A Novel Local Alignment-Based Approach to Motif Extraction in Polyphonic Music

Tiange Zhu[1], Danny Diamond[2], James McDermott[2], Raphaël Fournier-S'niehotta[1], Mathieu Daquin[3], and Philippe Rigaux[1] ⋆

[1] CNAM Paris
[2] University of Galway
[3] Université de Lorraine
tiange.zhu@lecnam.net
d.diamond1@nuigalway.ie

Abstract. The paper provides a novel approach to musicologically-informed intra-opus motif detection within polyphonic music scores. We extract diatonic interval sequences from each voice of a score; sequence segmentation is performed via pairwise local alignment between each pair of voices. From the output of this step, string-based approaches are used for motif discovery.
Specifically, a weighted directed acyclic graph is constructed, giving a custom measurement of motif importance. A selection and filtration procedure is applied according to a set of rules and music structural information, to generate a final selection of music motifs.
The ground truth annotated JKUPDD dataset is used for evaluation of the proposed methodology. The results demonstrate that this algorithm is capable of extracting musically meaningful motifs with high precision and recall.

Keywords: Music Information Retrieval, Pattern Discovery, Computational Musicology

1 Introduction

A musical motif is "*the smallest structural unit possessing thematic identity*" within a piece of music [1]. The detection of frequent musical patterns is a long-standing area of work in the field of Music Information Retrieval (MIR). The existing pattern discovery research covers both audio and symbolic music, adopting methods generally falling into three broad categories of 1) string or sequence-based [2, 3], 2) geometric pattern discovery, [4], and 3) machine-learning based methods [5].

⋆ This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004746 (Polifonia: a digital harmoniser for musical heritage knowledge, H2020-SC6-TRANSFORMATIONS).

The main goal of this work is to develop a methodology that, while informed by musicological knowledge, is not specialized to a single genre or musical tradition. A second aim is to produce a short and focused set of output motifs, reducing the requirement for time-intensive human validation of the results.

We designed a novel approach to achieve these aims, by working with interval sequences extracted from digital music scores, and segmenting a composition based on pairwise local alignments [6] between all possible voice pairs.Alignment has been commonly applied to the task of similarity between pieces of music [7, 8], but not in works on detection of local patterns. The outputs of the segmentation are taken as the input for a string-based motif discovery process. The overall importance of a pattern is measured based on its frequency of occurrence, using a graph that represents the relationship between patterns. The top-ranked patterns are further analyzed and filtered according to their musical (metrical) structure, generating a final set of motifs. In the context of this paper, motifs are defined as short recurring melodic patterns within a piece of music which contain important or characteristic thematic material; it must repeat at least two times throughout a composition, and contain at least three intervals.The proposed method is proven to generate satisfying results for an intra-opus pattern detection task based on the JKUPDD dataset [9], discussed in Section 4. The results exhibit a high degree of accuracy, broadly comparable to state-of-the-art pattern detection algorithms.

Identified motifs are of importance in use cases which range from thematic analysis of the piece of music, or musicological study of the body of work of a composer [10], to characterisation of a musical tradition, genre or period. Apart from being applied to polyphonic melodies as in this paper, the introduced methodology can also be applied to detect motifs between multiple related monophonic scores, which is potentially of use in the study of tune families or regional styles within folk traditions [11].

2 State of the art

Musical pattern detection tasks in MIR can be either "*intra-opus*" (within a single piece of music) and/or "*inter-opus*" (across multiple pieces of music). Input data is typically either audio or symbolic music representation. The following discussion mainly covers work on symbolic music inputs, with the exception of [12] and [13].

Pattern detection studies on symbolic music tend to break down into string-based, geometric or machine learning approaches. String-based pattern detection studies are the most common of all approaches. They range from *n*-grams and NLP-based work such as [14] to tree models of pattern relationships, subsuming and compressing many unique pattern instances to a smaller set of 'maximal' patterns. The latter approach to pattern detection has been influential on the work presented in this paper. It has most commonly been used in monophonic inter-opus applications [15, 16]; some polyphonic applications exist in the literature [2] but differ to our work significantly in the specific structural model applied.

