

# An Evaluation of Virtual Lenses for Object Selection in Augmented Reality

Julian Looser\*  
Mark Billingham†  
Raphaël Grasset‡

HIT Lab NZ, University of Canterbury, New Zealand

Andy Cockburn§

Department of Computer Science and Software Engineering  
University of Canterbury, New Zealand

## Abstract

This paper reports the results of an experiment to compare three different selection techniques in a tabletop tangible augmented reality interface. Object selection is an important task in all direct manipulation interfaces because it precedes most other manipulation and navigation actions. Previous work on tangible virtual lenses for visualisation has prompted the exploration of how selection techniques can be incorporated into these tools. In this paper a selection technique based on virtual lenses is compared with the traditional approaches of virtual hand and virtual pointer methods. The Lens technique is found to be faster, require less physical effort to use, and is preferred by participants over the other techniques. These results can be useful in guiding the development of future augmented reality interfaces.

**CR Categories:** H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Interaction styles

**Keywords:** augmented reality, selection technique, lens

## 1 Introduction

Augmented Reality (AR) [Azuma 1997] offers a way to merge virtual graphics with the real world in real time. Virtual objects can be overlaid onto tracked physical objects to augment the user's view and support tangible user interaction [Ishii and Ullmer 1997].

Tangible Augmented Reality (TAR) [Kato et al. 2001] is an interaction concept that combines Tangible User Interface (TUI) input techniques with augmented reality. Like traditional user interfaces, those based on Tangible Augmented Reality require a set of basic interaction techniques the user can engage to perform specific actions in the interface. In 3D environments, such as those using AR and VR, interaction techniques are generally categorised into selection, manipulation, navigation and system control [Bowman et al. 2004].

Selection is fundamental because it precedes other actions. The user selects the objects they wish to manipulate and the locations to which they wish to navigate. For example, to orient and position a house in an immersive urban planning tool, the user must first select the component (or set of components) that make up the house from

the set of all components in the scene. Only then will subsequent manipulation actions be applied to the correct objects.

Performance with a particular selection technique can be measured via a set of metrics such as time taken to make the selection, accuracy of selection, and number of errors. Qualitative measures such as mental and physical effort are also important. The features of the environment that can affect selection performance include the target object's size and distance (objects that are small, due either to geometrical size or range from the viewpoint, are more difficult to select), the density of surrounding objects (distractors) and visibility (occluders) [Bowman et al. 2004].

Previous work has investigated the use of virtual lenses as a visualisation technique in AR. Following the TAR approach, a virtual magnifying glass has been attached to a physical handheld prop [Looser et al. 2004]. The user's view is now partitioned into a primary view and a secondary view seen through the lens of the magnifier. The lens view is generalised in the style of Magic Lenses [Bier et al. 1993], such that the content seen through it, and the style in which it is visualised, can be configured independently of the primary view.

The combination of Tangible Augmented Reality and Magic Lenses has many intriguing applications. Previous research has begun to explore these applications, however until now there have been few formal evaluations conducted. In this paper the filtering and visualisation aspects of the tool are temporarily set aside and the extent to which the virtual lens supports object selection is concentrated upon. Selection is an important addition to the virtual lens tool because it expands its role in the user interface beyond visualisation support, to that of a more general purpose instrument.

In the next section, work related to selection in virtual environments is reviewed. The experiment we conducted is then presented, followed by a discussion of the results and suggestions for further research.

## 2 Related Work

A large number of selection techniques have been proposed and implemented for immersive virtual environments. Although the field is large, it is generally well understood due to the many rigorous evaluations that have been carried out (e.g. Bowman, Johnson and Hodges [2001]) and the development of taxonomies to structure our understanding. For example, the classification-by-metaphor taxonomy of Poupyrev et al. [1998] is treated as a de facto standard. At the highest level, it partitions the space of selection techniques into either *exocentric*, those that operate from a third-person perspective, or *egocentric*, those that operate from a first-person perspective.

A good example of an exocentric technique is World-in-Miniature (WIM) [Stoakley et al. 1995]. The WIM technique displays a small copy of the virtual environment in front of the user who can then use it as a proxy for object selection and manipulation and as a navigation aid.

Egocentric techniques are generally more interesting for AR be-

\*e-mail: julian.looser@hitlabnz.org

†e-mail: mark.billinghurst@hitlabnz.org

‡e-mail: raphael.grasset@hitlabnz.org

§e-mail: andy@cosc.canterbury.ac.nz

cause AR interfaces are anchored to the user's real view, and therefore favour a first-person perspective. Egocentric techniques are further categorised into those that follow the Virtual Hand metaphor and those that follow the Virtual Pointer metaphor. Virtual Hand techniques involve directly touching target objects (either through close proximity or collision) whereas Virtual Pointer techniques involve indirectly designating targets from a distance (such as by a virtual ray).

The most basic Virtual Hand technique is a direct mapping between real and virtual hand motion. This mapping can be manipulated to create new techniques, such as Go-Go [Poupyrev et al. 1996], which introduces a non-linear relationship between the offset of the user's physical hand and the offset of the virtual hand to greatly increase the user's reach within the virtual environment. In AR, Virtual Hand techniques can be implemented by tracking the user's fingers [Piekarski 2004], or through a tracked handheld tool.

