

Balloon Selection: A Multi-Finger Technique for Accurate Low-Fatigue 3D Selection

Hrvoje Benko*
Department of Computer Science
Columbia University

Steven Feiner†
Department of Computer Science
Columbia University

ABSTRACT

Balloon Selection is a 3D interaction technique that is modeled after the real world metaphor of manipulating a helium balloon attached to a string. Balloon Selection allows for precise 3D selection in the volume above a tabletop surface by using multiple fingers on a multi-touch-sensitive surface. The 3DOF selection task is decomposed in part into a 2DOF positioning task performed by one finger on the tabletop in an absolute 2D Cartesian coordinate system and a 1DOF positioning task performed by another finger on the tabletop in a relative 2D polar coordinate system. We have evaluated Balloon Selection in a formal user study that compared it to two well-known interaction techniques for selecting a static 3D target: a 3DOF tracked wand and keyboard cursor keys. We found that Balloon Selection was significantly faster than using cursor keys and had a significantly lower error rate than the wand. The lower error rate appeared to result from the user's hands being supported by the tabletop surface, resulting in significantly reduced hand tremor and arm fatigue.

Keywords: 3D interaction, selection technique, augmented reality, virtual reality, world-in-miniature, multi-touch interaction, cross-dimensional interaction, hybrid user interface.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Input devices and strategies, Interaction Styles*; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—*Artificial, augmented, and virtual realities*; I.3.6 [Computer Graphics]: Methodology and Techniques—*Interaction techniques*.

1 INTRODUCTION

In their survey paper, Hinckley et al. [23] identify many of the challenges associated with designing 3D interactions, dividing the design space into two broad categories: those that deal with human perception and those that deal with ergonomics. The authors suggest that spatial references (props or the user's own body), two-handed interaction, multi-sensory feedback, and physical constraints are invaluable for helping users to perceive and interact with a 3D object and environment. They also strongly recommend the reduction of extraneous degrees-of-freedom (DOF) to simplify the 3D task when possible. For example, in tasks requiring translation, but not rotation, rotation should be disabled. Furthermore, they suggest that providing a clear control metaphor (e.g., *eyeball-in-hand* [2, 9] camera or *ray-casting*) significantly

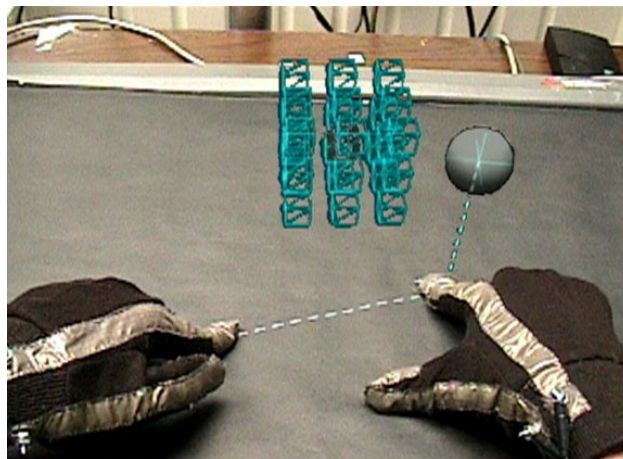


Figure 1: Balloon Selection in action. 3D spherical cursor above the surface is controlled by multi-finger interaction on the touch-sensitive surface. This image is taken through a video-mixed augmented reality system on a black paper-covered projection surface.

improves the effectiveness of the interface and enhances the ability of the user to perceive the task at hand.

In this paper, we present a novel interaction technique, called *Balloon Selection* (Figure 1), which follows these guidelines and allows for precise and accurate 3D selections in a constrained within-reach 3D environment. We are particularly interested in scenarios in which the user wishes to interact with a very small scale 3D environment, such as a model of the city, university campus, or an archaeological dig site [4, 25], seen from an outside-in (exocentric) view. Such environments, also known as 3D maps or world-in-miniature (WIM) models [37], are particularly well suited for collaborative scenarios in which multiple users discuss and analyze a given model over an augmented tabletop (e.g., [4, 35]). Unfortunately, objects in small scale WIMs tend to be very hard to select, due to their small size, closeness, and potential overlap (Figure 2).

We designed Balloon Selection to fully exploit the benefits of a firm touch-sensitive surface to assist the user in accurate 3D selections. First, the surface is able to provide passive haptic support for the user's hands, which would otherwise need to be held continuously in mid air. By having the user rest their hands on the surface, we improved selection accuracy by significantly reducing hand tremor and arm fatigue. Second, when the user's hands are on the surface, they tend to remain below the 3D model and thus do not obscure the 3D objects being manipulated. Third, we took advantage of the multi-touch surface by designing a technique that decomposes the 3DOF selection task into a 2DOF positioning task performed by one finger on the tabletop in an absolute 2D Cartesian coordinate system and a 1DOF positioning task performed by another finger on the tabletop in a relative 2D polar coordinate system.

* e-mail: benko@cs.columbia.edu

† e-mail: feiner@cs.columbia.edu

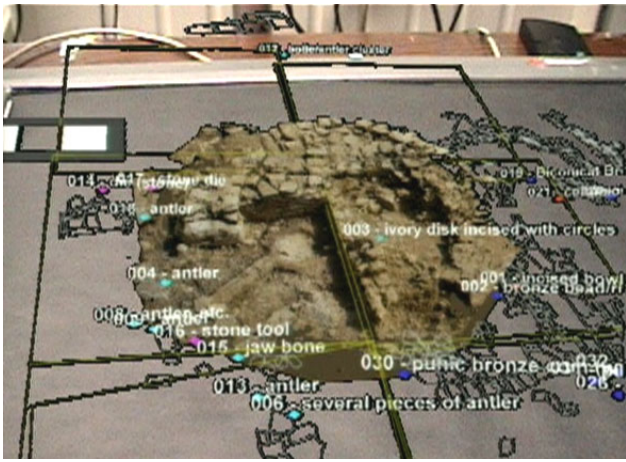


Figure 2: A world-in-miniature model of an archaeological excavation site at Monte Polizzo, Sicily, annotated with representations of excavated finds. Note that the objects representing the excavated finds tend to be extremely small and grouped together, making their selection difficult at this scale.

In the remainder of this paper, we first provide an overview of related work, followed by a detailed description of our technique. We then present a user study that compares our technique with two well-known interaction techniques for selecting a static 3D target: a 3DOF tracked wand (chosen for its intuitive and direct mapping to the task) and keyboard cursor keys (chosen for accuracy and insensitivity to hand tremor and arm fatigue). We conclude with a discussion of our results and some potential future applications of our work.

