Augmented Practice Mirror: A Self-learning Support System of Physical Motion with Real-Time Comparison to Teacher's Model

Itaru Kuramoto[1], Yoshikazu Inagaki[2], Yu Shibuya[1], and Yoshihiro Tsujino[1]

[1] Kyoto Institute of Technology,
Matsugasaki, Sakyo-ku, Kyoto 606-8585 Japan
[2] Kyoto University of Education,
Fujinomori-cho, Fukakusa, Fushimi-ku, Kyoto 612-8522 Japan
ent@hit.is.kit.ac.jp

Abstract. An effective way to learn some physical motions such as dancing, playing sports, making traditional crafts, and so on, is to mimic teacher's motion. In this style of learning, it is important for the learner to recognize the difference between the teacher's motion and his/her one. We propose Augmented Practice Mirror (APM) learning support system. APM shows the mirror image of learner's motion overlapped teacher's one, and the difference between them. These three images are shown simultaneously on a large screen as virtual mirror in real time. As a result of the experimental evaluations, it was found that APM was better in recognizing the difference between the participant's motion and the teacher's one than two common methods, and that the hybrid interface of voice recognition and gesture was better than the single interface of voice recognition or gesture for operating APM.

Keywords: mirror interface, gesture, physical motion tracking, learning support, human model, voice recognition, augmented reality.

1 Introduction

An effective and simple way to learn some physical motions such as dancing, playing sports, making traditional crafts, and so on, is to mimic teacher's motion. It is not often that a learner can be taught by a teacher of the physical motion which he/she want to learn because there might be only a few teachers, or all of them might have already died in case of making some traditional craftworks. In such a case, a learner typically practices in front of a large mirror while watching the teacher's motion by movie in order to compare his/her motion with the teacher's one. When he/she notices that his/her motion is different from the teacher's one, he/she tries to decrease such different motion by changing his/her motion.

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2009, LNCS 5620, pp. 123–131, 2009.

In this style of learning, it is important for the learner to recognize the difference between the teacher's motion and his/her one. In order to support for recognizing it, three following points are needed;

1. To show teacher's motion model overlapped with learner's motion
2. To show both teacher's motion and learner's one simultaneously, in real-time
3. To make the difference between teacher's motion and learner's one easy to understand

Some self-learning support systems had already proposed, but they cannot deal with some of these points. Interactive Video Mirror[1] is a system for learning the physical motion of playing sports. It shows the mirror image of a learner's motion in real-time. However, it cannot overlap a teacher's motion on the image. Paravie[2] shows a learner's mirror image and a teacher's motion simultaneously in real-time, but they are displayed in juxtaposition, not overlapping.

We propose Augmented Practice Mirror (APM) self-learning support system which covers all of these points. APM shows the mirror image of learner's motion, overlapped teacher's motion, and the difference between them. These three images are shown simultaneously on a large screen as virtual mirror.

2 Implementation

2.1 Displaying Motions

Figure 1 shows the overview of APM. When using APM, a learner can see three following images in real-time.

- The learner's motion as mirror image with learner's wire human model.
- The wire human model tracing teacher's motion overlapped with the learner's motion. The model has already been recorded.
- The difference of each joint position (ex. knee, elbow, and so on) between the learner's motion and the teacher's one. Each difference is shown by dotted lines, and the color of them is determined by the degree of difference. In other words, the larger the difference is, the warmer the color is.

The learner can play teacher's motion from any point of time, and he/she can also control the speed of the motion. Table 1 shows the functions APM has. Using these functions, he/she can see the motion and/or the difference in detail. In addition, APM has the function to adjust the size of teacher's model. Generally, the body size of a teacher is often different from that of the learner. He/she can enlarge or shrink the size of teacher's motion model in order to fit it to his/her body. This adjustment is important when showing the difference between the teacher's motion and his/her one.

