MUSICAL PATTERN EXTRACTION: FROM REPETITION TO MUSICAL STRUCTURE

Benoit Meudic

Music representation team
Ircam, Paris
meudic@ircam.fr

ABSTRACT

In the context of musical analysis, we propose an algorithm that automatically induces patterns from polyphonies. We define patterns as “perceptible repetitions in a musical piece”. We claim that a link can be established between the patterns that we extract and the musical structure of the piece. Our aim is that the patterns we extract are perceptively relevant although several perceptive criteria are not taken into account in our model: what we propose is an attempt to explore the limits of a system that do not considers, in a first step, the musical notions of expectation or temporal context (considering temporal context means that we take into account past events for the analysis of current events), but that integrates several other perceptive notions such as polyphonic context. This brings us to discuss several specific issues related to the extraction of patterns in a polyphonic context.
Our model proceeds as follows:
In a first step, we quantize a MIDI sequence with an algorithm described in [1], and we segment the music in “beat segments”.
Then, we compute a similarity matrix from the segmented sequence. The measure of similarity we use is described in [2]. The algorithm computes the similarity between all the pairs of sequences with same number of beats. It relies on features such as rhythm, contour and pitch intervals.
Last, a bottom-up approach is proposed for extracting patterns from the similarity matrix. It builds new patterns by concatenation of the consecutive cells that are on the same diagonal in the similarity matrix.
The algorithm was tested on several pieces of music, and interesting results were found. At the same time, new questions were raised on the notion of musical structure.

1 INTRODUCTION

Automatic music analysis is an increasingly active research area. Among the main tackled subjects, the search for *musical patterns* is at a central place. Indeed, most of the musical pieces are structured in various components (“phrases”, “motives”, “themes”, “generative cells”…) that can naturally be associated with the notion of pattern. Considering the only notion of “pattern” simplifies (or postpones) the issue of making the distinction between the different natures of the components (theme, motive etc…) of a musical piece. However, answering to the question “what is a pattern?” is still rather difficult. Often, the notion of pattern can be linked with the notion of repetition: patterns emerge from repetition. But a pattern could also be defined by its salient boundaries, and then patterns would emerge from discontinuities in the music. Last, patterns can also be characterized as independent and coherent entities. Providing a definition is all the more difficult as we place ourselves in a musical context. Most of the time, patterns are linked with perceptive notions, which raises one question: can we consider as “patterns” the structural (one would say mathematical) regularities of a musical sequence, even if theses regularities are not perceived? Inversely, do perceived repetitions only correspond to exact repetitions? We think that the notion of similarity between two sequences plays an important role in the perception of patterns and it should be part of a pattern extraction system.

In this article, we have chosen to focus on perceptible musical structures. Moreover, we assume that musical structures can be induced from the extraction of repeated sequences (which we call patterns) and thus we address the issue of extracting “perceptible repetitions” from a musical piece.

2 BACKGROUND

The literature is quite poor in algorithms that automatically extract patterns (perceptible repetitions) from polyphonic music.
An interesting method, starting from the audio, is proposed in [3]. It considers the signal as a succession of “states” (at various scales) corresponding to the structure (at various scales) of a piece of music. A similarity matrix is computed from feature vectors. The similarity matrix reveals large and fast variations in the signal that are analysed as boundaries of potential patterns. The patterns are then grouped according to their similarity. The method is relevant for pieces that contain salient transitions (timbre changes, important energy variations etc…), but could reveal less relevant for the detection of phrases in piano music.
Another method [4] starts from MIDI files and induces patterns by analysing the musical sequence in a chronological order. All the possible combinations of successive events contained in a temporal window of limited size are potential candidates for being a pattern. The notions of temporal context and expectation are modelized. However, if very promising, this method cannot analyse long sequences with too many events because it would require a too high computation cost. Moreover, polyphonic context (see 4.1 for definition)

is not considered.
Last, [5] proposes to segment a MIDI file and then to categorize the segments in clusters. The segmentation algorithm uses a boundary detection function that computes the similarity between all the possible patterns of the sequence. Then, the segments are clustered according to a threshold that depends on the different distances between the segments. Once the clusters are computed, the distances between segments are re-computed in order to optimize the current clustering, and the clustering function is called again until the clusters are found optimal. This method is interesting because it analyses whole sequences of polyphonic MIDI music, and the notion of context is originally used in one of the two steps of the analysing (the clustering). However, the segmentation step appears hazardous, and would require a high computational cost.

3 AIMS

The algorithm we present in this article is a new model for extracting patterns from polyphonic MIDI sequences. Our aim is that the patterns are some components of the musical structure. In order to make the system the more general (and primitive) as possible, we do not consider any knowledge based on tonality neither we consider particular styles. Several perceptive notions are taken into account in the algorithm. However, in a first step, we have chosen to challenge the limits of a system that does not modelize temporal context or expectation. Indeed, the integration in the system of all the cognitive mechanisms that play a role in our perception of music is far too complex. Thus, we have to draw the limits of the model. The consideration of temporal context and expectation should be possible in a second step, but we first prefer to explore the limited system, and we will rely on the results to show that our method offers promising applications. Besides, we will not try to extract all but a set of significant patterns from polyphonic music. If our aim is that the extracted patterns are the most relevant of all the sequence, we consider as more important the relevance of the pattern itself.

