Real-time Mapping of Periodic Dance Movements to Control Tempo in Electronic Dance Music

Lilian Jap
KTH Royal Institute of Technology
lilianj@kth.se

Andre Holzapfel
KTH Royal Institute of Technology
holzap@kth.se

ABSTRACT

Dancing in beat to the music of one's favorite DJ leads oftentimes to a powerful and euphoric experience. In this study we investigate the effect of putting a dancer in control of music playback tempo based on a real-time estimation of body rhythm and tempo manipulation of audio. A prototype was developed and tested in collaboration with users, followed by a main study where the final prototype was evaluated. A questionnaire was provided to obtain ratings regarding subjective experience, and open-ended questions were posed in order to obtain further insights for future development. Our results imply the potential for enhanced engagement and enjoyment of the music when being able to manipulate the tempo, and document important design aspects for real-time tempo control.

1. INTRODUCTION

In Electronic Dance Music (EDM), a DJ combines in advance planning and real-time decisions for the purpose of creating an intense and ecstatic dance experience. Taking such dance experience into account is a direction of high potential when finding new practices for interactive systems based on the ideas of embodied interaction [1]. Current musical/technical landscapes have shifted the focus away from the passive individual towards an active role in sound [2] with the impact of embodied interaction.

Enabling the user to find an intuitive way for controlling the sound parameters of a music playback could pave the way further for an interactive musical environment. When it comes to the mapping of music playback tempo, the idea of involving dance movements for manipulation might not be a novel exploration field of research, eg. [3] and [4] that both involved techniques of video and/or image analysis. The main contribution presented in this paper is however the study of dancers' experiences when interacting with the proposed system, an autonomous interactive system that expands the experience one has while dancing with it. This might open up possibilities for individuals wishing to physically interact with music on a personal level - without the requirement of having a musical background, or the expectation of having the actual physical ability to play a musical instrument.

In this paper, a proposed design system is presented where the following hypothesis is addressed: mapping of real-time measurements of a dancer's rhythmic movements to tempo manipulations in EDM can lead to a dance experience that compares positively to a standard playback of the same type of music. The definition of a dancer entails any possible user of the proposed system; whether the user possess the skill-sets of a dancer or not. A proof-of-concept prototype is presented, along with its design process and findings made after performing a user-study where participants got to interact with the prototype.

2. BACKGROUND

A number of studies involving experiments with music and embodiment have been conducted. In the context of musical performances, the related work appears to have a common feature; sound-producing and communicative musical gestures as an extension of the body [5]. A music performance may include gestures by those that produce sounds, and by those that perceive sounds (i.e. listeners and dancers) [6], with the larger body of research focusing on the former.

Motion sensors used with the objective of implementing different sound synthesis techniques have been implemented in various mixed interdisciplinary approaches, such as the sonification of body movements of contemporary circus artists [7], conducting [8], and music pedagogy [9]. Placing wireless motion sensors on hands and feet, it was shown that utilizing gestures as game content brings more substance to the game [10].

When it comes to hearing rhythm in music, bodily movements play an important role when developing the skill involved in rhythmic perception [11]. Audio feedback can induce more awareness as it brings deeper understanding in how the body moves [12]. Aligning rhythmic movements with rhythm in music have been used as an approach to improve the users own movement performance in the context of sports (running) as well as physical rehabilitation and health [13, 14]. The gestures of dancers, and their relation to underlying meter have been the subject of various studies employing sensor technologies [15, 16], and repetitive movements have been found to be more pronounced in hand gestures than the gestures of other body parts [15].

Marker-based infra-red motion capture (MoCap) technology provides accurate data of complex movement in a three-dimensional space [17], but is largely limited to applications in a laboratory space. Recently, Inertial Measurement Units (IMU) have been applied in movement-based interaction design [17–19]. Their miniature in size, mobile use and reported accurate essential values make them beneficial to use in more flexible contexts.

