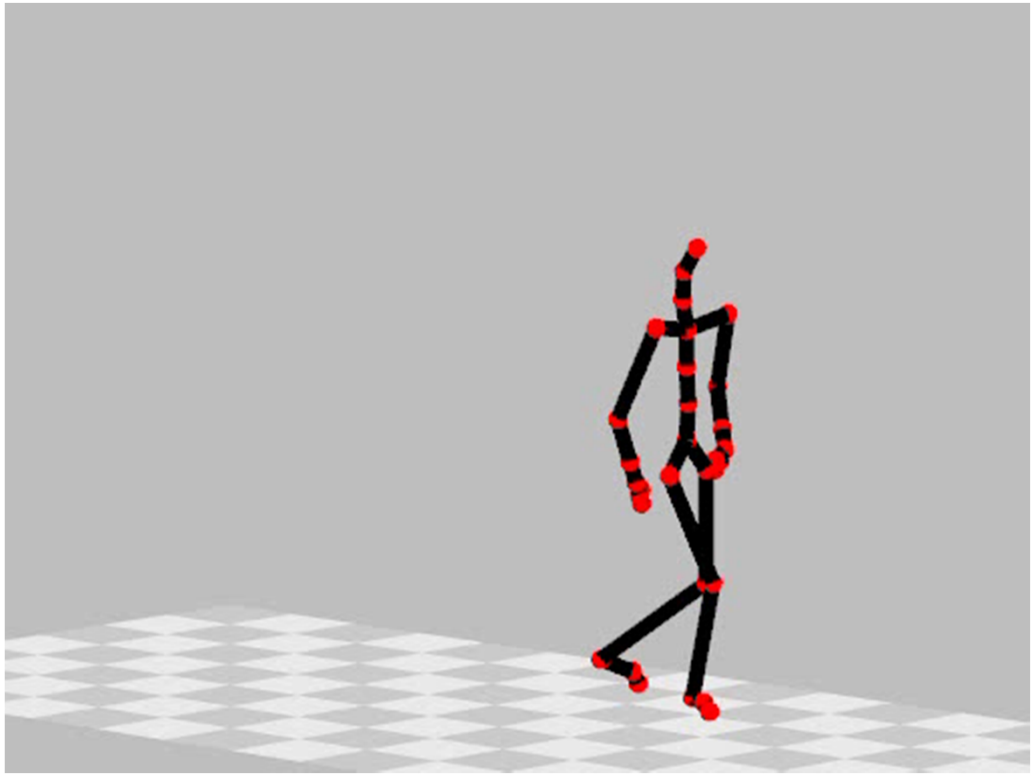


Overview

- Introduction
- Music Retrieval
- Motion Retrieval
- Music Structure Analysis
- Beat Tracking

Motion Capture Data

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

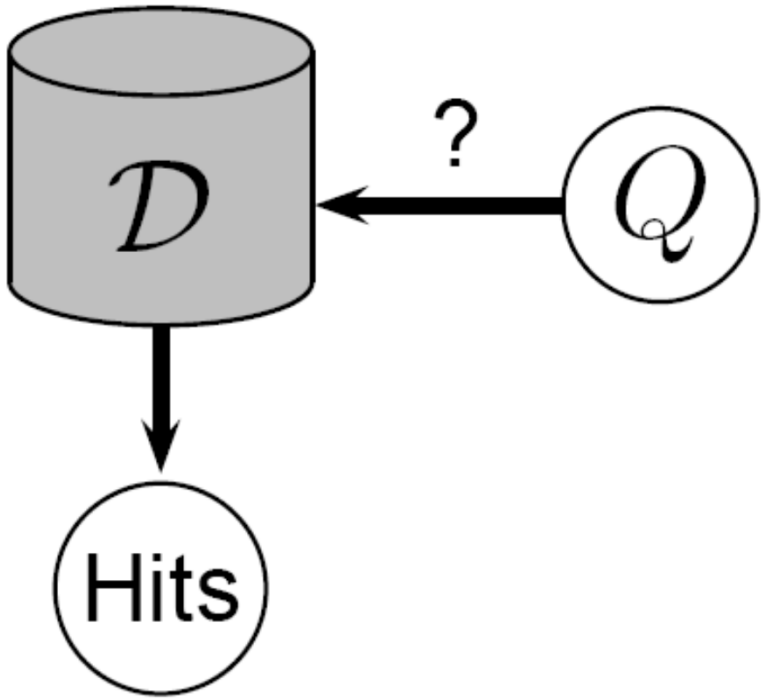


Motion Capture Data

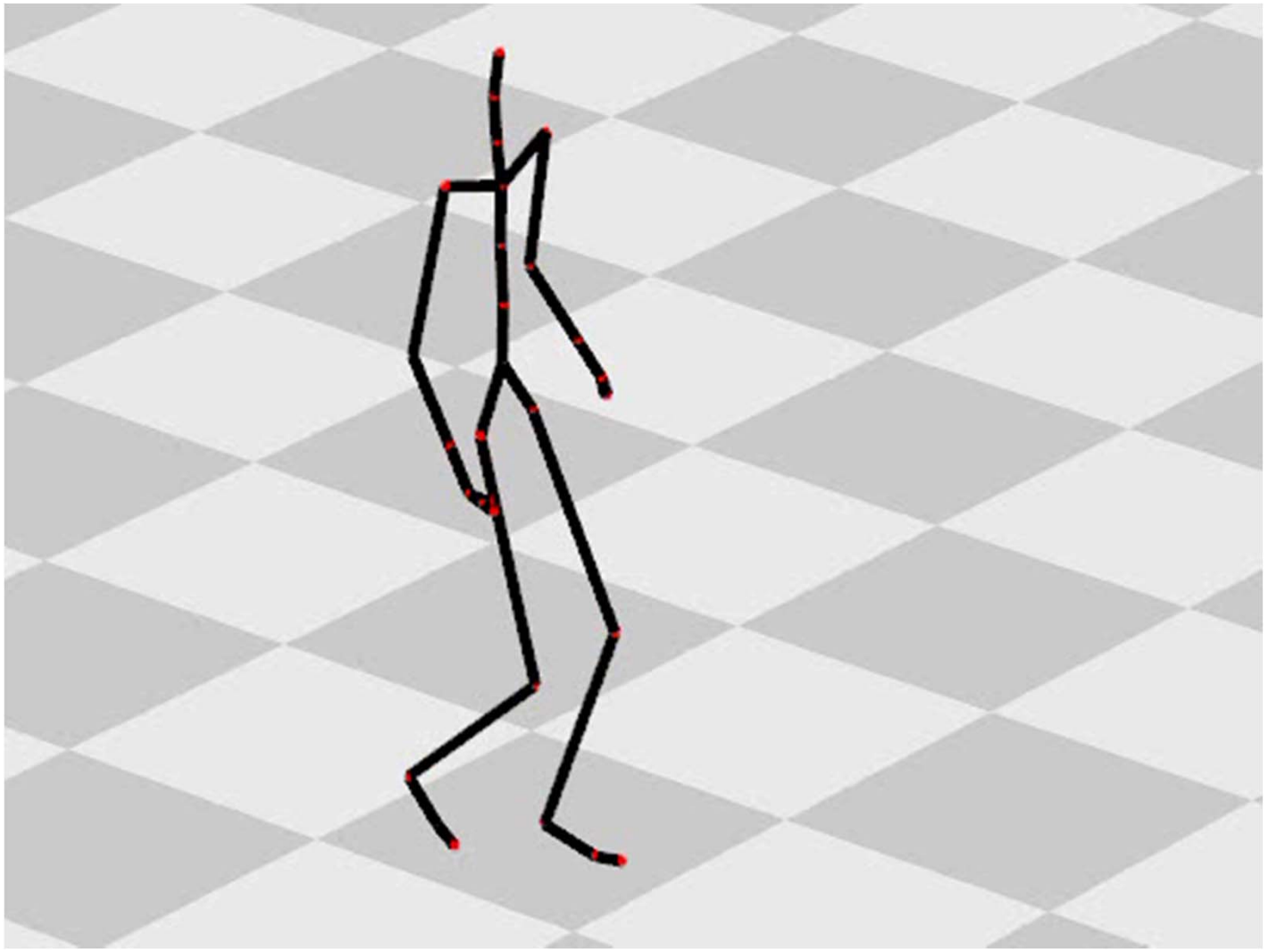
Optical System

Motion Retrieval

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

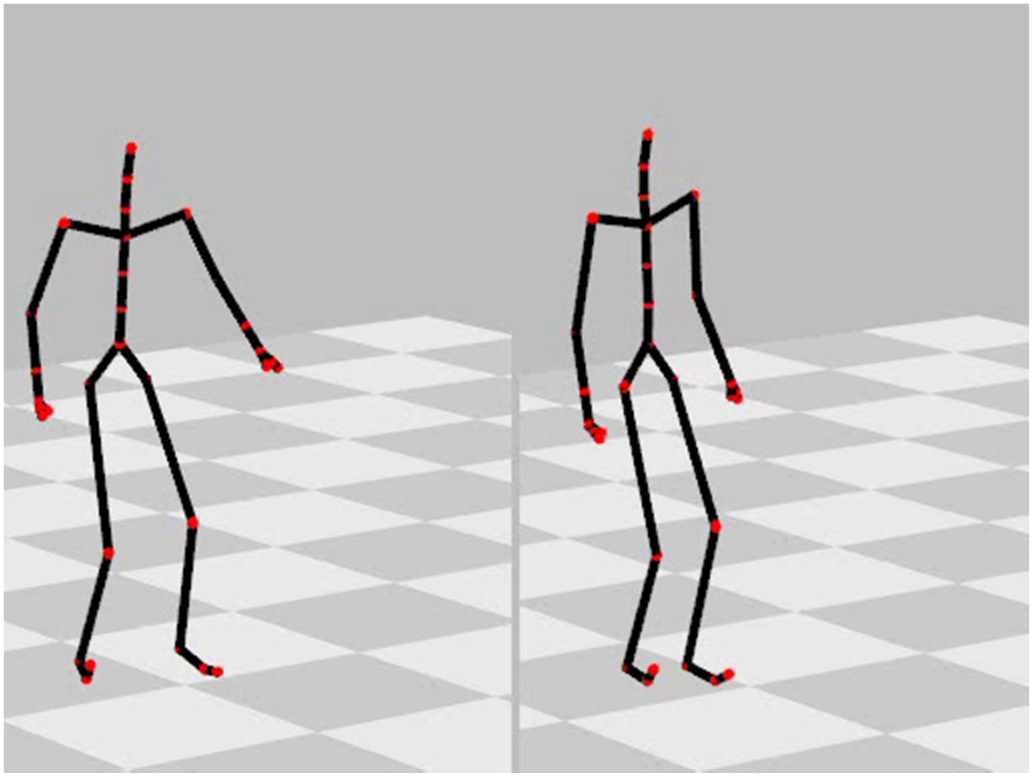


Motion Retrieval



Motion Retrieval

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



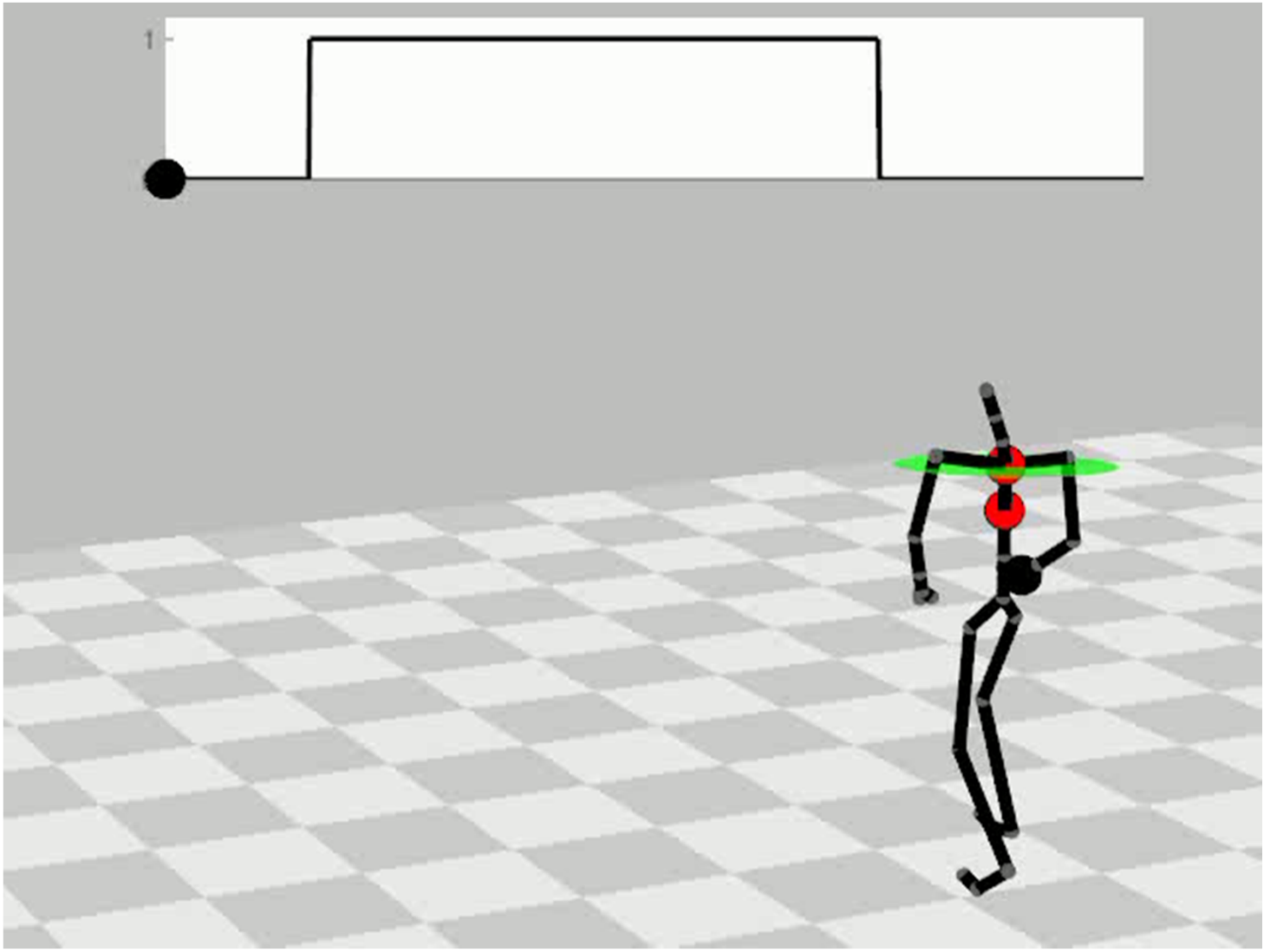
Relational Features

- Exploit knowledge of kinematic chain
- Express geometric relations of body parts
- Robust to motion variations

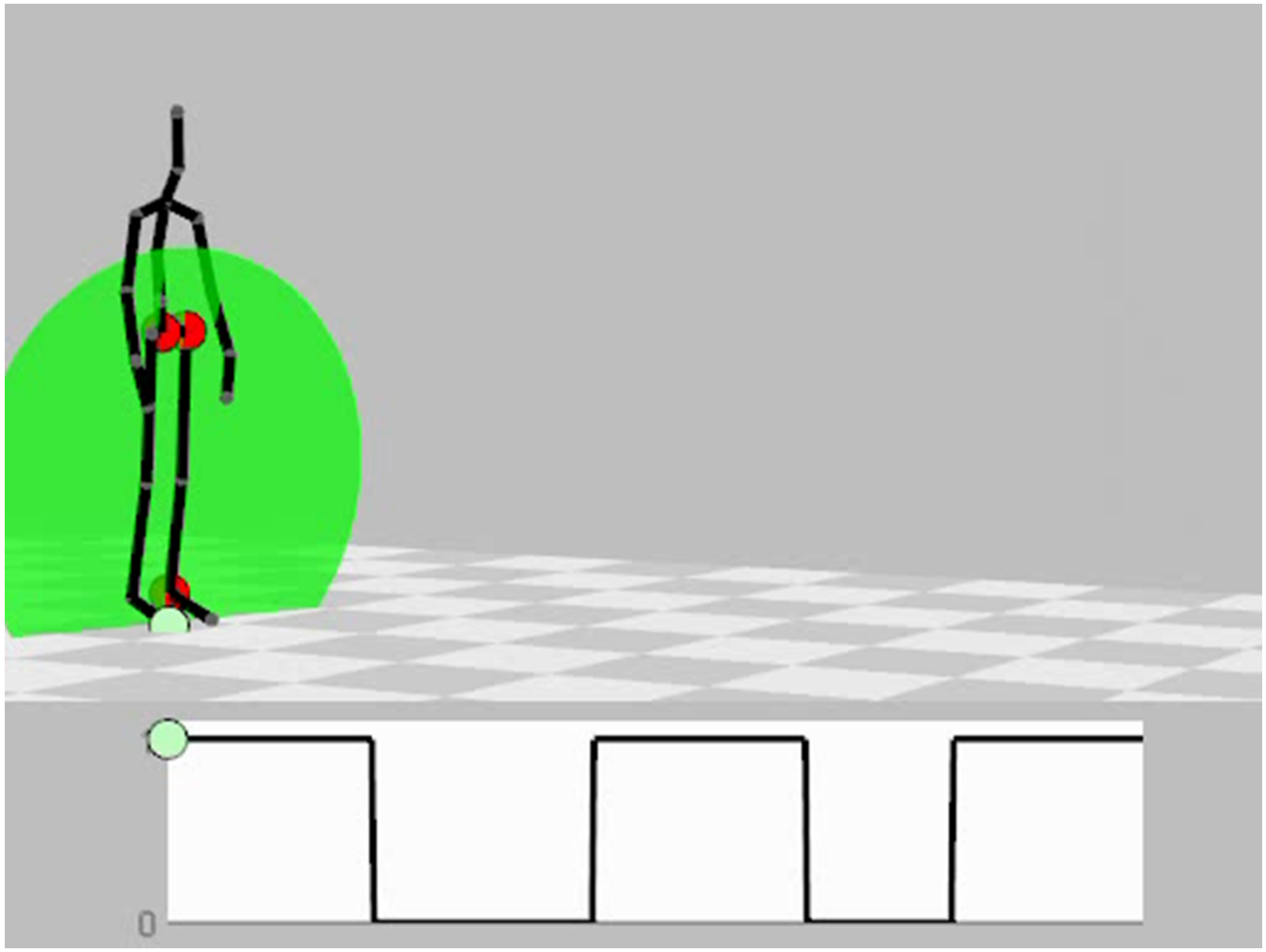
Meinard Müller, Tido Röder, and Michael Clausen
Efficient content-based retrieval of motion capture data.
ACM Transactions on Graphics (SIGGRAPH), vol. 24, pp. 677-685, 2005.

Meinard Müller and Tido Röder
Motion templates for automatic classification and retrieval of motion capture data.
Proceedings of the 2006 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA), Vienna, Austria, pp. 137-146, 2006.

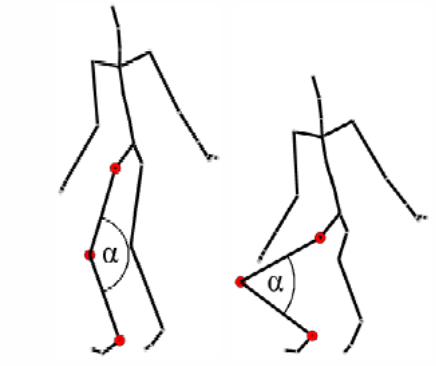
Relational Features



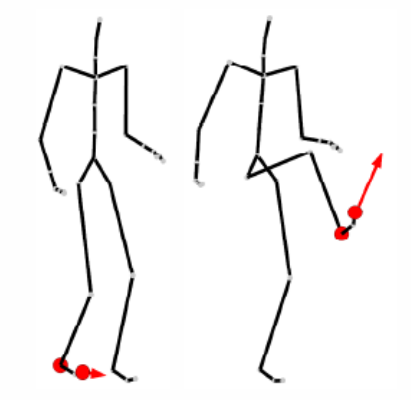
Relational Features



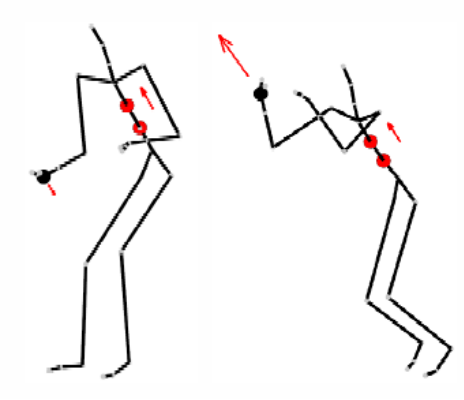
Relational Features



Right knee bent?

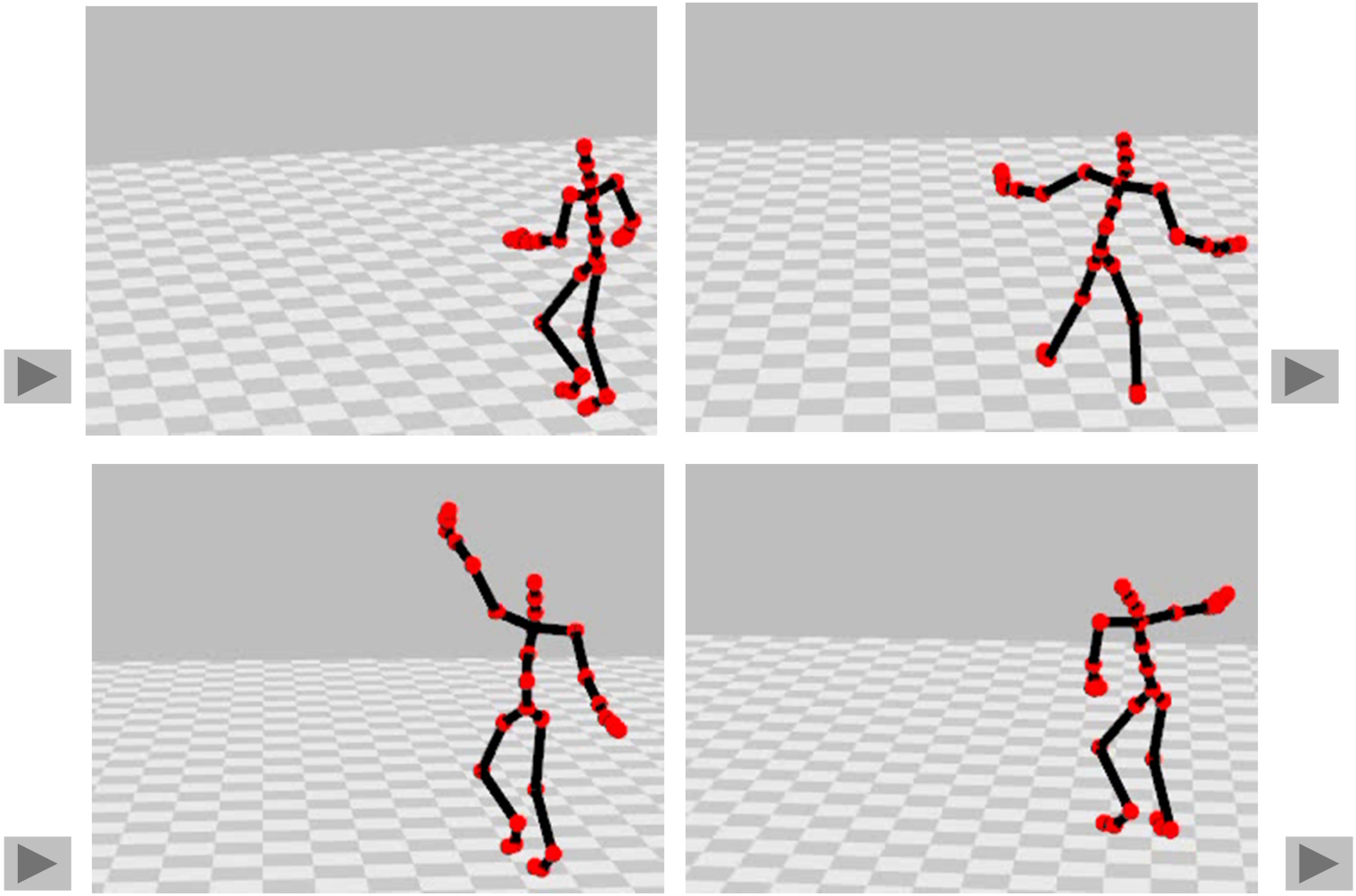


Right foot fast?

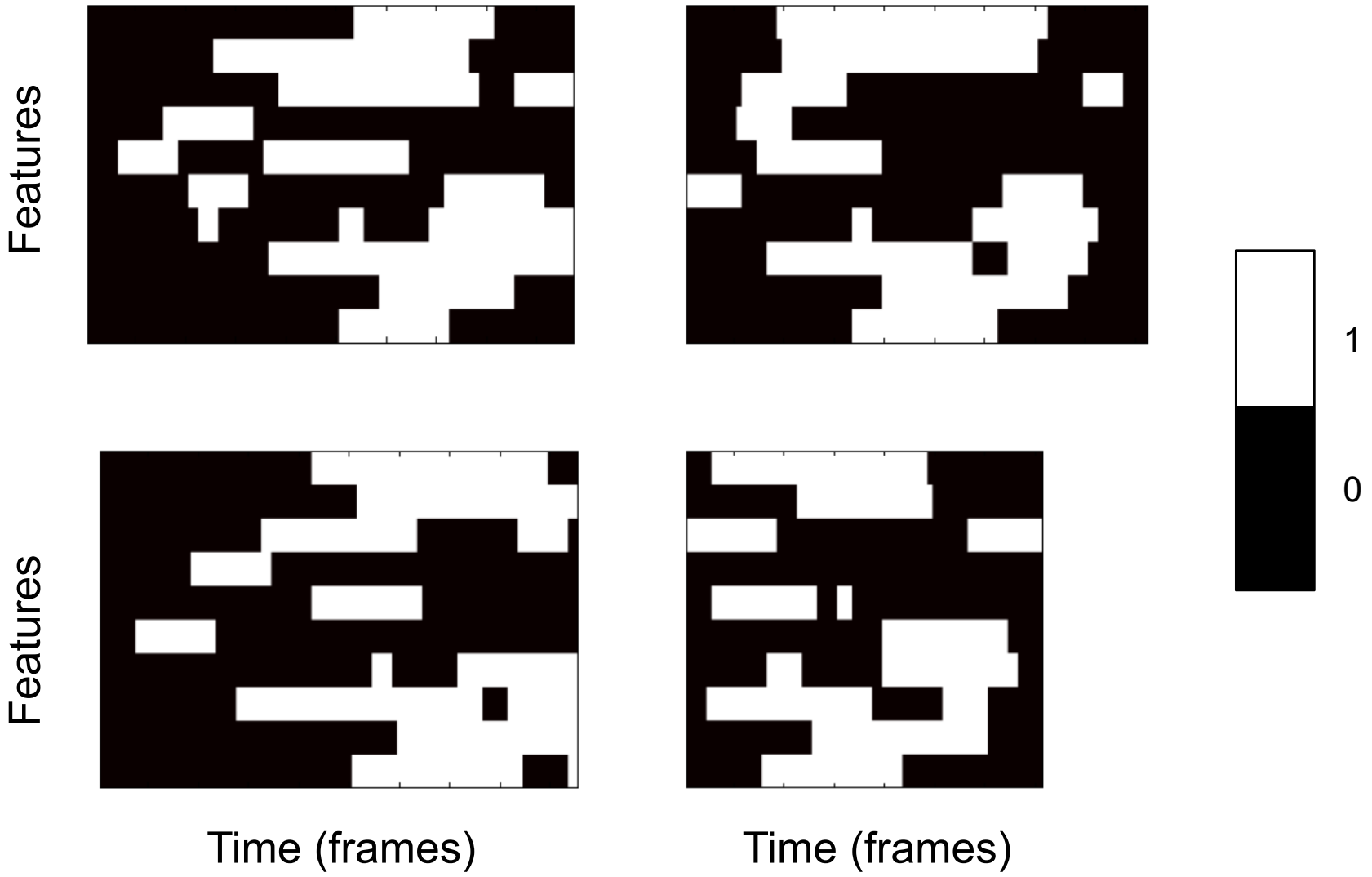


Right hand moving upwards?

