INTERNATIONAL AUDIO LABORATORIES ERLANGEN
A joint institution of Fraunhofer IIS and Universität Erlangen-Nürnberg

AUDIO LABS

Learning with Music Signals: Technology Meets Education

Meinard Müller

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

Jahrestreffen GIBU, Dagstuhl
03. April 2023

FAU Friedrich-Alexander-Universität Erlangen-Nürnberg

Fraunhofer IIS

Meinard Müller

- Mathematics (Diplom/Master, 1997)
  Computer Science (PhD, 2001)
  Information Retrieval (Habilitation, 2007)
- Senior Researcher (2007-2012)
- Professor Semantic Audio Processing (since 2012)
- Former President of the International Society for Music Information Retrieval (MIR)
- IEEE Fellow for contributions to Music Signal Processing

universitätbonn

max planck institut informatik

FAU

ISMIR

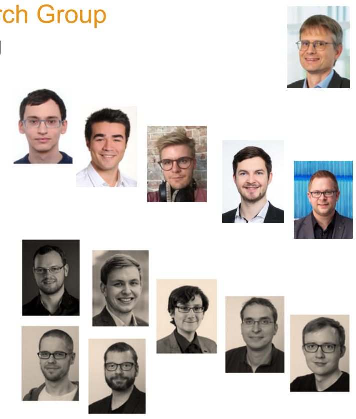
IEEE

Meinard Müller: Research Group

Semantic Audio Processing

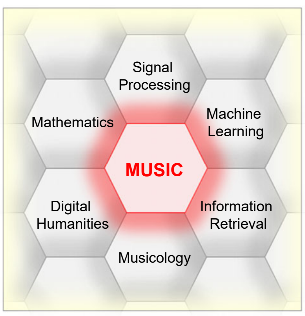
- Michael Krause
- Yigitcan Özer
- Simon Schwär
- Johannes Zeitler
- Peter Meier (external)

- Christof Weiß
- Sebastian Rosenzweig
- Frank Zalkow
- Christian Dittmar
- Stefan Balke
- Jonathan Driedger
- Thomas Prätzlich
- …



Music Processing

Music Processing: A Multifaceted Research Area

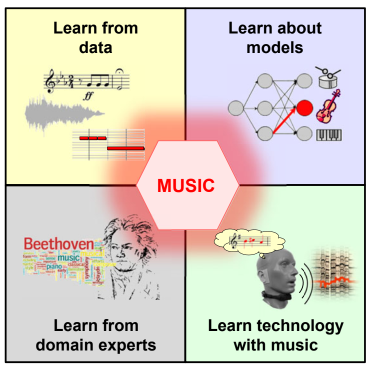


Music …

- important part of our lives …
- … Spotify, Pandora, iTunes, …
- interdisciplinary research
- intuitive entry point to education

Reinhart Koselleck-Projekt: LEARN

Learning with Music Signals: Technology Meets Education

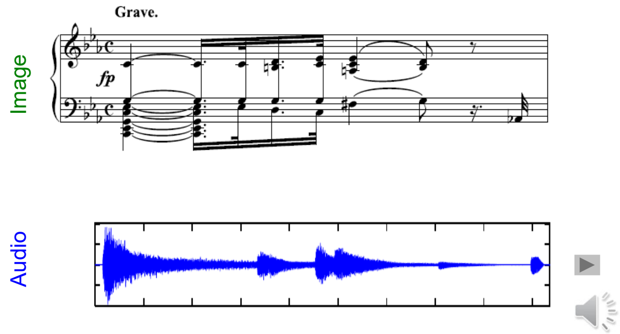


- Machine learning for music signal processing
- Interpretable models and knowledge integration
- Music understanding and applications
- Interactive learning in engineering through music

