Chord Extraction Method in Development of a Score Click Playback System

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

Naoki Morita
School of Information Telecommunication Engineering
Tokai University
Tokyo, Japan
e-mail: morita@tokai.ac.jp

Chiharu Nakanishi
Kunitachi College of Music
Tokyo, Japan
e-mail: nakanishi.chiharu@kunitachi.ac.jp

Chiaki Sawada
Kunitachi College of Music
Tokyo, Japan
e-mail: sawada.chiaki@kunitachi.ac.jp

Kazue Kawai
Faculty of Literature
Seitoku University
Chiba, Japan
e-mail: kawai.kazue@wa.seitoku.ac.jp

***Abstract*—The purpose of this study is to recognize the locations of the chords from sound data of recorded piano performances. Previous methods have been unable to recognize all the musical scales contained within a chord and to identify the location of the chord. To identify the chord points, we adopted two ideas. The first idea is that, in instances where multiple notes of the same musical scale are present, if even one note reaches the predetermined threshold, it is considered as representing the respective scale. The second idea involves moderating the determination of whether the threshold has been exceeded. To validate the efficacy of the proposed method, piano performances of scores containing chords were recorded, and the method's ability to identify chord locations from the recorded data was assessed. The results demonstrated a notable improvement in identifying chord points with the proposed method, achieving an approximately 88% identification rate of musical notes, which is significantly superior to the approximately 14% achieved by the preceding method.**

***Keywords*—Chord, Score Analysis, Video Analysis, Piano Lesson.**

I. Introduction

Piano practice is crucially centered around repetitive practice. Students, for instance, record their performances and, through subsequent review, identify areas needing improvement. They then engage in continual practice, making necessary corrections based on these identified areas. Previous studies have proposed various methods to support piano lessons [1]–[8].

In this study, we have been developing a system that enables users to view a recorded piano performance from a specified bar by selecting a bar within a piano score displayed on-screen.

In the system we are developing, it is necessary to synchronize the visual onset of notes in the score with those in the performance video. To synchronize the visual onset, identification of the correspondence between notes in the score and the auditory output in the recorded performance is required. Wakiyama et al. have proposed a method for identifying which measure in a musical score corresponds to a played piano sound [9]. Piano performances encompass a variety of techniques, including chords, which involve striking two or more notes simultaneously using not only the fingers but also the pedals. This method is capable of identifying the musical scale of a note when only a single note is sounded momentarily. However, it was unable to identify the musical scale for sounds produced by striking two or more notes simultaneously, such as chords. The objective of this study is to recognize the scales of chords from recorded piano sounds.

In this paper, section II describes the previous method and its associated challenges. The proposed method will be explained in section III, and the validation of its efficacy and its consideration will be described in section IV. Finally, section V gives a summary of this study.

II. Previous Method

In this section, the preceding method for identifying musical scales is introduced, and the reasons for its inability to recognize chords are elucidated.

A. The Previous Method for Identifying Musical Scales of Piano Performances

The procedure utilized by the previous method to determine which notes in the musical score correspond to the sounds of the piano in the recorded performance is introduced herein. This method consists of five steps.

1) Recognize musical notes from images of the scores using OpenCV [10].
2) Retain the recognized notes as sequence data in chronological order.
3) Time-frequency analysis using Constant-Q Transform (CQT) [11] for sound data of recorded piano performance.
4) Extract notes one by one from the sequence data in step 2, and use the frequency corresponding to the musical scale as a search key to search the results of step 3 in chronological order.

 ISBN: 978-1-68558-087-2

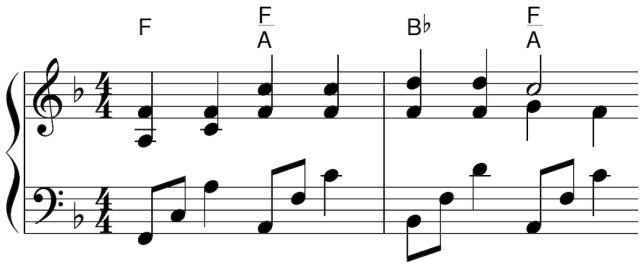

Figure 1. A score sample.