Ray-casting is the simplest Virtual Pointer technique. A ray originating at the user's virtual hand shoots out in the direction they are pointing. Typically the first object to be hit by the ray is selected, however often selecting the first object is not ideal. Recently Grossman and Balakrishnan [2006] explored various disambiguation mechanisms for multiple target intersections for 3D volumetric displays, finding an enhancement called *Depth Ray* to perform faster and with fewer errors.

A weakness of ray-casting is that a slight change in angle at the origin of the ray equates to an increasingly large change in angle along the ray. Therefore, selecting small or distant objects can be difficult. There are several variations of ray-casting that address this problem. Cone-casting [Liang and Green 1994] uses a cone to select objects based on their relative distances from the ray. Objects that are far from the user are allowed to be further from the ray and still be selected. Shadow Cone [Steed and Parker 2004] is a further refinement that selects objects that remain continuously within the cone while selection is active. This provides the user with finer control for complex selection tasks with a high level of occlusion as they can modify their selection on the fly.

Aperture selection [Forsberg et al. 1996] is a cone-based technique where the cone originates the user's eye-point and passes through a circle defined by a tool held in the user's outstretched hand. The direction of the cone is controlled by moving the tool left, right, up and down, and the spread of the cone is controlled by moving the tool nearer or farther away.

Image Plane selection techniques [Pierce et al. 1997] reduce 3D object selection to a 2D task by operating on the 2D projection of the 3D scene. The Sticky Finger version of this technique casts a ray from the user's eye, through their finger on the 2D projection plane, and out into the scene, selecting the first object to be hit. This can be considered a limiting case of aperture selection, where the spread of the cone is effectively zero.

For a more detailed survey of selection techniques for 3D virtual environments refer to Bowman et al. [2004]. Most techniques designed for VR are easily adapted for AR interfaces. In contrast, some selection techniques have emerged as a result of features or needs specific to AR. For example, in AR interfaces that use fiducial tracking the act of occluding certain markers can indicate selection [Lee et al. 2004].

Recently, de Boeck et al. [2006] carried out a formal evaluation comparing three selection techniques (direct touch, ray-casting and aperture) in a virtual environment. They tested each technique's performance when controlled by both the dominant or non-dominant hand and found aperture selection to be the significantly fastest technique, even performing faster in the non-dominant hand

than the other two techniques in the dominant hand. There is a shortage of similar evaluations for AR interfaces, which has motivated the experiment we present in this paper. Our experiment, described in the following section, is an incremental step towards better understanding object selection in augmented reality.

### 3 Experiment

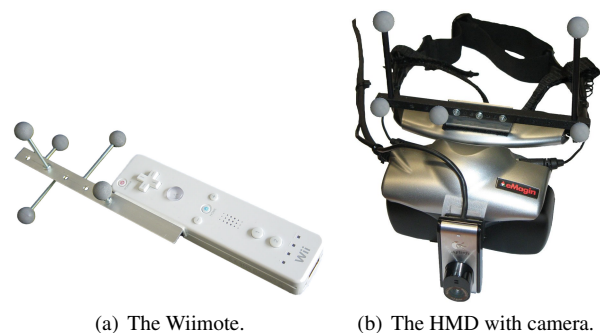
The goal of this experiment is to compare a selection technique built for existing AR virtual lens interfaces with two traditional techniques based on the approaches of virtual hand and virtual pointer: direct touch and ray-casting.

#### 3.1 Apparatus

The experiment was run on what has become a familiar desktop AR configuration: a webcam attached to a head-mounted display, each connected to a computer running ARToolKit based software. In this case, the camera was a Logitech Notebook Pro (640x480 pixels at 30FPS), the HMD was an eMagin Z800 (800x600 pixels, 40° field of view) and the computer was a 3.2GHz PentiumD Shuttle PC. The view provided by this system was monocular.

