Real-Time MIR Algorithms for Music-Reactive Game World Generation

PETER MEIER, Deggendorf Institute of Technology, Germany and International Audio Laboratories Erlangen, Germany

SIMON SCHWÄR, International Audio Laboratories Erlangen, Germany

SEBASTIAN ROSENZWEIG, International Audio Laboratories Erlangen, Germany

MEINARD MÜLLER, International Audio Laboratories Erlangen, Germany

In music education, an important learning objective is to acquire a sense of rhythm and melody, e.g., by tapping or singing along to a musical performance. In this context, music games are becoming increasingly popular due to their accessibility and interactivity. However, existing games are often tailored to a specific musical repertoire. In this paper, we explore the potential of music information retrieval (MIR) algorithms for generating dynamic game worlds that are reactive to musical input in real-time, e.g., the music from a stereo system or a live performance. Such a game may be used to explore musical performances in an interactive and playful fashion. As a case study, we develop a jump-and-run game prototype that challenges players to tap along the beat of the music in their environment. Its music-reactive game world relies on past and predicted beat positions estimated by a real-time beat tracking algorithm.

Additional Key Words and Phrases: Game, MIR, real-time, beat tracking

1 INTRODUCTION

Musical education lays the foundation for music literacy already at an early age. With basic tasks like clapping along to the beat or singing simple melodies, children are encouraged to actively experience musical performances. However, when practicing at home without direct feedback from a teacher, it is sometimes difficult to stay motivated. Music games have become a promising complement to music lessons since they constitute an accessible and playful way to learn. Indeed, educational music games have shown a positive effect on the musical skills of their players [8]. Such games often use algorithms from music information retrieval (MIR) to analyze the user input. For example, MIR algorithms for beat tracking and fundamental frequency (F0) estimation have been used in rhythm training [1, 6, 10] or singing games [5, 9], respectively. However, except for an experimental study [4], most games rely on pre-generated game worlds tailored to a fixed musical repertoire. In this paper, we explore the potential of real-time MIR algorithms for dynamically generating game worlds. The underlying idea is that the game world reacts to the music in the player's environment, may it be their favorite song on the radio or a live performance. In this way, the game enables players to interact with their musical environment in an interactive and playful way while improving their musical skills. In an experimental case study, we developed a jump-and-run game prototype for rhythm training called "Rock Your Beats" where the player needs to tap along the estimated beat of a musical performance. The game uses an existing real-time beat tracking algorithm [7] that estimates the current beat of the recorded music and provides a prediction of future beat positions. By visualizing a buffer that contains past, current, and future beat positions, our game world provides orientation and guidance to players for tapping along with the beat.

In the remainder of this paper, we will first explain the real-time beat tracking algorithm (Section 2) and subsequently describe our game prototype (Section 3). Finally, we summarize our findings and outline future work (Section 4).

