A System for Training Stuffed-Suit Posing Without a Suit

Ryo Nakayama[1(✉)], Tsutom Terada[1,2], and Masahiko Tsukamoto[1]

[1] Graduate School of Engineering, Kobe University, Kobe, Japan
nakayamaryou@stu.kobe-u.ac.jp, tsutomu@eedept.kobe-u.ac.jp, tuka@kobe-u.ac.jp

[2] PRESTO JST, Japan Science and Technology Agency, Kawaguchi, Japan

Abstract. People who perform while wearing stuffed suits are popular among people of all ages; however, the performers need to train themselves stuffed-suits on their posing before performing. Many performers are forced to train themselves to pose without wearing a stuffed suit because there are few environments where they can train with a stuffed suit, which makes pose training difficult for them. This paper describes a system we propose that enables performers without a stuffed suit to pose train themselves by observing images of the same type of stuffed suits that performers actually wear. Using our system enables users to train themselves with the same sensations they would feel when wearing stuffed suits, which enables them to perform the posing smoothly in a stuffed suit. We carried out a preliminary study to verify the difficulties performers face when wearing a stuffed suit and implemented a prototype of our proposed system. Evaluation results confirmed that using our system enabled performers to improve their posing skills compared with conventional training methods.

Keywords: Stuffed suit · Training · Motion capture system
Visually feed back

1 Introduction

Stuffed suits have been widely used in various theme parks and events because people wearing stuffed suits provide performances popular among people of all ages and make them smile and be happy. Performers wearing stuffed suits need to perform as the character they play because stuffed suits have a role to make the characters in the virtual world appear in the real world.

To perform in an expert manner, performers need to perform the posing like the character they portray. Stuffed suits alone cannot change their voices or facial expressions. If the stuffed suits do not correctly portray the character being represented, the people watching will feel uncomfortable. Therefore, before each performance performers need to thoroughly practice their posing for the stuffed suit they will wear.

K. Murao et al. (Eds.): MobiCASE 2018, LNICST 240, pp. 183–200, 2018.
https://doi.org/10.1007/978-3-319-90740-6_11

In training while wearing a stuffed suit, many performers are forced to train without wearing one because there are few environments where they can train while wearing one. In general training environments, they imagine the appearance of the suits and imitate the appearance they will convey by posing in front of a mirror. However, it is difficult for them to imagine how they will perform without wearing a stuffed suit because of the differences between their own bodies and those that people will see when they are wearing a stuffed suit.

To address this issue, we propose a system that enables performers who are not wearing a stuffed suit to be trained in stuffed-suit posing by visually presenting them with the same type of stuffed-suit images that they would use as performers. In our system, the user first needs to create a database that consists of several images of stuffed-suit posing and user's skeleton data in the stuffed suit. Then, the user uses the database and trains himself or herself in stuffed-suit posing in front of a Kinect device and a display. Our system visually presents images of stuffed-suit posing that matches those with the database on the display in real time, based on the user's skeleton data measured by Kinect. In addition, we implemented a function that enabled users to learn whether or not they could pose in the same way as they could when wearing a stuffed suit. More specifically, we implemented a function to feed back to users which body parts were out of the range of motion that stuffed suits allow when they posed in a manner that they could not have done without wearing a stuffed suit. We carried out a preliminary study to verify the difficulties in training stuffed-suit posing and implemented a prototype application of our proposed system. The evaluation results indicated that our system is effective for enabling users to pose without wearing a stuffed suit.

2 Related Work

There are several studies on improving stuffed-suit performances. Okazaki et al. proposed a system for performers in stuffed suits [1]. The system has two functions, i.e., posing support and vision extension. In the posing support system, users can pose like the character they wish to portray by using displayed images of stuffed suits they get in a head mounted display (HMD). In the vision extension system, they can see the camera images the HMD gets around the eye level of the user wearing a stuffed suit. The system enables performers to check the surroundings easily and react to the action of surrounding people quickly. Tei et al. proposed a multimodal interface that supports stuffed-suit performance by visual information using an HMD as well as auditory information using speaker and tactile information from a vibration motor provided to the user according to circumstances [2]. Slyper et al. proposed a system in which a person wearing a stuffed suit can talk with other people [3]. Users wearing stuffed suits can talk to others by operating a mouth input device with their tongue and selecting the sound of the character's voice. Although many studies have been reported that support stuffed-suit performances, none have been reported that support the training of users not wearing a stuffed suit, as our proposed system does.

Kinect [4] is widely used for measuring body motion in systems that support human motion acquisition. Since Kinect is relatively inexpensive and users do not need to wear any sensors on their body when using it, they use it in several motion acquisition systems such as those reported by Kyan et al. ([5], Dimitrios et al. [6], and Saha et al. [7]). Because of Kinect's high versatility, we used it in our study to measure the data obtained for users when they were training while wearing a stuffed suit. In such cases, however, Kinect cannot effectively measure data. Since we needed to measure the data for users wearing stuffed suits, we sought measuring methods other than those from Kinect.

