INTERNATIONAL AUDIO LABORATORIES ERLANGEN

AUDIO LABS

Music Information Retrieval

When Music Meets Computer Science

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

IWR-Colloquium, Heidelberg
01.02.2017

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

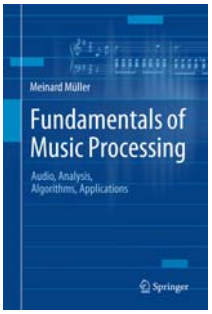
Meinard Müller

- 2001 PhD, Bonn University
- 2002/2003 Postdoc, Keio University, Japan
- 2007 Habilitation, Bonn University
  "*Information Retrieval for Music and Motion*"
- 2007-2012 Senior Researcher
  Max-Planck Institut für Informatik, Saarland
- 2012: Professor
  Semantic Audio Processing
  Universität Erlangen-Nürnberg

Group Members

- Stefan Balke
- Christian Dittmar
- Patricio López-Serrano
- Christof Weiß
- Frank Zalkow

Book: Fundamentals of Music Processing



Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., 30 illus. in color, hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

International Audio Laboratories Erlangen

Fraunhofer IIS

mp3

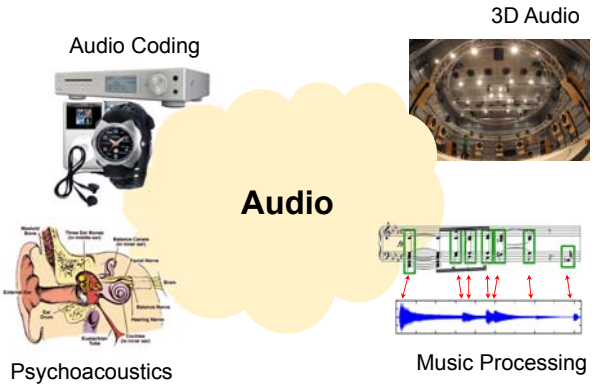
FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

AUDIO LABS

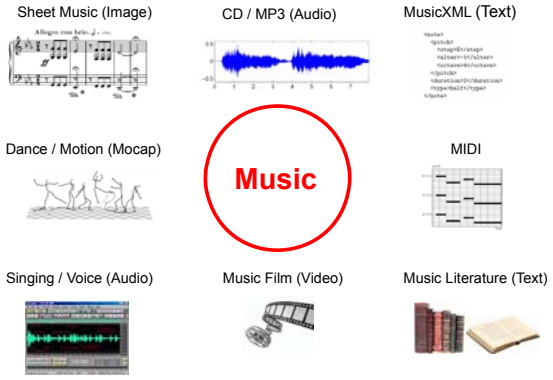
International Audio Laboratories Erlangen

Audio

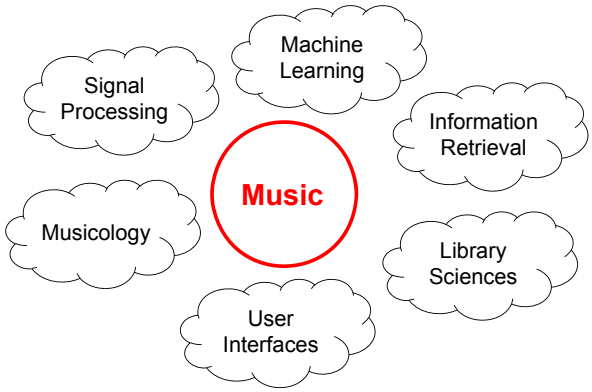
International Audio Laboratories Erlangen



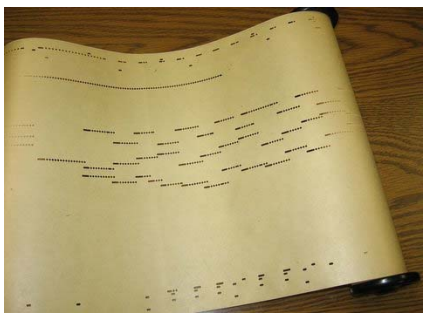
Music Information Retrieval



Music Information Retrieval



Piano Roll Representation

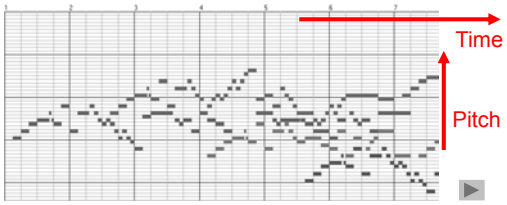


Player Piano (1900)

Piano Roll Representation (MIDI)

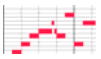
J.S. Bach, C-Major Fuge

(Well Tempered Piano, BWV 846)

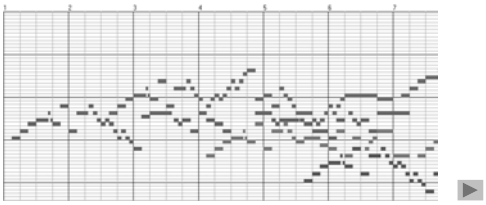


Piano Roll Representation (MIDI)

Query:

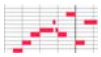


Goal: Find all occurrences of the query



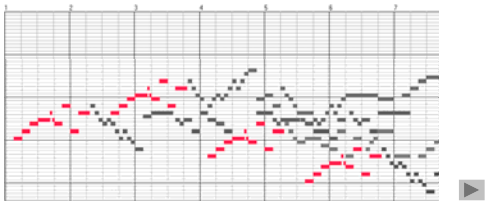
Piano Roll Representation (MIDI)

Query:

