A Real-time Drum-wise Volume Visualization System for Learning Volume-Balanced Drum Performance

Mitsuki Hosoya[1], Masanori Morise[1], Satoshi Nakamura[1], Kazuyoshi Yoshii[2]

[1] Meiji University, 4-21-1 Nakano, Nakano-ku, Tokyo, Japan
[2] Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto-shi, Kyoto, Japan
cs202017@meiji.ac.jp

Abstract. To improve drum performance, it is important to consider the volume balance of the bass drum, snare drum, and hi-hat. However, it is difficult for players to evaluate the balance of each volume of these instruments while playing. In addition, while it is possible to self-diagnose by recording drum performances, it is not always efficient to re-record and re-play drum performances based on the correction points. Therefore, we developed a system that uses semi-supervised non-negative matrix factorization (SSNMF) to separate a player's drum performance, recorded with a unidirectional microphone installed in front of the drum kit, into the bass drum, snare drum, and hi-hat sound sources in real-time, and estimates each volume at the time of beating. In addition, the system also visualizes their volume balance and enables the player to control the power of beating. We experimented using this system for actual drum performance and clarified its usefulness and points for improvement based on the feedback obtained from the experiment participants.

Keywords: Drum Performance, Sound Source Separation, Volume Balance.

1 Introduction

Drums express rhythm by playing multiple instruments simultaneously. Therefore, to improve one's drumming, in addition to rhythmic accuracy and timbre, it is important to consider the volume balance of the three main instruments: bass drum (BD), snare drum (SD), and hi-hat (HI) (see Fig. 1).

In basic rhythms such as 8-beat and 16-beat, the ideal volume balance for drum performance is generally to play the bass drum, snare drum, and hi-hat in the order of decreasing volume (BD > SD > HI) as perceived by the audience (exceptions are depending on the song and genre). The reason for this is that if the volume of the hi-hat is louder than the other instruments, it may sound like a performance with no intonation, and if the volume of the bass drum is very soft, it may sound like a performance with no stability. Therefore, drummers want to master the ideal volume balance.

When practicing the volume balance of drum performance, drummers who attend drum schools can have their performances listened to objectively by lecture. Still, drummers who practice by themselves, which is a large proportion, cannot listen to their performance objectively, making it difficult to understand the volume balance of their drum performance. Thus, there is a gap of perceived volume balance between a drummer and audiences because audiences listen to drum play a little bit far from the drum (see Fig. 2). Moreover, drummers can check the volume balance independently by recording their performance and checking it, but this method is time-consuming and labor-intensive. In contrast, by practicing with MIDI (electronic) drums, one can check the velocity, which is a numerical value that indicates the intensity of the sound, but this method is not suitable for learning the volume balance because the sensation of beating electronic drums is very different from that of acoustic drums.

In this study, we focused on the problem that drummers do not know the volume balance perceived by the audience during the performance when practicing basic rhythms on acoustic drums. It takes time and effort to check the volume balance. The goal of this research is to develop a system that enables drummers to quickly and efficiently learn the ideal volume balance independently. Therefore, we propose a system that estimates the volume balance of a drum performance by using a unidirectional microphone placed in front of the drum kit with a sound source separation method. In addition, the system visualizes the estimated volume balance in real-time and feeds it back to the drummer (see Fig. 1). We believe our system helps drummers learn the ideal volume balance, as they can learn it quickly by playing while checking the volume balance perceived by the audience.

The contribution of this paper is to focus on the problem of balance in drum practice and to realize a method for estimating volume balance using source separation and visualizing it in real-time.

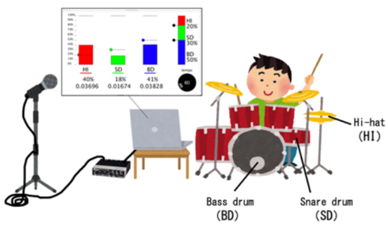

Fig. 1. Assumed system environment.