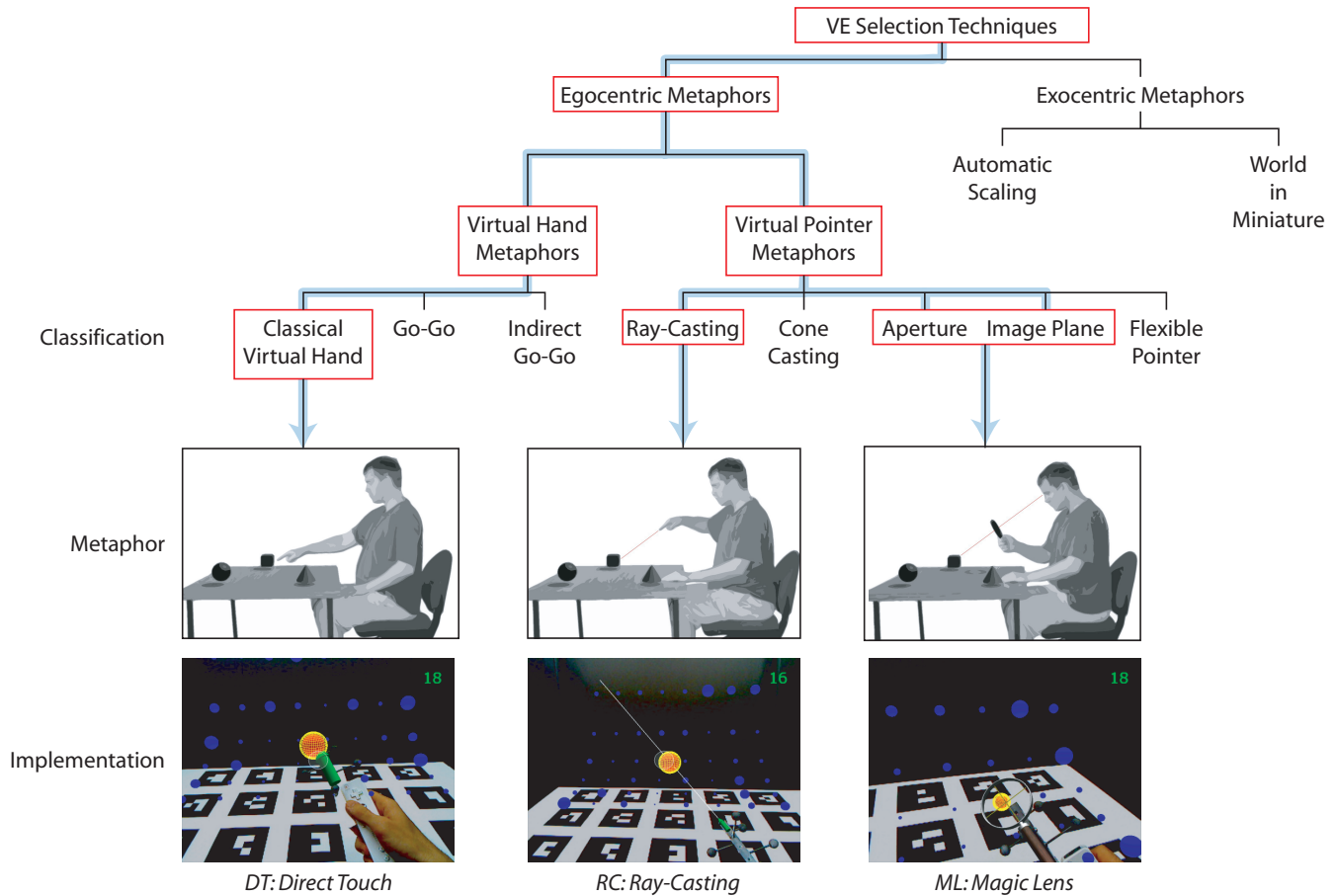
The test application was built on top of the osgART library [Looser et al. 2006]. The osgART software integrates vision-based trackers, such as ARToolKit, with the Open Scene Graph [Osfield 2007] to simplify and expedite the development of interactive AR applications.

Initially, it was intended to solely use ARToolKit for tracking, but it became apparent that the limitations of fiducial tracking would compromise the experiment. The particular problems were frequent marker occlusion and the need to track objects outside the field of view of the user's head-mounted camera. To remedy these problems our experiment was run within a visualisation center equipped with an ART infrared optical tracking system [Advanced Real-time Tracking GmbH 2007]. The ART tracker uses high resolution cameras to track constellations of retro-reflective spheres. While ARToolKit was still used to track the desk at which the user sat, the ART system was used to track the handheld tool and the user's HMD (so that their head movements could be recorded relative to the room). The two tracked peripherals are shown in Figure 1.



**Figure 1:** *The Wiimote controller and HMD with tracking constellations attached.*

The handheld tool was a Nintendo Wiimote. The Wiimote is a wireless Bluetooth controller with a number of buttons, orientation sensors, a speaker and a vibro-tactile actuator built in. The Wiimote and the cWiiMote library [Forbes 2007] was used to communicate with the device.



**Figure 2:** An overview of the three techniques being evaluated. Direct Touch (left) is an implementation of the classic Virtual Hand technique. Ray-Casting (middle) is the most basic Virtual Pointer technique. The Lens technique (right) is also a Virtual Pointer technique. The Lens can implement either the Aperture or Image Plane selection, although Image Plane was chosen for this experiment. All three techniques are classified as egocentric. This diagram was adapted from that of Poupyrev et al. [1998].

## 3.2 Techniques

The three evaluated selection techniques are described in this section. Figure 2 illustrates the classification of these techniques, the metaphor on which they are based, and a screenshot taken of the participant's view during the experiment.

### 3.2.1 Direct Touch

Direct Touch is a Virtual Hand technique that allows the user to select an object by simply reaching out touching it. In this interface, a handheld tool was used to approximate the user's actual hand. A virtual arrow was rendered on top of the tool in the augmented reality view. There was a 1:1 mapping between the user's hand position and the position of the virtual arrow. To make a selection, the user simply needed to reach out until the tip of the arrow intersected their desired target, and pull the trigger button on the controller.

### 3.2.2 Ray-Casting

Ray-Casting is a Virtual Pointer technique that creates a virtual line originating at, and aligned with, the user's hand and tests whether it intersects with objects in the scene. In this interface, out of all

objects hit by the ray, the closest one to the user was selected. The ray was rendered as a thin white line extending from the controller into the scene.

Ray-Casting offers several potential advantages over Direct Touch. Firstly, it decouples selection from the user's viewpoint so that the user can observe and select objects from different vantage points. Secondly, Ray-Casting permits selection at a distance, although objects become more difficult to select the further away they are because a small change in angle at the ray's origin can move the ray selection point a great distance.

### 3.2.3 Lens

The Lens is a Virtual Pointer technique designed for use with existing AR Magic Lens interfaces. Typically the view through an AR Magic Lens is enhanced by some effect such as X-Ray vision or the addition of different data layers. The user accesses these alternate views by looking through the lens as they pan it over the virtual scene, moving it nearer and farther from their eyes to reduce and enlarge its size simply via perspective. This style of usage, as well as the lens tool's circular shape, suggest that Aperture would be an appropriate selection technique. Another option is to employ a variation of the Sticky Finger Image Plane technique where the ray

from the user’s eye, passing through the center of the lens, is used to determine object selection. The Image Plane option was chosen for this experiment because it does not require any movement of the lens in the Z-axis (depth). This was considered important because such movement is already used to control the size of the lens, and is also the means to adjust magnification, if provided by the lens.

