A Motion Analysis Method for emotional performance on the snare drums

Masanobu Miura,*1 Yuki Mito#2, and Hiroshi Kawakami#3

* *Dept. of Media Informatics, Ryukoku University, Japan*

Dept. of Music, Nihon University, Japan

[1]miura@mail.ryukoku.ac.jp, [2]mitotic@hotmail.com, [3]kawakami.hiroshi@nihon-u.ac.jp

ABSTRACT

This study proposes a method for averaging several motions in order to analyze and synthesizing motions of musical performance. The averaged motion is expected to be useful for obtaining the feature of specified motions by just observing visually. Targeted motion here is the snare drum performance with emotion. This method is named "Motion-Averaging-Method (MAM)". Motion data are recorded by a motion capture system for performances by trained percussionists expressing each of five basic emotions or non-emotion. Recorded motion data have some deviations due to the variability of position and/or angle of each player when recording. Thus, the proposed method adjusts position and angle of the player in each recorded motion. Adjusts motion data are expanded or contracted based on impact time of drumstick obtained from acoustic waveform of recorded performance, and then an averaged motion is obtained by observing several motions adjusted. Quantitative features of averaged motion are extracted from stroke motions and their ratios of parameters of arm motions among emotions, as well as collecting up features of motion among emotions. A subjective experiment was conducted to evaluate the appropriateness of obtained features. Results showed the existence of motion related to a 2D-emotional space. The results show that several motions are dependent to the 2D emotional space and emotional performance has several features of motion not related to musical sound. We found that professional percussionists are representing emotion on the motion of the performance dependent to the 2D space and independent to its acoustic signal.

I. INTRODUCTION

The important aspects of musical performance are both the sound and its motion. Sounds of musical performance have been analyzed in order to describe the artistic deviation for the performance [Seashore, 1938]. The analysis of motion has been conducted, and a study pointed out the importance of motion on informing the style of performance to listeners [Davidson, 1993]. Concerning the snare drum performance, several studies are conducted such as the extraction of motion of arms [Dahl, 2006] and measurement of physical loads for players with respect to the rebound of snare head when stroking [Miura, 2012]. Although the studies have been clarified several aspects of drummers, the average features of performance among recorded has not been discussed. Here discusses a method for obtaining the average feature for several performances, by obtaining the averaged data of motion capture. We name the method as "Motion Averaging Method (MAM)". By using the MAM, we will understand the entire list of feature of emotional performance visually, and then it is expected that we will specify the characteristic feature of motion.

Here we focus on the motion of snare drum performance. Performers are asked to play an etude of solo snare drum with expression of each of five basic emotion (Juslin and Sloboda, 2001) or non-emotion. We record the snare drum performance by using a motion capture system and obtain an averaged motion for recorded multiple performances by using the MAM. Then we will obtain the characteristic of emotion in the averaged motion. The obtained characteristic is independent to the player, tempo, or other aspects, so it is thought as to depend on only the difference of emotion.

II. RECORDING THE SNARE DRUM PERFORMANCE

A. Outline

Snare drum performances are recorded by a motion capture system. Employed players are a professional and two students majoring percussion instruments. The employed etude is extracted from a famous textbook in Japan [Imamura and Tsukada, 2005]. Musical score of it is shown in Fig.1. We ask all the players to play it under each of six emotions such as tenderness, happiness, sadness, fear, anger, or non-emotion. The tempo of each performance is up to each player, but they are asked to keep it in each performance as much as possible. Number of trial for each emotion is two, so that we obtain 36 patterns of performance, comprised of 3 players, 6 patterns of emotion, and 2 trials. The obtained data is named as "emotional performance", derived from the fact that each performance is assumed to represent one of each emotion.

Figure 1. Performance Task.

B. Recording environment

We record the performances at a studio in Department of Music, Nihon University. Employed snare drum is SQ2 by SONOR Corp.. Employed motion capture system is the MAC 3D System of Motion Analysis Corp., which is an optical motion capture system. Recoding rate is 200 Hz with a shutter speed of 1 msec. The 30 markers in total are comprised of 23 markers in player's body and 4 for the edges of two drumsticks and 3 located on the drumhead are employed. Positions of markers are shown in Fig.2.