Motion Templates (MT)

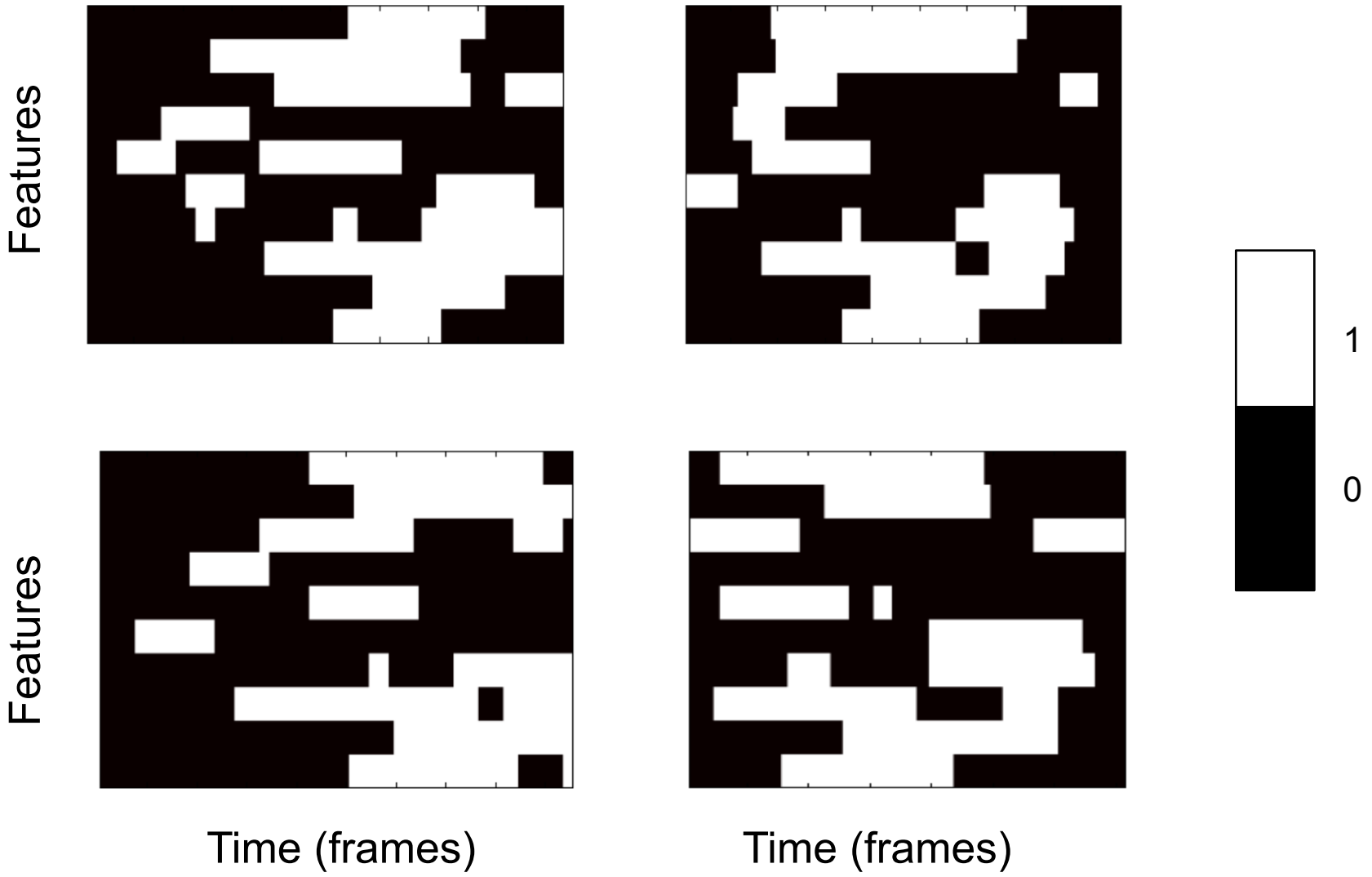


Motion Templates (MT)



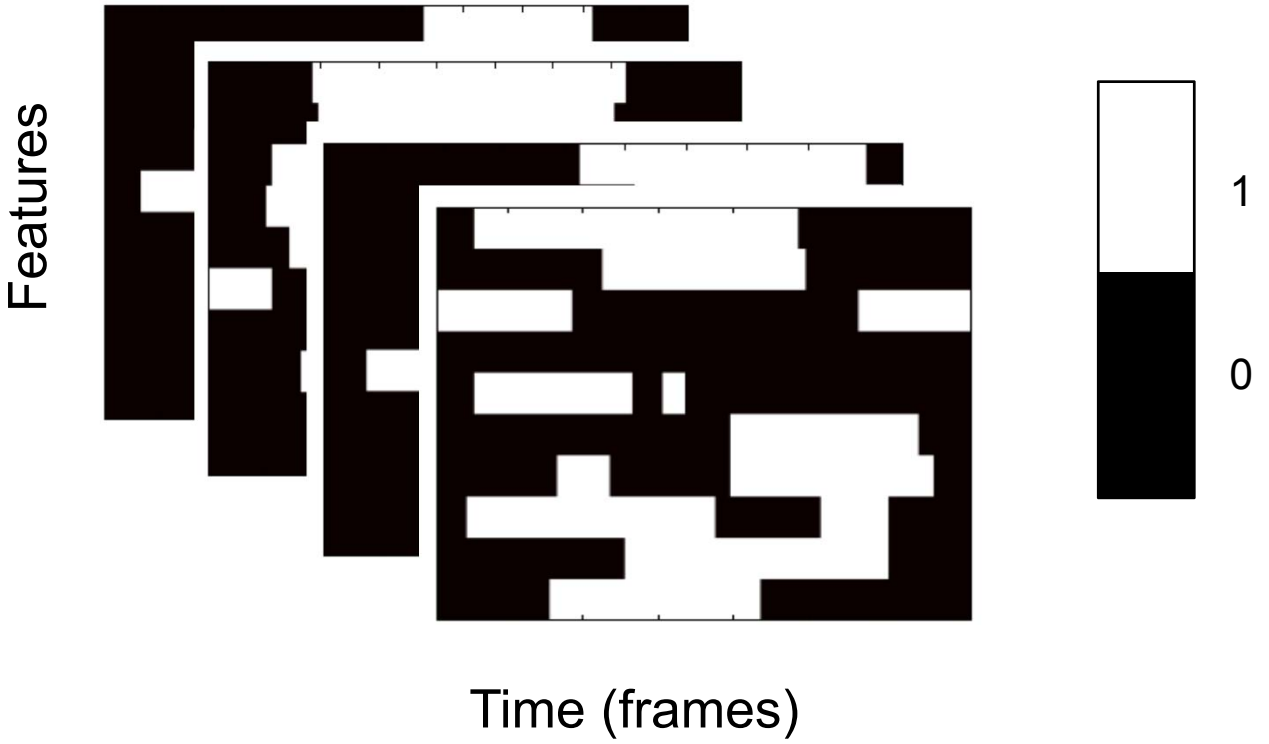
Motion Templates (MT)

Temporal alignment



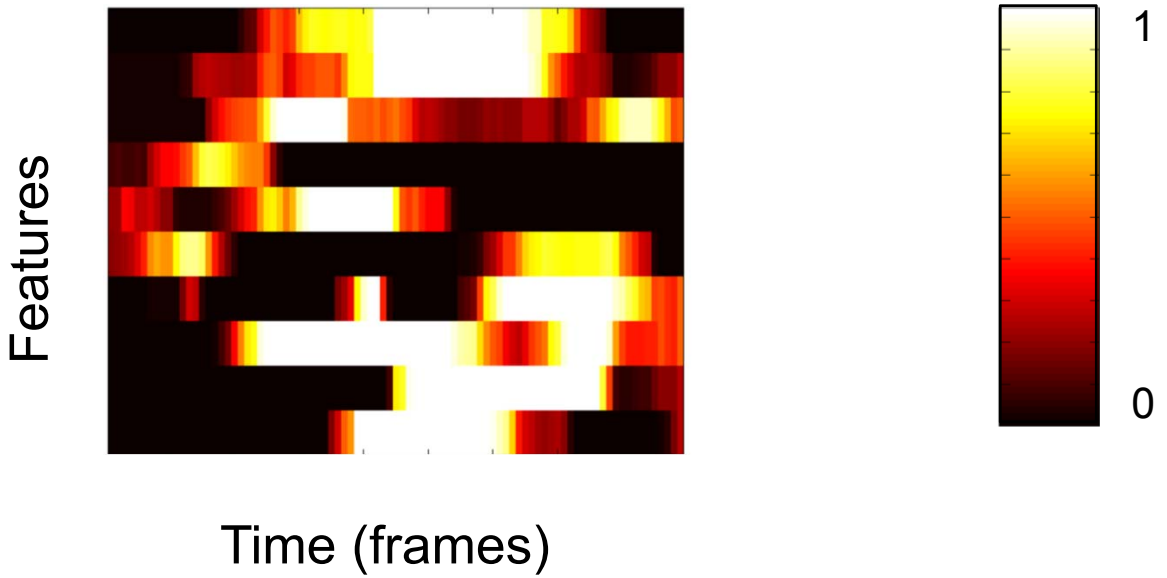
Motion Templates (MT)

Superimpose templates

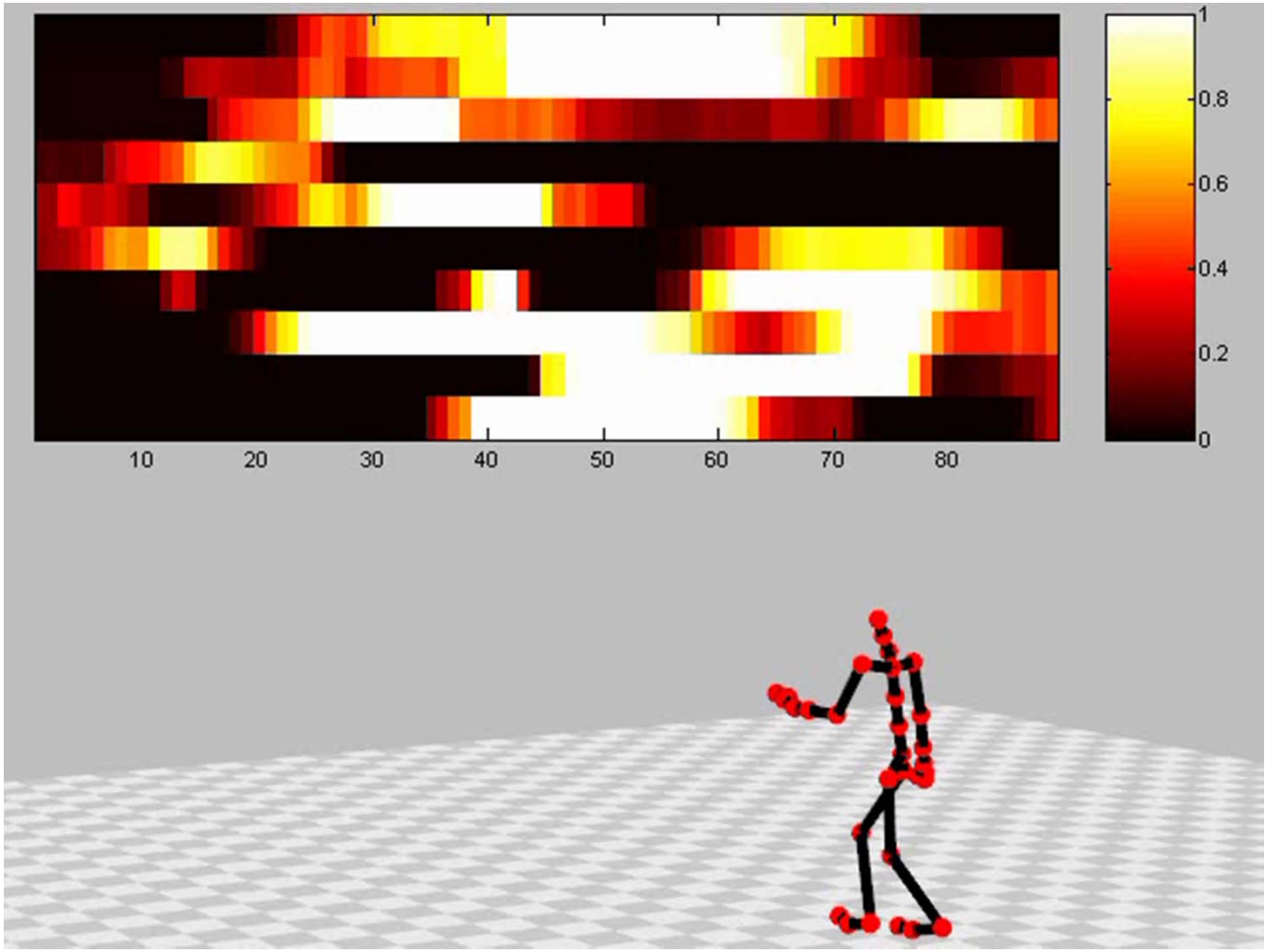


Motion Templates (MT)

Compute average

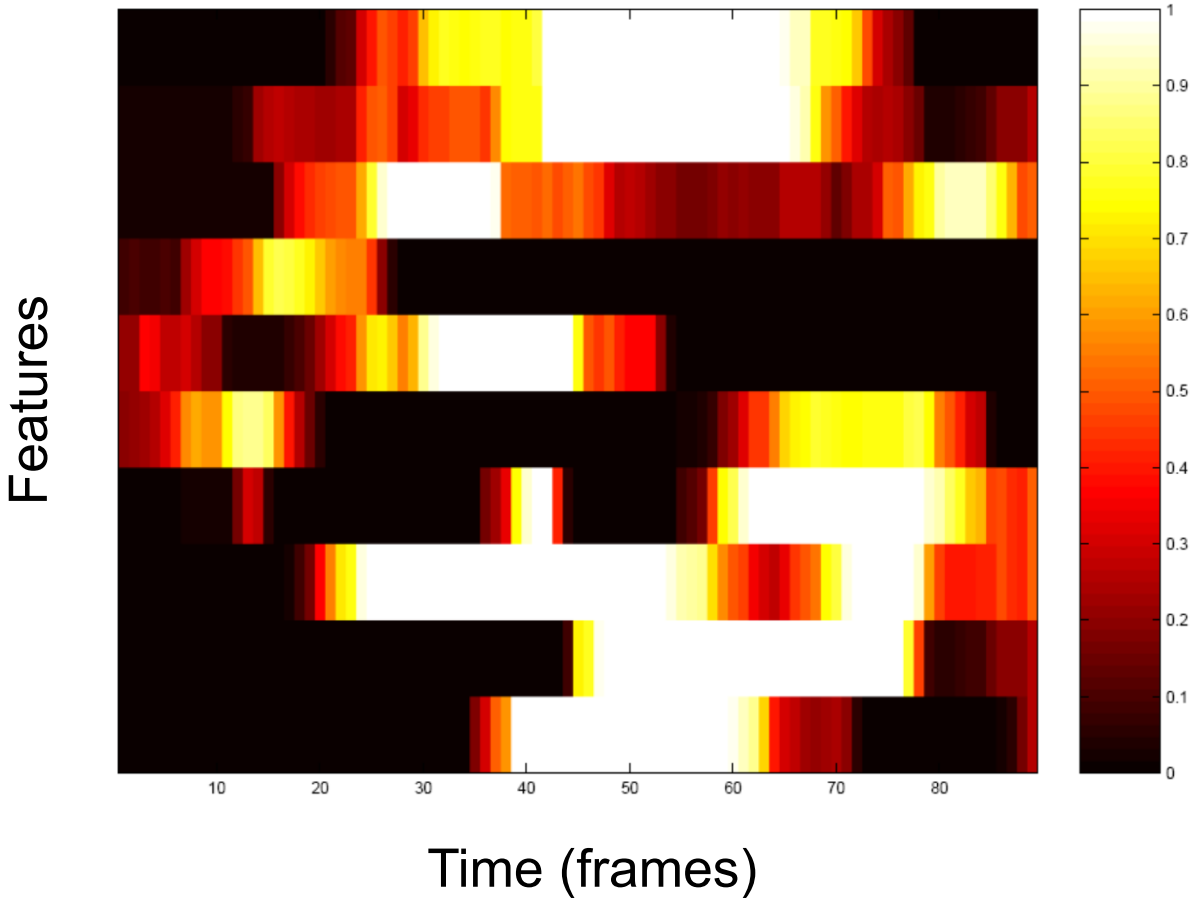


Motion Templates (MT)



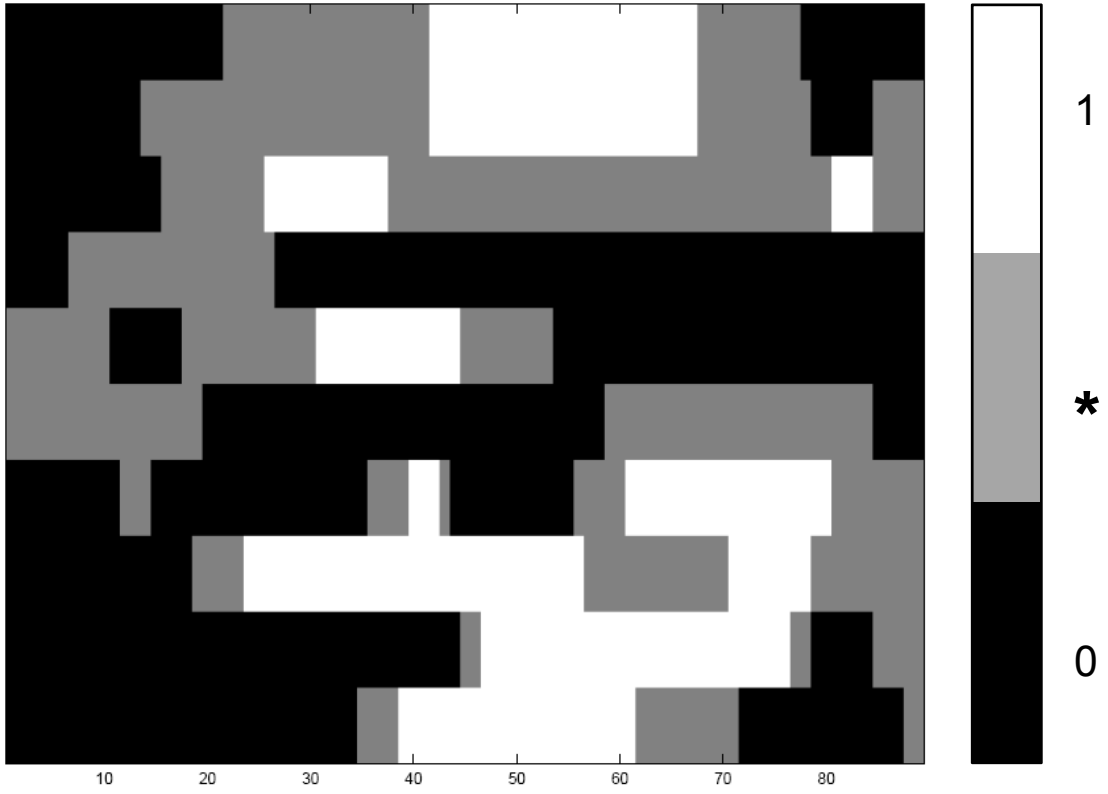
Motion Templates (MT)

Average template



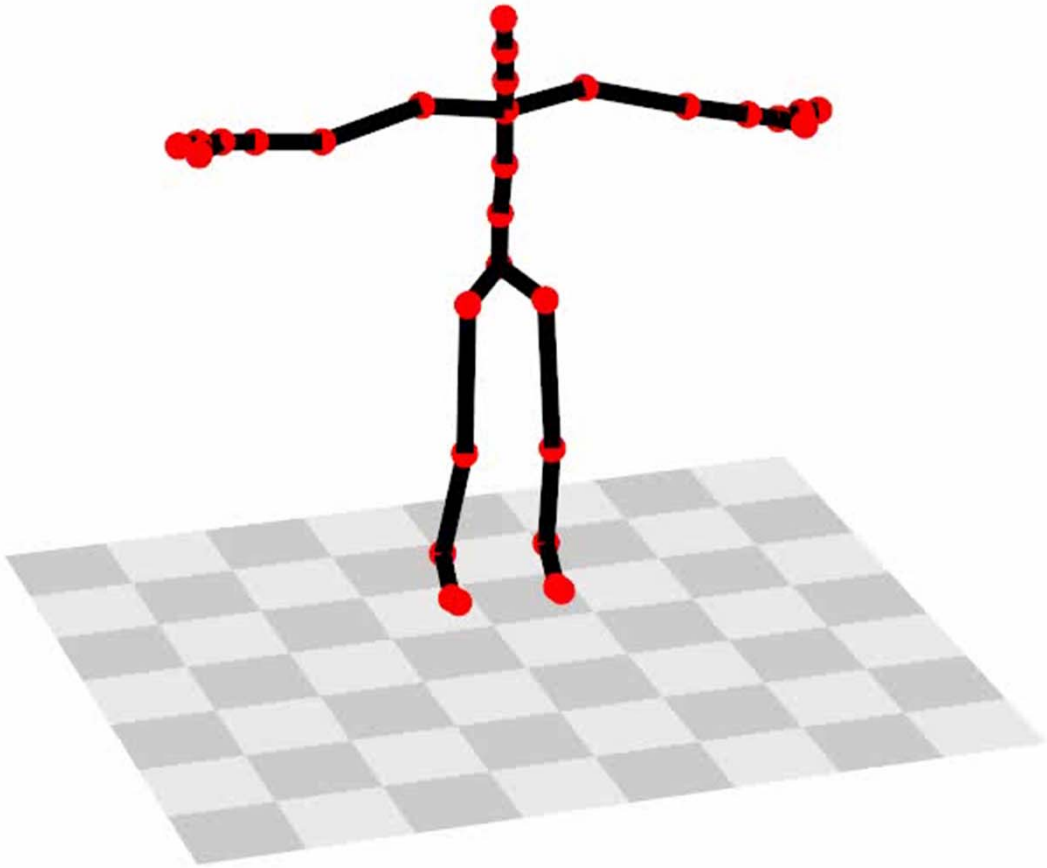
Motion Templates (MT)

Quantized template

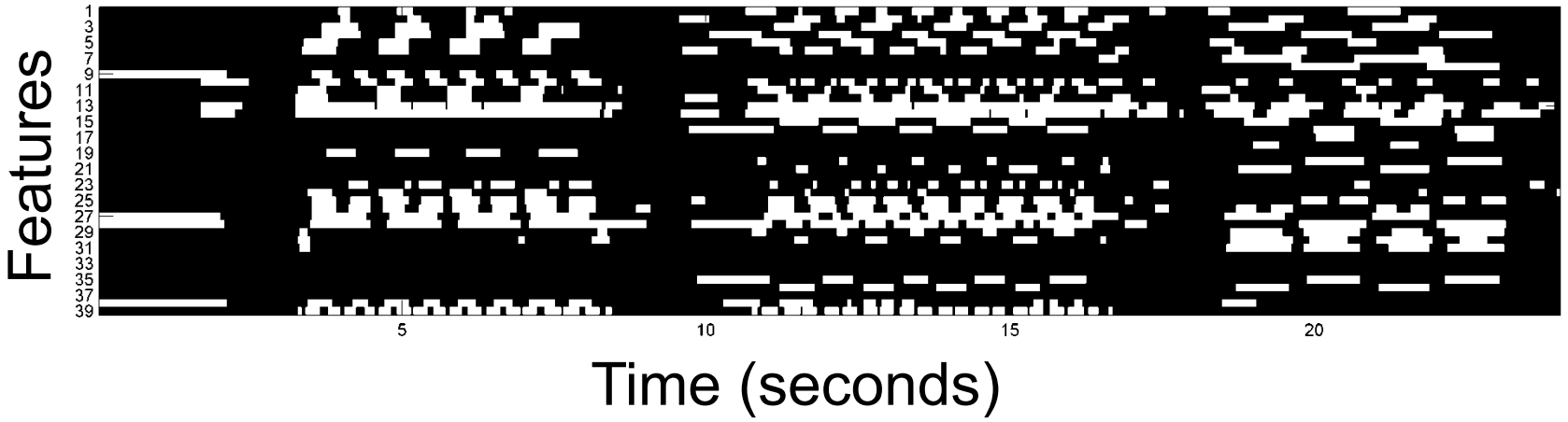


- Gray areas indicate inconsistencies / variations
- Achieve invariance by disregarding gray areas

