

DETECTING EPISODES WITH HARMONIC SEQUENCES FOR FUGUE ANALYSIS

Mathieu Giraud

LIFL, CNRS, Université Lille 1
INRIA Lille, France

Richard Groult

MIS, Université Picardie Jules Verne
Amiens, France

Florence Levé

MIS, Université Picardie Jules Verne
Amiens, France

ABSTRACT

Fugues alternate between instances of the *subject* and of other patterns, such as the *counter-subject*, and modulatory sections called *episodes*. The episodes play an important role in the overall design of a fugue: detecting them may help the analysis of the fugue, in complement to a subject and a counter-subject detection. We propose an algorithm to retrieve episodes in the fugues of the first book of Bach's *Well-Tempered Clavier*, starting from a symbolic score which is already track-separated. The algorithm does not use any information on subject or counter-subject occurrences, but tries to detect *partial harmonic sequences*, that is similar pitch contour in *at least two voices*. For this, it uses a substitution function considering "quantized partially overlapping intervals" [14] and a strict length matching for all notes, except for the first and the last one. On half of the tested fugues, the algorithm has correct or good results, enabling to sketch the design of the fugue.

1. INTRODUCTION

A fugue is a polyphonic piece built in imitation, where all voices appear successively sharing the same initial melodic material: a subject and, in most cases, a counter-subject. These patterns are repeated throughout the piece, either in their initial form or more often altered or transposed, building a complex harmonic texture. Many composers wrote fugues, or included fugal parts in larger pieces. The two books of Bach's *Well-Tempered Clavier* are a particularly consistent corpus, exploring the 24 major and minor tonalities in 48 preludes and fugues.

Fugues are often viewed as one of the pinnacle forms of Western music, and they are also used for pedagogical purposes, in music analysis as in composition. Their structure may look very formal, but still enable high levels of creativity. There are many treatises on fugues, or, more generally, on counterpoint, as for example [13] or [18]. Some of them include a complete musicological analysis of Bach's *Well-Tempered Clavier*, as the books of S. Bruhn [3, 4]. The fugues are thus perfect candidates for Music Informa-

tion Retrieval (MIR) research, stimulating the development of algorithms on symbolic scores.

A first way to analyze fugues can be to use generic tools detecting repeating patterns or themes, possibly with approximate occurrences. Similarity between parts of a piece may be computed by the Mongeau-Sankoff algorithm [17] and its extensions, or by other methods for approximate string matching [6, 7, 19], allowing a given number of restricted mismatches. Several studies focus on finding *maximal repeating patterns*, limiting the search to *non-trivial* repeating patterns, that is discarding patterns that are a sub-pattern of a larger one with the same frequency [10, 12, 15]. Other studies try to find musically significant *themes*, with algorithms considering the number of occurrences [20], but also the melodic contour or other features [16].

More specifically, some MIR studies already focused on fugues. The study [21] builds a tool to decide if a piece is a fugue or not, with a method to find occurrences of thematic materials. The bachelor thesis [2] contains methods to analyze fugues, including voice separation. It proposes several heuristics to help the selection of repeating patterns inside the algorithms of [10] which maximizes the number of occurrences. The web site [9] also produces an analysis of fugues, extracting sequences of some repeating patterns, but without precise formal analysis. Finally, we proposed in [8] a method to detect subjects and counter-subjects, based on an analysis of repeating patterns with a diatonic substitution function and a specific length matching. This method finds the precise ends of these patterns in the majority of the fugues of the first book of Bach's *Well-Tempered Clavier*.

The subject and the counter-subject are focus of musical cognition, and will often be what is remembered from a fugue. However, the link between the different expositions of these patterns occurs in transitional sections called *episodes* that modulate from one tonality to another [13, 18]. The episodes have a part in the development of tension during the fugue.

