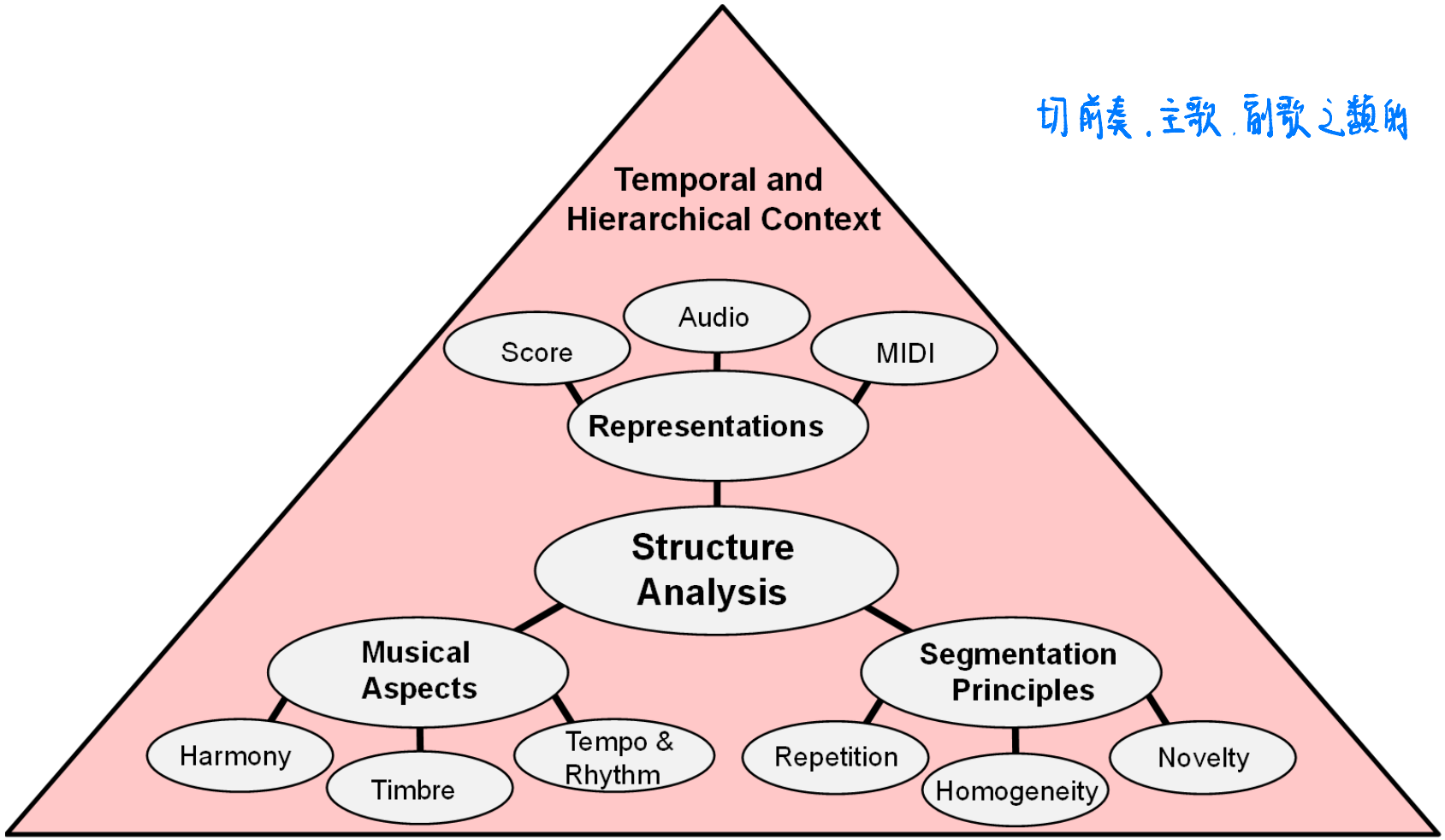
Structural segmentation

Li Su

2019/05/07

The structure of music 音樂結構分析

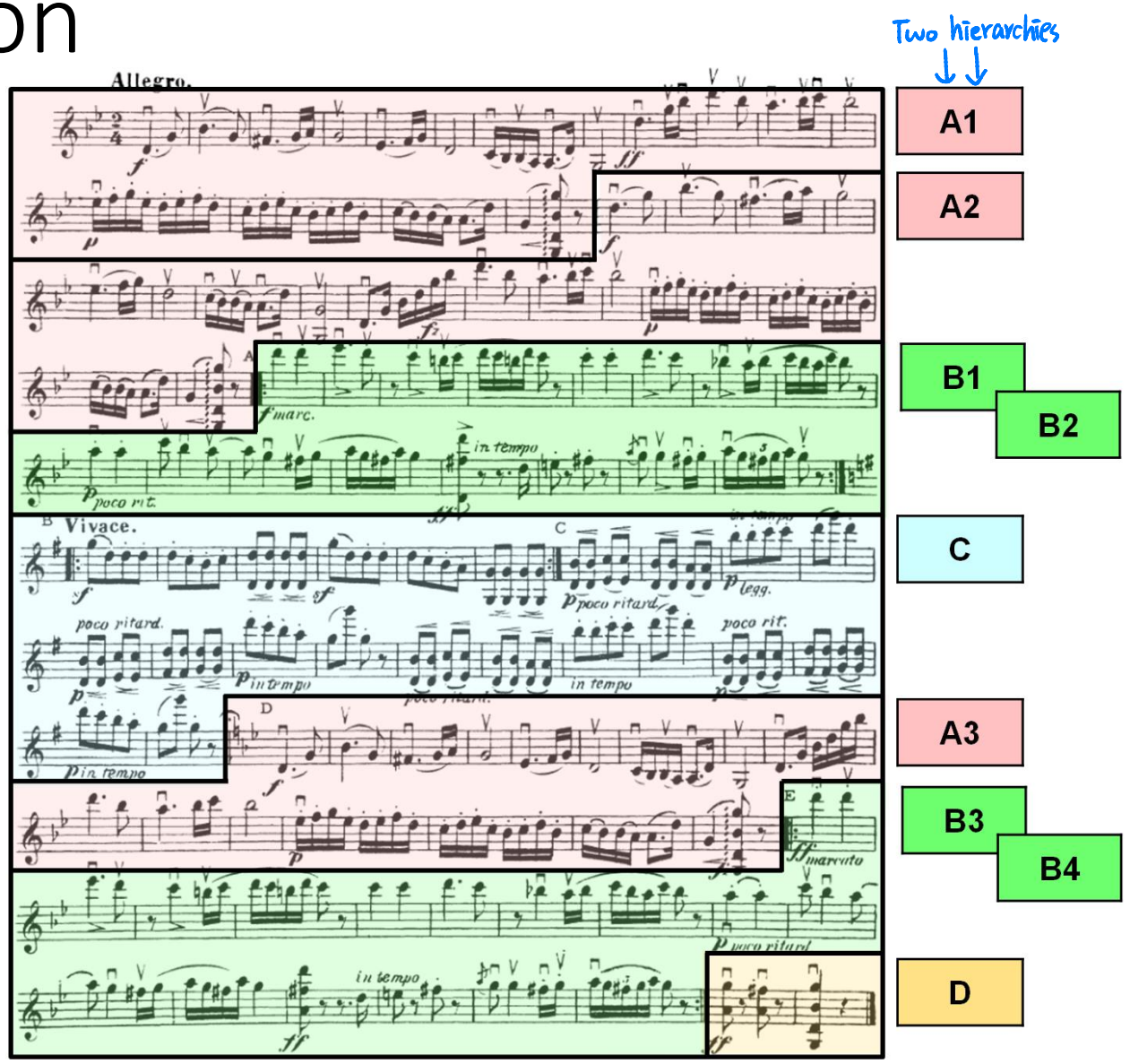


切前奏、主歌、副歌之類的

From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Introduction

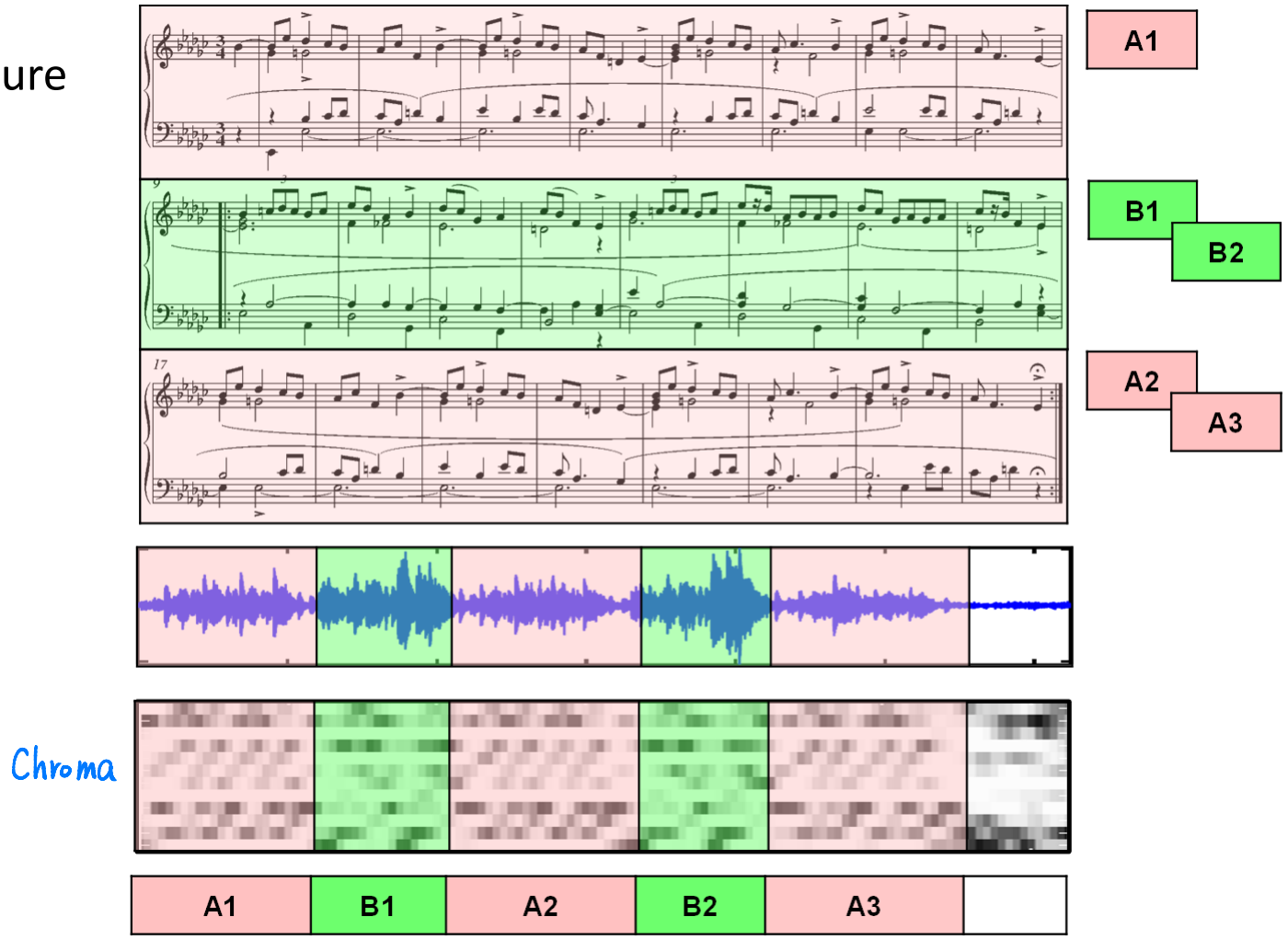
- The structure of Brahms' Hungarian Dance no.5



