Musicae Scientiae
Discussion Forum 4A, 2007, 281-314

Motivic matching strategies for automated pattern extraction

OLIVIER LARTILLOT AND PETRI TOIVIAINEN
University of Jyväskylä, Department of Music, Finland

• **ABSTRACT**

This article proposes an approach to the problem of automated extraction of motivic patterns in monodies. Different musical dimensions, restricted in current approaches to the most prominent melodic and rhythmic features at the surface level, are defined. The proposed strategy of detection of repeated patterns consists of an exact matching of the successive parameters forming the motives. We suggest a generalization of the multiple-viewpoint approach that allows a variability of the types of parameters (melodic, rhythmic, etc.) defining each successive extension of these motives. This enables us to take into account a more general class of motives, called heterogeneous motives, which includes interesting motives beyond the scope of previous approaches. Besides, this heterogeneous representation of motives may offer more refined explanations concerning the impact of gross contour representation in motivic analysis. This article also shows that the main problem aroused by the pattern extraction task is related to the control of the combinatorial redundancy of musical structures. Two main strategies are presented, that ensure an adaptive filtering of the redundant structures, and which are based on the notions of closed and cyclic patterns. The method is illustrated with the analysis of two pieces: a medieval *Geisslerlied* and a Bach *Invention*.

1. INTRODUCTION

Motives are musical structures that constitute one of the most characteristic descriptions of music. The perception of the motivic structure is generally governed by two main heuristics. Firstly, discontinuities of the sequential structure of music along its different dimensions imply the inference of segmentations (Lerdahl & Jackendoff, 1983). The strength of each segmentation depends on the size of the corresponding discontinuities. A local maximum of inter-pitch and/or inter-onset interval amplitude, or the accentuation of one particular note, are common examples of such local discontinuities. These segmentations result in a rich structural configuration. The multiple principles ruling these segmentations, such as Lerdahl and Jackendoff's Grouping Preference Rules (Lerdahl & Jackendoff, 1983), can be ordered relative to their perceptive salience (Deliège 1987). The second general principle, on which this article is focused, is motivic extraction based on the concept of pattern repetition. Contrary to local segmentation, the structures extracted as a result of the pattern heuristics are associated to concepts (the description of these repeated patterns) that form a lexicon of characteristic elements.

The motivic structure is often highly complex. The most salient and characteristic motives define the *themes*. A more detailed analysis shows the existence of deeper motivic structures that proliferate throughout the work. Some of these cells are specific material created in the context of the piece, while others are common stylistic features, also known as "signatures", that are used in a particular musical style (Cope, 1996). Detailed analysis of the deeper motivic structures contained in music has been undertaken during the twentieth century (Reti, 1951). In previous works, systematic approaches have been suggested, with a view to augmenting the analytic capabilities, both in quantitative and qualitative terms (Ruwet, 1966-1987; Nattiez, 1975-1990; Lerdahl and Jackendoff, 1983).

Computational modelling offers the possibility to automate the process, enabling the fast annotation of large scores, and the extraction of complex and detailed structures without much effort. One major difficulty here is to ensure the musical interest of the computer-based analyses, and in particular their perceptual relevance[1]. It is assumed here that analyses produced by alternative strategies or algorithms cannot all be considered as equally valuable, and should instead be evaluated according to their musical relevance. Yet no consensus seems to have been reached among musicologists as to the criteria by which this questioned notion of musical relevance should be defined. On the contrary, the analysis of a single piece by different musicologists may show important variability, expressing the subjectivity of the musicologists' approaches. The aim of computational modelling here would be to make explicit the spectrum of strategies that musicologists may choose to use for their analysis. Due to the experimental aspect of current computational approaches, including the one presented in this article, this complex question cannot be answered for the moment. As a first approach, the analysis may focus mainly on the simplest and most evident musical structures, whose automated discovery remains a scientific challenge.

This article proposes a solution to the problem of automated extraction of motivic patterns, restricted to the study of monodies and simple musical transformations. Section 2 presents different musical dimensions, restricted to the most prominent melodic and rhythmic features at the surface level. The strategy of detection of repeated patterns is explained in section 3. It consists of an exact matching of the successive parameters forming the motives. We suggest a generalization of the multiple-viewpoint approach by allowing a variability of the types of parameters

(1) Structuralism-based approaches (such as serialism) will not be considered in this article.

(melodic, rhythmic, etc.) defining each successive extension of these motives. This enables us to take into account a more general class of motives, called heterogeneous motives, which include salient motives that have remained outside the scope of previous approaches. Besides, this heterogeneous representation of motives may offer more refined explanations of the impact of gross contour representation in motivic analysis. This article also shows that the main problem caused by the pattern extraction task is related to the control of the combinatorial redundancy of musical structures. Section 4 presents two main strategies, which enable an adaptive filtering of the redundant structures based on the notions of closed and cyclic patterns. Results offered by this model are presented in section 5, and compared with analyses by Nicolas Ruwet and Jeffrey Kresky. Current and future directions of research are discussed in section 6.

2. Definition of the Parametric Space

This section presents the different musical dimensions currently integrated into our model. The study is restricted to monodies and does not take into account more complex polyphonic relations between notes[2].

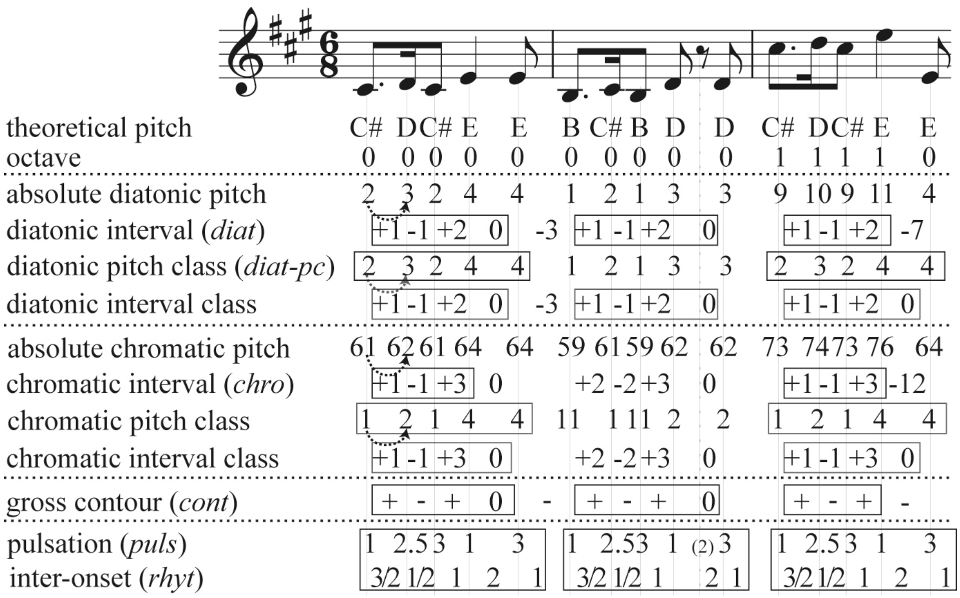

Figure 1.

Descriptions of a monody. Repeated sequences of values, forming patterns, are enclosed in boxes.

(2) See section 6.3. for a brief evocation of the generalization of the approach to polyphony.

The diatonic pitch representation indicates the height of each note with respect to the implicit tonal scale. This information can be directly obtained from the score when the tonality of a piece strictly follows the indication given by the key signature. In more general cases — not yet considered in our approach — local modulations need to be taken into account through a proper harmonic analysis. In particular, when analysing MIDI files where no tonality is specified explicitly, diatonic pitch representations need to be reconstructed using pitch-spelling algorithms (Cambouropoulos, 2003; Chew and Chen, 2005; Meredith, 2006). The result of the automated pitch spelling can be directly imported into the pattern extraction algorithm.

Absolute diatonic pitch values are represented on a numeric scale whose origin (0) is set at one tonic of the scale (see figure 1). *Diatonic pitch class* values are obtained by applying a modulo 7 operation to the absolute diatonic pitch values. The integer obtained by dividing the absolute diatonic pitch values by 7 gives the *octave position*. Alternatively, *absolute chromatic pitch classes* are represented on a chromatic scale, where, following the MIDI convention, the value of 60 is associated with pitch C4. Similarly to diatonic pitch, *chromatic pitch class* values are obtained by applying a modulo 12 operation to the absolute chromatic pitch values. In our system, due to the automated management of parametric dimensions according to their specificity relationships (as will be explained in section 4.2), the simple addition of the absolute pitch information automatically enables the discovery of transposition-invariant sub-classes, such as pattern *A* in the example of section 5.2.2.

Relative pitch configurations are modelled by defining the position of each successive pitch with respect to its direct neighbours within the monody, defining interval-based dimensions. Intervals can be defined either between absolute pitches or between pitch classes, resulting in two separate dimensions called respectively *absolute interval* and *interval class*. This distinction can be drawn for both diatonic and chromatic dimensions. For instance, the *chromatic interval class* dimension is used in Pitch-Class-Set theory (Forte, 1973), where interval classes are more important than absolute intervals. Absolute pitch intervals can be perceived more simply as *gross contours*, i.e., simple successions of ascending, descending or constant pitches. Studies have shown the perceptive importance of gross contour dimensions (White, 1960; Dowling and Harwood, 1986): distorted repetitions of the same motive can be recognised even if the interval values have been significantly changed, as long as the gross contour remains constant. On the other hand, due to the small alphabet of this dimension, repetition of gross contour motives cannot be perceptually detected if the occurrences are too distant in time (Dowling and Harwood, 1986). The impact of gross contour in pattern extraction will be further discussed in section 3.3.

Metrical position indicates the phase of each note with respect to the metrical structure. In a first approach, the metrical structure is represented by a main pulsation (defining onbeats) subdivided into another pulsation (defining offbeats), which is generally either twice or three time fasters (corresponding to so-called binary and ternary rhythms). Onbeats are indicated by a value of 1, whereas offbeats are indicated by a value of 2 in binary rhythm, and 2 and 3 in ternary rhythm. This dimension is a first attempt to represent metric hierarchy. This information can be directly obtained from the score. However, when analysing MIDI files, the metrical structure is not specified explicitly, and needs to be reconstructed using beat-tracking (Toiviainen, 1998; Large and Kolen, 1994; Dannenberg and Mont-Reynaud, 1987), quantization (Desain and Honing, 1991; Cemgil and Kappen, 2003), and meter induction algorithms (Toiviainen and Eerola, 2006; Eck and Casagrande, 2005). The result of these algorithms can be directly integrated as input of the pattern extraction algorithm.

The metric position dimension plays an important role in rhythmic identification. In particular, a rhythmic pattern is generally neither detected nor recognized when its phase is altered with respect to metrical structure (Povel and Essens, 1985; Ahlbäck, this volume). This constraint has been integrated into the model: a filter excludes any rhythmic repetition that does not agree with the metrical structure of the original motive.

