

# 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

# Meinard Müller

- 2007 Habilitation  
Bonn University
- 2007 – 2012  
Senior Researcher  
Saarland University & MPI Informatik
- Since 2012  
Professor: Semantic Audio Processing  
Erlangen-Nürnberg University



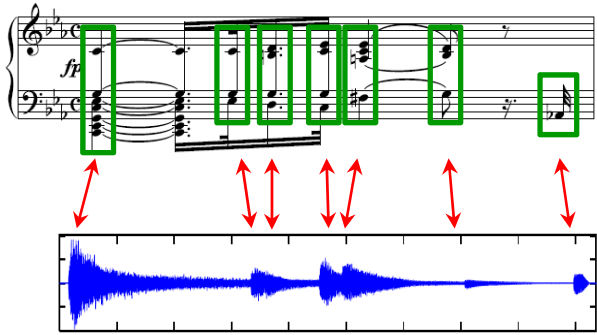
# Group Members

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

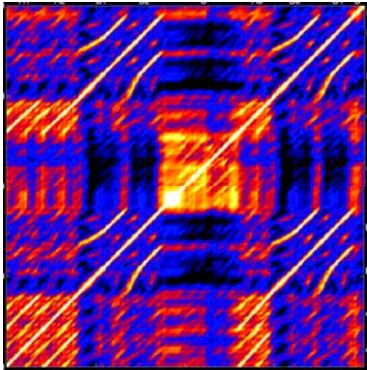


# Music Processing

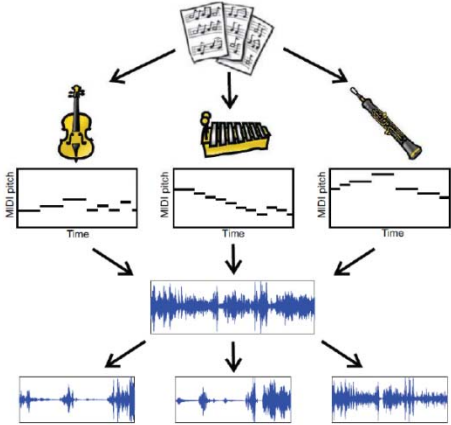
## Music Synchronization



## Structure Analysis



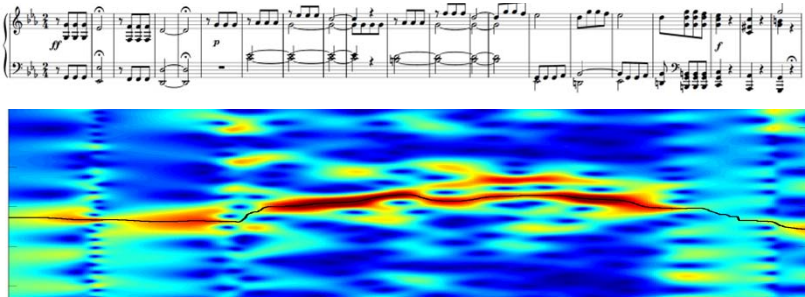
## Score-Informed Source Separation



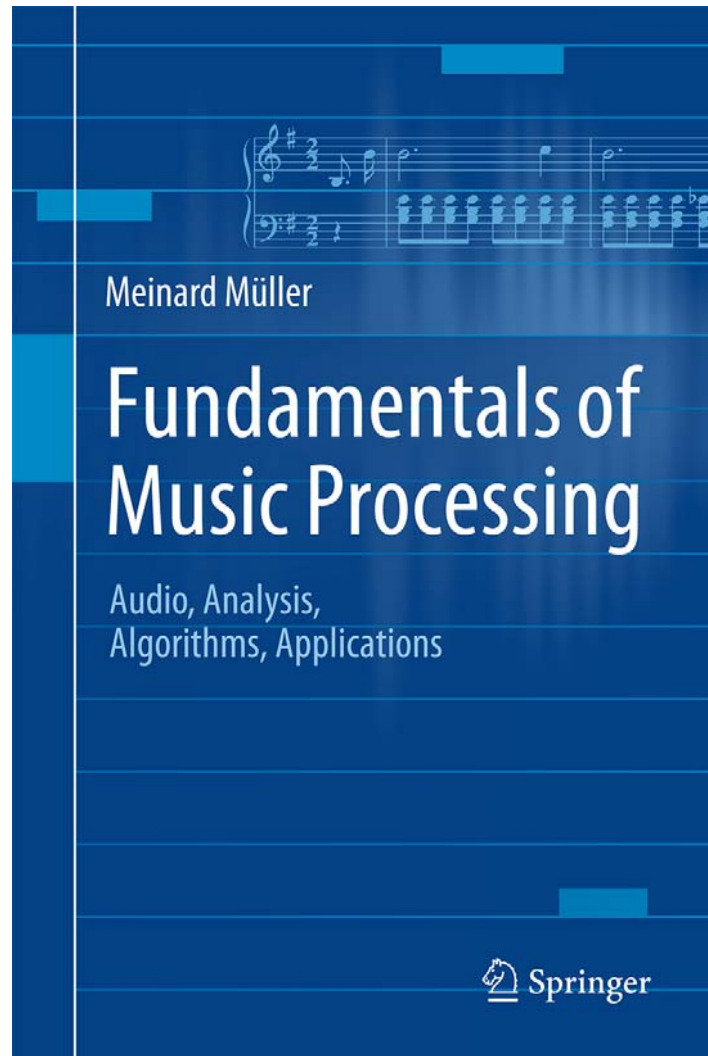
## User Interfaces



## Tempo Estimation and Beat Tracking



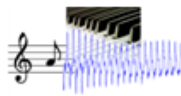

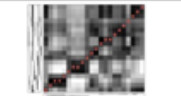
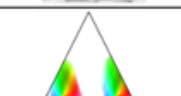

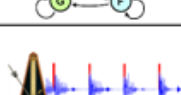


# Book: Fundamentals of Music Processing



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](http://www.music-processing.de)

# Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Representations
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

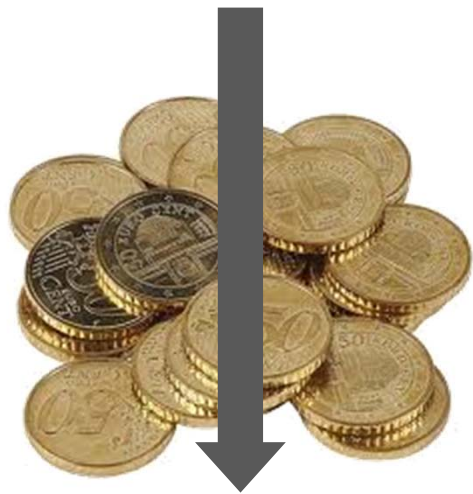
Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

# International Audio Laboratories Erlangen

 **Fraunhofer**  
IIS



**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



**Audio**

# International Audio Laboratories Erlangen

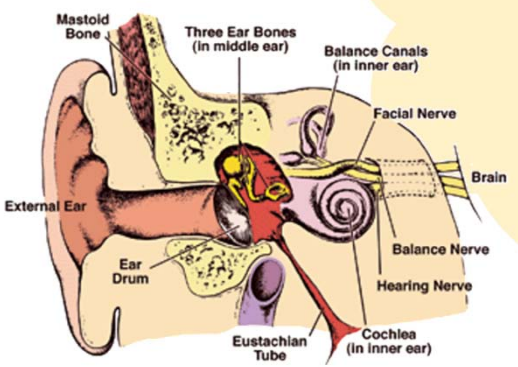
Audio Coding



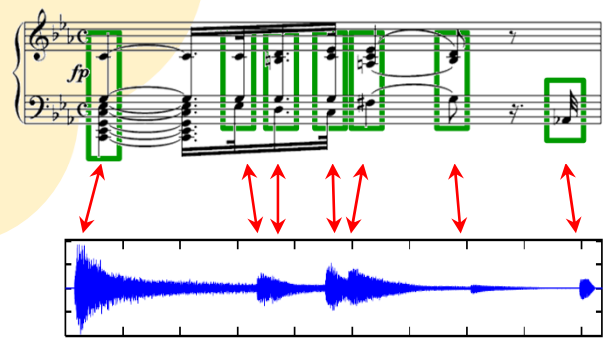
3D Audio



Audio



Psychoacoustics



Music Processing

---

Music

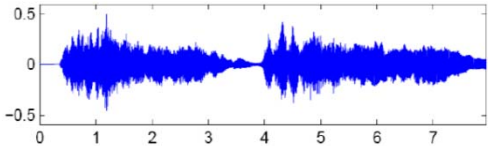


# Music Processing

Sheet Music (Image)



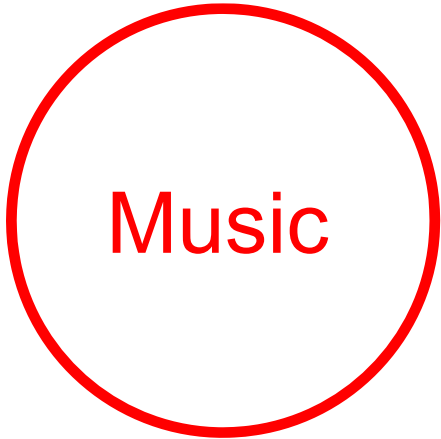
CD / MP3 (Audio)



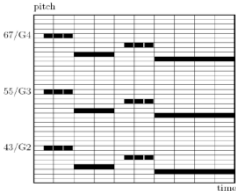
MusicXML (Text)

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

Dance / Motion (Mocap)



MIDI



Singing / Voice (Audio)



Music Film (Video)



Music Literature (Text)



# Research Goals

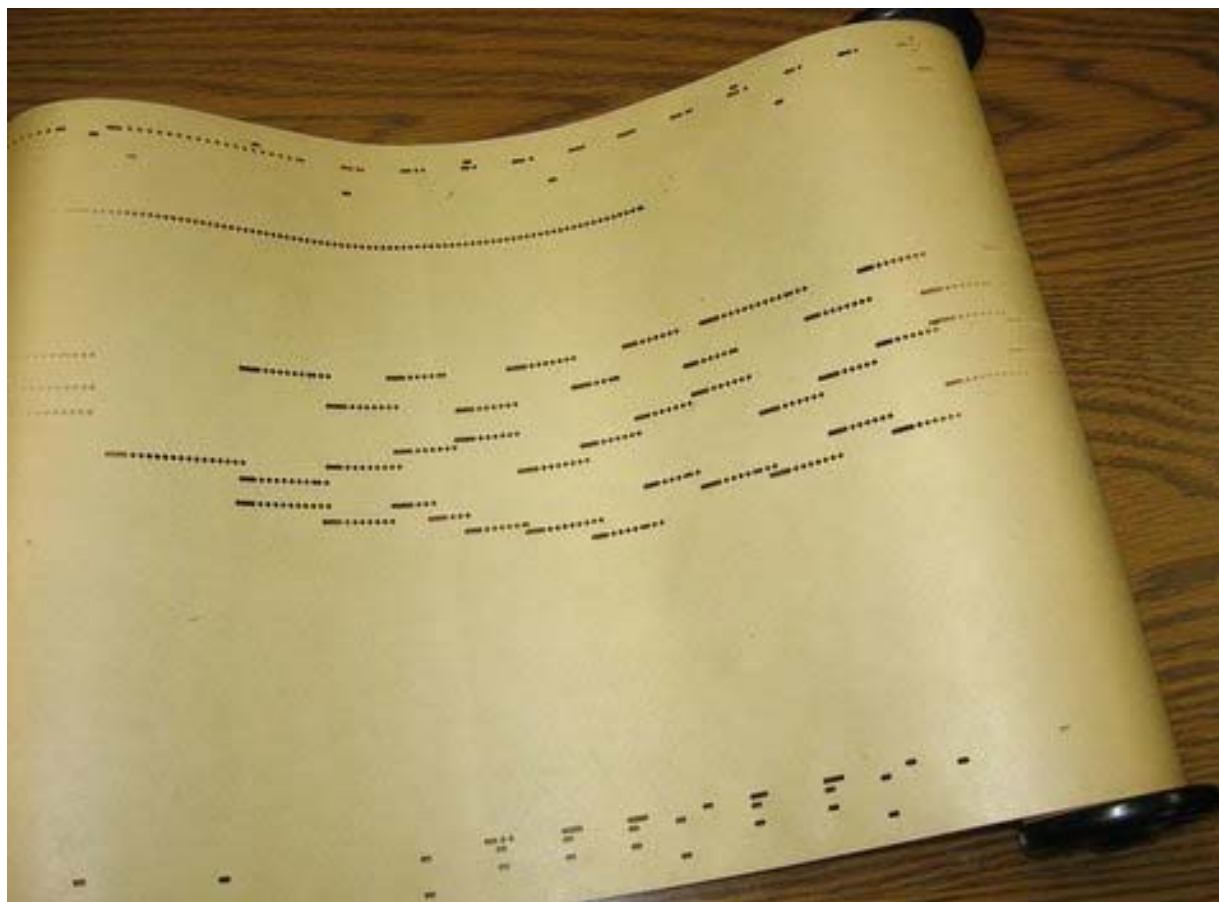
- Music Information Retrieval (MIR) → **ISMIR**
- Analysis of music signals  
(harmonic, melodic, rhythmic, motivic aspects)
- Design of musically relevant audio features
- Tools for multimodal search and interaction

---

# Overview

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

# Piano Roll Representation



# Player Piano (1900)

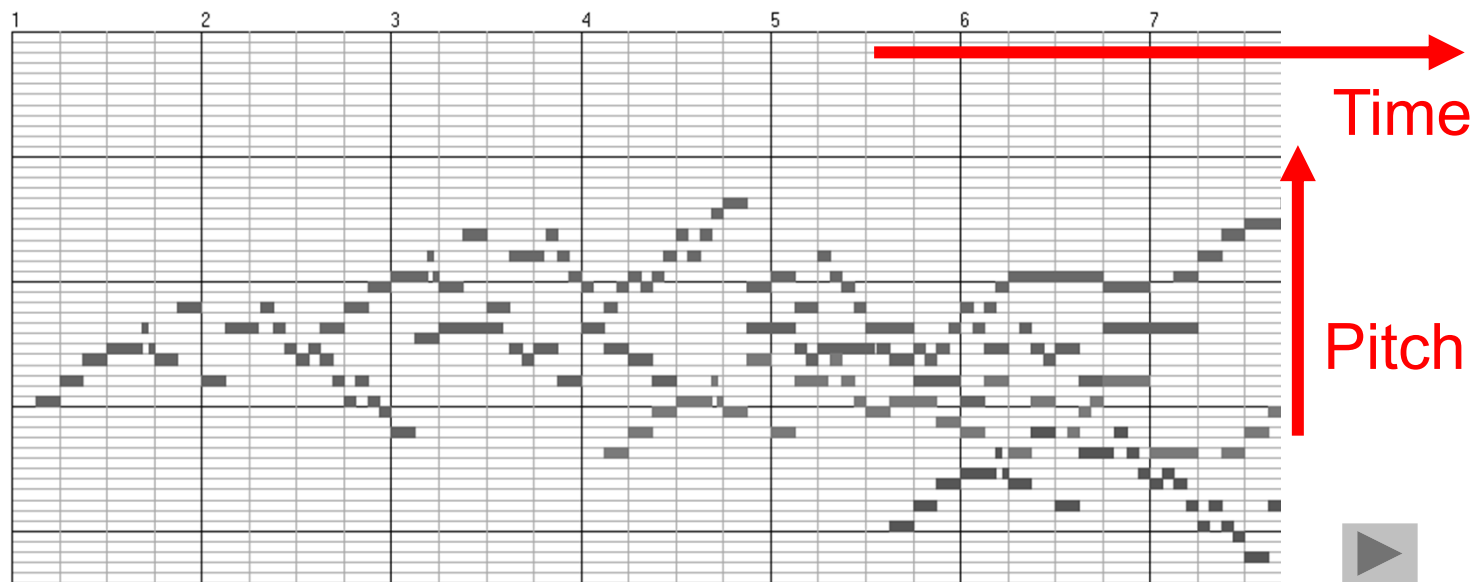
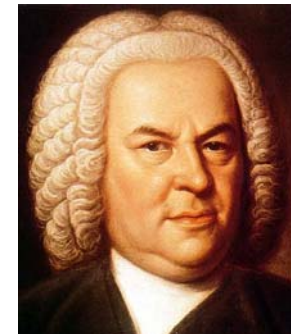




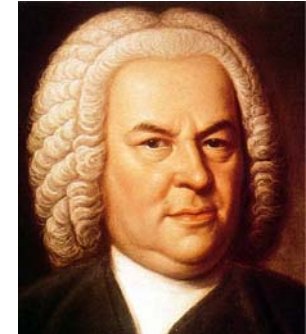
# Piano Roll Representation (MIDI)

J.S. Bach, C-Major Fuge

(Well Tempered Piano, BWV 846)



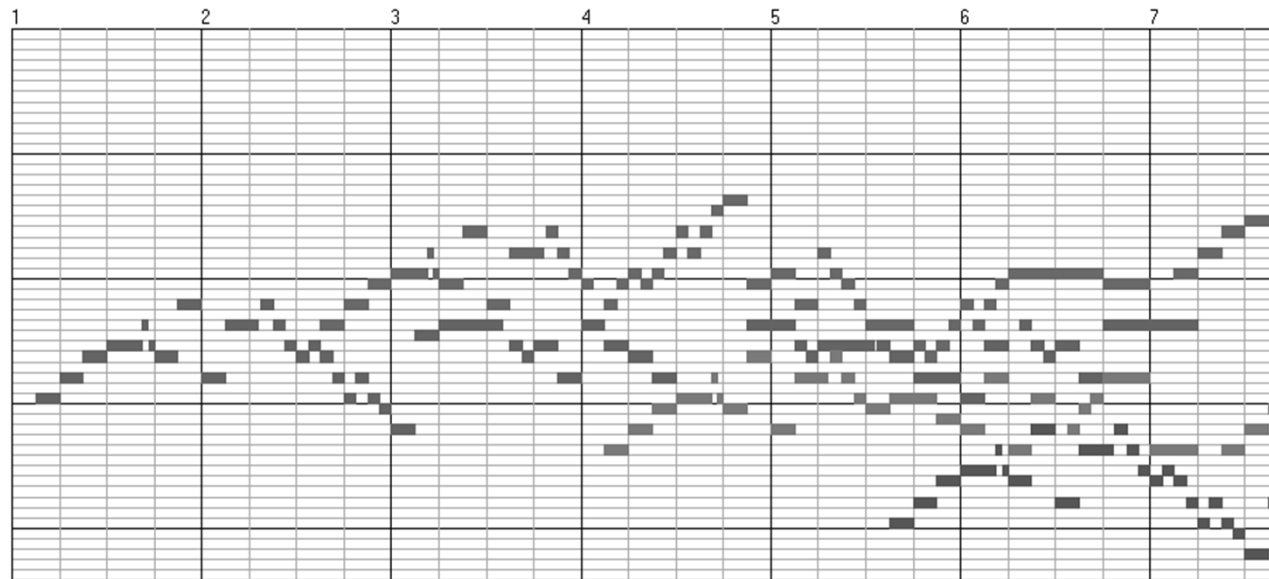
# Piano Roll Representation (MIDI)



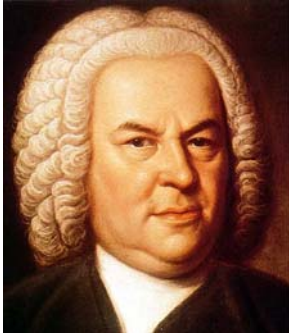
Query:



Goal: Find all occurrences of the query



# Piano Roll Representation (MIDI)

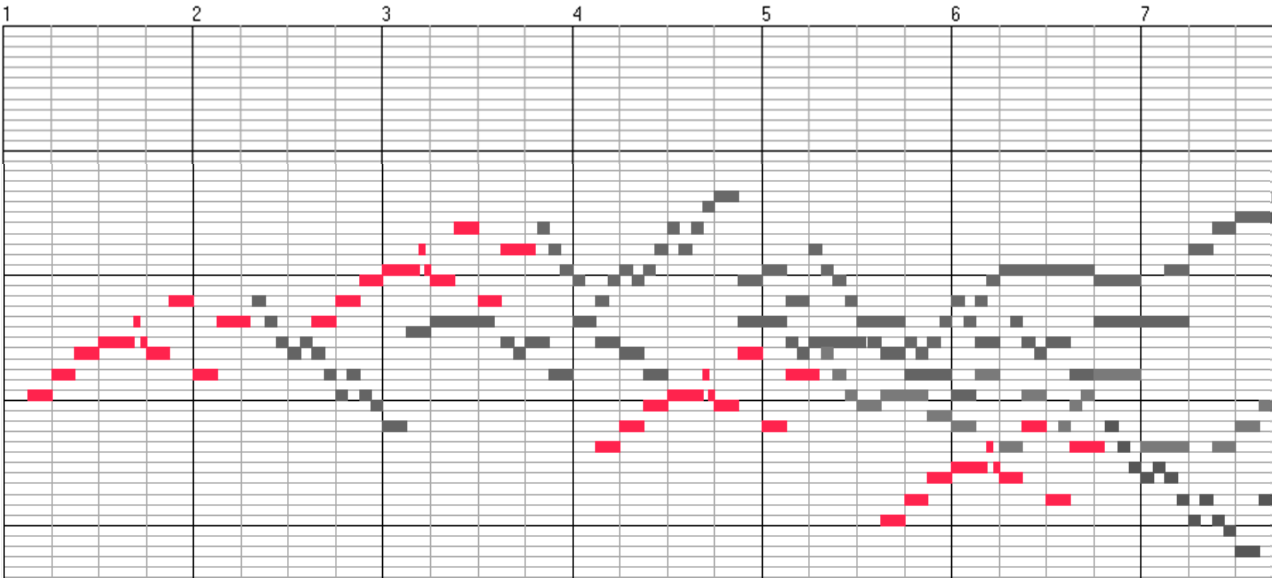


Query:

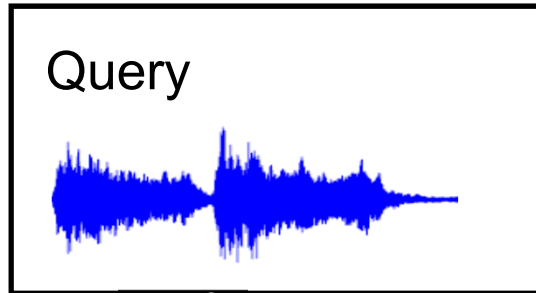


Goal: Find all occurrences of the query

Matches:



# Music Retrieval

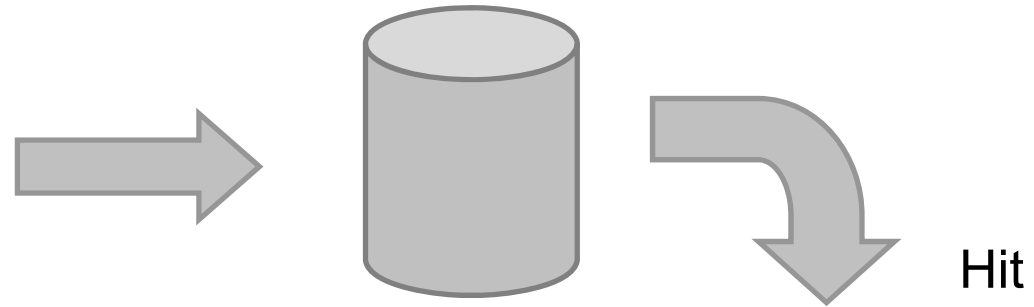


Audio-ID

Version-ID

Kategorie-ID

Database



Bernstein (1962)  
Beethoven, Symphony No. 5

Beethoven, Symphony No. 5:

- Bernstein (1962)
- Karajan (1982)
- Gould (1992)

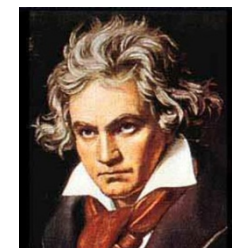


- Beethoven, Symphony No. 9
- Beethoven, Symphony No. 3
- Haydn Symphony No. 94



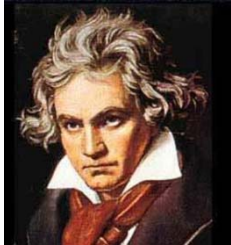
# Music Synchronization: Audio-Audio

Beethoven's Fifth

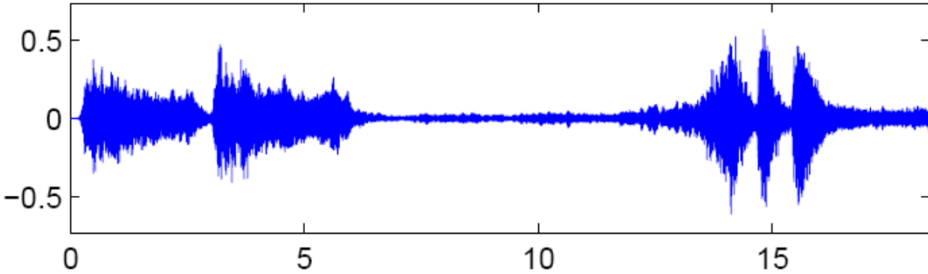


# Music Synchronization: Audio-Audio

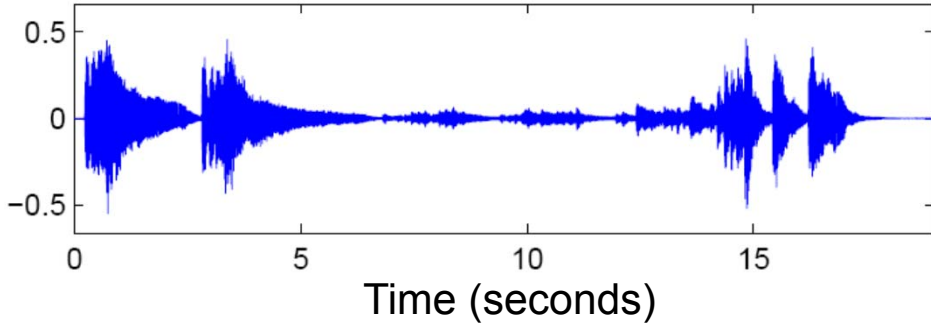
Beethoven's Fifth



Orchester  
(Karajan)

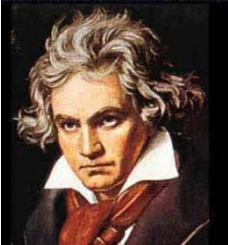


Piano  
(Scherbakov)

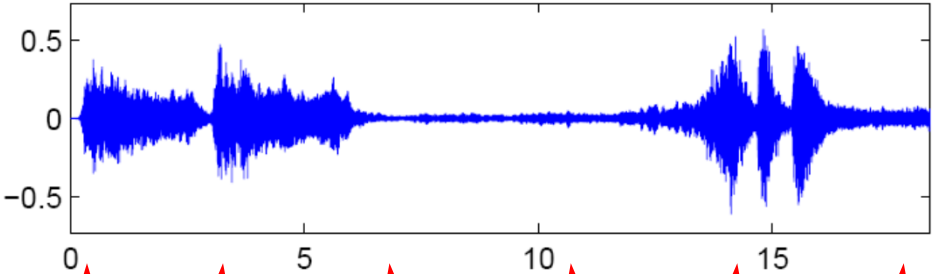


# Music Synchronization: Audio-Audio

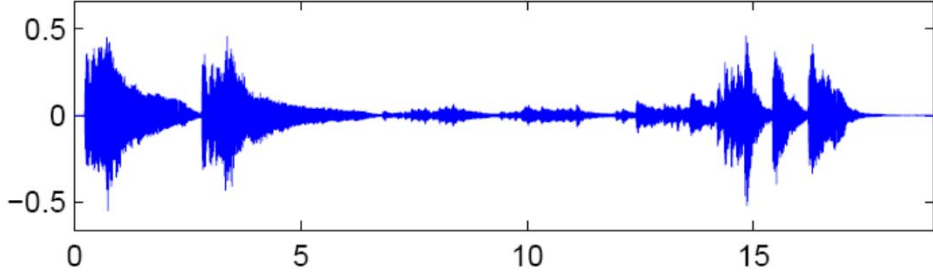
Beethoven's Fifth



Orchester  
(Karajan)



Piano  
(Scherbakov)



Time (seconds)

# Application: Interpretation Switcher

The screenshot shows the 'Interpretation Switcher' application window. The title bar reads 'Interpretation Switcher' and the subtitle is 'Beethoven, Op067-1\_Symphony5'. The main area displays four horizontal progress bars for different interpretations: 'midi', 'Bernstein', 'Sawallisch', and 'Scherbakov'. Each bar is divided into three colored segments: blue, red, and green. A white playhead marker is positioned on the red segment of each bar. The total duration for each interpretation is shown to the right of the bar: 'midi' (00:44.18), 'Bernstein' (01:00.64), 'Sawallisch' (00:58.35), and 'Scherbakov' (00:52.45). On the right side, there is a list of checkboxes: 'midi', 'Bernstein', 'Sawallisch', and 'Scherbakov', all of which are checked. Below this list is a 'Deselect all' button. At the bottom of the window, there is a control bar with the following elements: a dropdown menu showing 'Absolute' with a checkmark, 'Relative', and 'Reference'; three icons (pause, stop, and a triangle); the text 'Movement selection'; a checkbox for 'Interval Repeat'; and an 'Info' button with a question mark icon. A large grey play button is visible on the right side of the slide.



# Music Synchronization: Image-Audio

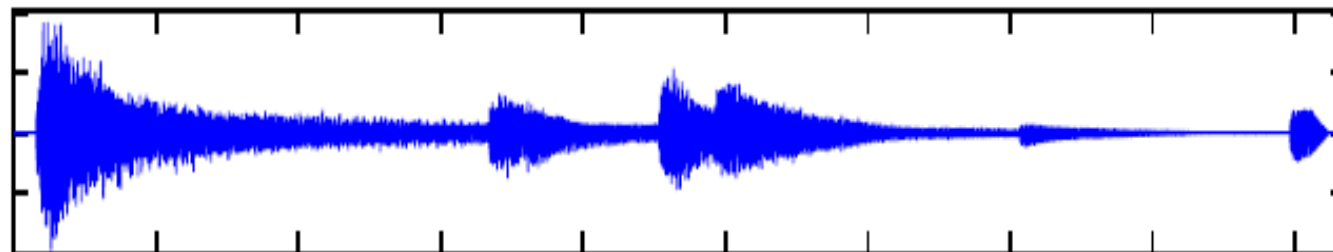
Image

Grave.

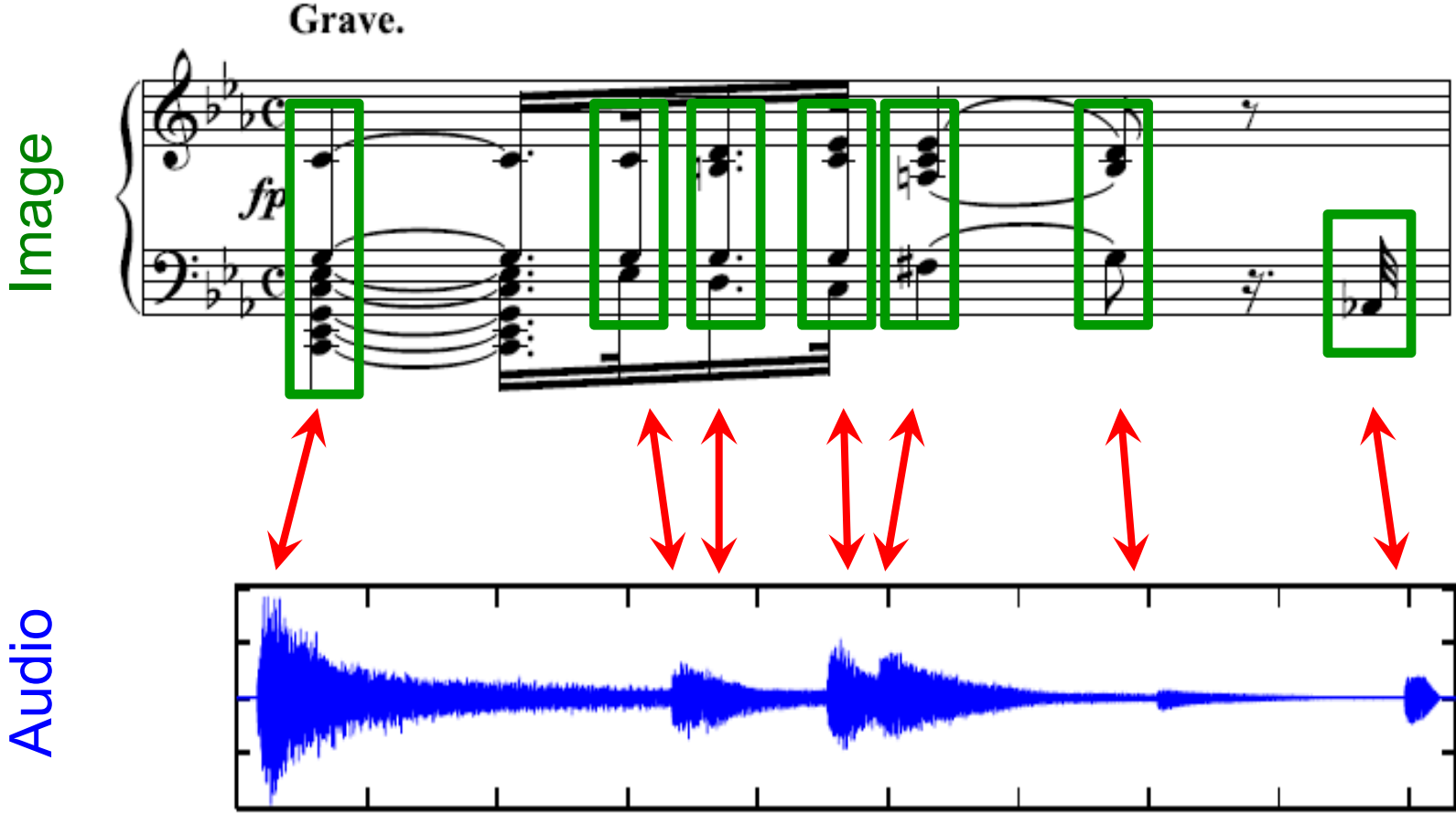


The image shows a musical score for piano, marked "Grave." and "fp". The score is written in G major (one sharp) and common time (C). It consists of two staves: a treble clef staff and a bass clef staff. The music features a slow, somber tempo with a focus on sustained chords and melodic lines. The treble staff begins with a half note chord (G4, B4, D5) followed by a series of chords and a melodic line. The bass staff features a series of chords and a melodic line that complements the treble staff. The overall mood is one of solemnity and gravity.

Audio



# Music Synchronization: Image-Audio

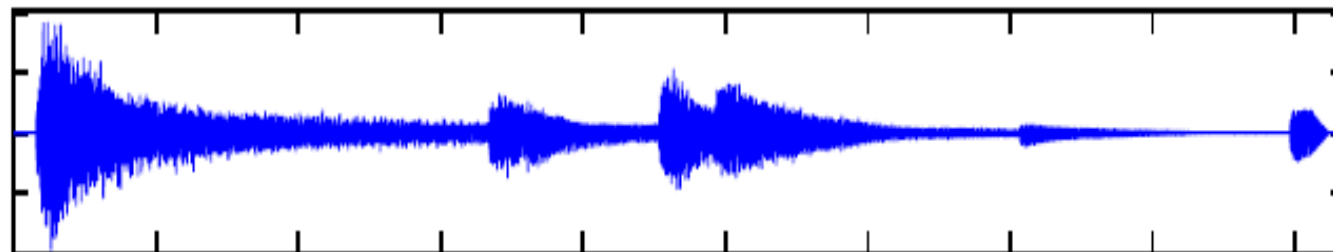


# How to make the data comparable?

Image



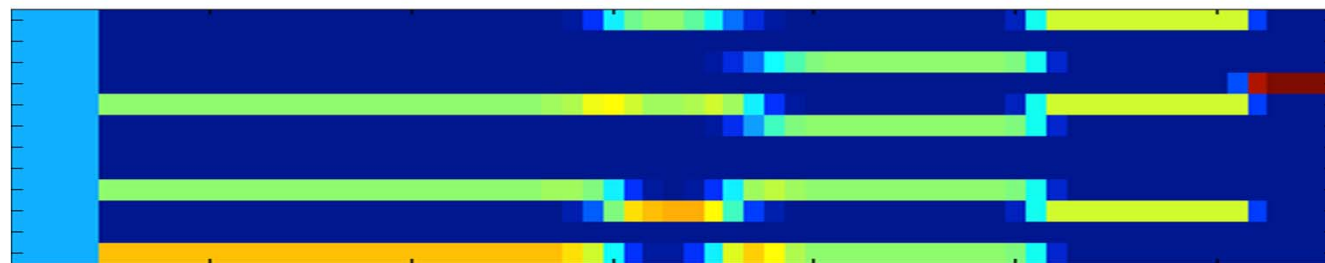
Audio



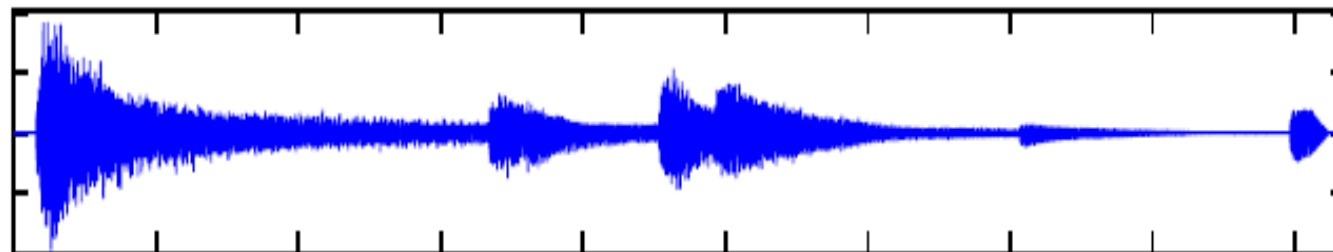
# How to make the data comparable?

## Image Processing: Optical Music Recognition

Image



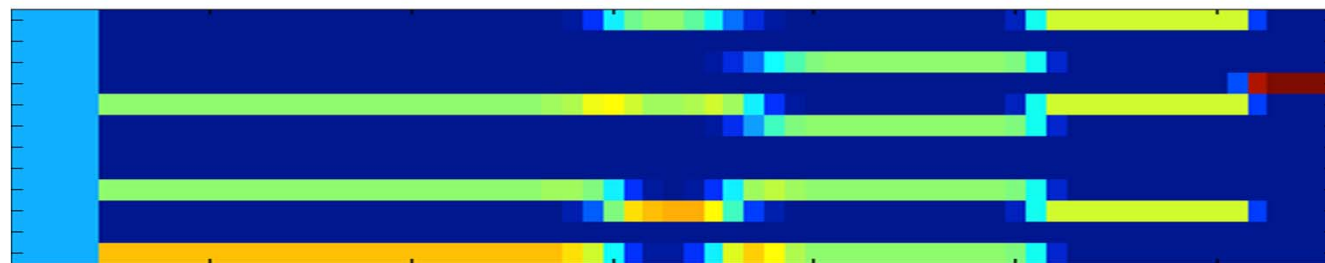
Audio



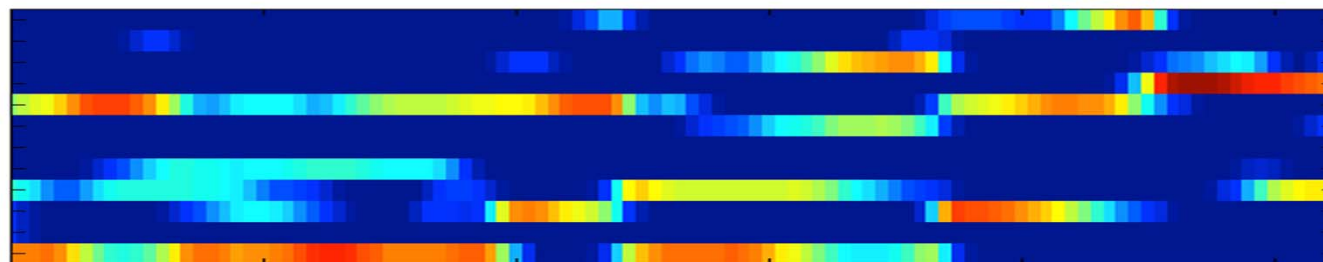
# How to make the data comparable?

