Design and Evaluation of 3D Multiple Object Selection Techniques

John F. Lucas
Center for Human-Computer Interaction
Virginia Tech
jolucas@vt.edu

Doug A. Bowman
Center for Human-Computer Interaction
Virginia Tech
bowman@vt.edu

Jian Chen
Center for Human-Computer Interaction
Virginia Tech
jichen8@vt.edu

Chadwick A. Wingrave
Center for Human-Computer Interaction
Virginia Tech
cwingrav@vt.edu

ABSTRACT

Few researchers have addressed the important issue of three-dimensional multiple object selection (MOS) in immersive Virtual Environments (VEs). We have developed a taxonomy of the MOS task as a framework for exploring the design space of these techniques. In this paper, we describe four techniques for selecting multiple objects in immersive VEs. Of the four techniques, two are serial (where only one object can be indicated per operation), and two are parallel (where one or more objects may be indicated per operation). Within each of the two categories we also investigated two metaphors of interaction: a 3D spatial metaphor and the pen and tablet metaphor. Two usability studies were used to evaluate the four techniques, iterate their designs, and gain a deeper understanding of the design space of MOS techniques. The results from our studies show that parallel MOS techniques can be far more effective than serial techniques as the number of target objects increase. We also show that effective techniques for MOS in immersive VEs can be created using both pen and tablet and 3D metaphors.

ACM Classification Keywords: H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Keywords: Virtual Reality; 3D Interaction; User Testing and Evaluation

1 INTRODUCTION

Much research on 3D user interfaces has dealt with designing techniques for interacting with one object at a time [4]. Immersive design [3] is the process of creating, reviewing and modifying objects in an immersive Virtual Environment (VE). In this application domain environments typically contain more manipulable objects than in other common VE applications. Desktop 3D modeling and CAD applications allow for interaction with many objects in one operation. In order for immersive design applications to be successful, new techniques need to be developed to interact with many objects at one time.
Consider Sally, an interior designer, who is designing a café for a client. Sally's employer has requested that the café feel small but still have adequate space for 15 tables. Sally designs her environment on her desktop computer and imports her design into an immersive design application. Using the immersive application Sally gains a better understanding of the space. She quickly realizes that the tables in her café are much too large and need to be resized. Unfortunately, her immersive design application does not allow her to modify all of the tables at once. Because resizing each table individually would be very time consuming, Sally must return to her desktop to make this simple modification. Sally's immersive design application did not provide a tool for resizing multiple tables at once because few such techniques for Multiple Object Interaction (MOI) in VEs exist, of which none have been thoroughly examined. Without this crucial class of interaction techniques, designers like Sally are forced to switch back and forth between their desktop and VE applications.

Cloning is the process of creating multiple spatially distributed copies of an object [5]. The first sub-task of the cloning operation is selecting the objects to be cloned. This sub-task, called Multiple Object Selection (MOS) is common to all MOI tasks and involves distinguishing a set of objects to the system for future interaction. When cloning large structures the MOS task may take a significant amount of time using one-at-a-time selection techniques. With these *serial* selection techniques, selection of large groups of objects can be very time consuming. Here the lack of powerful MOS techniques can cause the user to become fatigued while attempting to perform a simple task. *Parallel* MOS techniques, such as 2D brushes and lassos, are very effective at cutting down time to select groups of objects in desktop environments. Extending these techniques to immersive 3D applications, however, is a challenging task as issues such as occlusion, fatigue, and environment clutter become more prominent in immersive systems. Overcoming these issues and creating effective 3D parallel MOS techniques is an important hurdle that must be passed before powerful interfaces can be created for interacting with the ever increasing number of objects in today's VE applications.

In this paper, we examine the design space of 3D MOS techniques, focusing on the serial/parallel distinction. To aid in this exploration we present a taxonomy of 3D MOS techniques. We implemented four techniques representing

UIST'05, October 23–27, 2005, Seattle, Washington, USA.

different branches of our taxonomy. Two separate studies evaluate these techniques and compare their performance. We then discuss the strengths and weaknesses of the techniques and present some guidelines for the design of future 3D MOS techniques.

2 RELATED WORK

There are many existing MOS techniques for 2D graphical user interfaces (GUIs). In Willis' [21] framework, the task of area differentiation (the process of defining the area within which objects are indicated), is divided into two branches. *Brushes* are persistent areas where objects are continuously checked against the selection criteria. *Lassos* are single-use areas where the shape is cleared after the selection operation is performed. This simplification of the area differentiation task is appropriate for 2D WIMP interfaces but is not descriptive enough to adopt for 3D object selection, since there are many other factors in 3D volume differentiation that need to be examined.

Many researchers have investigated 3D single-object selection techniques, such as ray-casting [12] and image-plane selection [13] for immersive VEs. These techniques can, of course, be used in a serial fashion for MOS as well.

Previous work on differentiation of 3D volumes for object selection has been rather limited in VEs. 2D techniques for desktop VEs have been successful in 3D modeling and design applications due to their ability to work with multiple views of a scene at one time and user familiarity with WIMP selection techniques. Applications that allow for volume definition with 3D input devices are rare. Liang and Green [10] implemented cone-casting where a cone is projected from a 3D input device. The object closest to the center of the cone is selected. Steed and Parker [17] extend this idea with shadow cone selection. Here, the user moves a cone through the environment and only objects that were surrounded by the cone throughout the manipulation are selected. Darrah et al. [6] designed a MOS technique for the PHANTOM haptic pen that creates the selection volume using a convex hull algorithm. All these techniques have been effective in their given applications but none have been used for 3D multiple object selection in immersive VEs.

Szalavári and Gervaultz [18] presented an interface using a tracked tablet for 2D interaction with a 3D VE. Other researchers have used this metaphor to assist in defining 3D volumes for selection on workbenches [16; 8]. This approach cannot be easily extended to more immersive setups for two reasons: the user can only effectively define volumes within arm's reach and the techniques do not scale well for defining very large or very small volumes.

