

Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques

I. Poupyrev^{1,2}, S. Weghorst², M. Billinghamurst² and T. Ichikawa¹

¹Information Systems Laboratory, Faculty of Engineering, Hiroshima University, 1 - 4 - 1 Kagamiyama, Higashi-Hiroshima 739, Japan
{poup, ichikawa}@isl.hiroshima-u.ac.jp

²Human Interface Technology Laboratory, University of Washington, Box 352142, Seattle 98195-2142, USA
{poup, weghorst, grof}@hitl.washington.edu

Abstract

The acceptance of virtual environment (VE) technology requires scrupulous optimization of the most basic interactions in order to maximize user performance and provide efficient and enjoyable virtual interfaces. Motivated by insufficient understanding of the human factors design implications of interaction techniques and tools for virtual interfaces, this paper presents results of a formal study that compared two basic interaction metaphors for egocentric direct manipulation in VEs, virtual hand and virtual pointer, in object selection and positioning experiments. The goals of the study were to explore immersive direct manipulation interfaces, compare performance characteristics of interaction techniques based on the metaphors of interest, understand their relative strengths and weaknesses, and derive design guidelines for practical development of VE applications.

1. Introduction

With the rapid increase in performance of high-end computer graphics systems and the transition of 3D graphics onto fast and inexpensive PC platforms, virtual environment (VE) interfaces have become feasible enough to be practically used in areas such as industrial design, data visualization, training, and others¹. Development of useful VE applications, however, requires optimization of the most basic interactions, in particular object manipulation, so users can concentrate on high-level tasks rather than on low level motor activities².

Currently, there is little understanding of how manipulation interfaces should be designed to maximize user performance in immersive virtual environments³. Research that systematically investigates the human factors and design implications of immersive manipulation tasks, 3D devices, interaction metaphors and techniques remains sparse^{2, 4, 5}; consequently, VE designers have had to rely on their intuition and common sense rather than on the guidance of established theory and research results. However, as Brooks⁶ has noted, "in watching many awful interfaces being designed ... I observed that the uninformed and untested intuition of the designer is almost always wrong."

In this paper we present results of an experimental study of two generic interaction metaphors for object selection and manipulation in immersive VEs: *virtual hand* and *virtual pointer*. The specific goals of the study were to 1) compare user performance characteristics of the interaction techniques based on these metaphors, 2) understand their relative strengths and weaknesses, and 3) derive general guidelines to aid designers in practical development of immersive manipulation interfaces. Although object manipulation is among the most important interactions in VEs³, we are not aware of any formal experimental studies that systematically evaluate and categorize interaction techniques for immersive object manipulation. Prior research relates primarily to assessment of user performance as a function of the properties of input and output devices^{7, 8}. In contrast, the focus of this study is on the human factor aspects of different mappings between user input (captured by input devices) and resulting actions in VEs — i.e., interaction techniques³.

The organization of the paper is as follows: After a brief discussion of related work, we introduce a taxonomy of interaction techniques for object manipulation in VEs. This taxonomy categorizes techniques according to their underlying metaphors and provides a rationale for the choice of techniques evaluated in this study. We then de-

scribe the experiments and report their results. Finally, we discuss design implications of the study and directions for future research.

2. Related Work

Object selection and positioning are among the most fundamental interactions between humans and environments, whether it is a “desktop” of 2D direct manipulation interface, 3D virtual environment, or the physical world^{9, 10}. Prior research on manipulation in VEs relates primarily to assessment of user performance as a function of input and display devices and their properties. For example, a pioneering study by Ware¹¹ demonstrated applicability and ease of use of a 3D input device for a six degree of freedom (6DOF) placement task. A study by Zhai and Milgram⁷, comparing isometric versus isotonic input devices in various conditions of spatial manipulation, suggested that isometric devices are preferable for rate control and isotonic for position control. Studies of stereoscopic versus monoscopic display devices suggest that stereoscopy improves user manipulation performance^{12, 13}. The effects of system performance characteristics (such as lag and frame rate) on user manipulation performance have also been extensively studied^{8, 14}.

Investigation of the human factors related to input and output devices has considerable value; however, the lack of systematic research on interaction techniques, which map the user’s actions captured by input devices into resulting actions in the VE³, may significantly limit their appropriate use in VE design. Interaction techniques essentially define the “look and feel” of VEs. A wide variety of techniques can be implemented using the same input devices, and quite a few techniques for spatial manipulation have been demonstrated. Still, there have been few attempts to formally evaluate them, to assess and compare their functional capabilities under various circumstances.

A number of surveys have summarized and classified various approaches for designing techniques for spatial input and identified problems and possible solutions^{3, 15}. Zhai, Buxton and Milgram¹⁶ evaluated application of volumetric semitransparent cursors (“silk” cursors) in a 3D target acquisition task, and reported user performance improvement as compared with traditional cursors. A study by Hinckley, Tullio, Pausch, *et al.*¹⁷ evaluated and compared several spatial rotational techniques. And in a recent study by Mine, Brooks and Sequin¹⁸, automated world scaling techniques were evaluated. More relevant here is the pioneering usability study reported by Bowman and Hodges¹⁰ that evaluated several VE techniques for manipulation at a distance. Although no quantitative data were collected, this study provided useful preliminary observations of techniques.

Starting with early techniques that simply mapped the position and orientation of the user’s hand onto the posi-

tion and orientation of manipulated objects^{19, 20}, the field has been expanding with more sophisticated techniques such as flash light²¹, aperture²², Go-Go^{10, 23}, World-In-Miniature²⁴, image plane²⁵, scaled-world grab¹⁸ and many others. This variety of techniques, however, is also a source of difficulty. How do all these techniques relate to each other? Which interaction techniques should be chosen for particular task conditions? Which among the parameters of interaction techniques, tasks, and environments should be considered to design efficient VE interfaces? These questions persist and merit careful scrutiny by researchers and practitioners.

3. Interaction Techniques for Immersive Object Manipulation

Straightforward evaluation and comparison of manipulation techniques is difficult. There are a multitude of different techniques; even for the same technique performance varies depending on the particular implementation; and studies of a particular technique implementation may not be readily generalized to other implementations of the same technique, thus limiting their external validity.

On the other hand, many techniques apparently relate to each other and share many common properties. For example, there are more similarities between ray-casting and flashlight techniques than there are between ray-casting and techniques that use non-linear mappings to extend the user’s area of reach (as in Go-Go²³). While evaluation of ray-casting might provide insight into techniques similar to ray-casting, such as flashlight, it probably would not help in understanding techniques like the Go-Go. A taxonomy of techniques, classifying them according to their common properties, can be instrumental in understanding relations between techniques and directing their design and experimental evaluation.

