Music Information Retrieval

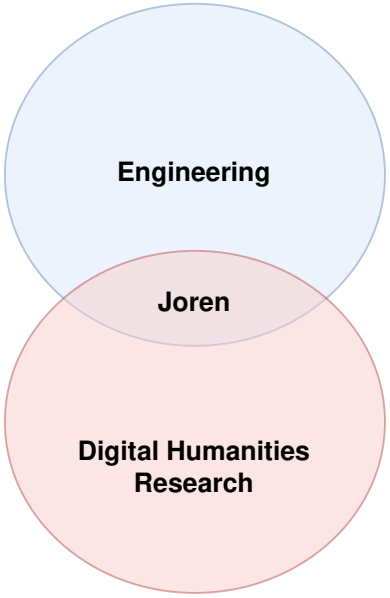
Opportunities for digital musicology

November 2023 - Ghent

Joren Six

Who?

- Studied computer science
- Researcher at Ghent Conservatory
- **Phd at IPEM** *Engineering Systematic Musicology*
- Involved as Post Doc:
  - Nano4Sports *Low impact runner*
  - CONBOTS *COnnected throug roBOTS*
  - AMPLE *the Augmented Movement Platform for Embodied Learning*
  - **PaPiOM** ***Pa****tterns in* ***Pi****tch* ***O****rganization in* ***M****usic*
- Now at Ghent Center for Digital Humanities



What?

- Music Information Retrieval
  - Introduction
  - Music information - Tasks
  - Methods - Tools
- Pitch organisation: PaPiOM
  - Introduction
  - Music information
  - Methods
  - Case study
- Duplicate detection
  - Introduction
  - Music information
  - Methods
  - Applications

MIR introduction

Goal

An overview of the Music Information Retrieval research field while focusing on opportunities for digital musicology.

MIR introduction

Definition

*Music Information Retrieval is the **interdisciplinary** science of extracting and processing **information** from music.*

MIR combines insights from **musicology**, computer science, library sciences, psychology, machine learning and cognitive sciences.

MIR introduction

MIR tasks process information on music. **Music information** can be captured by **signals or symbols**.

Definition

Signals are representations of analog manifestations and replicate perception. Symbols are discretized, limited and replicate content.

Music information

Signal

- Recorded performances
  - Video
  - Audio
  - Motion capture
  - MIDI
- Scans of scores

Symbols

- Meta-data
  - Artist
  - Title
  - Album-name
  - Label
  - Composer
  - Instrumentation ...
- Lyrics
- Rags, reviews, ratings
- Digitized scores

Music information

Scan

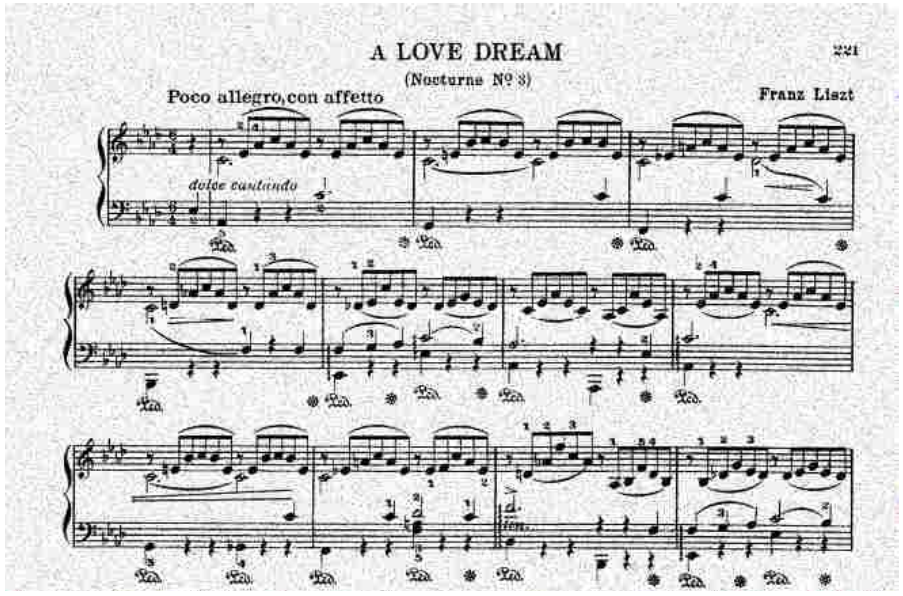


Fig: Scanned score

MusicXML

```
...
<note default-x="390.29" default-y="-40.00">
   <pitch>
      <step>E</step>
      <alter>-1</alter>
      <octave>4</octave>
      </pitch>
   <duration>3</duration>
   <voice>1</voice>
   <type>eighth</type>
   <stem>up</stem>
   <staff>1</staff>
   <beam number="1">end</beam>
   <notations>
      <slur type="stop" number="1">
   </slur></notations>
</note>
...
```

Code: MusicXML Digitized score

Music information: Signal or symbol?

Signal or symbol?

Video

0:00 / 0:20

MIDI

0:00 / 0:18

Recorded MIDI

0:00 / 0:17

Arthur Rubinstein

0:00 / 0:13

Daniel Barenboim

0:00 / 0:13

Music information

A score can be seen as a model of a performance.

> ***Quote***
>
> *"Essentially, all models are wrong, but some are useful"*
> *- George E.P. Box*

Models aim to reduce dimensions, complexity and improve understanding and readability.

MIR tasks: music transcription

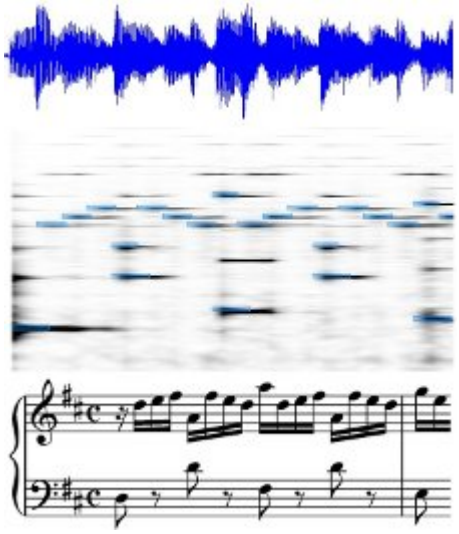


Fig: music transcription

- Source separation
- Instrument recognition
- Pitch estimation and segmentation
- Tempo and rhythm extraction

Task type: signal → symbolic

MIR tasks: structural analysis

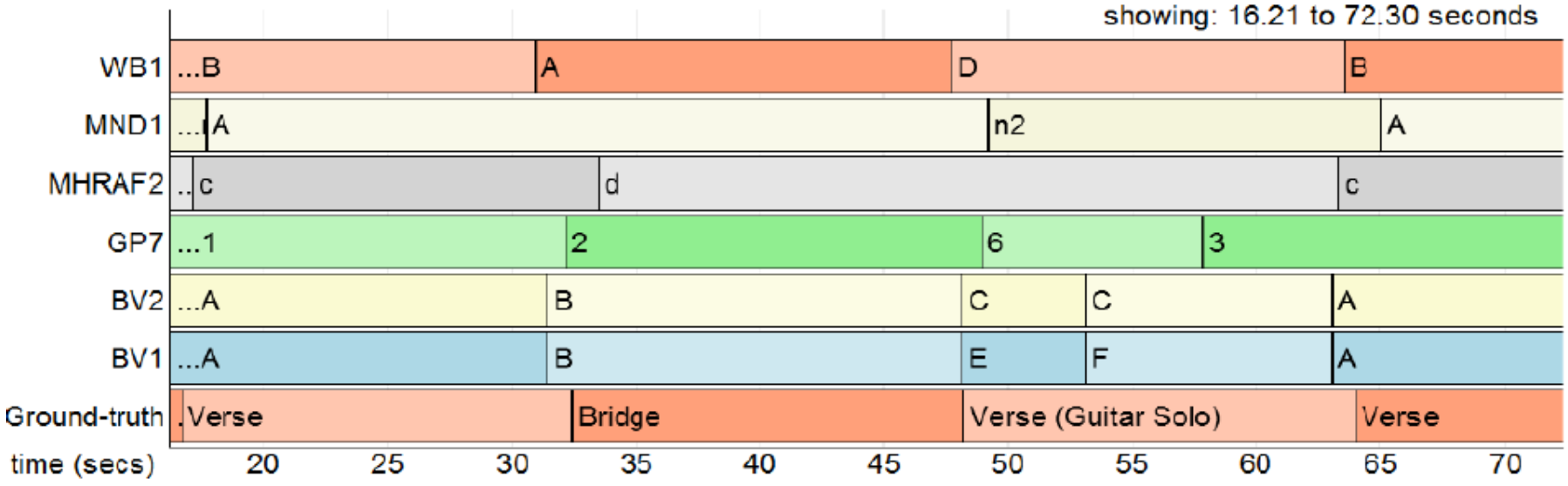


Fig: structural analysis

Task type: signal → symbolic

MIR tasks: music recommendation

Fig: Spotify automatically generates playlists based on listening behaviour

Music recommendation

- Content based: signal → symbolic
- Based on (listening) behaviour: symbolic → symbolic

MIR tasks: other tasks

- Score following: page turning based on musical content
- Music emotion recognition
- Automatic cover song identification
- Optical music recognition: convert images of scores into digital scores
- Symbolic music retrieval
- Automatic genre recognition

MIR Tasks

Most tasks enable to browse, categorize, query or discover music in large databases of music.

MIR tasks: ± Solved

- Monophonic pitch estimation

  De Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. The Journal of the Acoustical Society of America, 111(4), 1917-1930.