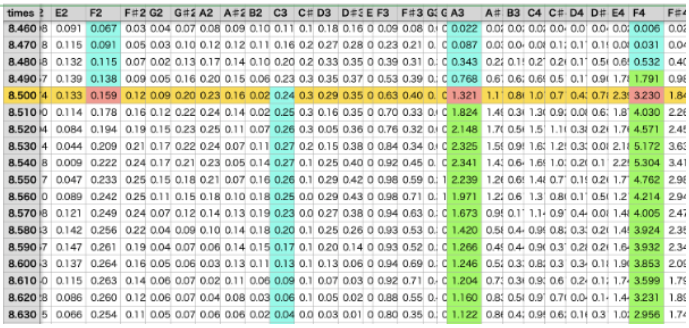

Figure 2. An example of CQT result.

TABLE I
THE ORDER DATA CONVERTED.

No.	Musical scale		
1	F4	A3	F2
2	C3		
3	F4	C4	A3
4	F4	A2	C5
5	F3		
6	F4	C4	C5
⋮	⋮		

TABLE II
A EXAMPLE TIMESTAMP LIST.

No.	Musical scale			Time
1-1	F4	A3	F2	8.500
1-2	C3			8.830
1-3	F4	C4	A3	9.200
1-4	F4	A2	C5	9.960
1-5	F3			10.310
1-6	F4	C4	C5	10.690
2-1	A#2	D5	F4	11.490
⋮	⋮			⋮

5) If the value is searched in chronological order and exceeds the threshold value, it is assumed that the searched note is found, and the appearance timing of the note in the performance data is recorded in a timestamp list. Steps 4 and 5 are repeated until the last note in the score is reached.

In step 1, the PDF or scanned music score is converted to image data, and the notes and lines of the score are recognized using recognition technology, specifically through template matching. The musical scale of the recognized note is identified by the position of the lines and the notes.

In step 2, the recognized musical scales are converted into sequence data in order from the left. Figure 1 presents a piano score sample, while Table I shows the sequence data converted by recognizing the musical scale from the score in Figure 1.

In step 3, a CQT is performed on the sound of the recorded piano performance. CQT is frequently utilized in the analysis of musical signals. By using CQT, the strength of frequency at each point of sound data can be obtained as a value. Figure 2 illustrates the result of applying the CQT of the music data of a certain performance.

Each musical scale is determined by frequency, and the horizontal axis of this table represents the music scale. The vertical axis indicates time and progresses downward.

In step 4, based on the converted ordinal data, each item is extracted in order and used as a search key. In the case of the table in Table I, the search is performed in order starting with No. 1. When No. 1 is selected, three scales, F4, A3, and F2, are subjected to the search. The search examines the frequency portion corresponding to the musical scale in chronological order for the values of the angular frequency band obtained by CQT.

In step 5, if the search identifies a point that exceeds the set threshold, it is assumed that that musical scale was sounded at that time. Then, the appearance timing of the note is recorded in a timestamp list. Table II is an example of a timestamp list. The number to the left of the No. is the bar number in the musical score, and the number to the right indicates the sequence in which the notes appear within that bar. The times in the table represent the times when it is assumed that the musical scale was pressed because the threshold was exceeded.

If the search process does not find the point that exceeds the set threshold within a certain amount of time, the musical scale of the next item is searched.

B. Some Problems of The Previous Method

Previous methods were unable to discern the timing of lower octave key presses when scales of identical types were played simultaneously. A chord is the sounding of multiple scales by striking two or more notes. In Figure 1, the initial note comprises three distinct notes: F4, A3, and F2. In this case, the key press timing for F2 could not be ascertained using the previous methods.

III. PROPOSED METHOD

In order to detect chords from the sound of recorded piano performance, two ideas are introduced.

The first idea is that if multiple notes within a chord share the same musical scale, and at least one of them surpasses a predefined threshold, all notes of that particular musical scale are deemed to have been pressed. For example, consider a scenario where F4 and F3, which share the same musical scale,

 ISBN: 978-1-68558-087-2

TABLE III
RECOGNITION RATE OF THE PROPOSED METHOD AND PREVIOUS METHODS.

	Proposed method				Previous method
Threshold	1.0	1.5	2.0	2.5	1.0
Chord only	29.55	63.64	88.64	0.00	11.36
ALL note	27.94	60.29	88.24	0.00	14.71

occur simultaneously in a score. If F4 exceeds the threshold, F3 is also considered to have been pressed, regardless of its actual value.