In our work, we take a systematic approach to the design and evaluation of 3D MOS techniques for immersive VEs. We consider existing single-object techniques as well as novel volume differentiation techniques.

3 DESIGN SPACE OF MOS TECHNIQUES

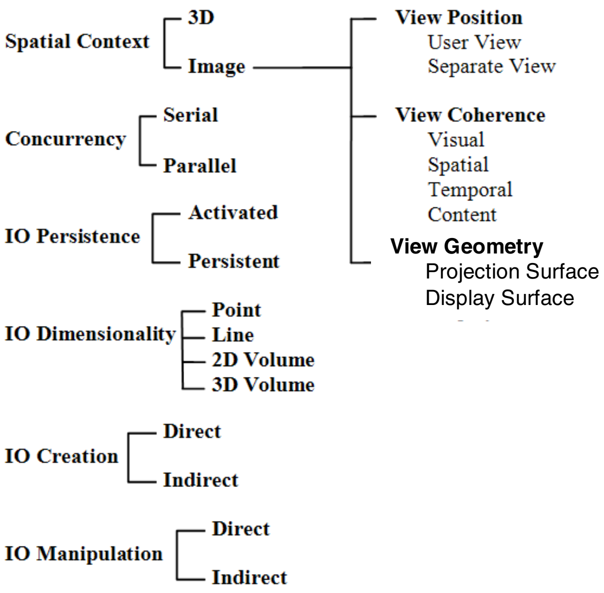

Figure 1. A Taxonomy of 3D MOS Techniques

Understanding of the design space for a class of interaction techniques is important when designing new techniques. Bowman et al. [4] used taxonomies to explore selection, manipulation and navigation techniques. Taxonomies are valuable as they enable the designer to understand what parameters of the technique can be modified to create new techniques. We present a taxonomy of the task of area differentiation for 3D MOS techniques. We refer the reader to Willis' [21] complete taxonomy of 2D selection to understand the possibilities of how objects can be chosen for selection after they have been indicated. We narrow our taxonomy to describe how objects are spatially indicated for selection in a 3D environment.

3D MOS techniques may have varying levels of *concurrency*. *Serial* techniques only allow one object selected per selection operation whereas *parallel* techniques allow one or more objects to be considered for selection. The properties of the indication object (IO) are also important for MOS. The IO may be *persistent* in the world and continuously consider indicated objects for selection or the selection operation may only be *activated* at the user's explicit command. The *dimensionality* of the IO is very important as IOs with a higher dimensionality may be able to select more objects but are also more difficult to constrain in three dimensions. Lastly the IO may be created and manipulated either *directly* or *indirectly*.

The *spatial context* of the MOS technique refers to the characteristics of the input used to specify the IO. Techniques using the *3D context* specify the IO directly in 3D space. Techniques using *image context* use a view of the environment mapped onto a 2D surface to specify the IO. There are several design choices when designing techniques using image context. Stoev and Schmalstieg [19] outline the possibilities for through-the-lens techniques but we simplify the space for MOS into *view position* and *view coherence*. Selection can be done either through the user's current view of the world [13], or by using a separate view of the environment. Views can vary along several parameters: *visual* (lighting, color palette, rendering style), *spatial* (exploded view), *temporal* (slow motion, still image), and *content* (a subset of objects in the world are displayed). Mapped views can also be created using non-planar geometry for both the *projection surface* and for the *display surface*.

4 EXPLORING THE DESIGN SPACE

		Concurrency	
		Serial	**Parallel**
Spatial Context	**3D**	Raycasting	Selection Box
	Image	Tablet Tapping	Tablet Lasso

Table 1. 3D MOS Techniques Evaluated in this Work

In order to explore the design space laid out in the previous section we designed four MOS techniques that represented different categories of our taxonomy. We were primarily interested in evaluating the effectiveness of parallel and serial techniques as we hypothesized that parallel techniques would be more effective at selecting large numbers of objects. We also wanted to explore the tradeoffs between techniques in both sub-branches of spatial context. Table 1 shows which category each our techniques fall under for concurrency and spatial context. Techniques in the image context branch can easily borrow 2D selection techniques while techniques using the 3D context can take advantage of true 3D input. All techniques were designed to be used in a head-tracked HMD with a 6-DOF wand device and a 6-DOF tracked tablet as a surface for displaying 2D interface objects. We used the Simple Virtual Environment (SVE) library to implement the four techniques as well as our testing framework [9]. Here we present the four techniques as originally designed. We later discuss our evaluation of these techniques and the improvements made on the designs.

4.1 Serial Techniques

Serial techniques were adapted directly from existing single object selection techniques. Minimal modifications were needed to these techniques to turn them into MOS techniques.

4.1.1 Ray-casting Ray-casting [12] was an obvious choice for a serial technique in the 3D spatial context. Ray-casting has been shown to be efficient for selecting objects from a distance with good accuracy [4]. Ray-casting has the advantage over arm extension techniques (such as the Go-Go technique [14]) for MOS as it only requires 2-DOF control and does not require the user to hold his arm extended out from the body during long selection tasks. Ray-casting allows users to "shoot from the hip" and select objects with their hand close to their body thus minimizing fatigue. Our design of ray-casting was limited to select (or deselect) only the first object the ray intersects, classifying it as a serial MOS technique.

4.1.2 Tablet Tapping One of the powerful features of desktop 3D modeling and CAD software is the ability to interact with multiple views of an environment at one time. We adopted a similar approach by mapping a separate view onto the tablet to facilitate remote 2D selection. We made use of the Pen and Tablet metaphor [1] to interact through the separate views. A tracked Plexiglas tablet provided a hard physical surface on which to place 2D widgets and extend 2D interaction techniques into 3D. In our design the user placed a camera object in the environment and positioned the camera to obtain the desired view of the target objects. The view of the camera object was mapped onto the user's tablet as a flat 2D image and was updated continuously. The user could then use the input device to interact with the camera view using 2D selection techniques similar to those found in modern desktop 3D modeling applications.

