Tempo, Beat, and Rhythm

Li Su

2019/04/16

Motivation

- Temporal structure is important in music, and it can even exist without melody and harmony
- Intriguing and fun
- MIR goals: detecting the onsets of notes and accents (重音), and tapping the beats like human beings
- Onset detection, tempo estimation, beat tracking, ... 目前算已成熟 難再突破


Tempo, beat, and rhythm

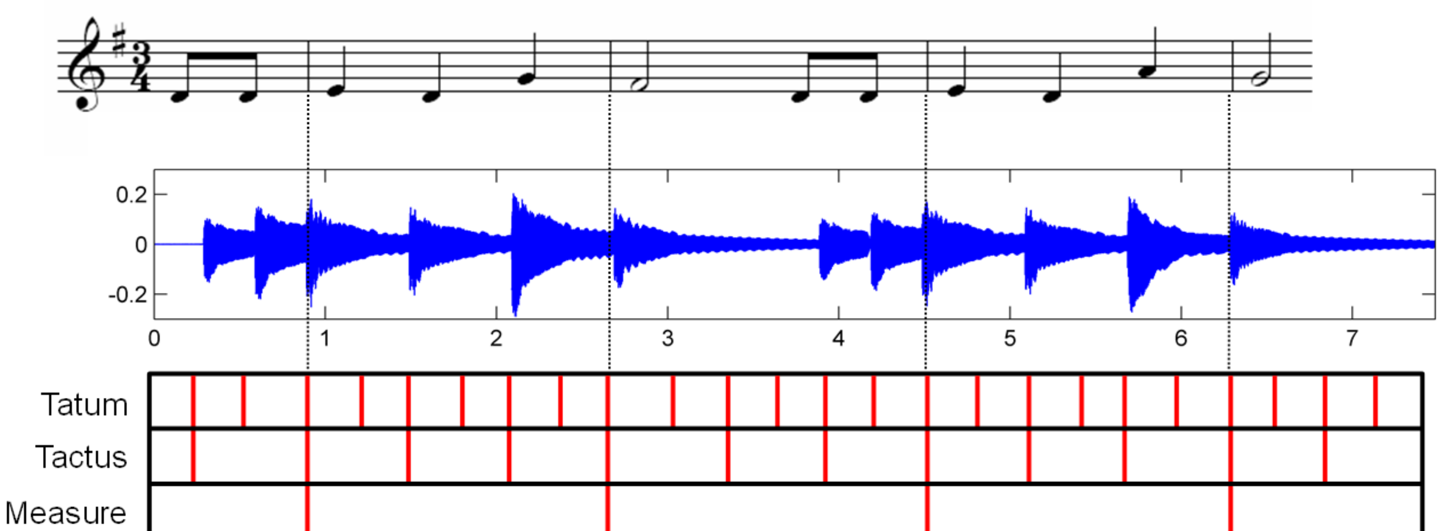
- Tempo: the speed or pace of music
- How to describe tempo?
  - Italian tempo markings: Largo, Andante, Presto, …
  - Note-level markings: e.g., ♩=120
  - Beats per minute (BPM): how many “pulses” per minutes

Beat track = ① beat + ② downbeat

- Beat: the unit of time of music
  - Beat-level: just “beat”
  - Measure-level: downbeat (the first beat in one measure) 每小節第一拍
- Rhythm: a characteristic pattern of beats in different level
  - Meter: a rhythmic structure identifying the relation between downbeats (accent, measure) and beats (e.g., $\frac{2}{4}$, $\frac{3}{4}$, …) 音較可能出現的位置
  - Sometimes you can think rhythm as “genre”
  - Cha Cha, Rumba, Tango, Waltz, …

Tatum, tactus, and measure

- Rhythm is hierarchical: there are various levels presumed to contribute to the human perception of tempo and beat
  - Tatum: the fastest repetition rate of musically meaningful accents 最小單位(最短的音符)
  - Tactus: typically the foot tapping rate (and quarter note level)
  - Measure: typically the rate of one cycle of counting beats
- BPM (beats per minute)



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Tasks and problems

- Tempo/meter analysis (on BPM)
  - BPM = 80, 160, 240 整數倍
  - Same as in pitch detection, tempo also has <u>octave ambiguity</u>
  - "Tempo harmonics" of a song with BPM τ: 2τ, 3τ, …
  - "Tempo subharmonics": $\tau/2$, $\tau/3$, …
  - 某時間點可能有哪些 tempo
  - Basic feature: **tempogram**, an indicator of the local relevance of a specific tempo for a given music recording for each time instance
- Beat tracking and downbeat tracking
  - Backbeat, syncopation
  - 後半拍 切分音
  - Rubato, swing
  - Other genre-/subject-dependent issues
- Rhythm classification

Fourier tempogram (1)

= novelty curve 的 stft

- Given the novelty function $\Delta(m)$ of a musical recording, given a window function $w(m)$, the tempogram , denoted as $F(n,\omega)$, is the STFT of the novelty function:

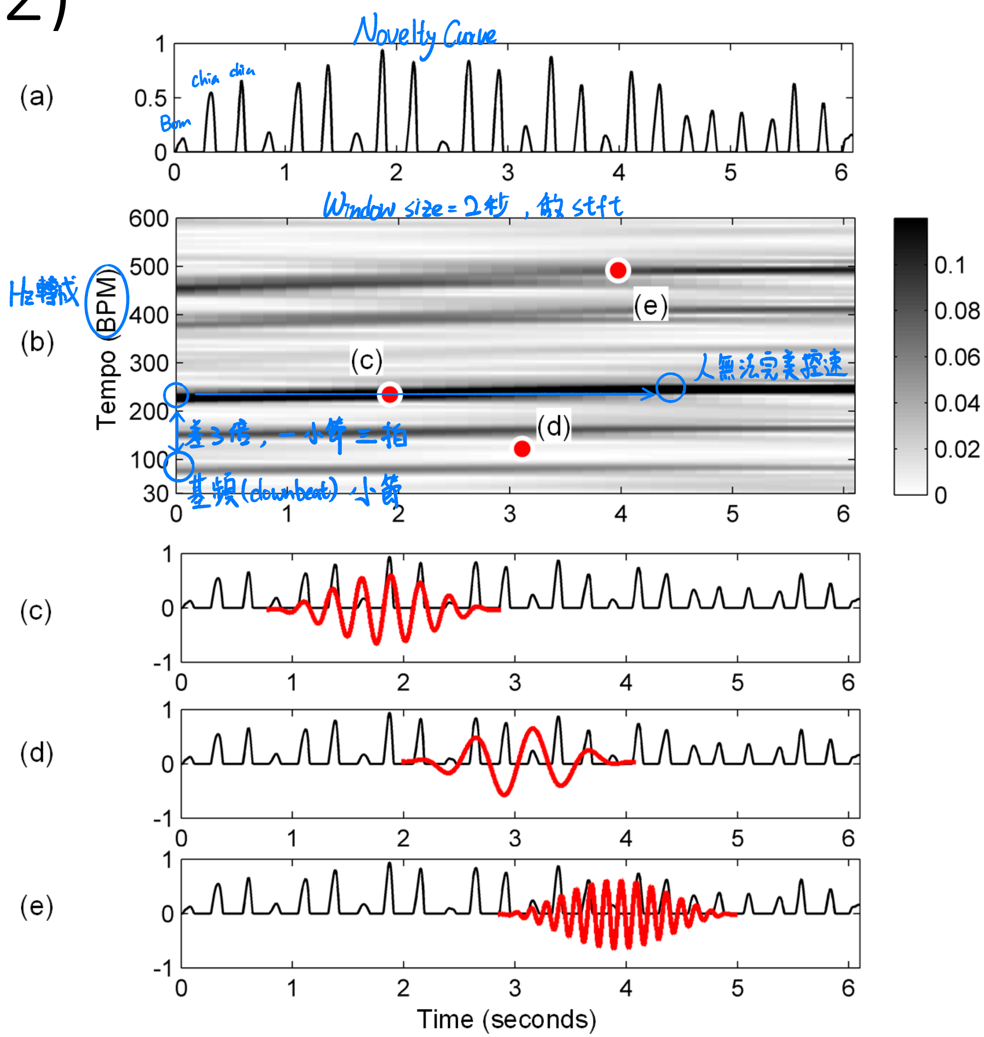
$$F(n,\omega) := \sum_{m \in Z} \Delta(m) w(m-n) e^{-2\pi i \omega n}$$

- Beat rate $\omega = 1/T$ (beats per second)
- BPM: $\tau = 60\omega$
- Discrete Fourier tempogram

$$T^F(n,\tau) = |F(n,\tau/60)|$$

Fourier tempogram (2)

- Example: Shostakovich's Waltz No. 2



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Autocorrelation tempogram (1)

- Recall the issue of finding frequency and finding periodicity in previous lectures in pitch detection
- Alternative way in finding tempo: taking time-varying autocorrelation function (ACF) on the novelty
- Recall ACF:

$$R_{xx}(l) = \sum_{m \in Z} x(m)x(m-l)$$

不一定用stft,亦可用 auto correlation

- Short-time ACF of the novelty curve (windowed ACF):

