International Journal of u- and e- Service, Science and Technology
Vol.8, No. 7 (2015), pp.99-108
http://dx.doi.org/10.14257/ijunesst.2015.8.7.10

Motion Tracking System for Multi-User with Multiple Kinects

HwanIk Jo, HuiBeom Yu, Karam Kim and Jung Hwan Sung[1]

[1]Information Science Department, Soongsil University,
Sangdo 1-dong, Dongjak-gu, Seoul, Korea
{vivid, cutyboyjin, kkr, artbysung}@crossdesignlab.com

Abstract

User motion recognition through image recognition has attracted considerable attention with the development of the Kinect senor by Microsoft. The Kinect exhibits excellent motion recognition performance at a small expense via a motion tracking method utilizing an RGB-D camera. Further, the Kinect offers the advantage that applications can be easily written for it using the Kinect SDK. However, the Kinect also suffers from the disadvantage that it is unable to recognize an image "shielded" by an object because of the application of the vision recognition method by the camera. Consequently, it is difficult to obtain accurate data concerning user movement. Studies on observing an object from various directions with multiple Kinects are being conducted to compensate for this problem, and it is possible to acquire the movement data for only one user in these studies. This study implements the human body trace system for multiple users. We suggest a method to compensate for the location of the user and motion data that cannot be recognized by one Kinect using data acquired by multiple Kinects.

Keywords: *Motion Tracking, Position Tracking, Multi-User, Multi-Kinect, Unity3D, Skeleton Data*

1. Introduction

The interaction method using natural movement without direct sensors or hardware input has recently become an important factor that enables users to play and enjoy the content of video games. Since Kinect of Microsoft was developed, motion tracking by image recognition, among interaction methods using natural movement, have received special consideration.

Although the Kinect sensor was originally developed as a game interface for the XBOX, which is a game console developed by Microsoft, it is now utilized actively in the motion tracking of a user since the Kinect software developer kit (SDK) allows developers to easily write Kinect apps. Further, the Kinect offers an excellent motion tracking performance. In the interaction method using Kinect, there are options to either recognize user gestures using the data for human joints or to manipulate the UI using the movement of the hands.

However, the Kinect uses a vision recognition method with a camera and recognizes only two-dimensional data acquired through the camera; it cannot recognize three-dimensional data. Because of this problem, researchers are now focusing on methods to compensate for the data that one Kinect cannot recognize; this involves the use of a second Kinect that acquires motion data from other directions. In these studies, [6, 18, 19, 20] researchers have suggested an approach to extract data with increased accuracy for each joint from the data obtained through multiple Kinects using various matching methods.

However, in these researches, motion tracking with increased accuracy can be achieved for only one user. In the case of multiple users, user motion may be "concealed" from the recognition range of the user's Kinect due to the overlapping body movements of multiple

ISSN: 2005-4246 IJUNESST

users. In this case, the “missing” motion data can be traced by using the motion data acquired by the other Kinects since one Kinect cannot recognize (or decipher) the motion data of all users. In this study, we develop an approach to continuously trace the movement of each user and extract the motion data within the user interaction space using the data acquired from multiple Kinects.

2. User Trace Technique with Multiple Kinects

Many studies have used multiple Kinects in order to compensate for joint data that cannot be recognized by one Kinect [6, 18-20]. These studies using multiple Kinects have considered various parameters such as the advantages, disadvantages, and data matching methods to address the issue of accurate motion tracking. In the study by Brian M. Williamson (2012) [20], the Skeleton Data of an user were integrated together using four Kinects. Further, Jinho Lee (2014) integrated the Skeleton Data of the user in question using three Kinects. This approach made it possible to more precisely trace the user’s motion, with the application of weights to each joint of the skeleton.

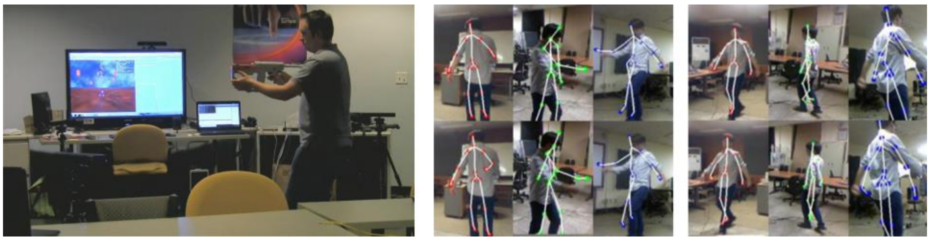


Figure 1. Multi-Kinect Tracking for Dismounted Soldier Training (Left) Advanced Tracking Method Using Multiple Kinects (Right)

Many studies have proposed advanced tracking methods for tracking the human body. However, most studies have thus far considered tracking of only one user; multi-user tracking still utilizes a two-dimensional space in these studies, and the problem observed in experiments using one Kinect still remains unresolved.

