



ISMIR
2017, SUZHOU, CHINA

Tutorial T3
A Basic Introduction to Audio-Related
Music Information Retrieval

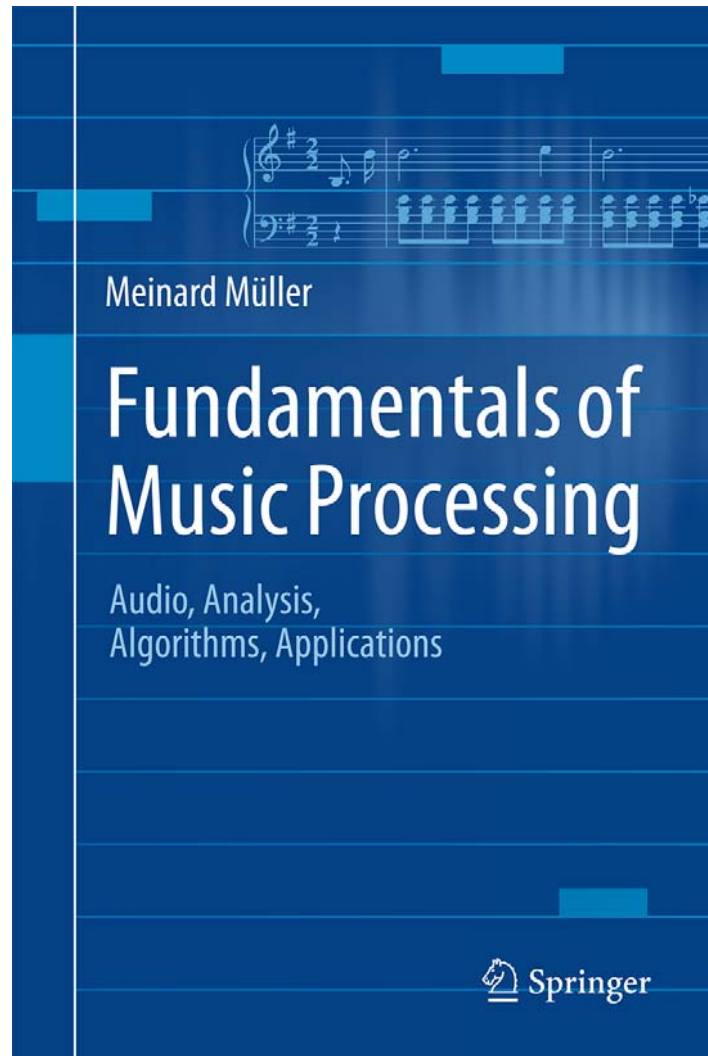
Audio Structure Analysis

Meinard Müller, Christof Weiß

International Audio Laboratories Erlangen

meinard.mueller@audiolabs-erlangen.de, christof.weiss@audiolabs-erlangen.de

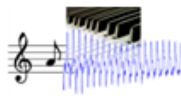

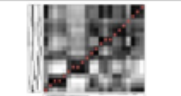


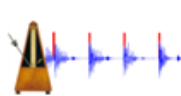
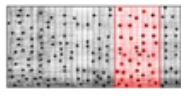
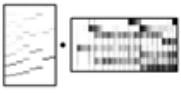
Book: Fundamentals of Music Processing



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 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

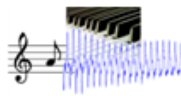

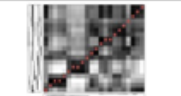


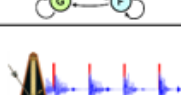


ISBN: 978-3-319-21944-8

Springer, 2015

Accompanying website:

www.music-processing.de

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

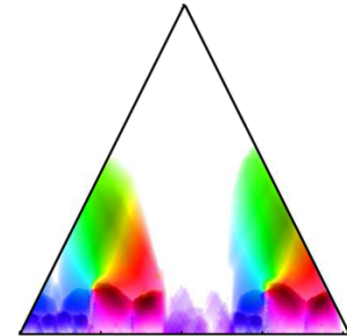
Springer, 2015

Accompanying website:

www.music-processing.de

Chapter 4: Music Structure Analysis

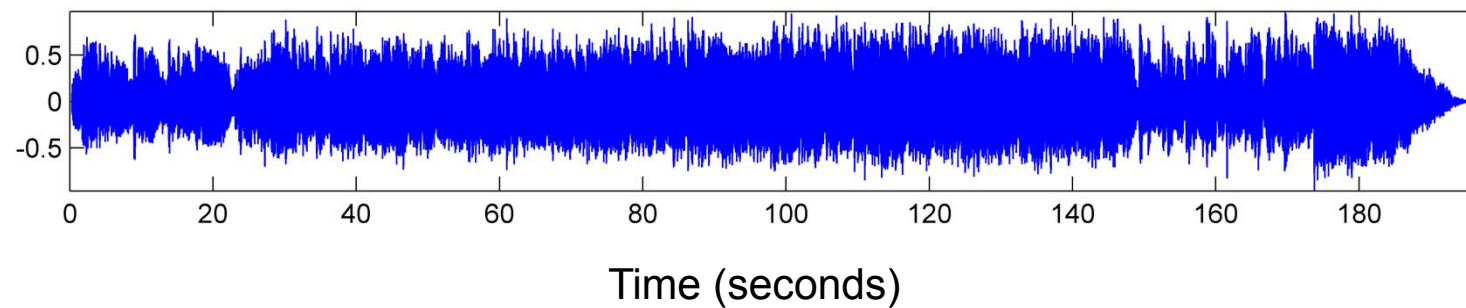
- 4.1 General Principles
- 4.2 Self-Similarity Matrices
- 4.3 Audio Thumbnailing
- 4.4 Novelty-Based Segmentation
- 4.5 Evaluation
- 4.6 Further Notes



In Chapter 4, we address a central and well-researched area within MIR known as music structure analysis. Given a music recording, the objective is to identify important structural elements and to temporally segment the recording according to these elements. Within this scenario, we discuss fundamental segmentation principles based on repetitions, homogeneity, and novelty—principles that also apply to other types of multimedia beyond music. As an important technical tool, we study in detail the concept of self-similarity matrices and discuss their structural properties. Finally, we briefly touch the topic of evaluation, introducing the notions of precision, recall, and F-measure.

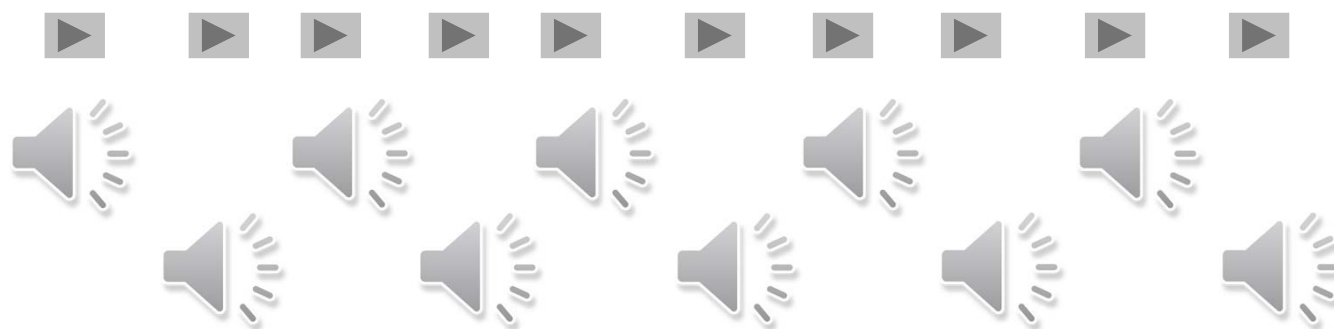
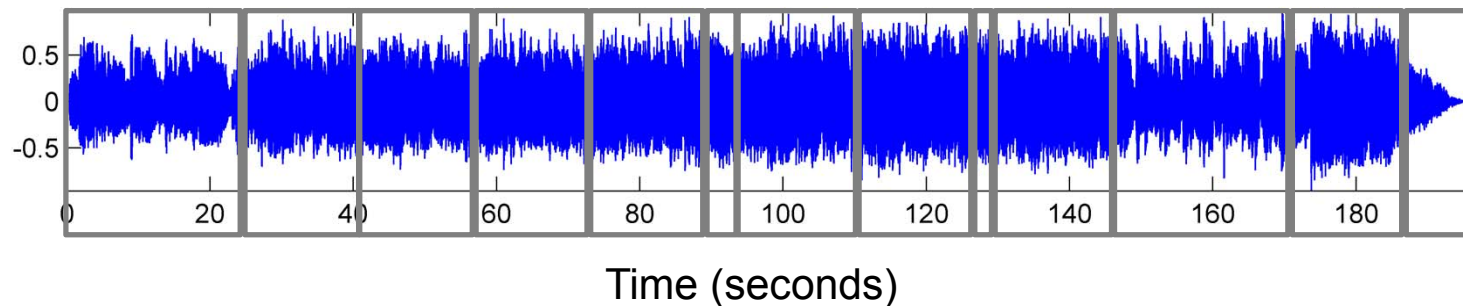
Music Structure Analysis

Example: Zager & Evans “In The Year 2525”



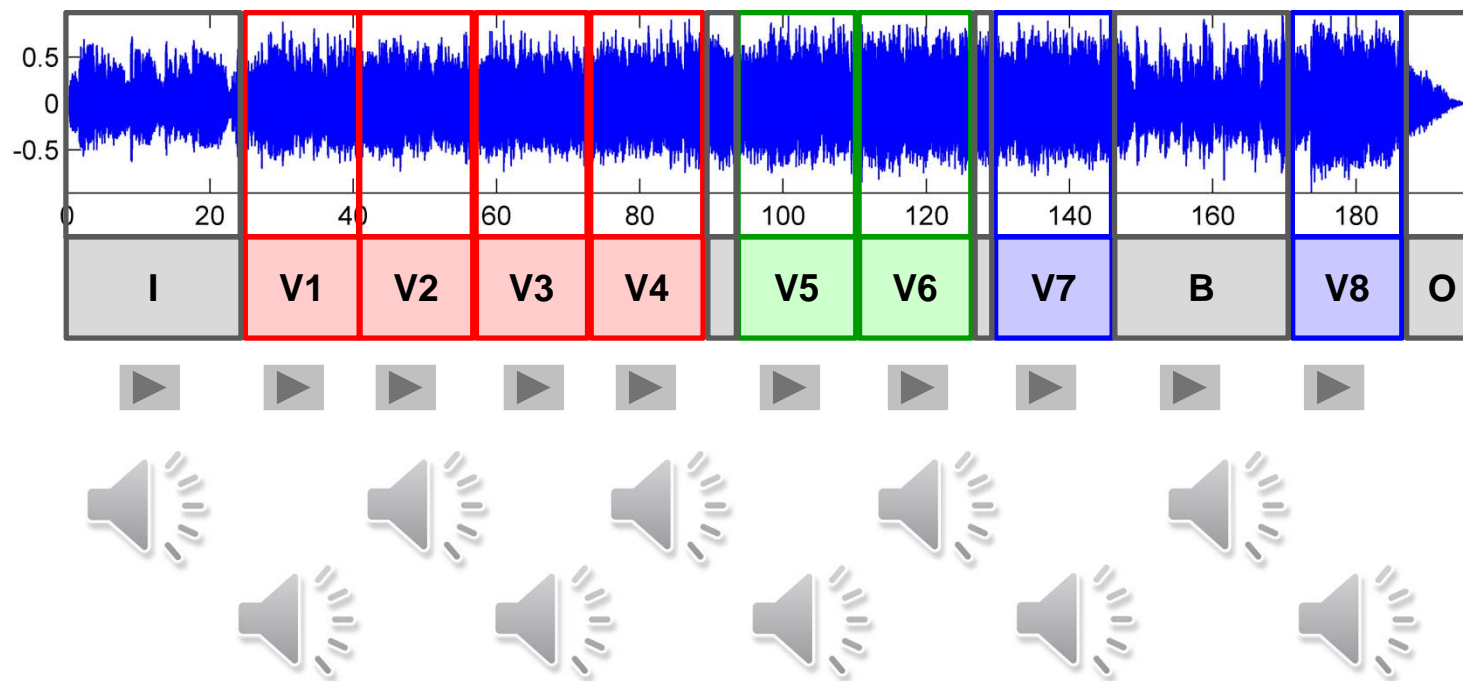
Music Structure Analysis

Example: Zager & Evans “In The Year 2525”



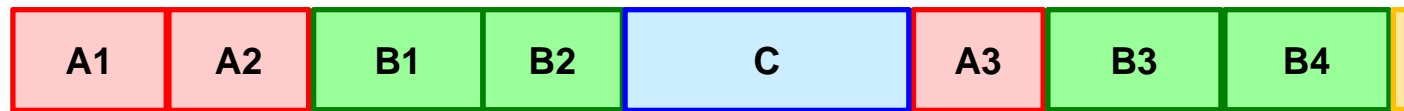
Music Structure Analysis

Example: Zager & Evans “In The Year 2525”

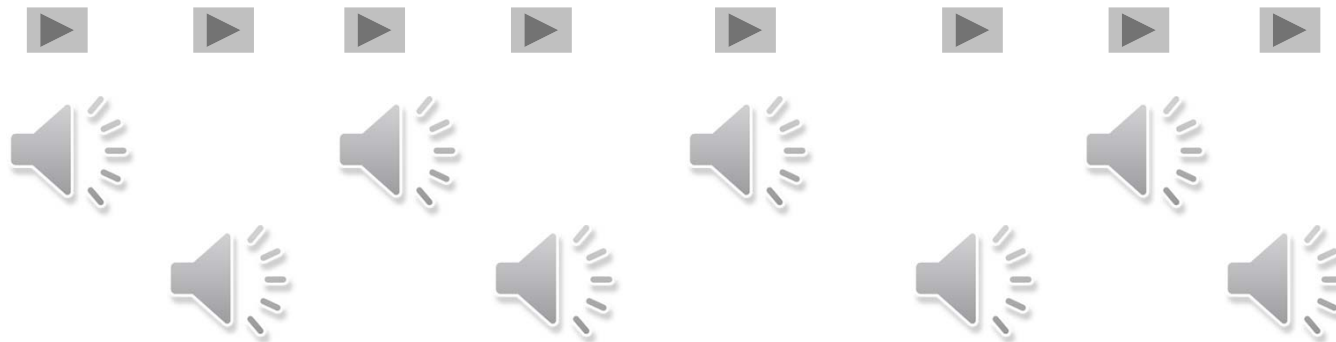


Music Structure Analysis

Example: Brahms Hungarian Dance No. 5 (Ormandy)

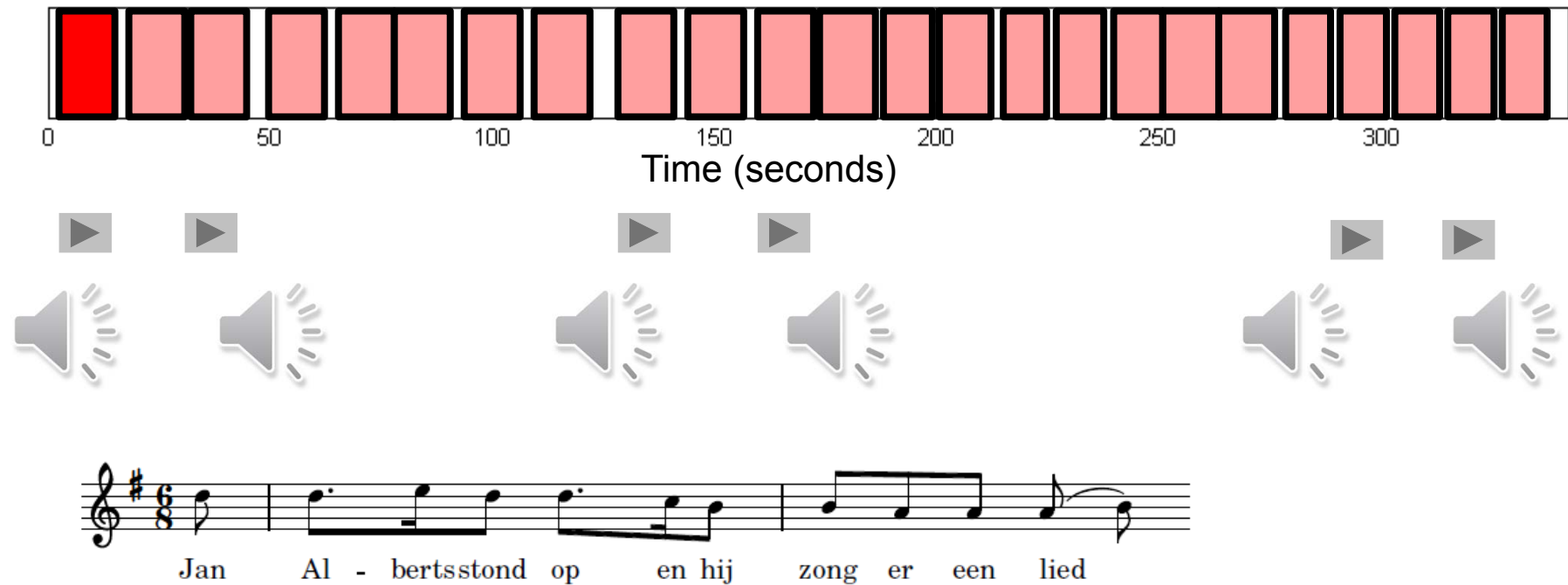


Time (seconds)



Music Structure Analysis

Example: Folk Song Field Recording
(Nederlandse Liederbank)



Music Structure Analysis

Example: Weber, Song (No. 4) from “Der Freischütz”

Introduction

Stanzas

Dialogues

Flauti piccoli.
Oboi.
Fagotti.
Violino I.
Violino II.
Viola.
Caspar.
Violoncello e Basso.