Best-in-class geometric work includes the family of "point-set" geometric compression algorithms set out in [17, 18]. This family of algorithms have performed well on tasks ranging from intra-opus pattern detection in the JKUPDD dataset [17], to an intra-opus tune family classification task in [18]. Other interesting work, which both builds on and evaluates the "point-set" approach includes [4, 19] and [20].

Works based on machine-learning are increasingly prominent in recent years. Chai Wei [13] uses self-similarity and Dynamic Time Warping (DTW) to detect repeating structural sections in audio corpora. Unsupervised machine learning is adopted by Jacopo de Berardinis et al. [12] to build graph-based music structure hierarchies adapted for segment audio-derived feature sequences into structural sections. Matevž Pesek et al. [5] uses unsupervised machine learning, in order to construct a compositional hierarchical model for analysis and discovery of pattern in symbolic music.

3 Methodology

3.1 Framework

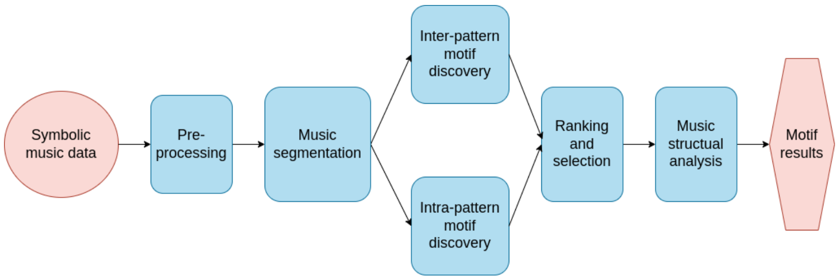

Fig. 1. A framework of motif discovery from polyphonic symbolic music

The framework of the proposed method for discovering motifs in polyphonic symbolic music is illustrated in Figure 1. Taking a symbolic music score as an input, key-invariant diatonic interval sequences are extracted from each voice, and encoded. Music segmentation is then applied to the encoded sequences via local alignment [6]. A set of patterns are gathered for further discovery, to find a set of potential motifs. The motifs are ranked and filtered based on specific rules and a customized measure of importance, and then analyzed by their music structure information, to select a final list of results.

3.2 Pre-processing

A polyphonic music score is taken as the input for pre-processing. Using music21 [21], we extract the melodic pitch sequence from each voice of the score, represented as a sequence of MIDI note numbers. From these pitch sequences we calculate diatonic intervals, then normalise them to the range of a single octave.

Definition 1 (Melody). *Let the kth voice in a score be v_k. Let $M(v_k) = [m_1, ...m_n]$ be the melody of v_k, in which m_i denotes the pitch of the ith note of v_k.*

Definition 2 (Diatonic Intervals). *Let $dia(\upsilon_k) = [d_1, ...d_n]$ be a sequence of diatonic intervals, d_i is the diatonic interval between two consecutive pitches, m_i and $m_i + 1$ in υ_k.*

$$d_i = (m_{i+1} - m_i)\%7 \tag{1}$$

For example, from incipit of the first voice of the Bach fugue BWV889 in Figure 2, the key-invariant diatonic interval sequence $dia(\upsilon_1)$ can be extracted as $[-2, +3, -6, +4, -2, +3, -2, -2, -2, +5, -2, -2, -2, +5]$.

Definition 3 (Encoding).

A function Dict(x) presents a set of rules for encoding, which maps an integer in [-6, 6] to a distinctive character. To specify in details, $Dict(-6) = A, Dict(-5) = B, Dict(-4) = C, Dict(-3) = D, Dict(-2) = E, Dict(-1) = F, Dict(0) = M, Dict(1) = G, Dict(2) = H, Dict(3) = I, Dict(4) = J, Dict(5) = K, Dict(6) = L$.