Motion capture systems are frequently used for measuring body motion. They measure and record movements in time sequences in accordance with position information in spaces for human body parts. Van der Linden et al. proposed a system that supports the learning of dance skills by mimicking the movement of avatars created on the basis of familiar motion data acquired using an inertial motion capture system [8]. Tommi and Marc are trying to convey South African traditional dances to people of different cultures using an optical motion capture system [9]. Van der Linden et al. proposed a system that helps to improve violin playing skills by showing beginners how to hold the violin and how to improve their bowing technique by using an inertial motion capture system [10]. In this way, motion capture systems have been effectively used in extraction, storage and inheritance of human motions. Using an inertial motion capture system makes it possible to measure human motion data obtained inside stuffed suits.

In our study, we aimed to construct a highly versatile system by separately using Kinect, which is widely popular and easy to use, and an inertial motion capture system that can measure human motion data obtained inside stuffed suits.

3 Preliminary Study

In general, many performers have to train to perform without wearing a stuffed suit because they cannot always prepare the suit in advance. When training, they imagine their appearance in the suit and imitate the pose they wish to convey in front of a mirror. However, in many cases, when they actually wear the stuffed suit, they cannot pose like they did in the training. In this section, we describe an experiment we conducted to explore why they are unable to do so and how to solve the problem. In the experiment we investigated two types of stuffed suits with a large head (Fig. 2). Stuffed suits of this type are the most often used. Suits A and B comprise three parts: the head, torso and leg parts. Suit A is larger in size and heavier than suit B because it is for a male character while suit B is for a female one. We used these suits because we felt the differences between them might affect the obtained results (Fig. 1).

3.1 Experiment Environment

In this experiment, we measured the skeleton data by using a motion capture system when the subject posed with the same appearance in three states: non-

Fig. 1. Two types of stuffed suits.

wearing, wearing-A and wearing-B (Fig. 2). The experiment subject had had no previous experience in wearing a stuffed suit. His actions were adjusted so that they would be the same for all three states. The motion capture system we used to measure the skeleton data was 3-Space Sensors, which enables measurements to be taken through 17 inertial sensors attached to the subject with a band and enables skeleton data to be visualized by using dedicated software [11]. Using a camera enabled us to obtain the appearance image of each state at the same time as we obtained the measured skeleton data. After the measurements, we compared the obtained skeleton data.

Fig. 2. Identical pose for non-wearing, wearing-A and wearing-B states.

3.2 Results and Discussion

Figure 3 shows the skeleton data measurement results obtained with the dedicated software and the subject's appearance image in each state. Clear differences were found in the skeleton data between the wearing and non-wearing states. The difference from the right shoulder to the elbow was especially clear; when wearing a suit the subject was unable to raise the shoulder as well as he could when not wearing one. We consider that this is because the head part of the suit covers the shoulder and makes it harder for the subject to lift it. There were also clear differences in the skeleton data between the wearing-A and the wearing-B cases. That is, it was harder for the subject to raise the shoulders and

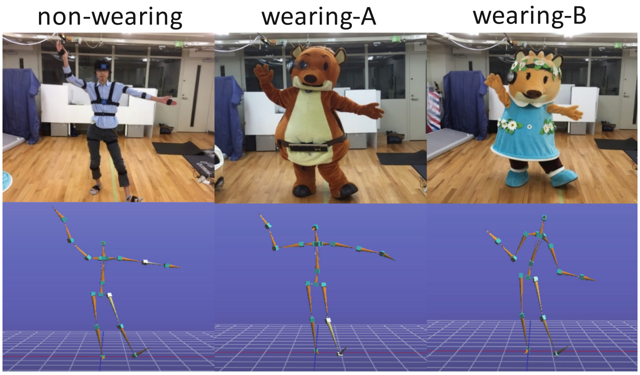

Fig. 3. Skeleton data measurement results for the non-wearing, wearing-A and wearing-B states

upper body in the latter case. We consider that this is because the smaller suit B restricts the range of movement more than for suit A. From this, we concluded that the type of suit will affect freedom of movement, even for suits having the same shape.

In this way, we concluded that performers cannot effectively make the motions of the stuffed suit's character because of the differences between the suit's appearance and the actions the performer makes when wearing it. Therefore, in actual performances they cannot pose in the same manner that they could when wearing a stuffed suit in training. Thus, in actual performances their poses do not come across effectively and this has a negative effect on the popularity and reliability of people wearing stuffed suits. We also concluded that the type of suit worn brings about differences in the performer's range of motion. That is, even performers who have worn suit A may not be able to perform in the same manner when wearing suit B because of the differences in the range of motion the suits allow them.

These results indicate that it is important for performers to understand the range of motion that suits will allow them during the time they are training. It is hard for performers who have never worn a stuffed suit to judge from the suit's appearance how it will restrict their freedom of movement. Even performers who have worn stuffed suits previously find it hard to imagine how a suit they have never worn will affect their freedom of movement. Beginning with the next section, we will describe the differences between the appearance of a stuffed suit and the actions the performer makes when wearing it. These differences make it difficult to train people on how to perform when wearing a stuffed suit when they are not actually wearing one.

