Gazing Point Analysis of Experts and Beginners DJ for Acquiring Basic Mixing Skills

Kazuhiro Minami, Takayoshi Kitamura, Tomoko Izumi, Yoshio Nakatani

***Abstract* - In recent years, the number of people interested in club music has increased, along with the demand for equipment and schools for new DJ (Disc Jockey). However, in the small Japanese DJ industry, there are few skilled teachers to help beginners to improve their skills. In this research, we set up a propose that beginners can effectively improve their skills if they can understand the intention and actions of experts. We created a system that can record observer's viewpoint in a situation close to an actual DJ playing using a head mounted display with a high immersion sense. We found differences between experts and beginners gazing points.**

***Index Terms* – Learning Support, Eye Tracking, Virtual Reality, Self-study**

I. INTRODUCTION

In the field of sports and music, imitating and repeating the behavior of experts is a common practice among beginners to acquire some skills. However, in some fields that are not well known in Japan, people have less opportunities to develop teaching methods because of less population of stakeholders. Acquiring skills on DJ playing has same problems.

DJ have the skills to select music from sound source (i.e. CDs or records) and do a smooth and exciting transition between different tracks. Club music have increase in popularity in recent years in Japan. Additionally, the demand for equipment and schools of DJs have increased. DJ have been getting familiar with people in Japan in the past few years. However, a lot of beginners have to learn basic skills by themselves because there are DJschools in urban areas. In the field of DJ, there are some distinctive terms and required sensibility. Understanding “DJing” by self-studying is a hard task for beginners. Therefore, beginners often face a harsh environment to improve their skills. Beginners who get some equipment and begin DJ as a hobby often quit DJ soon after. Therefore, it is hard for companies to expand the market sell DJ equipment in Japan.

We proposed a system in which beginners will be able to proceed effectively if they can understand the intentions and action of expert’s gazing points. To verify that, we recorded multiple subjects gaze points and compare their gazing points with these from experts.

K. Minami, T. Kitamura and Y. Nakatani are with Collage of Information Science and Technology, Ritsumeikan University, 1-1-1 Nojihigashi, Kusatsu-shi, Shiga 525-8577, Japan(e-mail: K. Minami: is0237kv@ed.ritsumei.ac.jp,T. Kitamura: ktmr@fc.ritsumei.ac.jp, Y. Nakatani: nakatani@is.ritusmei.ac.jp).

T. Izumi is with Information Science and Technology, Osaka Institute of Technology, 1-79-1 Kitayama, Hirakata City, Osaka 573-0196, Japan. (e-mail: tomoko.izumi@oit.ac.jp).

II. RELATED RESEARCH

It is often said that efficiency is poor when expert guidance lacks the verbal skills to teach beginners. Hence, a lot of learning support methods specialize on specific instruments. Research is being carried out to improve the guidance efficiency by supporting musical score reading using gaze measuring instruments in piano performances [1]. Also, in drum performances, there is a research which creates an animation of the performer's viewpoint using augmented reality and visualize the gaze information of the performers by changing colors of the instruments playing to catch the attention to important objects [2]. However, these systems aim to teach beginners a basic way to play. However, in DJ mixing, there is no routines and the sensibility of the individual is regarded as an important asset to discover an original way to play showcasing the various experiences of the DJ by using senses such as vision and touch.

Along with the DJ trend in Japan, there are some research on DJ skills acquisition assistance using ICT (Information and Communication Technology). Ishisaki et al [3] proposed an automated DJ mix which requires expertise and a lot of equipment. They proposed system that allows unexperienced users to enjoy song mixing using songs provided by the users. This system is designed for users who have never done DJ mixing. In addition, there are other research which target people who already have experience as DJ. Tomibayashi et al [4] developed a wearable DJ system that uses sensors. DJs cannot move from the DJ booth where the equipment is installed so their performance movement is limited. This system enables DJs to do their performance easily even if they are away from the DJ booth by using wearable sensors and a gesture recognition technology that assigns arbitrary functions to the DJ equipment.

The work described above has been often explored. However, there are no studies to focus on assisting beginners who have few DJ experience. In this research, we compared the difference between the gaze points of experts and beginners DJs and we proposed a system that reproduces a DJ’s point of view in virtual reality.

