ASyMMuS (Audio-Symbolic Music Similarity Modelling) Workshop – 2015/07/08

Geoffroy Peeters
STMS IRCAM CNRS UPMC

MUSIC STRUCTURE: WHAT IS SIMILAR ?

Geoffroy Peeters Is partly founded by the French government Programme Investissements d'Avenir (PIA) through the Bee Music Project, and by the European Commission through the SKAT-VG (618067) Project.

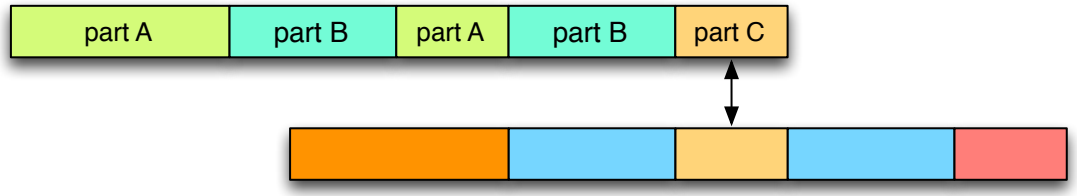
Why interest in music similarity/ structure ?

2001 Research on Audio Identification by fingerprint

-> Closest Matches = Music Similarity

-> But Similarity is not homogeneous within a track

-> Music Structure



Plan

What is Similarity ?

What is Music Structure ?

Sequence approach

- Factor Oracle
- DTW

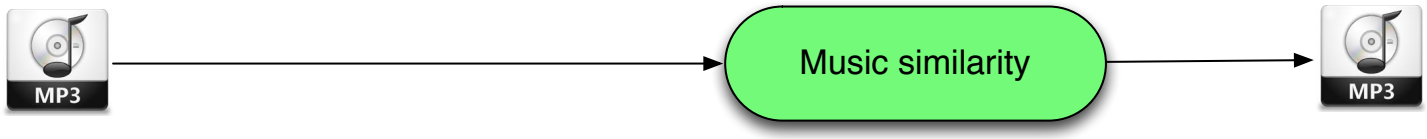
State approach

- Summarizing time-evolution
- Hidden Markov Model

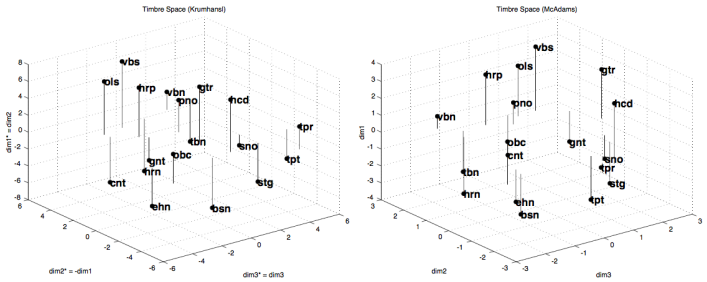
Model-based

- Joint key/chord/structure estimation

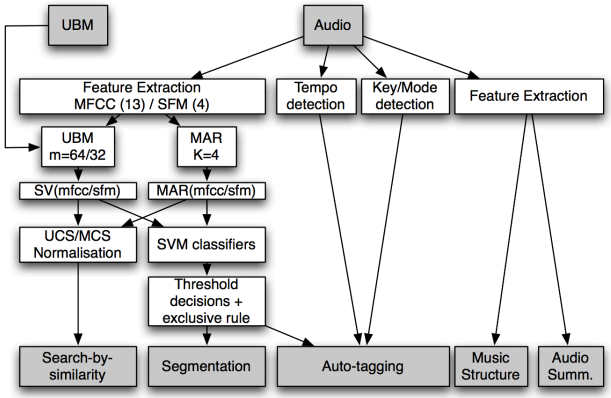
What is similarity ?



- The sound timbre approach
  - Experiment (Lakatos, McAdams)
    - MDS sound timbre space
    - Then find the best acoustic correlates (audio features)
- The music similarity approach
  - Develop the technology then validate
  - Validation ?
    - Perceptual Experiment ?
    - What is the spread of the considered music spac
    - Results are good if same Genre, Artist
    - Perform experiment: Find Outlier
    - Use Pandora recommendation



