

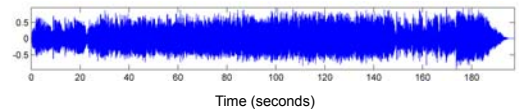
Lecture  
Music Processing

### Music Structure Analysis

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

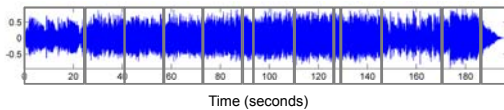
### Music Structure Analysis

Example: Zager & Evans "In The Year 2525"



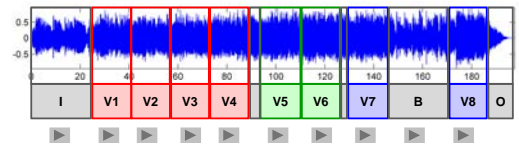
### Music Structure Analysis

Example: Zager & Evans "In The Year 2525"



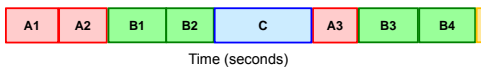
### Music Structure Analysis

Example: Zager & Evans "In The Year 2525"



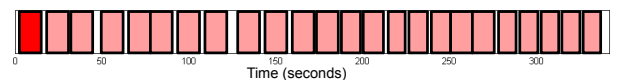
### Music Structure Analysis

Example: Brahms Hungarian Dance No. 5 (Ormandy)



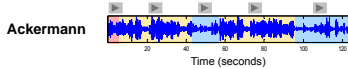
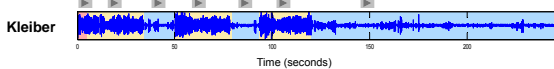
### Music Structure Analysis

Example: Folk Song Field Recording  
(Nederlandse Liederbank)



## Music Structure Analysis

**Example:** Weber, Song (No. 4) from "Der Freischütz"



## 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.

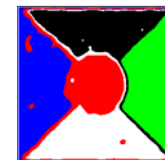
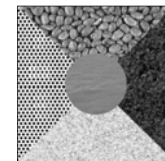
- **Homogeneity:** Consistency in tempo, instrumentation, key, ...
- **Novelty:** Sudden changes, surprising elements ...
- **Repetition:** Repeating themes, motives, rhythmic patterns, ...

## Music Structure Analysis

Novelty

Homogeneity

Repetition



## Overview

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

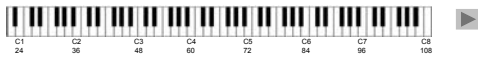
## Feature Representation

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

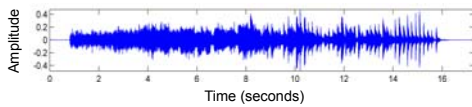
- Timbre / Instrumentation
- Tempo / Rhythm
- Pitch / Harmony

