A Study on the control Method of 3-Dimensional Space Application using KINECT System

Jong-wook Kang, Dong-jun Seo, and Dong-seok Jung,

Dept. Electronic Engineering, Inha University, Korea

Summary

This paper suggests how to control 3D application using KINECT system introduced by Microsoft. The conventional system uses a mouse and a touch sensor to control 3D application. Recently, the size of the display became so large that the existing methods are more difficult to control the application. To solve this problem, I propose the method which controls the application intuitively by using the distance information and the joints' location information.

Key words:

KINECT, 3D application, gesture

1. Introduction

Because of the dramatic developments on the graphic technology, the demands on the virtual reality technology and the augment reality technology has been inclined. Consequently, the 3 dimensional applications which implement the 3D graphics are fostered led by the computer game market. Controlling the 3 dimensional technologies has been researched in a various point of views.

There was no trouble using the keyboard and the mouse to control the 3D application in the formal systems because these systems used a small display; however, the recent development of the PC industries led the users to use the bigger display. Therefore, the difficulties of controlling the 3D applications using only a keyboard and a mouse have been a big issue. For the recent study, the gesture recognition became the center of the 3D spatial recognition technology researches.

The conventional gesture recognition technology was mainly focused on the Pointing System which recognizes the shape of the human's hand and controlling the mouse pointer [1]. The Pointing System is handy for the window applications; however, in the perspective of the proposed 3D spatial application, the Pointing System loses its agility. 3D spatial applications almost act as a first person view point. For those first person point of view applications, the most important features for controlling the system is the change on the point of views such as moving forward, backward, left, and right. In order to promote the natural changes on the point of view, rather than using a keyboard and a mouse, we need to implement more intuitive method.

This project will be performed with the KINECT system made by Microsoft Company. The KINECT system can perform better than the ordinary single lens camera systems and the stereo vision system which have the low gesture recognition ability. KINECT system consists of one color image camera, one infrared camera and infrared array transmission. The arithmetic operation is relatively small and recognition rate is extremely high because the KINECT system extracts the distance information by using the infrared sensor transmitter and the infrared camera performing the image processing using color image camera. Furthermore, if SDK (Software Development Kit) provided by Microsoft Company is used, it could get skeleton information that recognizes and chases joint regions of the human body.

This paper proposes the better recognition of gestures for controlling the mouse by identify the body motions such as moving forward, backward, left and right using the KINECT system's skeleton recognition technology.

2. Related Work

2.1 KINECT Sensor

The KINECT sensor made by Microsoft Company is the motion recognition system for Xbox360. The KINECT sensor is used for playing games by recognizing motions of the human. Moreover, Microsoft Company offers SDK (Software Development Kit) for the KINECT sensor and opens the infinite possibilities of the KINECT sensor. It advocates developments in speed and accuracy.

Fig.1 KINECT Sensor of Microsoft.

Manuscript received September 5, 2011
Manuscript revised September 20, 2011

Figure1 is the KINECT sensor. The KINECT sensor consists of a motor for the infrared light sensors, a camera part which takes color image as an input and a stereo microphone for the speech recognition. As an infrared transmission module from infrared sensor parts shoots the object in a cross stripes pattern, it gathers the distance information from a staff to measure a distortion degree by receiving infrared light with an infrared camera.

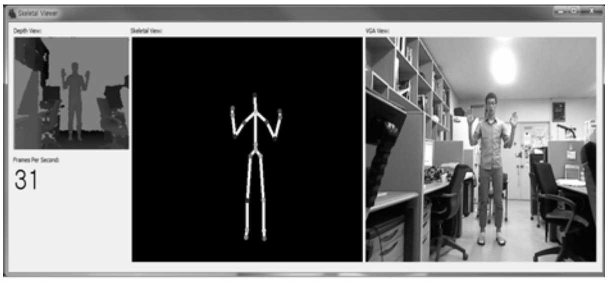

Fig. 2 KINECT Example program

From left to right, the figure 2 shows the Depth Map screen by measuring the distance information, a skeleton information screen extracted by recognizing the shape of the man, and a screen through the color image camera

Since the distance information map is processed only with an infrared ray's sensor, it is dramatically less time consuming compare to the older methods such as a stereo vision technique or the multi-image system.

The leftmost figure is an image which is taken by the infrared ray's sensor which varies the color based on the distance information of objects.

When you see the figure in the middle, after recognizing the shape of the man using distance information Map and applying these to the skeleton algorithm, the system shows the skeleton information of the human body extracted from the left most figures.

By utilizing extracted information, it would be possible to recognize the gesture of the man highly accurately with small arithmetic operation volume.

3. Proposed Algorithm

This paper suggests how to determine the five types of gestures which will be proposed by extracting information from the each joint point.