$$A(n,l) := \sum_{m \in Z} \Delta(m)w(m-n)\Delta(m-l)\,w(m-n-l)$$

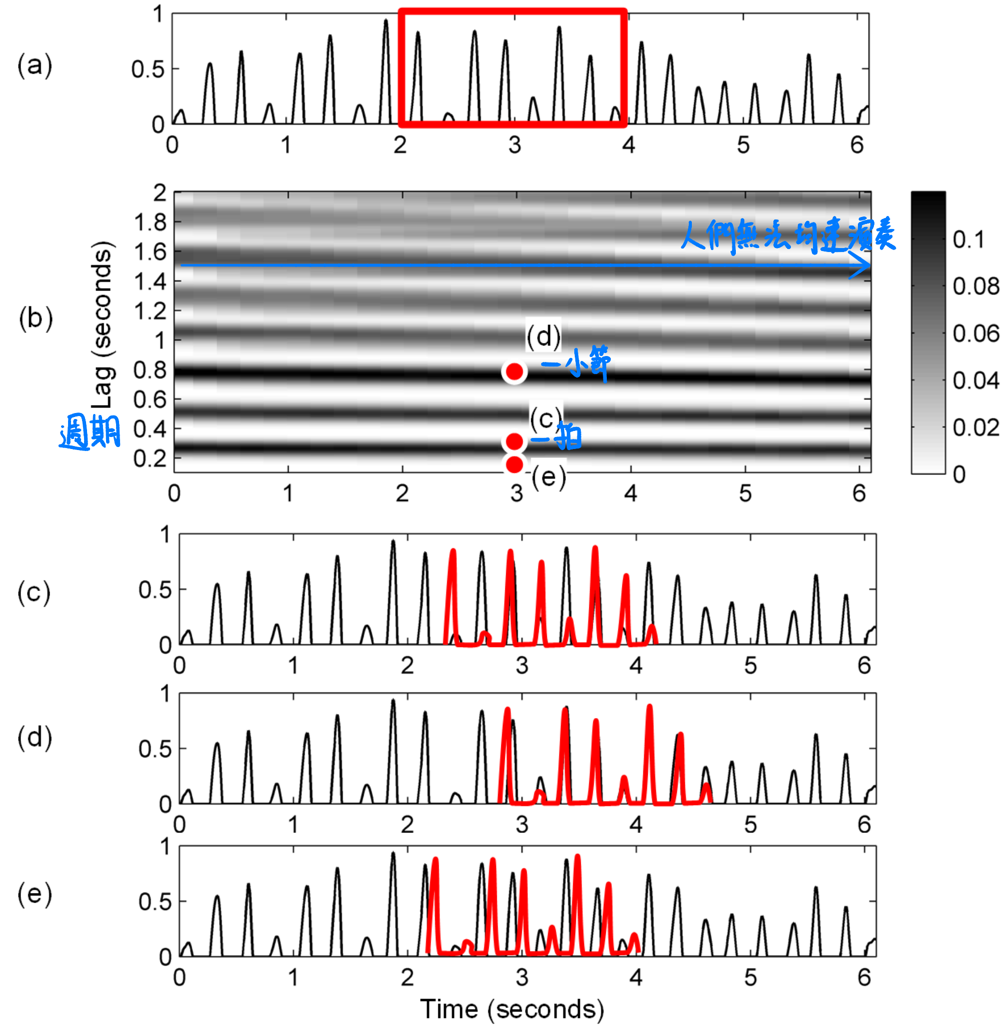
- Alternative way: $A(n,l) = IFFT\{|FFT\{w(m)\Delta(m)\}|^2\}$

Autocorrelation tempogram (2)

- ACF tempogram

$$T^A(n,\tau) := A(n,l)$$

- BPM $\tau = 60/(rl)$
- r: the “sampling period” of the novelty curve
- Comparison
  - STFT: Fourier basis
  - ACF: self basis
- Example: Shostakovich’s Waltz No. 2



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

wave → stft → spectrogram → spectral flux
↓
tempogram ← stft/auto correlation ← novelty curve

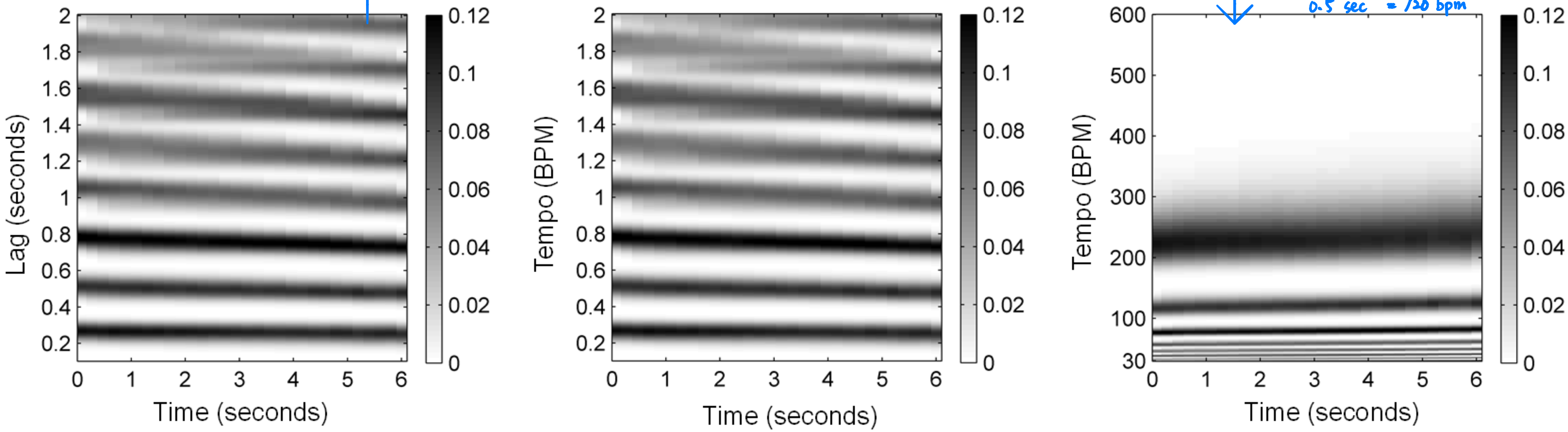
stft : 瞬時 (23 ms) 狀態
spectral flux : onset detection, novelty curve 的 fs 由 stft 的 hop 決定
stft : window size 要能 cover 數拍 (4~8 sec)
要抓 downbeat 要 cover 數小節 (12~ sec)
by Li Su

Autocorrelation tempogram (3)

- Conversion from lag to tempo
  - From time-lag representation to time-frequency representation
- Example:

Interpolation 內差
or filterbank

lag 1 second = 60 bpm
1.2 sec = 50 bpm
0.5 sec = 120 bpm



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Global and local tempo

- For a music recording with N time instances:
  - Average (global) tempo: let Θ the expected interval of tempo

$$T_{\text{Average}}(\tau) \coloneqq \frac{1}{N} \sum_{n \in [1,N]} T(n, \tau)$$

$$\hat{\tau} \coloneqq \max \left\{ T_{\text{Average}}(\tau) \middle| \tau \in \Theta \right\}$$