4 Proposed System

In our study, we designed and here propose a system that enables performers not wearing a stuffed suit to be trained as if they were wearing one, by visually presenting them with images of performers wearing stuffed suits posing in the same way as they will. The system enables users not wearing a stuffed suit to feel the same sensations in training as they would when wearing one. The results we obtained in a preliminary study confirmed the need to provide two system requirements for training purposes.

Requirement 1: users can learn the range of motion stuffed suits provide them. Posing training by watching one's image in a mirror induces users to train themselves in posing that cannot be performed when wearing a stuffed suit because the users imagine their appearance in wearing a stuffed suit during training. In such training, the users perform incomplete posing when wearing the stuffed suit because they cannot pose in the same way as they could during training. Therefore, in the posing training without a stuffed suit, if users perform posing that they could not while wearing the stuffed suit, the system needs immediate presentation to the user. For example, when the users train themselves in stuffed-suit posing B (Fig. 2), if they raise their shoulder, the system visually tells them they cannot perform this pose when wearing stuffed suit B. Through a trial-and-error procedure this enables users to find poses they can perform with a stuffed suit and learn the range of motions the stuffed suit will allow them.

Requirement 2: users can in real time see how they perform when wearing a stuffed suit. In several training systems, such as those for sports and dances, there are many ways to feed users' training data, such as graphs and video, back to the user after training. However, in this method, we considered that it is hard for performers to memorize their body sensations in posing by associating them with their training data. Therefore, our proposed system visually presents stuffed-suit posing to users in real time, which enables them to memorize the posing by associating it with their own body sensations.

4.1 System Structure

The structure of our proposed system is shown in Fig. 4. The system consists of a Kinect device, a motion capture system, a PC and a display. The system consists of two phases: a database-creation phase and a posing-training phase. In the database-creation phase, an expert performer with experience in wearing stuffed suits will acquire the appearance image of the stuffed suits and his/her skeleton data and will save them in a PC database. In the posing-training phase, users use the database and train themselves on how to pose in front of the Kinect device and a display. In the posing training, our system visually presents stuffed-suit posing to the user in real time.

4.2 Database-Creation Phase

It is desirable for an expert performer to carry out this Database-Creation phase because abundant databases cannot be created from the performances of beginners wearing stuffed suits. In this phase, expert performers wear a stuffed suit

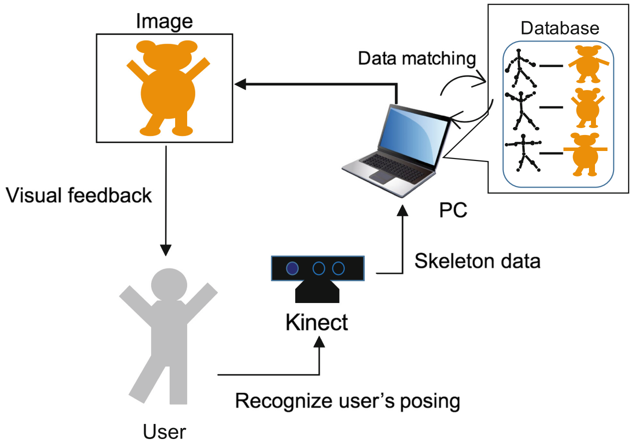

Fig. 4. System configuration.

and their skeleton data are measured by using the motion capture system. The use of a camera simultaneously enables them to see the images of themselves wearing the stuffed suit. Data measurement and appearance image acquisition are performed in synchronism with each other. After the measurements, the users create a database by associating the skeleton data with the appearance image as shown in Fig. 5, then saves them in the PC. The skeleton data is three-dimensional coordinates of 17 joints' position of the human body. The data is $17 * 3 = 51$-dimensional data. The procedure of this phase is shown below.

1. An expert performer wears a stuffed suit and inertial sensors of a motion capture system and performs various poses.

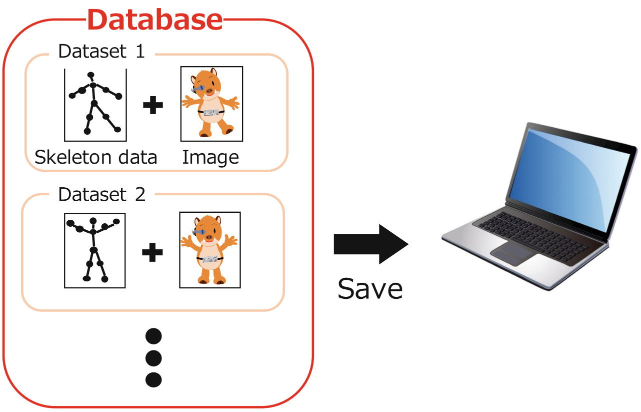

Fig. 5. Database-creation phase.

2. Skeleton data is measured by using a motion capture system with the sampling number set to 10 Hz. The data is acquired continuously.
3. At the same timing as the acquisition of skeleton data, appearance images of the stuffed-suit posing are recorded by using a camera.
4. The skeleton data and the appearance image of the stuffed suits acquired at the same timing were saved on the PC as a database.