Let an encoding sequence of $dia(\upsilon_k)$ be $enc(dia(\upsilon_k))$,

$$enc(dia(\upsilon_k)) = [Dict(d) : d \in dia(\upsilon_k)] \tag{2}$$

For instance, a sequence $[-1, +2, +1, -4, +2]$ can be encoded as "$FHGCH$". The encoding sequences are used as input for the music segmentation process.

3.3 Music Segmentation

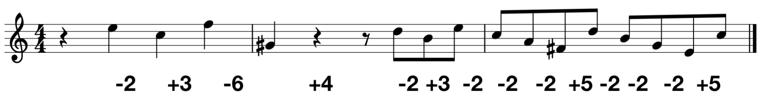

Fig. 2. Incipit of the first voice of Bach fugue BWV889

Using *swalign* [22], for each pair of voices υ_x, υ_y, a Smith-Waterman local alignment [6] between $enc(\upsilon_x)$ and $enc(\upsilon_y)$ is generated. This step detects locally-aligned segments between $enc(\upsilon_x)$ and $enc(\upsilon_y)$, and outputs an alignment of the two sequences. If the similarity score of the alignment between υ_x, υ_y is below 0.2, it should be omitted from further segmentation. This step ensures that the output of alignments between highly dissimilar voices are excluded from the motif detection process.

Figure 3 presents the alignment between the openings of voices 1 and 2 of the Bach BWV889 fugue, with locally-aligned segments identified by the Smith-Waterman algorithm boxed in red. In this representation, characters represent diatonic intervals, "-" represents a gap due to a mismatch between the two sequences, and "." represents a permitted mismatch. From the alignment of each pair of voices υ_x, υ_y, we break the sequences on each "-" character) to obtain a set of segments. We further investigate

```
Voice 1:  IAJEIEEEKEEEKEIFAM---FFFF-----HFFFFGGMFFFFGGMFFFFGGMFGG----MGGGFFFFM-
          ||||||||||||||...|   ||||     |||||||||||||||||||||||||    |  |    |
Voice 2:  IAJEIEEEKEEEKEGGMMGGGFFFFGFFFFHFFFFGGMFFFFGGMFFFFGGMFGGDEIEM--G----MF
```

Fig. 3. Alignment between the openings of voice 1 and 2 of Bach BWV889 fugue

this set of segments to take possible patterns. Furthermore, a filter is implemented to remove all segments of less than 3 elements in length, according to the definition of motif mentioned in Section 1.

Definition 4 (Pattern). *Let the filtered set of segments outputted by the alignment between υ_x, υ_x be $A(enc(\upsilon_x), enc(\upsilon_y))$. An element in $A(enc(\upsilon_x), enc(\upsilon_y))$ is a pattern.*

Definition 5 (Pattern set). *Let a score of m voices be $S = [\upsilon_1, ..., \upsilon_m]$. From the alignment between every possible pair of voices in S, we construct a set of all possible patterns*

$$P(S) = \bigcup_{(\upsilon_x, \upsilon_y) \in S \times S} A(enc(\upsilon_x), enc(\upsilon_y)) \tag{3}$$

As a valid pattern in P may appear multiple times in the course of the segmentation process, the sum of its occurrences is defined as $occ(p)$.

3.4 Intra-pattern discovery of motifs

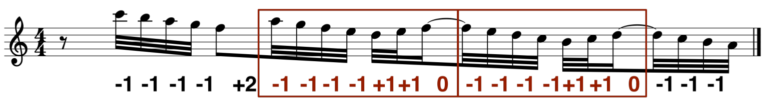

Fig. 4. Intra-pattern discovery example (from Bach BWV889 fugue)

String-based approaches are used to uncover additional motifs which are not well-captured in the segmentation process, including those which occur exclusively in one voice. For patterns of greater than 11 intervals in length in P, we identify and extract from them the longest frequent substring which occur two or more times. The choice of 11 as a length threshold is informed by previous use of a maximum pattern length of 12 notes in the literature on *n*-gram-based Music Information Retrieval [23], which is equivalent to 11 intervals. It also follows the definition of motif in this paper, favouring relatively short motivic patterns over longer patterns, which potentially correspond to musical sections or themes. The lengths of such musical structural units are not defined in absolute terms, so the length threshold of 11 elements is proposed as a working heuristic rather than a formal definition of maximum motif length. Figure 4 illustrates a case where a pattern repeatedly appears in a sequence of intervals.

