INTERNATIONAL AUDIO LABORATORIES ERLANGEN

Tutorial
Automatisierte Methoden der Musikverarbeitung
47. Jahrestagung der Gesellschaft für Informatik

Music Structure Analysis

Meinard Müller, Christof Weiss, Stefan Balke

International Audio Laboratories Erlangen
{meinard.mueller, christof.weiss, stefan.balke}@audiolabs-erlangen.de

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

Book: Fundamentals of Music Processing

Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

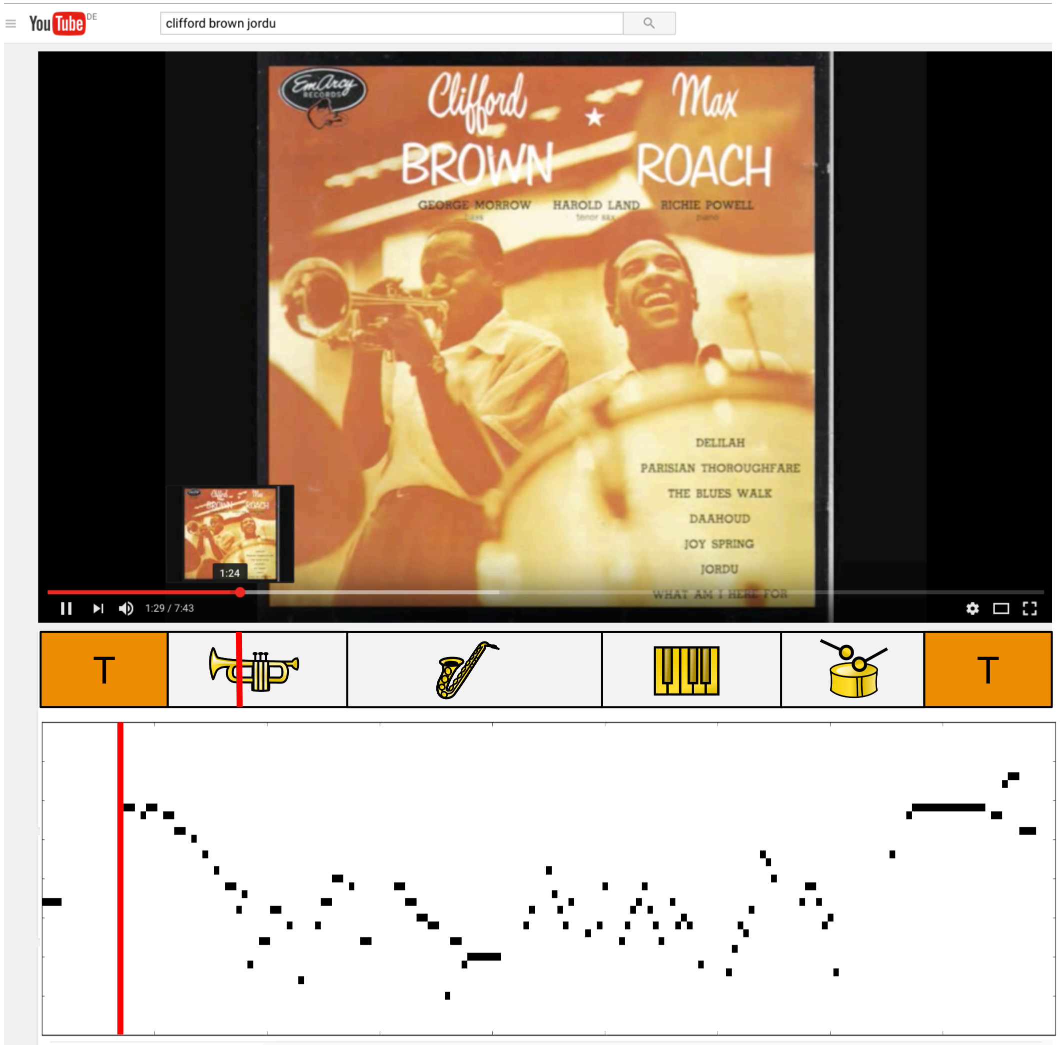
Book: Fundamentals of Music Processing

Chapter	Music Processing Scenario
1	Music Representations
2	Fourier Analysis of Signals
3	Music Synchronization
4	Music Structure Analysis
5	Chord Recognition
6	Tempo and Beat Tracking
7	Content-Based Audio Retrieval
8	Musically Informed Audio Decomposition

Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

Motivation



clifford brown jordu
Clifford BROWN ★ Max ROACH
GEORGE MORROW
HAROLD LAND
RICHIE POWELL
DELILAH
PARISIAN THOROUGHFARE
THE BLUES WALK
DAAHOUD
JOY SPRING
JORDU
WHAT AM I HERE FOR
1:24
1:29 / 7:43
T
T

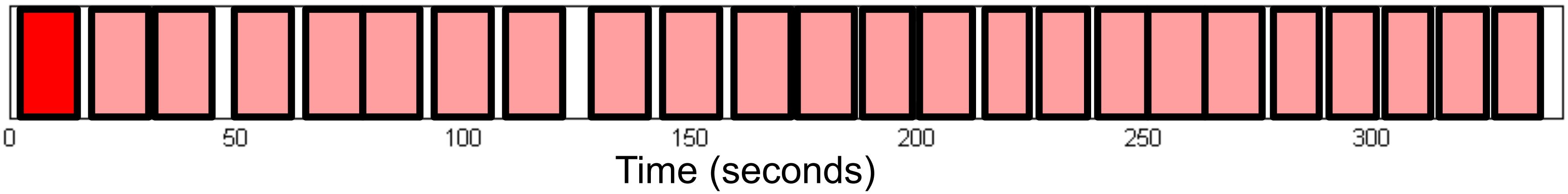
Music Structure Analysis

- **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
  - Solo parts in a jazz recording