A model of a magnifying glass was rendered at the position of the handheld tool in the augmented view. The lens area of the model had a crosshair drawn inside to indicate the center point that would be used for selection.

### 3.3 Participants

Sixteen participants (15 male and 1 female), ranging from ages 23 to 39, were recruited from our lab. All participants had previous experience with AR interfaces. This was a deliberate decision as we wished to investigate “expert” performance with various selection techniques and wanted to minimise the “wow factor” that routinely occurs with a participant’s first encounter with augmented reality. All but one participant were right handed. The handheld controller is symmetrical and all participants used their dominant hand during the experiment.

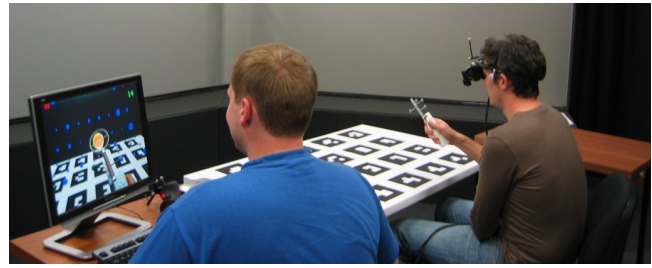
### 3.4 Procedure

The task in the experiment was to select a single object from a set of static targets. The targets were virtual blue spheres that appeared to hang in space above the table at which the participant was sitting. They were arranged in a curved grid. This arrangement was chosen to limit the amount of occlusion between targets.

In 2D user interfaces Fitts’ Law is a standard tool for evaluating selection performance. Fitts’ Law relates human movement time to the distance and size of a target [Fitts 1954]. These parameters are used to compute an Index of Difficulty (IoD) for each target. In this experiment six IoD values were chosen. The targets, which were distributed spatially in front of the user, were randomly assigned across the six groups. Once a target was assigned to a particular group, its radius was adjusted to satisfy the difficulty for that group, via the Fitts’ Law equation.

In each condition of the experiment, the participant was assigned one of the three selection techniques and carried out eighteen selection tasks. Each such task began with the participant selecting a starting target. This was a red virtual sphere that always appeared in the same place: the front-center of the virtual scene. The center of the starting target was the point from which the distance to each target was measured for Fitts’ Law calculations. Once the starting target was successfully selected, it would disappear and one of the many blue targets would turn red. The participant then had to select that target as quickly and accurately as possible. A miss caused a short error sound to play and the participant had to select again. A hit caused a success sound to play and the scene was reset to show the starting target once again, ready for the next task. A photograph of a participant performing tasks in the experiment is shown in Figure 3.

Participants had the technique demonstrated to them and carried out a practice round with that technique. They completed a questionnaire before, during and after the experiment. An initial questionnaire collected demographic data. A questionnaire consisting of seven questions was completed for each of the six conditions. The questions, shown in Table 3, were answered on a seven-point Likert scale ranging from ‘disagree’ to ‘agree’. A final questionnaire collected overall preference, summary information and comments. Participants were compensated for their time with a gift voucher.

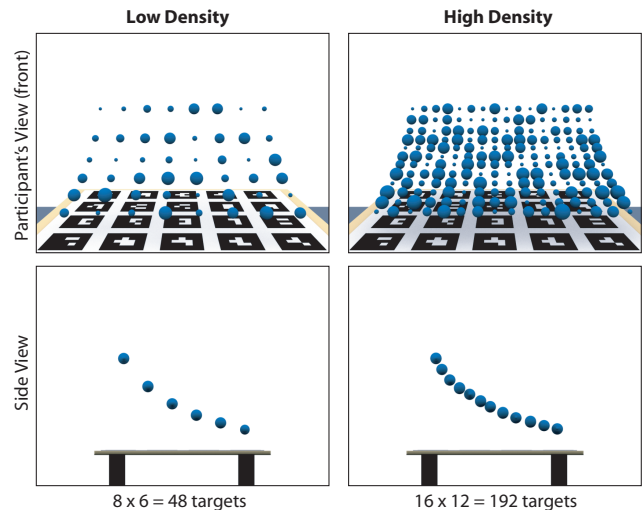


**Figure 3:** A participant performing the experiment using the Lens interface.

### 3.5 Design

The experiment followed a 3x2 repeated measures design. The independent variables were selection technique ST (direct touch, ray-casting, lens) and target density TD (low, high). Participants worked through six conditions comprising all combinations of selection technique and target density. Within each condition participants carried out a block of eighteen selection tasks following the procedure described in the previous section. The eighteen tasks were divided into six sets of three, where each set contained tasks from one of six difficulty levels. All the low density conditions were done first, in a counterbalanced fashion, followed by the high density conditions, also counterbalanced.

The difference between the low and high density conditions was the number of targets. In the low density condition the grid of targets was 8x6 giving a total of 48 targets. In the high density condition the grid was 16x12 giving a total of 192 targets, four times as many as low density. In both densities, the virtual scene occupied the same volume of approximately 0.5m x 0.3m x 0.5m. Therefore, in the high density condition, the targets were more closely packed than in the low condition. The arrangement of targets is shown in Figure 4.



