Visualization of Swiping Motion of Competitive Karuta using 3D Bone Display

Risa Kitagawa
Ochanomizu University
Tokyo, Japan
kawanishi.mami@is.ocha.ac.jp

Takayuki itoh
Ochanomizu University
Tokyo, Japan
itot@is.ocha.ac.jp

Abstract—Each player of competitive Karuta has a unique posture and a style of swiping the cards, which are important factors in winning a game that requires the swift acquisition of all 50 cards. It is essential to analyze the characteristics of a player's posture in order to achieve the swift acquisition. While previous studies on the analysis of posture of competitive Karuta required special devices to be attached to the body for measuring brain activity or analyzing wrist acceleration to determine the speed of a player. Such approaches were problematic since they may interfere with the natural movement of the players.

To address this issue, we are developing a visualization system that displays the three-dimensional skeletal movements of multiple players in response to a read the poetry of Karuta. We utilize skeletal information extracted from videos using Mediapipe, developed by Google, to compare and analyze the movements of different players. This paper presents the results of our analysis, which compared the swiping movements of multiple players using our proposed system.

Index Terms—Visualization, Competitive Karuta, Motion capture.

I. Introduction

Competitive Karuta [1] is a one-on-one game played with 50 randomly selected cards out of a set of 100 Ogura Hyakunin Isshu Karuta cards. The winner is the player who has no cards in their own territory.

It is essential to analyze own playing style to win in competitive Karuta, as players must pick up the fifty cards on the playing field faster than their opponents. Players can improve their speed and accuracy by analyzing their performance, such as their strengths and weaknesses in card position, posture, and swiping technique.

Many players analyze their playing style by exchanging subjective advice with opponents or instructors; however, this method has two problems. First, human observation does not always provide accurate knowledge, as individual senses are prone to various biases. Second, it is usually difficult to compare oneself with others, as the only way to do so is to compare one's own swiping motion with two-dimensional video information.

To address these issues, we are developing a system that enables players to compare their own actions with those of others by measuring and visualizing their swiping motions

[1] Rules in English: http://karuta.game.coocan.jp/simplerule-e.html

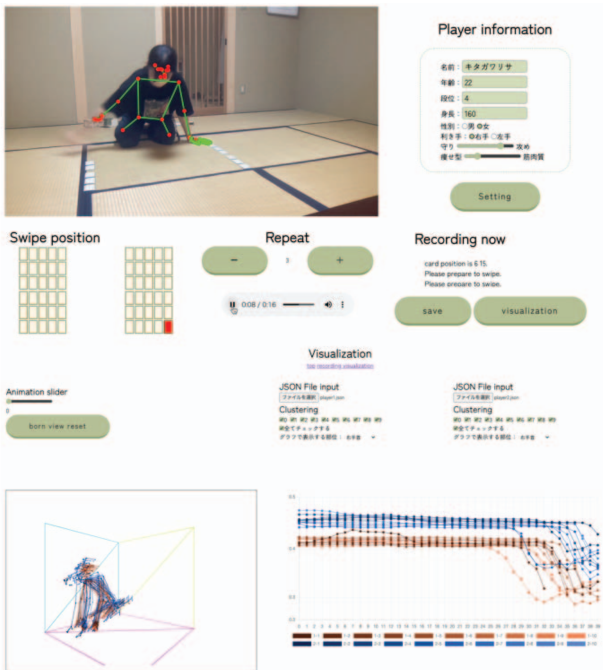

Fig. 1. Screenshot of the proposed system. (Upper) The system for measurement of swiping motion. (Lower) The system for visualization of swiping data.

in competitive Karuta games as shown in Figure 1. Our proposed method involves taking multiple videos of swiping motions of players at different positions and extracting motion information from each body part. This motion information is displayed using an interactive visualization system that features a 3D bone display of the motion of the players and a chart of the positional information of each body part. By comparing and analyzing the results of these visualizations, differences between multiple players can be evaluated.

This paper presents the results of a comparison of swiping motions between multiple players and examines the differences that can be observed. Specifically, we discuss a comparison between two expert players of different body

2375-0138/23/$31.00 ©2023 IEEE
DOI 10.1109/IV60283.2023.00065

sizes and genders, as well as a comparison between a beginner and an expert player.

II. Related Work

A. Studies on competitive Karuta

Studies on competitive Karuta are still few in number because they have been conducted only in Japan. This section introduces two studies on data collection and analysis of competitive Karuta.

