Time-Span Tree Analyzer for Polyphonic Music

Masatoshi Hamanaka[1], Keiji Hirata[2] and Satoshi Tojo[3],

[1] University of Tsukuba
[2] Future University Hakodata
[3] Japan Advanced Institute of Science and Technology
hamanaka@iit.tsukuba.ac.jp

Abstract. We have been developing a music analysis system called a polyphonic music time-span tree analyzer (PTTA). A time-span tree assigns a hierarchy of 'structural importance' to the notes of a piece of music on the basis of the Generative Theory of Tonal Music (GTTM). However, the theory only accepts homophonic music. To solve this problem, we first record the composers' processes for arranging from polyphony to homophony because the processes show how a musician reduces ornamental notes. Using the recording of the arrangement process with the time-span tree of the homophony, we manually acquire a time-span tree of polyphony. Then we attempt to develop a PTTA that semi-automatically acquires a time-span tree of polyphony by implementing an additional novel rule for time-span analysis. Experimental results show that the PTTA using our proposed rules outperforms the baseline.

1 Introduction

Our goal is to create a system that will enable a musical novice to manipulate a piece of music, which is an ambiguous and subjective media. For example, it is difficult for musical novices to manipulate music with commercial music sequencers that only operate on the surface structure of music, that is, the pitch and on-timing of each note. On the other hand, Garageband [1] can create a piece of music though simple manipulations, i.e., by just concatenating pre-stored phrases. However, when we want to arrange a portion of a melody in a phrase, we have to manipulate the surface structure of the music, and a musical novice will have difficulty finding the software to mirror his or her intentions in such a case.

Previous music systems [2, 3] have their own music analysis methods, from which deeper musical structures are difficult to acquire, and thus these systems are difficult for the users to manipulate at their will. However, a representation method and primitive operation for time-span trees in the Generative Theory of Tonal Music (GTTM) [4] have been proposed, and these developments indicate the potential for constructing a melody-arranging algorithm [5]. An example of arranging algorithms that use a time-span tree is the melody morphing method [6], which generates an intermediate melody between one melody and another with a systematic order in accordance with a certain numerical measure.

However, GTTM is limited to homophonic music, which consists of a single note sequence with chords. To overcome this limitation, we extend GTTM by proposing and implementing an additional novel rule for time-span analysis and attempt to develop a polyphonic music time-span tree analyzer (PTTA) that enables a time-span tree to be acquired semi-automatically from polyphony. This extension enables the time-span tree to represent the relationships between parts. For example, there are completely independent parts that form independent trees. In contrast, there are parts that progress in the same rhythm to form an overlapped tree.

This paper is organized as follows. We discuss the problem of how to construct a system that can acquire a time-span tree from polyphony in Section 2 and whether a time-span tree can actually be acquired from polyphony or not in Section 3. We describe the polyphonic music time-span tree analyzer in Section 4 and present experimental results and a conclusion in Sections 5 and 6.

2 GTTM and its Analyzer

GTTM is composed of four modules, each of which assigns a separate structural description to a listener's understanding of a piece of music. These four modules output a grouping structure, a metrical structure, a time-span tree, and a prolongational tree, respectively (Fig. 1).

In our previous work [7], we extended original theory by full externalization and parameterization and proposed a machine-executable extension of GTTM, exGTTM. The externalization includes introducing an algorithm to generate a hierarchical structure of the time-span tree in a mixed top-down and bottom-up manner. The parameterization includes introducing a parameter for controlling the priorities of rules to avoid conflicts among the rules, as well as parameters for controlling the shape of the hierarchical time-span tree.

We implemented the exGTTM on a computer called ATTA (Fig. 2). The ATTA only treats monophony because several rules in the theory only allow monophony. This limitation is too narrow because users may want to manipulate polyphonic or homophonic music. Therefore, we have been constructing a music analysis system that enables a time-span tree to be acquired from polyphony and homophony. Here, we discuss the problems when extending the system that enables polyphony and homophony to be treated.

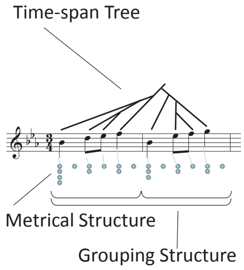

Fig. 1. Structures of GTTM.

