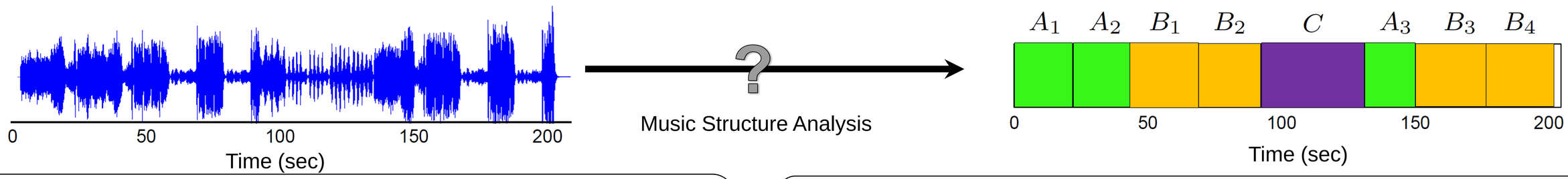
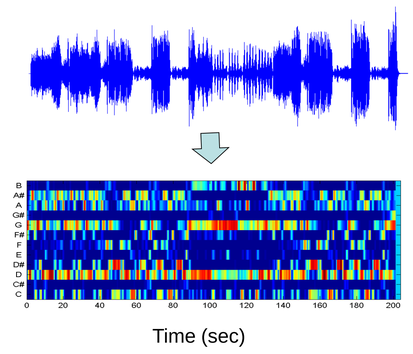
Music Structure Analysis: Audio Thumbnailing and Structure Scape Plot

Nanzhu Jiang
Saarland University and MPI Informatik, Germany



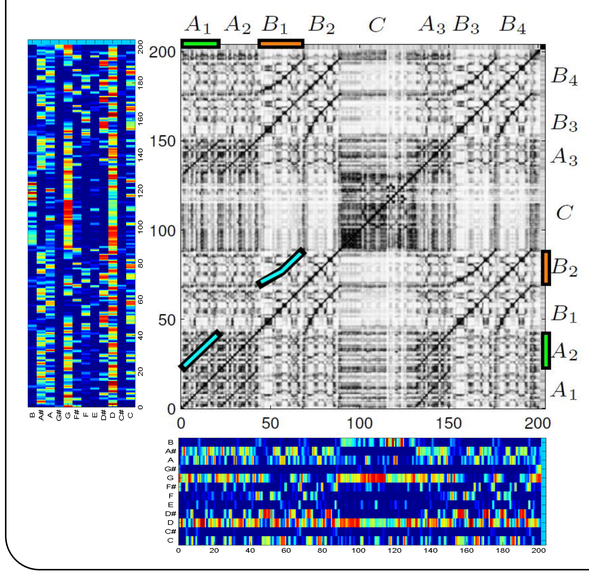
Feature Extraction



Convert a signal into musical meaningful features:

- Indicate note presence (via Fourier Transform / Filter Bank)
- Indicate pitch class presence (collapse all notes in the same pitch class, more robust features)

