Musical pattern extraction for melodic segmentation

Article · January 2003

1 author:

Emilios Cambouropoulos
Aristotle University of Thessaloniki

MUSICAL PATTERN EXTRACTION FOR MELODIC SEGMENTATION

Emilios Cambouropoulos

Department of Music Studies
Aristotle University of Thessaloniki
emilios@mus.auth.gr

ABSTRACT

Despite the fact that musical parallelism is considered as an important factor for musical segmentation, there have been very few attempts to describe systematically how exactly it affects grouping processes. The main problem is that musical parallelism itself is very difficult to formalise. In this paper a computational model will be presented that extracts melodic patterns from a given melodic surface. Following the assumption that the beginning and ending points of 'significant' repeating musical patterns influence the segmentation of a musical surface, the discovered patterns are used as a means to determine probable segmentation points of the melody. 'Significant' patterns are defined primarily in terms of frequency of occurrence and length of pattern. The special status of non-overlapping immediately-repeating patterns is examined. All the discovered patterns merge into a single 'pattern' segmentation profile that signifies points in the surface that are most likely to be perceived as points of segmentation. The effectiveness of the algorithm is tested against a series of musical surfaces illustrating both strengths and weaknesses of the approach.

1. INTRODUCTION

Music becomes intelligible to a great extent through self-reference, i.e. through the relations of new musical passages to previously heard material. Structural repetition and similarity are crucial devices in establishing such relations. Similar musical entities are organised into musical categories such as rhythmic and melodic motives, themes and variations, harmonic progression groups etc. Musical similarity not only establishes relationships between different musical entities but enables - in the first place - the definition of such entities by directly contributing to the segmentation of a musical surface into meaningful units.

Despite the importance of musical parallelism, even the most elaborate contemporary musical theories avoid tackling the problem of parallelism in a formal way. Temperley, that has developed one of the most sophisticated computational models of musical cognition admits that "despite the clear role of parallelism in meter, it would be very difficult to incorporate parallelism into a computational model. The program would have to search the music for patterns of melodic and rhythmic repetition. Since this seems to me a huge and complex problem, I am not addressing it formally in this book." (Temperley, 2001:51).

Pattern-matching techniques have been employed in attempts to formalise musical similarity. Most such research, however, has focused on algorithms for comparing melodic sequences (i.e. finding the best possible alignment between two given melodic excerpts) or for melodic recognition (i.e. finding instances of a given melodic excerpt in a larger musical database). Only very rarely have there been attempts to tackle the difficult issue of pattern extraction (i.e. extracting important patterns in one or more musical sequences) – one such very interesting model has been developed by (Rolland 1999, 2001). Overviews of the application of pattern processing algorithms on musical strings can be found in (Crawford et al. 1998; Cambouropoulos et al. 1999, Rolland et al. 1999).

In this study melodic pattern extraction is used as a means to segment a melodic surface. The current study is a continuation of the research presented in (Cambouropoulos 1998).

2. PATTERN EXTRACTION & SELECTION

An efficient algorithm that computes *all* the repetitions in a given string is described in (Crochemore, 1981; see also description in Iliopoulos et al., 1996). For a given string of simple or complex symbols, the matching process starts with the smallest pattern length (1 element) and ends when the largest pattern match is found. This algorithm takes $O(n \cdot \log n)$ time where n is the length of the string – this is the fastest algorithm possible.

It is apparent that such a procedure for the discovery of all identical melodic patterns for many melodic parametric strings will produce an extremely large number of possible patterns most of which would be considered counter-intuitive and non-pertinent by a human musician/analyst.

A procedure has been devised whereby a prominence value is attached to each of the discovered patterns based on the following factors: a) prefer most frequently occurring patterns, b) prefer longer patterns, c) avoid overlapping. A *selection function* that calculates a numerical strength value for a single pattern according to the these principles can be devised, for instance:

$$f(L,F,DOL)=F^{a} \cdot L^{b}/10^{c \cdot DOL}$$

where: L: pattern length; F: frequency of occurrence for one pattern; DOL: degree of overlapping[1]; a, b, c: constants that give different prominence to the above principles.

