

Etude d'algorithmes pour la détection de la tonalité de fichiers musicaux et implémentation en Clojure

Antoine Passemiers

Université Libre de Bruxelles
apassemi@ulb.ac.be

Résumé

Le projet consiste en la discussion de différents algorithmes relatifs à la détection automatisée de tonalité de fichiers musicaux, la conception d'un de ces algorithmes à l'aide du paradigme fonctionnel, et la comparaison de son efficacité avec celle obtenue avec une approche impérative. L'algorithme de détection sera évalué sur les critères de la rapidité d'exécution, ainsi que la précision de la détection. De fait, les approches impératives et fonctionnelles seront comparées sur ces deux aspects.

Introduction

La tonalité d'une oeuvre musicale se caractérise par l'ensemble des sons formant une même gamme diatonique. A la différence de la gamme, où les sons se succèdent de façon contiguë, la tonalité (ou ton) regroupe des sons qui peuvent être disjoints et/ou superposés (Danhauser, 1929). En conséquence, nous nous intéressons à l'analyse de mélodies polyphoniques, où plusieurs notes peuvent être jouées en même temps.

L'algorithme utilisé pour la détection de la tonalité repose en partie sur la solution proposée par Ibrahim Sha'ath lors de la conception du logiciel KeyFinder (Sha'ath, 2011).

TODO

La précision de la détection est évaluée à l'aide d'une base de données, constituée de 1000 fichiers musicaux au format wav, dont les tonalités sont connues et inscrites dans un fichier csv. Ces fichiers sont identiques à ceux utilisés par Ibrahim Sha'ath dans le cadre de sa recherche.

TODO

(Pauws, 2004) (Takeuchi, 1994) (Peeters, 2006)

Considérations théoriques

Le signal audio est premièrement extrait du fichier wav, puis la moyenne entre les deux canaux est effectuée si le fichier a été enregistré en stéréo. En effet il n'est pas nécessaire de prendre en compte le panoramique puisque celui-ci n'a que peu d'influence

sur la mélodie dans le domaine spectral. Etant donné que les notes jouées sont uniquement caractérisées par leur fréquence fondamentale, il n'est pas nécessaire de considérer l'entièreté du spectre du fichier musical. De fait, la fréquence d'échantillonnage est abaissée à un dixième de la fréquence standard (4410 Hz), mais ce sous-échantillonnage est susceptible de provoquer des phénomènes d'aliasing. Contrairement à la solution de Sh'ath, qui gère les problèmes d'aliasing sonore par l'application d'un filtre passe-bas, une approche plus simpliste et plus rapide se limiterait à l'application d'une moyenne mobile sur une fenêtre glissante de taille arbitraire. L'avantage de la méthode est de bénéficier d'effets passe-bas sans devoir effectuer de calculs dans le domaine fréquentiel. Pour ce qui est de la taille de la fenêtre temporelle, il s'agit d'un paramètre fixé durant l'évaluation de l'algorithme.

TODO - suite

Implémentation en Clojure

TODO

Implémentation en Python ?

TODO - Ajouts de modèles de Markov cachés pour la gestion de l'aspect temporel ? Approche machine learning ?

Résultats

Rapidité d'exécution

TODO

Précision des algorithmes

TODO

References

Danhauser, A. (1929). *Théorie de la Musique*.

Pauws, S. (2004). Musical key extraction from audio. *Proceedings of the 9th International Conference on Digital Audio Effects, DAFx-06, Montreal, Canada*, pages 96–99.

- Peeters, J. (2006). Musical key estimation of audio signal based on hidden markov modeling of chroma vectors. *Proceedings of the 9th International Conference on Digital Audio Effects, DAFx-06, Montreal, Canada.*
- Shah, I. (2011). Estimation of key in digital music recordings. pages 1–64.
- Takeuchi, A. H. (1994). Maximum key-profile correlation (mkc) as a measure of tonal structure in music. *Perception & Psychophysics*, pages 335–346.