Self-Training System for Tennis Shots with Motion Feature Assessment and Visualization

Masaki Oshita Takumi Inao
Kyushu Institute of Technology
Iizuka, Fukuoka, Japan
Email: oshita@ces.kyutech.ac.jp

Tomohiko Mukai
Tokyo Metropolitan University
Hino, Tokyo, Japan
Email: tmki@acm.org

Shigeru Kuriyama
Toyohashi University of Technology
Toyohashi, Aichi, Japan
Email: sk@tut.jp

***Abstract*—This paper describes a prototype self-training system for tennis forehand shots that allows trainees to practice their motion forms by themselves. Our system uses a motion capture device to record a trainee's motion, and visualizes the differences between the features of the trainee's motion and the correct motion as performed by an expert. This system enables trainees to understand the errors in their motion and how to reduce or eliminate them. In this study, we classified the motion features and corresponding visualization methods using one-dimensional spatial, rotational, and temporal features based on the key sporting poses. We also developed a statistical model for the motion features, allowing the system to assess and prioritize all features of a trainee's motion. This research focuses on the motion of a tennis forehand shot and evaluates our prototype through several user experiments.**

***Keywords*-training system, sports form, motion feature, visualization, motion capture**

I. Introduction

The training of motion forms is important in activities such as sports, dance, and fighting, where the trainees must learn how to perform basic motion forms correctly. However, it is not easy for trainees to evaluate the quality of their motion and identify which part of their motion should be improved. Normally, coaching by a skilled instructor is required for the training of motion forms, but it is difficult and expensive to employ a skilled instructor all the time. Moreover, it is not always easy for a skilled instructor to accurately identify the problems in the trainee's motion and describe how to correct them.

In this research, we have developed a prototype of a self-training system for tennis forehand shots that allows trainees to practice their motion forms by themselves (Fig. 1). Our system acquires a trainee's motion using a motion capture device. It then visualizes the differences between the features of the trainee's motion and the correct motions, as performed by experts, so that the trainee can understand the problems with their motion and determine how to make improvements. The trainee can repeat the motion and review the visualized problems until they are able to perform the motion form correctly.

There are many potential motion features and corresponding visualization techniques. The current system uses 1D spatial, rotational, and temporal features based on the key sporting poses, and then employs an appropriate visualization technique for each type of motion feature.

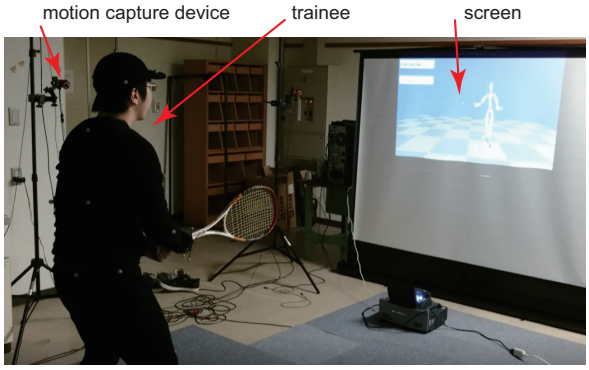


Figure 1. Environment of our training system.

A statistical model of the motion features has been developed to enable the system to assess and prioritize all features of a trainee's motion. In the visualization stage, it is not practical to show all of the motion features at the same time. Thus, it is necessary to choose which motion features to present based on the degree to which each motion feature deviates from the correct motion form. The error tolerance varies depending on the motion form and feature. A Gaussian model is used to represent the normal range of each motion feature based on a number of correct motions performed by experts.

Although our approach and system are applicable to all types of motion forms, this research focuses on the motion form of a tennis forehand shot. Our prototype was evaluated through several user experiments, and the experimental results and ideas for future research are discussed in this paper.

The remainder of this paper is organized as follows. Section II reviews some related studies before Section III describes our prototype self-training system for sports forms. Sections IV, V, and VI provide details of the definition, visualization, and assessment of motion features, respectively.

Section 7 presents our experimental evaluation, and Section 8 concludes this paper.

II. Related work

There are several training systems for motion forms that use motion capture systems [1], [2], [3]. However, these previous systems simply display the trainee's motion and the target motion performed by a skilled performer. Calculating the differences between the specific features of two motions and their visualizations has not been explored. Tang et al. [4] developed a physiotherapy system that visualizes motion features and provides interactive guidance on the camera image seen by the user. Although they used similar motion features and visualization techniques as our system, the visualized information is limited and must be manually specified. Their system was not designed for determining problems in a user's motion and correcting them.

Various motion features have been introduced for motion data processing. For example, kinematic-based features [5], geometrical features [6], and physics-based features [7] have been employed for motion or pose deformation. However, these features were not used to evaluate and identify the differences in the motions.

