Affordable Immersive Projection System for 3D Interaction

C. Andújar   M. Fairén   P. Brunet
Centro de Realidad Virtual, U.P.C.
c/ Llorens i Artigas 4-6, 1ª planta
E08028 Barcelona
crv@lsi.upc.es

Abstract

This paper describes an affordable Virtual Reality system designed and developed by a group of researchers at the Polytechnic University of Catalunya (UPC). The system allows direct selection and manipulation of virtual 3D objects. The interaction is based on stereoscopic images projected over the user's working space and on devices tracking the user's natural movements. The system includes a screen being adjustable both in orientation and height, sensors tracking the head and hand movements, and a tactile device for the forefinger providing touch sense. A prototype of the system is currently exhibited at the Virtual Reality Center of Barcelona and it is being used in different application fields like architecture, medicine and industrial design.

Keywords

Interaction, immersive systems, stereoscopic projection system

1 INTRODUCTION

A growing number of applications in computer graphics are based on the interactive inspection and manipulation of 3D objects. In such applications, most of the time is spent on providing the system with proper input for exploring a geometric model from different viewpoints, and for selecting and manipulating 3D objects [1].

Selection refers to the act of specifying or choosing an object for some purpose, e.g. to restrict the scope for further actions. Manipulation is the task of changing the position and orientation (and possibly other features such as shape) of a selected object.

Since target objects can be out of the user's field of view, and in order to allow comprehensive inspection from arbitrary viewpoints, such applications also provide interaction mechanisms to allow the user to navigate through the model. Viewpoint control motion refers to two separate tasks: the decision process involved in determining the desired direction and target of travel (wayfinding), and the actual movement of the user's viewpoint into a different location.

Object manipulation and inspection are still tedious and time-consuming tasks. The growing complexity of geometric models involving thousands of objects further increases the need for efficient Human-Computer Interaction (HCI) techniques and devices.

Virtual Environments (VE) offer a new human-computer interaction paradigm where users are no longer simply external observers of images on a computer screen but are active participants on a computer-generated virtual world [1].

Among the different interaction paradigms for the above tasks, direct or implicit techniques are specially convenient, since they are completely intuitive and require no training [2]. Direct selection and manipulation derive from the real-world metaphor of in-hand manipulation. The user selects an object by "touching" it with a virtual representation of his hand, and manipulates it directly by moving his hand using the same gestures as with real objects [1]. Direct viewpoint control also relies on a real-world metaphor. The system tracks the user's head movements so that the virtual camera is updated accordingly.

There are several components that a system must have in order to provide direct manipulation and inspection of 3D models [3]:

- Real-time stereoscopic visualization,

- Sensor devices to capture user's actions,

- Haptic devices for providing the user with contact feedback whenever an object is touched.

Over the last few years, several Virtual Reality solutions for implicit interaction have been proposed, including completely immersive systems (e.g. Head-Mounted Displays, HMD) and semi-immersive systems (e.g. CAVE [4, 5], Responsive Workbench [6, 7]).

The most serious limitation of these systems is the high price of hardware and projection devices for stereoscopic visualization. Other limitations include special space requirements (e.g. CAVE-like systems do not fit in conventional office locations), ergonomics (e.g. most HMD cannot be user over long periods of time).

This paper describes a semi-immersive interaction system based on two key elements. On one hand, a projection screen which can be adjusted both in orientation and height and which receives real-time stereoscopic images. On the other hand, a tracking system which captures user's head and finger movements through optical recognition. Our system has several points in common with the system described in [6], but our system uses more economical components, introduces new interaction techniques and allows the user to easily adjust the projection screen so that objects are project into the user's workspace.

The rest of this paper is organized as follows. Section 2 describes the objectives of the project; Section 3 reviews current virtual reality systems, focusing both on projection systems and interaction techniques. The design and construction of the system are described in Sections 4 and 5, respectively. Features and main application areas are discussed in Section 6. Finally, Section 7 is devoted to concluding remarks and future work.