- Content based audio search

  Six, J., & Leman, M. (2014). Panako: a scalable acoustic fingerprinting system handling time-scale and pitch modification. In 15th International Society for Music Information Retrieval Conference (ISMIR-2014).

MIR tasks: challenges

Un-mix the mix

Decomposing a mixed audio signal is very very hard.

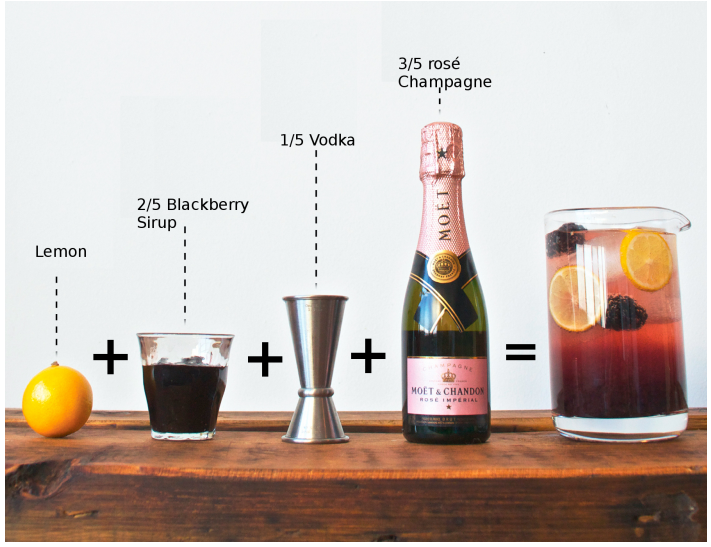
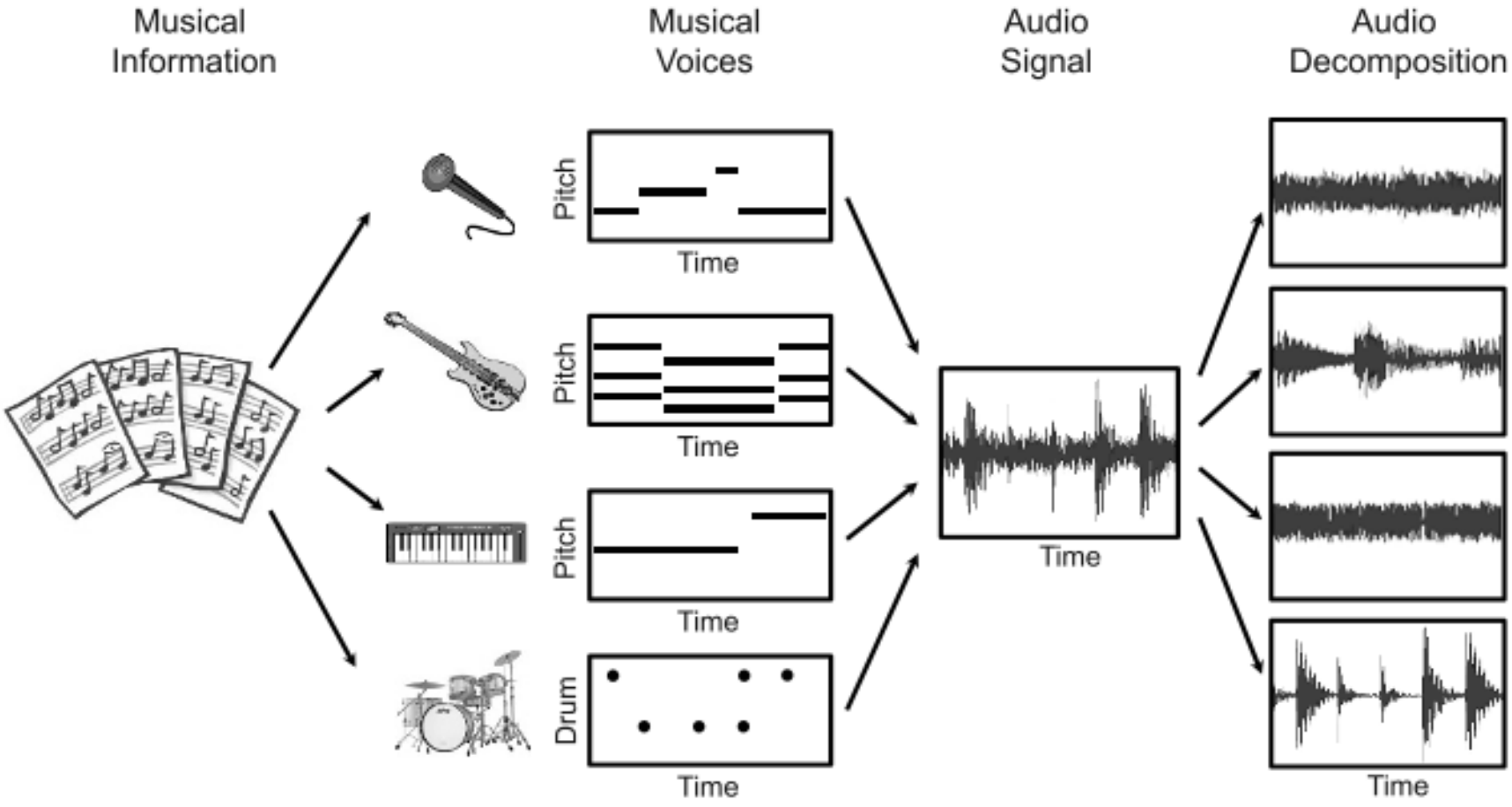


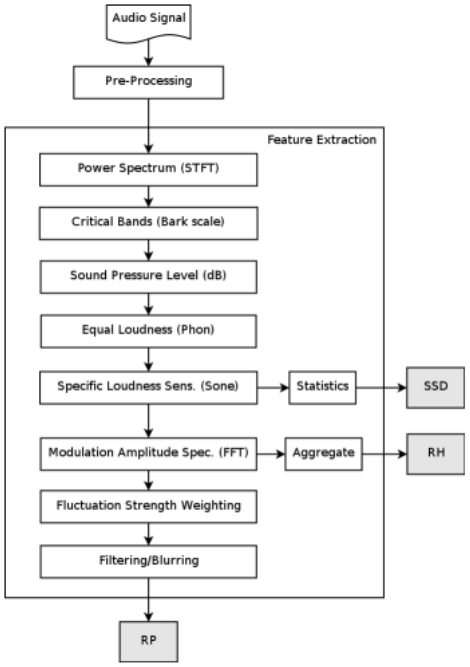
Fig: Mixing is easy,...

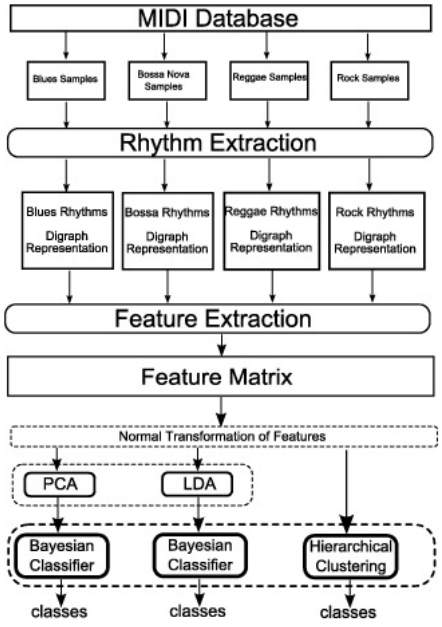
unmixing?

MIR tasks: challenges



MIR Methods - Bag of features





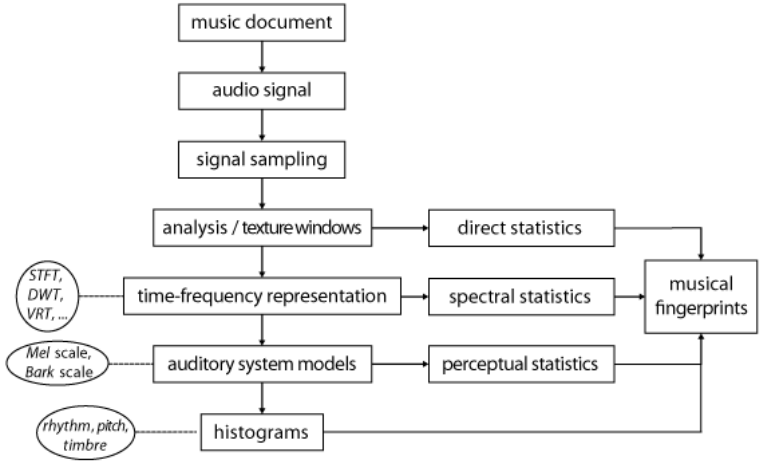


Fig: input → feature(s) → feature processing → output

MIR Methods - Bag of features

Bag of features and classifier to represent e.g. a **musical genre**.

- MFCC, timbral characteristic
- Spectral centroid
- Spectral moment
- Zero crossing rate
- Number of low energy frames
- Autocorrelation lag
- ....

MIR Methods - Data based

System learns a solution from (many) correct examples.

- Denoising
- Decomposition
- Transcription
- Genre detection
- AI-music generation
- ....