Lakatos, McAdams Sound Timbre MDS Space



Ircam Music Similarity System, ISMIR 2010

0. What is music structure/ similarity ?

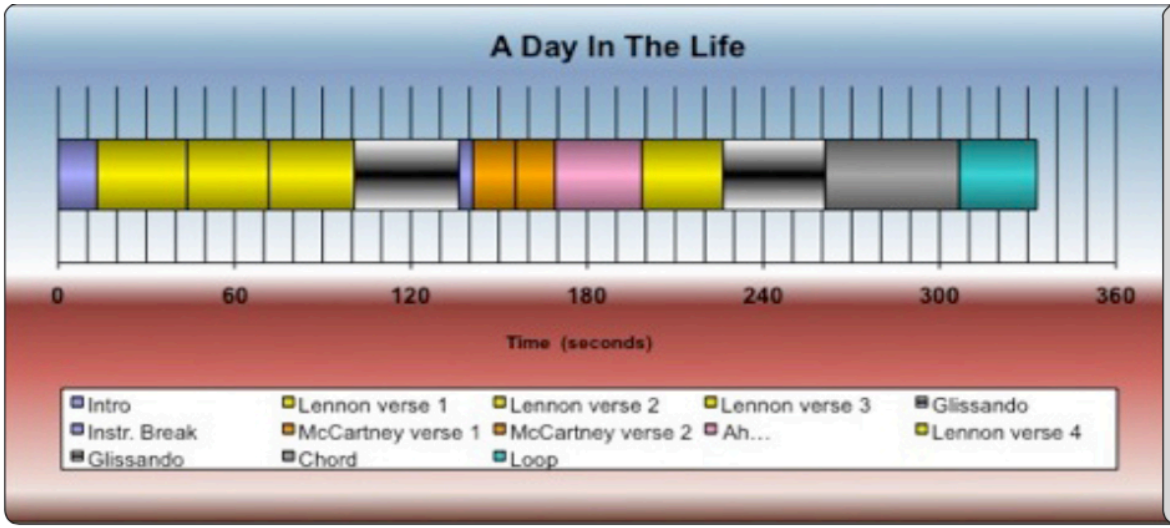


Fig. 9: John Lennon – Paul McCartney, 'A Day In The Life', as recorded by the Beatles (Sgt. Pepper's Lonely Hearts Club Band, Parlophone PCS 7027, 1967, timings from cd version)

Franco Fabbri, « Verse, Chorus (Refrain), Bridge: Analysing Formal Structures of the Beatles' Songs »

0. What is music structure/ similarity ?

Various possible definitions of Structure

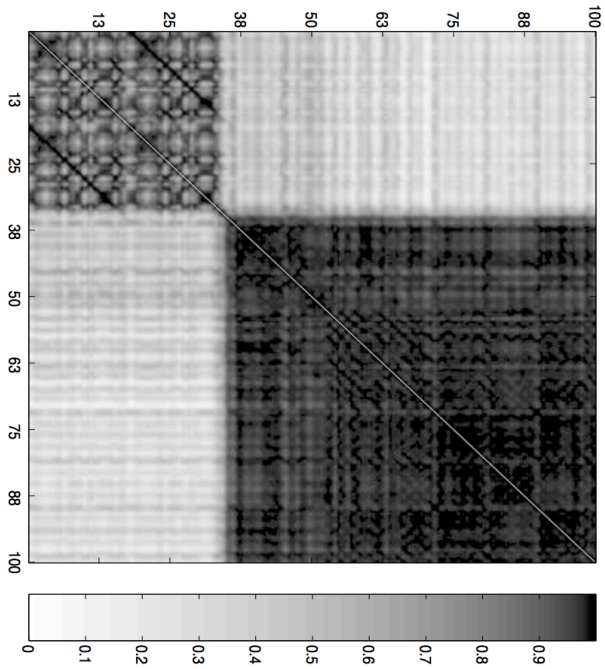
[Peeters, Deruty, 2009]

- 1) based on Musical Role
  - Music role that a part plays in a song
  - Introduction, Verse, Chorus, Bridge, Ending
  - Problems:
    - Intro, Outro= time position, can be the Chorus
    - Several version of Chorus and Verse
    - Definition of Chorus and verse not clear (Rap, R'n'B)
- 2) based on Acoustic Similarity
  - Acoustic similarity between parts
  - Problems:
    - Similar= identical, What about if small variation ?
    - How to quantify ?
    - How to put the threshold ?
    - What criteria for similarity ? Timbre ? Harmony ? Rhythm ? Lyrics ?
- 3) based on Instrument Role
  - Location of lead singer, location of solo guitar
  - Problems:
    - Few insights into the global structure
    - Identify instrument: huge number of labels (guitar= classical ? Folk ? Electric ? WhaWha ?)
  - Solution: Describe the instrument role: Primary Lead, Secondary Lead
- 4) based on Perceptual Tests
    - Average human perception of structure
  - Problems:
    - Very costly !
    - Labels used by people are usually not shared

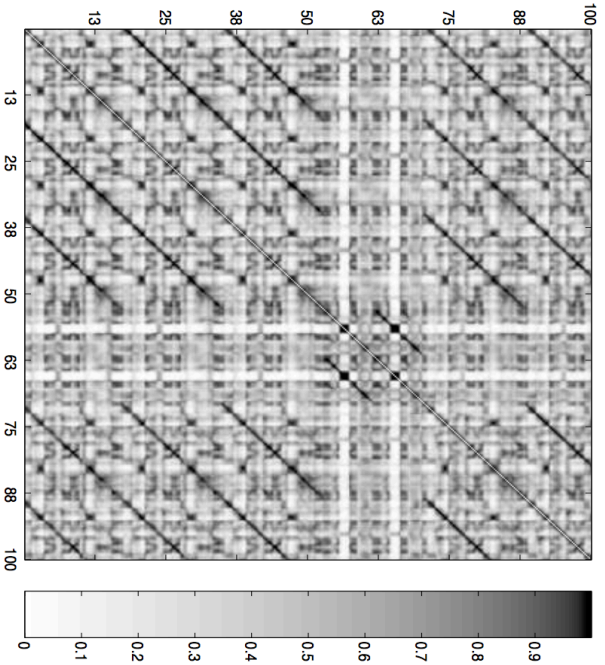
0. What is music structure/ similarity ?
Various view-points on the content