Some studies have extracted the detailed data of one person using multiple Kinects, and these experiments utilize a method to compensate for the joint data that are “shielded” by the human body from the other directions. In the case of multiple users, because one user sometimes covers or shields the recognition range of the other users, the Skeleton Data of this user cannot be recognized by all the Kinects. The Skeleton Data should be mapped suitably to each user because the amount of data acquired increases when using four Kinects as the number of users increase. In order to solve this problem, in our study, we use a system that tracks the position and movement of each user using the Skeleton Data obtained by all the Kinects in use. In position tracking, our method clubbed Skeleton Data in the same position and direction as one group. Only one User Id is assigned to the grouped data to enable continuous tracking during one Kinect to recognize the user. This approach also allows for recognition of the movement information and motion data of each user via integration of the data mapped to one user id into one joint data set.

3. Development of Human Body Tracking System for Multiple Users

3.1. Configuration of Multiple-Kinect-tracking Environment

In our study, the three-dimensional tracking space was configured by positioning four Kinects along the four corners of a square with a side of 5 m. The network was configured such that each Kinect was connected to its own PC and the data from all the Kinects were sent to one server. The client PC recognizes the Skeleton Data of the user via the Kinect SDK. Each joint is weighted in the client PC using the “Tracked, Inferred, Not Tracked” data of the Kinect SDK and the data corresponding to the position of the joint, its rotation, its weighted value, and type were transmitted to the server through the OSC (Open Sound Control) protocol.

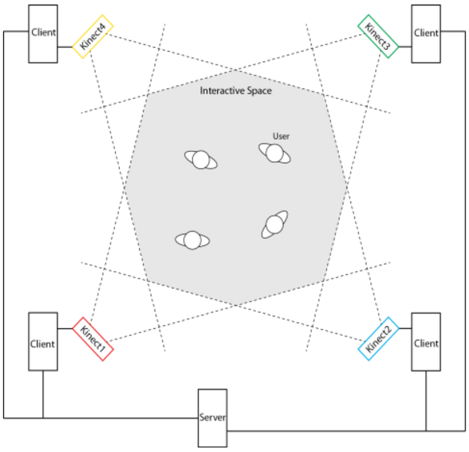


Figure 2. Human Body Tracking System for Multiple Users

The server PC reconfigures the Skeleton Data into three-dimensional images in the Unity3D platform using data obtained from the four client PCs. The data of four client PCs are integrated into the virtual space of Unity3D using the rotation and position for the each Kinect in the interaction space environment.

In order to extract the integrated Skeleton Data of each user obtained through the four Kinects, the method uses positioning data of Skeleton Data. It integrates all Skeleton Data that are in close distance using the positioning data. The integrated data are transmitted again to the client PC through the OSC, and each client PC is able to use the user joint data received from the server PC.

3.2. System Data Process

The user recognized by the four Kinects is transformed into Skeleton Data in the client PC and the data corresponding to each joint are transmitted to the server through the OSC.