Allegro feroce, ma non troppo presto.

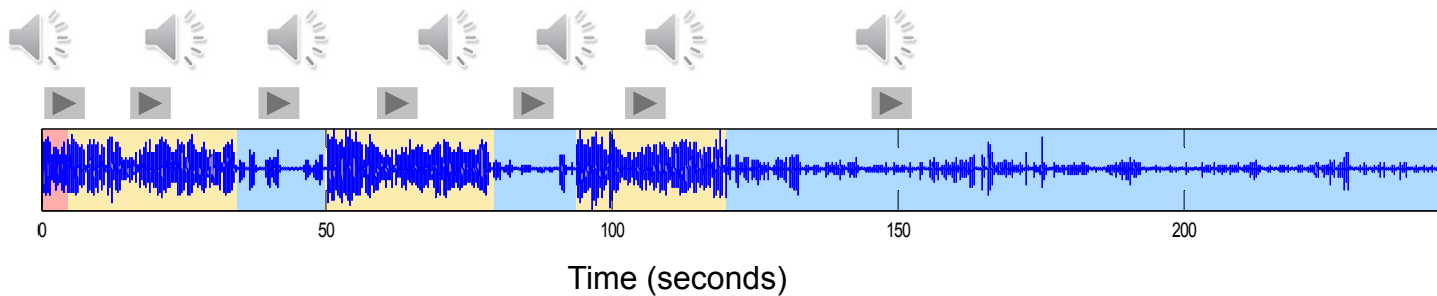
1. Hier im ird'schen Jammer . thal
2. Eins ist Eins und Drei sind Drei!
3. Oh . ne dies Tri . fo . li . um

(Nach der ersten Strophe wird gesprochen.)
Caspar. Ei, da mußt auch mit singen. (Trinke!)
Max. Lass mich!
Caspar. Jang'ler Agathe soll leben! Wer die Gesundheit seiner Braut ausschlägt, wir' doch wahrlich ein Schuft!
Max. Ja wirt' unverschäm't, die einen an und trinken.

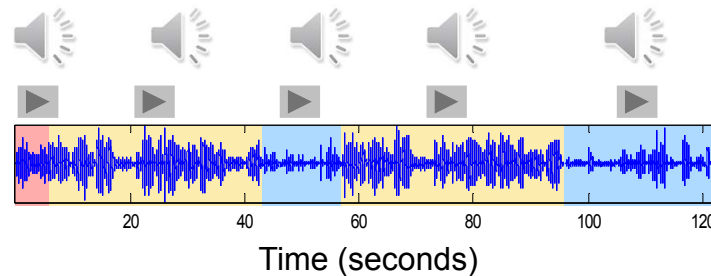
(Nach der zweiten Strophe.)
Caspar. Mit dir ist aber auch gar nichts anzufangen. (Trinke!)
Max. Wie kannst du mir summen, in so etwas einzustimmen.
Caspar. Unser Herr Fürst soll leben! Wer nicht' dabei ist, wir' ein Zufall!
Max. Nun denn, aber dann auch keine Tropfen mehr. Wo stehen sie mit trinken. Max weilt sich mit dem Hefe Laß, so und gibt nicht an erkennen, das ihn heile sei!

(Nach der dritten Strophe.)
Max. (ausbrüllend.) Bube! Agathe hat Recht, wenn sie mich immer vor dir warnen will fort. Du bist besessen!
Caspar. Wie kannst du auch gleich so in Harnisch geraten. Bräuerher! Ich diene noch als Bote in der letzten Foh. de. Untern Kriegevolk lernt man solche Schmeiseldien. Ob es dich nicht für. Max steht auf. Willst du schon nach Hause?
Max. Ja, es wirt' Zeit. Du schlag Stöhen.
Caspar. Zu Agathe? Das rath' ich doch nicht. Du könntest sie erschrecken. Weiss du nicht, dass sie auf einem Gewiss als gute Vorbedeutung für morgen hofft?
Max. Ach, die Armut und ich selbst! Morgen!
Max. Was machst du, wir' mir doch ganz schauerlich. Was hast du geladen? Was war das für eine Kugel?
Caspar. Das keine Kugel. Nur oben. Eine rechtliche Blindschleibe, die trifft allemal.
Max. (zitternd) Ich denk', sehr klein ich heranzieh'! So etwas ist mir begegnet. Caspar! Ich bitte dich, ich beschwöre dich, laßst ihn Caspar, ich bring' dich um! Sag, was war das für eine Kugel?
Caspar. Hast du verwirrt vor Freuden? Ich theile sie mit dir. Caspar! Das war eine Schale. Laß' mich los!
Max. (läst ihn los. Wo hast du die Kugel her?)

Kleiber



Ackermann



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.

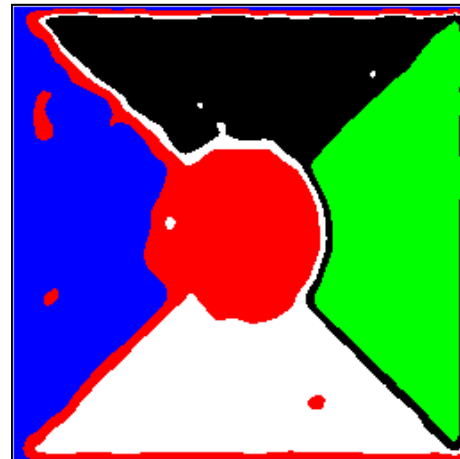
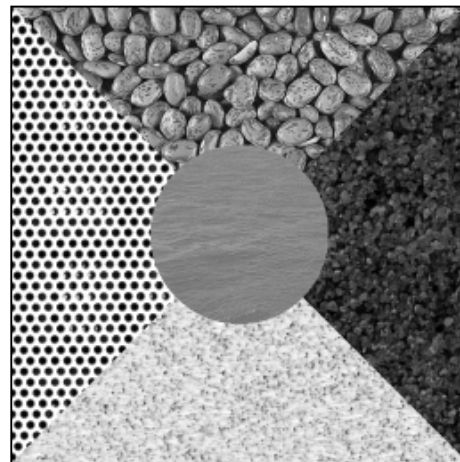
- **Homogeneity:** Consistency in tempo, instrumentation, key, ...
- **Novelty:** Sudden changes, surprising elements ...
- **Repetition:** Repeating themes, motives, rhythmic patterns,...

Music Structure Analysis

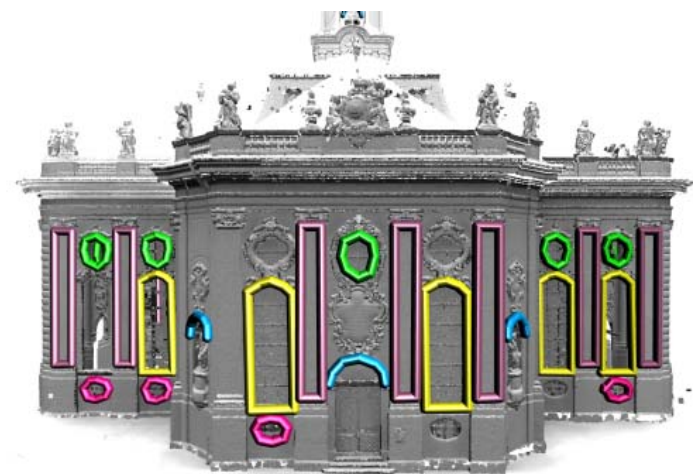
Novelty



Homogeneity



Repetition



Overview

- Introduction
- Feature Representations
- Self-Similarity Matrices
- Novelty-Based Segmentation

Thanks:

- Clausen, Ewert, Kurth, Grohgan, ...
- Dannenberg, Goto
- Grosche, Jiang
- Paulus, Klapuri
- Peeters, Kaiser, ...
- Serra, Gómez, ...
- Smith, Fujinaga, ...
- Wiering, ...
- Wand, Sunkel, Jansen
- ...

Overview

- Introduction
- **Feature Representations**
- Self-Similarity Matrices
- Novelty-Based Segmentation

Thanks:

- Clausen, Ewert, Kurth, Grohgan, ...
- Dannenberg, Goto
- Grosche, Jiang
- Paulus, Klapuri
- Peeters, Kaiser, ...
- Serra, Gómez, ...
- Smith, Fujinaga, ...
- Wiering, ...
- Wand, Sunkel, Jansen
- ...

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

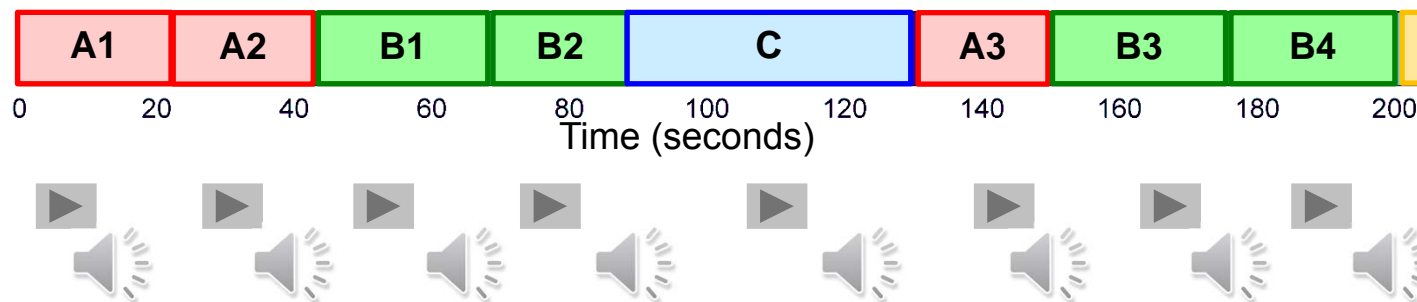
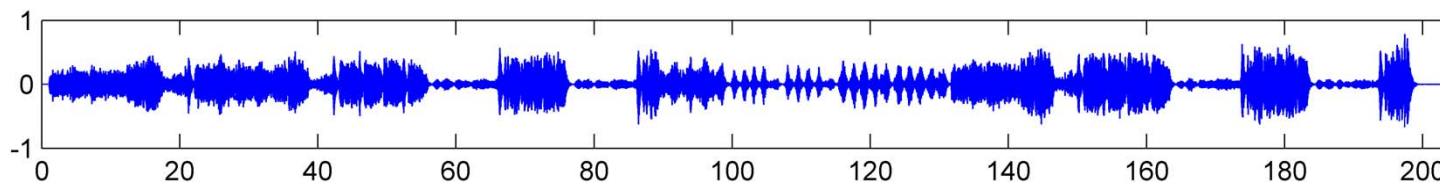
Feature Representation

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**

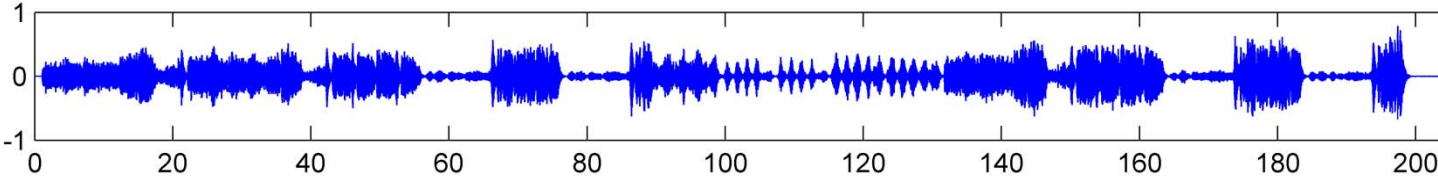
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



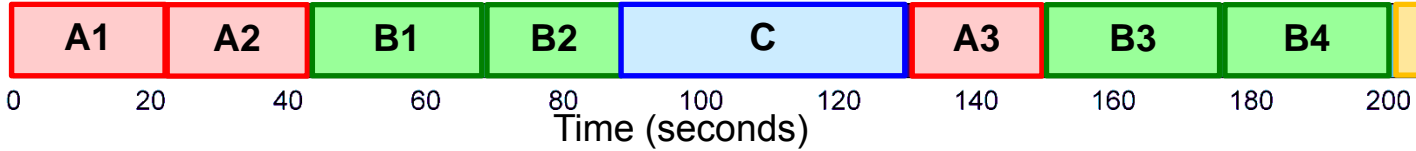
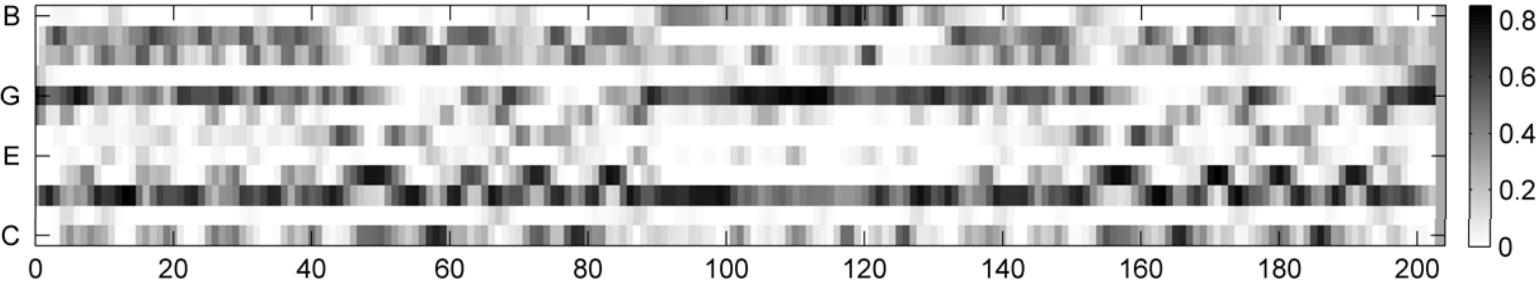
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



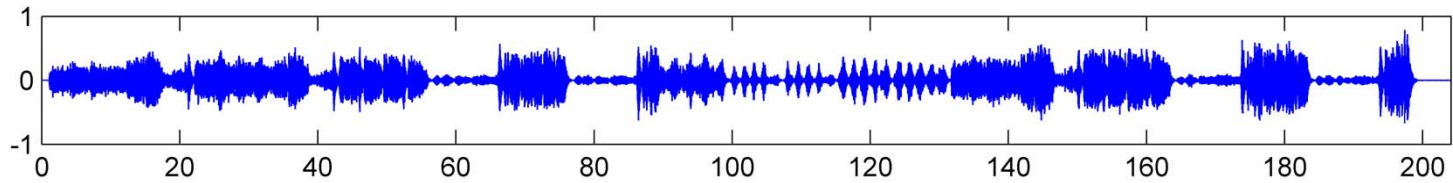
Feature extraction

Chroma (Harmony)