3.1. A Taxonomy of Manipulation Techniques

Analysis of current VE manipulation techniques suggests 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 (affordances), and what they cannot do (constraints) when using the technique^{19, 26}. Particular techniques are essentially implementations of the basic metaphors, often extending them in order to overcome some of the metaphor’s shortcomings and constraints. For example, the flashlight technique enhances ray-casting by using a spotlight to ease selection of small objects²¹. These improvements often result in new constraints; for example, with the flashlight technique an ambiguity might occur if several small objects fall into the spotlight²².

In Figure 1 we present a simple classification of current VE manipulation techniques according to their basic inter-

VE MANIPULATION TECHNIQUES

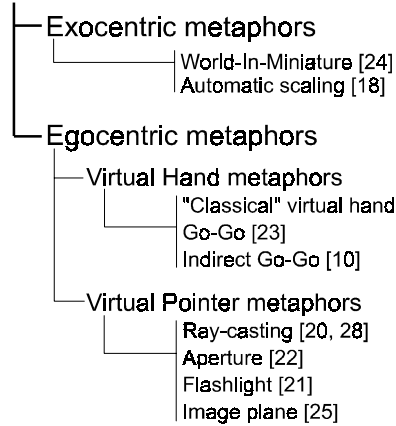


Figure 1: Classification of VE manipulation techniques depending on their underlying metaphors.

action metaphors. We divide the whole variety into exocentric and egocentric techniques. Originating in studies of cockpit displays²⁷, these terms are used now to distinguish between two fundamental frames of reference for user interaction with VEs. With the exocentric interaction, also known as the God's eye viewpoint, users interact with VEs from the outside (the outside-in world referenced display²⁷). An example is the World-In-Miniature technique, which allows manipulation of objects by interacting with their representations in a miniature model of the environment held by the user²⁴. Another example is the automatic scaling technique¹⁸, which scales down the world so the user can access objects located at a distance. Although the exocentric techniques are interesting and important, their evaluation is outside the scope of this work.

With *egocentric* interaction, which is the most common for immersive VEs, the user is interacting from inside the environment — i.e., the VE embeds the user²⁷. Currently there are two basic metaphors for egocentric manipulation: virtual hand and virtual pointer^{3, 10, 25}. With the *virtual hand*, users can grab and position objects by “touching” and “picking” them with a virtual representation of their real hand. The major design factor that defines a particular technique are the choice of input devices and mappings between the real hand's position and orientation and the virtual hand's position and orientation. For example a “classical” virtual hand technique provides one-to-one mapping between the real and virtual hands, while the Go-Go technique²³ employs non-linear mapping functions to extend the user's area of reach (Figure 2).

With the *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 and manipulated³. The major design aspects that distinguish techniques based on this metaphor

are definition of virtual pointer direction, shape of the pointer (selection volume), and methods of disambiguating the object the user wants to select. In the simplest case, the direction of the virtual pointer is defined by the orientation of the virtual hand, the pointer is a “laser ray,” and no disambiguation is provided²⁸ (Figure 3). 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^{22, 25}. Volumetric pointers are also used to ease selection of objects²¹.

This suggested taxonomy identifies only the basic, most general metaphors which can be further subdivided to reflect particular aspects of each technique. Also, techniques based on different metaphors can be combined together to form new manipulation techniques. For example, techniques that combined virtual pointer and virtual hand metaphors have been reported by Bowman and Hodges¹⁰ and by Cohen and Wenzel²⁹.

3.2. Interaction Techniques to Study

The primary goal of the study was to understand the usability characteristics of the virtual pointer and virtual hand metaphors. Thus, for this study we elected to evaluate those techniques that implement these basic metaphors as closely as possible. Focusing on the basic metaphors allows us to limit the number of studied techniques and to generalize results beyond their specific implementations so as the results of the evaluation can be applied to all techniques based on these metaphors. In this section we describe the implementations of techniques that were evaluated in this study.

3.2.1. Virtual pointer metaphor

We used the ray-casting technique for evaluation of the virtual pointer metaphor. Direction of the virtual pointer is defined by position and orientation of the virtual hand (Figure 3). The working volume of the technique is an invisible infinite ray emanating from the user's hand; a short segment of the ray is attached to the hand to indicate the direction of pointing. To select an object, the user points at it and presses a button on the button device. Two variations of the technique were evaluated: with and without visual feedback. When visual feedback is applied, the color of an object changes when the ray intersects with it.

3.2.2. Virtual hand metaphor

Two major variations of the virtual hand metaphor were investigated: the “classical” virtual hand technique and the Go-Go interaction technique. The user is provided with the virtual hand, which position and orientation is controlled by the tracker attached to the user's real hand (Figure 4). To select and pick a virtual object, the user intersects (“touches”) the object with the virtual hand and presses a

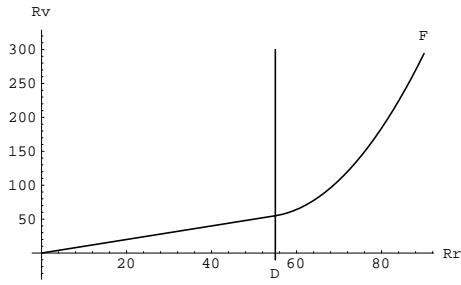


Figure 2: The Go-Go technique: while the real hand is within the distance D ($R_r < D$) the mapping is linear and the movements of the virtual hand correspond to the movements of the real one. When the user extends the hand further than D ($R_r > D$) the mapping becomes non-linear and the virtual arm “grows.”

button on the button device to “pick” the object. The virtual hand uses one-to-one mapping between real and virtual hands, simulating the way we manipulate objects in the real world. In contrast, the Go-Go technique uses a non-linear mapping function to translate the measured distance to the real hand into the controlled distance to the virtual one (Figure 2)²³. This allows for significant expansion of the user’s area of reach. Similarly to the virtual pointer, two variations of both hand techniques were evaluated: with and without visual feedback. With visual feedback, the object changes color when the virtual hand intersects with it.

4. Hypothesis and Objectives of the Study

The current study was designed to compare user performance with these basic interaction techniques on virtual object selection and repositioning tasks. The main variables of interest were *object distance*, *object size* and *visual feedback*. Although the effect of distance and object size on user performance has been extensively studied, prior studies relate either to real world target acquisition³⁰ or non-immersive object manipulation in 2D and 3D user interfaces^{7, 13}. In both cases, manipulation occurs only within the natural reaching distance, as opposed to immersive VEs in which users often need to access and manipulate objects located both nearby and far outside the normal area of reach¹⁸. The task of selecting and manipulating objects located far away is often referred as “action at-a-distance”³. Development of effective means of object manipulation across a wide range of distances has been recognized as an important problem in virtual interface research and development⁵.