- Moby – « Natral Blues »
  - Different views on the content highlight different structure

Timbre Self-Similarity-Matrix

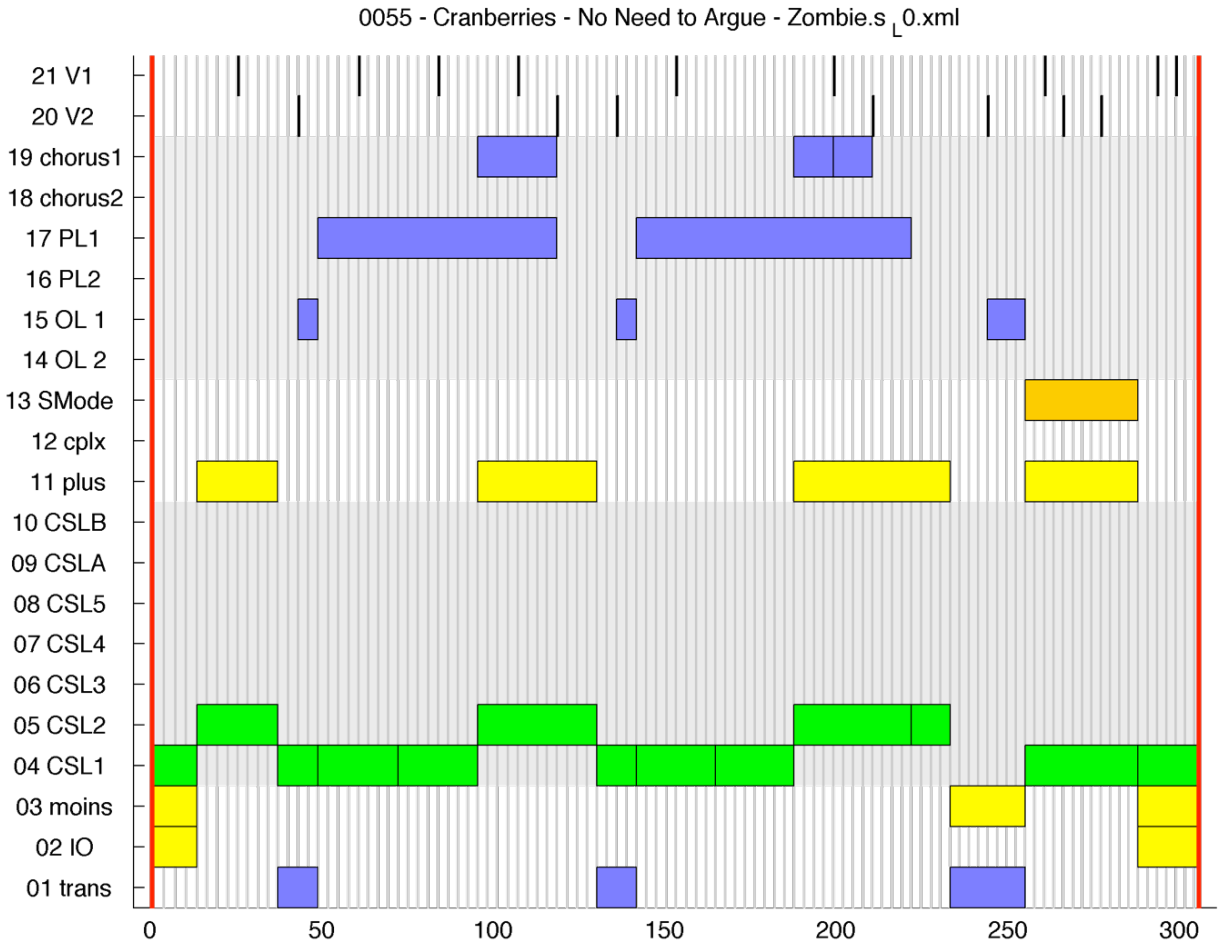


Harmonic Self-Similarity-Matrix



0. What is music structure/ similarity ?

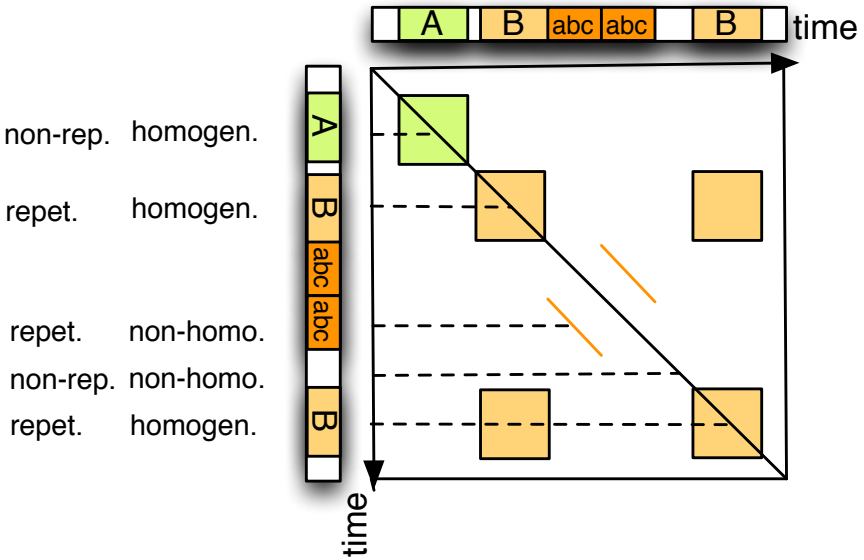
Proposal of a Multi-dimensional annotation system [Peeters, Deruty, 2009]



ircam Centre Pompidou

1. Sequence approach

Sequence approach: what is it ? $[a, b, c] == [a, b, c]$



	homogen.	non-homogen.
repetition	state	sequence
non-repetition	state	state

1. Sequence approach

Factor Oracle

[Laburthe, Peeters 2002]

