

Séparation de sources informée

Mots clés :

- **Directeur de thèse** : GAEL RICHARD
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire Traitement et Communication de l'Information
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

La problématique de cette thèse s'inscrit dans le cadre d'un domaine relativement récent du traitement du signal : la séparation de sources. La séparation de sources consiste à estimer des signaux sources inconnus à partir de l'observation de mélange(s) de ces signaux. Des techniques relativement efficaces (quoiqu'imparfaites) existent quand le nombre de signaux mélangés observés est supérieur ou égal au nombre de signaux à séparer. Les choses se compliquent singulièrement dans le cas contraire, appelé cas « sous-déterminé » ou « dégénéré ». Dans cette thèse, nous proposons de développer une technologie d'écoute active pour des signaux stockés sur un banal CD-audio stéréo, ne contenant donc pour chaque œuvre musicale que deux canaux audio, les traditionnelles voies gauche et droite de la stéréo. En général, ces deux voies sont redondantes et, en toute généralité, elles peuvent contenir chacune un grand nombre de voix et d'instruments, selon le type de musique et de formation musicale l'interprétant. La tâche de séparation des différents éléments de la scène sonore est alors très difficile, voire impossible à effectuer avec une approche dite « aveugle », c'est-à-dire sans connaissances a priori sur ces sources ou leur procédé de mélange (la faible quantité d'information disponible dans les observations par rapport à celle présente dans les sources est liée au degré de « superposition » des différentes sources dans le mélange, qui peut être très important dans certains cas de figure). Pour mener à bien ce défi d'extraction des différentes pistes dans le CD-audio, on propose d'exploiter la possibilité d'assister cette tâche de séparation d'un ensemble d'informations fournies en amont. Ces données peuvent en premier lieu être les signaux sources eux-mêmes, enregistrés séparément en studio ainsi que le contrôle du processus de mixage utilisé. Un autre exemple est le travail d'un opérateur fournissant aux algorithmes des informations pertinentes pour la séparation. Parmi ces dernières on peut trouver la hauteur des notes jouées, la nature des instruments présents à chaque instant, etc. Il s'agit alors d'extraire de ces données disponibles en amont une information pouvant être exploitée par le processus de séparation. Cette information, plus ou moins riche et volumineuse, peut par exemple décrire plus ou moins précisément la structure des différentes sources ou leur contribution au mélange dans le plan temps-fréquence, de façon à lever les difficultés de la séparation. Nous introduisons alors le concept de séparation de sources informée. Dans le cadre du CD-audio, l'absence de canal spécifique pour stocker cette information nous amènera à la coder dans un signal de tatouage (inaudible) inséré dans les deux pistes audio, ce qui imposera des contraintes fortes sur le débit disponible. Enfin, la séparation de sources et le tatouage n'ont de sens qu'en regard des transformations que l'on désire faire subir au son musical. On pourra ainsi s'intéresser aux effets utilisés par les ingénieurs du son dans les studios d'enregistrement et plus particulièrement ceux qui sont accessibles à un auditeur « grand public ». La transformation de base est la modification des coefficients de mixage permettant le contrôle individuel du volume : atténuer (voire supprimer) ou amplifier (voire isoler) une voix ou un instrument du mélange. L'égalisation permet, elle, de modifier ce volume en fonction du contenu fréquentiel de chaque source. Mixage et égalisation sont à la base du processus de « mastering » réalisé en studio avant la gravure du CD-audio. Ces effets sont souvent utilisés pour simuler la spatialisation des sources sonores : l'information spatiale (typiquement des différences d'amplitude et de phase) est encodée dans les deux canaux stéréophoniques. D'autres transformations sont également envisageables, telles que la transposition (changement de hauteur) d'une source, ou même l'éirement temporel (ralentissement ou accélération du rythme de la musique).

Résumé du projet de recherche (Langue 2)

En raison de la nouveauté de l'angle de vue du problème de séparation, les défis scientifiques sont nombreux et demanderont de développer des connaissances dans plusieurs domaines connexes (séparation, représentations des signaux, codage et tatouage et transformations sonores).

Informations complémentaires (Langue 1)

Participation au projet collaboratif ANR-Dream. Participation possible au projet international QUAERO et au réseau européen 3Dlife.