



**Tutorial T3**  
**A Basic Introduction to Audio-Related Music Information Retrieval**

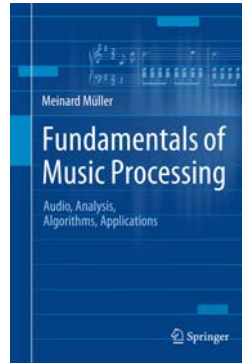
## Audio Decomposition

**Meinard Müller, Christof Weiß**

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



## 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](http://www.music-processing.de)

## Book: Fundamentals of Music Processing

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

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

Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

## Book: Fundamentals of Music Processing

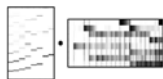
Chapter	Music Processing Scenario
1	Music Representations
2	Fourier Analysis of Signals
3	Music Synchronization
4	Music Structure Analysis
5	Chord Recognition
6	Tempo and Beat Tracking
7	Content-Based Audio Retrieval
8	Musically Informed Audio Decomposition

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

Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

## Chapter 8: Audio Decomposition

- 8.1 Harmonic-Percussive Separation
- 8.2 Melody Extraction
- 8.3 NMF-Based Audio Decomposition
- 8.4 Further Notes



In the final Chapter 8 on audio decomposition, we present a challenging research direction that is closely related to source separation. Within this wide research area, we consider three subproblems: harmonic-percussive separation, main melody extraction, and score-informed audio decomposition. Within these scenarios, we discuss a number of key techniques including instantaneous frequency estimation, fundamental frequency (F0) estimation, spectrogram inversion, and nonnegative matrix factorization (NMF). Furthermore, we encounter a number of acoustic and musical properties of audio recordings that have been introduced and discussed in previous chapters, which rounds off the book.

## Why is Music Processing Challenging?

**Example:** Chopin, Mazurka Op. 63 No. 3

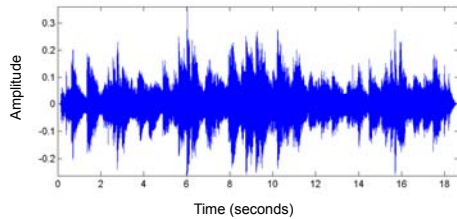
**Mazurka.**  
F. CHOPIN, Op. 63, No. 3

Allegretto.

## Why is Music Processing Challenging?

**Example:** Chopin, Mazurka Op. 63 No. 3

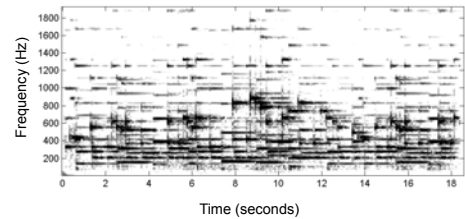
- Waveform



## 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”

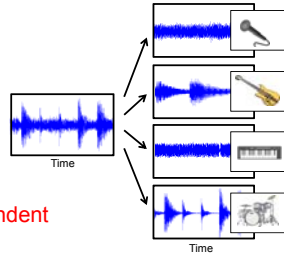


## Source Separation

- 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)

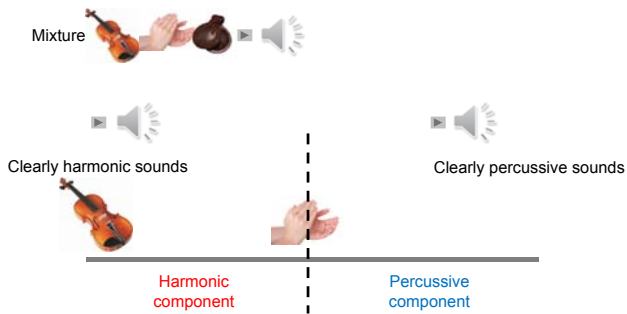
- Main melody, accompaniment, drum track
- Instrumental voices
- Individual note events
- Only mono or stereo
- Sources are often highly dependent