Self Similarity Matrix

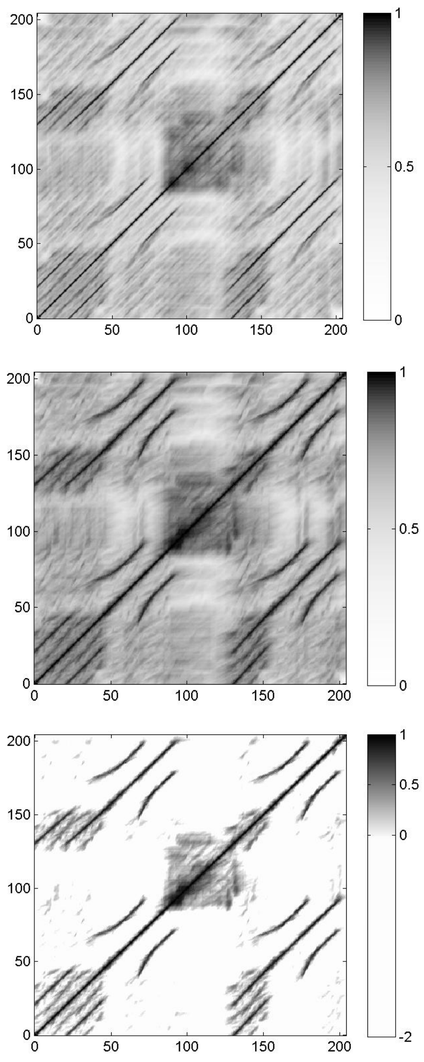


Self Similarity Matrix (SSM):

- Elements of SSM encode similarities of features.
- Strong paths in SSM (black stripes) indicate high similarity of feature sequences. (e.g. A1 ~ A2)
- Pair of similar feature sequence represent pair of similar segment: **repetitions**.

Path Detection Enhancement

Problem: musical variations often lead to noisy and fragmented paths.



Diagonal Filtering

Aimed Problem:
- Small gaps in paths.
- Noisy structure around paths.

Strategy:
Smoothing filter along the direction of the main diagonal.

Tempo-aware Filtering

Aimed Problem:
- Small gaps in paths.
- Noisy structure around paths.

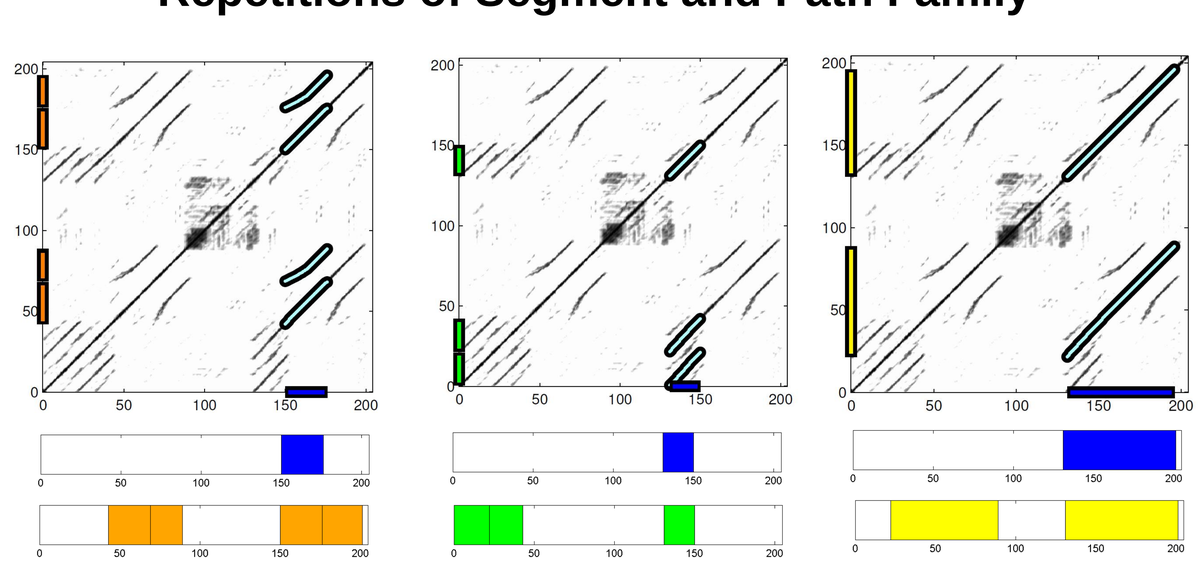
Strategy:
Smoothing filter covers tempo variations along different gradients.

Thresholding

Aimed Problem:
- General noise
- Weak paths

Strategy:
Emphasize score of strong paths and penalize all gaps inside or around paths.