MIR Tools - Sonic visualiser

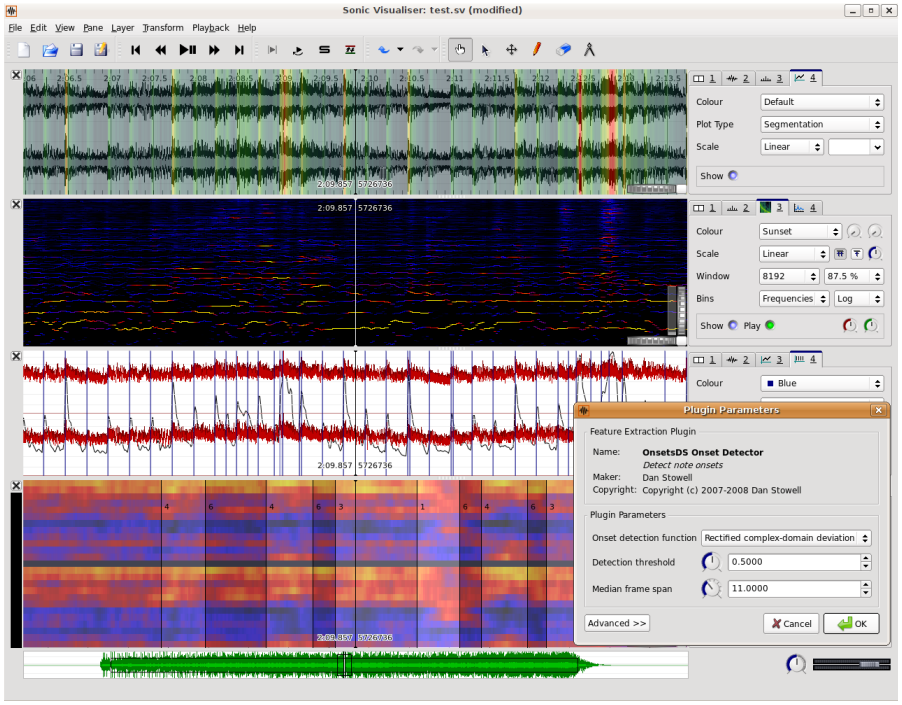


Fig: sonic visualiser

Sonic Visualiser is an application for viewing and analyzing the contents of audio files. It has support for:

- Beat tracking
- Cord estimation
- Melody detection
- Onset detection
- Annotations

Sonic visualiser

MIR Tools - Tartini

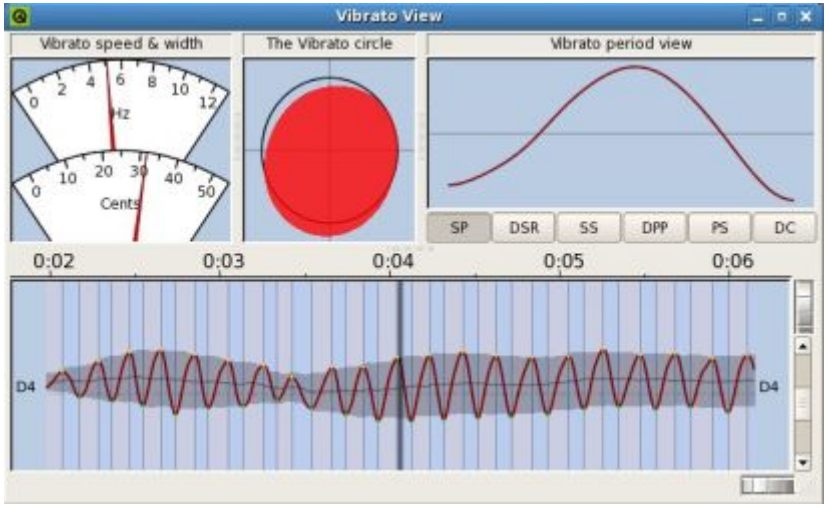


Fig: Tartini software

Specialized tool for (violin) pitch analysis

- Vibrato analysis
- Pitch contour
- Transcription

Tartini

MIR Tools - Music 21

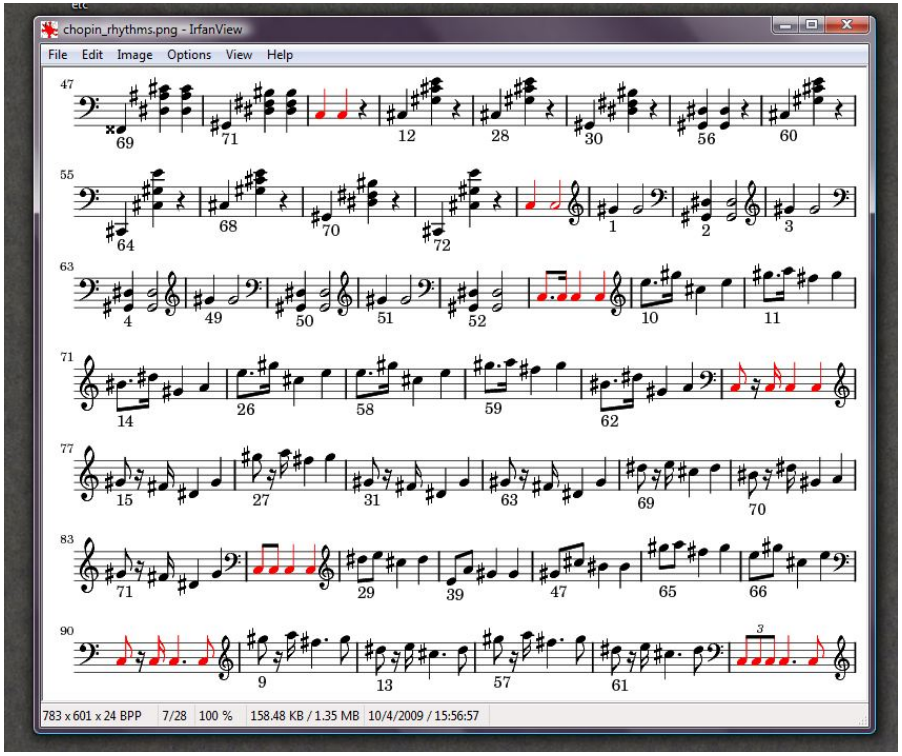


Fig: Music 21

A programming environment for symbolic music analysis

- Query rhythmic features
- Melodic contours
- Chord progressions

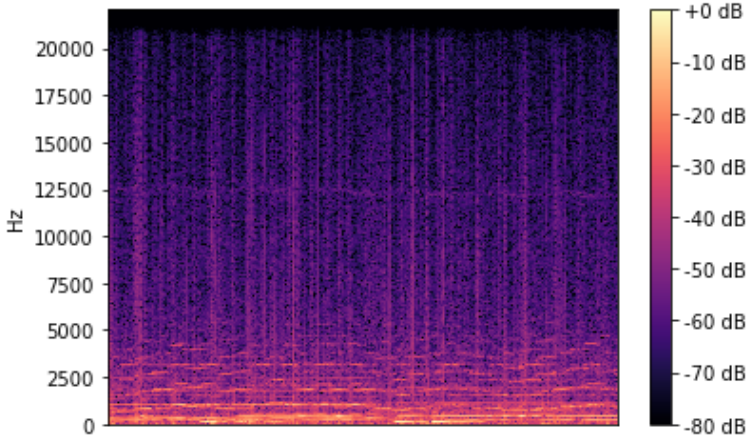
Music21

MIR Tools - MusicLM: Generating Music From Text

Caption	Generated audio
The main soundtrack of an arcade game. It is fast-paced and upbeat, with a catchy electric guitar riff. The music is repetitive and easy to remember, but with unexpected sounds, like cymbal crashes or drum rolls.	0:00 / 0:30
A fusion of reggaeton and electronic dance music, with a spacey, otherworldly sound. Induces the experience of being lost in space, and the music would be designed to evoke a sense of wonder and awe, while being danceable.	0:00 / 0:30
A rising synth is playing an arpeggio with a lot of reverb. It is backed by pads, sub bass line and soft drums. This song is full of synth sounds creating a soothing and adventurous atmosphere. It may be playing at a festival during two songs for a buildup.	0:00 / 0:30
Slow tempo, bass-and-drums-led reggae song. Sustained electric guitar. High-pitched bongos with ringing tones. Vocals are relaxed with a laid-back feel, very expressive.	0:00 / 0:30

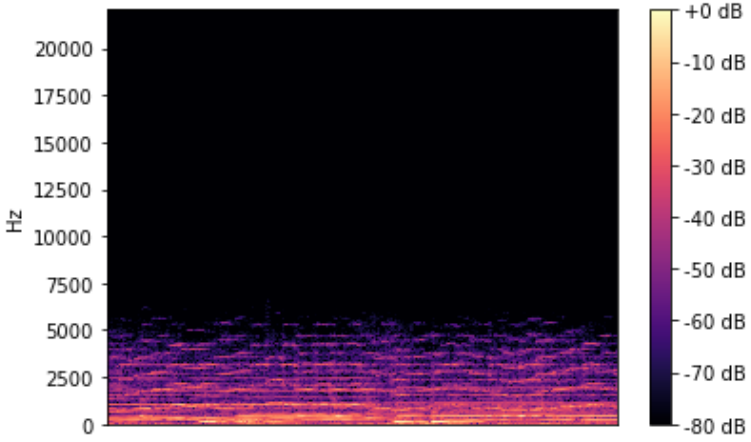
MIR Tools - Audio Denoising

Noisy input



0:00 / 0:05

Denoised



0:00 / 0:05

MIR tools - Apple Music Sing

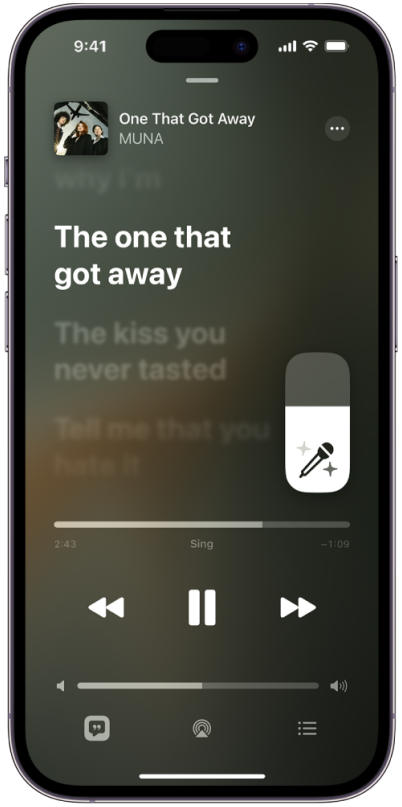


Fig: Apple Music Sing, surpress singing voice from any song

MIR tools - Moises.ai - Tools for musicians

Tools for musicians or analysis:

- Chord detection
- Source separation
- Tempo estimation

See: Moises.ai

Vid: source separation with Moises.ai

MIR Methods - Problems

MIR research is often limited by (over?) simplification:

- MIR focuses mainly on classical western art music or popular music with ethnocentric terminology like scores, chords, tone scale, chromagrams, instrumentation, rhythmical structures.
- It is mainly goal oriented and pragmatic (MIREX) without explaining processes. More engineering than science?
- Unclear which features correlate with which cognitive processes.
- It is mainly concerned with a limited, disembodied view on music: disregarding social interaction, movement, dance, the body, individual or cultural preferences.

