



HAL
open science

Description structurelle automatique de documents musicaux

Olivier Lartillot

► **To cite this version:**

Olivier Lartillot. Description structurelle automatique de documents musicaux. IC - 17èmes Journées francophones d'Ingénierie des Connaissances, Jun 2006, Nantes, France. pp.81-90. hal-01026244

HAL Id: hal-01026244

<https://inria.hal.science/hal-01026244>

Submitted on 21 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Description structurelle automatique de documents musicaux.

Olivier Lartillot

Music Cognition Group, Université de Jyväskylä, PL 35(A), 40100, Finlande, lartillo@campus.jyu.fi

Résumé

Une nouvelle méthodologie d'extraction de motifs répétés au sein de partitions musicales est présentée. L'automatisation de la recherche de répétitions permet un approfondissement de l'analyse musicale, et rend possible la conception d'outils de gestion des masses de données musicales. L'approche proposée consiste en une explicitation des facteurs de redondance structurelle sous-jacents à une telle problématique. Cette redondance est maîtrisée par l'intermédiaire d'heuristiques de contrôle assurant une meilleure concordance du comportement algorithmique avec les attentes musicologiques et les intuitions perceptives. Ces heuristiques se basent pour une partie sur des principes mathématiques ou informatiques, et pour une autre sur des règles de *Gestalt*. En particulier, un filtrage des motifs non-fermés et une factorisation des motifs cycliques sont assurés dans un espace paramétrique multi-dimensionnel. L'algorithme proposé offre une description à la fois compacte et détaillée des documents musicaux, présentant une pertinence musicale et perceptible forte, offrant ainsi une avancée significative par rapport aux recherches antérieures. De nombreux mécanismes doivent cependant être pris en compte afin d'intégrer l'ensemble des transformations possibles et d'envisager l'expression musicale dans toute sa généralité polyphonique.

Mots clés : Analyse de corpus musical, fouille de données multi-dimensionnelles, masses de données musicales, extraction de motifs répétés, motifs fermés, motifs cycliques.

1 Introduction

Une nouvelle méthodologie d'extraction de motifs répétés au sein de séquences symboliques temporelles est proposée. La problématique générale d'extraction de motifs étudiée en intelligence artificielle [24, 37] est envisagée ici dans le contexte musical. La recherche de motifs musicaux permet de mettre en évidence, au sein de documents musicaux, un réseau de structures généralement perçues de manière plus ou moins explicite. L'analyse musicale, en offrant ainsi une explicitation graphique des structures, peut guider et enrichir l'écoute musicale. Une automatisation de cette recherche de répétitions permet l'émergence de nouveaux outils d'analyse et de visualisation des struc-

tures musicales, et de gestion et de navigation au sein de larges bases de données musicales.

Les documents musicaux sont envisagés ici suivant leur composante symbolique, et non suivant leur dimension sonore, car c'est au niveau du langage discrétisé – et non du son – que réside la logique musicale et qu'émerge la problématique de redondance combinatoire des structures musicales. L'approche proposée inclut une résolution de la problématique envisagée au niveau symbolique, assurant en retour l'exhaustivité et la compacité des résultats. L'analyse des documents sonores pourra être envisagée par la suite par l'intermédiaire d'une opération préalable de transcription automatique (cf. paragraphe 5.3).

La tâche de recherche de motifs répétés, malgré son apparente simplicité, s'avère problématique, que ce soit dans sa mise en œuvre intuitive par le musicologue, ou systématique par la modélisation informatique. En effet, les tentatives musicologiques et sémiologiques de description motivique détaillée [30, 32, 33, 36] ont rencontré des difficultés à démontrer la pertinence des motifs découverts, et se voient également dans l'incapacité d'assurer une exhaustivité de l'analyse. C'est ici que la modélisation informatique peut apporter une aide précieuse, en permettant une description précise de l'ensemble des heuristiques de découverte mises en œuvre, et en assurant une mise en application exhaustive de ces mécanismes [22]. Cependant les travaux informatiques sur ce sujet [4, 9, 11, 13, 14, 26, 27, 29, 34, 35, 38] présentent des résultats d'assez faible qualité d'un point de vue aussi bien musicologique que perceptif. En particulier, comme le reconnaissent les auteurs de ces approches eux-mêmes, la phase d'extraction motivique résulte généralement en une grande combinatoire de structures qui, par son aspect peu synthétique, présente un intérêt analytique limité. Les méthodes de filtrage proposées en aval de la phase d'extraction ne permettent généralement pas une amélioration sensible de la qualité des résultats.

L'approche proposée consiste en une explicitation détaillée de tous les facteurs de redondance structurelle, et en l'élaboration d'heuristiques de contrôle structurel permettant une meilleure concordance du comportement algorithmique avec les attentes musicologiques et les intuitions perceptives. Certaines de ces heuristiques – telles que le concept de motifs clos – s'apparentent directement aux travaux actuels en algorithmique d'extraction de connaissance ; d'autres – telles que la règle de *Gestalt* de figure/fond – offrent une modélisation concrète de phéno-

mènes étudiés dans les sciences cognitives ; d'autres enfin ouvrent de nouvelles problématiques encore peu étudiées jusqu'alors, telles que la gestion du réseau de spécificité [22].

La mise en œuvre informatique d'une telle modélisation d'inspiration cognitive donne lieu à des résultats surpassant les approches informatiques contemporaines. L'algorithme proposé offre en particulier une description à la fois compacte et détaillée des documents musicaux, présentant une pertinence musicale et perceptive forte. Ce modèle n'est toutefois qu'une première ébauche d'un système complet d'analyse motivique, car il ne peut analyser actuellement que des œuvres musicales suffisamment simples, et ne peut assurer non plus une totale pertinence des résultats. De nombreux mécanismes doivent être pris en compte afin d'intégrer l'ensemble des transformations possibles et d'envisager l'expression musicale dans toute sa généralité polyphonique.

