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

**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,…
- **Variation:** Modification and transformation
Overview

- Introduction
- Self-Similarity Matrices
- Audio Thumbnailing
- Novelty-based Segmentation
- Converting Path to Block Structures

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

Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Chroma (Harmony)

MFCC (Timbre)

Tempo
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Blocks: Homogeneity
Paths: Repetition
Corners: Novelty
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Chroma (Harmony)

Idealized SSM

G major G minor G minor

SSM Enhancement

Block Enhancement
- Feature smoothing
- Coarsening

Path Enhancement
SSM Enhancement

Path Enhancement
- Diagonal smoothing

SSM Enhancement

Path Enhancement
- Diagonal smoothing
- Multiple filtering

SSM Enhancement

Path Enhancement
- Diagonal smoothing
- Multiple filtering
- Forward-backward

SSM Enhancement

Path Enhancement
- Diagonal smoothing
- Multiple filtering
- Forward-backward
- Thresholding (binary)

SSM Enhancement

Path Enhancement
- Diagonal smoothing
- Multiple filtering
- Forward-backward
- Thresholding (relative)
- Scaling & penalty
SSM Enhancement

Example: Zager & Evans “In The Year 2525”

Missing relations because of transposed sections

Idea: Cyclic shift of one of the chroma sequences

One semitone up

Two semitones up

SSM Enhancement

Example: Zager & Evans “In The Year 2525”

Idea: Overlay & add up  Transposition-invariant SSM
SSM Enhancement

Example: Zager & Evans "In The Year 2525"
Note: Order of enhancement steps important!

Adding up

Smoothing & adding up

Similarity Matrix Toolbox

Meinard Müller, Nanzhu Jiang, Harald Grohganz
SM Toolbox: MATLAB Implementations for Computing and Enhancing Similarity Matrices

http://www.audiolabs-erlangen.de/resources/MIR/SMtoolbox/

Overview

- Introduction
- Self-Similarity Matrices
- Audio Thumbnailing
- Novelty-based Segmentation
- Converting Path to Block Structures

Thanks:
- Jiang, Grosche
- Peeters
- Cooper, Foote
- Goto
- Levy, Sandler
- Mauch
- Sapp

Audio Thumbnailing

General goal: Determine the most representative section ("Thumbnail") of a given music recording.

Example: Zager & Evans "In The Year 2525"

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Thumbnail is often assumed to be the most repetitive segment

Audio Thumbnailing

Two steps
1. Path extraction
   - Paths of poor quality (fragmented, gaps)
   - Block-like structures
   - Curved paths

2. Grouping
   - Noisy relations (missing, distorted, overlapping)
   - Transitivity computation difficult

Both steps are problematic!

Main idea: Do both, path extraction and grouping, jointly

- One optimization scheme for both steps
- Stabilizing effect
- Efficient

Main idea: Do both path extraction and grouping jointly

- For each audio segment we define a fitness value
- This fitness value expresses "how well" the segment explains the entire audio recording
- The segment with the highest fitness value is considered to be the thumbnail
- As main technical concept we introduce the notion of a path family
Fitness Measure

Enhanced SSM

Fitness Measure

Path over segment
- Consider a fixed segment
- Path over segment
- Induced segment
- Score is high

Fitness Measure

Path over segment
- Consider a fixed segment
- Path over segment
- Induced segment
- Score is not so high
- A second path over segment
- Induced segment
- Score is not so high

Fitness Measure

Path over segment
- Consider a fixed segment
- Path over segment
- Induced segment
- Score is high
- A second path over segment
- Induced segment
- Score is not so high
- A third path over segment
- Induced segment
- Score is very low

Fitness Measure

Path family
- Consider a fixed segment
- A path family over a segment
- The induced segments do not overlap.
Consider a fixed segment
A path family over a segment is a family of paths such that the induced segments do not overlap.

This is not a path family!

Consider a fixed segment
A path family over a segment is a family of paths such that the induced segments do not overlap.

This is a path family!

(Even though not a good one)

Consider a fixed segment
Consider over the segment the optimal path family, i.e., the path family having maximal overall score.
Call this value: Score(segment)

Furthermore consider the amount covered by the induced segments.
Call this value: Coverage(segment)

Note: This optimal path family can be computed using dynamic programming.
Fitness Measure

Consider a fixed segment
Self-explanation are trivial!
Subtract length of segment
Normalization

Fitness

P := Normalize(Score(segment) - length(segment)) ∈ [0,1]
R := Normalize(Coverage(segment) - length(segment)) ∈ [0,1]

F := 2 * P * R / (P + R)
Note: Self-explanations are ignored → fitness is zero

**Thumbnail**: segment having the highest fitness

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

- Introduction
- Self-Similarity Matrices
- Audio Thumbnailing
- Novelty-based Segmentation
- Converting Path to Block Structures

Thanks:
- Foote
- Serra, Grosche, Arcos
- Goto
- Tzanetakis, Cook

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Example: Zager & Evans “In The Year 2525”
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 function using

Structure features
- Enhanced SSM
- Time-lag SSM
- Cyclic time-lag SSM
- Columns as features
Novelty-based Segmentation

**Example:** Chopin Mazurka Op. 24, No. 1

Overview

- Introduction
- Self-Similarity Matrices
- Audio Thumbnailing
- Novelty-based Segmentation
- Converting Path to Block Structures

Thanks:

- Grohganz, Clausen
- Kaiser
- Peeters
- Dubnov, Apel
- Serra, Grosche, Arcos

Converting Path to Block Structures

**Motivation**

- Perform joint analysis using repetitive as well as homogeneous aspects
- Make homogeneity-based methods applicable to repetition-based analysis
Converting Path to Block Structures

**Motivation**

- Homogeneity
- Repetition

**Procedure**

- Enhanced SSM
- Thresholding & image processing
- Eigenvalue decomposition
- Weighting
- Clustering & smoothing

Converting Path to Block Structures
Converting Path to Block Structures

**Procedure**
- Enhanced SSM
- Thresholding & image processing
- Eigenvalue decomposition
- Weighting
- Clustering & smoothing
- Columns as features

Final matrix shows paths as blocks

Conclusions

- Repetition, Homogeneity, Novelty
- Combined Approaches
- Hierarchical Approaches
- Evaluation
- Explaining Structure

PhD Projects (Final Stage)

- Nanzhu Jiang
  Universität Erlangen-Nürnberg  
  Supervisor: Meinard Müller

- Harald Grohganz
  Universität Bonn  
  Supervisors: Michael Clausen, Meinard Müller

- Jordan Smith
  Queen Mary University of London  
  Supervisor: Elaine Chew

- Oriol Nieto
  New York University  
  Supervisor: Juan P. Bello

Book Project

**A First Course on Music Processing**

Textbook (approx. 500 pages)

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. Music Transcription

To appear (plan): End of 2015
References

- R. B. DANNENBERG AND M. GOTO, Music structure analysis from acoustic signals, in Handbook of Signal Processing in Acoustics, D. Hewick, S.

References


References


References