

Transformée de scattering temps-chroma-octave

9 mars 2015

Un défi majeur de la classification automatique de sons repose sur une modélisation efficace de leur structure transitoire sur des échelles temporelles aussi longues que possible. De par sa bonne localisation temps-fréquence et sa faculté de régularisation des signaux modulés, un opérateur non-linéaire tel que le module de la transformée en ondelettes est un premier pas naturel dans ce sens. Cependant, celui-ci est incapable de capturer, par simple intégration temporelle, des éléments acoustiques plus riches tels que les variations de fréquence fondamentale (*chirps*) ou de profil formantique (coarticulations, attaques instrumentales). Or, si le cas des *chirps* et de la variabilité harmonique ont été abordés indépendamment ([Gribonval], [Peeters et al.]), il n'existe pas d'approche systématique qui rende compte de la dynamique jointe de ces deux facteurs.

Dans cette communication, nous introduisons une nouvelle représentation des sons, construite à partir du module de la transformée en ondelettes, visant explicitement à caractériser les changements de hauteur et de timbre. Dans une première partie, nous montrons comment enrouler l'axe fréquentiel en une spirale des hauteurs de sorte qu'un tour complet correspond à une transposition d'une octave, afin de séparer hauteur relative et registre global. Par la suite, nous définissons un opérateur unitaire et multi-échelles sur la spirale obtenue, construit comme une cascade de trois transformées en ondelettes à valeurs complexes.