INFLUENCE OF MUSICAL SIMILARITY ON MELODIC SEGMENTATION: REPRESENTATIONS AND ALGORITHMS

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

Costas Tsougras
Department of Music Studies
Aristotle University of Thessaloniki
tsougras@mus.auth.gr

ABSTRACT

In this paper a computational model is presented that extracts patterns from a given melodic surface and, then, 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 pattern length. 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 proposed melodic representations and algorithms is tested against a series of melodic examples.

1. INTRODUCTION

Musical similarity not only establishes relationships between different musical entities (such as rhythmic and melodic motives, themes and variations, harmonic progression groups etc.) but also enables - in the first place - the definition of such entities by directly contributing to the segmentation of a musical surface into meaningful units.

Models of melodic segmentation are often based on local Gestalt-based factors that identify points of local maximal change in various musical parameters such as IOIs, pitch intervals, dynamic changes and so on. Higher-level processes, however, play an important role as well. In this study, a central assumption is that similar musical patterns tend to be highlighted and perceived as units/wholes whose beginning and ending points influence the segmentation of a musical surface.

Pattern-matching techniques have been employed in attempts to formalise musical similarity. There have been, however, relatively few attempts to tackle the difficult issue of pattern extraction (i.e. extracting important patterns in one or more musical sequences). Overviews of the application of pattern processing algorithms on musical strings can be found in (Crawford et al. 1998; Rolland et al. 1999; Cambouropoulos et al. 2001; Meredith et al. 2002). Recent research that directly links pattern extraction to melodic segmentation via memory-based modeling is presented by Ferrand et al. (2003) and Bod (2001); additionally, Temperley et al. (2002) link parallelism to metrical structure and, indirectly, to melodic segmentation.

The aim of this paper is to examine the relation between musical parallelism and segmentation via computational modeling. This study does not provide a comprehensive stand-alone computer program for melodic segmentation; it rather explores melodic surface representation issues and issues relating to the pattern extraction mechanism itself through the application of a series of different representations and algorithm variants on a number of progressively more ‘difficult’ melodic examples. The main goal is not to provide a comprehensive solution to the problem of melodic parallelism but rather to shed light on various aspects of the problem and to enable a better understanding of it. Throughout the paper a number of melodic examples illustrate the strengths and weaknesses of the overall approach. The current study is a continuation of the earlier research presented in (Cambouropoulos 1998, 2003).

2. SEGMENTATION AND MUSICAL SIMILARITY

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 model is the *Local Boundary Detection Model (LBDM)* proposed by (Cambouropoulos 2001a). Higher-level processes, however, also affect the segmentation of a musical surface. 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 is herein assumed that 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.

From a cognitive point of view, it is suggested that elaborate pattern extraction processes are more likely to be applied on relatively short melodic excerpts due to the heavy computation involved. Such activity is usually more intense at the beginning of a musical piece/section where new musical materials are introduced and established. Once a number of such musical ideas have been extracted links to further new instances (varied or not) can be made more efficiently.

It is herein assumed that similarity processes for melodic segmentation tasks are confined 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.

The musical examples presented in this paper for testing the proposed algorithms have been selected on the basis that the segmentation process for these cases relies primarily on melodic parallelism and not on local detail grouping factors (actually, local gestalt-based factors provide clearly incorrect boundaries). It is common that these two segmentation components (i.e. local gestalt-based factors and parallelism) reinforce each other but, for the sake of clarity, examples have been selected that illustrate a conflict between the two approaches and a clear predominance of the parallelism factor.

In this paper, the pattern extraction algorithm is applied at parametric profiles of the melodic surface for pitch intervals (diatonic intervals, a step-leap representation and some further more refined representations) and for interonset intervals (IOI ratios). An important aspect of the paper is to discover which of these parameters (or combination of them) is more appropriate for the segmentation task and to show how a 'balanced' representation that is neither too specific nor too general may yield better results in more cases. It should be noted, however, that the issue of representation is examined primarily in order to show how important it is and how better representations can be devised rather than to propose a 'best' solution.