3.1 A forward movement and backward movement

The figure 3 is gesture considered as a forward movement. It is an ordinary posture which steps the right foot that in the front. If the differences between the z axis value of the right knee joint point and the left knee joint point exceed the certain threshold, this posture is considered as moving forward.

$$z_{kneeright} - z_{kneeleft} \geq Th_{knee} \qquad (1)$$

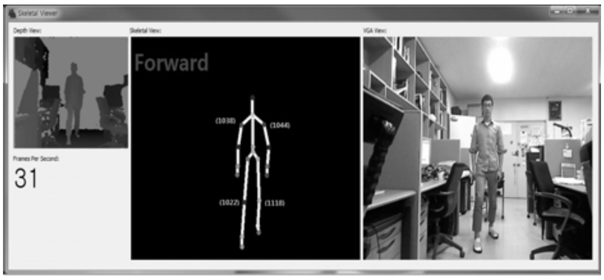

Fig. 3 Forward gesture

Contrary to the forward movement, it is considered as moving backward, as shown in the figure 4, if the object places the left foot forward and the satisfies the following formula.

$$z_{kneeleft} - z_{kneeright} \geq Th_{knee} \qquad (2)$$

Th_{knee} is decided through several experiments with the mathematical formula (1), (2).

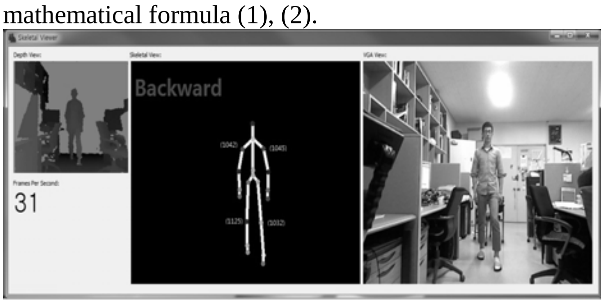

Fig. 4 Backward gesture

3.2 Turning left and Turning right

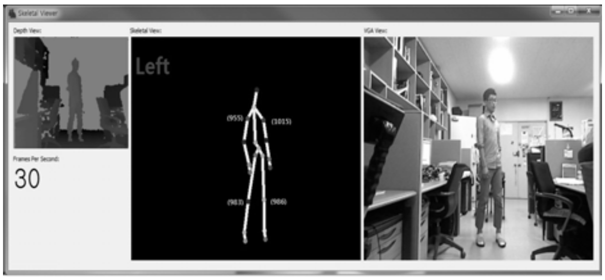

Fig. 5 Turning left gesture

The figure 5 and figure 6 are gestures representing turning left and turning right. Based on changes the changes in the point of view from left to right, it conducts the ability of changing directions in 3D space. These are the gestures when the subject changes its direction of the movement by shifting the shoulder to the corresponding direction.

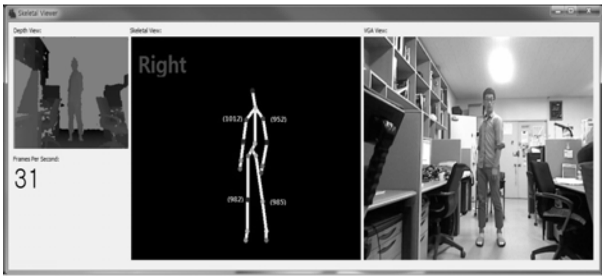

Fig. 6 turning right gesture

The standard for the judgment is the difference between z axis val ue of the right and left shoulder joint point.

If difference for each z-axis value of joint point exceeds the thresh old, we judge that the gesture means turning left or right.

$$z_{shoulderright} - z_{shoulderleft} \geq Th_{shoulder} \quad (3)$$

$$z_{shoulderleft} - z_{shoulderright} \geq Th_{shoulder} \quad (4)$$

The value for $Th_{shoulder}$ for the mathematical formula (3), (4) is determined by the several experiments.

3.3 Mouse Pointer Control

The mouse pointer control is represented by the movement of the hands explicit shapes. Sometimes, in order to control the 3D application, pointing motion would be needed. Therefore this functionality is suggested. The gestures are divided into three parts. There is a standby status to lose the control over the pointer, a movement status to move the mouse pointer, and a clicking status to click the mouse pointer.

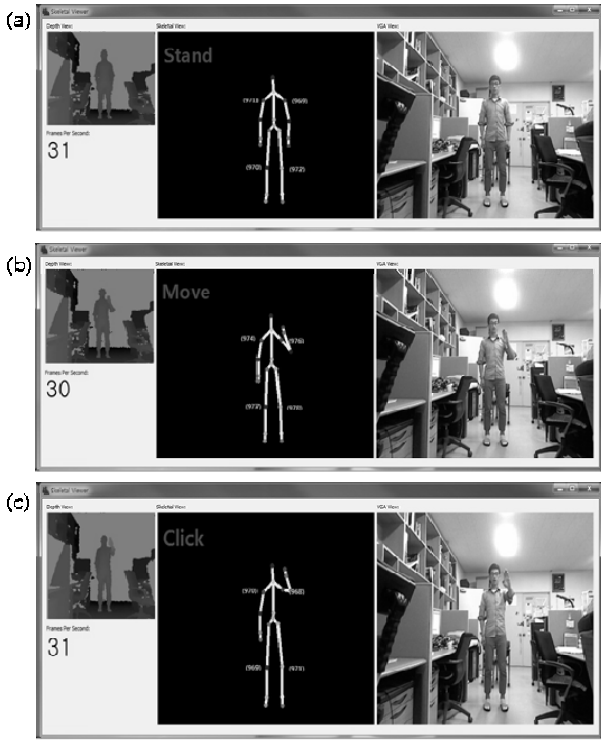

Fig. 7Turning right gesture (a) Standby gesture (b) Mouse moving gesture (c) Click gesture

The standby status is the posture the both hand is placed under the waist as shown in the figure 7(a). The point of the determination i s that y axis value of the wrist joint point is lower than y axis valu e of the lower back joint point.