## Feature Representation

**Example:** Chromatic scale

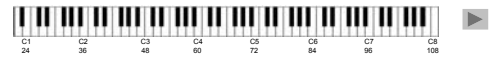


**Waveform**

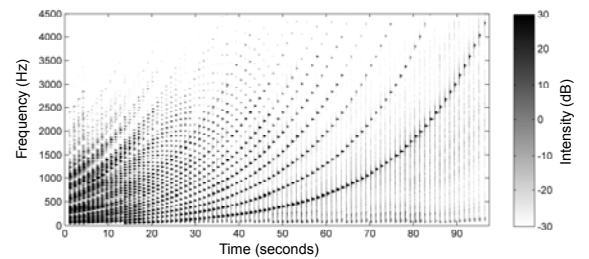


## Feature Representation

**Example:** Chromatic scale

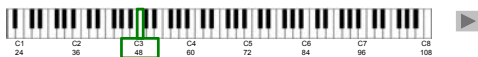


**Spectrogram**

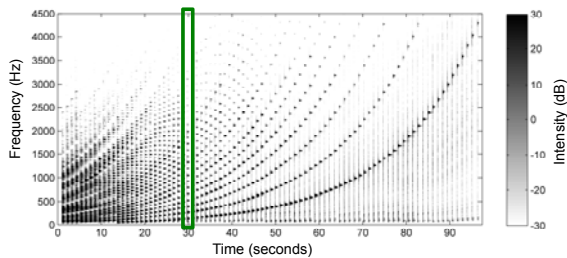


## Feature Representation

**Example:** Chromatic scale

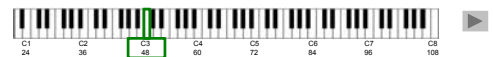


**Spectrogram**

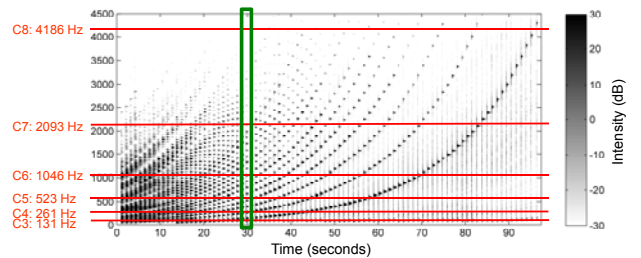


## Feature Representation

**Example:** Chromatic scale

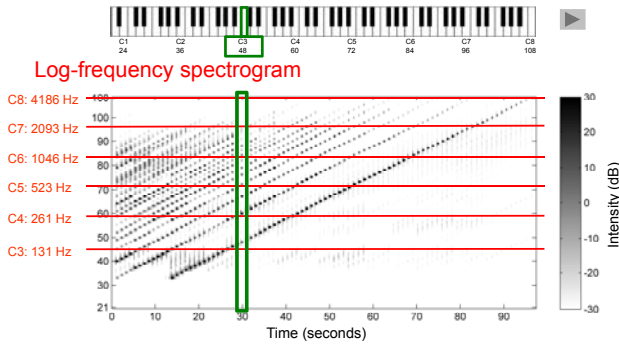


**Spectrogram**



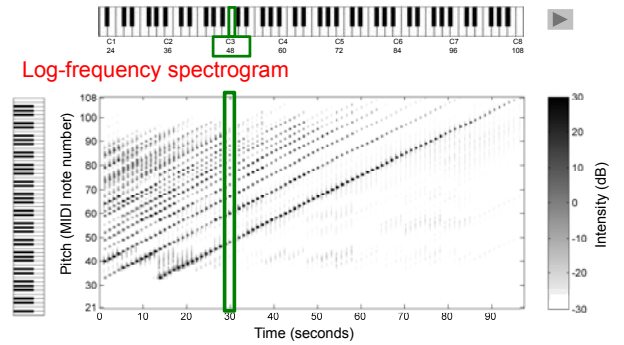
## Feature Representation

Example: Chromatic scale



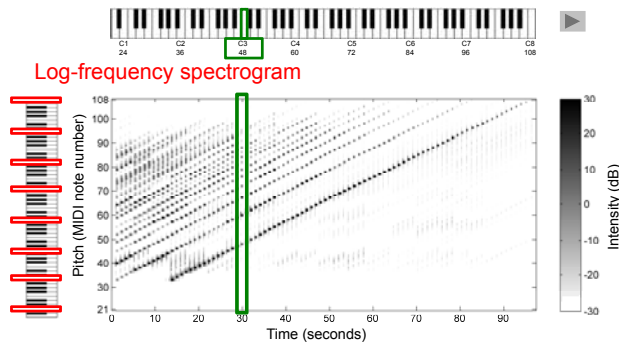
## Feature Representation

Example: Chromatic scale



## Feature Representation

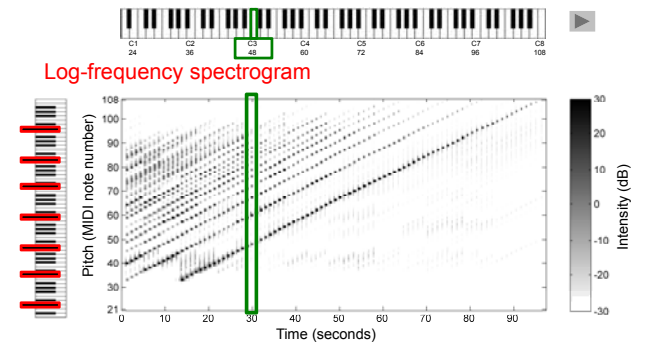
Example: Chromatic scale



Chroma C

## Feature Representation

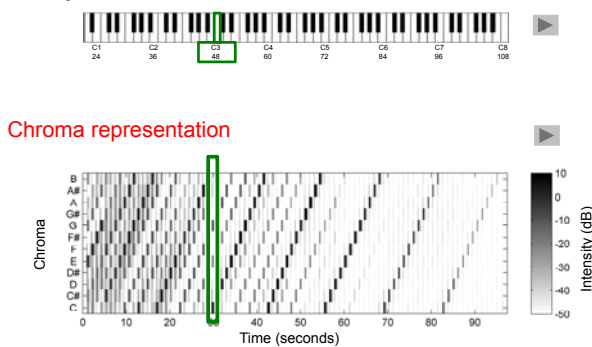
Example: Chromatic scale



Chroma C#

## Feature Representation

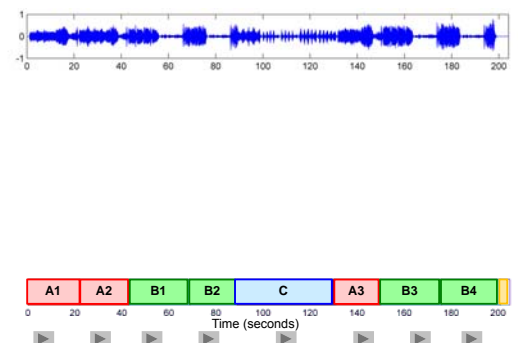
Example: Chromatic scale



Chroma representation

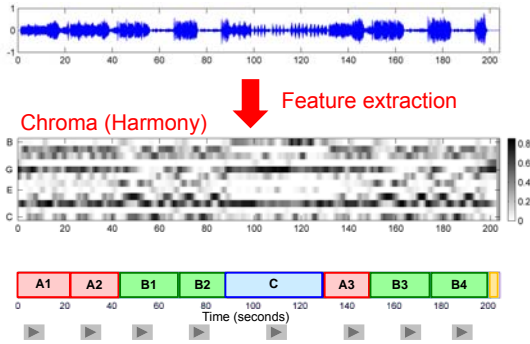
## Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



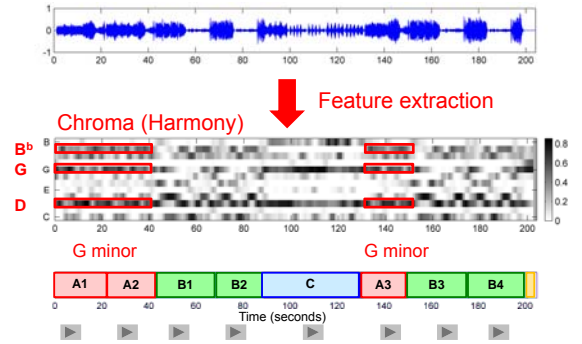
## Feature Representation

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



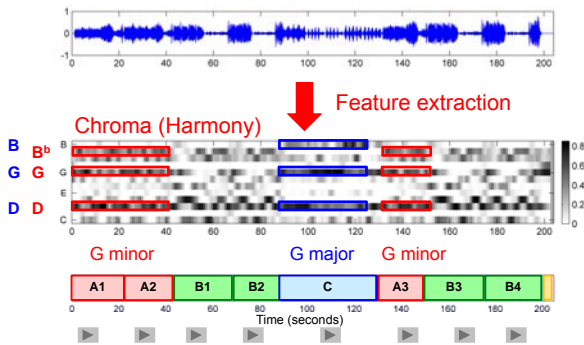
## Feature Representation

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



## Feature Representation

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



## Overview

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

## Self-Similarity Matrix (SSM)

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

→ Quadratic self-similarity matrix

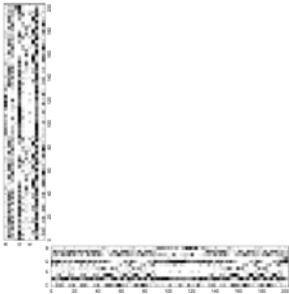
## Self-Similarity Matrix (SSM)

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



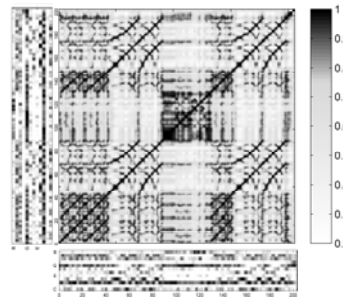
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



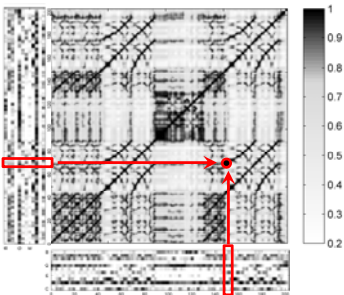
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



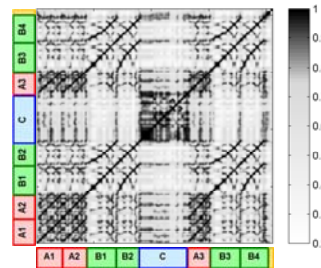
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



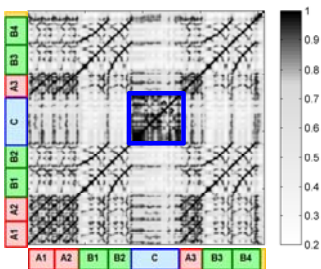
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



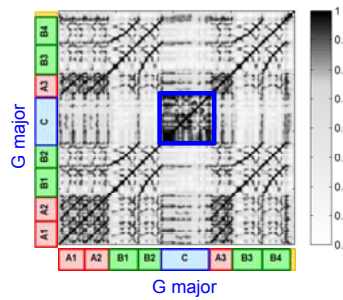
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



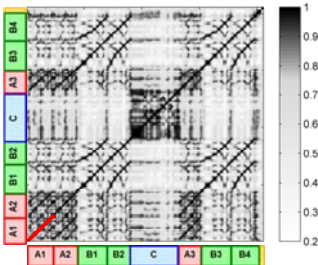
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



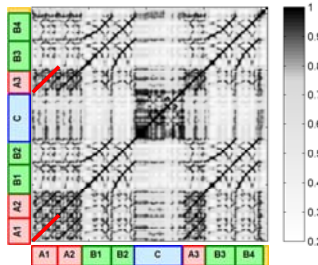
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



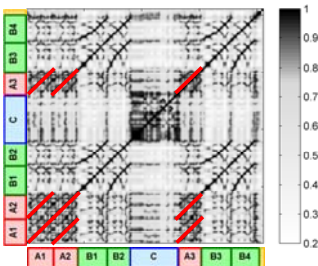
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



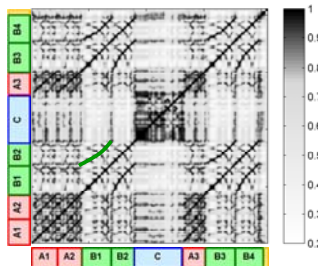
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



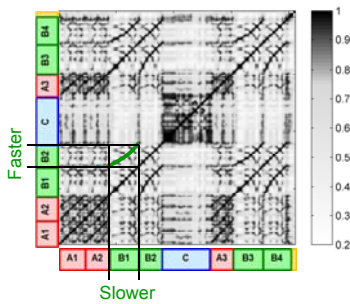
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



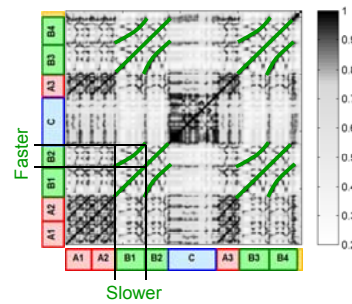
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



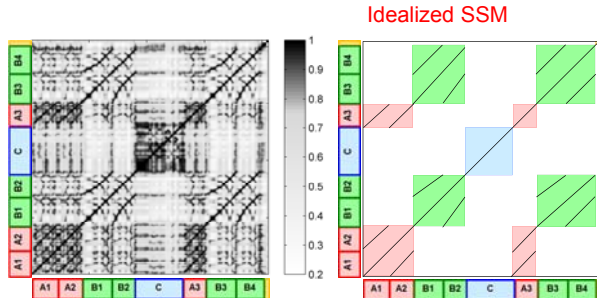
### Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