The rhythmic description of notes is generally expressed along two main distinct parameters: note duration and inter-onset intervals. Durations are rhythmic values explicitly associated with each note, whereas inter-onset intervals correspond to the temporal distance between successive note onsets in the monody. Inter-onsets might be considered as more prominent than note durations because note onsets are perceptually more salient than offsets. For instance, in figure 1, the inequality between occurrences of bars 1 and 2 in terms of rhythmic value — the quarter note in bar 1 transformed into a succession of an eighth note and an eighth-note rest in bar 2 — is a detail that does not mask the inter-onset identity. For this reason, the duration parameter has been discarded from the analysis. Since, in this paper, we are dealing with monophonic sequences, inter-onset interval is defined as the interval between successive note onsets. The specification of this parameter is more complex in polyphony as it requires a voice separation.

The set of musical dimensions considered in the current version of the model is hierarchically ordered. In particular, the contour dimension is considered as more general than both the diatonic and chromatic representations. No logical relations have been set between the diatonic and chromatic representations due to ambiguities of translation between the two representations: for instance, the augmented fourth and diminished fifth degrees of a diatonic scale have identical chromatic values. New musical dimensions can be added to the framework, provided that the logical dependences, if any, between the new dimensions and the previously defined representations are specified.

3. Specifications for melodic comparison

3.1. Fuzzy vs. exact matching

Once a range of musical parameters has been defined, the heuristics for motivic identification must be specified. The simplest strategy would consist of inferring identifications only when parameters of compared entities are strictly equal. An alternate strategy hypothesizes the existence of a large range of similarities that can be perceived between melodies, but that cannot be described through exact parametric identifications. A fuzzy definition of pattern matching can be used for that purpose: a numerical distance is defined, and a matching is made when the similarity distance is lower than a pre-specified threshold.

The fuzzy approach offers a way of avoiding the integration of musical dimensions that require extensive and complex computation. In particular, the diatonic pitch-interval dimension might be avoided by adopting a fuzzy approach along the chromatic pitch-interval dimension, since a one-semi-tone threshold theoretically allows the merging of major and minor intervals (Cambouropoulos et al., 2002; Cope, 1996). However, this threshold also tolerates other transformations that are not directly related to the major/minor configuration: for instance, a category that contains major third and perfect fourth intervals, but that excludes fifth intervals, cannot be easily explained using traditional musical concepts. More generally, the fuzzy approach can be considered as a clustering method that allows new identifications between musical entities. For instance, in the melodic dimension, the use of numerical similarity enables the identification of motives whose respective intervals are similar but not identical. The size and content of each cluster is highly determined by the value of the dissimilarity threshold. Yet no heuristic for precisely fixing this value has been proposed. Hence, the determination of the threshold value relies entirely on the user's intuitive choices.

Due to the difficulties created by the fuzzy approach, another solution consists of restricting more simply to exact matching along multiple musical dimensions (Conklin and Anagnostopoulou, 2001). For instance, concerning the melodic dimensions, patterns can be identified along their chromatic and diatonic pitch-interval, and contour dimensions. The computational model presented in this paper follows this exact matching heuristic.

3.2. Adaptive matching in a multi-parametric space

We propose a generalisation of the multiple viewpoint approach by allowing some *variability* in the set of musical dimensions used during the construction of each musical pattern. This enables us to take into consideration a more general type of pattern, called *heterogeneous patterns*, which despite their structural complexity seem to catch an important aspect of musical structure. An example of a heterogeneous pattern is the first theme of Mozart's *Sonata in A*, K. 331 (Fig. 2), which contains two phrases that repeat the same pattern. This pattern, enclosed in a solid box in the figure, is decomposed into two parts: a melodico-rhythmic antecedent, and a rhythmic consequent. The antecedent itself contains an exact repetition with transposition of a short cell (indicated with dotted lines). In the actual piece, the ending of the first phrase contains a little melodic ornamentation, that we have indicated in the score of Figure 2 by grace notes.

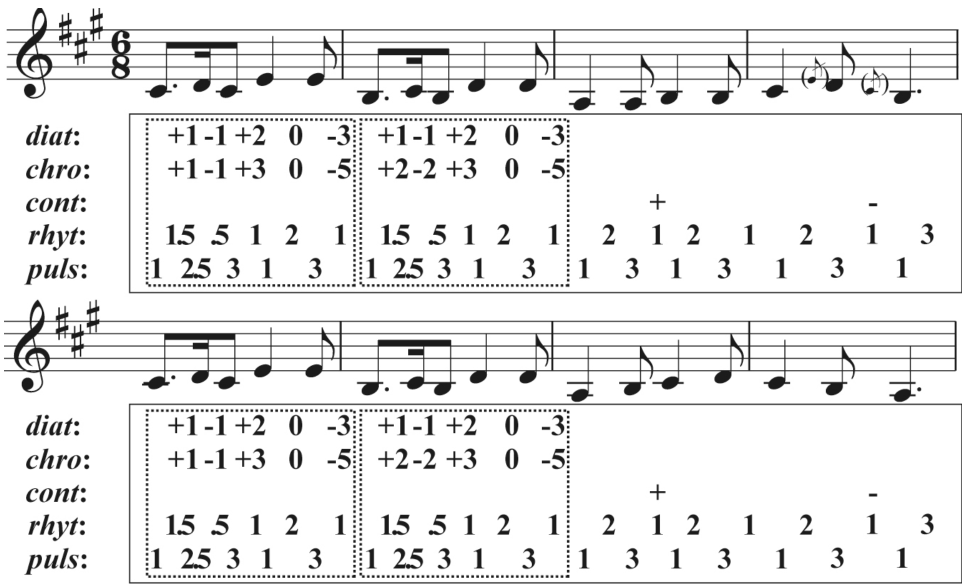

Figure 2.

Analysis of the first theme of Mozart's Sonata in A, K. 331, bars 1-8. The reduced melodic phrase in bar 4, where ornamentations are shown as grace notes, suggests a rhythmic similarity with bar 8.

Another example is the finale theme of Beethoven's *Ninth Symphony* (Fig. 3), which begins with an antecedent/consequent repetition of a phrase, with identities both in pitch and time domains, except a slight modification of the ending of each phrase.

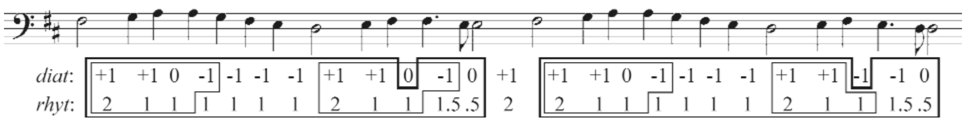

Figure 3.

Analysis of Ode to Joy, from Beethoven's Ninth Symphony.

Another famous example, that will be further discussed in the next section, is the four-note pattern of Beethoven's *Fifth Symphony*, shown in figure 4. This pattern is actually subdivided in our approach into a hierarchy of pattern descriptions of diverse levels of specificity. The most specific one is the complete melodico-rhythmic pattern *a* repeated twice at the very beginning of the Symphony. This specific pattern is progressively generalised during the piece through a disintegration of the different parameters constituting its description: the modification of the last descending third interval into a simple descending contour (pattern *b*), the modification of the third pitch (pattern *c*), the variation of the general contour pattern (pattern *d*), etc.

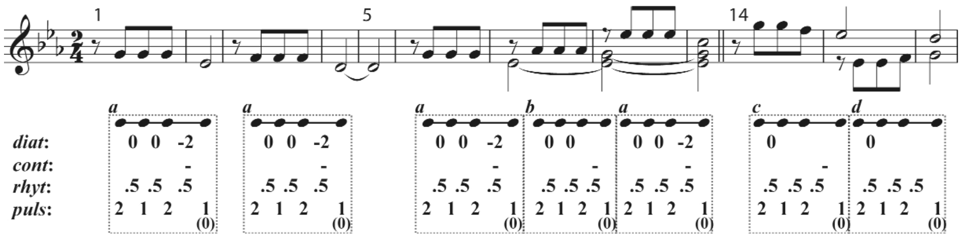

Figure 4.

The development of the famous four-note motive in the first movement of Beethoven's Fifth Symphony, in term of a succession of patterns of descending order of specificity: a, b, c and d. In an alternate and more precise description of the metrical dimension, the last note of each pattern is located on a downbeat represented by a 0 value.

3.3. A SOLUTION FOR THE CONTOUR PARADOX

As mentioned in section 2, due to the very limited degree of specificity of the gross contour parameter, patterns made of ascendant and descendant intervals are not easily recognised. It has been suggested, therefore, that repetition of gross contour sequences can be identified only when sufficiently close in time that, when the second occurrence is heard, the first one remains in short-term memory. Indeed, gross contour sequences can more easily be searched in short-term memory due to the limited size of this memory store, and availability of its content. On the other hand, a search in long-term memory seems cognitively implausible because of this store's large size, the resulting combinatory explosion of possible results, and the insufficient specificity of the query (cf. Dowling and Harwood, 1986). The contour dimension is all the more restricted to short-term memory since successive contour patterns are hardly perceived for long patterns (15 notes for instance) (Edworthy, 1985). However, this restriction leads to paradoxes (Dowling and Fujitani, 1971; Dowling and Harwood, 1986): if gross contour has no impact on long-term memory, how could the different occurrences of the familiar four-note theme throughout the first movement of Beethoven's *Fifth Symphony* (see for instance figure 4) actually be detected? One suggested explanation is that the numerous repetitions of the motive enable a memorisation of the contour pattern in long-term memory. Yet, could not this motive be detected, due to its intrinsic construction, even when repeated only a couple of times throughout the piece (as in figure 5)?

The heterogeneous pattern representation may offer an answer to this question, by enabling, as we saw in section 3.2, a decoupling of the choice of musical dimensions and the construction of patterns. A full understanding of the perceptive properties of motivic patterns requires a chronological view of the construction of these structures, in terms of an incremental concatenation of successive intervals. The dependency of such constructions upon long- and short-term memory may be understood in this incremental approach. More precisely, the initiation of a new occurrence of a pattern requires, as previously, a matching in long-term memory along interval dimensions. However, in this framework, it may be suggested that the further extensions of a discovered new occurrence do not require such a demanding computational effort: Once the first intervals have been initiated, the discovery of the progressive extensions simply requires a matching of the successive intervals with the corresponding successive intervals in the pattern. In other words, the proposed heuristic enables contour identification between temporally distant repetitions only when the contour value is related to the continuation of a pattern featuring more specific representations for the first intervals. This heuristic enables a selective filtering of non-salient patterns.