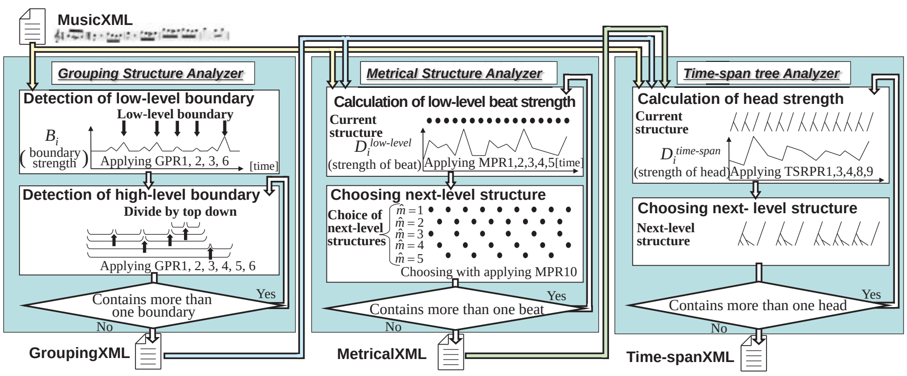

Fig. 2. Automatic time-span tree analyzer (ATTA).

No ground truth exists. If we analyze polyphony using GTTM, there are no ground truth data because the theory is limited to treating homophony and the Lerdahl and Jackendoff's book contains no example that uses polyphony.

The time-span reduction represents the intuitive idea: if we remove ornamental notes from a long melody, we obtain a simple melody that sounds similar. An entire piece of Western tonal music can eventually be reduced to an important note. For example, the left-hand side of Fig. 3(a) depicts a simple monophony and its tree. The time-span (designated as <--->) is represented by a single note, called a head, which is designated here as "C4".

We believe this intuitive idea of GTTM is applicable to polyphonic and homophonic music. For example, the left-hand sides of Fig. 3(b) and (c) depict a simple example of polyphony or homophony consisting of two parts and its tree. These time-spans can also represent a single note as on the right side. In Fig. 3(b), the connection between notes in a chord is important, but in Fig. 3(c), that in a melodic part is important.

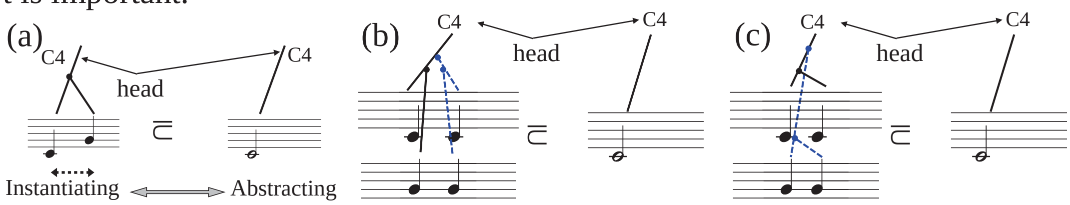

Fig. 3. Subsumption relationship of monophony, homophony, and polyphony.

The rules of GTTM cannot treat polyphony. Even if we acquire the ground truth of GTTM analysis results for polyphony, it is difficult to implement a system which can retrieve a time-span tree from a polymorphic piece of music. For example, the grouping analysis is limited in monophony because those rules imply a single note sequence.

We propose an algorithm to treat polyphony and propose an additional novel rule that enables a polyphonic music time-span tree to be acquired in section 4.

Less precise explanation of feedback link. The GTTM has rules for feedback links from higher to lower level structures, e.g. GPR7 prefers a grouping structure that results in a more stable time-span and/or prolongation reductions. However, no detailed description and only a few examples are given.

To solve this problem, we developed an interactive analyzer with which a user can acquire the target analysis results by iterating the automatic and manual processes interactively and easily reflects his or her interpretations on a piece of music.

3 Manual Acquisition of Time-span Tree for Polyphonic Music

We attempted to acquire the time-span tree from polyphony manually in order to investigate whether the time-span tree was actually acquirable from polyphony or not. We could easily consider a bottom up way to construct a time-span tree of polyphony; that is, we first constructed sub-trees of musical phrases in each part and then connected the heads of two trees and make a new head iteratively. However, this bottom-up approach did not work well when a musicologist attempted to acquire a time-span tree of polyphony because it is difficult to select two sub-trees to make a new head. Therefore, we consider another approach based on a polyphonic reduction process.