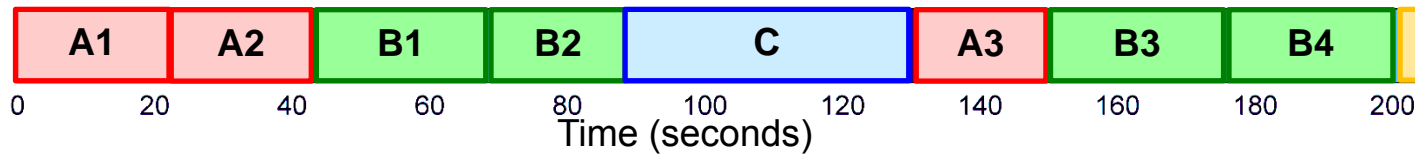
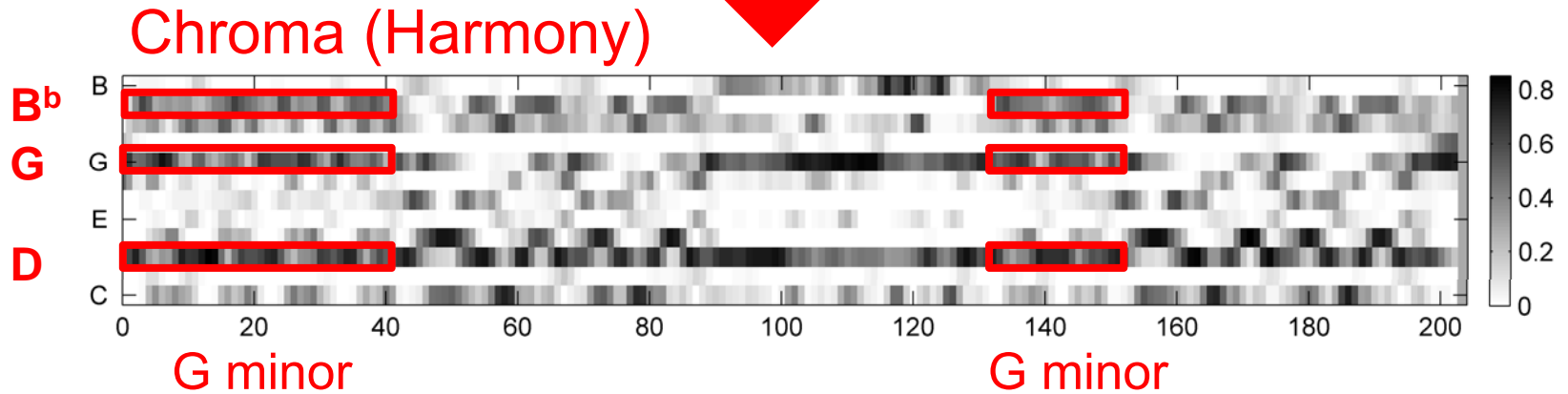


Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)

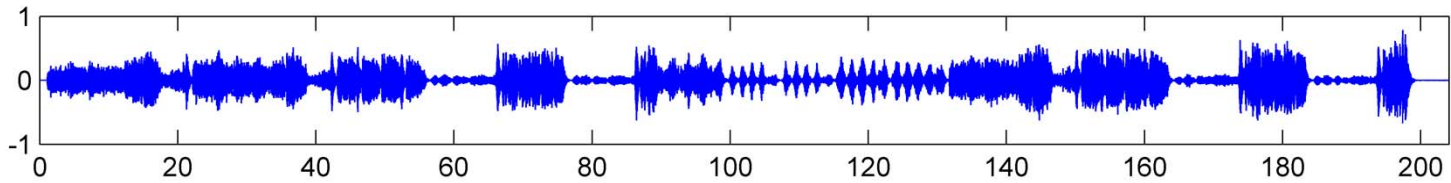


Feature extraction

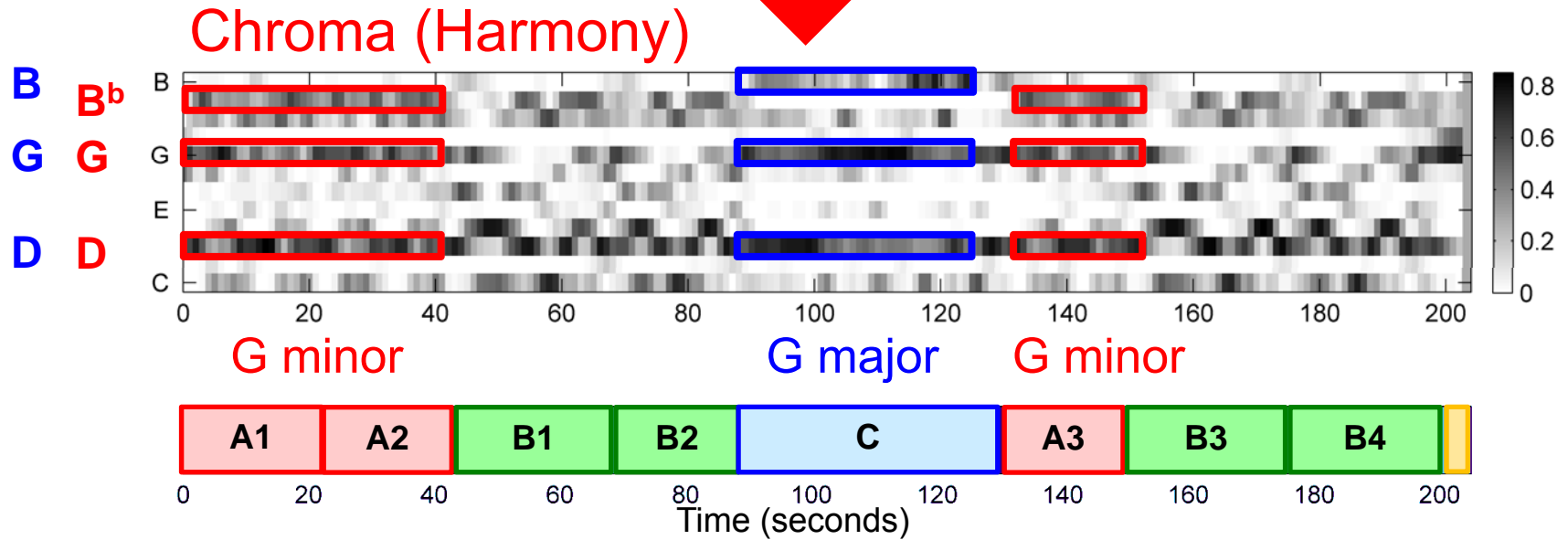


Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Feature extraction



Overview

- Introduction
- Feature Representations
- **Self-Similarity Matrices**
- Novelty-Based Segmentation

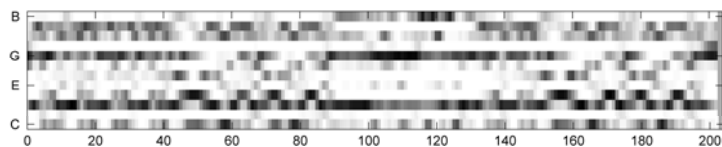
Self-Similarity Matrix (SSM)

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

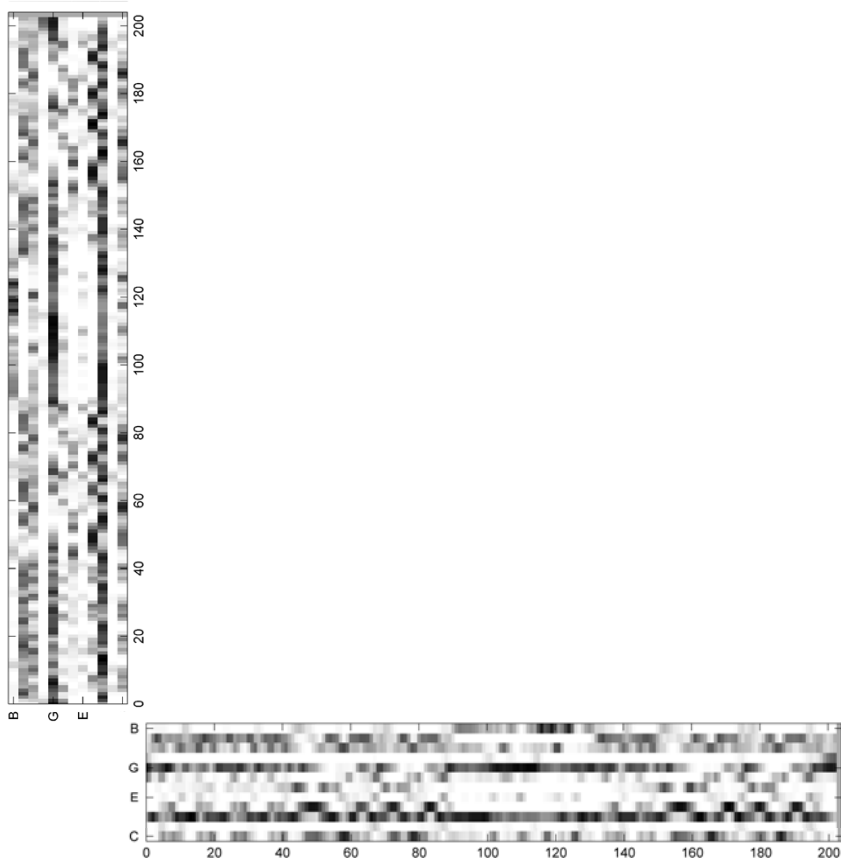
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



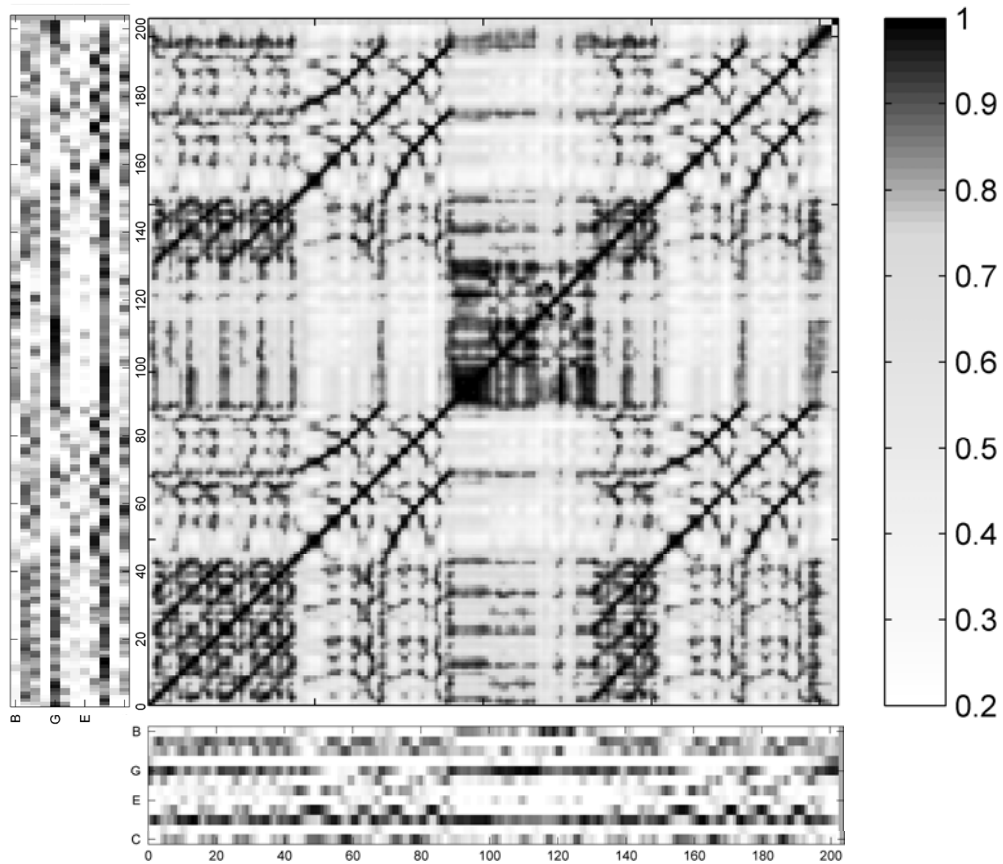
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



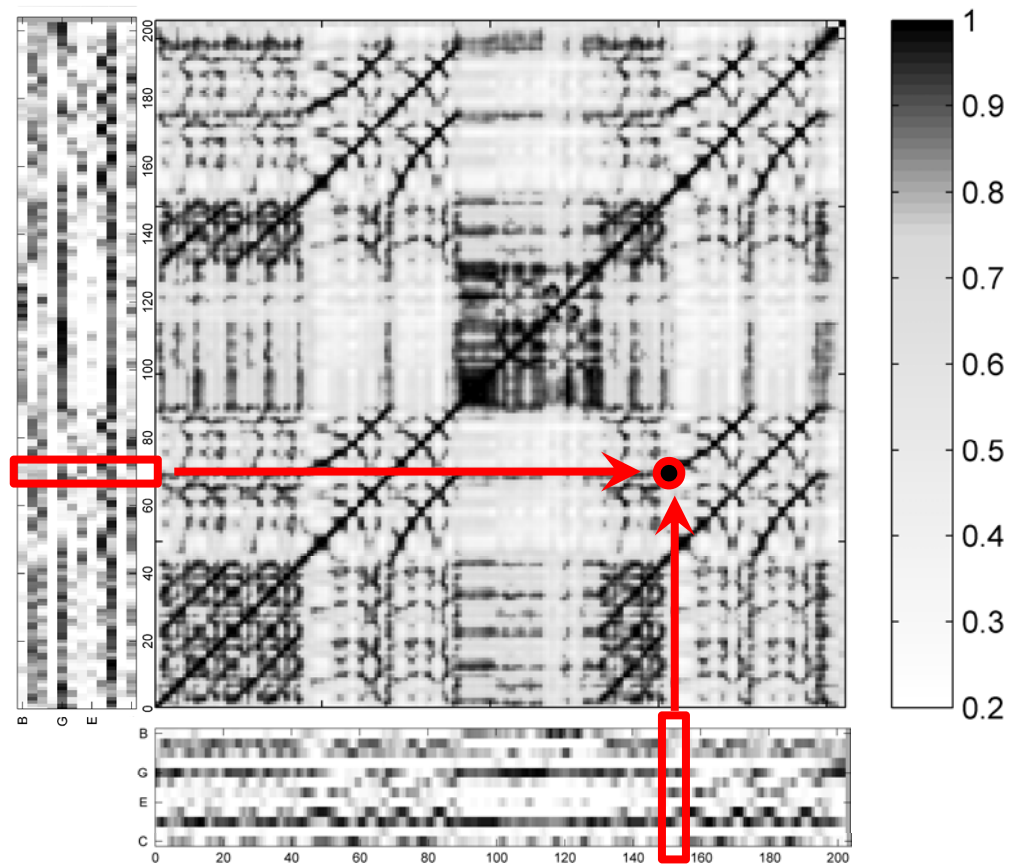
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



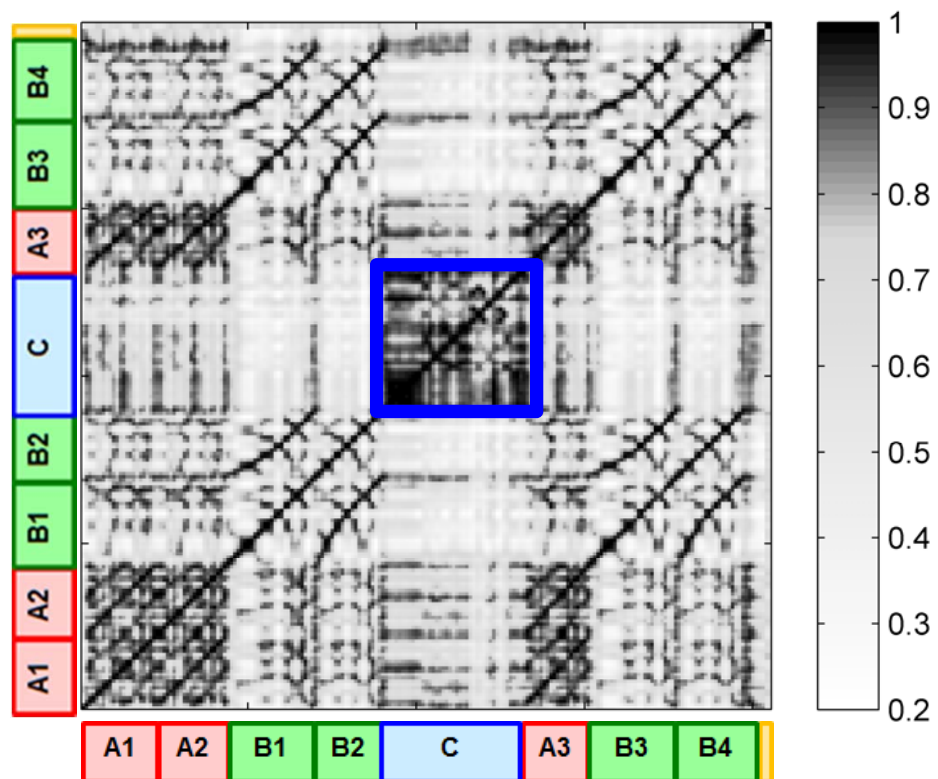
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



