

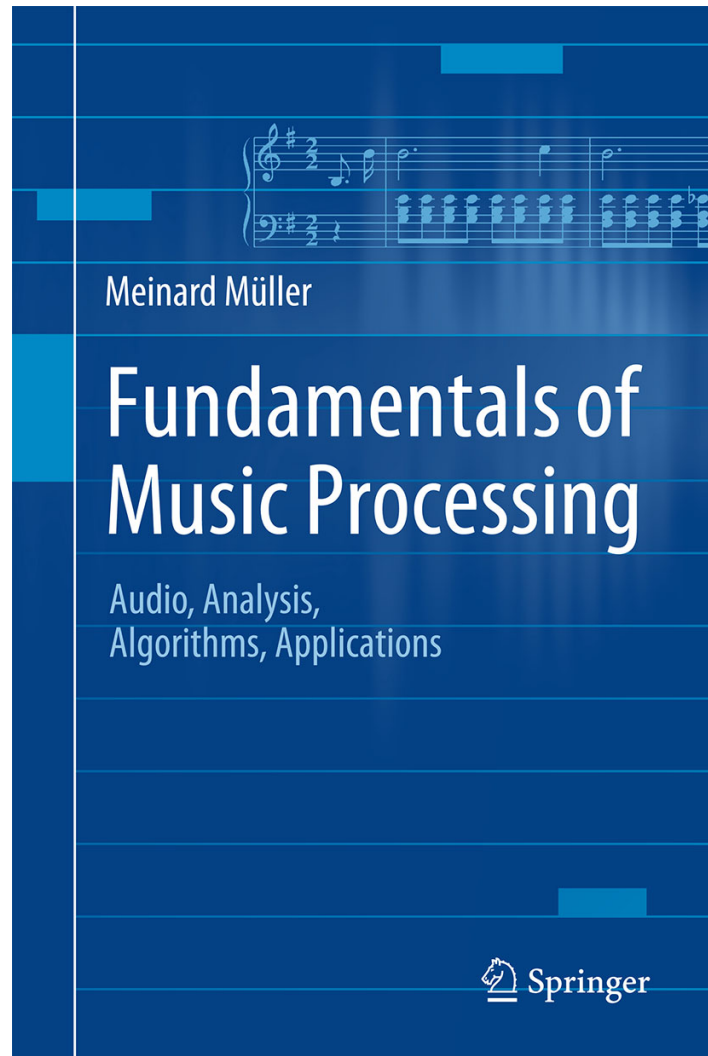
Lecture
Music Processing

Tempo and Beat Tracking

Meinard Müller

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

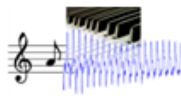

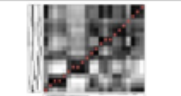


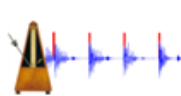
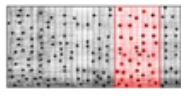
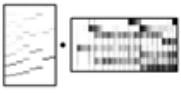
Book: Fundamentals of Music Processing



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

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

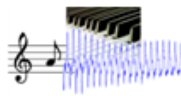

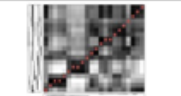
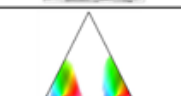

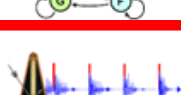


ISBN: 978-3-319-21944-8

Springer, 2015

Accompanying website:

www.music-processing.de

Book: Fundamentals of Music Processing

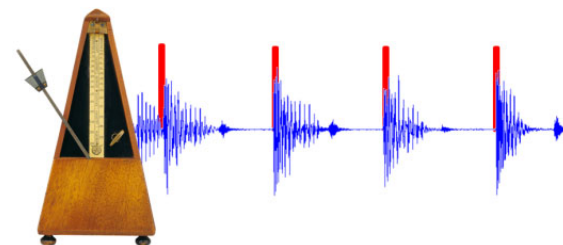
| Chapter | | Music Processing Scenario |
|---------|---|--|
| 1 |  | Music Representations |
| 2 |  | Fourier Analysis of Signals |
| 3 |  | Music Synchronization |
| 4 |  | Music Structure Analysis |
| 5 |  | Chord Recognition |
| 6 |  | Tempo and Beat Tracking |
| 7 |  | Content-Based Audio Retrieval |
| 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

Chapter 6: Tempo and Beat Tracking

- 6.1 Onset Detection
- 6.2 Tempo Analysis
- 6.3 Beat and Pulse Tracking
- 6.4 Further Notes



Tempo and beat are further fundamental properties of music. In Chapter 6, we introduce the basic ideas on how to extract tempo-related information from audio recordings. In this scenario, a first challenge is to locate note onset information—a task that requires methods for detecting changes in energy and spectral content. To derive tempo and beat information, note onset candidates are then analyzed with regard to quasiperiodic patterns. This leads us to the study of general methods for local periodicity analysis of time series.

Introduction

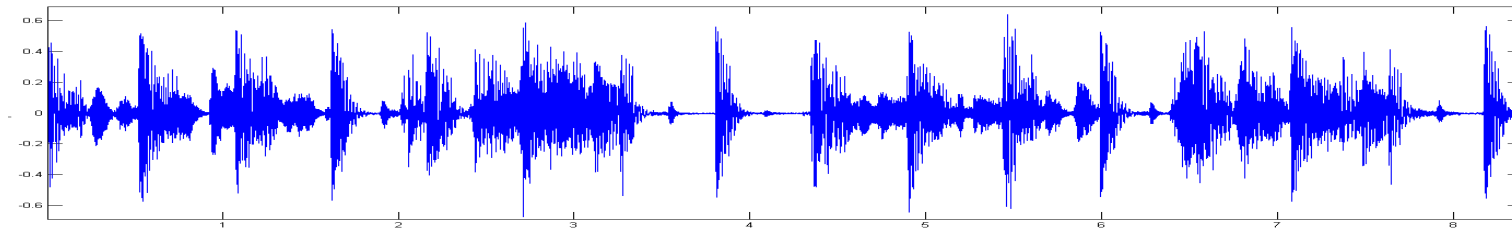
Basic beat tracking task:

Given an audio recording of a piece of music, determine the periodic sequence of beat positions.

“Tapping the foot when listening to music”

Introduction

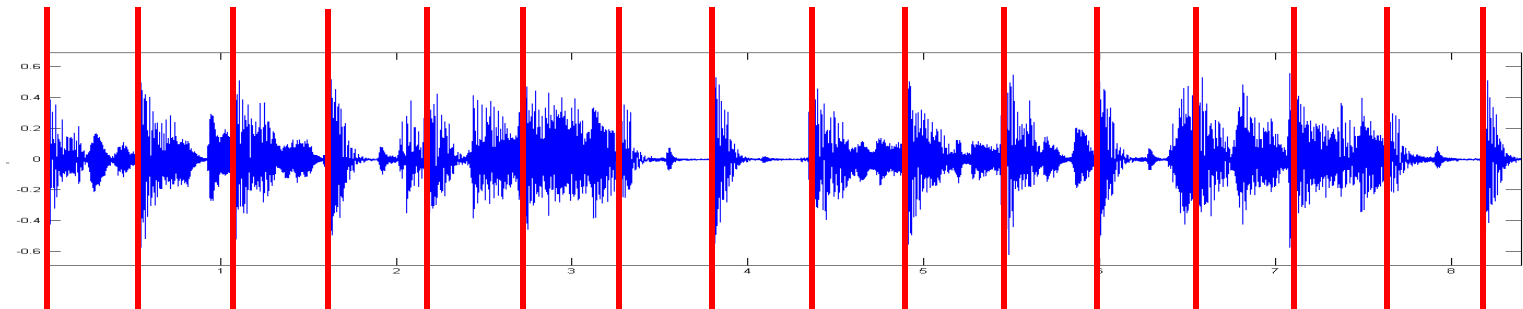
Example: Queen – Another One Bites The Dust



Time (seconds)

Introduction

Example: Queen – Another One Bites The Dust



Time (seconds)



Introduction

Example: Happy Birthday to you

Pulse level: **Measure**

The image shows two staves of musical notation for the song 'Happy Birthday to you'. The first staff contains the first two phrases: 'Hap - py Birth - day to you,' and 'Hap - py Birth - day to you, Hap - py'. The second staff contains the third phrase: 'Birth - day dear _____, Hap - py Birth - day to you!'. The music is in 3/4 time with a key signature of one sharp (F#). Four red arrows point downwards to the first note of each of the four measures in the first staff, indicating the pulse level at the measure level.

Introduction

Example: Happy Birthday to you

Pulse level: **Tactus (beat)**

The image shows a musical score for the song "Happy Birthday to you" in 3/4 time. The score is written on two staves. The first staff contains the melody for the first two phrases: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py". The second staff contains the melody for the final phrase: "Birth - day dear _____, Hap - py Birth - day to you!". Red arrows point down to the first beat of each measure in the first staff, indicating the pulse level (Tactus). The key signature is one sharp (F#) and the time signature is 3/4.

Introduction

Example: Happy Birthday to you

Pulse level: **Tatum (temporal atom)**

The image shows a musical score for the song "Happy Birthday to you" in G major and 3/4 time. The score is divided into two staves. The first staff contains the first line of the melody: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py". The second staff contains the second line: "Birth - day dear _____, Hap - py Birth - day to you!". Above the first staff, there are 24 red arrows pointing downwards to the notes, indicating the pulse level (Tatum) for each note. The arrows are placed above the notes in a regular, repeating pattern, corresponding to the 3/4 time signature.

Introduction

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: ???



Introduction

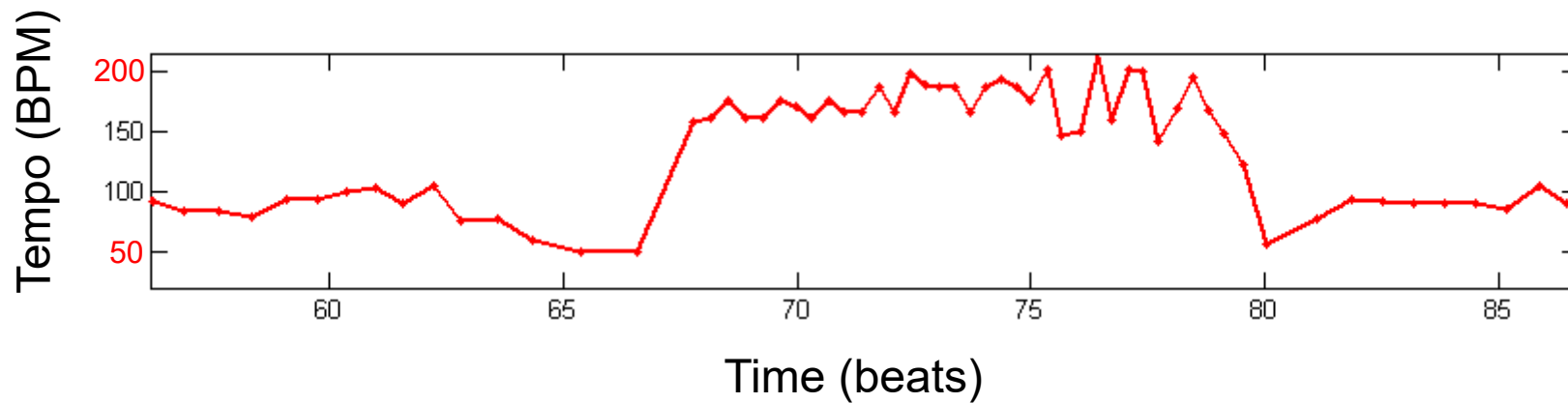
Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: **50-200 BPM**



Tempo curve



Introduction

Example: Borodin – String Quartet No. 2

Pulse level: Quarter note

Tempo: 120-140 BPM (roughly)

Beat tracker without any prior knowledge



Beat tracker with prior knowledge on
rough tempo range



Introduction

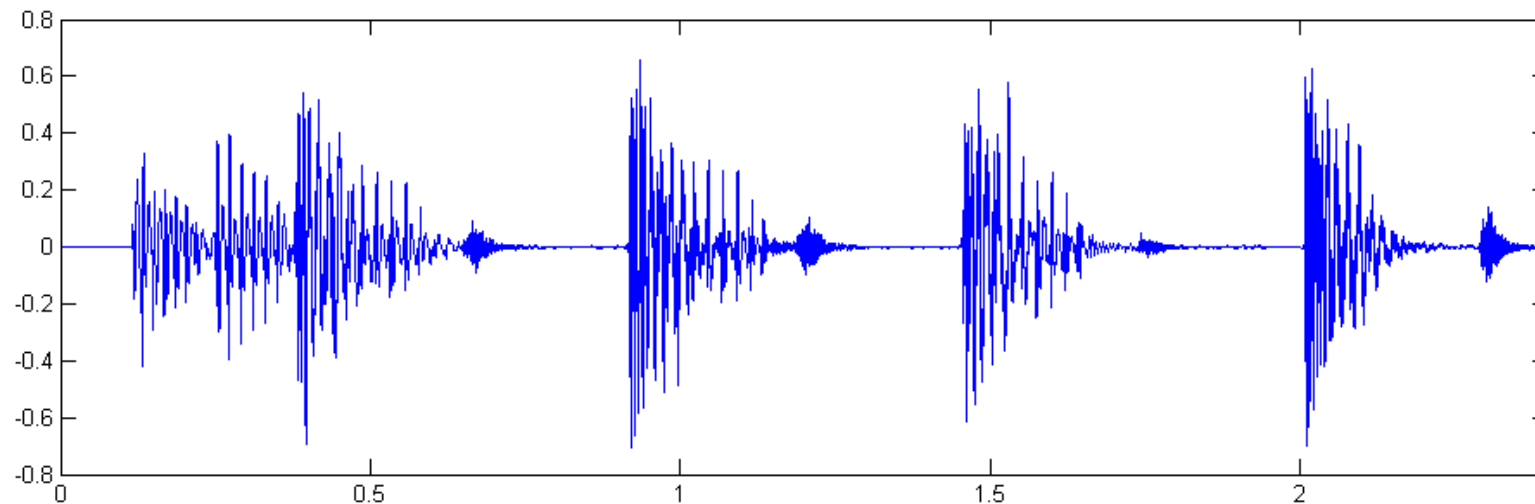
Challenges in beat tracking

- Pulse level often unclear
- Local/sudden tempo changes (e.g. rubato)
- Vague information
(e.g., soft onsets, extracted onsets corrupt)
- Sparse information
(often only note onsets are used)

Introduction

Tasks

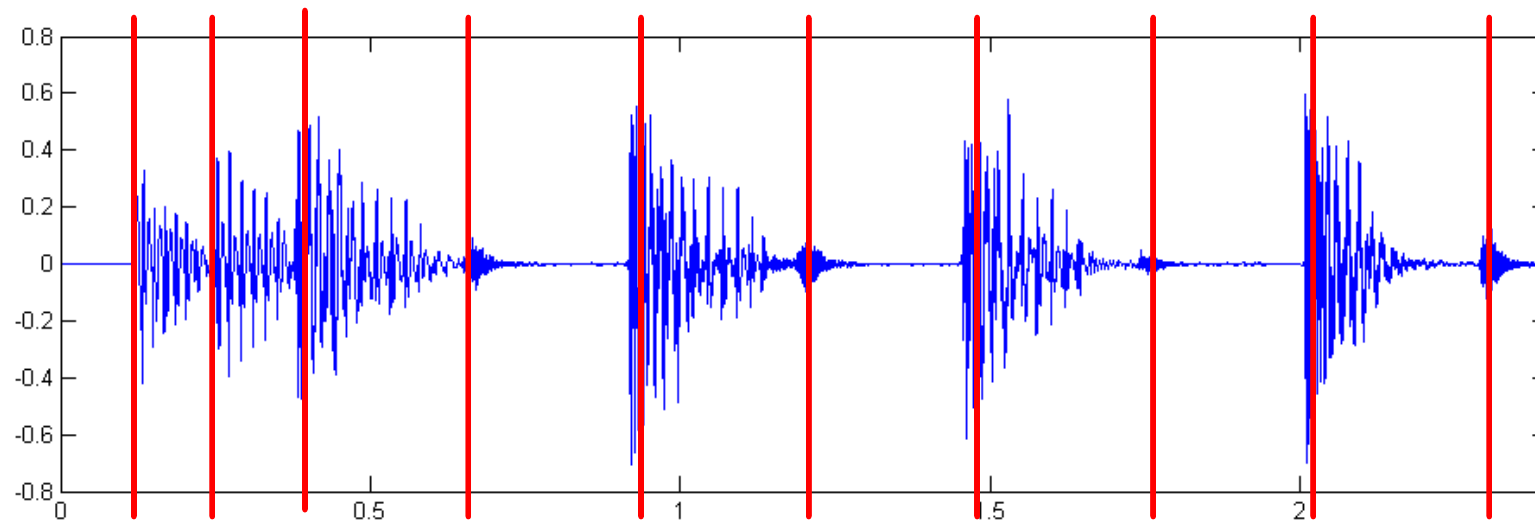
- Onset detection
- Beat tracking
- Tempo estimation



Introduction

Tasks

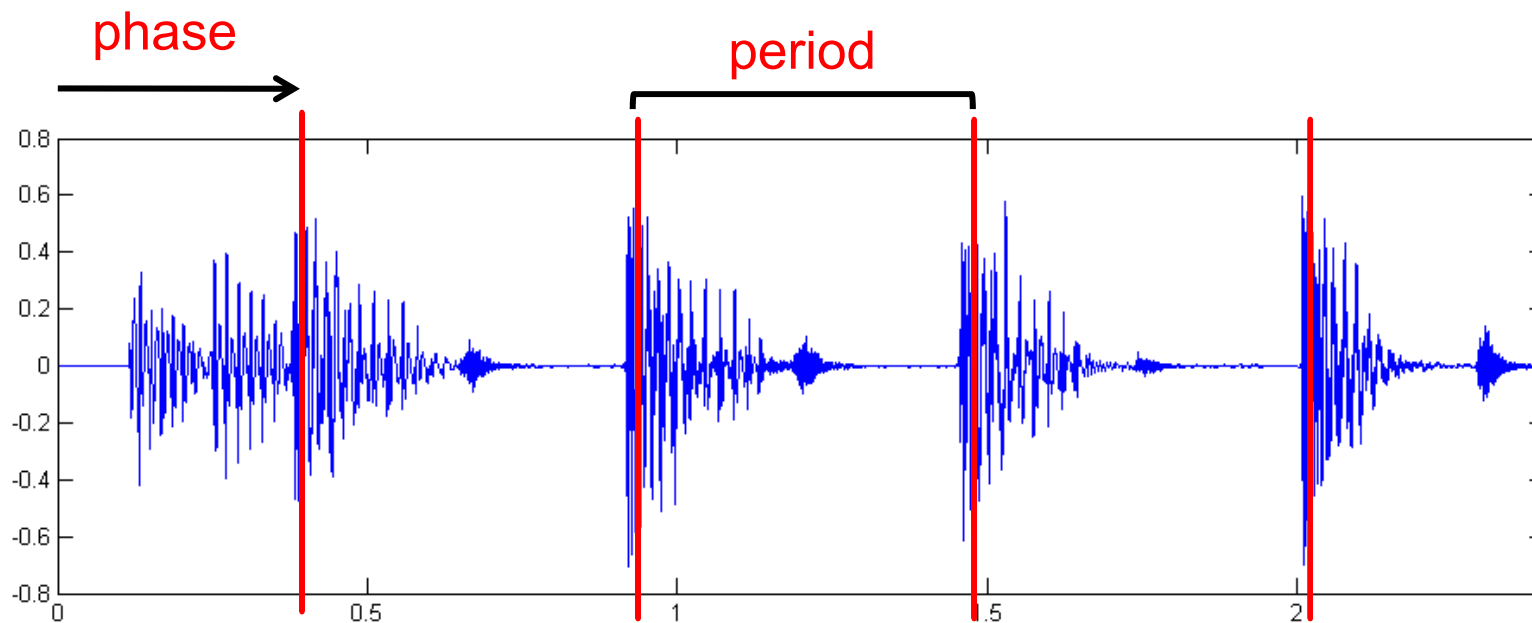
- Onset detection
- Beat tracking
- Tempo estimation



Introduction

Tasks

- Onset detection
- **Beat tracking**
- Tempo estimation



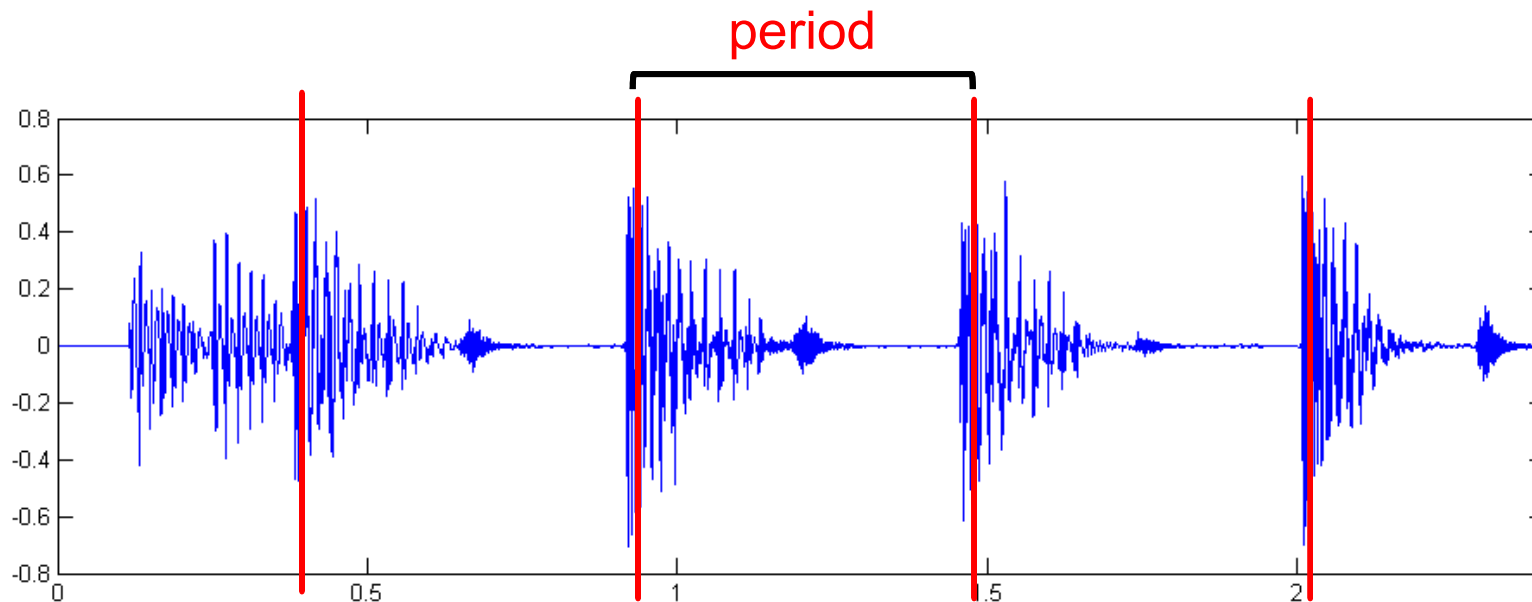
Introduction

Tasks

- Onset detection
- Beat tracking
- **Tempo estimation**

Tempo := 60 / period

Beats per minute (BPM)

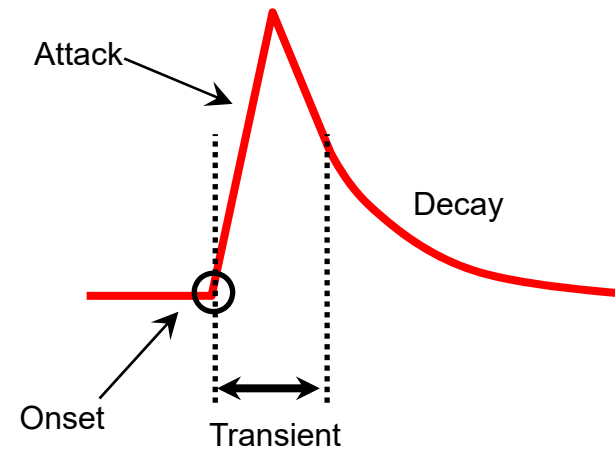
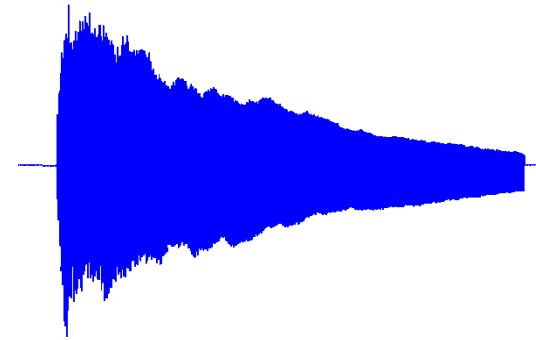


Onset Detection

- Finding start times of perceptually relevant acoustic events in music signal
- Onset is the time position where a note is played
- Onset typically goes along with a change of the signal's properties:
 - energy or loudness
 - pitch or harmony
 - timbre

Onset Detection

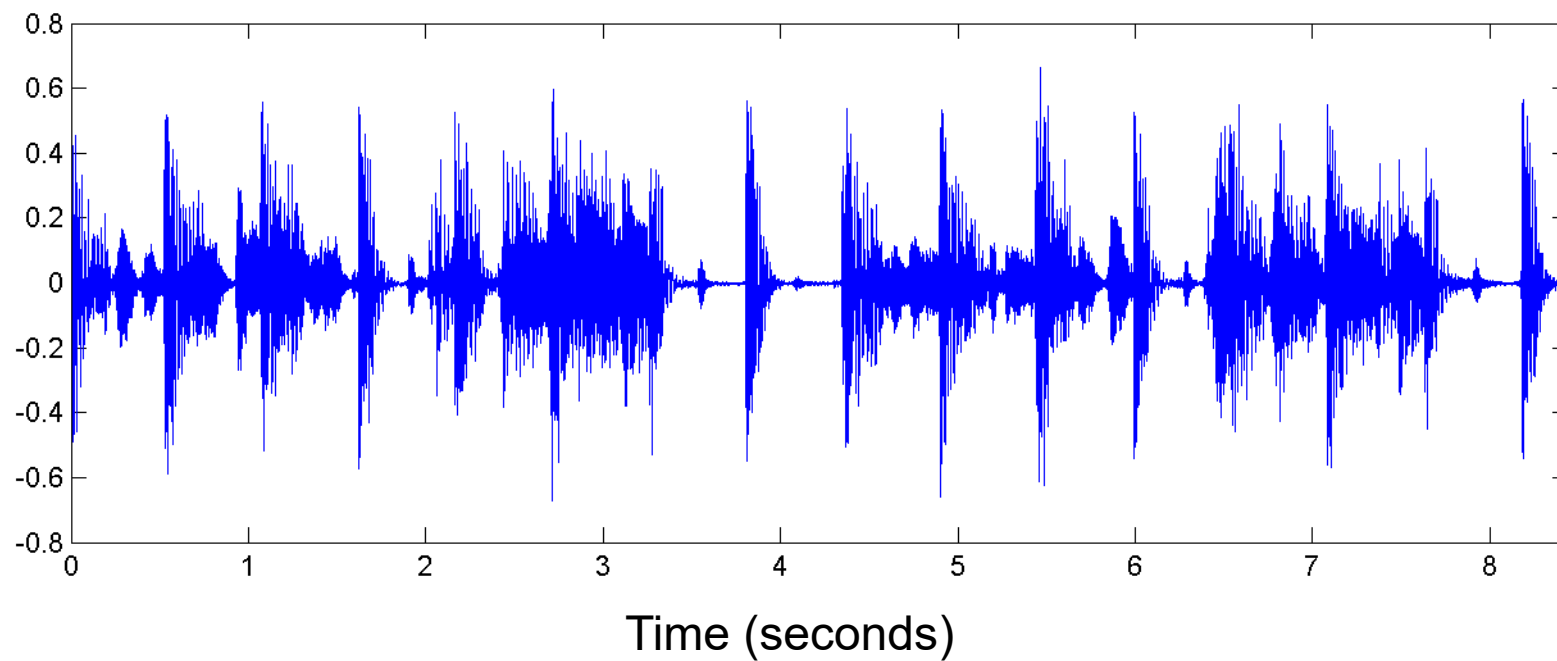
- Finding start times of perceptually relevant acoustic events in music signal
- Onset is the time position where a note is played
- Onset typically goes along with a change of the signal's properties:
 - energy or loudness
 - pitch or harmony
 - timbre



Onset Detection (Energy-Based)

Steps

Waveform

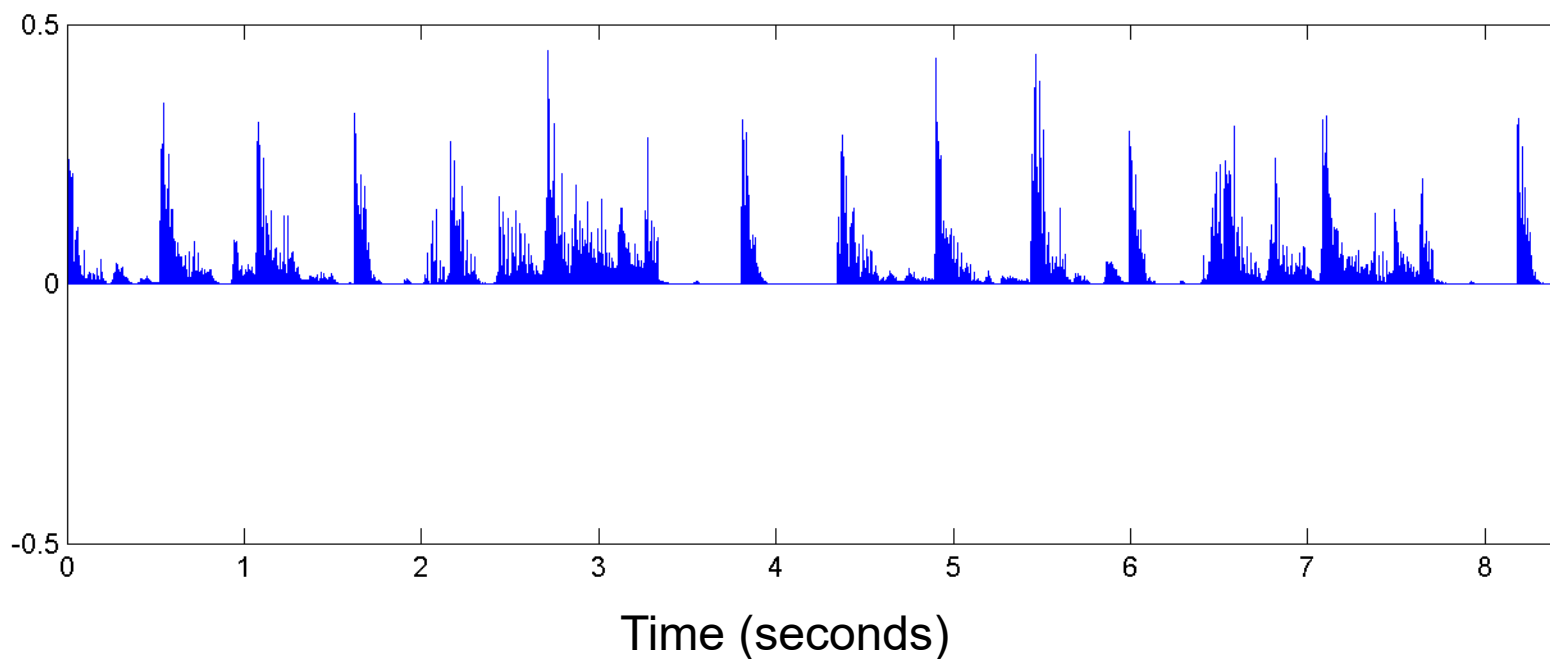


Onset Detection (Energy-Based)

