CHAPTER V

SELECTION AND MANIPULATION

5.1 Introduction and Definitions

Once a VE user has been given the ability to move about the 3D space effectively, via a viewpoint motion control technique, the next step is to interact with the environment in some way. In a mechanical design application, this might mean positioning various parts so that they fit together. A training system for rescue workers might require the user to open doors, move obstacles, or make use of tools. A virtual science lab could allow the user to build molecules from components or position electrical charges. All of these interactions fall under the broad heading of selection and manipulation.

Selection involves the specification of one or more virtual objects by the user for some purpose. The purpose might be to specify the object of a command (e.g. delete the selected object), to invoke a command (e.g. selecting a menu item), to change the system state (e.g. selecting a toggle switch that controls a rendering parameter), or to choose a new tool (e.g. selecting a tool that creates cubes). Often, however, selection is performed to set up *manipulation*, that is, setting the position and/or orientation of a virtual object. Obviously, unless the user is constantly manipulating a single object, she must first select the object she wishes to manipulate.

Since many VE developers believe that the best way for the user to interact with a VE is the most natural way (a position we do not hold), many VE systems utilize a naive *natural mapping* for selection and manipulation. The natural mapping simply maps the scale and location of the user's physical hand directly to the scale and location of a virtual hand, so that when the virtual hand touches an object in the VE, it may be selected, and selected objects are manipulated by attaching them to the virtual hand – in other words, the user simply reaches out and grabs an object to select or manipulate it. This basic metaphor has been extended so that users can have fingertip control of virtual objects (Kijima and Hirose, 1996).

The natural mapping does have the advantage that it is quite intuitive for almost all users, since it replicates the physical world. However, this metaphor is simply not powerful enough for most VE applications. First, the objects that may be selected are only those within a physical arm's reach of the user, and once an object is selected, it may only be manipulated within that relatively small space. This may not be a problem if the work environment is only the size of a tabletop, but makes manipulation in larger environments difficult. To allow selection of faraway objects or large-scale movement of objects, a travel technique must be used in conjunction with the natural mapping.

Secondly, manipulation of large objects is problematic with the natural mapping. In the physical world, the objects that we can manipulate in our hands are limited to a certain size, but there are no such restrictions in the virtual world. Imagine a city planning application where the user wished to reposition a skyscraper. If the user was within an arm's length of the building, it would inevitably obscure the user's view, so that precise placement would be impossible.

When careful consideration is taken, it should be obvious that a real-world technique would be inadequate for selection and manipulation tasks in VEs, since the tasks we wish to perform go beyond our real-world capabilities. In the same way, a travel technique based on physical walking will be completely inadequate if the application requires travel on a global scale. The power of VEs is not to duplicate the physical world, but to extend the abilities of the user to allow him to perform tasks not possible in the physical world. For these reasons, we will consider in this chapter techniques for selection and manipulation that go beyond the natural mapping. In particular, the techniques will allow selection of objects at a distance, and manipulation within a large space.

5.2 Related Work

5.2.1 Interaction Metaphors

A variety of interaction techniques have been proposed and implemented which address the problem of selecting and/or manipulating objects within a virtual space. Among techniques which can select and manipulate faraway objects, most techniques fall into three categories: arm-extension, ray-casting, and image plane techniques.

Arm-extension techniques address the problem of the user's limited reach directly – they allow the user to extend her virtual hand much farther than her physical hand, so that faraway objects can be "touched." An advantage of such techniques is that manipulation can still be done via hand motion, as in the natural mapping. However, selection of objects that are very far away or small may be difficult, because the hand must be positioned precisely. Such techniques differ in the way that the virtual arm is extended. Some map the physical hand motion onto virtual hand motion using a mapping function (Poupyrev et al, 1996). Others use more indirect means to extend and retract the virtual arm (Bowman and Hodges, 1997). Still others employ more arcane mapping functions, such as from physical hand position to virtual hand velocity (Bowman and Hodges, 1997).

Ray-casting techniques select faraway objects by extending an idea from the 2D desktop metaphor. Just as one positions the pointer over an icon on the desktop to select it, so in three-dimensions one can point a virtual light ray into the scene to intersect and select a virtual object (Mine, 1995). Generally, the direction of the light ray is specified by the orientation of the user's hand (e.g. the ray emanates from the user's outstretched index finger), so that selection becomes a simple task of pointing at the desired object. The common manipulation scheme is to attach the object to the light ray at the point of intersection, but this makes manipulation unwieldy (Bowman, 1996), so other manipulation schemes may be desired.

Image plane techniques (Pierce et al, 1997) are a combination of 2D and 3D interaction. Selection of objects is done, as the name suggests, in the viewplane, so that the dimension of depth into the scene is not considered. For example, in one technique the user selects an object by partially occluding it with his virtual hand. That is, the virtual hand covers the desired object in the displayed image. Actually, this is a ray-casting technique, since one can consider it to use a ray emanating from the user's eyepoint and going through the virtual hand position to select an object, but we list these techniques separately, preferring that the term "ray-casting" be used for pointing techniques where the ray emanates from the virtual hand. Again, selection is simple for these techniques, but manipulation of objects once they are selected is an open question. Pierce et al's implementation (1997) scales the user so that the virtual hand actually touches the selected object, at which point natural hand movements can be used to manipulate the object. When the object is released, the user is scaled back to normal size.

Finally, there are certain techniques which do not fit into any of these categories. Rather, they try to maintain the intuitiveness of the natural mapping while overcoming its inherent limitations by employing the natural mapping in a manner not consistent with the physical world (perhaps we could call these "unnatural mappings"). One of the most obvious of these techniques is to employ a scaling factor (make the user larger or the world smaller) so that the user can reach any object with the virtual hand. Mine, Brooks, and Sequin (1997) use scaling together with a framework that allows the user to exploit his proprioceptive sense for navigation and manipulation. This can be a powerful metaphor, but may also have side effects for viewing the effects of changes – since the scale of the user and world are different, a small motion by the user results in a large motion in the world. Another idea employing scaling is to have two copies of the world, one large and one small. In the World in Miniature (WIM) technique (Stoakley, Conway, and Pausch, 1995), the user manipulates small objects in a "dollhouse" world held in the hand, and the corresponding full-size objects move accordingly. This has been extended in the recent "voodoo dolls" technique (Pierce, Stearns, & Pausch, 1999), in which the user creates his own miniature parts of the environment (dolls), and may use two hands to manipulate these doll objects relative to one another.

