

Moving Objects in Space: Exploiting Proprioception In Virtual-Environment Interaction

Mark R. Mine, Frederick P. Brooks Jr.¹
University of North Carolina at Chapel Hill

Carlo H. Sequin²
University of California at Berkeley



Abstract

Manipulation in immersive virtual environments is difficult partly because users must do without the haptic contact with real objects they rely on in the real world to orient themselves and their manipulanda. To compensate for this lack, we propose exploiting the one real object every user has in a virtual environment, his body. We present a unified framework for virtual-environment interaction based on *proprioception*, a person's sense of the position and orientation of his body and limbs. We describe three forms of body-relative interaction:

- Direct manipulation—ways to use body sense to help control manipulation
- Physical mnemonics—ways to store/recall information relative to the body
- Gestural actions—ways to use body-relative actions to issue commands

Automatic scaling is a way to bring objects instantly within reach so that users can manipulate them using proprioceptive cues. Several novel virtual interaction techniques based upon automatic scaling and our proposed framework of proprioception allow a user to interact with a virtual world intuitively, efficiently, precisely, and lazily. We report the results of both informal user trials and formal user studies of the usability of the body-relative interaction techniques presented.

CR Categories and Subject Descriptors: I.3.6 [Computer Graphics]: Methodology and Techniques - Interaction Techniques; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism - Virtual Reality.

Additional Keywords: virtual worlds, virtual environments, navigation, selection, manipulation.

¹CB #3175 Sitterson Hall, Chapel Hill, NC 27599-3175. Tel: +1-919-962-1700. Email: {mine, brooks}@cs.unc.edu. URL: <http://www.cs.unc.edu/~mine>

²Computer Science Division, Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720. Tel: +1-510-642-5103. Email: sequin@cs.berkeley.edu

1 MANIPULATION IN A VIRTUAL WORLD: WHY IS IT HARD?

1.1 The Problem

The promise of immersive virtual environments is one of a three-dimensional environment in which a user can directly perceive and interact with three-dimensional virtual objects. The underlying belief motivating most virtual reality (VR) research is that this will lead to more natural and effective human-computer interfaces. Promising results in several key application domains have been demonstrated:

Domain	Example Applications
"Being There", experience for the sake of experience	Phobia therapy: [Rothbaum 1995] Aesthetics: [Davies 1996] Entertainment: [Pausch 1996]
Training and practice of different skills	Surgery: [Hunter 1993] Military : [Macedonia 1994] Maintenance: [Wilson 1995] Wayfinding: [Witmer 1995]
Visualization of unrealized or unseeable objects	Architecture: [Brooks 1986] Fluid Flow: [Bryson 1992] Nano-surfaces: [Taylor 1993]
Design	3D models: [Butterworth 1992] Cityscapes: [Mapes 1995]

Table 1: Successful Virtual-World Application Domains

The number of successful virtual-environment applications, however, still remains small, with even fewer applications having gone beyond the research laboratory. Why?

Many of these successes fall within the realm of spatial visualization. The applications exploit the intuitive view specification (via head tracking) offered by VR systems but make little use of direct virtual-object manipulation. Why is it difficult to do much more than look around in a virtual world?

Besides the well known technological limitations such as system latency and display resolution, several less obvious factors have hampered development of "real-world" virtual-environment applications.

1) *The precise manipulation of virtual objects is hard.* Although immersion, head-tracked view specification, and six degree-of-freedom (DoF) hand tracking facilitate the coarse manipulation of virtual objects, the precise manipulation of virtual objects is complicated by:

- *Lack of haptic feedback:* Humans depend on haptic feedback and physical constraints for precise interaction in the real world; the lack of physical work-surfaces to align against and rest on limits precision and exacerbates fatigue. Though there is considerable ongoing research in the area of active haptic feedback [Durlach 1995], general-purpose haptic feedback devices that do not restrict the mobility of the user are not yet practical or available.

- *Limited input information:* Most virtual-environment systems accept position and orientation (pose) data on the user's head and (if lucky) two hands. One also typically has a button or glove to provide signal/event information. This suffices for specifying simple 6 DoF motion and placement. In the real world, we do this and much more:
 - a) Object modification, usually with tools.
 - b) Directing the cooperation of helping hands, by spoken commands ("Put that there").
 - c) Measuring.
 - d) Annotating objects with text.

Today in most VR systems:

- a) Tool selection is difficult.
 - b) Voice command technology is marginal.
 - c) Measuring tools are rarely available.
 - d) Alphanumeric input is difficult.
- *Limited precision:* The lack of haptic and acoustic feedback, inaccurate tracking systems, and whole-hand input typical of current VR systems restricts users to the coarse manipulation of virtual objects. Fine-grained manipulations are extremely difficult using this "boxing glove" style interface. Shumin Zhai of the University of Toronto, for example, has demonstrated that users' task completion times were slower in a 3D docking task when using a 3D input device which excluded the use of the fingers (vs. a similar device that utilized the fingers) [Zhai 1996].

2) *Virtual environments lack a unifying framework for interaction*, such as the desktop metaphor used in conventional through-the-window computer applications. Without haptics, neither real-world nor desktop computer interaction metaphors are adequate in a virtual environment. Knowledge on how to manipulate objects or controls can no longer be "stored in the world" [Norman 1988], with the physical constraints of the devices giving the user clues as to their use (e.g. a dial can only be rotated about its axis).

