The notion of sound object
Time,
Frequency,
and Time-Frequency
Time-Frequency analysis:
Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Allocating, Detecting and Mining Sound Structures

Dörfler, Monika¹
monika.doerfler@univie.ac.at

28th Sept., 2012

¹Funded by Audio.Miner (WWTF Vienna)
The notion of sound object

AudioMiner: Vienna-based interdisciplinary project. Exploring applications of modern mathematical methods in MIR.
The notion of sound object

AudioMiner: Vienna-based interdisciplinary project. Exploring applications of modern mathematical methods in MIR.
Detection, classification and processing of sound objects
The notion of sound object

AudioMiner: Vienna-based interdisciplinary project. Exploring applications of modern mathematical methods in MIR.
Detection, classification and processing of sound objects
Mathematics $\leftrightarrow$ MIR - both ways!
The notion of sound object

AudioMiner: Vienna-based interdisciplinary project. Exploring applications of modern mathematical methods in MIR.
Detection, classification and processing of sound objects
Mathematics ←→ MIR - both ways!
Particular focus on identification and extraction of sound objects from (polyphonic) audio or acousmatic sound signals
The notion of sound object

AudioMiner: Vienna-based interdisciplinary project. Exploring applications of modern mathematical methods in MIR.
Detection, classification and processing of sound objects
Mathematics $\leftrightarrow$ MIR - both ways!
Particular focus on identification and extraction of sound objects from (polyphonic) audio or acousmatic sound signals $\rightarrow$ but what is a sound object?
The notion of sound object

Sound object, sound structure or Gestalt?

The notion of sound object

Sound object, sound structure or Gestalt?

- G. Monaci, P. Vandergheynst, ”Audiovisual Gestalts,” 2012

Object, Structure or Gestalt: interpretation depends on representation!
The notion of sound object

Object or Gestalt implies sparsity or compactness in representation.
The notion of sound object

Object or Gestalt implies sparsity or compactness in representation.
Representation by few as opposed to many elements.
Object or Gestalt implies sparsity or compactness in representation.
Representation by *few* as opposed to *many* elements.
What do we mean by representation?
The notion of sound object

Object or Gestalt implies sparsity or compactness in representation. Representation by *few* as opposed to *many* elements. What do we mean by representation?
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis:
- Principles
- Time-frequency Multipliers
- Adaptive signal representations
- Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

\[ f[n] = \sum_{n} c_n \delta_n = \sum_{k} \hat{c}_k (\cos 2\pi kn + \sin 2\pi kn) = \sum_{n,k} c_{n,k} \varphi_{n,k} \]
Analysis tool should render representation for visual display, reflecting a user’s acoustical impression.
The notion of sound object

Time, Frequency, and Time-Frequency

- Analysis tool should render representation for visual display, reflecting a user’s acoustical impression.
- Sound objects should be visible as distinct from a more textural background.
- More sophisticated representations of sound signals than classical representations do exist now:
  - Adaptivity in the transformation parameters: sharpen the visual and assure a perfect connection between signal and representation (invertibility).
  - Bayesian coefficient priors highlight particular structures by means of informed analysis.

Dörfler, Monika | http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

- Analysis tool should render representation for visual display, reflecting a user’s acoustical impression.
- Sound objects should be visible as distinct from a more textural background.
- More sophisticated representations of sound signals than classical representations do exist now:
  - Adaptivity in the transformation parameters: sharpen the visual and assure a perfect connection between signal and representation (invertibility).
  - Bayesian coefficient priors highlight particular structures by means of informed analysis.

Dörfler, Monika http://nuhag.eu
Time-Frequency analysis: Principles

- FFT-based time-variant analysis
- Mathematical frame theory always provides perfect reconstruction

\[ f = \sum_{n,k} \langle f, \varphi_{n,k}^t \rangle \varphi_{n,k}^t \]

\[ \varphi_{n,k} := M_{kb} T_{na} \varphi \]

- Choice of involved parameters \((\varphi, a, b)\) is vital for the quality of representation
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika
http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörrler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika | http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers

Adaptive signal representations

Sparsity and structured representations
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Dörfler, Monika

http://nuhag.eu
Time-frequency Multipliers

Dörfler, Monika | http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers

Adaptive signal representations

Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis:
Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers

Adaptive signal representations

Sparsity and structured representations

Dörfler, Monika  http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles

Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Centered Plot
Adaptive signal representations

The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis:
Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika
http://nuhag.eu
Adaptive signal representations

The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Dörfler, Monika http://nuhag.eu
Adaptive signal representations

