

Étude de similarité acoustique entre différents instruments, techniques de jeux et perception acoustique

En utilisant la représentation scattering

Christian EL HAJJ

Sous la supervision de :
Dr. Mathieu lagrange

16/09/2016



**CONSERVATOIRE
NATIONAL SUPÉRIEUR
DE MUSIQUE ET
DE DANSE DE PARIS**



Introduction

Fouille de données dans le domaine de la musique.

- Extractions des informations.
- Recherche plus facile.
- Applications utiles (Spotify, shazam ...).

Le problème qu'on a traité

- Outil de recherche basé sur les extraits musicaux.
- Donner des suggestions en se basant sur :
 - 1 Les instruments.
 - 2 Les modes de jeux.
 - 3 Les jugements perceptifs.

Sommaire

Première partie :

Classements en considérant les classes d'instruments et les modes de jeux.

- 1 Base de données
- 2 Algorithme
- 3 Extractions des descripteurs
 - MFCC
 - Scattering
 - Exemple de scattering
- 4 Traitement des descripteurs
- 5 Résultats

La base de donnée SOL

Composée de 25119 fichier qui peut être décomposé en :

- 1 16 classes** d'instruments.
- 2 32 classes** d'instruments avec variations.
- 3 498 classes** de mode de jeux.

Extraits sonores de la base de données

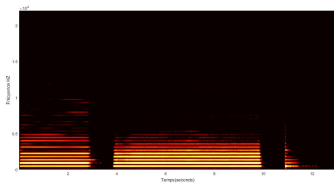


FIGURE – Accordéon , Flute et harpe joué en ordinario

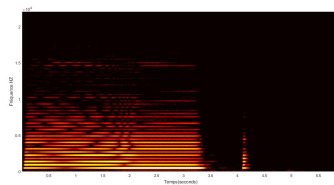


FIGURE – Accordéon joué en ordinario et accordéon joué en staccato

Algorithme

Procédure de calcul :

- Regroupement des fichiers.
- Extractions des descripteurs.
- Traitement des descripteurs.
- Calcul des distances.
- Calculs des métriques de classements.

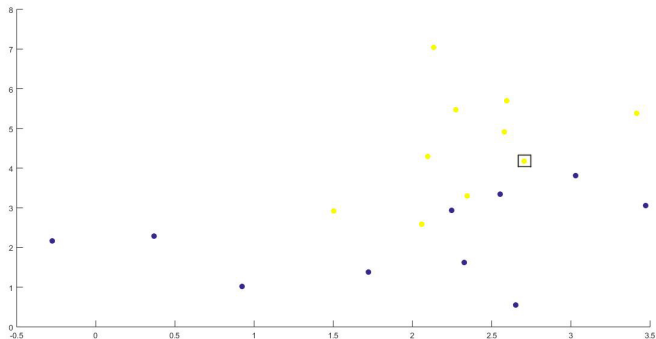
MFCC

MFCC

Algorithme d'extraction des MFCC :

- Fenêtrage.
 - Application de la transformé de Fourier.
 - Calcul du log de l'amplitude des fréquences.
 - Projection des fréquences vers l'échelle de mel.(échelle logarithmique)
 - Application de la transformée en cosinus discrète.
 - Troncature de l'espace.
- + Un espace compact et bien descriptif.
- Pertes d'informations relative aux longues variations temporelles.

Precision à 5



$$Precision \text{ à } k = \frac{\text{nombre de document de meme classe}}{k}$$

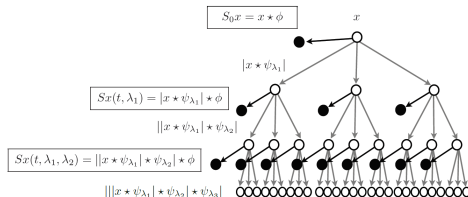
MFCC Choix des paramètres

Réglages des paramètres de MFCC	
Descripteurs	Precision à 5
16 classes d'instruments	
mfcc 12/27	86
mfcc 27/27	85
mel	63
32 classes d'instruments avec variation	
mfcc 12/27	84
mfcc 27/27	83
mel	60
498 classes de mode de jeux	
mfcc 12/27	44
mfcc 27/27	44
mel	33

Scattering

Algorithme de scattering :

- 1 Fenêtrage.
 - 2 Moyenne des trames en appliquant un filtre passe bas.
 - 3 Restauration des informations perdus en appliquant un banque de filtres d'ondelettes.
 - 4 Introduction d'une non linéarité en prenant le module.
 - 5 Prendre les résultats de l'étape 2 en sortie.
 - 6 Répéter les étapes 2 à 4 n fois.
- + Récupérations des informations par l'ajout des étapes.
- Espace de descripteurs à haute dimensionnalité.