2 Une identification adaptative et multi-dimensionnelle

Les travaux musicologiques ont mis en évidence la présence de multiples dimensions paramétriques entrant en jeu au sein du discours musical. Cette multi-dimensionnalité est une des raisons de la complexité de la structure musicale. Seront considérées, dans cette brève présentation, les deux dimensions musicales généralement considérées comme principales (FIG. 1) :

1. la dimension mélodique, exprimée par la différence de hauteurs entre notes successives, (correspondant à la dimension verticale des partitions),
2. la dimension rythmique, correspondant au positionnement temporel entre notes successives.



FIG. 1 – Une partition et sa description suivant les dimensions mélodique et rythmique. Les deux motifs encadrés sont deux occurrences d'un motif. Le motif ainsi que ses occurrences sont représentés sous la forme de chaînes d'états.

Lorsqu'une suite de descriptions est répétée à divers instants d'une pièce musicale, chacune de ses répétitions sera appelée une *occurrence* du *motif* que constitue cette suite de descriptions. Le motif pourra être représenté sous la forme d'une chaîne d'états, et les descriptions du motif

seront représentées suivant les transitions successives de cette chaîne (FIG. 1). L'ensemble des motifs peut ensuite être représenté sous la forme d'un arbre de préfixes.

2.1 L'identification de similarités

Les motifs ne sont généralement pas répétés de manière identique mais subissent différentes formes de transformations locales. Pour pouvoir détecter ces motifs, il est nécessaire de pouvoir identifier les différentes occurrences malgré la diversité de surface. Deux stratégies opposées peuvent être distinguées [22]. Une première méthode, basée sur la notion de *similarité numérique*, consiste à tolérer une certaine marge de dissimilarité [11, 34]. Toutefois, la décision d'identification dépend ici d'un seuil de dissimilarité, généralement fixé de manière arbitraire ou laissé à la discrétion de l'utilisateur, détériorant en conséquence la robustesse et l'objectivité de l'analyse.

Cette première stratégie est en outre remise en cause par certains travaux de référence en cognition musicale [15], suivant lesquels la similarité ne doit pas être envisagée en terme de minimisation d'une distance numérique, mais fondée au contraire sur une stratégie alternative d'*identification exacte* le long d'une multiplicité de dimensions musicales de divers niveaux de spécificité. Certaines approches informatiques suivent cette seconde stratégie de détection d'identités le long de multiples dimensions musicales [6, 10, 27] et recherchent des répétitions le long de chaque dimension ou produit de dimensions. L'approche proposée dans cet article s'inscrit dans cette démarche.

On peut alors montrer l'existence de motifs qui, parce que progressivement construits à partir d'une variabilité de dimensions musicales successives, ne peuvent pas être identifiés suivant de telles approches [21, 22]. Par exemple, chaque ligne de la partition de la FIG. 2 contient une répétition d'un même motif : dans sa première moitié, les paramètres mélodiques et rythmiques sont répétés alors que, dans sa seconde moitié, seuls les paramètres rythmiques sont répétés. La description de ce motif est donc, dans sa première partie, *mélodico-rythmique*, puis, dans sa seconde partie, *rythmique*. Une telle description suivant différents paramètres musicaux successifs sera qualifiée de *description hétérogène*. Le modèle présenté dans cet article, contrairement aux approches antérieures, est en mesure de détecter de telles descriptions hétérogènes.

Les choix effectués par notre approche dans ce cadre s'inscrivent dans une stratégie d'ingénierie des connaissances : la préférence du concept d'identification à celui de similarité est fondée sur des travaux développés en cognition musicale ; la définition du concept de description hétérogène, quant à lui, relève d'un questionnement musicologique.

description mélodique : $[+1 -1 +2 \ 0 \ -3 \ +1 -1 +2 \ 0 \ -3]$ $0 \ +1 \ 0 \ +1 \ +1 \ -2$
description rythmique : $1.5 \ .5 \ 1 \ 2 \ 1 \ 1.5 \ .5 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 1 \ 3$

description mélodique : $[+1 -1 +2 \ 0 \ -3 \ +1 -1 +2 \ 0 \ -3]$ $+1 \ +1 \ +1 \ -1 \ -1 \ -1$
description rythmique : $1.5 \ .5 \ 1 \ 2 \ 1 \ 1.5 \ .5 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 1 \ 3$

FIG. 2 – Répétition d'un motif hétérogène (encadré) : la première moitié de ce motif est mélodico-rythmique (les intervalles mélodiques et valeurs rythmiques sont répétés) et la seconde moitié simplement rythmique (seules les valeurs rythmiques sont répétées).

2.2 L'algorithme basique de découverte

L'algorithme de découverte au cœur de la modélisation proposée se fonde sur le principe de *mémoire associative*, c'est-à-dire la capacité de mettre en relation des entités mémorisées présentant des propriétés communes. Cette mémoire associative est modélisée par l'intermédiaire de listes inversées associées à chaque paramètre musical (ici : mélodique et rythmique), et à chaque phase d'extension successive des motifs.