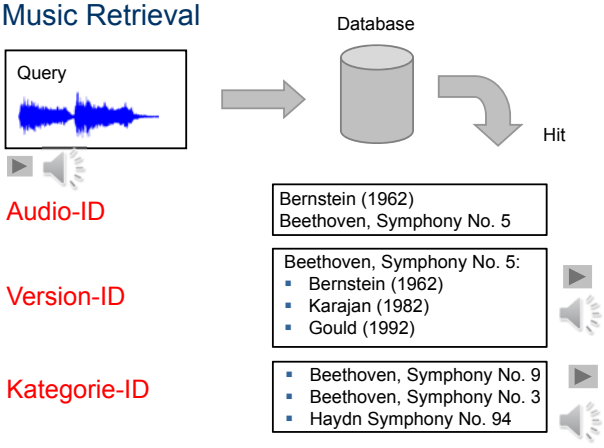


Goal: Find all occurrences of the query

Matches:

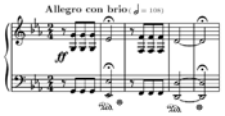


Music Retrieval



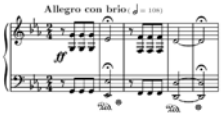
Music Synchronization: Audio-Audio

Beethoven's Fifth

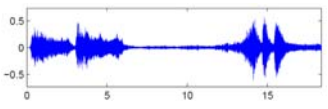


Music Synchronization: Audio-Audio

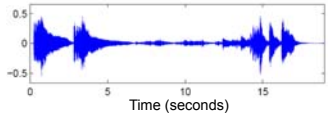
Beethoven's Fifth



Orchester (Karajan)

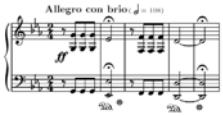


Piano (Scherbakov)



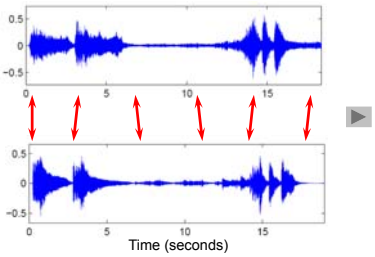
Music Synchronization: Audio-Audio

Beethoven's Fifth



Orchester (Karajan)

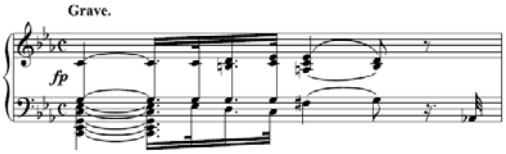
Piano (Scherbakov)



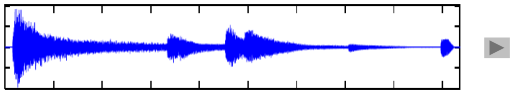
Application: Interpretation Switcher

Music Synchronization: Image-Audio

Image

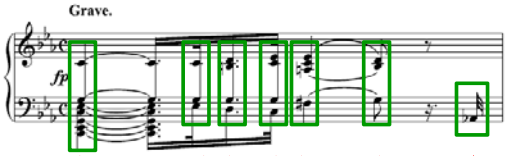


Audio

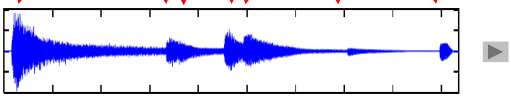


Music Synchronization: Image-Audio

Image

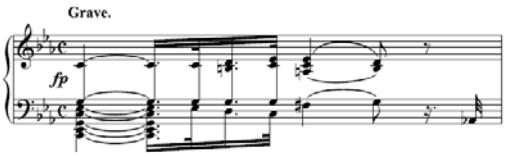


Audio

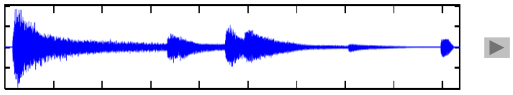


How to make the data comparable?

Image



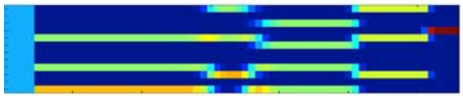
Audio



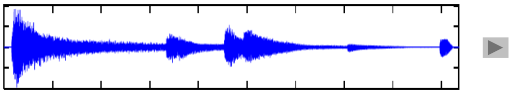
How to make the data comparable?

Image Processing: Optical Music Recognition

Image



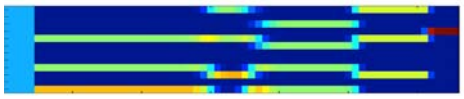
Audio



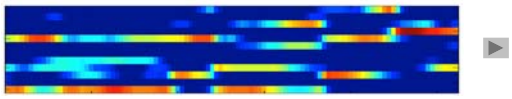
How to make the data comparable?

Image Processing: Optical Music Recognition

Image



Audio

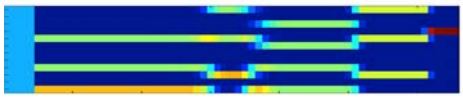


Audio Processing: Fourier Analyse

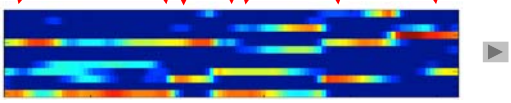
How to make the data comparable?

Image Processing: Optical Music Recognition

Image

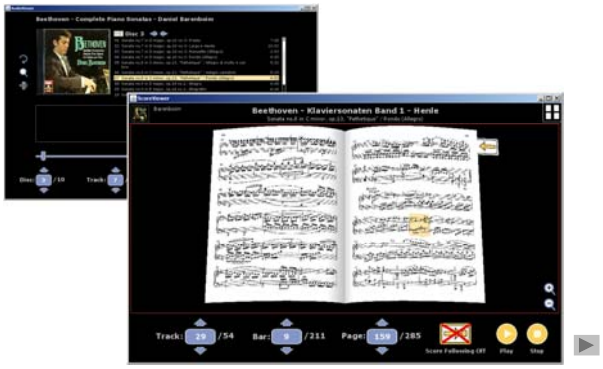


Audio



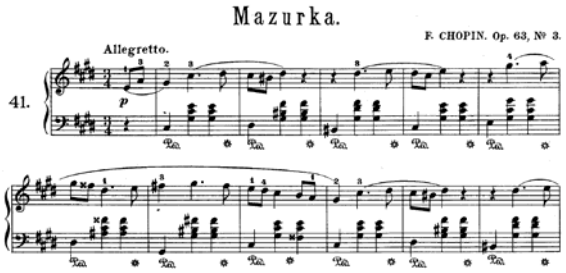
Audio Processing: Fourier Analyse

Application: Score Viewer



Why is Music Processing Challenging?

