Polyphonic Music Analysis Database Based on GTTM

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

[1] RIKEN
[2] Future University Hakodata
[3] JAIST
masatoshi.hamanaka@riken.jp

Abstract. This paper presents a method for manually analyzing polyphonic music on the basis of the generative theory of tonal music and the analysis database obtained with the tool. A time-span tree of the theory assigns a hierarchy of 'structural importance' to the notes of a piece of music. However, the theory only accepts homophonic music. To solve this problem, we first recorded musicologists' processes for arranging from polyphony to homophony because the processes show how a musician reduces ornamental notes. Using the recordings of the arrangement process with the time-span tree of the homophony, we attempted to manually acquire a time-span tree of polyphony. We hope that the database we will publish will be used for various kinds of musical creativity computer simulation.

Keywords: The generative theory of tonal music (GTTM), time-span tree, polyphonic music, musicologist' processes.

1 Introduction

This paper describes a polyphonic music analysis database we have compiled that is based on the generative theory of tonal music (GTTM) [1]. In constructing a musical analyzer or a system for computer simulation of musical creativity, test data from musical databases is very useful for evaluating and improving the performance of the system. The Essen folk song collection is a database for folk-music research that contains score data on 20,000 songs along with phrase segmentation information and also provides software for processing the data [2]. The Répertoire International des Sources Musicales (RISM), an international, non-profit organization with the aim of comprehensively documenting extant musical sources around the world, provides an online catalogue containing over 850,000 recordings , mostly for music manuscripts [3]. The Variations3 project provides online access to streaming audio and scanned score images for the music community with a flexible access control framework [4], and the Real World Computing (RWC) Music Database is a copyright-cleared music database that contains the audio signals and corresponding standard MIDI files for 315 musical pieces [5]. The Digital Archive of Finnish Folk Tunes provides 8,613 Finish folk song midi files with annotated meta data and Matlab data matrix encoded by midi toolbox [6]. The Codaich contains 20,849 MP3 recordings, from 1941 artists, with high-quality annotations [7], and the Latin Music Database contains 3,227 MP3 files from different music genres [8].

We previously published a monophonic version of our GTTM-based analysis database and analysis tool [9]. However, we were not able to develop and publish a polyphonic version because the GTTM theory is limited to homophonic music.

In the work described in this paper, we attempted to acquire a time-span tree from polyphony manually to determine whether one could actually be acquired. Results of GTTM-based analysis showed that 27 out of 30 polyphonic pieces could acquire a time-span tree of polyphony with multiple parts. However, the remaining three pieces, which had independent parts, could not.

The rest of this paper is organized as follows. We discuss the problem of acquiring a time-span tree from polyphony in Section 2 and whether one can actually be acquired from polyphony in Section 3. We present experimental results in Section 4 and conclude with a summary of key points and a mention of future work in Section 5.

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 respectively output a grouping structure, a metrical structure, a time-span tree, and a prolongational tree (Fig. 1).

The grouping structure is intended to formalize the intuitive belief that tonal music is organized into groups that are in turn composed of subgroups. These groups are graphically presented as several levels of arcs below a musical staff. The metrical structure describes the rhythmical hierarchy of the piece by identifying the position of strong beats at the levels of a quarter note, half note, a measure, two measures, four measures, and so on. Strong beats are illustrated as several levels of dots below the staff. A time-span tree is a binary tree, which is a hierarchical structure describing the relative structural importance of notes that differentiate the essential parts of the melody from the ornamentation. Figure 2 is an example of abstracting a melody by using a time-span tree. The figure shows a time-span tree from melody *A* that embodies the results of GTTM analyses. In the time-span tree, the important notes are connected to a branch nearer the root of the tree. In contrast, the unimportant notes, that is, the ornaments, are connected to the leaves of the tree. We can obtain an abstracted melody *B* by slicing the tree in the middle and omitting notes that are connected to branches under line *B*. In the same manner, if we slice the tree higher up at line *C*, we can obtain a more abstracted melody *C*.

