An Exploration of Immersive Virtual Environments

Martin Usoh and Mel Slater,

Department of Computer Science, and
London Parallel Applications Centre,
Queen Mary and Westfield College,
University of London,
Mile End Road,
London, E1 4NS, U.K.

Immersive Virtual Environment (IVE) technology has made it possible for computers to accommodate to humans in a way which is more natural, in terms of human behaviour, than most other human-computer interfaces. A virtual world allows human participants to perform tasks as "naturally" as they would do in everyday reality. Therefore no special learning of skills is required in order to carry out common place activities (such as picking something up) in such environments. To go through a door in the virtual world would be similar to the real world. This intuitive approach is a key factor which has moved virtual environments closer to the ultimate human-computer interface, allowing anyone to effectively communicate with the environments that computers display, irrespective of their age, experience, class, or cultural background.

Introduction

Steven Ellis of NASA Ames defined *virtualisation* as "the process by which a human viewer interprets a patterned sensory impression to be an extended object in an environment other than that in which it physically exists" [1]. In a virtual environment, the patterned sensory impressions are delivered to the senses of the human participant through computer generated displays (visual, auditory, tactile and kinesthetic). An ideal *immersive* virtual environment is one where the *totality* of inputs to the participant's senses are continually supplied by the computer generated displays, though in practice this is the case only for a subset of modalities (such as visual and auditory). In Ellis' terms, this provides the possibility of an egocentric frame of reference, where the self-representation of a participant in the virtual environment (VE) coincides with the position from which the virtual world is displayed and experienced.

In this description we emphasise the word "totality" for two reasons: first, ideally all senses should be involved. Second, however, we also mean that there is total immersion for each sense separately. Unlike the situation when looking at a screen or a film projection, when observers turn their head they must continue to perceive the computer generated environment - that is, it is all around them.

This idea of immersion in a virtual reality (VR) is not new. It was proposed by Ivan Sutherland in 1965 [2], and realised in 1968 [3] with a head-mounted display (HMD) that could present the wearer with a stereoscopic three dimensional view of a computer generated world. This was coupled to a six degree of freedom mechanical sensing device which tracked the wearer's head movements. This "ultimate display" presented simple wire frame models of the computer generated world.

With the recent increase in computing power it is now possible to cope with the processing demands of rather more sophisticated real-time interactive three dimensional environments. In such environments objects are represented as 3D geometric data together with information corresponding to the radiant and physical characteristics of their material properties. These underlying abstract representations are converted into the media of the human senses, primarily visual, but in principal also auditory, tactile and kinesthetic. Moreover, the sensory data is presented in a surrounding fashion, as mentioned above. Visual and auditory immersion may be achieved using HMDs, and to a degree tactile immersion with gloves and "body suits".

Such immersive virtual environments offer people the ability to carry out actions which may not be possible, or perhaps are too dangerous to do in everyday reality. Thus, on the one hand they provide a good training environment for hazardous occupations, and on the other offer great potential for many tasks ranging from complex data visualisation to entertainment and leisure.

In this paper we will discuss our research in the understanding and utilisation of VR. We give a brief account of the technology which has made the creation of virtual environments possible. We explain how this emerging field is significantly different from traditional interaction with computers through screens, keyboards, and pointing devices such as a mouse. We also present and discuss the research issues and examine the applications of IVEs along with the concept of *presence*, the fundamental feature of immersive systems. We take the standpoint that all interactions and behaviour in an IVE should not compromise the participant's sense of presence, that is their sense of being in the environment specified by the displays. We conclude by noting the importance of on-going research into the scientific understanding of human behaviour during a VR experience as a basis for constructing effective virtual environments.

Virtual Reality Technology

Todays head-mounted displays are generally helmet-like with two liquid crystal display TVs positioned in such a way that each is only visible to one eye - one for the left and one for the right [4], [5]. Therefore, as with normal vision a slightly different image is seen by each eye, and these stereoscopic views are integrated by the human visual system and interpreted as three dimensional (Figure 1). The position of the wearer's head is tracked using electromagnetic (wireless) constructions to determine the position and orientation of sensors on the HMD in relation to an external source [6].

In order to influence the computer generated world we also need a device which will allow us to interact with it. This device often comes in the form of a special glove or 3D mouse. The *dataglove* when worn allows the computer to track the wearer's hand and finger movements. From the position of the subject's hand in the real world the computer can determine whether the subject is touching an object in the virtual world. The 3D mouse is a similar interaction device which tracks the holder's hand and relays additional information to the virtual world through the subject pressing its buttons. A typical HMD and 3D mouse is shown in Figure 2.