MIR Methods - Problems

> ***Quote***
>
> *Essentially, all MIR-research is wrong, but some is useful*

PART II

PaPiOM

Patterns in Pitch Organization in Music

PaPiOM: Introduction

Patterns

“Patterns are fundamental in music around the world”

Why study cross-cultural patterns in music?

PaPiOM: Introduction

> ***Patterns***
>
> *"Patterns are fundamental in music around the world"*

Why study cross-cultural patterns in music?

- **Origins** of music

PaPiOM: Introduction

> ***Patterns***
>
> *"Patterns are fundamental in music around the world"*

Why study cross-cultural patterns in music?

- **Origins** of music
- **Evolution** of music

PaPiOM: Introduction

> ***Patterns***
>
> *"Patterns are fundamental in music around the world"*

Why study cross-cultural patterns in music?

- **Origins** of music
- **Evolution** of music
- **Non-human** musicality

PaPiOM: Introduction

> ***Patterns***
>
> *"Patterns are fundamental in music around the world"*

Why study cross-cultural patterns in music?

- **Origins** of music
- **Evolution** of music
- **Non-human** musicality
- **Nature-nurture** debates

PaPiOM: Introduction

> ***Patterns***
>
> *"Patterns are fundamental in music around the world"*

Why study cross-cultural patterns in music?

- **Origins** of music
- **Evolution** of music
- **Non-human** musicality
- **Nature-nurture** debates
- **Definition of music**

PaPiOM: Introduction

Action and Perception

Corpora

PaPiOM: Introduction

Action and Perception

- Context-poor

Corpora

- Context-rich

PaPiOM: Introduction

Action and Perception

- Context-poor
- Data-poor

Corpora

- Context-rich
- Data-rich

PaPiOM: Introduction

Action and Perception

- Context-poor
- Data-poor
- Controlled

Corpora

- Context-rich
- Data-rich
- Wild

PaPiOM: Introduction

Action and Perception

- Context-poor
- Data-poor
- Controlled
- Many studies

Corpora

- Context-rich
- Data-rich
- Wild
- **Few large scale**

PaPiOM: Introduction

PaPiOM: Introduction

Goal

*PaPiOM: perform **large-scale corpus-based studies** to identify patterns in pitch use in music. **Link** corpus-based findings with other findings.*

Study of patterns in music

Potential patterns common in music around the world:

Study of patterns in music

Potential patterns common in music around the world:

- Distinctness of pitches

Study of patterns in music

Potential patterns common in music around the world:

- Distinctness of pitches
- Octave equivalence

Study of patterns in music

Potential patterns common in music around the world:

- Distinctness of pitches
- Octave equivalence
- Number of pitch classes

Study of patterns in music

Potential patterns common in music around the world:

- Distinctness of pitches
- Octave equivalence
- Number of pitch classes
- Intervals between pitch classes

PaPiOM: Introduction

PaPiOM: Music Information

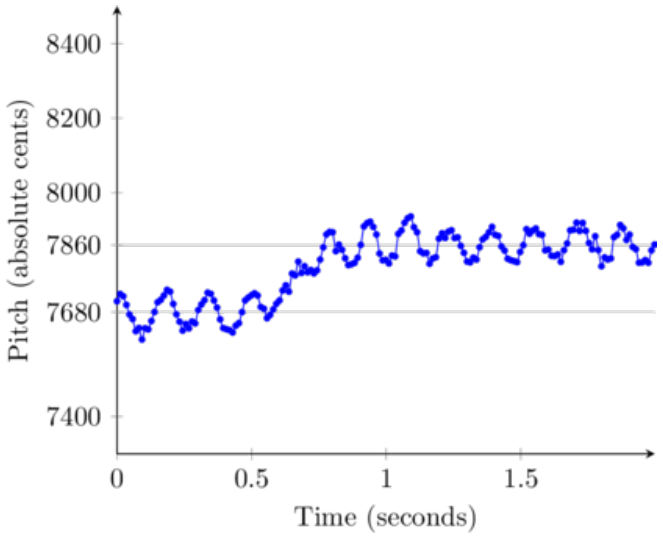
Starting from a recording we need a semi-automated way to extract:

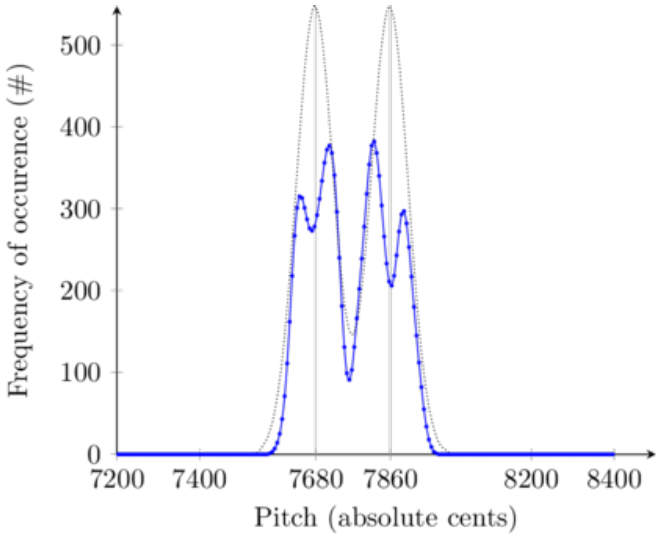
- Main pitch contour
- Pitch Class Set (scale) To interpret, other data is needed
- Meta-data: recording place, date, people, language, ...

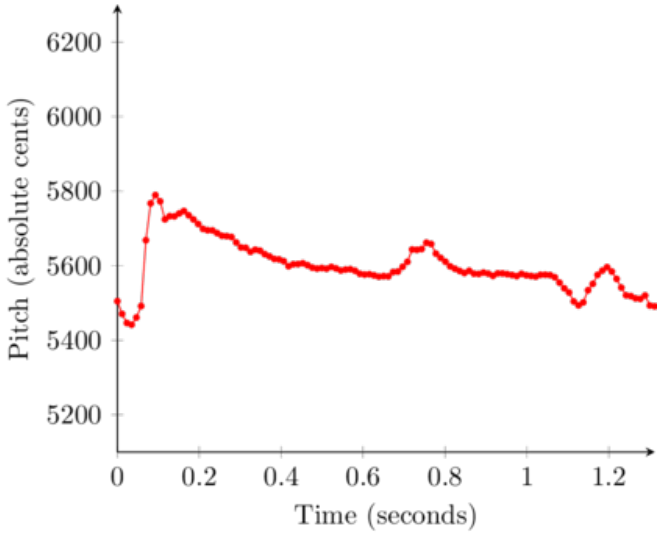
Methods: pitch tracking

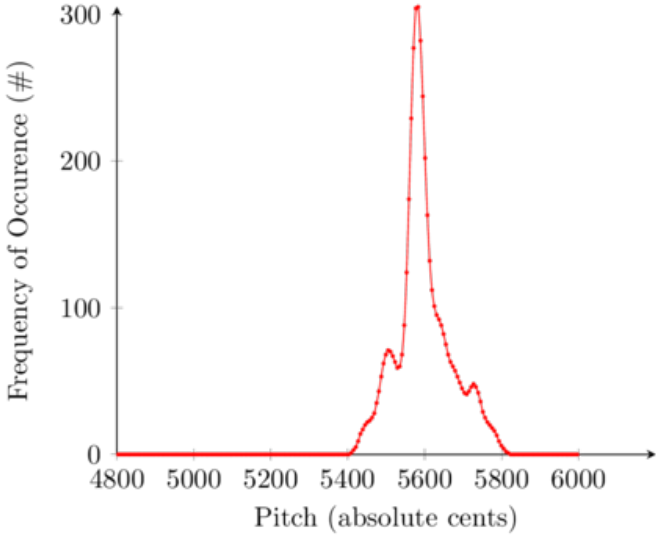
Methods: pitch tracking

Methods: pitch tracking - Pitch Class?