To our knowledge, no previous MIR study was devoted to analysis of episodes. Episodes can be detected by the absence of subjects and counter-subjects: A perfect detection of subjects and counter-subjects should yield a perfect episode detection. In this paper, we try to retrieve episodes without using any information on subject or counter-subject occurrences. We thus look for a *positive identification of*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2012 International Society for Music Information Retrieval.

episodes. Starting from a symbolic score which is already track-separated, we propose an algorithm to retrieve episodes containing *partial harmonic sequences*, that is similar pitch contour in *at least two voices*. Harmonic sequences are commonly used to modulate, and are thus an essential feature of many episodes.

As in [8], the algorithm uses a strict length matching for all notes, except for the first and the last one. We tested several substitution functions to have a sensible and specific approximate matching. Our best results use the “quantized partially overlapping intervals” (QPI), introduced by Lemström and Laine in [14], that can be also seen as one case of the “General Pitch Interval Representation” defined by Cambouropoulos in [5].

The paper is organized as follows. Section 2 gives definitions and some background on fugues, Section 3 details our algorithm for episode detection through partial harmonic sequences, and Section 4 details the results on 21 fugues of the first book of Bach’s *Well-Tempered Clavier*. These results were evaluated against a reference musicological book [4]. On half of the tested fugues, the algorithm has correct or good results, enabling to sketch the design of the fugue. The other cases are fugues where the episodes do not show enough harmonic sequences, or where the sequences are too short or too much altered.

2. PRELIMINARIES

A *note* x is described by a triplet (p, o, ℓ) , where p is the pitch, o the onset, and ℓ the length. The pitches can describe diatonic (based on note names) or semitone information. We consider ordered *series of notes* $x_1 \dots x_m$, that is $x_1 = (p_1, o_1, \ell_1), \dots, x_m = (p_m, o_m, \ell_m)$, where $1 \leq o_1 \leq o_2 \leq \dots \leq o_m$ (see Figure 1). The series is *monophonic* if there are never two notes sounding at the same onset, that is, for every i with $1 \leq i < m$, $o_i + \ell_i \leq o_{i+1}$. To be able to match transposed patterns, we consider relative pitches, also called *intervals*: the interval series is defined as $\Delta x_2 \dots \Delta x_m$, where $\Delta x_i = (\Delta p_i, o_i, \ell_i)$ and $\Delta p_i = p_i - p_{i-1}$.



Figure 1. A monophonic series of notes (start of Fugue #2, see Figure 4), represented by (p, o, ℓ) or $(\Delta p, o, \ell)$ triplets. In this example, onsets and lengths are counted in sixteenthths, and pitches and intervals are counted in semitones through the MIDI standard.

Fugue. We now introduce some notions about fugue analysis. These concepts are illustrated by Fugue #2 of the first book of Bach’s *Well-Tempered Clavier*. This fugue has a very regular construction.

A *fugue* is given by a set of *voices*, where each voice is a monophonic series of notes. In Bach’s *Well-Tempered Clavier*, the fugues have between 2 and 5 voices, and Fugue #2 is made of 3 voices.

The fugue is built on a theme called *subject*. The first three *occurrences* of the subject in Fugue #2 are detailed in Figure 4: the subject is *exposed* at one voice (the alto), beginning on a C, until the second voice enters (the soprano, measure 3). The subject is then exposed at the second voice, but is now transposed to G. Meanwhile, the first voice continues with the first *counter-subject* that combines with the subject. Figure 3 shows a sketch of the entire fugue. The fugue alternates between other instances of the subject together with counter-subjects and development and modulatory sections called *episodes*.

Episodes and sequences. The episodes “effect a smooth transition from one key to the next [and] provide variety, as well as relief from a constant emphasis on the subject as a whole” [13]. They are often built on portions of material from the subjects of counter-subjects. S. Bruhn lists three roles for an episode in the design of the fugue: “It can link two subject statements by leading from one towards the next; it can be conclusive by resolving tension that was built up by the preceding subject statement; it can represent a different register, appearing basically independent of its surroundings and serving as a color contrast.” [4].

The Figure 3 shows the two first episodes of Fugue #2. Note that the term “episode” can also be restrained to the ones after the exposition of all voices, the first episode being called *codetta* [18].

