INTERNATIONAL AUDIO LABORATORIES ERLANGEN

AUDIO LABS

Automatisierte Methoden zur Strukturierung von Musikaufnahmen

Meinard Müller

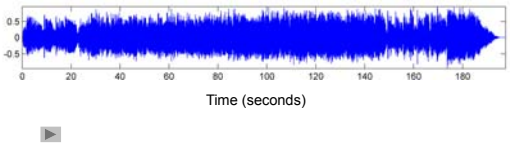
Deutsche Gesellschaft für Musikpsychologie e.V.

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

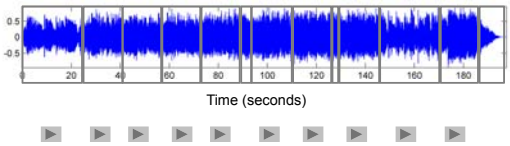
Music Structure Analysis

Example: Zager & Evans “In The Year 2525”



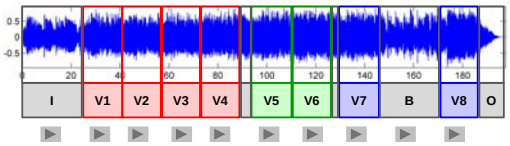
Music Structure Analysis

Example: Zager & Evans “In The Year 2525”



Music Structure Analysis

Example: Zager & Evans “In The Year 2525”

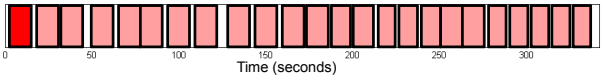


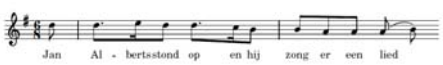
Music Structure Analysis

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Music Structure Analysis

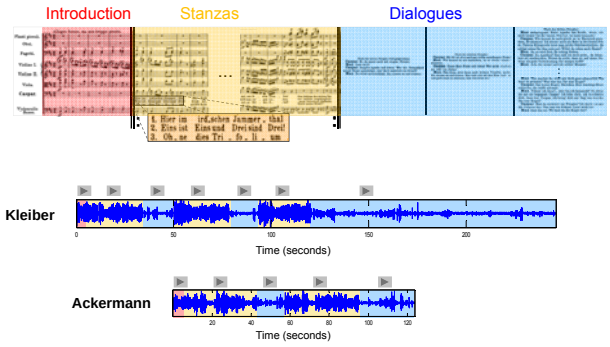
Example: Folk Song Field Recording (Nederlandse Liederenbank)





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

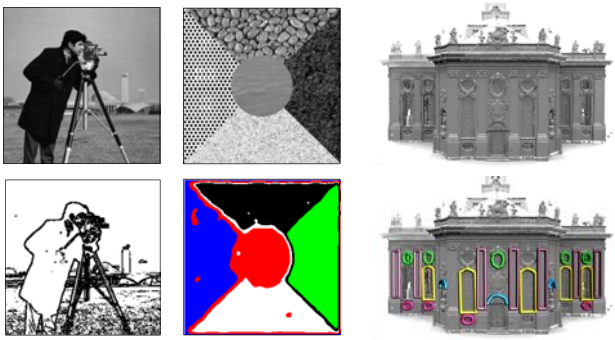
General goal: Divide an audio recording into temporal segments corresponding to musical parts and group these segments into musically meaningful categories.

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
- Music Synchronization
- Beat Tracking

Thanks:

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

Overview

- Introduction
- Feature Representations
- Self-Similarity Matrices
- Music Synchronization
- Beat Tracking

Feature Representation

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

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

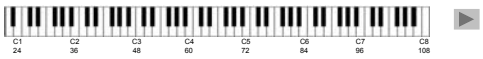
Feature Representation

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

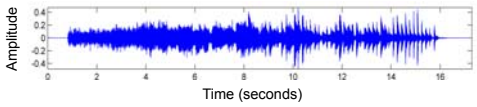
- Timbre / Instrumentation
- Tempo / Rhythm
- Pitch / Harmony