Fig. 4 shows the process for acquiring the time-span tree from polyphony manually. First, we asked musicologists to arrange polyphony to homophony (Fig. 4(a)). This arrangement process is very similar to time-span reduction, which removes ornamental notes and acquires the abstracted melody. Then we asked the musicologists to analyze the homophony and acquire the time-span tree on the basis of GTTM (Fig. 4(b)). Finally, we tried to acquire the time-span tree from polyphony by tracing the inverse process of the arrangement that instantiates the ornamental note and adds the branch of the note to the time-span tree (Fig. 4(c)).

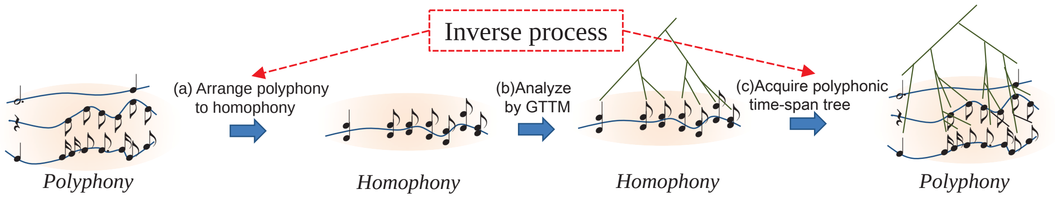

Fig. 4. Manual process for acquiring time-span tree of polyphony.

Record Arrangement Process. We recorded five musicologists' processes for arranging a polyphonic orchestral score into a homophonic piano score. To video a composer's visual points, we used an eye mark camera. We also used microphones to record each musicologist's voice when he/she was thinking aloud. After the music arranging processes had been completed, we asked the musicologists about the details of their arrangement processes using the video from the eye mark camera. For example, we asked composers the following questions: Why did you focus on this section for a long time? What bothers you now about your arrangement? Do you have any better ideas for the arrangement?, and so on.

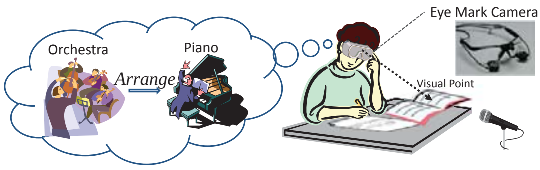

Fig. 5. Recording arrangement process.

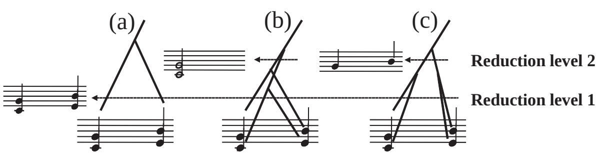

Fig. 6. Refinement of time-span tree.

Refinement of time-span tree. By refining the time-span tree of homophony, we can acquire a time-span tree applicable to polyphony. Fig. 6(a) is a time-span tree of homophony in which each branch of the tree is connected to a chord. Fig. 6(b) and (c) show the refined time-span trees of homophony.

If we slice these three time span trees by using reduction level 1, all the results are the same: two chords of quarter notes. When we abstract the tree in Fig. 6(b) by using reduction level 2, the result is a chord of half notes. Thus, the time-span tree like that in Fig. 6(b) will be formed when there are important chords in the phrase. On the other hand, when we abstract the tree in Fig. 6(c) using reduction level 2, the result remains one voice consisting of two quarter notes. Thus, the time-span tree in Fig. 6(c) will be formed when there is an important voice in the phrase, such as unison.

The same musicologists refined the time-span tree of homophony and decided the type of time-span tree. As a result, the refinement was applicable only in the smallest group in which the grouping boundaries of multiple voices are the same.

Manual Acquisition of Polyphonic Music Time-span Tree. If we assume a time-span tree can be acquired from polyphony, a subsumption relationship is formed between the time-span tree of polyphony and that from homophony acquired and refined by the musicologists. To acquire the time-span tree from polyphony, we manually add the omitted ornamental notes and its branches one by one by tracing the inverse process of the arrangement (Fig. 7). The note that connects to the omitted note by the branches can be detected as either of the following.

