

Identification automatique de fichiers audio.

Johan MATHE - Linda CALVARIN

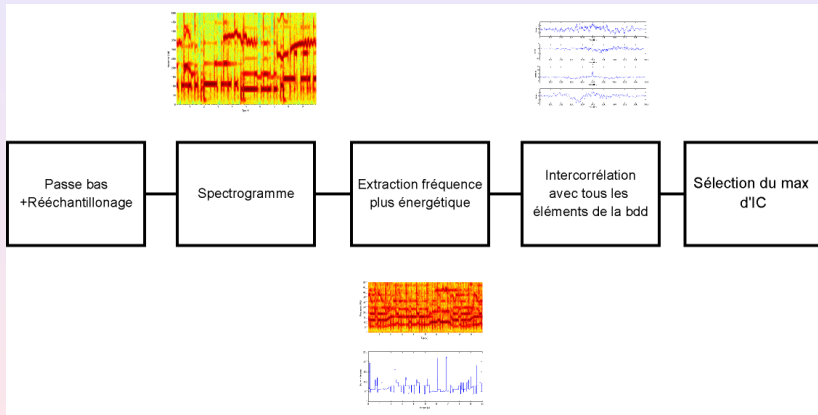
SY16

11 janvier 2008

Problématique - définition des objectifs

- Objectif : mettre en place un système de reconnaissance de chansons
- Par rapport au contenu
- Indépendamment du taux de compression, bruit, échantillonnage

Synoptique de l'algo



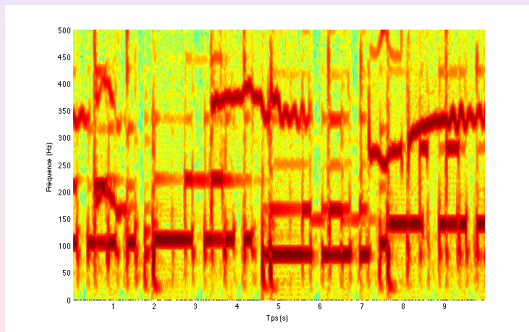
Rééchantillonnage

Après avoir fait quelques expériences, ils s'avère qu'un rééchantillonnage à basse fréquence (1KHz) donne de très bons résultats pour plusieurs raisons :

- Moins complexe : moins de valeurs
- Plus grande stabilité de l'algorithme d'extraction de fréquence la plus intense
- Suppression des bruits HF. Robustesse au bruit.

Spectrogramme

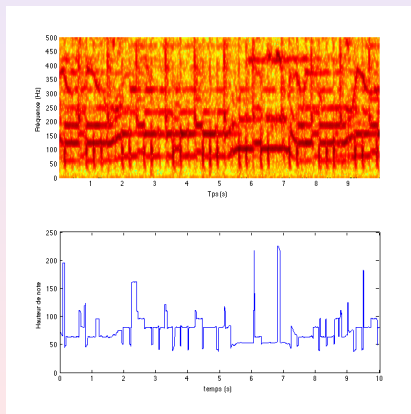
- Une dizaine de secondes au milieu du morceau



- Détection facile de la fondamentale entre 0 et 500 Hz

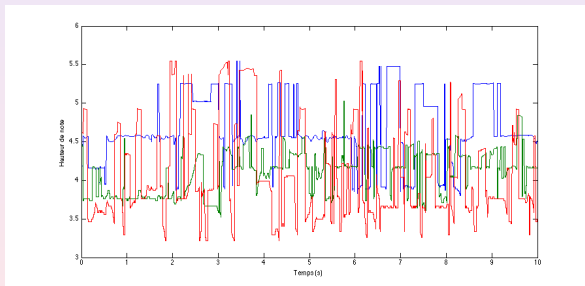
Détection de la fréquence dominante

- Selection du max dans le spectre



Construction d'une base de données de musiques

- Mini BDD empreintes



Sélection de la meilleure chanson

- Impératif de méthode résistante aux décalages temporels
- Intercorrélation
- Max d'intercorrelation
- Permet de détecter les décalages

Graphes

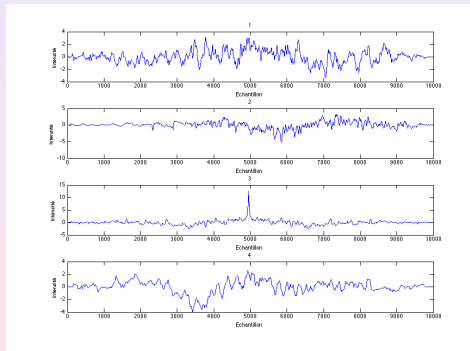


FIG.: Intercorrélations entre signaux de la bdd

Ajout d'un bruit

Le premier test de robustesse de robustesse a été l'ajout d'un bruit blanc en vue de mesurer l'impact de celui ci sur la mesure

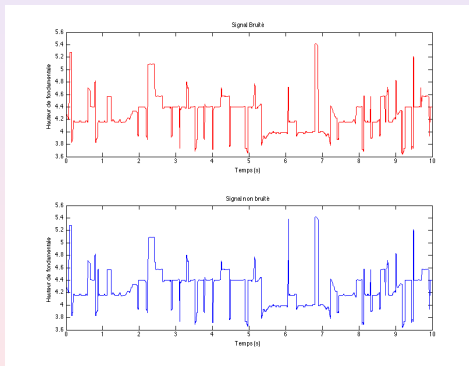
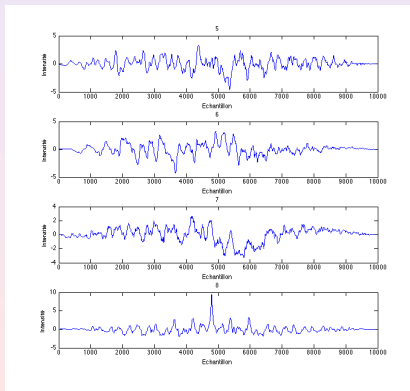


FIG.: Comparaison signal bruité/ Non bruité

Test avec deux musiques de deux sources différentes

- Une musique trouvée sur youtube et l'autre en MP3
- Détection de décalage



Complexité et algorithmique

- Complexité mémoire OK : $5K_{ech} \times NE$ on reste donc en $O(n)$ au niveau mémoire.
- Complexité en temps pour chaque morceau $O(n^2)$ par rapport à la taille de l'échantillon ($5 \cdot 10^3 \cdot 5 \cdot 10^3 = 25 \cdot 10^6$ opérations).
- $O(n)$ par rapport au nombre de musiques dans la bibliothèque.
- $O(nm^2)$ n le nombre de morceaux dans la base de données et m la taille de l'échantillon.
- Moyenne et pire

Optimisations

- Critère d'ordre
- Base triée : accès en $\log(n)$ à la donnée recherchée
- Voir du côté du Data Mining