Stereo-Vision-Based Human-Computer Interaction with Tactile Stimulation

Ho-joong Yong, Jongwon Back, and Tae-Jeong Jang

If a virtual object in a virtual environment represented by a stereo vision system could be touched by a user with some tactile feeling on his/her fingertip, the sense of reality would be heightened. To create a visual impression as if the user were directly pointing to a desired point on a virtual object with his/her own finger, we need to align virtual space coordinates and physical space coordinates. Also, if there is no tactile feeling when the user touches a virtual object, the virtual object would seem to be a ghost. Therefore, a haptic interface device is required to give some tactile sensation to the user. We have constructed such a human-computer interaction system in the form of a simple virtual reality game using a stereo vision system, a vibro-tactile device module, and two position/ orientation sensors.

Keywords: Human-computer interaction, haptic interface, tactile stimulation, stereo display, vibro-tactile device.

Manuscript received Oct. 16, 2006; revised Mar. 28, 2007.

Ho-joong Yong (phone: + 82 33 257 3657, email: tt95k@hotmail.com), Jongwon Back (email: harukib@mail.kangwon.ac.kr), and Tae-Jeong Jang (phone: + 82 33 250 6348, email: jangtj@kangwon.ac.kr) are with Department of Electronic and Telecommunication Engineering, Kangwon National University, Gangwon-do, Korea.

I. Introduction

A virtual environment can be represented more realistically by using a stereo vision system. However, if a user wants to input commands by using two dimensional input devices in this virtual space, the user will be inconvenienced because the virtual space is three dimensional. To diminish this inconvenience, the user may also have to use three dimensional input devices. However, the user should also be well trained to input the right commands. One of the easiest methods of inputting commands in virtual space is direct pointing using the fingertip.

In HoloSketch [1] and Swordplay [2], the user feels as if a virtual object is directly pointed out by his/her own fingertip using a hand-held six-axis mouse/wand and 6-DOF tracker. But, if some tactile devices in addition to visual displays are used to give some tactile feeling on the user's fingertip when certain points on a virtual object are touched, the sense of reality could be greatly heightened. To realize this, we need to align the virtual and physical space coordinates and perform the appropriate view control. Also, some tactile devices should be used to deliver appropriate tactile stimulation to the user.

Systems which map virtual and physical locations using head-tracked stereo-rendering have been continually researched [1], [3]-[6]. To track the positions of the user's eyes and a fingertip we use two position/orientation magnetic sensors; and we control the view volume according to eye position in real-time.

In haptic interfaces, there are force feedback devices such as PHANToM from Sensible Technologies Inc. and tactile devices such as vibrator motors which only give tactile sensations. While force feedback devices are superior to tactile devices in the sense of reality, they are bulky and their power consumption is high. On the other hand, tactile devices usually are small and consume less power than force feedback devices. Our research team has already developed a vibro-tactile device module which gives tactile sensation to the user by generating some vibration on the user's skin.

Recently, a lot of research has been done to use existing vibro-tactile devices for tactile feedback in virtual reality systems to increase the sense of reality [7]-[9]. Scheibe and others [7] proposed a new tactile feedback system, which uses shape memory alloy (SMA) wires, for finger-based interactions in immersive virtual reality applications. Roegenbrecht and others [8] used several existing vibro-tactile actuators like electromagnetic alarm buzzers, vibro-motors, and piezoelectric bending actuators in combination with magnetic and optical tracking devices to construct prototype systems, namely, TactilePointer and TACTool. They tried to enhance the user's interaction capability with virtual or augmented environments by offering additional three-dimensional tactile feedback with their systems. Ryu and Kim [9] used vibro-motors as vibro-tactile displays to enhance the sense of presence by simulating collisions between the user and the virtual environment.

Since most of the existing vibro-tactile devices used in the above research have their own uses, that is, they originally were designed for their own purposes, such as cell-phone vibro-motors, they may not be suitable for vibro-tactile displays for other purposes. Our vibro-tactile device has been specially designed for tactile display in wearable environments, and it has some advantages in size and power consumption compared with other existing vibro-tactile devices. We used this device to construct a realistic human-computer interaction system in the form of a virtual bubble popping game to provide tactile stimulation to the user.

