Recherche de Documents Musicaux par Similarité Mélodique

Pierre Hanna — Pascal Ferraro — Matthias Robine — Julien Allali

Laboratoire Bordelais de Recherche en Informatique LaBRI (CNRS - INRIA)
351 Cours de la Libération, F-33405 Talence Cedex
prenom.nom@labri.fr

RÉSUMÉ. *Dans un monde numérique en croissance permanente, la navigation dans de grandes bases de données musicales, comme des catalogues de musique en ligne ou de partitions, ne peut plus se limiter à une recherche par mot-clé textuel, comme un nom d'interprète ou le titre d'une chanson. Les utilisateurs ont besoin de nouvelles méthodes de recherche adaptées à la spécificité de la musique, comme la requête par extrait qui permet de retrouver un morceau à partir d'un court extrait (éventuellement approximatif).*

Ces nouvelles méthodes de navigation reposent en grande partie sur l'estimation de la similarité musicale entre deux morceaux. La notion de similarité est cependant très difficile à définir et le problème de l'estimation de la similarité musicale reste un des problèmes majeurs de la recherche en système d'informations musicales. D'un point de vue informatique, cela consiste à déterminer des algorithmes qui calculent une mesure indiquant le degré de similarité entre deux segments de morceaux. Plusieurs approches ont déjà été proposées en ce sens. Mais elles sont en majorité basées sur la similarité du timbre, évaluée par des statistiques sur des descripteurs bas niveau du son. La prise en compte d'autres informations musicales, telles que la mélodie par exemple, doit aussi être étudiée.

Pour cela, nous proposons dans un premier temps d'étudier les différentes représentations de la mélodie d'un morceau. Nous discutons des limitations de chacune, en fonction des applications souhaitées. Nous proposons ensuite d'adapter des méthodes couramment employées dans le domaine de la BioInformatique, notamment dans la recherche sur les séquences ARN, au contexte de la recherche d'informations musicales. Ces méthodes utilisent des algorithmes d'édition ou d'alignement qui cherchent à déterminer le nombre optimal d'opérations nécessaires pour transformer une séquence musicale en une autre. Pour cela, un coût est associé à chaque opération élémentaire d'édition. Des améliorations sont détaillées, permettant par exemple d'effectuer des comparaisons indépendemment d'une éventuelle transposition.

Nous avons développé un système de recherche d'informations musicales à partir d'un extrait

2e soumission à *Document Numérique: Les Documents Musicaux*, le 23/01/09

basé sur des algorithmes d'alignement. Des évaluations à partir de bases de données réelles montrent que la précision de ce système est tout à fait satisfaisante : l'évaluation internationale effectuée dans le cadre du MIREX (Music Information Retrieval Evaluation eXchange) en 2006 et 2007 a mesuré sa précision au niveau des meilleurs algorithmes proposés.
Nous présentons enfin une application originale à la recherche automatique de plagiats. Les résultats d'expériences effectuées à partir de cas célèbres de plagiats montrent alors un des intérêts de la comparaison de mélodies.

ABSTRACT. *The number of digital musical documents available on the Internet are always increasing but browsing is only based on names (song, interpret). Users need new research methods dedicated to the specificity of the music, like for example research from a short excerpt (eventually approximated).*
These new browsing methods mainly rely on the estimation of the musical similarity between two musical pieces. However, the notion of similarity between musical data is very difficult to precisely define, and the estimation of the musical similarity is one of the main open problem in the musical information retrieval research area. From a computational point of view, it consists of determining algorithms that calculate a measure which indicates the degree of similarity between a pair of musical segments. Several approaches for computing such measures have already been proposed. But the major part of these methods is based on timbral similarities, mainly evaluated with statistics in low-level audio features. The consideration of other musical information such as melody has to been studied.
Firstly, we propose to investigate the different representations for melodies. We then discuss limitations of each one of these representations, depending on the applications. We propose to adapt algorithms applied in BioInformatics or string recognition to the musical context. These methods apply alignment or editing algorithms which determine the optimal number of operations necessary to transform one sequence into the other. A cost is associated to each elementary editing operation. Improvements are presented, allowing for example comparisons that take into account possible transpositions.
We have developed a music retrieval system from an short excerpt, based on alignment algorithms. Evaluations with real databases show that the accuracy of the system is good: international evaluation performed during the MIREX (Music Information Retrieval Evaluation eXchange) in 2006 and 2007 concludes that its precision is similar to the one of the best algorithms proposed.
We then present an original application for the automatic research of plagiarisms. The results of experiments performed with famous cases of plagiarisms show one of the interests of the comparison of melodies.

MOTS-CLÉS : *Similarité mélodique symbolique, alignement local, séquences musicales, comparaison de mélodies, informatique musicale, algorithmique du texte.*

KEYWORDS: *symbolic melodic similarity, local alignment, edition, musical sequences, melody comparison, computer music, string matching.*

1. Introduction

Le nombre de documents musicaux numériques disponibles en ligne (Internet, mobile, ...) ne cesse de croître. Les utilisateurs ne disposent actuellement que d'une navigation basée sur les noms comme par exemple les titres des morceaux ou le nom des interprètes. De nouvelles méthodes de navigation et de classification, plus intuitives et plus pratiques, doivent être proposées aux utilisateurs. Ces méthodes reposent essentiellement sur le problème majeur et encore ouvert de l'estimation automatique de la similarité musicale. Ce problème nécessite une expertise pluridisciplinaire dans des domaines variés tels que la théorie musicale, le traitement du signal audio, l'algorithmique de comparaison de structures de données, etc. Nous nous intéressons à ce difficile problème dans cet article, en nous restreignant à une composante particulière : la mélodie. Les travaux de recherche effectués dans le domaine de la recherche d'informations musicales (*Music Information Retrieval*) concerne généralement la musique occidentale. Pour ce type de musique, la mélodie est un des paramètres perceptifs les plus importants (Selfridge-Field, 1998).

Le problème de l'estimation de la similarité mélodique a notamment été mis en avant lors du développement d'applications musicales telles que la requête par chantonnement ou sifflement (*query-by-humming/singing*). Il est néanmoins difficile de définir avec précision la notion de similarité entre deux mélodies. D'un point de vue informatique, estimer la similarité entre deux morceaux consiste à calculer une mesure relative au degré de ressemblance entre ces deux morceaux. Pour certaines applications telles que la requête par fredonnement, il est impératif que cette estimation puisse prendre en compte certaines transformations musicales. Par exemple, comme un extrait peut être chanté dans une tonalité différente par un utilisateur, le système d'estimation doit être invariant par transposition. De même, un extrait peut être chanté plus ou moins rapidement, le système doit donc aussi être invariant par changement de tempo.

Plusieurs techniques d'évaluation de la similarité mélodique ont été introduites pendant les dernières années. Les algorithmes dits géométriques considèrent des représentations géométriques de la mélodie et calculent une distance entre ces objets. Certains de ces systèmes (Ukkonen *et al.*, 2003; Lubiw *et al.*, 2004) utilisent la représentation *piano-roll*, où les notes sont représentées par des segments de droites horizontales dont la longueur est relative à la durée, et dont les coordonnées correspondent au temps de début et à la hauteur de la note. D'autres approches géométriques représentent les notes par des masses (Typke *et al.*, 2004). Le poids est alors lié à la durée des notes.