The set of data comprised of 30 points of 3 dimension M_{Epi} is obtained, where E means ID of emotion (E = T (tenderness), H (happiness), S (sadness), F (fear), A (anger), N (non-emotion), p means ID of performance (p =1, 2, …, 6), and i means ID of frame (i =1, 2, …, I). Several lacked data are interpolated by a linear interpolation method. As well, the sound is obtained by a microphone (SM-58, SHURE Inc.,) and a sound recorder (HD-P2, TASCAM Inc.,).

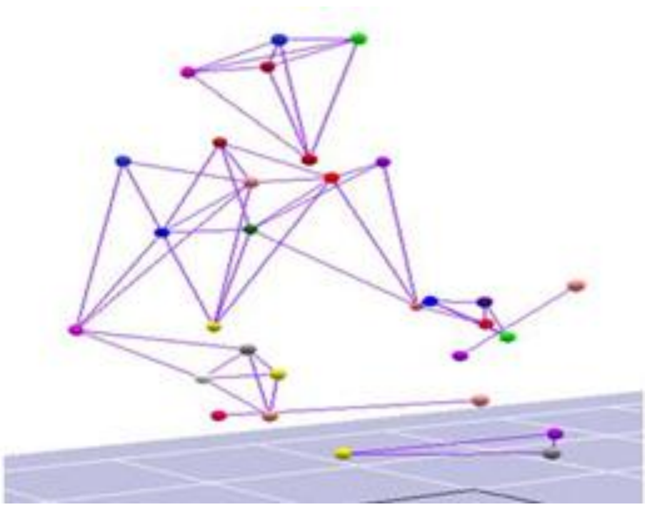

Figure 2. Position of marker.

III. MOTION AVERAGING METHOD

A. Outline

The motion averaging methods (MAM) provides an average of each motion from a viewpoint of time and space. In other words, the MAM conducts an adjustment for time axis and for the location of each maker on the 3-dimensional space. To do so, it requires adjusting position of players and expansion of motion data for time axis.

B. Averaging for the location of each maker on 3-dimensional space

Subtle differences due to the difference of position for player are usually shown in the recorded data of motion capture system. In order to adjust them, we obtain the average location of the three markers located on the drumhead, named as (x, y, z). Then, we assume the (x, y, z) as (0, 0, 0), and all the markers are shifted based on the amount of the shift of the average of three markers. After that, we need to consider the difference of angle of player. Then we obtain the angle by observing the average of two markers located in each shoulder and (x, y, z). Therefore we determine a standard from obtained 36 patterns and the other 35 patterns are rotated.

C. Averaging for time axis

We use the acoustic signal recorded with the motion capture data, and we estimate each impact time for recorded signal by using a previous method based on the spectral flux of obtained signal (Shimazu & Miura, 2011). Then the sequence of impact time O_{Epj} is estimated, where j means ID of impact (j =1, 2, …, J). Then we made a correspondence relation for each of O_{Epj} to a frame of motion capture data M_{Epj}. Since the motion data is recorded at a rate of 200 Hz, the impact time should be corresponded to the nearest frame of motion capture data. Then frames of motion capture data are relabeled as M_{Epj}.

We obtained the average of impact sequence, in other words, by using the O_{Epj}, we obtain the average impact time for each performance, which is named as O_j, and it is obtained as

$$O_j = \frac{1}{36}\sum_p O_{Epj}. \quad (1)$$

After obtaining O_j, we obtain the expanded or contracted motions by observing the average sequence of impact for all performances.

Firstly we obtain the IOI between two consecutive impacts, O_j and O_{j+1}, and all motion data is expanded or contracted for all impacts and for all performance. In other words, the data of M_{Epj} and M_{Epj+1} are expanded or contracted. This is done for all motion capture data. Then the modified motion capture data for all the recorded performance M_{Epk} is then obtained, where k means ID of frame (k =1, 2, …, K). Outline of expansion or contraction of motion data is shown in Fig.3.

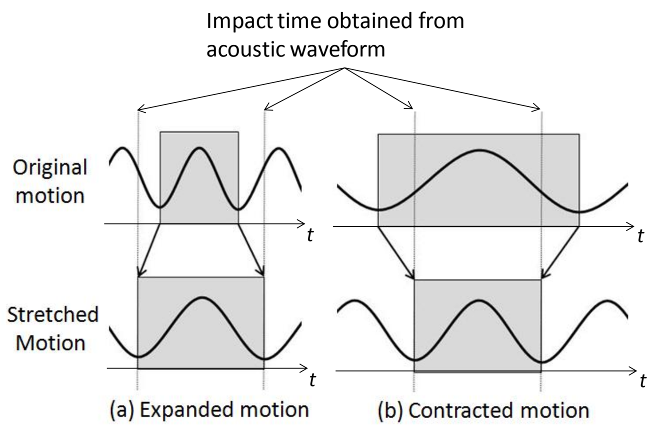

Figure 3. Stretch for motion data.

