Effects of a Synchronized Scoring Interface on Musical Quality

Yuji Takai, Masao Ohira, and Ken-ichi Matsumoto

Graduate School of Information Science, Nara Institute of Science and Technology
8916-5, Takayama, Ikoma, Nara, Japan
+81-743-72-5312
{yuji-t,masao,matumoto}@is.naist.jp

Abstract. Collaborative music composition among casual users has the potential of creating advanced music that cannot be composed by a single user since the users can complement a shortage of musical knowledge each other. Although some studies have proposed music composition interfaces which synchronously visualize the composition data in real time, their effects on the quality of music are not still clear. As a result of our experiment, we found that the EMD value (0 is the best score) in using the interface *Marble* we proposed was lower than that of asynchronous music composition systems and *Marble* has increased the total number of utterances among users.

Keywords: Synchronized Scoring Interface, EMD, music quality, Computer-Supported Cooperative Work.

1 Introduction

In recent years, computer music software such as Apple's GarageBand has been widely used from novice users to professional musicians, to efficiently compose music and/or effectively perform music since it can complement with the lack of real musical instruments with computational ones. One of the advantages to compose with computer music software is that it can give an ability of collaborative music composition to casual users who have knowledge of a limited number of musical instruments. Such the users can compose high-level music by combine each other's musical knowledge.

In the past study, we have developed a system called *Marble* which supports collaborative music composition by the casual users [1]. *Marble* has a synchronized scoring interface to help users collaboratively compose music online. We have confirmed that the system could make collaborative music composition among users easier and faster than the traditional environment[1] where users have to share score data (e.g., MIDI), communicate with each other about musical ideas, request each other's modified/edited score data over again and again, via e-mail or IRC (Internet Relay Chat).

[1] We use a composition system with an asynchronous interface.

A.A. Ozok and P. Zaphiris (Eds.): Online Communities, HCII 2011, LNCS 6778, pp. 363–372, 2011.

The easy and fast way of music composition is important in collaborative music composition among casual users. However, we believe the more important thing is the quality of composed music: "how are users satisfied with the quality of music composed through the collaborative music composition with *Marble* (synchronized scoring interface)?"

This paper forces on the above question and discusses the effects of synchronized scoring interface on music quality and collaborative composition process. We quantitatively evaluate the quality of music composed through the collaborative music composition with *Marble*. We also conduct qualitative evaluation on user's satisfaction with composed music.

2 Online Collaborative Music Composition

2.1 Types

Table1 shows that the online collaborative music composition can be classified by members and their DTM composition skills. In Table 1, each type of the online collaborative music composition is briefly describes. Due to the space limitation, we cannot discuss each type in detail, but our target of this study is on the online collaboration among acquaintance, casual users.

Table 1. Classification of Online Collaborative Music Composition

Member	DTM Composition Skill	Collaborative Method and Feature of Composed Music
Acquaintance	Casual	A melody composer mainly composes based on melody by him/herself as the people in charge of each musical instrument part append the musical note in the accompaniment part. Variety of music is composed to the course of member; however, maturity is not always high.
	Advance	In spite of take professional advice from friends and modify, individual user basically composes all the musical instrument parts. High-level music is composed including the genre (e.g., Orchestras) is needed professional skill.
Stranger	Casual	There are various forms; Decided a theme and a role (composition or writing the lyrics of a song), arranged to composition individually composed by high-level users.
	Advance	Multimedia composition is derivatively composed by high-level composer through mutual citation. There are high quality compositions with CG.

2.2 Process

In general, collaborative music composition among casual users proceeds as follows:

Step 1. A melody composer composes a melody part of music based on her knowledge of a particular music instrument (e.g., piano). She cannot compose accompaniment parts by herself because of her limited knowledge on music instruments even if she has her images of the accompaniments.
Step 2. She asks other users who are knowledgeable about other instruments to make the accompaniment parts (e.g., bass guitar and drum).
Step 3. After taking the accompaniments of music from the other users, She tries to fine-tune to bring the accompaniments to her images.
Step 4. If there is a big gap between the composed accompaniments and her images, she asks the others to modify them again.

