Effect of Body Movement on Music Expressivity in Jazz Performances

Mamiko Sakata[1], Sayaka Wakamiya[2], Naoki Odaka[2], and Kozaburo Hachimura[3]

[1] Faculty of Culture and Information Science, Doshisha University,
1-3 Tatara Miyakodani, Kyotanabe City, 610-0394 Japan
msakata@mail.doshisha.ac.jp
[2] Graduate School of Human Development and Environment, Kobe University
inumonjiyaki@gmail.com, odaka@kobe-u.ac.jp
[3] College of Information Science & Engineering, Ritsumeikan University
hachimura@media.ritsumei.ac.jp

Abstract. In this study, we tried to examine empirically how body motion contributes to music expressivity, both in terms of intensity and manners, during impromptu jazz performances. Psychological rating experiments showed that music expressivity in jazz performances are assessed in two aspects, namely power and aesthetic quality. In the assessment of musical performances, the music itself basically contributed to how observers evaluated its expressivity. However, it was also shown that body motion had a greater influence on assessing the quality of music in terms of “hard or soft” and “light or heavy.”As a result of the three-dimensional motion analysis using motion capture, we learned that the characteristics of the player’s body motions changed with the playing style and the playing dynamics. The player, therefore, is making music not only by producing the “sound,” but by also showing “body motions” for creating that sound.

Keywords: Jazz Performances, Music Expressivity, Body Movement, Motion Capture.

1 Introduction

When performing classical music, the musicians share a general, predetermined understanding of the expression of their music and timing and play under the conductor’s direction. In impromptu jazz performances, however, the musicians begin to play their music only after agreeing on minimal requirements, such as the chord progression and structure of the musical composition they are about to perform. This means that each player must understand, on a real-time basis, the music expressivity of the other players and respond to or go along with them.

Many studies on music expressivity have been conducted concerning its technical aspects, such as performance methods and their relationship with the impressions aroused by music, but hardly any empirical research has been conducted up to this point concerning body movements as representative of music expressivity; in other

J.A. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2009, LNCS 5610, pp. 159–168, 2009.

words, the visual expression of body movements. Recently, some studies (those by Davidson [1], Okada [2], and Maruyama [3] and others, or example) have made some reference to the role played by the body in musical performances, but such works are sporadic, and the discussion has really only just begun.

This study aims to illuminate “the visual roles of body movements during impromptu performances of jazz music” and empirically show the modes and intensity of body movements that contribute to music expressivity.

For this purpose, we employ “Kansei” information processing techniques, motion capturing, feature extraction from motion data and some statistical analyses. “Kansei” is a Japanese word whose meaning is close to “feeling” or “sensibility” in English. Kansei information processing is a method of extracting features related to Kansei conveyed by the media we receive. Conversely, it is also a method of adding or generating some Kansei factors to media produced by computers [4].

We employ motion capturing techniques for obtaining images of human body motions. This technique is now used commonly in movie and CG animation production. Several systems are commercially available nowadays. This study uses motion capturing to analyze jazz performances and quantitatively analyze the roles played by body movements.

2 Study Subjects

For study subjects, we prepared materials with the help of professional jazz musicians in order to study the role of body movements in music expressivity. We asked a 10-year veteran male alto-saxophone player (24 years old) to play the jazz standard “Summertime.” He played the front theme[1], an ad-lib solo in the middle and the back theme for two choruses (with each chorus containing 16 bars). He was asked to play them in three different modes: “ordinary,” “expressionless” and “over expressive.” In this study, these three different modes of expression are defined as “expression dynamics.” In order to retain the characteristic feature of freestyle jazz performance, which is to perceive each other’s music expressivity in real-time, respond to them or follow them, we asked other players to join in the performance of our subject. The back band consisted of a drummer, a wood base player and a guitarist. The drummer was asked to keep the BPM=120 tempo, and all the other players were asked to follow the alto-saxophonist’s performance.

