PLAYING RUBIK'S CUBE IN MIXED REALITY

Makoto SATO, Yasuharu KOIKE

Precision and Intelligence Laboratory, Tokyo Institute of Technology, Japan

{msato, koike}@pi.titech.ac.jp

Somsak WALAIRACHT

Department of Computer Engineering, Faculty of Engineering,

King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

somsakw@ce.kmitl.ac.th

Abstract This paper describes a prototype system for virtual objects manipulation in mixed reality environment. SPIDAR-8, the multi-fingers string-based haptic interface device, is used to display force feedback at four fingertips on each hand of the user, i.e., thumb, index, middle, and ring finger. Fingertip positions measured by the haptic interface device are also used to model the 3D virtual hands. Then image sequences of real hands of the user taken by video camera are superimposed on the silhouette of the virtual hands. By employing the technique of sensors fusion and the successive adjustment algorithm, correct geometric registration of image of real hands and the virtual objects can be achieved. The user can see his or her real hands manipulating the virtual objects in the simulated virtual world and perceive force feedback in the same way as manipulating real objects. The system is aimed to enhance the sensation of immerse into the virtual world.

Keywords: SPIDAR, string-based, two-handed, multi-fingered, haptic interface device, mixed reality environment, virtual image registration

1. Introduction

Merging real and virtual images requires correct geometric alignment of both images in one single reference frame. Such issue is the research projects on the augmented reality (AR) systems, which try to augment computer graphics with real world scenes. Meanwhile, a virtual reality (VR) environment is a completely synthetic world. The merging of a live image of the user's interaction of the VR system enhances the user's sense of immersion inside the virtual world. Adding a genuine video scene adds a significant degree of richness and realistic detail at minimal computational expense to the completely modelled

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35660-0_65](https://doi.org/10.1007/978-0-387-35660-0_65)

R. Nakatsu et al. (eds.), *Entertainment Computing*

© IFIP International Federation for Information Processing 2003

environment. The integration of haptic sensation realized in VR into AR to become a mixed reality (MR) system is an attractive improvement. In this paper, we propose a prototype system for virtual object manipulation in a mixed reality environment. We have developed a two-handed, multi-fingered, string-based, haptic interface device. By using this interface device, the user can perceive force feedback at fingertips. When manipulating an object, the user can see his or her real hands in the virtual environment displayed on the screen. A sense of immersion into the virtual world is perceived. The depth cues between the hands and virtual objects can be determined using the sensed fingertip positions. Therefore, image sequences of the real hands can be superimposed on the virtual hands with correct alignment of the real and virtual images.

2. SPIDAR-8 — Two-handed multi-finger string-based haptic interface device

To construct a mixed reality environment where the user see a live image of his or her hands manipulating virtual objects, the haptic interface device should not be seen in the composite scene of the virtual world. Many haptic interface devices are not suitable, because either they attach the user's hand with complex structure of actuators, or they cover or hide the user's hand in the glove. We use our prototype haptic interface device that transparently attaches the user's fingers to light-weight strings. Attaching our device to the hands does not interfere the view of the user's hands or the virtual objects in the virtual world.

Figure 1. SPIDAR-8

SPIDAR-8 is a kind of string-based haptic interface device designed for multi fingers of each hand, as shown in Fig. 1. It is used to provide force feedback while the user is manipulating virtual objects in a simulated virtual world. Forces are displayed at four fingertips—the thumb, index, middle, and ring finger—on the left and right hand of the user. The device calculates each fingertip position in the virtual space by using the length of three strings. When the position of any fingertip comes to the contact with the position of a virtual object, force feedback is generated by controlling the tension of the strings.

3. System setup

The mixed reality environment setup consists of our prototype haptic interface device, SPIDAR-8, a CCD video camera, and an 18-inch LCD flat-screen display, as shown in Fig. 2.

Figure 2. System configuration

4. Implementation of mixed reality environment

First, the virtual hands are modelled from the fingertip positions measured by SPIDAR-8. Next, the virtual hands are compared with the projected image of the user's real hands and adjusted in position to try to make them coincide. Then, the image of the real hands is merged into the virtual world. In this way, correct occlusion of the user's real hands and the virtual objects can be achieved.

4.1. Model the virtual hand

A human’s hand has a very complex structure and is difficult to model, because on each finger alone there are many joints and linkages. The total number of degrees of freedom found on one hand is 26, which includes all joints on all fingers and 6 DOFs of motion at the wrist. In this system, the number of DOFs on a virtual hand is reduced to 17 to make it possible to estimate the virtual hand pose using the eight fingertip positions measured by SPIDAR-8. By the following criteria, we can reduce the number of DOFs of a virtual hand:

Figure 3. Model of the real hand and the reduced version of the virtual hand

- It is almost impossible to move the joint of the last link nearest to the fingertip without moving the next-to-last joint, and vice-versa. Therefore, the DOF of the joint of the last link is assigned to be proportional to the next-to-last joint on the thumbs and each of the fingers.
- It is assumed that the middle finger cannot be moved side-to-side in most cases of grasping an object. Hence, the joint at the link of the palm to the middle finger is reduced by 1 DOF.
- The little finger is not measured by SPIDAR-8, therefore, the joint motion of the little finger is assumed to resemble the joint motion of the ring finger.

However, only fingertip positions sensed by SPIDAR-8 are insufficient to solve the registration problem. Therefore, additional hand position information is necessary; the following data issued for this purpose are: the projected wrist position, and the joint-ball.

4.2. Projected wrist position

A projected wrist position can be found from screen data when an image of the user's real hand taken by the camera is displayed on the screen. The user wears a red bracelet marker on the wrist, as shown in Fig. 4. The screen data is raster-scanned and then the marker can be extracted by the color-filtering technique. Using the ends of the extracted marker, a position at the center is computed. This position represents a projected wrist position of that hand.