Both metaphors evaluated in this study allow for selection at-a-distance. Informal evaluation by Bowman *et al.*¹⁰ suggested that ray-casting might be more effective for object selection, while the Go-Go technique might be superior for object manipulation. Forsberg *et al.*²² reported

difficulties in using a virtual pointer for selection of small objects at-a-distance; no quantitative findings, however, was reported. In this study we systematically evaluated and compared the selection performance of virtual pointer and virtual hand metaphors for objects of different sizes located both close to the user and at-a-distance.

Repositioning objects has been reported as difficult using the virtual pointer techniques^{10, 18}. The classical virtual pointer implementation does not permit change in ray length; therefore, moving the object from a close to a far distance or vice versa can be accomplished only through iterative object picking, moving, releasing, re-picking and so on¹⁸. This method is obviously very inefficient. In our pilot study it was about 3 times slower and required 2.5 times more iterative movements than the Go-Go technique. Are there any cases in which virtual pointer might be an efficient metaphor for object manipulation? In this study, we evaluate and compare virtual pointer and virtual hand metaphors in two conditions: 1) object repositioning at constant distances from the user and 2) object repositioning within the area of reach.

5. The Experiments

Experiments that evaluated three interaction techniques were conducted within the framework of a Virtual Reality Manipulation Assessment Testbed (VRMAT)³¹. The VRMAT is a tool that facilitates rapid design and implementation of a variety of studies of immersive manipulation. It provides definitions of tasks and their properties, suggests experimental procedures including relevant independent and dependent variables, defines metrics and units for their measurements, and so on. In this paper we describe only those aspects of the VRMAT that are relevant to this study.

5.1. Experimental Tasks and Design

Subjects were immersed in a VE consisting of a large checked ground plane and a virtual representation of their hand. Participants wore a 6DOF tracking sensor on their dominant hand and held a button device (used for selecting and picking targeted objects) in the other hand. To reduce the number of variables affecting subjects’ manipulation performance, they were not allowed to move in the VE. We restricted their physical movement by placing them on a platform about 1.5 meter in diameter.

Experimental tasks required subjects to select or position test objects (stimuli) using the technique under investigation. Stimuli for the selection task were solitary virtual objects located in the user’s field of view (Figure 3). After successful selection the test object disappeared, informing the subject that the task was completed. All stimuli were simple geometric objects such as spheres, cubes, and cylinders. More elaborate shapes, such as real world objects, were not used in order that knowledge about their real

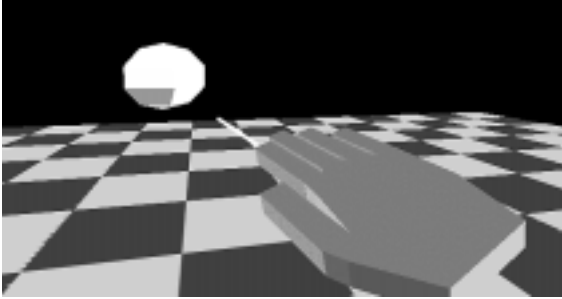


Figure 3: Selection task: the user selects a solitary test object. The ray-casting technique is being evaluated.

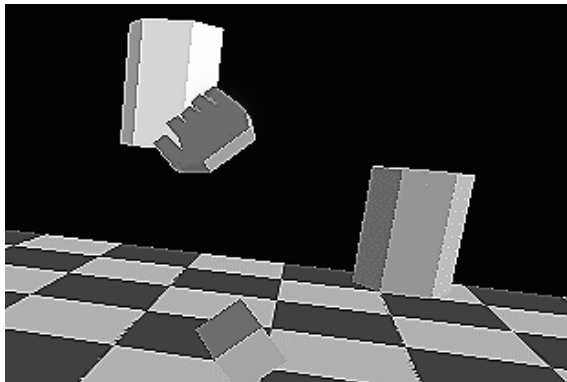


Figure 4: Position task: the user puts a test object on top of the terminal object, indicated by a different color, using the Go-Go technique (the cube in the foreground represents the position of the subject's physical hand).

sizes and proportions would not affect subjects' perception of sizes, proportions and distances in the VE ³².

The positioning task required the subject to pick and place a test object on top of a terminal object indicated by a different color (Figure 4). The shapes for both test and terminal objects were cylinders with equal radii, and subjects were asked to align the manipulated cylinder precisely on top of the other. The positioning could be performed using iterative movements, i.e., subjects could pick, move, and release the object several times. Each time the object was released the VRMAT calculated the error of positioning according to Equation 1. These equations define positioning error as percent of target object displacement relative to the terminal in horizontal and vertical directions. The researcher can control the required accuracy of positioning by specifying maximal vertical and horizontal displacements for task conditions. For example, 0% displacement means that the target object must be aligned on top of the terminal object without any positional error. When the error of positioning falls below the specified threshold the trial is completed and both objects disappear, cueing the subject that the task is successfully accomplished. The next test trial is then presented.

The VRMAT testbed used for the experiments was implemented using a custom VR software toolkit developed as an extension of the Sense8 World Toolkit 6.0. An SGI Onyx RE2 workstation, equipped with a Virtual Research VR4 head-mounted display and Polhemus Fastrak 6DOF sensors, were used. A mouse was used as a button device. The frame update rate was controlled at 15Hz.

$$E_H = \frac{\sqrt{(x_0 - x_t)^2 + (y_0 - y_t)^2}}{D_s} \cdot 100\%$$

$$E_V = \frac{\sqrt{(y_0 - y_t)^2 + (z_0 - z_t)^2}}{H_s} \cdot 100\%$$

Equation 1: The error of positioning, where E_H, E_V — represents horizontal and vertical displacement of the manipulated object relative to the terminal; x_t, y_t, z_t and x_o, y_o, z_o represents coordinates of the target and terminal objects respectively, and D_s, H_s represents diameter and height of stimuli (equal for both the terminal and target objects).

5.2. Independent Variables

The main independent variables of interest for the selection task were *distance* to the object, *object size*, *interaction techniques*, and *visual feedback*. The VRMAT defines objects' positions and sizes in a user-centered coordinate system similar to that used in Kennedy's classic study of the reaching and grasping envelope of seated U.S. Air Force operators³³. Position of a stimulus in VE is defined as the length d and orientation α, β of the vector pointing from the user's chest to the object (Figure 5). Distance d from user to stimulus is defined in terms of *virtual cubits*, a unit of distance introduced in the VRMAT³¹. One virtual cubit is equivalent to the length of the user's maximum reach (Figure 5). It is named after the classic cubit of ancient Rome — the distance between the elbow and the tip of middle finger.