Takeda et al. [1] used an optical brain imaging system to measure brain changes in competitive Karuta players. The results showed that the concentration of oxygenated hemoglobin increased from the beginning of the reading of the bottom phrase to the end of the reading of the top phrase, as well as when the cards on the field were read. This allows us to observe the information processing cycle of the brain that repeats the reception, recognition, processing, and agile movement of sophisticated auditory information when players swipe the cards.

Yamada et al. [2] measured the timing of card acquisition in competitive Karuta by attaching acceleration and angular velocity sensors to the wrist. Acceleration and angular velocity of the swiping collected in advance, and the timing of swiping the cards at that time are used as correct data. Then, DTW (Dynamic Time Warping) distance between the test and correct data is calculated. The time when DTW distance is the shortest is calculated as the acquisition time of the card in the test data. The difference in acquisition timing between players is several tens of milliseconds. But even, this method could estimate 99.0% of cases with an error of 20 milliseconds or less.

These methods are similar to our method in that they measure human body movements and analyze competitive Karuta as a sport. Differently from these methods, our method aims to visualize the movements of the whole body during a game so that users can understand the problems of the players' own motions, not only aiming to measure the timing of swiping a card.

B. Study on Skeletal Extraction

Various technologies have been developed to extract skeletons from the human body for various purposes.

1) Motion Capture using a Camera: Motion capture technology that utilizes infrared cameras and machine learning to extract human motion has been popular. Microsoft's Azure Kinect is an example of such systems, which uses a markerless motion capture system to capture human motions. Azure Kinect has advantages such as high positional accuracy, low cost, ease of installation, and tracking without using markers or sensors.

Kawanishi et al. [3] used Azure Kinect to acquire motion data from repetitive practice of dancers and visualizes the results by applying clustering to each body part. Melios et al. [4] also used Azure Kinect to capture motion data of dance movements, and have created a system that allows users to enjoy a game while practicing dances.

2) Pose Estimation from Video Images Using Machine Learning: Posture estimation is another technology that has been developed recently, which estimates the positions of human joints from ordinary two-dimensional video images by machine learning, without the use of specialized equipment such as markers, sensors, or infrared cameras. This technology has the advantage of being able to estimate the posture of multiple persons and can be applied to videos recorded without the purpose of motion capture. Thanks to the above convenience, we can use in a variety of fields, including team sports such as soccer.

Bridgeman et al. [5] applied a fast method to two videos of a soccer match to improve the accuracy of whole-body pose estimation and tracking of multiple players. Ghasemzadeh et al. [6] proposed a method for detection of skeletons of overlapping basketball players.

Although we do not study on machine learning of videos itself, we would like to explore methods for motion extraction from video images to improve the accuracy of posture estimation for competitive Karuta.

III. Processing Flow

This section presents the proposed visualization system consisting of the following two processes.

- Measurement of swiping motion
- Visualization of swiping data

In the first step (step 1), we record the video of the swiping card motion, and extract the skeletal information to obtain positional data for each body part. This data is referred to as "swiping data". The second step (step 2) involves visualization of the "swiping data" to observe the swiping motions of the players. Specifically, Sections 3.1 and 3.2 outline the method used to acquire the "swiping data", while Section 3.3 provides more detailed information about the system used to extract the "swiping data". Finally, Section 3.4 explains the visualization system.

A. Measurement of swiping motion

This section explains the method for measuring the motion of the swiping cards in competitive Karuta. In this study, we use pose estimation by machine learning from video footage because it has the advantage of collecting information on a large number of players, since the skeleton can be extracted if only video footage can be captured without using dedicated equipment. Although motion capture using Azure Kinect was also attempted, the accuracy of the measurement results in the swiping motion did not differ significantly from that of pose estimation from video footage. Therefore, we concluded that posture estimation from video footage was a reasonable solution.

At first, we intended to extract the skeleton of the entire body. However, both posture estimation from video footage and measurement using Azure Kinect had low detection accuracy for the lower body. This is due to the nature of the competitive Karuta posture, such as sitting in “seiza”, which obscures the lower body from the camera and clearly reduces accuracy. Since observing the swiping motion mainly focuses on the movements of the upper body, arms, and hands rather than the lower body, we decided to proceed with the study mainly analyzing the upper body.

As a result, we used Mediapipe developed by Google for pose estimation. It is lightweight compared to other pose estimation models. We tried other pose estimation models, but some of them were not compatible with the performance of personal computers or produced inaccurate results. We aim to collect information from more players in the future, so we decided to use Mediapipe, which can be run on a web browser regardless of the runtime environment.

