

Scruter la forme de l'informel Automatiser l'analyse du free jazz

Marie Tahon, Zaher Belghith



Analyse musicale et modélisation informatique des activités cognitives
14 novembre 2019

Sommaire

- Contexte musical: le Free Jazz des années 50-60
- Problématique
- Proposition pour une segmentation automatique
- Analyse acoustique
- Perspectives

Le Free Jazz (50-60)



Problématique

- Le Free Jazz: « le refus des codes conventionnels » [Ekkehard Jost]
- L'analyste cherche des catégories musicales
- Il existe des « partitions » graphiques montrant implicitement l'existence d'une structure musicale
- Peut-on (re)trouver une structure musicale dans les enregistrements de Globe Unity ?

Structure musicale

cas d'étude: Mozart

- Motifs harmoniques ou rythmiques
- Sections AABA
 - A: abac
 - B: deabfg
- Exemple de l'allegretto K545 (Mozart)

The image shows a musical score for the Rondo Allegretto grazioso (K. 545) by Mozart. The score is divided into sections labeled with letters a through g, illustrating the AABA structure. The sections are:

- a**: First section, marked *p*, highlighted in green.
- b**: Second section, marked *p*, highlighted in blue.
- c**: Third section, marked *mf*, highlighted in blue.
- d**: Fourth section, marked *mp*, highlighted in brown.
- e**: Fifth section, marked *p*, highlighted in purple.
- f**: Sixth section, marked *p*, highlighted in pink.
- g**: Seventh section, marked *poc*, highlighted in yellow.

The score also includes dynamic markings such as *p*, *mf*, *mp*, and *poc*, and performance instructions like *res.* and *crese.* The tempo is marked *Allegretto grazioso* with a quarter note equal to 104 beats per minute.

Structure musicale

vers une extraction automatique ?

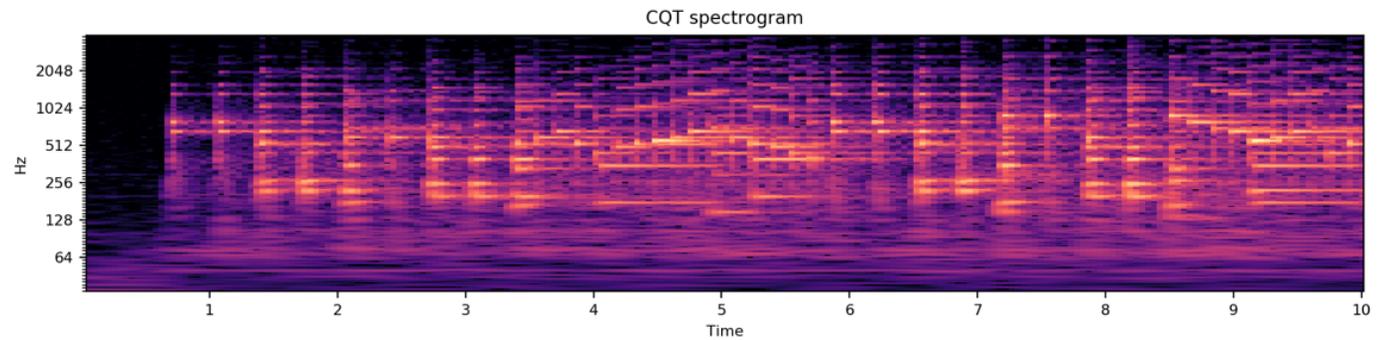
- Méthodes de partitionnement et regroupement spectral [McFee&Ellis2014]
 - Segmenter le signal musical et regrouper ces segments suivant leur similarité
 - Méthode non-supervisée (absence de corpus conséquent)
- Quelle représentation du signal musical ?
- Quelle résolution temporelle ?
- Combien de structures cherche-t-on ?

Résolution temporelle (1/2)

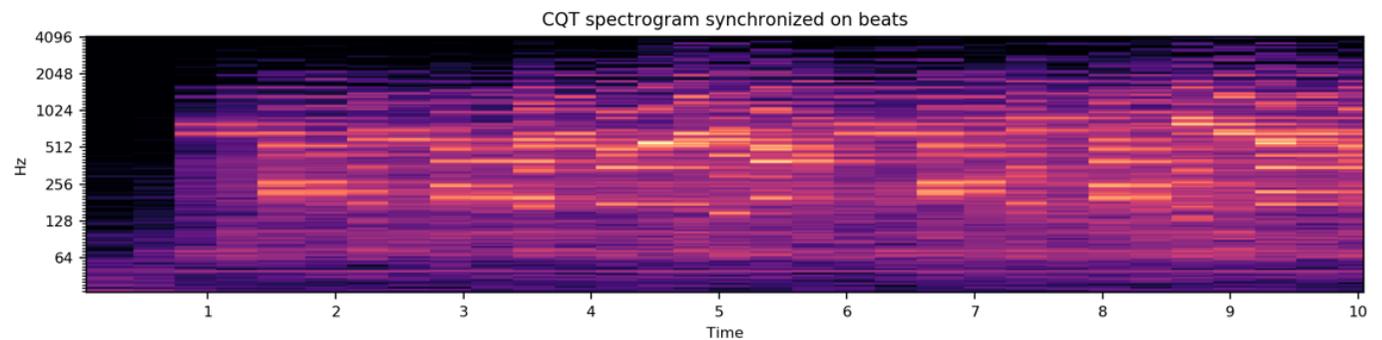
- Résolution pour l'analyse spectrale $\approx 30\text{ms}$ (FFT)
- Résolution pour l'analyse musicale ?
 - Les temps ? détection automatique de beat
 - Les début de notes ? détection automatique d'onset
 - La phrase musicale ? $\approx 1\text{s}$ (comment la segmenter ?)
 - Une taille fixe ? 30ms
 - Définie manuellement

Résolution temporelle (2/2)

