

## Musically Informed Audio Decomposition

Meinard Müller and Jonathan Driedger

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

Department of Computational Perception  
Johannes Kepler University Linz

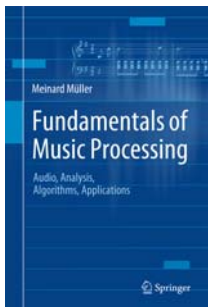
14.12.2016

## 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

## 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

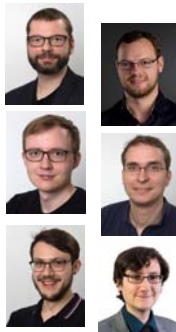
Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

## International Audio Laboratories Erlangen



## Research Group Semantic Audio Signal Processing

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



## Research Group Semantic Audio Signal Processing

- Jonathan Driedger

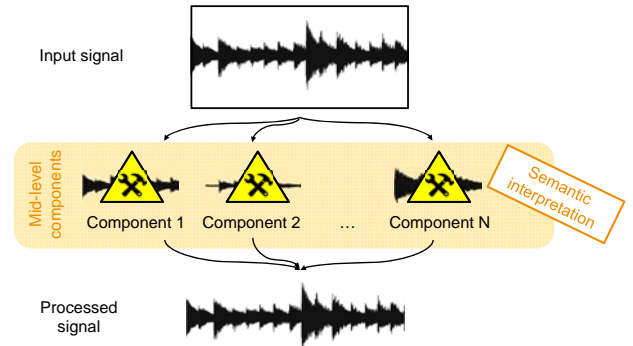


PhD Thesis  
**Processing Music Signals Using Audio Decomposition Techniques**  
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)  
May 2016

## Music Signal Processing

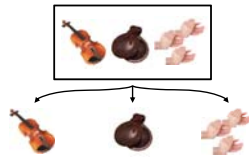


## Why Audio Decomposition?



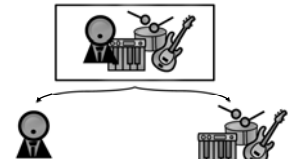
## Structure of this Talk

- Harmonic-percussive-residual decomposition



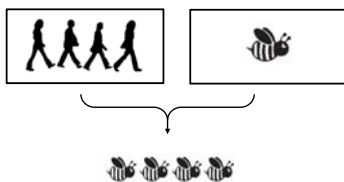
## Structure of this Talk

- Harmonic-percussive-residual decomposition
- Singing voice extraction

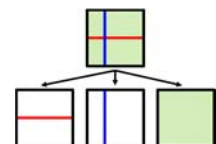


## Structure of this Talk

- Harmonic-percussive-residual decomposition
- Singing voice extraction
- Audio mosaicing

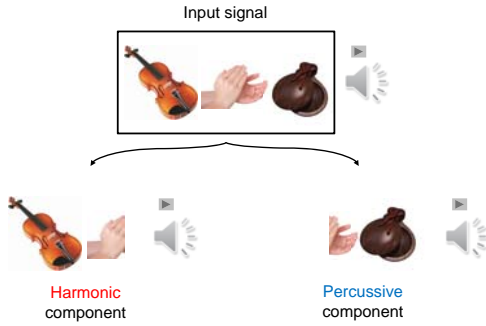


## Harmonic-Percussive-Residual Decomposition



## Harmonic-Percussive Decomposition

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



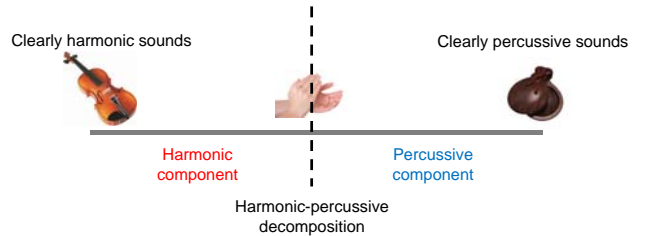
© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Harmonic-Percussive Decomposition

The "harmonic-percussive scale":



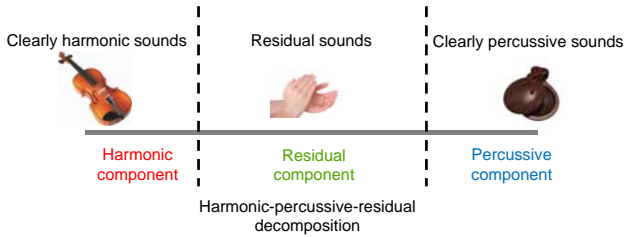
© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

The "harmonic-percussive scale":



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

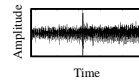
Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

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

Audio



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

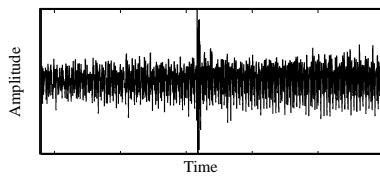
Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

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

Audio



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

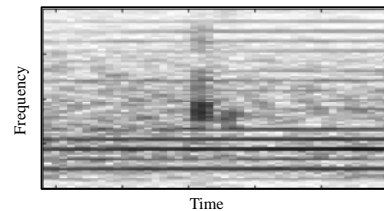
Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