3. AN ALGORITHM FOR SEGMENTATION VIA PATTERN EXTRACTION

3.1. The PAT Algorithm

The pattern extraction model (see Appendix), which 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)
A pattern extraction algorithm 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 at 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 for which local maxima appear are more likely to be perceived as boundaries because of musical similarity.

Table 1

In the melodic example of Frère Jacques (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.

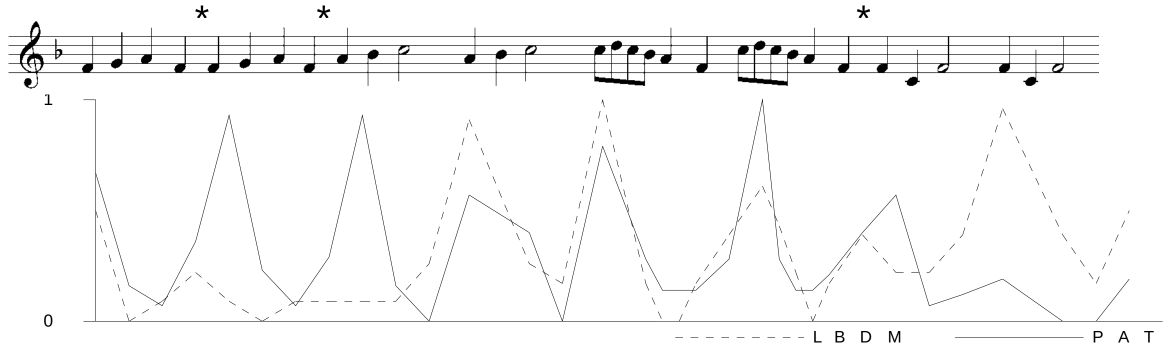


Figure 1 *Frère Jacques* - Segmentation profile according to the Local Boundary Detection Model (LBDM) and the Pattern Boundary Detection Model (PAT) for the *diatonic pitch interval profile* – local maxima indicate positions that may be considered as points of segmentation.

3.2. The PAT algorithm (revised)

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 Lerdahl and Jackendoff, 1983, p.51). 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 the revised PAT model detects correctly the beginning of the repeated phrases (the initial PAT algorithm inserts spurious peaks at the endings of the exactly repeating parts of the phrases but the revised PAT algorithm correctly identifies the beginning of the second phrase).

