PIANO PRECISION: ADVANCING MUSIC PERFORMANCE ANALYSIS BY INTEGRATING USER-CORRECTABLE AUDIO-TO-SCORE ALIGNMENT

Yucong JIANG (yjiang3@richmond.edu) [1], **David M. WEIGL** (weigl@mdw.ac.at) [2], and **Werner GOEBL** (goebl@mdw.ac.at) [2]

[1] *Department of Computer Science*, **University of Richmond**, VA USA
[2] *Department of Music Acoustics*, **University of Music and Performing Arts**, Vienna, Austria

ABSTRACT

Audio-to-score alignment is a core task in MIR, but its application to music performance analysis is underexplored and it has not been developed for prospective users. Moreover, there have been very few user studies to shed light on the software needs of performance researchers. To help fill this gap, we present an open-source software application with a graphical UI that integrates audio-to-score alignment, designed to facilitate performance analysis. Working with Vamp plugins, this tool can also serve as a suitable platform for developing and comparing new alignment algorithms. In addition, in this paper we offer a use-case example of this tool comparing multiple performances of the same piece. We have also evaluated this tool with 15 prospective real-world users and received highly positive responses. Our user study suggests that tools like this can help bridge disciplinary divides between MIR and performance research.

1. INTRODUCTION

Audio-to-score alignment is a core task in Music Information Retrieval (MIR). Given a digital musical score and a performance recording thereof, an alignment algorithm automatically finds the onset times of notes from the score in the audio. However, despite its natural link to the field of music performance analysis, there are no software applications integrating audio-to-score alignment for performance research. Indeed, attempts to understand the software needs of relevant performance researchers are scarce. In this paper, we present an open-source software tool, *Piano Precision*, with an intuitive graphical UI and integrated audio-to-score alignment, designed to facilitate performance analysis. We also report the findings from interviewing 15 scholars in this highly specialized user group.

Music scholars doing performance analysis are interested in examining performance characteristics such as style and expressiveness. A common scenario involves comparing multiple performances of the same piece. For example, [1] analyzed recordings of the same composition over almost four decades to illustrate historical trends in performing styles, cross-examined with social and cultural contexts. Another common scenario is tracing a particular artist's stylistic changes over time (e.g., [2–4]). One essential aspect of performance style and expressiveness is timing, and it is among the most researched [5]. Typical analyses examine tempo fluctuations, or microtiming such as in *ritardando* and *rubato* [6].

Arguably the most popular software among music scholars for analyzing music recordings is Sonic Visualiser (SV) [7–9]. SV can automatically calculate low-level audio attributes through various specialized plugins, called Vamp plugins, which can then be displayed onscreen (e.g., as a spectrogram). Other Vamp plugins can directly generate high-level musical annotations about, e.g., pitch [10], beat [11], and articulation [12], often optimized for certain kinds of music. What makes SV particularly helpful is that these attributes and annotations, either manually entered or auto-generated, can be visualized together, and they all can be exported as CSV files for further quantitative analysis. Although a number of timing-related Vamp plugins are available for automatic annotation, however, the results are unreliable; none of them "can afford robust enough data for musicological research... [and] working with SV thus remains manual work" [13]. For example, from experiments with four note-onset detection plugins, [14] reports that the accuracy of auto-detected onsets is too low for any meaningful analysis of timing in music.

Moreover, SV does not incorporate musical score information, often a vital element for analyzing performances. In *Piano Precision*, note onsets are recognized by the score aligner component, which runs an audio-to-score alignment algorithm. In contrast to the onset detectors mentioned before, the score aligner is aware of the score and thus always returns the expected number of onsets. In addition, *Piano Precision* allows users to correct any alignment errors with intuitive visual cues, e.g., highlighted notes in the rendered score.

There has been voluminous MIR research on audio-to-score alignment, such as [15–20]. (Our focus here is offline alignment, not real-time.) However, there are no unified platforms for evaluating different alignment algorithms, nor are there platforms for inspecting alignment errors intuitively [21], hindering advancement of this task. *Piano Precision* can serve as such a platform on both counts: its score aligner component is a score-aware Vamp plugin, and therefore MIR researchers can use alternate alignment algorithms within the same platform by developing and in-

ISBN 978-3-200-10642-0

stalling new plugins.