Mediapipe can estimate x, y, and z three-dimensional coordinates from a single camera input. This enables the visualization of bones in three dimensions, making it possible to observe one’s own posture and movements from various angles and distances.

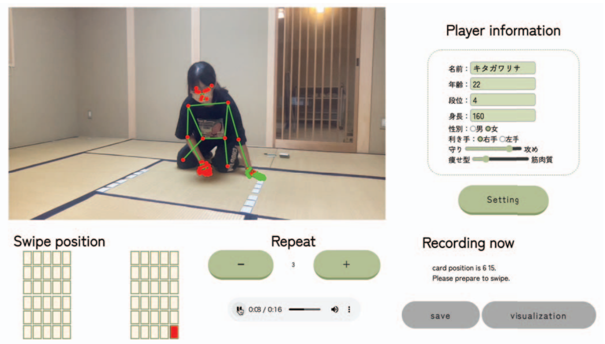

Fig. 2. Screenshot of the system for measurement of swiping motion. It consists of five components.

We developed a JavaScript-based Web system to obtain the swiping data as shown in Figure 2. The system is composed of the following five components.

1) Display of video and bones
2) Input form of player information
3) Selection panel for the position
4) Control panel for voice reading
5) Message box displaying current status

This system displays the bones in real-time on the screen alongside the video captured by the webcam. The system also provides a form for entering player information to distinguish each player on the right-hand side of the video display. Once all fields are filled out, the player information is saved by clicking the “Setting” button.

We suppose a competition field of 87 cm × 44.5 cm is set on the “Tatami”, which is the same size as the space used for the competition, and cards are placed at arbitrary locations, while measuring the player’s motion. The user is free to select where to place the cards on the system. The location information of the swiped cards obtained here is recorded together with the swiping data, in order to apply clustering in the visualization process described below.

The system starts reading the “Waka” poems and preparing to record them simultaneously, when the “Playback” button is pressed. Only one type of poem is read, and the cards are repeatedly read as many times as set by the counter on the control panel. The poems are read randomly in actual competitions, but in this method, we suppose the same poems are read repeatedly because the focus is only on the body of Waka poems that can be determined by listening to the first character (known as the “one-character rule”), and the participant is instructed to swipe the card at the moment they are read. Recording of body position information starts one second before the first character is read and stops three seconds later to minimize recording time and data size.

The “Save” button is activated once the number of swiping data specified by the counter has been measured. The swiping data, including player information, is exported in JSON format and saved as a local file.

B. Visualization System

The visualization system comprises four major components, as shown in Figure 3.

- The Bone Display Control Panel (upper left)
- The File Loading and Clustering Panel (upper right)
- The Bone Display (bottom left)
- The Time-Series Display (bottom right)

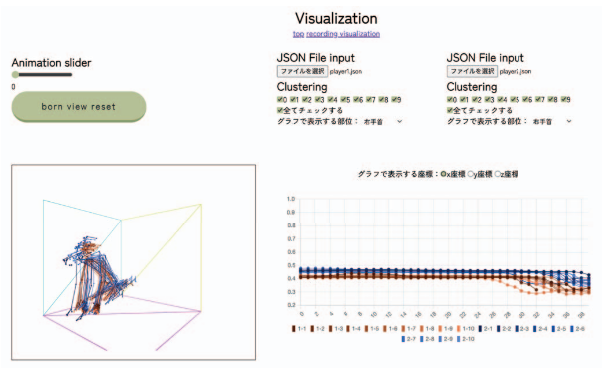

Fig. 3. Screenshot of the visualization system. It consists of four components.

The system is composed of two processes: the visualization process and the clustering process. In the visualization process, the system first loads the JSON files for the two players by clicking on the buttons located on the right side of the upper control panel. After loading is complete, the system displays the bones and line charts.

Users can animate the motion of the player’s swiping by sliding the bar on the upper side of the Bone Display Control Panel. The skeletal information for the same spot is superimposed, allowing users to compare and observe the movements of the entire body. The system can display bones from an arbitrary angle or position by dragging or scrolling with two fingers on the canvas of the bone display.

The Time-Series Display screen shows temporal changes in the coordinate values of each body part. The horizontal axis shows the time (frame ID), and the vertical axis shows the coordinate values of the body part selected by the radio buttons on the graph. Since Mediapipe acquires coordinates in three directions (x, y, z) and outputs normalized values between 0 and 1, the vertical axis of the line chart ranges from 0 at the minimum to 1 at the maximum. Like the Bone Display, users can scale and shift the line chart by dragging it on the canvas or scrolling it with two fingers.