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Introduction

- Musical structure and signals



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Structure segmentation

- General goal:
  - **Divide** an audio recording into temporal segments corresponding to musical parts
  - **Group** these segments into musically meaningful categories
- Examples
  - *Intro*, *verse*, *chorus*, *bridge*, *outro* of a pop song
  - *Exposition*, *development*, *recapitulation*, *coda* of a sonata

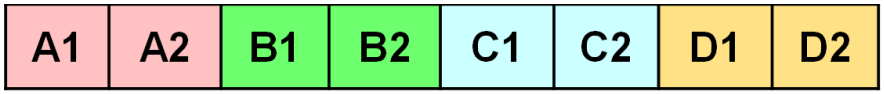
Types of musical structures

MIREX 有 structure segmentation 比賽（目前做得不好）

Strophic form

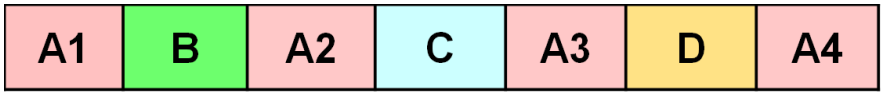
Sonata form 奏鳴曲式

Chain form



"Tell Me Why" by Beatles

Rondo form 輪旋曲



"Yesterday" by Beatles

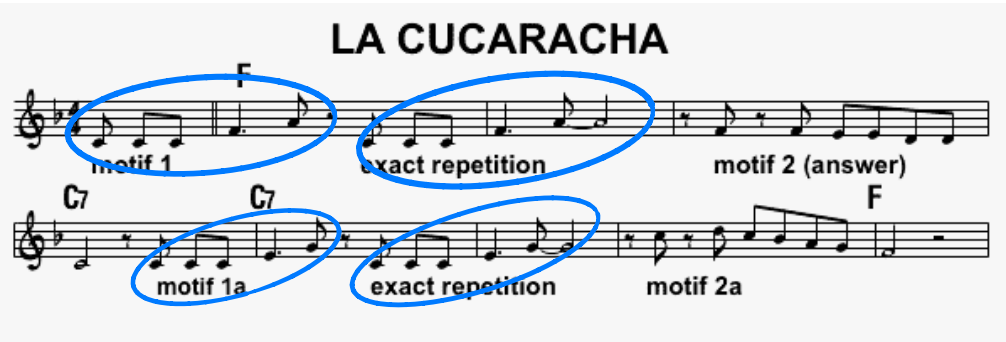
From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Music segmentation (1)

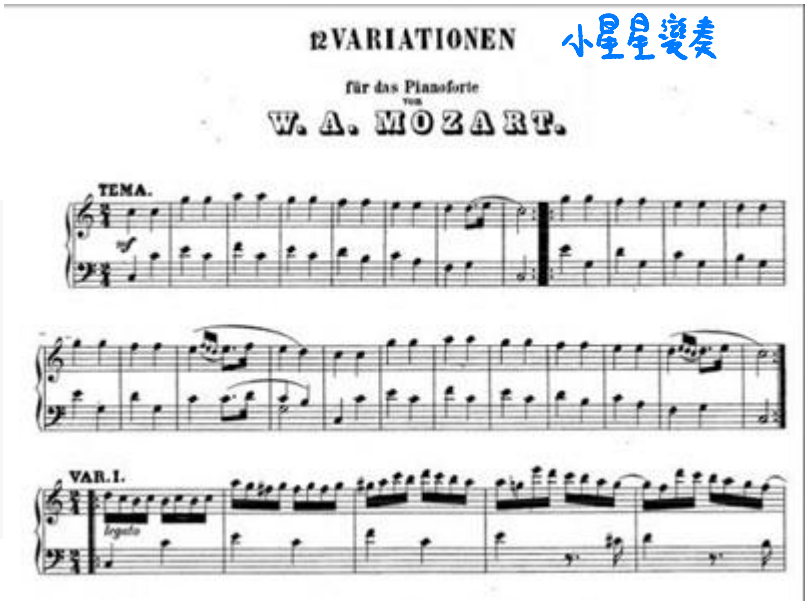
中型結構

- Motifs, riffs, phrases, themes, and variations…
  - A motif is a recurring pattern in a writing work that helps to strengthen the theme
  - A riff is a short, repeated, memorable musical phrase (or melody figure)

找重覆出現不難，找有義意的重覆不容易

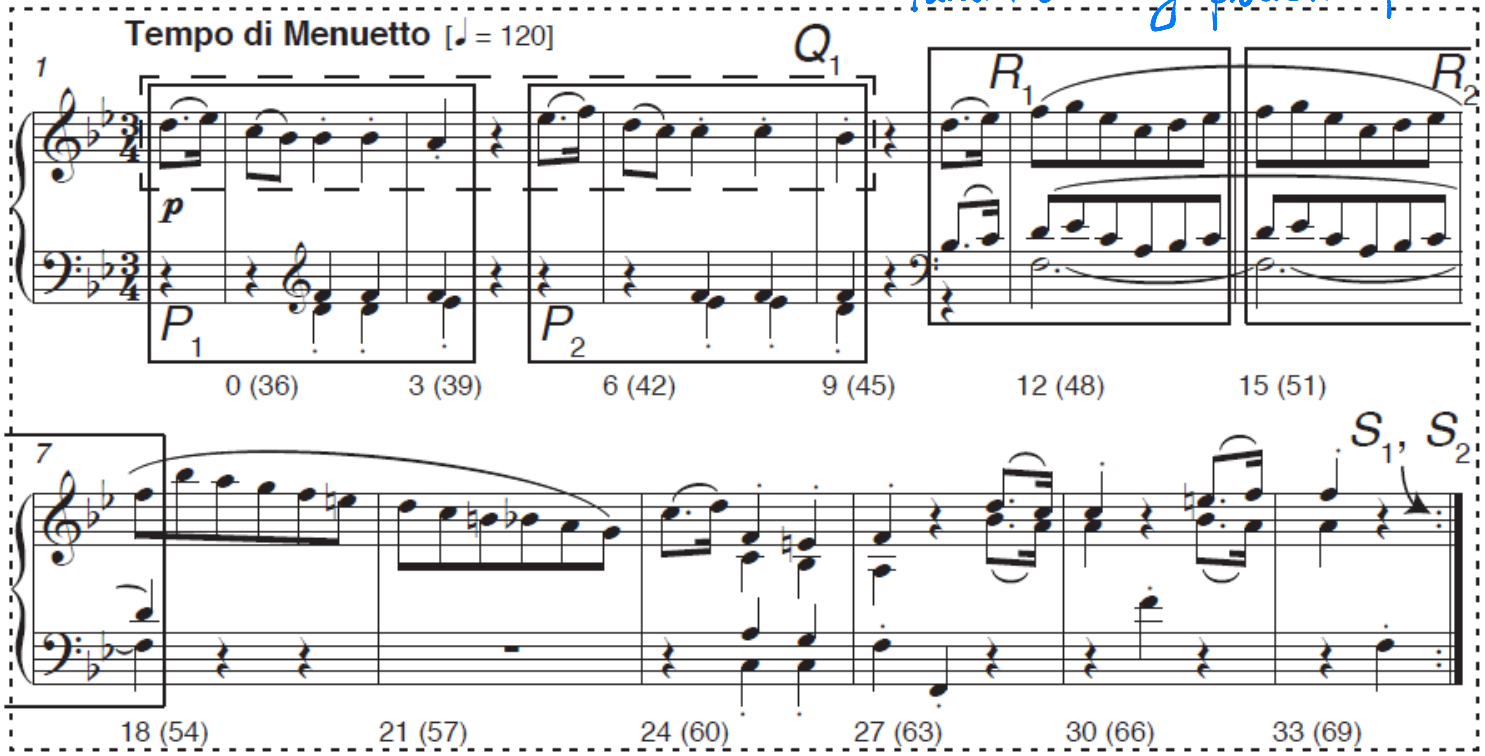


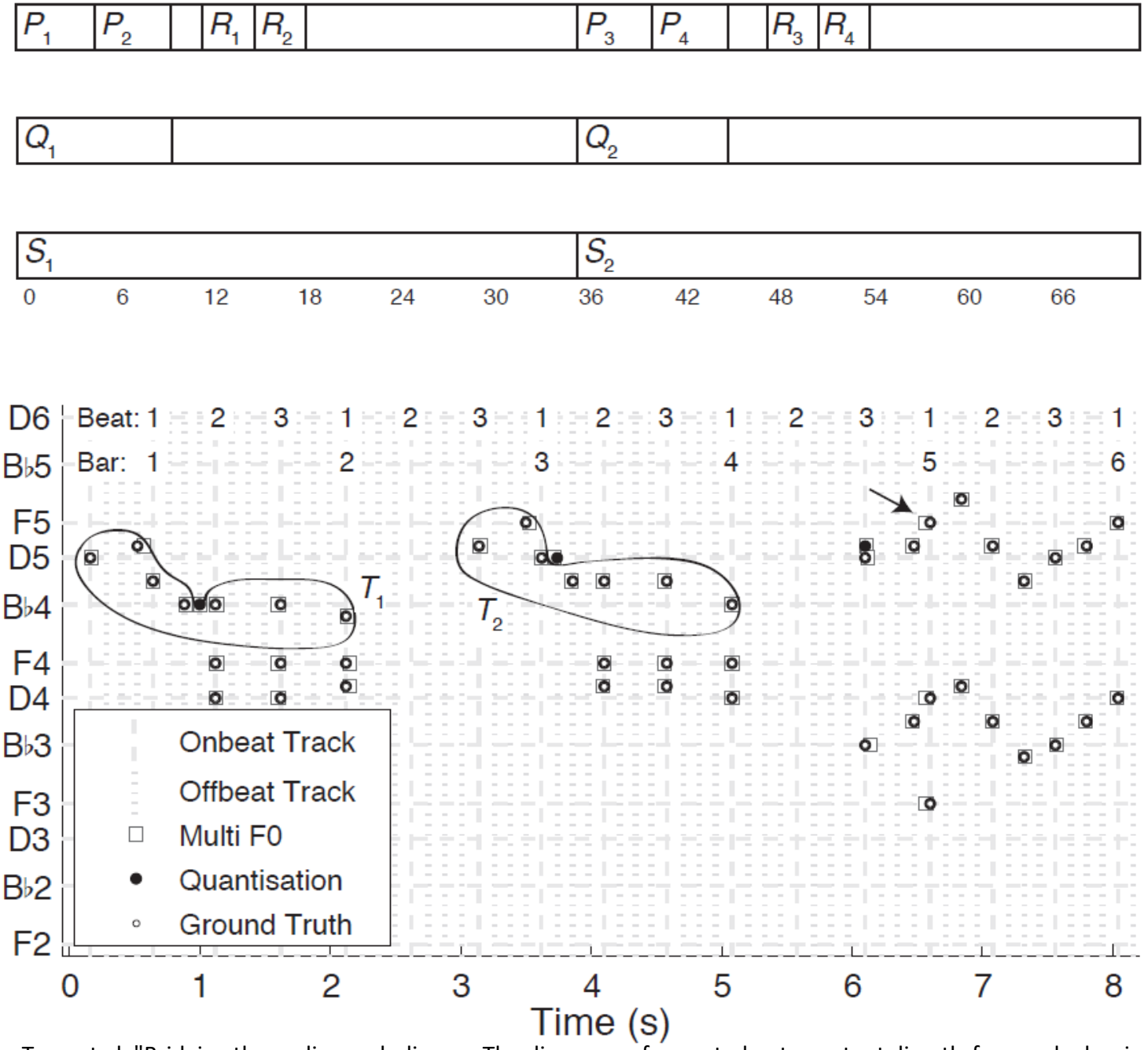
https://tamingthesaxophone.com/composition-motif