Certains algorithmes d'estimation de la similarité mélodique sont des adaptations d'algorithmes classiques en traitement du texte (Doraisamy *et al.*, 2003; Uitdenbogerd, 2002). Les techniques *N-grams* proposent de compter le nombre de termes distincts que les deux morceaux comparés ont en commun. Si cette approche est très simple, elle s'avère être assez efficace (voir section 3). Toutefois, cette mesure de similarité ne prend pas en compte les propriétés perceptives de la musique puisque

seulement deux cas sont envisagés : soit deux sous-séquences correspondent, soit elles ont au moins une différence et ne correspondent pas. Néanmoins, le processus de perception de la musique semble bien plus complexe. En effet, deux morceaux peuvent être perçus comme très semblables tout en présentant de fortes différences dans la séquence de notes. Cette simplification induit donc la principale limitation de ce type de méthode.

Dans cet article, nous proposons une étude détaillée des algorithmes d'édition développés principalement dans le contexte de la reconnaissance de séquences ARN (Gusfield, 1997), ainsi que leur adaptation à la mesure de la similarité mélodique pour de la musique monophonique (Cambouropoulos *et al.*, 2005). La musique monophonique est définie par le fait qu'une seule note est perçue à chaque instant. Les algorithmes d'édition déterminent le score optimal d'opérations nécessaires à la transformation d'une séquence musicale en une autre séquence musicale. Nous présentons dans la section 2 l'algorithme général, ainsi que les adaptations nécessaires pour une application aux morceaux de musique. Ensuite, dans la section 3, des applications à la recherche de documents musicaux dans des bases de données à partir d'extraits et à la détection automatique de plagiats sont détaillées. Des expériences sont aussi proposées. Enfin, les conclusions ainsi que des perspectives sont présentées dans la section 4.

2. Formalisation du problème

Les algorithmes développés pour les systèmes de recherche musicale basée sur la similarité mélodique peuvent être décomposés en deux grandes étapes. La première étape représente une pièce de musique monophonique par une séquence de symboles. La deuxième étape permet le calcul d'un score de similarité entre deux représentations. Ces deux étapes sont détaillées dans cette section.

2.1. *Représentation de pièces de musique monophonique par des séquences*

Les pièces de musique monophonique peuvent être représentées par des arbres de hauteurs de notes (Rizo *et al.*, 2002). Cette représentation suppose une hiérarchie relative à la mesure induite par le chiffrage de la mesure d'une partition. Différents arbres peuvent alors représenter la même mélodie (même suite de notes). Par exemple, deux mêmes mélodies, composées des mêmes notes, avec deux chiffrages différents sont représentées par deux partitions différentes. Même si elles sont alors représentées de manière différente, elles seront perçues comme étant très fortement similaires.

En nous basant sur les travaux de Mongeau et Sankoff (Mongeau *et al.*, 1990), toute partition monophonique peut être représentée par une séquence de paires ordonnées avec la hauteur de la note en première composante, et la durée de la note (ici indiquée en nombre de doubles croches) en deuxième composante. Ainsi, la séquence

```
(B4 B4 r4 C4 G4 E2 A2 G8)
```

représente l'exemple illustré par la figure 1.

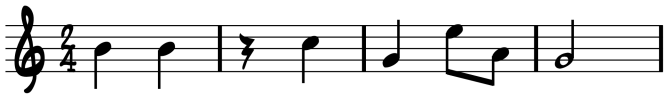

Figure 1. *Exemple de mélodie monophonique.*

Plusieurs alphabets de caractères et ensembles de nombres ont déjà été proposés pour représenter les hauteurs et durées (Uitdenbogerd, 2002; Lemström, 2000). Nous en présentons seulement quelques uns que nous pensons être les plus pertinents pour les applications de recherche automatique de musique envisagées.

Le contour mélodique indique la variation entre des notes successives. Seules trois valeurs sont possibles : Haut (plus aigu), Bas (plus grave) et Pareil. Ainsi, la séquence correspondant à la figure 1 est :

$$PHBHBB.$$

La hauteur absolue indique simplement la hauteur selon l'encodage *pitch* MIDI. Par exemple, la mélodie de la figure 1 est représentée par la suite de nombres :

$$71, 71, 72, 67, 76, 69, 67.$$

Si l'on souhaite réduire le nombre de caractères employés pour la représentation, la hauteur exacte peut être simplifiée par des valeurs modulo-12, sachant que 12 est le nombre de demi-tons composant une octave. Le contour mélodique peut aussi être pris en compte par l'ajout d'un signe ($+$ ou $-$) selon l'évolution des variations de hauteurs successives. La séquence de la figure 1 devient alors :

$$+11, +11, +0, -7, +4, -9, -7.$$

Pour des applications telles que la requête par fredonnement, cette représentation présente l'inconvénient majeur de ne pas être invariante par transposition.

À la différence de cette représentation, les représentations par intervalle ou relative à la tonalité sont invariantes par transposition. La représentation par intervalle indique simplement le nombre de demi-tons entre deux notes successives. La représentation de la séquence de la figure 1 est donc :

$$0, 1, 5, 9, 7, 2.$$

L'alphabet de cette représentation peut aussi être limité par une opération modulo-12. L'information relative à la direction mélodique peut également être indiquée :

$$0, +1, -5, +9, -7, -2.$$

Les représentations relatives à la tonalité indique la différence en demi-tons entre les notes et la tonique de la mélodie. Toujours dans le cas de la figure 1, comme la tonalité correspond à Do Majeur, la séquence associée devient :

$$11, 11, 0, 7, 4, 9, 7.$$

L'alphabet de cette représentation peut aussi être limitée par une opération modulo-12 et l'information de la direction mélodique peut également être indiquée :

$$11, 11, +0, -7, +4, -9, -7.$$

Les limitations de cette représentation sont évidemment fortement liées au choix judicieux de la tonalité. Cette tonalité doit être connue à priori pour pouvoir représenter correctement une mélodie.

Des représentations semblables sont envisageables concernant la durée des notes et des silences d'une mélodie. Le contour général (plus Court, plus Long, Pareil) indique la variation de durée entre deux notes consécutives. Ainsi, la représentation des durées des notes de la mélodie présentée dans la figure 1 est :

$$PPPPCPL.$$

La représentation absolue indique simplement la durée des notes en nombre de double croches :

$$4, 4, 4, 4, 4, 2, 2, 8.$$

Il est important de remarquer que cette représentation n'est pas invariante par changement de tempo, contrairement à la représentation relative. Le changement de durée entre deux notes successives peut être exprimé par des différences de durée :

$$0, 0, 0, 0, 2, 0, 6$$

ou par des rapports de durée :

$$1, 1, 1, 1, \frac{1}{2}, 1, 4.$$

Dans la suite, nous avons fait le choix de considérer les représentations de notes par intervalle, et les représentations de durée relatives. A partir de ces représentations, chaque élément (note) d'une séquence peut donc être formellement représenté par un symbole appartenant à un ensemble infini Σ d'étiquettes. Nous considérons une fonction de score d'édition s sur cet ensemble d'étiquettes. Cette fonction assigne un nombre réel $s(x, y)$ à chaque paire d'étiquettes (x, y) dans $\Sigma \cup \{\lambda\}$ où λ représente le symbole vide[1], de telle façon que :

$$\begin{aligned} s(x, x) &> 0 \qquad \forall x \in \Sigma, \\ s(x, y) &< 0 \qquad \forall x \neq y, (x, y) \in \{\Sigma \cup \{\lambda\}\}^2. \end{aligned}$$

Cela signifie que le score entre deux symboles x et y devient d'autant plus important que leur similarité croît.