Her images is often unclear and vague because her limited knowledge (e.g., how to render) about accompaniments [2]. So, the other users anticipate her images based on a composed melody part and conversations with her. At the same time, they have to compensate the lack of information to compose accompaniment parts, based on their knowledge and experience. In general, there is a big gap of knowledge between users because expertise in music instruments is different from each other. Therefore, it is too difficult for a melody composer to exactly convey her images to others.

All the members involved in the collaborative music composition need to spend much time and efforts to complete satisfying compositions because they must repeatedly send and take the music data to construct mutual understandings among them.

3 Marble

Marble is a system which can modify synchronized musical data mutually among members. *Marble* consists of two parts; the user interface to compose music among multiple user simultaneously, and the data control part to synchronize musical data as users can modify music in real-time.

Figure 1 shows the interface of *Marble*. Track provides a function to alter tune and volume of musical notes. Score Edit Space is to modify musical notes. Musical score editor adopted the Piano roll interface which is widely used for DTM software. It can be used to externalize user's musical images and to share them among members. Users can easily find who is composing which parts because users can compose music in real-time while watching others' composition process. Moreover, users can intervene in other users' composition process on musical score editor.

3.1 Proceed of Marble

The collaborative music composition with *Marble* proceeds as follows;

Step 1. Users respectively compose music parts based on their musical images on each musical editor.
Step 2. Users compare composed music with her images while viewing each composition process listening to music as needed.
Step 3. Users intervene in other user process of composition to directly adjust others' music parts to get close to their images

Users repeat Step1~3 until they agree with the quality of composed music.

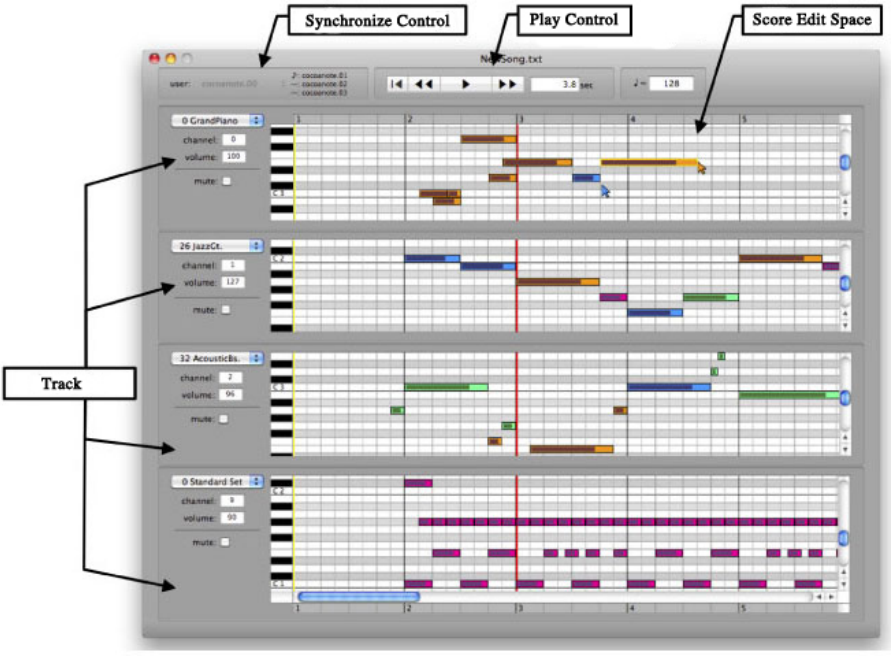

Fig. 1. Collaborative Music Composition with *Marble* interface

3.2 Problem of Marble about Evaluation

In the previous study, we have focused on the time to be required to compose music with *Marble*. From the experiment, we have found that *Marble* could help degrease the time to compose music in a process of the online collaborative music composition.

We believe the more important thing is whether users can satisfied with the quality of music composition through the collaborative composition. In this paper, we quantitatively evaluate the quality of music composed with EMD (Earth Mover's Distance).