Music segmentation (2)

- The key to bridge the audio-to-semantic gap
- Example: Mozart's Piano Sonata no.4 in E-flat major K282 by Friedrich Gulda



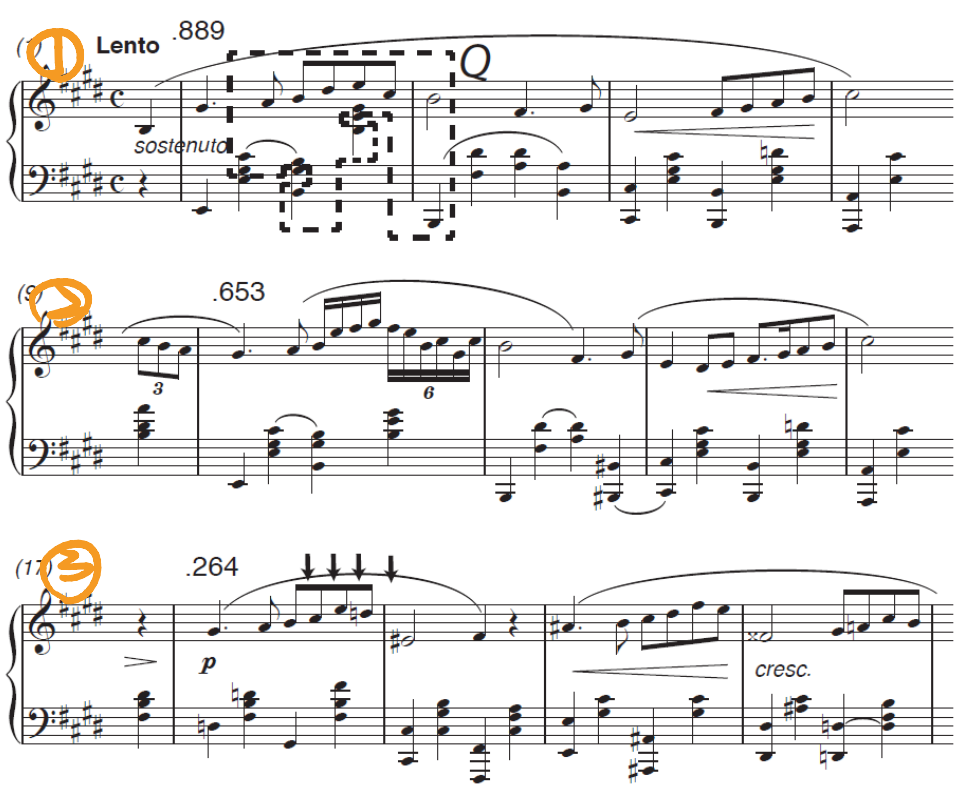


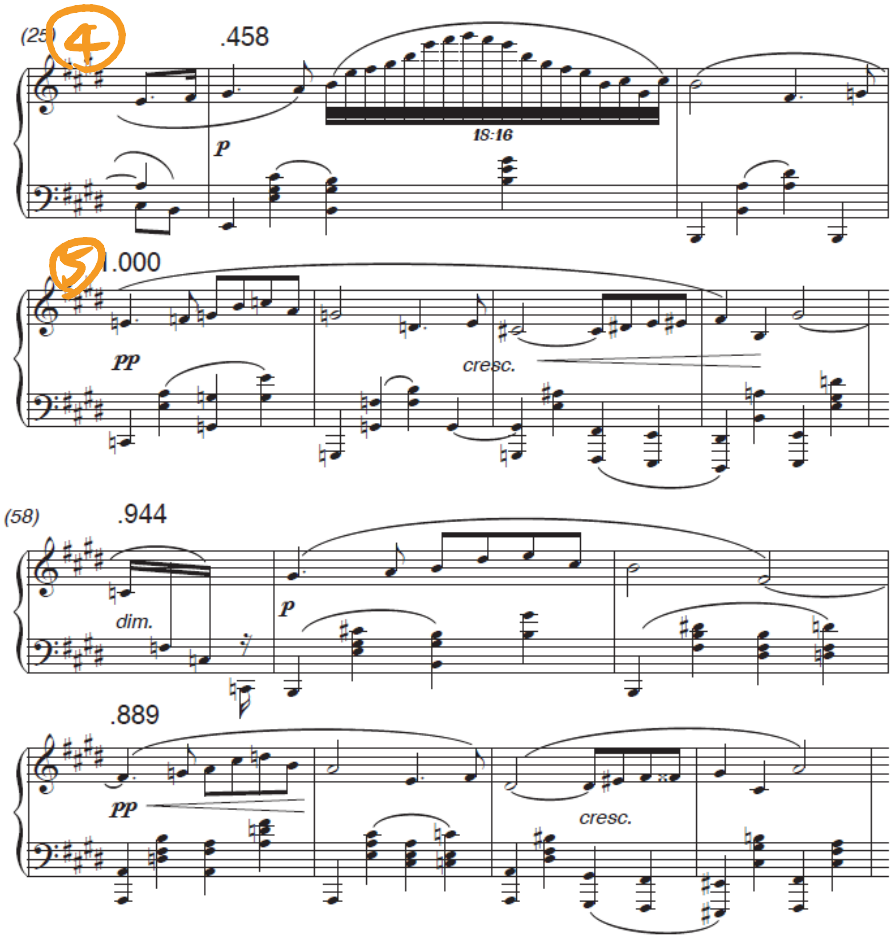
Collins, Tom, et al. "Bridging the audio-symbolic gap: The discovery of repeated note content directly from polyphonic music audio." *Audio Engineering Society Conference: 53rd International Conference: Semantic Audio*. Audio Engineering Society, 2014.

Challenging case

重點出現好多次，但每次呈現都差很多

- Chopin's Nocturne Op. 62, No. 2 in E major





Collins, Tom, et al. "SIARCT-CFP: Improving Precision and the Discovery of Inexact Musical Patterns in Point-Set Representations." *ISMIR*. 2013.

Tasks

- **Segmentation**
  - The process of **dividing** a given audio stream into acoustically meaningful segments
  - Notes > Beat intervals > Sentences (4 or 8 measures) > Sections
- **Structure analysis**
  - Find the relations between the segments
- **Structure labeling**
  - Computational music analysis

MIREX competition

目前performance低

- Structural Segmentation 常常切太小段
- Discovery of Repeated Themes & Sections

找曲子的動機(未必週期性出現)

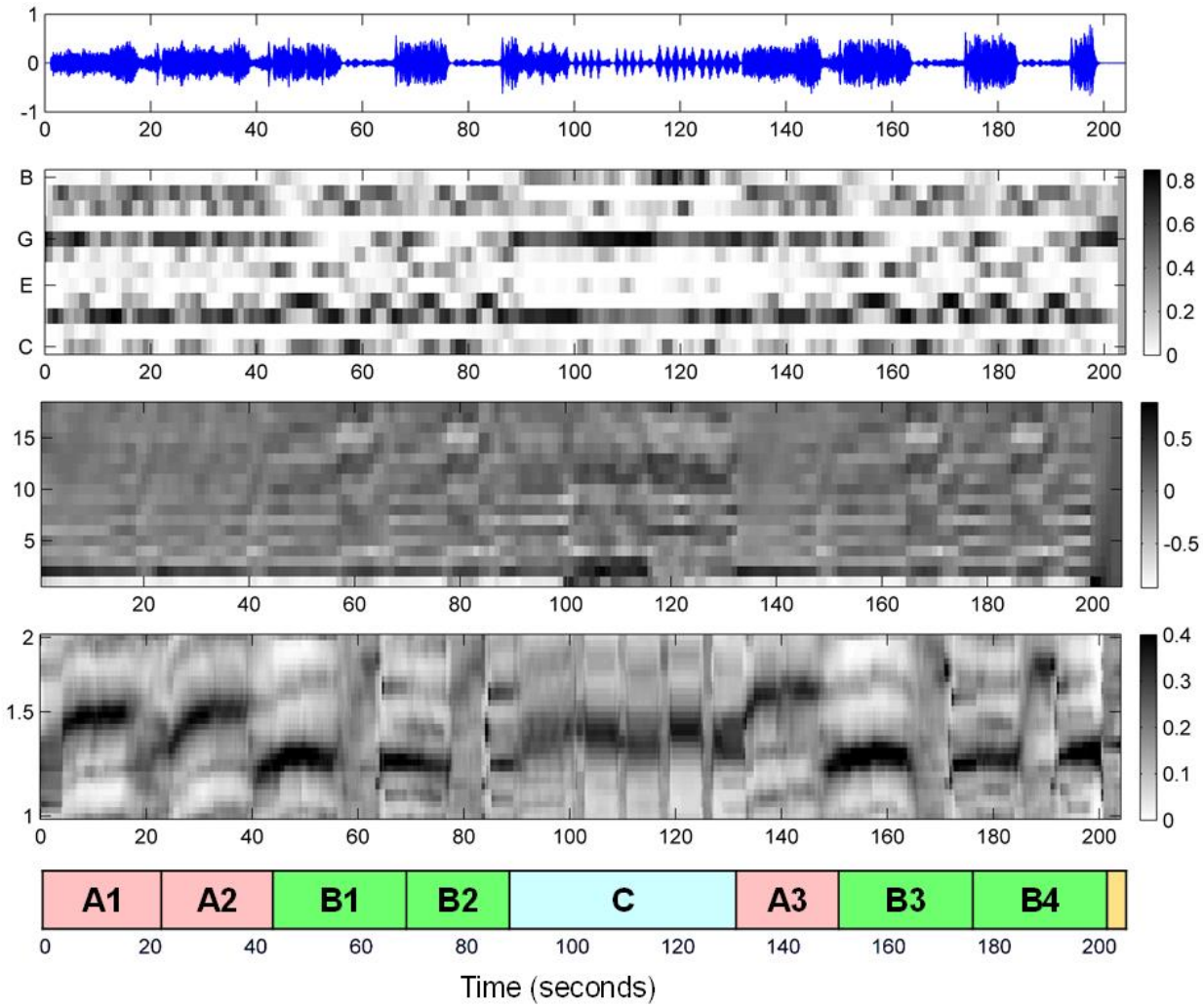
songle.jp

可思考怎麼唱合切歌問題

General principle

- In terms of signal processing, we wish to find
- 1) Novelty 訊號跟過盤不同，有新東西
- 2) Homogeneity
- 3) Repetition
- For a music signal in a macro scale