We also note that a good deal of work has been done in the area of aiding the user to position objects correctly. Most of these methods use some type of constraints to reduce the number of degrees of freedom that must be controlled by the user, or to reduce the required precision on the part of the user. For example, one can constrain an object to move only in one dimension (Bowman and Hodges, 1995), model an object's collisions with other parts of the world (Kitamura, Yee, and Kishino, 1996) or place some intelligence in the object so that it naturally seeks to be aligned correctly with the world and other objects (Bukowski and Sequin, 1995).

5.2.2 Evaluation of Techniques

There has been little work in the evaluation of selection and manipulation techniques for immersive VEs, but some studies have been reported in the areas of 3D selection and manipulation. Ware (Ware and Jessome, 1988, Ware and Balakrishnan, 1994) has carried out several investigations into the use of a tracked hand or input device for object placement in 3D environments. Also, Zhai and Milgram (1993) compared different input devices in a principled manner based on a proposed taxonomy of manipulation in 3D space.

One piece of work in immersive VEs deserves special mention. Poupyrev (1997) has implemented a "testbed" for the evaluation of selection and manipulation schemes, which incorporates our goals of systematic evaluation and multiple performance measurements. Unlike our proposed testbed, however, this work is more of a tool for those who would wish to perform experiments to compare various techniques. The user of the system can design and implement experiments quickly based on a text description of the interaction techniques, outside factors, and performance measurements. Our testbed, on the other hand, is a more generalized set of experiments that attempts to model all of the important variables and measurements.

5.3 Initial Evaluation and Design

Our first work in this area was inspired by a talk given at SIGGRAPH '96 on a new interaction technique for virtual object manipulation: the Go-Go technique (Poupyrev et al, 1996). The technique seemed intuitive and easy-to-use, and it promised to have wide application. However, no indications of performance were given, and no studies compared this technique with the many others that had been proposed for the same task. The technique had novelty and elegance, but we felt that this was not enough to proclaim it a cure-all. It needed to be tested and understood.

Therefore, we produced our own implementation of the Go-Go technique and several others and evaluated them with a simple user study (Bowman and Hodges, 1997). Our goal was to understand the characteristics of the task and the techniques, in an attempt to discover what makes a technique "good" for virtual object manipulation.

5.3.1 Techniques Considered

The techniques we studied fell into two categories: arm-extension and ray-casting. As we have noted, arm-extension techniques, including Go-Go, allow the user to select faraway objects by providing a mechanism by which the virtual arm may be made much longer than the physical arm. Users can then manipulate the objects directly with their hand, in a natural manner. Ray-casting techniques (Mine, 1995), on the other hand, use a pointing metaphor. A virtual light ray extends from the user's hand, and objects are selected by intersecting them with the light ray. The object is attached to the light ray for manipulation.

Within each of these categories, we investigated several variants. For arm-extension techniques, the main differentiator was the mapping technique used to determine the length of the virtual arm. The mapping function for the Go-Go technique, shown in figure 5.1, has two parts. When the user's physical hand is within a threshold distance D from the body, there is a one-to-one relationship between physical and virtual arm length. However, outside this threshold, the virtual arm length follows a non-linear function relative to the distance of the physical arm from the user's body.

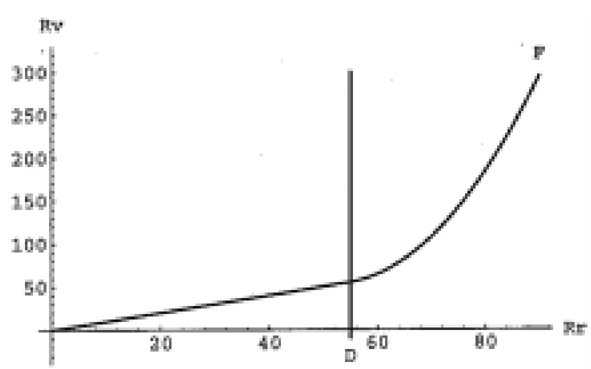

Figure 5.1 Mapping Function for the Go-Go Technique: R_r=Physical Hand Distance, R_v=Virtual Hand Distance. Reproduced from (Poupyrev et al, 1996)

We also looked at two other mapping functions. One is similar to Go-Go, except that there is no area of one-to-one growth – the virtual arm grows according to the non-linear function at every position ("fast Go-Go"). This allows the user's reach to extend to a greater, though still bounded, distance.

Second, we explored the possibility of mapping physical hand position to virtual hand *velocity*, in a technique we called "stretch Go-Go." This was done by defining three concentric regions of space about the user. When the physical hand is within the medium-range region, the virtual arm length is constant. If the physical arm is stretched far from the body, into the outer region, the virtual arm grows at a constant rate. Similarly, with the physical hand in the inner region, the virtual arm shrinks at a constant rate. This has the advantage that the user can reach any object, no matter its distance. To help the user visualize the mechanism, we provided a graphical gauge showing the three regions and the user's current hand position (figure 5.2).

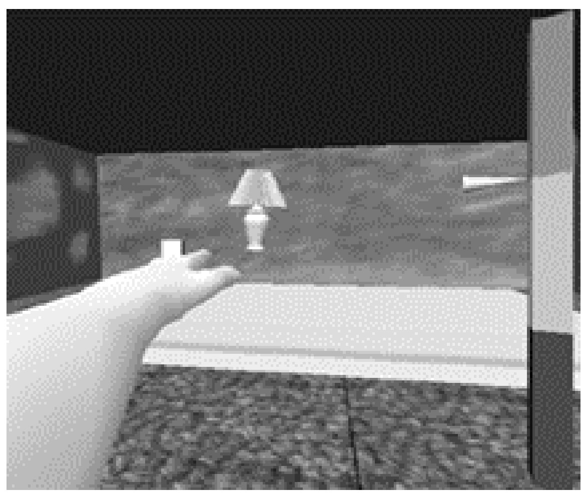

Figure 5.2 Stretch Go-Go Technique, with Gauge

Finally, we considered a technique that does not use a mapping function at all, but rather specifies the virtual arm length in a more indirect manner. This technique simply uses two mouse buttons to grow or shrink the virtual arm at a constant rate. Again, this technique has unlimited reach, although it may lack the intuitive characteristics of techniques where the arm is stretched out to make it longer.

