



**Tutorial T1**  
**Fundamentals of Music Processing:**  
**An Introduction using Python and Jupyter Notebooks**

### Audio Decomposition

**Meinard Müller, Frank Zalkow**

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



### Source Separation

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

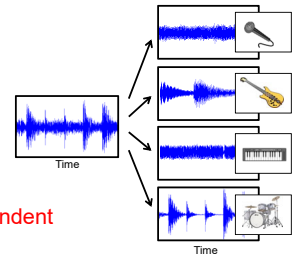


### Source Separation

- 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)

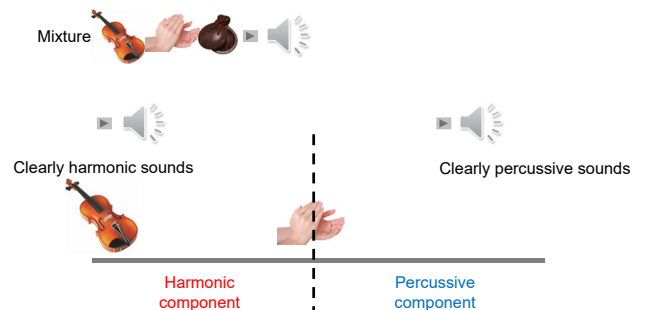
- Main melody, accompaniment, drum track
- Instrumental voices
- Individual note events
- **Only mono or stereo**
- **Sources are often highly dependent**



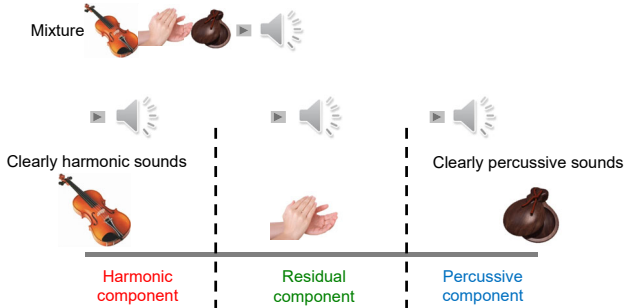
### Harmonic-Percussive Decomposition



### Harmonic-Percussive Decomposition



## Harmonic-Percussive Decomposition

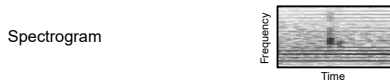


[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

## Harmonic-Percussive Decomposition

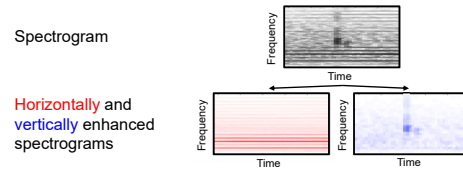


## Harmonic-Percussive Decomposition



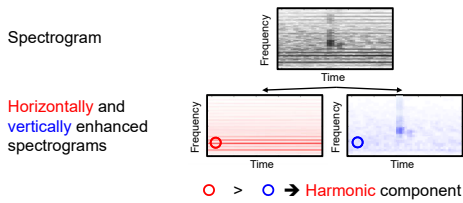
[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

## Harmonic-Percussive Decomposition



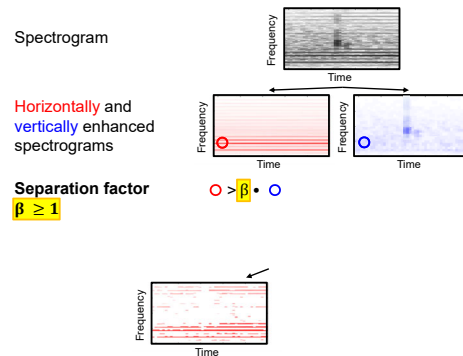
[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

## Harmonic-Percussive Decomposition



[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

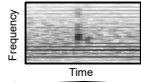
## Harmonic-Percussive Decomposition



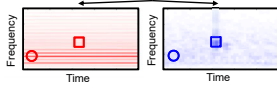
[Driedger et al. ISMIR 2014]