Virtual reality is a whole collection of results of the interaction between humans and computers in which, through the use of computer applications, the user's senses are tricked into perceiving as if nonexistent environments and circumstances are set in the real world. The game industry is a good experimental arena for the application of virtual reality. We have developed a stereo-vision-based human-computer interaction system as a form of a virtual reality game using two position and orientation sensors and a vibro-tactile device. This demonstration game has a simple form. Some bubbles seem to come out and move slowly toward the user and he/she can pop those bubbles. The bubbles are automatically generated and move randomly. Because the bubbles are represented as stereoscopic images, they seem to come out from the monitor screen. This is achieved by controlling the parameters of the stereo vision system. When the user's fingertip touches the surface of a bubble, the bubble pops and the vibro-tactile stimulation is instantly delivered to the fingertip with a graphic display and sound.

The structure of the human-computer interaction system is described in section II including the new vibro-tactile device module. The bubble popping game, coordinate alignment, and view control are described in section III. Conclusions are given in section IV.

II. Structure of the Human-Computer Interaction System

As previously stated, we use a vibro-tactile device and two magnetic-type position/orientation sensors. We also use stereo shutter glasses and a high speed CRT monitor for stereo vision. The 120 Hz field rate of this monitor is enough for the user to perceive in stereo vision without flickering. Figure 1 shows the structure of a human-computer interaction system with tactile stimulation.

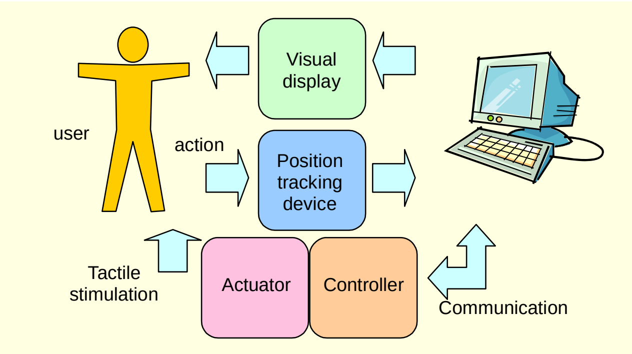

Fig. 1. Structure of a human-computer interaction system with tactile stimulation.

1. Vibro-Tactile Device Module

Various types of tactile devices have been developed: electro-tactile, vibro-tactile, pneumatic-tactile, and so on. The vibro-tactile device used in our proposed system was designed and fabricated by our research team. The size of the pin-type vibro-tactile module is 5 mm×5 mm×5 mm as shown in Fig. 2. The module weighs 0.32 g. It is typically actuated by 3.3 to 5 V with power consumption lower than 0.4 W per module. We combined the vibro-tactile device module with a position/orientation sensor, as shown in Fig. 2, so that a fingertip can be inserted into the device.

Previously, there have been several similar combinations of position tracking devices and vibro-tactile devices for fingertips or hand use. CyberGlove™ combined with CyberTouch™ [10] might be the earliest example. In that system, several vibration motors were used as vibro-tactile devices. The system proposed by Scheibe and others [7] is quite similar to our system. They made their own vibro-tactile devices to be worn on the fingertips and combined them with a commercial finger

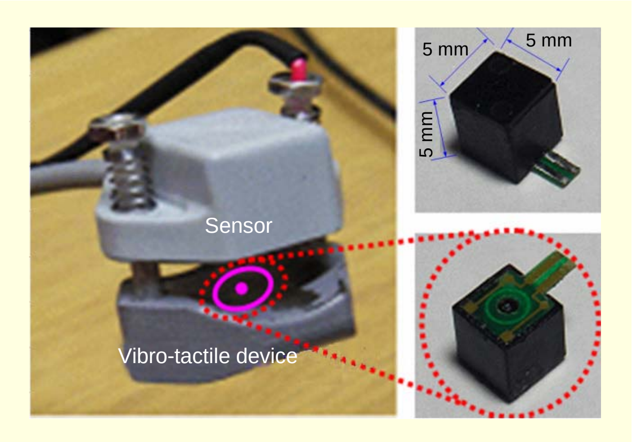

Fig. 2. A vibro-tactile device tied together with a position/orientation sensor.

tracking system. However, it is hard to control SMA wires to follow sudden changes in the contact situation of the fingers. Moreover, the power consumption and circuits required are not suitable for a wearable environment which requires small size and low power consumption.

To make the pins vibrate, a control board is also needed. Before this research, we made several kinds of control boards which can actuate various pin-type vibro-tactile devices and are capable of wireless communication to a personal computer.

This pin-type vibro-tactile device module delivers tactile stimulation to the human skin. The tip of the pin which vibrates lightly stimulates the skin. The fingertip is a suitable place to deliver tactile feeling, especially in this game feature. To deliver a more realistic tactile sensation, we compared the tactile sensations which are generated by the signal of various frequencies and amplitudes and chose the proper frequency and amplitude to give a sensation which is comparable to the sensation delivered when an actual bubble is popped.