## Image Processing: Optical Music Recognition

Image



Audio

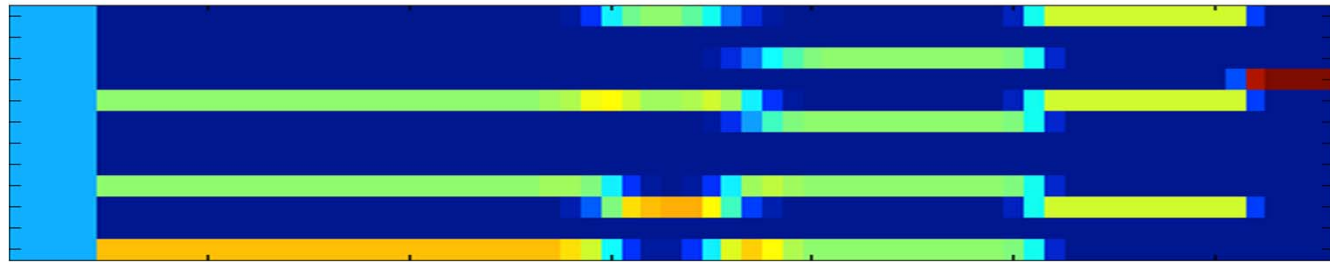


## Audio Processing: Fourier Analyse

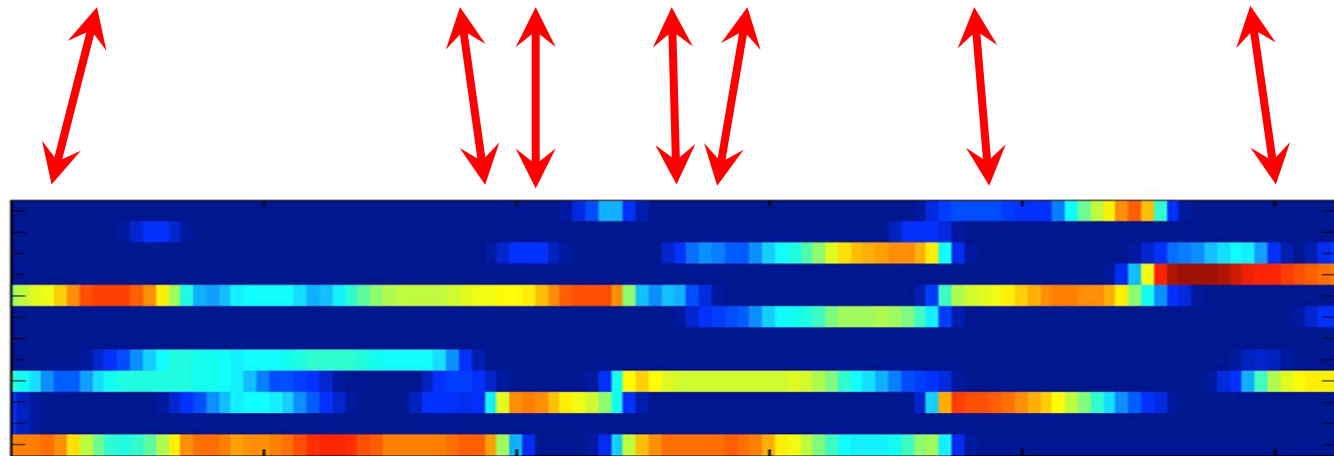
# How to make the data comparable?

## Image Processing: Optical Music Recognition

Image



Audio




## Audio Processing: Fourier Analyse



# Application: Score Viewer

AudioViewer

Beethoven - Complete Piano Sonatas - Daniel Barenboim



Disc 3

01 Sonata no.7 in D major, op.10 no.3: Presto	7:08
02 Sonata no.7 in D major, op.10 no.3: Largo e mesto	10:02
03 Sonata no.7 in D major, op.10 no.3: Menuetto (Allegro)	2:53
04 Sonata no.7 in D major, op.10 no.3: Rondo (Allegro)	4:05
05 Sonata no.8 in C minor, op.13, "Pathetique" / Allegro di molto e con brio	9:32
06 Sonata no.8 in C minor, op.13, "Pathetique" / Adagio cantabile	5:19
07 Sonata no.8 in C minor, op.13, "Pathetique" / Rondo (Allegro)	4:53
08 Sonata no.9 in E major, op.14 no.1: Allegro	6:48
09 Sonata no.9 in E major, op.14 no.1: Allegretto	4:16
10 Sonata no.9 in E major, op.14 no.1: Adagio	


Disc: 3 / 10    Track: 7

ScoreViewer

Barenboim

Beethoven - Klaviersonaten Band 1 - Henle

Sonata no.8 in C minor, op.13, "Pathetique" / Rondo (Allegro)



Track: 29 / 54    Bar: 9 / 211    Page: 159 / 285

Score Following Off    Play    Stop



# 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

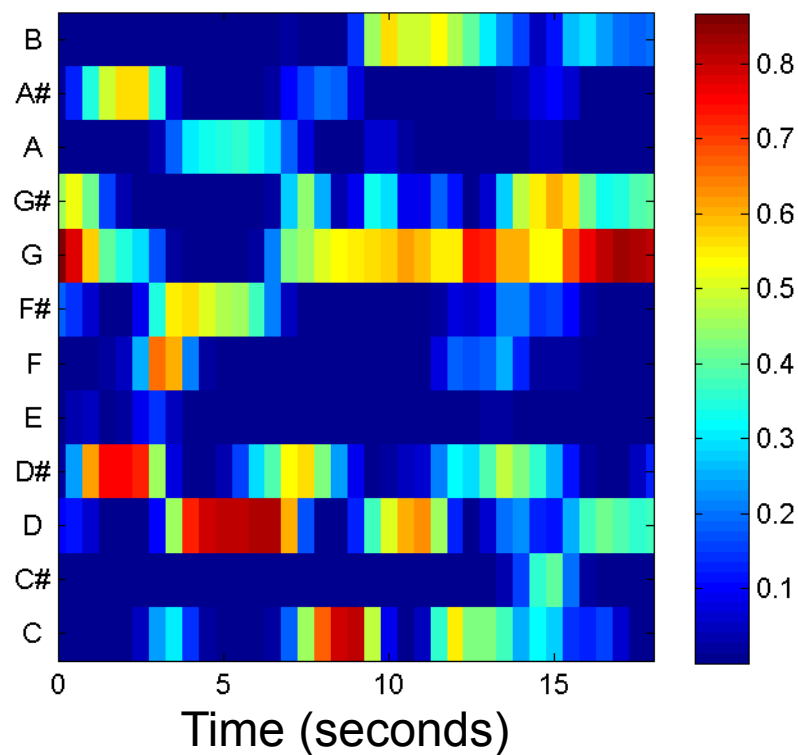
## 2.) Matching procedure

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

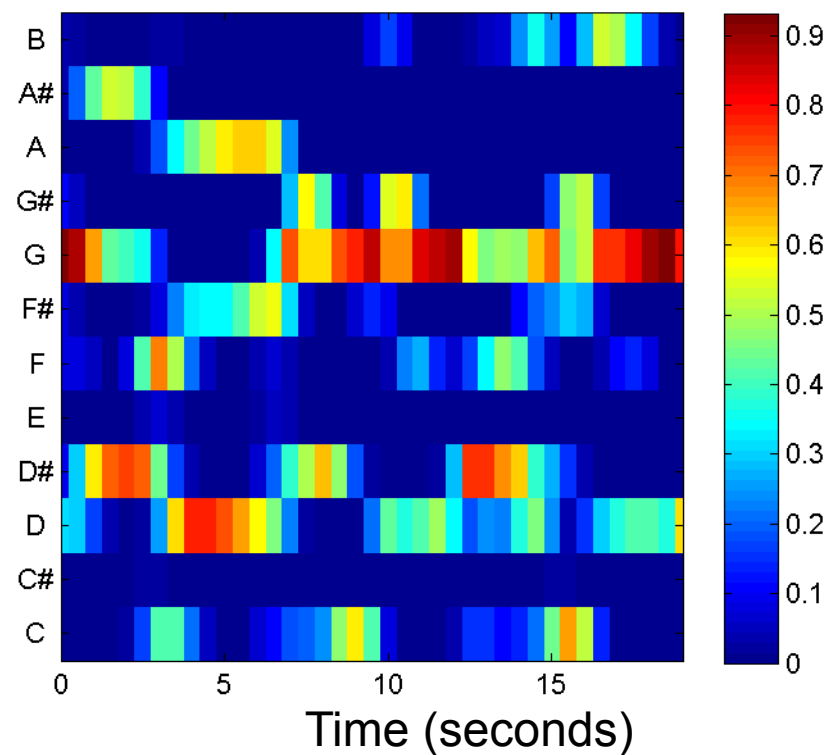
# Audio Features

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

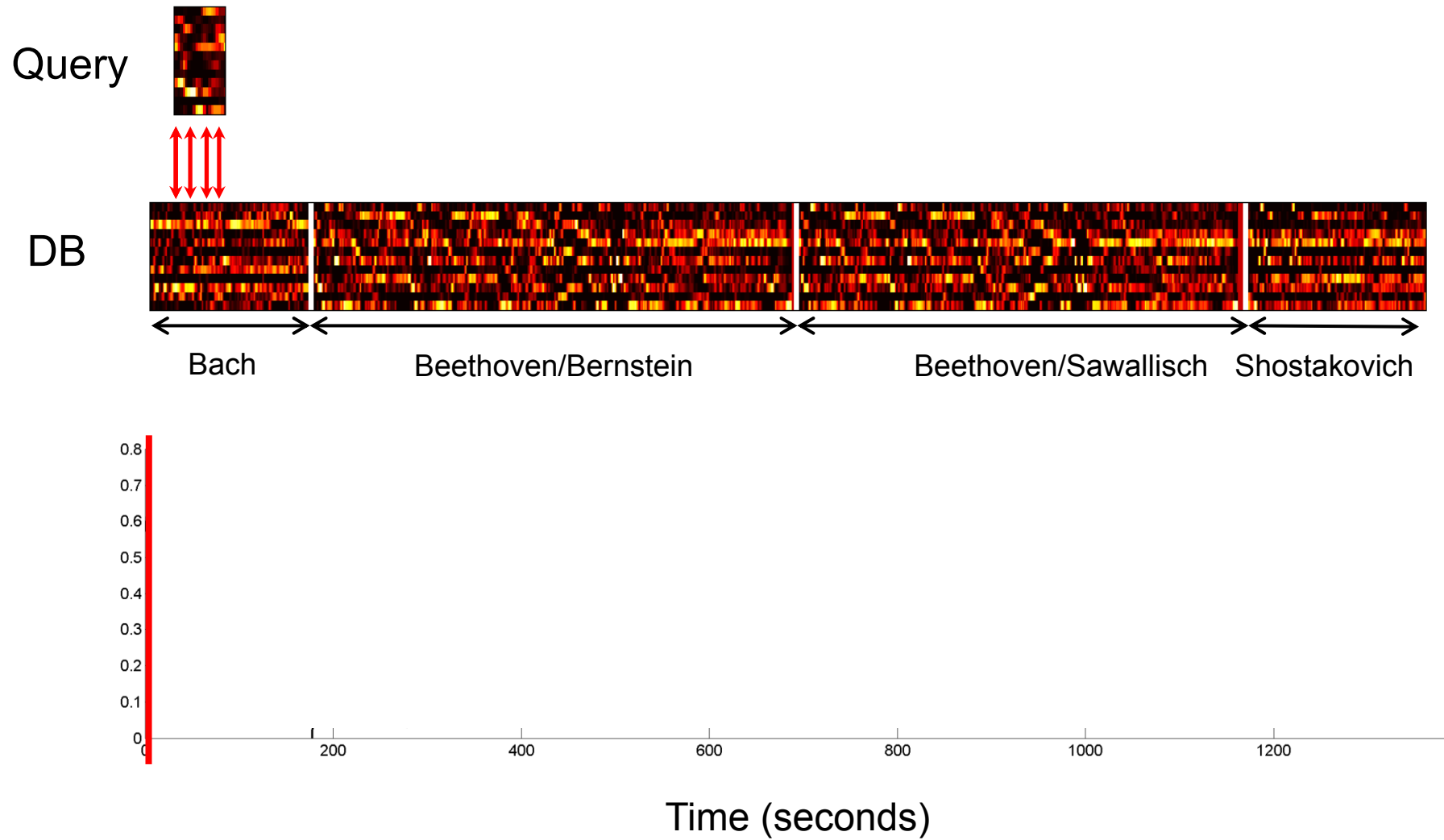
Karajan



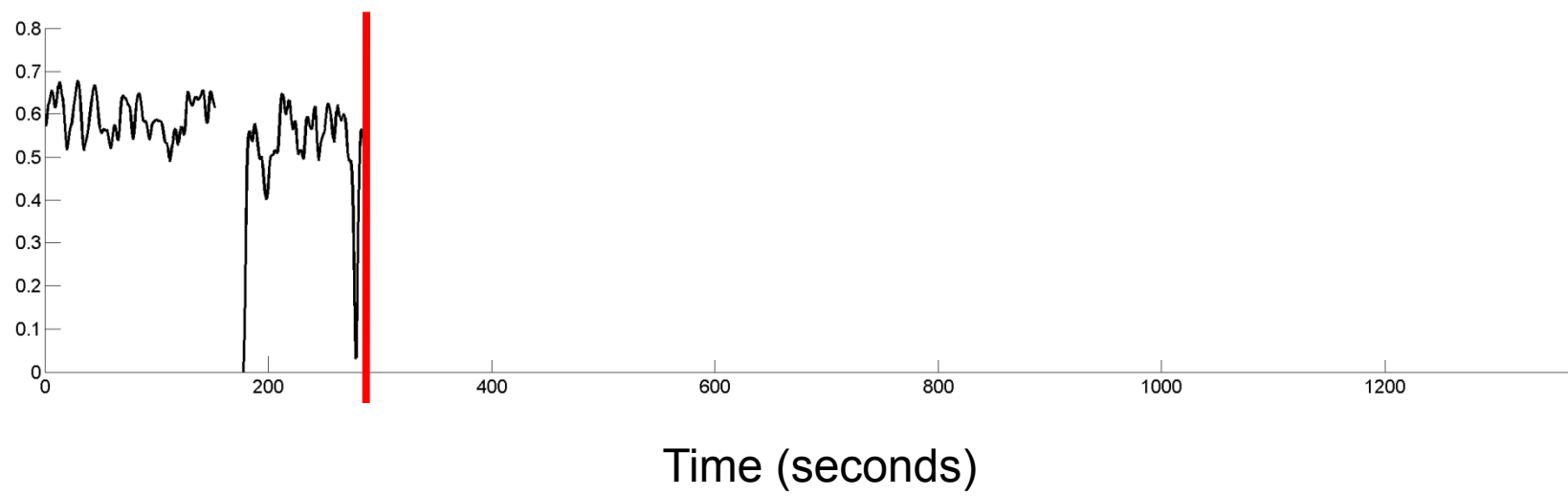
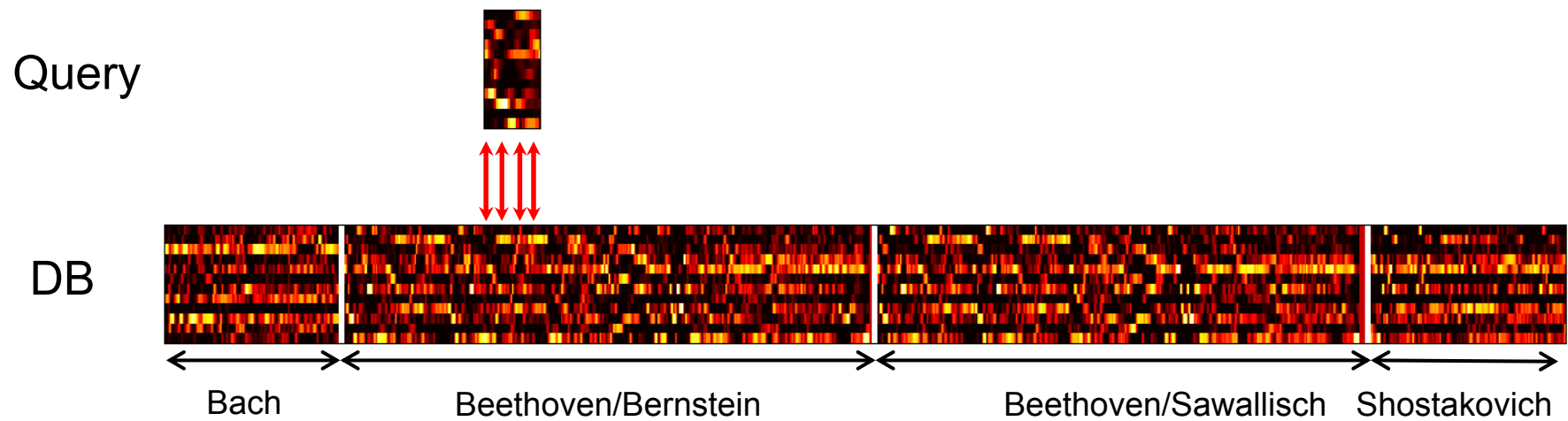
Scherbakov



# Matching Procedure



# Matching Procedure

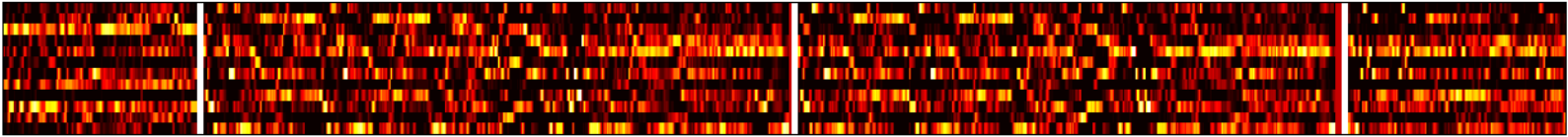


# Matching Procedure

Query



DB

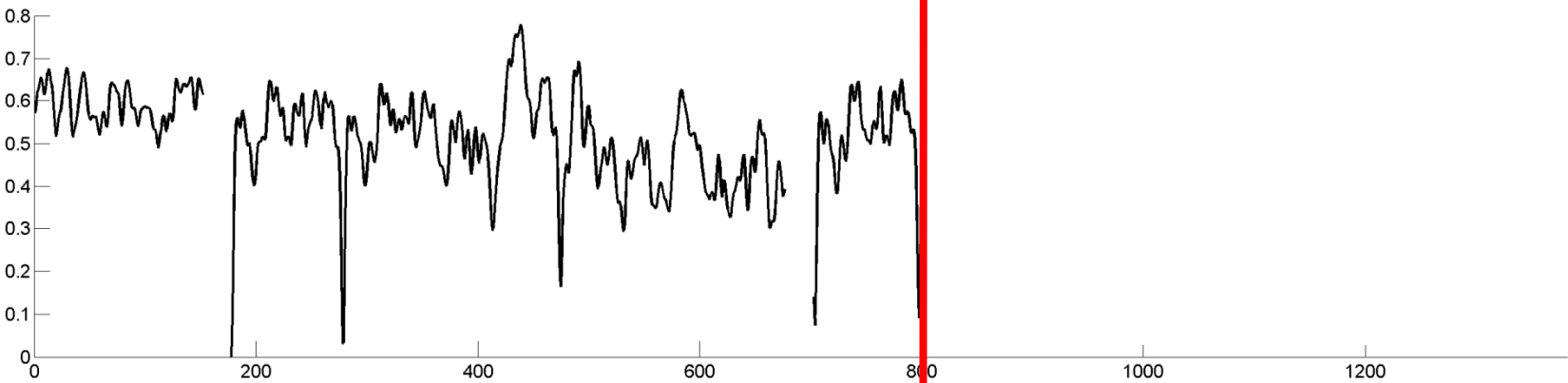


Bach

Beethoven/Bernstein

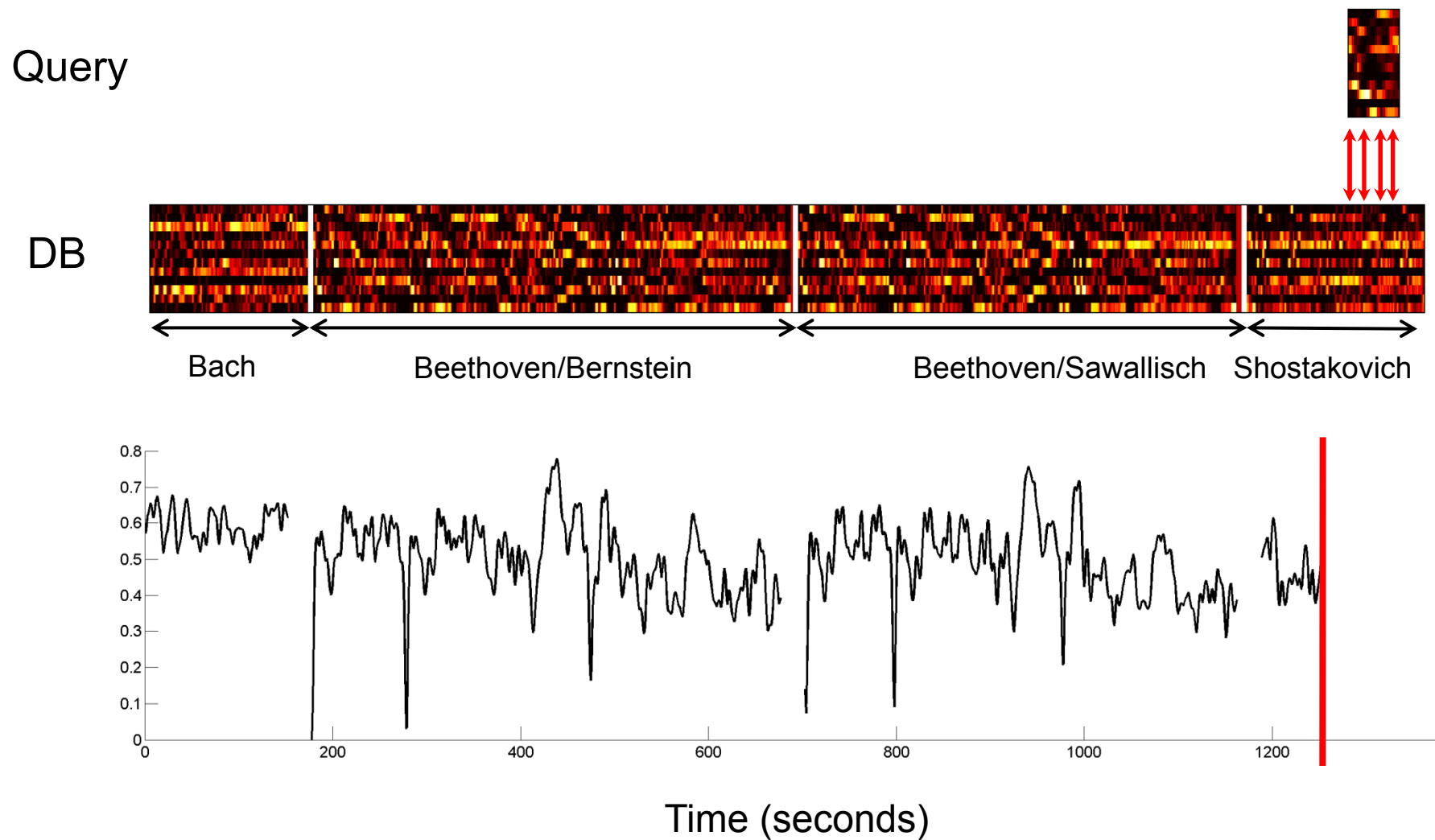
Beethoven/Sawallisch

Shostakovich



Time (seconds)

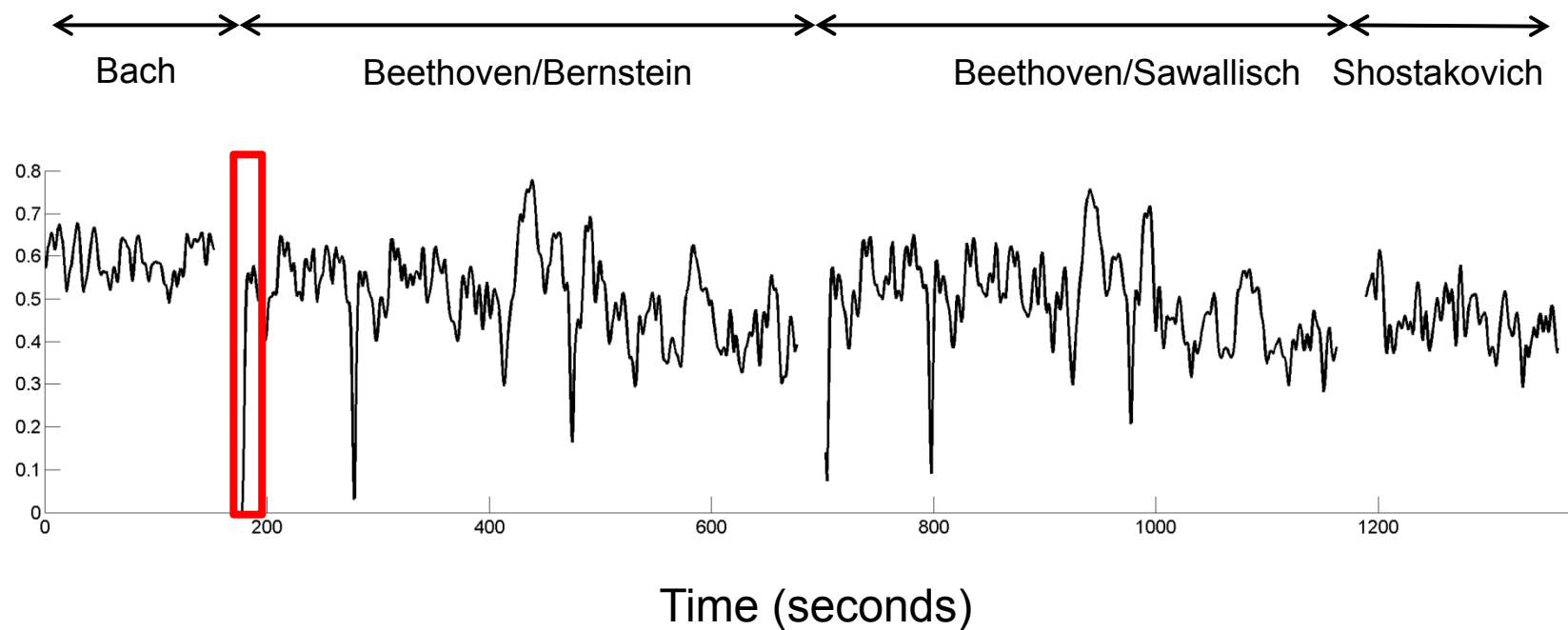
# Matching Procedure



# Matching Procedure

## Matching curve

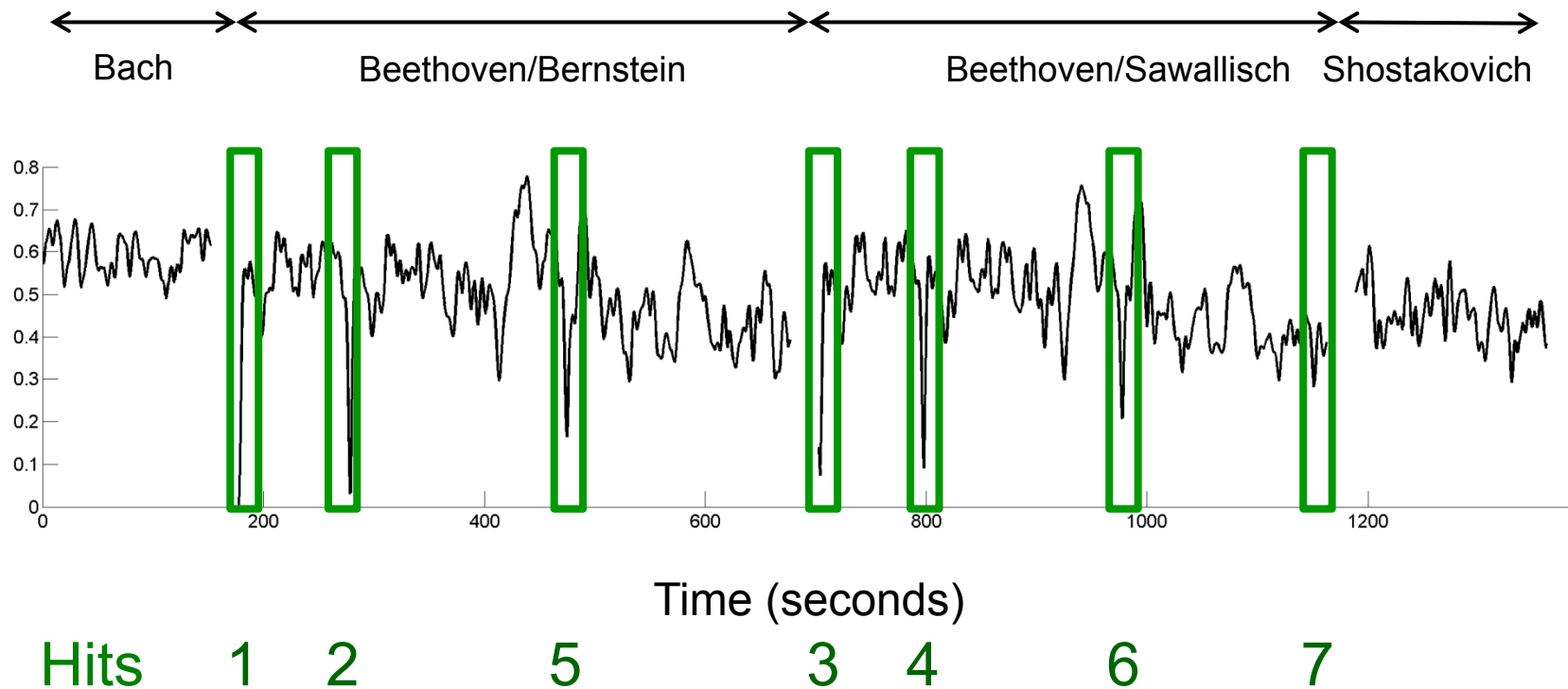
**Query:** Beethoven's Fifth / Bernstein (first 20 seconds)



# Matching Procedure

## Matching curve

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

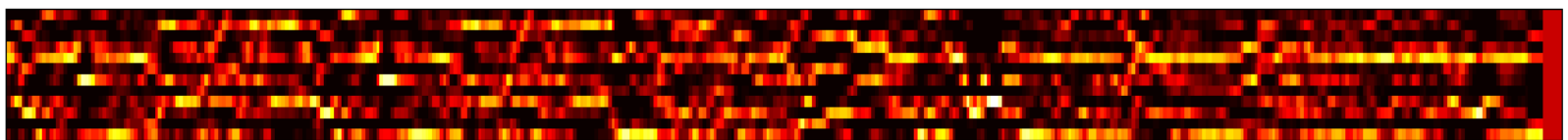
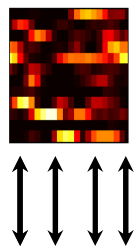




# Matching Procedure

Problem: How to deal with tempo differences?

Karajan is much faster than Bernstein!



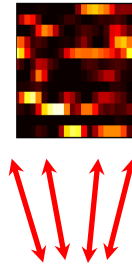
Beethoven/Karajan



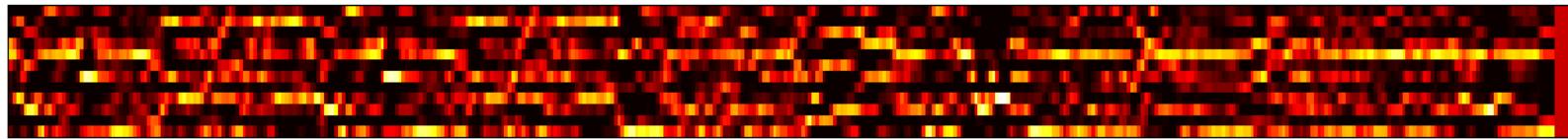
# Matching Procedure

## 1. Strategy: Usage of local warping

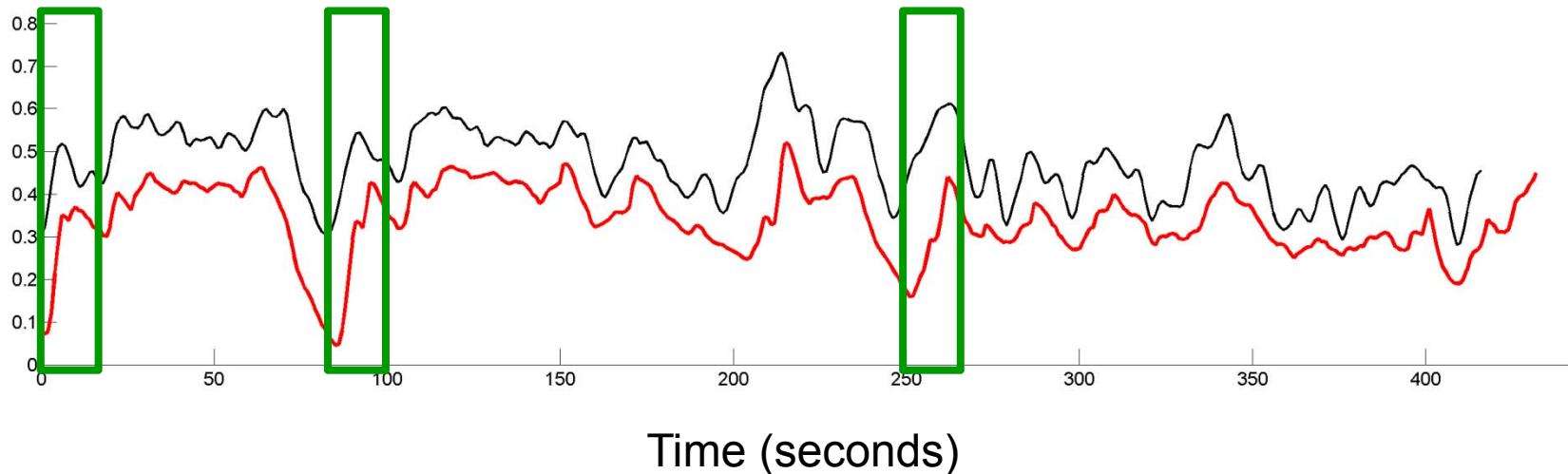
Karajan is much faster than Bernstein!



Warping strategies are computationally expensive and hard for indexing.