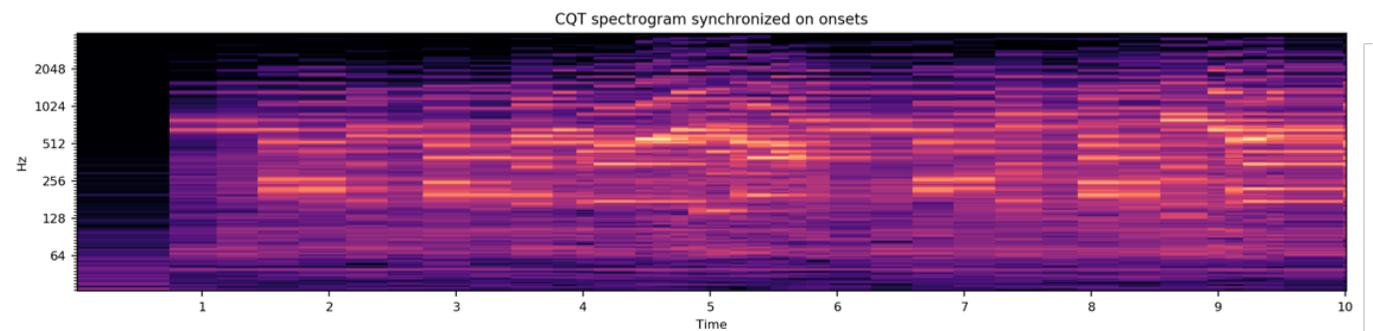
Taille fixe



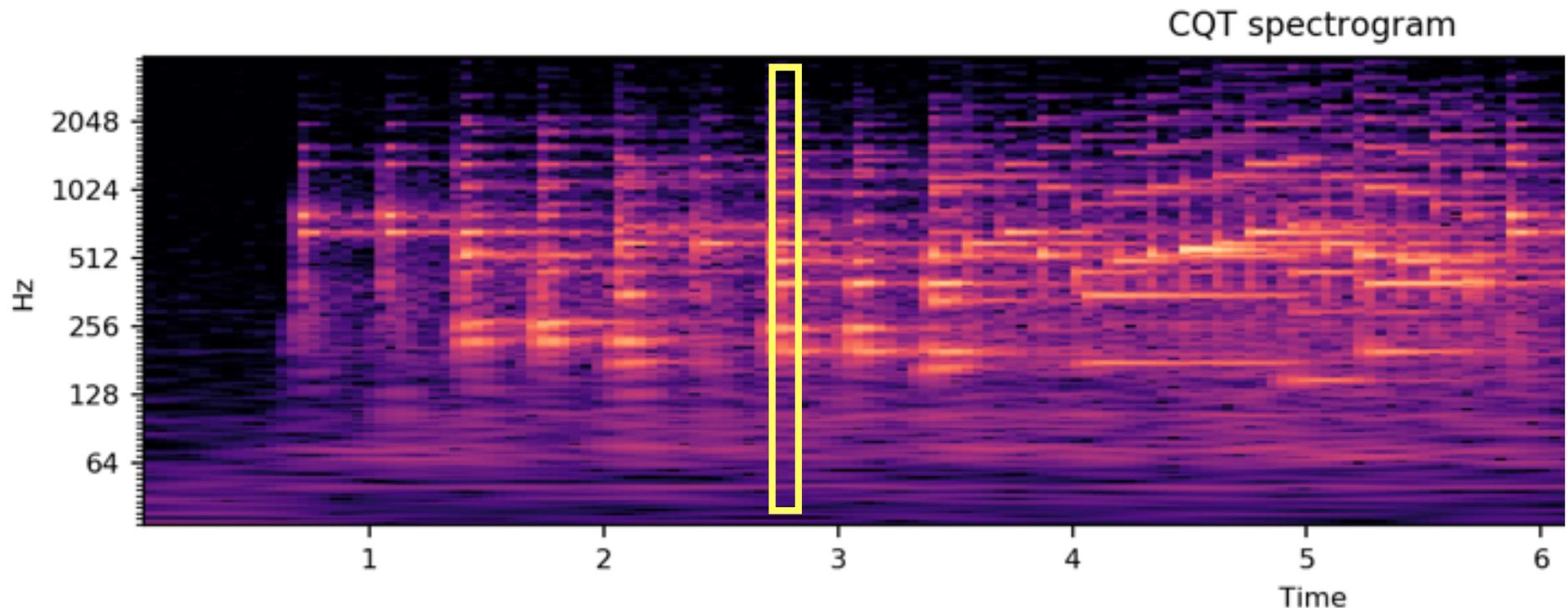
Beat detection



Onset detection

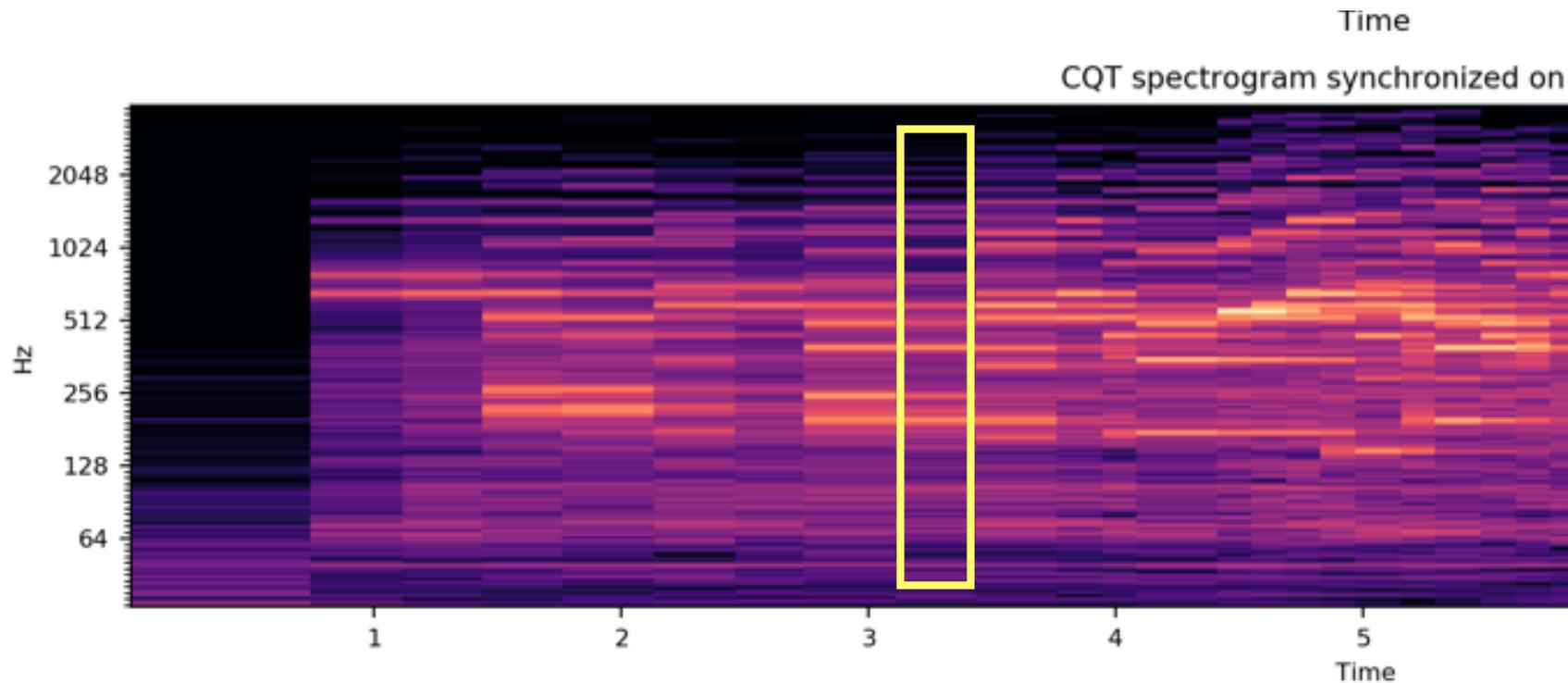


Représentation spectrale (1/4)



- Matrice de similarité à court terme: R_{court}
 - Entre 2 trames de 30ms
 - Représentation acoustique: coefficients cepstraux

Représentation spectrale (2/4)

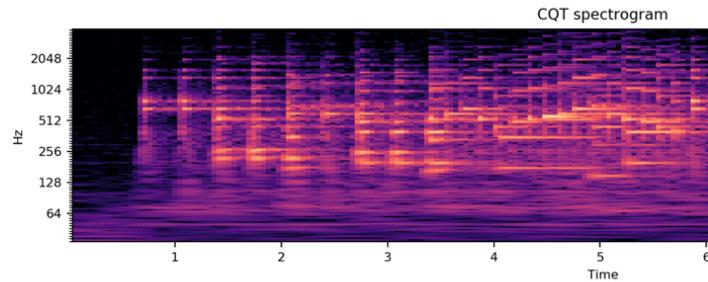


- Matrice de similarité à long terme:

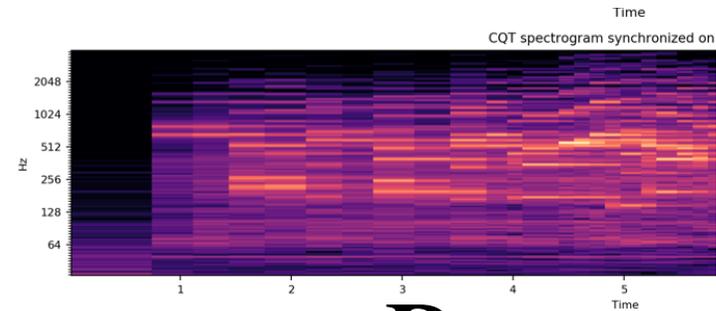
R_{long}

- Résolution à définir: onset, beat detection
- Représentation acoustique: cepstraux (MFCC), spectraux, chroma

Représentation spectrale (3/4)



