Guidelines for choosing VR Devices from Interaction Techniques

Jaime Ramírez
Computer Science School
Technical University of Madrid
Campus de Montegancedo. Boadilla del Monte. Madrid
Spain
http://decoroso.ls.fi.upm.es

Abstract: - This paper presents some guidelines to choose VR devices from the design of a virtual environment. This design is assumed to specify the interaction techniques that the users will employ to select and manipulate objects, and navigate in the virtual environment. Then, the proposed guidelines will pose different alternative devices for each interaction technique. In the case of the input devices, the devices will be grouped into two sets, desired devices and sufficient devices. While the desired devices represent the best option from the immersion point of view and, at the same time, the most expensive one, the sufficient devices stand for a cheap low quality solution.

Key-Words: - **Virtual Reality, Interaction techniques, VR devices, Software Engineering, Navigation, Selection, Manipulation**

1 Introduction

With the rapid increase in performance of high-end computer graphics systems and the transition of 3D graphics onto fast and inexpensive PC platforms, virtual environment (VE) interfaces have become feasible enough to be practically used in areas such as industrial design, data visualization, training, and others [1]. Development of useful VE applications, however, requires optimization of the most basic interactions, in particular object manipulation, so that users can focus on high-level tasks rather than on low level motor activities [2].

So far, some works have been published comparing different interaction techniques, sometimes drawing conclusions from experiments in VEs. However, if the state of the art in Virtual Reality (VR) technology is examined, a gap related to the election of VR devices according to the design specification of a VR system may be noticed. From our point of view, the correct election of the VR devices is a crucial issue in the final success of the VR system. Hence, it would be interesting that the VE designer can follow some guidelines that make easier this election. This paper pretends to be a first step in this direction.

In our approach, the proposed guidelines mainly take into account the interaction techniques chosen in the VE design. In addition, other issues such as the budget constraints and the desired degree of immersion are considered. The system performance issue, nevertheless, has not been considered, because it is very strongly coupled with the availability of powerful enough hardware resources, and the quality of the software. A study of the influence of these factors in the system performance is beyond the scope of this paper. Basically, as it will be shown, the proposed guidelines map the chosen interaction techniques into the most proper VR devices.

The organization of this paper is as follows: in section 2, some interaction techniques for selection/manipulation and navigation, drawn from the sate of the art in VE technology, are presented. Next, in section 3, the most representative VR devices are mentioned and classified. Then, in section 4, the guidelines to elect VR devices from interaction techniques are explained with special attention to the election of input devices. Finally, we end with some conclusions and future work.

2 Interaction Techniques

The fundamental forms of interaction between a human and a VE are selection/manipulation and positioning. Logically, an ideal scenario would be one in which the human actions in the real world have a perfect correspondence in the simulated VE. For example, it would be very natural if any movement of the human in the real world yields an equivalent movement of his/her avatar in the VE. However, the technology limitations (in the case of positioning, related to the maximum scope of the trackers), the budget constraints, or the impossibility of using a real scenario equivalent to the VE, make infeasible the aforementioned ideal scenario in many applications. Hence, in order to bridge the gap between the actions requested by the user and the simulation of these actions in the VE, as naturally as possible, some interaction techniques have been developed.

Next, two taxonomies of interaction techniques will be outlined: a taxonomy of manipulation techniques [3], and a taxonomy of positioning techniques [4].

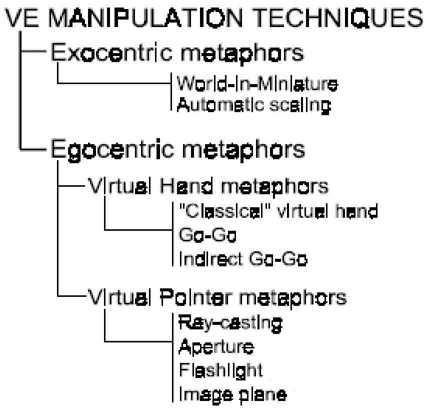

Fig.1: Manipulation techniques taxonomy

2.1 Taxonomy of Manipulation Techniques

This taxonomy is shown in the figure 1 (taken from [3]). Basically, the most common manipulation techniques have been classified according to their basic interaction metaphors. All the techniques can fall into either exocentric or egocentric techniques. These categories are used to distinguish between two fundamental frames of reference for user interaction with VEs. While under *exocentric* interaction (or God's eye viewpoint), users interact with VEs from outside the VE, under *egocentric* interaction, the user is interacting from inside the VE.

