

Annotations and Analyses in Computational Musicology: Separate it or not?

Meinard Müller, Vlora Arifi-Müller, Christof Weiß

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

Possibilities and Limitations of Digital Annotation Tools for
Audio-Visual Material with a focus on Sound and Music
Mainz, May 2–4, 2022

Meinard Müller



- Mathematics (Diplom/Master)
Computer Science (PhD)
Information Retrieval (Habilitation)
- Since 2012: Professor
Semantic Audio Processing
- Former President of the International Society for
Music Information Retrieval (MIR)
- IEEE Fellow for contributions to Music Signal Processing



Meinard Müller: Research Group

Semantic Audio Processing

- Christof Weiß
- Vlora Arifi-Müller
- Sebastian Rosenzweig
- Michael Krause
- Yigitcan Özer
- Simon Schwär
- Peter Meier (external)



International Audio Laboratories Erlangen



- Fraunhofer Institute for Integrated Circuits IIS
- Largest Fraunhofer institute with ≈ 1000 members
- Applied research for sensor, audio, and media technology



**AUDIO
LABS**



- Friedrich-Alexander Universität Erlangen-Nürnberg (FAU)
- One of Germany's largest universities with $\approx 40,000$ students
- Strong Technical Faculty

International Audio Laboratories Erlangen



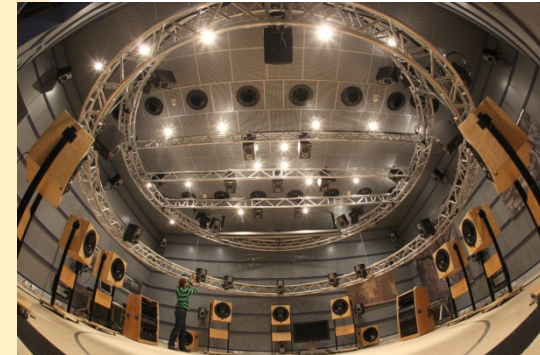
Audio

International Audio Laboratories Erlangen

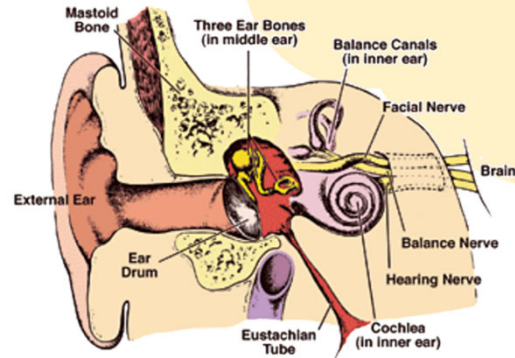
Audio Coding



3D Audio



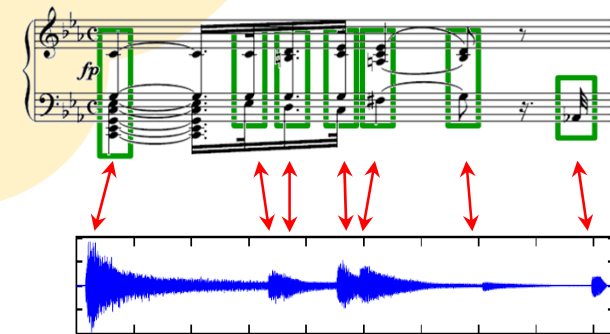
Audio



Psychoacoustics



Internet of Things



Music Processing

AudioLabs – FAU

- Prof. Dr. Jürgen Herre
Audio Coding
- Prof. Dr. Bernd Edler
Audio Signal Analysis
- Prof. Dr. Meinard Müller
Semantic Audio Processing
- Prof. Dr. Emanuël Habets
Spatial Audio Signal Processing
- Prof. Dr. Nils Peters
Audio Signal Processing
- Dr. Stefan Turowski
Coordinator AudioLabs-FAU

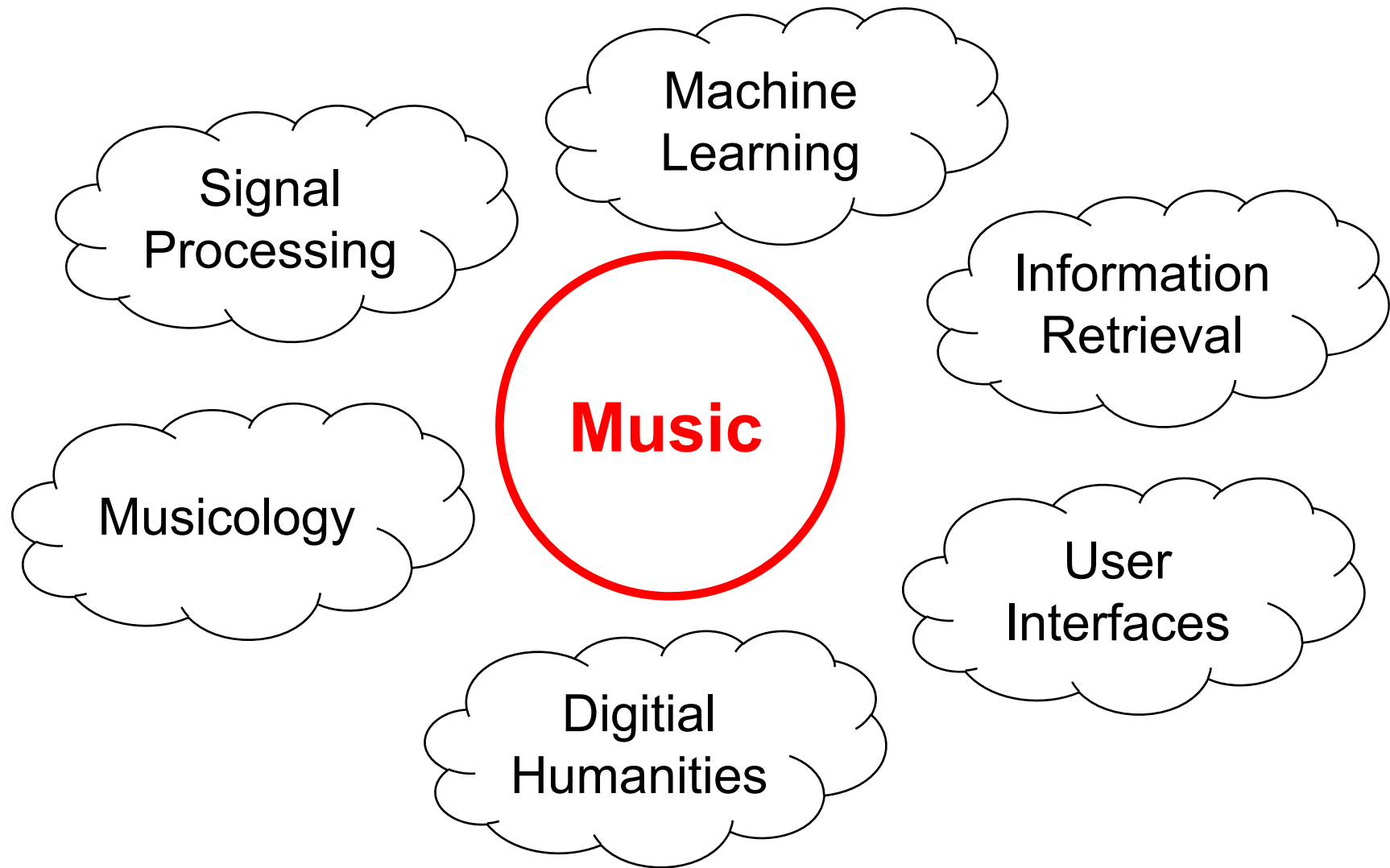




Music



Music Information Retrieval (MIR)

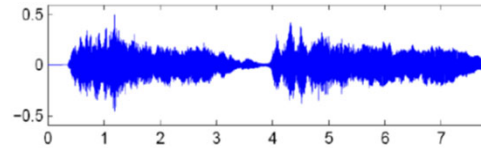


Music Information Retrieval (MIR)

Sheet Music (Image)



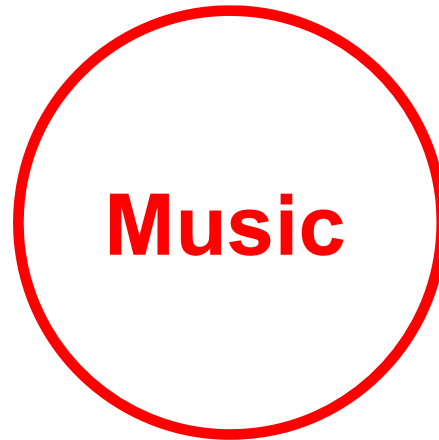
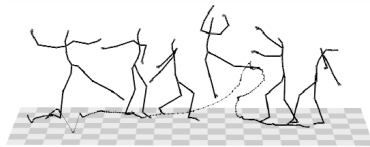
CD / MP3 (Audio)



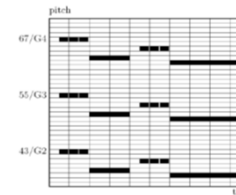
MusicXML (Text)

```
<note>  
  <pitch>  
    <step>E</step>  
    <alter>-1</alter>  
    <octave>4</octave>  
  </pitch>  
  <duration>2</duration>  
  <type>half</type>  
</note>
```

Dance / Motion (Mocap)



MIDI



Singing / Voice (Audio)



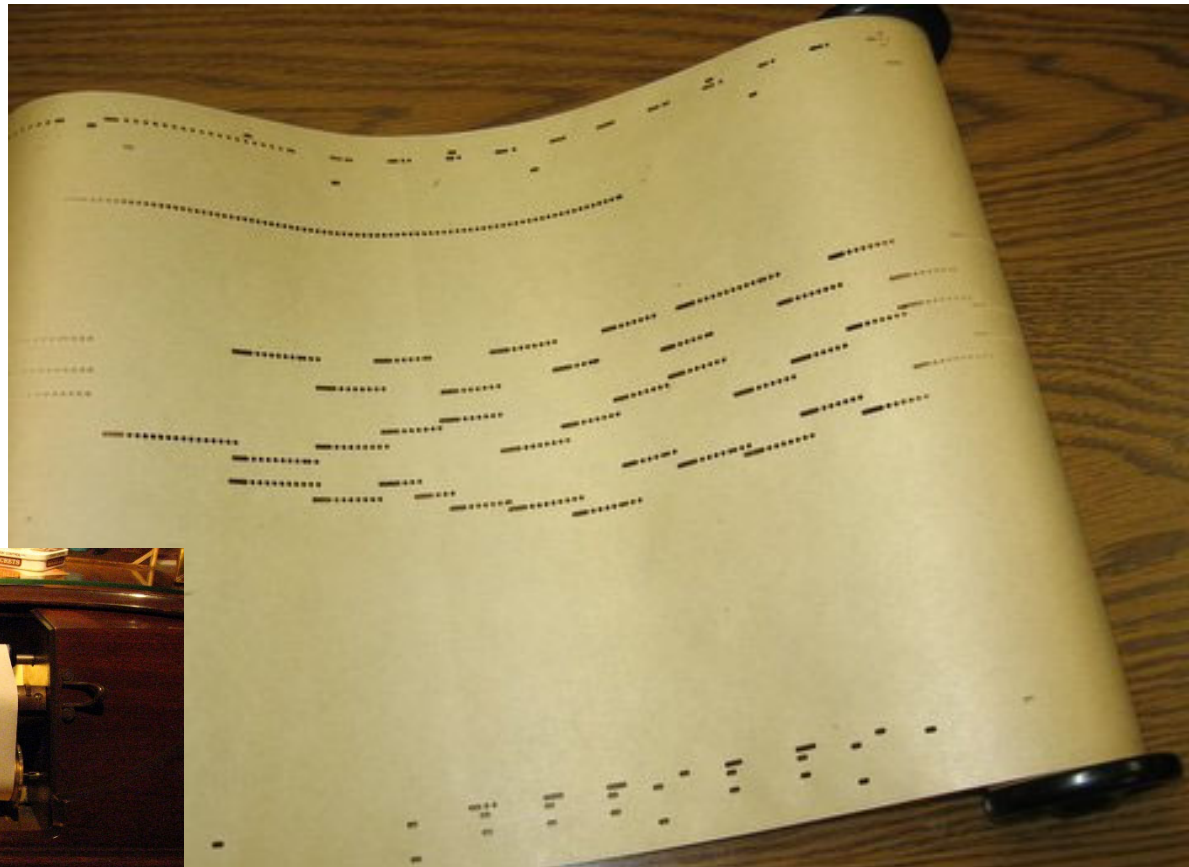
Music Film (Video)