[1] DOL is defined as the number of elements shared by some patterns divided by the number of all the elements in those patterns or more precisely: DOL = (T-U)/U where: T is the total number of elements in all the instances discovered for a pattern (T=F·L); U is the number of elements in the union set of all the instances discovered for a pattern (this definition allows DOL to be in some cases greater than 100%).

For every pattern discovered by the above pattern induction algorithm a value is calculated by the selection function. The patterns that score the highest should be the most significant ones.

3. SEGMENTATION AND PARALLELISM

Segmentation of a musical surface is a central part of musical analysis; an initial selected segmentation can seriously affect subsequent analysis as a great number of inter-segment musical structures are excluded *a priori*. The most commonly acknowledged (and perhaps most prominent) factors in musical segmentation relate to the perception of local discontinuities of the surface (e.g. longer note in between shorter ones or larger pitch interval in between smaller intervals etc.) – one such successful model is the *Local Boundary Detection Model (LBDM)* proposed by (Cambouropoulos 1997, 2001a). The segmentation of a musical surface, however, is also affected by higher-level processes as well. Perhaps the most important of these higher-level mechanisms is *musical similarity*, i.e. similar musical patterns tend to be highlighted and perceived as units/wholes whose beginning and ending points influence the segmentation of a musical surface. For instance, a model for determining local boundaries would select the interval between the 3rd and 4th notes of *Frère Jacques* (Figure 1) as a local boundary (larger pitch interval in between smaller ones) whereas it is obvious that a boundary appears between the 4th and 5th notes because of melodic repetition.

The focus of this study is primarily a special case of melodic similarity, namely immediate repetition of melodic passages. Such repeating passages often diverge towards their endings, contain small variations and the repeated passage may be transposed. David Lidov (1979) calls this kind of repetition *formative repetition.* Its function is to establish or to 'form' motives and phrases. It involves fundamental pattern discovery processes primarily at the melodic surface (not reductions of the surface) and essentially is independent of more abstract learned idiom-specific schemata (e.g. harmony, tonality, meter). This kind of melodic similarity is omnipresent in music – a number of such examples are presented in this paper.

It is herein assumed that similarity processes for melodic segmentation tasks are restrained essentially to the melodic surface in contrast to melodic categorisation tasks (i.e. creating motivic/thematic categories *after* segments have been defined) which require similarity measurements at deeper levels of musical structure as well (see Cambouropoulos 2000, 2001b for a computational model of melodic categorisation). This seems to be necessary because extracting patterns at reduced versions of the melodic surface would result in ambiguous segmentations as it would not be possible to define where exactly the boundaries of the repeated patterns should be placed (since there are notes missing from the reduced version). This problem, in some sense, defeats the point of using pattern extraction at reduced versions of the surface for melodic segmentation. Of course, musical similarity appears in many guises at deeper levels of musical structure but in such cases it is likely that this sort of abstract similarity is not the most crucial factor in segmentation tasks – other factors such as gestalt-based local boundary detection factors or learned schemata (e.g. harmonic cadences) are responsible for segmenting the surface and only then are more sophisticated comparisons of segments made possible at more abstract levels of description.

In this paper, the pattern extraction algorithm is applied at parametric profiles of the melodic surface for pitch intervals (diatonic intervals and step-leap intervals) and for interonset intervals. An important aspect of the paper is to discover which of these parameters (or combination of them) is most appropriate for the segmentation task (see discussion below).

The pattern extraction model described in section 2, that consists of the exact pattern extraction algorithm and selection function, provides a means of discovering 'significant' melodic patterns. There is, however, a need for further processing that will lead to a 'good' description of the surface (in terms of exhaustiveness, economy, simplicity etc.). It is likely that some instances of the selected pitch patterns should be dropped out or that a combination of patterns that rate slightly lower than the top rating patterns may give a better description of the musical surface.

