An Interactive Tool for Exploring Score-Aligned Performances: Opportunities for Enhanced Music Engagement

Caitlin Sales
caitlin.sales@richmond.edu
University of Richmond
Richmond, Virginia, USA

Peiyi Wang
john.wang@richmond.edu
University of Richmond
Richmond, Virginia, USA

Yucong Jiang
yjiang3@richmond.edu
University of Richmond
Richmond, Virginia, USA

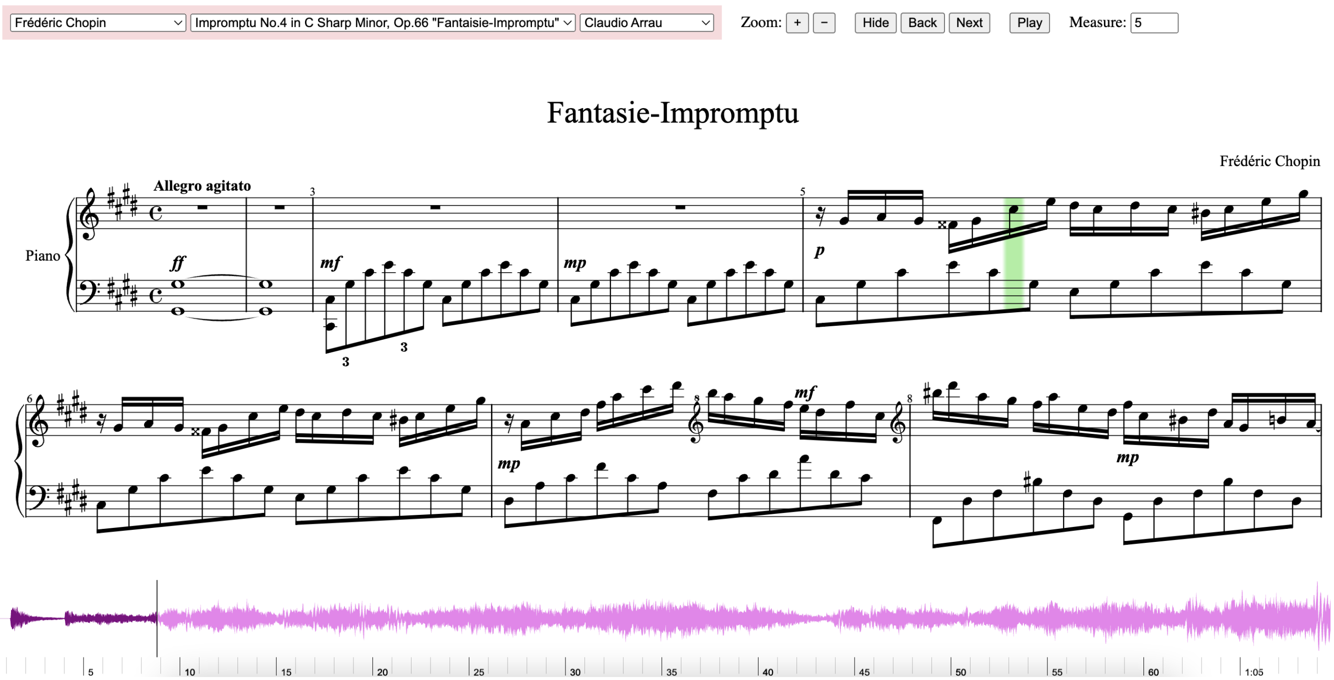

Figure 1: Web-based interface for exploring classical music with score-aligned performances.

ABSTRACT

Music scholars and enthusiasts have long been engaged with both performance recordings and musical scores, but inconveniently, these two closely connected mediums are usually stored separately. Currently, digital music libraries tend to have fairly traditional user interfaces for browsing music recordings, and more importantly, performance recordings are organized separately from their musical scores. In recent years, however, the same technological advances that have made vast troves of sound recordings and musical scores more widely available have also created tremendous potential for innovative new interfaces that can facilitate enhanced engagement with the music. In this paper, we present a web-based prototype tool that allows users to navigate classical piano recordings interactively in conjunction with their scores. We describe the technologies involved, and provide access to the actual website. Our pilot testing results are very positive, confirming the usefulness and potential of such a tool, especially in the areas of music education and scholarly research. We also discuss future development of this prototype.

CCS CONCEPTS

• **Human-centered computing → Web-based interaction**; • **Applied computing → Sound and music computing**.