The longest frequent substring extracted from a long pattern may become a substitution of the long pattern, according to rules defined as follows:

Definition 6 (Pattern substitution). *Let p be a pattern of length $|p|$, and let $sub(p)$ of length $|sub(p)|$ be the longest substring that repeated at least two times in p. Let $r_{sub(p),p}$ be the number of times $sub(p)$ appears in p without overlapping. sub(p) takes place of p in P if certain conditions are met, such as:*

$$p = \begin{cases} sub(p), & \textit{if } |sub(p)| > 3 \textit{ and } |sub(p)| * r_{sub(p),p} >= 0.6|p| \\ p, & \textit{otherwise} \end{cases} \tag{4}$$

In which, $|sub(p)| > 3$ ensures that $sub(p)$ is a non-trivial substring, following the logic discussed above in section 3.3, with the aim of removing frequent-but-insignificant short patterns. $|sub(p)| * r_{sub(p),p} >= 0.6|p|$ ensures that $sub(p)$ meets the required threshold to substitute for p.

Consider that $sub(p)$ could substitute for more than one pattern in P, let p_i be a pattern in P that is substituted by $sub(p)$, then

$$occ(sub(p)) = \sum_{p_i \in \mathcal{P}} occ(p_i) \times r_{sub(p),pi} \tag{5}$$

3.5 Inter-pattern discovery of motifs

Definition 7 (All possible pairs of long patterns). *Let (pi, pj) be a distinctive pair of patterns in P, and the set of all possible pairs of patterns in P that are longer than 11 elements be $longpairs(P)$, then*

$$longpairs(P) = \{(p_i, p_j) | p_i, p_j \in P \textit{ and } |p_i| > 11, |p_j| > 11\} \tag{6}$$

We take the longest common substrings (LCSS) of each pattern pair in $longpairs(P)$. Unique substrings discovered in this step are added as patterns for further selection in Section 3.6.2.

3.6 Ranking and selection of patterns

Graph-based pattern importance measure for ranking A weighted directed acyclic graph is used to capture the relationship between patterns in P, in order to measure the overall importance of a pattern based on its frequency of occurrence. The graph is constructed from the set of patterns P, in which each pattern in P is a node of the graph, and the directed edges represent substring relationships between patterns. The weights on the edges reflect the strength of the relationship.

Let a graph be $G = (P, E, w)$, P be a set of nodes, E be a set of directed edges, and w be the weight function. For each pair of patterns p_i and p_j, if p_i is a substring of p_j, we add a directed edge from p_i to p_j and a directed edge from p_j to p_i, weighted according to the weight function.

The weight function w is defined as follows:

If p_j is a substring of p_i, then the weight of the edge e_{ij} from p_i to p_j is 1, denoted as $\pi(e_{ij})$. The weight of the edge e_{ji} from p_j to p_i is the frequency of non-overlapping