Motion recognition and assessment are important research topics. Motion features can be extracted from a given motion and combined with a pattern recognition technique to recognize the motion type or subject. Recently, Khan et al. [8] developed a system for assessing the skill level of motion based on discretized acceleration data and pattern recognition. Kyan et al. [3] employed a spherical self-organizing map to recognize and score ballet dance motions, whereas Laraba and Tilmanne [9] introduced hidden Markov models to evaluate dance moves. Tao et al. [10] also adopted hidden Markov models of motion dynamics to evaluate the safety of motion. However, the motion features and models in these previous studies are used for machine learning, rather than visualizing the difference between an input motion and the target motion.

Some researchers have developed methods for constructing motion models from a set of example motions [11], [12]. These models represent a standard motion with statistical variations and are designed for motion synthesis; thus, they are not applicable for motion assessment and visualization.

III. Self-Training System for Sports Forms

In this section, we describe the overview and design of our self-training system.

A. System Environment

Our training system is designed in such a way that trainees can practice their motion forms by themselves. The system consists of a motion capture system and a large screen, as shown in Fig. 1. The trainee's motion is captured by the motion capture system on the fly. The captured motion and visualized motion features are displayed on the large screen.

Although any kind of real-time motion capture system can be integrated with our system, we used the Optitrack [13] optical motion capture system with 12 cameras. We also used a ring-type remote controller with three buttons. This allows trainees to operate the system while practicing their motion forms without returning to the computer to use other input devices. One button is assigned to mode transition and two buttons are assigned to "next" and "previous" actions, whose specific roles depend on the current mode.

B. Training Process

The training system has five modes, as shown in Fig. 2. The trainee can move back and forth between these modes.

1) Target motion display mode: A target motion performed by an expert is first displayed to the trainee as an instruction. In this mode, the trainee can observe the target motion and control the viewing angle using the remote controller.

2) Motion capture mode: The trainee performs the motion form, which is acquired by the motion capture device. The trainee can start and stop the motion capture process using the remote controller.

3) Input motion display mode: The captured motion is displayed in the same way as the target motion display mode so that the trainee can confirm the motion has been captured.

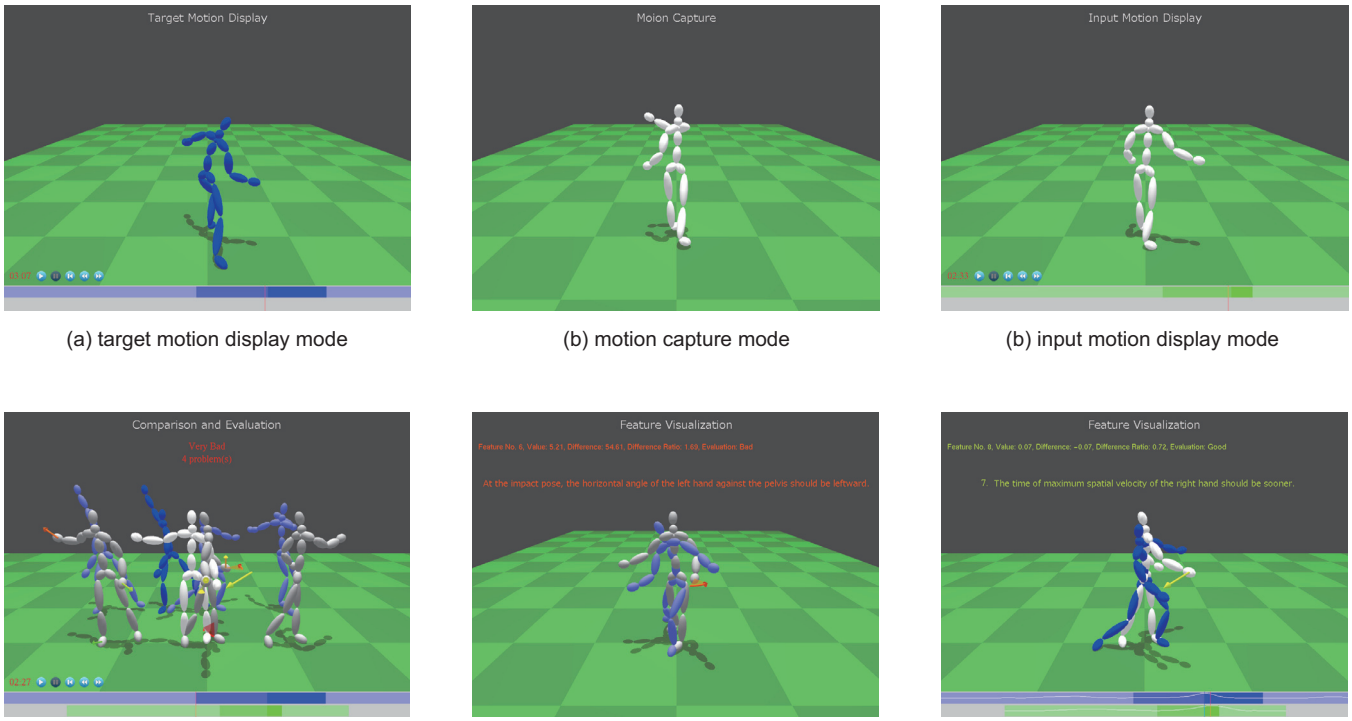
4) Comparison and evaluation mode: The input motion is displayed alongside automatically detected key poses. The target motion is also displayed next to the captured motion for comparison. The input motion is evaluated textually (see Section VI), and the trainee can again control the viewing angle using the remote controller.