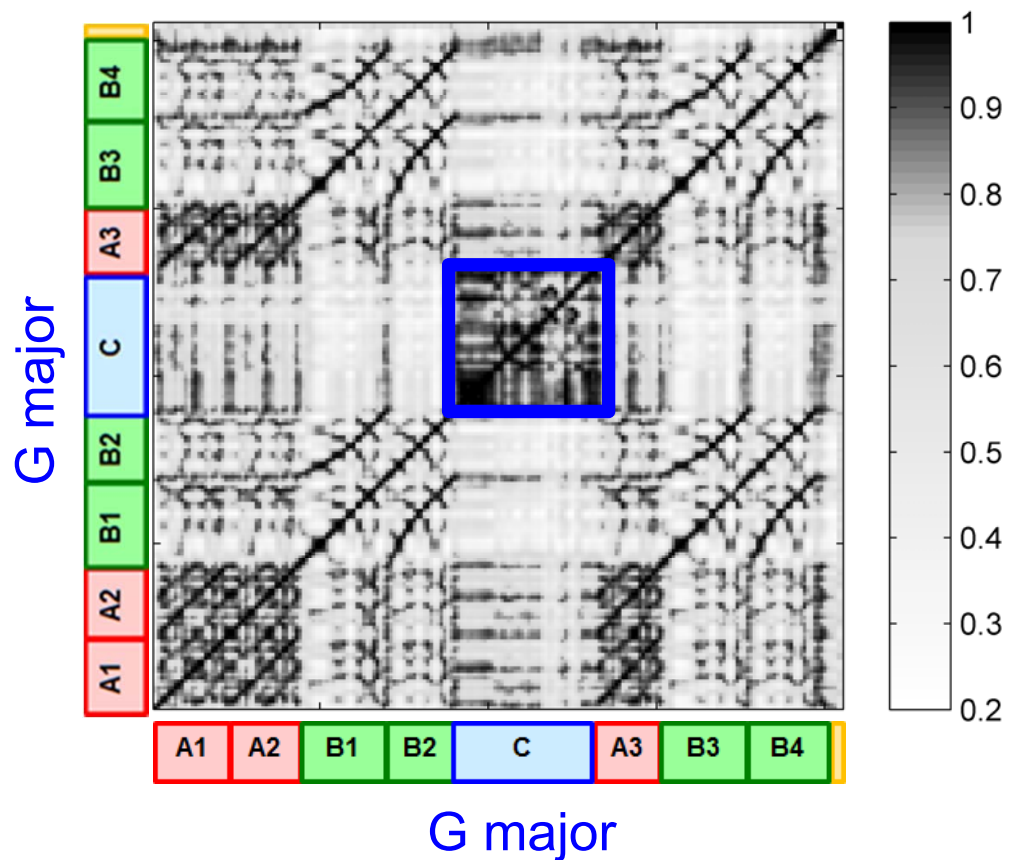
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



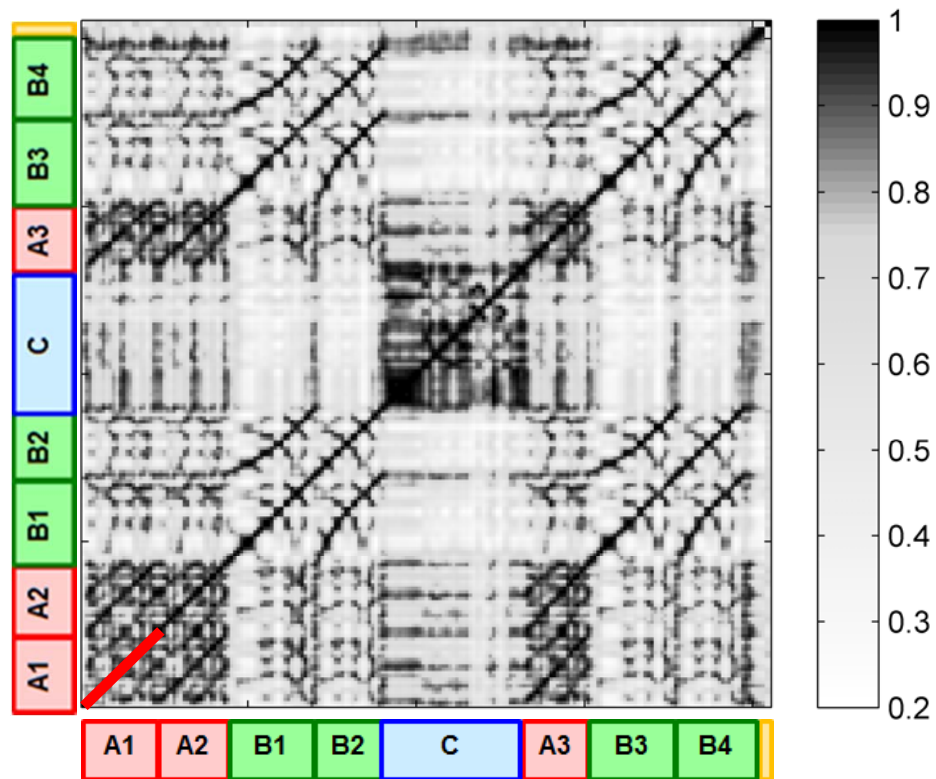
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



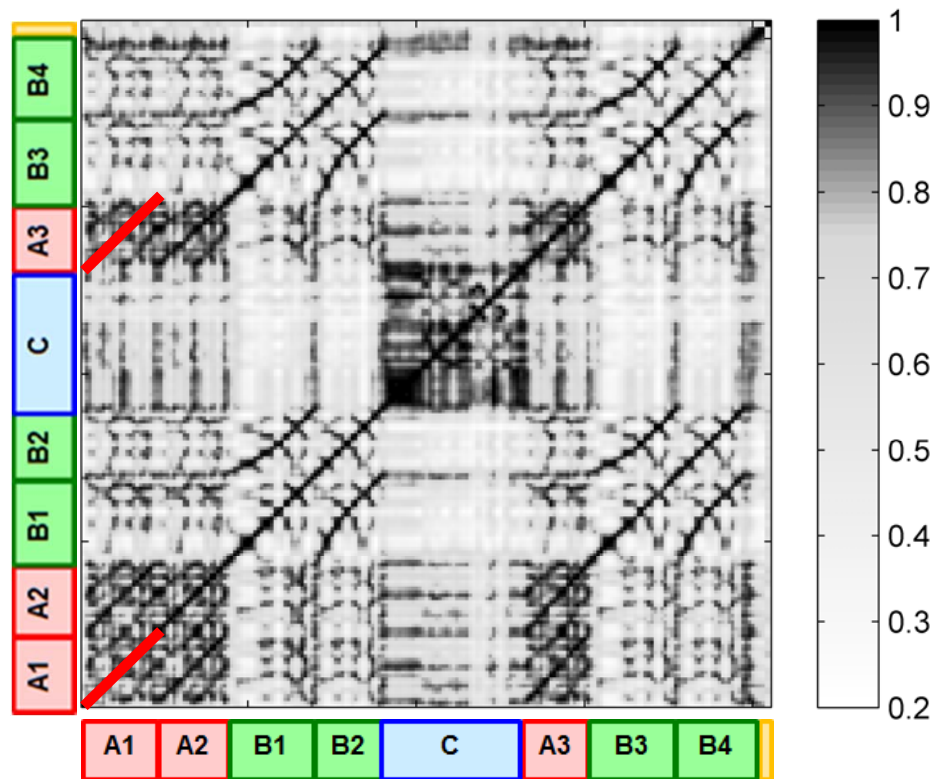
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



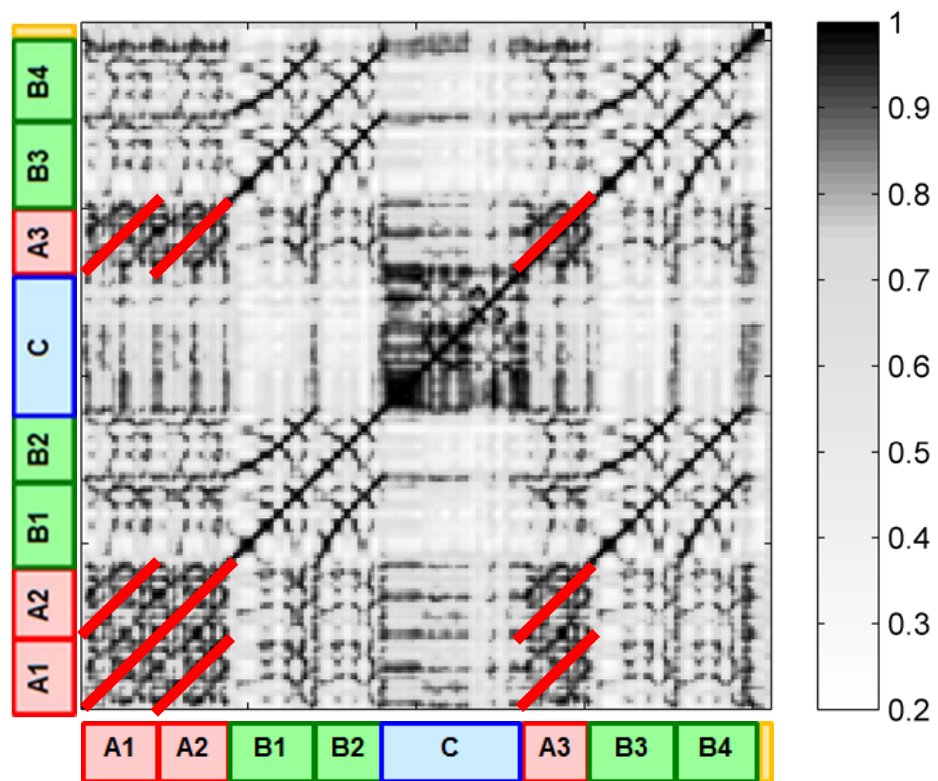
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



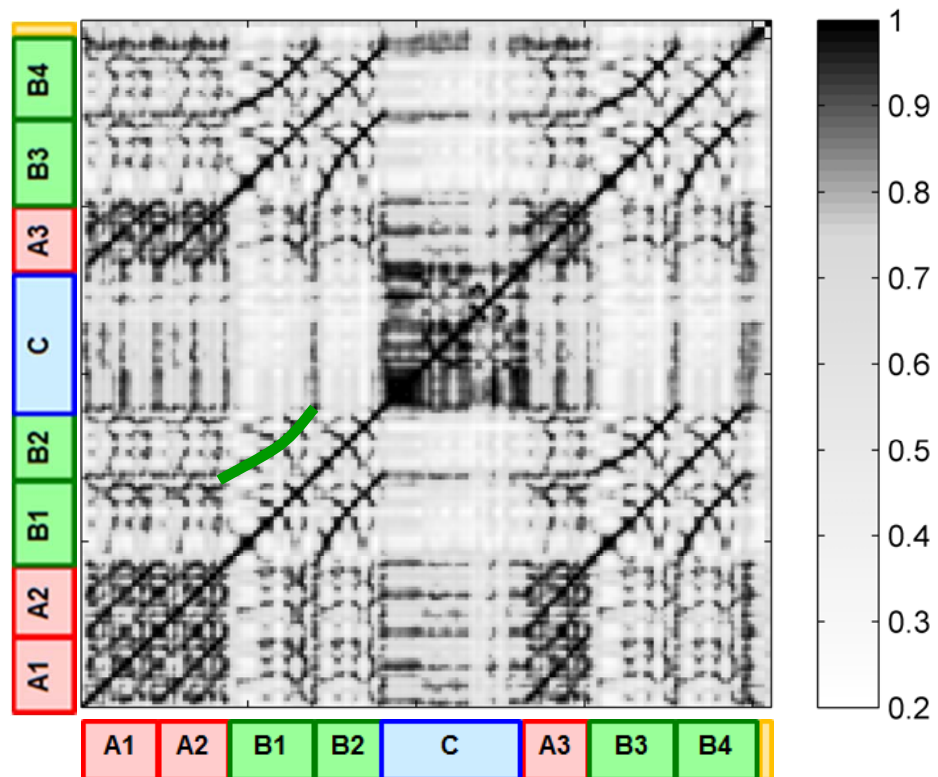
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



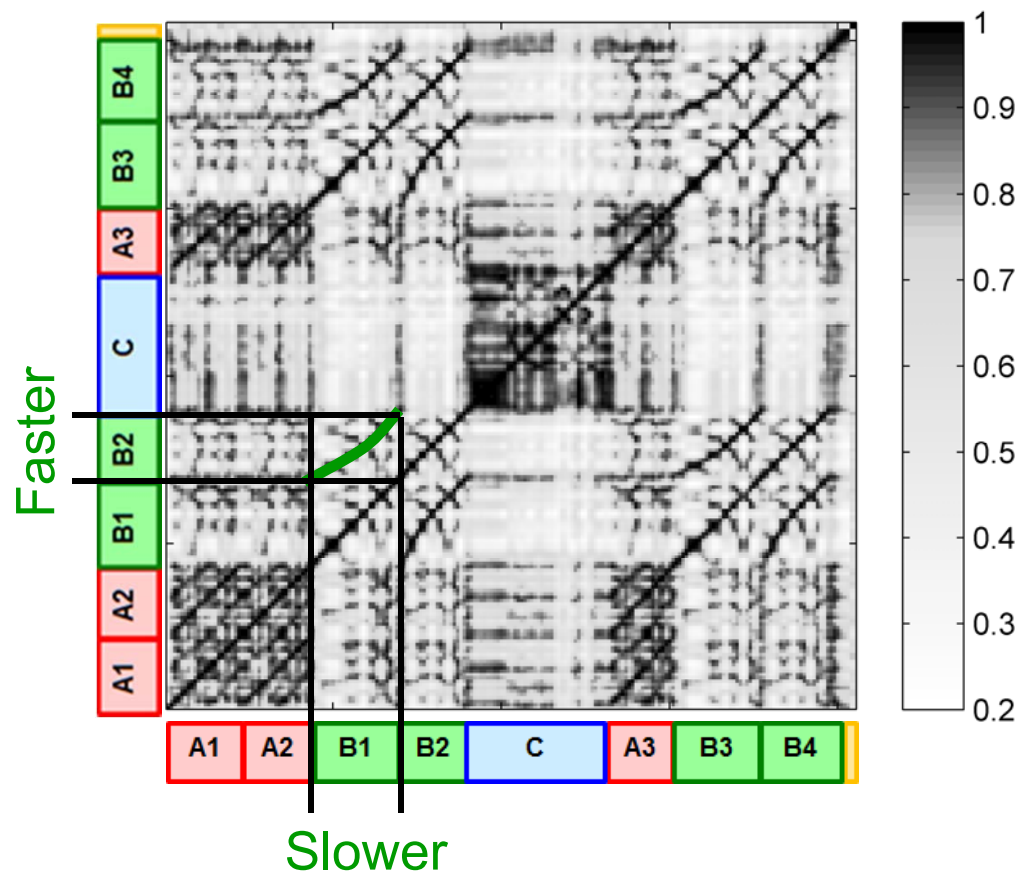
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