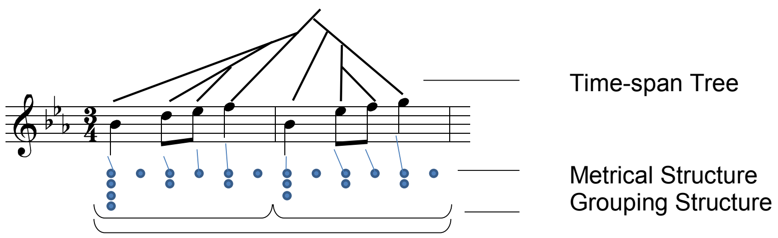

Fig. 1. Grouping structure, metrical structure, time-span tree of GTTM.

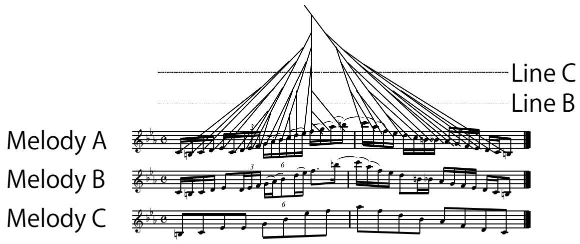

Fig. 2. Abstraction of melody.

In our previous work, we developed GTTM analyzers called ATTA [10], FATTA [11], σGTTM [12], and deepGTTM [13]. However, all those systems only allow monophony. This limitation is too narrow because users may want to manipulate polyphonic or homophonic music. Here, we discuss the problems in developing a system that enables polyphony and homophony to be treated.

If we analyze polyphony using GTTM, we find there are no ground truth data because the theory is limited to treating homophony. An indication of this is the fact that Lerdahl and Jackendoff's original published work on GTTM [1] contains no examples of polyphony being treated.

The time-span reduction represents an intuitive idea: if we remove ornamental notes from a long melody, we obtain a simpler, similar sounding melody. An entire piece of Western tonal music can eventually be reduced to one 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.

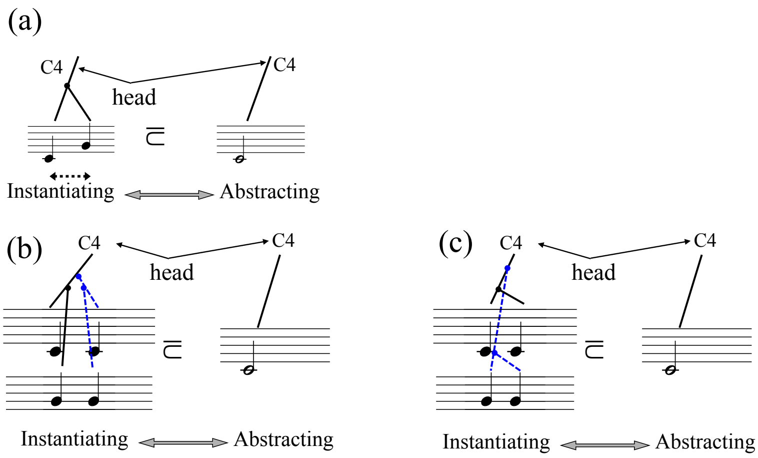

Fig. 3. Subsumption relationship of monophony, homophony, and polyphony.

3 Manual Acquisition of Time-span Tree for Polyphonic Music

It was easy for us to 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 made 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 considered another approach based on a polyphonic reduction process.

Figure 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 notes and adds their branches to the time-span tree (Fig. 4(c)).

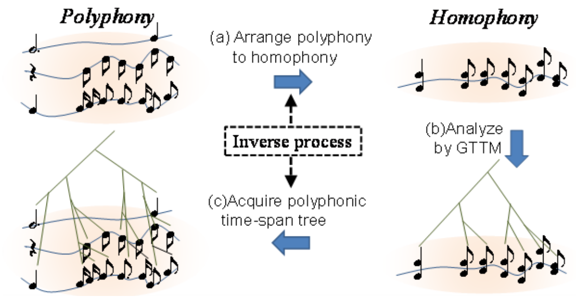

Fig. 4. Manual process for acquiring time-span tree of polyphony.