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

Spectrogram



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

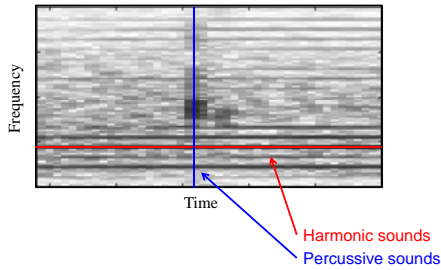
Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

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

Spectrogram



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

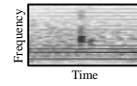
Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

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

Spectrogram



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

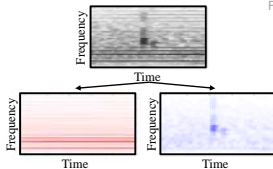


## Harmonic-Percussive-Residual Decomposition

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

Spectrogram

Horizontally and vertically enhanced spectrograms



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

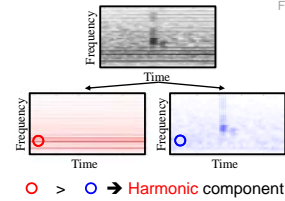


## Harmonic-Percussive-Residual Decomposition

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

Spectrogram

Horizontally and vertically enhanced spectrograms



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



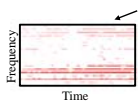
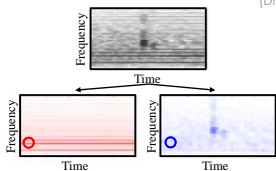
## Harmonic-Percussive-Residual Decomposition

[Driedger et al. ISMIR 2014]

Spectrogram

Horizontally and vertically enhanced spectrograms

Separation factor  $\beta \geq 1$



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



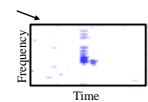
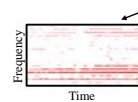
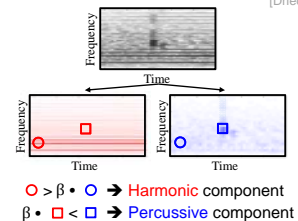
## Harmonic-Percussive-Residual Decomposition

[Driedger et al. ISMIR 2014]

Spectrogram

Horizontally and vertically enhanced spectrograms

Separation factor  $\beta \geq 1$



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

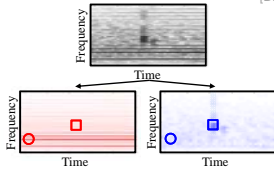


## Harmonic-Percussive-Residual Decomposition

[Driedger et al. ISMIR 2014]

Spectrogram

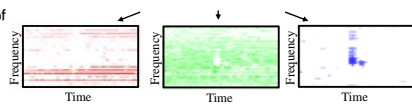
Horizontally and vertically enhanced spectrograms



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



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

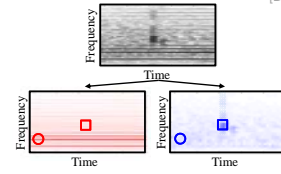


## Harmonic-Percussive-Residual Decomposition

[Driedger et al. ISMIR 2014]

Spectrogram

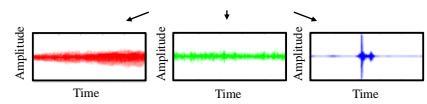
Horizontally and vertically enhanced spectrograms



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



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

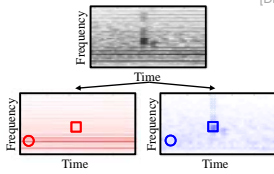


## Harmonic-Percussive-Residual Decomposition

[Driedger et al. ISMIR 2014]

Spectrogram

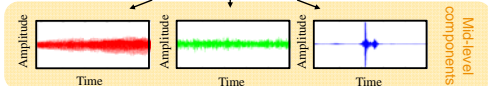
Horizontally and vertically enhanced spectrograms



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



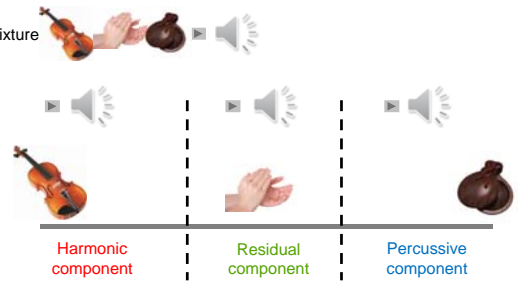
© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

Mixture



$\beta = 2$

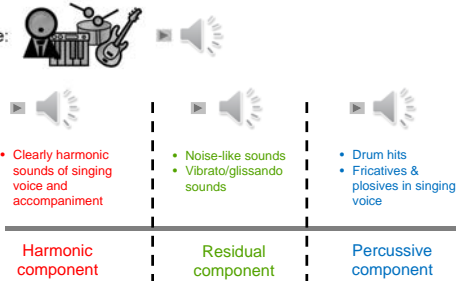
© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Harmonic-Percussive-Residual Decomposition

Mixture:



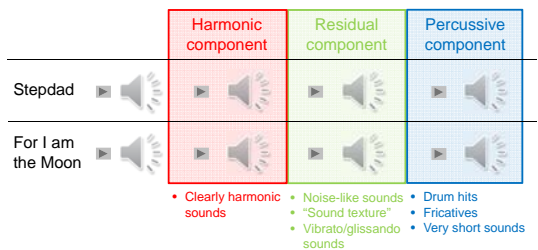
$\beta = 2$

© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Additional Examples

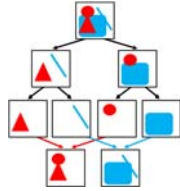


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

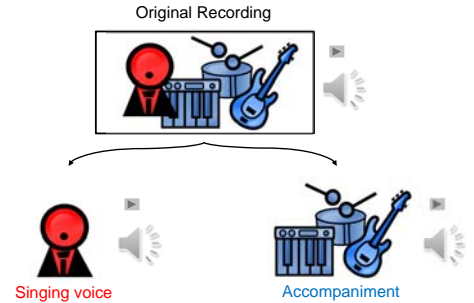
Musically Informed Audio Decomposition



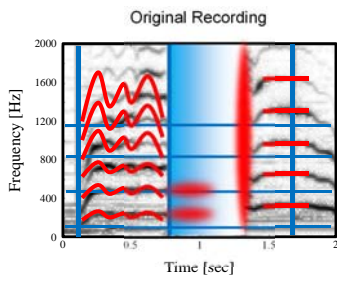
# Singing Voice Extraction



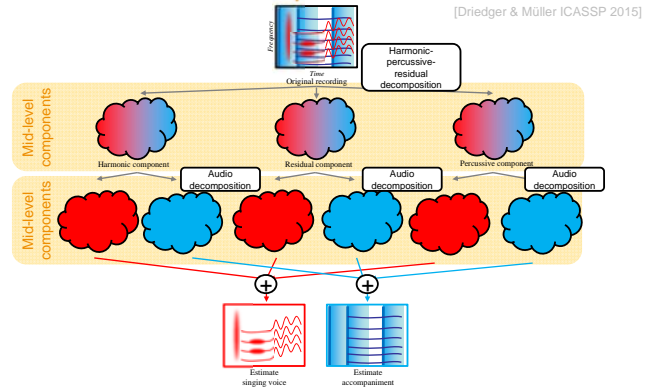
# Singing Voice Extraction



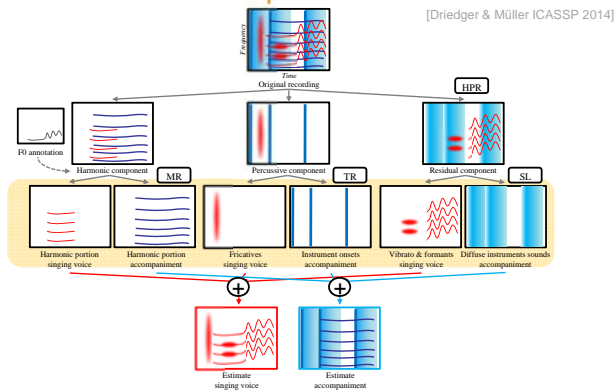
# Singing Voice Extraction



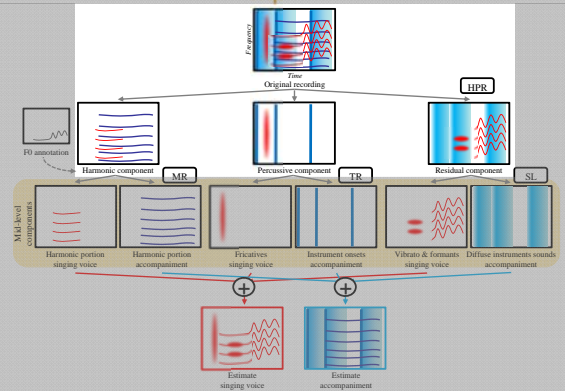
# Cascaded Audio Decomposition



# Cascaded Audio Decomposition

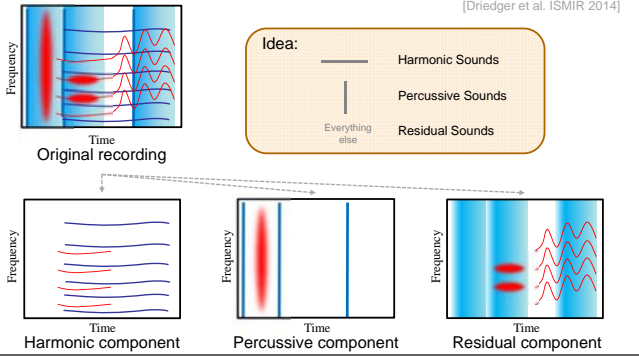


# Cascaded Audio Decomposition



## HPR – Harmonic-Percussive-Residual Decomposition

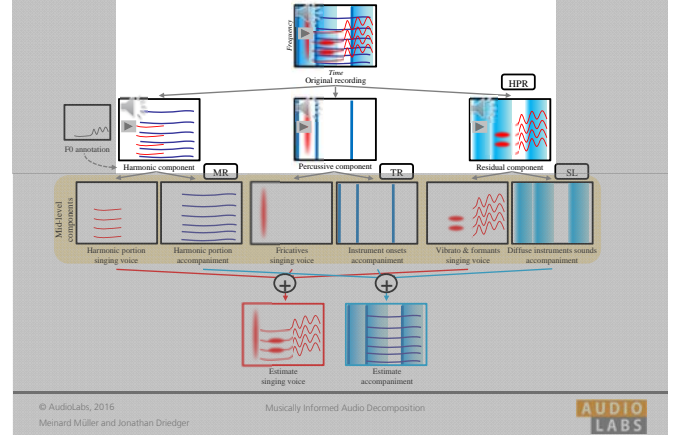
[Tachibana et al. IEEE-TASLP 2013]  
 [Jeong & Lee IEEE-SPL 2014]  
 [Driedger et al. ISMIR 2014]