1. $s(x, \lambda)$ est le score de suppression du symbole x dans Σ et $s(\lambda, y)$ est le score d'insertion de y.

Dans l'article (Hanna *et al.*, 2007), une étude des fonctions de coûts adaptés au domaine musical est proposé, s'appuyant notamment sur les premiers travaux décrits dans (Mongeau *et al.*, 1990). Pour les coûts d'insertion et de suppression, il est pertinent de prendre en compte la durée des notes, puisqu'insérer une longue note doit être plus pénalisant qu'insérer une courte note. Concernant le coût de substitution, les meilleurs résultats sont obtenus en prenant en compte la consonance entre deux notes. En effet, remplacer un Do par un Do# doit être plus pénalisé que le remplacement d'un Do par un Sol (quinte).

2.2. *Similarité locale entre séquences*

Le calcul d'une mesure de similarité entre séquences est un problème qui a souvent été abordé en informatique théorique au travers de nombreuses applications (Gusfield, 1997; Sankoff *et al.*, 1983) comme la biologie, l'analyse du texte, la détection d'erreur, la reconnaissance de forme, etc.

Au début des années 70, Needleman et Wunsh (Needleman *et al.*, 1970) ou Wagner et Fisher (Wagner *et al.*, 1974) ont proposé des algorithmes de calcul d'une mesure de similarité entre deux chaînes de symboles. Étant données deux chaînes de symboles S_1 et S_2 de longueurs respectives $|S_1|$ et $|S_2|$, un ensemble d'opérations élémentaires sur les chaînes, appelées opérations d'édition, et un score associé à chacune des opérations, le score entre S_1 et S_2 est défini formellement comme le score maximum d'une séquence d'opérations qui transforme S_1 en S_2. Cette mesure de similarité est calculée en utilisant le principe de programmation dynamique qui permet de résoudre le problème d'optimisation en un temps quadratique, *i.e.* en $O(|S_1| \times |S_2|)$.

Afin de comparer des séquences musicales, nous considérons uniquement trois opérations d'édition : la substitution, la suppression et l'insertion. Soit e une opération d'édition, un score s est associé à chacune des opérations comme suit :

– si e est une substitution, x_i (le $i$ème caractère de S_1) est substitué par y_j (le $j$ème caractère de S_2) et $s(e) = s(x_i, y_j)$

– si e est la suppression de x_i alors $s(e) = s(x_i, \lambda)$

– si e est l'insertion de y_j alors $s(e) = s(\lambda, y_j)$.

Le score s est étendu à la séquence d'opérations d'édition $E = (e_1, e_2, \ldots, e_n)$ en posant $s(E) = \sum_{k=1}^{n} s(e_k)$. Cette extension permet de définir le score $\sigma(S_1, S_2)$ entre deux séquences S_1 et S_2 comme le score maximum parmi les séquences d'opérations d'édition transformant S_1 en S_2 :

$$\sigma(S_1, S_2) = \max_{E \in \mathcal{E}} \{s(E)\}$$

où $\mathcal{E}$ représente l'ensemble des séquences d'opérations d'édition transformant S_1 en S_2.

Cependant, dans de nombreuses applications, deux chaînes peuvent ne pas être très similaires mais partager des *régions ayant une forte similarité*. Ceci est particulière-

ment remarquable lorsqu'on compare des séquences de longue taille n'ayant a priori pas de relation, mais dont certaines sections internes peuvent correspondre. Dans ce cas, le problème consiste à déterminer et à extraire une paire de régions de très forte similarité, chacune dans l'une des deux chaînes. Ce problème est appelé *problème de similarité locale* (Smith *et al.*, 1981) et il est formellement défini de la façon suivante : étant données deux chaînes S_1 et S_2, déterminer les deux sous-chaînes ρ_1 et ρ_2 respectivement de S_1 et S_2, dont la similarité parmi toutes les paires de sous-chaînes de S_1 et S_2 est maximum.

Le calcul d'une similarité locale permet de détecter les sous-régions conservées durant l'édition des deux séquences. La solution à ce problème est basée sur la comparaison des suffixes des deux séquences. Étant donnée une paire (x_i, y_j) de symboles, on recherche un suffixe (éventuellement vide) ρ_1 de la sous-séquence $S_1[x_i]$ (définie depuis le premier symbole de S_1 à x_i) et un suffixe (éventuellement vide) ρ_2 de la sous-séquence $S_2[y_j]$ (définie depuis le premier symbole de S_2 à y_j) tels que le score d'édition entre ρ_1 et ρ_2 est maximum parmi toutes les paires de suffixes de $S_1[x_i]$ et $S_2[y_j]$. Ce score optimal est noté $\mathrm{LS}(x_i, y_j)$:

$$\mathrm{LS}(x_i, y_j) = \max\{\sigma(\rho_1, \rho_2), (\rho_1, \rho_2) \text{ suffixes de } S_1 \text{ and } S_2\}.$$

La similarité locale entre deux séquences est alors définie comme le score maximum de la paire de suffixes entre les séquence S_1 et S_2 :

$$\mathrm{LS}(S_1, S_2) = \max\{\mathrm{LS}(x_i, y_j), (x_i, y_j) \in S_1 \times S_2\}.$$

Finalement, afin d'évaluer la similarité locale, l'algorithme doit déterminer les suffixes de similarité maximum pour toutes les paires (x_i, y_j) de $S_1 \times S_2$, puis déterminer la meilleure paire $x_1^{\max}$, $y_2^{\max}$ de S_1 et S_2.

On peut toujours choisir un suffixe vide $\mathrm{LS}(x_i, y_j) \geq 0$: en effet $\mathrm{LS}(x_i, \theta) = 0$ et $\mathrm{LS}(\theta, y_j) = 0$, où θ est une séquence vide. Enfin, pour toute paire (x_i, y_j), l'équation de récurrence permettant de calculer $\mathrm{LS}(x_i, y_j)$ est donnée par :

$$\mathrm{LS}(x_i, y_j) = \max \begin{cases} 0 \\ \mathrm{LS}(x_{i-1}, y_j) + s(x_i, \lambda) \\ \mathrm{LS}(x_i, y_{j-1}) + s(\lambda, y_j) \\ \mathrm{LS}(x_{i-1}, y_{j-1}) + s(x_i, y_j) \end{cases}$$

où x_{i-1} et y_{j-1} représentent respectivement les symboles avant x_i et y_j.