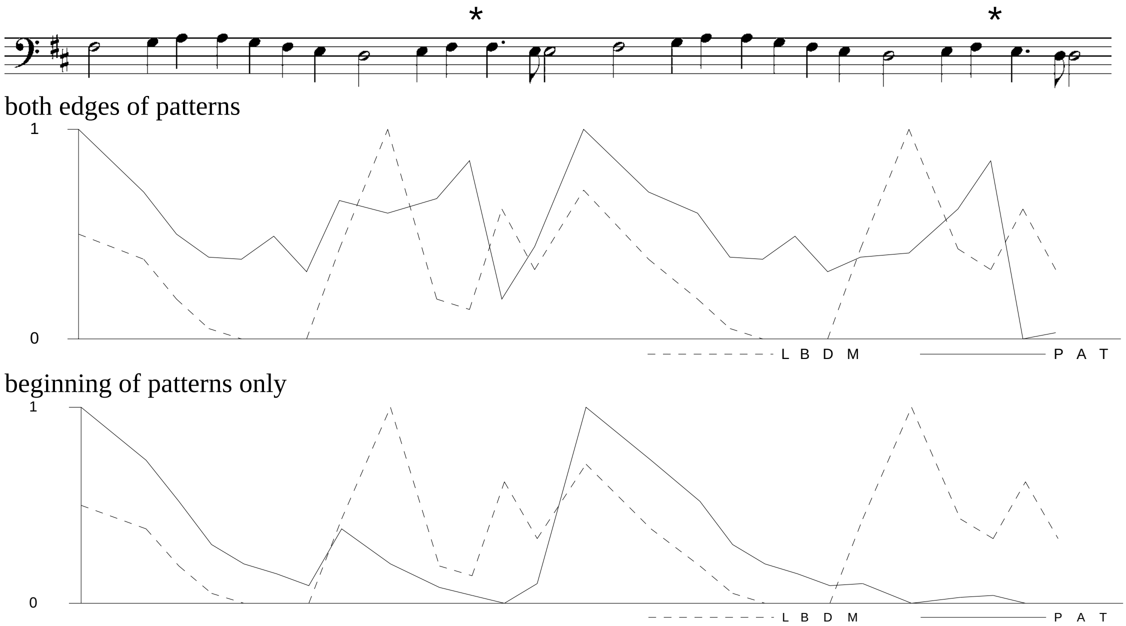


Figure 2 Beginning of the finale theme of Beethoven's 9^{th} *Symphony*- Segmentation profile according to the Local Boundary Detection Model (LBDM) and the Pattern Boundary Detection Model (PAT) for the diatonic pitch interval profile. The strong pattern boundaries that indicate the end points of the 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.

4. REPRESENTATION OF THE MELODIC SURFACE

4.1. Abstract representation

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, for instance, the first intervals of the repeating phrases in Figure 3)? 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 the 'refined contour' representation (or step-leap representation), especially if it is coupled with duration information. The step-leap encoding comprises of 5 distinct symbols (+step, +leap, -step, -leap, same) whereby an interval of one diatonic step (i.e. minor or major 2nd) is a 'step', a larger interval is a 'leap' and an interval of zero steps (between repeating notes) is 'same' - this 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 the previous examples presented in this paper.

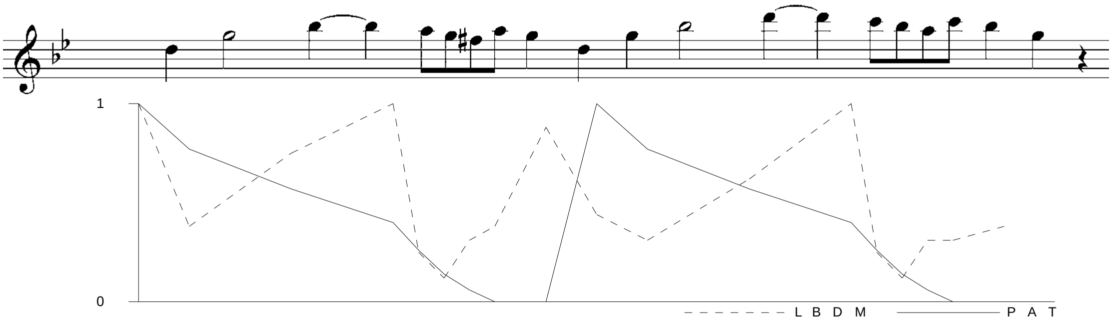


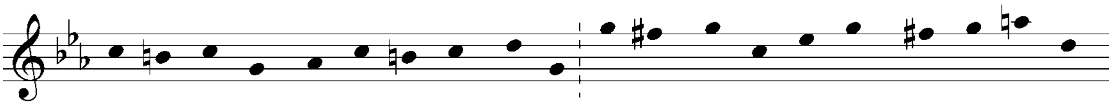
Figure 3 Theme of Mozart's *G minor Symphony K550,* section III. Segmentation profile according to LBDM and the PAT model for the combined step-leap and duration ratio profile. The diatonic pitch interval matching fails as the first interval of the repeating phrase is a 3rd rather than a 4th interval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

4.2. A variant of the PAT algorithm for further flexibility

Approximation can be introduced into an exact pattern-matching process by using a more abstract representation at the level of the initial string of symbols. For instance, a pitch interval representation such as the *step/leap* representation (or even *step/small-leap/medium-leap/large-leap* etc.) allows different size leaps to be matched. A problem, however, is that the abstract categories in the representation have sharp boundaries and no instance may belong to more than one category - this way, borderline members can never be matched to other 'similar' members of other categories (e.g. a 3rd interval as a member of *leap* can never be matched to a 2nd interval which is a *step*).

