PIANO CONCERTO ACCOMPANIMENT CREATION

Yigitcan Özer **Simon Schwär** **Meinard Müller**
International Audio Laboratories Erlangen, Germany
{yigitcan.oezer, simon.schwaer, meinard.mueller}@audiolabs-erlangen.de

ABSTRACT

Creating orchestral accompaniments for solo piano performances of concertos is challenging, especially when aiming for a system that adapts to the pianist's tempo and dynamics. In this work, we present a semi-automatic offline approach using source separation, alignment, and time-scale modification techniques to generate orchestral accompaniments from public-domain recordings. Our approach separates the orchestral part, aligns it with a new solo-piano performance played by the user, and adjusts the tempo and dynamics to match the user's interpretation. While some processes are automated, significant manual adjustments are needed for temporal alignment and sound production. Our results demonstrate the feasibility of this method while highlighting the need for further automation. We have made the multi-track recordings of the created mixes publicly available for research purposes at `https://www.audiolabs-erlangen.de/resources/MIR/PCD_AudioLabs`.

1. INTRODUCTION

Building real-time computer accompaniment systems that can analyze and synchronize with a performer's tempo and dynamics has long been a dream and a complex challenge in music technology and music information retrieval (MIR) [1–4]. Unlike pre-recorded backing tracks, such as those offered by music publishers, [1] an automated accompaniment system aims to interactively follow and adapt to a performer's interpretation. In an ideal scenario, it would replicate the real-time acoustic and visual interaction between musicians, ensuring synchronization and cohesion, similar to the communication between soloist and orchestra during a live performance [5, 6].

This paper presents a semi-automatic offline approach to creating orchestral accompaniments for piano concertos, a significant genre in Western classical music. Given that only elite pianists typically have the opportunity to perform with orchestras, there is a need for accessible orchestral accompaniments for pianists of all levels. Using music source separation (MSS) techniques, we extract orchestral tracks from public-domain recordings of piano concertos and align them with new solo piano performances. While aiming for a fully adaptive system, our semi-automatic approach falls somewhere between static play-along tracks and real-time interactive accompaniment. In Section 2, we describe the pipeline of our offline approach and then discuss experiments in Section 3. We applied this method to two movements of famous piano concertos and made the multi-track recordings publicly available for further research and development. [2]

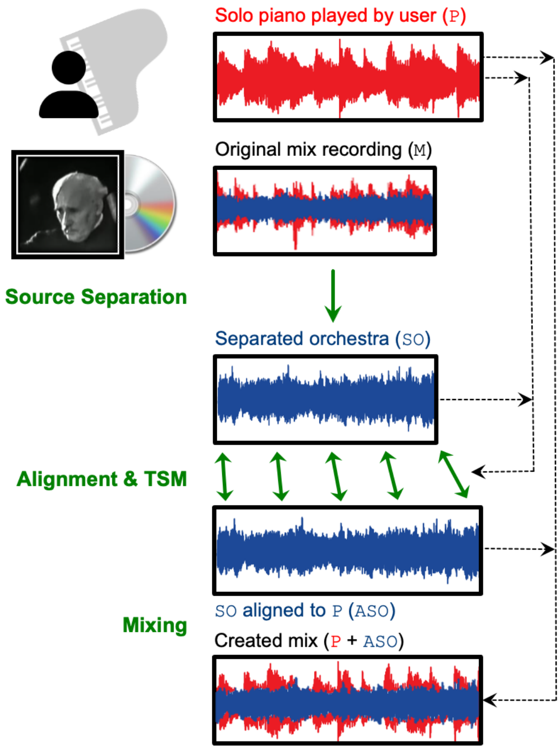

Figure 1. Processing steps for creating a piano concerto accompaniment.