Steps

1. Amplitude squaring

Squared waveform

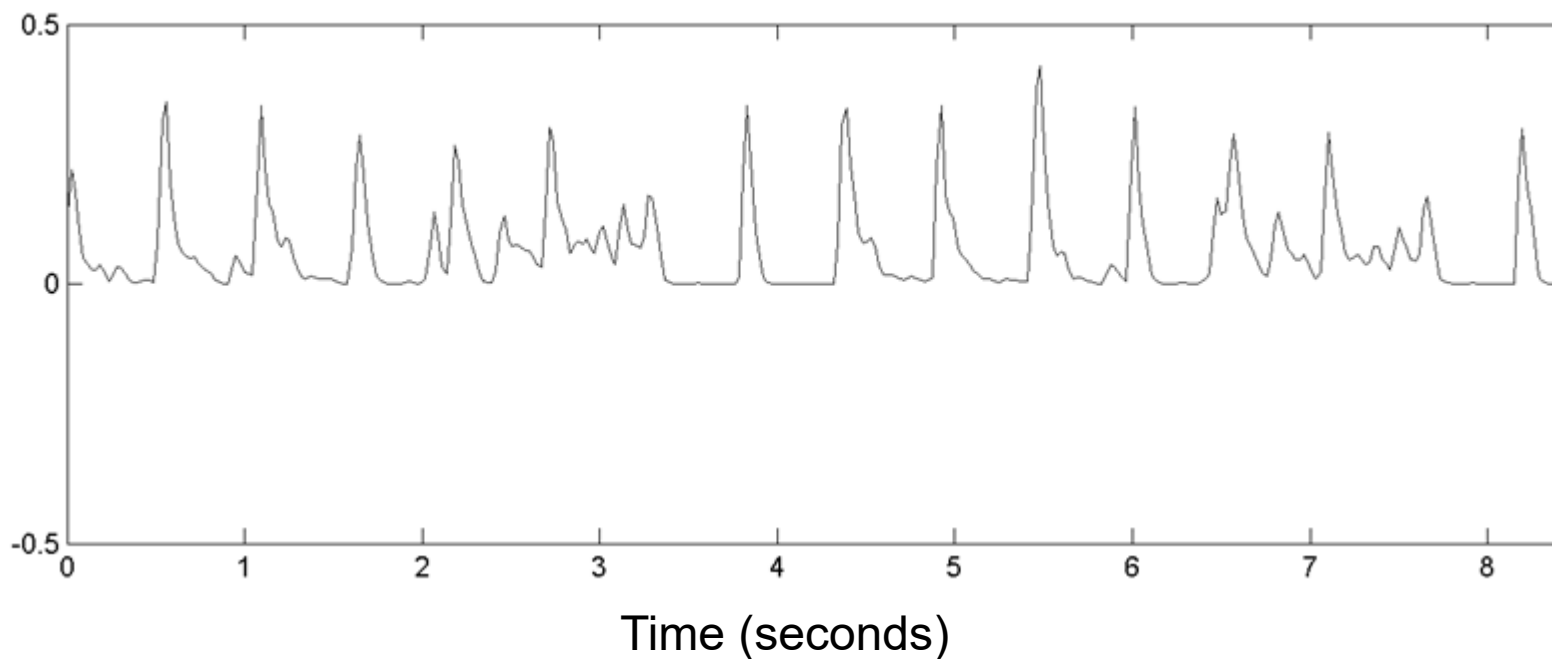


Onset Detection (Energy-Based)

Steps

1. Amplitude squaring
2. Windowing

Energy envelope



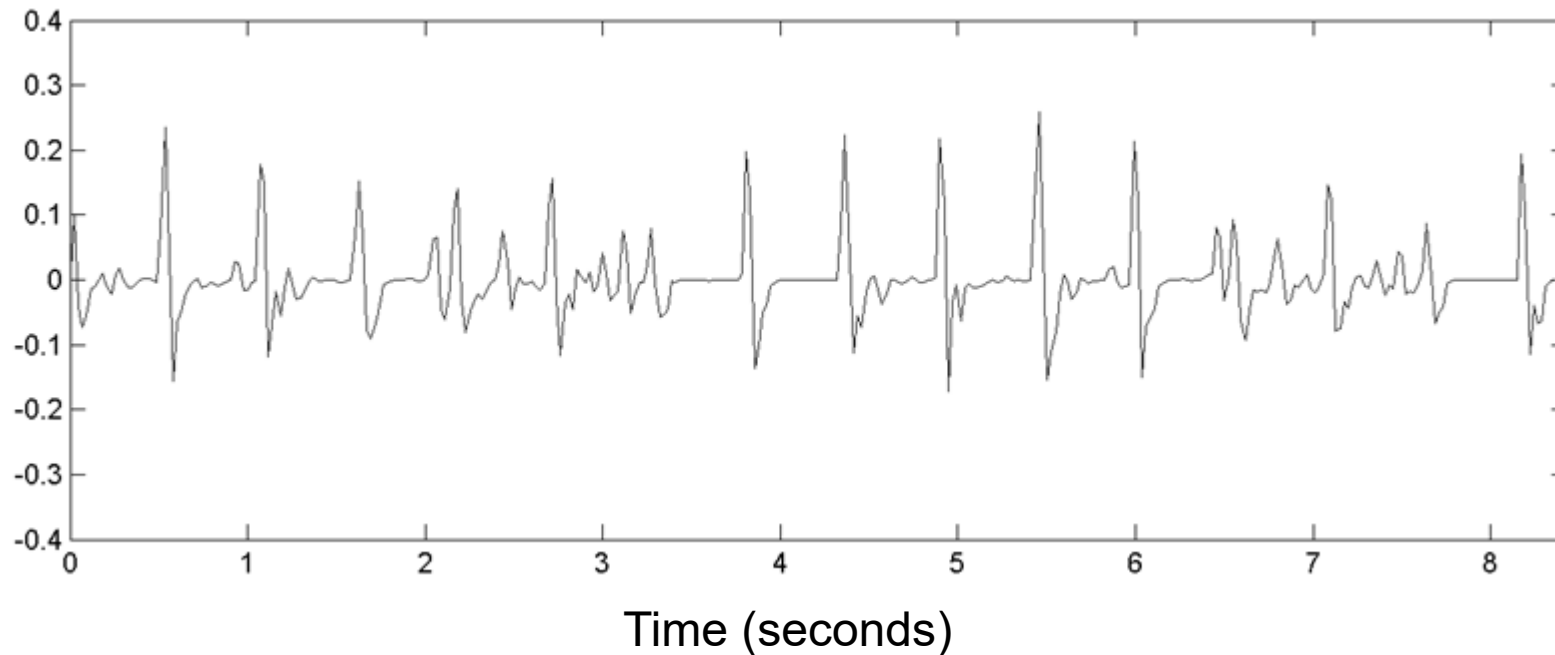
Onset Detection (Energy-Based)

Steps

1. Amplitude squaring
2. Windowing
3. Differentiation

Capturing energy changes

Differentiated energy envelope



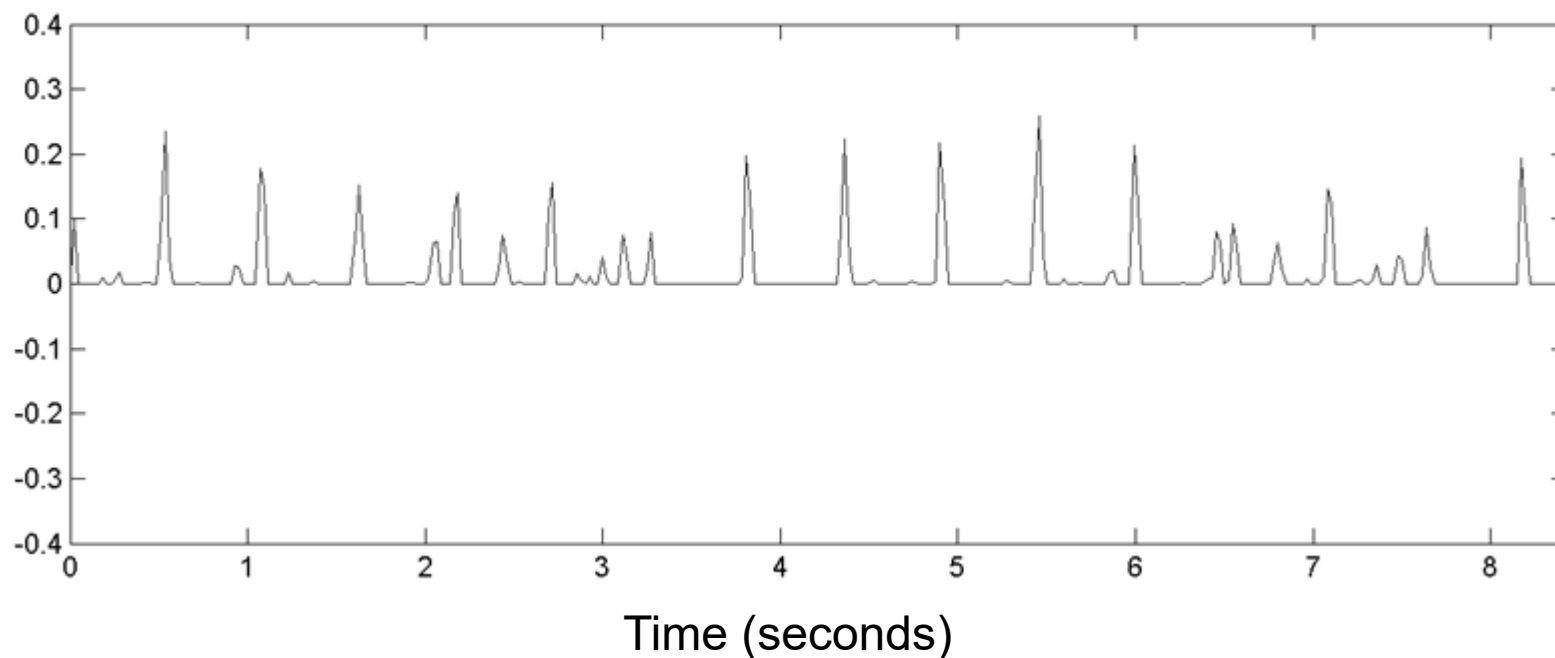
Onset Detection (Energy-Based)

Steps

1. Amplitude squaring
2. Windowing
3. Differentiation
4. Half wave rectification

Only energy increases are relevant for note onsets

Novelty curve

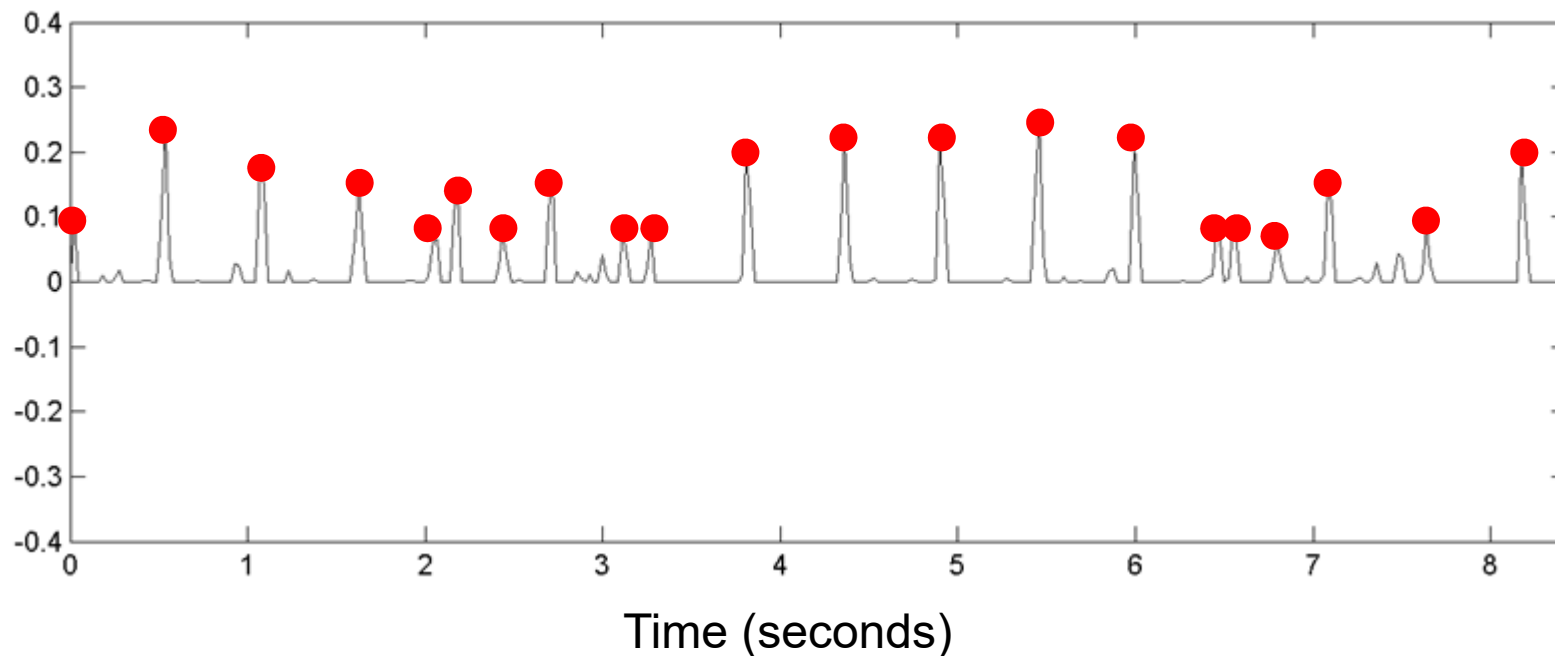


Onset Detection (Energy-Based)

Steps

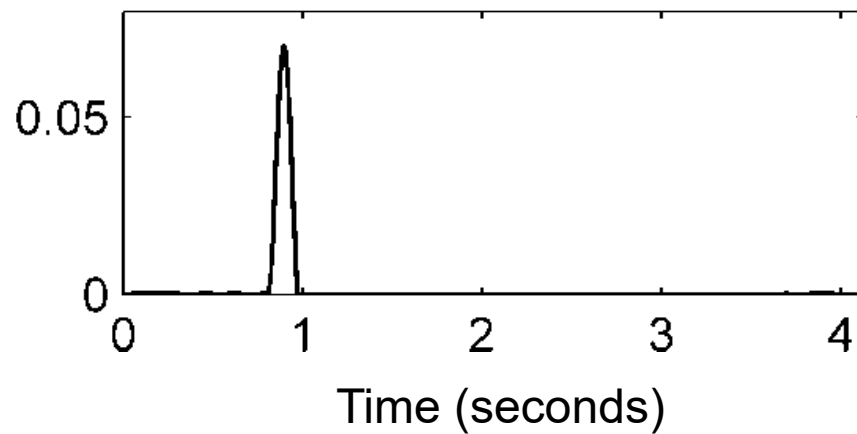
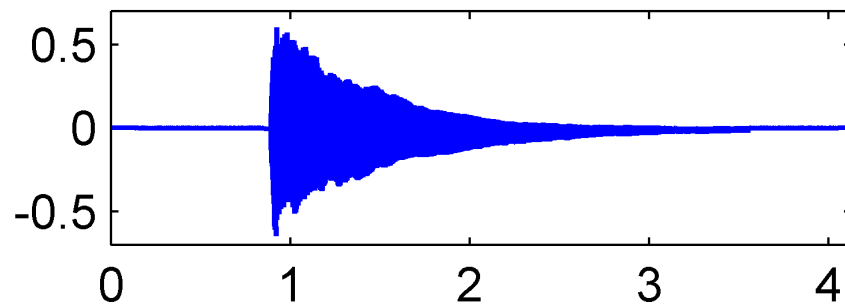
1. Amplitude squaring
2. Windowing
3. Differentiation
4. Half wave rectification
5. Peak picking

Peak positions indicate
note onset candidates



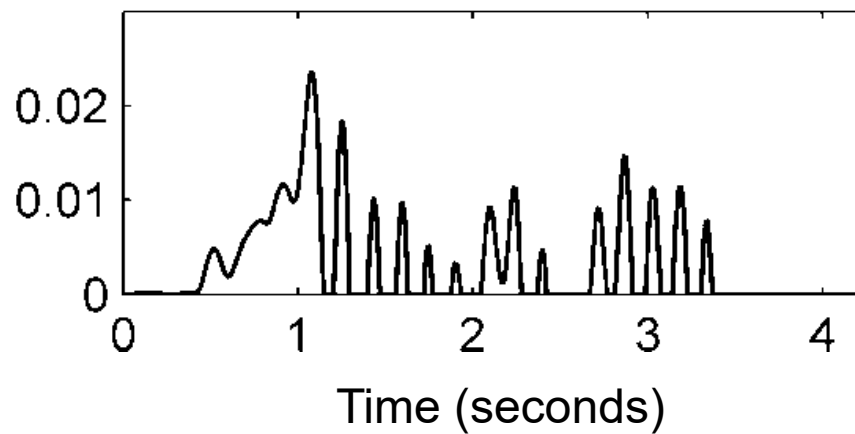
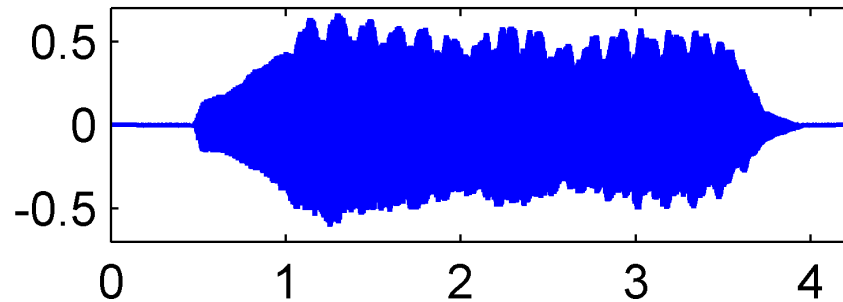
Onset Detection (Energy-Based)

Example: C4 played by piano



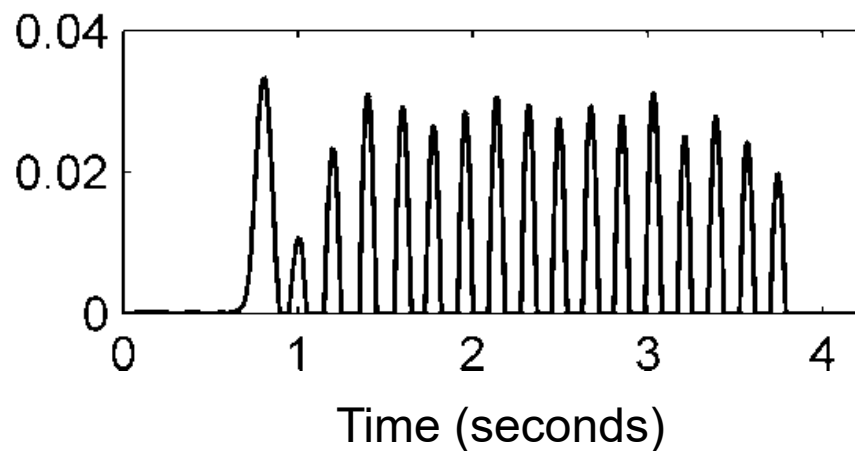
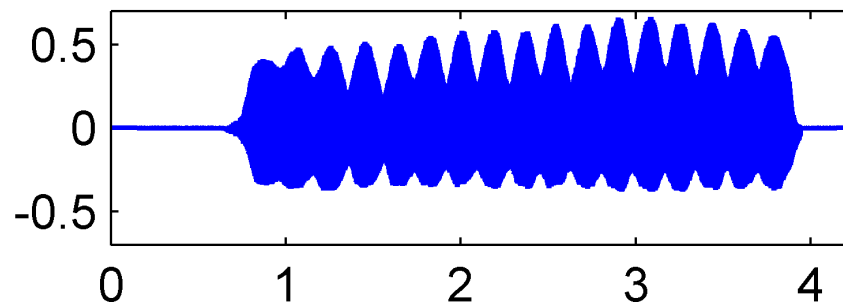
Onset Detection (Energy-Based)

Example: C4 played by violin



Onset Detection (Energy-Based)

Example: C4 played by flute



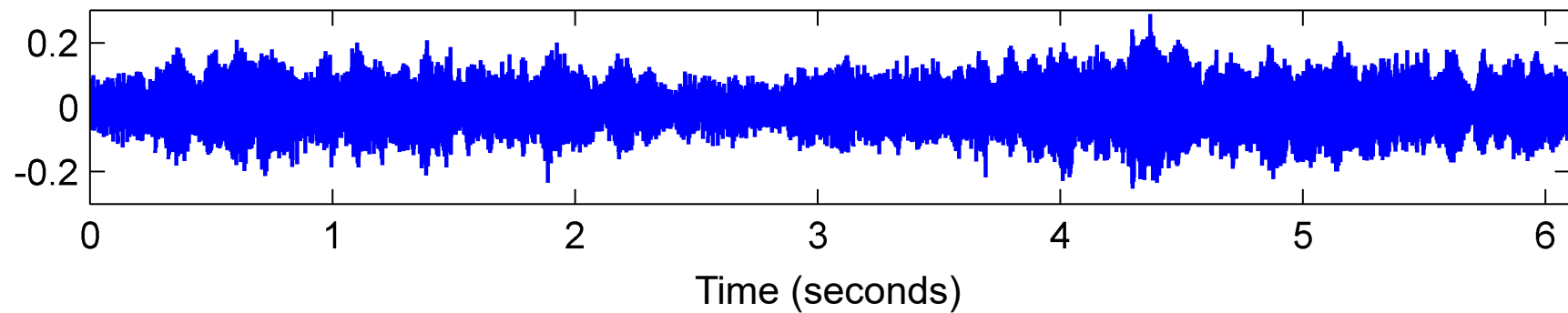
Onset Detection

- Energy curves often only work for percussive music
- Many instruments such as strings have weak note onsets
- No energy increase may be observable in complex sound mixtures
- More refined methods needed that capture
 - changes of spectral content
 - changes of pitch
 - changes of harmony

Onset Detection (Spectral-Based)



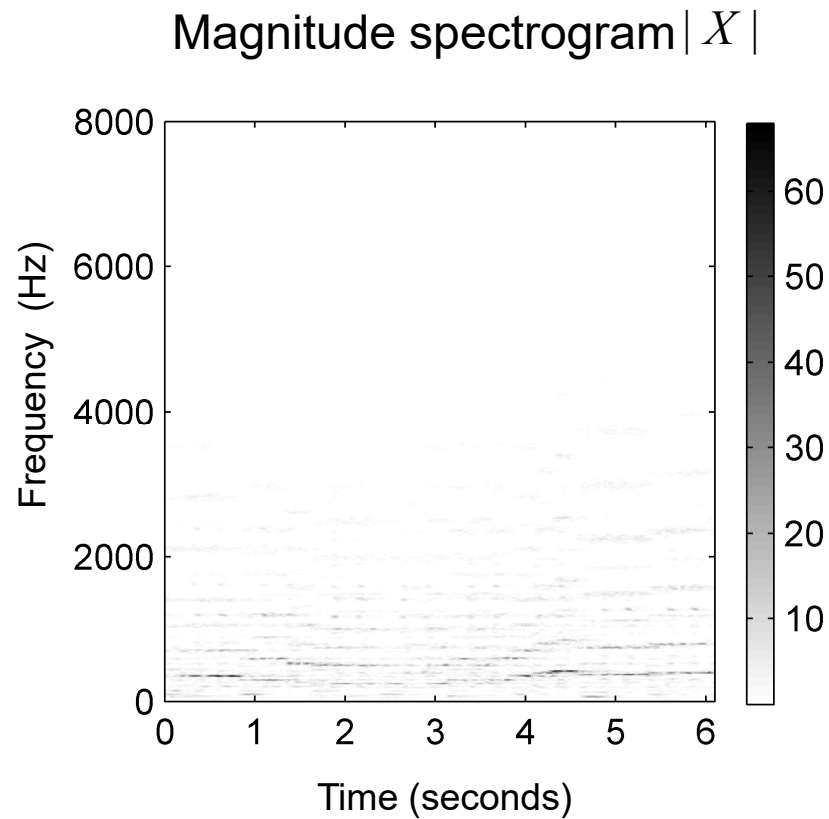
Audio recording



Onset Detection (Spectral-Based)

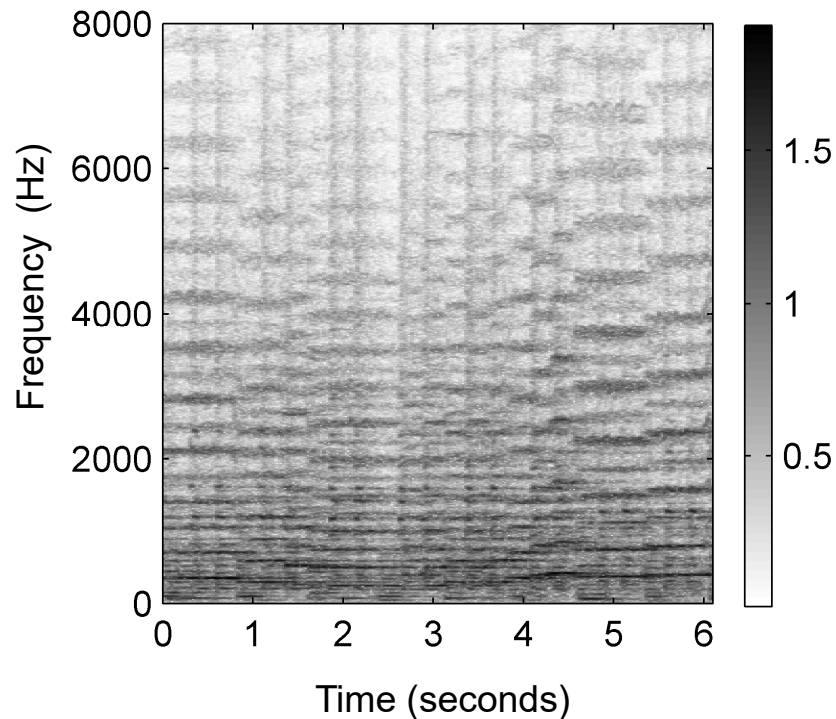
Steps:

1. Spectrogram



Onset Detection (Spectral-Based)

Compressed spectrogram Y

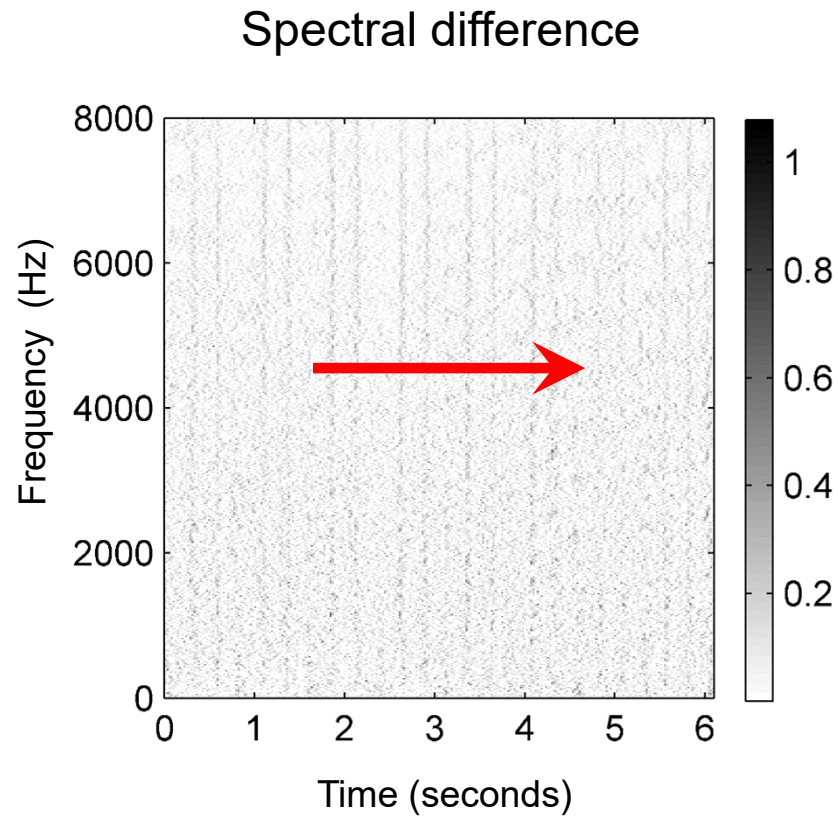


Steps:

1. Spectrogram
2. Logarithmic compression

$$Y = \log(1 + C \cdot |X|)$$

Onset Detection (Spectral-Based)



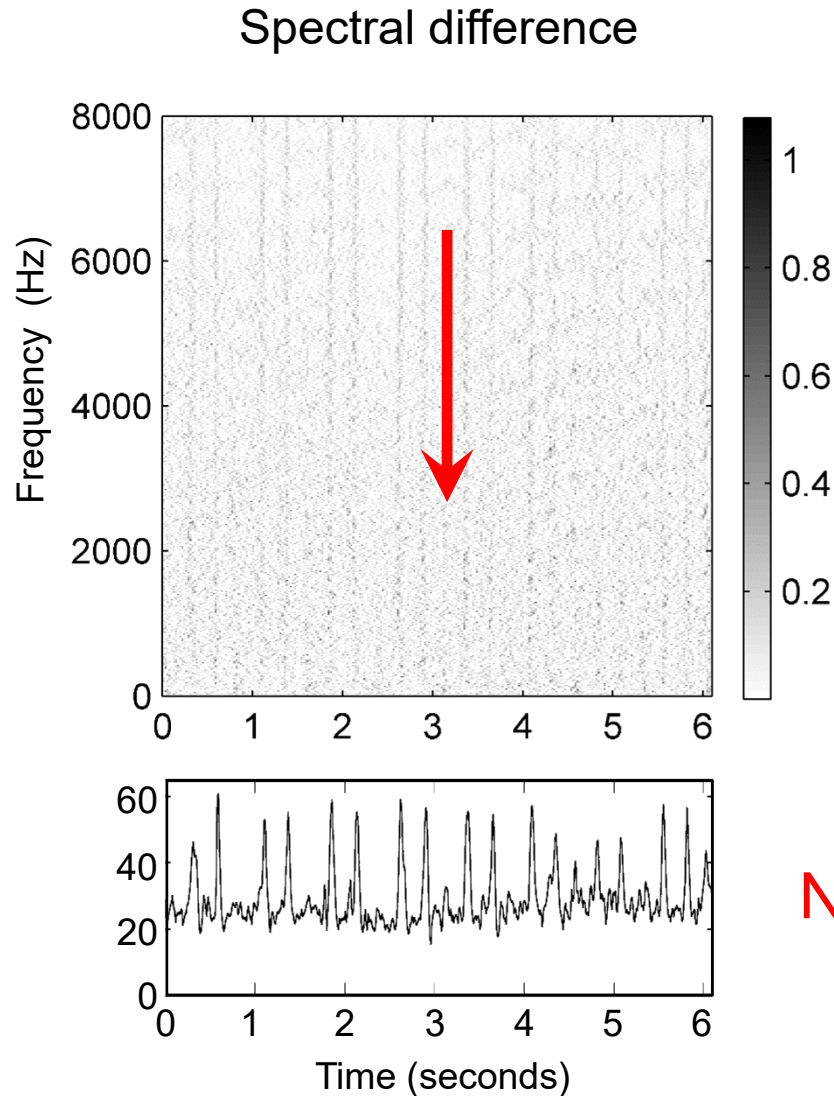
Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation & half wave rectification

Onset Detection (Spectral-Based)

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation & half wave rectification
4. Accumulation



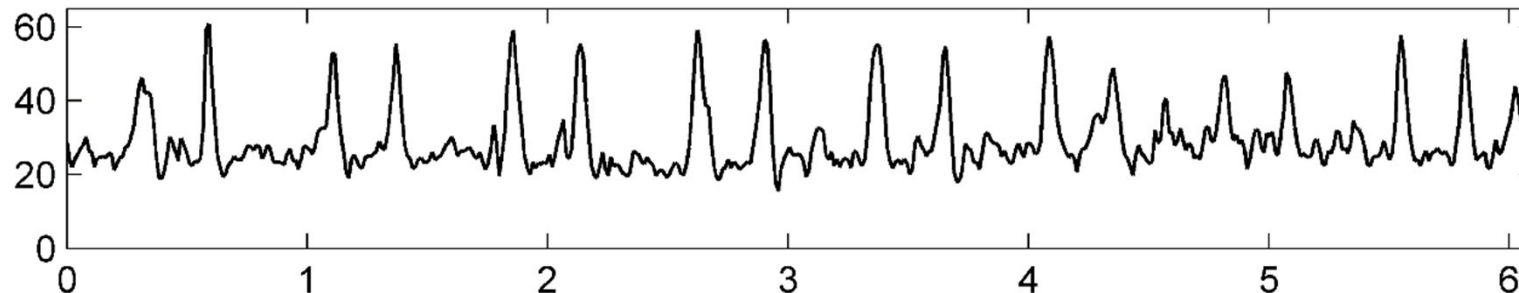
Novelty curve

Onset Detection (Spectral-Based)