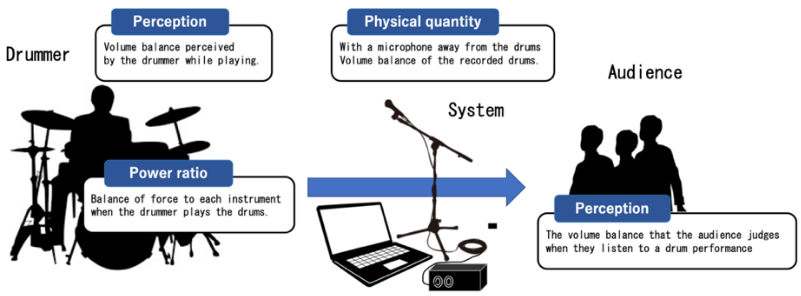

Fig. 2. The volume balance of drum performance in each positon.

2 Related Research

2.1 Support for Practicing Musical Instruments

There have been many studies on support for practicing musical instruments. Smoliar et al. [1] proposed a system that visualizes key strength, tempo, articulations such as staccato and legato, and timing consistency between the right and left hands based on MIDI data of piano performance. Rogers et al. [2] proposed a system that projects a piano roll on a physical keyboard by projection mapping to present the key position. Beici et al. [3] proposed a system that detects and visualizes the piano player's sustain pedal movement. In addition, Marky et al. [4] proposed a system that provides visual guidance via LEDs attached to a fretboard to help guitar players learn chords, scales, and melodies. Doi et al. [5] proposed a system that uses projection mapping to present information on the strings and body of a koto (a Japanese string instrument), such as picking position, picking direction, and symbols representing playing techniques.

As shown in these studies, visualization of performance information is used in various methods to support instrument practice. Since visualization of performance information makes it possible to practice musical instruments while recognizing one's own performance and the performance of one's target, we adopted the method of visualizing the volume balance of drum performance in this study.

As in this study, there have been various studies on assisting the practice of drumming. Ikenoue et al. [6] proposed a system that corrects the control of drumsticks by using a small delayed auditory feedback, and experiments showed that the use of the system could correct playing movements that use the extensor muscles. Holland et al. [7] proposed a haptic bracelet (haptic drum kit) to support the learning of complex drum patterns by presenting vibrations to the drummer from a device worn on both hands and feet. Imada et al. [8] proposed a drum practice support system that utilizes Kinect-based performance motion detection. In a study similar to our own, Iwami et al. [9] proposed a method to visualize and present the variation of sound intensity and tempo of MIDI (electronic) drum performance to users during a performance.

However, these studies have focused on supporting the rhythm and movement of drumming, and few have considered the volume balance of acoustic drums.

2.2 Sound Source Separation Technology

In this study, we use sound source separation to estimate the volume balance of a drum performance in real-time. Source separation is a technique to separate and recognize each sound source from an acoustic signal that contains multiple sound sources. A variety of sound separation methods have been proposed depending on the purpose and conditions. Examples include Universal Sound Separation [10], which separates specific sounds such as animal cries and door creaks to help users distinguish multiple sounds with machines, and Harmonic/Percussive Sound Separation (HPSS) [11], which separates harmonic and percussive sounds of a sound source without using prior knowledge of specific instruments. In addition, an example of the application of sound source separation technology to music is the use of Deep Neural Networks (DNN) [12] to extract vocals from a song [13] and separate them by parts of the song (vocal, bass, drums, etc.) [14].

For sound source separation of drum performances, various methods have utilized Non-negative Matrix Factorization (NMF) [15, 16] and Convolutional Non-negative Matrix Factor Deconvolution (NMFD) [17]. These methods are based on an algorithm for separating sources by decomposing the amplitude spectrum of the source to be separated, which is a single channel (monaural signal), into basis spectra representing the frequency components and activation matrices representing the temporal information corresponding to the individual basis spectra. NMF aims to decompose the magnitude spectrogram of a drum signal into a set of basis spectra corresponding to drum instruments and a set of the corresponding temporal activations. NMFD is a convolutional extension of NMF that uses a set of basis spectrograms for approximating the input mixture spectrogram in a patchwork manner. The main application of such sound source separation for drum performance is automatic music notation, which is the automatic generation of music scores from acoustic signals using a computer, and research to improve the accuracy of music notation is being actively conducted [18-22].

In the studies above, sound source separation for drums was mainly used for automatic music notation and not for estimating the volume balance during drum performance. In the current study, we utilize NMF, which is expected to provide highly accurate separation, as a method for source separation because the basis of the drum instruments to be played can be registered in advance.