Feature Representation

Example: Chromatic scale

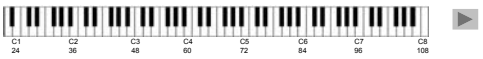


Waveform

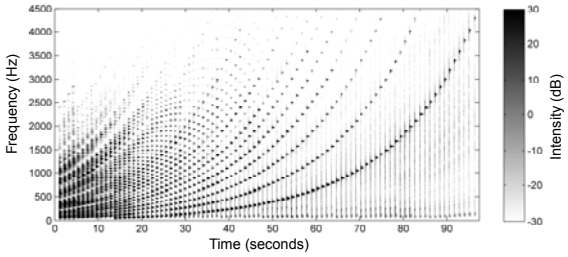


Feature Representation

Example: Chromatic scale

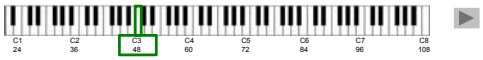


Spectrogram

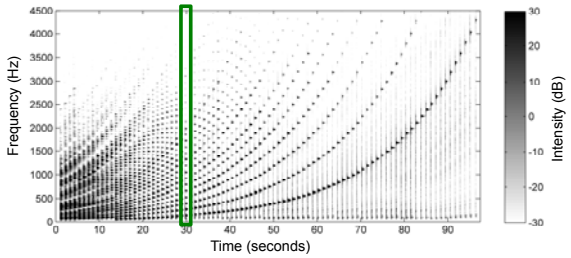


Feature Representation

Example: Chromatic scale

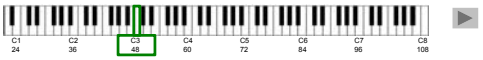


Spectrogram

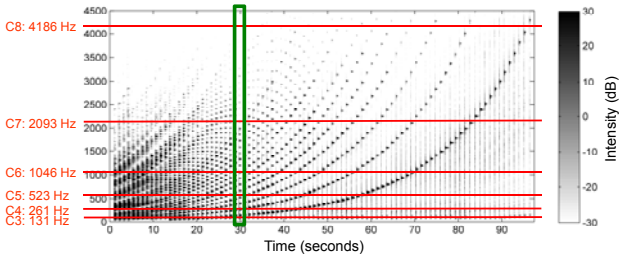


Feature Representation

Example: Chromatic scale

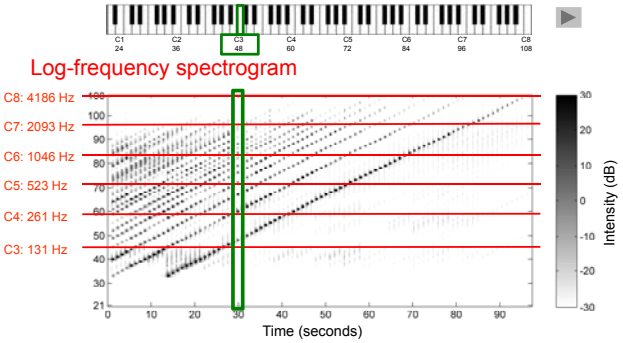


Spectrogram



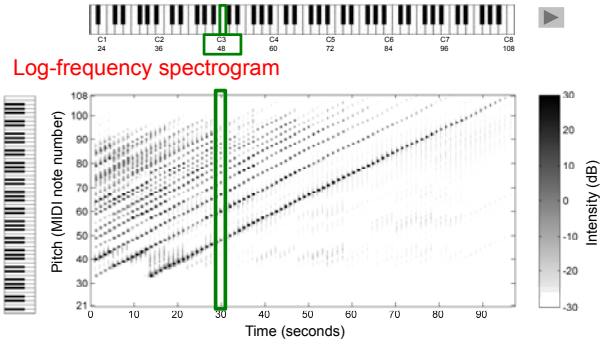
Feature Representation

Example: Chromatic scale



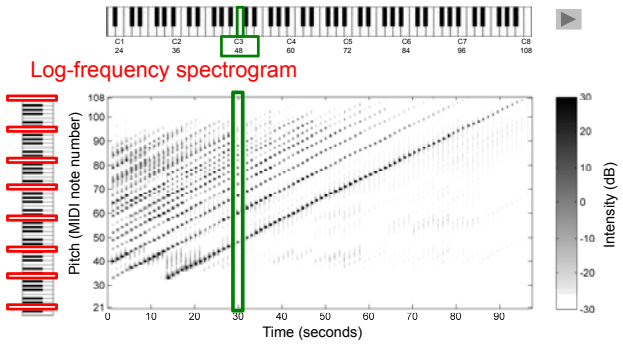
Feature Representation

Example: Chromatic scale



Feature Representation

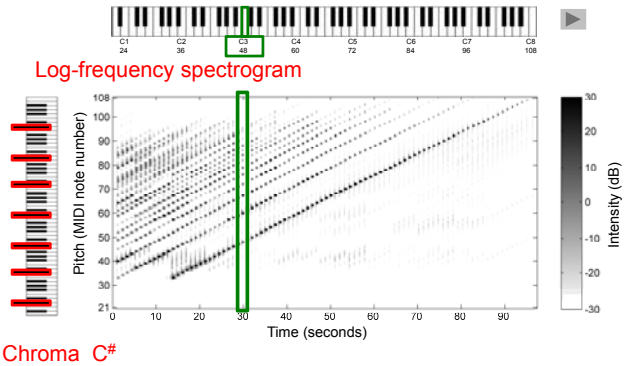
Example: Chromatic scale



Chroma C

Feature Representation

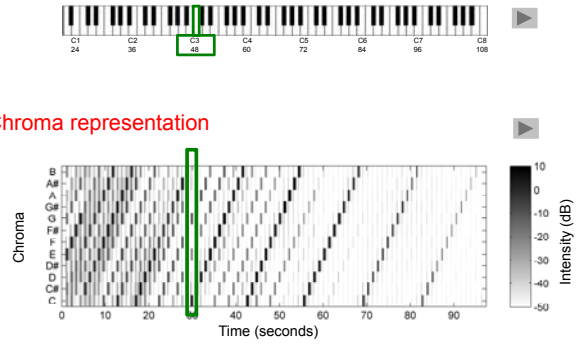
Example: Chromatic scale



Chroma C#

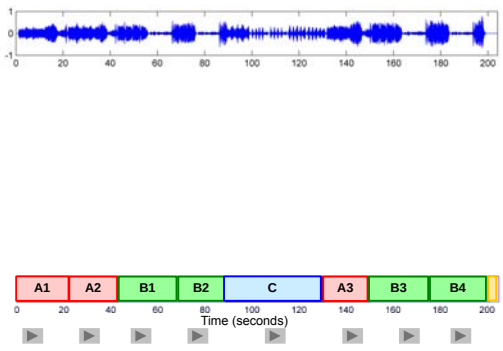
Feature Representation

Example: Chromatic scale



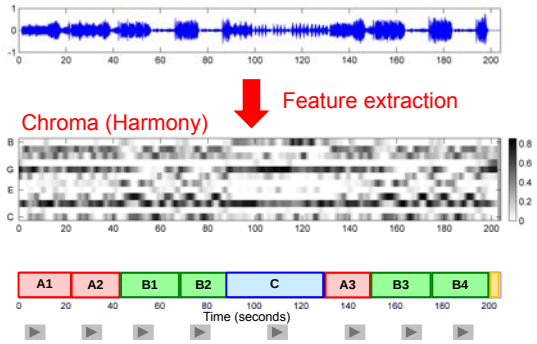
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