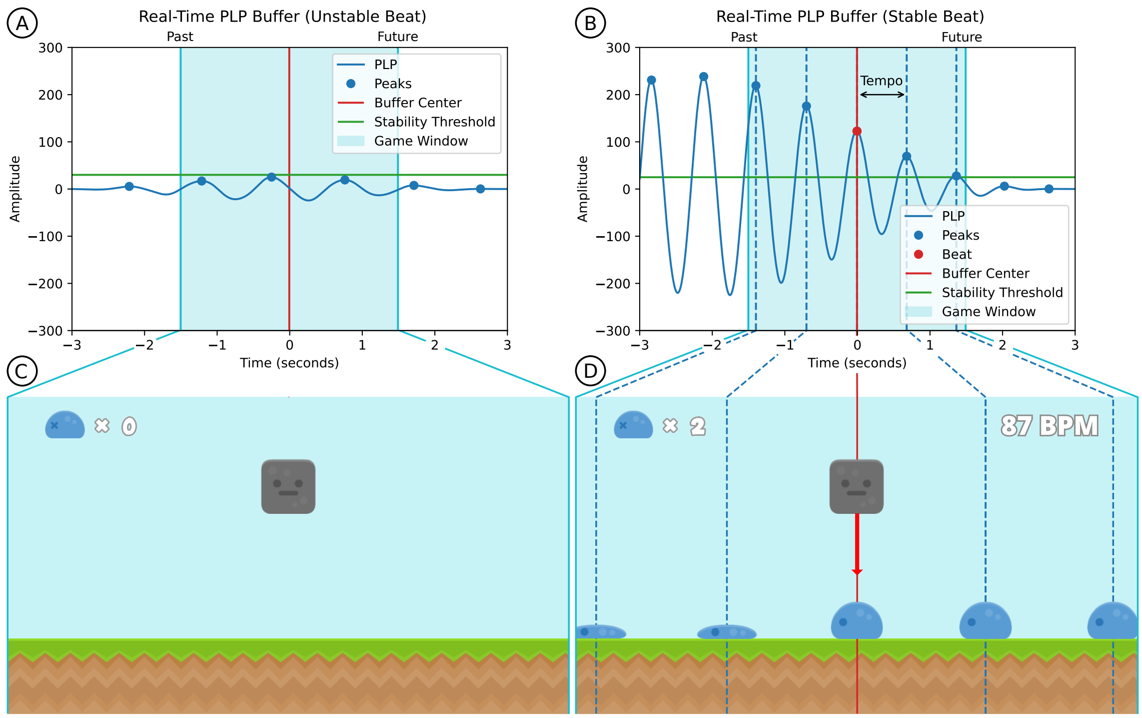

Fig. 1. Real-Time PLP buffer and world generation of the game "Rock Your Beats." **(A)** Real-Time PLP buffer with an unstable beat. **(B)** Real-Time PLP buffer with a stable beat. **(C)** Empty game world with a rock (black squared element). **(D)** Game world with a rock (black squared element) hitting moving beat creatures (blue rounded elements).

2 REAL-TIME BEAT TRACKING

Beat tracking is a long-studied task in the field of MIR [2]. In the following, we focus on the approach based on predominant local pulse (PLP), which was first introduced by Grosche et al. [3]. PLP is a method for computing local pulse information from music recordings, assuming that the beat positions correspond to the note onsets of the music. To this end, the PLP algorithm analyzes the musical input signal to estimate possible note onset candidates encoded by a novelty function that measures spectral changes over time. Note onsets typically go along with beat positions that are periodically spaced when there is a prominent rhythm in the music. To account for this, one compares local sections of the novelty function with windowed sinusoidal kernels that optimally match the local tempo structure of the music signal. These optimal pulse kernels get overlap-added over time to form a continuous PLP function, where each peak represents a beat candidate. A stable tempo results in constructive interference between successive kernels, resulting in higher peaks in the PLP function. A change in tempo results in a varying distance between the peaks of the PLP function.

This PLP algorithm has been adapted to support real-time beat tracking [7], which forms the basis for this work. In the real-time case, the PLP function is calculated in a buffer with a total size of about 4-8 seconds that is constantly updated to represent the current beat structure of the music (see Fig. 1A and B). The center of the PLP buffer is always aligned with the current point in time (t = 0 seconds). The left half of the buffer contains the pulse structure of the last three seconds based on the current data. The right half of the buffer extrapolates the pulse structure to the following three seconds, thus predicting future beat positions. The predictions are particularly helpful for generating music-reactive game worlds, as we demonstrate in Section 3. As time progresses, the data within the buffer flows from the right (future) to the left (past) and is shifted each time the buffer is updated at a given frame rate. The PLP buffer can be used to derive additional information about the beat and tempo of the analyzed music signal:

- The primary task of a real-time beat tracker is to determine whether or not there is a beat detected at the current time frame. As soon as a peak intersects the read position (in our example the buffer center illustrated by the red line in Fig. 1B), a beat is detected. Note that latency in both signal processing and user interaction can be problematic in such a time-critical application. To compensate for latency, one may shift the read position a few frames to the right from the buffer center to detect beats earlier.
- A tempo estimate can be derived from the period of the PLP function, as shown by the black arrow in Fig. 1B. In particular, the current tempo is the inverse of the time interval between the current beat and the next peak position.
- The PLP buffer also gives information about the beat stability of the music. Low amplitude values of the peaks in the PLP function represent an unstable beat (Fig. 1A), high amplitude values represent a stable beat structure (Fig. 1B). For further details, we refer to [3, 7].

3 MUSIC-REACTIVE GAME WORLD GENERATION

In the following, we show how the real-time beat tracking algorithm from Section 2 can be used for generating music-reactive game worlds. To this end, we have developed a prototype jump-and-run game named "Rock Your Beats" which is illustrated in Fig. 1C and D. The goal for the player is to tap along to the beat of music in the environment (e.g., by pressing a button on a touch screen or keyboard) and thereby hitting moving "beat creatures" with a dropping rock located in the center of the game world. Each beat creature represents one beat, and one point is awarded for each hit.