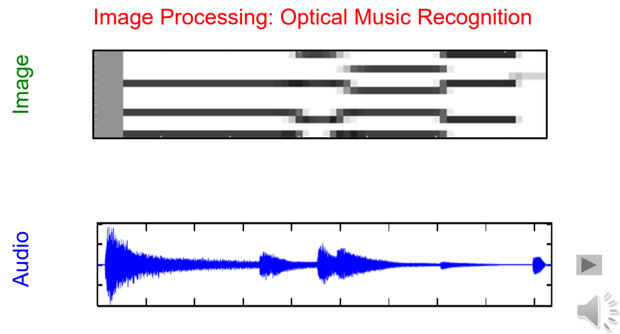
Multimodal Music Processing

Multimodal Music Processing

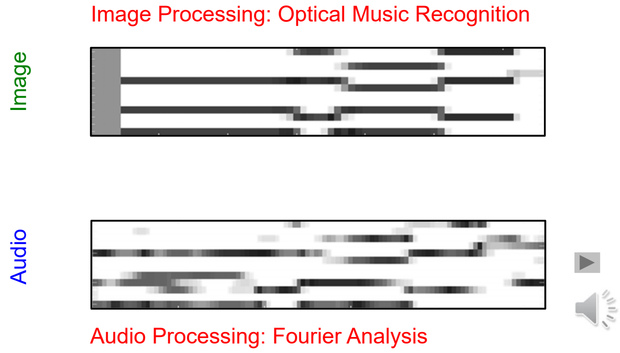
Multimodal Music Processing



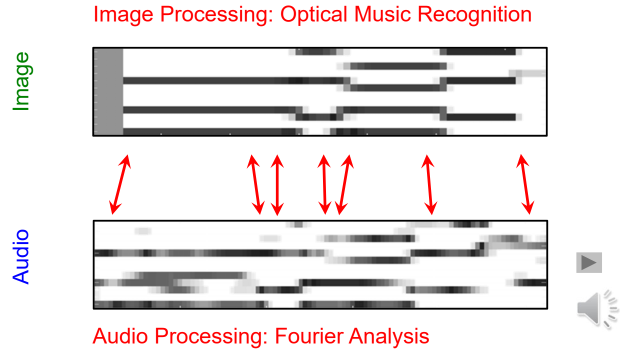
Multimodal Music Processing



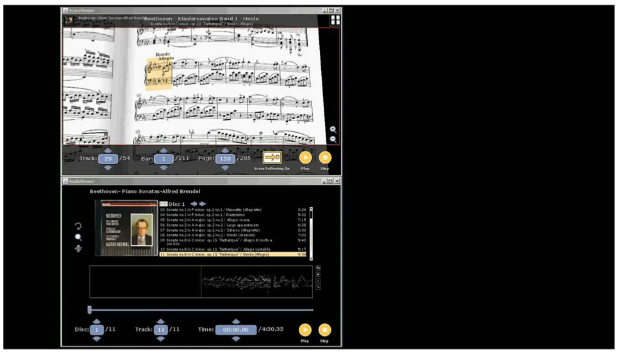
Multimodal Music Processing



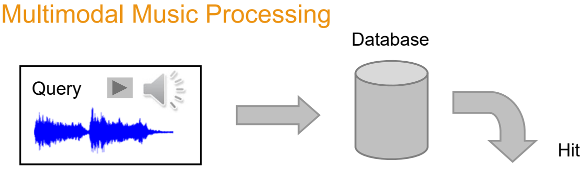
Multimodal Music Processing



Multimodal Music Processing



Multimodal Music Processing



Audio ID

Bernstein (1962) Beethoven, Symphony No. 5

Version ID

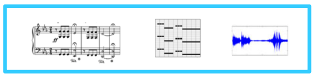
Beethoven, Symphony No. 5:
- Bernstein (1962)
- Karajan (1982)
- Gould (1992)

Category ID

- Beethoven, Symphony No. 9
- Beethoven, Symphony No. 3
- Haydn Symphony No. 94

Multimodal Music Processing

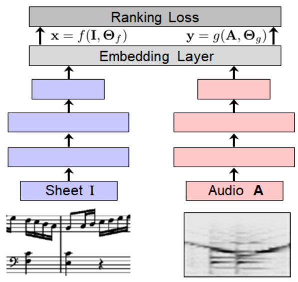
Modalities



Retrieval tasks:	**Specificity**	**Granularity**
Audio ID	High specificity	Fragment-based retrieval
Version ID	↕	↕
Category ID	Low specificity	Document-based retrieval

Reinhart Koselleck-Projekt: LEARN

- Representation learning
- Embedding techniques
- Weak annotations
- Loss functions
- …

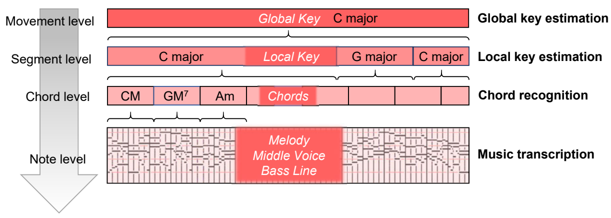


Computational Musicology

- Cooperation:
  - Rainer Kleinertz (Saarbrücken)
  - Stephanie Klauk (Saarbrücken)
  - Christof Weiß (Würzburg)
- Objectives
  - Harmony-based structural analysis
  - Beethoven Sonatas & Wagner's Ring
  - Interdisciplinary dialogue
- Since 2014: DFG-funded project

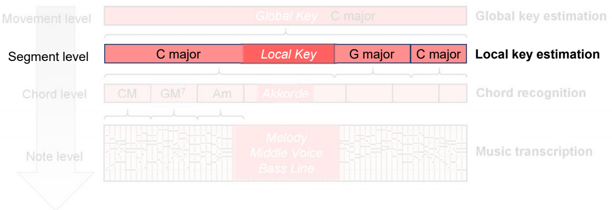
Computational Musicology: Harmony Analysis

- Different concepts
- Different temporal levels



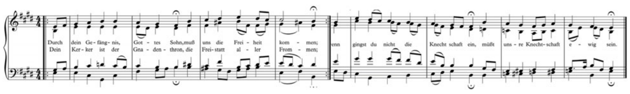
Computational Musicology: Harmony Analysis

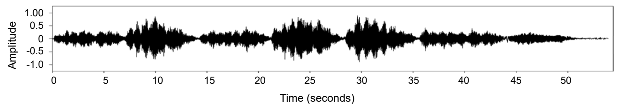
- Different concepts
- Different temporal levels



Local Key Estimation

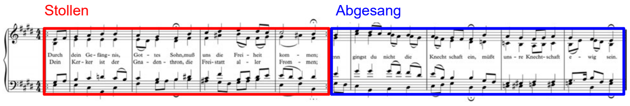
Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)

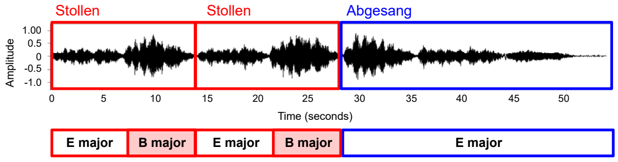




Local Key Estimation

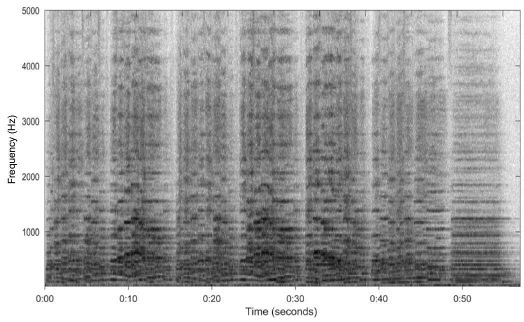
Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)