2 RELATED WORK

Research on selection techniques in 3D immersive environments can be divided into two broad categories. First, the ray-based selection techniques, and their aperture-based superset, use a projected ray, either visible or not, from the user’s hand or head to select targets that the ray intersects [6, 16, 29, 31, 38]. Second, virtual hand techniques, or 3D cursors, provide a direct or offset mapping between the hand position and the location of a 3D cursor [17, 29, 34, 38]. For example, thirty-five years ago, Vickers [38] described the use of a hand-held 3D position-tracked wand that could be used to select vertices within a small virtual cube at its tip (3D cursor selection); buttons on the wand allowed the selection geometry to move to an otherwise out-of-reach vertex along the vector defined by the right eye and the tip of the wand (ray-based selection).

While ray-based selection has been shown to achieve faster selection times than the use of a 3D cursor for general virtual reality (VR) tasks [7], it is not as well suited to selecting targets that are small, closely grouped together, or partially or completely occluded, such as the ones found in many WIM models. It is possible to enhance the WIM metaphor by showing just a relevant subset of the environment [41] and then scrolling or scaling through it; however, this technique tends to lose the context provided by the complete WIM, which is one of the benefits of using a scaled model of the environment.

Several researchers have tried to address the problems posed by nearby and occluding objects in ray-based selection. Hinckley et al. [23] and Bowman and Hodges [8] discuss methods for choosing one of multiple selections along a ray. Work by Olwal et al. on *SenseShapes* [31] uses multimodal input and the history of interaction with the selection geometry to disambiguate among potential selections, while Wyss et al. [42] present a two-handed technique based on the intersection of two rays. Olwal and Feiner

[32] describe a flexible pointer that allows the user to bend the selection ray to avoid other objects. Recently, the use of flexible and adjustable rays to select 2D widgets in 3D environments has been further explored by Andujar and Argelaguet [1] and de Haan et al. [20].

Improvements to 3D cursor techniques usually focus on two categories: extending the user’s reach and making the user’s virtual hand or 3D cursor more accurate. Poupayev et al. developed *GO-GO* [34], which nonlinearly maps the distance between the user’s body and their real hand to the distance that the virtual hand moves from the user. Pierce et al. developed the *Voodoo Dolls* technique [33], which brings scaled copies of distant objects within reach for fine manipulation. Frees and Kessler [17] adaptively adjust the control-display ratios of a 3D cursor to assist in precise position and rotation interactions. Herndon et al., in their work on *Interactive Shadows* [22], manipulate 3D objects by means of their 2D projections on principle planes, inspiring our manipulation of the balloon on the projection surface.

Bimanual interaction has been explored in both 2D and 3D contexts. In their pioneering work, Buxton and Myers [11] demonstrated that users tend to parallelize interaction tasks between hands, thus gaining significant performance improvements. Guiard highlighted the asymmetric roles of human hands in many bimanual tasks [19]. Cutler et al. [13] describe bimanual interactions for manipulating both the user’s perspective and 3D models. Mine et al. employ two-handed interaction for travel [29], while Hinckley et al. [24] argue for use of tangible props held with a non-dominant hand and manipulated with a dominant hand. Grossman et al. [18] evaluated several two-handed and multi-finger interactions for 3D volumetric displays.

The recent emergence of many multi-touch–sensing tabletop surfaces [14, 21, 36, 39, 40] has inspired numerous techniques that explore the benefits of multi-touch interactions. Moscovich et al. [30] demonstrate the use of multi-touch methods to interactively animate 2D figures. Benko et al. [5] use multiple fingers to adaptively scale cursor movement and increase precision in a cursor selection task. In many ways, our Balloon Selection technique follows the ideas of Benko et al. [3], whose *Cross-Dimensional Gestures* use the multi-touch surface in conjunction with a 3D tracked glove to enable seamless transitions between displays and dimensions in a hybrid augmented environment.

3 BALLOON SELECTION TECHNIQUE

Balloon Selection was inspired by how people play with a helium balloon on a string. A tethered helium balloon floats straight up from the point where its string is being held. If one holds the string tightly with one hand, with the string passing loosely through the fingers of the other hand and moves the hands relative to each other, the balloon will change in height along a vector passing perpendicularly through the hand closest to the balloon (Figure 3). This simple, yet powerful, metaphor demonstrates one way in which a 3DOF positioning task can be effectively decomposed into two separable tasks [26]: a 2DOF positioning task in

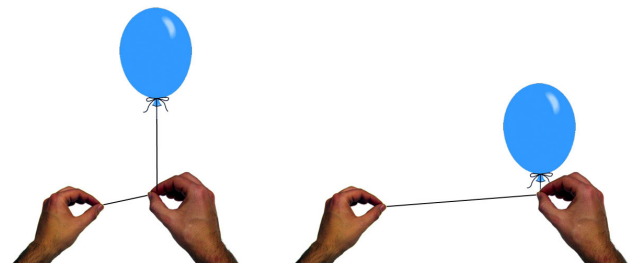


Figure 3: An illustration of the basic principle of Balloon Selection.

the ground plane and a 1DOF “string-pulling” task for regulating the balloon height. This basic metaphor has the interesting property that the user’s hands only need to move in a single plane to control the 3D location of the balloon. This inspired us to use a multi-touch surface to control and select 3D objects above the surface.

While we consider Balloon Selection to be a 3D cursor technique, it is also possible to analyze it as a modified ray-casting technique in which the ray always originates on the surface, points in a fixed direction (vertical), and has variable finite length (string length). Not being able to change the (always vertical) angle of the ray restricts the volume in which Balloon Selection can operate; however, this constraint eliminates some of the hand tremor difficulties associated with ray-casting techniques where small angular movement at the origin results in large physical movement at a distance.

3.1 Interaction Implementation

We implemented Balloon Selection in a hybrid multi-display augmented reality (AR) environment consisting of a tracked head-worn display and a multi-touch-sensitive projected tabletop display. A pair of instrumented gloves allows the locations on the tabletop of four fingers (the index fingers and thumbs) to be individually tracked. (A detailed description of our environment can be found in Section 4.2.) The physical balloon is replaced by a virtual balloon, consisting of three-axis crosshairs within a semi-transparent sphere. The physical string is replaced by a pair of dashed virtual lines. In our environment, a horizontal line is displayed on the projected tabletop, while the cursor and a vertical line are displayed in the head-worn display. (Alternatively, both lines could have been displayed in the head-worn display.)