**Figure 4:** The arrangements of targets in the low and high density conditions.

The dependent variables of interest were Task Completion Time, measured in seconds, and Error Rate, measured as the number of misses made when trying to select a target. Task Completion Time was the time interval beginning when the participant selected the

starting target and ending when they successfully selected the target object, regardless of how many times they missed. Participants were instructed to be as quick as possible but also to be accurate. Targets could be selected at any point on their surfaces – a hit near the edge counted equally to a hit in the center.

Other data recorded was movement of the participant’s hand and head in three-dimensions, relative to the room, at intervals of approximately 10ms. From this data, measures for average head and hand movement distance and velocity were computed.

### 3.5.1 Hypotheses

This experiment examines the performance differences between three different object selection techniques for augmented reality user interfaces. A technique’s performance is typically evaluated by selection time and accuracy. It is expected that there will be statistically significant differences in these measures between the three techniques.

The techniques are evaluated across two object density levels. Object density is a known factor in determining selection performance. Therefore, it is expected that the measures of selection time and error rate will increase as density increases. This being the case, the rate of increase is also of interest. A technique where performance degrades slowly as density increases is preferable.

### 3.6 Results

The sixteen participants each performed eighteen tasks in six conditions for a total of 108 trials per participant. Each trial yielded a time, miss count and 3D movement paths for the head and hand. Out of all trials, fifteen were removed because of invalid data caused by tracking failures, hardware faults (including drained batteries in the Wiimote) and participant discomfort (such as needing to adjust the HMD). The remaining data was summarised per condition. Times and miss counts were averaged, and movement paths were analysed to produce measures for average distance traveled and average velocities.

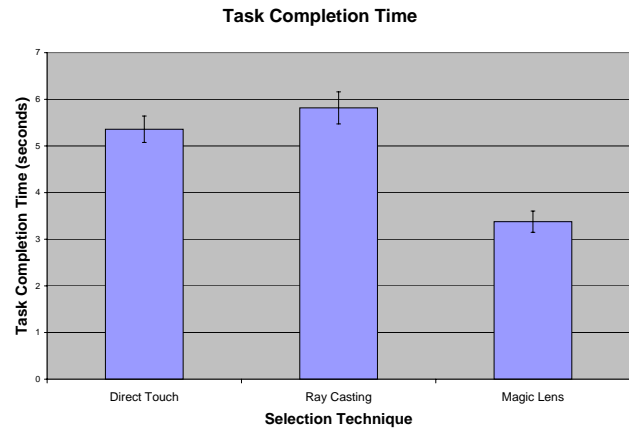
In the following sections, the results for Task Completion Time and Error Rate are presented first, followed by an analysis of the movement path data, and finally subjective feedback collected via questionnaires.

#### 3.6.1 Objective Measures

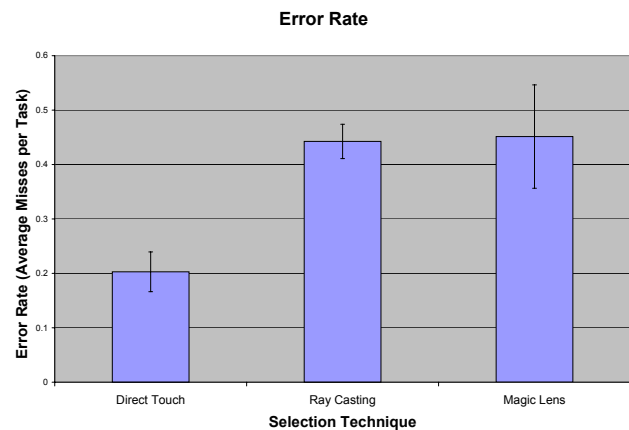
There was a significant main effect of Selection Technique on Task Completion Time ( $F_{2,30} = 64.0, p < .05$ ). There was no significant effect of Target Density on Task Completion Time. Therefore, the times for low and high densities were averaged to give the overall means for each Selection Technique.

Pairwise comparison (with Bonferroni correction) revealed significant differences between Lens and the other two techniques, with no significant difference between Direct Touch and Ray-Casting. The Lens technique performed faster than both other techniques, with a mean selection time of 3.38 seconds ( $sd = 0.911$ ). The results are listed in Table 1 and shown in Figure 5.

There was a significant main effect of Selection Technique on Error Rate ( $F_{2,30} = 7.53, p < .05$ ), but no significant effect from Target Density. Figure 6 shows the overall mean Error Rate for each condition. Pairwise comparison revealed significant differences between Direct Touch and the other two techniques, with no significant difference between Ray-Casting and Lens. Direct Touch had a significantly lower Error Rate, with a mean of 0.203 misses per trial ( $sd = 0.146$ ).



