3D Interaction in Virtual Worlds

Daniel del Pozo Nieto

12.3.2001

University of Joensuu
Department of Computer Science
Master's Thesis

ABSTRACT

Virtual environments constitute a new human-computer interaction paradigm where users are not simply external observers of images on a computer screen but are active participants with a computer-generated three dimensional virtual world. As virtual environment applications become more complex, it is clear that a better understanding of the principles of virtual environment interaction is needed and designers need a guidance in choosing three dimensional interaction techniques. Also virtual environments are supposed to be systems that will enhance the communication between humans and computers. If virtual systems are to be effective and well received by their users, considerable human factors research needs to be accomplished.

This paper is a literature survey which gives some solutions and a guidance of how to achieve these goals. First a set of taxonomies of interaction techniques are presented. According to these taxonomies, some current navigation, selection and manipulation techniques in which the interaction techniques are split into, are described. Also, this paper presents various types of interaction devices and visual displays which also play an important role in the success of virtual environment systems.

This paper provides an overview of many of the human factors issues that influence the acceptance of virtual reality systems, including interaction techniques usability; human performance efficiency in virtual worlds, which is likely influenced by tasks characteristics, user characteristics, human sensory and motor physiology, multimodal interaction, and the potential need for new design metaphors; health and safety issues of which cybersickness and deleterious physiological aftereffects may pose the most concern; and the social impact of the technology.

ACM-classes (ACM Computing Classification System, 1998 version): I.3.0, H.1.2, H.5.2

Contents

List of Abbreviations

DOF Degrees of Freedom: the number of ways a body can move in space, including all translations and rotations.

FOV Field of View: the viewer's visual field when using a visual display.

HCI Human-Computer Interaction: the exchange of information between human beings and computers during a task sequence for the purpose of controlling the computer (from the point of view of the human) or informing the user (from the point of view of the computer). This interaction usually has the goal of increasing human productivity, satisfaction, or ability.

HMD Head Mounted Display: a computer graphics display that is worn on the head of the user, so that the display graphics are continuously in front of the eyes of the user. HMD may use cathode ray tube (CRT) or liquid crystal display (LCD) technology, and usually incorporate optical lenses to widen the displayed image and move it farther from the user's eyes. Many HMDs include headphones for audio, and most are used in conjunction with trackers (a device that measures 3D position, and sometimes orientation, relative to some known source).

HVEI Human Virtual Environment Interaction: involves HCI but while HCI is generally from an exocentric frame of reference, HVEI is generally from an egocentric perspective.

IT Interaction Technique: a method by which the user performs a task on a computer via the UI. An IT may be as simple as clicking a mouse, or as complex as a series of gestures. There may be many possible ITs for any given interaction task. The IT may be influenced by the input device used, but is separate from it. The same input device may be used for many ITs for the same task. Conversely, it may be possible to implement a given IT using several different input devices.

IVE Immersive Virtual Environment: a VE is described as immersive when the computer generated environment appears to enclose or surround the user, and when the parts of the physical world that are not integral system components are blocked from view.

UI User Interface: the hardware and software that mediate the interaction between humans and computers. The UI includes input and output devices such as keyboards, gloves, visual displays, and speakers, as well as software entities such as menus, windows, toolbars, or 3D widgets.

VE Virtual Environment: a three dimensional (3D) model of a space displayed to a human user from an egocentric point of view using real time 3D computer graphics. Motion and point of view orientation are generally controlled by the user, not by the system. VEs often include other sensory information, such as auditory or haptic cues.

VR Virtual Reality: the experience of being within a VE.

1 Introduction

Virtual reality (VR) is the experience of being within a three dimensional (3D) model of a space displayed to a human user from an egocentric point of view using real time 3D computer graphics. Motion and point of view orientation are generally controlled by the user and not by the system. Virtual interfaces include multi sensory transfer of information involving visual, auditory, tactile, kinesthetic, and occasionally even olfactory cues. For its potential advantages VR is believed to be beneficial in many applications such as military, education, 3D CAD, architectural design, space travel, medicine, entertainment, and industry (Bowman, 1999).

Virtual reality now offers to exploit the human motor capabilities more fully (Coomans, 1999). Some new devices make it possible to track 3D hand and head gestures, while others track the 3D forces that are performed on them. This enables us to build user interfaces (UI) that let a user input data by moving objects, connecting objects to each other, moulding objects, and so on. Of course, the success of such an innovative interaction paradigm stands or falls with what is usually called the interaction metaphor. This is the mapping between the application's semantics and the user's actions.

This paper does not discuss tracking techniques for user's strength and limb position measures which are more technological issues, but it considers a set of taxonomies for better designing of interaction techniques (IT). Some current techniques that rely on few interaction metaphors are described. For example, the Go-Go immersive IT (Poupyrev & al., 1996) uses the metaphor of interactively growing the user's virtual arm and non linear mapping for reaching and manipulating distant objects.

Traditionally, computers use input devices ranging from switches and punched cards to keyboards and mice. However, it is obvious that such devices are insufficient for interacting with the virtual world. For example, how would the simple action of drinking a cup of water using a mouse be expressed? Issuing the command with the keyboard is a cumbersome solution since things like which cup to drink from or how much to drink must be specified. This paper examines various types of input devices that have been developed for use with VR.

Ongoing developments in information technology have led to an increasing demand for 3D displays and techniques for enhancing the visualization of information have commanded the attention of researchers (Stanney 1997). These methods have generally been confined to the flat 2D surfaces (e.g., movie screens or the computer monitor). Recently, however, progress has been made toward overcoming this limitation by introducing the third dimension of vision. Typically, visual displays that capture this third dimension of visual experience appear to the viewer as greatly enhanced in their vividness, clarity, realism, and depth (Morris, 1997). Even though there are some technological failures (e.g., lag of rendering time or graphic resolution for the 3D graphic scene) that are potential causes of visual strain and limit the comfortable viewable depth volume, this paper does not try to address these technological issues, but describes various types of visual displays.

Although the ability of virtual reality to immerse a user in a virtual world (through the use of 3D real time computer graphics and advanced displays devices such as head mounted displays or caves) is beneficial for many applications (e.g., education, entertainment, simulations), research needs to focus efforts on addressing a number of human factors issues if VR systems are to be effective and well received by their users.

Designing usable and effective interactive virtual worlds is a new challenge for system developers and human factors specialists. Due to the close bond between the user and the system within the virtual environment (VE), it may be impossible to follow past traditions and suggest human factors from design issues when trying to enhance the potential of VE technology (Stanney & al., 1998a). It is the capabilities and limitations of the user that many times will determine the effectiveness of virtual worlds (e.g., what type of metaphors will enhance the user's performance in VEs, which user characteristics will influence VE performance, will the user be adversely affected by VE exposure, will the user perceive system limitations as flicker or lag, etc.). An understanding of human factors issues can thus be used to provide a systematic basis by which to direct future VE research aimed to better meet the needs of its users. In this thesis, some of the human factors that affect the usabiltity of the VE technology are discussed.

Chapter 2 goes through few interaction techniques taxonomies and describes how some navigation, selection, and manipulation techniques are used. Then, current interaction and visual display devices used in VE systems are presented. Chapter 3, considers several human factors including technology limitations, human sensory issues and different VE design aspects, human performance efficiency, health issues, and the social impact of the VR technology. Finally, Chapter 4 contains conclusions.

2 Interaction Techniques

First, in order to understand the techniques and their effects, and their design and evaluation, a taxonomy for interaction tasks is established. The taxonomy categorizes techniques and partitions the design space into more easily understandable pieces. Then, the taxonomy will be used to design new techniques, by combining one component for each of the lowest level substask from the taxonomy. Then, some specific navigation, selection, and manipulation techniques in which IT techniques are divided into, are described. Finally, some interaction devices and visual displays, which can be used in conjunction with the interaction techniques to interact with the system, are also described.

2.1 Taxonomy

Virtual reality (VR) opens a door to new manners of human-machine interaction that exploit the possibilities offered by presence simulation creating a palpable, concrete illusion for users to manipulate real, physical objects positioned into the VE (Gobbetti and Scateni, 1998). This capability to manipulate visual objects interactively, rather than simply viewing a passive environment is desirable in many VR applications and is typically accomplished using a real-world metaphor: the user selects an object by touching it with his virtual hand, and manipulates it directly by moving his hand using natural physical motions (Bowman and Hodges, 1997). Although intuitive, this metaphor does not consider the fact that in VEs users may perform actions which are impossible in the real world (Poupyrev, 1997). For example using some VR interaction techniques users can manipulate objects located far outside normal human reach.

Analysis of current manipulation techniques (Ware and Jessome, 1988; Erickson, 1990) suggest that most of them are based on a few interaction metaphors. Each of these basic metaphors forms the fundamental mental model of a technique —a perceptual manifestation of what users can do, how they can do it, and what they can not do when using the technique. Poupyrev & al. (1998) present a simple classification of current manipulation techniques according to the basic interaction metaphors these techniques are based on (Figure 1). They divide the whole variety into exocentric and egocentric techniques. With the exocentric interaction (so called God's eye viewpoint) users interact with VEs from the outside. Examples are World-In-Miniature and Automatic scaling techniques, which will be described later.

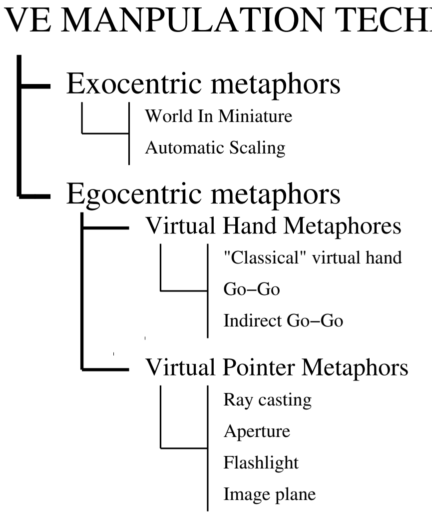

Figure 1: Classification of VE manipulation techniques depending on their underlying metaphors (Poupyrev & al., 1998).

With egocentric interaction, which is the most common for immersive VEs (IVE), the user is interacting from the inside of the environment, i.e., the VE embeds the user (Wicknes and Baker, 1995). Currently there are two basic metaphors for egocentric manipulation: virtual hand and virtual pointer (Mine, 1995; Bowman and Hodges 1997; Pierce & al., 1997). With the *virtual hand*, users can grab and position objects by touching and picking them with a virtual representation of their real hand. The outstanding design factors that define a particular technique are the choice of input devices and mappings between the real and virtual hand's position and orientation. For example a "classical" virtual hand technique provides ono-to-one mapping between the real and virtual hands, while the Go-

Go technique employs non-linear mapping functions to extend the user's area of reach (Poupyrev & al., 1996).

With *virtual pointer* metaphor, the user selects and manipulates objects by pointing at them. When the vector emanating from the virtual pointer intersects with an object, it can be picked an manipulated (Mine, 1995). Design aspects that distinguish techniques based on this metaphor are definition of virtual pointer direction, shape of the pointer, and methods of disambiguating the object the user wants to select. In the basic case, the direction of the virtual pointer is defined by the orientation of the virtual hand in the virtual scene, the pointer is a laser ray, and no disambiguation is provided (Jacoby, 1994). Some techniques define the direction of the virtual pointer by using two points: position of the user's dominant eye and location of the tracker manipulated by the user (Forsberg and Zeleznik, 1996, 1997). Other techniques use volumetric pointers to facilitate object selections.

These basic and most general metaphors can be subdivided to reflect particular aspects of each technique. Also, techniques based on different metaphors can be combined together to form new manipulation techniques.

Many studies (e.g., Nielsen, 1993) intuitively suggest that interaction in VEs should replicate our interaction with the physical world. Bowman and Hodges (1999) claim that such interaction is never completely realistic so that VEs should instead enhance the physical, cognitive, and perceptual capabilities of users, allowing them to do things that are impossible in the real world.

There are multitude of different techniques and tasks (desired user actions) and many of them are domain-specific. For better analysis and understanding of interaction techniques Bowman and Hodges (1999) reduce the space by recognizing that there are a few basic interaction "building blocks" that most complex VE interactions are composed of. These basic interactions are divided into three task categories: viewpoint motion control, selection, and manipulation.

Viewpoint motion control, or travel, refers to a task in which the user interactively positions and orients his viewpoint within the environment, in other words, refers to the movement of one's viewpoint between different locations in the VE. Travel is part of the larger task of navigation, which includes the actual movement and the decision process involved in determining the desired direction and target of travel (wayfinding).

Bowman and Hodges's analysis of this task identified three basic components that must be included in any travel technique (Figure 2): direction/target selection, velocity/acceleration selection, and conditions of inputs. Direction/target selection is the mean by which the user indicates the direction of motion or the end point of motion. Examples of this are gaze-directed steering, pointing steering, gesture steering, discrete selection (e.g., menus) or 2D pointing. Velocity/acceleration selection is the way in which the user indicates the speed and acceleration of the motion, like constant velocity/acceleration, gestured-based, explicit selection (e.g., discrete, continuous range), user or environment scaling, and automatic or adaptive velocity/acceleration. Conditions of input are the means by which the user begins, continues, and ends the motion (e.g., constant travel/no input, continuous input, start and stop inputs or automatic start or stop).

Selection refers to the act of specifying or choosing an object for some purpose. *Manipulation* is the task of setting the position and orientation (and possibly other features such as scale or shape) of a selected object. Manipulation requires a selection technique, but the converse is not always true, selection techniques can be used alone for tasks for other purposes such as choosing a menu item or deleting an object.