2. PROCESSING STEP

For a lonely solo pianist performing the piano part of a concerto, our vision is to create an orchestral accompaniment that automatically adapts to the pianist's interpretation. In this work, similar to [7, 8], we experiment with a semi-automatic offline approach, utilizing computational

[1] e.g., Music Minus One, `https://www.halleonard.com/series/MMONE`

[2] `https://www.audiolabs-erlangen.de/resources/MIR/PCD_AudioLabs`

tools from signal processing and machine learning. The main steps of our approach, illustrated in Figure 1, can be summarized as follows:

1. **Pianist's Performance:** The pianist freely performs and records a piano track (possibly with multiple takes and subsequent editing), creating a personal interpretation (referred to as `P`).
2. **Source Separation:** We select an existing recording of a piano concerto and apply source separation techniques [9] to decompose the recording into a piano track (referred to as `SP`) and an orchestral track (referred to as `SO`). Note that in our scenario, only `SO` is used.
3. **Temporal Alignment:** The separated orchestral track `SO` is then temporally aligned with the pianist's recording `P` using alignment [10, 11] and beat-tracking techniques [12].
4. **Time-Scale Modification:** Based on the alignment, we apply time-scale modification techniques to warp the `SO` track, resulting in `ASO`, which temporally synchronizes with the `P` track [13, 14].
5. **Dynamic Adjustment and Mixing:** Finally, we adjust the dynamics of the `ASO` track and mix it with the `P` track, creating an orchestral accompaniment tailored to the pianist's interpretation.

It is important to note that while some of these steps can be assisted by automated methods, much of the process required careful manual fine-tuning to achieve musically and acoustically acceptable results. We will further explore these issues in the subsequent section.

3. EXPERIMENTS

In our experiments, we considered two movements from famous piano concertos, with the piano tracks performed by two non-professional pianists (both co-authors of this work), and the orchestral tracks derived through source separation from older, public-domain piano concerto recordings. Specifically, we used the following data (see also our website):

- **Beethoven Piano Concerto No. 1** in C Major, Op. 15, first movement. Original mix recordings (1958): Symphony of the Air Orchestra conducted by Josef Krips with pianist Arthur Rubinstein. Lonely pianist (user): Meinard Müller.
- **Mozart Piano Concerto No. 20** in D Minor, KV 466, first movement. Original mix recordings (1960): Orchestre des Concerts Lamoureux conducted by Igor Markevitch with pianist Clara Haskil. Lonely pianist (user): Yigitcan Özer.

The solo piano recordings were created following the conventions of a traditional classical music production, where the lonely pianists freely performed the piano parts in several takes. In a post-processing step, the best segments were then combined into a coherent single performance, balancing technical and musical aspects during the editing. This recording process was completely independent of any orchestral accompaniment and was primarily coordinated by Simon Schwär, the sound engineer in charge. The resulting performance (`P`) served as the input to the overall processing pipeline as described above. However, several manual interventions were required to optimize the final mix, as detailed in the following:

- The quality of source separation heavily depends on the selected pieces, recordings, and MSS approaches [15]. By experimenting with several MSS methods,[3] we achieved comparatively good separation quality in the two mix recordings used in this paper. Furthermore, in these cases, separation artifacts were often masked in the final mix. Additionally, we applied source separation only where the orchestra and piano played together, using the original mix for orchestra-only sections and muting the orchestra during solo piano passages.
- Good temporal alignment between the orchestra and piano is crucial for the final musical quality of the generated mix. Starting with a measure-wise alignment, we refined this by setting anchor points at musically and rhythmically critical time points, aiming to strike a balance between precision and natural tempo fluctuations. The process of fine-tuning the timing between critical onsets in both `P` and `ASO` was iterated through careful listening and further corrections in the mixing stage.
- Finally, we created a mix from `P` and `ASO` in a digital audio workstation. At this stage, we applied dynamic range compression to `ASO` and manually adjusted the relative levels for cohesive dynamics. Furthermore, we carefully applied artificial reverb to match the spatial impression of `P` and `ASO`.