KEYWORDS

web-based interaction design, digital music libraries, audio-to-score alignment, piano music performance

AM '23, August 30–September 01, 2023, Edinburgh, United Kingdom

ACM ISBN 979-8-4007-0818-3/23/08.
https://doi.org/10.1145/3616195.3616212

ACM Reference Format:
Caitlin Sales, Peiyi Wang, and Yucong Jiang. 2023. An Interactive Tool for Exploring Score-Aligned Performances: Opportunities for Enhanced Music Engagement. In *Audio Mostly 2023 (AM '23), August 30–September 01, 2023, Edinburgh, United Kingdom.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3616195.3616212

1 INTRODUCTION

Music scholars and enthusiasts have long been engaged with both performance recordings and musical scores, but inconveniently, these two closely connected mediums are usually stored separately. In the digital era, however, vast troves of both sound recordings and musical scores have become widely available, presenting a great need (and corresponding opportunity) to create a coordinated digital music infrastructure that enriches the experience of engaging with music. (One recent effort along these lines is the TROMPA project [10], which deploys multiple web and music technologies for better access to musical content from selected repositories.) Currently, digital music libraries (e.g., the British Library's classical music sound archive [20] and the Vienna Symphonic Library [21]) have a fairly traditional user interface for listening to music recordings: users can only start, pause, or jump to a new starting point of the playback. More importantly, performance recordings are organized separately from their musical scores. The goal of our project, therefore, is to explore the potential of an interactive tool that allows users to navigate recordings interactively in conjunction with their scores. This kind of interactive interface could offer benefits not only to archivists, but also music educators, the public, and anyone interested in the nuances of musical performances.

Various innovative interfaces have been developed for searching, analyzing, or visualizing music data. For instance, [12] describes a system where users can search for music by content instead of by metadata (per tradition); [22] proposes a new interface for visualizing and browsing scores. Sonic Visualiser [1] is a mature software tool for analyzing and visualizing acoustic features of sound recordings (including music recordings), and Sonic Lineup [19] is a related tool for comparing different recordings of the same composition. [4] and [8] are examples of visualizing music performances, and [9] is a recent survey of music visualization techniques. However, none of these interfaces explicitly shows the connection between performances and scores.

Connecting performances and scores at the technical level is a familiar topic in Music Information Research (MIR), and the most relevant challenge is known as *audio-to-score alignment*, whereby performance recordings are aligned to their scores, associating each note in the score with an onset timestamp in the audio. Numerous approaches have been proposed within this topic, including [13][18][2]. In response to this challenge, various open datasets of score-aligned performances have been created, such as the data in [6] and [25]. Some open datasets don't contain alignments at the note level, but annotations of beat, harmony, phrases, or other score-related features [5][7]. In our project, the score-aligned performances come from [25].

There are currently two commonly used open-source tools for representing and rendering musical scores on the web: OpenSheetMusicDisplay (OSMD) [15], which facilitates rendering of sheet music from the MusicXML score format, and Verovio [17], which renders sheet music from the MEI score format. For our project we employ the former, as the MusicXML format is presently more widely used than MEI. Two demo projects closely related to ours, [16] and [24], also have interfaces that allow users to interactively navigate professional recordings while automatically scrolling rendered sheeting music, but both use Verovio with MEI. The first of these is designed for easy comparison among different performances of the same composition, and the second for easy switching among audio tracks during choir rehearsals. Both demos use only a fairly small repertoire (three and five scores respectively), and neither provides any user evaluations.

Table 1: Piano Music in Our System.

Composer	# of Scores	# of Performances
Claude Debussy	5	15
Franz Liszt	2	4
Franz Schubert	3	6
Frédéric Chopin	5	14
Johann Sebastian Bach	6	7
Ludwig van Beethoven	3	5
Wolfgang Amadeus Mozart	1	3

Another category of related work is educational software for instrument learning, such as SmartMusic [11], where notes played by the user are compared to an interactive score. Besides differences in purpose, however, the sound synchronized to the score in such applications is either synthesized or provided by the user, whereas ours uses professional recordings.

2 METHOD

In this section, we describe the interaction design of the tool, our data preparation procedure, and the technical details involved in building this prototype.

2.1 Interaction Design

Figure 1 shows a screenshot of our web-based prototype tool allowing users to explore piano classical music interactively. The three drop-down menus at the top allow the user to select a composer, a score, and a performance of the score (by performer's name). Table 1 summarizes the music data in our system so far. Once a score is selected, its sheet music is rendered on the screen; once a performance is selected, an audio waveform of the recording displays at the bottom of the screen, along with a time scale indicating minutes and seconds into the recording. During playback, a green cursor will move along on the sheet music, identifying the current note(s) being played. The sheet music will scroll automatically, centering the staves currently being played. The audio waveform spans only about 70 seconds on the screen (the exact length depends on the screen size); the waveform and the time scale underneath will also automatically scroll as the music plays.