The advantage of using virtual cubits is the ease of generalization of results from experimental studies to practical VE development. A stimulus located at a distance of one virtual cubit in the test environment would be located on the boundary of the user's reach for any user and any other VE, independently from the computational platform and software used. Virtual cubits also eliminate bias due to anthropometrical differences between subjects.

Size of the stimulus is defined as its non-occluded *visual size*: the vertical and horizontal angles φ, ϕ the object occupies in the user's field of view (Figure 5)³¹. Visual angles are also user-centered units. The geometrical size of test objects is recalculated before each trial depending on current position of the user, in order to maintain the objects' visual size as specified a priori by the experimenter. The benefit of visual angles is the separation of influence of distance and object size on user performance: when an

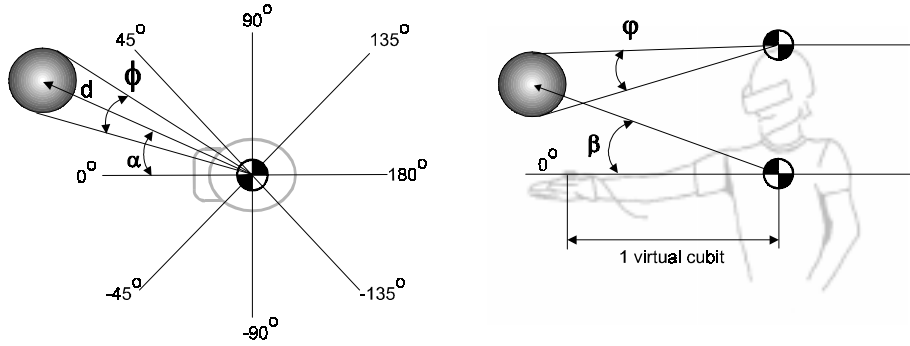


Figure 5: Object position is defined as distance d and direction α , β in user-centered coordinate system. Object size is defined in terms of vertical (ϕ) and horizontal (β) angles of the visual field subtended by the object.

object's size is defined in terms of visual angles, it has the same visual size at different distances. Visual angles also allow for easy generalization of results beyond the particular test VE.

The main independent variables of interest for manipulation tasks were *initial distance* to the stimulus, *distance to the terminal position*, *required accuracy* of positioning, and *interaction technique*. Both initial and final distances are defined in terms of virtual cubits. Required accuracy is defined according to Equation 1.

5.3. Performance Criteria

Completion time, the time taken to successfully accomplish each task, was used as a primary performance criterion. For the selection task this is the time from the moment the stimulus appears until the moment it is successfully selected by the subject. For positioning tasks, completion time is measured from the moment the subject picks a test object until the moment it is positioned with the required accuracy. Because position tasks allow iterative manipulation, i.e., the subject can reposition the object after dropping it, we also measured the *number of iterations* it took to complete positioning, as well as *"net" manipulation time*, i.e., completion excluding the time required for each selection between repositioning. Subjective criteria, such as subject satisfaction, were assessed through post-experimental questionnaires.

5.4. Subjects

Two groups of subjects were recruited from the laboratory subject pool. Ten males and three females served as a subjects for selection task experiments; eight males and four females served as subjects for positioning task experiments. Subjects ranged in age from 19 to 32; all subjects were right handed, as determined by the Edinburgh inventory. In order to reduce the variability in subject performance we chose subjects that had moderate prior experience with virtual reality.

5.5. Procedure

A balanced within-subject (repeated measures) design was used for each task. After donning the HMD subjects were asked to momentarily extend their tracked hand to its full natural reach for "virtual cubit" calibration. The environment then was re-calibrated according to the length of the virtual cubit. Following a demonstration and explanation of the interaction techniques and test tasks, subjects had in average three minutes to practice them. During studies of the selection each subject completed 18 experimental sessions with 15 trials in each session, manipulating each of three different object sizes (4, 6 and 9 degrees) and five different distances (0.7, 1, 2, 4 and 6 virtual cubits). For each of three interaction techniques three sessions were completed with using visual feedback and three sessions without it. Studies of the positioning task consisted of three sessions for each interaction technique, with six trials in each session: four trials for repositioning at constant distances (0.7, 2.2, 3.5, and 6 virtual cubits) and two trials for repositioning close to the user with moderate distance changes (from 0.7 to 1 and from 1 to 0.7). All conditions were defined with 20% required accuracy.

The order of trials within sessions was randomized, trials were presented one after the other, with a four-second delay between them, until all trials were completed. The order of interaction techniques studied was also randomized. In addition to the on-line performance data, an informal questionnaire was administered after completion of the tasks to assess subjects' preferences and opinions.

5.6. Results

5.6.1. Selection task

We begin discussion of our experimental results with a comparison of the usability characteristics of the ray-casting and Go-Go techniques for the object selection task. A repeated measures multiple-way ANOVA was performed with completion time as the dependent variable and distance, size, visual feedback and interaction technique as

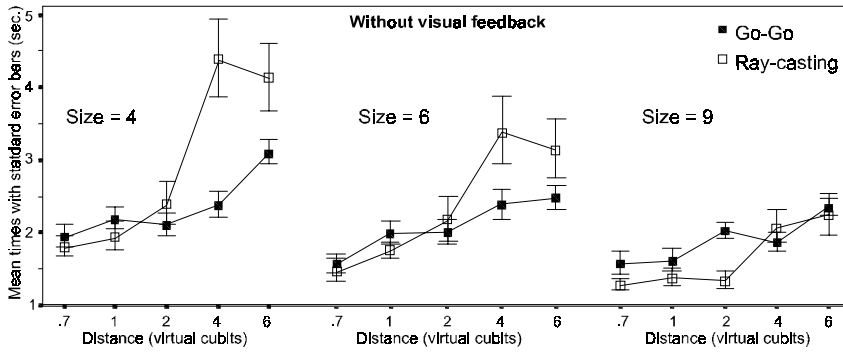


Figure 6: Means of selection time for ray-casting and Go-Go interaction techniques (without visual feedback).