The music recordings on our website result from a combination of recent automated methods and manual refinements using traditional audio production techniques. Obviously, any musical issues in the piano parts are due to the limitations of the non-professional pianists.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a semi-automatic offline approach for creating orchestral accompaniments for solo pianists performing a piano concerto. We have made all data, including multitrack recordings of the final mixes, publicly available, hoping this constitutes a valuable resource for various MIR tasks, including source separation and performance analysis. Achieving the desired quality required significant manual tweaking. Future work could focus on further automating the process by refining alignment and synchronization, improving dynamic adjustments, acoustic coherence between tracks (see, e.g., [16]), and enhancing the overall mix quality. Increased automation could reduce the time and effort needed to produce high-quality accompaniments, making this technology more accessible to a broader range of users, from students to professional musicians.

[3] In our experiments, *AudioShake* (`https://www.audioshake.ai/`) turned out to be the best model for our purposes.

Acknowledgements: This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Grant No. 521420645 (MU 2686/17-1) and Grant No. 328416299 (MU 2686/10-2). The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institute for Integrated Circuits IIS.

5. REFERENCES

[1] R. B. Dannenberg, "An on-line algorithm for real-time accompaniment," in *Proceedings of the International Computer Music Conference (ICMC)*, Paris, France, 1984, pp. 193–198.

[2] B. Vercoe, "The synthetic performer in the context of live performance," in *Proceedings of the International Computer Music Conference (ICMC)*, 1984, pp. 199–200.

[3] R. B. Dannenberg and C. Raphael, "Music score alignment and computer accompaniment," *Communications of the ACM, Special Issue: Music Information Retrieval*, vol. 49, no. 8, pp. 38–43, 2006.

[4] C. E. C. Chacón, S. Peter, P. Hu, E. Karystinaios, F. Henkel, F. Foscarin, N. Varga, and G. Widmer, "The ACCompanion: Combining reactivity, robustness, and musical expressivity in an automatic piano accompanist," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Macao, China, 2023, pp. 5779–5787.

[5] J. W. Davidson, "Visual perception of performance manner in the movements of solo musicians," *Psychology of Music*, vol. 21, no. 2, pp. 103–113, 1993. [Online]. Available: https://doi.org/10.1177/030573569302100201

[6] W. Goebl and C. Palmer, "Synchronization of timing and motion among performing musicians," *Music Perception: An Interdisciplinary Journal*, vol. 26, no. 5, pp. 427–438, 2009. [Online]. Available: http://www.jstor.org/stable/10.1525/mp.2009.26.5.427

[7] C. Raphael and Y. Gu, "Orchestral accompaniment for a reproducing piano," in *Proceedings of the International Computer Music Conference (ICMC)*, Montreal, Quebec, Canada, 2009, pp. 501–504.

[8] T. Tsai, S. K. Tjoa, and M. Müller, "Make your own accompaniment: Adapting full-mix recordings to match solo-only user recordings," in *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, Suzhou, China, 2017, pp. 79–86.

[9] Y. Özer and M. Müller, "Source separation of piano concertos using musically-motivated augmentation techniques," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 32, pp. 1214–1225, 2024.

[10] S. Dixon and G. Widmer, "MATCH: A music alignment tool chest," in *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, London, UK, 2005, pp. 492–497.

[11] 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 (JOSS)*, vol. 6, no. 64, pp. 3434:1–4, 2021.

[12] M. E. P. Davies, P. M. Brossier, and M. D. Plumbley, "Beat tracking towards automatic musical accompaniment," in *Proceedings of the AES International Conference on Semantic Audio*, Barcelona, Spain, 2005. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=13124

[13] J. Driedger, M. Müller, and S. Ewert, "Improving time-scale modification of music signals using harmonic–percussive separation," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 105–109, 2014.

[14] S. Rosenzweig, S. Schwär, J. Driedger, and M. Müller, "Adaptive pitch-shifting with applications to intonation adjustment in a cappella recordings," in *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, Vienna, Austria, 2021, pp. 121–128.

[15] Y. Özer, H.-U. Berendes, V. Arifi-Müller, F. Stöter, and M. Müller, "Notewise evaluation for music source separation: A case study for separated piano tracks," in *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, San Francisco, USA, 2024.

[16] J. Im and J. Nam, "DiffRENT: A diffusion model for recording environment transfer of speech," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Seoul, South Korea, 2024, pp. 7425–7429.