© AudiLabs, 2016  
 Meinard Müller and Jonathan Driedger

AUDIO LABS

## Cascaded Audio Decomposition

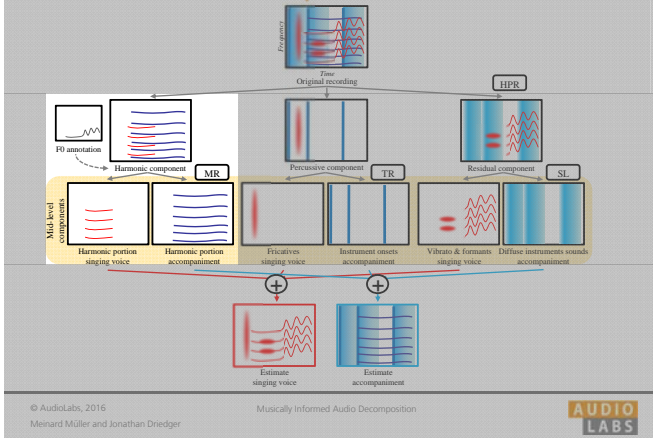


© AudiLabs, 2016  
 Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

AUDIO LABS

## Cascaded Audio Decomposition

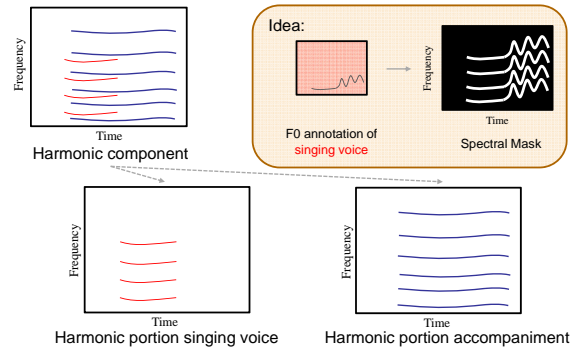


© AudiLabs, 2016  
 Meinard Müller and Jonathan Driedger

AUDIO LABS

## MR – Melody/Residual Decomposition

[Virtanen et al. ISCA 2008]  
 [Salamon & Gómez IEEE-TASLP 2012]

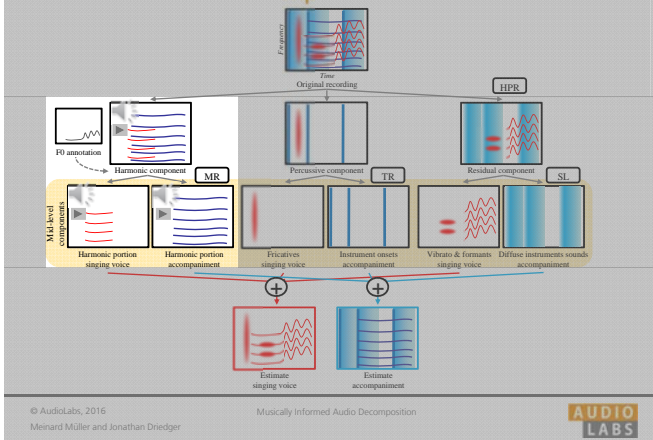


© AudiLabs, 2016  
 Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

AUDIO LABS

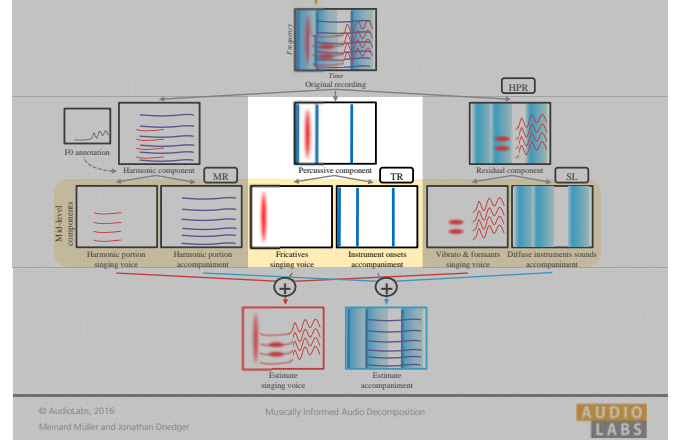
## Cascaded Audio Decomposition



© AudiLabs, 2016  
 Meinard Müller and Jonathan Driedger

AUDIO LABS

## Cascaded Audio Decomposition



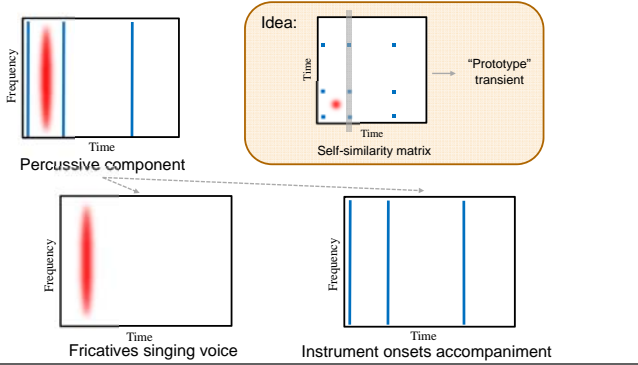
© AudiLabs, 2016  
 Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition

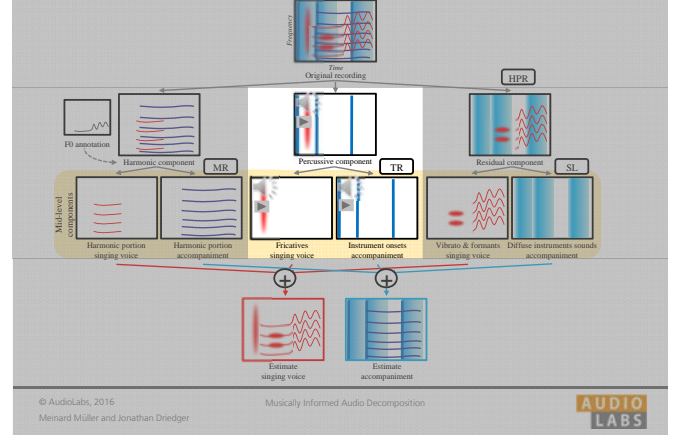
AUDIO LABS

## TR – Transient/Residual Decomposition

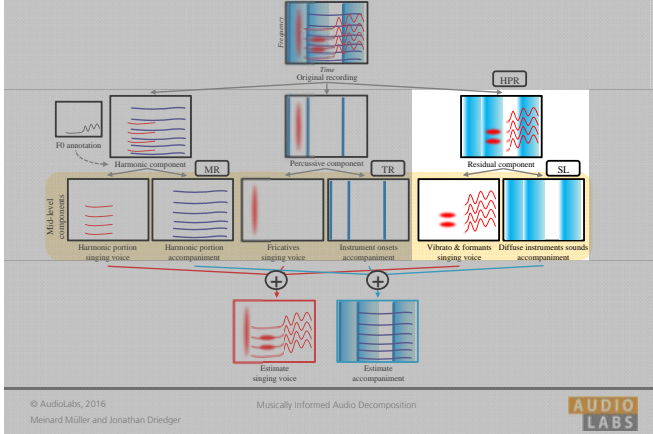
[Talmon et al. IEEE-TASLP 2011]



## Cascaded Audio Decomposition

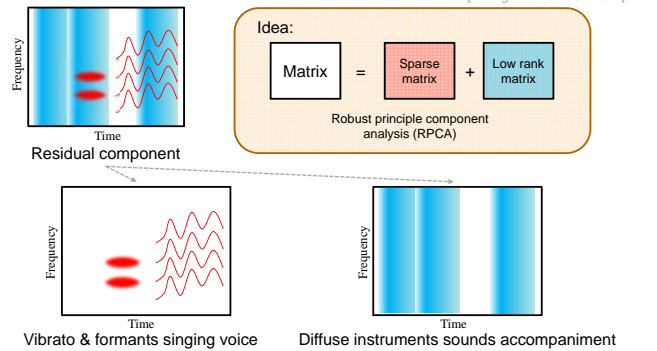


## Cascaded Audio Decomposition

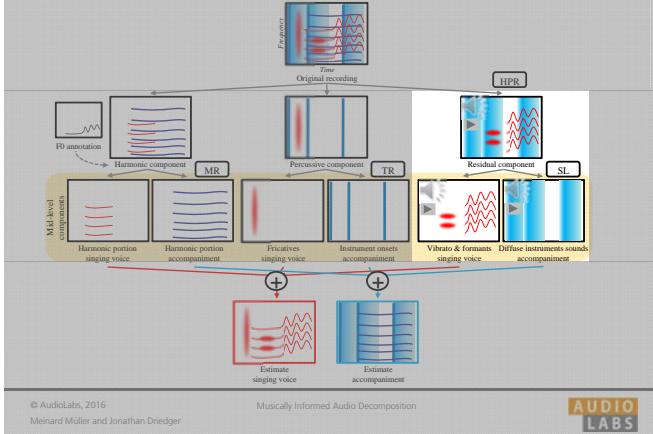


## SL – Sparse/Low Rank Decomposition

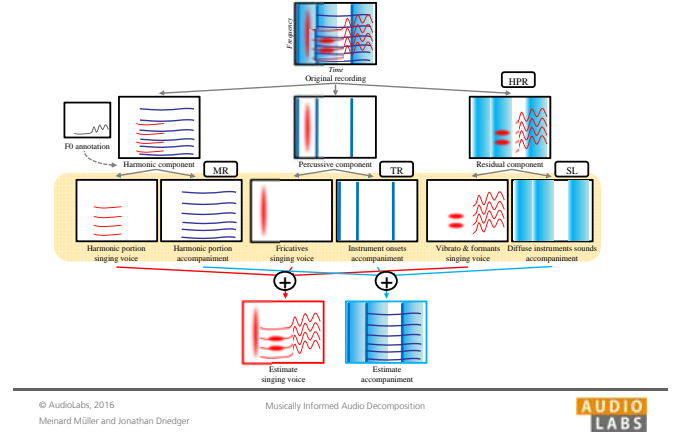
[Candès et al. JACM 2011]  
[Huang et al. ICASSP 2012]