4.3 Posing-Training Phase

In the posing-training phase, users use the database and train themselves on the posing in front of a Kinect device and a display. If another person conducts the database-creation phase, the user uses that database. In the training of stuffed-suit posing without a stuffed suit, the display visually presents stuffed-suit posing images that match those in the database. The procedure of this phase is shown below.

1. The user gets posing training without a stuffed suit in front of a Kinect device and a display.
2. User's skeleton data is measured by using Kinect.
3. The measured skeleton data is matched with skeleton data in the database.
4. Stuffed-suit images corresponding to skeleton data in the database presented to the user on the display.

Recognition Method. The flow of data processing was as follows. Figure 6 shows the data processing flow. First, we used the nearest neighbor method to search for the data in the database that was closest to that in the Kinect device. We set the learning data as the skeleton data acquired in the database-creation phase. The test data was set as the skeleton data acquired from Kinect in the

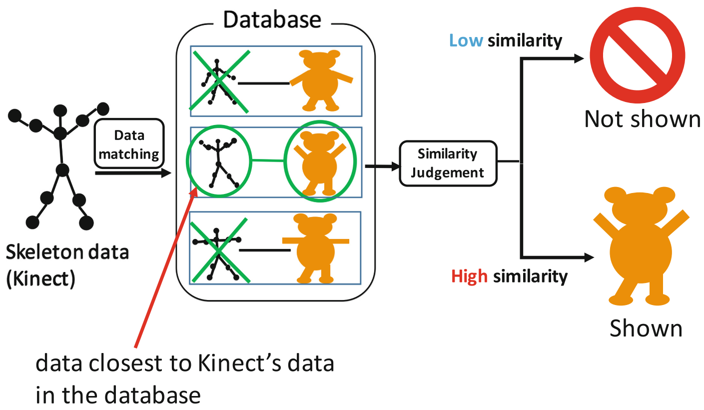

Fig. 6. Recognizing the processing flow of the posing-training phase.

posing-training phase. One label was assigned to one dataset with the skeleton data in the database and the appearance image of the stuffed suits. We allocated the labels so that they would match the number of skeleton data elements in the database. Therefore, the number of labels depended on that of the datasets in the database. For determining labels, we calculated the distance between the learning data and the test data at each joint and took the total of the distances as the total distance of data for one label. Skeleton data of the label with the smallest total calculated distance was selected as the skeleton data closest to that from Kinect. Next, we judged the similarity between the skeleton data from Kinect and the selected data. If there was even one joint whose distance was at least 10 cm longer than the skeleton data from Kinect in the selected data, we regarded it as having low similarity and so did not display the appearance image of the stuffed suit. If there were no joints whose distance was at least 10 cm longer than the skeleton data from Kinect, we regarded them as having high similarity and displayed the image of the stuffed suits. This function enables users to determine whether they could perform the posing they intended whether or not they were wearing the stuffed suit.

5 Implementation

On the basis of what we reported in the previous section, we implemented a prototype of our proposed system. In the database-creation phase, we used the Xsens MVN motion capture system [12] and logicool HD PRO WEBCAM C 920 R [13]. This motion capture system can measure a user's motion data by wearing a dedicated suit with 17 motion sensors on the body. The camera can acquire the appearance images of the stuffed suits at the same time it acquires the motion data by connecting to a PC. The sampling frequency of the motion sensor was set to 10 Hz and the PC used Lenovo's ThinkPad X1 Carbon (CPU: COREi 7-4600 U 2.10 GHz, 2.69 GHz, memory: 8 GB). As the software for the motion capture system and the camera, we used Microsoft's Visual C ++ 2013 and OpenCV [14]. In the posing-training phase, we also developed applications using Microsoft's Kinect and Visual C ♯ 2013. The sampling frequency of Kinect was set to 30 Hz.

5.1 Application

The application displays the appearance image of the stuffed suits that is the closest to a user's posing in real time. If the user performs posing that cannot be performed with the stuffed suit, the application immediately presents this information to the user. In the posing training, the user performs the posing in front of a display and the Kinect device as shown in Fig. 7. In this application, we implemented a skeleton correction function for users to use our system with the database from others. In the skeleton correction function, the user's skeleton data is linearly corrected and matched with the skeleton data of the database. By using this function, the user does not have to perform the database-creation

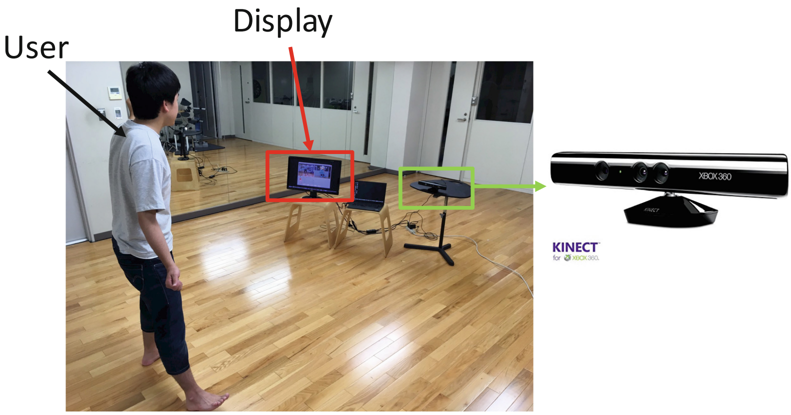

Fig. 7. Posing-training environment.