4 INTRODUCTION TO THE PATTERN EXTRACTION MODEL

4.1 General considerations on patterns

We have defined patterns as “perceptible repetitions in a musical sequence”. This notion has to be refined before we describe in details the model.

An issue arises when refining the term ‘repetition’ of the definition. How to define the similarity between two sequences? An attempt to answer to this question is proposed in 5.2.
Another issue stems from the polyphonic context: is a pattern a single melodic line inside a polyphony, or is it itself a polyphonic component of a polyphony? For instance, let’s consider the following extract of the 9nth Beethoven’s sonata (Figure 1).

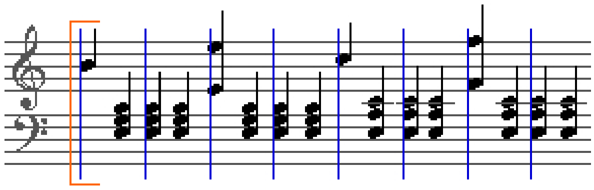

Figure 1. Beginning of the Sonata for piano No9 EM from Beethoven. The first note and the three chords form a pattern. Transforming the three chords into three independent voices would not be perceptively relevant.

The sequence formed by the first four events can be called a pattern as it is repeated several times. This sequence is polyphonic (one note followed by three chords). Should we consider that the sequence of chords is formed by three different monophonic patterns that are superposed, or is it one polyphonic pattern? Would it be relevant to dissociate the chords in as many components as there are notes inside, or should we consider them as entities? This also asks the question of the independence of a melodic line from its polyphonic context. Can we always isolate one or several monophonic lines, or do the superposition of several notes form a unique entity that cannot be decomposed? In several musical examples, there would be no reason to extract a single monodic line (see also Figure2).

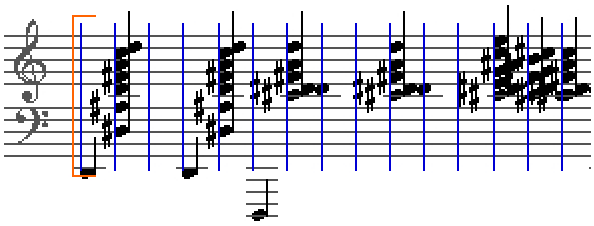

Figure 2. A pattern extracted from the Sacral dance of the Rite of the spring from Stravinsky. A monodic voice could hardly be isolated from the polyphonic pattern.

We also claim that considering a single melodic voice without its polyphonic context is often nonsense, as the polyphonic context often plays an important role in the perception of the structure of the piece. For instance, if we consider the beginning of the Sonata AM d664 op121, 1rst Part from Schubert (Figure3), one can extract the melody from the voice at the top. However, the segmentation of the melody in different patterns (that correspond in musical analysis to the usual “antecedent-following” repeated one time with variation) is not trivial without information on the polyphonic context. Similarly, if we consider the beginning of the Intermezzo op117 no1 from Brahms (Figure 4), we could extract the melody by hand with difficulties (it is not at the top of the voices, and it is repeated but with big variations), but it would not characterize the excerpt. Indeed, one could imagine the same simple melody in very different polyphonic contexts, divided in several different voices.

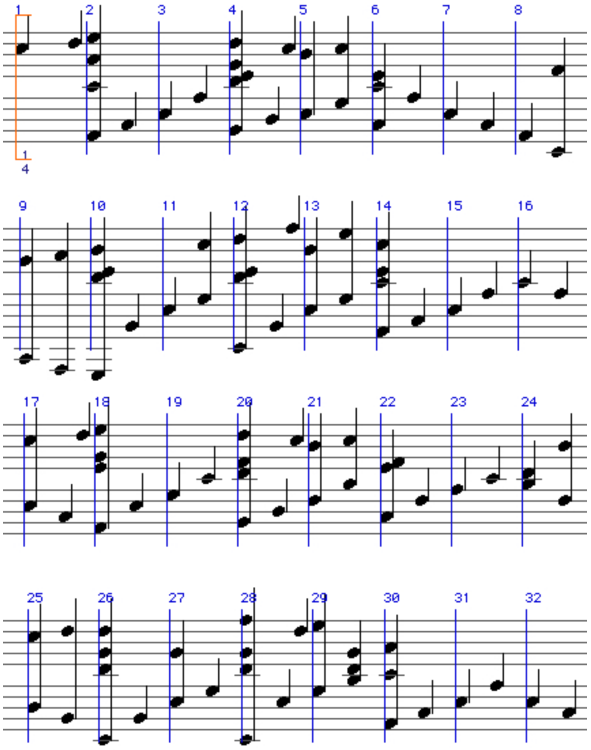

Figure 3. Beginning of the Sonata AM d664 op121, 1rst Part from Schubert. The four patterns correspond to the structure antecedent1-following1-antecedent2-following2. This structure could not be extracted from the consideration of the only melodic line.

However, and this raises one of the main issues, the only melodic line should also be taken into account, for instance when it follows the model 'melody + accompaniment'. Indeed, in this case the melodic line could be separated from the accompaniment and considered as a monophonic pattern that would be repeated in the following music sequence, but with different accompaniment, or with melodic variations. This is mainly a perceptive issue that is difficult to solve with only one rule.

