

Detection de la structure d'une chanson

Michel Buffa
buffa@unice.fr

20 février 2017

Nombre d'étudiants souhaités : 1-4

Description du sujet

Dans le cadre du projet de recherche WASABI, nous avons construit une base de données de 2 millions de chansons, avec pour chaque chanson de multiples métadonnées (paroles, artistes, compositeurs, orchestration, émotions, analyse audio complète etc.).

Dans ce projet, l'équipe Wimmics de l'INRIA est en charge de l'analyse des paroles, et à l'intérieur de cette tâche figure la détection de structure (couplet, refrain, chorus, solo, intro, outro etc.). On trouve de la littérature sur le sujet, mais dans notre cas précis, nous avons prévu une approche originale mêlant l'analyse audio, effectuée par l'IRCAM à Paris (partenaire du projet de recherche) et l'analyse textuelle. En effet, par pure analyse audio on est capable de détecter les parties chantées d'une chanson (et donc celles où il y a des paroles) et aussi la "structure harmonique et temporelle de la chanson". Chaque analyse de structure (audio et textuelle à partir des paroles) a ses erreurs (exemple : confondre un couplet et un refrain) alors que l'approche mixte (paroles et audio interchangeant leurs données respectives) devrait donner de meilleurs résultats.

Le sujet de TER propose donc de travailler sur ce sujet précis : effectuer une analyse de structure de chansons basée sur à la fois l'analyse du texte des paroles et l'analyse audio (en collaboration avec IRCAM, le TER ne consistera pas à faire l'analyse audio mais plutôt à utiliser les logiciels d'analyse de l'IRCAM ou peut-être juste des résultats d'analyse fournis par eux sur un ensemble de chansons).

Le travail consistera dans un premier temps en un état de l'art des techniques existantes, puis dans un second temps en des implémentations de test (technologies : libres à priori mais si possible : JavaScript/NodeJS/librairies externes...). Dans un troisième temps il faudra réfléchir à un modèle original permettant de stocker les données de structure en base (mongoDB mais aussi base de données de graphe RDF + ontologie).

Ne vous inquiétez pas si vous ne connaissez pas les technologies évoquées, vous travaillerez avec 1) deux ingénieurs qui développent les logiciels 2) des chercheurs en charge du projet 3) un étudiant en thèse spécialiste de l'analyse des paroles, mais qui de son côté travaille sur d'autres aspects de l'analyse. Il vous sera d'une aide précieuse.

Lieu

Le projet sera développé sur le site des Lucioles et aussi des Templiers, lieu où réside l'équipe de recherche Wimmics. Les logiciels seront développés en suivant les règles de l'open source (projet GitHub notamment).

Le projet peut déboucher sur un apprentissage en vue de faire une thèse par la suite.

Prérequis

Rien de spécial.

Informations complémentaires