The desktop metaphor further breaks down when the user is inside the user interface. Interface controls and displays must move with the user as he moves through the environment and be made easy to locate. The differences between working in a conventional computer environment and working immersed are analogous to the differences between a craftsman at a workbench and one moving about a worksite wearing a toolbelt. His toolbelt had better be large and filled with powerful tools.

1.2 A Solution: Use What You Have

Without touch, a user can no longer feel his surroundings to tell where he is nor use the felt collision of a manipulandum with stationary objects to refine spatial perception. It is imperative, therefore, to take advantage of the one thing a user can still feel in the virtual world, his body.

A person's sense of the position and orientation of his body and its several parts is called *proprioception* [Boff 1986]. We propose that proprioception can be used to develop a unified set of interaction techniques that allow a user to interact with a virtual world intuitively, efficiently, precisely, and lazily.

We describe several novel body-relative interaction techniques based on the framework of proprioception. These techniques provide better control, precision, and dynamic range during manipulation and take advantage of additional input information such as data from the user's head or second hand.

We present automatic scaling as a means of instantly bringing objects in reach so that users can manipulate them using proprioceptive cues.

Finally, we present the results of formal users studies designed to evaluate several core aspects of body-relative interaction. In addition, throughout the paper, we provide anecdotal data from the tens of demonstrations that were given to students, faculty and visitors at the University of North Carolina during development of the interaction techniques presented in this paper.

2 PROPRICEPTION AND BODY-RELATIVE INTERACTION

In a series of user observations, we have found that body-relative interaction techniques (exploiting proprioceptive feedback) are more effective than techniques relying solely on visual information. Such body-relative interaction techniques provide:

- a physical real-world frame of reference in which to work
- a more direct and precise sense of control
- "eyes off" interaction (the user doesn't have to constantly watch what he's doing)

A user can take advantage of proprioception during body-relative interaction in at least three ways:

- *Direct manipulation:* If a virtual object is located directly at the user's hand position, the user has a good sense of the position of the object (even with eyes closed) due to proprioception, and thus a greater sense of control. It is easier to place an object precisely by hand than when it is attached to the end of a fishing rod.
- *Physical mnemonics:* Users can store virtual objects, in particular menus and widgets [Conner 1992], relative to his body. Since a user can no longer feel the world around him, it can be difficult to find, select, and use virtual controls in world space, especially if the user is free to walk about the environment. If, however, controls are fixed relative to the user's body, he can use proprioception to find the controls, as one finds his pen in his pocket, or his pliers in his tool belt. If controls are attached to the user's body, they move with him as he moves through the environment and are always within reach. Finally, controls can be stored out of view (behind the user's back for example), reducing visual clutter, yet remaining easily accessible (like an arrow from a quiver).
- *Gestural actions:* Just as a user's body sense can be used to facilitate the recall of objects, it can be used to facilitate the recall of actions, such as gestures used to invoke commands or to communicate information.

3 WORKING WITHIN ARM'S REACH

Interacting within a user's natural working volume (i.e. within arm's reach) has these advantages:

- takes advantage of proprioceptive information
- provides a more direct mapping between hand motion and object motion
- yields stronger stereopsis and head-motion parallax cues
- provides finer angular precision of motion

Often the target of manipulation lies outside of the user's reach. Though he can move to reach it, constantly switching between object interaction and movement control breaks the natural rhythm of the operation and adds significant cognitive over-

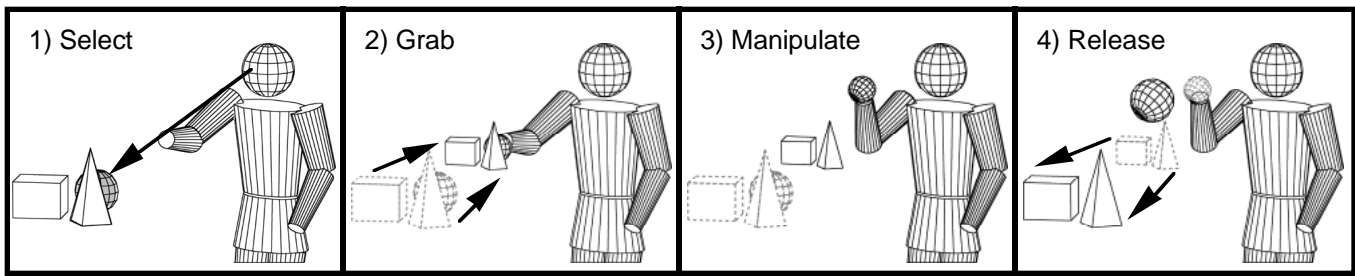


Figure 1: Automatic scaling of the world when the user grabs and releases an object.

head. We developed a convenient automatic scaling mechanism as a way to allow the user to interact instantly with objects lying at any distance as though they were within arm's reach.

Selected objects that fall outside of the reach of the user are brought instantly into reach by automatically scaling down the world about the user. For example, if the user's arm is extended 0.5 meters, the application brings a selected object that is 5 meters away to the user's hand by scaling down the world by a factor of 10 (see Figure 1)¹. Scaling takes place at the start of each manipulation and is reset when the user is done (when the user grabs and then releases an object, for example).

The scaling factor used to scale down the world (or conversely, scale up the user) is equal to the ratio of the distance of the object being manipulated to the distance of the user's hand:

$$\frac{\|\text{head_object}\|}{\|\text{projection of head_hand onto head_object}\|}$$

where *head_object* is the vector from the user's head (defined to be the midpoint between the user's eyes) to the object, and *head_hand* is the vector from the user's head to his hand.