## Harmonic-Percussive Decomposition

Spectrogram

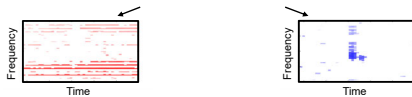


Horizontally and vertically enhanced spectrograms



Separation factor  $\beta \geq 1$

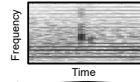
$\circ > \beta \cdot \circ \rightarrow$  Harmonic component  
 $\beta \cdot \square < \square \rightarrow$  Percussive component



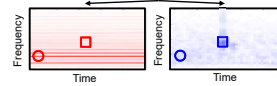
[Driedger et al. ISMIR 2014]

## Harmonic-Percussive Decomposition

Spectrogram



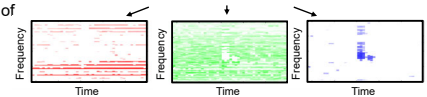
Horizontally and vertically enhanced spectrograms



Separation factor  $\beta \geq 1$

$\circ > \beta \cdot \circ \rightarrow$  Harmonic component  
 $\beta \cdot \square < \square \rightarrow$  Percussive component  
 otherwise  $\rightarrow$  Residual component

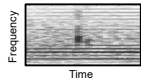
Spectrograms of the harmonic, residual and percussive components



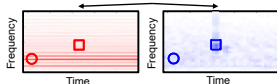
[Driedger et al. ISMIR 2014]

## Harmonic-Percussive Decomposition

Spectrogram



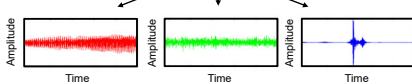
Horizontally and vertically enhanced spectrograms



Separation factor  $\beta \geq 1$

$\circ > \beta \cdot \circ \rightarrow$  Harmonic component  
 $\beta \cdot \square < \square \rightarrow$  Percussive component  
 otherwise  $\rightarrow$  Residual component

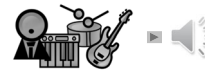
Harmonic, residual and percussive components



[Driedger et al. ISMIR 2014]

## Harmonic-Percussive Decomposition

Mixture:



Clearly harmonic sounds of singing voice and accompaniment

Harmonic component

Noise-like sounds  
 Vibrato/glissando sounds

Residual component

Drum hits  
 Fricatives & plosives in singing voice

Percussive component

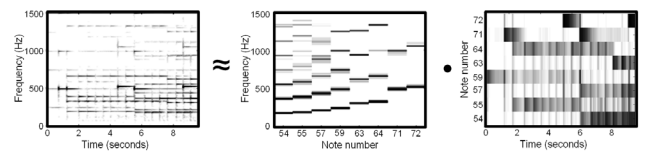
Demo: <https://www.audiolabs-erlangen.de/resources/2014-ISMIR-ExTHPSep/>

[Driedger et al. ISMIR 2014]

## Score-Informed Audio Decomposition

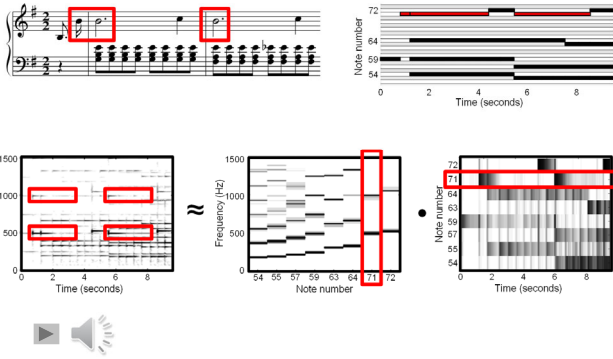
## Score-Informed Audio Decomposition

Exploit musical score to support separation process

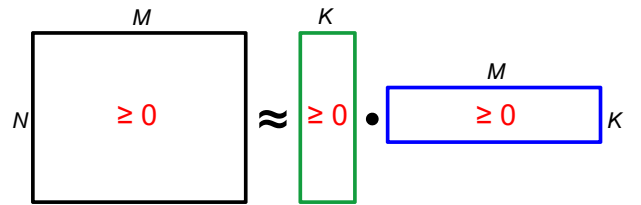


## Score-Informed Audio Decomposition

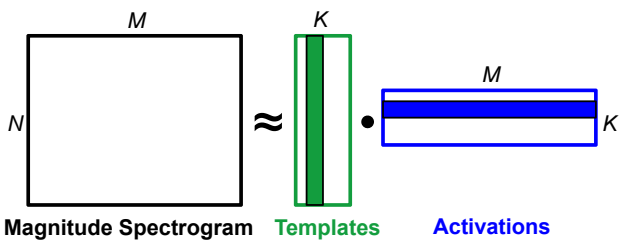
Exploit musical score to support separation process