Intuitively a set of techniques for these interactions is the real-world metaphor, where the user selects an object by making the gesture of touching it, in the virtual scene, with his virtual hand and manipulates it directly by moving his hand. This is cognitively simple, but limited practically. For example, if a virtual object is too large to allow easy placement, or if the object is not within arm's reach. Therefore, designs should focus on techniques that allow selection and manipulation at a distance. The entire task of selection and manipulation is broken down initially

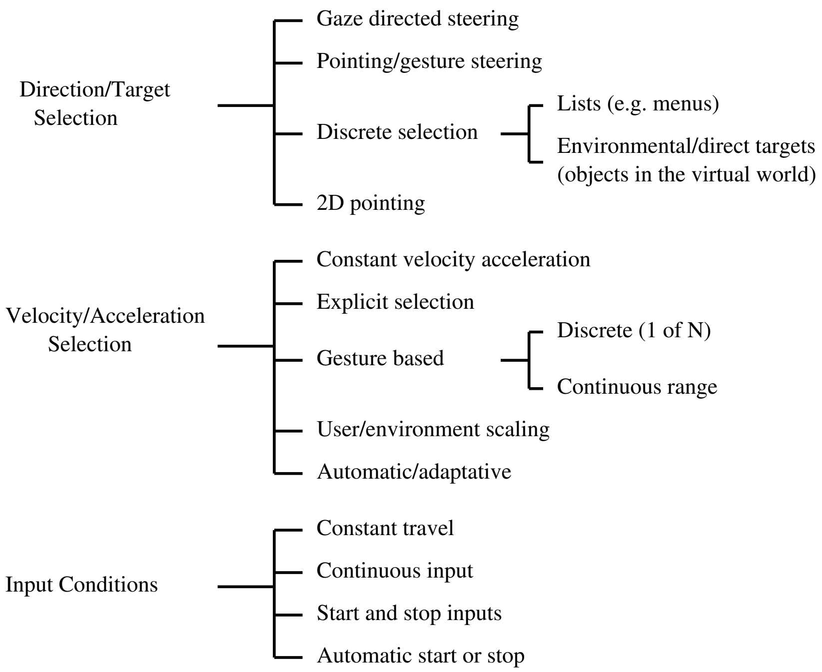

Figure 2: Taxonomy of travel techniques for IVE (Bowman and Hodges, 1999).

into three substasks: selection, manipulation, and release (Figure 3). These subtasks are further divided. Selection is divided into feedback, indication of object, and indication to select. Manipulation splits in object attachment (hand, gaze,...), object position, object orientation, and feedback. Finally, release is decomposed in indication to drop (gesture, button,...) and object final location.

Similar to this approach, Poupyrev & al. (1997) suggest that the VR manipulation tasks are the same tasks that we perform in the real world when we make positioning movements. We make positioning and orientation movements every time we reach for or move something to another location. These movements are a combination of reaching or grabbing, moving, and orienting of objects. According to this, a division of the manipulation task into several manipulation tasks, consistent with the taxonomy of interaction tasks suggested by Foley (1992), is: position (the task of positioning an object), selection (the task of identifying an object), orient (the task of orienting an object), text (the input of a string of characters), and quantify (the input of a numerical value).

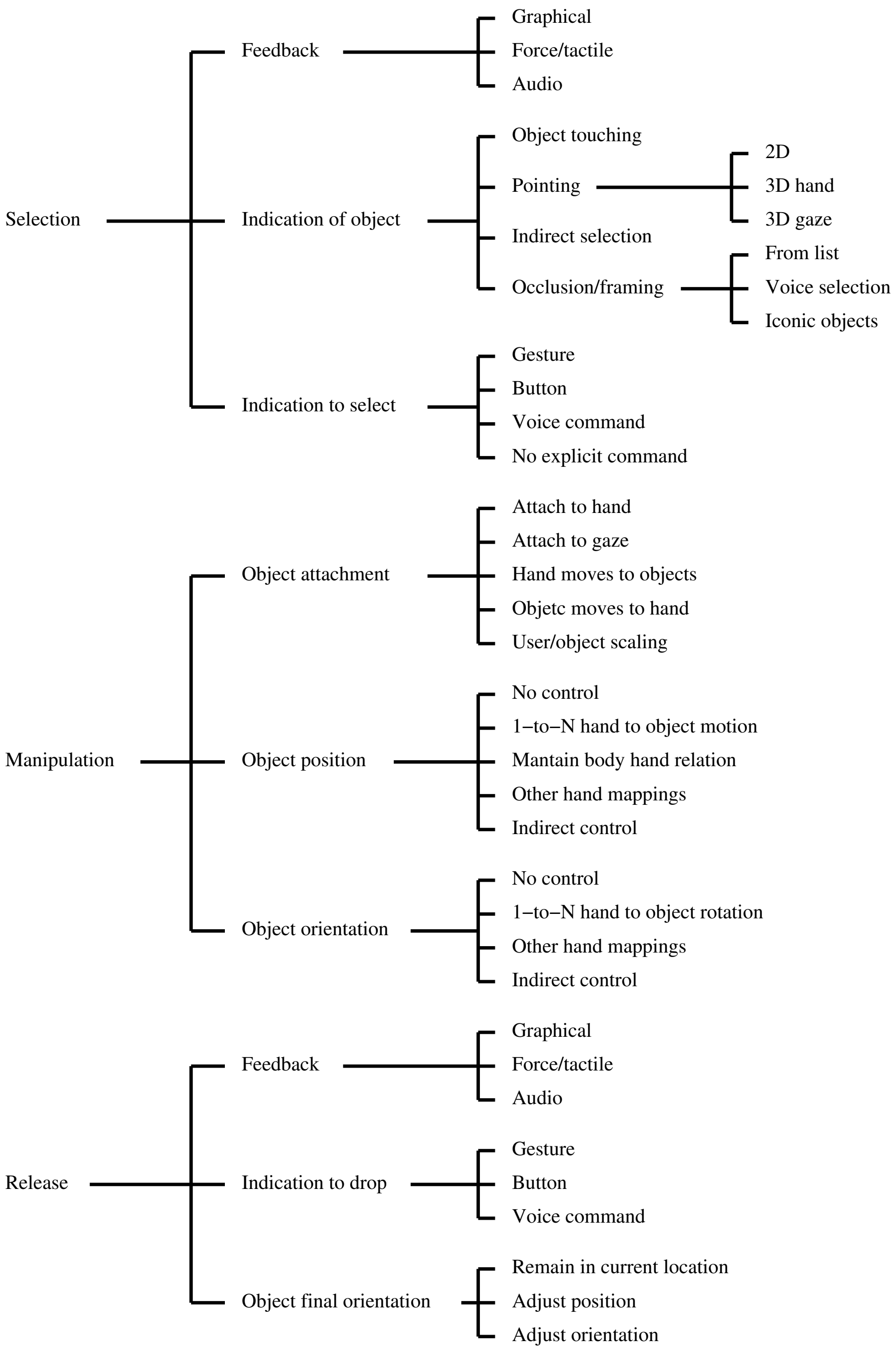

Figure 3: Selection and manipulation taxonomy (Bowman and Hodges, 1999).

Continuing with Bowman's and Hodges's approach, a fourth interaction task could be system control which comprises other commands that the user gives to accomplish work within the application (e.g., delete the selected object, save the current location, load a new model), but at a low level, this task can be compound by selection and/or manipulation tasks.

For each of these universal interaction tasks, there are many proposed interaction techniques. For example, one technique could accomplish a selection technique in a very indirect way, by choosing an entry from a list of selectable objects. Other technique could use a direct technique, where the user moves his (tracked) virtual hand so that it touches the virtual object to be selected.

2.2 Selection and Manipulation Techniques

Selecting and manipulating virtual objects is an important user interaction for VE. The human hand is an ideal manipulation input device for VR because it allows natural and intuitive interaction with virtual worlds. However, using the hand as an input device, users can not interact directly with objects out of their reach. This section presents various techniques that allow manipulation of remote objects.

World in Miniature (WIM) (Stoakley & al., 1995). This technique provides the user with a hand-held miniature copy of the virtual environment. This copy is a representation of the life-size world and serves as an additional viewpoint of the virtual environment (Figure 4) and enables the user to interact with the environment by direct manipulation through both the WIM and the life-size world. Moving an object on the model the user moves the corresponding life-size representation of that object and vice versa. Because WIM is a hand-held model, it allows the user quickly establish different points of view by rotating the WIM in his both hands.

If the IVE is large enough, there will be objects that are out of physical arm's reach. With WIM the user has the choice of selecting objects either by pointing

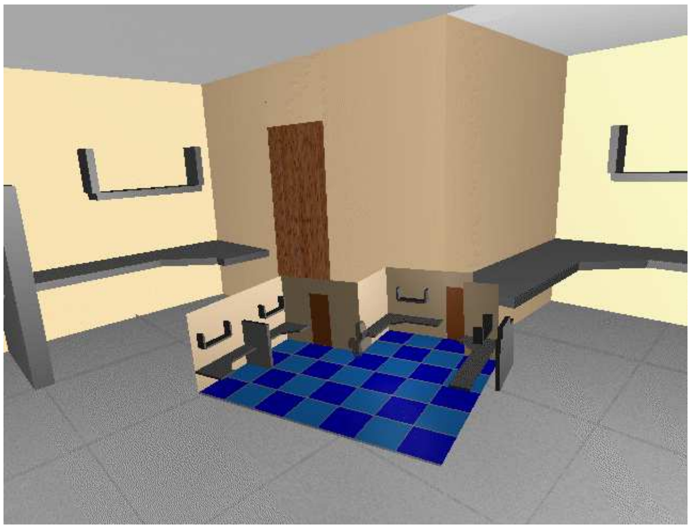

Figure 4: A miniature copy of the room where the user stands and a series of common furniture pieces (Stoakley & al., 1995).

to the object itself or by pointing to its proxy in the WIM. By turning the model in his hands, the user can view and pick objects that are obscured by his current line of sight from the immersive (camera) viewpoint. The WIM provides a second point of view (usually "bird's eye") from which to observe the world.

The user manipulates objects by grasping them as he would do with any other (graphical) object in the virtual scene. As long as the user continues the grasping gesture, the WIM follows the position and orientation of the user's hand and when the user releases, the object remains in the space wherever it is dropped. Due to the lack of haptic feedback rotations become difficult and to assist user's manipulation the technique provides a "clipboard" and a "tennis ball" (Figure 5). The clipboard allows users to rotate the WIM using a two-handed movement passing the clipboard quickly from one hand to other and back when the rotation is greater than can be done comfortably with one hand. The tennis ball has installed two buttons. By pressing one button users can select objects while the other button is for application specific manipulation actions. If the scale of the WIM is smaller than that of the immersive world, manipulating objects on the WIM gives the user far-reaching coarse-grained control of objects. The WIM can also display objects at a greater one-to-one scale (in which most of the world is out of the user's immediate reach) giving the user fine grain control of objects at the expense of range. The user can choose the most appropriate scale for any given subtask, and even switch scales in the middle to accomplish the requirements of the task.

Figure 5: A user in action holding the clipboard and the tennis ball (Stoakley & al., 1995).

This technique can consider more than one active instance of a WIM at a time. This could be used to display: widely separated regions of the same environment, several totally different environments, worlds at different scales, the same world with different filters applied to the representation, or the same world displayed at different points in time. These multiple views into the virtual world allow the user to visually compare different scales or different locations. Selecting a WIM will cause the immersive environment to change (jump) to that particular world.

Aperture (Forsberg & al., 1996). Aperture uses a conic selection volume, i.e., all the objects within the volume could be selected. The vertex of this conic volume (the from point) is set to the location of the user's dominant eye and the direction vector of the cone is the vector from that eye through the tracker's location (represented in the VE by a cursor). The aperture cursor is a circle of fixed radius and crosshair that is aligned with the film plane (image displayed on the screen) (Figure 6). The size of the selection volume is determined by the distance between the eye point and the aperture cursor. A user can adjust the scope of the selection volume by moving the hand in and out, with hand movement changing the distance. With this technique, even using a disambiguation metric, to select a single object among the multiple candidates on a set of closely spaced objects, is difficult.

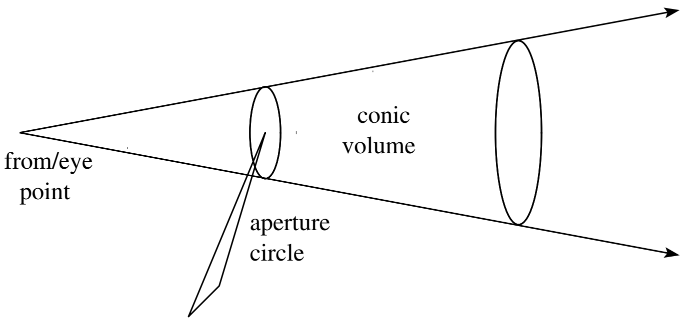

Figure 6: The conic volume of the aperture selection techniques described by the eyepoint and aperture geometry (Forsberg & al., 1996).

In the real world, we orient our hands to match the orientation of a target object in some way before manipulating it. An extended technique, called *aperture with orientation*, uses this 3D orientation information for selecting objects. If multiple objects fall into the conic volume, the object or objects, whose orientation most nearly match(es) the orientation of the tracker is/are selected. Orientation information provides a primary disambiguation metric and if various candidate objects have similar orientations, a single object is reselected with a basic disambiguation metric.

Go-Go (Poupyrev, & al., 1996). The technique is based on the arm-extension metaphor which means that the user's reach is expanded allowing him to grasp far away objects by making the virtual arm longer than the user's physical arm, i.e., with Go-Go Technique[1] users can manipulate both close objects and those at a distance (Figure 7) unlike traditional techniques (Figure 8). In the Go-Go technique a local area is defined around the user at some distance. While the

[1] The name is due to the cartoon character Go-Go Gadget, who had a powerful ability to extend his arms much further than normal.

user's hand stays within that physical distance, the virtual hand moves in a one to one correspondence with the physical hand (linear mapping), i.e., the virtual hand matches real hand movement, thus allowing fine manipulations to be made in normal manner. When the physical hand exceeds the defined threshold the virtual hand begins to grow longer than the physical hand, following a non-linear function. Then the user selects an object by touching it with his virtual hand and manipulates it with hand movements (real-world metaphor). A boundary box drawn around an object provides visual feedback when the object is selected. To return to the normal manipulation the user just brings the hand within the threshold distance.