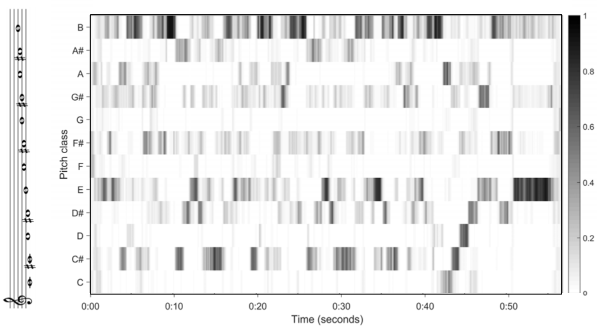
Local Key Estimation

Spectrogram



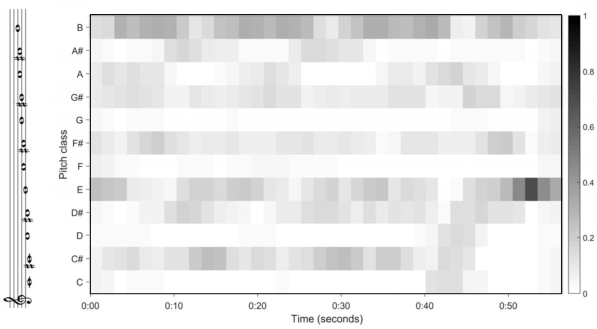
Local Key Estimation

Chromagram



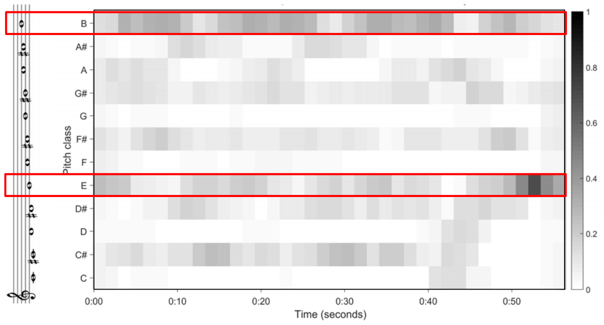
Local Key Estimation

Chromagram after smoothing



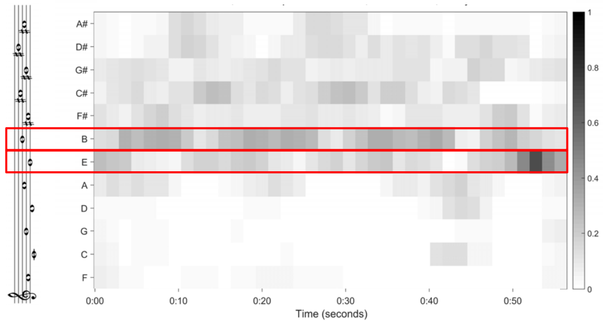
Local Key Estimation

Arrange pitch classes according to **perfect fifth** series



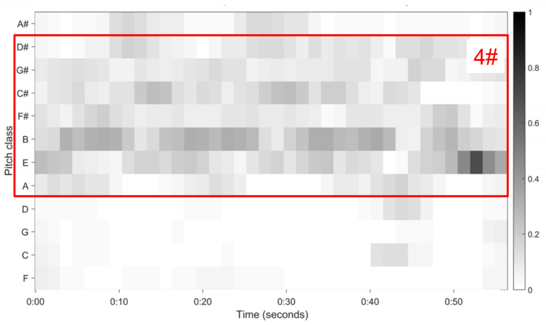
Local Key Estimation

Arrange pitch classes according to **perfect fifth** series



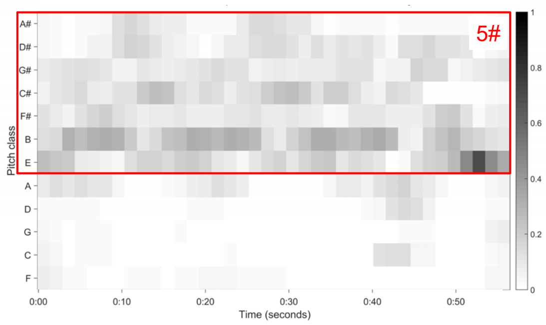
Local Key Estimation

Summarize pitch class content according to **diatonic scales**



Local Key Estimation

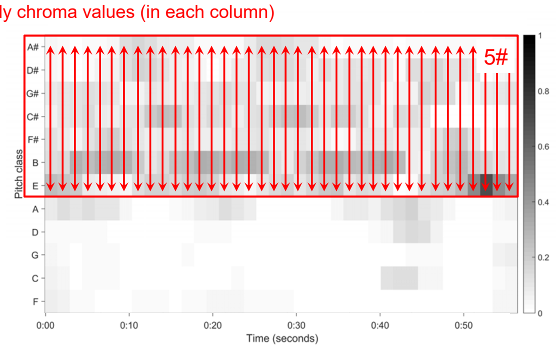
Summarize pitch class content according to **diatonic scales**



Local Key Estimation

Summarize pitch class content according to **diatonic scales**

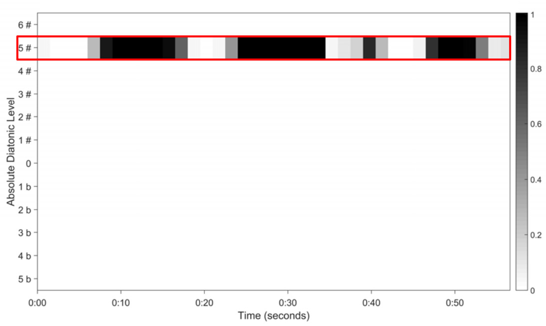
Multiply chroma values (in each column)