To instantiate the balloon, the user touches the tabletop with one sensed finger (*anchor*) and places a second sensed finger (*stretching finger*) immediately adjacent to it. This dual-and-

adjacent finger state triggers the creation of a 3D balloon sphere that now resides on the surface at the location of the anchor (Figure 4a). It also allows for a simple yet powerful way to disambiguate between regular (single touch) pointing and invocation of the Balloon Selection technique.

By moving the stretching finger away from the anchor, the user “stretches” the virtual string between the fingers (Figure 4b). When the user reverses direction, the string’s length becomes fixed and the balloon rises from the surface (Figure 4c).

At this point, the user’s anchor finger controls the position of the balloon in the horizontal plane (2DOF) while the stretching finger controls its elevation from the surface (1DOF). It is important to note that only the relative distance between the anchor and the stretching finger matters, and not the absolute location of the stretching finger on the surface, thus making the specification of the height a pure 1DOF task. Moving the two fingers in parallel translates the balloon in a horizontal plane above the surface. Maintaining a constant offset between the fingers can be challenging, so we have designed some variants that make high precision possible (see Section 3.3).

Additional modifications of the balloon are possible by employing a separately sensed third finger. The user can adjust the size of the balloon by changing the distance between the third finger and the anchor, which is used as a scaling factor: moving the fingers apart scales up the balloon, while moving them closer together scales it down (Figure 4d–e). To perform a selection, the user again uses the dual-and-adjacent finger method, this time between the stretching finger and a third finger, making the equivalent of a “clicking” action (Figure 4f).

3.2 Visual and Audio Feedback

In initial testing, we discovered that users had difficulty understanding the string stretching state, since it did not map well to a physical metaphor. Therefore, we decided to add visual and audio

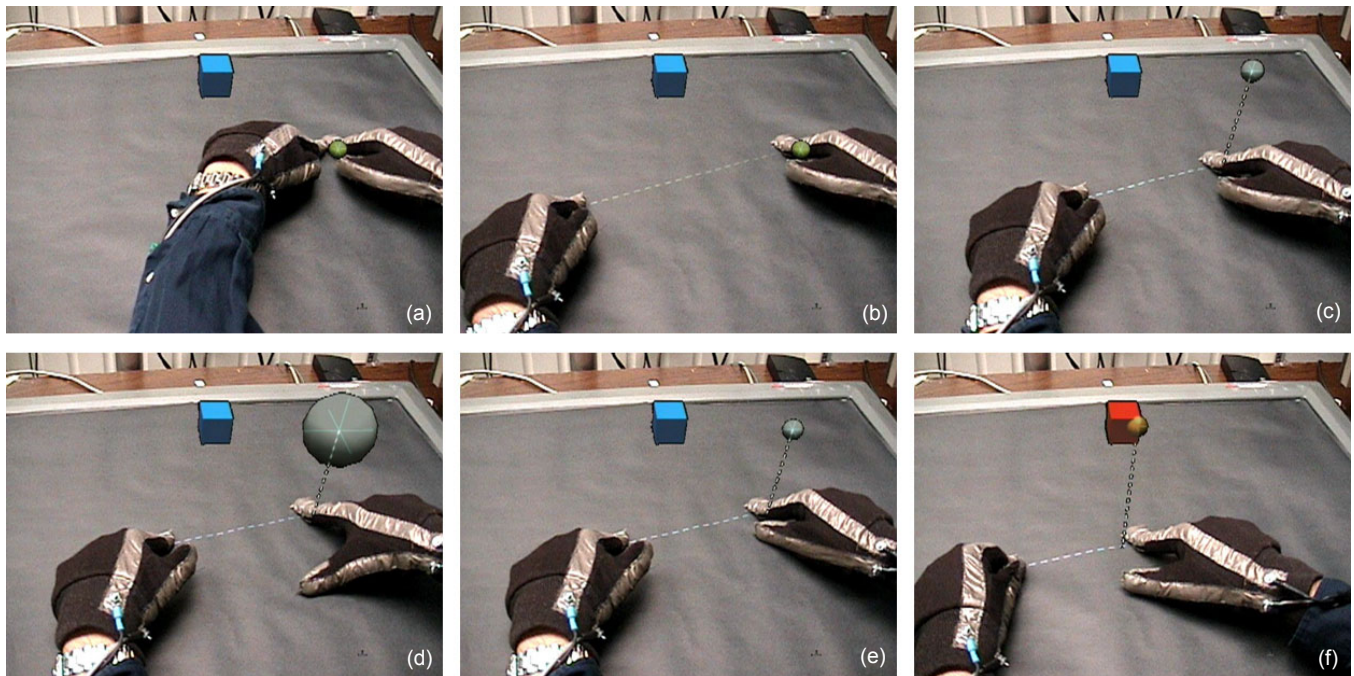


Figure 4: Frames from a Balloon Selection interaction sequence (captured through a video-mixed display). (a) Placing two fingers (the anchor and the stretching finger) adjacent to each other on the tabletop instantiates the 3D cursor (balloon). (b) Moving the stretching finger away from the anchor stretches the virtual string between the two fingers. (c) Moving the fingers closer together raises the balloon from the surface. (d–e) Moving the thumb on the anchor’s hand towards or away from the anchor scales the balloon up (for easier) or down (for more precise) selection. (f) Moving the anchor on the surface translates the balloon parallel to the plane of the table, while varying the distance between the anchor and the stretching finger determines the balloon’s height. Placing the thumb of the stretching finger’s hand adjacent to the stretching finger triggers the selection of the target cube.



Figure 5: Interaction devices used in our experiment (top to bottom): A 3DOF tracked three button wand (Logitech TrackMan Live), a pair of specially constructed gloves with two active fingers on each hand, and the keyboard. The devices are laid out on the DiamondTouch surface.

cues to attempt to guide the user through successful use of Balloon Selection. Currently, both the string and the balloon change color when the string is being stretched (orange during stretching, white while in use). In addition, when stretching, the user hears a rapid sequence of ratcheting clicks to enhance the physicality of the string/chain that holds the balloon. Interestingly, despite the mixed metaphors, our test users often highlighted this feature as the key aspect in making the balloon-on-a-string metaphor believable.

Additionally, we change the cursor and string color and add a different single clicking sound during selection to provide clear feedback for the triggering action.