**Figure 5:** Task Completion Time for each Selection Techniques. Error bars show the standard error.



**Figure 6:** Error Rate for each Selection Technique. Error bars show the standard error.

Technique	Time (seconds)		Error Rate (misses)	
	mean	sd	mean	sd
DT	5.358	1.132	0.203	0.146
RC	5.816	1.376	0.442	0.126
ML	3.377	0.911	0.451	0.378

**Table 1:** Overall summary statistics for the dependent variables Task Completion Time and Error Rate.

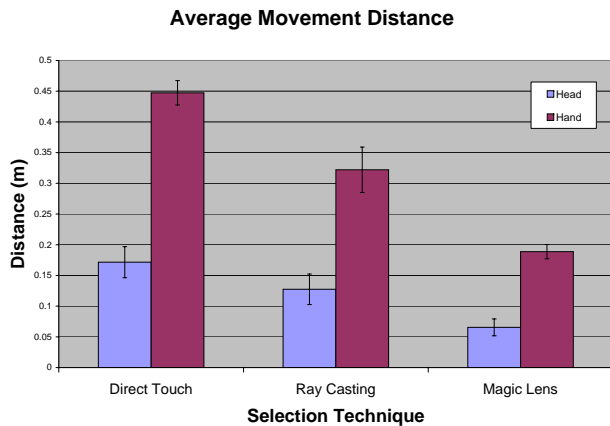
### 3.6.2 Analysis of Motion

Motion data was recorded for the head and hand of each participant. This data was processed to produce metrics of Average Head Distance, Average Hand Distance, Average Head Speed and Average Hand Speed for each condition. The results of this analysis are summarised in Table 2 and reported below.

Movement distance results are shown in Figure 7. The means for Average Head Distance were significantly different for all Selection Techniques ( $F_{2,30} = 21.12, p < .05$ ). Direct Touch had the longest head movement distance and Lens the shortest. Target Density had no significant effect on this metric. Selection Technique had a significant main effect on Average Hand Distance ( $F_{2,30} = 23.81, p < .05$ ). All three means were significantly different, with Direct Touch taking most movement (mean = 0.172m, sd = 0.101m), and Magic Lens the least (mean = 0.066m, sd = 0.055m). Again, Target Density was not significant.

Speed results are shown in Figure 8. Selection Technique had a significant main effect on Average Head Speed ( $F_{2,30} = 23.831, p < .05$ ). All three means were significantly different with the Lens technique exhibiting the least speed (mean =  $0.017\text{ms}^{-1}$ , sd =  $0.013\text{ms}^{-1}$ ) and Direct Touch the most (mean =  $0.033\text{ms}^{-1}$ , sd =  $0.020\text{ms}^{-1}$ ). Average Hand Speeds were significantly different ( $F_{2,30} = 56.81, p < .05$ ). Direct Touch was significantly higher than the other two techniques (mean =  $0.090\text{ms}^{-1}$ , sd =  $0.010\text{ms}^{-1}$ ).

Target Density did not have a significant effect on any of the motion measures. The measures were averaged across both densities and are shown in Table 2.



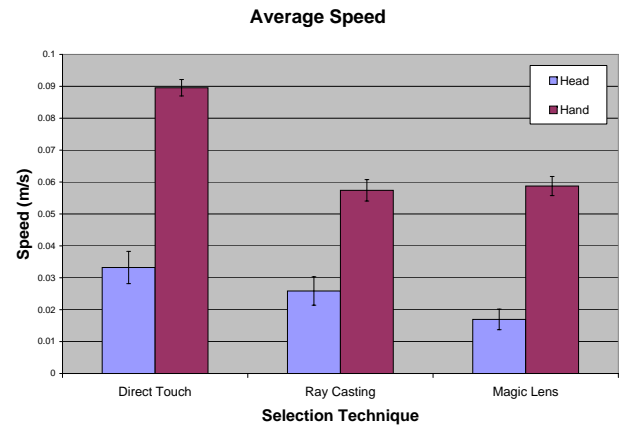
**Figure 7:** Movement distances for each Selection Technique. Error bars show the standard error.

### 3.6.3 Subjective Measures

The answers to the questionnaire are summarised in Table 3. We address the results from each question in detail below.

All techniques were well understood by the participants. When asked if they found the technique easy to understand, participants rated all three selection techniques higher than 6.2 out of 7. There was a significant main effect ( $F_{2,30} = 4.34, p < .05$ ), although we considered this to have negligible practical significance.

It was easiest to select targets with the Lens (mean = 6.125, sd = 0.793), followed by Direct Touch (mean = 4.031, sd = 1.062) and Ray-Casting (mean = 3.500, sd = 1.459), producing a significant



**Figure 8:** Movement speeds for each Selection Technique. Error bars show the standard error.

main effect ( $F_{2,30} = 47.4, p < .05$ ). Post-hoc analysis with Bonferroni correction ( $p < .05$ ) showed pairwise differences between Lens and the other two techniques, which were not significantly different from each other.

Participants felt they performed fastest with the Lens (mean = 6.000, sd = 0.803), followed by Direct Touch (mean = 4.063, sd = 1.366) and Ray-Casting (mean = 3.625, sd = 1.362), giving a significant main effect ( $F_{2,30} = 36.4, p < .05$ ). Post-hoc analysis with Bonferroni correction ( $p < .05$ ) showed the Lens to be rated significantly higher than the other two techniques, which were not significantly different from each other.

When asked if they would appreciate having a particular selection technique available when using AR interfaces, participants were most enthusiastic about the Lens (mean = 5.844, sd = 1.505), followed by Direct Touch (mean = 3.656, sd = 1.558) and finally Ray-Casting (mean = 3.344, sd = 1.658). These results gave a significant main effect ( $F_{2,30} = 43.5, p < .05$ ), and post-hoc analysis with Bonferroni correction ( $p < .05$ ) showed pairwise differences between the Lens and the other two techniques, which were not significantly different from each other.

The Lens technique was found least physically demanding (mean = 3.031, sd = 1.062), followed by Ray-Casting (mean = 4.156, sd = 1.322) and then Direct Touch (mean = 5.438, sd = 1.105), giving a significant main effect ( $F_{2,30} = 24.5, p < .05$ ). Post-hoc analysis with Bonferroni correction ( $p < .05$ ) showed pairwise differences between all techniques.

