

Melodic Pattern Segmentation of Polyphonic Music as a Set Partitioning Problem

Tsubasa Tanaka and Koichi Fujii

Abstract In polyphonic music, melodic patterns (motifs) are frequently imitated or repeated, and transformed versions of motifs such as inversion, retrograde, augmentations, diminutions often appear. Assuming that economical efficiency of reusing motifs is a fundamental principle of polyphonic music, we propose a new method of analyzing a polyphonic piece that economically divides it into a small number types of motif. To realize this, we take an integer programming-based approach and formalize this problem as a set partitioning problem, a well-known optimization problem. This analysis is helpful for understanding the roles of motifs and the global structure of a polyphonic piece.

1 Motif Division

In polyphonic music like fugue-style pieces or J.S. Bach's *Inventions and Sonatas*, melodic patterns (motifs) are frequently imitated or repeated. Although some motifs are easy to find, others are not. This is because they often appear implicitly and/or appear in the transformed versions such as inversion, retrograde, augmentations, diminutions. Therefore, motif analysis is useful to understand how polyphonic music is composed.

Simply speaking, we can consider the motifs that appear in a musical piece to be economical if the number of types of motif is small, the number of repetitions is large, and the lengths of the motifs are long. Assuming that this economical efficiency of motifs is a fundamental principle of polyphonic music, we propose a new method of analyzing a polyphonic piece that efficiently divides it into a small number of types of motif. Using this division, the whole piece is reconstructed with a

Tsubasa Tanaka
Tokyo University of the Arts, e-mail: tsubasa@y4.dion.ne.jp

Koichi Fujii
NTT DATA Mathematical Systems Inc. e-mail: fujii@msi.co.jp

small number of types of motif like a video game puzzle *Tetris*[1] (In tetris, certain domains are divided with only seven types of piece). We call such a segmentation a *motif division*.

If a motif division is accomplished, it provide us a simple and higher-level representation whose atom is a motif, not a note and it will be helpful to clarify the structures of polyphonic music. The representation may provide knowledge about how frequent and where each motif is used, the relationships between motifs such as causality and co-occurrence, which transformations are used, how the musical form is constructed by motifs, and how the long-term musical expectations are formed. This analysis may be useful for applications such as systems of music analysis, performance, and composition.

Studies about finding boundaries of melodic phrases are often based on human cognition. For example, [2] is based on grouping principles of gestalt psychology, and [3] is based on short memory model. While these studies deal with relatively short range of perception and require small amounts of computational time, we focus on global configuration of motifs on the level of compositional planning. This requires us to solve an optimization problem that is hard to solve. To deal with this difficulty, we take an integer programming-based approach [4] and show that this problem can be formalize as a set partitioning problem [5]. This problem can be solved by IP-solvers that use efficient algorithms such as the branch and bound method.

2 Transformation Group and Equivalence Classes of Motif

In this section, we introduce equivalence classes of motif derived from a group of motif transformations as the criterion for the identicalness of motifs. These equivalence classes are used to formulate the motif division in Sect. 3.

Firstly, a motif is defined as an ordered correction of notes $[N_1, N_2, \dots, N_k]$ ($k > 0$), where N_i is the information for the i th note, comprising the combination of the pitch p_i , start position s_i , and end position e_i ($N_i = (p_i, s_i, e_i)$, $s_i < e_i \leq s_{i+1}$). Next, let \mathcal{M} be the set of every possible motif, and let T_p, S_t, R, I, A_r be one-to-one mappings (transformations) from \mathcal{M} to \mathcal{M} , where T_p is the transposition by pitch interval p , S_t is the shift by time interval t ($p, t \in \mathbb{R}$), R is the retrograde, I is the inversion, and A_r ($r > 0$) is the r -fold argumentation (diminution, in the case of $0 < r < 1$). These transformations generate a transformation group \mathcal{T} whose operation is the composition of two transformations and whose identity element is the transformation that does noting. Each transformation in \mathcal{T} is a *strict imitation* that preserves the internal structures of the motifs.

Here, a binary relation between a motif m ($\in \mathcal{M}$) and $\tau(m)$ ($\tau \in \mathcal{T}$) can be defined. Due to the group structure of \mathcal{T} , this relation is an equivalence relation (i.e., it satisfies reflexivity, symmetry and transitivity[6]). Then, it derives equivalence

classes in \mathcal{M} . Because the motifs that belong to a same equivalence class share the same internal structure, they can be regarded as identical (or the same type)¹.