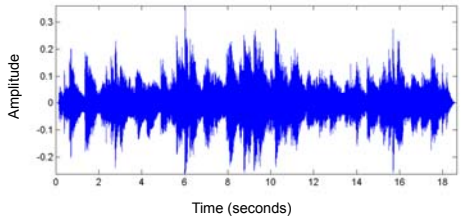
Example: Chopin, Mazurka Op. 63 No. 3



Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3

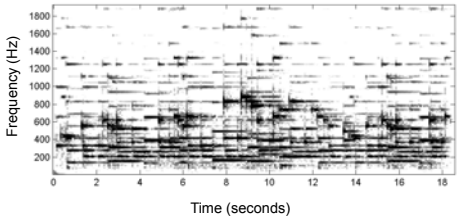
- Waveform



Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3

- Waveform / Spectrogram



Why is Music Processing Challenging?

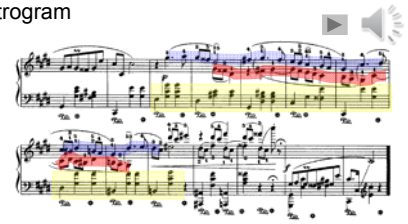
Example: Chopin, Mazurka Op. 63 No. 3

- Waveform / Spectrogram
- Performance
  - Tempo
  - Dynamics
  - Note deviations
  - Sustain pedal

Why is Music Processing Challenging?

Example: Chopin, Mazurka Op. 63 No. 3

- Waveform / Spectrogram
- Performance
  - Tempo
  - Dynamics
  - Note deviations
  - Sustain pedal
- Polyphony



Main Melody
Additional melody line
Accompaniment

Source Separation

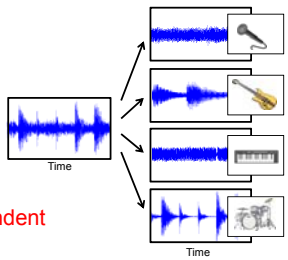
- Decomposition of audio stream into different sound sources
- Central task in digital signal processing
- “Cocktail party effect”

Source Separation

- Decomposition of audio stream into different sound sources
- Central task in digital signal processing
- “Cocktail party effect”
- Several input signals
- Sources are assumed to be statistically independent

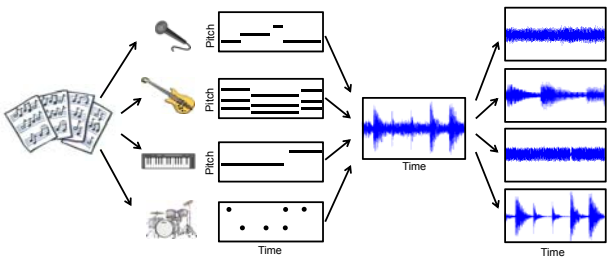
Source Separation (Music)

- Main melody, accompaniment, drum track
- Instrumental voices
- Individual note events
- Only mono or stereo
- Sources are often highly dependent



Score-Informed Source Separation

Exploit musical score to support seperation process



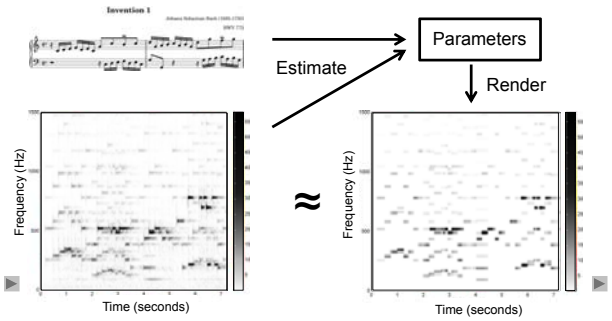
Score-Informed Audio Decomposition

Strategies

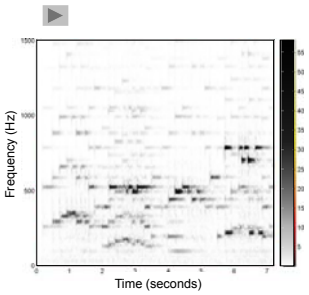
- Parametric model: Rebuild spectrogram
- NMF: Decompose spectrogram
- Melody tracking