The Lens technique was found least mentally demanding (mean = 2.781, sd = 1.039), followed by Direct Touch (mean = 4.656, sd = 1.285) and Ray-Casting (mean = 5.031, sd = 1.282), giving a significant main effect ( $F_{2,30} = 25.2, p < .05$ ). Post-hoc analysis with Bonferroni correction ( $p < .05$ ) showed the Lens to be rated significantly less demanding than the other two techniques, which were not significantly different from each other.

The Lens technique was reported as causing the least frustration (mean = 2.469, sd = 1.367), followed by Direct Touch (mean = 4.250, sd = 1.391) and Ray-Casting (mean = 4.813, sd = 1.281). These results gave a significant main effect ( $F_{2,30} = 34.9, p < .05$ ). Post-hoc analysis with Bonferroni correction ( $p < .05$ ) showed the Lens to be rated significantly lower than the other two techniques, which were not significantly different from each other.

The Lens technique was the significantly preferred technique in

Technique	Head Movement				Hand Movement			
	Distance (m)		Speed (ms <sup>-1</sup> )		Distance (m)		Speed (ms <sup>-1</sup> )	
	mean	sd	mean	sd	mean	sd	mean	sd
DT	0.172	0.101	0.033	0.020	0.447	0.079	0.090	0.010
RC	0.127	0.099	0.026	0.018	0.322	0.148	0.057	0.013
ML	0.066	0.055	0.017	0.013	0.189	0.046	0.059	0.012

**Table 2:** Summary of Motion Variables.

Low Density, High Density and Overall (Friedman Test  $\chi_r^2 = 24.1$ ,  $df=2$ ,  $N=16$ ,  $p < 0.05$ ). In each case, the next preference was Direct Touch, followed by Ray-Casting.

## 4 Discussion

The Lens technique was significantly faster than both Direct Touch and Ray-Casting, verifying the hypothesis that there would be a difference across techniques. It also required the least head and hand movement, followed next by Ray-Casting and then Direct Touch. These results are likely due to the fact that the experiment did not use stereoscopic AR. The lack of the stereo depth cues appears to have had the least effect on the Lens because it is based on an Image Plane technique, which reduces 3D selection to a 2D task anyway. In contrast, Direct Touch and Ray-Casting suffered without stereo. The increased head and hand movements in these conditions may suggest that participants moved more to exploit other depth cues such as monocular movement parallax and overlapping.

On average, participants moved their head more slowly in the Lens Technique, and moved their hand more quickly with Direct Touch. We believe the lower movement speeds and distances traveled for the Lens indicate that it requires less physical effort. This hypothesis is supported by the subjective results, which strongly indicate that the Lens technique is the least physically demanding, the least mentally demanding, and overall the least frustrating of the three techniques. There is the possibility that the subjective results unfairly favoured the Lens because participants were aware it was the new technique being evaluated.

Direct Touch had significantly fewest errors. The Lens and Ray-Casting had more errors, but were not significantly different from each other. One source of errors was tracking noise, which had a greater effect of the two ray-based techniques because jitter at the ray's origin can greatly affect targeting. Also, pulling the controller's trigger to make a selection often moved the controller enough to cause a miss - a situation that participants found most frustrating.

Interestingly, target density had no significant effect on selection time, error rate, movement speed or movement distance. This is counter to the hypothesis that performance would degrade as density increased. This may suggest that the choices for density levels were not extreme enough, or that there is some other factor at work that overshadows the differences between densities.

## 5 Future Work

There are several factors that could be considered in future evaluations of lens selection techniques. The use of stereoscopic cameras and displays will make many selection techniques more viable. Many participants complained about the lack of depth perception hindering their performance, especially in the Direct Touch conditions. It could be argued that the lack of stereo vision in this experiment biased it towards the Lens technique, however this argument only highlights the fact that since most current tabletop AR

setups are not stereoscopic, selection and manipulation techniques that rely on reasonable depth perception should be avoided. Other modalities could be employed to compensate for limited depth perception. For example, vibro-tactile feedback in the controller, or auditory feedback, could be used to indicate proximity to a target.

In this experiment, the chance of targets occluding each other was intentionally minimised. This is obviously not the typical scenario in many 3D virtual environments and future work will need to address how these techniques cope under increased object density and occlusion.

This experiment also only tested single object selection. Multiple object selection is often required, and is a cause of frustration if the user is forced to make many single selections rather than using an appropriate multiple object selection technique. In terms of virtual lens interfaces, this could be a reason to investigate Aperture selection rather than the Image Plane technique used in this experiment, as Aperture would permit the selection of a volume containing many objects.

## 6 Conclusion

This paper reports on a formal evaluation of selection techniques for table-top augmented reality user interfaces. Object selection is a crucial task in user interfaces because it precedes almost all other tasks the user might wish to perform. The three selection techniques compared in this study were Direct Touch, Ray-Casting, and an Image Plane technique intended for use with virtual Magic Lenses, which have been explored for visualisation purposes in other research.

The Lens technique was found to be faster than the other techniques and although it was not the most accurate, it required less head and hand movement, and lower head movement speed. This suggests that it is a more enjoyable technique, which is supported by subjective results. These would suggest that the Lens technique was least physically and mentally demanding, least frustrating, and preferred in all cases.