D. Averaged motion

The M_{Epk} contains no difference for position, angle, and impact times. Therefore the averaged motion for the performance is newly able to be obtained. Here we obtained the average of M_{Epk}, named as $\boldsymbol{M}_{\bar{E}pk}$, is obtained. The $\boldsymbol{M}_{\bar{E}pk}$ is represented as a matrix of 90 (30 markers for 3-dimension) and number of frame (K =4,119). We name it here as an "Averaged motion". The outline of obtaining the averaged motion is shown in Fig. 4.

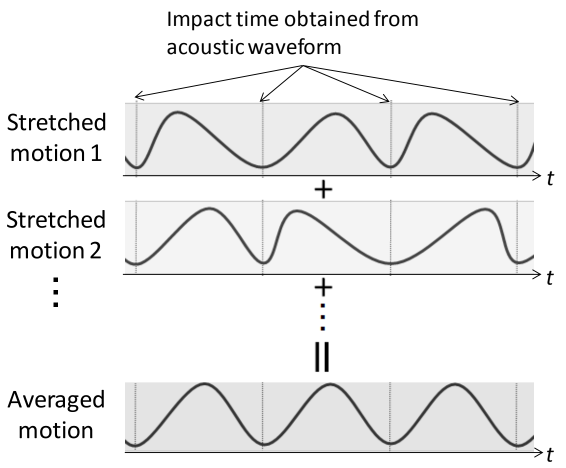

Figure 4. Calculation for averaged motion for each emotion.

IV. REPRODUCTION EXPERIMENT

A. Outline

Here tries to clarify the possibility of correctly informing the emotion to the listener by the averaged motion. Subjects are asked to answer each of five emotions for movie generated by the averaged motion. We also employ the recorded motion and video data. If the rate of correct answer is significantly larger than the chance level (20%, in this case), we can confirm the validity of averaged motion from a cognitive viewpoint.

We present the subjects each of stimuli such as recorded motion, averaged motion, or recorded video for five emotions. Thus the number of patterns is 15, comprised of 5 emotions and 3 types of stimulus. We present them randomly two times for each, and then the subjects are required to answer for given 30 patterns. Since the video is expected to answer for other stimuli, they are presented as final set of stimuli. Subjects are five students majoring informatics engineering.

B. Results

The correct ratio for recorded, averaged, and video are obtained, where each ratio is calculated based on the number of correct answer. All of the results for recorded, averaged and video are evaluated by the test of significant difference such as Tukey HSD, Dunnet T3, and Dunnet C test, so no significant difference among recorded, averaged and video is confirmed. Moreover, correct ratios for all of them are confirmed as significantly higher than the chance level. Therefore, we can say that the averaged motion is comparative to recorded and video motion in terms of informing the emotion to people.

V. EXTRACTION OF SPECIFIC FEATURES FROM THE AVERAGED MOTION

For the averaged motions, we extract manually 14 features, where the 13 features are concerning stroke motion of player, and the other is concerning the variation of motion among players calculated by the deviation from the averaged motion to each 6 patterns of motion. Statistical value for each features are obtained and listed for each emotion in Fig. 5 and Fig.6. As shown in Fig.5 and Fig.6, most of features correspond to the axis of activity. However, several features do not correspond to each axis of Juslins's emotional plane, but to specific emotion such as tenderness and fear. Therefore, by using the MAM, we can easily understand visually the characteristic of motion and obtain the features of motion.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we propose a method of averaging the motion capture data for emotional performance by a snare drum. For the data obtained by a motion capture system, we obtained an averaged motion by shifting and rotating in a 3-dimentional space, and expanding or contracting the data in time axis. Then the characteristic features of emotional performance are manually extracted by observing the averaged motion. So we can say that the characteristic of the motion can be visually confirmed by the averaged motion. Moreover, the result of the MAM is that the averaged motions are newly obtained and the feature of motions can be discussed.

Feature works are to do an experiment with a mixture of emotion, such as anger motion with sad sound for example, and we will obtain results of superior-inferior relation among audio-visual perception. Moreover, in case of recognizing emotion from music, which is dominant between sound and motion will be discussed in near future.

ACKNOWLEDGMENT

This work is partly supported by MEXT, Kakenhi (22700112).