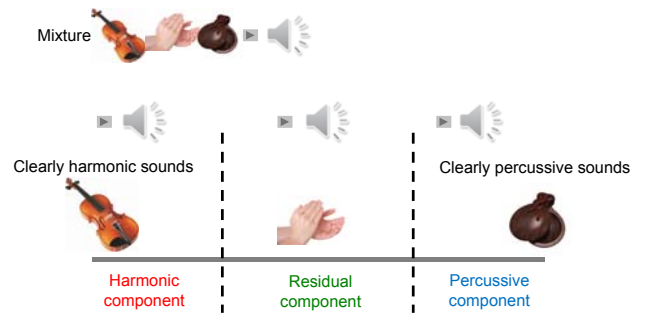
## Harmonic-Percussive Decomposition



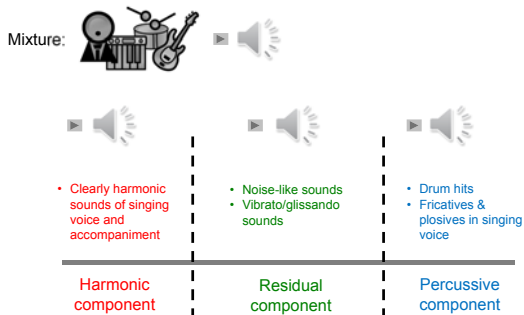
## Harmonic-Percussive Decomposition



## Harmonic-Percussive Decomposition

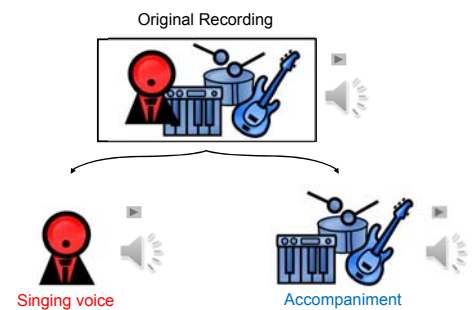


## Harmonic-Percussive Decomposition

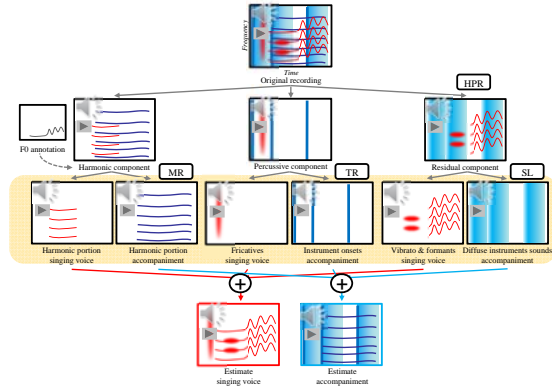


Literature: [Driedger/Müller/Disch, ISMIR 2014]  
 Demo: <https://www.audiolabs-erlangen.de/resources/2014-ISMIR-ExtHPSep/>