Parametric Model Approach

Rebuild spectrogram information

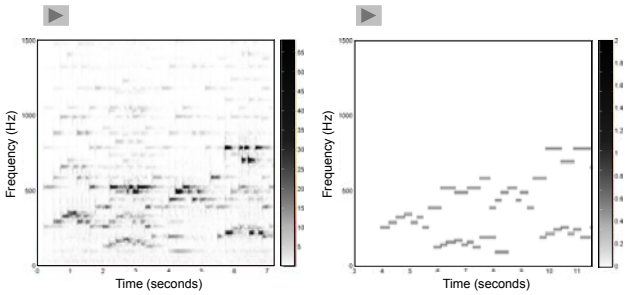


Parametric Model Approach



Original audio recording

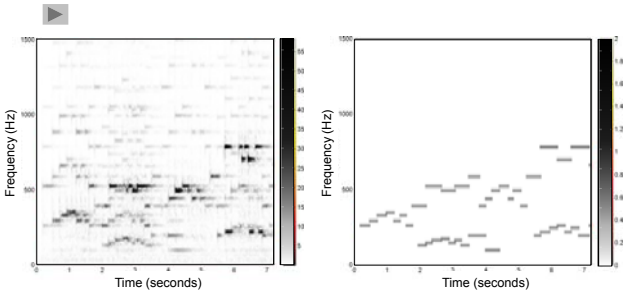
Parametric Model Approach



Original audio recording

Model initialized with score information

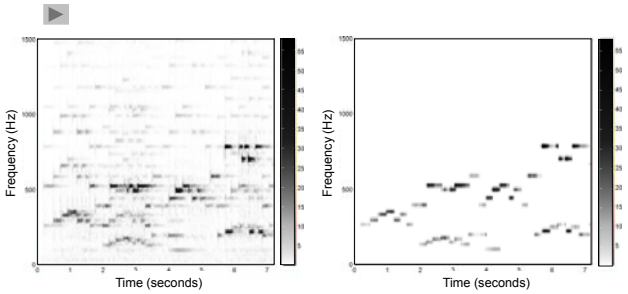
Parametric Model Approach



Original audio recording

Temporal synchronization

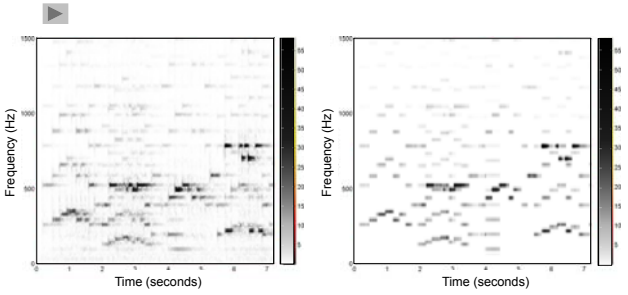
Parametric Model Approach



Original audio recording

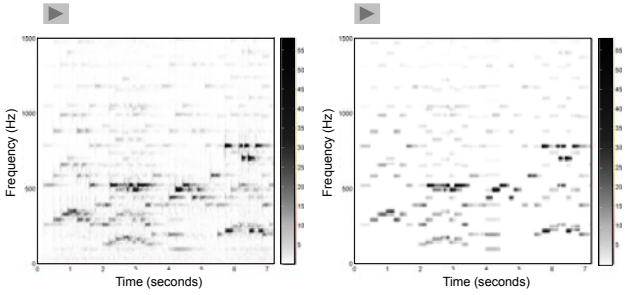
Dynamics parameters

Parametric Model Approach



Original audio recording | Timbre parameters

Parametric Model Approach



Original audio recording | Model after several iterations

Parametric Model Approach

Idea:

- Each note parameterizes a portion of the spectrogram
- Explicit model for
  - pitch + timing
  - dynamics
  - timbre + instrumentation

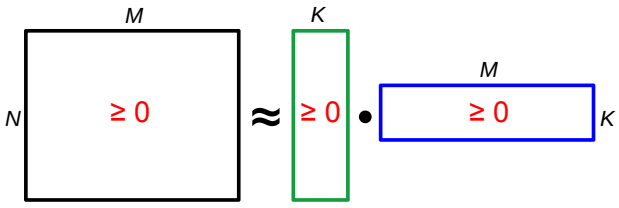
Advantages:

- Integration of musical knowledge is easily possible
- High degree of robustness

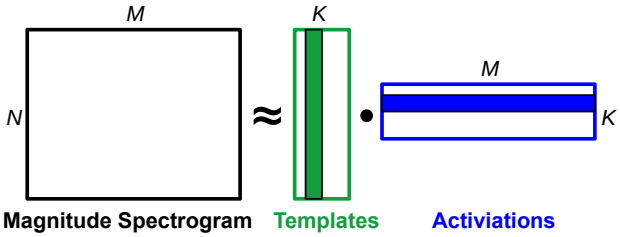
Problems:

- Inaccurate if model assumptions are violated
- Computationally expensive

NMF (Nonnegative Matrix Factorization)



NMF (Nonnegative Matrix Factorization)

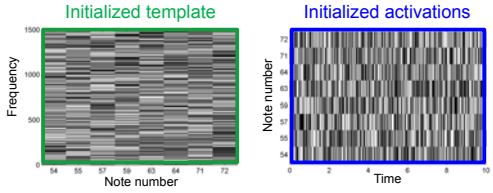


Magnitude Spectrogram | **Templates** | **Activiations**

Templates: **Pitch + Timbre** **"How does it sound"**

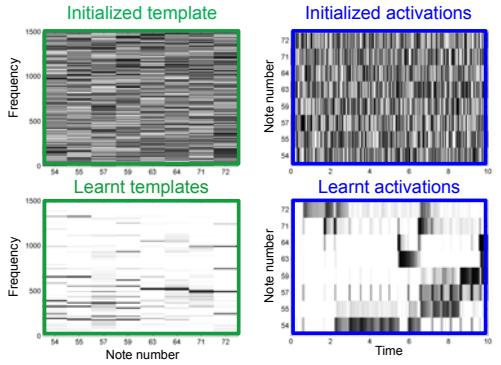
Activiations: **Onset time + Duration** **"When does it sound"**

NMF-Decomposition



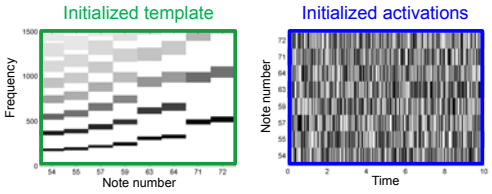
Random initialization

NMF-Decomposition



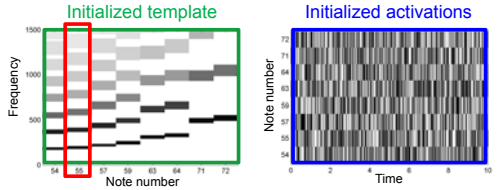
Random initialization → No semantic meaning