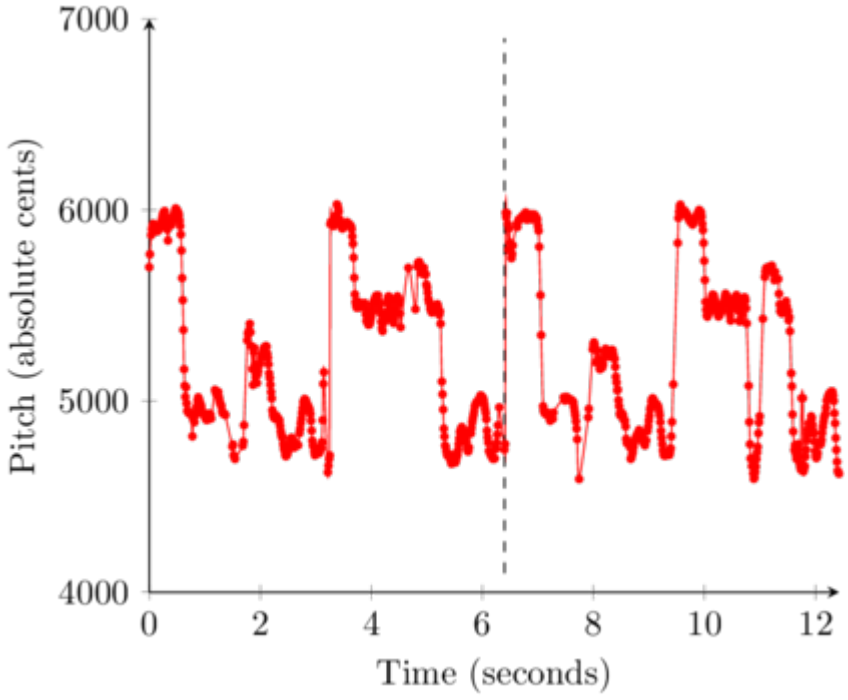


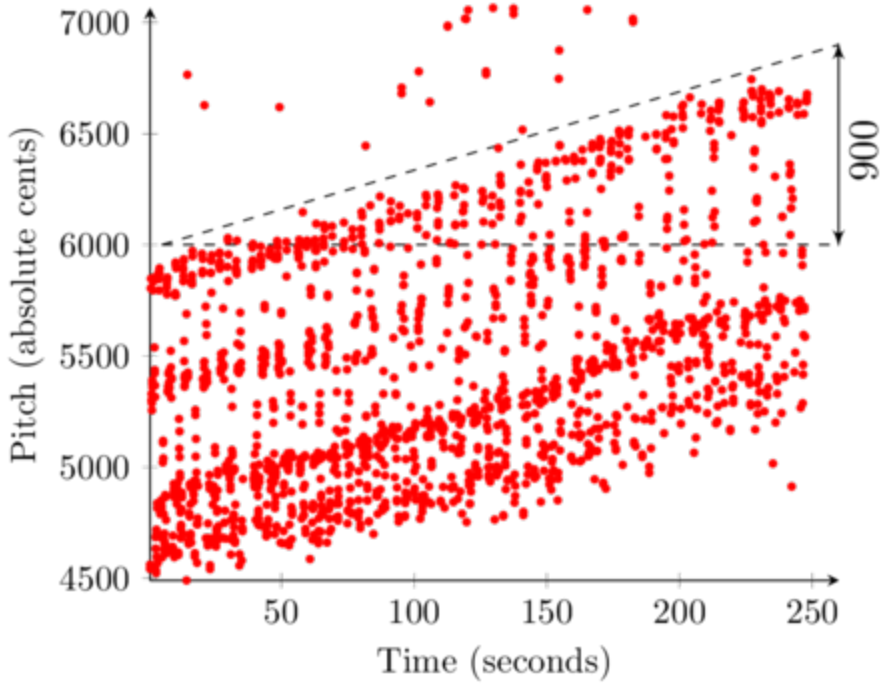



Pitch (absolute cents)
6200
6000
5800
5600
5400
5200
0
0.2
0.4
0.6
0.8
1
1.2
Time (seconds)


Frequency of Occurence (#)
300
200
100
0
4800
5000
5200
5400
5600
5800
6000
Pitch (absolute cents)

Methods: pitch tracking - Pitch Class?





PaPiOM: Case study - Wax Cylinder Recordings

+ Culturally diverse
+ Geographically spread
+ 'Uninfluenced'

- Noisy
- Short
- Unbalanced

Fig: Wax Cylinders, most poplar around 1896–1916 with a capacity of about 2 minutes.

PaPiOM: Case study - Wax Cylinder Recordings

IU - K.E. Laman, 1911, French Equatorial

0:00 / 0:07

IU - G. Herzog, 1930, Liberia

0:00 / 1:49

RMCA - P. Tempels, 1944, RDC

0:00 / 3:02

Moliner, E., & Välimäki, V. (2022, May). A two-stage u-net for high-fidelity denoising of historical recordings. ICASSP 2022-2022. IEEE.

RMCA - P. Tempels, 1944, RDC - Denoised

0:00 / 3:02

PaPiOM: Case study - Wax Cylinder Recordings

- The concept of pitch class is present
- 170 or 240 cents as building blocks
- 2 to 8 pitch classes
- Fifth is almost always present.

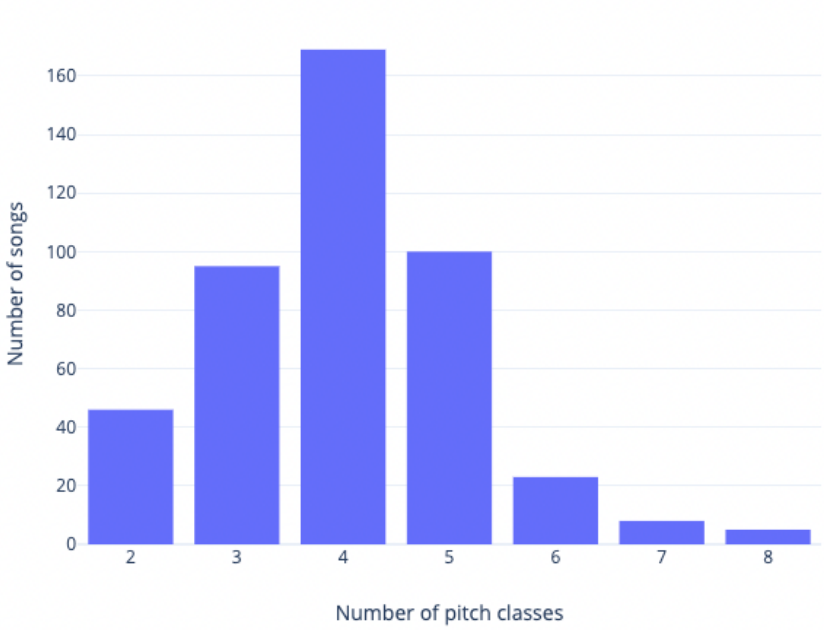


Fig: The number of songs for each pitch class set size.

PaPiOM: Case study - Wax Cylinder Recordings

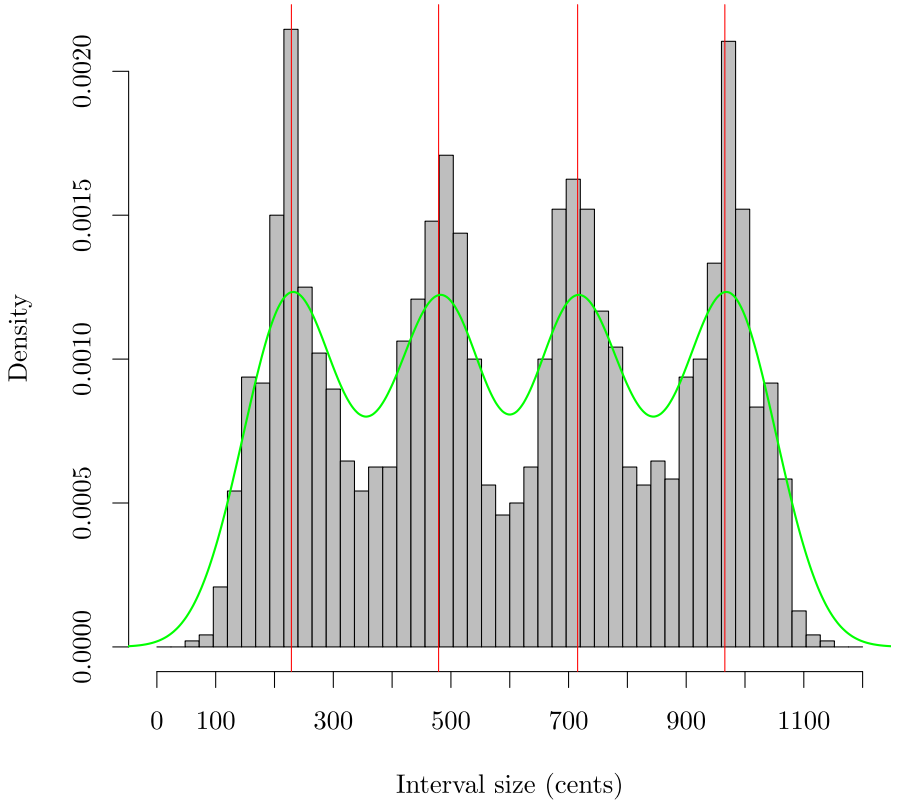


Fig: Pitch intervals between all pitch classes for recordings with **5** identified pitch classes.