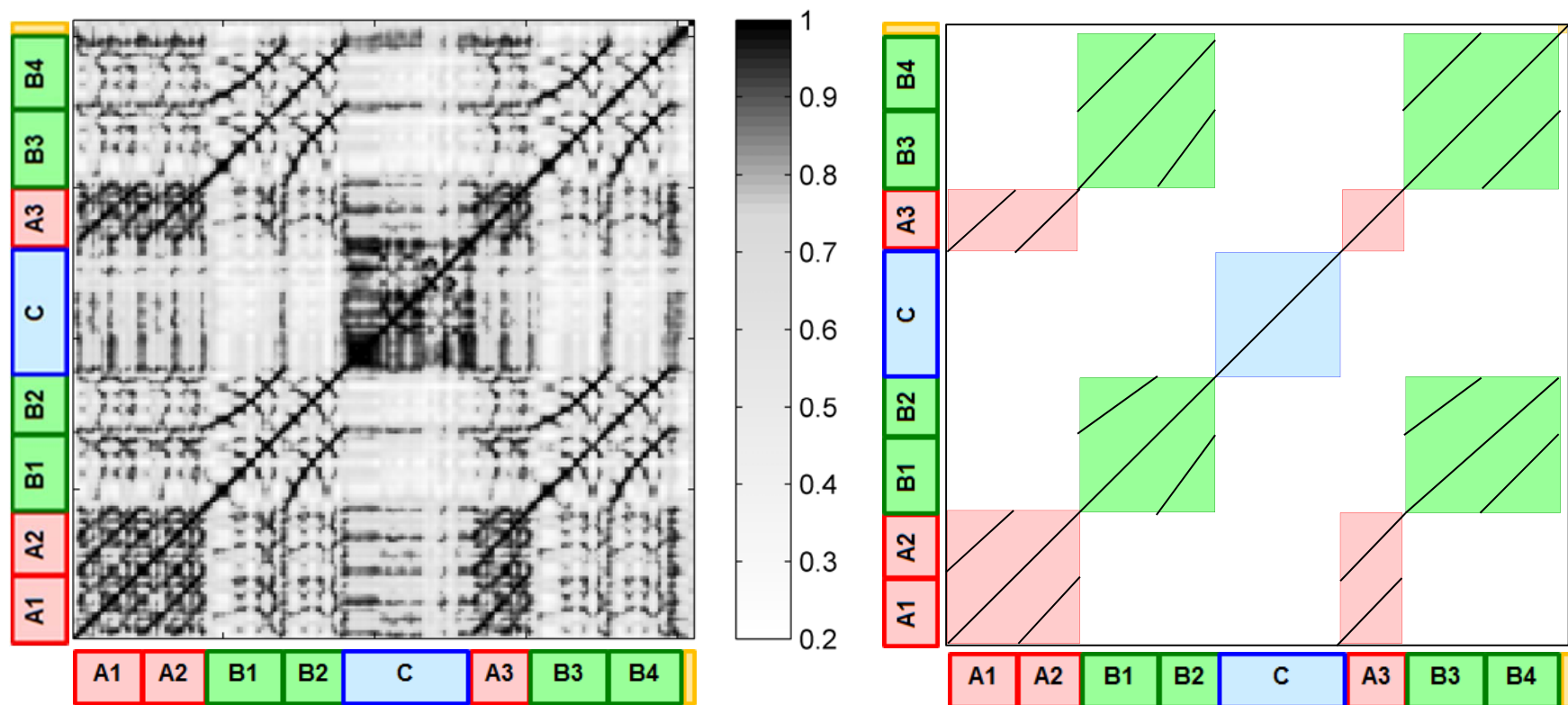
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Self-Similarity Matrix (SSM)

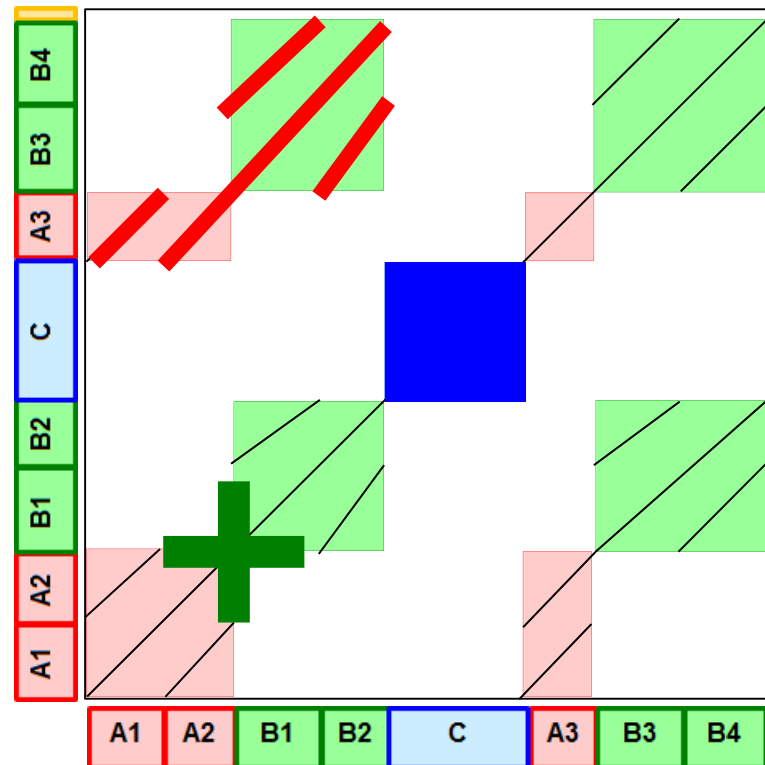
Example: Brahms Hungarian Dance No. 5 (Ormandy)

Blocks: Homogeneity

Paths: Repetition

Corners: Novelty

Idealized SSM



SSM Enhancement

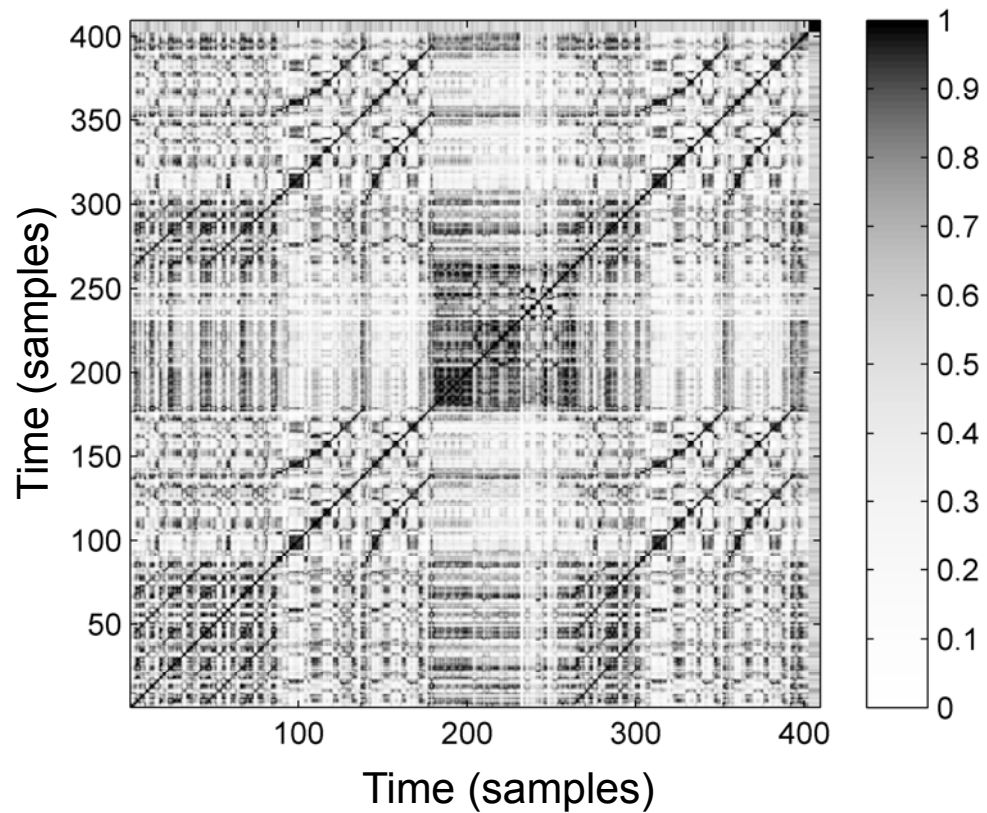
Challenge: Presence of musical variations

- Fragmented paths and gaps
- Paths of poor quality
- Regions of constant (low) cost
- Curved paths

Idea: Enhancement of path structure

SSM Enhancement

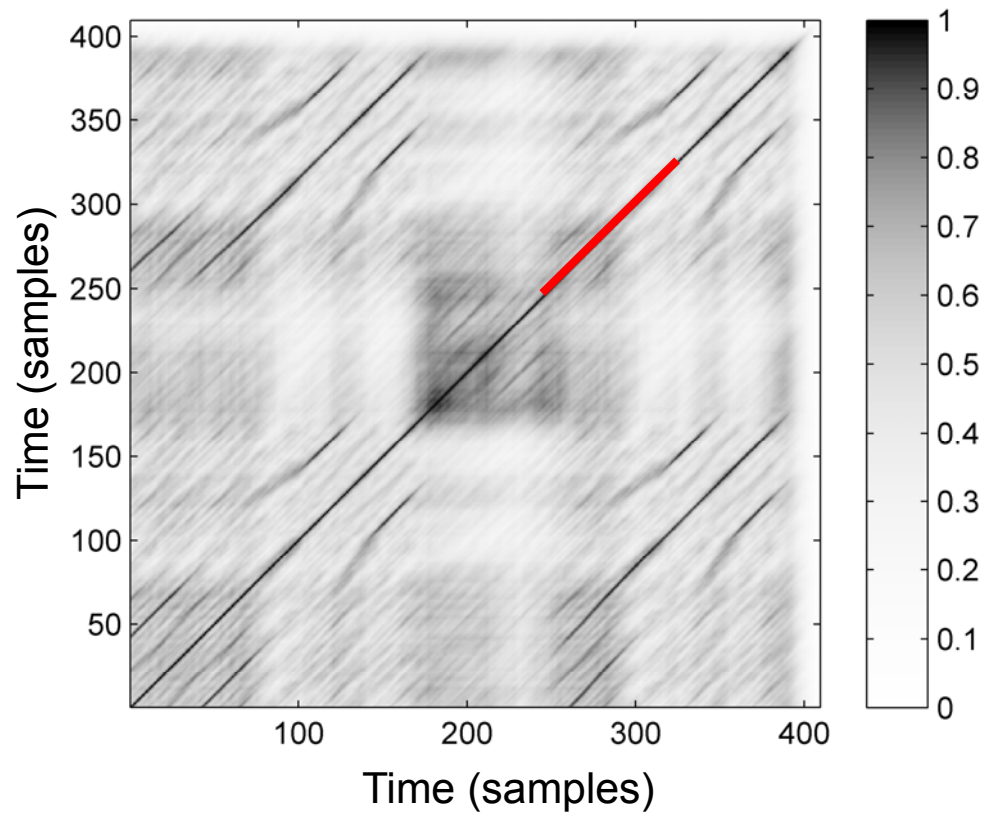
Path Enhancement



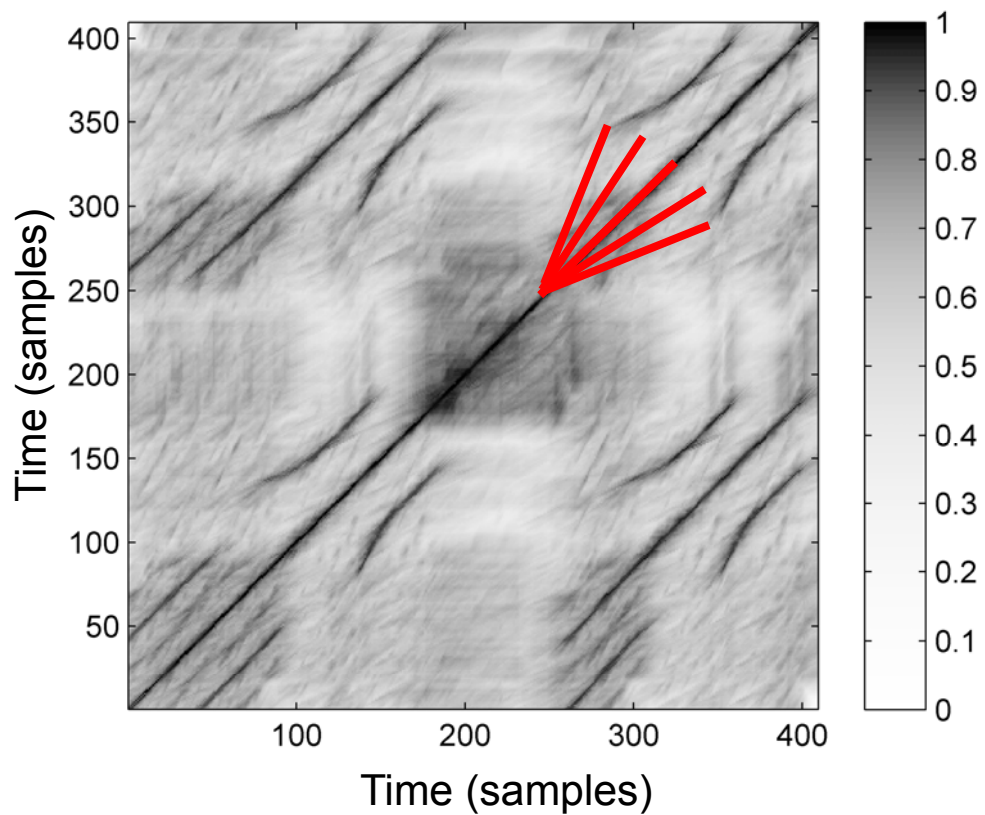
SSM Enhancement

Path Enhancement

- Diagonal smoothing



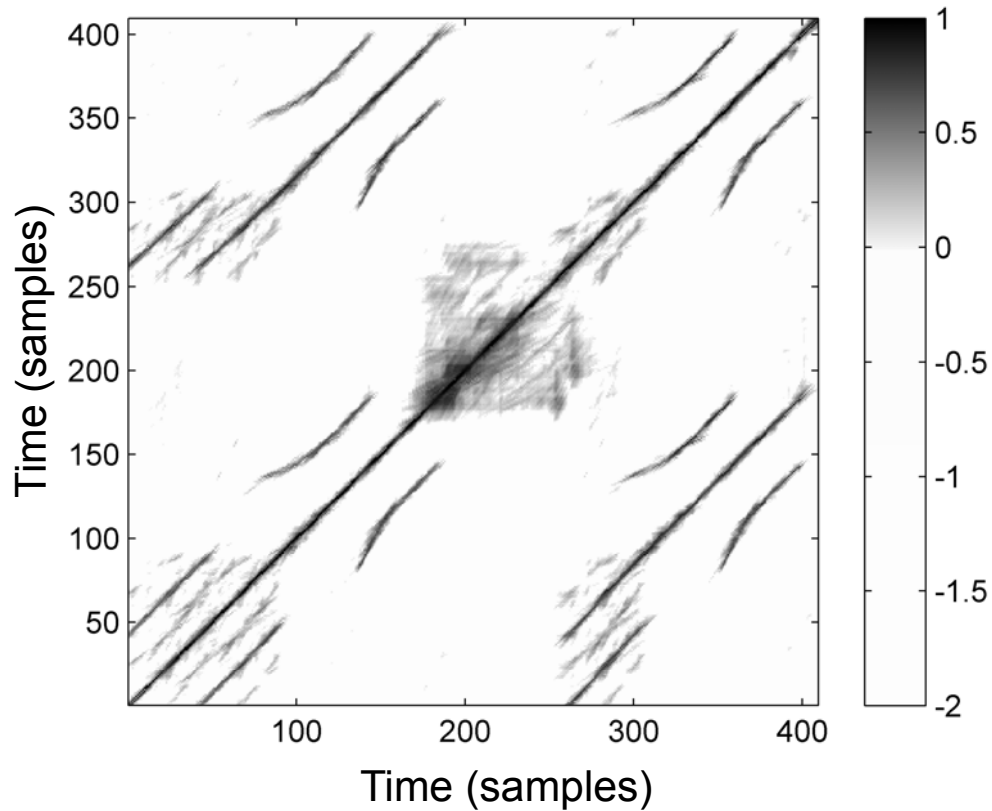
SSM Enhancement



Path Enhancement

- Diagonal smoothing
- Multiple filtering

SSM Enhancement



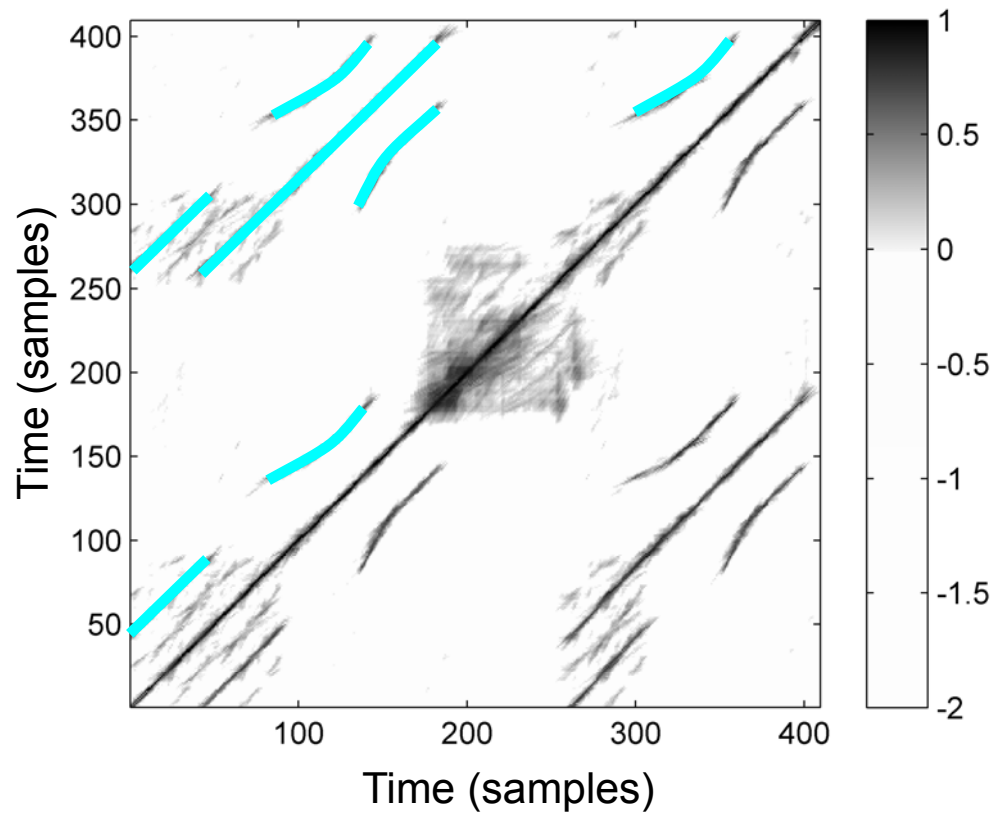
Path Enhancement

- Diagonal smoothing
- Multiple filtering
- Thresholding (relative)
- Scaling & penalty