One of the considered exocentric techniques is the *World-In-Miniature* (WIM) technique [5]. This technique augments an immersive head tracked display with a hand-held miniature copy of the virtual environment. In addition to the first-person perspective offered by a virtual reality system, a WIM offers a second dynamic viewport onto the virtual environment. Objects may be directly manipulated either through the immersive viewport or through the three-dimensional viewport offered by the WIM.

Another considered exocentric technique is the *automatic scaling* [6], which scales down the world, when the user wants to select or manipulate a far object.

With egocentric interaction, essentially, there are two basic metaphors, virtual hand and virtual pointer.

Using the *virtual hand* metaphor, the user is equipped with a virtual hand whose position and orientation is controlled by a tracker attached to the user's real hand. In order to pick up a virtual object, the user intersects the object with the virtual hand and presses a button. This is basically the approach employed by the "classical" virtual hand technique. However, if the user wants to pick up an object that is out of his scope of reach, a way to reach it may consist on stretching the virtual arm so that the virtual hand can pick up these far objects. This technique has already been proposed, and it is called *Go-Go*. Here, a local area is defined around the user at some distance. While the user's hand stays within that physical distance, the virtual hand moves in a one-to-one correspondence with the physical hand. When the physical hand goes beyond the threshold, however, the virtual hand begins to move outward faster than the physical hand, following a non-linear increasing function. There exist variants of the Go-Go technique [7], for example the *Indirect Go-Go*. With the Indirect Go-Go, the length of the arm is controlled by a wheel device or a two-buttons device.

Using the *virtual pointer* metaphor, the user selects and manipulates objects by pointing at them. The simplest technique based on this metaphor is *ray casting* [8, 9], in which the selection of objects is carried out by pointing at them with an invisible infinite ray emanating form the user's hand. When the user wants to select an object, he/she points at it and presses a button. The ray casting allows the user to select easily objects at any distance, but it is not easy to use with occluded objects, small objects or with far objects, and it does not permit the manipulation with more than one degree of freedom (rotation in the axis of the ray). There exists another technique called *fishing reel* [10] that enhances the ray casting technique with traslation of the picked object towards or away from the user. There exist other techniques based on the virtual pointer metaphor that rely upon a cone pointer instead of a ray pointer, such as the aperture or the flashlight. The *aperture* technique [11] utilizes a cone with a variable size whose direction is determined from the position of the user's head, and whose size can be modified by moving the user's hand forward or backward. With aperture, all the objects included in the cone are selected, so if the user wants to focus on a particular object, he/she must reduce the size of the cone. The *flashlight* technique [12], on the other hand, employs a cone with a constant size, and if there are more than one object inside the cone, the elected one is the closest object to the axis of the cone. Another technique based on the virtual pointer metaphor is the *Image Plane* technique [13]. In this technique, the user can interact with the 2D projections of the objects. In this way, the manipulation is limited to be 4DOF.

So far, the interaction techniques mentioned in the taxonomy of [3] have been explained briefly. Now, another three interaction techniques will be considered, two of which combine the virtual hand metaphor and the virtual pointer metaphor, HOMER and Voodoo Doll, and the third one is voice recognition.

The *HOMER* technique [14] combines ray casting with the virtual hand metaphor. With HOMER, the selection is accomplished by using a ray, and then, upon selecting the object, the ray converts into a hand, so that the manipulation can be performed. In addition to the drawbacks inherited from the ray casting, this technique has the drawback of not offering an easy manipulation of far objects. The *Voodoo Dolls* technique [15] allows a user to manipulate objects at a distance by creating miniature copies of objects or dolls. These dolls have different properties and affect the objects they represent differently when held in the user's right or left hand. The doll in the left hand provides a stationary frame of reference for the right hand to work in. This simplifies working relative to moving objects, and allow for working at multiple scales without explicitly resizing objects or changing modes. Moreover, both visible and occluded objects can be manipulated using this technique by either creating the doll directly or grabbing it from a previously defined context.

The *Voice Recognition* technique may fall into the group of techniques based on the exocentric metaphor or on the egocentric metaphor. In any case, this technique assumes the computer can recognize oral sentences including pre-defined actions and names of objects that are in the VE. Some of the parameters of the pre-defined actions are set automatically to the default value. As a result, the interaction is simpler, but the precision of the manipulations is lower. Moreover, the user is required to know the names of all the objects he/she is going to interact with.