- Head of the smallest time-span that includes the omitted note
- Head of another voice time-span that is the same as or similar to the time-span that includes the omitted note

Fig. 8(a) shows an example of the former, where the smallest time-span that includes omitted note 2 in time-span *b*. Therefore, note 2 is connected to note 3. Fig. 8(b) shows an example of the latter, where omitted note 3 in time-span *d* is connected to note 1 in time-span *c* because time-spans *c* and *d* are the same. When connecting the branch of note 1, note 3 is at a higher position than note 2; otherwise, note 3 in time-span *d* is connected to note 1 in time-span *a*, and the time-spans *d* and *a* are different.

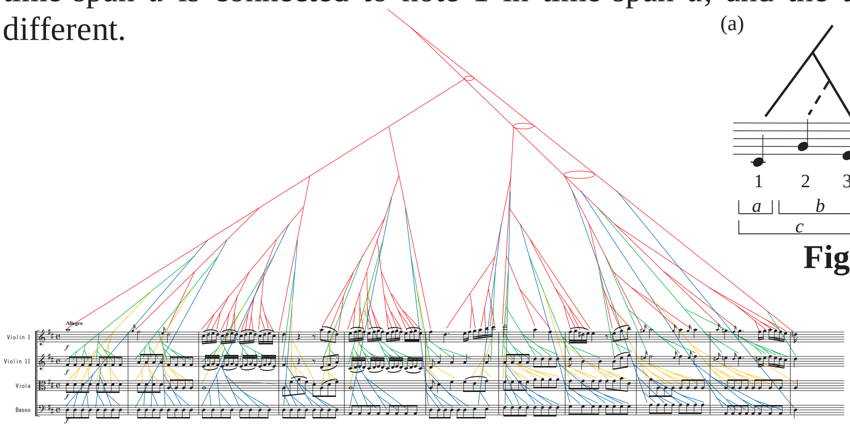

Fig. 7. Polyphonic time-span trees.

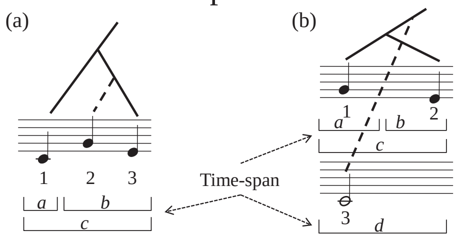

Fig. 8. Adding a note to time-span tree.

4 Implementing PTTA

In this section, we describe the system to acquire a time-span tree from a polymorphic piece of music. Fig. 9 shows an overview of our polyphonic music time-span tree analyzer (PTTA).

In the initial input we use MusicXML of polyphony, in which each voice is separated. We also need MusicXML of homophony and its time-span tree that is manually arranged and analyzed by musicologists.

The part divider splits each voice and outputs the MusicXML of monophonies. The harmonic analyzer analyzes the harmony using Tonal Pitch Space [8], which was written by Lerdahl, an author of GTTM, and is implemented the same way as in the work of Sakamoto and Tojo [9].

We designed the PTTA to separate the input polyphony to each voice part of the monophonic melody and to analyze in a parallel manner using a grouping structure analyzer and a metrical structure analyzer because the grouping and metrical structures of each part of polyphony sometimes differ. The contents of the grouping structure analyzer and metrical structure analyzer are the same as those of the monophonic version in the ATTA.

On the other hand, the time-span analyzers analyze all the parts of polyphony together because only one time-span tree is acquired from polyphony.

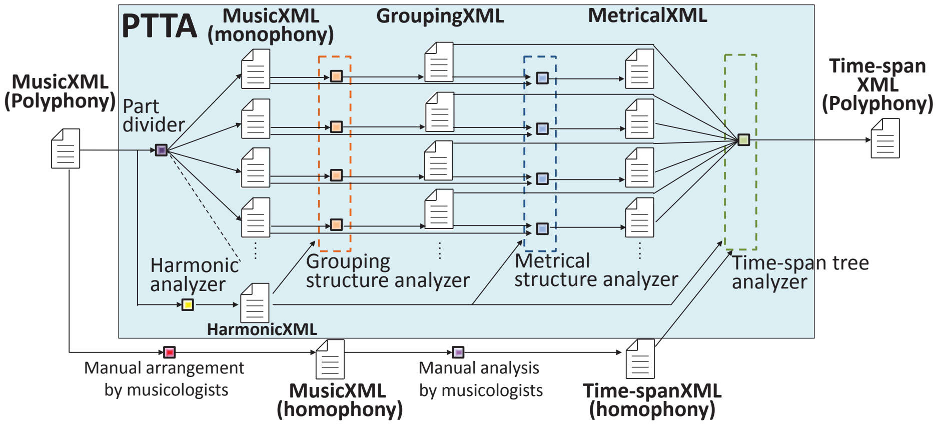

Fig. 9. Overview of PTTA.