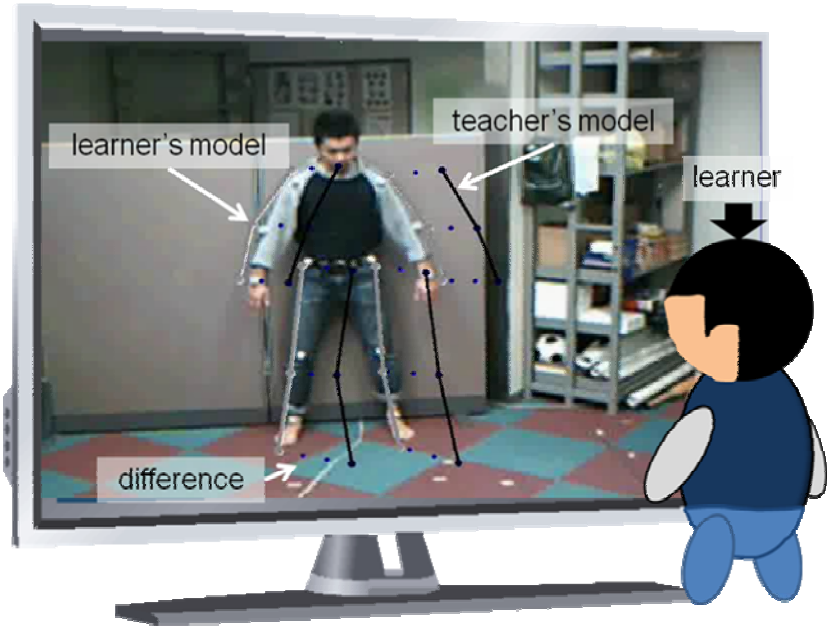

Fig. 1. Overview of APM

Table 1. The operations of APM

operation	*detail*
(re-)start	(re-)start teacher's motion with a normal speed
stop	pause teacher's motion
go to begging point	go back teacher's motion to the beginning point (scene #1)
rewind	start teacher's motion reversal with x2 speed
fast-forward	start teacher's motion with x2 speed
set the start point	set teacher's start point where a learner want to start
change play speed	change the speed of teacher's motion

2.2 Interface

When using APM, a learner must stand in front of the virtual mirror screen. In such a case, the learner cannot manipulate conventional PC devices such as mouse, keyboard, and so on, in order to operate APM. Some hand-held devices are also not available because he/she might want to learn the motion with precise hand movement. Therefore, it is considered that there are two interfaces for operating APM, that is, voice recognition and body gesture.

In the voice recognition interface, a voice command should be defined for each function mentioned Table 1. A learner wears a wireless hands-off microphone. In the interface, the learner can operate any functions with any poses. However, some functions such as "rewind" and "change play speed" may require many commands to operate, and they are time-consuming.

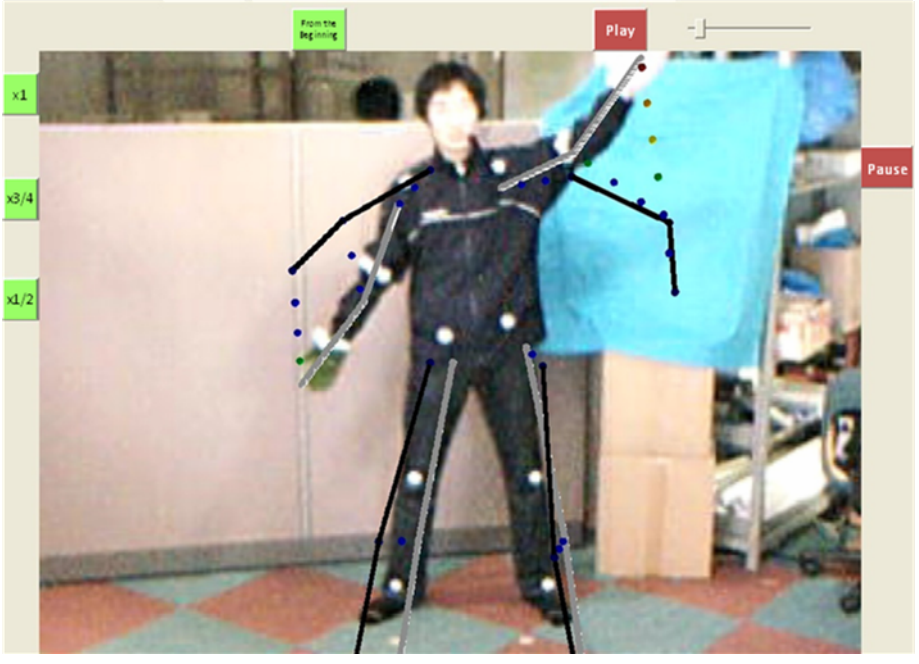

Fig. 2. Gesture interface

By contrast, the learner can operate simply by gesture like as HyperMirror[3]. Figure 2 shows the gesture interface of APM. In the interface, he/she touches or manipulates the control objects on the virtual mirror screen with his/her hands. The position of them is captured by the motion tracking system which APM uses for tracking the motion of his/her whole body. The interface has a slider for setting the start point of teacher's motion, so he/she can start from the desired scene of the motion directly and intuitively. However, the interface forces him/her to get out of the correct position when he/she invokes some commands.