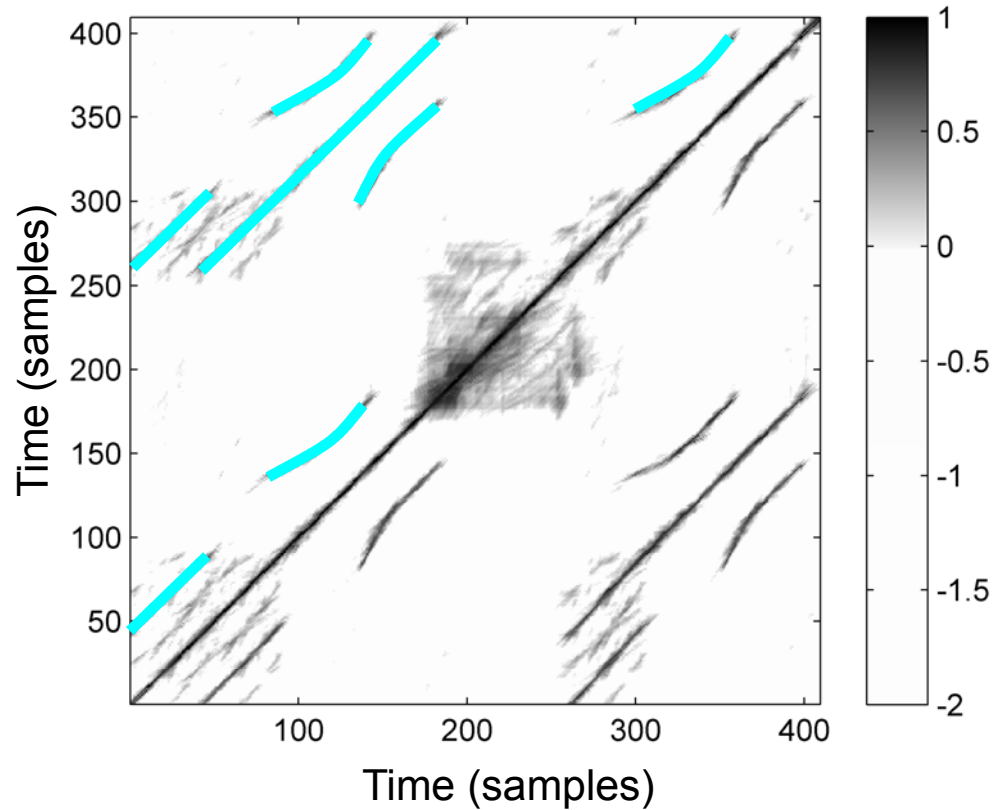
SSM Enhancement

Further Processing

- Path extraction

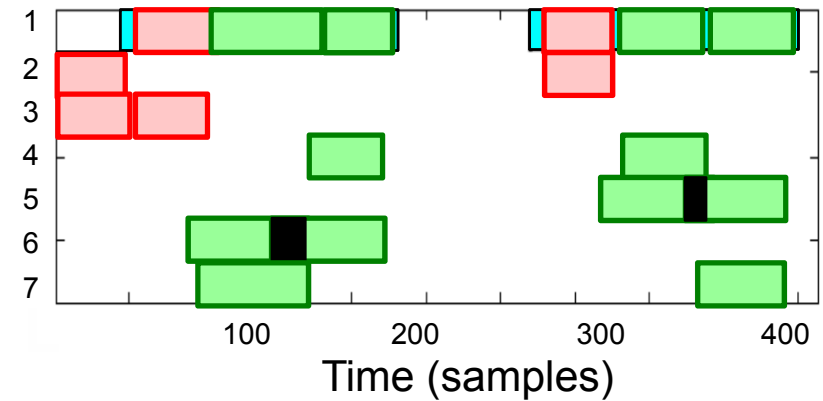


SSM Enhancement

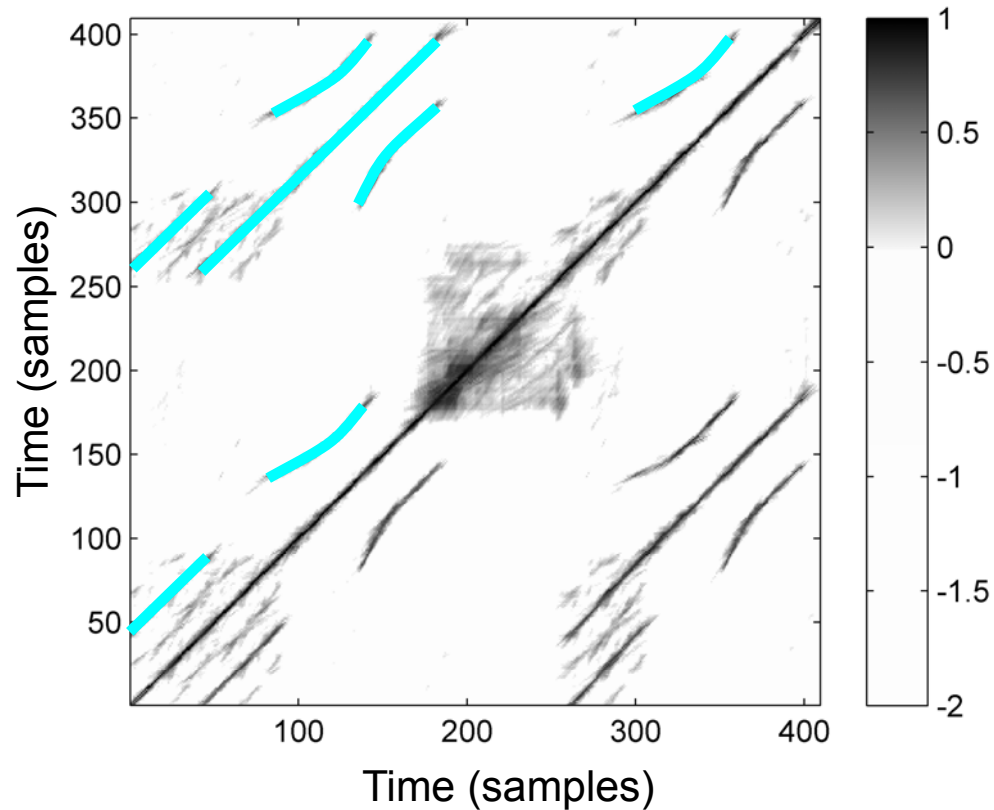


Further Processing

- Path extraction
- Pairwise relations
- Grouping (transitivity)

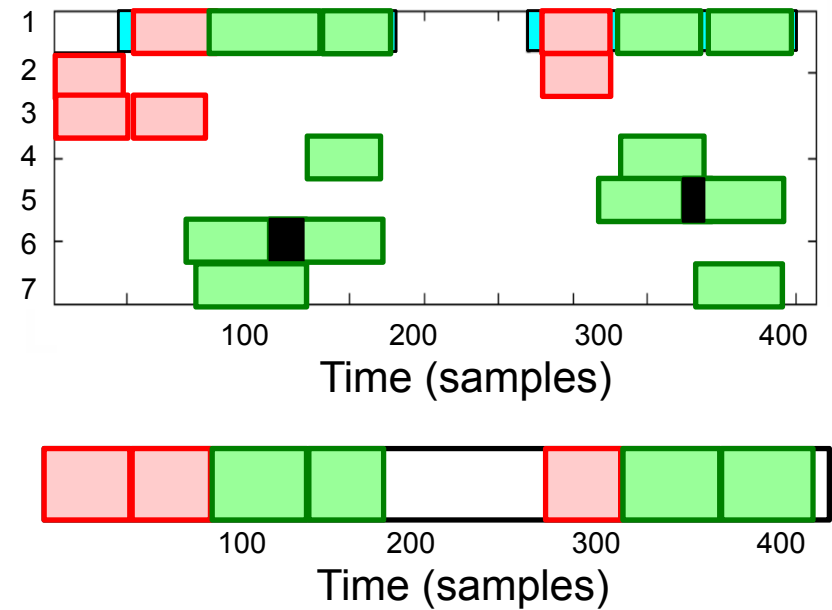


SSM Enhancement



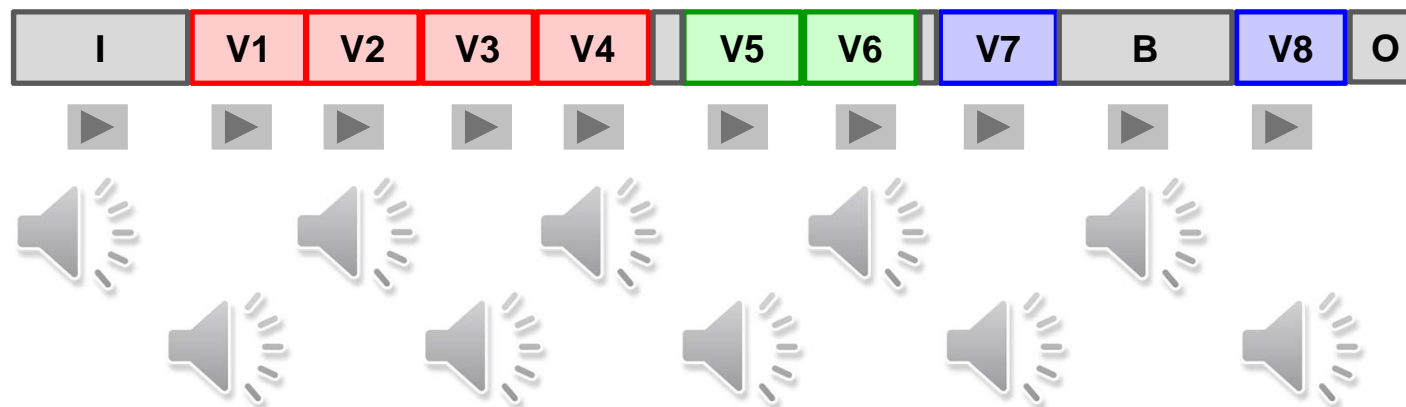
Further Processing

- Path extraction
- Pairwise relations
- Grouping (transitivity)



SSM Enhancement

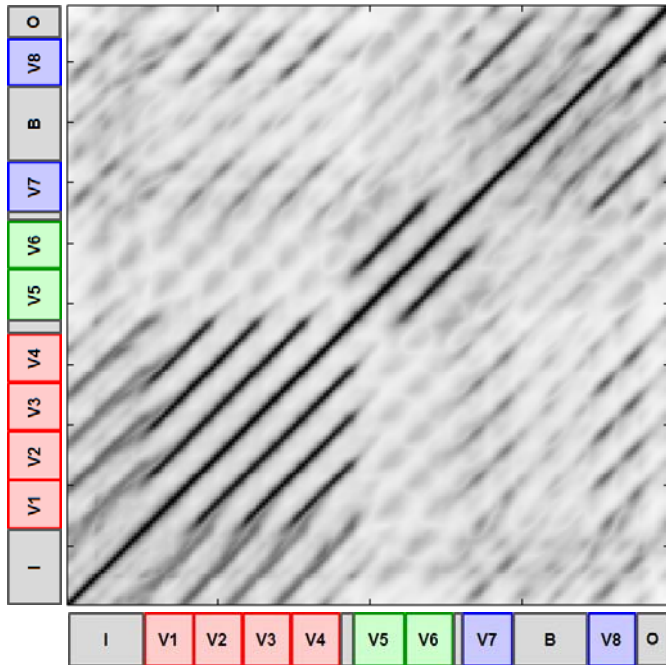
Example: Zager & Evans “In The Year 2525”



SSM Enhancement

Example: Zager & Evans “In The Year 2525”

Missing relations because of transposed sections

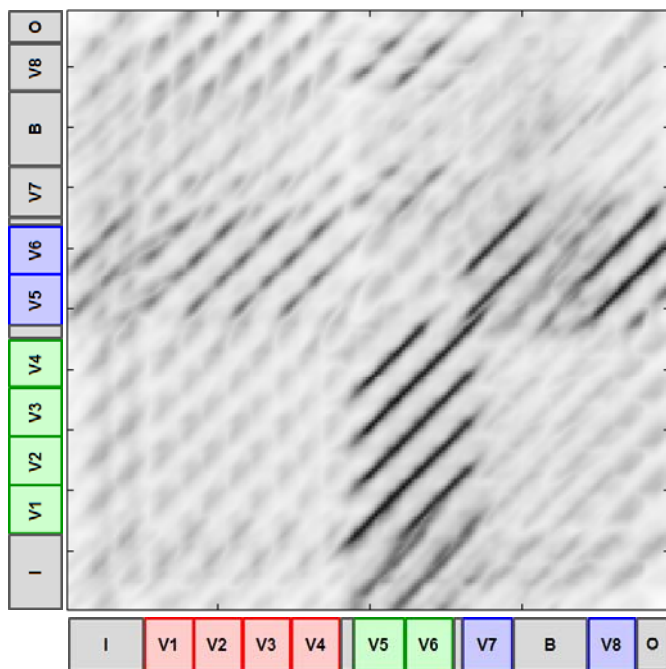


SSM Enhancement

Example: Zager & Evans “In The Year 2525”