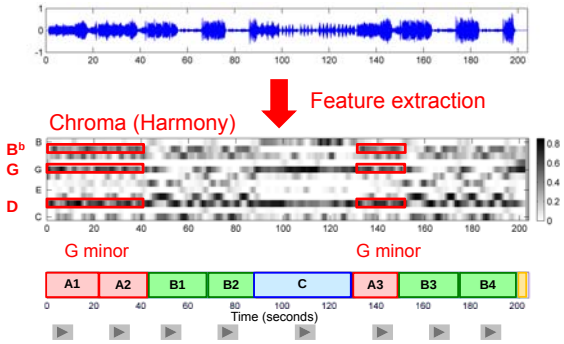
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



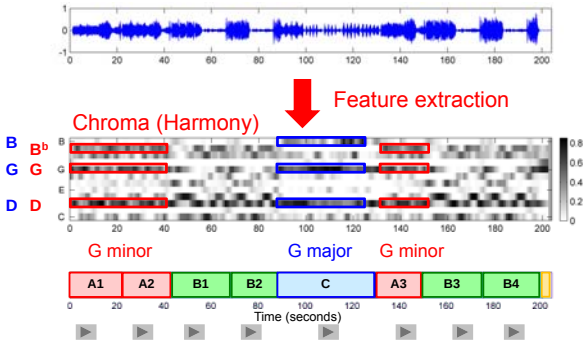
Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Feature Representation

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Overview

- Introduction
- Feature Representations
- Self-Similarity Matrices
- Music Synchronization
- Beat Tracking

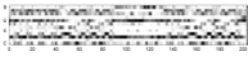
Self-Similarity Matrix (SSM)

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

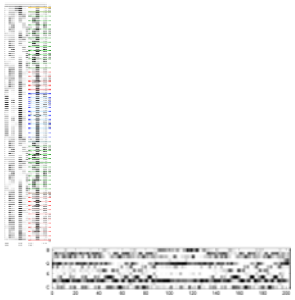
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



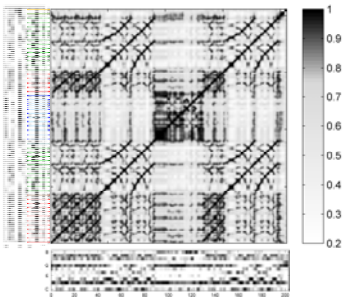
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



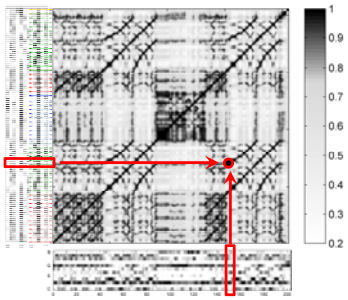
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



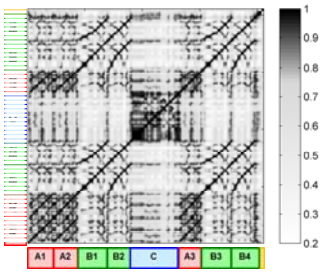
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



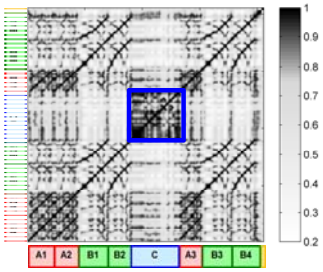
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



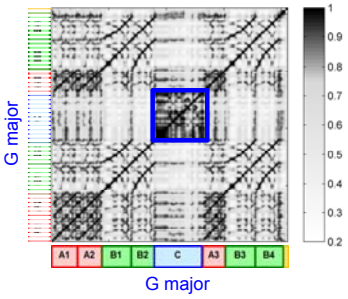
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



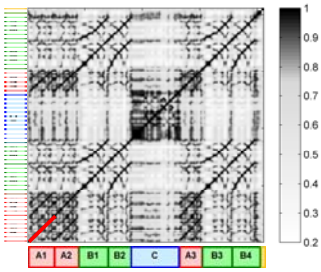
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



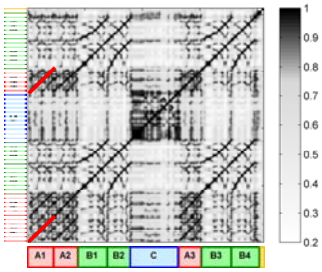
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



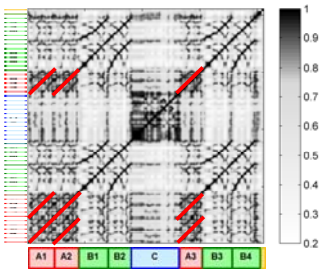
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



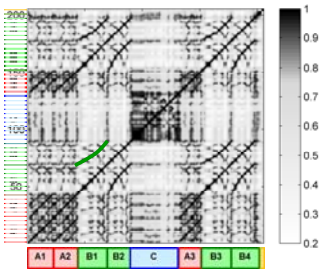
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



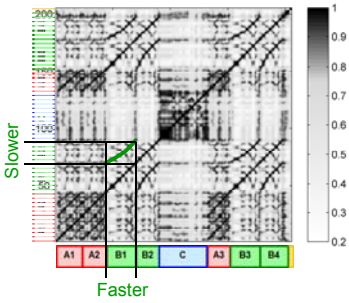
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



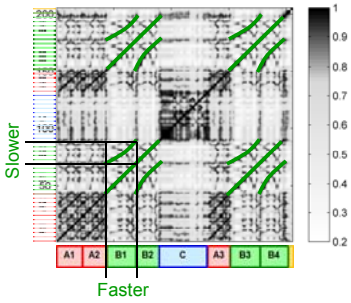
Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)



Self-Similarity Matrix (SSM)