The four-note pattern of Beethoven's *Fifth Symphony*, figure 5, may be considered in this respect as a concatenation of two specific unison intervals (or three repetitions of the same note), followed by a less specific descending contour. Each new occurrence of the pattern can be easily perceived due to the high specificity of its three first notes (leading to an interval-based matching in long-term memory). The integration of these principles in the model enables a reconstitution of this phenomenon.

4. Pattern extraction

This section deals with the core problem of motivic extraction: modelling the mechanisms which ensure the discovery of repeated structures.

4.1. Related works

Designing robust algorithms for automated motivic analysis is a very difficult problem. Cambouropoulos (2006) searched for exact pattern repetition, using Crochemore's (1981) approach, in different parametric descriptions of musical sequences. The obtained large set of extracted patterns was not taken into consideration directly. Instead, an estimation of the segmentation points was computed through a weighted average of the segmentations implied by the different patterns.

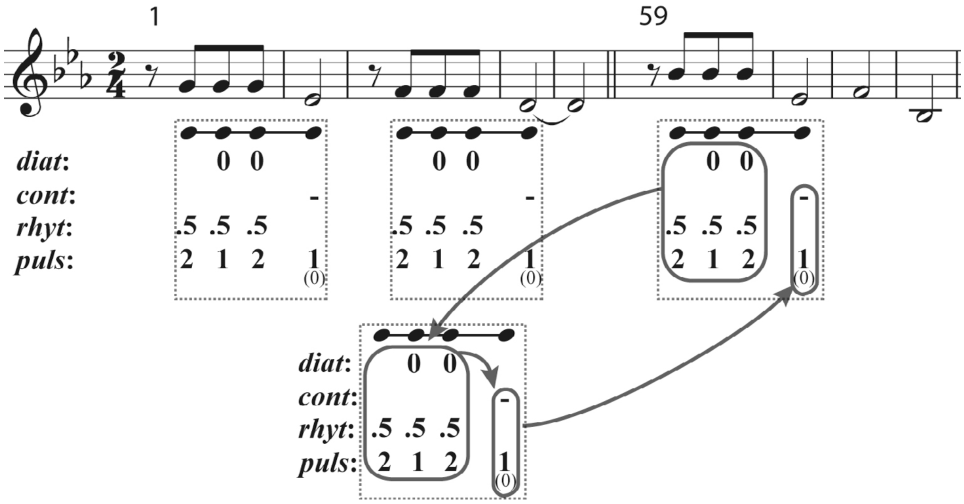

Figure 5.

The famous four-note pattern at the beginning of the first movement of Beethoven's Fifth Symphony is considered here as a concatenation of two unison intervals (diat = 0) and a decreasing contour, (cont = -) in a uniform eighth-note rhythm (rhyt = .5), which begins on an offbeat (puls = 2) and ends on a onbeat (puls = 1), or, more precisely, on a downbeat (puls = 0, in brackets). Any new instance of this pattern, such as the instance far later at bar 59 is detected following a two-step process: The specific description of the first two intervals triggers the matching, whereas the extension of the matching can follow a less specific contour description.

In Conklin & Anagnostopoulou (2001), pattern discovery was performed by building a suffix tree data structure along several parametric dimensions. Once again, due to the large size of the set of discovered patterns, a subsequent step selected patterns that occured in a specified minimum number of pieces, and that satisfied a statistical significance criterion. A further filtering step globally selected the longest significant patterns within the set of discovered patterns.

Rolland (1999) defined a numerical similarity distance between sub-sequences based on edit distance. In order to extract patterns, similarity distances were computed between all possible pairs of sub-sequences of a certain range of lengths, and only similarity exceeding a user-defined arbitrary threshold was selected. From the resulting similarity graph, patterns were extracted using a categorisation algorithm called *Star center*. The set of discovered patterns was reduced even further using offline filtering heuristics. In particular, only patterns repeated in a minimum number of musical sequences were selected.

Meredith, Lemström and Wiggins (2002) generalised the pattern extraction task to polyphony. Notes of musical sequences were represented by points in a two-dimensional (pitch/time) space, and maximal repetitions of point sets were searched. However, this geometrical strategy did not apply to melodic repetitions that presented rhythmic variations. Post-processing techniques were added that performed global selection in order to enhance the precision factor.

In all of these approaches, in order to reduce the combinatory explosion of the results obtained by the pattern extraction process, filtering heuristics are added that select a sub-class of the result based on global criteria such as pattern length, pattern frequency (within a piece or among different pieces), etc. The main limitation of this method comes from the lack of selectivity of these global criteria. Hence, by selecting longest patterns, one may discard short motives (such as the 4-note Beethoven pattern) that may nevertheless be considered as highly relevant for listeners. On the other hand, patterns repeated only twice may be considered as highly relevant by listeners, as long as these repetitions are sufficiently close in time that the first occurrence remains available in the short term memory when the second occurrence is heard[3].

The present study was primarily aimed at discovering the reasons for these failures, and at building as simple a model as possible that would be able to closely mimic the listeners' structural perception. We propose heuristics ensuring a compact representation of the pattern configurations without any loss of information, thanks to an adaptive and lossless selection of most specific descriptions.

4.2. Closed pattern mining

The problem of reducing the combinatorial complexity of pattern structure is also studied in current research in computer science, where several strategies have been tried. The *frequent pattern mining* approach is restricted to patterns that have a number of occurrences (or *support*) exceeding a given minimum threshold (Lin et al., 2002). We explained in the previous section the limitation of such heuristics for musical purposes.

Another approach is based on the search for *maximal patterns*, i.e. patterns that are not included in any other pattern (Zaki, 2005; Agrawal and Srikant, 1995). This heuristic enables a more selective filtering of the redundancy. For instance, in figure 6, the suffix *aij* can be immediately discarded following this strategy, since it is included in the longer pattern *abcde*: its properties can be directly induced from the long pattern itself. However, this approach still leads to an excessive filtering of important structures. For instance, in figure 7, the same 3-note pattern *aij* presents a specific configuration that cannot be directly deduced from the longer pattern *abcde*, for the simple reason that its *support* (or number of occurrences) is higher than the support of *abcde*. This corresponds to the concept of *closed patterns*, which are patterns whose support is higher than the support of the pattern in which they are included (Zaki, 2005). A filtering of non-closed patterns is therefore more selective than a filtering of non-maximal patterns. In fact, it ensures a more compact representation of the pattern configuration *without any loss of information*.

(3) This method of pattern segmentation through simple repetition corresponds to a common compositional strategy, in particular by Claude Debussy (Ruwet, 1962).

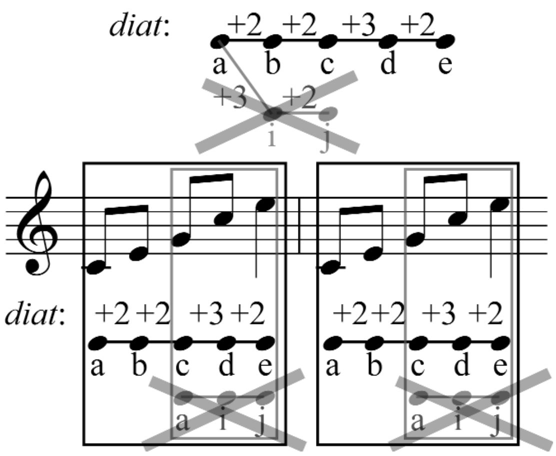

Figure 6.

Occurrences of patterns are extracted from the score. These occurrences are displayed below the score, and the patterns are grouped in a tree above the score. Pattern aij, a suffix of abcde with the same support, is therefore non-closed, and should not be explicitly represented.

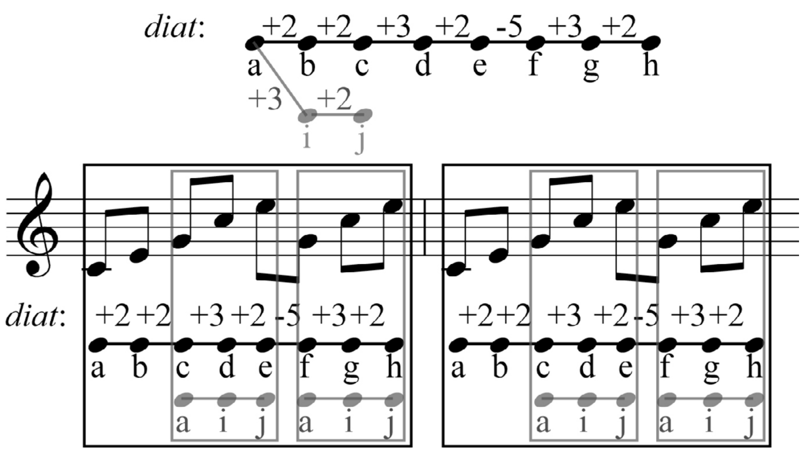

Figure 7.

Pattern aij, now featuring more occurrences than abcde or abcdefgh, is explicitly represented.

Figure 8 shows another illustration of the closed pattern paradigm. Pattern *a* is a maximal pattern, and therefore closed. Pattern *b* is included in pattern *a*, but has a larger support (four) than pattern *a* (two): it is therefore also a closed pattern. Pattern *c*, on the other hand, has a support equal to pattern *a* (two). Pattern *c* is therefore non-closed and should be discarded.

The model presented in this article looks for closed patterns in musical sequences. For this purpose, the notion of inclusion relation between patterns

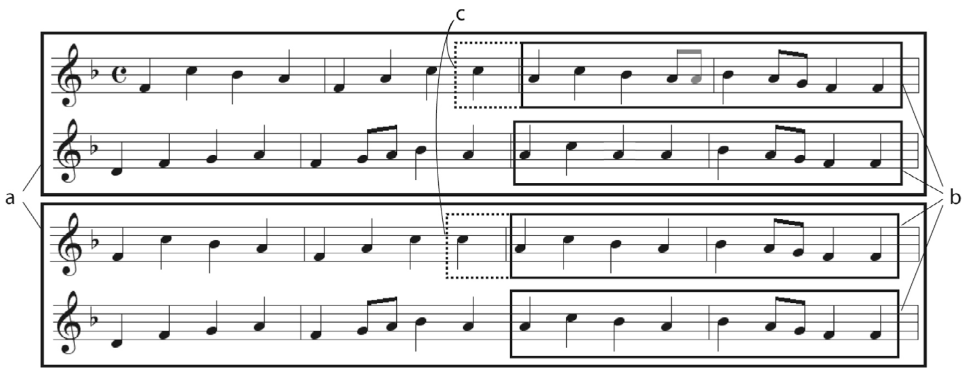

Figure 8.
Beginning of the Geisslerlied "Maria muoter reinû maît". A complete analysis will be presented in section 5.1. The little ornamentation displayed in grey is not taken into consideration. Patterns a and b are closed, whereas pattern c is non-closed and therefore discarded.

founding the definition of closed patterns is generalized to the multi-dimensional parametric space of music, defined in section 3.2. A mathematical description of this operation can be formalised using the Gallois correspondence between pattern classes and pattern description (Ganter and Wille, 1999; Lartillot, 2005a). For instance, pattern *abcde* (in Figure 9) features melodic and rhythmical descriptions, whereas pattern *afghi* only features the rhythmic part. Hence pattern *abcde* can be considered as more specific than pattern *afghi*, since its description contains more information. When only the first two occurrences are analyzed, both patterns have the same support, but only the more specific pattern *abcde* should be explicitly represented. But the less specific pattern *afghi* will be represented once the last occurrence is discovered, as it is not an occurrence of the more specific pattern *abcde*.