2.2 Taxonomy of Positioning Techniques

There are two key parameters which must be specified to fully define the user's movement through the VE: *speed* and *direction of motion*. Next, different techniques to specify these parameters will be outlined.

2.2.1 Techniques for specifying direction of motion

There are different techniques to specify the direction of motion [4]: Hand directed, Gaze directed, Dynamic scaling, Physical controls, Virtual controls, Object driven and Goal driven.

With the *Hand directed* technique, the position and orientation of the user's hand determine the direction of motion. There exist two variations or modes of this technique: pointing mode and crosshairs mode. In *pointing mode*, the direction of motion through the virtual space depends upon the current orientation of the user's hand or hand held input device. The drawback of this mode is that it can be confusing for novices. For that reason, the crosshairs mode was developed. In the *crosshairs mode*, the user simply positions the cursor (typically attached to the user's hand) so that it visually lies on top of the object that he/she wishes to fly towards.

If the direction of motion is *Gaze directed*, the user will fly following the direction that the user is looking.

Another technique to specify a movement in the VE is called *Dynamic scaling*. This technique consists of scaling down the world until the desired destination is within reach; then, moving the centre of scaling (the location in three-space that all objects move away from when scaling up and move towards when scaling down) to the desired destination; and finally, scaling the world back up again.

Physical controls (joysticks, mice, etc.) or *Virtual controls* (virtual buttons, virtual steering wheels, etc.) can also govern the movements of the user through the VE. Both of them result in unnatural interaction, mostly when using virtual controls, due to the lack of haptic feedback.

The movement of the user in the VE is *Object driven* if it can be induced by virtual objects (for example, an elevator) included in the VE. In this way, when the user situates his/her avatar in one of them, the next movement of the user's avatar is determined by the driven object.

Another technique that can be used to move is the *Goal driven* technique, where the user chooses one destination point from a map or a list of accessible points or naming the destination point by using the voice. After that, the user is automatically moved to the chosen point.

2.2.2 Techniques for specifying speed

Several options for the specification of speed of motion can be used [4]: Constant speed, Constant acceleration, Hand controlled, Physical controls and Virtual controls. While with Constant speed, the speed is the same during the whole virtual session, with Constant acceleration, in contrast, the speed grows exponentially with movement duration. The use of hand position as a throttling mechanism is an adaptable form of speed control. In addition to control direction of motion, Physical controls, Virtual controls, and Voice Recognition can also be employed to modify the speed.

3 VR devices

Two kinds of VR devices will be considered, input devices and output devices.

3.1 Input Devices
Input devices can fall into one of these two categories: immersive devices and desktop devices.
The immersive devices are devices that help to produce a feeling of immersion in the user. Immersion is the feeling of being deeply engaged. Some examples of immersive input devices are: tracker, glove (data-glove or pinch-glove), wand (and variations of it like hornet, dragonfly, and mike), mechanical arm, etc.
In the other hand, the desktop devices come from multimedia environments, and although they do not produce the same feeling of immersion as the immersive devices, they are cheaper and the user is more familiarized with their handling. Some examples of desktop input devices are: mouse (desktop mouse or 3D mouse), joystick, keyboard, trackball, etc.

3.2 Output Devices
Within this group, two subgroups will be distinguished too: immersive devices and desktop devices.
The immersive devices can be visual or haptic. Within the immersive visual devices we can find: projection systems like CAVE (Cave Automatic Virtual Environment) or Powerwall, BOOM (Binocular Omni-Orientation Monitor), and Head Mounted Displays (HMDs). Moreover, Mechanical arms and some models of gloves are immersive haptic devices.
The desktop devices can also be visual or haptic. The desktop visual devices are desktop monitor and stereo glasses, whereas the desktop haptic devices are some models of joysticks, mice and pens.

4 Choosing VR devices

4.1 Choosing Output Devices
The election of output devices does not depend on the interaction techniques chosen in the design of the VE, since no interaction technique requires a certain visual device, and haptic feedback can only be considered as desirable, but never as mandatory. Some interaction techniques, like aperture selection or gaze directed positioning, seem to need a HMD. Nevertheless, they do only require a tracker attached to the user’s head rather than a HMD carrying on a tracker.
Besides, some combinations of input and output devices may be possible, yet uncomfortable. For instance, the combination of desktop monitor plus glove, because the monitor may mean an obstacle for the user’s hand movements. Another example of an uncomfortable combination of devices would be a HMD plus physical controls (mice, keyboards, etc.), as the HMD would not allow the user to see the physical controls in front of him.
Thus, the designer of the VE should choose the output devices taking into account the desired degree of immersion and the budget constraints, and should avoid some undesirable combinations of input-output devices. The output devices may be grouped into three levels according to their provided degree of immersion, and at the same time, their price:

1. CAVE
2. Other projection systems (Curved screen, Powerwall, etc.), BOOM, HMD, haptic glove, haptic mechanical arm.
3. Desktop output devices (monitor, haptic joystick, haptic pen, etc.).