Cyclic tempogram (1)

來處理 octave ambiguity 看看就好

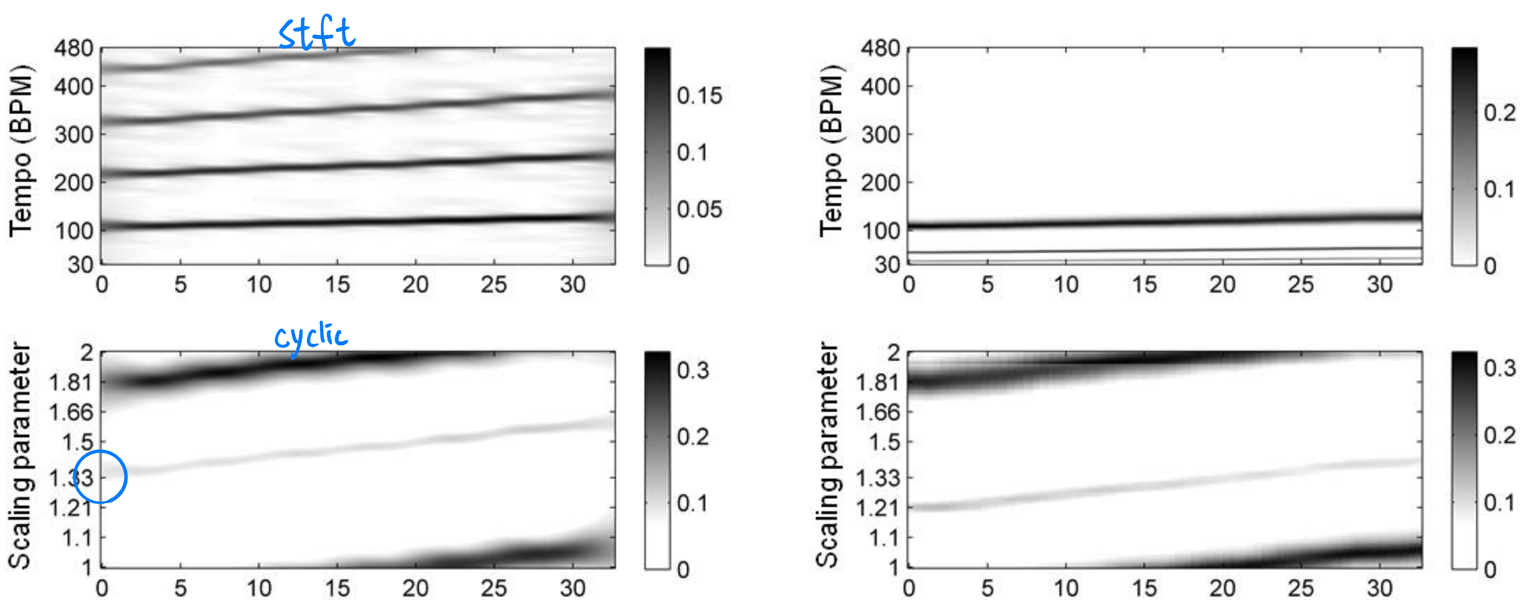
- Motivation: like the pitch scale, the tempo scale also has "octave equivalence" 像 chroma，把 octave 合併起來
  - A song with BPM=120 can usually be interpreted as BPM=240 (or BPM=60)
- Define $[\tau] := 2^k\tau, k \in Z$
  - Example: for $\tau = 120$ one obtains $[\tau] = \{\dots, 30, 60, 120, 240, 480, \dots\}$
- The cyclic tempogram

$$C(n, [\tau]) := \sum_{\lambda \in [\tau]} T(n, \lambda)$$

- The cyclic tempogram referred to τ_0: $C_{\tau_0}(n, s) := C(n, [s, \tau_0])$
  - s: scaling parameter
  - Notice that $C_{\tau_0}(n, s) := C_{\tau_0}(n, 2^k s)$ for $k \in Z$

Cyclic tempogram (2)