3.3 Multi-Finger Considerations

While our technique requires the use of multiple fingers, the fingers are not assigned a predefined role. Rather, we assign the finger role based strictly on the *order* in which the fingers touch the surface (following the approach outlined in Benko et al. [5], and inspired by the complementary notion of determining handedness from the spatial relationship of two tracked devices described by Kurtenbach et al. [28]). Thus, rather than requiring the particular assignment of fingers shown in Figure 4, we generalized our technique by classifying the first contact with a tabletop surface as the *primary finger (anchor)*, the second contact as the *secondary finger (stretching finger)*, and the third contact as the *tertiary finger (scaling finger or selection finger)*. Therefore, our technique does not depend on the tabletop surface identifying the specific fingers or hands that are currently touching the surface, even though the glove-based sensing technology that we use can provide that information.

This *order-based* approach has several significant benefits. First, it makes our technique more widely applicable, given that most touch screens cannot identify individual touches as belonging to a particular user or hand. Second, both left-handed and right-handed users have no problems adopting their preferred finger usage, since Balloon Selection allows either hand to be the dominant positioning hand. Additionally, allowing the user to choose which finger/hand to use for the particular role ensures that the entire table surface is easily reachable. For example, it is possible to control the position of the cursor with the left hand when close to the left border, to avoid crossing the hands. Third, this approach has an added benefit, in that Balloon Selection can

be used as both a two-handed and a single-handed technique. (We note here that our current glove design supports only two touches per glove, necessitating the use of two gloves to trigger the selection.)

3.4 String Height Clutching

The requirement of always needing to hold at least two fingers on the surface can be relaxed once the balloon has been created. In fact, by lifting the secondary finger off the surface, the user can “declutch,” temporarily fixing the height of the balloon. Repositioning the secondary finger on the surface automatically reacquires the string, with the height of the balloon remaining constant, while the horizontal portion of the string is lengthened or shortened to fit between the primary and secondary fingers.

Clutching has two major benefits. First, when further height adjustment is not desired, the user can eliminate this DOF (and potential effects of additional hand tremor) by lifting the secondary finger off the surface. Second, by clutching, the string can be extended beyond the maximum length allowed by the surface diagonal. This makes it possible to extend the balloon’s reach far above the surface and for the total length of the balloon string to be lengthened or shortened as desired.

While string height clutching and declutching might not be necessary in a two-handed approach in which the distance between the primary and secondary fingers can be sufficiently great, we believe that it would be very beneficial in single-handed operation where the physical distance between the fingers of a single hand is a serious limiting factor. Our system supports string height clutching and declutching; however, we disabled it for our user experiment to minimize learning time.

4 USER EVALUATION

We performed a user study to determine how Balloon Selection compared with other selection techniques. Twelve paid participants (10 male and 2 female), ages 20–27, were recruited by mass email to graduate and undergraduate Computer Science students at our university. All subjects were frequent computer users, but had little previous experience with VR or AR techniques or technology. Two participants reported that they actively play 3D computer games (several hours per week). All subjects identified themselves as right-handed. Four participants wore corrective contact lenses, while the others reported no vision impairment. While all subjects reported using touch-screens regularly (mostly ATM machines), none had experience with multi-touch screens.

4.1 Baseline Comparison Techniques

To evaluate the performance of Balloon Selection (BALLOON), we decided to compare it to two existing 3D selection techniques for selecting static targets: a direct interaction technique using a 3DOF tracked wireless wand (WAND), and a keyboard technique (KEYBOARD). Each of these baseline techniques is known for having certain benefits with which we wanted to compare Balloon Selection. WAND has the advantage of being relatively fast and using a direct, absolute interaction device that relies on well trained 3D human motor skills. In contrast, KEYBOARD would be expected to offer the most accuracy, eliminating the effect of hand tremor and arm fatigue. All three techniques offered the user the same capabilities: 3DOF positioning of a spherical cursor, 1DOF cursor size (radius) control, and selection triggering. For each technique, all scene objects intersecting the cursor were selected when the trigger was fired.

4.1.1 WAND Technique

WAND was implemented with a wireless Logitech TrackMan Live device whose trackball was removed (Figure 5). The 3D

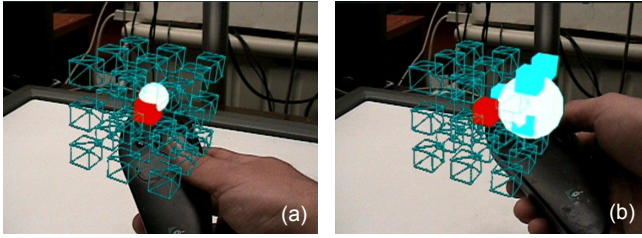


Figure 6: The “cube-of-cubes” target setup being used with a WAND technique (3D cursor is a semi-transparent white sphere). (a) The successful selection of a goal target (red middle target). (b) The unsuccessful selection of more than one target by a larger 3D cursor.

cursor tracked the position of the wand in space, which was accomplished by using an Origin Instruments DynaSight optical tracker to track a passive retroreflective target affixed to the wand. Cursor size was controlled by pressing the buttons on the wand (the left and middle buttons made the cursor incrementally smaller and larger, respectively), while selection was performed by pressing on the right button.

4.1.2 KEYBOARD Technique

KEYBOARD used the number pad on a standard keyboard to provide cursor position and size control. Keys 4 and 6 moved the cursor left and right, keys 8 and 2 moved the cursor out and in, and keys 7 and 1 moved the cursor up and down. The participant could adjust the size of the cursor by pressing 3 (smaller) and 9 (larger), and select by pressing the *space* key. Positioning was done in the absolute world-coordinate system and was therefore consistent with WAND and BALLOON. Movement was performed at a fixed speed of 4 cm / second when a key was depressed, which we found to offer a good balance between speed and accuracy in our preliminary pilot experiments. Our implementation supported moving in more than one dimension simultaneously; however, we observed that participants rarely took advantage of this.

4.2 Experimental Setup

The experiment was performed on a PC running Windows XP Pro, with two NVIDIA Quadro FX 4500 graphics cards. One display output was connected to a Proxima Ultralight x350 DLP projector (1024 × 768 resolution), which projected onto a Mitsubishi Electric Research Laboratory DiamondTouch table [14], a multi-user, multi-touch interaction surface. The other display output was connected to a Sony LDI-D100B stereo, see-through, head-worn display (800 × 600 resolution). The projected and head-worn displays are driven by separate application modules, both developed in Java and Java3D, which communicate through the Adaptive Agent Architecture [27] to facilitate easy connection to, discovery of, and communication among the modules.