phase and is trained in posing by using a database prepared in advance from others. In addition, we implemented a posing correction feedback function that feeds back to the user the part of the body that is the cause of the posing that cannot be done. By using this function, the users can revise their posing to an appropriate posing. Figure 8 shows the application UI. The skeleton data from Kinect and that in the database are matched by using the nearest neighbor method and the appearance image of the stuffed suit that is the closest to the user's posing is visually presented. If the user performs posing that cannot be done with the stuffed suits or posing that is unmeasured during the database-creation phase, the appearance image of the stuffed suit is not displayed as shown in Fig. 9. Then, the parts of the body that are out of the motion range of the stuffed suit are displayed in red circles and fed back to the user. In addition, the

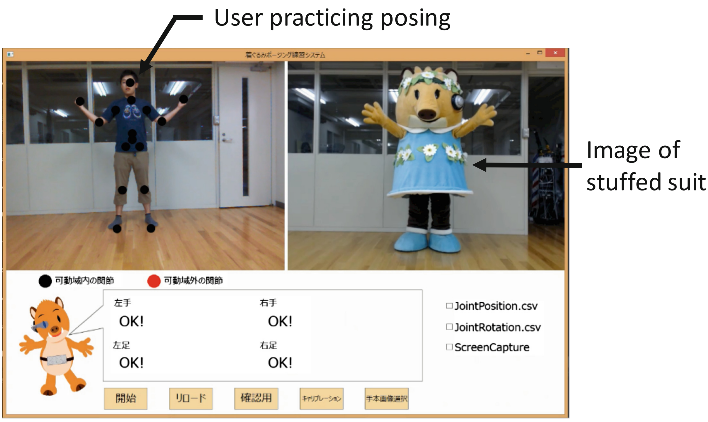

Fig. 8. Application UI.

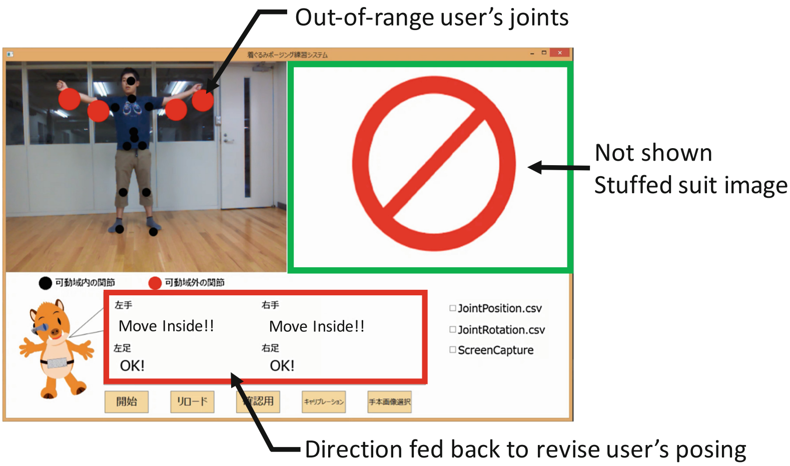

Fig. 9. Posing that cannot be done with a stuffed suit. (Color figure online)

parts shown by the red frame in the figure recognize which direction the body parts deviating from the motion range of the stuffed suit are deviated in the upward, downward, inward or outward directions. This application feeds back an instruction for pose correction to the user. Thus, users can learn the direction in which the pose should be corrected. By using this application, the users can be trained in posing while confirming the appearance image of the stuffed suits that is the closest to their posing in real time and can learn the motion range the stuffed suits allow.

6 Evaluation

6.1 Database Creation Phase

To use the system in the performed evaluation experiments, it was necessary to prepare the database in advance. In this experiment the principal author, whose height is about 175 cm, created the database that the subjects used to use the system. The procedure is shown below. First, the author wore the Xsens suit and the stuffed suit shown in Fig. 10 and performed the posing shown in Fig. 12. At this time, the author's skeleton data was measured by the motion capture system and the image of the stuffed suit was taken by a web camera. Data measurement was performed for 30 seconds, and a database of 300 data sets was created.

6.2 Experiment Environment

We conducted an experiment to evaluate whether the posing skill was improved by using our proposed system compared with the conventional method using

Fig. 10. The stuffed suit used in this experiment.

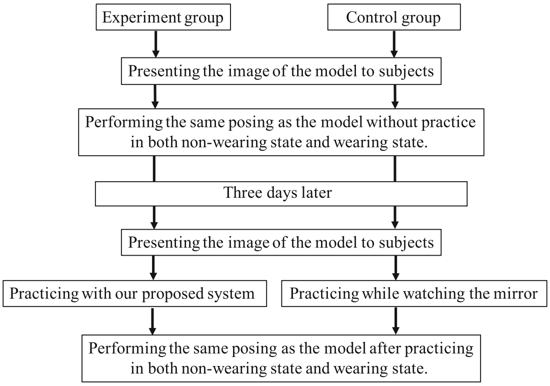

Fig. 11. Experimental procedure.

a mirror. In this experiment, the experiment group trained by using our proposed system and a control group were trained by watching a mirror. We evaluated whether the subjects' posing approached the models before and after the training.