We also included two ray-casting techniques in our survey. Both techniques use the same virtual light ray idea for object selection, and both manipulate the object by attaching it to the light ray. The techniques differ in their expressive power. With the basic ray-casting technique, there is no way to change the distance of the object from the user – the object must move along a sphere centered at the user whose radius is the object's original distance from the user. Thus, in the second of these techniques, we added a "reeling" feature, which allows the user to move the object closer or farther away along the light ray, similar to reeling a fishing line in or out.

5.3.2 User Study

Armed with these six techniques (four arm-extension and two ray-casting), we conducted a simple user study to assess their performance and applicability. Eleven student volunteers (two females and nine males) participated in the study. The equipment used included a Virtual Research VR4 head-mounted display, Polhemus Fastrak trackers, an SGI Indigo2 Max Impact, and a custom-built 3-button joystick. Users were immersed in a virtual room containing several pieces of furniture and given several minutes to practice and use each of the six techniques.

We did not collect any quantitative data in this study, but instead observed the performance and errors of the users, and collected their comments about the relative merits of each of the interaction techniques. This information led to a much more thorough understanding of the tasks of selection and manipulation, and of the techniques themselves.

None of the six techniques proved adequate for selection and manipulation of faraway objects. The favorite techniques were Go-Go and the indirect arm-extension technique, but problems were noted with each of these as well. There were difficulties with precision of selection, precision of manipulation, speed of use, user comfort, and expressiveness of the technique. We made three general observations about the tasks and techniques, which can be expressed as guidelines (hypothesis 2).

First, naturalism is not always a necessary component of an effective technique. Users almost unanimously found Go-Go to be the most natural technique, but many evaluators preferred other techniques. Indirect stretching was more effective for several subjects because it offered more precise control of the hand location, and less physical work on the part of the user. Several users also liked ray-casting with reeling because of the lack of physical effort required: they could support their arm and simply point with their wrists and press joystick buttons. This goes against common intuition regarding VE interaction: the most natural technique is not always the best in terms of performance or preference. This indicates that more formal methods are necessary to determine appropriate ITs (hypothesis 1).

Second, physical aspects of users were important in their evaluation of the techniques. For example, those users with shorter arms were less likely to prefer the go-go technique because their reach was more limited. Also, all of the arm-extension techniques depend on the specification of a point at the center of the user's torso. The virtual hand in these techniques is kept on a line defined by this torso point and the location of the physical hand. Although we defined this point relative to the user's head position, the height of the user made a difference. If the torso point is not approximated well, the hand will appear lower or higher than it should be, and grabbing and manipulation will be more difficult. In short, techniques that are dependent on the user will require user modeling in order to be most effective.

Our most important finding, however, was that grabbing and manipulation must be considered separately for overall usability. Although only two of our users preferred a ray-casting technique overall, almost every user commented that it was easier to grab an object using ray-casting than with any of the arm-extension techniques. This result agreed with our earlier observations on the use of ray-casting in VE applications (Bowman, 1996, Bowman, Hodges, and Bolter, 1998). It requires no arm stretching and less precision on the part of the user: one simply points the ray and releases the button. With the arm-extension techniques, one must place the hand within the object, which can be quite difficult at a great distance or when a small physical motion maps to a large translation of the virtual hand.

On the other hand, no users preferred ray-casting techniques for object manipulation, as arbitrary rotations of an object are practically impossible using these techniques. With an arm-extension technique, objects can be rotated in their own coordinate system, and their position can be controlled easily as well. None of the current techniques, then, were universally acclaimed, because none of them were easy to use and efficient throughout the entire interaction: grabbing, manipulating, and releasing the object.

5.3.3 HOMER Technique

In response to these results, it was clear that a hybrid technique combining the best features of both the arm-extension and ray-casting metaphors could provide gains in efficiency, accuracy, and usability. This simple observation led to the implementation of the HOMER (Hand-centered Object Manipulation Extending Ray-casting) family of techniques. These techniques simply use the better-performing metaphor for each part of the task: ray-casting for object selection and in-hand object manipulation.

The basic technique works like this (see figure 5.3): the user activates the virtual light ray and intersects the desired object with it by pointing, just as in the ray-casting technique. Upon releasing the button, the virtual hand immediately moves to the center of the selected object, so that manipulation can be performed directly with the hand, and so that any rotation can be achieved. When the drop command is given, the virtual hand returns to the location of the physical hand.

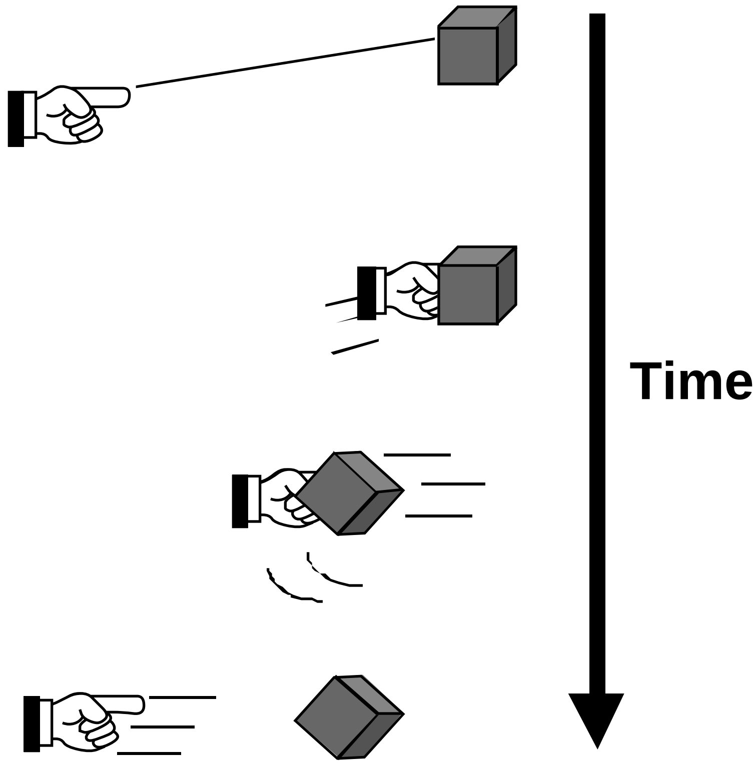

Figure 5.3 Time Sequence of the HOMER Technique

The HOMER techniques exhibit both ease of selection and ease of manipulation, since they use well-performing technique components for both of these tasks. There is one issue that must be addressed, however, to make the HOMER techniques completely expressive (that is, to ensure that they allow a user to place an object at any position and orientation). This is the question, again, of object distance from the user. In the basic HOMER technique, hand motions are mapped one-to-one onto the object, so there is no way the object could be placed twice as far away from the user, or brought very near for inspection. Thus, we need a mechanism for controlling object depth once the object has been selected.

