

Augmented Music Composition Support as Active Mining

Akinori ABE[†], Rodney BERRY^{††}, Masami SUZUKI^{††}, and Norihiro HAGITA[†]

[†] ATR Intelligent Robotics & Communication Labs

^{††} ATR Media Information Science Labs

2-2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288 JAPAN

E-mail: †{ave,rodney,msuzuki,hagita}@atr.co.jp

Abstract Since the middle of the 20th century, various types of automatic music composition have been attempted. Usually, these attempts utilized random number or stochastic information to generate music. In these trials, the researches focused on the automatic generation of music. Indeed, their trials are important for art and science. However, from the viewpoint of entertainment, it is more important to generate music by ourselves with the proper assistance or guidance from a computer. We are now developing a symbolically rich interactive music composition system, the Augmented Composer, that uses atomic music patterns and controls their duration, pitch, and velocity to generate music. Its interface is quite simple and easy, so even a non-expert can compose “music”. Through analysis of composed music (MIDI data), we found an explicit difference between music by expert composers and novices composers. In this paper, we briefly introduce the Augmented Composer. We show the difference of induced composition patterns between experts and non-experts. Finally, we propose a music composition support (suggestion) system from the viewpoint of active mining.

Key words augmented music composition, active mining

Augmented Music Composition Support as Active Mining

阿部 明典[†] Rodney Berry^{††} 鈴木 雅実^{††} 萩田 紀博[†]

[†] ATR 知能ロボティクス研究所

^{††} ATR メディア情報科学研究所

〒 619-0288 京都府相楽郡精華町光台 2-2-2

E-mail: †{ave,rodney,msuzuki,hagita}@atr.co.jp

あらまし 20世紀の中頃から様々な自動作曲手法が試されてきた。大抵は、乱数や統計的情報が音楽を生成する為に利用されてきた。確かに、確率的に音楽を自動的に作曲出来るかという試みは芸術的、科学的観点からは重要であった。しかしながら、作曲をエンタテインメントとして考えると、コンピュータによる適当な支援やガイドにより、自分で作曲を行うことが更に重要である。我々は、現在、Augmented Composer というインタラクティブ作曲システムを開発している。Augmented Composer では音のパターンを用い、それらの持続、ピッチ、強さなどを簡易なインタフェースを用いて調節することで、音楽を生成出来る。従って、作曲に習熟していない者でも「音楽」を作ることが出来る。曲は、MIDIの形で得られるが、これをコンピュータで解析した所、プロとアマチュアの間に際立った違いを発見した。本論文では、先ず、Augmented Composer の概要を示し、得られたデータを解析した結果によるプロとアマチュアの作曲の間の違いを示す。その違いに基づいて、我々は、アクティブマイニングから見た、作曲データのマイニングを利用したよりプロに近い作曲に導くような作曲支援システムを提案する。

キーワード augmented music composition, active mining

1. Introduction

Since the middle of the 20th century, various types of automatic music composition have been attempted. The first trial, string quartet No.4 “The Illiac Suite”, was composed by Hiller and Isaacson in 1957 [4]. They used their experience with the ILLIAC for computing polymer conformations and adapted the approach to develop rules for the composition of music. Then, according to the rule, the suite was automatically composed by a probabilistic

method. Xenakis proposed stochastic music [16]. This is a mathematical method to construct unknown, but overall controlled complex sound movements from the repeated random trials. Without any musical premises such as theme, tone row or rhythmic patterns, using only Poisson distributions and other probability functions, one can determine every phase from the pitch, duration, density (number of sound within a unit time), timbre distribution to the overall structure (macro composition) [15]. Thus, they utilized random number or stochastic information to

generate music. In these trials, they focused on the automatic generation of music. That is, these trials aim to make the computer perform like a human beings. Indeed, their trials are important for art and science. However, from the viewpoint of an entertainment, it is more important to generate music by ourselves with the proper assistance or guidance of a computer.