4.2 Choosing Input Devices
In this section, some input devices will be associated with each interaction technique shown in section 2. Except for some cases in which the interaction technique forces to use a certain input device, two subsets of input devices will be distinguished, desired devices and sufficient devices, for each interaction technique. While the desired devices represent the best option from the immersion point of view and, at the same time, the most expensive one, the sufficient devices stand for a cheap low quality solution.
Further on, we will refer to 6DOF devices as desktop input devices that permit the interaction with an object in 3D, for example, joysticks, 3D mice and keyboards. Moreover, a 4DOF device will be a desktop input device that permits the interaction with an object in 2D, for example, a desktop mouse or a keyboard.

4.2.1 Input devices for manipulation techniques
When using techniques based on the exocentric metaphor, the user must be able to select small objects in a miniature scenario. For that purpose, some desired devices would be a glove, a pen and a mechanical arm, all of them equipped with a tracker and a button device. However, if manipulation of objects is also required, the glove will be the most natural option among the desired devices, since it will provide the most natural form of manipulation. Within the sufficient devices, any 6DOF device fulfills the requirements of these techniques, as long as it can be employed to move a cursor in the miniature VE, as well as to manipulate an object in a complex manner.

All the techniques based on the virtual hand metaphor require a glove, therefore it does not make sense to consider other alternative devices for these techniques.
Within the group of the virtual pointer techniques, the ray casting and fishing reel techniques must be supported by devices that allow the user to point at an object in the VE. For doing that, the desired devices will be a glove, a pen or a wand, equipped with a button and a tracker. Additionally, in the case of the fishing reel technique, a wheel device or a pair of buttons will be necessary to move backward or forward the selected object. In default of some of the desired devices, any of the 6DOF devices can be used as a sufficient device.
Aperture and flashlight techniques require two controls, one for specifying the direction of the conic pointer, and another for varying the size of the cone. Ideally, different combinations may be used for managing these two controls:

- HMD (with tracker) + glove (with tracker),
- HMD (with tracker) + wheel device,
- glove (with tracker) + wheel device, and
- wand (with tracker) + wheel device.

The wheel device may be replaced with a two-buttons device. Again, the 6DOF devices can be used as sufficient devices with aperture and flashlight. In general, aperture and flashlight techniques are more suitable than ray casting when the tracking system has a significant jitter [16].
With the image plane technique, the user must be able to interact with the 2D projections of the objects. The authors of this technique present it so that two tracked gloves are employed in selection, manipulation and navigation operations. However, a tracked pen or a tracked wand with a button device can also be used in order to accomplish the same operations, with a similar immersion sensation as using the two gloves. In default of some of these devices, any 4DOF device can be considered a sufficient device.
The HOMER technique, as we have seen before, combines the ray casting technique with the virtual hand metaphor. For that reason, the unique valid device would be a tracked glove accompanied by a button device.
The Voodoo Doll technique is intended, as stated by its authors, to be performed by using two pinch gloves with a tracker in the index finger and another one in the thumb. Given the operation of this technique, other alternative devices are difficult to imagine for this technique.
Finally, if the voice recognition technique is employed, the user must carry a microphone, so that he/she can order actions for selecting or manipulating objects.

4.2.2 Input devices for navigation techniques

Using the hand directed technique with pointing mode, the user must carry either a tracked glove or a tracked wand. However, under crosshairs mode, the best option would be a tracked HMD, though other cheaper alternatives (sufficient devices) would be possible, which would consist of using a 6DOF device (or a 4DOF device if the user is not allowed to go up or down in the VE) to specify the destination point.
In the dynamic scaling technique, the alternative devices are the same as the alternative devices in the techniques based on the exocentric metaphor.
In the gaze directed technique, it is only possible to use a tracked HMD.
With physical controls the options are clearly 4DOF or 6DOF devices, whereas with virtual controls or object driven, the chosen devices will depend on the manipulation methods specified in the design for working with objects in the VE.
If the navigation technique is goal driven, the possible devices to be chosen will depend on the manner of representing the list of destinations. If the different destinations are represented by means of widgets, desktop devices must be employed. Nevertheless, if the list of destinations is represented somehow as part of the VE, that is, as virtual controls, the chosen devices will be determined by the chosen manipulation techniques, as it has already been explained previously.