exemple de scattering

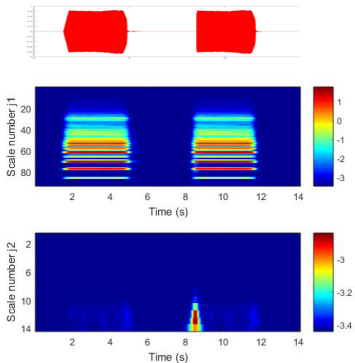


FIGURE – attaque scattering

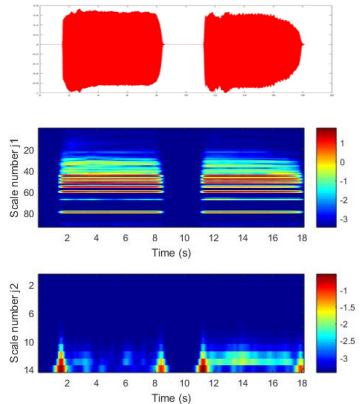


FIGURE – vibrato scattering

Traitement des descripteurs

MFCC

- Standardisation.

Scattering

- Standardisation.
- Normaliser les descripteurs (std and median proposé par Vincent LOSTANLEN) :
 - a Eliminer les petites variances.
 - b Diviser par la mediane.

Résultats 1/3

Ranking metrics pour 16 classe d'instruments

Descripteurs	Traitements	Precision à 5
MFCC	Brut	86
MFCC	Stand.	87
scattering	Brut	70
scattering	Stand.	79
scattering	Std and median	94

Résultats 2/3

Ranking metrics pour 32 classe d'instruments avec variations		
Descripteurs	Traitements	Pat5
MFCC	Brut	84
MFCC	Stand.	85
scattering	Brut	67
scattering	Stand.	75
scattering	Std and median	91

Résultats 3/3

Ranking metrics pour 498 classes d'instruments avec différentes techniques de jeux

Descripteurs	Traitements	Pat5
MFCC	Brut	44
MFCC	Stand.	45
scattering	Brut	37
scattering	Stand.	47
scattering	Std and median	58

Sommaire

Deuxième partie :

Classements en considérant les jugements perceptifs.

- 1 Introduction
- 2 Base de donnée et algorithme
- 3 Large Margin Nearest Neighbor (LMNN)
- 4 Résultats de LMNN sur les classes physiques d'instruments et mode de jeux
- 5 Résultats de LMNN sur les jugements perceptives

La partie câblé.

- la cochlée et les premier étapes de traitement.
- Aspect physique du son.
- problème traité par le scattering.

La partie cognitive.

- La psychologie du cerveau.
- La culture et l'éducation.
- Solution avec des LMNN basés sur des jugements perceptifs.



base de donnée

- 78 fichiers de différents instruments avec différents mode de jeux.
- Jugement par 32 utilisateurs donnant 32 différents vecteur de labels.

Algorithme

- Extractions des descripteurs pour les 78 fichiers.
- Traitement des descripteurs.
- Introduction d'une méthode d'apprentissage supervisé nécessaire pour la résolution du problème perceptive.
- Calcule des distance Euclidienne deux à deux pour chaque jugements.
- Calcule des métriques de classements pour chaque jugement.
- Moyenner sur les résultats des métriques de classements.

LMNN

objectives

- K plus proche voisins appartiennent à la même classe.
- Éloigner les exemples de différents classes d'une certaine marge.
- Trouver la Matrice de manahobis qui assure ces contraintes

$$D_M(\vec{x}_i, \vec{x}_j) = (\vec{x}_i - \vec{x}_j)^T M (\vec{x}_i - \vec{x}_j),$$

$$M = L^T L$$

- Optimisation à pénalisation double.

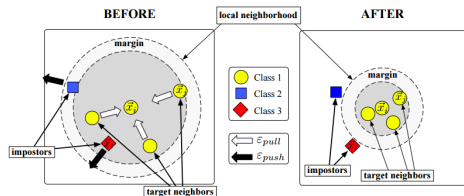


FIGURE – Illustration de LMNN par Weinberger

Pénalisation double

Pénalisation double

- Optimisation à pénalisation double.
- $\epsilon_{pull}(L) = \sum_{j \rightsquigarrow i} ||L(\vec{x}_i - \vec{x}_j)||^2$
- $\epsilon_{push}(L) = \sum_{i,j \rightsquigarrow i} \sum_l (1 - y_{il}) [1 + ||L(\vec{x}_i - \vec{x}_j)||^2 - ||L(\vec{x}_i - \vec{x}_l)||^2]_+$
- $\epsilon(L) = (1 - \mu)\epsilon_{pull}(L) + \mu\epsilon_{push}(L)$
- SDP(semi definite programming).

LMNN

- LMNN utilisé pour renforcer la précision pour la partie cognitive.
- Efficacité du LMNN testé sur l'aspect physique.