Repetitions of Segment and Path Family



Segment [150:176]: Musical Part B_3

Four Repeated Segments: B_1, B_2, B_3, B_4.

Segment [131:150] : Musical Part A_3

Three repeated segments to A_1, A_2, A_3

Segment [131:150] : Musical Part $A_3B_3B_4$

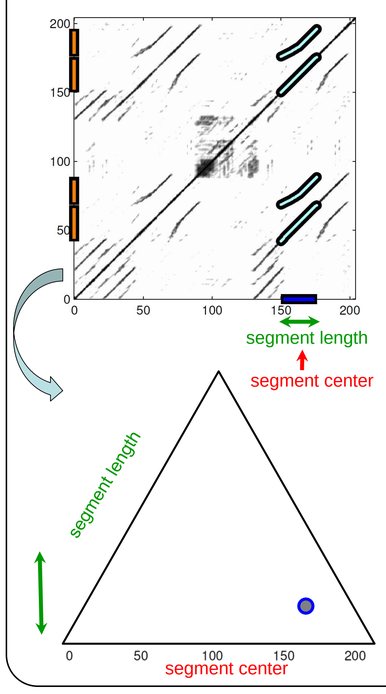
Two repeated segments: $A_2B_1B_2$, $A_3B_3B_4$

Audio Thumbnailing

Audio Thumbnail: most **repeated** and **representative** segment of a recording.
Whether a segment can be considered as an thumbnail depends on:
- coverage of all repetitions in the whole recording.
- quality of repetitions in the whole recording.

Compute a value which capture these two properties together: **fitness measure**

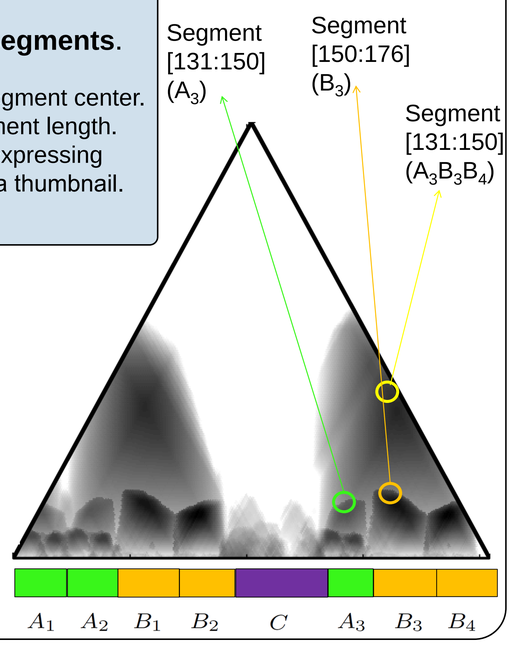
Fitness Scape Plot



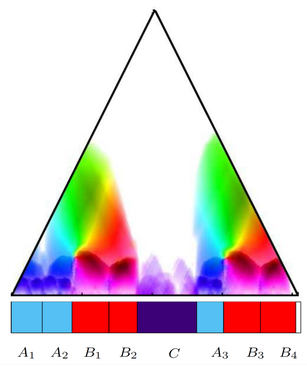
Visualize all segments.
Use **points** to represent **segments**.

Point horizontal coordinate: segment center.
Point vertical coordinate: segment length.
Point lightness: fitness value expressing the segment's ability of being a thumbnail.

fitness measure computing for all segments



Structure Scape Plot:



Color coding of visualization.

- Segment relationship:
  similar segment s with similar color.
  different segments with far different color.
- Segment representativeness:
  lightness of the color to represent fitness.

References:

• Müller, Grosche and Jiang : *A segment-based fitness measure for capturing repetitive structures of music recordings"* ISMIR 2011.

• Müller and Jiang : "*A scape plot representation for visualizing repetitive structures of music recording*" ISMIR 2012

max planck institut informatik

UNIVERSITÄT DES SAARLANDES

International Max Planck Research School for Computer Science