If the center of the scaling operation is chosen to be the point midway between the user's eyes, he will be unaware, usually, that scaling has taken place, due to the ambiguity of perspective projections. This is particularly true if the inter-pupillary distance used to compute stereo images is also adjusted by the same scaling factor. This saves the user's having to reconverge the eyes. The most noticeable change is an apparent change in the size of the user's hand (which was not scaled). This can be offset by using non-realistic hand representations such as 3D crosshairs whose size is harder for the user to estimate visually.

A more implicit effect of the scaled-down world is the more dramatic effects of head movements; a small movement left-to-right may enable the user to see a big object such as a house from different sides, as though it were a dollhouse. While, in general, this is desirable, in some cases head motion can result in distracting movement of small, nearby objects.

¹ More precisely the object will move to a point on the surface of a sphere whose radius is equal to the user's current arm extension. The object will move to the user's hand only if his hand lies along the vector from the scaling center (the point midway between the user's eyes) to the object. This is the case when the user's hand visually occludes the object.

4 SAMPLE INTERACTION TECHNIQUES

4.1 Direct Manipulation

4.1.1 Scaled-world grab for manipulation

An example of the power of automatic scaling is *scaled-world grab*. In scaled-world grab, the world is automatically scaled down about the user's head every time he grabs an object and scaled back up when he releases it. With the object at the user's hand he can exploit proprioception, stereopsis, and head-motion parallax as he grabs an object and moves it.

Scaled-world grab is a powerful technique with an important property: it minimizes user work for a given result. With scaled-world grab the user can bring the most remote object in the scene to his side in a single operation; he doesn't have to fly (or worse, walk) to reach it or repeatedly grab, drop, and re-grab the object to reel it in. Furthermore, since the scale factor is automatically set, the user can manipulate near and far objects with equal ease. Scaled-world grab makes excellent use of a user's proprioceptive information for radial object movement, too: if the user halves his arm extension, the distance to the object will be halved. Movement of an object is easy for the user to control, predict and understand. Scaled-world grab is a surprising yet intuitive technique. In our informal user trials we have observed that users are often surprised to learn that scaling has taken place, and that they have no problem using the technique.

Related Work

The principles on which scaled-world grab is based have their foundations in the lessons we learned while exploring other forms of remote object manipulation.

Originally, for example, we tried the remote manipulation of objects via laser beams [Mine 1996, Mine 1997] (and later spotlights, following [Liang 1994]). We found, however, that even though these beams extend a user's reach, they are effective only for translations perpendicular to the beam direction and rotations about the beam axis. While it is easy to move an object about in an arc, translations in the beam direction and arbitrary rotations are much more difficult, requiring the user to repeatedly grab, move, drop, and re-grab the object.

A very effective way to specify arbitrary rotations is to use an *object centered interaction* paradigm [Wloka 1995, Mine 1997] in which changes in pose of the user's hand are mapped onto the center of a remote object. One technique we developed that grows out of this paradigm we call *extender grab*. Changes in orientation of the user's hand are applied 1:1 to the object's orientation. Translations are scaled by a factor which depends upon the distance of the object from the user at the start of the grab. The further away the object, the larger the scale factor.

By automatically setting the scale factor based on object distance, extender grab enables a large dynamic range of manipulation. No matter how far away an object lies, it can be brought to the user's side in a single operation. A key distinction between scaled-world grab and extender grab is that in the latter, manipulanda are not necessarily co-located with a user's hand as they are in scaled-world grab. This makes it harder for the user to exploit proprioception to determine object position and orientation. Similar techniques have been presented in [Bowman 1997] and [Pierce 1997]. In Pierce's techniques, the user interacts with the two-dimensional projections of 3D objects on the image plane.

A closely related technique for extending a user's reach called *go-go interaction* has been developed by Poupyrev et al. at the University of Washington [Poupyrev 1996]. In go-go interaction a user's virtual arm extension is a function of his physical arm extension, with a 1:1 mapping applied close to the user's body and a nonlinear function used further out. The maximum distance a user can reach depends upon the length of the user's arm and the scaling function used. Go-go interaction may require different scaling functions in scenes with different distributions of objects (i.e. mostly nearby or faraway).

Scaled-world grab has some common features with the Worlds-in-Miniature (WIM) paradigm (see [Pausch 1995, Stoakley 1995, Mine 1996, Mine 1997] and related earlier work in [Teller 1991]), in which objects are brought into reach in the form of a miniature copy of the environment floating in front of the user. WIMs have shown excellent promise in areas such as remote object manipulation and wayfinding. With a WIM, users can perform large scale manipulations of remote objects (moving a chair from one room in the house to another, for example) simply by manipulating the corresponding miniature copy in the WIM.

One drawback we have found with WIMs is that they force one to split limited display real estate between the miniature copy and the original environment. In addition, we have found that fine-grained manipulations can be difficult, particularly if the user is forced to hold a copy of the entire environment in his hand (as was the case in our system). If the entire environment has been scaled down to WIM size, individual scene elements may be quite small, and thus difficult to see, select, and manipulate. Note that manipulations at arbitrary resolutions would be easier if the user could interactively select a subset of the environment to view in the WIM (choosing to look at a single room instead of the whole house, for example). In that case the WIM could be thought of as a more general three-dimensional windowing system.

4.1.2 Scaled-world grab for locomotion

Automatic world-scaling also yields a useful locomotion mode, in which the user transports himself by grabbing an object in the desired travel direction and pulling himself towards it. With scaled-world grab the user can reach any visible destination in a single grab operation.

