LF-5

Applying Voronoi Diagrams in the Automatic Grouping of Polyphony

Masatoshi Hamanaka† Keiji Hirata‡

Abstract

This paper presents a method of grouping that provides a way to automatically analyze the grouping structure of polyphony. Sets of grouping rules have been applied in previous methods, but these methods were not capable of resolving contradictions among the rules. In order to solve this problem, we propose a method where Voronoi diagrams of the notes on the piano-roll format are used to form the groups.

1． Introduction

The purpose of this study is to automatically divide the notes in polyphony into groups. Two kinds of grouping appear in polyphony, one in the direction of pitch interval and the other in the direction of time. Grouping in the interval direction divides polyphony into multiple homophony (Figure 1a). In the time direction, notes are grouped from time gap to time gap (Figure 1b). Automation of these two forms of division not only allows us to more deeply analyze musical structures [1]. Moreover, creating a database of the groups obtained by division gives us the ability to reuse these groups in systems for *automatic music composition* and *jam session system* [2].

GTTM (generative theory of tonal music) [3] is a theory of music which includes a concept of grouping and thus can be used for deriving the set of rules for the division of notes into groups. We think that GTTM is the most promising theory of music in terms of computer implementation, but there is no strict order for applying the rules of GTTM. This may lead to ambiguities in terms of the results of analysis. The implementation of GTTM as a computer system has been attempted [4], but the resulting system was only capable of dealing with the limited polyphony which is made up of two monophonies.

In this paper, we propose a method of grouping which is based on applying the Voronoi diagram. We have developed a method of grouping rather than naively implementing GTTM so that a result by our method is equifinal to that of GTTM. We compare the results of grouping by this method with the results of grouping by a human according to GTTM.

a: Grouping in the direction of pitch interval.

b: Grouping in the direction of time.

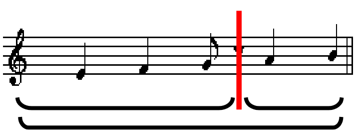

Figure 1: Examples of grouping

† University of Tsukuba
‡ NTT Communication Science Laboratories

2． Generative Theory of Tonal Music

The generative theory of tonal music is composed of four modules, in each of which assigns a separate part of the structural description of a listener's understanding of music. The four modules of GTTM are the *grouping structure*, the *metrical structure*, the *time span reduction* and the *prolongational reduction.*

Grouping structure is intended to formalize the intuition that tonal music is organized into groups, which are in turn composed of subgroups. The rules for grouping of GTTM are of two kinds: *grouping well-formedness rules* and *grouping preference rules.* Grouping well-formedness rules are necessary conditions for the assignment of a grouping structure and restrictions on the generated structures. When more than one structure may satisfy the grouping well-formedness rules, grouping prefernce rules only suggest the superiority of one structure over another, not a deterministic procedure. This is a problem of ambiguity, as described above.

3. Voronoi diagrams used for grouping

To overcome the problem, we propose a method of grouping based on the use of Voronoi diagrams. The result of GTTM is a binary tree that indicates the hierarchical structure of a piece of music. In our method, the hierarchical structure of a piece of music is represented by Voronoi diagrams on a piano-roll.

3.1 Voronoi diagram

A Voronoi diagram partitions a plane into cells, designated by the set of points P, that is, $\{p_1, p_2, ..., p_n\}$, such that for each cell which corresponds to point p_i, all points q in that cell are nearer to p_i than to any other point in P (Figure 2)[5].

$$\mathrm{dist}\,(q, p_i) < \mathrm{dist}\,(q, p_j) \quad (p_i, p_j \in P, j \neq i) \qquad (1)$$

We can form the Voronoi diagram for a given set of points in a plane as a connected set of segments of half-plane boundaries, where each of the half-planes is formed by partitioning the plane into two half-planes, one on either side of the bisector of the line between each adjacent pair p_i and p_j.

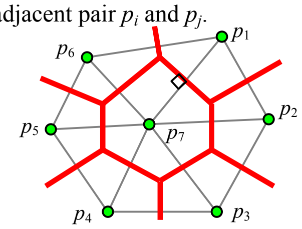

Figure 2: A Voronoi diagram.

3.2 Voronoi diagram for two notes

Our method adopts the piano-roll format as a score, and thus notes are expressed as horizontal line segments on a piano-roll. To construct a Voronoi diagram on the score, we need to consider the Voronoi diagram for multiple horizontal line segments, which is constructed of linear segments and quadratic segments.

When two notes are sounding at the same time or no note is sounding, the corresponding part of the Voronoi diagram is a linear segment (Figure 3a, Figure 3c). When a single note is sounding, the Voronoi diagram becomes a quadratic segment (Figure 3b).

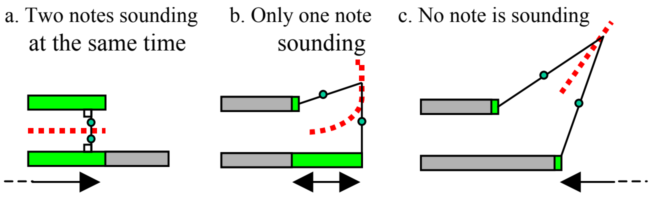

Figure 3: Voronoi diagrams for pairs of notes.