Recording Arrangement Process. We recorded three musicologists' processes for arranging a polyphonic orchestral score into a homophonic piano score. To video a musicologist'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 queried the musicologists about the details of their arrangement processes using the video from the eye mark camera. For example, we asked them a number of 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.

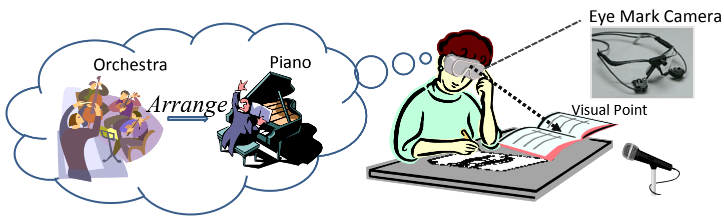

Fig. 5. Recording arrangement process.

Refinement of Time-span Tree. By refining the time-span tree of homophony, we can acquire a time-span tree applicable to polyphony. Figure 6(a) is a time-span tree of homophony in which each branch of the tree is connected to a chord. Figure 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, a 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 to determine the type of time-span tree it was. The refinement was found to be applicable only in the smallest group in which the grouping boundaries of multiple voices were the same.

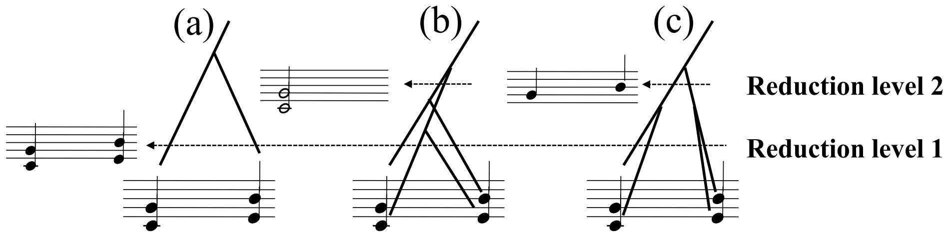

Fig. 6. Refinement of time-span tree.

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 the homophony acquired and refined by the musicologists. To acquire the time-span tree from polyphony, we manually add the omitted ornamental notes and their branches one by one by tracing the inverse process of the arrangement (Fig. 7). A note whose branch connects it to an omitted note 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

Figure 8(a) shows an example of the former, where the detected note is the head of the smallest time-span that includes omitted note 2 in time-span *b*. Therefore, note 2 is connected to note 3. Figure 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.

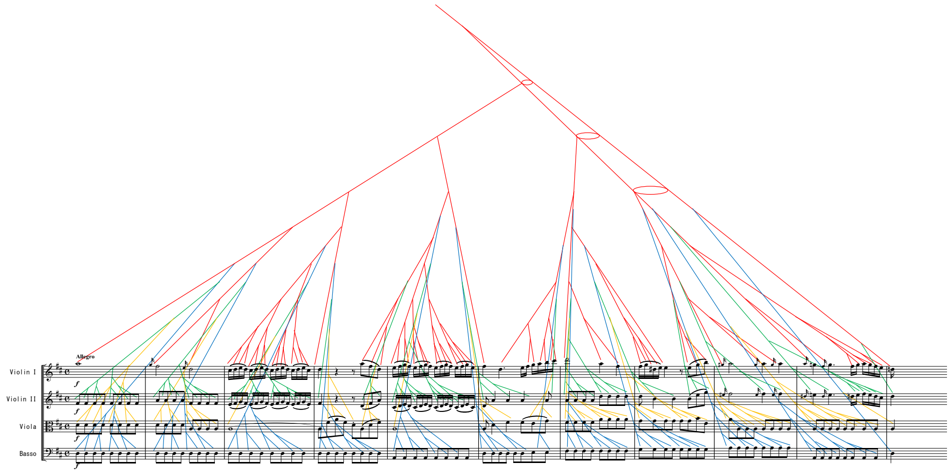

Fig. 7. Polyphonic time-span trees.

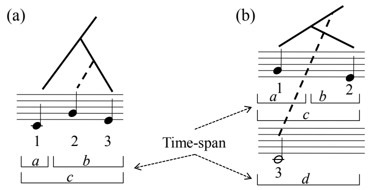

Fig. 8. Adding a note to a time-span tree.