Besides SV, other performance analysis applications used among music scholars without requiring programming skills include SKYLINE [22], specialized for timing, Vmus.net [23] for performance visualization, Sonic Lineup [24] and Listen Here [25] for computer-assisted close listening, and Tony [26] for solo vocal music. Other relevant projects include MPM [27] for editing encoded performance markups, and Parangonada [28] and Nakamura [29] for aligning MIDI performances. However, except for MPM, no existing application integrates musical scores in its UI.

The contribution of this paper is three-fold: i) introducing *Piano Precision*, to our knowledge the first open-source tool designed for the study of music that features audio-to-score alignment in a graphical UI; ii) a platform for audio-to-score alignment research in MIR; and iii) a user study confirming the value of such tools for performance research and providing insights for future development.

2. SOFTWARE OVERVIEW

The *Piano Precision* UI is adapted from Sonic Visualiser, with an additional area displaying a score, as shown in Figure 1. The interface is designed to help streamline the processes of loading a digital score and a performance recording, editing and visualizing annotations (note onsets and tempo curve in this initial version), and exporting score-aware annotations. The software installs easily on a Mac or Windows machine. The installers and source code are published on GitHub at `https://github.com/yucongj/piano-precision`. That page also includes a detailed user manual.

Piano Precision emphasizes audio-to-score alignment functionality. Estimated onsets ("time instances") are represented by the vertical lines on top of the spectrogram, as shown in the upper pane in Figure 1. The lower pane shows the waveform of the audio and a tempo curve calculated from the onsets. Users can add more panes and data layers just as in SV. As a user navigates different parts of the recording, the corresponding notes in the score are highlighted with a light-blue marker. Conversely, the user can click on notes in the score to quickly jump to the targeted playback place.

Of course, no alignment algorithm is error-free. If estimated onsets are misaligned with the actual performance (which can be discovered with the spectrogram and onset clipping sound), users can correct them easily by dragging the onset indicators, with the assistance of highlighted notes in the score and text labels representing the position of selected notes. As the onsets are edited, the tempo curve is updated in real time. Users can then export these validated onsets and the tempo curve (and any user-added annotations) to CSV.

2.1 MEI Score Format

The score input format is MEI (Music Encoding Initiative) [30,31], an increasingly popular format among music scholars [32]. MEI has the advantage of being semantic-oriented and thus well-equipped for music research. Technically, what makes MEI especially useful is that it can be rendered to graphical SVG output while retaining its semantically meaningful XML hierarchy and unique identifiers for every score element (e.g., ties, notes). *Piano Precision* renders the scores into SVG output using Verovio [33, 34], an open-source toolkit for visualizing MEI data. This makes it easy to locate interesting objects in a score and, e.g., highlight them in the rendering. The result is interactive images in which users can flip pages, zoom in and out of a score, and click on elements in a score to jump to their playback positions. Increasing numbers of scores are becoming available in MEI, e.g., Mozarteum [35], which hosts the critical edition of Mozart's works. In addition, the popular crowd-sourced score platform MuseScore [36] supports conversions from MusicXML to MEI (as of MuseScore 4.2). The recently developed MEI score editor mei-friend [37] also handles format conversions, and may be used for final correction and validation to produce high-quality scores.

2.2 Vamp Plugins and Audio-to-Score Alignment Vamp Plugins

Vamp plugins are audio-analysis plugins that generate low-level features or high-level annotations about a signal, which can then be visualized in host applications. They can be designed to process all kinds of audio signals, although most existing plugins are developed for analyzing music recordings, especially with the popular host application SV. The open-source Vamp website [38] includes detailed documentations about developing new plugins. In our host application, *Piano Precision*, audio-to-score alignment is conducted by special Vamp plugins designed to generate estimated onset times of notes in a given score. Such aligner plugins take an input of a musical score (in the MEI format) provided by users through the host. The host receives an alignment result from a plugin and visualizes it for users. All installed aligner plugins reside in a shared library and users can switch among them from within the host at any analysis stage. On the GitHub page, we have provided documentation for anyone who would like to develop new aligner plugins for *Piano Precision*. Any new plugin needs to parse scores in the same way as the host (code provided), and it needs to provide the alignment result with a particular output id (example code provided).

In any audio-to-score alignment, a score is represented as a sequence of events ordered by their score times, and the algorithm estimates the onset time of each event in the audio recording. The default aligner plugin in *Piano Precision* adopts the classic hidden Markov model from [39]; we have published the plugin code at `https://github.com/yucongj/piano-aligner`. As typical for polyphonic music, notes sharing the same score time are treated as a single event.