More specialised gloves and devices are available which enable the wearer to actually feel a sensation corresponding to touching an object in the virtual reality. This may be done through pressure capsules located on the glove which are subsequently filled with air, or the subject placing their hand inside some form of exoskeleton. To further add to the realism of these environments advanced hardware can be used to generate directional sound that seem to come from outside the person's head - that is, from the virtual reality itself [7]. The interested reader is referred to [8] for a detailed account of current virtual reality technology and systems.

The Virtual World

The objects which inhabit a VE can be programmed to exhibit a number of physical properties and behaviours - for example, a virtual object will bounce when dropped, a virtual TV will show video pictures along with the appropriate sounds. This is achieved by the world being controlled by several independent computer modules which process information concurrently [9].

A graphics module is responsible for the conversion of the object's geometric representation into visual images on the graphics display. This module requires certain parameters of a virtual camera or eye such as the viewpoint and the gaze direction which are obtained from the computer tracking the person's head movements. It is then able to determine which objects will be visible and displays these using standard graphics algorithms. The bottleneck of most VR systems is the display rate of these images. It must be at least 25 image-frames per second to give the illusion of real-time movement and continuity.

Another module is responsible for determining collisions between virtual objects (including the geometric representation of the participant) by continually testing for intersection of geometric volumes of objects against each other. An intersection of these volumes signifies a collision and this information is relayed to other modules such as an audio module which will generate the appropriate sounds. Given a virtual representation of the participant's hand the system can make it possible for the person to grasp and pick up a virtual object when it detects a collision between the virtual object and the participant's virtual hand.

The Immersive Environment

In this paper we are strictly concerned with immersive virtual environments. These are virtual reality systems where the human participant becomes "immersed" in the world created by the computer and has an egocentric view of the world. We make this distinction from exocentric systems where the observer interacts with a virtual world presented on a separate screen. In this case the observer is not directly part of the VE in that the world is seen through a computer screen rather than from inside the environment. Note that we choose to adopt the term *participant* for the human who is immersed in the virtual world since, through immersion, they become part of that world.

Since the action of the participant turning their head to look at the objects around them in the virtual space gives the impression that the objects actually exist in that space, it follows that by looking down they should see their own body. This body, however, is not their physical body but is in fact a computer representation whose movements are correlated with the movements of their real body through information relayed by sensors worn by the participant. Therefore, just as there is a virtual environment, the environment formed by the computer system and displays, so there is a "virtual body" - the computer representation of the human participant and his or her activities. It is this computer representation of the human participant which is one of the distinguishing features of an immersive system. If the participant "believes" that the computer generated objects, along with the virtual body exist in the virtual space, and they associate the virtual body with their real body, then it follows that they believe that they are themselves *inside* the virtual space. Therefore immersion strictly means that a representation of the participant's body is placed in a surrounding VE.

The immersive virtual environment can also be shared simultaneously by several people, each one with a distinct virtual body. The participant is then free to meet, interact and form a working relationship with another person never met in real life. This colleague might in fact be miles away - the relationship taking place entirely in an imaginary space created by computers linked across a high speed network.

Applications of IVEs

Virtual Reality is a technology which is likely to have an important social impact, perhaps on the same scale as television. Its uses and applications are diverse and range from entertainment and leisure to training, design, simulations, and teleoperation (i.e. an operator having the ability, from a remote location, to work in a physical environment which may be dangerous or impractical for humans). Architectural walkthrough and design is an example of an IVE which can allow architects and their clients to experience a building, from the inside, before a single brick has been laid. Unlike looking at images of a building interior on a flat screen, we are able to enter and walk through the building and consider it from the inside. This is the major advantage that an IVE has over other traditional visualisation techniques: it allows a person to evaluate by direct experience whether a tool or environment is suitable for their requirements before it is built in reality. There wouldn't be much point in designing a fantastic looking building which people are unable to use.

