INTERNATIONAL AUDIO LABORATORIES ERLANGEN

Musically Informed Audio Decomposition

Meinard Müller and Jonathan Driedger

International Audio Laboratories Erlangen

Meinard Müller

- 2007 Habilitation, Bonn
- 2007 MPI Informatik, Saarbrücken
  Senior Researcher
  Music Processing & Motion Processing
- 2012 W3-Professur, AudioLabs Erlangen
  Semantic Audio Processing

Thanks

- Jonathan Driedger
- Thomas Prätzlich
- Harald Grohganz
- Sebastian Ewert

International Audio Laboratories Erlangen

International Audio Laboratories Erlangen

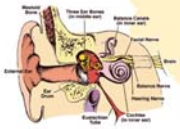
Audio

International Audio Laboratories Erlangen

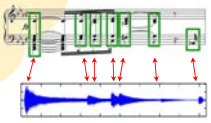
Audio Coding

3D Audio

Audio



Psychoacoustics



Music Processing

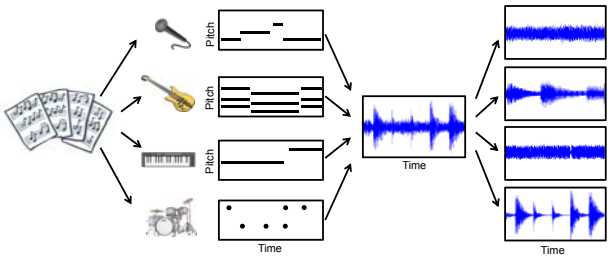
Musically Informed Audio Decomposition

- Extraction of main melody
- Separation of drum track
- Separation of instrumental voices
- Decomposition into individual note events
- Harmonic-percussive separation

Exploit musical knowledge to support decomposition process

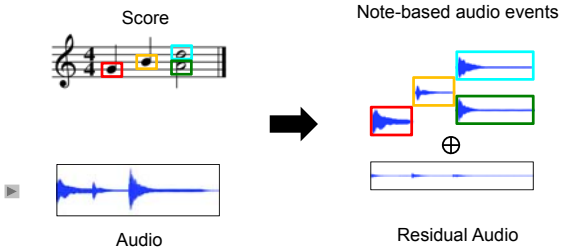
Score-Informed Source Separation

Exploit musical score to support separation process



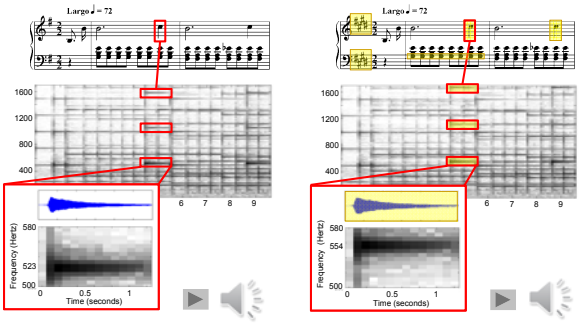
Score-Informed Audio Decomposition

Parameterize audio signal using score's note events



Score-Informed Audio Decomposition

Application: Audio editing



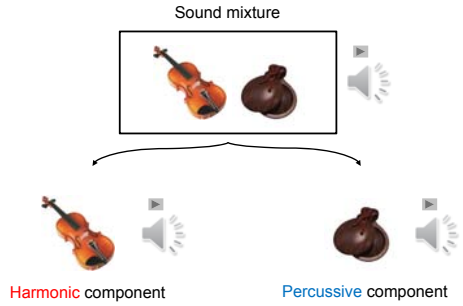
Musically Informed Audio Decomposition

- Harmonic-percussive-residual separation
- NMF-based audio decomposition
- Main melody separation

Musically Informed Audio Decomposition

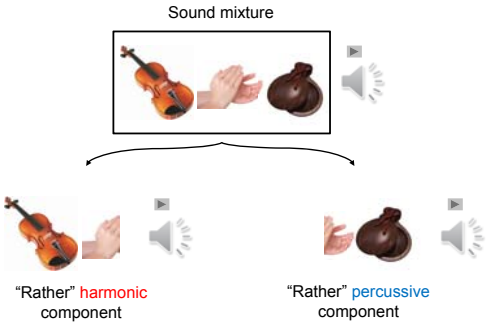
- Harmonic-percussive-residual separation
- NMF-based audio decomposition
- Main melody separation