Since the point of interest is attached to the user's hand he can quickly view it from all sides by simply torquing his wrist. Alternately, if the virtual world stays oriented with the laboratory (which aids wayfinding), the user can swing himself about the point of interest, in a fashion similar to orbital mode (discussed later), by holding it in front of his face while he turns around (the world pivoting about his hand).

A similar movement metaphor called *virtual walking* is used in MultiGen's SmartSceneTM application [MultiGen 1997]. With

virtual walking users can pull themselves through the environment hand-over-hand, like climbing a rope. Virtual walking, however, is more suitable for the exploration of nearby objects since the extent of the pulling operation is limited to the reach of the user. To go much further the user must either invoke a separate scaling operation, to scale down the world until the desired destination is within reach, or switch to an alternate movement mode such as two-handed flying.

4.2 Physical Mnemonics

We call the storing of virtual objects and controls relative to the user's body *physical mnemonics*.

4.2.1 Pull-down menus

A thorny problem is the management and placement of virtual menus. If menus are left floating in space they are difficult to find. If they are locked in screen space they occlude parts of the scene. One solution is to keep the menu hidden and use a virtual button (like a menu bar) or a physical button to invoke the menu. However, small virtual buttons that minimally occlude are difficult to hit, and the cost of dedicating a physical button just for menu activation is high, since the number of buttons available on an input device is inherently limited.

As an alternative, we propose that one can hide virtual menus in locations fixed relative to the user's body, just above his current field of view for example. To access a menu the user simply reaches up, grabs it, and pulls it into view. The user can then interact with the menu using his other hand (if two hands are available) or through some form of gaze-directed interaction. Once the user is done with the menu he lets go, and it returns to its hiding place. This obviates a dedicated menu button, avoids occlusion by the menu, uses an existing operation for menu invocation, and keeps menus easy to find and access. In informal trials we have found that the user can easily select among three menus from above his field of view; one up and to the left, one just above, and one up and to the right.

The user's body can similarly be used to locate other tools or mode switches. Widgets for changing the viewing properties can be stored by the user's head; widgets for manipulating objects can be stored by the user's hands. The user's own body parts act as physical mnemonics which help in the recall and acquisition of frequently used controls. Furthermore, since the user is interacting relative to his own body, controls can remain invisible until acquired and can snap back to their hiding place when no longer needed. This minimizes occlusion of the scene by the virtual controls.

4.2.2 Hand-held widgets

In [Conner 1992], widgets (such as handles to stretch an object) were attached directly to the objects they control. To use such object-bound widgets in an immersive environment requires either the ability to reach the widget or some form of at-a-distance interaction. As an alternative we developed *hand-held widgets*: 3D objects with geometry and behavior that appear in the user's virtual hand(s). Hand-held widgets can be used to control objects from afar like using a TV remote control (see Figure 2).

We prefer hand-held widgets to object-bound ones for several reasons. First, we have observed in both formal user studies (Section 5) and informal user trials that users can select and work with hand-held widgets (assisted by proprioceptive in-

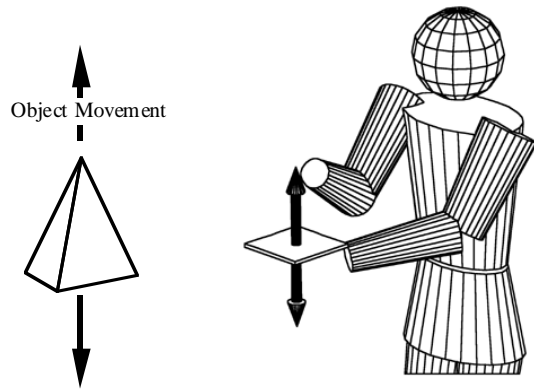


Figure 2: Using a hand-held widget.

formation) more easily than they can with object-bound widgets (whose position can be deduced only visually). Second, hand-held widgets enable a user to interact with selected objects from afar; he doesn't have to reach an object to use its widgets. Third, hand-held widgets reduce visual clutter, since each object doesn't have to have its own set of widgets, only a single copy of each kind of widget is needed. Finally, hand-held widgets eliminate obscuration of an object by its widgets. As a result of these factors, we find that it is preferable to work with widgets held in the hand, even though at-a-distance interaction with object-bound widgets could be accomplished using image-plane interaction techniques or automatic scaling.

4.2.3 FoV-Relative mode switching

One knows when one's hand or foot is in one's field of view (FoV), and one knows intuitively how to bring it into view. We have found that this can be effectively used for mode switching. Applications can, for example, switch between different forms of object selection. Occlusion selection, the selection of objects that visually lie behind a hand-attached cursor [Pierce 1997], requires the user's hand to be visible. The selection of objects pointed at by a laser beam or spotlight attached to the user's hand, does not. A logical and intuitive form of body-relative mode switching is to automatically change between occlusion selection to ray casting whenever the user's hand moves out of/into his current field of view.

4.3. Gestural Actions

We demonstrate how intuitive gestures can be augmented in powerful ways.

4.3.1 Head-butt zoom

Promising results have been reported on the potential of head pose as an auxiliary channel into the system. In orbital mode, for example, the user's head orientation is tracked and mapped so as to move the viewpoint of the user about the surface of a virtual sphere surrounding an object [Chung 1994]. Orbital mode is an excellent example of a technique not possible in the real world that gives the user power in the virtual environment. Chung found radiologists preferred it over six other methods of view-direction control such as mouse, joystick, and walk-around, and hypothesized that that was because it minimized work.