The episodes can include *cadential passages* for the release of tension. However, they are often composed with *harmonic sequences*, which are passages where a pattern is consecutively repeated starting on a different pitch. Figure 2 shows a simple harmonic sequence, outside of a fugue. Sequences can be diatonic (keeping the same key signature, possibly modulating to a neighbor tonality) or real (possibly gaining or losing some sharps or flats, often modulating to some other tonality).

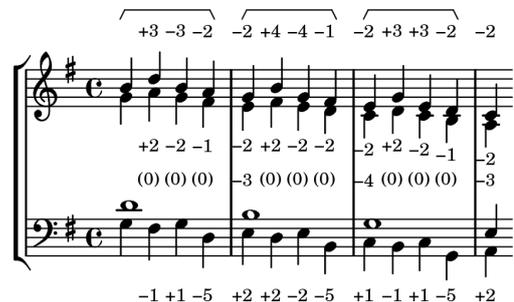


Figure 2. A simple diatonic sequence [1]. The values indicate the intervals from the preceding note of the same voice, in number of semitones. The occurrences #1 and #3 have exactly the same semitone intervals. The occurrence #2 is identical to these occurrences when one considers only diatonic information.

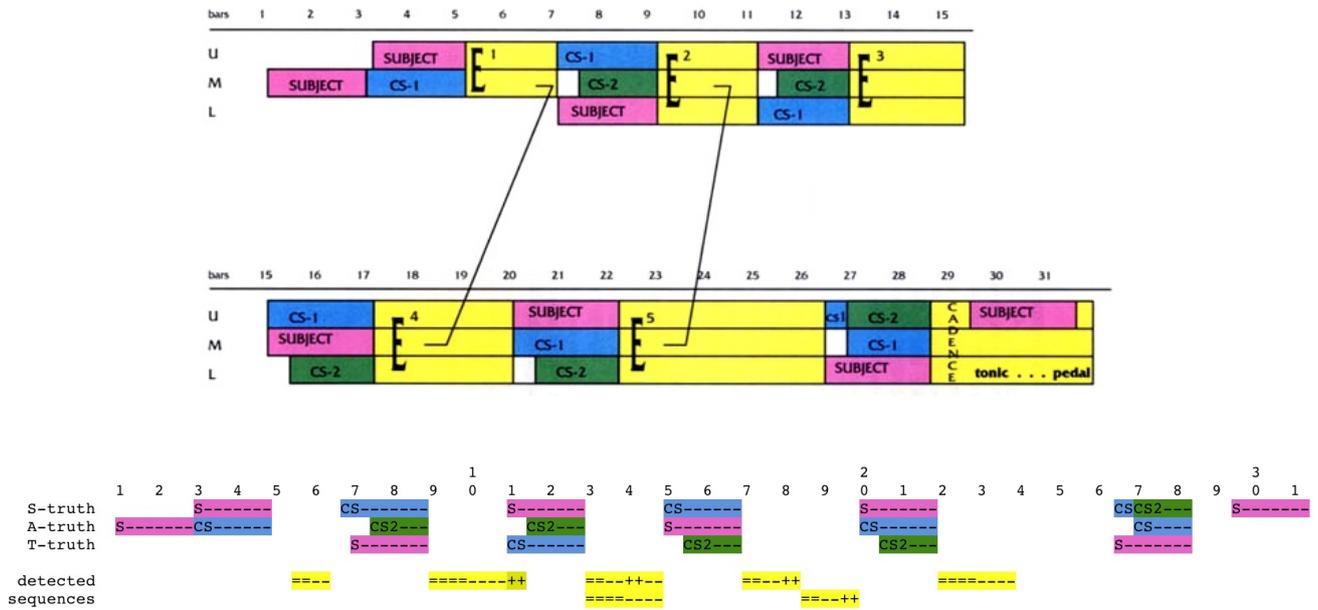


Figure 3. Analysis of Fugue #2 in C minor in the first book of Bach’s *Well-Tempered Clavier* (BWV 847). Top: ground truth (analysis by S. Bruhn, used with permission [4], [4, p. 80]). Bottom: the two lines named “detected sequences” show the output of the proposed algorithm, detecting partial harmonic sequences in 5 out of the 6 episodes and 68% of the concerned measures. The only false positive is the end of the second episode: at measure 11, it overlaps with the next occurrence of the subject (S) and counter-subject (CS).