## Cascaded Audio Decomposition

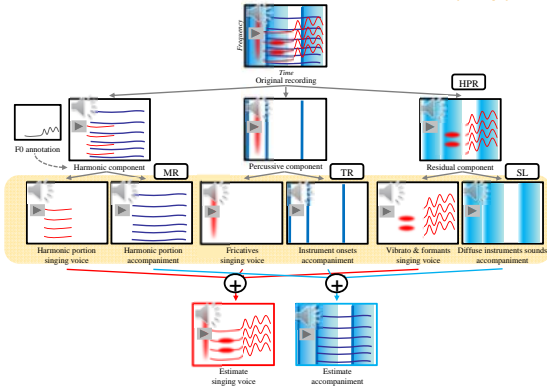


## Cascaded Audio Decomposition





## Cascaded Audio Decomposition: Bearlin (Pop)

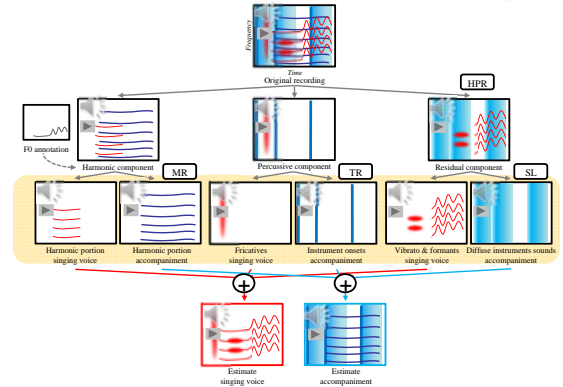


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Cascaded Audio Decomposition: Acheronita (Metal)

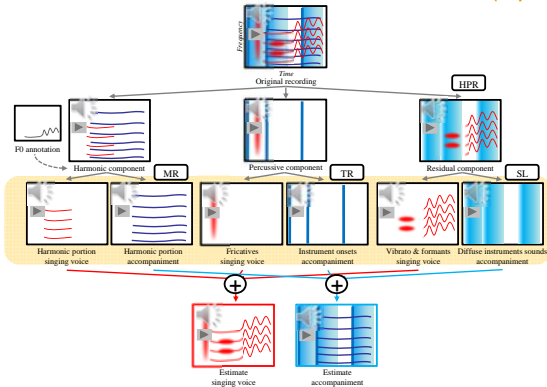


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Cascaded Audio Decomposition: Freischütz (Opera)

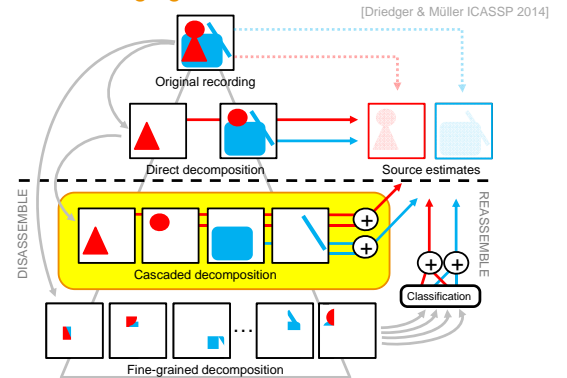


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Approaches to Singing Voice Extraction



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Audio Mosaicing

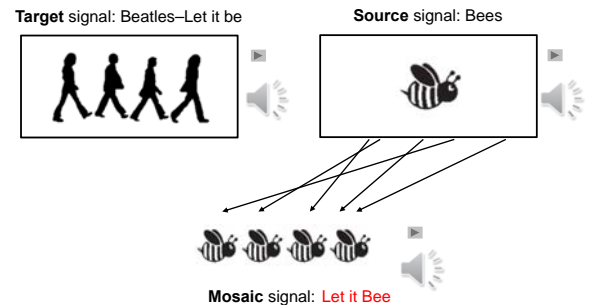


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Audio Mosaicing



© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

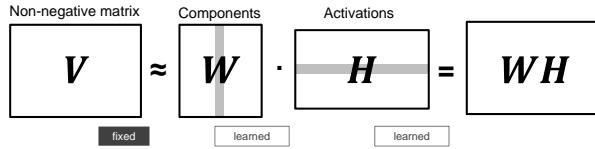
Musically Informed Audio Decomposition



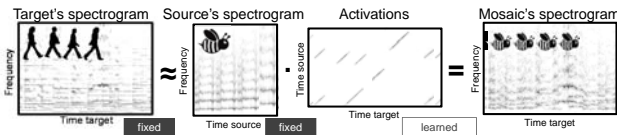
## NMF-Inspired Audio Mosaicing

### Non-negative matrix factorization (NMF)

[Driedger et al. ISMIR 2015]