Feature representation



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Self-similarity matrix (SSM)

- Given a sequence of feature vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]$
- Self-similarity matrix (SSM) $\mathbf{S} \in R^{n \times n}$:

$$\mathbf{S}(i,j) = s(\mathbf{x}_i, \mathbf{x}_j), \quad i,j \in [1:n]$$

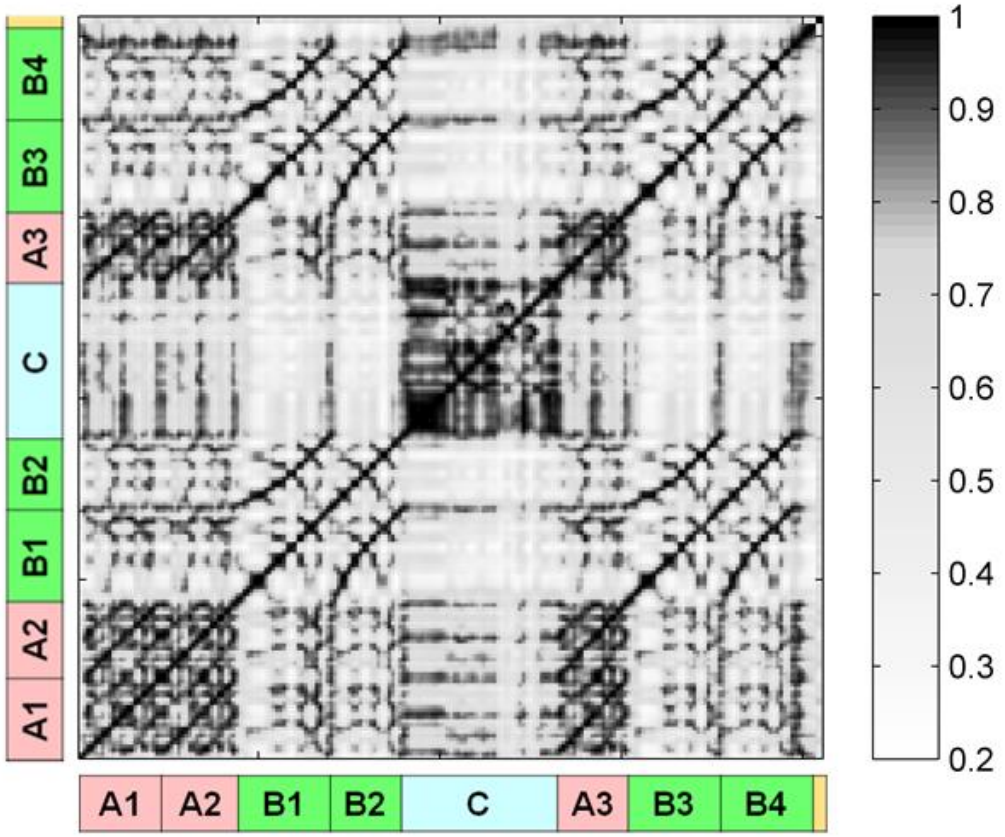
- Cosine-similarity

可試試別的metric

$$s(\mathbf{x}_i, \mathbf{x}_j) = \frac{\mathbf{x}_i \cdot \mathbf{x}_j}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|}$$

Self-similarity matrix (SSM)

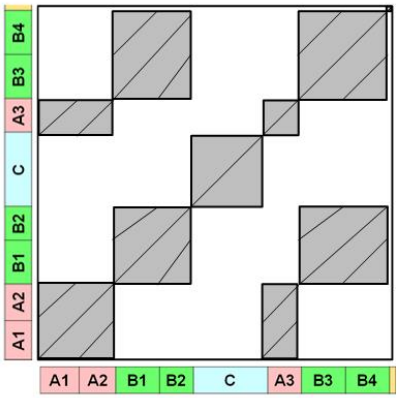
- Cosine similarity of the ith and the jth frames

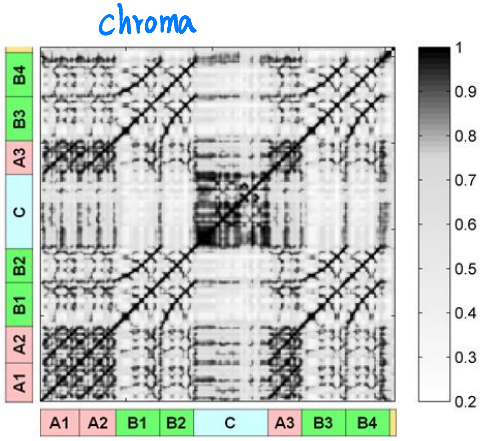


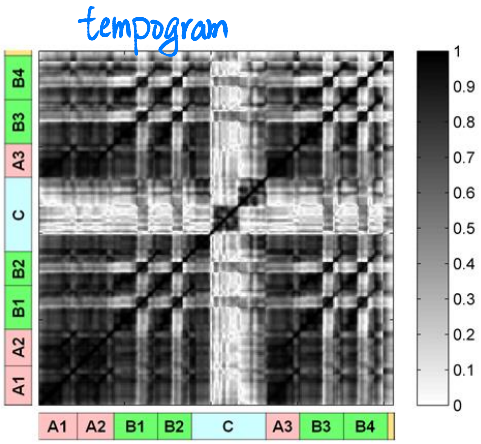
From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

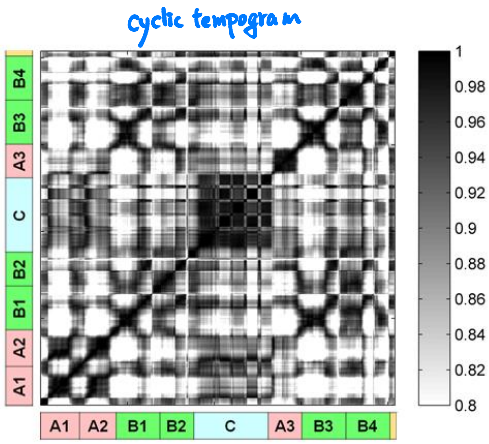
Self-similarity matrix (SSM)

也可用 mfcc





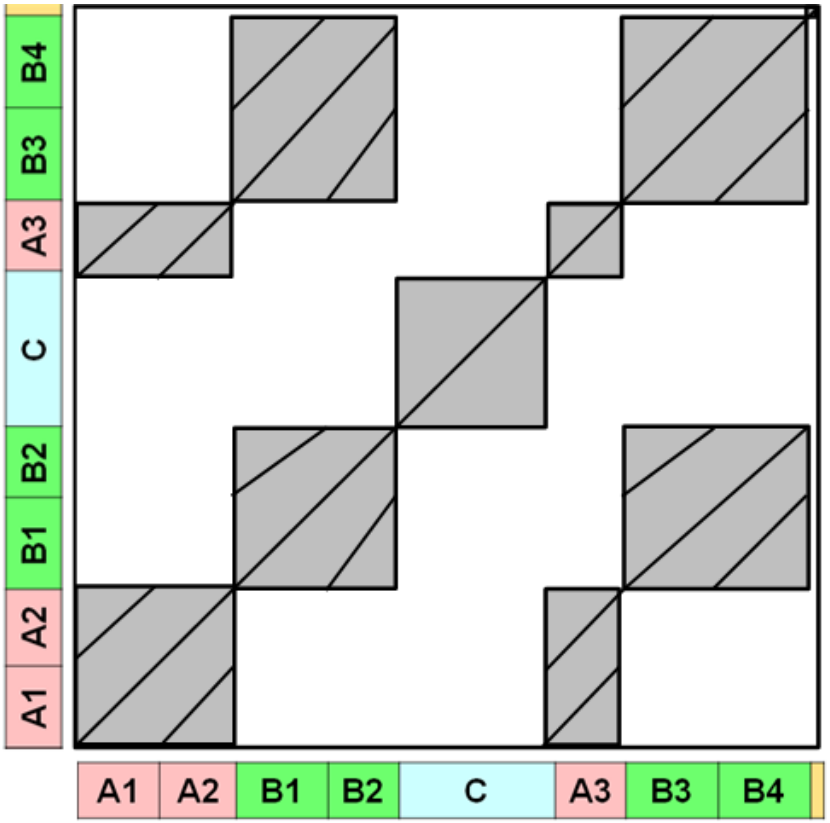




From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Self-similarity matrix (SSM)

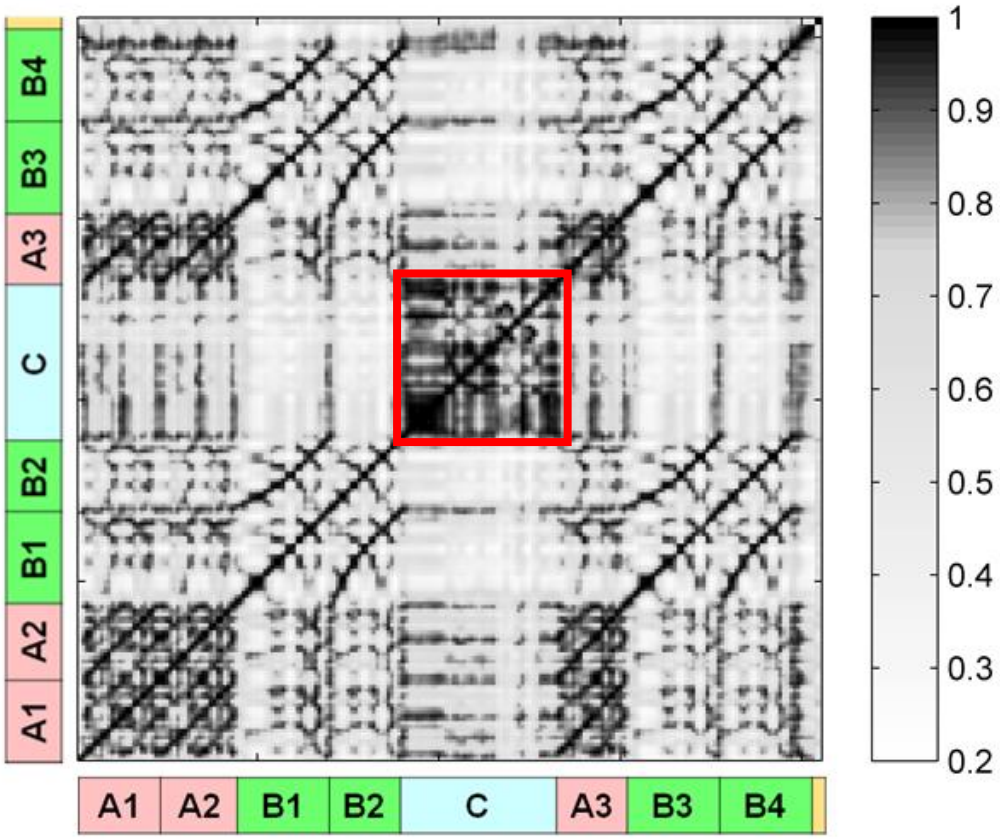
- How to interpret the meaning of a SSM?