Positioning of the virtual cameras was accomplished by first traveling to the desired location of the camera. Pressing a button on the tablet started a three second countdown that was displayed on the viewplane. When the countdown finished the camera was moved to the exact position and orientation of the user's view. The three second countdown allowed the user time to precisely set their view before the camera was moved. We chose not to use camera in hand techniques [20] because in our pilot testing they did not prove effective for use in large environments where navigation was required.

The serial selection technique for the tablet is easily adapted from 2D single object selection. As the user moves the input device close to the camera view on the tablet, the virtual representation of the device becomes semi-transparent and a red crosshair appears over the camera view on the tablet surface. The crosshair is positioned at the closest point on the camera view plane to the device and moves along with the device as it hovers above the tablet surface. In our interface the crosshair appeared on the tablet so long as the device was within 10cm of the tablet; this gives the user the option of either touching the device to the tablet surface or hovering above the tablet. When the user presses a button on the device the object under the crosshair is selected (or deselected); this is implemented by casting a ray into the world.

4.2 Parallel Techniques

Our serial techniques both defined a one-dimensional ray to select objects in the environment. A ray can be used for selecting multiple objects in parallel but only if the target objects are aligned correctly. We designed two techniques that define 3D volumes to investigate parallel MOS techniques.

4.2.1 Selection Box Designing an effective 3D parallel MOS technique is a difficult task. To better understand the dimensions of a volume the user needs to be able to view the volume from multiple angles. When creating IOs that are 2D or 3D objects it is important to allow the user to use several viewing angles so that they may make an accurate selection. With this in mind we created a parallel 3D MOS technique using a persistent box as the IO. The box is semi-transparent and any objects touching or within the box are selected. We chose to implement this technique with a box but other shapes could work equally well depending on the structure of the environment and the shape of the objects to be selected.

The user may manipulate the box using 3D object manipulation techniques. Go-Go object selection and manipulation is used for adjusting the position and orientation of the selection box. The Go-Go technique scales the movement of the virtual hand only after the user's physical hand moves past a set threshold.

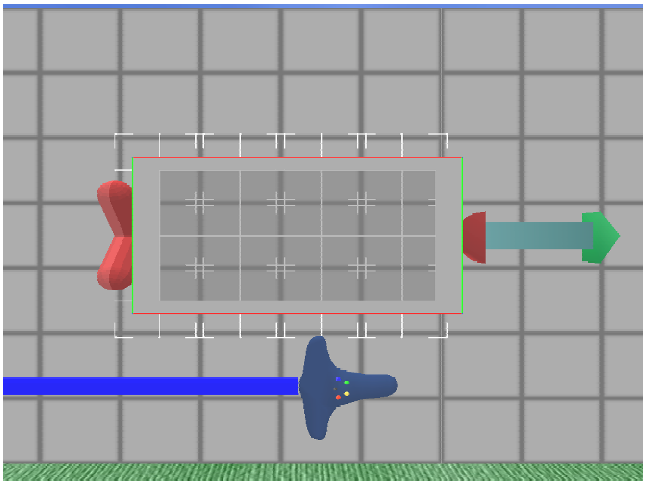

Figure 2. Pointer Orientation-based Resize Technique (PORT)

We developed a new object resize technique called PORT (Pointer Orientation-based Resize Technique) [11] to independently position all six faces of the selection box. PORT (Figure 2) allows only one axis to be manipulated at a time and uses the orientation of the pointer to determine which direction to resize. PORT does not uniformly resize the object but stretches the object in the direction the pointer is oriented. This gives the user the sensation that they are pulling out or pushing in the sides of the object. The system determines which of the object's axes the pointer is most closely aligned with by taking the dot product of the pointer's direction vector and the six direction vectors of the object's positive and negative x, y and z axes. The vector which is most coincident with the pointer's direction vector determines the axis on which resizing will occur. In practice the user simply needs to imagine that their pointer is contained inside the object and point at the face they wish to be stationary and the opposite face will be resized. Visual cues that indicate which side of the box will move and which side will not move are placed at the outer bounds of the object. Once the user has the desired axis selected the resize operation is performed one of two ways: by moving a joystick on the input device or by holding a button and translating the pointer along the pointer's direction vector. Once a set was selected the user could lock the selection and move the selection box without de-selecting the set. This feature allows the selection box to select any possible combination of objects.

4.2.1 Tablet Freehand Lasso Our final technique allows freehand 2D lassos (Figure 3) to be drawn on the camera view for parallel selection. To provide the user with maximum flexibility when creating 2D shapes freehand lassos were used rather than the standard box lassos found in most desktop applications. With the freehand lasso tool the user presses a button and moves the crosshair to draw an arbitrary shape on the camera view screen. Jonathan Shewchuk's Triangle [15] was used to create a 2D triangulation of the lasso outline. This 2D shape is then projected out from the camera into the environment to create a 3D volume. UNC's SWIFT++ library [7] was used to ensure quick and accurate collision detection between the lasso volume and the objects in the world. Like the ray-casting techniques, the selection operation toggles selection for indicated objects. We provided visual feedback from the last selection operation in the form of a wire frame "spider web" projected from the camera object (Figure 3).

