

Sujet de thèse: Accélération des méthodes de recherche de musique sur les nouvelles architectures.

Directeurs de thèse : Pierre Hanna / Pascal Ferraro

Courriels : hanna|ferraro @labri.fr

Équipe: Image et Sons/MaBioVis

Nature du sujet: Traitement Informatique de la Musique

Présentation du sujet

De plus en plus de documents musicaux numériques sont aujourd'hui disponibles en ligne (internet, mobile, ...) ou facilement accessibles. Pour permettre une navigation entre ces documents ne se basant pas seulement sur les métadonnées (noms des morceaux et des albums, interprètes, compositeurs...), de nouvelles méthodes de recherche ont vu le jour, adaptées à la spécificité de la musique. Par exemple, la *requête par fredonnement* (Query-by-Humming) cherche à retrouver un morceau à partir d'un extrait fredonné, et la *requête par l'exemple* a pour but de retrouver des morceaux ressemblant à un extrait donné. Ces requêtes doivent faciliter aux utilisateurs l'accès à de nouvelles pièces musicales.

Ces nouvelles méthodes de navigation reposent en grande partie sur l'estimation de la **similarité musicale** entre deux morceaux. La notion de similarité est très difficile à définir avec précision : c'est un enjeu de modélisation encore largement ouvert dans les recherches en système d'informations musicales. D'un point de vue informatique, estimer la similarité se fait par des algorithmes qui calculent une mesure de ressemblance entre deux segments de morceaux. Plusieurs approches ont été proposées pour calculer une mesure de similarité entre deux morceaux de musique. Mais la majorité de ces approches sont basées sur la similarité du timbre, évaluées par des statistiques sur des descripteurs bas niveau du son. Aller plus loin demande d'avoir accès à des **séquences musicales symboliques**, représentant un morceau comme suite de différentes notes (hauteur, durée, intensité).

Les méthodes basées sur l'alignement de séquences symboliques sont généralement plus **expressives** car elles tiennent compte des erreurs présentes dans la requête ou dans les morceaux de musique de la base de données. En effet, une requête chantée peut être transposée, jouée plus ou moins vite, sans pour autant dégrader la mélodie. Les systèmes de recherche doivent donc être à la fois invariants pour la transposition et le tempo. Les méthodes d'alignement, originellement développées dans le contexte de la bioinformatique, peuvent s'adapter aux applications musicales pour tenir compte de ces propriétés d'invariance. Même si certaines méthodes ont été validées depuis 2007, notamment lors des différents concours MIREX (Music Information Retrieval Evaluation eXchange), l'expressivité de ces méthodes est toujours à améliorer, en prenant en compte des paramètres musicaux pertinents.

Enfin, ces méthodes ont un coût calculatoire très important, et ne sont pas encore assez **efficaces** pour être utilisées sur des bases de données musicales dont la taille augmente de manière exponentielle. Pour les applications biologiques, des heuristiques à base de graines (BLAST, FASTA, puis les graines espacées) ont été proposées pour accélérer la recherche de similarités dans des bases de données biologiques. De telles heuristiques n'ont pas encore été mises en œuvre pour les séquences musicales.

Objectifs du projet

Les limites actuelles en expressivité comme en efficacité empêchent les méthodes symboliques de

s'être vraiment diffusées. Le succès de certaines applications, bien que très imparfaites, travaillant uniquement sur l'audio (telles que Shazam) montre que toute amélioration de ces recherches peut avoir une grande importance sociétale et économique. L'objectif de ce travail de thèse est donc travaillé dans plusieurs directions orthogonales mais complémentaires :

1) *Des recherches plus expressives.* Passer par une représentation symbolique de la musique permet l'utilisation de différents critères (mélodie, harmonie, rythme,...), qui vont au-delà du timbre analysé par les solutions uniquement audio. Une excellente manière d'améliorer l'expressivité serait de considérer des séquences musicales représentées sous forme d'arbres. À court terme, nous aimerions étendre les idées proposées au LaBRI sur les structures arborescentes à l'échelle d'un morceau.

2) *Des recherches plus efficaces.* Les relations de programmation dynamique sur les séquences musicales sont assez similaires à certaines utilisées en bioinformatique. Cependant, des différences dans la taille des alphabets tout comme dans les définitions des similarités font que les techniques "standard" d'indexation fonctionnent mal.

3) *Des implémentations plus efficaces.* L'une des motivations de ce projet sera de proposer des outils informatiques utilisables sur de grandes bases de données. Une solution est pour cela d'exploiter l'incroyable puissance de calcul des processeurs graphiques modernes pour le développement de solutions hautes performances pour les applications de recherche de musique.