2   OBJECTIVES

This project was undertaken to build an affordable visualization system featuring simple but efficient interaction mechanisms for 3D model inspection and manipulation. The system is intended for medium-size companies, research institutes and hospitals with not enough resources to be able to acquire an expensive virtual reality system. In order to cover the needs for such organizations, the system has to fulfill the following extra requirements.

On one hand, intuitive and effective manipulation requires direct interaction mechanisms, so that users can handle virtual objects using the same actions as when working with real-world objects. Direct manipulation brings extra requirements which are discussed in next sections.

On the other hand, the system must support several users viewing the stereoscopic images without imposing barriers in the communication with other people, so that the system can be used for collaborative work (e.g. several specialists discussing about a design), for training (the instructor supervises the student's

actions) and for experimentation and demonstration (providers show interactively different options to the customer).

Finally, the system must be affordable, i.e., the overall cost must be of the same order of the graphic workstations required for 3D model visualization. Moreover, the system has to fit on conventional office spaces and must be easy to transport; otherwise the final cost would be increased.

3   PREVIOUS WORK

This section reviews both existing stereoscopic visualization systems and interaction devices from the point of view of their suitability for the desired system described above.

3.1   Stereoscopic visualization systems

State-of-the art stereoscopic systems include basically two visualization paradigms: stereoscopic helmets and projection-based systems.

Stereoscopic helmets are completely immersive in the sense that they completely replace the images from real world with computer-generated images from the virtual world. These devices can be classified into two categories: Head-Mounted Displays, HMD [8], and Head-Coupled Displays. These devices are not suitable for the desired system because they are completely immersive and avoid the vision of the user's environment including his own body. Several problems arise with such a complete immersion. Direct 3D object manipulation on empty space requires a high degree of coordination between sight and hand movements, but this coordination in HMD is too weak because the user only sees a virtual representation (avatar) of his hand. The user of a stereoscopic helmet cannot see the other members of the group, therefore communication is more difficult. For the above reasons stereoscopic helmets are not suitable for collaborative work, which was one of our system's requirements.

In projection-based systems, images are projected in one or more screens (usually back-projection screens) which can adopt different configurations varying their number, shape and arrangement. The most important commercial configurations are the Computer-Animated Virtual Environment, CAVE [4, 5] and the Workbench [6]. A CAVE consists of up to six 3x3 m projection screens forming a cubic room, whereas the Workbench is equipped with a horizontal screen similar to a desk. Other configurations are the stereoscopic walls (a single vertical projection screen), Holobench (two L-shaped screens), domes (semi-spherical screens) and cylindrical walls.

Projection-based systems have several advantages with respect to HMD. The user has no need to carry out the display device, which isolates him from the
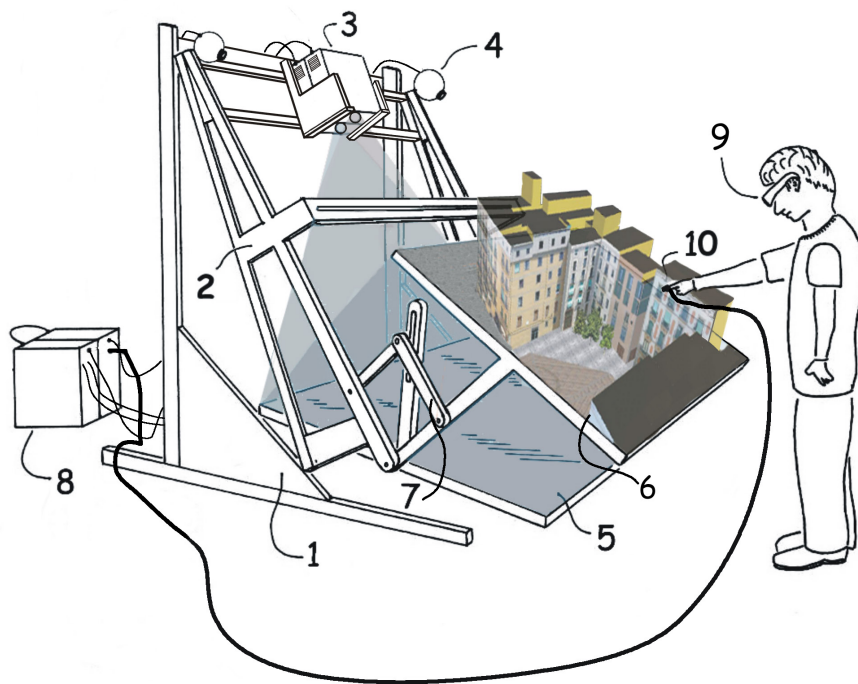
Figure 1. Architecture of the proposed system