Immersive virtual reality technology is useful for applications where the human participant must have a sensation of presence, a sense of "being there" in the worlds created by the computer. An obvious requirement is that it must be economical or safer for participants to experience virtual worlds rather than the corresponding real worlds. For example, learning to operate a nuclear power station might be an ideal future application of VR, since mistakes are rather less costly than mistakes in the corresponding real situation. Of course, not every possible application fits these requirements: virtual reality will never be a universal panacea for all aspects of human-computer interaction. The following areas are examples of potentially useful applications:

- walkthroughs - where potential clients can experience what it is like to be inside structures through immersion in virtual models.
- the construction of mock-ups, such as in the automobile construction industry - avoiding the costs and delays in building real physical mock-ups. Designers, their managers, and eventual customers can experience at first-hand what it would be like to be in the eventual real vehicle (Figure 3).
- virtual meetings over wide area networks.
- education and training, where trainees can experience and interact at "first hand" with representations of concepts, systems, ideas and cultures. Training in hazardous situations would be included (e.g. fire fighting).
- medicine and counselling, where disabled individuals can regain a sense of mobility through virtual exploration (for example, shopping), or counselling patients can meet their fears in controlled and safe environments.
- cultural acclimatisation, where potential visitors to a new culture can begin to have an immersed experience of being in that culture before the real visit.

Presence

Immersion can lead to presence, the participant's sense of "being there" in the virtual environment. The psychological sense of presence may be considered as an emergent property of an IVE. It is important to understand the factors that contribute to this, and a means of quantifying the concept of presence itself.

We have distinguished between two kinds of factors that contribute to the state of presence - external and internal [10]. External factors are those which are created by the virtual reality system or manifest themselves through the media. These factors have been discussed by [11]-[15]. Their exact nature and degree with which they can influence a participant's sense of being in the virtual environment is yet unclear. However, they include the extent to which the environment looks and behaves realistically. For example, the images generated by the system should be produced in real-time and be of high quality to induce a higher presence level. It should behave "correctly" in the context of the application. In architectural walkthrough a door should be opened by pulling on the handle, not by selecting the option "open" from some form of menu. Also, the environment should be consistent across all the displays and devices and exhibit minimal lag, e.g. when a virtual object falls to the ground the visual and aural cues should coincide at the point of collision.

Internal factors, however, determine the responses of different people to the same externally produced stimuli. These concern how the perceptions generated by the IVE are mediated through the mental models and representation systems that structure the participant's subjective experiences [16]. Generally a person who processes information in a visually dominant manner will experience the world in a different way than a person who is auditorially dominant. This can affect their degree of presence in the same virtual environment which has say, high quality graphics but low quality sound reproduction.

The Virtual Body

The sense of presence and degree of immersion are important features of IVEs. This sense of being somewhere other than where the real body is located, a kind of technological "out of the body experience", is fundamental to creating effective virtual environments.

To utilise the virtual environment to its maximum potential, the subject must suspend belief in the real world in order to allow themselves to be immersed in the virtual world. We said earlier that this is possible through the subjects' belief that the virtual objects exist in an extended space. If, however, they can identify with a computer representation of their real body in this extended space then they can believe themselves to be part of that space. This has led us, through communications with subjects experiencing immersive virtual spaces, to investigate the relation between presence and the virtual body [10].

We have some experimental evidence that the operation of the virtual body (VB) within the environment has a strong positive effect on the sense of presence for most people. Here we take *virtual body* as a self-representation of the human which responds to the movements of their own body. We have found that, on the whole, people who entered the virtual world with a VB reported a higher degree of presence than subjects who had a 3D cursor arrow in place of their hand as their self-representation. The degree with which the VB accurately reflects the movements of the subject also relates to the sense of presence.

The strength of the association which the participant makes with the virtual body is surprising considering its typical polygonised "block like" structure when compared to the participant's real body. However, we suspect that functionality of the VB is more important than its appearance - that is, the participant must quickly learn the relationship between his or her sensed movements and the observed movements of the VB. To this end the virtual hand was constructed with fingers that responded with the participant's real fingers to different button presses on the 3D mouse. Similarly the virtual hand was able to mimic all wrist movements of the the real hand. It is important that this part of the virtual body behaves correctly since when performing tasks in the virtual world the participants are more likely to have their virtual hand in view, for example in picking up an object, than any other part of the virtual body (Figure 4).

Some Research Issues

The current state of virtual reality systems has several limitations which must be overcome in order for it to become widely used as a tool for productive work. These research issues include the limitations of the technology as well as limitations of knowledge of human-computer interfaces in extended space. As stated earlier we take the standpoint that all interactions and behaviour in an IVE should not compromise the participant's sense of presence.