3 Motion Capture System

We used an optical motion capture system (Motion Analysis Corporation, EvaRT with Eagle cameras) to measure body movements during a jazz performance. Fig. 1 shows a scene from the motion capturing session in our studio. Reflective markers were attached to the joints on the player’s body, and several high-precision, high-speed video cameras were used to track the motion. In our case, 33+2(on the instrument) markers were put on the player's body (see Fig. 2), and the movement was

[1] The pre-composed part of a jazz number is called “theme”. In common jazz performances, musicians play the theme first, then the solo, and then the theme again.

measured with 10 cameras. The acquired data can be observed as a time-series using the three-dimensional coordinate values (x, y, z) of each marker in each frame (frame rate is 120 fps).

Fig. 1. Motion Capture

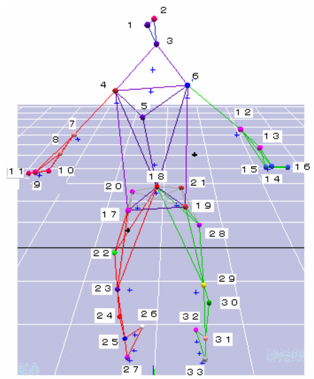

Fig. 2. Positions of markers

4 Psychological Rating Experiments

In order to determine what kind of impressions are perceived from jazz performances, the object of our study, and how the different modalities, namely the "sound" and the "body," contribute to musical expressivity, we conducted psychological rating experiments. Thirty-eight observers (20 men and 18 women) participated in this experiment. The mean and standard deviations of age among the 38 observers were 20.5 and 1.43, respectively. All observers had some training in jazz.

4.1 Stimuli for Experiments

The motion measured by the motion capture system described above was filmed by a digital camera (SONY) and then edited to produce experimental stimuli. The stimuli were obtained by editing the performances of the front theme and ad-lib solo played with three different expression dynamics, and then further edited using the three modalities of "sound only," "visual images only" and "sound and visuals." Table 1 shows the order in which the stimuli were presented (the modalities, dynamics and styles) and the length of each stimulus.

4.2 Procedure

We briefed the observers on the experiment, asked them to answer questions concerning their personal attributes and then presented the stimuli, one type at a time, in the order of "visual images only" -> "sound only"-> "sound and visuals." We provided an interval after showing one type of stimulus twice. In the first showing of the stimuli, the subjects were asked to closely and carefully observe the stimuli. In the second showing, they were asked to fill in the Answer Sheet using the Assessment Words on a scale from 1 to 7 for each word in the adjective pair, which are shown in Table 2[2]. The videotaped recording was temporarily stopped during the intervals between showing the different types of stimuli. The new stimulus was presented after making sure all subjects finished filling in their Answer Sheet.

Table 1. Order of stimuli

Order of display	Modality	Expression dynamics	Style	Duration
1	visual images only	ordinary	solo	64 sec
2	visual images only	expressionless	theme	66 sec
3	visual images only	over expressive	theme	64 sec
4	visual images only	over expressive	solo	62 sec
5	visual images only	expressionless	solo	63 sec
6	visual images only	ordinary	theme	65 sec
7	sound only	expressionless	solo	63 sec
8	sound only	over expressive	theme	64 sec
9	sound only	ordinary	theme	65 sec
10	sound only	over expressive	solo	62 sec
11	sound only	expressionless	theme	66 sec
12	sound only	ordinary	solo	64 sec
13	sound and visuals	over expressive	theme	65 sec
14	sound and visuals	expressionless	solo	63 sec
15	sound and visuals	over expressive	solo	62 sec
16	sound and visuals	expressionless	theme	66 sec
17	sound and visuals	over expressive	theme	64 sec
18	sound and visuals	ordinary	solo	64 sec