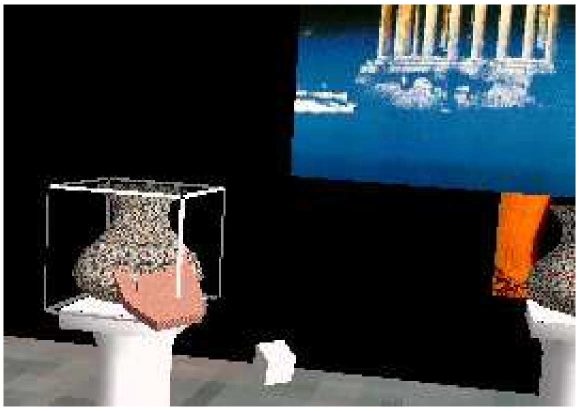

Figure 7: A user selecting a vase out of his arm's reach using Go-Go technique. The white cube shows the real hand position (Poupyrev, & al., 1996).

Two extensions to this technique (Bowman and Hodges 1997) are *fast go-go* which has no local area and a more quickly growing function, and *stretch go-go* which allows infinite stretching, so that, any arm length can be obtained permitting the arm to be placed at any distance.

Ray Casting (Bowman and Hodges 1997). Ray casting makes use of a virtual light ray which emanates from the user's virtual hand (Figure 9). The ray's direction is specified by the hand of the user. To select an object, the user intersects the object with the light ray and performs a grab action (button pressing). Releasing the button the object will be attached to the end of the ray and the user can then

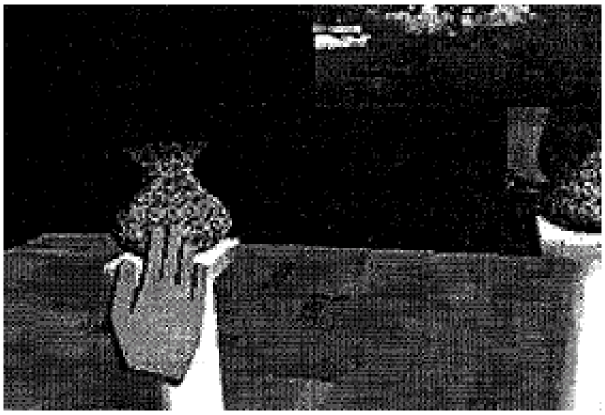

Figure 8: With traditional VR manipulation techniques the user cannot reach the vase; reach is limited by the arm length (Poupyrev, & al., 1996).

manipulate the object using the ray and moving his wrist or arm.

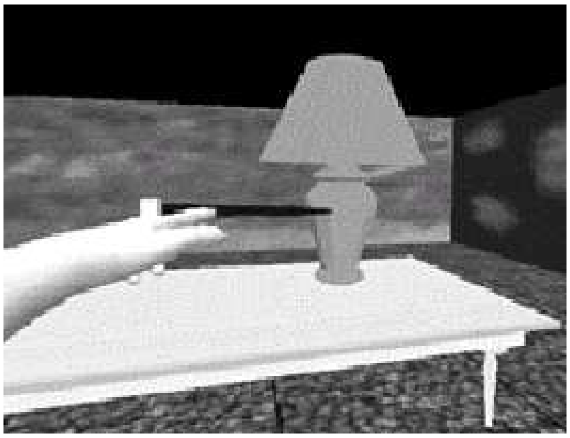

Figure 9: Grabbing an object via Ray Casting. The figure shows the virtual light emanating from the user's virtual hand (Bowman and Hodges 1997).

Manipulation is not hand-centered, the object is attached to the light ray instead of directly to the user's virtual hand. This makes it more difficult to manipulate the object: there is not a simple method for rotating an object, except around the axis of the ray itself, and no method for controlling the object's distance from the user.

Scaled World (Mine & al., 1997). This technique uses an automatic scaling mechanism as a way to allow the user to interact instantly with objects placed at any distance in the VE. Selected objects far of the user's reach are brought instantly into reach every time the user grabs an object by automatically scaling down the world about the user. For example, if the user's arm is extended 0.5 meters, the selected object that is 5 meters away is brought to the user's hand by scaling down the world by a factor of 10 (Figure 10).

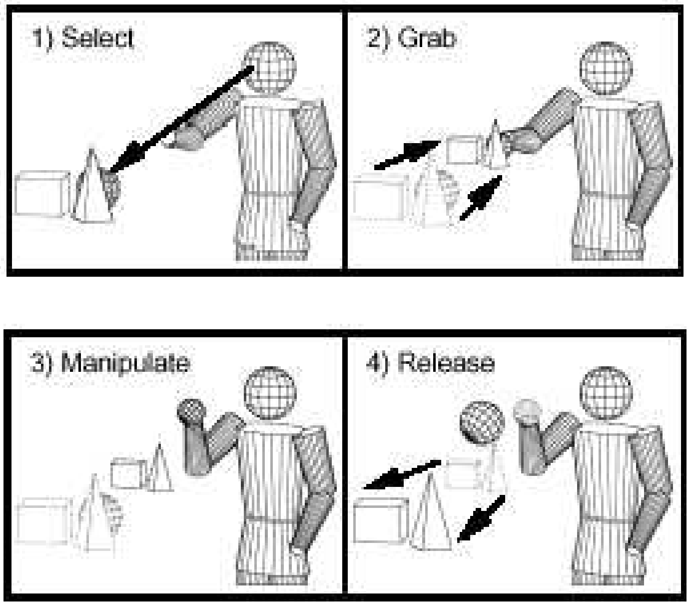

Figure 10: Automatic scaling of the world when the user grabs and releases an object (Mine & al., 1997).

With scaled world the user is able to bring the most remote object in the scene to his side in a single operation, the user does not need to "fly" to reach it or repeatedly grab, drop, and regrab to obtain the object. Once the object is released the world is scaled back up.

Image Plane (Pierce & al., 1997). With image plane interaction techniques the user interacts with the 2D projections that 3D objects in the scene make on his image plane (image displayed on a large 2D screen). Using these techniques a user can select an object by positioning his hand in the 3D scene so that the projection of it on his image plane is positioned in some way around the projected image of the object. The different ways of positioning the hands gives name to different techniques. These techniques need to find the position of the user's hand on his image plane and use this point as a 2D cursor to cast a pick ray into the scene.

Four of these image plane interaction techniques are: Head Crusher, Sticky Finger, Lifting Palm, and Framing Hands. In the *Head Crusher* technique the user positions his thumb and index finger above and below the desired object in the 2D image. Figure 11 shows the user selecting a chair in the scene using this technique.

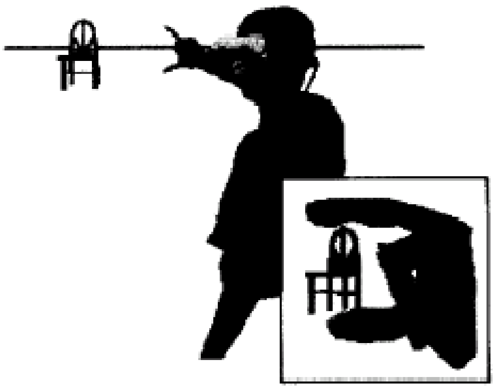

Figure 11: A third person point of view of the Head Crusher Technique. The inset shows the first person view (Pierce & al., 1997).

The *Sticky Finger* technique provides a very easy gesture when picking a very large or close objects by using a single out stretched finger to select objects. The object underneath the user's finger in the 2D image is the object that will be selected. In Figure 12 the user selecting a distant object using the Sticky Finger is shown.

With *Lifting Palm* technique the user selects an object by stretching his flattened hand out and positioning his palm so that it appears to lie below the desired object on his image plane (Figure 13).

Framing Hands technique uses both hands to select the objects. The user must position his hand to form two corners of a frame in the 2D image. The user then positions this frame surrounding the object to be selected (Figure 14). Using this

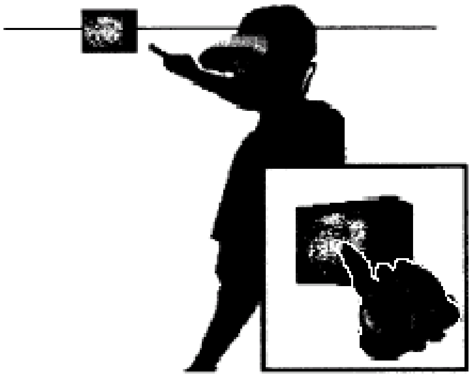

Figure 12: A third person point of view of the Sticky Finger Technique. The inset shows the first person view (Pierce & al., 1997).

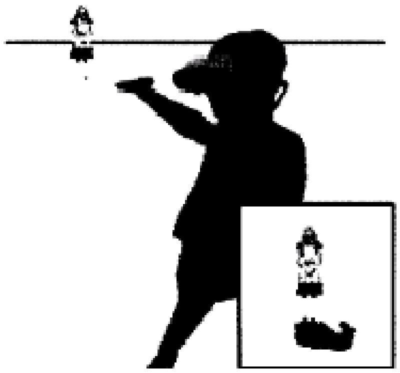

Figure 13: A third person point of view of the Lifting Palm Technique. The inset shows the first person view (Pierce & al., 1997).

technique the user can select several objects by selecting all of the objects that fall into the frame formed by the user's hands. The user can arbitrarily rotate or resize the frame with a motion of his hands.

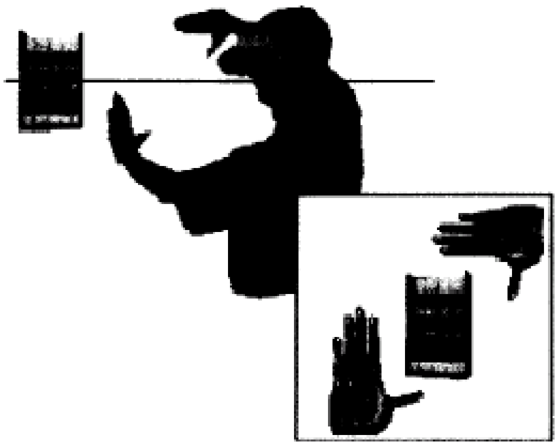

Figure 14: A third person point of view of the Framing Hands Technique. The inset shows the first person view (Pierce & al., 1997).

The user can confirm that he has correctly positioned his hand for a desired object to select, before issuing a selection command, using the feedback provided by the system. This feedback is given by highlighting or showing the bounding box of the object that is currently candidate for selection. These techniques also provide a disambiguation metric to identify a single object when there are several candidates. The metric is the same as used in the aperture technique, the object with the closest matching orientation in the user's image plane can be chosen. The user's finger(s) provide this orientation information for the head crusher, sticky finger, and framing hands techniques. The normal to the user's palm gives the disambiguating orientation for the lifting palm technique.

To allow users to manipulate objects beyond their arm's reach the techniques should provide mechanisms that move a distant object to them. One option is to instantaneously translate the desired object to the selection point and to scale the object down so that the object projection in the image plane remains unchanged. This might leave the user with a too small selected object for manipulating conveniently. To address this problem the object could be resized automatically to a suitable working volume. The object could also be resized dynamically based on the distance from the user's hand to his hand's natural working position and as the user moves his hand to this position the object scales to a convenient working size. Another alternative is to leave the object small allowing the user to scale the object explicitly by making a two handed gesture indicating the desired working size.

Rather than translating just the selected object, the entire world can be scaled down and translated so that the selected object is moved to selection point. This approach brings the world within reach so that the user can directly manipulate selected objects. When the object is released, the world scales back and translates to its original size and position.

2.3 Navigation Techniques

Some of the selection and manipulation techniques we described before can be used for navigating or traveling within a VE. This section describes how these techniques are used to travel and also presents some new travel techniques.

World in Miniature (WIM) (Stoakley & al., 1995). This technique includes a representation of the user as an object in the scene, so the user can simply reach into the WIM and pick himself up (select) to change his location in the environment. WIM technique uses three means of when to update the immersive world as objects (e.g., the user) in the WIM are manipulated. Usually the real and proxy object are moved simultaneously (*immediate update).* When occlusion occurs *post-factor update* is used, where the immersive environment updates only after the user is finished with the WIM interaction and has released the proxy object. With *batch update* the user makes several changes to the WIM and then issues an explicit command to cause the immersive environment to change to the current location in the WIM. Before the user issues the commitment command, he has two independent views of the environment, the "as-is" picture in the immersive world and the "proposed" picture in the WIM.

Scaled World (Mine & al., 1997). Here the user transports himself by grabbing an object in the desired travel direction and pulling himself towards it. With scaled world the user can reach any visible destination in a single grab operation.

Since the point of interest is attached to user's hand he can quickly view it from all sides by simply torquing his wrist. The user can also swing himself about the point of interest by holding it in front of his face while he turns around and the world pivots about his hand.

Two-handed Flying (Mine & al., 1997). The direction of the flight is specified by the vector between the user's two hands, and the speed of the travel is proportional to the user's hand separation. A minimum hand separation (e.g., 0.1 meters) enables users to stop their current motion quickly by bringing their hands together. Flying backwards requires the user to simply swap his hands.

Image Plane (Pierce & al., 1997). With image plane technique the travel is relative to a selected object. A constraint is placed on the positions of the user's hand and selected object in the 2D image so that they remain in the same position relative to each other on the image plane. For example, if the user selects a distant object with the head crusher technique (described in Section 2.2) the constraint will keep the object between the user's fingers in the 2D image.

The distance of the user from the object can be hold constant to let the user orbit the object. This allows him to quickly change his point of view around the object by simply moving his hands. Also user's distance from the object can vary as a function of the user's hand motion. For example, if the user moves his hand half a distance to his eye point he moves half a distance to the selected object.