4 EMD (Earth Mover's Distance)

4.1 Outline of EMD

EMD is a method to evaluate the similarity between two signatures. A signature is a set of features and their corresponding weights. Computing the EMD is based on a solution to the transportation problem for which efficient algorithms are available [3]. Given two signatures $P = \{(p_1, w_{p1}), \ldots, (p_m, w_{pm})\}$ and $Q = \{(q_1, w_{q1}), \ldots, (q_m, w_{qm})\}$, where p_i and q_j are the points in some Euclidean space, the musical feature in our study, and

w_{pi}, w_{qj} are the weights of the points. We want to find a flow C, with C_{ij} the flow between p_i and q_j, that minimizes the overall cost

$$\sum_{i=1}^{m}\sum_{j=1}^{n} C_{ij}\,|p_i - q_j|$$

subject to the following constraints;

$$C_{ij} \geq 0, \qquad 1 \leq i \leq m, \qquad 1 \leq j \leq n \tag{1}$$

$$\sum_{i=1}^{m} C_{ij} \geq w_{pi}, \qquad 1 \leq i \leq m \tag{2}$$

$$\sum_{j=1}^{n} C_{ij} \geq w_{qj}, \qquad 1 \leq j \leq n \tag{3}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n} C_{ij} = \min\,(w_{p,} w_q) \tag{4}$$

Constraint (1) allows shipping of supplies from a supplier to a consumer and not vice versa. Constraint (2) forces the consumers to fill up all of their capacities and constraint (3) limits the supply that a supplier can send as a total amount. Constraint (4) force to move the maximum amount of the flow possible. The EMD can be defined as:

$$\mathrm{EMD}(p, q) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} C_{ij}\left|p_i - q_j\right|}{\sum_{i=1}^{m}\sum_{j=1}^{n} C_{ij}}$$

$$= \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} C_{ij}\left|p_i - q_j\right|}{\min\left(w_p, w_q\right)}$$

Then, we will illustrate the application of EMD to this study in the following.

4.2 Application of This Study

In our study, the features are the musical features (Note Position, Pitch, and Transition of Pitch) and the weights are length of a musical note. “Note Position” signifies where the musical note from start point is, a beat is a musical note having the time value of an eighth of a whole note. "Pitch" signifies the difference of pitch from front note. "Transition of Pitch" signifies the transition of pitch from the pitch of note on start place. We define that the chromatic scale is 1 on pitch and transition of pitch.

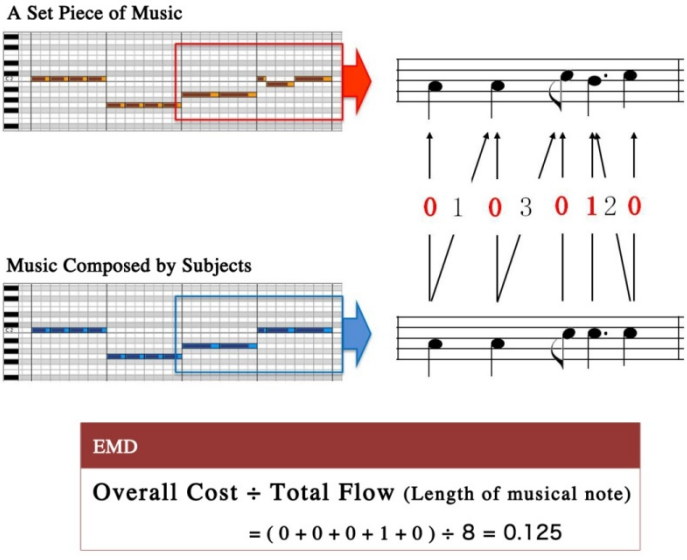

Fig. 2. Example of the calculation of EMD (Simplified values)

5 Experimentation

We conducted an experiment for comparing the quality of music composed using *Marble* and existing computer music software without a synchronized scoring interface.

