Feel What You Don't Hear: A New Framework for Non-Aural Music Experiences

Aoi Uyama
Keio University Graduate
School of Media Design
aoi727@kmd.keio.ac.jp

Danny Hynds
Keio University Graduate
School of Media Design
dannyhynds@kmd.keio.ac.jp

Dingding Zheng
Keio University Graduate
School of Media Design
zheng208@kmd.keio.ac.jp

George Chernyshov
Keio University Graduate
School of Media Design
chernyshov@kmd.keio.ac.jp

Tatsuya Saito
Keio University Graduate
School of Media Design
tatsuyas@kmd.keio.ac.jp

Kai Kunze
Keio University Graduate
School of Media Design
kai@kmd.keio.ac.jp

Kouta Minamizawa
Keio University Graduate
School of Media Design
kouta@kmd.keio.ac.jp

ABSTRACT

Just as the way a performer is moved differs even among audiences who have the same impression of the performance (Oode et al. 2009), the sensations and experiences felt by the performers themselves and the audiences' experiences also differ. The purpose of this research is to create a new listening experience by analyzing and extracting the performer's introspection of rests, groove, and rhythm, and physically presenting it to the audience. Although these elements are important in shaping music, they are not always directly expressed as auditory sounds. Our hypothesis is that this introspection, such as a sense of rhythm and groove, is latent and observable in physiological states such as breathing and heartbeat. By sensing and presenting them to the audience, music appreciation that includes introspection could become possible. In other words, by sensing and presenting introspection to the audience, the music listening experience itself can be redesigned to include a physicality that is closer to the performer's experience of the music, rather than being passive in an auditory sense (Morisawa 2014). In this study, preliminary experiments were conducted on the extraction of the performer's introspection, and a device was designed to present it to the audience.

Author Keywords

Haptics, Music, Feeling Share, Experience

CCS Concepts

• **Human-centered computing** → **Systems and tools for interaction design;** • **Applied computing** → *Arts and humanities; Performing arts;*

NIME'23, 31 May-2 June, 2023, Mexico City, Mexico.

1. INTRODUCTION

Melody, harmony, and rhythm are considered as important elements that constitute music regardless of the genre (Kagawa, Tezuka, and Inaba 2015). Beside these aural elements of music, during a performance there are also elements that are not transmitted by sound but are vital in expressing the music (Davidson and Correia 2002). For instance, Watanabe et al. pointed out that the sense of groove felt by the players is not directly shown in the musical score (Watanabe and Chikayama 2006). We will refer to the subjective elements from the performers as "non-aural elements". These elements usually stem from subjective aspects of how the performers perceive the music and it's emotional flow and imagery. These subjective aspects are also reflected in physiological responses such as respiration and heartbeat patterns during the music performance (Cervellin and Lippi 2011). In this project, we explore non-aural elements such as the performer's perception of beat, rhythm, and rests. We also explore the emotional flow not explicitly presented in the musical score as well as aural aspects of the performance.

By extracting and sharing the differences in interpretation, individuality, and musicality of the performers from various perspectives rather than analyzing the structure of the piece, we may be able to allow the audience to feel the physical experience of the performers more strongly and sincerely.

This work explores a means of sharing the personal sensations that are often difficult to share. Since the sensations to be shared are of the bodily experience, the recorded introspection will be replaced by a physically approachable method. We propose a new way of music performances by adding the musicality and introspection of the performer that usually is not directly passed to the audience, and sharing performers psychophysiological responses during the performance. We believe this direction could be a new form of expression during musical performance for both the audience and the performers (Kanato et al. 2016).