- Converting audio to symbols

$$[x(1), x(2) \cdots x(t) \cdots x(T)] \in R \rightarrow [a, b \cdots c \cdots a]$$

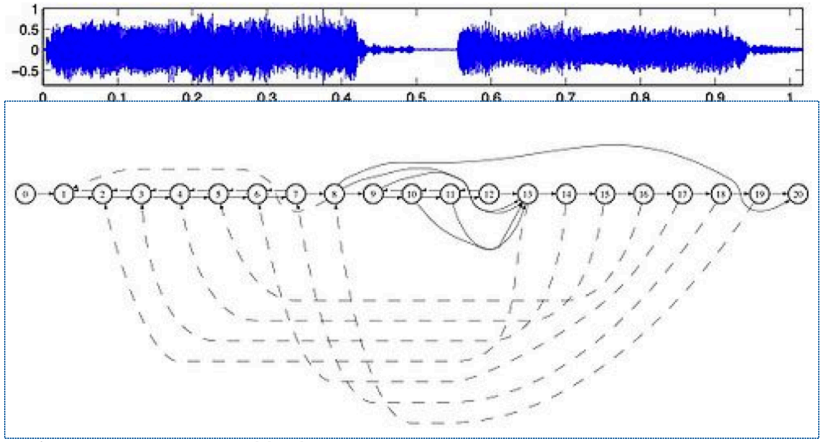
$$a == a$$

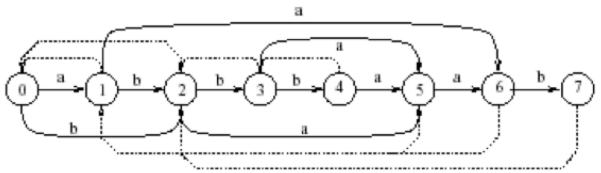
$$a \neq b$$

- Compute Factor Oracle [Alauzen, Crochemore, 1999]

Fonction add_letter(Oracle($p = p_1 p_2 \ldots p_m$), σ)

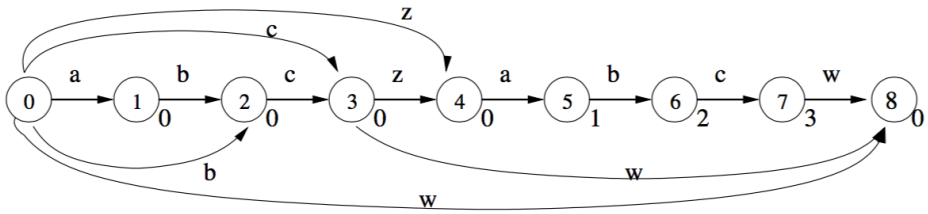
1. Create a new state $m+1$
2. Create a new transition from m to $m+1$ labeled by σ
3. $k \leftarrow S_p(m)$
4. **While** $k > -1$ and there is no transition from k by σ **Do**
5. Create a new transition from k to $m+1$ by σ
6. $k \leftarrow S_p(k)$
7. **End While**
8. **If** $(k = -1)$ **Then** $s \leftarrow 0$
9. **Else** $s \leftarrow$ where leads the transition from k by σ.
10. $S_{p\sigma}(m+1) \leftarrow s$
11. **Return** Oracle($p = p_1 p_2 \ldots p_m \sigma$)





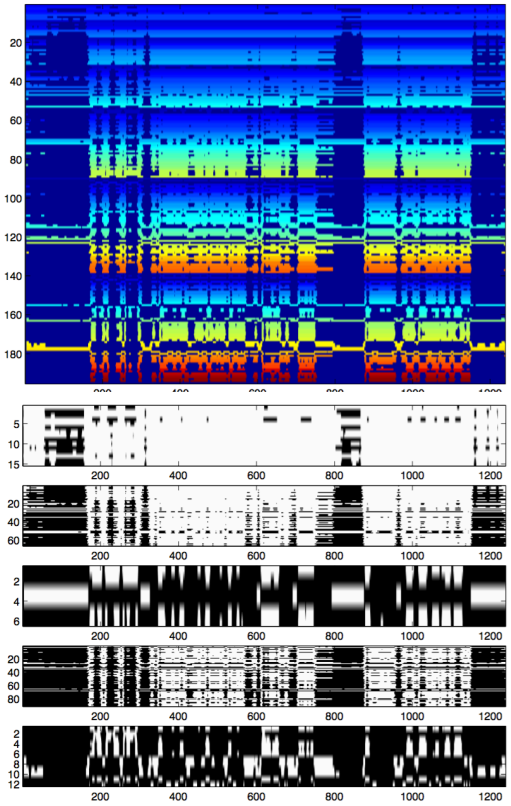
1. Sequence approach
Factor Oracle
[Laburthe, Peeters 2002]