5 Conclusions and Future Work

This paper has reviewed the most relevant interaction techniques for VEs, and has outlined the current technology in VR devices. After that, a correspondence between interaction techniques and VR devices has been established. In this correspondence, each interaction technique has been related to a set of suitable VR devices, in which, if possible, two subsets have been distinguished, desired devices and sufficient devices. To the best of our knowledge, no other paper has proposed a similar correspondence. This correspondence pretends to be a guide for the VE designer, so that he/she can elect the VR devices from the interaction techniques he/she has chosen. Moreover, this guide may be used in the opposite way, so that the designer may restrict the set of all the existing interaction techniques to a subset of applicable techniques taking into account the VR devices that are already available in the organization, or may be available in the future.

We think the proposed guide may be improved by extending the information on devices with software and hardware requirements, and ranges of market prices, so that the VE designer can work out an approximation to the cost of implementing some interaction techniques using a certain combination of devices. This information would provide a more exact criterion to compare different interaction possibilities.

References:

[1] M. Göbel, Industrial applications of VEs, *IEEE Computer Graphics & Applications*, Vol.16, No. 1, 1996, pp. 10-13.

[2] K. Stanney, Realizing the full potential of virtual reality: human factors issues that could stand in the way, *Proceedings of VRAIS'95*, 1995, pp. 28-34.

[3] I. Poupyrev, S. Weghorst, M. Billinghurst and T. Ichikawa, Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques, *Computer Graphics Forum*, Vol.17, No.3, 1998, pp. 41-52.

[4] M. Mine, *Virtual environment interaction techniques*, UNC Chapel Hill CS Dept., Technical Report TR95-018, 1995.

[5] R. Stoakley, M. Conway and R. Pausch, Virtual reality on a WIM: interactive worlds in miniature, *Proceedings of CHI'95*, 1995, pp. 265-272.

[6] M. Mine, F. Brooks and C. Sequin, Moving objects in space: exploiting proprioception in virtual environment interaction, *Proceedings of SIG-GRAPH'97*, 1997, pp. 19-26.

[7] I. Poupyrev, M. Billinghurst, S. Weghorst and T. Ichikawa, Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR, *Proceedings of UIST'96*, 1996, pp. 79-80.

[8] R. Bolt, Put-that-there: voice and gesture at the graphics interface, *Computer Graphics*, Vol.14, No.3, 1980, pp. 262-270.

[9] R. Jacoby, M. Ferneau and J. Humphries, Gestural Interaction in a Virtual Environment, *Proceedings of Stereoscopic Display and Virtual Reality Systems: The Engineering Reality of Virtual Reality*, 1994, pp. 355-364.

[10] D. Bowman and L. Hodges, An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments, *Proceedings of Symposium on Interactive 3D Graphics*, 1997, pp. 35-38.

[11] A. Forsberg, K. Herndon and R. Zeleznik, Aperture based selection for immersive virtual environment, *Proceedings of UIST'96*, 1996, pp. 95-96.

[12] J. Liang, JDCAD: A Highly Interactive 3D Modelling System, *Computers and Graphics*, Vol.18, No.4, 1994, pp. 499-506.

[13] J. Pierce, A. Forsberg, M. Conway, S. Hong, R. Zeleznik and M. Mine, Image plane interaction techniques in 3D immersive environments, *Proceedings of Symposium on Interactive 3D Graphics,* 1997.

[14] D. Bowman and L. Hodges, An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments, *1997 Symposium on Interactive 3D Graphics,* 1997, pp. 35-38.

[15] J. Pierce, B. Stearns, R. Pausch, Voodoo Dolls: seamless interaction at multiple scales in virtual environments, *Proceedings of the 1999 Symposium on Interactive 3D Graphics*, 1999, pp. 141-145.

[16] A. Steed, C. Parker, 3D Selection Strategies for Head Tracked and Non-Head Tracked Operation of Spatially Immersive Displays, *Proceedings of the 8th International Immersive Projection Technology Workshop*, 2004.