NMF-Decomposition



Template constraints

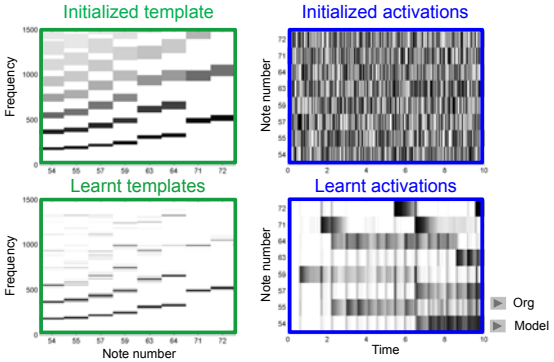
NMF-Decomposition



Template constraint for p=55

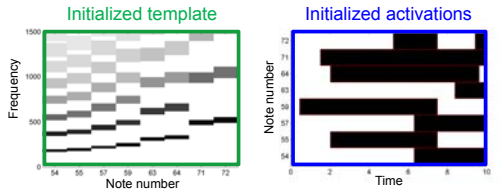
Template constraints

NMF-Decomposition



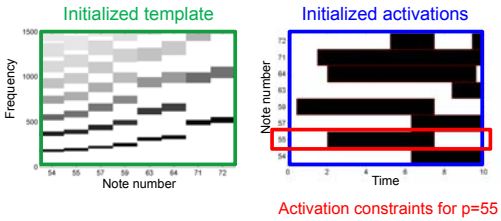
Template constraints → Semantic decomposition

NMF-Decomposition



Activation constraints

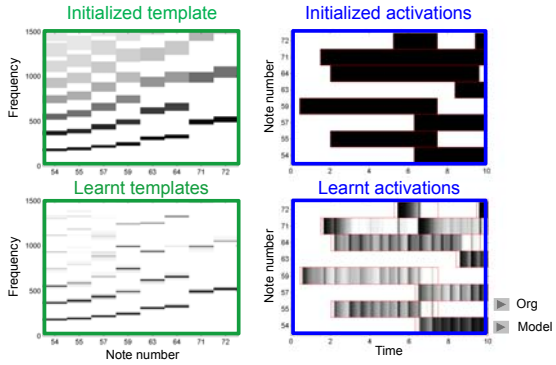
NMF-Decomposition



Activation constraints for p=55

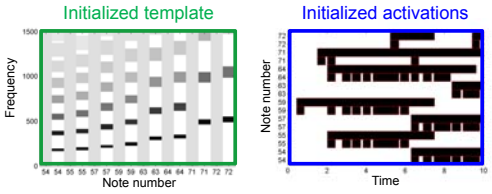
Activation constraints

NMF-Decomposition



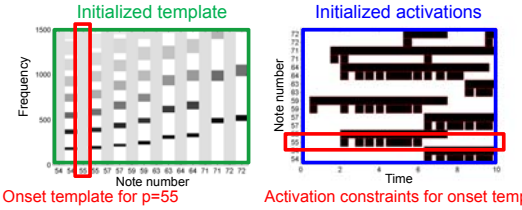
Activation constraints → NMF as refinement

NMF-Decomposition



Additional onset models → NMF as refinement

NMF-Decomposition

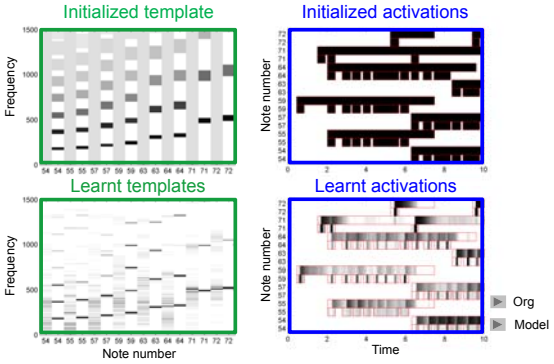


Onset template for p=55

Activation constraints for onset template

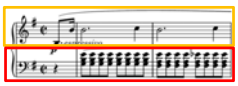
Additional onset models → NMF as refinement

NMF-Decomposition

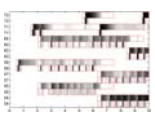


Additional onset models → NMF as refinement

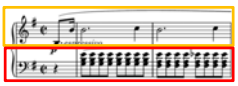
Score-Informed Source Separation



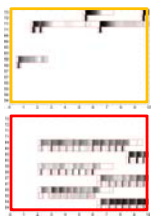
1. Split activation matrix



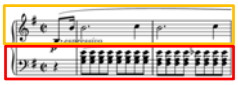
Score-Informed Source Separation



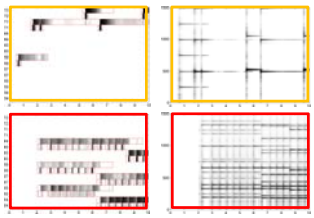
1. Split activation matrix



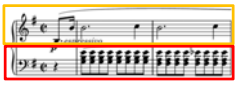
Score-Informed Source Separation



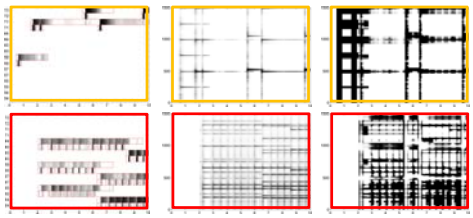
1. Split activation matrix
2. Model spectrogram for left/right



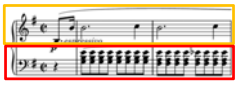
Score-Informed Source Separation



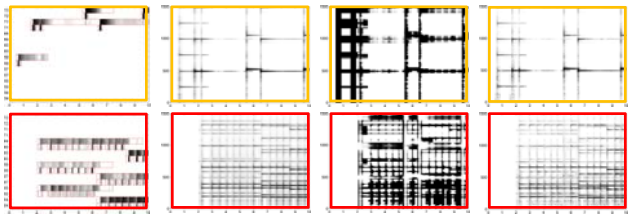
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right



