Human-Computer Interaction ─ INTERACT'03
M. Rauterberg et al. (Eds.)
Published by IOS Press, (c) IFIP, 2003, pp. 1107-1108

3D Model Viewer with Real-time Viewpoint Tracking System

Jun Iio, Tomoyuki Yatabe & Kazuo Hiyane

Mitsubishi Research Institute, Inc., Otemachi 2-3-6, Chiyoda-ku Tokyo, JAPAN

iiojun@mri.co.jp

Abstract: Today, it has become very popular to represent 3D CG models onto the screen of a workstation or personal computer. In almost all software that displays 3D models, users can control their viewpoint of projection with pointing devices such as a mouse or trackball. However it is not intuitive to control the viewpoint with such devices, because there is an insufficient degree of freedom. Although locating viewpoint needs 6 degrees of freedom (3 degrees in position, and 3 degrees in direction), off-the-shelf pointing devices are basically two-dimensional controllers. We propose a 3D model viewer with a real-time user's face tracking system. This system provides an easy method to make observations of 3D models from any direction and to any scale that users may choose.

Keywords: 3D model, face tracking, viewpoint, pointing device

1 Introduction

There are many software packages; which handle objects represented by three-dimensional data, such as 3D modeller, 3D-CAD (Computer Aided Design), and some kind of games. However, a device such as stereoscopic vision is still expensive. Generally we observe 3D data as a projected image on screen.

The system needs to fix a viewpoint that gives a basis of projection for rendering silhouette of objects on projection plane. There is no intuitive way to control the viewpoint using mouse, trackball, or some standard pointing devices. Difficulty to locate viewpoint using such devices result from a lack of flexibility. A mouse has two-degrees of freedom in its control capability, but this is insufficient to specify viewpoint position and direction in 3D space.

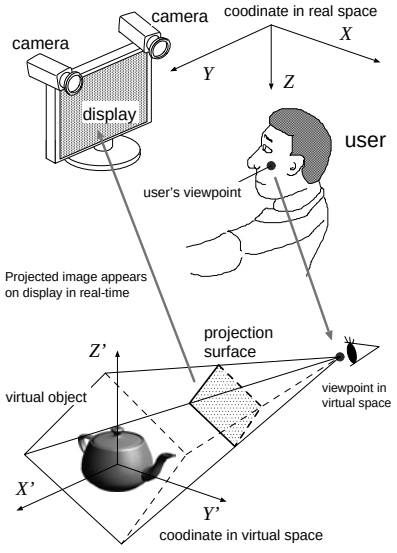

Figure 1: The concept of the proposed system.

Ordinary software realizes viewpoint control by modal interface. This involves changing control-mode using a menu in a GUI, or by concurrent keyboard use.

This solution suffers from at least two problems. First, it is not intuitive and it requires users to learn its usage. In addition, if users want to change the model, they cannot modify the model while viewing it from various scales and directions, simultaneously.

2 Overview

In our demonstration, we propose a new 3D model displaying system, which provides an easy pointing method for 3D model projection. In our system, controlling the viewpoint in model space is based on measurement of observer's eye point in real space.

Three-dimensional information of a user's face position is measured in real-time, and used to create a new projection image of the model. The concept of proposed system is shown in Figure 1.

3 System Architecture

The proposed system has two cameras, and the three-dimensional position of the user's face is calculated from two images. Two cameras are set up to such that the user is placed in the center of its image. The picture of using our system is illustrated

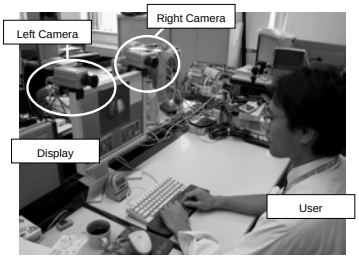

Figure 2: User sitting in front of display and cameras.

in Figure 2.

Cameras are connected with the PC via the IEEE1394 serial bus. The size of the image is QVGA (320x240 pixels), the image data format is YUV (or RGB) and 30 fps images can be transferred without compression.

The software is built on a Linux platform. Recognising face position and model rendering processes run concurrently as multi-threaded applications.

A face recognition process is implemented for each camera and can be carried out independently. This image recognition part is developed with MAlib (Iio et al., 2001, 2002).

MAlib is a general-purpose image handling software library written in C-language. We are offering it as open source software and can be found at **`http://www.malib.net/`**.

4 Algorithm

Our approach to detecting the user's head position is based on a `disparity map' that is calculated from stereoscopic image pairs.

First, input image { EMBED Equation.3 } is separated from the background part { EMBED Equation.3}and the user part {EMBED Equation.3} , accord-ing to the following equation;

{EMBED Equation.3}

where { EMBED Equation.3 } is the disparity value at pixel {EMBED Equation.3}, and {EMBED Equation.3} is a threshold value.

The system we proposed previously (Iio et al., 2002) used skin-color base face tracking. But an object tracking method using color or brightness information is basically sensitive to lighting conditions. The result is that the system is not so robust to changes in the system environment.

Next, a labeling process is applied to remove noises and to find the largest area of conjunction pixels. We assume that the largest and the closest to the center area represents an image of the user's face. It is considered reasonable because the user is sitting in front of the cameras and a shape of the user occupies almost all of the whole image.

Finally, a median point {EMBED Equation.3} in face candidate area is computed. Then we can calculate the 3D position of the user's head in real space as follows;

{EMBED Equation.3}.

5 Application

A prototype system was used to display 'virtual globe' (Figure 3). The main window shows a CG globe image observed from the user's position. Two small windows on the right side display the status of the user face recognition.

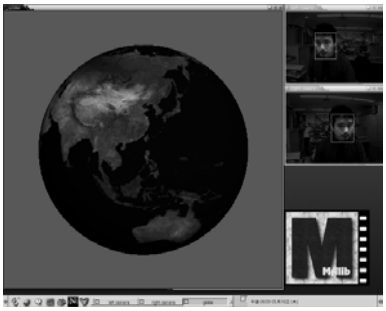

Figure 3: Screenshot of application 'Virtual Globe'.

References

Iio, J., Yatabe, T. & Hiyane, K. (2001), Generic Video Image Processing Library - MAlib -, *Linux Conference 2001,* http://lc.linux.or.jp/lc2001/papers/malib-paper.pdf [In Japanese]

Iio, J., Yatabe, T. & Hiyane, K. (2002), Video Image Processing with Linux and Its Applications, *Linux Conference 2002,* http://lc.linux.or.jp/lc2002/papers/iio0920h.pdf [In Japanese]

Iio, J., Yatabe, T. & Hiyane, K. (2002), An User Interface Using 3D Information of User's Head Position, *Proceedings of the eighth Symposium on Sensing via Image Information,* pp.573-576. [In Japanese]