Music Literature (Text)

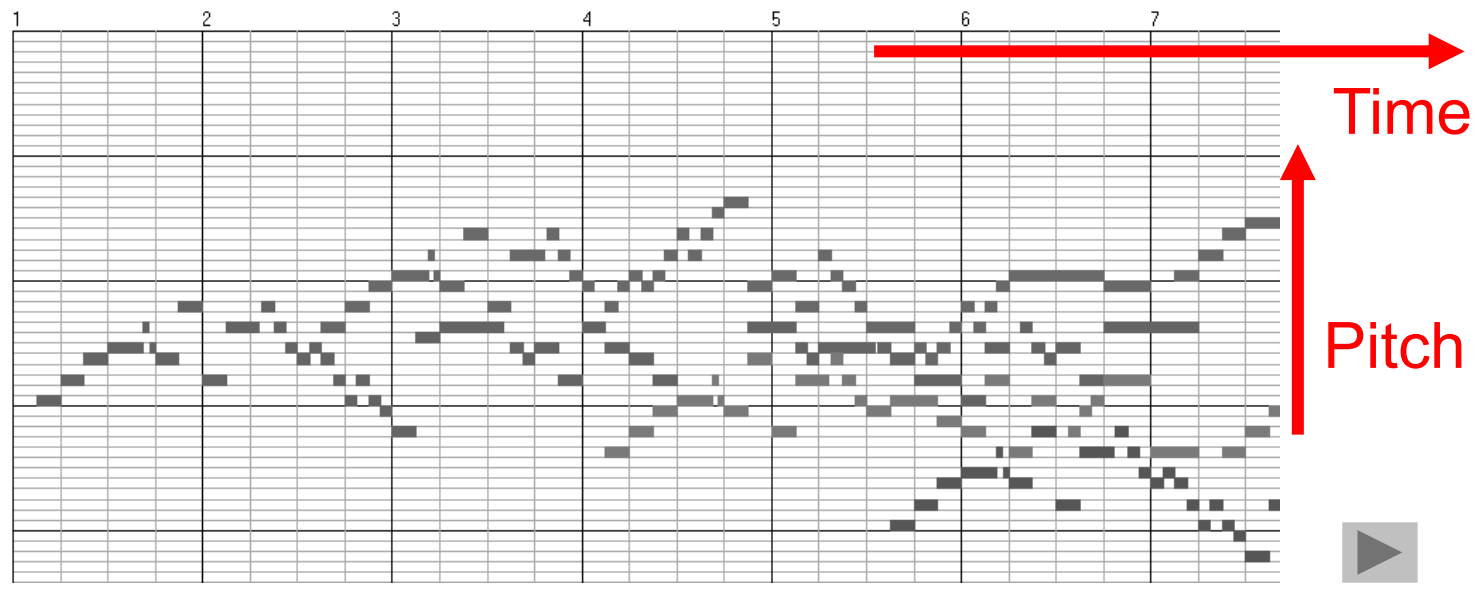
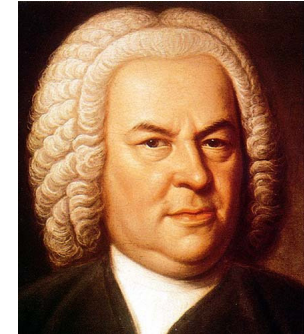


Piano Roll Representation (1900)



Piano Roll Representation

J.S. Bach, C-Major Fuge
(Well Tempered Piano, BWV 846)

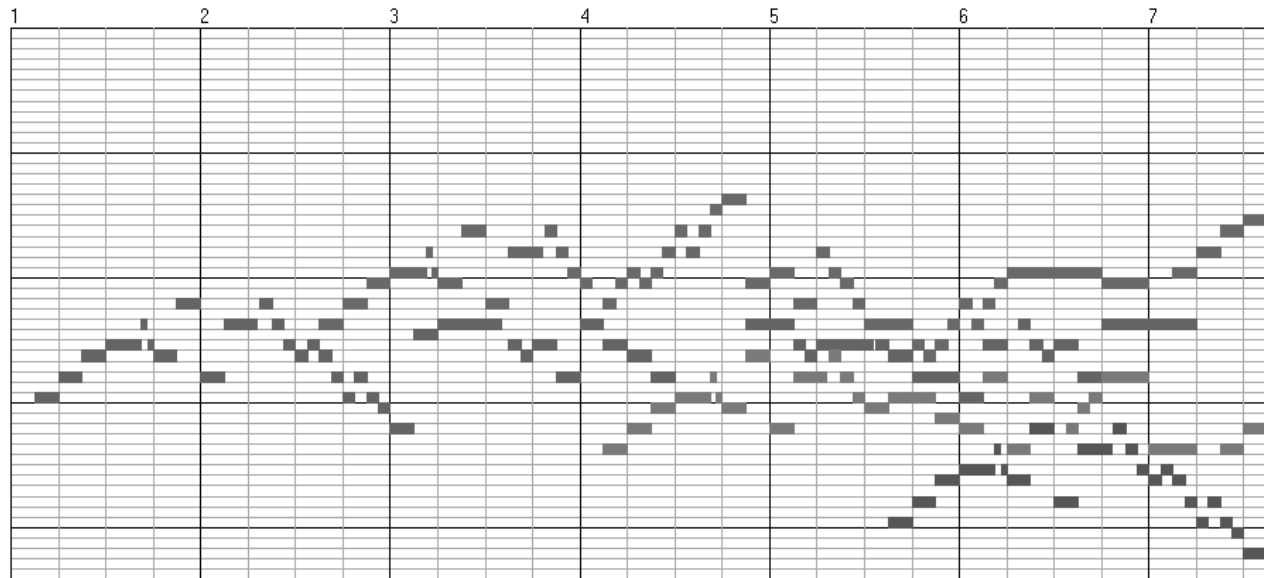
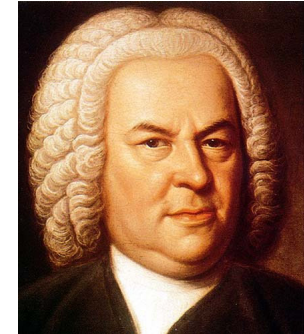


Piano Roll Representation

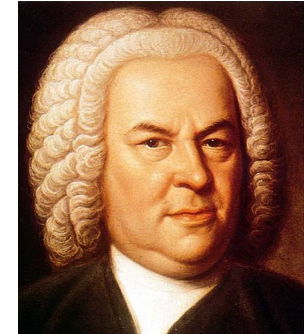
Query:



Goal: Find all occurrences of the query



Piano Roll Representation

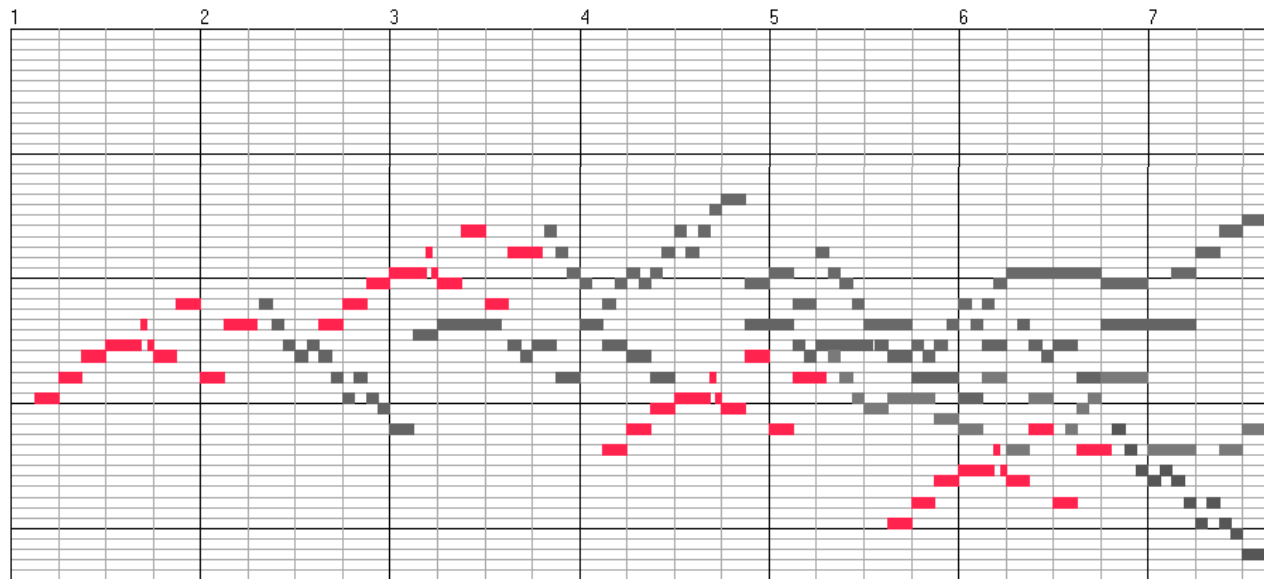


Query:

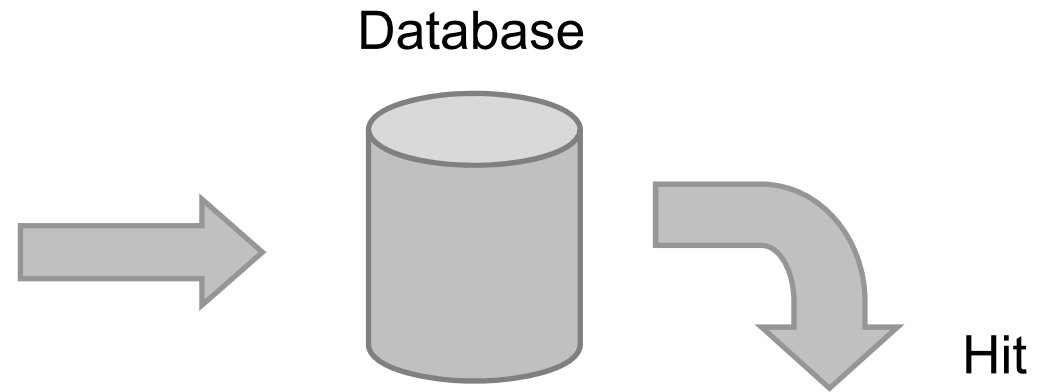


Goal: Find all occurrences of the query

Matches:




Music Retrieval




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
- 

Score Following

The screenshot displays two windows from a music software application. The top window, titled "ScoreViewer", shows a digital score for "Beethoven - Klaviersonaten Band 1 - Henle". The score is for "Sonata no.8 in C minor, op.13 'Pathétique' / Rondo (Allegro)". The score is currently on page 159 of 285, track 29 of 54, and bar 1 of 211. The score is highlighted in yellow, and the text "Rondo Allegro" is visible. The bottom window, titled "AudioViewer", shows a tracklist for "Disc 1" of the album "Beethoven - Piano Sonatas - Alfred Brendel". The tracklist includes 11 tracks, with track 11, "Sonata no.8 in C minor, op.13 'Pathétique' / Rondo (Allegro)", selected. The tracklist table is as follows:

Track	Track Name	Duration
03	Sonata no.1 in F minor, op.2 no.1 / Menuetto (Allegretto)	3:24
04	Sonata no.1 in F minor, op.2 no.1 / Prestissimo	5:32
05	Sonata no.2 in A major, op.2 no.2 / Allegro vivace	7:15
06	Sonata no.2 in A major, op.2 no.2 / Largo appassionato	6:28
07	Sonata no.2 in A major, op.2 no.2 / Scherzo (Allegretto)	3:30
08	Sonata no.2 in A major, op.2 no.2 / Rondo (Gravioso)	7:03
09	Sonata no.8 in C minor, op.13 "Pathétique" / Allegro di molto e con brio	9:40
10	Sonata no.8 in C minor, op.13 "Pathétique" / Adagio cantabile	5:17
11	Sonata no.8 in C minor, op.13 "Pathétique" / Rondo (Allegro)	4:30

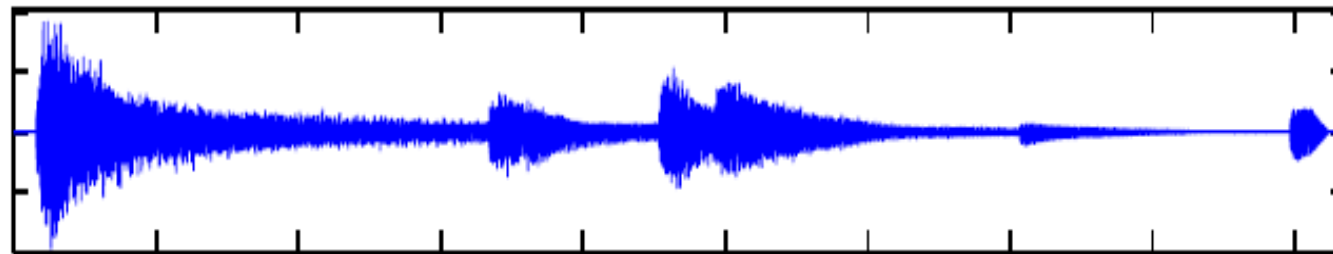
The AudioViewer window also shows a progress bar and controls for "Disc: 1 / 11", "Track: 11 / 11", and "Time: 00:00.00 / 4:30.35".

Music Synchronization: Image-Audio

Image



Audio



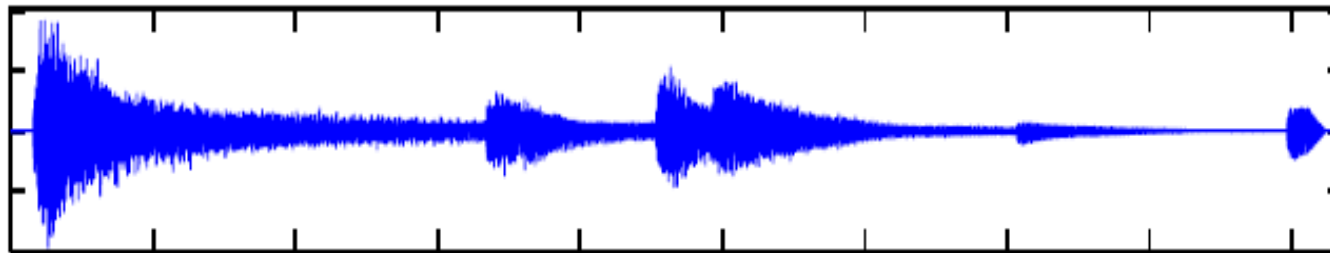
Music Synchronization: Image-Audio

Image Processing: Optical Music Recognition

Image



Audio



Music Synchronization: Image-Audio

Image Processing: Optical Music Recognition

Image



Audio



Audio Processing: Fourier Analysis

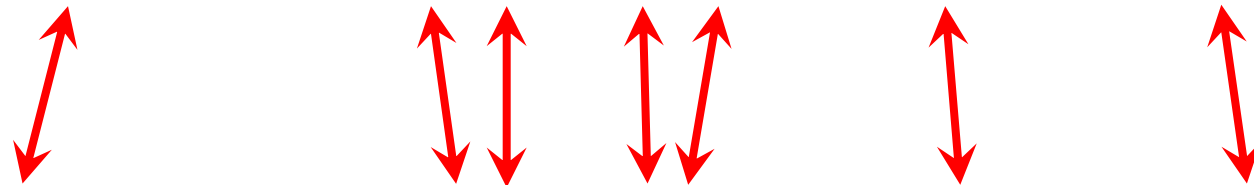
Music Synchronization: Image-Audio

Image Processing: Optical Music Recognition

Image



Audio



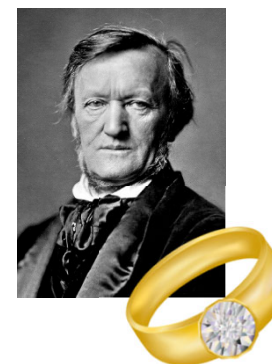
Audio Processing: Fourier Analysis

Music Scenarios

- Freischütz Digital



- Wagner's Ring



- Georgian Music



- Schubert Winterreise



Scenario: Freischütz Digital



- BMBF (2012 – 2016)
- Detmold/Paderborn
(Prof. Veit, Digital Editions)
- Frankfurt
(Prof. Betzwieser, Musicology)
- Erlangen
(Prof. Müller, Computer Science)



Scenario: Freischütz Digital



Audio

- BMBF (2012 – 2016)
- Detmold/Paderborn
(Prof. Veit, Digital Editions)
- Frankfurt
(Prof. Betzwieser, Musicology)
- Erlangen
(Prof. Müller, Computer Science)








Scenario: Freischütz Digital



Recordings

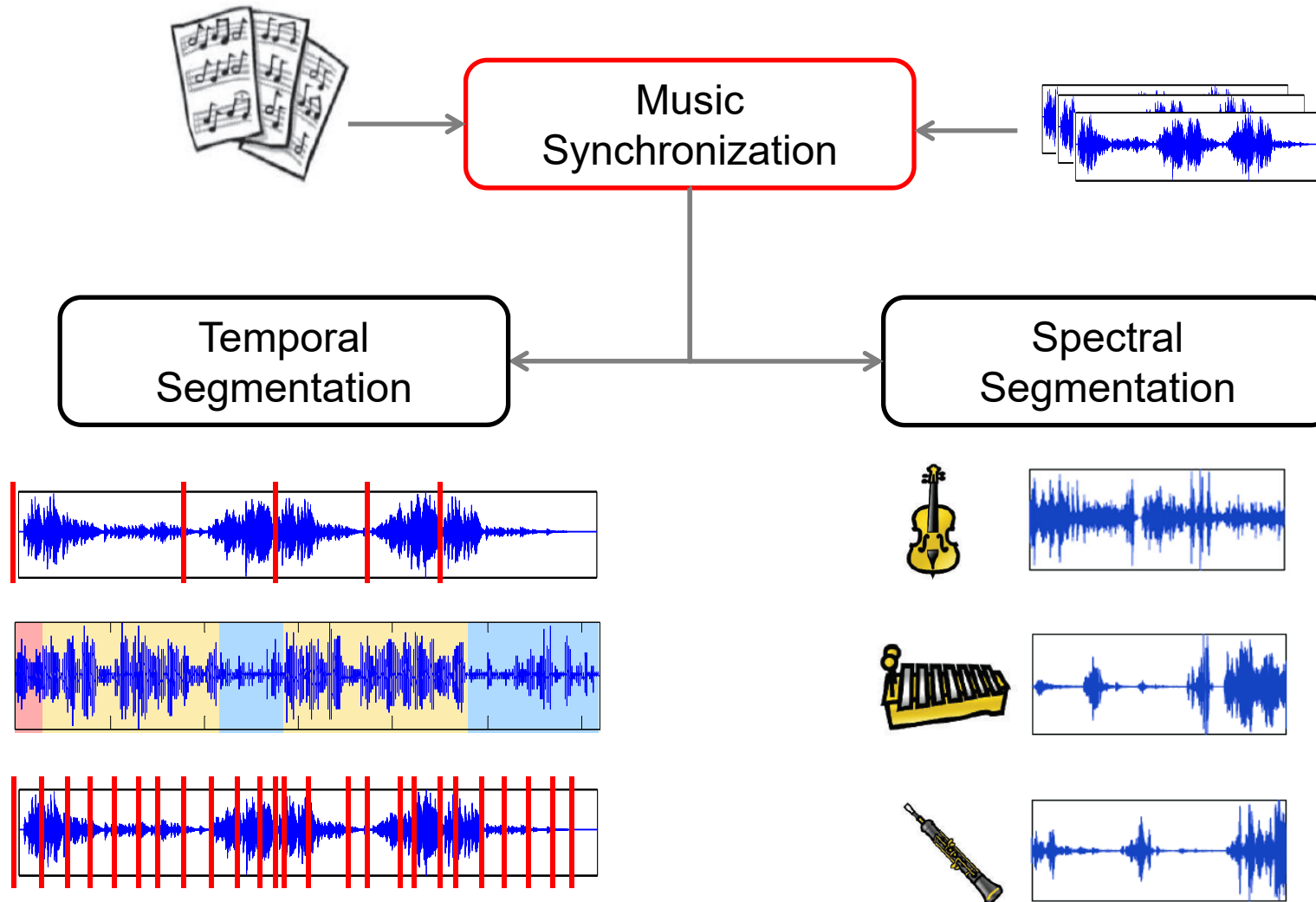
- 23 mostly complete recordings
- 10 abridged/short versions
- Recorded between 1933 and 2001

Example: Song (No. 4) from “Der Freischütz”