2.3. *Améliorations proposées*

À partir du système présenté, basé sur l'alignement de symboles représentant les mélodies, plusieurs améliorations ont été récemment proposées. Un algorithme a ainsi

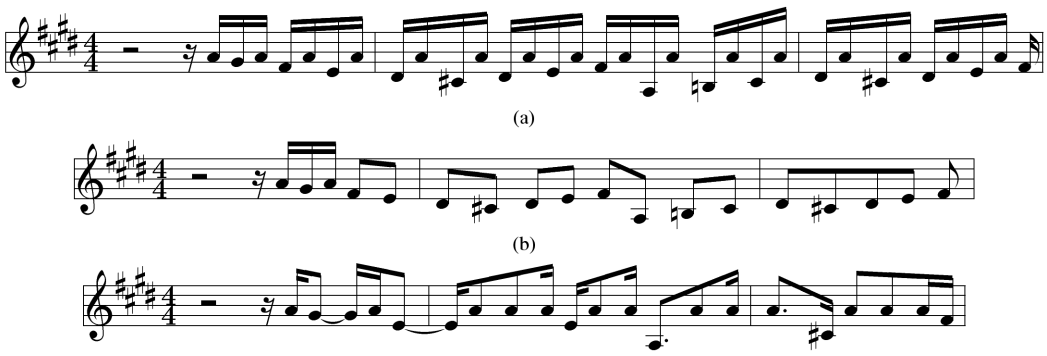

Figure 2. *(a) Morceau* Toccata et Fugue en Ré mineur *(Bach) (b) Une variation où de nombreuses notes ont été supprimées, sauf les notes sur les temps. (c) Au contraire, cette variation a été définie en supprimant les notes sur les temps. Le nombre de notes est à peu près le même dans les variations (b) et (c), alors que la perception est très différente : seulement la variation (b) est perçue comme étant similaire à la version originale.*

été proposé pour prendre en compte d'éventuelles transpositions locales (ou modulations) au cours d'une même mélodie (Allali *et al.*, 2007). Pour cela, plusieurs matrices de programmation dynamique (1 matrice par transposition possible, soit 12 au total) sont calculées en parallèle. Une nouvelle opération d'édition est prise en compte autorisant, avec un certain coût fixé, les passages entre différentes matrices de transposition. La principale application concernée est la requête par fredonnement. Il arrive en effet fréquemment qu'un utilisateur commence à chanter un extrait dans une certaine tonalité, et termine dans une autre. L'algorithme proposé permet de retrouver des similarités mélodiques malgré ces éventuelles transpositions.

Par ailleurs, d'autres améliorations algorithmiques basées sur la théorie musicale ont été proposées dans (Robine *et al.*, 2007a). Des éléments musicaux pris en compte par les auditeurs lors de la discrimination de morceaux de musique ont été étudiés. Par exemple, les notes placées sur les temps forts d'une mesure semblent devoir être considérées comme plus importantes que les notes placées sur des temps faibles. La figure 2 donne un exemple d'un morceau et de deux variations obtenues en supprimant des notes. Si les notes supprimées sont sur les temps, la différence devient très importante, alors que si les notes supprimées ne sont pas sur les temps, la similarité reste forte. Le système proposé permet donc de prendre en compte cette propriété en ne pénalisant pas les suppressions ou insertions de notes placées sur des temps faibles. En suivant cette nouvelle règle, les morceaux (a) et (b) sont bien estimés comme très similaires.

D'autres exemples de prise en compte d'informations musicales sont illustrés par la figure 3, présentant un extrait ainsi que des variations du thème du morceau *D*ay Tripper des *B*eatles. Nous proposons de prendre également en compte les notes tonales, plus importantes sur le plan de la perception que les autres notes. Ainsi, même

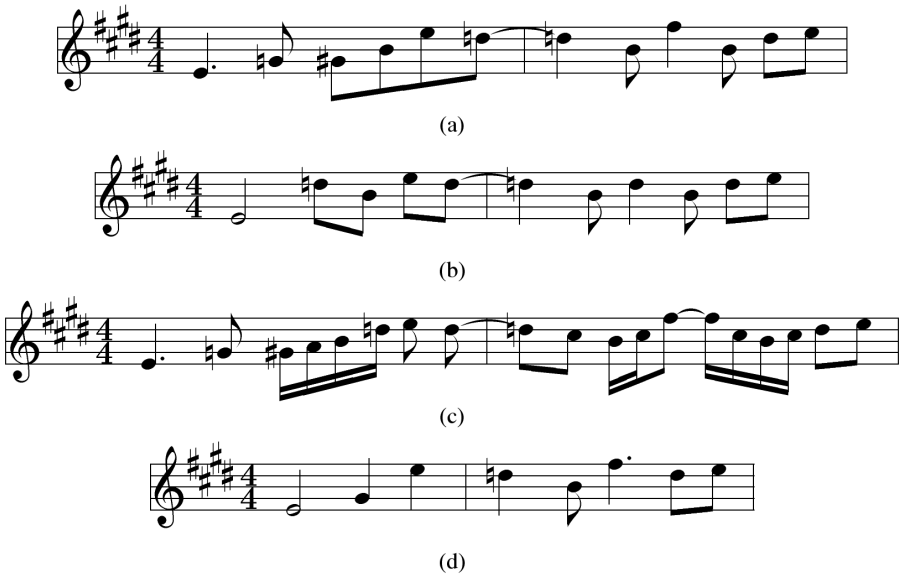

Figure 3. *(a) Extrait de* D*ay Tripper (The Beatles) (b) Variation avec conservation de la tonique Mi et de la dominante Si de la tonalité Mi mineur (c) Variation avec insertion de notes de passage entre les notes de la mélodie (d) Variation avec conservation des notes placées sur les temps.*

si plusieurs notes ont été supprimées dans l'extrait (b) de la figure 3, le morceau reste très proche de l'original car les notes relatives à la tonique et à la dominante ont été préservées. De la même façon, les variations (c) et (d) restent également très similaires à l'extrait original, puisque seules des notes de passage ont été insérées (variation (c)) ou puisque les notes correspondant aux temps forts ont été préservées (variation (d)).

3. Applications et expériences

Dans cette partie, nous présentons deux applications importantes reposant sur les algorithmes d'estimation automatique de la similarité mélodique présentés ci-dessus. La première application concerne la recherche de mélodies dans une base de données musicale à partir d'un extrait. La seconde application propose de rechercher dans une base de données d'éventuels plagiats. Pour chacune de ces applications, nous présentons quelques expériences d'évaluation des algorithmes proposés sur des bases de données musicales.

3.1. *Recherche à partir d'un extrait*

Les méthodes de recherche de musique existantes reposent essentiellement sur une recherche par le nom (de l'interprète, du titre). Il est toutefois courant qu'un utilisateur souhaite retrouver un morceau de musique sans posséder ces informations. Un utilisateur peut par exemple souhaiter retrouver un morceau de musique dont il vient d'écouter un extrait à la télévision ou à la radio. La seule donnée dont il dispose est une mélodie qu'il est capable de fredonner ou de siffler, parfois approximativement. Dans ce cas, la ressemblance entre le thème fredonné et le morceau recherché est uniquement basée sur la mélodie. Il en découle une application des algorithmes d'estimation automatique de la similarité mélodique : la recherche de morceaux à partir d'un extrait chantonné ou fredonné. Ces systèmes de recherche, appelés systèmes *query-by-humming*, sont en plein développement. Ainsi de nombreuses interfaces, plus ou moins précises, apparaissent régulièrement sur Internet. On peut citer par exemple `watzatsong`[2], `musipedia`[3], `c-Brahms`[4], etc.