In order to balance the trade-off between operation time and forcing some unnecessary body movements, we propose the hybrid interface of voice recognition and gesture. In functions shown on Table 1, it is considered that "(re-)start" and "stop" by gesture are particularly obstructive for learning. Therefore, in the hybrid interface, these two commands are executed by voice recognition, and the other commands are by gesture. However, the interface may be confusing because it combines two different interfaces so the learner cannot do the correct action immediately.

2.3 Hardware

Figure 3 shows the system overview of APM. APM is implemented on two PCs with Windows XP Professional. One PC tracks learner's motion with Radish the motion tracking system. Another PC captures learner's image by an USB camera, and composes the mirror image. DirectX technology is used to capture images from the camera and represent the wire human model in a virtual 3-D space. In addition, Microsoft Speech SDK 5.1 is used for voice recognition.

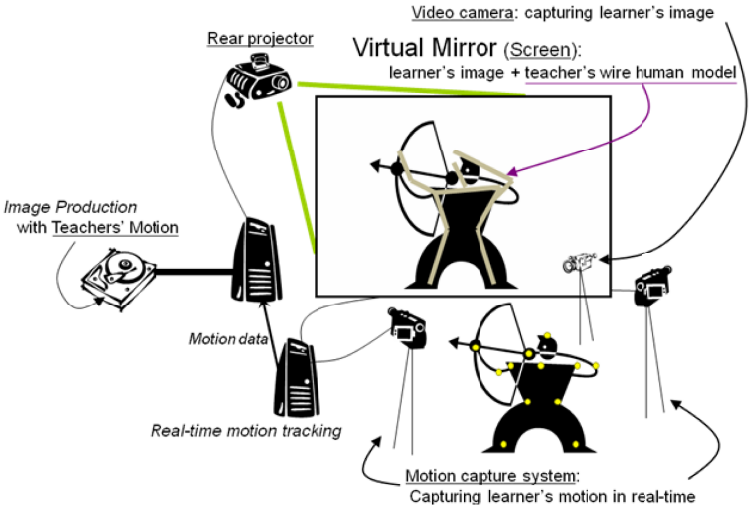

Fig. 3. System overview

3 Evaluation

3.1 Usability

In order to evaluate the usability and efficiency of APM, we held laboratory-based experiment. We compared three types of representation, which were:

- Md: The proposed method. APM showed the difference, the wire human model of teacher's motion, that of learner's motion, and learner's image. All of them were overlapping.
- Mo: APM showed the wire human model of teacher's motion, that of learner's motion, and learner's image. They were overlapping.
- Mp: The wire human model of teacher's motion and learner's image were shown in juxtaposition.

We asked 9 participants to learn three types of motion, which were throwing a dart, shot-putting, and robotic dance. In the experiment, one of the authors operated APM to exclude the influence of the usability of its interface. The learning time was 15 minutes each. After each learning, they were asked to answer the questionnaire, which consists of the questions described below;

- Q1. I could recognize the motion of the teacher easily. (-2: strongly disagree, 2: strongly agree)
- Q2. I could recognize the motion of myself easily.
- Q3. I could recognize the difference between the motions of the teacher and me.