3 Proposed Method

3.1 Calculation Procedure

In this study, we aim to estimate the volume balance of drums. Therefore, a drum performance input by a unidirectional microphone is separated into the sounds of each instrument, and the volume of each is estimated. We use the RMS value, which is an index of the sound pressure perceived by a person, calculated from the amplitude of the acoustic signal to calculate the volume. The calculation procedure is as follows.

1. Recognize a drum performance from a unidirectional microphone (see Fig. 3).
2. Separate the sound sources of the drum performance into a hi-hat, snare drum, and bass drum only for each second of the last musical bar (4/4 beat) (see Fig. 4).
3. Detect the onset timing of each separated sound source [23] (see Fig. 5).
4. Calculate the peak (maximum amplitude) from the frames of $t1$ seconds before and after the detected onset, and adopt the peak that exceeds the threshold T as the correct striking timing of each instrument (see Fig. 6).
5. Calculate the RMS value from the $t2$-second frames around each peak, and calculate the average value.

6. Multiply the average RMS values calculated in (5) by the correction values (ω_HI, ω_SD, and ω_BD) to obtain the volume balance that a person feels when listening to a drum performance, and calculate the volume of each instrument.
7. Determine the volume balance as the ratio of the volume of each instrument (6).

Since we felt it would take a certain amount of time for a person to judge the volume balance after listening to a drum performance, we performed the calculation for every frame of the last measure (4/4 beat) in seconds (2). While the peak is a numerical value that indicates the instantaneous loudness of a sound, the RMS value, which is calculated by averaging the energy of a sound, is an index that takes into account a person's perception of sound pressure, so we adopted the RMS value as the volume of each instrument (5). The RMS values were not converted into dB units but were used directly in the calculations. The reason for multiplying the RMS value by a correction value is that there is a difference between the volume balance estimated from the RMS value of the pronunciation timing of each instrument detected and the volume balance felt by a person when listening to a drum performance, and it is necessary to tune it (6).

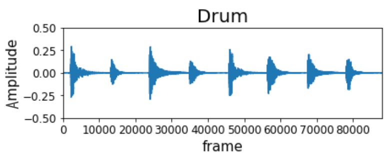

Fig. 3. Waveform of a drum performance input from a microphone.

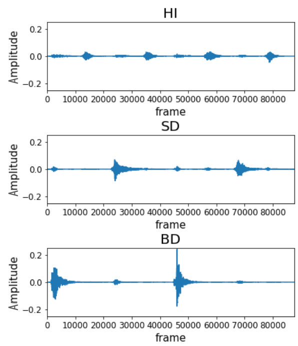

Fig. 4 Waveform after sound source separation.

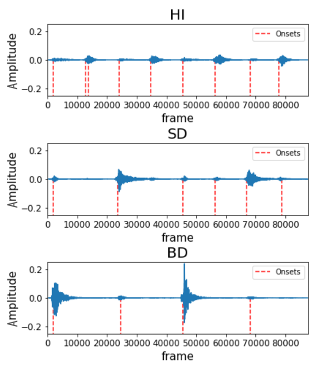

Fig. 5. Detected onsets.

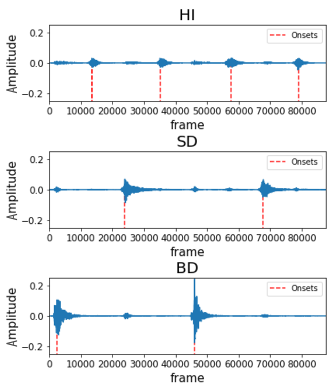

Fig. 6. Timing when the amplitude exceeds the threshold.

3.2 Source Separation Used for Volume Balance Estimation

In this study, SSNMF [24], a source-separation method that can fix the basis spectrum (a matrix representing the frequency components) of the sources separated by NMF, was utilized for volume balance estimation of drum performances. These are algorithms that decompose one non-negative matrix into two non-negative matrices. These methods perform source separation by utilizing the amplitude spectrogram obtained from the speech data, which can be regarded as a non-negative matrix and represented as a product of a basis matrix representing the frequency components and an activation matrix representing the temporal information. In addition, SSNMF is used to deal with the Eigen-differences of the instruments, such as tone and resonance. Utilizing SSNMF enables us to fix the basis in advance, implement the basis created from the recorded sound of the drum instrument used in the system, and deal with the intrinsic difference of the instrument.