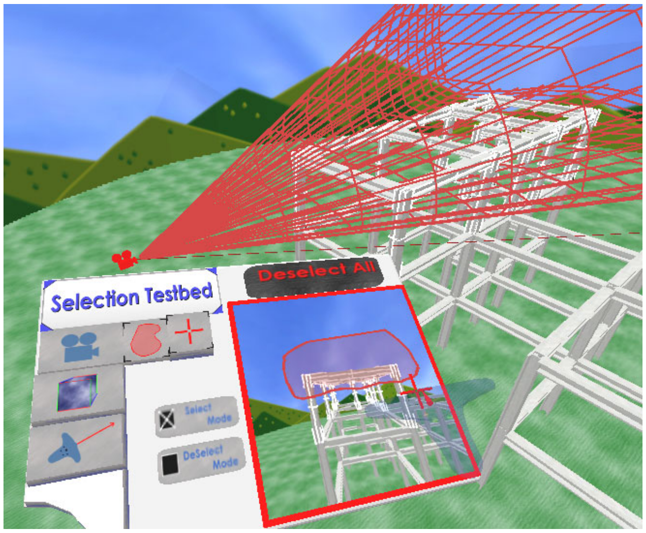

Figure 3. Tablet Lasso Technique

5 FORMATIVE EVALUATION

Exploration and refinement were the two main goals in our first evaluation of the techniques. Ray-casting is a well established technique that had been evaluated many times before and has been used in many applications. Though it is very common as a single object selection technique there had been no formal evaluation of ray-casting as a MOS technique. There had been little evaluation of tablet techniques as multiple selection techniques and there was no previous evaluation of the selection box. In order to be sure that our implementations of all four techniques were representative of the potential the techniques posses they all required evaluation and iteration.

5.1 Evaluation Platform

The experiment used a Virtual Research V8 head-mounted display (HMD) with 640x480 resolution and a 60° diagonal field of view. An Intersense IS900 tracking system provided tracking for the head, wand and tablet in a 10' x 10' area. The main input device was a tracked wand with four buttons and an analog joystick. For tablet interaction we used a custom designed 21" by 14" sheet of plexiglass that is held in the user's non-dominant hand. Graphics were rendered on a Pentium IV 2.4 GHz PC with a GeForce 4 TI 4600 128Mb graphics card running Windows XP.

5.2 Environments and Tasks

MOS techniques are applicable in an extremely wide variety of domains and applications. These applications can vary from being sparsely populated with only a few large objects to having millions of tiny data points. It is highly unlikely that one MOS technique will be the optimal technique for every possible immersive application. Therefore it is important to evaluate the techniques in a range of environment and task situations so that VE designers can choose the appropriate technique or set of techniques for their applications. We designed four different environments within which to evaluate our techniques. Motivation for each environment came from different domains and selection scenarios.

Building Environment: The initial inspiration for our MOS work came when designing interfaces for Virtual-SAP [3]. Our first environment was similar to the environments used in Virtual-SAP. The building environment was comprised of 98 rectangular beams and columns arranged to form the frame of a 3-story structure (see figure 3).

Point Box Environment: Vertex editing in 3D modeling and CAD applications can be a challenging task as the objects are very small and are often positioned close together. To simulate this challenge our second environment was a point mesh that represented the faces of a 15m x 10m x 6m box. Points in the mesh were represented by 15cm cubes and were spaced 1.5m apart.

Tile Wall Environment: In our experience, the volume based MOS techniques had difficulty accurately selecting objects that were adjacent or very close together. We were unsure if this was due to the nature of volume techniques or due to choices we made in our implementations. For our third environment we created a 9m x 3m wall made of 50cm x 50cm tiles set adjacent to each other (figure 2). All tasks were performed on the broad-side of the wall which was only 5cm thick.

Semi-Random Environment: The fourth environment had no structure or consistent theme. In our experience, the parallel MOS techniques tended to take advantage of the structure or patterns in an environment. We hypothesized that the lack of such structure would hinder performance of parallel techniques. Objects in the environment were placed, oriented and scaled at random in a 20m x 20m x 15m area. The objects were not allowed to overlap but could be adjacent. The environment is considered semi-random as we did not implement a random shape generator but instead used four basic shapes, scaled them and scattered them about the environment.

Participants used all four techniques to select two sets of objects in each environment for a total of 32 trials. The number of objects in the target sets ranged from 5 to 30 objects with an average of about 10 objects. The target sets in each environment were the same for all subjects and were chosen to reflect common sets of objects that might be selected in a real application.

5.3 Participants

We recruited 5 participants with VE experience ranging from VE expert to beginner. Participants self-reported their experience with VR, 3D games, and 2D MOS and had ranges from expert to beginner in all categories. One participant was female and four were male; all were graduate students. Their mean age was 24.

5.4 Procedure

Before the experiment participants was asked to fill out a pre-questionnaire with questions regarding their age, occupation, VE experience and visual acuity. Participants were given training on the techniques in a generic environment until they were reasonably proficient with all techniques. Participants selected two target sets in each of the four environments using all four techniques for a total of 32 trials.

Participants were encouraged to think aloud during the trials and were told to explore different selection strategies rather than complete the trials quickly. After each trial, participants were asked to provide a subjective rating of the effectiveness of the technique for selecting the target set. A post-questionnaire was used to collect the participant's final ratings of the techniques. All participants were interviewed to discuss the specifics of the techniques and to gain a deeper understanding of the challenges of each environment. With the training, trials and interview the experiment lasted approximately two hours.

5.5 Results

All subjective ratings were taken on a scale from 1 to 7 with 7 being the most preferable. In general, subjects preferred the serial techniques over their parallel counterparts in both the between-trial ratings and the ratings collected in the post-questionnaire. Ray-casting was the clear winner in every category (Figure 4); participants found it easy to learn and versatile in all environments. Though ray-casting was the most preferred of the techniques participants often said that it was "tedious" or as one user commented "this technique requires lots of pointing". Surprisingly, ray-casting was still rated as the least tiring of the techniques even though all users commented on the amount of pointing and clicking it required. Positive comments for ray-casting included "This ray makes me feel like I'm part of the world." and "It's easier to correct a problem if you mess up".

Participants felt that the selection box could be a very effective technique but all had difficulties precisely orienting the box. The selection box was rated as the second most frustrating technique and participants found that it was especially difficult to use when objects were very close together. It was also difficult for users to tell how far the back face extended when viewing the box from one side. There was often confusion when an object was behind the box but not selected. The transparency of the box allowed the user to see the desired objects but gave no sense of their distance from the far side of the box. Participants felt that selection box was a very intuitive concept but that it was limited by the lack of a precise manipulation technique.