L'application proposée offre donc la possibilité à un utilisateur non seulement de retrouver le morceau recherché, mais aussi de retrouver des morceaux dont les mélodies, ou les thèmes, sont ressemblants. Pour cela, le principe est relativement simple. Le morceau fredonné ou sifflé par l'utilisateur est considéré comme une requête, c'est-à-dire le morceau référence pour la recherche. Le système informatique proposé doit alors calculer les scores de similarité entre cette requête et tous les morceaux de la base de donnée musicale. Les scores les plus importants doivent correspondre aux morceaux de la base de données qui ressemblent le plus à la requête, notamment le morceau de score le plus important devrait être le morceau recherché par l'utilisateur.

Un des problèmes principaux du domaine de recherche lié à l'informatique musicale est le problème de l'évaluation des systèmes proposés. Le premier concours international *Music Information Retrieval Evaluation eXchange* (MIREX 2005) (Downie *et al.*, 2005) a été proposé dans le but de comparer les algorithmes existants dans des domaines variés. Un des concours proposés durant la première édition du MIREX concernait l'estimation de la similarité mélodique à partir de données symboliques. A cette occasion, les participants ont discuté le processus d'évaluation et ont proposé une méthode. Les résultats que nous allons présenter reposent sur cette méthode, que nous détaillons ici.

Tout d'abord, il est nécessaire de considérer une importante base de données de morceaux de musique. La collection RISM A/II du répertoire international des sources musicales est composé d'un demi-million de partitions de compositions existantes. Chaque début de morceau est représenté symboliquement, dans une notation proche d'une partition de musique. Chaque extrait est monophonique et composé d'entre 10 et 40 notes de musique. 11 extraits ont été aléatoirement choisis dans cette collection. Une base de vérité de similarité a été établie (Typke *et al.*, 2005) en combinant

2. `www.watzatsong.com`
3. `www.musipedia.org`
4. `www.cs.helsinki.fi/group/cbrahms/demoengine/`

les listes ordonnées données par 35 experts en musique. La base de vérité résultante prend donc la forme de groupes ordonnés d'extraits. Chaque groupe contient des extraits dont les différences entre eux n'ont pas été jugées comme statistiquement significatives. Par contre, l'ordre des groupes est statistiquement significatif.

Chaque système testé retourne une liste ordonnée d'extraits estimés comme similaires sur le plan mélodique à la requête proposée. Quelques mesures pour l'évaluation sont alors utilisées pour produire un unique score relatif à la correspondance avec la base de vérité établie par les experts musiciens. Une mesure spécifique a ainsi été introduite : le rappel dynamique moyen (*Average Dynamic Recall* ou ADR) (Typke *et al.*, 2006a). Cette mesure prend en compte les groupes ordonnés de la base de vérité en notant le nombre de documents de ces groupes qui apparaissent avant ou à une position donnée de la liste ordonnée, sans tenir compte de l'ordre à l'intérieur des groupes (puisqu'il n'est pas pertinent). L'ADR prend une valeur dans l'intervalle $[0; 1]$. Plus l'ADR est importante, plus le système est proche de la base de vérité établie par les experts musiciens, et plus le système est supposé précis.

Puisque notre algorithme n'a pas pu participer au MIREX 2005, les premières expériences que nous avons menées concernent les données d'entraînement du MIREX 2005. Ces données sont composées de 11 morceaux requêtes pour une base de données de 580 extraits. Dans le tableau 1, nous présentons nos résultats et ceux obtenus par les participants au concours. Il est important de rappeler ici que cette comparaison n'est pas suffisante en elle-même, puisque l'algorithme proposé n'a pas participé au concours. Les deux bases de données (entraînement et test) sont toutefois issues de la même base de données, et donc les résultats obtenus sur la base d'entraînement peuvent être comparés à ceux obtenus sur la base de test. Les algorithmes proposés par Grachten et Lemström sont basés sur la distance d'édition, et considèrent donc une approche très similaire à la méthode présentée dans cet article.

Les résultats obtenus sur les données d'entraînement du MIREX 2005 sont très bons : l'ADR moyen obtenu par l'algorithme d'alignement présenté est 0.79 alors que le meilleur ADR moyen obtenu par un des candidats du MIREX 2005 n'est que de 0.66. De plus, l'ADR minimum reste important (0.63) alors que l'ADR maximum (0.96) s'approche quasi parfaitement du choix des experts musiciens.

Lors de la deuxième édition du MIREX, un autre concours sur la similarité mélodique a été proposé[5]. Une partie du concours consistait à comparer les algorithmes de recherche. La base de données, sous-ensemble de la collection RISM A/II était plus conséquente puisqu'elle était composée de plus de 15000 extraits monophoniques. La requête était un de ces extraits. La moitié des requêtes a été générée à partir d'extraits fredonnés ou sifflés, convertis en format symbolique. Ces requêtes étaient donc caractérisées par des imperfections rythmiques (légères accélérations ou légers ralentissements) et de hauteur (présence de fausses notes). Certaines requêtes étaient également transposées, d'autres avaient un tempo différent du morceau original. Contrairement à ce qui avait été effectué pour le MIREX 2005, aucune base de vérité n'avait été éta-

5. http://www.music-ir.org/mirex/2006/index.php/Symbolic_Melodic_Similarity_Results

Algorithme	Auteur	ADR moyen	ADR min	ADR max
Distance d'édition I/R	Grachten	0.66	0.29	0.88
N-grams	Orio	0.65	0.31	0.91
N-grams de base	Uitdenbogerd	0.64	0.31	0.91
Géométrique	Typke	0.57	0.29	0.86
Géométrique	Lemström	0.56	0.34	0.72
Distance d'édition	Lemström	0.54	0.29	0.84
Hybride	Frieler	0.52	0.34	0.81

Tableau 1. *Résultats de l'évaluation des systèmes de recherche lors du MIREX 2005. En comparaison, l'algorithme proposé obtient de meilleurs résultats (ADR moyen de* 0.79*, ADR min de* 0.66 *et ADR max de* 0.96*).*

blie à l'avance. Les résultats des différents algorithmes ont été jugés a posteriori par des experts, selon une évaluation générale (ressemblance forte, moyenne ou nulle) et un score de similarité (entre 0 et 10). La précision de chaque système a ensuite été estimée selon plusieurs mesures, dont l'ADR moyen.

Les résultats sont présentés dans la figure 4. Notre algorithme est noté FH (selon les initiales de certains de ses auteurs) et a obtenu des résultats similaires aux résultats obtenus par l'algorithme évalué comme le plus précis soumis par Rainer Typke (Typke *et al.*, 2006b).