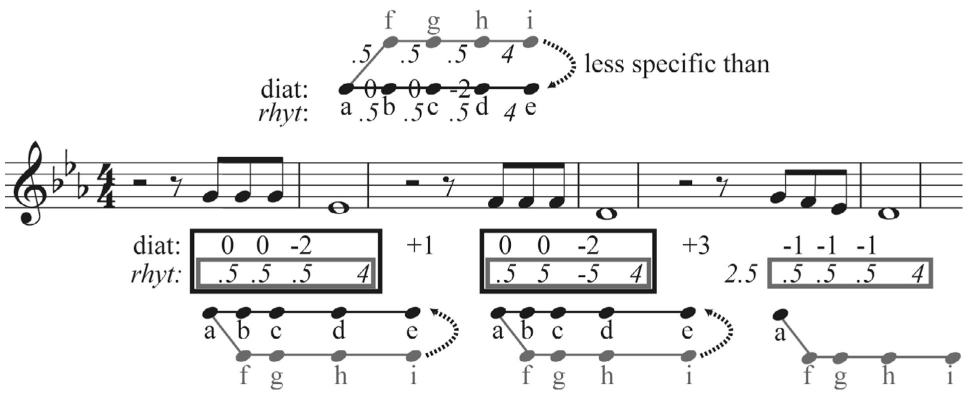

Figure 9.
The rhythmic pattern afghi is less specific than the melodico-rhythmic pattern abcde.

4.3. Cyclic patterns

Combinatorial explosions can be caused by another common phenomenon provoked by successive repetitions of a single pattern (for instance, in figure 10, the simple rhythmic pattern *abcd*, a succession of one quarter note and two eighth notes forming two ascending and one descending intervals). As each occurrence is followed by the beginning of a new occurrence, each pattern can be extended (leading to pattern *e*) by a new interval whose description (an ascending quarter-note interval) is identical to the description of the first interval of the same pattern (i.e., between states *a* and *b*). This extension can be prolonged recursively (into *f*, *g*, *h*, *i*, etc.), leading to a combinatorial explosion of patterns that are not perceived due to their complex intertwining (Cambouropoulos, 1998).

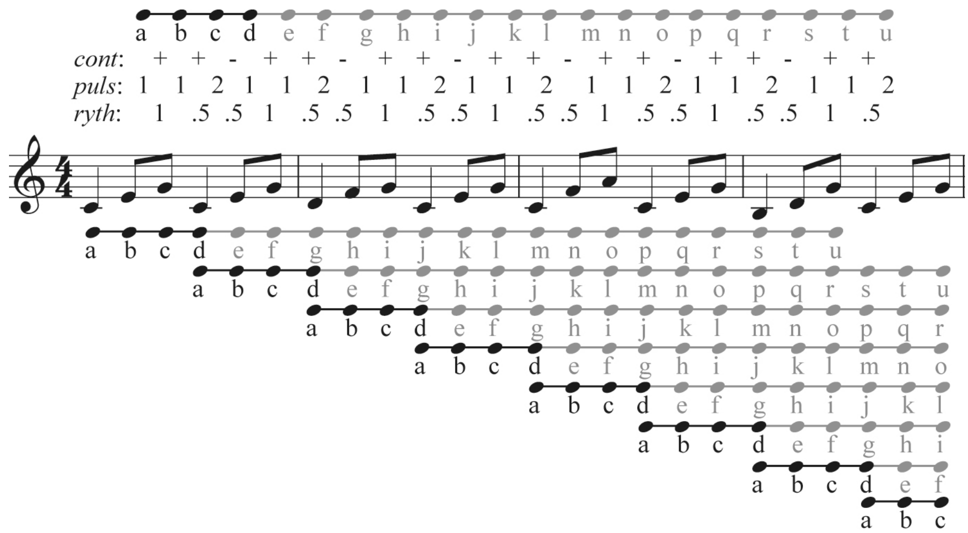

Figure 10.
Multiple successive repetitions of pattern abcd logically lead to extensions into patterns e, f, etc. which form a complex intertwining of structures.

The graph-based representation (Figure 10) shows that the last state of each occurrence of pattern *d* is synchronised with the first state of the following occurrence. Listeners tend to fuse these two states, and to perceive a loop from the last state (*d*) to the first state (*a*) (Figure 11). The initial acyclic pattern *d* leads, therefore, to a cyclic pattern that oscillates between three *phases b'*, *c'* and *d'*. Indeed, when listening to the remainder of the musical sequence, we actually perceive this progressive cycling. Hence this cycle-based modelling seems to explain a common listening strategy, and resolves the problem of combinatorial redundancy.

This cyclic pattern (with three phases *b'*, *c'* and *d'* at the top of Figure 11) is

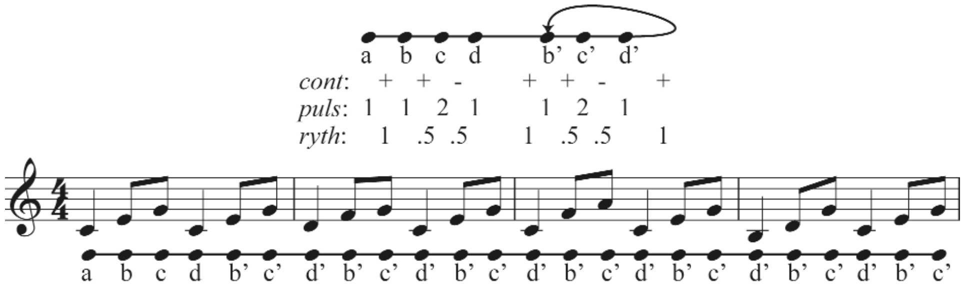

Figure 11.

Listening to successive repetitions of pattern abcd leads to the induction of its cyclicity, and thus to an oscillation between states b', c' and d'.

considered as a continuation of the original acyclic pattern *abcd*. Indeed, the first repetition of the rhythmic period is not perceived as a period as such but rather as a simple pattern: its successive notes are simply linked to the progressive states *a*, *b*, *c* and *d* of the acyclic pattern. On the contrary, the following notes extend the occurrence, which cannot be associated with the acyclic pattern anymore, and are therefore linked to the successive states of the cyclic pattern (*b'*, *c'* and *d'*). The whole periodic sequence is therefore represented as a single chain of states representing the traversal of the acyclic pattern followed by multiple rotation in the cyclic pattern.

This additional concept immediately solves the redundancy problem. Indeed, each type of redundant structure considered previously is a non-closed suffix of a prefix of the long and unique chain of states, and will therefore not be represented anymore. But this compact representation will be possible only if the initial period (corresponding to the acyclic pattern chain) is considered and extended before the other possible periods. This implies that scores need to be analysed in a chronological fashion.

Heterogeneous descriptions, as presented in section 2.2, can be associated with cyclic patterns too. For instance, in Figure 12, the cyclic pattern is a little more specific than the cyclic pattern presented in Figure 11, since the first note of each period is always C, and the interval between the second and third notes is always an ascending third. This can therefore be added to the representation of the pattern, as shown in the figure.

A mechanism has been added that unifies all the possible rotations of the periodic pattern (*b'c'd'*, *c'd'b'*, *d'b'c'*) into one single cyclic pattern. For instance, in Figure 13, the periodic sequence beginning in a different phase than previously (on an upbeat instead of a downbeat) is still identified with the same cyclic pattern.

By construction of the cyclic pattern, no segmentation is explicitly represented between successive repetitions. Indeed, the listener may be inclined to segment at any phase of the cyclic PC (or to not segment at all). Then it may be interesting to estimate the positions in the cycle where listeners would tend to segment. Several

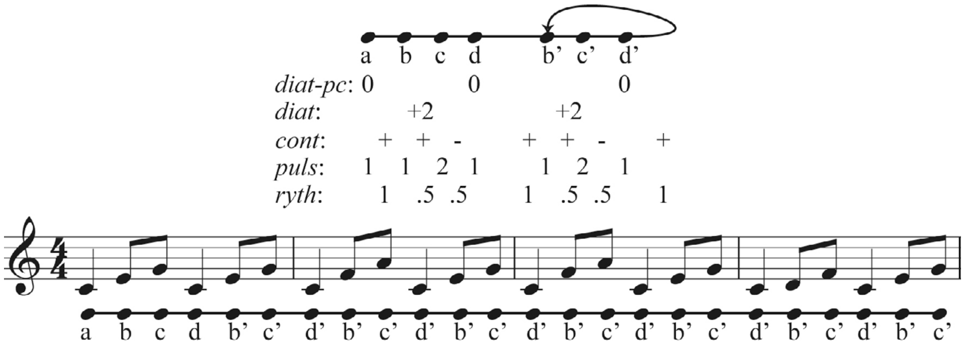

Figure 12.

Heterogeneous cyclic pattern, including two complete layers of rhythmic and contour descriptions, plus two local descriptions: the absolute pitch value C associated with the first note of each period (diat-pc = 0), and the constant pitch interval value of ascending major third between the second and the third note of each period (diat = + 2).

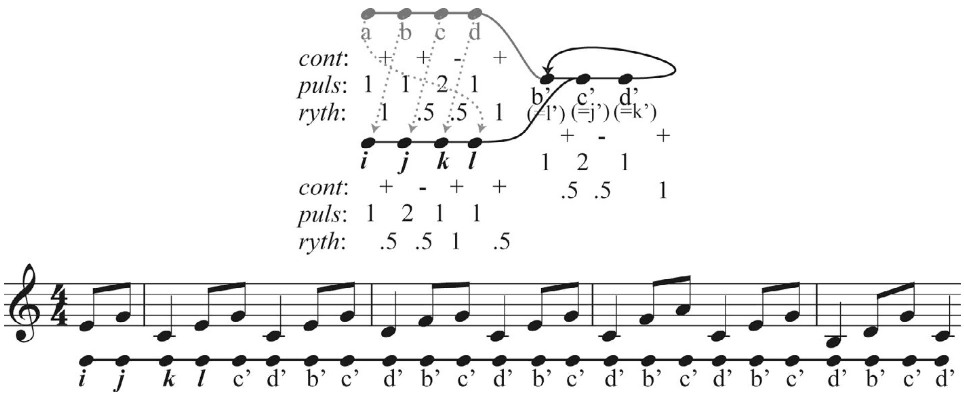

Figure 13.

