Proceedings of the International MultiConference of Engineers and Computer Scientists 2013 Vol I,
IMECS 2013, March 13 - 15, 2013, Hong Kong

Development of a Camera Control System Using Human Gestures Recognition

Youji Ochi, and Yuya Takeda

***Abstract*— Recently, a distance learning using a real-time video streaming is more popular. A communication gap between the camera staff and the teacher makes the camera control more difficult. We focus on Microsoft Kinect sensor and develop the camera-control system using the gesture. Our system uses the infrared ray information of Kinect to control a camera in according to teacher's intention. In this paper, we describe the camera control framework, the system implementation and the trial use.**

***Index Terms*— Camera Control, Image Processing, Motion Analysis, Kinect,**

I. Introduction

In recent years, due to the spread of broadband networks, distance learning courses using a real-time video streaming is more popular. Distance learning can be considered two types. One type shoots the whole a screen, a blackboard and a teacher with a fixed camera angle. Another type shoots the lecture image to follow the teacher's movement or instructions with an active the camera angle. Our study assumes distance learning in the latter. We tried a distance-learning course [1]. However, in the case, we faced the difficulty of adjusting the teacher's intention. Therefore we focus on an approach to deliver the lecture video automatically in accordance with the intent of the teacher. Then we have developed a camera control system that uses the pointing stick as previous study [2]. This system controls the camera using the color information recognition of the pointing stick. However, there are some difficulty factors to a precisely camera-control Ambient light often disturbs the recognition of the pointing stick. In this study, we focus on infrared ray information and adopt the Kinect sensor (hereinafter, Kinect). Kinect is a game device for the Xbox360 that was released by Microsoft in November 2010. We can get three-dimensional coordinate information of the body movement using Kinect. Kinect is examined in terms of its affordances of technical interactivity, which is an important aspect of pedagogical interactivity [3]. Therefore we develop the camera-control system using the gesture. In this paper, we describe the camera control framework, the system implementation and the trial use.

Manuscript received Jan 8, 2013; revised Jan 30, 2013. This work was supported by Grant-in-Aid for Scientific Research(C) 23501122 from the Ministry of Education, Culture, Sports, Science & Technology in Japan.

Youji Ochi is with Faculty of Science and Engineering, Kinki University, Higashi-Osaka, 5778502 Japan (e-mail: ochi@ ele.kindai.ac.jp).

Yuya Takeda is with Graduate School of Science and Engineering Research, Kinki University, Higashi-Osaka, 5778502 Japan.

II. Our issues and approach of camera control

A. Issues of Camera Control

Our research assumes a real-time distance-learning environment where the teacher uses multiple media such as a blackboard, whiteboard and projection screen at the same time. The objective of our study is a controlling the network camera with pan and tilt, and zoom in/out function for distance learning. The camera control means that the change of focus to the specified coordinates and performs a zoom in and out. General video camera requires a camera staff to change the angle according to a teacher's position. The incomplete communication between the camera staff and the teacher makes difficult to control the camera correctly. However, the camera control reflected the teacher's intent request an operation instruction to the camera staff. The instruction often disturbs the lecture process.

B. Our Approach

We have defined for the following items as required specifications.

1. Teacher can control the camera angle by her/himself
2. The system does not require a specific device to a teacher

There some researches that are automatic camera control [4][5][6]. Richard [7] introduces a set of natural camera control and multimedia synchronization schemes based on the individual object interaction. On the other hand, there are some researches about hand gesture recognition and finger detection for sign language [8][9]. We propose the camera control approach in order to control camera angle using image-processing technology. Specifically, we adopt the following approach.

1) Camera control using teacher's gestures

Our study makes use of the gesture, which is determined by the teacher's body movements in order to control the camera angle without camera staff. We defined the gestures based on the human skeletal information that are the coordinate, displacement, angle and tilt.

Our system recognizes the operation for camera angle control by combining the gesture.

2) Prepare camera for gesture recognition

Generally, a teacher always moves on the podium, because teaching materials are not always beside the teacher. The camera must be able to shoot the whole of the podium in order to recognize her/his gesture. Therefore we prepare a camera for gesture recognition in addition of the camera for shoot the lecture.

3) Two operation modes

We have prepared two kinds of operation modes which are "point mode" and "angle mode" to send gesture commands to the network camera from the teacher.

The former mode is for changing the camera angle to the coordinates teacher points out. When the teacher points out a location on the blackboard, the system changes the angle of the camera system to the point. Teacher can send the important point of the teaching material through the camera using this function.

The latter is a mode for changing the camera angle relatively. The teacher can change the camera angle to the location that s/he can not point out directly such as her/his inaccessible location by this mode.