2. DESIGN CONCEPT

Thomas Clifton addressed that "music is the actualization of the possibility of any sound whatever to present to some human being a meaning which he experiences with his body—that is to say, with his mind, his feelings, his senses, his

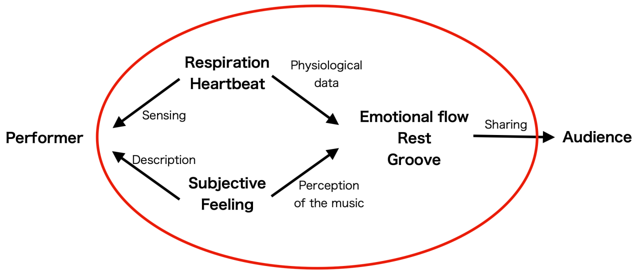

Figure 1: Design Concept.

will, and his metabolism" (Clifton 1984). However, during performances, the performers usually experience and express their "feelings, sense, will, metabolism"beyond sounds (Davidson and Correia 2002). Since the focus of this work is to offer the audience a glimpse of the performer's subjective perception and introspective experience, we tried to quantify and record the experience itself and then designed a way to render and reproduce it (see Fig.1).

In this work, we try to extract the information of the performers subjective perceptions of the music pieces as well as the non-aural expressions during performance from two perspectives: self-reported subjective perception regarding the performed piece and measurable biophysical data during the performance. The first part consists of physiological data (e.g. heartbeat and respiration) and physical phenomena (e.g. movement and sound). The second part highly relies on subjective and individual perceptions of the performer. To add these new dimensions to the audience's experience while listening to the music, we extracted these two sets of prerecorded data into vibration and force feedback from a backrest in addition to the musical recording.

2.1 Subjective Interpretation and Experience Recording

In order to test the designed system in a controlled lab environment, we prerecorded a piano performance with a performer as well as her biometric response during the performance and her subjective report of perception of the music piece. First, an investigator explained to the performer the expected procedures and helped her wear the respiration tracking device (see Fig.2 a) and heartbeat tracking device (see Fig.2 b and c). For this study we asked the performer to play a 1 minute excerpt from Debussy's Suite Bergamasque "Clair de Lune" and recorded the sound with a setup of YAMAHA Grandpiano, Audio-Technica AT2035 microphone, Yamaha AG06 interface. Once the performance started, the investigator recorded the sound of the performance as well as the biometric response. After the performance, the player was asked to add labels to the recorded performance that would be descriptive of the player's interpretation of the piece and personal introspective experience in MIDI data using LogicPro on the sound source while listening to the recording (see Fig.3 and Fig.6 lower).

The performer chose to use two types of labels and described one label as "heavy","deep", or "down"; and the other as "light" or "up". The interview was conducted in Japanese, so we present the literal translation of these term s. However, they may carry slightly different meanings in English. For the sake of clarity in this study, these terms can be seen as "forte" for heavy and "piano" for light. Although we recognize the subjective nature of these terms and labeling methods, we consider this method to be ap-