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation & half wave rectification
4. Accumulation

Novelty curve



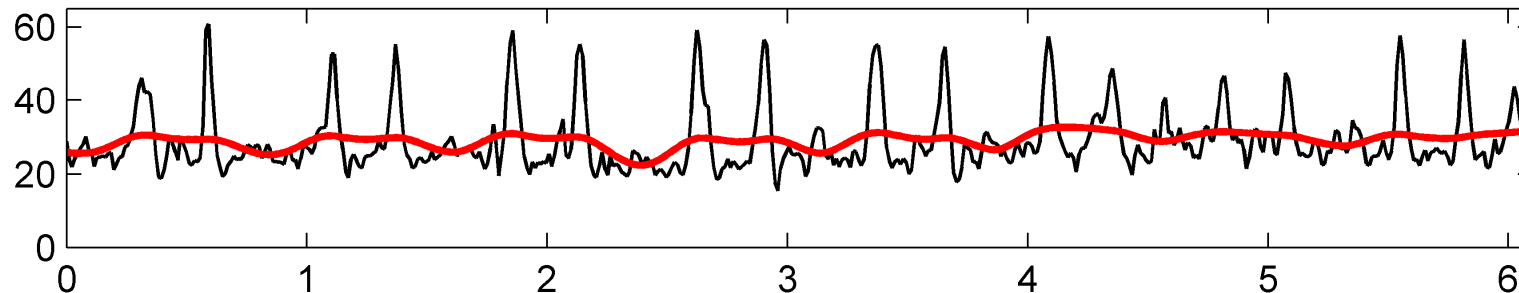
Onset Detection (Spectral-Based)

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation & half wave rectification
4. Accumulation
5. Normalization

Novelty curve

Substraction of local average

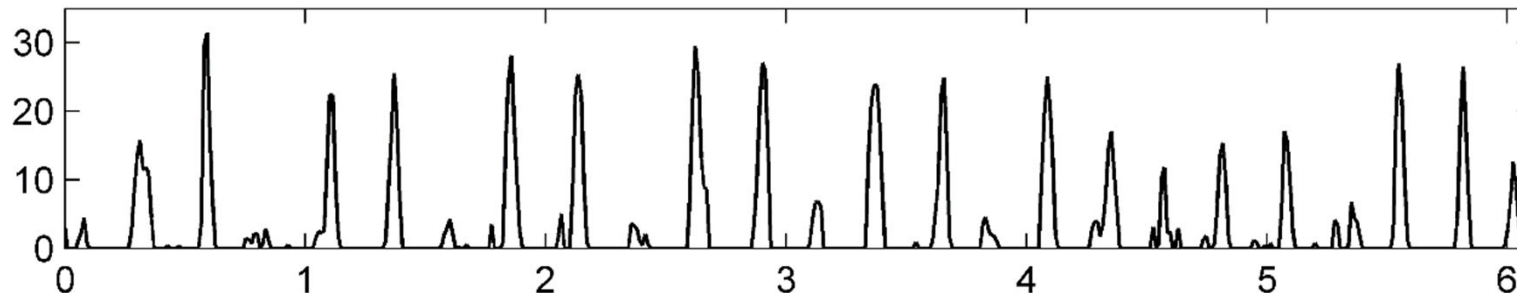


Onset Detection (Spectral-Based)

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation & half wave rectification
4. Accumulation
5. Normalization

Normalized novelty curve



Onset Detection (Spectral-Based)

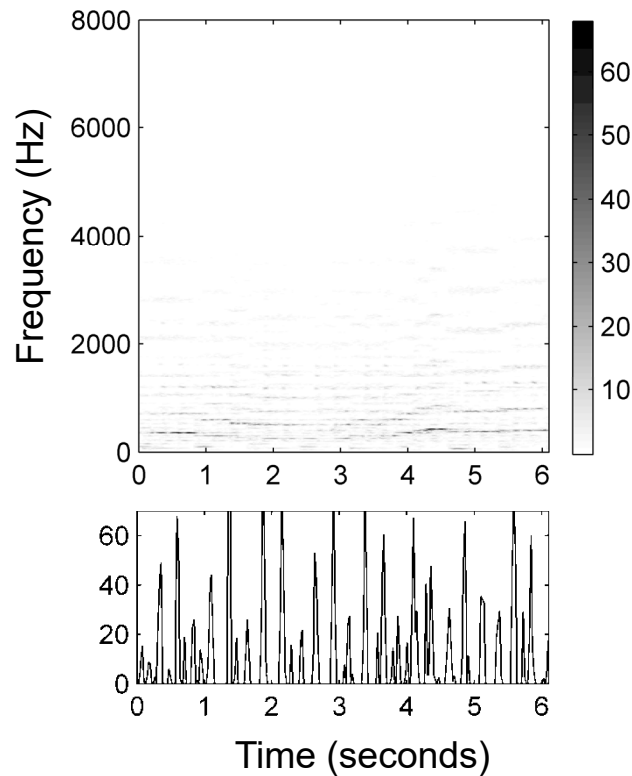
- Spectrogram $X = (X(t, k))_{t, k}$ $t \in [1 : T]$
 $k \in [1 : K]$
- Compressed Spectrogram $Y := \log(1 + C \cdot |X|)$ $C > 1$.
- Novelty curve $\Delta : [1 : T - 1] \rightarrow \mathbb{R}$:

$$\Delta(t) := \sum_{k=1}^K |Y(t+1, k) - Y(t, k)|_{\geq 0}$$

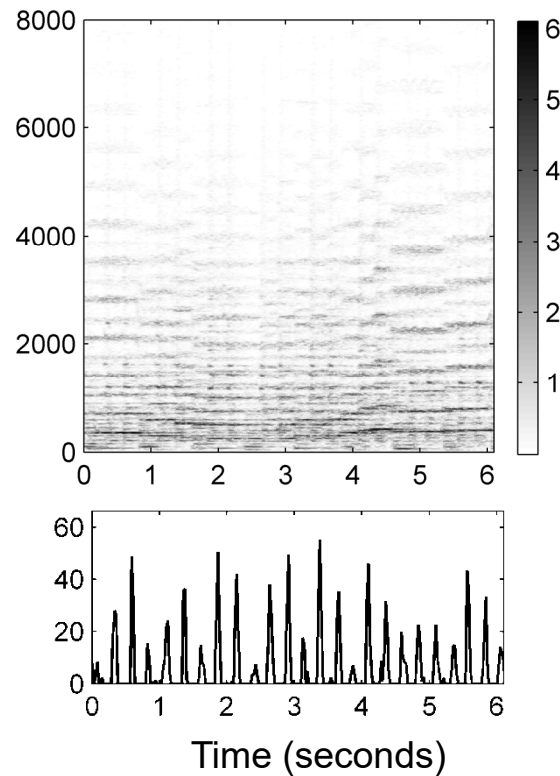
Logarithmic Compression

$$Y = \log(1 + C \cdot |X|)$$

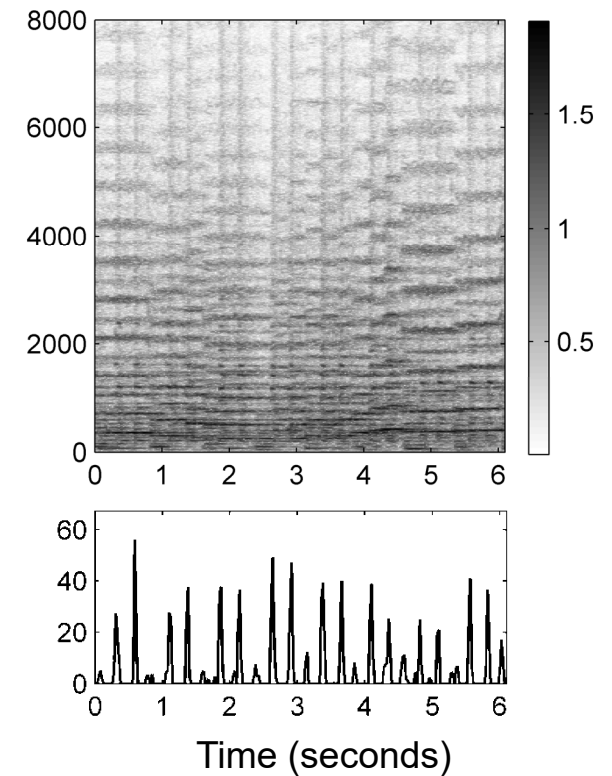
No compression



$C = 1$

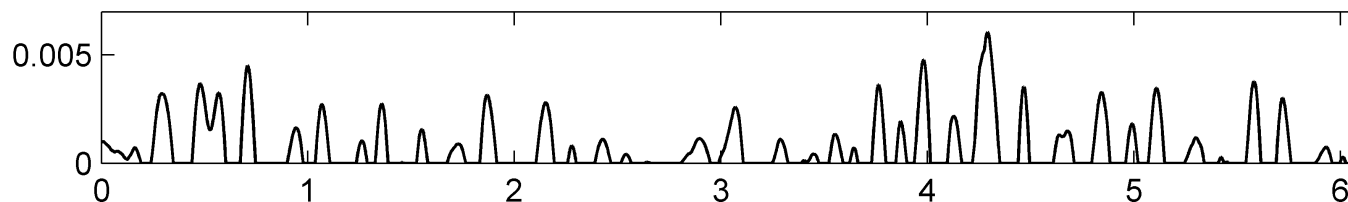


$C = 100$

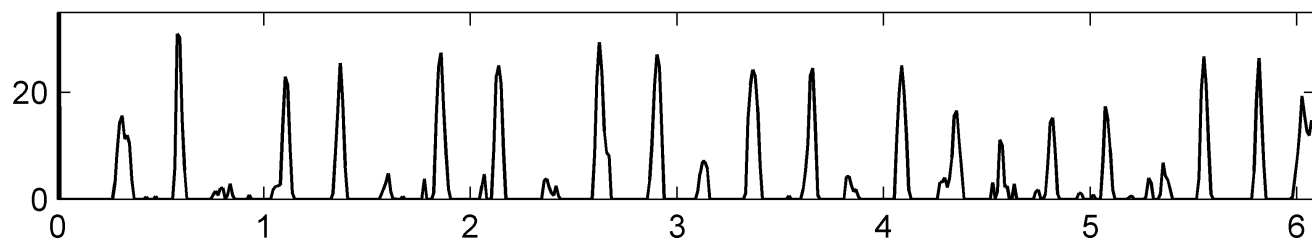


Onset Detection

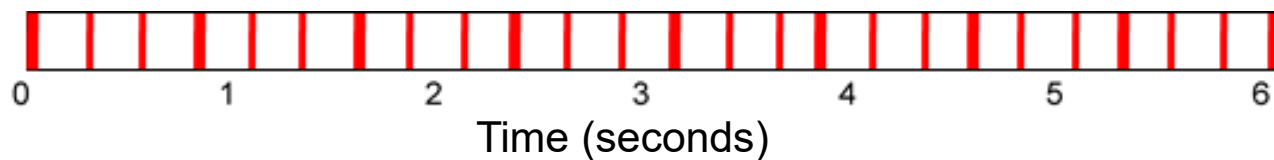
Energy-based novelty curve



Spectral-based novelty curve

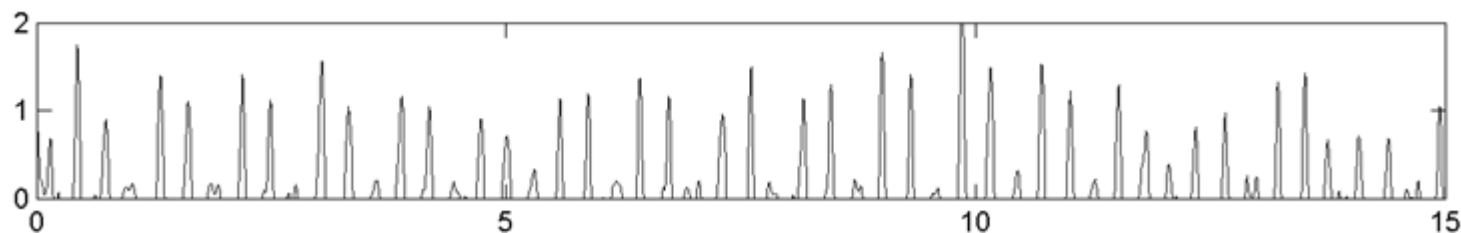


Manual onset annotations

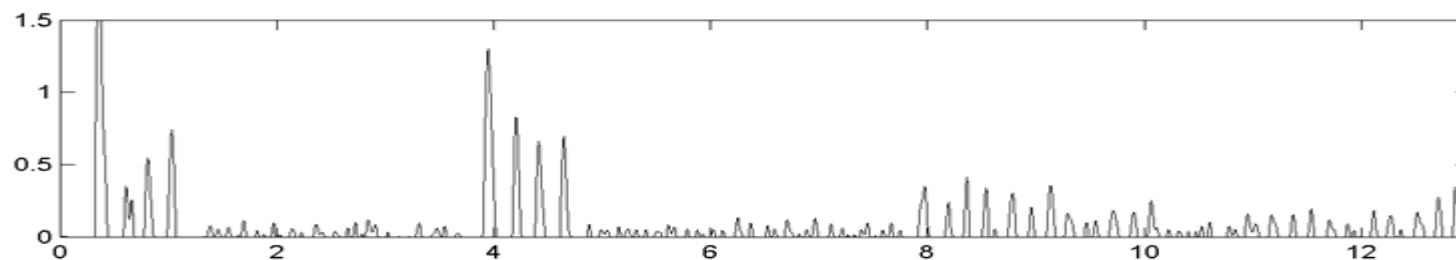


Onset Detection

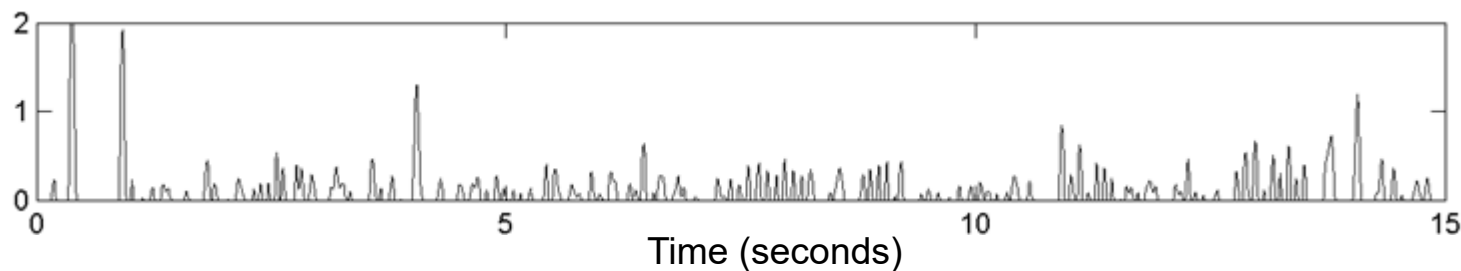
Shostakovich – 2nd Waltz



Beethoven – Fifth Symphony



Borodin – String Quartet No. 2



Onset Detection

Drumbeat



Going Home



Lyphard melodie



Por una cabeza



Donau



Beat and Tempo

What is a beat?

- Steady pulse that drives music forward and provides the temporal framework of a piece of music
- Sequence of perceived pulses that are equally spaced in time
- The pulse a human taps along when listening to the music

[Parncutt 1994]

[Sethares 2007]

[Large/Palmer 2002]

[Lerdahl/ Jackendoff 1983]

[Fitch/ Rosenfeld 2007]

The term **tempo** then refers to the speed of the pulse.

Beat and Tempo

Strategy

- Analyze the novelty curve with respect to reoccurring or quasi-periodic patterns
- Avoid the explicit determination of note onsets (no peak picking)

Beat and Tempo

Strategy

- Analyze the novelty curve with respect to reoccurring or quasi-periodic patterns
- Avoid the explicit determination of note onsets (no peak picking)

[Scheirer, JASA 1998]

Methods

- Comb-filter methods
- Autocorrelation
- Fourier transform

[Ellis, JNMR 2007]

[Davies/Plumbley, IEEE-TASLP 2007]

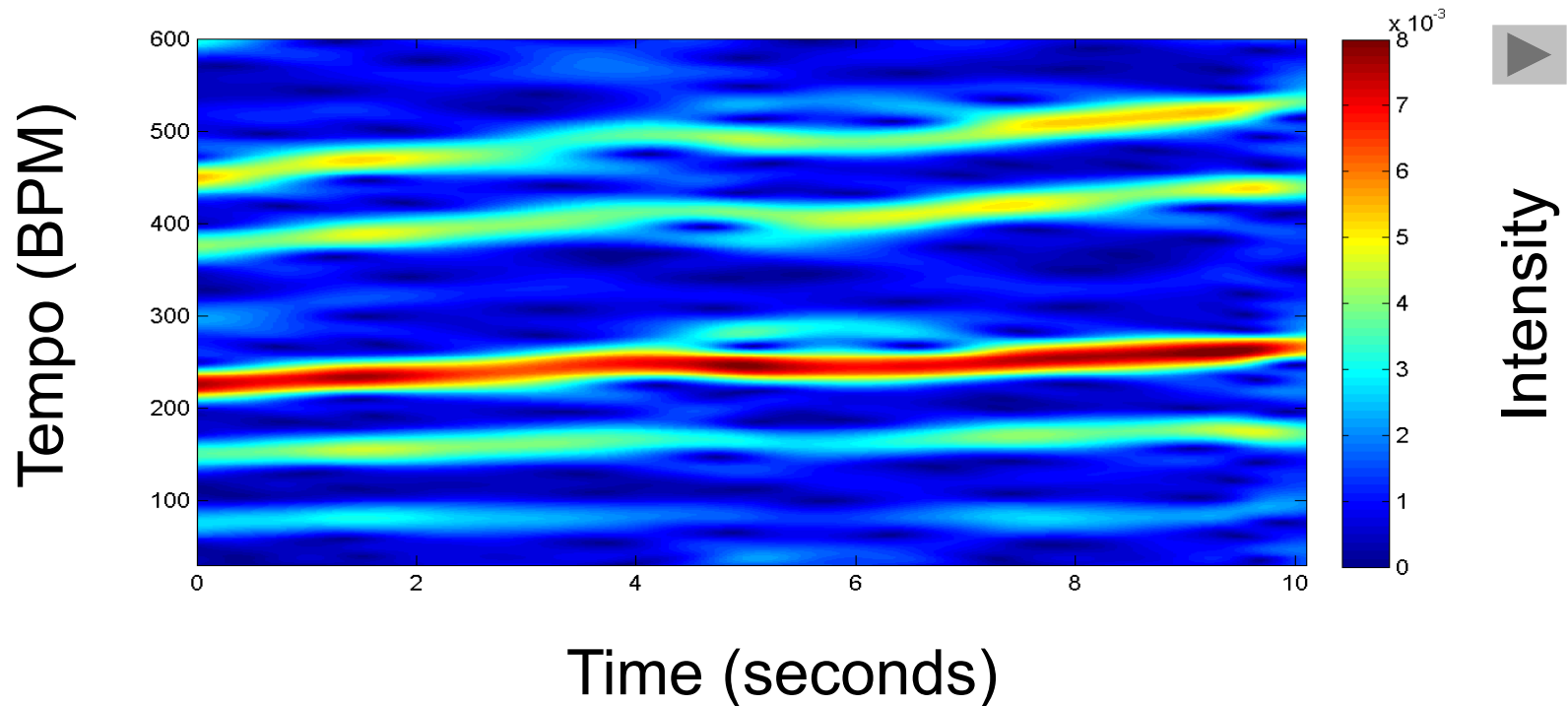
[Peeters, JASP 2007]

[Grosche/Müller, ISMIR 2009]

[Grosche/Müller, IEEE-TASLP 2011]

Tempogram

Definition: A **tempogram** is a time-tempo representation that encodes the local tempo of a music signal over time.



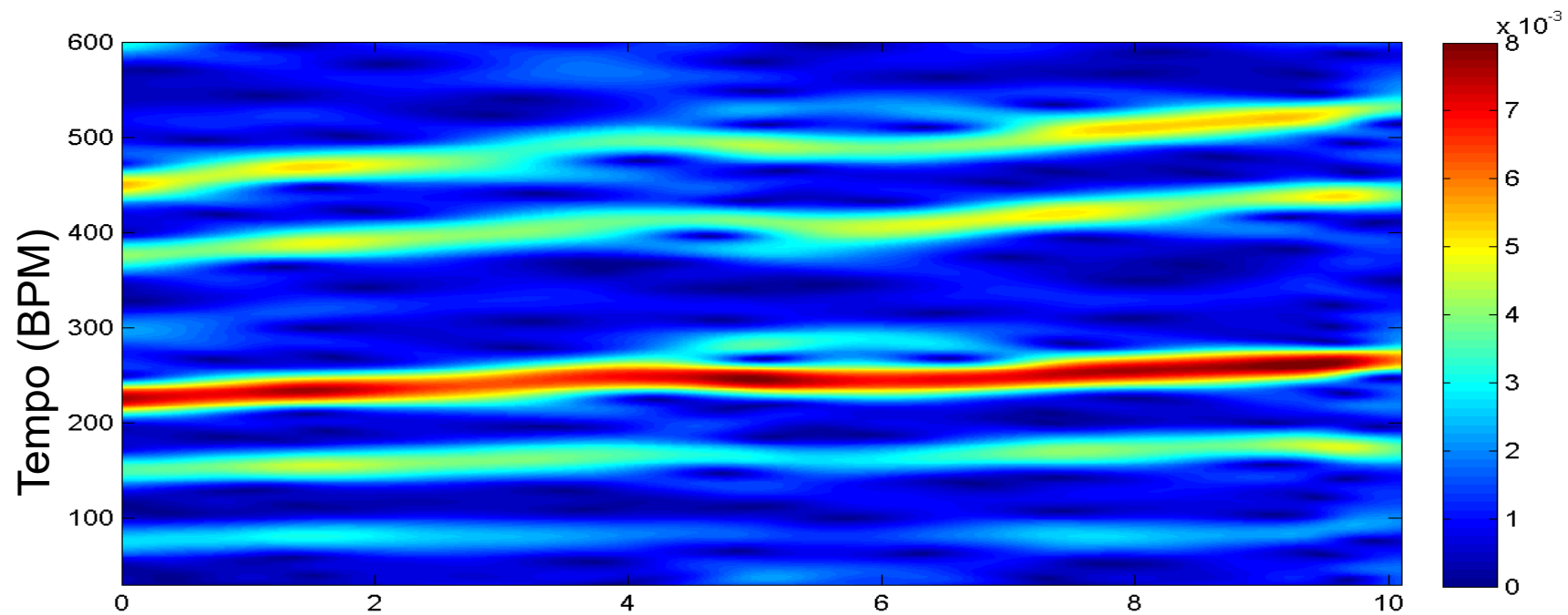
Tempogram (Fourier)

Definition: A **tempogram** is a time-tempo representation that encodes the local tempo of a music signal over time.

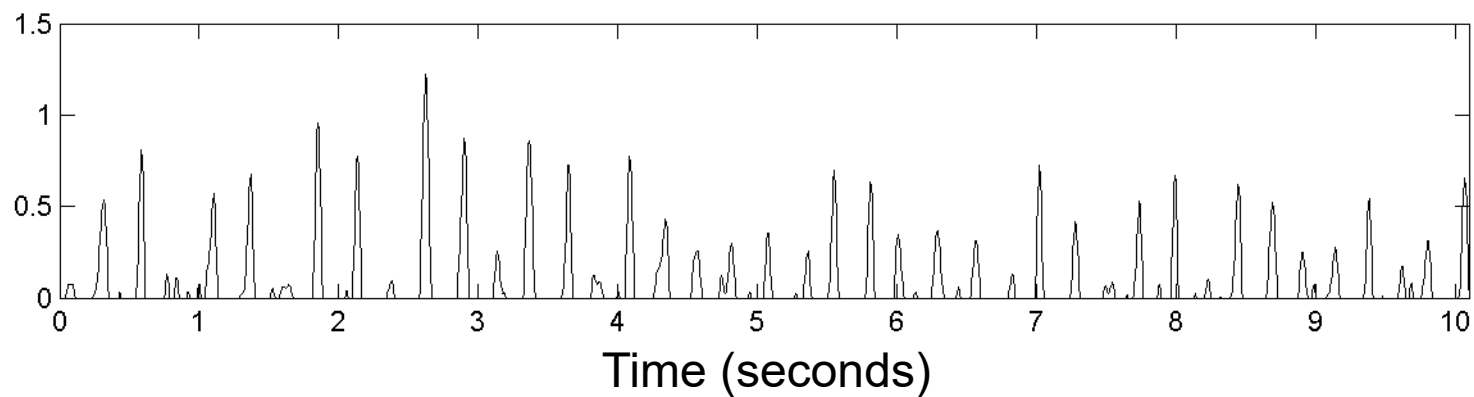
Fourier-based method

- Compute a spectrogram (STFT) of the novelty curve
- Convert frequency axis (given in Hertz) into tempo axis (given in BPM)
- Magnitude spectrogram indicates local tempo

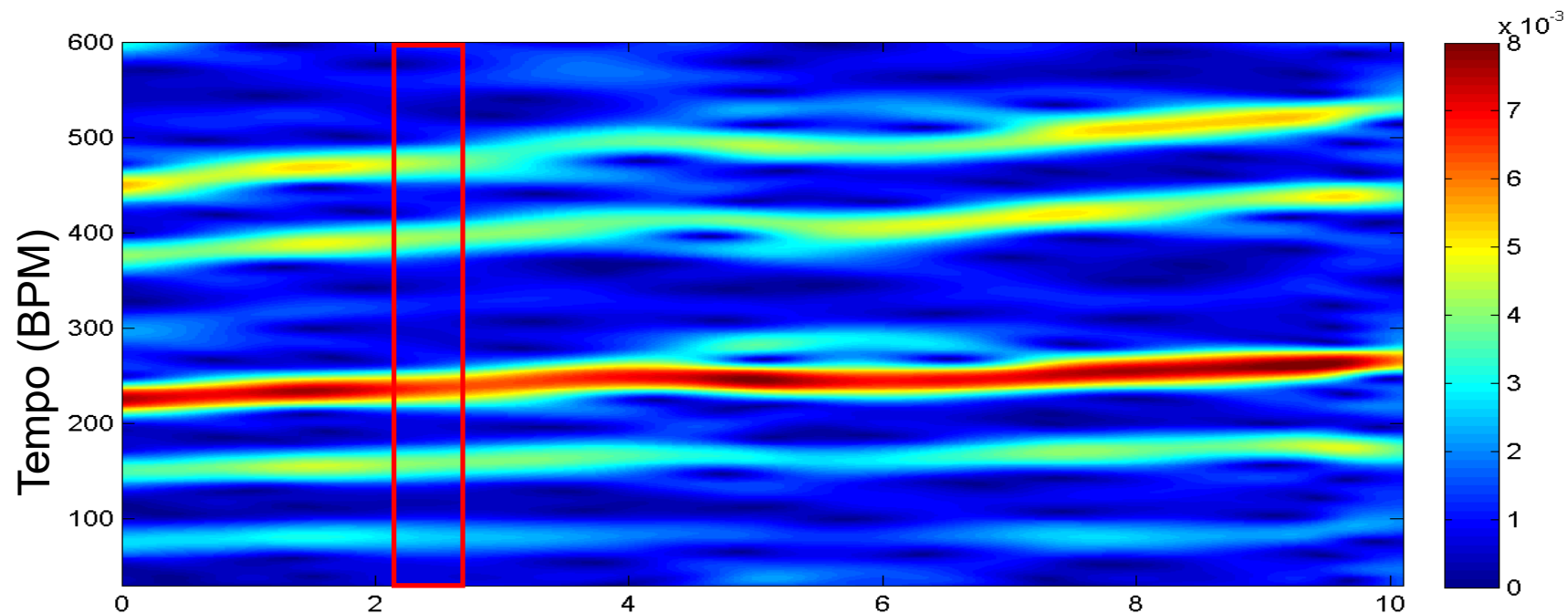
Tempogram (Fourier)



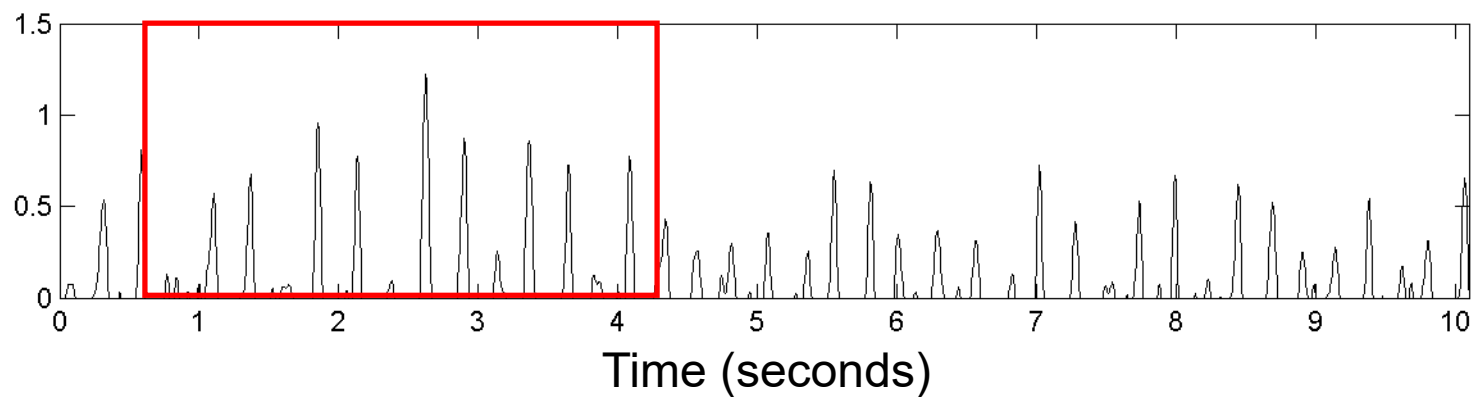
Novelty curve



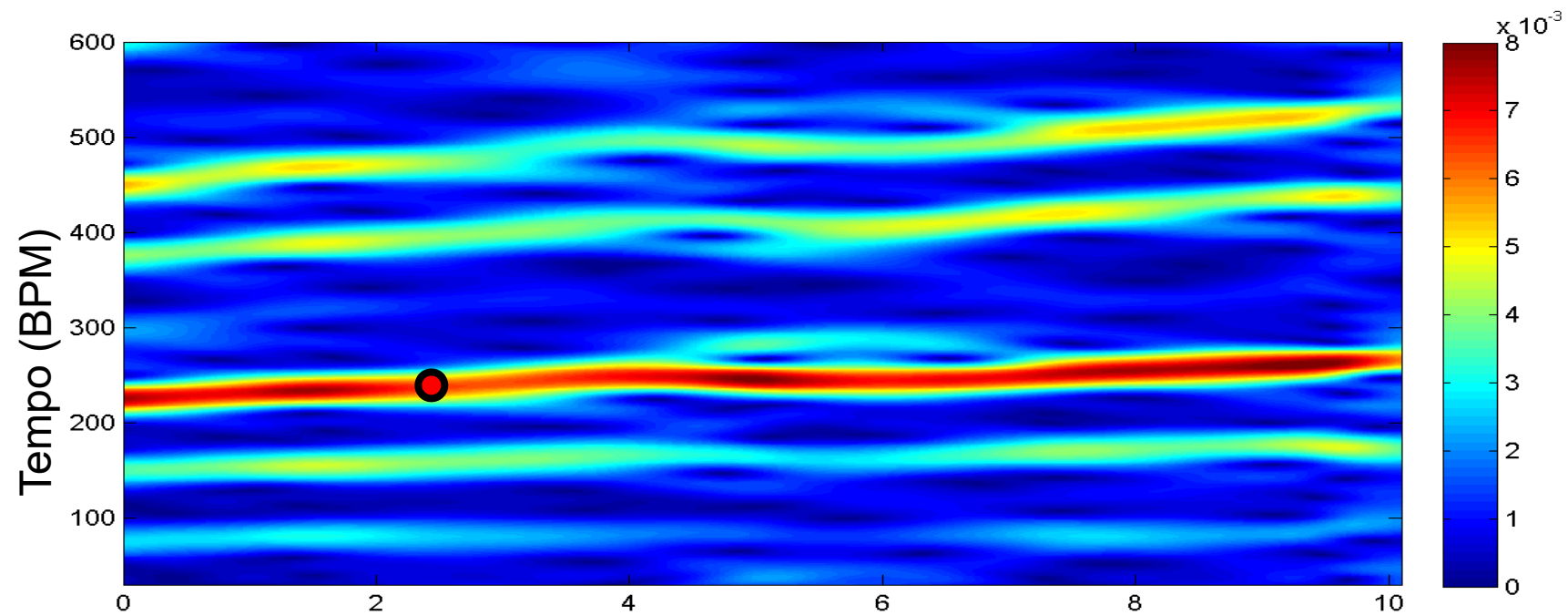
Tempogram (Fourier)