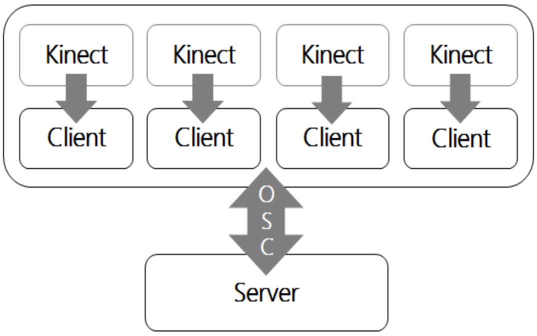


Figure 3. System Data Process

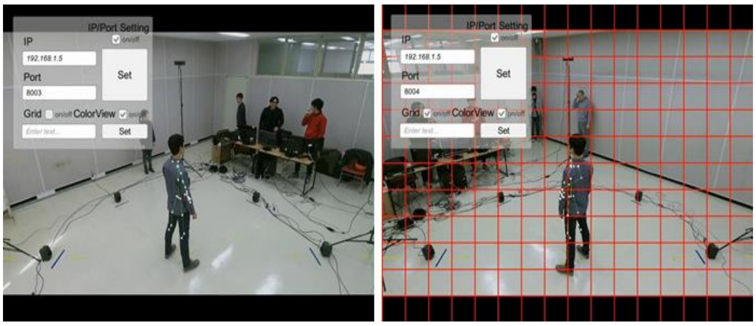
The server PC extracts the data of the position and joint of each user in the interaction space through the four sets of three-dimensional data obtained through the OSC as acquired by the Kinects.

All the user data extracted are transmitted to the four client PCs, and each client PC enables position tracking through continuous connection of the user id using this data.

3.2.1. Client

The client PC saves and transmits the Skeleton Data id, user id, types of each joint, joint position, and weighted data of the object recognized by the Kinect. The basic data supplied by Kinect SDK include the "Tracked, Inferred, Not Tracked" with weights of 1.0, 0.1, and 0.01, respectively. The Inferred and Not Tracked data are assigned low weighted values because it is possible to obtain the error data based on the estimated result for the joint position and rotation that were not recognized by the Kinect SDK.

For the Not Tracked data, the correct weight to assign is zero; however, we used a weight value of 0.01 instead because the "Not Tracked" joint data becomes zero and cannot be integrated when the Skeleton Data of the user is recognized by only one Kinect. Further, when one object is recognized continuously through the Kinect, the client PC matches the skeleton id and user id and transmits it to the server PC.



IP/Port Setting
IP ☑ on/off
192.168.1.5
Port
8003
Set
Grid □ on/off ColorView ☑ on/off
Enter text...
Set

IP/Port Setting
IP ☑ on/off
192.168.1.5
Port
8004
Set
Grid ☑ on/off ColorView ☑ on/off
Enter text...
Set

Figure 4. Views from Client4 (Left), Client3 (Right)

The object recognized by the Kinect has its own skeleton id. The skeleton id is randomly determined by the Kinect SDK, and it removes with the Skeleton Data in case the object cannot be recognized because it disappeared from the line of sensor sight or is out of the range of recognition. Each skeleton has its unique skeleton id, and the user id can be continuously tracked using this skeleton id.

3.2.1. Server

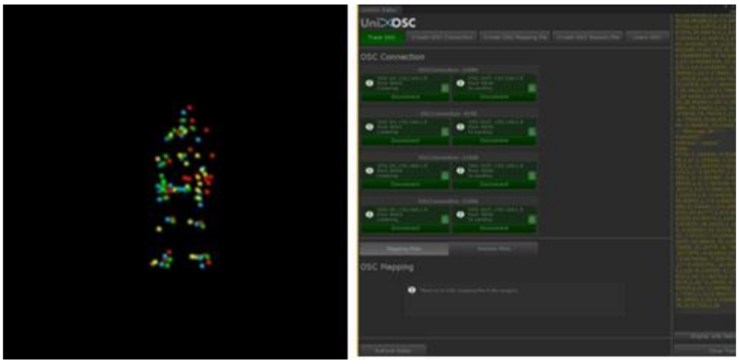


Figure 5. Visualization of the Data from four Kinects in Unity3D and OSC Console Windows

The server receives data from each client through the OSC, and it marks the object data corresponding to Kinect1 with red color. The object corresponding to Kinect2 is marked in blue, that for Kinect3 is marked in green, and that for Kinect4 is marked yellow. Each of these data streams are then visualized in the virtual space. In the OSC, the IP and corresponding port were configured to "OSC In and Out," which enables communication with the four client PCs. The server PC first performs an initialization that matches the three-dimensional space obtained from the four Kinects to the Unity3D platform using the joint data of one user who is positioned at the center of the square and equidistant from the four Kinects. This initialization creates the data space for each Kinect using the angles designated as up/down/left/right from both the actual environment in which the Kinect is installed and the joint data of each user, and further, the initialization enables integration of the space data from the four Kinects with one Kinect. The server calculates the number of users in the current space using the position of each Skeleton Data set from the integrated data and saves the Skeleton Data of each user within a certain range as a Skeleton Data group. These several stored Skeleton Data sets are integrated into one Skeleton Data set, and larger weighted values are applied to the Kinect data corresponding to position in which the user directly faces the Kinect.