## Singing Voice Extraction

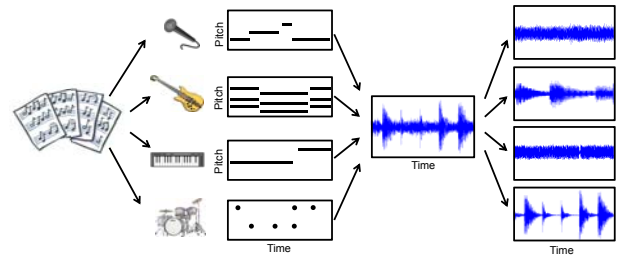


## Singing Voice Extraction



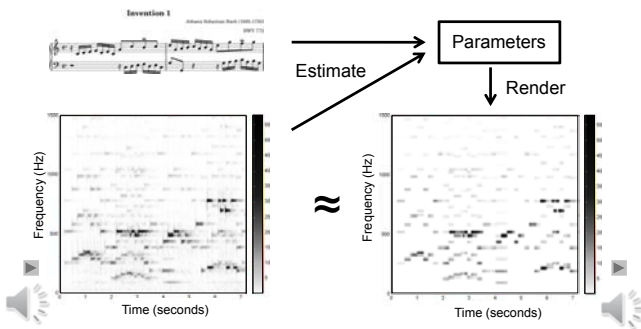
## Score-Informed Source Separation

Exploit musical score to support separation process

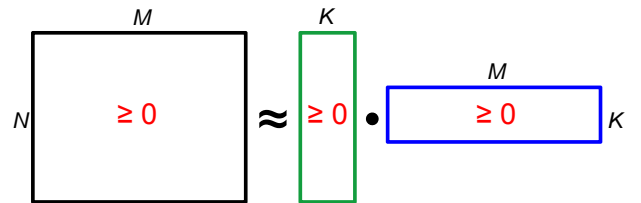


## Parametric Model Approach

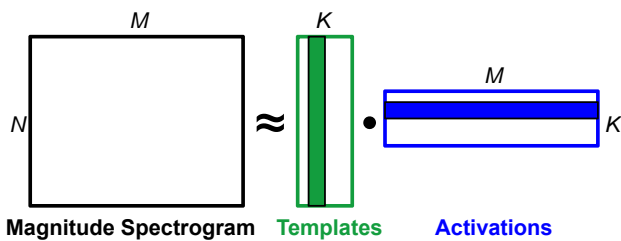
Rebuild spectrogram information



## NMF (Nonnegative Matrix Factorization)



## NMF (Nonnegative Matrix Factorization)



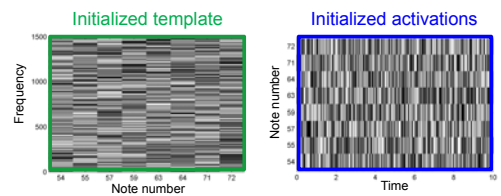
Templates: Pitch + Timbre

“How does it sound”

Activations: Onset time + Duration

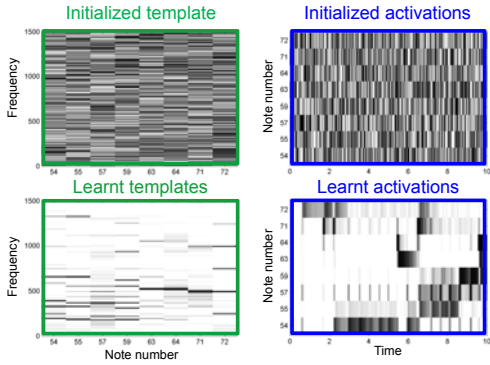
“When does it sound”

## NMF-Decomposition



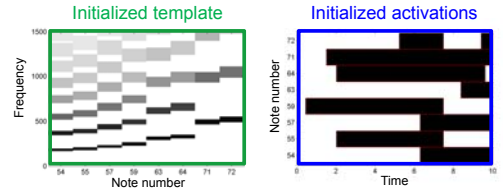
Random initialization

## NMF-Decomposition



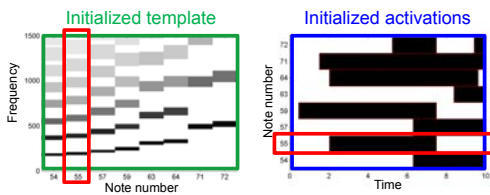
Random initialization → No semantic meaning