In order to overcome this problem a very simple methodology has been devised – see Table 1.

Construction of the pattern boundary strength profile (PAT)
The pattern extraction procedure is applied to one (or more) parametric sequences of the melodic surface as required. No pattern is disregarded but each pattern (both the beginning and ending of pattern) contributes to each possible boundary of the melodic sequence by a value that is proportional to its Selection Function value. That is, for each point in the melodic surface all the patterns are found that have one of their edges falling on that point and all their Selection Function values are summed. This way a pattern boundary strength profile is created (normalised from 0-1). It is hypothesised that points in the surface in which local maxima appear are more likely to be perceived as boundaries because of musical similarity.

Table 1

In the melodic example of Figure 1 the pattern boundary strength profile (PAT) has been calculated by applying the pattern extraction model to the diatonic pitch interval profile – notice the strong pattern boundaries at the points indicated by asterisks where no local boundaries are detected by *LBDM* or other local detail grouping models.

The rhythmic profile is not used as it is poor in terms of information content, i.e. the number of available duration values that form the alphabet of this parametric profile is too small and repetitions too many. It is obvious that the smaller an alphabet is (in *Frère Jacques* only three durational values) the larger the number of repeating patterns is. This low information content (high redundancy) means that it is unlikely that this profile will convey non-trivial pattern information (see below for a method to incorporate rhythmic parameters whilst retaining a higher information content).

The above example consists only of exact full repetitions. This, however, is not usually the case. A very frequently encountered situation is when two patterns diverge towards their ends (see example in Figure 2).

In general, the beginning of melodic patterns is of paramount important into discovering parallel passages. This intuition has been incorporated into the current model by making a very simple modification to the method described in Table 1: *only the beginnings of patterns contribute to the strength of the pattern boundary profile*.

In the example of Figure 2 (further examples not included in this paper due to paper length restrictions) the adjusted PAT model detects correctly the beginning of the repeated phrases (the initial PAT model inserts spurious peaks at the endings of the exactly repeating parts of the phrases). It should be noted that in this example the repeated phrases are 5 (i.e. 3+2) bars long which is very unusual – the positions given by local segmentation processes are wrong – Lerdahl and Jackendoff (1983, p.206) take this 5 bar length grouping structure for granted (no systematic procedure for detecting it is given).

The pattern boundary detection model, as described to this point, can discover repeating patterns in the diatonic pitch interval domain that may or may not diverge towards their endings (patterns may be transposed). What happens if some intervals are not exactly the same (as the first intervals of the repeating phrases in Figure)? How can rhythmic information be also taken into account?

It is suggested that a more abstract representation for pitch intervals may be useful, such as a step-leap profile, especially if it is coupled with duration information. The step-leap encoding comprises of 5 distinct symbols (+step, +leap, -step, -leap, same) which is a rather too limited alphabet. If it is combined with duration symbols (or duration ratios) then the alphabet becomes rich enough to capture all the necessary information so that the pattern boundary detection model may operate effectivelly. In this encoding each interval of a melody is represented as a tuple [step-leap interval, duration ratio].

This further adjustment to the model enables it to segment correctly more difficult cases such as the one depicted in Figure 3 giving correct results, at the same time, for all the previous cases studied in this paper.

4. FURTHER IMPROVEMENTS

The computational attempt presented in this paper for capturing melodic similarity with a view to achieving melodic segmentation is still a long way from providing a robust, flexible and general model of melodic parallelism. It does, however, show its potential and further research is necessary to improve the model and also to evaluate it on a much larger scale.

