A Musical Performance Evaluation System for Beginner Musician based on Real-time Score Following

Yoonchang Han, Sejun Kwon, Kibeom Lee, Kyogu Lee
Music and Audio Research Group (MARG)
Seoul National University
Seoul, Republic of Korea
{yoonchanghan, saejun1234, kiblee, kglee}@snu.ac.kr

ABSTRACT
This paper proposes a musical performance feedback system based on real-time audio-score alignment for musical instrument education of beginner musicians. In the proposed system, we do not make use of symbolic data such as MIDI, but acquire a real-time audio input from on-board microphone of smartphone. Then, the system finds onset and pitch of the note from the signal, to align this information with the ground truth musical score. Real-time alignment allows the system to evaluate whether the user played the correct note or not, regardless of its timing, which enables user to play at their own speed, as playing same tempo with original musical score is problematic for beginners. As an output of evaluation, the system notifies the user about which part they are currently performing, and which note were played incorrectly.

Keywords
Music performance analysis, Music education, Real-time score following

1. INTRODUCTION
Recently, smart portable devices such as smartphone and tablet PC became commodities, and it allows people to do various kinds of tasks anytime, anywhere. Smart devices do not have limitations comes from location and time, such that it can be used as an useful tool for practicing musical instrument which needs to be repeated frequently. When people practicing musical instruments, professional musicians normally focus on richer emotional expression of their performance. However, unlike experts, important part of beginner's practice is actually playing the note correctly. Because they were using the fingering what they thought correct, they sometimes believe that the wrong-sounding note must be correct. [1] Also, beginners tend to play slower than the original speed of music, as they are not proficient enough.

Music education system for beginners has to be different from for experts, such that our system focuses on providing necessary features to satisfy the needs of beginner musicians' practice. The rest of the paper is organized as follows: in the next section, we briefly present existing systems related to our work. In the following section, we explain out proposed system and implementation method in detail, followed by conclusion and directions for future work.

NIME'13, May 27-30, 2013, KAIST, Daejeon, Korea.

2. RELATED WORKS
There are several attempts to build an automatic system for musical instrument education. First of all, The Piano Tutor [2] is one of the first interactive piano teaching systems. It combines an expert system and multimedia technology to form an interactive piano teaching system. This system uses score following to turn pages automatically, also to synchronize an accompaniment to give ensemble experience to students. Also, it analyzes performance of student to give feedback about mistakes they made.

Secondly, Tonara [3] is an iPad application for music education. It makes use of score following to notify user which part they are currently playing, and it automatically turns page for performer. However, this system does not give any feedback to user. Another recently released mobile application for music education is SmartMusic [4]. This system gives feedback about wrongly played note, but does not make use of score following to show which part they are playing.

As shown above, previous works has made significant contribution for automatic music education system. However, some systems such as The Piano Tutor require their own equipment such as two monitors, computer, audio mixer and video discs. Our system does not require extra equipment, but exist as a software form, which can be easily downloaded on smart devices. In addition, major difference in our system from existing systems is real-time score following with feedback about note that are played wrong. Tonara includes score following without feedback, and SmartMusic returns feedback, but it is not in real time.

3. SYSTEM DEISGN
The system architecture of proposed system is shown in Figure 1. In order to align sound from musical instrument and ground truth from original musical score, onset and pitch of the note was extracted from input audio. From musical score side, we extracted note event regardless of its timing for same reason. Next, information from audio was aligned with the one from score in order to find current position on the score, also to analyze whether it was correct pitch or not. If the played note was correct, the system just progresses to the next note, whereas the input was wrong pitched note, it is rendered on the music score with red color to indicate user about which note is wrong, and what was actually played note.

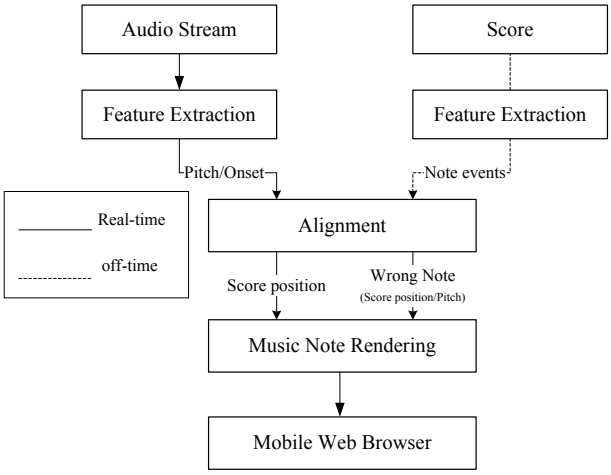

Figure 1. System architecture of proposed system

4. IMPLEMENTATION

This section is about how we implemented the proposed system. Our proposed system should work in mobile, also in real-time environment, such that the most of algorithm used in the system is light and computationally efficient, while still able to provide the information what we need.