## NMF-Decomposition



Constrained initialization

## NMF-Decomposition

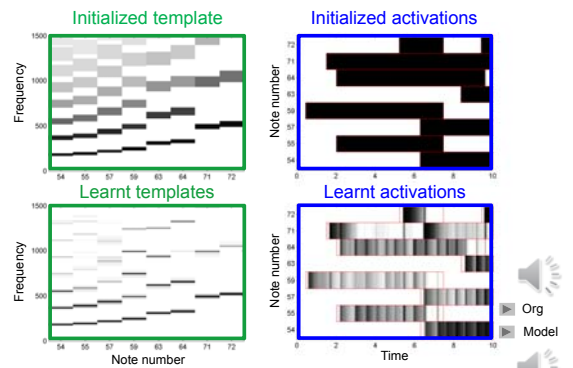


Template constraint for  $p=55$

Activation constraints for  $p=55$

Constrained initialization

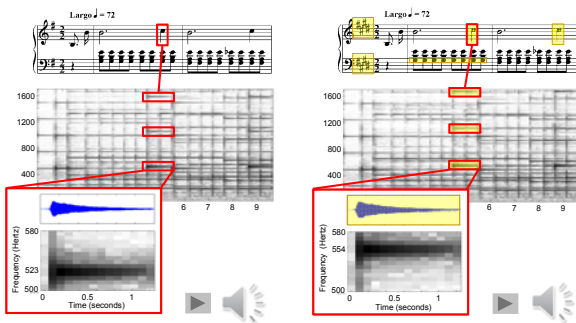
## NMF-Decomposition



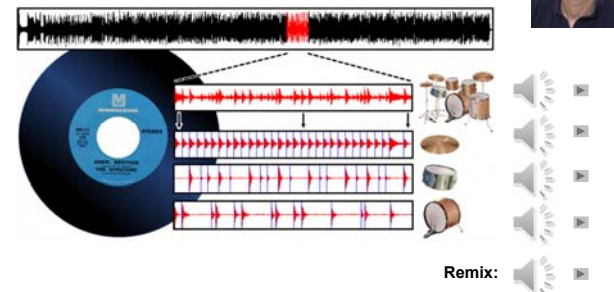
Constrained initialization → NMF as refinement

## Score-Informed Audio Decomposition

Application: Audio editing



## Informed Drum-Sound Decomposition

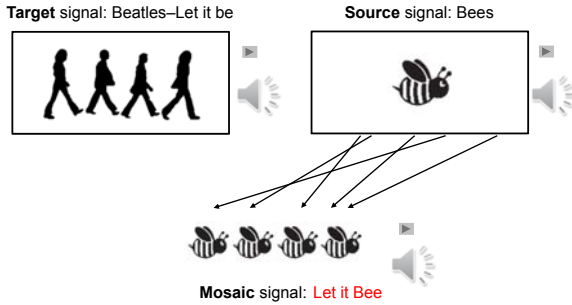


Remix:

Literature: [Dittmar/Müller, IEEE/ACM-TASLP 2016]

Demo: <https://www.audiolabs-erlangen.de/resources/MIR/2016-IEEE-TASLP-DrumSeparation>

## Audio Mosaicing

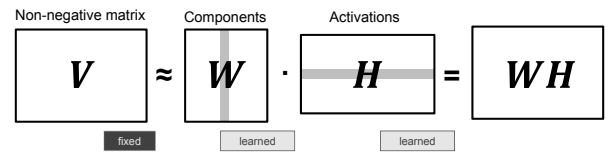


Literature: [Driedger/Müller, ISMIR 2015]

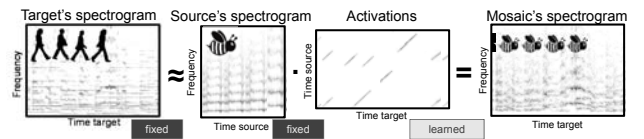
Demo: <https://www.audiolabs-erlangen.de/resources/MIR/2015-ISMIR-LettBee>

## NMF-Inspired Audio Mosaicing

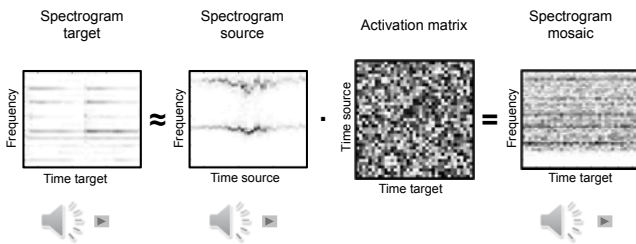
### Non-negative matrix factorization (NMF)