The experimental procedure is shown in Fig. 11. Eight subjects who had had no experience with stuffed suits were randomly assigned to each of the experiment groups who received posing training by using our proposed system and the control group who received posing training with the conventional method by watching the mirror. The subjects' heights ranged from 165 to 180 cm. First, both groups were presented images for the three kinds of poses shown in Fig. 12. Second, the subjects who were not wearing a stuffed suit imagined the posing inside the stuffed suit and imitated the model posing. Then, the posing data of each subject was measured by using the motion capture system. Data measurements were performed at 17 joints of the body. Third, the subjects wore a stuffed suit and imitated the model posing. Finally, the posing data was measured by using the motion capture system and the appearance image was taken by using a camera.

Three days later, the experiment group received model posing training by using our proposed system and the control group was trained by watching the mirror. The control group trained in a comfortable environment to check their entire body because the mirror was positioned on a wall. We explained the function and usage of this application to the experiment group and they understood them and were trained in posing accordingly. Subjects performed various poses without wearing a stuffed suit so that images of the models could be displayed on the application. Posing training went on until the subject was satisfied. After practicing, subjects who were not wearing a stuffed suit performed the posing

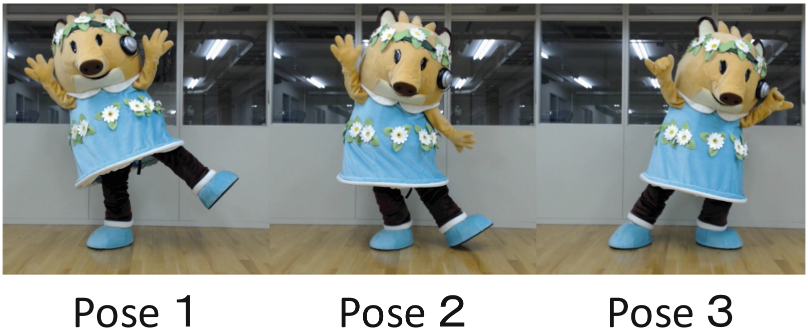

Fig. 12. Models for the experiment.

that they were trained in and the posing data of each subject was measured by using the motion capture system. Then, the subjects wore the stuffed suit and performed the posing that they were trained in. Finally, the posing data was measured by using the motion capture system and the appearance image was taken by using a camera.

We used two ways to evaluate our method, one a non-wearing state (i.e., a state when a stuffed suit was not worn) and the other a wearing state in which a stuffed suit was worn. In the non-wearing state, we evaluated whether subjects can learn the range of motion the stuffed suits allow them by practicing the posing with our proposed system. To be specific, we quantitatively evaluated whether the subject's posing approached that of the model data before and after the training by comparing the posing data of subjects and that of the model. Finally, we compared the results in terms of the transition of their numerical values before and after the training.

In the wearing state, we evaluated whether or not the subject's posing skill improved by using our proposed system. In addition to the quantitative evaluation described above, we carried out a subjective evaluation of the evaluators. In this evaluation, we conducted a questionnaire survey with a seven-level Rickard scale about how closely a subject's posing approached that of the models by using our proposed system. We compared subjects' posing scores before and after the training on a scale from one to seven, where one meant "close agreement" and seven meant "agreement not totally consistent." Ten different evaluators from among the subjects answered the questionnaire.

6.3 Results

Table 1 shows the obtained quantitative evaluation results and also the subjective evaluation results obtained from the evaluators. The quantitative evaluation values are the total differences between the joint angle data of the subject's posing and that of the model posing. Small values mean that the subjects' posing performance was close to that of the model posing. The subjective evaluation

Table 1. Quantitative and subjective results.

			Quantitative				Subjective					Quantitative				Subjective	
			Non-wearing		Wearing		Wearing					Non-wearing		Wearing		Wearing	
		Sub	Before	After	Before	After	Before	After			Sub	Before	After	Before	After	Before	After
Exp[1]	Pose 1	A	305	187	214	204	3.9	5.4	Con[2]	Pose1	E	457	263	258	207	5.9	4.8
		B	376	270	345	226	2.1	3.6			F	375	372	263	252	5.1	4.4
		C	329	253	309	271	3.7	4.8			G	336	362	162	206	1.9	2.2
		D	306	251	289	281	3.0	4.1			H	339	389	419	369	3.2	3.1
	Pose 2	A	276	252	365	290	2.9	3.8		Pose 2	E	490	315	332	251	2.3	3.7
		B	387	313	502	435	2.9	4.1			F	399	331	383	328	4.1	4.0
		C	308	243	309	303	3.6	4.2			G	318	257	407	282	5.2	5.2
		D	403	323	481	410	3.3	3.3			H	175	228	285	223	2.4	4.3
	Pose 3	A	569	359	385	362	3.5	3.8		Pose 3	E	530	480	367	402	4.1	4.0
		B	554	376	464	367	4.3	5.4			F	768	583	522	611	2.6	2.7
		C	550	387	368	428	2.8	4.1			G	511	436	392	370	2.8	3.1
		D	678	457	440	459	3.8	4.1			H	553	581	570	484	2.5	3.3

values are the average of the score of subjects' posing before and after the training. Large values mean that the posing the subjects performed was close to the model posing.