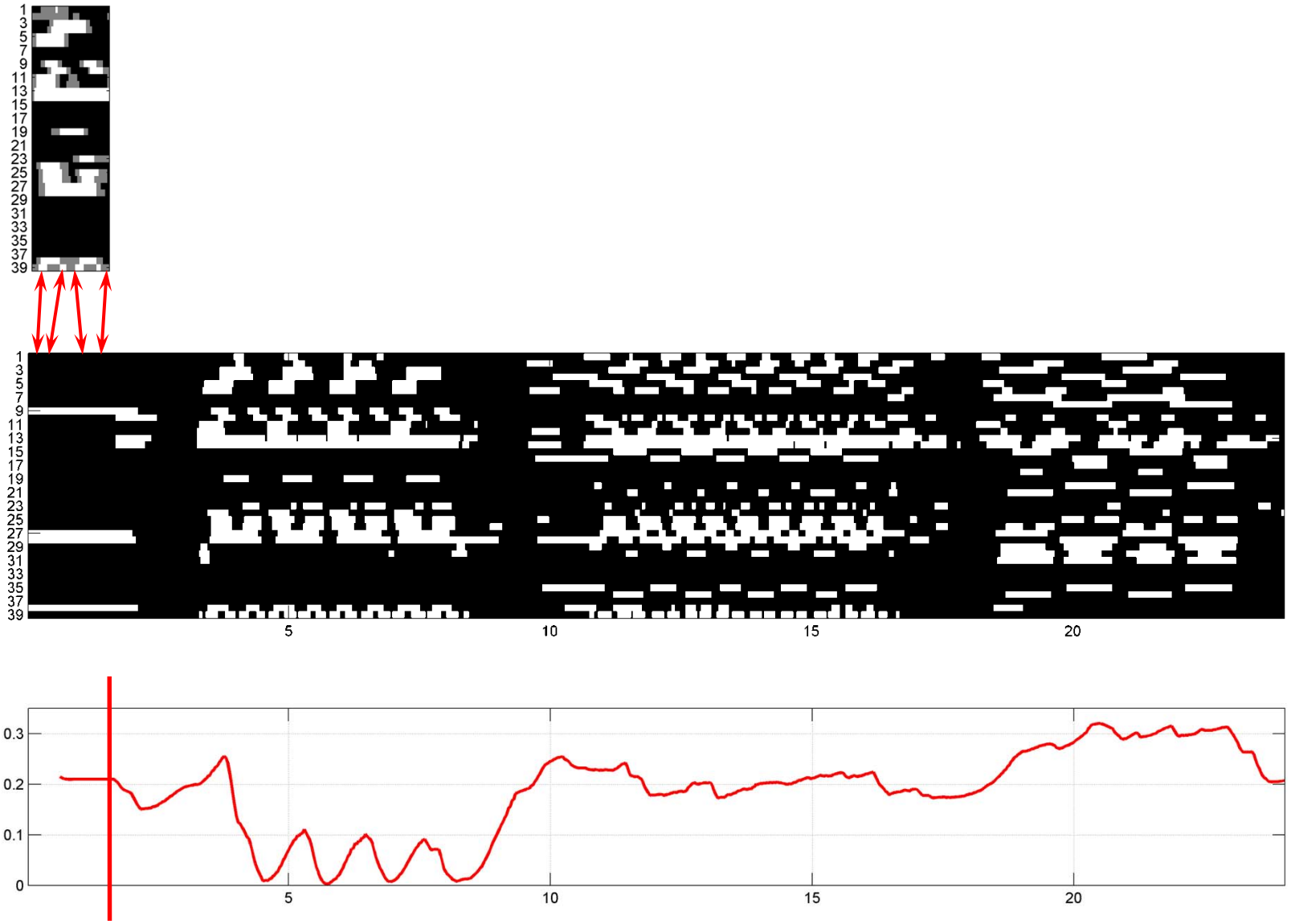
MT-based Motion Retrieval



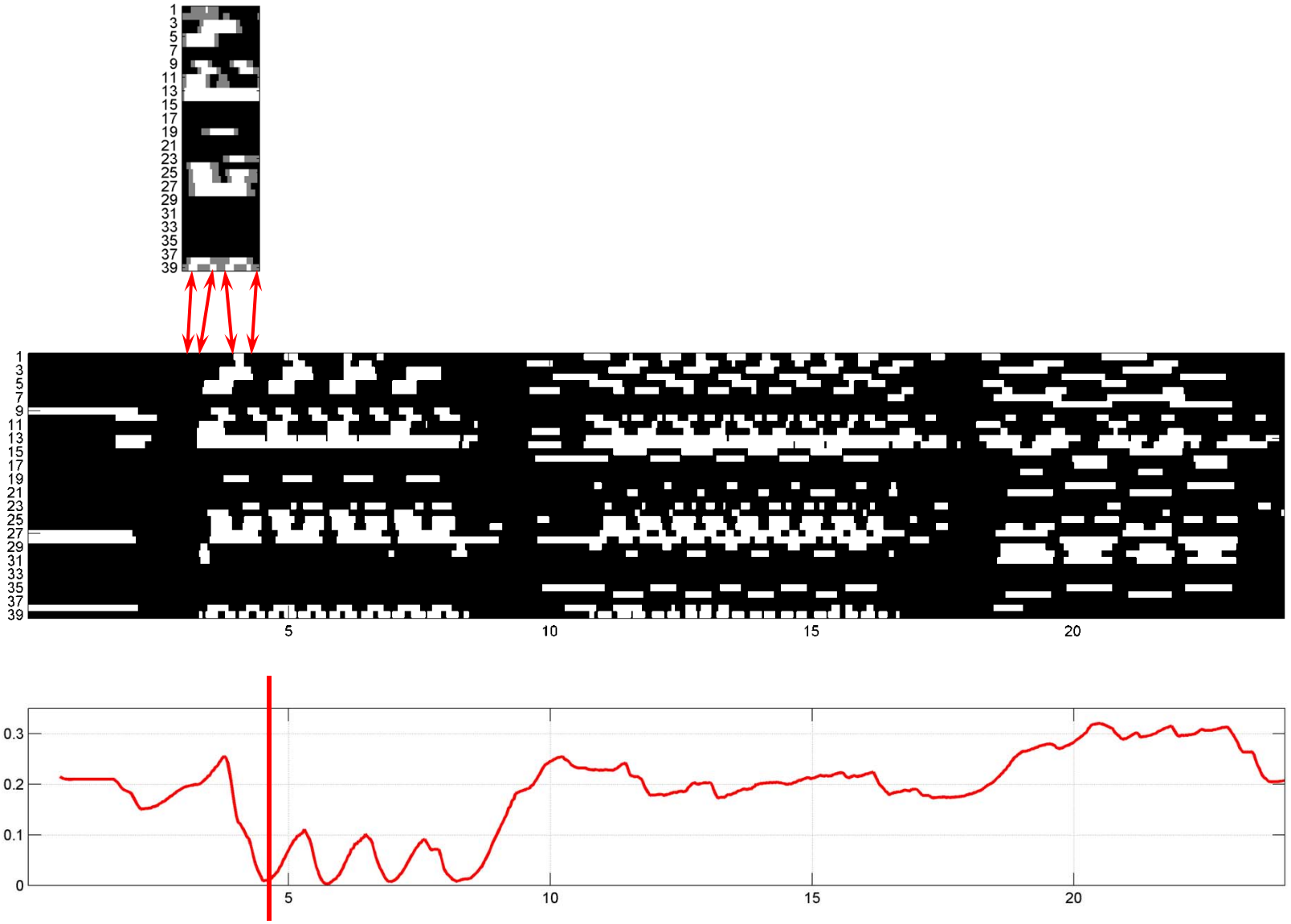
MT-based Motion Retrieval



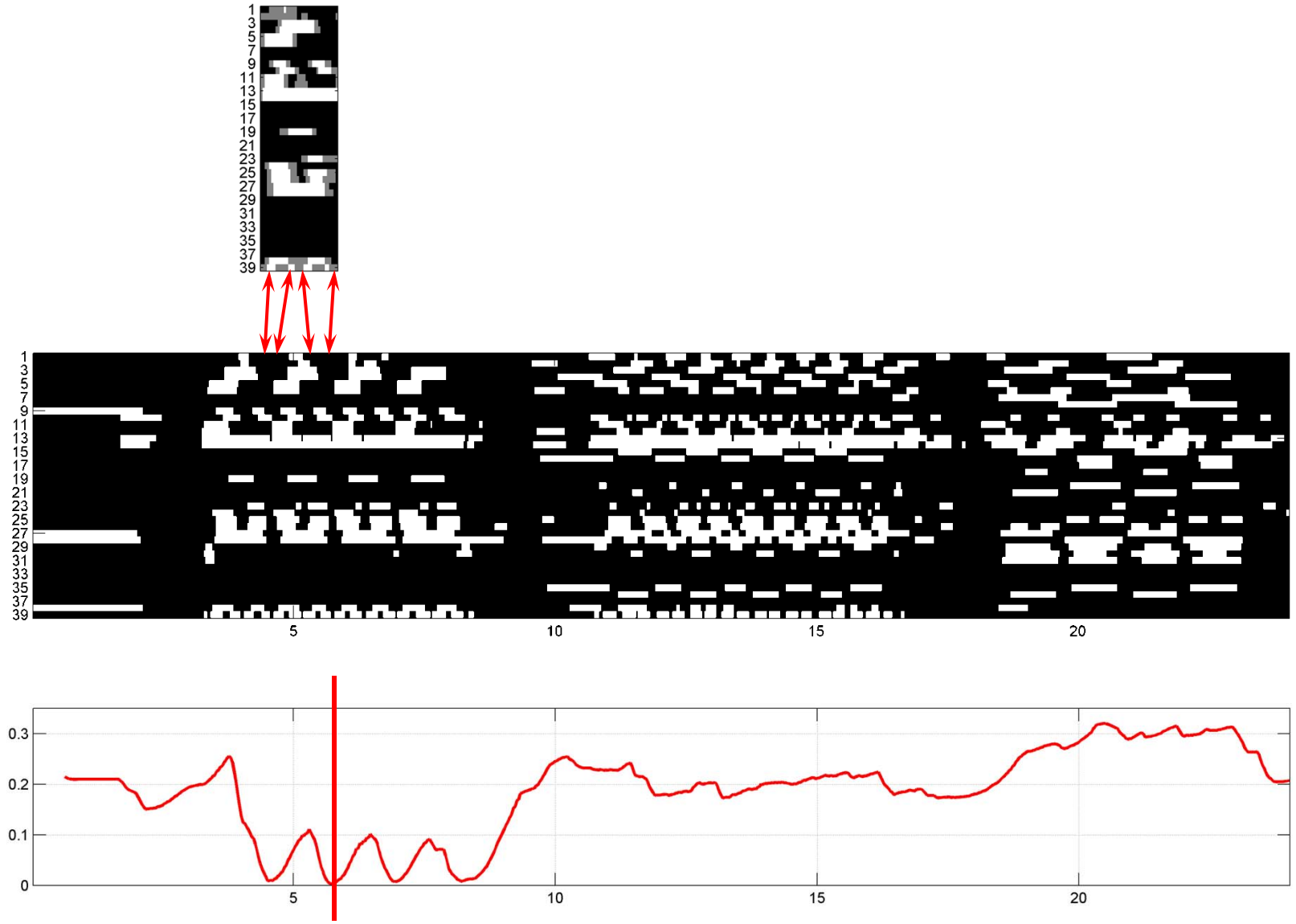
MT-based Motion Retrieval: Jumping Jack



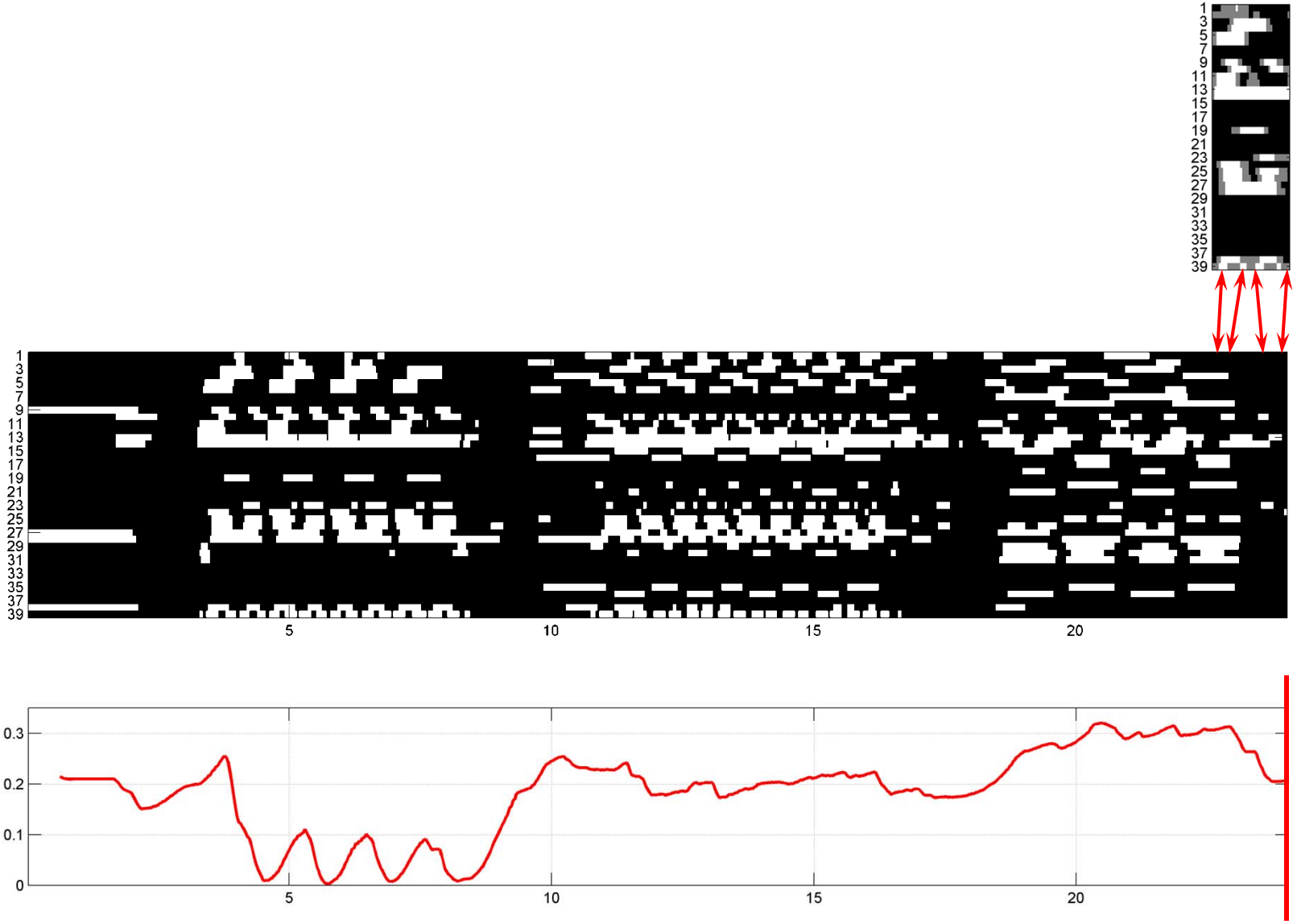
MT-based Motion Retrieval: Jumping Jack



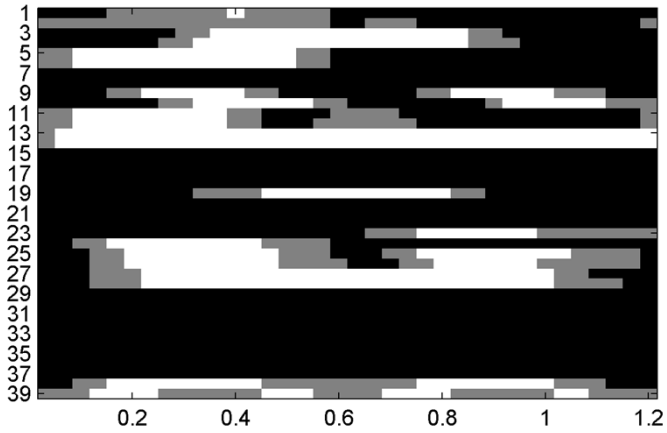
MT-based Motion Retrieval: Jumping Jack

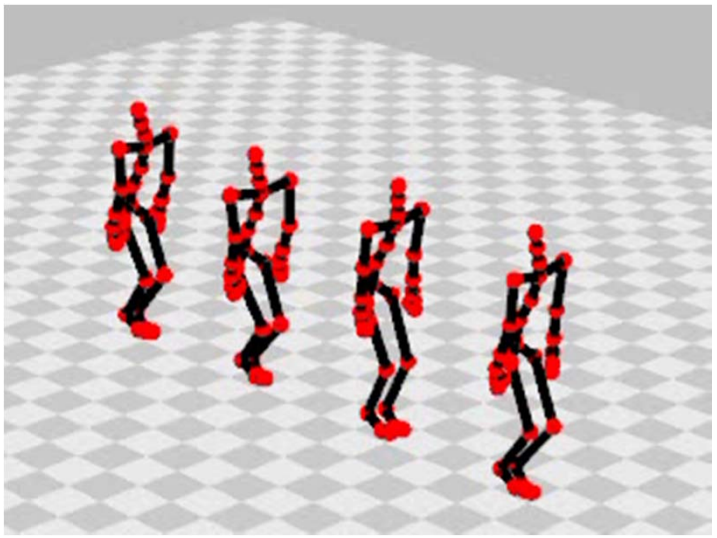


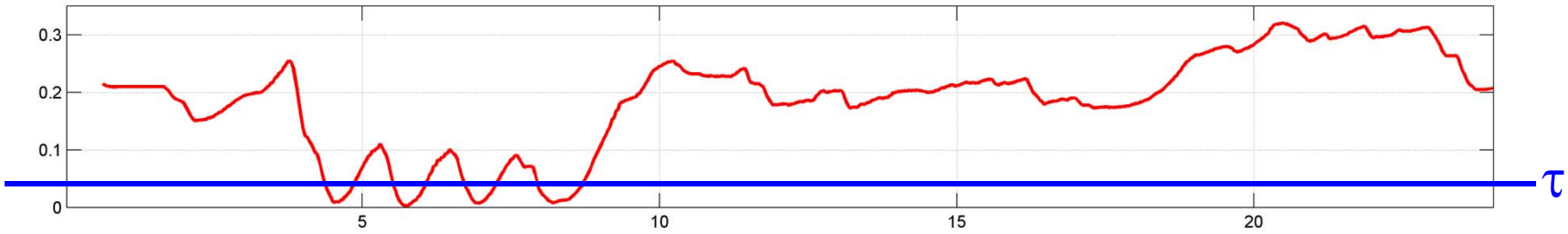
MT-based Motion Retrieval: Jumping Jack



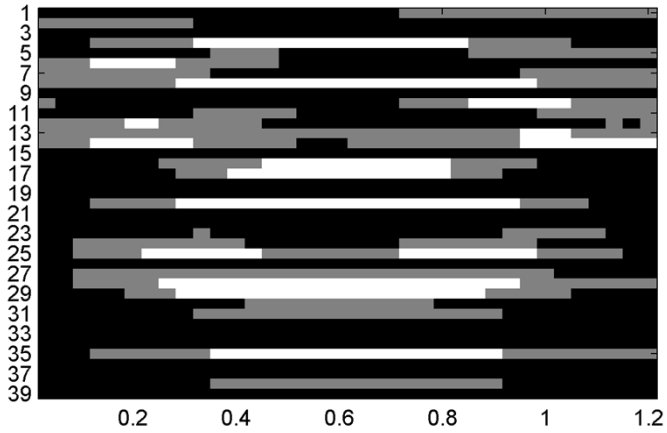
MT-based Motion Retrieval: Jumping Jack

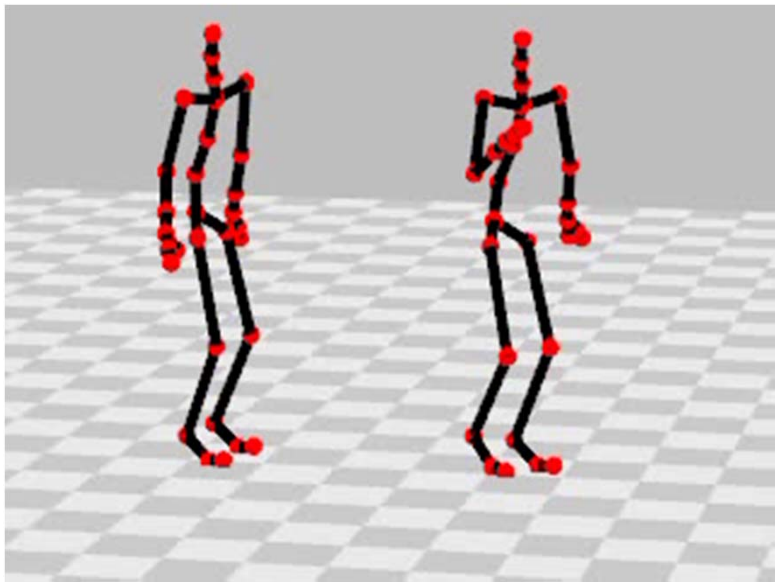


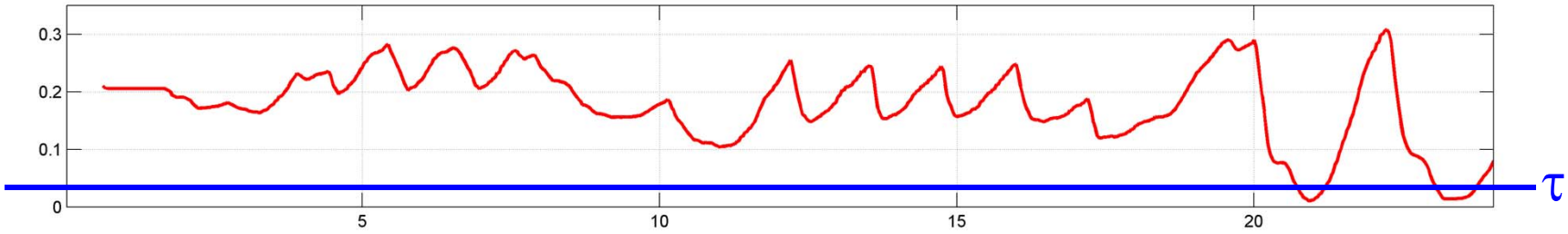




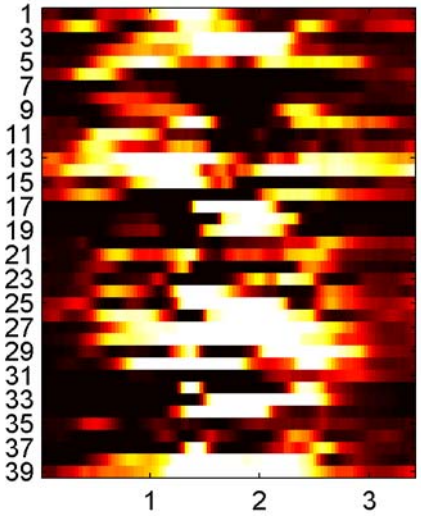
MT-based Motion Retrieval: Elbow-To-Knee