R_{court}

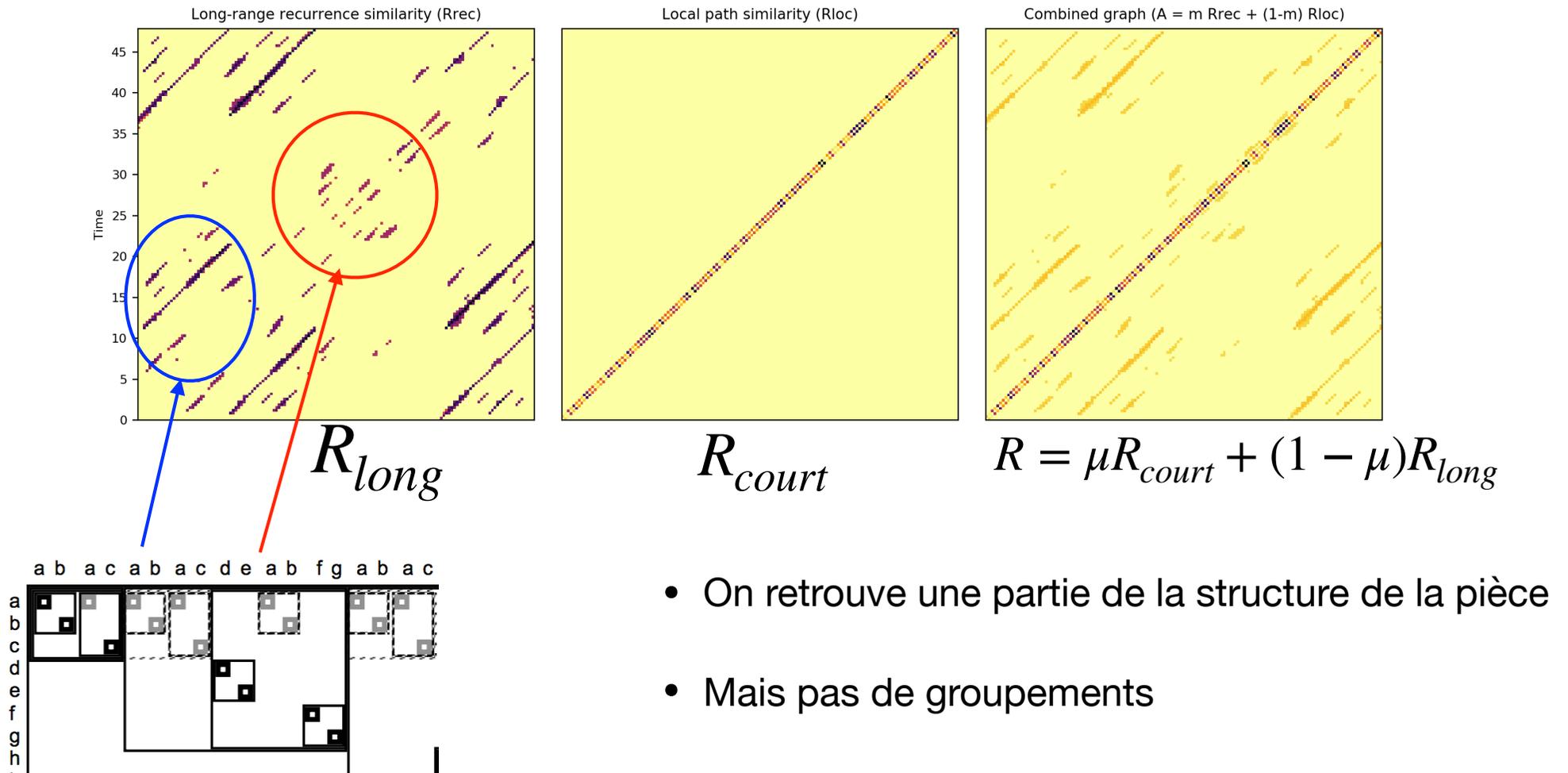


R_{long}

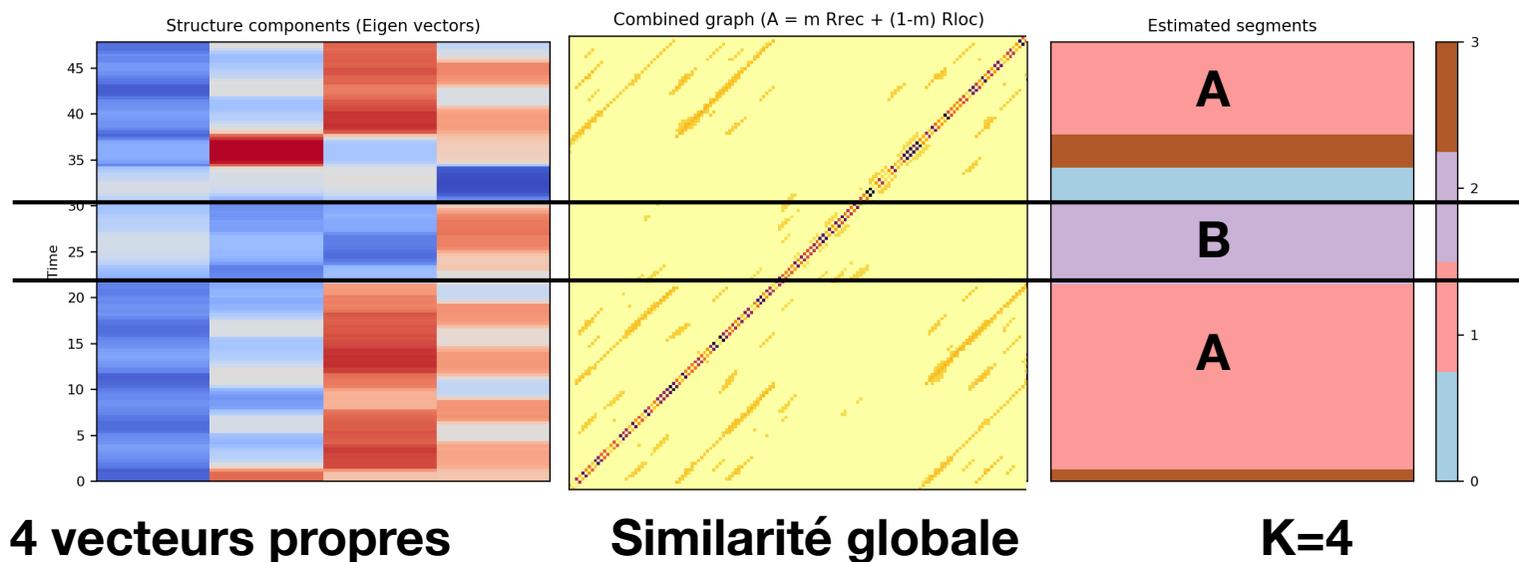
$$R = \mu R_{court} + (1 - \mu) R_{long}$$