The game world is generated on the fly from music in the player's environment by using a microphone signal as the input of the real-time PLP algorithm as described in Section 2. The center of the game world (the rock) represents the current point in time and thus relates to the center of the PLP buffer. The beat creatures are placed at the peak positions of the PLP function and therefore move from the right (future) to the left (past), crossing the player position in sync with the beat of the music. If no stable beat structure is detected in the input signal, the peaks of the PLP function have a rather low amplitude, as illustrated in Fig. 1A. If this amplitude is below a certain stability threshold, no beat creatures are created in the game world (see Fig. 1C). This way, beat creatures only appear when music with a stable beat is playing in the player's environment. The time range that is visually represented in the game (the *game window*) can have maximally the length of the PLP buffer but can be chosen smaller to avoid the visualization of inaccurate beat predictions in the distant future. The size of the game window also determines how much visual support the player gets, and makes it possible to adapt to smaller devices such as smartphones.

Real-time game world generation with MIR algorithms has several advantages. First of all, it encourages players to think about the music in their surroundings, which to some degree counteracts the cliché of disconnecting from the real world by playing computer games. If there is no music playing around the player, there are no points to be awarded. Furthermore, through the displayed future beat positions, the player not only has to feel the beat but also can see it through the game. This visual guidance could be interesting especially in a music education setting so that children can playfully learn how to interact with the beats of their favorite music and train their rhythmic timing.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we showed how real-time MIR algorithms can be used for generating music-reactive game worlds with the example of a jump-and-run style game for rhythm training based on beat tracking. The real-time approach allows players to explore the beat of the music in their environment interactively. In future work, we plan to expand our prototype by employing other MIR algorithms that enable further adaptation of the game world to the musical contents, e.g., by recognizing the melody or using real-time F0 estimation. In this way, more advanced game mechanics are conceivable, where a player may have to sing one of the pitches in the currently sounding musical scale to reach a goal. Such a game would further encourage creative engagement with the musical surroundings, being adaptable to different skill levels and accounting for different musical dimensions.

ACKNOWLEDGMENTS

This work was supported by the BayWISS Joint Academic Partnership “Digitalisation.” The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institut für Integrierte Schaltungen IIS.

REFERENCES

[1] José Jaime Ariza, Isabel Barbancho, Cristina Urdiales, Lorenzo J. Tardón, and Ana M. Barbancho. 2016. A new multiformat rhythm game for music tutoring. *Multimedia Tools and Applications* 75, 8 (2016), 4349–4362. https://doi.org/10.1007/s11042-015-2478-8

[2] Simon Dixon. 2007. Evaluation of the audio beat tracking system BeatRoot. *Journal of New Music Research* 36 (2007), 39–50.

[3] Peter Grosche and Meinard Müller. 2011. Extracting Predominant Local Pulse Information from Music Recordings. *IEEE Transactions on Audio, Speech, and Language Processing* 19, 6 (2011), 1688–1701. https://doi.org/10.1109/TASL.2010.2096216

[4] Christian M. Hahn and Paul J. Diefenbach. 2010. Surge: an experiment in real-time music analysis for gaming. In *ACM SIGGRAPH ASIA 2010 Posters*. ACM, Seoul, Republic of Korea, 16:1. https://doi.org/10.1145/1900354.1900371

[5] Rubén Jesús García Hernández, Isabel Barbancho Pérez, Lorenzo José Tardón García, Jon Arrambarri, Milán Magdics, and Mateu Sbert. 2014. E-cecilia: implementation of a music game. In *Proceedings of the Congreso de la Sociedad Española para las Ciencias del Videojuego (CoSECivi)*, Vol. 1196. Barcelona, Spain, 66–77. http://ceur-ws.org/Vol-1196/cosecivi14_submission_12.pdf

[6] Matthias Husinsky and Patrik Lechner. 2015. Das Rhythmus-Radar. In *Mensch und Computer 2015 - Workshopband*. De Gruyter Oldenbourg, Stuttgart, Germany, 339–344. https://dl.gi.de/20.500.12116/8124

[7] Peter Meier, Gerhard Krump, and Meinard Müller. 2021. A Real-Time Beat Tracking System Based on Predominant Local Pulse Information. In *Demos and Late Breaking News of the International Society for Music Information Retrieval Conference (ISMIR)*. Online. https://archives.ismir.net/ismir2021/latebreaking/000021.pdf

[8] Matevz Pesek, Lovro Suhadolnik, Peter Savli, and Matija Marolt. 2020. The Rhythmic Dictator: Does Gamification of Rhythm Dictation Exercises Help?. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*. Montreal, Canada, 497–503. http://archives.ismir.net/ismir2020/paper/000048.pdf

[9] Voxler SAS. 2022. Let’s Sing – Karaoke Application. Retrieved 2022-05-30 from https://letssing2022.com/

[10] Tzu-Chun Yeh and Jyh-Shing Roger Jang. 2020. AutoRhythm: A Music Game With Automatic Hit-Timing Generation and Percussion Identification. *IEEE Transactions on Games* 12, 3 (2020), 291–301. https://doi.org/10.1109/TG.2019.2936033