We hope the convenient graphical UI of Piano Precision, with its modular integration of alignment plugins, will alleviate today's difficulty in comparing alignment algorithms due to the lack of clear benchmarks. By providing a clearer qualitative sense of strengths and weaknesses, our tool

ISBN 978-3-200-10642-0

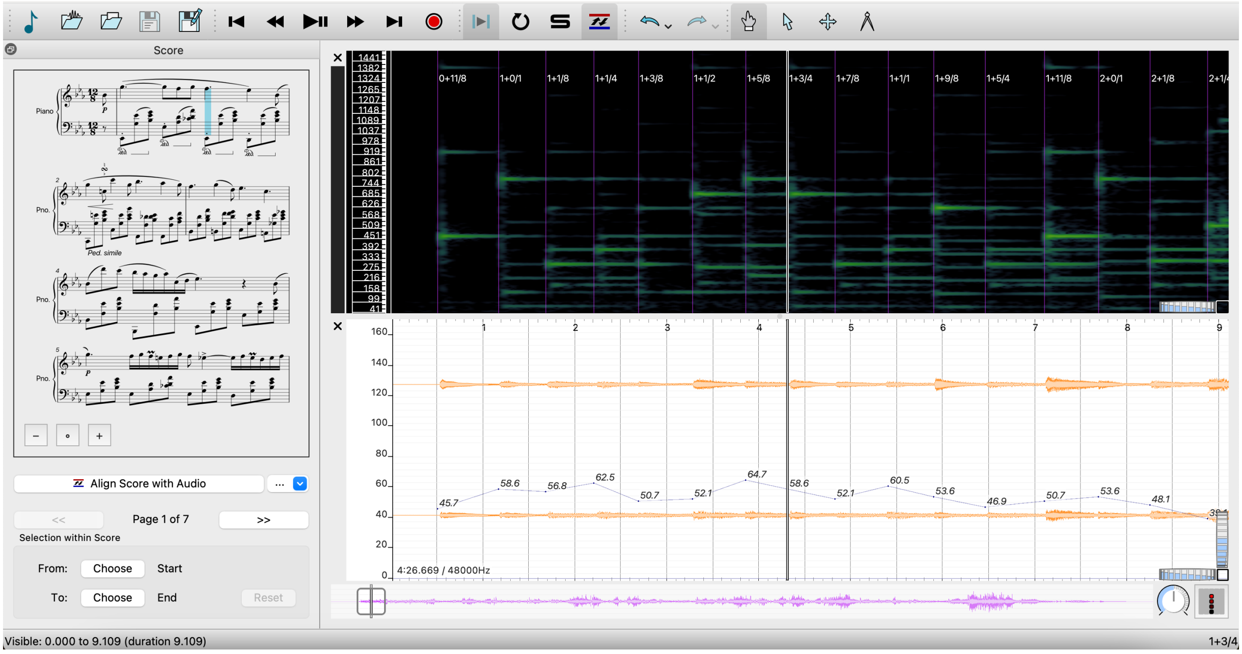

Figure 1. User interface of *Piano Precision* (adapted from Sonic Visualiser).

could ease the path to developing better alignment plugins with more accurate algorithms, perhaps also optimized for different instruments or genres. Such improvements would directly benefit performance analysis researchers.

2.3 Exporting Alignments and Other Functionalities

The ability to export and import score alignments is an important feature of *Piano Precision*. Exported alignment data are stored in CSV files in a straightforward way: one column containing the score time of an event, and the other column containing its onset time. Table 1 shows the first six rows of the exported alignment file for the performance in Figure 1. A score time is represented by a text label specifying the measure number and the position within the measure. For example, the label "1+0/1" stands for the beginning of the first measure, and the label "1+3/4" stands for *after* three quarter notes worth of time in the first measure; the latter score position is highlighted in Figure 1. These labels are also indicated in the spectrogram pane, next to the vertical line of each onset.

Label	Time
0+11/8	0.512
1+0/1	1.168
1+1/8	1.68
1+1/4	2.208
1+3/8	2.688
1+1/2	3.28
...	...

Table 1. Example of exported alignment data containing score labels and onset times (in seconds).

In cases where only timing at the measure level is of interest to the user—a common scenario in performance analysis—one can easily filter rows by label (using Excel) to get "1+0/1, 2+0/1, 3+0/1, . . . ". Similar strategies can be applied to filter rows at the beat level. A CSV alignment file can be imported back to *Piano Precision*; if only a subset of rows are kept, then only those onset lines display on screen. The use case in Section 3 examines tempo changes at both the measure level and the beat level.