We provide two such mechanisms, one direct and one indirect. The indirect HOMER technique simply uses the “reeling” feature discussed earlier, where two mouse buttons are used to move the object nearer or farther away. This provides complete expressiveness, but may be slow or cumbersome. The direct HOMER technique uses a linear mapping function to control object depth. A linear function was chosen because it is more predictable and easier to control than a non-linear function, no matter the distance from the user. The virtual object moves N meters in or out for every one meter of physical hand motion in or out, where N is the ratio between the original object-to-user distance and the original hand-to-user distance. Therefore, if the user moves his physical hand twice as far away from his body, the object will move to twice its original distance from the body as well.

This technique also allows the user to have direct control of the mapping function, since it depends on the distance between the user's physical hand and her body at the time of selection. If a large N is needed, the user can place her hand very close to her body, but if more control is desired, the hand can be positioned farther away.

5.4 Formal Evaluation Framework

5.4.1 Categorization of Techniques

The initial user study provided us with a good understanding of the tasks of selection and manipulation, and of the space of possible techniques for realizing these tasks. Our original categorization of techniques into arm-extension, ray-casting, image-plane, and "other" techniques is useful at a high level, but there may be large performance differences within a category. Therefore, this categorization does not allow us to make generalizations such as, "arm-extension techniques provide greater accuracy of placement," since this depends on the implementation of the arm-extension technique.

Therefore, we have re-categorized ITs for selection and manipulation based on a more formal task analysis, as we did for travel techniques. This taxonomy is shown in figure 5.4.

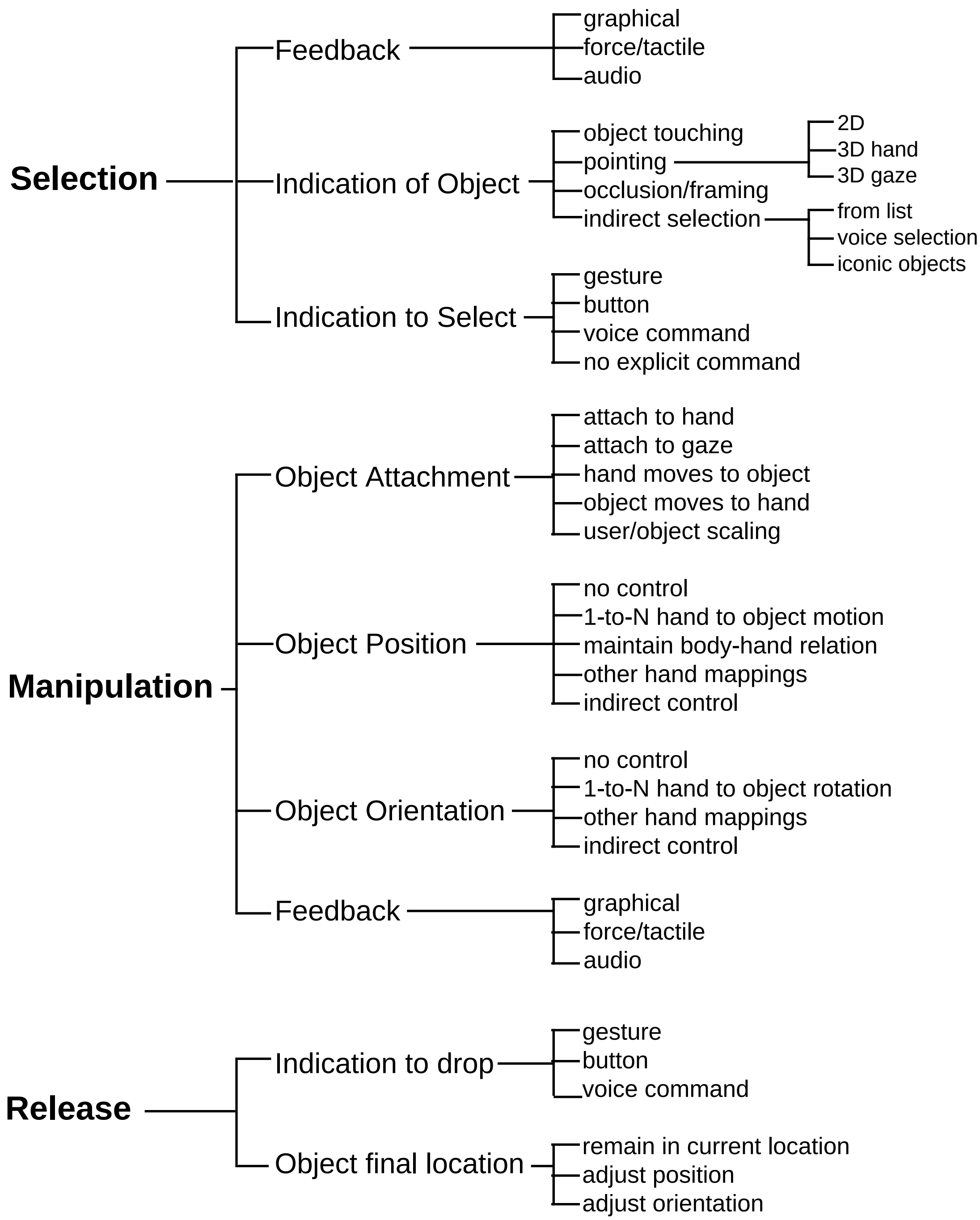

Figure 5.4 Taxonomy of Selection/Manipulation Techniques

The first thing that should be noted about the taxonomy are its three main branches, which break the task into its component parts: selection, manipulation, and release of the object. For selection-only tasks, the top branch of the taxonomy may be used alone. This division stems from the observation we made in our user study that selection and manipulation should be considered separately for optimal performance.

The main subtasks within the selection branch are the indication of the object and the indication to select the object. These subtasks are listed separately since the indication of the object does not necessarily imply that the object should be selected. For example, in a simple technique where the user touches objects, the user may touch many objects with his virtual hand, but only selects an object when a button is pressed while the object is being touched. Feedback is also given as a subcomponent of selection, but this is purely an interaction issue, and does not correspond to an actual user goal.