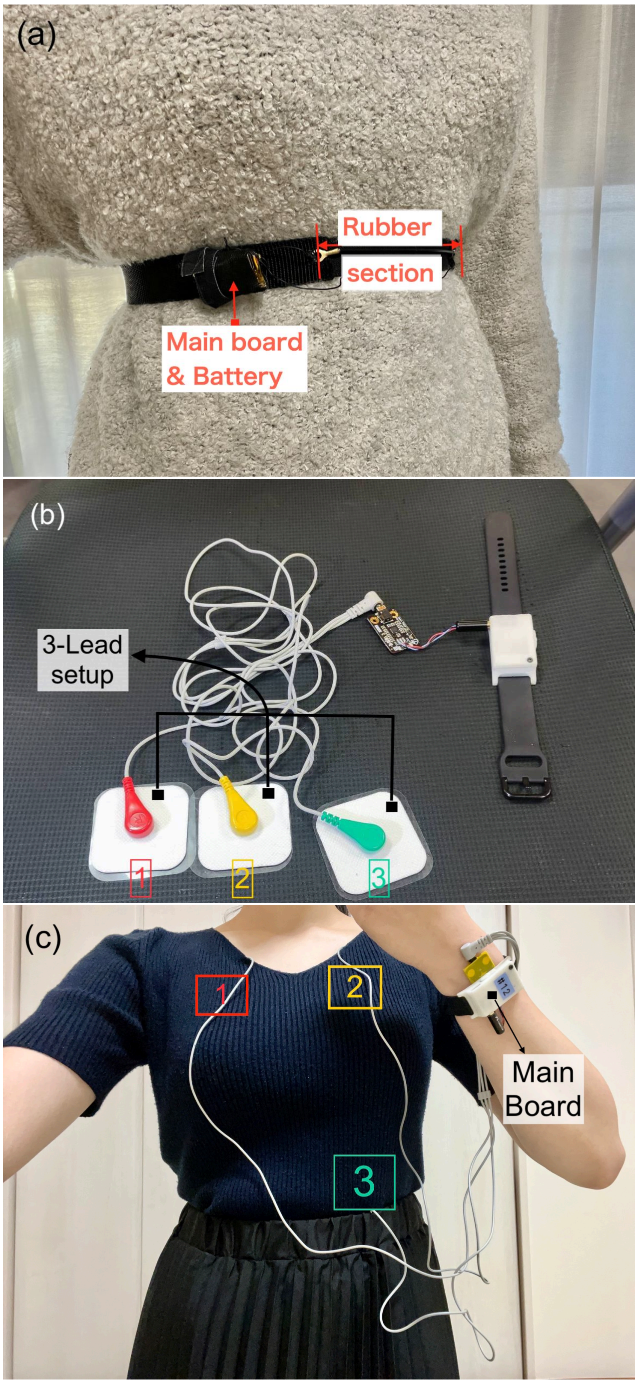

Figure 2: (a) Respiration sensor on a model. (b) ECG setup. (c) ECG setup on a model.

propriate for the cause.

2.2 Relationship Between Respiration and Song Structure

The possibility of conveying breath through vibration may lead to the presentation of the performer's intentions to the audience. This is analyzed from the perspective of the relationship between breath and the structure of the piece. Here, specific specific sections from the score are used as examples, and the analysis and verification are based on the data of respiration and heartbeat during these sections.

Based on the premise that transmitting breathing and

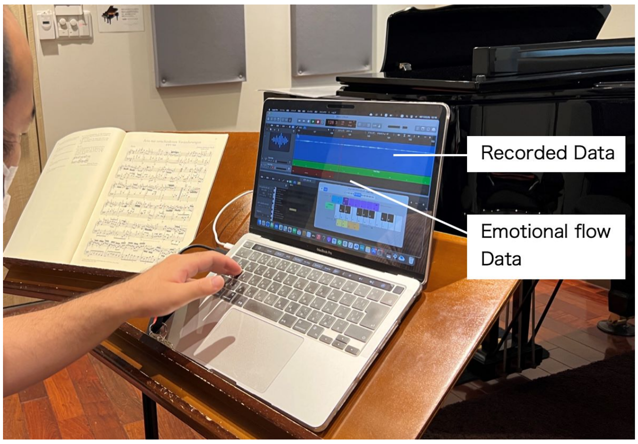

Figure 3: An example of having the performer marking the subjective emotional flow of the music piece with a keyboard. The result example is shown in Fig.6 (upper).

heartbeats through vibration will lead to the presentation of the performer's introspection, we will continue to detect and share biometric information in this experiment. The subjects were told to describe how they thought they were breathing when they played, and once they listened to their own performance once they were asked to write a descriptive graph on the score.

The blue line graph above is the sensed breath data and the red line is the performer's own description. Although there are differences in the number of breaths, the recorded waveforms can be seen to be similar. As shown in Fig.4 and Fig.5, the frequency of the reparation from self-reported breath markings are less than the detected data. The performers pointed out they were not aware that they were breathing that much during the performance. Although differences are expected depending on the performer's emotions, the results suggest that there may be a correlation between unconscious breathing and introspection during performance. This could suggest that unconscious breathing patterns play a significant role in the expression and interpretation of a musical performance.