Harmonic-Percussive Separation



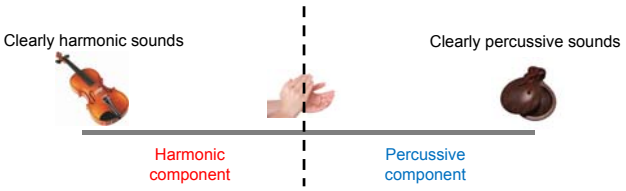
[Ono et al. ISMIR 2008]
[Fitzgerald DAFx 2010]

Harmonic-Percussive Separation



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

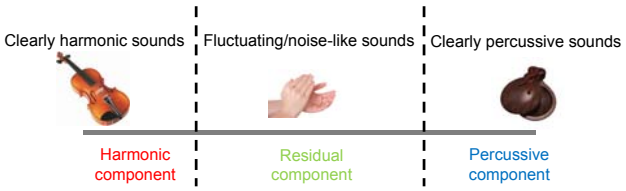
Harmonic-Percussive Separation



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

Harmonic-Percussive-Residual Separation

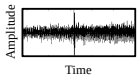
Idea: Introduction of a third component



[Driedger et al. ISMIR 2014]

Harmonic-Percussive-Residual Separation

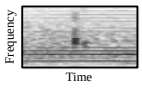
Audio



[Fitzgerald DAFx 2010]

Harmonic-Percussive-Residual Separation

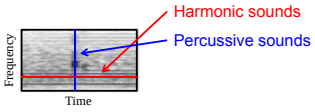
Spectrogram



[Fitzgerald DAFx 2010]

Harmonic-Percussive-Residual Separation

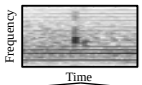
Spectrogram



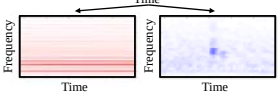
[Fitzgerald DAFx 2010]

Harmonic-Percussive-Residual Separation

Spectrogram



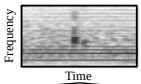
Horizontally and vertically enhanced spectrograms



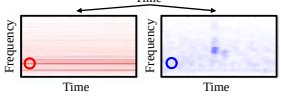
[Fitzgerald DAFx 2010]

Harmonic-Percussive-Residual Separation

Spectrogram

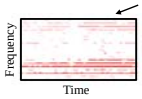


Horizontally and vertically enhanced spectrograms



Separation factor
$\beta \geq 1$

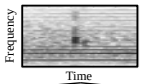
$\circ > \beta \cdot \circ$ ➔ Harmonic component



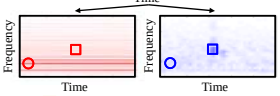
[Fitzgerald DAFx 2010]
[Driedger et al. ISMIR 2014]

Harmonic-Percussive-Residual Separation

Spectrogram



Horizontally and vertically enhanced spectrograms



Separation factor
$\beta \geq 1$

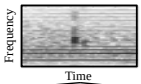
$\circ > \beta \cdot \circ$ ➔ Harmonic component

$\beta \cdot \square < \square$ ➔ Percussive component

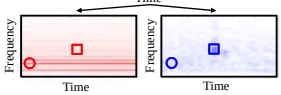
[Fitzgerald DAFx 2010]
[Driedger et al. ISMIR 2014]

Harmonic-Percussive-Residual Separation

Spectrogram



Horizontally and vertically enhanced spectrograms



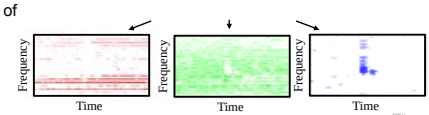
Separation factor
$\beta \geq 1$

$\circ > \beta \cdot \circ$ ➔ Harmonic component

$\beta \cdot \square < \square$ ➔ Percussive component

otherwise ➔ Residual component

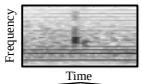
Spectrograms of the harmonic, residual, and percussive components



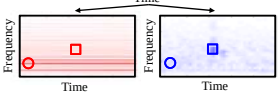
[Fitzgerald DAFx 2010]
[Driedger et al. ISMIR 2014]

Harmonic-Percussive-Residual Separation

Spectrogram



Horizontally and vertically enhanced spectrograms



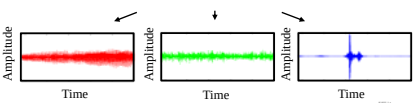
Separation factor
$\beta \geq 1$

$\circ > \beta \cdot \circ$ ➔ Harmonic component

$\beta \cdot \square < \square$ ➔ Percussive component

otherwise ➔ Residual component

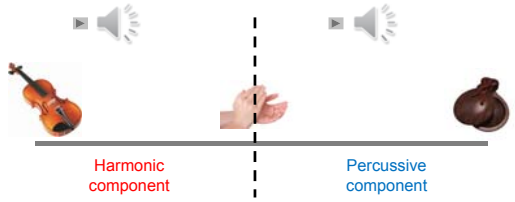
Harmonic, residual and percussive components