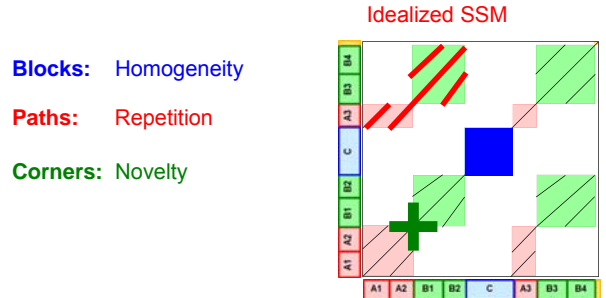
## Self-Similarity Matrix (SSM)

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



## Self-Similarity Matrix (SSM)

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

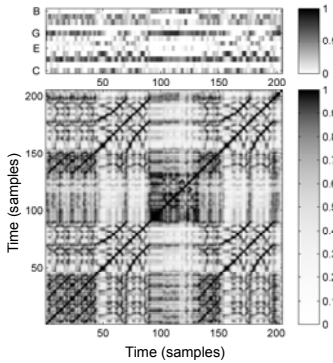


**Blocks:** Homogeneity

**Paths:** Repetition

**Corners:** Novelty

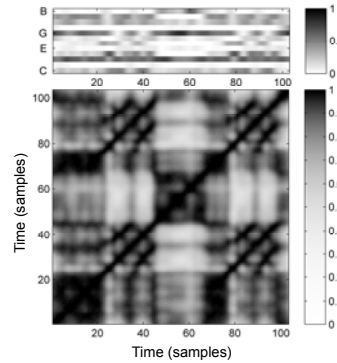
## SSM Enhancement



**Block Enhancement**

- Feature smoothing
- Coarsening

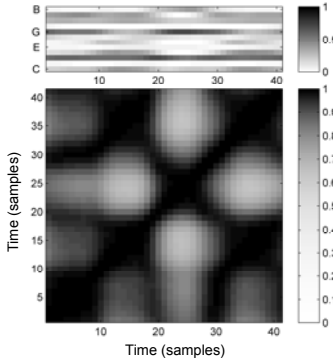
## SSM Enhancement



**Block Enhancement**

- Feature smoothing
- Coarsening

## SSM Enhancement



**Block Enhancement**

- Feature smoothing
- Coarsening

## SSM Enhancement

**Challenge:** Presence of musical variations

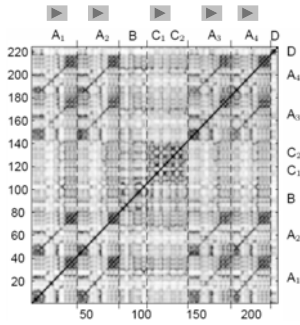
- Fragmented paths and gaps
- Paths of poor quality
- Regions of constant (low) cost
- Curved paths

**Idea:** Enhancement of path structure

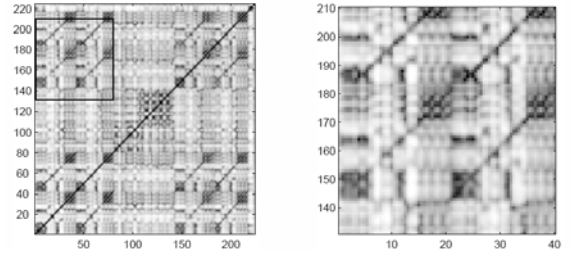


## SSM Enhancement

Shostakovich Waltz 2, Jazz Suite No. 2 (Chailly)

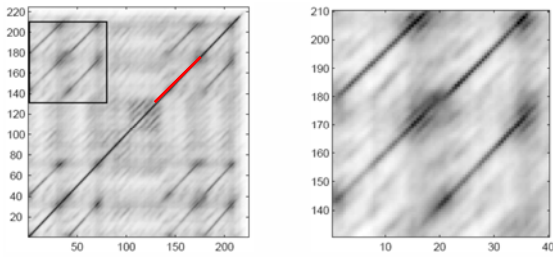


## SSM Enhancement



Cost matrix  $C$

## SSM Enhancement



Enhanced cost matrix  $C_L$   
Filtering along main diagonal

## SSM Enhancement

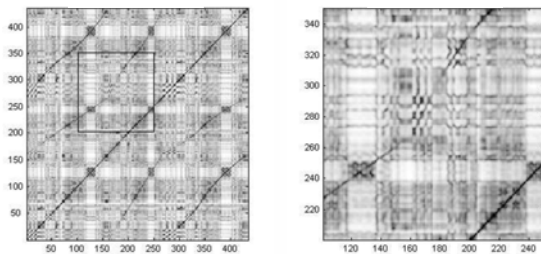
Idea: Usage of contextual information (Foote 1999)

$$C_L(n, m) := \frac{1}{L} \sum_{\ell=0}^{L-1} c(v_{n+\ell}, v_{m+\ell})$$

- Comparison of entire sequences
- $L$  = length of sequences
- $C_L$  = enhanced cost matrix

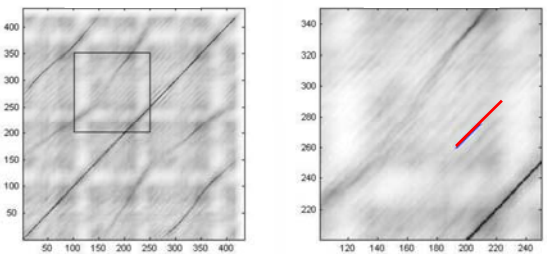
↪ smoothing effect

## SSM Enhancement



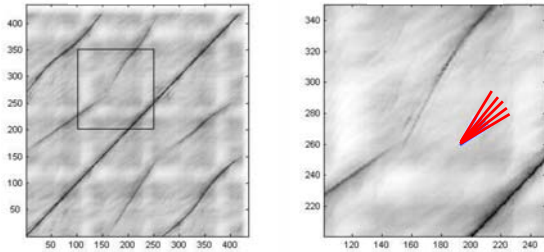
Cost matrix  $C$

## SSM Enhancement



Cost matrix  $C_L$  with  $L = 20$   
Filtering along main diagonal

## SSM Enhancement



Cost matrix  $C_L^{\min}$  with  $L = 20$

Filtering along 8 different directions and minimizing

## SSM Enhancement

Idea: Smoothing along various directions and minimizing over all directions

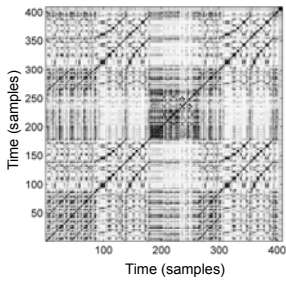
$$C_L^{\min}(n, m) := \min_k C_L^{\text{slope}_k}(n, m)$$

- $\text{slope}_k$  =  $k$ th direction of smoothing
- $C_L^{\text{slope}_k}$  = enhanced cost matrix w.r.t.  $\text{slope}_k$
- Usage of eight slope values

↔ tempo changes of -30 to +40 percent

## SSM Enhancement

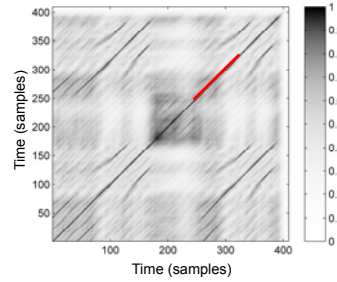
Path Enhancement



## SSM Enhancement

Path Enhancement

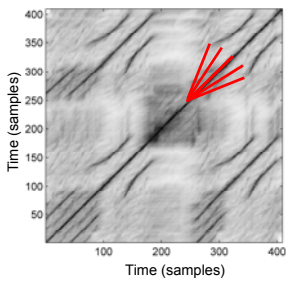
- Diagonal smoothing



## SSM Enhancement

Path Enhancement

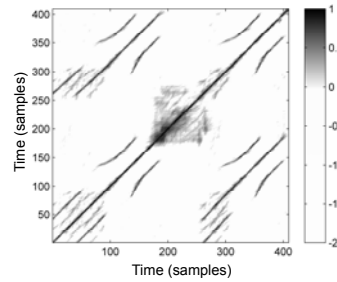
- Diagonal smoothing
- Multiple filtering