This procedure is performed because many errors occur with respect to the joint data when the Kinect recognizes the user at oblique angles. The Kinect obtains the overlapped depth data when the user’s arm and body overlap when it “looks” at the user at oblique angles. This allows the extraction of the incorrect estimated data for each joint.

3.2.1.1. Position Tracking of User

Several researches have focused on user tracking with a camera. In this study, we track the user by utilizing the positioning data corresponding to the middle of the spine from the Skeleton Data.

When the movements of two users overlap, the assigning of the user id enables its continuous tracking through the position recognition of other cameras. In this study, the user id of each user is continuously mapped by the three-dimensional position of the Skeleton Data obtained by the four Kinects.

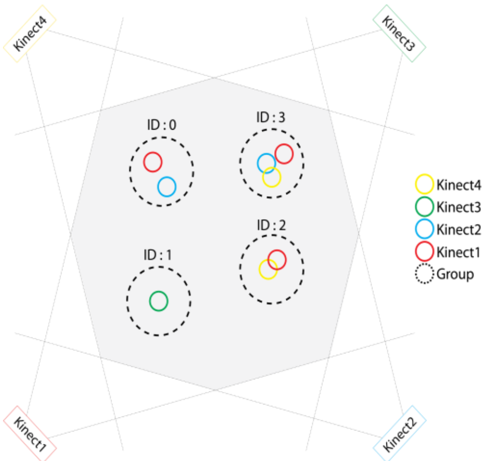


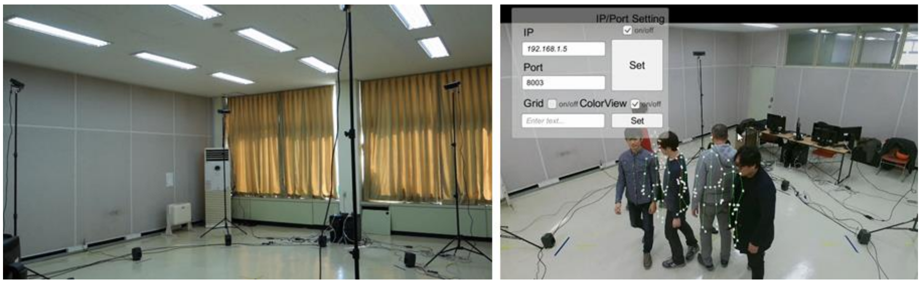
Figure 6. User Id Recognition of Skeleton Data Group using the Position of Skeleton Data

The approach recognizes the Skeleton Data of users looking in the same direction within a specified range as one group in order to continuously track the positions of the users. The specified range is obtained through the movement of one user within the interaction space after integration of the data space of the four Kinects. Each Skeleton Data group has a unique user id that is transmitted to the client PC corresponding to the specific skeleton id. Once connected, the user id transmits the data to the server PC by connecting the skeleton id of the client PC and user id until the server sends modified data. If the users “overlap,” with two users in the designated range, it is possible to track the location of two users as the unique user id is preserved even when the Skeleton Data of only one user is available. A user id is generated when the Kinect first recognizes the user, and except for the case when four Kinects cannot recognize one user, the Skeleton

Data integrated with the position of each user can be obtained since the skeleton id and user id are continuously connected.

3.3. System Experiment

3.3.1. Experiment Environment



IP ☑ on/off
192.168.1.5
Port
8003
Grid □ on/off ColorView ☑ on/off
Enter text...

Figure 7. Experiment Environment

Since the range of the Kinect V2 is a maximum of 4.5 m, four Kinects are positioned at angles of 45°, 135°, 225°, and 315° (Figure 7) within a square with a side of 4 m. Each Kinect is additionally connected to a rotation table for fine-tuning the up and down angles. We examined two cases: single user and multi-user. In the single-user experiment, the accuracy of position tracking and the change in bone length was displayed as a graph by tracking the body position under movement and extracting the bone length variation. The multi-user experiment involved the movement of four users. The position tracking of each user and the change in bone length were displayed as a graph by tracking the position of each user and extracting the bone length change in a circular space with a diameter of 4 m. In the experiment, the tracking accuracy was judged as follows: the position tracking data were compared with the recorded movement video, and the acquired tracking data were considered as correct if the bone length variation as per the video and the data were identical.