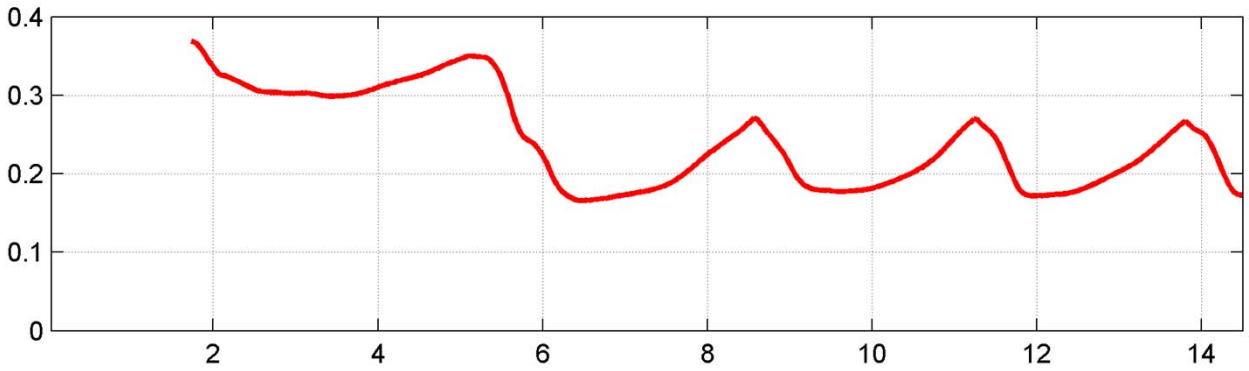


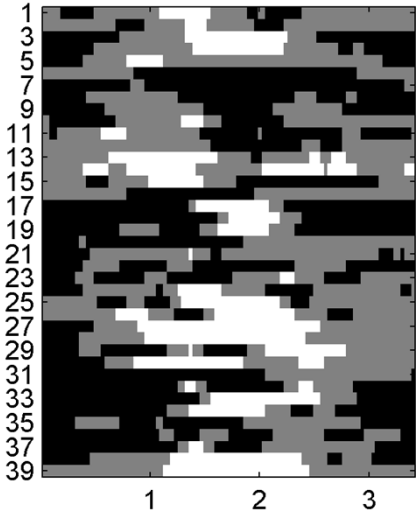


MT-based Motion Retrieval: Cartwheel

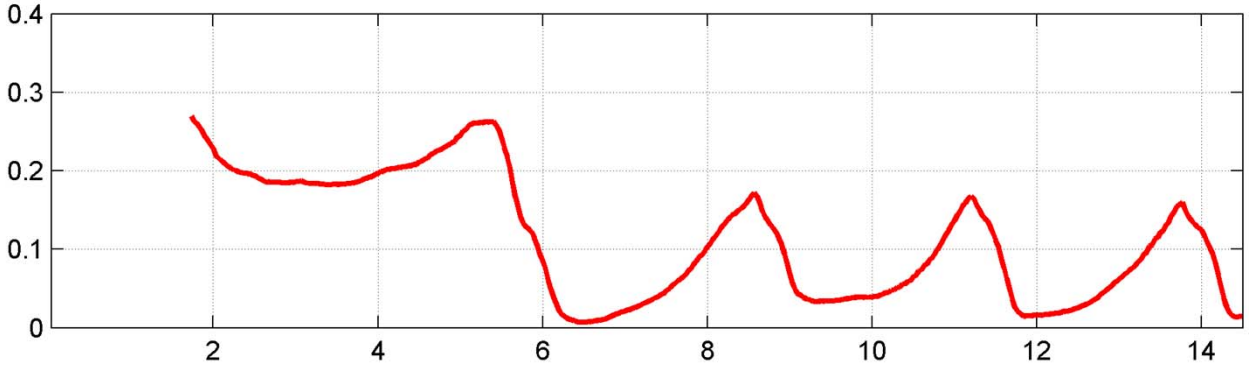


Matching curve using average MT

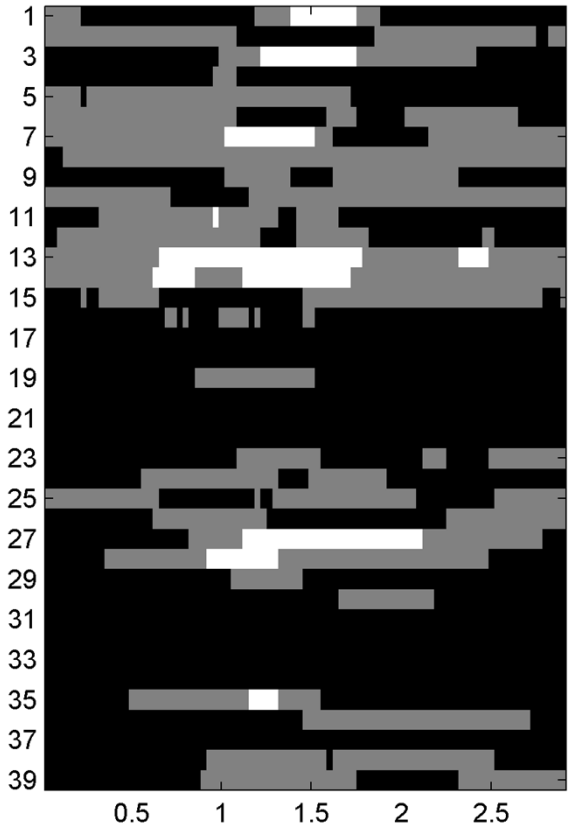


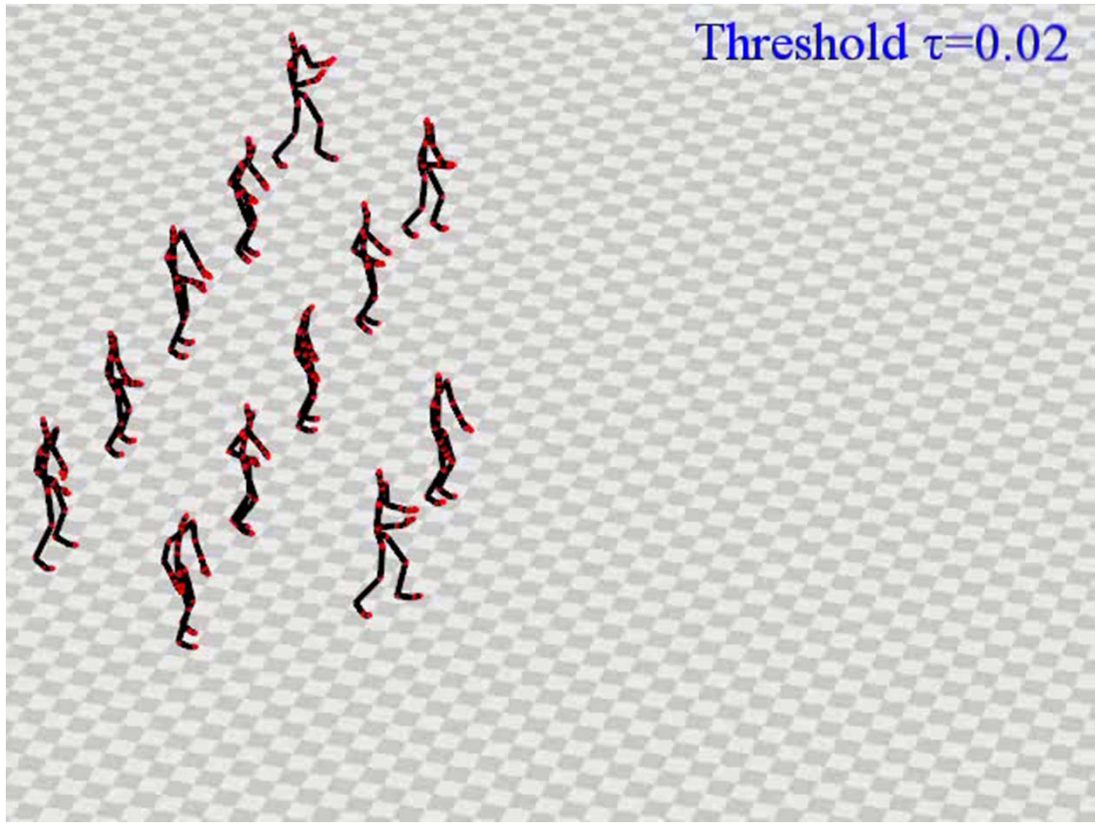


Matching curve blending out variations

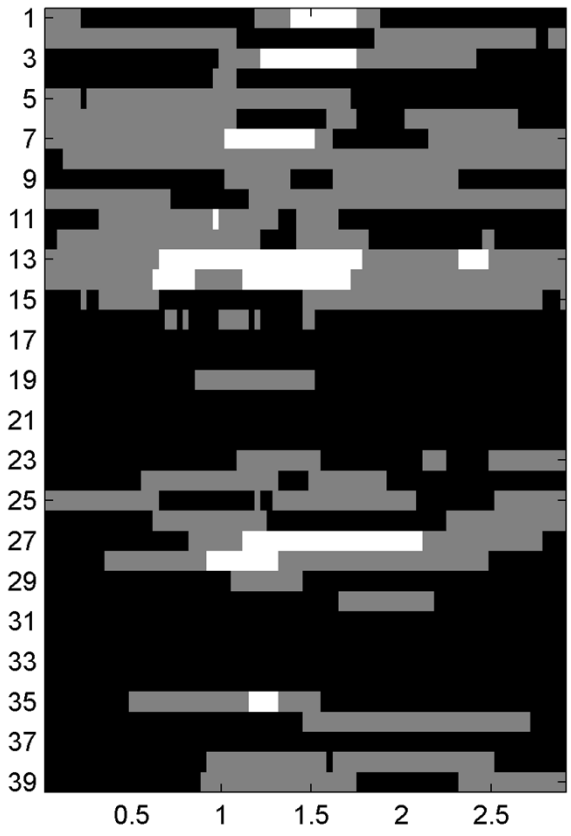


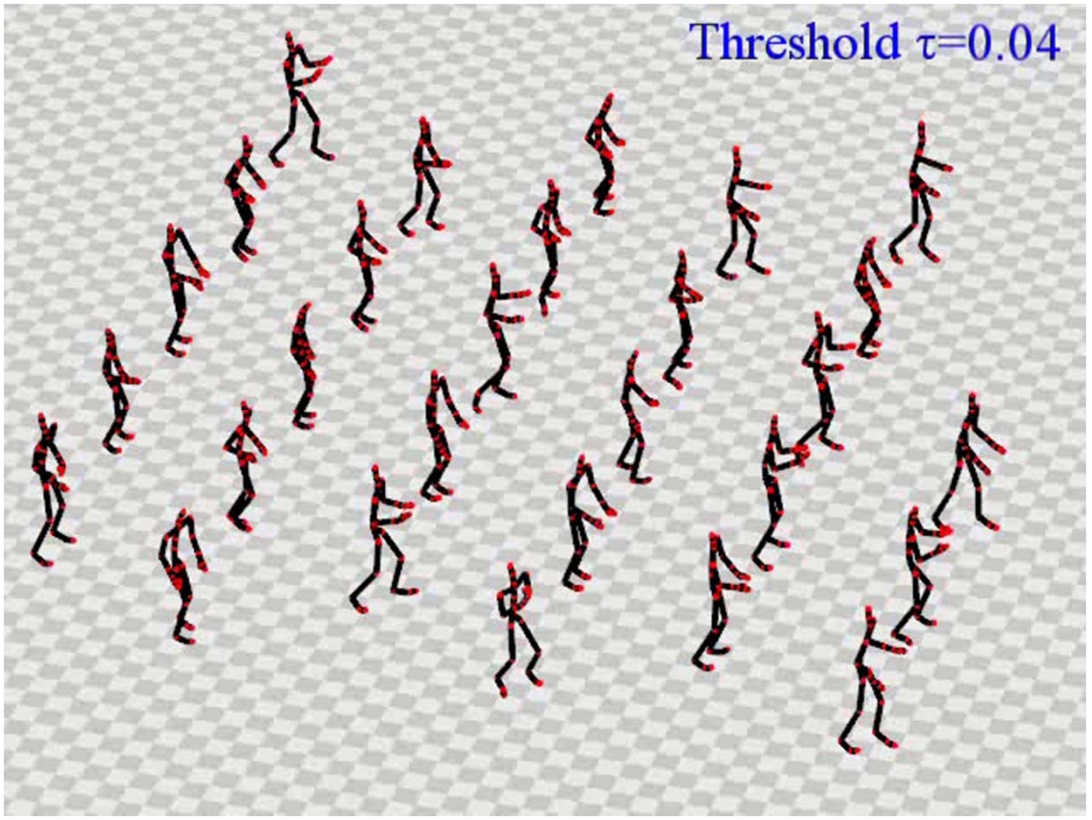
MT-based Motion Retrieval: Throw



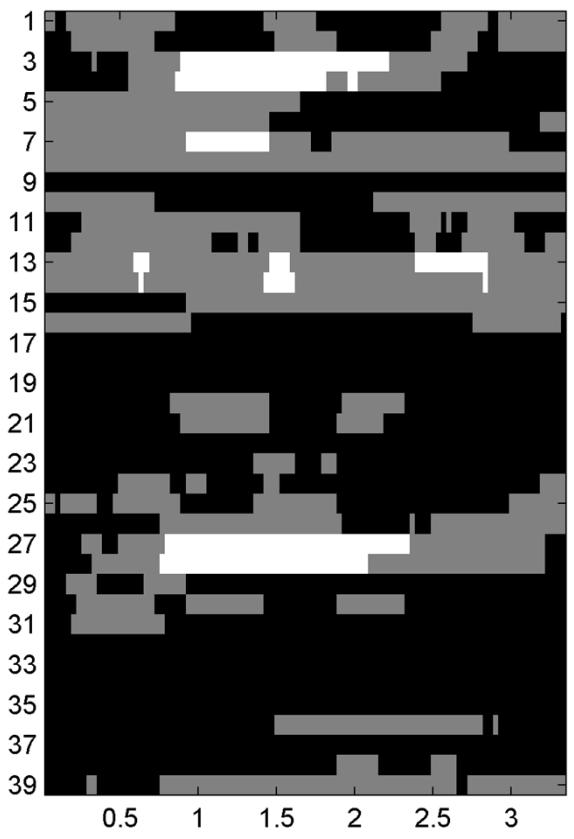


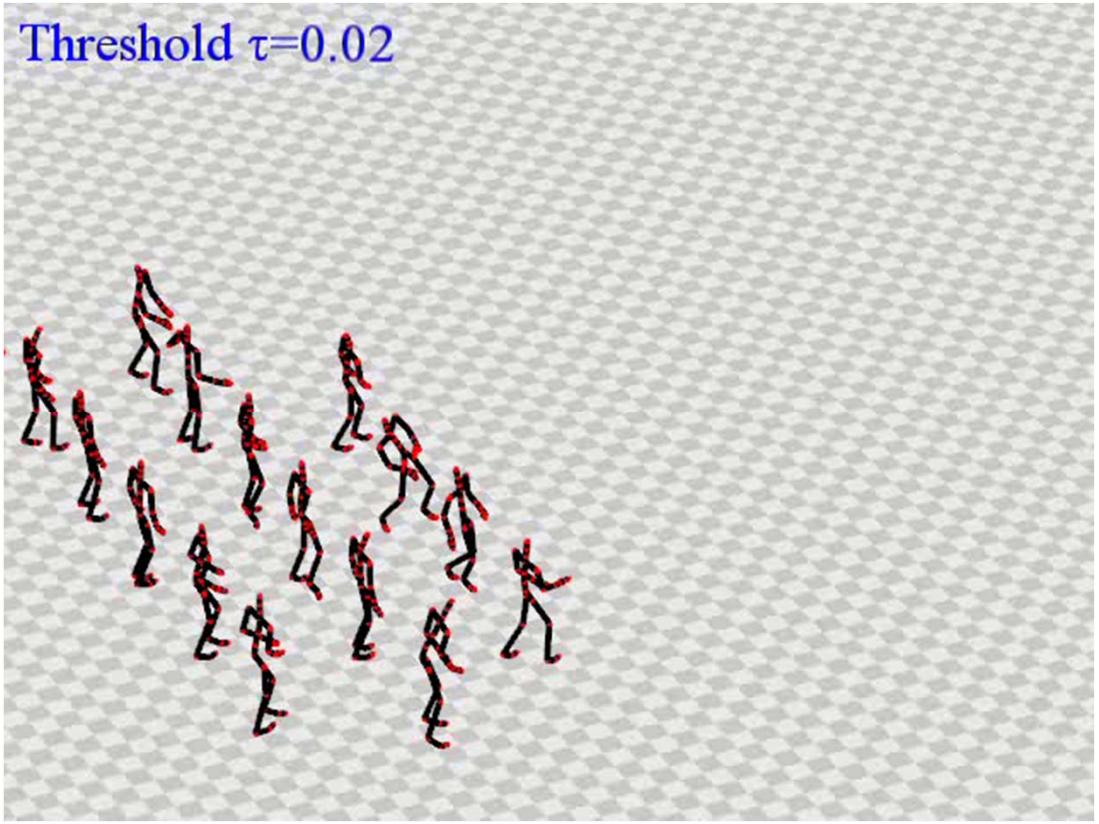
MT-based Motion Retrieval: Throw



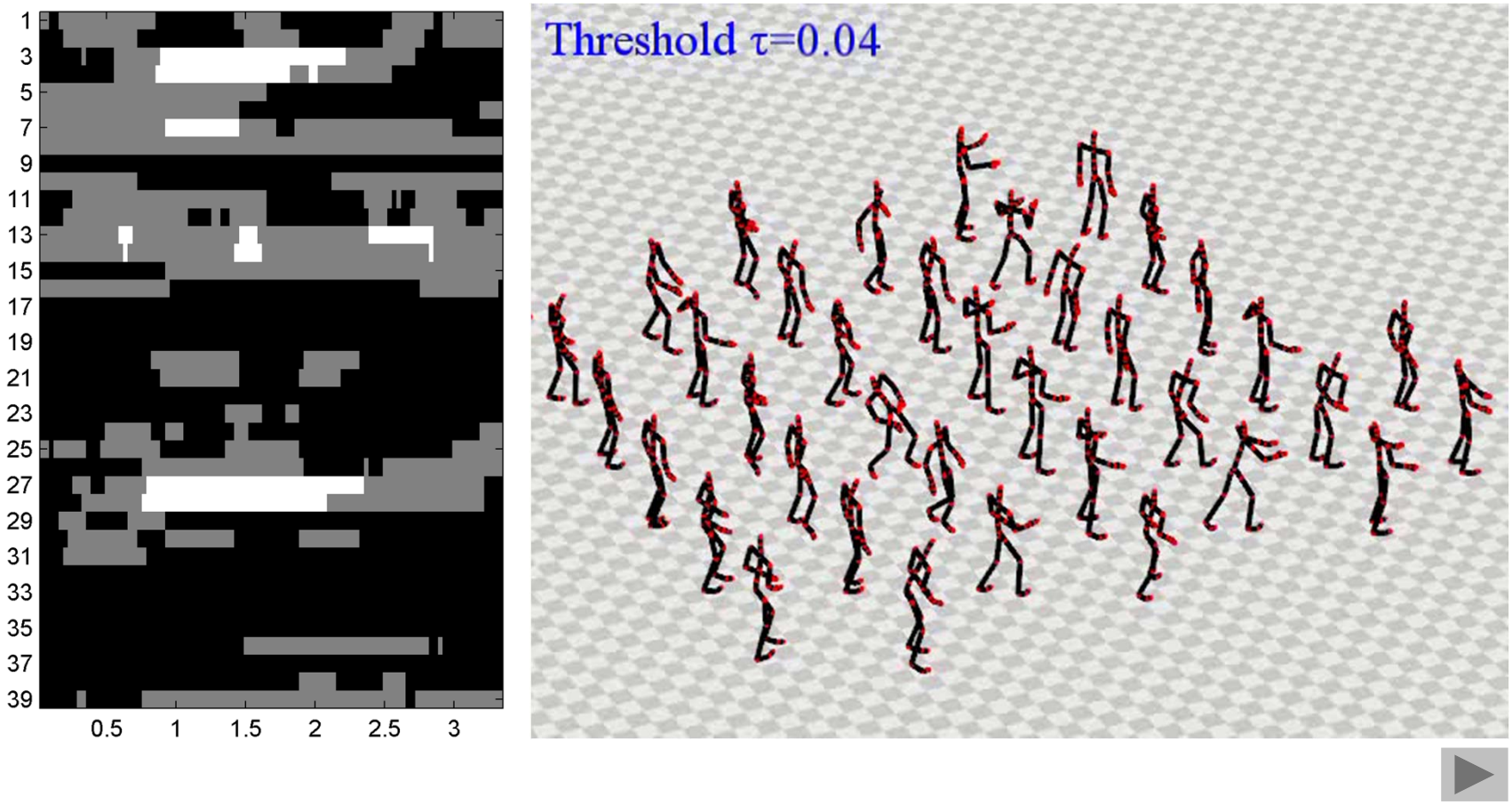


MT-based Motion Retrieval: Basketball

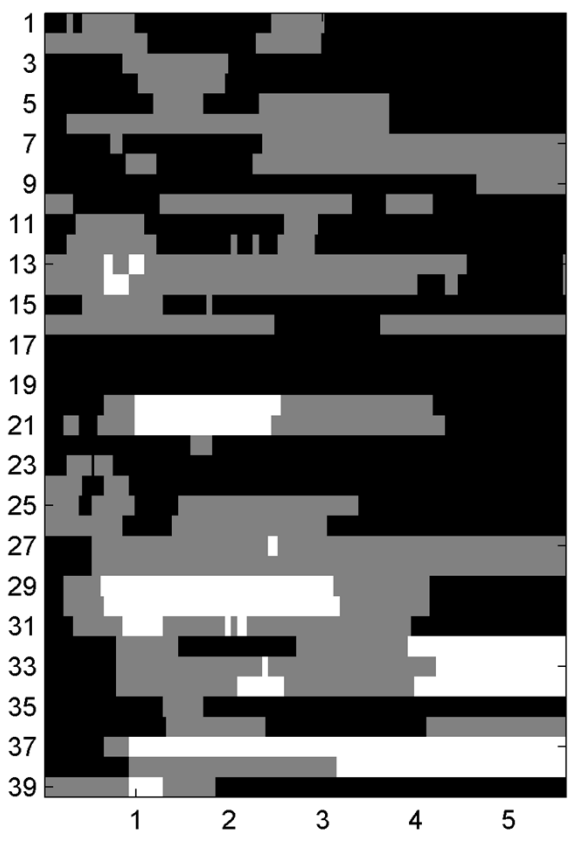


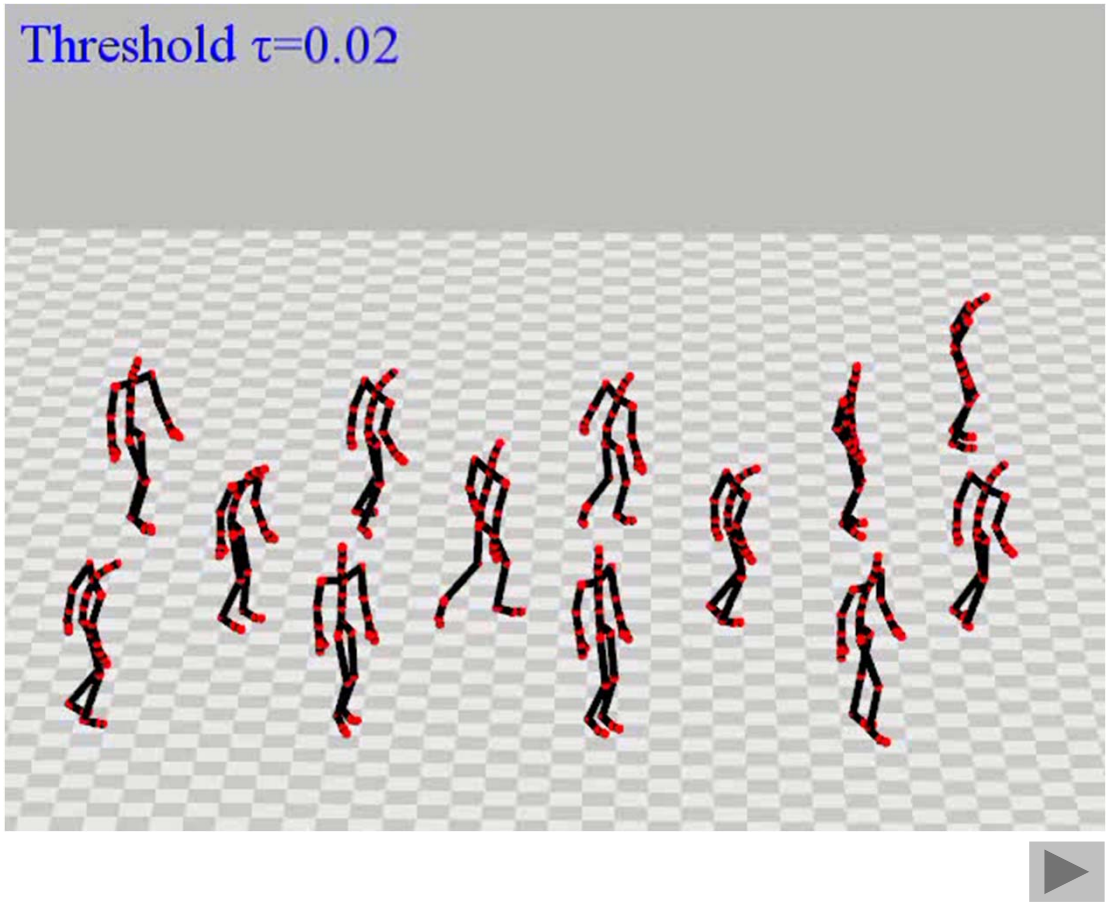


MT-based Motion Retrieval: Basketball

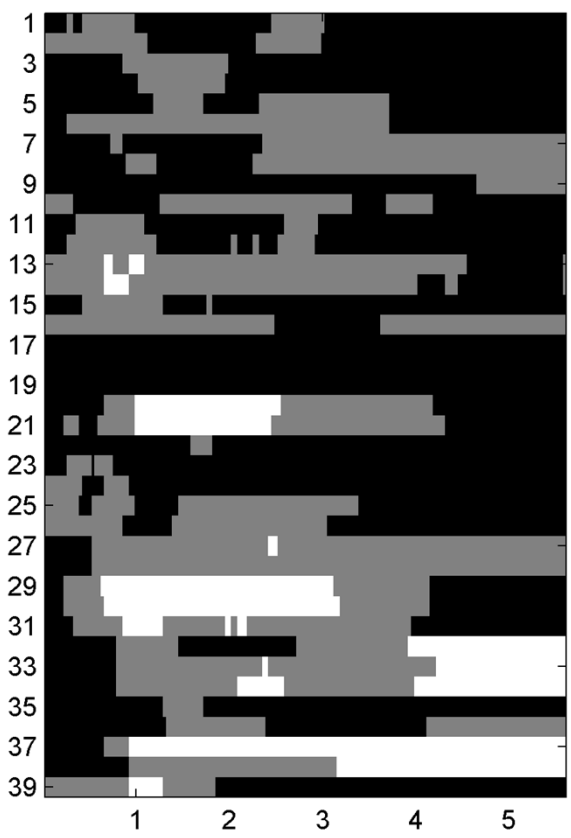


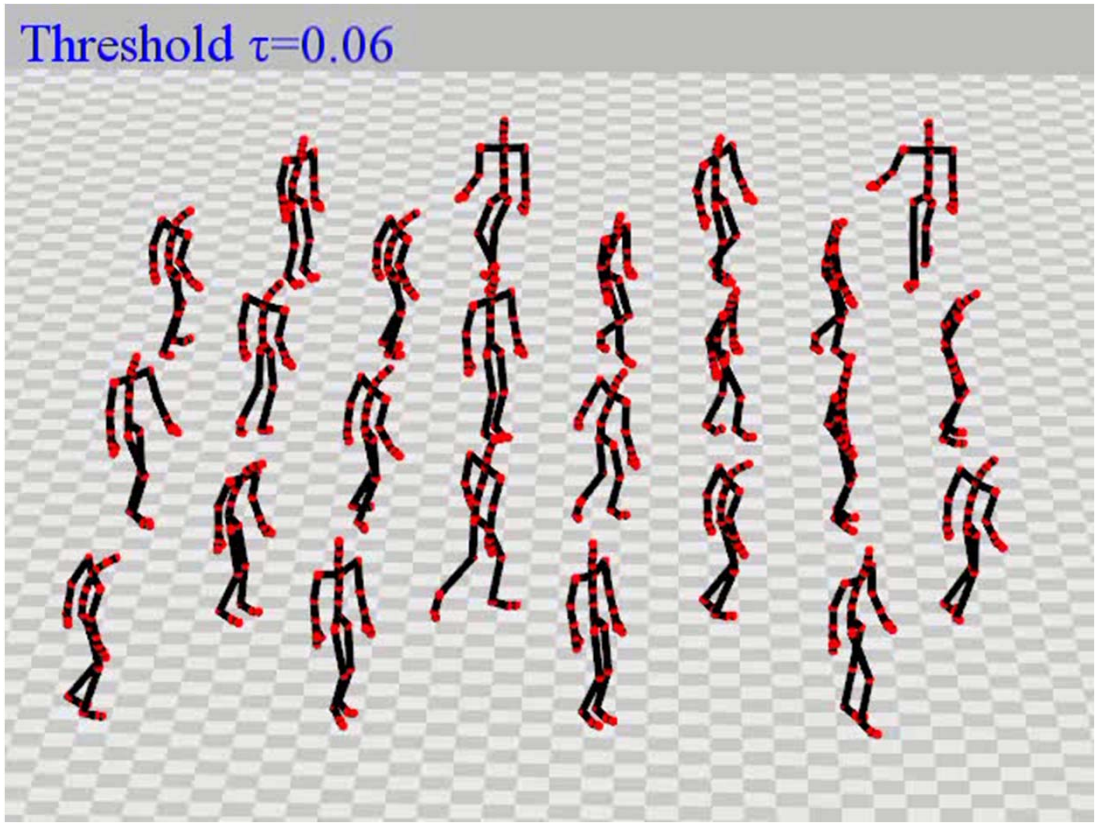
MT-based Motion Retrieval: Lie Down Floor





MT-based Motion Retrieval: Lie Down Floor



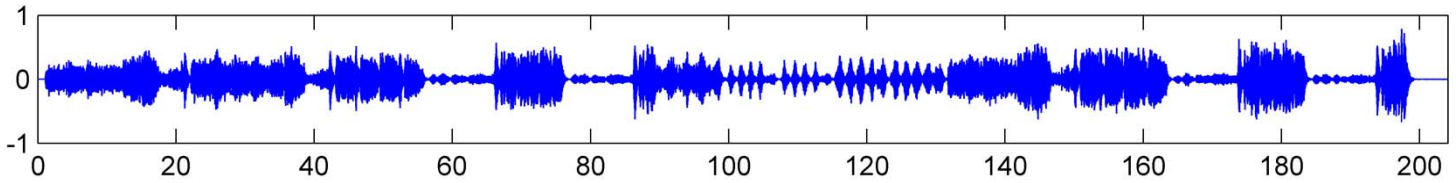


Overview

- Introduction
- Music Retrieval
- Motion Retrieval
- Music Structure Analysis
- Beat Tracking

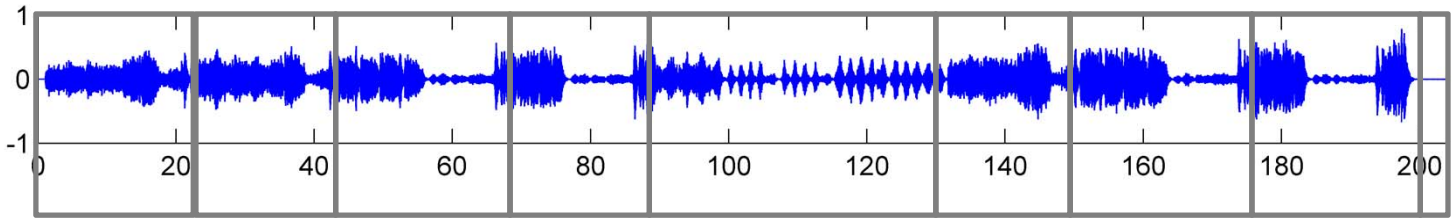
Music Structure Analysis

Example: Brahms Hungarian Dance No. 5 (Ormandy)



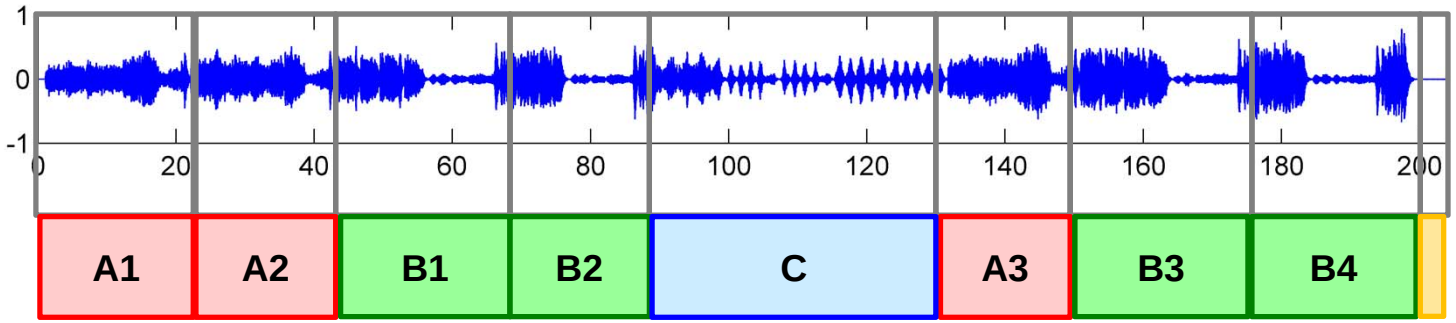
Music Structure Analysis

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Music Structure Analysis

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Music Structure Analysis

General goal: Divide an audio recording into temporal segments corresponding to musical parts and group these segments into musically meaningful categories.

Examples:

- Stanzas of a folk song
- Intro, verse, chorus, bridge, outro sections of a pop song
- Exposition, development, recapitulation, coda of a sonata
- Musical form ABACADA … of a rondo

Music Structure Analysis

General goal: Divide an audio recording into temporal segments corresponding to musical parts and group these segments into musically meaningful categories.

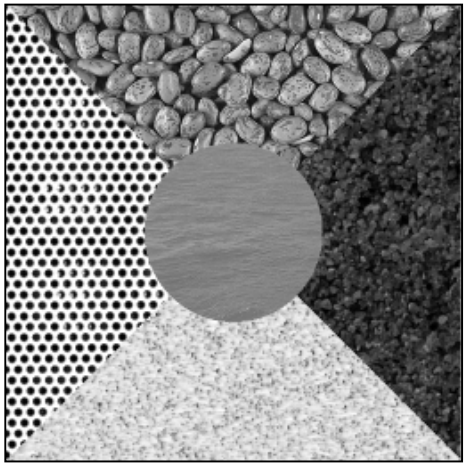
Challenge: There are many different principles for creating relationships that form the basis for the musical structure.

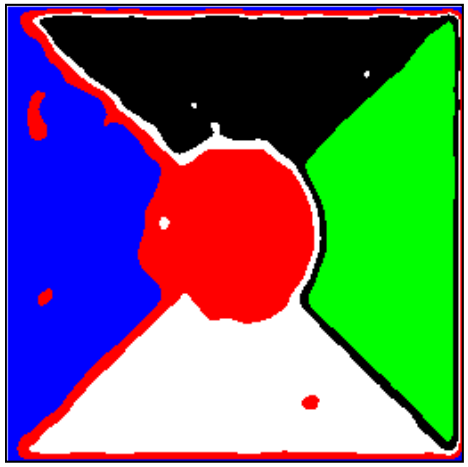
- Homogeneity: Consistency in tempo, instrumentation, key, …
- Novelty: Sudden changes, surprising elements …
- Repetition: Repeating themes, motives, rhythmic patterns,…

Music Structure Analysis

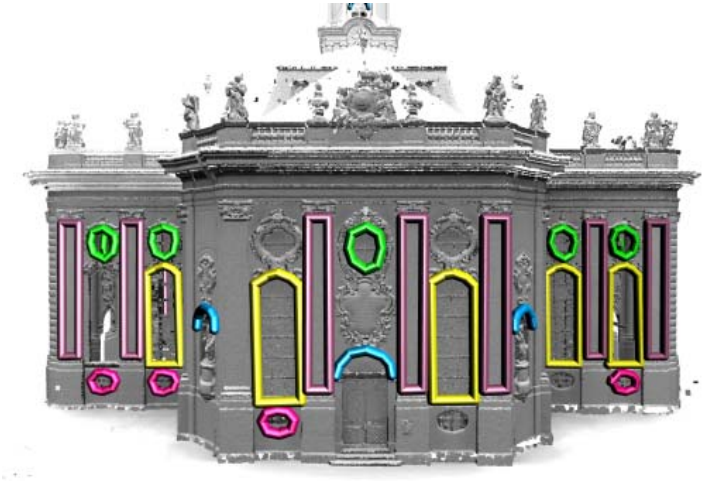
Novelty

Homogeneity





Repetition



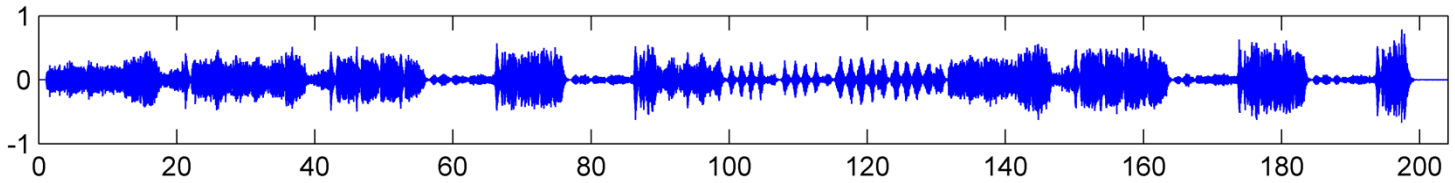
Feature Representation

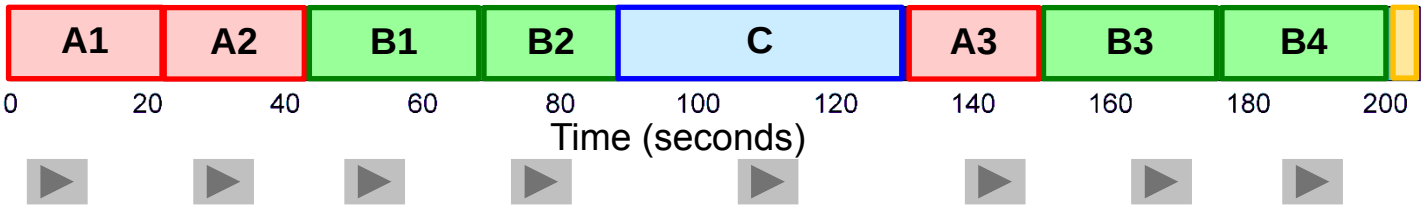
General goal: Convert an audio recording into a mid-level representation that captures certain musical properties while supressing other properties.

- Timbre / Instrumentation
- Tempo / Rhythm
- Pitch / Harmony

Feature Representation

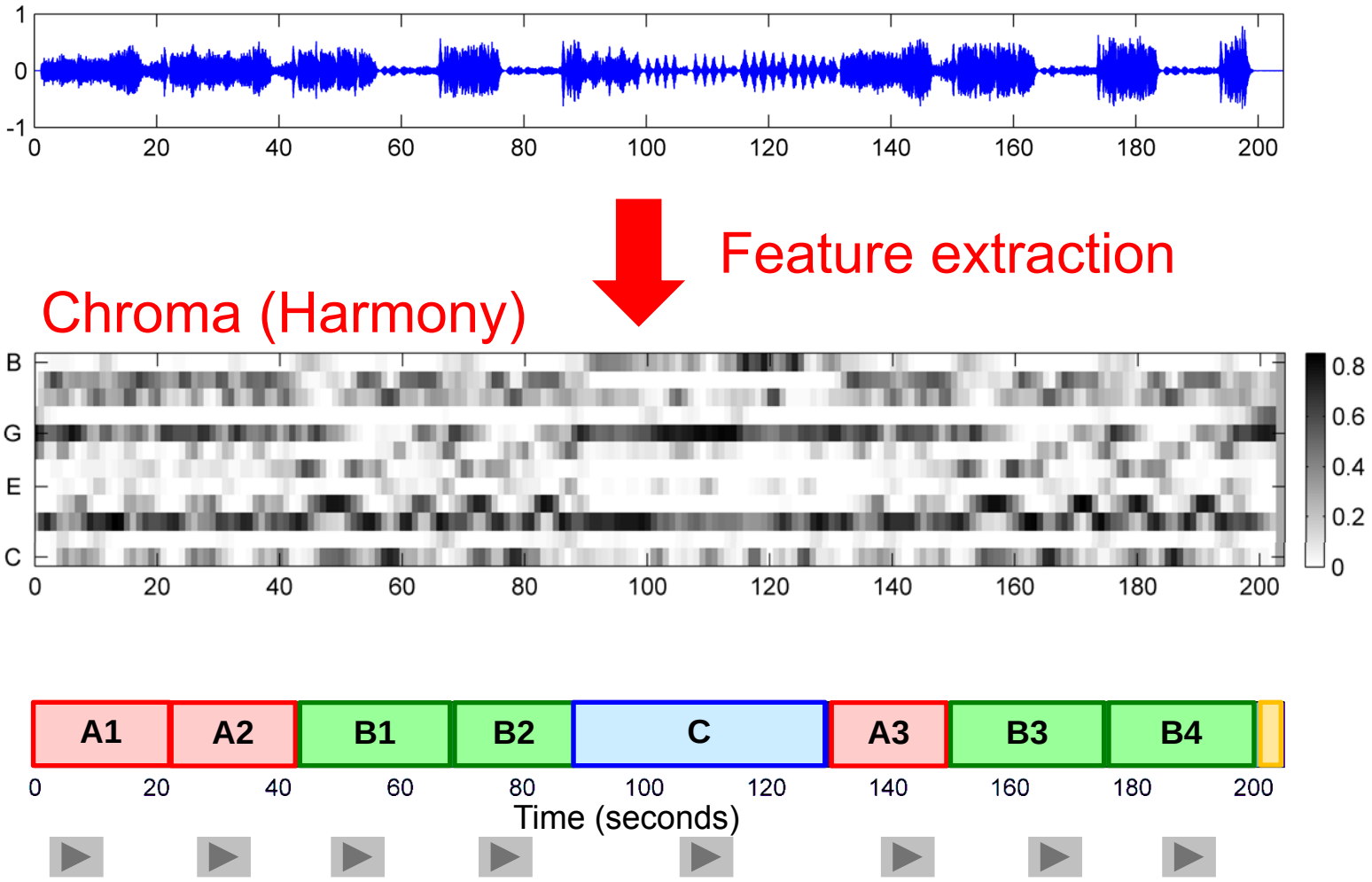
Example: Brahms Hungarian Dance No. 5 (Ormandy)





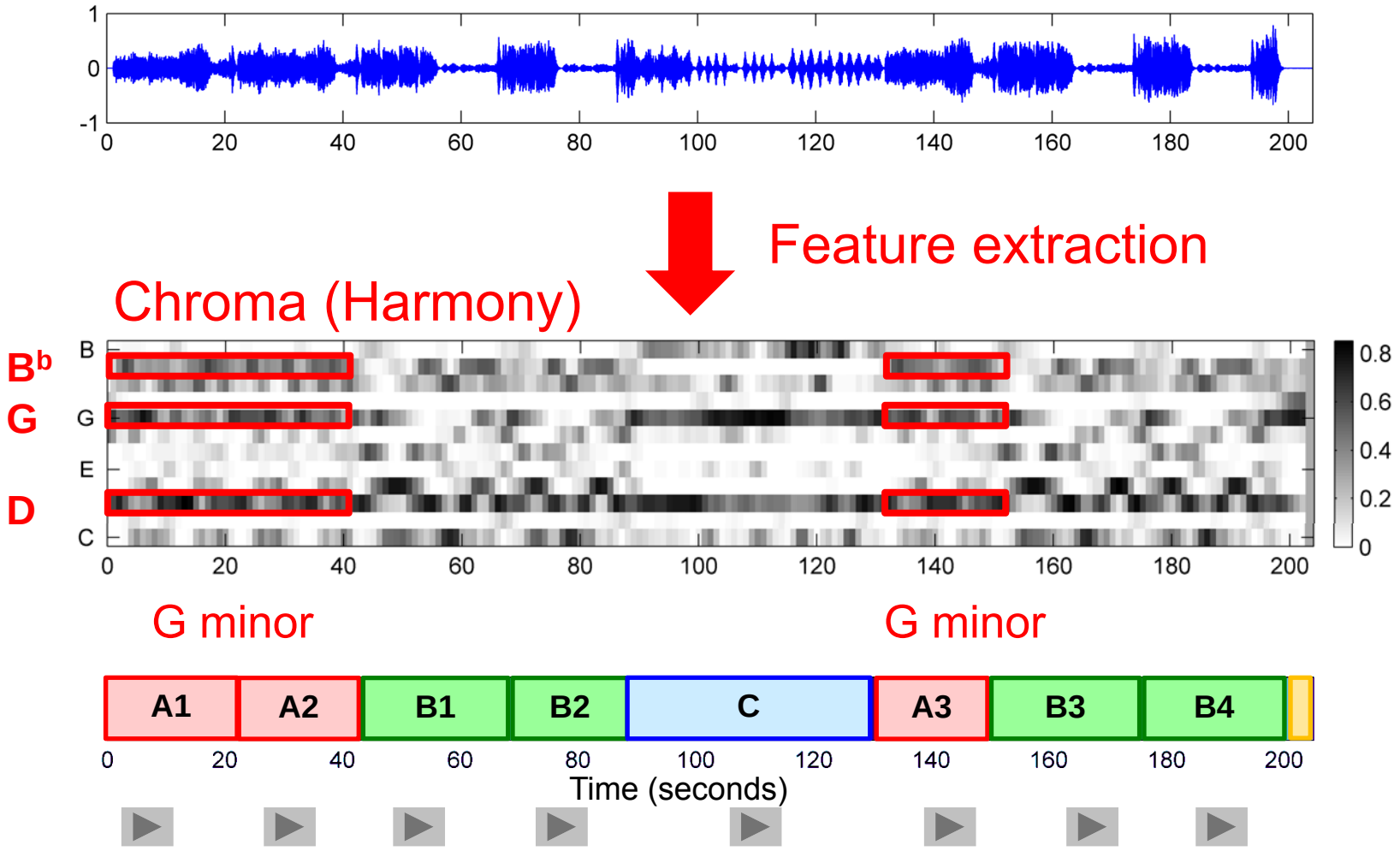
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