- Compute Length Repeated Suffix (LRS) using FO [Lefevre, Lecroq, 2000]



- Structural matrix based on LRS
(from longest to shortest LRS)

- Grouping matrix based on pattern similarity

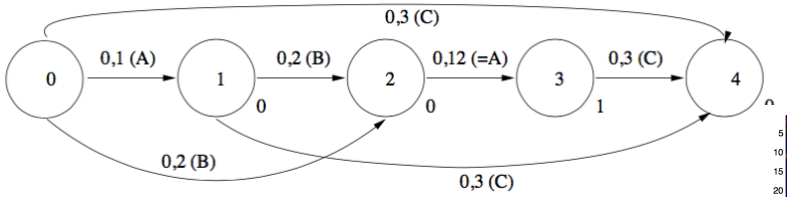


1. Sequence approach
Factor Oracle
[Laburthe, Peeters 2002]

- Converting distance to equivalence $d(x(t), x(t')) \leq \tau \rightarrow t == t'$

  $d(x(t), x(t')) > \tau \rightarrow t \neq t'$

- Adaptive Factor Oracle



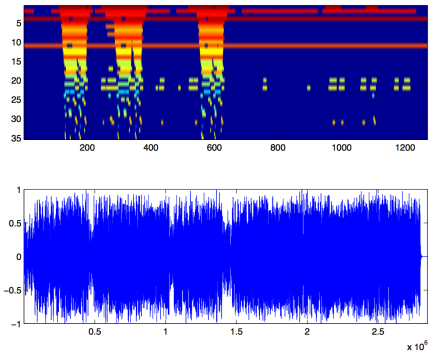


FIG. 4.17 – *Matrice de décomposition sur 'Day Tripper' des Beatles: Oracle sur valeurs réelles seuil de ressemblance=0.05*

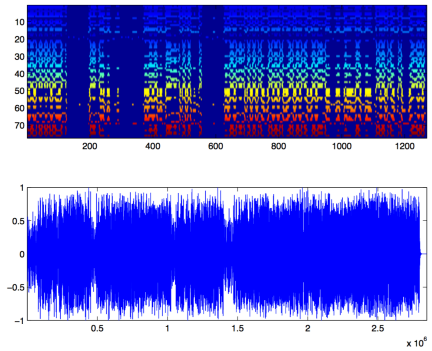


FIG. 4.18 – *Matrice de décomposition sur 'Day Tripper' des Beatles: Oracle sur valeurs réelles seuil de ressemblance=0.0025*

Dynamic Time Warping [Mueller, 2012]

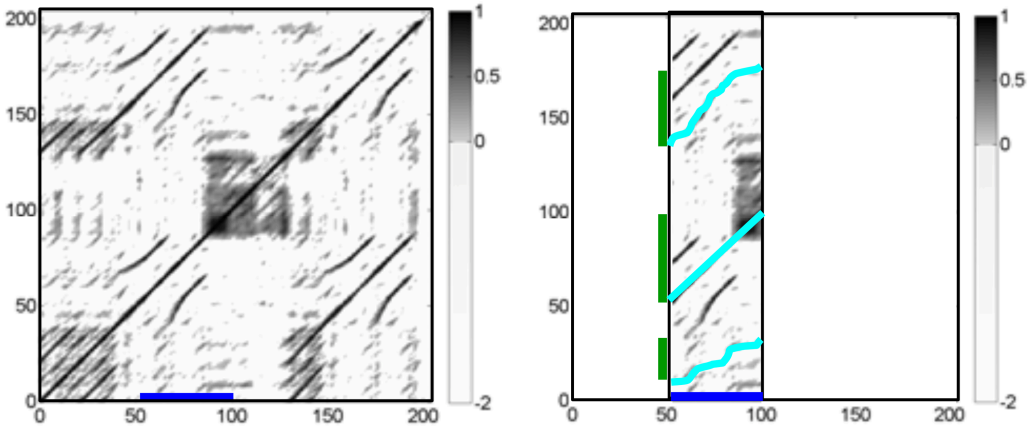
- Compute distance between continuous values

$$[x(1), x(2) \cdots x(t) \cdots x(T)] \in R$$

$$d(x(t), x(t'))$$

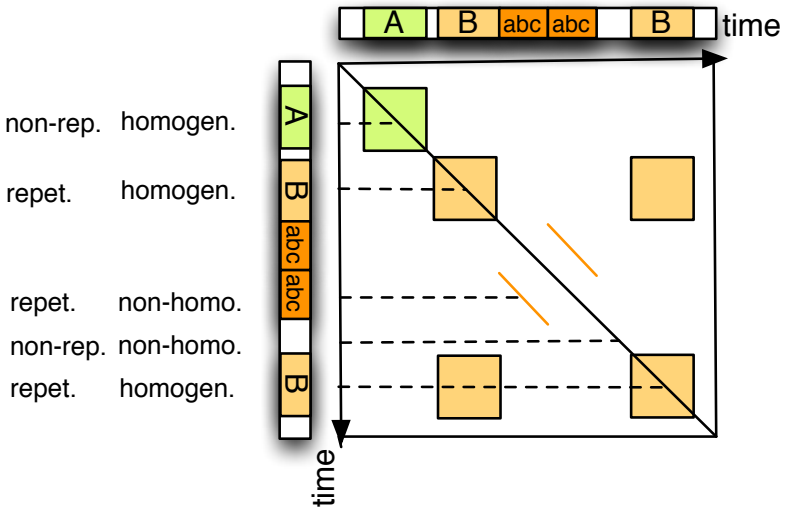
$$insertion, deletion, minimum - cost - path$$