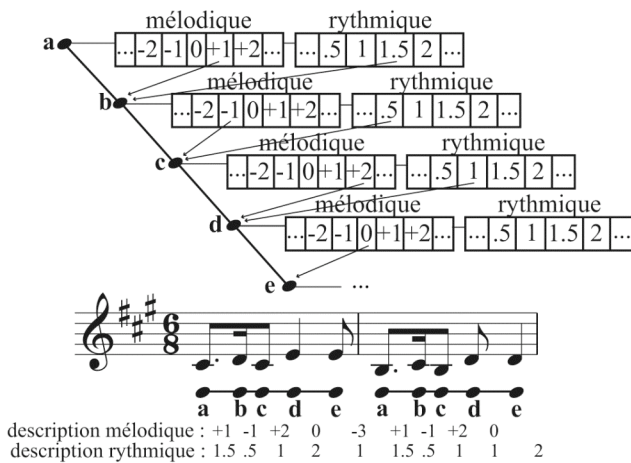


FIG. 3 – Construction progressive du motif abcde (au-dessus de la partition) et de ses deux occurrences (sous la partition), par mémorisation des intervalles dans des listes inversées associées à chaque phase successive du motif.

Des premières tables permettent de mettre en relation les descriptions d'une œuvre musicale présentant un ou plusieurs paramètres en commun (FIG. 3). Ces descriptions sont alors reliées à un nouveau nœud de l'arbre, en descendance directe de la racine, représentant ce ou ces paramètres communs. Associées à ce nœud, de nouvelles tables mettent en relation, cette fois-ci, les descriptions qui succèdent aux descriptions précédentes, et qui présentent un

ou plusieurs paramètres en commun. Ces nouvelles descriptions sont alors reliées à un nouveau nœud de l'arbre, en descendance directe du nœud précédent, et ainsi de suite. Ceci permet ainsi une découverte progressive des extensions successives de chaque motif. L'algorithme assure une découverte de motifs aussi bien homogènes qu'hétérogènes, car aucune contrainte n'est fixée *a priori* quant aux paramètres musicaux associés aux états successifs des motifs.

3 Le contrôle de la redondance combinatoire

Une mise en œuvre directe de l'algorithme de base sur des exemples musicaux, même simples, donne lieu à la découverte d'un grand nombre de motifs, qui ne correspondent pas, pour la plupart d'entre eux, à des structures effectivement perçues, et conduisent à une explosion combinatoire. Les approches actuelles dans le domaine musical envisagent une réduction de la complexité par l'intermédiaire d'un filtrage des résultats suivant des critères appliqués globalement, tels qu'une sélection des motifs les plus longs ou les plus fréquents [6, 10, 26, 27, 29]. Toutefois, un tel filtrage ne permet pas d'améliorer la pertinence perceptive des résultats, mais entraîne en revanche leur appauvrissement.

3.1 Motifs fermés multi-dimensionnels

Une première stratégie de réduction combinatoire consiste en une sélection des motifs maximaux [1, 41], c'est-à-dire des motifs non-inclus dans d'autres motifs. Cette heuristique permet une réduction importante de redondance, mais conduit également à une perte importante d'information. En effet, un sous-motif de motif ne doit pas être considéré comme structure purement redondante si son nombre d'occurrences est supérieur à celui du motif qui le contient (FIG. 4). Dans ce cas, ces motifs, dits *fermés*, doivent être conservés au sein de la description car ils appor-

tent une information spécifique [41]. Notre modélisation applique au cadre musical l’heuristique de recherches de motifs fermés [22]. Pour cela, il a fallu généraliser la relation d’inclusion entre motifs – qui fonde la définition du motif fermé – en l’intégrant dans l’espace paramétrique multi-dimensionnel caractérisant le domaine musical. Ce problème peut être résolu en considérant la correspondance de Gallois entre les espaces de classes et de descriptions des motifs, étudiée notamment dans les travaux d’Analyse de concepts formels (*Formal Concept Analysis*) [17].

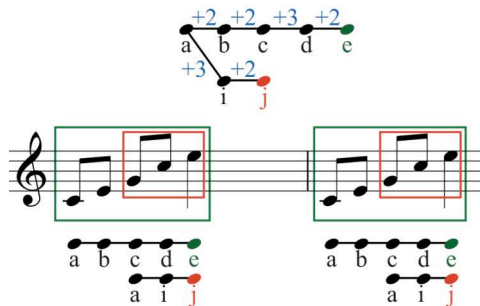


FIG. 4A – Le motif *aij* est un motif non-fermé, car présentant le même nombre d’occurrences (deux) que le motif *abcde* dans lequel il est inclus. Il peut donc être rejeté.

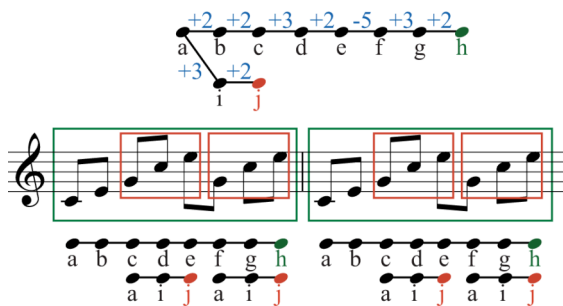


FIG. 4B – Le motif *aij* est fermé, car il présente un nombre supérieur d’occurrences (quatre), par rapport au nombre d’occurrences des motifs *abcde* et *abcdefgh* (deux) dans lesquels il est inclus.

Cet effort d’optimisation de la description motivique a été prolongé par l’ajout d’un principe de description maximale spécifique des occurrences de motifs : lorsque une occurrence d’un motif est découverte (par exemple, le motif *abcde*), les occurrences des motifs inclus (par exemple, le motif *aij*) ne seront pas superposées sur cette occurrence, car elles n’apportent aucune information nouvelle, et peuvent être directement déduites à partir de l’occurrence du motif plus spécifique (*abcde*). Les descriptions moins spécifiques doivent toutefois être prises en compte implicitement car leur extension peut parfois mener à des descriptions spécifiques [22].