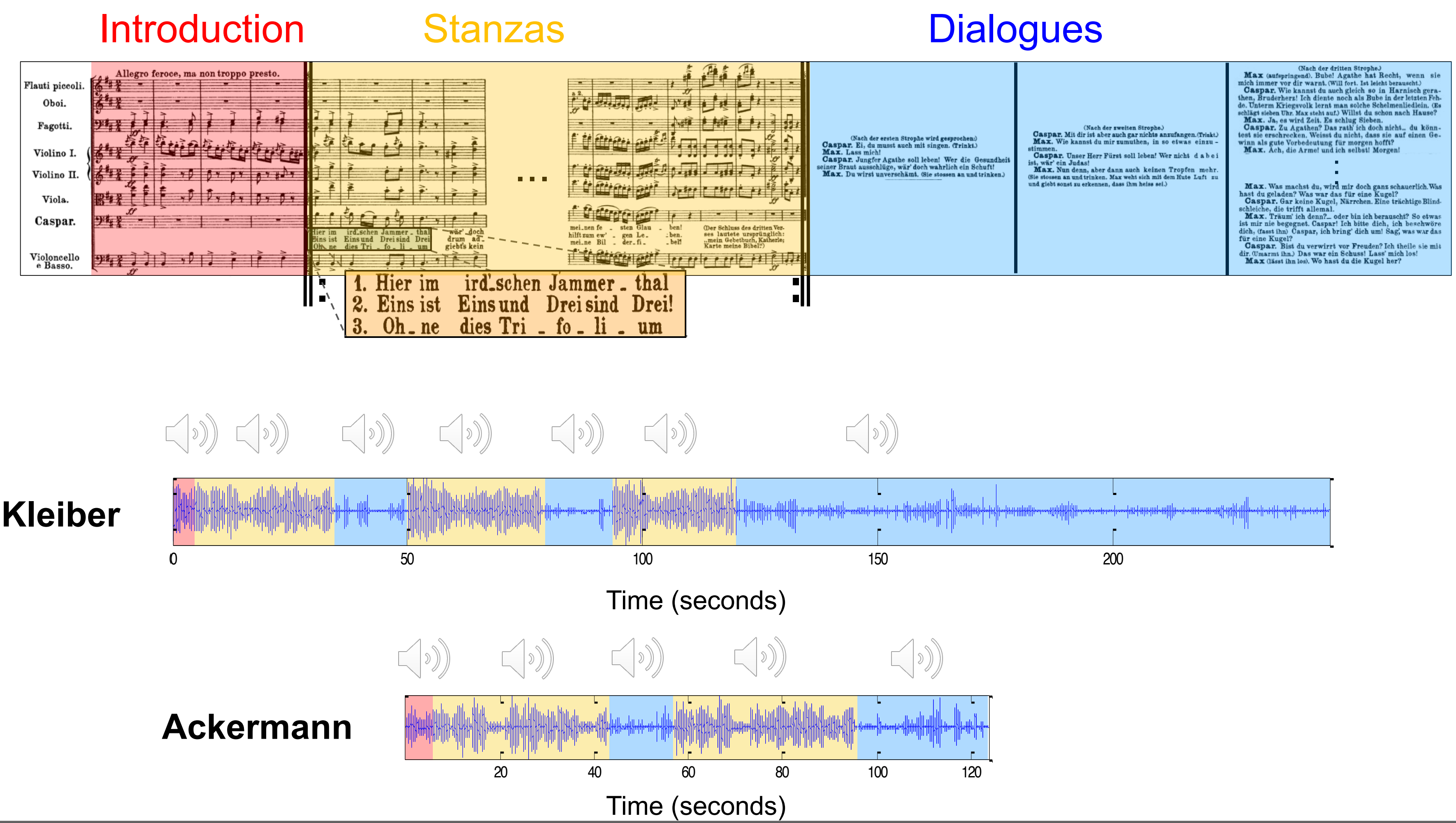
Example: Folk Song

Nederlandse Liederenbank

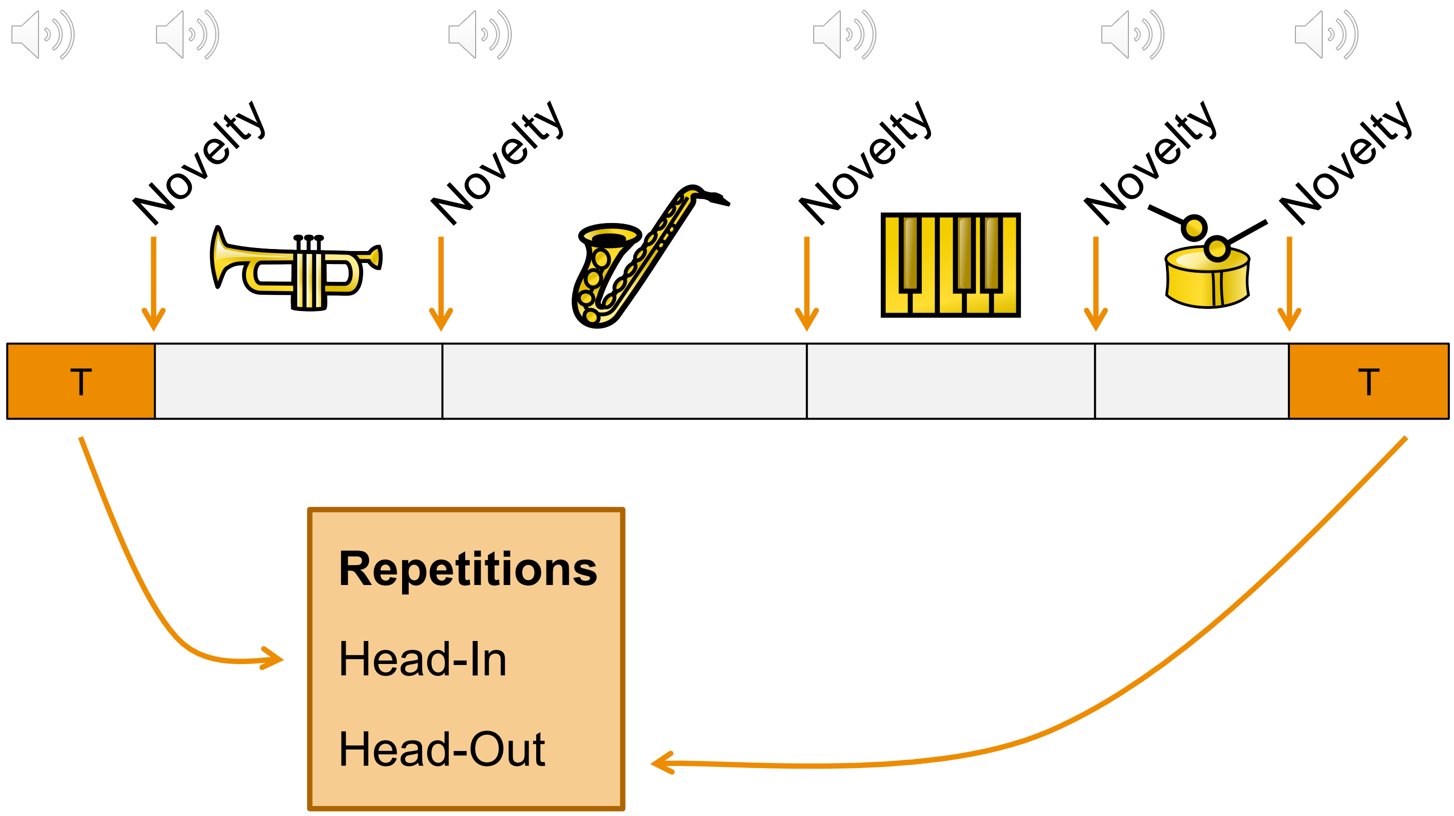


Example: Opera

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



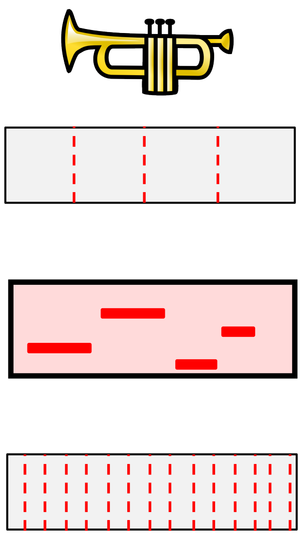
Example: Jazz Recording
Clifford Brown - Jordu



Instrument Comics by H. Grohganz: https://mir.sechsachtel.de/orchpics/

Weimar Jazz Database (WJD)

http://jazzomat.hfm-weimar.de



Transcription

Beats

| E^7 A^7 | D^7 G^7 | …

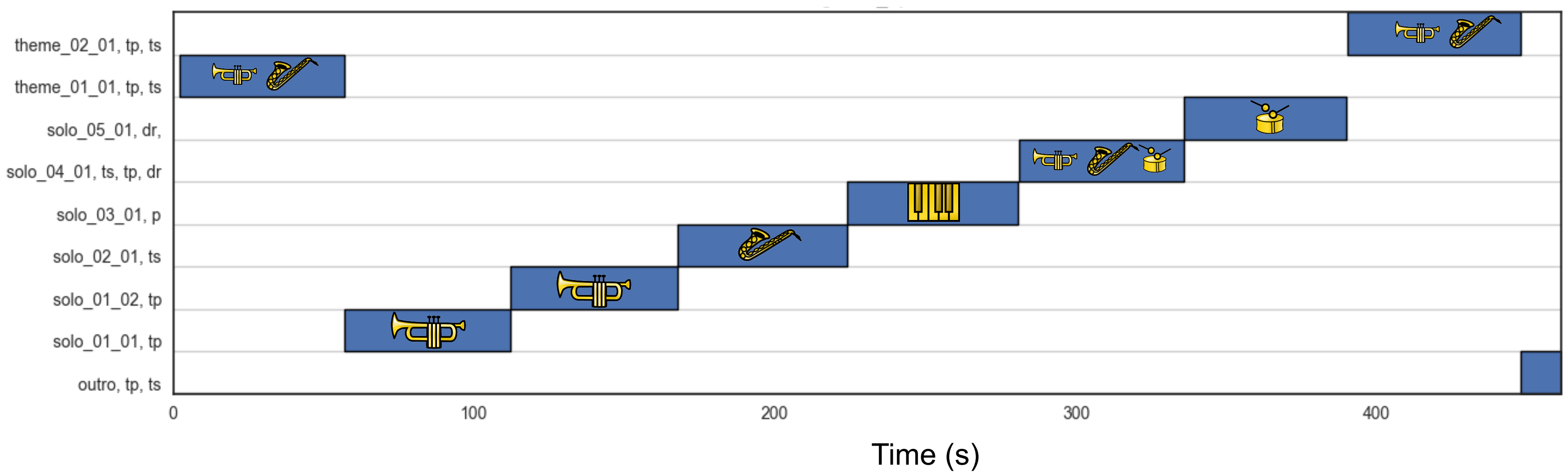
Chords

…

- 456 transcribed jazz solos of monophonic instruments.
- Transcriptions specify a musical pitch for physical time instances.
- 810 min. of audio recordings.
- **Soon available:** Track structure

Thanks to the Jazzomat research team: M. Pfleiderer, K. Frieler, J. Abeßer, W.-G. Zaddach

Annotation Example from the WJD
Clifford Brown - Jordu



Song-Centric Annotations:

- Chorus boundaries
- Solo choruses
- Theme repetitions

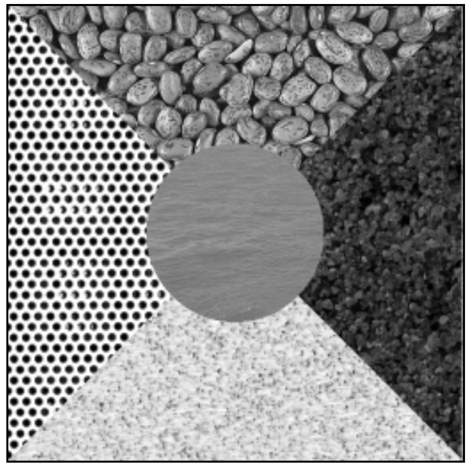
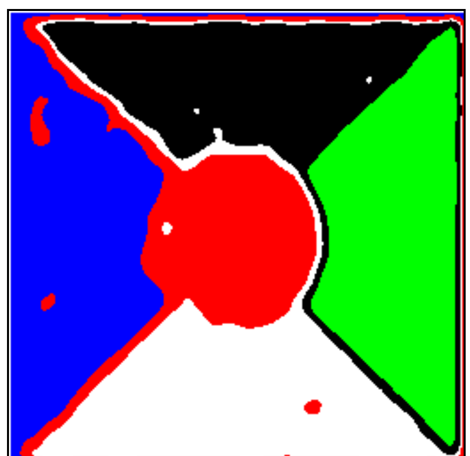
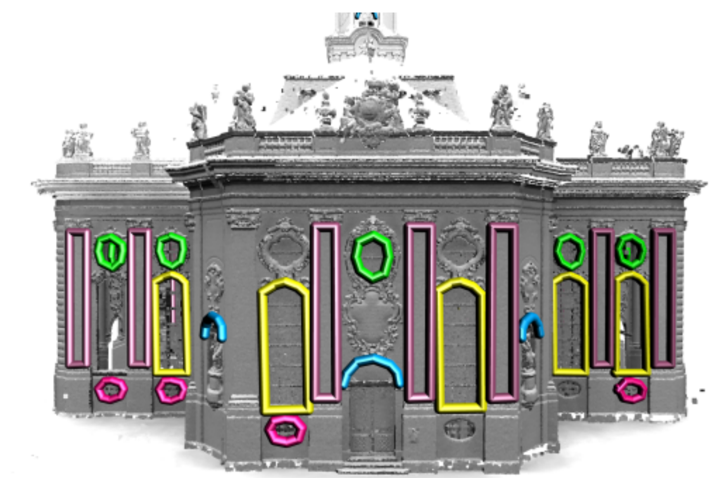
Many Thanks to
Moritz Berendes and
Julian Reck!

Music Structure Analysis

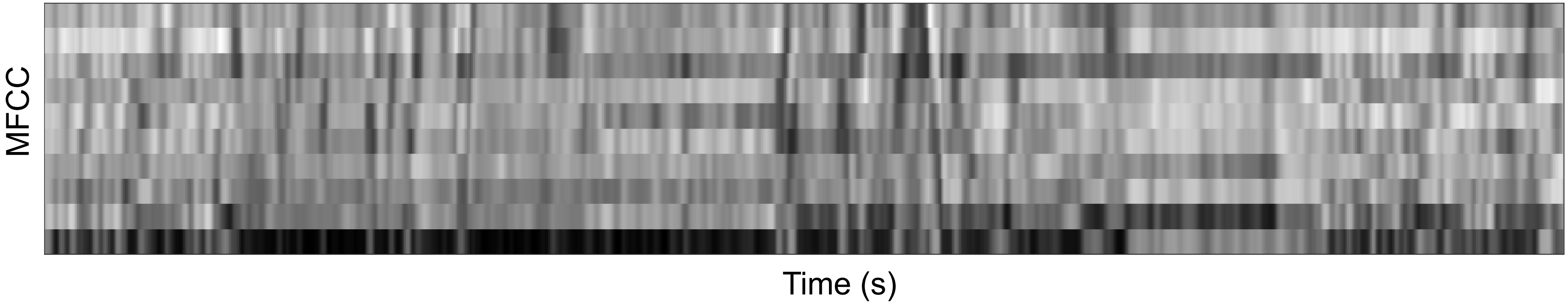
- Main principles:
  - **Repetition**-based Structure Analysis
  - **Homogeneity**-based Structure Analysis
  - **Novelty**-based Structure Analysis

Music Structure Analysis

Image Analogy

Novelty	Homogeneity	Repetition
		
		

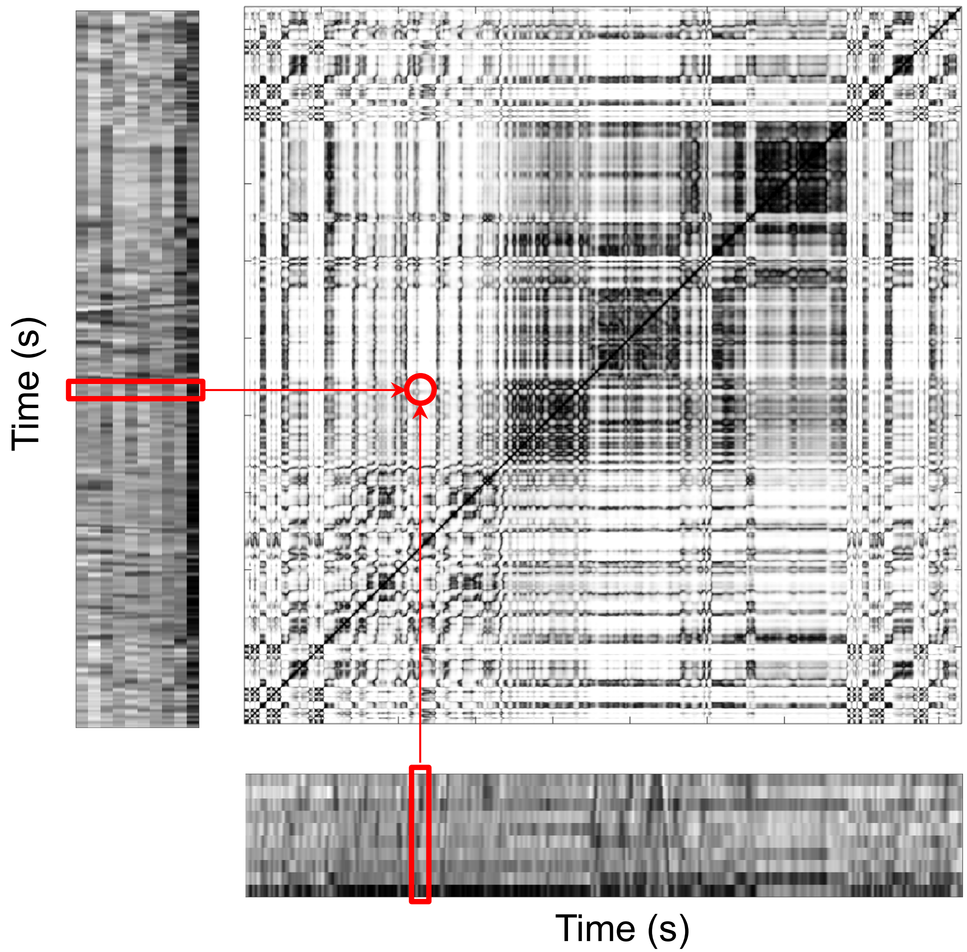
Self-Similarity Matrix
1. Step: Extract Audio Features



- Mel Frequency Cepstral Coefficients (MFCC) correlate to the timbre.