- Find the best alignments between sub-sequences



2. State approach

State approach: what is it ? $[A = A = A] == [A = A = A]$

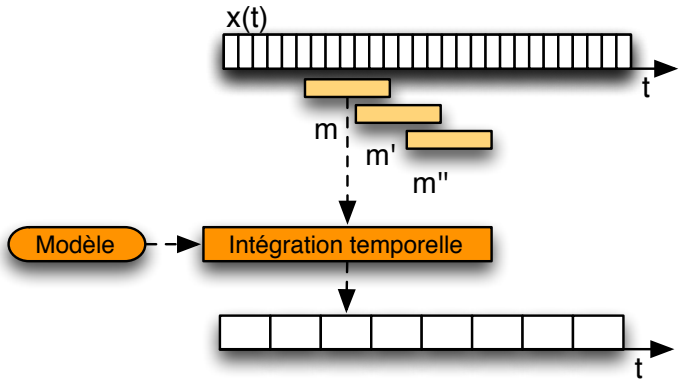


	homogen.	non-homogen.
repetition	state	sequence
non-repetition	state	state

Modeling Time Evolution

Instead of comparing distances between sequences $[x(1), x(2) \cdots x(t) \cdots x(T)]$

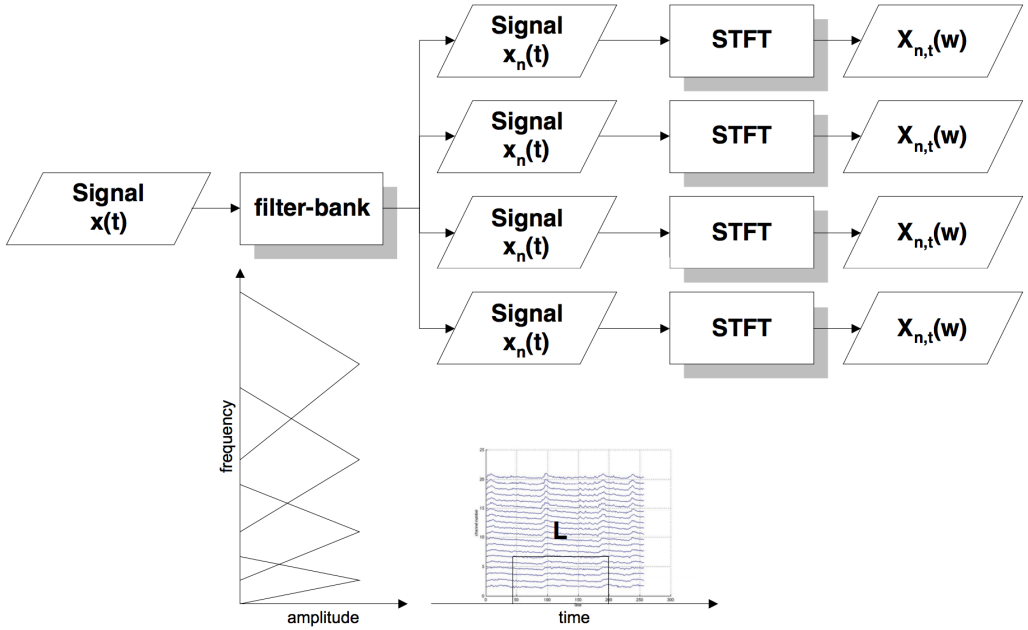
- We model the evolution of $x(t)$ over time:
$$f([x(t-\delta) \cdots x(t) \cdots x(t+\delta)]) \rightarrow x'(t)$$
- Then compute the distance between the models (discriminant, invariant)
$$d(x'(t), x'(t'))$$
- Two similar sequences will have two similar models -> state representation



2. State approach
Modeling Time Evolution

Some time evolution models:

- Modulation Spectrum [Peeters, Rodet, 2002]