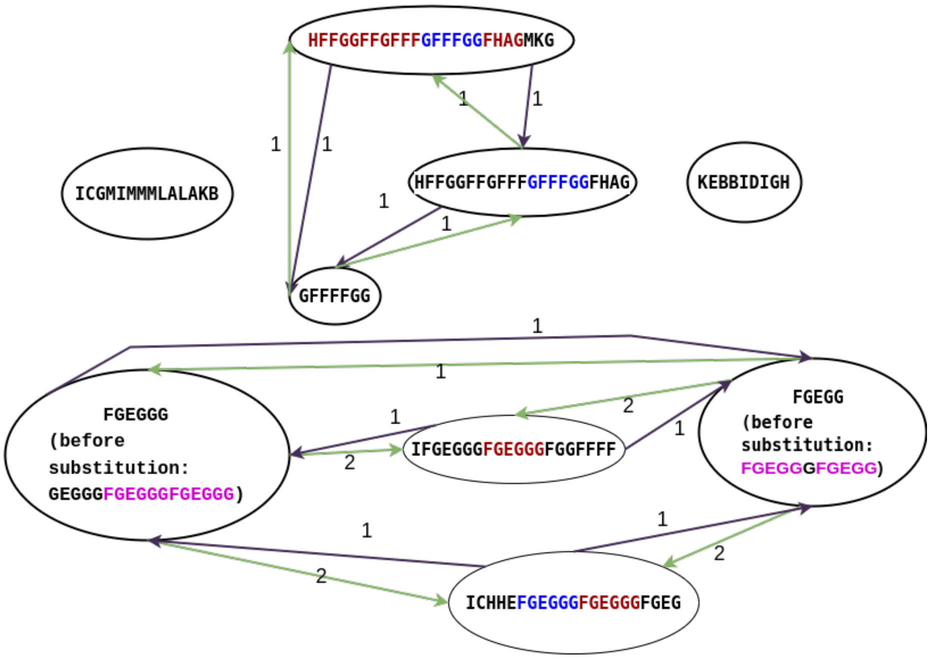

Fig. 5. A weighted directed acyclic graph of 9 patterns

occurrence of p_j in p_i, denoted as $\pi(e_{ji})$. Let $I(p_i)$ be the custom importance measure of a pattern p_i, and $[e_{i1}, ..., e_{in}]$ be the set of directed edges from p_i, then

$$I(p_i) = occ(p_i) + \sum_{k=0}^{n} \pi(e_{in}) \tag{7}$$

Figure 5 shows an example of a DAG constructed from a P of 9 patterns. Edges are added when two patterns have a substring relationship. The green edges are edges from substrings to their parent string, with a number representing their weight, while the purple edges are from parent strings to their substring. Pattern "FGEGGG" is the substring of "IFGEGGGFGEGGGFGGFFFF" which appeared 2 times, thus the edge from the former to the latter is weighted as 2. Both "FGEGGG" and "FGEGG" have the highest out-degree of 4, which indicates their importance in this set of patterns.

A set of rules for selection The patterns in P are ranked by their importance measure. The top-20 ranked patterns are selected. The set of longest common substrings generated in inter-pattern discovery of motif process are not ranked along with patterns in P, as they are extracted substrings of patterns in P. Instead, we consider those which are longer than 3 elements and repeated more than twice as valid patterns. The patterns outputted in this selection process are retained and inputted to the music structural analysis step.

3.7 Music structural analysis

In addition to fundamental attributes such as duration and pitch, we use music21 to extract a *beatStrength* value for every note in an input musical score. *beatStrength* is an encoding of the degree of rhythmic emphasis associated with each note or item in a score. It takes the form of a float value between 0 and 1. The first note of every bar can be assumed to be heavily rhythmically accented, and is assigned a value of 1 by default. *beatStrength* values are extracted for all first notes of each pattern occurrence.

If the first note of a pattern has a *beatStrength* value of 1, it indicates the pattern onset coincides with the beginning of a bar, i.e. the pattern aligns with the metric structure of the piece of music. Such patterns are retained, while patterns which begin on less rhythmically-emphasised notes are filtered out of the results.

There is one exception to this rule: As it is often the case that motifs occur at or near the beginning of a score, the above metric filtering step is not applied to patterns which occur in the opening 8 bars of a score. A threshold of 8 bars is chosen as this is the most common length for the opening period (the opening two phrases) in common western musical practice. Within this subsection of the score, patterns which begin on a less-heavily emphasised note (i.e. which are not coincident with the metric structure) are retained.

