Co-Musicator : A Web-Based System for Musicians to Support Recording, Playing, Indexing and Discussing

Takehiko Sakamoto, Shin Takahashi, and Jiro Tanaka

University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan
sakamotot@iplab.cs.tsukuba.ac.jp,
{shin,jiro}@cs.tsukuba.ac.jp

Abstract. Musicians discuss how to represent the music in their performance. In previous work, we have developed a web-based system that enables online discussion by commenting on music score. This paper introduces a new extended system – Co-Musicator that supports musicians to record their performance and make good use of it in their discussion. The system is being implemented as a web application using recent HTML5 technologies. We state various future prospect of the system for better usefullness.

1 Introduction

Most music ensembles, groups of musicians, discuss musical expression to make their performance more impressive. They consider every detail of the piece of music they perform, i.e. the length, the dynamics, the articulation such as accent or staccato, and so on. They also discuss together what they feel and imagine, and how the music is written by the composer. When they meet together, however, they prefer taking time to practice playing together rather than discussing. Therefore, there is a great demand for a system that enables the discussion about musical expressions asynchronously and remotely.

As a previous work, we proposed a method for discussing musical expression on the web-based system[1]. The system enables distant and asynchronous discussion especially for amateur ensembles that cannot spent enough time for practicing. In the system, users can discuss their interpretation by commenting, drawing symbols, linking videos, and depicting "articulation shapes" *on music scores.*

However, the previous system did not handle recording of their practices. There is a need to record ensemble practices and use the recorded sound data in discussions. By using the recorded data in discussions, the users can discuss and communicate more smoothly and deeply after the practice. They are also helpful for absentees of practice to learn the details. In addition, the group members may perceive their problems by listening their practice recordings.

This paper presents a new extended discussion system that also incorporates the logging of ensemble practices including the recording sound data of practices. The challenges in designing and implementing the system are as follows:

Decreasing the effort to log practices. When players practicing together, they focus on it. Extra effort to logging and recording should be avoided as much as possible. In our system, the players only have to touch the current location in the music score displayed on the tablet PC before playing a instrument.

Visualizing logged practice data and linking to music scores Logs of practices should be easily used in the discussion system. Currently, we assume a log is an entire sound recording of each performance practice. In discussions, the user must be able to refer to the specific part of it and comment on it. In our system, the recorded sound data is linked semi-automatically to a music score utilizing the user's touch data during practices. Visualization of each sound data is also shown to users so that they can understand which part corresponds to which part in the music score.

Combining logging practices and asynchronous discussions enables more powerful support for musicians. We designed and prototyped a web-based system that enables ensemble players to record practices easily and use them naturally.

2 System Design

2.1 Overview

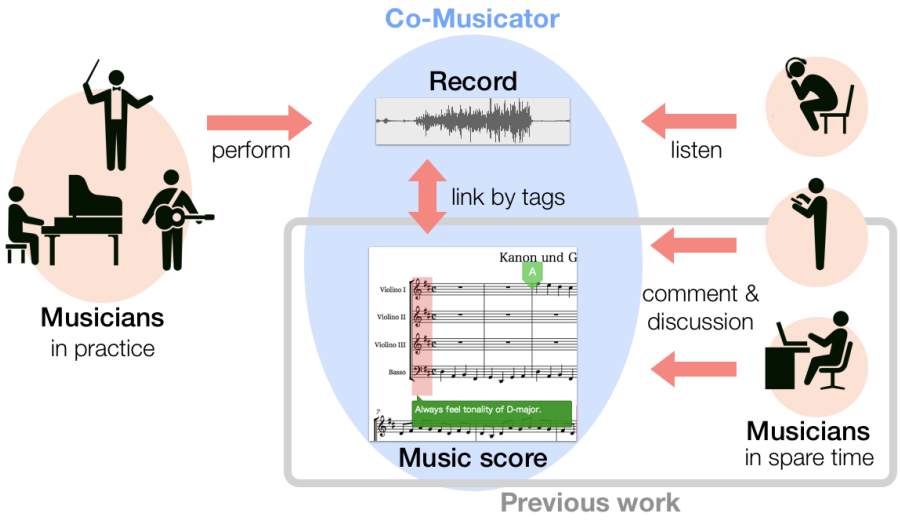

Fig. 1. System overview

Co-Musicator is a web-based system for musicians to record their performance, share the sound data with others, and discuss music expressions. Fig. 1

shows the overview of our system. We basically assume two different scenes — co-located practice and spare time. In co-located situations, musicians perform practice and record its sound into the system. On the other hand, after the practice, musicians in their spare time discuss the performance. The music score and the recordings are linked to each other with attached tags.