[Fitzgerald DAFx 2010]
[Driedger et al. ISMIR 2014]

Harmonic-Percussive-Residual Separation

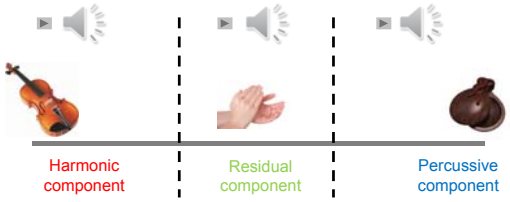
Mixture:



$\beta = 1$

Harmonic-Percussive-Residual Separation

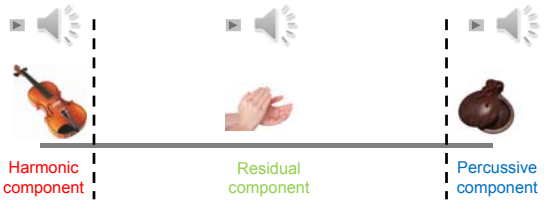
Mixture:



$\beta = 2$

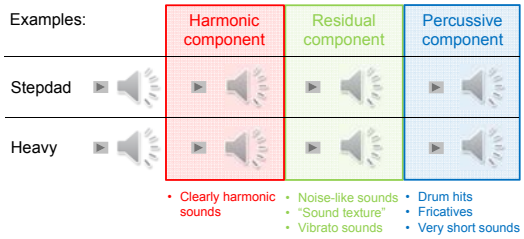
Harmonic-Percussive-Residual Separation

Mixture:



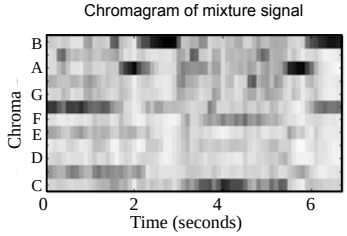
$\beta = 5$

Harmonic-Percussive-Residual Separation



Additional audio material:
http://www.audiolabs-erlangen.de/resources/2014-ISMIR-ExtHPSep/
[Driedger et al. ISMIR 2014]

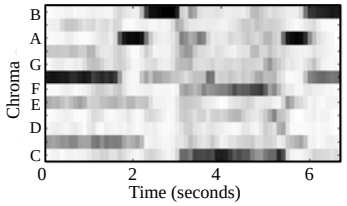
Application: Chroma Enhancement



Application: Chroma Enhancement

Harmonic-percussive decomposition

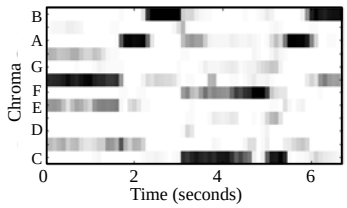
Chromagram of harmonic component



Application: Chroma Enhancement

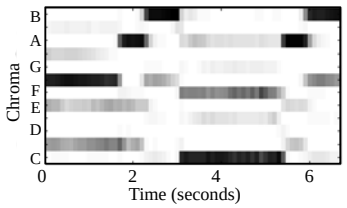
Harmonic-percussive-residual decomposition

Chromagram of harmonic component ($\beta = 2$)



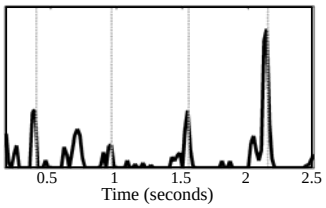
Application: Chroma Enhancement

Chromagram of voline source



Application: Novelty Enhancement

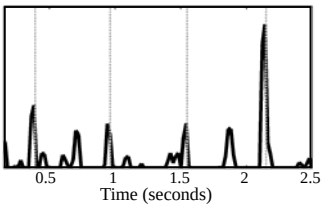
Novelty curve of mixture signal



Application: Novelty Enhancement

Harmonic-percussive decomposition

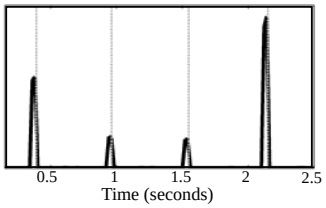
Novelty curve of percussive component



Application: Novelty Enhancement

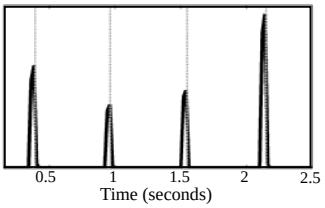
Harmonic-percussive-residual decomposition