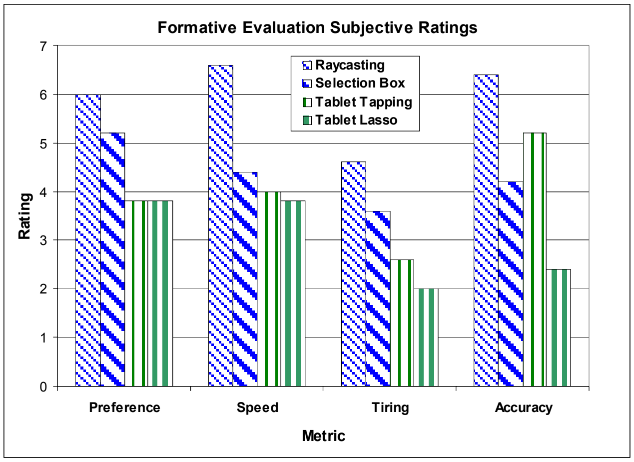

Figure 4. Formative Evaluation Subjective Ratings

Tablet techniques suffered from three main problems: the accuracy of the tracking system, the size of the camera screen, and the difficulty of positioning the camera. Precise manipulation on the tablet was very challenging due to jitter, which created large problems when selecting small objects or when trying to draw straight lines with the lasso. The limited size of the camera screen required users to position the camera very close to small objects to make selections with the tablet tapping technique. This limited the power of the techniques as fewer objects could be seen in one view and required more camera manipulation to select all targets. Camera positioning received the largest number of complaints because the user had to travel to the desired location and call the camera to this position when they arrived. This technique did not take advantage of the camera's ability to obtain a view of the environment that is different from the user's perspective. Aside from general camera issues the two tablet techniques both performed well. Participants were able to understand the mapping from the camera's view to the screen on the tablet and none had problems interacting with the screen as a 2D surface.

5.6 MOS Components

While watching users select objects with our techniques we made an important observation. 3D MOS tasks are composed of three components:

- *Navigation* to a location in the environment from where the technique can be performed
- *Manipulation* of the IO to indicate the target objects
- *Execution* of the selection operation

Depending on the specifics of the technique the time required to accomplish each component can vary greatly. Parallel techniques that specify a volume for selection require much more time to manipulate the IO than serial techniques such as ray-casting. The number of repetitions of this procedure will also vary greatly depending on the technique. Serial techniques will repeat this procedure once for every object selected whereas parallel techniques may only perform the procedure once for the same set. The amount of navigation required and the number of repetitions of the procedure are highly dependent on the characteristics of the target set. Sets whose elements are very far apart will likely require more navigation than sets that are densely packed.

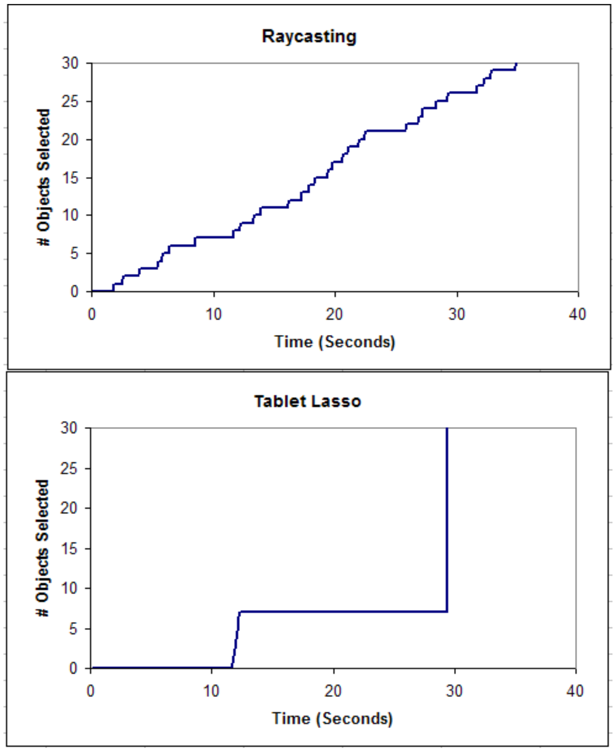

Figure 5. Two MOS Component Graphs

In order to visualize when navigation, manipulation and execution were taking place we instrumented our testbed to graph the behavior of our MOS techniques. Figure 5 shows the graphs of an expert user performing tasks with two of our techniques. The graphs show the number of correctly selected target objects as time goes by. Vertical movements of the line indicate execution of a selection operation where horizontal movements indicate navigation and manipulation. Examining these graphs, a designer can determine where the bulk of the interaction time is spent and can concentrate their efforts on improving a particular component of the technique.

5.7 Design Iteration

Using our observations and the participants' comments to improve our techniques we made several changes. Ray-casting required little iteration as its simple design limits the number of changes that can be made without transforming it into a different technique. We simply added more visual feedback to the ray by brightening its color when the ray intersected an object.

A new method for positioning and orienting the box was also needed. Go-Go did not allow the user to effectively manipulate the box at a distance and the box was almost completely useless when close to the user as it often occluded their view. We chose to implement HOMER [2] for selection box manipulation as it allows the user to quickly select the box with ray-casting and precisely move and orient the box from a distance. To counteract the selection box's lack of occlusion cues we implemented a variant of the silk cursor [22]. In our variant objects behind the box were completely occluded, objects inside the box were seen through a layer of silk, and objects in front of the box were viewed as normal.

The size of the camera screen for the tablet techniques was increased by 24% in order to counteract some of the problems caused by tracker jitter. The lasso previously used a selection toggle operation and participants found this non-intuitive as well as frustrating. Select and de-select modes were added to the lasso technique; modes were changed with buttons on the tablet.