Figure 13 shows the quantitative results obtained for the transition of the differences from the model data in the non-wearing state. The values represent the average of the differences from the model data before and after the training. ◎ is $p < 0.01$, ◯ is $p < 0.05$ and △ is $p < 0.1$. The result with two-way ANOVA (analysis of variance) showed that the main effect of the training was significant ($F_{(1,22)} = 34.51, p < 0.01$) and interaction was significant ($F_{(1,22)} = 4.32, p < 0.05$). The simple main effect of interaction was as follows. The simple main effect of the method was significant only after the training ($F_{(1,22)} = 4.17, p < 0.1$). Therefore, the differences in the values from those of the model data for the experiment group were significantly smaller than those of the control group. Next, the simple main effect of the training was significant in the experiment

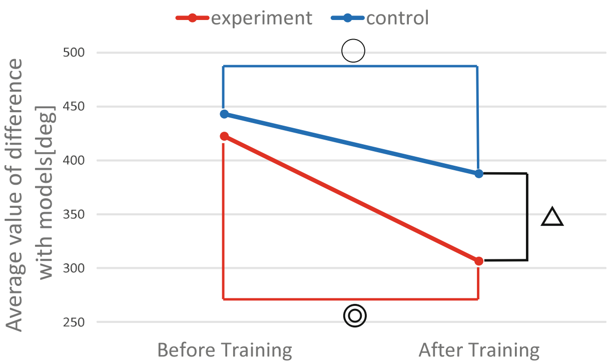

Fig. 13. Quantitative posing results obtained in non-wearing state.

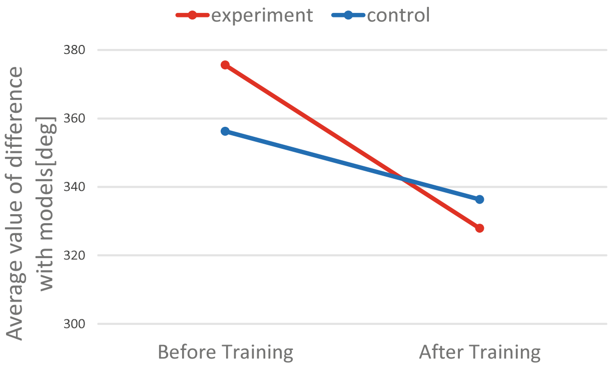

Fig. 14. Quantitative posing results obtained in wearing state.

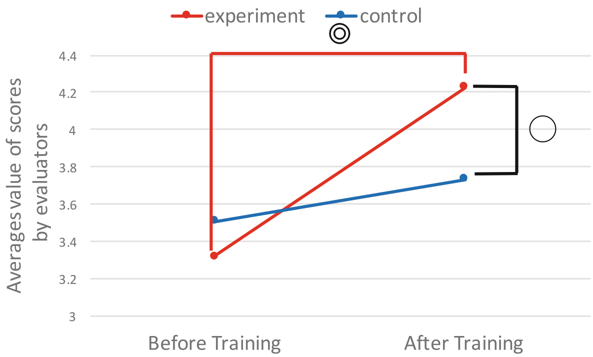

Fig. 15. Subjective posing results obtained in wearing state.

group ($F_{(1,22)} = 31.63, p < 0.01$). Therefore, the differences in values from the model data after the training were significantly smaller than those before the training. In addition, the simple main effect of the training was significant in the control group ($F_{(1,22)} = 7.20, p < 0.05$). Therefore, the differences in values from the model data after the training were significantly smaller than those before the training.

Figure 14 shows the quantitative results obtained for the transition of the differences from the model data in the wearing state. The values represent the average differences in the values from the model data before and after the training. The results obtained with two-way ANOVA showed that the main effect of the training was significant ($F_{(1,22)} = 11.70, p < 0.01$) but that interaction was not significant. Figure 15 shows the transition in the questionnaire scores before and after the training in the wearing state. The values represent the average difference in the questionnaire scores before and after the training. The results obtained with two-way ANOVA showed that the main effect of the training was significant ($F_{(1,6)} = 17.71, p < 0.01$) and that the interaction was also significant ($F_{(1,6)} = 7.17, p < 0.05$). The only significant differences we observed between the methods were those found after the training ($F_{(1,22)} = 6.33, p < 0.05$). The difference in values from the model data for the experiment group was significantly higher than that for the control group. In addition, the simple main effect of the training was significant only in the experiment group ($F_{(1,22)} = 23.71, p < 0.01$). The scores obtained with the experiment group after the training were significantly higher than those obtained before the training.

6.4 Discussion

The results obtained for the quantitative evaluation in the non-wearing state and those obtained through a subjective evaluation made by a third party confirmed that our proposed system is useful for posing training without a stuffed suit.