PaPiOM: Case study - Wax Cylinder Recordings

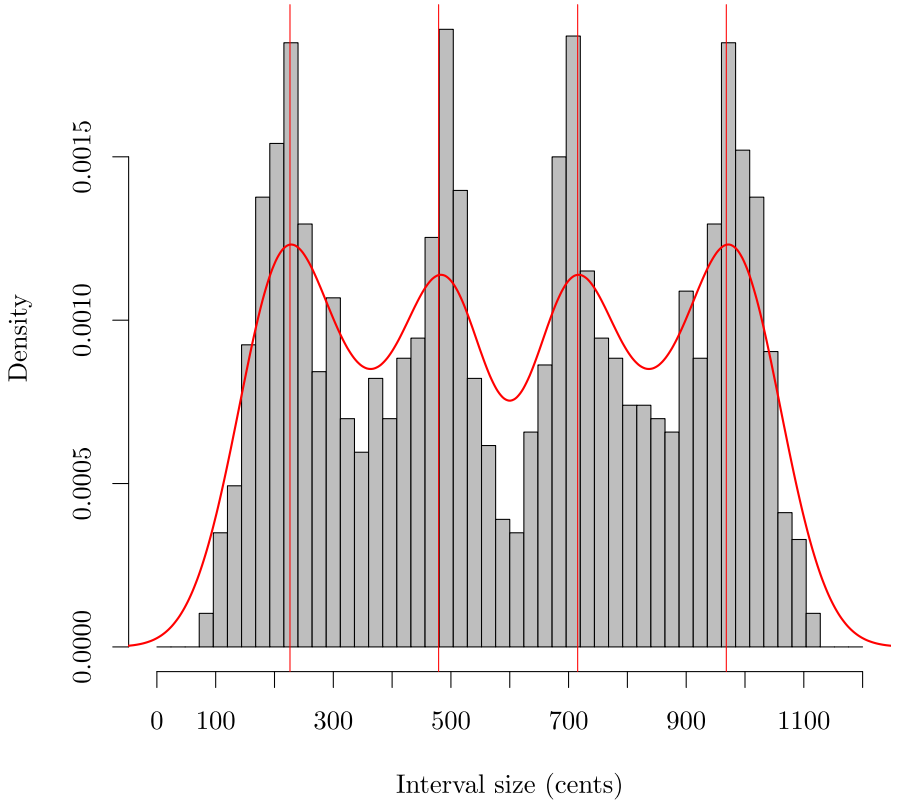


Fig: Pitch intervals between all pitch classes for recordings with **4** identified pitch classes.

PaPiOM: Case study - Wax Cylinder Recordings

- Bias: selection or recording technology
- Analysis assumes octave equivalence (498 = 702)
- Absolute pitch unclear
- Timbre ignored
- Unbalanced dataset
- Release dataset without audio
- Separating scale origins difficult
  - perception
  - production
  - information content minimization
  - transmission

PaPiOM: Summary

PaPiOM: Summary

- MIR task: find patterns in pitch use

PaPiOM: Summary

- MIR task: find patterns in pitch use
- Music information: signal to symbolic pitch classes

PaPiOM: Summary

- MIR task: find patterns in pitch use
- Music information: signal to symbolic pitch classes
- Feature based method: pitch tracking and processing

PaPiOM: Summary

- MIR task: find patterns in pitch use
- Music information: signal to symbolic pitch classes
- Feature based method: pitch tracking and processing
- Case study: 400 historic recordings reveal patterns

PART III

Duplicate detection for digital musicology

Duplicate detection: Introduction

What if we have an easy way to detect duplicate audio?

Duplicate detection: Introduction

0:00 / 4:03

Start or stop audio

Match found: ☐ Estimated reference time (s): `0.00`

```
Press the start button to begin audio fingerprinting.
```

Duplicate detection: introduction

0:00

Duplicate detection: Music Information

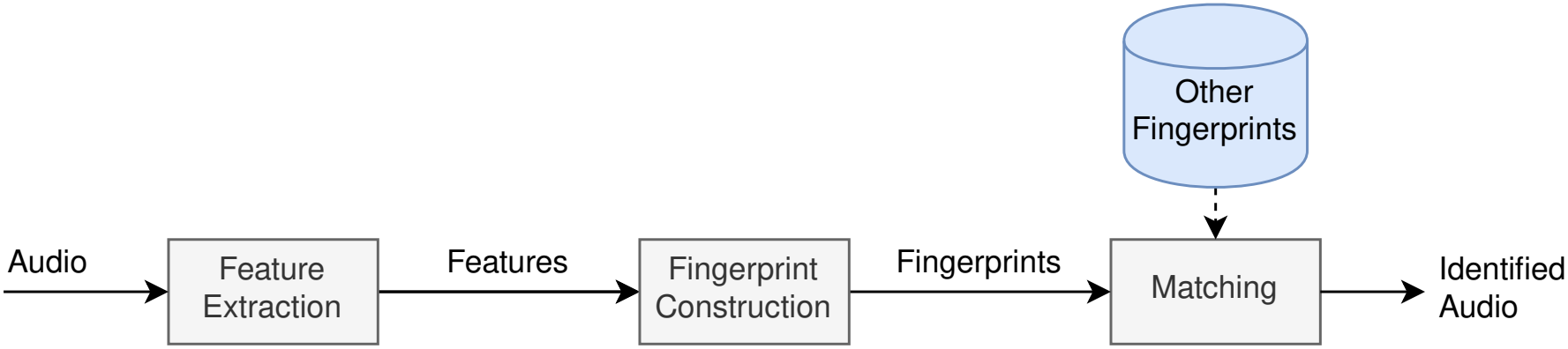


Fig: General acoustic fingerprinting schema. Audio to fingerprints.