real world and limits his interaction with other users. In projection-based system, the user can see his own body, which avoids the disorientation problem, quite common when working with HMD. In both cases user's head position has to be captured by tracking devices in order to compute correct stereo perspectives. In projection-based systems this information is used only to introduce a subtle correction of the view frustum, but instability of HMD during head rotation can produce the so called flight simulator syndrome.

Projection-based systems allow one or more users to perceive computer-generated stereoscopic images with depth sense, using any of the existing mechanisms for separating left-eye and right-eye channels, e.g. shutter, polarized or anaglyph glasses and autostereoscopy. Stereoscopic images cause virtual objects to be perceived in front, on or behind the projection screen (depending on adjustable parallax values). Stereoscopy provides a better interpretation of 3D data sets and enables direct interaction with virtual objects.

3.2   Devices for implicit interaction

Virtual reality uses implicit or direct interaction in opposition to explicit or classic interaction. In classic interaction paradigms, the user has to communicate explicitly his will to the computer, adopting the communication scheme defined by the application's interface, e.g. a WIMP interface (window-icon-menu-pointing device).
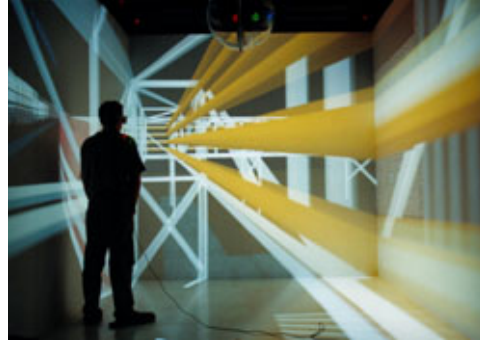
Unlike classic interaction, direct interaction captures the user's will which is implicit in his natural movements. The most common example is the viewpoint control, which is continuously updated according to user's head movements. The main revolution in the human-computer interface is that the user's does not perceives the computer interface and interacts directly with the computer scene. The perceptual difference is that the user does not look through a window but is immersed in a virtual world.

There are plenty of interaction devices which have been designed for implicit interaction. Input devices are sensors that capture user's actions and send this data to the system. Output devices are effectors that generate stimuli for user's senses, translating computer-generated digital signals into full-color images and sound.

Most common sensors include position tracker, flexion trackers (usually mounted on hand fingers) and voice recognition. Position tracker systems capture continuously user's movements, mainly head position (to compute correct perspective) and hand position (for direct interaction). Position tracker can be based on a broad range of technologies: optical recognition, magnetic fields, acoustics, mechanical links, and inertial devices. In our system, user interaction is based on a novel haptic device mounted on a user's finger.

(a) Head-Mounted Display

(b) Interior of a CAVE system

Figure 2. Immersive and semi-immersive virtual reality systems

4 DESIGNING OUR SYSTEM

The most important decision in the design steps relies on choosing the visualization paradigm to use. Since we want a multiuser system allowing collaboration among them for the design revision of 3D models, our design must avoid any sort of barriers between users, facilitating their communications. In this sense thus we decided to use a semi-immersive system based on back projection instead of using totally immersive devices like HMD.

Almost all the rest of the design decisions are guided by the requirement of making the total cost of the system affordable by the sort of organizations it is designed for. In this sense we discarded those projection systems including more than one screen (CAVE, Holobench, etc.) because the screen multiplicity affects to the number of projectors, graphical pipes and physical space required by the system, causing the system cost to become much higher. For all these reasons, we decided to make our system with only one screen.