Sometimes, the solution is trivial. For instance, melodic lines are often difficult to follow in Bach's Fugues, but fortunately, they appear at the beginning of the fugue one after the other, which let us time to memorize them. In this case, the pattern first appears in a monophonic context (the Fugue's subject) and then is repeated and varied in a polyphonic context. The issue is not to extract a pattern (monophonic or polyphonic) from polyphony but to recognize a known (or memorized) monophonic pattern inside polyphony. However, in some pieces such as canons, only the beginning of the pattern appears in a monophonic context. The repetition covers a part of the pattern. Thus, in a polyphonic context, music cannot always be segmented in individual successive segments (that is often proposed in several musical theories) and possible coverings between the structural components must be taken into account.

Whatever the situation would be, we believe that the polyphonic context plays an important role. Even if a single melodic line could be extracted from the polyphony, the polyphony should be associated to the melody in most of the cases. One could say that when remembering a polyphonic excerpt, we often sing a single melodic line, but this is due to our physical impossibility to sing a polyphonic sequence, and often when singing the melodic line, we hear (but don't sing) the polyphonic context (at least the harmony) that was associated with it. It means that we have memorized it, and that we take it into account in our comparisons with other sequences.

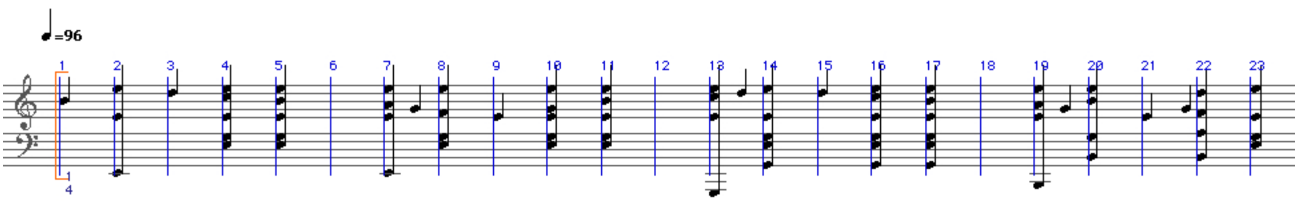

Figure 4. Beginning of the Intermezzo op117 no1 EM from Brahms. The sequence can be divided in four similar patterns (1-6, 7-12, 13-18 and 19-24) Polyphonic context is part of the structure of the sequence. The only melody line with another polyphonic context would not be stated as similar to this one.

In this article, we propose to define a pattern as a polyphonic component of a polyphonic sequence.

4.2 General architecture of the model

Our model analyses a MIDI file that contains the onset (in milliseconds), the pitch (in midicents) and the duration features. The dynamic and the channel features are not considered.

The model is composed of two main algorithms:

- The first algorithm computes several similarity matrices from a quantized MIDI file (Section5).
- The second algorithm extracts patterns from the similarity matrices (Section 6).

In the first algorithm, we first quantize and segment the MIDI sequence with an algorithm proposed in [1]. The boundaries of the segments correspond to the downbeats of the music. The initial sequence of MIDI events is thus considered as a sequence of beat-segments. Each beat-segment (b.s) is itself a sequence of MIDI events (see Figure 5).

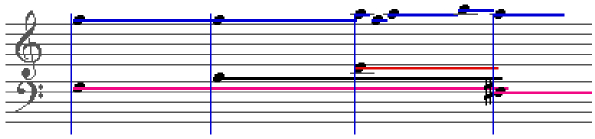

Figure 5. A sequence of beat-segments (b.s) extracted from the "Variations Goldberg" from Bach. The vertical lines delimit the beat segments. Horizontal lines are the durations of each event.

Then, given a length L (in number of beats), each sequence of b.s of length L is compared with all the other sequences of same length. For each comparison, three similarity values are computed (corresponding to three different features: pitch intervals, pitch contour and rhythm) and associated with the events of the two sequences that have been found similar (we call them "templates"). All the similarity measures are then stored in similarity matrices. The measures of similarity we use are described in [2]. They will be presented in 5.2.

In the second algorithm (section 6), patterns are extracted from the similarity matrices. First, the matrix cells (a matrix cell is a couple of two sequences of b.s of same length) are filtered and clustered according to their similarity value. Then, patterns are extracted in a bottom-up approach that starts from the different clusters of cells and then groups them in new (and often smaller) clusters with longer cells.

5 The similarity matrices

5.1 The computation of a matrix

A similarity matrix (Figure 6) stores the similarity values between the pairs of b.s (beat.segment) sequences of same length L (in number of beats). The units of the vertical and horizontal axis are expressed in number of beat-segments. We consider the matrix as symmetric (see discussion in 4.4.5). Each cell refers to a pair of b.s sequences. The vertical and horizontal positions of each cell are the two beginning positions (in number of beats) of the two corresponding b.s sequences.
Each similarity measure provides a real value between 0 and 1 that states how similar are the compared sequences (1 is for identical).
In our model of similarity, sequences of b.s are of same length (length is expressed in number of b.s), so that each position of b.s in a given sequence can be matched with the same position of b.s in the other sequence.