## NMF (Nonnegative Matrix Factorization)



## NMF (Nonnegative Matrix Factorization)



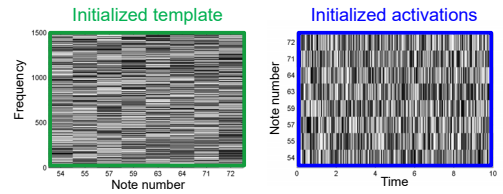
**Templates:** Pitch + Timbre

"How does it sound?"

**Activations:** Onset time + Duration

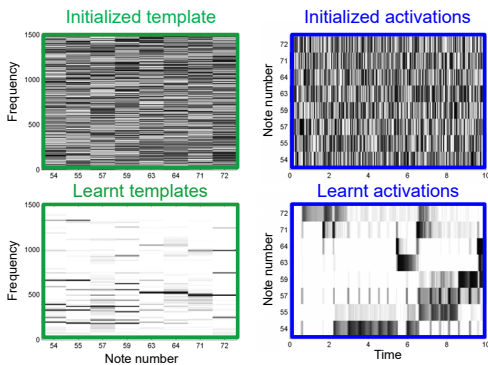
"When does it sound?"

## NMF-Decomposition



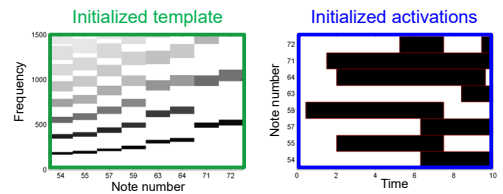
Random initialization

## NMF-Decomposition



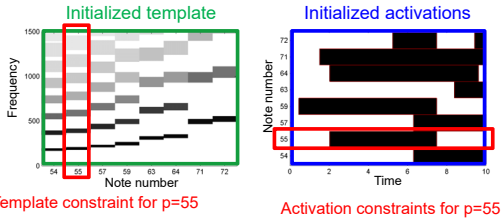
Random initialization → No semantic meaning

## NMF-Decomposition



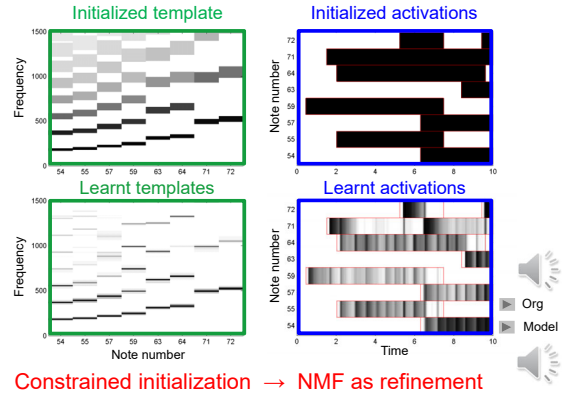
Constrained initialization

## NMF-Decomposition



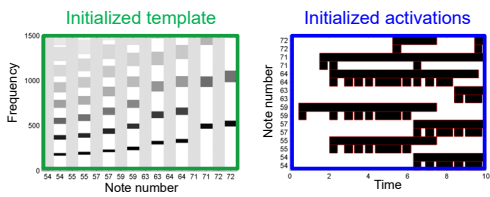
Constrained initialization

## NMF-Decomposition



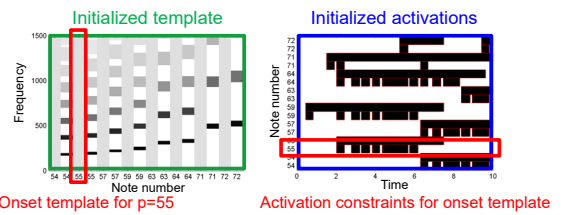
Constrained initialization → NMF as refinement