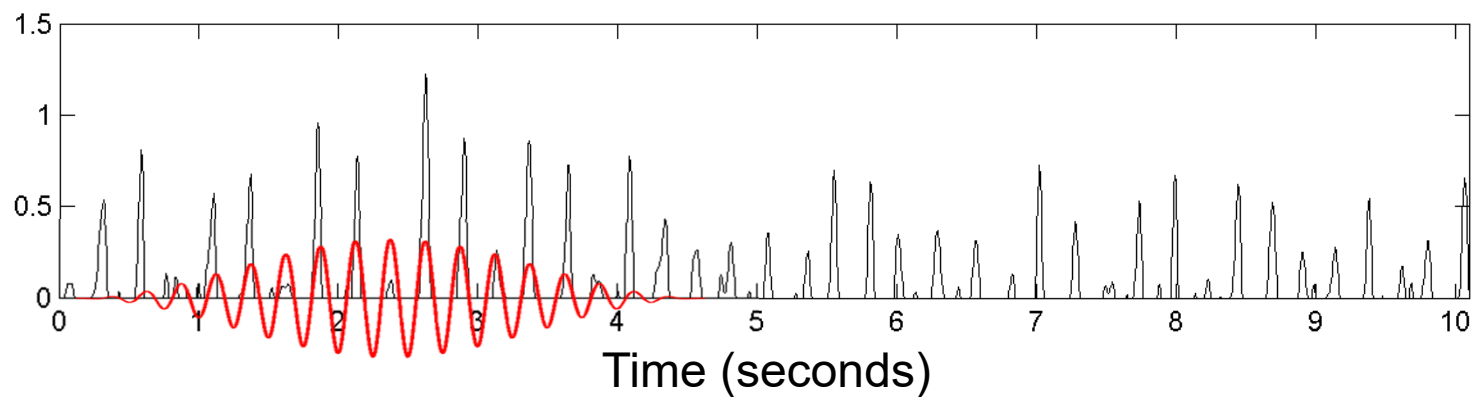
Novelty curve (local section)



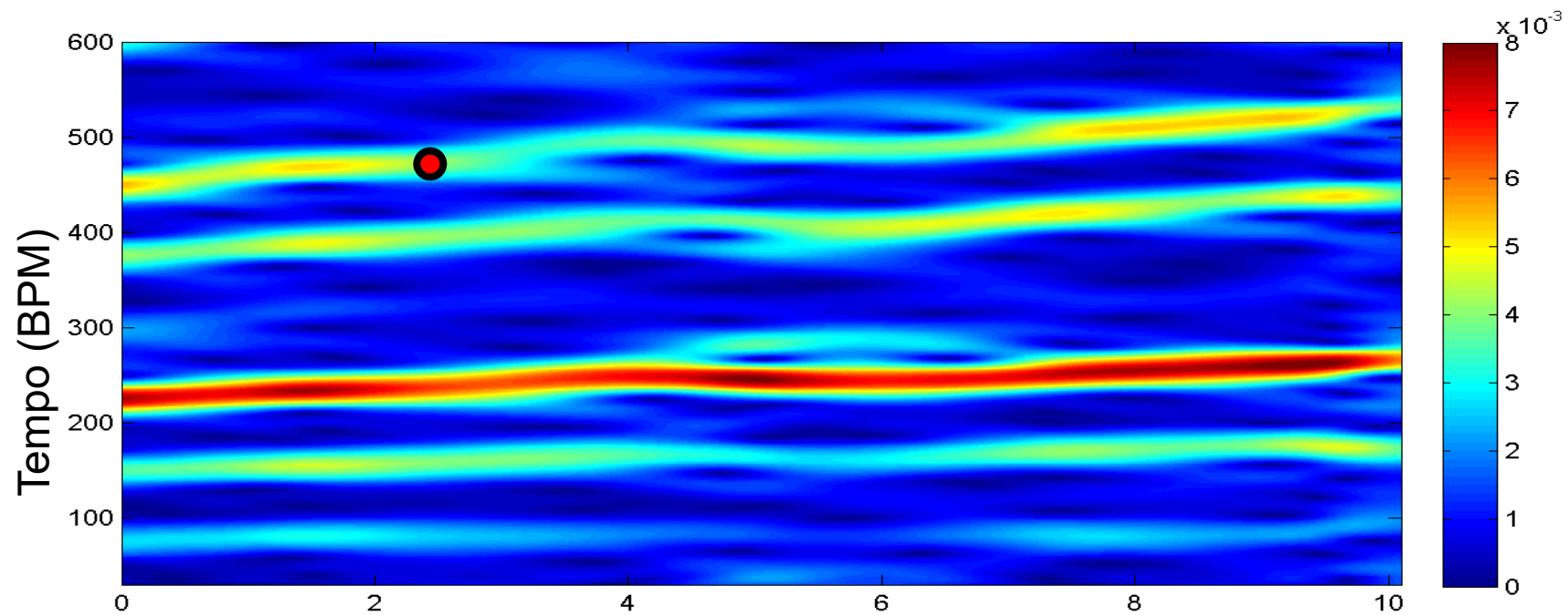
Tempogram (Fourier)



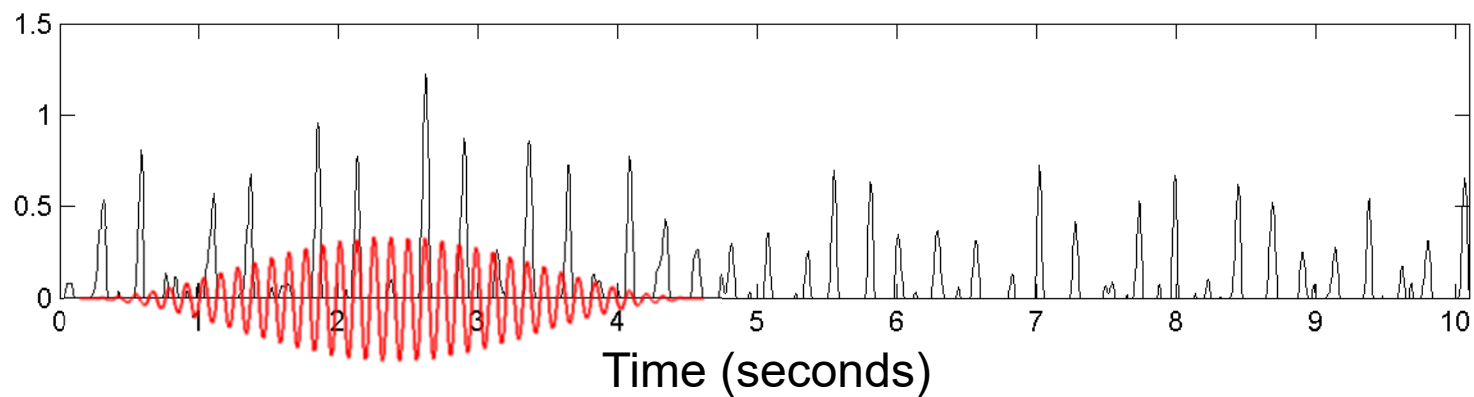
Windowed sinusoidal



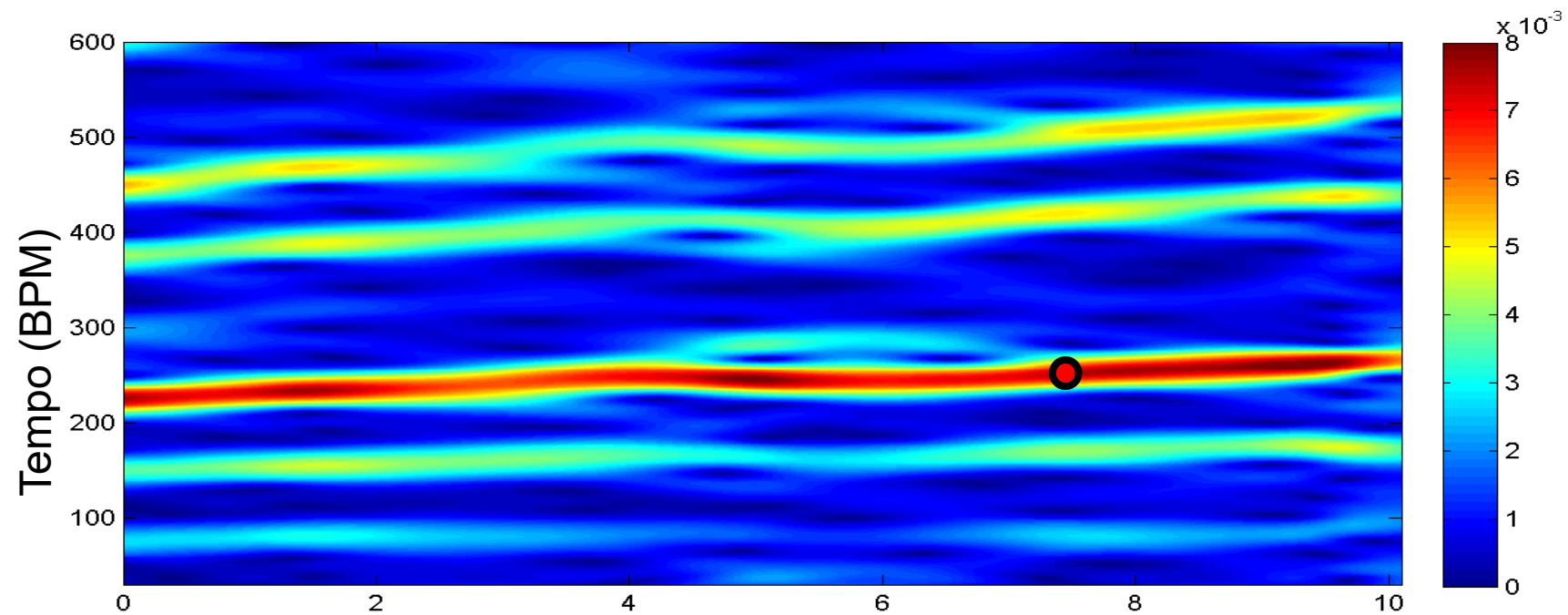
Tempogram (Fourier)



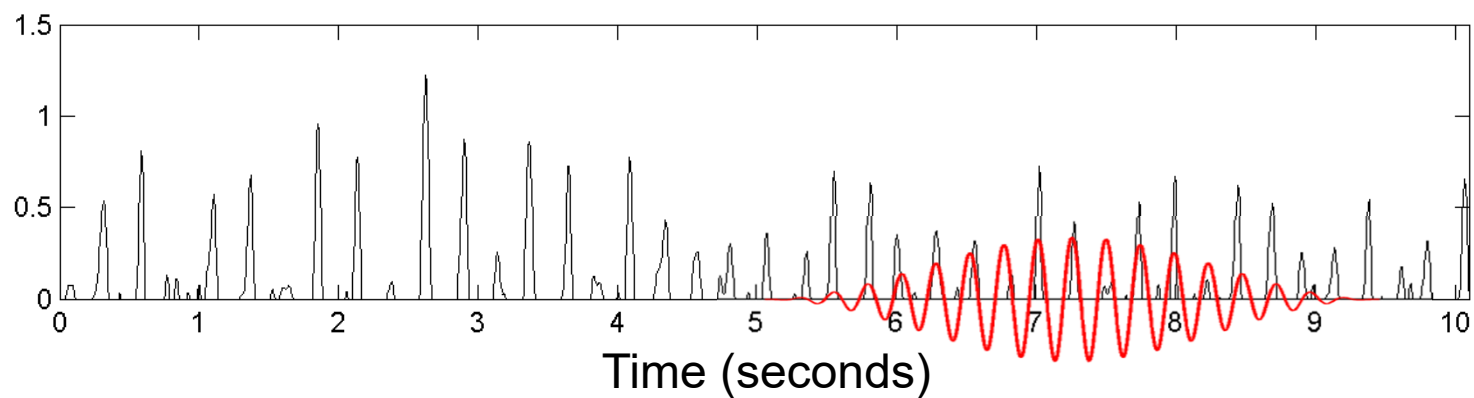
Windowed sinusoidal



Tempogram (Fourier)



Windowed sinusoidal



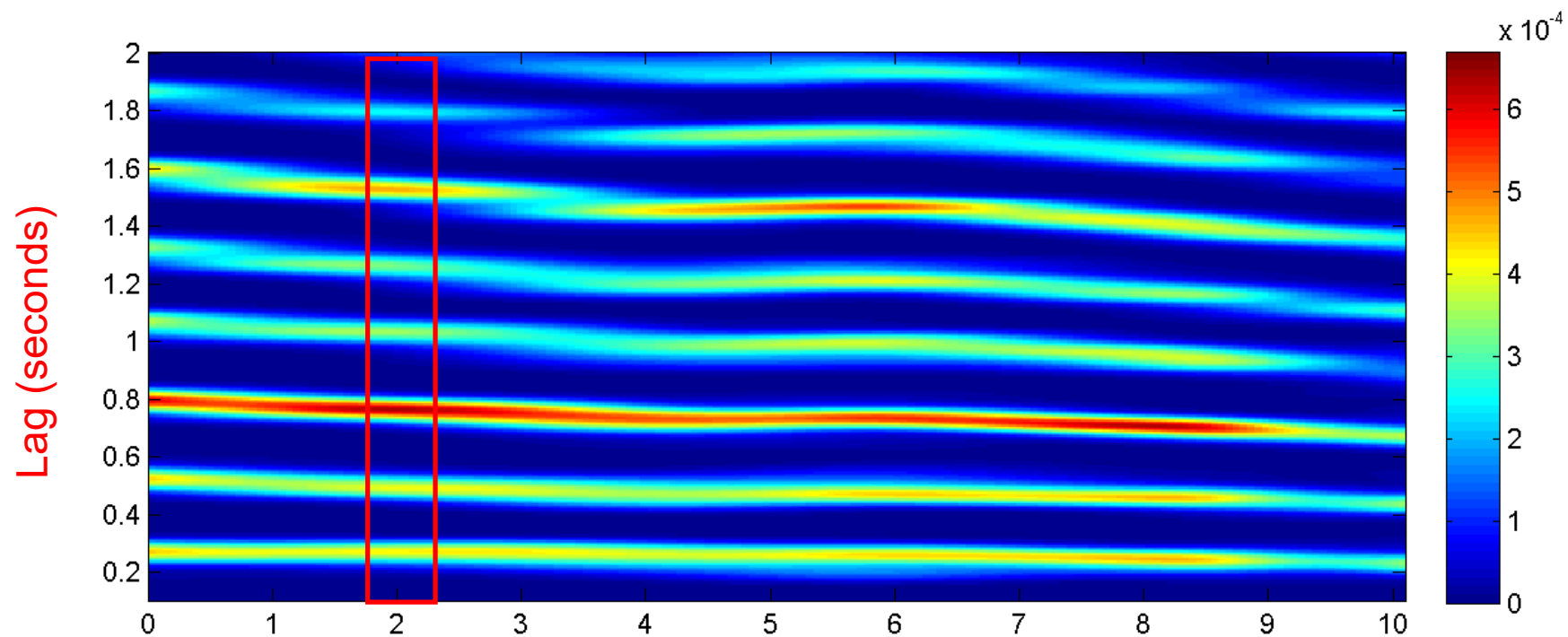
Tempogram (Autocorrelation)

Definition: A **tempogram** is a time-tempo representation that encodes the local tempo of a music signal over time.

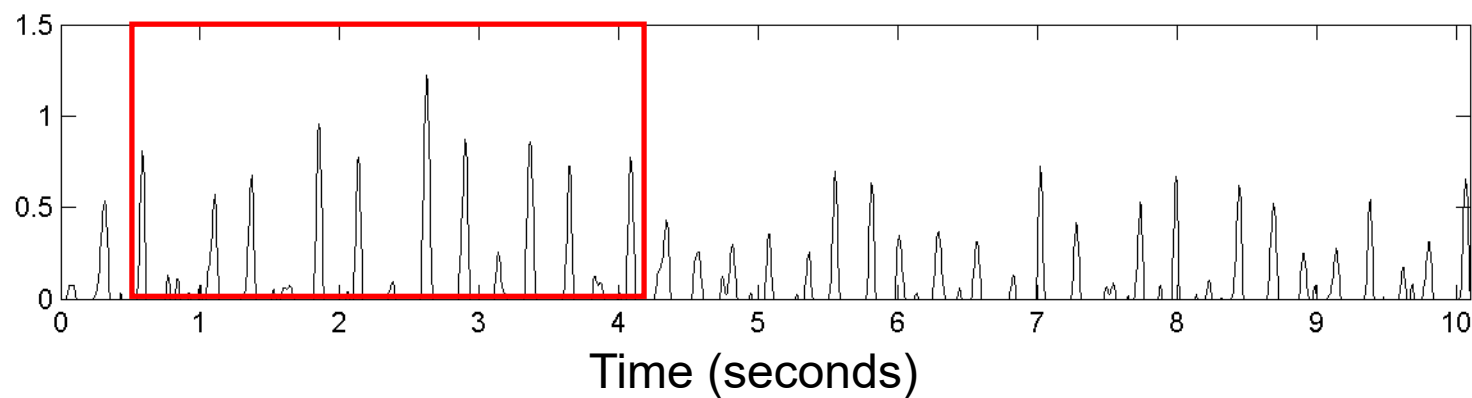
Autocorrelation-based method

- Compare novelty curve with time-lagged local sections of itself
- Convert lag-axis (given in seconds) into tempo axis (given in BPM)
- Autocorrelogram indicates local tempo

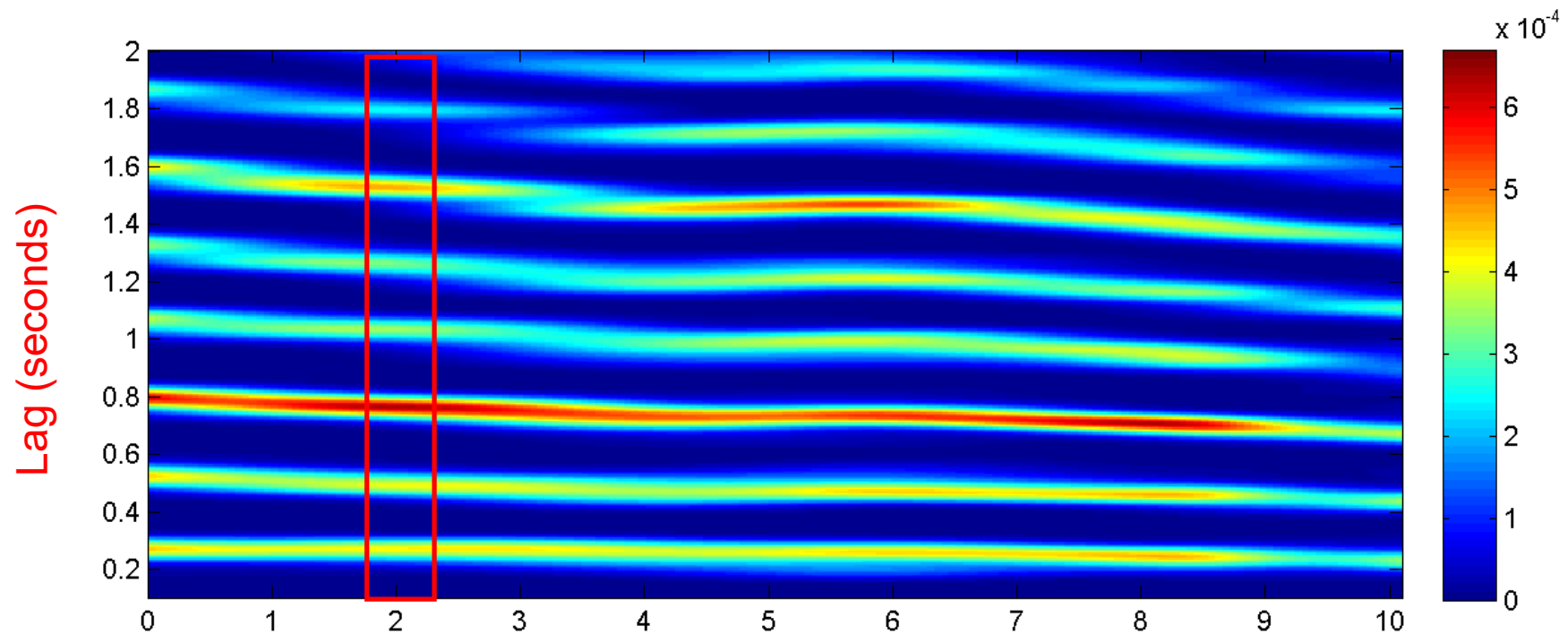
Tempogram (Autocorrelation)



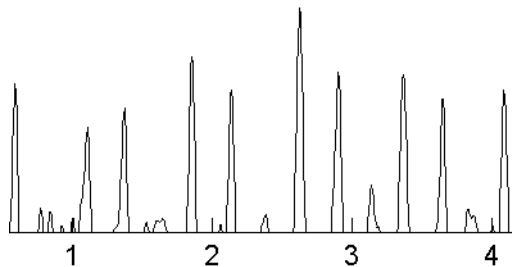
Novelty curve (local section)



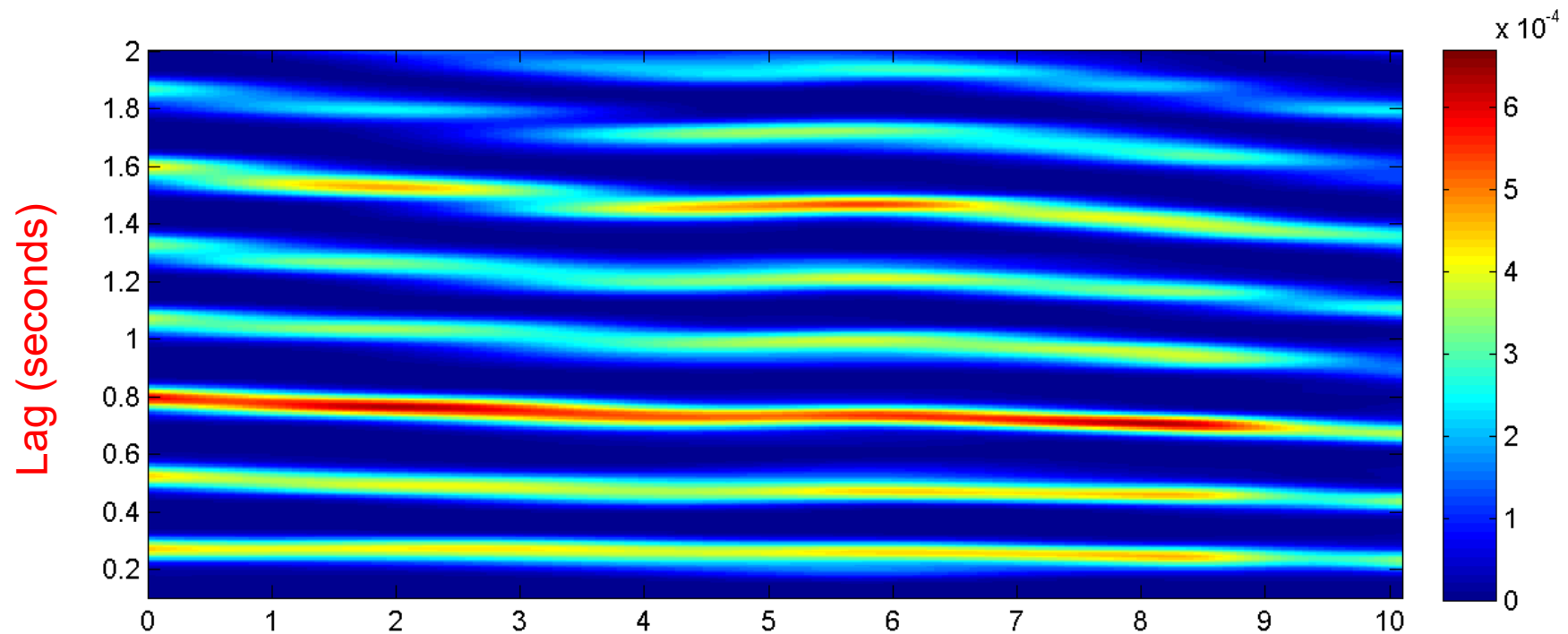
Tempogram (Autocorrelation)



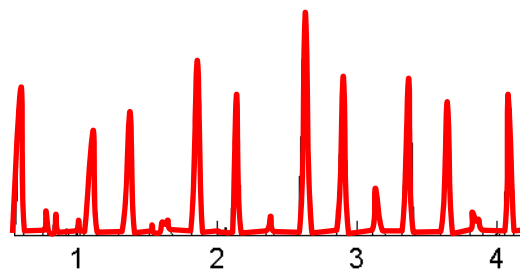
Windowed autocorrelation



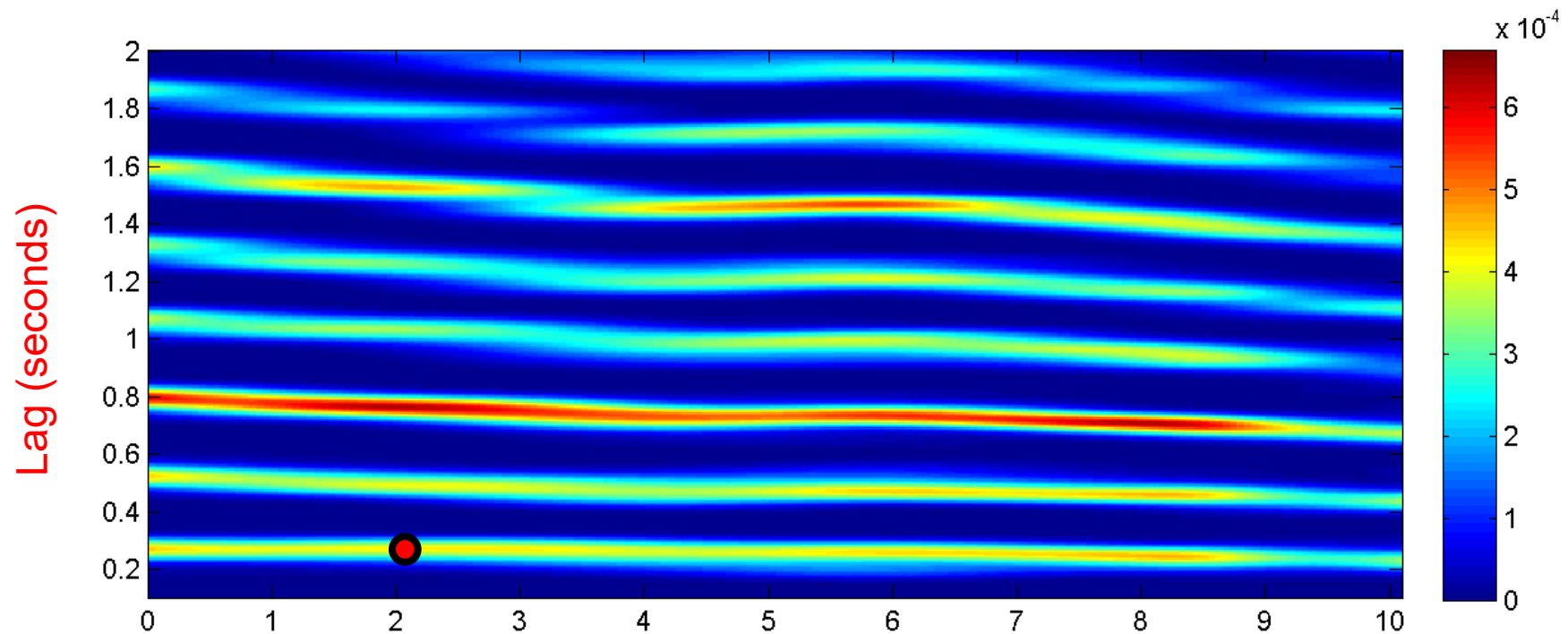
Tempogram (Autocorrelation)



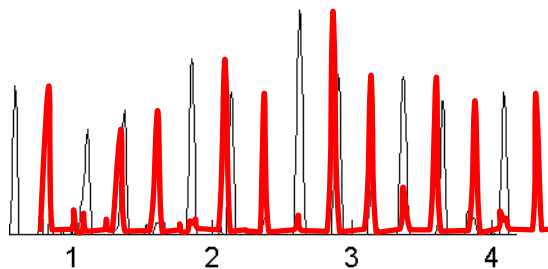
Lag = 0 (seconds)



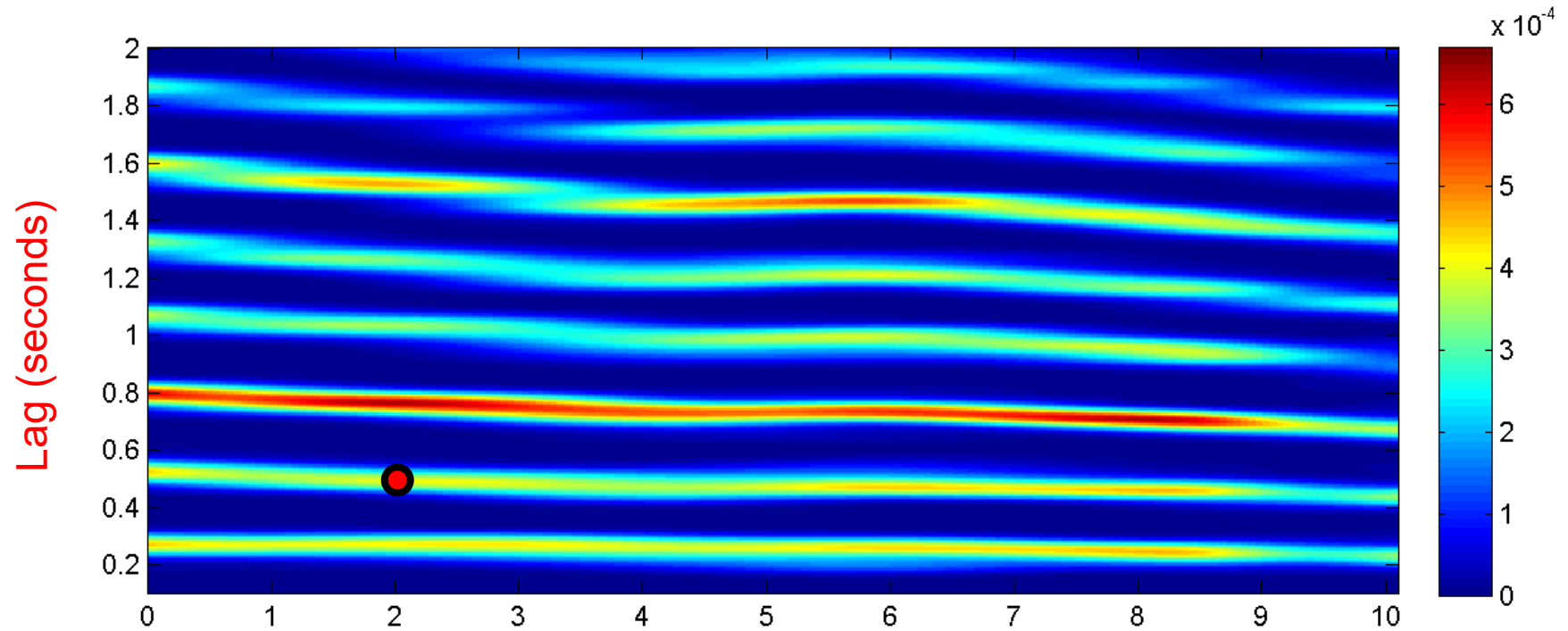
Tempogram (Autocorrelation)