La mise en œuvre de tels principes nécessite une gestion soignée du réseau des motifs, et notamment un parcours de l’ensemble des motifs et de l’ensemble des extensions possibles orienté du plus spécifique vers le plus général. Un tel mécanisme permet ainsi d’assurer une description motivique optimale, par compression sans perte de la taille de la description, suivant un filtrage de la redondance. Cette optimisation est nécessaire d’une part afin d’assurer la clarté des résultats, et d’autre part afin de limiter la complexité combinatoire des calculs.

Cette modélisation fondée sur les motifs fermés ne relève ni d’une problématique musicologique *stricto sensu*, ni de principes de *Gestalt*. Mais elle met en évidence un ensemble de conditions nécessaires à une modélisation de l’extraction motivique, permettant un fonctionnement optimal du système, tant dans son élaboration computationnelle que dans la pertinence musicologique des résultats.

3.2 Une modélisation par cycles des répétitions successives de motifs.

Cette section présente un autre facteur général de redondance combinatoire, maîtrisé par l’intermédiaire d’une approche de type *Gestalt*. Cette explosion combinatoire survient lors de la répétition successive d’un même motif (par exemple le motif rythmique *abc* à la FIG. 5). Comme chaque occurrence est suivie du début d’une nouvelle occurrence, chaque motif peut être étendu (aboutissant au motif *abcd*) d’un intervalle dont la description est identique à celle du début de ce même motif (c’est-à-dire *ab*). Une telle extension peut alors être prolongée de manière récursive (en *abcde*, *abcdef*, etc.), aboutissant en une explosion combinatoire de motifs qui ne sont pas perçus en tant que tels en raison de leur enchevêtrement complexe [4].

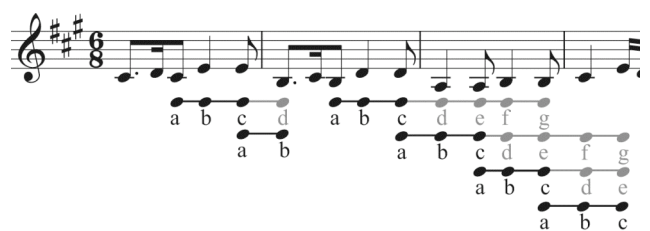


FIG. 5 – La répétition successive multiple du motif *abc* entraîne en toute logique une extension de ce motif en *abcd*, *abcde*, etc., formant un enchevêtrement complexe de structures qui ne sont pas perceptibles en tant que telles.

Une telle configuration structurelle est au contraire généralement perçue de manière beaucoup plus simple sous la forme d’un principe *Gestalt* de cyclicité motivique. La représentation par graphe de la FIG. 5 met en effet en évidence le fait que le dernier état de chaque occurrence du motif *abc* se superpose au premier état (*a*) de l’occurrence suivante. Ces deux états peuvent alors être fusionnés, abou-

tissant à un rebouclage du dernier état (c) du motif vers le premier (a) (FIG. 6). Se constitue ainsi un motif cyclique oscillant entre deux états a' et b' , à partir du motif acyclique initial abc . Une telle modélisation à l'aide de cycles correspond à une stratégie effective d'écoute musicale, tout en réduisant de manière optimale la redondance combinatoire.

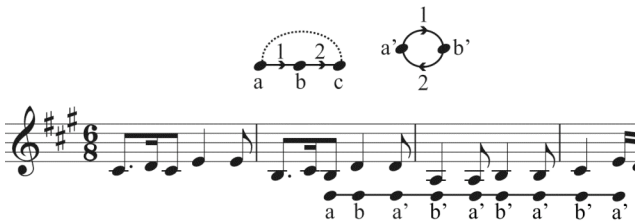


FIG. 6 – L'écoute d'une répétition successive du motif abc induit la perception d'une cyclicité de ce motif, aboutissant en une oscillation entre deux états a' et b' .

Ce phénomène de répétition successive, bien que très fréquent notamment dans l'expression musicale, a été peu étudié. Une stratégie simple consiste en l'interdiction de tout chevauchement entre motifs [37]. Cette heuristique se fonde cependant sur une contrainte forte – de segmentation stricte et unidimensionnelle de la séquence temporelle – qui n'est généralement pas valable dans le domaine musical. De manière moins contraignante, la minimisation des chevauchements de motifs a été envisagée séparément pour chaque classe de motif [4]. Mais une telle sélection, effectuée sur les motifs de manière globale, provoque une excessive suppression de motifs pertinents, alors que notre approche fondée sur les configurations locales entre occurrences de motifs permet un filtrage plus précis. La problématique de recherche automatique de motifs périodiques au sein de séquences temporelles a été étudiée dans des travaux récents [18, 19, 25, 40]. Ces travaux ne prennent

cependant pas en compte l'articulation importante entre motifs périodiques et motifs non-périodiques, et ne permettent donc pas de maîtriser le phénomène d'explosion combinatoire.

L'approche proposée ici permet au contraire une articulation étroite entre motifs périodiques et non-périodiques, assurant une réduction sans perte de la redondance combinatoire. Une telle articulation nécessite une lecture strictement *chronologique* de la séquence musicale [22]. La problématique de périodicité a été généralisée à la multidimensionnalité de l'espace paramétrique musical. À cette fin, il a été nécessaire d'étendre la notion de relation de spécificité aux motifs cycliques, et de développer des mécanismes de généralisation et spécification progressives des cycles motiviques, afin de permettre une description la plus compacte possible de la séquence musicale. Cette formalisation a permis de mettre en évidence d'autres problématiques structurelles de première importance et d'en proposer des méthodes de résolution simples. On a ainsi montré la nécessité d'intégrer des heuristiques apparentées à la règle de *Gestalt* figure/fond, et proposé des mécanismes permettant d'unifier toutes les rotations possibles d'un même motif cyclique sous une seule et même classe [22].