Figure 4. Start of Fugue #2 in C minor (BWV 847), showing the ground truth for the first two episodes. Non-episodic parts are grayed. The notes starting the initial patterns and the occurrences of the sequences are circled.

3. EPISODE DETECTION

We propose here to detect episodes containing partial harmonic matches in at least two voices. For this, we consider consecutively repeating patterns under a substitution function using a relaxed similarity for pitch intervals, and enforcing length equalities of all notes but the first one and the last one. These are very conservative settings, to have as few false positives as possible.

Consecutively repeating patterns. Formally, in a given voice x , we look for consecutively repeating patterns of p notes, starting at note x_e . The pattern $x_e \dots x_{e+p-1}$ has a candidate second occurrence $x_{e+p} \dots x_{e+2p-1}$, and, for larger episodes, we also check for a third ($x_{e+2p} \dots x_{e+3p-1}$) and fourth ($x_{e+3p} \dots x_{e+4p-1}$) occurrences.

The score $I(x, e, p, r)$ between the pattern $x_e \dots x_{e+p-1}$ and its candidate occurrence number r ($r = 2, 3$ or 4) is defined by the number of intervals matched between the pattern and its candidate occurrence:

$$\begin{aligned} I(x, e, p, r) &= \delta(\Delta x_{e+1}, \Delta x_{e+(r-1)p+1}) \\ &+ \delta(\Delta x_{e+2}, \Delta x_{e+(r-1)p+2}) \\ &+ \dots \\ &+ \delta_f(\Delta x_{e+p}, \Delta x_{e+rp}) \end{aligned}$$

As in [8], we propose to use a strict length matching for all notes, except for the first and the last one – the length of these notes, at the extremities of the pattern, being more frequently altered. The substitution function δ checks thus pitch intervals and lengths, whereas the substitution function, δ_f , for the last note, only considers pitch intervals:

$$\begin{aligned} \delta((\Delta p, o, \ell), (\Delta p', o', \ell')) &= \\ \begin{cases} +1 & \text{if } \Delta p \approx \Delta p' \text{ and } \ell = \ell' \\ 0 & \text{if } \Delta p \not\approx \Delta p' \text{ and } \ell = \ell' \\ -\infty & \text{otherwise } (\ell \neq \ell') \end{cases} \\ \delta_f((\Delta p, o, \ell), (\Delta p', o', \ell')) &= \\ \begin{cases} +1 & \text{if } \Delta p \approx \Delta p' \\ 0 & \text{otherwise } (\Delta p \not\approx \Delta p') \end{cases} \end{aligned}$$

The actual comparison of lengths ($\ell = \ell'$) also checks the equality of the rests that may be immediately before the compared notes. The length of the first note of the pattern (x_e against $x_{e+(r-1)p}$) is never checked, as the score actually compares the series of intervals $\Delta x_{e+1} \dots \Delta x_{e+p}$ against $\Delta x_{e+(r-1)p+1} \dots \Delta x_{e+rp}$.

The relation \approx is a similarity relation on pitch intervals. We use here the ‘‘quantized partially overlapping intervals’’ (QPI) [14], that defines *short* intervals (from one to three semitones), *medium* intervals (from three to seven semitones), and *large* intervals (starting from six semitones).

These classes can be considered for upward or downwards intervals, giving, with the unison intervals, a total of 7 different interval classes. Two pitch intervals Δp and $\Delta p'$ will be considered as similar if there exists one class containing both of them.

There is an *exact occurrence* of the consecutively repeating pattern if $I(x, e, p, r) = p - 1$. For example, on the sequence depicted on Figure 2, for any of the four voices x and for $r \in [2, 3]$, we have $I(x, 1, 4, r) = 3$, since intervals are perfectly similar under the QPI similarity relation. An *approximate occurrence* can be detected if $I(x, e, p, r)$ is at least equal to a given threshold $\tau(p)$.

Here the score $I(x, e, p, r)$ only considers substitution operations, and can be computed in time $O(p)$. The score can be extended to consider other edit operations, with computation through dynamic programming.