Résultats sans division train test

Résultats de LMNN sur les classes physiques d'instruments et mode de jeux sans division en train test		
descripteurs	p@5 avant LMNN	p@5 après LMNN
16 classes d'instruments		
mfcc	87	87
scattering	94	100
32 classes d'instruments avec variation		
mfcc	85	86
scattering	91	100
498 classes de mode de jeux		
mfcc	45	46
scattering	58	88

Résultats avec division train test

Résultats de LMNN avec division en train test pour les 16 classes d'instruments		
descripteurs	P@5 avant LMNN	P@5 après LMNN
Division en 80% for en-trainer et 20% pour tester		
mfcc	74	75
scattering	83	98
Division en 50% pour en-trainer et 50% pour tester		
mfcc	81	82
scattering	90	99

Comparaison entre MFCC et scattering avec même nombre de descripteurs

Égalisations de nombre de descripteurs

- Ajout de 12 descripteurs de delta MFCC
- Ajout de 12 descripteurs de delta delta MFCC
- Création de nouveaux descripteurs en multipliant deux à deux les descripteurs.
- Extraire les 666 descripteurs de scattering en se basant sur la méthode de Std and median.

Résultats de LMNN sur les MFCC et scattering avec même nombre de descripteurs

features	p@5 avant LMNN	p@5 après LMNN
mfcc	58	78 (87)
scattering	93	100

Avis moyen

algorithme

- Effectuer un alignements des labels en utilisant '*Normalised Mutual Information*' NMI.
- Effectuer une vote majoritaire pour avoir un seul label représentative.
- Calculer les LMNN sur ce vecteur de label.(un seul calcul)
- Projeter vers le nouveau espace (une seul projection).
- Calculer les distances euclidiennes (une seul matrice de distance)

Résultats de LMNN sur les jugements perceptives en effectuant un alignement des labels

Descripteurs	P@5	P@5 après LMNN
mfcc	56	62
scattering	44	61

LMNN par somme des distances

algorithme

- Calculer les LMNN pour chaque jugement.
- Projeter vers le nouveaux espace pour chaque jugement.
- Calculer les distances euclidiennes dans les nouveaux espaces.
- Sommer les distances.

Résultats de LMNN sur les jugements perceptives en sommant les distances

Descripteurs	P@5	P@5 après LMNN	LMNN sur avis moyen
mfcc	56	60	62
scattering	44	66	61

Somme des distances avec pondérations

algorithme

- Calculer les LMNN pour chaque jugement.
- Projeter vers le nouveaux espace pour chaque jugement.
- Calculer les distances euclidiennes dans les nouveaux espaces.
- Calculer pour chaque jugement une valeur de précision en utilisant (Normalised mutual information NMI)
- Sommer les distances après multiplications respectivement par la précision.

Résultats de LMNN sur les jugements perceptives en sommant les distances après pondération

features	P@5	P@5 après LMNN	LMNN sur avis moyen
mfcc	56	61	62
scattering	44	67	61

Somme sur les distances après normalisation

Algorithme

- Calculer les LMNN pour chaque jugement.
- Projeter vers le nouveaux espace pour chaque jugement.
- Calculer les distances euclidiennes dans les nouveaux espaces.
- Normaliser les distances.
- Sommer les distances normalisées.

Résultats de LMNN sur les jugements perceptives en sommant les distances après normalisation

features	P@5	P@5 après LMNN	LMNN sur avis moyen
mfcc	56	60	62
scattering	44	67	61

Conclusion








Conclusion

Les résultats sont très satisfaisants. On a présenté :

- Solutions du problème physique avec une très grande précision.
- Une approche originale pour combiner l'aspect physique et cognitive.

Ouverture

- Une étude approfondie des coefficients de scattering pour un problème bien posé et original.
- Étudier la convergence des résultats pour un plus grand nombre d'experts et d'objets d'études.

-  Beth Logan *Mel Frequency Cepstral Coefficients for Music Modeling*. 2000
-  J. Andén and S. Mallat. *Multiscale scattering for audio classification*.. ISMIR 2011
-  J. Andén *Time and frequency scattering for audio classification*. January 7, 2014
-  Hermann Ludwig Ferdinand von Helmholtz *On the sensations of tone as a physiological basis*. 1895
-  K. Q. Weinberger, L. K. Saul. *Distance Metric Learning for Large Margin Nearest Neighbor Classification*. Journal of Machine Learning Research (JMLR) 2009
-  P. Xing, A. Y. Ng, M. I. Jordan, and S. Russell *Distance metric learning, with application to clustering with side-information*. Cambridge, MA, 2002.
-  Kraskov Alexandre and al. , *Hierarchical Clustering Based on Mutual Information* Grassberger, Peter (2003)