

Détection de tempo pour la transcription musicale

proposition de stage de recherche

Contexte

La détection du tempo à partir d'une performance musicale est un problème difficile dans le domaine de l'extraction d'information musicale (MIR). Plusieurs approches ont été proposées, parmi elles le modèle perceptif de [1], fondé sur un modèle physique de synchronisation d'oscillateurs, et validé par des expériences avec des musiciens. Ce modèle a été appliqué en particulier au problème du suivi de partition [2] (suivi en temps réel d'un instrumentiste humain par une machine, afin par exemple, de jouer un accompagnement au bon tempo).

Objectifs

Ce stage vise à intégrer un module de détection de tempo implantant [1] dans la plateforme de transcription musicale automatique qparse [3]. Ce système converti une séquence d'événements musicaux datés (messages MIDI) en une partition structurée en notation traditionnelle occidentale, dans un format XML. La suite d'événements en entrée est traitée mesure par mesure, quasiment en ligne, à l'aide de techniques de parsing étendues à des modèles appropriés de la théorie des automates. La partition en sortie est sélectionnée par optimisation du rapport entre sa lisibilité (au sens de la complexité de la notation rythmique) et la proximité des durées de notes avec la performance en entrée. La détection de tempo, par [1], se fera suivant les durées de notes estimées pour la transcription, en ajoutant la minimisation de variations de tempo au critère d'optimisation précédant.

Une évaluation sera faite sur des extraits monophoniques et des voix issues d'enregistrements de Disklavier, si le temps le permet.

Mots clef

extraction d'information musicale, estimation de tempo, transcription musicale

Lieu de travail et gratification

Le stage se déroulera au laboratoire CÉDRIC du Conservatoire des Arts et Métiers (CNAM), Paris, et sera encadré par Florent Jacquemard. Une gratification incluant le remboursement de la moitié du coût de transport sera envisagée suivant la durée du stage.

Contact

Merci d'envoyer votre demande accompagnée d'un CV à florent.jacquemard@inria.fr

References

[1] Edward W. Large, M. R. Jones
The dynamics of attending: how people track time-varying events
Psychological review, 106(1), 119–159, 1999.
<https://doi.org/10.1037/0033-295X.106.1.119>

[2] Arshia Cont

A coupled duration-focused architecture for realtime music to score alignment

IEEE Transaction on Pattern Analysis and Machine Intelligence 32 (6), 2010.

<https://hal.archives-ouvertes.fr/hal-01161284>

[3] <https://qparse.gitlabpages.inria.fr>

[4] Florent Jacquemard, Lydia Rodriguez de La Nava

Symbolic Weighted Language Models, Quantitative Parsing and Automated Music Transcription

International Conference on Implementation and Application of Automata, CIAA 2022.

<https://hal.archives-ouvertes.fr/hal-03647675>