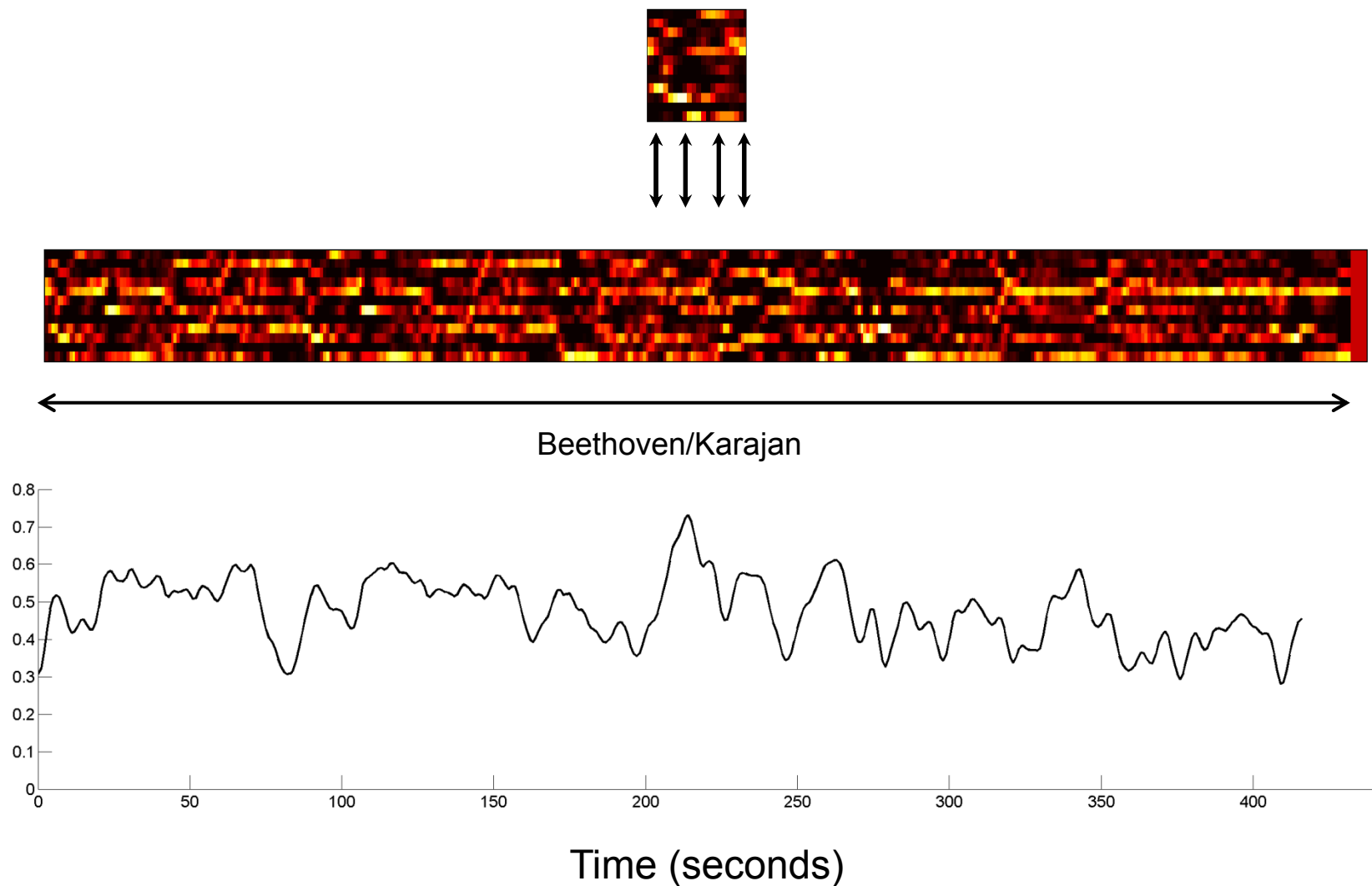


Beethoven/Karajan



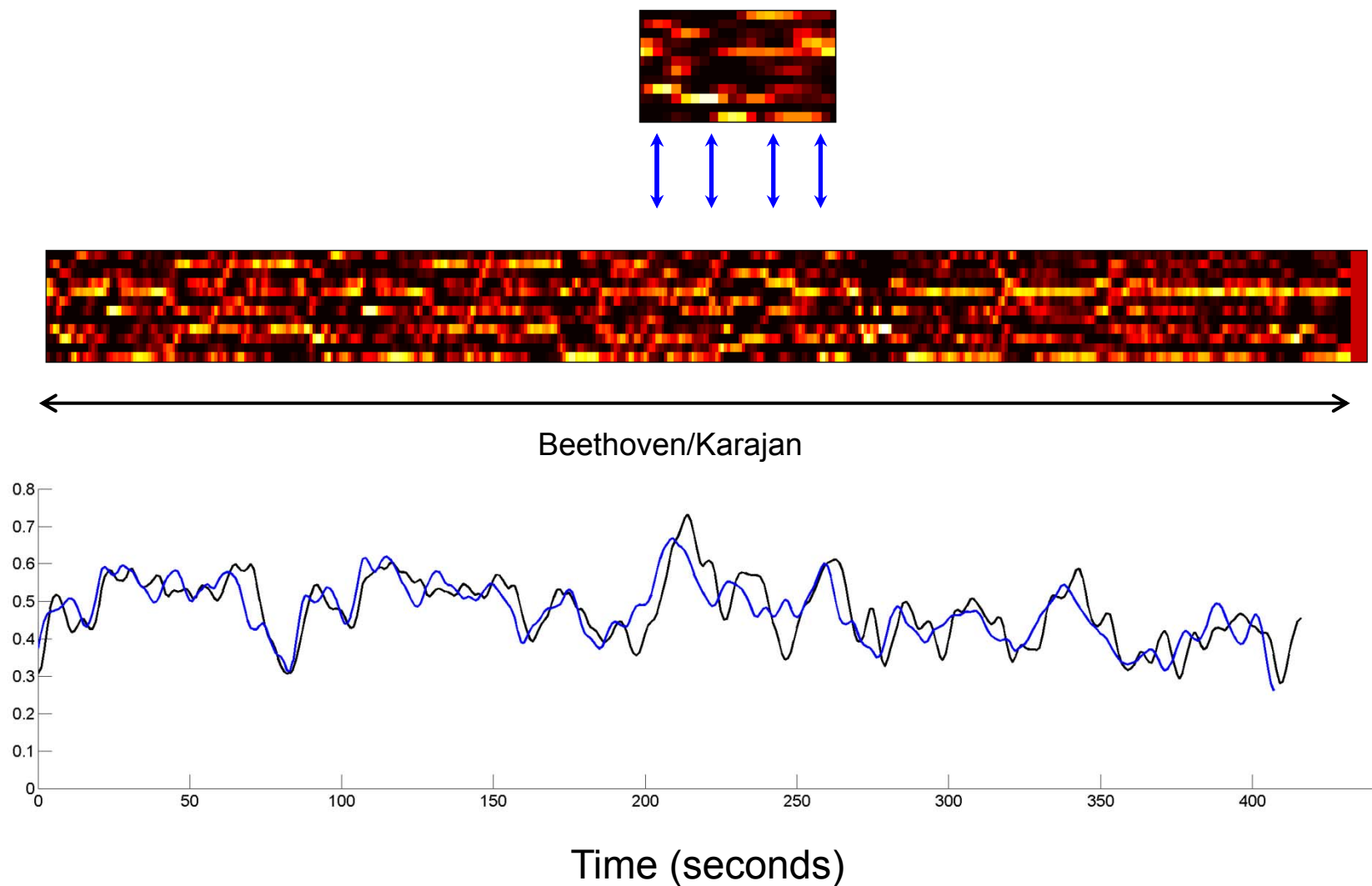
# Matching Procedure

## 2. Strategy: Usage of multiple scaling



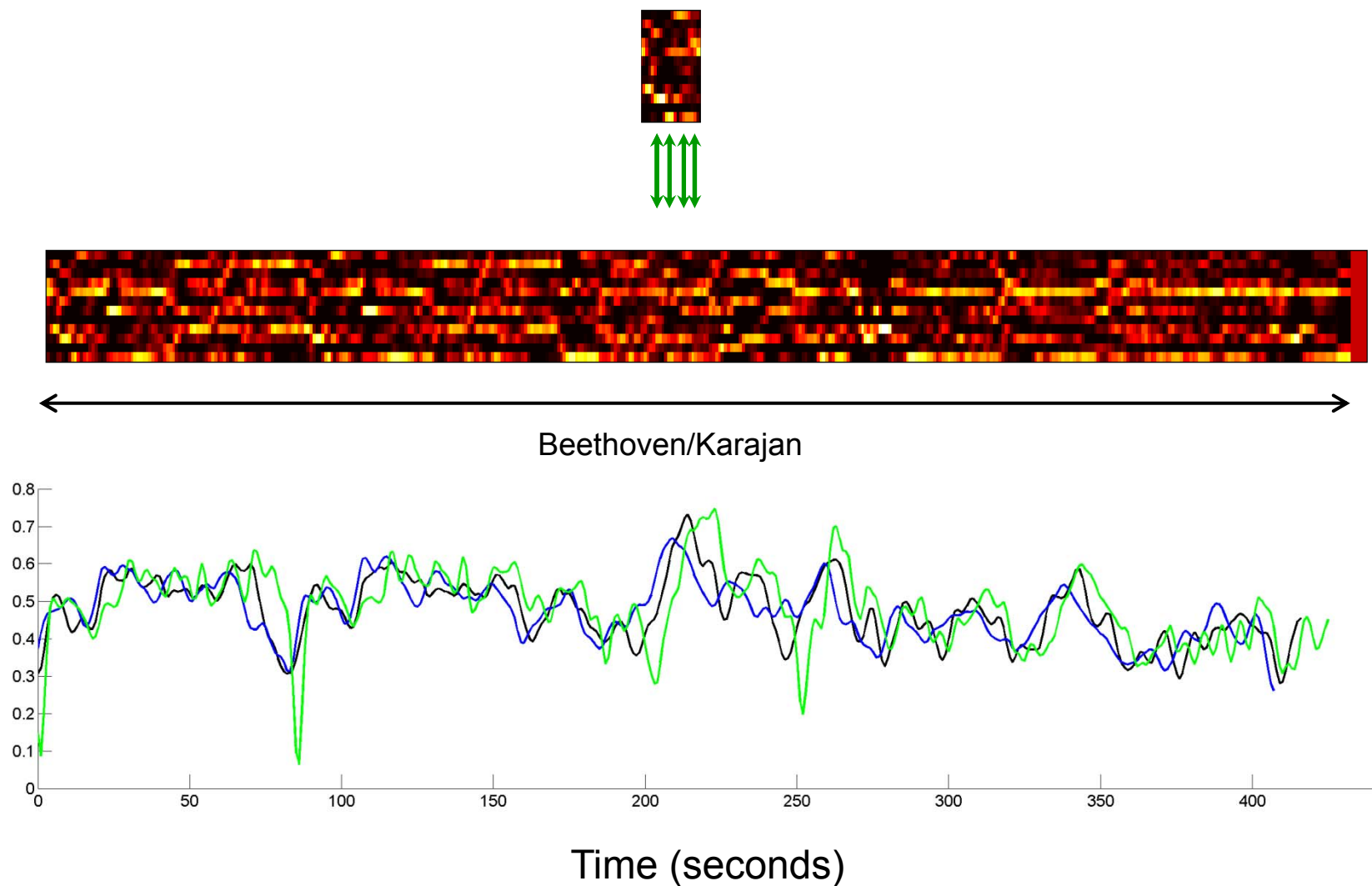
# Matching Procedure

## 2. Strategy: Usage of multiple scaling



# Matching Procedure

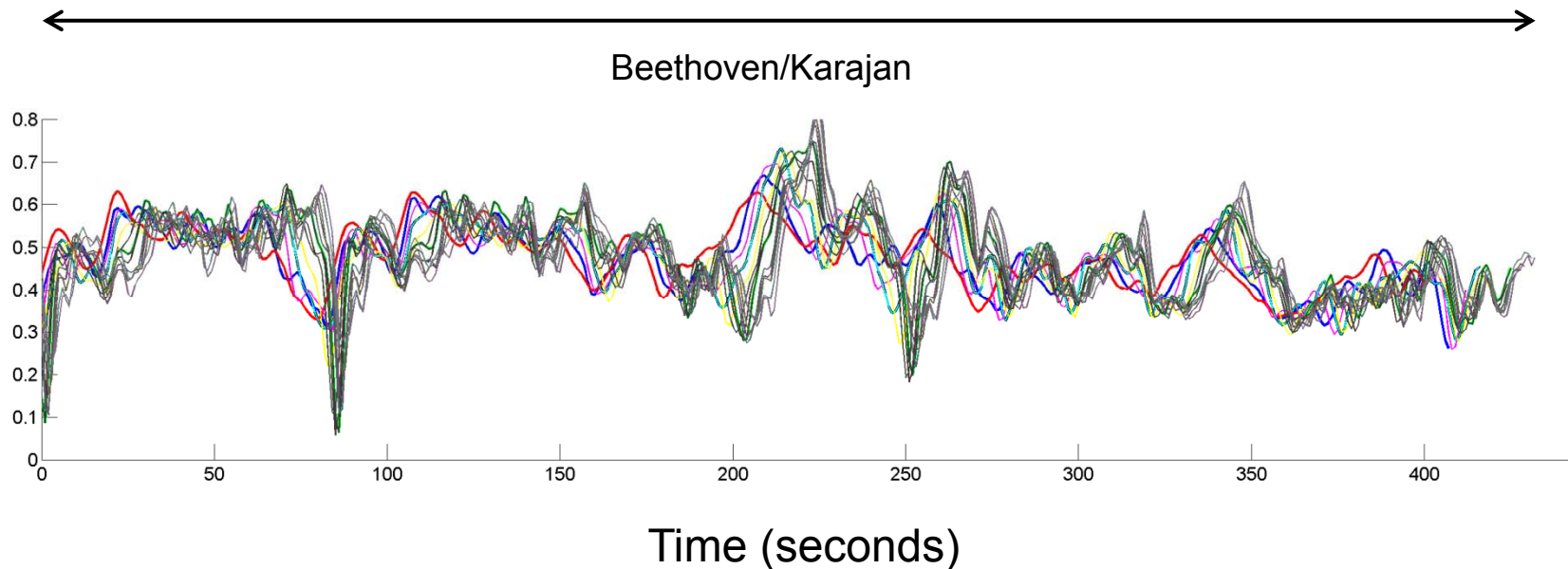
## 2. Strategy: Usage of multiple scaling



# Matching Procedure

## 2. Strategy: Usage of multiple scaling

Query resampling simulates tempo changes

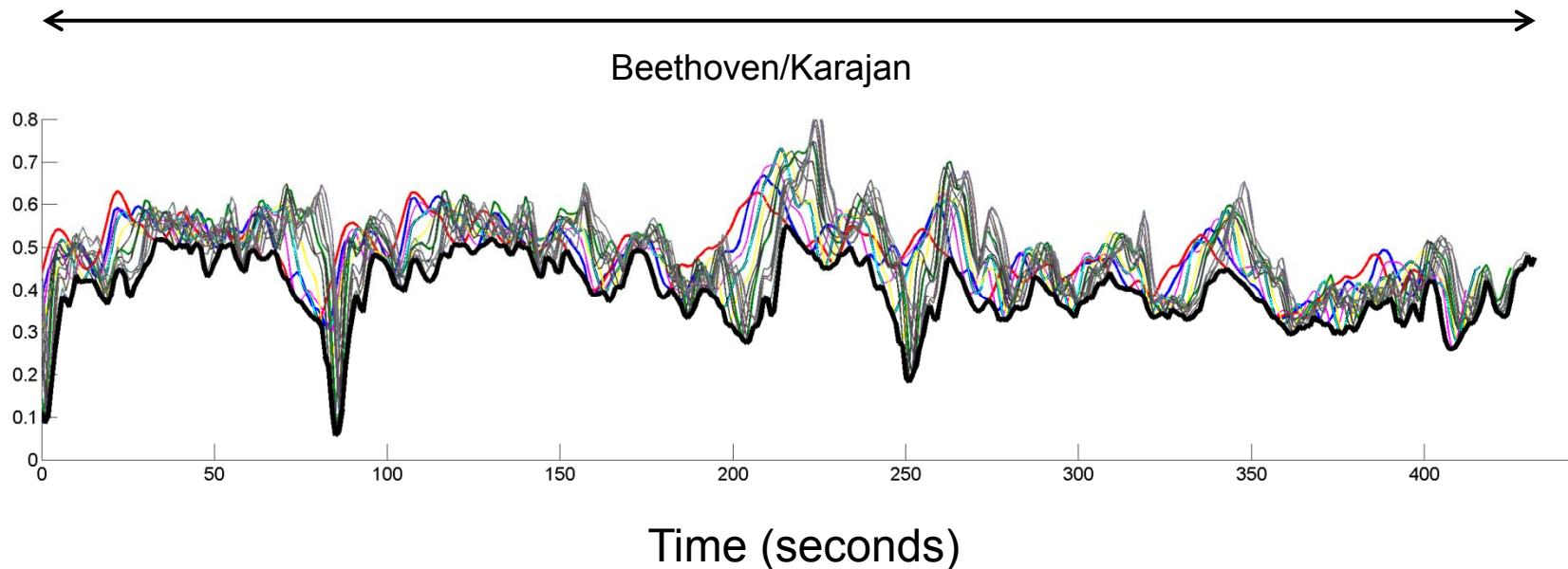


# Matching Procedure

## 2. Strategy: Usage of multiple scaling

Query resampling simulates tempo changes

**Minimize over all curves**



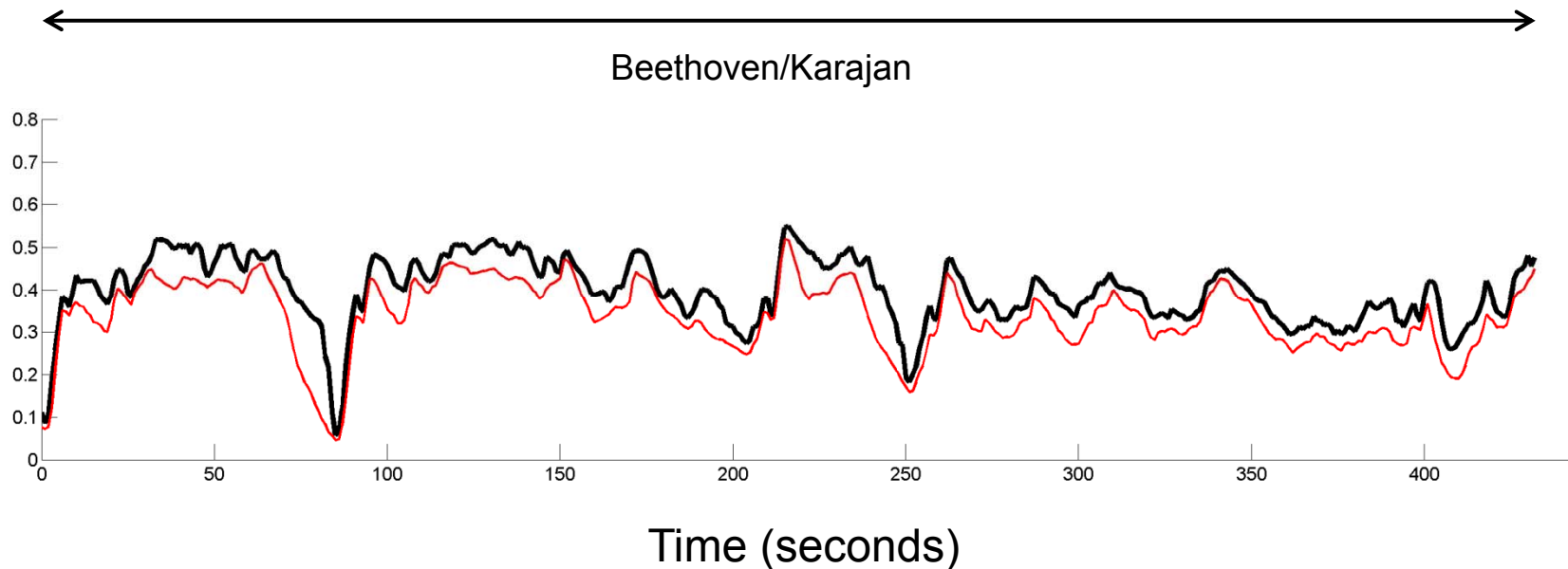
# Matching Procedure

## 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

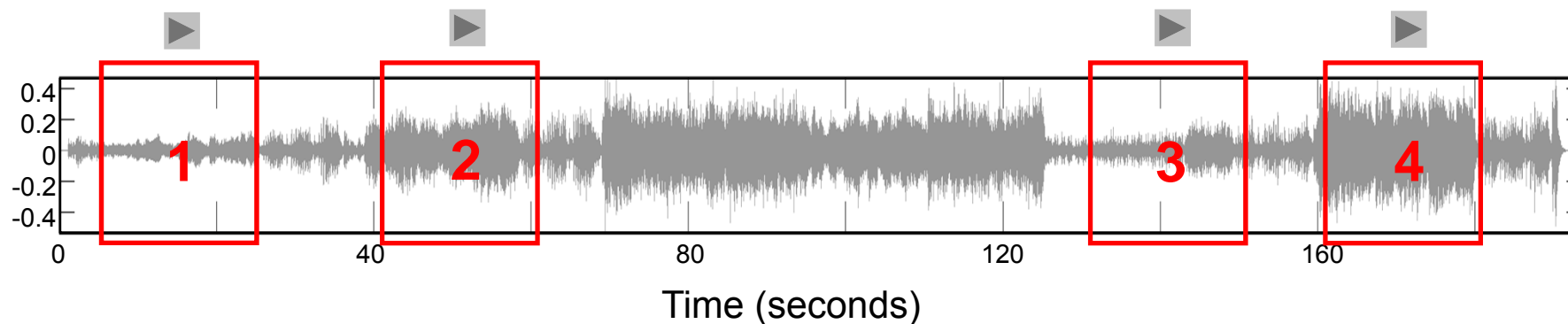
## Idea:

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

- Relative comparison
- Compact
- Intuitive
- Quantitative evaluation

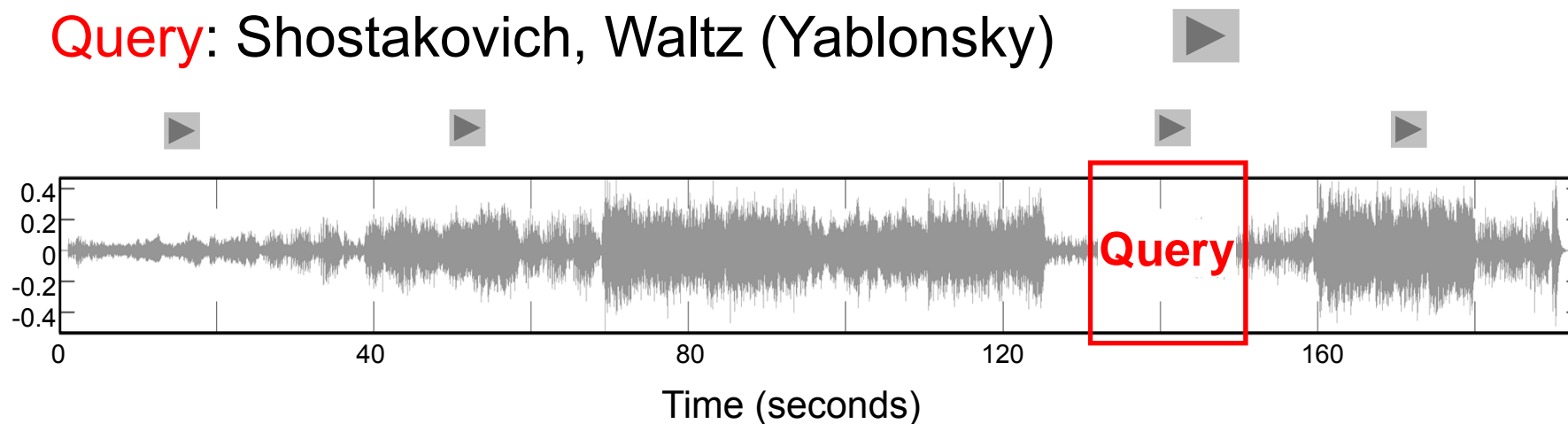
# Audio Matching for Feature Analysis

Example: Shostakovich, Waltz (Yablonsky) 



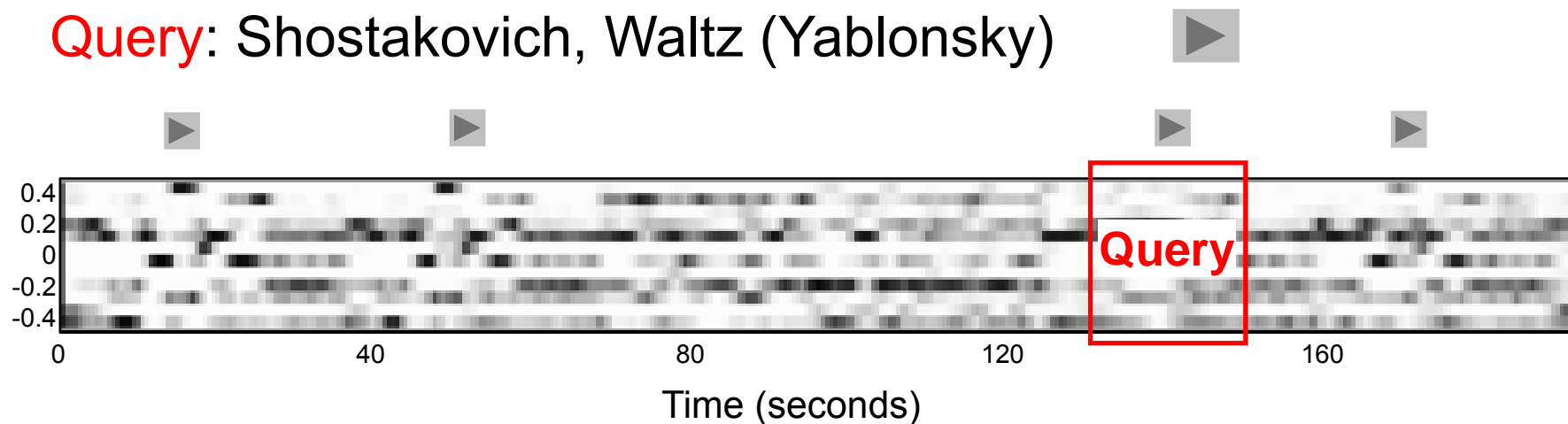
# Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



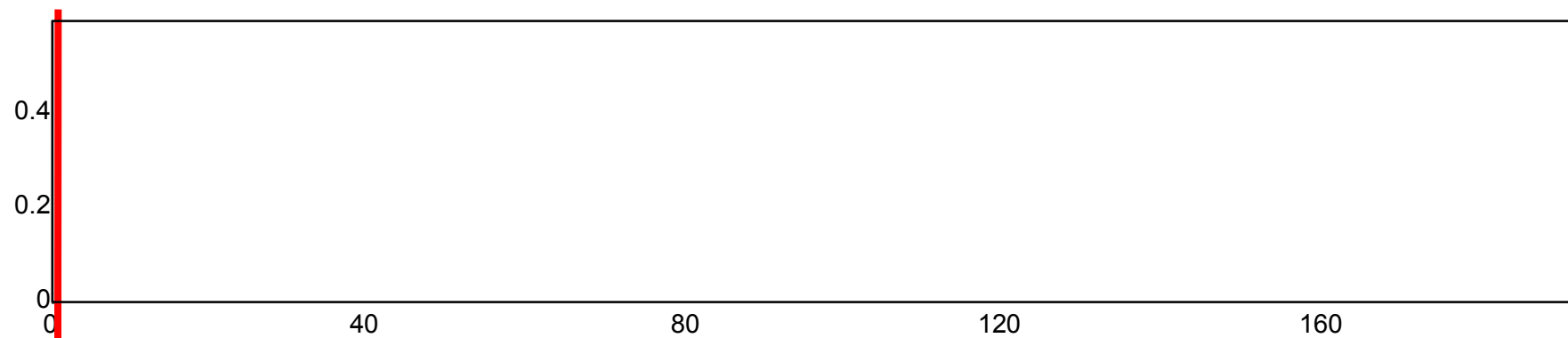
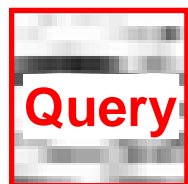
# Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



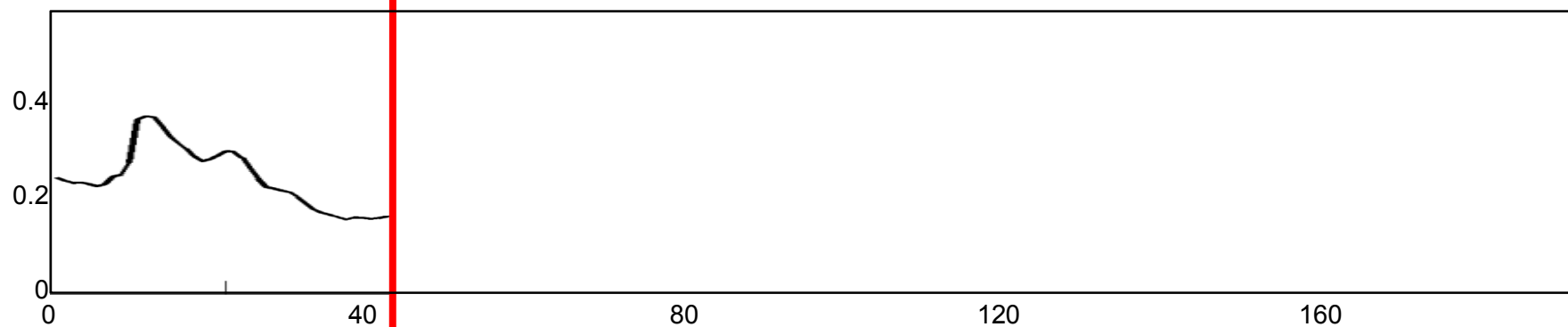
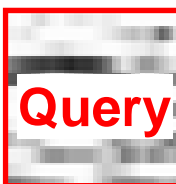
# Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



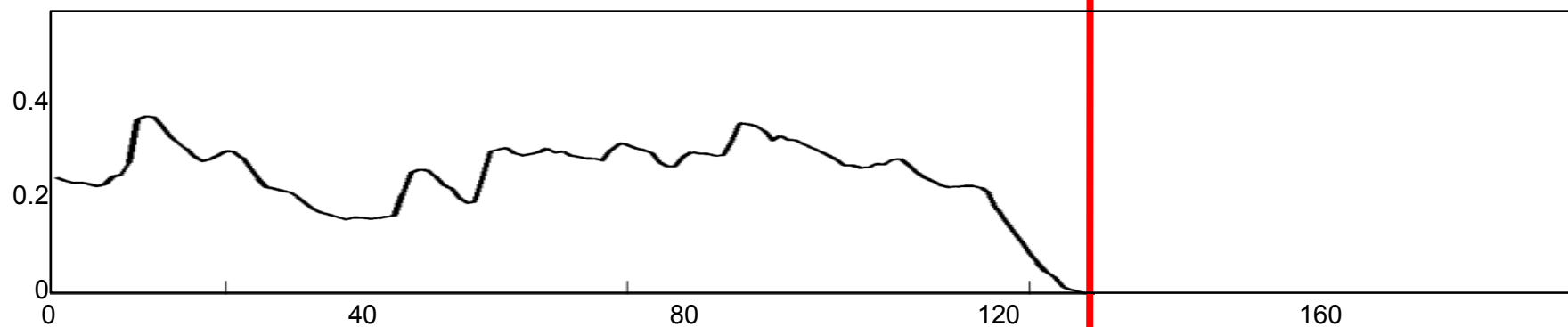
# Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



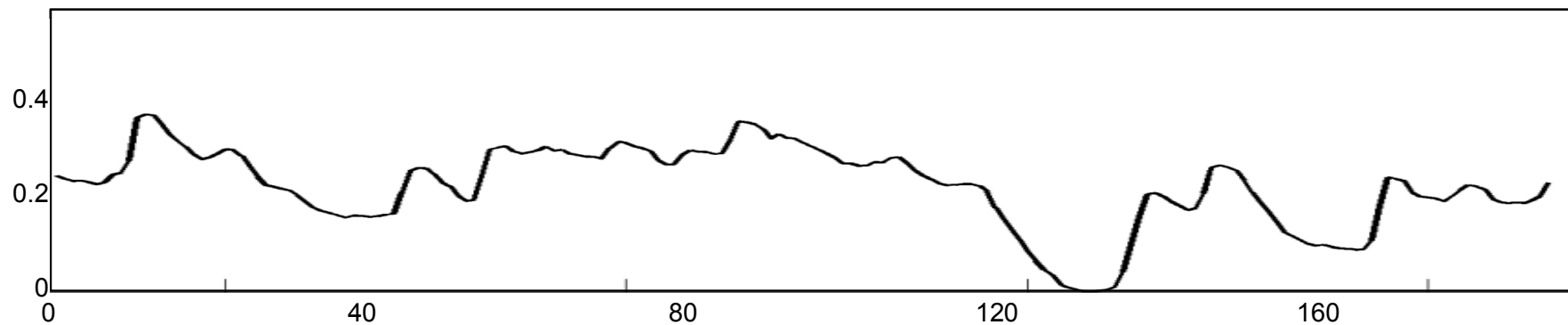
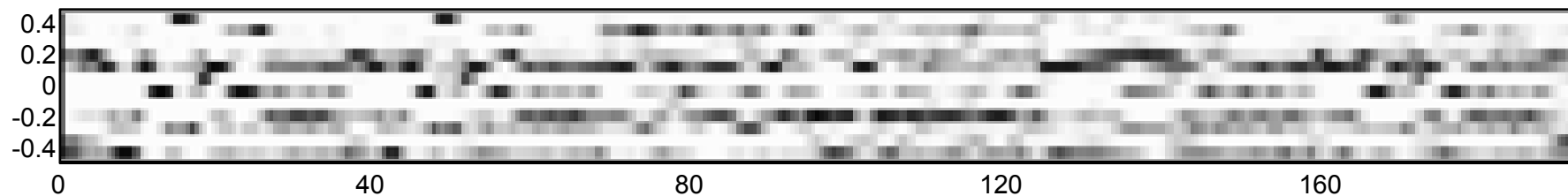
# Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



# Audio Matching for Feature Analysis

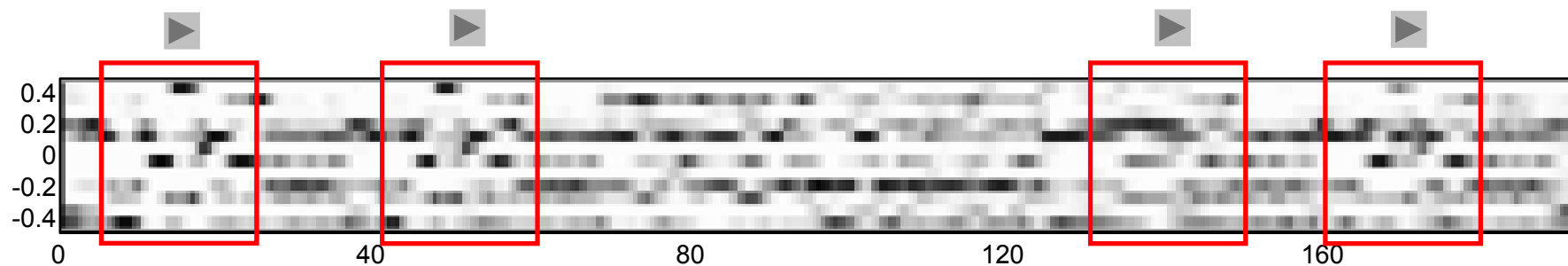
Query: Shostakovich, Waltz (Yablonsky)



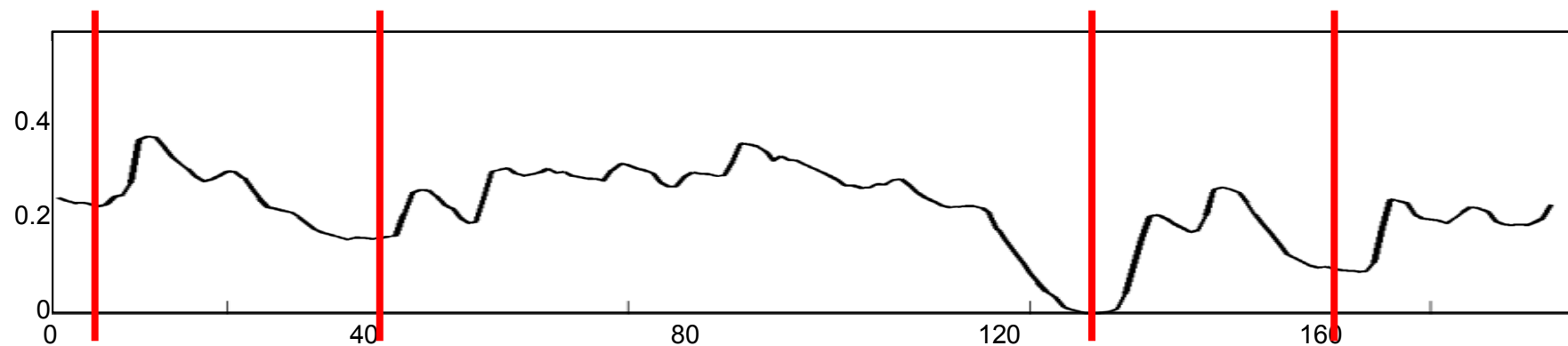


# Audio Matching for Feature Analysis

Query: Shostakovich, Waltz (Yablonsky)



Expected matching positions (should have local minima)

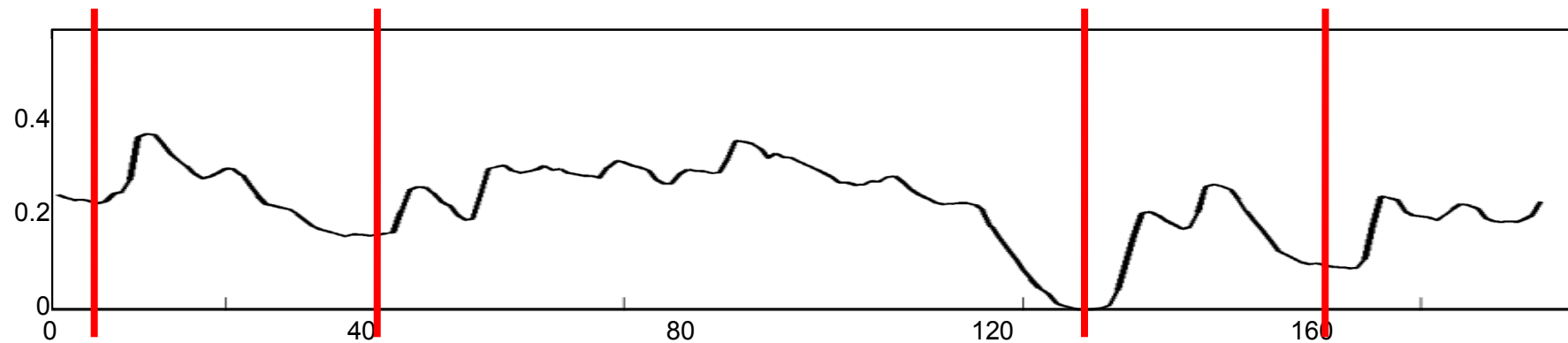


# Audio Matching for Feature Analysis

Idea:

- Use matching curve for analyzing feature properties

Expected matching positions (should have local minima)

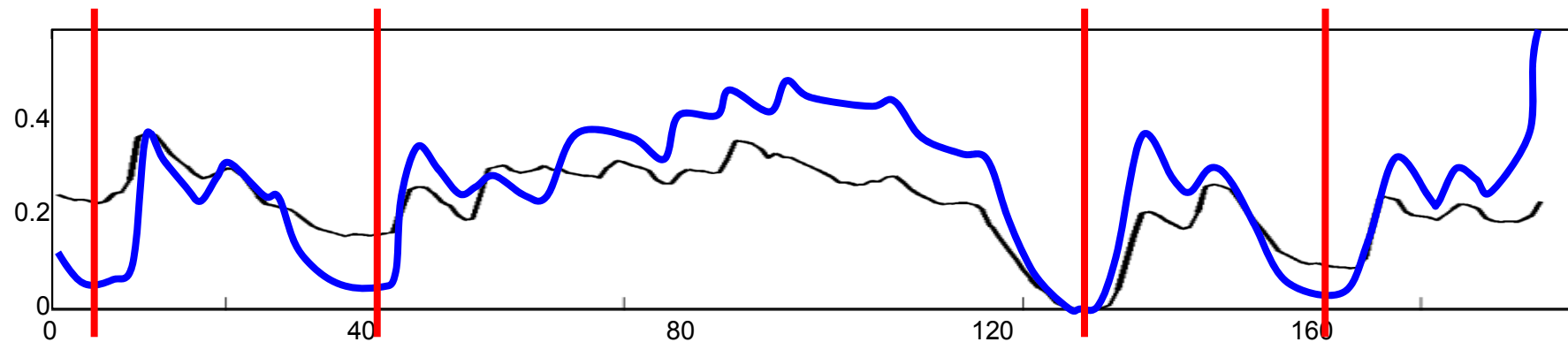


# Audio Matching for Feature Analysis

Idea:

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

Expected matching positions (should have local minima)