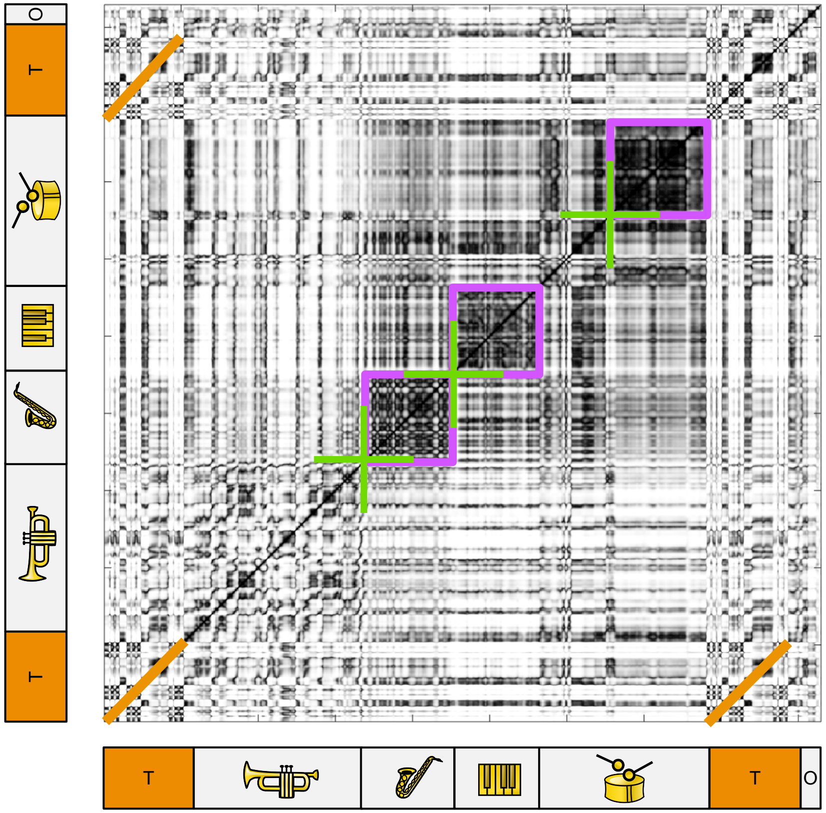
Self-Similarity Matrix

2. Step: Calculate Pairwise Similarity

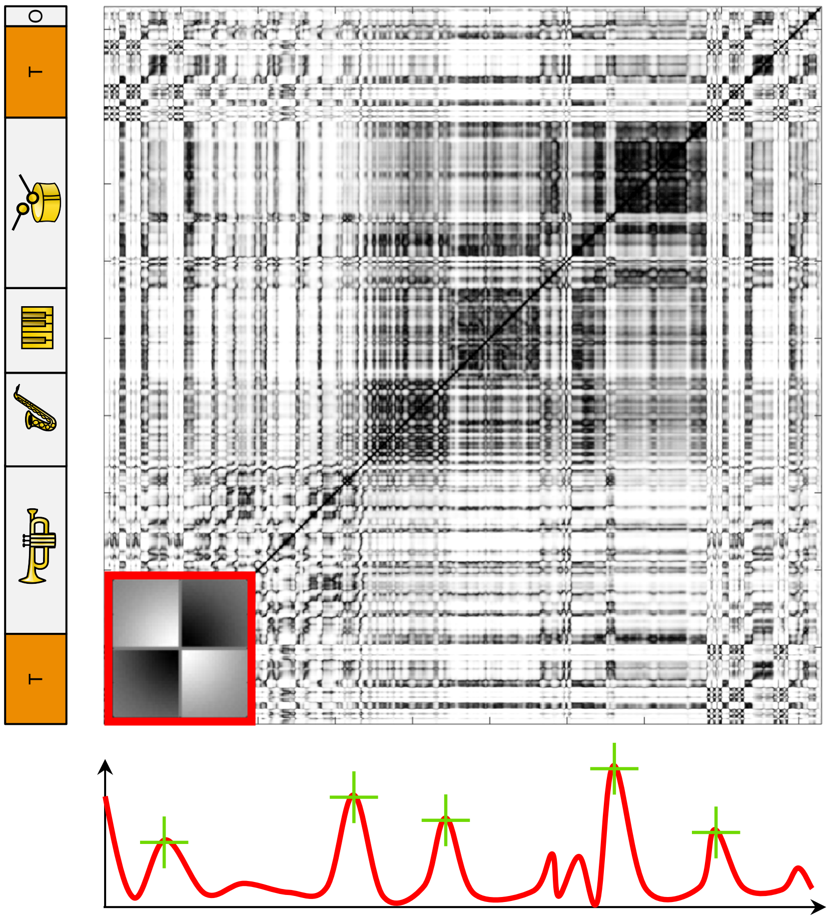


Self-Similarity Matrix
Analysis

- Repetitions:
  Path-like structures
- Homogeneity:
  Block-like structures
- Novelty:
  Corners



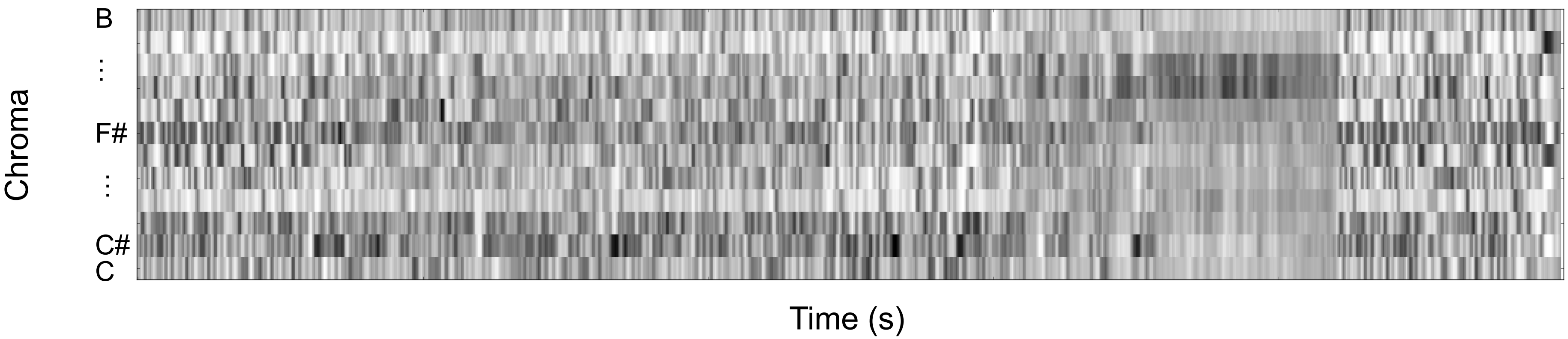
Self-Similarity Matrix
Novelty Detection



Idea (Foote):

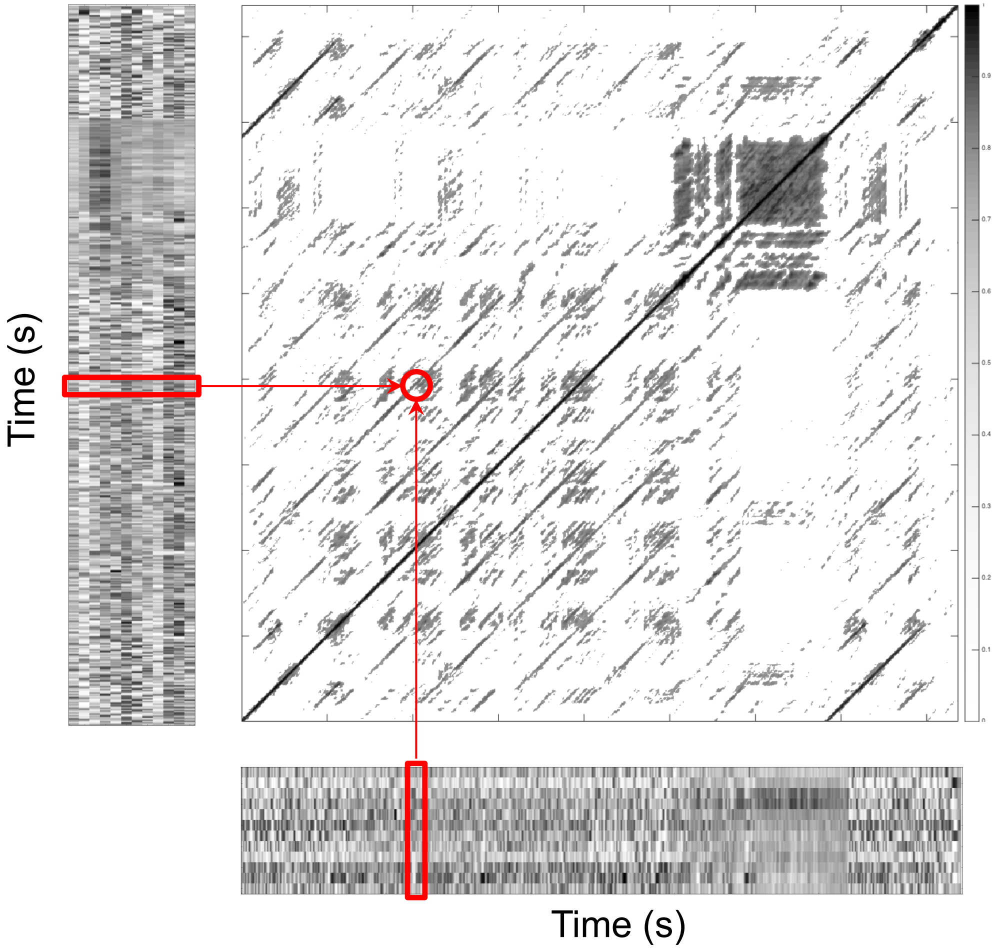
Use checkerboard-like kernel function to detect corner points on main diagonal of SSM.

Self-Similarity Matrix
Chroma Features



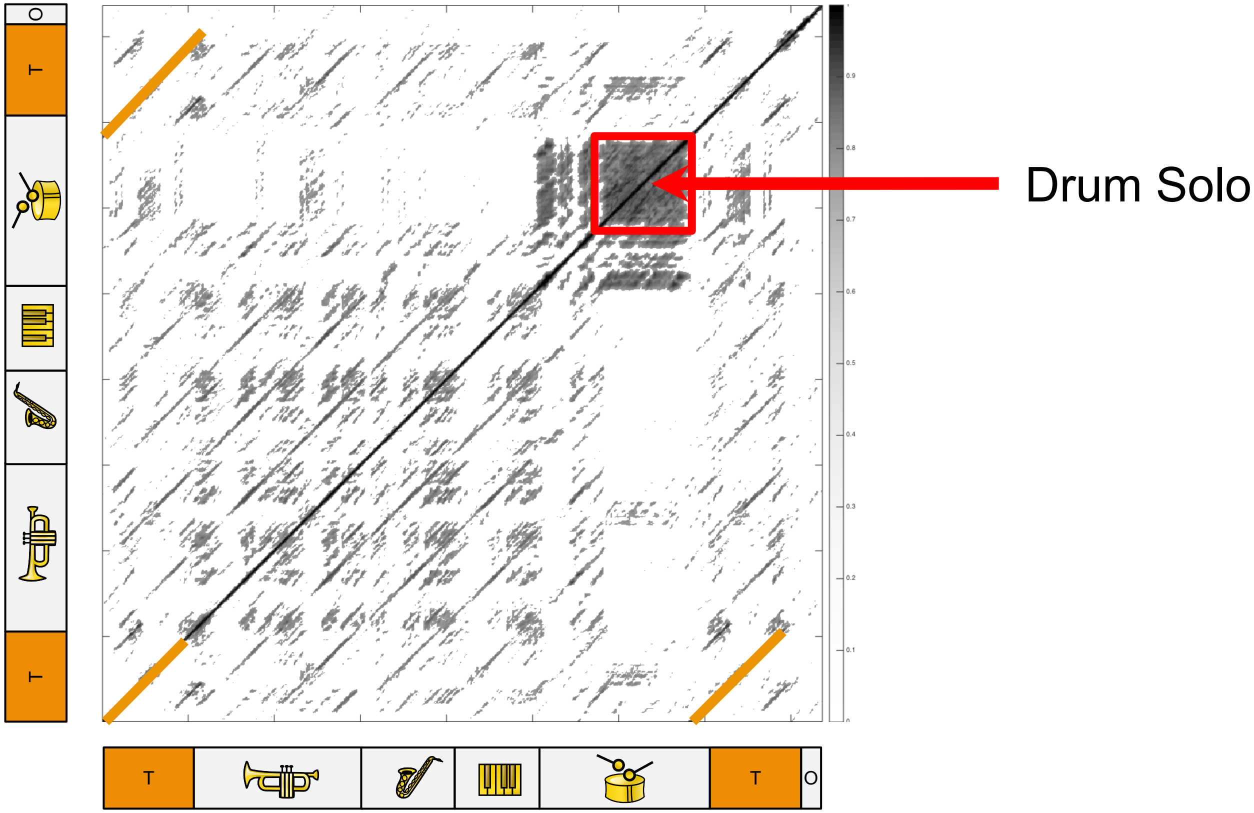
- Chroma Feature correlate to harmonic and melodic progressions.