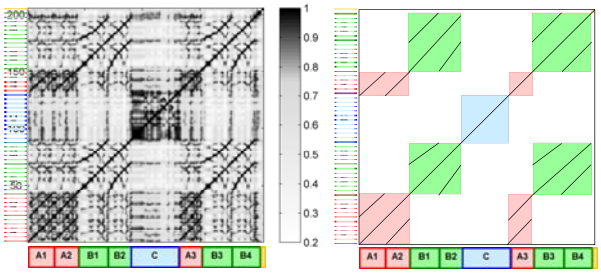
Example: Brahms Hungarian Dance No. 5 (Ormandy)



Self-Similarity Matrix (SSM)

Example: Brahms Hungarian Dance No. 5 (Ormandy)

Idealized SSM



Self-Similarity Matrix (SSM)

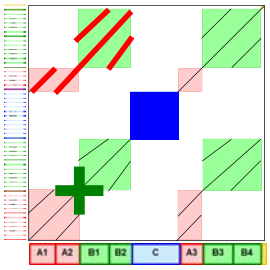
Example: Brahms Hungarian Dance No. 5 (Ormandy)

Idealized SSM

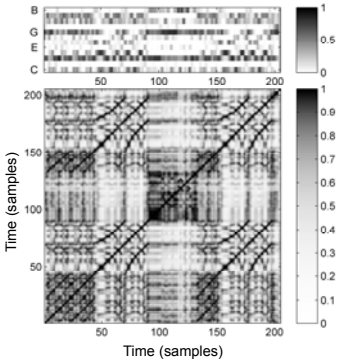
Blocks: Homogeneity

Paths: Repetition

Corners: Novelty



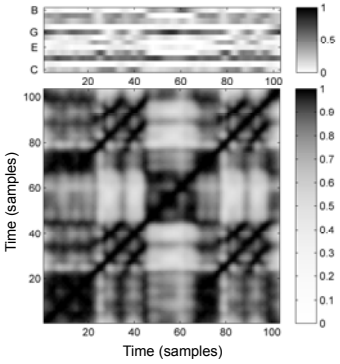
SSM Enhancement



Block Enhancement

- Feature smoothing
- Coarsening

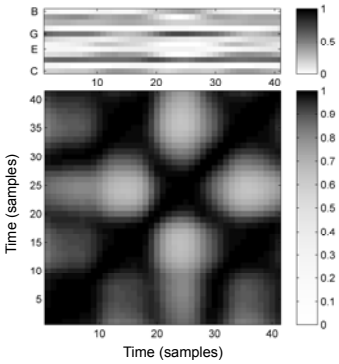
SSM Enhancement



Block Enhancement

- Feature smoothing
- Coarsening

SSM Enhancement

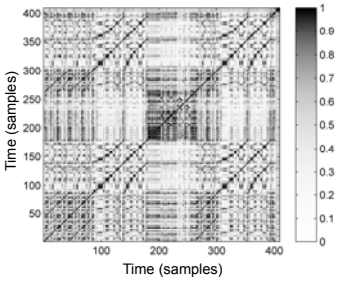


Block Enhancement

- Feature smoothing
- Coarsening

SSM Enhancement

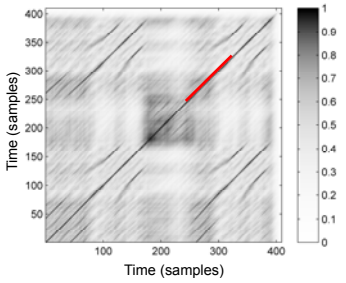
Path Enhancement



SSM Enhancement

Path Enhancement

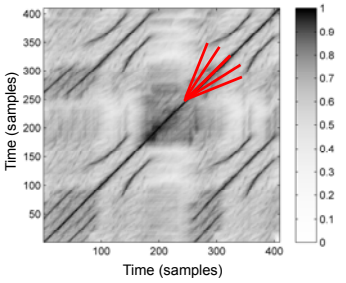
- Diagonal smoothing



SSM Enhancement

Path Enhancement

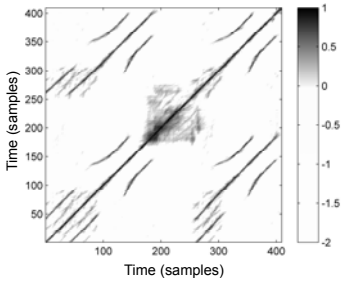
- Diagonal smoothing
- Multiple filtering



SSM Enhancement

Path Enhancement

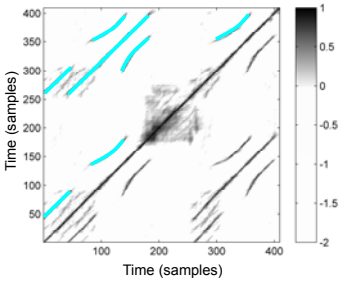
- Diagonal smoothing
- Multiple filtering
- Thresholding (relative)
- Scaling & penalty