The following is a detailed step-by-step description of the method of sound source separation adopted in this study. First, to create the basis for NMF, single notes of the hi-hat, snare drum, and bass drum need to be recorded without clipping (when the input of the recording is too large, the peak of the signal waveform is saturated and collapses at a certain level). The recoded data for each instrument must then be cut out, normalized, and registered in the program. The program then performs a short-time Fourier transform on the recorded data of each instrument and applies NMF to the obtained amplitude spectrogram to get the basis matrix. Then, the sound source of the drum performance recorded from the microphone is subjected to a short-time Fourier transform in real-time, and SSNMF is applied to the obtained amplitude spectrogram with the previously obtained basis of each registered instrument.

In this way, the sound source of the drum performance can be represented in the basis matrix of only the hi-hat, snare drum, and bass drum, and the activation matrix and the sound source separated for each instrument can be obtained by multiplying the basis and the activation matrices of each instrument and then performing the inverse Fourier transform. Then, the sound source of the drum performance recorded from the microphone is subjected to a short-time Fourier transform in real-time, and SSNMF is applied to the obtained amplitude spectrogram with the previously obtained basis of each registered instrument.

4 Proposed System

As stated earlier, the purpose of this study is to help drummers learn the ideal volume balance quickly and efficiently by themselves. Therefore, we developed a system that inputs drum performances from a unidirectional microphone installed in front of the drum kit, extracts only the bass drum, snare drum, and hi-hat by sound source separation, and estimates the volume of each instrument to visualize and present the volume balance of the drum performance to the drummer in real-time. Our assumed usage environment is shown in Fig. 1. A microphone is placed in front of the drum kit, and a PC is placed where the drummer can check the screen while playing. The microphone and PC are connected via an audio interface. The system runs on the PC, estimates the volume balance of the drum performance in real-time, and visualizes the results.

4.1 Implementation

The system was operated on a PC and implemented using Python for the acoustic processing part and Processing for the visualization part. The system visualizes the information shown in Fig. 7. The red, green, and blue bars in the volume balance visualization zone represent the hi-hat, snare drum, and bass drum volume, respectively, estimated in real-time from the drum performance input from the microphone. The volume balance (ratio of each instrument) that the user aims for is set by operating the target volume balance setting bar (right side of the figure). The value of the target volume balance reference bar in the volume balance visualization zone also changes in conjunction with the value of the ratio of each instrument set in the target volume balance setting bar. In addition, the tempo (BPM) setting knob (lower right of the figure) allows the user to adjust the tempo when practicing.

To calculate the volume balance estimation described in Section 3.1, we set $t1 = 0.15$ and $T = 0.8$ so as to recognize the correct striking timing of each instrument. Since the time constant of the Fast characteristic (JIS C 1509-1[25]) in the measurement of noise in the sound level meter is 125 ms, we set $t2 = 0.125$. The correction values, ω_HI, ω_SD, and ω_BD, were set subjectively in this study in order to bring the volume balance calculated by the RMS values closer to the volume balance perceived by a person when listening to a drum performance. Specifically, we set $\omega\_HI = 5$, $\omega\_SD = 1$, and $\omega\_BD = 0.8$ (5 times for hi-hat and 0.8 times for bass drum), as the waveform of the hi-hat was quite small, and the input waveform of the bass drum was rather large due to recording via the audio interface.

4.2 Usage

The procedure for using the system is described below. First, a user sets up a unidirectional microphone in front of the drum kit and places the PC's position to check the display easily (see Fig. 1). Next, the user starts the system on the PC and sets the volume balance of the drum performance (hi-hat, snare drum, bass drum) and the tempo of the performance. Next, the user starts drumming and plays while watching the system screen (see Fig. 7), where the visualization of the volume balance is updated every second for one musical bar in 4/4 beat (this will vary depending on the tempo that is set, which is 4 seconds for BPM 60). The reason for this frequency of updates is because we thought it would take some time for people to judge the volume balance after listening to a drum performance. In Fig. 7, the red bar indicating the volume of the hi-hat is higher than the target volume balance setting bar, indicating that the volume of the hi-hat is higher than the target volume balance set by the user. Therefore, by modifying the performance to hit the hi-hat volume of the drum performance at a lower level, it is possible to play the drums at a volume balance close to the intended target.