Flying Chair (Bowman and Hodges, 1998). This technique does not allow the user move freely inside the environment. The user sits in a physical chair with his feet above the ground so that the illusion of no support can be maintained.

With this technique the user sees a virtual tablet on which different menu items (locations) are presented. By selecting a widget on the tablet the user can travel being automatically transported (via flying chair) to the selected location within the environment. The motion is smoothly animated in a straight line from the current location to the destination, and the virtual chair rotates to face the most interesting view at the new position. Views and locations in the menu are chosen so that traveling does not cause flying through solid objects.

With this technique users may feel restricted, since they are only allowed to travel to a set of discrete locations presented on the tablet, but if these predefined positions are wisely chosen, users will be able to visit each place and obtain views of all the interesting parts of the environment.

System-automated (Bowman, & al., 1999). The system automated technique gives a user no control over his path, so that the technique makes the user a simple passenger. The system moves the user from the current location to the destination placing him all the time about the center of the path in the scene.

In this technique translation of the user's viewpoint is virtual (the user cannot change his point of view in the virtual scene by moving his physical body), but the user's heading is controlled by physical head rotations. A map of the scene could be available before world traversal for specifying the path through the scene, as well as a velocity control feature with which the user could speed up or slow down his travel movement.

In System-automated techniques travel is accomplished using a tracked stylus (a tracker receiver embedded inside a pen). Users see a virtual representation of the stylus in the VE that moves in synchronization with the physical stylus so that the direction of pointing can be visualized. When the user wants to begin moving in the predefined path through the scene, he simply presses a stylus button. Even though the system controls translation of user's viewpoint, he is still allowed to look in any direction while moving. At any moment during the travel the user can click a button to stop moving and could travel through solid objects.

Two extensions to System-automated technique are *Route-planning* and *Pointing*. With Route-planning the users set a path before moving, and then the system moves them along the path. Pointing technique allows users to continuously specify the direction of the motion by pointing to that direction, then the system automatically moves the user to the pointed direction so that the user is allowed to "drive".

These two techniques also use a stylus for traveling in the IVE. For Route-planning technique, the users use the stylus to place markers on the 3D map of the scene to define the path. Once the user has placed the markers on the map, he clicks a stylus button and starts moving from the current position in a straight line to the first marker, from there to the second, and so on. With pointing technique the user presses a stylus button to start moving and then points the stylus in the desired direction of travel. In these techniques the user is also free to look in any direction while being translated by the system.

2.4 Interaction Devices

To interact effectively with a 3D user interface, input devices with high degree of freedom (DOF) are needed. DOF refers to how many ways a body can move in the 3D space, including all translation and rotations (Bowman and Hodges, 1997). For example a two DOF device senses X-axis and Y-axis displacements (MacKenzie, 1997). Six is the minimum number of degrees of freedom required to simulate the motion of a free floating body in the virtual space. A six DOF motion system can translate along and rotate around three axes: it can translate east-west, norht-south, and up-down, and it can also rotate about those same axes. The translations and rotations are independent of each other. This section describes various input devices with different DOF.

Two DOF Input Devices. 3D Mice and Joysticks are the most common input devices with two DOF. They are sufficient for navigating around a simple virtual world and for performing simple tasks by using the buttons on the devices. *3D Mice* (Perry & al., 1997) are intuitive to use, the user pushes the mouse in the direction he wants to move. Mice have buttons the user presses for selecting. However, mice are not very useful for interaction other than navigation and selection of objects. Mice usually have two DOF, although there are mice designed with six DOF. A *joystick* (Perry & al., 1997) basically works the same as a conventional joystick, but it is not fixed to a base that sits on a table top. Instead, the joystick is equipped with an orientation tracker so the user simply holds it in the air and tilts it. Joysticks also have buttons on the stick for clicking and selecting, similar to a mouse.

Six DOF Input Devices. Wands, force balls, gloves and dexterous hand master are various types of six DOF interaction devices that have been developed for use with VR. A *Wand (Perry & al., 1997)* is like a joystick with an unrestrictive base. Moving the wand in the space moves a 3D pointer in the VE. Often wands include buttons that allow the user to pick up objects in the virtual world, thus user's main actions when using a wand involve pointing and clicking with the wand.

A variation of the wand is the *Virtual Stick* (Rorke & al., 1998) designed to replace more complicated input devices, like Datagloves, while still providing the same usability. A virtual stick is a device comprised of a small handle with four switches built into it. The stick is held in the user's hand with a button assigned to each finger. The four buttons are used to represent the state of the user's fingers to the system, i.e., different combinations of on/off states of the buttons, representing different combinations of open and closed fingers on the user, are used to implement simple gesture recognition in the system. For example, one gesture consist of all the buttons being pressed, i.e, all the fingers of the user's hand closed around the stick. Using this gesture, the user is able to pick up and manipulate objects in the world. Other usual and important gesture in VEs is the pointging gesture represented by closing all the fingers, except the index finger.

Force (Space) Balls (Perry & al., 1997) enable the user to pan, zoom, and rotate 3D models with simultaneous six DOF control. With a spaceball (Figure 15) the user can easily move or rotate his 3D environment by pushing, pulling, or twisting the ball, but the ball cannot be moved. The model then moves in the direction of the force or twist applied.

The spaceball 3D motion controller provides a smooth and dynamic model manipulation, i.e., the greater pressure applied to the ball, the faster the model moves or rotates. The force ball is not needed to be hold in midair, reducing hand and arm fatigue.

Figure 15: Spaceball, a 3D interactive motion control device.

With *Gloves* (Perry & al., 1997) users are able to point, grab and manipulate objects in the virtual world as they would do in the natural world. To achieve this purpose gloves must capture the movement (finger flex and extension) and location (position and orientation) of the hand. When the user moves the hand, the glove picks up the movement and sends an electrical signal to the computer which represents the hand gesture within the virtual space where the user often is able to see his virtual hand. Different versions of gloves (e.g., *Datagloves or Powergloves*) differ on the method they use to perform the movement capturation functions. Gloves need readjusting and recalibration for different users.

A *Dexterous Hand Master (DHM)* (Perry & al. 1997) is an exoskeleton that attaches to the fingers with velcro straps. The differences with gloves is that DHM is able to measure side to side movement of fingers and it takes into account the fact that the human finger has three sections, rather than two, which gloves do measure. DHM has six DOF, however, there are DHMs developed with ten DOF.

2.5 Types of Visual Displays

In most applications of VR, visual feedback is required, in fact, visual cues are perhaps the most important feedback in the VR system. To achieve a sense of reality, the pictures sent to the display have to be real time to avoid discontinuity, i.e., visual displays used in VR context should guarantee continued updating of the visual display to reflect the user's movement through the environment (Perry & al., 1997). In this section different types of visual displays are presented.

Liquid Crystal Display (LCD) Flicker Lens (Perry & al., 1997). The LCD flicker lens has the appearance of a pair of glasses. A photosensor is mounted on these LCD glasses for reading a signal from the computer. This signal tells the glasses whether to allow light to pass through the left or right lens. When the light is allowed to pass through the left lens, the computer screen will show a left eye scene, which corresponds to what the user will see through his left eye. When light passes through the right lens, the scene on the computer screen is a slightly offset version of the view from the left eye.

LCD flicker lenses are cordless, but the user has to stare at the computer screen in order to see the 3D scene. Since the field of view is limited to the computer screen size, the user can notice the surrounding real world environment minimizing the immersive effect.

Head-mounted Display (HMD) (Perry & al., 1997). HMDs place a screen in front of both viewer's eyes at all times. The virtual environment view is controlled by orientation sensors mounted on the "helmet". Head movement is recognized by the computer, and a new perspective of the scene is generated. HMDs allow free head movement without loosing screen contact. The cable connecting the HMD and the computer restricts the user to move as far as the cable length is.

As real world surroundings are occluded from sight (visor closed) HMDs provide a feeling of total immersion in the scene displayed.

Binocular Omni-orientation Monitor (BOOM) (Perry & al., 1997). The BOOM is mounted on a jointed mechanical arm with tracking sensors located at the joints. A counterbalance is used to support the monitor, so that when the user releases the monitor, it remains in place. The viewer regards the stereo images as if looking through binoculars. To view the VE, the user must take the monitor, move it and put his face up to it. The computer will generate an appropriate scene based on the position and orientation of the joints on the mechanical arm.

Using BOOM the user is released from the encumbrance of wearing an HMD, allowing the user to switch between the virtual world and the real world by simply moving his eyes away from the BOOM.

Virtual Retinal Display (VRD) (Emerson, 1999). A VRD device actually consist of a very small (2cm x 1cm x 1cm) polished mirror placed on a mount of the size of conventional eyeglasses. This display device creates images by scanning modulated laser light directly on the retina of viewer's eye.

In contrast to a conventional display, with VRD no real image is produced, instead, a very small spot is focused on the user's retina. Large field of view with bright and high resolution images (close to that of human vision) are generated to be readily seen in ambient room (and even daylight). The image colors correspond exactly to what the viewer is seeing.

For 3D viewing an image is projected into both of user's eyes from a slightly different view point to create a stereo pair. With VRD it is also possible to vary the focus of of each pixel in the image such that a true 3D image is created.

CAVES (Jaswall, 1997). A CAVE is a surround screen (e.g., a CAVE can consist of three rear projection screens for walls and a down projection screen for the floor) and surround sound system that provides a VE with 3D graphics and 3D sound. Unlike HMDs, the images are projected onto screens that surround one or multiple simultaneous users. To see in stereoscopic 3D, users wear a lightweight LCD shutter glasses. One user also wears a special tracking device that the CAVE hardware uses to adapt the displays to that user's position and orientation.

3 Human Factors

Besides technology problems which could be solved in a near future because of rapid technology advance, there are many other problems that can cause an interaction technique (IT) to fail in its usability. The acceptance of VEs technology requires scrupulous optimization of the most basic interactions in order to maximize user performance and provide enjoyable virtual interfaces. In order to determine which human factors may be influential in VEs, we could examine past studies in human-computer interaction (HCI). But due to the complex nature of human virtual environment interaction (HVEI), erlier studies in HCI may not be directly applied to HVEI. VE measures need to focus on more than one task outcome to be effective and it may be essential to develop multicirteria measures to evaluate human performance in VEs. Eventhough, previous HCI studies could provide potential insights, this Chapter analyses studies directly applied to HVEI.

The ability of an individual to effectively complete a set of benchmark tests might provide a comprehensive assessment of (HVEI) performance. Some studies have been done toward benchmarking VE performance (Lampton & al., 1994) and the results from such tests could provide a baseline from which to judge the effectiveness of a VE.

This paper does not discuss about measurement methodologies of VE effectiveness but this chapter concentrates on the description of the main factors that affect efficiency of human task performance. The chapter also discusses the importance of health and social issues, which are usually neglected while developing a VR system to guarantee its success and usability. Finally, this chapter presents an overview of these health and safety issues.

3.1 Limitations of Technology

Users have difficulties controlling multiple DOF simultaneously, interacting in a volume rather than on a surface, and understanding 3D spatial relationships. Development of useful VE applications requires optimization of the ITs, so users can concentrate on high level tasks rather than on low level motor activities.

VR aims to allow users to interact with the system they are using on a more intuitive level (e.g., via gestures and movement as opposed to typing commands on a keyboard). These added interaction possibilities are what makes VR so exciting in its application but, unfortunately they also make it difficult to learn and use. The design of interaction techniques and user interfaces for VEs must be done with extreme care in order to produce useful and usable systems, and besides the technological limitations such as system latency and display resolution, the lack of interactive usability problems should be addressed (Rorke & al., 1998).

VE devices. Manipulation of objects in VEs is often difficult and inconvenient (Poupyrev & al., 1997). A lack of feedback, tracker noise, poor design of interaction techniques, and other factors can make a simple task of grabbing and moving a virtual object a frustrating experience. If humans can not act easily in VEs, then further pursuit of this technology may fail.

First, the *precise manipulation* of objects in the VE is difficult. While one is able to accurately track the positions of objects and represent this visually, the lack of haptic makes it difficult for the user to precisely manipulate objects. Haptic feedback is the term given to feel an object. This feel is a result of the weight of the object pulling the hand downwards, and the pressure exerted by the fingers on the objects, all of which the brain registers and uses to help accurately position or manipulate the object. At present, VR is able to reproduce only the visual information about the object, thus the users may see their hands holding an object, but, at no point they do actually believe that they are holding something real.

The *expensive equipment* used to capture data about the user for the VE (e.g., magnetic trackers, which provide a 3D value for their position and orientation) often prohibits acquiring complete equipment for the virtual system. Thus one is forced to use a limited number (usually 2 or 3) of them. This, in turn, limits the amount of input one is able to receive from the user and becomes other problem which restricts the usability of the system.

The *limited input information* provided to the system is another problem related to VEs (Mine & al., 1997). Most VE systems accept position and orientation data on the user's limb (usually head and, if lucky, both hands). A user also typically has a button or glove to provide signal/event information. This suffices for specifying simple six DOF motion and placement but in the real world, we do this and much more: object modification (usually with tools), directing the cooperation of helping hands, by spoken commands ("put that there"), measuring, and annotating objects with text. In most VR systems tool selection is difficult, voice command technology is marginal, measuring tools are rarely available, and alpha-numeric input is difficult.

The *limited precision* is another problem associated with many VR input devices. There is a limit to the accuracy of the readings that these devices produce, and they are also prone to an effect known as "drifting" whereby the value from the device changes when it should remain steady (e.g., if the tracker has not moved).

The *lack of physical work surfaces* in the VE is also a major interaction problem. People depend on naturally occurring physical constraints to give them some idea of the behavior of objects, for example a book pushed over the edge of a desk will fall to the ground (Mine & al., 1997). The physical workspaces also often provide some form of support for the user, alleviating fatigue and allowing more precise manipulation of limbs.