5) Visualization mode: The differences between the features of the trainee's motion and the target motion model are visualized. To avoid confusing the trainee by showing too many motion features, a single motion feature is visualized along with the corresponding key poses. A visualization method for displaying the motion features is introduced in Section V. This method selects motion features in order of importance (see Section VI). The trainee can switch to the next or previous motion feature using the remote controller.

Normally, the trainee would proceed through these five modes in the above order before returning to the first mode to perform the next trial. The trainee can repeat this process until the motion form is performed correctly. Switching between the modes is controlled using the remote controller.

C. Motion Data Representation

The trainee's skeletal model and motion are captured through the motion capture system. Each captured motion is represented as a series of poses (60 poses per second), and each pose is represented by the position and orientation



(a) target motion display mode

(b) motion capture mode

(b) input motion display mode

(d) comparison and evaluation mode

(e) visualization mode

Figure 2. Screen shots of the five modes in our training system. The trainee's motion is shown by the white figure and the target motion is shown by the blue figure. The key poses of the trainee's and target motions are denoted by the gray and light blue figures, respectively. When necessary, the timeline is displayed along the bottom of the screen, with colored segments representing the motion between key poses. During visualization, the trainee can switch between motion features in order of importance.

of the pelvis and the rotation of all joints; each joint rotation is represented by a quaternion.

Experts prepare a set of target motions for the specific motion form in advance. At least one target motion is required to display the instruction model. To compute the statistical models for motion features, a number of target motions are required; in our experiment, we tested 12 target motions of tennis forehand shots as captured from two experts.

Because the skeletal models of the target motions differ from that of the trainee, their ideal motion feature may also be slightly different. To compensate for differences between the skeletal models (e.g., lengths of body parts), we could apply motion retargetting [14] or a simple scaling [2]. However, such motion deformations may cause unexpected variations in the motion features. Moreover, the ideal forms and their motion styles often vary depending on the trainee. Examples from a different physique cannot be used as target motions if the retargetting method has no proper style deformation. Therefore, we did not apply any motion retargetting or deformation, and assumed that the target motions were performed by experts whose skeletal models were sufficiently close to that of the trainee. The development of a sophisticated motion retargetting or style deformation method is left for future work.

It is possible to capture the motion of sporting props along with the full body motion, e.g., the motion of the racket in the case of tennis. However, we did not consider the prop motion, and focus on full body motion features for reasons of generality.

IV. Motion Feature Definition

In this section, we define and classify the motion features that are computed from the captured motion and used by our system.

A. Key Poses

A motion consists of a series of time-varying poses that constitute high-dimensional data. Moreover, the length of the pose varies depending on the motion. To compute motion features, we consider the key poses of the motion form. Normally, a motion form contains a number of key poses. Short motion forms such as tennis shots have three primary key poses, as shown in Fig. 3. These motions have preparation and expiration phases. In the initial key pose, the person starts to move. The effective key pose occurs when the person is moving at maximum speed. Finally, in the terminal key pose, the person stops. Because complex

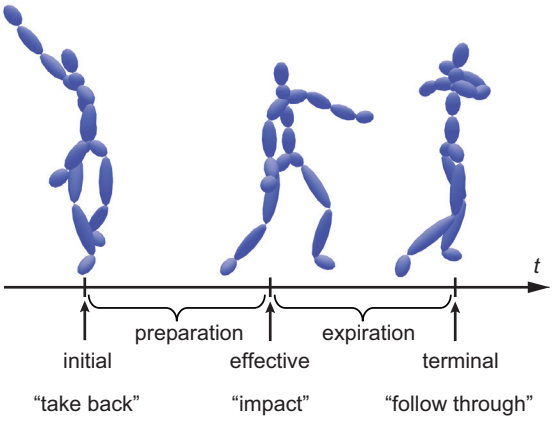


Figure 3. Example of key poses during the motion of a tennis shot. Three standard key poses (initial, effective, and terminal) are defined.

