Development of Gesture Recognition Sub-system for BELONG

Increasing the Sense of Immersion for Dinosaurian Environment Learning Support System

Mikihiro Tokuoka[1], Haruya Tamaki[1], Tsugunosuke Sakai[1], Hiroshi Mizoguchi[1], Ryohei Egusa[2,3], Shigenori Inagaki[3], Mirei Kawabata[4], Fusako Kusunoki[4] and Masanori Sugimoto[5]

[1]*Department of Mechanical Engineering, Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba-ken, Japan*

[2]*JSPS Research Fellow, Tokyo, Japan*

[3]*Graduate School of Human Development and Environment, Kobe University, Hyogo, Japan*

[4]*Department of Computing, Tama Art University, Tokyo, Japan*

[5]*Hokkaido University, Hokkaido, Japan*

Keywords: Kinect V2 Sensor, Immersive, Learning Support System, Gesture Recognition.

Abstract: We are developing an immersive learning support system for paleontological environments in museums. The system measures the body movement of the learner using a Kinect sensor, and provides a sense of immersion in the paleontological environment. Conventional systems are only able to recognize simple body movements, which is insufficient to completely immerse learners in the paleontological environment. On the other hand, when they need to perform complicated body movements, learners move their bodies eagerly while thinking. This emphasizes the importance of developing a sub-system capable of recognizing complicated body movements. In this paper, we describe a sub-system developed to recognize the body movement of the most important learner in the immersive learning support system.

1 INTRODUCTION

Museums are important places for children to learn about science (Falk, J. H., 2012). They also operate as centers for informal education in connection with schools, and they enhance the effectiveness of scientific education (Stocklmayer, S. M., 2010). However, because the main learning method within museums is to study the specimens on display and read their explanations, there are few opportunities for learners to observe or experience the environment about which they are learning. In particular, it is impossible to personally experience a paleontological environment, because it includes extinct animals and plants and their ecological environment (Adachi, T., 2013). It is difficult for children to learn about such environments merely by viewing fossils and listening to commentary. Overcoming this problem would qualitatively improve scientific learning within museums. As for these problems, a system that simulates a paleontological environment and transitions that would be impossible to experience in reality would solve the problem. The system would also be required to enhance learners' sense of immersion.This explains the need for such a system.

In order to enhance the sense of immersion, a full body interaction interface in which the movement of the whole body is linked to the operation of the system has been shown to be effectice (Klemmer S., 2006).

We aim to realize an immersive learning support system "BELONG" (Tokuoka M., 2017). This system is operated using complicated body movements as observational behavior, thereby allowing learners to enhance their sense of immersion in a paleontological environment. The most important function in this system is recognition of the learner's body movement. This required us to develop a sub-system specifically for this purpose. "BELONG" is intended to be implemented in a museum. In situations with limited funds and space such as a museum, a low-cost space-saving sub-system that can recognize body movements was required.

Hence, we thought to use a Kinect v2 sensor, which is a three-dimensional range image sensor that is highly affordable and does not consume much space, unlike motion capture equipment. Many

Tokuoka, M., Tamaki, H., Sakai, T., Mizoguchi, H., Egusa, R., Inagaki, S., Kawabata, M., Kusunoki, F. and Sugimoto, M.
Development of Gesture Recognition Sub-system for BELONG - Increasing the Sense of Immersion for Dinosaurian Environment Learning Support System.
DOI: 10.5220/0006357304930498
In *Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017) - Volume 1*, pages 493-498
ISBN: 978-989-758-239-4

conventional full body interaction interfaces use a three-dimensional distance image sensor. These interfaces achieved recognition by computing the trajectory of the skeleton when recognizing body movement. However, this existing method could easily perform recognition in the case of simple body movement, but was less effective when complicated body movements were carried out. Moreover, it is difficult to recognize complicated body movements only from the trajectory of the skeleton.

In this paper, we summarize a sub-system capable of recognizing the body movement of the most important learner in the immersive learning support system named "BELONG."