Overview of time-span tree analyzer in PTTA. The time-span tree analyzer of PTTA differs from that of ATTA. However, the algorithm for acquiring a time-span tree is the same.

(1) Consider all the notes as a head.

(2) Apply time-span tree preference rules (TSRPR) to local-level heads $D_i^{TSRPR_R}$, where R is an index of preference rule in GTTM [4].

(3) Calculate the head strength $D^{\mathrm{timespan}}(i)$ at a local level.

(4) Select the next-level head from each time-span.

(5) Iterate steps 2 to 4 as long as the time-span contains more than one head.

Application of time-span reduction preference rules. We implemented seven out of nine time-span reduction preference rules (TSRPR). We did not implement those concerning feedback loops: TSRPR5 (metrical stability) and TSRPR6 (prolongational stability). $D_i^{\mathrm{TSRPR}_R}$ indicates whether TSRPR_R holds. Since the priority among these TSRPRs is not shown in GTTM [4], we introduce adjustable parameter $S_{\mathrm{TSRPR}R}$.

For example, TSRPR7 prefers that a head i appear at a cadence. D_i^{TSRPR7} returns 1 if the head is at the cadence position and 0 otherwise:

$$D_i^{\mathrm{TSRPR7}} = \begin{cases} 1 & i \text{ appear at cadence} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Novel rule of TSRPR. To treat polyphonic music on the basis of GTTM, we propose an additional novel rule that we call TSRPR10.

TSRPR10 (Time-span Trees Interaction) Prefer a time-span analysis of polyphonic that minimizes the conflict between the time-span tree of polyphony and the time-span tree of homophony that is arranged and analyzed by a musicologist.

We express the degree of application of the TSRPR10 as follows:

$$D_i^{\mathrm{TSRPR10}} = h_i \Big/ \max_j h_j \tag{2}$$

where i is note transition and h_i is the number of time-spans whose head is i from the refined time-span tree of homophony. If i is an important note, note i can be a head in several hierarchies of the time-span. The denominator max h_j is for normalization. We added the adjustable parameter S_{TSRPR10} to control the strength of the rule.

Generation of time-span tree. We calculate the plausibility of the head $D^{\mathrm{timespan}}(i)$ by $D_i^{\mathrm{TSRPR}_R}$ and adjustable parameter $S_{\mathrm{TSRPR}R}$.

$$D^{\mathrm{timespan}}(i) = \sum_{R} D_i^{TSRPR_R} \times S_{TSRPR_R} \qquad (3)$$

A hierarchical time-span tree is constructed by iterating the calculation of the plausibility of the head $D^{\mathrm{timespan}}(i)$ for the current heads and choosing the heads of the next level.

Interactive analyzer for PTTA. Because GTTM has a feedback link from higher to lower level structures, the analyzing process is not straightforward. We therefore developed an interactive analyzer that enables the use of analyzers of the grouping structure, metrical structure, and time-span tree and of a manual editor for each structure in alternative orders.

Fig. 10 shows a screen snapshot of the interactive analyzer, where polyphonic sequences are displayed in a piano roll format. Each part of the sequence is shown in a different color. When a user selects one part, the grouping and metrical structures of the sequence are displayed below it.

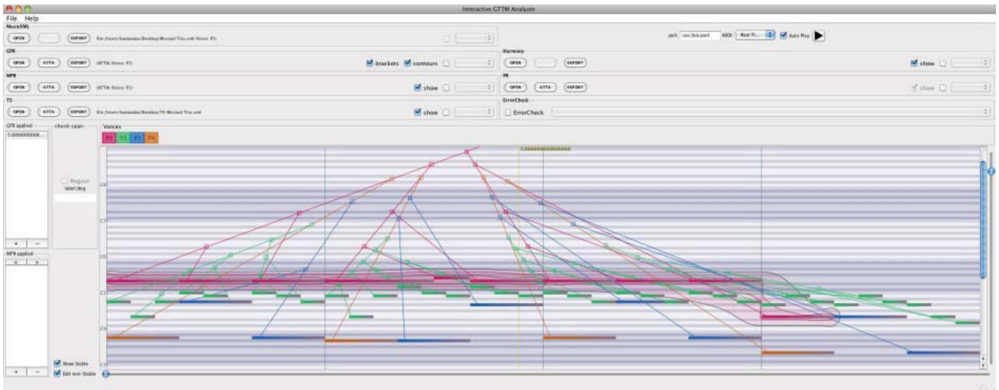

Fig. 10. Interactive analyzer for PTTA.