3. METHOD

In the present study, a real-time prototype was designed for estimation of periodicity in a user's body movements, and tempo manipulation of audio recordings based on these measurements. In order to establish the building blocks and parameters for the prototype, a pre-study was conducted (Section 3.1). An overview of the prototype's components is outlined in Section 3.2. The participant groups, the experimental setup, and the evaluation method based on questionnaires are described in Sections 3.3 to 3.5, respectively.

3.1 Pre-study

An initial study was conducted with three participants to establish the design of the prototype. Different placements of IMU sensors on the body were examined in terms of usability as well as pronouncement in the movements (hand wrist, ankle joint, hip and the back). One sensor was used and based on the notion that the placement should serve a practical fit.

In the participants' expressed preferences, the ankle joint and hand wrist were the preferred placements as it made the sensor less noticeable or could enable more control in the tempo manipulation. However, the clearest pronunciation of measured periodic movement in the initial study was observed in the hand wrist. This corroborates findings by Leman and Naveda [15], motivating our decision to place the sensor on the hand wrist.

Adjustments in the implementation of the prototype were made as well since high latency in combination with sudden tempo changes were encountered and a confusion from the participants was expressed. The analysis frame size for estimating the periodicity of the arm movement was set to obtain a sufficiently reactive system, while still facilitating reliable periodicity estimation. Other system parameters, such as the form of tempo changes in the audio playback were also determined in this pre-study. After testing various sensor and audio processing approaches, the system design as depicted in Figure 1 emerged.

3.2 Prototype

3.2.1 Equipment and platform

A IMU sensor from x-io Technologies Ltd [1] was used, which makes use of the Open Sound Control (OSC) protocol. This opens up compatibility with other software applications, for instance Max/MSP [2], which was applied to collect and process incoming data from the sensors. The real-time communication was performed via Wi-Fi using TP-Link AC750 travel router as a separate 5Ghz wireless network instead of the sensors internal antennae, allowing for future extensions using multiple sensors.

[1] http://x-io.co.uk/ngimu/
[2] https://cycling74.com/products/max/

3.2.2 System design

Parts of the operations were computed using JavaScript, within the Max/MSP environment. The chain of operations as depicted in Figure 1 can be described as follows:

1. OSC messages about the accelerometer data from the NGIMU sensor are received.

2. Raw accelerometer data from x-, y- and z-axis are smoothed through low-pass filtering in a Max/MSP [slide]-object, filtering with slide value $S = 10$ according to Equation 1. A given sample output y_n is equal to the previous value y_{n-1} plus the difference between the input x_n and the previous value divided by the slide value S. Given a slide value of $S = 1$, the output will therefore always equal the input. Given a slide value of $S = 10$, the output will only change 1/10th as quickly as the input [3].

$$y_n = y_{n-1} + \frac{x_n - y_{n-1}}{S} \qquad (1)$$

3. The fundamental frequency F_0 of the movements in the data stream in each axis is estimated by the [pipo.yin]-object from Mubu for Max-toolbox [4]. The object makes use of the YIN-algorithm [20], which also provide the quality factor of the detected periodicity. Several values were tested for different attributes of the yin-object in the initial study, and significant for the interaction were the sample size of frame (N), hop size (N/16), and frame-rate (sample rate of sensors). N = 100 was found to provide sufficiently accurate results while being short enough to track a speed up/speed down of the acceleration. Data streams are sliced into windowed frames of size N using the [pipo.slice]-object from the same toolbox and with the sensor's default sample send rate of 50 Hz, data streams are processed over the last 2 seconds.

4. Values of the computed quality factor - in each axis - are smoothed using a second-order moving average filter with subsets of 10 and 5 sample values, respectively.

5. The highest obtained quality factor determines which estimated F_0 to be used, i.e. a choice between x-, y-, and z-axes was made based on which expresses the most consistent periodic movement.