2 SUB-SYSTEM FOR BELONG

2.1 BELONG

We aim to realize the immersive learning support system "BELONG" that simulates a paleontological environment and transitions that are impossible to experience in reality for efficient learning at the museum (Tokuoka M., 2017). Figure 1 illustrates the concept of "BELONG." This system accepts body movements as input for observational behavior. The movements of the whole body are linked to the operation of the system; hence, it is possible to enhance the sense of immersion in the paleontological environment. This sense of immersion improves if the system can be operated in conjunction with complicated body movements as compared with a case in which the system is operated with simple body movements. In this system, we utilize Microsoft’s Kinect v2 sensor, a range-image sensor originally developed as a home videogame device. This enables us to provide a low-cost immersive learning experience within a small space, because "BELONG" comprises only this Kinect v2 sensor, a projector, and a control PC. This arrangement has the advantage that it is possible to easily change the learning contents. Moreover, we recognize the body movements of learners by gesture recognition using the Kinect v2 sensor. The gesture recognition sub-system, which also has the ability to recognize more complex body movements, registers the body movement the creator wishes to recognize and judges whether it is recognized by assessing the similarity with the body movement.

Figure 1: Concept of BELONG.

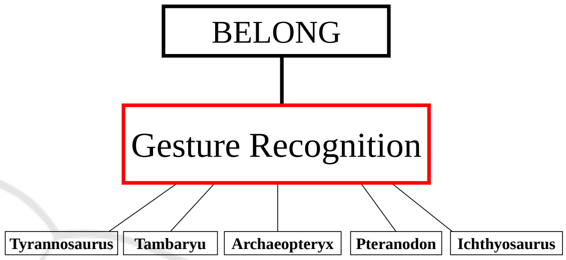

Figure 2: Sub-system for BELONG.

2.2 Sub-system

We aim to realize the immersive learning support system "BELONG" that simulates a paleontological environment and transitions that are impossible to experience in reality. As a first step towards the realization of this system, we developed a sub-system that recognizes complicated body movements. The immersive learning support system "BELONG," allows children to learn about five dinosaurs, Tyrannosaurs, Tambaryu, Archeoptery, Pteranodon, and Ichthiosaurus. Realization of this learning exercise required the ability to recognize complicated body movements. Conventional systems can only recognize simple body movements, which is insufficient to completely immerse learners in the paleontological environment. On the other hand, by performing complicated body movements, learners move their bodies eagerly while thinking. We thought that this would enhance the sense of immersion and result in efficient learning. Therefore, the development of a sub-system that recognizes complicated body movements is important. Consequently, we adopted the movement associated with excavation as a complicated body movement for the sub-system. We thought that learners' interest would rise by virtually excavating the fossils being exhibited. When the excavation proceeds successfully, videos showing the characteristics of the particular dinosaur are shown. This enables learners to learn about the kind of dinosaur to which the excavated fossils belong.

"BELONG" is intended to be implemented in a museum. In the case of limited funds and space as in a museum, a subsystem that can recognize body movements with low cost and space saving is required.

We therefore created a sub-system to achieve these objectives. The sensors used in this sub-system and their functions are described below.

2.2.1 Kinect V2

Recognizing human body movements requires us to use a sensor to recognize these movements as well as the position of a person. There are a number of sensors that could be used for this task. For example, there are motion capture sensors and also a three-dimensional range image sensor.

A motion capture sensor is expensive and requires a large space for installation. First, we would need to surround the space we want to recognize with multiple cameras. Next, we would have to calibrate the space. The next step would involve attaching a marker that reflects infrared rays in the wavelength range from 3 [mm] to 200 [mm] to a person, thus recognizing the person's body movement and position. However, the sensor is expensive and requires a wide space for installation, neither is it suitable for implementation in a museum.

The Kinect v1 sensor of a three-dimensional distance image sensor is inexpensive and it can be mounted simply by installing a single sensor. Moreover, a person's position can be estimated by the depth sensor, and a total of 20 measurements relating to skeletal information are possible. This information is then used to calculate the trajectory of the skeleton. Thus, the system can be developed based on this information.

(Example of system) Suppose you raise your hand when HAND - RIGHT is above HEAD.

This is an example of a condition for recognizing the simple body movement of raising a hand. Although, as mentioned above, simple bodily movements can easily be recognized, it is difficult to recognize complex body movements when the only information that is available is that relating to the 20 skeletal points. Moreover, it is time consuming to decide the conditions for implementation. Thus, the Kinect v1 sensor is not suitable for recognition of complicated body movements of learners who are the most important part of the immersive learning support system "BELONG."

The Kinect v2 sensor of the three-dimensional range image sensor is as inexpensive as Kinect v1, and it can be mounted by simply installing one sensor. Moreover, the position of a person can be estimated by the depth sensor, and information relating to a total of 25 skeletal points can be measured (Figure 3). Therefore, by recording this skeletal information and by performing machine learning, it is possible to easily create a discriminator of gesture recognition by using the Kinect v2 sensor (Shotton, J., 2011). This discriminator can detect various body movements upon request. Hence, we used the Kinect v2 sensor to enable the sub-system to recognize complicated body movements, because it is inexpensive, space saving, and offers fast implementation.