Consider, for instance, the sequence of pitch intervals in figure 4. The *step-leap* representation allows the extraction of the two different underlined patterns (see representation A in figure 4). A musician, however, would consider the second half of the sequence as a (near-exact) repetition of the first half (the pitches of this example are taken from Bach's *Well-Tempered Clavier, Book I, Fugue in F# major* – see figure 7). This match can be achieved only if the first 3rd interval in the second half of the pitch sequence can be matched with the corresponding 2nd interval of the first half.

An abstract symbolic representation can become more flexible in terms of category gradedness and membership if instances are allowed to be members of more than one category. In the following examples, a 3rd interval is allowed to be an instance of either *step* or *leap (s/l)* - see representation B in Figure 4. Whichever of the alternative abstractions (step or leap) allows the longest patterns to emerge is selected (the first 3rd interval of the second half of the melody is taken to be a member of *step* and is thus matched to the corresponding 2nd interval of the first half as this gives a longer melodic repetition).



Rep A: -s s -l s l -s s s -l l -s s -l l l -s s s -l

Rep B: -s s -l s s/l -s s s -l l -s s -l s/l s/l -s s s -l

Figure 4 The *step-leap* representation allows the extraction of two patterns repeating twice each (single- and double-line underlined patterns in representation A). The proposed representation that allows overlapping of pitch categories - in this case, a 3rd interval can be a member of either *step* or *leap (s/l)* - allows the matching of the second half of the sequence to the first half (see rep. B).

The case where a 2nd and a 3rd interval should be considered similar is not simply some rare exception in music. It is a rather common phenomenon especially when themes appear in their dominant form (see, for instance, the tonal answers of almost half of Bach's fugue themes from the two books of the *Well-Tempered Clavier*). In figures 5, 6 & 7 some examples are presented (NB: Bach fugue themes and their tonal aswers are presented as belonging to the same auditory stream - this is not musically correct but is not cognitively implausible – a streaming algorithm could generate tentative streaming options including the ones presented in the examples).

The exact pattern matching algorithm (described above) that extracts all repeating patterns can be adjusted so as to cope with representations that allow alternative symbols for elements of the initial string (ongoing research is examining issues of algorithm efficiency). Examples of the application of the new version of the PAT algorithm are given in Figures 5, 6, & 7. With this new version of the pattern extraction algorithm it is now possible to adopt more elaborate representations of the melodic surface that allow overlapping among abstract categories (such as the more 'sophisticated' pitch interval representation shown in Figure 7).

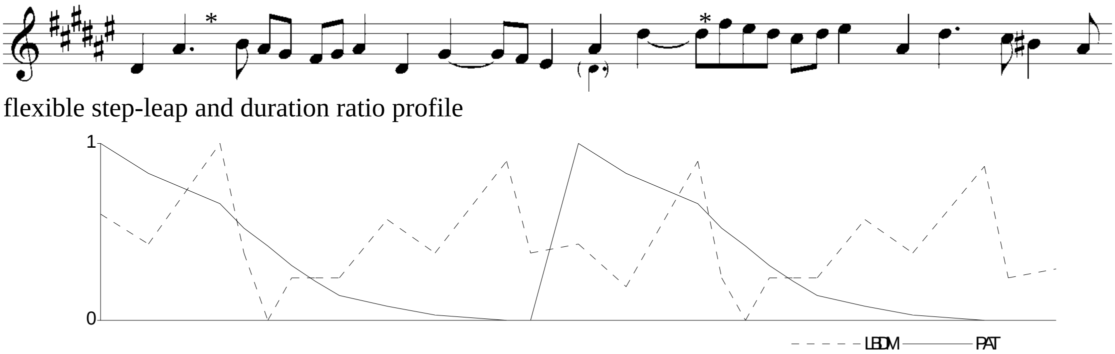


Figure 5 Upper voice (theme and tonal answer as one melodic 'stream') from the opening of Bach's *Well-Tempered Clavier, Book I, Fugue in F# major*. Segmentation profile according to LBDM and the PAT variant for the combined step-leap and duration ratio profile that allows additionally a 3rd interval to be a member of either *step* or *leap* (the repeated pattern is correctly identified).