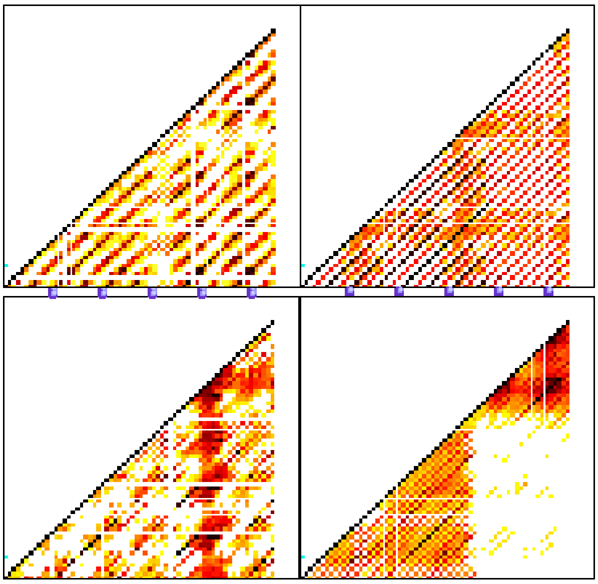

Figure 6. Four different similarity matrices from first 30 seconds of (from left to right and top to bottom): Sonata AM d664 op121 2nd Part from Schubert, 1rst Gymnopedie from Satie, 1rst and 3rd part of the Sacral dance of the Rite of the Spring from Stravinsky. White areas correspond to dissimilar cells. The unit for both vertical and horizontal axis is expressed in number of beat-segments (one cell per beat-segment).

The choice of the length L can appear somewhat arbitrary. The issue is to find the "right level" between the maximum length of patterns that cannot be divided in several parts and the minimum length of patterns that can be divided. Indeed, we think that some patterns must be considered as whole entities (see sub-section 5.2.1) and not as the concatenation of smaller patterns. Two such patterns (sequences of beat-segments) can be similar while their components (the beat-segments), if considered individually, are different. We think that the maximal length of such patterns could be linked to the limits of our short-term memory: when listening to a musical sequence, we compare what we are hearing with what we have already listened. If what we are hearing is a variation of something already listened, we will match the two sequences. We match the sequences because they share common aspects. The question is: what is the minimal number of common aspects, and the maximal distance between the aspects, that is required to initiate a match between the sequences. In other words, what is the maximal size of a pattern (for us, the size is in number of beat-segments) that must be considered for matching it with another pattern?
In our tests, we have decided to define a "blurred" length of pattern L: we compute the similarity between pairs of sequences of length L, L-1, L-2 and L-3, and we choose the best (the higher) similarity value (a decreasing coefficient is applied to the similarity value when the length of the sequence decreases, in order to support longer sequences).

5.2 The similarity measures

In this part, we define several similarity measures between two b.s sequences of given length.
We compute three different similarity values by considering three different sets of features: pitches (chords, pitch intervals etc…), pitch contours (contour at the top and at the bottom of the polyphony) and rhythm. The similarity values are computed in respect with some cognitive aspects (see sub-section 5.2.1). More details can be found in [2]. Each time a similarity value is computed between two sequences seq1 and seq2, it is associated with two "templates", that is to say the events of seq1 similar to seq2 (template1) and the events of seq2 similar to seq1 (template2). These templates will be used to refine the kind of similarity relation that exists between the two sequences (see sub-section 6.1).

5.2.1 Cognitive aspects of the similarity measures

First, a musical sequence of b.s is considered as a whole entity (it may contain an abstract cognitive structure), and not solely as the concatenation of smaller entities. Indeed, we think that several relations between non-adjacent events emerge from the whole entity. Theses relations play a role in the cognitive processes for recognizing the similarity between two sequences. To integrate this aspect, the similarity value between two sequences will not be computed from the addition of the similarity values between the smaller components:

$$S(x, x') + S(y, y') \neq S(xy, x'y') \qquad (1)$$

S(x, x') designs the similarity value between sequences x and x', and xy designs the concatenation of sequence x and y.
Another cognitive aspect is that our similarity measure is not symmetric in a polyphonic context:

$$S(x, x') \neq S(x', x) \qquad (2)$$

If x is approximately included in x', x will be very similar to x'. But x' will not automatically be very similar to x because some events in x' may not be included in x.
Last, according to [6], the similarity measure is not transitive, that is to say the triangular inequality is not true:

$$S(x, y) + S(y, z) \leq \text{ or } \geq S(x,z) \qquad (3)$$

For instance, z can be a variation of y that is a variation of x. But z can be very far from x and thus not judged as a variation of x.

5.2.2 Similarity measure for pitches

We consider here the chords and the pitch intervals features. A similarity value is computed from two b.s sequences seq1 and seq2 of same length.
The only events falling on the downbeats are considered. This may be arguable, but two reasons have conducted this choice:

- Considering all the polyphonic events would require too much running time.
- The downbeats are often perceived as salient temporal position. Two sequences whose pitches coincide on the downbeat but differ elsewhere are often recognised as very similar (this has been confirmed in our experiments).