Piano Precision supports partial alignments, aligning only one section of the score. This feature is useful when the whole recording of a score is not available, or when users are interested in analyzing only a particular section. Exported alignment files from partial alignments still contain all the score labels, and the ones outside of aligned regions would have "NA" as their onset times.

The controls for correcting and fine-tuning alignments are designed to emulate (and expand on) the controls in SV, to enhance user familiarity and ease-of-use. Many features (e.g., playback speed, navigation, selection, editing, and zoom) can be executed using either on-screen icons or keyboard commands.

3. USE CASE EXAMPLE: SEVEN FAMOUS PIANISTS

To provide a use case example of *Piano Precision*, we collected recordings of seven famous pianists playing *Chopin Nocturne Op. 9 No. 2*. Table 2 lists these pianists, sorted chronologically. (We can find only a range of years for two of the recordings, and we use their middle years for the purpose of sorting.) This piece is written in $\frac{12}{8}$ meter and has 34 measures. Our tempo analysis focuses on the first 31 measures, where the left hand follows the same pattern with four groups of three beamed eighth notes in each measure; the pattern can be observed in Figure 1, displaying the first five measures. We conduct analysis at both the measure level and the beat level (i.e., at every eighth note).

We used *Piano Precision* to generate alignment data

ISBN 978-3-200-10642-0

Pianist Name	Year Recorded
Nelson Freire	2009
Maurizio Pollini	2005
Daniel Barenboim	1982
Vladimir Ashkenazy	1974-1984
Vladimir Horowitz	1957
Arthur Rubinstein	1950-1958
Sergei Rachmaninoff	1927

Table 2. Information about the performance recordings.

for the first 31 measures in each recording. We typically spent about 50 minutes correcting and fine-tuning the alignment of each recording, though the ones by Rachmaninoff and Freire took longer, mainly due to chords played asynchronously. To determine the onset for such an arpeggiated chord (whose notes share the same score time), we deferred first to left-hand notes, and then to notes played last. We determined measure- and beat-level onsets by filtering rows as described in Section 2.3. We have published all these data, including the web links to each recording (copyrighted), the MEI score, and the alignment files at the note-, beat-, and measure-level, at `https://github.com/yucongj/alignment-data-chopin-op9-no2`.

Figure 2 shows the tempo curves at the measure level for all seven performances. We have also marked the composition structure of an AABABAC form, each section occupying exactly four measures except for Section C (coda). It is clear from the figure that these pianists chose a variety of general tempi, with Rachmaninoff playing markedly slower than the others, and Ashkenazy playing faster. In fact, some tempo curves almost never overlap with those two, e.g., Barenboim, until Section C. At a closer look, we can observe a consistent pattern where each A and B section generally starts with a bit of speeding up and ends with slowing down—we can spot the dips at the last measure of those sections (i.e., Measures 4, 8, ..., 24). Comparing repetitions of the same section, it seems later A sections are played slightly slower than earlier ones, while the later B section tends to be a bit faster than the earlier one, presenting an increasing contrast between these two alternating sections.

As an example of examining beat-level timing, we compared how three pianists played the ending measure of Section A (repeating at Measures 4, 8, 16, and 24), illustrated in Figure 3. All four places share the same melody, as first seen in the third system of the score in Figure 1, but with more ornaments over time. At first glance, all three pianists created a "tempo dip" around Beat 6, and a "tempo dome" peaking around Beat 9. Ashkenazy shows a deeper dip each time this melody is repeated, expressing an increasing sense of drama; the first two times played by Freire and the first time by Rachmaninoff don't present clear dips at all until later repeats. In addition, Freire and Ashkenazy's dips are always at Beat 6, while Rachmaninoff has two at Beat 5 and only one at Beat 6.

4. USER INTERVIEWS AND EVALUATIONS

To evaluate this version of *Piano Precision* and gather insights for future development, we conducted individual user studies with 15 potential users. We recruited mainly music scholars who conduct research in performance analysis (sometimes called performance studies), as well as two music psychologists who analyze performances. The interviews were conducted over Zoom and the procedure consisted of three parts: i) the participants describe their research background related to performance analysis, including the kinds of research questions they investigate and any digital tools they use to assist research; ii) via screen-sharing, the interviewer demonstrates the key functionalities of *Piano Precision*, and then the participants use the software (pre-installed on their own computers) to generate precise alignment of a performance excerpt; iii) the participants share their user experience, discussing any aspects of the software they find easy or difficult to use, any additional features they wish to have, and ways they envision that this tool could assist research. The interviews were semi-structured; for example, if participants mentioned having used SV, they would be asked to share their levels of experience and their workflow. A typical interview lasted 60 to 90 minutes.