III. SYSTEM OVERVIEW

Over repetitive, we record the gazing point information of a subject and created a measurement system using a head mounted display with a high immersion feeling. According to Sakata [5], obtaining a gaze measurement during exercise is difficult because the DJ is unable to fix its eyes to a specific point their limited movement range. In many cases, DJs often use their whole body during their performances, this is problem in the case of gaze measurement. We constructed a situation close to actual DJ play in VR (Virtual Reality). Users gaze measurement can be done without using any specialized eye gaze measuring equipment.

As shown in Fig.1, in the system there are two main videos

ISBN: 978-988-14047-8-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

recorded: the video directly taken from the expert DJ point of view and the operation of the software used by the DJ. By placing these videos on top of each other and displaying them through the screen on a Head Mount Display (HMD), we built a situation close to an actual DJ view in VR.

The position of the mouse cursor (center of the field of view) linked to the movement of HMD is measured. The upper left corner of the system screen (x, y) is set to (0, 0) by default. The measured coordinates and the playback time of the movie are stored in the database every second. In addition, the movement of the user's natural line of sight is measured and the mouse cursor which may hinder the visibility is not displayed. A total of 13 areas are set between the two videos in the system and the color of the area where the mouse cursor is positioned changes to white in order to show the user current viewing position. As shown in Fig. 2, areas are named depending their roles, and their names are recorded in the database every second along with the coordinates and the reproduction time (Mixer 2ch).

For example, in FIG. 1 the area "g" is the current viewing position.

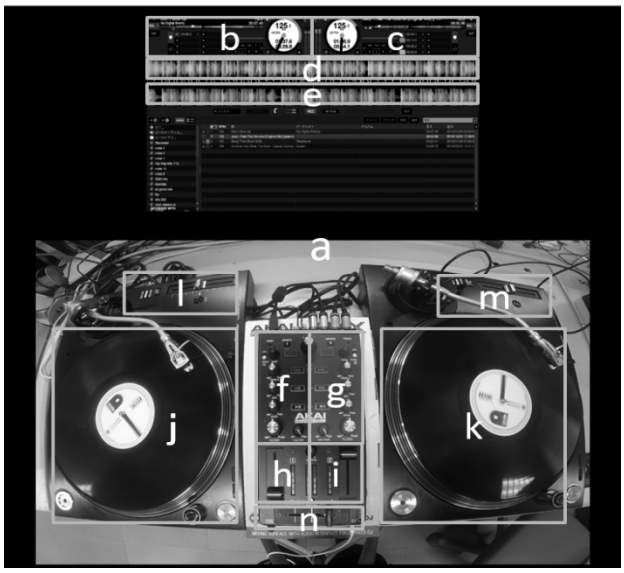

Fig.1 Screen displayed on the HMD.

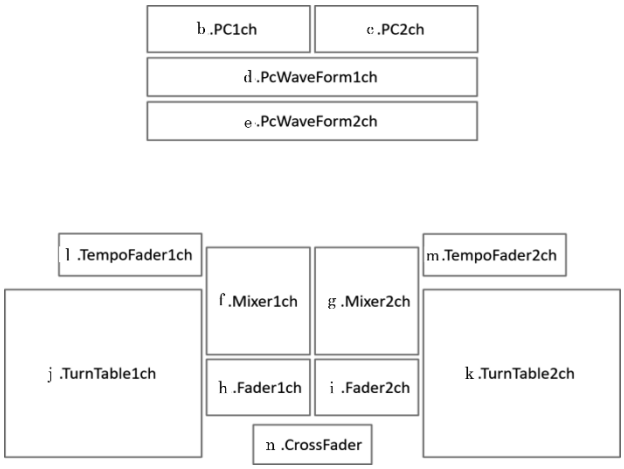

Fig.2 Names of each area.

IV. EXPERIMENT

By using the gaze point recording system proposed in the previous chapter, we conducted experiments to evaluate the differences between 5 beginners and 5 experts DJs. We asked them to wear the HMD as shown in Figure 3 and asked them to use the system standing. We recorded the areas they closely watched and their timing, as well as the screen of the system. Mixing is an operation in which specific parts of a song are connected to any other parts of a different song. In our experiments an experienced DJ performs mix operations three times in total. We then analyzed their gazing point tendency.