La mise à disposition des utilisateurs d'outils efficaces d'analyse de la musique étant aujourd'hui devenu un besoin indispensable à la communauté, ce travail de thèse s'articulera principalement autour de ce dernier point. La thèse s'appliquera à mettre en place des définitions « parallèles » des algorithmes de recherche en vue d'implémentation sur des architectures (éventuellement) hétérogène. En effet, depuis quelques années, les processeurs GPU (General Purpose computation GPU) ont reçu une attention particulière en raison de leurs importantes ressources de calcul. Avec les évolutions récentes des architectures GPU dans un processeur programmable unifié et massivement parallèle, et le développement d'outils de programmation et de langages de haut niveau tel que Cuda de NVidia (*Compute Unified Device Architecture*), les GPU sont devenus une alternative très attractive et à faible coût aux microprocesseurs traditionnels pour des applications demandant de grandes ressources de calculs et dont les algorithmes peuvent s'exprimer sur des flots de données parallèles (*i.e.* le même programme peut s'exprimer sur beaucoup d'éléments de données en parallèle). Ce type de parallélisme paraît très adapté aux applications de type recherche par fredonnement sur de grandes bases de données musicales. Cependant, manipuler de telles quantités de données structurées pose de véritables challenges pour la gestion des mémoires et de l'allocation des ressources de calcul. De manière intéressante ces nouvelles ressources de calcul n'ont pas encore été étudiées au sein de la communauté MIR (Music Information Retrieval), qui s'intéressent aux systèmes d'informations musicales, comme en témoignent les travaux présentés lors des différentes conférences ISMIR (conférence qui fait référence dans le domaine depuis 10 ans). Dans ce projet, afin d'analyser les avantages et les désavantages des cartes graphiques récentes comme nouvelle plateforme de calcul, nous proposons donc d'explorer de nouvelles implémentations des variantes musicales des méthodes d'alignement.

Contexte scientifique du sujet

Ce projet de recherche s'appuiera sur un groupe de travail, issu de la collaboration entre de **jeunes chercheurs du LaBRI à Bordeaux** (Julien ALLALI, Pascal FERRARO, Pierre HANNA) dont les recherches consistent à développer des systèmes d'estimation de la similarité musicale en se basant sur de nombreux critères, essentiellement musicaux, comme la mélodie, l'harmonie, la tonalité, le rythme, etc. Ce groupe de recherche, transdisciplinaire, concerne aussi bien des membres de l'équipe Image et Son (thème **Modélisation du Son et de la Musique**) que des membres de l'équipe **Modèles et Algorithmes pour la Bioinformatique et la Visualisation d'Informations**. Il propose de développer des méthodes couramment employées dans le domaine de la Bioinformatique, et notamment dans la

recherche de séquence d'ADN.

References

Comparaison de séquences musicales

Polyphonic Alignment Algorithms For Symbolic Music Retrieval, Julien Allali, Pascal Ferraro, Pierre Hanna, Matthias Robine, CMMR - Auditory Display, in Lecture Notes in Computer Sciences (LNCS). S. Ystad, M. Aramaki, R. Kronland-Martinet and K. Jensen Editors. Springer Verlag Berlin Heidelberg, vol. 5954, pages 466-482, April 2010.

Toward a General Framework for Polyphonic Comparison Julien Allali, Pascal Ferraro, Pierre Hanna, Costas Iliopoulos, Matthias Robine in the Fundamenta Informaticae Journal FI vol. 97, n°3, pages 331-346, December 2009.

On Optimizing the Editing Algorithms for Evaluating Similarity Between Monophonic Musical Sequences Pierre Hanna, Pascal Ferraro, Matthias Robine in the Journal of New Music Research JNMR vol. 36, n°4, pages 267-279, December 2007.

Indexing Musical Pieces Using their Major Repetition Benjamin Martin, Pierre Hanna, Matthias Robine, Pascal Ferraro in: ACM/IEEE Joint Conference on Digital Libraries (JCDL) Ottawa, Canada, 13-17 June, (2011). (to appear)

Filtres pour les séquences

Bioinformatics Algorithms: Techniques and Applications, chapter A survey of seeding for sequence alignment, D. G. Brown., pages 126–152. Wiley-Interscience (I. Mandoiu, A. Zelikovsky), 2008.

Basic local alignment search tool, S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman *Journal of molecular biology*, Vol. 215, No. 3. (5 October 1990), pp. 403-410.

Rapid and sensitive protein similarity searches, Lipman DJ, Pearson WR. *Science*. 1985 Mar 22;227(4693):1435-41.

Q-gram based database searching using a suffix array (QUASAR). S. Burkhardt, A. Crauser, P. Ferragina, H.-P. Lenhof, E. Rivals, and M. Vingron. In Annual Conference on Research in Computational Molecular Biology (RECOMB 99), pages 77–83, 1999.

Index sur les séquences

Algorithms on Strings, Trees, and Sequences - Computer Science and Computational Biology, Dan Gusfield, Cambridge University Press 1997

Suffix Arrays: A New Method for On-Line String Searches. Udi Manber, Eugene W. Myers. *SIAM J. Comput.* 22(5): 935-948 (1993)

Indexing gapped-factors using a tree. Pierre Peterlongo, Julien Allali, Marie-France Sagot. *International Journal of Foundation of Computer Science* - 2007

The Gapped-Factor Tree. Pierre Peterlongo, Julien Allali, Marie-France Sagot. Prague Stringology Conference (PSC) - 2006