Because Co-Musicator is a web-based system, various kind of devices with a browser can be used for running Co-Musicator. However, a relatively large display tablet PC is suited for co-located practice. It can display a music score for performing as usual. In addition, the touch display is especially useful for pointing the current location of performance during practices. Thus performing, recording and discussing will be more seamless.

2.2 Two views of Co-Musicator

Fig.2 shows two views of Co-Musicator: the score view (left) and the recorded data view (right). The score view shows music scores, and various information for performance and discussion. In this view, every information is shown on the music score. Users can put annotations such as texts, musical symbols, articulation shapes, and references to recorded data. This view is basically similar to our previous system, therefore, refer to our previous paper[1] for more details.

On the other hand, the recorded data view shows a list of practice logs where each of them corresponds to one meeting of practice. In this view, a log (a sound data) is visualized as a sound wave. Users can select a point in it to playback the sound data from the specified position.

The users can record practices from both of views. The microphone icon at the top-right corner is the button to start recording. During the recoding, the realtime sound input is displayed at the top.

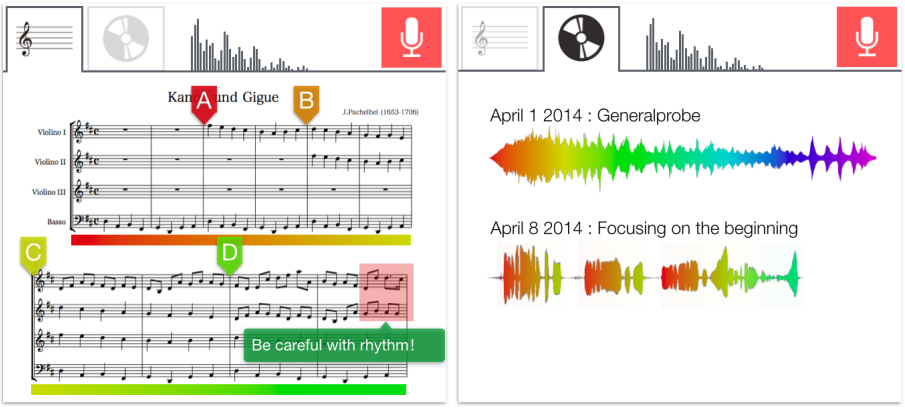

Fig. 2. Two views of Co-Musicator

2.3 Rehearsal Letters and Tags

In general, it is time-consuming to find the point that you want to listen from a lot of long sound recordings. In our system, the users can easily get there by using tags that links sound data to music scores.

Most music scores have "rehearsal letters" and "bar numbers" (Fig. 3) for helping musicians find quickly a certain point on the score in performance practice. Musicians are accustomed to these letters and numbers, so this system makes use of them. Tags in our system are automatically created by the system by using music structure data, but the users can also create and edit them.

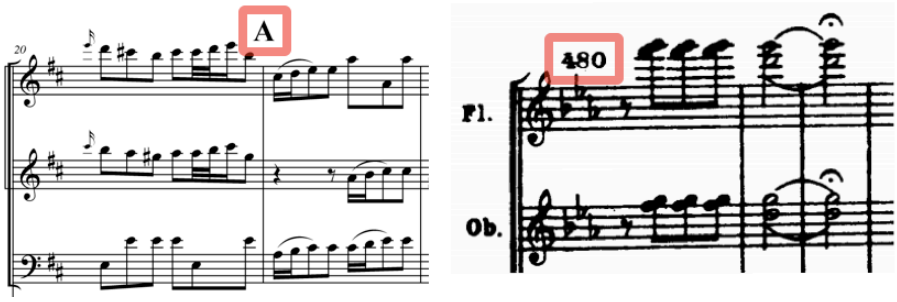

Fig. 3. Reference image of a rehearsal letter and bar number in existing music score (encircled by red rectangle).