Fig.3 User with the gazing point analysis system

In general, a DJ repeatedly performs a mix operation. In the most basic mix operation (FIG. 4), the song to be played next is selected first, and the position at which the song starts is determined. Then, the DJ use the headphones to control the pitch of both songs while adjusting the volume to shift to the next song while fading out the preceding song and in some cases to jump on specific point of a song. There are various methods for this mix operation, and it differs from DJ to DJ and the genre of music played. In the experiment, to make an accurate comparison we adopted a basic mix operation in which any skilled DJ performs the same procedure.

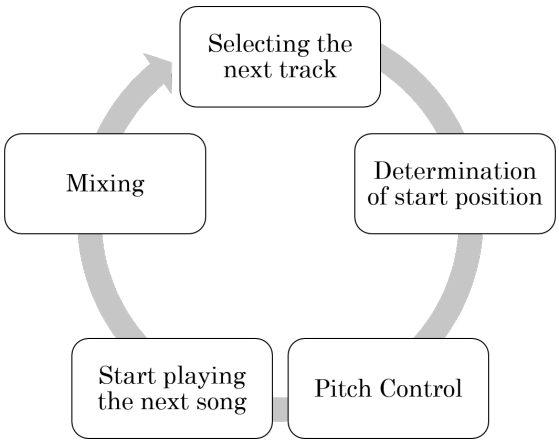

Fig.4 A basic DJ mixing

V. RESULT

The results of the experiment are compared and analyzed using two evaluation methods.

(1) Analysis of gaze point
- Extract the areas that users watch
  (Use N-gram analysis)
- Obtained the most viewed area
- Record the timing in which each area is gazed

(2) Comparison of the gaze points in an actual performance

ISBN: 978-988-14047-8-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

- Obtain the number of matches between the reference data created and the user gazing points data.
- Verify significant differences between expert DJs and beginner DJs.
  (Using Chi-square test)

In our experiments, we took 3 basic mixing operations that any skilled DJ performs exactly the same way as a point of reference. This data was created in collaboration with several experts DJ.

A. GAZING POINT EXTRACTION BY USING N-gram ANALYSIS

We used a gazing point extraction method based on the line of sight movement N-gram analysis [6]. Originally, N-gram is an analysis method used for finding the frequency of appearance of words in a character string (Fig 5).

We replaced each of the 13 area names in this system with one letter of the alphabet. In the database, the attention areas of the users are measured every second forming a string of characters. We define a gazing point as an area that was watched for more than 2 seconds by the DJ (a substring of two of more of the same characters). Using N-gram analysis on the gazing points is possible to obtain their frequency. In our experiments, we used the top 2 gazed points with the highest frequency.

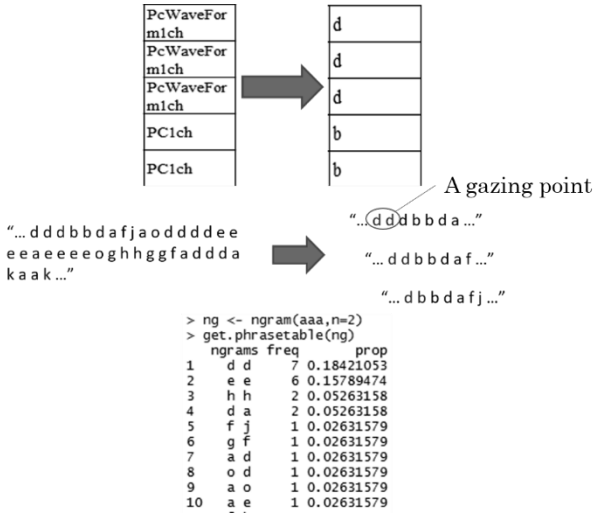

Fig.5 Flow chart of gazing point extraction by using N-gram analysis

B. ANALYSIS RESULT OF GAZING POINTS

N-gram analysis was used to find areas with a high degree of attention for each mixing operation performed three mixing operations. From the analysis we obtained the difference between experts and beginners.

