Lecture

Music Processing

Audio Decomposition

Meinard Müller
International Audio Laboratories Erlangen
meinard.mueller@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
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Music Processing Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Music Representations</td>
</tr>
<tr>
<td>2</td>
<td>Fourier Analysis of Signals</td>
</tr>
<tr>
<td>3</td>
<td>Music Synchronization</td>
</tr>
<tr>
<td>4</td>
<td>Music Structure Analysis</td>
</tr>
<tr>
<td>5</td>
<td>Chord Recognition</td>
</tr>
<tr>
<td>6</td>
<td>Tempo and Beat Tracking</td>
</tr>
<tr>
<td>7</td>
<td>Content-Based Audio Retrieval</td>
</tr>
<tr>
<td>8</td>
<td>Musically Informed Audio Decomposition</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Music Processing Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Music Representations</td>
</tr>
<tr>
<td>2</td>
<td>Fourier Analysis of Signals</td>
</tr>
<tr>
<td>3</td>
<td>Music Synchronization</td>
</tr>
<tr>
<td>4</td>
<td>Music Structure Analysis</td>
</tr>
<tr>
<td>5</td>
<td>Chord Recognition</td>
</tr>
<tr>
<td>6</td>
<td>Tempo and Beat Tracking</td>
</tr>
<tr>
<td>7</td>
<td>Content-Based Audio Retrieval</td>
</tr>
<tr>
<td>8</td>
<td>Musically Informed Audio Decomposition</td>
</tr>
</tbody>
</table>

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
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
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

Mixture
Harmonic-Percussive Decomposition

Mixture:
- Clearly harmonic sounds
- Clearly percussive sounds

Harmonic component

Percussive component
Harmonic-Percussive Decomposition

Mixture

Clearly harmonic sounds

Harmonic component

Residual component

Clearly percussive sounds

Percussive component
Harmonic-Percussive Decomposition

Mixture:
- Clearly harmonic sounds of singing voice and accompaniment
- Drum hits
- Fricatives & plosives in singing voice
- Noise-like sounds
- Vibrato/glissando sounds

Harmonic component
Residual component
Percussive component

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

Original Recording

Singing voice

Accompaniment
Singing Voice Extraction

Original recording

Harmonic component

Percussive component

Residual component

Harmonic portion

singing voice

Harmonic portion

accompaniment

Fricatives

singing voice

Instrument onsets

accompaniment

Vibrato & formants

singing voice

Diffuse instruments sounds

accompaniment

F0 annotation

Time

Frequency

Estimate

singing voice

Estimate

accompaniment
Score-Informed Source Separation

Exploit musical score to support separation process
Parametric Model Approach

Rebuild spectrogram information
NMF (Nonnegative Matrix Factorization)

\[
\begin{align*}
N & \quad M \\
K & \quad K \\
\geq 0 & \quad \geq 0
\end{align*}
\]
NMF (Nonnegative Matrix Factorization)

Magnitude Spectrogram $\approx$ Templates

Templates: Pitch + Timbre
Activations: Onset time + Duration

“How does it sound”
“When does it sound”
NMF-Decomposition

Random initialization
NMF-Decomposition

Initialized template

Learnt templates

Initialized activations

Learnt activations

Random initialization → No semantic meaning
NMF-Decomposition

**Initialized template**

**Initialized activations**

**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

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-LetItBee
NMF-Inspired Audio Mosaicing

Non-negative matrix factorization (NMF)

Non-negative matrix Components Activations

\[ V \approx W \cdot H = WH \]

Proposed audio mosaicing approach

Target's spectrogram Source's spectrogram Activations Mosaic's spectrogram
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

Spectrogram target

Spectrogram source

Activation matrix

Spectrogram 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
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/