Local Key Estimation

Summarize pitch class content according to **diatonic scales**

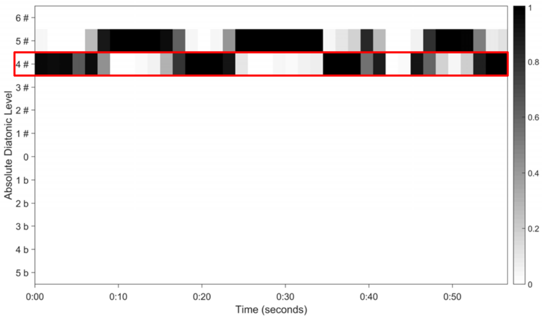
Multiply chroma values



Local Key Estimation

Summarize pitch class content according to **diatonic scales**

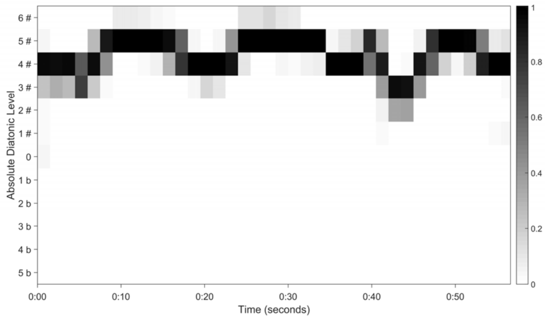
Multiply chroma values



Local Key Estimation

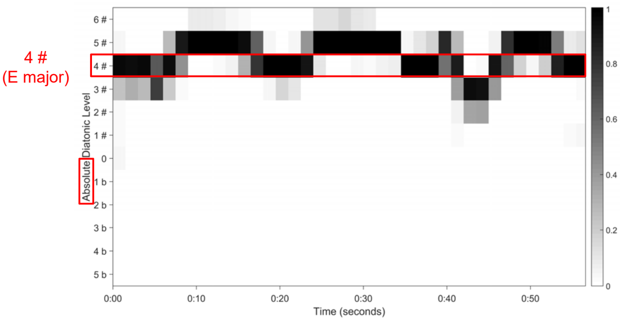
Summarize pitch class content according to **diatonic scales**

Multiply chroma values



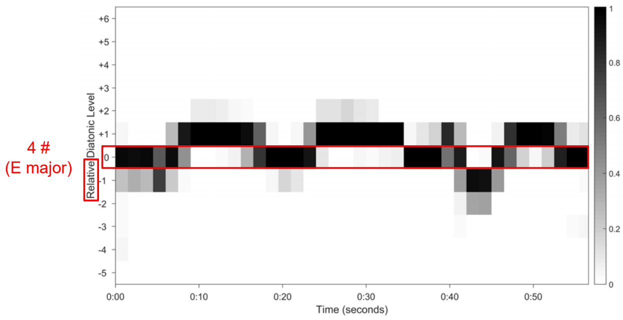
Local Key Estimation

Normalize representation relative to global key



Local Key Estimation

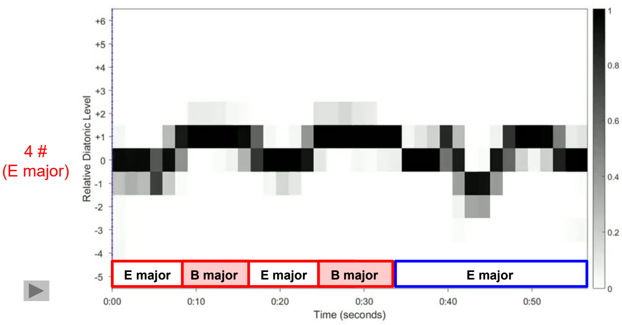
Normalize representation relative to global key



Local Key Estimation

J.S. Bach: Choral "Durch Dein Gefängnis" (*Johannespassion*)

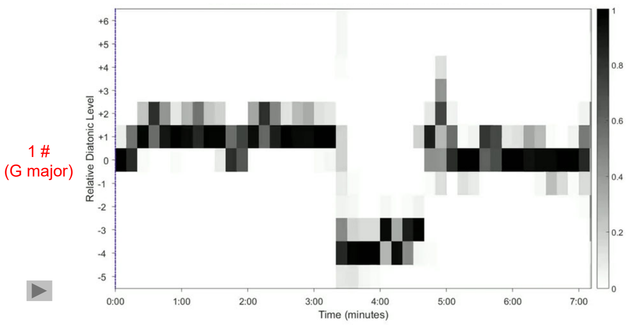
Recording: Scholars Baroque Ensemble, Naxos 1994



Local Key Estimation

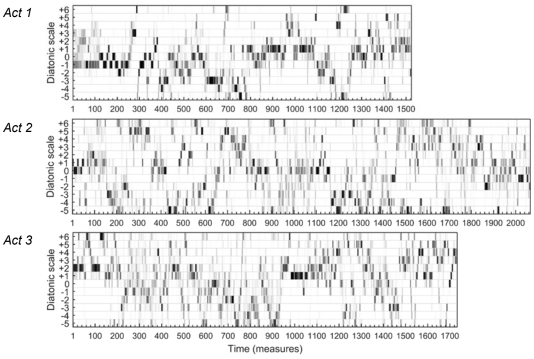
L. v. Beethoven: Piano Sonata No. 10 (Op. 14 Nr. 2), 1. Allegro