We developed head-butt zoom as another way for head motion to be used in controlling interaction. We have observed that users routinely and frequently switch between close-up local (and detailed) views and pulled-back global (and simplified) views when using interactive design systems, whether architectural CAD, molecular map tracing, or technical illustration preparation. Head-butt zoom enables a user to switch quickly between these two types of views as simply as leaning forward for a closer look.

Setting up head-butt zoom is similar to using a zoom tool in a conventional through-the-window application. The user frames the chosen detailed subset of his current view using a screen-aligned rectangle in front of his face. He sizes the rectangle like a movie director framing a shot; the position of his hands setting the corners of the rectangle (Figure 3). The size of this rectangle determines the zoom factor; its placement in world space determines the transition point. To remind the user that he is in head-butt zoom mode, a semi-transparent rectangle is left floating in space.

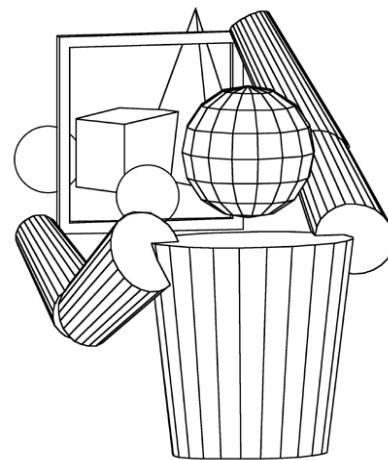


Figure 3: Setting up head-butt zoom

The user switches between the two views simply by leaning forward and backward. Lean forward (across the plane of the rectangle) to get a close-up and detailed view; lean back to return to the normal view.² If the user wishes to remain in the close-up view, he simply steps forward, at which point he will translate to the point of view of the zoomed-in view. Stepping back, he will return to the original view.

Instead of having the user explicitly frame a region of interest, the current zoom factor can be based upon the currently selected object (chosen so that when he leans forward the object fills his field of view). This mode makes it easier to integrate head-butt zoom with other forms of interaction, since the user does not have to interrupt his current operation in order to switch modes and specify a zoom rectangle.

Head-butt zoom makes good use of an additional input channel, i.e., head position. Users can change zoom levels without having to interrupt the current operation they are performing with their hands. Head-butt zoom also makes good use of limited display space, since one no longer needs to share screen space between versions of the same scene at different scales, as in World-In-Miniature.

² Note that head-butt zoom can also be used to switch between other kinds of viewing modes. E.g., the user could lean forward to get a wireframe view, lean back to get a full shaded representation.

4.3.2 Look-at Menus

Head orientation can be used instead of the traditional hand position to control the cursor used to select an item from a menu. To select one turns the head instead of moving the hand. The pick ray is fixed relative to the head (thus tracking head motion, see Figure 4). This gives an intuitive way to select an item simply by looking at it. To confirm selection, the user presses a physical button or, with pull-down menus, releases the menu.

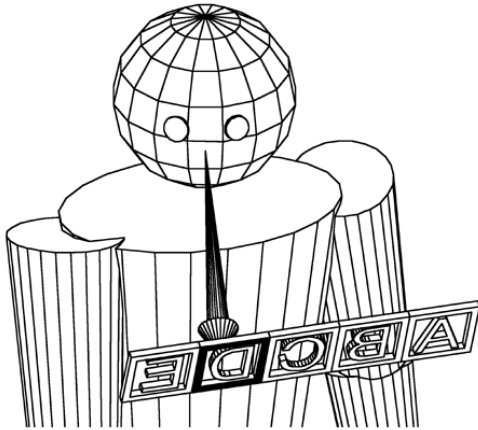


Figure 4: Look-at menu.

4.3.3 Two-handed flying

Numerous results describe the benefits of two-handed input in interactive applications ([Buxton 1986, Bier 1993, Shaw 1994, Goble 1995, Mapes 1995, Cutler 1997, Zeleznik 1997] and [Guiard 1987] for more theoretical foundations). We have found two-handed flying an effective technique for controlled locomotion. The direction of flight is specified by the vector between the user's two hands, and the speed is proportional to the user's hand separation (see Figure 5)³. A dead zone (some minimum hand separation, e.g. 0.1 meters) enables users to stop their current motion quickly by bringing their hands together (a quick and easy gesture). Two-handed flying exploits proprioception for judging flying direction and speed.

Two-handed flying is easier ergonomically than conventional one-handed flying in which the user's hand orientation specifies direction and arm extension specifies speed. Flying backwards using one-handed flying, for example, requires the user to regrab the input device or to turn his hand around awkwardly. With two-handed flying the user simply swaps his hands. Moreover, speed control based on hand separation is less tiring than speed control based on arm extension, the user doesn't have to hold his hands out in front of his body.

4.3.4 Over-the-shoulder deletion

A common operation is deletion; users need an easy way to get rid of virtual objects. Over-the-shoulder deletion is an intuitive gesture that exploits body sense. To delete an object the user simply throws it over his shoulder. It is easy to do, easy to remember, and it does not use up any buttons or menu space. It is

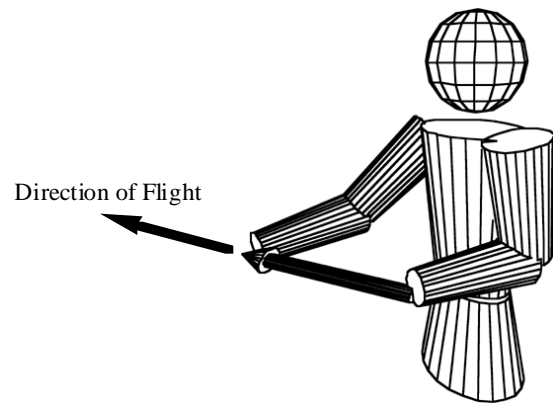


Figure 5: Two-handed flying.

unlikely to be accidentally invoked, since users do not typically manipulate objects in that region.