The second branch lists components and techniques for manipulation. Subtasks that are purely task-related are the indication to start manipulating the object (often the same as the indication to select, but not necessarily), indication of the center of rotation (not required), and the technique(s) to control object position and orientation. Object attachment is a technique consideration that may or may not apply – it refers to the way the object is attached to the manipulator (often the virtual hand). Feedback is also listed as an interaction component.

The final main branch concerns the release of a manipulated object. The only task-related component here is that the user must give some indication to drop the object (stop manipulation). From a technique point of view, however, the most important components of a release technique are what happens to the object and/or the virtual hand after release. For example, virtual gravity might be implemented which causes the object to fall naturally to a surface below. Also, in a technique where the virtual hand is displaced from the location of the physical hand (e.g. HOMER), the virtual hand position may need to be adjusted so that it once again coincides with the physical hand's position.

This taxonomy does not have the intuitive appeal of the broad technique categories mentioned above, but it is much more complete and general. It allows us to make interesting comparisons between various components of techniques, and general statements about performance. Perhaps even more important is the fact that this taxonomy encourages the guided design of new techniques because of its task-oriented structure.

5.4.2 Performance Measures

Like viewpoint motion control, selection and manipulation techniques can be evaluated for performance with a large number of possible metrics. Some techniques may trade off performance on one measure for better performance on another, and different applications may perform best with very different interaction techniques, due to different performance requirements. Again, we need to consider both quantitative and qualitative metrics, and those relating to the task as well as those relating to the user.

As in the case of ITs for travel, we have defined a list of metrics with which performance of techniques can be measured. Application designers can specify requirements for selection and manipulation in terms of those metrics, and choose ITs which meet those requirements.

Our list of performance metrics for immersive selection and manipulation techniques includes:

1. *Speed* (efficiency of task completion)
2. *Accuracy of Selection* (the ability to select the desired object)

3. *Accuracy of Placement* (the ability to achieve the desired position and orientation)
4. *Ease of Learning* (the ability of a novice user to use the technique)
5. *Ease of Use* (the complexity of cognitive load of the technique from the user's point of view)
6. *Presence* (the user's sense of immersion within the environment while using the technique)
7. *Expressiveness of Selection* (the number and distance of objects that can be selected)
8. *Expressiveness of Manipulation* (the ability to position and orient the object at any desired location in the environment)
9. *User Comfort* (lack of physical discomfort, including simulator sickness)

Speed and accuracy are important to many of the target applications, but more user-centric metrics such as user comfort can also play a major role. Many of the techniques which allow complete 6 DOF manipulation of virtual objects can force the user to assume awkward arm, wrist, or hand positions, for example. Also note that accuracy and expressiveness play a double role here, having different meanings for selection vs. manipulation.

5.4.3 Outside Factors

The final component of our formalized evaluation framework for selection and manipulation techniques is the consideration of other factors that could affect the performance of a technique. These factors were explicitly modeled in the evaluation testbed, so that performance differences could be attributed to the proper source. As before, we separate these outside factors into four categories: task, environment, user, and system characteristics.

5.4.3.1 Task Characteristics

A technique may perform very well for certain selection/manipulation tasks, but poorly on others. To determine these relationships, we can consider the following set of task characteristics:

- distance from the user to the object
- degrees of freedom required to be manipulated
- accuracy required
- task complexity (cognitive load induced)

5.4.3.2 Environment Characteristics

The environment (3D virtual world) surrounding the user can also have an effect on selection and manipulation. Interesting variables include:

- visibility
- number of objects
- size of objects
- shape of objects
- density of objects
- activity (motion)
- size of environment

- level of detail
- randomness/structure in the environment

5.4.3.3 User Characteristics

The individual user is also quite important for selection/manipulation techniques. For example, the Go-Go technique is less powerful for users with shorter arms. We have identified these user characteristics for consideration:

- age
- gender
- length of reach
- spatial ability
- height
- VE experience
- visual acuity
- manual dexterity
- ability to fuse stereo images
- technical/non-technical background

5.4.3.4 System Characteristics

Finally, the hardware and software comprising the VE system may themselves have effects on performance of selection/manipulation tasks. Such characteristics include:

- rendering technique
- use of shadows
- virtual body representation
- frame rate
- latency
- display type
- use of collision detection or constraints
- realism of physics model (e.g. gravity)

5.4.4 Guided Design

The selection and manipulation taxonomy has also proven useful as a framework for the design of new techniques. Because there are such a large number of techniques described in the literature, most of the techniques that arise from guided design are variants of techniques already available. However, small changes to certain subtasks can have a large effect on performance.

We have taken the guided design of selection and manipulation techniques to the next logical step by “implementing” the taxonomy in software. Five low-level subtasks (selection, attachment, positioning, orientation, and release), along with a large number of technique components for each of these subtasks, have been implemented in a modular fashion so that they can be arbitrarily combined automatically. In other words, a designer can create a new IT immediately simply by entering five codes into a program. Currently, there are 8 x 6 x 6 x 4 x 4 = 4608 possible combinations of technique components. However, because of dependencies and constraints in the design space, the number of possible techniques is reduced to 667.

Through experimentation with this system, a number of interesting possibilities have emerged. For example, a HOMER-like technique which uses gaze direction instead of pointing direction for selection frees the hands for other tasks until an object is selected. It also seems useful in some cases to separate positioning and orientation of objects by using two trackers instead of one that controls all six degrees of freedom. We can also combine techniques such as HOMER and Pierce's (1997) "sticky finger" technique, to use the best aspects of each. For example, occlusion selection might prove easier than 3D ray-casting, and so it could be used in a technique along with HOMER-style object manipulation.

5.5 Selection/Manipulation Testbed

The three components of the formal framework (taxonomy, performance measures, and outside factors) come together in the evaluation testbed for selection and manipulation. This testbed is a set of tasks and environments that measure the performance of various combinations of technique components for each of the performance metrics. Ideally, this testbed would vary all of the outside factors listed above, but such an experiment would not be completed for decades.

Therefore, we designed and implemented a simpler testbed system that can evaluate techniques in a number of what we consider to be the most important conditions. The analysis of importance is based on our experiences with real applications, our more informal study of selection and manipulation, and the requirements of our target application.

The testbed was designed to support the testing of any technique that can be created from the taxonomy. The tasks and environments are not biased towards any particular set of techniques. We have evaluated nine techniques, but others can be tested at any time with no loss of generality.