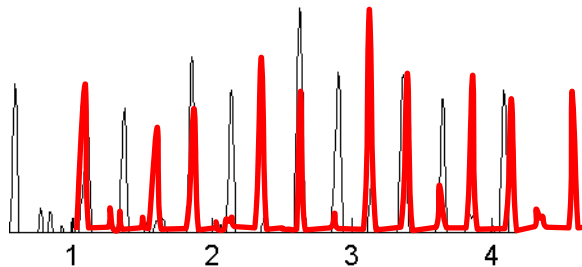
Lag = 0.26 (seconds)



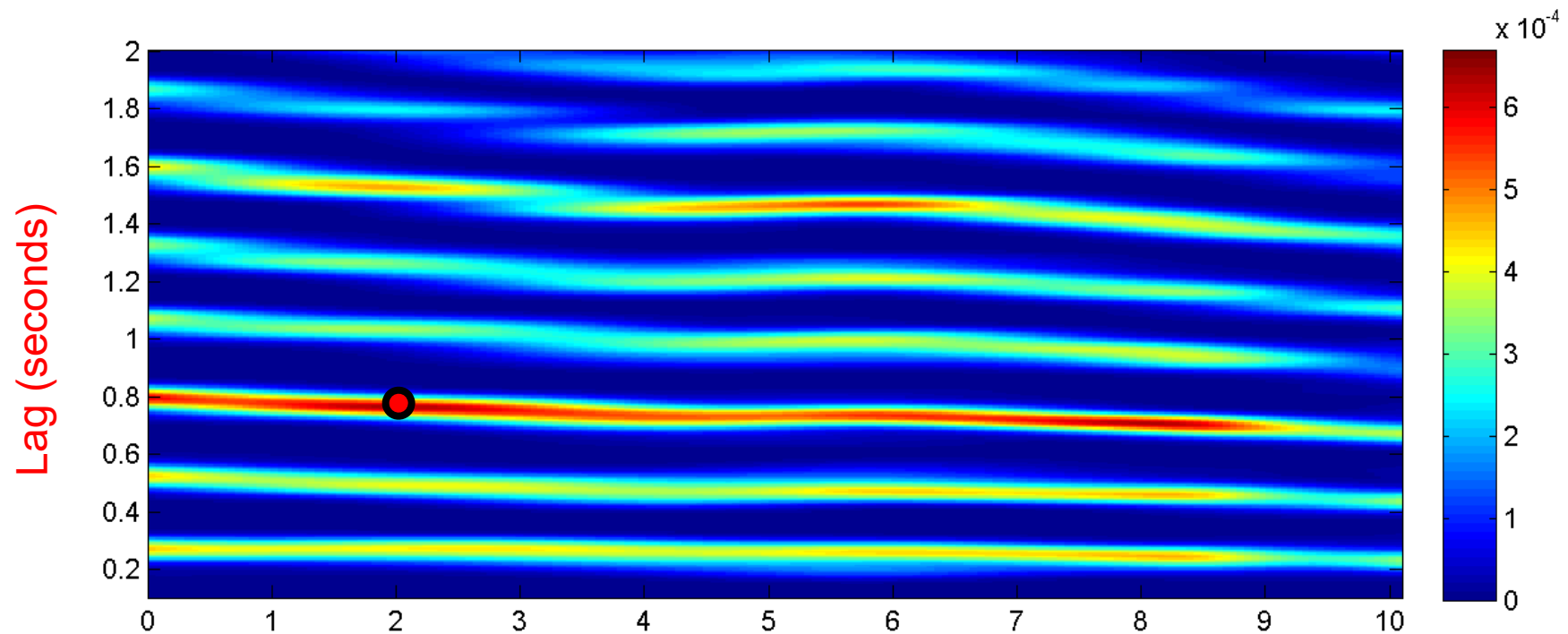
Tempogram (Autocorrelation)



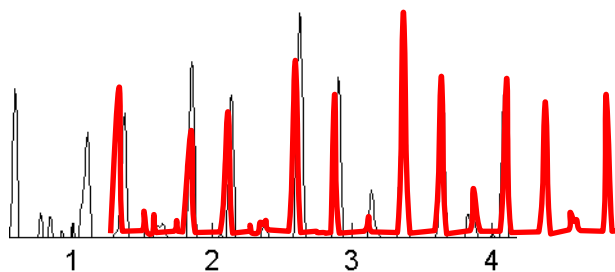
Lag = 0.52 (seconds)



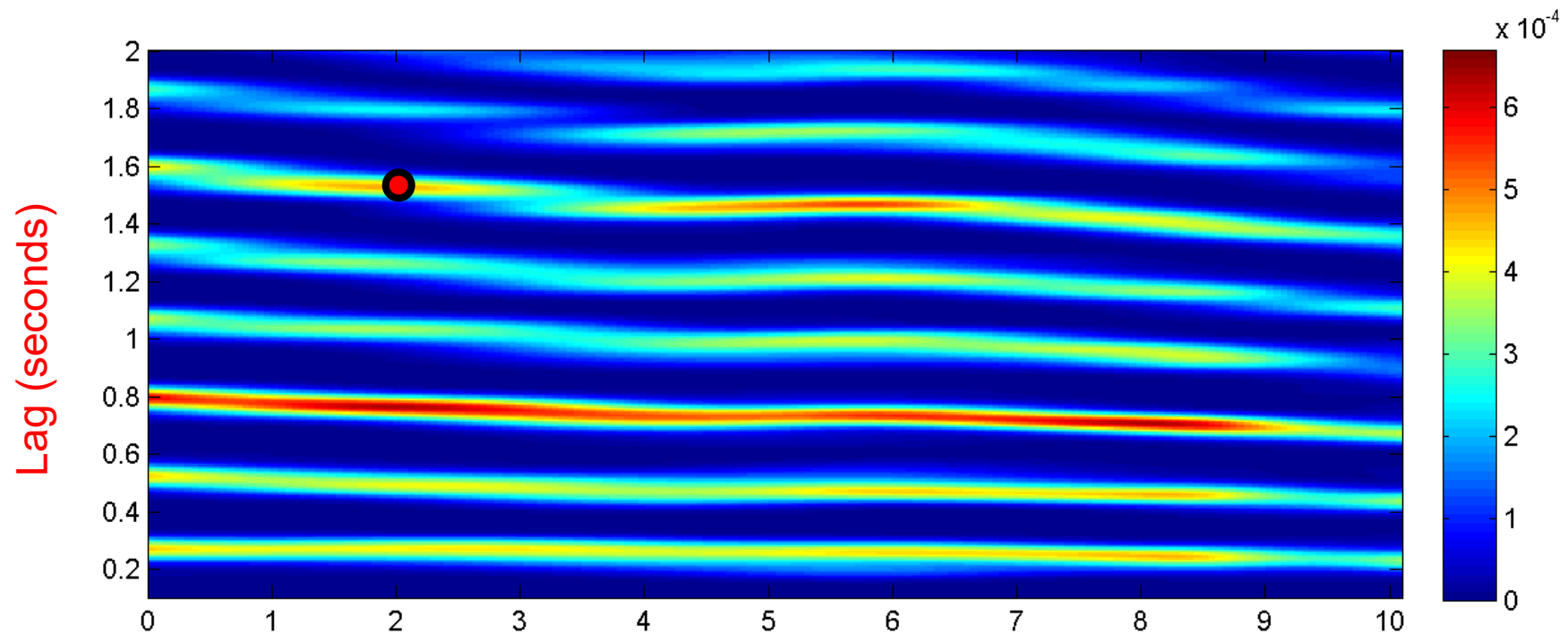
Tempogram (Autocorrelation)



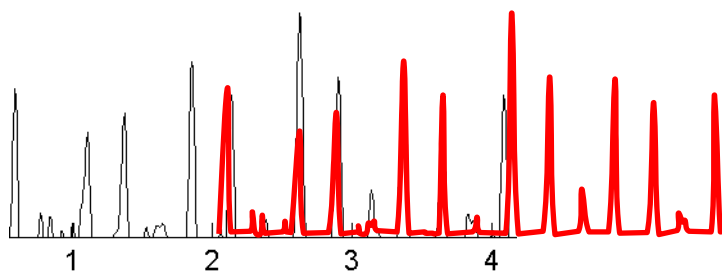
Lag = 0.78 (seconds)



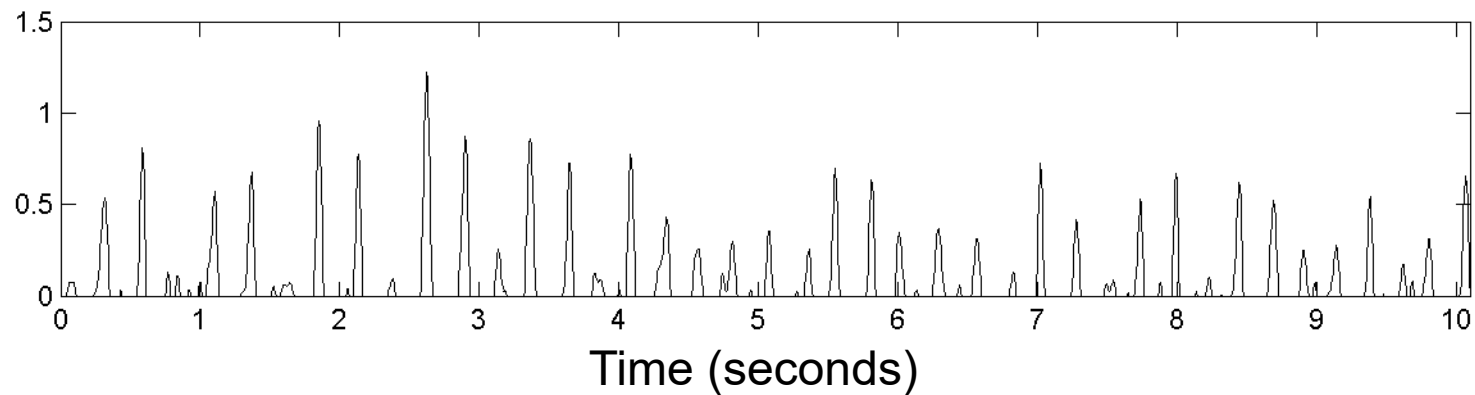
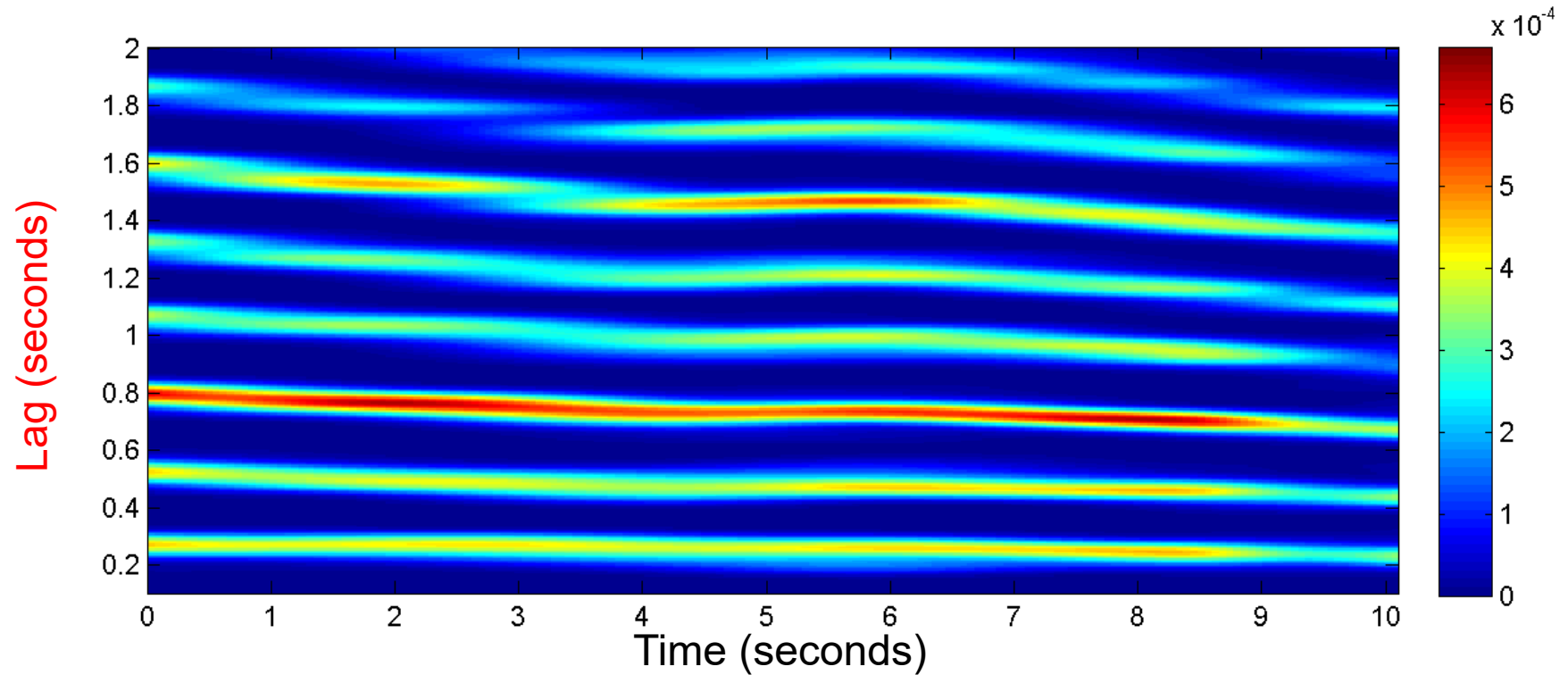
Tempogram (Autocorrelation)



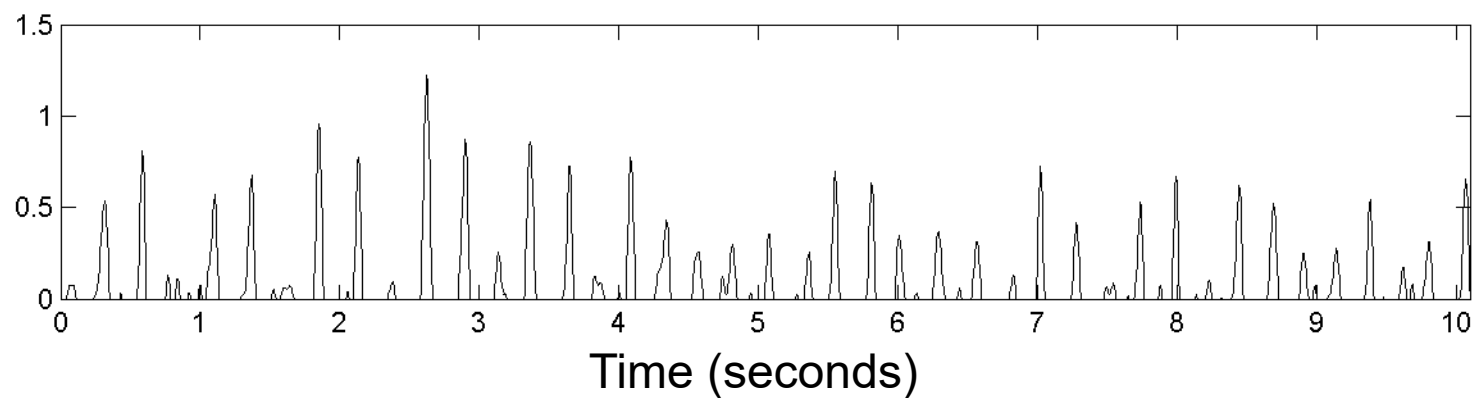
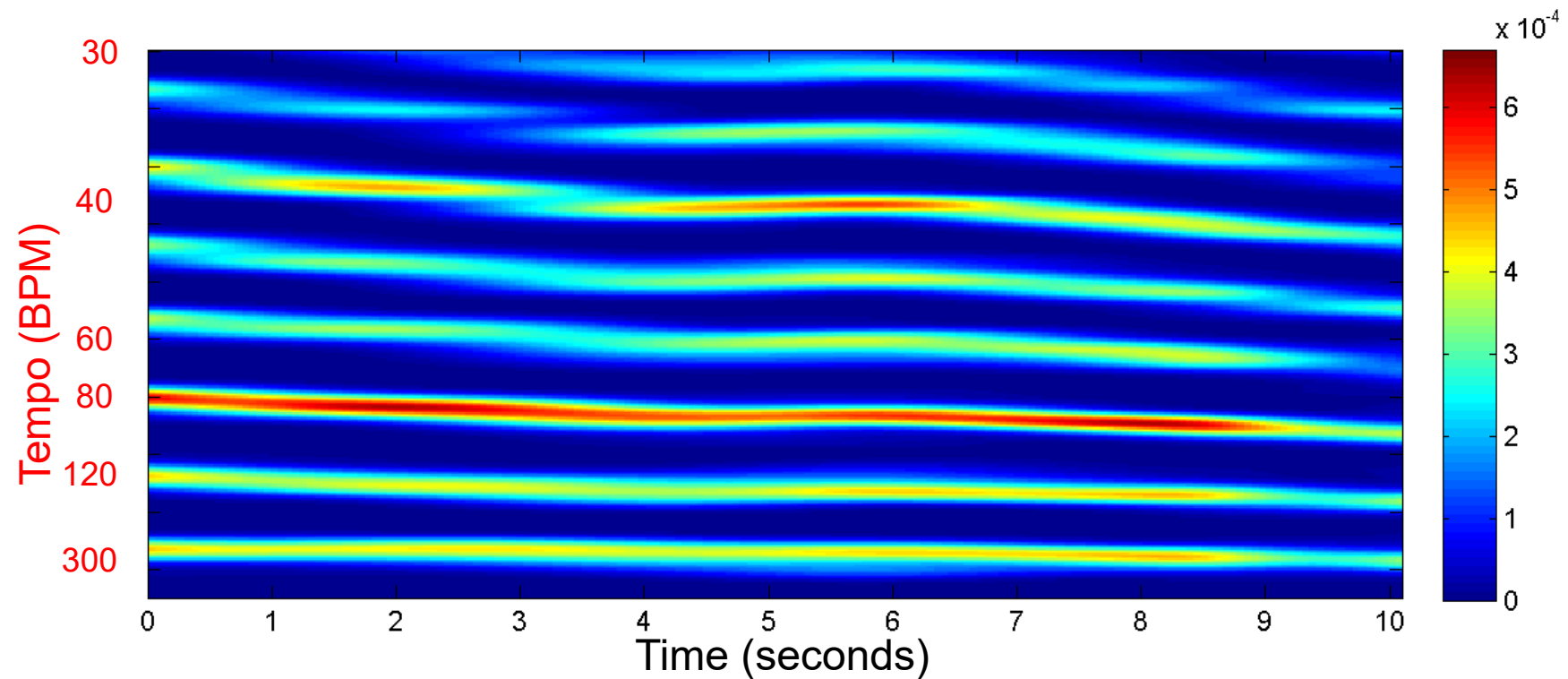
Lag = 1.56 (seconds)



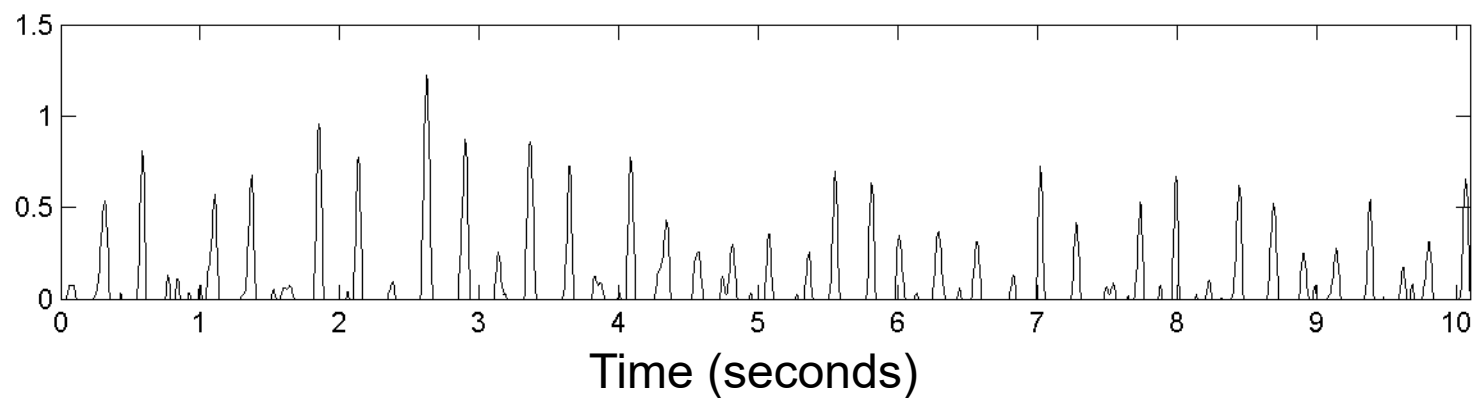
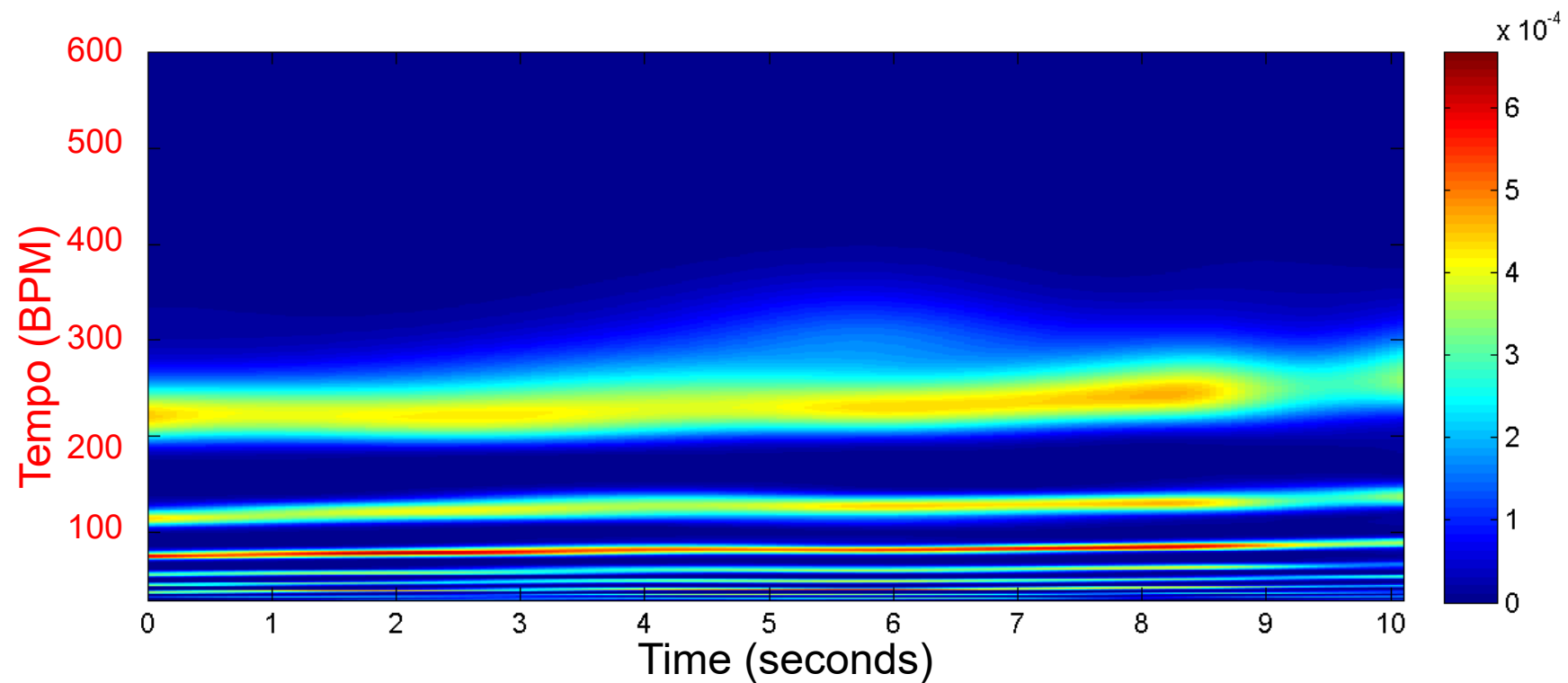
Tempogram (Autocorrelation)



Tempogram (Autocorrelation)

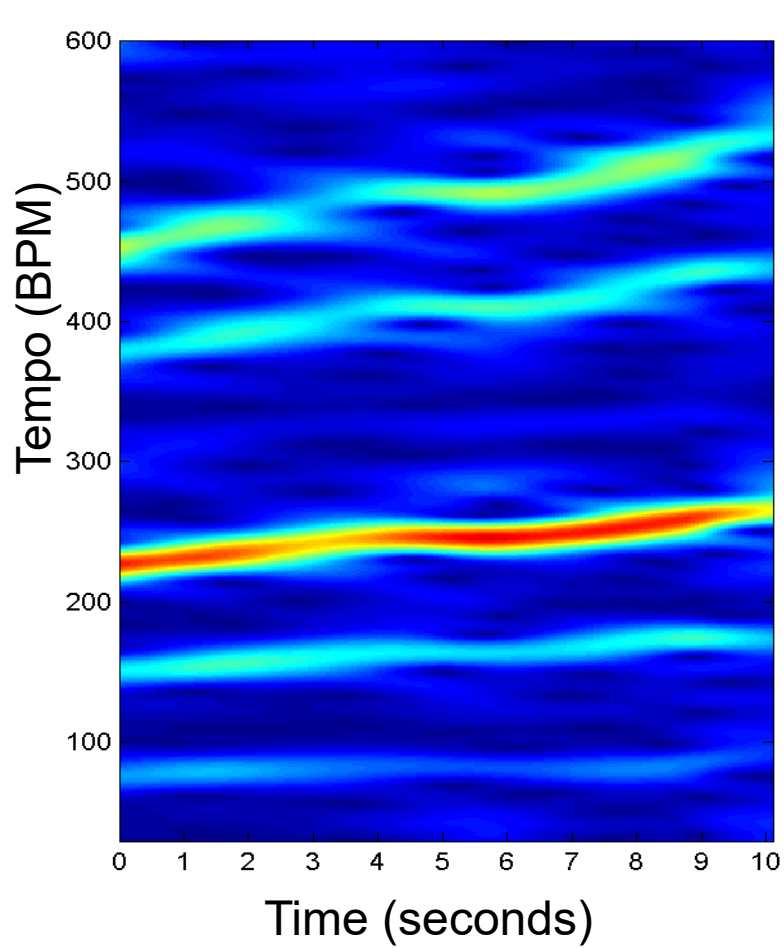


Tempogram (Autocorrelation)