2.4 Recording and tagging sound data

Our system provides easy way to record musical performance. The users only have to access the system with a browser and click/tap the 'Rec' button. The system shows frequency spectrum of surrounding sound at the top of the view, so the users easily recognize the state of the recording. Another click/tap of the 'Rec' button end the recording. The recorded data is shown in the recorded data view immediately.

Currently, we assume that users record one meeting of practice entirely. Therefore, the users basically only have to click/tap the recording button at the beginning and the end of the practice. However, to estimate the correspondence between each time slot in the recorded sound data and the position in the music score, users are encouraged to add tags to sound data during practices. That is, if the user clicks/taps a tag on the music score just before start playing that part, the tag is also attached to the time position in the recorded sound data. By using those tagging information, the system can estimate the correspondence between the sound data and the music score.

To make it easy for users to understand which position in recorded data is performing which position in music score, the system colors each log of practice

(sound wave) using tags information (Fig. 2). The reference color gradation bar is displayed under the music score. The beginning of the music is colored red. The color gradually changes to purple with getting close to the end of the music. Therefore, for example, if a part of sound wave is colored red, that part must be the beginning or around of the recorded data. The system estimates the color of intervals between tags by using the music structure and its tempo. Furthermore, if the volume of a recorded section is smaller than the threshold, it is regarded as not-playing section, and the section is colored gray.

3 Implementation

We are developing a prototype system using the web technologies such as HTML/CSS, JavaScript and PHP. Only a web browser is necessary for using our system, which is advantageous for multi-device and multi-platform for the future.

To record sound, we use Web RTC (Real Time Communication) API and Web Audio API technology. Web RTC API is a new API of HTML5, that can access camera and microphone stream data. Fig.4 shows the sound data flow in our system.

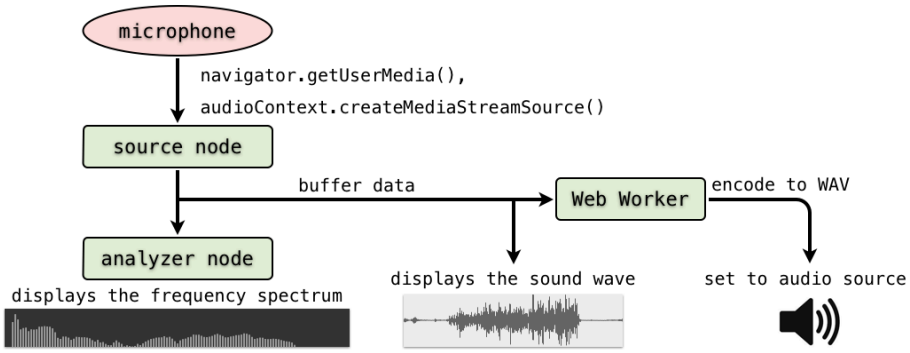

Fig. 4. Sound data flow. The sound data from microphone is sent to Web Audio instance. Web Worker encodes it to WAV in background.

While the system captures sound data, the data is sent to instances of AudioContext for display the frequency spectrum. After finishing recording, the data is transcoded to the WAV format, and sent to the source of the HTML audio element as an URL of a binary large object. At this time, Web Workers API is utilized to process the transcoding asynchronously. A sound wave is drawn by calculating maximum volume every unit time from the buffers data.

A tag is simply represented by a div element, and its innerHTML value and an input element is used to edit a tag name seamlessly. All tag name is stored to an array and referred to make a list of linked tags when a tag is clicked.

Commenting function is implemented in similar method with previous work[1]. All of the recorded data, tags, comments and comment regions are stored to the database immediately by using so-called Ajax. Music score is also stored to database in Base-64 format.

4 Related Work

Web-Based Collaboration A lot of web-based collaborative tools and systems have been created for many purpose. Heer et al. developed *sense.us* for collaborative information visualization with census data of the U.S. on web[2]. *Sense.us* provides various annotations on the visualized data for discussion and analyzing. Lang and Minker developed a collaborative web-based help-system[3]. It helps users, especially for novice, achieve goals with an avatar that provides information visually and aurally. Goldman et al. developed *Collabode* that is a web-based IDE for collaborative coding[4]. Collabode has efficiency for various type of coding and they stated the probability of HTML5. Marion and Jomier implemented real-time collaborative visualization with WebGL and WebSocket[5]. These two new technologies allowing faster and easier way to remotely collaborate with 3D datasets.