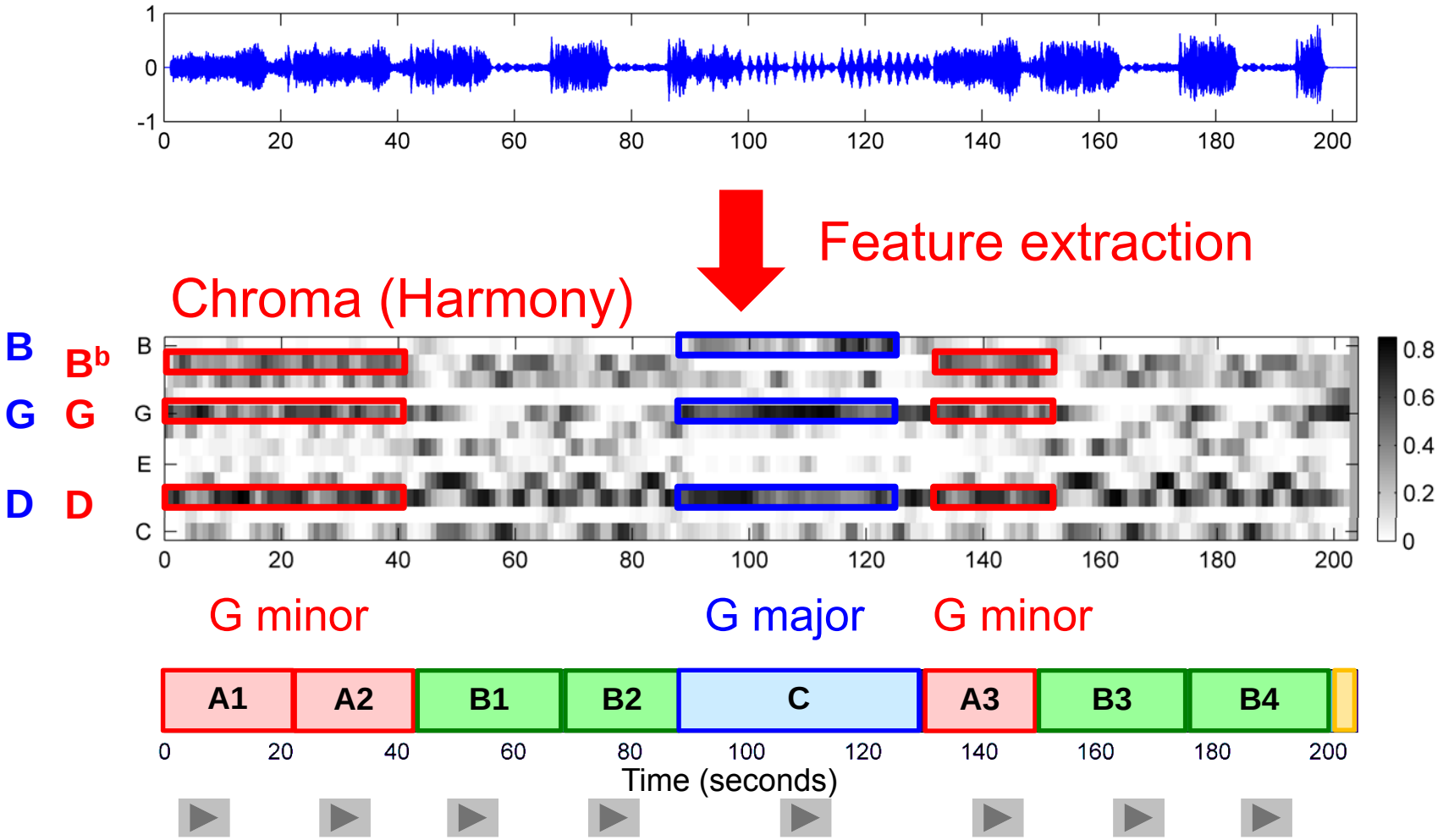
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



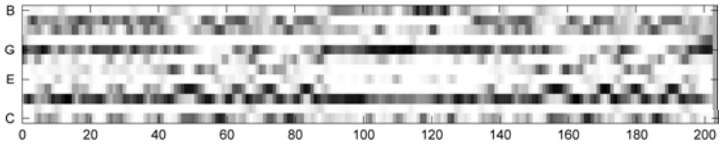
Self-Similarity Matrix (SSM)

General idea: Compare each element of the feature sequence with each other element of the feature sequence based on a suitable similarity measure.

→ Quadratic self-similarity matrix

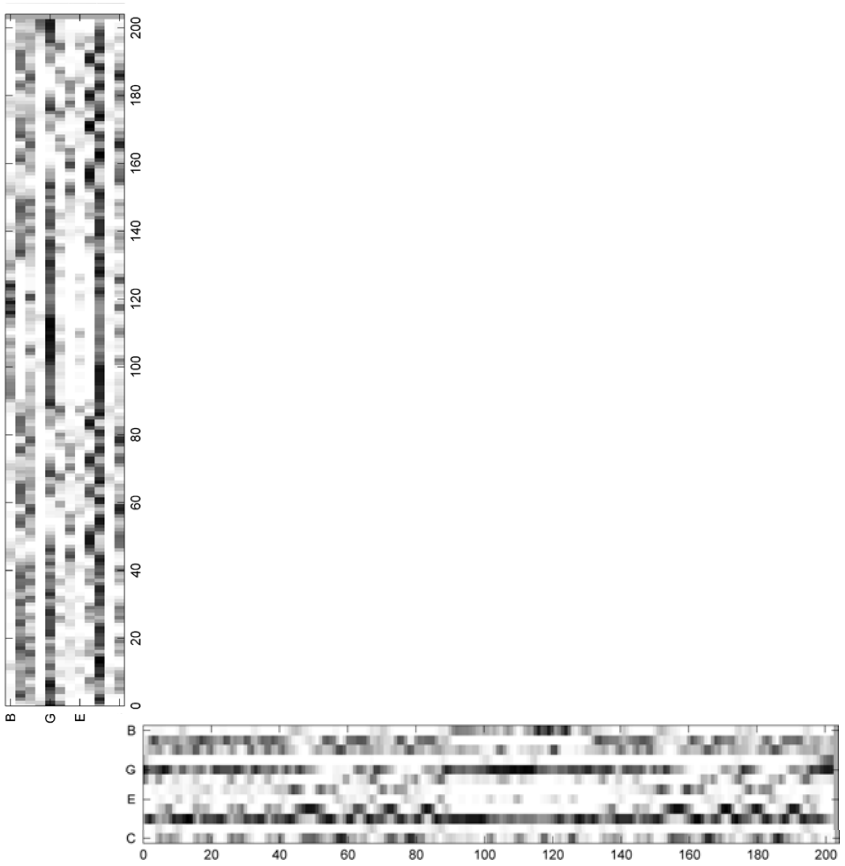
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



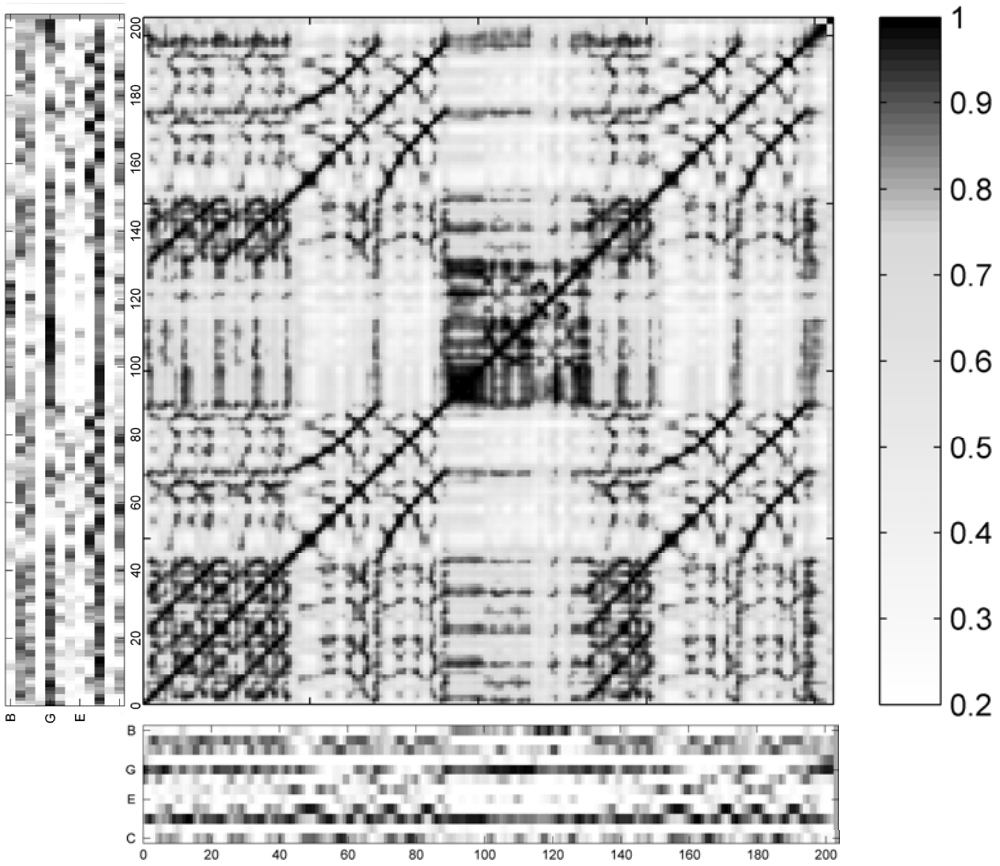
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



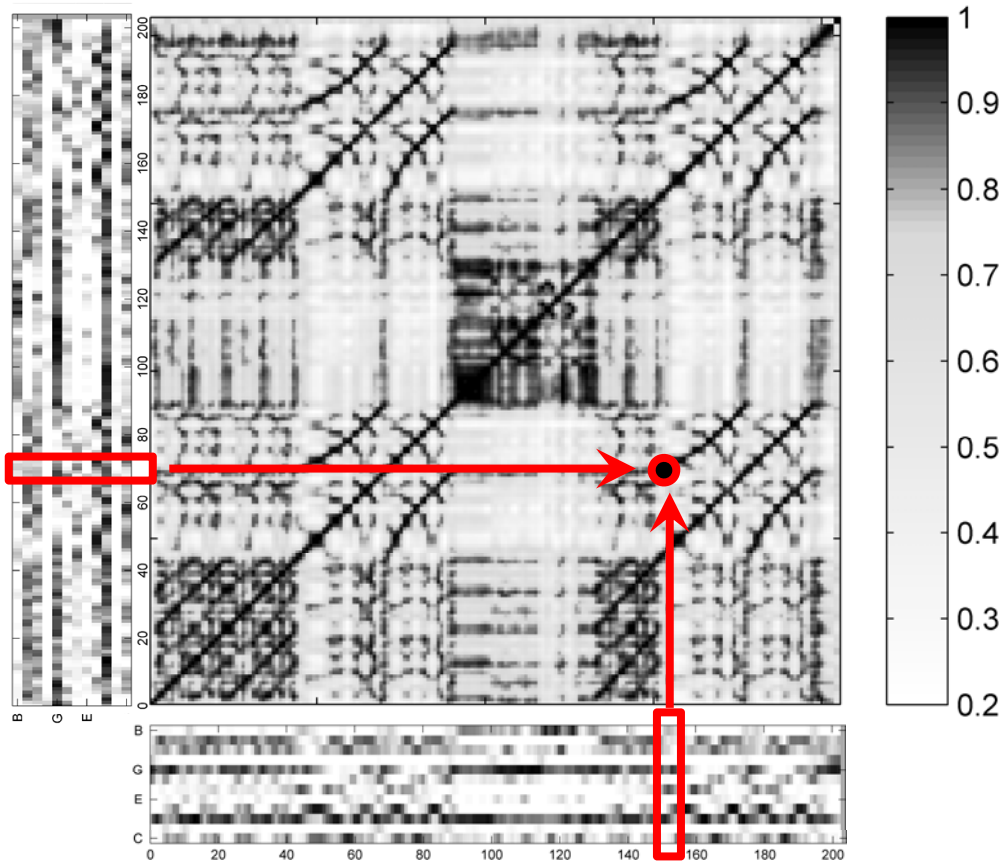
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



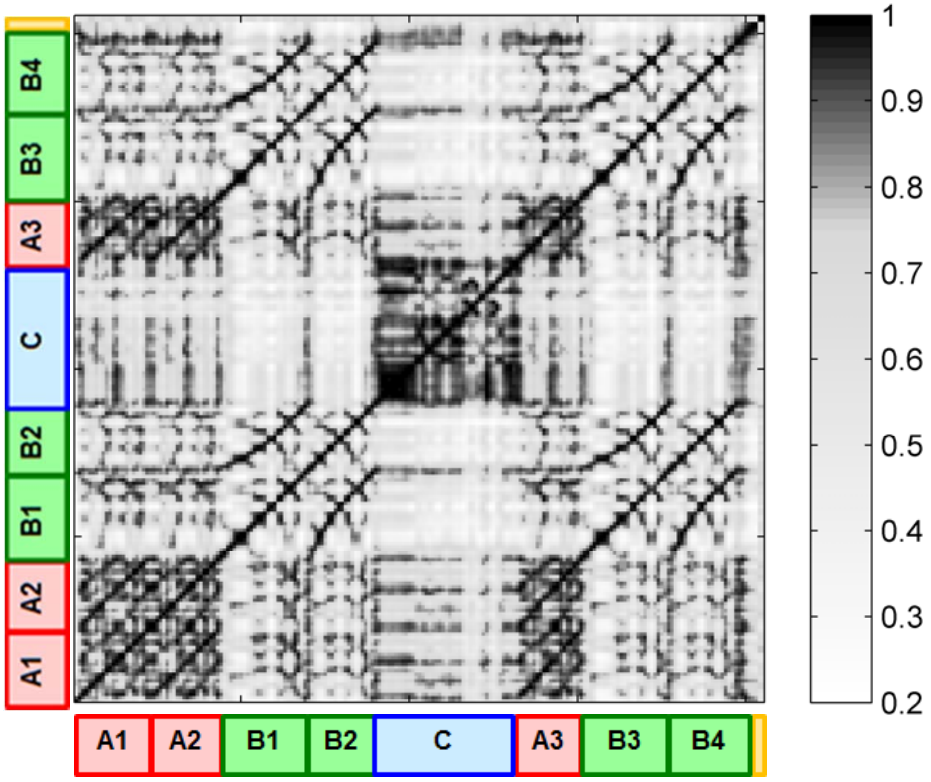
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



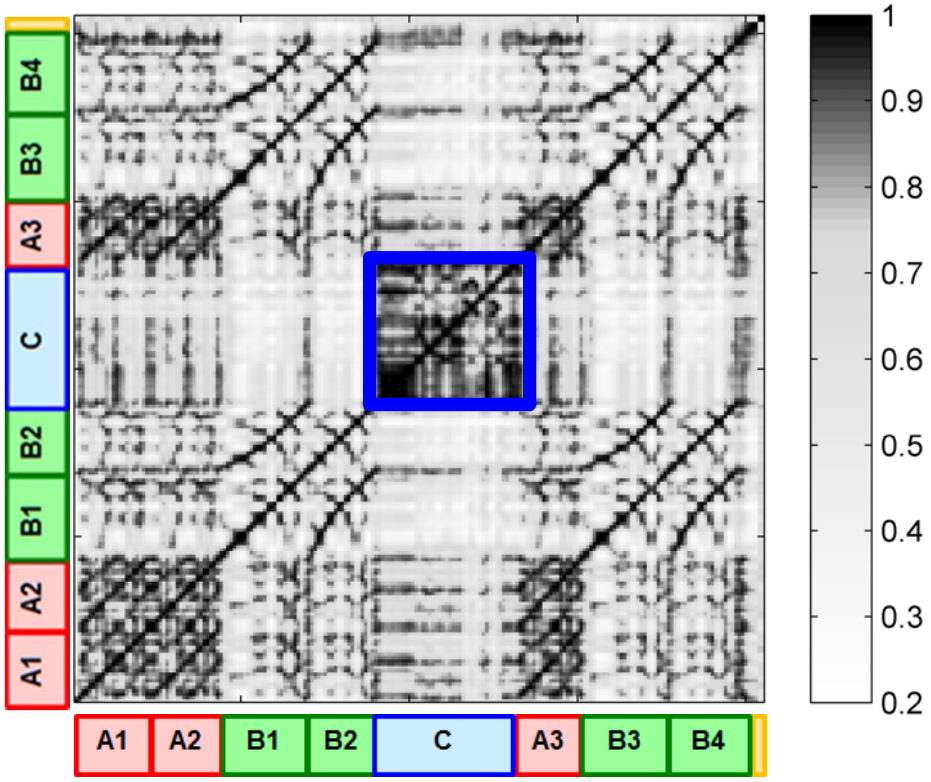
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



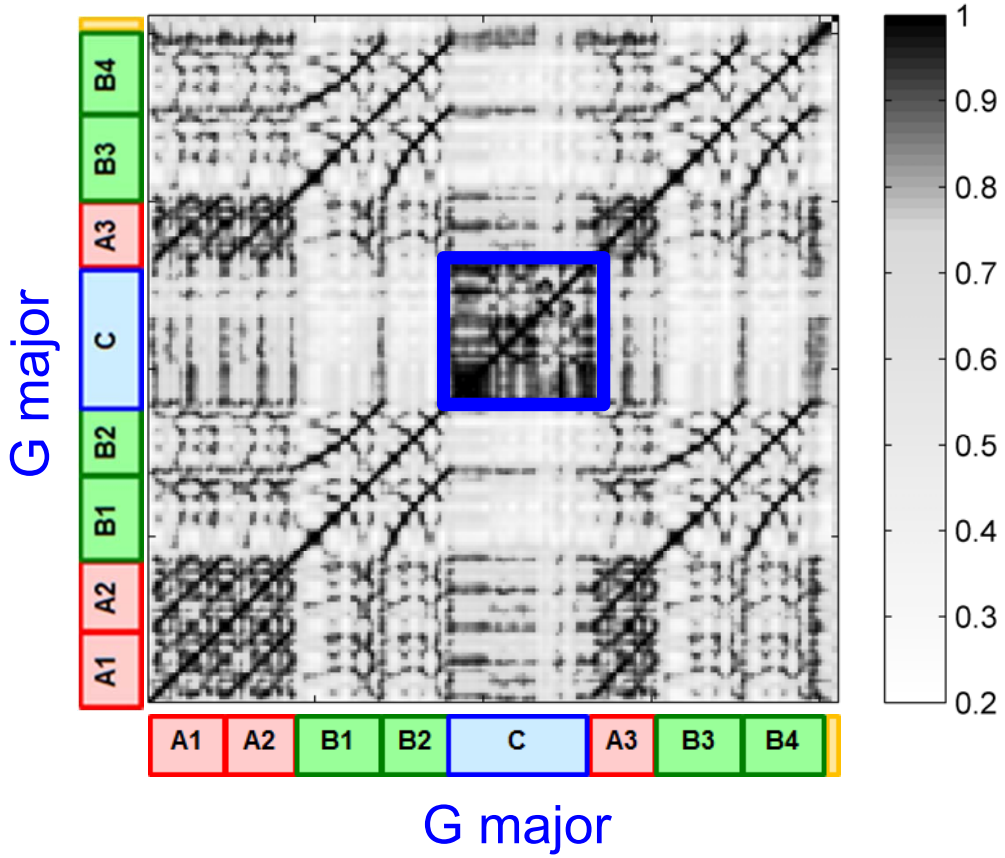
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



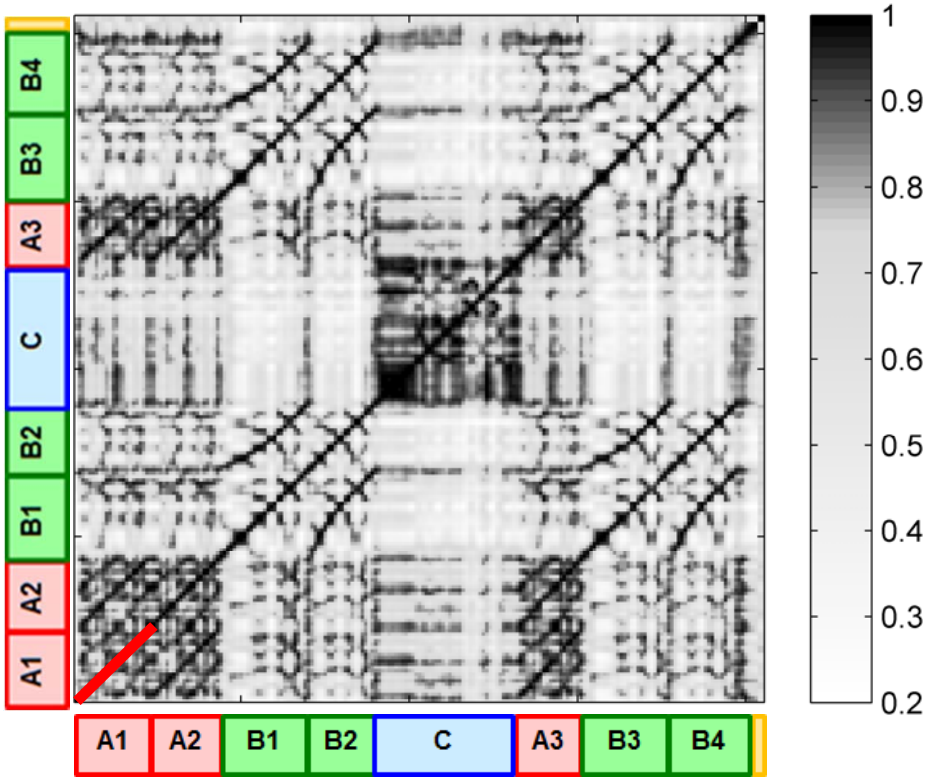
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



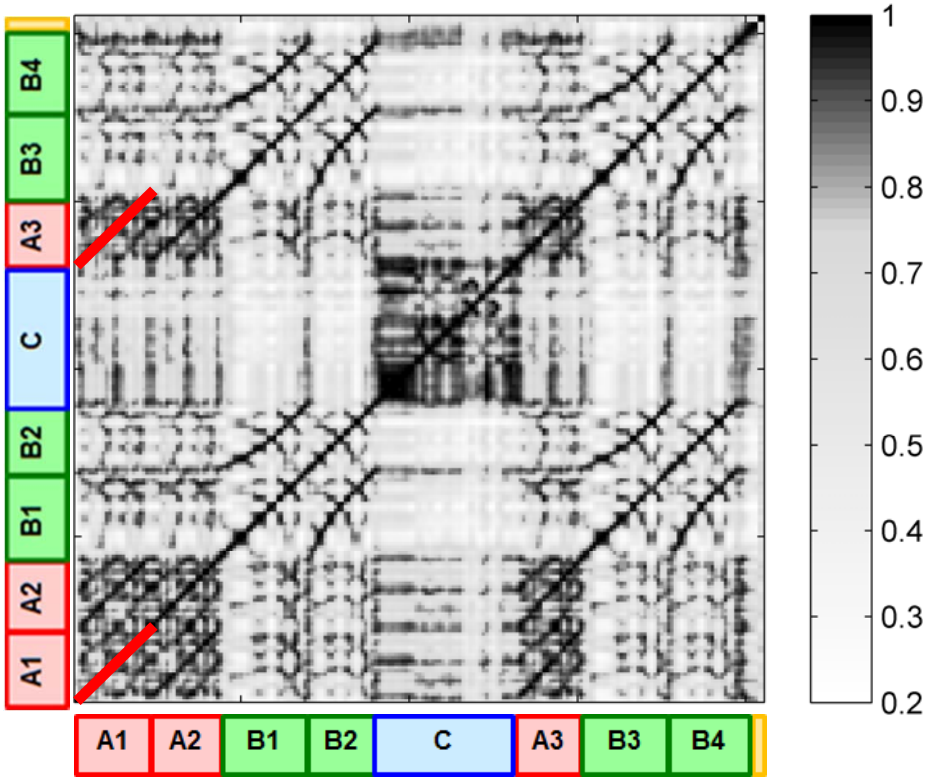
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



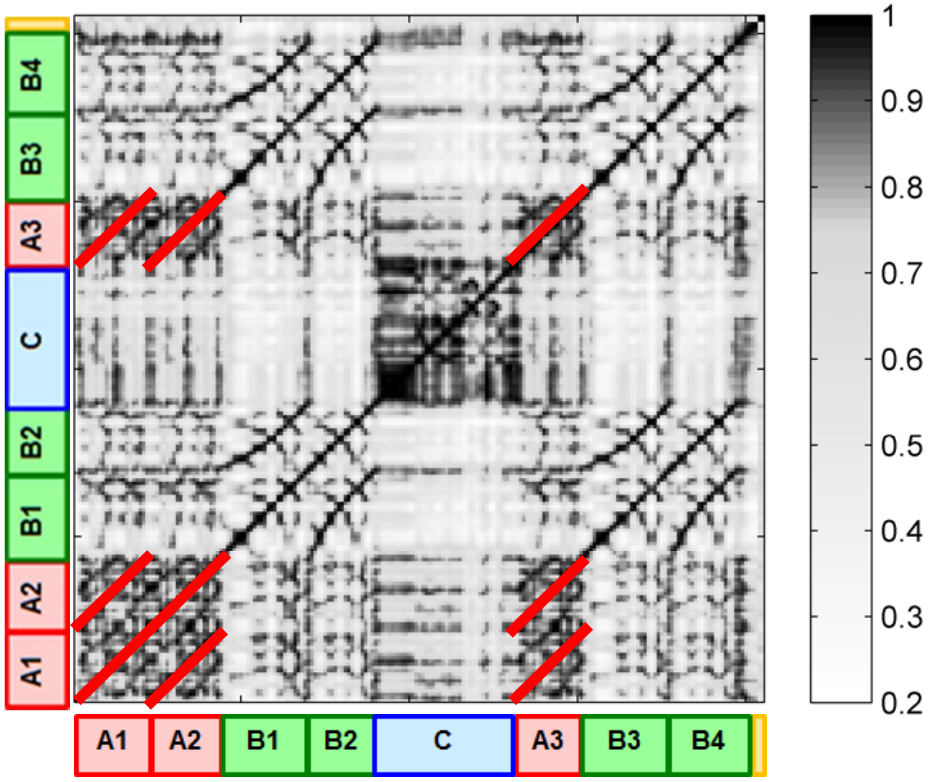
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



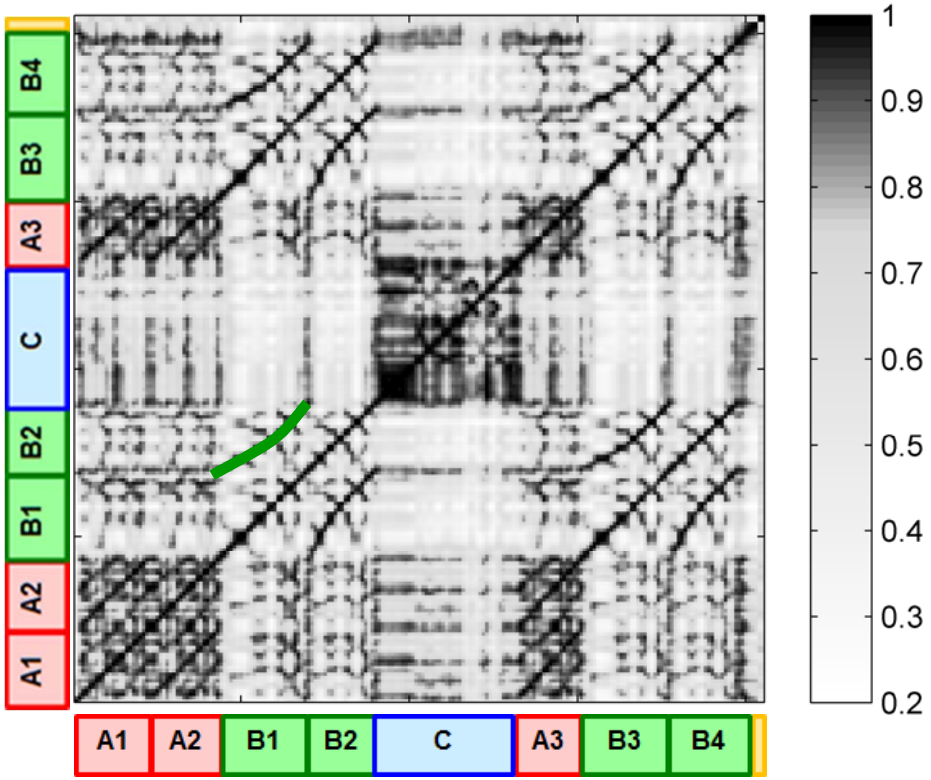
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