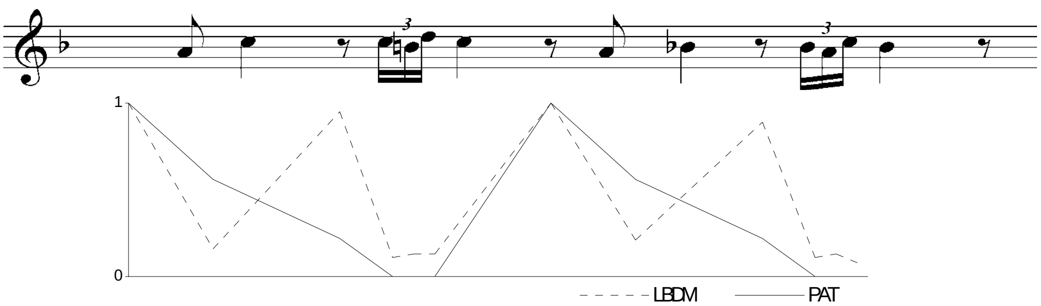


Figure 6 Opening melody of Beethoven's *Piano Sonata Op.10, No.2.* Segmentation profile according to LBDM and the Pattern Boundary Detection Model (PAT) variant for the combined step-leap and duration ratio profile that allows additionally a 3rd interval to be a member of either *step* or *leap*.

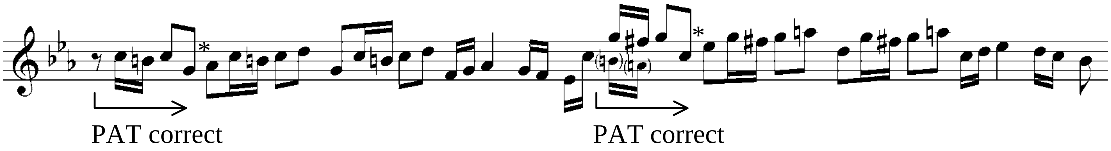


Figure 7 Upper voice 'stream' (theme and tonal answer) from the opening of Bach's *Well-Tempered Clavier, Book I, Fugue in C minor*. The Pattern Boundary Detection Model (PAT) variant detects correctly the beginning of the repetition (tonal answer) for the combined step-leap and duration ratio profile that allows additionally a 3rd interval to be a member of either *step* or *leap* (NB: the two intervals indicated by the asterisks are matched).

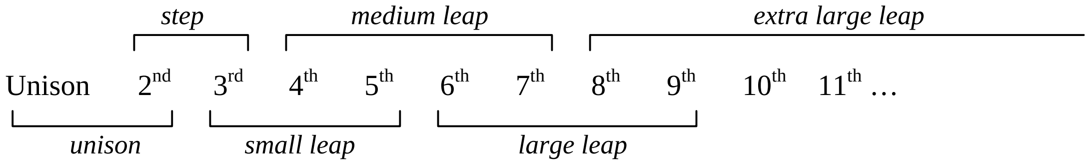


Figure 8 A possible abstract representation for pitch intervals – in this representation overlap between categories is allowed. Such a representation could prove to be more powerful than the more standard step-leap or contour representations as it allows rather high discriminability between intervals and also significant flexibility. It was tested on all the examples in this paper giving correct results.

5. CONCLUSION

Overall, the methods and results presented in this study provide an informing attempt to address formally the difficult issue of musical parallelism and its links to melodic segmentation. The examples against which the proposed algorithm was tested were known to pose serious problems for local detail grouping algorithms – additionally these examples contain gradually more difficulties regarding melodic similarity. The proposed model is quite successful in tackling all these problems yet it requires further experimentation and development before it can be integrated in a comprehensive model of melodic segmentation.