Novelty curve of percussive component ($\beta = 3$)

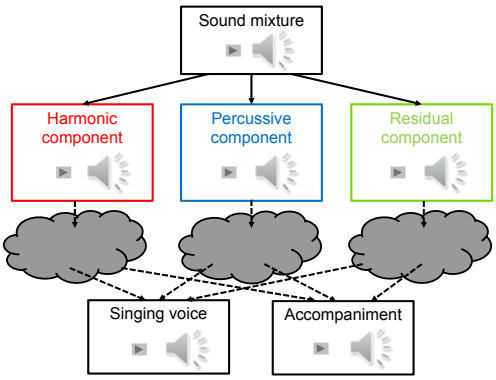


Application: Novelty Enhancement

Novelt curve of isolierten castanet source



Application: Singing Voice Separation



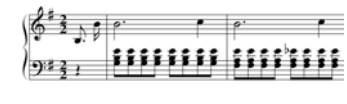
[Driedger & Müller ICASSP 2015]

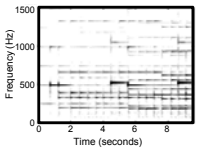
Musically Informed Audio Decomposition

- Harmonic-percussive-residual separation
- NMF-based audio decomposition
- Main melody separation

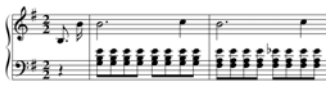
NMF-based Audio Decomposition

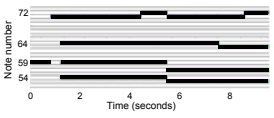
NMF-based Audio Decomposition

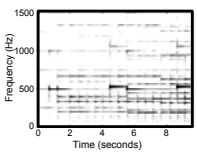




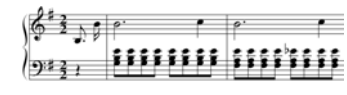
NMF-based Audio Decomposition

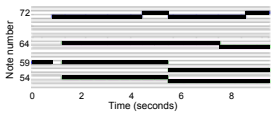


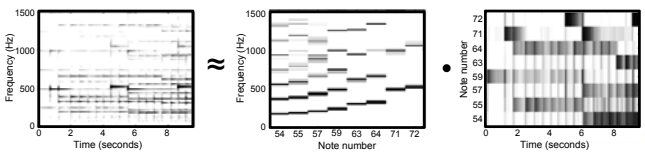




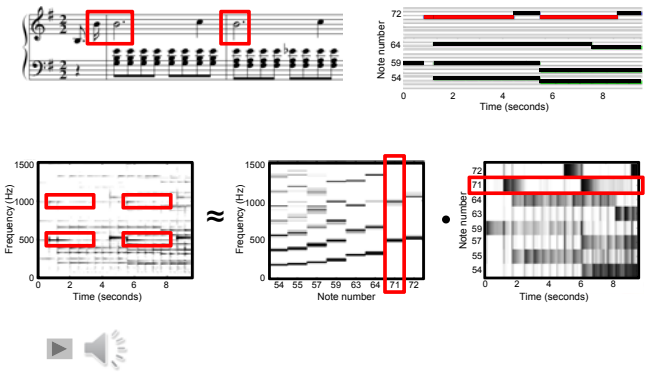
NMF-based Audio Decomposition







NMF-based Audio Decomposition



NMF-Decomposition

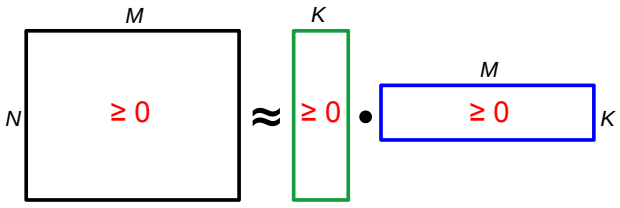
Idea:
- Factorization of spectrogram

Advantages:
- Flexible and easy to implement
- Efficient

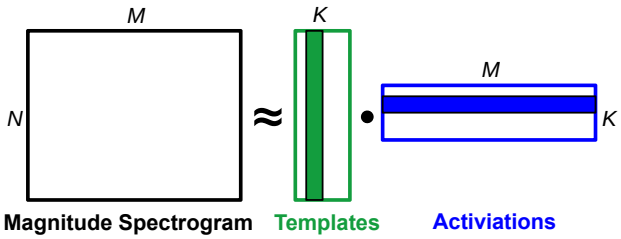
Problems:
- Decompostion difficult to control
- Often no semantic meaning

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

NMF (Nonnegative Matrix Factorization)