The periodic sequence is initiated with a different phase, since it begins on an upbeat instead of the downbeat. Due to this rotation, the first period of the cycle is built from a new pattern, called ijkl, whose prefix ijk corresponds to the suffix bcd of the period of the initial cycle (abcd)[4]. The rotated cyclic sequence can be related to the same cyclic pattern (b'c'd', that can be also denoted j'k'l').

(4) Each pattern can accept multiple possible extensions forming a pattern tree (Lartillot, 2005b). Thus a suffix of a pattern (such as *ijk*, in our example, which corresponds to the suffix "*bcd*" of the pattern *abcd*) is not designated with the same letters that are used in the original pattern (*abcd*), because the multiple possible extensions of the new pattern might be different from those of the original pattern: patterns *ijk* and *abc* forms therefore two distinct branches of the total pattern tree related to the whole piece (Lartillot, 2005b).

factors need to be taken into consideration, such as primacy, local segmentation (as defined in the introduction), and global context. For instance, a primacy-based segmentation will favour the period that appears first in the sequence, which depends on the phase at which the cyclic sequence begins. Global context corresponds to the general segmentation of the piece, based on the major motives and the metrical structure. This will be considered in future work.

4.4. A COMPLEX SYSTEM

The general model is decomposed into modules dedicated to the different underlying problems, each of them further decomposed into basic building blocks focussing on specific sub-problems. All these blocks can easily be redesigned and articulated with each other in a flexible way, offering the possibility to test various hypotheses. The data representation itself has been designed with the view of offering maximum of flexibility in the choice of structure representation. The main principle of the methodology consists of progressively building the computational system through a careful design of each sub-module. At each progressive step of the construction, the general behaviour of the system is controlled, and unwanted behaviours are listed. The overall results of the system are improved by determining the reasons for each unwanted behaviour identified, and subsequently fixing these problems, either through the modification of sub-modules, or the creation of new ones.

These redundancy-filtering mechanisms ensure an optimal pattern description. Information is compressed without any loss, since all the discarded structures can be implicitly reconstructed. The filtering of redundant structures ensures clear results and at the same time decreases the combinatorial complexity of the process. Other rules have been integrated, based on cognitive heuristics. One rule in particular controls the combinatorial explosion that may be caused by the superposition of specific patterns on more general cyclic patterns, with the help of the *Gestalt* Figure/Ground principle (Lartillot, 2005b).

5. RESULTS

This model, called *kanthus*, will be included in the next version of *MIDItoolbox* (Eerola & Toiviainen, 2004). The model can analyse monodic pieces, and highlights the discovered patterns on a score. Rhythmic values are obtained through simple quantification operations, and scale degree parameters are computed through a straightforward mapping between pitches values and scale degrees. In the current state of the model, only repetition of patterns formed by series of strictly contiguous notes can be detected. The model has been tested using different musical sequences taken from several musical genres (classical music, pop, jazz, etc.), and featuring various level of complexity. The experiment has been run using version 0.8 of *kanthus*. This section presents the analysis of two pieces: a medieval *Geisslerlied* and a Bach *Invention*. The whole set of musical parameters defined in section 2 has been taken into account, and all patterns longer than 3 notes are displayed.

5.1. ANALYSIS OF A MEDIEVAL *GEISSLERLIED*

We first present an analysis of a forteenth-century German *Geisslerlied*, "Maria muoter reinû maît", proposed by the linguist Nicolas Ruwet (1966-1987), as a first application of his famous method of systematic motivic analysis. We will then show in comparison the results offered by the computational modelling.

5.1.1. Ruwet's analysis

Figure 14 presents Ruwet's analysis of the piece, which offers a hierarchical decomposition in three successive levels, enumerated from I to III. On the higher level, the piece features two repetitions (with slight variation) of a I-level unit *A* and two repetitions of another I-level unit *B*. Unit *A* is decomposed into two phrases[5]: the first phrase is further decomposed into two II-level units *a* and *b,* and the second phrase into two other II-level units *c* and *b'*, which is a slight variation of unit *b*. The second occurrence *A'* differs only by the fact that the two units *b* are identical. The second I-level unit *B* is another phrase formed by two II-level units: *d* and *b*'. Each II-level unit is decomposed into a succession of two III-level units. For instance, *a* is decomposed into two units *a1* and *a2*. One particularity of *d* is that its two III-level units are identical (*d1*), and *c* is decomposed into *c1* and *d1*. Moreover, *a1*, *b1*, *b1'*, *c1* and *a2* are considered as melodic transformations of a same rhythmical structure and another similarity is proposed between *b2* and *d1*. Finally, shorter units, composed of 2 to 5 notes, are suggested: three of them are shown in grey in the figure, another consists in two quarter notes forming a decreasing major third interval A-F, and a last one is composed of two quarter notes forming an increasing minor third interval. Ruwet's analysis concludes with a modal analysis.

Ruwet's methodology consists of a mostly top-down hierarchical segmentation of the piece: first, the two repetitions of *B*, being exactly identical, are discovered; then the leftover (bars 1 to 16) is segmented, leading to the extraction of the two units *A* - *A'*. This strategy would not have worked if there were slight variations between the two occurrences of pattern B*,* for instance: *B* = *d* + *b* and *B'* = *d* + *b'*. It can be shown that a systematic application of the methodology can produce, for the same musical piece, alternative analyses that are contradictory and counter-intuitive (Lartillot, 2004a). Hence Ruwet's analysis of the *Geisslerlied* is not strictly and uniquely guided by the systematic methodology he introduces in this paper, but is rather deeply influenced by his implicit intuitions.

(5) The decomposition in two phrases is not explicitly stated in Ruwet's representation.

Figure 14.

Analysis of the Geisslerlied "Maria muoter reinû maît" following Nicolas Ruwet's approach (1966-1987). Although no time-signature was given in the original transcription of the medieval piece, bar lines have been added here. The representation of the analysis follows the same convention that will be used for the computational results in Figure 15, in order to facilitate the comparison between the two methods. Each one- or two-bar-long unit is represented by a line below the corresponding stave, and labelled with a letter. Each four- or eight-bar-long unit is represented by a line on the left of the stave, also labelled with a letter.

5.1.2. Computational analysis of the piece

A complete motivic analysis of the *Geisslerlied* has been carried out with the computational model. Figure 15 shows the result of the analysis[6].

Unit *A* has been retrieved by the computer due to its repetition. However, as the current model cannot take into consideration ornamentations, the eighth-note repetition of bar 3, displayed in grey in the figure, had to be removed from the score, as it prevents the detection of the complete pattern *A*[7]. On the other hand, the model is able to take into account the slight variation concerning the varying pitch

(6) The computational results have been filtered manually, as explained in the end of section 5.1.2.

(7) The consideration of ornamentation is the object of current work (see section 6.2).

value, which is an A in the first repetition, and a Bb in the second repetition. Pattern *A* is heterogeneous in that it is described along all the musical dimensions of the parametric space, except in regard to the varied note which is described by the gross contour parameter, only for the preceding interval, and by the rhythmic parameters[8].

Pattern *b*, which corresponded in Ruwet's analysis to the identical ending of each line of the score, has been extracted, too. The aforementioned pitch variation, since repeated also several times (here, three times), is described by another pattern *b'*. Both pattern classes *b* and *b'* can be unified into a more general pattern that contains the two possible variations of the endings, and, as a consequence, leaves the variable note undescribed, similarly to the description of pattern *a*. Patterns *a* and *c*, on the contrary, are not explicitly represented, because they do not convey additional information concerning the pattern structure of the piece. Following the terminology introduced in section 4.2, patterns *a* and *c* are non-closed subsequences of pattern *A*. In Ruwet's analysis, the selection of these patterns is based on a segmentation process: patterns *a* and *c* are the leftovers after the extraction of the endings *b* from the bigger phrase *A*. No segmentation process has been integrated into the model yet.

Pattern B has not been correctly extracted by the system. This is due to the fact that pattern *A* itself is concluded by a suffix of *B* (shown by the *B'* line in the figure). Following the incremental approach of the algorithm, the two complete repetitions of the *B'* pattern are first discovered, leading to a cyclic pattern whose starting points are indicated by the *B'* letters in the figure. The inference of the *B* segmentation proposed by Ruwet would require the incorporation of new mechanisms, as explained in section 4.3. In the example, the initial cyclic pattern *A* implies a segmentation at the beginning of the third stave and also at the beginning of the fifth stave. We may suppose that the prolongation of this first segmentation would be expected by listeners. This expectation is reinforced by the repetition of pattern *b* in the last two staves, which seems to induce a generalisation of cycle *A*, that should be studied in future works. Pattern *d*, like patterns *a* and *e*, is not explicitly represented in the computational analysis since it is a redundant subsequence of pattern *B* (and *B'*). Its extraction would thus require segmentation heuristics. Among the III-level units, the pattern extraction algorithm can only discover unit *d1*, due to its intrinsic repetition. Other III-level units resulted either from segmentation processes — *c1* is the leftover after the extraction of *d1* from the unit *c* — or for purely symmetrical reasons,

(8) The successive repetition of pattern *A* leads to the creation of a cyclic pattern, each cycle being the repetition of a new occurrence. The cyclic pattern implies the expectation of a third occurrence (indicated by the third *A* graduation) finally aborted. As explained in section 4.3, the concept of cyclic pattern enables to avoid the extension of each occurrence, which would lead to an overlapping of the occurrence and to a combinatorial explosion of structures. For instance, the pattern obtained by shifting of pattern *A* to the right by one note is filtered out as it is a non-closed suffix of one phase of the cyclic pattern *A*. Indeed, the support of this candidate pattern is not higher than the support of the corresponding phase in the original cyclic pattern.

relatively to the relative size of each unit, and cannot therefore be detected by the algorithm.

Among the shortest units proposed by Ruwet, the three-note conjunct lines (grey arrows in the figure) are formalized as cyclic successions of second intervals. On the other hand, the two other units displayed in grey in the figure cannot be detected because they are repeated through retrogradation, which is a musical transformation not yet taken into account in the model. The two last units proposed by Ruwet are only composed of two notes. But as a huge number of interval repetitions can be found in any musical piece, the selection of a particular interval requires further justifications, not given by Ruwet for these particular structures. On the other hand, short patterns, such as *e* and *f*, are proposed by the algorithm, that have no correspondence with Ruwet's analysis. The assessment of their perceptual or musical relevance will require further study.

Figure 15.

Analysis of the Geisslerlied "Maria muoter reinû maît" using our approach. Each motivic repetition is represented by a line below the corresponding stave, each labelled with a letter. Each repetition of motive A is represented by a line on the left of the stave.