Table 2. Assessment Words

```
               1  2  3  4  5  6  7
loose        --+--+--+--+--+--+--   tight
soft         --+--+--+--+--+--+--   hard
powerful     --+--+--+--+--+--+--   weak
clear-cut    --+--+--+--+--+--+--   unclear
impressive   --+--+--+--+--+--+--   unimpressive
have presence-+--+--+--+--+--+--   no presence
neat         --+--+--+--+--+--+--   messy
plain        --+--+--+--+--+--+--   passionate
happy        --+--+--+--+--+--+--   sad
light        --+--+--+--+--+--+--   heavy
unique       --+--+--+--+--+--+--   ordinary
fantasy-like--+--+--+--+--+--+--   realistic
rich         --+--+--+--+--+--+--   poor
beautiful    --+--+--+--+--+--+--   ugly
fast         --+--+--+--+--+--+--   slow
warm         --+--+--+--+--+--+--   cold
subdued      --+--+--+--+--+--+--   bright
inarticulate--+--+--+--+--+--+--   articulate
favorable    --+--+--+--+--+--+--   unfavorable
good         --+--+--+--+--+--+--   bad
```

[2] In defining the Assessment Words used in our experiment, we referred to Iwamiya [5] and added our own 10 pairs of adjectives to create a group of Assessment Words and started our preliminary experiments using these words. After running a factor analysis, we deleted the terms that were not affecting any of the factors, deleted one of the highly-correlated pairs and finally selected 20 pairs of adjectives.

4.3 Results of Kansei Assessment Experiment

The results of the assessments of each stimulus were converted into scores from 1 to 7 using the SD method. We also obtained the average for the adjectives.

Extraction of KANSEI information from the stimuli. After conducting a principal component analysis based on the Kansei Assessment Scores obtained, we extracted two principal components with a characteristic value greater than 1. (The cumulative contribution rate was 0.879 up to the second principal component.) Table 3 shows the values of factor loading of each word pair to the two principal components. The shaded areas in the Table indicate the significant image word pair ratings to each principal component with a magnitude larger than 0.8.

Table 3. Results of PCA for the rating experiment

Assessment Words	PC1	PC2
loose-tight	.882	.122
soft-hard	.650	.025
powerful-weak	.959	-.141
clear-unclear	.948	.252
impressive-unimpressive	.964	.108
have presence-no presence	.974	.158
neat-messy	-.649	.737
plain-passionate	.932	.102
happy-sad	.749	-.524
light-heavy	.673	-.385
unique-ordinary	.974	-.076
fantasy-like-realistic	.859	.036
rich-poor	.789	.589
beautiful-ugly	.308	.942
fast-slow	.743	-.653
warm-cold	.618	.640
subdued-bright	-.701	.706
inarticulate-articulate	-.464	.853
favorable-unfavorable	.490	.846
good-bad	.468	.865
Eigenvalue	11.708	5.877
Variance (%)	58.539	87.925

Table 4. Result of multiple regression analysis

Assessment Words	Standardized Coefficients		Adjusted R^2
	Sound	Visuals	
loose-tight	0.851	0.204	0.931
soft-hard	0.593	0.601	0.897
powerful-weak	0.926	0.055	0.876
clear-unclear	0.942	0.081	0.932
impressive-unimpressive	0.999	-0.031	0.928
have presence-no presence	0.931	0.077	0.926
neat-messy	0.773	0.222	0.914
plain-passionate	0.841	0.193	0.959
happy-sad	0.617	0.433	0.847
light-heavy	0.587	0.656	0.959
unique-ordinary	-	-	-
fantasy-like-realistic	-	-	-
rich-poor	1.028	-0.047	0.991
beautiful-ugly	0.943	0.051	0.945
fast-slow	0.670	0.348	0.861
warm-cold	-	-	-
subdued-bright	0.657	0.405	0.910
inarticulate-articulate	0.687	0.341	0.969
favorable-unfavorable	0.915	0.102	0.898
good-bad	0.884	0.145	0.988

From Table 3, one can interpret the PC1 to be the variable concerned with the "power" of a musical performance, and PC2 to be the variable concerned with "aesthetic quality."

We plotted the principal components of PC1 and PC2 on the x- and y-axes and also plotted 18 types of stimuli on a graph (see Fig. 3). On this graph, the presence or power of a performance increases as you move toward the right, while the aesthetic quality increases as you move upward.