Idea: Cyclic shift of one of the chroma sequences

One semitone up

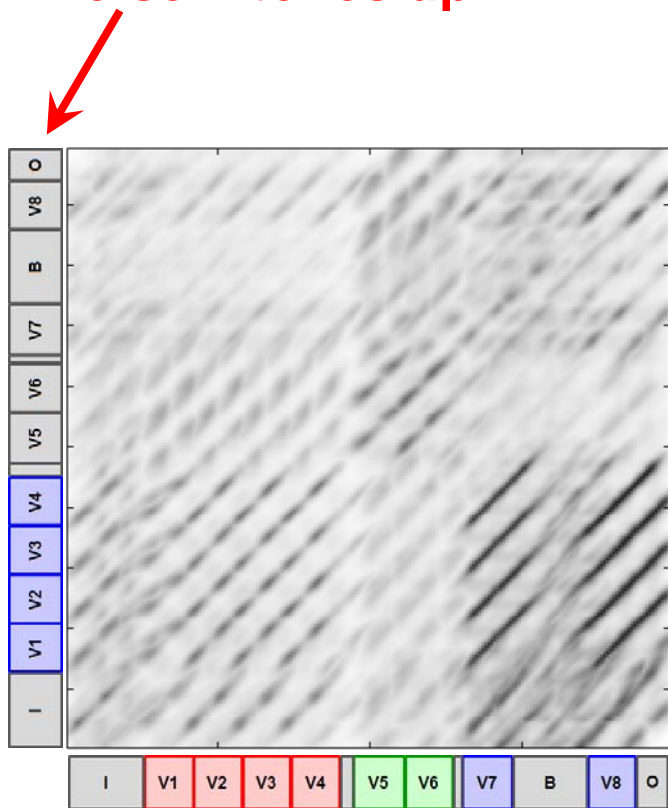


SSM Enhancement

Example: Zager & Evans “In The Year 2525”

Idea: Cyclic shift of one of the chroma sequences

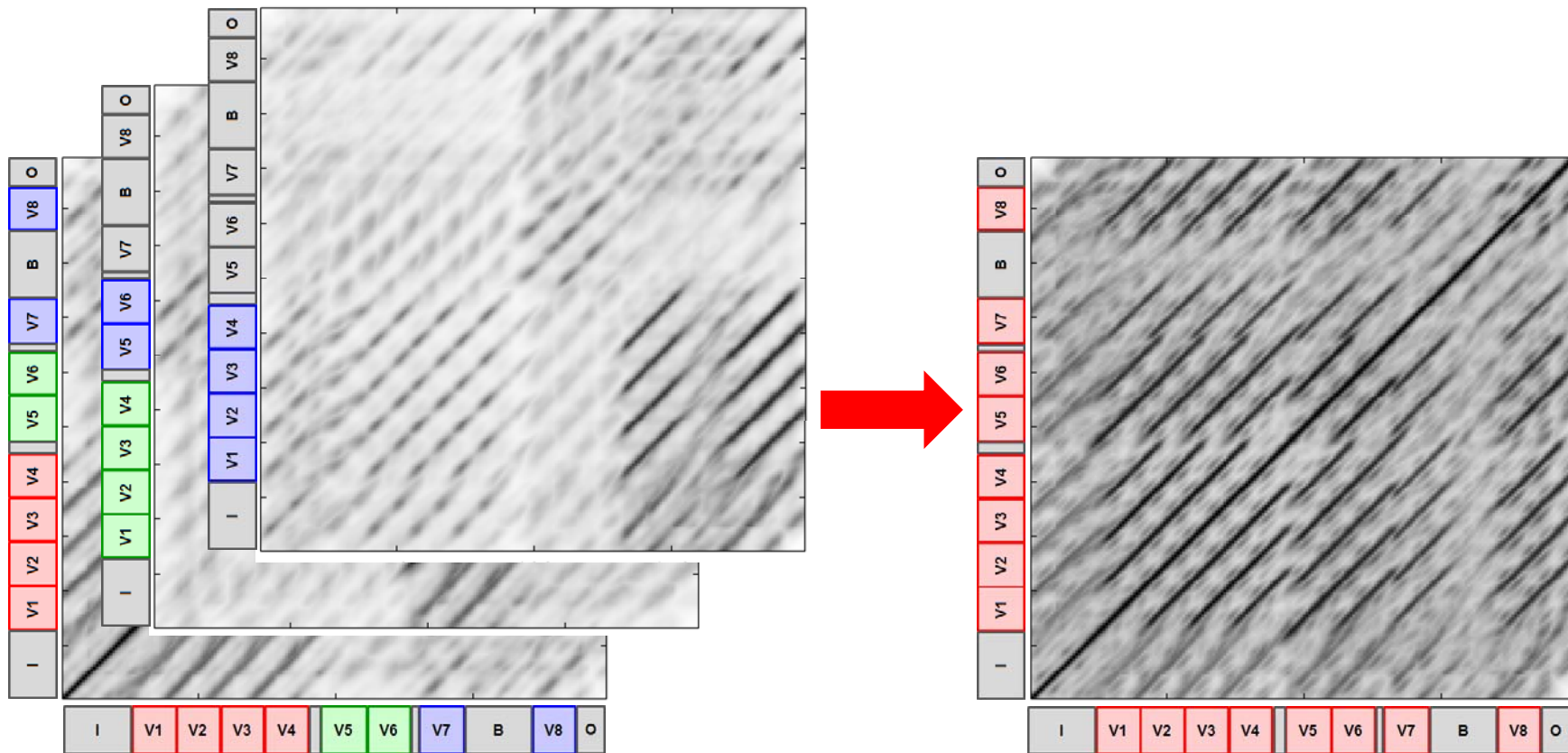
Two semitones up



SSM Enhancement

Example: Zager & Evans “In The Year 2525”

Idea: Overlay & Maximize \longrightarrow Transposition-invariant SSM



Overview

- Introduction
- Feature Representations
- Self-Similarity Matrices
- **Novelty-Based Segmentation**

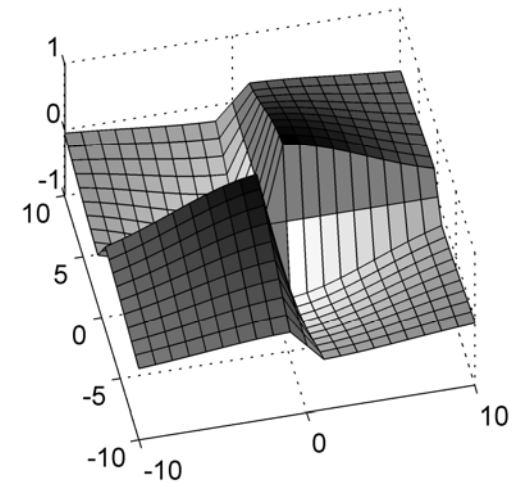
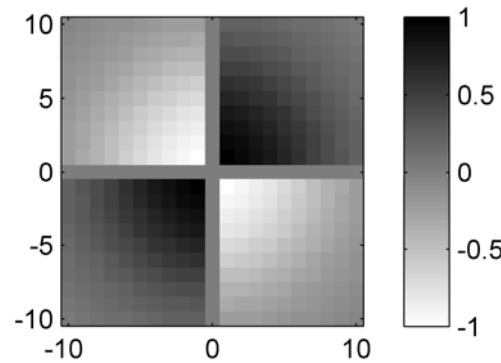
Novelty-Based Segmentation

General goals:

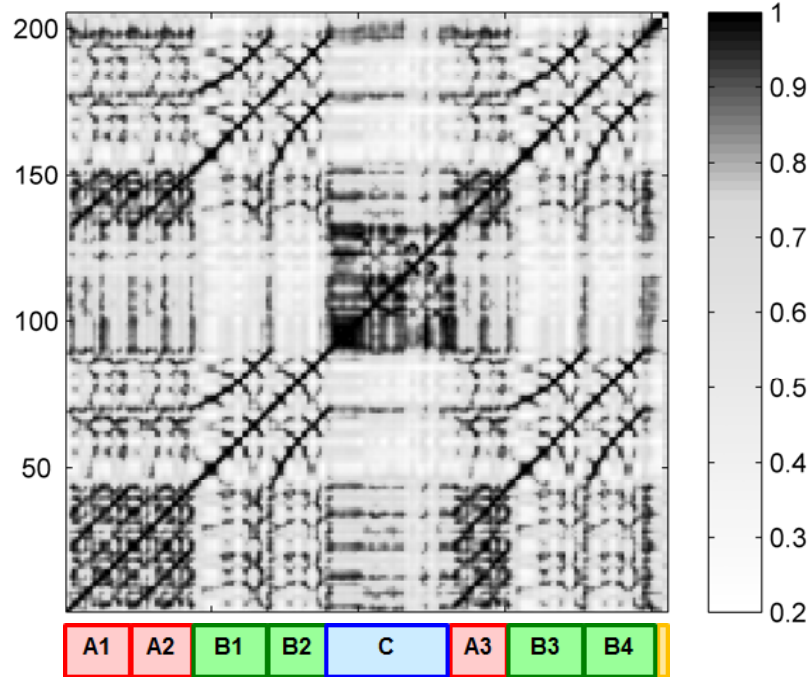
- Find instances where musical changes occur.
- Find transition between subsequent musical parts.

Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.