4 Résultats

Cette modélisation a été développée sous la forme d'un système d'analyse motivique automatisée écrit en *Common Lisp* et intitulé *OMkanthus*, et a été intégrée au sein du logiciel *OpenMusic* de programmation graphique de structures musicales [2]. Cette bibliothèque est en mesure d'analyser des pièces musicales monodiques – c'est-à-dire constituées d'une suite de notes temporellement non superposées – et de mettre en évidence sous forme graphique les structures découvertes sur une partition.

TABLE 1 – Résultats d'analyses effectuées par *kanthus* version 0.6.8.

Nom du morceau	Taille du morceau	Classes de motifs découvertes	Classes de motifs pertinentes	Temps de calcul
<i>Geisslerlied</i> (chanson médiévale)	108 notes	6	5 (83%)	2.2 secondes
Mozart, <i>Sonate K 331</i> 1 ^{er} mouvement, 1 ^{er} thème	36 notes	14	10 (71%)	0.8 seconde
Bach, <i>Invention BWV 775</i>	283 notes	49	34 (69%)	37.6 secondes
<i>Au clair de la lune</i>	44 notes	26	20 (77%)	5.6 secondes
The Beatles, <i>Obla Di, Obla Da</i>	390 notes	14	10 (71%)	28.1 secondes

FIG. 7 – Analyse motivique de l’Invention en ré mineur BWV 775 de J.S. Bach (21 premières mesures). Les occurrences de chaque classe de motifs sont représentées de manière distincte.

4.1 Evaluation musicologique

Le modèle a été testé sur un ensemble d’œuvres musicales de styles variés. Quelques résultats sont présentés dans le tableau 1. Les résultats ont été comparés avec une analyse manuelle réalisée par un musicologue. Un exemple d’analyse est présenté à la FIG. 7. Une procédure de test est en projet d’élaboration sur une base de donnée musicale de grande taille, afin d’établir une mesure quantitative précise du comportement algorithmique.

Les approches informatiques actuelles ne sont pas en mesure de présenter un tel degré de précision, ni d’assurer la pertinence perceptive de leurs découvertes. Les analyses produites par le modèle proposé restent cependant encore assez limitées, car de nombreuses composantes de l’expression musicale n’ont pas encore été prises en compte.

4.2 Complexité algorithmique

La complexité algorithmique s’exprime sous deux formes. D’une part, par la complexité des structures découvertes : la prolifération de motifs redondants, par exemple, induit une explosion combinatoire, puisque chaque nouvelle structure nécessite des processus permettant d’estimer leur impact sur les autres structures et d’envisager leur éventuelle extension. Assurer une description la plus compacte possible des structures permet ainsi tout à la fois d’offrir des résultats clairs et proches des motifs perçus, et de réduire les risques d’explosion combinatoire. La complexité algorithmique s’exprime d’autre part à travers la mise en œuvre technique de la modélisation. De ce point de vue, la réalisation proposée se présente sous la forme d’un prototype qui demande à être davantage optimisé. La modélisation a toutefois été conçue dans un souci constant de minimisation des coûts computationnels. Ainsi la détection de répétitions d’intervalles s’effectue par simple identification au sein de listes inversées, ce qui permet de simplifier la procédure élémentaire de recherche de motifs.

4.3 Autres approches

La problématique d’extraction de motifs répétés au sein de séquences musicales symboliques a été envisagée suivant de multiples approches. Cambouropoulos [5] procède à une recherche de répétitions exactes au sein de projections suivant diverses dimensions (ou conjonction de dimensions) de l’espace musical. L’ensemble de motifs extraits n’est pas considéré tel quel : au contraire, une segmentation globale de la séquence musicale est estimée par l’intermédiaire d’un moyennage des multiples segmentations élémentaires, aboutissant à une compression avec forte perte de l’information structurelle.

Une autre approche [10] consiste en la construction d’un arbre de suffixes suivant une multitude de dimensions musicales distinctes. Ici aussi, en raison de la taille importante de l’ensemble des motifs découverts, un post-traitement sélectionne les motifs présentant un nombre minimal d’occurrences et un degré de pertinence statistique suffisant.

Rolland [34] définit une distance numérique de similarité fondée sur une distance d’édition entre sous-séquences. Des distances de similarité sont estimées entre tous les couples possibles de sous-séquences, et seules les similarités dépassant un certain seuil prédéfini par l’utilisateur sont sélectionnées. À partir du graphe de similarité résultant, des motifs sont extraits suivant l’algorithme de catégorisation *Star center*. L’ensemble de motifs découverts est réduit encore davantage par l’intermédiaire d’heuristiques de filtrage *offline*. En particulier, seuls les motifs répétés au sein d’un nombre minimum de séquences musicales sont sélectionnés.

Une généralisation de l’extraction motivique à la polyphonie a également été proposée [27]. Les notes des séquences musicales sont représentées sous forme de points au sein de l’espace bi-dimensionnel hauteur / temps, et les répétitions maximales d’ensemble de points sont extraits. Cette stratégie d’inspiration géométrique ne s’applique toutefois pas aux répétitions mélodiques présentant des

variations rythmiques. Des post-traitements ont été ajoutés afin de réduire la redondance combinatoire.

Dans l'ensemble de ces approches, l'explosion combinatoire produite par la phase d'extraction motivique *stricto sensu* est filtrée *a posteriori* suivant des critères globaux basés sur la longueur des motifs, leur fréquence, etc. La limitation principale de telles approches est due au manque de sélectivité de ces critères globaux.

Notre approche s'est attaché en particulier à mettre en évidence les raisons profondes de ces redondances combinatoires. Les heuristiques proposées permettent une représentation compacte, factorisée, mais sans perte, des configurations motiviques. Les analyses résultantes présentent un degré de clarté et de compacité jamais égalé par les approches antérieures.