There are also some problems directly related with the IT. For example, image plane interaction or Go-Go technique can cause arm fatigue to the user working with his arms constantly extended. Similarly, worlds in miniature can cause fatigue when holding the physical clipboard.

VE design metaphors. VE *lacks a common interface*. The standard windows, icons, mouse, and pointer (WIMP) interface is now common on all computer platforms, and the user is easily able to identify common interface elements and begin productive work with little or no learning curves. Unfortunately, no such common interface exists in VR, for example, there is no set of standard input devices. So the user is forced to start from a very basic level whenever a new piece of software or hardware is encountered. Designing effective VE metaphors could likely enhance human performance in virtual worlds.

For example, affordances may be designed to assist users in interacting with the virtual world as they would interact with the (multimodal) real world (Stanney & al., 1998a). Affordances are cues and environmental properties that direct the behavior in an environment to achieve a goal. For example, if the task of a user in a VE is to open a file, the environment could highlight a 3D archive shelf in some point of the virtual world so the user knows that he has to walk in that direction. Thus, the environment, in some way, is addressing the user's behavior. To be effective in managing human performance, VE affordance properties need to be presented by both an inflow of controllable human behavior (i.e., an input or cue in the environment showing the user how to intuitively interact in the virtual scene) and an outflow of detectable information. (i.e, a detectable optic, haptic, olfactory or auditory output of the system). Reactions to this detectable information could then feed back into the cycle making them to be the inflow of additional controlled human behavior, directing in this manner, the behavior of users in every scene.

Well designed affordances could reduce the perceived navigational complexity of a VE and enhance user's performance, but unfortunately actually the human VE interface designers are using old affordances and metaphors (e.g., windows, toolbars), that may be inappropriate for HVEI.

Although VE designers have virtually endless design possibilities and they are not bound by the design object toolkits as designers were in the past, they also currently have limited guidance in designing usable human virtual environment interfaces. Consequently, research into the design of new metaphors and guidelines to develop such metaphors are needed. VE design will need to be multisensory, providing guidance for the design of effective visual, auditory, and haptic metaphors which provide a means of effectively presenting multisensory information to the users. VR sliders (3D equivalents of scroll bars), map cubes (3D maps that show space in a viewer's vicinity), tow planes (where a viewer's navigation is tied to a virtual object that tows him about the VE), and portals and spirals as replacement for windows are being investigated as potential visual metaphors for VEs (Laurel, 1994). The key to the design of these new metaphors may be to identify expectations developed through real world experiences in common and familiar environments and to use this knowledge to design VE interaction metaphors and thereby construct a more intuitive VE interface (Eberts, 1994).

3.2 Human Sensory Issues

Human beings are well equipped to interact in the world they live in, thus, for VR technology success, research should strive to make users interact with virtual worlds in the same way they interact with the real world, allowing them to exploit their cognitive and motor skills for interacting with the VE in a range of sensory modalities (Gobbetti and Scateni, 1998). In other words, 3D UIs should be able to use more than just the visual channel for human-machine communication. In this manner, Human performance in VEs will likely be influenced by VE design constraints related to human sensory limitations and integration issues with multimodal interaction (Stanney & al., 1998a).

VE design constraints related to human sensory limitations. To maximize human efficiency, VEs should be designed taking into consideration the abilities and limitations of human sensory and motor physiology. Without deep understanding on these areas, VEs systems could not be compatible with their users. The physiological and perceptual issues that directly affect the design of VEs include visual perception, auditory perception, and haptic and kinesthetic perception. Also, olfactory perception might have leverage on the design of VEs

(Gobbetti and Scateni, 1998).

Visual perception. The design of visual presentations for VEs is complicated because the human visual system is very sensitive to any anomalies in image perception (Larijani, 1994). When motion is introduced into a virtual scene, and if the VE is unable to generate approximate optical flow patterns, the user becomes very aware that the experience is not natural. Currently VE systems commonly provide the same image to both eyes not presenting a real stereo image but when they provide stereo, adjustments for the interpupillary distance are critical, and these adjustments can cause problems during post-exposure readaptation to normal conditions.

Other important factor is the viewer's visual field provided by the visual display (Biocca, 1992a). For example, HMDs reduce the field of view (FOV) of a user, decreasing the sense of motion in the peripheral vision. Current systems are generally limited to a FOV which does not provide enough peripheral vision and many VE tasks may require complete FOVs in order to achieve an immersion feeling, but while FOV is expanded the resolution of the projected images diminishes.In addition, wider FOVs have been shown to produce greater sensation of motion sickness. It is thus essential to determine what FOV is optimal to perform different kinds of VE tasks effectively.

Stereopsis, another physiological issue, is the portion of the visual field shared by both eyes (known as binocular field of vision). Stereopsis is important to study because it is a functional component of depth perception, and depth perception has bee shown to benefit virtual world performance (Ellis and Mengers, 1995). When two different images are presented to a viewer, the image of the dominant eye often dominates the visual system. When the dominant image alternates from one eye to another binocular rivalry effect occurs producing (extremely) disturbing sensations. Differences in size, display scene rendering and complexity, brightness and colors can lead to binocular rivalry effects. Such disturbing effects need to be eliminated or at least moderated for VEs to receive wide user acceptance.

Auditory perception. In order to synthesize a realistic auditory environment, it is important to obtain better understanding of how the ears receive sound, particularly focusing on 3D audio localization. Audio localization assists listeners in distinguishing separate sound sources and is affected by the presence of other sounds and the direction from which these sounds are originated. The aural realism of virtual spaces requires replicating the spatial characteristics of sounds (Biocca, 1992), thus, in order to effectively develop aural displays the ability of users to track and focus in on a particular auditory source needs to be well understood.

Haptic and kinesthetic perception (Smith & al., 1997). Haptics is the study of how to couple the human sense of touch with a computer generated world. It is important to incorporate haptic feedback in VEs because it substantially enhances performance. By adding an independent input channel, the amount of information processed by the brain increases and the increase of information reduces the error and time taken to complete a task.

Haptic perception can divided into two subfields: force (kinesthetic) feedback and tactile feedback. *Force (kinesthetic) feedback* is the area that deals with devices that interact with the muscles and tendons that give the human the sensation of a force being applied. These devices mainly consist of robotic manipulators that push back against the user. *Tactile feedback* deals with the devices that interact with the nerve endings in the skin which indicate heat, pressure, and texture. These devices typically are used to indicate if the user is in contact with a virtual object. In order to effectively design a haptic interface, the anatomy an physiology of the human body must be taken into consideration. In force feedback, the proportions and strengths of average joints must be considered. For example, since the hands are most often used for haptic interfaces, the properties of the hand should be considered when designing the interface. In tactile feedback, the interface must track several variables of the human sense of touch. For example, since the fingers are one of the most sensitive parts of the surface of the skin, forces on individual fingers should be such that they do not provide experience of discomfort or fatigue.

Humans use two different forms of haptic exploration: active and passive. *Active haptic exploration* occurs when the user controls his own actions. *Passive haptic exploration* occurs when the user's hand or finger is being guided by another person. Passive haptics experiences show to be more accurate. When the participants are guided around an object they are more likely to be able to identify the object because the entire attention can be paid on the task at hand. Active users make more distracting mistakes and have difficulty differentiating between the erroneous paths and the correct paths of exploration.

Olfactory perception. The main problem about simulating the human olfactory system is the little knowledge about how it works (Gobbetti and Scateni, 1998). Olfactory perception is the least mature perception technology and currently is almost limited to research applications. Eventhough, some studies (Yagel & al., 1996) reported that a VR environment giving olfatory cues should provide, to give sense of reality, the possibility to diffuse the odors when needed and purify and filter the air when the cue is no longer required.

Integration issues with multimodal interaction. While efforts are focused on developing effective visual, auditory, and haptic representations of virtual worlds, it is also important to determine how to effectively integrate this multimodal interaction. This characteristic of VEs to present the user with multiple inputs and outputs might help users to achieve better scores on performance for certain tasks presented in virtual worlds (Smith & al., 1997). There is, however, a limited understanding on how to effectively provide such sensorial parallelism (Burdea and Coiffet, 1994). When sensorial redundancy is provided to users it is essential to consider the design of the integration of these sources of feedback.

The use of redundant user inputs could potentially be useful in HVEI (Eberts, 1994) because humans usually use simultaneously haptic, verbal, and gesture actions for effective communication. Redundant input capability could also support user preferences (i.e., some users, for example, may prefer verbal interaction, while others may find direct manipulation more convenient). When providing redundant feedback it is fundamental to consider the limitations of human information processing to avoid sensorial overload.

Another factor that could affect user performance is the sensorial transposition which occurs when a user receives feedback through other senses than those expected. This may occur because a VE designer has substituted an unavailable system sensory feedback (e.g., force feedback) with other modes of feedback (e.g., visual or auditory). The sensorial substitution schemes may be one-to one (e.g., sound for force, visual for force, visual for sound) or more complex (e.g., visual for force and auditory, visual and auditory for force). The design of these substitution schemes should be consistent throughout the virtual world to avoid sensorial confusion.

3.3 Human Performance on VE Tasks

Computer speed and functionality, image processing, synthetic sound, and tracking mechanisms have been joined together to provide realistic virtual worlds. But as Stanney & al. (1998a) say, for VEs to be effective it is still required to determine how to maximize the efficiency of human task performance in virtual worlds. Maximizing the efficiency will require developing a set of guiding design principles that enable the user intuitive and efficient interaction. In order to determine the effectiveness of a VE, means of assessing human performance efficiency in virtual worlds is first required. Due to the complex nature of human virtual environment interaction (HVEI), it may be essential to develop effective multicriteria measures to evaluate human performance in virtual environments. Factors contributing to human efficiency in VEs include spatial orientation, and selection and manipulation techniques effectiveness.

Spatial orientation. If individuals cannot effectively navigate in VEs, then their ability to perform required task will be severely limited. Due to the current issue of users becoming lost in VEs, the design and development of standardized navigational techniques that assist individuals in maintaining spatial orientation within

VEs may be one of the most important issues to resolve in VEs design. Without adequate means of moving around the VE, developers cannot expect to maximize human performance. One potential approach of addressing this issue is to develop a means of measuring the navigational complexity of a VE (i.e., how challenging it is to move from a point A to a point B). Such a measure would likely be dependent on the sensorial cues and spatial mediators (e.g., maps) provided by an environment, as well on the spatial ability of the navigator.

There are many features in a VE that might affect user performance during the travel (Figure 16). Bowman & al. (1999) propose an experiment for the design and evaluation of VE travel techniques. The experiment presents a set of performance metrics and outside factors to explore the effects of a travel technique on spatial orientation in an IVE.

Performance metrics are response variables by which one measures the performance of a travel technique in a given situation. Application designers could test various travel techniques by implementing them in the target application and then running usability studies, but the results would be only valid for the tested application. The experiment proposes instead that travel techniques can be evaluated more generally measuring their performance in terms of the metrics provided. Application designers can specify the desired levels of performance for one or more of the metrics, then experimentally test the travel techniques and choose that one which best shows to achieve those desired levels. For travel, there are a large set of possible performance metrics representing different application goals and different uses for the travel interaction, but in the the most significant ones are listed in Table 1.

Since there is a lack of systematic methods for measuring or evaluating subjective metrics (such as presence and comfort) (Stanney & al., 1998a), the VE community has attempted to standardize questionnaires (Bowman & al. 1999) for these important subjective metrics.

Outside factors, besides travel technique, that might affect one or more of the

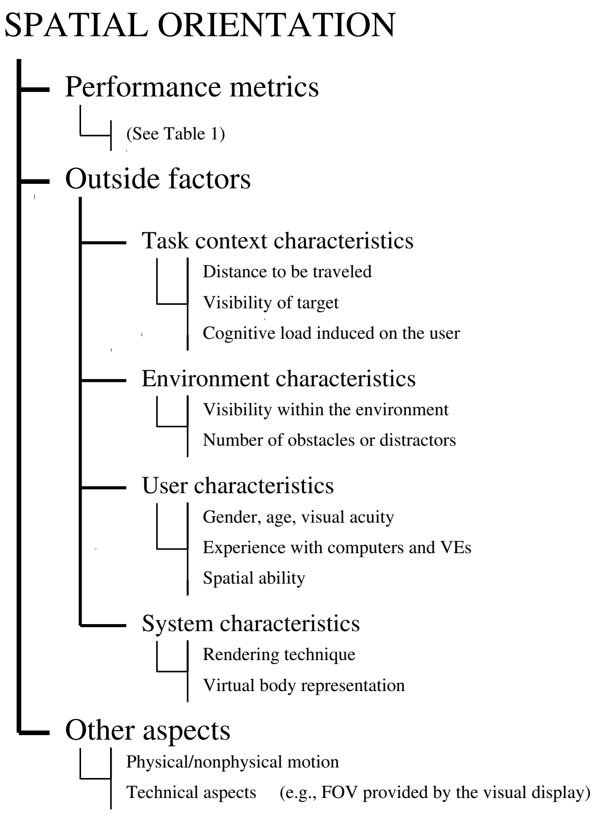

Figure 16: Factors to take into account while designing a travel technique to achieve optimal maintenance of spatial orientation.

performance metrics are divided into four groups:

- *task characteristics*
- *environment characteristics*
- *user characteristics*
- *system characteristics*

Because outside factors may strongly affect on performance, it is important to describe them in detail (Bowman & al., 1998). The nature of the tasks (**task characteristics**) being performed directly influence in how effectively humans can operate in virtual worlds. To justify the use of VE technology for a given task, when compared to alternative approaches, Stanney & al. (1998a) say that the use of a VE should improve task performance when transferred to the real world task because the VE system emphasize human sensory, perceptual, information processing, or cognitive capability. It is important to determine the types of task for which these types of capabilities are enhanced by using VE technology. Viable tasks can be specified by understanding the relationship between task characteristics and the corresponding VE characteristics (i.e., stereoscopic 3D visualization, real time interactivity, immersion, multisensory feedback) that effectively support their performance within the VE.