Analyzing Tool for Polyphonic Music. Because GTTM has a feedback link from higher to lower level structures, the analyzing process is not straightforward. We therefore developed an analyzing tool 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.

Figure 9 shows a screen snapshot of the analyzing tool, 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.

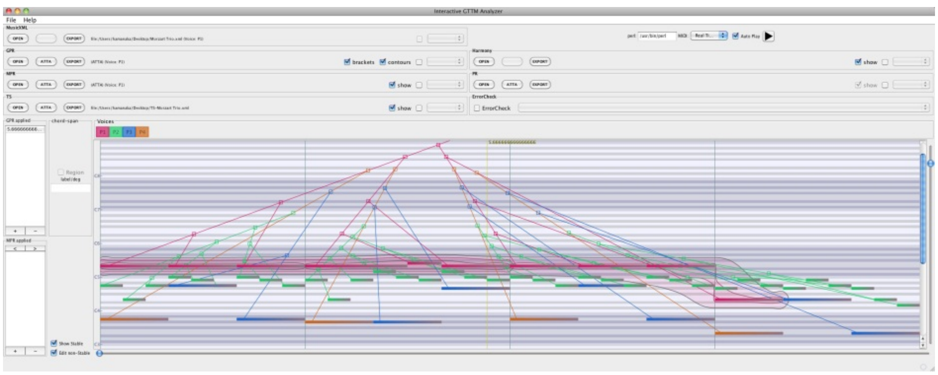

Fig. 9. Interactive analyzer for PTTA.

4 Experimental Results

We attempted to manually analyze polyphonic pieces of music on the basis of the method described in Section 3. Three musicologists each collected 10 32-bar long, polyphonic, classical music pieces including notes, rests, slurs, accents, and articulations entered manually with music notation software called Finale [14]. Analysis results showed that 27 out of the 30 pieces could acquire a time-span tree that integrates multiple parts in the polyphony. Figure 10 shows one piece out of the remaining three: the second movement of Beethoven's String Quartet Op. 18 No. 4. This piece is in fugue form, that is, the first part is the subject, the second part is the free counterpoint, the third part is the counter-subject, and a new subject is introduced near the end of the fourth part. In a fugue, there is no master-slave relationship between the voices because each voice is completely independent. We plan to investigate more about the parts in a piece of music moving independently.

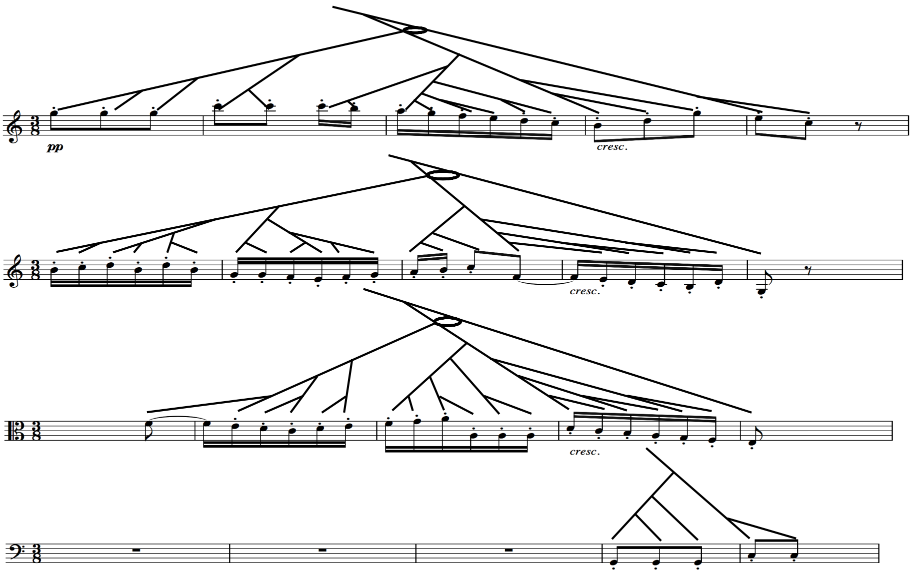

Fig. 10. Analysis result of Beethoven's String Quartet Op. 18 No. 4.