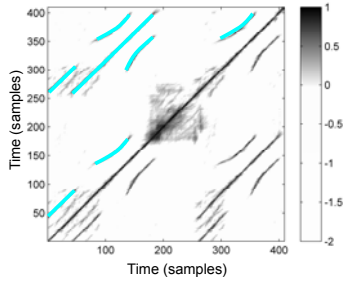
## SSM Enhancement

Path Enhancement

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



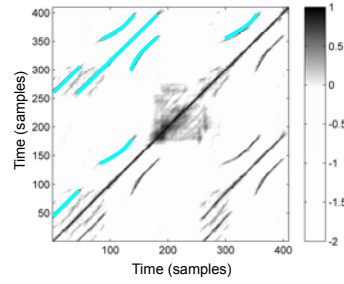
## SSM Enhancement



### Further Processing

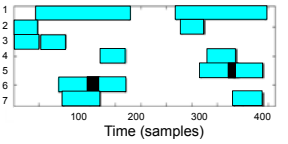
- Path extraction

## SSM Enhancement

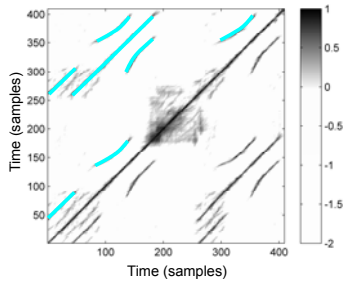


### Further Processing

- Path extraction
- Pairwise relations

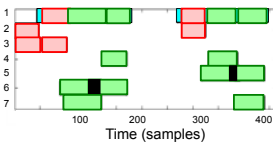


## SSM Enhancement

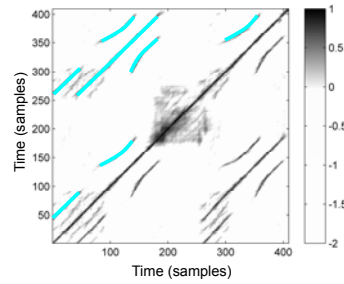


### Further Processing

- Path extraction
- Pairwise relations
- Grouping (transitivity)

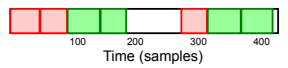


## SSM Enhancement



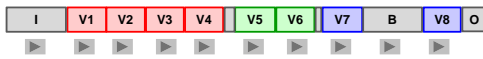
### Further Processing

- Path extraction
- Pairwise relations
- Grouping (transitivity)



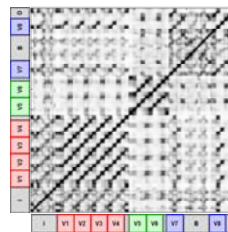
## SSM Enhancement

Example: Zager & Evans "In The Year 2525"



## SSM Enhancement

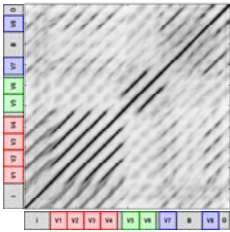
Example: Zager & Evans "In The Year 2525"



## SSM Enhancement

**Example:** Zager & Evans "In The Year 2525"

Missing relations because of transposed sections

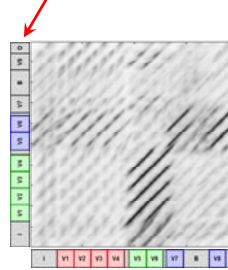


## SSM Enhancement

**Example:** Zager & Evans "In The Year 2525"

Idea: Cyclic shift of one of the chroma sequences

One semitone up

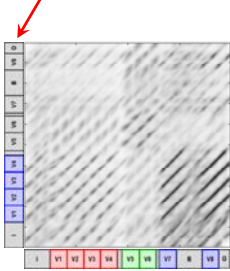


## SSM Enhancement

**Example:** Zager & Evans "In The Year 2525"

Idea: Cyclic shift of one of the chroma sequences

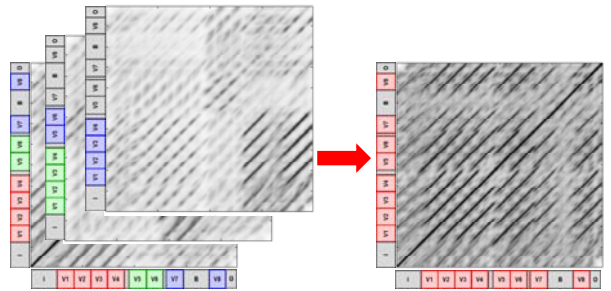
Two semitones up



## SSM Enhancement

**Example:** Zager & Evans "In The Year 2525"

Idea: Overlay & Maximize → Transposition-invariant SSM

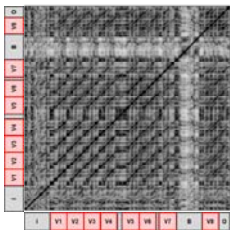


## SSM Enhancement

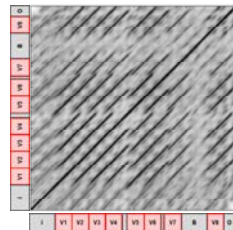
**Example:** Zager & Evans "In The Year 2525"

Note: Order of enhancement steps important!

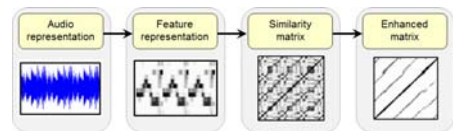
Maximization



Smoothing & Maximization



## 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
- Feature Representations
- Self-Similarity Matrices
- Audio Thumbing**
- Novelty-based Segmentation
- Converting Path to Block Structures

### Thanks:

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

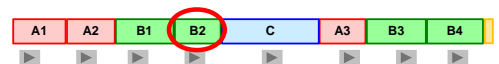
## Audio Thumbing

**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 Thumbing

### Two steps

1. Path extraction
2. Grouping

### Both steps are problematic!

- Paths of poor quality (fragmented, gaps)
- Block-like structures
- Curved paths
- Noisy relations (missing, distorted, overlapping)
- Transitivity computation difficult

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

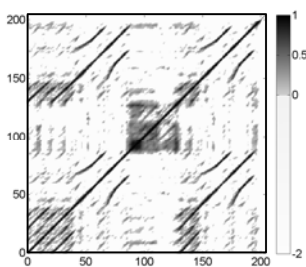
- One optimization scheme for both steps
- Stabilizing effect
- Efficient

## Audio Thumbing

Main idea: Do both path extraction and grouping jointly

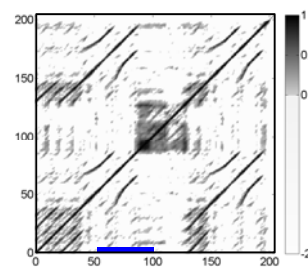
- For each audio **segment** we define a **fitness** value
- This fitness value expresses "how well" the segment explains the entire audio recording
- The segment with the highest fitness value is considered to be the **thumbnail**
- As main technical concept we introduce the notion of a **path family**

## Fitness Measure



Enhanced SSM

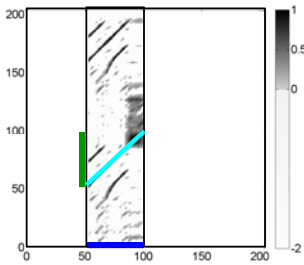
## Fitness Measure



Path over segment

- Consider a fixed **segment**

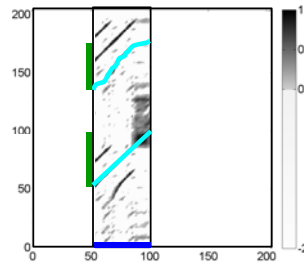
## Fitness Measure



### Path over segment

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

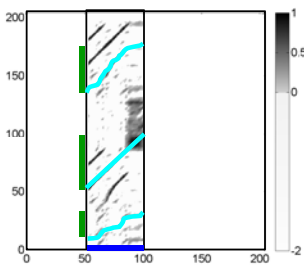
## Fitness Measure



### Path over segment

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

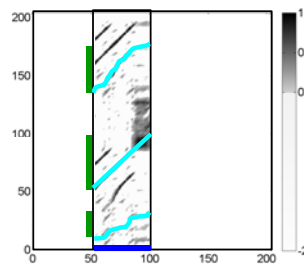
## Fitness Measure



### 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

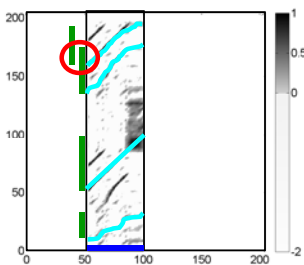
## Fitness Measure



### Path family

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

## Fitness Measure

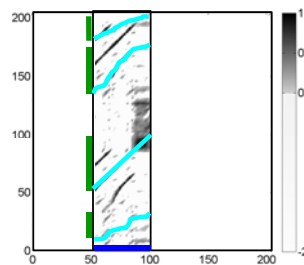


### Path family

- 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 **not** a path family!