2. Head-Tracked Stereo Display

A 3-dimensional stereo image is more realistic than a 2-dimensional plane image. A 3-dimensional stereo image is made on the 2-dimensional plane using the effect of binocular disparity with a factor which expresses a cubic effect using a far and near law and a perspective drawing law. The reason for the binocular disparity of the cubic effect is that the two eyes of an individual are about 6.5 cm apart on average. The two eyes look at images of a given object from different planes. The two images are transmitted to the brain, and through synthesis, the user can recognize a reconstructed 3-dimensional stereo image [11]-[13].

The stereo image supplies a sense of reality, a sense of presence, and a sense of depth to the user. However, even though virtual space is simply seen using a fixed stereo display, the sense of reality is not greatly enhanced. That is, if an equal image, independent of user's movement, is continually shown, the user might feel the cubic effect, but the sense of reality and the feeling of movement would decrease. Therefore, tracking the user's eyes is essential. Moreover, the virtual and physical coordinates should be aligned. Also, the distance between the eyes of each person should be considered in the design of the stereo display.

3. Displaying Stereoscopic Images and Tracking Methods

There are many methods of displaying 3-dimensional stereoscopic images. Our method uses shutter glasses with a high speed CRT monitor. When the monitor shows the left image frame, the shutter glasses shut the right eye and when the monitor shows the right image frame, the shutter glasses shut the left eye. As a result, each eye sees separate images. The vertical sync frequency used in the monitor is 120 Hz and the image frame of 60 Hz for each eye is seen alternatively. For synchronization between the shutter glasses and the monitor, an emitter which is synchronized to the monitor sends infrared signals to the shutter glasses.

To track the position and orientation of a particular object in 3-dimensional space, various types of sensors can be used. A typical method uses ultrasonic or magnetic tracking devices. For our purposes, we selected a magnetic sensor system. This device is quite good because of its high resolution capability, although it has some defects due to the limited distance between the source and sensors and it makes some errors due to other magnetic fields or metallic objects. To diminish errors, we set the source as close to the sensors as possible. This is because the 3-dimensional position/orientation of a magnetic-type sensor is determined by the relative location of the sensor from the source that is the origin of the sensor.

III. Realization of a Stereo-Vision-Based Human-Computer Interaction System

The virtual bubble popping game is naturally controlled by the user's movement. To point at a virtual bubble directly, coordinate alignment and view control processes are needed. Coordinate alignment converts the physical position and orientation of a sensor based on the source into a value based on the virtual coordinate frame. The view control sets a virtual object to be seen as a real object. As a result of this process, a user can point out the desired location on the virtual object with his/her own fingertip. Also tactile stimulation and the head-tracked stereo system increase the sense of reality of the human-computer interaction even more.

1. Coordinate Alignment

There are two coordinate frames in our system. One is a virtual coordinate frame for representation of the virtual environment and virtual objects, and the other is a physical coordinate frame for the measurement of the sensor position based on the source. Because each of the two coordinate frames has its own origin, we must align the two coordinates through a simple coordinate conversion. Actually, coordinate alignment is implemented by converting the physical coordinate frame into the virtual coordinate frame. Figure 3 shows the relative positions of the monitor, the source, and the sensor. One sensor tracks the fingertip and the other is attached to the shutter glasses to track the user's eye position.

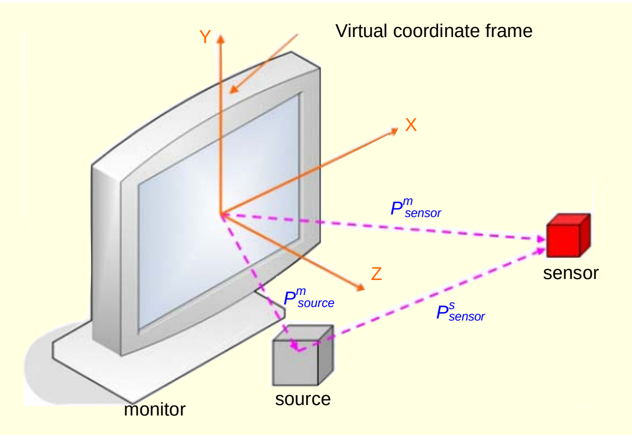

Fig. 3. Realationship between source and sensor for eye-tracking.