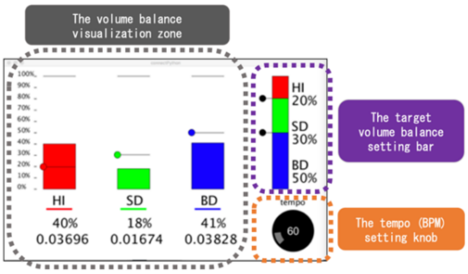

Fig. 7. Screenshot of our system.

5 Experiment

To investigate whether drummers can quickly and efficiently learn the ideal volume balance using the proposed system, and to obtain feedback on its usage, we conducted an experiment in which participants played acoustic drums while using the system (see Fig. 8). The experiment participants were four males in their 20s who had been playing the drums for more than five years and wanted to improve their volume balance.

In preparation for the experiment, we set up the system and equipment (PC, microphone, and audio interface) and created the basis for each drum instrument (HI, SD, and BD) to be used for sound source separation. In creating the basis, we instructed the participants to record one note for each instrument struck at a normal volume (left to the discretion of the participants). Then, we completed the preparation of the basis by performing onset trimming, normalization, and NMF of the recorded sound source of each instrument and registering the basis to the program.

After the preparation was completed, we instructed the usage of the prototype system and the experiment. Then, we asked participants to set a goal for the volume balance and to play the music with a backing 8-beat rhythm. The backing 8-beat rhythm is based on eighth notes and can be played without the timing of the notes of each instrument overlapping (see Fig. 9). Since we wanted to investigate the system's usefulness rather than its accuracy, we limited our experiments to the backing 8-beat, which we judged to be easy to separate from the sound source. After that, we asked the participants to set a target volume balance using the system and to try their drum performance closer to the target volume balance while watching the visualization of the volume balance presented by the system.

During the experiment, we made recordings. When the participants judged that they were able to play at the target volume balance, we asked them to stop using the system and then listen to the recording to judge whether they could play at the target volume balance. After completion of the task, we administered a questionnaire to the participants and interviewed them. The specific questions are shown in Table 1.

Fig. 8. Scene of the experiment.

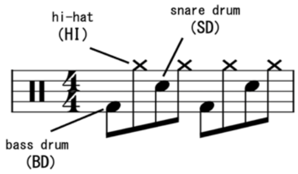

Fig. 9. Sheet music for backing 8-beat.

Table 1. Questionnaire items.

	Question	Answer method
Q1	How satisfied were you with the system?	5-pt evaluation (–2 to +2)
Q2	Would you want to use the system in the future?	5-pt evaluation (–2 to +2)
Q3	Were you able to perform at the set volume balance?	5-pt evaluation (–2 to +2)
Q4	What are your thoughts on using the system?	Free answer

6 Results and Discussions

The average evaluation values were 0.75 for Q1 “How satisfied were you with the system?”, 1 for Q2 “Would you want to use the system in the future?”, and 1.25 for Q3 “Were you able to perform at the set volume balance?”. The average of the evaluation values from Q1 to Q3 were all positive, indicating that the participants’ impression of the system was not bad.

Table 2 and Table 3 list the responses to Q4 “Thoughts on using the system.” Table 2 shows favorable opinions, and Table 3 shows comments for improvement. An overall favorable opinion of the system was obtained in Q4, suggesting that the proposed method can support the acquisition of volume balance in drum performance. However, as indicated by the feedback we received in Q4, several points about the system need to be improved. We will discuss them as follows.

The current system switches the visualization of the estimated volume balance every time it finishes the calculation for about one measure of drum performance, which means its real-time performance is lost to an extent. In addition, since the volume balance of the drum performance is displayed as a percentage of the overall volume, users had difficulty understanding the visualization. Therefore, in the future, we will consider a line graph type visualization method that allows the user to check how the volume of each instrument has been changing. It also became clear that the accuracy of the volume balance estimation was not sufficient. This was because the tuning of the volume balance of the drum performance, which was estimated based on the average of the RMS values of each instrument calculated by the sound source separation, and the volume balance of the drum performance judged by a person listening to the drum performance, was done based on the subjective judgment of the first author. Therefore, in the future, we plan to conduct experiments in which we ask people to judge the volume balance after listening to drum performances with various volume balances, and to review the method of estimating the volume balance perceived by the audience. In addition, the first author recorded, cut out, normalized, and registered the basis of each instrument used in the sound source separation into the program before using the system. Still, we believe the system can be made more efficient by incorporating these tasks into the system itself.