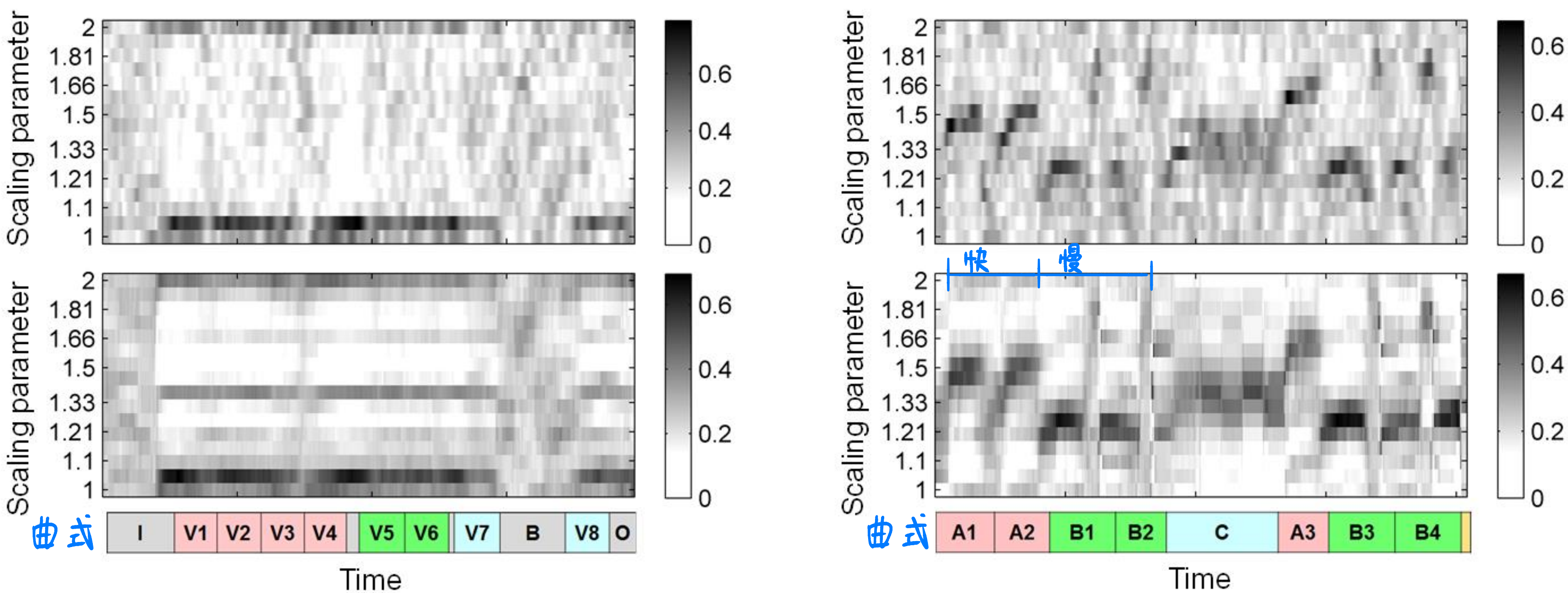
- Example: a pulse sequence with increasing tempo
- 110 to 130 BPM
- Left: Fourier tempogram
- Right: ACF tempogram



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Tempogram and musical structure

- Fourier tempogram (top) and ACF tempogram (bottom)
- “In the year 2525” (Zager and Evans)
- Hungarian Dance No.5 (Johannes Brahms)



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Comparison

對 novelty curve 做

Fourier tempogram	Autocorrelation tempogram
Comparison of novelty curve with windowed sinusoids with each sinusoid representing a tempo.	Comparison of novelty curve with time-shifted windowed sections of itself with each lag representing a tempo.
Conversion of frequency (Hertz) into tempo (BPM).	Conversion of lag (seconds) into tempo (BPM).
Measurement of novelty periodicities.	Measurement of novelty self-similarities.
Emphasis of tempo harmonics. 很容易高估成2,3,4倍 tempo	Emphasis of tempo subharmonics. 很容易低估成½, ⅓, ¼ tempo
Suitable to analyze tempo on tactus and tatum level.	Suitable to analyze tempo on tactus and measure level. or phrase

From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Cyclic tempogram and chromagram

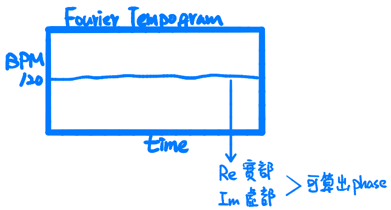
少被使用，看看即可

Steps	Chromagram	Cyclic tempogram
1.	Analysis of waveform.	Analysis of novelty curve.
2.	Computation of spectrogram.	Computation of tempogram.
3.	Usage of log-frequency axis.	Usage of log-tempo axis.
4.	Cyclic projection.	Cyclic projection.

From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Beat tracking by PLP (1)

- Tempo estimation: identifying the frequency (or periodicity) of the novelty curve
- Beat tracking: identifying the phase of the novelty curve at the local tempo frequency
- Predominant local pulse (PLP)



- Local tempo estimation at time n

$$\tau_n := \operatorname{argmax}_{\tau\in\Theta} T^F(n,\tau)$$

- Local phase estimation: the phase ϕ_n belonging to the windowed sinusoid of tempo τ_n

$$\phi_n = \frac{1}{2\pi}\arccos\left(\frac{Re\left(F(n,\frac{\tau_n}{60})\right)}{\left|F(n,\frac{\tau_n}{60})\right|}\right)$$

Beat tracking by PLP (2)

- The optimal window sinusoid "matching" the local novelty curve can be described by a sinusoid in phase with the novelty curve

$$\kappa_n(m) := w(m-n)\cos\left(2\pi\left(\frac{\tau_n}{60}\cdot m - \phi_n\right)\right)$$