6. F_0 is converted into beats per minute (bpm).

7. Based on the changes in speed of the movements, the tempo control value for the audio playback will change accordingly. If detecting a speed increase/ decrease in the dancer's movement, the current bpm value of the playback will increase/decrease with a value of 3bpm. If the movements are stopped, the playing tempo will decrease as the current design of

[3] https://docs.cycling74.com/max7/maxobject/slide
[4] http://forumnet.ircam.fr/product/mubu-en/

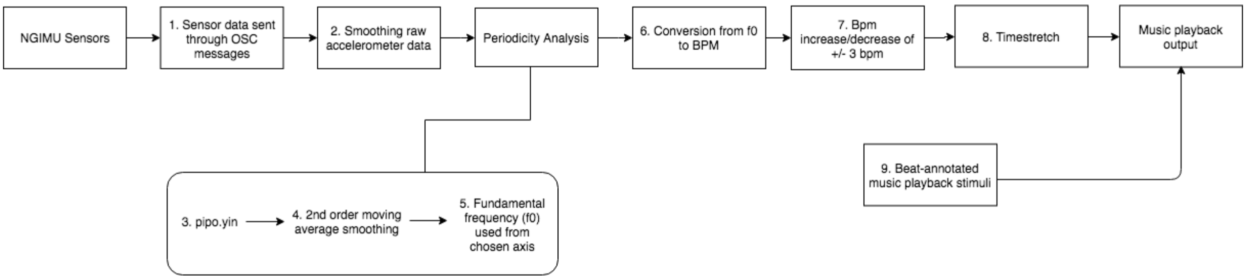

Figure 1. Block diagram of the developed prototype.

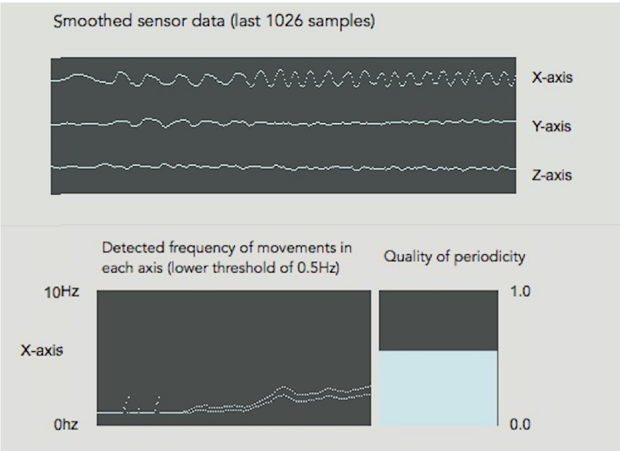

Figure 2. Screen-shots of recorded sensor data graph, detected frequency and the axis with highest estimated quality of detected periodicity. Recorded with one of the participants dancing to a EDM music sample.

the system is interpreting the movements to be in a slower state than previous data stream set. The minimum playback tempo was set to 60bpm. The initial tempo control value was set to the tempo of the original music recording.

8. Bpm-control values are sent to a sample-playback object to be time-compressed/time-stretched in real-time, based on the difference between the bpm-control value and the current playback tempo. The time compression/stretching make use of beat annotations of the played audio sample (see Block 9 in Figure 1). In future work, this may be replaced by a real-time beat estimation of the music audio signal.

Figures 2 provides an example for sensor data when a participant is moving in an intense and repetitive manner. The axis with the resulting highest estimated quality factor of detected periodicity (here, the X-axis) determines which detected frequency to be used for tempo manipulation, and the lower part of Figure 2 illustrate how a tempo increase in the oscillation on the X-axis leads to an increasing tempo estimate for the movement.

3.3 Participants

12 participants between the ages 22-31 participated in the study (8 men and 4 women, mean age 27 years). 9 participants have a background in dance or are working professionally with dance. All participants were reported to be in a healthy condition. Each participant was recorded individually and written consent was obtained before the experiment started.

Figure 3. Two of the participants testing the prototype in the main study.

3.4 Experimental Setup

The participants were offered to choose music samples themselves. However, since no participant preferred this option, music stimuli were randomly chosen for each session within a range of 110-140 BPM from a collection of 33 recent EDM productions [5]. The duration of each session was kept within 15 minutes as a way to keep the participant engaged.