SSM Enhancement

Further Processing

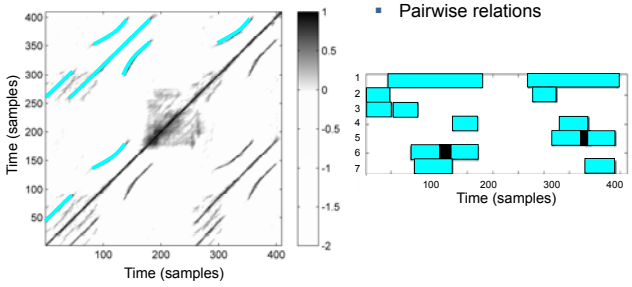
- Path extraction



SSM Enhancement

Further Processing

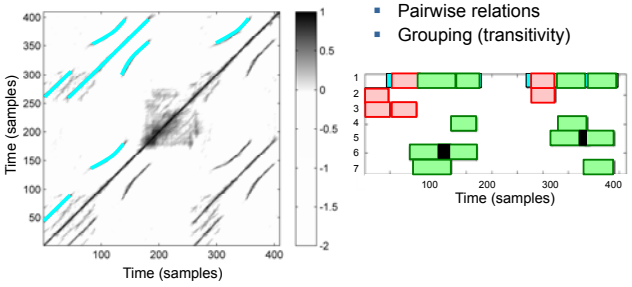
- Path extraction
- Pairwise relations



SSM Enhancement

Further Processing

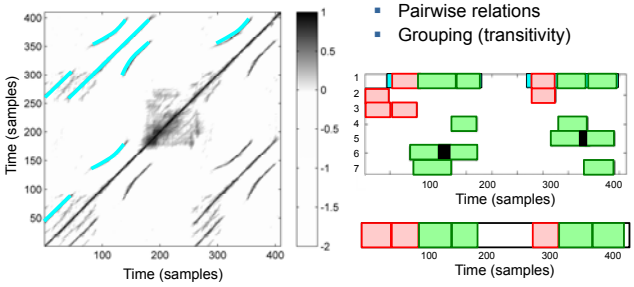
- Path extraction
- Pairwise relations
- Grouping (transitivity)



SSM Enhancement

Further Processing

- Path extraction
- Pairwise relations
- Grouping (transitivity)

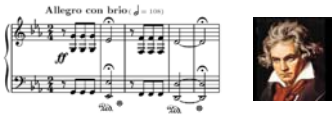


Overview

- Introduction
- Feature Representations
- Self-Similarity Matrices
- Music Synchronization
- Beat Tracking

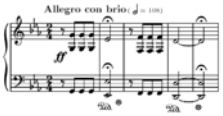
Music Synchronization

Beethoven's Fifth

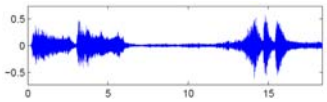


Music Synchronization

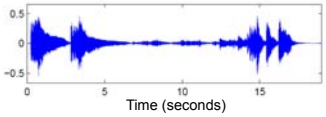
Beethoven's Fifth



Orchester (Karajan)

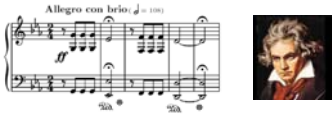


Piano (Scherbakov)



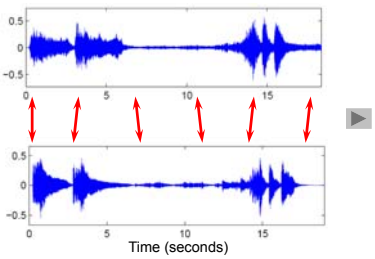
Music Synchronization

Beethoven's Fifth



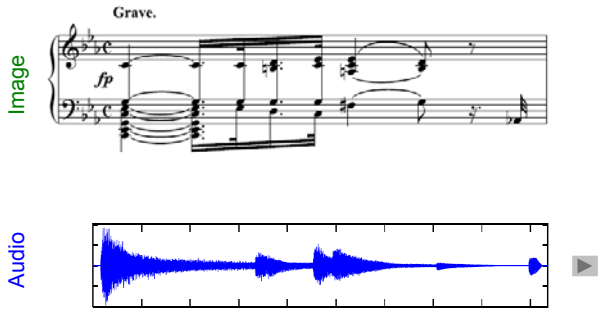
Orchester (Karajan)

Piano (Scherbakov)

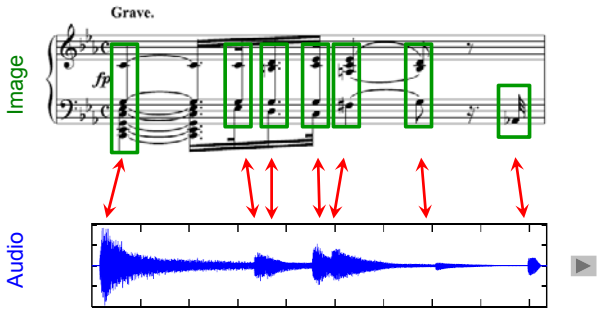


Application: Interpretation Switcher

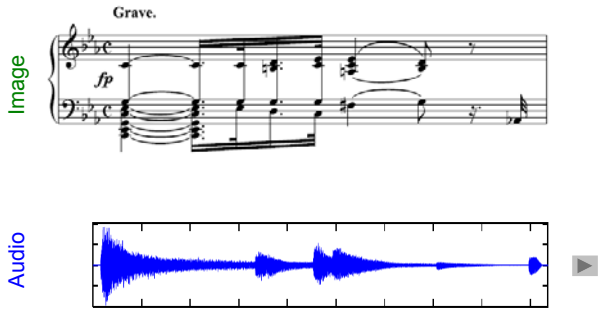
Music Synchronization: Image-Audio



Music Synchronization: Image-Audio

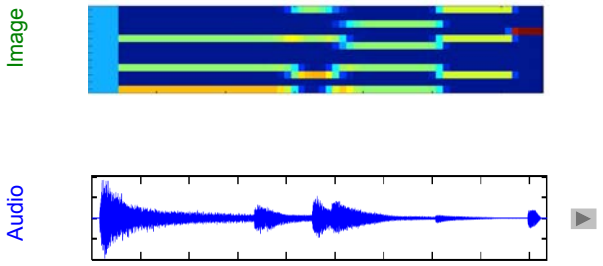


How to make the data comparable?



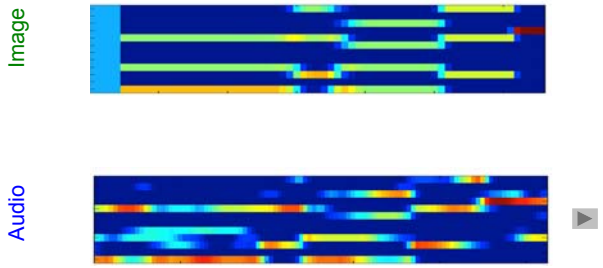
How to make the data comparable?