4.1 Onset and Pitch Extraction

In order to find onset, we need to find abrupt changes in the audio. Usually, onset can be found by measuring amount of changes in frequency domain by applying Fourier Transform on the signal, or observing sudden changes in phase of the signal. Study from Dixon [5] shows that Spectral Flux (SF) performs better than other existing onset detection methods across various onset styles. Also, it requires fairly small computational power, which is a suitable characteristic for mobile application. Therefore, SF was used as an onset detection algorithm in the system.

Regarding pitch detection, as multi-pitch estimation technology is not mature at present, the system only supports monophonic sound. Hence, we used Harmonic Product Spectrum (HPS) for pitch detection, which is simple to implement with low-computational power. Note that the pitch was only calculated on the moment of onset, not for whole input audio, in order to avoid unnecessary computations.

4.2 Score Alignment

Dynamic Time Warping (DTW) [6] and Hidden Markov Model (HMM) [7] are widely used and proven algorithm for real-time score following. DTW is an algorithm that measures similarity between two sequences, which may differ in speed or time. This is exactly what we want to achieve, hence our system makes use of DTW as it does not require data training, and suitable to find wrongly played notes.

4.3 Music note rendering

An example of rendered music note is shown in Figure 2. The music notation is rendered using VexFlow with HTML5. The information from the backend, such as the incorrect notes and current position, are sent to VexFlow to be rendered accordingly. This HTML page is displayed on the iOS through the UIWebView class, which is used to embed web content in applications. Any additional buttons used for user interactions are coded natively through Xcode.

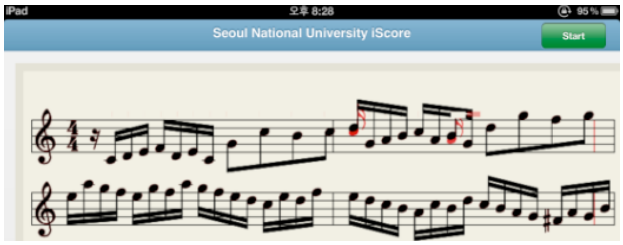

Figure 2. Screen shot of user interface. Red colored note indicates correct note for wrongly played note.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed music education system that is especially focused on beginners. The system obtain instrument sound as an audio input using built in microphone in smartphone, and finds onset and pitch of the input audio. Next, acquired information is aligned with original music score for score following purpose. Because the system follows the speed of playing speed of the user, it allows to user to play instrument at their own speed, which varies depends on their proficiency. As an output of the system, it returns which part is currently playing, and which note they played wrong. Also, it returns pitch of the wrongly played note on music score such that performer can easily recognize what mistake they made during the performance.

Currently the system only supports monophonic sound. However, many of musical instruments are actually polyphonic. Hence, in future, robust polyphonic pitch detection should be included in order to support instruments regardless of its type. In addition, if the system take other important features of music such as accent and dynamics into account, it would be gratefully helpful to users to develop emotional expression skills.

6. ACKNOWLEDGMENTS

This work was supported by the Korea Creative Contents Agency (KOCCA). Many thanks to the members of the Music and Audio Research Group (MARG), Seoul National University who helped further develop the ideas.

7. REFERENCES

[1] McPherson, Gary E., and James M. Renwick. "A longitudinal study of self-regulation in children's musical practice." *Music Education Research* 3.2 (2001): 169-186.

[2] Dannenberg, Roger B., et al. "Results from the piano tutor project." *Proceedings of the Fourth Biennial Arts and Technology Symposium*. 1993.

[3] http://tonara.com/

[4] http://www.smartmusic.com/SmartMusic/

[5] Dixon, S., "Onset detection revisited." Proceedings of the 9th International Conference on Digital Audio Effects. Vol. 120. 2006.

[6] Dixon, S., "Live tracking of musical performances using on-line time warping," in Proceedings of the 8th International Conference on Digital Audio Effects, 2005.

[7] Cont, A., "Realtime audio to score alignment for polyphonic music instruments using sparse non- negative constraints and hierarchical HMMs" in Proceedings of IEEE International Conferences on Acoustics, Speech, and Signal Processing (ICASSP 2006), May 2006.