6. REFERENCES

Bod, R. (2001). Memory-Based Models of Melodic Analysis: Challenging the Gestalt Principles. *Journal of New Music Research*, 30(3).

Cambouropoulos E. (2003) Musical Pattern Extraction for Melodic Segmentation. In *Proceedings of ESCOM5*, Hanover, Germany.

Cambouropoulos E. (2001a). The *Local Boundary Detection Model (LBDM)* and its application in the study of expressive timing. In *Proceedings of the International Computer Music Conference (ICMC01)*, Havana, Cuba.

Cambouropoulos, E. (2001b). Melodic Cue Abstraction, Similarity and Category Formation: A Formal Model. *Music Perception* 18(3):347-370.

Cambouropoulos, E., Crawford, T. and Iliopoulos C.S. (2001). Pattern Processing in Melodic Sequences: Challenges, Caveats and Prospects. *Computers in the Humanities*, 35(1):9-21.

Cambouropoulos, E. (1998). Musical Parallelism and Melodic Segmentation. In *Proceedings of the XII Colloquium of Musical Informatics*, Gorizia, Italy.

Cambouropoulos, E. and Widmer, G. (2000). Automated Motivic Analysis Via Melodic Clustering. *Journal of New Music Research* 29(4): 303-318.

Crawford, T., Iliopoulos, C.S. and Raman, R. (1998). String Matching Techniques for Musical Similarity and Melodic Recognition. *Computing in Musicology*, 11:71-100.

Crochemore, M. (1981). An Optimal Algorithm for Computing the Repetitions in a Word. *Information Processing Letters*, 12(5):244-250.

Ferrand, M. and Nelson, P. (2003). Unsupervised Learning of Melodic Segmentation: A Memory-based Approach. In *Proceedings of the 5th Triennial ESCOM Conference*, Hanover, Germany.

Iliopoulos, C.S., Moore, D.W.G. and Park, K. (1996) Covering a String. *Algorithmica*, 16:288-297.

Lerdahl, F. and Jackendoff, R. (1983). *A generative Theory of Tonal Music*, The MIT Press, Cambridge (Ma).

Lidov, D. (1979). Structure and Function in Musical Repetition. *Journal of the Canadian Association of University Schools of Music* 8(1):1-32.

Meredith, D., Lemström, K. and Wiggins G.A. (2002). Algorithms for Discovering Repeated Patterns in Multidimensional Representations of Polyphonic Music. *J. of New Music Research*, 31(4):321-345.

Rolland, P.Y. and 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.

Temperley, D. and Bartlette, C. (2002). Parallelism as a Factor in Metrical Structure. *Music Perception*, 20(2):117-149.

Melodic segmentation performed on symbolic representations of music (scores) is part of many music information and retrieval tasks. Large music databases, such as RÃ©per-toire International des Sources Musicales (RISM), often use the first phrase of a com-position as an identifying label.Â The value of the proposed algorithm, and its relevance in terms of precision and recall, will only become clearer after its testing on musical data. Bibliography. ALTMANN, Gabriel â€“ SCHWIBBE, Michael.Â CAMBOUROPOULOS, Emilios. Musical parallelism and melodic segmentation: A computational approach. Music Perception: An Interdisciplinary Journal, 2006, 23(3), pp. 249â€“268. CAMBOUROPOULOS, Emilios. â€œStructural Representations of Musical Pitch.â€ In The Psychology of Music, pp. 343-390. Academic Press.Â â€œChord Recognition with Application in Melodic Similarity.â€ In additional proceedings (online) of the 2nd ICMAI. Winograd, Terry (1968), â€œLinguistics and the Computer Analysis of Tonal Harmony.â€Â â€œMelodic Segmentation: Evaluating the Performance of Algorithms and Musical Experts.â€ In Proceedings of ICMPC 2002, p.7-14. Time Structures.