

LaBRI

Modélisation du son, de la parole et de la musique

**Myriam Desainte-Catherine, Pierre Hanna, Matthias Robine,
Jean-Luc Rouas**

*Laboratoire Bordelais de recherche en informatique (LaBRI)
351, cours de la Liberation
F-33400 Talence
prenom.nom@labri.fr*

Introduction

Le Laboratoire Bordelais de recherche en informatique (LaBRI) développe un partenariat avec le Studio de Création et de recherche en informatique et musique électroacoustique (SCRIME) (<http://scime.labri.fr>), dont l'objectif est de créer des outils pour la création artistique. Les recherches s'effectuent dans le thème « Modélisation du Son, de la Musique et de la Parole » (<http://www.labri.fr/index.php?n=ImageSon.ImgInformatique>) de l'équipe « Image et Son » et suivent actuellement deux grands axes que sont l'interactivité pour le spectacle vivant et la navigation dans les données.

Interactivité et systèmes temps réels pour le spectacle

Des modèles de partitions prenant en compte les types d'interactions possibles en temps réel sont actuellement en cours de développement, ainsi que diverses techniques permettant des interactions nouvelles dans le domaine du spectacle vivant multimédia.

Partitions interactives

Dans le cas des partitions écrites, le modèle des partitions interactives intégrant des informations temporelles utiles pour permettre l'interprétation, a été étendu à d'autres types d'interactions incluant notamment le parcours dans des œuvres ouvertes, avec l'introduction de la logique et de la rétroaction sur un même objet temporel. L'étude théorique du modèle a été basée sur les structures d'événements et le calcul de processus concurrents.

Ces travaux sont actuellement en cours d'implémentation dans le logiciel I-score développé dans le cadre du projet ANR VIRAGE et les extensions du modèle se développent dans le cadre des projets ANR OSSIA et INEDIT.

Projet « Fluxus »

Ce projet consiste à étudier le suivi vocal en temps réel d'un acteur permettant de modifier l'environnement. En se basant sur la double expertise d'algorithmique du texte et d'analyse musicale proposée par les chercheurs ayant collaboré sur le projet ANR SIMBALS, un algorithme avancé d'inférence structurelle a été mis en place et optimisé afin de fournir les résultats les plus cohérents possibles avec la perception de la structure musicale (couplets, refrains etc.).

Interaction et émotions

Suite à des travaux sur la modélisation sonore liée à la perception sonore, des études considèrent actuellement la prise en compte de la physiologie et en particulier des émotions pour l'interaction musicale. Ainsi, un système de génération de musique piloté par les émotions a été élaboré. Celui-ci a été utilisé en connexion avec le système eMotion du projet ANR CARE pour permettre à un danseur de piloter la musique par l'expression dansée de certaines émotions.

Instruments virtuels immersifs

D'autres recherches menées dans l'équipe explorent l'utilisation de la réalité virtuelle et de l'interaction 3D pour la performance musicale. Les environnements virtuels immersifs offrent de nouvelles possibilités de représentation, de manipulation et de construction des processus musicaux. Ainsi un dispositif d'interaction appelé « Piivert » a été développé ainsi que plusieurs techniques d'interaction, qui permettent de manipuler de nombreux objets 3D audiovisuels de manière efficace et expressive. Ce dispositif a été combiné avec des travaux précédents, notamment concernant la représentation de processus sonores, dans un instrument immersif : « Drile ». Cet instrument permet de faire évoluer la technique musicale du live-looping vers un live-looping hiérarchique. « Drile » apporte en effet des possibilités de construction de structures musicales complexes et de navigation dans ces structures.

Recomposition de la structure d'un morceau de musique

Cet axe de recherche est dédié à la reconstruction de grandes parties manquantes dans des flux audio (entre 1 et plusieurs secondes). Le système de reconstruction proposé est inédit en traitement du signal sonore et basé sur la répétition musicale, permettant ainsi de combler l'information manquante par l'extrait choquant le moins

la perception musicale dans le reste du morceau. Cette technique a donné lieu à une demande de dépôt de brevet.