3 Formulation as a Set Partitioning Problem

A Set partitioning problem, which is well known in the context of operations research, is an optimization problem defined as follows. Let N be a set that consists of n elements $\{N_1, N_2, \dots, N_n\}$, and let M be a family of sets $\{M_1, M_2, \dots, M_m\}$, where each M_j is a subset of N . If $\bigcup_{j \in X} M_j = N$ is satisfied, X , a subset of indexes of M , is called a cover, and the cover X is called a partition if $M_{j_1} \cap M_{j_2} = \emptyset$ is satisfied for different $j_1, j_2 \in X$. If a constant c_j called a cost is defined for each M_j , the problem of finding a partition X that minimizes the sum of the costs $\sum_{j \in X} c_j$ is called a set partitioning problem.

3.1 Condition of motif division

If N_i corresponds to each note of a musical piece to be analyzed and M_j corresponds to a motif, the problem to find the most economically efficient motif division can be interpreted as a set partitioning problem. The index i starts from the first note of a voice to the last note of the voice and from the first voice to the last voice. $M_j (1 \leq j \leq m)$ corresponds to $[N_1]$, $[N_1, N_2]$, $[N_1, N_2, N_3]$, \dots , $[N_2]$, $[N_2, N_3]$, \dots in this order. The number of notes in a motif is less than a certain limit number (Fig. 1).



Fig. 1 Possible motifs of the first voice of J.S. Bach's *Invention No. 1*. (In the case where the maximum number of notes in a motif is 4.)

This information can be represented by the following matrix A :

¹ Although the criterion for identical motifs defined here only deals with strict imitations, we can define the criterion in different ways to allow more flexible imitations, such as by (1) defining an equivalence relation from the equality of a shape type [7, 8, 9, 10] and (2) defining a similarity measure and performing a clustering of motifs using methods such as k-medoid method [11] (the resulting clusters derive an equivalence relation). In any case, making equivalence classes from a certain equivalence relation is a versatile way to define the identicalness of the motifs.

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} \quad (1)$$

, where each row corresponds to each note N_i and each column corresponds to which notes are covered by each motif M_j . This matrix is the case where the maximum number of notes in a motif is 4. Representing the element of A as a_{ij} , the condition that the whole piece is exactly divided by a set of selected motifs can be described by the following constraints that mean each note N_i is covered by one of M_j once and only once:

$$\forall i \in \{1, 2, \dots, n\}, \sum_{j=1}^m a_{ij} x_j = 1, \quad (2)$$

where x_j is a 0-1 variable that means whether or not M_j is used in motif division. These conditions are equivalent to the condition of partitioning.

3.2 Objective Function

The purpose of motif division is to find the most efficient solution from the many solutions that satisfy the condition of partitioning. Then, we must define efficiency of motif division. It is considered that the average length (the number of notes) of motifs used in the motif division is one of the simplest barometers that represent the efficiency of motif division. Also, the number of motifs and that of the types of motif used in motif division will be efficient if they are small. In fact, the average length of motifs is inversely proportional to the number of motifs. Therefore, if the number of types of motif is fixed, the number of motifs will be what we should minimize.

The number of motifs can be simply represented by $\sum_{j=1}^m x_j$. This is the cost function $\sum_{j=1}^m c_j x_j$ whose c_i is 1 for each i . We adopt this cost function. However, in the next subsection, we introduce additional variables and constraints to fix the number P , which is the number of equivalence classes that are used in motif division.

3.3 Controlling the Number of Equivalence Classes

Let C be the set of equivalence classes of motif, which is derived from M , which is the set of all possible motif classes that can be found in a piece (only the motif classes whose number of notes is less than a certain number is included in M).

This means that M is derived from \mathcal{M} by a restriction. Let y_k be a 0-1 variable that represents whether or not one of the members of C_k appears in X (the set of selected motifs), where each element of C is denoted as $C_k (1 \leq k \leq l)$. This means that statement “ $y_k = 1 \Leftrightarrow \sum_{j \in C_k} x_j > 0$ ” must be satisfied. This statement can be represented by the following constraints that use $\sum_{j \in C_k} x_j$, the number of selected motifs that belong to C_k :

$$\forall k \in \{1, 2, \dots, l\}, y_k \leq \sum_{j \in C_k} x_j \leq Qy_k, \quad (3)$$

where Q is a constant that is sufficiently large. Then, the statement that the number of equivalence classes is P can be represented by the following constraint:

$$\sum_{k=1}^l y_k = P. \quad (4)$$

If P is small to a certain degree, the motif division will tend to be simple. However, if P is too small, covering whole piece with few motif classes will be difficult and one note motif will be used too many times. This will lead to lose the efficiency of motif division. Therefore, we should find good balance between the smallness of the objective function and the smallness of P . Because knowing which number is adequate for P in advance is difficult, we will solve the optimization problems for respective P in a certain range. Then, we will find an adequate number for P , observing the solutions for respective P .