REFERENCES

Dahl, S. (2006). Movements and analysis of drumming, In Altenmuller, E., Wiesendanger, M., & Kesselring, J. (Ed.), Music Motor Control And the Brain (pp.125-138), New York: Oxford University Press.

Davidson, J. W. (1993), Visual perception of performance manner in the movement of solo musicians, Psychology of Music, 21, 103-113.

Y. Imamura and Y. Tsukada (2005), Textbook of percussion for snare drum and bass drum, Zen-On Music Company Limited.

Juslin, P. N & Sloboda, J. A. (2001). Music and Emotion, New York: Oxford University Press.

Miura, M. (2012). Playability of electric snare drums based on rebound feature, Acoustical Science and Technology (in press).

Repp, B. H. (1999). , A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major, JASA, 105, 3, pp. 1972-1988.

Seashore, C. E. (1938), Psychology of Music. New York, McGraw-Hill..

Shimazu, S. & Miura, M. (2011). Proficiency estimation for CZERNY piano performance from acoustic signal, Proc. of APSCOM, pp.115-119.

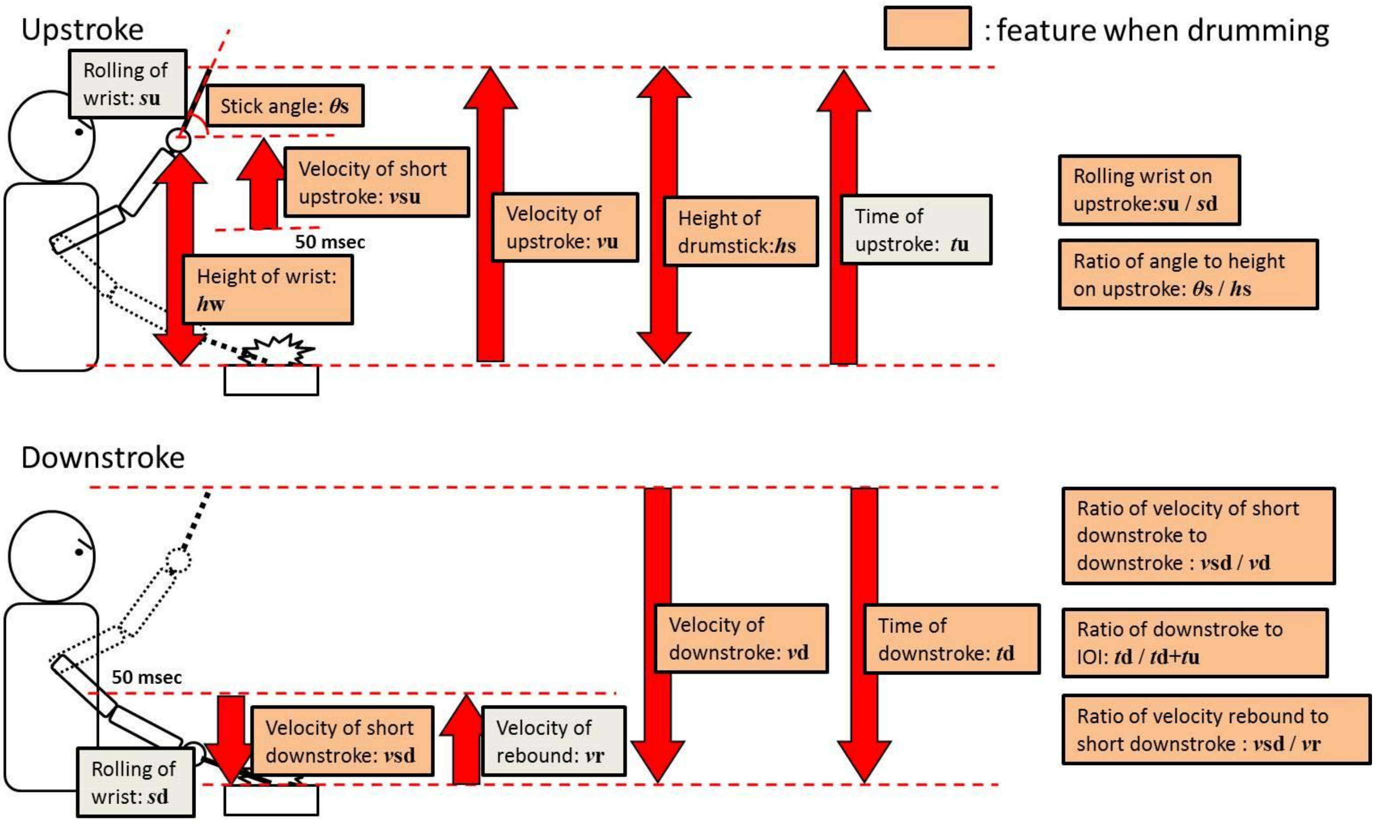

Figure 5. List of features..