Recording: Barenboim, EMI 1998



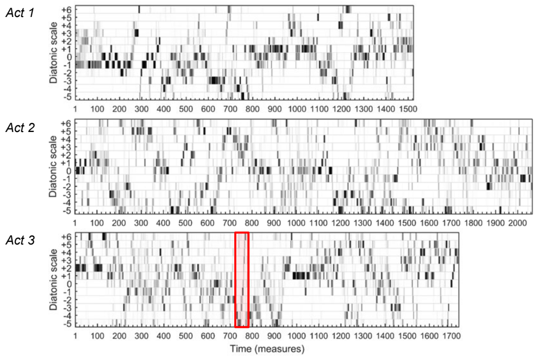
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)



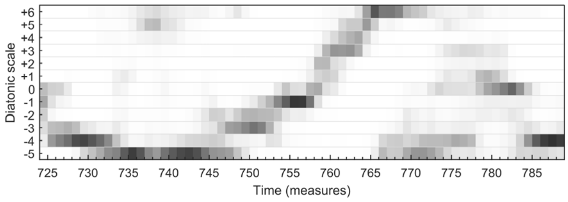
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)



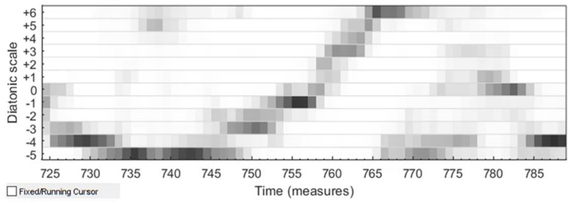
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)
Act 3, measure 724–789 (*Wotan's punishment*)



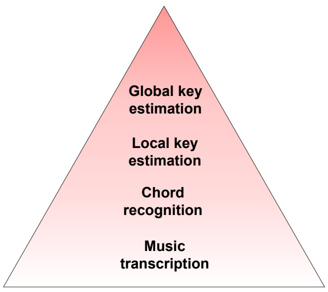
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)
Act 3, measure 724–789 (*Wotan's punishment*)



Reinhart Koselleck-Projekt: LEARN

- Interpretability & explainability
- Knowledge integration
- Hybrid models
- Multitask learning
- Hierarchical approaches
- …

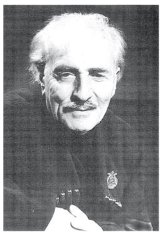


Computational Ethnomusicology: Traditional Georgian Vocal Music

- Interdisciplinary research project
  - Prof. Dr. Frank Scherbaum (Potsdam)
  - Dr. Nana Mzhavanadze (Tbilisi)
  - Sebastian Rosenzweig (FAU)
- Objective: Tonal analysis
- 2018 – 2022: DFG-funded project

Traditional Georgian Vocal Music

Example: Erkomaishvili corpus

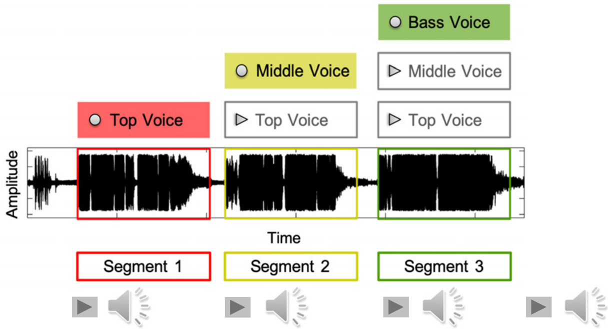


- Collection of traditional three-voice Georgian songs
- Performed by the former Georgian master chanter Artem Erkomaishvili (1887-1967)
- Recordings of 100 songs using tape recorders (1966)

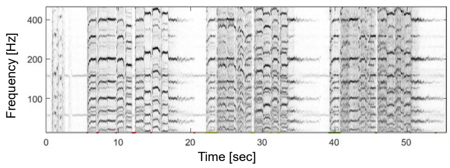
"*Original masterpieces of Georgian musical thinking.*" (Shugliashvili, 2014)

Traditional Georgian Vocal Music

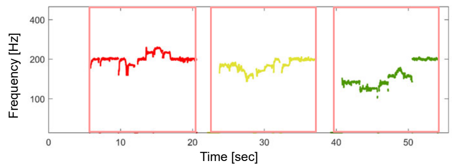
Example: Erkomaishvili corpus



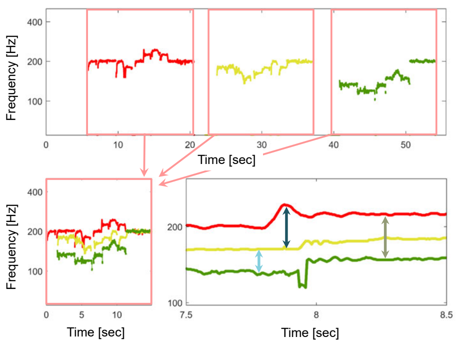
Traditional Georgian Vocal Music



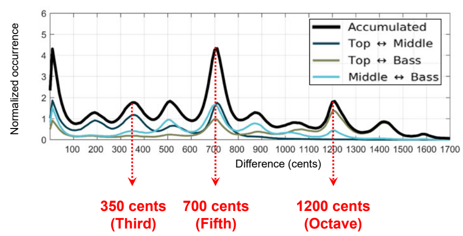
Traditional Georgian Vocal Music



Traditional Georgian Vocal Music



Traditional Georgian Vocal Music



- Peak at 350 cents (between minor and major third)
- **Non-western temperament**

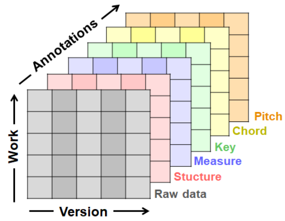
Traditional Georgian Vocal Music

- Recordings from field expedition in 2016
- 216 performances
- Multitrack audio + video
  - Room, HSM, LRX
- Total duration: 6 h