Figure 4. Bracelet marker

4.3. Joint-ball

In many cases the projected images of the user's real hands are not completely coincident with the virtual hands, especially at the finger joints. Such errors can be corrected by adjusting the finger joints of the virtual hands. By comparing the silhouettes of the real hand and the virtual hand in the camera image, a joint-ball is defined at each finger joint on the silhouette of the virtual hand as shown in Fig. 5.

Figure 5. Joint-ball

4.4. Successive adjustment of model parameters

A model of the real hand consists of four fingertip positions measured by SPIDAR-8 and the projected wrist position. A model of the virtual hand is constructed using the initial fingertip positions and wrist position. If all parameters of both hand models are matched, the projected image of real hand can completely coincide with the virtual hand. However, in practice, they are always different. The model of the virtual hand is adjusted at every update. The adjusting amounts can be found from the model of real hand at the current step and the approximately adjusted model of the virtual hand in the next step. The successive adjustment can be described as follows.

- Step 1.** Construct model parameters of the real hands from the updated fingertip positions measured by SPIDAR-8 and the extracted projected wrist position.
- Step 2.** Compute the differences of positions of all joint-balls.
- Step 3.** Compute the differences of model parameters in Step 1 with the model parameters of the virtual hands.
- Step 4.** Consider the integrated error, the differences of positions of all joint-balls and the differences of model parameters. If the integrated error is small enough, the process is finished. Otherwise, continue.
- Step 5.** Compute the adjustment amount of model parameters for the successive model parameters.
- Step 6.** Repeat the algorithm from Step 2.

4.5. Merging image of real hands into the virtual world

A CCD video camera is used to record the images of the user's real hands when performing a manipulation task. The chroma-key technique is used to extract only the image of the hands from a blue-screen background. We implemented a software routine for chroma-keying on the graphics server. After the model of the virtual hands has been adjusted as already described in the previous sections, the color pixels of the virtual hands are blitted to be transparent on the screen. From the three-dimensional position information from our model of the virtual hands, the virtual hands always occlude the virtual objects in the correct geometrical position. The extracted image of the real hands are then superimposed on the transparent virtual hands. In the virtual world scene, the image of the real hands can be merged with the virtual objects with correct occlusion.

5. Implementation result

Figure 6 shows the constructed mixed reality environment where the user is manipulating a virtual Rubik's cube with force sensations and can see his real hands in the virtual world manipulating the virtual cube.

Figure 6. The implemented system providing mixed reality environment

However, we can observe the registration errors in the mixed reality scenes, as in Fig. 7. Space appeared between the user's fingers and the virtual cube in some cases, and sometimes certain parts of the finger were covered by the virtual cube. These errors are suspected to be caused by the heavy computational requirements of this process. The estimation of the virtual hands pose is slow. It is very difficult to make the virtual hands coincide exactly with the real hands. In addition to the limited measurement position information provided by the interface device, the estimated position information using the joint-balls and the projected wrist position were employed. The estimation improved, however, some error still remains. Further, the graphics server currently has inadequate computation power to handle these tasks quickly enough to eliminate these errors. Replacing the graphics server with a higher performance computer could help reduce registration errors.

6. Conclusion

In this paper, we have proposed a mixed reality environment system in which an image of the user's real hands is merged into a virtual world when performing virtual object manipulation. The user can see his or her real hands and perceive force feedback at the fingertips when manipulating virtual objects. The composition of real and virtual images with correct registration can be achieved by a successive adjustment algorithm that we have introduced. The implementation results have shown a feasible mixed reality environment. Although the problem

Figure 7. Screens of mixed reality environment

of registration remains for some cases, the proposed system is promising for a feasible implementation of a mixed reality system.

References

- Feiner S., B. MacIntyre, M. Haupt, and E. Solomon. (1993) Windows on the World: 2D Windows for 3D Augmented Reality. *Proc. of UIST'93*, 145-155.
- Milgram P., and F. Kishino. (1994) A Taxonomy of Mixed Reality Virtual Displays. *IEICE Transaction on Information and Systems*. Vol. E77-D, No. 9, 1321-1329.
- Azuma R. T. (1997) A Survey of Augmented Reality. *Presence*, Vol. 6, No. 4, 355-385.
- Edwards E. K., J. P. Rolland, and K. P. Keller. (1993) Video See-through Design for Merging of Real and Virtual Environments. *Proc. of VRAIS'93*, 197-204.
- Okuma T., K. Kiyokawa, H. Takemura, and N. Yokoya. (1998) An Augmented Reality System Using a Real-time Based Registration. *Proc. of ICPR'98*, Vol. 2, 1226-1229.
- Tsai R. Y. (1986) An Efficient and Accuracy Camera Calibration Technique for 3D Machine Vision. *Proc. of CVPR*, 364-374.
- Yokohohji Y., R. L. Hollis, and T. Kanada. (1996) What You See is What You Can Feel - Development of a Visual/Haptic Interface to Virtual Environment. *Proc. of VRAIS'96*, 46-53.
- Yokoya N., H. Takemura, T. Okuma, and M. Kanbara. (1999) Stereo Vision Based Video See-through in Mixed Reality - Merging Real and Virtual Worlds. *Ohmsha Springer Verlag*, 131-145.
- Walairacht S., K. Yamada, Y. Koike, and M. Sato. (1999) Modeling Virtual Hands with Haptic Interface Device. *Proc. of ICAT'99*, 233-236.
- Walairacht S., M. Ishii, Y. Koike, and M. Sato. (2001) Two-Handed Multi-Fingers String-Based Haptic Interface Device. *IEICE Transactions on Information and Systems*, Vol. E84D, No. 3, 365-373.