4.1 Background of Participants

The 15 participants spanned diverse demographic backgrounds and research experiences. Their experience in relevant performance analysis research ranged from 3 to 40 years (mean=11.5, std=9.4); their ages ranged from 29 to 60 (mean=44.4, std=10.1); six were female (40%); 11 were from Europe, three from Asia, and one from North America. Four had no previous SV experience, one of whom was not familiar with spectrograms; of the rest, seven self-reported high levels of SV experience, three medium, and one low. Two had previously built their own digital tools for such research.

All but three mentioned their research involved analyzing tempo or timing; seven said that was their primary focus. Other aspects mentioned include dynamics, articulations, timbre, texts (lyrics), and balance (of instruments). The kinds of music they analyze are mainly piano, orchestral, and vocal music, almost equally distributed, and the vocal music sometimes involves orchestral or piano accompaniment as well.

4.2 Typical Workflow of Music Scholars

Among the 11 SV users, only two use onset-detector plugins for annotating onsets, and eight manually "tap" onsets into SV while listening to the recording (the other one uses SV solely for close listening). The shared sentiment (including the two plugin users) is that existing onset detectors make many mistakes which require a lot of effort to correct; in many cases, manual tapping is relatively easier. When tapping, they get into the flow of the music and tap at the measure, half-measure, or beat level. At some complex places where it's more difficult to predict timing, e.g., a change of rhythmic pattern, they often need to check

ISBN 978-3-200-10642-0

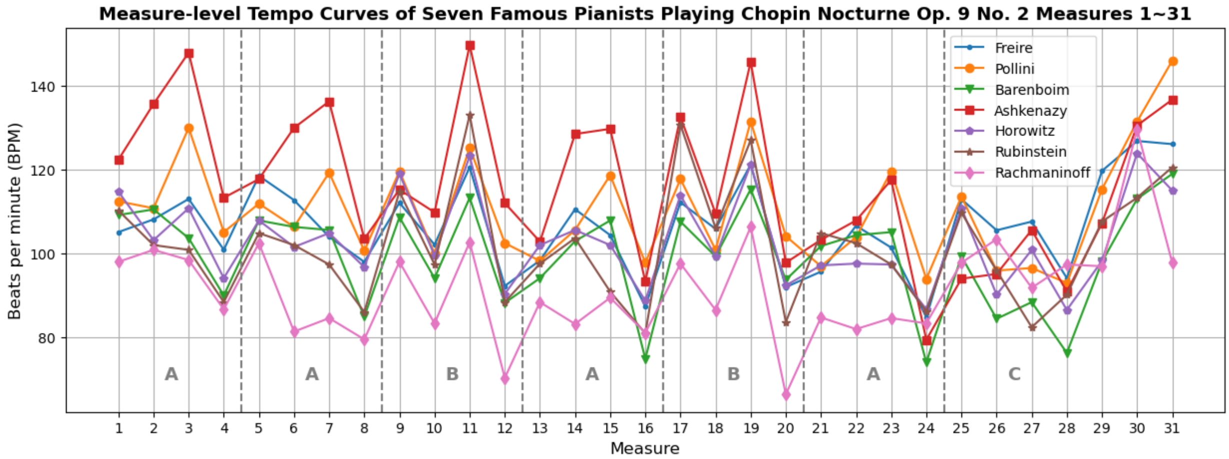

Figure 2. Comparing measure-level tempo curves. The composition structure is indicated by the vertical dashed lines.