6 SUMMATIVE EVALUATION

The most surprising result of our formative evaluation was the poor usability of the parallel techniques. Intuitively, parallel techniques should provide a distinct advantage when selecting a large number of objects. In our second experiment, we aimed to demonstrate these benefits and to determine why the parallel techniques were less successful in the first study.

6.1 Environment

A limitation of our first experiment was that the number of objects in the target sets was rather small. With an average of 10 objects per set the advantage of parallel selection over serial selection is minimal. What interested us most was what the performance of each technique would be among expert users. We did not have a large subject pool of expert users nor did we have enough time to train every participant to the performance level of an expert user. Therefore, in this study we attempted to create a task that was simple enough so that novice users could perform at the level of expert users with moderate training.

Observations from our first study showed that time required for the navigation component of selection could vary greatly depending on the technique and the target selection set. As much is already known about navigation techniques we decided to eliminate this variable from our second study and seat the user in one place. We chose to create an environment where there was no occlusion and objects were always in the user's view. This simplified task is still interesting because it still requires precise manipulation of the selection tools in order to avoid selecting the bounding set of objects.

With these goals in mind we designed an environment (figure 6) that consisted of a 13m x 13m wall made up of 100 cubes spaced 1.3m apart (center-to-center). The wall was positioned 10m away from the user so that it could be viewed without the user having to turn their head. The camera used in the tablet-based techniques was positioned so that all cubes were visible on the camera screen. The cubes were textured to contrast highly with the background.

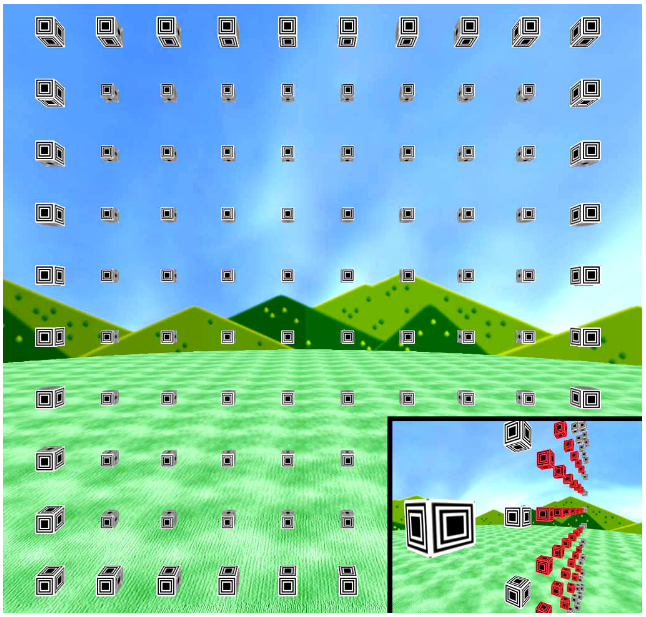

Figure 6: Summative Experiment Environment

6.2 Variables and Tasks

We chose four independent variables for our trials: concurrency, spatial context, number of targets and target size. The first two variables corresponded with the portion of the MOS design space we are exploring with our four techniques. By varying the number of targets we expected to demonstrate that as the number of targets increases the performance gap between parallel and serial tasks increases. The size of the objects seemed to have a significant effect on the ease of selection for the serial techniques. Selecting small objects from a distance required more precise movement of the input device than that required to select large objects. We had also observed that our parallel techniques had difficulty selecting objects when the targets were very close together. When the space between targets was small participants would often unintentionally select non-target objects. We chose to use target size as a variable in our experiment so that we could better understand its effect on the techniques.

The task for this experiment was simple; the participant was to select all of the red cubes on the wall and none of the grey cubes. When this condition was met the trial ended and a wall was placed in front of the user blocking their view. The target (red) cubes were always in the center of the wall forming a square. Grey cubes always bounded the target cubes so that the participants had to constrain the volume techniques for all tasks. Trials started when the participant pressed a button on the wand and ended upon selection of the entire target set. To cut down on experiment length we chose to use a mixed methods design. Participants were assigned either 3D or Tablet techniques and repeated all combinations of the other independent variables three times. Trials were blocked by technique; the ordering of the other variables was randomized.

The dependent variable for our study was time to select all targets. Values for the independent variables were as follows:

- Technique Concurrency: Serial, Parallel
- Technique Spatial Context: 3D, Tablet
- Number of targets: 9, 36, 64
- Target size: 30cm, 37.5cm, 45cm

6.3 Participants

Eighteen participants were recruited for the experiment. Nine subjects were assigned the tablet techniques and nine were assigned the 3D. Participant ages ranged from 20 to 35; six participants were female and twelve were male. Eleven subjects were computer science undergraduates and seven were from various other majors and professions.

6.4 Procedure

The procedure for the second experiment was identical to the first experiment with two exceptions. Participants were required to reach a minimum level of proficiency with both techniques before beginning the measured trials. Proficiency was measured by a minimum time to select a set of 36 medium-sized targets. During the experiment participants were asked to complete the trials as quickly as possible, holding questions or comments until breaks.

6.5 Objective Results

To estimate the expert performance of the techniques we analyzed the data using only the fastest trials of each condition for each participant. We used a multi-factor analysis of variance (MANOVA) on this data and found main effects of concurrency, target size, and number of targets. Figure 7 shows representative results from the medium object size condition. A very strong effect of concurrency ($p < .0001$) indicates that the parallel techniques were significantly faster than their serial counterparts. We also found a strong interaction between concurrency and number of targets ($p < .0001$). As the number of targets increased the performance gap between parallel and serial techniques widened. Both serial and parallel techniques performed well for the small number of targets (9), but as the number increased the advantage of using parallel techniques became clear.

We hypothesized that small objects would be harder to select using serial techniques. Our analysis showed that serial techniques were affected by changes in size whereas parallel techniques were not. A least squares means test did not show an effect of target size on parallel techniques ($p = .4108$) and a large effect of size on serial techniques ($p < .0001$). We found no effect of spatial context in our results. This indicates that neither 3D nor tablet techniques were superior for this MOS task.