The space behind a user's head can be treated as a virtual clipboard. A user can later retrieve the last object deleted by simply reaching over his shoulder and grabbing it.

5 USER STUDIES

To evaluate some of the core aspects of body-relative interaction we performed two formal user studies, a virtual-object docking task, and a widget interaction task.

5.1 Virtual-Object Docking

The goal of this experiment was to explore the differences between manipulating virtual objects that are co-located with one's hand and manipulating objects at a distance. The experiment was a repeated measures design with three primary experimental conditions: the manipulation of objects held in one's hand, objects held at a fixed offset, and objects held at an offset which varied with the subject's arm extension. These three conditions were abstractions of three forms of remote object manipulation: scaled-world grab, laser beam interaction, and extender grab. Our hypothesis was that the manipulation of objects co-located with one's hand is more efficient than the manipulation of objects held at an offset.

5.1.1 The Experiment

Subjects

Eighteen unpaid subjects (7 female, 11 male) were recruited from staff and students at the University of North Carolina-CH.

The Task

The virtual environment consisted of target shapes and hand-held docking shapes. Target shapes were semi-transparent red cubes floating in space in front of the subject. Hand-held docking shapes were fully opaque blue cubes attached to the subject's hand. Docking shapes were either co-located with the subject's dominant hand or at some random initial offset ranging from 0.1 - 0.6 meters.

The task given to the subjects was to align the hand-held docking cube with the target cube floating in space as quickly as possible. Target cubes were presented to the subject one at a time. Each time the subject successfully aligned the docking cube with the target cube, both the target cube and docking cube

³ A similar flying technique has been implemented by Mapes and colleagues in MultiGen's SmartScene™ [MultiGen 1997].

would disappear from their current locations and then reappear in their new locations.

The subject controlled the virtual docking cube using a small styrofoam cube held in his dominant hand.

5.1.2 The Results

Table 2 presents the overall means obtained in each experimental condition.

Condition	Mean (sec)	Std. Dev. (sec)
In hand	3.87	2.05
Fixed offset	5.10	3.83
Variable offset	4.96	4.48

Table 2: Mean trial completion time.

Results were analyzed using a one-way multivariate analysis of variances (MANOVA), repeated measures design. This analysis revealed significant differences among the mean trial completion times for the three manipulation techniques $F(2,16) = 7.86$; $p < 0.005$. Contrasts of trial completion times showed that the manipulation of objects co-located with one's hand was significantly faster than the manipulation of objects at a fixed offset ($F(1,17) = 16.70$; $p < 0.001$), and the manipulation of objects at a variable offset ($F(1,17) = 8.37$; $p = 0.01$). No significant difference was found between the manipulation of an object at a fixed offset and one at a variable offset ($F(1,17) = 0.25$; $p = 0.62$).

Analysis of a post-test questionnaire also revealed a preference for the manipulation of an object co-located with one's hand, subjects rated it significantly higher than the manipulation of an object with an offset.

5.2 Virtual-Widget Interaction

This experiment explored the differences between interacting with a widget held in one's hand and interacting with a widget floating in space. Our goal was to see if subjects could take advantage of proprioception when interacting with widgets held in their hand. Our hypothesis was that it would be easier for subjects to interact with hand-held widgets than it would be for them to interact with widgets floating in space.

5.2.1 The Experiment

Subjects

The same eighteen subjects used in the object-docking experiment were used in this experiment.

Task

The virtual environment consisted of a three-dimensional cursor, a virtual widget, a current-color indicator, and target objects. The three-dimensional cursor was a small sphere attached to the subject's dominant hand. The virtual widget consisted of three orthogonal rods with colored spheres at each end (for a total of six spheres). The widget was either attached to the subject's non-dominant hand or was fixed floating in space. The current-color indicator was an additional colored sphere that was fixed in the upper right hand corner of the subject's display. Finally, the target objects were red semi-transparent cubes which appeared at random positions and orientations about the subject.

Each trial was broken down into three phases. First, the subject moved the 3D cursor to the indicated point on a virtual widget (one of the six colored spheres). Next, he performed an

unrelated abstract task (moving his hand from the widget to a target cube and clicking on the input button). This made the target cube and widget disappear. Finally, he returned his hand as closely as possible to the original point of interaction on the virtual widget, without visual feedback. The dependent variable measured was the positional accuracy with which the user could return his hand to a point in space.

5.2.2 The Results

Table 3 presents the overall means obtained in each experimental condition.

Condition	Mean (cm)	Std. Dev. (cm)
Hand held	5.1	4.0
In space	10.4	6.1

Table 3: Mean positional offsets

Results were analyzed using a one-way analysis of variances (ANOVA), repeated measures design. This analysis revealed a significant difference in positional accuracy between widgets held in one's hand and widgets fixed in space $F(1,17) = 115.52$; $p < 0.001$. Contrasts of positional accuracy showed that subjects were able to return to a position relative to their own hand more accurately than a position fixed in virtual space.

Analysis of a post-test questionnaire also revealed a preference for interaction with widgets held in one's hand, subjects rated it as being easier, more precise and better overall.