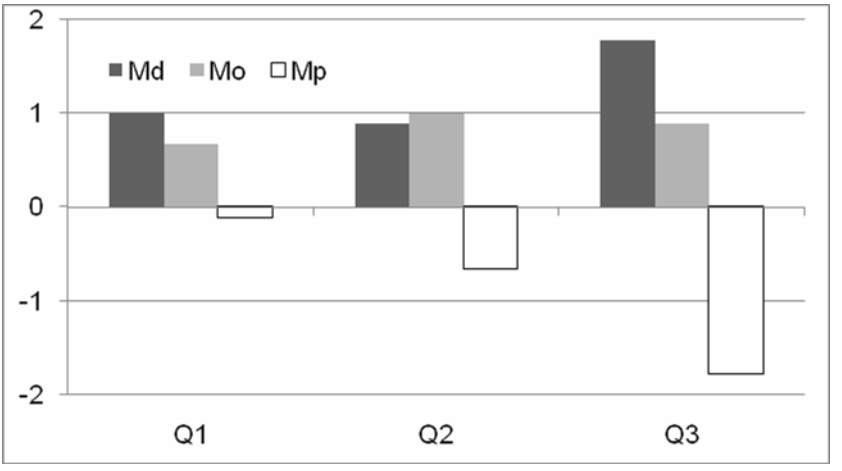

Fig. 4. The result of the usability experiment

Figure 4 shows the result of the questionnaire. As a result, the average score of Mp is significantly worse than those of Md and Mo in question 1. The result indicates that overlapping teacher's motion with that of participants is better to recognize the motions than juxtaposition. It suggests that the participants must concentrate both of motions simultaneously to recognize the difference, so the concentration in the case of parallel representation (Mp) is more difficult than the cases of overlapping representation (Mo and Md).

In addition, the score of Md is significantly better than those of Mo and Mp in question 3. It is found that the proposed method seems to be acceptable and useful.

3.2 Interface

In order to compare the performance of the interface of APM, we held an additional laboratory-based experiment. We compared three interfaces of APM, which are V: voice recognition, G: gesture and H: hybrid interface mentioned in section 2.2.

Table 2. The operations measured in the experiment

operation	*detail*
play	(re-)start teacher's motion
pause	stop teacher's motion when mimicking
rewind short	rewind a short period (or set the start point slightly before the current position in H and G)
rewind long	rewind a long period (or set the start point largely before the current position in H and G)
two levels slower	change the speed to x1/2 (in V, say ``slow'' two times)
one level faster	change the speed from 1/2 to x3/4 (in V, say ``fast'' once)

We asked 18 participants to learn a motion of robotic dance, which was the same one as the previous experiment in section 3.1. Each of the participants performed one task for each interface. We asked them to execute some operations which were displayed at the upper part of the virtual mirror screen (see Figure 5) during they were mimicking the motion of the dance. Table 2 shows the operations which we measured.

We measured the time interval from displaying an instruction to the execution of the operation correctly as the performance time of the operation. In addition, after each task we asked them to answer the questionnaire, of which questions are shown below;

- Q1. I could understand the interface easily. (-2: strongly disagree, 2: strongly agree)
- Q2. I felt a load when learning with the interface.
- Q3. I could control APM appropriately.

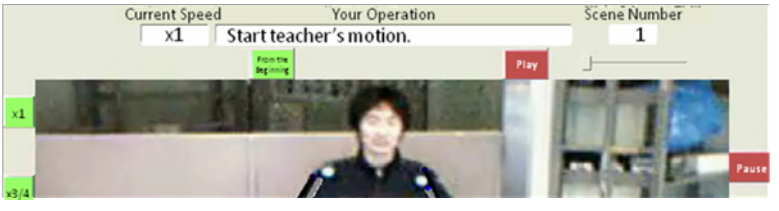

Fig. 5. Operation displaying field for experiment

Figure 6 shows the result of the performance time per operation. As a result, the G's performance time of "play" and "pause" is shorter than that of V and H. In addition, the V's time of "rewind long" and "two levels slower" is significantly longer than that of G and H. It is pointed out that the gesture interface is more effective than the voice recognition interface in the aspect of the performance time. It is the similar to the argument which Hämäläinen pointed out in the evaluation of Interactive Video Mirrors[1].

In the interview after the experiment, some participants said that they were confused in case of "play" and "stop" operations with H because they could not judge quickly which modes (voice or gesture) they should use. From the opinion, it is expected that the difference between H's operation time and G's becomes small as they are familiar to the hybrid interface.

Figure 7 shows the result of the questionnaire. In question 1 and 3, the average score of G and H is higher than V because the participants are required to remember all of the commands in the voice recognition interface. By contrast, the average score of V and H is higher than G in question 2. It means that operation by the gesture interface is an actual obstruction when they are learning.