Table 2. Favorable opinions about the system from responses to Q4.

Classification	Answer
Volume balance estimation accuracy	I thought it was great that the volume of the performance was reflected in the results.
	I deliberately played louder or softer, and it was reflected in the results.
	I intentionally changed the volume balance, and the result was reflected immediately.
Impressions	I was able to analyze my performance objectively, which was helpful.
	I was able to check my volume balance with concrete figures and graphs.
	It was good to be able to check my volume balance with objective numbers.
	It was good to know that an instrument was relatively loud or quiet.
	It was good practice to be aware of the volume balance while paying attention to the tempo.
	I thought it was important to be aware of the volume balance.

Table 3. Opinions about improvements to the system from responses to Q4.

Classification	Answer
Volume balance estimation accuracy	When I hit the snare hard, I felt it was not reflected in the result.
	I felt that the bass drum was perceived as louder than it was.
	The results were roughly linked to my performance, but I felt they were not linked when I made detailed changes.
	There were times when the display went to 0% while I was playing.
	There were times when the displayed results of SD and BD were confusing.
	I felt that the results were affected by the tone.
	I felt that the accuracy of the results decreased when the tempo was made faster.
How to visualize volume balance	I want to check the volume of each instrument, not the percentage.
	It was difficult to see when the results of the previous performance were displayed, so I would like the timing to switch to the beginning of a measure.
	When the results of a performance from a few seconds ago are fed back, they may differ from the current performance results, so I would like a more real-time feel.
UI	I would like to be able to freely input the ratio of the target volume balance setting.
	The value of the BPM change knob could not be set smoothly when using the mouse, so it would be better if it could be controlled with the arrow keys.
General impressions	It was difficult to adjust the performance volume while being aware of the metronome's tempo and watching the visualization.
	If you are not familiar with the system, it would be difficult to understand how much of a change in performance would be reflected in the volume balance display.
	Since I often play Japanese rock music, I would like to use the system more if it could support open hi-hat and fast tunes.
	I want to use it while playing in a band (with other instruments).

7 Conclusion

In this study, we focused on the problem that when practicing basic rhythms on acoustic drums, drummers do not know the volume balance perceived by the audience during the performance, and that it takes time and effort to check the volume balance. We proposed and implemented a system that separates drum performances recorded with unidirectional microphones installed in front of the drum kit into bass drum, snare drum, and hi-hat sound sources in real-time, which estimates the volume at the time of striking, and presents the results to the user in a visual form. We then conducted an experiment to obtain feedback from drummers who used the system.

The results showed that the proposed system has the potential to help drummers learn the target volume balance quickly and efficiently. There are also many areas for improvement regarding the method of visualizing the volume balance and the accuracy of the volume balance estimation. In future work, we will improve the system based on the feedback obtained from this experiment.

References

1. Smoliar, S., Waterworth, J. and Kellock, P.: pianoFORTE: A System for Piano Education Beyond Notation Literacy. Proc. 3rd ACM International Conference on Multimedia, pp.457–465 (1995).
2. Rogers, K., R¨ohlig, A., Weing, M., Gugenheimer, J., K¨onings, B., Klepsch, M., Schaub, F., Rukzio, E.,Seufert, T. and Weber, M.: P.I.A.N.O.: Faster Piano Learning with Interactive Projection. Proc. 9th ACM International Conference on Interactive Tabletops and Surfaces, pp.149–158 (2014).
3. Beici, L., György, F. and Mark, S.: Measurement, recognition, and visualization of piano pedaling gestures and techniques. Journal of the Audio Engineering Society, vol. 66, no. 6, pp. 448–456 (2018).
4. Marky, K., Weiß, A., Müller, F., Schmitz, M., Mühlhäuser, M. and Kosch, T.: Let's Frets! Mastering Guitar Playing with Capacitive Sensing and Visual Guidance. Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. New York, NY, USA, no. 169, pp.1–4 (2021).
5. Doi, M. and Miyashita, H.：Koto Learning Support Method Considering Articulations. In Proceedings of ACE, pp.368–383 (2017).
6. Ikenoue, A., Nishimoto, K. and Unoki, M.：iDAF-drum: Supporting Practice of Drumstick Control by Exploiting Insignificantly Delayed Auditory Feedback. Knowledge Information and Creativity Support Systems, pp. 483–498 (2016).
7. Holland, S., Bouwer, A. J., Dalgelish, M. and Hurtig, T. M.：Feeling the beat where it counts: fostering multi-limb rhythm skills with the haptic drum kit. Proceedings of the 4th International Conference on Tangible and Embedded Interaction, pp. 21–28 (2010).
8. Imada, Y. and Ochi, Y.：Hitting Arm Detection for Drum Performance Learning System Using Kinect. The Seventh International Conference on Informatics and Applications (ICIA2018), pp.8–10 (2018).
9. Iwami, N. and Miura, M.：A Support System for Basic Practice of Playing the Drums. International Computer Music Conference, pp. 364–367 (2007).