Usually, a downbeat event (dwb.event) is a chord, but it can also be a note or a rest.
All the intervals (horizontal and vertical) between all the pairs of dwb.event of one sequence are compared with all the intervals between the corresponding pairs of dwb.event of the other sequence (see Figure 7).
The similarity values between pitches or intervals are:

- 1 for equal pitches or equal intervals
- 0.5 for transposed vertical intervals
- 0 otherwise

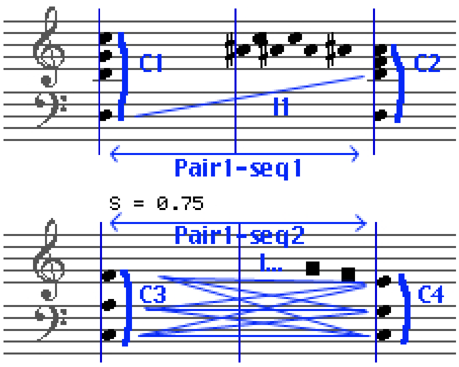

Figure 7. Excerpt from Sonat8Am-mvt2 from Mozart. A reference pattern (the above one) is compared to another pattern. Durations are not represented. All vertical and horizontal intervals between dwb.events C1 and C2 of Pair1-seq1 and dwb.events C3 and C4 of Pair1-seq2 are compared.

5.2.3 Similarity measure for contours

Our model compares the upper and lower contours of two b.s sequences seq1 and seq2 of same length.
As above, the only events falling on a downbeat (dwb.events) are considered. An up (down) contour is the sequence of the intervals between the upper (lower) pitches of the consecutive dwb.events. Each contour of each sequence is compared with the two contours of the other sequence. Contours are very similar (see Figure 8) if the intervals from one sequence are similar to the corresponding intervals from the other sequence (two intervals are similar if their difference is less than 5 half tones).

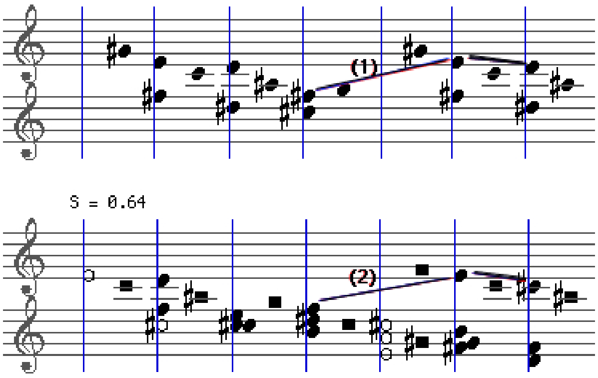

Figure 8. Two similar patterns in Pierrot Lunaire from Schoenberg. The events that determined the similarity between the patterns (for pitches, contours and rhythm) are represented in black. Lines (1) and (2) show similar contours.

5.2.4 Similarity measure for rhythm

Our model compares the rhythmic structure of two sequences of b.s seq1 and seq2 of same length. In a first step, seq1 and seq2 are normalized so that the total duration of the b.s will be the same for seq1 and seq2. Then, for each b.s, onsets (temporal positions) in seq1 are associated to the corresponding onsets in seq2. Two onsets of two b.s form a pair if they share similar temporal positions in the b.s. If an onset of one sequence does not form a pair with an onset of the other sequence, then it is deleted. The similarity between two sequences of b.s is the mean of the similarity between each corresponding b.s (as seq1 and seq2 have same length, each b.s of seq1 correspond to one b.s of seq2). The similarity between two corresponding b.s is the mean of the similarity between each pair of corresponding onsets. Corresponding onsets are already similar because they share the same temporal position in the b.s. The similarity increases with the length of the intersection of the durations of the events corresponding to the onsets of a pair (an approximation value of the durations is considered for the intersection).

5.2.5 Overall similarity measure

Each of the three above measures (pitches, contour and rhythm) computes a similarity value between each pair of sequences of length L. The results can be represented in three different similarity-matrices that can be analysed separately. Sometimes, it appears that rhythm, pitches and contour play a different role in the similarity measure, and sometimes the similarity matrices for the different measures are very similar (see Figure 10 and Figure 12).
The three measures can also be linearly combined into a global similarity measure (see Figure 9 and 11 for similarity between sequences and Figure 10 and 12 for similarity matrices). In this case, different weights can be applied to the different measures. An algorithm could be used to determine the best weightings by comparing the output of the similarity measures to a set of expected similarity-matrix.
In our experiments, we have chosen to give a higher weight to the rhythmic and the pitch based measures as they are more "selective" than the contour measure: the contour is expected to be common to more sequences than the pitch or the rhythmic successions of events.