In order to facilitate the portability, our system uses low-price, light-weight DLP or LCD projectors. Given the resolution of these projectors (1024x768 in our prototypes) and the dimensions of the screen we achieve an angular resolution which is enough for the human perception. Furthermore, we must think also in the total height, in order to allow its installation in a conventional room. Since there is a minimum distance required from the projector lens to the projection screen and we want to minimize the total space required by the system, we decided to use mirrors to reduce the distance needed between the projectors and the screen.

Other design decisions are related to the direct manipulation of virtual objects. In this sense, stereoscopic visualization, a tracking system and a tactile device are also very important features.

The possibility of manipulating virtual objects in the 3D space requires the user to feel the objects spatial distribution easily and also requires a high degree of coordination between sight and hand movements. The solution chosen for the tactile device and the tracking system will be discussed in the following section.

The stereoscopic vision allows virtual objects to be perceived in front, behind or on the screen plane. In our case we need to grant those objects will be perceived in the user working space, because the user must be able to *touch* them easily. By user working space we mean the space region easily reachable by the user while working normally with the system.

Other aspect to be taken into account with respect to the stereoscopic projection is the fact that when the user looks at a stereoscopic pair the natural relation between the crystalline accommodation and the convergence angle of the eye axes is broken. This is because the accommodation is adjusted to the distance from the observer to the screen (focusing distance), while eye convergence depends on the parallax of the object the user is looking at [9]. If this disarrangement is high the possibility of getting visual fatigue increases.

For all these reasons the screen must be adaptable to the user in order to increase the sense of presence of the virtual objects, so our design should include a screen being adjustable in orientation (from horizontal to vertical) and height.

Presently, there are systems where the position of the screen is just one, being vertical, horizontal or inclined. There are also systems where the screen is adjustable in orientation (like in Responsive Workbench of TAN), and systems where the screen is adjustable in both orientation and height as well (like in Baron of BARCO).

All projection systems based on adjustable screen need to solve the problem of granting the projector-generated image remain correctly positioned with re-

spect to the screen when the position of the screen is modified. There are solutions where the projectors, the mirror and the screen are joined in a rigid subsystem which can be moved altogether with respect to a structure laying on the floor. This is the solution adopted in systems like V-Desk 5 of Trimension or the Versabench system of Fakespace. In order to avoid the inconveniences coming from the alteration of the projectors orientation, we will use a solution based only in the modification of the orientation of the reflectant surface with respect to the screen.

The developed system is a new stereoscopic visualization system based on back projection over a screen which is adjustable in orientation and height using a new mechanism to transmit the angular movements of the screen to the angular movements of the reflectant surface. This allows the projectors to be in a fixed orientation and reduces the total space required by the system.

The problem of multi-user tracking has not been solved successfully yet, but this problem is not as important in a small system like this as in a CAVE where you can have a wide distortion in the edges of the screens because of another user is guiding the interaction. Moreover, in a stereoscopic table, users usually are nearer one each o ther, so the effect of not being the guiding user is almost unnoticeable.

5   DESCRIPTION OF THE SYSTEM

In this section we are going to describe the system in more detail and to concrete several aspects of it.

5.1   Global structure and the projection system

Figure 1 shows graphically the global structure of the system. First of all, the system consists of two wooden supports, the first one fixed (figure 1–1), and the second one mobile (figure 1–2), which is able to slide along the fixed one allowing the adjustment of the screen height without altering the right setting of the image in the screen. To facilitate this movement we use small wheels and a motor doing a pulling power over the mobile support.

The mobile support also includes all the needed parts for the projection system on the screen. It consists of two projectors (figure 1–3) projecting the images generated by two computers; two video cameras (figure 1–4) allowing an optical positioning and tracking system; the support for the mirror and the mirror itself (figure 1–5) where the images are projected from the projectors; the support of the screen and the screen itself (figure 1–6) where the images reflected in the mirror are projected; and two auxiliary pieces, each one making a parallelogram with the mobile support arm, the mirror support arm and the screen support arm (figure 1–7).