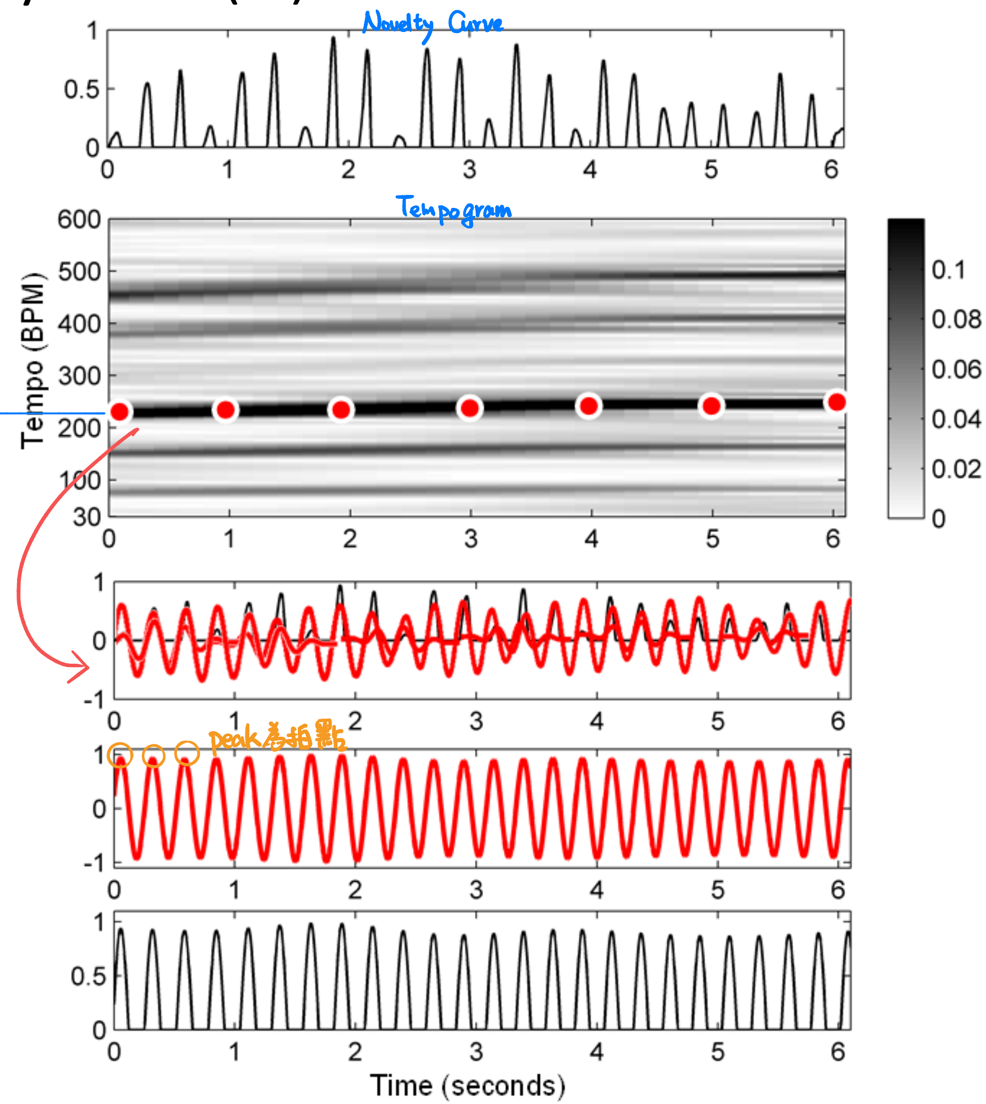
- The PLP function: applying the overlap-add (OLA) technique and reconstruct the sinusoids
  - Picking only positive values

$$\Gamma(m) = \left|\sum_{n\in Z}\kappa_n(m)\right|_{\geq 0}$$

Beat tracking by PLP (3)

投影到Cos(220)的投影量大

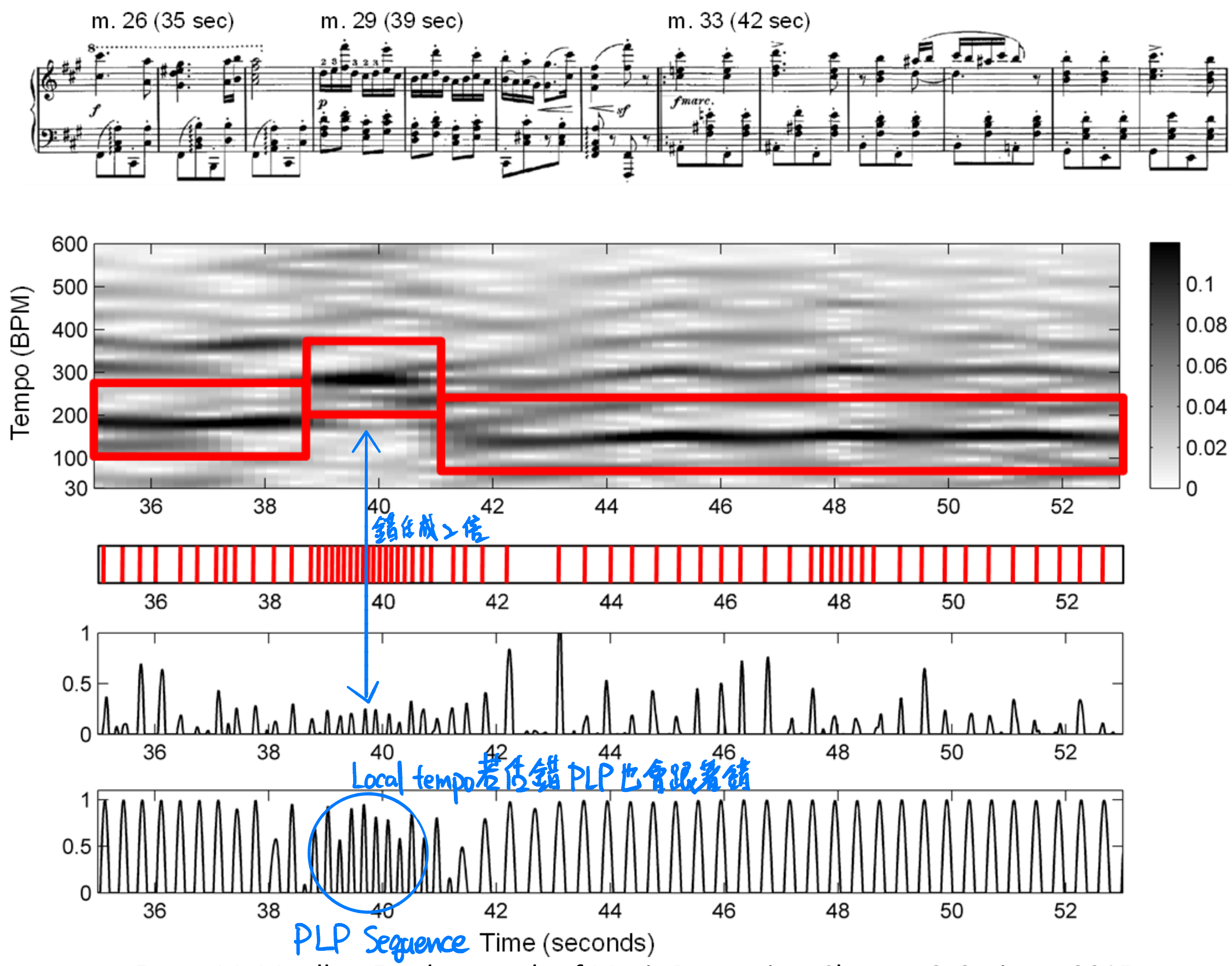
把這條線重建回時間軸
(即 inverse FT)