Head tracking was performed by a ceiling-mounted InterSense IS900 6DOF tracker, while the wand’s position was tracked by a Origin Instruments DynaSight optical tracker mounted 60 cm from the surface. This setup was deliberately chosen to maximize the measurement accuracy of the wand and therefore not provide a bias against this technique. The DynaSight documentation reports that the optimal measurement resolution of the DynaSight tracker at about 40 cm is 0.05 mm in three axes. While we did not attempt to verify this numeric claim, we did determine that the rendered cursor did not move when the wand was held stationary. This sub-pixel resolution was deemed more than adequate given that the smallest target in our test was 4 mm wide and projected to multiple pixels. (Note that because we were selecting virtual,

rather than real targets, we were concerned more with resolution than absolute accuracy.)

The DiamondTouch surface reports an interpolated touch resolution of 2032 × 1520 touch pixels. The physical dimensions of the DiamondTouch surface are 64 cm × 48 cm, which translates into a measurement resolution of about 0.3 mm in each dimension. Again, this was deemed more than satisfactory for our targets of 4 mm width and up. To receive individual reports for each of the user’s two thumbs and index fingers on the DiamondTouch surface, we created gloves with conductive fabric (Shieldex Bremen), inspired by those developed by Butler et al. for their HabisDraw DT system [10]. As shown in Figure 5, a separate strip of conductive fabric is run along each thumb and index finger, and each strip is connected by a cable to a separate DiamondTouch input.

4.3 Task

Participants were asked to perform a target acquisition task. The experiment environment consisted of a cubic array of 3 × 3 × 3 wireframe cubes, of which only the center (red) cube was the actual goal target and the other 26 (cyan) cubes were distracter targets (Figure 6). Cubes were rendered solid when they intersected the balloon. To successfully complete the trial, participants needed to select just the goal target, without selecting any of the distracter targets. Simple audio cues were used to signal to the participant whether or not their selection was successful. The sounds were chosen from the standard Windows XP sounds, with “ding.wav” signaling success and “chord.wav” signaling failure. This was done to ensure that the participants remained motivated to perform correct selections as the trials continued.

4.4 Procedure

A within-subjects, repeated measures design was used consisting of three techniques (WAND, BALLOON, and KEYBOARD) and four target sizes (cubes with edge lengths of 10 mm, 8 mm, 6 mm, and 4 mm). The experiment lasted approximately 45 minutes, and was divided into three blocks, with short breaks between blocks. Each block consisted of all trials for one of the three techniques, and the order in which the techniques were presented was counterbalanced across participants. Each block began with a short practice period in which the participant was taught a technique and given a set of practice trials in which to learn and experiment with that technique.

When the participant felt comfortable with a technique, they proceeded to complete the test of 40 target selections. The targets were presented in order of decreasing size, where the target size and the distance between targets were reduced by 2 mm in each dimension after each set of ten selections. The position and orientation of targets was randomly changed for every trial, but all targets were positioned between 10 to 30 cm above the tabletop, and never more than 10 cm apart. This design resulted in 120 different recorded trials per participant.

4.5 Hypotheses

Prior to our experiment, we postulated the following four hypotheses:

(H1) WAND would be the fastest technique because of its use of a direct interaction device for integrated, continuous 3D position control in an absolute 3D coordinate system, modeled on the familiar task of directly touching a 3D point.

(H2) KEYBOARD would be the slowest technique because of its use of keys to separately control velocity along each of the three axes.

(H3) BALLOON would not be as affected by hand tremor and arm fatigue as WAND, and therefore have a lower error rate than WAND.

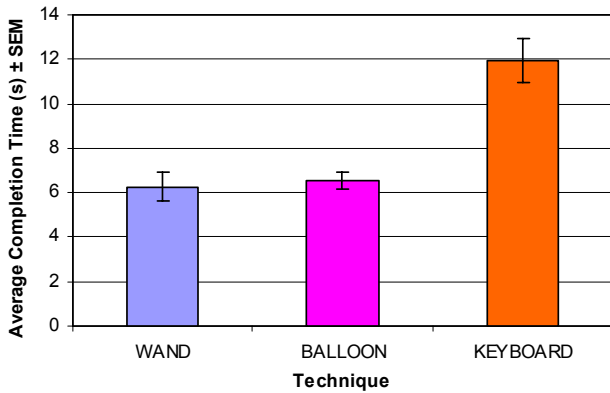


Figure 7: Average completion times (seconds) for the three techniques: KEYBOARD was significantly slower than BALLOON ($t(11)=5.07$, $p<0.001$), while WAND and BALLOON performed without statistically significant differences overall.

(H4) KEYBOARD would have the lowest error rate, since tracking precision, hand tremor, and arm fatigue would not affect it at all.

Overall, we were hoping to show that BALLOON would perform almost as fast as WAND, but have substantially lower error rate due to reduction of hand tremor and fatigue.

4.6 Results

Our data was first cleared by removing outliers, which accounted for 1.5% of all trials. The trials that ended in less than one second or more than one minute, or that resulted in a selection more than 5 cm away from the set of targets, were classified as outliers. Most of the outliers were caused by mistaken clicking when the participant inadvertently executed a selection command rather than moving or resizing the cursor. Outliers were present for all three techniques: 3 in WAND, 10 in BALLOON, and 9 in KEYBOARD, out of 480 selections for each technique (40 selections \times 12 participants). We analyzed our outlier-free results according to three factors: completion time, error rate, and subjective ratings.

4.6.1 Completion Time Analysis

We performed a 3 (Technique) \times 4 (Size) repeated measured ANOVA on mean selection times for the successfully completed trials (91% of outlier-free trials), with our subjects as a random variable. We found significant main effects across all conditions.

As expected, Technique had a significant main effect on completion times ($F_{(2,22)}=22.334$, $p<0.001$). KEYBOARD was on average more than 5 s slower than BALLOON ($t(11)=5.07$, $p<0.001$), while there was no statistically significant difference in

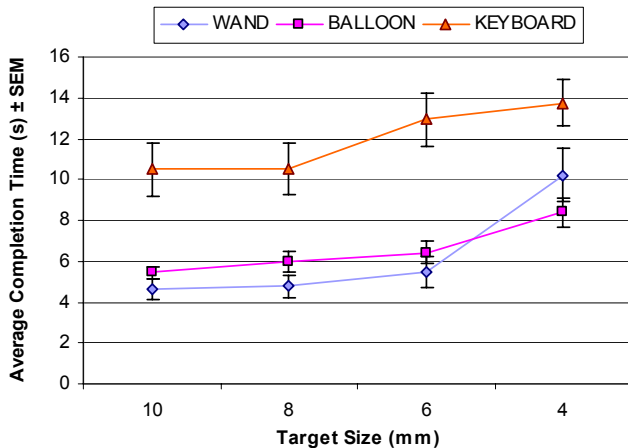


Figure 8: Performance of three techniques across four target sizes.