The parallelogram joining the mobile support, the mirror support and the screen support is the mechanism allowing the adjustment of the screen orientation (from the horizontal position to the vertical position). It works by transmitting changes on the screen orientation to changes on the mirror orientation in such way that the images projected by the projectors always go correctly to the screen. The orientation changes on the screen only affect to the mirror. The projectors remain static, so the rotating angle of the mirror must be half of the rotating angle of the screen. This solution is proved stable and rigid in terms of the movement of the screen. This rigidity allows a precision of 2mm (measured over the screen) on the image fit on the screen, which is not noticeable by the human eyes.

Apart from the physical structure, the system consists also of two computers (figure 1–8), which generate the images corresponding to both eyes simultaneously. The system, thus, implements a *passive stereo* visualization [9]. The passive stereo is based on the use of filters that polarize the images for the two eyes which are projected simultaneously on the screen. The polarization of the image for one eye is totally inverse to the polarization of the image for the other, so by using glasses also polarized in the right way, the *right image* is perceived only by the right eye and the *left image* is perceived only by the left eye. The passive stereo is implemented by producing the two images (this can be done using a computer for each image, as in our system, or using just one computer with a graphic card having two output channels) and sending each image to each one of the projectors. The main advantage of the use of passive stereo is its low cost. It does not require the use of projectors with very high vertical frequency (as it happens with active stereo), so the cost of the projectors becomes affordable.

With respect to the tracking subsystem, currently the system is able to work with two different possibilities:

- The first one is a *magnetic* tracking system, consisting of a transmitter of magnetic fields which produces three orthogonal magnetic fields, and two sensors sited at the positions to be detected which are able to detect these magnetic fields and give the computer the position and orientation of the sensor with respect to the transmitter of fields.

- The second one is an *optical* tracking system consisting of two cameras sited on the mobile support (figure 1–4), which detect infrared LEDs put on the positions to be detected. To facilitate this detection, the cameras use filters which only allow the infrared light to pass through. From the two images captured by both cameras and a previous calibration indicating the orientation and position of the screen, we can determine the po-

sition of the infrared LEDs in space with respect to the virtual model.

From these two tracking systems, the *magnetic* is the most precise and it is not sensitive to occlusion, but on the other hand it is limited in the distance accepted between the "sensor" and the transmitter and it is also more expensive than the *optical* tracking system.

5.2 Interaction system

The interacting devices taking part on the system are: polarized glasses to see the virtual model using passive stereo and a tactile device for the forefinger.

As already introduced in the previous section the glasses (figure 1–9) and the tactile device (figure 1–10) enclose either infrared LEDs emitting infrared light which will be detected by the cameras (*optical tracking*) or "sensors" which detect the magnetic fields created by the transmitter (*magnetic tracking*).

For the glasses, the system detects the position and orientation of the user's head, and having this information the computers are able to interpret the head movements and show the user the proper point of view.

For the tactile device, there is also a detection of its position and movement, so the system is able to interpret those movements in some way. For example, we can decide this device allows the user to directly manipulate virtual objects (moving them from one place to another, pushing them, rotating them, etc.). Just by moving the forefinger as you would do in a natural way. Another example can be to manage the application menus just by pushing in the option with the forefinger. Furthermore, this tactile device is also a *haptic* device, allowing the contact simulation with a virtual object in the scene. This behavior is achieved by including a lever mechanism inside the device, When the system detects there is a collision between the device position in virtual space and the virtual object this lever is activated touching the finger, allowing the reinforcement of the coordination between viewing and hand movements.

6 APPLICATIONS

Several applications of the system explained above are already being used. The system has been tested, thus, by some medical doctors, architects and also CAD designers from the automovile area. All of them agreed on evaluating the system as a very useful and promising tool.

In medicine, the most important applications are planning and training of surgical operations [10]. So far, medical doctors usually contrasted 2D data obtained through different methods. The possibility of having a 3D representation of these data allows them a better understanding of the patient part to be analyzed.