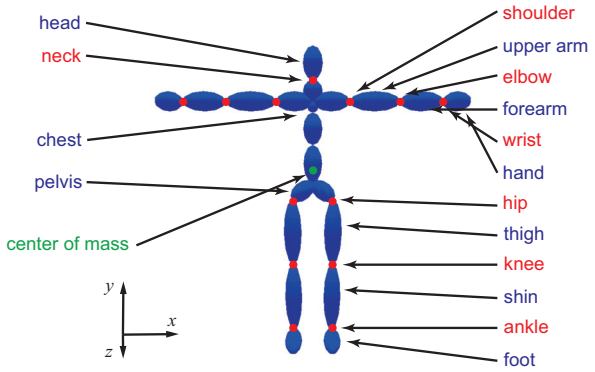


Figure 4. Standard skeleton model. The body parts include 15 body segments and 13 joints. Blue labels represent body segments, red labels represent joints. The center of mass (green) is treated in a similar way to the body parts.

motion forms are considered to have several combinations of these key poses, these three types of key poses are widely applicable.

For tennis shot motion forms, the key poses that are commonly used in coaching are called the “take back,” “impact,” and “follow through” poses, respectively. In the impact pose, the racket hits the ball.

B. Motion Feature Classification

The motion features computed from the key poses can be categorized as follows.

1) Features computed from a single key pose.
2) Features computed from the difference between two key poses.
3) Features computed from the whole motion or a segment of motion between key poses.

In addition, the motion can be categorized based on its spatial, rotational, and temporal features. The spatial and rotational features are computed based on the position and orientation of a body part at a key pose or their differences between two key poses. The temporal features are computed from the trajectory of positions or rotations of a body part over the whole motion or a segment of motion.

We describe the motion features based on the above classification in the following subsections.

C. Spatial and rotational motion features at a key pose

In this research, we use kinematics-based low-level motion features. Based on the skeletal model described in Section III-C, the position and orientation of body parts are computed from a pose using forward kinematics.

Body parts consist of body segments and joints. We define a standard skeletal structure and body parts as shown in Fig. 4. Although an actual skeletal structure depends on the motion capture system and its marker/sensor setup, which may have additional body parts, we use these standard body parts to specify the motion features for our system. Each joint has a position and a relative rotation between two connecting body segments. Each body segment has a center position and a global orientation. In addition, the center of mass is computed from the pose and the skeleton model using pre-specified segment weights. The center of mass is treated in the same way as the joints and is assigned a position. Some motion features are computed for a single body part (body segment or joint), whereas other motion features are computed for two (subject and reference) body parts.

According to the feature types and body structure, a motion feature can be defined by a combination of its key pose, feature type, subject body part, and reference body part (if applicable).

The positions and orientations are 3D features. However, through our preliminary experiments, we found that trainees struggled to grasp 3D motion, orientation, and rotation when visualized on a 2D screen. Therefore, we mainly use 1D motion features that are computed by decomposing 3D motion features. For example, the vertical, horizontal, right-left, and front-back distance between two body parts can be computed by projecting the difference vector between two body parts onto the y-axis, xz-plane, x-axis, and z-axis, respectively. We define a motion capture coordinate system in which the trainee faces the positive direction of the z-axis, as shown in Fig. 4. For many motion forms, the facing direction is important when performing the motion form. For example, in the case of tennis shot motion, the direction toward the opponent is given by the direction in which the trainee is facing.

D. Spatial and Rotational Motion Features between Two Key Poses

Motion features from two key poses can be computed by taking the difference between the features in the two poses.

In addition, the temporal difference (duration) between two key poses provides another motion feature.

These motion features are 3D vectors or rotations, or 1D distances or angles. They can be treated in the same way as the motion features extracted from a single key pose.

E. Temporal Motion Features over a Motion Segment

By analyzing the trajectory of a spatial or rotational feature of a body part, temporal features such as the point of maximum or minimum velocity can be computed. When performing a motion form, the synchronization of movements among multiple body parts is important. For example, in the motion of a tennis shot, the velocity of the right hand should be maximized at the point of impact, and the pelvis should attain its maximum velocity before the arm. Temporal features can be used to validate the trainee's motions in terms of such criteria.

A temporal feature is a 1D temporal difference. It is defined by the temporal difference between the maximum or minimum value of a body part and a specified key pose time, or the temporal difference between the maximum or minimum values of two body parts.