Duplicate detection: methods

Duplicate detection: Music Information

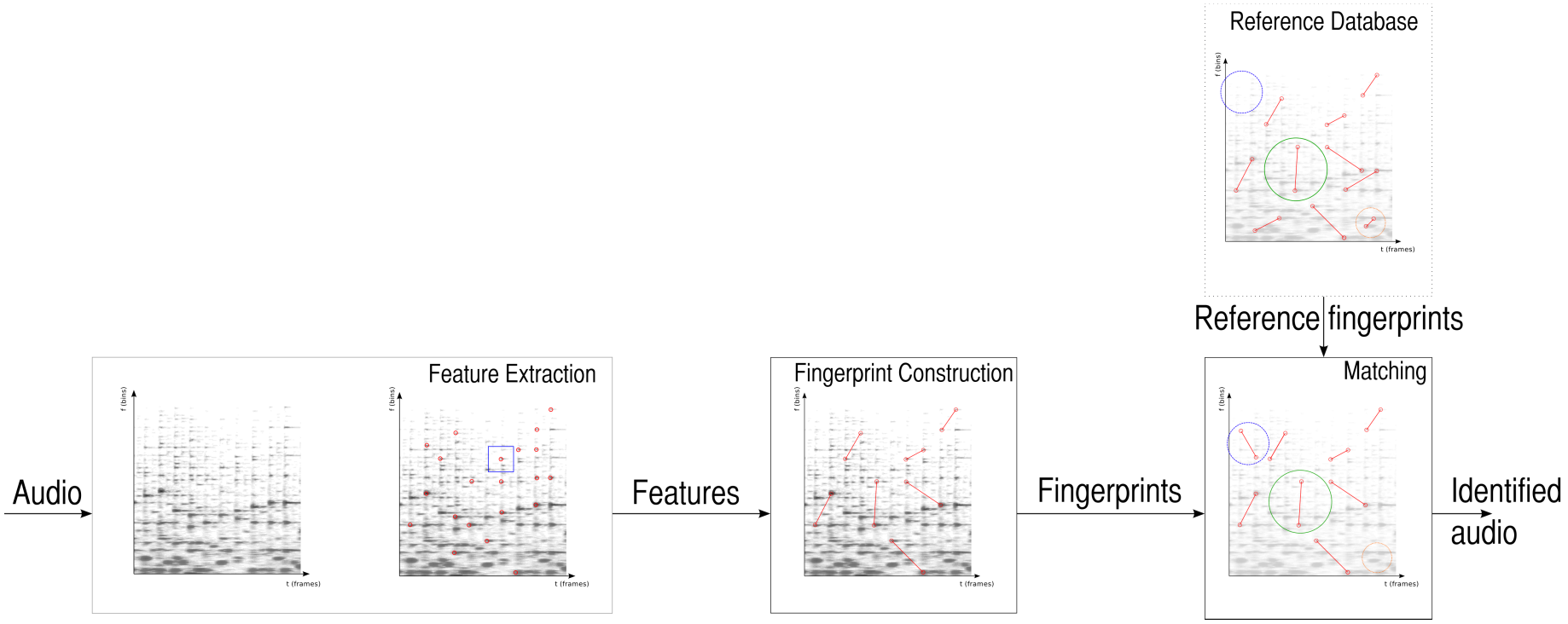


Fig: Spectral peak based acoustic fingerprinting schema

Duplicate detection: Applications - Musical structure

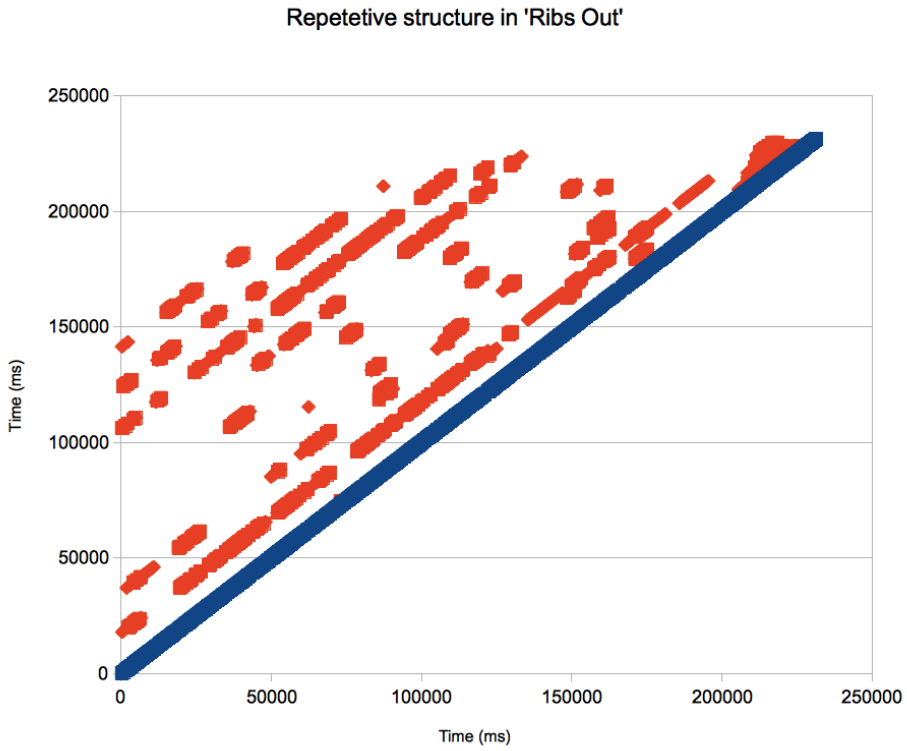


Fig: structure in 'Ribs Out' by Fuck Buttons

0:00 / 3:58

Duplicate detection: Applications - Exact repetition in music

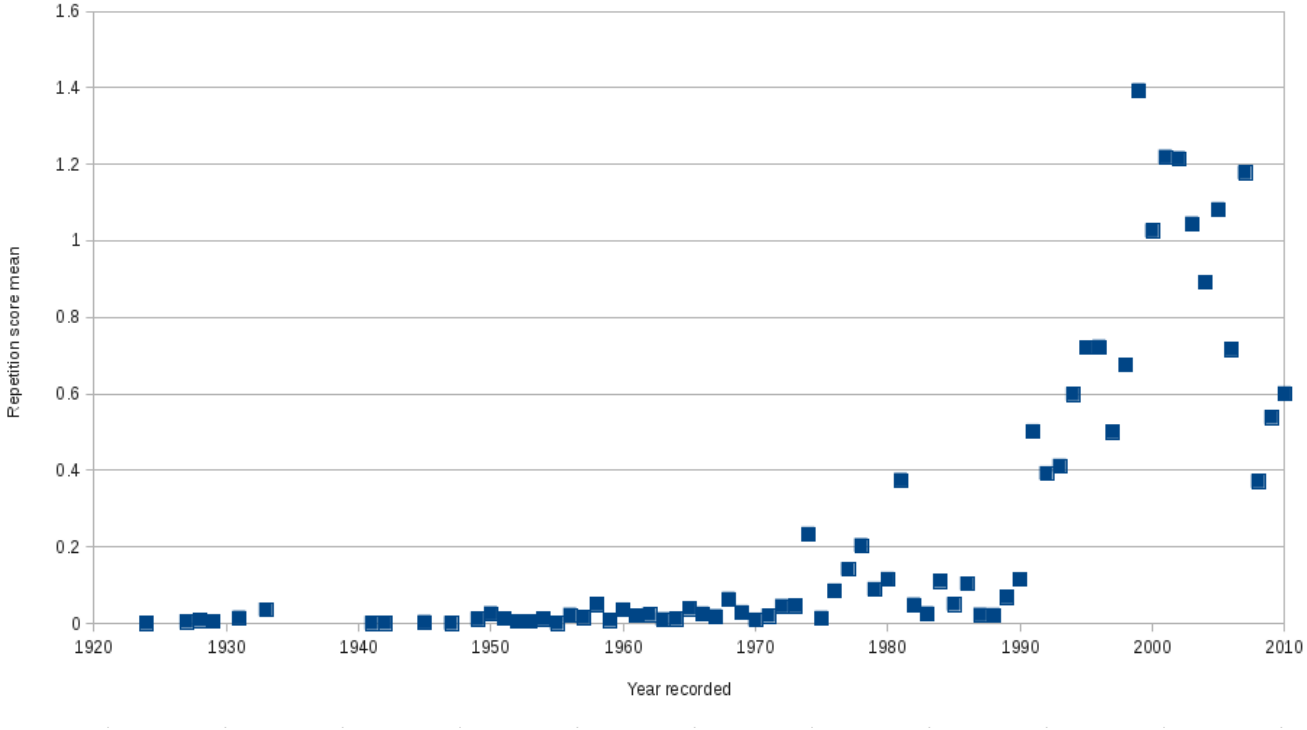


Fig: exact repetition in popular music over the years

Duplicate detection: versions

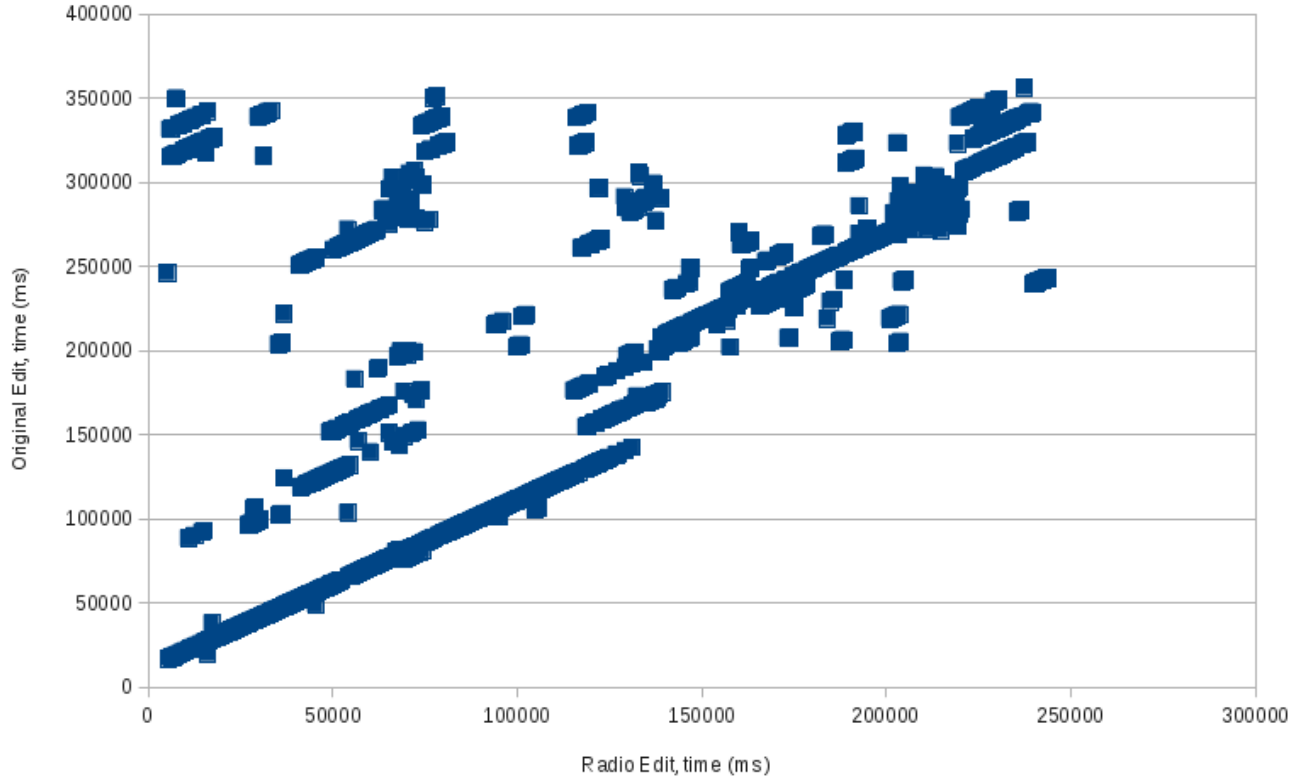


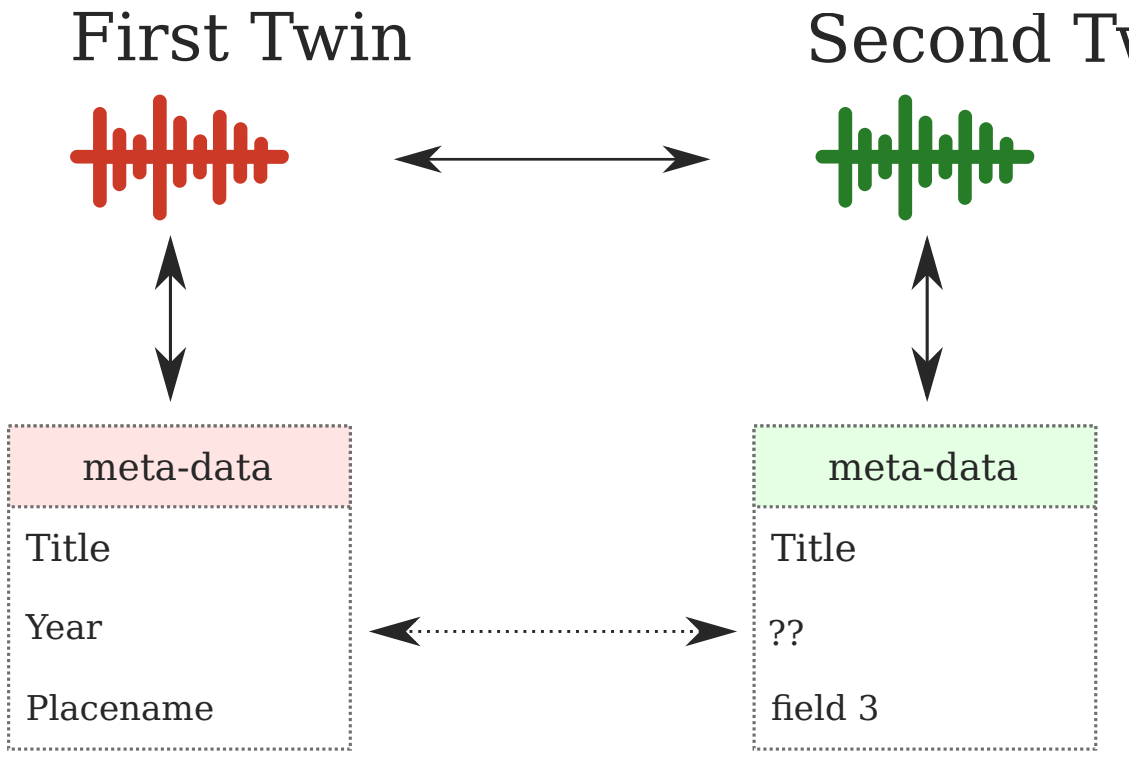
Fig: Radio vs original edit

Duplicate detection: Applications - DJ-set analysis

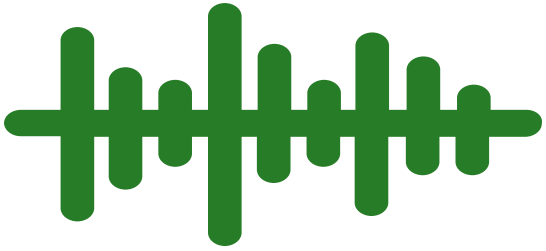
Duplicates after time-stretching, pitch-shifting and tempo change:

- Which parts of which songs were played and for how long
- Which modifications were applied (percentage modification of time and frequency)

Duplicate detection: Applications - Compare meta-data



First Tw



Second T

Duplicate detection: Applications - Twins

	First twin	Second twin
Audio	0:00 / 0:20	0:00 / 0:20
Year recorded	?	1949
Title	The daughter Mandega	?
People	Zezuru	Shona / Zezuru
Collector	Hugh Tracey	Hugh Tracey

Duplicate detection: Applications - Merge digital music archives

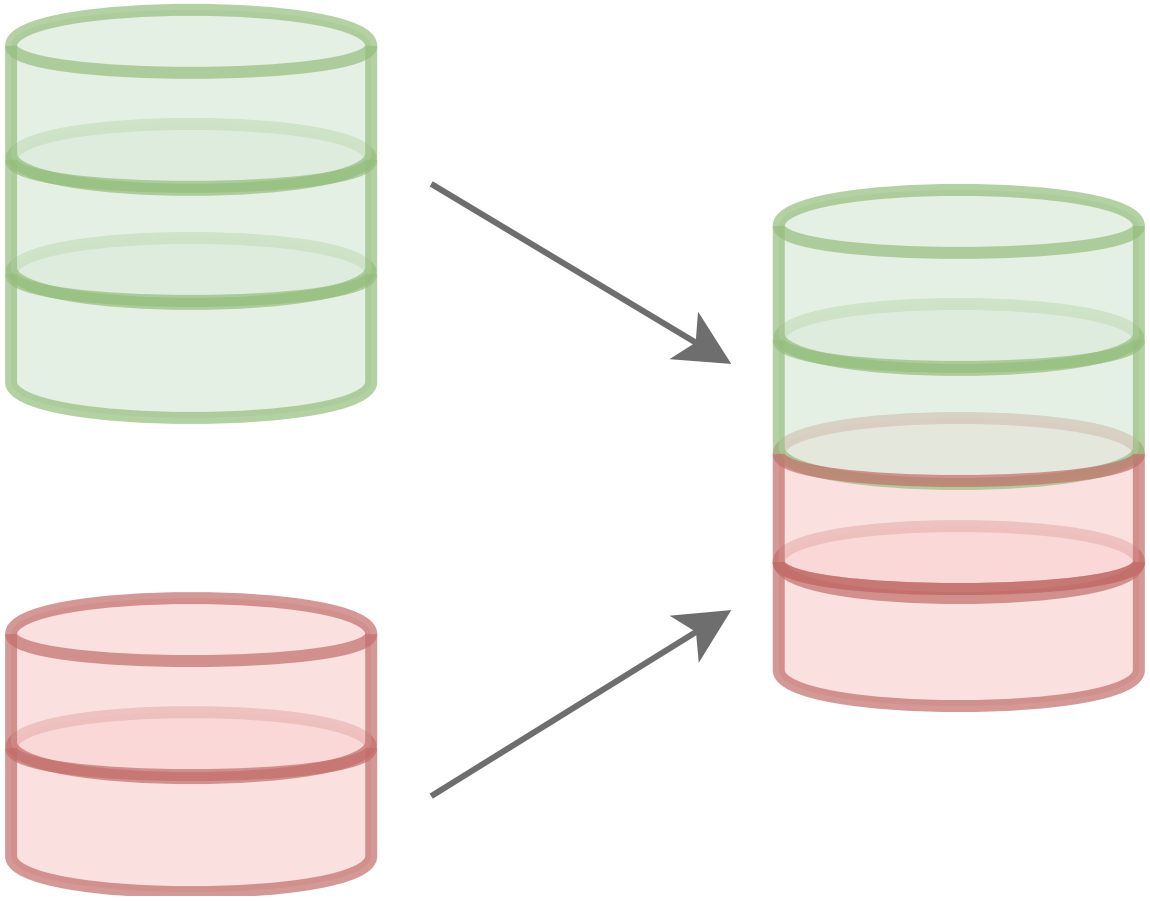


