A Natural User Interface to Control an Augmented Reality Application in the CAVE at the HTW Berlin

Hannes Helmholz, Thomas Jung, Rico Kenzler, Franz Simon

Hochschule für Technik und Wirtschaft Berlin
Wilhelminenhofstraße 75A
12459 Berlin
t.jung@htw-berlin.de

Abstract

This paper introduces a Natural User Interface (NUI) that enables interaction in a Cave Automatic Virtual Environment (CAVE). The tracking of the head and hand in a CAVE usually takes place with the help of magnetic tracking systems or marker-based optical tracking systems. In both cases the user has to wear a number of instruments. In contrast, the CAVE at the HTW Berlin uses two depth detection cameras (currently Microsoft Kinect) and the data provided is then merged. This records the user's entire skeleton without the use of instruments and it allows applications to be controlled by a NUI.

CAVE systems are usually used for virtual reality applications. The application described in this case visualises the CAVE itself, creating a spatial augmented reality system in which the user can interact directly with the real CAVE. The physical CAVE walls are aligned to the virtual walls so that the user can touch the walls and receive realistic haptic feedback.

1 Introduction

Paul Milgram and others define mixed reality environments as environments in which real and virtual objects are displayed together. Variants of mixed reality environments can be classified based on a reality/virtuality continuum [Milgram 1994]. At one end of the continuum is reality, in which we all live, while at the other end is virtuality, which consists solely of synthetically rendered computer models.

In virtual environments, the immersion of the user should give the impression of a virtual reality. In this case, immersion means the feeling of being "immersed" in a virtual world, which can also be identified by a reduced perception of one's self in the real world. Virtual reality applications generally use head-mounted displays [Sutherland 1968] or Spatially Immersive Display Environments (SIDE). The most well-known SIDE is the CAVE [CruzNeira 1993].

Augmented reality (AR) applications are at the centre of the reality/virtuality continuum and augment reality with additional, mostly visual, stimulations. Azuma defines augmented reality systems as those in which virtual reality and reality are combined (partly superimposed). Interactivity has to take place in real time, while the real and virtual objects are connected in 3D [Azuma 1997]. AR systems use hand-held displays, head-mounted displays or project in a real environment. To the authors' knowledge, AR applications using a CAVE have not previously been described.

CAVE systems have been used to create immersive virtual environments since the 1990s. The interaction initially took place using magnetic tracking systems and later using optical tracking systems for which users were generally equipped with active or passive markers.

Natural User Interfaces avoid the use of visible control elements to the greatest possible extent in order to ensure a more natural control. Depth detection cameras that determine the distance to the picture elements enable gestures to be identified without markers. The development of more cost-effective devices (Microsoft Kinect, Asus Wavi Xtion) means that the development of NUIs is gaining in importance.

2 Natural User Interface based on depth detection cameras

The HTW CAVE has had a NUI based on a Kinect camera since 2011. The sensor data supplied directly from the Kinect is suitable for head tracking and navigation interfaces, although the data has so far been too imprecise for selection tasks [Jung 2011]. In addition, a single depth detection camera cannot sufficiently cover the interaction space of a CAVE. As a result, two Kinect cameras are now used in the CAVE described here, in order to enlarge the range of detection and to improve the sensor data by merging the two.

The Kinect cameras each emit infra-red patterns into the scene, which are then recorded in order to calculate the depth readings. As the infra-red patterns of several cameras interfere with each other, the cameras have to be aligned so that they each cover different areas. Both cameras must be calibrated for the data fusion. Calibration using a calibration object is difficult due to the alignment of the cameras. In the system described here, a thin wooden plate is used to allow joint corner points and edges to be found without the need for the same areas to be in the camera images. The recorded scenes are registered based on the point clouds recorded in the depth images [Jung 2012] (see Figure 1). The skeletal data can be merged in situations where they are recorded with sufficient precision by both cameras in order to improve the accuracy. Otherwise, at least one camera is able to supply a sufficiently precise skeleton, whereby the interaction area is now suitable for natural interaction [Jung 2012]. The controller-free interaction enables innovative interaction scenarios as is explained in the following section.

The Unity3D engine is used to display the virtual world so that new 3D worlds can now be easily visualised in the CAVE, in contrast to earlier approaches based on special software.