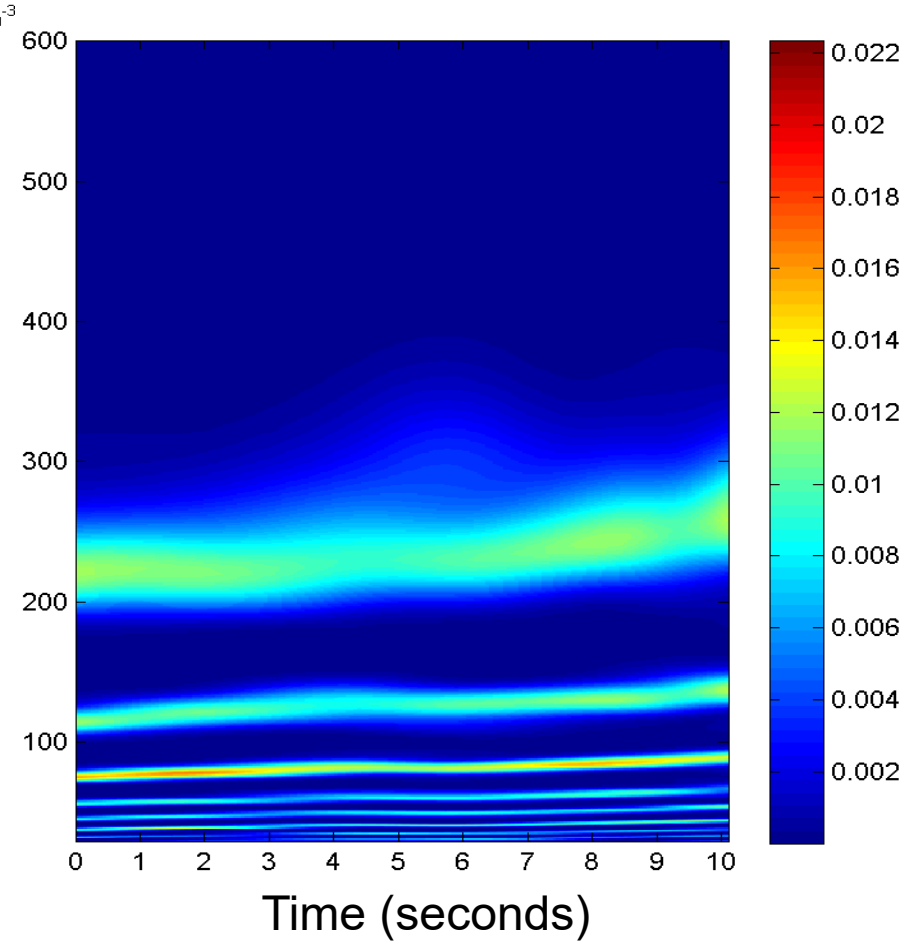


Tempogram

Fourier

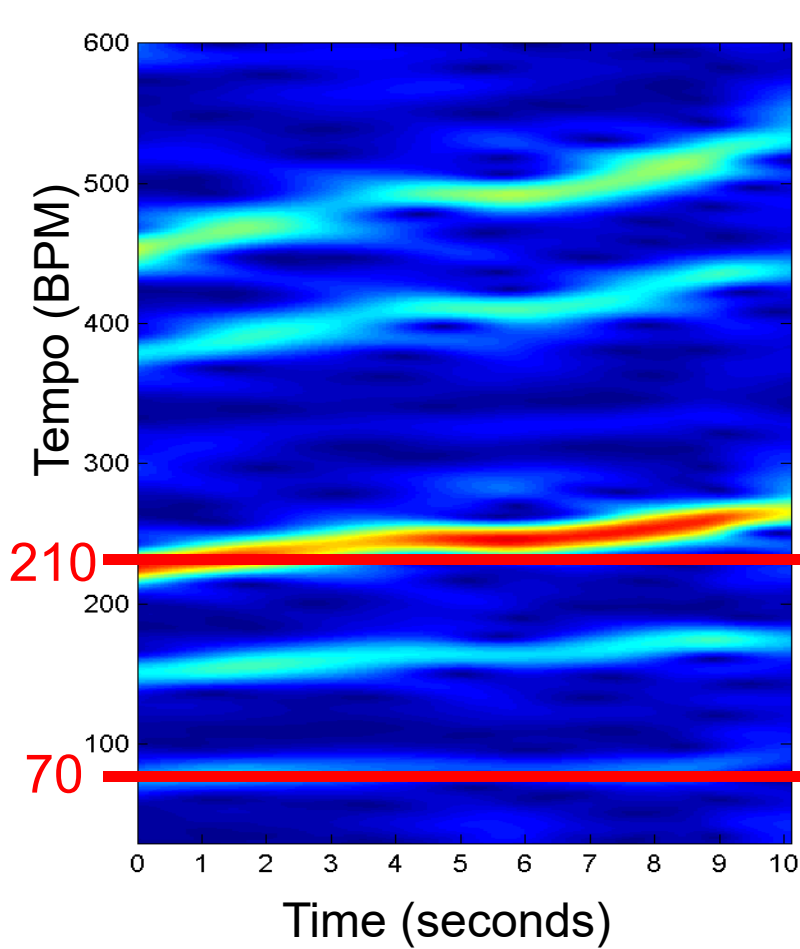


Autocorrelation

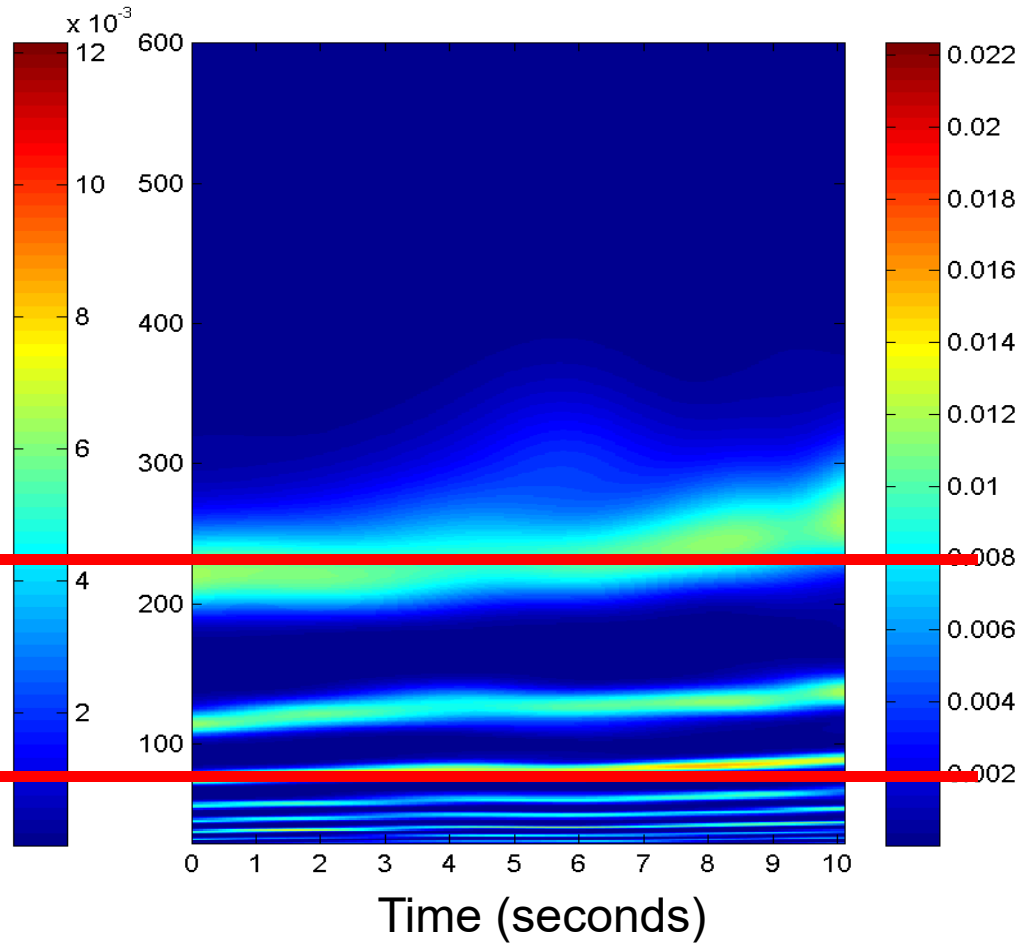


Tempogram

Fourier



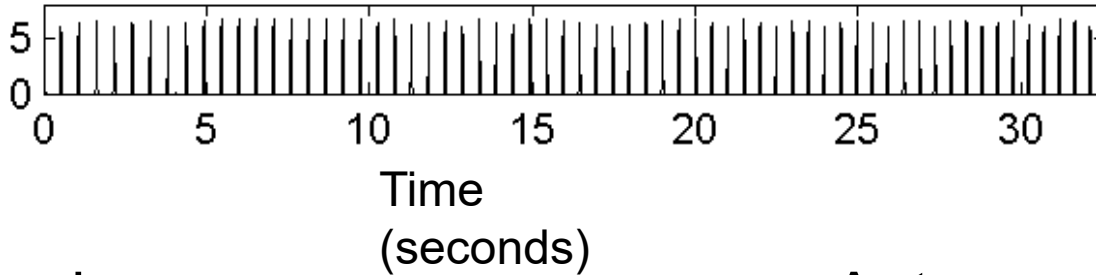
Autocorrelation



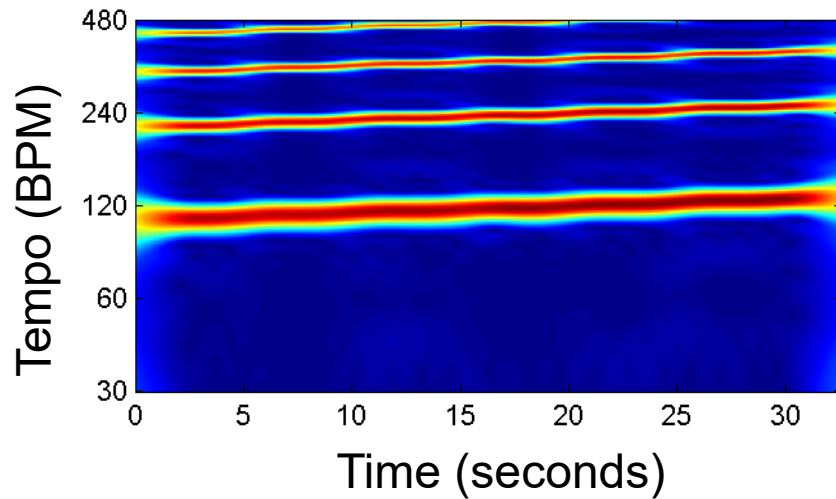
Tempo@Tatum = 210 BPM

Tempo@Measure = 70 BPM

Tempogram

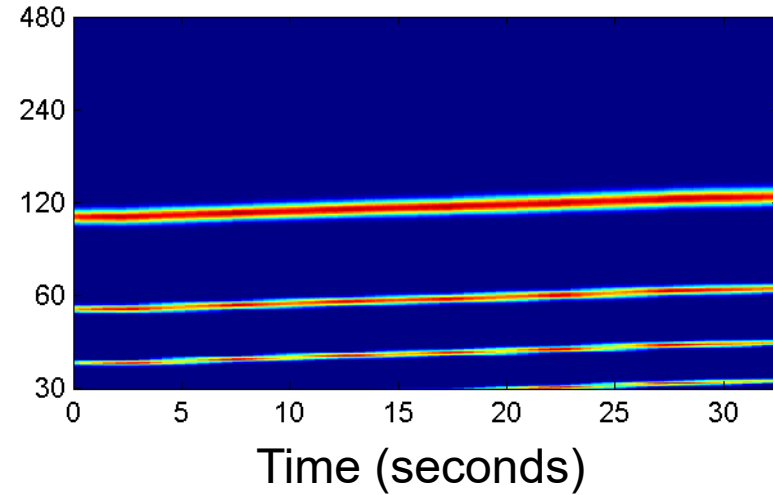


Fourier



Emphasis of tempo harmonics
(integer multiples)

Autocorrelation



Emphasis of tempo subharmonics
(integer fractions)

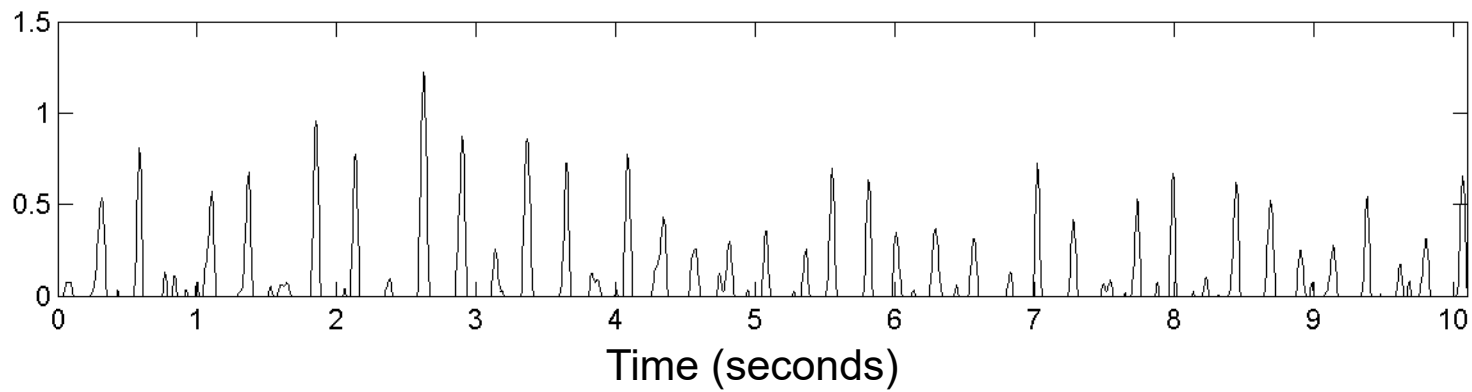
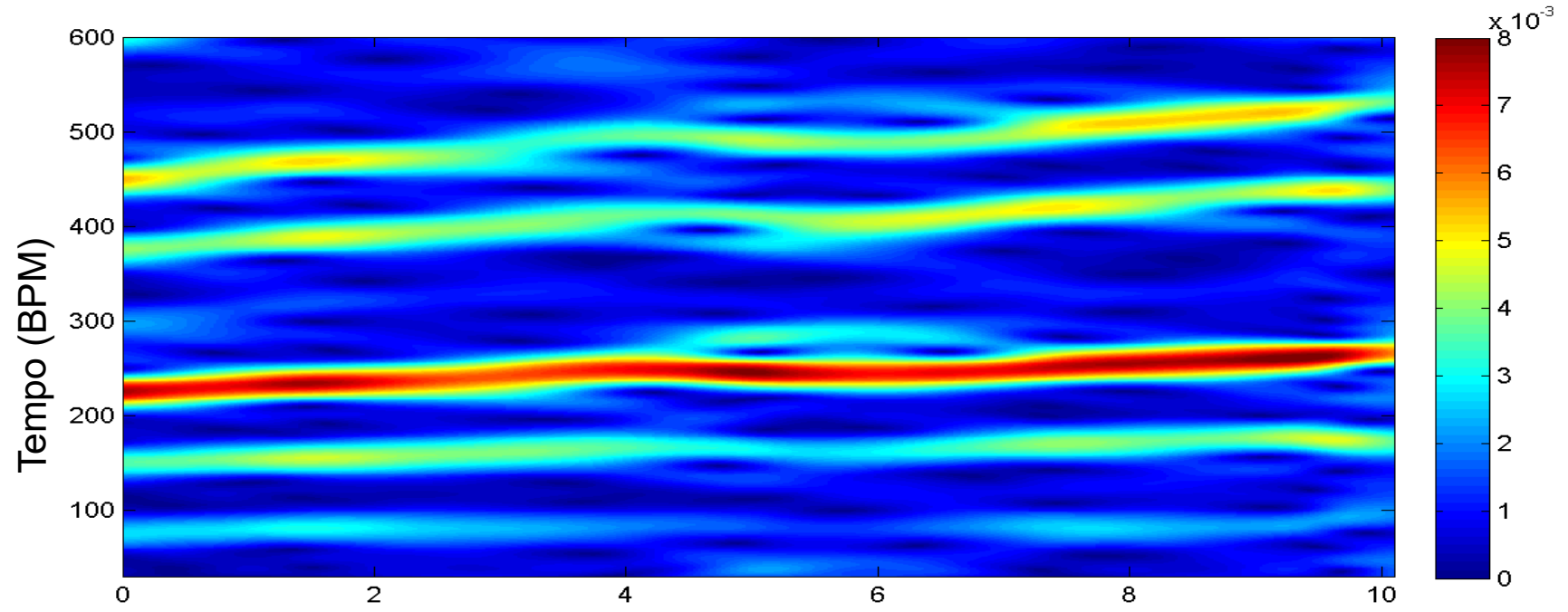
Tempogram (Summary)

| Fourier | Autocorrelation |
|--|--|
| Novelty curve is compared with sinusoidal kernels each representing a specific tempo | Novelty curve is compared with time-lagged local (windowed) sections of itself |
| Convert frequency (Hertz) into tempo (BPM) | Convert time-lag (seconds) into tempo (BPM) |
| Reveals novelty periodicities | Reveals novelty self-similarities |
| Emphasizes harmonics | Emphasizes subharmonics |
| Suitable to analyze tempo on tatum and tactus level | Suitable to analyze tempo on tactus and measure level |

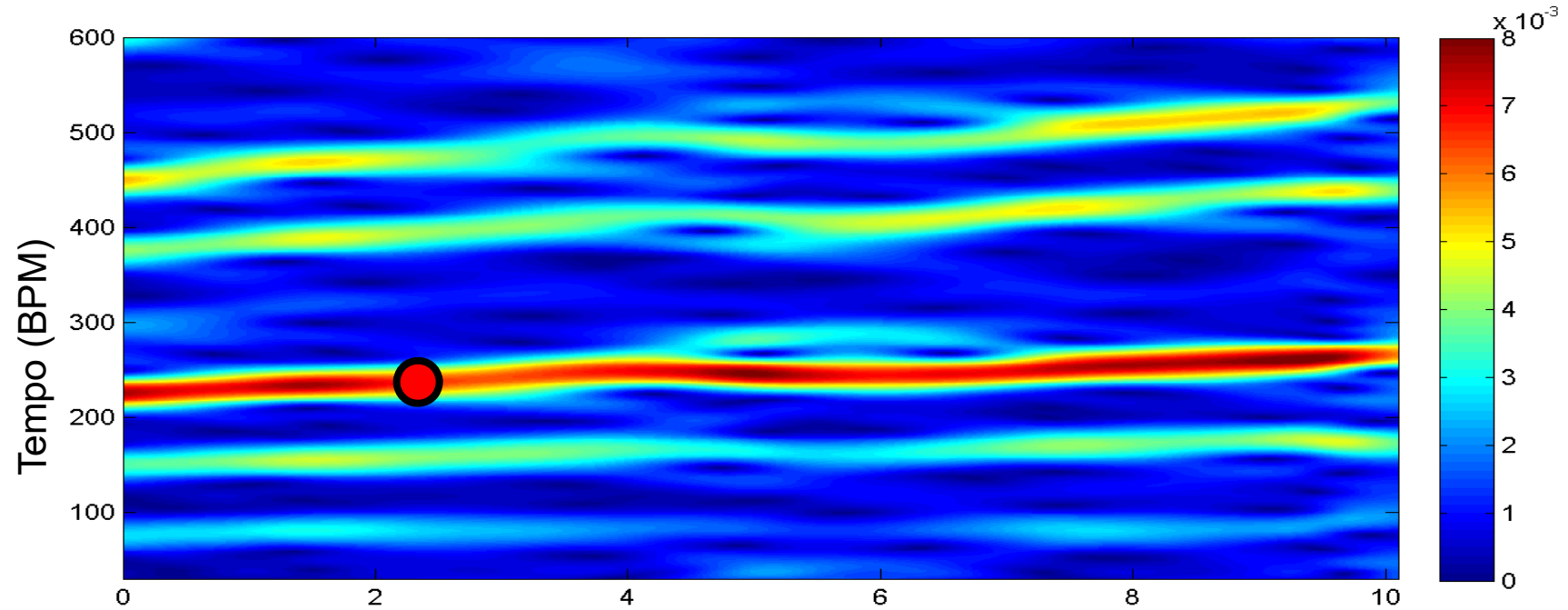
Beat Tracking

- Given the tempo, find the best sequence of beats
- Complex Fourier tempogram contains **magnitude** and **phase** information
- The **magnitude** encodes how well the novelty curve resonates with a sinusoidal kernel of a specific tempo
- The **phase** optimally aligns the sinusoidal kernel with the peaks of the novelty curve