The tasks used are simple and general. In the selection phase, the user selects the correct object from a group of objects. In the manipulation phase, the user places the selected object within a target at a given position and orientation. Figure 5.5 shows an example trial. The user is to select the blue box in the center of the three by three array of cubes, and then place it within the two wooden targets in the manipulation phase. In certain trials, yellow spheres on both the selected object and the target determine the required orientation of the object.

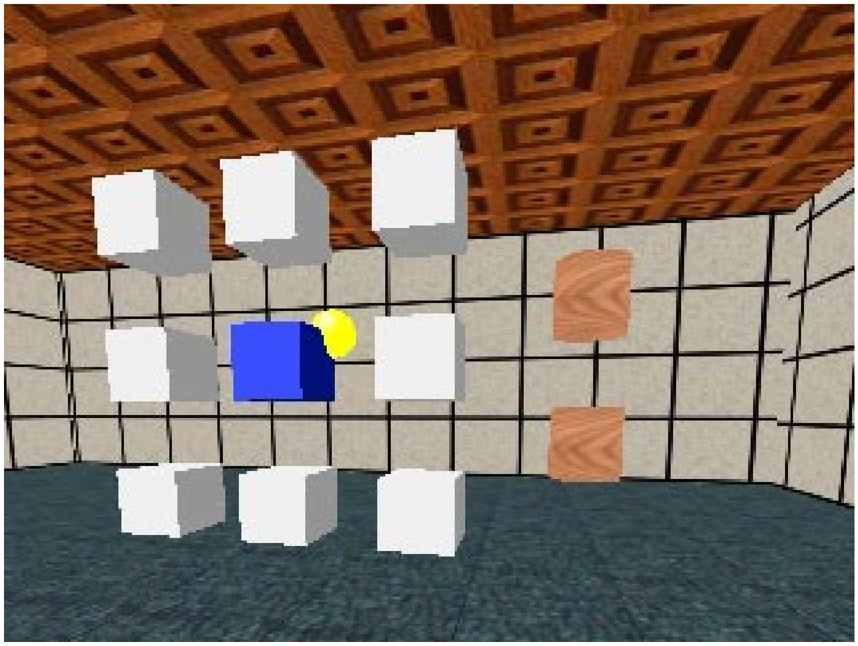

Figure 5.5 Example Trial Setup in the Selection/Manipulation Testbed

5.5.1 Method

Three within-subjects variables were used for the selection tasks. We varied the distance from the user to the object to be selected (three levels), the size of the object to be selected (two levels), and the density of objects surrounding the object to be selected (two levels).

The manipulation phase of the task also involved three within-subjects variables. First, we varied the ratio of the object size to the size of the target (two levels – this corresponds to the accuracy required for placement). Second, the number of required degrees of freedom varied (two levels), so that we could test the expressiveness of the techniques. The 2 DOF task only required users to position the objects in the horizontal plane (with constraints implemented that prevented the user from rotating the object or moving it vertically), while the 6 DOF task required complete object positioning and orientation. Finally, we changed the distance from the user at which the object must be placed (three levels), since this was a primary concern in our earlier user study.

Besides these explicit variables, we also included characteristics of the user in our analysis. We studied the effects of age, gender, spatial ability, VE experience, and technical background on the performance of techniques by having users fill out a pre-experiment questionnaire (Appendix A) and standardized spatial ability test (the ETS cube comparison test).

Response variables were the speed of selection, the number of errors made in selection, the speed of placement, and qualitative data related to user comfort (the same subjective reports as in the travel testbed – arm strain, hand strain, dizziness, and nausea on a ten-point scale; see Appendix B). We did not measure accuracy of placement; instead we required users to place the selected objects completely within the targets and within five degrees of the correct orientation on the six degree of freedom trials. Graphical feedback told the user when the object was in the correct location.

Forty-eight subjects (31 males, 17 females) participated in the study. Subjects were undergraduates from the Department of Psychology subject pool, and were given extra credit for their participation. Each subject completed 48 trials, except for three subjects who did not complete the experiment due to dizziness or sickness.

Nine different selection/manipulation techniques, taken from the taxonomy, were compared in a between-subjects fashion. Thus, there were five subjects per technique. First, we chose the Go-Go technique because of its importance and the fact that it was under consideration as the technique to be used in the Virtual Habitat application (chapter six). The other eight techniques were created by combining two selection techniques (ray-casting and occlusion), two attachment techniques (moving the hand to the object, scaling the user so the hand touches the object), and two positioning techniques (linear mapping of hand motion to object motion and the use of buttons to move the object closer or farther away).

Subjects wore a Virtual Research VR4 HMD, and were tracked using Polhemus Fastrak trackers. Input was given using a 3-button joystick. Subjects were allowed to practice the technique for up to five minutes in a room filled with furniture objects before the experimental trials began. Subjects completed four blocks of 12 trials each, alternating between trials testing selection and manipulation. After the practice session and after each block, subjective comfort information was taken.

5.5.2 Results

This complex experiment necessarily has a complex set of results. Here, we will present several major findings that emerge from the data. For complete results, see Appendix D. We performed a repeated measures analysis of variance (MANOVA) for both the selection and manipulation tasks.

First, results for selection of objects matched most of the experience that we had in our earlier informal study. Selection technique proved to be significant ($f(2,42)=13.6$, $p < 0.001$), with the Go-Go technique (mean 6.57 seconds per trial) proving to be significantly slower than either ray-casting (3.278 secs.) or occlusion selection (3.821 secs.) in post-hoc comparisons (LSD and Bonferroni). There was no significant difference between ray-casting and occlusion. This is because selection using ray-casting or occlusion is essentially a 2D operation, while the Go-Go technique requires users to place the virtual hand within the object in three-dimensional space.

We also found significant main effects for distance ($p < 0.001$) and size ($p < 0.001$), with nearer and larger objects taking less time to select. There were also several interesting significant interactions. As shown in figures 5.6 and 5.7, the effects of distance and size varied depending on the selection technique being used ($p < 0.001$ in both cases). Figure 5.6 shows that selection time for the Go-Go technique increases with distance, while the other two selection technique times remain approximately constant, regardless of object distance. Figure 5.7 indicates that the Go-Go technique benefits much more from larger object sizes as compared to ray-casting and occlusion selection.