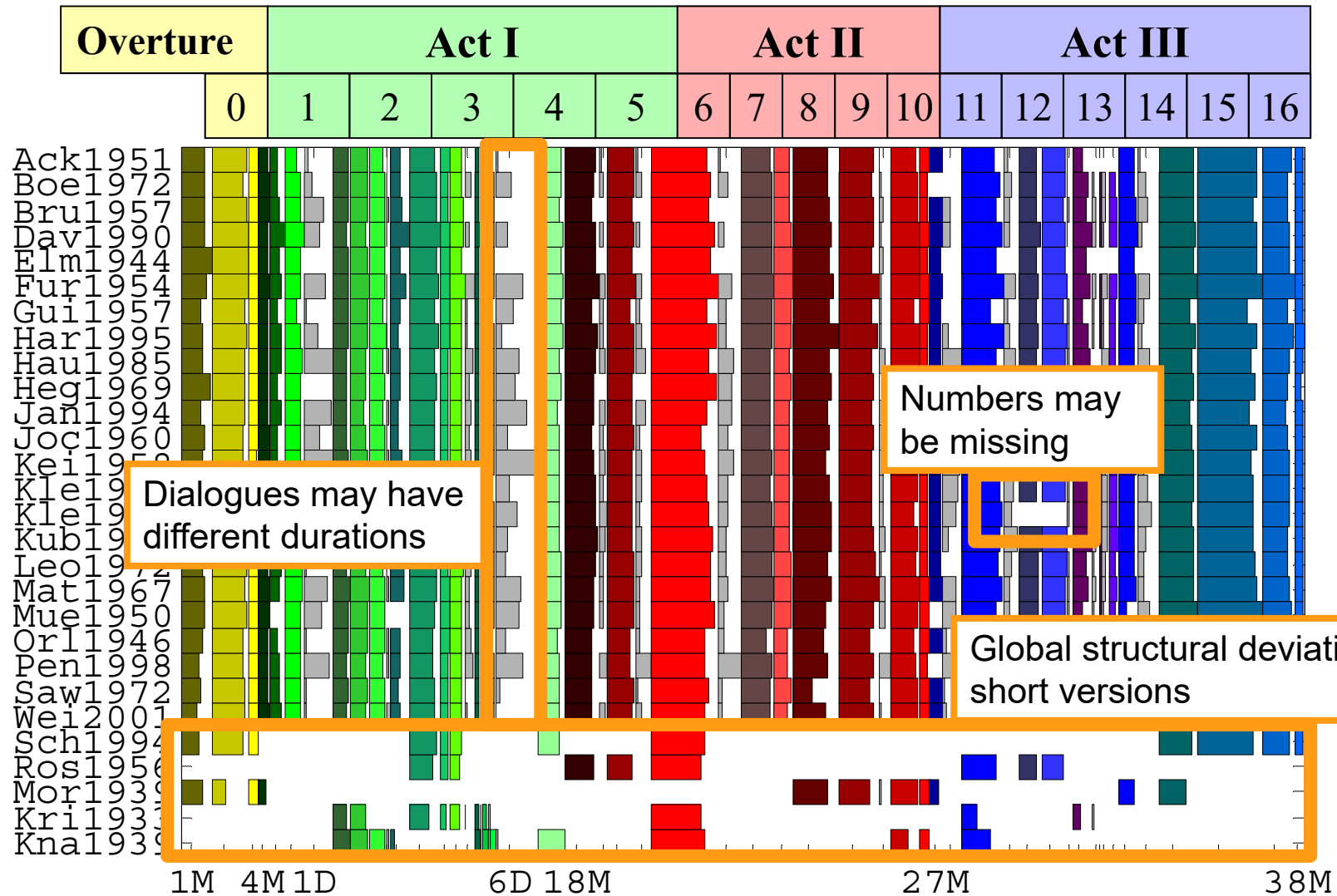
Variations	Performance
	Kleiber C. , 1973 
Tempo	Elmendorff, 1944 
Language	Penin (fr.), 1998 
Key	Orlov (russ.), 1946 
Sound quality	Gui (it.), 1957 



Scenario: Freischütz Digital



Scenario: Freischütz Digital



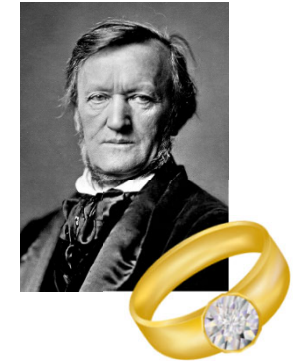
Scenario: Freischütz Digital



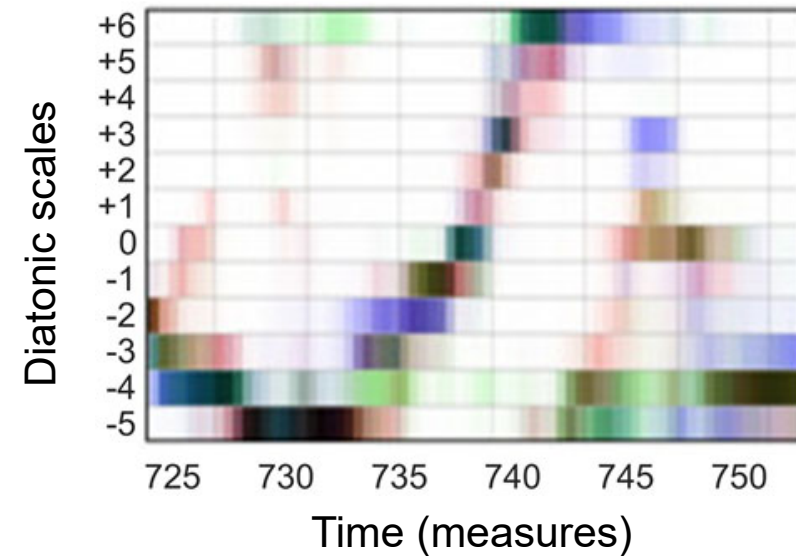
- Global inconsistencies and deviations
- Annotation process raises research questions
 - Structure analysis
 - Partial alignment
 - Language detection
 - Key detection
 - ...
- Annotation process becomes the subject of research

Daniel Röwenstrunk, Thomas Prätzlich, Thomas Betzwieser, Meinard Müller, Gerd Szwillus, Joachim Veit:
Das Gesamtkunstwerk Oper aus Datensicht — Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt Freischütz Digital‘. Datenbank-Spektrum, 15(1): 65–72, 2015.

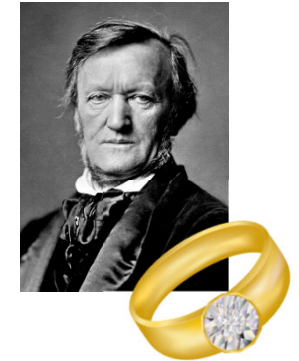
Scenario: Wagner's Ring



- DFG (2014 – 2023)
- Saarbrücken
(Prof. Kleinertz, Musicology)
- Erlangen
(Prof. Müller, Computer Science)
- Objectives
 - Harmony-based structural analysis
 - Visualization techniques
 - Exploration of interdisciplinary research



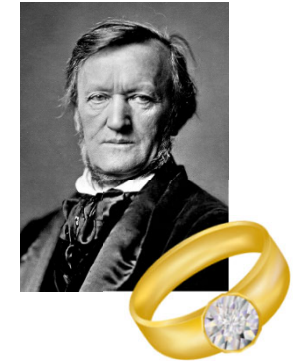
Scenario: Wagner's Ring



No.	Conductor	Recording	hh:mm:ss
1	Barenboim	1991–92	14:54:55
2	Boulez	1980–81	13:44:38
3	Böhm	1967–71	13:39:28
4	Furtwängler	1953	15:04:22
5	Haitink	1988–91	14:27:10
6	Janowski	1980–83	14:08:34
7	Karajan	1967–70	14:58:08
8	Keilberth/Furtwängler	1952–54	14:19:56
9	Krauss	1953	14:12:27
10	Levine	1987–89	15:21:52
11	Neuhold	1993–95	14:04:35
12	Sawallisch	1989	14:06:50
13	Solti	1958–65	14:36:58
14	Swarowsky	1968	14:56:34
15	Thielemann	2011	14:31:13
16	Weigle	2010–12	14:48:46

- Large-scale work
- Four operas
 - ca. 15 hours
 - 21941 measures
- 16 performances

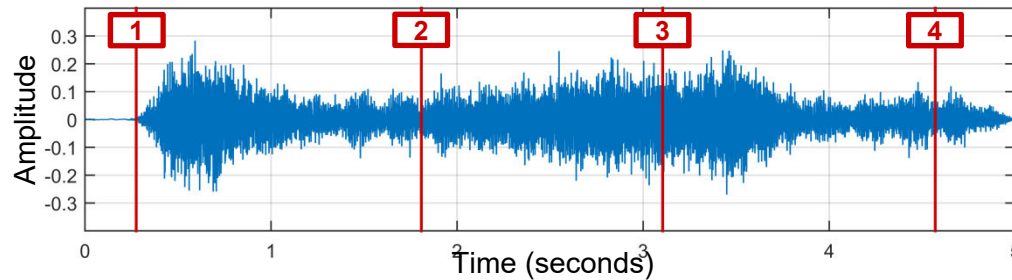
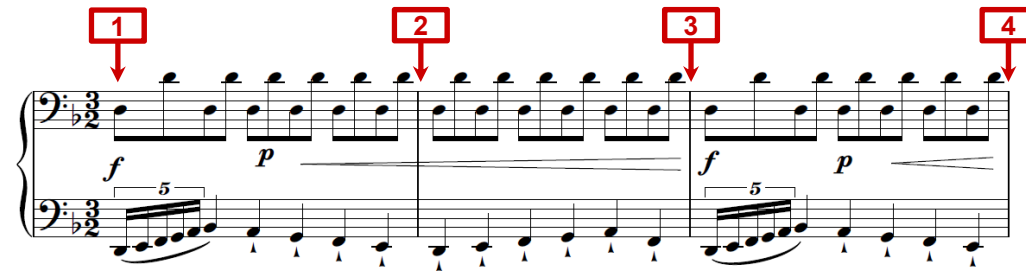
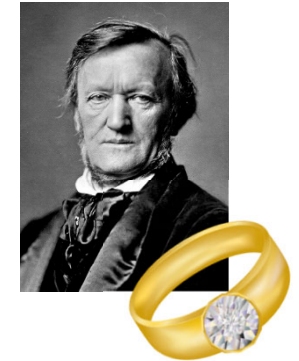
Scenario: Wagner's Ring



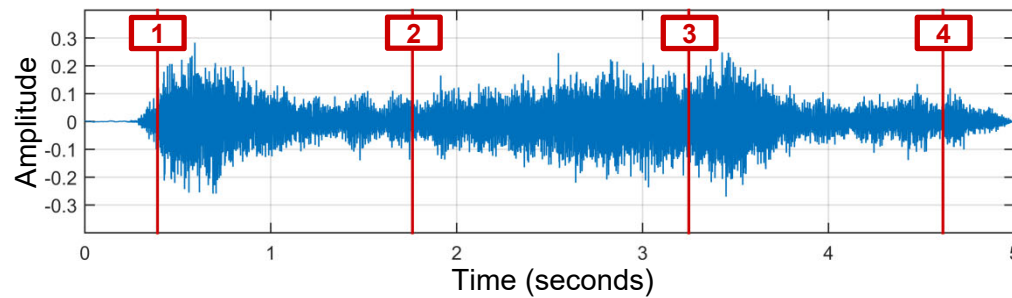
No.	Conductor	Recording	hh:mm:ss
1	Barenboim	1991–92	14:54:55
2	Boulez	1980–81	13:44:38
3	Böhm	1967–71	13:39:28
4	Furtwängler	1953	15:04:22
5	Haitink	1988–91	14:27:10
6	Janowski	1980–83	14:08:34
7	Karajan	1967–70	14:58:08
8	Kelbertn/Furtwängler	1952–54	14:19:56
9	Krauss	1953	14:12:27
10	Levine	1987–89	15:21:52
11	Neuhold	1993–95	14:04:35
12	Sawallisch	1989	14:06:50
13	Solti	1958–65	14:36:58
14	Swarowsky	1968	14:56:34
15	Thielemann	2011	14:31:13
16	Weigle	2010–12	14:48:46

- Large-scale work
- Four operas
 - ca.15 hours
 - 21941 measures
- 16 performances
- Manual measure annotations