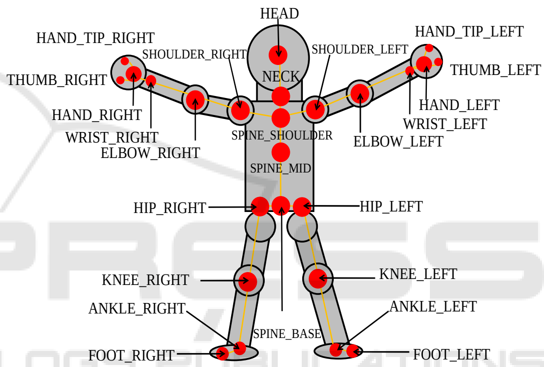

Figure 3: Information about the skeleton.

2.2.2 Gesture Recognition

The above-mentioned Kinect v2 sensor, which is used to recognize complicated body movements, provides the tools Kinect Studio and Visual Gesture Builder. These tools are used to perform complicated body movements recognition. Kinect Studio can record position information and motion information of a person's body, as acquired with the Kinect v2 sensor. Visual Gesture Builder is a tool that can create a discriminator for recognizing a gesture using machine learning based on information about the position and motion of a person`s body.

Kinect Studio functions as follows. Three-dimensional information about the position and motion of a person's body is acquired during arbitrary body movement, as shown in Figure 4. This three-dimensional information comprises the coordinates of a total of 25 skeletal points shown in Figure 3. When recording is started, these three-dimensional coordinates are automatically recorded as training data. This training data can be recorded by any number of people.

This training data is used in Visual Gesture Builder to create a discriminator of body movements. This training data can be used for machine learning by Visual Gesture Builder. First, the recorded data is played back in Visual Gesture Builder and then the frame of the motion you want to recognize is specified. Next, the body movements you want to recognize are clipped, as shown in Figure 5. After building the frame once it has been specified, the classifier is learned based on the three-dimensional information of the person's body. In this way, it is possible to learn discriminators of arbitrary gestures using the Visual Gesture Builder. By reading the discriminator learned using the Visual Gesture Builder, it is possible to create a program that recognizes the gesture with the Kinetic v2 sensor.

In this way, it is possible to easily recognize complicated body movements.

Furthermore, acquiring data of various people for inclusion in the training data serves to improve the ability of the discriminator to recognize body movements. Because the height and physique of the people who experience this system are various, it is necessary to perform body movement even if the three-dimensional information is largely different. Therefore, it is necessary to acquire various training data by implementing it at a museum, allowing it to be experienced by many learners, and acquiring data. As the number of learners increases, it becomes a good discriminator and the immersive feeling of learners can be enhanced. Thus, the Kinect v2 sensor can be used to easily develop the system that accurately recognizes complicated body movements.

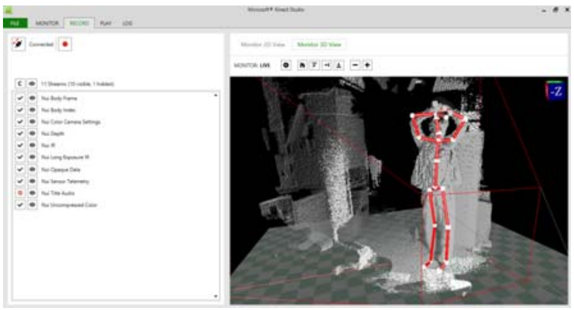

Figure 4: Kinect Studio.

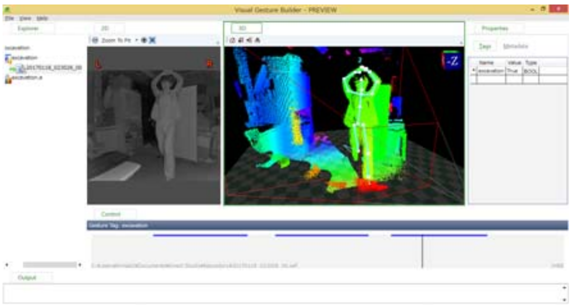

Figure 5: Visual Gesture Builder.