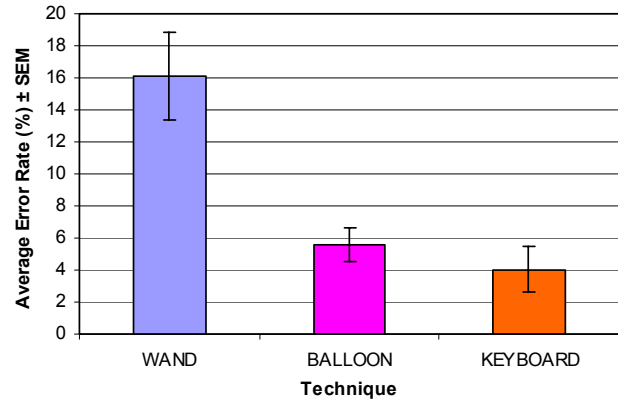


Figure 9: Average error rate (%) of the three techniques. We found no statistically significant difference between the error rates of BALLOON and KEYBOARD.

the performance of WAND and BALLOON (Figure 7). The average completion time difference between WAND and BALLOON techniques was very small (0.3s). This confirmed H2, but failed to confirm H1. Size contained significant effects ($F_{(3,33)}=20.228$, $p<0.001$) that confirmed the expected result: selection time increased as target size decreased.

While the interaction of Technique and Size did not contain significant effects overall, it is interesting to observe that WAND was more negatively affected when selecting the smallest target (4 mm) than BALLOON (Figure 8). Overall, when selecting the smallest target, BALLOON was the fastest technique.

4.6.2 Error Rate Analysis

To examine the performance of our techniques with regard to error rate, we performed a 3 (Technique) \times 4 (Size) repeated measured ANOVA on mean error rate data, with our subjects as a random variable. Significant main effects were found across all conditions.

The Technique factor contained a significant main effect ($F_{(2,22)}=18.707$, $p<0.001$). On average, WAND was more than three times more error-prone than BALLOON ($t(11)=4.635$, $p<0.001$), with an error rate of 16.1%, compared to 5.5% for BALLOON or 4.1% for KEYBOARD (Figure 9). This confirmed H3. Interestingly, while our data showed that BALLOON was slightly more error prone, this difference was not statistically significant. Therefore, H4 was not confirmed.

The Size factor also had a significant effect on error rate ($F_{(3,33)}=4.672$, $p<0.01$) (Figure 10). While all techniques performed somewhat consistently on 10, 8, and 6 mm targets, the smallest target (4mm) caused significantly more errors (approximately twice as many) than the others.

The interaction of Technique and Size was not found to be significant, with all techniques being equally affected by the target size decrease. However, we note that across all target sizes, WAND was consistently the most error-prone technique.

4.6.3 Subjective Evaluations

Subjects filled out a post-experiment questionnaire rating their experience with the three techniques on a five-point Likert scale (1 being most negative and 5 being most positive). The participants were asked to comment on the techniques with regard to ease of use, learning time, performance speed, mental effort, hand/arm fatigue, and enjoyment. Given the relatively small number of participants and some differences in their relative ratings, it is hard to draw significant conclusions from this data. However, we present these results as an indication of general trends of user preferences, rather than conclusive evidence.

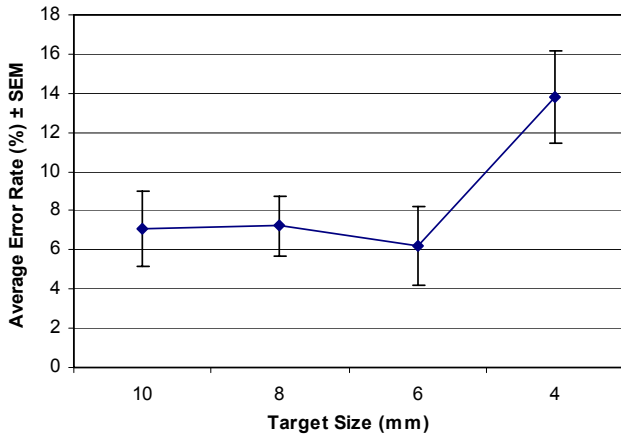


Figure 10: Average error rate (%) across target sizes. On average, selecting the smallest target (4mm) caused participants to commit twice as many errors as they did for the larger targets.

Overall, WAND was rated best for ease of use (3.75), learning time (4.41), and mental effort (3.91). In subjective evaluation of performance speed, the participants rated BALLOON (4.08) and WAND (4.00) as the fastest with almost identical scores, while KEYBOARD was given a much lower average score (2.58). These results are consistent with our quantitative experimental results, which did not find a statistically significant difference between the performance speeds of BALLOON and WAND for this task.

For hand/arm fatigue, KEYBOARD was given the best rating (4.50), with WAND rated the most fatiguing (3.00), and BALLOON rated 3.58. Subjects found BALLOON to be the most enjoyable to use (4.33), while KEYBOARD was the most frustrating (2.91). This might be skewed by the novelty and “coolness” factor of BALLOON, as noted by five participants. Overall, half of our participants (6 out of 12) rated WAND as their top choice for a 3D selection task, followed by BALLOON (4 out of 12). The preference for WAND is hardly surprising, considering the directness of this technique; however, many participants who preferred WAND overall, stated in their comments that for small targets they preferred BALLOON given that it “caused less fatigue and was easier to use more accurately than the wand” (Subject 8).

5 DISCUSSION AND FUTURE WORK

While we expected that BALLOON would achieve lower error rates than WAND, we were pleasantly surprised that there was no statistically significant difference in performance speeds between BALLOON and WAND in our study. Our data suggests that Balloon Selection provides the user with fast and slightly indirect interaction, while simultaneously reducing hand fatigue and lowering selection error rate.