F. Calculations of Motion Feature

A set of motion features are computed from a given motion. First, three key poses $\mathbf{p}_i, \mathbf{p}_e, \mathbf{p}_t$ (see Section IV-A) are detected from the captured motion. Although there are many methods for extracting key poses from motion sequences [15], they are intended to extract a number of visually important poses and may not extract the three key poses we desire. Therefore, we introduce an algorithm for detecting the three key poses in the motion of a tennis shot.

Our algorithm takes the captured motion from a single tennis shot, which starts from the initial pose and ends at the terminal pose. First, the effective pose $\mathbf{p}_e$ is detected based on the end-effector position. In tennis shot forms, at the point of the effective pose, the hand holding the racket should be just beside the pelvis. $\mathbf{p}_e$ is determined from the right hand and pelvis positions of all frame poses. Second, the initial pose $\mathbf{p}_i$ and terminal pose $\mathbf{p}_t$ are determined as the two poses that are farthest from $\mathbf{p}_e$. The two poses that satisfy the following condition are used as $\mathbf{p}_i$ and $\mathbf{p}_t$:

$$\mathbf{p}_i, \mathbf{p}_t = \text{argmax}_{\mathbf{p}_i, \mathbf{p}_t} D(\mathbf{p}_i, \mathbf{p}_e) + D(\mathbf{p}_t, \mathbf{p}_e) + D(\mathbf{p}_t, \mathbf{p}_i) \quad (1)$$

where $D(\mathbf{p}_0, \mathbf{p}_1)$ computes the distance between two poses based on the average distance of all body parts. This is computed as follows:

$$D(\mathbf{p}_0, \mathbf{p}_1) = \frac{1}{n} \sum_{i \in S} |\mathbf{p}_0^i - \mathbf{p}_1^i| \quad (2)$$

where S is the set of all n standard body parts in Fig. 4 and $\mathbf{p}_0^i, \mathbf{p}_1^i$ are their positions in the two poses.

Once the key poses have been determined, the motion features are computed. The global position and orientation of body parts are computed using forward kinematics. The 1D distances and angles are computed by projecting or decomposing the vectors and rotations, and the spatial and rotational velocities are given by the poses in successive frames.

As explained above, there are many possible motion features, and their importance depends on the type of motion form. Currently, we assume that the set of motion features to be used are manually specified in advance by an expert who is familiar with the target motion form. The automatic selection of important motion features based on the analysis of example motions in a specific motion form is a task for future work.

V. Motion Feature Visualization

The differences between the trainee's motion features and those of the correct motion model are computed and visualized to indicate how each motion feature should be corrected.

As explained in Section 4.2, we mainly use 1D spatial, rotational, and temporal differences. We use a straight arrow or curved arrow to represent each motion feature. The motion features are visualized along with the corresponding pose, and a description, specific values, and assessment of each feature are displayed as text, as shown in Fig. 2(e). The camera direction is adjusted automatically. Normally, the camera faces the human figure from the front. However, when the axis of the primitive faces toward the front or back, the camera faces the human figure from the right. The visualization indicates how the position or orientation of the specified body part should be corrected.

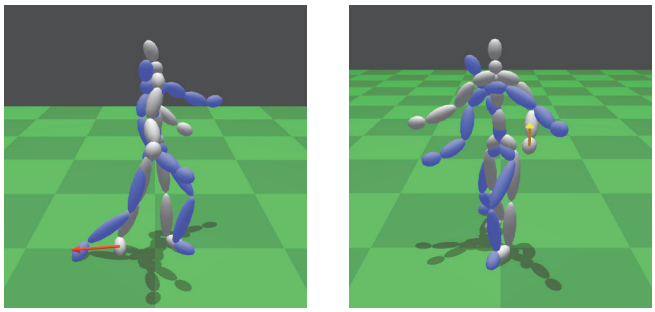
Color is used to represent additional 1D information. In particular, we use it to indicate the deviation of the motion feature from the correct motion model through a tone mapping. As the deviation increases, the color changes from green to red. Computing the deviation of motion feature is explained in Section VI.

A. Visualization of Spatial Difference

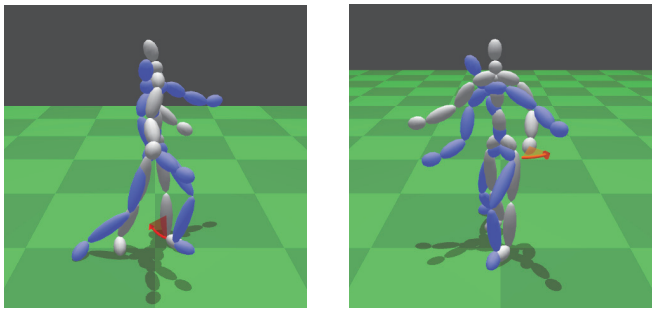
Spatial differences or distances are visualized using arrows, as shown in Fig. 5(a). Each arrow has a root position, direction, length, and color. The position of the subject body part at the key pose is used as the root position. If the motion feature is a 3D position, the direction and length are determined by the position of the reference body parts. If the motion feature is a 1D distance, the direction and length are determined according to the axis of the feature and the feature value.