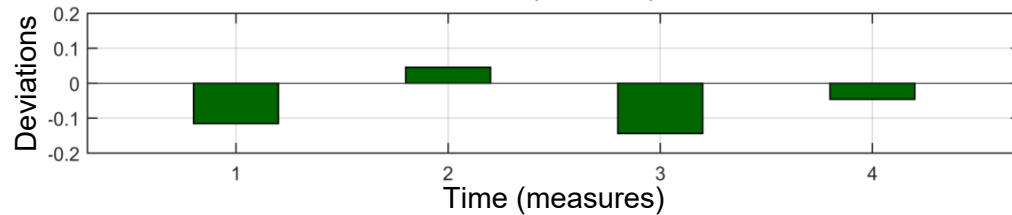
Scenario: Wagner's Ring



Annotator 1

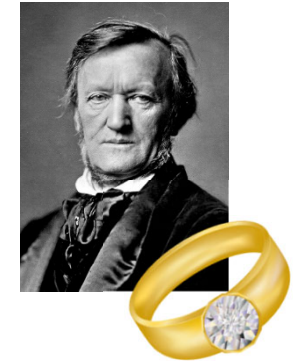


Annotator 2

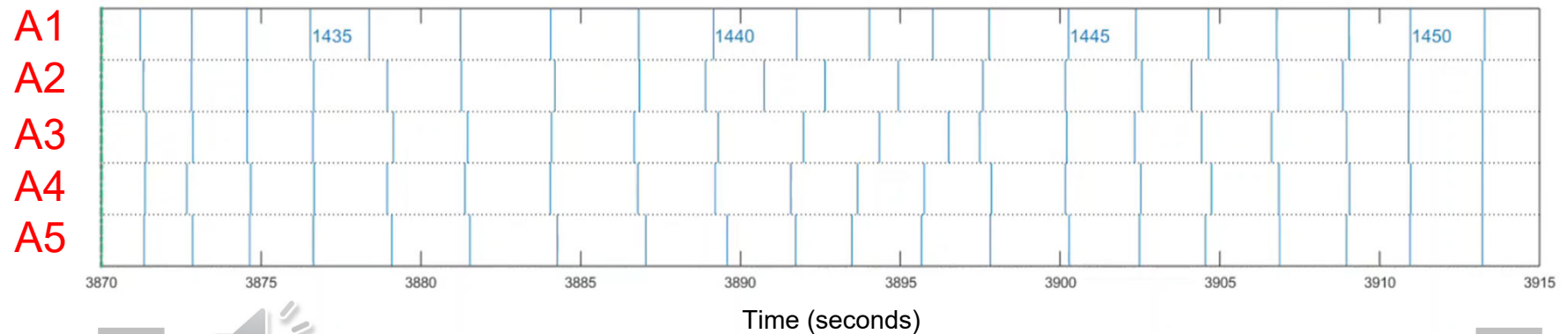
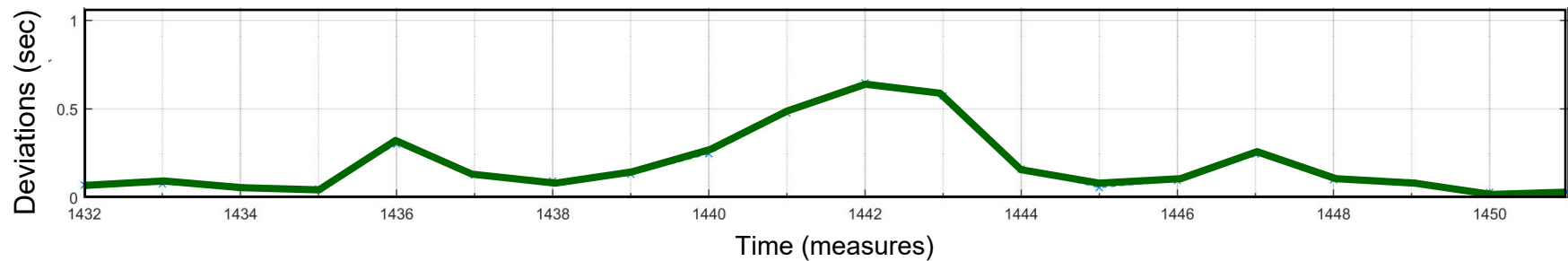


Deviations

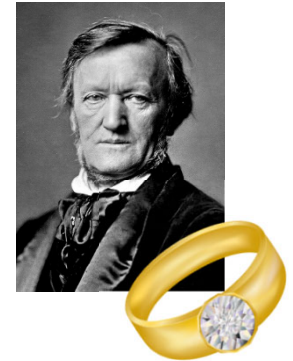
Scenario: Wagner's Ring



Standard deviations among **annotators**



Scenario: Wagner's Ring



- Measure position ambiguities
 - Rhythm or beat unclear
 - Vague note onset positions
 - Non-synchronous parts (e.g., singers and orchestra)
 - ...
- Introduce confidence measures
 - Cross-annotator agreement
 - Cost function based on novelty and homogeneity
 - ...

Christof Weiß, Vlora Arifi-Müller, Thomas Prätzlich, Rainer Kleinertz, Meinard Müller:
Analyzing Measure Annotations for Western Classical Music Recordings.
In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR): 517–523, 2016.

Scenario: Georgian Music



- DFG (2018 – 2022)
- Potsdam
(Prof. Scherbaum, Ethnomusicology)
- Erlangen
(Prof. Müller, Computer Science)
- Objectives
 - Harmonic and melodic singing analysis
 - New sensors (larynx microphones)
 - Digital humanities

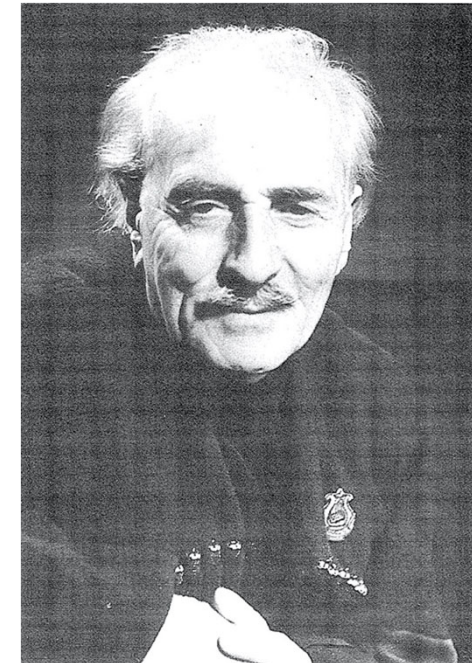


Scenario: Georgian Music

Erkomaishvili Dataset



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



“Original masterpieces of Georgian musical thinking.” (Shugliashvili, 2014)

Scenario: Georgian Music

Erkomaishvili Dataset



○ Bass Voice

○ Middle Voice

○ Top Voice

▶ Middle Voice

▶ Top Voice

▶ Top Voice

Amplitude

Time

Segment 1

Segment 2

Segment 3

▶ 🔊

▶ 🔊

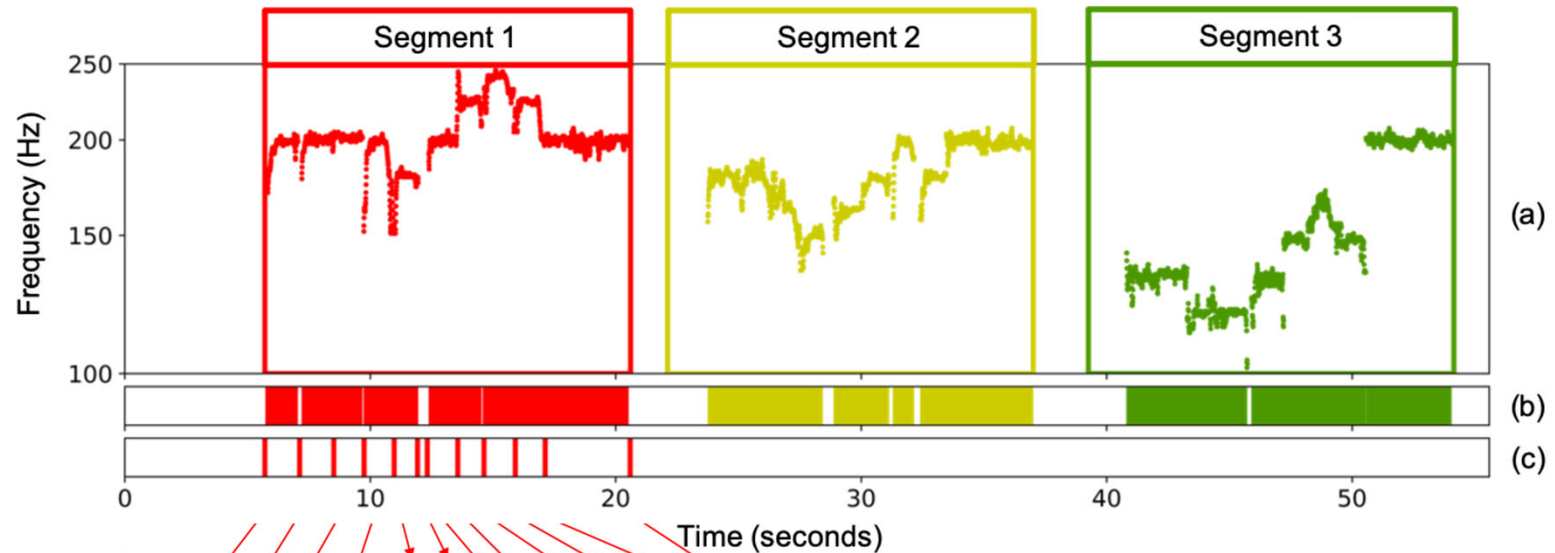
▶ 🔊

▶ 🔊

An audio waveform visualization with three segments highlighted in red, yellow, and green. Above the waveform are three rows of controls: 'Bass Voice' (green), 'Middle Voice' (yellow), and 'Top Voice' (red). Each row has a radio button and a play button. Below the waveform are three boxes labeled 'Segment 1', 'Segment 2', and 'Segment 3', each with a play button and a speaker icon.

Scenario: Georgian Music

Erkomaishvili Dataset



Top Voice (Segment 1)

Middle Voice (Segment 2)

Bass Voice (Segment 3)

Time (seconds)