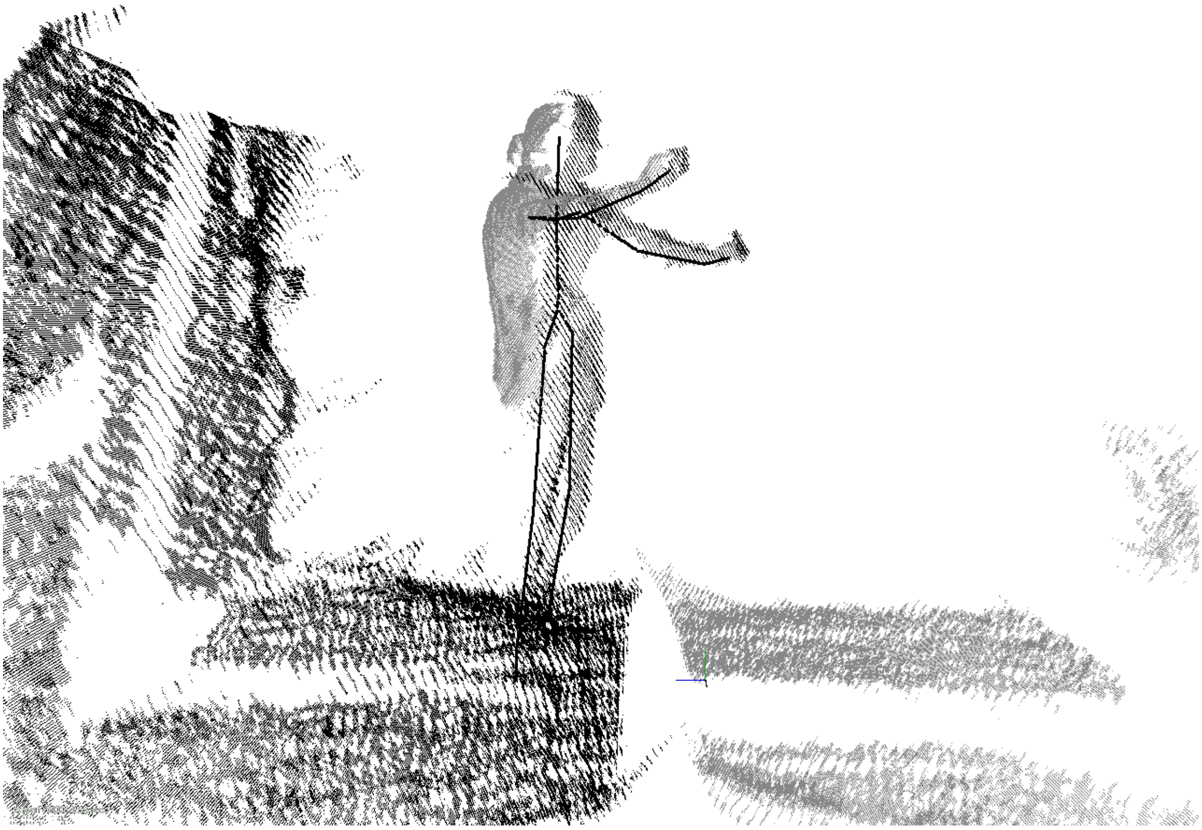

Fig. 1: point clouds in the CAVE as recorded by the two cameras (light grey/dark grey), the black lines show the merged skeleton

3 Augmented Reality in the CAVE

The objective of virtual reality applications is to completely block out reality in the user's perception, including all the devices used. This is virtually impossible using current technology. Head-mounted displays are too heavy and their field of vision is too narrow. In CAVE systems - even with controller-free interaction - the screens are still recognised as such, especially in the corners of the CAVE. In the system described here, the walls of the CAVE are also geometrically represented in virtual reality (see Figure 2, left side) so that the contrast between the perception of virtuality and reality is temporarily dissolved. 3D control elements can be virtually complemented on the real walls of the CAVE so that the user is supported by realistic haptic feedback during operation. A good example of an action that is triggered by these control elements is currently the virtual removal of the CAVE walls, which gives the user a view of the equipment room located behind the wall and explains the functions of the individual components.