Score-Informed Source Separation



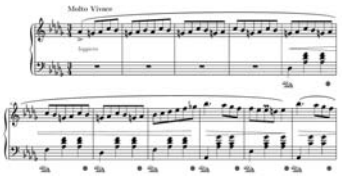
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right
4. Estimated spectrograms for left/right



Score-Informed Audio Decomposition

Application: Separating left and right hands for piano

Chopin, Waltz Op. 64, No. 1

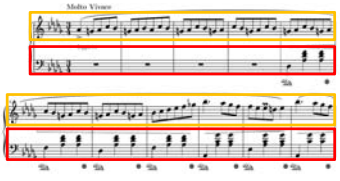


Original

Score-Informed Audio Decomposition

Application: Separating left and right hands for piano

Chopin, Waltz Op. 64, No. 1



Original

Left/right hand

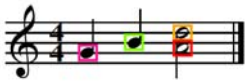
Right hand

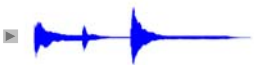
Left hand

Score-Informed Audio Decomposition

Parameterize audio signal using score's note events

Score

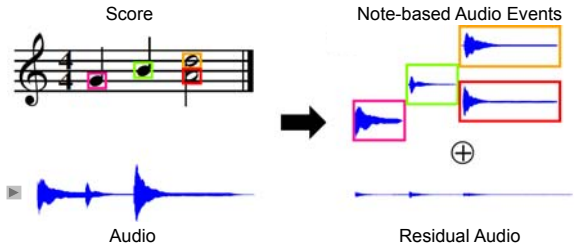




Audio

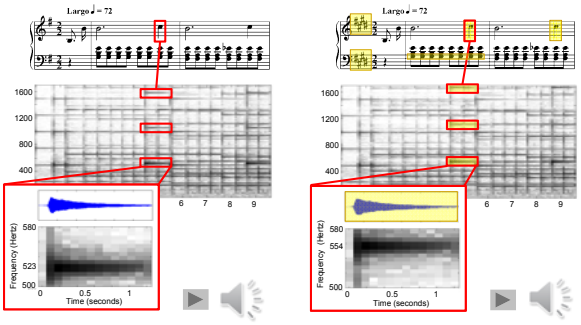
Score-Informed Audio Decomposition

Parameterize audio signal using score's note events



Score-Informed Audio Decomposition

Application: Audio editing



NMF-Decomposition

Idea:
- Factorization of spectrogram
- Implicit modeling of signal properties

Advantages:
- Flexible and easy to implement
- Efficient

Problems:
- Decompostion difficult to control
- Often no semantic meaning

Strategy: Multiplicative update rules allow for introducing hard constraints to control the decomposition

Audio Decomposition

Works reasonable

Audio Decomposition

Much more difficult

Audio Decomposition

Related problems:
- F0 estimation
- Melody tracking
- Human voice
- Vibrato

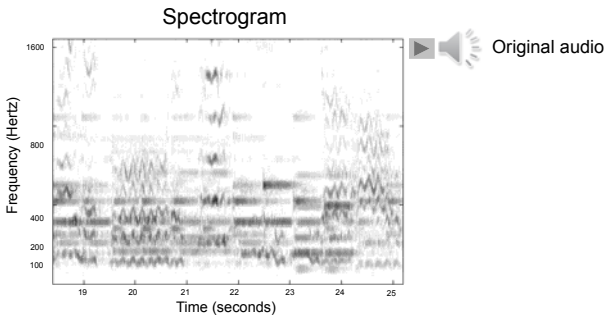
Much more difficult

Melody Tracking

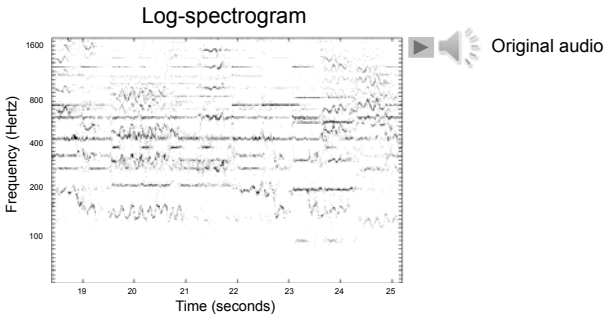
Justin Salamon and Emilia Gómez:
Melody extraction from polyphonic music signals using pitch contour characteristics.
IEEE-TASLP 2012