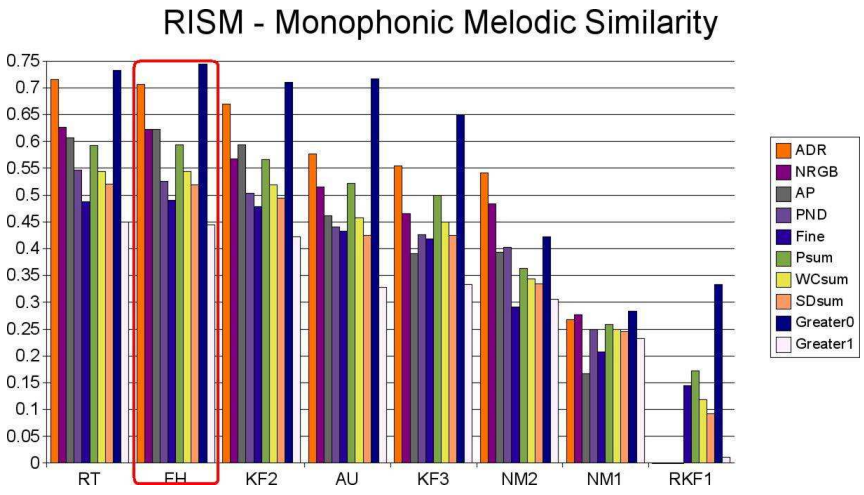

Figure 4. *Résultats du concours d'évaluation des systèmes d'estimation de la similarité mélodique lors du MIREX 2006 : FH représente les résultats obtenus par l'algorithme présenté dans l'article.*

Concernant la polyphonie, les résultats sont encore mitigés. Les résultats obtenus lors du MIREX 2006 prenaient en compte un système qui a subi depuis de nombreuses améliorations. Les résultats sont maintenant comparables aux résultats obtenus par le meilleur algorithme (Hanna *et al.*, 2008). Toutefois, la complexité de ces algorithmes devient très importante et semble limiter leur utilisation sur des bases de données importantes. C'est pourquoi de nouvelles méthodes basées sur l'estimation des informations tonales telles que les accords, qui permettraient donc une réduction monophonique, sont actuellement à l'étude.

Ces expériences réalisées dans le cadre de concours internationaux montrent assez clairement l'intérêt des méthodes d'alignement dans le cadre d'une application musicale. La flexibilité de ces méthodes permet d'effectuer des réglages spécifiques à l'information musicale (Ferraro *et al.*, 2007), et de prendre en compte d'éventuelles imperfections rythmiques ou tonales. Le système de recherche induit permet non seulement de retrouver un morceau à partir d'un extrait, mais aussi de trouver des morceaux mélodiquement ressemblants. D'autres applications proches de ces systèmes de recherche sont alors envisageables. Par exemple, estimer la similarité entre une interprétation et une partition pourrait aider un élève instrumentiste à s'entraîner et à corriger ses éventuelles imperfections. Ces systèmes peuvent aussi permettre de découvrir des ressemblances fortes entre des morceaux dans une base de données, sans avoir forcément besoin d'analyser musicalement *à la main* tous les morceaux deux à deux.

3.2. *Détection automatique de plagiats*

Le nombre de documents musicaux disponibles sur Internet est en forte croissance. Chaque année, plus de 10000 nouveaux albums de musique sont proposés à la vente dans le monde, et plus de 100000 nouveaux morceaux sont déposés à des organismes de protection des droits d'auteur (Uitdenbogerd *et al.*, 1999). Par exemple, en France, le nombre total de morceaux de musique enregistrés auprès de la SACEM (Société des Auteurs, Compositeurs et Éditeurs de Musique), organisme chargé de protéger les droits des artistes, dépassait 250000 en 2004 (Emberger, 2005). Un des rôles de cet organisme est d'aider la justice à prendre des décisions concernant des plaintes pour plagiats. Le plagiat est un acte délibéré de copie partielle ou totale d'une idée d'un autre auteur sans accord préalable. Il est d'ailleurs important de noter que seule la justice est habilitée à décider d'un cas de plagiat. Quelques exemples de procès célèbres pour plagiat musical ont eu lieu au cours des dernières années : Madonna et Salvatore Acquaviva en Belgique, Georges Harrison et le groupe The Chiffons au Royaume-Uni, *Les feuilles mortes* et *La Maritza* en France, etc. En 2004, la SACEM avait vérifié l'originalité de seulement 18000 morceaux sur son catalogue total de 250000 titres. Ce n'est uniquement qu'en cas de plainte déposée qu'une analyse musicale complète est effectuée par des experts. Étant donné le nombre important de nouveaux documents musicaux déposés chaque année, il apparaît difficile de détecter pour chaque morceau un hypothétique cas de plagiat, car il semble impossible d'écouter et de com-

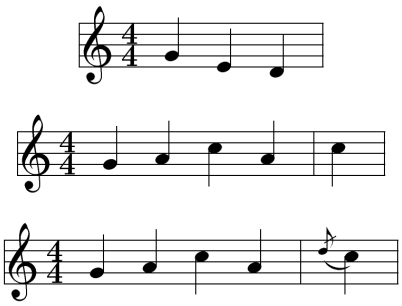

Figure 5. *Courts motifs musicaux composant la structure des chansons* My Sweet Lord *(G. Harrison) et* He's So Fine *(R. Mack) : motif A (haut), motif B (milieu) et motif C (bas).*

parer un à un tous les documents déposés. Un système informatique d'estimation de la similarité mélodique peut permettre d'effectuer cette comparaison de manière automatique. Nous avons effectué une première étude (Robine *et al.*, 2007b) à partir du système présenté dans les sections précédentes. Nous montrons ici quelques résultats pour quelques cas célèbres de plagiats.

Un des plus fameux procès pour un cas de plagiat musical concerne l'ex-Beatles Georges Harrison pour sa chanson *My Sweet Lord*, sortie en 1970 sur l'album *All Things Must Pass* (Benedict, 2008). Il a été accusé de plagiat du morceau *He's So Fine* composé en 1963 par Ronald Mack et interprété par le groupe The Chiffons. Malgré les explications d'Harrison sur le fait qu'il n'a pas eu conscience de s'être approprié la mélodie du morceau, la cour conclut en 1976 qu'il avait copié, peut-être de manière inconsciente, la mélodie de *He's So Fine*. Pour appuyer sa décision, la cour examina la structure des deux morceaux. La figure 6 montre deux extraits de la mélodie de chaque chanson. *He's So Fine* est composé de quatre variations d'un court motif musical (le motif A de la figure 5), suivies de quatre variations du motif B (figure 5). La deuxième utilisation de la série de motifs B inclut une appoggiature, comme le montre le motif C de la figure 5. *My Sweet Lord* possède une structure très similaire puisqu'il est composé de quatre variations du motif A, puis de trois variations du motif B. Une quatrième variation du motif B inclut l'appoggiature du motif C.

Les premières expériences que nous proposons concernent ces deux chansons dont la figure 6 montrent deux extraits. Nous pouvons remarquer que, même si les mélodies sont perçues comme très similaires, les extraits sont assez différents. Ces différences rendent évidemment l'analyse automatique plus difficile. Les deux requêtes pour le système de recherche dans une base de données musicales seront les extraits des deux mélodies de chaque morceau. Nous attendons que le système détecte le morceau correspondant comme le morceau le plus mélodiquement ressemblant, malgré les différences. La base de données considérée est la collection RISM A/II utilisée durant le