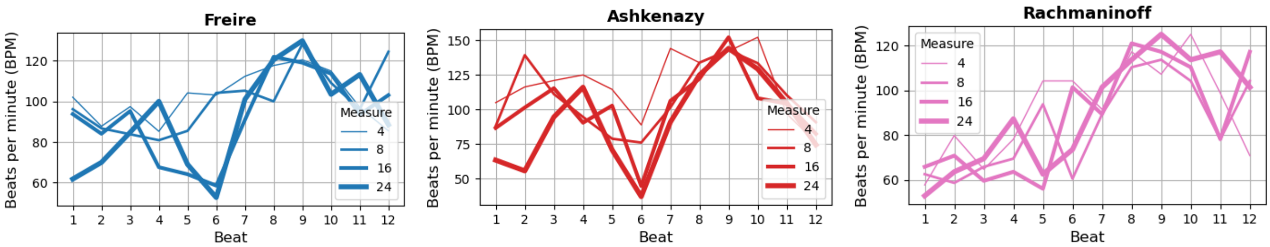

Figure 3. Comparing beat-level tempo curves at four different endings of Section A by three pianists.

back-and-forth with the score (either printed or digitally on a separate screen). After double checking and fine-tuning the entered taps, they export these annotations to a CSV file, and label each onset there, e.g., with a measure number.

4.3 Task for Participants

To give the participants real experience using the software, they were asked to load a provided score and a recording, align them, verify and correct an excerpt at the note level, and export the alignment file. The excerpt was the first 26 measures in *Clara Schumann Romanze in A minor*, containing 144 note onsets in total. Figure 4 shows the accuracy of our built-in aligner for this excerpt (vs. a timestamp-matched MIDI performance file as ground truth). The participants were also encouraged to pause at any time to share thoughts in the moment about the software (which might be lost if waiting until after completing the task). A common question from the participants was how precisely they should correct the alignment; for the sake of time, they were prompted to correct only obvious errors without fine-tuning, although this still could be subjective. Typically it took participants around 13 minutes to correct the alignments. Unsurprisingly, expert SV users tend to complete this task faster. The one participant who was not previously familiar with spectrograms still managed to complete this task (in 26 minutes), with guidance.

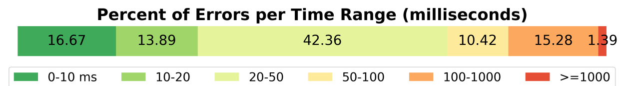

Figure 4. Alignment accuracy (error distributions) of the 144 note onsets in the first 26 measures of *Clara Schumann Romanze in A minor*.

4.4 User Evaluation and Feedback

12 out of the 15 participants expressed that they found this tool easy to use in general, including two participants who had no previous SV experience (while the other two expressed interests in learning more about SV). Helpful features they mentioned include visual connection between the score and the onsets, jumping to playback places by clicking on the score, auto-generated tempo curves, and other features such as spectrograms, onset clipping sound, and slowing down playback. As for aspects to improve, six participants mentioned that the label format for onsets was not intuitive, though some said they could get used to it. Five mentioned that they would like to be able to easily add or delete onsets, e.g., for additional onsets in rolled chords, common in historical performances. Grace notes are currently omitted in *Piano Precision*, and four participants said they would like to have the option to include them. With reference to MIR, two expressed hopes for improved alignment accuracy to lower the effort of manual correction, and two hoped for alignment plugins optimized for various instruments or kinds of music. Two

ISBN 978-3-200-10642-0

would like to annotate and visualize aspects beyond timing (e.g., dynamics and pedaling). Two would like to be able to stack multiple recordings of the same piece and align them not only to each other (as in Sonic Lineup), but also to the score (a feature the efficacy of which is demonstrated in [40]).

When asked about this tool's value, all participants confirmed that it could assist research in performance analysis, and three added it could also be useful for teaching. Eight participants said (unsolicited) that they would personally use *Piano Precision*, five of whom wished it was available for a past project. All but four elaborated at least one of three common reasons why this tool would benefit their annotation process. First, having the correct amount of onsets from the start makes it easier to check errors comparing to tapping or correcting results from onset detectors. Second, this tool is especially valuable for annotating micro-timing (e.g., studying degrees of rubato), complex notations (e.g., in post-tonal music), or unfamiliar scores; various visual features connecting onsets with the score tremendously help with such challenging cases. Third, the automatically generated labels in alignment files save effort for manually entering labels—especially labels for micro-timing—and are flexible for filtering, e.g., beat-level onsets. In addition, five participants appreciated the feature of jumping to an onset time by clicking on the score, which helps, e.g., listening to and comparing recurring themes at different places. Three participants also recognized this tool unites the two ways people engage with music: music as written and music as performed.

Two participants admitted that they would probably stick with their old workflow: one was concerned about this tool's suitability for orchestral music (but said it might work with piano reductions), and the other mainly preferred familiarity (but shared that other researchers, especially early-career, should try different tools to find their own workflow). Other concerns from participants include availabilities of scores, and alignment accuracy for more complex music.