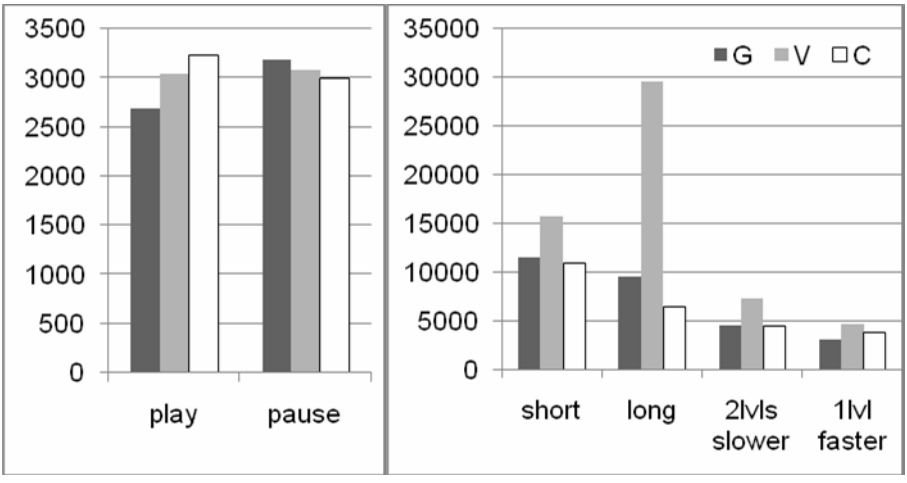

Fig. 6. The result of operation time (by millisecond)

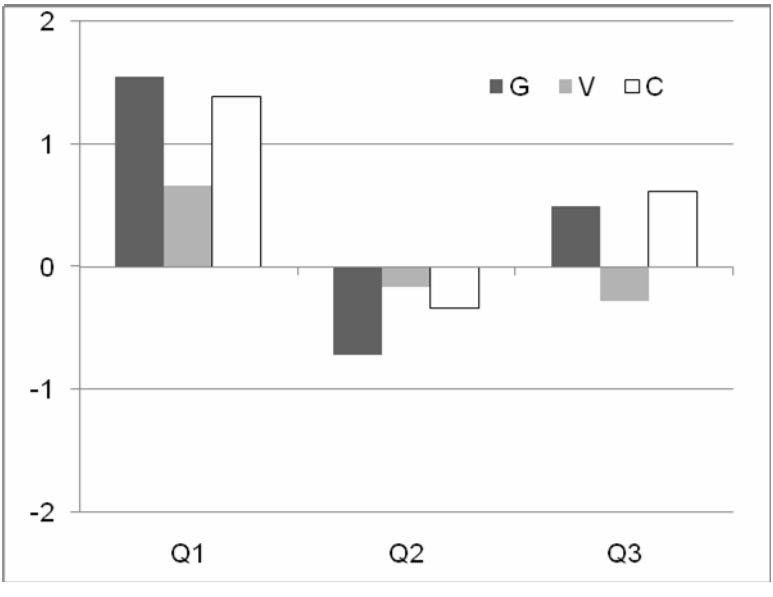

Fig. 7. The result of the questionnaire

In addition, the H's average scores of all questions are high, therefore it is considered that the hybrid interface is more effective than other two interfaces in the both aspect of performance time and subjective evaluation.

4 Conclusion

APM is a virtual mirror system for self-learning of physical motion. It has mirror-like screen which shows a learner's motion, a teacher's motion by wire human model, and differences between their motions simultaneously in real-time.

As a result of the experimental evaluations, it is found that APM is better in recognizing the difference between the participant's motion and the teacher's one. In addition, it is also found that the hybrid interface of voice recognition and gesture is better for operating APM than the voice recognition interface and the gesture interface.

Acknowledgement

This work was partially supported by KAKENHI (Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (B), 20300037, 2008).

References

1. Hämäläinen, P.: Interactive Video Mirrors for Sports training. In: The 3rd Nordic Conference on Human-Computer Interaction, pp. 199–202 (2004)
2. Usui, J., Hatayama, H., Sato, T., Furuoka, Y., Okude, N.: Paravie: dance entertainment system for everyone to express oneself with movement. In: The 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, p. 30 (2006)
3. Morikawa, O., Maesako, T.: HyperMirror: toward pleasant-to-use video mediated communication system. In: 1998 ACM conference on Computer supported cooperative work, pp. 149–158 (1998)