5. REFERENCES

1. Cambouropoulos E. (2001a) The Local Boundary Detection Model (LBDM) and its application in the study of expressive timing. In *Proceedings of ICMC01*, La Havana, Cuba.
2. Cambouropoulos, E. (2001b) Melodic Cue Abstraction, Similarity and Category Formation: A Formal Model. *Music Perception* 18(3):347-370.
3. Cambouropoulos, E. (1999) Pattern Processing in Melodic Sequences: Challenges, Caveats and Prospects. *Computers in the Humanities*, 35(1):9-21.
4. Cambouropoulos, E. (1998) Musical Parallelism and Melodic Segmentation. In *Proceedings of the XII Colloquium of Musical Informatics*, Gorizia, Italy.
5. Cambouropoulos, E. (1997) Musical Rhythm: Inferring Accentuation and Metrical Structure from Grouping Structure. In *Music, Gestalt and Computing - Studies in Systematic and Cognitive Musicology*. M. Leman (ed.), Springer-Verlag, Berlin.
6. Cambouropoulos, E. and Widmer, G. (2000) Automated Motivic Analysis Via Melodic Clustering. *Journal of New Music Research* 29(4): 303-318.
7. Crawford, T., Iliopoulos, C.S. and Raman, R. (1998) String Matching Techniques for Musical Similarity and Melodic Recognition. *Computing in Musicology*, 11:71-100.
8. Crochemore, M. (1981) An Optimal Algorithm for Computing the Repetitions in a Word. *Information Processing Letters*, 12(5):244-250.
9. Iliopoulos, C.S., Moore, D.W.G. and Park, K. (1996) Covering a String. *Algorithmica*, 16:288-297.
10. Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music*, MIT Press, Cambridge (Ma).
11. Lidov, D. (1979) Structure and Function in Musical Repetition. *Journal of the Canadian Association of University Schools of Music* 8(1):1-32.
12. Rolland, P.Y. (1999) Discovering Patterns in Musical Sequences. *Journal of New Music Research* 28(4):334-350.
13. Rolland, P.Y. (2001). FlExPat: Flexible Extraction of Sequential Patterns. In *Proceedings of the IEEE International Conference on Data Mining (IEEE ICDM'01)*. San Jose - Silicon Valley, California, USA.
14. Rolland, P.Y., Ganascia, J.G. (1999). Musical Pattern Extraction and Similarity Assessment. In Miranda, E. (ed.) *Readings in Music and Artificial Intelligence*. Gordon & Breach - Harwood Academic Publishers, London.
15. Temperley, D. (2001) The Cognition of Basic Musical Structures. MIT Press, Cambridge (MA).

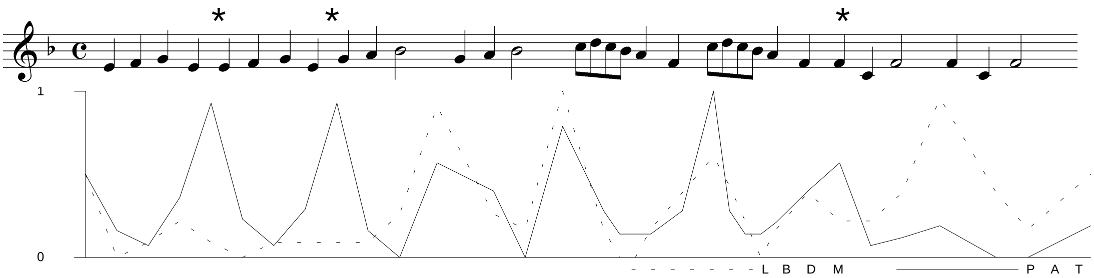

Figure 1 *Frère Jacques* - Segmentation profile according to the Local Boundary Detection Model (LBDM) and the Pattern Boundary Detection Model (PAT) for the *diatonic pitch interval profile* – local maxima indicate points of segmentation (N.B. strong pattern boundaries are detected at the points indicated by asterisks where no local boundaries are discovered by *LBDM*)