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

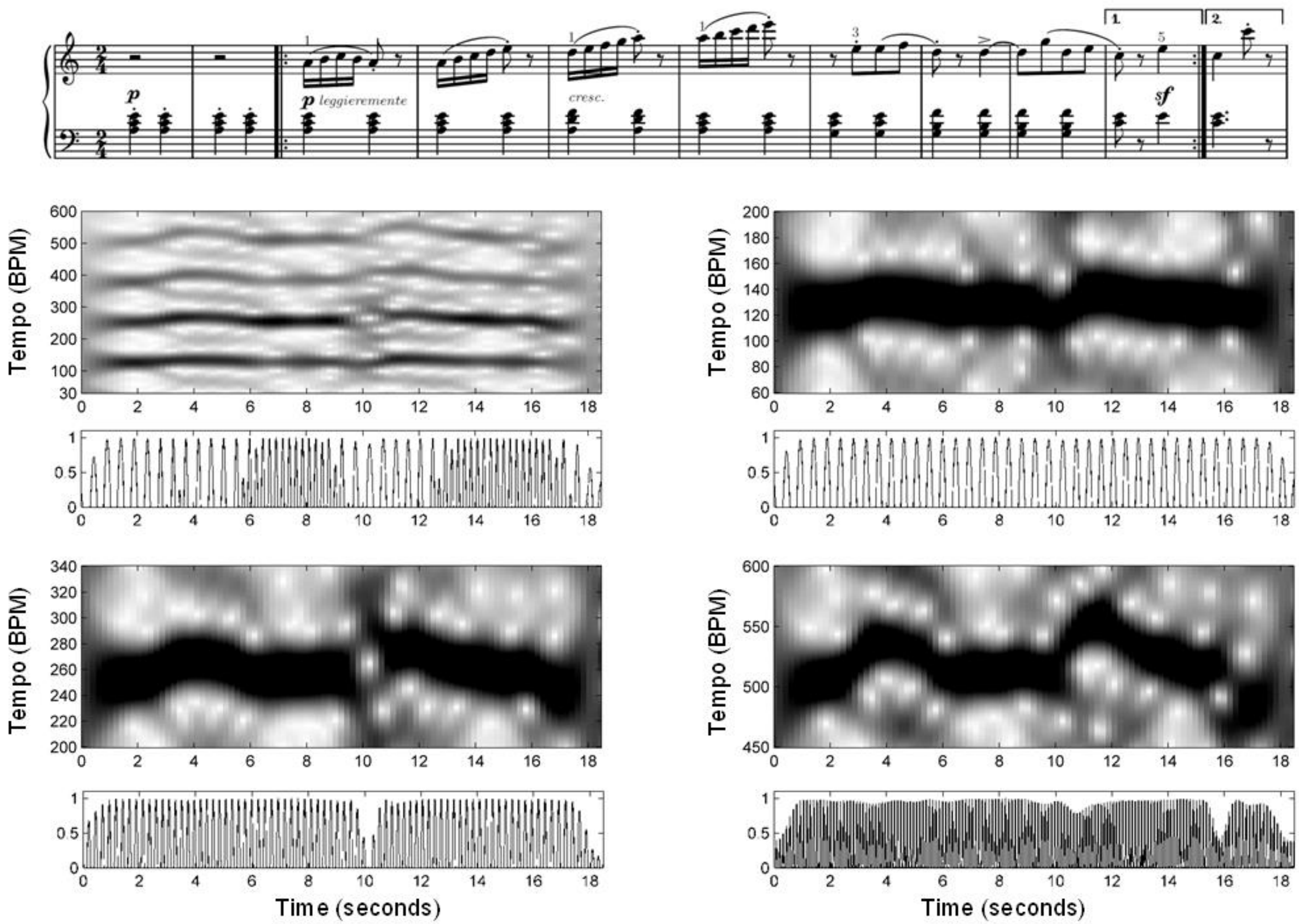
Example of PLP

- Hungarian Dance No.5 (Johannes Brahms)