Self-Similarity Matrix
Chroma Features



Self-Similarity Matrix
Chroma Features

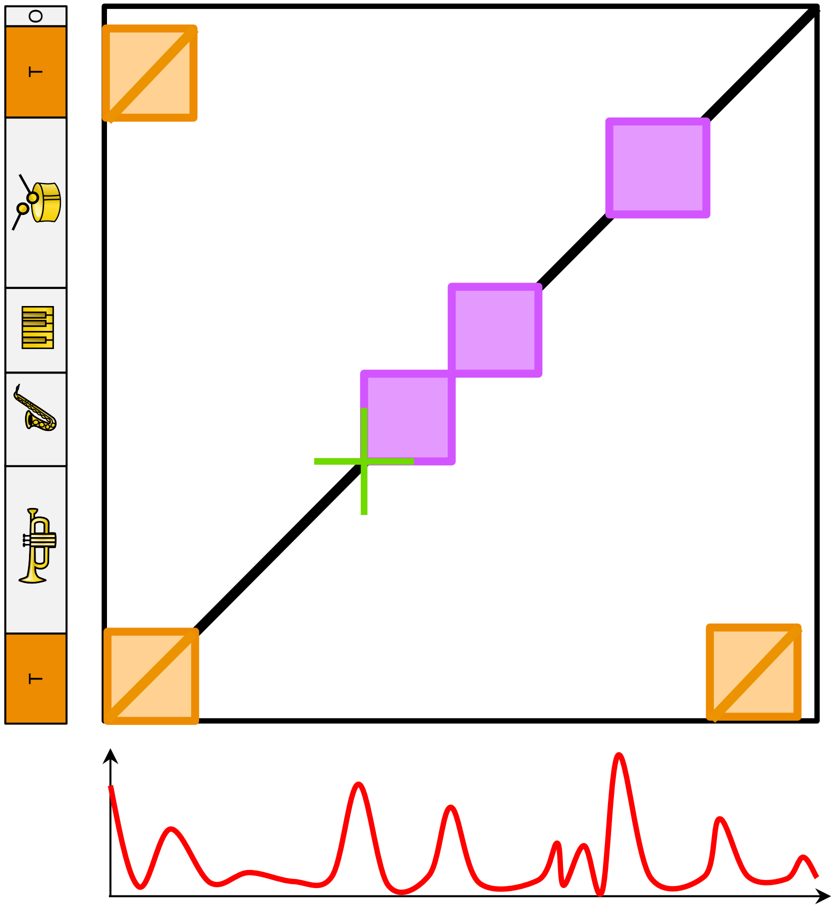
- Chroma instead of MFCC
- Repetitions result in path-like structures
- Head-In and Head-Out



Self-Similarity Matrix
Recap

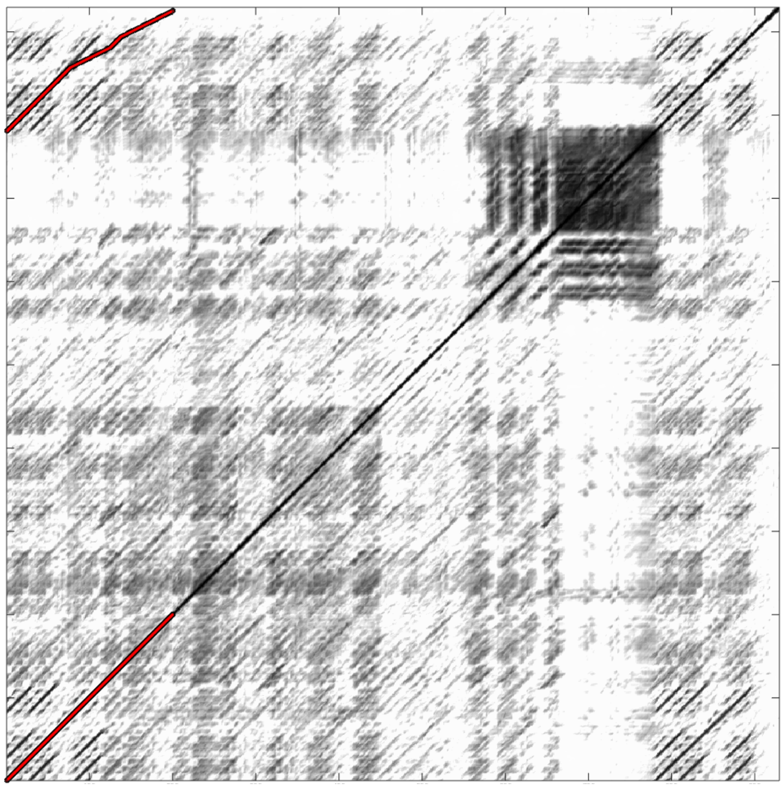
- Repetitions:
  Path-like structures
- Homogeneity:
  Block-like structures
- Novelty:
  Corners

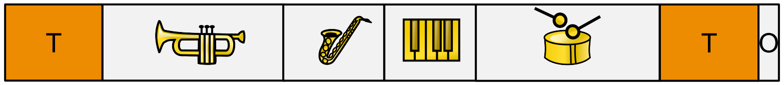
- Features are important!