## NMF-Decomposition



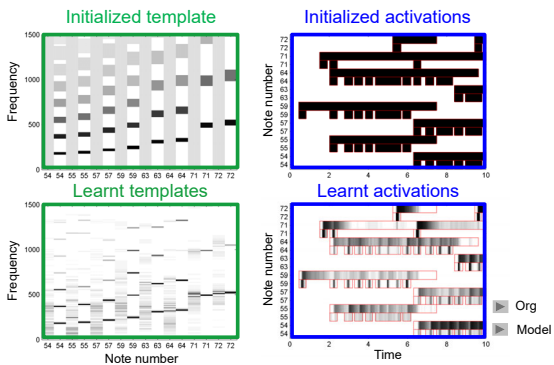
Additional onset models → NMF as refinement

## NMF-Decomposition



Additional onset models → NMF as refinement

## NMF-Decomposition



Additional onset models → NMF as refinement

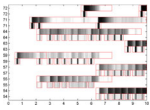
## Score-Informed Audio Decomposition



## Score-Informed Audio Decomposition



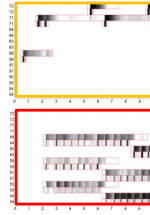
1. Split activation matrix



## Score-Informed Audio Decomposition



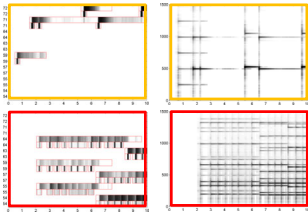
1. Split activation matrix



## Score-Informed Audio Decomposition



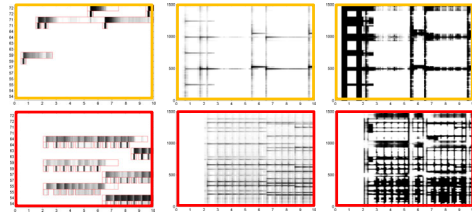
1. Split activation matrix
2. Model spectrogram for left/right



## Score-Informed Audio Decomposition



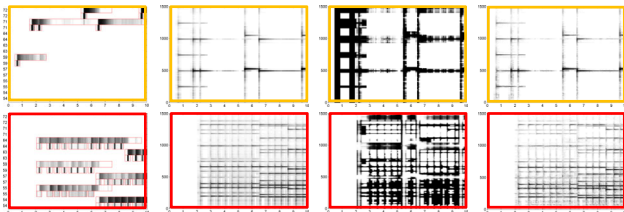
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right



## Score-Informed Audio Decomposition



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

Note-based Audio Events

Audio

Residual Audio

⊕

## Score-Informed Audio Decomposition

Application: Audio editing

Largo  $\lambda = 72$

Frequency (Hz)

Time (seconds)

## 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

Target signal: Beatles—Let it be

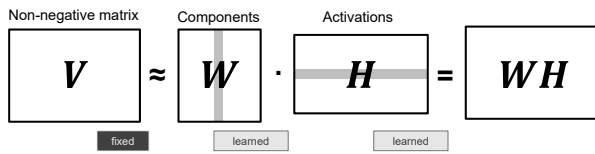
Source signal: Bees

Mosaic signal: Let it Bee

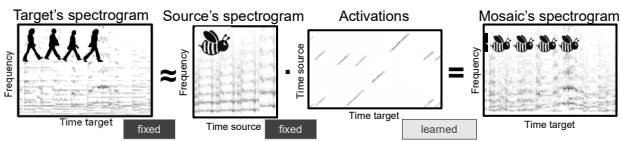
Literature: [Driedger/Müller, ISMIR 2015]  
 Demo: <https://www.audiolabs-erlangen.de/resources/MIR/2015-ISMIR-LettItBee>