5 Travaux actuels

5.1 Prise en compte de transformations musicales complexes

Un motif, formé d'une suite de descriptions le long de divers paramètres musicaux, est découvert lorsque une telle suite est explicitement présente à divers instants d'une séquence musicale. Or une caractéristique essentielle du discours musical est l'existence de mécanismes de transformation musicale, suivant lesquels une suite de descriptions, sous une forme originaire simple, se modifie en une suite de descriptions plus complexe. En particulier, un motif peut être *ornementé* par l'ajout de notes supplémentaires dans un voisinage proche (cf. FIG. 8). Ces ornements n'appartiennent pas à la description du motif, mais ont pour fonction de mettre en valeur les notes auprès desquelles elles sont situées.



FIG. 8 – Répétition d'un motif avec *ornementation* : la deuxième occurrence du motif est enrichie de notes supplémentaires qu'il est nécessaire de filtrer afin de retrouver la répétition de la structure sous-jacente.

Des solutions à cette problématique de recherche de motifs de notes non immédiatement successives ont été proposées [13, 31, 34]. Elles se basent principalement sur des techniques de programmation dynamique, qui permettent un alignement des différentes occurrences, fondé sur une distance d'édition entre séquences musicales. Or, comme expliqué au paragraphe 2.1, l'utilisation d'une distance numérique introduit des difficultés supplémentaires. Cela conduit notamment à un assujettissement de la décision d'identification à un seuil de similarité numérique, qui ne

peut être défini que de manière subjective et remet donc en cause la robustesse de l'analyse.

Une nouvelle méthode de résolution est actuellement en cours d'élaboration, fondée sur une identification adaptative, et sur une réduction de l'espace de solutions par l'intermédiaire d'heuristiques basées sur des considérations d'ordre perceptuel. Pour cela, sur la *chaîne syntagmatique* – que constitue la succession apparente de notes de la séquence musicale – est construit un *graphe syntagmatique* contenant l'ensemble de toutes les connexions perceptibles entre notes temporellement proches (cf. FIG. 9). Des hypothèses d'ordre cognitif et des expérimentations informatiques permettent de mettre en place un ensemble de règles limitant la combinatoire du graphe syntagmatique. Dans un second temps, l'algorithme d'extraction motivique, initialement conçu pour l'analyse de chaînes syntagmatiques, sera généralisé à ce nouveau cadre, à travers un parcours de toutes les chaînes contenues au sein du graphe syntagmatique. L'élaboration d'un système complet et fiable de règles de réduction du graphe syntagmatique nécessitera un effort de recherche important. D'autre part, ce nouveau paradigme induit de nouveaux facteurs de redondance combinatoire qu'il est nécessaire de contrôler de manière optimale.



FIG. 9 – Afin de détecter la répétition de la structure profonde, un graphe syntagmatique (indiqué en gris) est construit, et la recherche de répétition motivique est effectuée le long des différentes branches

5.2 Généralisation vers la polyphonie

La musique a été considérée, dans notre approche ainsi que dans la plupart des approches actuelles, sous la forme d'une séquence de notes successives, ou *monodies*. Mais de manière générale, la musique doit être envisagée en terme d'un flux complexe de notes, une *polyphonie*, pouvant contenir en particulier des superpositions de notes. Dans un tel flux peuvent être perçues, d'une part, des superpositions de monodies simultanées. En d'autres termes, à partir du réseau initial que constitue le flux polyphonique émerge un graphe de chaînes syntagmatiques monodiques, lequel pouvant présenter des configurations complexes telles que des embranchements et des entrelacements (FIG. 10). Suivant ces diverses chaînes monodiques peut alors être opérée la recherche de répétitions de motifs assurée par ma modélisation. Une généralisation de l'approche au cadre polyphonique nécessitera donc la formalisation et modélisation de ce processus d'émergence du réseau monodique à partir du flux polyphonique.



FIG. 10 – Construction d'un graphe syntagmatique à partir d'une polyphonie.

Au sein d'un flux polyphonique émergent non seulement des chaînes monodiques, mais également des groupements de notes qui s'agglomèrent en raison de leur proximité temporelle, que l'on pourra appeler des *agrégats*, et qui forment diverses constructions musicales telles que les accords, les degrés ou les modes. Des exemples d'accords sont encadrés dans la FIG. 10. De même que les motifs résultent de répétitions de chaînes monodiques, la répétition d'agrégats induit la formation de classes d'agrégats. D'autre part, les agrégats se succédant forment eux-mêmes des chaînes d'agrégats. Là où la description de chaînes monodiques consistait simplement en une suite de distances entre notes successives, la description de chaînes d'agrégats aboutit à un réseau complexe de distances entre notes d'agrégats successifs, ainsi que de distances entre classes d'agrégats, qu'il faudra formaliser et modéliser. Il s'agira par la suite de généraliser aux chaînes d'agrégats la problématique de recherche de répétition de chaînes. La résolution de l'ensemble de ces problèmes permettra d'assurer une modélisation complète des mécanismes en jeu dans l'analyse motivique.

La problématique de recherche de motifs polyphoniques a été peu étudiée [10, 14, 27, 29]. L'approche la plus avancée [27] procède par une recherche de répétitions exactes de suites de coordonnées exprimées dans l'espace bi-dimensionnel temps / hauteurs. En conséquence, cette approche ne peut pas découvrir des répétitions de motifs présentant des fluctuations le long de certaines des différentes dimensions musicales – transformation du rythme, ou des hauteurs, etc. –, lesquelles pourtant ne remettent pas en cause, chez les auditeurs, l'identification des motifs. L'approche envisagée ici permettra justement une telle identification adaptative le long des différentes dimensions, car ma démarche est fondée sur une modélisation la plus précise possible des stratégies d'écoute, et non sur une méthode fondée sur des critères simplement géométriques ou statistiques.

