

Development of a Score Click Playback System

Motoya Wakiyama, Megumi Wakao, Naoki Morita
 School of Information Telecommunication Engineering
 Tokai University
 Tokyo, Japan
 e-mail: {9bjt2116@cc, 9bjt2103@cc, wv062303@tsc}.
 u-tokai.ac.jp

Kazue Kawai
 Miyagi University
 Miyagi, Japan
 e-mail: kawaik@myu.ac.jp

Chiharu Nakanishi, Chiaki Sawada
 Faculty of Music Studies
 Kunitachi College of Music
 Tokyo, Japan
 e-mail: {nakanishi.chiharu, sawada.chiaki}@kunitachi.ac.jp

Kenta Morita
 Faculty of Medical Engineering
 Suzuka University of Medical Science
 Mie, Japan
 e-mail: morita@suzuka-u.ac.jp

Abstract— In a piano lesson, the instructor returns feedback to the student for each set of measures using a score. In this case, we would like to use a video to facilitate the process of instantly returning feedback to the student. Specifying the playback start position for each section using the seek bar in the video is difficult. In the present study, we propose a method by which to connect a score and a video by analyzing the score and video and comparing sound changes. The effectiveness of the proposed method was verified using scores practiced by beginning piano students.

Keywords- Score Analysis; Video Analysis; Piano Lesson; Without Seek Bar

I. INTRODUCTION

Sound is important in playing music, and knowing how to perceive and perform music is important. In piano playing in particular, tone changes depending on how players use their arms, legs, and body. Therefore, it is important to feedback how to use their arms, legs, and body on video. In such cases, the performance may be feedbacked by means of a video.

Prior research has studied various ways to support piano lessons [1]-[6]. For example, improvement of remote assistance using neural networks and multiple angles [7] and bad habits developed when using multiple cameras [8]. It is possible to add more information and look back at the way the score was played when the video was recorded. However, it is not possible to instantly project the points that the instructor wants to point out.

In a piano lesson, the instructor returns a feedback to the student using a score. In that case, when using a video for feedback, it is necessary to playback the video from the beginning of each musical section. However, it is difficult to specify the playback position for each passage using the seek bar control.

The purpose of the present study is to support piano lessons using video so that instructors can smoothly return comments to their students. Specifically, we propose a system that enables playback from the corresponding starting point by clicking on a measure in the score. Although some piano lessons involve the repetition of the same section of music, the

present study targets videos played through an entire piece of music.

The remainder of the present paper is organized as follows: Section II presents the development system. Section III describes the experiments conducted and presents the results and considerations, and Section IV provides our conclusions.

II. DEVELOPMENT SYSTEM

The developed system consists of a user interface (UI) module, a score analysis module, and a video analysis module.

In this system, the UI module is first used to upload the score and the video. When a music score is uploaded, the music analysis module extracts the musical scale from the score. When a video is uploaded, the video analysis module extracts the musical scale from the video, and then combines this scale with the musical scale extracted from the score. The video can then be played back from the corresponding time by clicking on a measure in the music score display screen of the UI module. The following sections describe the UI module, score analysis module, and video analysis module.

A. User Interface Module

The UI module provides a screen for uploading scores, a screen for uploading videos, and a screen for displaying scores and playing videos. Figure 1 shows the score display and video playback screens.

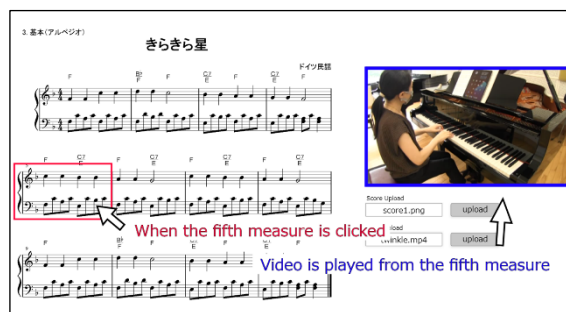


Figure 1. Displaying a score and playing a video.

This screen consists of the music notation screen on the left-hand side and the video screen on the right-hand side. In Figure 1, when the fifth measure is clicked, the video is played from the fifth measure.

B. Score Analysis Module

The score analysis module creates a scale list from an uploaded score, recording measure numbers and scales for each clef. This module identifies clefs, staff, bars, sharp, natural, flat, and note head symbols in the score. This module then finds the relative coordinates of the notehead of each note relative to the staff and detects the scale for each note stem. Sharp and flat symbols are unified as sharps during identification. Figure 2 is a scale list consisting of measure numbers and scales generated when the score in Figure 3 is uploaded.

1,F4,F3
1,C4
...
1,C4
2,D5,F3
2,C3
2,D5,A#3
...

Figure 2. Scale list.



Figure 3. Scores to be analyzed.

In the present study, we used OpenCV library [9] to identify a number of symbols.

C. Video Analysis Module

The video analysis module creates a time stamp list that records the start time of each measure from the uploaded video. The method for creating a time stamp list is as follows.

First, a constant-Q transformation is performed to generate a list of constant-Q values from the video. The constant-Q transform is a frequency analysis method that works well with pitch, chord, and melody analysis of musical signals [10].

Then, the scale list is read for each line, and the time of the beginning of the playing of each note stem is acquired from the list of constant-Q values. The acquisition condition is when the change in the constant-Q value exceeds a certain threshold value. After obtaining the start playing time of the first note, the time of the second note is searched. If the search process does not find the start playing time within a certain period of time, the next note is searched again. When the measure of the scale list to be searched changes, the measure number and the start time are recorded in the time stamp list.