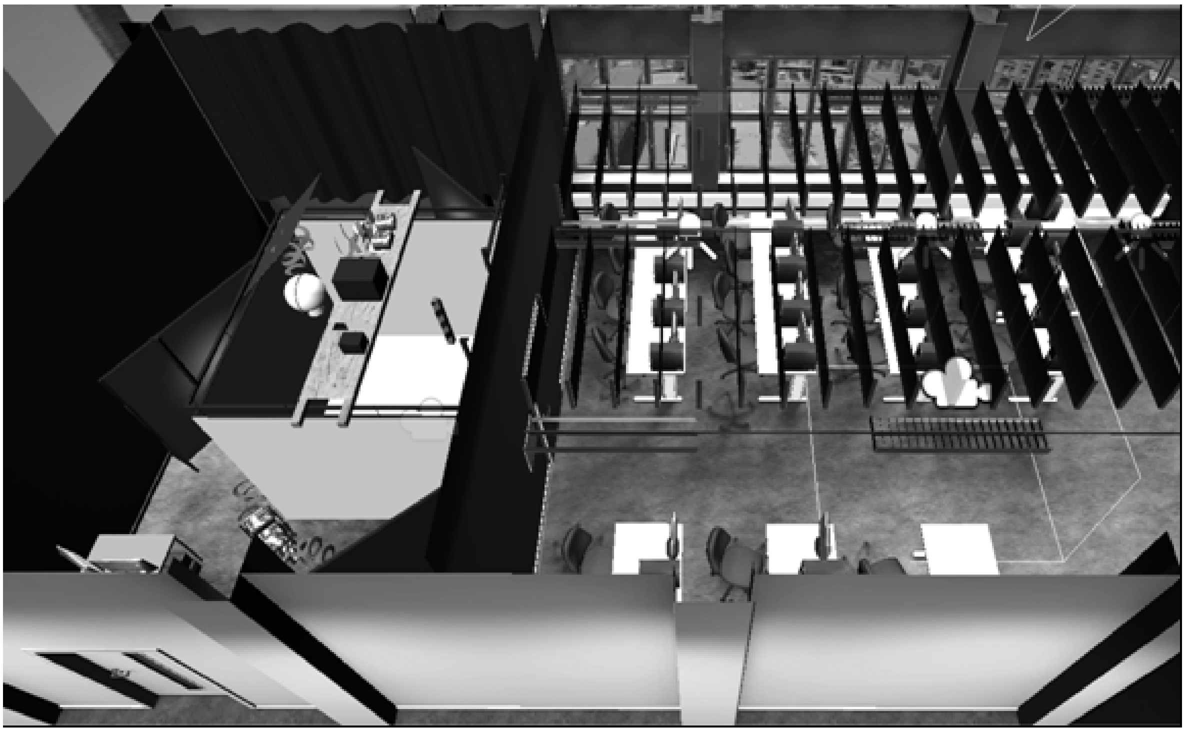

Fig. 2: 3D model of the CAVE (left) and the adjacent laboratory

Literature

[Milgram1994] Milgram P., Takemura, H., Utsumi, A., Kishino, F.; Augmented Reality: A class of displays on the reality-virtuality continuum, SPIE Vol. 2351, Telemanipulator and Telepresence Technologies (1994) pp. 282-292

[Sutherland1968] Sutherland, I. E. A Head-Mouted Three-Dimensional Display. Proceedings of AFIPS 1968, pp. 757-764

[CruzNeira1993] Cruz-Neira C., Sandin D.J., DeFanti T. A., “Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE”, Computer Graphics, SIGGRAPH Annual Conference Proceedings, 1993. pp. 64–72

[Azuma1997] R. Azuma: A Survey of Augmented Reality. In: Presence: Teleoperators and Virtual Environments. 6, Nr. 4, 1997, pp.355–385

[Jung2011]. Jung T., Krohn S., Schmidt P., “Ein Natural User Interface zur Interaktion in einem CAVE Automatic Virtual Environment basierend auf optischem Tracking”, In Proc Workshop 3D-NordOst 2011, Berlin, Germany, December 2011, pp 93-102

[Jung2012] Jung T., Simon F., Interaktion in einem CAVE Automatic Virtual Environment unter Verwendung mehrerer Tiefensensor-Kameras, In Proc Workshop 3D-NordOst 2012, Berlin, Germany, December 2012, pp 199-208