3.3.2. Experiment Result

3.3.2.1. Single-user Experiment

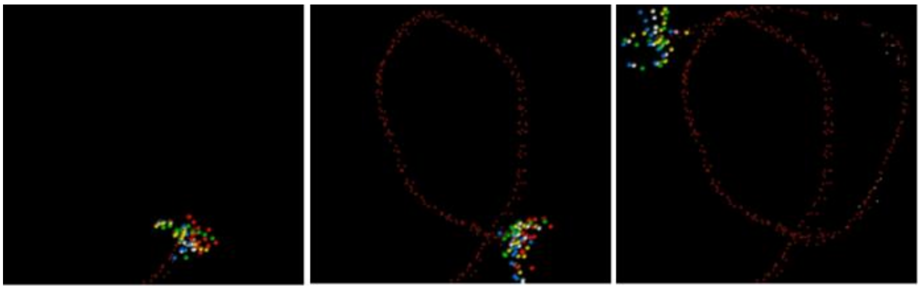


Figure 8. Situation Depicting the Movement of a Single-user as Experiment Image

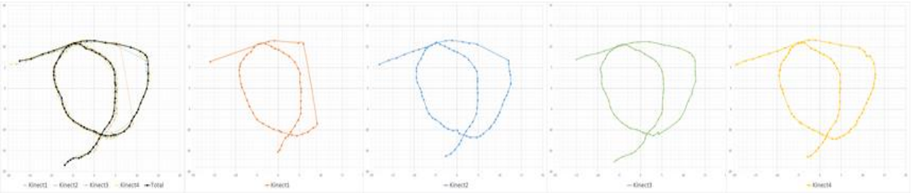


Figure 9. Movement Tracking Data in the Case of Single user (Result of Integration of Data Acquired by Kinect1, Kinect2, Kinect3, and Kinect4)

Figure 9 shows the position tracking data in the single-user case in the form of a graph. The correctly recognized data for one user could be obtained from all the Kinects except Kinect1. The unique user id of the user could subsequently be continuously maintained, and this enabled exact position tracking. Since the movement data of one user is not affected by "shielding," the absence of positioning data indicates that the Kinect in question did not recognize the user's back or that Kinect could not generate the Skeleton Data of the user due to the depth data error.

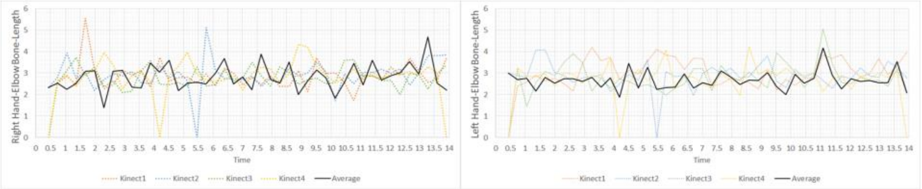


Figure 10. Bone Length Change Mapped in Single-user Case

Figure 10 depicts the data concerning the change in bone length during arm movement within the recognition range in the form of a graph. The average of the four sets of data is relatively accurate when compared with the data from each Kinect.

3.3.2.2. Experiment with Four Users

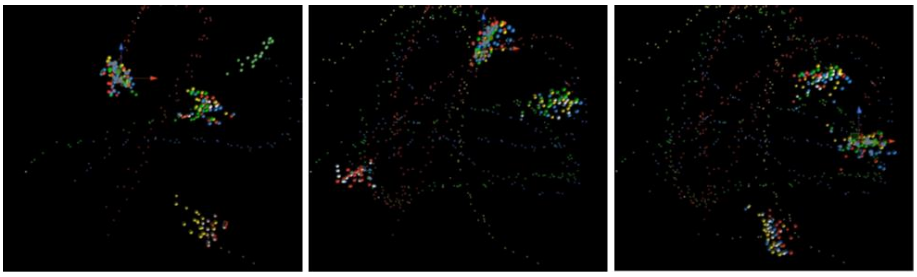


Figure 11. Experimental Image Depicting Movement of Four Users

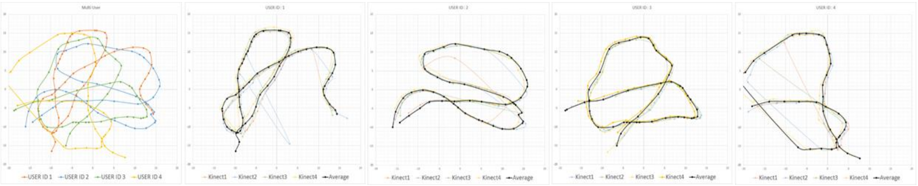


Figure 12. Movement Tracking Data for Four Users

As regards the results of extracting the position tracking data of the user, it was possible to continuously track the user id of each user. In the case of multiple users, the problem of the "shielding" of the Skeleton Data often occurs. Thus, the Skeleton Data is acquired and lost repeatedly. During such an error situation, if one of four Kinects retains the user id of the user, then it is possible to track the position of each user because the user id for multiple users does not reassigned.