3. APPARATUS

We used two sets of apparatus in order to make this work possible. One was used to input the data from the performer, while the other output this information and mimicked the perceptions and biometric responses of the performer.

3.1 Biophysical Recording Apparatus

Since this study is rather exploitative we opted for a more traditional setting in order to reduce it's complexity. We recorded a piano player performing several classical pieces and recorded the performance, the performer's breathing, ECG, and asked the performer to post-hoc label the audio piece with some introspective information.

For the respiration recording we used a chest-worn strap with a conductive rubber section (see Fig.2 a). Stretching of that section results in changes in it's resistance. Resistance is sampled at 50Hz rate with a 12-bit ADC analog front-end which is made of a Wheatstone bridge, an RC-filter and an instrumentation amplifier with adjustable gain. Later on, the signal was cleaned with a digital band-pass filter (0.2-0.5Hz) to remove the physical movement-induced noise and

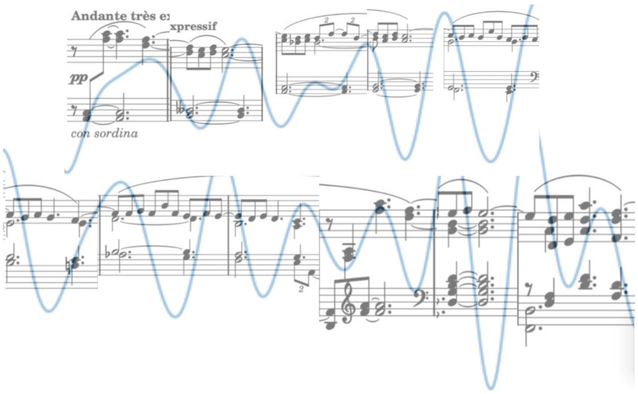

Figure 4: Scores and Detected respiration data (blue line).

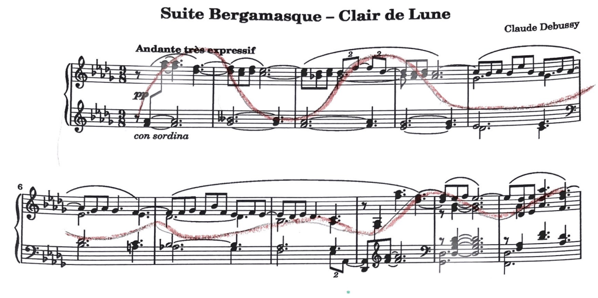

Figure 5: Scores and breath description by performer (dark red line).

smooth-out the curve (See an example in Fig.4).

ECG was measured at a rate of 200Hz with 12-bit resolution with a custom built wrist-worn data acquisition device (See at Fig.2b). For the recording we used the Analog Devices AD8232 ECG fronted in the manufacturer recommended configuration and 3-lead setup. Post-hoc the raw signal was smoothed with a (0.7-3.5Hz) band-pass filter. Since for our purposes we needed the exact timing of each beat, we used the heartpy library (Gent et al. 2018) to detect all the R-peaks in the ECG waveform and used their timings.

Both devices (Fig.2) were battery powered and wirelessly operated for the performer's comfort.

3.2 User Test Setup Apparatus

As shown in Fig.8, we conducted 2 different sessions in the user test. In the device wearing session, we prepared the following apparatus.

The labels related to personal introspection given by the performer were presented using vibration on the hands. Although the performer stated that there is no association between given labels (Heavy/Forte and Light/Piano) and left-right hands, for consistency we presented the vibrations for "heavy/forte" on the left hand and for "light/piano" on the right. An example of the vibration onsets and timing is in Fig.6. Vibration was produced using two Acouve VP2 voice-coil actuators driven through an audio amplifier (Fig.8 b).