Représentation spectrale (4/4)

cas d'étude: Mozart

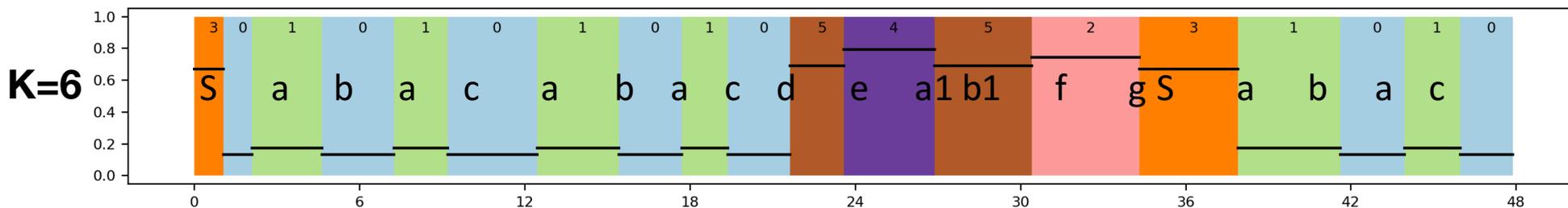
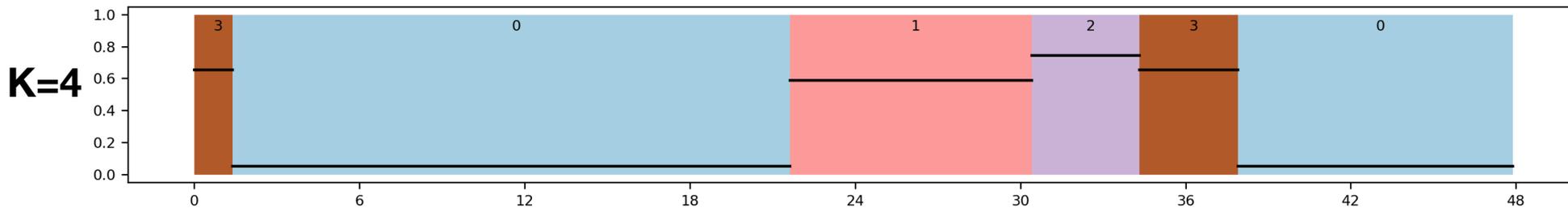
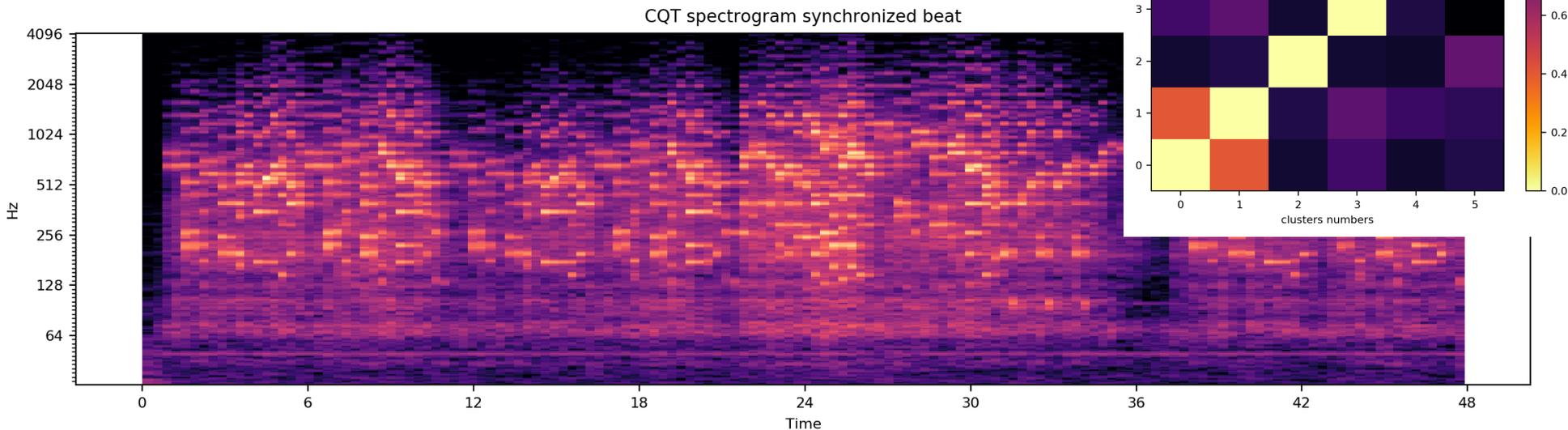