The results obtained from the quantitative evaluation in the non-wearing state confirmed that using our proposed system enabled the experiment group to learn the range of motion that the stuffed suit allowed better than the control group. Figure 13 indicates that the differences between the experiment group's posing data and that of the control group were significantly reduced before and after the training. We consider that compared with the control group, the experiment group was able to perform posing closer to that of the models. The experiment group was able to do so by learning the range of motion the stuffed suit allows with our proposed system. This indicates that users wearing a stuffed suit can perform the posing with the same body sensations as they felt in practicing the posing. They can also grasp the rough structure of stuffed suits, thus enabling them to perform the posing in the suit in accordance with the image they wish to convey.

The subjective evaluation results we obtained confirmed that the experiment group improved their posing skill better than the control group. Although no significant differences were found in the quantitative evaluation results obtained in the wearing state (Fig. 14, we found that for the subjective evaluation results (Fig. 15 the experiment group increased their scores much more than the control group did before and after the training. This indicates that the experiment group, by using the proposed system and learning the range of motion the stuffed suit allows them, was better able than the control group to perform posing close to the sample posing. This indicates that the subjects who trained with our system can pose appropriately in performances while wearing stuffed suits. Although the quantitative evaluation results obtained for the groups showed no significant differences, we consider that subjective evaluations from evaluators are important because the performances people see are those presented by people wearing stuffed suits. This is why we feel that it was important to find that using our proposed system enabled users in the experiment group to improve posing skills more than the control group could. We also confirmed that subjects can train properly even with a database provided by others (i.e., the author).

7 Conclusion

In this paper, we described a system we propose that enables users not wearing a stuffed suit to train themselves in posing by visually presenting them with images of performers wearing stuffed suits posing in the same way as the users will. In a prototype application of the system, we implemented a function that allows users to learn whether they can perform the posing they intend whether or not they are wearing a stuffed suit. Using our system allows users to train themselves in how to pose while getting the same body sensations whether or not they are wearing stuffed suits. Experiments with the system confirmed that it was better able than the conventional method to help users to improve their posing skills.

We used only one type of stuffed suit in our work, but there are various types of stuffed suits having many different kinds of structures. The more the costume structure differs from the human structure, the more our proposed system can be expected to be useful because it is difficult to imagine how persons will pose when they are inside the costume. We need to investigate whether our proposed system is useful for costumes other than the one we used in our work. We also plan to demonstrate our proposed system to professional performers in stuffed suits and get opinions and impressions from them.

Acknowledgement. This research was supported in part by a Grant in aid for Precursory Research for Embryonic Science and Technology (PRESTO) and CREST from the Japan Science and Technology Agency.

References

1. Okazaki, T., Terada, T., Tsukamoto, M.: A system for supporting performers in stuffed suits. In: Nijholt, A., Romão, T., Reidsma, D. (eds.) ACE 2012. LNCS, vol. 7624, pp. 85–100. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34292-9_7
2. Tei, Y., Terada, T., Tsukamoto, M.: A multi-modal information presentation method for performers in stuffed suits. In: Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia (MoMM 2014), pp. 77–84, November 2014
3. Slyper, R., Lehman, J., Forlizzi, J., Hodgins, J.: A tongue input device for creating conversations. In: Proceedings of ACM Symposium on User Interface Software and Technology (UIST2011), pp. 117–126 (2011)
4. Microsoft Kinect. http://www.xbox.com/ja-JP/xbox-one/accessories/kinect
5. Kyan, M., Sun, G., Li, H., Zhong, L., Muneesawang, P., Dong, N., Elder, B., Guan, L.: An approach to ballet dance training through MS Kinect and visualization in a CAVE virtual reality environment. ACM Trans. Intell. Syst. Technol. (TIST) **6**(2), 23 (2015)
6. Alexiadis, D.S., Daras, P., Kelly, P., O'Connor, N.E., Boubekeur, T., Moussa, M.B.: Evaluating a dancer's performance using Kinect-based skeleton tracking. In: Proceedings of the 19th ACM International Conference on Multimedia (ACM 2011), pp. 659–662, November 2011

7. Saha, S., Ghosh, S., Konar, A., Nagar, A.K.: Gesture recognition from Indian classical dance using Kinect sensor. In: Proceedings of Computational Intelligence, Communication Systems and Networks (CICSYN 2013), pp. 3–8, June 2013
8. Chan, J.C.P., Leung, H., Tang, J.K.T., Komura, T.: A virtual reality dance training system using motion capture technology. IEEE Trans. Learn. Technol. **4**(2), 187–195 (2011)
9. Tommi, H., Marc, R.T.: Learning and synchronising dance movements in South African songs: cross-cultural motion-capture study. Dance Res. **29**(2), 303–326 (2011)
10. Van Der Linden, J., Schoonderwaldt, E., Bird, J., Johnson, R.: Musicjacket-combining motion capture and vibrotactile feedback to teach violin bowing. IEEE Trans. Instrum. Measur. **60**, 104–113 (2011)
11. 3-Space Sensor. https://yostlabs.com/
12. Xsens motion capture system. https://www.xsens.com/
13. logicool HD PRO WEBCAM C920R. https://www.logicool.co.jp/
14. OpenCV. http://opencv.jp/