From the PCA results, it became clear that musical expressivity in jazz performance is perceived from two aspects, namely the "power" and "aesthetic quality."

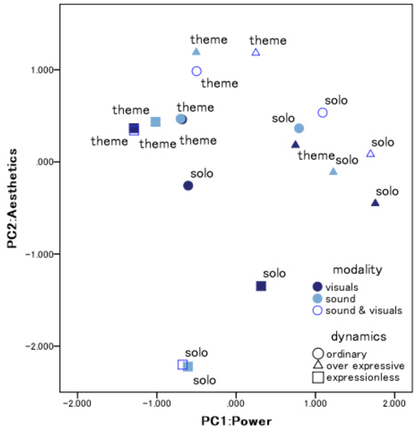

Fig. 3. Plot of PCA score for each motion

The Body and the Music in Music Expressivity. To examine which of the two factors, namely "body motions" and "sound," contributes more to music expressivity, we made a multiple regression analysis for each of the adjectives listed in Table 2, using Kansei Assessment Scores obtained by showing "visual images only" (in other words, body motions only) and "sound only" (in other words, music only) as independent variables. The Kansei Assessment Scores obtained by showing the "sound and visuals" were treated as dependent variables. As a result, we obtained a multiple regression equation ($p<0.05$) with a high contribution rate for many of the adjectives, as shown in Table 4.

Based on the information in Table 4, we see that music contributes more than the body motion (visual images) in the Kansei assessment, as expressed by the words "loose-tight," "powerful-weak," "clear-unclear," "impressive-unimpressive," "have presence-no presence," "neat-messy," "plain-passionate," "happy-sad," "rich-poor," "beautiful-ugly," "fast-slow," "subdued-bright," "inarticulate-articulate," "favorable-unfavorable" and "good-bad." On the other hand, the body motion (visual images) contributes more than the music in the Kansei assessment, as expressed by words like "soft-hard" and "light-heavy."

5 Feature Values for Body Motion

In the current study, the angles of each body part (back, sides, and knee), velocity (finger tips, elbow, sacral, head, and toe) and the distance moved on the floor (heel movement per unit time) were adopted as "feature values for body motion."

5.1 Extracting Physical Parameters

Angle. This parameter shows how the various body parts change in a time-series manner during musical performances. In our study, we measured the angles of the back, the body's side and the knee. In Fig. 2, the angle created by marker numbers 5, 18 and 31 shows the angle of the back. The angle created by marker numbers 7, 4 and

18 is the angle of the side of the body. The angle created by marker numbers 19, 21 and 31 is the angle of the knee. For example, in the case of the back, we set the origin at marker no.18 (x_2, y_2, z_2) and measured the angleθ created between marker no. 5 (x_1, y_1, z_1) and marker no. 31 (x_3, y_3, z_3). Then, using Equation (1) below, we calculated $\cos\theta$ and returned the obtained cosine radian to a radian using the Arc Cosine. Then we used the Degrees function to convert the angles in degrees into numerical values of the angle.

$$\cos\theta = \frac{(x_1 - x_2) * (x_3 - x_2) + (y_1 - y_2) * (y_3 - y_2) + (z_1 - z_2) * (z_3 - z_2)}{\sqrt{\{(x_1 - x_2)^2 * (y_1 - y_2)^2 * (z_1 - z_2)^2\} * \{(x_3 - x_2)^2 * (y_3 - y_2)^2 * (z_3 - z_2)^2\}}} \quad (1)$$

Velocity. This parameter shows the time-series change in the movement of the body parts during a musical performance. In the current study, we measured the speeds of the fingertips of the right hand (no.11), elbow (no.7), sacral (no.18), head (no.2) and the toe (no.27). For each marker, we obtained the Euclidian distance in the frames from the data expressed by the x, y and z coordinates. We then multiplied the Euclidian distance with the frame rate to obtain the time-series data of the velocity. For example, when the x, y and z coordinates of the marker in Frame i are expressed as x_i, y_i and z_i, we can obtain the distance d from Equation (2).