4 Results

4.1 Evaluation process

The algorithm is tested on the JKUPDD dataset [9], a set of 5 polyphonic classical scores with ground-truth annotation of repeating patterns drawn from academic sources [24–26]. This database has been previously used for testing and evaluation of other pattern detection work, notably as input data for the Music Information Retrieval Evaluation eXchange (MIREX) 2017 Discovery of Repeated Themes & Sections task [19].

Diatonic interval sequences are extracted from the labelled ground truth patterns for evaluation. The scores are manually checked to identify and annotate the exact diatonic interval occurrences of the ground truth patterns.

The pattern annotation in the JKUPDD dataset covers a wide variety of pattern types, including motifs, themes, phrases, and sections. They range in length from three elements to more than 150 elements. As our method specializes towards detection of short motif patterns, we elected to omit ground-truth annotated periods and sections from our results scoring. For the same reason, we also chose to score incomplete matches of at least 4 pattern elements as positive results in the precision and recall calculations.

According to documentation, patterns are labelled with alphabetic identifiers: “A”, “B” and so on for each score, named in order of their importance. We will make reference to this hierarchical ordering in the discussion.

4.2 Results

Precision and recall scores of the testing are presented in Table 1.

Work	**Precision** (%)	**Recall** (%)
Bach: BWV889 fugue	66.7	61.5
Beethoven: Op. 2, No. 1, Mvt. 3	100.0	45.0
Chopin: Op. 24, No. 4	50.0	50.0
Gibbons: The Silver Swan	87.5	84.6
Mozart: K282, Mvt. 2	60.0	100.0
Average	**72.8**	**68.2**

Table 1. Results: precision and recall for all JKUPDD scores

4.3 Discussion

Fig. 6. Exact match of ground truth pattern "A", occurrence 4, in Bach BWV889 fugue

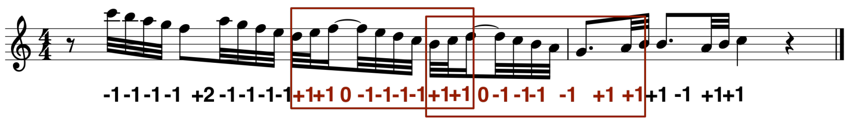

Fig. 7. Pattern "B" from Bach BWV889 fugue with two overlapping partial matches highlighted and boxed in red.

Bach: BWV889 fugue Patterns A and B are the most frequent and most significant patterns in this score. The algorithm returned A exactly. It is the opening musical motif of the entire piece and the key musical idea behind the composition. The result is illustrated in Figure 6, and detailed in the following sections. Although B matched only partially, the matching subsequence repeats twice within B. This may suggest we are capturing a core or fundamental motif within pattern B, per Figure 7.

Beethoven: Op. 2, No. 1, Mvt. 3 We fail to identify pattern A but match the opening 11-element subsequence of pattern B, which is the second-most important musical pattern in the piece per annotation.

Chopin: Op. 24, No. 4 We found a robust partial match to ground truth pattern A. The found motif occurs at the start of pattern A and repeats twice within it, in a similar manner to 7 above. Thus, the motif may be core content within pattern A, which is the most musically important/distinctive in the piece.

Gibbons: The Silver Swan Pattern A, which occurs early and repetitiously in 4 of the 5 voices, has been detected in full. Overall, our precision and recall scores are very high for this composition. It is possible that the shortness of the ground truth patterns for *The Silver Swan* play to the strengths of our tool, as it is tailored towards shorter motif pattern detection.

Mozart: K282, Mvt. 2 The sole ground truth pattern detected is a significant subsequence of pattern A. This is an incomplete but positive result, capturing the last 6 notes of this significant 10-note pattern. The detected pattern does not appear in the 'definitive' opening occurrence of pattern A, but occurs in 10 of the 11 other noted variant occurrences of pattern A in the course of the score.