### Proposed audio mosaicing approach

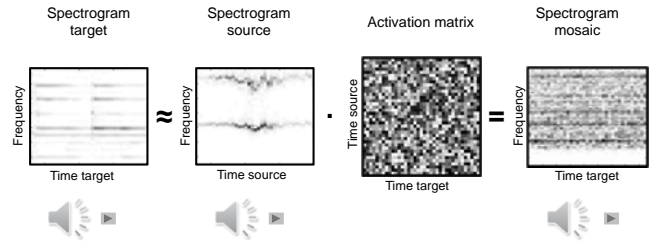


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Basic NMF-Inspired Audio Mosaicing

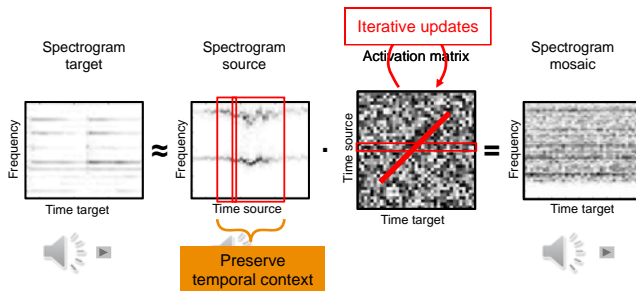


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Basic NMF-Inspired Audio Mosaicing



Core idea: support the development of sparse diagonal activation structures

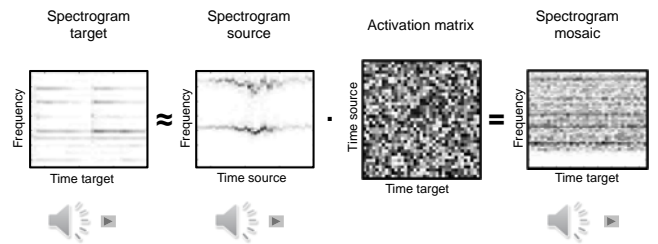
© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Audio Mosaicing with Extended Set of Update Rules

[Driedger et al. ISMIR 2015]



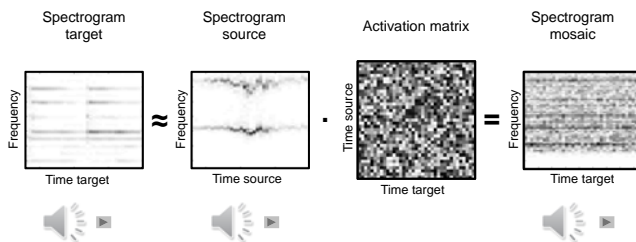
© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## Audio Mosaicing with Extended Set of Update Rules

[Driedger et al. ISMIR 2015]

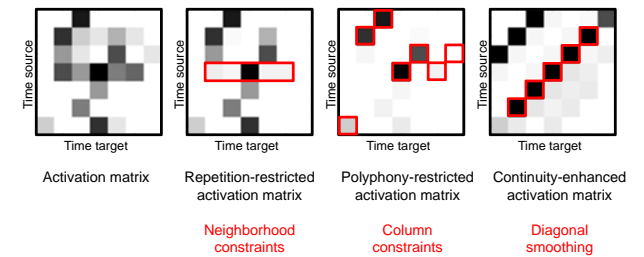


© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

Musically Informed Audio Decomposition



## NMF with Extended Set of Update Rules



Neighborhood constraints

Column constraints

Diagonal smoothing

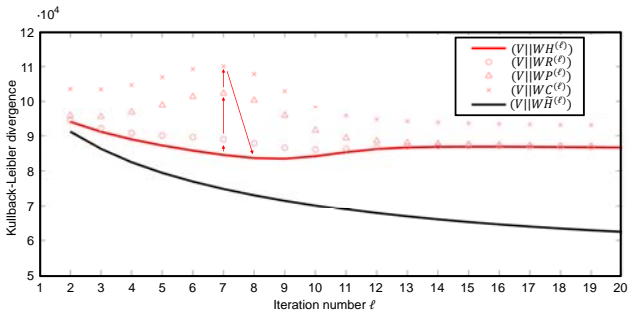
- Constraints are enforced by additional update rules
- Additional rules are interleaved with standard NMF update rules
- Soft alternative to NMF

© AudiLabs, 2016  
Meinard Müller and Jonathan Driedger

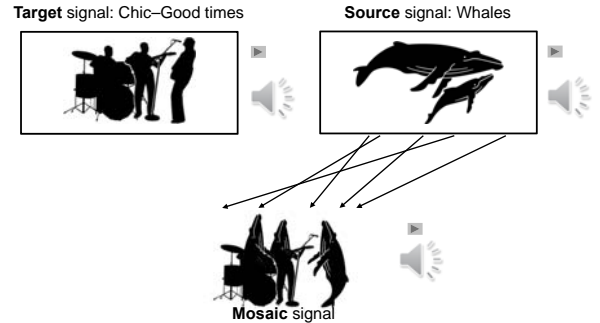
Musically Informed Audio Decomposition



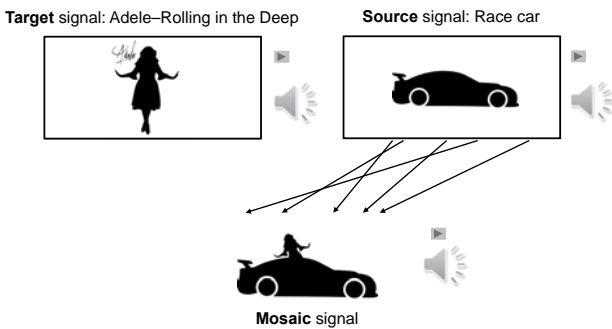
## Kullback-Leibler Divergence between Target and Mosaic