Image Processing: Optical Music Recognition



How to make the data comparable?

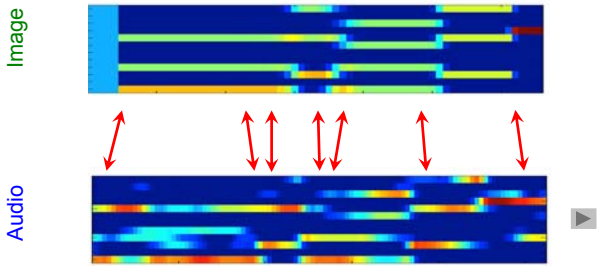
Image Processing: Optical Music Recognition



Audio Processing: Fourier Analyse

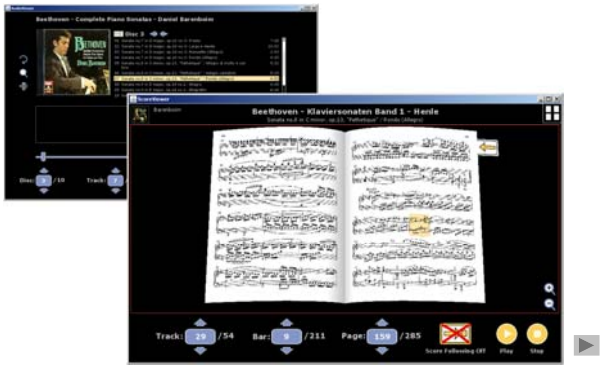
How to make the data comparable?

Image Processing: Optical Music Recognition



Audio Processing: Fourier Analyse

Application: Score Viewer



Overview

- Introduction
- Feature Representations
- Self-Similarity Matrices
- Music Synchronization
- Beat Tracking

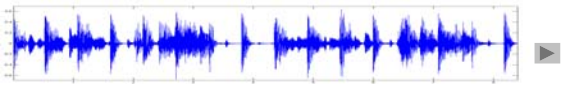
Beat Tracking

Basic task: "Tapping the foot when listening to music"

Beat Tracking

Basic task: "Tapping the foot when listening to music"

Example: Queen – Another One Bites The Dust

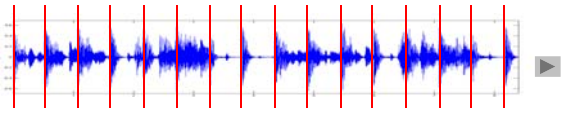


Time (seconds)

Beat Tracking

Basic task: "Tapping the foot when listening to music"

Example: Queen – Another One Bites The Dust

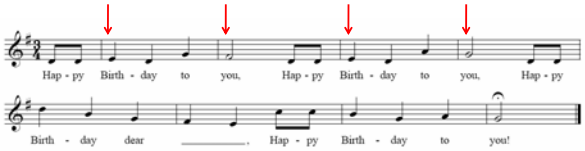


Time (seconds)

Beat Tracking

Example: Happy Birthday to you

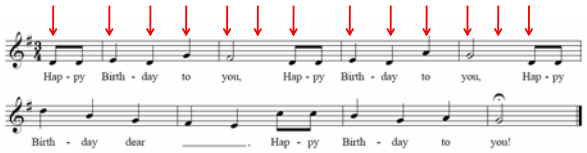
Pulse level: **Measure**



Beat Tracking

Example: Happy Birthday to you

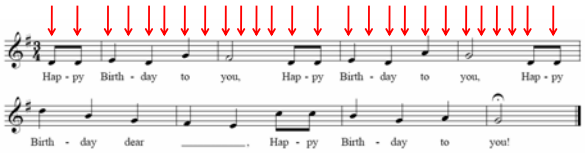
Pulse level: **Tactus (beat)**



Beat Tracking

Example: Happy Birthday to you

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



Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: ???

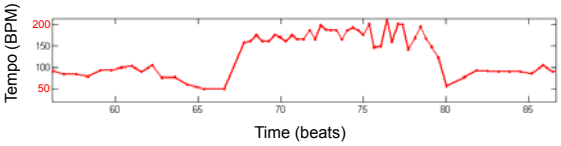
Beat Tracking

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: 50-200 BPM

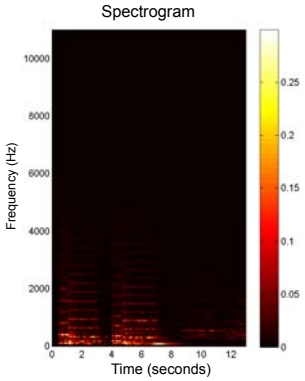
Tempo curve



Beat Tracking

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

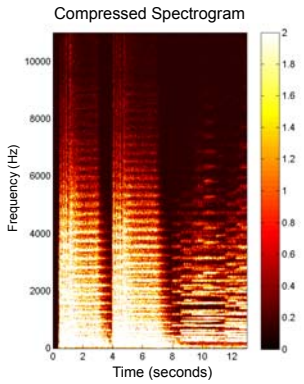
Beat Tracking



Steps:

1. Spectrogram

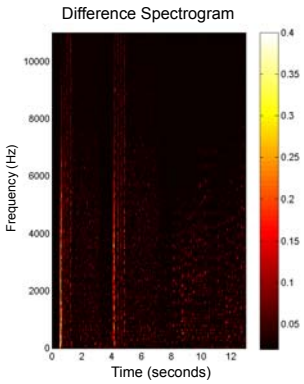
Beat Tracking



Steps:

1. Spectrogram
2. Log Compression

Beat Tracking



Steps:

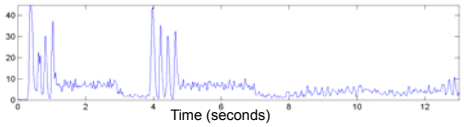
1. Spectrogram
2. Log Compression
3. Differentiation

Beat Tracking

Steps:

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

Novelty Curve

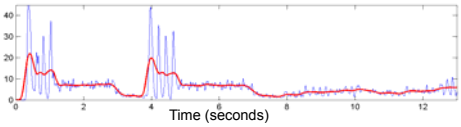


Beat Tracking

Steps:

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

Novelty Curve
Local Average

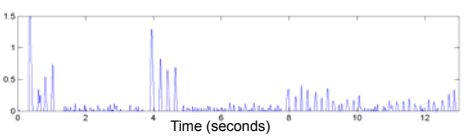


Beat Tracking

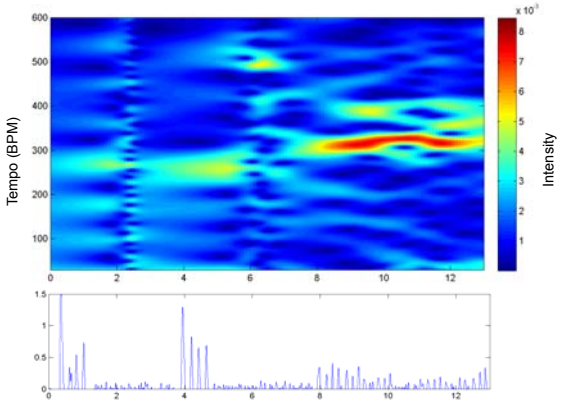
Steps:

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

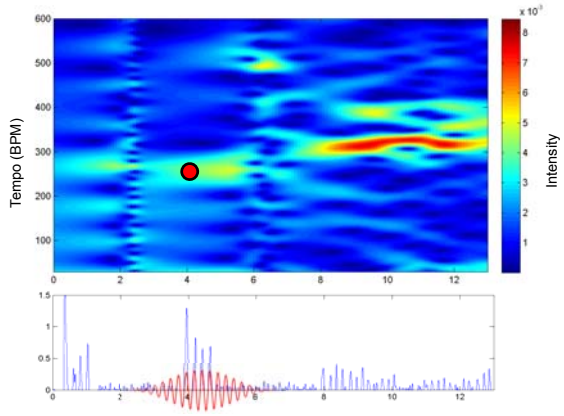
Novelty Curve



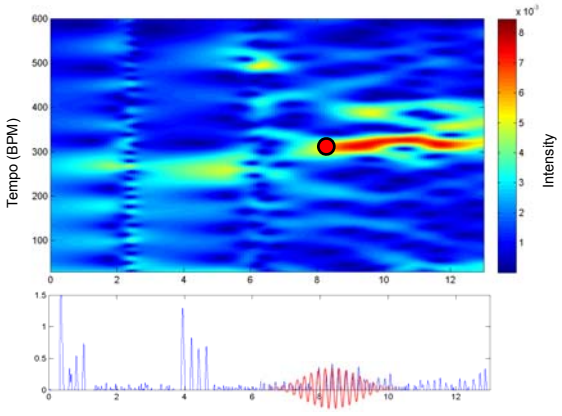
Beat Tracking



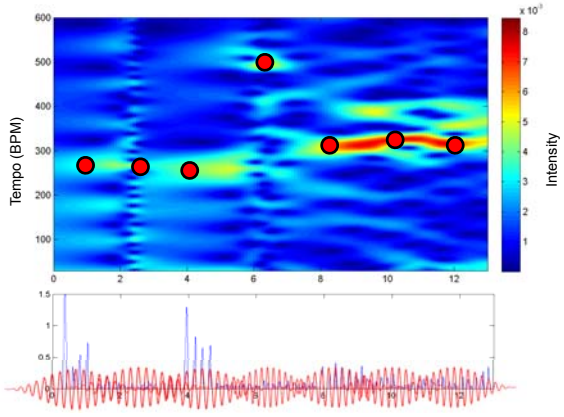
Beat Tracking



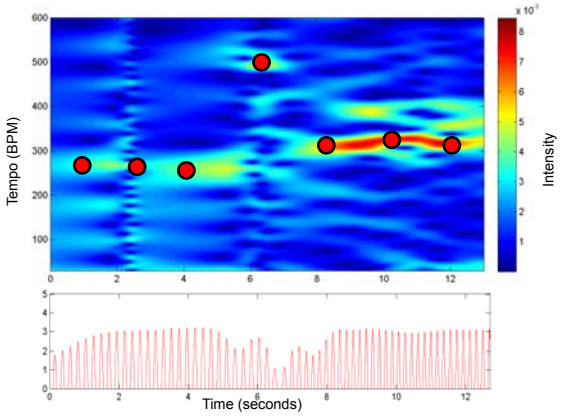
Beat Tracking



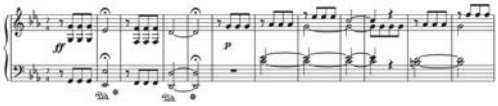
Beat Tracking



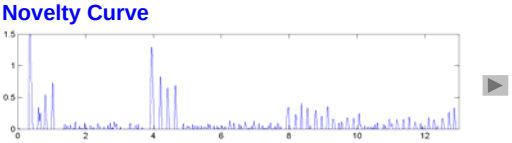
Beat Tracking



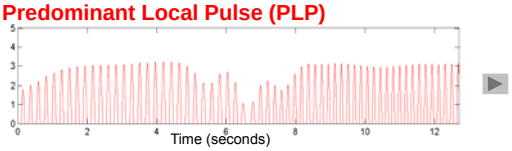
Beat Tracking



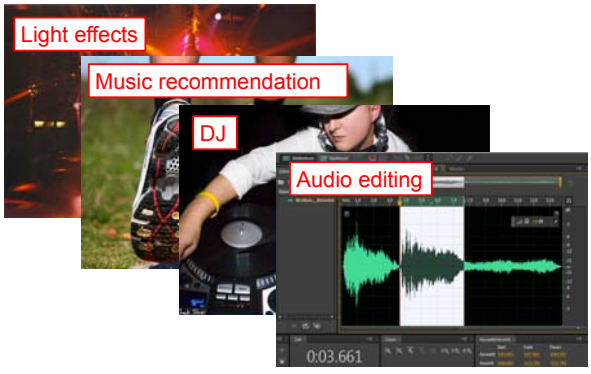
Novelty Curve



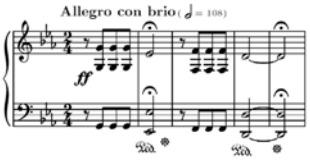
Predominant Local Pulse (PLP)