We are now developing a symbolically rich interactive music composition system, the **Augmented Composer** [3], that uses atomic music patterns and controls their duration, pitch, and velocity to generate music. Since the **Augmented Composer** is a real time music generator, the output data is not a graphical score (pattern) but MIDI (Musical Instrument Digital Interface) sequence data (symbol). Thus, we can easily obtain information of a composition as MIDI sequence data that can be regarded as symbols. As shown later, the data from the system can be dealt with by data mining methods with a little modification. Though MIDI data is quite low level, we first analyzed the data for comparison between compositions by experts and non-experts. After analysis of the composed music data, we were impressed to find an explicit difference between music by expert composers and non-expert composers. In fact, we can distinguish music composed by Mozart from that composed by Beethoven. This has been achieved by expertized or intuitive induction. Though there will be some limitations, this type of analysis can also be computationally achieved. As to an analysis of literature, there has been much study for a quantitative analysis that can determine the author of a certain article [6]. Since articles are symbolically written, it is easy to analyze them statistically. For instance, by referring to statistically analyzed results, analyzers can determine whether the work was written by Shakespeare or not. As to an analysis of music, some research has been done. For example, Paelo [10] adopted statistical methods, rule-based analysis, and constraint derivation to analyze a score. The object of analysis is to provide a database search that is a sort of index. It does not aim to find a style of composers automatically. Thus, as for music, there has been study into data mining as library data base support. However, there does not seem to be any research to detect a difference between the style or manner of professionals and amateurs. Furthermore, there does not seem to be any research to automatically find the style or manner of a particular composer in a logical manner. Although, Pachet proposed a playing style leaner and imitator by using Hierarchical Markov models [9].

This paper points out the importance of data mining not for the library but for individual reference as entertainment and education. In addition, this paper points out that the above data mining can be achieved under the framework of active mining. This paper proposes music composition support from the viewpoint of active mining. That is, this paper proposes a user-friendly support system. This paper also analyzes the necessary features for a music composition supporting system and sets forth current problems.

In Section 2, we briefly overview the **Augmented Composer**. In Section 3, we show the analyzed results of the outputs (composition data) from the **Augmented Composer**. In Section 4, we propose a music composition suggestion system from the viewpoint of active mining.

2. Symbolically rich interactive music composition environment

In this section, we briefly introduce a symbolically rich interactive music composition system, the **Augmented**

Composer. The Music Table is the first of a planned series of interface designs under the umbrella title of the **Augmented Composer**. The Music Table is a system that enables a player to compose musical patterns by arranging cards on a tabletop. An overhead camera allows the computer to track the movements and positions of the cards and to provide immediate feedback in the form of music and on-screen computer generated images. The Music Table provides both tactile and visual representations of music that can be easily manipulated to make new musical patterns. In this way, inexperienced music makers can experience their own music as patterns in musical space.

The overlapping tactile and visual representations of the music help reinforce one another in a way that can provide a form of interaction and representation not possible with mouse, keyboard and screen (Fig. 1). By giving the player a physical model of the music, the abstract nature of music can be experienced on an intuitive level.



Fig. 1 Augmented Composer

2.1 The system overview

The software component of the system is in two main parts. The Music Table program itself handles the camera-based tracking of the cards, the rendering of VRML objects, and the compositing of CG graphics into the original video image. To do this, it uses the Augmented Reality Toolkit programming libraries [5] devised by Kato and supported by a lively community of programmers. The sequencer for MIDI events is built in the PD music programming environment [12] and receives data from the Music Table program via MIDI and via UDP sockets. This use of MIDI is a carry-over from the Augmented Groove [11], an earlier project that used the same marker tracking technology to mix live dance music tracks and effects via MIDI.

When a card is placed on the table, and its tracking pattern becomes visible to the computer, the software recognizes the position of the card on the table. A card's position on the axis running toward or away from the user determines the pitch of the note to be played. Its position from left to right determines its position on a looping timeline running across the table. Several cards combined make up new musical patterns that follow a sequence, as if the cards on the table were notes on a musical score.

By rotating a card clockwise, the loudness of its note is increased, as is the spikiness and movement of the animated character associated with that card. Counter-clockwise rotation makes the note less loud. Tilting a card slightly to the left or the right causes the length of the note to be increased or decreased. When this happens, the creature on the screen also becomes longer or shorter (Fig. 2 and Fig. 3).

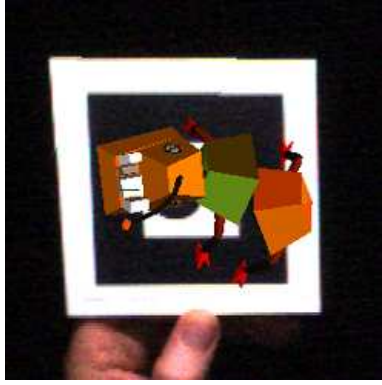


Fig.2 A short, smooth creature corresponding to a short, quiet note



Fig.3 A long, spiky creature corresponding to a long, loud note

The onscreen animated characters serve to complement the physical representation on the table by giving extra information about the notes being played. When a note's turn comes to actually make a sound, the onscreen character will suddenly jump to give the user some indication of the state of the timeline.