da su - li - sa - tsa she - ni - sa - ta - na.
1 2 3 4 5 6 7 8 9 10

da su - li - sa - tsa she - ni - sa - ta - na.
1 2 3 4 5 6 7 8 9 10

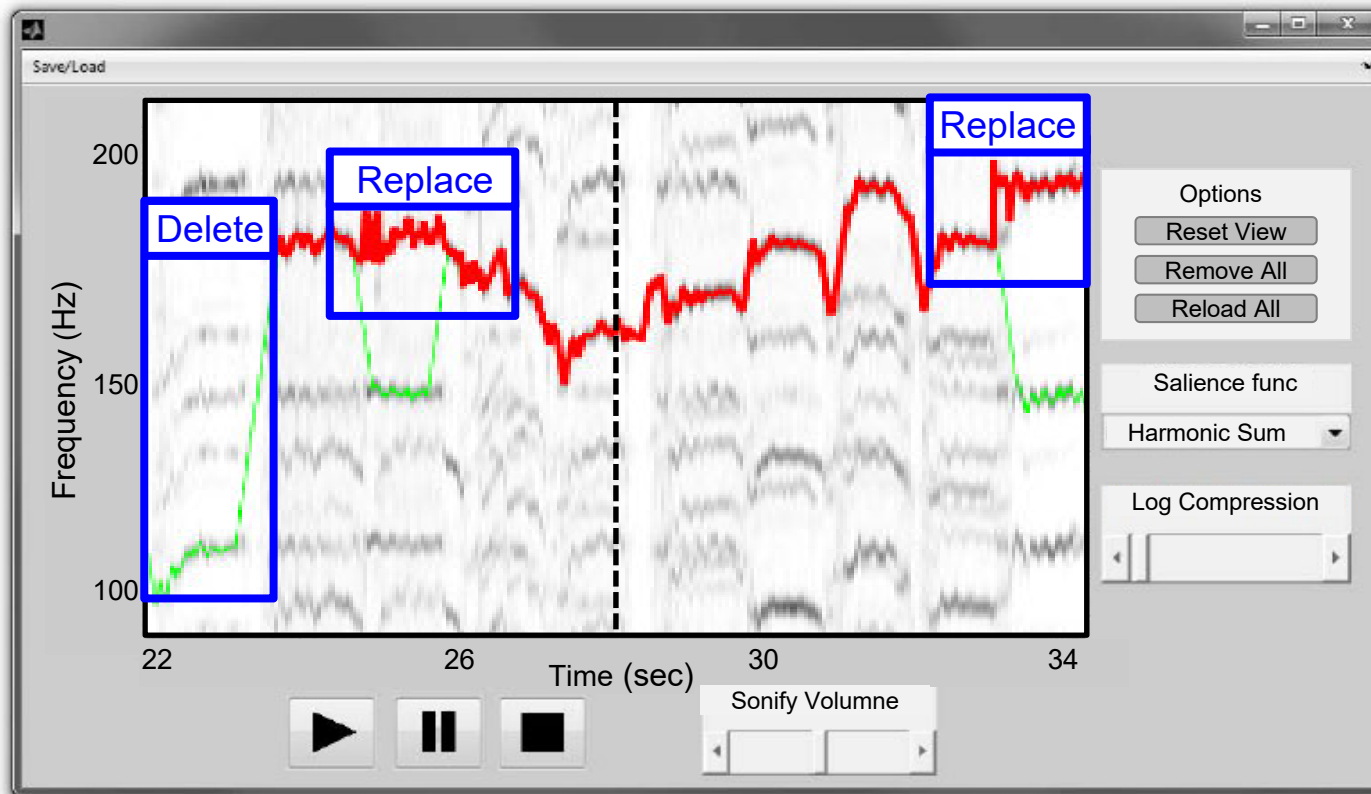
da su - li - sa - tsa she - ni - sa - ta - na.
1 2 3 4 5 6 7 8 9 10

The figure shows three staves of musical notation. The top staff is for the Top Voice (Segment 1), the middle for the Middle Voice (Segment 2), and the bottom for the Bass Voice (Segment 3). The lyrics are 'da su - li - sa - tsa she - ni - sa - ta - na.' with syllables numbered 1 to 10. Red arrows point from the lyrics to the corresponding notes in the top staff.

- F0-annotations
- Recording structure annotations
- Onset annotations
- Digital sheet music

Scenario: Georgian Music

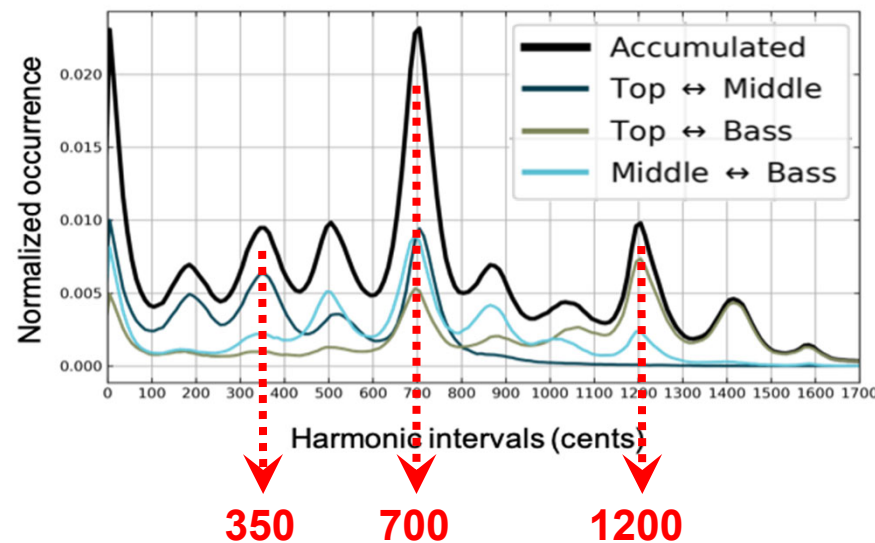
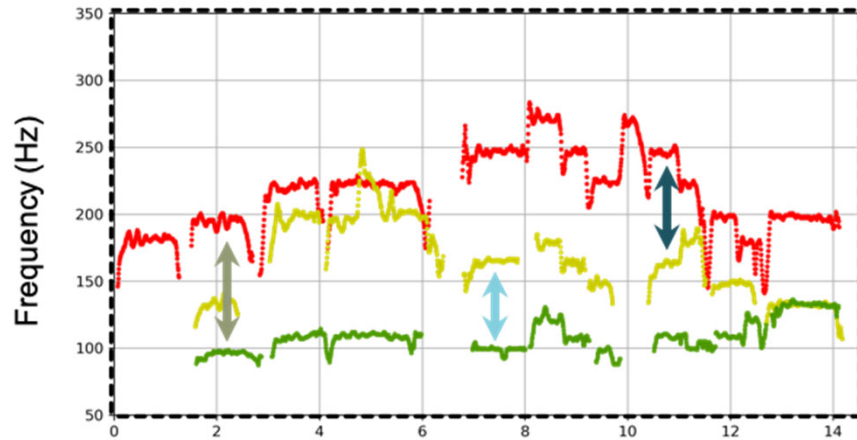
Interactive F0 Annotation Tool



Meinard Müller, Sebastian Rosenzweig, Jonathan Driedger, and Frank Scherbaum:
Interactive Fundamental Frequency Estimation with Applications to Ethnomusicological Research.
In Proceedings of the AES Conference on Semantic Audio, 2017.

Scenario: Georgian Music

Erkomaishvili Dataset



- Superimposed F0-trajectories
- Measuring harmonic intervals
- Peak at 350 cents (between minor and major third)
- Non-western temperament

Scenario: Georgian Music

Erkomaishvili Dataset



The interface displays a table of recordings and a musical score player. A red circle highlights the 'Link' column in the table, with a red arrow pointing to the score player.

ID	English Title	Georgian Title	Website
001	Christ is risen from the dead	Qrist'e aghsdga	Link
002	The Angels in the Heaven	Aghdgomasa shensa	Link
003	Christ is risen from the dead	Qrist'e aghsdga	Link
004	Christ is risen from the dead	Qrist'e aghsdga	Link
005	The Day of Resurrection	Aghdgomisa dghe ars	Link
006	Let us purify our senses	Ganvits'midnet satsnobelni	Link
007	For meet is it that heavens	Tsani q'ovlad ghirsabit	Link
008	O, come, let us quaff a beverage new	Movedit da vsvat	Link
009	Now are filled with all the light	Ats' q'ovliturt aghivso	Link
010	Yesterday, O Christ	Gushin shentana	Link

The score player shows three staves of music with lyrics: agh-dgo-ma - sa she-n-sa, 1 2 3 6 7 9 10 11 12 13. The score includes a treble clef, a key signature of one sharp (F#), and a common time signature (C). The lyrics are: agh-dgo-ma - sa she-n-sa, 1 2 3 4 5 7 9 10 11 12 13. The score player also features a playback control bar with a play button, a progress bar, and a time display of 00:00:03:639 / 00:00:57:263. Below the score player, there are four radio button options: First Segment (Top Voice), Second Segment (Top + Middle Voice), Third Segment (Top + Middle + Bass Voice), and Mix of all Segments.

<https://www.audiolabs-erlangen.de/resources/MIR/2019-GeorgianMusic-Erkomaishvili>

Scenario: Georgian Music

Erkomaishvili Dataset



- Temporal organization
 - No notion of meter
 - Continuous note transitions (glissando)
 - Voices not synchronous
- Tonal organization
 - Non-western temperament
 - Harmonic vs. melodic intonation
 - Transcription problematic
- Poor recording conditions

Sebastian Rosenzweig, Frank Scherbaum, David Shugliashvili, Vlori Arifi-Müller, and Meinard Müller:
Erkomaishvili Dataset: A Curated Corpus of Traditional Georgian Vocal Music for Computational Musicology.
Transactions of the International Society for Music Information Retrieval (TISMIR), 3(1): 31–41, 2020.

Scenario: Schubert Winterreise

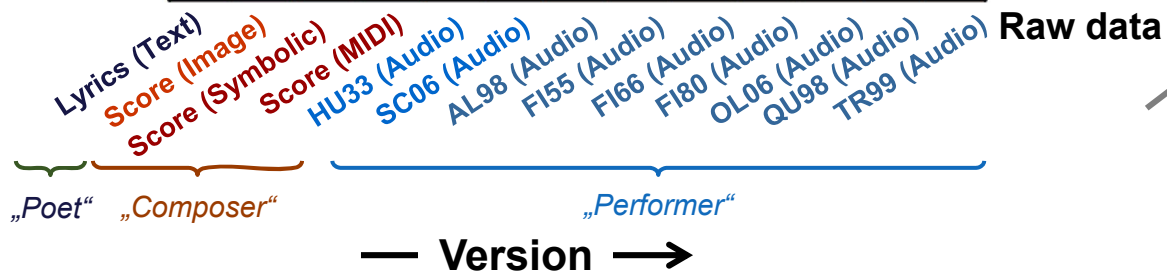


- Winterreise
 - Song cycle for voice and piano
 - Music: Franz Schubert (1828)
 - Poems: Wilhelm Müller

- MIR Objectives
 - Music synchronization
 - Structure analysis
 - Harmonic analysis
 - Activity detection (singing, lyrics, ...)
 - ...

— Song —
↓

- 01. Gute Nacht
- 02. Die Wetterfahne
- 03. Gefrorne Tränen
- 04. Erstarrung
- 05. Der Lindenbaum
- 06. Wasserflut
- 07. Auf dem Flusse
- 08. Rückblick
- 09. Irrlicht
- 10. Rast
- 11. Frühlingstraum
- 12. Einsamkeit
- 13. Die Post
- 14. Der greise Kopf
- 15. Die Krähe
- 16. Letzte Hoffnung
- 17. Im Dorfe
- 18. Der stürmische Morgen
- 19. Täuschung
- 20. Der Wegweiser
- 21. Das Wirtshaus
- 22. Muth
- 23. Die Nebensonnen
- 24. Der Leiermann