## Fitness Measure

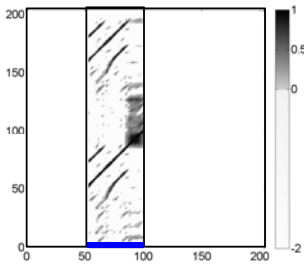


### Path family

- 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)

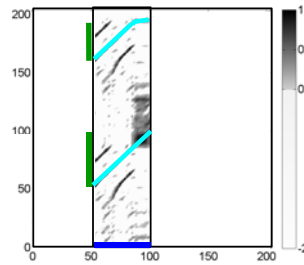
## Fitness Measure



### Optimal path family

- Consider a fixed **segment**

## Fitness Measure

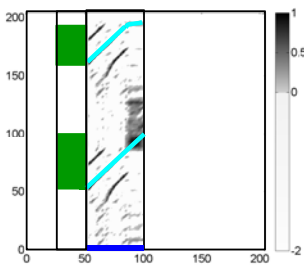


### Optimal path family

- Consider a fixed **segment**
- Consider over the **segment** the **optimal path family**, i.e., the path family having maximal overall score.
- Call this value:  $\text{Score}(\text{segment})$

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

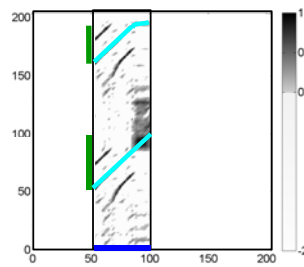
## Fitness Measure



### Optimal path family

- Consider a fixed **segment**
- Consider over the **segment** the **optimal path family**, i.e., the path family having maximal overall score.
- Call this value:  $\text{Score}(\text{segment})$
- Furthermore consider the amount covered by the **induced segments**.
- Call this value:  $\text{Coverage}(\text{segment})$

## Fitness Measure

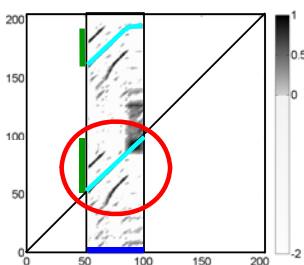


### Fitness

- Consider a fixed **segment**

$P := \text{Score}(\text{segment})$   
 $R := \text{Coverage}(\text{segment})$

## Fitness Measure

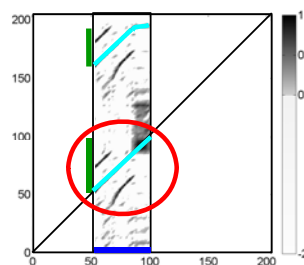


### Fitness

- Consider a fixed **segment**
- Self-explanation are trivial!**

$P := \text{Score}(\text{segment})$   
 $R := \text{Coverage}(\text{segment})$

## Fitness Measure

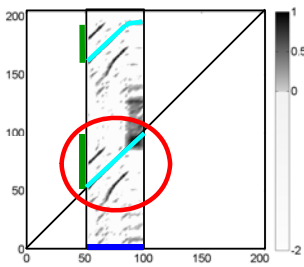


### Fitness

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

$P := \text{Score}(\text{segment}) - \text{length}(\text{segment})$   
 $R := \text{Coverage}(\text{segment}) - \text{length}(\text{segment})$

## Fitness Measure



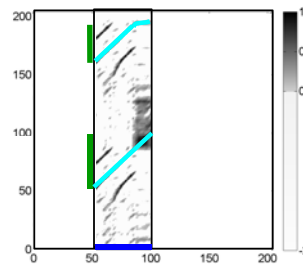
### Fitness

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

$$P := \text{Normalize}(\text{Score}(\text{segment}) - \text{length}(\text{segment})) \in [0,1]$$

$$R := \text{Normalize}(\text{Coverage}(\text{segment}) - \text{length}(\text{segment})) \in [0,1]$$

## Fitness Measure



### Fitness

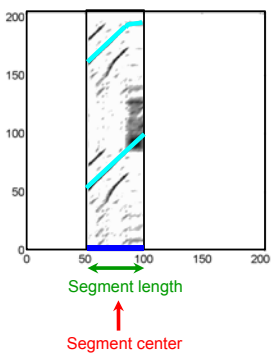
- Consider a fixed **segment**

$$F := \frac{2 \cdot P \cdot R}{P + R}$$

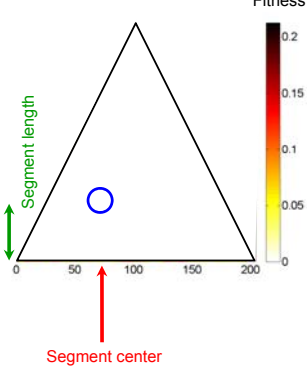
$$P := \text{Normalize}(\text{Score}(\text{segment}) - \text{length}(\text{segment})) \in [0,1]$$

$$R := \text{Normalize}(\text{Coverage}(\text{segment}) - \text{length}(\text{segment})) \in [0,1]$$

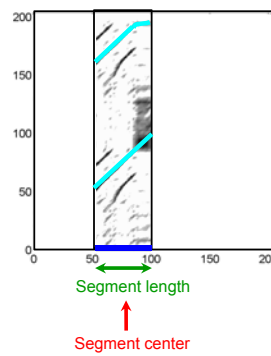
## Thumbnail



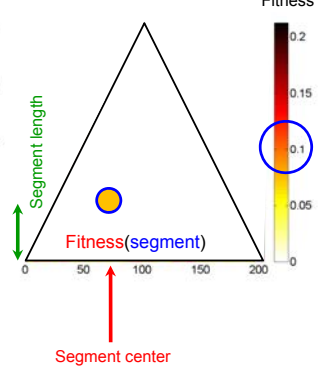
### Fitness Scape Plot



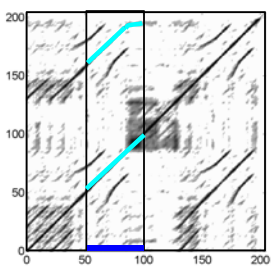
## Thumbnail



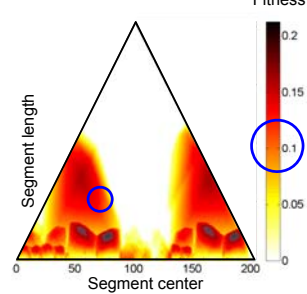
### Fitness Scape Plot



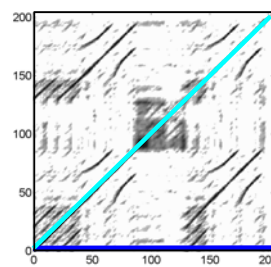
## Thumbnail



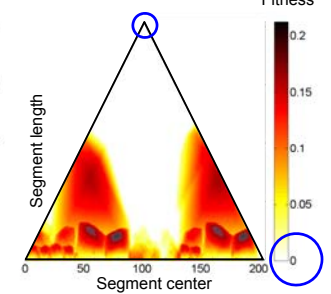
### Fitness Scape Plot



## Thumbnail



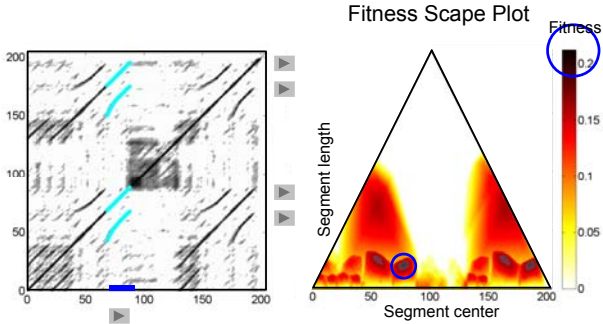
### Fitness Scape Plot



Note: Self-explanations are ignored → fitness is zero

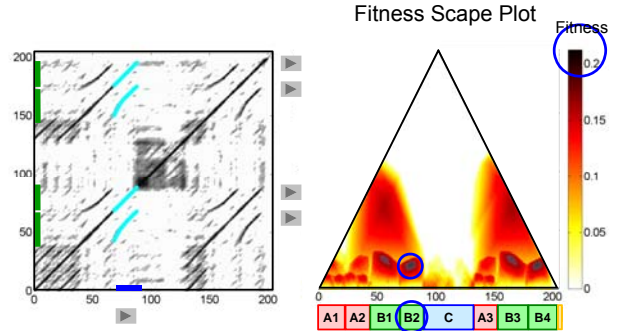


### Thumbnail



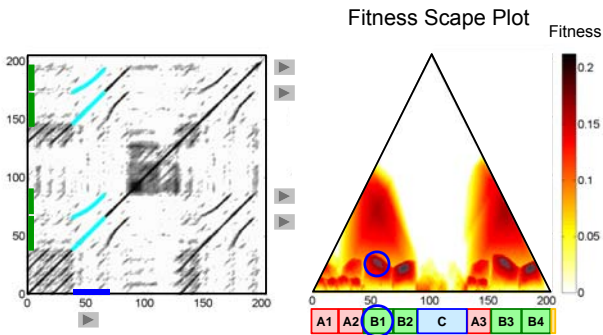
Thumbnail := segment having the highest fitness