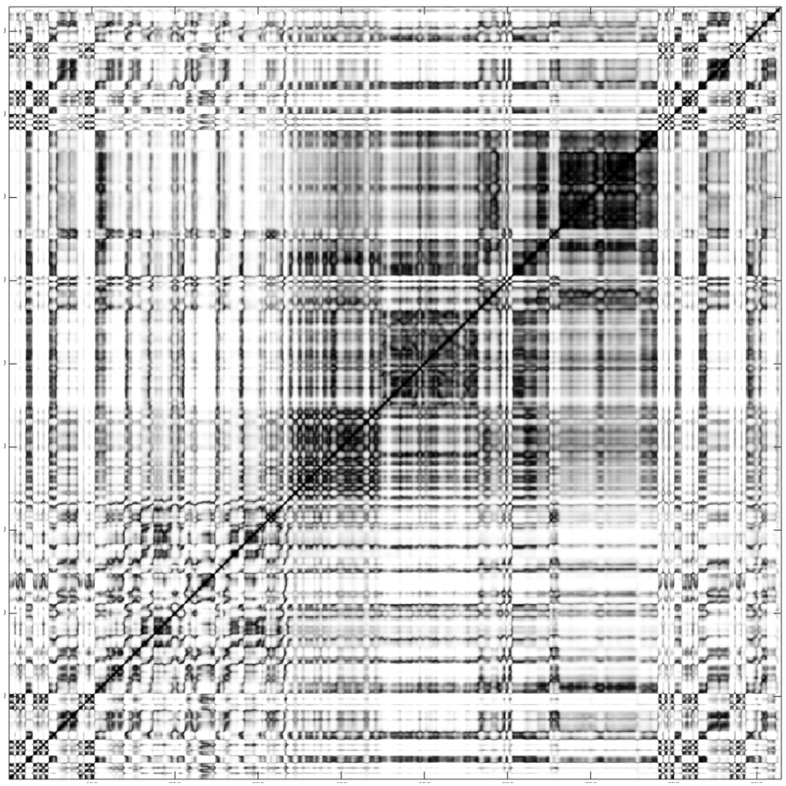
Audio Examples
Clifford Brown – Jordu

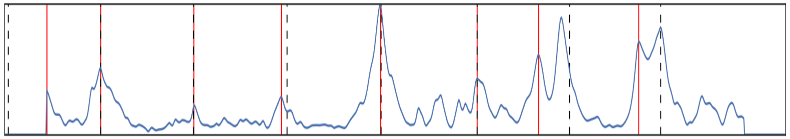
Chroma





MFCC

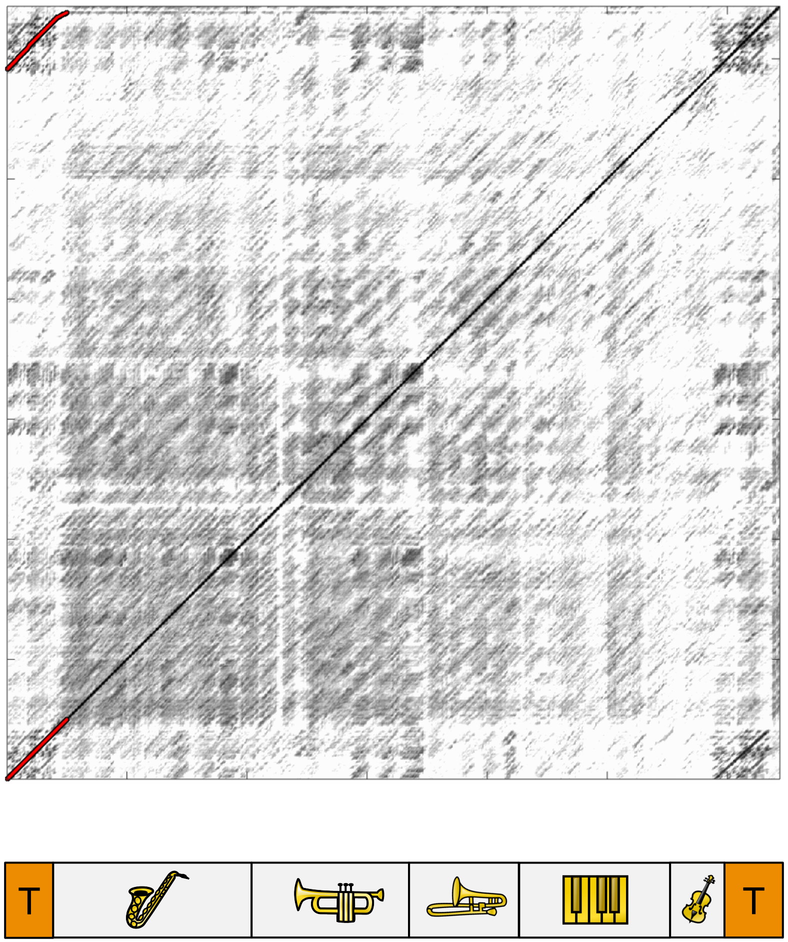




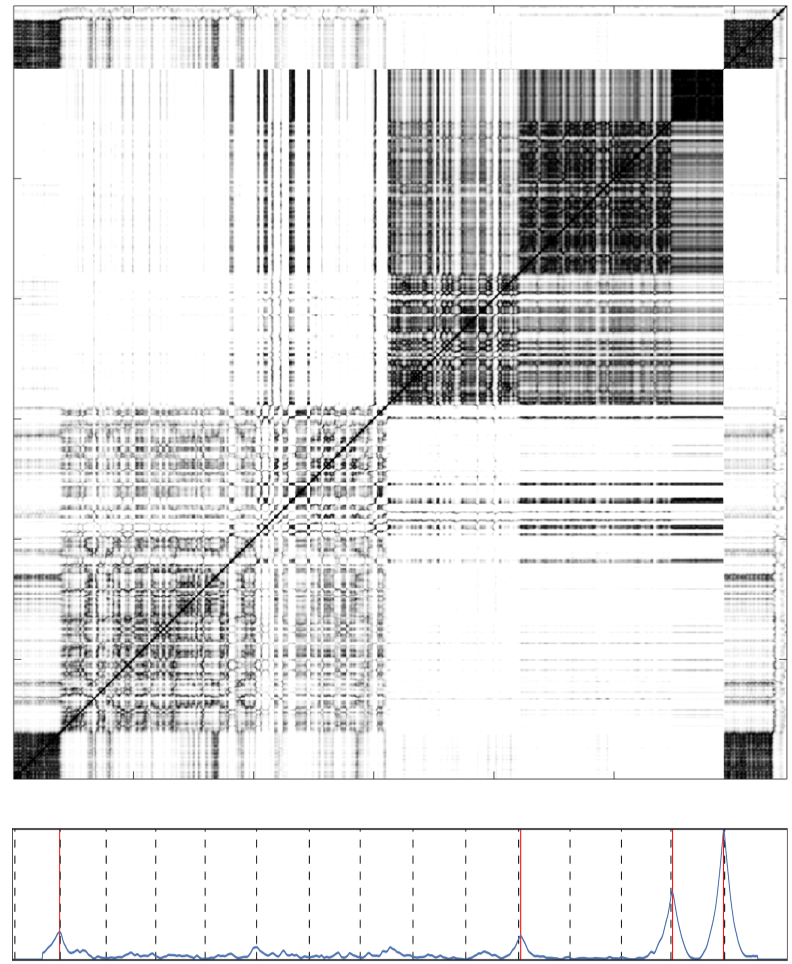
Audio Examples

John Coltrane – Blue Trane

Chroma



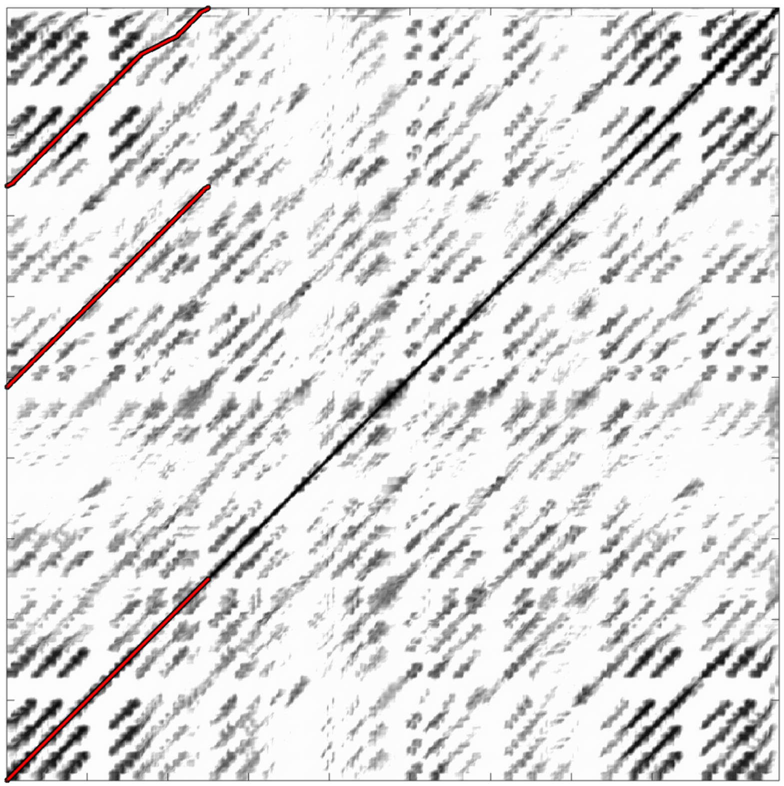
MFCC



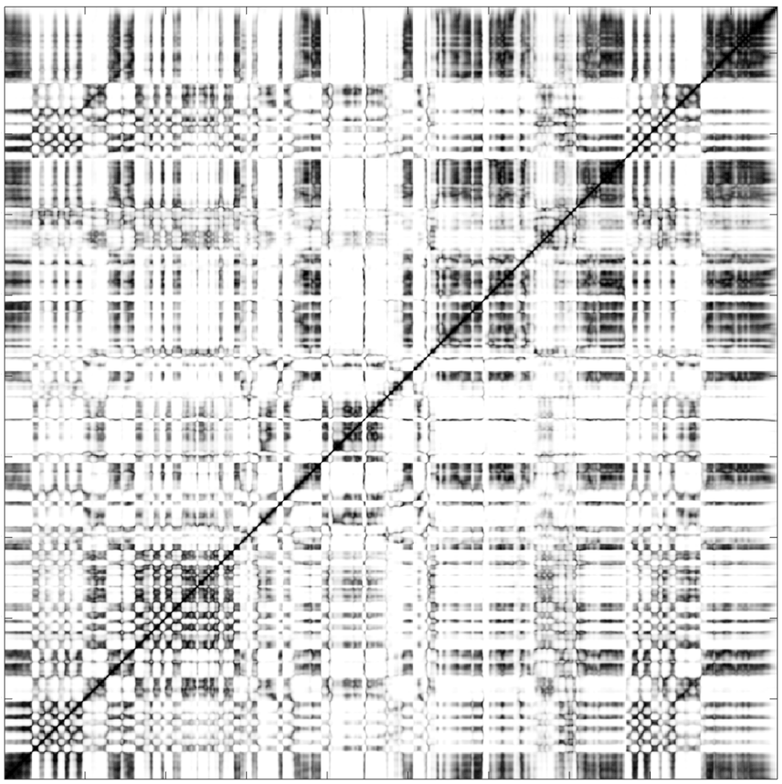
Audio Examples

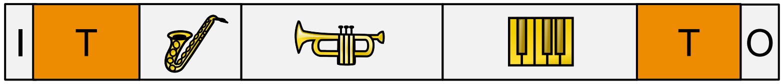
Herbie Hancock – Maiden Voyage

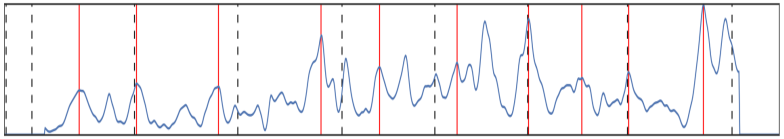
Chroma



MFCC

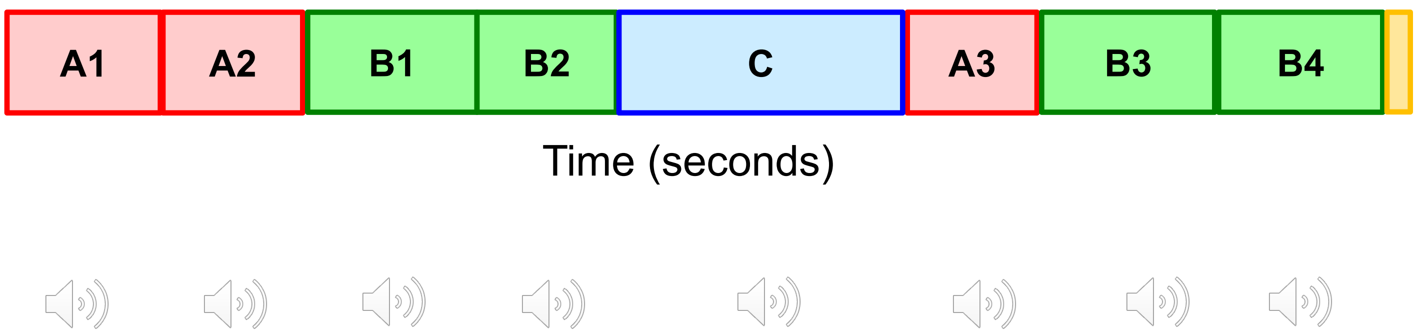






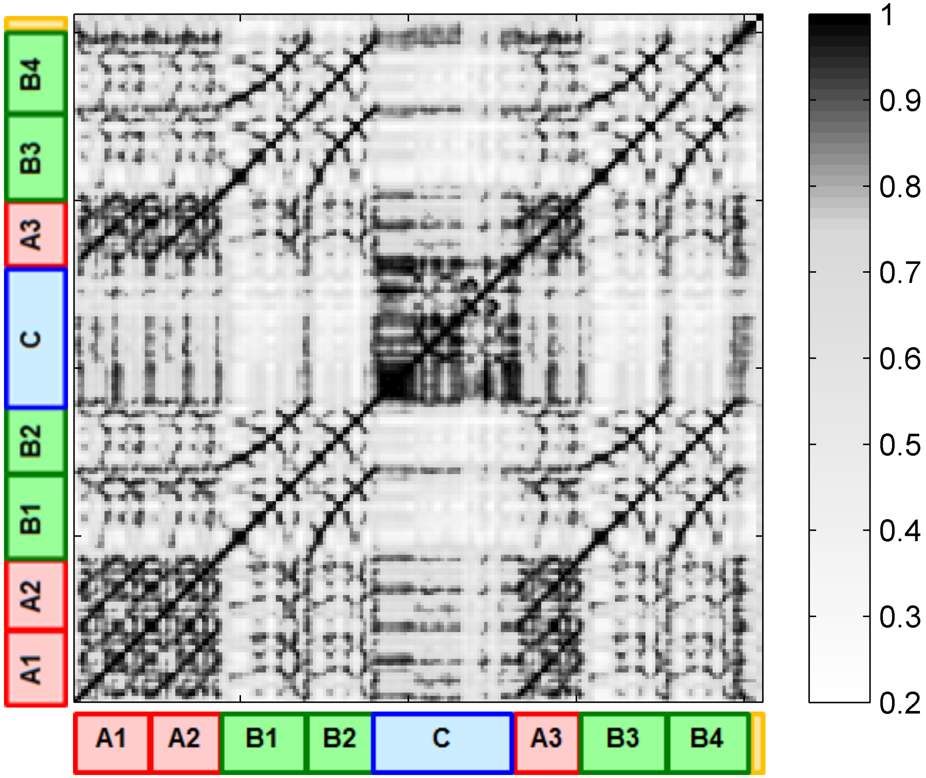
Self-Similarity Matrix

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



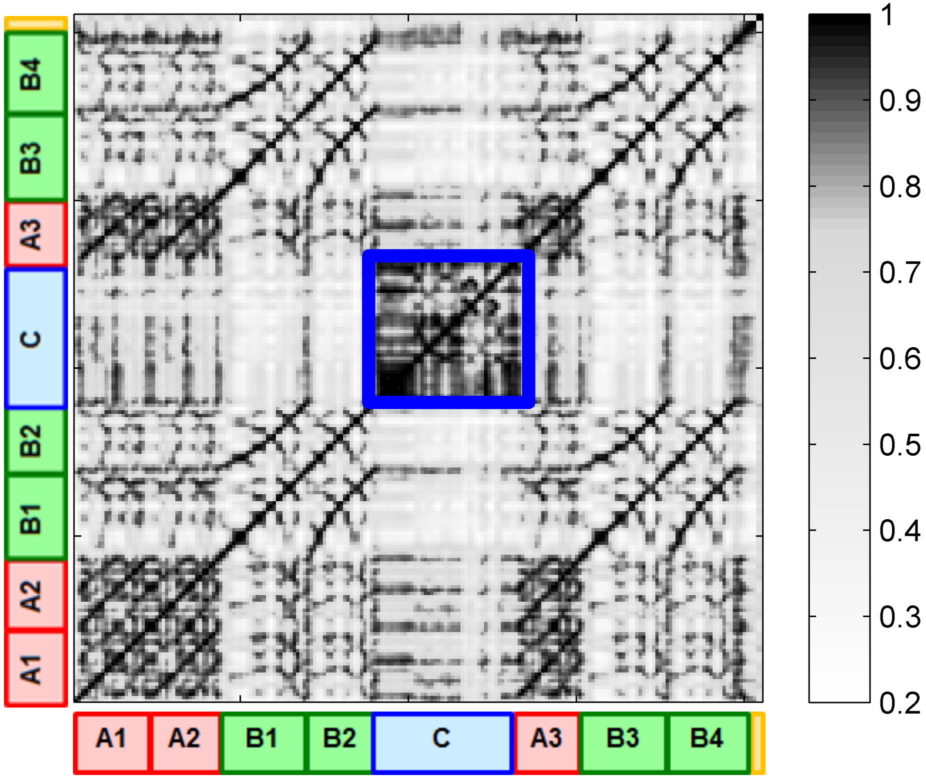
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