Moreover, having also the needed tools to interact with this 3D model, the training sessions and planning of surgical operations become feasible in a virtual environment. After several discussions with some medical doctors of the Valle Hebrón Hospital in Barcelona, we can assure that the system is very useful in this area and that the best orientation of the screen for this sort of applications is the horizontal one (simulating a stretcher).

In the industrial area the most relevant application is the maintenance tasks training of industrial equipment. As an example, the system allows a worker to operate over several pieces of a complex equipment, an engine for example, in order to do assembly, substitution and repairing tasks, and at the same time the supervisor can observe what the worker is doing.

In the architecture, interior design and town planning areas, the main application is the possibility of experimenting with different distributions and configurations of the different elements of the building, furniture or urban scenery.

Our system is specially indicated for those applications requiring intensive 3D object manipulation. Of course, in case the interaction weight was higher for other aspects different from the manipulation, like navigation or model interrogation, a simpler visualization system would be more suitable.

The system allows to work with real scale objects if their dimensions do not surpass the screen dimensions. For applications requiring a real scale visualization of bigger models (for example the visual evaluation of the static aspects of a car body), systems like CAVE or powerwall (vertical screen of big dimensions) would be more appropriate.

Another limitation to take into account is a consequence of the visual field. The presented system is not appropriate for those applications requiring the presentation of a surrounding virtual world (like the evaluation of ergonomics aspects inside a car), because other systems like CAVE or cylindrical screens give a much wider visual field.

7 CONCLUSIONS

In this work we have presented an affordable immersive virtual reality system for 3D interaction. The interaction part of the system allows the direct manipulation of virtual objects in a 3D geometric model. The interaction consists of the projection of stereoscopic images on the user working space and the natural movements of the user. The system includes a screen, with dimensions comparable to an office desk, adjustable in orientation and height, a tracking system, which can be optical or magnetic, and a tactile device for the forefinger providing touching sense. The cost of this system is, compared to the existing similar ones, around 1/4 the price of them.

(a)　　　　　　　　　　　　　　(b)

Figure 3. System prototype in horizontal position (a) and almost vertical position (b)

From our experience until now, we can affirm the system is specially useful in those cases where the application requires an exhaustive inspection of the 3D model with a certain necessity of object manipulation, like for example scaling, translating or rotating them in a very intuitive way.

The last prototype of the system is presently being exhibited in the Virtual Reality Center in Barcelona, where is it being used in areas like architecture, medicine and industrial design.

8 Acknowledgements

References

[1] D. Bowman and L. Hodges. Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments. *The Journal of Visual Languages and Computing*, 10(1):37–53, 1999.

[2] D. Bowman and L. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, 1997.

[3] M. Mine, F.P. Brooks, and C Sequin. Moving objects in space: Exploiting proprioception in virtual-environment interaction. In *Proceedings of SIGGRAPH 97*, Los Angeles, CA.

[4] T.A. DeFanti, D.J. Sandin, and C. Cruz-Neira. A room with a view. *IEEE spectrum*, pages 30–33, 1993.

[5] C. Cruz-Neira, D.J. Sandin, and T.A. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Computer Graphics Proceedings of SIGGRAPH 93.*, pages 135–142, 1993.

[6] W. Kruger et al. The responsive workbench: A virtual work environment. *IEEE Computer*, July 1995.

[7] L. Rosenblum. Applications of the responsive workbench. *IEEE Computer Graphics and Applications*, 17(4):10–15, July-August 1997.

[8] I.E. Sutherland. A head-mounted three-dimensional display. In *AFIPS Conference Proceedings*, volume 33, pages 757–764, 1968.

[9] StereoGraphics Corporation. *Stereo-Graphics Developers' Handbook.*, 1997. http://www.stereographics.com.

[10] M. Montgomery and S. Brewster. Virtual reality for veterinary teaching. In *First International Workshop on Usability Design and Evaluation for Virtual Environments*, September 1998.