- F0 estimation
- Voice detection
- Pitch contour creation
- Melody selection

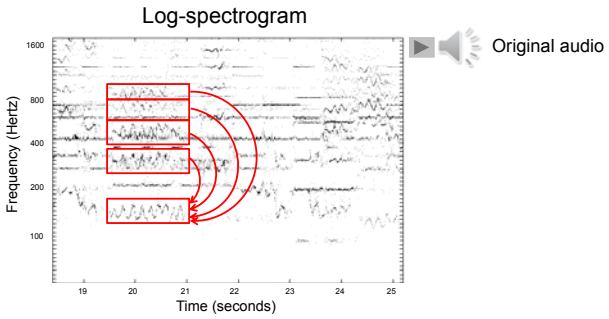
F0 Estimation



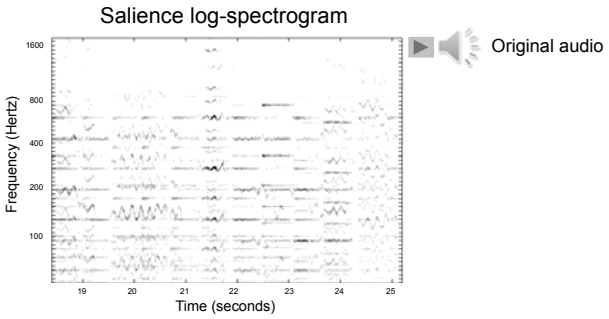
F0 Estimation



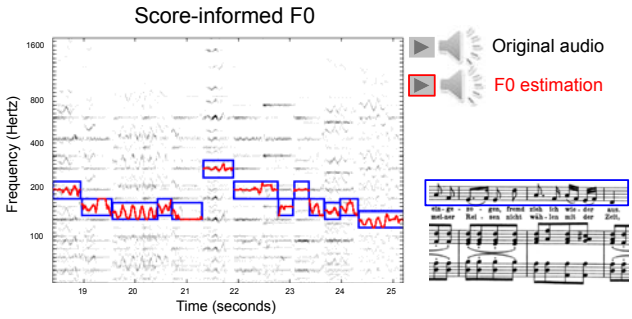
F0 Estimation



F0 Estimation

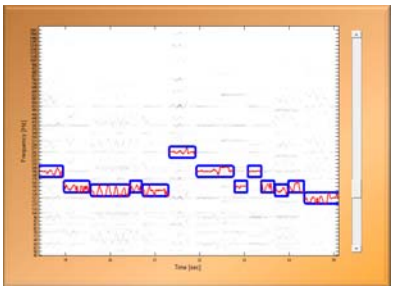


F0 Estimation



Score-Informed Source Separation

Application: Voice separation and editing

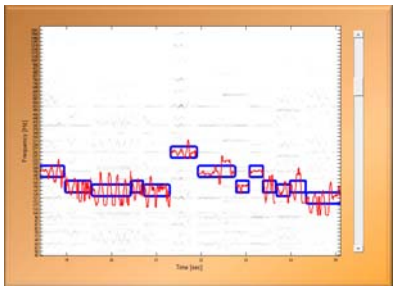


Original audio

Separated voice

Score-Informed Source Separation

Application: Voice separation and editing



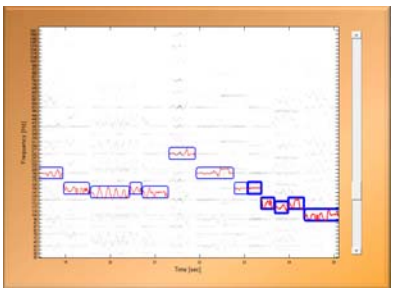
Original audio

Separated voice

Amplified vibrato

Score-Informed Source Separation

Application: Voice separation and editing



Original audio

Separated voice

Amplified vibrato

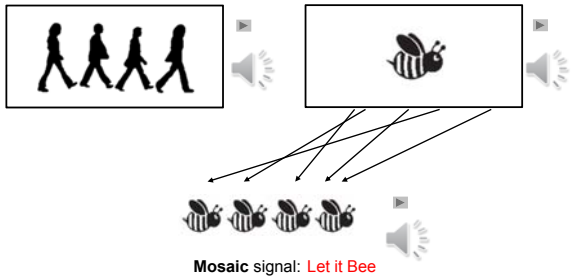
Pitch shift

Audio Mosaicing

Audio Mosaicing

Target signal: Beatles–Let it be

Source signal: Bees

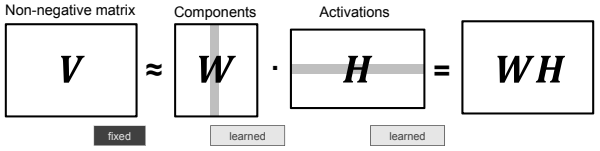


Mosaic signal: Let it Bee

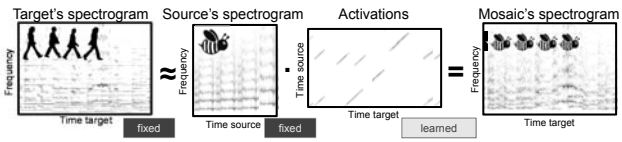
NMF-Inspired Audio Mosaicing

[Driedger et al. ISMIR 2015]

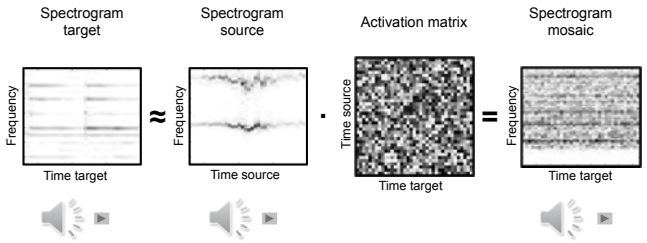
Non-negative matrix factorization (NMF)



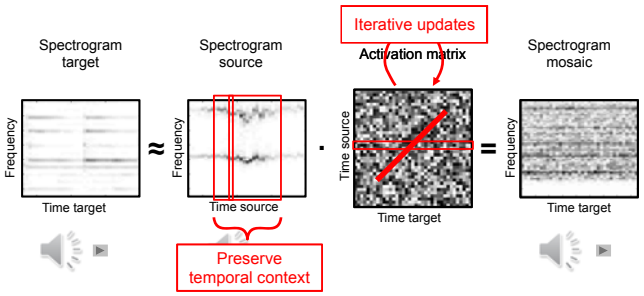
Proposed audio mosaicing approach



NMF-Inspired Audio Mosaicing



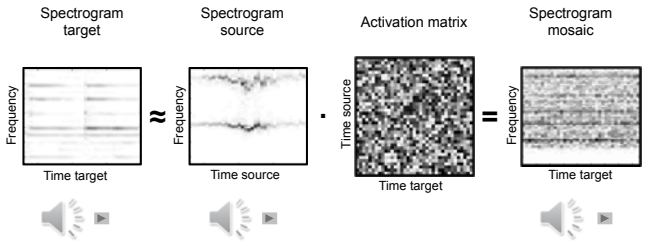
NMF-Inspired Audio Mosaicing



Core idea: support the development of sparse diagonal activation structures

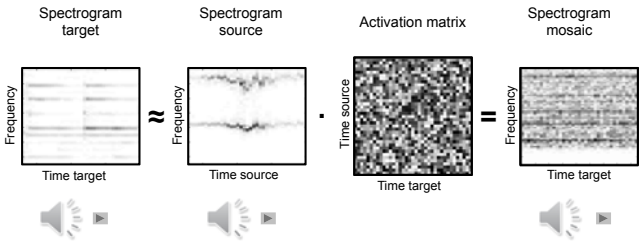
NMF-Inspired Audio Mosaicing

[Driedger et al. ISMIR 2015]