In the Clustering Panel, users can select the motion data they want to display from among all the measured swiping data, and they can freely change the display and compare the line charts.

The system draws the data of the first player in orange, and the data of the second player in blue in both the Bone Display and the Time-Series Display, to differentiate between the two players. The system represents the time series of the motion of each player by a gradation of each color. In other words, the measurements near the beginning of all measurements are drawn in darker colors, while those near the end are drawn in lighter colors.

IV. Examples

This section presents the two examples of results using the proposed method.

A. Comparison of Expert Players

We first compared the stance and swiping of two expert players to find any differences. Both players were right-handed, had played competitive Karuta for over five years, and were "A-level 4-dan" players. Player 1, represented in orange, is a 160-cm-tall woman, while Player 2, represented in blue, is a 180-cm-tall man. We examined whether differences in playing style could be observed between two players of similar ability based on their physique and gender. During the experiment, the players were filmed while swiping a card in the lower right corner of their own territory side. The results of the experiment are presented below.

1) Analysis of the Bone Visualization: Figure 4 shows an example of the bone visualization of the two players. The most significant difference between the two players was the height of their heads. As shown by the circles in Fig. 4, the head of Player 1 was taller than the head of Player 2, despite Player 2 being taller overall.

Bone visualization is a useful tool for displaying the swiping motion data in competitive Karuta and for comparing multiple players. Moreover, three-dimensional bone visualization allows for the observation of angles that are not normally visible, enabling more detailed data analysis. This method is particularly effective for comparing the swiping actions of players across multiple rounds, which is difficult to achieve with ordinary video.

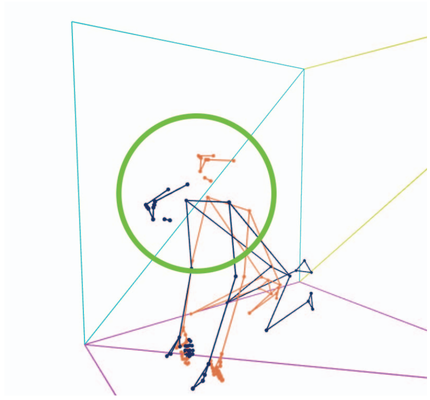

Fig. 4. Comparison of head height.

2) Analysis of the time series: This section presents a more detailed discussion by comparing the line charts while observing the bone display. We selected one or two movements from ten movements of each player and compared them by changing the display. Since positions were not extracted for areas where the accuracy of the Mediapipe estimation was less than 95%, the line charts may be incomplete depending on the measurement results. Therefore, in this section, we selected movements with no deficits in the line chart as the first priority.

We showed that the most significant difference between the movements of the two players was head height in the previous section. This is evident from the line chart of the y-coordinate of the nose as shown in Figure 5. Although we had expected head height to vary in proportion to height, this expectation was overturned since we found the variation according to playing style.

Next, we compared the right wrist, which moves the most during the swiping action. In this experiment, we focused on the "beginning" and the "end" of the swipe, but the timing of the "beginning" of the swipe varied among the players, and no significant difference was observed between the two players. On the other hand, there was a significant difference between the movements of the two players at the end of the swipe. The following is a discussion of the posture at the end of the swipe.

As shown in Figure 6, Player 1 raises her right hand while Player 2 places his right hand close to the floor when he finishes swiping. We found that player 1 swung her arm up like a pendulum using her shoulder as a fulcrum, while player 2 swung his arm in a compact motion, as if he moved parallel to the right, from the bone animation display.

Fig. 7 shows the line chart of the y-coordinate of the right wrist. This line chart illustrates that both players moved upward in the same manner up to frame 34, but

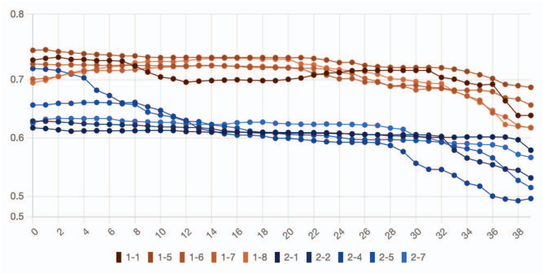

Fig. 5. Comparison of y-coordinates of nose.