B. Visualization of Rotational Difference

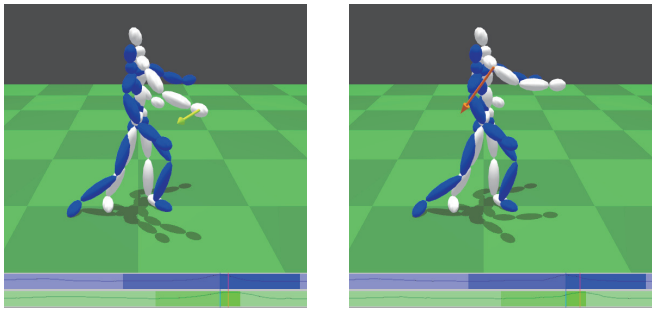
Rotational or angular differences are visualized using curved arrows, as shown in Fig. 5(b). Each curved arrow has a root position, axis direction, arrow direction, angle, and color. In the same way as for the spatial difference, the



(a) using an arrow for visualizing spatial difference



(b) using an arched arrow for visualizing rotational difference



(c) using an arrow for visualizing temporal difference

Figure 5. Example of the visualization of spatial, rotational, and temporal features using an arrow or curved arrow. (a) Visualizations of the distance between the right foot and left foot (right) and the height of the left hand (right). (b) Visualizations of the vertical angle of the left leg against the left thigh (left) and the horizontal angle of the left hand against the pelvis. (c) Visualizations of the maximum spatial velocity of the right hand (left) and the maximum rotational velocity of the right shoulder (right). The trajectories of the velocity are plotted on the timeline. The color indicates the deviation from the correct motion model.

root position is determined by the subject body part position. The axis and arrow directions are also determined based on the motion feature, and the angle is determined according to the feature value.

C. Visualization of Temporal Difference

Temporal features are different from the spatial and rotational features, as they do not exist in the spatial domain. To visualize them, we project the temporal difference into the spatial domain and use arrows to visualize differences in the timings of the input and target motions. We draw the pose at each timing and place an arrow going from the body position at the timing of the input motion toward the timing of the target motion model.

In addition, when visualizing temporal features, our system presents the timeline and corresponding trajectories and indicates the timings of the two motions at the bottom of the screen, as shown in Fig. 5(c).

VI. Motion Feature Assessment Based on Statistical Models

In this section, we describe our statistical model for motion features that allows the system to assess and prioritize the motion features of a trainee's motion. The statistical model is constructed from a number of correct motions performed by experts for each motion feature.

A. Statistical Model of Motion Feature

We employ a Gaussian model to represent the range of each motion feature. To validate this approach, we computed motion features from 12 tennis forehand shot motions performed by two experts and applied the Shapiro–Wilk test [16] to determine whether these features exhibit normality and fit a Gaussian model. The Shapiro–Wilk test is a common method for validating the normality of a dataset. We applied this test on 28 spatial and rotational motion features from the key poses listed in Table I. For all tests, the p values were between 0.920 and 0. 01102, and thus greater than 0.05. This implies that these motion features follow a normal distribution and can be represented by Gaussian models. Although we only tested the motion form and features for a tennis forehand shot, we believe that this approach is applicable to other sports motion forms and motion features in general.

A Gaussian model is represented by its average μ and standard deviation σ. In our system, these values are computed for each motion feature from a given set of target motions performed by experts.

B. Motion Feature Assessment

Given a trainee's motion and its features, each motion feature value v_i is evaluated and assessed as a_i, that is, its deviation from the Gaussian model of correct motions, which is computed by

$$a_i = \frac{|v_i - \mu_i|}{\sigma_i} \tag{3}$$

When a_i is less than 1.0, the motion feature is considered to be within the error margin of the correct motion model. The higher the value of a_i above 1.0, the more the motion feature is considered to deviate from the correct motion model. The motion features are prioritized according to their a_i value in order to determine which motion features should be presented first.

VII. Experiments and Evaluation

We conducted user experiments to evaluate our prototype when applied to the tennis forehand shot form. Two subjects, one trainee and one expert (instructor), participated in the experiments. The trainee had no prior tennis experience. We asked the trainee to evaluate whether the motion features selected by our system match the problems identified by the expert. We also evaluated whether the trainee subject could learn to perform the tennis shot form by himself, without the help of the instructor, using our prototype. As explained in Section VI, we used the 28 motion features listed in Table I and their statistical models.