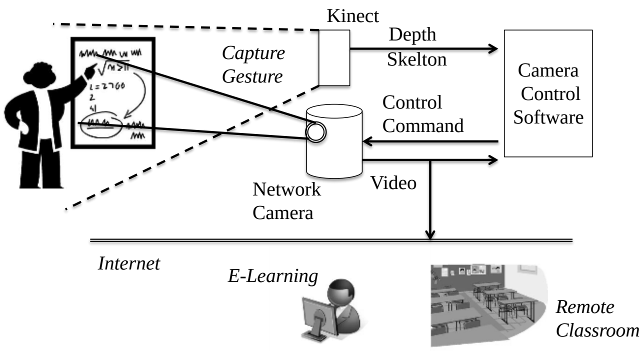

Fig. 1. System configuration

III. System Configuration

Figure 1 shows the system configuration. This system consists of a network camera, camera control software and Kinect.

1) Network Camera

We adopt the DG-SC385 (Panasonic, Figure 2) in order to shoot the lecture video. This camera has HTTP Request APIs in order to control the angle (pan, tilt, zoom) and various setting. Network camera has a camera control interface through HTTP in order to control the pan, tilt, and zoom. The camera sends the video image by motion jpeg or mpeg4. We adopt motion jpeg format in order to receive the image.

Fig.2. Network camera

2) Microsoft Kinect

Kinect is for recognizing the teacher's body motion. Kinect is set up in the panoptic location where it can shoot the teacher's movement and teaching material (screen and black board etc.). It sends the depth and skeleton information to the camera control software. The system receives the image information present in the human body from the Kinect and determines the camera control command from teacher's gesture.

3) Camera Control Software

This software is our original software to determine the physical behavior pattern based on the information from the Kinect. It sends camera control commands to the network camera through HTTP requests. We implement it using Microsoft Visual C# and Open CV Library [10].

IV. Principle of Human Gesture Recognition

A. Definition of movement pattern

We define the four movement patterns using coordinates pointed of teacher's hand. The direction patters are up, down, right and left. Our system calculates the delta-degree from inter-frame deference and determines the direction pattern according to Table 1.

TABLE I
Delta Degree for Direction Recognition

Direction	Delta-degree
Up	-45 - 45
Down	45 - 135
Right	135 -225
Left	225 -315

TABLE 2
Gesture Pattern for Camera Control

Camera Control	Gesture Pattern
Up	Raise one's right arm to shoulder level, and bend the arm vertically
Down	Raise one's left arm to shoulder level, and bend the arm vertically
Right	Extend the right arm horizontally in a sideling direction
Left	Extend the right arm horizontally in a sideling direction
Zoom In	Rotate ones' hand clockwise
Zoom Out	Rotate one's hand counterclockwise
Default position	Extend one's both arm horizontally in a sideling direction, and bend the arms vertically.

B. Definition of gesture patterns

1) Switching the operation mode

We define the gesture for switching point mode and angle mode. It must be a conscious gesture, and low false-recognition rate. Specifically, we define the gesture of the switching mode bending and stretching motion of the arm using the coordinate information of depth from Kinect.

2) Angle mode

Table 2 shows the definition of the gesture patterns in the angle mode. Angle mode is for changing any direction (pan, tilt) angle of the camera according to the teacher's intention. Determination conditions are as follows.

Cond.1) When you meet the conditions of 90% or more to the pattern instruction in 30 frames, the system counts the pattern.

Cond.2) When the same instruction pattern appeared three times in a row, the system sends the operation command to the network camera.

Cond.3) If the above conditions are not met, the system resets the count.

Determination of zoom operation is using rotation movement of one's hand. The direction of rotation is determined by the direction pattern. For example, clockwise occur in order of Down, Left, Up and Right.