Clustering (1/2)



- Objectif: regrouper les zones de la matrice de similarités en K groupes homogènes
- Définir le nombre de groupes K
 - Détermination automatique (plusieurs méthodes issues du clustering)
 - Itération sur plusieurs valeurs de K
- Obtenir une représentation compacte: K premiers vecteurs propres.
- Ajout d'une distance de similarité entre deux clusters

Clustering (2/2)

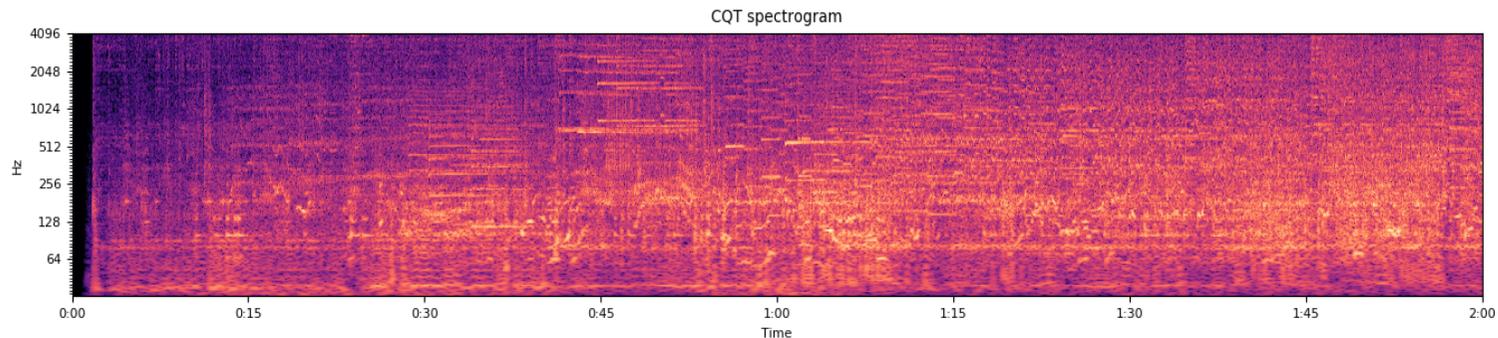


Segmentation automatique

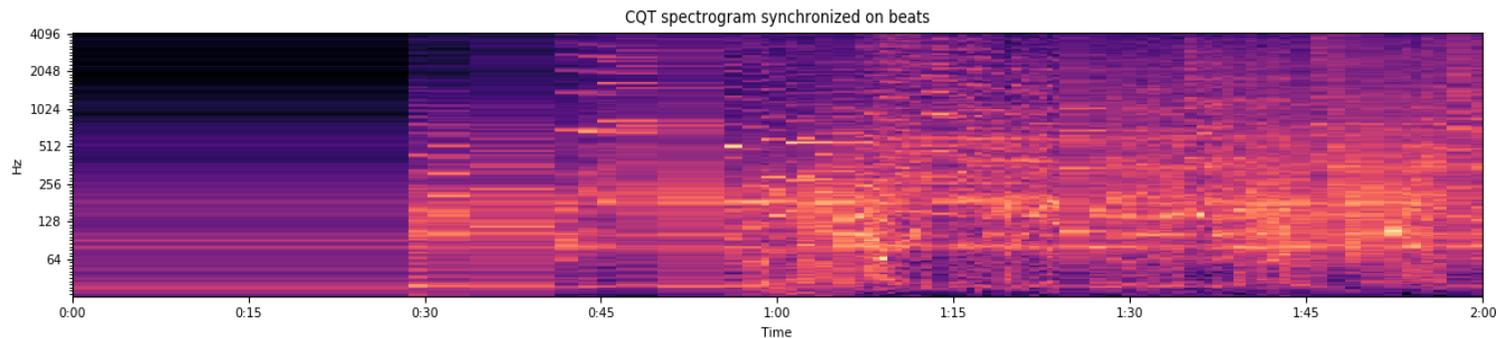
- On retrouve à peu près la structure
- Les clusters peuvent être mélodiques, rythmiques ou timbraux
- Plusieurs paramètres à choisir 😞 mais outil flexible 😊
- Disponible à l'adresse:
<https://git-lium.univ-lemans.fr/tahon/spectral-clustering-music>
- Passons à l'analyse du Free Jazz ...