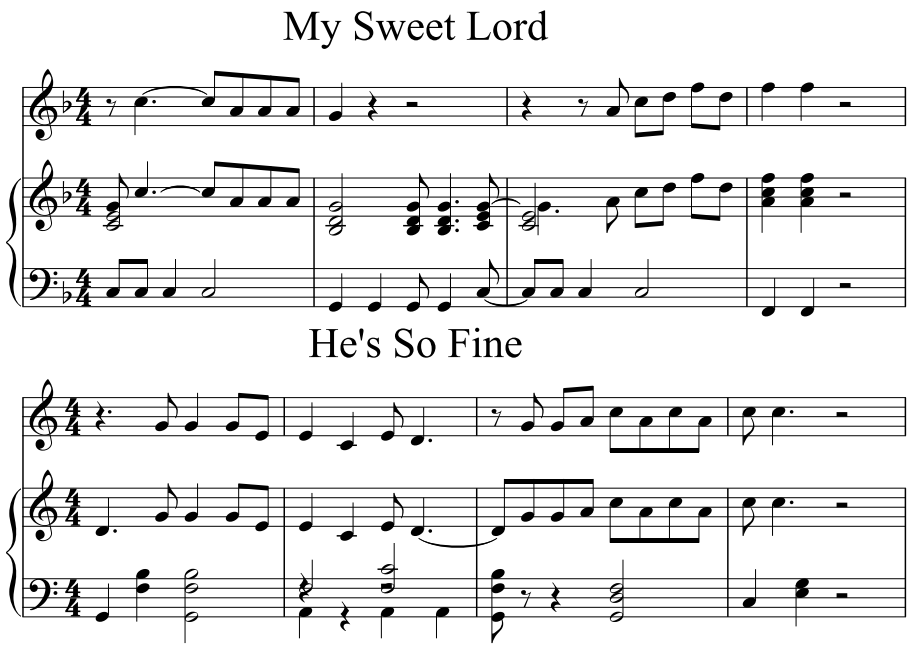

Figure 6. *Transcription manuelle d'extraits (correspondant aux motifs A et B de la figure 5) des chansons* My Sweet Lord *(G. Harrison) et* He's So Fine *(R. Mack).*

MIREX 2006, présentée dans la section précédente. Cette base de données contient aussi bien entendu les mélodies monophoniques de *My Sweet Lord* et *He's So Fine*.

Le tableau 2 montre les résultats obtenus pour chacune des requêtes. Dans les deux cas, les noms, ainsi que les scores de similarité, des pièces de la base de données sont donnés. Les résultats obtenus sont ceux escomptés puisque la mélodie détectée comme la plus similaire est bien la mélodie du morceau correspondant, ce qui montre que le système est bien entendu capable de retrouver un morceau à partir d'un extrait. Plus intéressant, la seconde pièce estimée comme la plus similaire est bien la mélodie de *He's So Fine* pour la requête *My Sweet Lord*, et inversement. Malgré les différences entre les deux mélodies, le système s'avère donc intéressant pour rechercher de manière automatique les hypothétiques cas de plagiat. Il est aussi important de remarquer les différences significatives entre les deuxième et troisième scores (83 par rapport à 52 ou 45). La similarité mélodique entre la requête et le morceau le plus similaire de la base de données, exception faite du plagiat, est indiquée par le troisième score et reste peu importante comparée à la similarité entre les morceaux *He's So Fine* et *My Sweet Lord*.

Pour confirmer les résultats de ces premières expériences, nous avons considéré une autre base de données, composée de morceaux monophoniques de durée plus im-

Requête	**rang** 1 **score** 1	**rang** 2 **score** 2	**rang** 3 **score** 3
Sweet Lord mélodie	Sweet Lord 178.9	So Fine 83.0	X 52.2
So Fine mélodie	So Fine 199.7	Sweet Lord 83.0	X 45.5

Tableau 2. *Résultat d'expériences sur la détection automatique de plagiats musicaux : exemples des morceaux* My Sweet Lord *et* He's So Fine *(X indique un morceau de la base de données peu similaire à la requête).*

portante. Cette base de données regroupe plus de 1650 fichiers MIDI monophoniques collectés sur Internet. Quatre autres cas de plagiats célèbres sont également considérés, décrits sur le site (Columbia Center for New Media Teaching and Learning, 2008). Ainsi, pour chacun des cinq cas de plagiats, la mélodie monophonique est considérée comme une requête, et le système de recherche calcule les scores de similarité avec tous les morceaux de la base de données (qui contient également les mélodies des plagiats). Le tableau 3 montre les résultats obtenus par le système présenté, avec notamment les morceaux retrouvés aux trois premières positions, ainsi que leur score de similarité. Comme attendu, le morceau estimé comme étant le plus similaire est bien la requête elle-même. Le score du morceau en première position correspond donc au score maximum. L'information la plus intéressante pour notre évaluation concerne le morceau en deuxième position. Idéalement, ce morceau devrait correspondre à la mélodie déterminée par la justice comme un plagiat. Le tableau 3 montre que c'est toujours le cas, à l'exception du cas *Fantasy vs Fogerty*. Cette erreur montre les limites du système actuel. Ces limites sont discutées dans la section 4. Pour tous les autres cas, le système de recherche obtient les résultats escomptés. Pour des cas tels que les plagiats *Selle vs Gibb* ou *Heim vs Universal* par exemple, la similarité est même estimée comme importante. Néanmoins, les limites apparaissent également puisque l'on observe les petites différences de scores pour les morceaux en deuxième et troisième position pour le cas *Repp vs Webber*. Le faible score du morceau en deuxième position, considéré comme plagiat, implique une faible différence entre ce score et le score de similarité obtenu avec les autres morceaux de la base de données.

Ces expériences sur la détection de cas de plagiats sont encore des études préliminaires, mais montrent l'intérêt d'adapter les méthodes issues de l'alignement de texte au contexte musical. Les limites actuelles restent liées à la gestion de la polyphonie. L'analyse musicale permettant de mener à une réduction monophonique en prenant en compte des accords semblent être une piste prometteuse.

Requête	**rang** 1 **score** 1	**rang** 2 **score** 2	**rang** 3 **score** 3
R. Mack vs G. Harrison (1976)			
Sweet Lord	Sweet Lord 178.9	So Fine 83.0	X 77.5
So Fine	So Fine 199.7	Sweet Lord 83.0	X 75.3
Fantasy vs Fogerty (1994)			
Road	Road 168.9	X 87.6	Jungle 75.9
Jungle	Jungle 146.3	Road 75.9	X 75.5
Heim vs Universal (1946)			
Vagyok	Vagyok 248.6	Perhaps 123.5	X 92.8
Perhaps	Perhaps 215.5	Vagyok 123.5	X 76.8
Repp vs Webber (1997)			
Till You	Till You 135.5	Phantom 50.8	X 50.4
Phantom	Phantom 145.8	Till You 50.8	X 49.7
Selle vs Gibb (1984)			
Let It End	Let It End 192.4	How Deep 118.1	X 68.9
How Deep	How Deep 202.8	Let It End 118.1	X 83.8

Tableau 3. *Résultats d'expériences sur la détection de plagiat sur quelques cas célèbres.*

4. Conclusions et Perspectives

Nous avons présenté un système d'estimation automatique de la similarité automatique entre deux mélodies, basé sur une adaptation d'algorithmes couramment appliqués en BioInformatique et en traitement du texte. Deux applications ont été proposées et testées, l'une permettant la recherche de mélodies à partir d'un extrait, l'autre proposant d'aider à la recherche automatique de plagiats. Les expériences menées indiquent que le système développé semble très prometteur.