Study	**Precision** (%)	**Recall** (%)
VM1 [27]	84.0	89.0
VM2 [27]	76.0	80.0
SymCHM [5]	67.9	45.4
SymCHMMerge [5]	68.0	51.0
Chen & Su [28]	50.0	69.6
Zhu & Diamond	**72.8**	**68.2**

Table 2. Average establishment precision and recall results for a selection of work evaluated on the JKUPDD database. Standard precision and recall results for our work included for comparison.

Comparison with other studies Table 2 compares our scores against establishment precision and recall values reported in other studies tested on the JKUPDD database. The establishment precision and recall defined in MIREX task guidelines [29] allows for the validity of a partial match, which is similar to our use of standard precision/recall with positive scoring of partial matches.

Although the results in Table 2 allow informal comparison of our results against similar work, it is important to note that our use of diatonic interval sequences rather than MIDI pitch sequences, our omission of sections from the ground truth, and our use of standard precision/recall all differ from the approach set out in the MIREX task documentation.

In Table 2 our approach compares favorably against all studies other than Velarde and Meridith's VM1 and VM2 studies [27]. Both VM1 and VM2 extract short pitch sequence 'segments' directly from MIDI; VM2 also filters these sequences via wavelet transform. Contiguous segments are concatenated, clustered via city block distance and ranked by the length of their occurrences in the ground truth. This building up from an initial set of short patterns contrasts against our work in which long patterns are compressed in multiple passes to produce shorter motific output patterns.

5 Contribution and future work

This paper introduces a motif extraction approach that makes novel use of local alignment for string segmentation. Patterns are detected by employing string-based methods, and a custom graph model for similarity scoring has been developed, combined with a musicologically-informed analysis and filtering step. The results presented exhibit a high degree of accuracy, broadly comparable to best-in-class pattern detection algorithms. To aid reproducibility, the source code is available on GitHub [30].

The proposed method supports related musicological tasks, such as the analysis of characteristic motifs in composition styles, or the classification of music corpora. It also has potential applications in various domains in MIR including music generation.

In the future, we plan to improve the algorithm via encoding more musicological knowledge. We also intend to apply the algorithm to inter-opus pattern detection in a corpus of monophonic Irish traditional folk tunes on *The Session* [31], which will help gain greater insight into the role of motifs in defining *tune families* [32] within the corpus.

References

1. J. D. White, *The Analysis of Music*. Prentice-Hall.
2. D. Conklin and M. Bergeron, "Feature set patterns in music," *Computer Music Journal*, vol. 32, no. 1, pp. 60–70.
3. B. Janssen, W. de Haas, A. Volk, and P. Van Kranenburg, "Finding repeated patterns in music: State of knowledge, challenges, perspectives," in *Sound, Music, and Motion*. Springer International Publishing, vol. 8905, pp. 277–297.
4. T. Collins, A. Arzt, S. Flossmann, and G. Widmer, "SIARCT-CFP: Improving precision and the discovery of inexact musical patterns in point-set representations," in *International Society for Music Information Retrieval Conference*.
5. M. Pesek, A. Leonardis, and M. Marolt, "SymCHM—an unsupervised approach for pattern discovery in symbolic music with a compositional hierarchical model," *Applied Sciences*, vol. 7.
6. T. F. Smith and M. S. Waterman, "Identification of common molecular subsequences." *Journal of molecular biology*, vol. 147 1.
7. P. van Kranenberg, "A computational approach to content-based retrieval of folk song melodies," 2010.
8. R. Hillewaere, B. Manderick, and D. Conklin, "Alignment methods for folk tune classification," in *Data Analysis, Machine Learning and Knowledge Discovery*. Springer International Publishing, pp. 369–377.
9. T. Collins, "The johannes kepler university patterns development database." [Online]. Available: https://www.music-ir.org/mirex/wiki/2017:Discovery_of_Repeated_Themes_\%26_Sections
10. M. Giraud, R. Groult, and F. Levé, "Subject and counter-subject detection for analysis of the well-tempered clavier fugues," in *From Sounds to Music and Emotions*, M. Aramaki, M. Barthet, R. Kronland-Martinet, and S. Ystad, Eds. Springer Berlin Heidelberg.
11. A. Volk and P. Van Kranenburg, "Melodic similarity among folk songs: An annotation study on similarity-based categorization in music," *Musicae Scientiae*, vol. 16.
12. J. de Berardinis, M. Vamvakaris, A. Cangelosi, and E. Coutinho, "Unveiling the hierarchical structure of music by multi-resolution community detection," *Transactions of the ISMIR*, vol. 3, no. 1, pp. 82–97.