Episode detection through partial sequences. On the beginning of the Fugue #2, the consecutively repeating patterns are as follows:

- the second episode fits perfectly into a sequence: $I(\text{soprano}, 58, 7, 2) = 6$, $I(\text{alto}, 76, 7, 2) = 6$ and $I(\text{tenor}, 21, 16, 2) = 15$.
- the first episode has two complete occurrences, as $I(\text{soprano}, 21, 5, 2) = 4$ and $I(\text{alto}, 41, 6, 2) = 5$. There is no complete third occurrence, as the lengths do not match for one voice: $I(\text{alto}, 41, 6, 3) = -\infty$.

The complete algorithm computes $I(x, e, p, r)$ for every voice x , every note x_e starting right after a quarter beat, several periods (1 quarter, and 1/2, 1 and 2 measures) and for $r \in \{2, 3, 4\}$ occurrences. The algorithm reports an episode every time that at least two different voices contain a consecutively repeating pattern after the same onset (with $\tau(p) = 0.5 \times p$). Overlapping episodes with the same period are merged into an unique episode. The result on the Fugue #2 is depicted at the bottom of Figure 3.

For testing purposes, we used a naive implementation running in $O(n^2)$ worst time, where n is the total number of notes in the fugue. Even if similarities between occurrences in a sequence could be computed with tools in existing frameworks (such as the `simil` tool in the Humdrum toolkit [11, 19]), we coded our own implementation to be able to handle some specificities (specific length matching, partial detection in two voices).

4. RESULTS AND DISCUSSION

Results can be asserted in two different ways:

- We can count the sequences that are located completely or partially in episodes of the ground truth;
- More precisely, we can look at the total length of detected sequences, and compare it to the total length of all ground truth episodes, computing a ratio called *length sensibility*. This sensibility can be seen as a *coverage of episodes by harmonic sequences*: it will not reach 100%, as some episodes do not have sequences, and as the sequences may not be spanning all the episodes. We also compute a *length specificity* as the ratio between the lengths of true positive measures and of detected measures.

The result on Fugue #2 is shown at the bottom of Figure 3. Here 5 episodes out of 6 are detected with partial harmonic sequences. This is musically relevant, since the last episode (measures 29-31) is a cadential end, with a last exposition on the subject on a bass pedal. The ground truth has 14 1/2 measures of episodes. The algorithm outputs 10 measures (length sensibility of 68%) and falsely marks one half measure (2 quarters) as an episode (length specificity of 96%).

The false negatives are: 1 measure at the codetta (due to the shift between the two voices, only 2 occurrences are detected), 2 measures and a half at measure 24 (including a change of voices, see below), and all the 3 measures of the last episode (discussed above). The only false positive is the end of the second episode, which is extended 2 quarters below the next subject occurrence at measure 11, the soprano and the bass voices continuing the sequence (see Figure 4, last measure).

The complete test contains 21 fugues of the first book of Bach's *Well-Tempered Clavier* (fugues #1, #4 and #9 not showing significant episodic material). We started from .krn Humdrum files [11], available for academic purposes at <http://kern.humdrum.org/>. The output of the algorithm on all these 21 fugues is available at <http://www.lifl.fr/~giraud/fugues>. We checked all detected episodes, and Table 1 summarizes the results. On the 1098 measures of this test set, the algorithm labels about 20% of the measures as episodes, and finally identifies 43% of all episodes. A subjective quality assessment on the predictions, looking on the detailed output of each run, gives a *good* mark on 6 fugues, and a *correct* mark for 5 out of the 21 fugues.

False positives. There are very few false positives: less than 5% of the partial harmonic sequences overlap with subject and counter-subject occurrences. As for the measure 11 in Fugue #2, this is often because the texture of the episode fades into the next section.