Contrary to our expectations, Target Density was found to have no significant effect on any of the measures analysed in this experiment. Further evaluation is required to determine the cause. It is possible that the density levels were simply not different enough. Another possibility is that an unaccounted for factor in the tabletop AR setup hinders selection performance by some large factor (such as a large preparation cost), and that this outweighs the effect of increased density.

Overall, this experiment successfully showed the performance benefits and preference for the lens-based selection technique over traditional techniques. Magic Lenses have previously been shown to have intriguing visualisation applications in augmented reality. These applications can now be enhanced with efficient object selection capabilities to provide users with more powerful analysis and visualisation tools.

Question	DT		RC		ML	
	mean	sd	mean	sd	mean	sd
1. I found the selection technique easy to understand	6.375	0.492	6.219	0.659	6.500	0.568
2. I found it easy to select the target	4.031	1.062	3.500	1.459	6.125	0.793
3. I feel that I performed quickly with this technique	4.063	1.366	3.625	1.362	6.000	0.803
4. If I had to use AR interfaces like this regularly, this is a technique I would appreciate having available	3.656	1.558	3.344	1.658	5.844	1.505
5. I found using this technique physically demanding	5.438	1.105	4.156	1.322	3.031	1.062
6. I found using this technique mentally demanding	4.656	1.285	5.031	1.282	2.781	1.039
7. I found this technique frustrating	4.250	1.391	4.813	1.281	2.469	1.367

**Table 3:** Summary of questionnaire responses. Questions were posed on a seven-point Likert scale between 1 = Disagree and 7 = Agree.

## References

- ADVANCED REALTIME TRACKING GMBH, 2007. ART website: <http://www.ar-tracking.de>.
- AZUMA, R. T. 1997. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (August), 355–385.
- BIER, E. A., STONE, M. C., PIER, K., BUXTON, W., AND DEROSE, T. D. 1993. Toolglass and magiclenses: The see-through interface. In *Proceedings of Siggraph93, Computer Graphics Annual Conference Series*, 73–80.
- BOECK, J. D., WEYER, T. D., RAYMAEKERS, C., AND CONINX, K. 2006. Using the non-dominant hand for selection in 3d. In *3DUI '06: Proceedings of the 3D User Interfaces (3DUI'06)*, 53–58.
- BOWMAN, D., JOHNSON, D., AND HODGES, L. 2001. Testbed evaluation of virtual environment interaction techniques.
- BOWMAN, D. A., KRUIJFF, E., JNR, J. J. L., AND POUPYREV, I. 2004. *3D User Interfaces Theory and Practice*. Addison-Wesley.
- FITTS, P. 1954. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. 381–391.
- FORBES, K., 2007. cWiiMote website: <http://simulatedcomicproduct.com/2006/12/cwiimote-02.php>.
- FORSBERG, A., HERNDON, K., AND ZELEZNIK, R. 1996. Aperture based selection for immersive virtual environments. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, 95–96.
- GROSSMAN, T., AND BALAKRISHNAN, R. 2006. The design and evaluation of selection techniques for 3d volumetric displays. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, 3–12.
- ISHII, H., AND ULLMER, B. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press, 234–241.
- KATO, H., BILLINGHURST, M., POUPYREV, I., AND TETSUTANI, N. 2001. Tangible augmented reality for human computer interaction. In *Proceedings of Nicograph*.
- LEE, G. A., BILLINGHURST, M., AND KIM, G. J. 2004. Occlusion based interaction methods for tangible augmented reality environments. In *VRCAI '04: Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry*, 419–426.
- LIANG, J., AND GREEN, M. 1994. JDCAD: A highly interactive 3d modeling system. *Computers and Graphics* 18, 4, 499–506.
- LOOSER, J., BILLINGHURST, M., AND COCKBURN, A. 2004. Through the looking glass: the use of lenses as an interface tool for augmented reality interfaces. In *GRAPHITE '04: Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, 204–211.
- LOOSER, J., GRASSET, R., SEICHTER, H., AND BILLINGHURST, M. 2006. OSGART - a pragmatic approach to MR. In *In Industrial Workshop at ISMAR 2006*.
- OSFIELD, R., 2007. Open scene graph website: <http://www.openscenegraph.org>.
- PIEKARSKI, W. 2004. *Interactive 3D Modelling in Outdoor Augmented Reality Worlds*. PhD thesis, School of Computer and Information Science, University of South Australia.
- PIERCE, J. S., FORSBERG, A. S., CONWAY, M. J., HONG, S., ZELEZNIK, R. C., AND MINE, M. R. 1997. Image plane interaction techniques in 3d immersive environments. In *SI3D '97: Proceedings of the 1997 symposium on Interactive 3D graphics*, 39–ff.
- POUPYREV, I., BILLINGHURST, M., WEGHORST, S., AND ICHIKAWA, T. 1996. The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, 79–80.
- POUPYREV, I., ICHIKAWA, T., WEGHORST, S., AND BILLINGHURST, M. 1998. Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques. *Computer Graphics Forum, EUROGRAPHICS 98* 17, 3, 41–52.
- STEED, A., AND PARKER, C. 2004. 3d selection strategies for head tracked and non-head tracked operation of spatially immersive displays. In *8th International Immersive Projection Technology Workshop*.
- STOAKLEY, R., CONWAY, M. J., AND PAUSCH, R. 1995. Virtual reality on a WIM: Interactive worlds in miniature. In *Proceedings of CHI'95 Conference on Human Factors in Computing Systems*, 265–272.