For immersive travel, some of the task characteristics that Bowman & al. (1999) consider to affect performance are: distance to be traveled, amount of curvature or number of turns in the path, visibility of target, number of degrees of freedom of motion required, accuracy required, complexity of the task (cognitive load induced on the user), and information required to the user.

Table 1: Performance metrics of travel techniques.

Performance metric	*Definition*
Completion time	The time for task completion and appropriate velocity.
Accuracy	The proximity to the desired target or path.
Spatial orientation	The user's knowledge of position and orientation within the environment.
Ease of learning	The ability of a novice user to learn the technique.
Information gathering potential	The user's ability to acquire information while traveling.
Sense of presence	The user's sense of immersion within the environment.
User comfort	The lack of simulator sickness, dizziness, or discomfort while traveling.

The environment in which the user travels can also have an effect on performance. The same task in different environments may produce surprisingly different results on one or more the performance metrics measurements. **Environment characteristics** are: visibility within the environment, number of obstacles or distractors, activity or motion within the environment, size of the environment, level of visual detail and fidelity, homogeneity (amount of variation) in the environment, structure, and alignment with the standard axes.

Varying one or more of these environment variables may allow to see some significant differences between travel techniques.

It is also important to consider the differences in users of VE applications when evaluating performance. This can be a significant factor in the performance of various techniques, because the designers of techniques often assume something implicitly about users. **User characteristics** that significantly influence VR experiences need to be identified in order to design VE systems that accommodate these unique needs of users. User differences have already been reported to influence the sense of presence (Barfield and Weghorst, 1993) and cybersickness experienced by VE participants (Parker and Harm, 1992). But what needs more research is which user characteristics have little impact and which highly affect virtual experiences.

Bowman & al. (1999) consider the following user characteristics to be relevant: age, gender, visual acuity, height, reach, ability to fuse stereo images, experience with VEs, experience with computers, technical/non-technical background, and spatial ability. Stanney & al. (1998a) consider that user personality also affect HVEI becoming increasingly important during more complex interactions.

Finally, different aspects of hardware and software used to realize the VE application must be taken into account. It is quite possible that decisions made by system developers or hardware designers may affect the performance of techniques for virtual travel. However, the fact that these factors are not always under control of the technique designer does not mean that they should not be considered in the design. For best performance, designers may need to create techniques which perform in a robust manner under a wide variety of system conditions. The **system characteristics** identified by Bowman & al. (1999) are: rendering technique, lighting model, frame rate, latency, display characteristics (stereo/mono, field of view, resolution, brightness, etc), collision detection, and virtual body representation.

These factors can cause differences in the usefulness of many ITs. Studies on the effects of varying frame rate and latency for various task have been performed (Ware and Balakrishnan, 1994), but there is still much work to be done.

Besides performance metrics and outside factors, studies have found that other aspects of the travel technique might have effect on spatial orientation. For example, a recent study (Chance & at., 1998) found that, in general, a technique more similar to real walking (physical translations and rotations) produced better spatial orientation in participants than those in which self-motion was virtual (the participant's view translated or rotated while the participant remained in place) and physical motion has a positive influence on reported levels of presence in a VE.

Bowman's & al. experiment (1999) using a HMD and comparing three different travel techniques (system-automated, pointing, and route-planning techniques) indicates that virtual travel techniques, in which users do not physically translate their bodies, can allow the maintenance of a user's spatial orientation. Other results of the same experiment show that path dimension significantly affects user performance. Furthermore there are no significant differences in spatial awareness (orientation) due to different levels of velocity or acceleration, except in the case where velocity is infinite (the user is instantly teleported to the target location) where this jumping technique seems to cause more disorientation. Finally, the strategies used by participants to perform the task at hand show to be significant, indicating that performance depends not only on the technique, environment, and task, but also on the sophistication of the user. Bowman & al. observed six main strategies during the experiment used by the subjects for minimizing error in realizing the tasks:

- *Stop and look:* The user simply stops moving at various points within the path and turn to look in other directions than the direction of motion.
- *Proprioceptive pointing:* Users physically point in the directions of objects they have already seen, to give themselves a proprioceptive cue for later recall.
- *Backing in:* Users turn around just before the end of the path and move backwards to the end, ensuring that they see the path just before it disappears.
- *Path retracting:* Users go back along the path they have just traversed through the scene, both to remind themselves of what they have seen, and to see the scene from another viewpoint.
- *Moving through walls:* If allowed, users move through walls in order to better understand the relationships between adjacent passageways.
- *3D overview:* Users fly up above the path to obtain a single view of the complete scene and objects within it.

Selection and manipulation techniques effectiveness. Evaluation of manipulation techniques also involves measuring user performance using some metrics (Figure 17). Some of the performance metrics (Table 2) proposed by Poupyrev & al. (1997) to evaluate the effectiveness of selection and manipulation techniques are the same as proposed by Bowman & al. (1999) to evaluate travel techniques but this time are measured in the context of selection and manipulation.

Some of the metrics defined in Table 2 require to distinguish between selection, and positioning and orienting. **Completion time** for a selection task is the time from the moment the stimulus triggers the user to select an object until the moment when it has been successfully selected. For a positioning and orienting task,

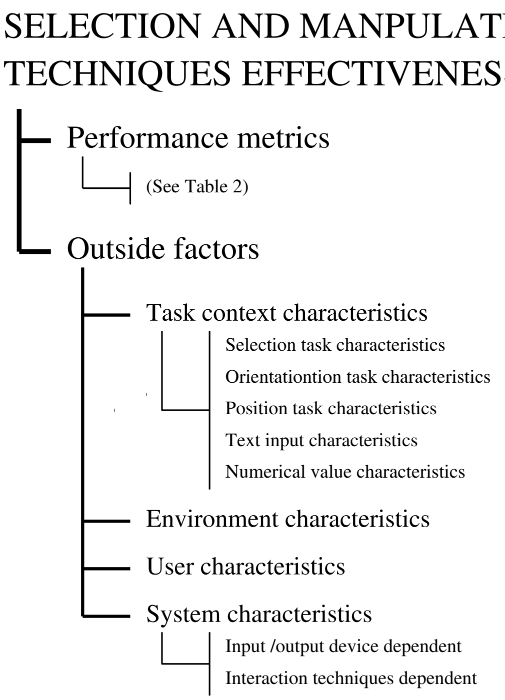

Figure 17: Criteria to take into account while designing and evaluating selection and manipulation techniques to achieve optimal performance.

Table 2: Performance metrics of selection and manipulation techniques.

Performance metric	*Definition*
Completion time	The time taken to successfully accomplish the task.
Accuracy	The proximity to the desired position and/or orientation of the target object.
Error rate	The number of failed attempts to accomplish the task.
Ease of use	The cognitive load on the user while using the technique.
Ease of learning	The ability of the user to improve performance with experience.
Sense of presence	The user’s sense of immersion and spatial awareness.

completion time is measured from the moment when the user picks (selects) an object until the moment it has been positioned with the required accuracy. The **error rate** for a selection task is the number of failed attempts to select an object. For a positioning and orienting task it is the number of iterative movements required for positioning or orienting an object with the required accuracy.

As in the case of travel techniques, there are many factors which affect user performance while accomplishing a task. These factors can be classified as (Poupyrev & al., 1997) *task context dependent* (required precision, initial and final conditions of the task, task constraints, etc.), *application dependent* (configuration of the VE, size, shape, and locations of objects, color lighting and others), *user dependent* (experience, perceptual and motor abilities, anthropometrical differences and so on), *input/output device dependent* (attributes of the device such as DOF, resolution, field of view, supported depth cues and others), and *interaction techniques dependent* (underlaying metaphors of techniques, their design and implementation).

Because of the similarity to Bowman's & al. (1998) classification and description of outside factors (task, environment, user and system characteristics) affecting on travel techniques performance, application, user, and device parameters are not addressed. Instead, task parameters which greatly vary in the context of selection an manipulation are described in more detail. As described in Chapter 2, the taxonomy of Poupyrev & al. (1997) suggested five basic interaction tasks: selection, position, orientation, text and quantify. Because of their relevance we will next concentrate on selection, position and orientation task parameters:

Selection task parameters. Selection is the task of identifying an object. The parameters to consider are:

- *number of objects to be selected*
- *size and distance to the target object*
- *direction to the target object*

- *occlusion of the target object*
- *other parameters*

It is important to consider the **number of objects to be selected**. In the simple case we need to select only one object but many times more than one object can be selected requiring a disambiguation method that could affect user performance.

Size and distance to the target object can also influence on performance. The time required to select an object depends on the ratio of the distance to the size of the object, i.e., a large object located close is easier to select than a small one located far away. In IVE, however, due to perspective effects, a large object located far away could be more difficult to select than a small one located close by.

Also, depending on the **direction to the target object** relatively to the user, different performance is achieved as different body parts and muscles are brought into the action to select the desired object.

Occlusion of the target object influences user performance in two ways. First, when an object is occluded its visual size diminishes. In this respect, selection of an occluded object is the same as selection of a non occluded one with smaller size. Second, occlusion partially blocks access to the target object and makes it difficult for the user to access it, and the close proximity of other objects increases the possibilities of selecting a wrong object.

There are more characteristics which might be considered, such as dynamics of the target, density of objects around the target, bounding volume of the target object, or visual feedback indicating that an object has already been selected.

Position task parameters. Positioning is the task of moving an object to a desired final terminal location. The execution of position tasks can be generally split in three phases: selection of the object, gross (or primary) travel, and a corrective motion to position the object on the terminal with the desired accuracy. Parameters to be considered are:

- *initial distance to the manipulated object*
- *initial direction to the manipulated object*
- *distance to the terminal object*
- *direction to the terminal object*
- *required precision of the positioning*
- *other parameters*

Initial distance and direction to the manipulated object, and **distance** and **direction to the terminal object** may affect on performance because the time and accuracy of positional movements depend on the object's initial position, the travel distance, and the direction from the user to the terminal. **Required precision of the positioning** can also influence on performance. Precise positioning is more difficult than imprecise for unconstarined movements. Required accuracy mostly affects the last phase of the positioning movement, corrective motion close to the terminal. Other position task characteristics to consider are visibility of the terminal (blind positioning), dynamics of the terminal (positioning on a moving terminal), occlusion of the terminal, size of the manipulated object, etc.

Orientation task parameters. Orientating is the task of changing orientation of an object from an initial to a desired orientation. Orienting consists of object selection, gross orienting and then fine adjustment to achieve desired orientation precision. The task characteristics to consider are:

- *direction to the object*
- *distance to the object*
- *initial orientation of the object*

- *final orientation of the object*
- *other parameters*

Orientation of an object located close to the user and one located far off is different and may require different interaction techniques. Therefore, **direction** and **distance to the object** might cause different levels of user performance while realizing the the positioning task. The **initial** and **final orientation of the object** define the direction of the object rotation during task execution, thus, causing an effect on performance if the user rotates more naturally in one direction than in the other. Another orientation task characteristic that may affect user performance is the required orientation accuracy. High precision could result tedious for the user.

3.4 Health and Safety Issues

It is essential for users who interact with VE to feel that their welfare and health is ensured while interacting with the VE. With no means of achieving this goal VR systems are not expected to succeed. If human element in these systems is not taken into consideration, it could result in discomfort, harm, or even injury. There are many health and safety issues that affect VEs users. Stanney & al. (1998a) divide these issues into direct and indirect effects (Figure 18). The direct effects can be subdivided into microscopic (e.g., user's tissue) and macroscopic effects (e.g., trauma, sickness). The indirect effects include physiological aftereffects and physiological disturbances.

Direct microscopic effects. There are several microscopic effects that could affect the tissues of VEs users. The eyes, which will be closely coupled to the visual displays used in VEs, have the risk of being harmed. The eyes may be affected by the electromagnetic field (emf) of VEs if the exposure is intense or prolonged enough. For example, emf exposure could cause cataracts if the visual display uses X-rays or laser lights. Also visual distortions could be caused by poor adjustment of the

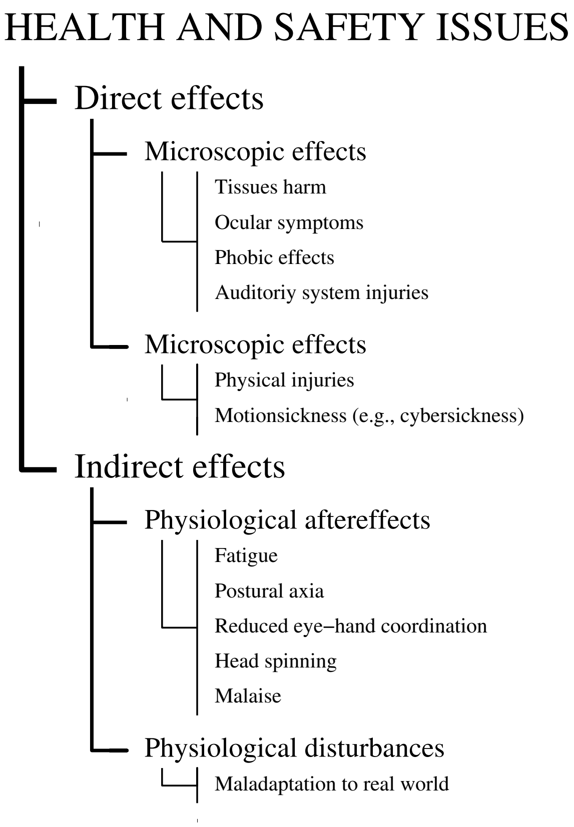

Figure 18: Categorization of effects due to VE exposure.

visual displays (e.g., eye strain, flicker, glare). These effects associated with incorrect settings (poor illumination, poor contrast, and close proximity of the working distance) can be mitigated by a careful visual display design. If not considered, however, visual stress from visual displays exposure can lead to ocular symptoms (e.g., unstable binocular vision and reduced visual acuity) and related visual aftereffects may provoke a user unsafe to interact with the real world (e.g., drive a car) after VE exposure. Other results show that strong emfs can also cause cellular and genetic damage in the brain as in other tissues.