The object of the experiment is to evaluate the complete of music composed with a synchronized scoring interface and find the factor from process of composition. Then, we examined whether it is different of complete music composed by the environmental condition about synchronization with a set piece of music we prepared. Further, we analyzed the activation of argument from utterance data in experiment.

Materials. The experimental materials are to compose a accompaniment part which a melody has composed, by a couple of users. The set piece of music which we prepared consists of two parts (a melody and an accompaniment) and 4～5 bar The music is original to affect the materials. We attempted to compose the set piece of music without the difficulty of its, we consider to user the same number in each condition with environment not to depend the result of experiment by the difficulty.

Table 2. Definition of Subjects

DTM Skill	Feature
Novice User	User hasn't experience of musical performance and using DTM. User can know the usage of the piano roll, but cannot compose by herself because user does not have musical knowledge.
Medium User	User has experience of musical performance or using DTM. But user does not have the skill to compose the high-level music because he does not only have a specific musical skill.

Environment. We compared *Marble* (Synchronous System) with a traditional system (Asynchronous System). We utilized Score Editor of *Marble* as a traditional system to remove other factors as much as possible. *Marble* has the synchronized function, but we switched off the function for the traditional system. Subjects used the Skype video conference to communicate with others in remote sites.

Subjects. 10 couples of 20 subjects in total joined the experiment. Each couple of subjects well knows each other and consists of a novice and a medium-skill user of DTM in accordance with the definition of subjects (see Table 2). A novice subject played the role of a melody composer and the medium subject' role was an accompaniment composer.

Procedure. After we fully explained each role, the tasks, and how to operate *Marble*, the subjects performed a trial task to accustom to the operation of *Marble*. Then the subjects started to experimental tasks as follows;

Step 1. A melody composer memorizes an accompaniment image after he listens to music which is given by experimenters.

Step 2. Both subjects take music data only consisting of a main melody.

Step 3. Both subjects communicate each other using the Skype video conference.

Step 4. An accompaniment composer listens to a melody part, start to experiment.

Step 5. The melody composer tells her accompaniment image to the accompaniment composer. The accompaniment composer composes and modifies the musical note.

Step 6. One task finished when subjects reported that composition was over, or it took the maximum 15 minutes from the beginning of the task.

6 Results

We evaluated the quality of music composed using *Marble* and the traditional system and analyzed the number of subjects' utterances through the collaborative music composition.

6.1 Music Quality

Table 3 shows the results of the evaluation based on the EMD value (0 is the best score). It shows that the average and median of EMD values (means music quality) are lower in *Marble* than the traditional system. Table 3 also shows that the standard deviation of the EMD value of traditional system is larger than *Marble*. There was also a significant difference of EMD values between *Marble* and the traditional system ($p < 0.05$). It indicates that *Marble* can successfully help users compose music as users can get close to their musical images, compared to the traditional music composition system.

Table 3. Difference of EMD values between Marble and the traditional system

	Marble	Traditional System
Average	1.40	2.06
Median	1.16	2.22
Standard Deviation	1.45	2.28

6.2 Utterances

We found that the music quality is higher in *Marble*. However, it is necessary to clarify what kind of situation in the collaborative composition process has an effect on the EMD value. So, we analyzed discussions among subjects in the collaborative composition from using video data.

We counted the number of utterances when speakers changed or remained silent more than constant distance (around 3 seconds). In the following, we describe the three sorts of utterances: (1) All utterance, (2) Convey message, (3) Proposal message.

(1) "All utterance" is the utterance of all that a subject utters including a beck and an instructions word, regardless of the length of an utterance.
(2) "Convey message" is the utterance conveying the musical image from a melody composer to other users; humming and "Higher." "I think it is correct." etc.
(3) "Proposal message" is the utterance proposing to confirm the musical images and urge a main composer to modify from a melody composer to other users; "is it OK?" and "Is it correct although this musical note is disharmony?" etc.