In the first mixing operation, analysis is carried out with the two areas that had the most DJ attention time. The results are shown in Table 1. In the mixing operation, since it is necessary to match the tempo of the current song playing with the time and tempo of the next song, Experts E1, E2, E4, E5 who were familiar with the procedure, closely watched this area at the moment of selecting songs and determining their starting position. In addition to this, the area "f" (Fig 6) is used for mixing two songs, and by operating several knobs in the area "f", it is possible to change the volume to high, middle and low tones in the song. Experts E2, E4, E5 and the Beginner B3 watched this area in scenes where these knobs were operated in the video. Expert E3 frequently watched outside the studied areas and three beginners moved often between areas during the mix operation, therefore, they had no gazing point watched frequently.

Table1 Most frequent areas in the first mixing operation.

Experts	E1	E2	E3	E4	E5
Areas	e	f, d	None	e, f	e, f
Beginners	B1	B2	B3	B4	B5
Areas	e	None	f	None	None

Then we analyzed the area where the attention was the highest in the second mix operation. In the second mix operation, the mix operation is performed on the turn table contrary to the first mix operation. The role of each area also changes in the software. For this reason, the first time and second time, the area "d" showed the waveform of the next song the area "d" attracted attention of the users. Experts E1, E2, E4, E5 gazed at this area before playing the next song in order to confirm the match in the tempo of the two songs. Beginners B1 and B5 tended to go back between area "d" and other areas during the second mixing operation. Also, as in the first mixing operation, expert E3 gazed outside the available areas frequently and beginners B2 and B4 also moved often between areas and had multiple gazing points.

Table.2 Most frequent areas in the second mixing operation.

Experts	E1	E2	E3	E4	E5
Areas	d, e	d, e	None	D	d, g
Beginners	B1	B2	B3	B4	B5
Areas	d	None	f	None	d, e

In the third mixing operation, a basic mixing operation is performed in the same fashion as the first and second mixing operations, but because the speed is different between the songs, an additional matching process is necessary. By operating the pitch control provided at the top of each turntable, the speed of the song can be changed. In the software, the display area "b" corresponds to the pitch. Experts E1, E4, E5 and beginners B1, B5 gazed at area "b" where information such as the speed of the song is described. Also, in the third mixing operation many DJs closely watched the area showing the next song to be played. Fig. 6 shows the most gazed areas in the three mixes.

Table.3 Most frequent areas in the third mixing operation.

Experts	E1	E2	E3	E4	E5
Areas	b, e	e, d	e	b	b, e
Beginners	B1	B2	B3	B4	B5
Areas	b, e	None	None	b, f	e, d

ISBN: 978-988-14047-8-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

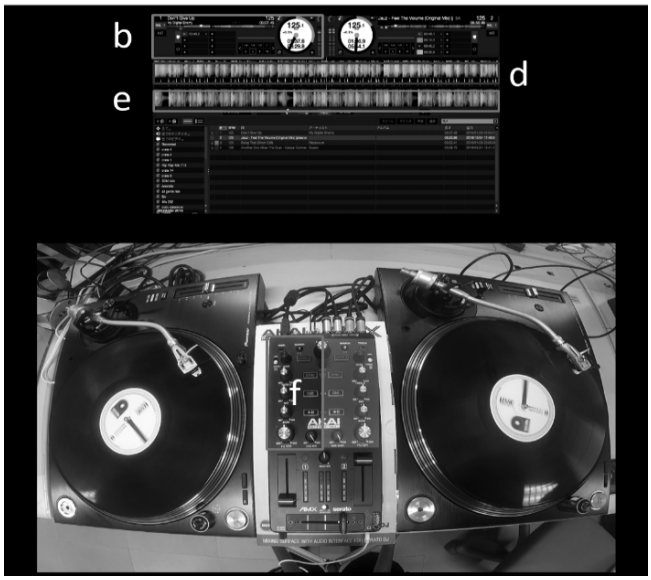

Fig.6 Areas with the most attention during 3 mix operations.