The analysis shown in Figure 15 results from a manual filtering of the output of the computational analysis:

• The output of the algorithm catalogues the progressive construction of patterns during the incremental and chronological scanning of the score. This trace contains, therefore, much redundancy since it shows for each pattern the list of successive extensions. In particular, prefixes of patterns are not discarded, even if they are non-closed, because they form the successive states of the chronological construction of the pattern (Lartillot, 2005a). Figure 15, on the contrary, shows the final state of the analysis, which simply consists of the set of all the motives that have been discovered. The transformation of the chronological analysis into this compact list is carried out manually for the moment.

• Some evident motivic structures have not been shown in the score: for instance, the simple succession of eighth notes, or the successive repetition of a same gross contour value.

• The mechanism based on the *Gestalt* rule of figure against ground, mentioned at the end of section 4, enables a filtering of a large set of redundant structures: it prevents each pattern (such as *b*) from being extended by a simple rhythmic succession of quarter-notes if this succession already existed before the pattern: pattern b is considered as a figure above the background formed by the succession of quarter-notes. This rule does not currently work when the pattern is preceded by a succession of eighth notes instead of quarter notes, but would work if the model were able to infer the implicit succession of quarter notes hiding underneath. This will be implemented in future work.

• The alternation of series of quarter notes and series of eighth notes should be formalized as a cycle (the alternation) between two cycles (the series). This concept of cycle of cycles is, however, not yet implemented in the model.

Hence the computational analysis reveals for the most part trivial motivic structures that can, in most cases, be easily perceived by listeners. The interest of this model, in its current state, is in the automation of the process, which enables an exhaustive analysis. Moreover, these results show that the model is able to offer a compact and significantly relevant description of musical structures. The refinement of the results of the computational analyses may now be planned through an enrichment of the modelling process.

5.2. BACH *INVENTION* IN *D* MINOR

5.2.1. Kresky's analysis

Jeffrey Kresky proposed a detailed analysis of Bach *Invention* in *D* minor (Kresky, 1977). His approach is founded on a close interconnection between the tonal and motivic dimensions of music. As our study focuses solely on motivic analysis, the large part of Kresky's analysis related to the tonal evolution of the *Invention* is not considered in this review, the scope of which is further restricted to the first 15 bars of the piece. Figure 16 presents a tentative explicit reconstruction of the analysis, that

Figure 16.
Reconstitution of Kresky's analysis of Bach Invention in D minor (Kresky, 1977).

Kresky originally presented in a mostly textual manner, without explicit categorization of the motivic classes.

The piece begins with an "opening motive" (here *A*), composed of an "ascending line" (indicated by a grey arrow in the figure) followed by a decreasing interval Bb - C#, forming altogether a motive shape (*a*). This motive shape "reverses itself in the second measure" of motive *A* (motive *a1*). A "direct imitation of the motive" (*A*) "first appears in the bass part" at bars 3 and 4, followed by an "octave-higher repeat" in the treble part at bars 5 and 6.

During the repetitions of the opening motive in bars 3 to 6, the alternate voice presents an "accompanying figure" (denoted *B* in the figure), containing two repetitions of a variant of motive shape *a* (denoted *a2*). This "shape directly clarifies the hidden triadic structure of the first measure (by isolating the triad notes and eliminating the passing tones)", and "the shape of the accompaniment is just that of the motive — an ascent spanning a sixth". More precisely, most occurrences of motive *a2* picks up the whole pattern of the motive shape *a*, "as an ascent through a sixth followed by a large skip in the opposite direction".

From measure 7 to measure 10 is built a first sequence, whose unit (denoted *A'*) here also can be considered as a variation of the opening motive *A*: the second half of *A'* is identical to the ending of *A* (*a1*), whereas the first half of *A'* (denoted *a3*) is a variation of the first half of *A* (*a*) as the scalar climb "starts one attack late, a minor third having been added before it"[9]. Another matching between *a* and *a3* comes from the fact that they both feature "the same sequence of pitch classes (D-E-F-G-A-Bb)" but "reshaped".

From measure 11 follows a second sequence, whose unit (denoted *a5*) "also derives from the original motive shape [*a*], for the rising six-note scale [grey arrays] is here represented by two adjacent three-note scales" (denoted *b*), which "ends with the skip in the opposite direction". The three-note scale is formed by the "first three notes of the original scale form" (denoted *b*), and the octave position of *b* in measure 11 is the same than in measure 5 (both denoted *b'* for this reason).

The first sequenced unit, when played at the treble voice between measures 7 and 10, is accompanied by two successive similar "bass lines" (denoted *a4*) that also shows close similarity with the motive shape *a*, since it contains the same three-note shape *b"* in the second half of the line, and the whole line can be identified with the six-note scale (grey array). We may notice that this scale here also ends with a large skip in the opposite direction (dotted line under the score), and that the six-note scale is prolonged in the next measure (dotted arrow).

Finally, the pitches found at the downbeat of each first two bars of the sequence between measures 7 and 14 forms a descending scale F-E-D-C (circled in the figure) that is prolonged in the next measures of the piece and that plays an important tonal role.

Hence Kresky's analysis precisely shows how "the opening motive [*a*], through various imitative means, seems to generate the entire fabric of the invention" (p. 63).

5.2.2. Computational analysis

Figure 17 shows the results of the computational analysis of the first 14 bars of the *Invention*, and Table 1 lists the set of discovered motives. The opening motive *A*, with

(9) Since this minor third interval can be found after the end of each occurrence of motive *A*, Kresky considers it as an extension of motive *A* (represented in dotted lines in the figure).

all its occurrences throughout the piece, has been exactly retrieved by the machine. The ascending and descending lines within the motive (grey arrows) have been detected, too. On the other hand, the inversion *a1* of the motive shape *a* has not been detected since inversions are not taken into account in the model yet[10]. As the opening motive *A*, throughout the analyzed extract, is not transposed in the pitch-class domain, it can therefore be expressed in this dimension, as indicated in Table 1 under column titled "Melodic"[11].

The accompanying figure *B* has been detected too, and is decomposed into a succession of two successive and similar shapes. They are noted *c* instead of *a2*, however, because the similarity of these shapes with motive *a* has not been discovered. Indeed, the model is neither able to automatically retrieve the hidden triadic structure in motive *a*, nor reduce the ascending figure to an ascending sixth interval[12].

The first sequence unit *A'* is detected and identified as a variation of the opening motive *A*, but not exactly for the same reasons: the scalar climb *a* has not been detected as delayed of one sixteenth-note, but rather modified on its first interval (motive *a3*). Contrary to the opening motive *A*, the sequence unit *A'* is constantly transposed, and is therefore expressed along the pitch-interval dimension instead of the absolute-pitch description.

The second sequence unit a5 has been detected and identified as a variation of the original motive shape *a*. The repeated three-note scale *b* is also identified with the first three notes of the original scale form at the beginning of the motive shape *a*.

The repeated bass line during the first sequence is detected, but cannot be identified with the motive shape *a*, therefore denoted *C* instead of *a4*. Indeed, in Kresky's analysis the three-note shape *b'''* is extracted from the bass line following a heuristic, not implemented in our framework, based on a segmentation induced by the metrical structure. Besides the six-note scale (grey array in figure) has not been identified either due to the irregularity of the second note.

Finally, the descending scale F-E-D-C (circled in the figure) has not been detected due to the incapacity of the algorithm to consider motivic configurations between distant notes.

The computational analysis also includes additional motivic structures that do not seem to offer much musical interest or perceptive relevance. Pattern *D*, featuring

(10) The little extension of motive *A* with the following third interval (dotted lines in figure 16), due to the variability of its contour, is not detected either.

(11) The motive spans the whole range of one single octave (from C# to Bb), therefore no indication of relative octave position is given here.

(12) In the two occurrences of the accompanying figure, the first shape is built on the same series of pitch class (F, A, D), where D is one octave higher, therefore noted in the table (F, A, D +). On the other hand, the second shape is variable, and can be described only in the contour dimension as (+, +, -). The whole melodic description of the accompanying figure is thus (F, A, D +, -, +, +, -).

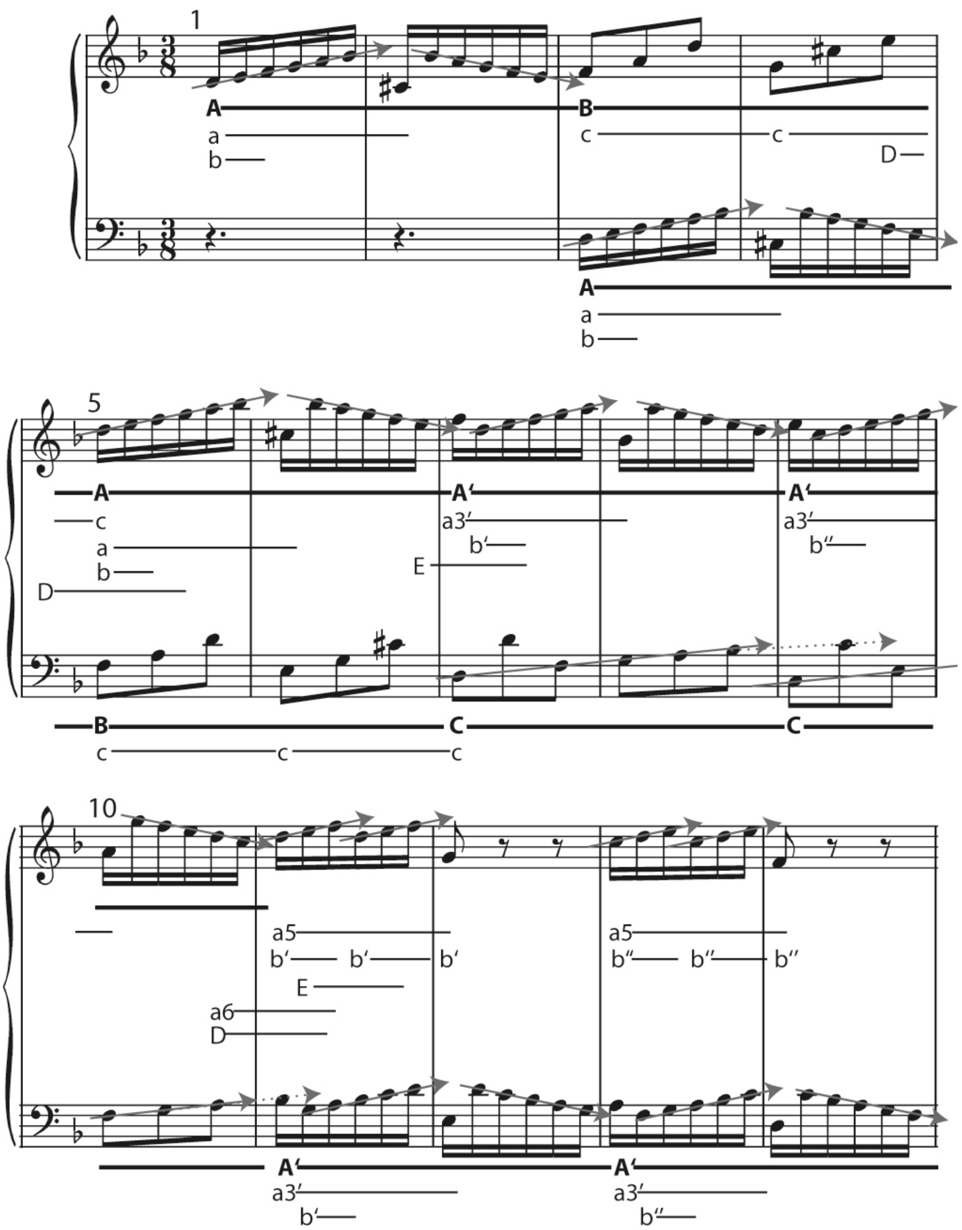

Figure 17.