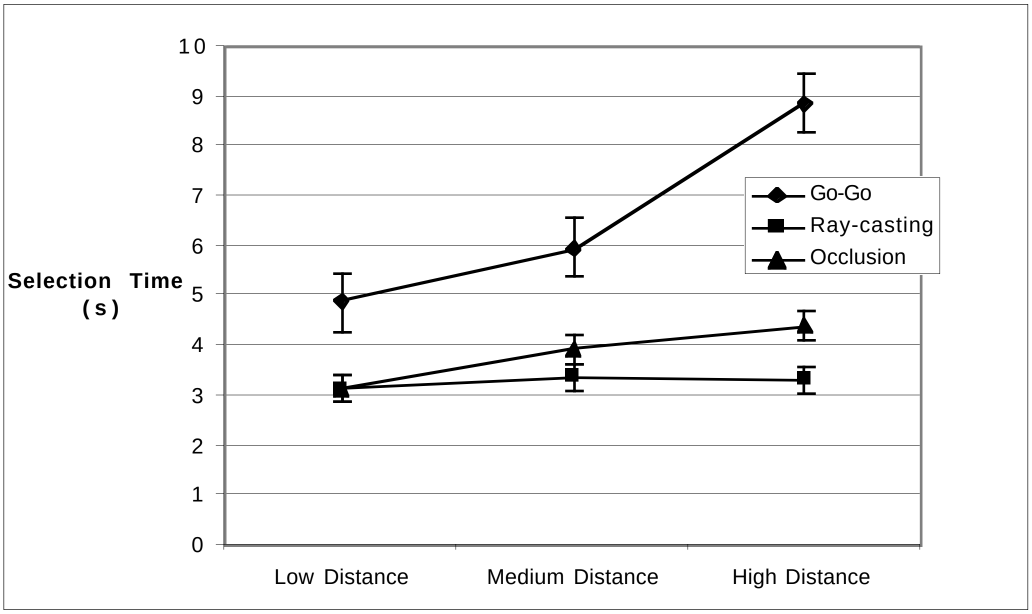

Figure 5.6 Interaction of Selection Technique with Object Distance for Selection Time Measure

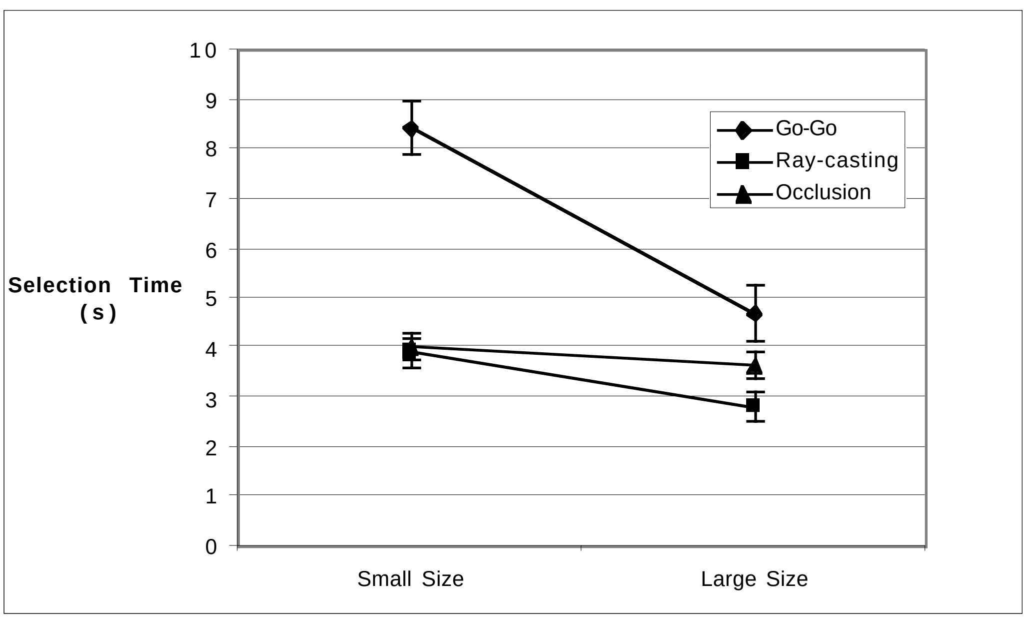

Figure 5.7 Interaction of Selection Technique with Object Size for Selection Time Measure

We found that the number of errors made during selection (errors included both selecting the wrong object and selecting no object) were significantly affected by both distance ($p < 0.001$) and size ($p < 0.001$). Interestingly, however, selection technique had no significant effect on errors.

It appears from this data that either ray-casting or occlusion is a good general-purpose choice for a selection technique. However, this is tempered by our findings with regard to user comfort. We found that selection technique had a high correlation to the reported final level of user arm strain (after all trials had been completed, approximately thirty minutes of use). Occlusion selection produced significantly higher levels of arm strain than ray-casting, because ray-casting allows the user to "shoot from the hip," while occlusion selection requires that the user's hand be held up in view. When selection takes a long time, as in the case of small or faraway objects, this can lead to arm strain of unacceptable levels.

The results for manipulation time were more difficult to interpret. Once the object had been selected, many of the techniques produced similar times for manipulation (table 5.1 shows the results for the nine techniques). We did find a significant main effect for technique ($f(8,36)=4.3$, $p < 0.001$) where technique is the combination of selection, attachment, and manipulation components. The only combinations that were significantly worse than others in the post-hoc tests were the two combinations that combined ray-casting with the attachment technique that scales the user, and this was likely due to poor implementation, from our observations of users. We found no significant effects of technique when attachment and manipulation techniques were considered separately.

Table 5.1 Mean Manipulation Time Results by Technique from Testbed Evaluation

(The linear mapping used in these cases was a one-to-one physical to virtual hand mapping)*

Tech	Selection	Attachment	Manipulation	Mean Time (s)
1	Go-Go	Go-Go	Go-Go	26.551
2	Ray-casting	Move hand	Linear mapping	32.047
3	Ray-casting	Move hand	Buttons	30.970
4	Ray-casting	Scale user	Linear mapping*	40.683
5	Ray-casting	Scale user	Buttons	39.851
6	Occlusion	Move hand	Linear mapping	31.800
7	Occlusion	Move hand	Buttons	22.537
8	Occlusion	Scale user	Linear mapping*	24.780
9	Occlusion	Scale user	Buttons	20.528

One interesting fact to note from table 5.1 is that for each pair of techniques using the same selection and attachment components, the technique using indirect depth control (button presses to reel the object in and out) had a faster mean time. Though this was not statistically significant, it indicates that an indirect, unnatural positioning technique can actually produce better performance. These techniques are not as elegant and seem to be less popular with users, but if speed of manipulation is important, they can be a good choice.