### Thumbnail



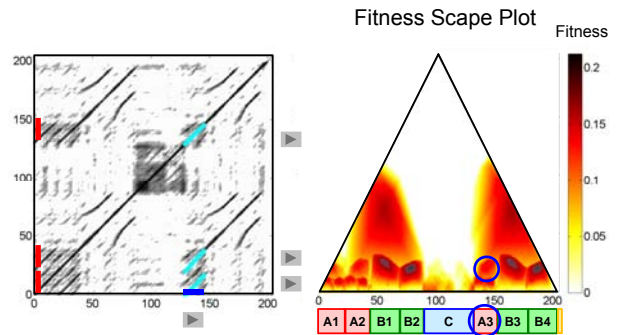
Example: Brahms Hungarian Dance No. 5 (Ormandy)

### Thumbnail



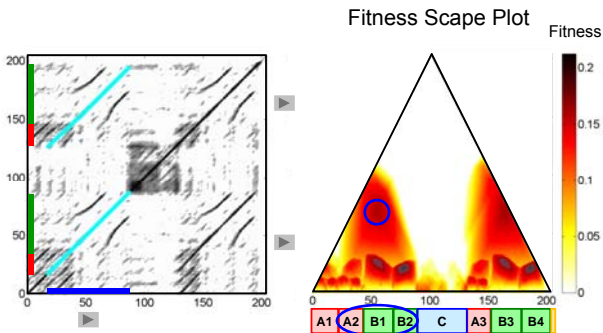
Example: Brahms Hungarian Dance No. 5 (Ormandy)

### Thumbnail



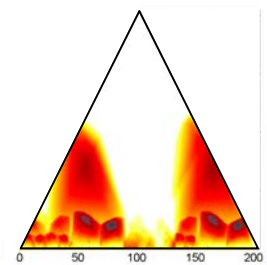
Example: Brahms Hungarian Dance No. 5 (Ormandy)

### Thumbnail



Example: Brahms Hungarian Dance No. 5 (Ormandy)

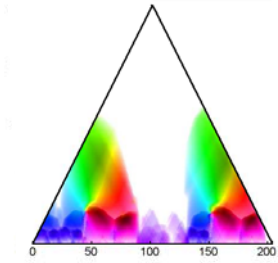
### Scape Plot



Example: Brahms Hungarian Dance No. 5 (Ormandy)

## Scape Plot

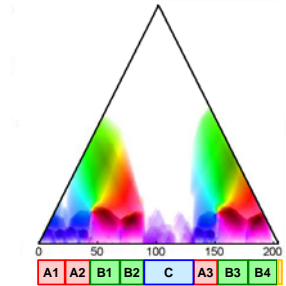
Coloring according to clustering result (grouping)



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

## Scape Plot

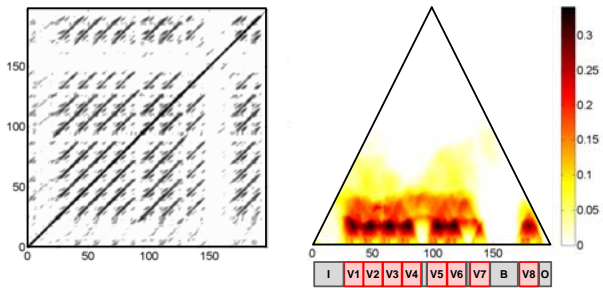
Coloring according to clustering result (grouping)



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

## Thumbnail

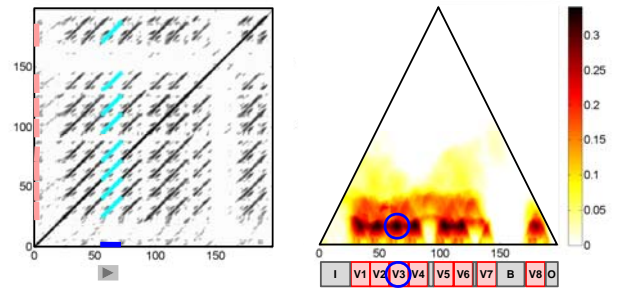
Fitness Scape Plot



**Example:** Zager & Evans "In The Year 2525"

## Thumbnail

Fitness Scape Plot



**Example:** Zager & Evans "In The Year 2525"

## Overview

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

**Thanks:**

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

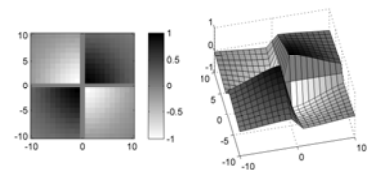
## Novelty-based Segmentation

**General goals:**

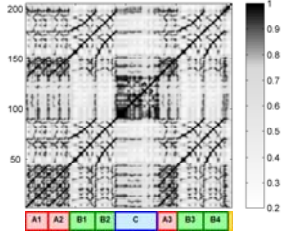
- Find instances where musical changes occur.
- Find transition between subsequent musical parts.

**Idea (Foote):**

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



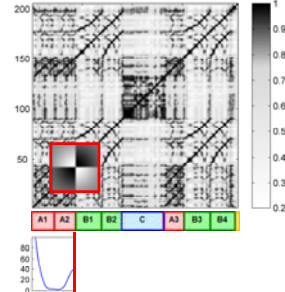
## Novelty-based Segmentation



### Idea (Foote):

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

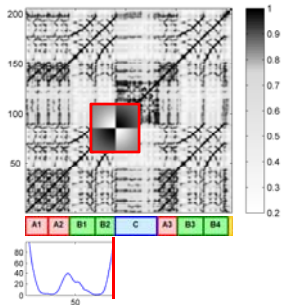
## Novelty-based Segmentation



### Idea (Foote):

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

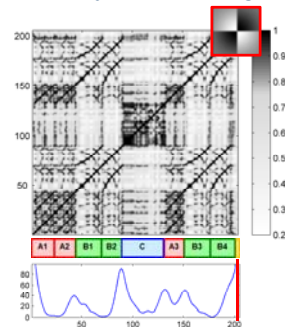
## Novelty-based Segmentation



### Idea (Foote):

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

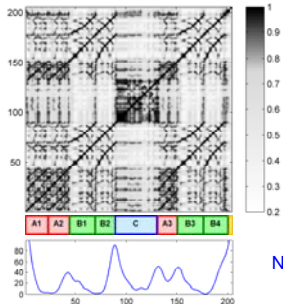
## Novelty-based Segmentation



### Idea (Foote):

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

## Novelty-based Segmentation



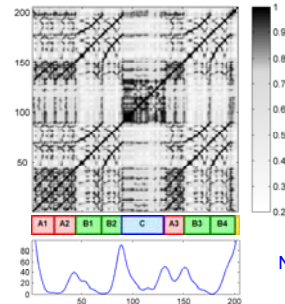
### Idea (Foote):

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



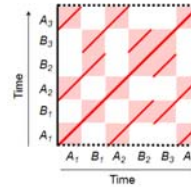
## Novelty-based Segmentation

### Idea:

- Find instances where **structural** changes occur.
- Combine **global** and **local** aspects within a unifying framework