We believe that much further work can be done to optimize Balloon Selection for specific tasks. For example, we plan to investigate ways to further eliminate spurious selections, as well as evaluate the clutching enhancements outlined in Section 3.4. Furthermore, for environments in which a wide range of selection heights must be supported, Balloon Selection could be extended to include GOGO-like nonlinear mapping of the distance between the anchor and stretching finger, or translation of the working volume.

While constraining the point of origin to the surface ensures that the user’s hands are not occluding the 3D object, it also means that the balloon’s first appearance will often be behind the virtual scene and therefore invisible to the user. One simple extension would be to support a user-settable initial offset from the

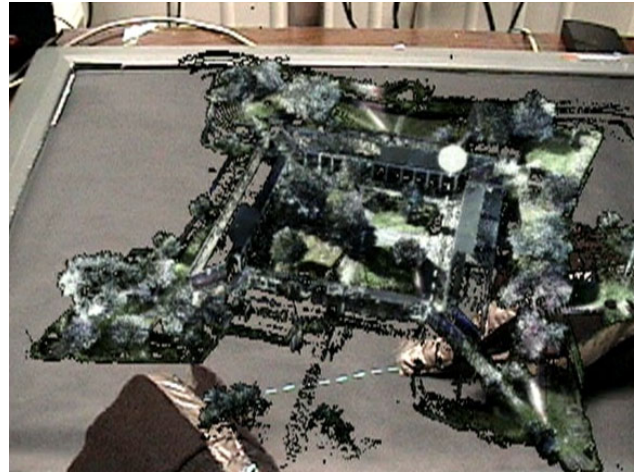


Figure 11: Balloon Selection being used in a dense point cloud model of Fort Jay at Governors Island, NY. (Range scan data courtesy of Columbia University Robotics Lab.)

surface, which would address sparse scenes with a raised ground plane. Additional visual feedback could be provided (e.g., cut-away views [12, 15] or transparency) to allow the user to see the balloon upon creation or as it moves within a densely populated environment. Furthermore, not being able to see one’s hands while interacting might be disturbing to the user. This effect should be evaluated in a separate study.

We are currently planning to further evaluate Balloon Selection in a hybrid AR environment for visualizing large point cloud data sets, such as the model of the Fort Jay complex at Governors Island, NY, shown in Figure 11. Because of the size of the Fort Jay complex (about 22,500 m²), each of the buildings in its scaled model is only several cm tall, but still contains millions of distinct points that could be of interest. We plan to explore whether the use of Balloon Selection can reduce the number of scaling operations that need to be performed on the model to be able to select features of interest with some desired degree of precision.

6 CONCLUSIONS

We have presented a 3D selection technique that decomposes a 3DOF precise positioning task into a set of 2DOF and 1DOF positioning tasks on the tabletop. Our study did not show that users were penalized for this decoupling in terms of interaction speed and simplicity, but rather that they gained accuracy while reducing hand fatigue. Although we do not believe that Balloon Selection is the best method for 3D selection in every circumstance, we are hopeful that for many 3D tasks that operate on the volume above an interaction surface, this technique will provide a good metaphor for 3D selection. In particular, Balloon Selection appears to work well for static targets, such as the ones described in this paper, and may also prove useful for targets whose movement is perceived by the user as being decomposed into 2D and 1D components. Our exploration of this single technique barely scratches the surface of the rich interaction space of multi-finger and multi-touch techniques for improving interactions on and above hybrid touch-sensitive surfaces.

7 ACKNOWLEDGEMENTS

We would like to thank Mitsubishi Electric Research Lab for generously donating the DiamondTouch table. This work was funded in part by NSF Grant IIS-01-21239, Office of Naval Research Contract N00014-04-1-0005, and gifts from Autodesk, Inc., NVIDIA, and Microsoft Research.