Our system enables the users to look at music score, record sound, play it and discuss in web browser. We also utilize new technologies : Web RTC and Web Audio. There is extensive collaboration of musician with this system.

Computer-Supported Ensemble As the study of computer-supported musical performance, Bellini et al. researched supporting of real-time music performance[6]. They developed *MOODS*: a cooperative editor for musical scores that can automatically synchronize written notes on their sheets of music. Sawchuk et al. developed *DIP* that enables players to participate their practice at a distance[7]. Akoumianakis et al. developed the prototype toolkit for the purpose of distant and asynchronous ensemble practicing[8]. The system *DIAMOUSES* records performance every part of the music and enable the users to enjoy distant ensemble by using the recorded data. Gurevich developed *JamSpace* : a real-time collaborative music environment that provides interaction for novice[9].

Our system does not support music performance directly but supports utilization of their performance. We focused on discussions that is a significant part of ensemble practice, and had an eye to indirect supporting musical performance.

5 Future Directions

Our system is still under development, and we have several future directions. First of all, we have to conduct user study of musicians for qualitative evaluation. We are planning usability testing and long-term evaluation in effective conditions.

It may be also possible to make better use of structure data of music scores. In other words, if the system understand every note, rest, and any other musical symbol, rehearsal letters and bar numbers will automatically be recognized as tags. Comments on music score will point out what they mention in great detail. This function may be implemented with the use of a music engraving language such as GNU Lilypond or MusicXML.

In addition, speech recognition function should be implemented in our system. Now our system is able to record all sound that contains performance, voice and noise but not able to distinguish them. In a practice of a music group, musicians of course perform and make sound but they also talk and discuss about their music. In this paper, we have focused on the performance side. If the system can recognize words from voices, there are various probabilities. For instance, the users can use the system without touching devices. The words 'rec', 'stop' or 'add tag A here' will be triggers of operation and their voices of a discussion will be comments immediately. Though it may require high-precision API or library of natural language processing.

If we can use structure data of music score and speech recognition, the system may automatically maps recording sound to specific part of music score. In that case, tagging is no longer required.

6 Conclusion

We have designed a web-based system that supports musicians to record their performance and make good use of it. In previous work, the users can comment on music score for their discussion. Our new system – Co-Musicator provides several functions for simple recording, linking it to music score and discussion using them. It is implemented only by web technology using recent HTML5 APIs.

References

1. Takehiko Sakamoto, Shin Takahashi and Jiro Tanaka. A Method for Discussing Musical Expression between Music Ensemble Players using A Web-Based System. Proceedings of the 15th International Conference on Human-Computer Interaction (HCI International2013), Lecture Notes in Computer Science Volume 8007, 2013, pp.730-739.
2. Jeffrey Heer, Fernanda B. Viegas, and Martin Wattenberg. 2007. Voyagers and voyeurs: supporting asynchronous collaborative information visualization. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07). pp.1029-1038.
3. Helmut Lang and Wolfgang Minker. 2012. A collaborative web-based help-system. In Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics (WIMS '12). Article 60, 5 pages.
4. Max Goldman, Greg Little, and Robert C. Miller. 2011. Collabode: collaborative coding in the browser. In Proceedings of the 4th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE '11). pp.65-68.

5. Charles Marion and Julien Jomier. 2012. Real-time collaborative scientific WebGL visualization with WebSocket. In Proceedings of the 17th International Conference on 3D Web Technology (Web3D '12). pp.47-50.
6. P. Bellini, P. Nesi, and M. B. Spinu. 2002. Cooperative visual manipulation of music notation. ACM Transactions on Computer-Human Interaction (TOCHI) Volume 9 Issue 3, September 2002, pp.194-237.
7. A. A. Sawchuk, E. Chew, R. Zimmermann, C. Papadopoulos, and C. Kyriakakis. 2003. From remote media immersion to Distributed Immersive Performance. In Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence (ETP '03). pp.110-120.
8. Demosthenes Akoumianakis, George Vellis, Ioannis Milolidakis, Dimitrios Kotsalis, and Chrisoula Alexandraki. 2008. Distributed collective practices in collaborative music performance. In Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts (DIMEA '08). pp.368-375.
9. Michael Gurevich. 2006. JamSpace: a networked real-time collaborative music environment. In CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06). pp.821-826.