Navigation dans les données

Distances musicales et algorithmes de comparaison de données musicales

Le nombre de documents musicaux disponibles en ligne est de plus en plus important et nécessite le développement de nouveaux moyens de recherche par le contenu musical. Ces nouvelles méthodes soulèvent le problème de l'estimation automatique de la similarité musicale. Pour cela, un morceau de musique peut être représenté par des séquences de symboles, obtenus depuis le signal audio, et représentant l'information musicale. Des adaptations d'algorithmes utilisés en bioinformatique permettent la comparaison de ces séquences et autorisent ainsi la comparaison rapide de morceaux de musique. Les applications sont nombreuses : détection de plagiat, systèmes de recommandation musicale, aide à la recherche de documents musicaux, etc.

Logiciel de navigation musicale à partir de l'extraction des informations

Dans le cadre des thèses de Thomas Rocher et Benjamin Martin, des études ont porté sur l'analyse automatique d'informations musicales de haut niveau telles que la structure musicale ou les propriétés tonales, depuis un signal musical audio. Ces informations extraites doivent permettre de proposer de nouvelles méthodes de navigation et de recherche dans les bases de données musicales audio.

L'étude du rythme est également abordée. Le rythme est une propriété importante de la musique Occidentale. Des travaux ont porté sur l'analyse des propriétés rythmiques d'un morceau sous la forme d'un descripteur original. Ce descripteur permet de mettre en relation les morceaux présentant les mêmes propriétés métriques.

Détection de zones d'intérêts

La détection de zones sonores permet de repérer dans un flux audio la nature et la localisation temporelle de composantes primaires pouvant faire l'objet de traitements spécifiques : les zones parlées et les zones contenant de la musique. Cette détection a récemment été étendue pour déterminer à l'intérieur des zones de musique quels sont les moments où le chant est présent. Par ailleurs, les méthodes d'estimation de la similarité musicale ont permis de proposer de nouvelles méthodes pour localiser les parties d'un même morceau les plus répétées, et ainsi permettent d'envisager la détection des parties importantes d'un morceau comme par exemple le refrain.

Caractérisation et identification automatique des langues et des accents

L'approche développée au LaBRI pour l'identification des langues consiste à exploiter les différences de prosodie entre celles-ci. La prosodie peut être définie en quelque sorte comme la *musicalité* d'une langue : elle s'exprime par ses caractéristiques mélodiques (intonation, énergie) et rythmiques. Cette approche permet également de considérer des variations plus fines et d'effectuer des comparaisons entre langues, dialectes ou accents au niveau prosodique, domaine rarement exploré.

Ainsi, lors d'une collaboration avec Mélissa Barkat du Laboratoire ICAR-Praxiling (Montpellier), le système a été appliqué sur une base de données constituée de dialectes arabes, peu étudiés du point de vue de la prosodie. Les résultats obtenus sur cette base de données ont permis de faire l'hypothèse que de fortes différences prosodiques existent entre les différentes zones dialectales de l'arabe.

De la même manière, en collaborant avec Philippe Boula de Mareüil du LIMSI (Orsay), les différences et les similitudes entre différents accents ouest-africains en français ont pu être mises en évidence. Un travail a débuté également depuis peu sur les réalisations prosodiques de Français apprenant le japonais (ou inversement), en collaboration avec Takaaki Shochi du CLLE-ERRSàB (Université Bordeaux 3).

Caractérisation et identification de la voix chantée ou actée

La poursuite du travail sur la détection de zones de voix chantée consiste actuellement à déterminer précisément quelles sont les différences entre la voix parlée et la voix chantée, en tentant de prendre en compte des états supposés intermédiaires tels que la voix actée. Une étude a également commencé sur la possibilité de distinguer automatiquement différents styles de chant.

Remerciements

Ces recherches se déroulent en partenariat avec le SCRIME (scrimelabri.fr), soutenu par la DGCA du ministère de la Culture et le Conseil régional d'Aquitaine. Le SCRIME est une convention de coopération entre l'Université Bordeaux 1, l'IPB (Institut Polytechnique de Bordeaux) et le conservatoire de musique de Bordeaux.