10. Kavalerov, I., Wisdom, S., Erdogan, H., Patton, B., Wilson, K., Le Roux, J. and Hershey, J.: Universal sound separation. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), pp. 175–179 (2019).
11. Ono, N., Miyamoto, K., Kameoka, H., Le Roux, J., Uchiyama,Y., Tsunoo, E., Nishimoto, T. and Sagayama, S.: Harmonic and Percussive Sound Separation and its Application to MIR-related Tasks. Advances in Music Information Retrieval, vol. 274, pp. 213–236 (2010).
12. Grais, E. M., Sen, M. U. and Erdogan, H.: Deep neural networks for single channel source separation. ICASSP, pp. 3734–3738 (2014).
13. Hennequin, R., Khlif, A., Voituret, F. and Manuel, M.: Spleeter: a fast and efficient music source separation tool with pre-trained models. Journal of Open Source Software, vol.5, no. 50, p.2154 (2020).
14. Stöter, F. R., Uhlich, S., Liutkus, A. and Mitsufuji, Y.: Open-Unmix - A Reference Implementation for Music Source Separation. Journal of Open Source Software. The Journal of Open Source Software, vol. 4, no. 41 (2019).
15. Lee, D. D. and Seung, H. S.: Learning the parts of objects with non-negative matrix factorization. Nature, vol. 401, pp. 788–791 (1999).
16. Cichocki, A., Zdunek, R., Phan, A.H. and Amari, S.: Non-negative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation. Wiley, New York (2009).
17. Smaragdis, P.: Non-negative Matrix Factor Deconvolution; Extraction of Multiple Sound Sources from Monophonic Inputs. International Conference on Independent Component Analysis and Signal Separation (ICA), pp. 494–499 (2004).
18. Ueda, S., Shibata, K., Wada, Y., Nishikimi, R., Nakamura, E. and Yoshii, K.: Bayesian drum transcription based on non-negative matrix factor decomposition with a deep score prior. ICASSP, pp. 456–460 (2019).
19. Paulus, J. and Virtanen, T.: Drum transcription with non-negative spectrogram factorisation. European Signal Processing Conference (EUSIPCO) (2005).
20. Dittmar, C. and Gärtner, D.: Real-Time Transcription and Separation of Drum Recordings Based on NMF Decomposition. International Conference on Digital Audio Effects (DAFX), pp. 187–194 (2014).
21. Wu, C. and Lerch, A.: Drum Transcription Using Partially Fixed Non-Negative Matrix Factorization with Template Adaptation. International Society for Music Information Retrieval Conference (ISMIR), pp. 257–263 (2015).
22. Vogl, R., Widmer, G. and Knees, P.: Towards multi-instrument drum transcription. Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18) (2018).
23. Böck, S. and Widmer, G.: Maximum filter vibrato suppression for onset detection. International Conference on Digital Audio Effects (DAFX) (2013).
24. Lee, H., Yoo, J. and Choi, S.: Semi-Supervised Non-negative Matrix Factorization. IEEE Signal Processing Letters, vol. 17, no. 1, pp. 4–7 (2010).
25. C 1509-1 : 2017 (IEC 61672-1 : 2013), http://kikakurui.com/c1/C1509-1-2017-01.html, last accessed 2021-7-26.