To allow users to customize their experience, we have implemented a cluster of options at the top of the screen. Besides toggling between "Play" and "Pause", the user can choose to "Show" or "Hide" the green cursor, and to zoom the sheet music in or out in order to fit fewer or more measures on the screen. By specifying a measure number in the "Measure" text box, the user can jump to a new playback position starting from the first beat in that measure. For example, if the user specifies "5" in the example in Figure 1, the cursor would jump to the first C on the bass clef in Measure #5, and start the playback from there. If the user wants to start from a beat other than the first beat in a measure, he or she can additionally click on the "Back" or "Next" buttons to move the cursor backward or forward note by note.

The waveform at the bottom is also interactive. If the user wants to jump to a specific starting time in the recording, he or she just needs to click on the target timestamp on the waveform. If the target timestamp is beyond the region shown on the screen, the user can manually scroll the waveform to the left or right to find the target timestamp.

This prototype tool can be accessed through the link https://facultystaff.richmond.edu/~yjiang3/papers/am23/.

2.2 Score-Aligned Performances

Our system currently contains 54 performances played by 20 world-renowned pianists; the compositions come from 25 scores by seven famous piano composers, as listed in Table 1. Typically, a score has two or three performance interpretations. Each performance recording is preprocessed to align to its score, and this alignment is represented in an *alignment file* by a sequence of <score time, audio time> pairs indicating when each score time occurs in the audio. For example, the alignment for the highlighted C note in Figure 1 is represented as <5+3/8, 8.935>, meaning the seventh sixteenth note (calculated as (7-1)/16) in Measure 5 has an onset time at 8.935 seconds in the audio. Notes in a chord share one alignment pair. Below, we describe the preprocessing steps for generating the alignment files.

Both the digital scores and the performances come from (a small portion of) the ATTEP dataset [25]; the scores are in the MusicXML format, and we downloaded the audio performances in the WAV format through the YouTube links provided by ATTEP. Additionally and importantly, ATTEP provides a music transcription file for each performance, in the MIDI format, indicating what pitches are being played and when. The transcriptions, however, do not contain information about how played pitches (and their timestamps) are associated with notes in the scores. To derive this information, we used the MIDI-to-score alignment tool from [14], which takes a performance MIDI file (*aka* the transcription) and a MusicXML score and generates a *matching file* for them. A matching file not only contains the onsets of each note played correctly according to the score, but also information about incorrect, extra, and missing notes in the performance. When determining the audio time for each score time in the alignment file described above, we discard onsets of extra notes, average onsets of notes from the same chord, and estimate onsets for missing notes according to their predecessors and successors. It is worth mentioning that the matching file uses a score time format different from the one in the alignment file. The latter adopts a common format (i.e., the same as in OSMD, described in Section 2.3), and one can easily extract all such score times in a MusicXML score using a popular score-processing Python tool called music21 [3]. We nevertheless were able to translate the format used in the matching file to the latter format reliably.

2.3 Implementation Details

For representing and rendering sheet music on the web from MusicXML files, we use the open-source JavaScript library OSMD [15], as mentioned in the Introduction. OSMD maintains the general hierarchical structure of a MusicXML score, consisting of systems, measures, notes, and other score elements. This library provides convenient functions for controlling visual effects in a rendered score, including drawing/moving the green cursor and scrolling the score.

The interactive audio waveform at the bottom of our interface is powered by wavesurfer.js [23], an open-source JavaScript library that helps generate customizable waveform visualizations of audio files. One of the main features of wavesurfer.js is its ability to create highly interactive visualizations of audio waveforms, allowing users to easily visualize and navigate through audio files. For example, we customize the height and color of the waveform, display a time scale that moves together with the waveform, create regions within the audio for individual notes, and allow users to fast-forward or rewind by clicking on any desired portion of the waveform.

Based on the information in the alignment file described in Section 2.2, we create a cause-and-effect loop between the cursor position in the sheet music and the playback timestamp in the audio. Whenever the user moves the cursor (either by entering a measure number or by clicking on the "Back" or "Next" buttons), our program searches for that score time and sets its onset as the new playback time. Likewise, whenever the user clicks on a timestamp in the waveform, our program searches for the corresponding score time (i.e., the latest score time with an onset earlier than the selected timestamp) and moves the cursor to that score time.