### Structure features

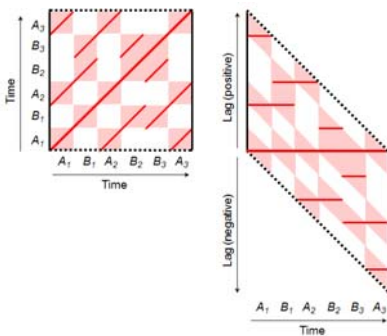
## Novelty-based Segmentation



### Structure features

- Enhanced SSM

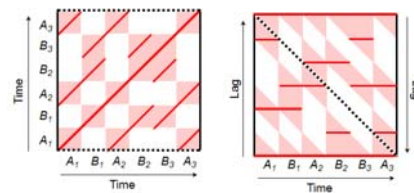
## Novelty-based Segmentation



### Structure features

- Enhanced SSM
- Time-lag SSM

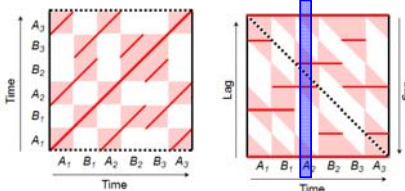
## Novelty-based Segmentation



### Structure features

- Enhanced SSM
- Time-lag SSM
- Cyclic time-lag SSM

## Novelty-based Segmentation

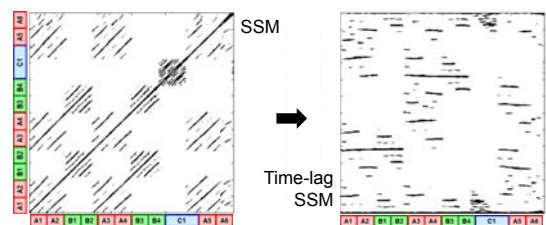


### Structure features

- Enhanced SSM
- Time-lag SSM
- Cyclic time-lag SSM
- Columns as **features**

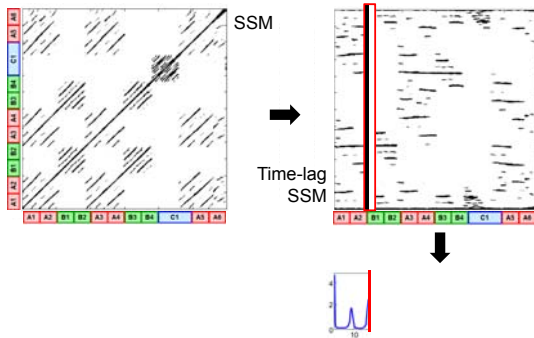
## Novelty-based Segmentation

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



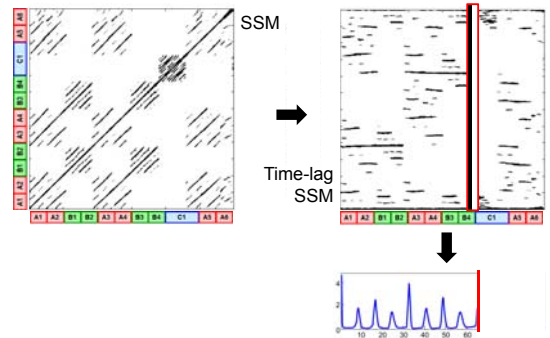
## Novelty-based Segmentation

Example: Chopin Mazurka Op. 24, No. 1



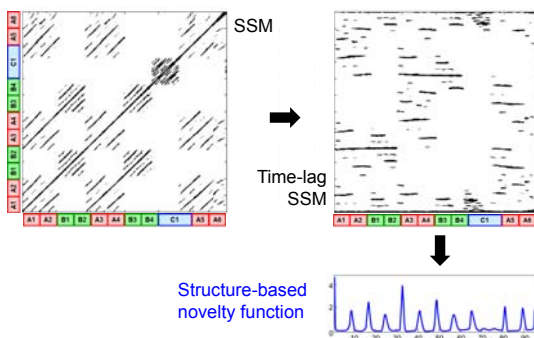
## Novelty-based Segmentation

Example: Chopin Mazurka Op. 24, No. 1



## Novelty-based Segmentation

Example: Chopin Mazurka Op. 24, No. 1



## Overview

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

Thanks:

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

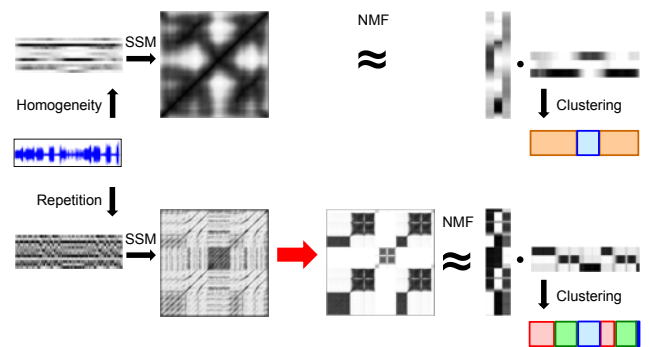
## Converting Path to Block Structures

### Motivation

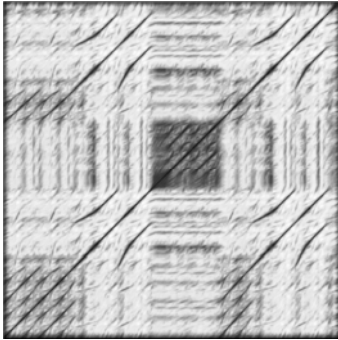
- Perform joint analysis using repetitive as well as homogeneous aspects
- Make homogeneity-based methods applicable to repetition-based analysis

## Converting Path to Block Structures

### Motivation



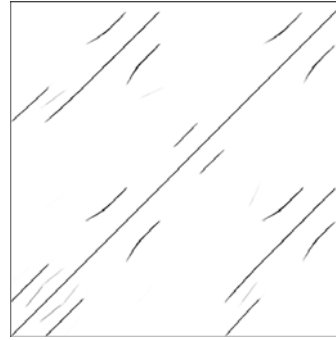
## Converting Path to Block Structures



### Procedure

- Enhanced SSM

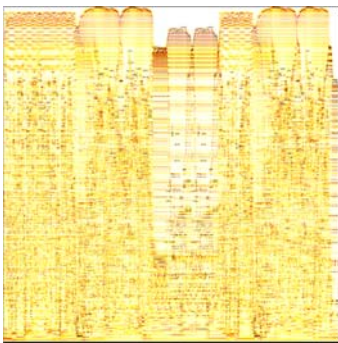
## Converting Path to Block Structures



### Procedure

- Enhanced SSM
- Thresholding & image processing

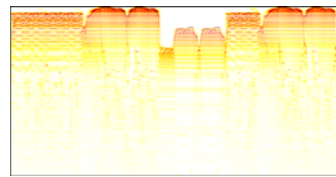
## Converting Path to Block Structures



### Procedure

- Enhanced SSM
- Thresholding & image processing
- Eigenvalue decomposition

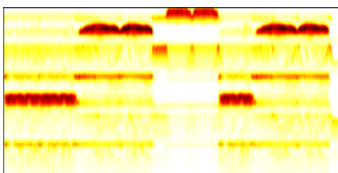
## Converting Path to Block Structures



### Procedure

- Enhanced SSM
- Thresholding & image processing
- Eigenvalue decomposition
- Weigthing

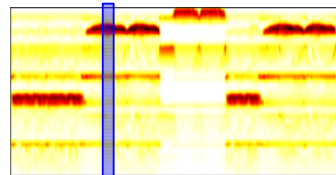
## Converting Path to Block Structures



### Procedure

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

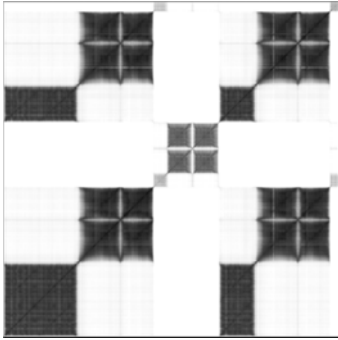
## Converting Path to Block Structures



### Procedure

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

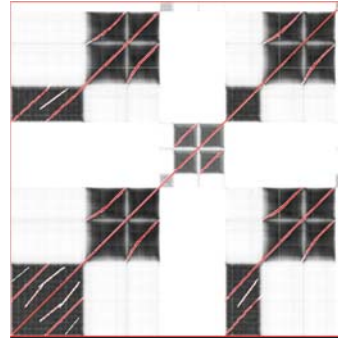
## Converting Path to Block Structures



### Procedure

- Enhanced SSM
- Thresholding & image processing
- Eigenvalue decomposition
- Weighing
- Clustering & smoothing
- **Columns as features**
- SSM from these features