Other cards are used to save and copy patterns and to change instrument sounds for making multi-layered musical compositions. For the purposes of this particular research however, just the basic note cards were used to simplify the data collection.

For each iteration of the timeline, the positions of the note cards and their respective MIDI values were recorded as a line in a text file. The resultant files represent around five minutes of interaction by one person.

3. Obtaining user models from composed data

3.1 Data from the Augmented Composer

As shown in the previous section, the **Augmented Composer** works as music composition support, so even a novice can enjoy music composition. The result from the **Augmented Composer** is not a traditional score, but an MIDI sequence data. Usually, the system is connected to an MIDI apparatus to play music. In fact, the data is an ASCII data and not a standard MIDI file.

For example, a part of the MIDI data is shown in Fig. 4, where “loop” means the number of cycles of the musical phrase (all lines prefixed by loop 1 belong to the same musical pattern), “card” means the ID (identification) of card (1–10), “pitch” means the MIDI note number on the pitch axis, “position” means the note’s position on the time axis range (0–7), “on_off” means whether the

```

loop 1 card 1 pitch 61 position 5 on_off 0
      velocity 124 duration 17;
loop 1 card 2 pitch 55 position 2 on_off 0
      velocity 0 duration 24;
...
loop 1 card 10 pitch 59 position 3 on_off 0
      velocity 121 duration 17;
loop 2 card 1 pitch 61 position 5 on_off 0
      velocity 124 duration 17;
loop 2 card 2 pitch 55 position 2 on_off 0
      velocity 0 duration 24;
.....

```

Fig.4 Output from the Augmented Composer

card is seen by the camera or not (0 or 127), “velocity” means the loudness of the note (MIDI velocity number), and “duration” means the length of the note’s playing time (0–127).

Since the data is represented electronically as symbols, it is suitable for computers. Thus, MIDI data can be convenient for data analysis as well as to play music on various MIDI-supporting players.

3.2 Analysis of result from the Augmented Composer

As shown above, since the output data from the **Augmented Composer** is MIDI data, the data from the system can be dealt with by a certain data mining method with a little modification. Though MIDI data is quite low level, first, we analyzed the data for a comparison between compositions by experts and non-experts. We adopted two popular analysis systems. One is C4.5 [13], which can analyze continuous or discrete data to obtain logical relationships. Another is KeyGraph [7], which analyzes discrete data to obtain relationships between elements in a graphical manner. Part of the result from KeyGraph can be regarded as association rules.

The result by C4.5 is shown in Fig. 5. “E” means expert and “NE” means non-expert.

```

Pitch ≤ 51 :
| duration ≤ 1 : E
| duration > 1 :
| | velocity ≤ 75 :
| | | duration ≤ 24 : NE
| | | duration > 24 : E
| | velocity > 75 :
| | | ON/OFF = 0: NE
| | | ON/OFF = 127:
| | | | velocity ≤ 97 : E
| | | | velocity > 97 : NE
Pitch > 51 :
| Pitch ≤ 66 :
| | velocity ≤ 8 :
| | | duration ≤ 23 :
| | | | Card ≤ 7 : NE ..... (omitted)

```

Fig.5 Generated decision tree by C4.5 (part of)

Each cluster of the tree in Fig. 5 can be interpreted as follows, for example:

$$E : \text{pitch}(\leq 51), \text{dura}(\leq 1). \quad (1)$$

creativity and education, we think there is a certain relationship. For example, for education, we need a certain reference or teacher. For students, it is very important to be shown a proper reference in a proper manner. For teachers, it is very important to show proper examples to students in a proper manner. Usually, this type of education is discussed and achieved only by an expert teacher. This type of education is exactly what active mining aims at.

First, it is important to select the proper data to be shown. Second, it is important to modify the data to a user-friendly style or according to the users' capabilities. Third, the users can refer to and compare with a reference in real time to improve their compositions. These types of tasks can be discussed in the framework of active mining.

4.2 Music composition support as active mining

We define one of the music composition supports as showing a proper reference. A reference will be a rule for good music. Usually, composers learn music composition rules in schools or from text books. In addition, they study the style of great composers to obtain composition rules. Usually, this type of education is provided by a human teacher. In this paper, we propose a computational music teacher who analyzes good music styles and rules to provide a guidance for a better music composition.

In Section 3, we pointed out that composition style can be obtained by inductive methods such as C4.5. As shown in Fig. 5, if we collect many patterns (composition data) by experts and non-experts, we can generalize the difference of composition rules between them. As discussed in Section 3, we think these results can be used as powerful references or guidance for all users. For instance, Fig. 8 shows a generated decision tree where the data of non-experts is classified as data by NE1. That is, there are three classes, i.e. E, NE, and NE1 (individual).