C. COMPARISON OF GAZING POINTS DURING ACTUAL PLAY

In addition to the first evaluation, we compared the users gazing points in each step with the gazing points when actually performing. We created a reference data of the basic mixing operation with 5 experts, and the areas to watch are set for each basic mix operation. In FIG. 7, we compared the users gazing points in each step with the reference data. In addition, we examined whether there was a significant difference between experts and beginners using the *Chi-square Test*.

Mixing Operation	Reference	E1	E2	E3	E4	E5	B1	B2	B3	B4	B5
Selecting the new track	e,o	e	b	o	e	o	d	a	f	c	a
	e,o	d	o	e	e	e	d	g	d	e	a
	e,o	e	e	d	e	e	d	f	o	k	o
	e,o	e	e	d	l	o	c	a	e	a	e
Determination of start position	l,b,c,e	a	e	e	a	l	c	l	e	l	o
	l,b,c,e	l	e	c	b	l	c	a	e	l	a
	l,b,c,e	l	o	c	b	l	b	a	o	l	e
	l,b,c,e	a	a	b	b	b	b	o	e	j	e
	l,b,c,e	b	a	e	b	b	e	d	o	j	e
	b,c,l	b	f	a	b	b	e	c	l	l	e
	b,c,l	b	g	o	b	b	e	b	a	l	o

Matches
Not matches

Fig.7 Comparing the users gazing points with the reference data.

Table 4 shows the total numbers that the users gazing points did not match with the reference data and the total numbers that match with the reference data.

Our results show that experts can see different gazing points when actually performing than when using the system. However, the number of unexperienced beginners were not consistent with the number of gazing points in the reference data. Also, when comparing experts and beginners, it is conceivable that the experiences of the DJs have a certain influence on the gaze tendency during system usage. These results were verified using a *Chi-square test* to confirm whether there is a difference could be confirmed between experts and beginners. A 5% difference was found between the first and third mixes.

Table.4 Result of Chi-squared test.

The First Mix Operation

	Number of matches	Number not matched
5 Experts	95	95
5 Beginners	64	126

p value = 0.03431 (P<0.5) Significant difference

The Second Mix Operation

	Number of matches	Number not matched
5 Experts	103	92
5 Beginners	78	117

p value = 0.1684 (P>0.5) No significant difference

The Third Mix Operation

	Number of matches	Number not matched
5 Experts	133	92
5 Beginners	80	145

p value = 4.938e-05 (P<0.5) Significant difference

In gazing point analysis using N-gram, there are areas commonly watched by many experts at the same time. Also, many of the beginners moved from area to area often during the mixing operation, and we had to record their gazing points a few times.

Subsequently, the gazing points in an actual performance, we found that experts also performed different movements when actually playing live than when using the system. However, in the first evaluation, there are areas commonly watched by experts at with same timing during the mixing operation, and since they match with the reference data, we can conclude that the gazing tendency is similar to the points considered important in the mixing operation. In addition, we could not find significant difference in the *Chi-square test* conducted. We had a problem with one of the experts in the experiment. The expert gazed by moving only his eyes, therefore, gazing areas were often registered outside, so it was impossible to measure adequately. We performed an additional test without one of the experts who had the lowest performance in our experiments (Table 5). Results improved significantly.

Table.5 Result of Chi-squared test without 1 expert.

The First Mix Operation

	Number of matches	Number not matched
4 Experts	88	64
5 Beginners	64	126

p value = 0.000489 (p<.05) Significant difference

The Second Mix Operation

	Number of matches	Number not matched
4 Experts	97	59
5 Beginners	78	117

p value = 0.001887 (p>.05) Significant difference

The Third Mix Operation

	Number of matches	Number not matched
4 Experts	114	66
5 Beginners	80	145

p value = 3.181e-06 (p<.05) Significant difference

D. QUESTIONAIRES AFTER EXPERIMENT

Questionnaires were given to the users after the experiment. The questionnaires asked the following questions.

(1) Gazing areas where the users were particularly

ISBN: 978-988-14047-8-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

interested during the experiment
(Selection formula, multiple answers possible)

(2) How they learn DJ playing.
(3) What was the hardest thing when they started DJ playing.

We also asked the experts the following points.

(4) Existence of teaching experience to beginners
(5) Points that they felt difficult to teach.