Beat Tracking

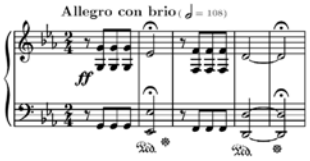


Motivic Similarity



Beethoven's Fifth (1st Mov.)

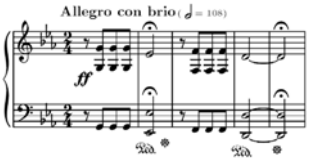
Motivic Similarity



Beethoven's Fifth (1st Mov.)

Beethoven's Fifth (3rd Mov.)

Motivic Similarity

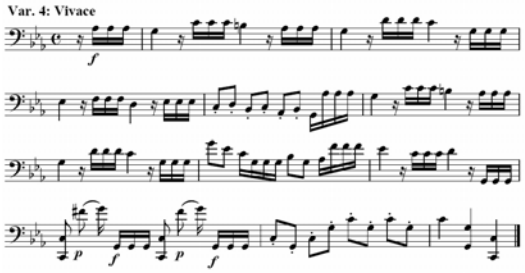


Beethoven's Fifth (1st Mov.)

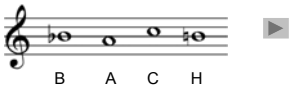
Beethoven's Fifth (3rd Mov.)

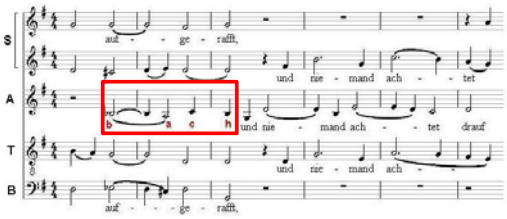
Beethoven's Appassionata

Motivic Similarity



Motivic Similarity





Book Project

A First Course on Music Processing

Textbook (approx. 500 pages)

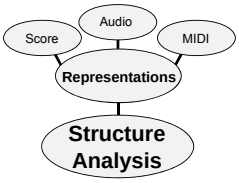
1. Music Representations
2. Fourier Analysis of Signals
3. Music Synchronization
4. **Music Structure Analysis**
5. Chord Recogntion
6. Temo and Beat Tracking
7. Content-based Audio Retrieval
8. Music Transcription

Springer

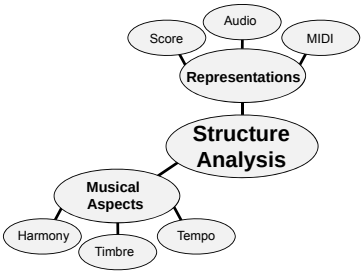
To appear (plan):
End of 2015

Conclusions

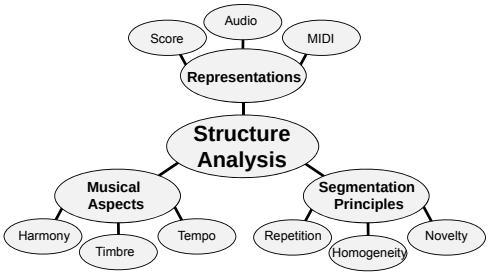
Conclusions



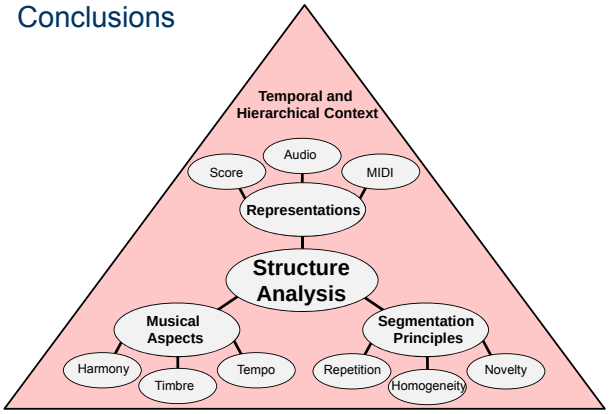
Conclusions



Conclusions



Conclusions



Conclusions

- Combined Approaches
- Hierarchical Approaches
- Evaluation
- Explaining Structure

- MIREX
- SALAMI-Project
- Smith, Chew

Projects with Musicology

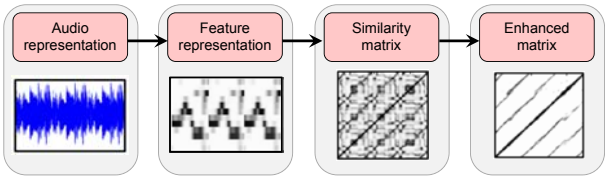
Computergestützte Analyse harmonischer Strukturen

Kooperationspartner:
Prof. Rainer Kleinertz
Universität des Saarlandes
Institut für Musikwissenschaft

Freischütz Digital

Kooperationspartner:
Prof. Joachim Veit, Universität Paderborn / Detmold
Prof. Thomas Betzwieser, Universität Frankfurt
Prof. Gerd Szwillus, Universität Paderborn

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/

PhD Projects (Final Stage)

- Nanzhu Jiang
  Universtät 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

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. 452–455.
- 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 8th 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 similiarty 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 IEEEWorkshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Platz, NY, USA, 1999, pp. 103–106.