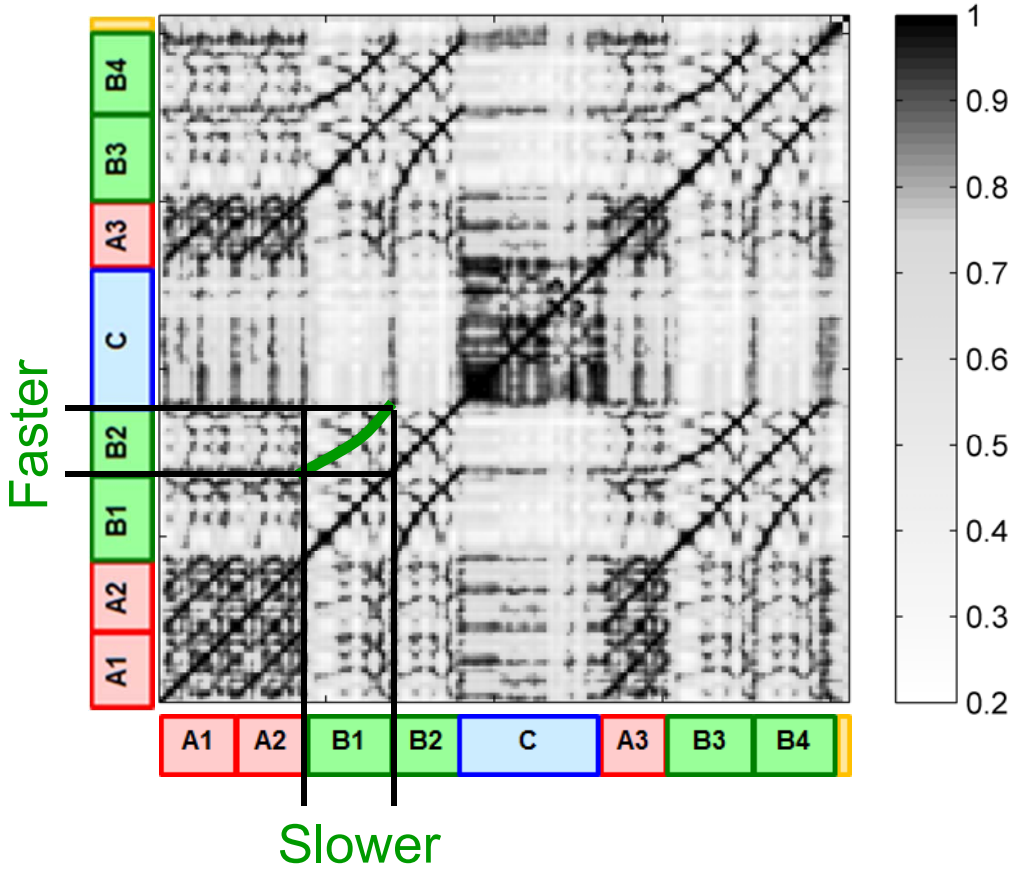
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



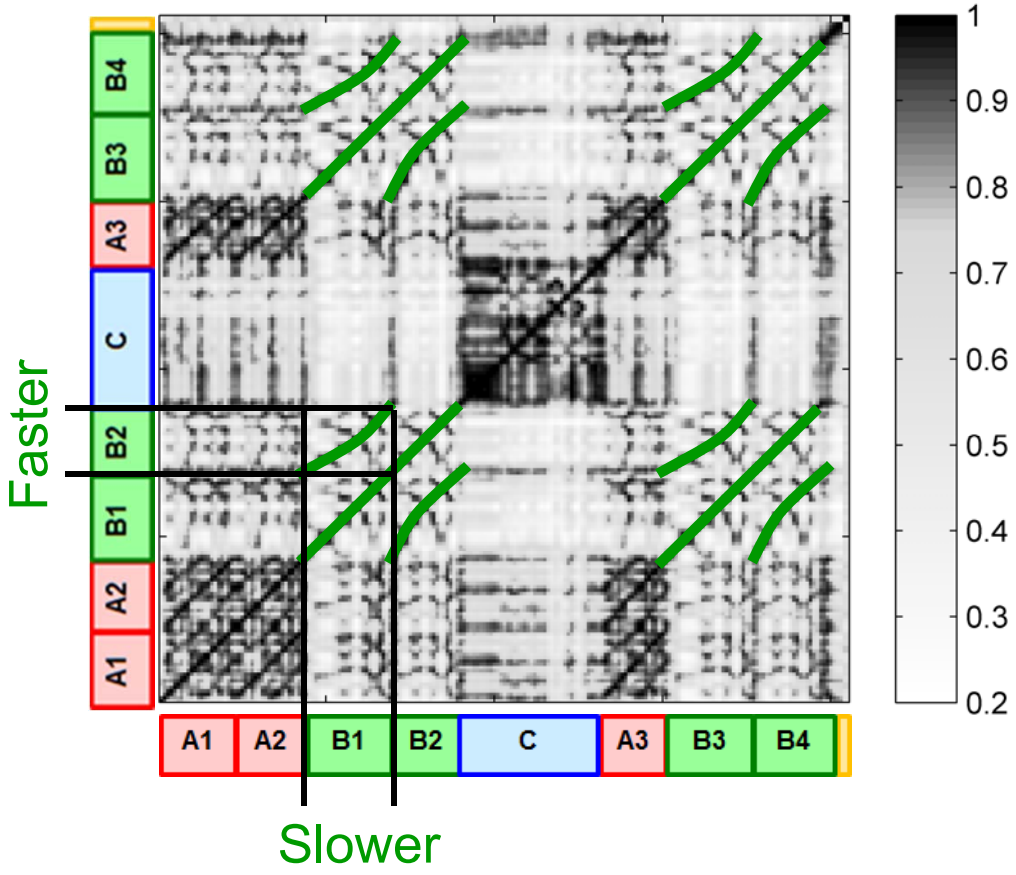
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



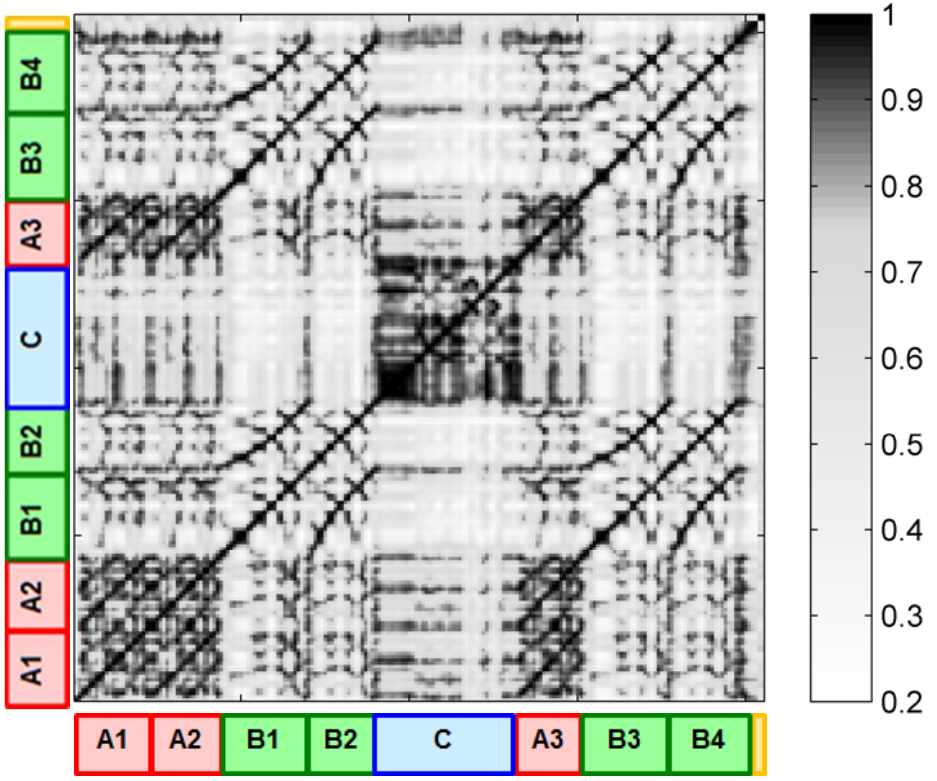
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)

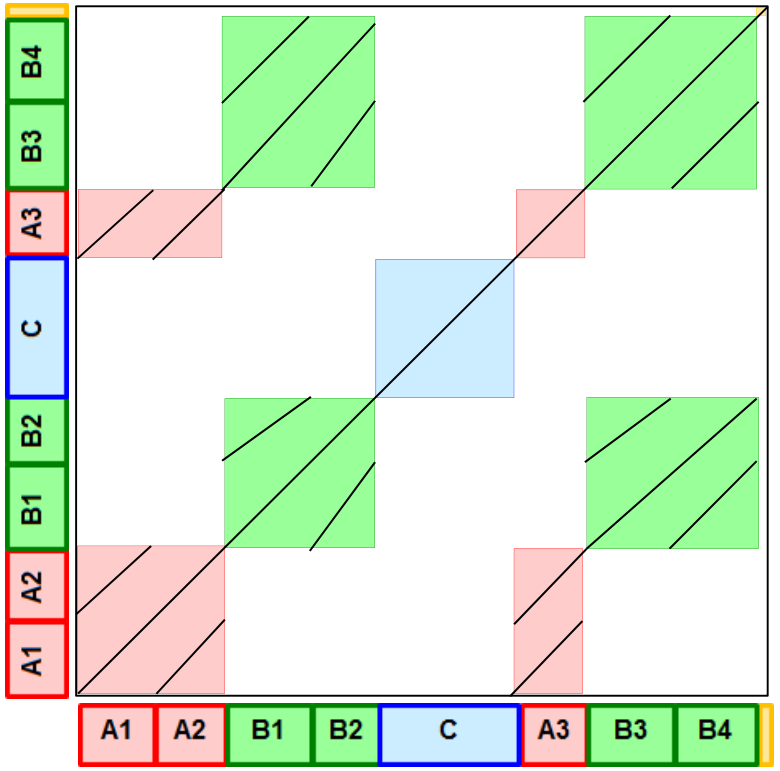


Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Idealized SSM



Self-Similarity Matrix (SSM)

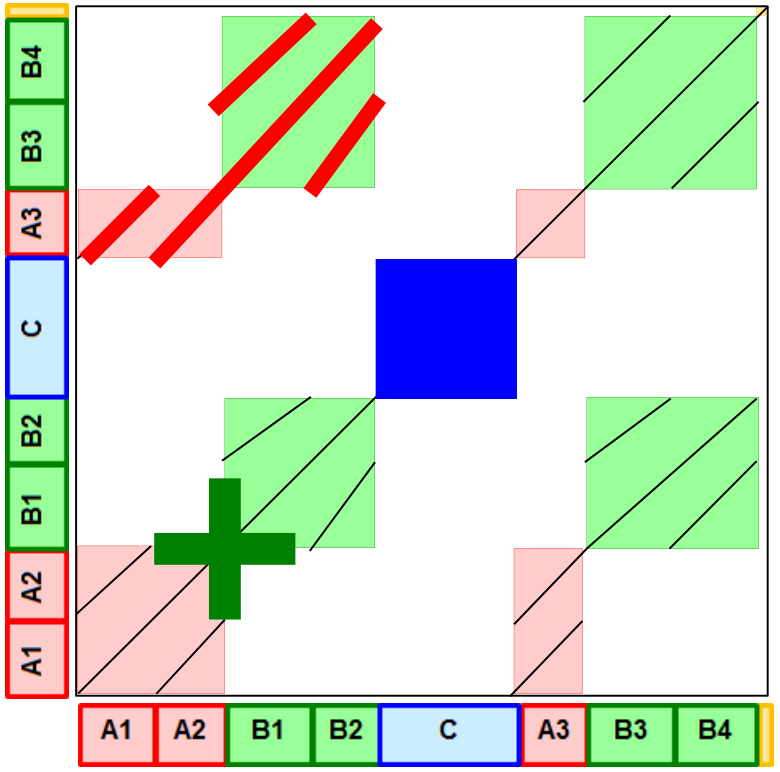
Example: Brahms Hungarian Dance No. 5 (Ormandy)

Blocks: Homogeneity

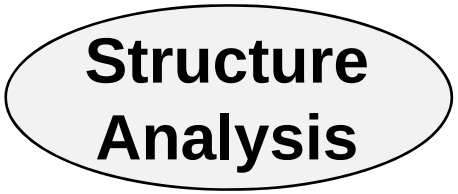
Paths: Repetition

Corners: Novelty

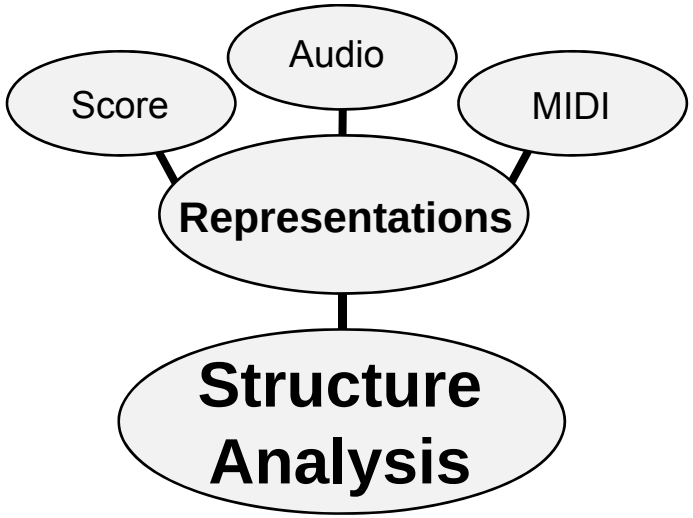
Idealized SSM



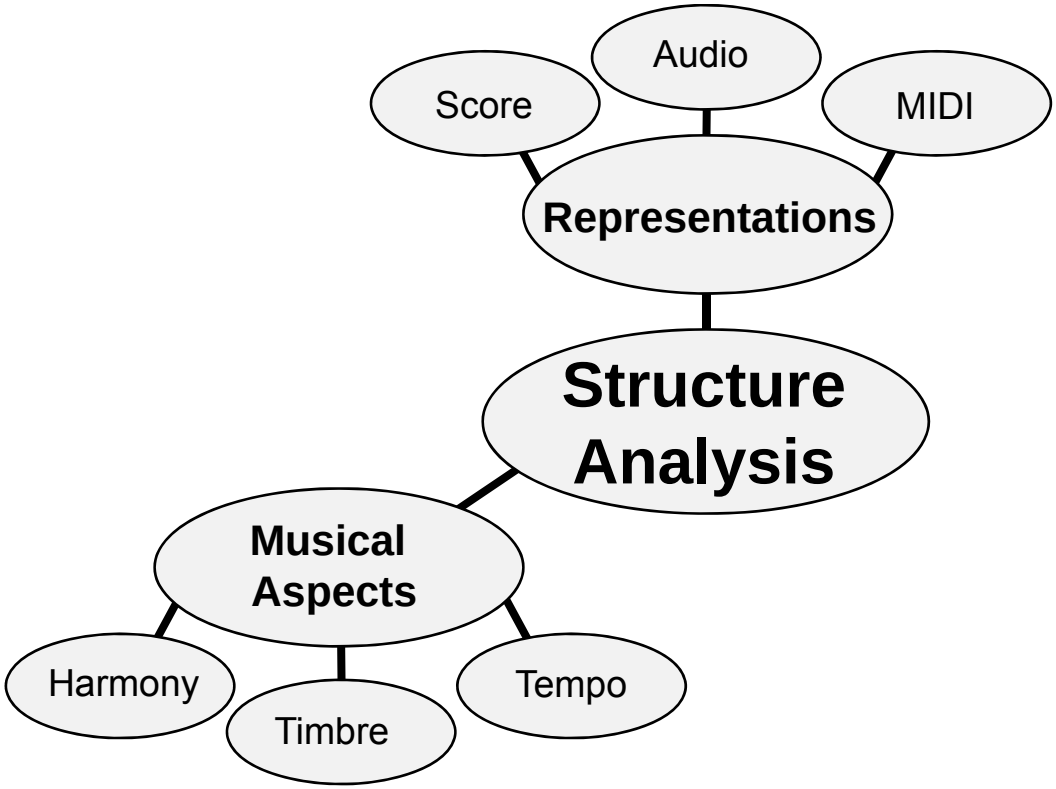
Music Structure Analysis



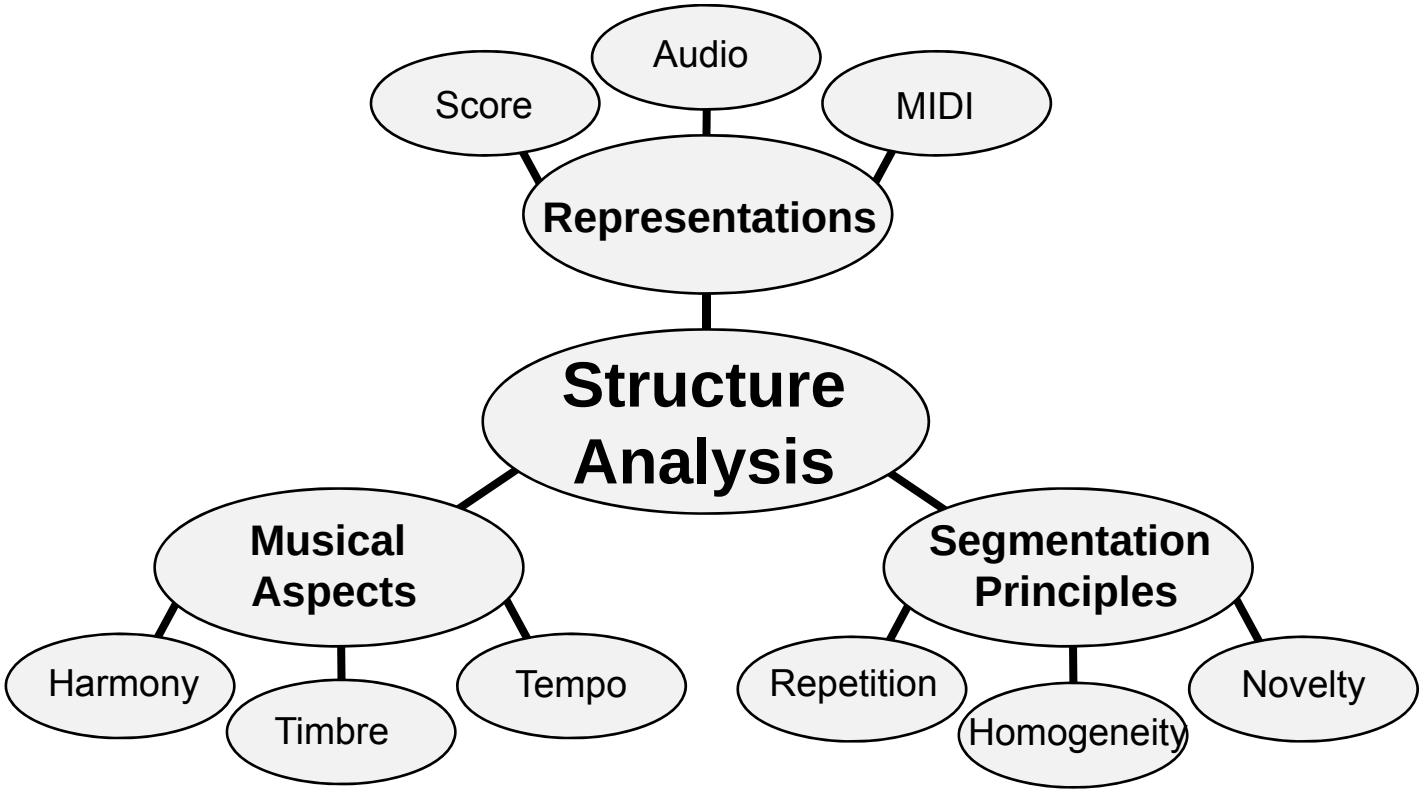
Music Structure Analysis



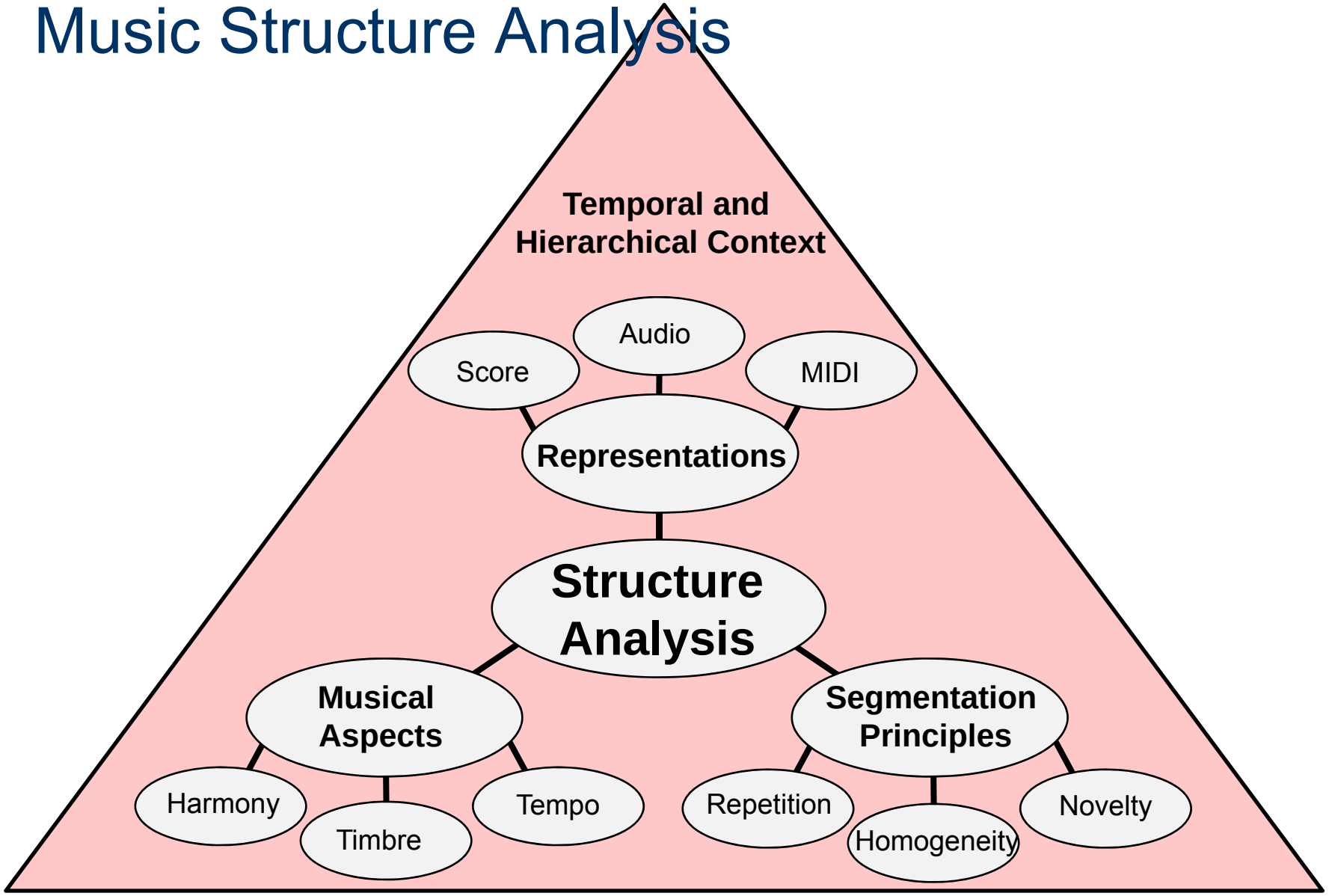
Music Structure Analysis



Music Structure Analysis



Music Structure Analysis



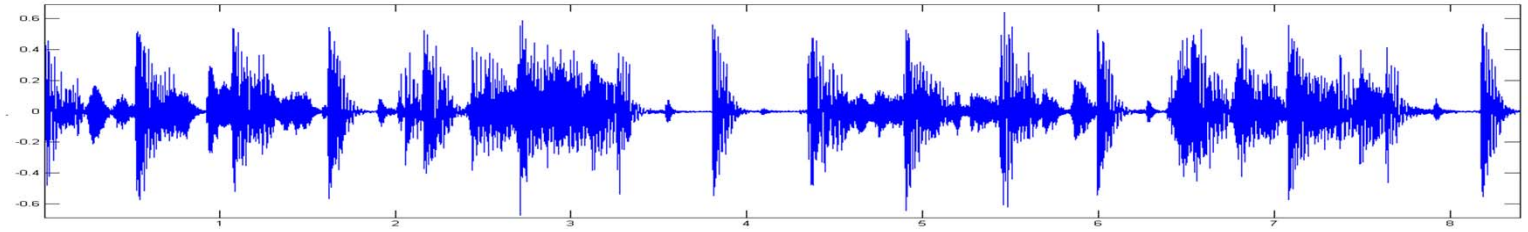
Overview

- Introduction
- Music Retrieval
- Motion Retrieval
- Music Structure Analysis
- Beat Tracking

Beat Tracking

Basic task: “Tapping the foot when listening to music”

Example: Queen – Another One Bites The Dust

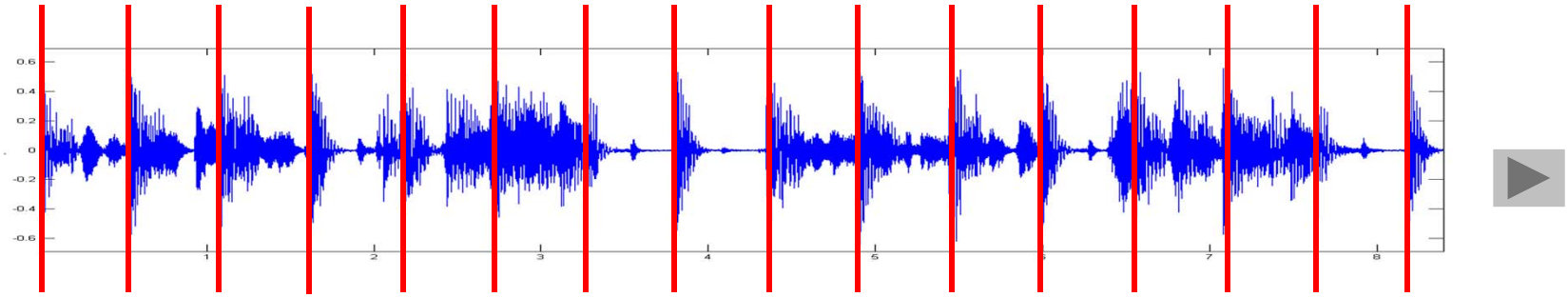


Time (seconds)

Beat Tracking

Basic task: “Tapping the foot when listening to music”

Example: Queen – Another One Bites The Dust

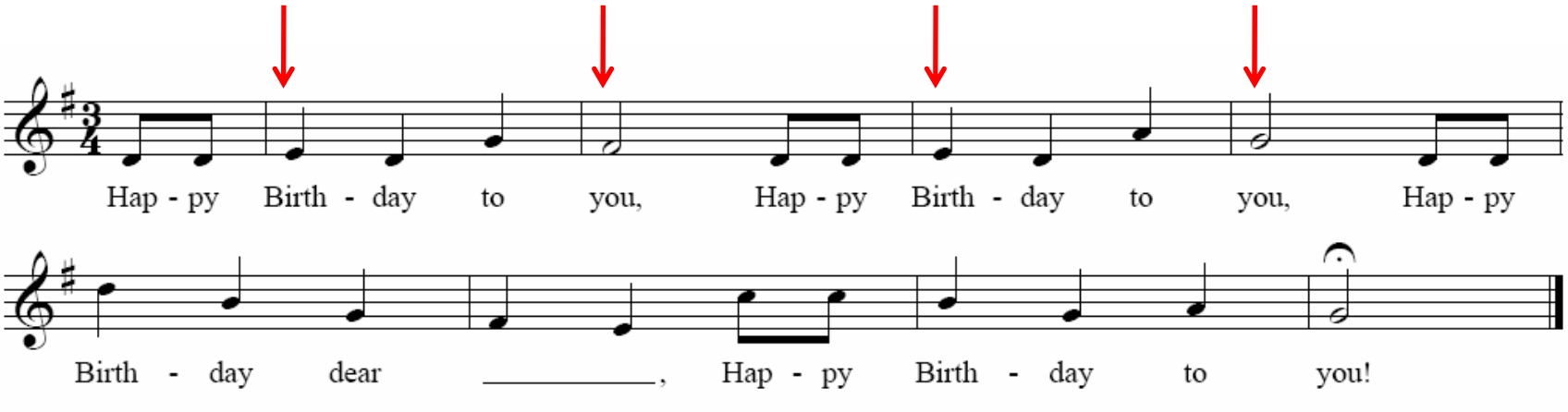


Time (seconds)

Beat Tracking

Example: Happy Birthday to you

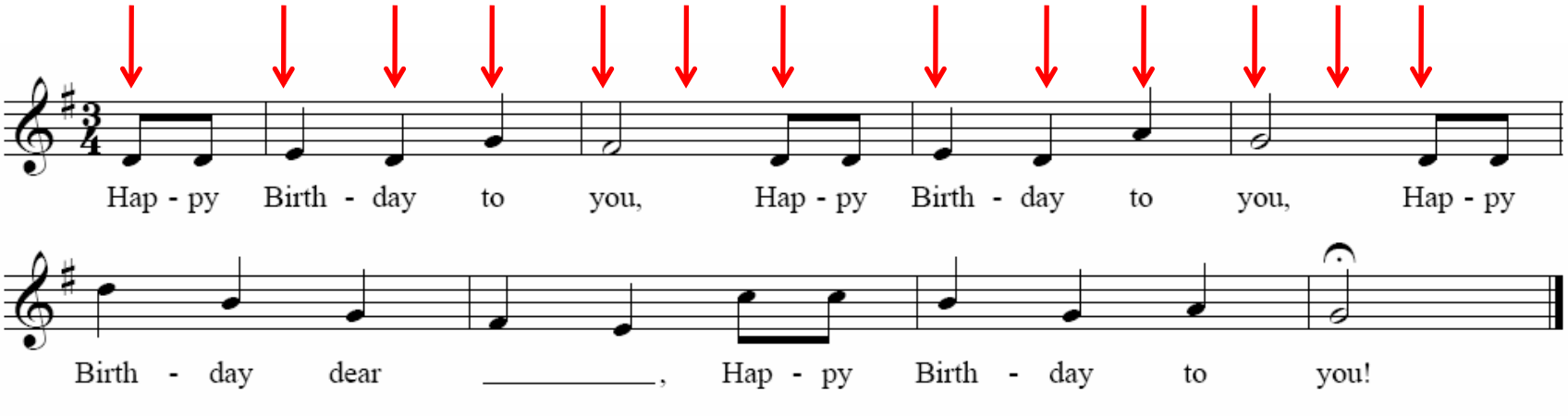
Pulse level: **Measure**



Beat Tracking

Example: Happy Birthday to you

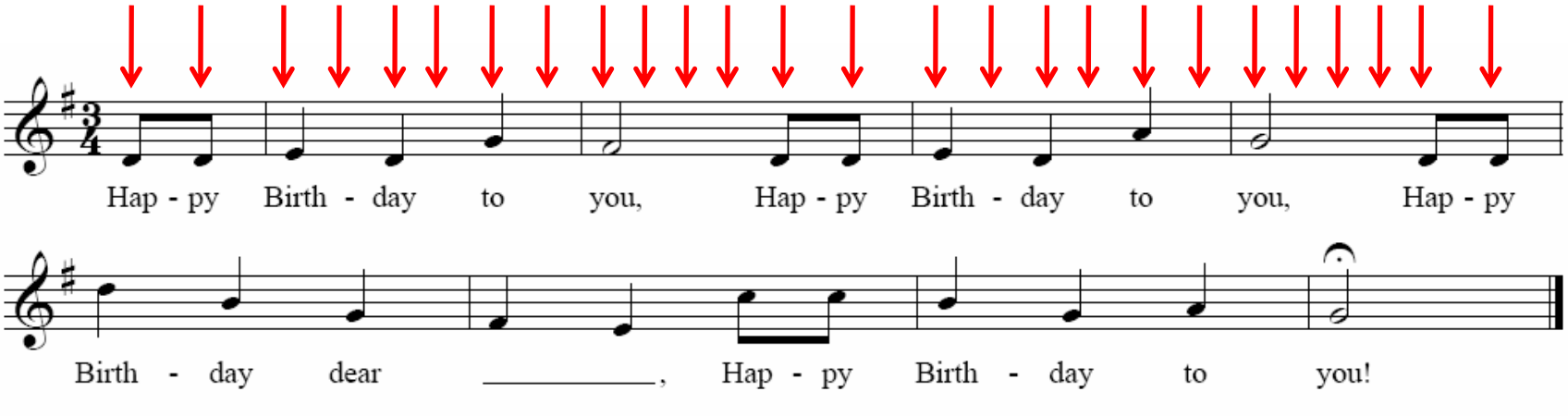
Pulse level: **Tactus (beat)**



Beat Tracking

Example: Happy Birthday to you

Pulse level: **Tatum (temporal atom)**



Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: ???