Hardware issues tend to be most evident during an immersive experience. The update rate of the graphics is usually the bottleneck of most systems and is usually of most importance. It can introduce lags between the actual movement of the participant in the real world and the reproduction of this movement in the virtual world. This lag is of prime concern in terms of human factors requirements. It can induce feelings similar to sea sickness due to the mismatch of information between actions of the participant and the results of these actions in the VE. This can be overcome by additional computing power. However, this may only be a temporary solution since the system will most likely be given more complex scenes which will again slow it down. A more permanent but less realistic solution is to work in less complex environments. At present we still do not know to what extent the quality and complexity of the graphics is necessary for a given application. The resolution of present head-mounted displays and trackers are also limiting factors. Current HMDs are generally of low resolution (e.g. 360 x 240 picture elements) and the electromagnetic trackers tend to be limited to a range of about two metres.

There has been numerous work in the area of 2D interfaces, e.g. window based desktop systems. In terms of human factors requirements however, research in interfacing with immersive environments has been limited. This area of work is of major importance since it may be used to overcome certain hardware issues. For example, since current tracking systems limit the participant in terms of range it means that they cannot simple walk around the work area unconstrained. It was through investigations into the human factors of matching proprioceptive data with sensory feedback and hence maintaining or enhancing the sense of presence [17] that led us to a solution for this hardware problem. We have derived a technique for navigation based on whole body motions termed the *virtual treadmill* [18]. The technique is an attempt to simulate body movements associated with walking and is suitable in applications where the participant is constrained to ground level, for example while exploring a virtual building. The participant "walks in place" to move through the virtual environment. This does not require any additional hardware. The technique works through the fact that while physically walking, or walking in place, a person has a specific pattern of body and head movements. For example, their head may move up and down, or from side to side. This pattern of head movement ranges between individuals. The change in position of the head can be obtained from the sensor mounted on top of the HMD. This information is fed into a neural network pattern recognisor which is able to determine whether or not the person is walking in place. When it determines such a motion the participant is transported along their line of sight in the VE (Figure 5). Hence, since the subjects are able to manoeuvre by walking in place some of the problems associated with the limited range of the sensors can be overcome.

Conclusions

A very important part of our research at QMW has been in the scientific understanding of what happens to people immersed in VEs. Our programme includes studies of the reactions of humans to the IVE experience, and the construction of a theory and practical model of the concept of "presence". Since presence is what is uniquely offered by such systems, we believe such a theory enhances the utility of our practical work. An important measure of the success of a system is the degree of "presence" that it induces - without such presence the application may as well have been based on the use of a standard computer screen, keyboard and mouse.

It is important to realise that virtual technology is very much in its infancy. It is therefore unrealistic to expect successful and long-lasting applications of virtual reality systems for commercial and industrial exploitation without a thorough understanding of what happens when people are immersed in these computer generated worlds. For example, we do not know today the extent to which the graphics must be "realistic", in the context of any particular application, in order for the participant to experience a strong sense of presence, or the effect on efficiency of task performance. Also there remain fundamentally unsolved and inherent problems with the technology for which viable solutions must be found.

Acknowledgements

The work described in this paper was funded by the U.K. Engineering and Physical Sciences Research Council (EPSRC), and Department of Trade and Industry, through grant CTA/2 of the London Parallel Applications Centre. Thanks to Anthony Steed and Anna Bara for their help with experimentation and to Marianna Gaston for commenting on earlier drafts of this paper.

Biographies

Martin Usoh (BSc, PhD) is a Research Assistant in the Department of Computer Science at Queen Mary and Westfield College, University of London. He has a background in Biochemistry and a PhD in Computer Vision for work in parallel processing for machine vision at QMW. He has been involved in virtual reality since 1992 initially in the context of architectural walkthroughs. He has investigated natural interfacing and presence in immersive systems and is currently working on the London Parallel Applications Centre's MIVE project (Modeling in Virtual Environments). Previous research involved SIMD processing in model-matching, raytracing and Artificial Intelligence.

Mel Slater (BSc, MA, MSc) is Reader in Computer Science, Queen Mary and Westfield College, University of London. He became Head of Department of Computer Science in October 1993. He was Visiting Professor in the Computer Science Division, University of California at Berkeley, for the spring semesters 1991 and 1992. His research and teaching areas are computer graphics and virtual reality. He is principal investigator of the London Parallel Applications Centre project, Modeling in Virtual Environments, and the EPSRC/ROPA project Distributed Extensible Virtual Reality Laboratory (joint with Universities of Lancaster and Nottingham). He leads the FIVE (Framework for Immersive Virtual Environments) Working Group, funded by ESPRIT Basic Research.

References

[1] Ellis, S.R. *Comp. Syst. Eng.* **2**(4), 321-347, 1991.