A. Evaluation of Motion Feature Assessment

The first experiment evaluated the motion feature assessment of our model. It is important to choose problematic motion features that must be corrected in a given motion, similar to those that an instructor would point out during coaching. We collected three example forehand shot motions from the trainee subject. The expert was then asked to observe the captured motion and identify problems with the trainee's motion. He was asked to choose six motion features from among the 28 listed in Table I. Because it is even difficult for an expert to specifically prioritize problematic motion features, we asked him to simply choose six motion features with no priority. Our system assessed all of the motion features based on their statistical models and prioritized them. We then examined whether the first six motion features given by our system matched those selected by the expert for the three example motions. As a result, the overall matching rate (the percentage of motion features that appear on both lists) across all three example motions was 61.0%. Overall, our system and the expert subject chose similar motion features.

We also asked the expert to review the motion features chosen by our system. He commented that the motion features given by our system, including those that the expert did not choose, were reasonable. Moreover, he admitted that the system identified some problems that he had overlooked. For example, for the first result, although the instructor did not choose motion feature no. 11, he said that this was indeed an important feature. These results show that our system can assess the motion features and accurately identify problems with the trainee's motion.

B. Evaluation on Self-Training

A second experiment was conducted to evaluate the effectiveness of our self-training system. To simplify the training process for the evaluation, we visualized only one motion feature from the trainee's motion, namely that for which the deviation from the correct motion model was the greatest. The trainee was asked to correct the visualized problem for the next trial. The trials were repeated until all features of the trainee's motion satisfied the correct motion models.

Table I
Motion features that are used in our experiment for tennis forehand shot form.

no	timing	feature type	subject body part	reference body part
1	take back	height	right ankle	
2	take back	height	center of mass	
3	take back	distance	right hand	left hand
4	take back	distance	right ankle	left ankle
5	take back	vertical angle	left shin	left thigh
6	take back	vertical angle	right forearm	right upper arm
7	impact	height	left hand	
8	impact	height	right ankle	
9	impact	height	center of mass	
10	impact	distance	right hand	pelvis
11	impact	distance	right hand	left hand
12	impact	distance	right ankle	left ankle
13	impact	distance	left hand	left thigh
14	impact	vertical distance	right hand	pelvis
15	impact	vertical distance	right forearm	right upper arm
16	impact	horizontal distance	right hand	pelvis
17	impact	horizontal distance	left hand	pelvis
18	impact	vertical angle	left thigh	pelvis
19	impact	vertical angle	left hand	left upper arm
20	impact	vertical angle	left shin	left thigh
21	impact	horizontal angle	right hand	pelvis
22	impact	horizontal angle	left hand	pelvis
23	impact	horizontal angle	left thigh	pelvis
24	impact	horizontal angle	right forearm	right upper arm
25	follow through	height	center of mass	
26	follow through	distance	right hand	heat
27	follow through	distance	right ankle	left ankle
28	follow through	horizontal angle	left upper arm	pelvis

It took 18 trials until all motion features were within the margin of error of the correct motion models. During the series of trials, the subject was able to fix the problem specified in the previous trial. However, some problems were identified several times over the course of the experiment. This means that the trainee repeated the same mistake during the series of trials. The subject commented that it was difficult to satisfy all motion features for the correct motion models. Actually, this is the hardest aspect of motion form training, and requires practice to overcome. The subject also commented that the visualization produced by our system was clear and helpful in improving the motion form. Therefore, our system facilitated the training process by visualizing problems with the trainee's motion.

C. Discussion

The experimental results show that, although our motion features and visualizations are relatively simple, our system is useful for the self-training of motion forms.

Our system uses kinematics-based motion features such as positions, distances, and angles at key poses. There are many other motion features that could be used. Physics-based features such as joint torques and ground contact forces could be computed by applying a suitable dynamic model, but these features are difficult to visualize and trainees may not be able to understand them. Examining the introduction of new motion features and corresponding visualization techniques will be considered in future work.

There are correlations between several motion features. When there is some problem in the trainee's motion, several related motion features are likely to deviate from the

correct motion model. Currently, our system first presents the motion feature that is farthest from the motion model from among all the corresponding motion features. However, this may not be the best way to indicate problems with the trainee's motion. When an instructor coaches a trainee, the instructor often tries several ways of explaining how to fix one problem, because it is unclear which will be best understood by the trainee. A possible extension to our system involves finding these correlations between motion features and visualizing the related motion features together rather than presenting them individually.

VIII. Conclusion

We have developed a prototype self-training system for practicing tennis forehand shots. We classified the motion features and corresponding visualization methods, and introduced a statistical model that allows the system to assess and prioritize the motion features. We applied our approach to a tennis forehand shot motion form and evaluated our prototype through some user experiments.