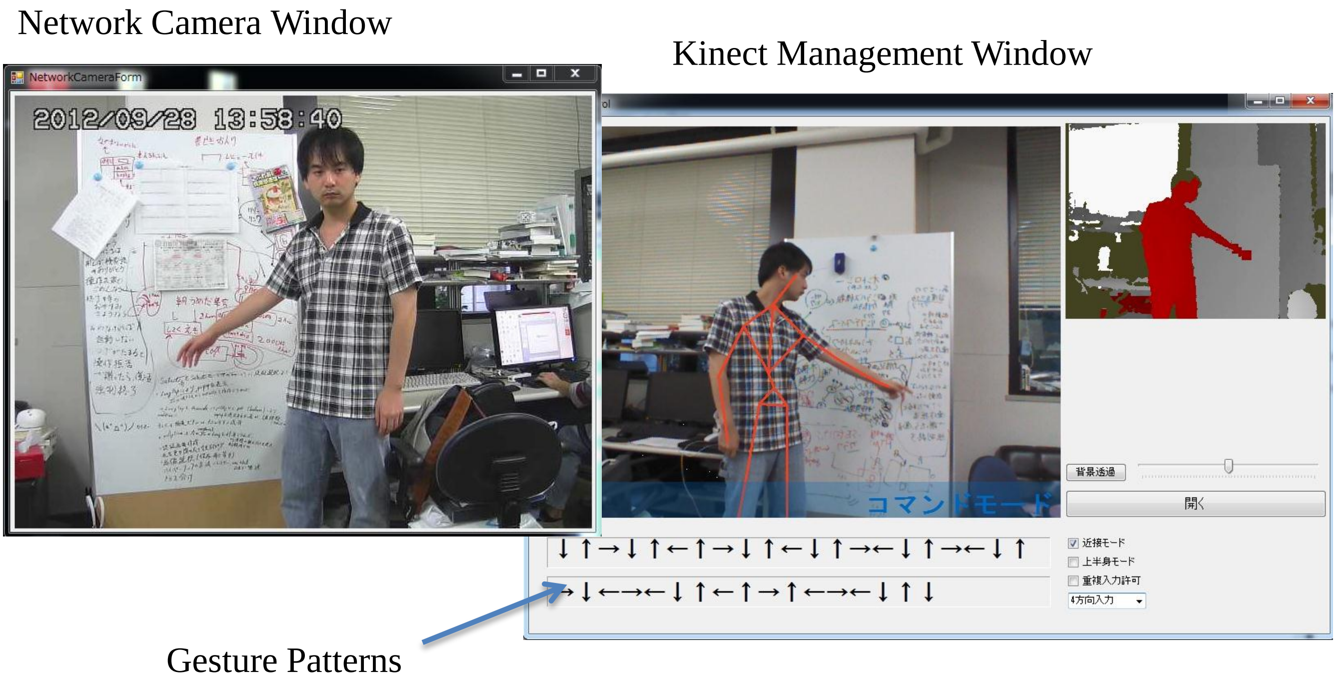

Fig.4. User interface

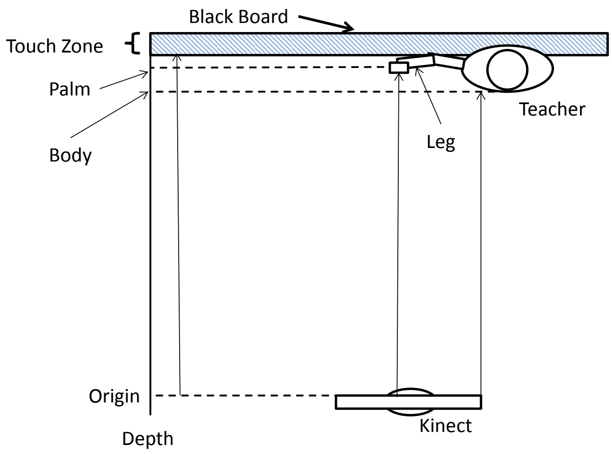

Fig.3. Principle of judging the touch gesture

Our system changes the angle by one step per one command and not receives the other command until the operation is complete in order to prevent malfunction.

3) Point mode

This mode changes the camera angle to the location where the teacher points out. We develop a touch gesture use the depth information from Kinect in order to specify the location according to the teacher's intention.

Figure 3 show the principle of judging the touch gesture. At first, our system records the depth value of black board or background. We define the area within a few centimeters of the black board as Touch Zone. The width of the Touch Zone is variable. When the teacher`s wrist enters and keeps within the touch zone, the system judges the movement as touch gesture. Then the system picks up the coordinates of teacher's wrist and changes the angle to there. The way of zooming is the same as the angle mode.

V. User Interface

We prepare two user interfaces for our system management. They are Kinect manager and network-camera controller.

A. Kinect management Window

This interface displays the RGB color image, skeleton information and depth information from the Kinect. Then the user can set the recognition mode that is whole body mode and upper body mode. The latter mode is designed if the teacher's lower body disappeared by the teacher's desk.

B. Network Camera Window

This interface displays a video image that be delivered to the remote classroom. It gets the video image (motion jpeg format) via HTTP. These interfaces are used for the initial configuration of the system.

VI. Trial Use and Evaluation

A. Trial Use on Campus Open Day

We have demonstrated the system on our campus open day. This event is for high school students to visit a university. 70 participants came to our demonstration for five hours. We explained the basic operation of each mode and demonstrated our system. Then 15 participants of them tried to use our system.