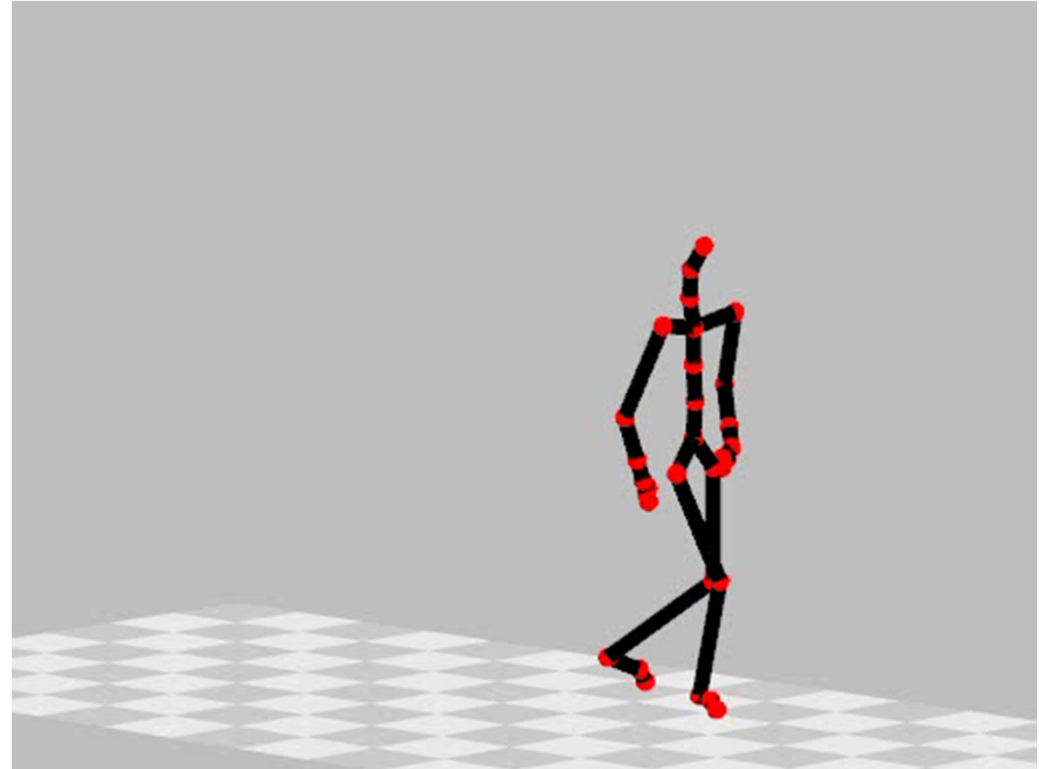
---

# Overview

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

# Motion Capture Data

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



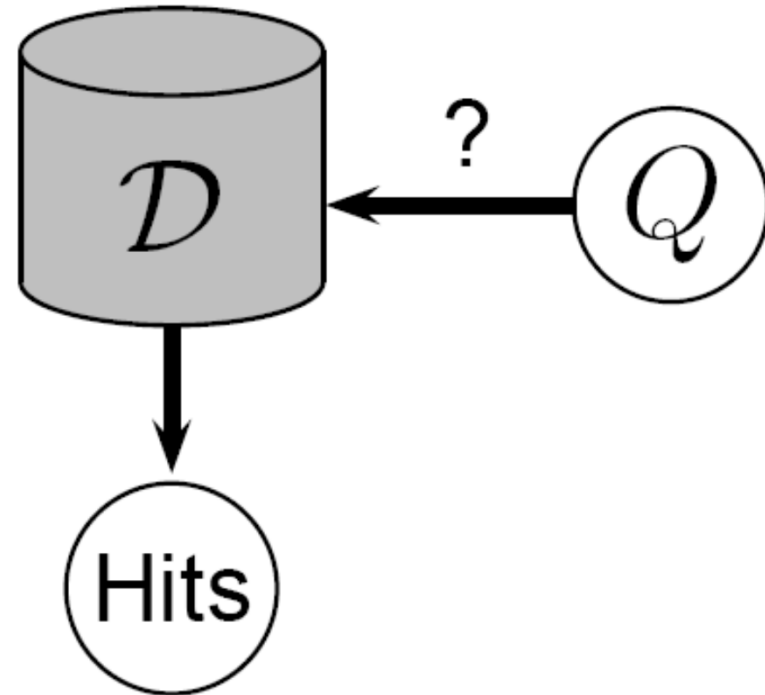
# Motion Capture Data

## Optical System

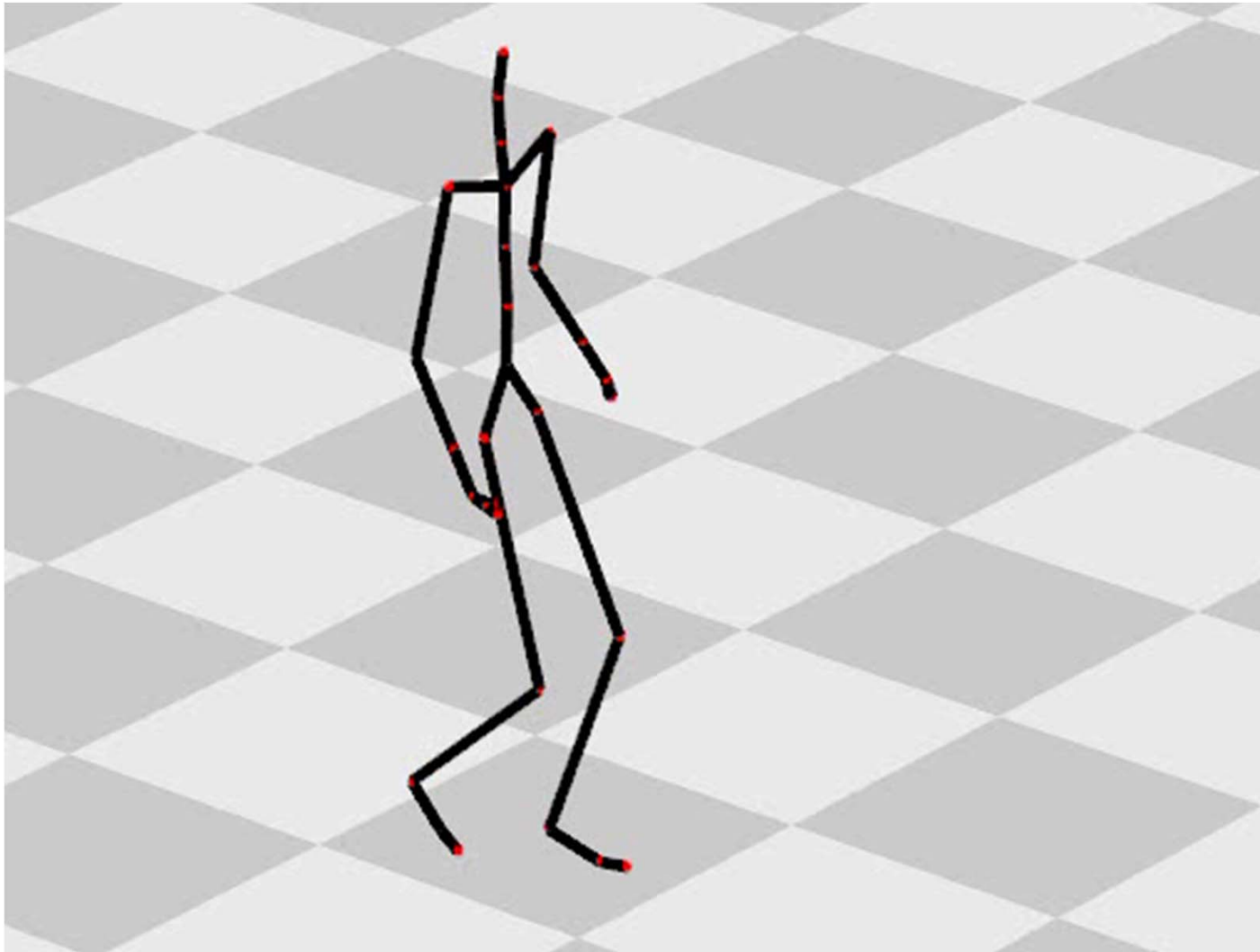


# Motion Retrieval

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



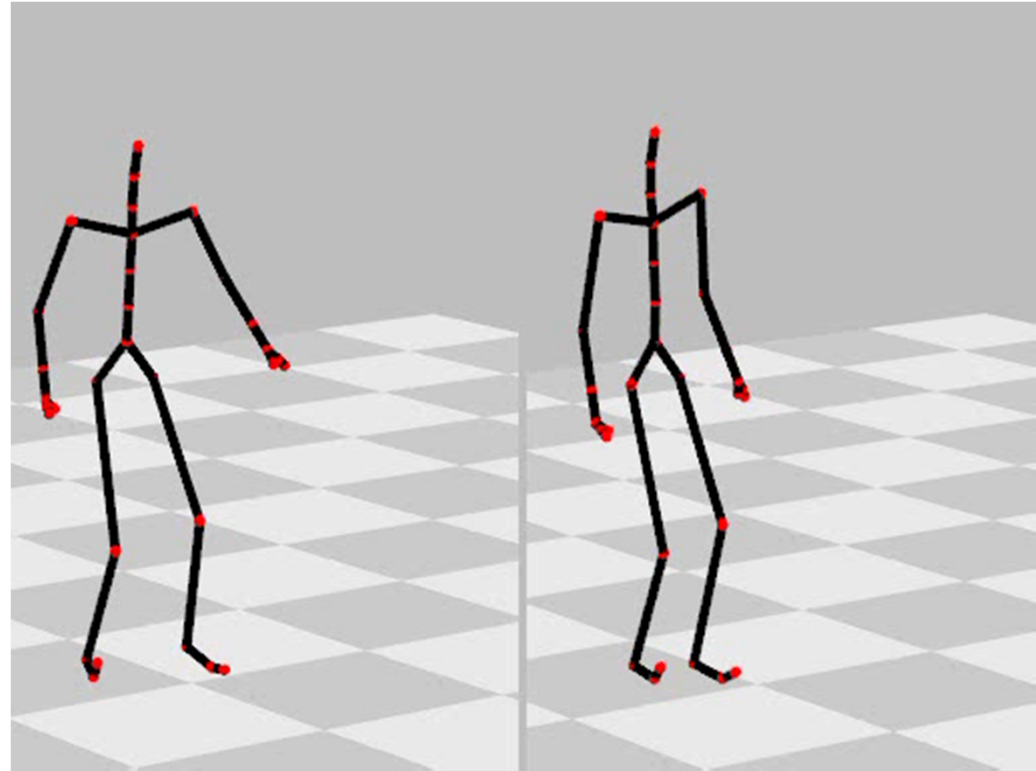
# Motion Retrieval





# 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.**

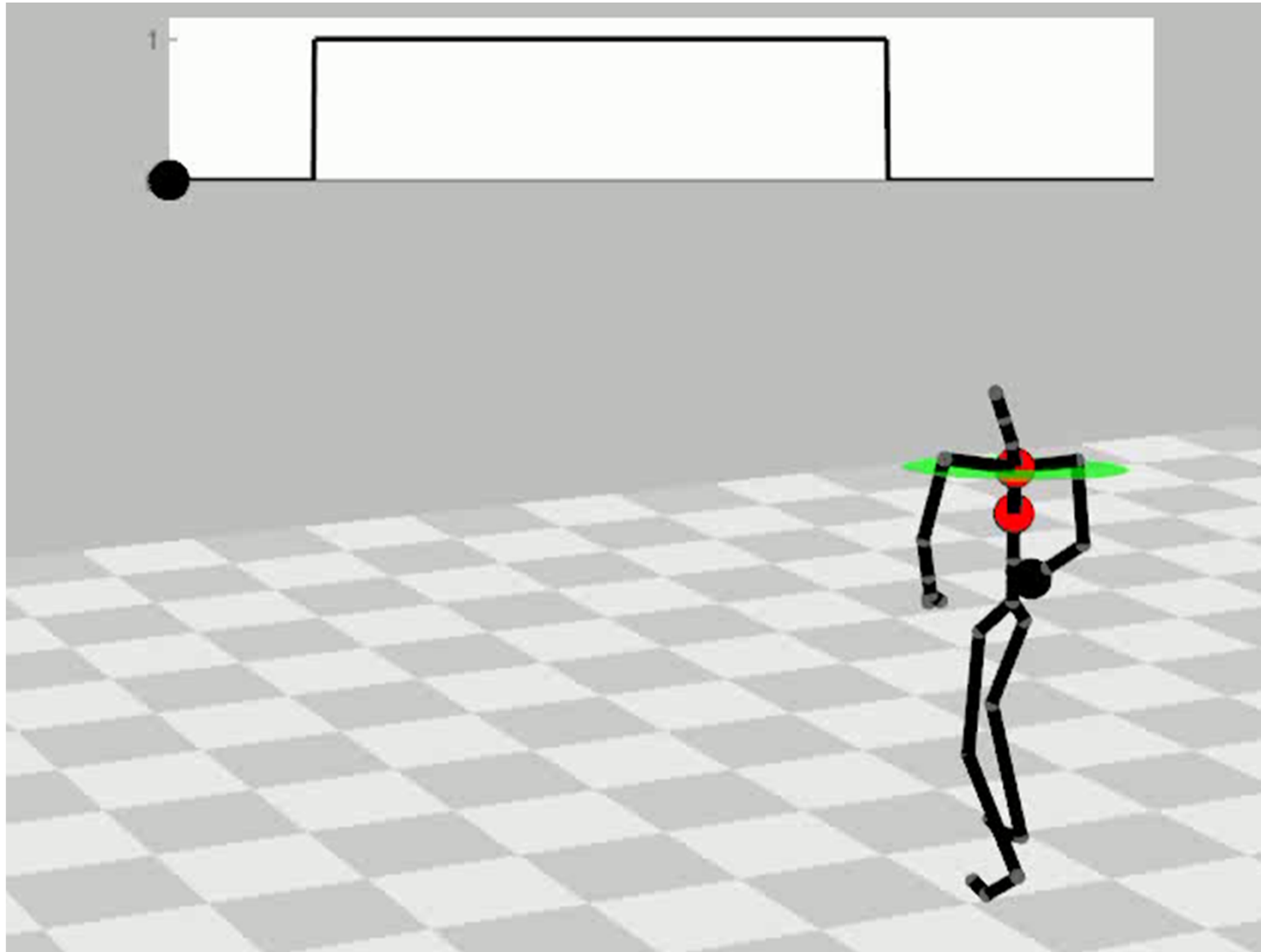
ACM Transactions on Graphics (SIGGRAPH), vol. 24, pp. 677-685, 2005.

Meinard Müller and Tido Röder

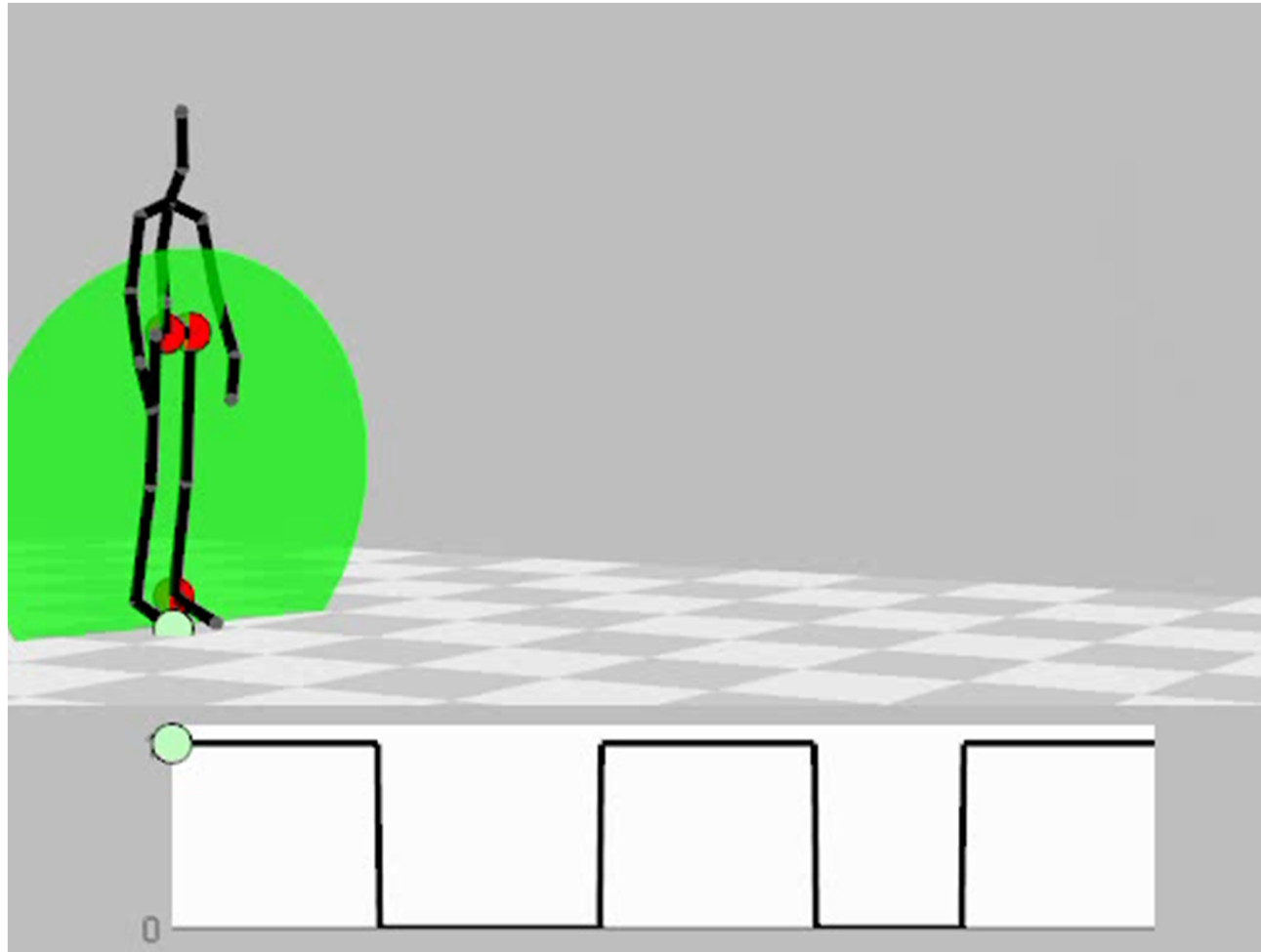
**Motion templates for automatic classification and retrieval of motion capture data.**

Proceedings of the 2006 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA), Vienna, Austria, pp. 137-146, 2006.

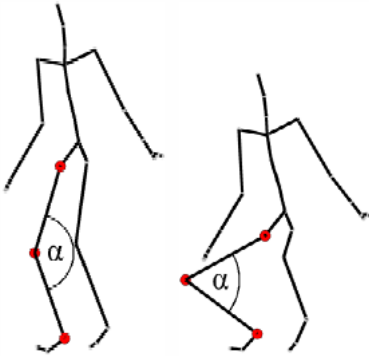
# Relational Features



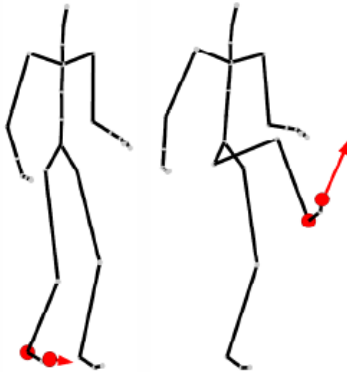
# Relational Features



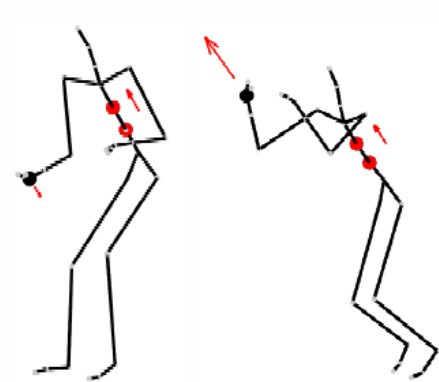
# Relational Features



Right knee bent?

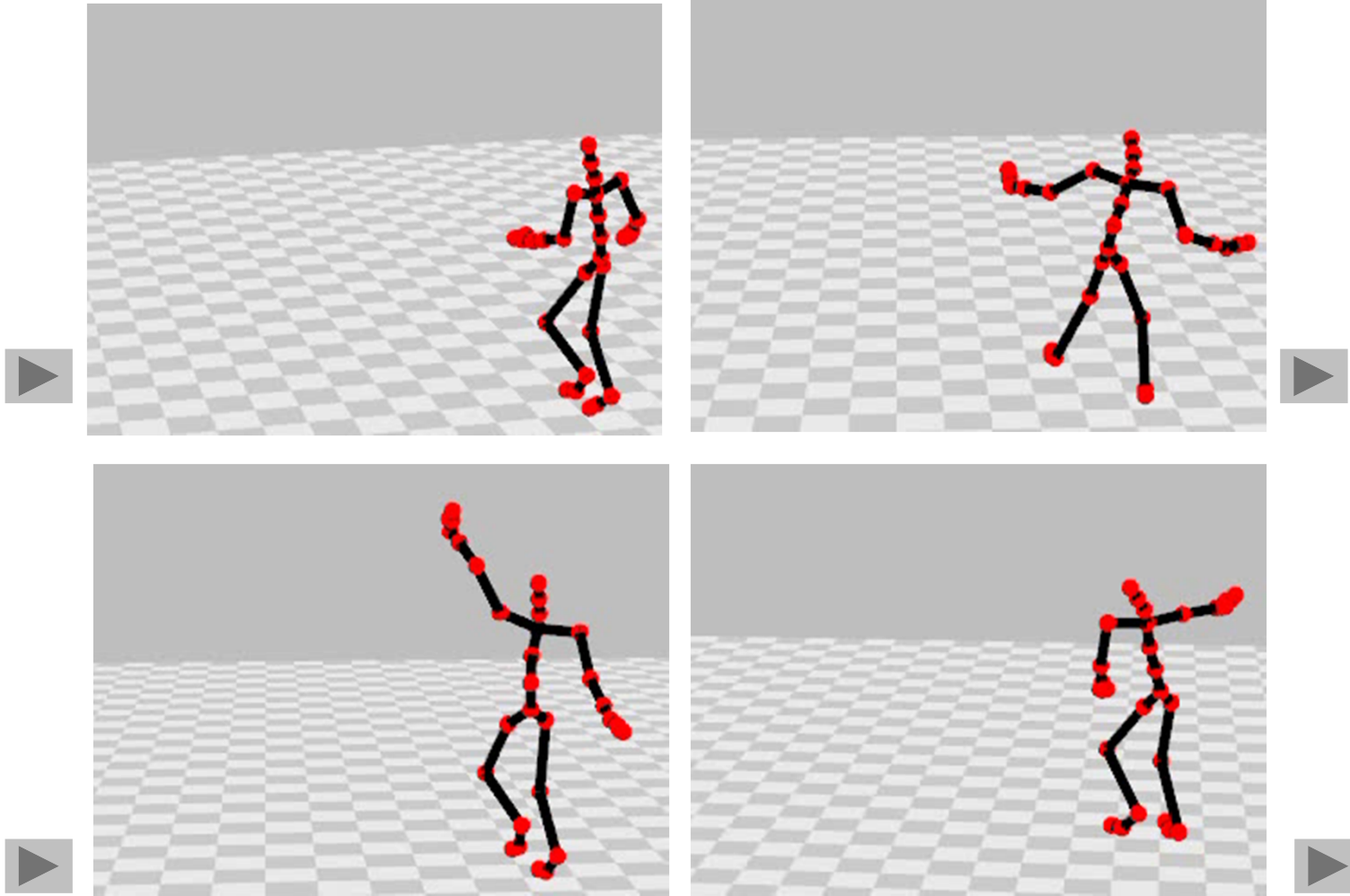


Right foot fast?

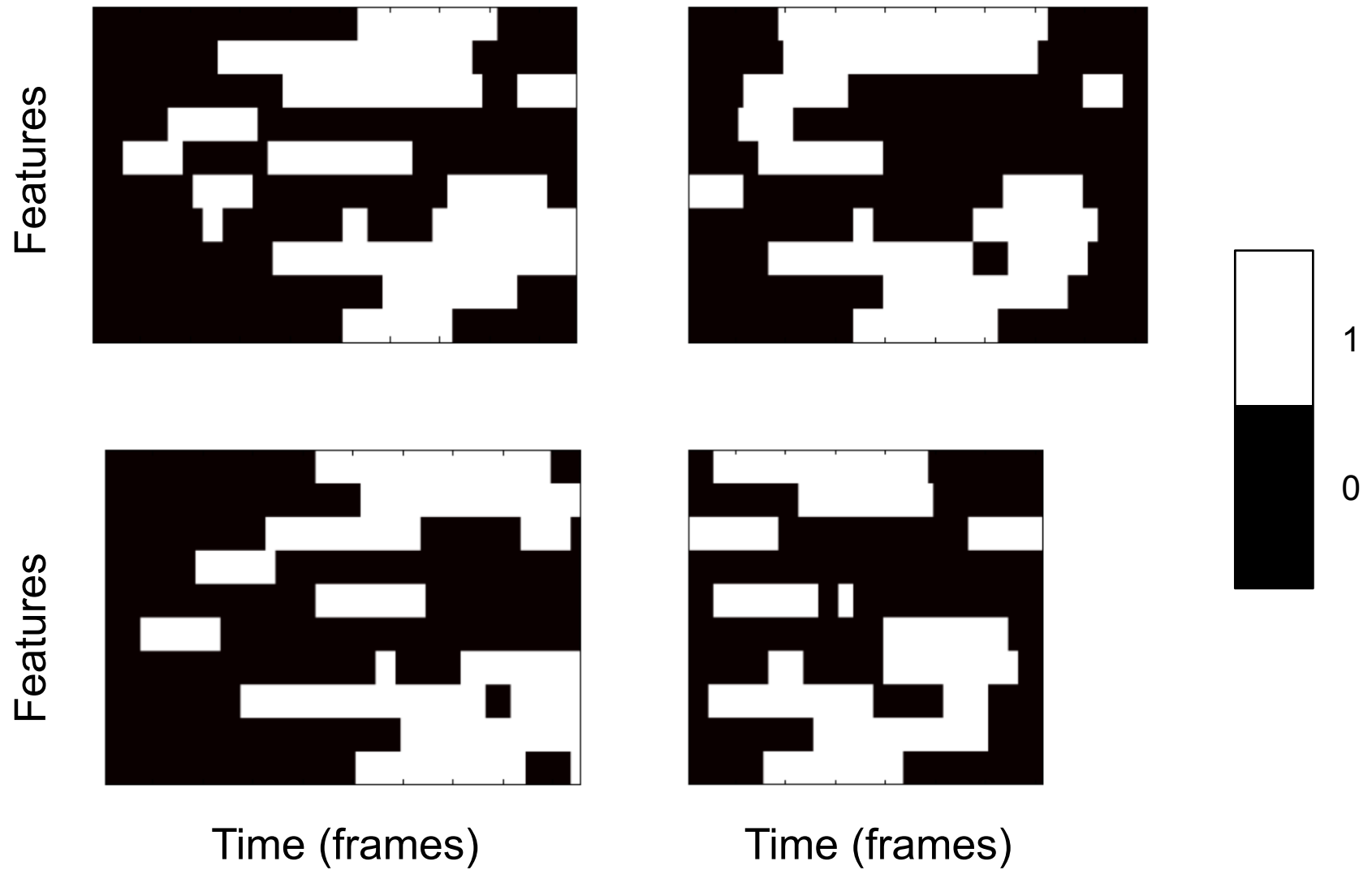


Right hand moving upwards?

# Motion Templates (MT)

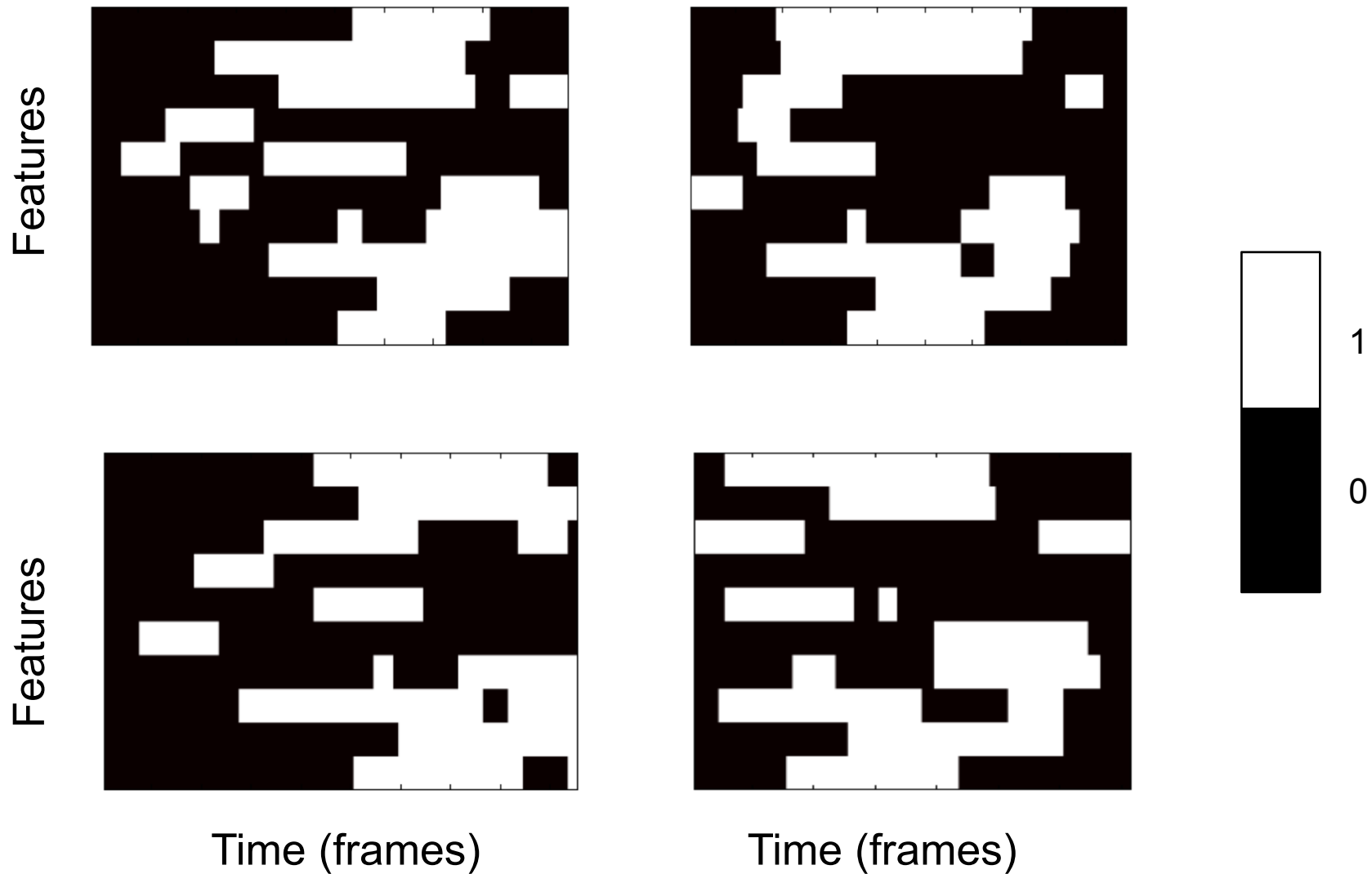


# Motion Templates (MT)



# Motion Templates (MT)

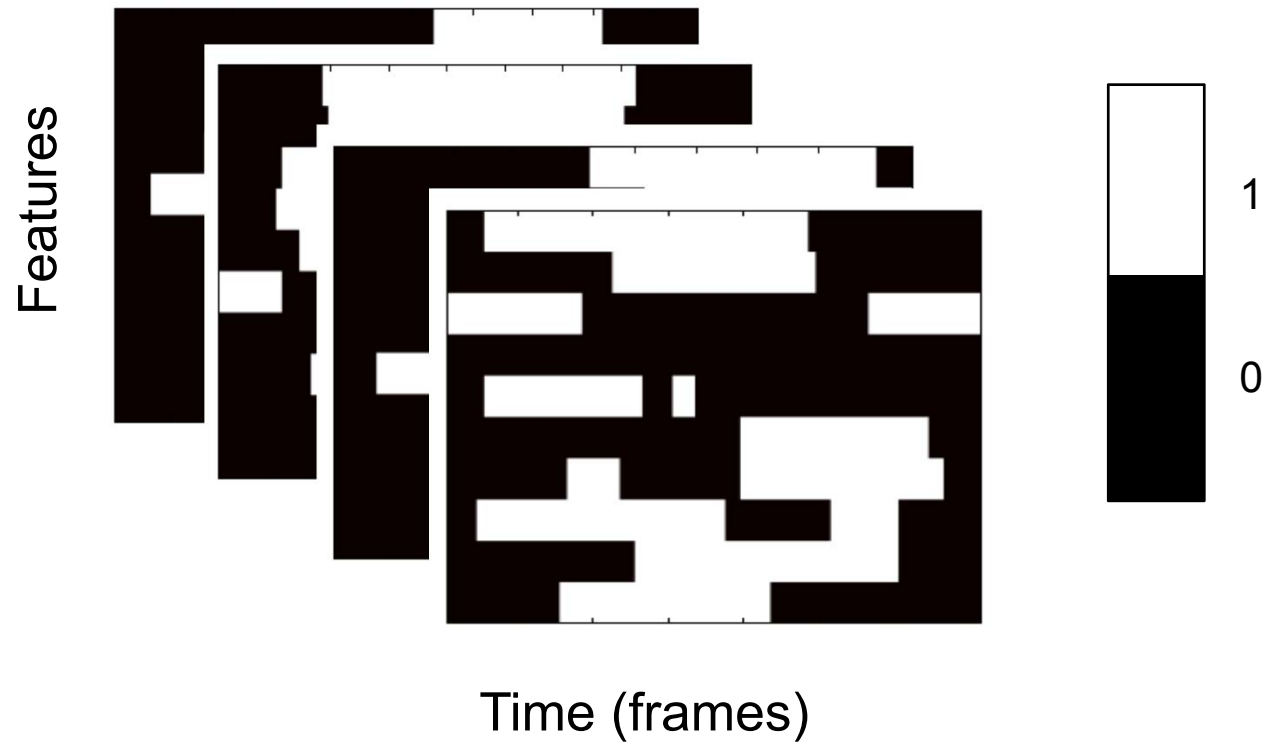
## Temporal alignment





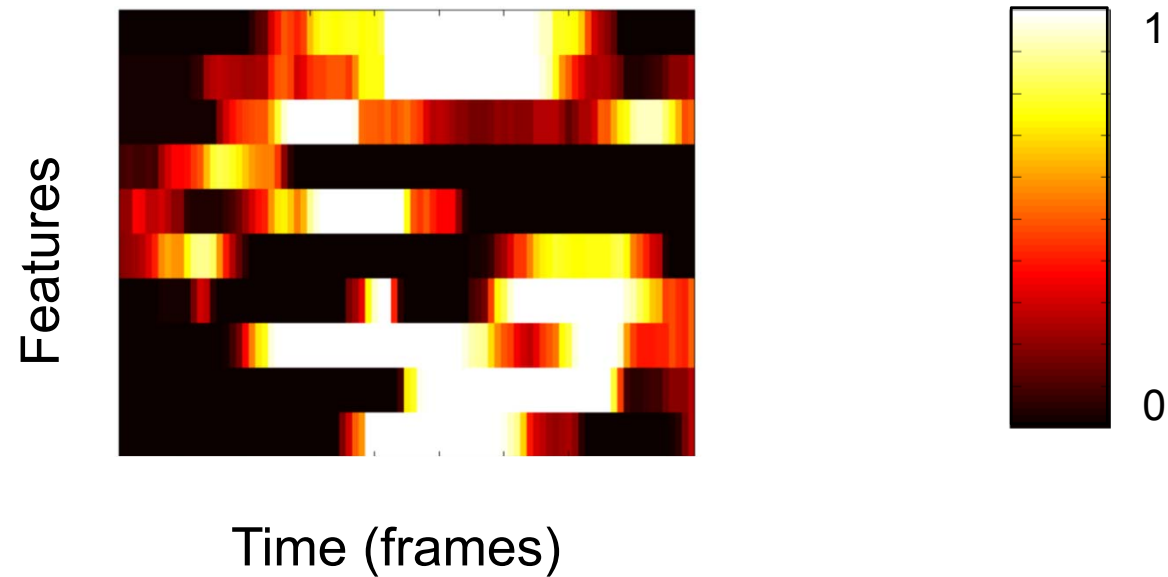
# Motion Templates (MT)

Superimpose templates

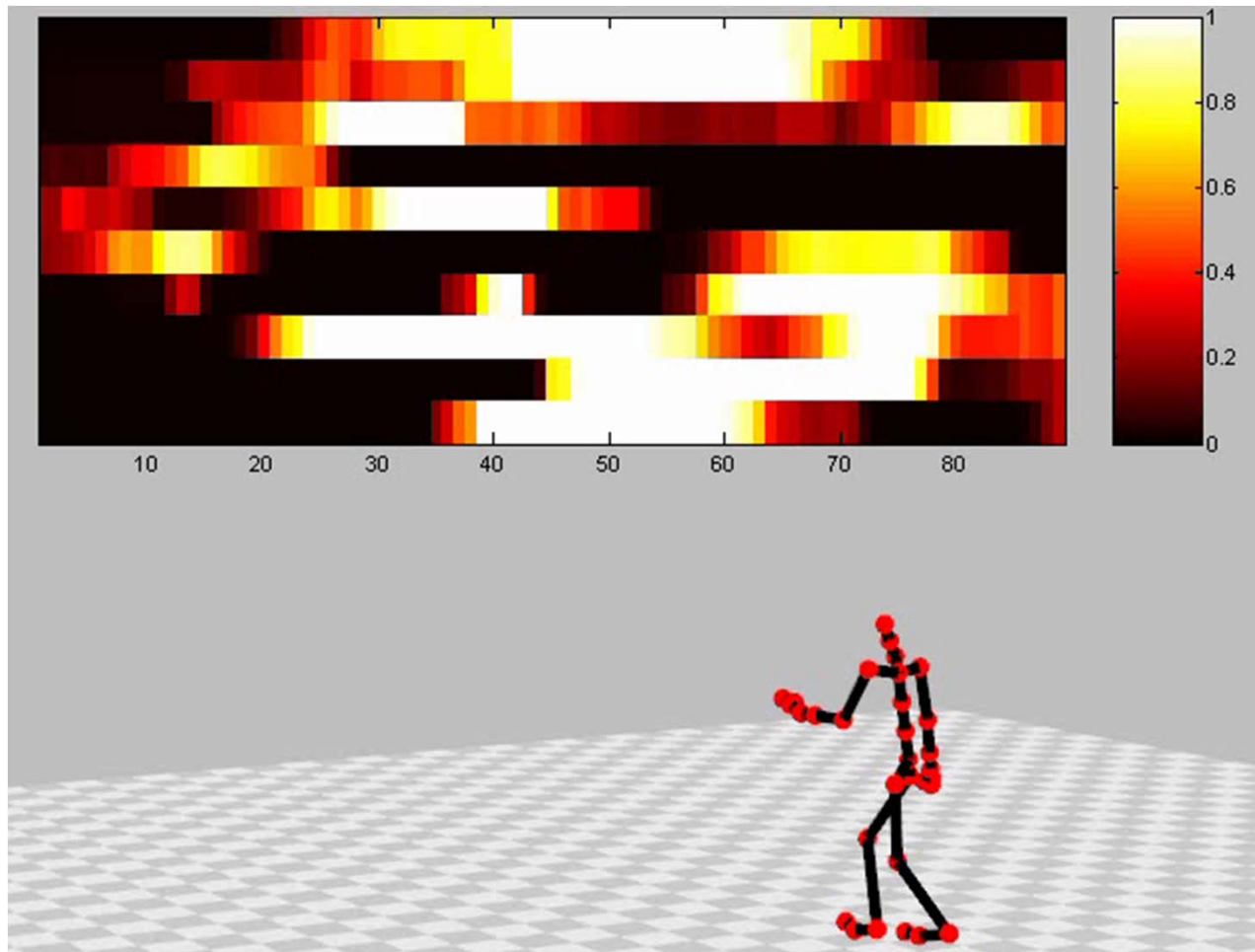


# Motion Templates (MT)

Compute average

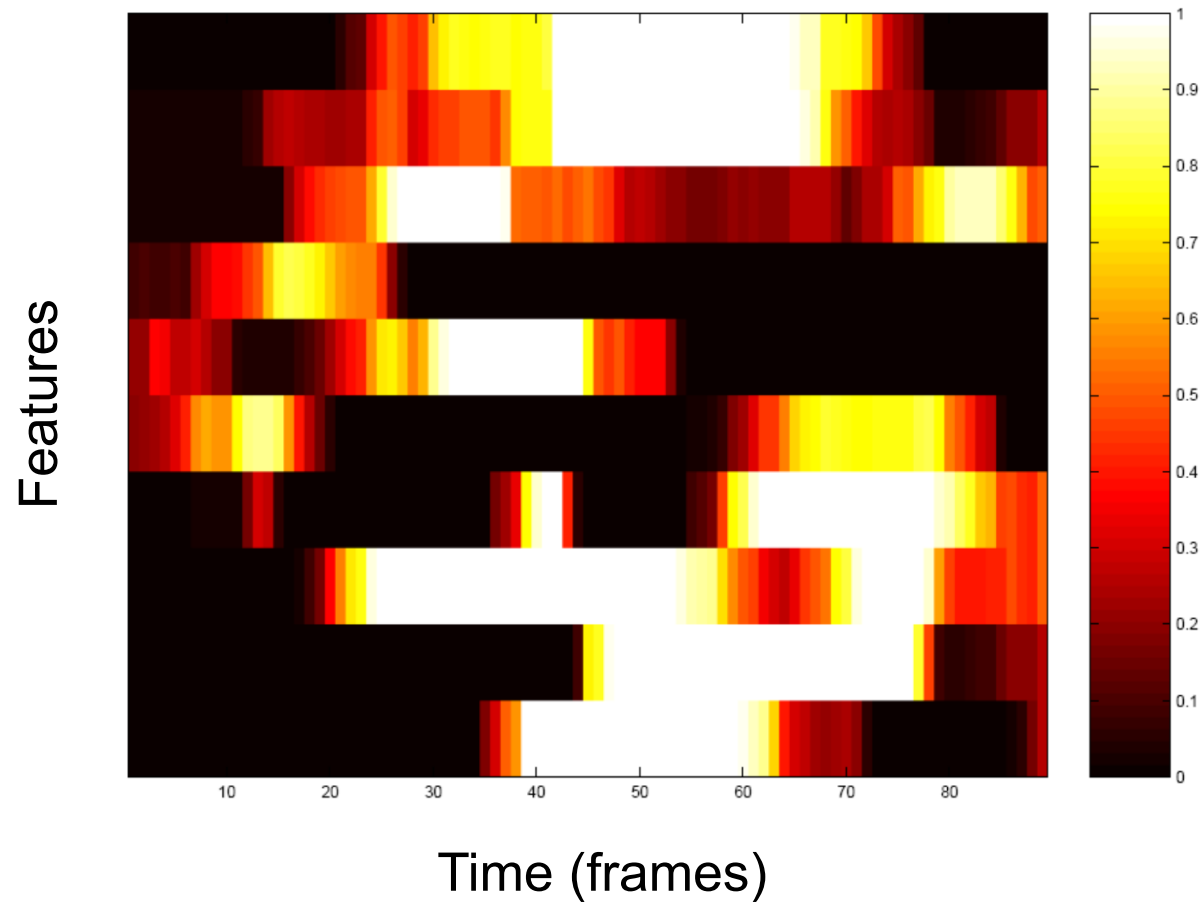


# Motion Templates (MT)



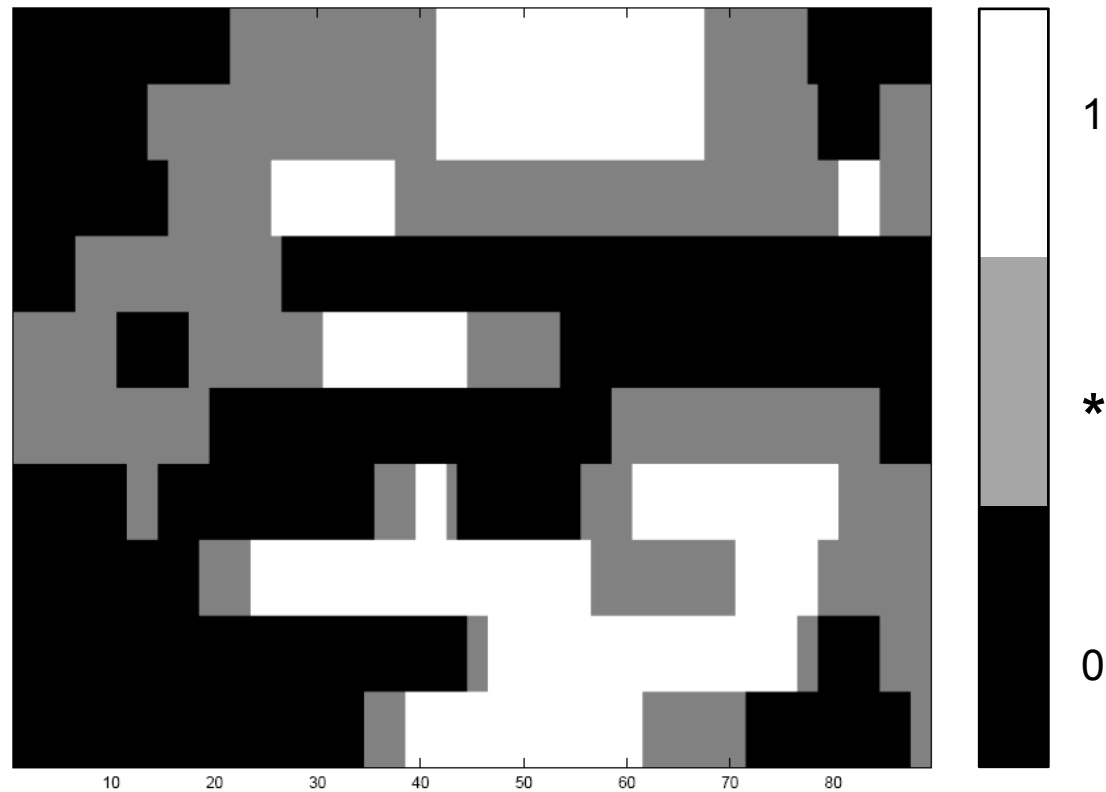
# Motion Templates (MT)