False negatives. The length sensibility, that is the coverage of the episodes (in the ground truth) by the prediction of harmonic sequences is, in average, only 36%. These false negatives can be explained by several facts:

- As mentioned above, the sequences often not cover all the episodes. Moreover, there are some episodes with no harmonic sequence: It is often the case for the last episode, which is thus almost always missed by the proposed method;
- There are some episodes with *changes of voices* (Figure 5), in which the consecutive occurrences of a pattern are not in a same voice;
- Finally, the algorithm fails to detect some partial harmonic sequences that are too much altered to be recognized with the current threshold, or too short to be discovered.



Figure 5. Partial sequence with a change of voices: the pattern is heard at the soprano, then, transposed, at the alto (measure 24 of Fugue #2).

5. CONCLUSIONS

We proposed an algorithm retrieving some episodes in the first book of Bach's *Well-Tempered Clavier*, starting from a symbolic score which is already track-separated. To our knowledge, this is the first MIR study on episodes in fugues. The algorithm, relying only on partial harmonic sequences detection, gives very few false positives, and already gives good results on some fugues. Enabling voice changes inside harmonic sequences should naturally detect more episodes, but may produce more false positives.

Many improvements are possible to have a better analysis of episodes. Detection of other positive features of the episodes (such as cadential passages) or, most of all, of some negative features (subject and counter-subject occurrences) could probably lead to a complete fugue analysis pipeline with better results.

The algorithm could also be tested on other corpus of fugues. As an example, the web page <http://www.lifl.fr/~giraud/fugues> shows the output of the proposed algorithm in the fugue of Mozart's *Adagio and Fugue in C minor*, K 546. Finally, partial or full harmonic sequence detection could be used to help the analysis of other genres.

Acknowledgements. The authors would like to thank the anonymous reviewers for their valuable comments.

6. REFERENCES

- [1] Marcel Bitsch. *Précis d'harmonie tonale*. Alphonse Leduc, 1957.
- [2] Lisa Browles. Creating a tool to analyse contrapuntal music. Bachelor Dissertation, Univ. of Bristol, 2005.
- [3] Siglind Bruhn. *J. S. Bach's Well-Tempered Clavier. In-depth Analysis and Interpretation*. 1993. ISBN 962-580-017-4, 962-580-018-2, 962-580-019-0, 962-580-020-4. Available online at <http://www-personal.umich.edu/~siglind/text.htm>.
- [4] Siglind Bruhn. *J. S. Bachs Wohltemperiertes Klavier, Analyse und Gestaltung*. Edition Gorz, 2006. ISBN 3-938095-05-9.
- [5] Emiliios Cambouropoulos. A general pitch interval representation: Theory and applications. *Journal of New Music Research*, 25(3):231–251, 1996.

#	BWV	tonality	voices	ℓ	ground truth		found by the proposed algorithm						
					nb	ℓ	nb	ℓ (TP)	ℓ (FP)	sens	spec	quality	
2	847	C minor	3	31m	6	14m2q	5	10m	2q	68%	96%	+	
3	848	C# major	3	55m	7	31m2q	4	17m	1m	54%	94%	+	
5	850	D major	4	27m	5	16m+1q	4	7m	–	41%	100%	=	
6	851	D minor	3	44m	8	12m	3	7m	1m+1q	57%	84%	=	
7	852	Eb major	3	37m	9	22m+2q	7	16m+1q	–	72%	100%	+	
8	853	D# minor	3	87m	10	22m+2q	1	3m+1q	2q	15%	100%	–	
10	855	E minor	2	42m	4	25m	4	18m+2q	1m	75%	95%	+	
11	856	F major	3	72m	6	31m+1q	4	10m	1m	31%	87%	=	
12	857	F minor	4	58m	8	26m	5	12m+3q	7m	48%	65%	–	
13	858	F# major	3	35m	6	17m+2q	4	7m+1q	2m+2q	42%	75%	=	
14	859	F# minor	4	40m	5	10m	4	5m+3q	3m	56%	66%	=	
15	860	G major	3	86m	9	36m	6	11m	15m+1q	30%	42%	–	
16	861	G minor	4	34m	6	12m+1q	2	3m+1q	1m+2q	27%	71%	–	
17	862	Ab major	4	35m	7	24m+2q	4	5m+1q	1q	21%	100%	–	
18	863	G# minor	4	41m	6	17m	4	9m+3q	1m	57%	92%	+	
19	864	A major	3	54m	7	27m	1	2m+3q	–	10%	100%	–	
20	865	A minor	4	87m	16	19m	3	1m+2q	10m+2q	9%	14%	–	
21	866	Bb major	3	48m	3	14m+1q	3	12m	1q	83%	96%	+	
22	867	Bb minor	5	75m	6	34m+2q	1	2m+1q	2m	6%	60%	–	
23	868	B major	4	34m	7	10m+1q	0		–	0%		–	
24	869	B minor	4	76m	12	38m+1q	3	7m+2q	3m	20%	76%	–	
					1098m	153	457m+20q	72	164m+24q	48m+14q			+ : 6
												= : 5	
												– : 10	