We compared the answer of question No.1 (Fig. 8) with the area that each user closely watched based on the gazing point information recorded while using the system. As a result, for the five experts, 16 answers were matched out of 22 answers. On the other hand, only five beginners matched on 7 out of 20 answers. From this, we concluded that experts can have intentional gazing points when using this system, compared with beginners. Also, the most frequent answer to Question1 was the waveform information of the song, which is consistent with the result of the first evaluation using N-gram analysis. Results 3 and 5 in Fig. 8 correspond to those parts.

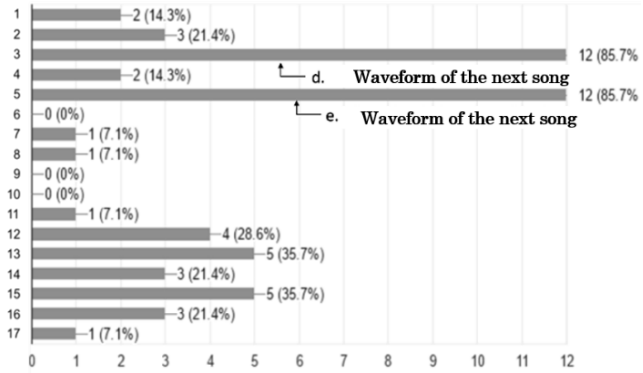

Fig.8 Answers of question 1.

To the second question, experts responded that they learned by watching other DJ play at events, self-study, or learning from experienced people. 2 beginners answered that they were learning by themselves and the other 3 were learning from experienced people. When comparing these questions with the attributes of each user, many users have acquired experience participating in club events, dance events, etc., and they have often see experienced DJ and in environments susceptible to the guidance from experts. Although some DJs have motivation for improvement, two beginners did not have the basic skills because they were not in an environment where they can easily receive some guidance.

This shows the necessity for support. In Question No.3, 7 out of 10 users responded “during the mixing operation”, and did not understand the sense of the bar and beat of the song, so they found out that they had difficulties with the timing necessary to synchronize the next song. Moreover, referring to question No.5, most experts felt that they had difficulty teaching using musical sheets to beginners.

The results of these questionnaires showed the importance of gazing at the waveform information. We demonstrated that beginners can carry out an efficient self-study if they were able to learn the sense of rhythm and the use of musical sheets.

VI. CONCLUSION

In this research, we compared gazing points of DJs watching other DJs performance gazing, we developed a gaze point measurement system using HMD with a high immersion sense and conducted experiments on 10 subjects. We created a comparative analysis using two evaluation methods. As a result, we found some interesting discoveries worth examining in future support methods. In the future, we consider developing a support system aimed to teach basic skills acquisition of DJ for beginners.

REFERENCES

[1] Wataru Toma, Katsuko T.Nakahira: descriptive analysis of process for acquiring piano playing skill by means of eye-movement tracking while reading music scores, The 11th Forum on Information Technology (2012).

[2] Kazuki Hayakawa, Dai Hasegawa, Hiroshi Sakuma: Performer's Viewpoint on Drum Performance Support Effect of Visualization of AR Animation and Gaze Information, The 76th Information Processing Society of Japan (2014).

[3] Hiromi Ishizaki, Keiichiro Hoashi, Takishima Yasuhiro: Automatic DJ Mixing System Based on Measurement Function of User Discomfort on Music, Forum on Information Technology Journal, Vol.52, No.2, pp. 890-900 (2011).

[4] Yutaka Tomibayashi, Yoshinari Takekawa, Tsutomu Terakawa, Masahiko Tsukamoto: Development and Actual Use of a Wearable DJ System using Wearable Sensors, IPSJ special Interest Group on Music And Computer (MUS), pp. 39-44 (2008).

[5] Mamiko Sakata: One’s Eyes are Often More Eloquent than One’s Mouth: Study Cases on Eye Movement Analysis, SCBDMMT Kobe University, 6(1), pp. 103-116 (2006)

[6] Hironobu Takagi: Recognizing User’s Uncertainty on the Basis of Eye Movement Patterns: A Step toward an Effective Task Assistance System. Forum on Information Technology Journal, vol. 41, No. 5, pp.1317-1327 (2000).

ISBN: 978-988-14047-8-7
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)