In future work, we intend to evaluate our system on other kinds of sport motions. As discussed in Section VII-C, future research will also seek to introduce new motion features and corresponding visualization techniques, identify correlations between motion features and visualize the related motion features together, and construct more accurate models of the correct motion.

Our system has several limitations. We assume that there is an ideal motion form for the trainee's physique and use a set of example motions provided by experts. Although we believe that this is a reasonable assumption, preparing example motions is sometimes difficult. As explained in Section III-C, a motion retargetting method should be developed to solve this issue. Moreover, in practice, there are various styles of correct motion. For example, the style of tennis shots varies depending on the trajectory of the incoming ball and how the player wants to hit it back. Developing a sophisticated motion model to represent such manifold motion forms is another topic for future work.

Acknowledgments

This work was supported in part by a Grant-in-Aid for Scientific Research (No. 15H02704) from the Japan Society for the Promotion of Science (JSPS).

References

[1] J. C. Chan, H. Leung, J. K. Tang, and T. Komura, "A virtual reality dance training system using motion capture technology," *IEEE Transactions on Learning Technologies*, vol. 4, no. 2, pp. 187–195, 2011.

[2] F. Anderson, T. Grossman, J. Matejka, and G. Fitzmaurice, "Youmove: Enhancing movement training with an augmented reality mirror," in *ACM Symposium on User Interface Software and Technology (UIST) 2013*, 2013, pp. 311–320.

[3] M. Kyan, G. Sun, H. Li, L. Zhong, P. Muneesawang, N. Dong, B. Elder, and L. Guan, "An approach to ballet dance training through ms kinect and visualization in a cave virtual reality environment," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 6, no. 2, p. Article No. 23, 2015.

[4] R. Tang, X.-D. Yang, A. Tang, S. Bateman, and J. Jorge, "Physio@home: Exploring visual guidance and feedback techniques for physiotherapy patients at home," in *SIGCHI Conference on Human Factors in Computing Systems (CHI) 2015*, 2015, pp. 4123–4132.

[5] M. Neff and Y. Kim, "Interactive editing of motion style using drives and correlations," in *Eurographics/ ACM SIGGRAPH Symposium on Computer Animation 2009*, 2009, pp. 103–112.

[6] M. Guay, M.-P. Cani, and R. Ronfard, "The line of action: an intuitive interface for expressive character posing," *ACM Transactions on Graphics (SIGGRAPH Asia 2013)*, vol. 32, no. 6, p. Article No. 205, 2013.

[7] K. W. Sok, K. Yamane, J. Lee, and J. Hodgins, "Editing dynamic human motions via momentum and force," in *Eurographics/ ACM SIGGRAPH Symposium on Computer Animation 2010*, 2010, pp. 11–20.

[8] A. Khan, S. Mellory, E. Berlin, R. Thompsony, R. McNaneyy, P. Oliviery, and T. Plötz, "Beyond activity recognition: Skill assessment from accelerometer data," in *ACM International Joint Conference on Pervasive and Ubiquitous Computing (UBICOMP) 2015)*, 2015, pp. 1155–1166.

[9] S. Laraba and J. Tilmanne, "Dance performance evaluation using hidden markov models," *Computer Animation and Virtual Worlds*, vol. 27, no. 3-4, pp. 321–329, 2016.

[10] L. Tao, A. Paiement, D. Damen, M. Mirmehdi, S. Hannuna, M. Camplani, T. Burghardt, and I. Craddock, "A comparative study of pose representation and dynamics modelling for online motion quality assessment," *Computer Vision and Image Understanding*, vol. 148, pp. 136–152, 2017.

[11] M. Lau, Z. Bar-Joseph, and J. Kuffner, "Modeling spatial and temporal variation in motion data," *ACM Transactions on Graphics (SIGGRAPH Asia 2009)*, vol. 28, no. 5, p. Article No. 171, 2009.

[12] J. Min, Y.-L. Chen, and J. Chai, "Interactive generation of human animation with deformable motion models," *ACM Transactions on Graphics*, vol. 29, no. 1, p. Article No. 9, 2010.

[13] Natural Point, "Optitrack motive," http://www.naturalpoint.com/, 2013.

[14] M. Gleicher, "Retargetting motion to new characters," in *SIGGRAPH '98 Proceedings*, 1998, pp. 33–42.

[15] J. Assa, Y. Caspi, and D. Cohen-Or, "Action synopsis: pose selection and illustration," *ACM Transactions on Graphics (SIGGRAPH 2005)*, vol. 24, no. 3, pp. 667–676, 2005.

[16] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.