Local Keys
Chords
Measures
— Annotations —→

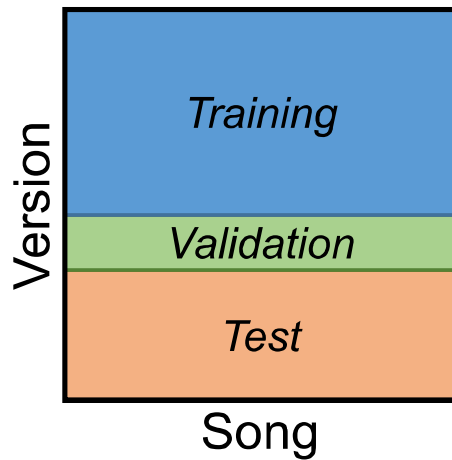


Scenario: Schubert Winterreise

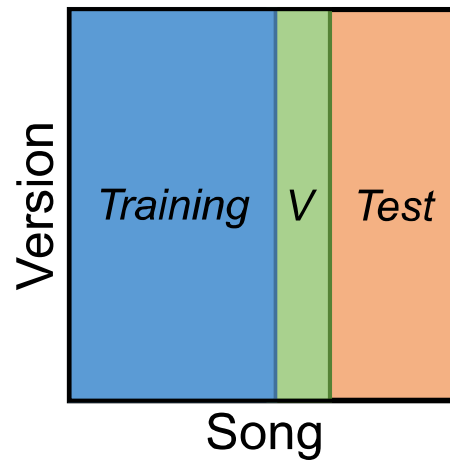
Cross-Version Evaluation



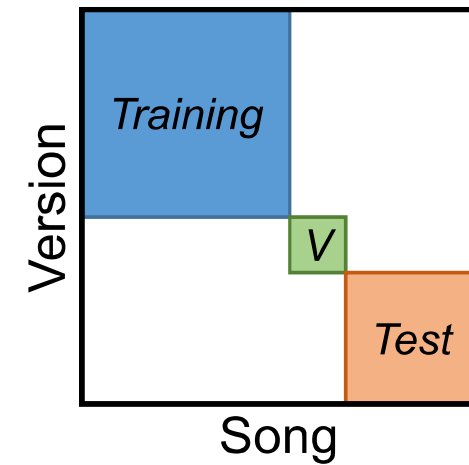
Version split



Song split

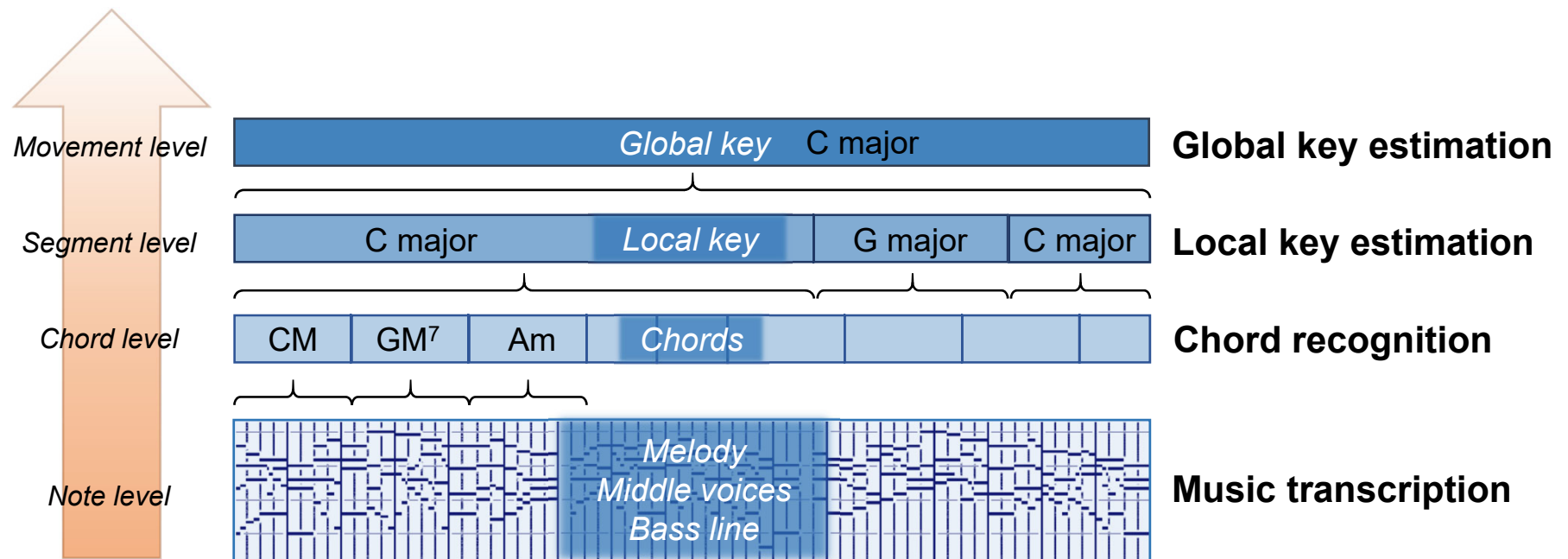


Neither split



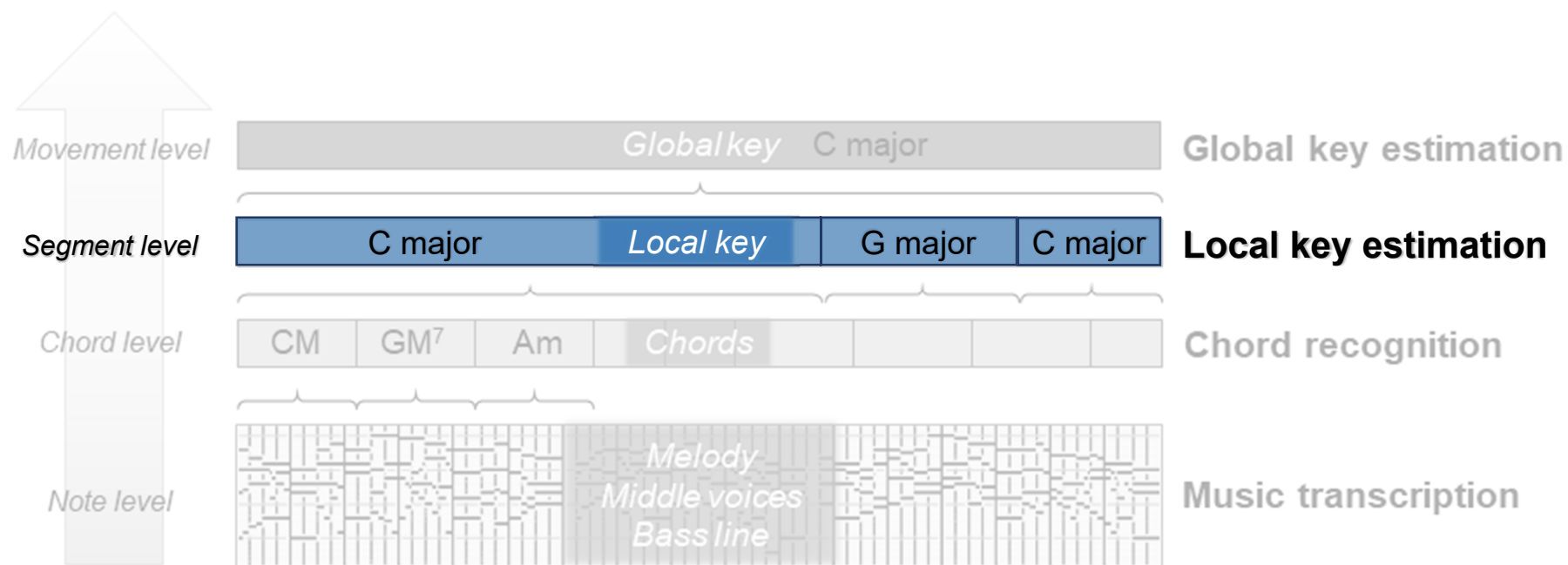
Scenario: Schubert Winterreise

Harmony Analysis



Scenario: Schubert Winterreise

Harmony Analysis



Scenario: Schubert Winterreise

Harmony Analysis



21 22 23 24 25 26 27 28 29 30 31

Voice
 Ach! wer wie ich so e-lend ist, gibt gern sich hin der bun-ten List, die hin-ter Eis und Nacht und Graus ihm weist ein

Piano
cresc. *p*



Annotations

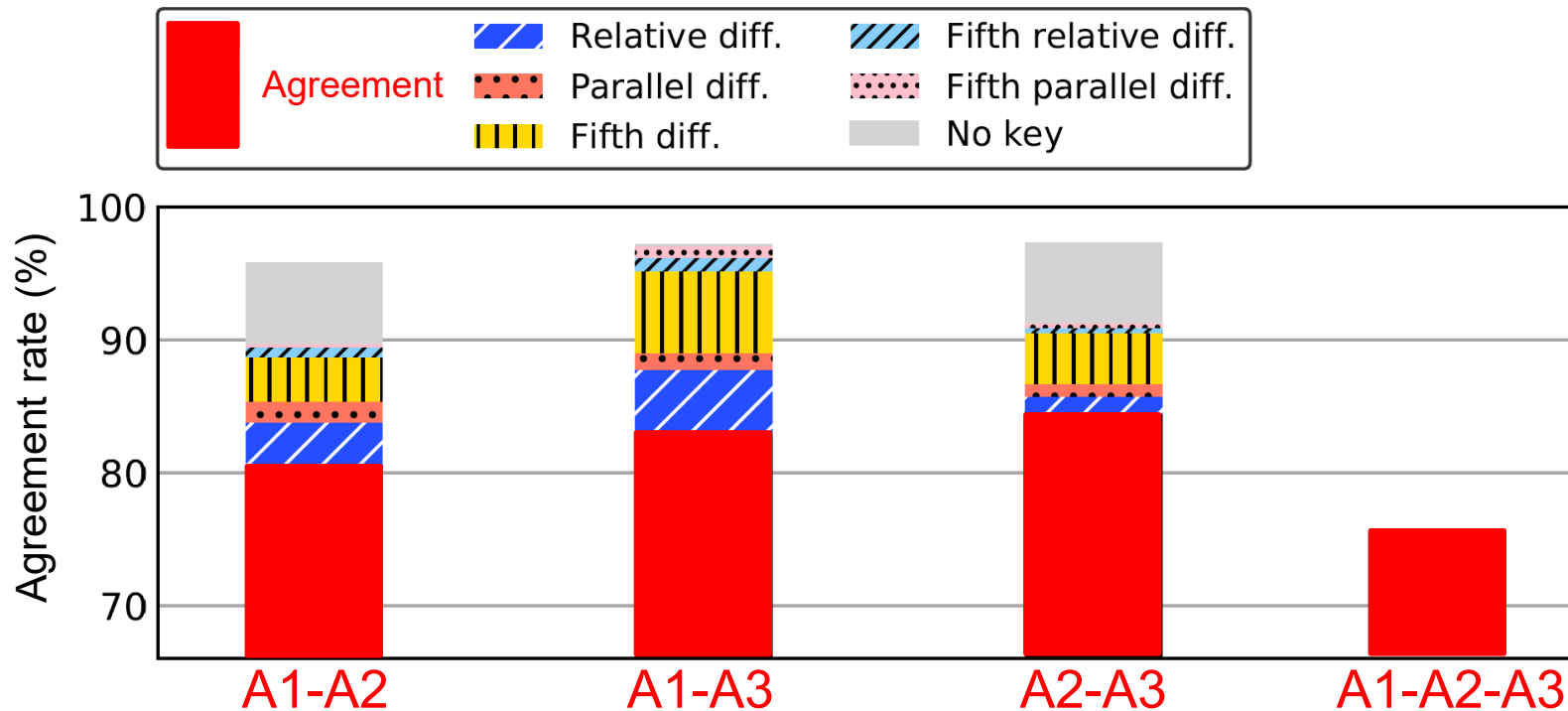
A1	A major	A minor	E major	A major
A2	A major			
A3	A major	A minor		A major

Scenario: Schubert Winterreise

Harmony Analysis



Annotator **agreements** and differences



Scenario: Schubert Winterreise

Harmony Analysis

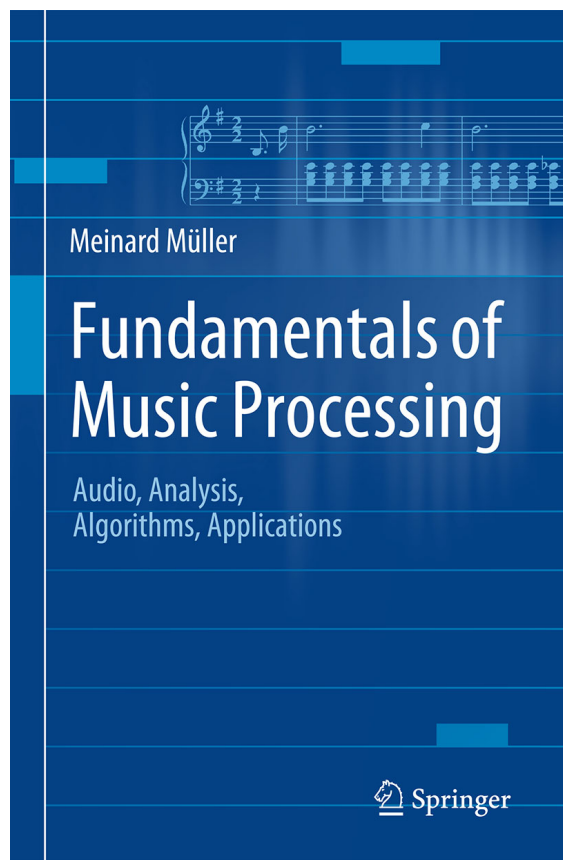
- Harmony-related annotations
 - Hierarchical nature of musical structures
 - High degree of subjectivity
 - Dependence on user needs and applications
- ...