In one session, two experiments were performed for each participant. In the first experiment no tempo manipulations were conducted, and the participant was instructed ”to move freely, but repetitively to the presented music stimuli”. The second experiment included the same task but the ability to control the tempo of the music stimuli through the implemented prototype. The NGIMU was in both sessions placed on the right hand wrist of the participant. The experiments were conducted in a personal living room using 2.0 stereo speakers with Bluetooth for playback of music. The participants were also video recorded to facilitate further analysis of spontaneous reactions and interactions. Figure 3 shows screen-shots from two participants' recordings.

3.5 Questionnaires

After each session, the participants were asked to fill a questionnaire [6], which contained both open- and closed-

[5] The list of songs is provided here: https://bit.ly/2DhrLO9

[6] The questionnaire, including all responses, can be obtained from https://bit.ly/2WGJ0RU

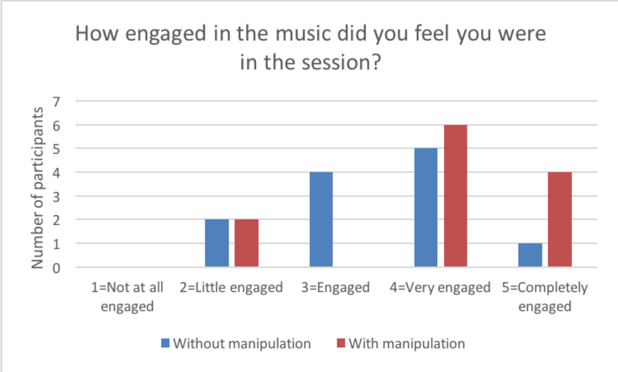

Figure 4. Results of participant ratings of how *engaged* in the music they were in the first (no tempo manipulation) and second (with tempo manipulation) session.

ended questions to gather qualitative and quantitative data. The questionnaire focused on enabling comparison of the participants experience to standard playback to an interactive setting. In addition, several questions were provided to gather more qualitative suggestions regarding the system design.

4. RESULTS

4.1 Engagement in music

Figure 4 shows the participants' ratings when asked about how engaged in the music they were, comparing first (no tempo manipulation) and second session (with tempo manipulation). The x-axis represents the distribution of the participants rating, where 1 = not at all engaged, 2 = little engaged, 3 = moderately engaged, 4 = very engaged and 5 = completely engaged. The y-axis represents the number of participants. The first session without tempo manipulation gave mean value rating of 3.42, while the second session with the ability to manipulate tempo gave a mean value rating of 4.00. Two participants gave a lower rating in the second session (both from 3 to 2), five provided the same ratings, and another five increased ratings.

Participants who had given a lower rating (2 - little engaged) would give the motivation that dancing alone in a room made them feeling little engaged as to whenever he/she would be in a club-like situation: "I think it was not easy to be dancing alone."(P1), "I was all alone. Dancing is kind of a social experience for me."(P8)

Participant (P8) gave this first session a rating of 2 (little engaged) while the second session got a rating of 5 (completely engaged). For the reason that:

> I can't say that my dancing improved, but it was really engaging when you could control the tempo with your movements. (P8)

Another participant who gave a different rating the second session (value of 4 – very engaged), in comparison to the first session (value of 2 – little engaged) explained the difference as:

> Because I was in control of the music. It made it more of a "game" than just dancing to music. (P6)

In relation to the first session, where the motivation for the rating was explained as:

> Hard to lose yourself in the music when you are alone in room like this. You feel watched even though the room is empty. It's easier to dance when you're in a room full of people dancing, or at home, where you feel completely relaxed. (P6)

Another participant who had given a higher rating the first session (value 3 – moderately engaged) but had a different engagement the second session (value 2 – little engaged) explained it as:

> The changing tempo made it hard for me to enjoy the music and dance to it. As I moved to the pace of the music it somehow did not catch my movement and began to slow down which made me have to wave my arm fast to make the music speed up again (...) On the other hand, I felt like I got to interact with the music in a new way. The ability to adjust the music as if I was dj-ing was cool, as I could play with it. The songs also sounded cool when switching the tempo. (P3)