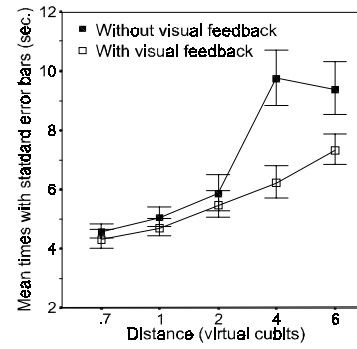


Figure 7: Means of selection time when using the ray-casting (collapsed over object size)

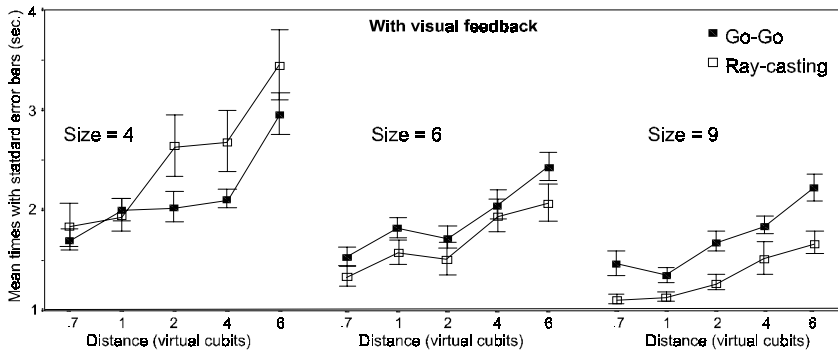


Figure 8: Means of selection time for ray-casting and Go-Go interaction techniques (with visual feedback).

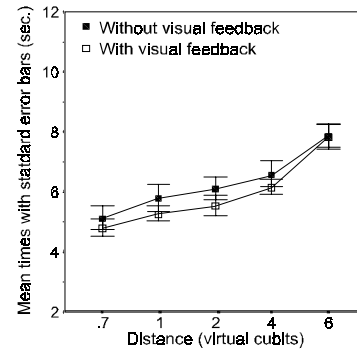


Figure 9: Means for selection time when using the Go-Go (collapsed over object size)

independent variables. A significant main effect was found for distance ($F_{4,48} = 54.23, p < 0.0001$), object size ($F_{2,24} = 92.25, p < 0.0001$) and visual feedback ($F_{1,12} = 15.4, p < 0.002$). And significant interactions between technique and task conditions were also found, i.e., between technique and object size ($F_{2,24} = 47.95, p < 0.0001$), between technique and distance ($F_{4,48} = 6.9, p < 0.0001$), and between technique and visual feedback ($F_{1,12} = 8.19, p < 0.01$). These interactions suggest that neither the Go-Go nor the ray-casting technique was universally preferable in all studied conditions - their comparable weaknesses and strengths depend on the particular task conditions.

Figure 6 summarizes the effects of distance and size on object selection performance for ray-casting and Go-Go techniques without using visual feedback. For both techniques we see that with a decrease in object size or an increase in distance, the target object is increasingly harder to "hit". This conclusion is also supported by our ANOVA analysis (Table 1). This finding is consistent with expectations and appears to represent a "Fitt's Law" phenomenon³⁰. A significant interaction between size and distance was also found for both techniques (Table 1). This interaction suggests that the effect of distance is stronger in those conditions that require more accurate selection, i.e., selec-

tion of smaller objects. This trend is demonstrated in the performance data for the ray-casting technique (Figure 6).

Comparison of the techniques showed that for local selection conditions (within the area of one virtual cubit) both techniques exhibited comparable performance for all object sizes, with slightly better performance for ray-casting (Table 2). However, with increasing distance, the performance of ray-casting was more degraded compared to Go-Go especially when higher selection accuracy was required, i.e., selection of smaller objects. As shown in Figure 6, in selection at-a-distance conditions the Go-Go technique performed significantly better for small objects and exhibited comparable performance for selection of large objects. Table 2 summarizes the effects of interaction technique separately for local selection and selection at-a-distance when the visual feedback was not used.

The introduction of the visual feedback improved selection performance for the ray-casting technique (Figure 7). ANOVA revealed a significant effect due to visual feedback, as well as an interaction between visual feedback and distance for interaction technique (Table 1). Separate analyses for local and remote selection reveal that while visual feedback significantly improves performance for

	Distance (D)	Size (S)	Visual feedback (VF)	Interaction (S*D)	Interaction (S*VF)	Interaction (D*VF)
<i>Ray-casting</i>	$F_{4,48}=28.15$ $p<0.0001$	$F_{2,24}=79.84$ $p<0.0001$	$F_{1,12}=18.3$ $p<0.001$	$F_{8,96}=5.9$ $p<0.0001$	$F_{2,24}=7.97$ $p<0.46$	$F_{4,48}=8.41$ $p<0.0001$
<i>Go-Go</i>	$F_{2,24}=46.8$ $p<0.0001$	$F_{2,24}=71.3$ $p<0.0001$	$F_{1,12}=2.7$ $p<0.13$	$F_{8,96}=4.9$ $p<0.0001$	$F_{2,24}=1.89$ $p<0.17$	$F_{4,48}=7.47$ $p<0.57$

Table 1: The effect of object distance (D), object size (S), and visual feedback (VF) on user selection performance, and interactions between variables for ray-casting and Go-Go techniques.

selection at a distance ($F_{1,12}=16.1$, $p < 0.002$), the effect of visual feedback was not significant for local selection ($F_{1,12}=2.789$, $p < 0.12$). Surprisingly, although visual feedback seems to improve performance of the Go-Go technique (Figure 9) this effect was not statistically significant ($F_{1,12}=2.7$, $p < 0.13$; see also Table 1).

Comparison of the techniques (Figure 8) shows that ray-casting enhanced with visual feedback generally results in better performance, except for conditions when selection of small objects is required (Table 3). When high accuracy of selection is required, such as selection of small objects, the Go-Go technique exhibited better performance than ray-casting ($F_{1,12}=8.96$, $p < 0.01$). Table 3 summarizes the comparison of ray-casting and Go-Go techniques for various object sizes separately for local selection and selection at a distance when the visual feedback was applied.

Finally, we compared the ray-casting and Go-Go techniques with the “classical” virtual hand technique. Because selection with the classical virtual hand is limited by the user’s natural reaching envelope, we compared techniques only in local selection conditions (0.7 virtual cubits) with and without visual feedback. Statistical analysis did not reveal a significant treatment effect due to technique ($F_{2,24}=2.25$, $p < 0.13$). There also was no significant effect of visual feedback on completion time for the classical virtual hand ($F_{1,12}=1.55$, $p < 0.24$).

5.6.2. Positioning task

A repeated measures multiple-way ANOVA was conducted with distance and interaction technique as independent variables and absolute positioning time, “net” positioning time (i.e., with selection time subtracted) and number of iterative positioning movements required to complete a trial as dependent variables. Figure 10 through Figure 12 compare user performance using ray-casting and Go-Go interaction techniques in repositioning objects at constant distances. A significant effect due to distance was found for all dependent variables ($F_{3,33} = 48.5$, $p < 0.0001$, $F_{3,33} = 39.22$, $p < 0.0001$ and $F_{3,33} = 25.83$, $p < 0.0001$, respectively). No significant effect due to interaction technique was found for either absolute or net positioning times ($F_{1,11} = 0.132$, $p < 0.72$ and $F_{1,11} = 0.747$, $p < 0.41$,

respectively). However, a significant effect due to interaction technique was found for number of iterative movements, ($F_{3,33} = 5.47$, $p < 0.039$); the distance by distance analysis reveals a significant effect due to interaction technique at the farthest distance (6 virtual cubits, $F_{1,11} = 11.5$, $p < 0.006$) and does not show a significant effect at closer distances (0.8, 2.2 and 3.5 virtual cubits from the user, $F_{1,11} = 0.576$, $p < 0.46$).