The notion of sound object
Time, Frequency, and Time-Frequency
Time-Frequency analysis: Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

Glockenspiel – dB-scaled Gabor transform

Glockenspiel – dB-scaled NSGT

(Listen!) (Listen!)
Adaptive signal representations

Exploring non-stationarity on the frequency side

Figure: STFT with adapted window

Dörfler, Monika http://nuhag.eu
Adaptive signal representations

Exploring non-stationarity on the frequency side

hk - dB-scaled Gabor transform

hk - dB-scaled CQ-NSGT

Dörfler, Monika  http://nuhag.eu
Adaptive signal representations

Exploring non-stationarity on the frequency side

Figure: STFT with adapted window

Redmanns - dB-scaled Gabor transform

Redmanns - dB-scaled CQ-NSGT

Dörfler, Monika

http://nuhag.eu
DISCUSSION

- The meaning of reconstruction.

Dörfler, Monika
http://nuhag.eu
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis: Principles
Time-frequency Multipliers
Adaptive signal representations
Sparsity and structured representations

DISCUSSION

- The meaning of reconstruction.
- The meaning of adaptivity: salience of sound objects/Gestalts.
Sparsity and structured representations

http://homepage.univie.ac.at/monika.doerfler/StrucAudio.html

Regression problems with mixed-norm priors on time-frequency coefficients lead to structured, sparse representations of audio signals.
Regression problems with mixed-norm priors on time-frequency coefficients lead to structured, sparse representations of audio signals.

Systematic formulation of thresholding operators allows for weighting in the time-frequency domain.
The notion of sound object

Time, Frequency, and Time-Frequency

Time-Frequency analysis:
Principles

Time-frequency Multipliers
Adaptive signal representations

Sparsity and structured representations

Sparsity and structured representations

http://homepage.univie.ac.at/monika.doerfler/StrucAudio.html

- Regression problems with mixed-norm priors on time-frequency coefficients lead to structured, sparse representations of audio signals.

- Systematic formulation of thresholding operators allows for weighting in the time-frequency domain.

- Related iterative algorithms are evaluated on synthetic and real-life audio signals in the context of denoising and multi-layer decomposition.
Regression problems with mixed-norm priors on time-frequency coefficients lead to structured, sparse representations of audio signals.

Systematic formulation of thresholding operators allows for weighting in the time-frequency domain.

Related iterative algorithms are evaluated on synthetic and real-life audio signals in the context of denoising and multi-layer decomposition.

Shape of the weighting masks influences results.
Regression problems with mixed-norm priors

\[ \Delta(f) = \left\| \sum_{n,k} c_{n,k} \varphi_{n,k}^t - \hat{f} \right\|_2^2 + \mu \| c \|_{\ell^1} \]
Regression problems with mixed-norm priors

\[
\Delta(f) = \| \sum_{n,k} c_{n,k} \varphi_{n,k}^t \hat{f} \|_2^2 + \mu \| \mathbf{c} \|_{\ell^1}
\]

...becomes...

\[
\Delta(f) = \| \sum_{n,k} c_{n,k} \varphi_{n,k}^t \hat{f} \|_2^2 + \mu \| \mathbf{c} \|_{p,q}
\]

algorithmic solution is iterative soft thresholding.
Thresholding operators with weighting in the time-frequency domain

\[ S_\omega(c_\gamma) = c_\gamma (1 - \frac{\mu}{\left(\sum_{\gamma' \in U_\gamma} \nu_{\gamma}(\gamma')|c_{\gamma'}|^2\right)^{1/2}}) + \]

Dörfler, Monika  
http://nuhag.eu
Application to denoising-Bird
Application to denoising - Lasso, unstructured
Application to denoising - Windowed Lasso
Application to denoising - Persistent Elitist Lasso
Application to multi-layer decomposition.
Application to multi-layer decomposition.
Application to multi-layer decomposition.
Application to multi-layer decomposition.

- Listen to residual!

(Reconstruction Ton+Trans+Res)
Influence of weighting mask shape

Figure: Iterated WGL shrinkage results for different shapes (i.e. weightings) of the neighborhood on an excerpt including piano, double-bass and drums and on an snare drum hit excerpt. The result of unsmoothed shrinkage, Rectangular, Simplet (= simple tent, starting at 1 and then linearly decaying to zero), Telpmis (= time-reversed simple tent).
Perception of visual representation is highly dependent on analysis tools.

Also in higher level processing, this influence should be considered.

Prior knowledge can also be used in low level processing to extract meaningful information.