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

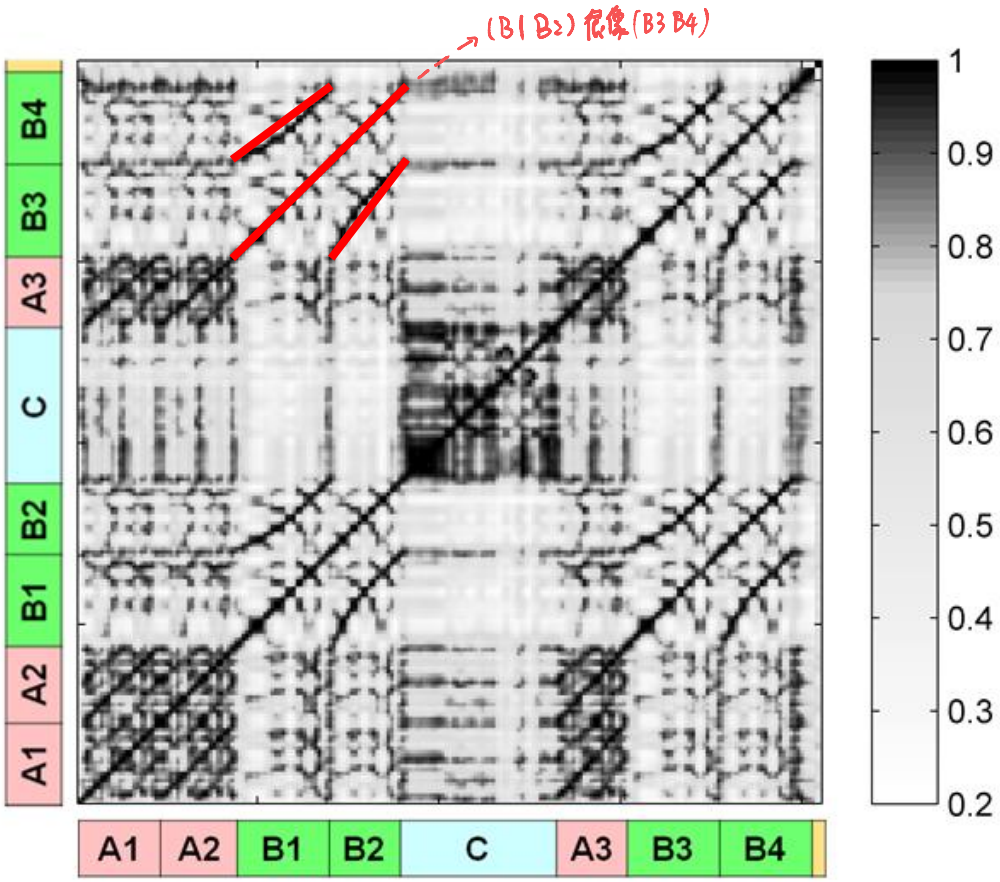
Blocks: homogeneity

自我重複，跟別人不像



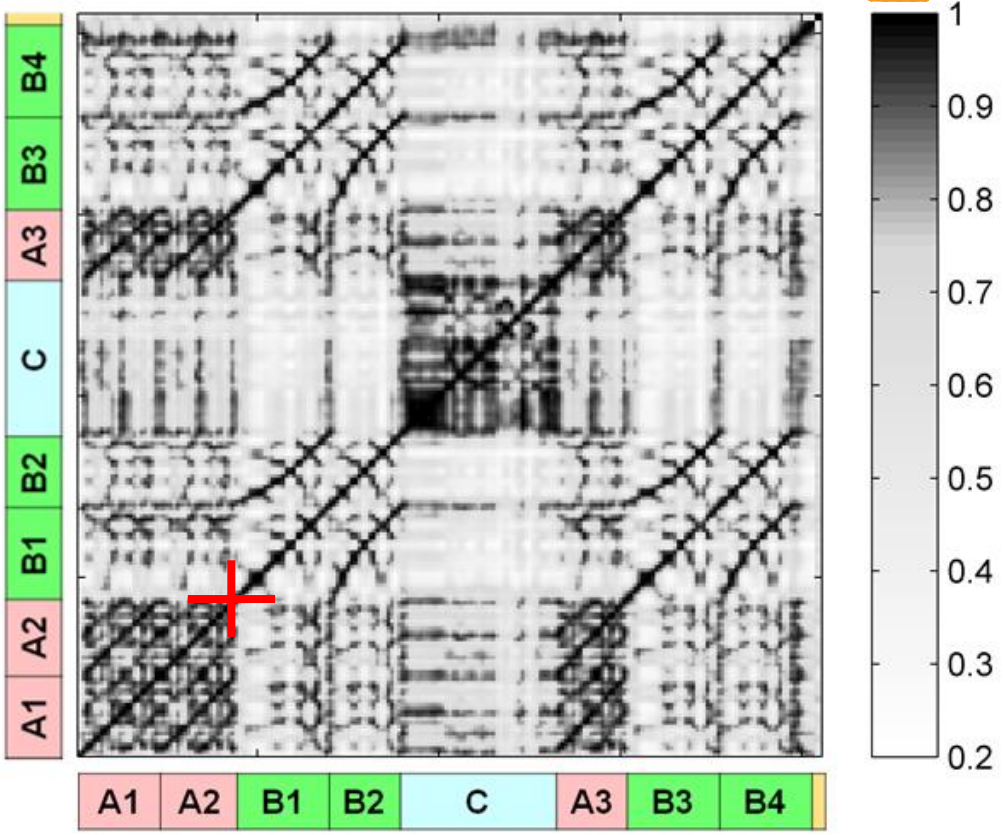
From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Paths: repetition