In the demonstration and trial use, the participants could operate the camera by gesture. Interior lighting and sunlight did not affect the operation. Moreover, there were no reductions of the response to the operations. However, some user caused the malfunction of the coordinate detection. We think their clothing that hides the outline of their body caused the incorrect coordinate

B. Evaluation of the touch gesture recognition

We evaluated the accuracy of the touch gesture recognition. In this evaluation, we prepared a white board and tried touch gesture action 10times in 3positions. These positions are 2m, 2.5m, 3m, 3.5m and 4m away from Kinect. Then we assessed the success rate of touch gesture. In this evaluation, we set the width of touch zone 10cm.

Figure 5 shows the success rate of touch gesture. The result means that the distance does not have much influence on the success rate. Our system could not run in the case of over 4m, because Kinect could not recognize human body more than 4m away. Also our system could not judge the touch gesture, when the Kinect failed to recognize the human skeleton.

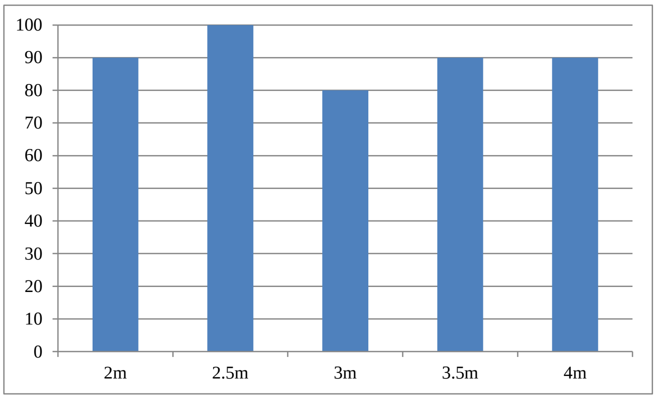

Fig. 5. Success rate of touch gesture

VII. Conclusion

We focused on skeletal coordinate and depth information of human body from Kinect. We have proposed the camera control system using teacher's gesture. Recognition of her/his gesture using the coordination and depth information of Kinect is useful as a control device of the camera. We demonstrated our system on campus open day and confirmed the utility.

However, we can expect a more extensive range of camera control to increase the recognition pattern of behavior. As future research, we are considering a combination of voice-controlled in order to achieve advanced control of the network and improve the accuracy of camera motion analysis.

References

[1] Youji Ochi , Yuh Etoh, Nobukazu Iguchi,Shouji Mizobuchi, Yoshiaki Shiraishi,Masatoshi Mori, Byon-Chol So, Yoshinobu Kurose, Itaru Sano1, Sonoyo Mukai, A Practice of Interactive English Class On a Multipoint/Interactive Remote Lecture Environment,8th International Conference on Information Technology Based Higher Education and Training, pp.294-297,2007.

[2] Youji Ochi, Nobukazu Iguchi, Development of a Camera Control System for Lecture Recording Using Pointing Stick, International Multiconference of Engineering and Computer Science、pp.924-927, 2012.

[3] Hui-mei Justina Hsu, The Potential of Kinect as Interactive Educational Technology, 2nd International Conference on Education and Management Technology,pp.334-338,2011

[4] Shimada Atsushi, Suganuma Akira, Taniguchi Rin-ichiro, Automatic Camera Control System for a Distant Lecture Based on Estimation of Teacher's Behavior, Proceedings of the Seventh IASTED International Conference on Computers and Advanced Technology in Education, pp.106-111, 2004.

[5] Yong Rui, Anoop Gupta, Jonathan Grudin and Liwei He, Automating lecture capture and broadcast: technology and videography, MULTIMEDIA SYSTEMS,Volume 10, Number 1, 3-15, 2004

[6] Nishigori, S. and Suganuma, A., Automatic Camera Control System for a Distant Lecture with Videoing a Normal Classroom. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002, pp. 1892-1897, 2002.

[7] Richard Y. D. Xu, Jesse S. Jin, Camera Control and Multimedia Interaction using Individual Object Recognition, pp.77-85, JOURNAL OF MULTIMEDIA, VOL. 2, NO. 3, JUNE 2007

[8] Prateem Chakraborty, Prashant Sarawgi, Ankit Mehrotra, Gaurav Agarwal, Ratika Pradhan, Hand Gesture Recognition: A Comparative Study, Proceedings of the International MultiConference of Engineers and Computer Scientists 2008 Vol.1, pp388-393, March, 2008

[9] Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari, Finger Detection for Sign Language Recognition, Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I, pp489-493, March 2009

[10] Open Computer Vision Library, http://sourceforge.net/projects/opencvlibrary/