Phobic effects may result from VE use, such as claustrophobia (e.g., HMD enclosure) and anxiety. The auditory system and inner ear could be affected by VE exposure to high volume audio. One possible effect of such exposure is noise induced hearing loss and continuous exposure could lead to nerve deafness. Prolonged repetitive VE movements could also cause over use injuries to the body (e.g., carpal tunnel syndrome, tenosynovitis, epicondylitis). Such effects can be moderated by emphasizing ergonomics in VE design and judicious usage procedures. The head, neck, and spine could be harmed by the weight or position of specific visual displays, such as HMDs. This overweight additionally leads to symptoms of motion sickness.

Direct macroscopic effects. VE equipment interferes with normal sensory perception and body movements, and can cause physical injury or trauma while interacting with the virtual world. Limited (or eliminated) vision of real surroundings when wearing a visual display or imbalance of body position due to the weight of VE equipment or cables (not visible in the virtual world) that link equipment to the computers can cause the user to fall.

If haptic feedback systems fail a user might be accidentally physically wounded. A haptic study made by Smith (1997) suggests that the user must be able to overpower the system and feel like he can not be injured by the device if control system fails. In attempting to replicate physical forces, robotic systems which are much stronger than the human joints must be designed to account for the flexion and extension strengths and flexibility of the joints. If an interface for the average user is being considered, it should be taken into account that the size of the hand affects the extension and flexion of fingers. For example, if the UI uses a glove for interaction, a hand that is larger than the average user hand will not be able to flex its finger as designed and could be injured by the interface.

Another direct macroscopic effect is that users may become motionsick. ***Cybersickness,*** a form of motion sickness, has become an essential issue of recent concern.

There is a tendency for some users to exhibit symptoms (that parallel symptoms of classical motion sickness) during and after the VE experience, that poses a serious threat to the usability of VR systems. Symptoms due to cybersickness and motion sickness include eye strain, headache, pallor, sweating, dryness of mouth, fullness of stomach, disorientation, vertigo, ataxia (i.e., postural instability), nausea, and vomiting. Although both motion sickness and cybersickness produce the same type of symptoms, they are not necessarily the same thing. Cybersickness is distinct from motion sickness in that the user is often stationary but has a compelling sense of self motion through moving visual imaginery and can occur strictly with visual simulation and no vestibular simulation while vestibular simulation alone can be sufficient to induce motion sickness. La Viola (2000) proposes three main theories as to the cause of cybersickness: sensory conflict theory, the poison theory, and the postural instability theory.

The *sensory conflict theory* is based on the premise that discrepancies between the senses which provide information about the body's orientation and motion cause a perceptual conflict which the body does not know how to handle. When the user moves in the VE the visual system tells the user that he is moving in a certain direction, however, since the user is not actually moving, he does not get the expected response from the vestibular system, thus a conflict occurs and cybersickness may result.

The *poison theory* suggests that the adverse simulation found in some VEs can affect the visual and vestibular systems in such a way that the body misreads the information and thinks it has ingested some type of toxic substance thus causing disturbing symptoms which led to an emetic response.

As the conflict theory, the poison theory lacks predictive power in determining if cybersickness will occur given a certain situation or how severe it will be. Also, it makes no determination for why people who get sick in VEs do not always have an emetic response. It also, as the conflict theory, does not provide any explanation of why some people get cybersick in VEs but others with a given set of identical stimuli do not.

The *postural instability theory* is centered on the idea that one of the primary behavioral goals in humans is to maintain postural stability in the environment. For example, a user may use muscular force to resist the tilt of an angular acceleration during a virtual navigation. This angular acceration is visually perceived and since there was no physical tilt the user created an unintended divergence from a stable position causing postural instability.

The postural instability theory states that the cause of cybersickness is prolonged postural instability and the severity of the symptoms are straightly related to the duration of the instability, so, the longer the duration of the instability, the more severe the symptoms are. Also, symptoms will not occur if postural control is lost completely because the event is very brief.

There are many contributing factors to cybersickness in VEs that La Viola (2000) classifies into display and technology issues and user characteristics. Due to the imperfections in the technology, there are a number of problems (specially with visual displays) which have been associated with inducing cybersickness. Fortunately, with time technology improves so many of these problems could be avoided in the future. Some of these technical system factors are:

- *position tracking error*
- *lag*
- *flicker*
- *vection*

Position tracking error might derive into cybersickness. An accurate representation of the user in the virtual space relies on the ability of the VE technology to track the user's limbs in physical space. Also, the tracking information gives the user the correct perspective when viewing in the VE. Positions trackers are not completely accurate because of technology limitations, and depending on their inaccuracy rate it will determine if they will cause cybersickness symptoms.

The time between the user initiating an action and the action actually occurring in the VE represents **lag**. If lag is significant, the computer will not immediately update the display, and the user will have to wait for the images to appear where they are expected to be. This delay is very unsettling and cause cybersickness symptoms. **Flicker** is distracting, can cause eye fatigue, and could induce cybersickness symptoms. The perception of flicker differs between individuals and depends on the flicker fusion frequency threshold (the point at which flicker becomes visually perceptible). One of the goals of virtual reality is to surround the user's field of view with visual simulation and this goal becomes a problem since the likelihood that flicker will be perceived increases as the FOV increases. In order to reduce the possibility of flicker, the refresh rate of the system must be increased and this is likely to be achieved as the technology improves.

Finally, when someone is in a stationary position, he can still get the impression of self motion under certain conditions. This deceptive impression of self motion is called **vection**. Wide FOV can produce vection since a larger FOV stimulates more the retinal periphery. The optical flow rate (perceived speed) also contributes to induce vection since a faster flow rate will increase the perceived motion's speed thus making the illusion more intense. Another important factor to consider in the cause of vection is the apparent depth of the objects in the VE.

One of the most difficult questions about the causes of cybersickness is why some people gets sick under certain situations while others do not. However, some work in this area stands out the following individual factors:

- *Gender:* A study of Biocca (1992b) shows that women seem to be more susceptible to cybersickness than men. One of the reasons is that women generally have wider FOVs and thus higher likelihood of flicker perception.
- *Age:* Differences in age play a factor in cybersickness susceptibility. La Viola (2000) states that susceptibility is greatest between the ages of 2 and 12 years. It decreases rapidly from 12 to 21 years of age and more slowly thereafter. And around 50 years of age, cybersickness is almost nonexistent.
- *Position in the simulator:* Positioning the user in the VE can play a role

in the individual's susceptibility to cybersickness. Based on postural instability theory, sitting seems to be a better position to reduce cybersickness symptoms since it reduces the demands on postural control. Positioning a user in a VE where more than one person participates can also increase the susceptibility of cybersickness. Users who control the VE simulation are less susceptible to cybersickness than those who are passive participants.

Theoretical treatments of these system and user factors have failed to produce a general predictive model of cybersickness. Stanney & al. (in press) believe that the temporal aspects of exposure are factors that have been largely neglected in systematic evaluation and that there is a big influence of these temporal factors in the variance of motion sickness outcomes (i.e., incidence and severity). In their view temporal aspects have two main components: the amount of time that a person is exposed to a VE (exposure duration) and the distribution of the exposure over time (repeated exposures). Their evaluation confirmed that longer exposures produce greater sickness outcome and that total sickness subsides over repeated exposures due to the significant adaptations of the subjects to the conditions that produce cybersickness. This was found to be true over a variety of simulators, scenarios, and diverse subject sources.

La Viola (2000) discusses several methods that could be used to reduce cybersickness symptoms such as motion platforms, direct vestibular simulation, and adaptation.

Motion platforms. Basing on the conflict theory, a mean for reducing cybersickness can be to add motion platforms to the virtual simulator. With the motion platforms the user would get both vestibular and visual stimulation. Some users still get sick using motion platforms, maybe because the motion platform was not aligned correctly with the visual stimuli and there would still be a sensory conflict and so research needs still to be conducted to determine if motion platforms can be used to reduce the severity of cybersickness symptoms.

Direct vestibular simulation. This idea is similar to motion platform concept but instead of using a large motion base, the user wears a device which sends electrical signals to the eigth cranial nerve which tricks the vestibular system into believing that there is a linear or angular acceleration taking place. Currently, there has been no attempt to use this device in the context of cybersickness reduction and there are some questions like if the device can be accurate enough to eliminate the sensory conflict or how much electrical current is needed to provide the brain with self motion information that need to be answered.

Adaptation. One of the more common approaches that has been developed and used in practice is to provide VE users with some type of adaptation program to the VE. Sensory adaptation is possible because of the plasticity of the human nervous system that allows some individuals to adapt and become less ill with continued system interaction. By having users increase their exposure time gradually they can adapt to to the VE. However, adaptation strategies that appear to be the best method for cybersickness reduction do not help the user to readapt to the real world once the user is done with the VE. As a result, aftereffects (i.e., indirect effects that are described later) can still occur.

Stanney and Hash (1998) have found in their studies an evidence to suggest that cybersickness may be overcome or moderated by providing users with an optimal level of user control over their movements in the virtual world. If users are allowed sufficient degree of freedom over their movements within a VE, it is suggested that the contributing factors to cybersickness (i.e., lag, vection, position tracking errors, age, etc.) will not induce cybersickness as severely as that resulting from a passive control environment, i.e., an environment where the users are forced to traverse a predetermined path through the virtual world as passive observers with no control over their movements.

Their studies show that active an voluntary movements made when users have control over their own motion provide the key to efficiently adapting to sensory rearrangements in VEs. More specifically, active control (i.e., users can move about a VE with total freedom of motion) is superior to passive control when attempting to adapt to a rearranged environment, but it may not be the best solution to the cybersickness problem. Under active control, users can move about a VE with total freedom of motion. Under such conditions, users may not be able to handle efficiently the abundant amount of sensory information that is generated as a result of their unrestricted movements. While cybersickness may be overcome under active control conditions, because users can predict and thus adapt to their movements in the VE, this process may occur at slower rate than under coupled control conditions described below.

Coupled control (or passive-active control), where total control (i.e., full range of user movements) is restricted to prevent extraneous motion and where the user control is mapped to the needs of the task, is an effective method for minimizing cybersickness. By allowing users task oriented control within a VE, users can adapt rapidly to their movements in the altered world and not be overloaded with extraneous sensory information associated with active control.

Indirect effects. A less known but equally important consequence of HVEI is the potential deleterious physiological aftereffects from VE exposure. Any effect of a VE that is observed after the user has returned to the physical world is defined as an aftereffect, including head spinning, reduced eye-hand coordination, flashbacks, illusory climbing and turning sensations, perceived inversions of the visual field, reduced motor control, disturbed vestibulo-ocular reflex, postural disturbances, sickness, fatigue, malaise, and drowsiness. Aftereffects are dependent on individual factors and may persist for some time after simulator exposure varying from few hours to several weeks.

Stanney & al. (1998b) support the idea that many of the aftereffects identified result from the user's adaptation to the VE. With protracted or repeated exposures individuals become adapted, feeling less affected (i.e., less ill) when exiting the VE. Although such adaptation may seem advantageous, these users may actually be the ones at greatest risk because aftereffects make individuals maladapted for the return to the real world once the VEI concludes. Stanney & al. found evidence that when VE devices present altered visual conditions, adaptation does occur.

Adaptation to altered visual cues does not simply disappear immediately after exposure but can manifest itself in the real world for extended periods of time (Stanney & al., 1999). Users should be aware of the danger of driving, roof repair, or other machinery use after VE exposure.

Until the technological defects of most current VEs can be corrected, adaptation and aftereffects will be inevitable accompaniments of their use. Stanney & al. (1998b) suggest an alternative to the technological solution. If a VE does include discordances that trigger adaptation, but ***dual adaptation*** (i.e., the ability to seamlessly transition between two disparate sensory environments) has been produced within the user and is complete, aftereffects should either not occur or be extremely short lived. The concept of dual adaptation is that, as a result of repeated adaptation to a given VE and readaptation to the normal environment, users will eventually acquire the ability to move between the virtual and real environments with little or no interference (i.e., no aftereffects). Although Stanney & al. found evidence for dual adaptation, it has not been experimentally demonstrated how to systematically generate complete dual adaptation or if this adaptation will be achieved. The identification of dual adaptation strategies could be substantially beneficial to the advancement of VE technology.

3.5 Social Impact

While research is often taking into consideration human performance and health and safety issues when developing a new technology, an often neglected effect of new technologies is their potential social impact. VR is a new technology, which like its precedents (e.g., television, computers, video games, and simulations) has the potential for negative social implications through misuse and abuse (Kallman, 1993). We are already aware of the negative influences of television and video games, which are much less interactive than VEs. Rather than waiting for these negative issues to arise in VEs, it might be prudent and beneficial to address social issues before they result in crisis or harm.

Actually the potential negative social influences resulting from VE exposure are not well understood. For example, VE applications that are presented as being forms of entertainment but which in fact represent violence and destruction may become available at video games rooms in the near future (Sheridan, 1993). The concern is that users who are embedded and obsessed in what it seems like harmless violence in the virtual world may become desensitized to their violent virtual actions and then reproduce that behavior in the look alike real world. There is a concern that virtual interactions could provoke addiction and violence.

There are many open issues (Stone, 1993) that should be carefully analyzed to anticipate or perhaps prevent some of the problems of VEs: What will be the psychological and character effects of VE use? How will interaction in the virtual world modify behavior? Will individuals transfer violent virtual experiences to the real world? Will users experience traumatic physical or psychological consequences due to a virtual interaction? Will people avoid reality and real social encounters with peers and become addicted to escapism? Is continual exposure to violent virtual worlds similar to military training, which through continued exposure may desensitize individuals to the acts of killing and maiming? How will VE influence young children who are particularly liable to psychological and moral influence? Even with all of these concerns there is currently limited focus on the social implications of VE technology by its developers.