Table 1. Detection of episodes in 21 fugues of the first book of Bach’s *Well-Tempered Clavier*. The ground truth is the analysis of [4]. All lengths (ℓ) are given in number of measures (m) and quarters (q). The sensibility (sens.) and specificity (spec.) are computed on the lengths (see Section 4). The last column, “quality”, is a subjective assessment on the output of the algorithm: find almost all episodes with $\geq 50\%$ length sensibility, almost no false positives with $\geq 90\%$ length specificity (*good*, +), find most of the episodes, few false positives (*correct*, =), miss many episodes or produces many false positives (*bad*, –).

- [6] Raphaël Clifford and Costas S. Iliopoulos. Approximate string matching for music analysis. *Soft. Comput.*, 8(9):597–603, 2004.
- [7] T. Crawford, C. Iliopoulos, and R. Raman. String matching techniques for musical similarity and melodic recognition. *Computing in Musicology*, 11:71–100, 1998.
- [8] Mathieu Giraud, Richard Groult, and Florence Levé. Subject and counter-subject detection for fugue analysis. In *Computer Music Modeling and Retrieval (CMMR 2012)*, 2012.
- [9] J. Hakenberg. The Pirate Fugues. <http://www.hakenberg.de/music/music.htm>.
- [10] J. L. Hsu, C. C. Liu, and A. Chen. Efficient repeating pattern finding in music databases. In *Int. Conference on Information and Knowledge Management (CIKM 1998)*, 1998.
- [11] David Huron. Music information processing using the humdrum toolkit: Concepts, examples, and lessons. *Computer Music Journal*, 26(2):11–26, 2002.
- [12] Ioannis Karydis, Alexandros Nanopoulos, and Yannis Manolopoulos. Finding maximum-length repeating patterns in music databases. *Multimedia Tools Appl.*, 32:49–71, 2007.
- [13] Kent Kennan. *Counterpoint*. Prentice Hall, 4th ed., 1999.
- [14] Kjell Lemström and Pauli Laine. Musical information retrieval using musical parameters. In *Int. Computer Music Conference (ICMC '98)*, pages 341–348, 1998.
- [15] Chih-Chin Liu, Jia-Lien Hsu, and Arbee L.P. Chen. Efficient theme and non-trivial repeating pattern discovering in music databases. In *Int. Conf. on Data Engineering (ICDE 99)*, pages 14–21, 1999.
- [16] Colin Meek and William P Birmingham. Automatic thematic extractor. *Journal of Intelligent Information Systems*, 21(1):9–33, 2003.
- [17] Marcel Mongeau and David Sankoff. Comparaison of musical sequences. *Computer and the Humanities*, 24:161–175, 1990.
- [18] Hugo Norden. *Foundation Studies in Fugue*. Crescendo Publishing, 1977.
- [19] Keith S. Orpen and David Huron. Measurement of similarity in music: A quantitative approach for non-parametric representations. *Computers in Music Research*, 4:1–44, 1992.
- [20] Lloyd Smith and Richard Medina. Discovering themes by exact pattern matching. In *Int. Symp. for Music Information Retrieval (ISMIR 2001)*, pages 31–32, 2001.
- [21] Pei-Hsuan Weng and Arbee L. P. Chen. Automatic musical form analysis. In *Int. Conference on Digital Archive Technologies (ICDAT 2005)*, 2005.