3.3 Voronoi diagram for more than two notes

To construct the Voronoi diagram for more than two notes, we construct the Voronoi diagrams for all of the pairs of notes and delete the irrelevant segments. For example, to construct the Voronoi diagram for notes a, b, and c, we construct three Voronoi diagrams (Figure 4). The boundaries in the three Voronoi diagrams then intersect at a point which is equidistant from each of the notes. The Voronoi diagram for notes a and b is divided into two half-lines at the intersection. We then delete that half-line which is nearer to c than to a or b.

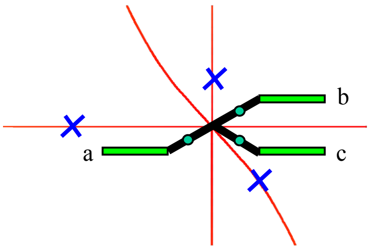

Figure 4: Voronoi diagram for more than two notes.

3.4 Scaling problem

A problem of scaling is inevitable when we construct a Voronoi diagram of notes in the piano-roll format. Since the vertical axis of the piano-roll score indicates the pitch (in semitones) and the horizontal axis indicates time (in ticks), the shape of a Voronoi diagram depends on the ratio of the pitch scale to the time scale. In this paper, we experiment with the following three scaling ratios, 1 semitone = 90 ticks, 1 semitone = 100 ticks, and 1 semitone = 110 ticks. The rules of GTTM in terms of time interval and register (GPR2 and GPR3) possibly contradict each other, but our proposed method avoids such a contradiction by setting the scaling ratio appropriately.

3.5 Making groups

Hierarchical groups of notes were constructed by converging adjacent notes iteratively. Here, we introduce the following principle for making groups; *the smallest Voronoi cell is first merged to an adjacent group.*

By changing the scaling ratio, a distinct Voronoi cell becomes smallest, and thus a merging pattern also changes.

4. Experimental Results

We have implemented our method of grouping and compare the results with correct data obtained by a human (Figure 5). We evaluate the performance of grouping by using a correctness rate that we define as follows.

$$\text{Correctness rate} = \frac{\text{The number of notes grouped correctly}}{\text{The number of notes}} \quad (2)$$

We gave three different scaling ratios, run the program, and calculated the correctness rate (Table 1). The correctness rate for each of the scaling ratios is above 70 percent. The tune that we used in this experiment was the MIDI data of *Turkish March*.

Table 1: The performance of our method with three scaling ratios.

Scaling ratio (ticks/semitone)	90	100	110
Correctness rate	78.5%	78.5%	72.8%

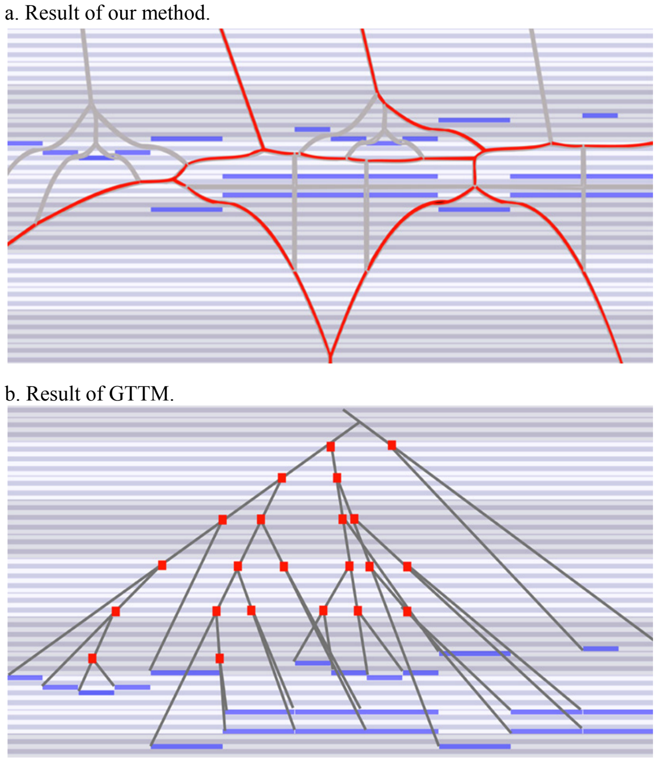

Figure 5: Results of our method and GTTM.

5. Conclusion

This paper described the grouping method based on the Voronoi diagrams. This method makes it possible to automatically analyze the grouping structure of polyphony. Experimental results showed that our method performed correctly, with a correctness rate being above 70 percent.

We now plan to use this method for the automatic creation of phrase databases for our jam session system and to apply it to the analysis of deep musical structure.

6. References

[1] Keiji Hirata, Tatsuya Aoyagi: "Representation Method and Primitive Operations for a Polyphony based on Music Theory GTTM," Journal of IPS Japan, Vol. 42, No. 3, 2002. (in Japanese)

[2] Masatoshi Hamanaka, Masataka Goto, Nobuyuki Otsu: "Learning-Based Jam Session System for a Guitar Trio," Proc. of ICMC, 2001.

[3] Fred Lerdahl, Ray Jackendoff: "A Generative Theory of Tonal Music," The MIT Press, 1983.

[4] Kentarou Ida, Keiji Hirata, Satoshi Tojo: "The Attempt of the Automatic Analysis of the Grouping Structure and Metrical Structure Based on GTTM," Trans. of IPS Japan, Vol. 2001, No. 42, 2001. (in Japanese)

[5] Franz Aurenhammer: "Voronoi Diagrams - a Survey of Fundamental Geometric Data Structure," ACM Computing Surveys, Vol. 23, 1991.