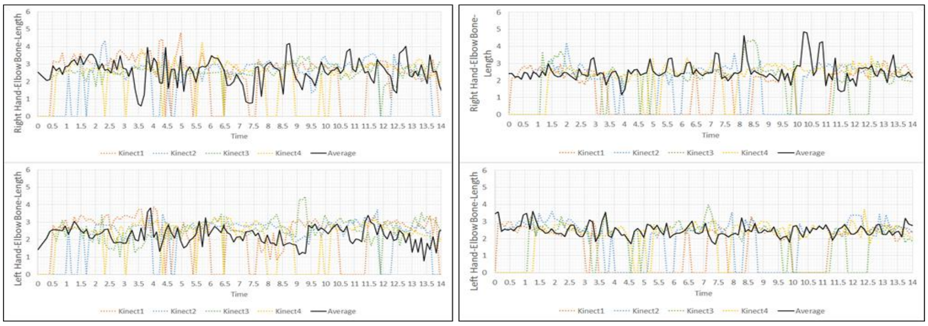


Figure 13. Bone Length Change (ID 1 (Left), ID 2 (Right))

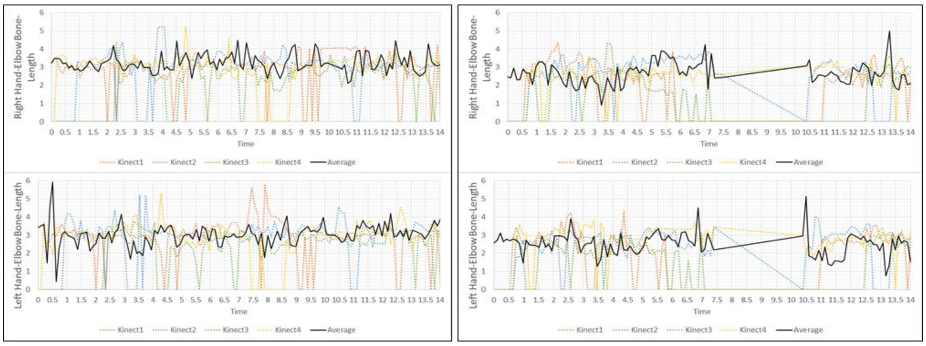


Figure 14. Bone Length Change (ID 3 (Left), ID 4 (Right))

We found that as regards the user with user id 1, we could obtain relatively accurate Skeleton Data from Kinect3, but we could not obtain a reasonable amount of data from Kinects1, 2, and 4. The error data were extracted from Kinect2, but these error data did not influence the accuracy of the movement path of the user. It was possible to remove the error since the weighted value of the Kinect2 data was assigned to a value smaller than that corresponding to the data acquired by the "observing" Kinect. The user with user id 2 obtained the correct data from Kinect4, and smaller amounts of data were obtained from Kinects1, 2, and 3. The recognition rate of the user with user id 3 was high in most of the Kinects. The user with user id 4 obtained the correct data from Kinects2 and 4, and the data corresponding to the user's left corner were lost because of shielding by another user. We speculate that data were lost in this case because the user was either shielded by another user or that the Kinect observed the user's back.

4. Conclusion

In this study, we proposed a tracking method for multiple users by utilizing multiple Kinects. We identified the problems in existing human body tracking research and resolved these problems using multiple Kinects. This method allows the continuous extraction of Skeleton Data through movement tracking using the data of the users of four Kinects. In our experiments, it was possible to integrate the Skeleton Data with via the continuous mapping of each user's user id using the three-dimensional data obtained from the four Kinects. Further, we were able to acquire the exact data in a small space with fewer users, and the data accuracy increased over a wider range of space as the number of users increased. We believe that this system allows the formation of an interaction space for multiple users. However, many problems encountered in our study require further investigation.