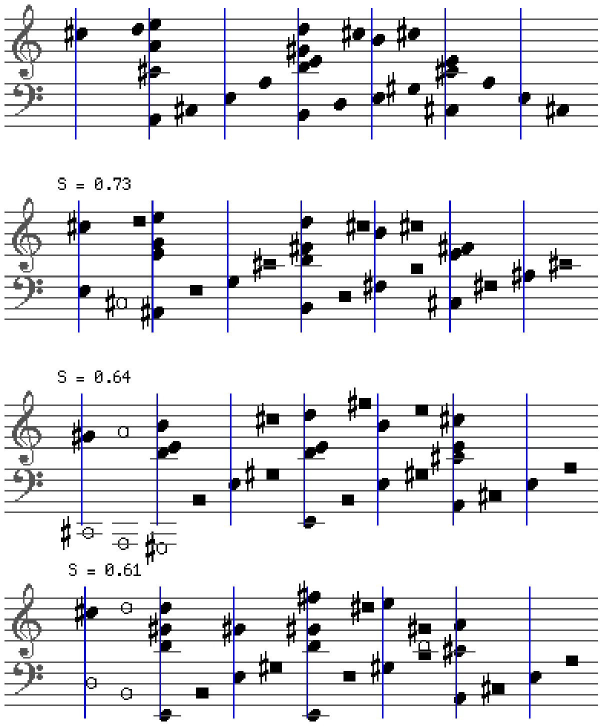

Figure 9. Three patterns similar to the above pattern in Sonata AM d664 op121, 1rst Part from Schubert. The similar patterns are sorted according to their decreasing global similarity value S. The square symbols only determined the similarity for rhythm. The oval symbols determined similarity for both pitches and contour.

Due to the non-symmetrical relation (see equation 2), the similarity value between two non-ordered sequences seq1 and seq2 is composed of two different values: S(seq1, seq2) and S(seq2, seq1). In our model, we have chosen not to consider the two different values and thus we only consider the greater value. Doing that, our similarity matrix becomes symmetric and we represent only half of it.

6 Pattern extraction from the similarity matrix

Several different types of information can be extracted from the matrices: general evolution for rhythm, pitch intervals or pitch contours, local repetitions of cells, areas of structural changes, etc…
All of theses informations could be linked with the notion of pattern. However, in this part, we will only focus on the extraction of "the most important" patterns.
Defining such patterns is quite difficult, as there are no objective criteria to characterize them. One would agree to say that the "most important" patterns are the ones that are perceived as the "most salient" in the

musical sequence: thus the attribute “most important” is related to perceptive criteria.

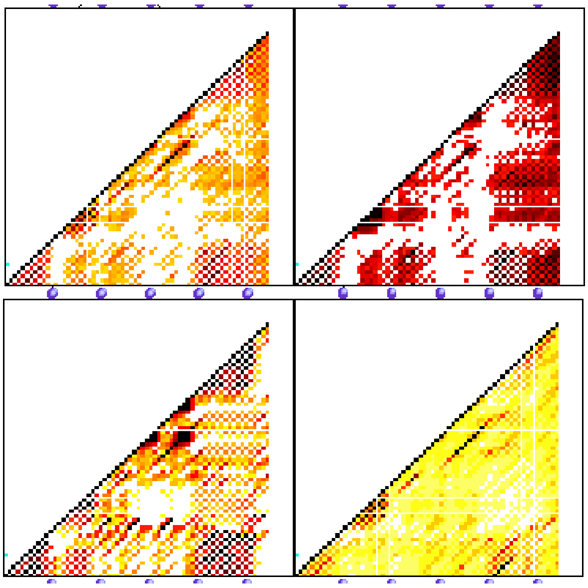

Figure 10. Four similarity matrices from Sonata for piano No9 EM from Beethoven (from left to right and top to bottom): overall similarity measure, similarity for rhythm, similarity for contour and similarity for pitches. The contribution of the different measures to the overall similarity measure is very different (see figure 12 for comparison).

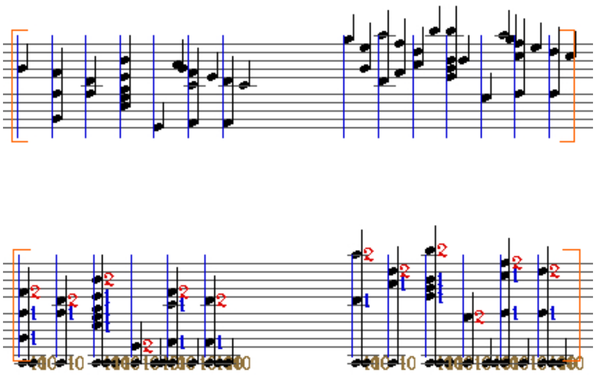

Figure 11. Detail of one cell of the similarity matrix (Figure 10). Above, the two compared sequences of the cell. Below, the similar events of each sequence (the templates). Numbers 1, 2 and 10 correspond to similar pitches, contour and rhythmical events. As rhythmical similarity do not considers pitches, it is represented by the events at the bottom of the sequence with static pitch C.

We think that the musical temporal context would play an important role in the definition of those criteria but once again, we will try not to consider it. Two other criteria are often found in the literature: the length of the pattern, and the number of its repetitions. However, it is quite difficult to combine the two criteria: the “most important” patterns often appear as a compromise between the two. For instance, the same accompaniment can be found all along a musical piece. This pattern is very often repeated and thus very important in regards with the second criterion. However, it rarely appears to us as the “most important” pattern. Inversely, musical pieces that are exactly repeated two times can be seen as two very long patterns but they are not the most relevant for us because the repetition is trivial.