Table 4. The number of utterance in collaborative composition of music with each system

Number	*Marble*	Traditional System	p-value
all utterance	71.65	63.15	0.05
Convey message	31.9	26.6	0.04
Proposal message	14.85	7.65	0.001
Number per minute	***Marble***	**Traditional System**	**p-value**
all utterance	5.81	4.48	0.001
Convey message	2.63	1.89	0.001
Proposal message	1.23	0.54	0.001

The result of table 4 is separated into two parts: the upper part is representing the average number of utterance in all time of experiment and the lower part is representing the average number of utterance per minute because collaborative composition time is different. Also, we eliminated the time to send and take the music data with Asynchronous system. It was found that *Marble* has made a significant difference (5%) for the number of all sort of utterance. Also

We found that *Marble* has increased the total number of utterances among users. In particular, the number of utterances to suggest music ideas to partners has been

doubled. Therefore, it seems that the subjects performed a constructive discussion in collaborative composition with *Marble*.

6.3 Discussion

In the collaborative composition with *Marble*, the degree of difficulty of tasks less influenced the quality of composed music. In general, casual users can convey the rhythm of music to others correctly, but they often face the difficulty in discussing the pitch of music. In many cases of our experiment, subjects tried to complete music while they discuss the pitch again and again. Compared with the traditional system, *Marble* seemed to ease the difficulty because the subjects with *Marble* could see other subject's operations on musical notes and then understand the intention of the operations. In this paper, we used relatively short music as tasks. In case of composing more real, longer music, the differences of interfaces for music composition would have a great impact on the collaboration among users, resulting in the music quality.

We also observed that each subject performed a task alone immediately after beginning the task. Especially in case of the traditional system, the subjects tend to individually compose each music part without communicating with other subjects since it was difficult to know what others were doing. In fact, the number of the "proposal utterance" was larger in using *Marble*. The awareness of other's operations would trigger to discuss the operations and then increase the number of proposals for changing positions of notes on other's musical notes. As a result, the subjects with *Marble* seemed to be able to compose the higher quality music.

7 Related Work

In related work of asynchronous system, Latta suggested Netjam[4], which one of first system supported collaborative composition through internet, supported to improve the give and take of MIDI file with e-mail and Jorda suggested FMOL[5], which the system supported the public users join collective composition on the internet by real-time based on e-mail. In related work of synchronous system, there is hardly a case study, as for many research only proposed the new interface[6][7]. These researches are differently circumstanced to this study, which we evaluate the musical quality and analyze the collaborative composition process, while the system to stimulate the musical idea and facilitate on the process of collaborative composition.

8 Conclusion

In this paper, we analyzed the effects of a synchronized music scoring interface of on the music quality. We conducted an experiment to measure the quality of music composed using a synchronized interface (*Marble*) and asynchronous interface (traditional system). As a result of the experiment, we found that the users with *Marble* can correctly convey musical images to each other and then can compose

better quality music, compared with a traditional system. We also found that the number of the "proposal utterances" in using *Marble* was twice larger than the traditional system, resulting in creating higher quality music. In the future, we would like to develop a system for supporting the collaborative music composition among strangers.

References

1. Ohira, M., Kimura, M., Mastumoto, K.-i.: Marble: A System for Supporting Online Collaborative Music Composition among Casual Users. Human Interface Journal 12(3), 219–229 (2010) (in Japanese)
2. Sloboda, J.A.: The Musical Mind: The Cognitive Psychology of Music. Oxford University Press, Oxford (1985)
3. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover's distance, multidimensional scaling, and colorbased image retrieval. In: Proc. of the ARPA Image Understanding Workshop, pp. 661–668 (1999)
4. Latta, C.: A New Musical Medium: NetJam. Computer Music Journal 15 (1991)
5. Jorda, S.: Faust Music On Line (FMOL): An approach to Real-time Collective Composition on the Internet. Leonardo Music Journal 9 (1999)
6. Barbosa, A., Kaltenbrunner, M.: Public sound objects: A shared musical space on the web. In: Proc. Intl. Conf. on Web Delivering of Music (2002)
7. SoundWIRE., `http://ccrma.stanford.edu/groups/soundwire/`