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

PLP extracted in different bands

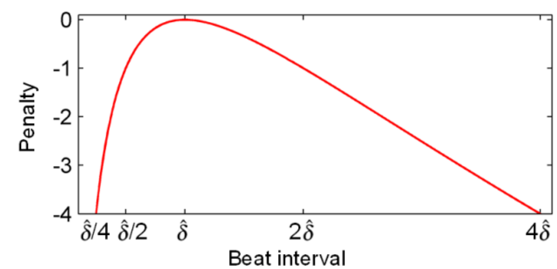


From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Dynamic programming for beat tracking

(Implemented in librosa)
需給 tempo

- Assumption: the tempo of music is more or less constant
- A beat sequence $B = [b_1, b_2, \ldots, b_L]$, estima



- Penalty function

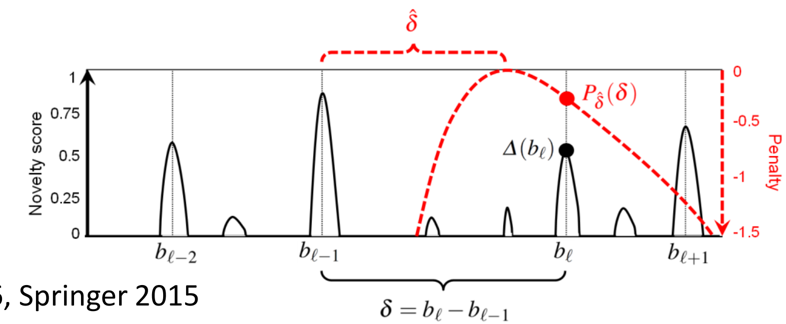
$$P_{\hat{\delta}}(\delta) := -\left(\log_2(\delta/\hat{\delta})\right)^2$$

- The objective function

$$S(B) := \sum_{l=1}^{L} \underbrace{\Delta(b_l)}_{\text{novelty curve 的 peak}} + \lambda \sum_{l=2}^{L} P_{\hat{\delta}}(b_l - b_{l-1})$$

$\hat{\delta}$ 估出來的 tempo，如果 $b_l - b_{l-1} \approx \hat{\delta}$，則給 reward
否則給 penalty

- Let B^N all possible sequence, optimize $B^* := \mathrm{argmax}_{B \in B^N} S(B)$



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 6, Springer 2015

Algorithm

Algorithm: OPTIMAL BEAT SEQUENCE

Input: Novelty function $\Delta : [1:N] \to \mathbb{R}$
Estimate $\hat{\delta}$ for the beat period (given in samples)
Weight parameter $\lambda \in \mathbb{R}$

Output: Optimal beat sequence $B^* = (b_1, b_2, \ldots, b_L)$

Procedure: Initialize $\mathbf{D}(0) = 0$ and $\mathbf{P}(0) = 0$. Then compute in a loop for $n = 1, \ldots, N$:

$$\mathbf{D}(n) = \Delta(n) + \max\left\{0, \ \max_{m \in [1:n-1]} \{\mathbf{D}(m) + \lambda P_{\hat{\delta}}(n-m)\}\right\}$$

If $\mathbf{D}(n) = \Delta(n)$ then set $\mathbf{P}(n) = 0$,
otherwise set $\mathbf{P}(n) = \mathrm{argmax}_{m \in [1:n-1]} \left\{\mathbf{D}(m) + \lambda P_{\hat{\delta}}(n-m)\right\}$

Set $\ell = 1$ and $a_\ell = \mathrm{argmax}_{n \in [0:N]} \mathbf{D}(n)$. Then repeat the following steps until $\mathbf{P}(a_\ell) = 0$:

Increase ℓ by one.
Set $a_\ell = \mathbf{P}(a_{\ell-1})$.

If $a_\ell = 0$, then set $L = 0$ and return $B^* = \emptyset$.
Otherwise let $L = \ell$ and return $B^* = (a_L, a_{L-1}, \ldots, a_1)$.

Recurrent neural networks for beat tracking

annotation 較多
較易取得

- Input: audio features
- Output: the likelihood being beat and downbeat (classification-based output)
- http://www.cp.jku.at/research/papers/Boeck_etal_ISMIR_2016.pdf
- Sebastian Böck, Florian Krebs and Gerhard Widmer, *Joint Beat and Downbeat Tracking with Recurrent Neural Networks*, Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR), 2016.
- Böck, Sebastian, et al. "Madmom: A new Python audio and music signal processing library." *Proceedings of the 2016 ACM on Multimedia Conference*. ACM, 2016.

建議瞧瞧

will be used in HW2

Remaining Challenge:
目前只能處理給定 meter (幾拍子)
難處理 dynamic 情況
副拍易難 (拍子的 hierarchy)

Further readings

- Tempogram toolbox
- Beat tracking by dynamic programming (Dan Ellis' code)
  - 經典
  - Dan Ellis, "Beat Tracking by Dynamic Programming," *J. New Music Research*, Special Issue on Beat and Tempo Extraction, vol. 36 no. 1, March 2007, pp. 51-60.
  - http://labrosa.ee.columbia.edu/projects/beattrack/
- Beat tracking with dynamic Bayesian networks
  - A. Srinivasamurthy, A. Holzapfel, A. Cemgil, and X. Serra. Particle filters for efficient meter tracking with Dynamic Bayesian networks. In ISMIR 2015.
- Downbeat tracking with neural networks
  - S. Durand, J. P Bello, B. David, and G. Richard, "Downbeat tracking with multiple features and deep neural networks," in ICASSP 2015, pp. 409–413.