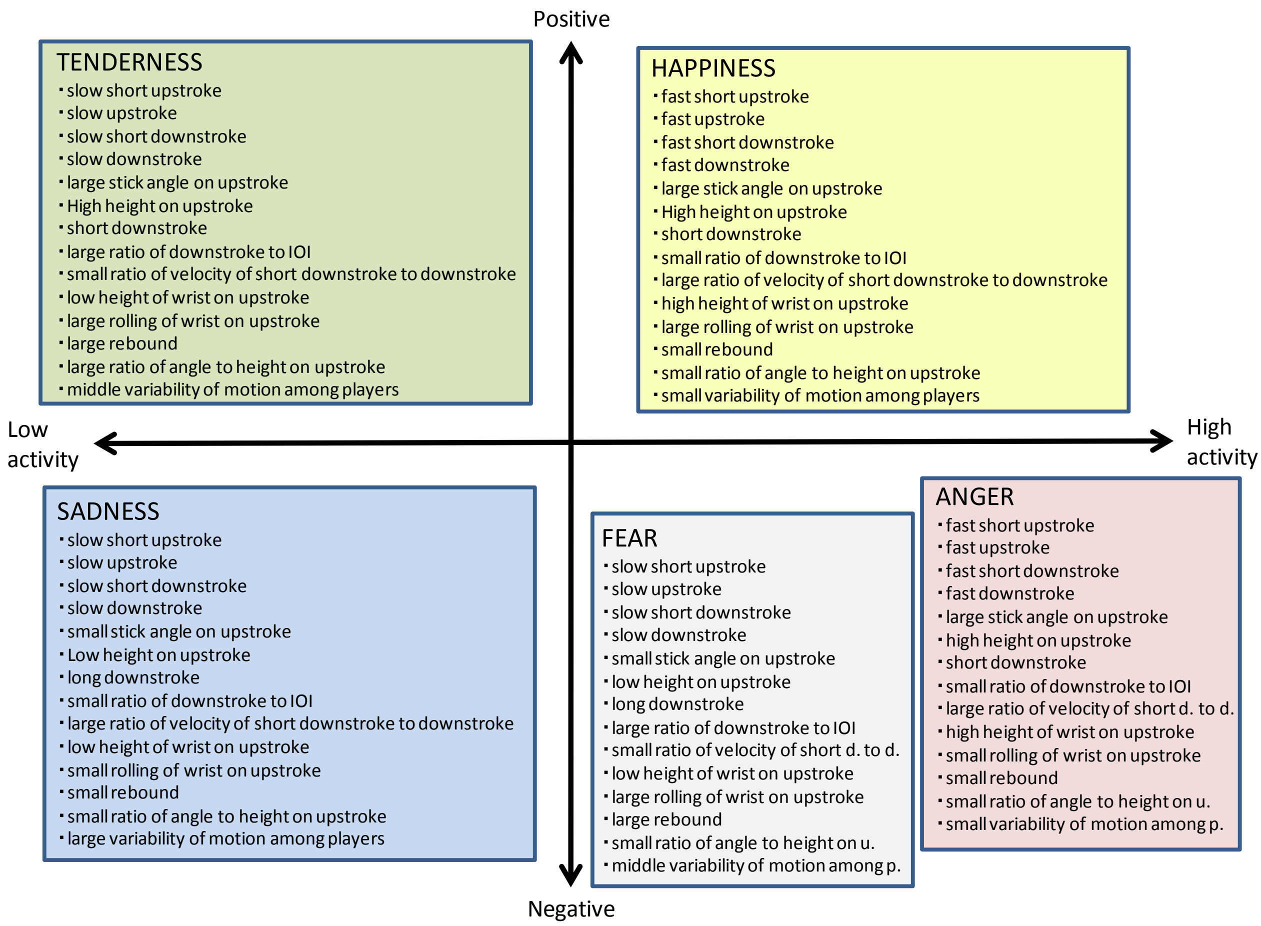

Figure 6. Movement features on snare drum.