Paramètres choisis pour le Free Jazz

30ms

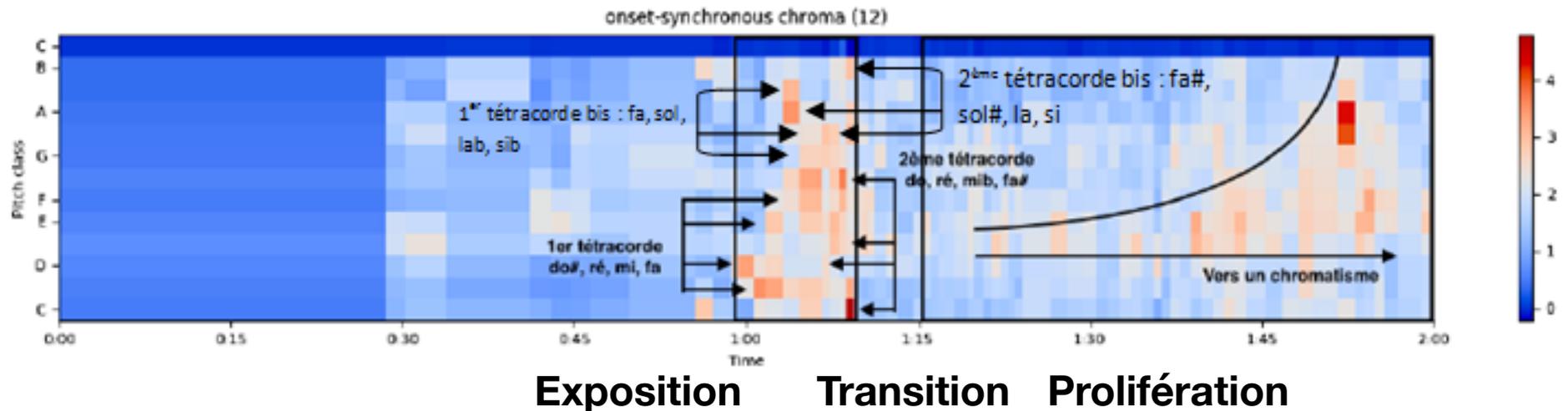


Segmentation
Manuelle



- Résolution temporelle donnée manuellement
- Représentation spectrale à long terme: chroma
- Nombre de clusters: $K=4, 6$