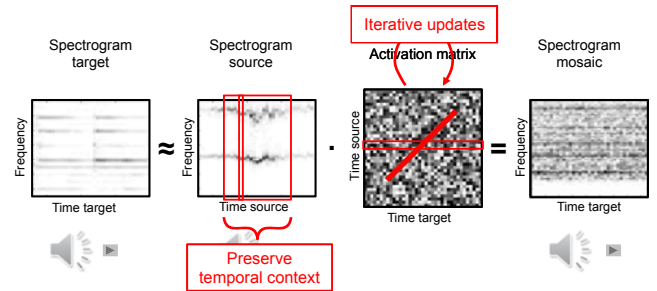
### Proposed audio mosaicing approach



## NMF-Inspired Audio Mosaicing

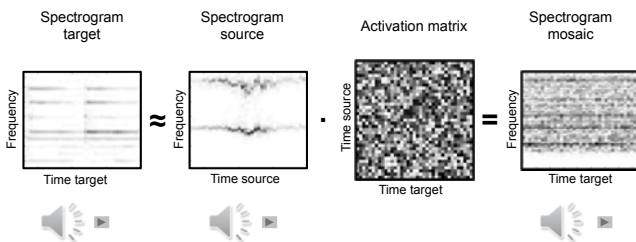


## NMF-Inspired Audio Mosaicing

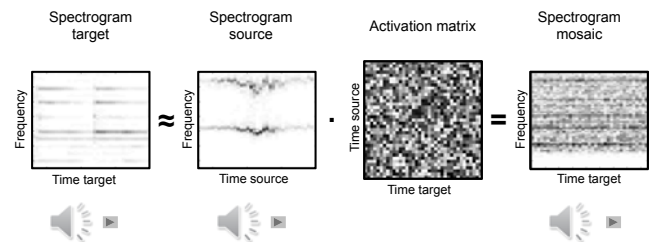


Core idea: support the development of sparse diagonal activation structures

## NMF-Inspired Audio Mosaicing



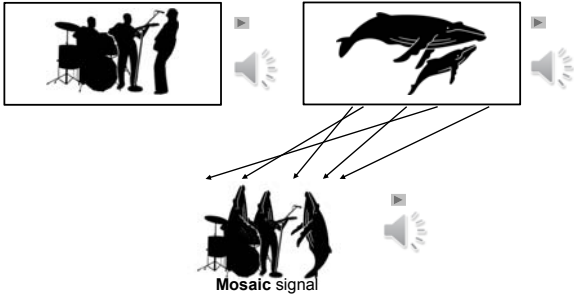
## NMF-Inspired Audio Mosaicing



## Audio Mosaicing

Target signal: Chic–Good times

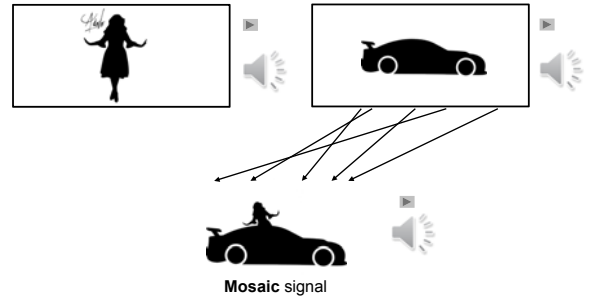
Source signal: Whales



## Audio Mosaicing

Target signal: Adele–Rolling in the Deep

Source signal: Race car



## Links

- SiSEC: Signal Separation Evaluation Campaign  
<https://www.sisec17.audiolabs-erlangen.de/>
- MedleyDB: A Dataset of Multitrack Audio  
<http://steinhardt.nyu.edu/marl/research/medleydb>
- LibROSA (Python)  
<https://librosa.github.io/librosa/>