Comparison of the ray-casting, Go-Go and classical virtual hand techniques in object repositioning at close range (within 1 virtual cubit) reveals that when task conditions required the user to bring stimuli closer or move them further away, all techniques exhibited essentially the same performance (absolute positioning time: $F_{2,22} = 2.9$, $p < 0.08$, net positioning time: $F_{2,22} = 1.36$, $p < 0.28$). However, for object repositioning at a constant distance (0.8 virtual cubits) a significant effect of technique was found

	Local (0.7, 1 virtual cubits)	At-a-distance (2,4,6 virtual cubits)
<i>small (4)</i>	$F_{1,12} = 19.40$ $p<0.001$	$F_{1,12} = 9.13$ $p<0.01$
<i>medium (6)</i>	$F_{1,12} = 5.57$ $p<0.04$	$F_{1,12} = 4.66$ $p<0.05$
<i>big (9)</i>	$F_{1,12} = 4.85$ $p<0.048$	$F_{1,12} = 2.47$ $p<0.142$

Table 2: Statistical significance of difference between ray-casting and Go-Go in local and remote selection for various object sizes (without visual feedback).

	Local (0.7, 1 virtual cubits)	At-a-distance (2,4,6 virtual cubits)
<i>small (4)</i>	$F_{1,12} = 0.136$ $p<0.719$	$F_{1,12} = 8.959$ $p<0.01$
<i>medium (6)</i>	$F_{1,12} = 12.1$ $p<0.005$	$F_{1,12} = 4.707$ $p<0.05$
<i>big (9)</i>	$F_{1,12} = 24.1$ $p<0.0001$	$F_{1,12} = 21.819$ $p<0.0001$

Table 3: Statistical significance of difference between ray-casting and Go-Go in local and remote selection for various object sizes (with visual feedback).

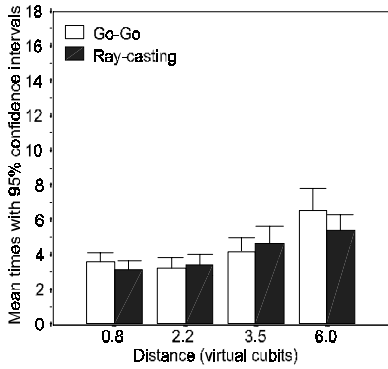


Figure 10: Means of the net time for object repositioning at constant distances.

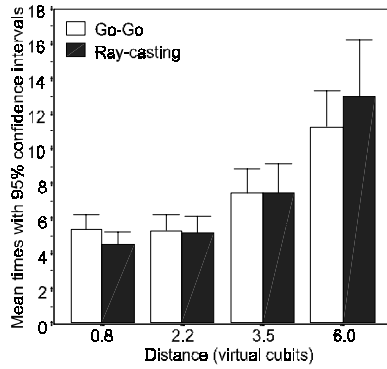


Figure 11: Means of the absolute time for object repositioning at constant distances.

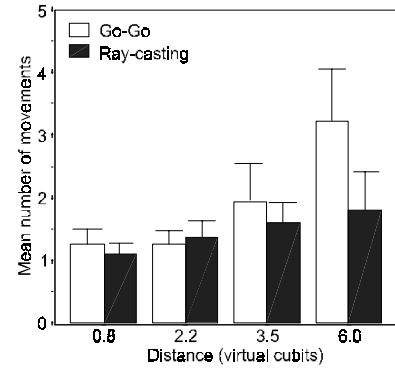


Figure 12: Means of the number of movements for object repositioning at constant distances.

(absolute positioning time: $F_{2,22} = 13.759$, $p < 0.0001$, net positioning time: $F_{2,22} = 8.8$, $p < 0.002$). In these conditions classical virtual hand was 22% faster in absolute positioning than Go-Go technique and 8% faster than ray-casting; ray casting was 15% faster than Go-Go technique. While user performance when using the Go-Go technique was essentially the same for all conditions (absolute positioning time: $F_{2,22} = 0.95$, $p < 0.4$) the ray-casting technique performed significantly better when the positioning did not involve a change of distance (absolute positioning time: $F_{2,22} = 17.77$, $p < 0.0001$).

5.6.3. Subject's comments

While none of the subjects had difficulties in using either Go-Go, ray casting, or virtual hand techniques, in general the Go-Go interaction technique was rated as the most intuitive and enjoyable, with ray-casting second, a finding which replicates results reported previously¹⁰. Three subjects, however, preferred the classical virtual hand, reporting that it was more familiar and more closely simulated real world interaction. All subjects were dissatisfied with decrease of performance of ray-casting in selection of small objects at far distances. These decrease in performance of ray-casting with increasing distance may be due to difficulties with hand-eye coordination, and the magnified effect of tracker noise. Several subjects commented on improvement in selection performance when enhancing ray with visual feedback. Subjects further reported that one of the main difficulties in positioning objects at a distance was the limited visual cues, rather than shortcomings of the techniques themselves. Subjects simply could not see if the object was being positioned correctly.

6. Discussion

These experiments demonstrate that there is no one "best" interaction technique among those studied. The strengths

and weaknesses of the techniques can be compared only in relation to the particular conditions of the spatial manipulation. We discuss below our findings as well as some of the design issues which arise from our studies.

6.1. Virtual Pointer vs. Virtual Hand in Object Selection Task

Performance of techniques based on virtual hand or virtual pointer metaphors depends on the task conditions in which the techniques are used. Within the area of local manipulation all of the techniques we studied demonstrated essentially comparable performance, with ray-casting exhibiting slightly better performance, especially when accurate selection was not required, i.e., selection of big objects. Therefore, in those applications where both local and remote selection are required, the classical virtual hand can be replaced by the Go-Go or ray-casting techniques without degrading user performance in local manipulation conditions. Our finding that there was no significant difference between the classical virtual hand and the Go-Go techniques in local selection conditions suggests that the Go-Go is a generalization of the classical virtual hand technique for selection at-a-distance.