In the 3-dimensional virtual coordinate frame, we set the center of the screen as the origin of each axis. We also set the +z axis as the direction from the screen to user. As previously mentioned, the position value of the magnetic sensor is determined by the sensor's position based on the source. Accordingly, since each axis of the source does not agree with the corresponding axis of the virtual coordinate frame, it must be converted into its corresponding axis of the virtual coordinate frame. If each axis of the source is paralleled to the axis of the virtual space, the coordinate value of the sensor may be used directly as P^s_{sensor}. Otherwise, as seen in (1), the rotation matrix (R_{source}) must be calculated. The coordinate matrix based on the source is denoted as P_{raw}.

$$P^s_{sensor} = R_{source} \times P_{raw} \quad (1)$$

The sensor coordinates based on the virtual coordinate frames are simply calculated by a transformation matrix:

$$P^m_{sensor} = P^m_{sensor} \times P^s_{sensor}. \quad (2)$$

If the physical position of the display surface or source is changed, R_{source} and P^m_{sensor} must be recomputed.

2. View Control

Since an object in the virtual coordinate should be perceived as having the same size and shape as a real object according to the user's eye position and orientation, the virtual and physical coordinates must have an equal length scale and the view volume must be controlled in real-time. This problem can be solved by fixing the screen surface plane and a corresponding plane of virtual space through information about the monitor screen and tracked user eye position. The edges of the monitor screen can be measured easily by magnetic sensors.

Since we assume that the user's gaze is directed at the screen surface of the monitor, if the z-axis value of a virtual object is a negative number, the object must be seen as if it were located at a depth behind the monitor screen; if the z-axis value of a virtual object is a positive number, the object must be seen as if it were located in front of the monitor screen. When composing the stereo image, each frustum in virtual space must be controlled to look onto the x-y plane with the value of the z-axis as 0. The two viewpoints which correspond to the user's left and right eyes actually do not lie at the center of each eyeball and change as the user's eyes move [3]. We assume that the relative position of each viewpoint from the magnetic sensor attached on the shutter glasses is fixed in this system.

Looking at the same objects from the same location, users can see different images because the distances between each user's eyes are different. Therefore, the distance between the two virtual cameras should be controlled to be the same as the distance between the user's eyes. A virtual sensor is drawn the same size in the virtual space as the real sensor. Through calibration, when the virtual sensor generated as a stereoscopic image in the virtual space is hidden by the real sensor worn on a finger, view control is achieved.

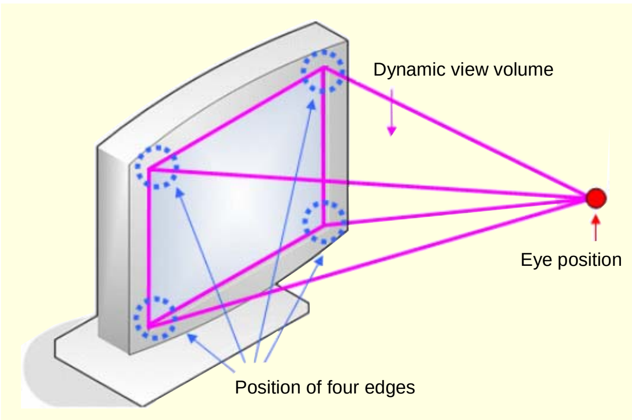

Fig. 4. Factors for view control.

3. Implementation of Virtual Bubble Popping Game

We implemented a virtual bubble popping game by using the device and technologies introduced in prior sections. Figure 5 shows a user playing this virtual bubble popping game and Fig. 6 shows a screenshot as seen by one eye. The user can feel some vibration on his/her index finger wearing the vibro-tactile display module when he/she successfully touches the virtual bubble, and then the virtual bubble pops.

Fig. 5. Playing the virtual bubble popping game.

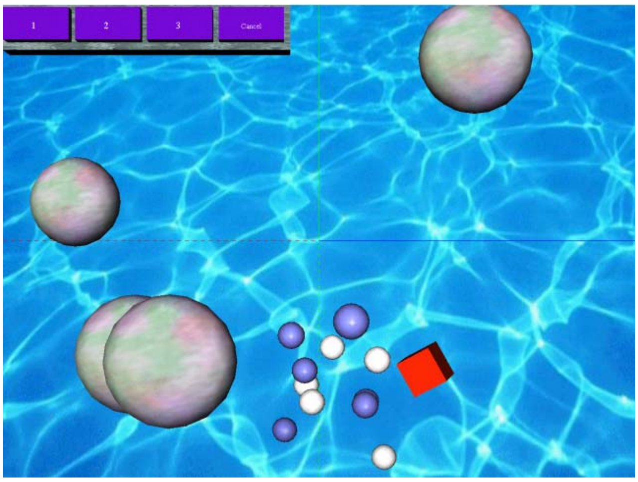

Fig. 6. Screenshot of the virtual bubble popping game.