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Corner: novelty

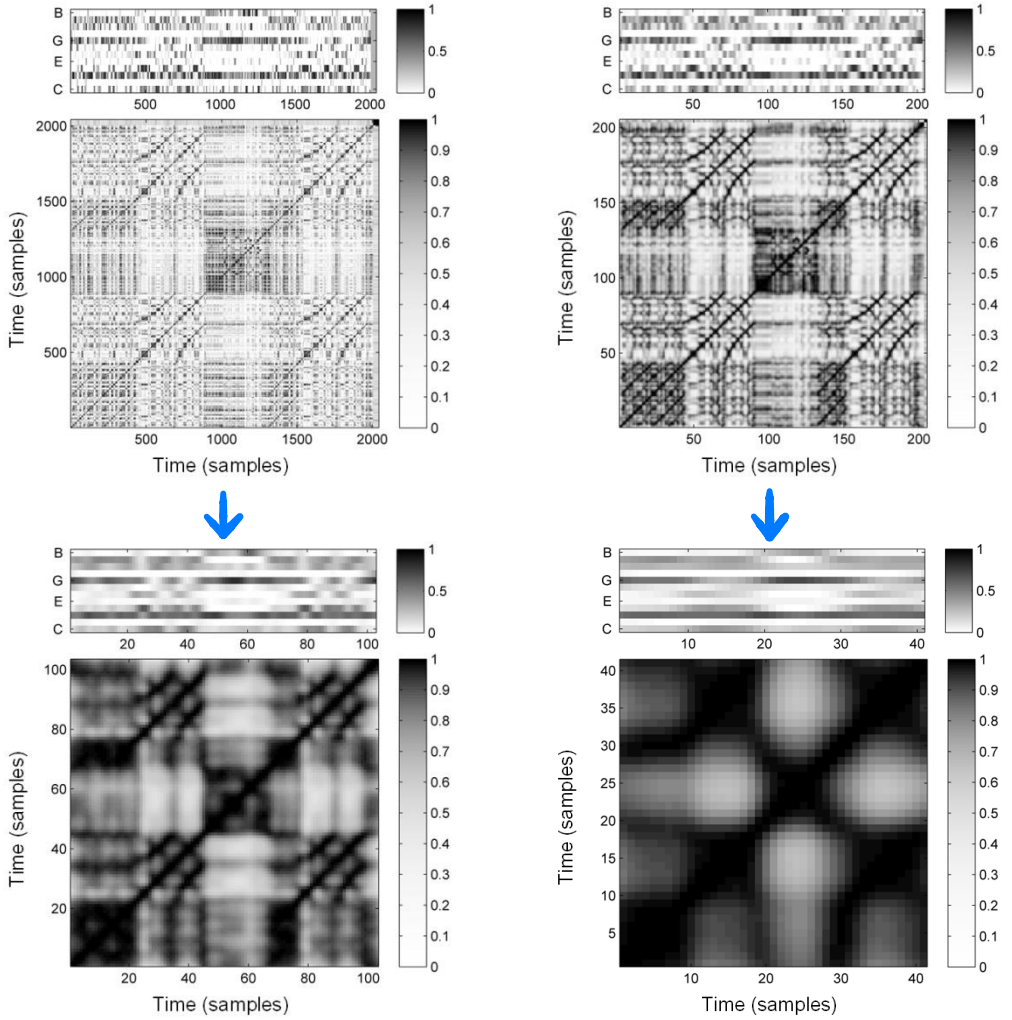


From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Block enhancement

smoothing on (1) feature (2) SSM

- Feature smoothing
- Coarsening



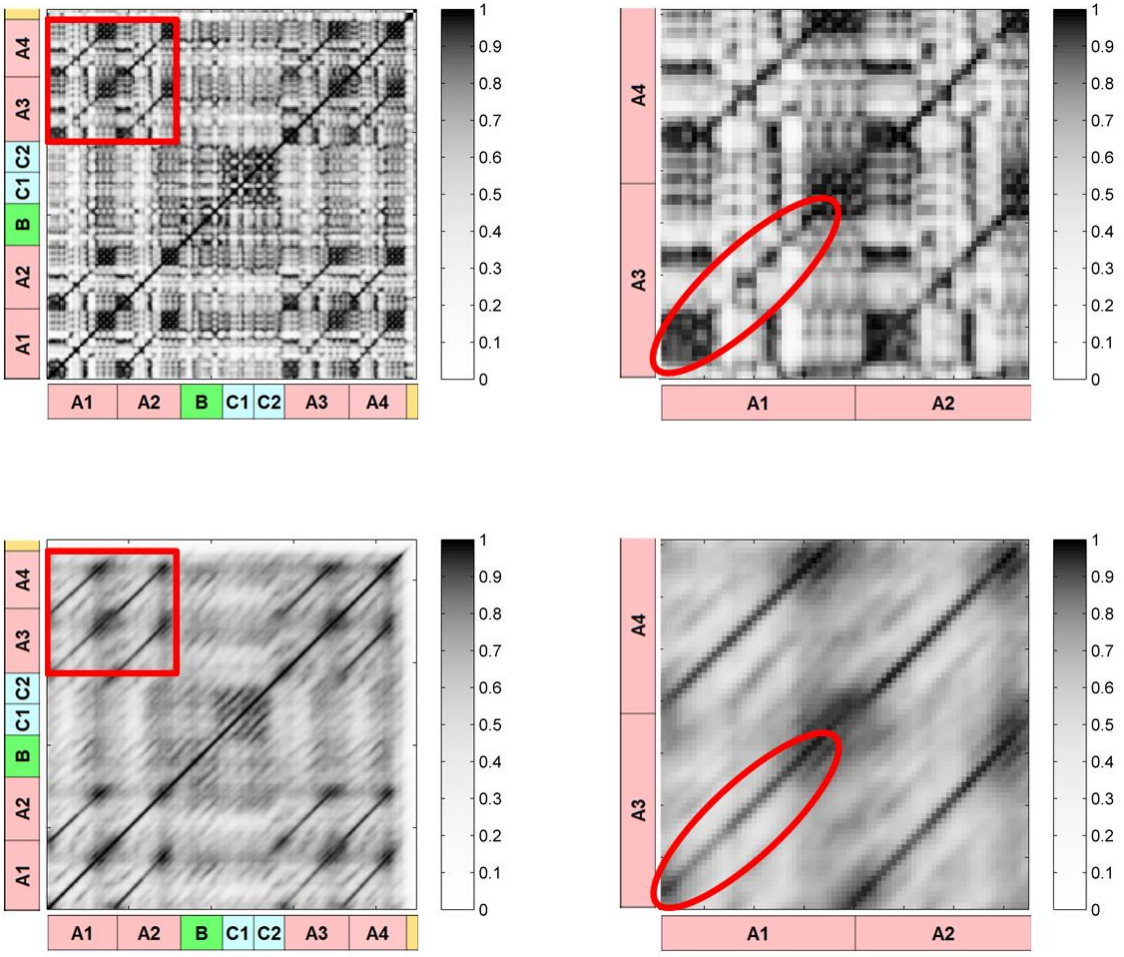
From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Path enhancement

- Filtering along diagonals

$$\mathbf{S}(i,j) = \frac{1}{L}\sum_{l=1}^{L} \mathbf{S}(i+l, j+l)$$

沿 45° 做 smoothing

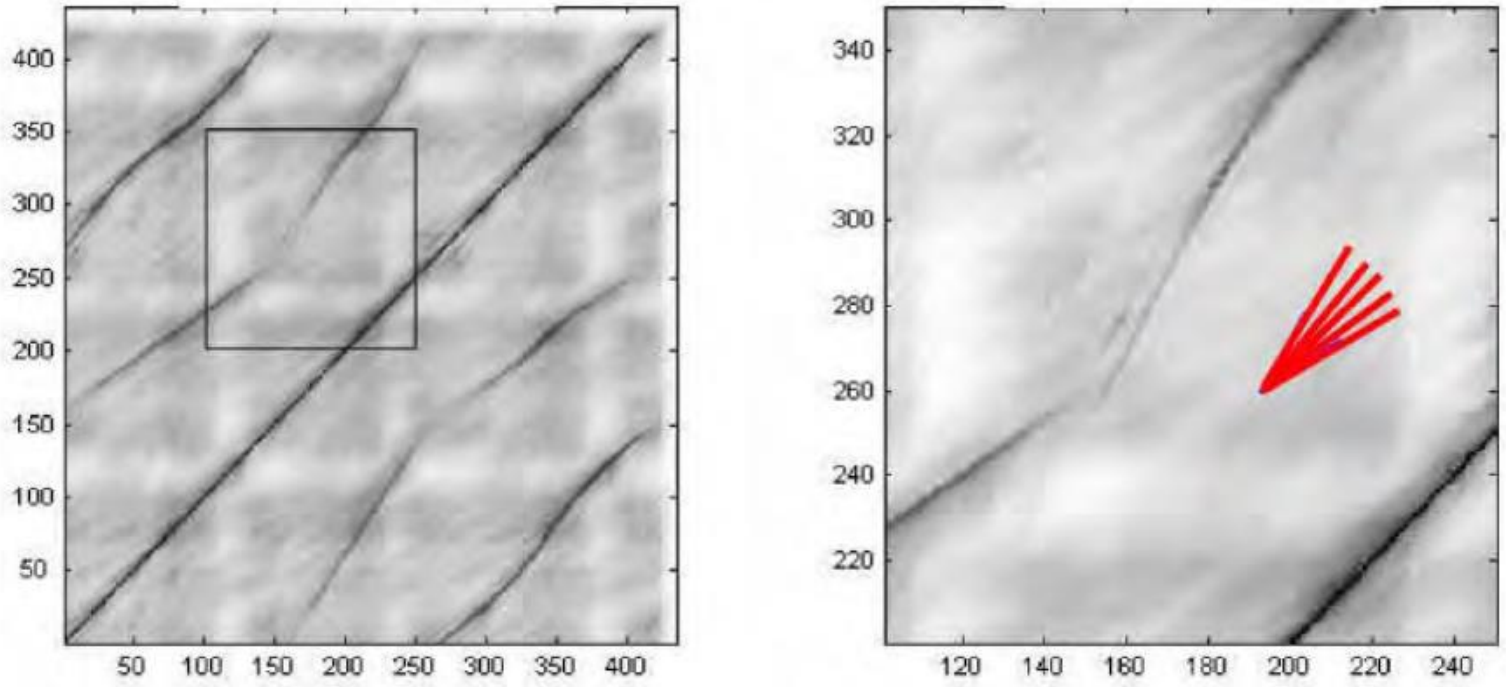


From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Path enhancement: multiple filtering

note 45°未必 always true · A1 to A2 可以不同速

- Filtering along **different directions** and **taking the maximal value,** to account for tempo changes

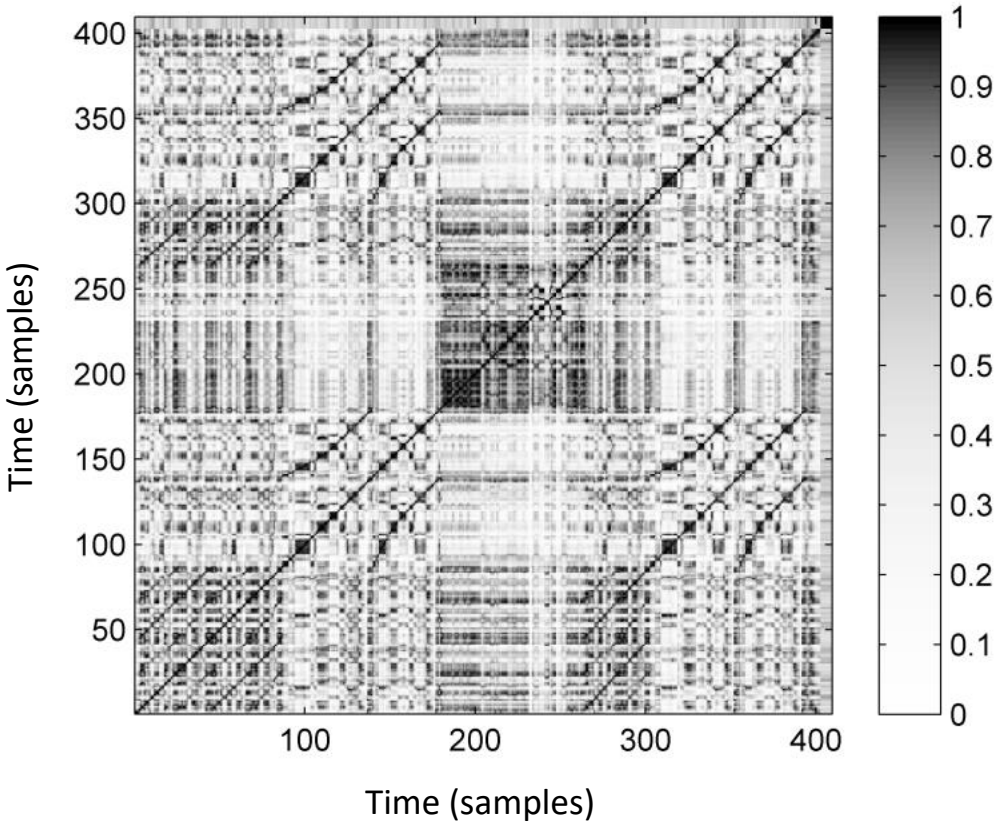


From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Path enhancement

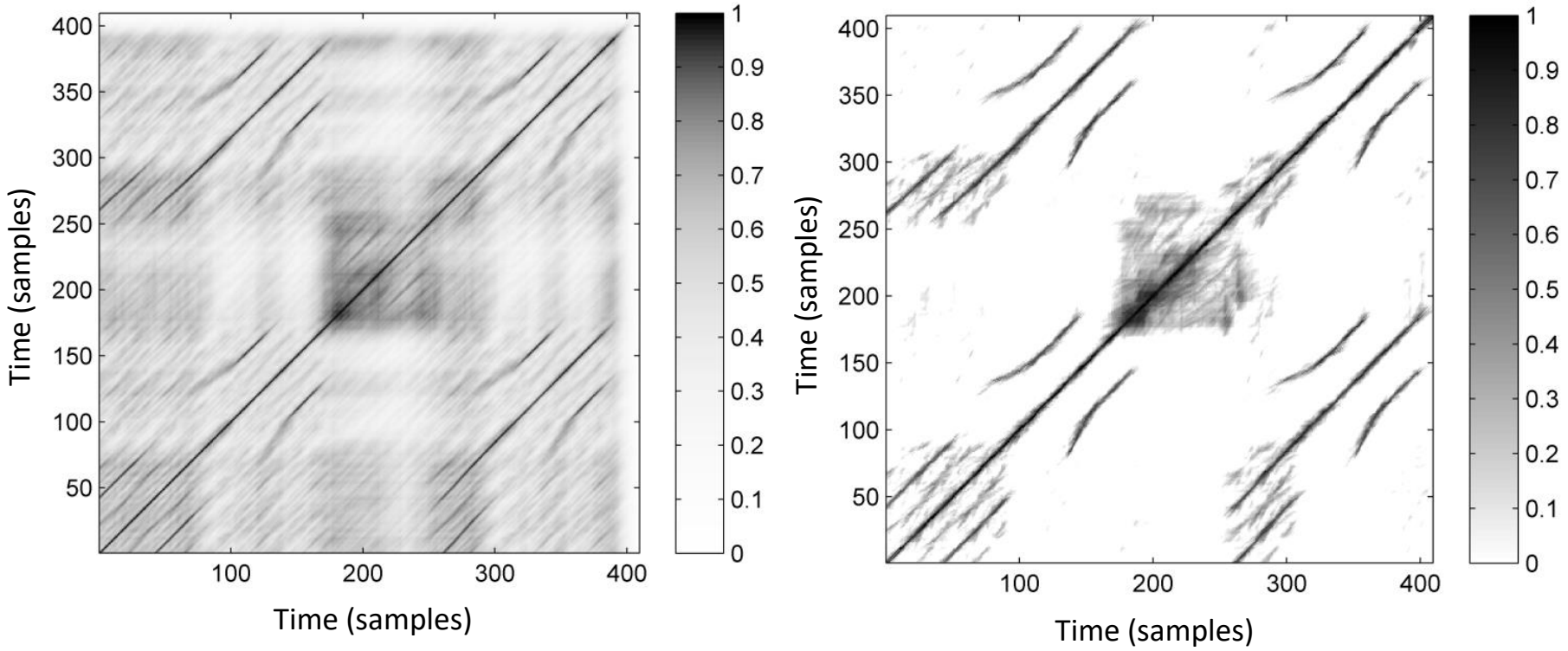
- Procedures:
  - Diagonal smoothing
  - Multiple filtering
  - Thresholding (relative)
  - Scaling and penalty
  - Path extraction
  - Transposition invariance