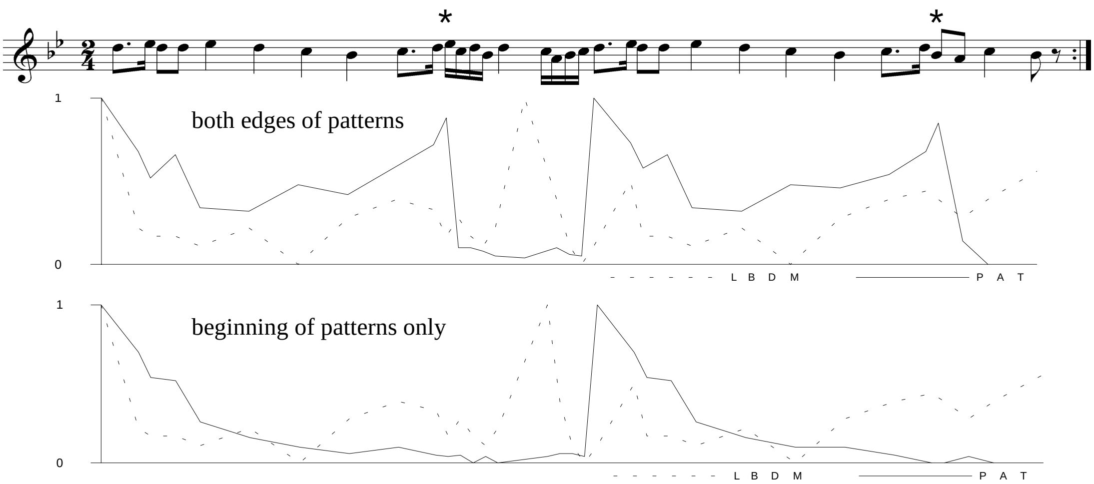

Figure 2 *Chorale St. Antoni* (arranged by Brahms in his *Haydn Variations op.56*). Segmentation profile according to LBDM and PAT for the diatonic pitch interval profile – N.B. the strong pattern boundaries that indicate the end points of the exactly repeating parts of the two phrases (indicated by asterisks) are eliminated in the version of the model that takes into account only the beginnings of patterns.

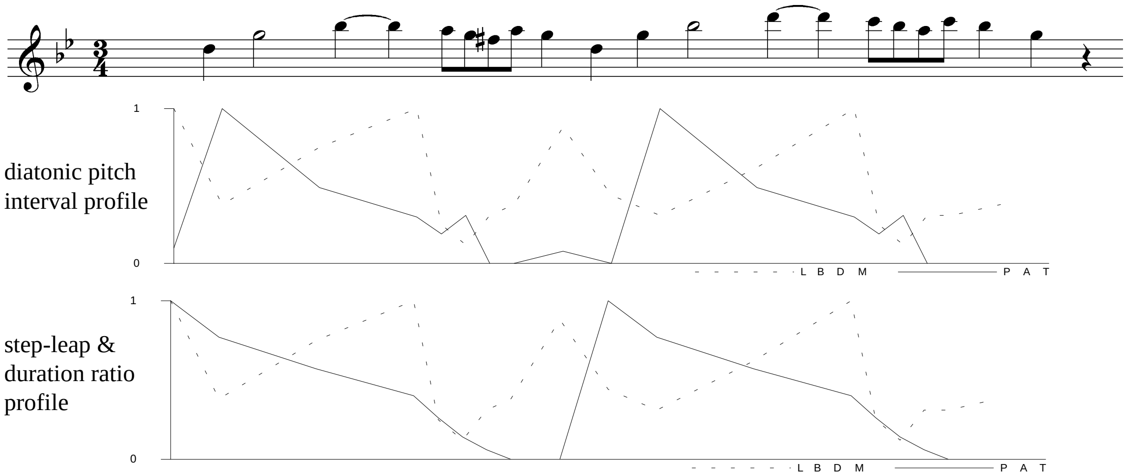

Figure 3 *Theme of section III of Mozart's G minor Symphony K550*. Segmentation profile according to LBDM and the Pattern Boundary Detection Model (PAT), firstly, for the diatonic pitch interval profile and, secondly, for the combined step-leap and duration ratio profile. The diatonic pitch interval matching fails as the first interval of the repeating phrase is a 3rd interval rather than a 4th interval – the combined step-leap and duration ratio encoding enables the correct segmentation of the two phrases - local boundaries are not capable of providing a correct segmentation.