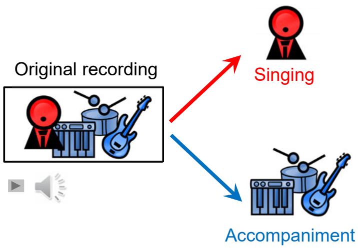
Room
Microphone

Reinhart Koselleck-Projekt: LEARN

- Non-standard datasets
  - Variety of music
  - Poor audio quality
  - Various sensor types
- Exploring DL models
  - Generalization
  - Overfitting
  - Data scarceness



Source Separation (Singing)



Reference
SP
AI

Reference
SP
AI

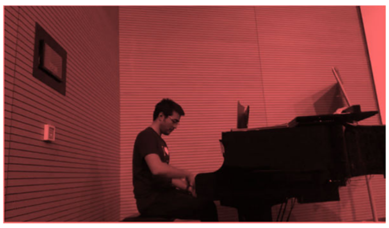
- Reference: Best possible result
- SP: Using traditional signal processing
- AI: Using data-driven approach

Source Separation (Piano Concerto)

- Yigitcan Özer
- PhD student in engineering
- Pianist

Source Separation (Piano Concerto)

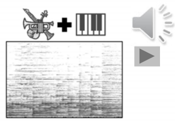
- Yigitcan Özer
- PhD student in engineering
- Pianist



Only Piano!

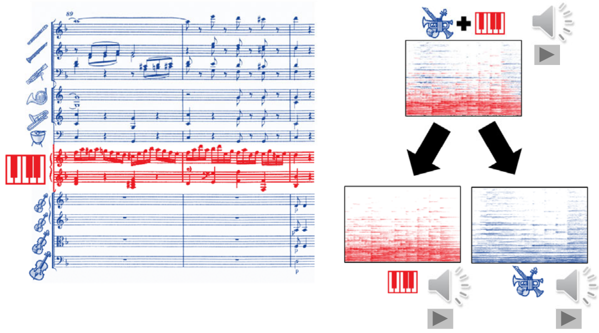
Where is the orchestra?

Source Separation (Piano Concerto)

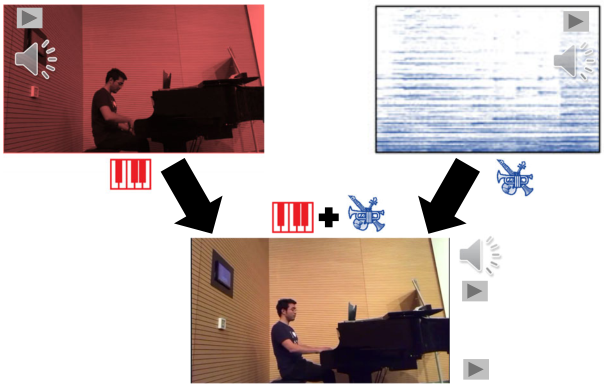


Source Separation (Piano Concerto)

Source Separation (Piano Concerto)

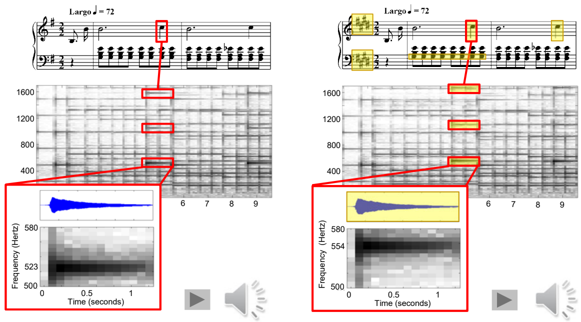


Source Separation (Piano Concerto)



Source Separation

Score-informed audio editing

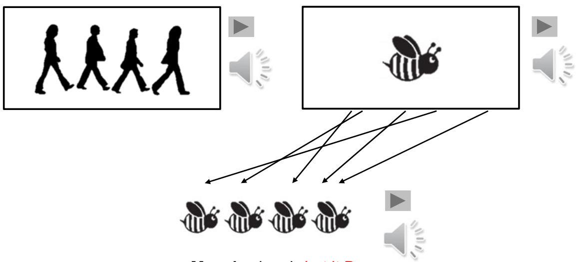


Source Separation

Audio mosaicing (style transfer)

Target signal: Beatles–Let it be

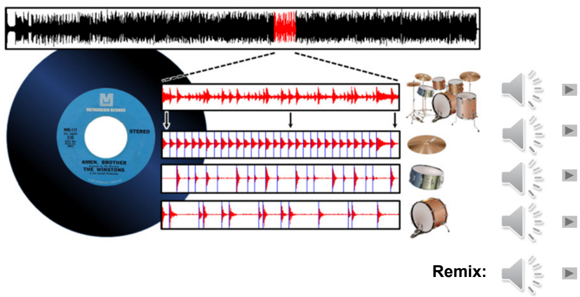
Source signal: Bees



Mosaic signal: Let it Bee

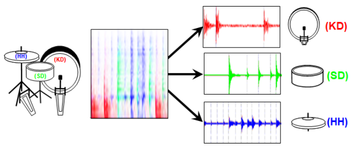
Source Separation

Informed Drum-Sound Decomposition

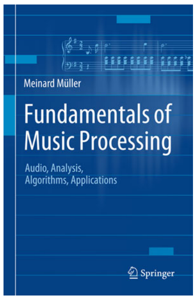


Reinhart Koselleck-Projekt: LEARN

- Reconstruction of sources
- Generative models
- Differentiable DSP
- Analysis by synthesis
- …