3 USER EVALUATION

We conducted a pilot test to evaluate our user interaction design and to gather feedback for future development. We recruited seven participants whose backgrounds broadly indicated they might be potential users of our tool—people who enjoy listening to classical piano music and who can also read sheet music. We started each interview session with a set of warm-up questions about the participant's musical background, what they like the most about classical music, their music listening habits, and some basic demographic information. The ages of the participants ranged from 20 to 60 (four are college students around 21), and only one participant was male. All participants had played one or more instruments (typically piano) for at least a decade, and two had music degrees. After the warm-up questions, we briefly demonstrated the functions of the interface before letting the participant explore the interface freely. To ensure privacy and minimize distractions, the participant was left alone in a quiet room, with instructions to try each functionality at least once, and also at least once to compare the experience of looking at the interface vs. not looking. We gave each participant at least 15 minutes to explore the interface thoroughly, more if requested. Afterward, we conducted semi-structured interviews with a set of follow-up questions regarding the users' experiences.

Six of the seven participants expressed that they enjoyed the experience or would use such a tool. Four of them mentioned that they appreciated being able to read along with the music and keep track of places in the score (especially helpful for fast pieces). The other two shared that they wouldn't use this interface for recreational listening, but would use it for intellectual engagement, such as teaching or learning a piece. The remaining participant found the listening experience stressful, mainly because of the cursor's motion: she would have preferred the cursor to move continuously rather than "jumping" from beat to beat. When asked to compare the experience of looking vs. not looking at the interface, most participants pointed out that when looking at the interface, they paid more attention to the musical elements in the score or thought more about the composer's intentions.

We also asked the participants whether such an interface would be useful for comparing different interpretations of a piano piece. All but one participant answered yes, and a few also described it as a great tool for learning to play a piece as it would make it easy to compare different interpretations of a score. The participant who answered *no* is a more specialized user with a background in musicology; she said she would prefer simultaneous representations of multiple performances instead of one performance at a time—for example, visualizing and comparing the volume changes of multiple performances simultaneously. Two other participants also suggested lining up multiple performances in some way for easier comparison.

From these interviews we learned that our current interface offers features and functions that serve distinct purposes appreciated by users. All the buttons, as well as the text box (for jumping to a measure), were valued by at least one participant. We also asked users about any additional functionalities or improvements that might be desirable; most suggested additions would be easy to implement (for example, being able to scroll ahead while the music is playing). Three suggested they would like to be able to click on the sheet music to jump to a measure, which would require more effort to implement. Three users noticed occasional errors in the score or in the alignments, and expressed that they would appreciate higher-quality music data.

4 DISCUSSIONS AND CONCLUSIONS

The overall pilot test results are positive: almost all participants found the tool valuable. Their responses confirm strong potential for applications that synchronize performances with scores, an underexplored area of music applications. One prospective use mentioned often by the participants would be education (e.g., for learning how to play a piece). Another, implicitly described by the musicology participant, would be scholarly research. For both cases, participants expressed interest in comparing multiple interpretations of the same composition; although almost all confirmed our interface would be useful for such purposes, some suggested additional features to facilitate easier comparisons.

Our next steps include three aspects. First, we plan to improve the current interface design based on the evaluation feedback: e.g., allowing users to choose the motion of the cursor. Second, we want to explore new ways to compare different interpretations of a composition: e.g., by visualizing multiple performances simultaneously and offering seamless switching among recordings. Finally, we want to expand the repertoire for our tool, and improve both the score quality and the alignment accuracy.

In conclusion, this paper presents an innovative tool with an interface that allows users to explore score-aligned performances. Positive user evaluations suggest that such a tool has great potential in music education and scholarly research. We anticipate that further development and testing will refine this tool's ability to meet the needs of a wide range of users.

REFERENCES

[1] Chris Cannam, Christian Landone, and Mark Sandler. 2010. Sonic Visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the 18th ACM international conference on Multimedia*. 1467–1468.

[2] Julio José Carabias-Orti, Francisco J Rodriguez-Serrano, Pedro Vera-Candeas, Nicolás Ruiz-Reyes, and Francisco J Canadas-Quesada. 2015. An Audio to Score Alignment Framework Using Spectral Factorization and Dynamic Time Warping.. In *International Society for Music Information Retrieval Conference*. 742–748.

[3] Michael Scott Cuthbert and Christopher Ariza. 2010. music21: A toolkit for computer-aided musicology and symbolic music data. In *International Society for Music Information Retrieval Conference*.

[4] Simon Dixon, Werner Goebl, and Gerhard Widmer. 2002. The Performance Worm: Real time visualisation of expression based on Langner's tempo-loudness animation. In *ICMC*.