5.3 Analyse de masses de données musicales

Un important domaine d'application de l'automatisation de l'analyse motivique concerne l'analyse des masses de données musicales. Sur l'internet se développent en effet des immenses catalogues de musique en ligne, mises à la

disposition du grand public, et qui occupent une part de marché de plus en plus dominante dans le commerce musical. La conception d'outils de gestion, de visualisation et de navigation au sein de ces bases de données, est une problématique de grande importance. Une communauté scientifique internationale s'est constituée autour de cette thématique, intitulée *Music Information Retrieval* [16], créant un lien étroit entre les recherches en informatique musicale et les applications industrielles.

Les travaux présentés dans cet article pourront être appliqués en particulier à deux problématiques majeures dans ce domaine : d'une part l'élaboration de mesures de distance de similarité entre œuvres musicales, et d'autre part le développement de méthodologies de recherche au sein de ces catalogues. La similarité musicale, tout d'abord, est envisagée dans les travaux récents à travers de multiples composantes de l'expression musicale, relative notamment à la texture sonore [3]. D'autres recherches étudient au contraire l'impact des structures musicales plus profondes sur les jugements de similarité musicale [12, 23]. La modélisation présentée ici, offrant une description précise de la dimension motivique de ces structures musicales, permettra en conséquence un approfondissement de ce champ de recherche, et pourra donc donner lieu à des applications intéressantes dans ce domaine.

Les systèmes de fouilles de données musicales, quant à elles, ont pour objectif de retrouver, au sein d'un catalogue, les œuvres correspondants à une requête spécifiée par l'utilisateur. Certaines approches dans ce domaine procèdent par un recoupement exhaustif de la requête avec chaque instant successif de chaque œuvre du catalogue musical [39]. À l'opposé, l'espace de recherche peut être réduit significativement par l'intermédiaire d'une analyse préalable de chaque œuvre du catalogue : la requête peut alors être simplement comparé avec les représentations compactes produites par les analyses préalables [20]. Cette focalisation sur les structures les plus saillantes des œuvres musicales assure en outre une amélioration qualitative de la pertinence perceptive des résultats. Dans ce contexte, notre système d'analyse automatique offrira une description compacte de la composante motivique de chaque œuvre du catalogue, permettant à terme une amélioration des performances de fouilles de données.

Dans notre modélisation, les séquences musicales sont décrites suivant des paramètres symboliques directement issus du texte musical, c'est-à-dire de la partition. Les bases de données musicales en ligne, quant à elles, sont généralement décrites de manière soit semi-symbolique, soit sonore. La description semi-symbolique, tout d'abord, suit généralement le protocole MIDI de contrôle informatique des instruments de musique, dans laquelle les paramètres musicaux (hauteur et rythme) ne s'identifient pas exactement aux données musicales symboliques, mais en représentent une quantification numérique. Afin de pouvoir analyser de telles bases de données, il est nécessaire de

mettre en œuvre une opération de transcription automatique des descriptions semi-symboliques en descriptions symboliques. Les méthodes proposées dans la littérature de quantification des hauteurs [5, 8, 28, 38] et des durées [7] ne répondent pas entièrement au problème, car cette opération de transcription se fonde partiellement sur la structuration motivique des œuvres, facteur qui a été peu pris en compte jusqu'alors. La modélisation de l'émergence de cette structuration offerte par notre approche permettra de prendre en compte cette interaction, afin d'envisager une amélioration des résultats.

Par contre, la majeure partie des bases de données musicales consiste en des enregistrements sonores. Une analyse de ces bases nécessite une opération de transcription automatique du sonore vers le symbolique, problématique d'une grande complexité et qui fait l'objet de nombreuses recherches actuelles. Sont proposés notamment des outils de transcription du sonore vers le semi-symbolique, que l'on pourra transcrire de nouveau vers une représentation symbolique. Mais ici encore, il semble exister une relation d'interdépendance entre l'opération de transcription à partir de la représentation sonore et la recherche de structuration musicale.

Les résultats actuels permettent de montrer l'intérêt d'une telle approche : d'une part, le système informatique qui résulte de cette étude offre au musicologue une description exhaustive des motifs d'une partition ; d'autre part, l'expérimentation algorithmique indique des conditions nécessaires relatives à la faisabilité d'une modélisation cognitive. Enfin, une telle automatisation suggère des applications très intéressantes dans l'étude des bases de données musicales.