5 Experimental Results

We evaluated the performance of the time-span analyzer of PTTA using an F-measure, which is given by the weighted harmonic mean of Precision P (proportion of selected heads that are correct) and Recall R (proportion of correct heads that were identified).

This evaluation required us to prepare correct data of the grouping structure, metrical structure, time-span tree, and harmonic progression of the polyphony and needed the time-span tree of homophony, which was arranged and analyzed by a musicologist. We collected 30 eight-bar-length, polyphonic, classical music pieces and asked the musicology experts to analyze them manually and faithfully with regard to GTTM. Three other experts crosschecked these manually produced results.

We compared the baseline performance where we fixed S_{TSRPR10} to zero and other parameters were configured by hand, which means the analyzer did not use TSRPR10.

It took us an average of about ten minutes per piece to find the plausible tuning for the set of parameters. As a result of configuring the parameters, the PTTA using TSRPR10 outperformed the baseline not using TSRPR10 (Table 1).

Table 1. F-measure for our method

Melody	Baseline not using TSRPR10	Our system using TSRPR10
1. Borodin Streichquartett No.2 3rd mov.	0.27	0.91
2. Mozart Eine Kleine Nachtmusik K525 2nd mov.	0.41	0.76
3. Beethoven Streichquartett Op.18 No.4 1st mov.	0.52	0.95
4. Haydn Quartett "Kaiser" Op.76 No.3 1st mov.	0.13	0.42
5. Brahms Streichquartett No.2 op.51-2	0.24	0.64
:	:	:
Total (30 Polyphonies)	0.36	0.89

6 Conclusion

We developed a music analysis system called a polyphonic music time-span tree analyzer (PTTA) that enables time-span trees to be acquired from polyphony. To treat polyphonic music on the basis of the generative theory of tonal music (GTTM), we propose a new preference rule called TSPRP10 for time-span reduction that prefers a time-span analysis of polyphony that minimizes the conflict between time-span trees of homophony, which is arranged and analyzed by musicologists. Experimental results showed that by using TSRPR10, our PTTA outperformed the baseline.

Since we hope to contribute to the research of music analysis, we will publicize our PTTA with an interactive analyzer and a dataset of a hundred pairs of a polyphonic score and musicologists' analysis results on our website:

http://music.iit.tsukuba.ac.jp/hamanaka/gttm.htm.

We plan to develop further systems using time-span trees of polyphony for other musical tasks, such as searching, harmonizing, voicing, and ad-libbing. Such systems will help musical novices to manipulate music.

References

1. Apple – Garageband: http://www.apple.com/ilife/garageband/ (2013)
2. Clausen, M., Engelbrecht, R., Meyer, D., and Schmitz, J.: PROMS: A web-based tool for searching in polyphonic music. In Proceedings of the 1st International Symposium on Music Information Retrieval, 10 pages, (2000)
3. Typke, R., Veltkamop, R., and Wiering, F.: Searching notated polyphonic music using transportation distance, In Proceedings of the 12th ACM international conference on Multimedia 2004, pp. 128-135, (2004)
4. Lerdahl F., Jackendoff R.: A generative theory of tonal music. Cambridge, Massachusetts: MIT Press (1983)
5. Hirata K., and Aoyagi T.: Computational music representation based on the generative theory of tonal music and the deductive object-oriented database. Computer Music Journal, 27(3)73-89 (2003)
6. Hamanaka M., Hirata K., Tojo S.: Melody extrapolation in gttm approach. In: Proceedings of the 2009 International Computer Music Conference (ICMC2009), pp. 89-92 (2009)
7. Hamanaka M., Hirata K., Tojo S.: Implementing "a generating theory of tonal music". Jornal of New Music Research (JNMR) 35(4) 249-277 (2007)
8. Lerdahl F.: Tonal pitch space. Oxford University Press, New York (2001)
9. Sakamoto S., Tojo S.: Harmony analysis of music in tonal pitch space. Information Processing Society of Japan SIG technical report, vol. 2009 (in Japanese) (2009)