5. DISCUSSION

User responses to *Piano Precision* are very positive overall. We confirmed via the interviews that annotating performance timing is vital to many music scholars, and that they need and want more advanced tools to streamline such workflows. As we hoped and anticipated, prospective users recognize multiple benefits of integrating score information and connecting it with performances. Combining these capabilities presents a tremendous opportunity to bring MIR advancements to music scholarship. When an analytical tool like *Piano Precision* is able to combine the representation of a score with its temporal connection to a performance, the rich symbolic information embedded in the score (e.g., pitches, rhythms, and directions) offers features otherwise absent from the recording. These features, in turn, can facilitate automated annotation of other aspects of a performance, such as articulation, dynamics, and more.

As this is a project in ongoing development, we value feedback and suggestions from both the performance analysis and the MIR communities—new functionalities are being added to future versions of the software, and we are in touch with developers of relevant digital tools, exploring possibilities of forming a software ecosystem.

Any such tool will still have limitations: most notably, it works only with notated music. Nonetheless, *Piano Precision* marks an important step to help draw new and illuminating connections between musical scores and performances, and we envision that it can also help bridge disciplinary divides between MIR and performance research.

Acknowledgments

First, we thank the user study participants for their enthusiasm and insightful feedback. Second, we thank Chris Cannam for his invaluable technical support. Last but not least, we thank our sponsors: the University of Richmond, EU's MSCA program, and the National Endowment for the Humanities.

6. REFERENCES

[1] A. Sung and D. Fabian, "Variety in performance: A comparative analysis of recorded performances of Bach's sixth suite for solo cello from 1961 to 1998," *Empirical Musicology Review*, 2011.

[2] D. Leech-Wilkinson, "Performance style in Elena Gerhardt's Schubert song recordings," *Musicae Scientiae*, vol. 14, no. 2, pp. 57–84, 2010.

[3] J. Yang, Y. Huang, and W. Everett, "Callas as violetta: A computer-assisted study on her recorded performances of la traviata," in *International Computer Music Conference (ICMC)*, 2017.

[4] F. Thalmann, E. Nakamura, and K. Yoshii, "Tracking the evolution of a band's live performances over decades," in *International Society for Music Information Retrieval (ISMIR)*, 2022, pp. 850–857.

[5] A. Lerch, C. Arthur, A. Pati, and S. Gururani, "An interdisciplinary review of music performance analysis," *Transactions of the International Society for Music Information Retrieval (TISMIR)*, Nov 2020.

[6] S. Davis, "Bring out the counterpoint: exploring the relationship between implied polyphony and rubato in Bach's solo violin music," *Psychology of Music*, vol. 37, no. 3, pp. 301–324, 2009.

[7] C. Cannam, C. Landone, and M. Sandler, "Sonic Visualiser: An open source application for viewing, analysing, and annotating music audio files," in *Proceedings of the 18th ACM international conference on Multimedia*, 2010.

[8] N. Cook and D. Leech-Wilkinson, "A musicologist's guide to Sonic Visualiser," *Centre for the History and Analysis of Recorded Music London, UK*, 2009.

ISBN 978-3-200-10642-0

[9] M. R. Thompson, "Sonic Visualiser: Visualisation, analysis, and annotation of music audio recordings," *Journal of the American Musicological Society*, vol. 74, no. 3, 2021.

[10] M. Mauch and S. Dixon, "pYIN: A fundamental frequency estimator using probabilistic threshold distributions," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 659–663.

[11] S. Dixon, "Evaluation of the audio beat tracking system BeatRoot," *Journal of New Music Research*, vol. 36, no. 1, pp. 39–50, 2007.

[12] "The Expressive Means Vamp plugin website," https://github.com/cannam/expressive-means.

[13] T. Wozonig, "Analysis of temporal phenomena in music with Sonic Visualiser," in *Softwaregestützte Interpretationsforschung: Grundsätze, Desiderate und Grenzen*. Königshausen & Neumann, 2023, pp. 73–98.

[14] N. Schüler, "Toward a higher accuracy of note onset detection," in *Softwaregestützte Interpretationsforschung: Grundsätze, Desiderate und Grenzen*. Königshausen & Neumann, 2023, pp. 419–424.