Average template



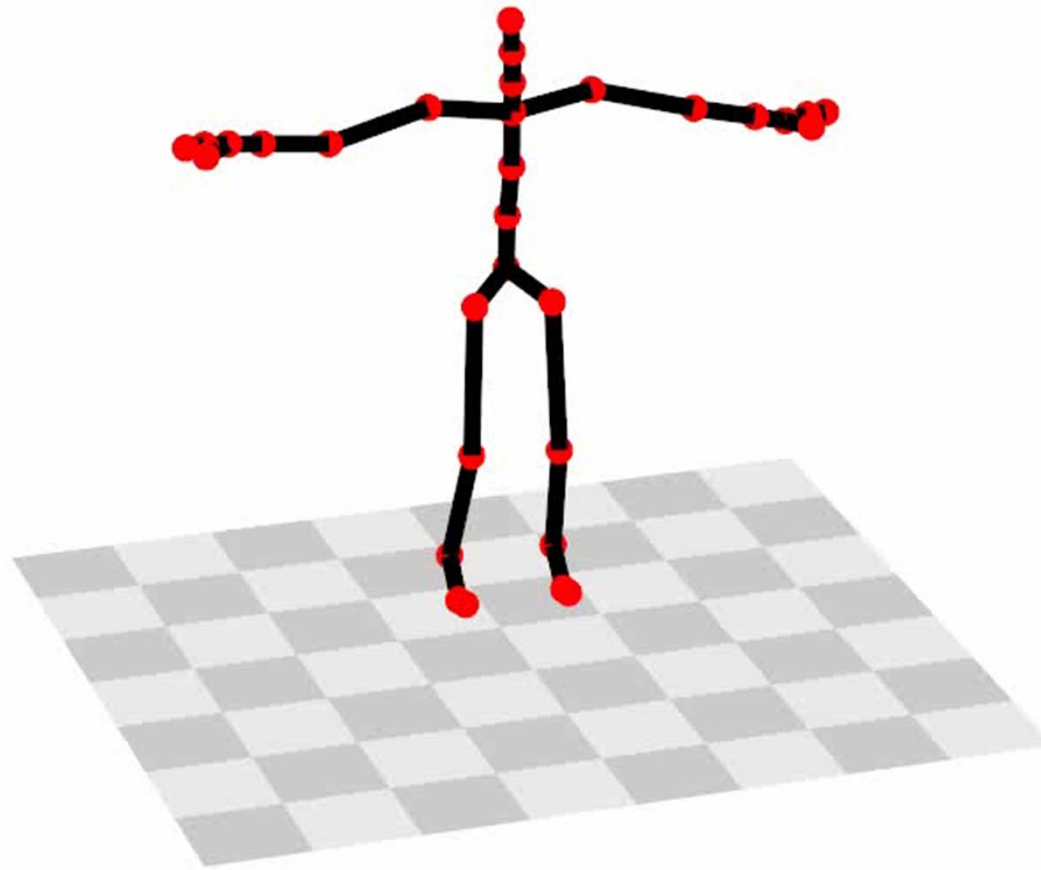
# Motion Templates (MT)

## Quantized template

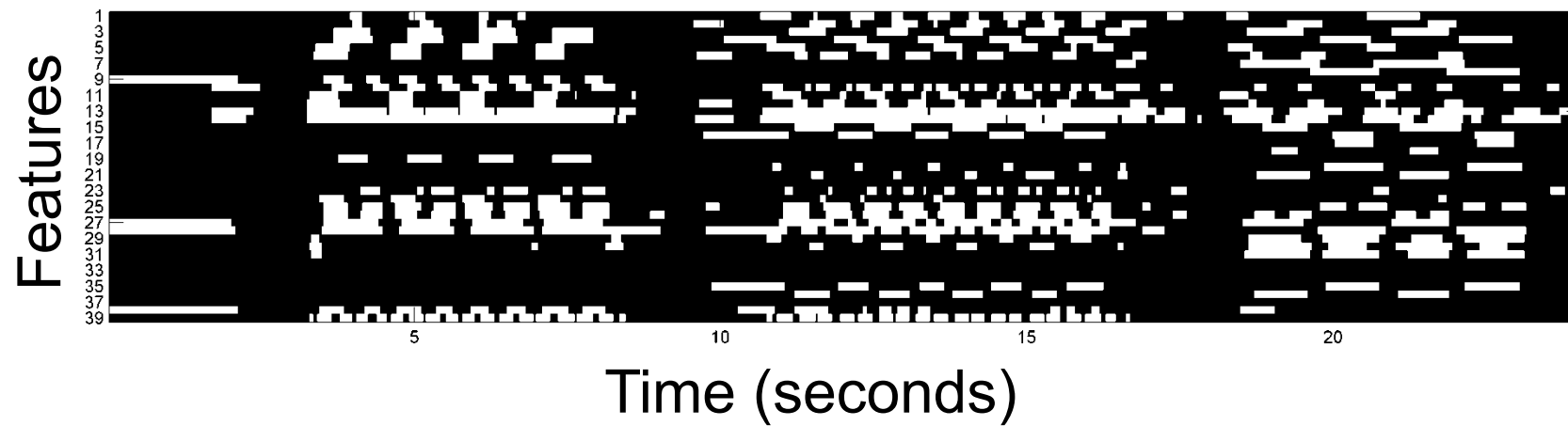


- Gray areas indicate inconsistencies / variations
- Achieve invariance by disregarding gray areas

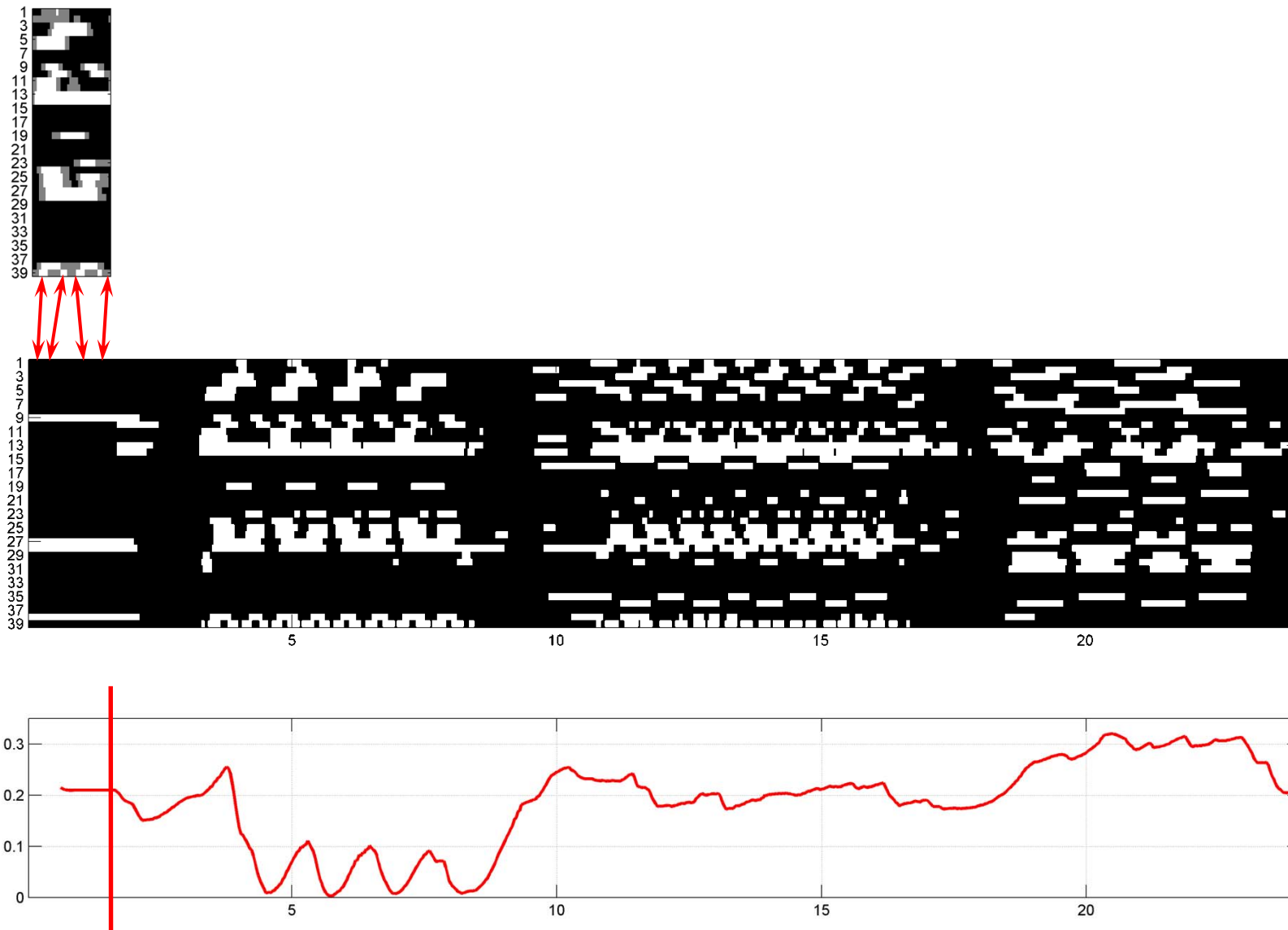
# MT-based Motion Retrieval



# MT-based Motion Retrieval

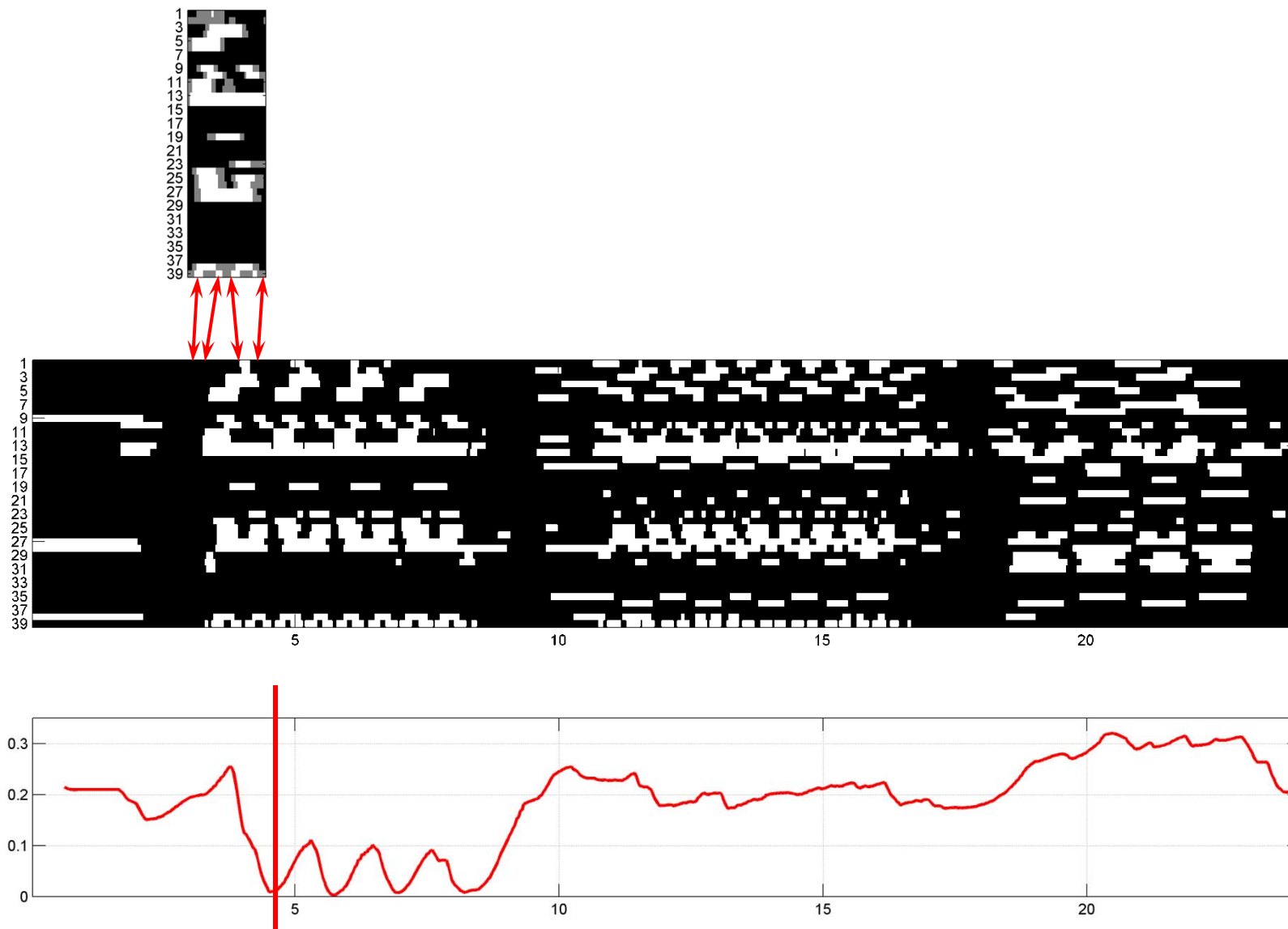


# MT-based Motion Retrieval: Jumping Jack

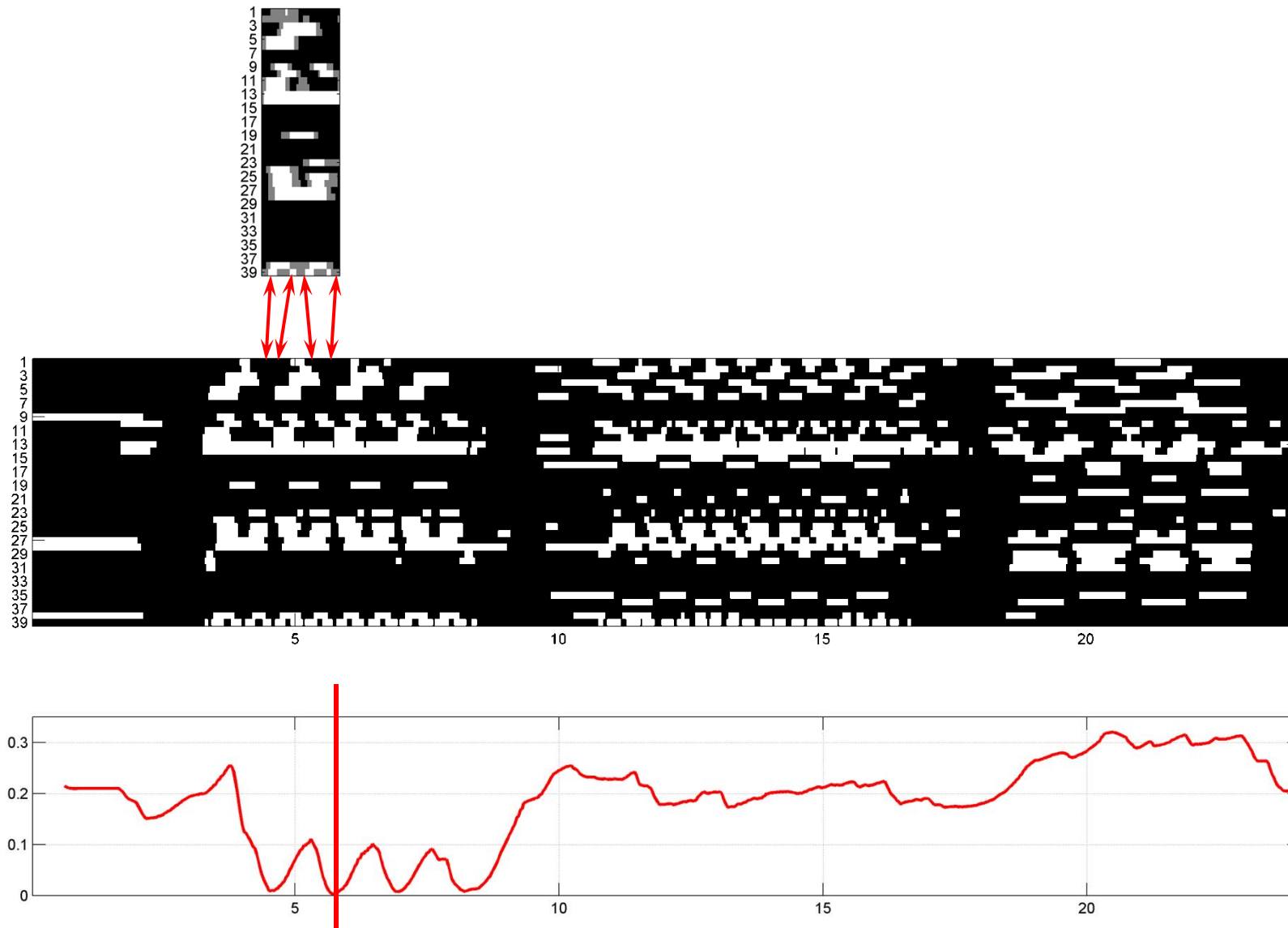




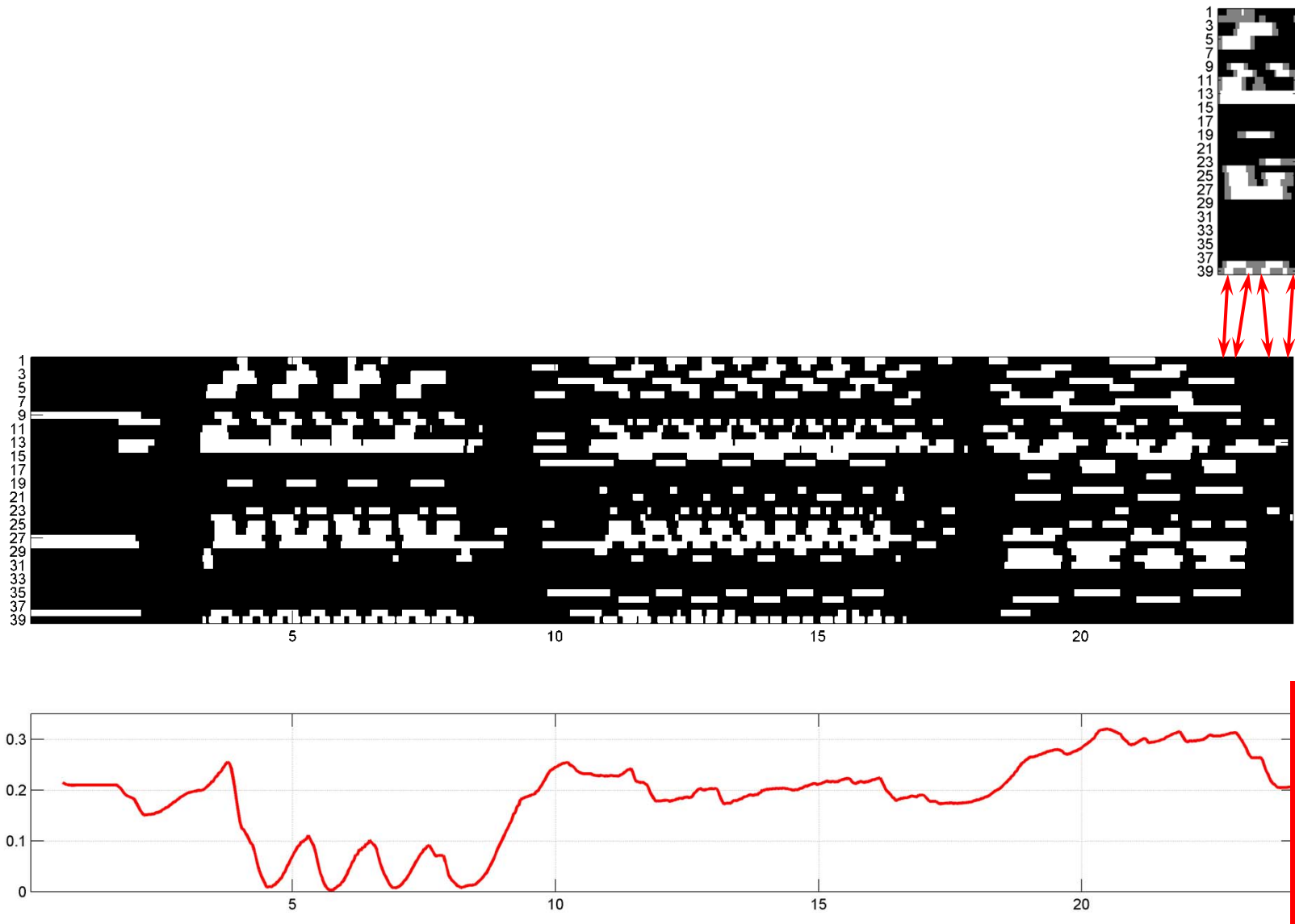
# MT-based Motion Retrieval: Jumping Jack



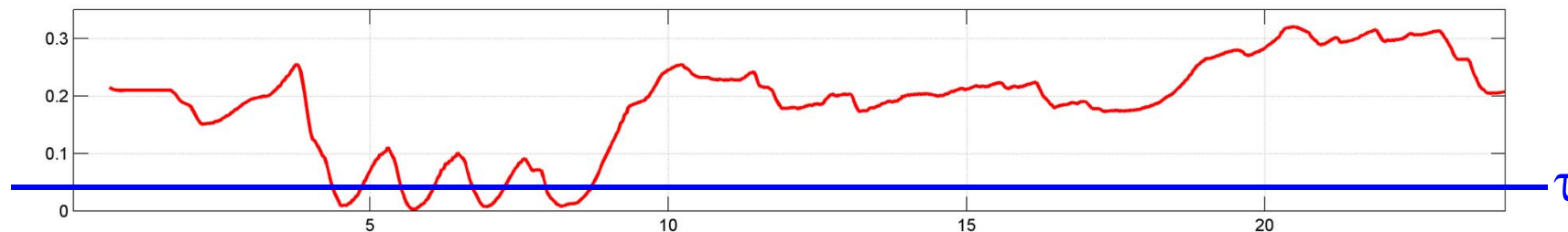
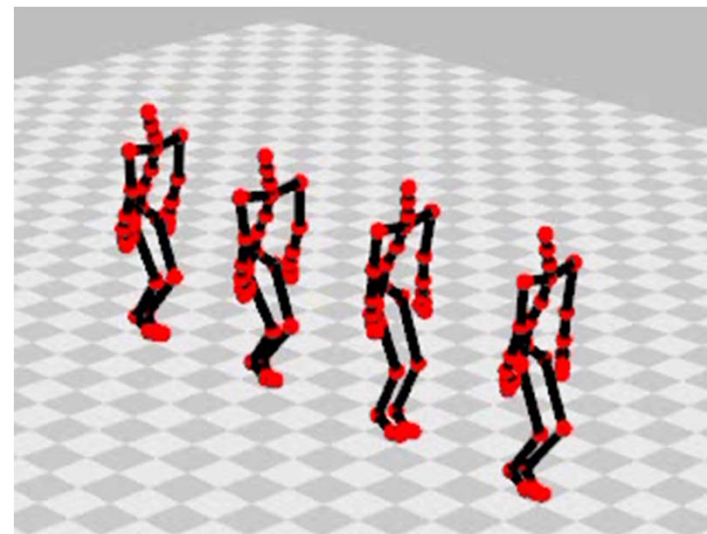
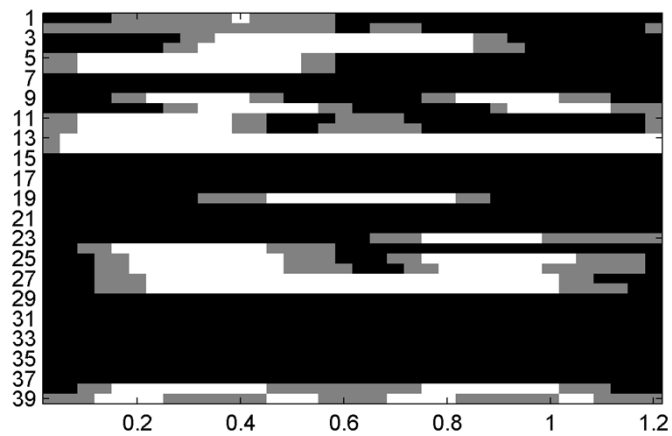
# MT-based Motion Retrieval: Jumping Jack



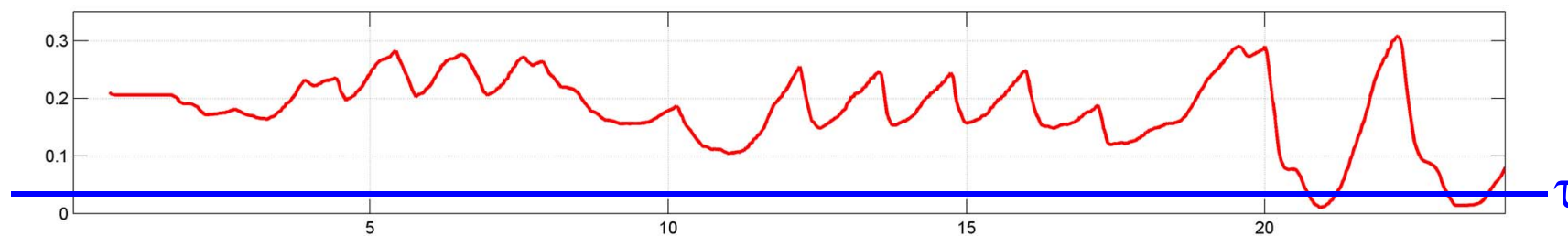
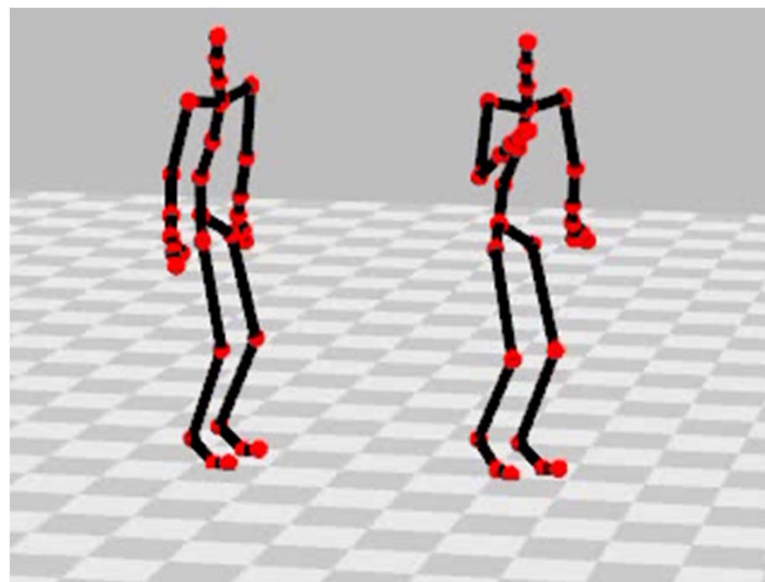
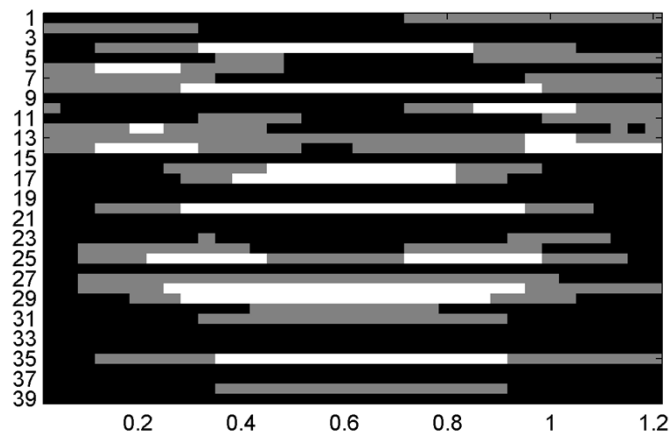
# MT-based Motion Retrieval: Jumping Jack



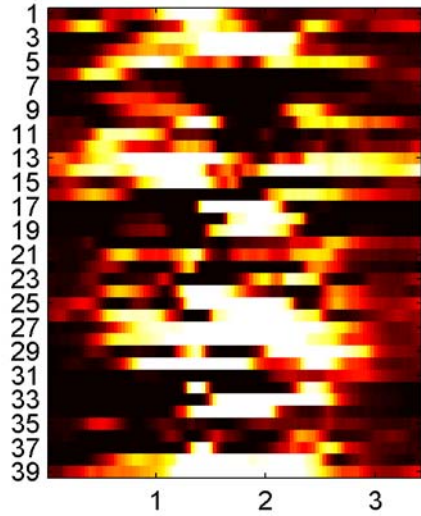
# MT-based Motion Retrieval: Jumping Jack



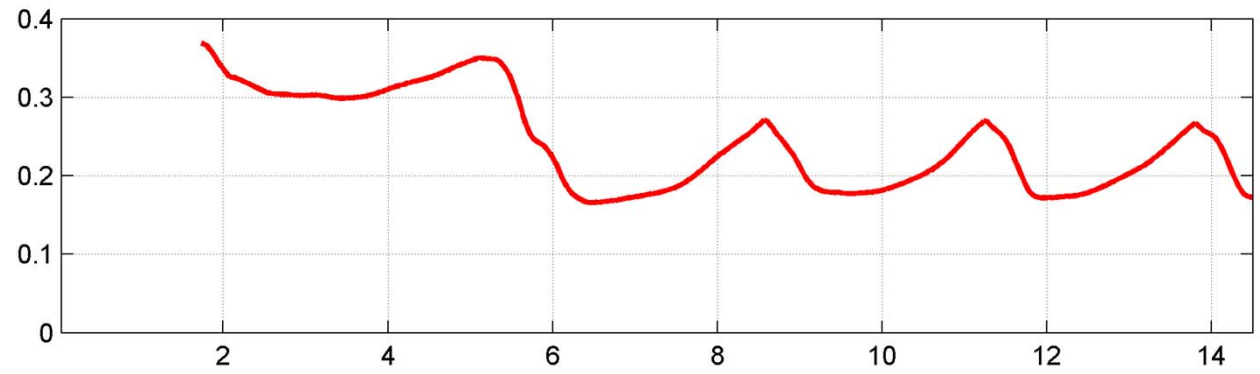
# MT-based Motion Retrieval: Elbow-To-Knee



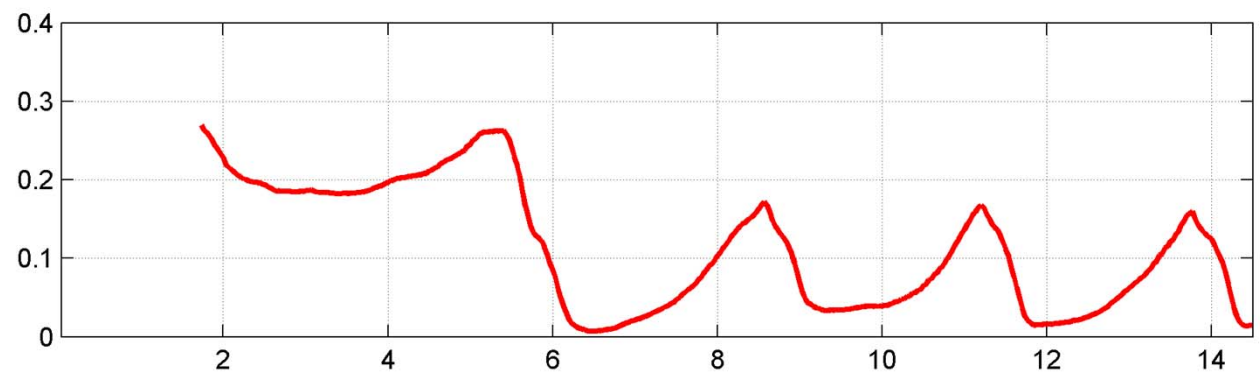
# MT-based Motion Retrieval: Cartwheel



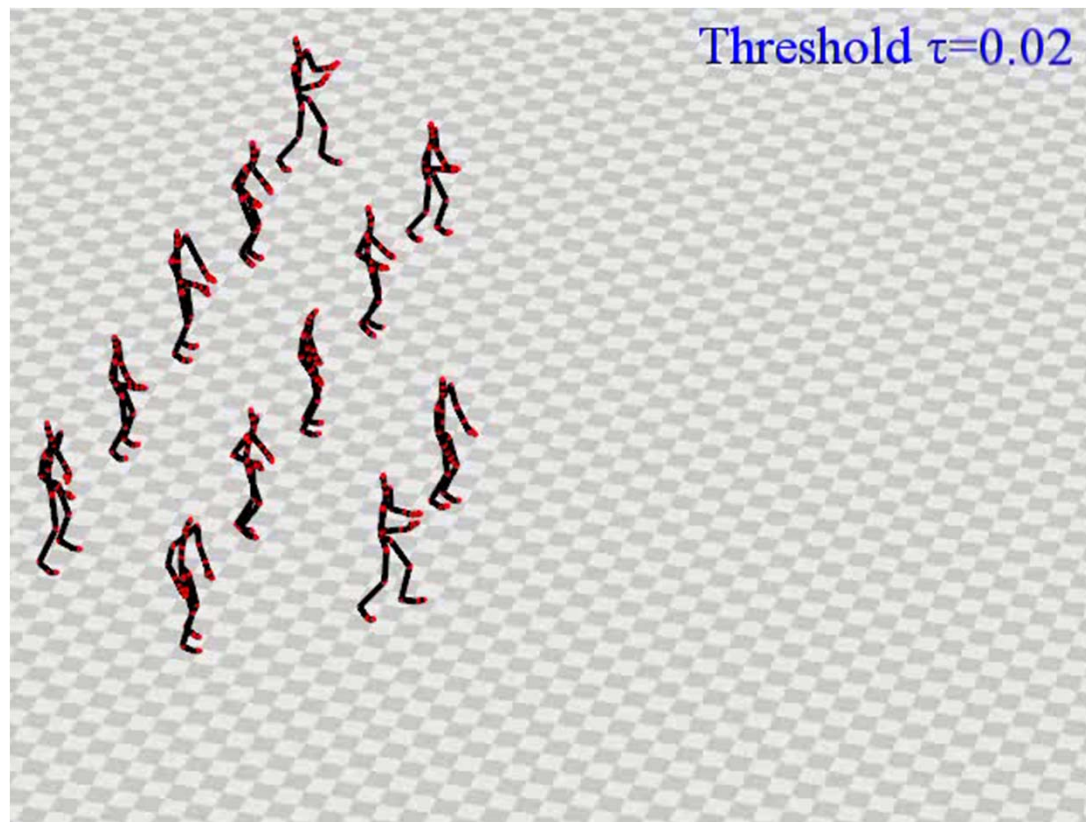
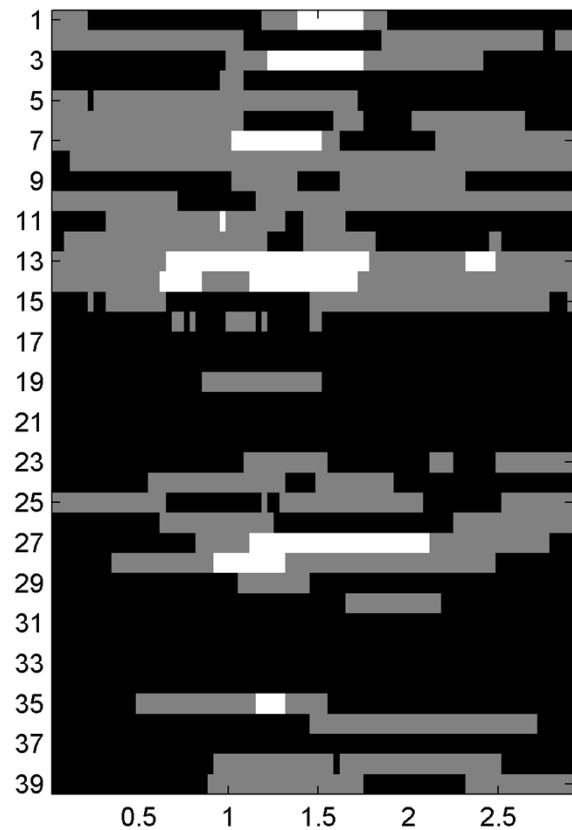
Matching curve using average MT