6 FUTURE WORK

We continue to explore additional means for compensating for the lack of haptic feedback in immersive virtual worlds. Though proprioception greatly enhances virtual-environment interaction, precise manipulation is still harder in virtual spaces than in real space. Several factors complicate fine-grained manipulation.

First, the lack of physical work surfaces and haptic feedback makes the controlled manipulation of virtual objects much more difficult. Users typically manipulate virtual objects by holding their arms out without support. In the real world, a person generally grounds the arm at the forearm, or elbow, or heel of hand to steady hand motions and to reduce fatigue when performing precise manipulation.

Second, humans depend upon naturally occurring physical constraints to help determine the motion of objects they are manipulating (sliding a chair along a floor, for example). Whereas it is possible to implement virtual equivalents of physical constraints [Bukowski 1995], it is more difficult for the user to take advantage of these constraints without haptic feedback. He can only see that the chair is on the floor, he can't feel the contact, hear it, or sense the vibration as the chair slides.

Third, users in a virtual world must typically do without the fingertip control they rely on for the fine-grained manipulation of objects in the real world. Instrumented gloves have shown some promise for the fine-grained manipulation of objects, but they have proven difficult to use in practice [Kijima 1996].

One approach we are exploring to give users a real surface on which they can work using haptic constraints is to provide a hand-held tablet (following the lead of [Sachs 1991] and [Stoakley 1995]). The tablet can be used as a two-dimensional drawing surface (to define detailed two-dimensional shapes) or it can be used as the input space for a two-dimensional menu

(allowing users to interact precisely with widgets and controls).

If the user interacts with the tablet using a hand-held stylus, he can take advantage of the user's fingertip control precision. In addition the friction between tablet and stylus and the grounding of the stylus against the tablet give the user better control.

To provide a larger work surface the tablet can be docked in a larger fixed physical surface such as a lectern or a drafting table which can also provide grounding and support during object manipulations (see related work in [Mapes 1995]).

7 ACKNOWLEDGMENTS

Special thanks to Robert Zeleznik of Brown University for his invaluable contributions to this paper. Thanks to Hans Weber of UNC; Andrew Forsberg, Seung Hong, Loring Holden of Brown University; and Jeff Pierce, and Dennis Cosgrove of the University of Virginia for numerous discussions about virtual-environment interaction and their many helpful comments and suggestions. Thanks to Todd Gaul of UNC for his assistance in the production of the associated video tape. As always, thanks to Sandra, Dylan, and Baby X.

This funding for this work was provided by DARPA Contract DABT63-93-C-0048, and Lockheed Missile and Space Co., Inc. (indirect DARPA).