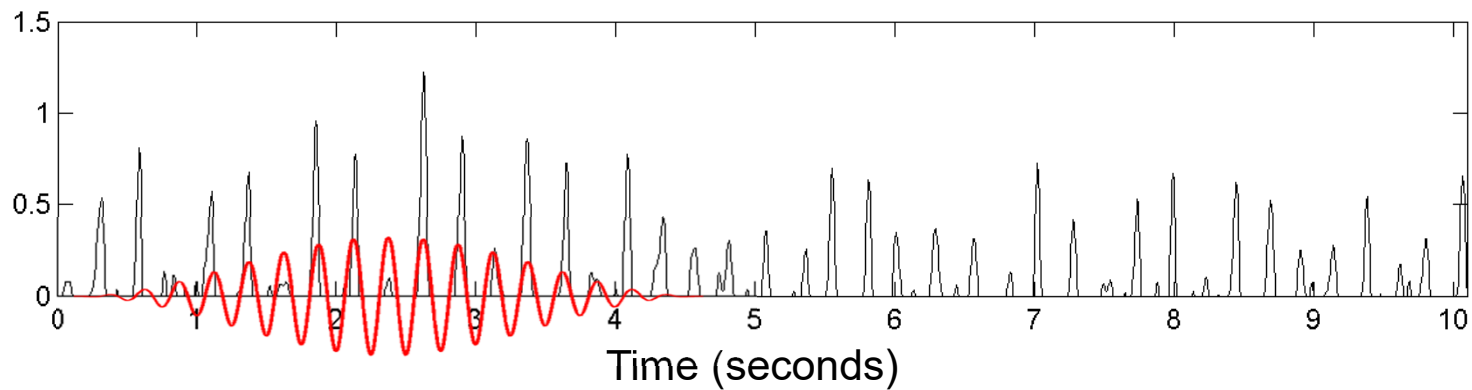
Local Pulse Tracking



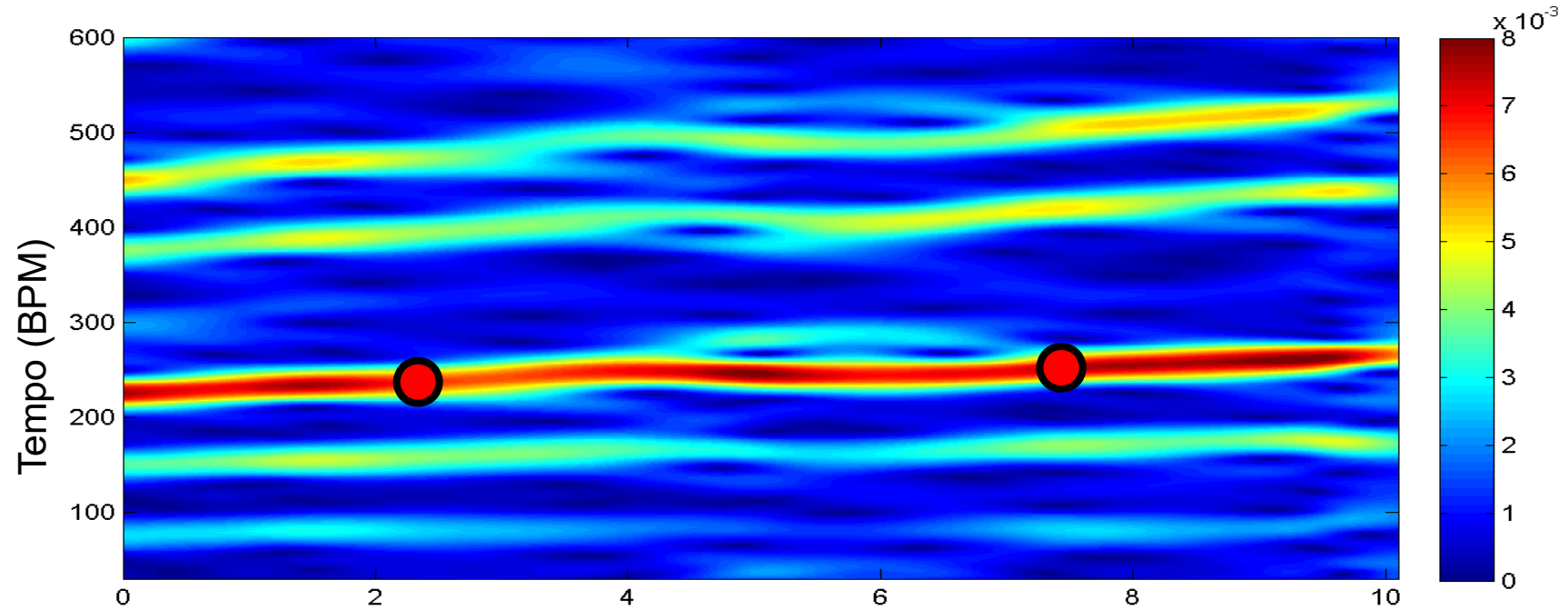
Local Pulse Tracking



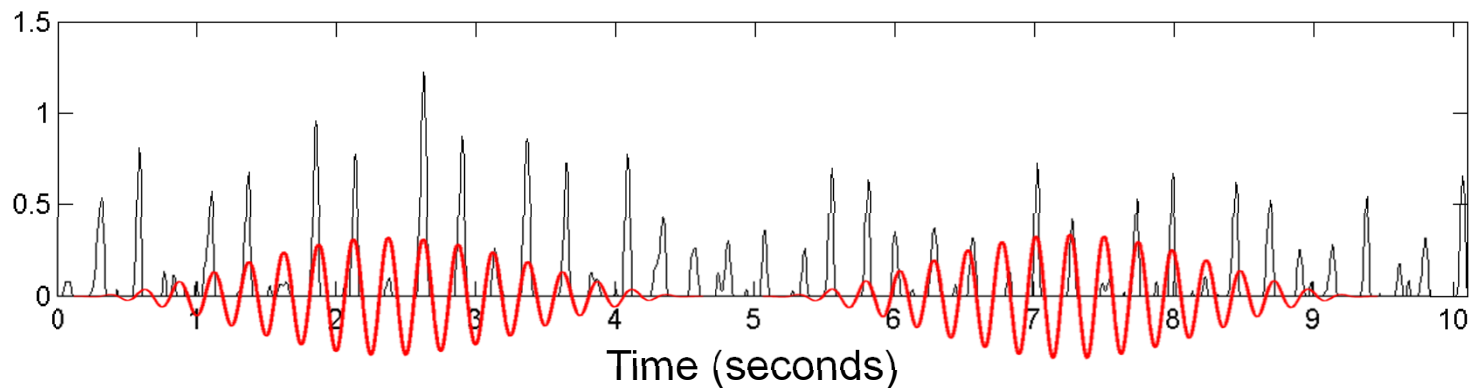
Optimizing local periodicity kernel



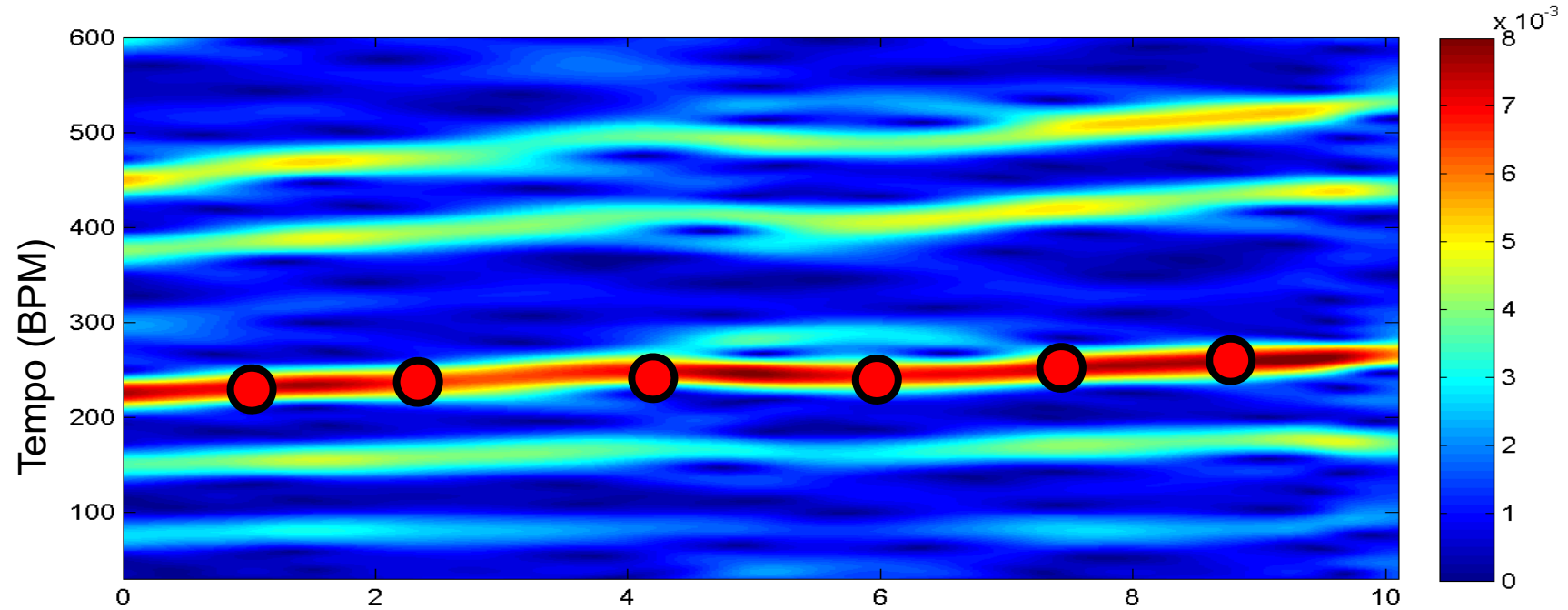
Local Pulse Tracking



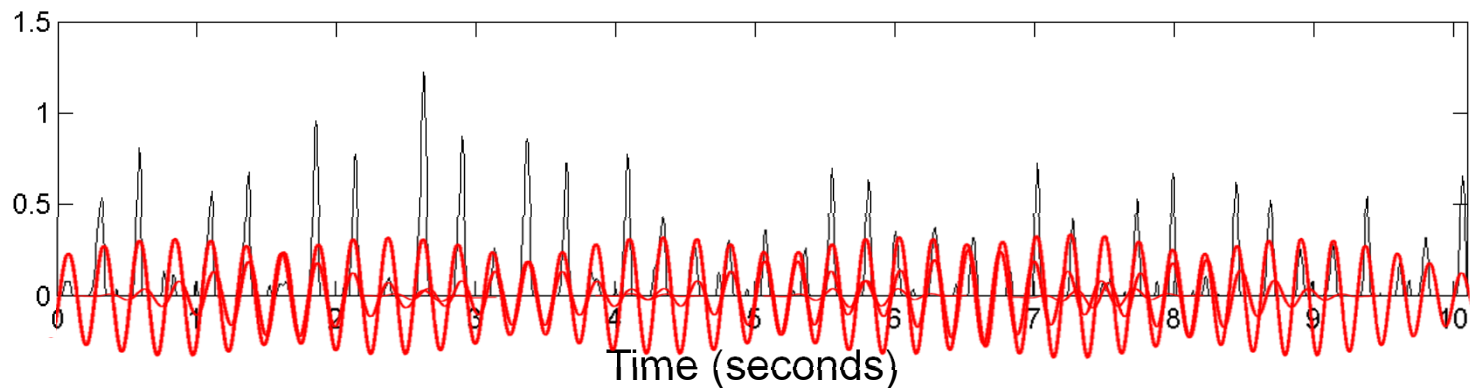
Optimizing local periodicity kernel



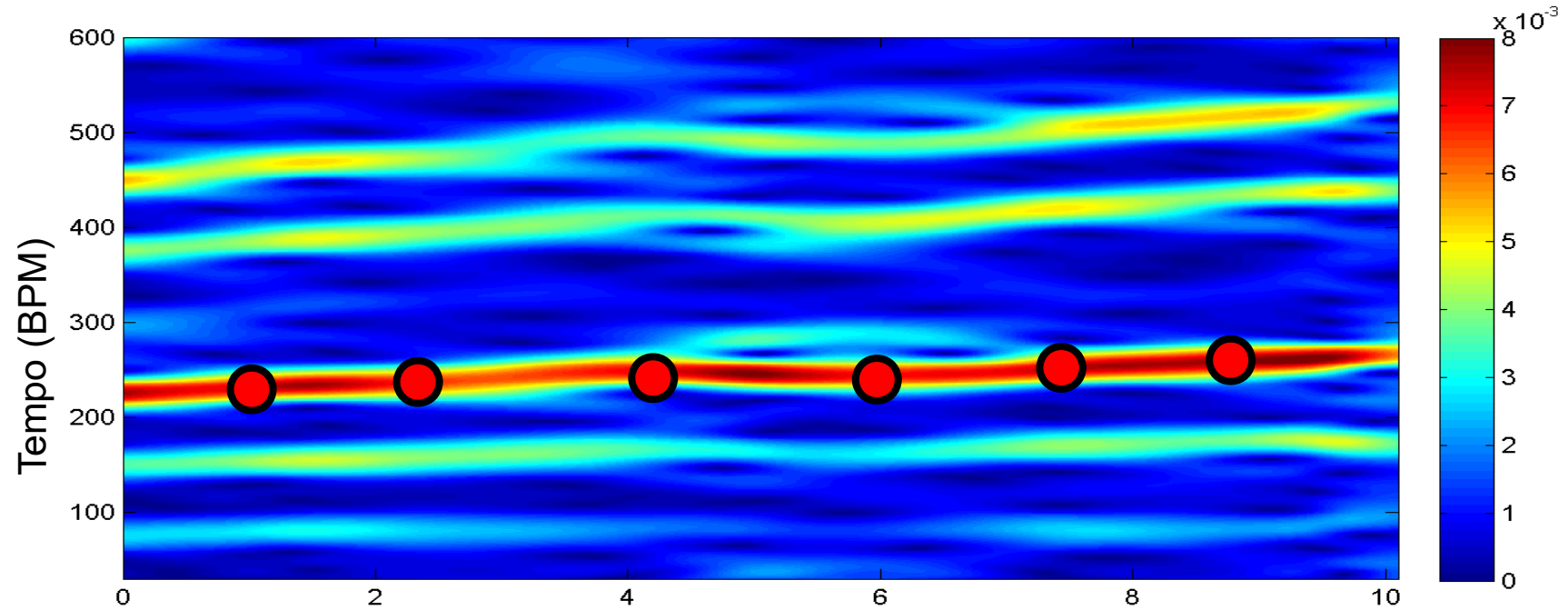
Local Pulse Tracking



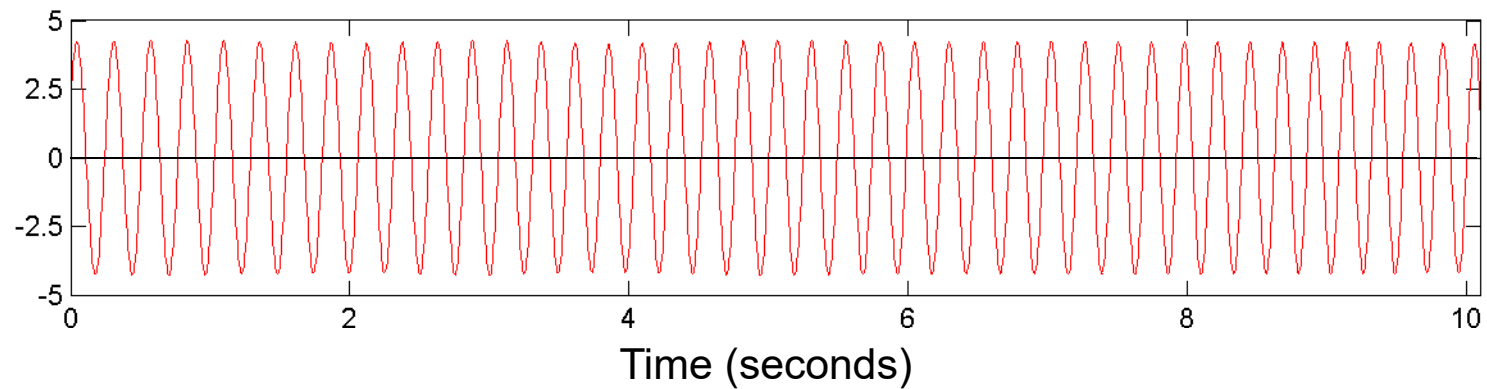
Optimizing local periodicity kernel



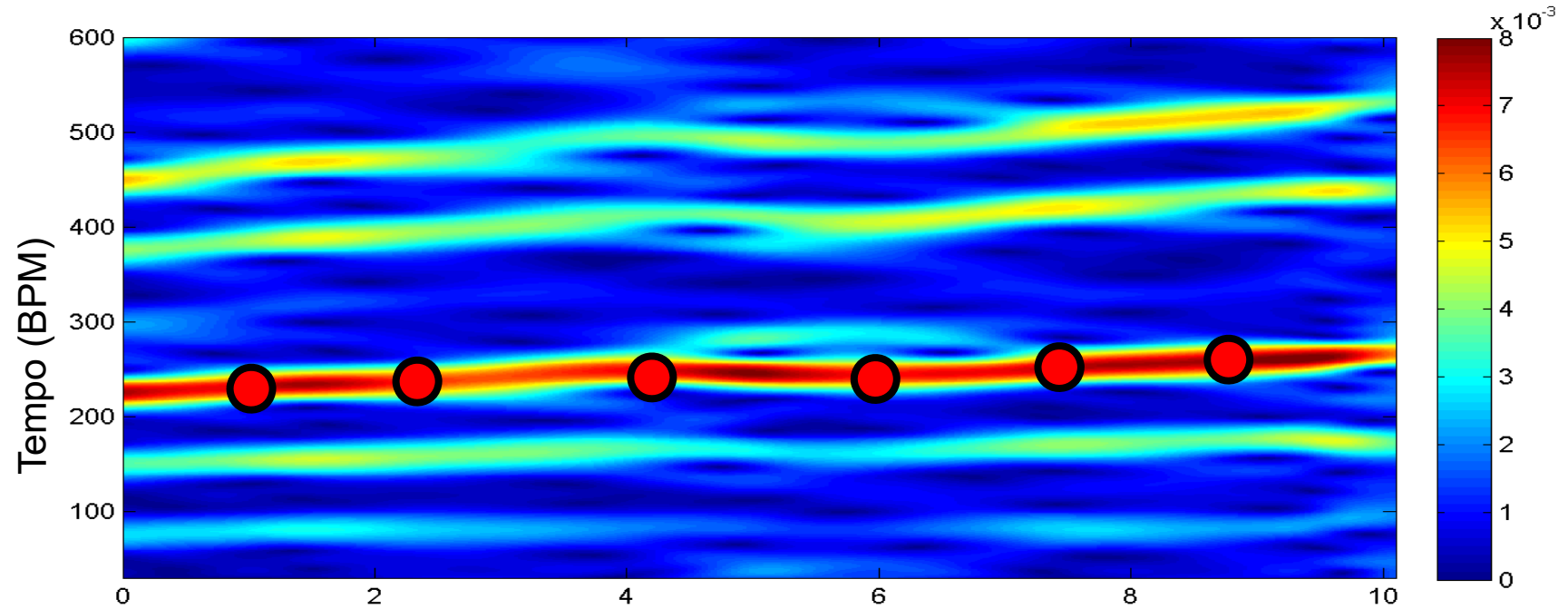
Local Pulse Tracking



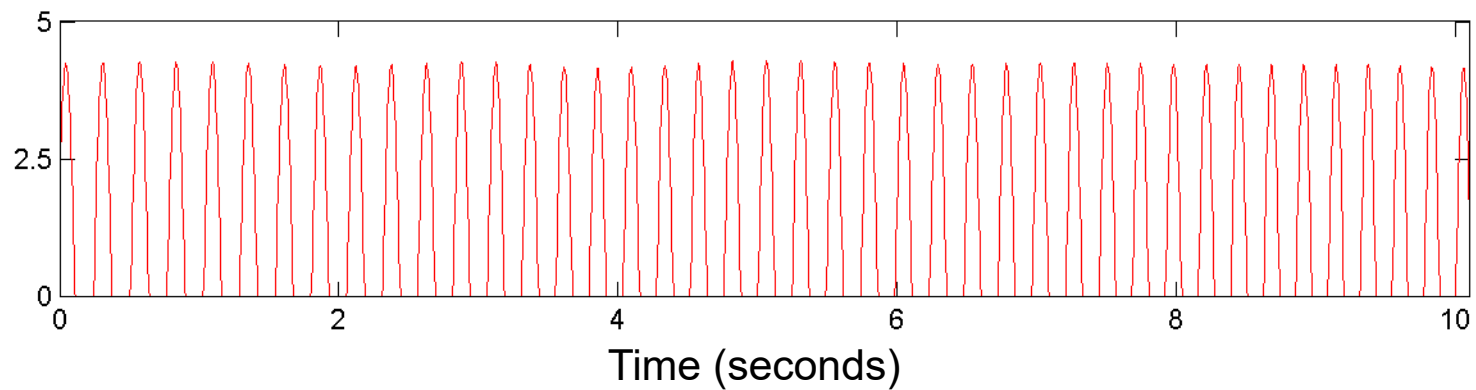
Accumulation of kernels



Local Pulse Tracking

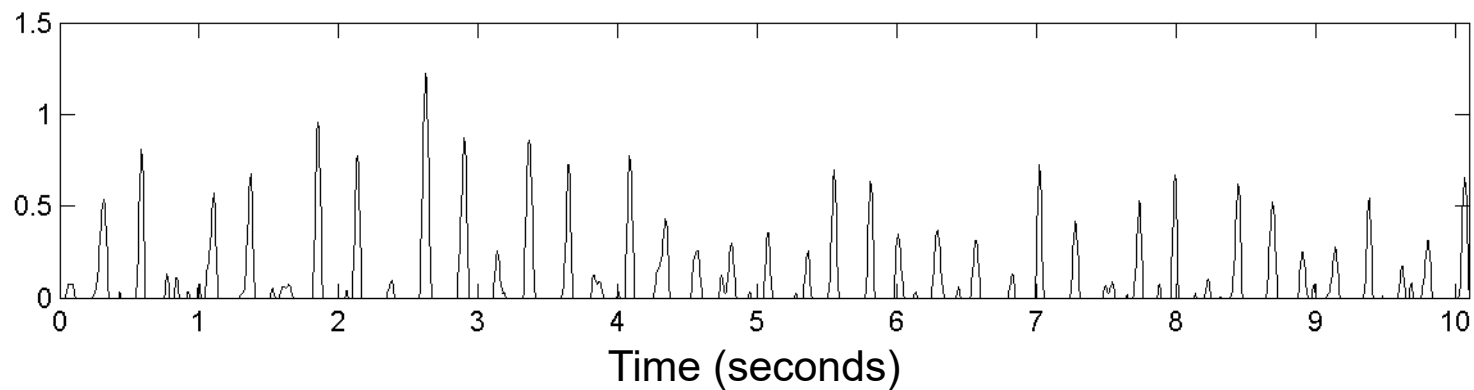


Halfwave rectification

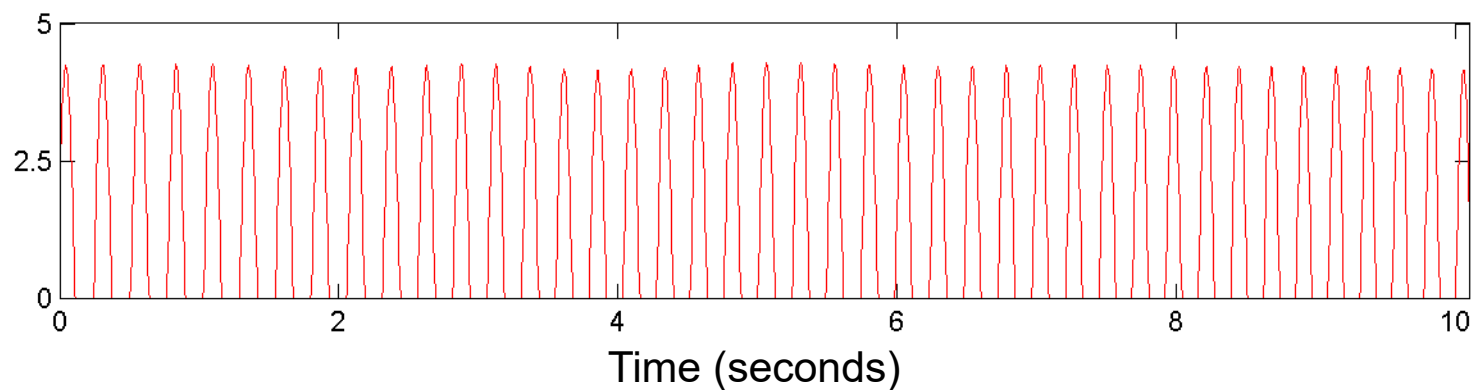


Local Pulse Tracking

Novelty Curve



Predominant Local Pulse (PLP)



Local Pulse Tracking

Novelty Curve

- Indicates note onset candidates
- Extraction errors in particular for soft onsets
- Simple peak-picking problematic



Predominant Local Pulse (PLP)

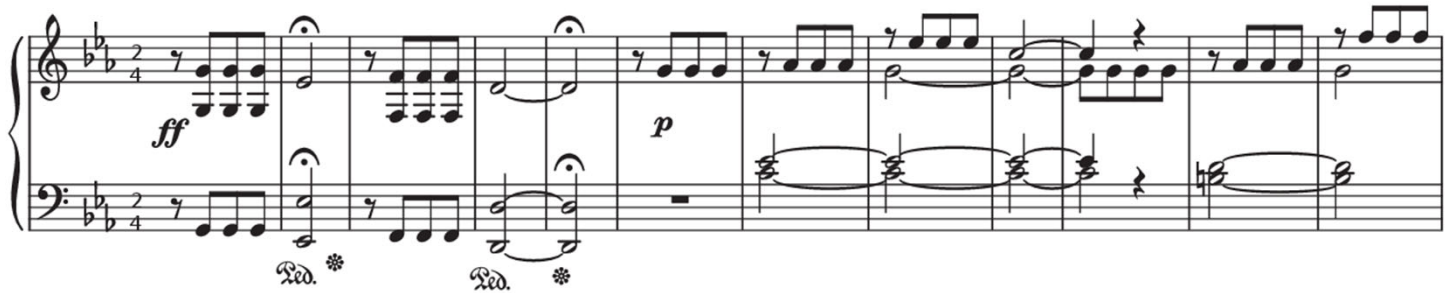
- Periodicity enhancement of novelty curve
- Accumulation introduces error robustness
- Locality of kernels handles tempo variations



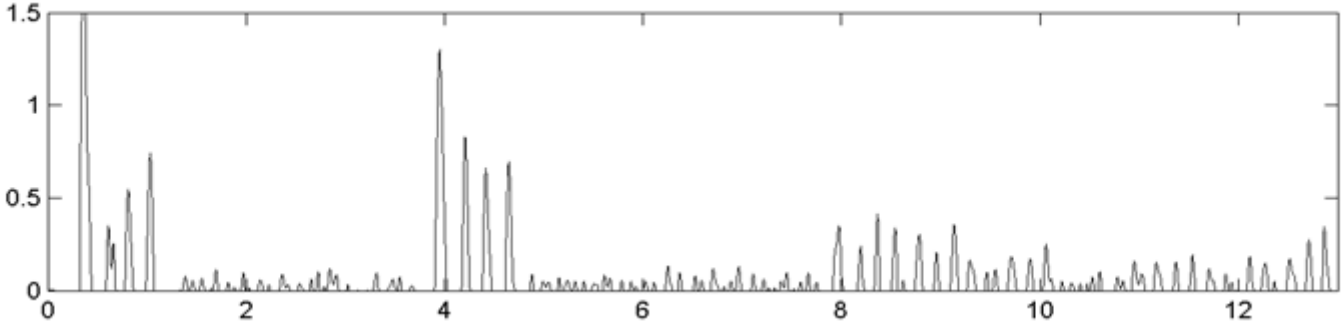
Local Pulse Tracking

- Novelty function $\Delta : \mathbb{Z} \rightarrow \mathbb{R}$
- Fourier representation $\mathcal{F}(n, \omega) = \sum_{m \in \mathbb{Z}} \Delta(m) \bar{w}(m - n) \exp(-2\pi i \omega m)$
- Fourier tempogram $\mathcal{T}^F(n, \tau) = |\mathcal{F}(n, \tau/60)|$
- Maximal tempo (per frame) $\tau_n := \operatorname{argmax}_{\tau \in \Theta} \mathcal{T}^F(n, \tau)$ Set of tempi, e.g.,
 $\Theta = [30:300]$ BPM
- Phase $\varphi_n = -\frac{1}{2\pi} \operatorname{angle}(\mathcal{F}(n, \tau_n/60))$
- Sinusoidal kernel $\kappa_n(m) := w(m - n) \cos\left(2\pi\left((\tau_n/60) \cdot m - \varphi_n\right)\right)$
- Periodicity curve (PLP) $\Gamma(m) = \left| \sum_{n \in \mathbb{Z}} \kappa_n(m) \right|_{\geq 0}$

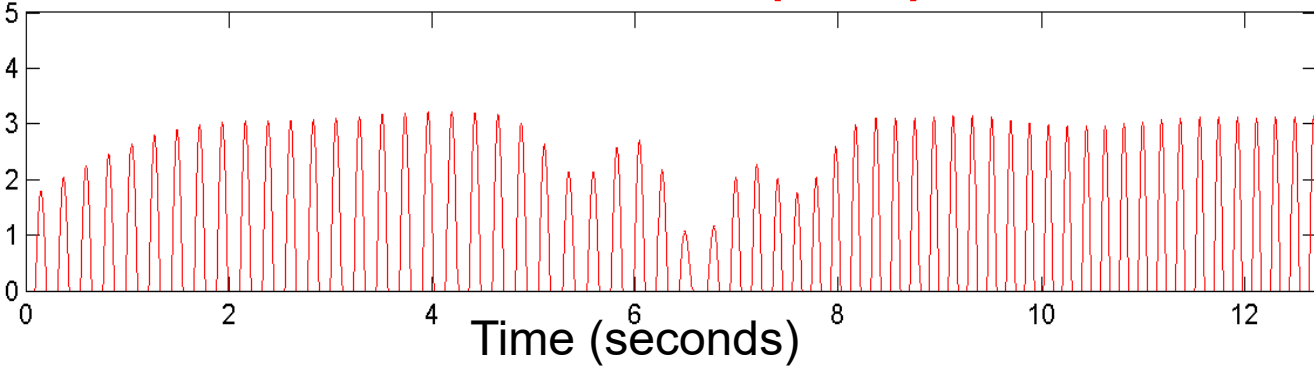
Local Pulse Tracking



Novelty Curve

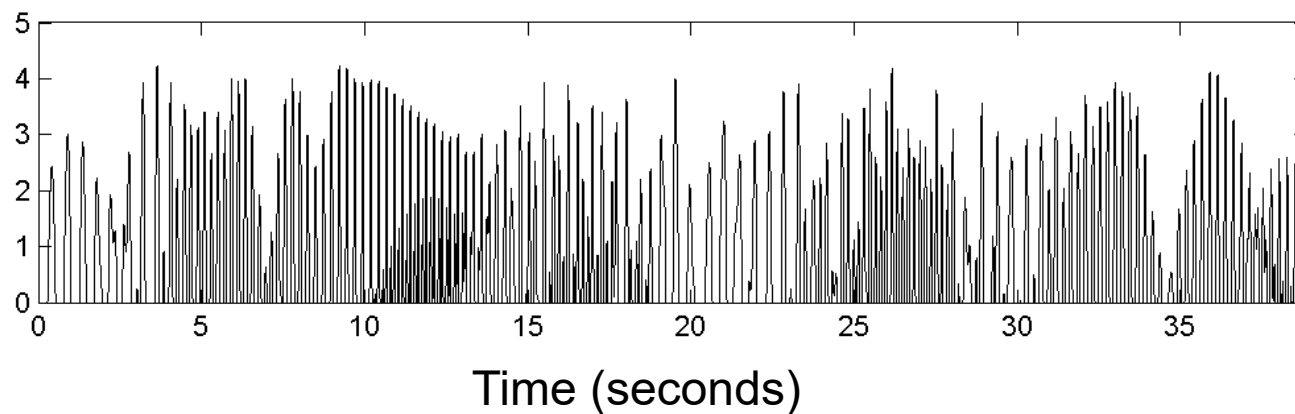
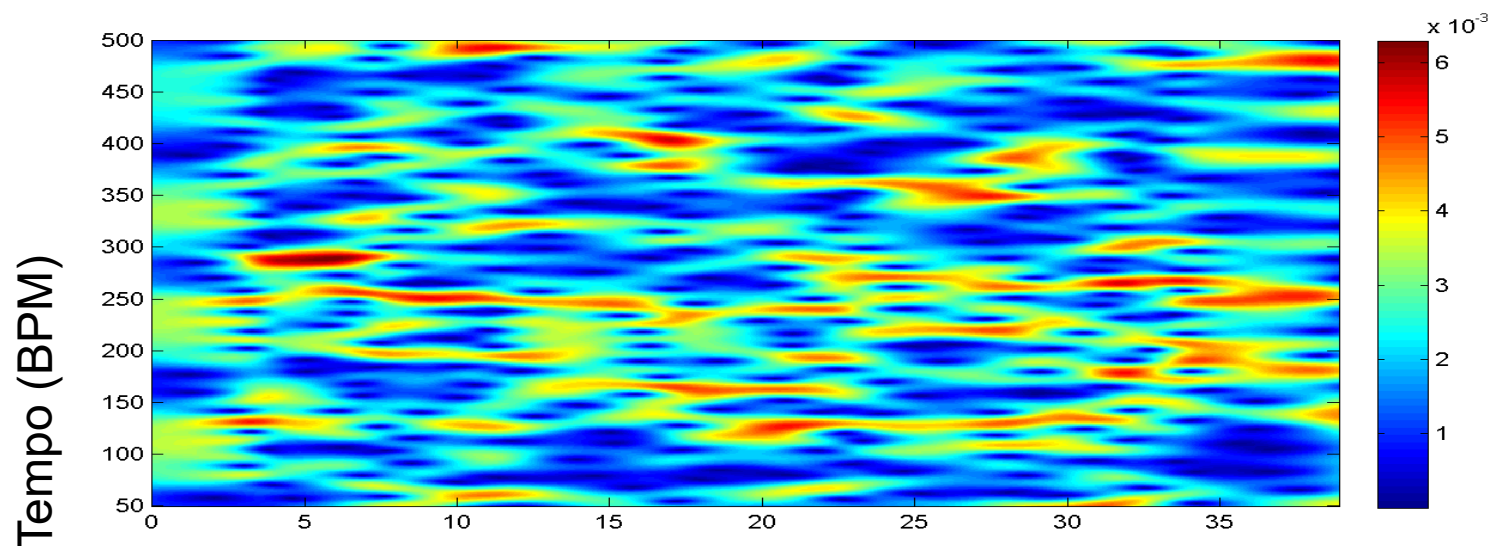


Predominant Local Pulse (PLP)



Local Pulse Tracking

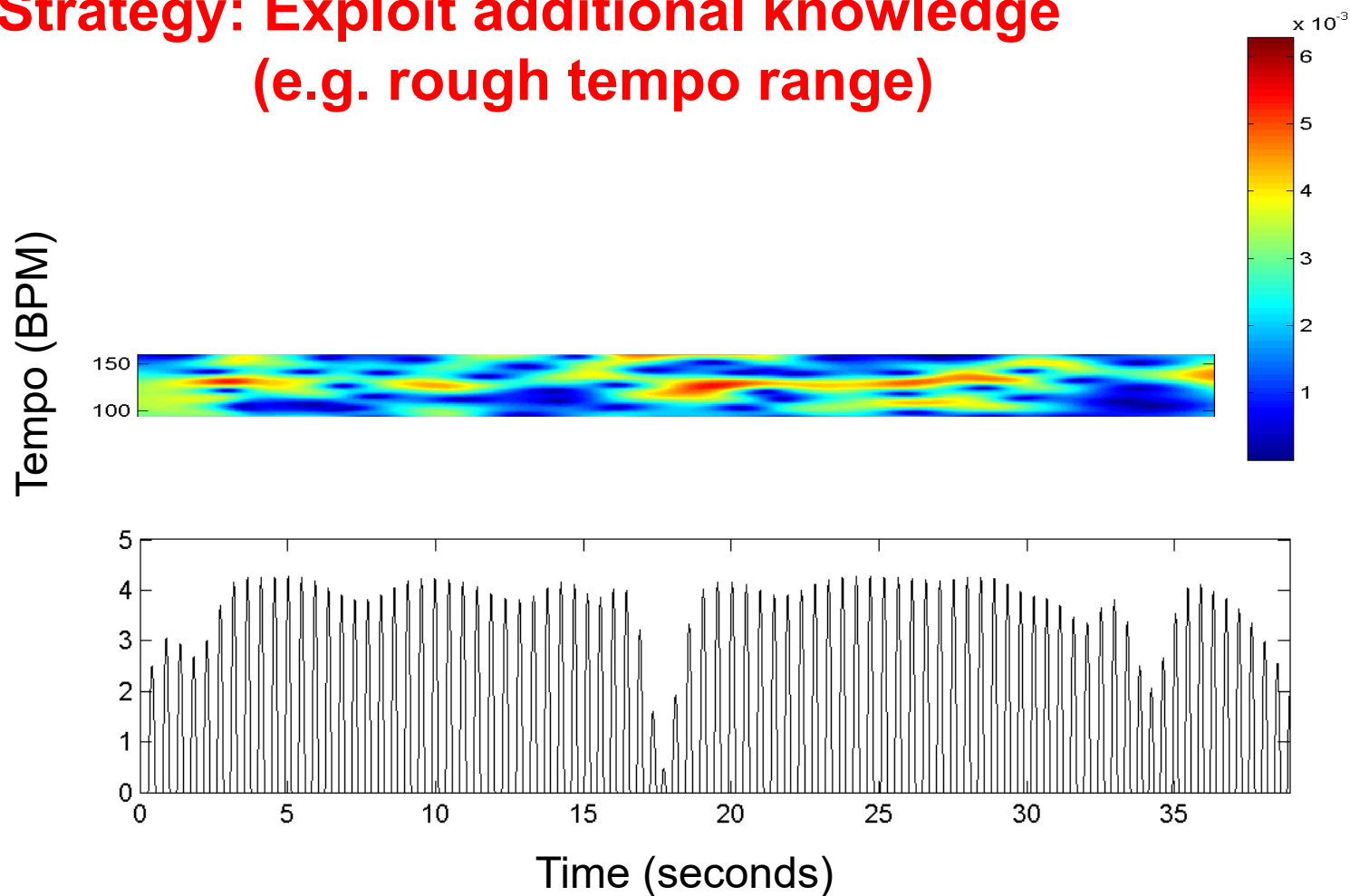
Borodin – String Quartet No. 2



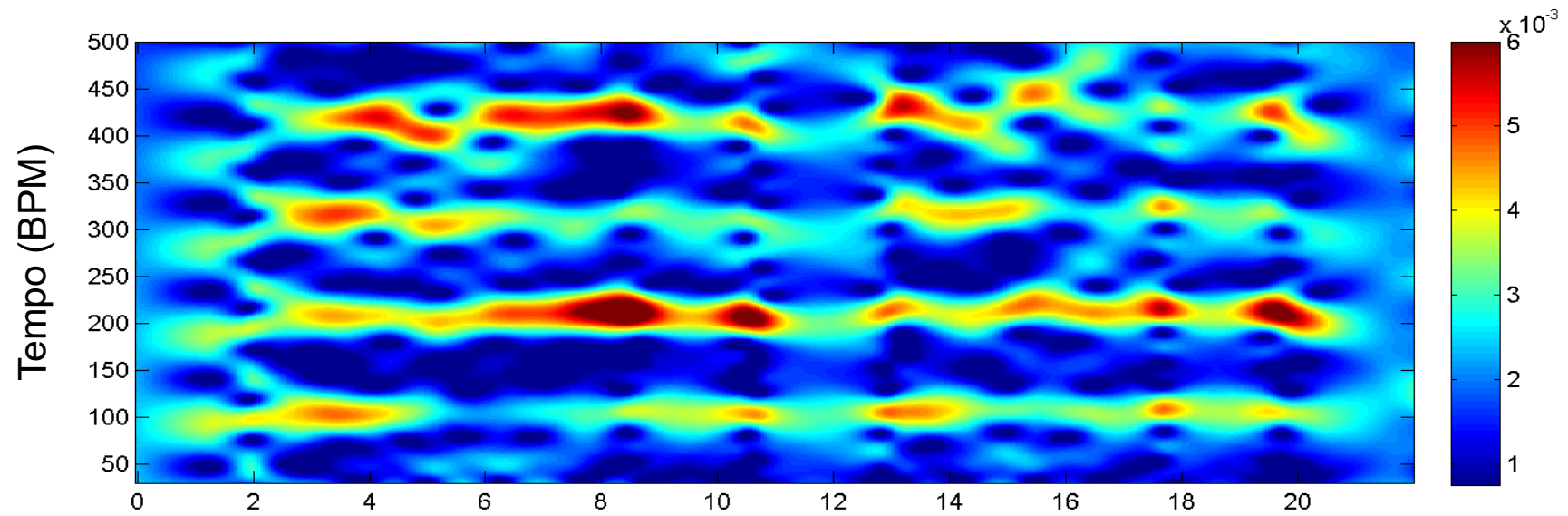
Local Pulse Tracking

Borodin – String Quartet No. 2

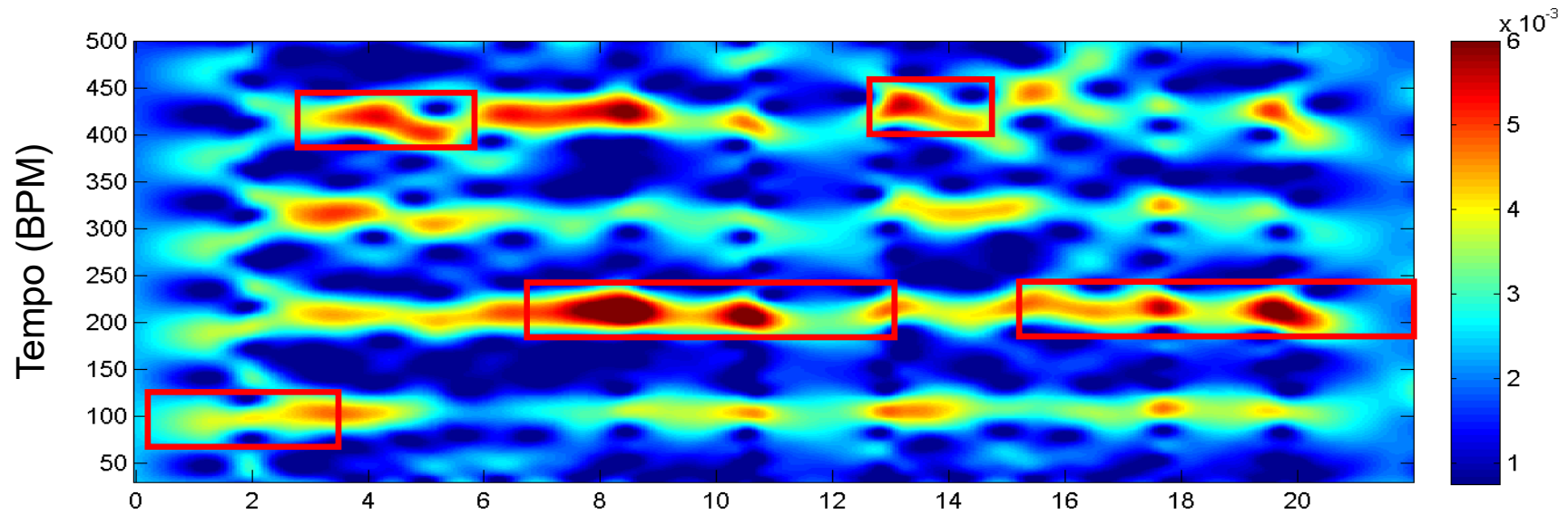
**Strategy: Exploit additional knowledge
(e.g. rough tempo range)**