Références

- [1] R. Agrawal et R. Skirant. Mining Sequential Patterns. *International Conference on Data Engineering*, Taipei, Taiwan, 1995.
- [2] G. Assayag, C. Rueda, M. Laurson, C. Agon et O. Delerue. Computer Assisted Composition at Ircam: From Patchwork to Openmusic. *Computer Music Journal*, 23(3), p. 59-72, 1999.
- [3] J.-J. Aucouturier et F. Pachet. Music similarity measures : What's the use ? *International Conference on Music Information Retrieval*, 2002.
- [4] E. Cambouropoulos. *Towards a General Computational Theory of Musical Structures*. Thèse de l'Université de Edinbourg, 1998.
- [5] E. Cambouropoulos. Pitch spelling: A computational model. *Music Perception*, 20, p. 411-429, 2003.
- [6] E. Cambouropoulos. Musical Parallelism and Melodic Segmentation: A Computational Approach. *Music Perception*, à paraître.
- [7] A.T. Cemgil et B. Kappen. Monte Carlo methods for tempo tracking and rhythm quantization. *Journal of Artificial Intelligence Research*, 18, pp. 47-81, 2003.
- [8] E. Chew et Y.C. Chen. Real-Time Pitch Spelling Using the Spiral Array. *Computer Music Journal*, 29(2), 2005.
- [9] D. Conklin. Representation and discovery of vertical patterns in music. Dans C. Anagnostopoulou, M. Ferrand et A. Smaill (éd.). *Music and Artificial Intelligence*, Springer-Verlag, pp. 32-42, 2002.
- [10] D. Conklin et C. Anagnostopoulou. Discovery of segmental patterns in music. *INFORMS Journal on Computing*, à paraître.
- [11] D. Cope. *Experiments in Musical Intelligence*, A-R Editions, 1996.
- [12] T. Crawford, C.S. Iliopoulos et R. Raman. String-matching techniques for musical similarity and melodic recognition. Dans W. Hewlett et E. Selfridge-Field (éd.). *Melodic similarity : Concepts, procedures and applications (Computing in Musicology 11)*, MIT Press, 1998.
- [13] R.B. Dannenberg et N. Hu. Pattern Discovery Techniques for Music Audio. *Journal of New Music Research*, 32(2), p. 153-164, 2002.
- [14] M.J. Dovey. A technique for "regular expression" style searching in polyphonic music. *International Conference on Music Information Retrieval*, 2001.
- [15] W.J. Dowling et D.L. Harwood. *Music Cognition*, Academic Press, 1986.
- [16] J.S. Downie. Music information retrieval. *Annual Review of Information Science and Technology*, 37, p. 295-340.
- [17] B. Ganter et R. Wille. *Formal Concept Analysis: Mathematical Foundations*, Springer-Verlag, 1999.
- [18] J. Han, W. Gong et Y. Yin. Mining Segment-Wise Periodic Patterns in Time-Related Databases. *International Conference on Knowledge Discovery and Data Mining*, 1998.
- [19] J. Han, G. Dong et Y. Yin. Efficient Mining of Partial Periodic Patterns in Time Series Database. *International Conference on Data Engineering*, 1999.
- [20] H. Hoos, K. Renz et M. Görg. GUIDO/MIR : an Experimental Musical Information Retrieval System based on GUIDO Music Notation. *International Conference on Music Information Retrieval*, 2001.
- [21] O. Lartillot. Efficient Extraction of Closed Motivic Patterns in Multi-Dimensional Symbolic Representations of Music. *IEEE/WIC/ACM International Conference on Web Intelligence*, Compiègne, IEEE Computer Society Press, 2005.
- [22] O. Lartillot. Multi-Dimensional Motivic Pattern Extraction Founded on Adaptive Redundancy Filtering. *Journal of New Music Research*, à paraître.
- [23] K. Lemström, G. Navarro et Y. Pinzon. Bit-Parallel Algorithms for Transposition-Invariant Multi-Track String-Matching. *Journal of Discrete Algorithms* 3(2-4), p. 267-292, 2005.
- [24] J. Lin, E. Keogh, S. Lonardi et P. Patel. Finding Motifs in Time Series. *International Conference on Knowledge Discovery and Data Mining*, 2002.
- [25] S. Ma et J. Hellerstein. Mining partially periodic event patterns with unknown periods. *International Conference on Data Engineering*, 2001.
- [26] C. Meek et W.P. Birmingham. Thematic Extractor. *International Conference on Music Information Retrieval*, 2001.
- [27] D. Meredith, K. Lemström et G.A. Wiggins. Algorithms for Discovering Repeated Patterns in Multidimensional Representations of Polyphonic Music. *Journal of New Music Research*, 31(4), p. 321-345, 2002.

Lartillot

- [28] D. Meredith et G. Wiggins. Comparing Pitch Spelling Algorithms. *International Conference on Music Information Retrieval*, London, 2005.
- [29] B. Meudic et E. Saint-James. Automatic Extraction of Approximate Repetitions in Polyphonic MIDI Files Based on Perceptive Criteria. Dans U.K. Will (éd.). *Computer Music Modeling and Retrieval*, Springer-Verlag, 2004.
- [30] L.B. Meyer. *Emotion and Meaning in Music*, The University of Chicago Press, 1956.
- [31] M. Mongeau et D. Sankoff. Comparison of Musical Sequences. *Computer and Humanities*, 24, p. 161-175, 1990.
- [32] J.-J. Nattiez. *Fondements d'une sémiologie de la musique*, Union Générale d'Éditions, 1975.
- [33] R. Reti. *The Thematic Process in Music*, Macmillan Publishing, 1951.
- [34] P.-Y. Rolland. Discovering Patterns in Musical Sequences. *Journal of New Music Research*, 28(4), p. 335-350, 1999.
- [35] R. Rowe. *Interactive Music Systems (Machine Listening and Composing)*, The MIT Press, 1993.
- [36] N. Ruwet. Méthodes d'analyse en musicologie. *Revue belge de musicologie*, 20, p. 65-90, 1966.
- [37] Y. Tanaka, K. Iwamoto et K. Uehara. Discovery of Time-Series Motif from Multi-Dimensional Data Based on MDL Principle. *Machine Learning*, 58, 2005.
- [38] D. Temperley. *The Cognition of Basic Musical Structures*, MIT Press, 1988.
- [39] E. Ukkonen, K. Lemström et V. Mäkinen. Geometric Algorithms for Transposition Invariant Content-Based Music Retrieval. *International Conference on Music Information Retrieval*, p. 193-199, 2003.
- [40] J. Yang, W. Wang et P.S. Yu. InfoMiner+: Mining Partial Periodic Patterns with Gap Penalties. *IEEE International Conference on Data Mining*, 2002.
- [41] M. Zaki. Efficient algorithms for mining closed itemsets and their lattice structure. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17 (4), p. 462-478, 2005.