In the present study, we used Librosa library [11] to audio analysis.

III. EXPERIMENTS

We confirmed the effectiveness of the proposed system. We conducted an evaluation of the score analysis module and the video analysis module. In this experiment, we prepared performance videos of playing each of the 12 measures of the “Twinkle, Twinkle Little Star” score (the C melody score, the arpeggio score, and the open harmony score) [12]. The evaluation of the score analysis module compares whether the identified scale is correct for each note stem with the score. The evaluation of the video analysis module compares the start time of each measure with the video.

Table I shows the aggregate results of the scale identification for each stem. The C melody score consists of 42 stems of 42 notes. The arpeggio score consists of 94 stems of 140 notes. Finally, the open harmony score consists of 68 stems of 164 notes. Pattern matching with scores resulted in a 100% notehead recognition rate for each of the three score stems.

TABLE I. RESULTS FOR IDENTIFICATION OF EACH STEM

	C Melody	Arpeggio	Open harmony
Detected/Total(stem)	42/42	94/94	68/68
Detected/Total(notehead)	(42/42)	(140/140)	(164/164)
Stem identification rate	100%	100%	100%

Table II shows the aggregate results for the match rate on the start time for each measure. The threshold of the constant-Q value for judging the beginning of a note in this experiment was set to 1.00. For the C melody score and arpeggio score, the start time of the video matched for all 12 measures out of 12. However, for the open harmony score, the start times did not match.

TABLE II. RESULTS FOR MATCH RATE ON THE START TIME

	C Melody	Arpeggio	Open harmony
Detected/Total(measure)	12/12	12/12	0/12
Measure match rate	100%	100%	0%

As a reason, the starting time was either undetectable or could be detected but was actually delayed in each measure. We focused on the constant-Q value to find out why the start time of any measure did not match the open harmony score. Open harmony contains overtone components of certain tones. The beginning of the first measure of the open harmony score is F4, A3, and F2, and the constant-Q value of F2 was 0.078 when F4 was 1.119 and A3 was 1.281. The maximum constant-Q value of F2 was 0.224 and never exceeded 1.00.

Therefore, in the open harmony section, instead of detecting all notes, it was necessary to detect either note as having started playing as being in an overtone relation.

IV. CONCLUSION

The purpose of the present study is to make it possible to playback videos from the starting point of each corresponding

measure by clicking on the measure. When looking back at the video, it was difficult to specify the playback position for each measure using the seek bar because fine control was required. In the present study, we developed a system that relates the starting point of each measure and the score and the video by analyzing the score and video and comparing the tone changes between the score and the video. As a result of experiments to verify the effectiveness of the system, we confirmed that the proposed method is effective for the C melody score and arpeggio score. On the other hand, through speech analysis of open harmony score, we found that there were cases in which overtone-related notes could not be detected by the fixed-Q transformation.

In the future, we intend to update the proposed system to relate the rise time of a note even when the note is an overtone and to validate the effectiveness of the proposed system using target music at a music academy.

ACKNOWLEDGMENT

The present study was supported by the Japan Society for the Promotion of Science (JSPS) through KAKENHI Grant Number 21K18528.

REFERENCES

- [1] Y. Takegawa, T. Terada, M. Tsukamoto, "Construction of a Piano Learning Support System considering Rhythm", *IPSJ Interaction 2012*, March 2012.
- [2] Y. Fukuya, K. Takegawa, H. Yanagi, "Design and Implementation of a Piano Learning Support System Considering Motivation", *IPSJ Interaction 2015*, March 2015.
- [3] T. Ishigami, T. Hamamoto, "A Piano Practice Support System Visualizing Correspondence Between Music Scores and Key Positions", *ITE Technical Report Vol.41, No14*, May 2017.
- [4] T. Nagai, K. T. Nakahira, M. Kitajima, "Analysis of the relationships between the proficiency levels of piano playing and the changes in visual behaviors while reading score and performing piano", *IPSJ SIG Technical Report, Vol.2017-CE-142 No.20*, December 2017.
- [5] T. Suzuki, K. Tanaka, R. Ogura, Y. Tsuji, "Practice of Beginners' Piano Skill Training Support Using "Visualization System for Piano performance (VSPP)"", *IPSJ SIG Technical Report, Vol.2018-MUS-119 No.16*, Jun 2018.
- [6] M. Hori, Christoph M. Wilk, S. Sagayama, "Visualizing deviations from exemplary performances for piano practice assistance (including retry detection)", *The 81st National Convention of IPSJ IT-02*, March 2019.
- [7] R. Matsui, K. Takegawa, K. Hirata, "Design, Implementation and Assessment of a remote piano lesson support system that automatically generates optimal multi-view camera work.". [Online]. Available from: https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_uri&item_id=177639&file_id=1&file_no=1 2023.01.30
- [8] R. Matsui, K. Takegawa, K. Hirata, "Design, implementation and Assessment of a Support System to Find Bad Fingering Habits for Piano Teachers", *2020 Information Processing Society of Japan Vol.61 No.4* 789-797, April 2020.
- [9] Open CV Library. [Online]. Available from: <https://docs.opencv.org/4.7.0/index.html> 2023.01.29
- [10] Constant-Q Transform (CQT) Description. [Online]. Available from: <https://www.wizard-notes.com/entry/music-analysis/constant-q-transform> 2023.01.30
- [11] Librosa Library. [Online]. Available from: <https://librosa.org/doc/latest/index.html> 2023.01.29
- [12] Score of "Twinkle Twinkle". [Online]. Available from: <https://atelier-music.com/sheetmusic/twinkle-twinkle-little-star> 2023.01.30