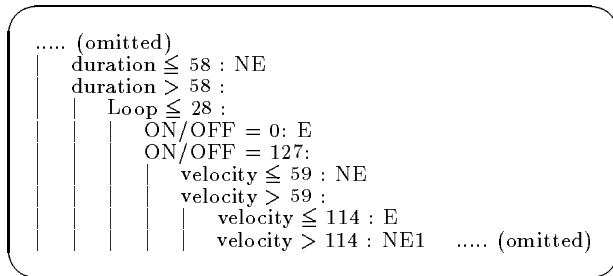


Fig.8 Generated decision tree by C4.5 (part of)

From Fig. 8, we can observe a certain difference between E and NE1, as well as a difference between NE and NE1. The latter classifications reflect the individual feature as a non-expert. A more important factor can be read from the former classifications that show a difference between the music composition style of experts and NE1 (individual). Thus, we think that if the result is properly displayed, though it depends upon the user's capability, the user can be aware of a difference between his/her composition style and the experts' composition style. If the users are aware of the difference, it will help them to improve their work. Of course, as shown in Fig. 6, since KeyGraph shows the relationship among each element in a graphical manner, we can use KeyGraph's result by analyzing the experts' data as a visual reference.

We propose a method dealing with a user-friendly reference or assistance for better composition under the concept of active mining. Under this context, "user-friendly" means that the users can easily ascertain a difference be-

tween works by experts and their own to improve their work.

4.3 Music composition support system

After taking the above analysis into consideration, we decided to build music composition supporting system with a user-friendly interface. In fact, there are many "user-friendly" applications for musicians. However, our conclusion is to provide an interface that guides users to be aware of a difference between work by experts and their own to improve their work.

We propose the system shown in Fig. 9. This figure shows a cyclic music composition by referring to an automatically generated reference. That is, a multiple sequence of composition, comparison, and improvement is a composition cycle for non-experts. This circulation will work as an automatic education for non-experts.



Fig.9 Music composition cycle

After composing music, the user can analyze his/her composition to compare with the experts' composition style. Then he/she can be made aware of the difference and improve his/her work according to the recognized difference. If the difference is displayed clearly and properly, the users can ascertain what need to be changed. The cycle will continue until the difference becomes considerably small. In this sense, the users can learn how to compose good music by themselves. If the users use the system in another way, they might be able to imitate or pretend other composers' styles.

From the viewpoint of active mining, this type of application is quite important. This application includes some problems regarding how to display the results in a user-friendly manner. Since this is a preliminary proposal of music composition support from the viewpoint of active mining, we currently have quite a few ideas regarding the above problems.

In the following section, we will point out problems from the viewpoint of art as well as active mining.

4.4 Problems in the music composition support system

First of all, since the obtained information is currently only MIDI output that is quite low level, any analysis we can do is quite limited. Nevertheless, we can derive certain tendencies from the data. The results are shown in Section 3. From the viewpoint of art, the data including pitch, duration, etc., is quite rough. However, from the viewpoint of computational data analysis, a small number of data seems to be suitable for focusing on significant data.

Besides the data structure problems, the following problems need to be solved.

1. How will the difference be shown?

Currently, we adopt C4.5 to generate a type of logical representation. This is because a logical representation is suitable for computational data and a tree-style representation is understandable as a type of visualization. However, whether a logical rule is suitable for representing music will be a problem. A logical presentation is suitable for computer and logical thinking. In fact, it seems suitable for analyzing music in a structural way. However, if we want to analyze music in an affective manner, is it still suitable to use logical representation? If not, we should find an alternative method to deal with affective representation.

2. Is the classification enough?

Currently, we divide samples into two categories, i.e., data by experts and data by non-experts. For an experimental brief check, this classification is enough. However, if we use this system in a real situation, more classification should be required. For instance, categorization such as beginner, teacher, professional composer, etc., might be required to give more detailed information.

3. How to close up necessary information?

Currently, we only show the result by C4.5, which is not user-friendly for a person who has never experienced C4.5. Of course, this problem will occur if we use another data mining tool. We must find an automatic close up or selection method.

4. Which part of music mainly influences music?

After reviewing all the data and music or interviewing professional composers, we will be able to focus on an exact data that influences musical appearance. This is so-called heuristic problem solving. However, there may be a computational method to detect the exact element that mainly influences music.

The problems shown above are key factors for making our system attractive.