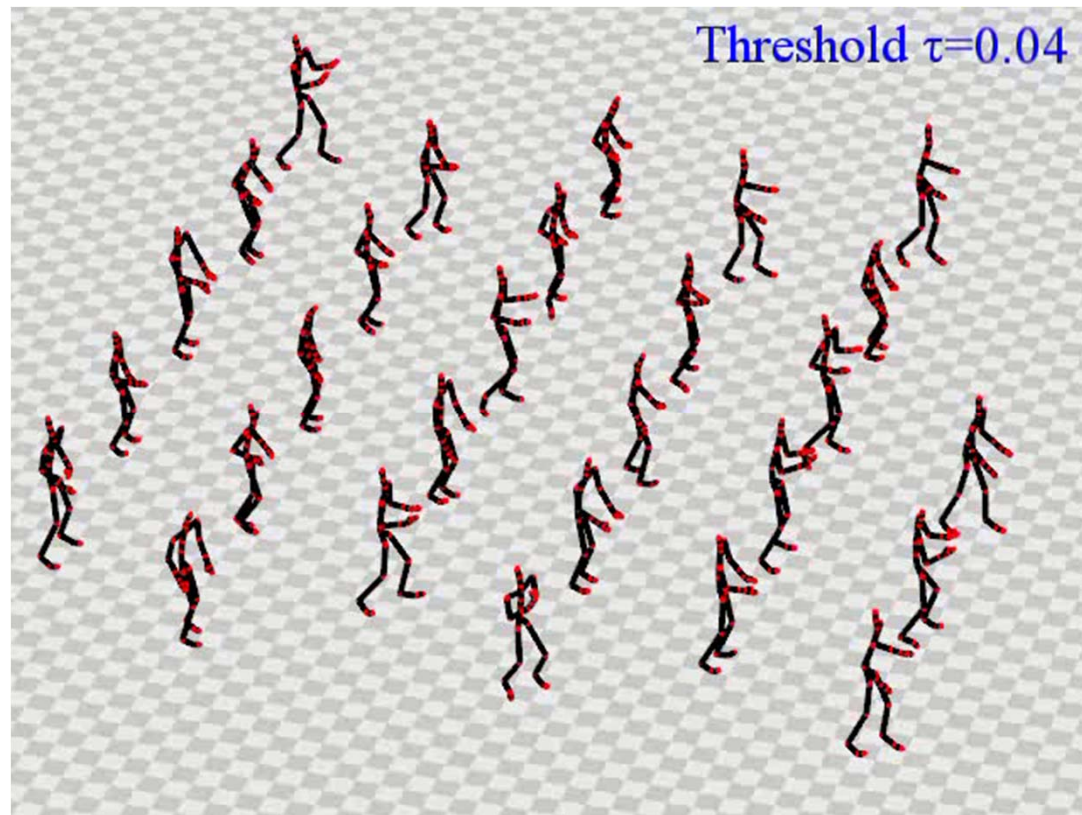
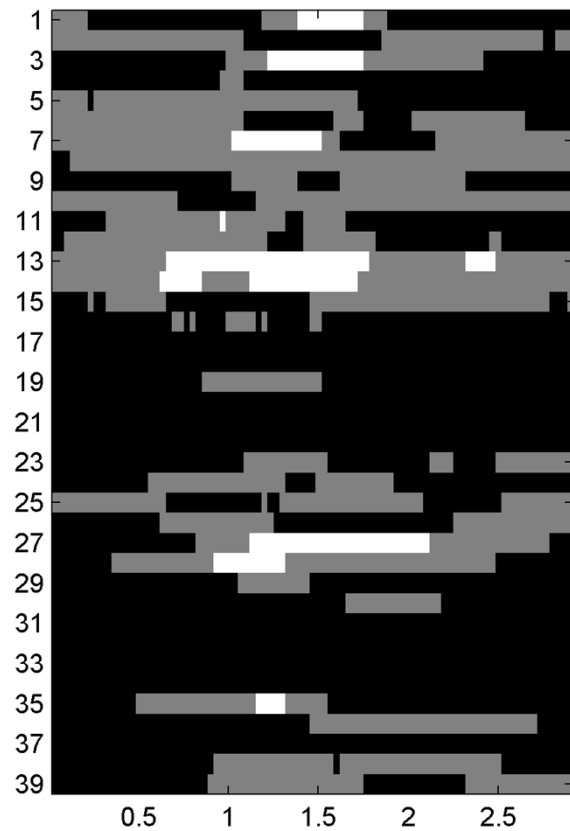
Matching curve blending out variations



# MT-based Motion Retrieval: Throw

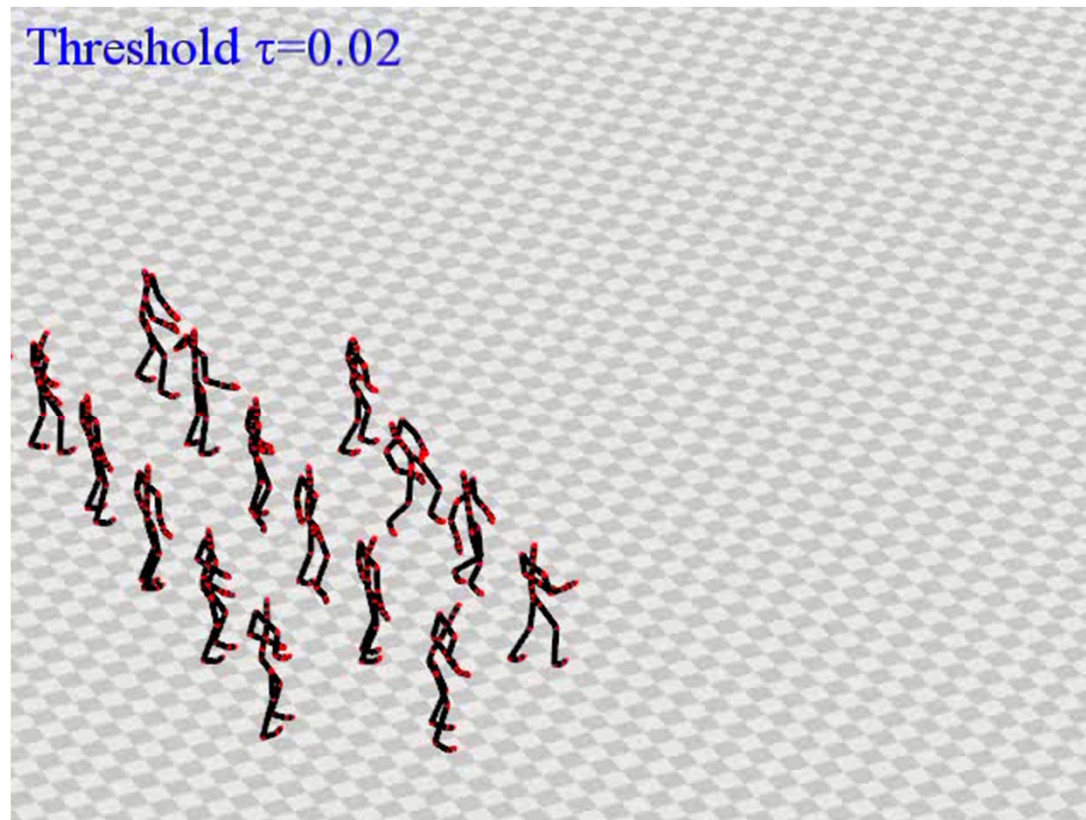
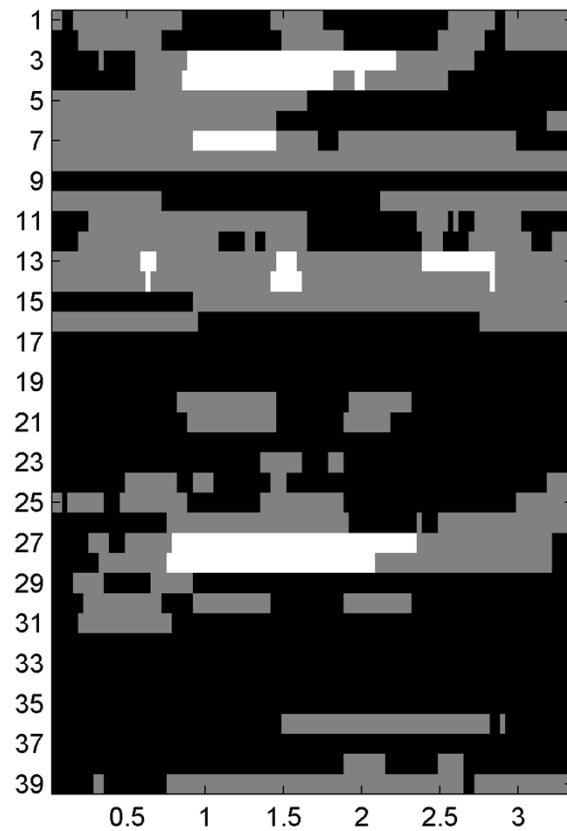


# MT-based Motion Retrieval: Throw

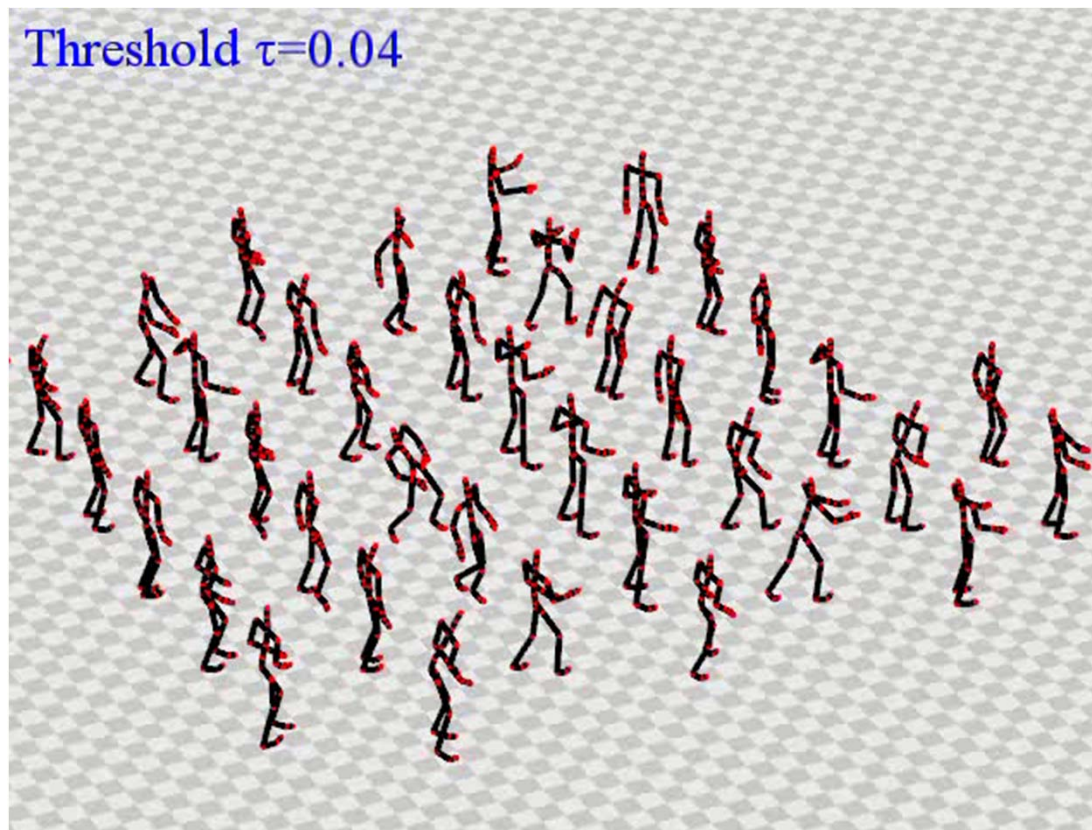
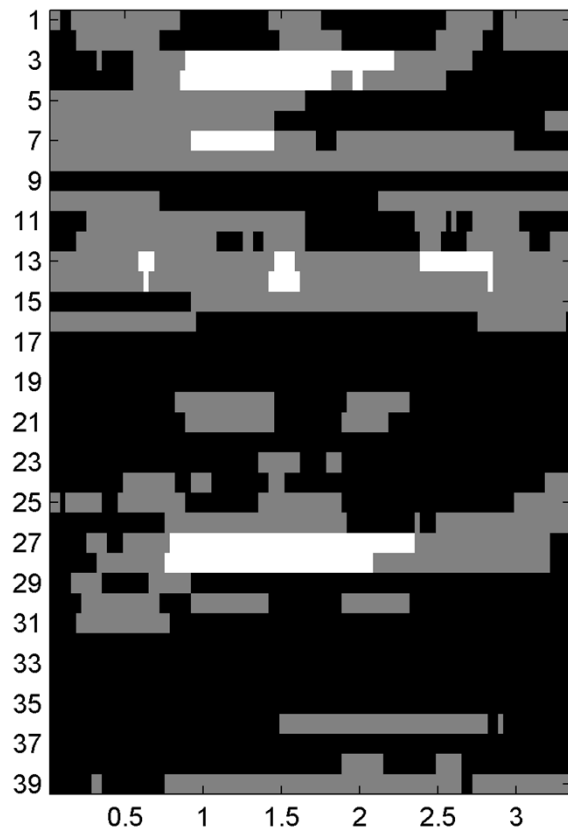




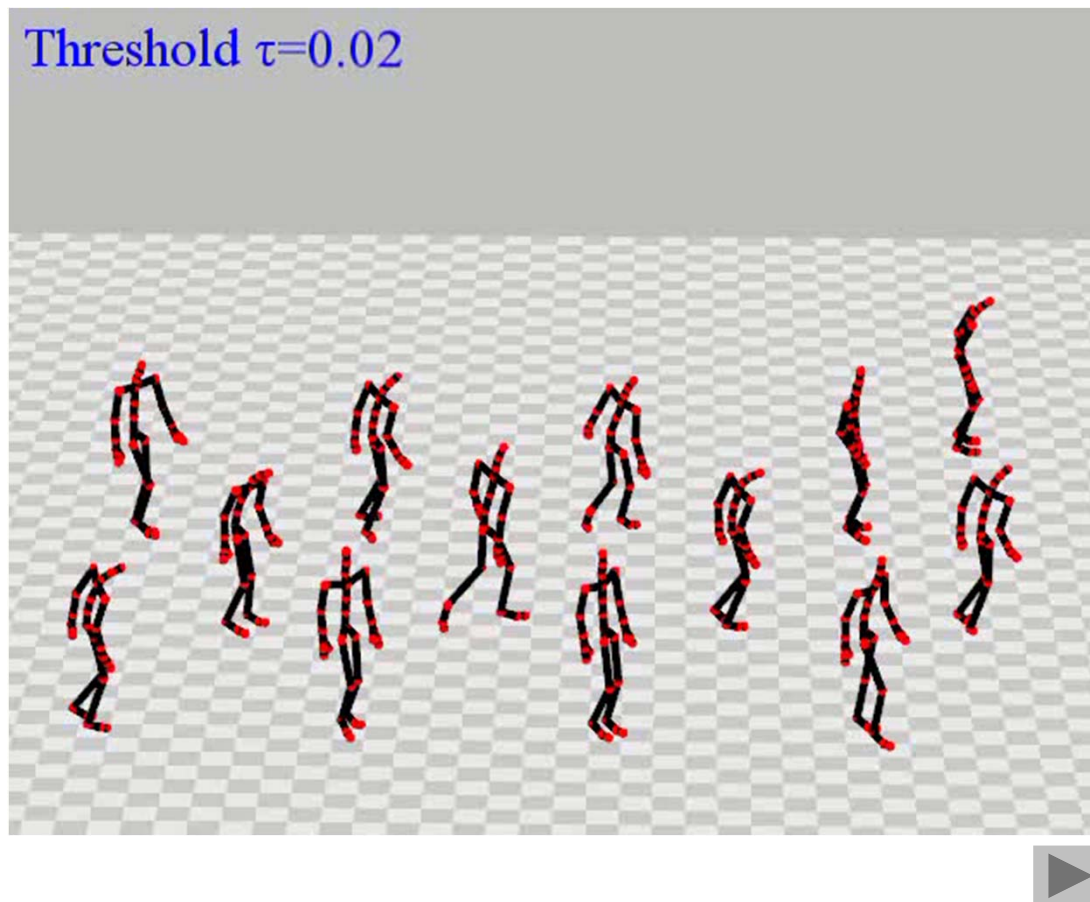
# MT-based Motion Retrieval: Basketball



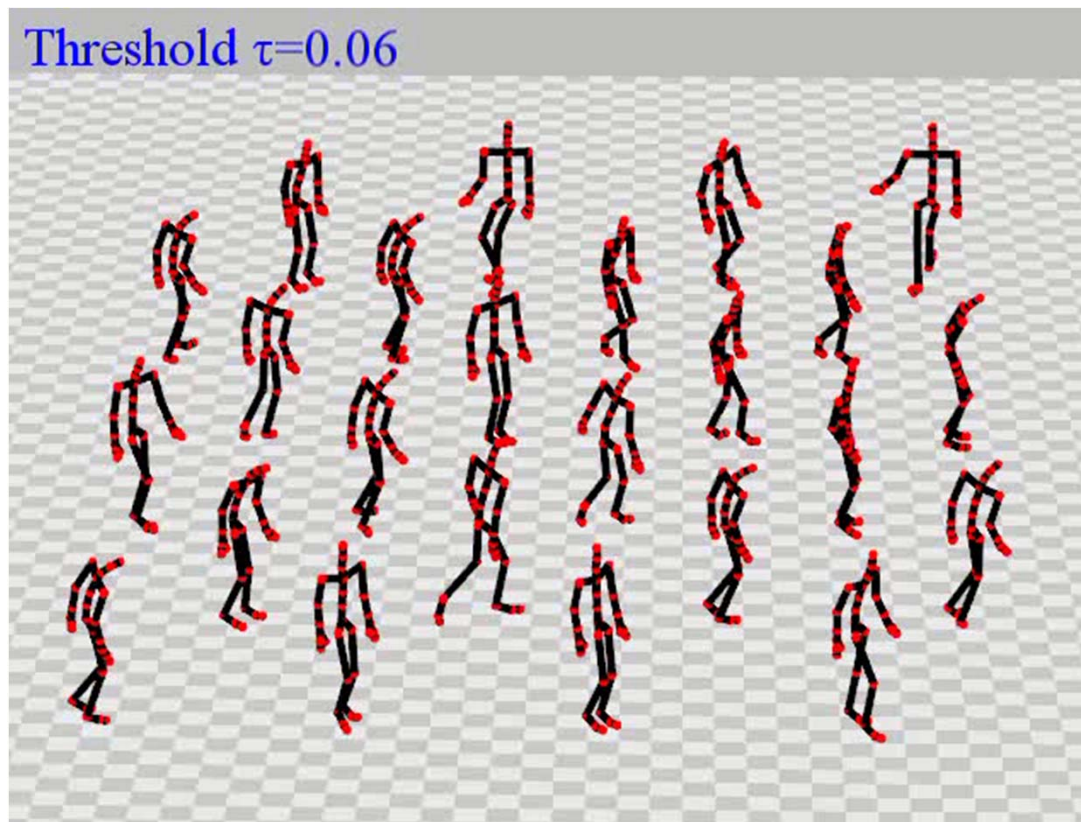
# MT-based Motion Retrieval: Basketball



# MT-based Motion Retrieval: Lie Down Floor



# MT-based Motion Retrieval: Lie Down Floor



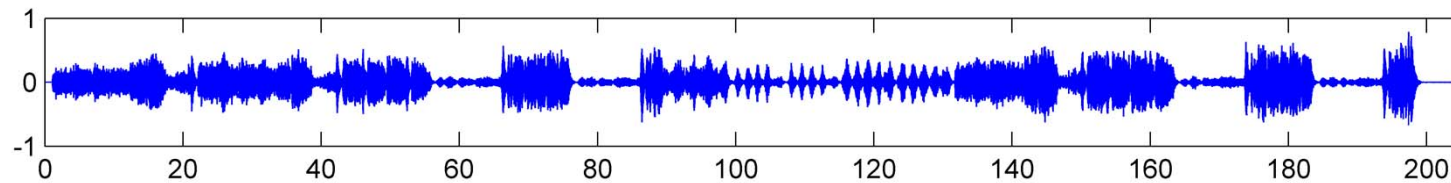
---

# Overview

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

# Music Structure Analysis

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)

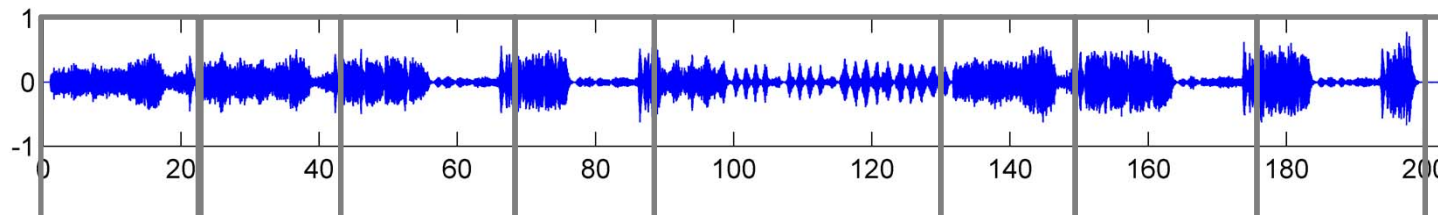


Time (seconds)



# Music Structure Analysis

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)

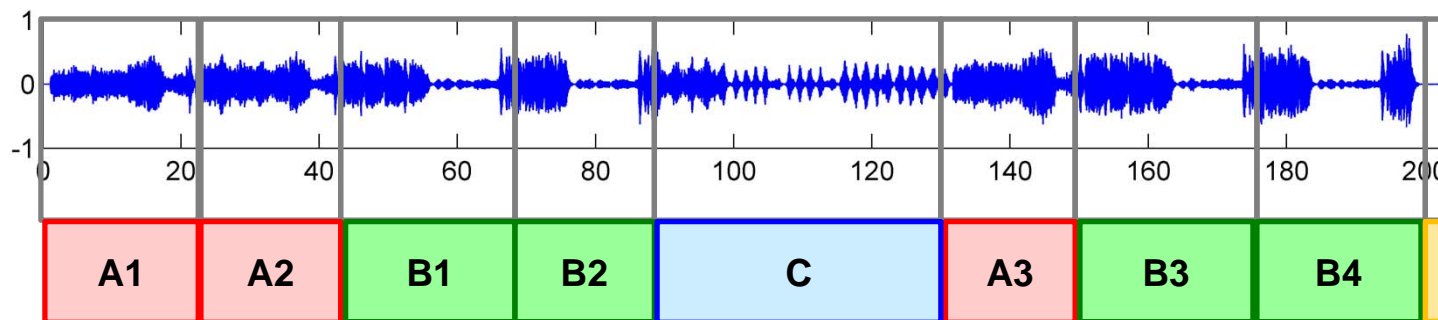


Time (seconds)



# 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.

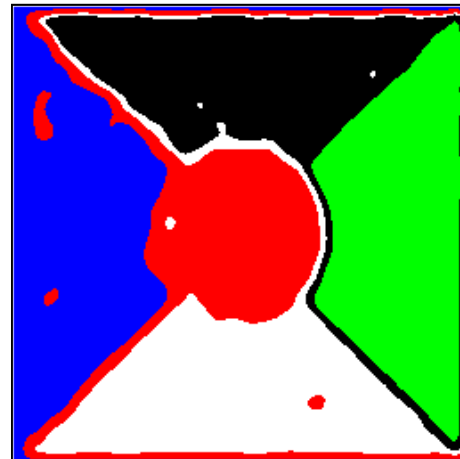
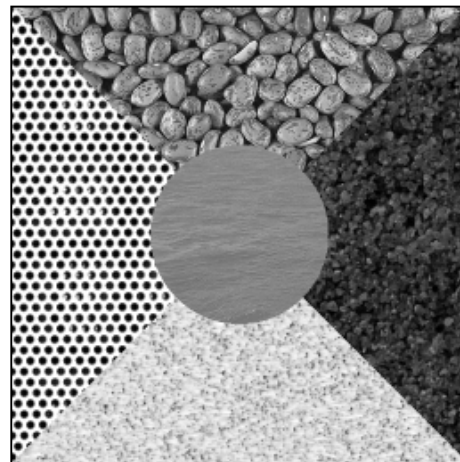
- **Homogeneity:** Consistency in tempo, instrumentation, key, ...
- **Novelty:** Sudden changes, surprising elements ...
- **Repetition:** Repeating themes, motives, rhythmic patterns,...

# Music Structure Analysis

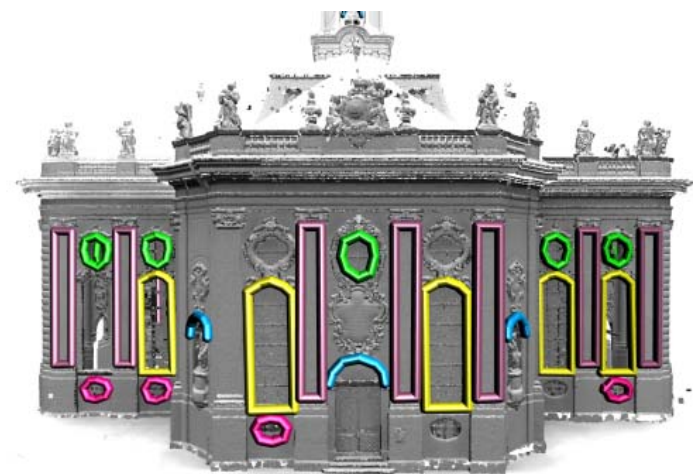
Novelty



Homogeneity



Repetition



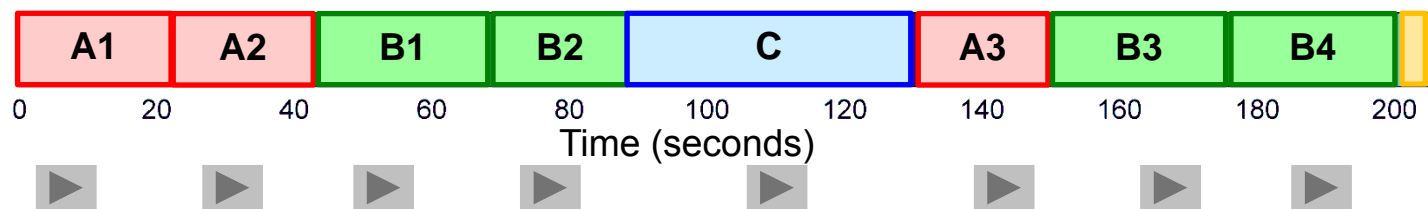
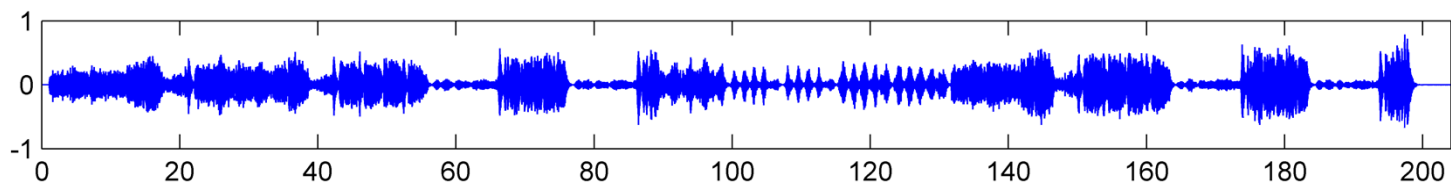
# Feature Representation

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

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

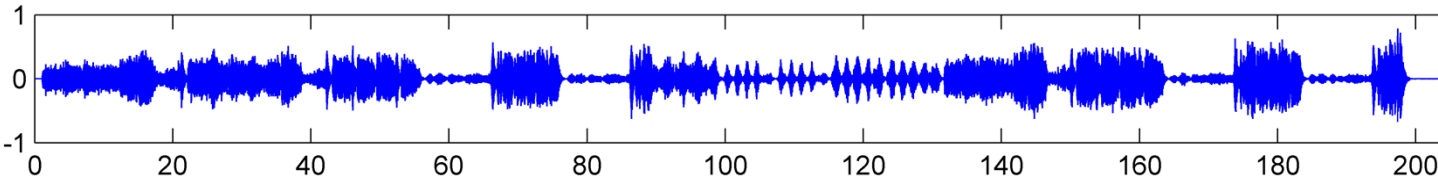
# Feature Representation

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



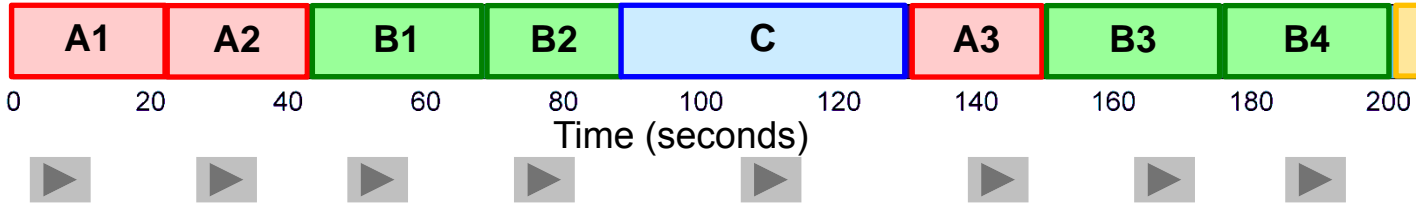
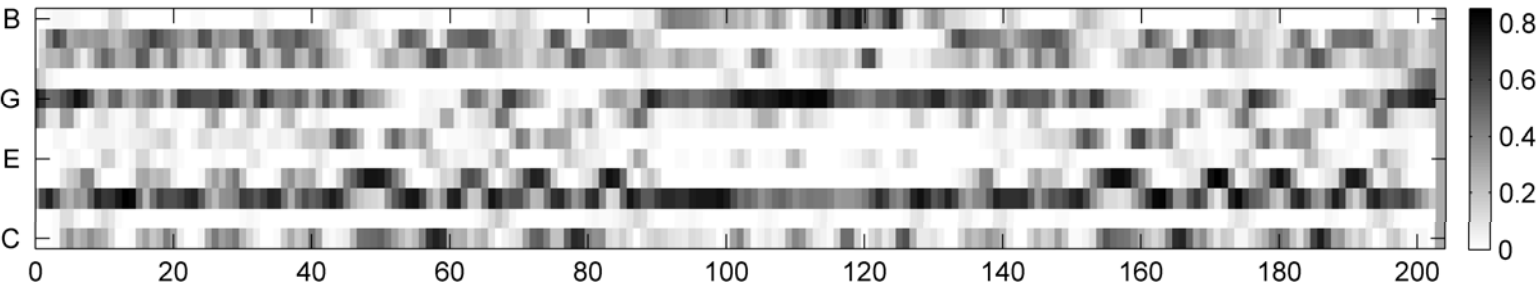
# Feature Representation

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



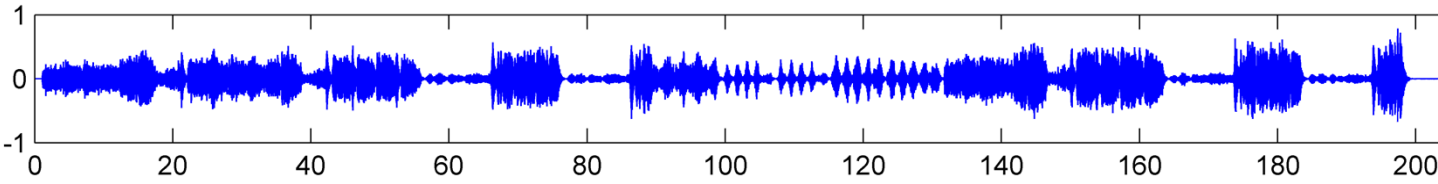
Feature extraction

Chroma (Harmony)

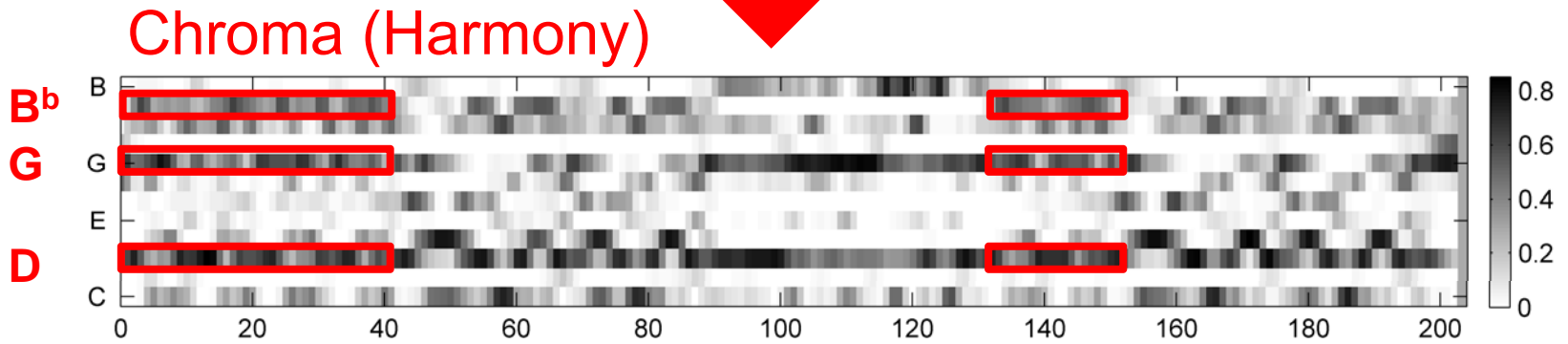


# Feature Representation

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



Feature extraction



G minor

G minor

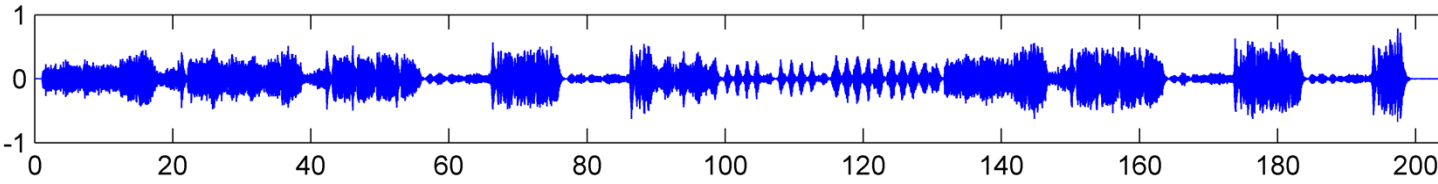


Time (seconds)

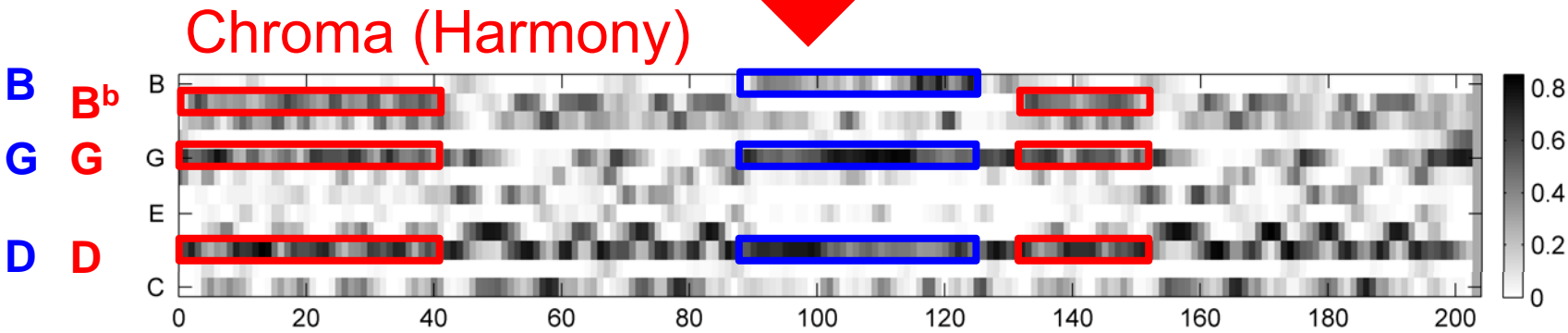


# Feature Representation

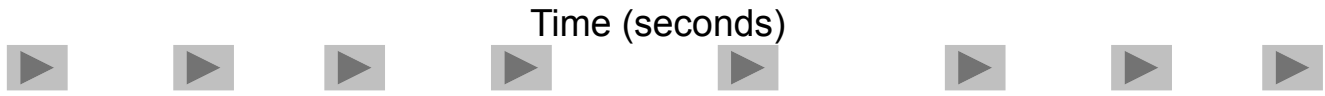
**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



Feature extraction



G minor G major G minor





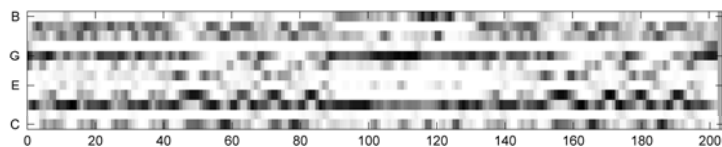
# Self-Similarity Matrix (SSM)

**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

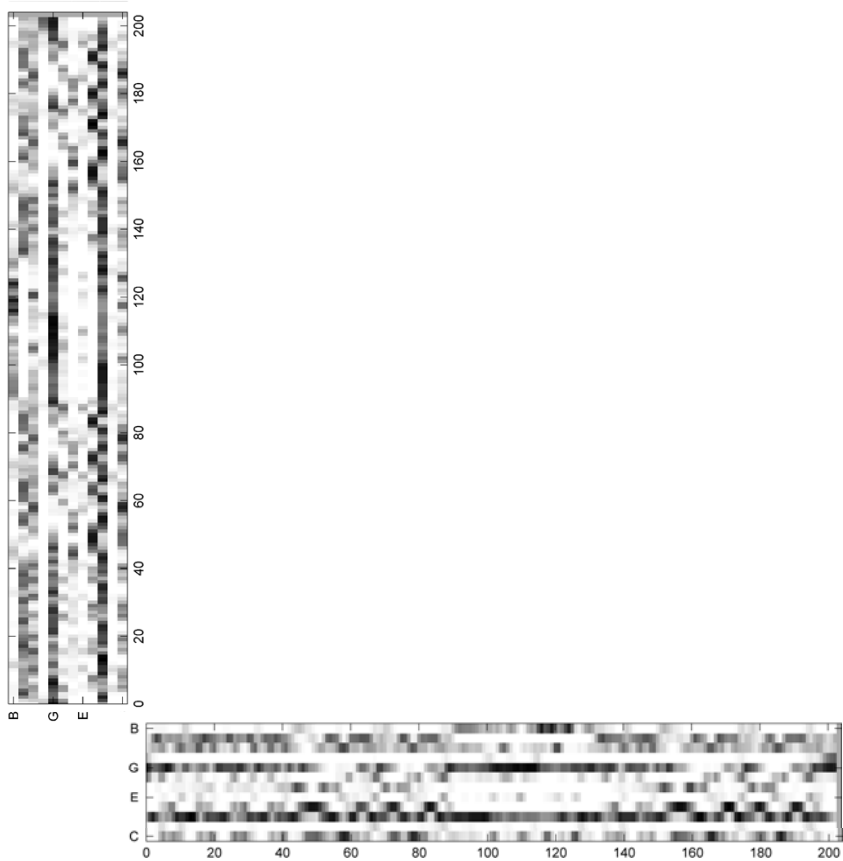
# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