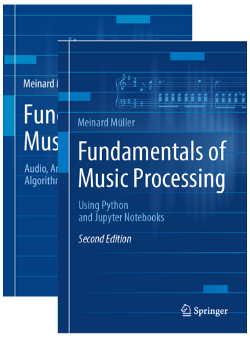
Fundamentals of Music Processing (FMP)



Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
Springer, 2015

Accompanying website:
www.music-processing.de

Fundamentals of Music Processing (FMP)



Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
Springer, 2015

Accompanying website:
www.music-processing.de

2nd edition
Meinard Müller
Fundamentals of Music Processing
Using Python and Jupyter Notebooks
Springer, 2021

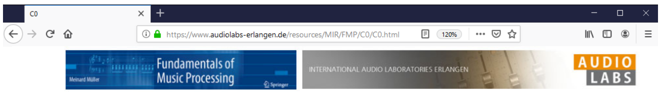
Fundamentals of Music Processing (FMP)

Chapter	Music Processing Scenario
1	Music Represenations
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
Springer, 2015

Accompanying website:
www.music-processing.de

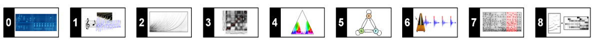
2nd edition
Meinard Müller
Fundamentals of Music Processing
Using Python and Jupyter Notebooks
Springer, 2021

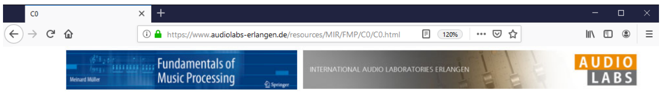


FMP Notebooks

Python Notebooks for Fundamentals of Music Processing

https://www.audiolabs-erlangen.de/FMP

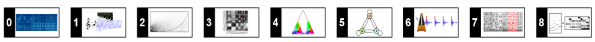




FMP Notebooks

Python Notebooks for Fundamentals of Music Processing

https://www.audiolabs-erlangen.de/FMP



Basics + 8 Chapters

Fundamentals of Music Processing

INTERNATIONAL AUDIO LABORATORIES ERLANGEN

FMP Notebooks

Python Notebooks for Fundamentals of Music Processing

https://www.audiolabs-erlangen.de/FMP

0 1 2 3 4 5 6 7 8

Basics + 8 Chapters

Tempo and Beat Tracking

Tempo and Beat Tracking

6

Tempo and Beat Tracking

6

Definition

We assume that we are given a discrete-time novelty function $\Delta : \mathbb{Z} \to \mathbb{R}$ ... indicate note onset candidates. The idea of Fourier analysis is to detect lo... in novelty curve by comparing it with windowed sinusoids. A high correlati... section of Δ with a windowed sinusoid indicates a periodicity of the sinus... (given a suitable phase). This correlation (along with the phase) can be co... short-time Fourier transform. To this end, we fix a window function $w : \mathbb{Z}$ ... length centered at $n = 0$ (e.g., a sampled Hann window). Then, for a frec... parameter $\omega \in \mathbb{R}_{\geq 0}$ and time parameter $n \in \mathbb{Z}$, the complex Fourier coe... is defined by

$$\mathcal{F}(n, \omega) := \sum_{m \in \mathbb{Z}} \Delta(m)\overline{w}(m - n)\exp(-2\pi i \omega m).$$

Tempo and Beat Tracking

6

Definition

Explanations

Theory

Mathematics

$$\mathcal{F}(n, \omega) := \sum_{m \in \mathbb{Z}} \Delta(m)\overline{w}(m - n)\exp(-2\pi i \omega m).$$

Tempo and Beat Tracking

6

Example: Shostakovich

In the following example, we consider an excerpt of a recording of Dimitri Shostakovich's Suite for Variety Orchestra No. 1. The score version of the excerpt.

0:00 / 0:06

We start with a spectral-based novelty function resampled to F_s^Δ ... Furthermore, we use a window size corresponding to 5 seconds (...