Other participants who had shifted from feeling moderately engaged to either very or completely engaged gave the following reasons:

> The possibility to change the dynamic through my movement was for me more exciting. As well as no need to stay repetitive. (P4)

> Because my moves and actions had an impact on the source/reason why I was originally moving. It created a little bubble in which a conversation with myself could happen. (P5)

4.2 Enjoyment when dancing

Figure 5 shows questionnaire responses when the participants were asked to compare enjoyment of dancing between the first and second session. In addition to comparing the engagement ratings for the individual sessions (previous subsection), these ratings provide an additional comparison from the perspective of the participant. The x-axis represents the rating from 1 (=much worse) to 5 (=much better). The y-axis represents the number of participants giving a certain rating. Following results gave a mean value rating of 3.08.

When asked about how they felt being able to modify the tempo, a majority described a positive feeling. But along with a positive feeling, some still expressed having a somewhat split feeling about the interaction.

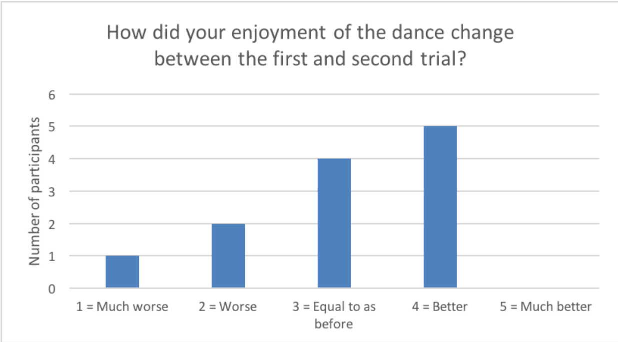

Figure 5. The rating when comparing the participants *enjoyment* of dancing between the first and second session.

> It was playful, an enjoyable negotiation. (P4)

> It was interesting and fun. At the same time there was a feeling of responsibility towards the tempo in comparison to the music that is not tempo modified. Like if it would be my mistake if the dance floor died. (P2)

> It felt interesting but exhausting to keep up as my arm had to stay in one tempo even though my body might have wanted to working in contradiction to the music sometimes. But it was also interesting to hear when I got tired and then realizing that I had physically changed tempo. (P11)

> Great but confusing at times. It made me move in a certain way to be sure to not mess up the tempo. Felt a bit restricted. (P12)

A common feeling of restriction as the last-mentioned could be identified among other participants as well.

> It was hard to use it, as there was a delay of a few seconds, and as I normally adjust to the beat and have a difficult time to set the pace for it to play, as I then need to move faster than the music. (P3)

> It slowed down to easily in my opinion. I often felt that the tempo was perfect, but it always slowed down a few seconds later. (P6)

> (...) I felt more engaged in one way, because I could control the pace of the music. Although I felt more restricted because I had to more repetitive and less instinctual. (P9)

A couple participants described a feeling of uncertainty in if he/she is doing right.

> (...) you realise that you sense of beat has gotten worse. (P7)

> (...) perhaps a slight misconception from my part when the vocals kicked in as I didnt feel that I had as much control over them. (P10)

4.3 Spontaneous reactions

Recognized among the participants was how a larger part of the participants generated more arm movements during their second session in comparison to their first session. Moving the arm as an indication of exploring possibilities in the tempo manipulation, going from one extreme to another (fast/slow, periodic/non-periodic) was often followed by reacting with a laughter. There were furthermore participants who appeared to shift between adjusting their dance to what they were hearing and interfere their dance by generating arm gestures to control the tempo. [7]

5. DISCUSSION

In this study, rhythmic movements of dancers were analyzed in real-time for their predominant periodicity, which was then mapped to manipulate tempo in the music playback the dancer was moving to. The emphasis in the user-study and resulting evaluation was put on the participants' subjective experience, both with and without the ability to control the tempo of the music. Even though no statistical significance emerges from our study, the results indicate the potential of positive dance experiences when improving the system based on the comments of our study.