It is considered as a mouse movement if as the wrist joint point is higher than the lower back joint point, the movement is made from left to right. Regardless to the left or right hand, any one hand raised takes the authority over the mouse control. In order to tell the difference between the clicking statuses, the mouse pointer move when the difference between the wrist joint point's z-axis value and corresponding hand's shoulder joint point's z-axis value is under certain threshold.

$$| z_{shoulder} - z_{handpoint} | \leq Th_{click} \quad (5)$$

For the clicking status, it is considered as a gesture for clicking when the difference between the wrist join point's z-axis value and corresponding shoulder joint point's z-axis value is under certain threshold while moving status.

$$| z_{shoulder} - z_{handpoint} | \geq Th_{click} \quad (6)$$

Th_{click} In the mathematical formula (5), (6) is determined by experiments.

4. Experiment and Result

This paper composed following two experiments to measure efficiency test about the gesture.

Fig. 8 Experiment Environment

The first experiment is composed with 10 students in 20's. For the testing background, one of commonly used 3D space application, DAUM map road view, is used. The time spent for moving 1 km from the certain starting point to the certain end point is measured. All tests were composed with an identical computer and following is the result.

Table 1: Time Measurement

	Mean Time	Improvement rate (%)
Mouse system	262s	100%
Proposed gesture	205s	127%

As it can be seen in Table 1, the amount of time spent to reach the destination by using the proposed gesture control system takes 27% less time compare to the conventional mouse control systems. Under consideration of the loading time for the web application and the delay phenomenon of web speed, it is easily recognized that significant amount of time is saved.

For the second experiment, the test for the gesture recognition has been performed. For this experiment, 10 test subjects pose each gestures 10 times. Therefore total of 400 test cases have been gathered and the result came out to be following.

According to the Table 2, among 400 trials, 98.75% trials were re cognized successfully.

Table 2: Recognition rate

	Number of success / Number of total test	Recognition rate (%)
Proposed system	395/400	98.75%

5. Conclusion

We have developed a gesture in order to control the 3-Dimensional Space Application. Unlike existing systems, KINECT system which can recognize human using infrared sensor and track each joint position is used for the gesture detection. As a result, we got a high recognition rate of 98.75 percent.

This result means that proposed system using KINECT can provide robust recognition rate using simpler algorithm than existing system..

In addition to the recognition rate, the 3-dimensional application control time using proposed gestures is 27 percent faster than the existing system using mouse. As a result, proposed gesture is more intuitive and effective than the method using mouse click.

References

[1] B. Min, H. Yoon, J. Soh, Y. Yang and T. Ejima, "Hand Gesture Recognition Using Hidden Markov Medels, " IEEE International Conference on Systems, Man, and Cybernetics, pp.4232-4235, 1997.

[2] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, " Proc. of the IEEE, vol. 77, no. 2, pp 267-296, 1989.

[3] Yamamoto, S. Mae, Y. Shirai, Y. Miura, J., "Realimte multiple object tracking based on optical flows ". Proc. IEEE Intl. Conf. on Robotics and Automation, vol.3, pp.2238-2333, 1995.

[4] P. Chen, C. Grecos, "A Fast Skin Region Detector, " ESC Division Research 2005, pp.35-38, Jan. 2005.

[5] Rubine, D., "Specifying Gestures by Example." ACM Computer Graphics, Vol. 25, No. 4, pp. 329-337, July 1991.

Jong-wook Kang received the B.S. degree in Electrical Engineering from Inha University in 2010. M.S. candidate in information engineering from Inha University, Korea, from 2010 to now.

Dong-jun Seo received the B.S. degree in Electrical Engineering from Inha University in 2010. M.S. candidate in information engineering from Inha University, Korea, from 2010 to now.

Dong-Seok Jeong became a Member (M) of IEEE in 1983, a Senior Member (SM) in 2000. He got his BSEE degree from Seoul National University in 1977 and MSEE and Ph.D. degree from Virginia Tech in 1985 and 1988 respectively. He is also the member of SPIE and HKN. From 1977 to 1982, he was a researcher at Agency for Defense Development of Korea and he is currently a professor at Inha University in Korea since 1988. Also he served as the president for the Institute of Information and Electronics Research from 2000 to 2004. His research interests include image and video processing, image and video signature and forensic watermarking.