Motion Templates based User Interface for Immersive Virtual Reality Environment

Hasup Lee, Yoshisuke Tateyama and Tetsuro Ogi

Graduate School of System Design and Management, Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8526 Japan
{hasups, tateyama, ogi}@sdm.keio.ac.jp

Abstract

In this paper, an unaffected interface for an immersive virtual reality environment using the motion templates method was proposed. We made background model with real photographs. The panoramic images were constructed using these photographs from real environment and texture-mapped to virtual sphere surround CAVE. As a user interface for this background model, the motion templates method of computer vision technologies is used. The computer vision module recognizes predefined user's gestures and sends commands to render module of CAVE system via internet using UDP protocols. Using this method, the users can manipulate the background model unaffectedly. This is efficient for VR applications like looking around a remote scene.

Keywords: *virtual reality, computer vision, user interface, CAVE, panorama*

1. Introduction

How to make the users more realistic is one of main problems in virtual reality (VR) research field. CAVE is one of the devices for this problem and its approach is visual. It contains several projectors and screens and covers user's view using them. The users can feel immersive by surrounding with VR contents. VR contents are not limited to synthesize graphic scenes but also be the real images from real world. For example, the images of other famous sites, deep seas, Moon and Mars can be also VR contents. Using real image as background of VR contents, users can feel them exist at other places.

In this paper, an unaffected interface for CAVE using the motion templates method [1] was developed. We applied this user interface to our real image-based background model for CAVE. As a user interface for this background model, the motion templates method of computer vision technologies is used. The computer vision module recognizes predefined user's gestures and sends commands to master/render modules of CAVE system via internet using UDP protocols. Using this method, the users can manipulate the background model unaffectedly.

There have been researches about real image-based background modeling of CAVE. One mono panoramic image is constructed using a digital camera and panoramic tripod head in [2]. This work extended to the stereo background modeling made by two panoramic images for left and right eye in [3]. The more rapid method for take picture of environment is improved with 3D panorama sweep function in [4]. The CAVE system with real image-based background modeling is shown in (Figure 1). A game pad or other wired devices are used for manipulate the background like (Figure 1). But they bother users to feel immersion. So, more unaffected interfaces are needed. The objects and background need to be moved, rotated, zoomed by these interface.

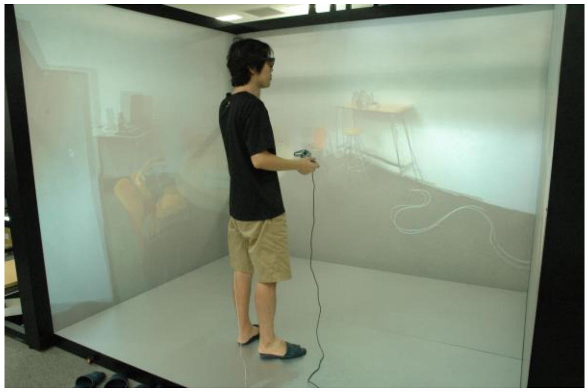

Figure 1. Real Image-based Background Modeling

As computer vision based interface for CAVE environment, the hand detection method was studied in [5]. Two camera attached in the corner was used and the direction of 3D pointing was detected. The direction was calculated using the user's head and kinematic constraints of a pointing gesture. The real and virtual objects were synchronized in [6]. They used the tangible gadget that vision marker attached and tracked it using computer vision technology. But the gadget bothers the users. Blob detection method was developed in [7]. This approach detected blobs of head and hands and detected gestures using these blobs' relative position and velocities. One built-in webcam of notebook PC was used for gesture recognition.

2. User Interface

We developed computer vision-based user interface using motion template method. The motion templates method was described in [1]. This method calculates whole gradients of motion history images which be accumulated through several time slices. Using this method, the user's whole body gesture can be recognized. This can be processed in real-time because it doesn't use a complex algorithm or operator but simple gradient. It is useful to analysis whole direction of body movement. This is also useful for arm or body gesture recognition under CAVE system because it is relatively robust to light change.

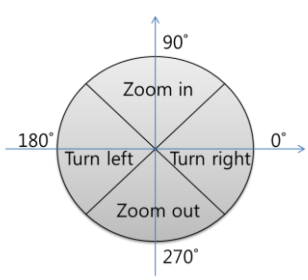

Figure 1. Gesture Commands

Table 1. Gesture Definition

Gesture command	Motion direction (degree d)
Turn left	$180\text{-}e < d < 180+e$
Turn right	$0 < d < e$ or $360\text{-}e < d < 360$
Zoom in	$90\text{-}e < d < 90+e$
Zoom out	$270\text{-}e < d < 270+e$

Our interface was designed for navigation or manipulation of background model, not objects. So, we defined several commands - turn left, turn right, zoom in and zoom out and gestures for these commands for this purpose. The definition of gestures is shown in (Table 1) and its coordination and range are shown in (Figure 2). In (Table 1), d is degree of gradient direction and *e* is tolerance. A vision module recognizes user's gesture and sends commands to renderer modules via internet. We use UDP protocols because we don't need buffering of commands and little loss of commands is acceptable.