轉調 Ex: 交響曲 or 副歌升key
要調整 chroma similarity metric



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

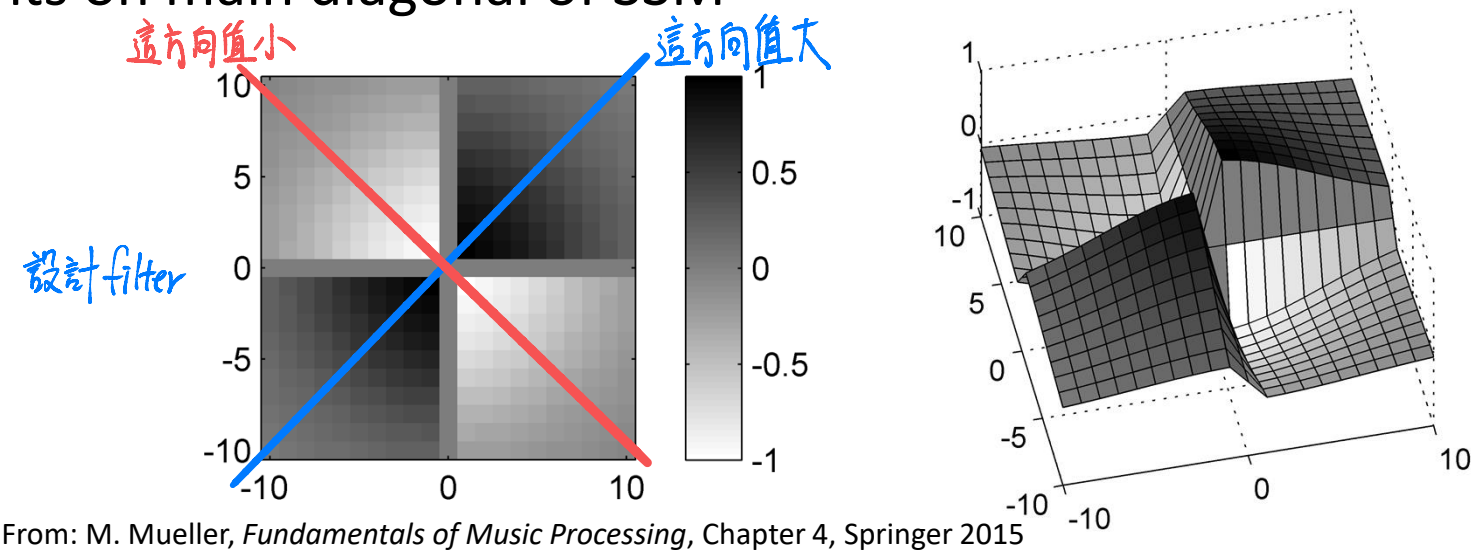
Filtering and thresholding



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Novelty-based Segmentation

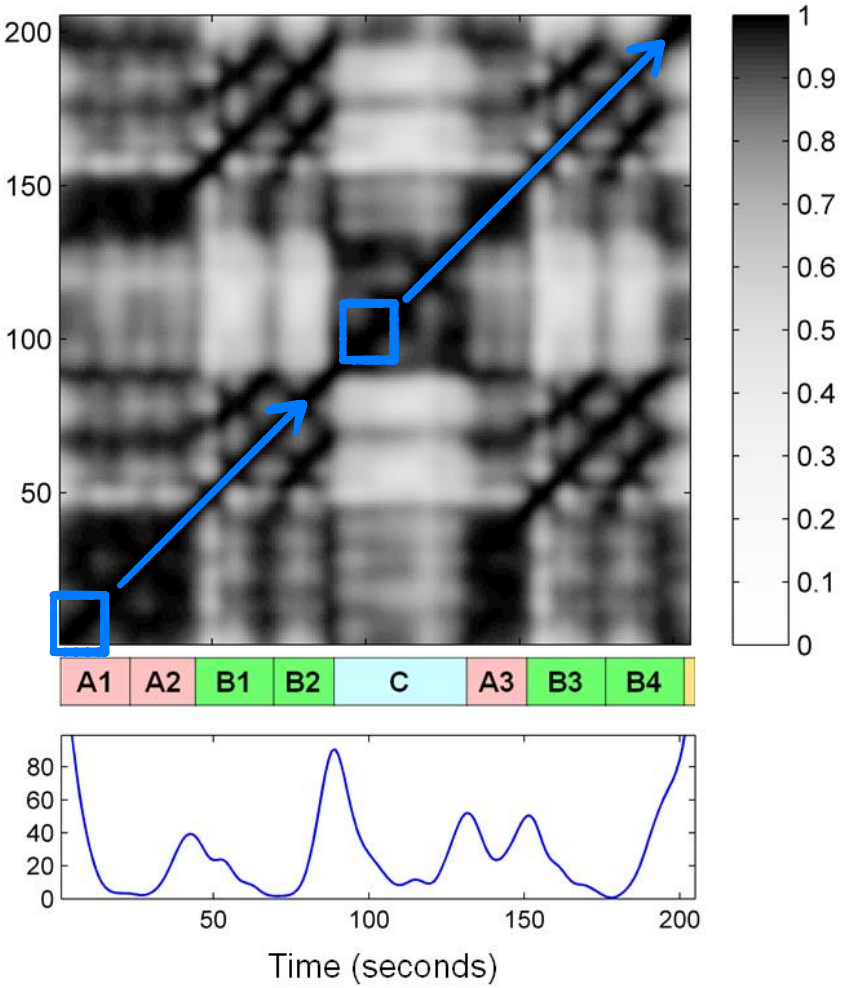
- General goals:
  - Find instances where musical changes occur
  - Find transition between subsequence musical parts
- Main idea:
  - Use checkboard-like kernel function to detect corner points on main diagonal of SSM



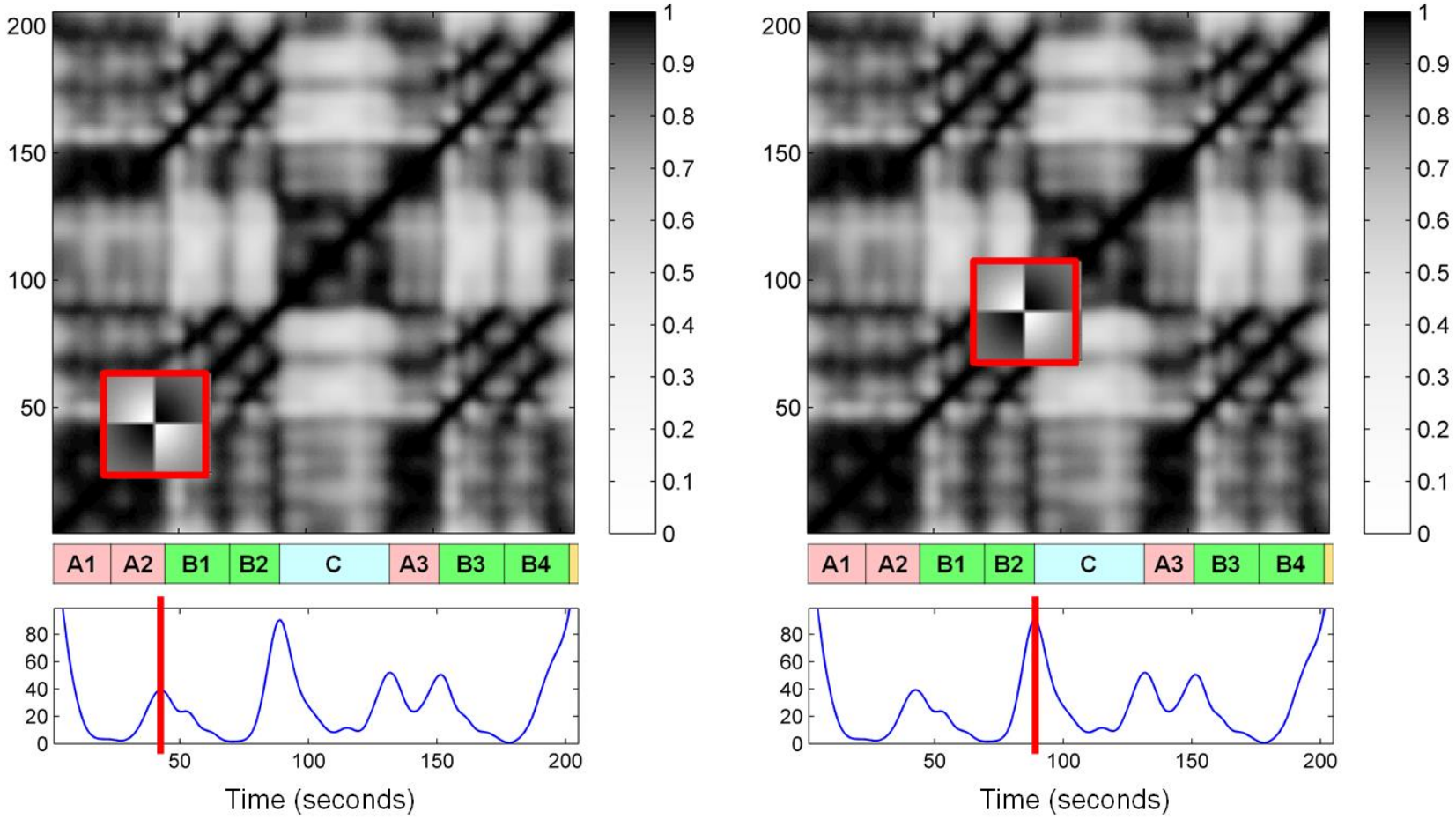
From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