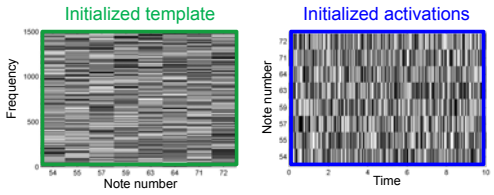
NMF (Nonnegative Matrix Factorization)



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

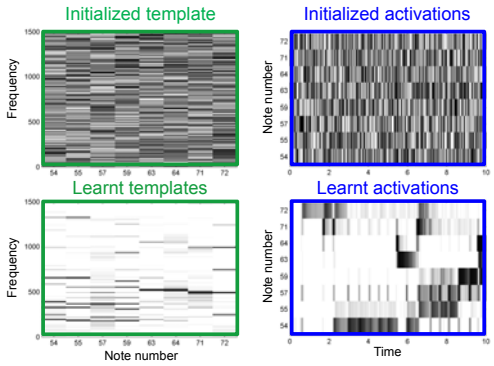
Activiations: **Onset time + Duration** **"When does it sound"**

NMF-Decomposition



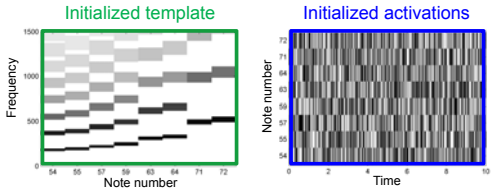
Random initialization

NMF-Decomposition



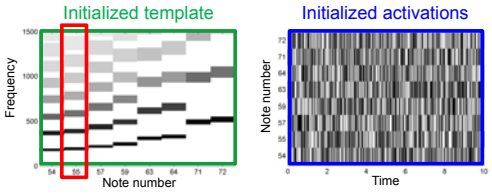
Random initialization → No semantic meaning

NMF-Decomposition



Template constraints

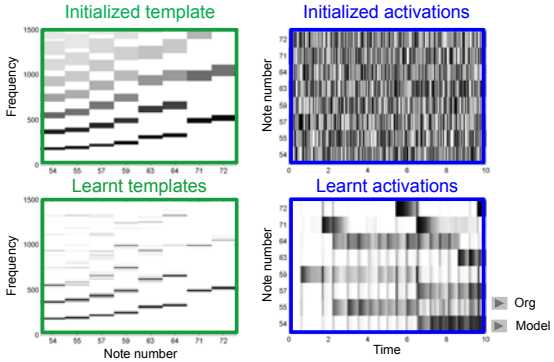
NMF-Decomposition



Template constraint for p=55

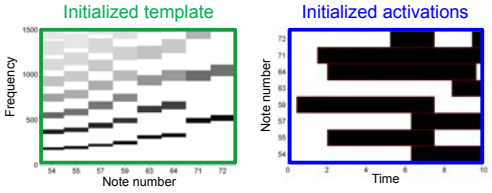
Template constraints

NMF-Decomposition



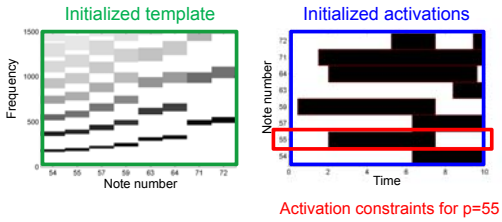
Template constraints → Semantic decomposition

NMF-Decomposition



Activation constraints

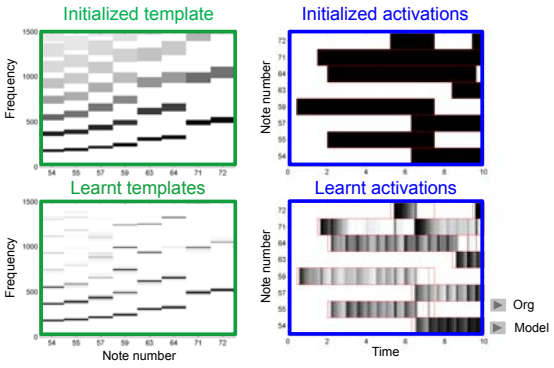
NMF-Decomposition



Activation constraints for p=55

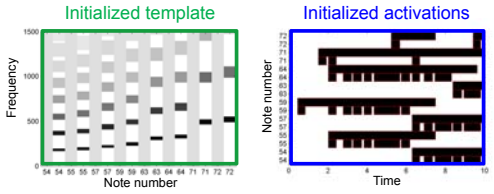
Activation constraints

NMF-Decomposition



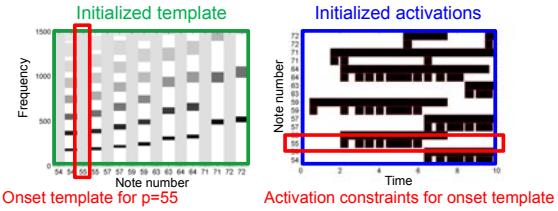
Activation constraints → NMF as refinement