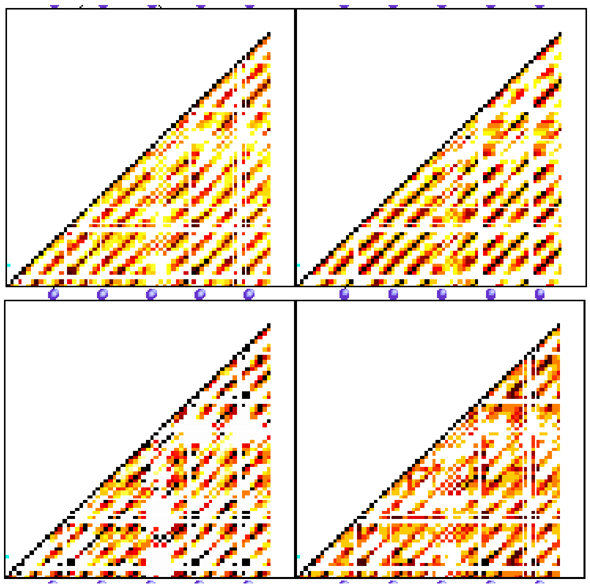

Figure 12. Four similarity matrices from Sonata DM d664 op121 2nd Part from Schubert (from left to right and top to bottom): overall similarity measure, similarity for rhythm, similarity for contour and similarity for pitches. The contributions of the different measures to the overall similarity measure are very similar (see Figure 10 for comparison).

We will now propose a method for extracting patterns that could be “the most important ones”. First, we propose to filter and cluster the cells of the matrix. Then, we propose to concatenate the cells in long patterns.

6.1 The filtering/clustering of the cells of the similarity matrices

As we are looking for repetitions, we focus on the patterns that are very similar. For that, we choose a threshold that selects the only cells with the higher similarity values (see Figure 13).

Then, we cluster the cells: each horizontal line of the similarity matrix represents all the similarity values between a reference sequence s(ref) and all the sequences of the line s(i). The high similarity values reveal high similarity with the reference sequence, but they do not reveal the kind of similarity, which has to be evaluated. For that, we compare the templates of the sequences. The template of seq1 (respectively seq2) contains the events of seq1 (resp seq2) similar to seq2 (resp seq1) (see Figure 11).

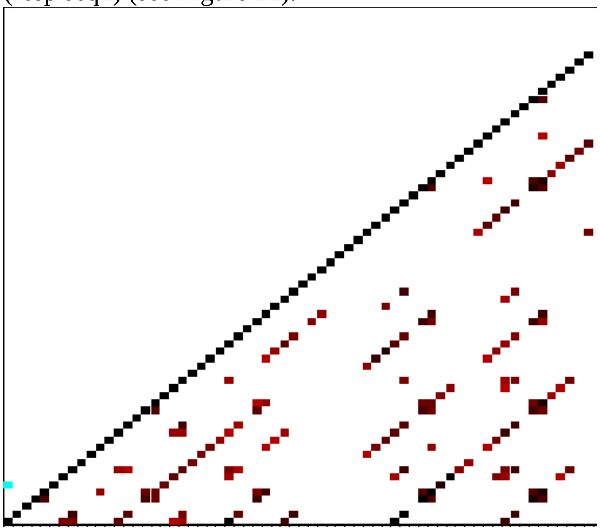

Figure 13. A filtered similarity matrix from Sonata DM d664 op121 2nd Part from Schubert (the initial matrix is in Figure 12). The most similar patterns appear (diagonal lines).

Sequences that have very similar templates are considered as sequences that have same kind of similarity.

6.2 The extraction of patterns

We define a bottom-up approach for pattern extraction that starts from the clustered cells and builds new patterns by concatenation of the consecutive cells of a same diagonal line.

The algorithm proceeds along the matrix from the lowest horizontal line to the upper one. For instance, in Figure 14, the lowest horizontal line contains 8 cells. Only 6 of theses cells are clustered together. If we consider the second horizontal line, 6 cells can be concatenated to the 6 clustered cells of the first line. Thus, we have extracted 6 patterns of length “2 cells”. As there are no cells in the third line, we stop the process.

In a second step, we associate to each pattern the corresponding horizontal lines of the matrix (see 6.1). Theses lines contain clustered cells that could be concatenated and that could form new patterns (see Figure 14). Theses patterns could then be added to the six other ones, depending on their kind of similarity with them. In Figure 14, we highlight the 6 extracted patterns and a pattern that has been added in a second step (the last on the right of the figure).