An example

From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015



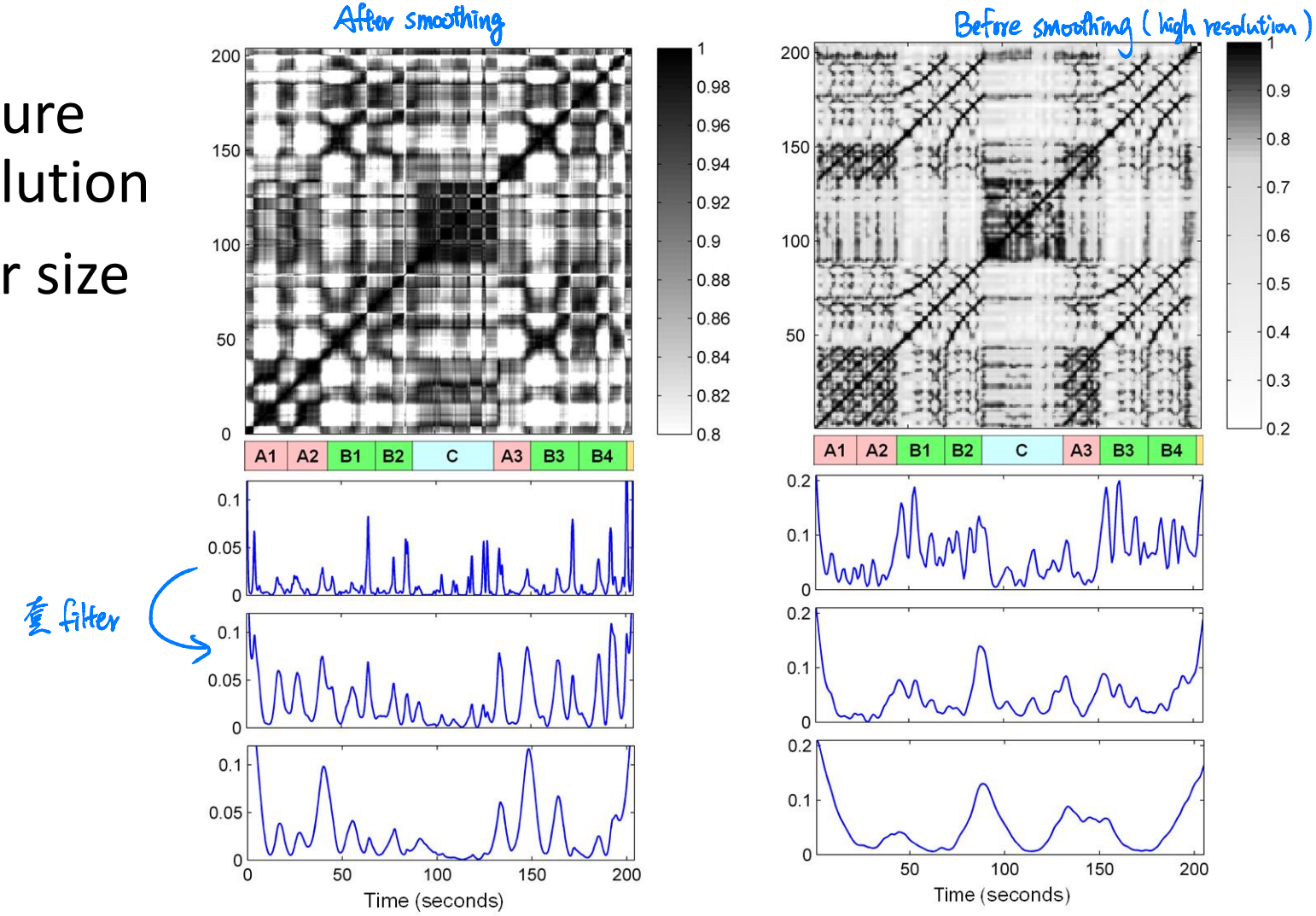
An example



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Comparison

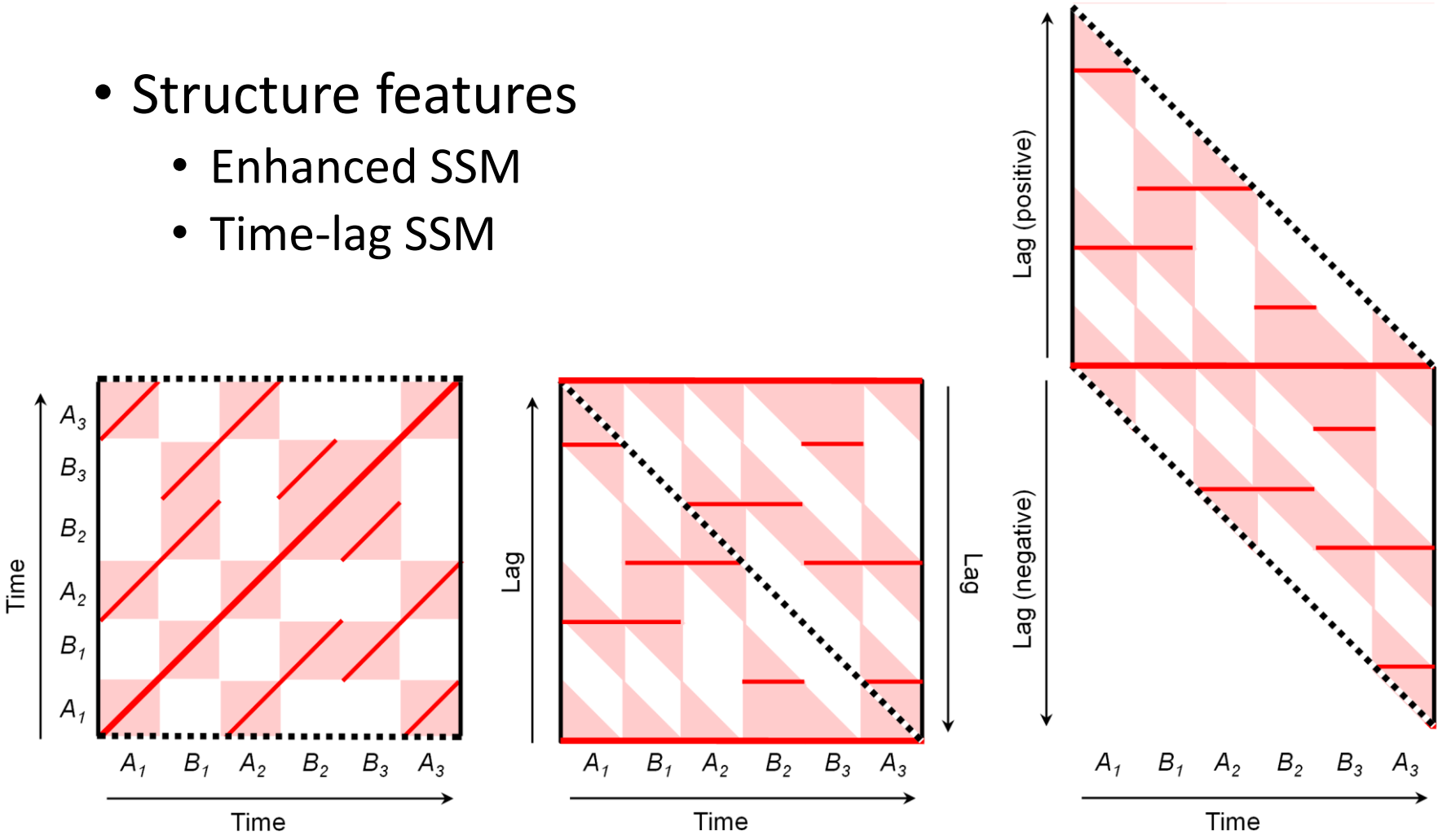
- Feature resolution
- Filter size



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

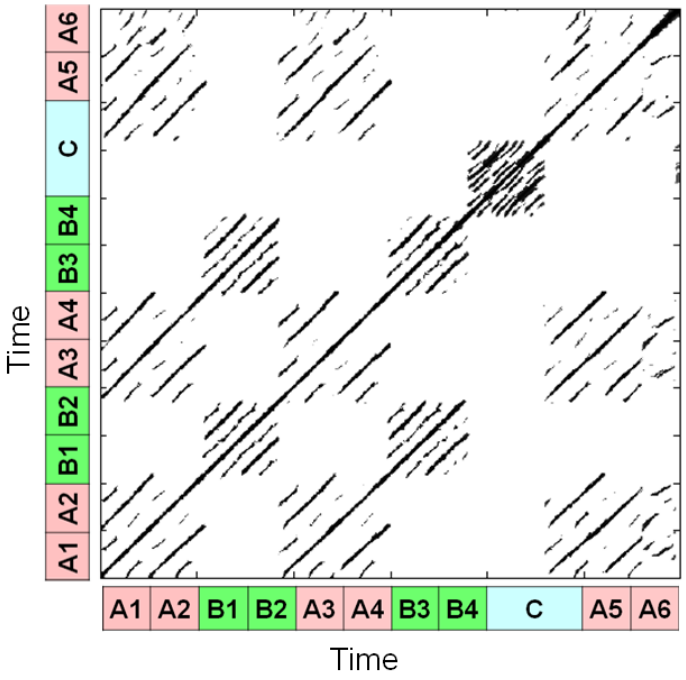
Novelty-based Segmentation: Method 2

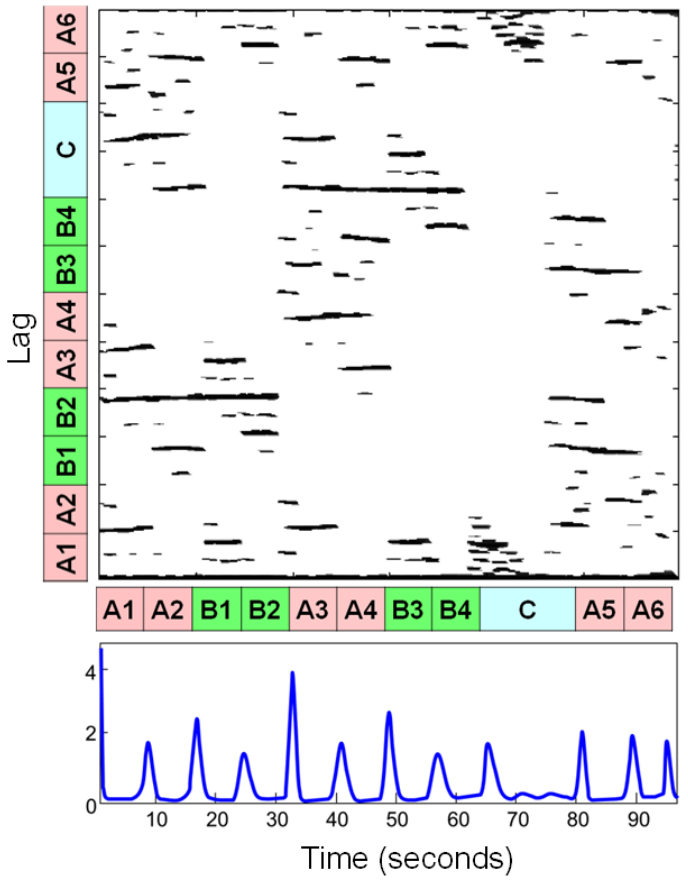
- Structure features
  - Enhanced SSM
  - Time-lag SSM



From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

An example





做 spectral flux

下個時間減上個時間，大於0留下，小於0設成0

From: M. Mueller, *Fundamentals of Music Processing*, Chapter 4, Springer 2015

Further readings

- Foote, Jonathan. "Visualizing music and audio using self-similarity." *Proceedings of the seventh ACM international conference on Multimedia (Part 1)*. ACM, 1999.
- Müller, Meinard, Nanzhu Jiang, and Harald Grohganz. "SM Toolbox: MATLAB implementations for computing and enhancing similarity matrices." *Audio Engineering Society Conference: 53rd International Conference: Semantic Audio*. Audio Engineering Society, 2014.

CNN / UNET 可改

Wrap-up

- The time-frequency structure of music: what have we learned in this semester?