Preserving the sound quality of the input audio was one of the main challenges. Granted that the prototype aimed to strictly change the bpm of the playing input audio – without affecting other sound parameters – the music's characteristics and the sound quality of it were still affected. A possible explanation for how most of the participants felt more engagement and/or enjoyment could be grounded in a feeling that it is music they themselves somewhat created. Thus, an enhanced feeling in their engagement and/or enjoyment in the interaction. Likewise, positive effect on engagement could be originating from the fact that they had to execute more control and needed to be more attentive to the details in the music playback.

Modifications that became significant for the user interaction were found to be situated in the functionality on how the playback tempo changes. The developed system applies small but noticeable changes within a short time span for the user to sense the agency in the interaction.

Among the participants' expressed opinions, the most common criticism was regarding the delay between a change in body movement and a tempo change in the playback. In the way that the current prototype is constructed, the detected frequency is always analyzed comparing current subset to previous subset. If the frequency is analyzed to be higher, the tempo will increase. If lower, the tempo will decrease. Thus, attempting to make a tempo change for a short time period might result in a playback tempo manipulation. It therefore requires the user to create faster movements for a longer time period in order to make movement changes noticeable in sound.

After having arrived at a desired tempo of the playback, the frequency and clarity of repetitive movements was fre-

[7] Video examples of users in the tempo-manipulation experiment: https://youtu.be/3toOXtS2bKI, https://youtu.be/i63UBMehWGs, https://youtu.be/IivNDDOkxeQ.

quently observed to decrease. This made P3 feel that the system did not catch the pace of the movements and therefore this participant pointed out decreased enjoyment of the dance in the interactive setup.

As a potential drawback of this study, a majority of the participants were educated or are professional dancers. For participants who have a deep background in dance performance, it is assumed that taking on the task given in this study can happen effortlessly and possibly feel more engaged to any music that is played for them. Having a larger as well as more diverse population, the distribution of feedback given from the participants may well have differed. However, involving experienced dancers enabled us to get a rich body of verbalizations that can guide further development.

5.1 Future research

Some of the participants expressed confusion about which sound parameters in the musical structure were controlled by body movement, even though it was limited to tempo alone. Further study on what other manipulations based on gestures can therefore be explored.

Studying how a beat synchronization would influence the interaction if it were to be done more "musically" can be of interest, *e.g.* changing the tempo only at the beginnings of bars. This can provide a solution for the user to feel more engaged in the dance if wanting to break off from the interaction and stay in the tempo. It is also likely to improve the interaction experience by detecting durations during which a clear signal is not received from the motion data, and deactivating the tempo manipulation in such phases.

In order for the system to be adaptive to as many gestural vocabularies as possible, additional features to the interaction can be considered as a way to give the user further choice and/or control in his or her movements. This allows the user to reconstruct the system as desired to accommodate his or her gestural preferences and/or capabilities, as was the case in Mulder's work of GRIP instruments [21].

Investigating how participants would interact in groups is worth exploring, and could add dimensions of social interaction through entrainment to the interaction.

6. CONCLUSION

The initial objective of this study was to investigate the subjective experience when users are given control of the decision-making in the music that is played for them. A proof-of-concept prototype was built and examined by a total of 12 participants. A user-study was conducted consisting of two sessions, one without tempo manipulations by the prototype and one with tempo manipulations controlled by periodic body movement. The proposed design is suggested to provide a dance experience that can compare positively to a standard playback of EDM music. Results imply giving an overall positive dance experience worth exploring further. For a number of the participants, the prototype indicated contributing to more engagement and enjoyment than to a standard playback of EDM involving not interacting with the prototype. The qualitative statements provide a rich set of directions to develop the prototype towards increased robustness and diversity of interactions.

7. REFERENCES

[1] P. Dourish, *Where the action is*. MIT press Cambridge, 2001.

[2] A. Tanaka, "Music one participates in," in *Proceedings of the 8th ACM conference on Creativity and cognition*. ACM, 2011, pp. 105–106.