First, as regards the position recognition of a user through a Kinect, depth data errors occur when the movements of two users overlap over long intervals. Errors also occur with regard to the continuous user id connection when one user is not recognized by all four Kinects. This requires a system which, in addition to user location recognition through the Kinect, can maintain the user's user id under any situation.

Second, the amount of obtainable data from a Kinect is reduced as the number of users increases; the user id was not continuously recognized in the experiment with more than five users. In order to solve this problem, we need to develop a system that enables continuous connection even in the situation that the user is not recognized by the Kinects.

References

[1] O. H. Jo, "A study on implementation of gesture recognition applying to interactive media art", Digital Design Research, **(2012)**.

[2] D. B. Kim, "Development in rhythm learning program using motion recognition device", Kunkook University, Educational Graduate School, **(2014)**.

[3] E. B. Oh and S. T. Ryu, "PC based drawing using Wii_remote", Korea Contents Academy Academic Conference Paper Books, **(2009)**.

[4] M. J. Kim, J. M. Hu, J. H. Kim and S. Y. Park, "Development & evaluation of game interface based on lip motion considering intuitive hand motion", Korea Game Academy Paper Magazine, **(2014)**

[5] J. Sin and S. Y. Lee, "Image based structured distance measurement sensor strong in lighting noise", Light, Electric Facility Academy Papers Book, **(2010)**.

[6] J. H. Lee, "Advanced human body tracking method using multiple Kinects", Sejong University Graduate School, **(2014)**.

[7] W. C. Park, D. H. Ryu and T. H. Choi, "Extraction and transmission of gesture information using ToF camera", Korea Electronic Communication Academy Paper Book, **(2014)**.

[8] C. W. Nam, "Spatial Augmented Reality Based 3D Tangram System Using Multiple Kinects", Kookmin University Graduate School, **(2012)**.

[9] J. C. Lee and M. S. Kim, "A study on implementation of paper keyboard using Kinect", Korea Computer Information Academy, **(2012)**.

[10] H. J. Yoon, G. I. Kim, J. H. Lee and H. Y. Lee, "Development in physical interactive motion recognition game using Kinect", Information Processing Academy Paper Book, **(2014)**.

[11] J. W. Lee and G. S. Oh, "Yoga learning contents using Kinect", Korea HCI Academy Conference, **(2012)**.

[12] C. Y. Kim, M. R. Kim and J. C. Kim, "Magic Mirror Fashion coordination system using Kinect", Multimedia Academy Paper Book, **(2014)**.

[13] B. S. Kim and E. H. Kim, "Gesture input method variety of users using multiple Kinects", Korea Precision Engineering, **(2014)**.

[14] J. Y. Park, S. M. Lee and M. H. Kim, "Silhouette-based head detection for multiple users position tracking", Korea Information Science Academy Paper Book, **(2005)**.

[15] J. H. Rye, Y. Y. Nam and W. D. Jo, "Human tracking in overlapping situation using multiple cameras", Korea Information Science Academy Paper Book, **(2008)**.

[16] T. Butkiewicz, "Low-cost coastal mapping using Kinect v2 time-of-flight cameras", Oceans-St. John's, IEEE, **(2014)**.

[17] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison and A. Fitzgibbon, KinectFusion: Real-time dense surface mapping and tracking. Mixed and augmented reality (ISMAR), 10th IEEE international symposium, IEEE, **(2011)**.

[18] K.-Y. Yeung, T.-H. Kwok and C. C. L. Wang, "Improved Skeleton Tracking by Duplex Kinects: A Practical Approach for Real-Time Applications", Journal of Computing and Information Science in Engineering, vol. 13, no. 041007, **(2013)**.

[19] M. Caon, *et al.*, "Context-aware 3d Gesture Interaction based on Multiple Kinects", AMBIENT, The First International Conference on Ambient Computing, Applications, Services and Technologies, **(2011)**.

[20] B. Williamson, "Multi-Kinect Tracking for Dismounted Soldier Training", Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, **(2012)**.

[21] "Korea Information Communication Technology Committee Information Communication Terminology Dictionary", Motion recognition, http://word.tta.or.kr/terms/terms.jsp.

[22] http://www.motionring.co.kr/.

[23] https://www.thalmic.com/.

[24] http://opensoundcontrol.org/introduction.