Respiration was presented using a flexible backrest slightly pushing the the participant forward on breathing out and leaning back on breathe out (See Fig.7). The backrest was actuated with a 300N linear actuator that was moving according to the recorded respiration waveform.

Heartbeats were presented using a large Acouve VP7 voice-

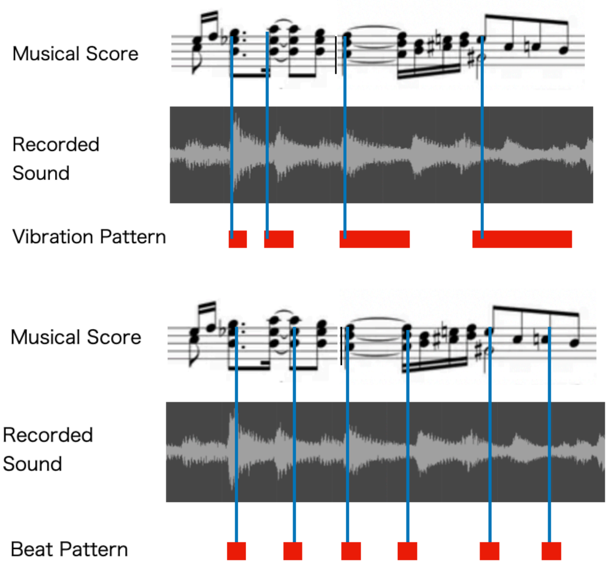

Figure 6: An example of the emotional flow described by performer (upper). Example of following the beat (lower).

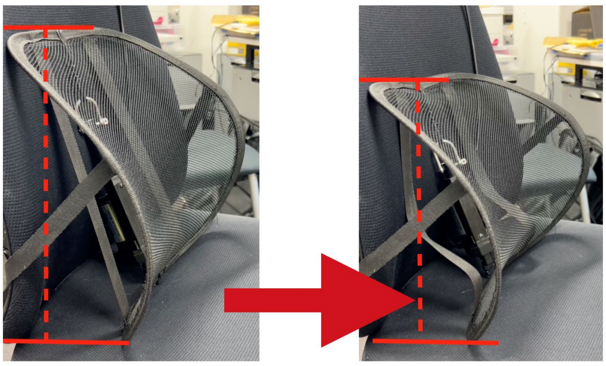

Figure 7: Back-chest type respiratory heart rate indication device.

coil haptic actuator placed on the participants' chest. At the time of each beat, the actuator was producing a heartbeat-like vibration pattern made from a low-pass filtered sound of an actual heartbeat.

To remove the sounds of the haptic transducers and the linear actuator in the backrest, participants were listening to the recording using Sony WH-1000XM4 noise canceling headphones in wired configuration. Using the headphones, the sounds of the actuators were nearly inaudible.

4. USER TEST

In order to explore how the subjective aspects of the performer may impact the experience of the audience alongside the audio recording, we conducted a user test applying light pushes and vibrations as simulations. We attempted to achieve at least a vague intuitive understanding of the performer's introspection. Considering it is hard to describe the subjective interpretation of the music piece as well as its perceived emotional connotations, at this stage we are not focusing on high fidelity representation of something we struggle to define.