Automated motivic analysis of J.S. Bach's Invention in D minor BWV 775, first 14 bars. The representation of the motives follows the convention adopted in the previous figures. The motives for each voice are shown below the respective staves. The class of each motive is indicated on the left side of the lines.

a succession of a descending second interval followed by a series of ascending second intervals, sounds poorly salient due to its weak position in the metrical structure, and the limited size of its description. The second occurrence of pattern *D* is also considered by the model as the beginning of a variant *a6* of motive *a3*, which is, once again, not very salient due to the weak position of this motive in the metrical structure. Pattern *E* shows similar limitations: it corresponds to a series of pitches (E5, F5, D5, E5, F5) starting on the offbeat. In order to filter patterns *D* and *E*, a higher-level metrical representation may be integrated in future works that would show the position of each pattern in the bar structure, and would, in particular cases, force pattern occurrences to conserve the same metrical information. The model also includes evident structural configurations, such as the oscillations between ascending and descending contour, and other patterns that are mostly redundant descriptions incorrectly filtered. This results partly from unsolved errors in the modelling process, but also shows the necessity of taking into account other heuristics.

Finally, in order to obtain the compact representation displayed in Figure 15, some manual ordering of the computational results was required. For more results, and discussions about the complexity of the algorithm, see Lartillot (2005b).

Thus, the computational system has been able to retrieve the most salient structures of the piece, which are congruent with Kresky's analysis. However, the subtlest configurations discovered by the musicologist cannot be detected with the algorithm. Indeed, the use of computers here is not aimed at replacing the musicologist's skills, but rather at experimenting with a formalisation of the basic principles of music

Table 1

Motives discovered in Bach's *Invention in D minor* BWV 775 both in Kresky's analysis and in the computational experiment. For each motive, Kresky's interpretation and the computational descriptions are given

Name	**Kresky's interpretation**	**Computational description**	
		Melodic	**Rhythmic**
A	Opening motive	(D,E,F,G,A,Bb,C#,Bb,A,G,F,E,F)	16th notes
A'	First sequence unit	(-2,+1,+1,+1,+1,-6,+6,-1,-1,-1,-1,+1)	16th notes
B	Accompanying figure	(F,A,D+,-,+,+,-)	8th notes
a	Motive shape	(prefix of A)	
a1	Reversed motive shape	(not detected)	
a2=c	Triadic structure	(+,+,-)	8th notes
a3=a3'	Retarded scalar climb	(-2,+1,+1,+1,+1,-6)	16th notes
a4=C	Bass line	(+7,-5,+1,+1,+1,-6)	8th notes
a5	Seconde sequence unit	(+1,+1,-2,+1,+1,-6)	16th notes
a6	*(not considered)*	*(-,+1,+1,+1)*	*16th notes*
b	Three-note scale	(+1,+1)	16th notes
b'	Identifying a5 with a3	(D5,E5,F5)	16th notes
b''	Same, transposed	(C5,D5,E5)	16th notes
b'''	Three shape in bass line	(not detected)	
D	*(not considered)*	*(-1,+1,+1,+1)*	*16th except 1st*
E	*(not considered)*	*(E5,F5,D5,E5,F5)*	*16th start. offbeat*

understanding. Besides, this automation may enable the annotation of large music databases.

6. Future work

6.1. More Detailed Validation of the Results

The current state of this research project is restricted to the most evident structures of music, as explained in the introduction. The analyses produced by the computational model are evaluated for the moment in a purely qualitative manner: the results are searched for the most important motives of the piece, and the additional unexpected structures offered by the algorithm have been assessed in a purely intuitive manner.

In future work, the computational results will be compared with analyses available in the music literature. A more precise and refined validation of the results will require the establishment of a "ground truth": a corpus of pieces of diverse style will be collected, on which manual motivic analyses will be carried out by a board of musicologists. The comparison of the manual analyses and the computational results will enable a more precise determination of the precision and recall factors offered by the model.

6.2. Musical transformations

One major limitation of the first version of the model, as presented in this article, is that only repetition of sequences of notes that are immediately successive could be detected. In music in general, repeated patterns are often ornamented: secondary notes can be added in the neighbourhood of the primary notes, in both time and pitch dimensions. Primary notes may be retrieved through an automated filtering of secondary notes, for instance, by focusing primarily on notes at metrically strong positions (Eerola, Järvinen, Louhivuori, Toiviainen, 2001; Conklin and Anagnostopoulou, 2001, etc.). This heuristic does not, however, work correctly for appoggiaturas (where ornamenting secondary notes are placed on metrically strong positions). Other approaches are based on optimal alignments between approximate repetitions using dynamic programming and edit distances (Rolland, 1999). We are developing algorithms that automatically discover, from the rough surface level of musical sequences, musical transformations revealing the sequence of pivotal notes forming the deep structure of these sequences.

As for absolute pitches, absolute representations of temporal positions play only a minor role in motivic analysis: motives more generally result from local temporal configurations. Many studies in computer music (Conklin and Anagnostopoulou, 2001; Cambouropoulos, 2006; Lartillot, 2004b, etc.) integrate an additional temporal dimension based on *inter-onset interval ratios*, i.e., ratios between successive inter-onsets intervals. This ratio would enable the detection of augmented or diminished rhythmic motives, i.e. motives that are globally dilated or contracted in time.

Another strategy consists of modelling metre as a multi-levelled hierarchy of beat-levels on the score, by which the rhythmic information is described. Hence, augmented or diminished repetitions can be detected as repetitions of the same rhythmic sequence on several different levels at several different positions in the score. This strategy is more satisfying since it restricts the domain of augmentations and diminutions to the different possible metrical subdivisions, a restriction that seems to correspond more closely to the listener's capacities. These multi-level rhythmic layers will be considered in future work.

6.3. Polyphony

Our approach is limited to the detection of repeated monodic patterns. Music in general is polyphonic, where simultaneous notes form chords and parallel voices. Research has been carried out in this domain (Dovey, 2001; Meredith et al., 2002), focusing on the discovery of exact repetitions along different separate dimensions. We plan to generalize our approach to polyphony following the syntagmatic graph principle. We are developing algorithms that construct, from polyphonies, syntagmatic chains representing distinct monodic streams. These chains may be intertwined, forming complex graphs along which the pattern discovery algorithm will be applied. The additional factors of combinatorial explosion resulting from this generalized framework will require further adaptive filtering mechanisms. Patterns of chords may also be considered in future works.

6.4. Articulating global patterns and local discontinuities

As explained in the introduction, this study is focused on the discovery of repeated patterns, and does not take into account the second heuristic of motivic analysis based on local discontinuities. Yet local discontinuities impose some important constraints on the pattern extraction process. The pattern extraction process needs, therefore, to be studied in interaction with local segmentation. Lerdahl and Jackendoff (1983) have proposed a coupling of the two principles, and Temperley (1988) has suggested a computational formalisation. But in these approaches, as acknowledged by their authors, pattern extraction (called here "parallelism") is theoretically considered without actual systematic modelling. Cambouropoulos (2006) proposed a way of modelling the interaction between the two principles. In a first step, both local segmentation and pattern extraction are performed in parallel, but only the boundaries of the segments and motives are taken into consideration. These boundaries are summed together, leading to a segmentation curve, and global segmentations are performed at local maxima of the curve. The resulting segments are then classified based on similarity measurement, following the paradigmatic analysis approach (Cambouropoulos & Widmer 2000). The description of the local discontinuities will be integrated in our model as a separate layer of description. The interactions between pattern extraction and local discontinuities, and particularly the constraints imposed by local discontinuities to pattern extraction, will be studied.

ACKNOWLEDGMENTS

This work has been supported by the Academy of Finland (Project No. 102253). The authors would like to acknowledge the valuable insights, observations and suggestions contributed by Emilios Cambouropoulos and David Meredith, and express their gratitude to Geoff Luck for his useful help during the writing of the paper.

Address for correspondence:
Olivier Lartillot
Petri Toiviainen
University of Jyväskylä
Department of Music
PL 35(M)
40014 University of Jyväskylä, Finland
tel.: + 358 14 260 1346
fax: +358 14 260 1331
e-mail: lartillo@campus.jyu.fi

• References

Agrawal, R. & Srikant, R. (1995). Mining sequential patterns. *International Conference on Data Engineering.*

Ahlbäck, S. (2007). Melodic similarity as determinant of melody structure. *Musicæ Scientiæ, Discussion Forum 4a.*

Cambouropoulos, E. (1998). *Towards a general computational theory of musical structure.* (Doctoral dissertation, University of Edinburgh).

Cambouropoulos, E., & G. Widmer (2000). Automated motivic analysis via melodic clustering. *Journal of New Music Research*, *29 (4)*, 347-70.

Cambouropoulos, E., M. Crochemore, C. Iliopoulos, L. Mouchard, & Y. Pinzon. (2002). Algorithms for computing approximate repetitions in musical sequences. *Journal of Computer Mathematics*, *79 (11)*, 1135-48.

Cambouropoulos, E. (2003). Pitch spelling: A computational model. *Music Perception*, *20 (4)*, 411-29.

Cambouropoulos, E. (2006). Musical parallelism and melodic segmentation: A computational approach. *Music Perception*, *23 (3)*, 249-68.

Cemgil, A.T., & B. Kappen. (2003). Monte Carlo methods for tempo tracking and rhythm quantization. *Journal of Artificial Intelligence Research*, *18*, 45-81.

Chew, E., & Y.C. Chen. (2005). Real-time pitch spelling using the spiral array. *Computer Music Journal*, *29 (2)*, 61-76.

Conklin, D., & C. Anagnostopoulou. (2001). Representation and discovery of multiple viewpoint patterns. *International Computer Music Conference.*

Cope, D. (1996). *Experiments in Musical Intelligence.* Computer Music and Digital Audio Series. Volume 12, A-R Editions.