IV. Conclusion

In this study, we implemented a stereo-vision-based human-computer interaction system in the form of a virtual reality game which enables a user to pop virtual bubbles with his/her own finger wearing a device as shown in Fig. 2.

We used our own vibro-tactile device module to imitate the tactile sensations experienced when popping real bubbles. Using a stereo vision system with position and orientation sensors, we achieved the alignment between the two coordinate frames and the alignment of a visual point by view control. The alignment of coordinate frames and the alignment of a visual point give the user the sense of reality as if actually popping real bubbles using a fingertip. We found that appropriate tactile stimulation and the use of a head-tracking system greatly increase the sense of reality.

References

[1] M.F. Deering, “HoloSketch: A Virtual Reality Sketching/ Animation Tool,” *ACM Trans. Computer-Human Interaction*, vol. 12, no. 3, 1995, pp. 220-238.

[2] M. Katzourin, D. Ignatoff, L. Quirk, J.J. LaViola Jr., and O.C. Jenkins, “Swordplay: Innovating Game Development through VR,” *IEEE Computer Graphic and Applications*, vol. 26, no. 6, 2006, pp. 15-19.

[3] M.F. Deering, “High Resolution Virtual Reality,” *Proc.the ACM SIGGRAPH*, 1992, pp. 195-202.

[4] R. Turner, E. Gobbetti, and I. Soboroff, “Head-Tracked Stereo Viewing with Two-Handed 3D Interaction for Animated Character Construction,” *Computer Graphics Forum*, vol. 15, no. 3, 1996, pp. 197-206.

[5] T. Meyer and A. Globus, “Direct Manipulation of Isosurfaces and Cutting Planes in Virtual Environments,” Technical Report CS-93-54, Dept. of Computer Science, Brown University, 1993.

[6] W. Piekarski, B. Avery, B.H. Thomas, and P. Malbezin, “Integrated Head and Hand Tracking for Indoor and Outdoor Augmented Reality,” *Proc. IEEE VR’04*, Mar. 2004, pp. 11-18.

[7] R. Scheibe, M. Moehring, and B. Froehlich, “Tactile Feedback at the Finger Tips for Improved Direct Interaction in Immersive Environments,” *Proc. the IEEE VR’07*, 2007.

[8] H. Regenbrecht, J. Hauber, R. Schoenfelder, and A. Maegerlein, “Virtual Reality Aided Assembly with Directional Vibro-Tactile Feedback,” *Proc. GRAPHITE’05*, 2005, pp. 381-387.

[9] J. Ryu and G.J. Kim, “Using a Vibro-tactile Display for Enhanced Collision Perception and Presence,” *Proc. VRST’04*, 2004, pp. 89-96.

[10] http://www.immersion.com/3d/products/cyber_touch.php

[11] K.T. Kham and J.H. Lee, “The Effect of Inter-Pillary Distance on Stereopis,” *Korean Journal of Cognitive Science*, vol. 14, no. 3, 2001, pp. 37-49.

[12] H. Kakeya, K. Oyama, and Y. Arakawa, “3D Display System for Reality-Enhanced Teleoperation,” *Proc. IEEE International Conference on System, Man, and Cybernetics (SMC’99)*, 1999.

[13] S. Miyake et al., trans. Eunsu Kim et al., *Basics of 3D Images*, Ohmsha Ltd. and Kidary, 2000 (in Korean).

Ho-joong Yong received the BS degree in electrical and computer engineering and the MS degree in electronic and telecommunication engineering, in 2005 and 2007, respectively, from Kangwon National University, Korea, where he is currently pursuing the PhD. His research interests include haptic devices, virtual reality, and augmented reality.

Jongwon Back received the BS and MS degrees in control and instrument engineering from Kangwon National University, Korea, 2000 and 2003, respectively. He is currently pursuing the PhD in the Department of Electronic and Telecommunication Engineering at Kangwon National University. His research interests include haptic device design, virtual reality, and augmented reality.

Tae-Jeong Jang received the BS, MS, and PhD degrees in control and instrumentation engineering from Seoul National University, Korea, in 1986, 1988, and 1994, respectively. He joined Kangwon National University in 1995 where he is currently a professor in the Department of Electrical and Electronic Engineering. His research interests include haptic interfaces, human interfaces, virtual reality, control theory and its applications, and digital communications.