[3] G. Castellano, R. Bresin, A. Camurri, and G. Volpe, "User-centered control of audio and visual expressive feedback by full-body movements," in *Affective Computing and Intelligent Interaction*, 2007, pp. 501–510.

[4] C. Guedes, "Controlling musical tempo from dance movement in real-time: A possible approach," in *Proceedings from the International Computer Music Conference*. ICMC, 2007, pp. 453–457.

[5] S. Dahl and A. Friberg, "Visual perception of expressiveness in musicians' body movements," *Music Perception: An Interdisciplinary Journal*, vol. 24, no. 5, pp. 433–454, 2007.

[6] A. R. Jensenius and M. M. Wanderley, "Musical gestures: Concepts and methods in research," in *Musical Gestures*. Routledge, 2010, pp. 24–47.

[7] L. Elblaus, M. Goina, M.-A. Robitaille, and R. Bresin, "Modes of sonic interaction in circus: Three proofs of concept," in *Sound and Music Computing Conference*, 2014.

[8] P.-J. Maes, D. Amelynck, M. Lesaffre, M. Leman, and D. Arvind, "The conducting master: an interactive, real-time gesture monitoring system based on spatiotemporal motion templates," *International Journal of Human-Computer Interaction*, vol. 29, no. 7, pp. 471–487, 2013.

[9] F. Bevilacqua, F. Guédy, N. Schnell, E. Fléty, and N. Leroy, "Wireless sensor interface and gesture-follower for music pedagogy," in *Proceedings of the 7th international conference on New interfaces for musical expression*. ACM, 2007, pp. 124–129.

[10] Y. Jung and B. Cha, "Gesture recognition based on motion inertial sensors for ubiquitous interactive game contents," *IETE Technical review*, vol. 27, no. 2, pp. 158–166, 2010.

[11] B. M. Costello, *Rhythm, Play and Interaction Design*. Springer, 2018.

[12] Q. Wang, P. Markopoulos, B. Yu, W. Chen, and A. Timmermans, "Interactive wearable systems for upper body rehabilitation: a systematic review," *Journal of neuroengineering and rehabilitation*, vol. 14, no. 1, p. 20, 2017.

[13] J. Hockman, M. M. Wanderley, and I. Fujinaga, "Real-time phase vocoder manipulation by runner's pace." in *NIME*. Citeseer, 2009, pp. 90–93.

[14] B. Moens, C. Muller, L. van Noorden, M. Franěk, B. Celie, J. Boone, J. Bourgois, and M. Leman, "Encouraging spontaneous synchronisation with d-jogger, an adaptive music player that aligns movement and music," *PloS one*, vol. 9, no. 12, p. e114234, 2014.

[15] M. Leman and L. Naveda, "Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in samba and charleston," *Music Perception: An Interdisciplinary Journal*, vol. 28, no. 1, pp. 71–91, 2010.

[16] P. Toiviainen, G. Luck, and M. R. Thompson, "Embodied meter: Hierarchical eigenmodes in music-induced movement," *Music Perception: An Interdisciplinary Journal*, vol. 28, no. 1, pp. 59–70, 2010.

[17] R. T. Solberg and A. R. Jensenius, "Optical or inertial? evaluation of two motion capture systems for studies of dancing to electronic dance music," in *Sound and Music Computing Conference (SMC)*, 2016.

[18] R. Niewiadomski, M. Mancini, S. Piana, P. Alborno, G. Volpe, and A. Camurri, "Low-intrusive recognition of expressive movement qualities," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ACM, 2017, pp. 230–237.

[19] F. Visi, E. Coorevits, R. Schramm, and E. R. Miranda, "Musical instruments, body movement, space, and motion data: Music as an emergent multimodal choreography," *Human Technology*, vol. 13, no. 1, 2017.

[20] A. de Cheveigné and H. Kawahara, "YIN, a fundamental frequency estimator for speech andmusic," *Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1917–1930, 2002.

[21] A. Mulder, "Getting a grip on alternate controllers: Addressing the variability of gestural expression in musical instrument design," *Leonardo Music Journal*, vol. 6, pp. 33–40, 1996.