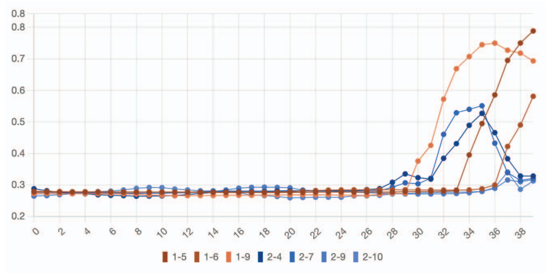

Fig. 7. Comparison of y-coordinates of the right wrist.

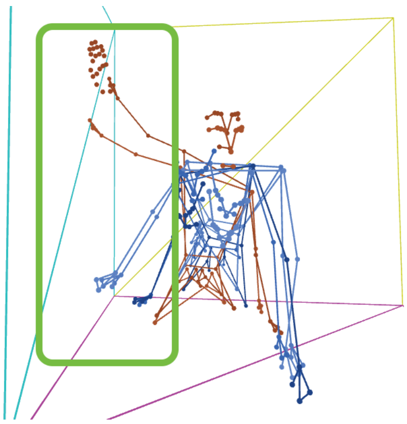

Fig. 6. Comparison of right hand height.

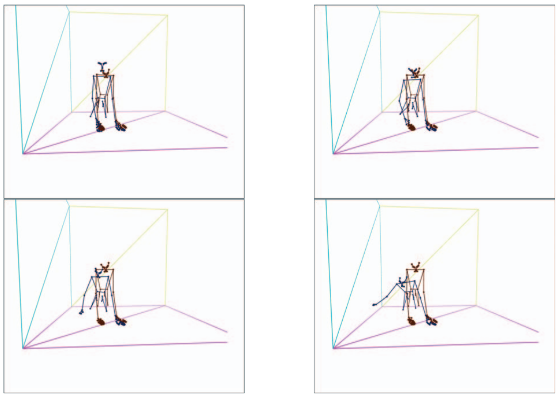

Fig. 8. Animation (1). (Upper-left) Before movement. (Upper-right) Start of movement of an expert. (Lower-left) Start of swiping of an expert. (Lower-right) End of swiping of an expert.

the right hand of Player 1 continued to move upward, while the right hand of Player 2 moved as if it were returning to the floor.

3) Summary of Analysis: From the above results, we found that Player 1 swung her arms up after swiping, while Player 2 landed on the floor with a compact motion.

As described above, we found that there were differences in movements among players of the same A-level 4-dan, depending on their physique and gender, by using the proposed system in combination with the bone display and line charts.

B. Comparison of Beginner and Expert Players

We conducted a comparison between beginner and expert players to investigate the differences in their stance and swiping movement. Player 1 (drawn in orange) was a right-handed woman, 155 cm in height, had little or no experience in competitive Karuta, and was considered a beginner. Player 2 (drawn in blue) was also a right-handed woman, 160 cm in height, and had over five years of experience in competitive Karuta as an A-level 4-dan player.

1) Comparison while swiping the lower right hand side of own territory: As in the previous experiment, we recorded the players swiping the lower right corner of their own territory, and generated a sequence showing the players with their bones displayed, as shown in Figure 8. We could observe in detail the differences in their movements by superimposing the bone of the two players and comparing them frame by frame.

Figure 8(upper-left) shows the two players before they start moving. We can see that Player 2, the expert player, has a higher head than Player 1. However, as we saw in the comparison of expert players in the previous section, there are individual differences in head height between the two players, suggesting that head height is not related to experience.

Player 1 had previously studied the posture of Player 2, and therefore, there were few differences in a posture other than head height.

Next, we compared Figures 8(upper-right), 8(lower-left), and 8(lower-right). During this period, Player 2, the expert player, made a swiping, while Player 1 could not begin moving until Player 2 had finished swiping. We recorded multiple swiping, and found that Player 2 started moving quicker in each recording.

The difference between Player 1 and Player 2 was significant in the speed of the start of the movement.

2) Comparison when swiping the lower right hand side of opponent's territory: We recorded the movements of

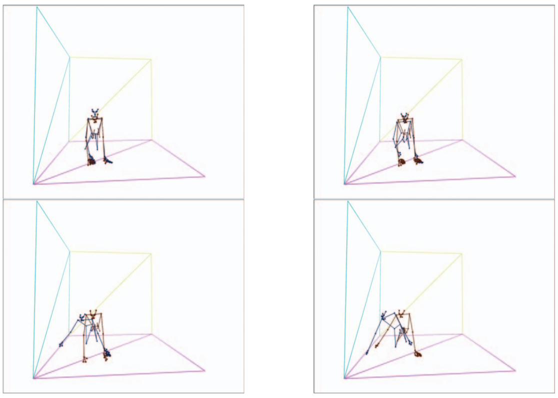

Fig. 9. Animation (2). (Upper-left) Before movement. (Upper-right) Start of movement of an expert. (Lower-left) Start of swiping of an expert. (Lower-right) End of swiping of an expert.

the players swiping the lower right-hand corner of the territory of the opponent. Figure 9 shows the sequence of the two players swiping the card.