REFERENCES

- [1] C. Andujar and F. Argelaguet, "Friction Surfaces: Scaled Ray-Casting Manipulation for Interacting with 2D GUIs," Proc. of *12th Eurographics Symposium on Virtual Environments*, 2006, pp. 101-108.
- [2] N. I. Badler, K. H. Manocheri, and D. Baraff, "Multi-Dimensional Input Techniques and Articulated Figure Positioning by Multiple Constraints," Proc. of *Workshop on Interactive 3D Graphics*, ACM, 1986, pp. 151-169.
- [3] H. Benko, E. Ishak, and S. Feiner, "Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments," Proc. of *IEEE VR '05*, Bonn, Germany, 2005, pp. 209-216.
- [4] H. Benko, E. W. Ishak, and S. Feiner, "Collaborative mixed reality visualization of an archaeological excavation," Proc. of *IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR '04)*, Arlington, VA, 2004, pp. 132-140.
- [5] H. Benko, A. Wilson, and P. Baudisch, "Precise Selection Techniques for Multi-Touch Screens," Proc. of *CHI '06*, ACM Press, 2006, pp. 1263-1272.
- [6] R. Bolt, "Put-That-There: Voice and Gesture at the Graphics Interface," Proc. of *SIGGRAPH '80*, ACM Press, 1980, pp. 262-270.
- [7] D. Bowman, D. B. Johnson, and L. F. Hodges, "Testbed Evaluation of Virtual Environment Interaction," Proc. of *ACM VRST '99*, 1999, pp. 26-33.
- [8] D. A. Bowman and L. F. Hodges, "An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments," In *Symposium on Interactive 3D Graphics (ISD '97)*. Providence, RI: ACM Press, 1997, pp. 35-38.
- [9] F. P. Brooks, "Walkthrough - A Dynamic Graphics System for Simulating Virtual Buildings," Proc. of *Workshop on Interactive 3D Graphics*, ACM Press, 1986, pp. 9-21.
- [10] C. G. Butler and R. S. Amant, "HabilisDraw DT: A Bimanual Tool-Based Direct Manipulation Drawing Environment," Proc. of *CHI '04 Extended Abstracts*, Vienna, Austria, 2004, pp. 1301-1304.
- [11] W. Buxton and B. Myers, "A Study in Two-Handed Input," Proc. of *CHI '86*, 1986, pp. 321-326.
- [12] C. Coffin and T. Hollerer, "Interactive Perspective Cut-away Views for General 3D Scenes," Proc. of *IEEE Symposium on 3D User Interfaces (3DUI '06)*, 2006, pp. 25-28.
- [13] L. D. Cutler, B. Fröhlich, and P. Hanrahan, "Two-Handed Direct Manipulation on the Responsive Workbench," Proc. of *ACM Symposium on Interactive 3D Graphics*, Providence, RI, 1997.
- [14] P. Dietz and D. Lehigh, "DiamondTouch: a Multi-User Touch Technology," Proc. of *UIST '01*, 2001, pp. 219-226.
- [15] S. Feiner and D. Seligmann, "Cutaways and Ghosting: Satisfying Visibility Constraints in Dynamic 3D Illustrations," *The Visual Computer*, vol. 8, pp. 292-302, 1992.
- [16] A. Forsberg, K. P. Herndon, and R. Zeleznik, "Aperture Based Selection for Immersive Virtual Environments," Proc. of *ACM Symposium on User Interface Software and Technology*, ACM Press, 1996, pp. 95-96.
- [17] S. Fries and G. Kessler, "Precise and Rapid Interaction through Scaled Manipulation in Immersive Virtual Environments," Proc. of *IEEE VR '05*, Bonn, Germany, 2005, pp. 99-106.
- [18] T. Grossman, D. Wigdor, and R. Balakrishnan, "Multi-Finger Gestural Interaction with 3D Volumetric Displays," Proc. of *UIST '04*, Santa Fe, NM, 2004, pp. 61-70.
- [19] Y. Guiard, "Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain Model," *Journal of Motor Behavior*, vol. 19, pp. 486-517, 1987.
- [20] G. d. Haan, E. J. Griffith, M. Koutek, and F. H. Post, "Hybrid Interfaces in VEs: Intent and Interaction," Proc. of *12th Eurographics Symposium on Virtual Environments*, 2006, pp. 109-118.
- [21] J. W. Han, "Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection," Proc. of *UIST '05*, ACM Press, 2005, pp. 115-118.
- [22] K. P. Herndon, R. C. Zeleznik, D. C. Robbins, D. B. Conner, S. S. Snibbe, and A. van Dam, "Interactive Shadows," In *Proc. Symp. On User Interface Software and Technology (ACM UIST 1992)*: ACM Press, 1992, pp. 1-6.
- [23] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell, "A Survey of Design Issues in Spatial Input," In *ACM UIST '94 Symp. on User Interface Software & Technology*, 1994, pp. 213-222.
- [24] K. Hinckley, R. Pausch, D. Proffitt, and N. F. Kassell, "Two-Handed Virtual Manipulation," *ACM Transactions on Computer-Human Interaction*, vol. 5, pp. 260-302, 1998.
- [25] H. Hua, L. D. Brown, C. Gao, and N. Ahuja, "A new collaborative infrastructure: SCAPE," *Proc. IEEE Virtual Reality (VR'03)*, 2003.
- [26] R. J. K. Jacob and L. E. Sibert, "The Perceptual Structure of Multi-dimensional Input Device Selection," Proc. of *CHI '92*, Monterey, CA, ACM Press, 1992, pp. 211-218.
- [27] S. Kumar, P. Cohen, and H. J. Levesque, "The Adaptive Agent Architecture: Achieving Fault-Tolerance Using Persistent Broker Teams," *Proc. Int. Conf. on Multi-Agent Systems*, pp. 159-166, 2000.
- [28] G. Kurtenbach, G. Fitzmaurice, T. Baudel, and W. Buxton, "The Design of a GUI Paradigm based on Tablets, Two-hands, and Transparency," Proc. of *CHI '97*, ACM Press, 1997, pp. 35-42.
- [29] M. Mine, R., F. Brooks, P. J., Jr., and C. Sequin, H., "Moving Objects in Space: Exploiting Proprioception In Virtual-Environment Interaction," In *Proc. ACM SIGGRAPH '97*. Los Angeles, CA, 1997, pp. 19-26.
- [30] T. Moscovich, T. Igarashi, J. Rekimoto, K. Fukuchi, and J. F. Hughes, "A Multi-finger Interface for Performance Animation of Deformable Drawings," Proc. of *User Interface Software and Technology (UIST '05)*, Seattle, WA, ACM Press, 2005.
- [31] A. Olwal, H. Benko, and S. Feiner, "SenseShapes: Using statistical geometry for object selection in a multimodal augmented reality system," *Proc. IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2003)*, pp. 300-301, 2003.
- [32] A. Olwal and S. Feiner, "The Flexible Pointer: An Interaction Technique for Selection in Augmented and Virtual Reality," Proc. of *Conference Supplement of UIST '03*, Vancouver, BC, Canada, ACM Press, 2003, pp. 81-82.
- [33] J. S. Pierce, B. Stearns, and R. Pausch, "Voodoo Dolls: Seamless Interaction at the Multiple Scales in Virtual Environments," Proc. of *ACM Symposium on Interactive 3D Graphics*, ACM Press, 1999, pp. 141-145.
- [34] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa, "The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR," In *Proc. UIST '96 (ACM Symposium on User Interface Software and Technology)*, 1996, pp. 79-80.
- [35] H. T. Regenbrecht, M. T. Wagner, and G. Baratoff, "MagicMeeting: A Collaborative Tangible Augmented Reality System," *Virtual Reality - Systems, Development and Applications*, vol. 6, pp. 151-166, 2002.
- [36] J. Rekimoto, "SmartSkin: an Infrastructure for Free-hand Manipulation on Interactive Surfaces," Proc. of *CHI '02*, 2002, pp. 113-120.
- [37] R. Stoakley, M. Conway, and R. Pausch, "Virtual Reality on a WIM: Interactive Worlds in Miniature," In *Proc. CHI '95*. Denver, CO: ACM Press, 1995, pp. 265-272.
- [38] D. L. Vickers, "Sorcerer's apprentice: head-mounted display and wand," in *Department of Electrical Engineering Salt Lake City, UT: University of Utah*, 1972.
- [39] A. Wilson, "PlayAnywhere: A Compact Tabletop Computer Vision System" Proc. of *UIST '05*, Seattle, WA, 2005, pp. 83-92.
- [40] A. Wilson, "TouchLight: An Imaging Touch Screen and Display for Gesture-Based Interaction," Proc. of *ICMI '04*, 2004, pp. 69-76.
- [41] C. Wingrave, Y. Haciahetoglu, and D. Bowman, "Overcoming World in Miniature Limitations by a Scaled and Scrolling WIM," Proc. of *IEEE Symposium on 3D User Interfaces (3DUI '06)*, 2006, pp. 11-16.
- [42] H. P. Wyss, R. Blach, and M. Bues, "iSith - Intersection-based Spatial Interaction for Two Hands," Proc. of *IEEE Symposium on 3D User Interfaces (3DUI '06)*, Alexandria, VA, USA, 2006, pp. 59-61.