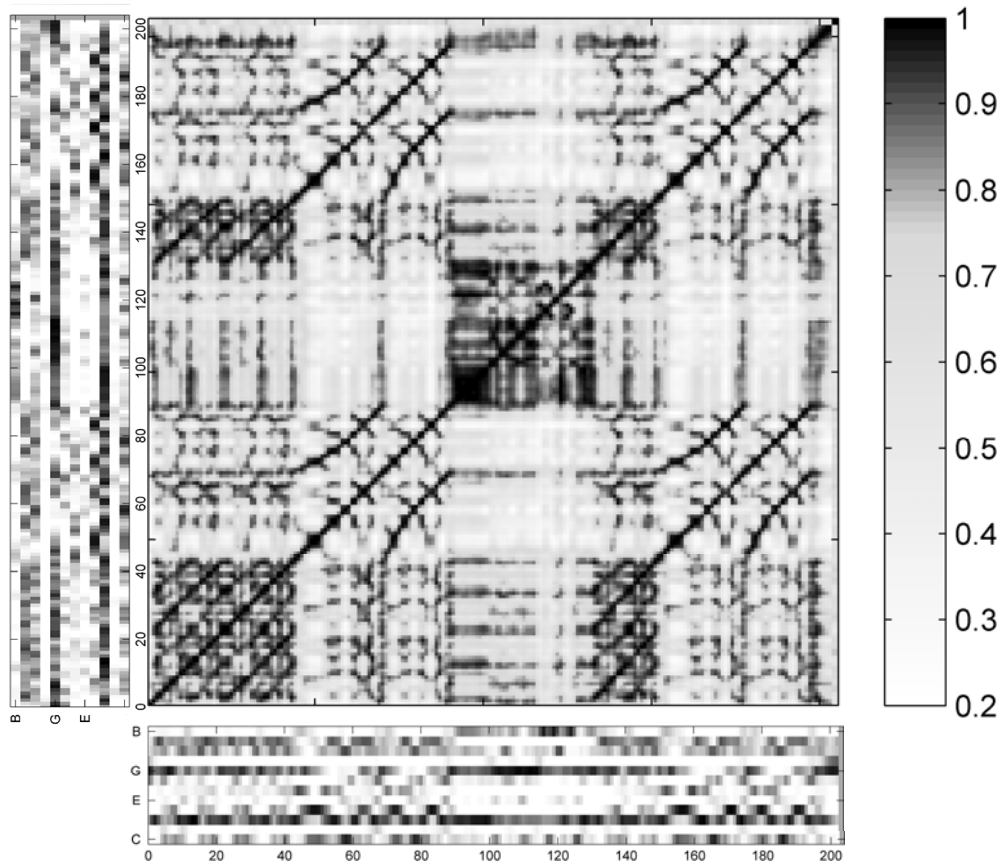
# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



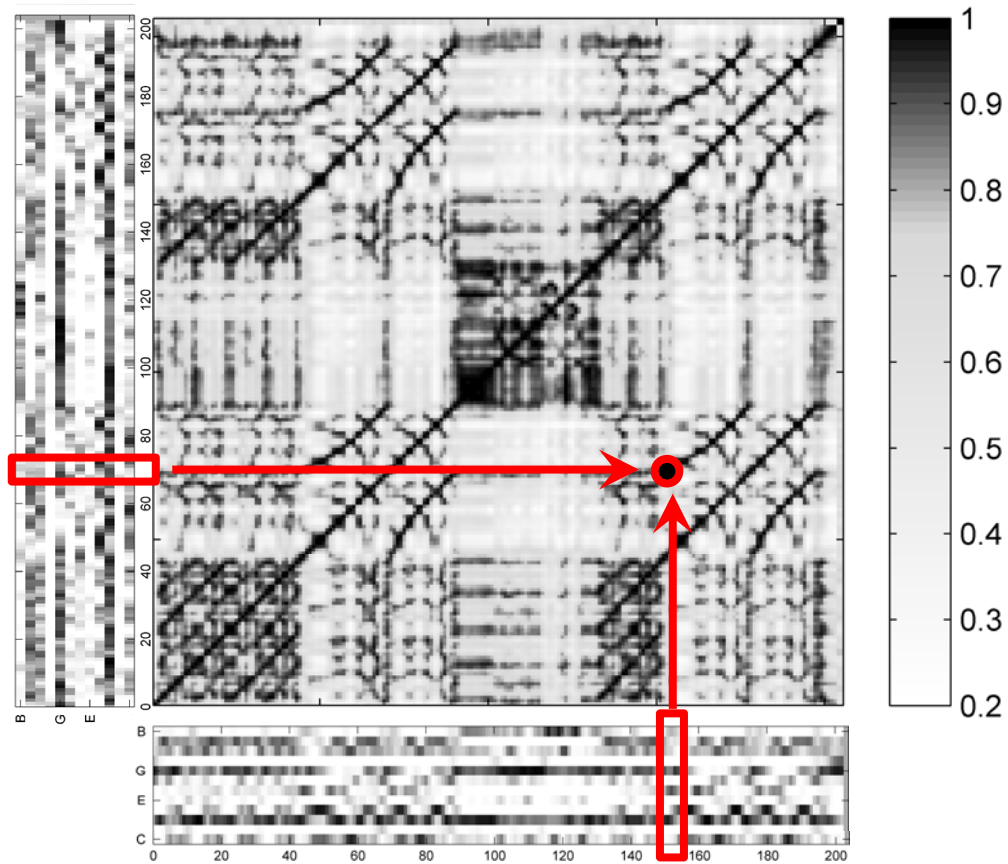
# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



# Self-Similarity Matrix (SSM)

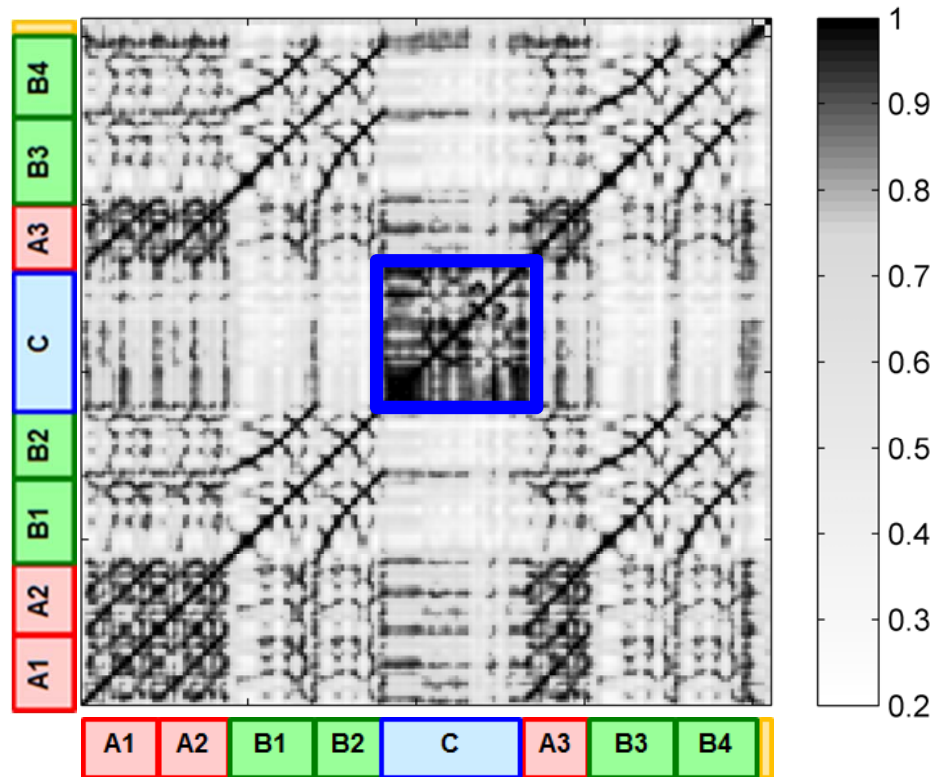
**Example:** Brahms Hungarian Dance No. 5 (Ormandy)





# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)

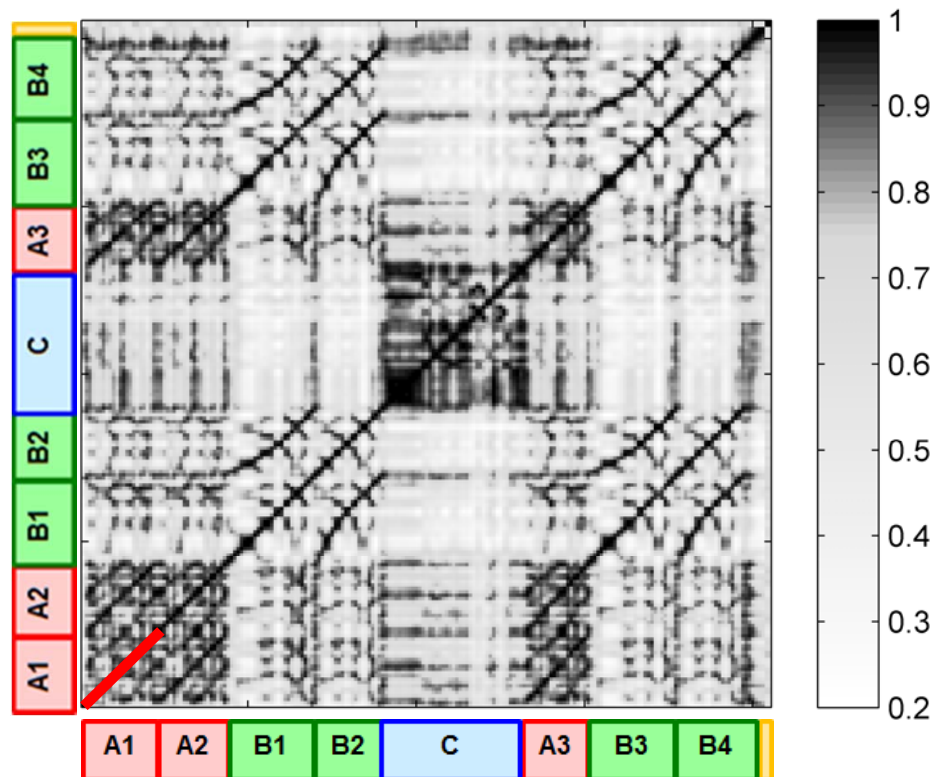






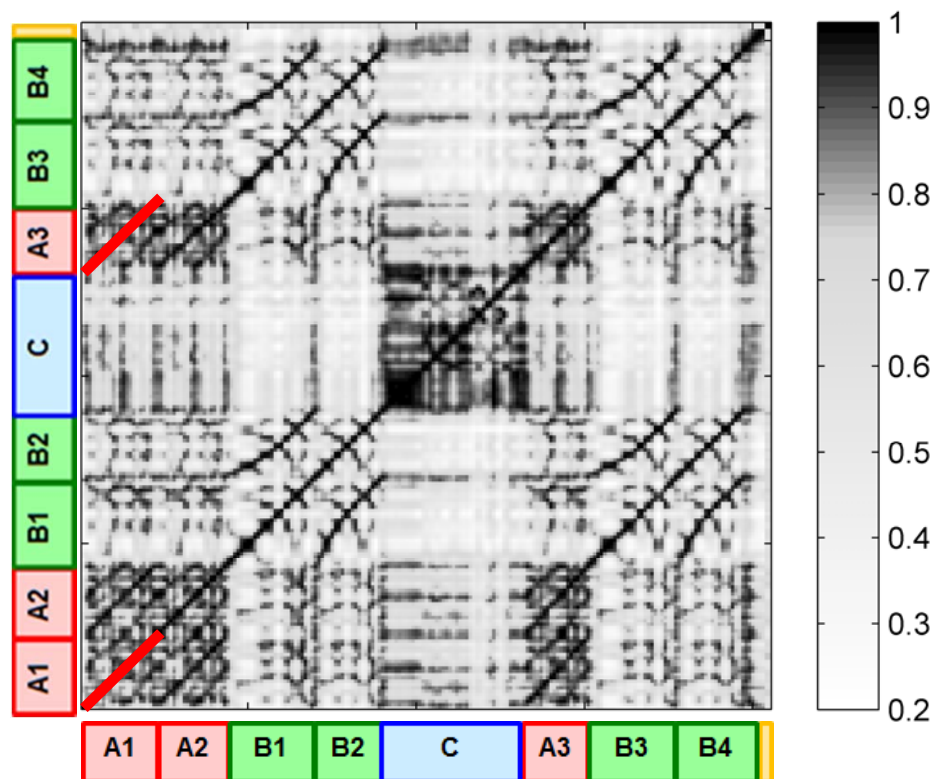
# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)

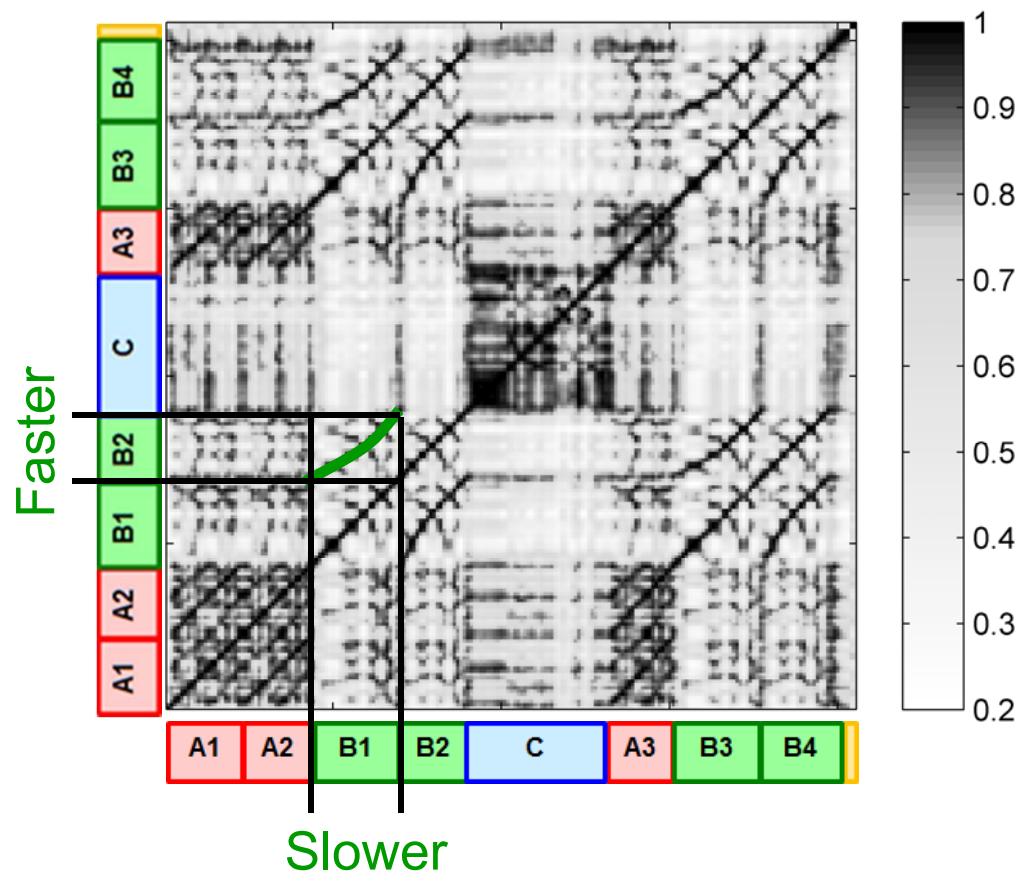






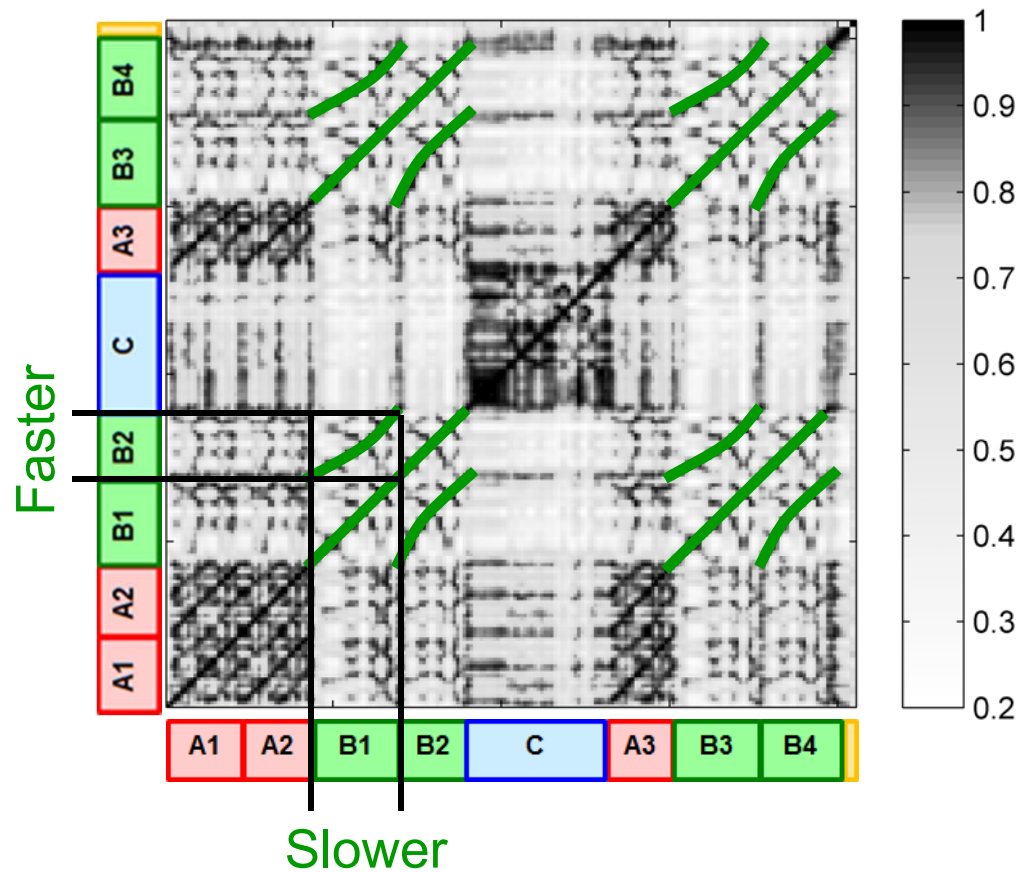
# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)



# Self-Similarity Matrix (SSM)

**Example:** Brahms Hungarian Dance No. 5 (Ormandy)





# Self-Similarity Matrix (SSM)

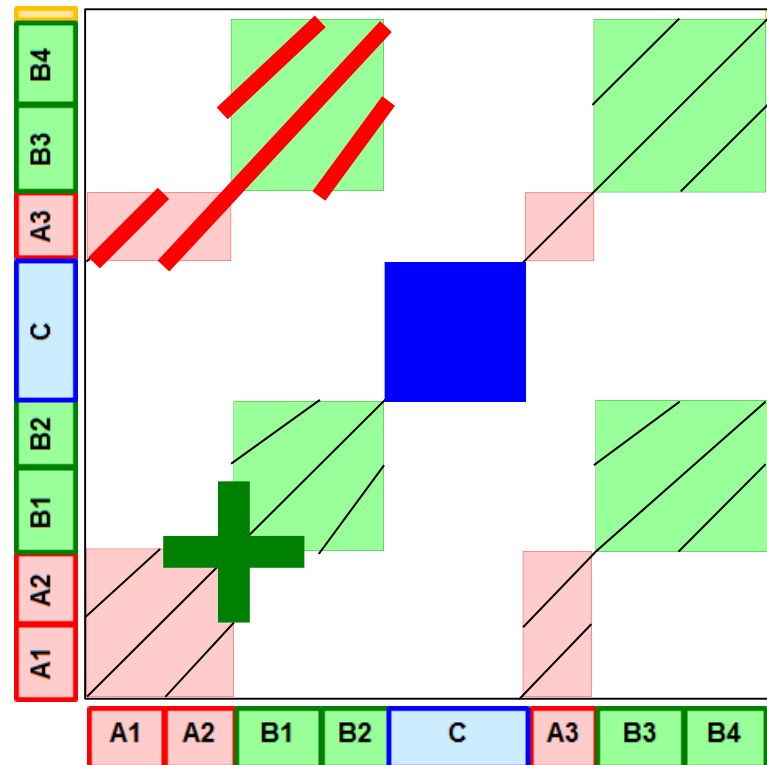
**Example:** Brahms Hungarian Dance No. 5 (Ormandy)

**Blocks:** Homogeneity

**Paths:** Repetition

**Corners:** Novelty

Idealized SSM



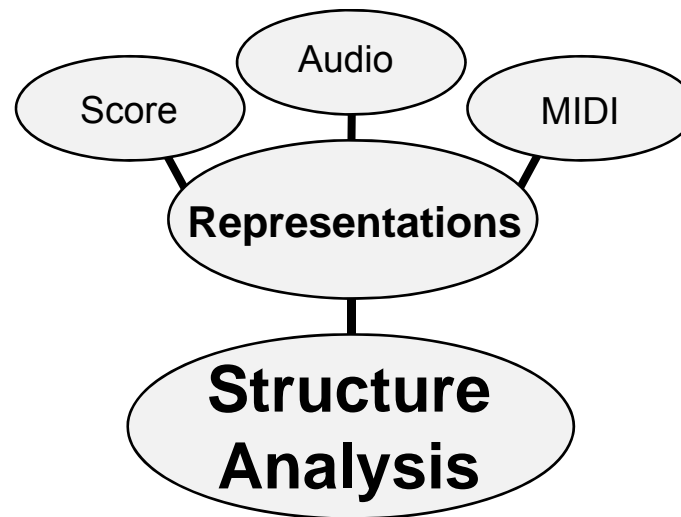


---

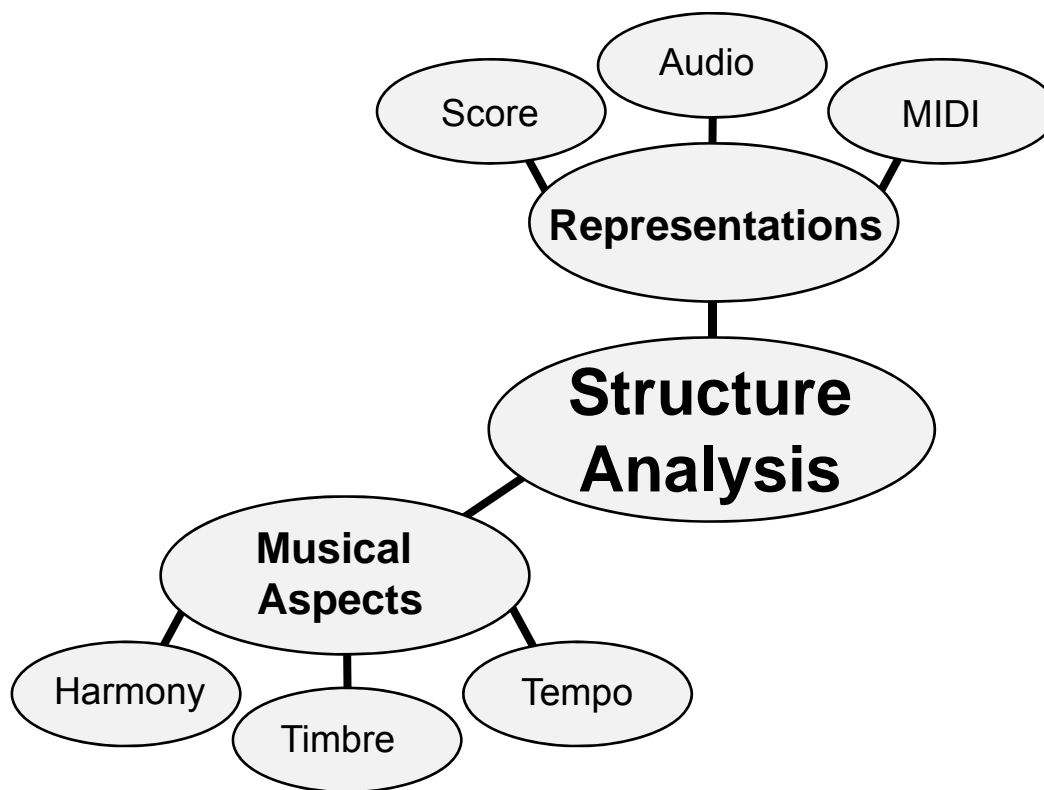
# Music Structure Analysis

**Structure  
Analysis**

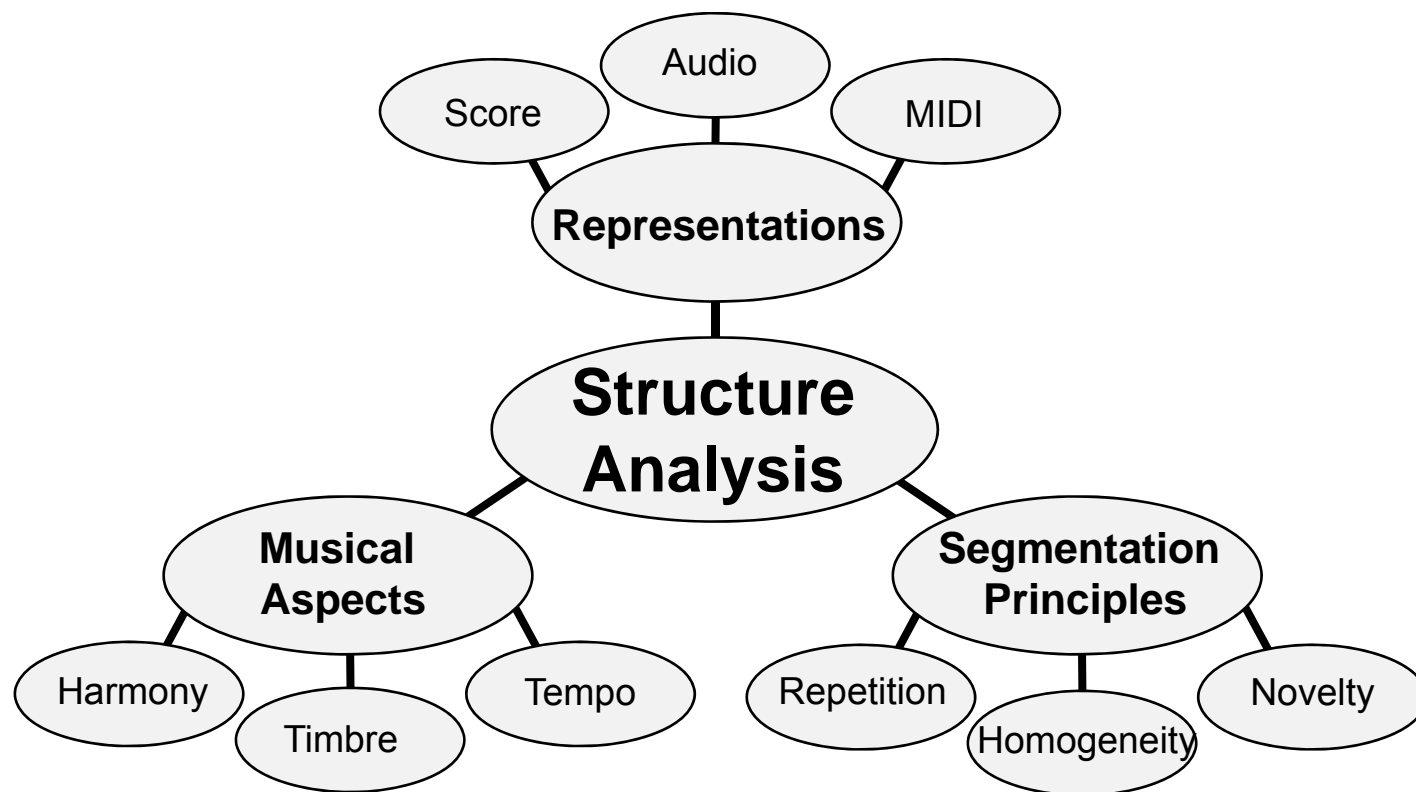
# Music Structure Analysis



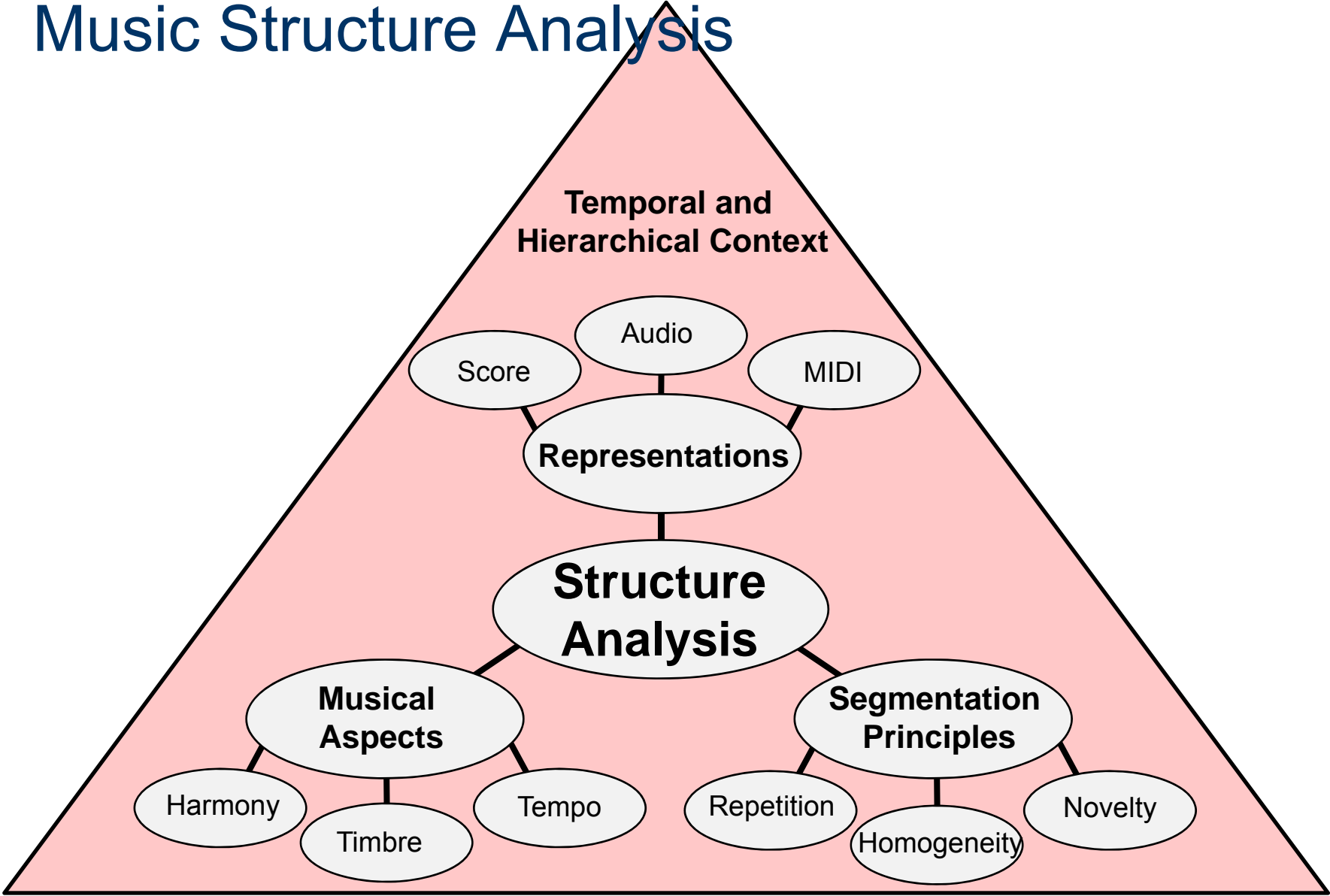
# Music Structure Analysis



# Music Structure Analysis



# Music Structure Analysis



---

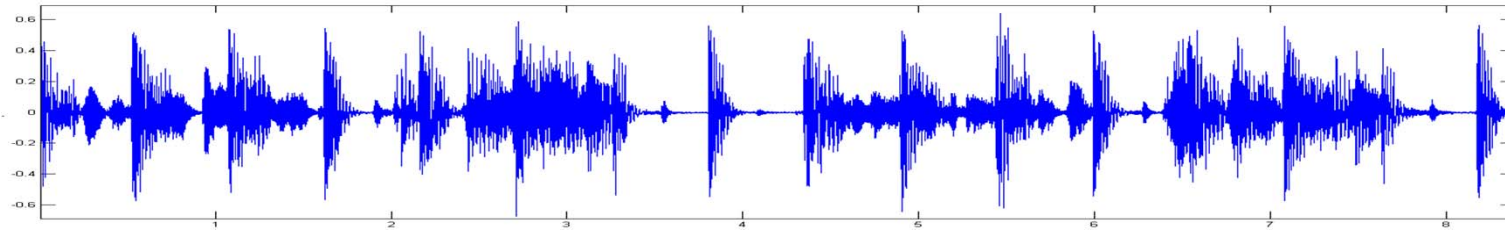
# Overview

- 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

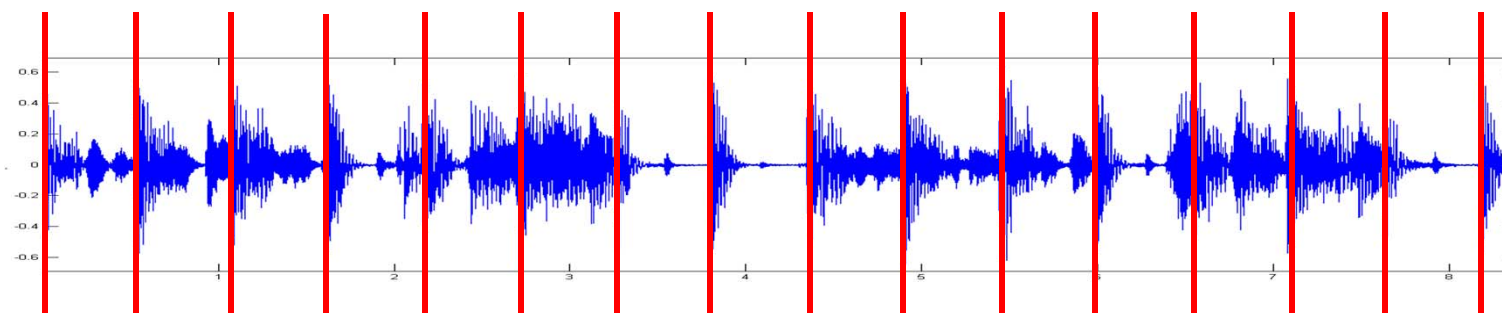


Time (seconds)

# 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**

The image shows two staves of musical notation for the song 'Happy Birthday to you'. The first staff contains the first two phrases: 'Hap - py Birth - day to you,' and 'Hap - py Birth - day to you, Hap - py'. The second staff contains the third phrase: 'Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!'. The music is in 3/4 time with a key signature of one sharp (F#). Four red arrows point downwards to the first note of each of the four measures in the first staff, illustrating the pulse level at the measure level.

# Beat Tracking

Example: Happy Birthday to you

Pulse level: **Tactus (beat)**

The image shows a musical score for the song "Happy Birthday to you" in 3/4 time. The score is written on two staves. The first staff contains the melody for the first two phrases: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py". The second staff contains the melody for the final phrase: "Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!". Above the first staff, there are 13 red arrows pointing downwards to the first note of each measure, indicating the pulse level (Tactus) for each beat. The key signature is one sharp (F#) and the time signature is 3/4.

# Beat Tracking

Example: Happy Birthday to you

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

The image shows a musical score for the song "Happy Birthday to you" in 3/4 time. The score is written on two staves. The first staff contains the melody for the first two phrases: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py". The second staff contains the melody for the final phrase: "Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!". Above the first staff, there are 24 red arrows pointing downwards, one for each note, indicating the pulse level (Tatum) for each note. The arrows are evenly spaced, corresponding to the 3/4 time signature.

---

# Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: ???