5. Conclusions

In this paper, we briefly introduced an interactive music composition system, the **Augmented Composer**. We then showed the analyzed result of MIDI data from the **Augmented Composer**. Though MIDI data is quite low level, we recognized that the analyzed data reflects techniques and styles of music composition. By reviewing the analyzed data, we saw a means of using the analyzed data for reference to music composition. This is because we could ascertain an explicit difference between the composition styles of experts and those of non-experts. Thus, we proposed a cyclic composition support system that shows the experts' composition pattern as a reference to educate non-experts in a cyclic manner. Part of the concept of this system is based on active mining, where we introduce user-friendly reference or assistance for better composition. "User-friendly" means that the users can be easily made aware of the difference between experts and themselves. Of course, a reference should be generated automatically.

Since this is a preliminary proposal of music composition support from the viewpoint of active mining, as shown in Section 4, we still face many problems and will likely uncover many problems during experiments. For instance, currently, we adopted a common induction or data mining system for a test. During the experiments, if we find problems in induction or data mining, we should improve or develop new induction or data mining methods.

In this paper, we did not address chance discovery [8]. However, since composition is a creative task, we should take into consideration an aspect of chance discovery. Our system will then become a better composition system. In fact, there have been several studies on creativity from

the viewpoint of chance discovery. Abe pointed out that when we use the internet as reference, a slightly different search result simulates our hidden creativity [1]. Shibata and Hori showed that related problems and ideas stimulate the user's creative thinking [14]. Though they deal with creativity in research, for artistic tasks, there should be a similar framework or mechanism to compose excellent music. Our future goal includes a method to stimulate the users' creativity by concepts such as chance discovery.

Acknowledgments

This research is supported by TAO (Telecommunications Advancement Organization of Japan).

References

- [1] Abe A.: User's Interests Change as Chance Discovery, *Proc. of KES2002*, pp. 1291-1295 (2002)
- [2] Active Mining Home Page: (in Japanese) <http://www.ar.sanken.osaka-u.ac.jp/activemining/>
- [3] Berry R.: Augmented Reality for Music, *Proc. of the International Computer Music Conference*, pp. 100-104 (2002)
- [4] Hiller L. and Issacson L.: ILLIAC Suite for String Quartet, *New Music Edition* (1957)
- [5] Kato H. and Billingham M.: Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System, *Proc. of 2nd Int. Workshop on Augmented Reality*, pp.85-94 (1999)
- [6] Murakami M.: Quantitative Analyses of Sentences - A New Methods for Document Analysis -, *JSAI, Vol. 17, No. 3*, pp. 331-334 (2002) in Japanese
- [7] Ohsawa Y., Benson N. E. and Yachida M.: KeyGraph: Automatic Indexing by Co-occurrence Graph based on Building Construction Metaphor, *Proc. Advanced Digital Library Conference (IEEE ADL'98)*, pp. 12-18 (1998)
- [8] Ohsawa Y.: Chance Discovery for Making Decision in Complex Real World, *New Generation Computing, Vol. 20, No. 2*, pp. 143-163 (2002)
- [9] Pachet F.: Interacting with a musical learning system: the continuator, In *Music and Artificial Intelligence (Anagnostopoulou C., Ferrand M., and Smail A. eds.)*, (LNAI 2445), pp. 119-132 (2002)
- [10] Pope S. T., Roy P., and Orio N.: Content Analysis and Queries in a Sound and Music Database, *Proc. of ICMC99* (1999)
- [11] Poupyrev I., Berry R., et al.: Augmented Groove: Collaborative Jamming in Augmented Reality, *Proc. of the SIGGRAPH'2000 Conference Abstracts and Applications*, pp. 77 (2000)
- [12] Puckette M: Pure Data, *Proc. of International Computer Music Conference*, pp. 269-272 (1996)
- [13] Quinlan J. R.: *C4.5: Programs for Machine Learning*, Morgan Kaufman (1993)
- [14] Shibata H. and Hori K.: An Approach to Support Long-Term Creative Thinking and Its Feasibility, *Post-Proc. of Joint JSAI Workshop (LNAI 2253)*, pp. 455-461 (2001)
- [15] Takahashi Y.: Xenakis in Kyoto, [http://www.ne.jp/asahi/kerbau/kerbau21/texts\(e\)/xenky.html](http://www.ne.jp/asahi/kerbau/kerbau21/texts(e)/xenky.html) (1997)
- [16] Xenakis I.: Formalized Music: Thought and Mathematics in Composition, *Pendragon Press* (1991) (original: Musiques Formelles, La Revue Musicales editions Richard-Masse (1963))