$$d = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2 + (z_i - z_{i+1})^2} \quad (2)$$

Then, by multiplying the d with the motion data frame rate, 120[Frame/sec], we can obtain the velocity $|v|$.

$$|v| = d * 120 \quad (3)$$

For the elbow, we used the relative coordinates by using the shoulders as the origin. The sacral were set as the origin to obtain the relative coordinates for the head and toe. This means that the velocity of the elbow, head and toe is expressed by a relative velocity based on the shoulder and the sacral. We obtained the velocity of the fingertips of the right hand and the sacral by using absolute coordinates using an origin determined in the capture area.

Floor Travel Distance. This parameter shows how much the player moved on the floor during the performance. We obtained the distance traveled by the left heel (no.32) frame by frame.

5.2 Feature Values for Body Motion

By conducting a principal component analysis after obtaining each parameter (raw data) described in Section 5.1 above, we extracted three components with characteristic values greater than 1. (The cumulative contribution rate up to the third principal component was 0.782.) From the values of factor loading shown by the Table 5, we can interpret that PC1 is the component showing the velocity of the upper part of the body, PC2 is the component showing flow travel distance, and PC3 is the component showing the bending of the body.

Table 5. Results of PCA for the motion capture data

Physical parameters	PC1	PC2	PC3
Angles of the back	-.088	.181	.801
Angles of the body's side	.240	.285	.601
Angles of the knee	-.396	.628	-.107
Speed of the hand	.925	.251	-.053
Speed of the elbow	.922	.278	-.093
Speed of the head	.922	.258	-.155
Speed of the sacral	.913	.065	.000
Speed of the toe	.883	-.312	.091
Floor travel distance	.422	-.744	.190
Eigenvalue	4.572	1.371	1.094
Variance (%)	50.798	66.029	78.186

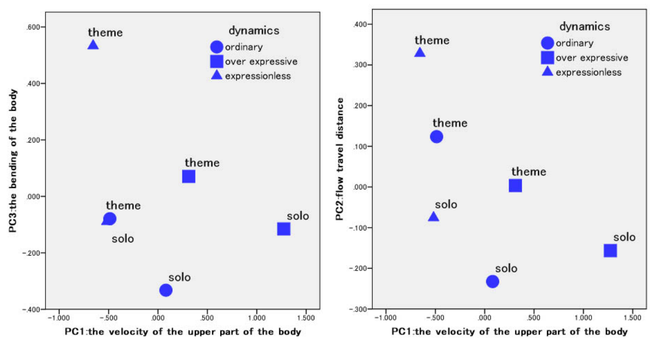

Fig. 4. Plot of PCA score for each motion by motion capture data

In the left graph of Fig. 4, the centers of gravity for PC1 and PC2 are plotted on the x and y axes, and six types of stimuli are plotted. In this figure, the more you move to the right, the greater velocity at which the top part of the body moves. The lower you go on the figure, the more distance covered in floor travel. Looking at each stimulus, the player showed greater floor travel when playing solo than when playing the theme. In terms of expression dynamics, the velocity of the top part of the body increased in this order: "ordinary" -> "expressionless" -> "over expressive."

Likewise, we plotted the centers of gravity for the principal components of PC1 and PC3 on the x and y axes and plotted each stimulus in a graph, which is shown in the right graph of Fig. 4. In this graph, the more you go to the right, the greater velocity of the top part of the body, and the lower you go, the more bending there is of the body. This graph shows that playing solo, rather than playing the theme, resulted in greater body bending. And when playing either solo or the theme, the player's body bending was greatest during the "ordinary" mode of playing.

5.3 Relationship between Kansei Assessment and Feature Values for Body Motion

We calculated the average and standard deviations for each stimulus for the nine parameters obtained in Section 5.1. In order to examine the relationship between the Kansei Assessment and the Feature Values for Body Motion, we calculated the coefficient of correlation between the principal component score of each performance obtained in Chapter 4 and the characteristic value of body motion of each performance(see table 6). The shaded area shows the combinations that showed a significant correlation of 5%. In the "Power Component," we found a significant correlation between all Power Components and body motion parameters, except for the parameters "average angle of body's side," "average foot travel distance" and "standard deviation of foot travel distance." For the Aesthetics Components, on the other hand, we could not find any correlation with any of the body motion parameters.