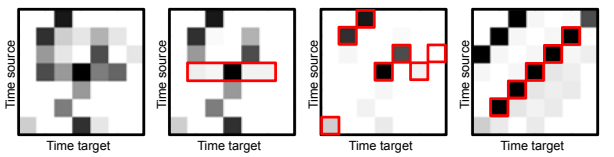


NMF-Inspired Audio Mosaicing

[Driedger et al. ISMIR 2015]



NMF with Additional Update Rules



Activation matrix

Repetition-restricted activation matrix

Neighborhood constraints

Polyphony-restricted activation matrix

Column constraints

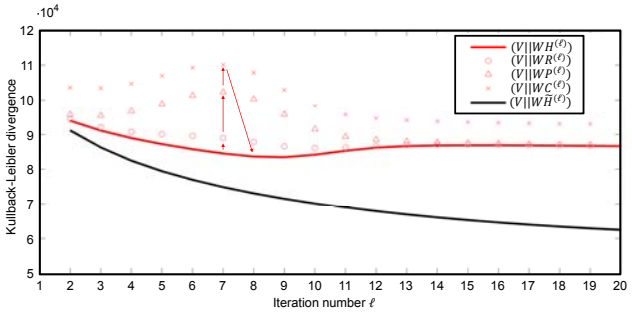
Continuity-enhanced activation matrix

Diagonal smoothing

- Constraints are enforced by additional update rules
- Additional rules are interleaved with standard NMF update rules
- Soft alternative to NMFD

NMF with Additional Update Rules

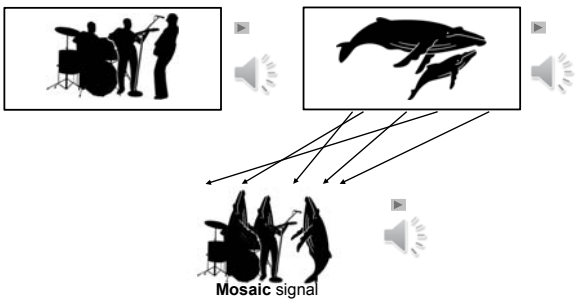
Kullback-Leibler Divergence between Target and Mosaic



Audio Mosaicing

Target signal: Chic–Good times

Source signal: Whales

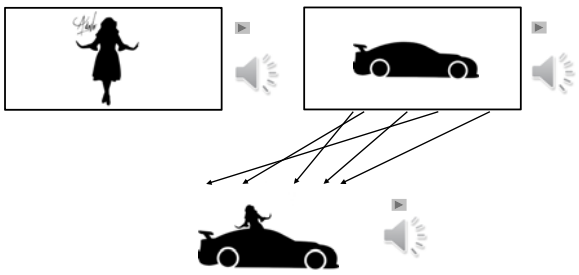


Mosaic signal

Audio Mosaicing

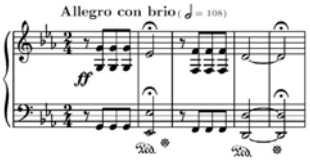
Target signal: Adele–Rolling in the Deep

Source signal: Race car



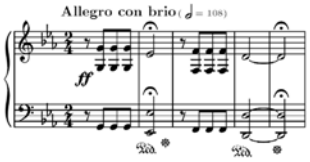
Mosaic signal

Motivic Similarity



Beethoven's Fifth (1st Mov.)

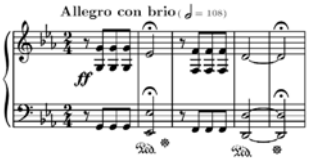
Motivic Similarity



Beethoven's Fifth (1st Mov.)

Beethoven's Fifth (3rd Mov.)

Motivic Similarity

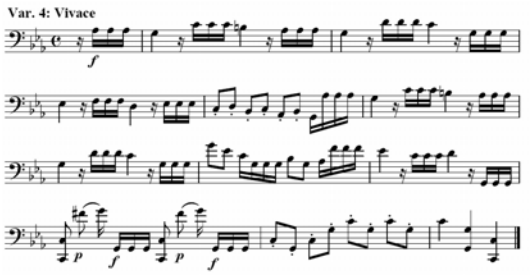


Beethoven's Fifth (1st Mov.)

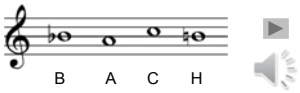
Beethoven's Fifth (3rd Mov.)

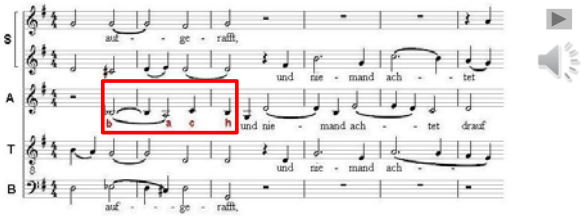
Beethoven's Appassionata

Motivic Similarity

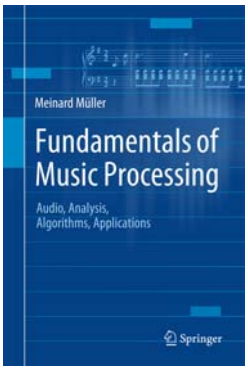


Motivic Similarity





Book: 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 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
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de