Figure 9(upper-left) shows the scene before the two players start moving. At this point, similar to the results in the previous section, there was little difference except for a slight difference in the height of the heads of players.

Next, we compare Figures 9(upper-right), 9(lower-left), and 9(lower-right). We found that Player 1 has just started moving, while Player 2, who is an expert player, is presumed to be in the process of swiping the card at the time of Figure 9(lower-left). Furthermore, in Figures 9(lower-left) and 9(lower-right), we found that Player 1 is only swinging her hand, whereas Player 2 is moving her body towards the right-hand side and shifting her gravity in that direction. Based on this, we found a difference not only in the speed of the start of the movement, but also in the presence or absence of a gravity shift.

3) Summary of Analysis: The results of our study showed no significant difference in the posture of player movement between beginner (Player 1) and expert (Player 2) players. However, Player 2 demonstrated a faster start of movement compared to Player 1. Furthermore, while Player 1 used only their arms to swipe the lower left-hand corner of the opponent, Player 2 used their entire body to shift her gravity. Player 2 could accelerate the motion by the gravity shift. This technique may be particularly useful when swiping at distant places, such as the lower right-hand side corner of the territory of the opponent.

We could visualize and analyze the differences between beginner and expert players by using the animation of the proposed system with a 3D bone display.

V. Conclusion

This paper presented the results of a comparison of movements between several players using the visualization system we developed for the swipe movements of competitive Karuta. Specifically, we could find differences between the movements of beginners and experts, in addition to comparisons between experts with different levels of experience.

Our developed system first measures the swipe motions of the player by applying posture estimation, and then visualizes the results using a system equipped with a bone display and line charts. This visualization system allows for detailed comparisons of the swiping motions of multiple players, and for the discovery of differences between different parts of the body. Interactive visualization with two functions, a three-dimensional animation display and a line chart of positional information of each part, is a suitable method for visualizing the movements of competitive Karuta. This visualization method can be used to build a system that helps beginners improve their skills by comparing them with advanced players.

As a future prospect, we would like to develop a system that can measure the swiping movements of more players and compare their swiping movements with those of others. If we can extract the characteristics of swiping motions of each player and calculate the players who have a high similarity to themselves, we expect that this system will be useful for improving the performance of players. In addition, it is easy to collect data from other players in different environments since the system developed in this study runs on a usual web browser. We are still searching for more methods to compare and evaluate the swiping actions of multiple players as future work.

While we have compared the lower right-hand side swipe action of players, we will also search for other differences among players by measuring other patterns. Additionally, we would like to apply a more accurate method to measure the timing of swiping the card.

References

[1] S. Takeda, Y. Hasegawa, Y. Hirai, T. Kosugi, T. Tsukui, S. Yamamoto, Research on brain information processing during competition by players of Hyakunin Isshu karuta, Kinki University Faculty of Biotechnology, 33-43 (2009, in Japanese).

[2] H. Yamada, K. Murao, T. Terada, M. Tsukamoto, A Method for Determining the Moment of Touching a Card Using Wrist-worn Sensor in Competitive Karuta, Journal of Information Processing (2018, in Japanese).

[3] M. Kawanishi, S. Tsuchida, T. Itoh, Visualization of the repetitive practice of dance motion: Case study with multiple genres of dance, 27th International Conference on Information Visualisation (IV2023) (in review).

[4] P. Melios, Creative Dance Learning Platform Using Microsoft Azure Kinect, Univ. Cyprus, Dept. Computer Science (2021).

[5] L. Bridgeman, M. Volino, J.-Y. Guillemaut, A. Hilton, Multi-Person 3D Pose Estimation and Tracking in Sports, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (2019).

[6] S. A. Ghasemzadeh, G. Van Zandycke, M. Istasse, N. Sayez, A. Moshtaghpour, C. De Vleeschouwer, D. Christophe, a Unified Framework for Ball Detection, Player Instance Segmentation and Pose Estimation in Team Sports Scenes, The 32nd British Machine Vision Conference, Creative Commons Attribution Non Commercial Share Alike 4.0 International (2021).