Table 6. Correlation matrix

Physical parameters		PC1 Power	PC2 Aesthetics
Angles of the back	Mean	-0.889*	0.130
	SD	0.880*	0.163
Angles of the body's side	Mean	0.377	0.040
	SD	0.879*	0.254
Angles of the knee	Mean	-0.949*	0.080
	SD	0.880*	0.211
Speed of the hand	Mean	0.894*	0.244
	SD	0.927*	0.242
Speed of the elbow	Mean	0.908*	0.241
	SD	0.954*	0.226
Speed of the head	Mean	0.931*	0.209
	SD	0.968*	0.194
Speed of the sacral	Mean	0.908*	0.247
	SD	0.944*	0.212
Speed of the toe	Mean	0.881*	0.256
	SD	0.859*	0.270
Floor travel distance	Mean	0.710	0.269
	SD	0.678	0.198

*$p<0.05$

This means that in musical performance, body motion contributes a large measure to the Power, but not to the Aesthetics.

6 Discussion and Conclusion

In this study, we tried to examine empirically how body motion contributes to music expressivity, both in terms of intensity and manners, during impromptu jazz performances.

Psychological rating experiments showed that music expressivity in jazz performances are assessed in two aspects, namely power and aesthetic quality. In the Kansei assessment of musical performances, the music itself basically contributed to how observers evaluated its expressivity. However, it was also shown that body motion had a greater influence on assessing the quality of music in terms of “hard or soft” and “light or heavy.”

As a result of the three-dimensional motion analysis using motion capture, we learned that the characteristics of the player’s body motions changed with the playing mode and the playing dynamics. The player, therefore, is making music not only by producing the “sound,” but by also showing “body motions” for creating that sound. It was found that body motions had a great role in creating “power,” but were not much related to the “aesthetics quality.”

Naturally, the Kansei emanated from the sound itself is central to music expressivity. However, we have shown empirically that the body motions people make when making music also contribute greatly in music expressivity. This study offers a basic examination of the role of body motions in musical performances; however, many challenges and problems still remain to be explored.

Acknowledgments. This work was supported in part by the Grant-in-Aid for Scientific Research, for Young Scientists (B) No. 19700493 of the Ministry of Education, Culture, Sports, Science and Technology, Japan. The authors would like to express their sincere gratitude to Mr. Junya Kondo for his cooperation with our research. Thanks are also due to Mr. Takahiro Yorino, Ms. Hitomi Toyoka and Mr. Naoyuki Okamoto for their kind help in motion capturing experiment.

References

1. Davidson, J.W.: Visual Perception of performance manner in the movements of Solo Musicians. Psychology of Music 21, 103–113 (1993)
2. Okada, A.: The body playing the piano. Shunjusha Publishing Company (2003) (in Japanese)
3. Maruyama, S.: The Embodied Sense of Music: Case Studies on the Rhetorical Function of Bodily Gestures by Highly Practiced Musicians. Cognitive studies: bulletin of the Japanese Cognitive Science Society 14(4), 471–493 (2007) (in Japanese)
4. Sakata, M., Hachimura, K.: KANSEI Information Processing of Human Body Movement. In: Smith, M.J., Salvendy, G. (eds.) HCII 2007. LNCS, vol. 4557, pp. 930–939. Springer, Heidelberg (2007)
5. Iwamiya, S.: Multimodal Communication of Music and Image. Kyushu University Press (2000) (in Japanese)
6. Iwamiya, S.: Design of Sounds. Kyushu University Press (2007) (in Japanese)
7. Nagashima, Y.: Drawing-in Effect on Perception of Beats in Multimedia. The Journal of the Society for Art and Science 3(1), 108–148 (2004)