## Audio Mosaicing



## Audio Mosaicing

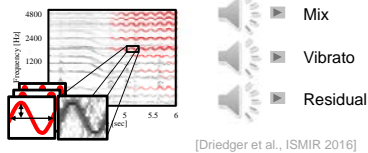


## Summary: Musically Informed Audio Decomposition

- Harmonic-percussive-residual decomposition
- Singing voice extraction
- Audio mosaicing

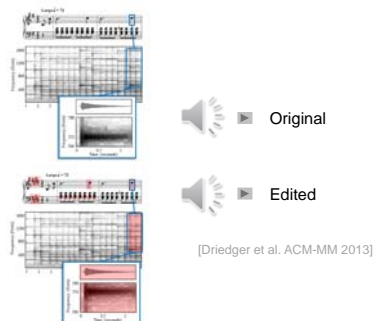
## Summary: Musically Informed Audio Decomposition

- Harmonic-percussive-residual decomposition
- Singing voice extraction
- Audio mosaicing
- **Vibrato-residual decomposition**



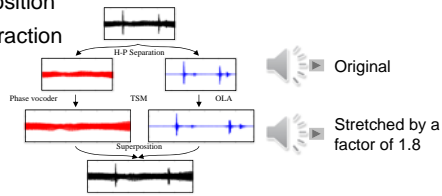
## Summary: Musically Informed Audio Decomposition

- Harmonic-percussive-residual decomposition
- Singing voice extraction
- Audio mosaicing
- Vibrato-residual decomposition
- **Score-informed audio editing**



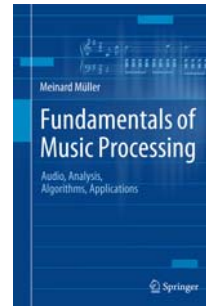
## Summary: Musically Informed Audio Decomposition

- Harmonic-percussive-residual decomposition
- Singing voice extraction
- Audio mosaicing
- Vibrato-residual decomposition
- Score-informed audio editing
- **Time-scale modification**



[Driedger et al. IEEE-SPL 2014, Driedger & Müller Appl. Sci. 2016]

## 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](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., 30 illus. in color, hardcover  
ISBN: 978-3-319-21944-8  
Springer, 2015

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

## References

- [Fitzgerald DAFX 2010] Derry Fitzgerald. Harmonic/percussive separation using medianfiltering. In Proceedings of the International Conference on Digital Audio Effects (DAFx), pages 246–253. Graz, Austria, 2010.
- [Tachibana et al. IEEE-TASLP 2013] Hideyuki Tachibana, Nobutaka Ono, and Shigeki Sagayama. Singing voice enhancement in monaural music signals based on two-stage harmonic/percussive sound separation on multiple resolution spectrograms. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 22(1):228–237, January 2013.
- [Driedger et al. ISMIR 2014] Jonathan Driedger, Meinard Müller, and Sascha Disch. "Extending harmonic-percussive separation of audio signals." in Proceedings of the 15th International Conference on Music Information Retrieval (ISMIR), Taipei, Taiwan, 2014.
- [Virtanen et al. ISCA 2008] Tuomas Virtanen, Annamaria Mesáros, and Matti Ryyänänen. "Combining pitch-based inference and non-negative spectrogram factorization in separating vocals from polyphonic music." in Proceedings of the ISCA Tutorial and Research Workshop on Statistical And Perceptual Audition (SAPA), 2008
- [Salamon & Gómez IEEE-TASLP 2012] Justin Salamon and Emilia Gómez. "Melody extraction from polyphonic music signals using pitch contour characteristics." IEEE Transactions on Audio, Speech & Language Processing, vol. 20, no. 6, pp. 1759–770, 2012.
- [Talmon et al. IEEE-TASLP 2011] Ronen Talmon, Israel Cohen, and Sharon Gannot. "Transient noise reduction using nonlocal diffusion filters." IEEE Transactions on Audio, Speech & Language Processing, 2011.
- [Candès et al. JACM 2011] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright. "Robust principal component analysis?," Journal of the ACM, vol. 58, no. 3, pp. 11:1–11:37, June 2011.
- [Huang et al. ICASSP 2012] Po-Sen Huang, Scott Deaann Chen, Paris Smaragdis, and Mark Hasegawa-Johnson. "Singing-voice separation from monaural recordings using robust principal component analysis," in Proceedings of ICASSP, 2012.

## Acknowledgements

This work has been supported by the German Research Foundation (DFG MU 2686/6-1). The International Audio Laboratories Erlangen are a joint institution of the Friedrich Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer IIS.