All three of our within-subjects variables proved significant. Distance ($f(2,72)=18.6$, $p < 0.001$), required accuracy ($f(1,36)=19.6$, $p < 0.001$), and degrees of freedom ($f(1,36)=286.3$, $p < 0.001$) all had significant main effects on manipulation time. As can be seen from the large f-value for degrees of freedom, this variable dominated the results, with the six degree of freedom task taking an average of 47.2 seconds to complete and the two degree of freedom task taking 12.7 seconds on average.

We also found a significant interaction between required accuracy and degrees of freedom, shown in table 5.2. The six degree of freedom tasks with a high accuracy requirement (small target size relative to the size of the object being manipulated) were nearly impossible to complete in some cases, indicating that we did indeed test the extremes of the capabilities of these interaction techniques. On the other hand, required accuracy made little difference in the 2 DOF task, indicating that the techniques we tested could produce quite precise behavior for this constrained task.

Table 5.2 Interaction Between Required Accuracy and Degrees of Freedom for Manipulation Time (seconds)

	2 DOFs	6 DOFs
Low Accuracy	11.463	40.441
High Accuracy	13.991	53.992

Unfortunately, these data cannot answer the question of whether there is a qualitative difference between the 2 DOF and 6 DOF tasks. Does the 2 DOF task have a constant slope regardless of the required accuracy or is its upward slope simply of lower magnitude than that of the 6 DOF task? In other words, does adding more degrees of freedom to a manipulation task create a different type of task, or does it simply add more of the same types of difficulty? The best way to answer these questions would be to include a middle condition with three degrees of freedom, and we propose this as future work. We can get some idea of the importance of this interaction by looking at these data on a log scale (figure 5.8). This graph does not appear to show an interaction, and thus we suggest that degrees of freedom may be additive, and not qualitatively different. This may be a fruitful topic for further research.

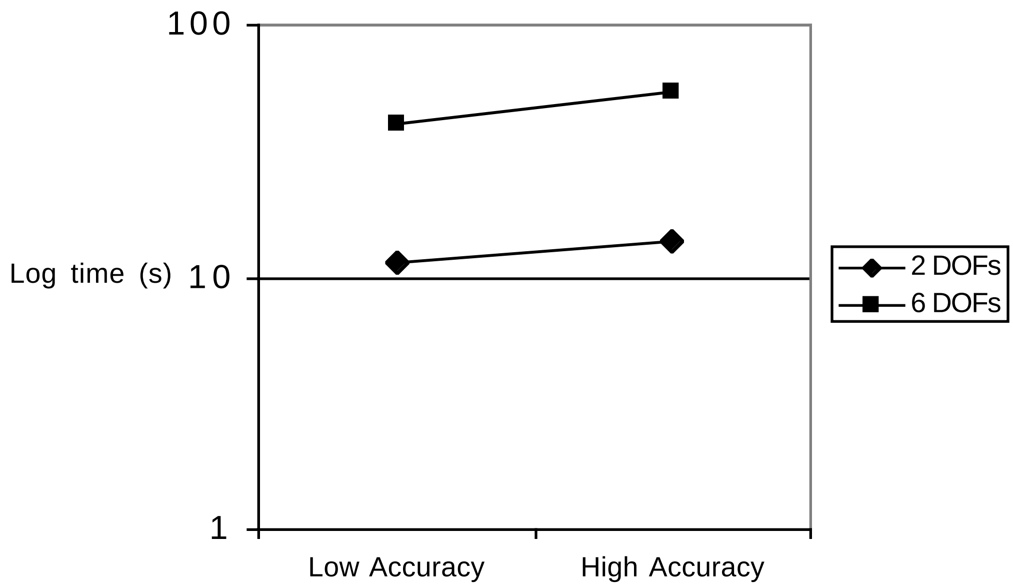

Figure 5.8 Logarithmic Scale Graph of Interaction Between DOFs and Accuracy

All of the significant results reported above have observed statistical power (computed using alpha = 0.05) of 0.92 or greater.

Finally, we found a demographic effect for performance. Males performed better on both the selection time ($p < 0.025$) and manipulation time ($p < 0.05$) response measures. Spatial ability and VE experience did not predict performance.

Again, looking at the results, we have any of a number of manipulation techniques to choose from which appear to have similar performance. The lowest mean times were achieved by techniques using occlusion selection and/or the scaling attachment technique (techniques 7, 8, and 9). The fact that the scaling technique produces better performance, especially on the six degree of freedom task, makes intuitive sense. If the user is scaled to several times normal size, then a small physical step can lead to a large virtual movement. That is, users can translate their viewpoint large distances while manipulating an object using this technique. Therefore, on the difficult manipulation tasks, users can move their viewpoint to a more advantageous position (closer to the target, with the target directly in front of them) to complete the task more quickly. We observed this in a significant number of users.

However, these techniques also have a price. We have already stated that occlusion selection increases arm strain. Similarly, scaled manipulation significantly increases the reported final level of dizziness relative to techniques where the user remains at the normal scale. Thus, an important guideline (hypothesis 2) is that such techniques should not be used when users will be immersed for extended periods of time.

5.6 Summary

In this chapter, we have used our design and evaluation methodology to study techniques for the selection and manipulation of objects in immersive VEs. These tasks will be found in most interactively complex VE applications, so it is crucial that we understand the performance characteristics of the various proposed ITs. Our initial user study of arm-extension and ray-casting techniques gave us useful information and understanding of these two metaphors, and allowed us to combine them for better performance in the HOMER techniques. We used this knowledge as a basis for our formal design and evaluation framework, including a taxonomy of selection and manipulation techniques, performance metrics, and outside factors that could influence performance. This framework was realized in our testbed evaluation, which produced complex but useful empirical results. In chapter six we apply these results to a complex VE application in order to increase its performance.

Several important principles come out of this research. Our user study showed that naturalism does not necessarily produce good performance on selection and manipulation tasks. Rather, magic techniques seemed to be easier, more efficient, and more acceptable to users. The testbed experiment showed that 2D selection metaphors based on ray-casting were more efficient, that the perceived size of virtual objects affects selection errors, and that scaled object manipulation can increase efficiency on difficult manipulation tasks. We also found user comfort to be a significant measure for selection and manipulation tasks. If speed were the only consideration, a technique such as Sticky Finger (occlusion selection combined with scaled object manipulation) would be an excellent choice. However, both of these components produced moderate to high levels of discomfort in users, which will not be acceptable in applications with longer exposure times.