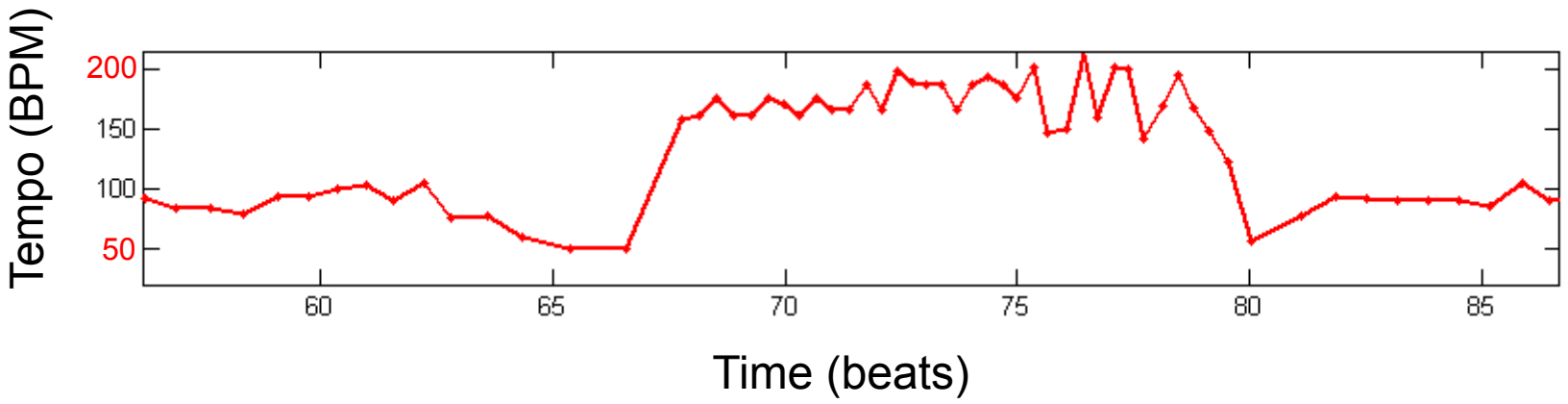
Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: 50-200 BPM

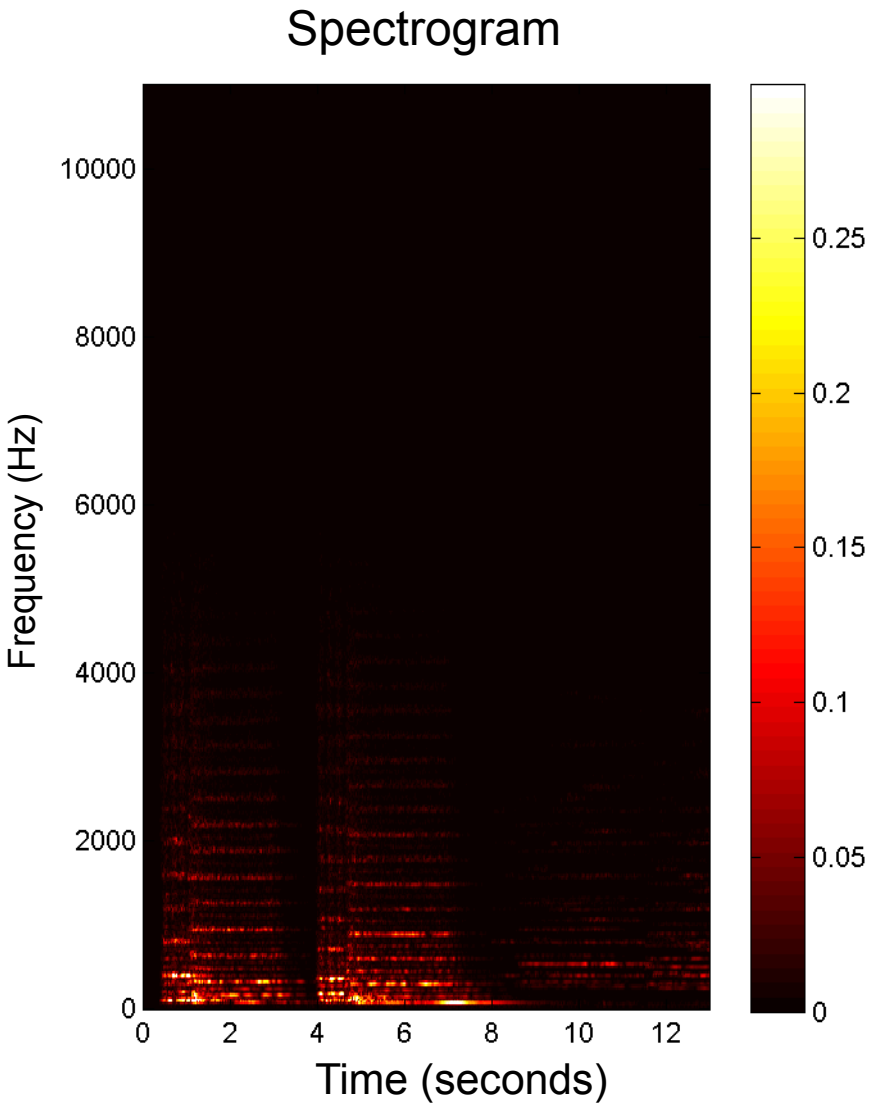
Tempo curve



Beat Tracking

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

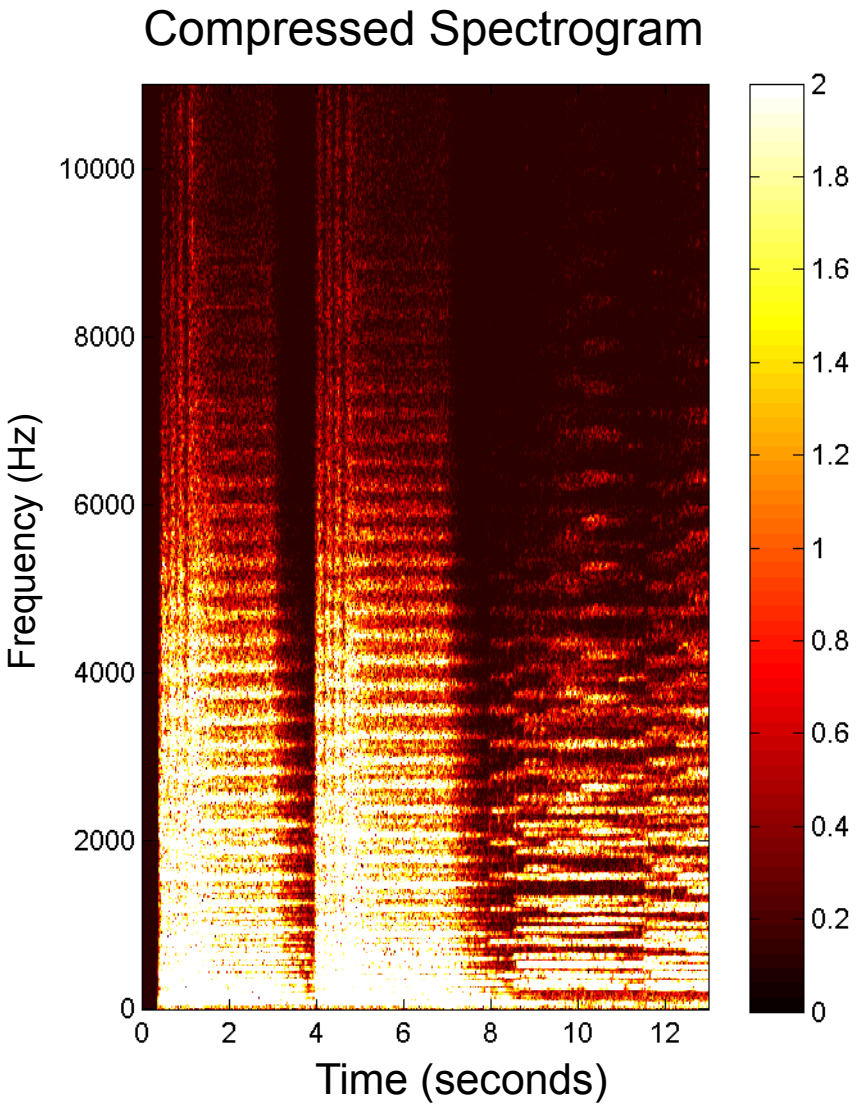
Beat Tracking



Steps:

1. Spectrogram

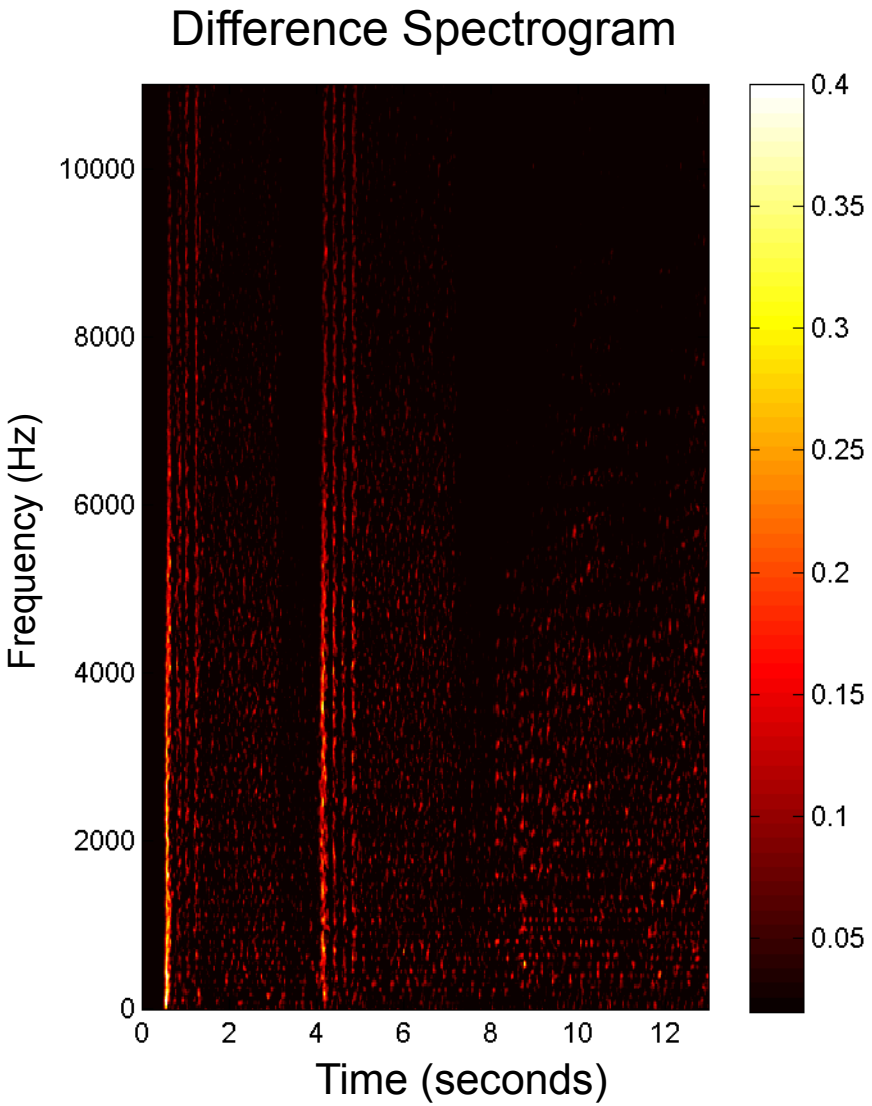
Beat Tracking



Steps:

1. Spectrogram
2. Log Compression

Beat Tracking



Steps:

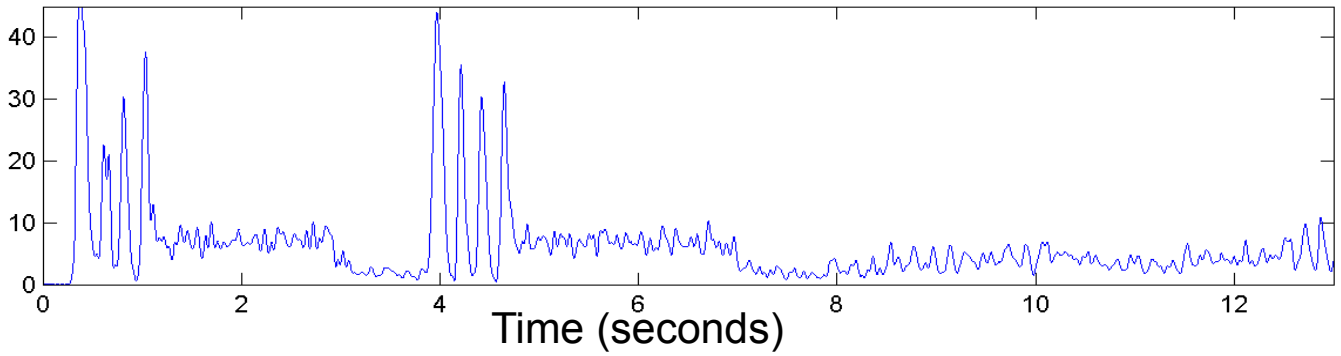
1. Spectrogram
2. Log Compression
3. Differentiation

Beat Tracking

Steps:

1. Spectrogram
2. Log Compression
3. Differentiation
4. Accumulation

Novelty Curve

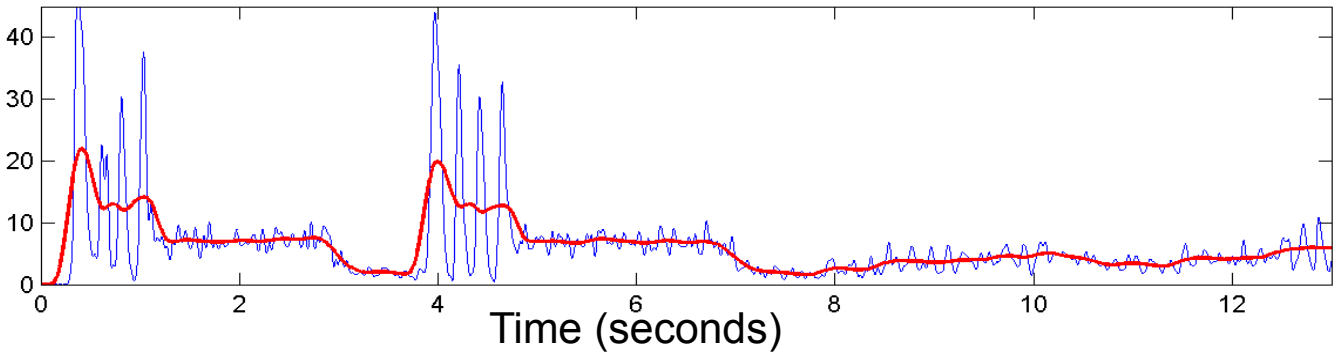


Beat Tracking

Steps:

1. Spectrogram
2. Log Compression
3. Differentiation
4. Accumulation

Novelty Curve
Local Average

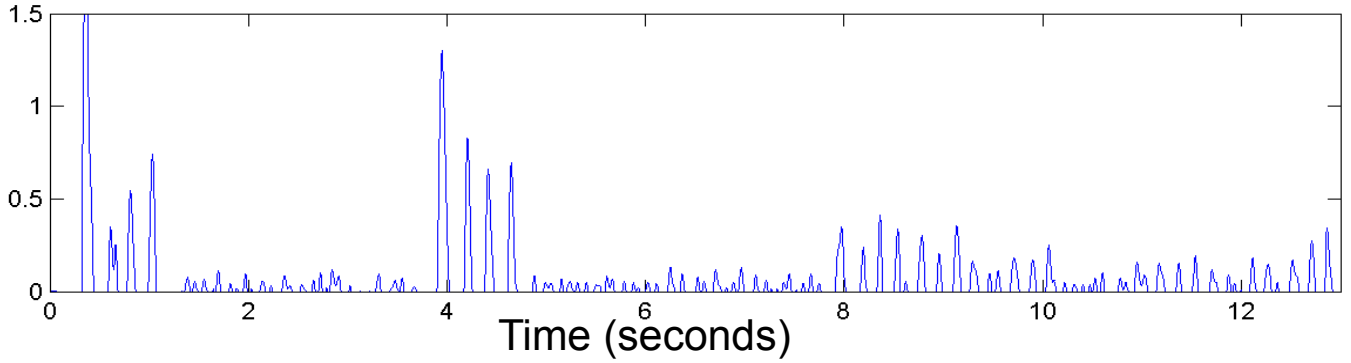


Beat Tracking

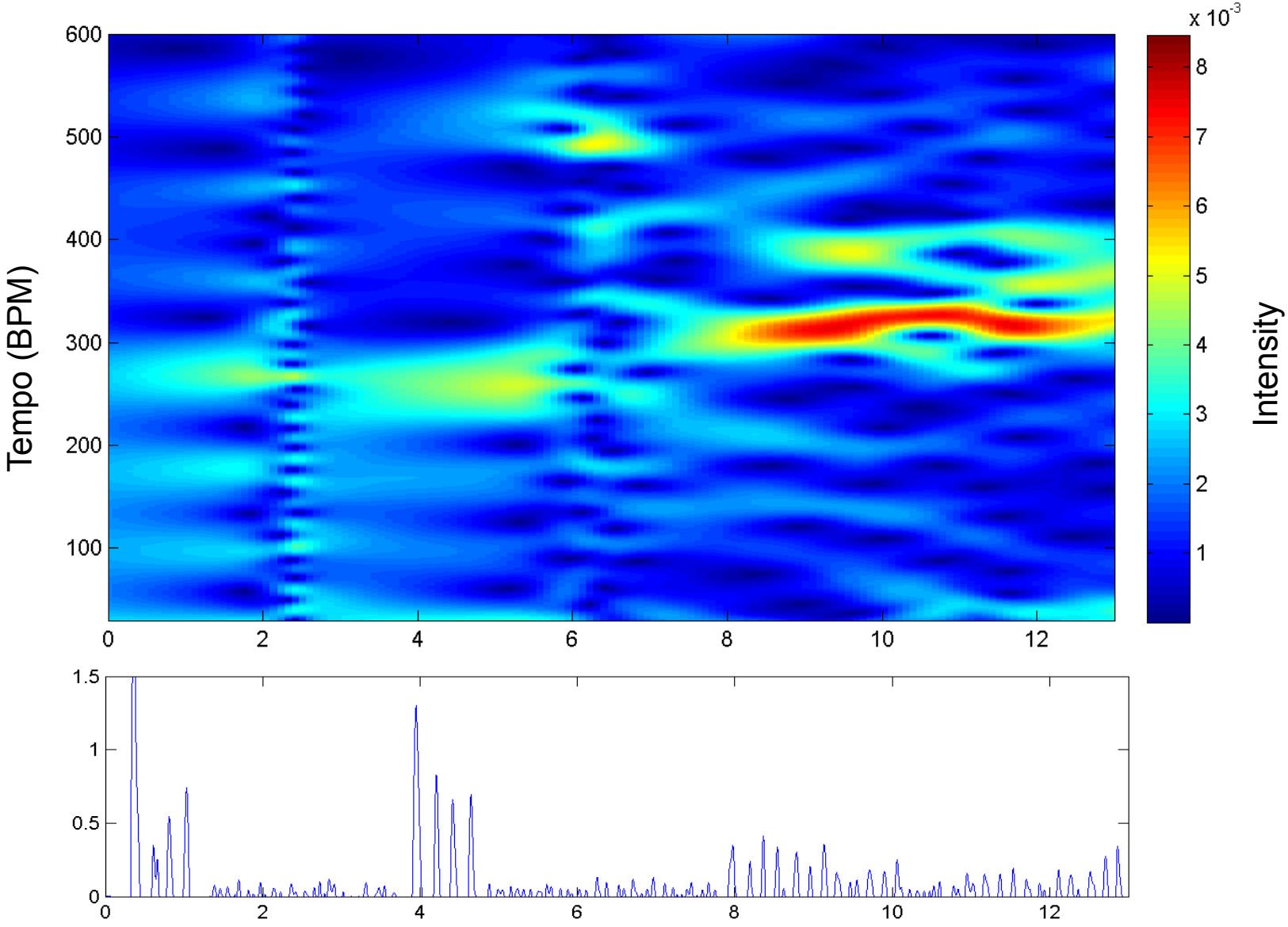
Steps:

1. Spectrogram
2. Log Compression
3. Differentiation
4. Accumulation
5. Normalization

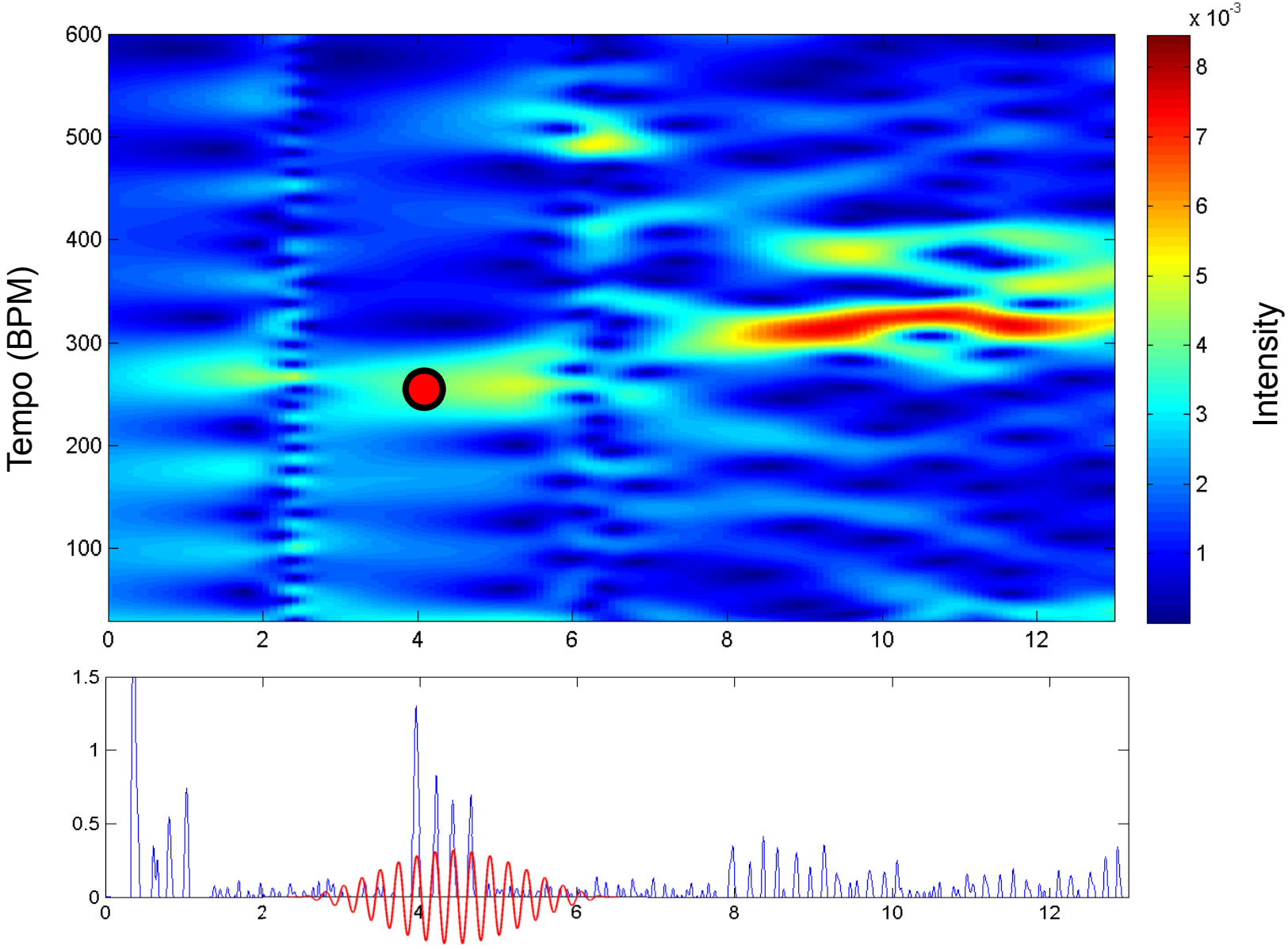
Novelty Curve



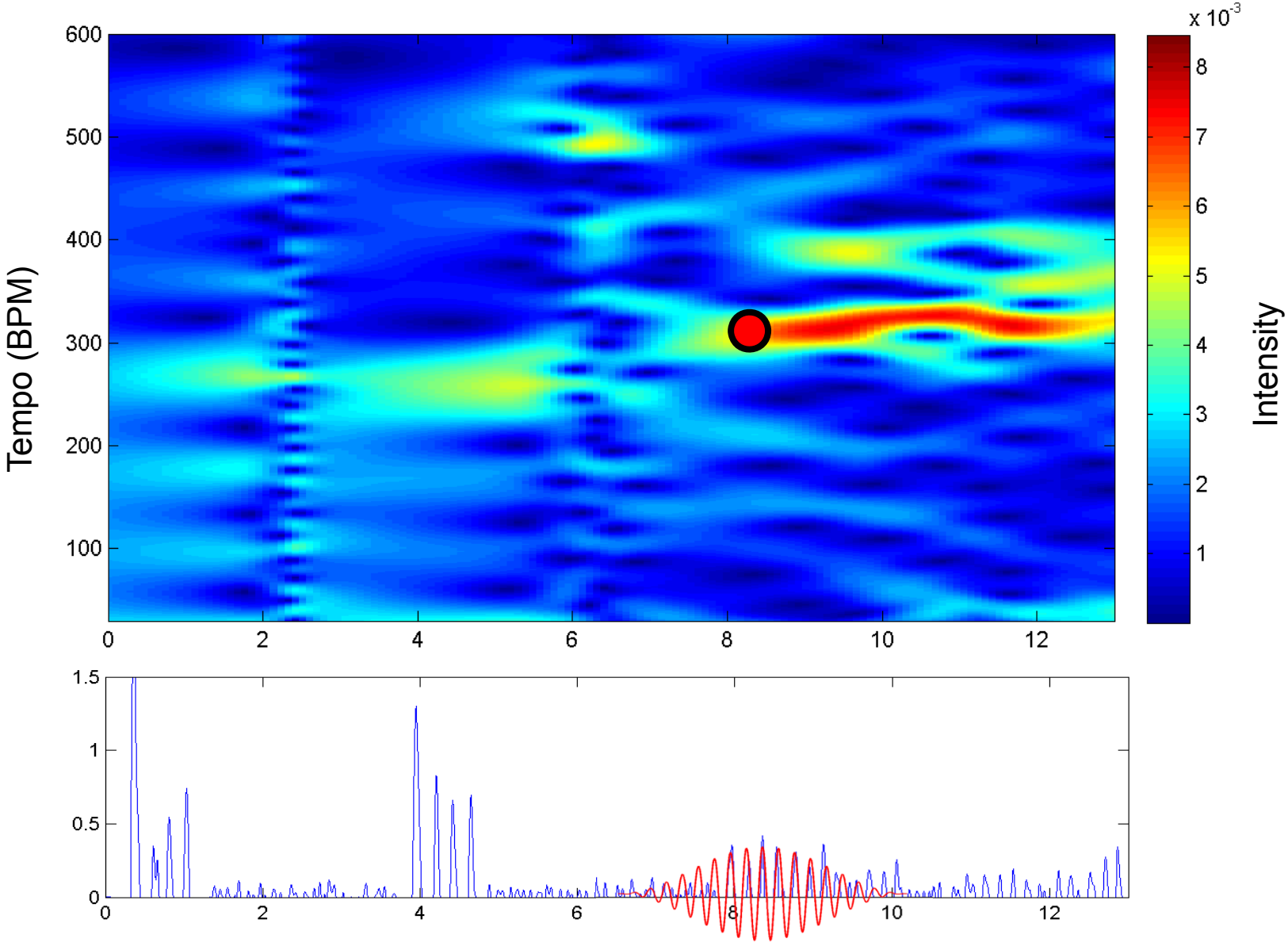
Beat Tracking



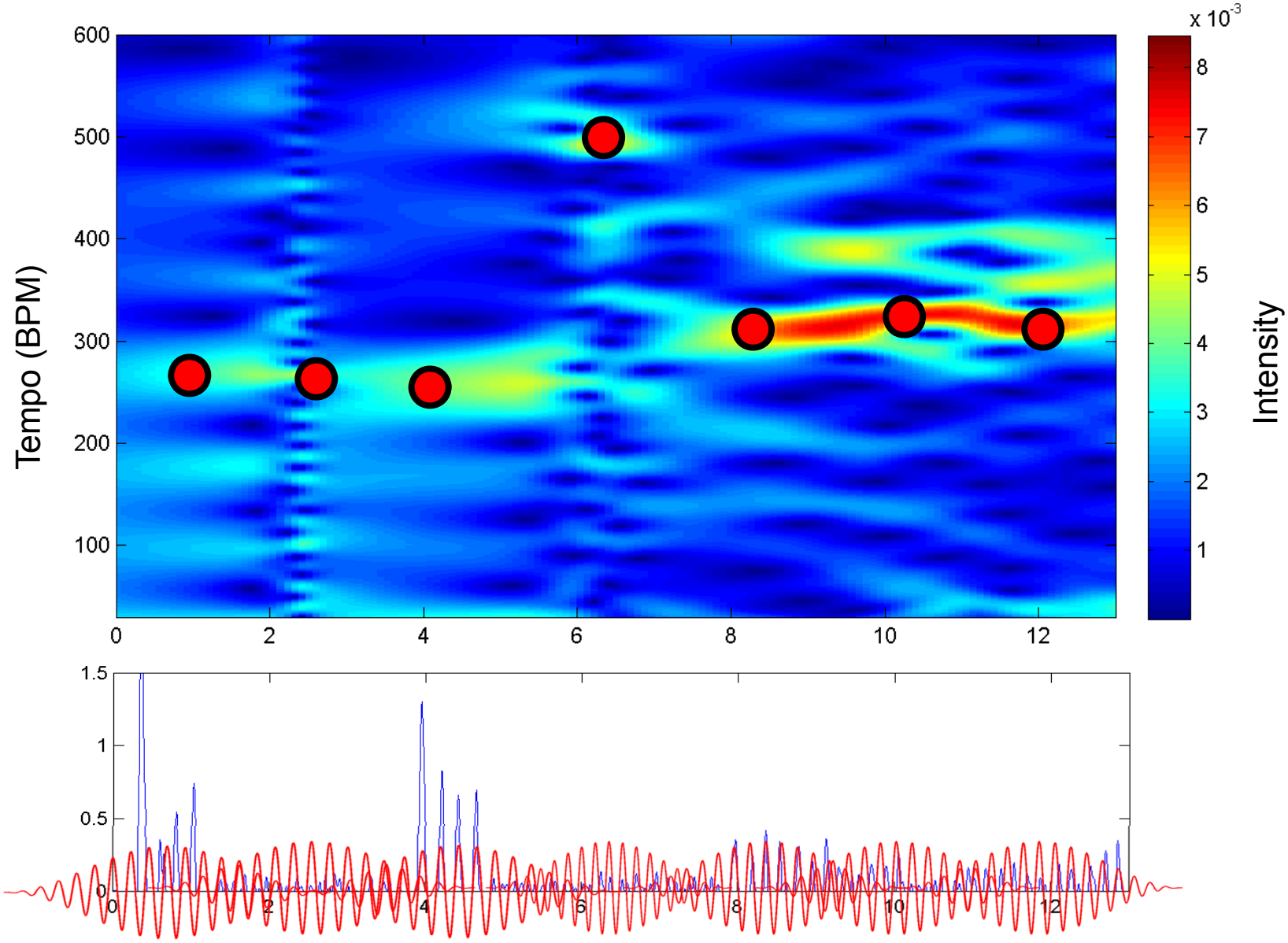
Beat Tracking



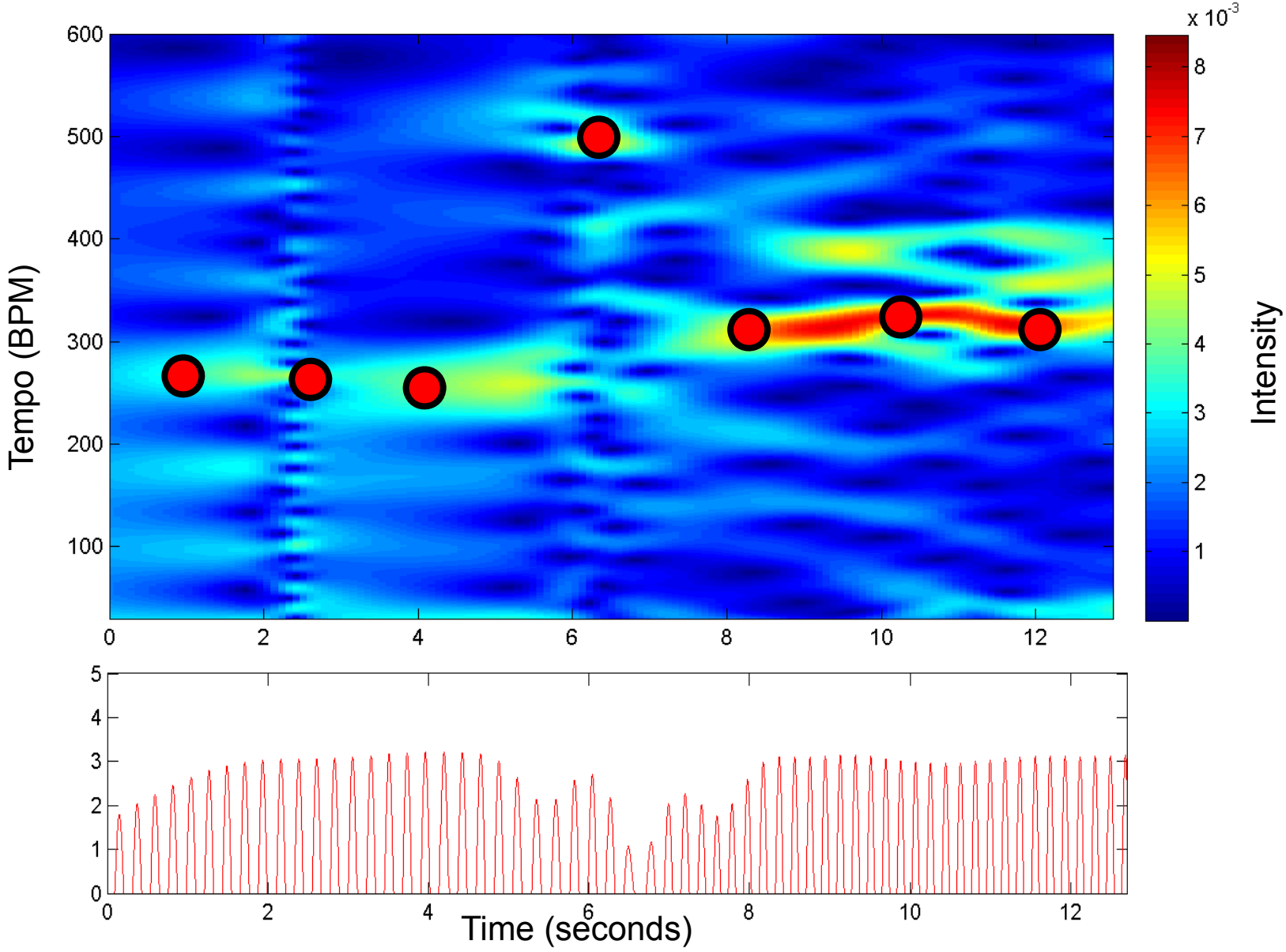
Beat Tracking



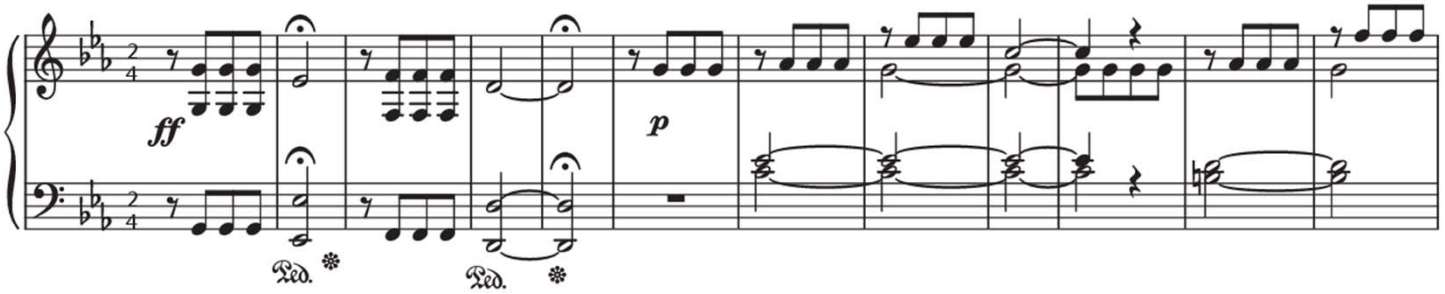
Beat Tracking



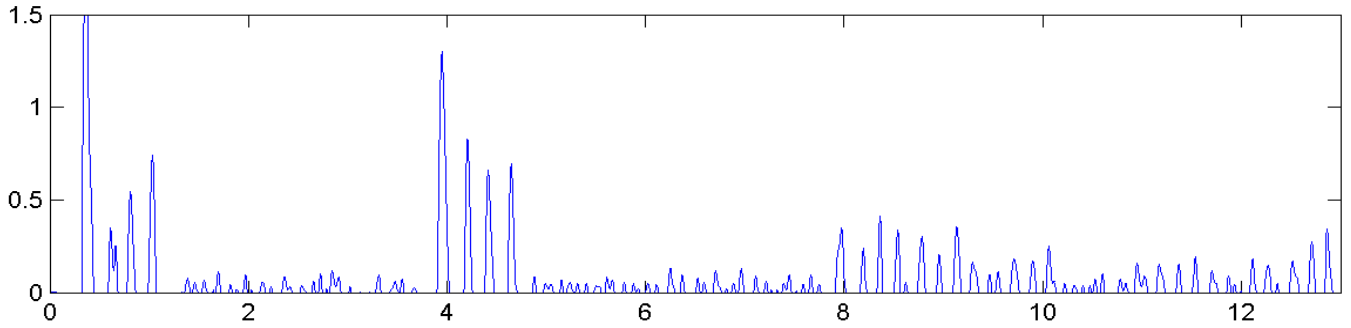
Beat Tracking



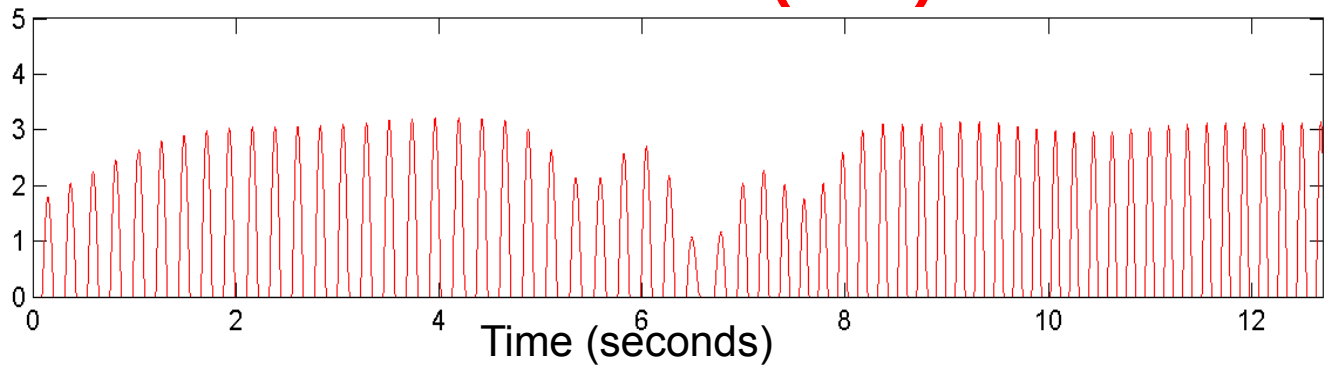
Beat Tracking



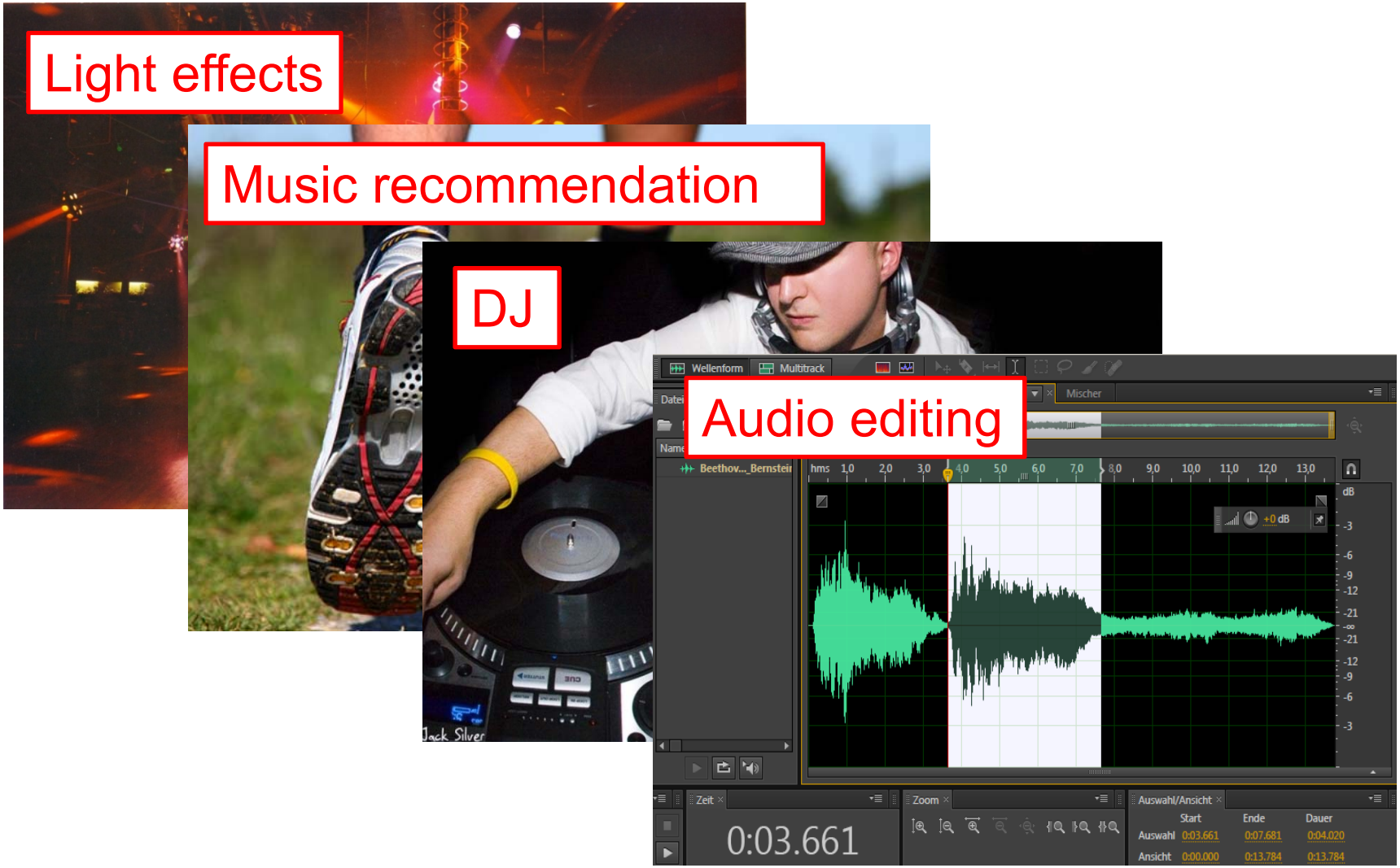
Novelty Curve



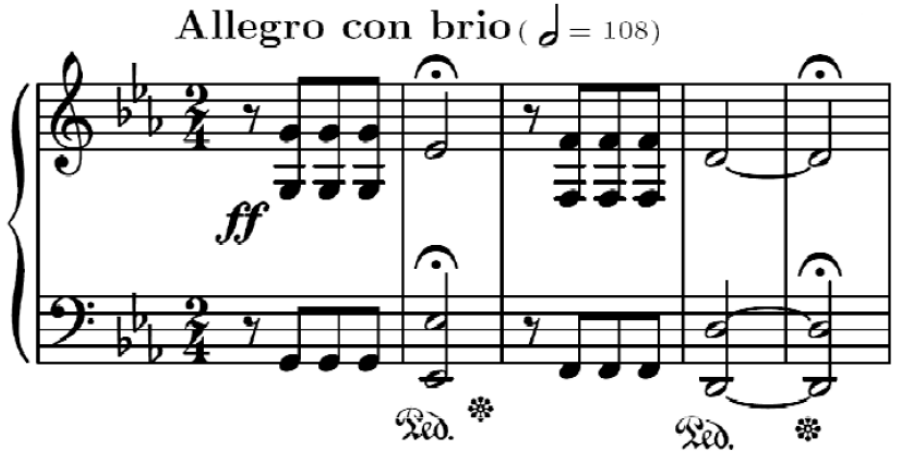
Predominant Local Pulse (PLP)



Beat Tracking

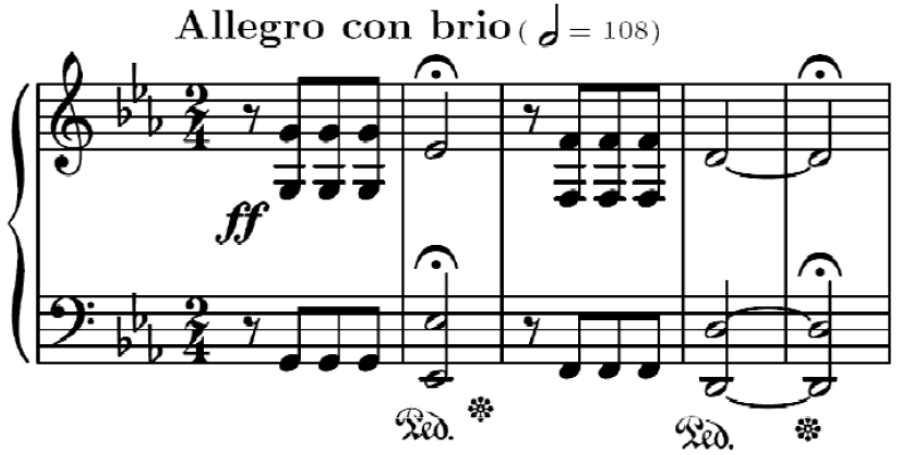


Motivic Similarity



Beethoven’s Fifth (1st Mov.)

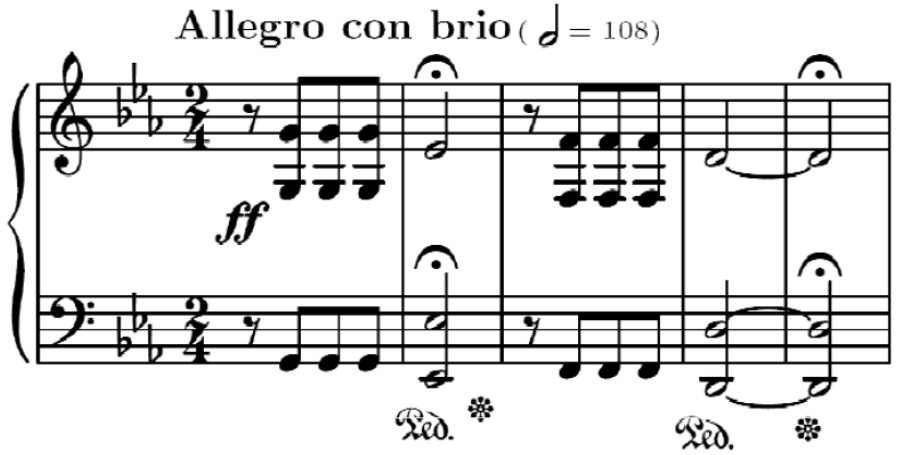
Motivic Similarity



Beethoven’s Fifth (1st Mov.)

Beethoven’s Fifth (3rd Mov.)

Motivic Similarity

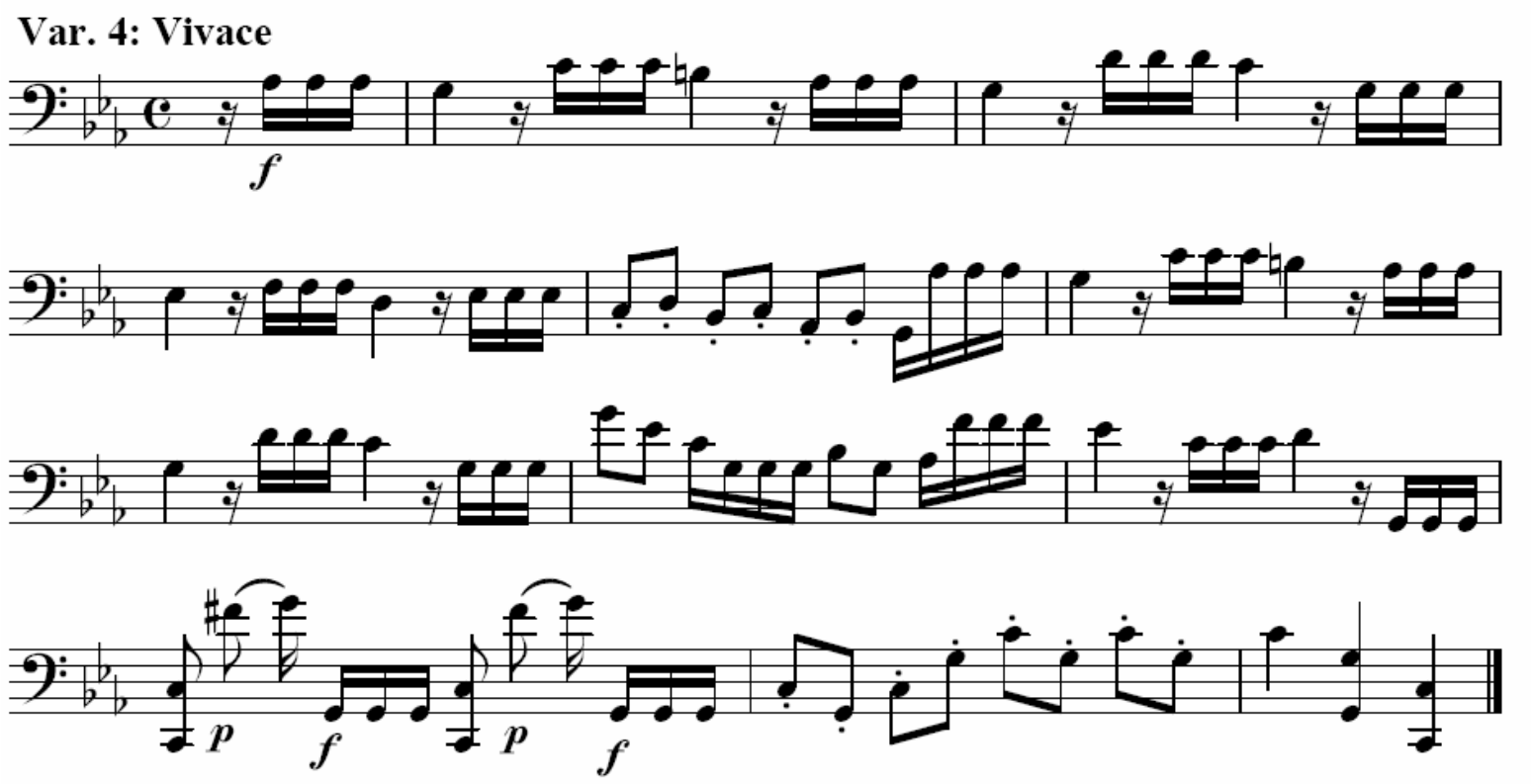


Beethoven’s Fifth (1st Mov.)

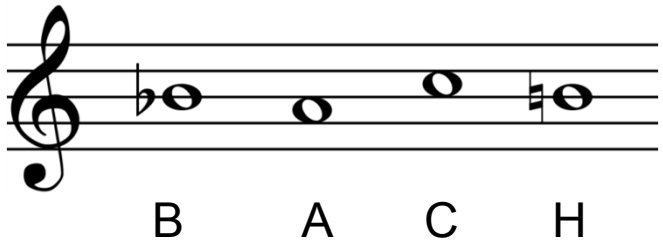
Beethoven’s Fifth (3rd Mov.)

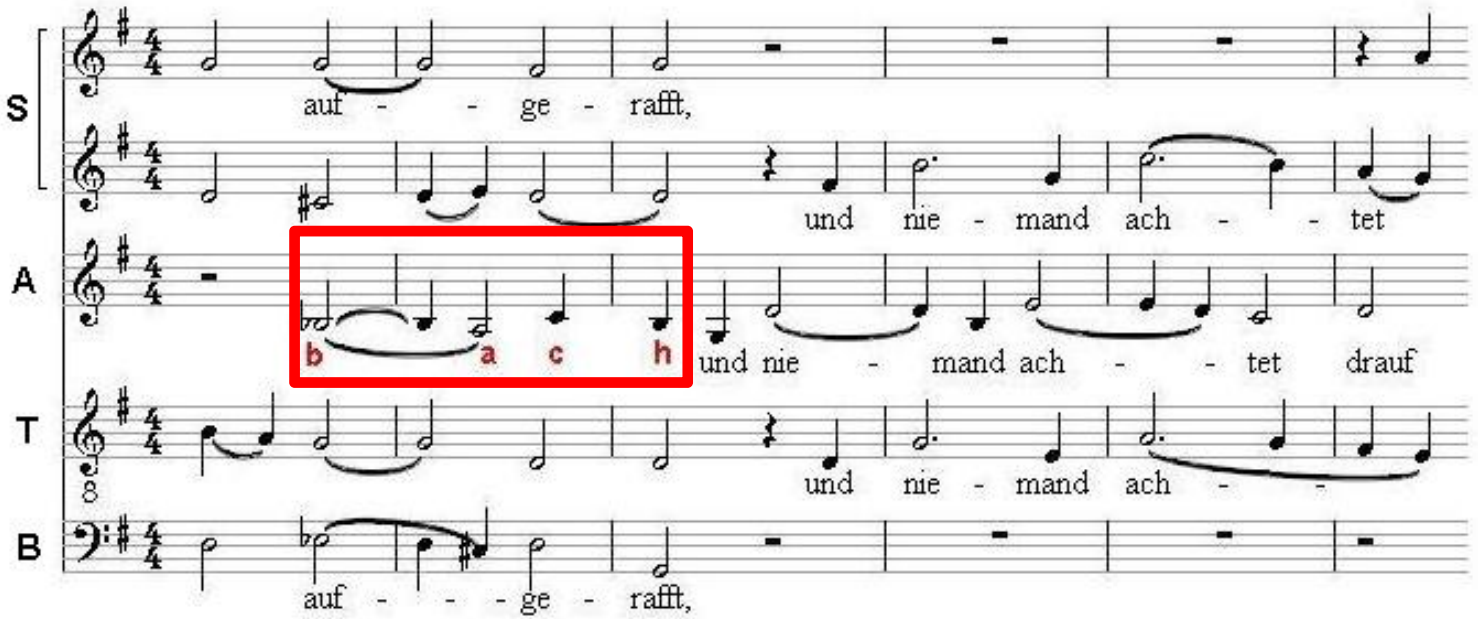
Beethoven’s Appassionata

Motivic Similarity

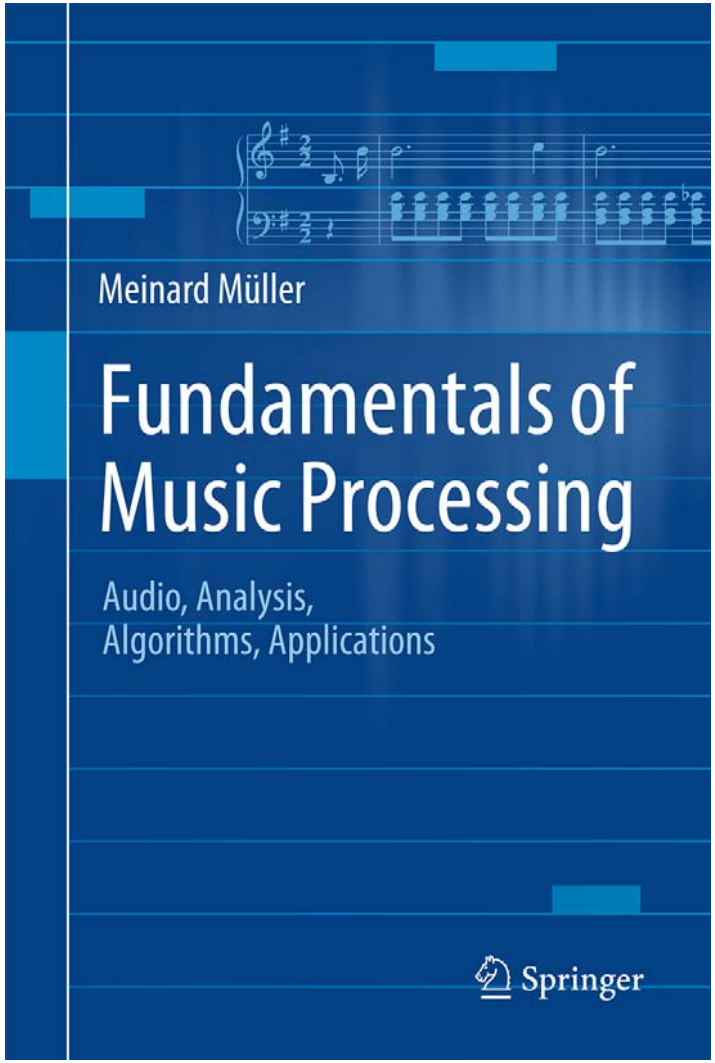


Motivic Similarity





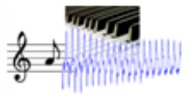
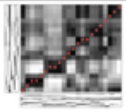
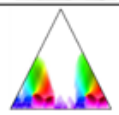
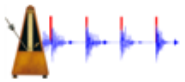
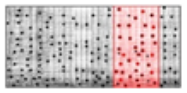
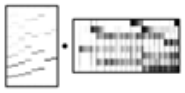
Book: Fundamentals of Music Processing



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

Accompanying website:
www.music-processing.de

Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Represenations
2		Fourier Analysis of Signals
3		Music Synchronization
4		Music Structure Analysis
5		Chord Recognition
6		Tempo and Beat Tracking
7		Content-Based Audio Retrieval
8		Musically Informed Audio Decomposition

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

Accompanying website:
www.music-processing.de