Christof Weiß, Frank Zalkow, Vlora Arifi-Müller, Meinard Müller, Hendrik Vincent Koops, Anja Volk, Harald Grohganz:
Schubert Winterreise Dataset: A Multimodal Scenario for Music Analysis.
ACM Journal on Computing and Cultural Heritage (JOCCH), 15(2): 1–18, 2021.

Conclusions

- Annotating music data is a challenge
 - Data inconsistencies
 - Underlying model assumptions often violated
 - High degree of subjectivity
 - Dependency on user needs and applications
 - **Never trust your annotations!**
- Annotations and analyses cannot be separated
 - Needs to be an interactive process
 - Requires a dialogue between domain experts and computer scientists
 - Requires an understanding and adaption of tools
- Opportunities
 - Annotation process becomes subject of research
 - Increasing appreciation of datasets
 - Great potential for interdisciplinary research

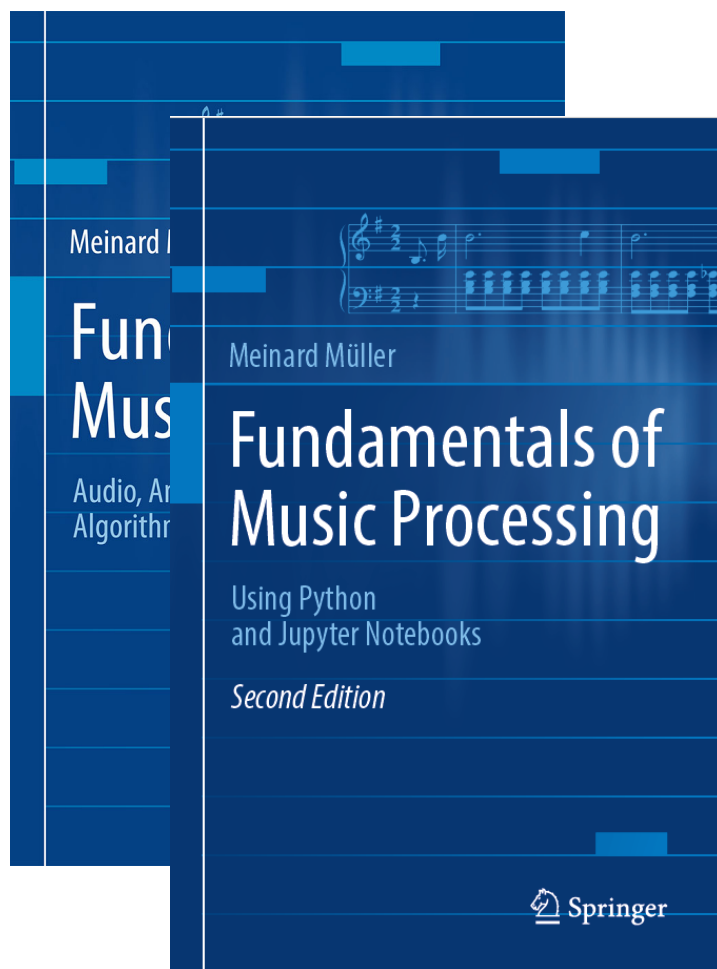
Fundamentals of Music Processing (FMP)



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

Accompanying website:
www.music-processing.de

Fundamentals of Music Processing (FMP)

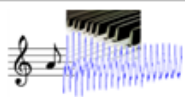

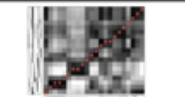
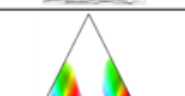

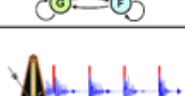




Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
Springer, 2015

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)

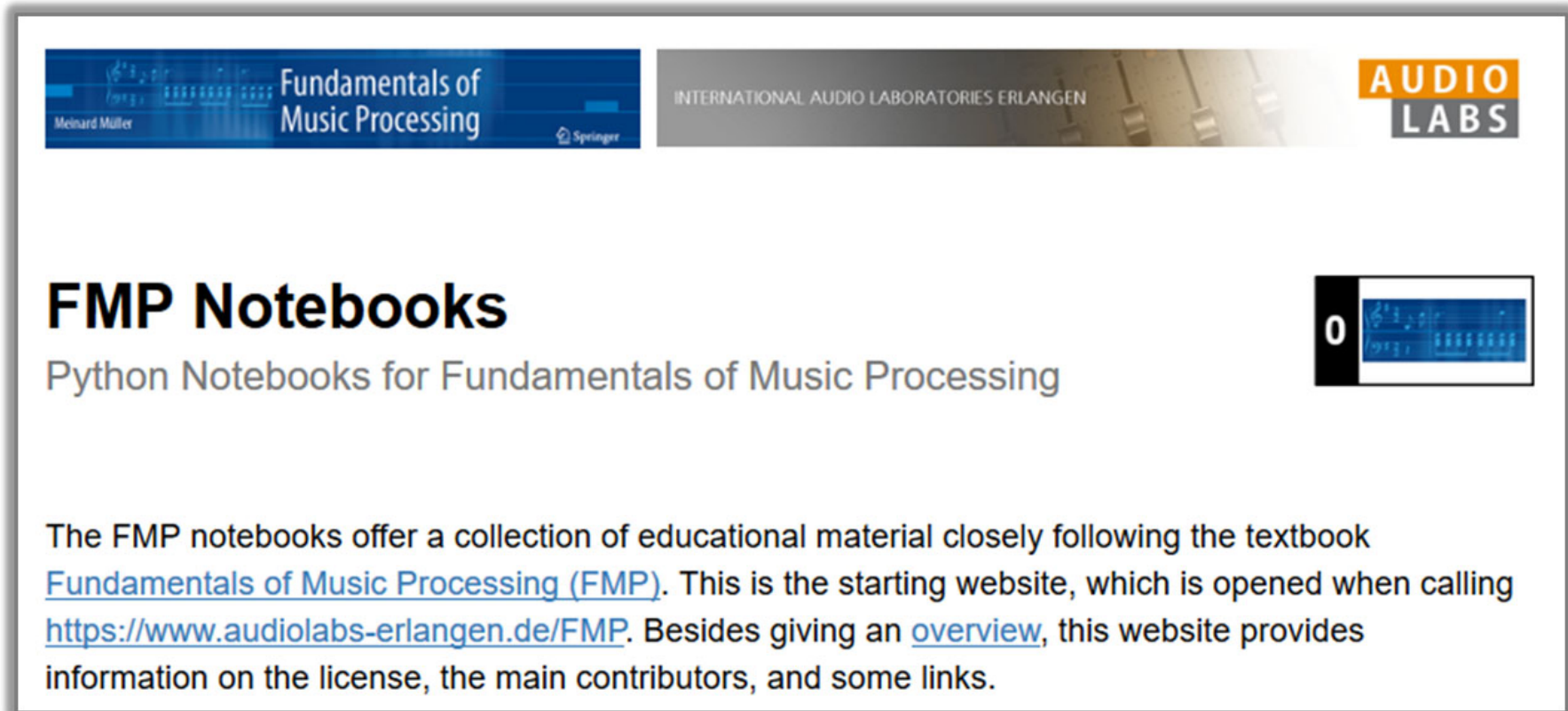
Chapter	Music Processing Scenario
1	 Music Representations
2	 Fourier Analysis of Signals
3	 Music Synchronization
4	 Music Structure Analysis
5	 Chord Recognition
6	 Tempo and Beat Tracking
7	 Content-Based Audio Retrieval
8	 Musically Informed Audio Decomposition

Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
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: Education & Research



The screenshot shows the header of the FMP Notebooks website. On the left, there is a blue banner for the textbook 'Fundamentals of Music Processing' by Meinard Müller, published by Springer. To the right of this banner is the text 'INTERNATIONAL AUDIO LABORATORIES ERLANGEN' and the 'AUDIO LABS' logo. Below the banner, the main heading reads 'FMP Notebooks' in a large, bold, black font, followed by the subtitle 'Python Notebooks for Fundamentals of Music Processing' in a smaller, grey font. To the right of the subtitle is a small thumbnail image of a notebook page with a black bar on the left containing the number '0'. Below the heading and subtitle, a paragraph of text describes the notebooks: 'The FMP notebooks offer a collection of educational material closely following the textbook [Fundamentals of Music Processing \(FMP\)](#). This is the starting website, which is opened when calling <https://www.audiolabs-erlangen.de/FMP>. Besides giving an [overview](#), this website provides information on the license, the main contributors, and some links.'

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

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>

References

- Meinard Müller: **Fundamentals of Music Processing – Using Python and Jupyter Notebooks**. 2nd Edition, Springer, 2021.
- Daniel Röwenstrunk, Thomas Prätzlich, Thomas Betzwieser, Meinard Müller, Gerd Szwillus, Joachim Veit: **Das Gesamtkunstwerk Oper aus Datensicht — Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt Freischütz Digital**. Datenbank-Spektrum, 15(1): 65–72, 2015.
- Christof Weiß, Vlora Arifi-Müller, Thomas Prätzlich, Rainer Kleinertz, Meinard Müller: **Analyzing Measure Annotations for Western Classical Music Recordings**. In Proceedings of the International Society for Music Information Retrieval Conference (ISMIR): 517–523, 2016.
- Sebastian Rosenzweig, Frank Scherbaum, David Shugliashvili, Vlora Arifi-Müller, and Meinard Müller: **Erkomaishvili Dataset: A Curated Corpus of Traditional Georgian Vocal Music for Computational Musicology**. Transactions of the International Society for Music Information Retrieval (TISMIR), 3(1): 31–41, 2020.
- Meinard Müller, Sebastian Rosenzweig, Jonathan Driedger, and Frank Scherbaum: **Interactive Fundamental Frequency Estimation with Applications to Ethnomusicological Research**. In Proceedings of the AES Conference on Semantic Audio, 2017.
- Christof Weiß, Frank Zalkow, Vlora Arifi-Müller, Meinard Müller, Hendrik Vincent Koops, Anja Volk, Harald Grohgan: **Schubert Winterreise Dataset: A Multimodal Scenario for Music Analysis**. ACM Journal on Computing and Cultural Heritage (JOCCH), 15(2): 1–18, 2021.
- Christof Weiß, Hendrik Schreiber, Meinard Müller: **Local Key Estimation in Music Recordings: A Case Study Across Songs, Versions, and Annotators**. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28: 2919–2932, 2020.