A second idea is to relax the determination of the threshold. Initially, a search is conducted for three scales. If this search is unfruitful, the number of scales is reduced to two, and subsequently to one if the search still yields no results. If the scale is not found in this way, the quantity of musical scales subjected to the search is reduced by one by one.

IV. Confirmation Of Extraction

In this section, the efficacy of the proposed method is substantiated through a structured experimental approach, aimed at ascertaining the capability of detecting chord points within the musical data of chord scores played on the piano.

A. Experimental Setup

To validate the efficacy, a comparative analysis between the note identification rates of the proposed method and the preceding method is conducted. The musical score used in this experiment contains not only singular notes but also chords [12], with performers utilizing pedals to actuate the keys during play The recognition of a note is determined by comparing the manually pre-verified start time of each measure with the adjudication time of each measure, as recognized by the system Both the proposed method and the previous method require a threshold setting as a parameter. The threshold of the previous method is designated at 1.0, whereas the proposed method is subjected to four threshold patterns: 1.0, 1.5, 2.0, and 2.5. The search timeout time was set to 2.6, correlating to the time span of one bar from the song.

B. Result

Table III shows the results of recognition rates for piano performances of musical scores containing chords. The proposed method achieved the highest note recognition rate at a threshold of 2.0. In addition, the recognition rate of the previous method was lower than that of the proposed method even when utilizing the same threshold.

In piano performances that incorporate chords, it was determined to be effective to assume a chord is identified when one or more of the notes sought are found. It was also observed that contingent upon the threshold, the position of the musical note might not be accurately recognized.

V. Conclusion

This study aimed to recognize chord points from the sound data of recorded piano performances. Previous methods were unable to recognize all scales within a chord, nor could they identify the point of the chord. To identify the chord points, we adopted two ideas. The first idea is that, when multiple notes of the same musical scale are present, if at least one reaches the threshold, it is considered as the same note. The second idea involves moderating the determination of whether the threshold has been exceeded. To validate the effectiveness, a piano performance of a musical score containing chords was recorded, and it was confirmed whether the proposed method could identify chord locations from the recorded data. As a result, the proposed method was able to identify chord points more effectively than the previous method. Future work will involve experimenting with other songs.

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, and M. Tsukamoto, “Construction of a Piano Learning Support System considering Rhythm”, IPSJ Interaction 2012, March 2012.

[2] Y. Fukuya, K. Takegawa, and H. Yanagi, “Design and Implementation of a Piano Learning Support System Considering Motivation”, IPSJ Interaction 2015, March 2015.

[3] T. Ishigami and T. Hamamoto, “A Piano Practice Support System Visualizing Correspondence Between Music Scores and Key Positions”, ITE Technical Report Vol.41, No14, pp. 71–76, 2017.

[4] T. Nagai, K. T. Nakahira, and 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, and 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, pp. 1–6, Jun 2018.

[6] M. Hori, Christoph M. Wilk, and S. Sagayama, “Visualizing deviations from exemplary performances for piano practice assistance (including retry detection)”, The 81st National Convention of IPSJ 1T-02, pp.337–338, March 2019.

[7] R. Matsui, K. Takegawa, and 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 [retrieved: 10, 2023]

[8] R. Matsui, K. Takegawa, and 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, pp. 789–797, April 2020.

[9] M. Wakiyama, M. Wakao, N. Morita, K. Morita, C. Nakanishi, C. Sawada, and K. Kawai,“Development of a Score Click Playback System”, Proceedings of The Eighth International Conference on Advances in Signal, Image and Video Processing, pp.21–23, 2023.

[10] Open_CV Library. [Online]. Available from:https://docs.opencv.org/4.7.0/index.html [retrieved: 10, 2023]

[11] Constant-Q Transform (CQT) Description. [Online]. Available from: https://www.wizard-notes.com/entry/music-analysis/constant-q-transform [retrieved: 10, 2023]

[12] Score of “Twinkle Twinkle”. [Online]. Available from:https://atelier-music.com/sheetmusic/twinkle-twinkle-little star [retrieved: 10, 2023]

 ISBN: 978-1-68558-087-2