[5] Francesco Foscarin, Andrew Mcleod, Philippe Rigaux, Florent Jacquemard, and Masahiko Sakai. 2020. ASAP: a dataset of aligned scores and performances for piano transcription. In *International Society for Music Information Retrieval Conference*. 534–541.

[6] Werner Goebl. 2001. Melody lead in piano performance: Expressive device or artifact? *The Journal of the Acoustical Society of America* 110, 1 (2001), 563–572.

[7] Johannes Hentschel, Markus Neuwirth, and Martin Rohrmeier. 2021. The annotated mozart sonatas: Score, harmony, and cadence. *Transactions of the International Society for Music Information Retrieval* 4, ARTICLE (2021), 67–80.

[8] Rumi Hiraga and Noriyuki Matsuda. 2004. Visualization of music performance as an aid to listener's comprehension. In *Proceedings of the Working Conference on Advanced Visual Interfaces*. 103–106.

[9] Hugo B Lima, Carlos GR Dos Santos, and Bianchi S Meiguins. 2021. A survey of music visualization techniques. *ACM Computing Surveys (CSUR)* 54, 7 (2021), 1–29.

[10] David M. Weigl, Werner Goebl, Tim Crawford, Aggelos Gkiokas, Nicolas F. Gutierrez, Alastair Porter, Patricia Santos, Casper Karreman, Ingmar Vroomen, Cynthia CS Liem, et al. 2019. Interweaving and enriching digital music collections for scholarship, performance, and enjoyment. In *6th International Conference on Digital Libraries for Musicology*. 84–88.

[11] MakeMusic 2023. *SmartMusic*. Retrieved April 27, 2023 from https://www.makemusic.com/

[12] Rodger J McNab, Lloyd A Smith, Ian H Witten, Clare L Henderson, and Sally Jo Cunningham. 1996. Towards the digital music library: Tune retrieval from acoustic input. In *Proceedings of the first ACM International Conference on Digital libraries*. 11–18.

[13] Meinard Müller, Frank Kurth, and Tido Röder. 2020. Towards an Efficient Algorithm for Automatic Score-to-Audio Synchronization.. In *International Society for Music Information Retrieval Conference*.

[14] Eita Nakamura, Kazuyoshi Yoshii, and Haruhiro Katayose. 2017. Performance Error Detection and Post-Processing for Fast and Accurate Symbolic Music Alignment. In *International Society for Music Information Retrieval Conference*. 347–353.

[15] OpenSheetMusicDisplay 2023. *The code repository of OpenSheetMusicDisplay*. Retrieved April 27, 2023 from https://github.com/opensheetmusicdisplay/opensheetmusicdisplay

[16] Thomas Prätzlich, Meinard Müller, Benjamin W Bohl, Joachim Veit, and Musikwissenschaftliches Seminar. 2015. Freischütz digital: demos of audio-related contributions. In *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR), Málaga, Spain*.

[17] Laurent Xavier Pugin, Rodolfo Zitellini, and Perry Roland. 2014. Verovio: A library for engraving MEI music notation into SVG. In *International Society for Music Information Retrieval Conference*. 107–112.

[18] Christopher Raphael. 2006. Aligning music audio with symbolic scores using a hybrid graphical model. *Machine Learning* 65, 2-3 (2006), 389–409.

[19] Sonic Visualiser 2023. *Sonic Lineup*. Retrieved April 27, 2023 from https://www.sonicvisualiser.org/sonic-lineup/

[20] The British Library 2023. *The British Library's classical music sound archive*. Retrieved April 27, 2023 from https://sounds.bl.uk/Classical-music/

[21] The Vienna Symphonic Library 2023. *The Vienna Symphonic Library's music archive*. Retrieved April 27, 2023 from https://www.vsl.co.at/en/Music

[22] Fumiko Watanabe, Rumi Hiraga, and Issei Fujishiro. 2003. Brass: Visualizing scores for assisting music learning. In *ICMC*.

[23] wavesurfer.js 2023. *The website of wavesurfer.js*. Retrieved April 27, 2023 from https://wavesurfer-js.org/

[24] Frank Zalkow, Sebastian Rosenzweig, Johannes Graulich, Lukas Dietz, El Mehdi Lemnaouar, and Meinard Müller. 2018. A web-based interface for score following and track switching in choral music. In *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR)*.

[25] Huan Zhang, Jingjing Tang, SR Rafee, Simon Dixon, George Fazekas, and Geraint A Wiggins. 2022. ATEPP: A dataset of automatically transcribed expressive piano performance. In *International Society for Music Information Retrieval Conference*. 446–453.