NMF-Decomposition



Additional onset models → NMF as refinement

NMF-Decomposition

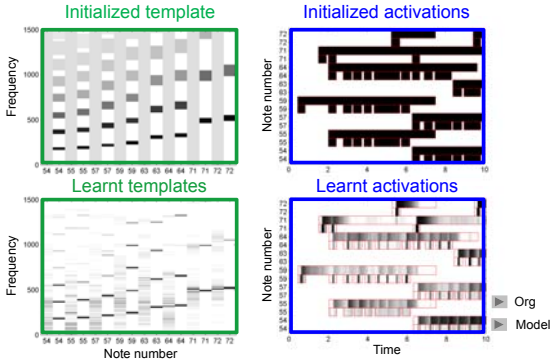


Onset template for p=55

Activation constraints for onset template

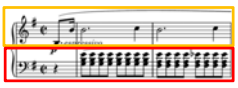
Additional onset models → NMF as refinement

NMF-Decomposition

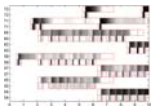


Additional onset models → NMF as refinement

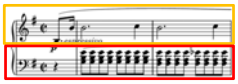
Score-Informed Source Separation



1. Split activation matrix

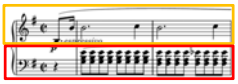


Score-Informed Source Separation

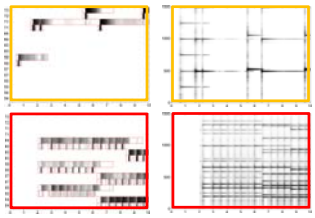


1. Split activation matrix

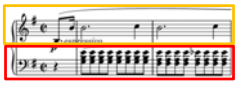
Score-Informed Source Separation



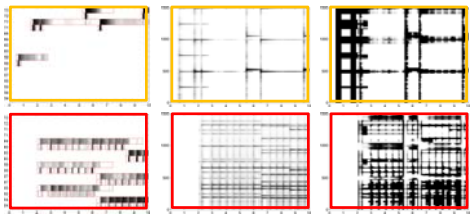
1. Split activation matrix
2. Model spectrogram for left/right



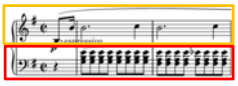
Score-Informed Source Separation



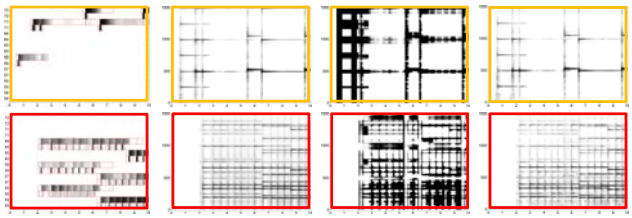
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right



Score-Informed Source Separation



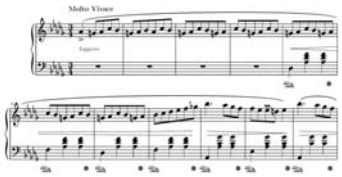
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right
4. Estimated spectrograms for left/right



Score-Informed Audio Decomposition

Application: Separating left and right hands for piano

Chopin, Waltz Op. 64, No. 1

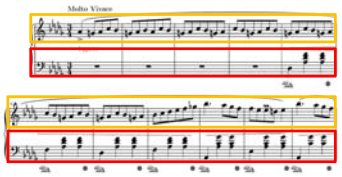


Original

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

Musically Informed Audio Decomposition

- Harmonic-percussive-residual separation
- NMF-based audio decomposition
- Main melody separation

Audio Decomposition

Works reasonable

Audio Decomposition

Much more difficult

Audio Decomposition

Related problems:

- F0 estimation
- Melody tracking
- Human voice
- Vibrato

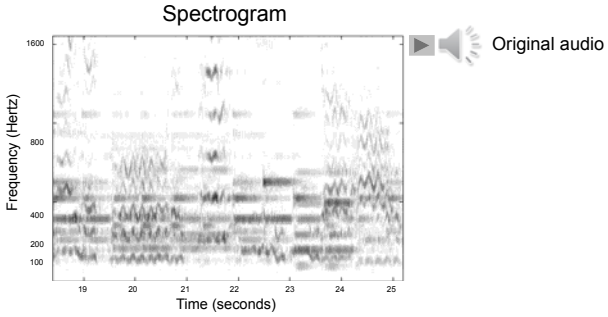
Much more difficult

Melody Tracking

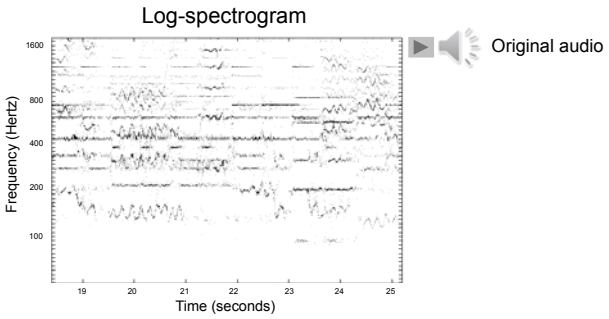
Justin Salamon and Emilia Gómez:
Melody extraction from polyphonic music signals using pitch contour characteristics.
IEEE-TASLP 2012