Crochemore, M. (1981). An optimal algorithm for computing the repetitions in a word. *Information Processing Letters*, *12 (3)*, 244-250.

Dannenberg, R., & Mont-Reynaud B. (1987). Following an improvisation in real-time. *International Computer Music Conference.*

Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl and Jackendoff's grouping preference rules. *Music Perception*, *4 (4)*, 325-50.

Desain, P., & Honing, H. (1991). Quantization of musical time: a connectionist approach. In Todd, P.M., & Loy, D.G. (eds), *Music and connectionism*, pp. 150-67. Cambridge, MA: MIT Press.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics*, *11*, 411-12.

Dovey, M.J. (2001). A technique for "regular expression" style searching in polyphonic music. *International Conference on Music Information Retrieval*, Bloomington.

Dowling, W.J., & D.S. Fujitani. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, *49*, 524-31.

Dowling, W.J., & D.L. Harwood. (1986). *Music cognition.* San Diego: Academic Press.

Eck, D. & Casagrande, N. (2005). Finding meter in music using an autocorrelation phase matrix and Shannon entropy. *Proceedings of the International Conference on Music Information Retrieval*, London.

Edworthy, J. (1985). Melodic contour and musical structure. In P. Howell, I. Cross and R. West, *Musical structure and cognition.* (chap. 7). San Diego: Academic Press.

Eerola, T., T. Järvinen, J. Louhivuori, & P. Toiviainen. (2001). Statistical features and perceived similarity of folk melodies. *Music Perception*, *18*, 275-96.

Eerola, T., & Toiviainen, P. (2004). MIR In Matlab: The MIDI Toolbox. *International Conference on Music Information Retrieval*, Barcelona, Spain.

Forte, A. (1977). *Structure of atonal music*. New Haven, NJ: Yale University Press.

Ganter, B. & Wille, R. (1999). *Formal concept analysis: Mathematical foundations*, Berlin: SpringerVerlag.

Kresky, J. (1977) *Tonal music: Twelve analytic studies*, Philadelphia: Indiana University Press.

Large, E.W. & Kolen, J.F. (1994). Resonance and the perception of musical meter, *Connection Science 6 (2/3)*, 177-208.

Lartillot, O. (2004a). *Fondements d'un système d'analyse musicale suivant une modélisation cognitiviste de l'écoute*. (Doctoral dissertation, University of Paris 6).

Lartillot, O. (2004b). A musical pattern discovery system founded on a modeling of listening strategies. *Computer Music Journal*, *28 (3)*, 53-67.

Lartillot, O. (2005a). Efficient extraction of closed motivic patterns in multi-dimensional symbolic representations of music. *Proceedings of the International Conference on Music Information Retrieval*.

Lartillot, O. (2005b). Multi-dimensional motivic pattern extraction founded on adaptive redundancy filtering. *Journal of New Music Research*, *34 (4)*, 375-93.

Lerdahl, F., & Jackendoff R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lin J., Keogh, J, Lonardi S. & Patel P. (2002). Finding motifs in time series, *International Conference on Knowledge Discovery and Data Mining*

Meredith, D., Lemström K., & Wiggins G. (2002). Algorithms for discovering repeated patterns in multidimensional representations of polyphonic music. *Journal of New Music Research*, *31 (4),* 321-45.

Meredith, D. (2006). The *ps13* pitch spelling algorithm. *Journal of New Music Research*, *35 (2)*, 121-59.

Nattiez, J.-J. (1975): *Fondements d'une sémiologie de la musique*. Paris: Union Générale d'Éditions.

Nattiez, J.-J. (1990). *Music and discourse: Towards a semiology of music*. Princeton: Princeton University Press.

Povel, D.J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411-40.

Reti, R. (1951). *The thematic process in music*. New York: Macmillan.

Rolland, P.-Y. (1999). Discovering patterns in musical sequences, *Journal of New Music Research*, *28 (4)*, 334-50.

Ruwet, N. (1962). Note sur les duplications dans l'œuvre de Claude Debussy. *Revue Belge de Musicologie*, *16*, 57-70.

Ruwet, N. (1966). Méthodes d'analyse en musicologie. *Revue Belge de Musicologie*, *20*, 65-90.

Ruwet, N. (1987). Methods of Analysis in Musicology. *Music Analysis*, *6 (1-2)*, 4-39.

Temperley, D. (1988). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.

Toiviainen, P. (1998). An Interactive MIDI Accompanist. *Computer Music Journal*, *22 (4)*, 63-75.

Toiviainen, P. & Eerola, T. (2006). Autocorrelation in meter induction: The role of accent structure. *Journal of Acoustical Society of America*, *119 (2)*, 1164-70.

White, B. (1960). Recognition of distorted melodies. *American Journal of Psychology*, *73*, 100-07.

Zaki, M. (2005). Efficient algorithms for mining closed itemsets and their lattice structure, *IEEE Transactions on Knowledge and Data Engineering*, *17*, 462-78.

• Estrategias de medición motívica para la extracción de patrones ("patterns") automatizados

Este artículo propone una aproximación al problema de la extracción automática de patrones motívicos en monodías. Se han definido diferentes dimensiones musicales, restringidas a aproximaciones actuales a los más destacados elementos melódicos y rítmicos del nivel superficial. La estrategia propuesta para la detección de repeticiones de patrones consiste en una medida exacta de los parámetros sucesivos que forman los motivos. Sugerimos una generalización de las múltiples aproximaciones diferentes que, de acuerdo con una variabilidad de tipos de parámetros (melódico, rítmico, etc.), definen cada extensión sucesiva de esos motivos. Esto nos conduce a considerar una clase más general de motivos, llamados motivos heterogéneos, que incluye motivos interesantes fuera del alcance de aproximaciones previas. Además, esta representación heterogénea de motivos puede ofrecer explicaciones más concretas sobre el impacto de la representación de contorno grueso en el análisis motívico. Este artículo muestra también que el principal problema revelado en la tarea de extracción de motivos está relacionado con el control de la redundancia combinatoria de las estructuras musicales. Se presentan dos estrategias principales, que aseguran un filtro adaptable a las estructuras redundantes, y que están basadas en las nociones de patrones cerrados y cíclicos. El método está ilustrado con el análisis de dos piezas : un *Geisslerlied* medieval y una *Invención* de J. S. Bach.

• Strategie di abbinamento motivico per l'estrazione automatica di schemi

Il presente articolo propone un approccio al problema dell'estrazione automatica di schemi motivici nelle monodie. Si definiscono differenti dimensioni musicali, limitate negli approcci correnti ai principali aspetti melodici e ritmici a livello superficiale. La strategia proposta per il rilevamento di schemi ripetuti consiste in un esatto abbinamento dei parametri successivi che formano i motivi. Suggeriamo una generalizzazione dell'approccio dal punto di vista multiplo, il quale permette una variabilità dei tipi parametrici (melodici, ritmici, ecc.) in grado di definire ogni successiva estensione di tali motivi. Ciò ci consente di prendere in considerazione una classe più generale di motivi, chiamati eterogenei, che include motivi interessanti al di là della sfera degli approcci precedenti. Oltre a ciò, questa rappresentazione eterogenea di motivi può offrire spiegazioni più sottili circa l'impatto della rappresentazione del contorno complessivo sull'analisi motivica. Il presente articolo mostra poi come il principale problema generato dal compito di estrazione degli schemi sia legato al controllo della ridondanza combinatoria delle strutture musicali. Si presentano due strategie principali, le quali garantiscono un filtro adattativo delle strutture ridondanti e si basano sulle nozioni di schemi chiusi e ciclici. Il metodo è illustrato attraverso l'analisi di due brani : un *Geisslerlied* medievale e un'*Invenzione* di Bach.

• Stratégies d'appariement pour l'extraction automatisée de motifs

Nous proposons ici une méthode pour aborder le problème de l'extraction automatisée de motifs dans une monodie. Nous définissons différentes dimensions musicales, limitées aux dimensions mélodiques et rythmiques les plus proéminentes dans les méthodes actuellement utilisées. La stratégie de détection de motifs répétés que nous proposons ici consiste à grouper avec précision les paramètres successifs qui forment les motifs. Nous proposons de généraliser la méthode des points de vue multiples, afin de pouvoir varier les types de paramètres (mélodique, rythmique, etc.) qui définissent chaque extension de ces motifs. Ceci nous permet de tenir compte d'une classe plus étendue de motifs, qu'on appelle motifs hétérogènes, incluant ceux qui n'étaient pas examinés par les méthodes précédentes. De plus, cette représentation de motifs hétérogènes offre une description plus fine de l'impact de la représentation de contours bruts dans l'analyse motivique. Nous montrons aussi que le problème principal que pose l'extraction d'empreintes est lié au contrôle de la redondance combinatoire d'une structure musicale. Nous présentons deux stratégies permettant le filtrage adaptatif des structures redondantes, fondées sur les concepts de motifs fermés et cycliques. Nous donnons comme exemple de la méthode l'analyse de deux pièces : un *Geisslerlied* médiéval et une *Invention* de Bach.

• Motivische Vergleichsstrategien zur automatischen Musterextrahierung

Dieser Artikel schlägt eine Herangehensweise an das Problem der automatischen Extraktion motivischer Muster in Monodien vor. Verschiedene musikalische Dimensionen werden definiert, die sich in jetzigen Herangehensweisen auf die auffälligsten melodischen und rhythmischen Merkmale in der Oberflächenstruktur beschränken. Die vorgeschlagene Strategie zur Erkennung wiederholter Muster umfasst eine exakte Abgleichung der sukzessiven Parameter, aus denen Motive gebildet werden. Wir schlagen eine Generalisierung der mehrdimensionalen Herangehensweise vor, die eine Variabilität in den Arten der Parameter erlaubt (melodisch, rhythmisch etc.), welche jeden sukzessiven Zusatz dieser Motive definieren. Dadurch können wir umfassendere Motivgruppen einbeziehen, so genannte heterogene Motive, somit auch interessante Motive jenseits bisheriger Forschungsansätze. Außerdem bietet diese heterogene Repräsentation der Motive genauere Erklärungen hinsichtlich des Einflusses der totalen Konturrepräsentation in motivischen Analysen. Dieser Artikel zeigt auch, dass das Hauptproblem in der Musterextrahierung mit der Kontrolle der kombinatorischen Redundanz der musikalischen Strukturen zusammenhängt. Zwei Hauptstrategien werden vorgestellt, die einen adaptiven Filter der redundanten Strukturen garantieren, und die auf Vorstellungen von geschlossenen und zyklischen Mustern basieren. Diese Methode wird anhand der Analyse zweiter Stücke dargestellt : eines mittelalterlichen Geißlerlieds und einer Invention von Bach.