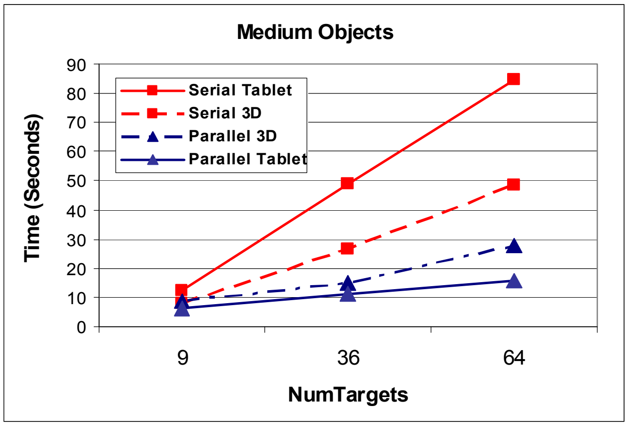

Figure 7. Best Trial Averages for Medium Targets

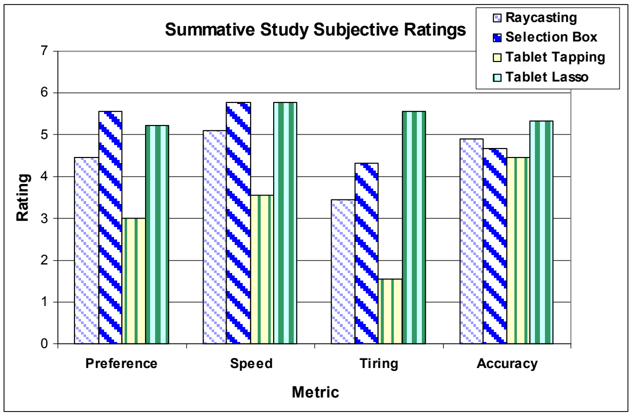

Figure 8. Subjective ratings in the summative evaluation

6.6 Subjective Results

Participants were asked to fill out a post-questionnaire identical to the questionnaire in the exploratory study. The results differed greatly from those of the first experiment in that the ratings of the parallel techniques had greatly improved (Figure 8). Parallel techniques received higher ratings than serial techniques in all categories except accuracy. From our observations we attribute the improvement of the parallel techniques to the iterated design of the techniques and to the optimal nature of the task for volume selection techniques. The lack of camera manipulation also contributed to the improvement of the tablet techniques. Comments were in line with the subjective ratings: “the lasso was quicker but a little more inaccurate”, “the selection box was much faster, but ray-casting was OK for a few boxes”. One participant commented that manipulating the selection box was mentally tiring and that ray-casting was physically tiring. We also recorded observations and user comments regarding the usability of the redesigned techniques. This information will be used in the future for continued iteration of our designs.

6.7 Discussion

Participants performed well when using both the 3D and tablet techniques. The tablet techniques were much easier for participants to learn due to their similarity with 2D desktop selection techniques. The tablet techniques also constrained the movement of the pointer to the handheld 2D surface simplifying movement as well as benefiting from the user's proprioceptive sense. Tracker jitter was a problem for both tablet techniques and should be accounted for by designers. Extended use of the tablet also caused back and neck discomfort for the participants making it less than ideal as the primary mode of interaction.

Serial MOS techniques, due to their nature, slow down linearly as the number of target objects increases. Serial selection is very effective for selecting small numbers of objects as it often requires very little manipulation per selection. As the number of targets increases the number of selections and the time to complete the task increases. Our serial techniques both specified a 1D ray that was simple to manipulate using a 6-DOF input device. Due to the accuracy required to precisely orient the ray our techniques were greatly affected by the size of the target objects. The lower the dimensionality of the IO, the greater the effect of object size on the difficulty of selection.

The parallel techniques that we implemented require much more navigation and manipulation before selection can occur. This is due partly to the higher dimensionality of the IOs specified with the parallel techniques. With both of the parallel techniques the IO is a 3D volume. Accurately describing a volume is a difficult task that requires powerful and expressive 3D interaction techniques. Parallel techniques can make up for their longer set-up time by requiring fewer repetitions to select a set. Parallel MOS techniques are not affected by object size but are affected by the shape of the target set and its proximity to non-target objects.

7 Conclusions and Future Work

3D multiple object selection is an important and relatively unexplored interaction task. In our research, we have designed novel techniques for 3D MOS, developed a taxonomy to represent the design space for these techniques, and used the taxonomy as a framework for evaluating the techniques both formally and informally. Our first study showed that serial techniques are easier to understand and require less initial setup time. The usability of parallel techniques is much more subject to the effects of user strategy and sophistication. In our second study, however, we showed that parallel MOS techniques can be superior to serial MOS techniques under certain conditions:

- The number of target objects is large.
- The technique takes advantage of the environment's structure.
- The target sizes make the use of serial techniques challenging.
- The time spent navigating or setting up camera views can be minimized.

Asking whether to use serial or parallel MOS techniques in an application is the wrong question. The important question is whether to use parallel techniques *in addition* to serial techniques. When the conditions outlined above are not met, parallel selection may add unneeded complexity to an application. On the other hand, in the right conditions adding parallel selection techniques to an application could greatly speed up selection tasks.

Our understanding of 3D MOS techniques is still very limited. The design space of 3D MOS techniques is very large and much exploration needs to be done in order to understand the advantages and disadvantages of the design choices for each branch. The techniques we examined have proven effective for the generic environments in our experiments. A key challenge for the design of usable parallel techniques is to minimize the time to navigate and manipulate the IO. Potential solutions to these problems include: more specialized manipulation techniques for working with objects at a distance, constraining manipulation of the IO to a grid or embedding more depth information into the IO itself. Much can be adapted from 2D interfaces as far as IO manipulation. Photo editing software allows for complex area selections with specialized lassos and masks. These concepts could be adapted to 3D selection for advanced interfaces.