Représentation temporelle: chroma



- Exposition: 2 tétracordes (+ 2 mineurs)
- Transition
- Prolifération

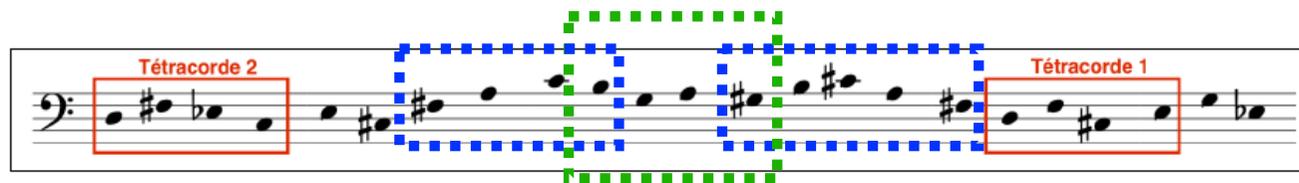
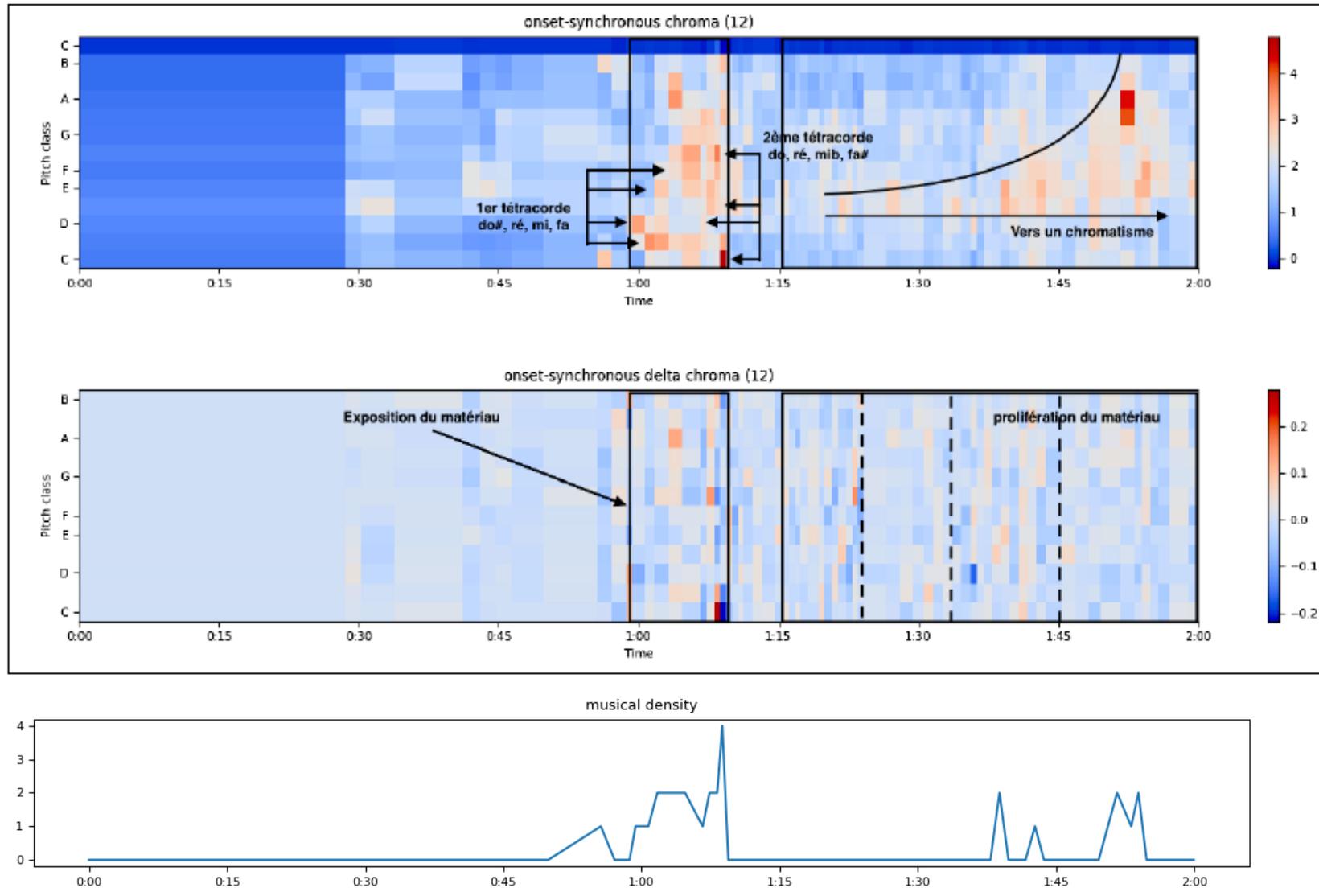
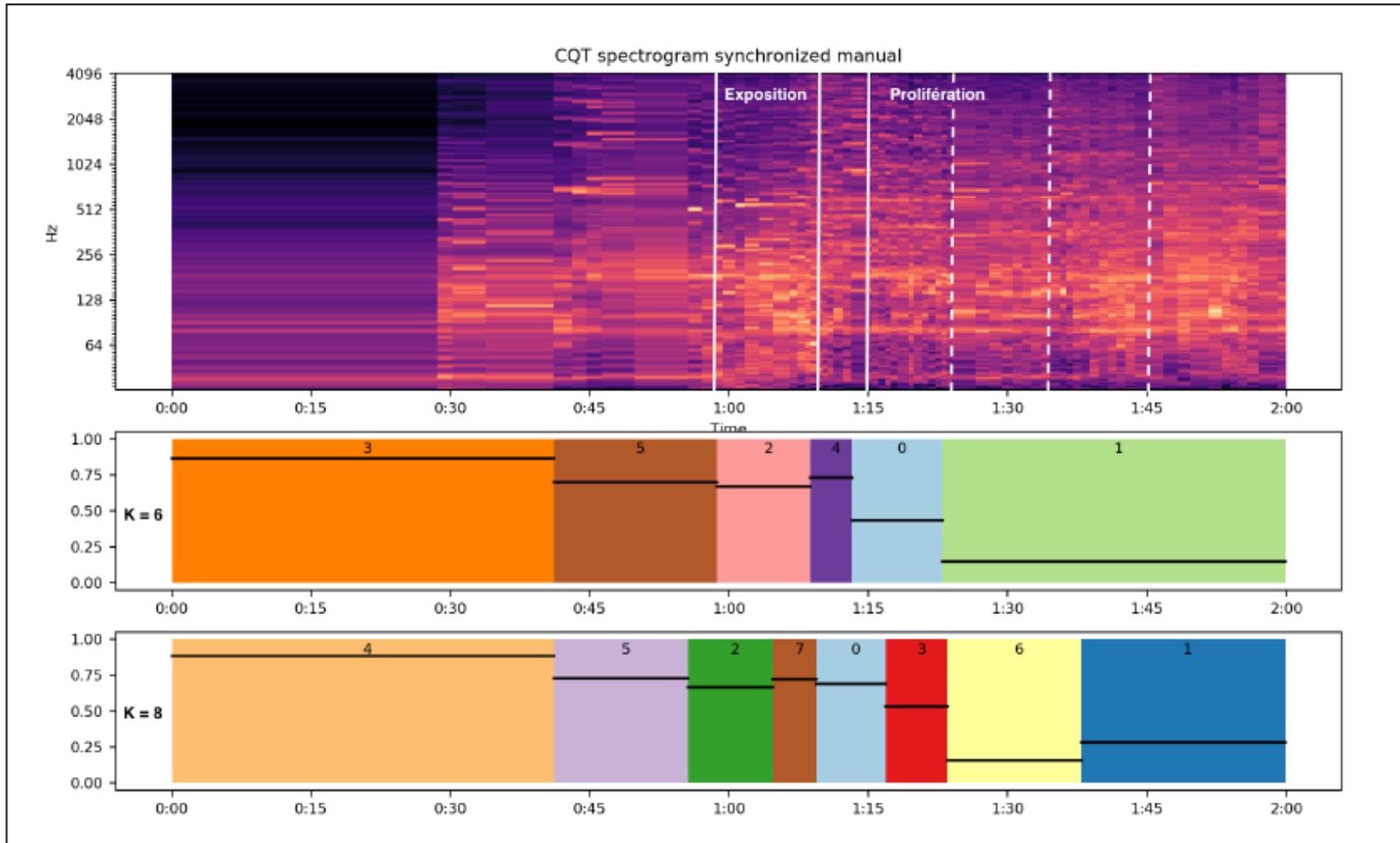