[2] Sutherland, I.E. *Proc. IFIPS Conf.*, **2**, 506-508, 1965.

[3] Sutherland, I.E. *Proc. Fall Joint Comp. Conf.*, **33**, 757-764, 1968.

[4] Teitel, M.A. *Proc. SPIE Stereoscopic Displays and Applications*, **1256**, 1990.

[5] Fisher, S., McGreevy, M., Humphries, J., and Robinett, W. *ACM Workshop 3D Int. Graph.*, 77-87, 1986.

[6] Ferrin, F.J. *In 'SPIE Proc. Large-Screen-Projection, Avionic, and Helmet-Mounted Displays'*, p. 1456, 1991.

[7] Wenzel, E. *Presence: Teleoperators and Virtual Environments*, **1**(1), 80-107, 1992.

[8] Kalawsky, R. *'The Science of Virtual Reality and Virtual Environments'*, Addison-Wesley, Reading, MA, 1993.

[9] Grimsdale, C. *In 'Proc. Computer Graphics '91 Conference'*, 1991.

[10] Slater, M. and Usoh, M. *Proc. IEEE Conf. VRAIS*, 90-96, 1993.

[11] Loomis, J.M. *Presence: Teleoperators and Virtual Environments*, **1**(1), 113-119, 1992.

[12] Sheridan, T.B. *Presence: Teleoperators and Virtual Environments*, **1**(1), 120-126, 1992.

[13] Zelter, D. *Presence: Teleoperators and Virtual Environments*, **1**(1), 127-132, 1992.

[14] Heeter, C. *Presence: Teleoperators and Virtual Environments*, **1**(2), 262-271, 1992.

[15] Steuer, J. *J. Comm.*, **42**(4), 73-93, 1992.

[16] Slater, M. and Usoh, M. *Presence: Teleoperators and Virtual Environments*, **2**(3), 221-233, 1994.

[17] Slater, M. and Usoh, M. *In 'Virtual Reality and Artificial Life'*, (Eds Thalmann, N. and Thalmann, D.), Prentice Hall, 1994.

[18] Slater, M., Steed, A. and Usoh, M. *In 'First Eurographics Workshop on Virtual Reality'*, (Ed Goebel, M.), 71-86, 1993.

Figures and Captions

Figure 1: A stereoscopic pair of left and rights eye images with a 10cm separation. When the images are viewed such that each is only visible to its respective eye the human visual system is able to integrate and interpret them as a single 3D scene. This can be achieved by partitioning the two images with a piece of card. It is also possible to obtain a stereo effect by crossing your eyes when viewing.

Figure 2: (a) A Virtual Research Flight Helmet™ showing the electromagnetic sensor which is mounted on the front of the HMD. The LCD displays mounted on the inside of the helmet supply the wearer with a stereoscopic view, and the headphones provide aural information from the VE. (b) The DIVISION 3D mouse (a hand held input device) contains a sensor which relays information about the holders hand position and orientation to the VR system. The buttons allow the participant to interact with the virtual world.

Figure 3: A participant testing the functionality of a virtual automobile. For example, when seated in the drivers seat are they able to reach all the instrument panels comfortably?

Figure 4: The viewpoint of a participant during a virtual table tennis training session. The virtual hand and arm come into view as their real hand is raised to hit the ball. The virtual hand responds correctly to all wrist movements of the real hand.

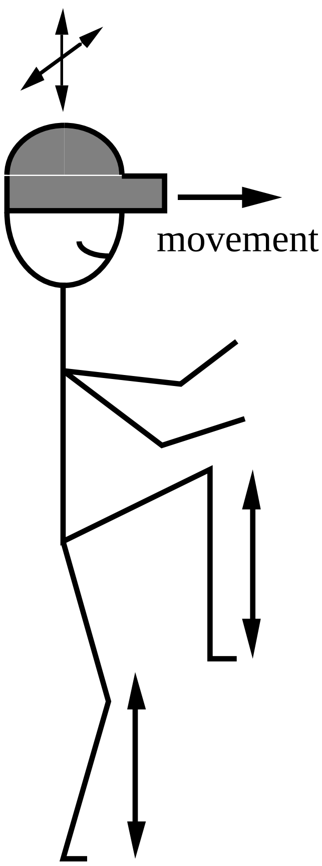

Figure 5: The virtual treadmill: the up and down movements of the subject walking in place are reflected through movements of the subject's head. The movements, up and down or from side to side, vary between individuals. This pattern of movement is recognised by the neural network program and the subject is moved in the VE along the direction 'movement'.