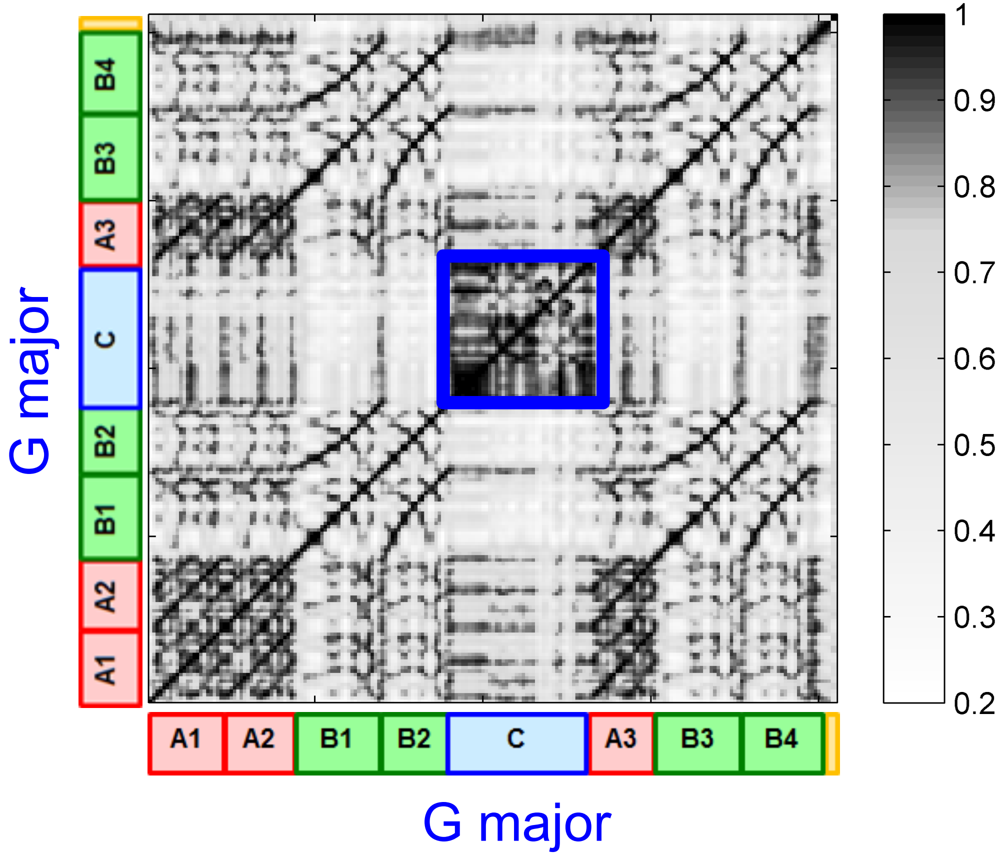
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



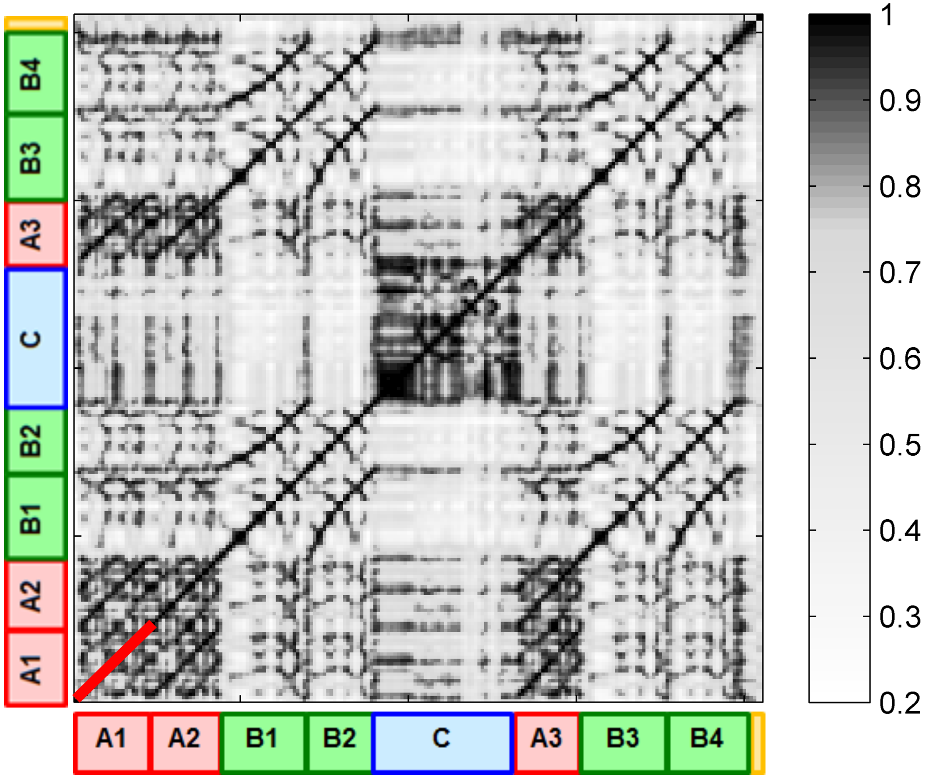
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



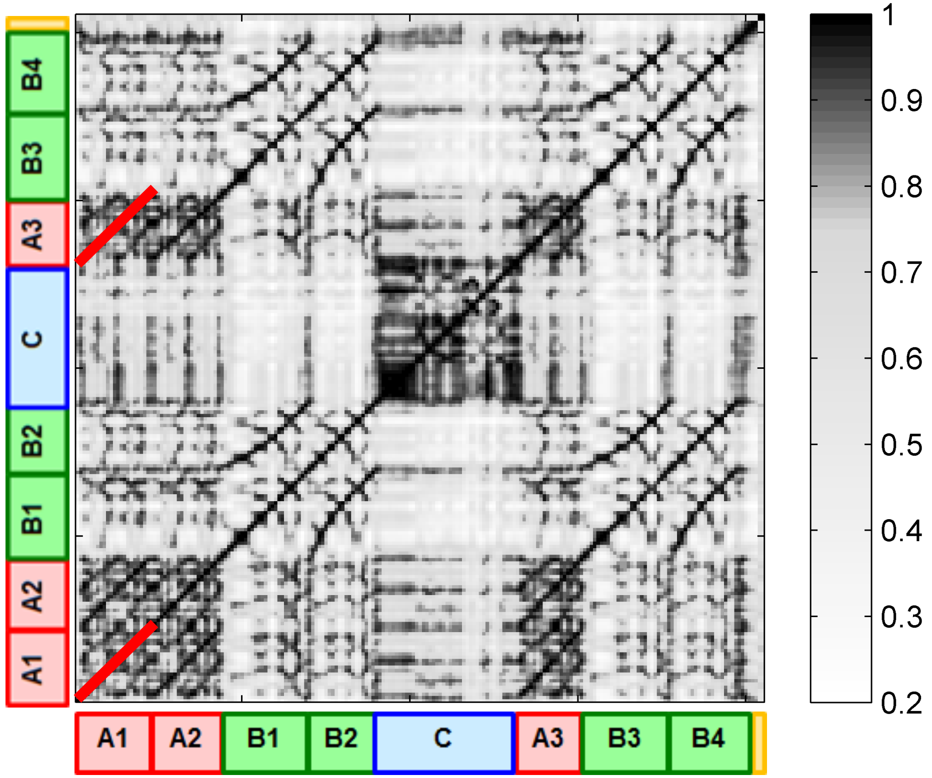
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



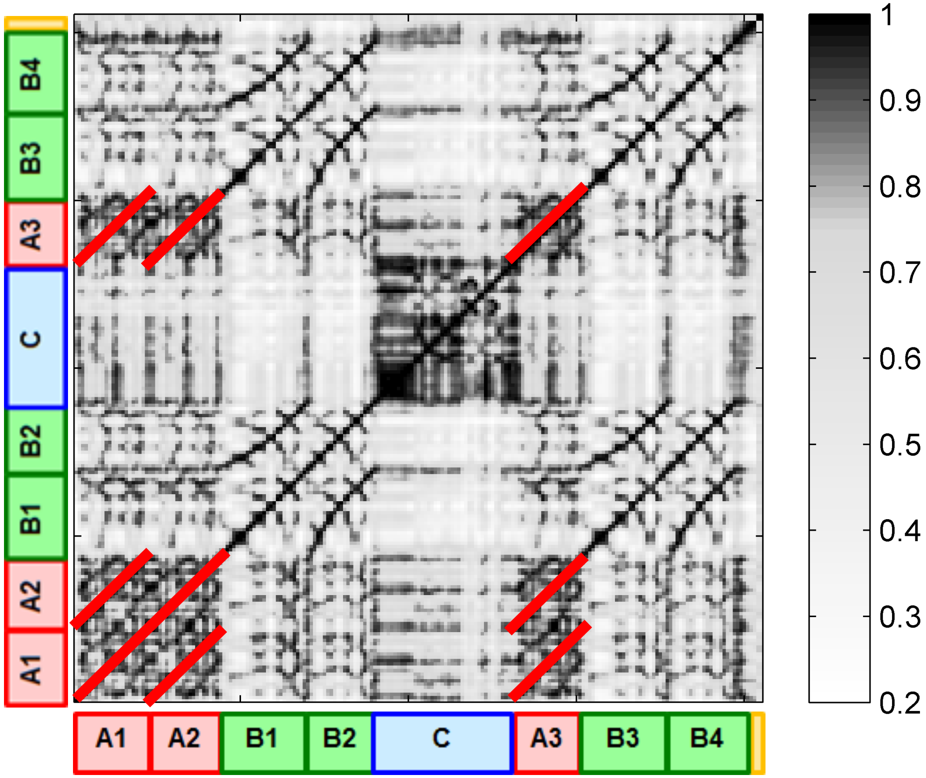
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



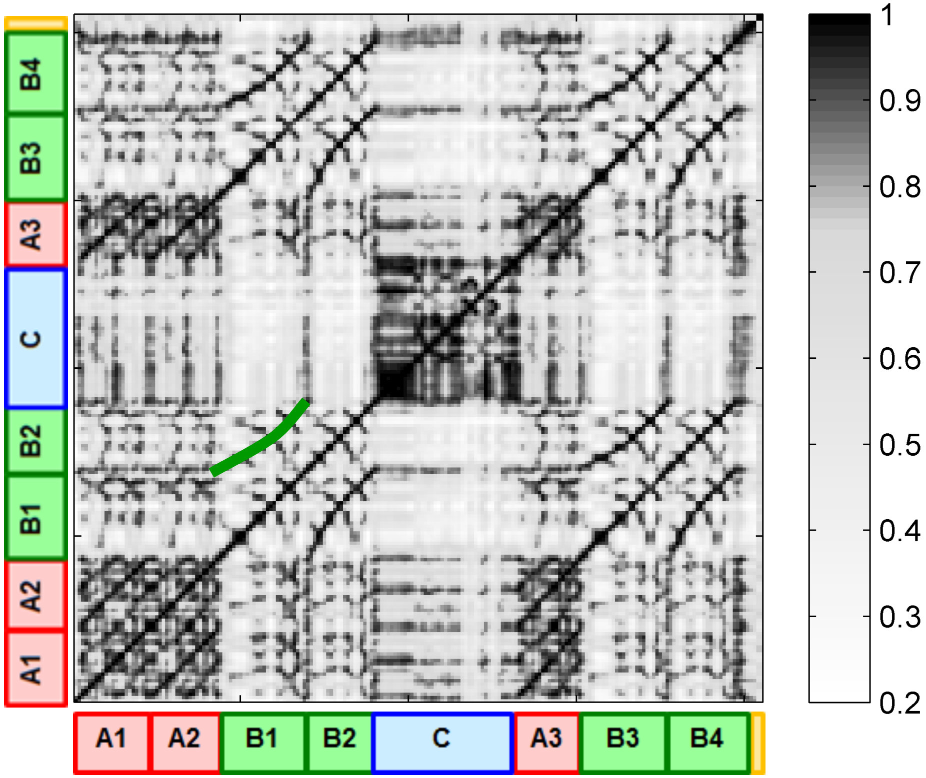
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