The purpose of our interface is not to detect exact pointing or distance of movement. It detects the tendency of the direction of whole body movement. For some applications like object manipulations, our interface is not useful but we focus user's unaffected movement. For example, if a user wants to move left in background contents, the user can by move hand or arm or shoulder or head or any combination of them. This is efficient for VR applications like looking around a remote scene.

3. Experiment

Our system was developed on our K-CAVE system [8] which has 4 screens, 8 stereo projectors and magnetic positioning sensor. It contains 1 master and 4 renderer Linux machines for each 2 stereo projectors. Our user interface by motion template based gesture recognition is shown in (Figure 3). The user puts on a stereo glasses and the position sensor with code is attached on it. Linux installed notebook PC and built-in webcam was used for input and processing vision interface module. Frame images were captured from built-in USB webcam. Vision module was developed using OpenCV library. The resolution of camera input is '640x480'.

There are screenshots of motion history images in (Figure 4). The white line in the circle indicates the gradient of whole movements and its direction. The vision module recognizes user's gesture and interprets into defined commands. These commands are applied to background of K-CAVE system (Figure 3). Only text commands are sent to master/renderer modules via wireless network using UDP protocols.

Figure 2. Configuration of K-CAVE System

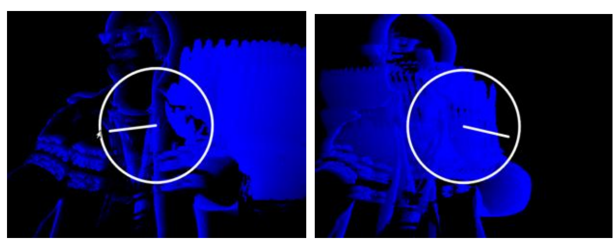

Figure 3. Left and Right Command Displayed on Motion History Images

4. Conclusion

We made a real image-based background modeling for CAVE system. It is developed for user to feel more immersion. And we add an affected interface to this. To make such interface, a device held or attached to body like a game pad, glove and hat is excluded. Because of seamless feature of screens of CAVE, standalone system like wireless network connected notebook, tablet pc is preferred for interface processing. As conclusion, we developed an interface of gesture recognition using motion templates method. It is one of computer vision algorithm robust to light change. Our approach can be applied to interface using whole body motion.

Acknowledgements

This work was supported by G-COE (Center of Education and Research of Symbiotic, Safe and Secure System Design) program at Keio University.

References

[1] G. Bradski and J. Davis, "Motion Segmentation and Pose Recognition with Motion History Gradients", International Journal of Machine Vision and Applications, vol. 13, no. 3, **(2002)**, pp. 174-184.
[2] H. Lee, Y. Tateyama and T. Ogi, "Realistic Visual Environment for Immersive Projection Display System", The 16th International Conference on Virtual Systems and Multimedia, **(2010)** October, pp. 128-132.
[3] H. Lee, Y. Tateyama and T. Ogi, "Panoramic Stereo Representation for Immersive Projection Display System", The 9th International Conference on VRCAI (VR Continuum and Its Applications in Industry), **(2010)** December, pp. 379-382.
[4] H. Lee, Y. Tateyama and T. Ogi, "Image-based Stereo Background Modeling for CAVE System", International Symposium on VR innovation (ISVRI) 2011, **(2011)** March, pp. 251-254.
[5] J. Kim, D. Gracanin, H. L. Singh, K. Matkovic and J. Juric, "A Tangible User Interface System for CAVE Applications", Virtual Reality Conference 2006, **(2006)** March, pp. 261-264.
[6] T. Moeslund, M. Storring and E. Granum, "Vision-Based User Interface for Interacting with a Virtual Environment", Proc. DANKOMB 2000, **(2000)** August.
[7] H. Lee, Y. Tateyama and T. Ogi, "Hand Gesture Recognition using Blob Detection for Immersive Projection Display System", ICECECE 2012 : International Conference on Electrical, Computer, Electronics and Communication Engineering, WASET, vol. 62, **(2012)** February, pp. 180-183.
[8] Y. Tateyama, S. Oonuki, S. Sato and T. Ogi, "K-Cave demonstration: Seismic information visualization system using the OpenCABIN library", Proceedings of the ICAT 2008, **(2008)**, pp. 363-364.