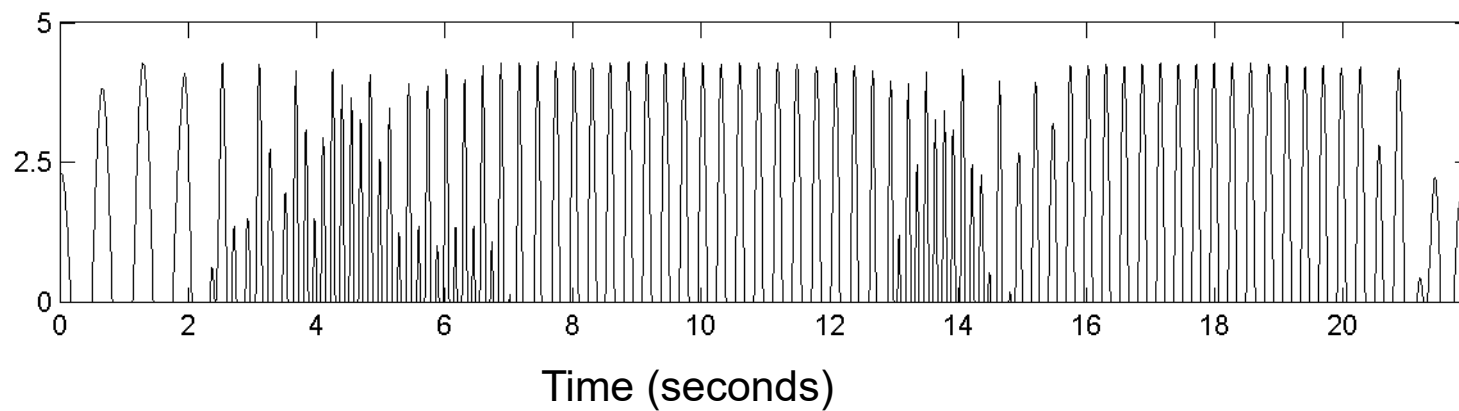
Pulse Levels



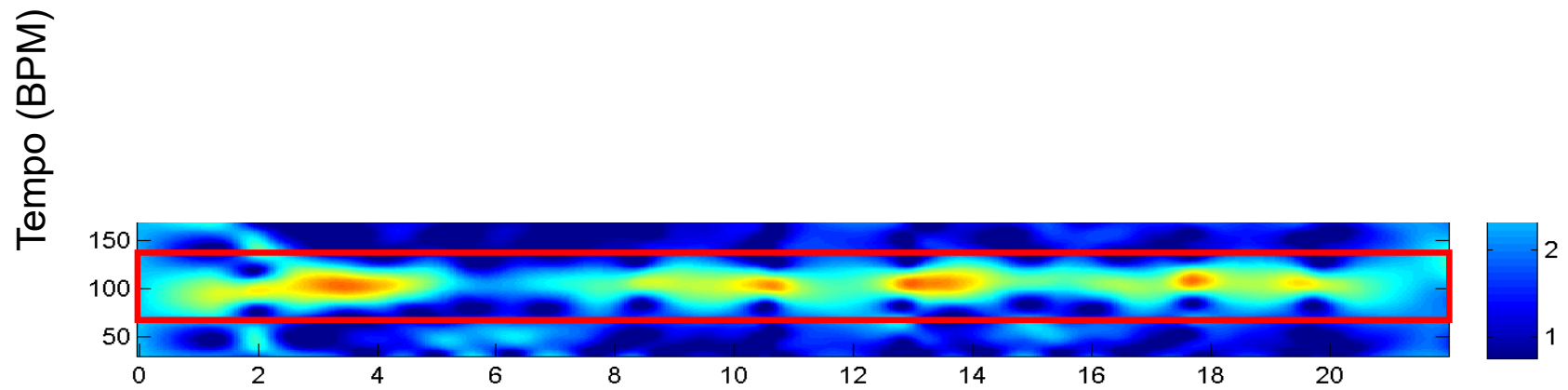
Pulse Levels



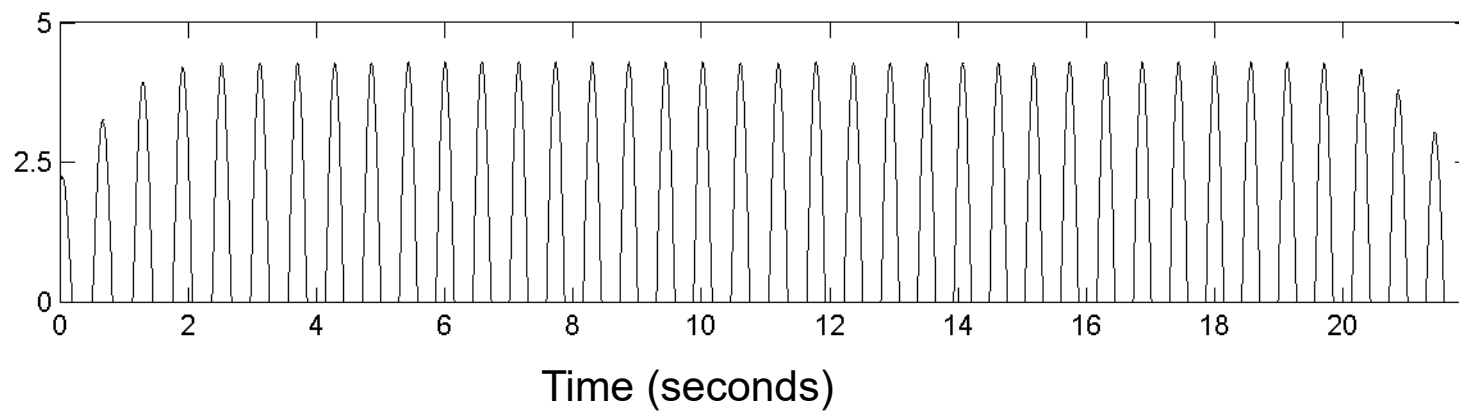
Switching of predominant pulse level



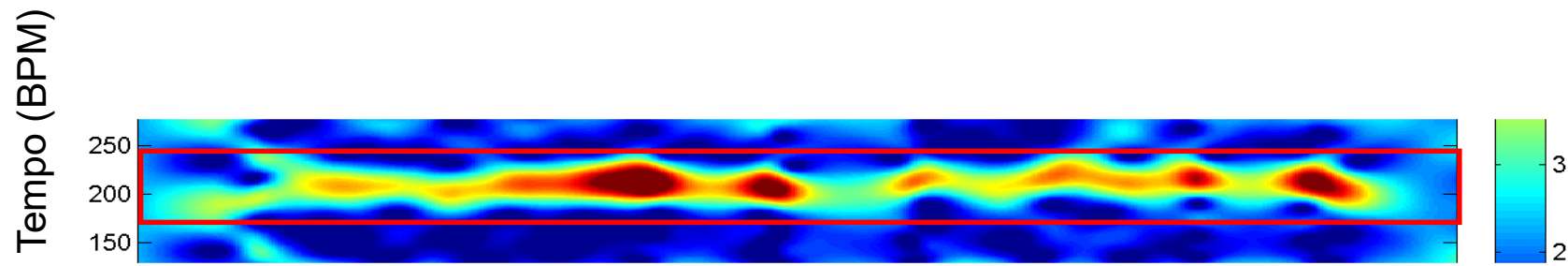
Pulse Levels



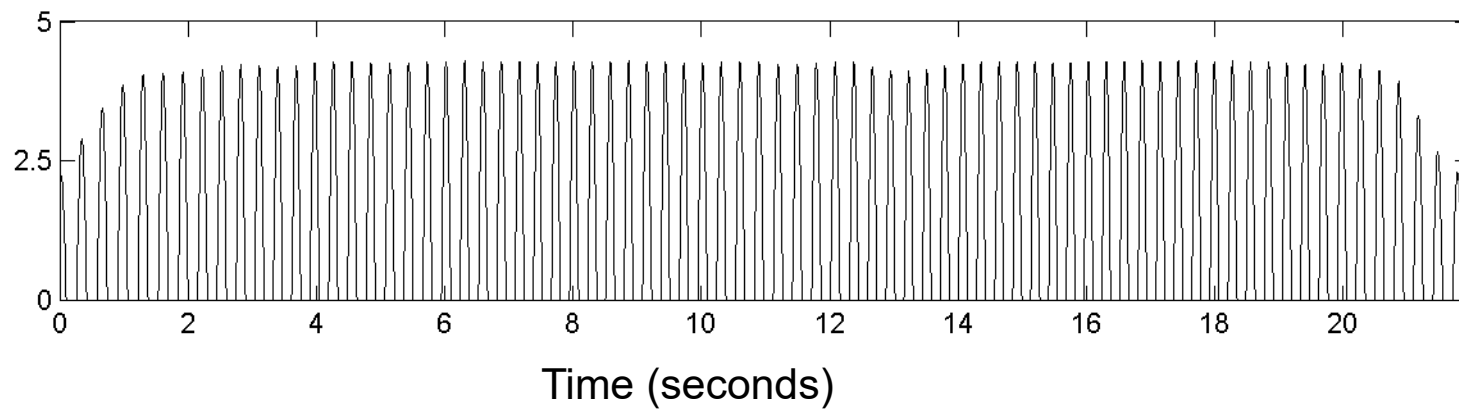
1/4 note pulse level



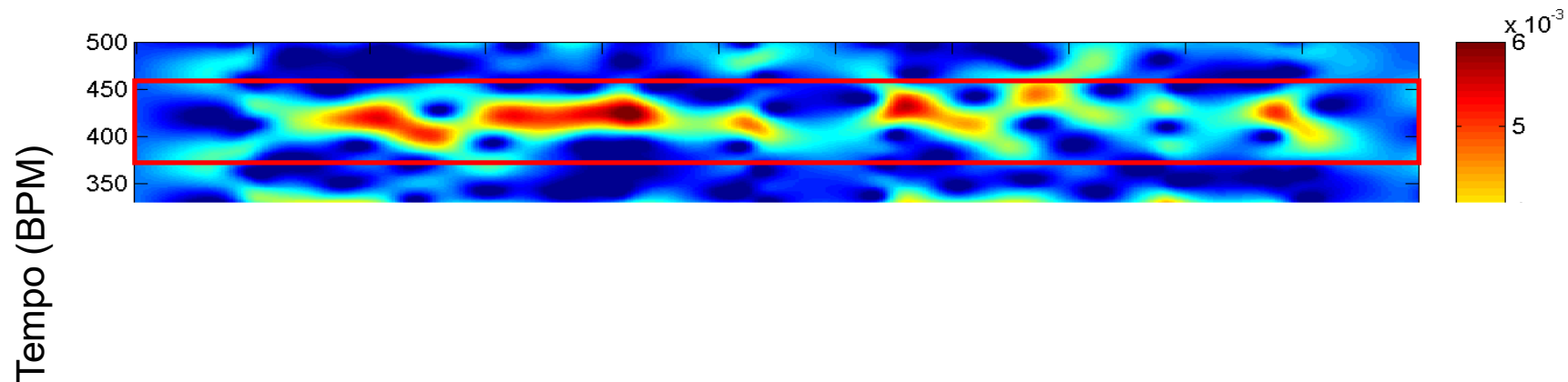
Pulse Levels



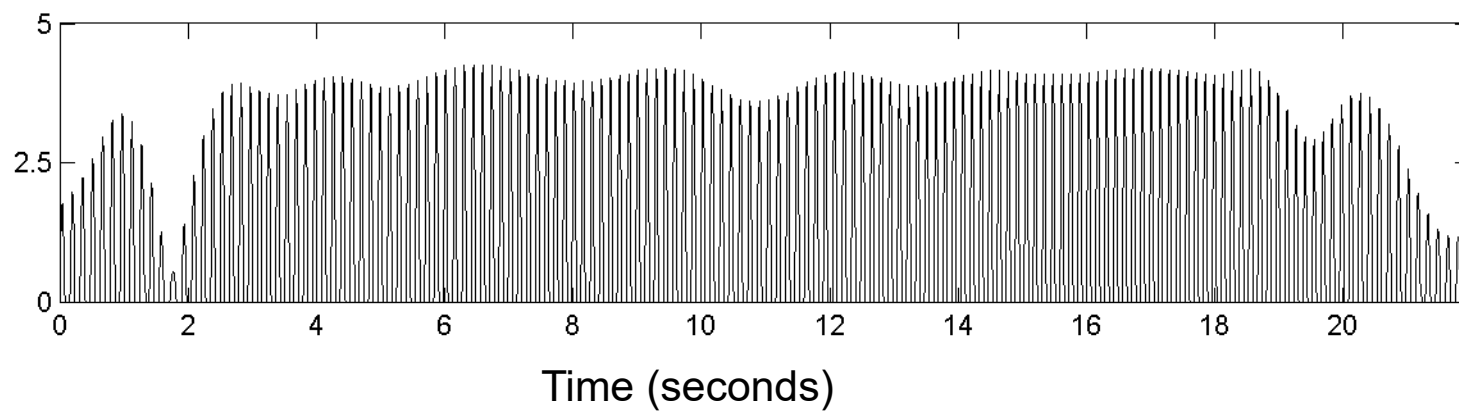
1/8 note pulse level



Pulse Levels

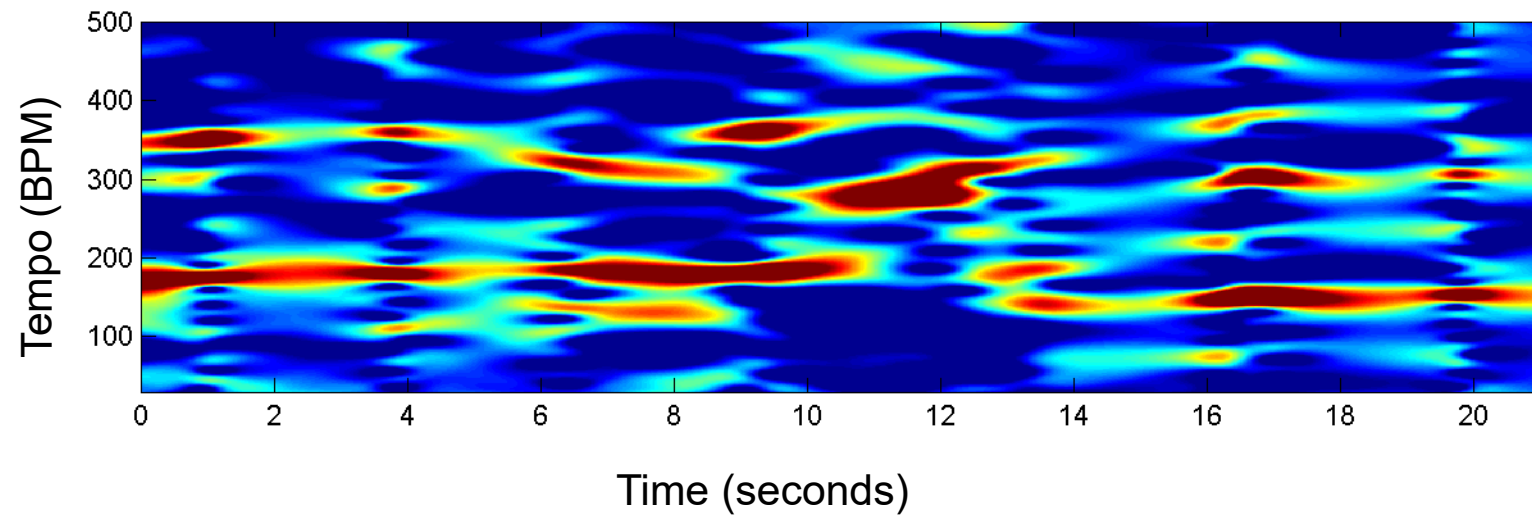


1/16 note pulse level



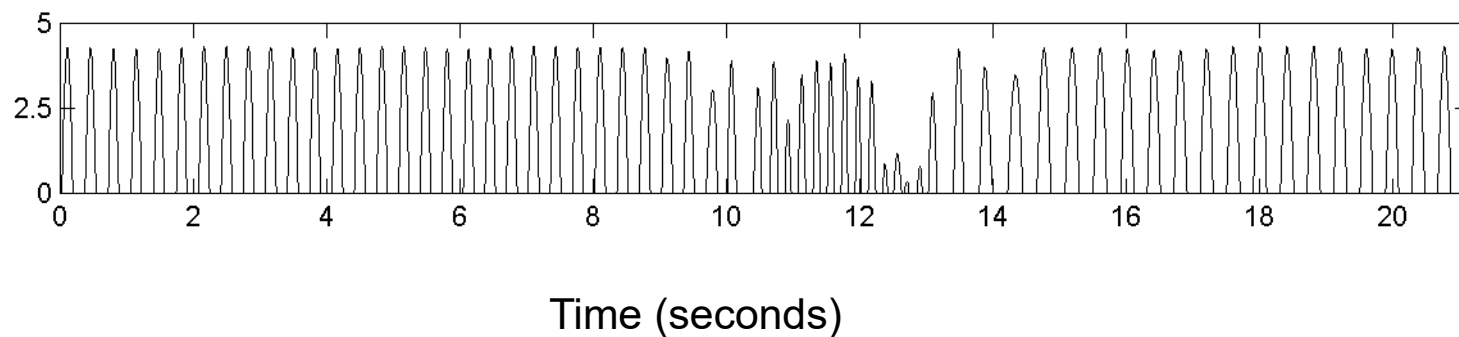
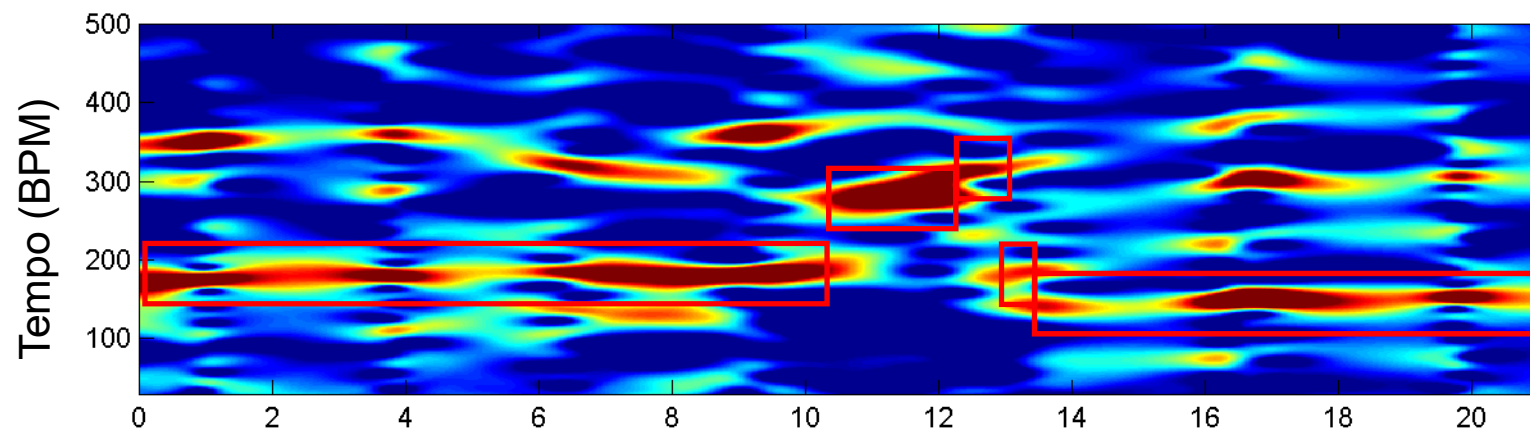
Local Pulse Tracking

Brahms Hungarian Dance No. 5



Local Pulse Tracking

Brahms Hungarian Dance No. 5

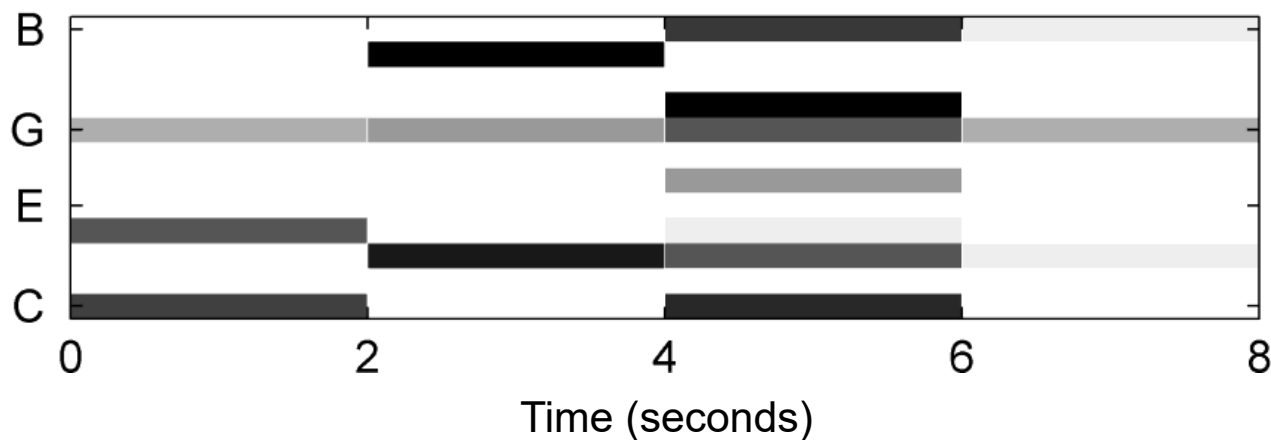
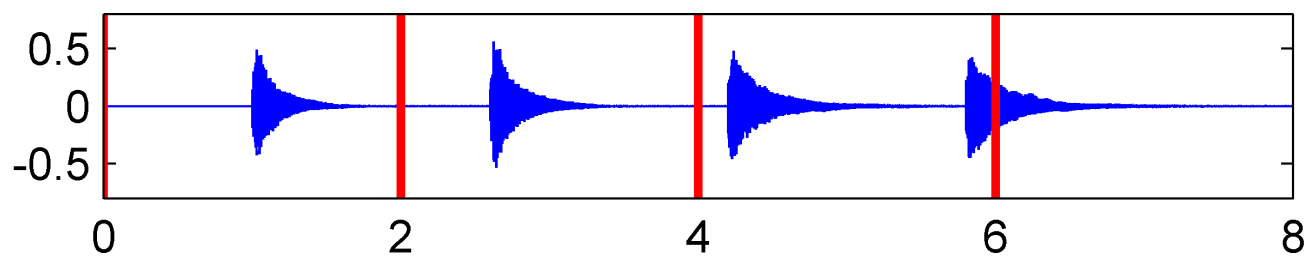
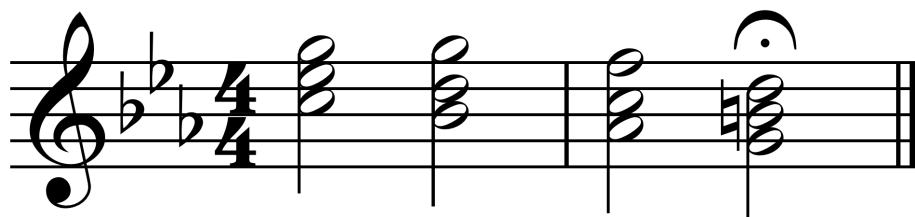


Applications

- Feature design
(beat-synchronous features, adaptive windowing)
- Digital DJ / audio editing
(mixing and blending of audio material)
- Music classification
- Music recommendation
- Performance analysis
(extraction of tempo curves)

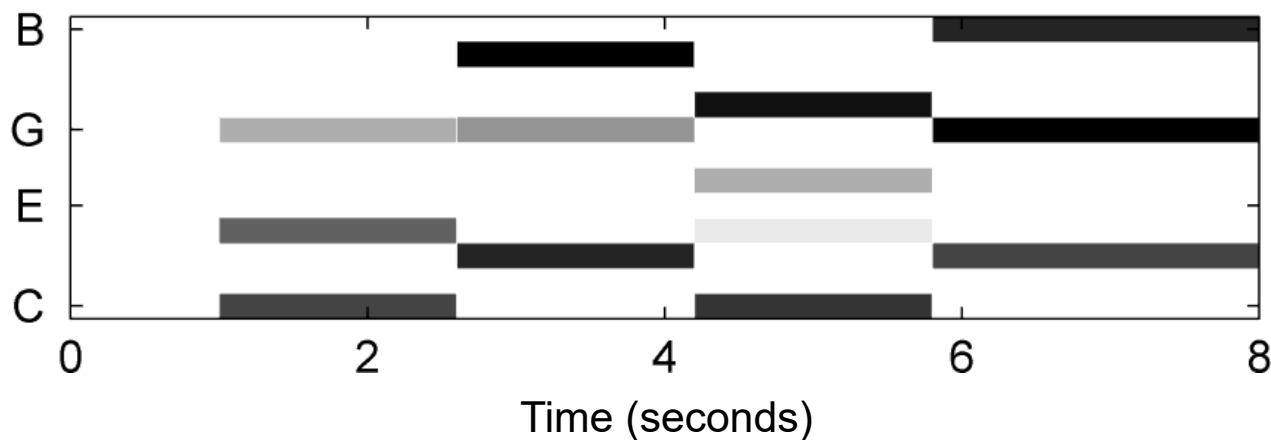
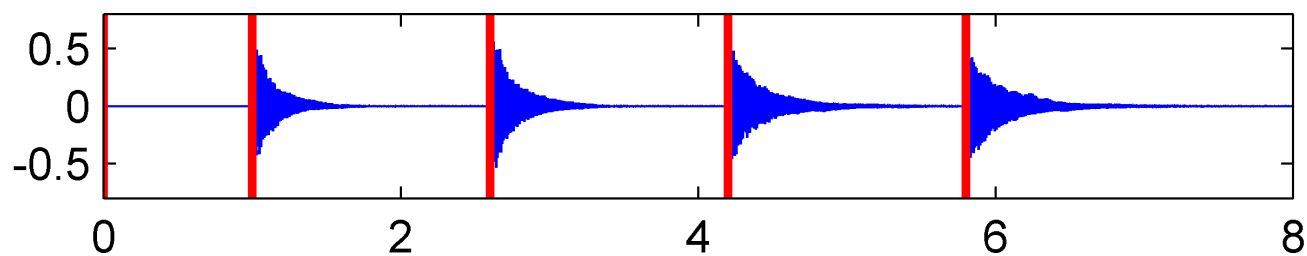
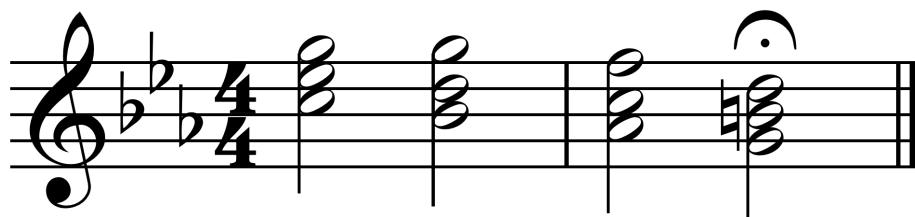
Application: Feature Design

Fixed window size



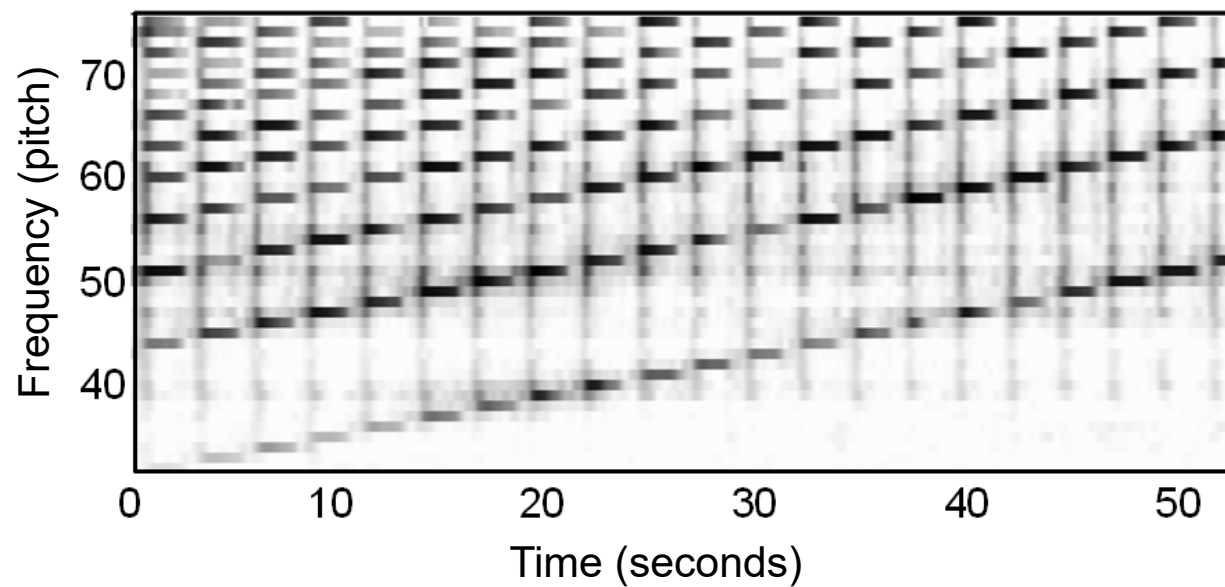
Application: Feature Design

Adaptive window size



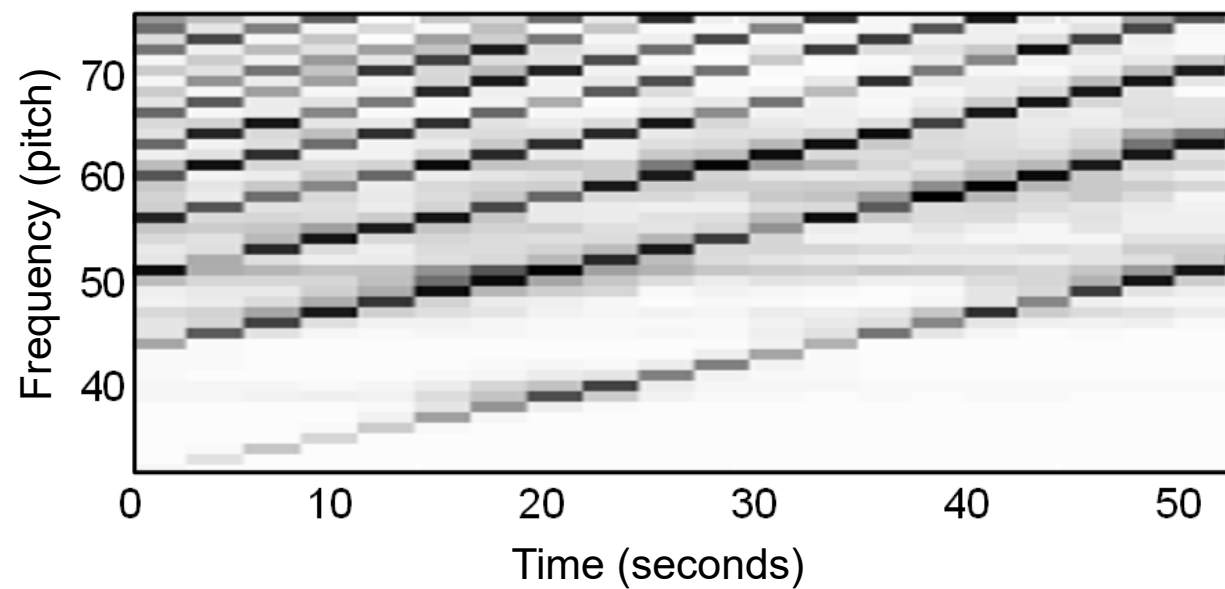
Application: Feature Design

Fixed window size



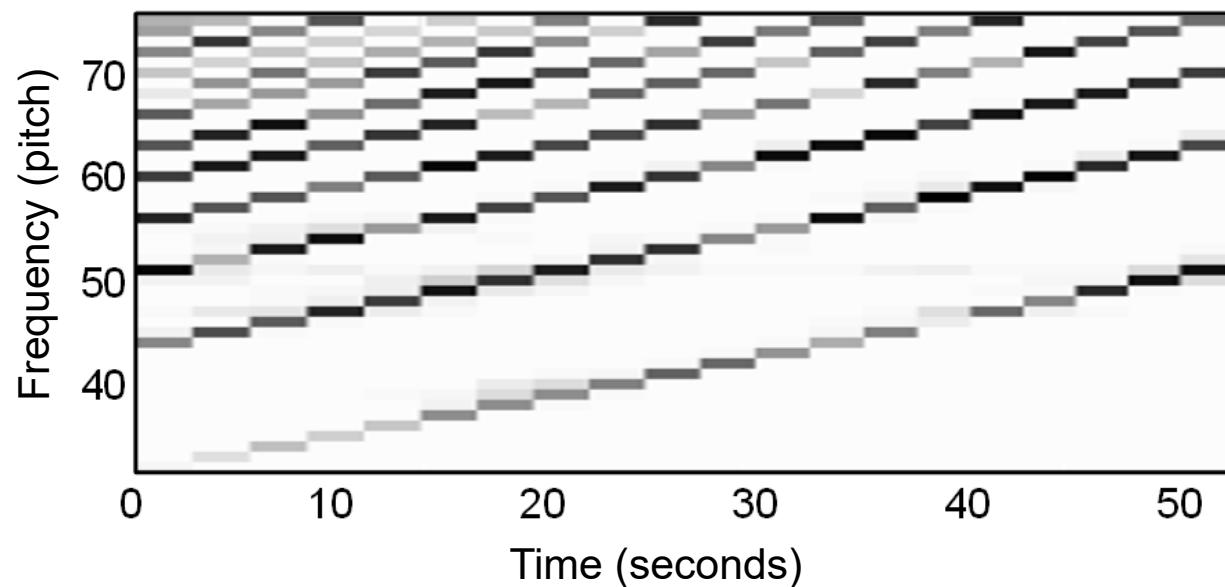
Application: Feature Design

Adaptive window size



Application: Feature Design

Adaptive window size



Denoising by excluding boundary neighborhoods

Application: Audio Editing (Digital DJ)

The screenshot displays the Mixxx 1.7.0 interface. At the top, the menu bar includes File, Library, Options, and Help. The main area is divided into two channels, CHANNEL 1 and CHANNEL 2. CHANNEL 1 is playing 'Alex Metric, Deadly On A Mission (Dub)' with a BPM of 123.98 and a duration of 00:24. CHANNEL 2 is playing 'Junior Boys, No Kinda Man (Chloé Remix)' with a BPM of 123.98 and a duration of 07:55. The interface features various controls such as CUE, REV, SYNC, and volume sliders. A playlist table is located at the bottom, listing tracks with columns for Artist, Title, Type, Length, kbit, BPM, and Comment. The track 'No Kinda Man (Chloé Remix)' by Junior Boys is highlighted.

| Artist | Title | Type | Length | kbit | BPM | Comment |
|--------------------------------|---|------|--------|------|-------|---------|
| Danger | 11h30 - Original Mix | mp3 | 3:40 | 320 | 132.3 | |
| Danger | 19h11 - Original Mix | mp3 | 5:54 | 320 | 122.9 | |
| Danger | 7h46 | mp3 | 5:25 | 160 | 118.0 | |
| Evolve | Safe To Dream Thrillseekers Re | mp3 | 7:32 | 0 | 139.5 | |
| Futurecop! | Class of 1984 (Anoraak Remix) | mp3 | 7:28 | 0 | 120.0 | |
| Global Deejays feat. Techno... | Get Up (Before The Night Is Over) (General Electric ... | mp3 | 6:35 | 0 | 128.2 | |
| Hardfloor | Murano | mp3 | 8:22 | 0 | 126.6 | |
| lio | Rapture | mp3 | 3:27 | 128 | 125.5 | |
| Junior Boys | No Kinda Man (Chloé Remix) | mp3 | 8:54 | 0 | 124.0 | |
| Justice | D.A.N.C.E. | mp3 | 4:02 | 0 | 113.0 | |
| Justice | Newjack | mp3 | 3:36 | 0 | 115.1 | |
| Justice | Waters of Nazareth | mp3 | ? | 0 | 0.0 | |
| Kavinsky | Wayfarer | mp3 | 4:29 | 128 | 125.4 | |
| Kavinsky | Testarossa SebAstian Remix | mp3 | 4:58 | 0 | 130.0 | |

Application: Beat-Synchronous Light Effects



Summary

1. Onset Detection

- Novelty curve (*something is changing*)
- Indicates note onset candidates
- Hard task for non-percussive instruments (strings)

2. Tempo Estimation

- Fourier tempogram
- Autocorrelation tempogram
- Musical knowledge (tempo range, continuity)

3. Beat tracking

- Find most likely beat positions
- Exploiting phase information from Fourier tempogram