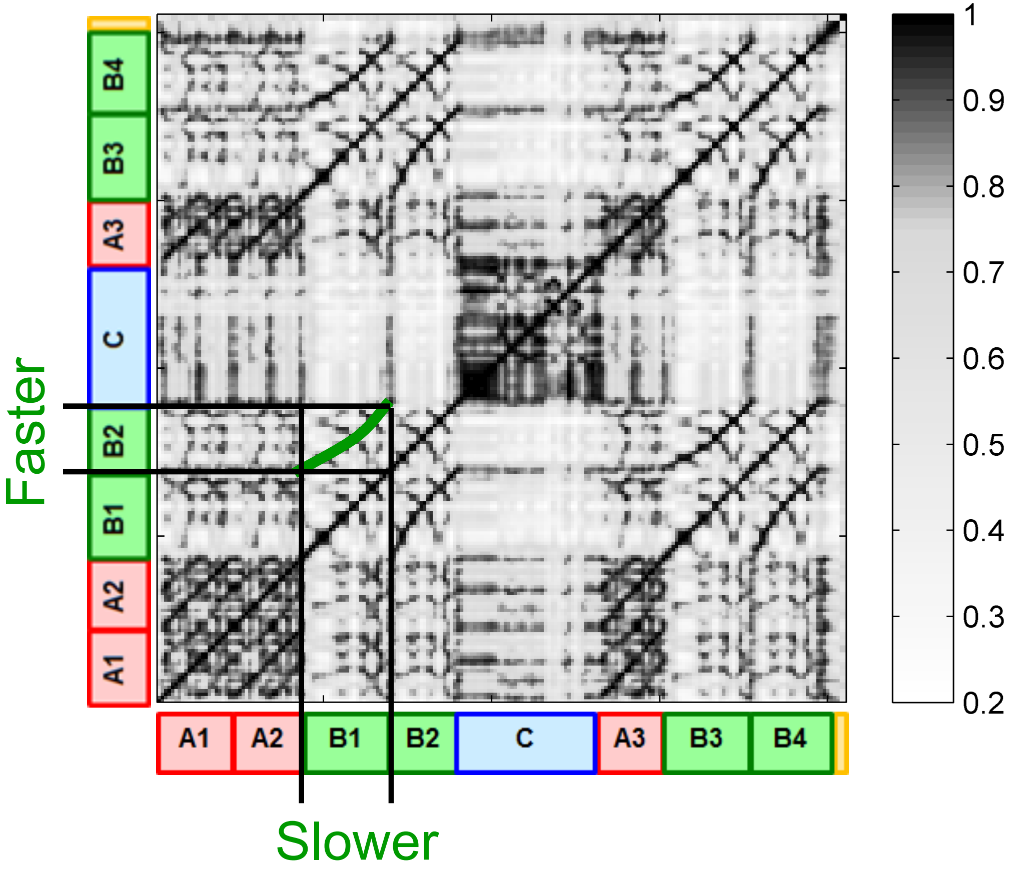
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



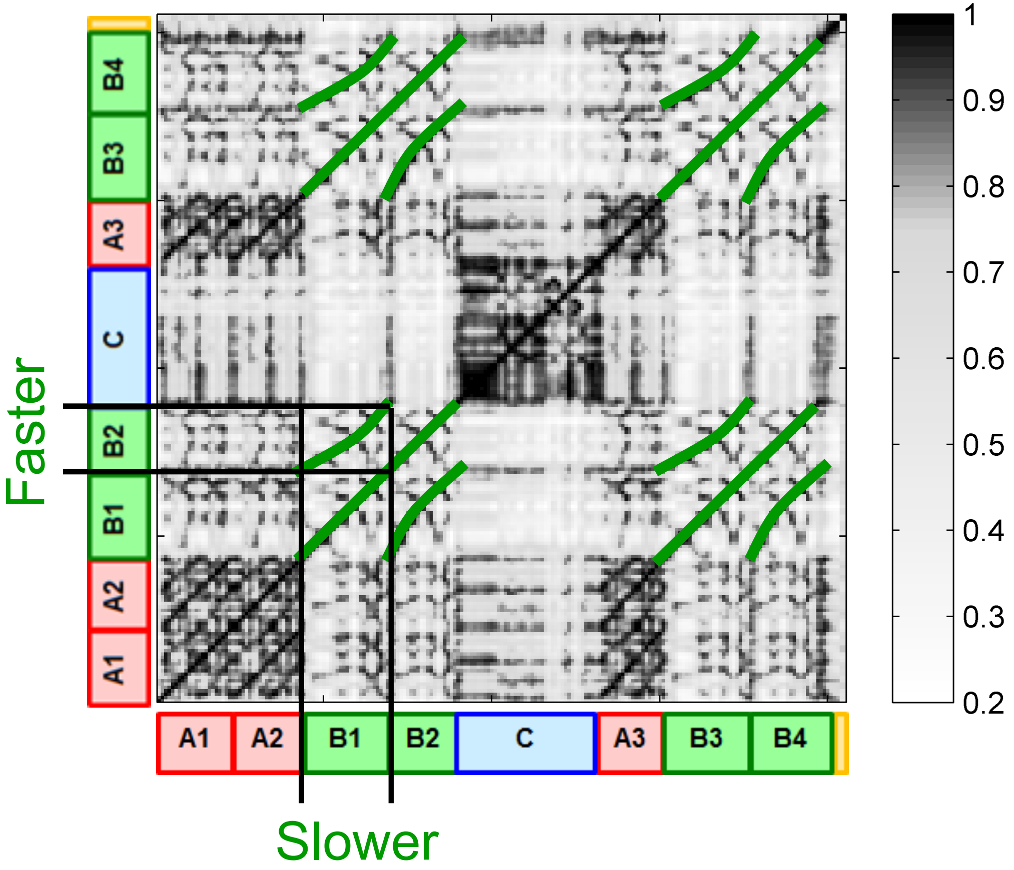
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)



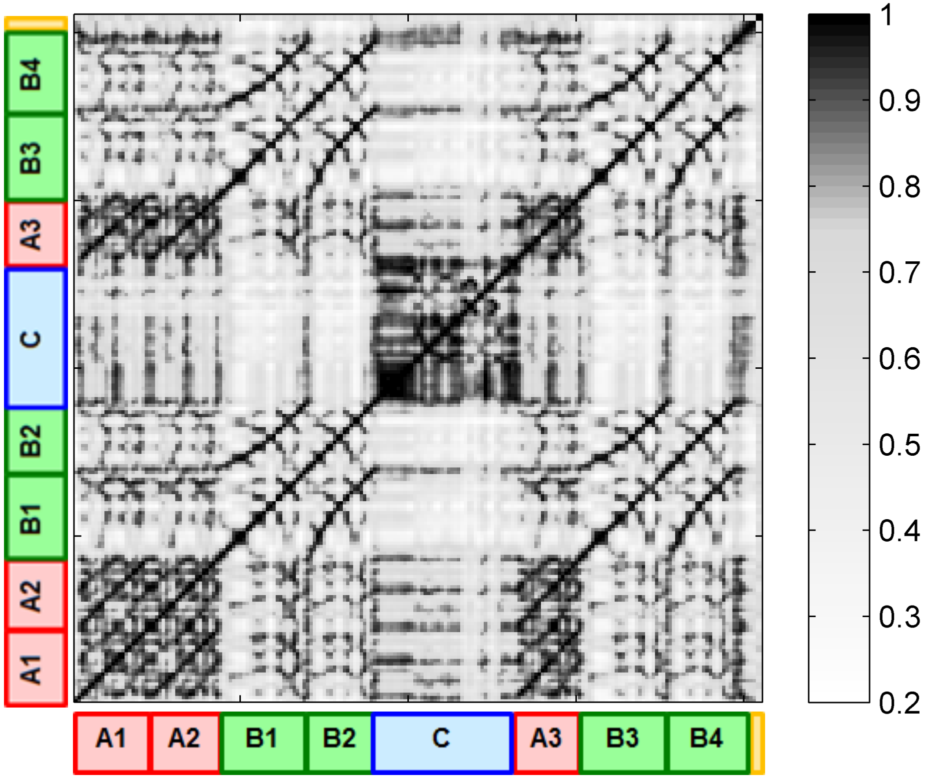
Self-Similarity Matrix

Example: Brahms Hungarian Dance No. 5 (Ormandy)

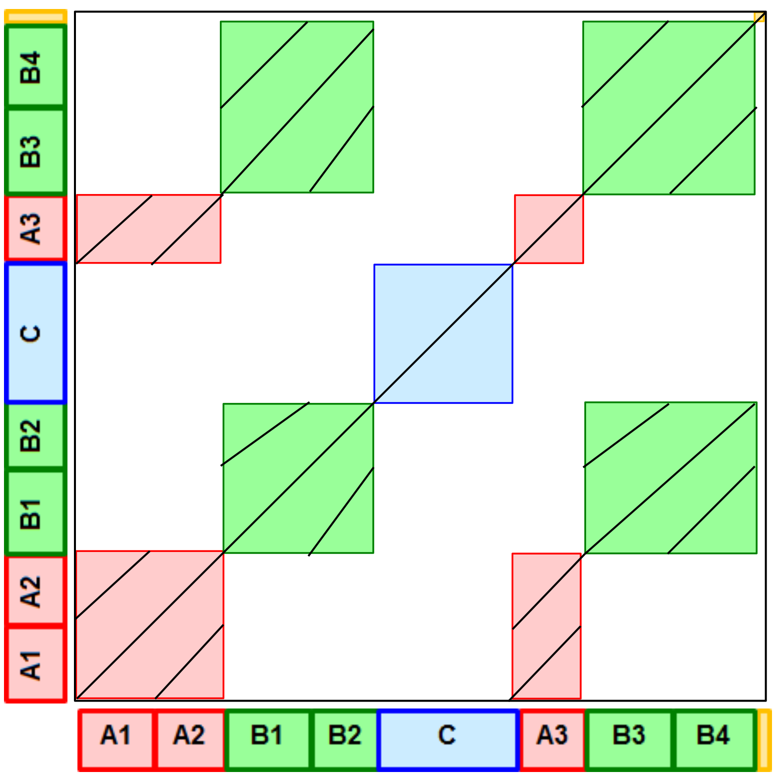


Self-Similarity Matrix

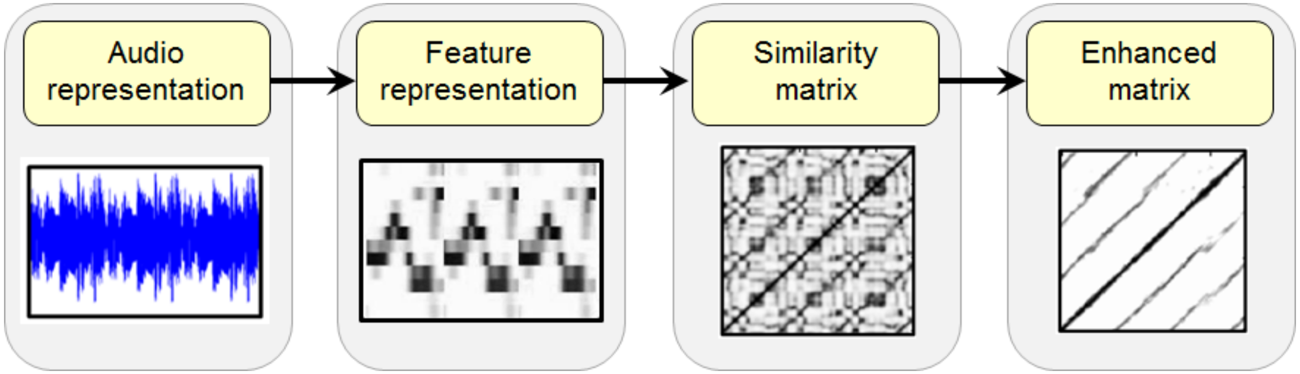
Example: Brahms Hungarian Dance No. 5 (Ormandy)