Fig: merge digital music archives: two + three = four

Duplicate detection: Applications - Improve listening experiences

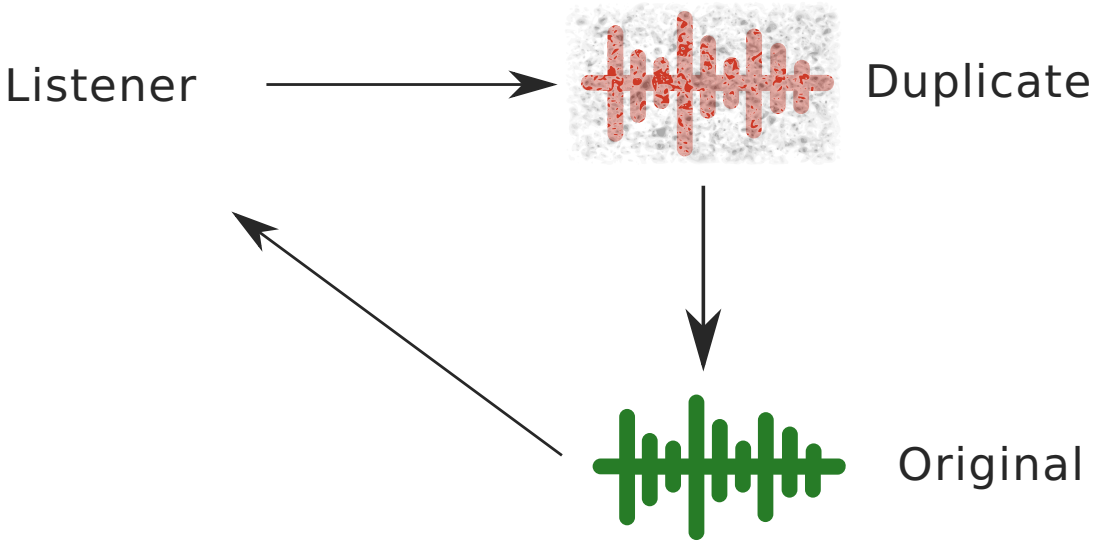


Fig: Redirect listeners to higher quality audio

Duplicate detection: Applications - Re-use segmentation

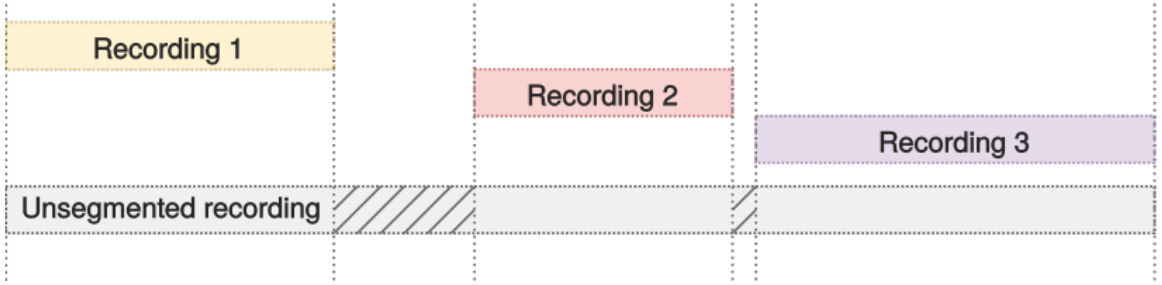


Fig: segmentation meta-data reuse.

0:00 / 1:04

Duplicate detection: Summary

Duplicate detection: Summary

- MIR task: find duplicate audio

Duplicate detection: Summary

- MIR task: find duplicate audio
- Music information: signal to symbolic, searchable fingerprints

Duplicate detection: Summary

- MIR task: find duplicate audio
- Music information: signal to symbolic, searchable fingerprints
- Feature based method: spectral peaks

Duplicate detection: Summary

- MIR task: find duplicate audio
- Music information: signal to symbolic, searchable fingerprints
- Feature based method: spectral peaks
- Many applications

General Summary

- **MIR** is the interdisciplinary science of extracting and processing information from music.
- **Symbols and signals** encode musical information.
- MIR offers opportunities for innovative (large scale) digital musicology and find patterns in music.
- Duplicate detection has many applications and can use spectral information to identify matches between audio.

Thanks!

joren.six@ugent.be