Figure 2. Série de 23 notes extraite de *Globe Unity*.

Transformée différentielle (Δ chroma)



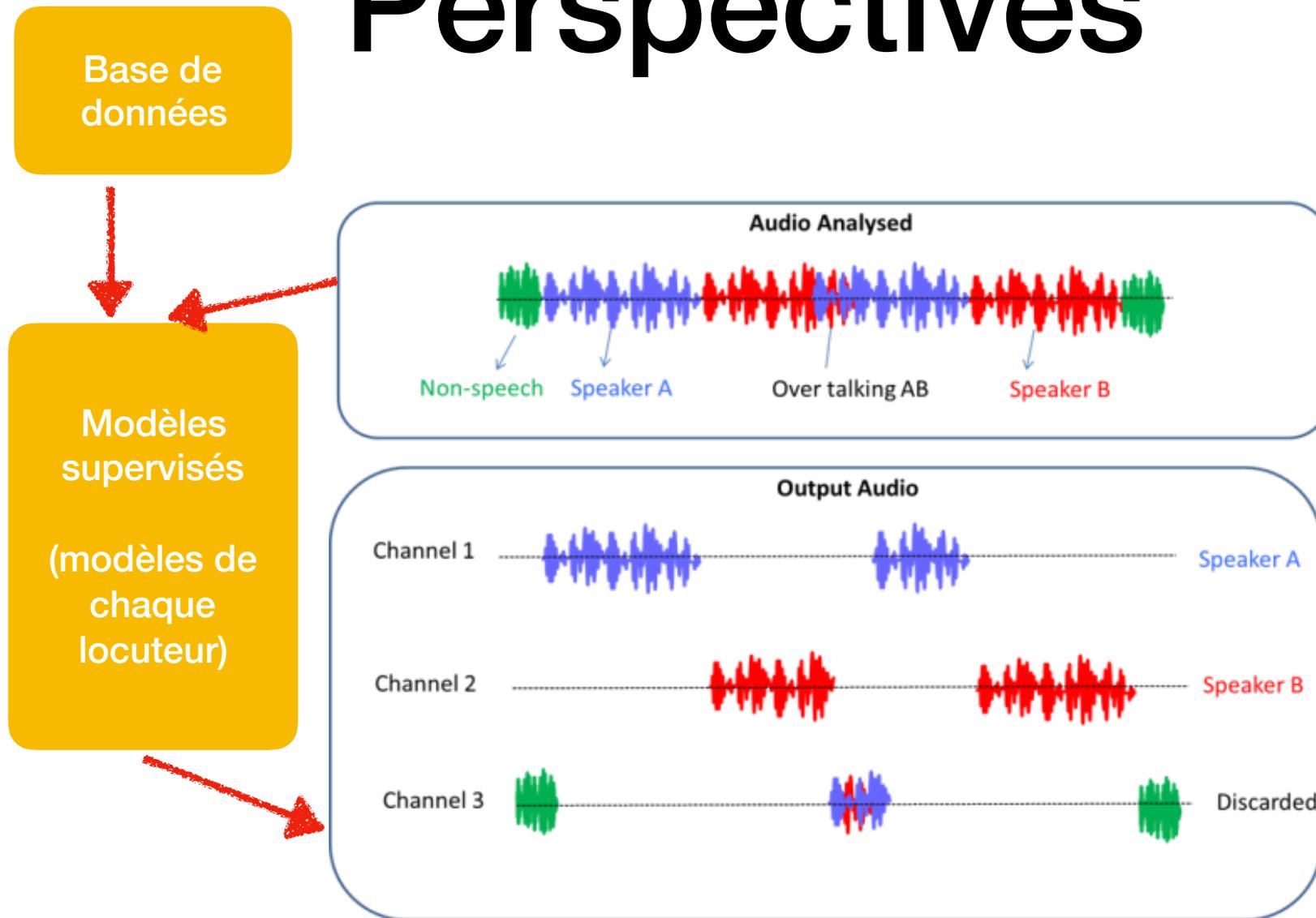
Clustering



Bilan

- Pour la segmentation
 - difficulté de la tâche
 - question de l'évaluation sur des musiques non mesurées.
 - évaluation sur d'autres types de geste: parole expressive
- Pour l'analyse:
 - ajout de mesures de densité harmonique et rythmique
 - la performance Free Jazz: déterminée / suite d'accidents ?
 - proposer un niveau de structure clair et dégager une *Gesthalt*
- Questions d'ordre esthétique
 - Hyper formalisme qui s'est développé à la fin des années 1950

Perspectives



Speaker diarization

Perspectives

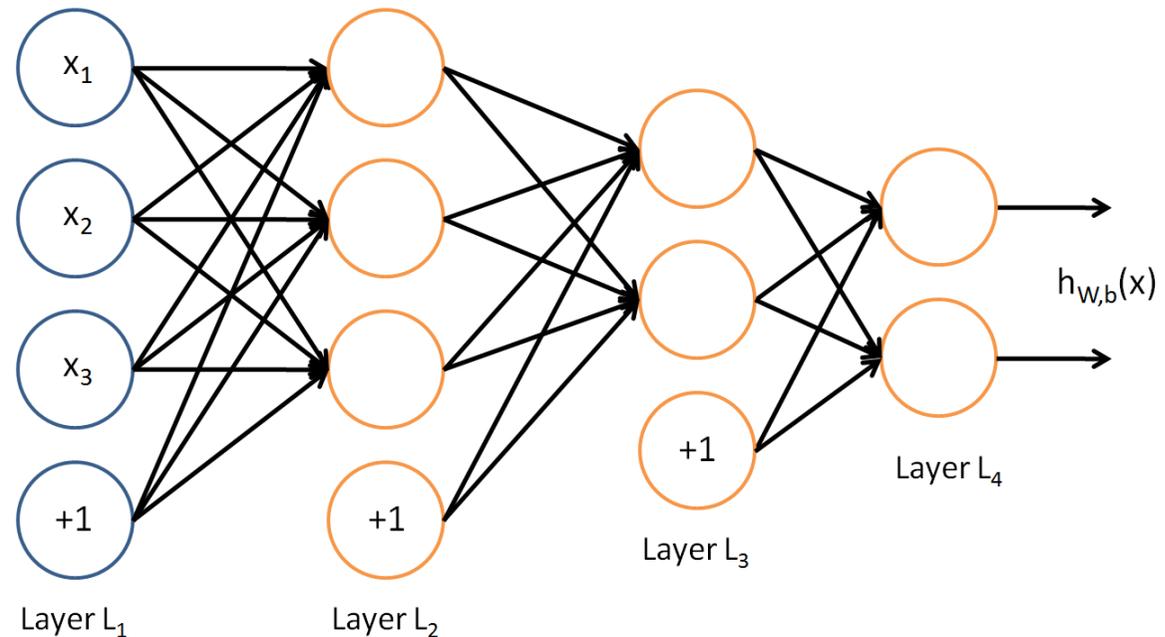
Base de données



Modèles supervisés

(modèles de chaque locuteur)

Modèles types réseaux de neurones



- Très performants mais besoin de beaucoup de données (100h)
- Travaux en cours sur l'interprétation
 - ▶ **Quelle représentation interne de la structure musicale ?**

Perspectives

- Modèles supervisés pour la recherche de structures musicales ?
 - ▶ Base de données segmentées et annotées manuellement
 - ▶ Segmentation + prédiction de l'appartenance à une classe.
- Quelle interprétation ?
 - ▶ Quelle représentation interne au système ?
 - ▶ Lien avec les activités cognitives humaines ?
- Comment évaluer la segmentation et les catégories obtenues ?
 - ▶ Annotation ?
 - ▶ Groupe d'experts ?

MFCCs

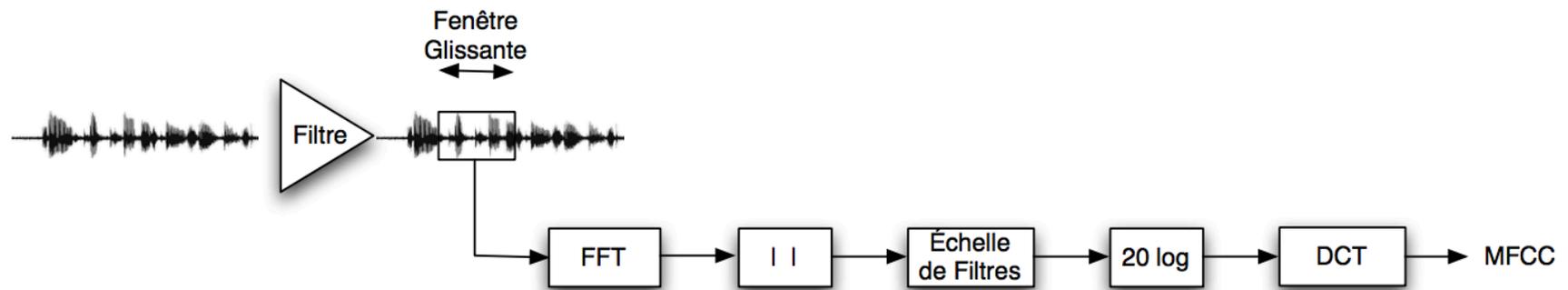


Figure 1.1 – Extraction des paramètres MFCC.

Larcher - HDR (2018)