Idealized SSM



Similarity Matrix Toolbox



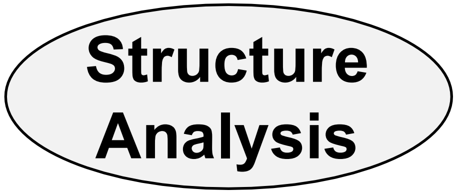
Meinard Müller, Nanzhu Jiang, Harald Grohganz
SM Toolbox: MATLAB Implementations for Computing and Enhancing Similarity Matrices

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

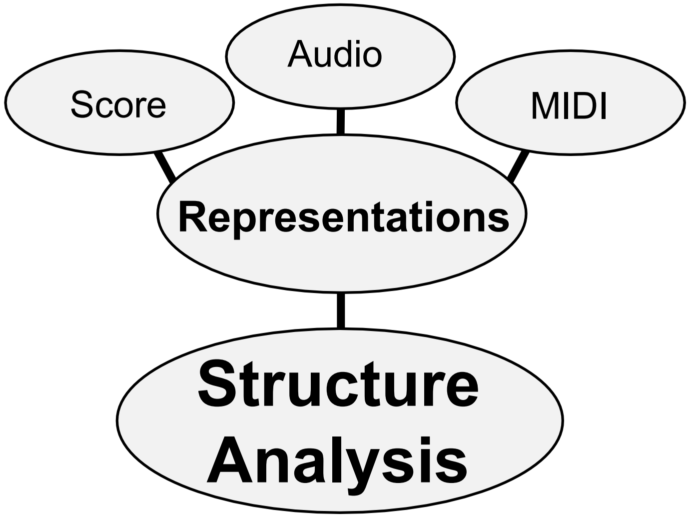
Demo

Code: https://github.com/stefan-balke/mpa-exc

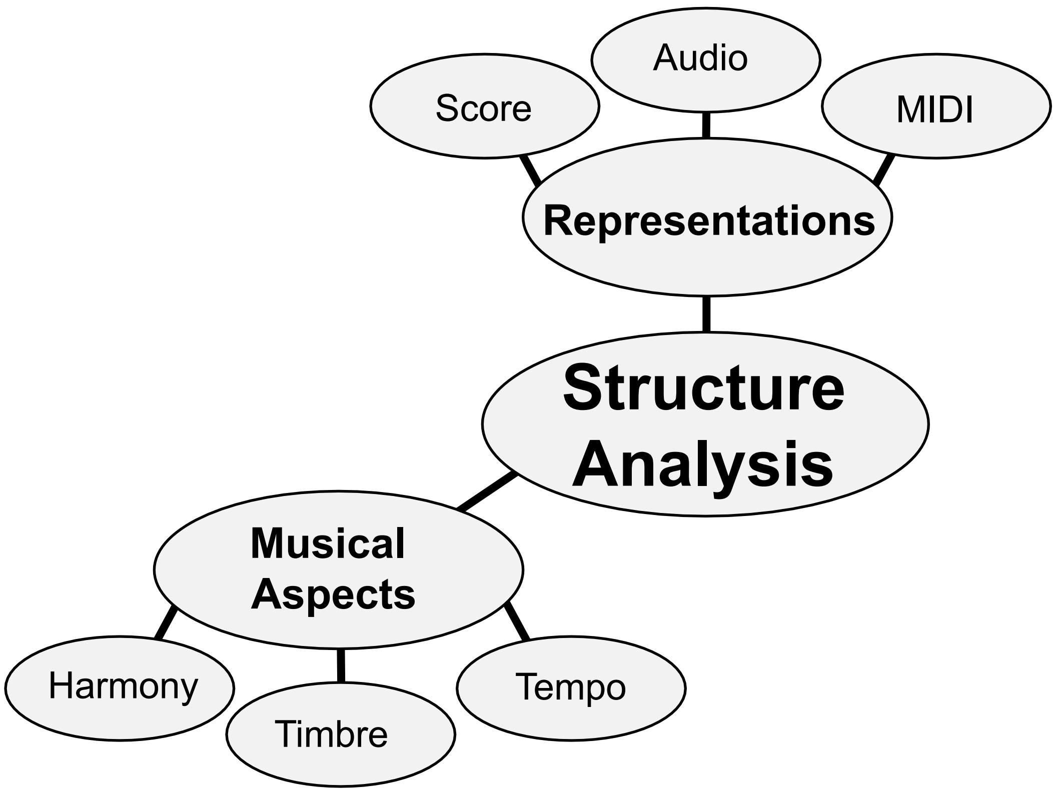
Conclusions



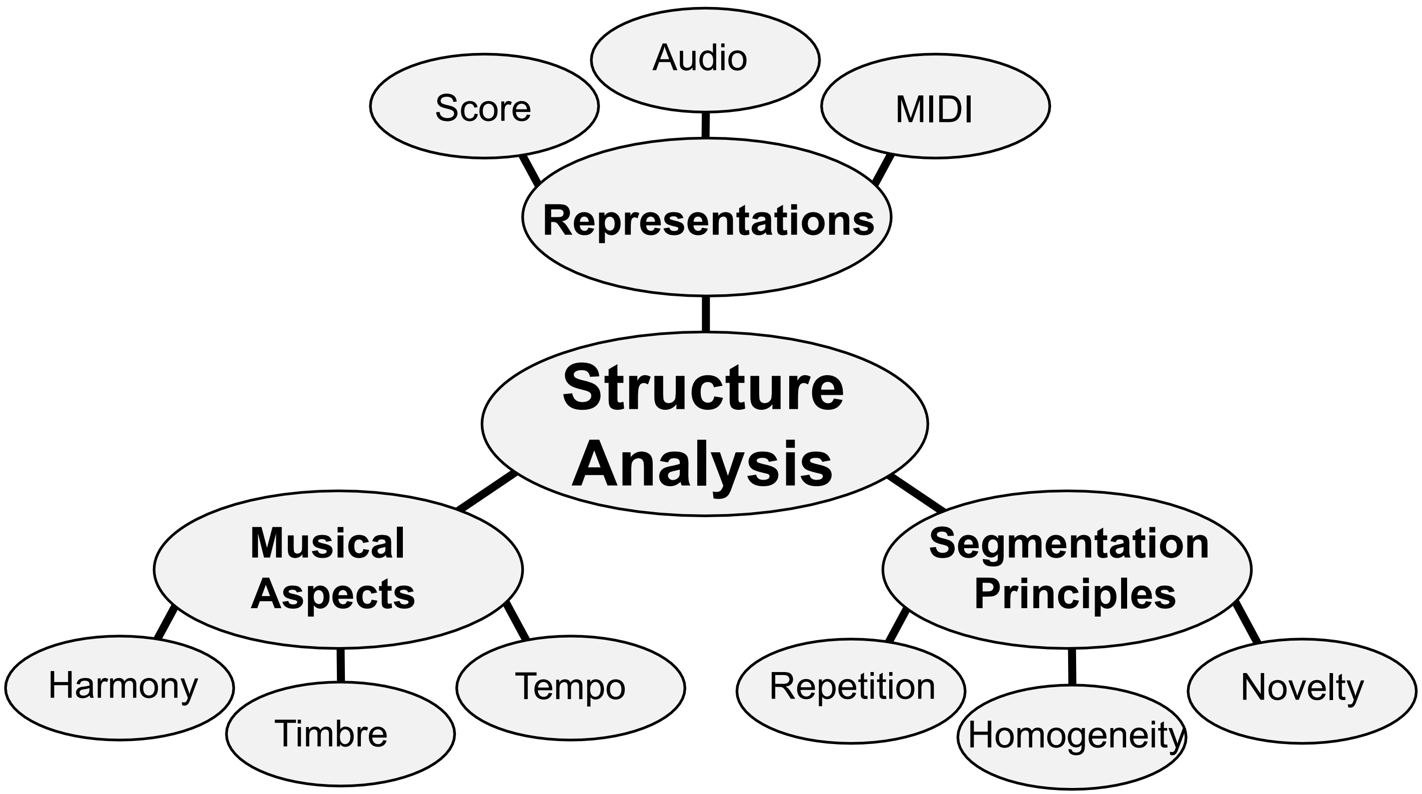
Conclusions



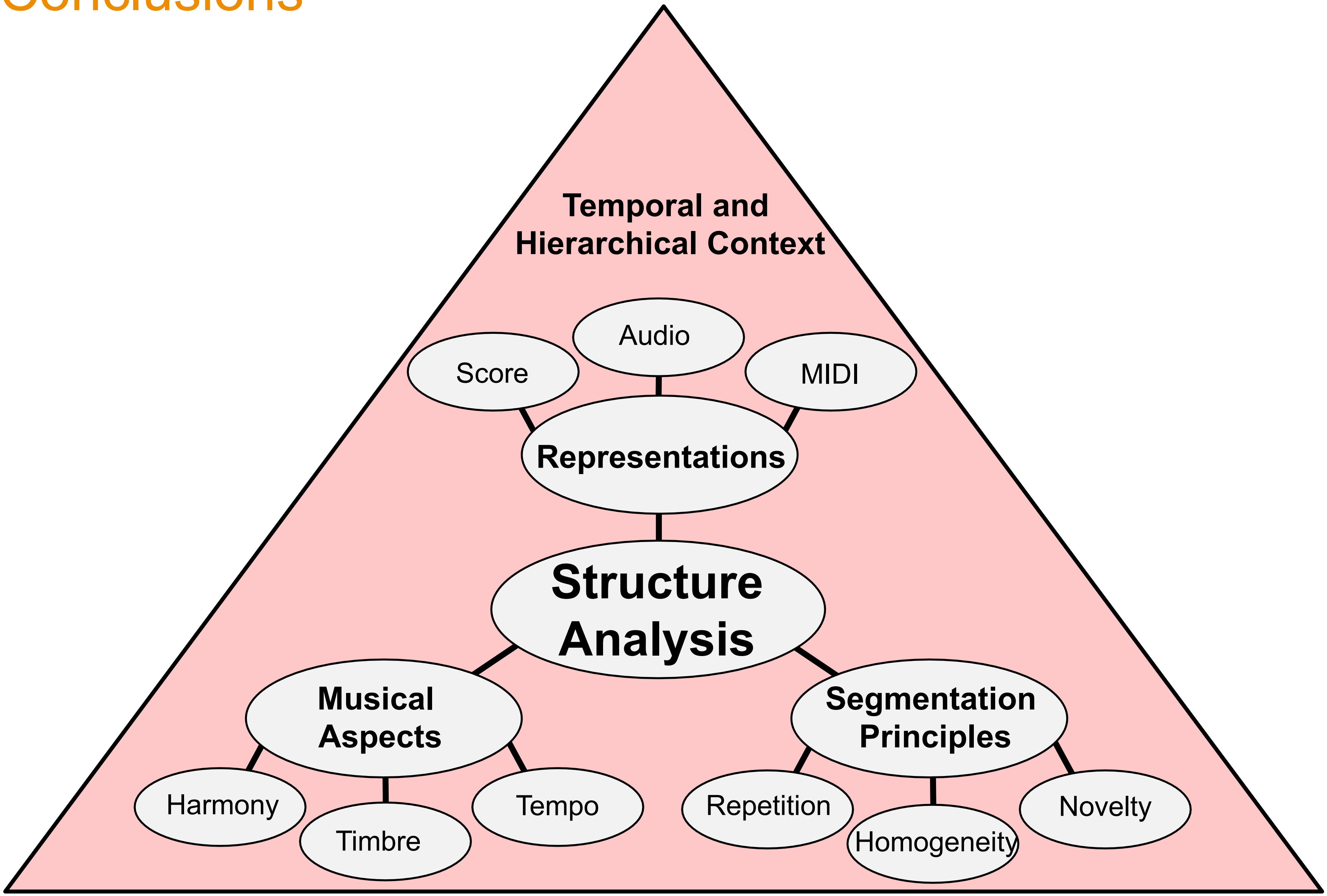
Conclusions



Conclusions



Conclusions



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.