4 Result

We analyzed J.S. Bach’s *Invention No. 1* by solving the optimization problem described in the previous section. The maximum length of motif was set as 7. An integer programming solver *Numerical Optimizer 16.1.0.* and a branch and bound method was used for searching the solution. From the observation of solutions for various values for P , P was set as 13. It took less than one minute to obtain a solution for $P = 13$. Fig.3 shows the result of motif division. The slurs represent motifs and the one-note motifs don’t have a slur. Fig.3 shows the representatives of 13 motif classes that are used in the motif division.

This result tells us many things. For example, 4th, 10th, and 11th motif classes in Fig.3 are slightly different but can be regarded as the same motif, which corresponds to the subject of this piece. Searching for the domains where the subject doesn’t appear, we find that there are three domains whose duration are one and half bars. The ends of these domains coincide with the places where the cadences exist. Therefore, we could detect three sections of this piece properly.

The last motif class in Fig.3 is a leap of octave. This motif class appears in all of the cadence domains and it is related to the ends of sections. It also co-occurs with 2nd motif class, which is a two-notes motif, in the cadence domains.

The image displays seven systems of musical notation for J.S. Bach's Invention No. 1. Each system consists of a treble clef staff and a bass clef staff. The music is written in C major and 3/4 time. The notation includes various rhythmic patterns, such as eighth and sixteenth notes, and melodic lines. Some measures contain rests, indicated by a 'z' symbol. The score is divided into measures, with some measures containing rests. The notation includes various rhythmic patterns and melodic lines. Some measures contain rests, indicated by a 'z' symbol.

Fig. 2 The representatives of motif classes that appear in the motif division of J.S. Bach's *Invention No. 1*, in the case that $P = 13$. Some flats are replaced by sharps for the purpose of programming.

The 12th motif class is a very characteristic one that includes a dotted note and large leap. This motif class only appears before the cadence domains. We can consider that this remarkable motif class plays an important role that tells listeners the end of the exposition of subject and the beginning of the cadence domain.



Fig. 3 The motif classes that appear in the motif division of J.S. Bach's "Invention No. 1" (the number of motif classes was set at 13).

The 9th zigzag motif class and the motif classes that are one-way slow movements shown by the arrows in Fig.3 only appear as the ascending form in the first 2 sections. In the final section, contrary, these motif classes appear only as the descending form. We interpret this contrast means that the ascending form creates a sense of continuation of the piece and the descending form creates a sense of conclusion. Thus, long-term musical expectations seems to be formed by the selections of transformation.

In such ways, motif division is useful to make us understand the roles of motifs and how the global musical structures are formed.

5 Conclusion

In this paper, we formulated the problem of motif division, which decompose polyphonic music into a small number of motif classes, as a set partitioning problem, and we obtained the solution using an IP-solver. It was shown that the motif division provides useful information to understand the roles of motifs and how the global musical structures are constructed from the motifs.

Future tasks include construction of a program that automatically analyze the global structures utilizing the obtained motifs and automatic composition of new pieces that use the same motifs as the original piece using the result of the analysis program. To create a criterion for determining adequate value of P automatically is also a remaining problem.

Acknowledgements This work was supported by JSPS Postdoctoral Fellowships for Research Abroad.

References

1. <http://tetris.com>
2. Cambouropoulos, E.: The Local Boundary Detection Model (LBDM) and its Application in the Study of Expressive Timing, Proc. ICMC, pp.290–293 (2001).
3. Ferrand, M., Nelson, P. and Wiggins, G.: Memory and Melodic Density: A Model for Melody Segmentation, Proc. the XIV Colloquium on Musical Informatics, pp.95–98 (2003).

4. Nemhauser, G.L. and Wolsey, L.A.: *Integer and Combinatorial Optimization*, Wiley (1988).
5. Balas, E. and Padberg, M.W.: *Set Partitioning: A survey*, *SIAM Review* 18.4, pp.710-760 (1976).
6. Armstrong, M.A.: *Groups and Symmetry*, Springer (2012).
7. Buteau, C and Mazzola G.: *From Contour Similarity to Motivic Topologies*, *Musicae Scientiae* Fall 2000 vol.4, no.2, pp.125-149 (2000).
8. Mazzola, G., et al.: *The Topos of Music: Geometric Logic of Concepts, Theory, and Performance*, Birkhäuser (2002).
9. Buteau, C.: *Topological Motive Spaces, and Mappings of Scores Motivic Evolution Trees*. In: Friepertinger, H., Reich, L. (eds.) *Grazer Mathematische Berichte, Proceedings of the Colloquium on Mathematical Music Theory*, pp.27—54 (2005)
10. Buteau, C.: *Melodic Clustering Within Topological Spaces of Schumann ' s Träumerei*, *Proceedings of ICMC*, pp.104-110 (2006).
11. Bishop, C.: *Pattern Recognition and Machine Learning*, Springer, pp.423-430 (2006).