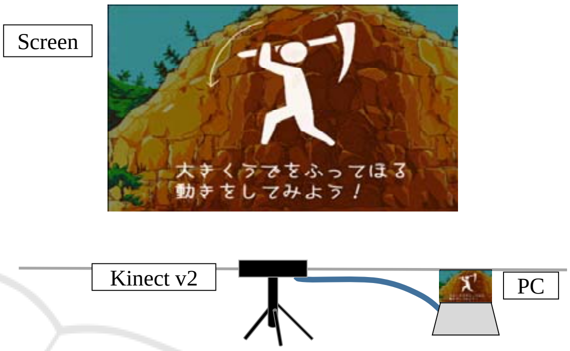

Figure 6: Installation status of sensor and projector.

3 EXPERIMENT

3.1 Experimental Environment

Using the sub-system we developed, we conducted an experiment to confirm whether it would be possible to use the gesture discriminator to recognize the complicated body movement associated with excavation. Figure 6 shows the Kinect v2 sensor, control PC, and screen that were installed. We asked children to carry out the excavation movements many times in front of Kinect v2. At this time, we confirmed whether the following functions were realized.

・Recognition of excavation movement.

・Changes on the screen according to recognition.

Figure 7 shows a successful example in which the excavation operation could be recognized using this sub-system.

3.2 Experimental Result

We evaluated whether the developed sub-system operated normally for learners. Figure 8 shows a situation in which children are experiencing this sub-system. As the figure shows, the children were excavating by making large gestures towards the screen.

The following opinion is that of one of the participating children.

・This fossil belonged to that kind of dinosaur.

・Fun to excavate.

・I wanted to know about that fossil by excavating it.

Thus, we recognized the possibility to encourage children to become interested in learning about fossils and dinosaurs. We could also confirm the effectiveness of the sub-system we developed by using gesture recognition.

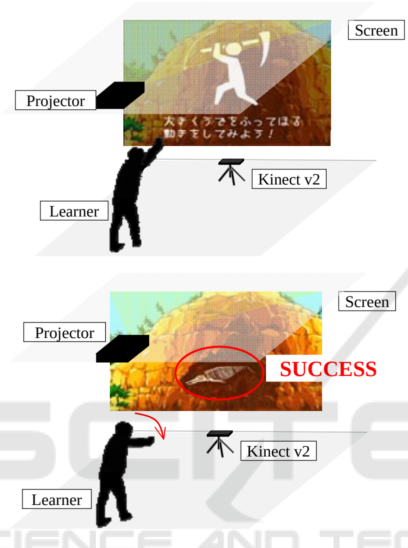

Figure 7: Sub-system.

4 CONCLUSIONS

In this paper, we described the sub-system we developed to recognize the body movements of the most prominent learner in the immersive learning support system named "BELONG." We carried out experiments to confirm whether the developed sub-system satisfies the function of recognizing the most significant complicated body movements using "BELONG." As a result, when the children performed complicated body movements to conduct the excavation, the sub-system responded as intended, thereby confirming that the system was operating successfully. This suggested that the sub-system intended to recognize the body motion of a learner using the gesture recognition function of the Kinect v2 sensor is effective.

ACKNOWLEDGEMENTS

This work was supported in part by Grants-in-Aid for Scientific Research (B). The evaluation was supported by the Museum of Nature and Human Activities, Hyogo, Japan.

REFERENCES

Falk, J. H., Dierking, L. D., 2012. *Museum Experience Revisited*. Left Coast Press. Walnut Creek California, USA, 2nd edition.

Figure 8: Evaluation experiment.

Stocklmayer, S. M., Rennie, L. J., Gilbert, J. K. 2010. The roles of the provision of effective science education, *Studies in Science Education*, 46(1), pp. 1-44, Routledge Journals, Taylor & Francis Ltd.

Adachi, T., Goseki, M., Muratsu, K., Mizoguchi, H., Namatame, M., Sugimoto, M., Takeda, Y., Fusako, K., Estuji, Y., Shigenori, I., 2013. Human SUGOROKU: Full-body interaction system for students to learn vegetation succession. *Proceedings of the 12th International Conference on Interaction Design and Children.* pp. 364-367, ACM.

Klemmer S., Hartmann B., Takayama L., How bodies matter: five themes for interaction design, *DIS '06 Proceedings of the 6th conference on Designing Interactive systems,* pp. 140-149.

Tokuoka, M., Tamaki, H., Sakai, T., Mizoguchi, H., Egusa, R., Inagaki, S., Kawabata, M., Kusunoki, F., Sugimoto, M., 2017. BELONG: Body Experienced Learning support system based ON Gesture recognition. *Proceedings of the 9th International Conference on Computer Supported Education, CSEDU 2017.* To appeared.

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A. 2011. Real-Time Human Pose Recognition in Parts from a Single Depth Image. *Communications of the ACM.* 56(1), pp. 116-124, ACM.