Tempo and Beat Tracking

6

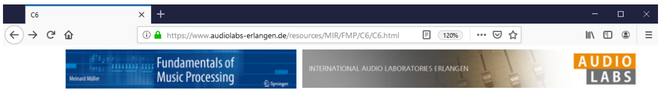
Example: Shostakovich

Annotations

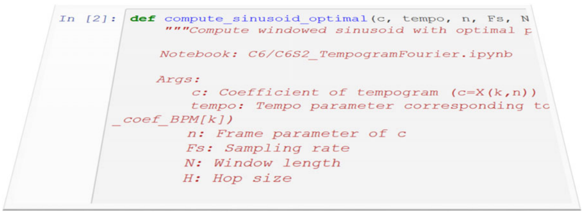
Music Examples

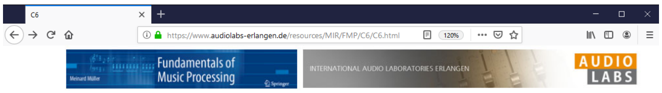
Audio

Links

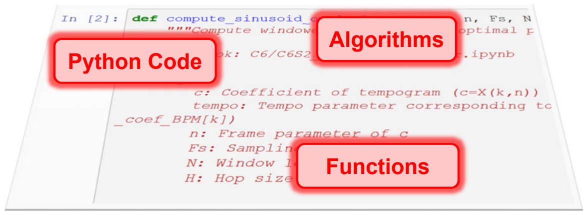


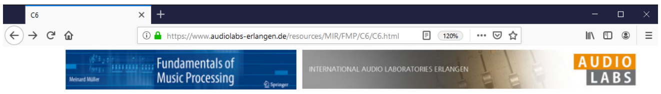
Tempo and Beat Tracking



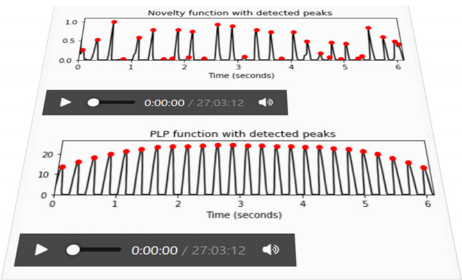


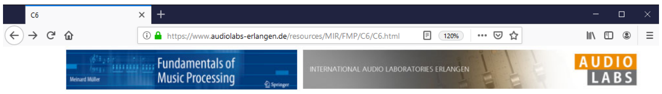
Tempo and Beat Tracking



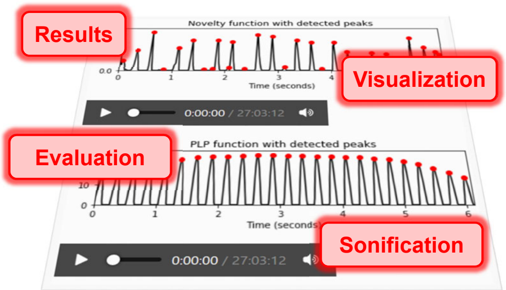


Tempo and Beat Tracking

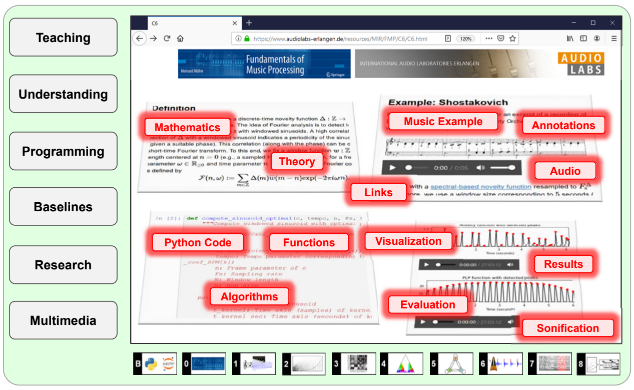




Tempo and Beat Tracking

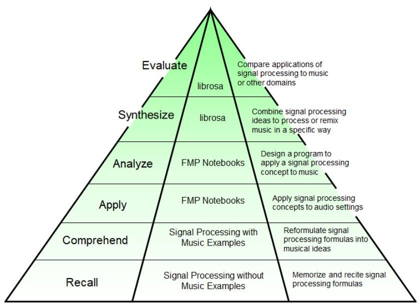


FMP Notebooks



Reinhart Koselleck-Projekt: LEARN

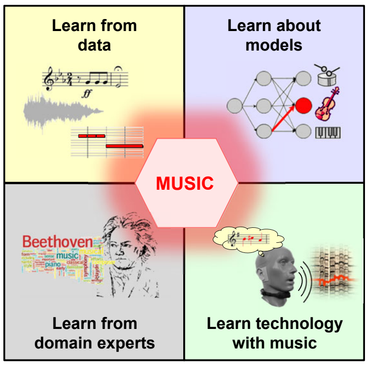
- Motivating examples
- Challenging domain
- Interdisciplinarity
- Structured learning
- …



Bloom's taxonomy

Reinhart Koselleck-Projekt: LEARN

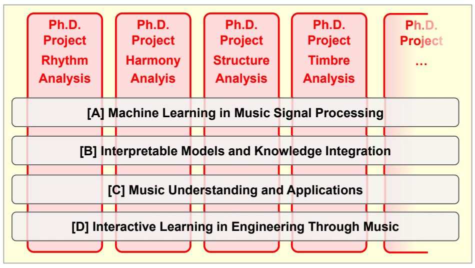
Learning with Music Signals: Technology Meets Education



- Machine learning for music signal processing
- Interpretable models and knowledge integration
- Music understanding and applications
- Interactive learning in engineering through music

Reinhart Koselleck-Projekt: LEARN

Learning with Music Signals: Technology Meets Education



References

- Meinard Müller: Fundamentals of Music Processing – Using Python and Jupyter Notebooks. 2nd Edition, Springer, 2021.
  https://www.springer.com/gp/book/9783030698072
- Meinard Müller and Frank Zalkow: libfmp: A Python Package for Fundamentals of Music Processing. Journal of Open Source Software (JOSS), 6(63): 1–5, 2021.
  https://joss.theoj.org/papers/10.21105/joss.03326
- Meinard Müller: An Educational Guide Through the FMP Notebooks for Teaching and Learning Fundamentals of Music Processing. Signals, 2(2): 245–285, 2021.
  https://www.mdpi.com/2624-6120/2/2/18
- Meinard Müller and Frank Zalkow: FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing. Proc. International Society for Music Information Retrieval Conference (ISMIR): 573–580, 2019.
  https://zenodo.org/record/3527872#.YOhEQOgzaUk
- Meinard Müller, Brian McFee, and Katherine Kinnaird: Interactive Learning of Signal Processing Through Music: Making Fourier Analysis Concrete for Students. IEEE Signal Processing Magazine, 38(3): 73–84, 2021.
  https://ieeexplore.ieee.org/document/9418542

Resources (Group Meinard Müller)

- FMP Notebooks:
  https://www.audiolabs-erlangen.de/FMP
- libfmp:
  https://github.com/meinardmueller/libfmp
- synctoolbox:
  https://github.com/meinardmueller/synctoolbox
- libtsm:
  https://github.com/meinardmueller/libtsm
- Preparation Course Python (PCP) Notebooks:
  https://www.audiolabs-erlangen.de/resources/MIR/PCP/PCP.html
  https://github.com/meinardmueller/PCP