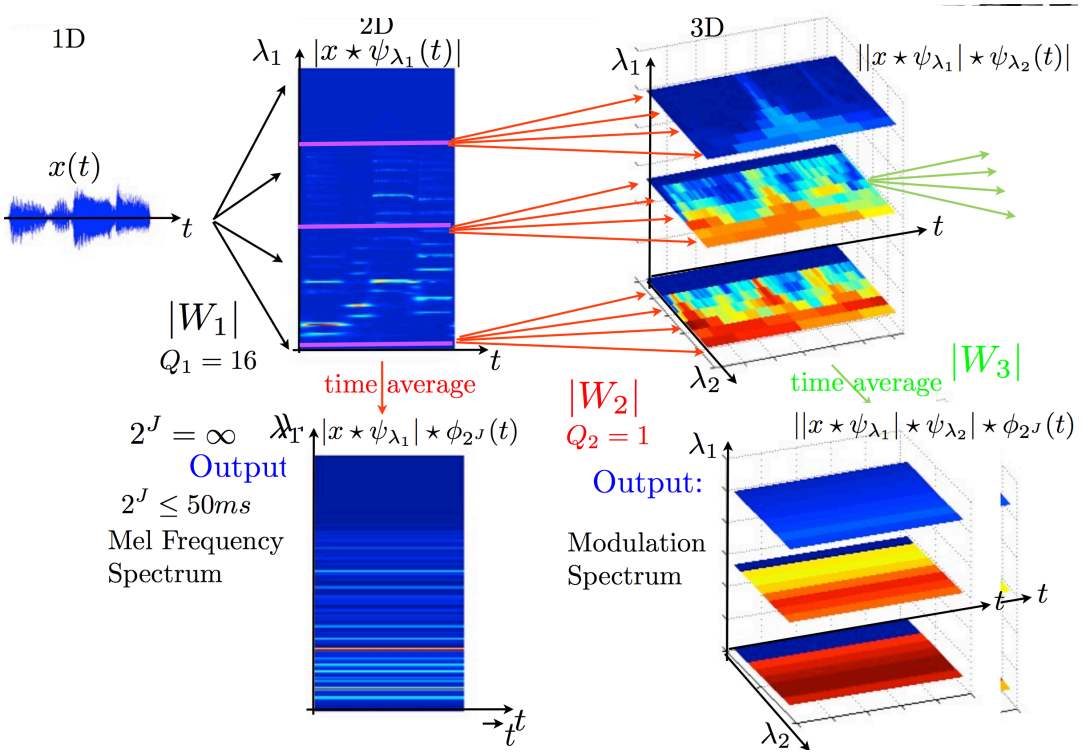


2. State approach

Modeling Time Evolution

Some time evolution models:

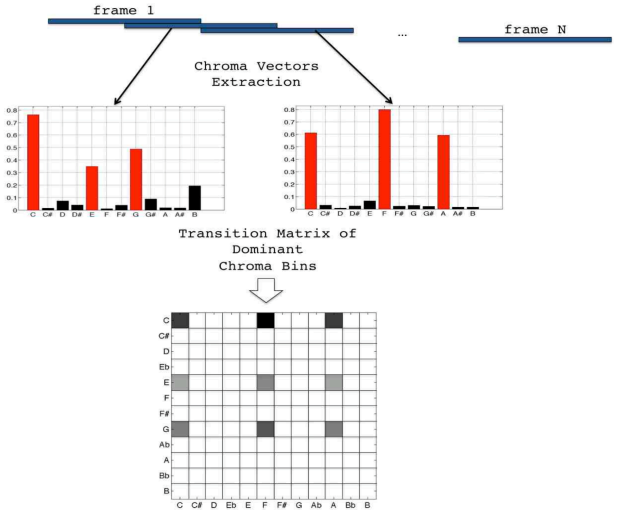
- Scattering Transform [Anden, Mallat, 2014]

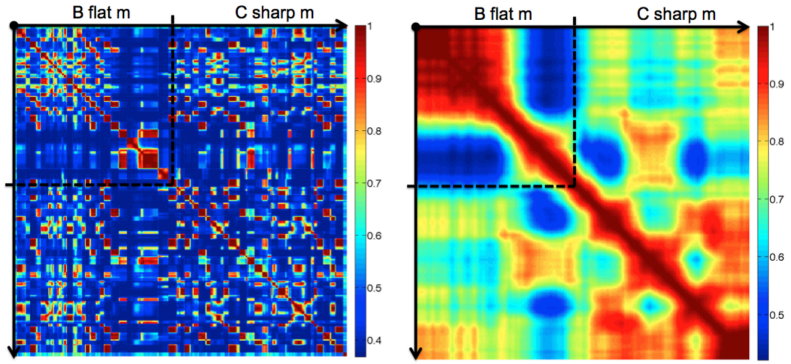


2. State approach
Modeling Time Evolution

Some time evolution models:

- Multi-Prob Histogram [Kaiser, Sikora, 2011]





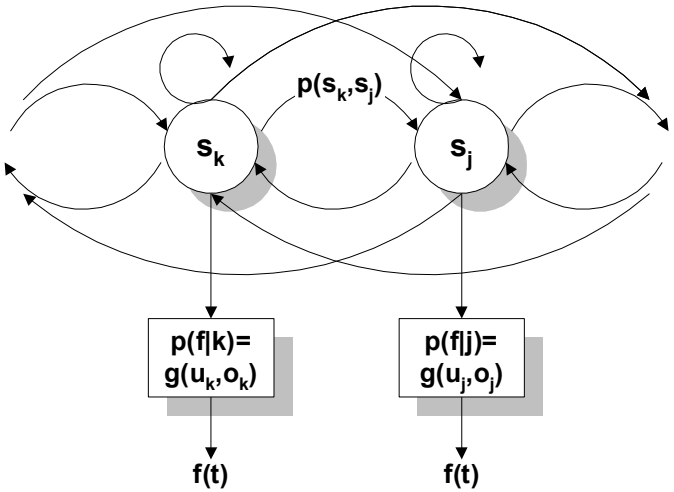
(a) Similarity Matrix with Chroma Vectors (b) MPH Similarity Matrix, L = 50 frames