13. W. Chai, “Semantic segmentation and summarization of music: methods based on tonality and recurrent structure,” *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 124–132.
14. T. Nuttall, M. G. Casado, A. Ferraro, D. Conklin, and R. C. Repetto, “A computational exploration of melodic patterns in arab-andalusian music,” *Journal of Mathematics and Music*, vol. 15, no. 2.
15. D. Conklin and C. Anagnostopoulou, “Comparative pattern analysis of cretan folk songs,” *Journal of New Music Research*, vol. 40, no. 2, pp. 119–125.
16. Jia-Lien Hsu, Chih-Chin Liu, and A. Chen, “Discovering nontrivial repeating patterns in music data,” *IEEE Trans. Multimedia*, vol. 3, no. 3.
17. D. Meredith, “COSIATEC and SIATECCompress: Pattern discovery by geometric compression,” in *Music Information Retrieval Evaluation eXchange (MIREX 2013)*. International Society for Music Information Retrieval.
18. ——, “Compression-based geometric pattern discovery in music,” in *2014 4th International Workshop on Cognitive Information Processing (CIP)*, pp. 1–6.
19. T. Collins, B. Janssen, and Y. Hao. Mirex 2017: Discovery of repeated themes & sections results. [Online]. Available: https://www.music-ir.org/mirex/wiki/2017:MIREX2017_Results
20. I. Y. Ren, H. V. Koops, A. Volk, and W. Swierstra, “In search of the consensus among musical pattern discovery algorithms,” in *International Society for Music Information Retrieval Conference*, 2017.
21. M. S. Cuthbert and C. Ariza, “music21: A toolkit for computer-aided musicology and symbolic music data,” in *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, pp. 637–642.
22. M. Breese, “swalign.” [Online]. Available: https://github.com/mbreese/swalign
23. T. Zhu, R. Fournier-S’niehotta, P. Rigaux, and N. Travers, “A framework for content-based search in large music collections,” *Big Data and Cognitive Computing*, vol. 6, no. 1, p. 23.
24. S. Bruhn, *J.S. Bach’s Well-Tempered Clavier: in-depth analysis and interpretation*. Mainer International.
25. H. Barlow and S. Morgenstern, *A dictionary of musical themes*. Crown Publishers.
26. A. Schoenberg, *Fundamentals of Musical Composition.* Faber and Faber.
27. G. Velarde and D. Meredith, “A wavelet-based approach to the discovery of themes and sections in monophonic melodies.”
28. T.-P. Chen and L. Su, “Discovery of repeated themes and sections with pattern clustering.” [Online]. Available: https://www.music-ir.org/mirex/abstracts/2017/CS3.pdf
29. T. Collins. Mirex 2017: Discovery of repeated themes & sections results. [Online]. Available: https://www.music-ir.org/mirex/wiki/2017:Discovery_of_Repeated_Themes_\%26_Sections
30. T. Zhu and D. Diamond, “motif_extraction: A novel local-alignment-based approach to motif extraction in polyphonic music.” [Online]. Available: https://github.com/TiangeZhu/motif_extraction
31. J. Keith. The session. [Online]. Available: https://thesession.org
32. S. P. Bayard, “Prolegomena to a study of the principal melodic families of british-american folk song,” *The Journal of American Folklore*, vol. 63, no. 247, publisher: University of Illinois Press.