Quelques améliorations et extensions doivent toutefois être proposées dans un futur proche. Tout d'abord, nous décrivons une représentation et des algorithmes dédiés

à la musique monophonique. Il semble essentiel de pouvoir considérer un contexte polyphonique. De nombreuses adaptations sont nécessaires. Certaines ont déjà été proposées et testées. Mais la complexité des méthodes devient beaucoup plus importante et semble limiter leur application sur des bases de données réelles (de taille importante). Il semble donc nécessaire de prendre en compte des informations musicales de plus haut niveau telles que les séquences d'accords, les modulations à l'intérieur d'un morceau ou la tonalité. Ces adaptations risquent de soulever de nouveaux problèmes algorithmiques telles que la comparaison d'accords par exemple. Des travaux de recherche sont en cours à ce sujet.

Il semble aussi intéressant de proposer des méthodes d'indexation, permettant d'éviter de comparer entièrement tous les morceaux d'une base de données. Des méthodes issues de la BioInformatique peuvent certainement être adaptées dans le futur.

De plus, nous avons restreint notre système de comparaison de musique aux propriétés mélodiques. Il semble évident que la similarité musicale peut reposer sur de nombreux autres facteurs tels que l'harmonie, le rythme, la structure, le timbre, etc. Là encore, de nouveaux problèmes apparaissent, notamment liés à la représentation informatique de ces propriétés musicales, et des éventuels liens entre elles.

5. Bibliographie

Allali J., Hanna P., Ferraro P., Iliopoulos C., « Local Transpositions in Alignment of Polyphonic Musical Sequences », *Proceedings of the 14th String Processing and Information Retrieval Symposium (SPIRE)*, October, 2007.

Benedict, *Copyright Website*. 2008, `http://www.benedict.com/Audio/Harrison/Harrison.aspx`.

Cambouropoulos E., Crochemore M., Iliopoulos C. S., Mohamed M., Sagot M.-F., « A Pattern Extraction Algorithm for Abstract Melodic Representations that Allow Partial Overlapping of Intervallic Categories », *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR'05)*, London, UK, p. 167-174, 2005.

Columbia Center for New Media Teaching and Learning, *Music Plagiarism Project Website*. 2008, `http://cip.law.ucla.edu/`.

Doraisamy S., Rüger S., « Robust Polyphonic Music Retrieval with N-grams », *Journal of Intelligent Information Systems*, vol. 21, n° 1, p. 53-70, 2003.

Downie J. S., West K., Ehmann A. F., Vincent E., « The 2005 Music Information retrieval Evaluation Exchange (MIREX'05) : Preliminary Overview », *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR'05)*, London, UK, p. 320-323, 2005.

Emberger P., « Dossier SACEM : Le Livre Blanc », *Keyboards Magazine*, vol. 200, Sep, 2005.

Ferraro P., Hanna P., « Optimizations of Local Edition for Evaluating Similarity Between Monophonic Musical Sequences », *RIAO 2007 : Proceedings of the 8th International Conference on Information Retrieval - Large-Scale Semantic Access to Content (Text, Image, Video and Sound), Pittsburgh, PA, USA*, 2007.

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

Hanna P., Ferraro P., Robine M., « On Optimizing the Editing Algorithms for Evaluating Similarity Between Monophonic Musical Sequences », *Journal of New Music Research*, vol. 36, n° 4, p. 267-279, 2007.

Hanna P., Robine M., Ferraro P., Allali J., « Improvements of Alignment Algorithms for Polyphonic Music Retrieval », *Proceedings of the International Computer Music Modeling and Retrieval Conference (CMMR)*, Copenhagen, Denmark, p. 244-251, 2008.

Lemström K., String Matching Techniques for Music Retrieval, PhD thesis, University of Helsinki, Dept. of Computer Science, 2000.

Lubiw A., Tanur L., « Pattern Matching in Polyphonic Music as a Weighted Geometric Translation Problem », *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR)*, Barcelona, Spain, p. 289-296, October, 2004.

Mongeau M., Sankoff D., « Comparison of Musical Sequences », *Computers and the Humanities*, vol. 24, n° 3, p. 161-175, 1990.

Needleman S., Wunsch C., « A General Method Applicable to the Search for Similarities in the Amino Acid Sequences of Two Proteins », *Journal of Molecular Biology*, vol. 48, p. 443-453, 1970.

Rizo D., Iñesta-Quereda J., « Tree-Structured Representation of Melodies for Comparison and Retrieval », *Proceedings of the 2nd International Workshop on Pattern Recognition in Information Systems (PRIS'02)*, Alicante, Spain, p. 140-155, 2002.

Robine M., Hanna P., Ferraro P., « Music Similarity : Improvements of Edit-based Algorithms by Considering Music Theory », *Proceedings of the ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR)*, Augsburg, Germany, p. 135-141, 2007a.

Robine M., Hanna P., Ferraro P., Allali J., « Adaptation of String Matching Algorithms for Identification of Near-Duplicate Music Documents », *Proceedings of the International SIGIR Workshop on Plagiarism Analysis, Authorship Identification, and Near-Duplicate Detection (PAN)*, Amsterdam, Netherlands, p. 37-43, 2007b.

Sankoff D., Kruskal J. B. (eds), *Time Wraps, Strings Edits, and Macromolecules : the Theory and Practice of Sequence Comparison*, Addison-Wesley Publishing Company Inc, University of Montreal, Canada, 1983.

Selfridge-Field E., « Conceptual and Representational Issues in Melodic Comparison », *Melodic Comparison : Concepts, Procedures, and Applications, Computing in Musicology*, vol. 11, p. 3-64, 1998.

Smith T., Waterman M., « Identification of Common Molecular Subsequences », *Journal of Molecular Biology*, vol. 147, p. 195-197, 1981.

Typke R., den Hoed M., de Nooijer J., Wiering F., Veltkamp R. C., « A Ground Truth For Half A Million Musical Incipits », *Journal of Digital Information Management*, vol. 3, n° 1, p. 34-39, 2005.

Typke R., Veltkamp R. C., Wiering F., « Searching Notated Polyphonic Music Using Transportation Distances », *Proceedings of the 12th ACM Multimedia Conference (MM)*, New-York, USA, p. 128-135, 2004.

Typke R., Veltkamp R. C., Wiering F., « A Measure for Evaluating Retrieval Techniques Based on Partially Ordered Ground Truth Lists », *Proceedings of the 7th International Conference on Multimedia and Expo (ICME'06)*, Toronto, Canada, p. 128-135, July, 2006a.

Typke R., Veltkamp R. C., Wiering F., « MIREX Symbolic Melodic Similarity and Query by Singing/Humming », *2nd Music Information Retrieval Evaluation eXchange (MIREX'06)*, Victoria, Canada, 2006b.

Uitdenbogerd A. L., Music Information Retrieval Technology, PhD thesis, RMIT University, Melbourne, Australia, July, 2002.

Uitdenbogerd A. L., Zobel J., « Matching Techniques for Large Music Database », *Proceedings of the 7th ACM International Conference on Multimedia (MM'99)*, Orlando, USA, p. 56-66, 1999.

Ukkonen E., Lemström K., Mäkinen V., « Geometric Algorithms for Transposition Invariant Content-Based Music Retrieval », *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR)*, Baltimore, USA, p. 193-199, October, 2003.

Wagner R. A., Fisher M. J., « The String-to-String Correction Problem », *Journal of the association for computing machinery*, vol. 21, p. 168-173, 1974.