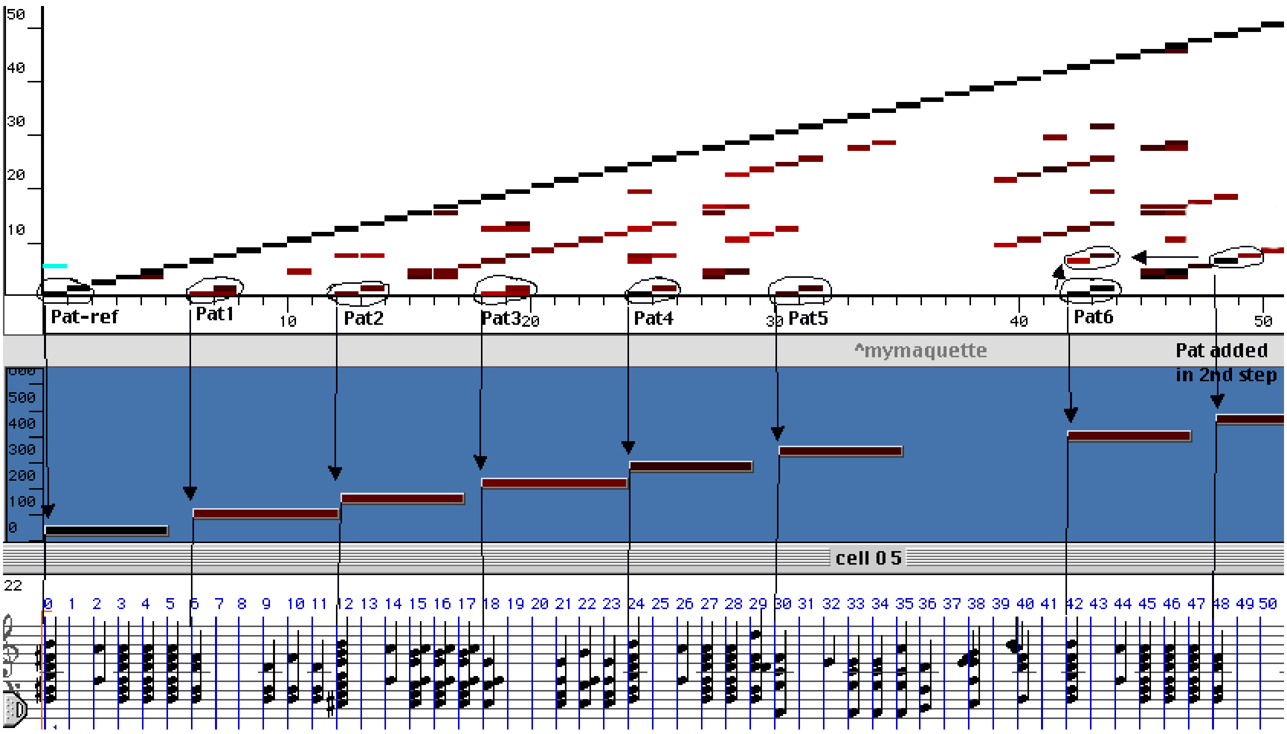

Figure 14. (from top to bottom) The filtered similarity matrix from Sonata DM d664 op121 2nd Part from Schubert, the extracted patterns (each horizontal segment is an instance of the same pattern) and the corresponding musical sequence (the graduation of the matrix and the musical sequence is in “number of beat-segments”).

We have tested several different musical pieces, and patterns that are part of the musical structure were often extracted. All of the musical examples that have been presented along the paper have been analysed, and the extracted patterns could always be associated with the musical structure of the piece. For other pieces, some patterns could be found that do not begin or end at the right temporal position in the sequence.

Together with this issue, the consideration of all the notes of the polyphonic context was sometimes a constraint for recognizing the repetition of the only melodic line (for instance in canons pieces). The relations between the two (polyphony and the melodic line) is one of the main issue of pattern extraction.

7 CONCLUSION

We have presented a general system for pattern extraction from polyphonic musical sequences. The notion of perceptible pattern has been discussed in the context of polyphonic music. A global similarity measure that takes into account rhythm, pitches and contour in a polyphonic context has been proposed. A method for extracting patterns from similarity matrices has been described and we have provided several musical examples for illustration. The method we propose could be used in several different musical applications (extraction of some components of musical structure, motive discovery, characterization of different musical pieces from their rhythmical or pitch components). The similarity measure could also be adapted for applications such as "query by constraints": instead of humming a pattern, one could specify it with constraints on pitches or rhythm. The algorithm could then extract several patterns from different musical pieces with same rhythmical or pitch profile.

In future work, we plan to integrate temporal context in the model, in such a way that both polyphonic and temporal context are taken into account in the computation of the similarity matrices. We think that our model considers interesting polyphonic aspects of music that are rarely taken into account, but several other polyphonic issues remain to be solved, particularly the relation between polyphony and the melodic line. However, we believe that solutions should emerge from such experimentations that come and go between perceptive considerations and modelization.

8 ACKNOWLEDGEMENTS

Part of this work was conducted in the context of the European project CUIDADO. I want to thank Olivier Lartillot for the interesting discussions that we have had on this subject.

9 REFERENCES

[1] Meudic B, "A causal algorithm for beat tracking", 2nd conference on understanding and creating music, Caserta, Italy, 2002

[2] Meudic B, "Musical similarity in a polyphonic context : a model outside time", submitted to the XIV Colloquium on musical informatics (CIM), Firenze, Italy, 2003

[3] Peteers G et al, "Toward Automatic Music Audio Summary Generation from Signal Analysis", Proc Ismir, Paris, 2002,

[4] Lartillot O, "Perception-Based Advanced Description of Abstract Musical Content", Proc. WIAMIS, London, 2003

[5] Cambouropoulos E, "Towards a General Computational Theory of Musical Structure", PhD, Edinburgh, Faculty of Music and Department of Artificial Intelligence, 1998

[6] Tversky A, "Features of similarity", journal of Psychological Review, p327-352 1977