While both ray-casting and Go-Go techniques allow for effective selection of objects at-a-distance, the Go-Go technique resulted in better performance when accurate selection is required, i.e., in selection of small objects. Ray-casting was found more efficient than Go-Go when high accuracy of selection is not required, i.e., selection of big objects. Introduction of visual feedback significantly improves the accuracy of ray-casting and is an important enhancement for this technique. However, even with visual feedback, the Go-Go was still faster in selection of small objects (Figure 8). The choice of technique for selection at-a-distance depends, therefore, on the accuracy of selection required in a particular application.

6.2. Virtual Pointer vs. Virtual Hand in Object Positioning Task

It is difficult to compare virtual pointer and virtual hand metaphors. Techniques based on virtual pointer metaphor do not allow for natural manipulation of the object distance, unless they are extended with some mechanism which allows the user to manipulate the length of the virtual pointer^{10, 18}. Virtual pointer, however, is an effective and efficient metaphor in those conditions where virtual pointer and virtual hand can be compared, i.e. in repositioning objects at a constant distance and object repositioning close to the user. Indeed, subject performance for object repositioning at constant distances using ray-casting was comparable to their performance using the Go-Go technique both locally and at-a-distance (Figure 10, Figure 11). For repositioning within the area of reach, where the change of distance can be easily accomplished, the ray-casting and Go-Go demonstrated comparable performance. Generally at close distances, the ray-casting was more efficient when a change of object distance was not required, while the Go-Go technique resulted in the same performance for all conditions of local manipulation.

6.3. Visual Feedback

Enhancing interaction techniques with visual feedback does not always improve user performance. Visual feedback did considerably improve ray-casting performance for selection at-a-distance, making selection of small objects significantly easier. The effect of visual feedback, however, was not significant in local selection conditions (Figure 7), and there was no significant effect of visual feedback for the Go-Go interaction technique either in local selection or at-a-distance. This result was somewhat surprising. Previous evaluation of the Go-Go technique indicated that because of the non-linear mapping used in the techniques, an increase in object distance often leads to "overshoot" of objects located far away (see Figure 2). Consequently, we expected that visual feedback would improve the Go-Go performance at far distances by minimizing this overshoot, however, this did not happen. One possible explanation for this result is that with the techniques based on the virtual hand metaphor the user can naturally see when the virtual hand intersects the object. Therefore, visual feedback is an inherent part of the virtual hand metaphor. Consequently, adding more visual feedback does not necessarily result in significant performance improvements. A second explanation is that because the VRMAT testbed defines the size of the objects in terms of visual angles, moving objects further from the user naturally increases their "real" geometrical size in order to maintain the same visual size³¹. Since the VRMAT uses symmetrical objects for stimuli, an increase in geometrical size leads to an increase in the volume of the stimulus which, in turn, counterbalances the effect of the overshoot. This situation, in fact, is very natural for interaction in VEs: in order for the object to be visible at a great distance

its geometrical size should be quite large. An important exception is in selection of flat objects. The depth of the stimulus, is therefore, another important variable which should be considered when evaluating techniques for immersive object selection.

6.4. Metaphor Affordances and Constraints: 2D vs. 3D Manipulation

The purpose of this study was to investigate two basic metaphors for immersive manipulation: virtual hand and virtual pointer. Our findings suggest that their basic affordances and constraints are defined by the number of degrees of freedom which can be effectively manipulated using the techniques based on these metaphors.

Indeed, the essence of immersive object selection and manipulation is a specification of a three-dimensional position within the virtual environment using the interaction techniques provided by the system⁹. Within the user-centered coordinate system used in this study, the position of an object is defined as three coordinates: distance to the object and two angles, pitch and yaw, that define direction to the object from the user's point of view³¹ (Figure 5). Results of our experiments suggest that while the virtual hand allows to effective manipulation in all three coordinates, the virtual pointer allows effective manipulation on only two of them: pitch and yaw (angles α and β in Figure 5); the virtual pointer technique is less effective in manipulating the third degree of freedom - the distance to an object. For example, the object selection performance decrement observed with an increase in object distance is significantly worse for ray-casting than for the Go-Go technique; ray-casting was inefficient for positioning tasks that required a change in distance; it was, however, very effective for repositioning at a constant distance. Thus, the virtual hand and virtual pointer can be categorized as 3D and 2D direct manipulation metaphors, respectively.

One of the interesting design implication arising from the proposed categorization is that the 2D nature of the virtual pointer makes the well-developed guidelines and techniques from 2D graphical user interfaces design suitable for development of effective immersive ray-based interaction dialogs, such as virtual menu systems. Indeed, as long as objects are located around the user at the same distance, ray-casting would be sufficient for effective interaction. In effect depending on the distance, the locus of interaction of virtual pointer forms a continuum of spherical surfaces of different size around the user.

Certainly, the virtual pointer metaphor can be enhanced to allow more direct control of the distance variable. For example, Bowman et al.¹⁰ extended classical ray-casting with a "fishing reel" metaphor: the user can change the length of the virtual pointer by pressing two additional buttons. The user performance implications of the metaphor extensions, however, are not clear. Will the reeling

mechanism provide performance comparable to the Go-Go technique in repositioning tasks? Would it improve the performance of the ray-casting technique in selection of small objects? Can we be sure that enhancing the metaphor would not degrade user performance in some task conditions? These questions are subject to systematic and careful human factors evaluation.

6.5. World-Centered vs. User-Centered Design of Virtual Interaction

Prior research and development of virtual user interfaces has been geared toward development of effective interaction techniques and tools. The developers' task was to create interaction techniques that would allow the user to interact effectively in any given virtual environment. Our findings, however, suggest that even for basic tasks, such as selection of virtual objects, and for a limited number of task variables, such as object size and distance, development of a single universally efficient technique is difficult, if not impossible. Thus, instead of developing of new interaction techniques, researchers and developers can take another route: improve the spatial design of VEs to allow for optimal performance using existing techniques. Following existing terminology we call these two approaches, respectively, *world-centered* and *user-centered* design of virtual reality interfaces. The categorization of VE design methods as user- or world-centered is, certainly, a generalization which cannot be practically implemented in its pure form. For some applications it is not possible to design the VE around the available techniques. Nevertheless, there are many application domains where designers do have the freedom to fit the VE to the interface; for instance, in many information visualization applications. Practical VE system development should probably use methods and principles based on both approaches, depending on the purpose of the particular application.

7. Conclusions

The growing acceptance of VE technology will require more attention to optimize immersive interaction in order to maximize user performance. This study systematically explores one of the most important aspect of immersive interaction - interaction metaphors for selecting and positioning objects in VEs. The paper presents an original taxonomy of interaction techniques for immersive manipulation, describes the methodical framework used in the experimental study, reports experimental results and draws design implications for the practical development of manipulation interfaces for VEs.