However, there are strong scientific reasons (Kallman, 1993) why the active engagement available in VE games rather than passive action of television watching can have better retention of learned skills. The implication is that continual act of virtual violence will facilitate a VE games user's (usually youth) ability to reason and think in the realm of violent behavior and this increased negative simulation may then be channeled into other activities in the real world. VR gaming thus represents an unknown risk of increased violence in the next generation, and for these reasons, the society should be aware about the impact playing these games may have on the behavior of youth. Research is needed that weights the risks and potential consequences associated with VE against the benefits. For example, the development of a VE system may improve the lives of several individuals (e.g., toy manufacturers, stockholders) yet harm others (e.g., children who become emotionally unhealthy due to the excessive exposure to violent VR games).

Research needs to be directed at developing a new conceptual model of the social implications of HVEI. Through such model, the social consequences of HVEI may be predicted. But there are problems with pursuing this research, including the lack of measures to quantify violent behavior or the considerable time required to establish cause-effect relationships.

4 Conclusions

VR holds promises of being the ultimate human interface which would incorporate a natural intuitive interface between a human and a machine generated environment.

Developing VR applications which enable actual use over a period of time requires optimization of interaction techniques, such as object selection and manipulation or navigation immersed in a VE. However, interaction in 3D environments is not well understood and there is a lack of empirical data regarding virtual environment interaction techniques. There is a need of formal evaluation of ITs, leading to easily applied guidelines and principles. Series of studies have been done in the direction of formal and systematic assessment of immersive ITs (Lampton & al., 1994; Poupyrev & al., 1997; Bowman & al., 1998) which provide testbed evaluations that can be effective methods for this purpose.

Even though this paper presents a set of taxonomies which give light for better designing of ITs, the goal was to describe how current ITs solve different VR interaction problems (like how to manipulate objects out of reach, orientation, etc.), how users interact with the VE using input devices based on the ITs, and several visual displays used for immersing into the VE.

In addition, further research and development on interaction and visual devices exploring possible ways to integrate multisensory (optic, haptic, olfactory, and 3D audio perception) cues into seamless virtual interfaces would help to build a more realistic interface metaphor.

On the other hand, the needs and abilities of VE users have not been adequately addressed in the design of virtual worlds. We have indicated variables that need to be considered when studying human factors issues in VE. Researchers have *a priori* limited knowledge of which parameters are the most influential and also further research in this area of VE is needed. Especially we have considered sensory issues, performance on VE interaction, safety and social impact issues.

As technology rapidly advances providing higher quality images with greater fields of view and systems with higher DOF, the use of VE will probably expand at rapid rate. Since these VEs will involve the interaction of people and machines, we need to apply human factors principles to their design and use.

References

Barfield, W. and Weghorst, S. (1993) The Sence of Presence within Virtual Environments: A Conceptual Model. *Human-computer Interaction: Software and Hardware Interfaces* (Salvendy, G. and Smith, M.), Elsevier Science Publishers, Amsterdam, Netherland, 699-704.

Biocca, F. (1992a) Virtual Reality Technology: A Tutorial. *Journal of Communication* **42**(4), 23-72.

Biocca, F. (1992b) Will Simulation Sickness Slow Down the Diffusion of Virtual Environment Technology. *Presence: Teleoperators and Virtual Environments* **1**(3), 334-343.

Bowman, D.A. (1999) *Interaction Techniques for Common Tasks in Immersive Virtual Environments: Design, Evaluation, and Application.* Ph.D. Thesis, Computer Science Department, Georgia Institute of Technology, USA.

Bowman, D.A. and Hodges, L.F. (1997) An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. *Proceedings of the Symposium on Interactive 3D Graphics*, ACM Press, New York, USA, 35-38.

Bowman, D.A. and Hodges, L.F. (1998) The Virtual Venue: User-computer Interaction in Information-rich Virtual Environments. *Presence: Teleoperators and Virtual Environments* **7**(5), 478-493.

Bowman, D.A. and Hodges, L.F. (1999) Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments. *Journal of Visual language and Computing* **10**(1), 37-53.

Bowman, D.A., Hodges, L.F. and Koller, D. (1998) A Methodology for Evaluation of Travel Techniques for Immersive Virtual Environments. *Research, Development, and Applications* **3**(2), 20-131.

Bowman, D.A., Davis, T.E., Hodges, L.F. and Badre, A.N. (1999) Maintaining Spatial Orientation During Travel in an Immersive Virtual Environment. *Presence: Teleoperators and Virtual Environments* **8**(6), 618-631.

Burdea, G. and Coiffet, P. (1994) *Virtual Reality Technology*, Wiley, New York, USA, 1-7.

Chance, S., Gaunet, F., Beall, A. and Loomis, J. (1998) Locomotion Mode Affects the Updating of Objects Encountered During Travel: the Contribution of Vestibular and Proprioceptive Inputs to Path Integration. *Presence: Teleoperators and Virtual Environments* **7**(2), 168-178.

Coomans, M.K.D. (1999) A Virtual Reality User Interface for a Design Information System. *CCAI: Journal for Integrated Study of Artificial Intelligence, Cognitive Science and Applied Epistemology*, Rijks University Press, Gent, Belgium, 24-38.

Eberts, R.E. (1994). *User Interface Design*, Prentice Hall, New York, 543-568.

Ellis, S.R. and Mengers, B.M. (1995) Judged Distance to Virtual Objects in the Near Visual Field. *Proceedings of the Human Factors Society 39th Annual Meeting*, Humans Factors Society Press, California, USA, 1400-1404.

Emerson, T. (1999) New Trends in Virtual Environment User Interface Research: Case Studies at the Human Interface Technology Laboratory. *Proceedings of Eurographics on Virtual Environments*, ACM Press, New York, USA, 17-25.

Erickson, T. (1990) Working with Interface Metaphors. *The Art of Human-computer Interface Design* (Laurel, B.), Addison-Wesley Publishing Company, MA, USA, 65-73.

Foley, J., Van Dam, A., Feiner, S. and Hughes, J. (1992) *Computer Graphics, Principles and Practice. 2nd Edition.* Addison-Wesley Publishing Company, MA, USA, 1175-1192.

Forsberg, A., Herndon, K. and Zeleznik, R. (1996) Aperture Based Selection for Immersive Virtual Environments. *Proceedings on User Interface Software and Technology*, ACM Press, New York, USA, 95-96.

Gobbetti, E. and Scateni, R. (1998) Vitual Reality: Past, Present, and Future. *Virtual Environments in Clinical Psychology and Neuroscience: Methods and Techniques in Advanced Patient-Therapist Interaction*, (Riva, G., Wiederhold, B. K. and Molinari. E.), Center for Advanced Studies, Research, and Development in Sardinia, Cagliari, Italy, 3-20.

Holbrook, M.B. (1997) Three Dimensional Stereographic Visual Displays in Marketing and Consumer research. *Academy of Marketing Science*, Electronic Journal at http://www.ams-web.org/

Jacoby, R., Ferneau, M. and Huphries, J. (1994) Gestural Interaction in a Virtual Environment. *Proceedings of Stereoscopic Display and Virtual Reality Systems: The Engineering Reality of Virtual Reality*, SPIE–The International Society for Optical Engineering, Washington, 355-364.

Jaswal, V. (1997) CAVEvis: Distributed Real-Time Visualization of Time-Varying Scalar and Vector Fields Using the CAVE Virtual Reality Theater. *Proceedings of the 1997 Conference on Visualization*, ACM Press, New York, USA, 301-308.

Kallman, E.A. (1993) Ethical Evaluation: A Necessary Element in Virtual Environment Research. *Presence: Teleoperators and Virtual Environments* **2**(2), 143-146.

Lampton, D.R., Knerr, B.W., Goldberg, S.L., Bliss, J.P., Moshell, J.M. and Blau, B.S. (1994) The Virtual Environment Performance Assessment Battery (VEPAB): Development and Evaluation. *Presence: Teleoperators and Virtual Environments*, **3**(2), 145-147.

Larijani, L.C. (1994). *The Virtual Reality Primer*, McGraw-Hill, New York, USA.

Laurel, B. (1994) Narrative and Interaction in Placeholder. *Computer Graphics*

28(2), 120-123.

La Viola, J.J. (2000) A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bulletin* **32**(1), 47-56.

MacKenzie, I.S., Soukoreff, R.W. and Pal, C. (1997). A Two-ball Mouse Affords Three Degrees of Freedom. *Conference on Human Factors in Computing Systems*, ACM Press, New York, USA, 303-304.

Main, M. (1995) *Virtual Environment Interaction Techniques*. University of North Carolina, Chapel Hill, Computer Science Department, Technical Report TR95-018, USA.

Mine, M.R., Brooks Jr., F.P. and Sequin, C.H. (1997) Moving Objects in Space: Exploiting Proprioception in Virtual-environment Interaction. *Proceedings of the 24th Annual Conference on Computer Graphics & Interactive Techniques*, ACM Press, New York, USA, 19-26.

Nielsen, J. (1993) Noncommand User Interfaces. *Communications of the ACM*, **36**(4), 83-99.

Parker, D.E. and Harm, D.L. (1992) Mental Rotation: A Key to Mitigation of Motion Sickness in the Virtual Environment? *Presence: Teleoperators and Virtual Environments* **1**(3), 329-333.

Perry, L.D.S., Smith, C.M. and Yang, S. (1997) An Investigation of Current Virtual Reality Interfaces. *ACM Crossroads*, **3**(see www.acm.org/crossroads), 23-28.

Pierce, J.S., Zeleznik, R., Forsberg, A., Conway, M.J. and Hong, S. (1997) Image Plane Interaction Techniques in 3D Immersive Environments. *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, ACM Press, New York, USA, 39-43.

Poupyrev, I., Weghorst, S., Billinghurst, M. and Ichikawa, T. (1996) Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. *Proceed-*

ings of User Interface Software and Technology, ACM Press, New York, USA, 70-80.

Poupyrev, I., Weghorst, S., Billinghurst, M. and Ichikawa, T. (1997) A Framework and Testbed for Studying Manipulation Techniques for Immersive Virtual Reality. *Proceedings of the 1997 ACM Symposium on Virtual Reality Software and Technology*, ACM Press, New York, USA, 21-28.

Poupyrev, I., Weghorst, S., Billinghurst, M. and Icikawa, T. (1998) Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. *Eurographics* **17**(3), 41-52.

Rorke, M., Bangay, S. and Wentworth, P. (1998) Virtual Reality Interaction Techniques. *Proceedings of the 1st Annual South African Telecommunication, Networks and Application Conference* SATNAC, South Africa, 525-532.

Sheridan, T.B. (1993) My Anxieties About Virtual Environments. *Presence: Teleoperators and Virtual Environments* **2**(2), 141-142.

Smith, C.M. (1997) Human Factors in Haptic Interfaces. *ACM Crossroads*, **3**(see www.acm.org/crossroads), 79-85.

Stanney, K.M. (1997) Virtual Marketing: Design Business Information Visualization through Cognitive and Visual Displays Theories. *International Journal of Cognitive Ergonomics* **1**(1), 77-99

Stanney, K.M. and Hash, P. (1998) Locus of User-initiated Control in Virtual Environments: Influences on Cybersickness. *Presence: Teleoperators and Virtual Environments* **7**(5), 447-459.

Stanney, K.M., Mourant, R.R. and Kennedy, R.S. (1998a) Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence: Teleoperators and Virtual Environments* **7**(4), 327-351.

Stanney, K.M., Salvendy, G., Deisigner, J., DiZio, P., Ellis, S., Ellison, E., Fogleman, G., Gallimore, J., Hettinger, L., Kennedy, R.S., Lackner, J., Lawson, B., Maida, J., Mead, A., Mon-Williams, M., Newman, D., Piantanida, T., Reevs, L., Riedel, O., Singer, M., Stoffregen, T., Wann, J., Welch, R., Wilson, J. and Witmer, B. (1998b) Aftereffects and Sense of Presence in Virtual Environments: Formulation of a Research and Development Agenda. *International Journal of Human Computer Interaction* **10**(2), 135-187.

Stanney, K.M., Kennedy, R.S., Drexler, J.M. and Harm, D.L. (1999) Motion Sickness and Proprioperceptive Aftereffects Following Virtual Environment Exposure. *Applied Ergonomics* **30**, 27-38.

Stanney, K.M., Kennedy, R.S. and Dunlap, W.P. (in press) Duration and Exposure to Virtual Environments: Sickness Curves During and Across Sessions. *Presence: Teleoperators and Virtual Environments.*

Stoakley, R., Conway, M.J. and Pausch, R. (1995) Virtual Reality on a WIM: Interactive Worlds in Miniature. *Proceedings of CHI'95*, ACM Press, New York, USA, 265-272.

Stone, V.E., (1993) Social Interaction and Social Development in Virtual Environments. *Presence: Teleoperators and Virtual Environments* **2**(2), 153-161.

Ware, C. and Balakrishnan, R. (1994) Reaching for Objects in VR Displays: Lag and Frame Rate. *ACM Transactions on Computer-Human Interaction* **1**(4), 331-356.

Ware, C. and Jessome, D.R. (1988) Using the Bat: A Six-dimensional Mouse for Object Placement. *IEEE Computer Graphics & Applications* **8**(6), 65-70.

Wickens, C.D., Baker, P. (1995) Cognitive Issues in Virtual Reality. *Virtual Environments and Advanced Interface Design* (Furness, T.A. and Barfield, W.), Oxford University Press, New York, 514-542.

Yagel, R., Stredney, D., Wiet, G. J., Schmalbrock, p., Rosenberg, L., Sessanna, D., J., Kurzion, Y., and King, S. (1996) Multysensory Platform for Surgical Simulation. *IEEE Virtual Reality Annual International Symposium 1996*, IEEE Computer Society Press, California, USA, 72-78.