## Converting Path to Block Structures



### Procedure

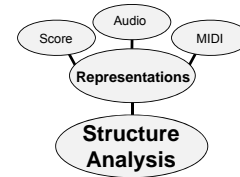
- Enhanced SSM
- Thresholding & image processing
- Eigenvalue decomposition
- Weighing
- Clustering & smoothing
- **Columns as features**
- SSM from these features

**Final matrix show paths as blocks**

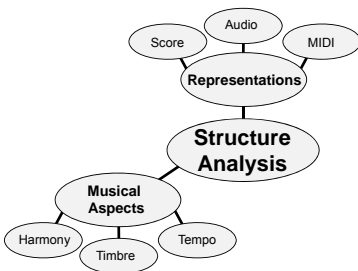
## Conclusions

Structure Analysis

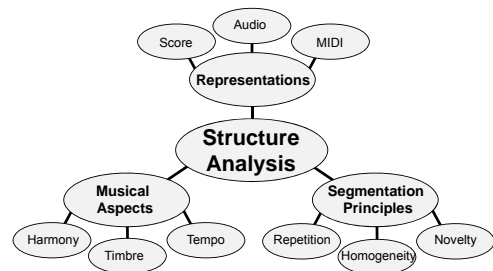
## Conclusions



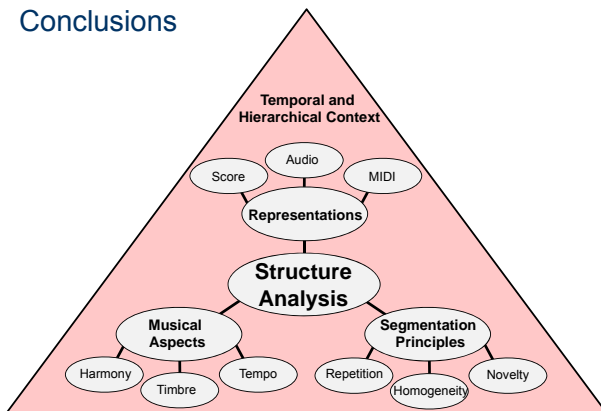
## Conclusions



## Conclusions

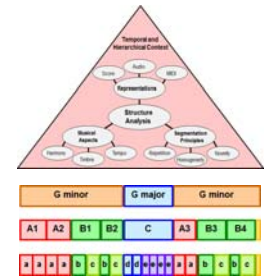


## Conclusions



## Conclusions

- Combined Approaches
- Hierarchical Approaches
- Evaluation
- Explaining Structure



- MIREX
- SALAMI-Project
- Smith, Chew

## References

- W. CHAI AND B. VERCOE, Music thumbnailing via structural analysis, in Proceedings of the ACM International Conference on Multimedia, Berkeley, CA, USA, 2003, pp. 223–226.
- M. COOPER AND J. FOOTE, Automatic music summarization via similarity analysis, in Proceedings of the International Conference on Music Information Retrieval (ISMIR), Paris, France, 2002, pp. 81–85.
- R. B. DANNENBERG AND M. GOTO, Music structure analysis from acoustic signals, in Handbook of Signal Processing in Acoustics, D. Havelock, S.
- J. FOOTE, Visualizing music and audio using self-similarity, in Proceedings of the ACM International Conference on Multimedia, Orlando, FL, USA, 1999, pp. 77–80.
- J. FOOTE, Automatic audio segmentation using a measure of audio novelty, in Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), New York, NY, USA, 2000, pp. 432–435.
- M. GOTO, A chorus section detection method for musical audio signals and its application to a music listening station, IEEE Transactions on Audio, Speech and Language Processing, 14 (2006), pp. 1783–1794.
- H. GROHGANZ, M. CLAUSEN, N. JIANG, AND M. MÜLLER, Converting path structures into block structures using eigenvalue decompositions of self-similarity matrices, in Proceedings of the 14th International Conference on Music Information Retrieval (ISMIR), Curitiba, Brazil, 2013, pp. 209–214.
- K. JENSEN, Multiple scale music segmentation using rhythm, timbre, and harmony, EURASIP Journal on Advances in Signal Processing, 2007 (2007).
- F. KAISER AND T. SIKORA, Music structure discovery in popular music using non-negative matrix factorization, in Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), Utrecht, The Netherlands, 2010, pp. 429–434.

## References

- M. LEVY, M. SANDLER, AND M. A. CASEY, Extraction of high-level musical structure from audio data and its application to thumbnail generation, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Toulouse, France, 2006, pp. 13–16.
- H. LUKASHEVICH, Towards quantitative measures of evaluating song segmentation, in Proceedings of the International Conference on Music Information Retrieval (ISMIR), Philadelphia, USA, 2008, pp. 375–380.
- M. MÜLLER AND M. CLAUSEN, Transposition-invariant self-similarity matrices, in Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR), Vienna, Austria, 2007, pp. 47–50.
- M. MÜLLER AND N. JIANG, A scape plot representation for visualizing repetitive structures of music recordings, in Proceedings of the 13th International Conference on Music Information Retrieval (ISMIR), Porto, Portugal, 2012, pp. 97–102.
- M. MÜLLER, N. JIANG, AND H. GROHGANZ, SM Toolbox: MATLAB implementations for computing and enhancing similarity matrices, in Proceedings of the 53rd AES Conference on Semantic Audio, London, GB, 2014.
- M. MÜLLER, N. JIANG, AND P. GROSCHE, A robust fitness measure for capturing repetitions in music recordings with applications to audio thumbnailing, IEEE Transactions on Audio, Speech & Language Processing, 21 (2013), pp. 531–543.
- M. MÜLLER AND F. KURTH, Enhancing similarity matrices for music audio analysis, in Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toulouse, France, 2006, pp. 437–440.
- M. MÜLLER AND F. KURTH, Towards structural analysis of audio recordings in the presence of musical variations, EURASIP Journal on Advances in Signal Processing, 2007 (2007).

## References

- J. PAULUS AND A. P. KLAPURI, Music structure analysis using a probabilistic fitness measure and a greedy search algorithm, IEEE Transactions on Audio, Speech, and Language Processing, 17 (2009), pp. 1159–1170.
- J. PAULUS, M. MÜLLER, AND A. P. KLAPURI, Audio-based music structure analysis, in Proceedings of the 11th International Conference on Music Information Retrieval (ISMIR), Utrecht, The Netherlands, 2010, pp. 625–636.
- G. PEETERS, Deriving musical structure from signal analysis for music audio summary generation: “sequence” and “state” approach, in Computer Music Modeling and Retrieval, vol. 2771 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2004, pp. 143–166.
- G. PEETERS, Sequence representation of music structure using higher-order similarity matrix and maximum-likelihood approach, in Proceedings of the International Conference on Music Information Retrieval (ISMIR), Vienna, Austria, 2007, pp. 35–40.
- C. RHODES AND M. A. CASEY, Algorithms for determining and labelling approximate hierarchical self-similarity, in Proceedings of the International Conference on Music Information Retrieval (ISMIR), Vienna, Austria, 2007, pp. 41–46.
- J. SERRÀ, M. MÜLLER, P. GROSCHE, AND J. L. ARCOS, Unsupervised detection of music boundaries by time series structure features, in Proceedings of the AAAI International Conference on Artificial Intelligence, Toronto, Ontario, Canada, 2012, pp. 1613–1619.
- J. B. L. SMITH, J. A. BURGOYNE, I. FUJINAGA, D. D. ROURE, AND J. S. DOWNIE, Design and creation of a large-scale database of structural annotations, in Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), Miami, FL, USA, 2011, pp. 555–560.
- J. B. L. SMITH AND E. CHEW, Using quadratic programming to estimate feature relevance in structural analyses of music, in Proceedings of the ACM International Conference on Multimedia, 2013, pp. 113–122.

## References

- M. SUNKEL, S. JANSEN, M. WAND, E. EISEMANN, H.-P. SEIDEL, Learning Line Features in 3D Geometry, in Computer Graphics Forum (Proc. Eurographics), 2011.
- D. TURNBULL, G. LANCKRIET, E. PAMPALK, AND M. GOTO, A supervised approach for detecting boundaries in music using difference features and boosting, in Proceedings of the International Conference on Music Information Retrieval (ISMIR), Vienna, Austria, 2007, pp. 51–54.
- G. TZANETAKIS AND P. COOK, Multifeature audio segmentation for browsing and annotation, in Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Platz, NY, USA, 1999, pp. 103–106.