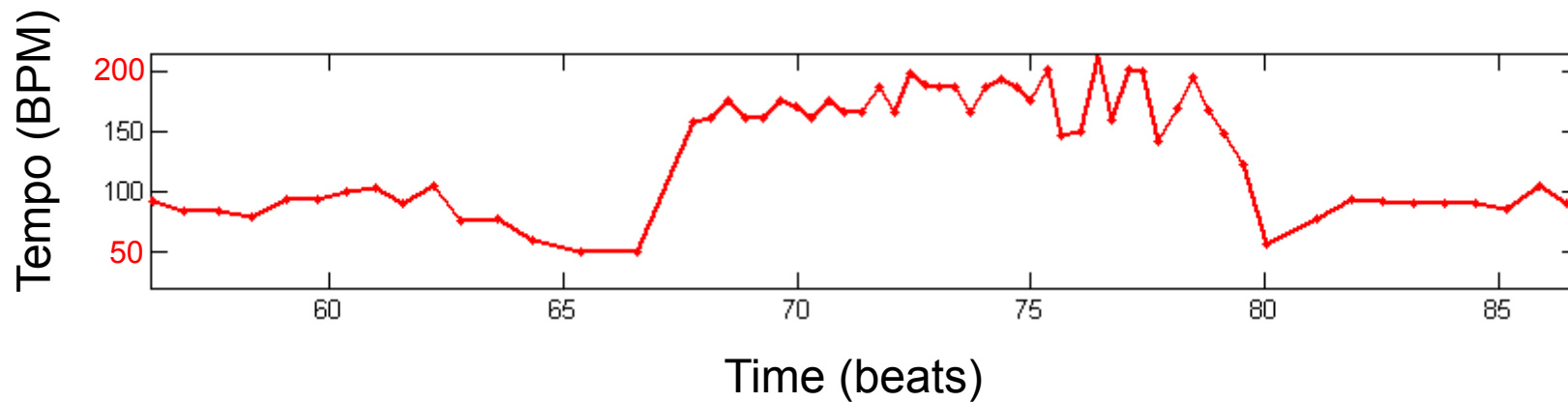
# Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: **50-200 BPM** 

Tempo curve

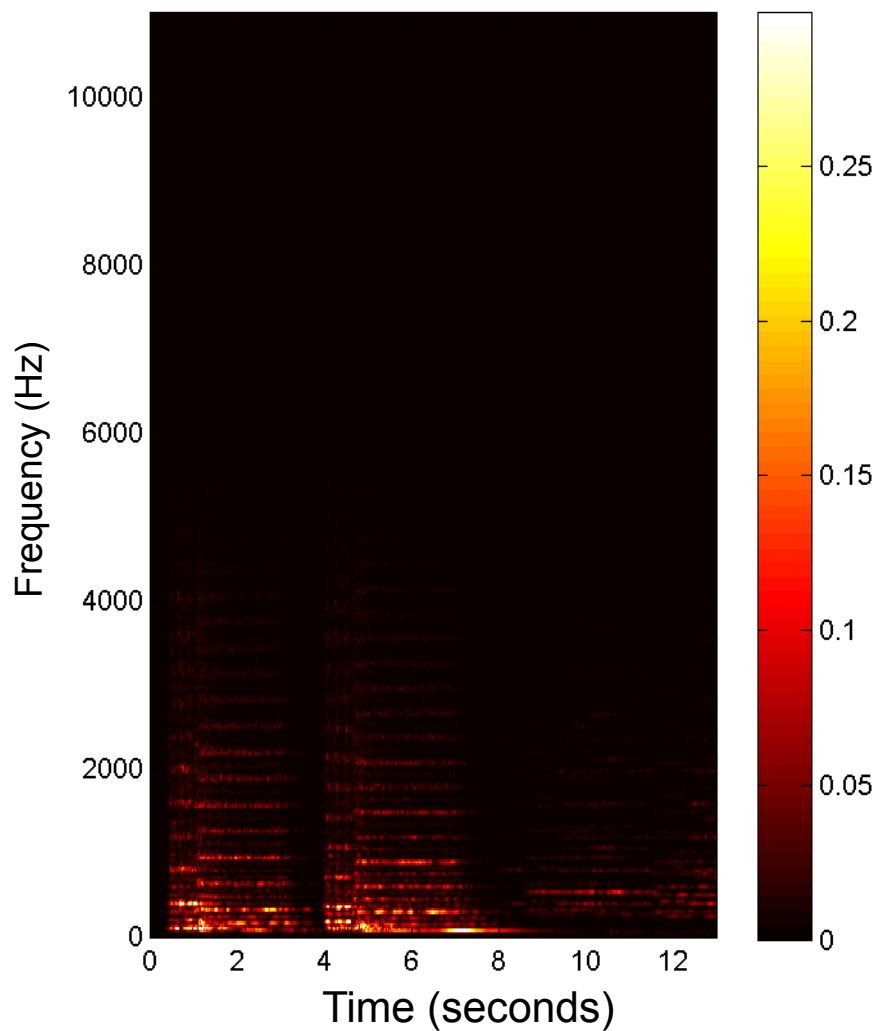


# Beat Tracking

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

# Beat Tracking

Spectrogram

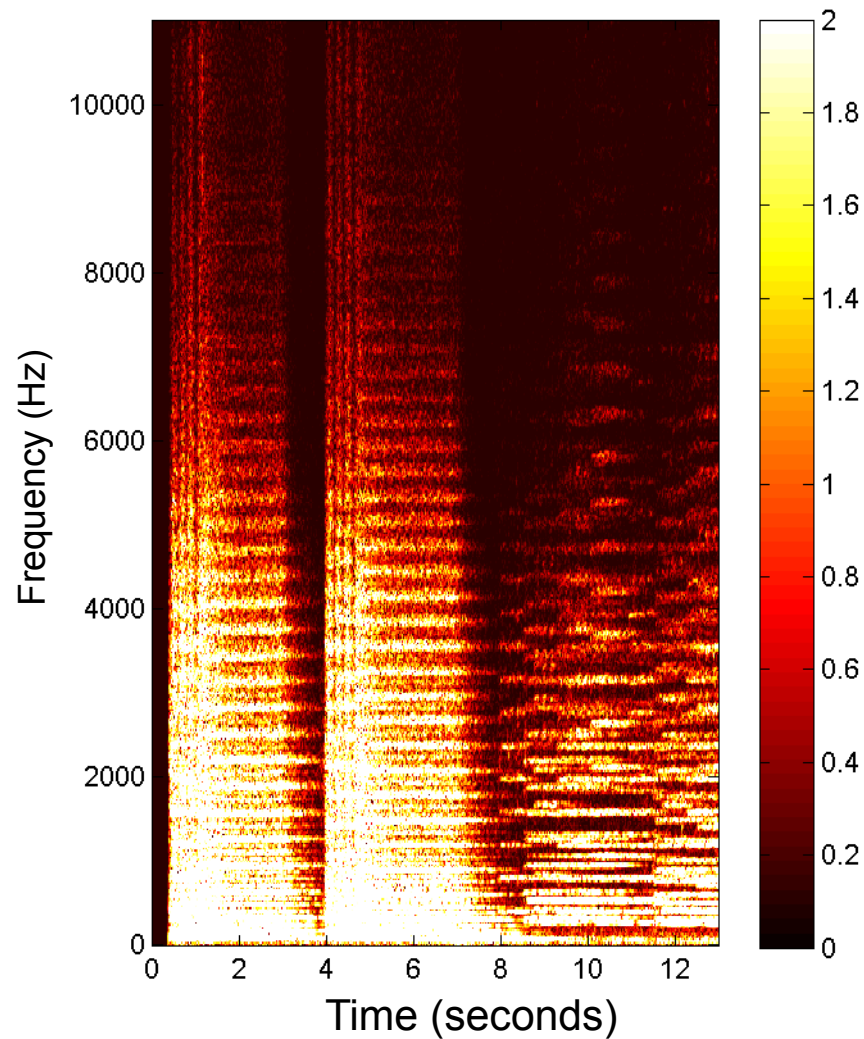


## Steps:

1. Spectrogram

# Beat Tracking

Compressed Spectrogram



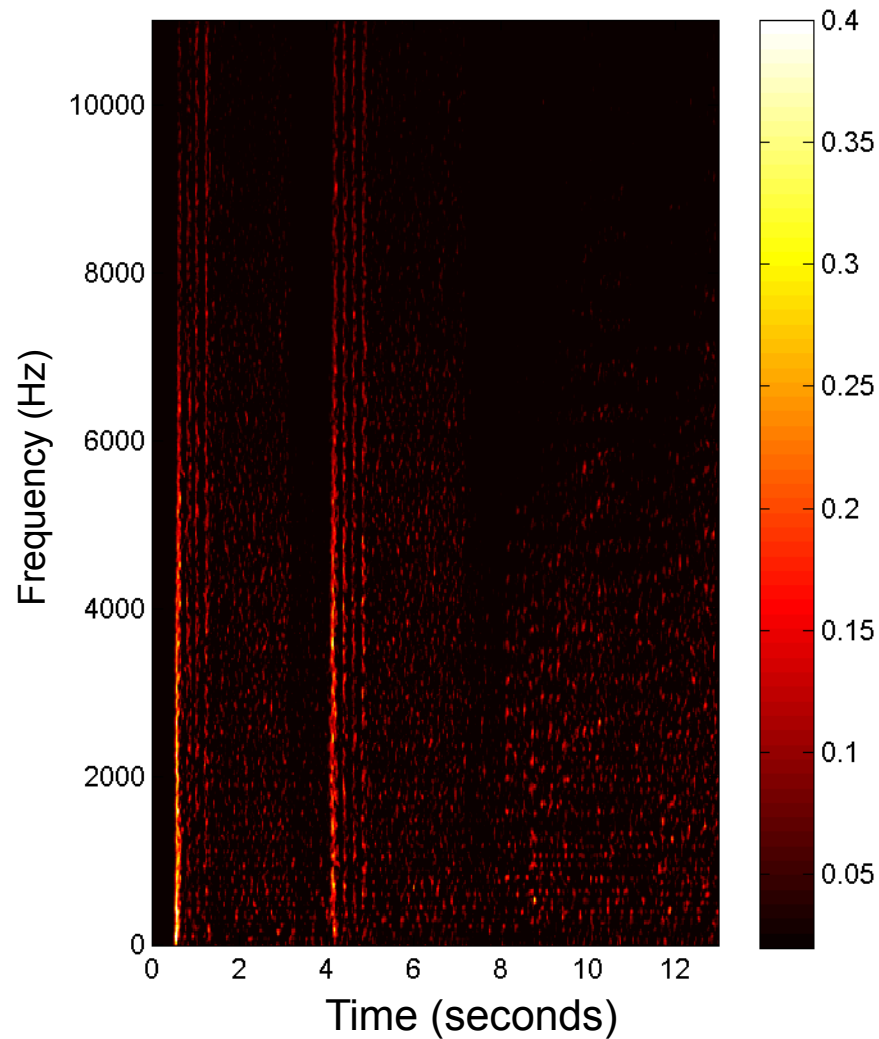
## Steps:

1. Spectrogram
2. Log Compression



# Beat Tracking

Difference Spectrogram



## Steps:

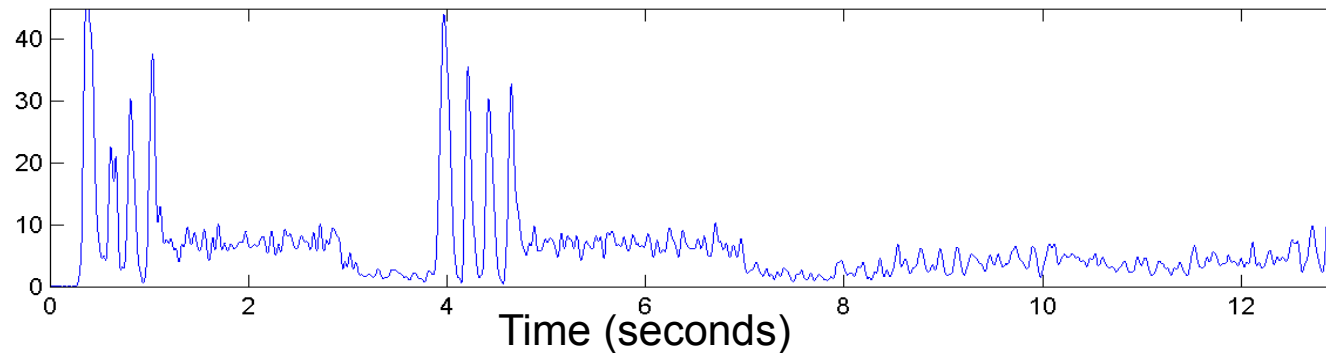
1. Spectrogram
2. Log Compression
3. Differentiation

# Beat Tracking

## Steps:

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

## Novelty Curve

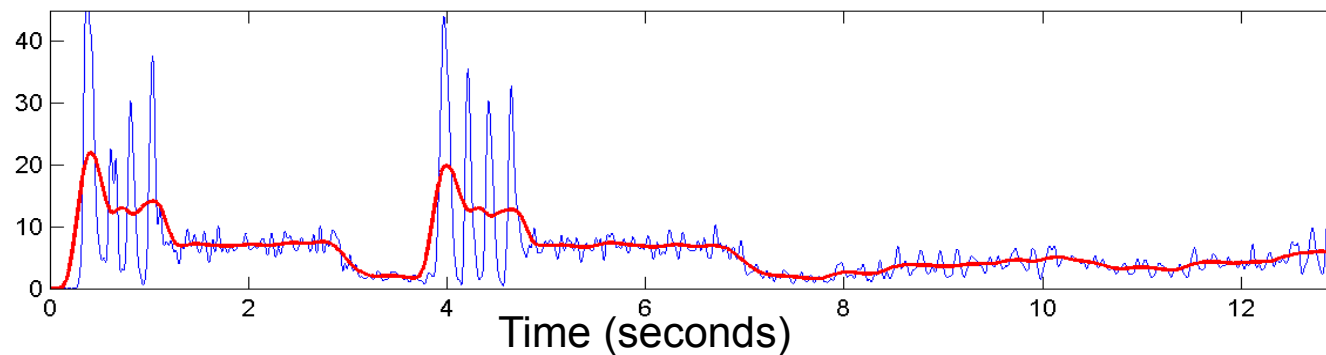


# Beat Tracking

## Steps:

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

## Novelty Curve Local Average

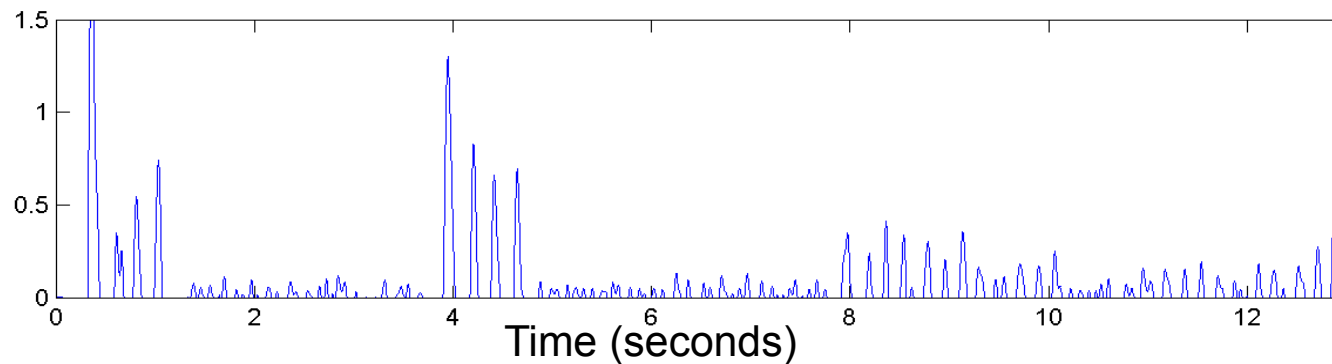


# Beat Tracking

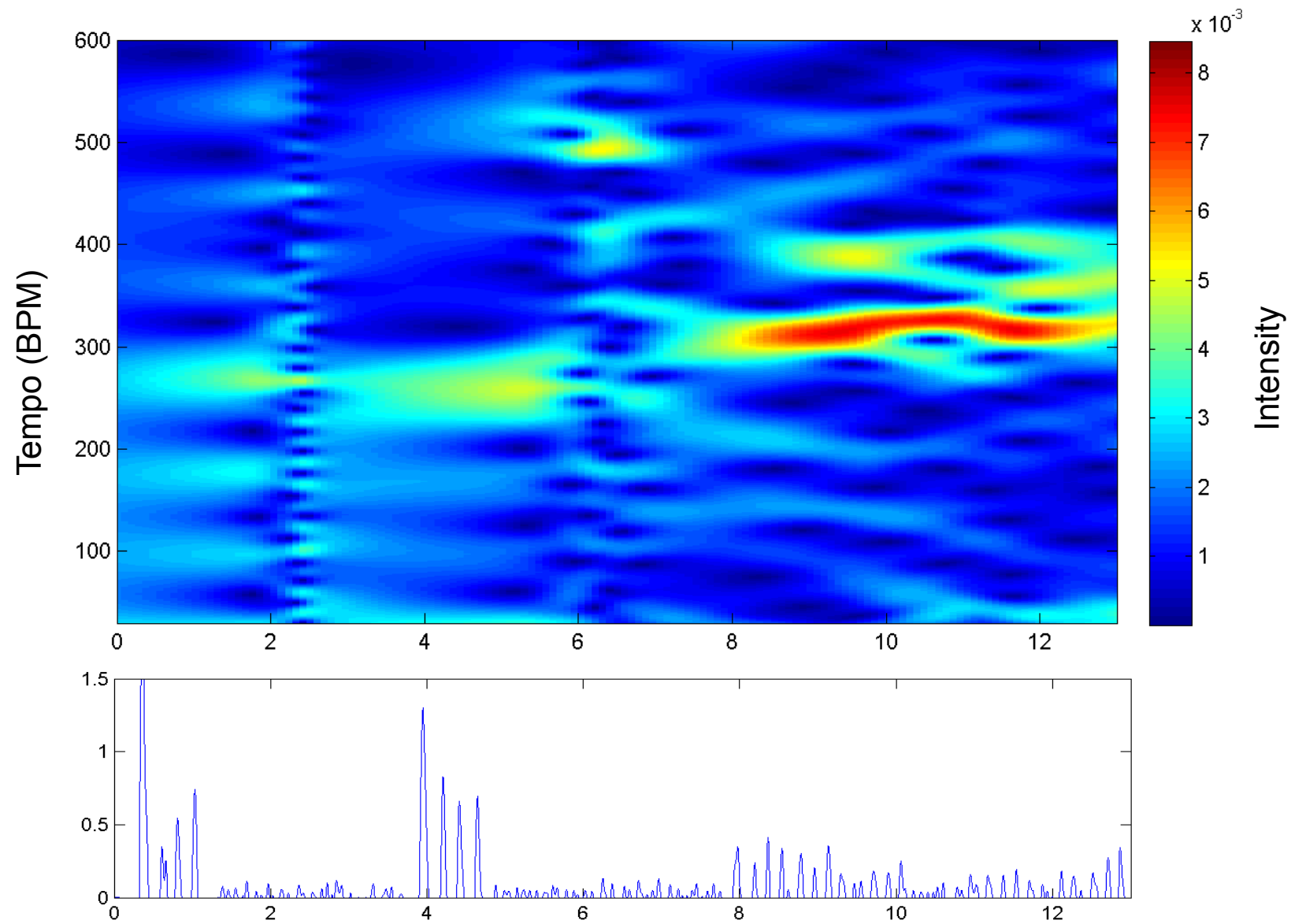
## Steps:

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

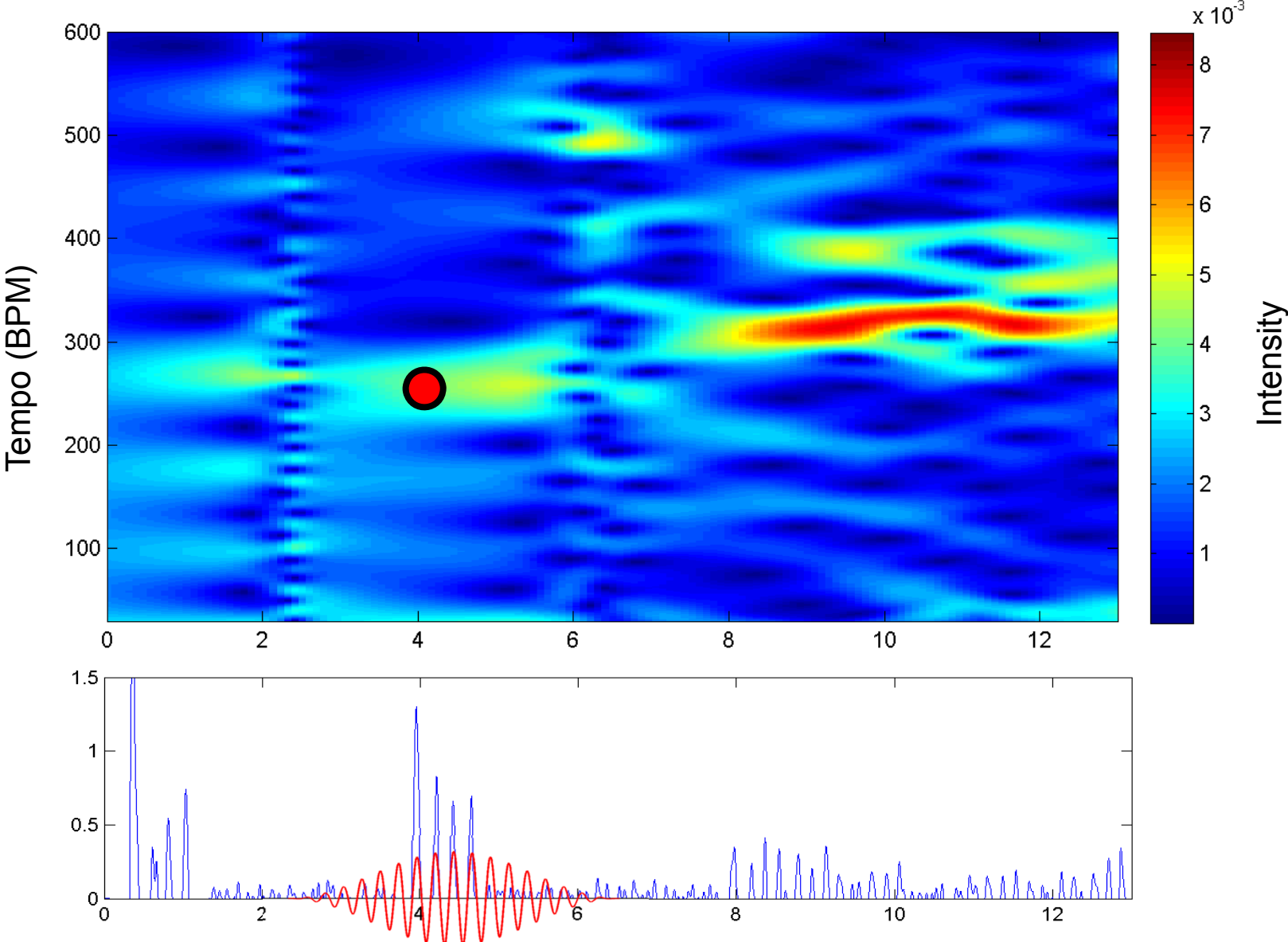
## Novelty Curve



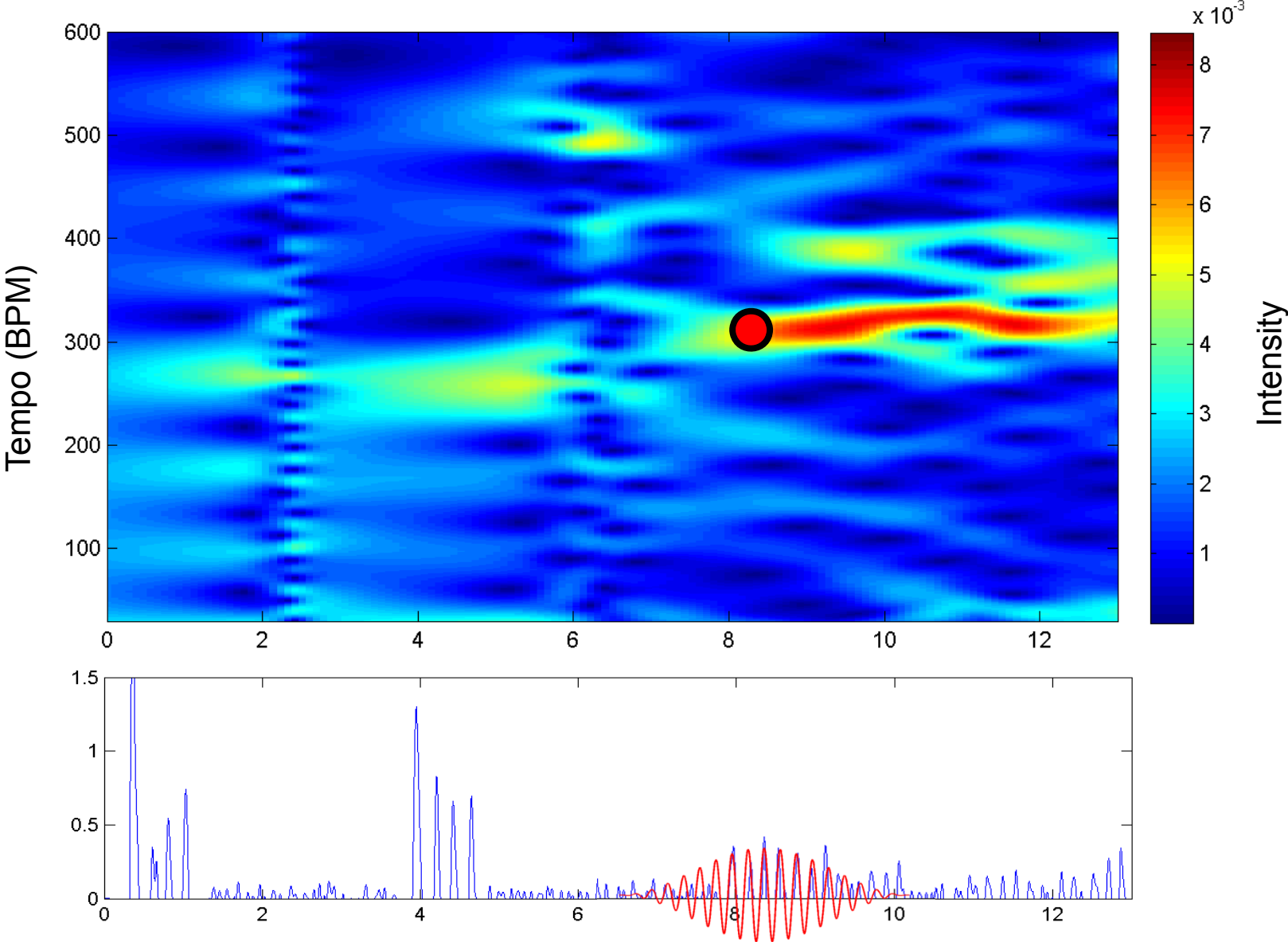
# Beat Tracking



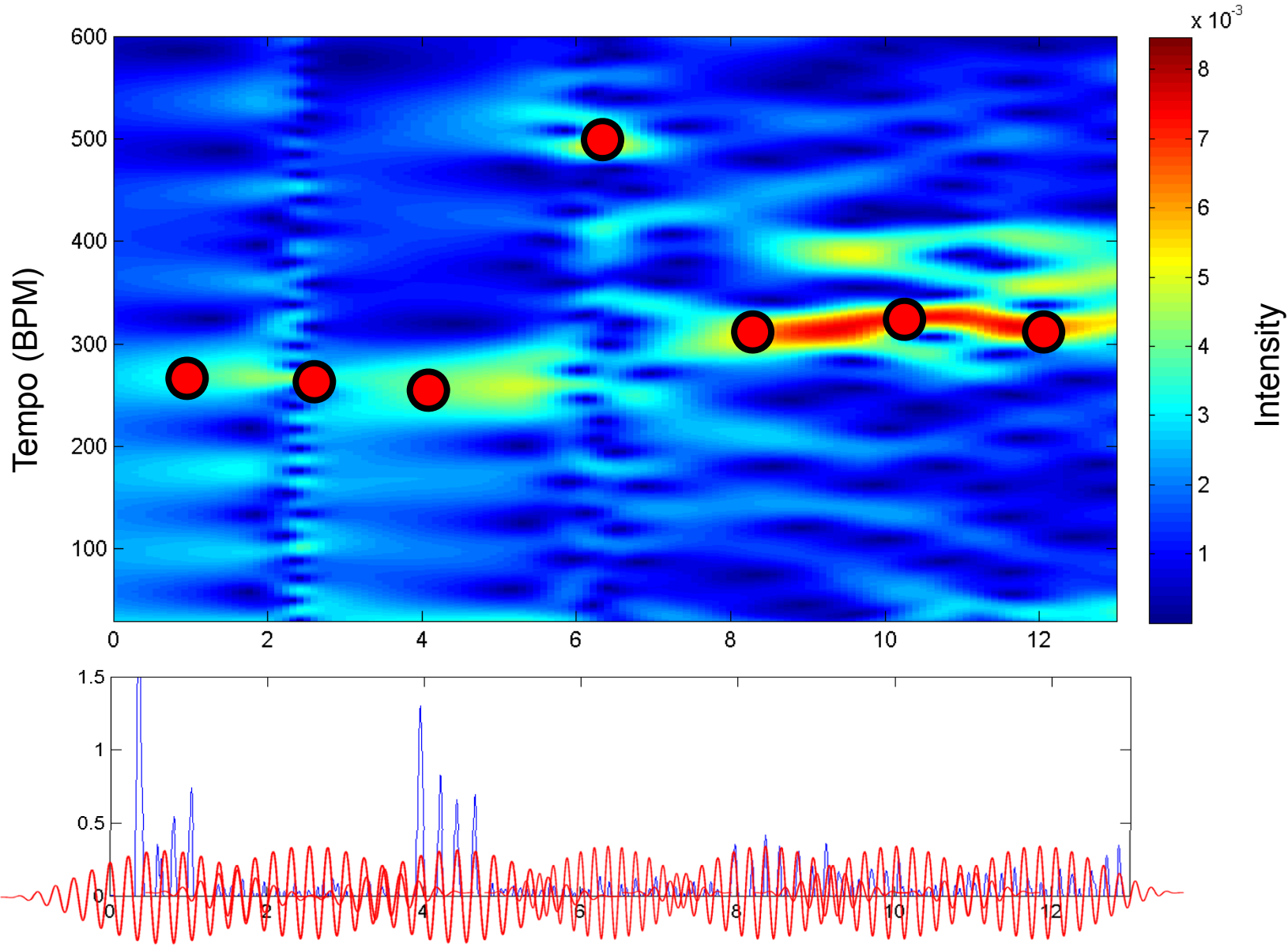
# Beat Tracking



# Beat Tracking

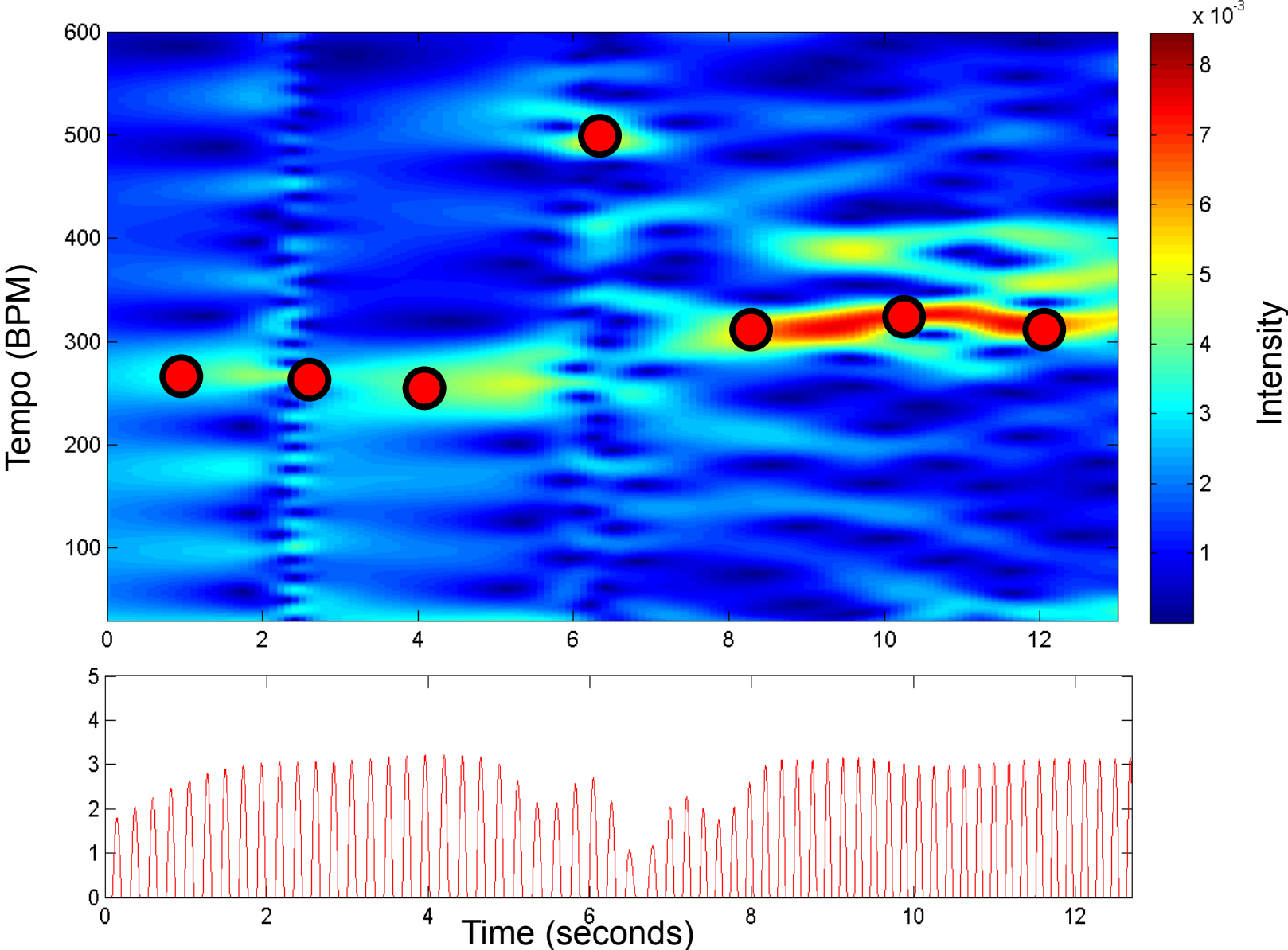


# Beat Tracking

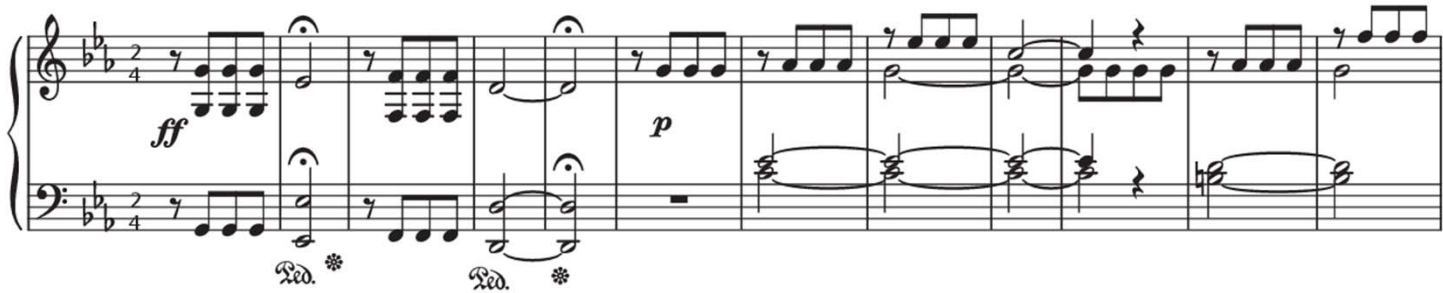




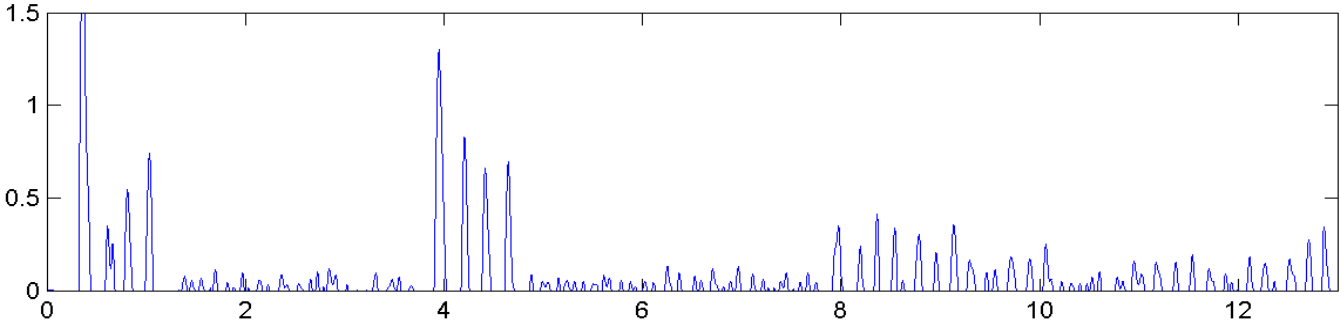
# Beat Tracking



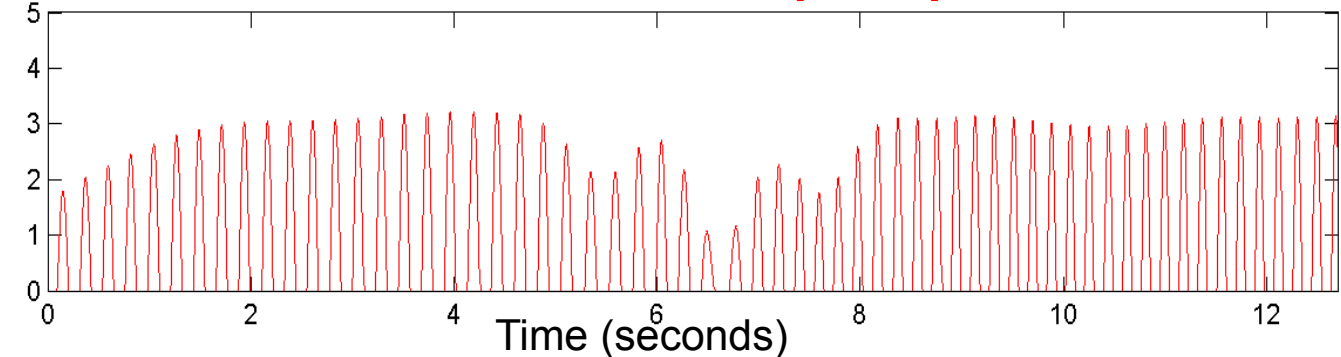
# Beat Tracking



## Novelty Curve



## Predominant Local Pulse (PLP)



# Beat Tracking





# Motivic Similarity

Allegro con brio (♩ = 108)

The image shows a musical score for the first movement of Beethoven's Fifth Symphony, marked 'Allegro con brio' with a tempo of 108 quarter notes per minute. The score is in 2/4 time and features a piano introduction with a forte (ff) dynamic. The notation includes a treble and bass clef, a key signature of two flats, and a 'Ped.' marking with an asterisk indicating a pedal point.

Beethoven's Fifth (1st Mov.)



Beethoven's Fifth (3rd Mov.)



# Motivic Similarity

Allegro con brio (♩ = 108)

The image shows a musical score for the first movement of Beethoven's Fifth Symphony. It is in 2/4 time, marked 'Allegro con brio' with a tempo of 108 quarter notes per minute. The score is for piano, starting with a fortissimo (ff) dynamic. The music features a prominent rhythmic motif of three eighth notes followed by a quarter note, which is a key element of the symphony's 'fate' motif. The score includes a piano introduction with a 'Ped.' (pedal) marking and a '\*' symbol under the first and third measures of the introduction.

Beethoven's Fifth (1st Mov.)



Beethoven's Fifth (3rd Mov.)



Beethoven's Appassionata



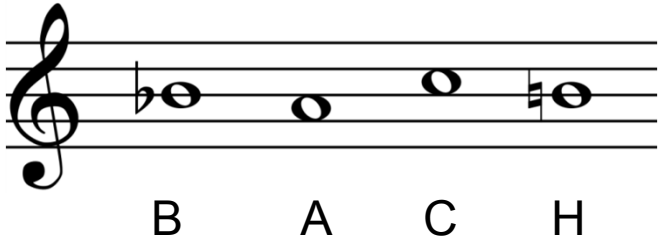
# Motivic Similarity

Var. 4: Vivace

The image displays a musical score for a piece titled "Var. 4: Vivace". The score is written in bass clef with a key signature of two flats (B-flat and E-flat) and a common time signature (C). The music is divided into four staves. The first staff begins with a dynamic marking of *f* (forte) and contains a series of rhythmic patterns, including eighth and sixteenth notes. The second and third staves continue these patterns with various rhythmic values and some slurs. The fourth staff features a more complex rhythmic structure with dynamic markings alternating between *p* (piano) and *f* (forte). The piece concludes with a double bar line.



# Motivic Similarity



A musical staff in treble clef showing a four-note motif: B (B-flat), A, C, and H (B-natural). The notes are spaced out across the staff.

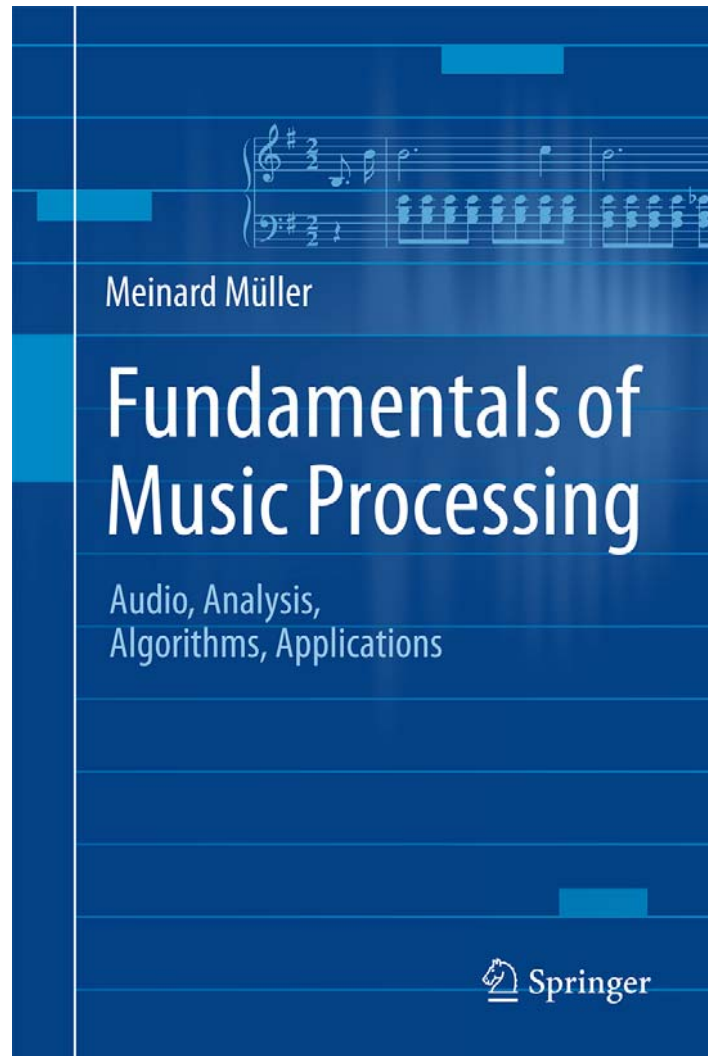


A musical score for Soprano (S), Alto (A), Tenor (T), and Bass (B) in 4/4 time. The lyrics are: "auf - - ge - rafft, und nie - mand ach - - tet und nie - mand ach - - tet drauf". The Alto part has a red box highlighting the notes b, a, c, h, which correspond to the motif in the first image.





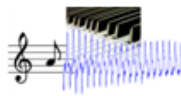

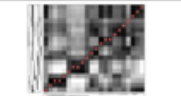
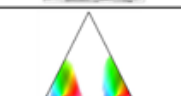

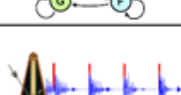


# Book: Fundamentals of Music Processing



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](http://www.music-processing.de)

# Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Representations
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](http://www.music-processing.de)