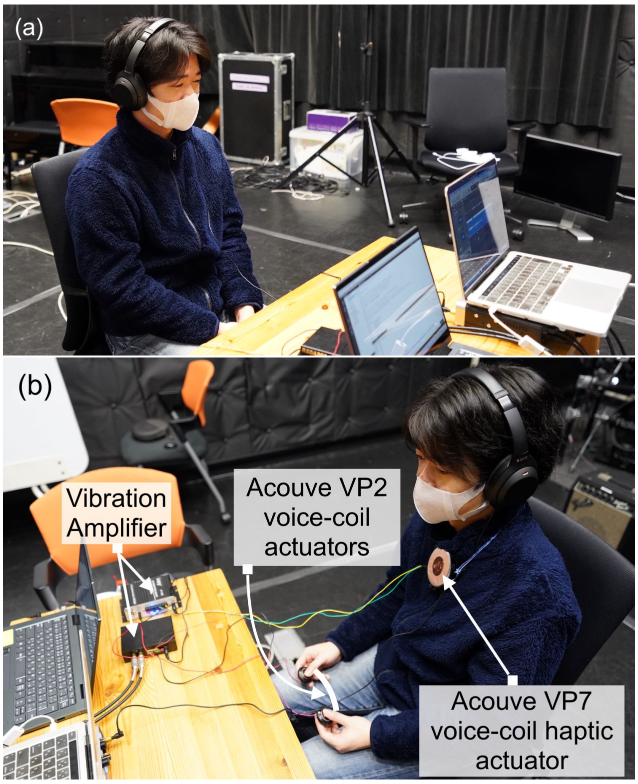

Figure 8: Two sessions of the user study: (a)the participant were listening to the music with noise canceling headphone only; (b) in the other session the participants were listening to the music with vibrations and light pushes from the backrest device (see at Fig.7).

4.1 Procedure

In presenting physical introspection, the following experimental procedure is used for each performer and audience member.

Seven volunteers (3 males and 4 females) between 20 to 32 years old participated in this test. Two had over 10 years of music training, three had less than 10 years of music training, and two of them had no musical training. No compensation was offered.

After reading the consent form and explaining the procedures by the investigator, all users signed the consent form. A demographic and music background questionnaires was then filled out. There were two main sessions, one in which the users were asked to listen to a 1 minute recording with our setup and then another without the setup. The order of the sessions are counterbalanced. After each session, the users were asked to fill out a feedback questionnaire regarding their experience. After these two testing sessions, all users were interviewed by the investigator regarding the experience of this user test. The whole procedure took about 20 minutes for each user. In the beginning of the user study we chose not to explain the meaning of the additional stimuli to the participants in order to observe their reactions and see what associations they have with our system. Besides the recorded performance, we used the respiration and heartbeat data we recorded from the performer, as well as the personal introspection or the emotional flow labels given by the performer. All this information was presented through force and haptic feedback.

4.2 Feedback and Discussion

We gathered mostly positive feedback from our seven participants. Three mentioned that when using the system, they were trying to focus on the connection between the feedback and the music performance. Two of them mentioned they had different experiences with and without the system. They described the session using the devices as positive and engaging. We also received some feedback which could help us to improve the user experience for future work. While three users found the devices fit naturally, four of them found themselves concentrated more on the tactile sensations caused by the devices. Three pointed out the distracting feelings may have been caused by stimulation from the devices being different from their own heartbeats, breathing patterns, and the emotional flows they perceive from the music recording. These differences are intended as the motivation of this work. However, this feedback encouraged us to improve the designed system for a more complete musical experience without sacrificing the immersion of the aural experience. Since we only tested a 1 minute excerpt with seven users, it is possible that the sessions end before users can get used to the feedback. Since most musical concerts are longer than 1 minute, we would like to investigate if longer experiences will lead to any new findings. We found no relation between participant's musical training and the feedback given.

Besides the users, we also interviewed the performer. She pointed out the physical movement of the body during performances also has an effect on the way a work or phrase is perceived. Moreover, some physical sensations such as the vibrations under the feet from the piano, are only perceived by the performer. In terms of sensing the beat, this is not just reliant on the BPM of the piece but also the way the performer expresses each beat and the correlation between patterns and rhythms within the score. Every performance can have expressive variations, which in turn presents a varying experience for the audience each time. Some rhythms within the score are oftentimes left vague in order to allow the audience to interpret them more freely. In fact, even the performer sometimes has trouble describing the beat due to the unconscious sensations occurring throughout the performance. Thus, it is necessary to provide a broad and flexible form of communication for the performers to utilize in order to allow them to express their introspection more freely. This would allow the performers to share their personal physical introspection in the same sincere manner as their aural expression allows. In addition, this new method not only enriches the musical experience for the audience, but knowing how their perception may be directly passed to the audience also changes how the performers may think about their own expressions and thus impact the performances.