The research reported here is just a small step toward understanding human factors behind manipulation in VEs. Future studies of VE manipulation should further investigate the design aspects of the particular techniques and their influence on user performance; assess usability of the techniques in other conditions of manipulation tasks; in-

vestigate combinations of manipulation and navigation techniques; and explore possible ways to integrate various techniques into seamless and intuitive interaction dialogues.

8. Acknowledgments

This research was partially sponsored by the Air Force Office of Scientific Research (contract #92-NL-225) and a grant from the HIT Lab Virtual Worlds Consortium. The authors want to especially thank Jennifer Feyma for her help with experiments. We would also like to thank Edward Miller, Jerry Prothero, Hunter Hoffman, Doug Bowman, Prof. Hirakawa and all subjects who participated in the experiments.

9. References

1. M. Göbel, "Industrial applications of VEs", *IEEE Computer Graphics&Applications*, **16**(1), pp. 10-13 (1996)
2. K. Stanney, "Realizing the full potential of virtual reality: human factors issues that could stand in the way", *Proceedings of VRAIS'95*, pp. 28-34 (1995)
3. M. Mine, "Virtual environment interaction techniques". UNC Chapel Hill CS Dept., Technical Report TR95-018 (1995)
4. K. Herndon, A. van Dam and M. Gleicher, "The challenges of 3D interaction: a CHI'94 workshop", *SIGCHI Bulletin*, **26**(4), pp. 36-43 (1994)
5. N. Durlach and A. Mavor, eds. *Virtual reality: scientific and technological challenges*, National Academy Press: WA, pp. 542, (1995).
6. F. Brooks, "Grasping reality through illusion - interactive graphics serving science", *Proceedings of CHI'88*, pp. 1-11 (1988)
7. S. Zhai and P. Milgram, "Human performance evaluation of manipulation schemes in virtual environments", *Proceedings of VRAIS'93*, pp. 155-61 (1993)
8. B. Watson, V. Spaulding, N. Walker and W. Ribarsky, "Evaluation of the effects of frame time variation on VR task performance", *Proceedings of VRAIS'96*, pp. 38-52 (1996)
9. D. Foley, V. Wallace and V. Chan, "The human factors of computer graphics interaction techniques", *IEEE Computer Graphics & Applications*, (4), pp. 13-48 (1984)
10. D. Bowman and L. Hodges, "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments", *Proceedings of Symposium on Interactive 3D Graphics*, pp. 35-38 (1997)

11. C. Ware, "Using hand for virtual object placement", *Visual Comp.*, **5**(6), pp. 245-253 (1990)
12. E. Spain and K. Holzhausen, "Stereoscopic versus orthogonal view displays for performance of a remote manipulation task.", *Proceedings of Stereoscopic Displays and Applications II*, SPIE, pp. 103-110 (1991)
13. J. Boritz and K. Booth, "A study of interactive 3D point location in a computer simulated virtual environment", *Proceedings of VRST'97*, pp. 181-187 (1997)
14. I. MacKenzie and C. Ware, "Lag as a determinant of human performance on interactive systems", *Proceedings of INTERCHI'93*, pp. 488-493 (1993)
15. K. Hinckley, R. Pausch, J. Goble and N. Kassell, "A survey of design issues in spatial input", *Proceedings of UIST '94*, pp. 213-22 (1994)
16. S. Zhai, W. Buxton and P. Milgram, "The "Silk cursor": investigating transparency for 3D target acquisition", *Proceedings of CHI'94*, pp. 459-464 (1994)
17. K. Hinckley, J. Tullio, R. Pausch, D. Proffitt and N. Kassell, "Usability analysis of 3D rotation techniques", *Proceedings of ACM UIST'97*, pp. 1-10 (1997)
18. M. Mine, F. Brooks and C. Sequin, "Moving objects in space: exploiting proprioception in virtual-environment interaction", *Proceedings of SIG-GRAPH'97*, pp. 19-26 (1997)
19. C. Ware and D.R. Jessome, "Using the bat: a six-dimensional mouse for object placement", *IEEE Computer Graphics & Applications*, **8**(6), pp. 65-70 (1988)
20. R. Bolt, "'Put-that-there': voice and gesture at the graphics interface", *Computer Graphics*, **14**(3), pp. 262-270 (1980)
21. J. Liang, "JDCAD: A Highly Interactive 3D Modeling System", *Computers and Graphics*, **18**(4), pp. 499-506 (1994)
22. A. Forsberg, K. Herndon and R. Zeleznik, "Aperture based selection for immersive virtual environment", *Proceedings of UIST'96*, pp. 95-96 (1996)
23. I. Poupyrev, M. Billinghurst, S. Weghorst and T. Ichikawa, "Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR", *Proceedings of UIST'96*, pp. 79-80 (1996)
24. R. Stoakley, M. Conway and R. Pausch, "Virtual reality on a WIM: interactive worlds in miniature", *Proceedings of CHI'95*, pp. 265-272 (1995)
25. J. Pierce, A. Forsberg, M. Conway, S. Hong, R. Zeleznik and M. Mine, "Image plane interaction techniques in 3D immersive environments", *Proceedings of Symposium on Interactive 3D Graphics*, (1997)
26. T. Erickson, "Working with interface metaphors" In *The art of human-computer interface design*, B. Laurel, Editor, Addison-Wesley Publishing Company. pp. 65-73 (1990)
27. C.D. Wickens and P. Baker, "Cognitive Issues in Virtual Reality" In *Virtual Environments and Advanced Interface Design*, T.A. Furness and W. Barfield, Editors, Oxford University Press: New York, NY. pp. 514-542 (1995)
28. R. Jacoby, M. Ferneau and J. Humphries, "Gestural Interaction in a Virtual Environment", *Proceedings of Stereoscopic Display and Virtual Reality Systems: The Engineering Reality of Virtual Reality*, pp. 355-364 (1994)
29. M. Cohen and E.M. Wenzel, "The Design of Multidimensional Sound Interfaces" In *Virtual Environments and Advanced Interface Design*, T. Furness and W. Barfield, Editors, Oxford University Press: New York, NY. pp. 291-348 (1995)
30. P. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement", *Journal of Experimental Psychology*, (47), pp. 381-391 (1954)
31. I. Poupyrev, S. Weghorst, M. Billinghurst and T. Ichikawa, "A framework and testbed for studying manipulation technique for immersive VR", *Proceedings of VRST'97*, pp. 21-28 (1997)
32. B. Gillam, "The perception of spatial layout from static optical information" In *Perception of Space and Motion*, Academic Press. pp. 23-67 (1995)
33. K. Kennedy, "Reach capability of the USAF population: Phase I. The outer boundaries of grasping-reach envelopes for the short-sleeved, seated operator". USAF, AMRL: Technical Report TDR 64-56 (1964)