[15] N. Orio and D. Schwarz, "Alignment of monophonic and polyphonic music to a score," in *International Computer Music Conference (ICMC)*, 2001.

[16] C. Raphael, "A hybrid graphical model for aligning polyphonic audio with musical scores," in *International Society for Music Information Retrieval (ISMIR)*, 2004.

[17] M. Müller, F. Kurth, and T. Röder, "Towards an efficient algorithm for automatic score-to-audio synchronization," in *International Society for Music Information Retrieval (ISMIR)*, 2004.

[18] C. Joder, S. Essid, and G. Richard, "A comparative study of tonal acoustic features for a symbolic level music-to-score alignment," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2010, pp. 409–412.

[19] R. Agrawal, D. Wolff, and S. Dixon, "Structure-aware audio-to-score alignment using progressively dilated convolutional neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 571–575.

[20] M. Müller, Y. Özer, M. Krause, T. Prätzlich, and J. Driedger, "Sync Toolbox: A Python package for efficient, robust, and accurate music synchronization," *Journal of Open Source Software*, vol. 6, no. 64, 2021.

[21] A. Morsi and X. Serra, "Bottlenecks and solutions for audio to score alignment research," in *International Society for Music Information Retrieval (ISMIR)*, 2022.

[22] H. Gottschewski, "Graphic analysis of recorded interpretations," *Computing in Musicology*, vol. 8, pp. 93–96, 1992.

[23] J. Yang, "The comprehensive application of Vmus.net for musical performance studies," in *Proceedings of the 4th International Workshop on Digital Libraries for Musicology*, 2017.

[24] "Sonic Lineup," https://www.sonicvisualiser.org/sonic-lineup.

[25] D. M. Weigl, C. VanderHart, D. Rammler, M. Pescoller, and W. Goebl, "Listen Here! a web-native digital musicology environment for machine-assisted close listening," in *Proceedings of the 10th International Conference on Digital Libraries for Musicology*, 2023.

[26] "Tony," https://www.sonicvisualiser.org/tony.

[27] A. Berndt, "The music performance markup format and ecosystem," in *International Society for Music Information Retrieval (ISMIR)*, 2021, pp. 50–57.

[28] S. D. Peter, C. E. Cancino-Chacón, F. Foscarin, A. P. McLeod, F. Henkel, E. Karystinaios, and G. Widmer, "Automatic note-level score-to-performance alignments in the ASAP dataset," *Transactions of the International Society for Music Information Retrieval (TISMIR)*, 2023.

[29] E. Nakamura, K. Yoshii, and H. Katayose, "Performance error detection and post-processing for fast and accurate symbolic music alignment." in *International Society for Music Information Retrieval (ISMIR)*. Suzhou, 2017, pp. 347–353.

[30] "The music encoding initiative website," https://music-encoding.org.

[31] A. Hankinson, P. Roland, and I. Fujinaga, "The music encoding initiative as a document-encoding framework," in *International Society for Music Information Retrieval (ISMIR)*, 2011, pp. 293–298.

[32] T. Crawford and R. Lewis, "Music encoding initiative (digital and multimedia scholarship)," *Journal of the American Musicological Society*, vol. 69, no. 1, pp. 273–285, 2016.

[33] L. X. Pugin, R. Zitellini, and P. Roland, "Verovio: A library for engraving MEI music notation into SVG," in *International Society for Music Information Retrieval (ISMIR)*, 2014, pp. 107–112.

[34] L. Pugin, "Interaction perspectives for music notation applications," in *Proceedings of the 1st International Workshop on Semantic Applications for Audio and Music*, 2018, pp. 54–58.

[35] "Movi: The digital Mozart score viewer," https://dme.mozarteum.at/movi/en.

[36] "The MuseScore website," https://musescore.org.

ISBN 978-3-200-10642-0

[37] W. Goebl and D. M. Weigl, "mei-friend: An interactive web-based editor for digital music encodings," *Journal of Open Source Software*, vol. 9, no. 98, p. 6002, 2024.

[38] "The Vamp plugin website," https://vamp-plugins.org.

[39] C. Raphael, "Automatic segmentation of acoustic musical signals using hidden Markov models," *IEEE transactions on pattern analysis and machine intelligence*, vol. 21, no. 4, pp. 360–370, 1999.

[40] Y. Jiang, "Performance precision: a software prototype for computer-assisted annotation and analysis of music performance," in *International Computer Music Conference (ICMC)*, 2025.

ISBN 978-3-200-10642-0