Selection of non-contiguous sets of objects is a large challenge for parallel MOS techniques. As parallel techniques often select an area they can be difficult to use when a small set of undesired objects resides in the selected area. Additional Boolean selections can be used to deselect the undesired objects after an initial selection is made, adding of course to the complexity of the technique as well as the total selection time. Future work should look to a) design techniques that are more flexible in the selection of non-contiguous sets and b) investigate how to add Boolean selection operations to 3D MOS.

In our experiments we only examined selection using spatial attributes. Filtering objects from consideration using non-spatial attributes such as color or shape has been proven to be effective in 2D information visualization applications, and should be investigated for 3D MOS techniques. Other information visualization techniques such as linked views could be useful for selection in dense environments.

We have presented a first look at the task of 3D multiple object selection with the hope that it will inspire future work in designing multiple object interaction techniques for immersive environments. With more powerful interfaces, both Virtual and Augmented Reality can grow into new domains that were previously impractical with existing interaction techniques.

Acknowledgements

Thanks to Chris North, Deborah Hix, Joe Gabbard, Bob Schulman, Susanne Aref and the Virginia Tech 3D Interaction Group for their assistance and encouragement throughout this project. This work was partially funded by National Science Foundation grant IIS-0237412.

References

1. Angus, I., Sowizral, H. 1995, Embedding the 2D interaction metaphor in a real 3D virtual environment. *In Proceedings of SPIE (Stereoscopic Displays and Virtual Reality Systems)*, 282-293.
2. Bowman, D., Johnson, D., Hodges, L. 2001, Testbed Evaluation of Virtual Environment Interaction Techniques. *Presence: Teleoperators and Virtual Environments*, 10(1), 75-95.
3. Bowman, D. Setareh, M., Pinho, M.S., Ali, N., Kalita, A., Lee, Y., Lucas, J., Gracey, M., Kothapalli, M., Zhu, Q., Datey, A., and Tumati, P. 2003, Virtual-SAP: an immersive tool for visualizing the response of building structure to environmental conditions. *Proceedings of the IEEE Virtual Reality,* 243-250.
4. Bowman, D., Kruijff, E., LaViola, J., Poupyrev, I. 2004, *3D User Interfaces: Theory and Practice, Addison-Wesley Professional*
5. Chen, J., Bowman, D., Lucas, J., Wingrave, C. 2004, Interfaces for Cloning in Immersive Virtual Environments. *Proceedings of the Eurographics Symposium on Virtual Environments*, 91-98,
6. Darrah, M., Kime, A., Van Scoy, F. 2002, A Tool for Editing 3-D Objects: Implemented Using a Haptics Device. *Proceedings of PHANTOM User Group Workshop.*
7. Ehmann, S., Lin, M. 2001, Accurate and Fast Proximity Queries between Polyhedra Using Surface Decomposition Computer Graphics Forum *(Proceedings of the Eurographics Symposium on Virtual Environments)*
8. Haan, G., Koutek, M., Post, F. 2002, Towards Intuitive Exploration Tools for Data Visualization in VR, *Proceedings of the ACM symposium on Virtual reality software and technology*, 105-112
9. Kessler, G., Bowman, D., Hodges, L. 2000, The simple virtual environment library: an extensible framework for building VE applications. *Presence: Teleoperators and virtual environments*, 9(2), 187-208.
10. Liang, J., and Green, M. 1994, JDCAD: A highly interactive 3D modeling system. *Computers & Graphics*, 18(4), 499-506.
11. Lucas, J., Kim, J., Bowman, D. 2005, Resizing Beyond Widgets: Object Resizing Techniques for Immersive Virtual Environments, *To be included in Proceedings of CHI 2005*
12. Mine, M., 1995, Virtual environment interaction techniques, *UNC Chapel Hill Computer Technical Report*, TR95-01.
13. Pierce, J., Forsberg, M., Conway, M., Hong, R., Zeleznik, R., Mine, M., 1997, Image plane interaction techniques in 3D immersive environments. *In Proceedings of SIGGRAPH Symposium on Interactive 3D Graphics*, ACM, 39-44.
14. Poupyrev, I., Billinghurst, M., Weghorst, S., Ichikawa, T. 1996, The Go-Go Interaction Technique: Non-linear mapping for Direct Manipulation in VR. *In Proceedings of User Interface Software and Technology*, 79 – 80.
15. Shewchuk, J. 1996, Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator, *In First Workshop on Applied Computational*, ACM, 124-133.
16. Schmalstieg, D., Encarnação, L., Szalavári, Z. 1999, Using transparent props for interaction with the virtual table, *Proceedings of the 1999 symposium on Interactive 3D graphics*, 147-153
17. Steed, A., Parker, C. 2004, 3D Selection Strategies for Head Tracked and Non-Head Tracked Operation of Spatially Immersive Displays. *8th International Immersive Projection Technology Workshop*
18. Szalav´ari, Z., Gervautz, M. 1997, The personal interaction panel - A two handed interface for augmented reality. *Computer Graphics Forum (Proceedings of EUROGRAPHICS'97),* 16(3), 335–346.
19. Stoev, S., Schmalstieg, D. 2002 Application and taxonomy of through-the-lens techniques, Proceedings of the *ACM symposium on Virtual reality software and technology*, 57-64
20. Ware, C., Osborne, S. 1990, Exploration and virtual camera control in virtual three dimensional environments. In *Proceedings of the 1990 Symposium on Interactive 3D Graphics*, *Special Issue of Computer Graphics*, Vol. 24, 175–183.
21. Willis, G. 1996, Selection: 524,288 ways to say "this is interesting", *Proceedings of the 1996 IEEE Symposium on Information Visualization*, 54-60.
22. Zhai, S., Buxton, W., Milgram, P. 1994, The "Silk Cursor": Investigating Transparency for 3D Target Acquisition, *In Proceedings of Computer Human Interaction*, 459-464.