## NMF-Inspired Audio Mosaicing

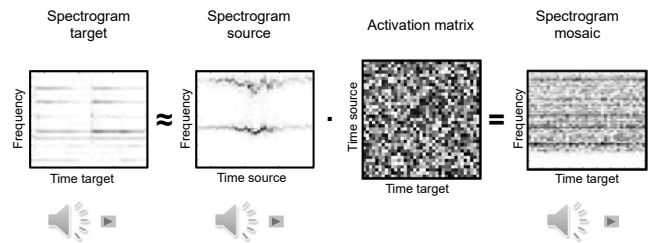
Non-negative matrix factorization (NMF)



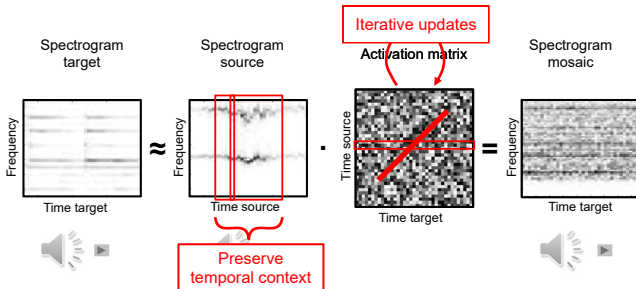
Proposed audio mosaicing approach



## NMF-Inspired Audio Mosaicing

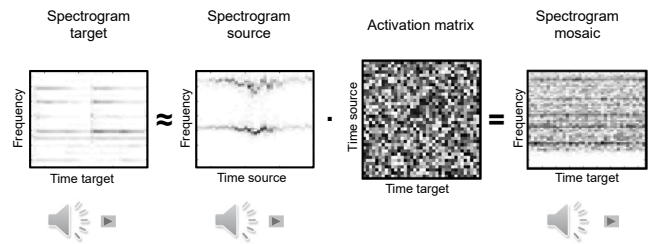


## NMF-Inspired Audio Mosaicing

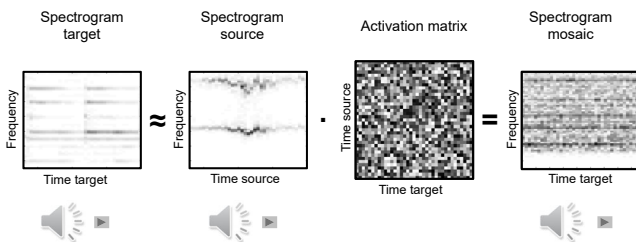


Core idea: support the development of sparse diagonal activation structures

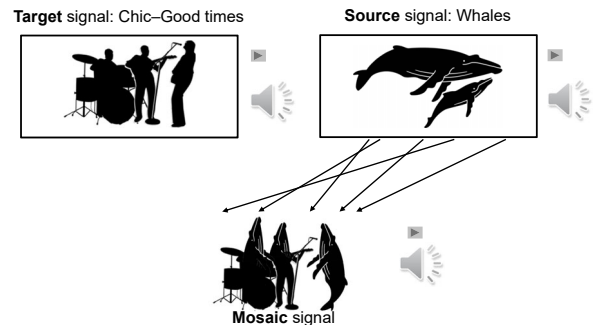
## NMF-Inspired Audio Mosaicing



## NMF-Inspired Audio Mosaicing



## Audio Mosaicing

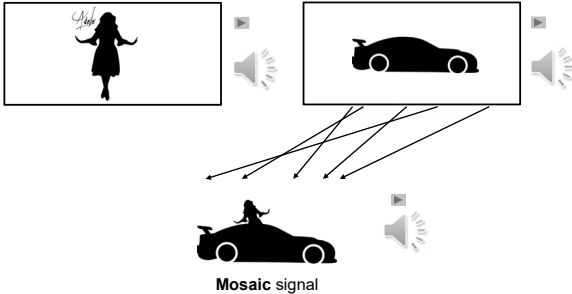




## Audio Mosaicing

Target signal: Adele–Rolling in the Deep

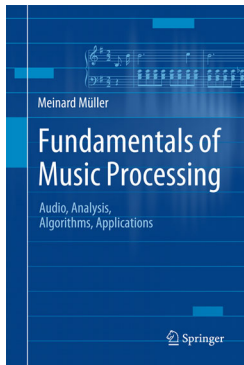
Source signal: Race car



## 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/>

## 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)