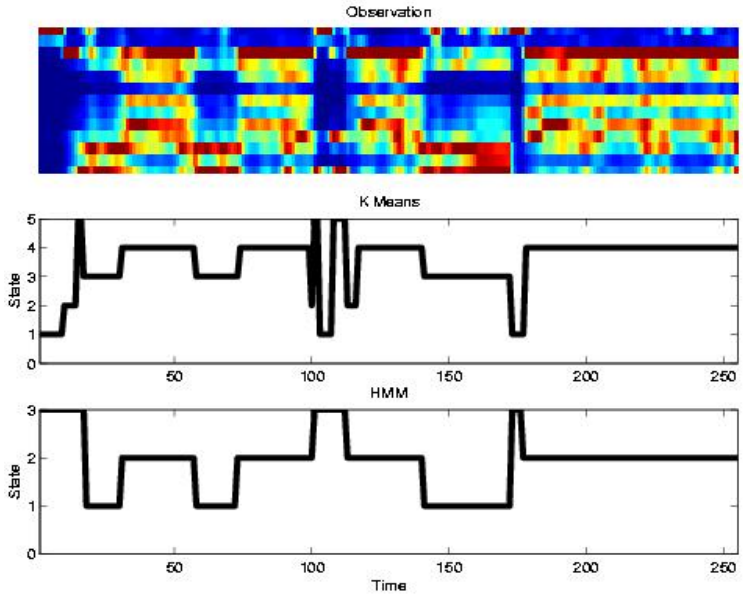
2. State approach

Hidden Markov Model for Structure Estimation

[Logan, 2000] [Aucouturier, 2001] [Peeters, 2002]

- The hidden states of the HMM represent the various parts (repeated or not) of the music track
- The observations are $x'(t)$



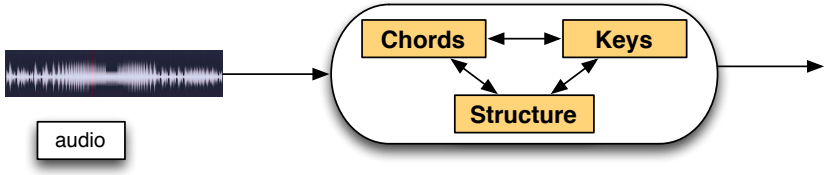


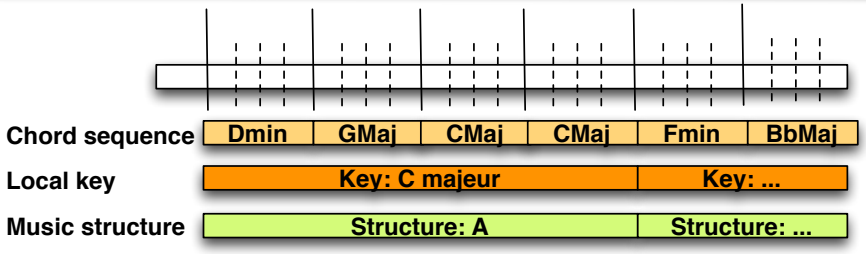
ircam Centre Pompidou

3. Model Approach

Joint Estimation: Structure/Chord/Key

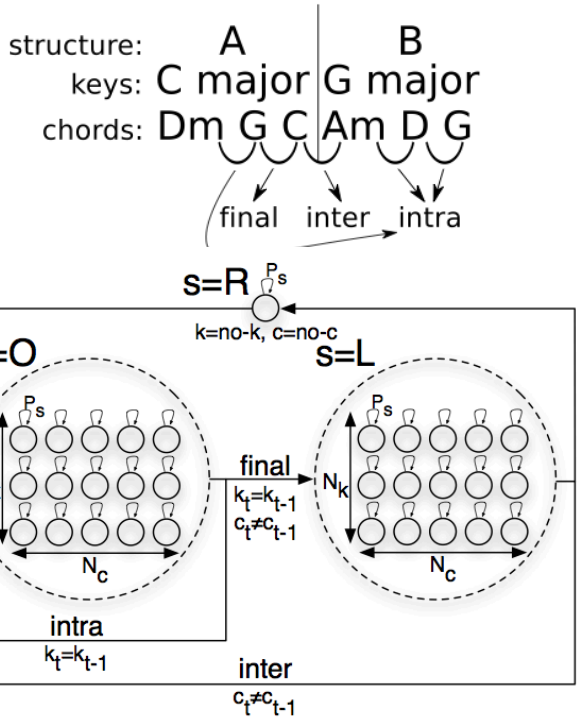
[Pauwels, Peeters, 2013]





Model

- Hidden State q=(s,k,c)
  - Key k in Nk
  - Chord c in Nc
  - Structure s=
    - L: last state of structural segment
    - O: not the last state
    - R: no chord, no-key state



$$\hat{S},\hat{K},\hat{C} = \arg\max \prod_{t=1}^{T} P\left(\mathbf{y_t}|s_t,k_t,c_t\right) P\left(\mathbf{z_t}|s_t,k_t,c_t\right) P\left(s_t,k_t,c_t|s_{t-1},k_{t-1},c_{t-1}\right)$$

Questions ?