5 Conclusion

We described a method we propose for manually analyzing polyphonic music on the basis of the generative theory of tonal music (GTTM). First, we recorded three musicologists' processes for arranging a polyphonic orchestral score into a homophonic piano score, using an eye mark camera to video their visual points. Then we asked them to analyze the homophony and acquire the time-span tree on the basis of GTTM. Finally, we tried to acquire the time-span tree from polyphony by tracing the inverse process of the arrangement that instantiates ornamental notes and adds their branches to the time-span tree.

Results obtained in analyzing 30 pieces of polyphonic music collected by the musicologists showed that 27 of them could acquire an integrated time-span tree of parts. However, the remaining three pieces, which had independent parts, could not. Since we hope to make further contributions to music analysis research, we will publicize our dataset of 27 pairs of polyphonic scores and musicologists' analysis results on our website:

http://gttm.jp/

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.

Acknowledgments. This work was supported by JSPS KAKENHI Grant Number 17H01847, 25700036, and 16H01744.

References

1. Lerdahl F., Jackendoff R.: A generative theory of tonal music. Cambridge, Massachusetts: MIT Press (1983)
2. Schaffrath H.: The Essen associative code: A code for folksong analysis. In E. Selfridge-Field (Ed.), *Beyond MIDI: The Handbook of Musical Codes*, Chapter 24, pp. 343–361. MIT Press, Cambridge, 1997.
3. RISM: International inventory of musical sources. In Series A/II Music manuscripts after 1600, K. G. Saur Verlag, 1997.
4. Riley, J., Hunter, C., Colvard, C., and Berry, A.: Definition of a FRBR-based Metadata Model for the Indiana University Variations3 Project , http://www.dlib.indiana.edu/projects/variations3/docs/v3FRBRreport.pdf, 2007.
5. Goto, M., Hashiguchi, H., Nishimura, T., and Oka, R.: RWC Music Database: Popular, Classical, and Jazz Music Databases, *Proceedings of the 3rd International Conference on Music Information Retrieval* (ISMIR 2002), pp. 287–288, October 2002.
6. Eerola, T., and Toiviainen, P.: The Digital Archive of Finnish Folk Tunes, University of Jyvaskyla, Available online at: http://wwwjyufi/musica/sks, 2004.
7. McKay, C., McEnnis, D., and Fujinaga, I.: A large publicly accessible prototype audio database for music research, *Proceedings of the 7th International Conference on Music Information Retrieval* (ISMIR 2006), pp. 160–164, October 2006.
8. Carlos, N., Alessandro, L., and Celso, A.: The Latin Music Database, *Proceedings of the 9th International Conference on Music Information Retrieval* (ISMIR2008), pp. 451–456, September 2008.
9. Hamanaka, M., Hirata, K., Tojo, S.: Musical Structural Analysis Database Based on GTTM, *Proceedings of the 15th International Conference on Music Information Retrieval* (ISMIR 2014), pp.325-330, October 2014.
10. Hamanaka, M., Hirata, K., Tojo, S.: Implementing A Generating Theory of Tonal Music, *Journal of New Music Research* (JNMR), Vol.35, No.4, pp.249-277, 2007.
11. Hamanaka, M., Hirata, K., Tojo, S.: FATTA: Full Automatic Time-span Tree Analyzer, *Proceedings of the 2007 International Computer Music Conference* (ICMC2007), Vol.1, pp.153-156, August 2007.
12. Hamanaka, M., Hirata, K., Tojo, S.:σGTTM III: Learning based Time-span Tree Generator based on PCFG, *Proceedings of The 11th International Symposium on Computer Music Multidisciplinary Research* (CMMR 2015), pp.303-317, June 16-19, 2015.
13. Hamanaka, M., Hirata, K., Tojo, S.: deepGTTM-II: Automatic Generation of Metrical Structure based on Deep Learning Technique, *13th Sound and Music Conference* (SMC2016), pp.221-249, 2016.
14. MakeMusic Inc., "Finale," 2017, http://www.finalemusic.com/.