REFERENCES

- [Bier 1993] Bier, E.A., M.C. Stone, K. Pier, W. Buxton and T.D. DeRose. "Toolglass and magic lenses: the see-through interface." *Proceedings of SIGGRAPH 93*, Anaheim, CA, ACM: 73-80.
- [Boff 1986] Boff, K.R., L. Kaufman and J.P. Thomas, Eds.. *Handbook of Perception and Human Performance*. New York, John Wiley and Sons.
- [Bowman 1997] Bowman, D. and L.F. Hodges. "An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments." *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, Providence, RI, ACM: 35-38.
- [Brooks 1986] Brooks, F.P., Jr. "Walkthrough-a dynamic graphics system for simulating virtual buildings." *Computer Graphics. 1986 Workshop on Interactive 3D Graphics* 21(1): 3.
- [Bryson 1992] Bryson, S. and C. Levit. "The Virtual Wind Tunnel." *IEEE Computer Graphics & Applications* 12: 25-34.
- [Bukowski 1995] Bukowski, R.W. and C.H. Sequin. "Object Associations: A Simple and Practical Approach to Virtual 3D Manipulation." *Proceedings of the 1995 Symposium on Interactive 3D Graphics*, Monterey, CA, ACM: 131-138.
- [Butterworth 1992] Butterworth, J., A. Davidson, S. Hench and T.M. Olano. "3DM: A Three Dimensional Modeler Using a Head-Mounted Display." *Computer Graphics. Proceedings 1992 Symposium on Interactive 3D Graphics* 25(2): 135-138.
- [Buxton 1986] Buxton, W. and B.A. Myers. "A Study in Two-Handed Input." *Human Factors in Computing Systems (ACM SIGCHI 86 Conference Proceedings)*, Boston, MA, ACM, New York: 321-326.
- [Chung 1994] Chung, J. *Intuitive Navigation in the Targeting of Radiation Therapy Treatment Beams*. University of North Carolina, Ph.D. Thesis
- [Conner 1992] Conner, D.B., S.S. Snibbe, K.P. Herndon, D.C. Robbins, R.C. Zeleznik and A. vanDam. "Three-dimensional widgets." *Computer Graphics (1992 Symposium on Interactive 3D Graphics)* 25(2): 183-188.
- [Cutler 1997] Cutler, L.D., B. Fröhlich and P. Hanrahan. "Two-Handed Direct Manipulation on the Responsive Workbench." *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, Providence, RI, ACM: 107-114.
- [Davies 1996] Davies, C. and J. Harrison. "Osmose: Towards Broadening the Aesthetics of Virtual Reality." *Computer Graphics* 30(4):
- [Durlach 1995] Durlach, N.I. and A.S. Mavor, Eds.. *Virtual Reality: Scientific and Technological Challenges*. Washington, D.C., National Academy Press.
- [Goble 1995] Goble, J.C., K. Hincley, R. Pausch, J. W. Snell and N.F. Kassell. "Two-handed spatial interface tools for neurosurgical planning." *Computer* 28,(7): p. 20-6.
- [Guiard 1987] Guiard, Y. "Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model." *The Journal of Motor Behavior* 19(4): 486-517.
- [Hunter 1993] Hunter, I.W., D.D. Tilemachos, S.R. Lafontaine, P.G. Charette, L.A. Jones, M.A. Sagar, G.D. Mallinson and P.J. Hunter. "A Teleoperated Microsurgical Robot and Associated Virtual Environment for Eye Surgery." *Presence* 2(4): 265-280.
- [Kijima 1996] Kijima, R. and M. Hirose. "Representative Spherical Plane Method and Composition of Object Manipulation Methods." *1996 Virtual Reality Annual International Symposium*, Santa Clara, CA, IEEE: 196-202.
- [Liang 1994] Liang, J. and M. Green. "JDCAD: a highly interactive 3D modeling system." *Computers & Graphics* 18(4): 499-506.
- [Macedonia 1994] Macedonia, M.R., M.J. Zyda, D.R. Pratt, P.T. Barham and S. Zeswitz. "NPSNET: A Network Software Architecture for Large Scale Virtual Environments." *Presence* 3(4): 265-287.
- [Mapes 1995] Mapes, D.P. and J.M. Moshell. "A Two-Handed Interface for Object Manipulation in Virtual Environments." *Presence* 4(4): 403-416.
- [Mine 1997] Mine, M.R. "ISAAC: A Meta-CAD System for Virtual Environments." *Computer-Aided Design* : to appear.
- [Mine 1996] Mine, M.R. "Working in a Virtual World: Interaction Techniques Used in the Chapel Hill Immersive Modeling Program." *University of North Carolina*, Technical Report TR96-029
- [MultiGen 1997] MultiGen Inc. "SmartScene™." for more information see: <http://www.multigen.com/smart.htm>
- [Norman 1988] Norman, D.A. *The psychology of everyday things*. New York, Basic Books.
- [Pausch 1995] Pausch, R., T. Burnette, D. Brockway and M.E. Weiblen. "Navigation and locomotion in virtual worlds via flight into hand-held miniatures." *Proceedings of SIGGRAPH 95*, Los Angeles, CA, ACM: 399-400.
- [Pausch 1996] Pausch, R., J. Snoddy, R. Taylor, S. Watson and E. Haseltine. "Disney's Aladdin: First Steps Toward Storytelling in Virtual Reality." *Proceedings of SIGGRAPH 96*, New Orleans, LA, ACM: 193-202.
- [Pierce 1997] Pierce, J.S., A. Forsberg, M.J. Conway, S. Hong, R. Zeleznik and M.R. Mine. "Image Plane Interaction Techniques in 3D Immersive Environments." *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, Providence, RI, ACM: 39-44.
- [Poupyrev 1996] Poupyrev, I., M. Billinghurst, S. Weghorst and T. Ichikawa. "The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR." *Proceedings of UIST 96*, Seattle, WA, ACM:
- [Rothbaum 1995] Rothbaum, B., L. Hodges, R. Kooper, D. Opdyke, J. Williford and M. North. "Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia." *American Journal of Psychiatry* 152(4): 626-628.
- [Sachs 1991] Sachs, E., A. Roberts and D. Stoops. "3-Draw: a tool for designing 3D shapes." *IEEE Computer Graphics and Applications* 11(6): 18-26.
- [Shaw 1994] Shaw, C. and M. Green. "Two-handed polygonal surface design." *Proceedings of UIST 94*, Marina del Rey, CA, p. 226, 205-12.
- [Stoakley 1995] Stoakley, R., M.J. Conway and R. Pausch. "Virtual Reality on a WIM: Interactive Worlds in Miniature." *Proceedings of CHI 95*, Denver, CO, ACM: 265-272.
- [Taylor 1993] Taylor, R.M., W. Robinett, V.L. Chi, F.P. Brooks Jr., W.V. Wright, S. Williams and E.J. Snyder. "The Nanomanipulator: A Virtual-Reality Interface for a Scanning Tunnel Microscope." *Proceedings of SIGGRAPH 93*, Anaheim, CA, ACM:
- [Teller 1991] Teller, S.J. and C.H. Sequin. "Visibility Preprocessing for Interactive Walkthroughs." *Computer Graphics (Proceedings of SIGGRAPH 91)* 25(4): 61-69.
- [Wilson 1995] Wilson, J.R., D.J. Brown, S.V. Cobb, M.M. D'Cruz and R.M. Eastgate. "Manufacturing Operations in Virtual Environments (MOVE)." *Presence* 4(3): 306-317.
- [Witmer 1995] Witmer, B.G., J.H. Bailey and B. Knerr, W. "Training Dismounted Soldiers in Virtual Environments: Route Learning and Transfer." *United States Army Research Institute for the Behavioral and Social Sciences*, Technical Report 1022
- [Wloka 1995] Wloka, M.M. and E. Greenfield. "The Virtual Tricorder: A Uniform Interface for Virtual Reality." *Proceedings UIST 95*, Pittsburgh, PA, ACM: 39-40.
- [Zeleznik 1997] Zeleznik, R.C., A.S. Forsberg and P.S. Strauss. "Two pointer input for 3D interaction." *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, Providence, RI, ACM: 115-120.
- [Zhai 1996] Zhai, S., P. Milgram and W. Buxton. "The Influence of Muscle Groups on Performance of Multiple Degree-of-Freedom Input." *Proceedings of CHI 96*, Vancouver, BC, Canada, ACM: 308-315.