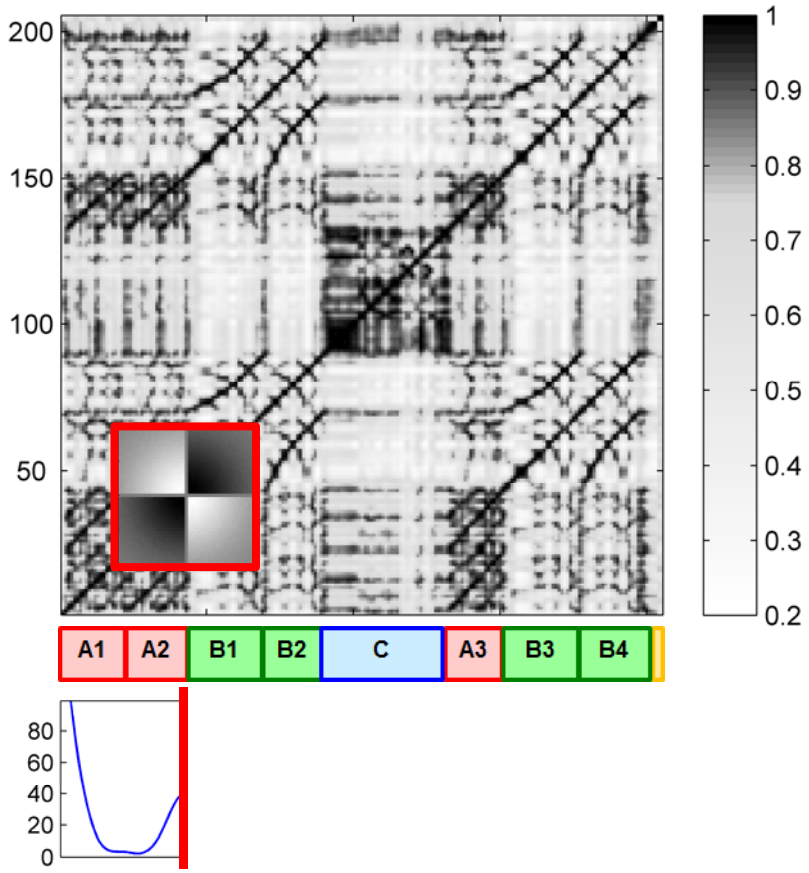
Novelty-Based Segmentation



Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

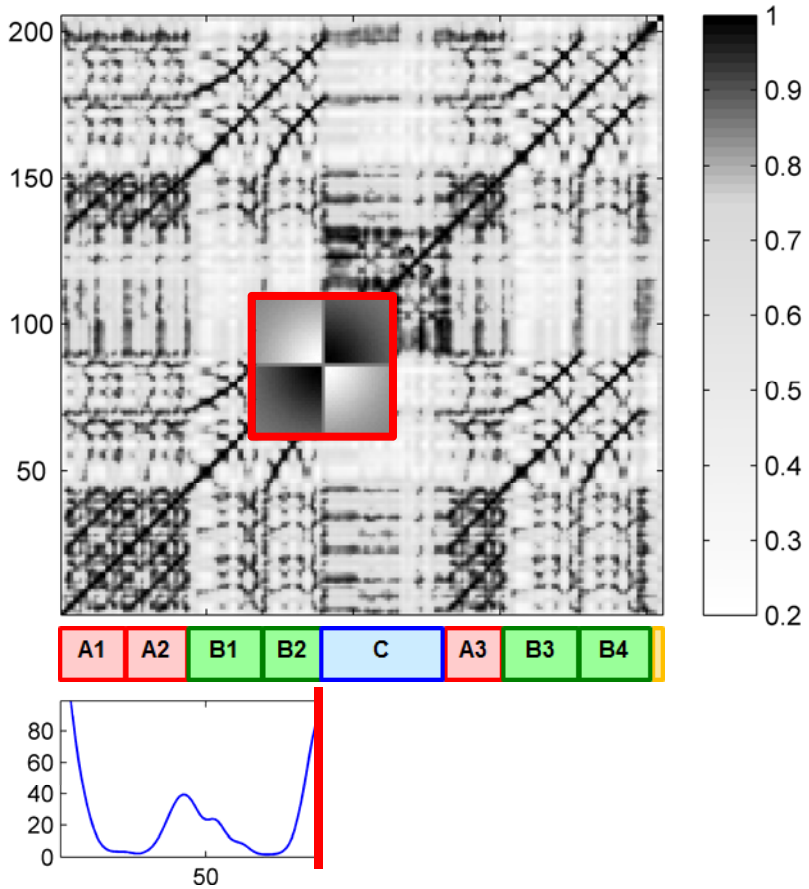
Novelty-Based Segmentation



Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

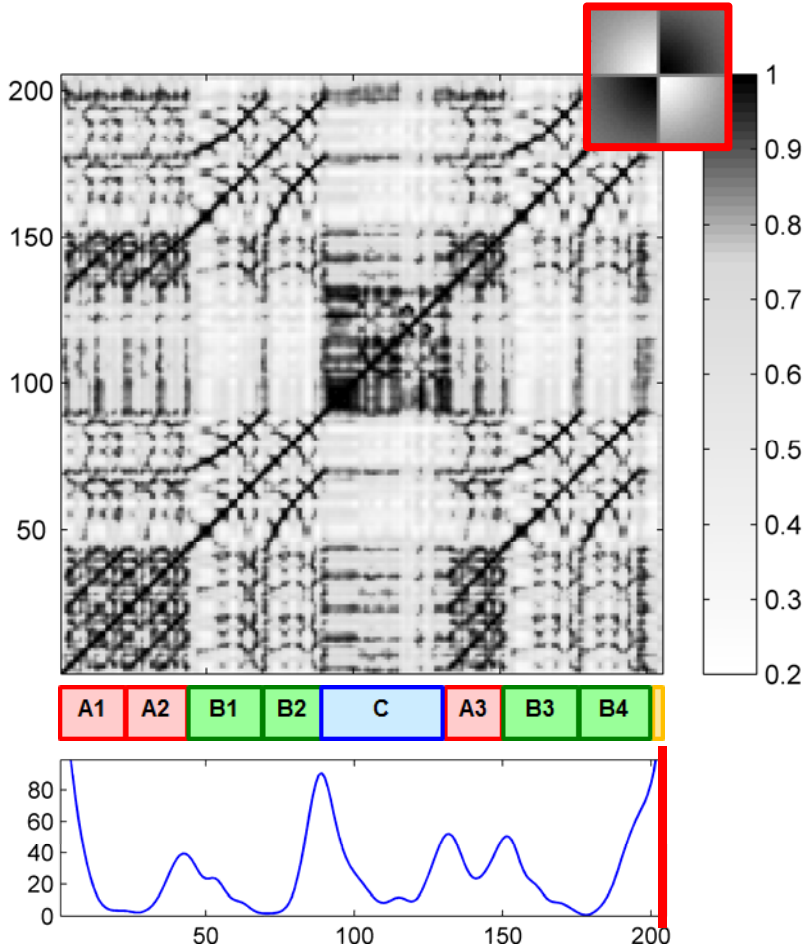
Novelty-Based Segmentation



Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

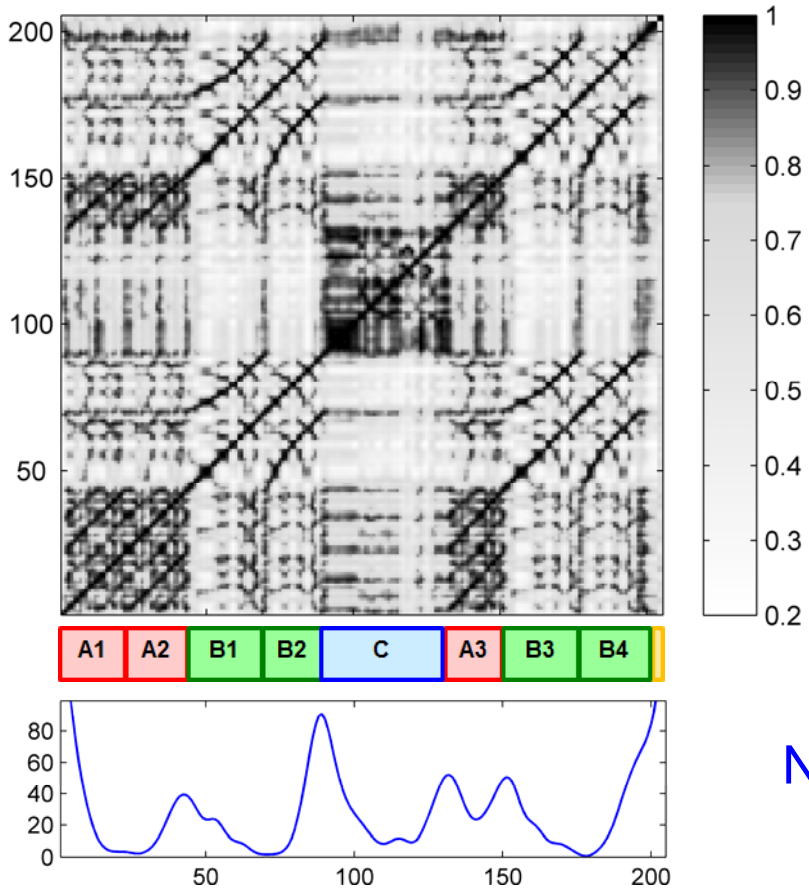
Novelty-Based Segmentation



Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

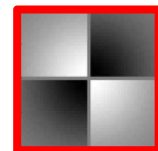
Novelty-Based Segmentation



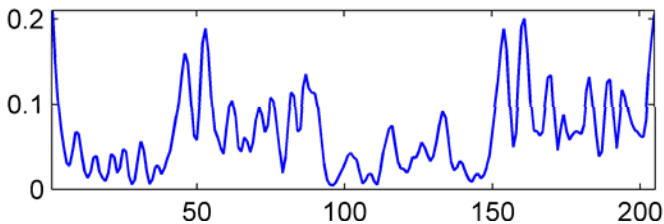
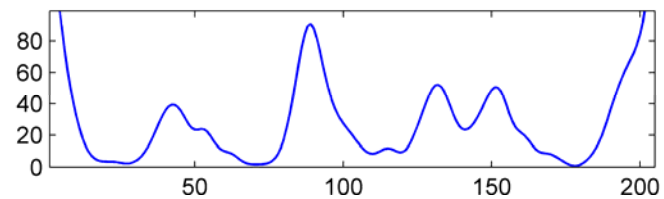
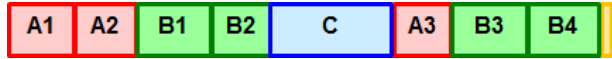
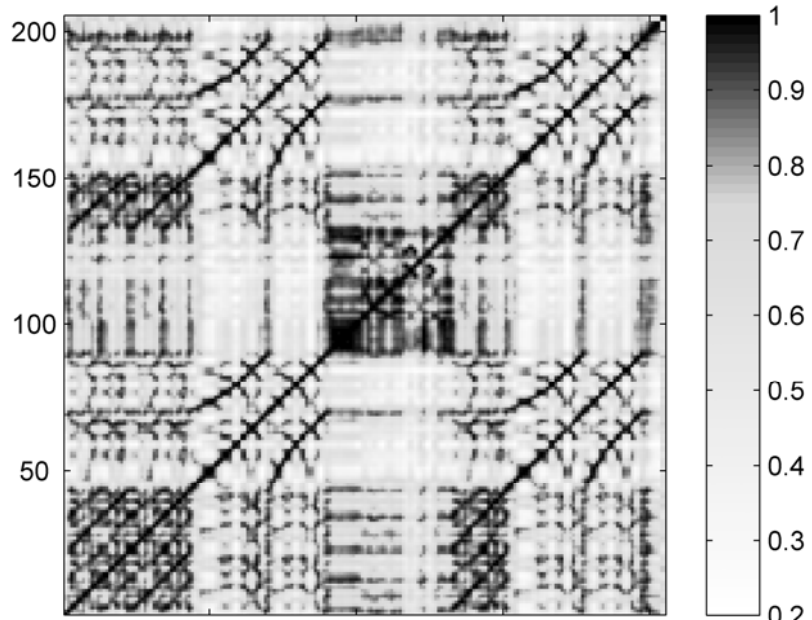
Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

Novelty function using



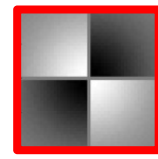
Novelty-Based Segmentation



Idea (Foote):

Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

Novelty function using



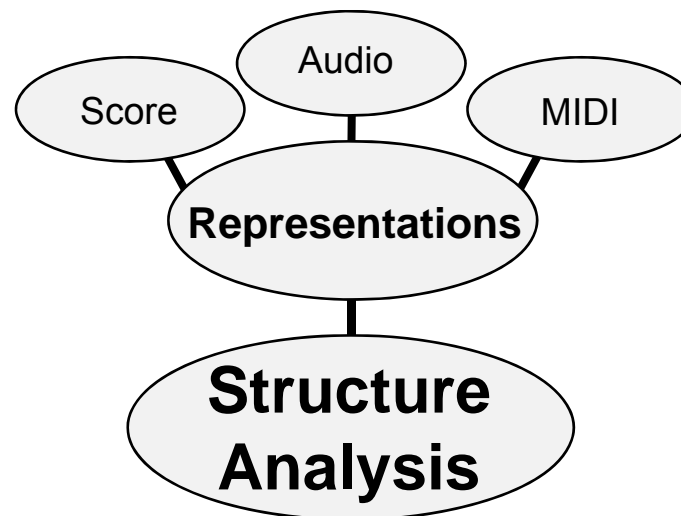
Novelty function using



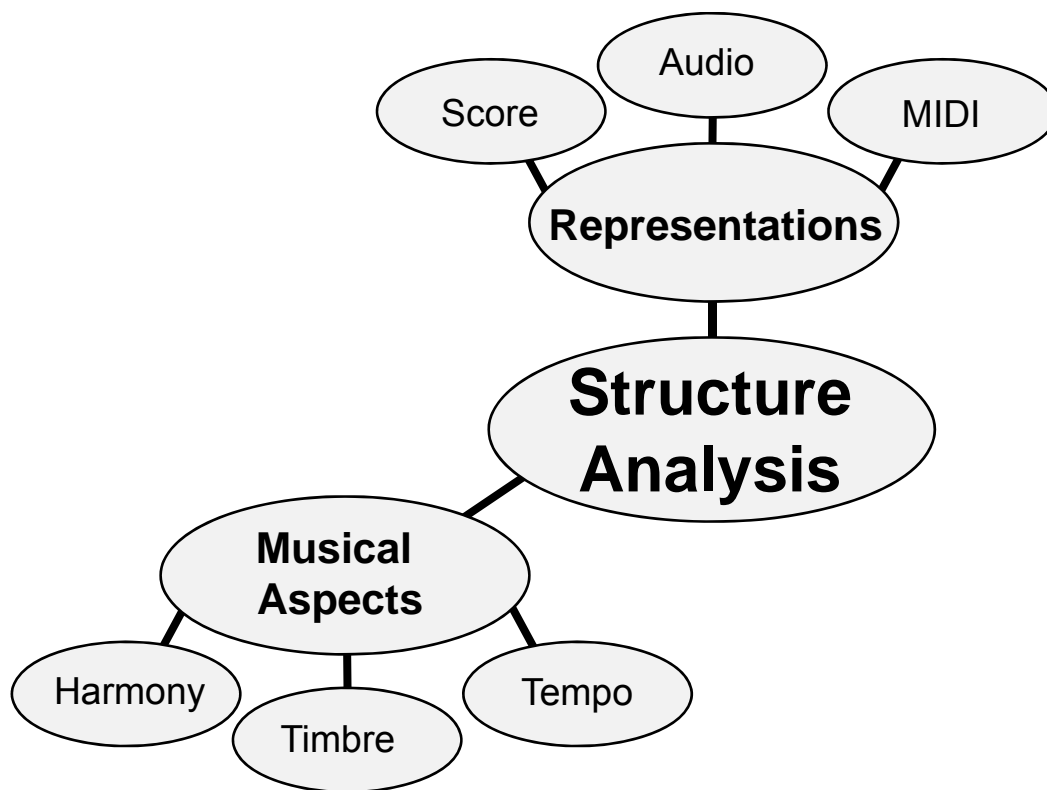
Conclusions

**Structure
Analysis**

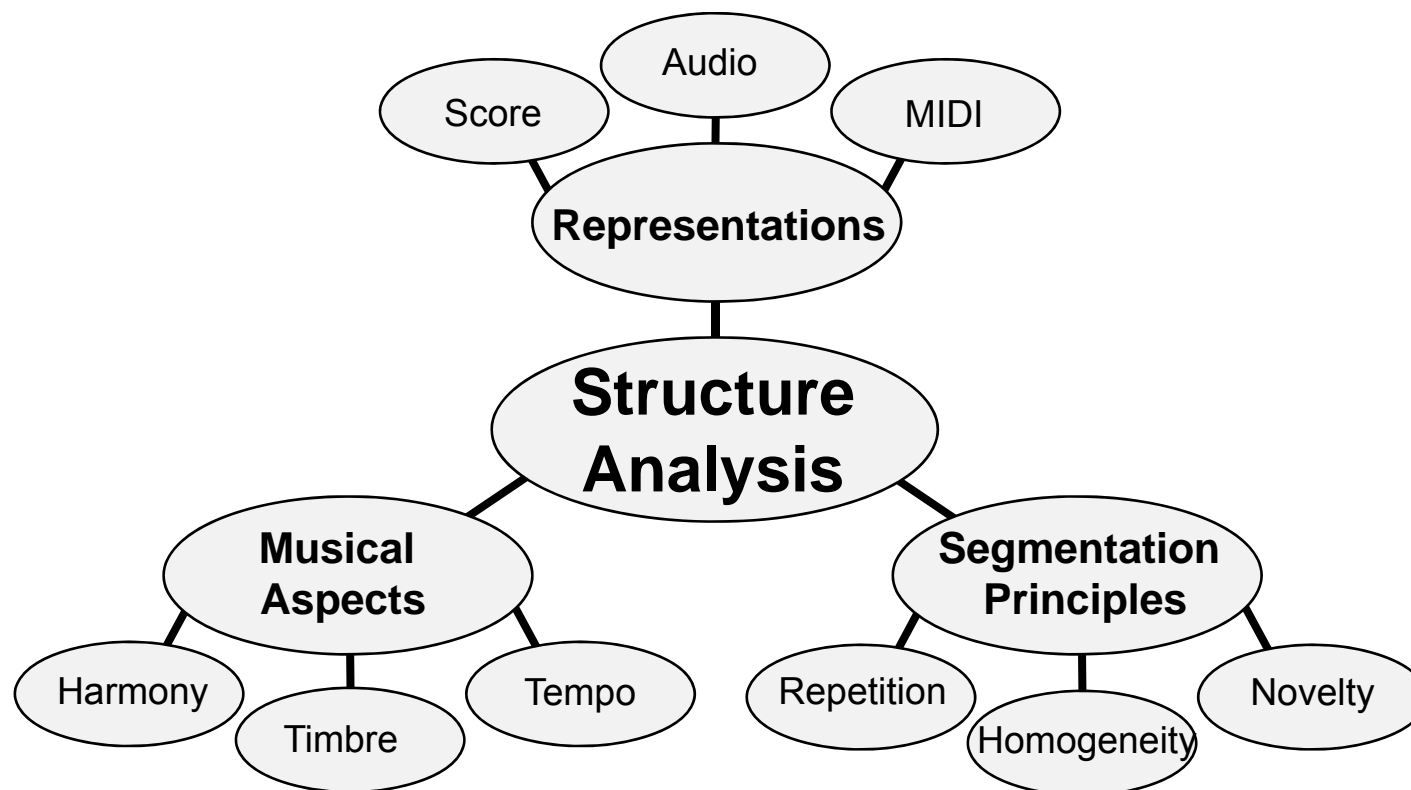
Conclusions



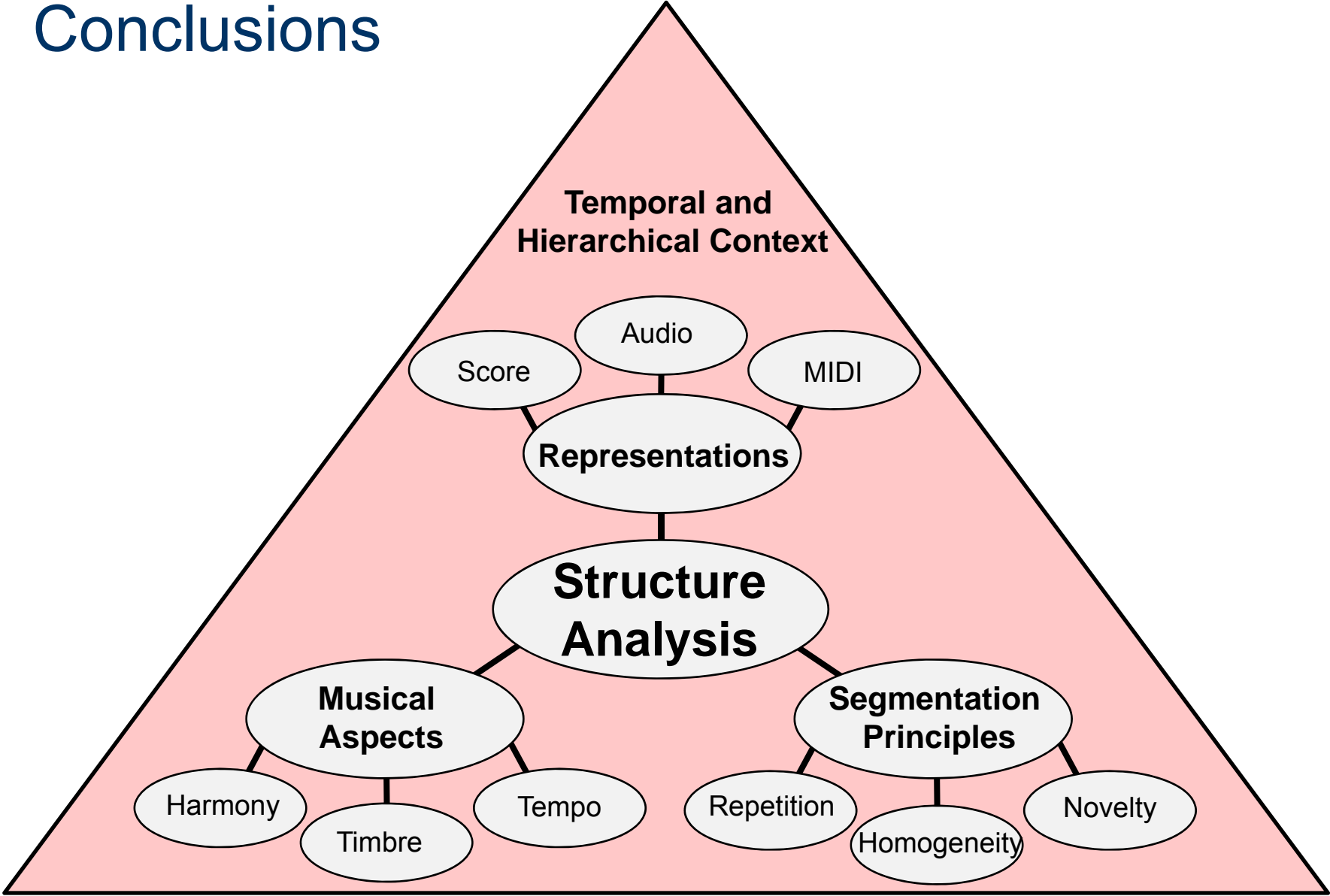
Conclusions



Conclusions



Conclusions



Links

- SM Toolbox (MATLAB)
<http://www.audiolabs-erlangen.de/resources/MIR/SMtoolbox/>
- MSAF: Music Structure Analysis Framework (Python)
<https://github.com/uriniето/msaf>
- SALAMI Annotation Data
<http://ddmal.music.mcgill.ca/research/salami/annotations>
- LibROSA (Python)
<https://librosa.github.io/librosa/>
- Evaluation: mir_eval (Python)
https://craffel.github.io/mir_eval/
- Deep Learning: Boundary Detection
Jan Schlüter (PhD thesis)