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

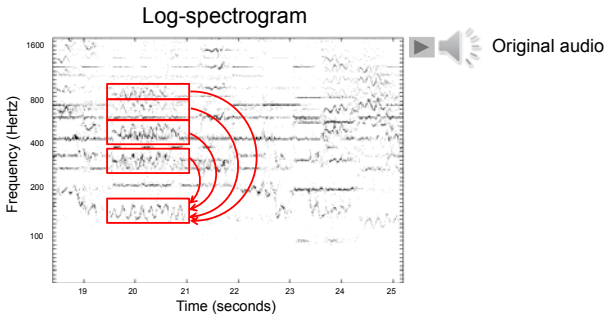
F0 Estimation



F0 Estimation

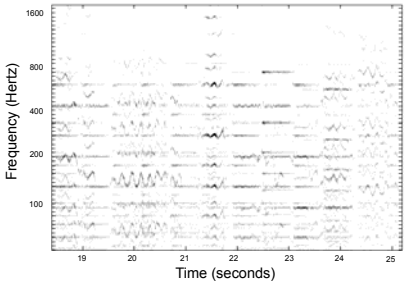


F0 Estimation



F0 Estimation

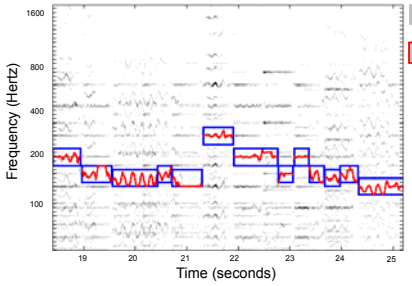
Salience log-spectrogram



Original audio

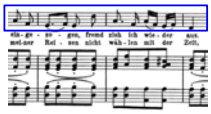
F0 Estimation

Score-informed F0



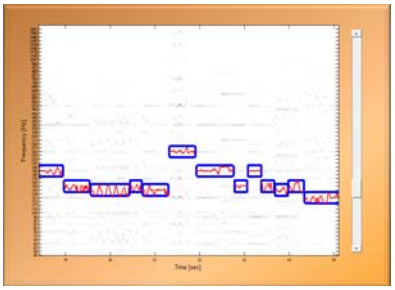
Original audio

F0 estimation



Score-Informed Source Separation

Application: Voice separation and editing

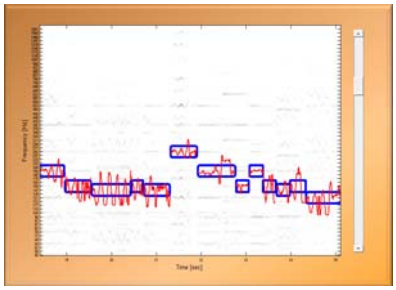


Original audio

Separated voice

Score-Informed Source Separation

Application: Voice separation and editing



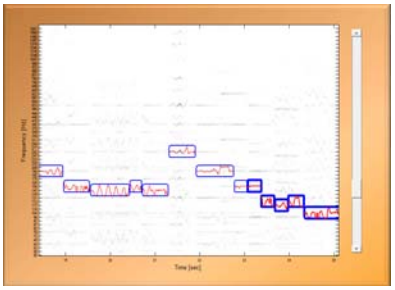
Original audio

Separated voice

Amplified vibrato

Score-Informed Source Separation

Application: Voice separation and editing



Original audio

Separated voice

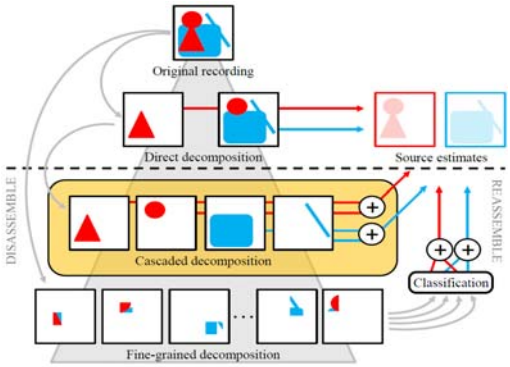
Amplified vibrato

Pitch shift

Conclusions

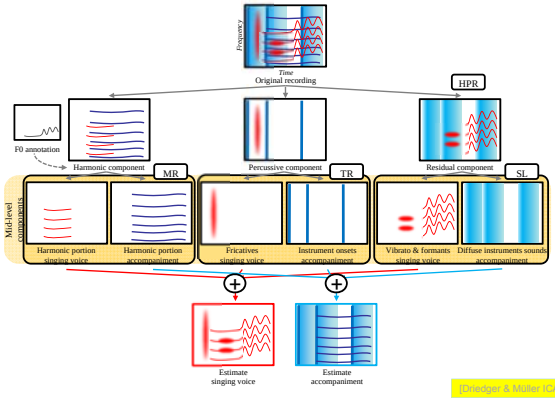
- Musically informed audio decomposition
  - Harmonic-percussive-residual separation
  - NMF-based audio decomposition
  - Main melody separation
- Separation of instrumental sources hard
- Extracting of human voices even harder
- Evaluation?
- Applications?

Cascaded Audio Decomposition



[Driedger & Müller ICASSP 2015]

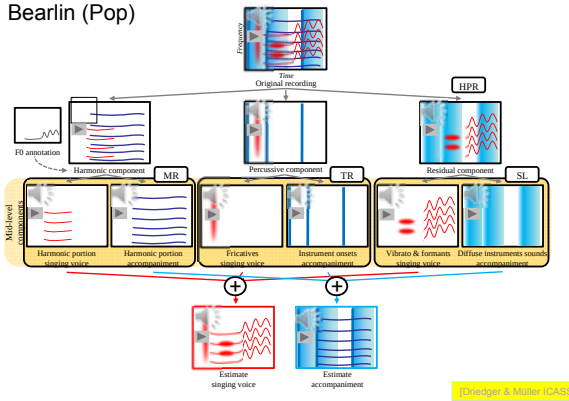
Cascaded Audio Decomposition



[Driedger & Müller ICASSP 2015]

Cascaded Audio Decomposition

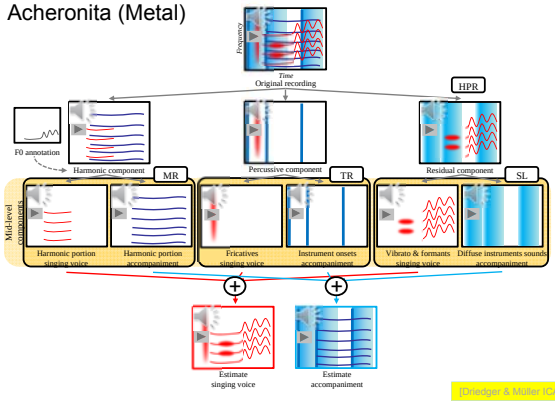
Bearlin (Pop)



[Driedger & Müller ICASSP 2015]

Cascaded Audio Decomposition

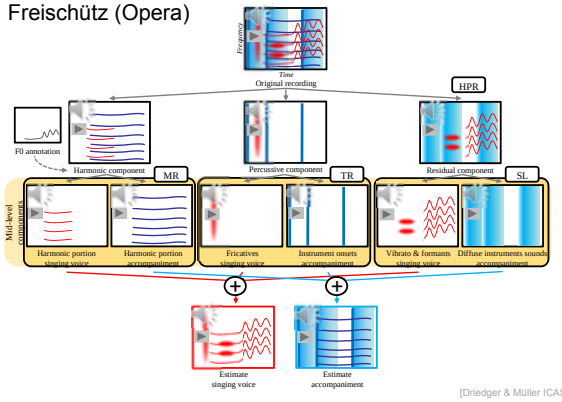
Acheronita (Metal)



[Driedger & Müller ICASSP 2015]

Cascaded Audio Decomposition

Freischütz (Opera)



[Driedger & Müller ICASSP 2015]

Textbook

A First Course on Music Processing

Audio, Analysis, Algorithms, Applications

- Approx. 500 pages
- Approx. 300 figures
- Exercises
- To appear: End of 2015

Springer

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

Acknowledgments

This work has been supported by

- International Audio Laboratories Erlangen, which are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer IIS
- German Research Foundation (DFG MU 2686/6-1)