Figure 9: Held type force feedback device.

Figure 10: Wearable type force feedback device.

5. CONCLUSIONS AND FUTURE WORKS

This system is designed so the audience could feel the performer's experience in real time; not through words or symbols, but through physical sensations. We chose to not use any visual cues as to not overwhelm the participants. Another consideration is that the interconnection between visual and acoustic neural pathways is well known (Cervellin and Lippi 2011), more so than haptics and tactile sensations. Although these associations can help in the future, at this point in the project they may introduce unnecessary variability. At this stage, we leave the feedback vague on purpose to avoid letting existing associations alter the experience in an undesired way and to reduce the complexity.

We plan to increase the number of subjects and the length of the music pieces. We hypothesize that this will highlight the differences in individual perception and aid us to design a more comprehensive approach to the representation of the artist's subjective introspection. For future system iterations we will use inflatable force feedback devices using air pressure held by the user (See Fig.9), and worn on the body (See Fig.10), to see if a more immersive experience can be achieved, and to see how placement affects this experience. We believe this would lead to a more organic feel of the feedback and be less distracting. We will continue this work with more artists and spectators in order to investigate the diversity of the performers' ways of expressing their personal introspective experience.

6. ACKNOWLEDGMENTS

We would like to thank the pianist who cooperated with us with her wonderful piano playing. We would also like to express our appreciation to all the participants who understood the purpose of this study and willingly cooperated with us. This work was supported by JST Moonshot R&D Program "Cybernetic being" Project (Grant number JP-MJMS2013).

7. COMPLIANCE WITH ETHICAL STANDARDS

The experiments were approved by the university ethics committee. Informed consent was obtained from all participants for being included in the study. All recordings were anonymous, and no personal data was stored.

References

Cervellin, Gianfranco and Giuseppe Lippi (2011). “From music-beat to heart-beat: a journey in the complex interactions between music, brain and heart”. In: *European journal of internal medicine* 22.4, pp. 371–374.

Clifton, Thomas (1984). “Music as heard: A study in applied phenomenology”. In: *Journal of Aesthetics and Art Criticism* 42.3.

Davidson, Jane W and Jorge Salgado Correia (2002). “Body movement”. In: *The science and psychology of music performance*, pp. 237–250.

Gent, Paul van et al. (2018). “Heart rate analysis for human factors: Development and validation of an open source toolkit for noisy naturalistic heart rate data”. In: *Proceedings of the 6th HUMANIST Conference*, pp. 173–178.

Kagawa, Toshimune, Hiroshi Tezuka, and Mari Inaba (2015). “Frequency-based key component extraction-automatic generation of instruction scores for music video games”. In: *Information Processing Society of Japan, Entertainment Computing*, pp. 326–333.

Kanato, Ai et al. (2016). “An automatic singing impression estimation method using factor analysis and multiple regression”. In: *CD review* 699, p. 372.

Morisawa, Yukihiro (2014). “Relationship of Creative Thinking and Feeling Shared Communication by Social media”. In: *Bulletin of Saitama Women’s Junior College* 29, pp. 45–61.

Oode, S et al. (2009). “Evaluation of Kandoh evoked by music: Relation between type of Kandoh and affective value of music”. In: *NHK STRL R&D* 50, pp. 1111–1121.

Watanabe, Tetsuro and Takashi Chikayama (2006). “Analysis of Groove Feeling of Drums Plays”. In: *IPSJ SIG Technical Reports* 2006.113 (2006-MUS-067), pp. 27–32.