Touching the Void Revisited: Analyses of Touch Behavior on and above Tabletop Surfaces

Gerd Bruder, Frank Steinicke, Wolfgang Stuerzlinger

Touching the Void Revisited: Analyses of Touch Behavior On and Above Tabletop Surfaces

Gerd Bruder, Frank Steinicke
Department of Computer Science
University of Würzburg
{gerd.bruder,frank.steinicke}@uni-wuerzburg.de

Wolfgang Stuerzlinger
Department of Computer Science and Engineering
York University
wolfgang@cse.yorku.ca

Abstract. Recent developments in touch and display technologies made it possible to integrate touch-sensitive surfaces into stereoscopic three-dimensional (3D) displays. Although this combination provides a compelling user experience, interaction with stereoscopically displayed objects poses some fundamental challenges. If a user aims to select a 3D object, each eye sees a different perspective of the same scene. This results in two distinct projections on the display surface, which raises the question where users would touch in 3D or on the two-dimensional (2D) surface to indicate the selection. In this paper we analyze the relation between the 3D positions of stereoscopically displayed objects and the on- as well as off-surface touch areas. The results show that 2D touch interaction works better close to the screen but also that 3D interaction is more suitable beyond 10cm from the screen. Finally, we discuss implications for the development of future touch-sensitive interfaces with stereoscopic display.

Keywords: Touch-sensitive systems, stereoscopic displays, 3D interaction.

1 Introduction

Recent exhibitions and the entertainment market have been dominated by two different technologies: (i) (multi-)touch-sensitive surfaces and (ii) stereoscopic three-dimensional (3D) displays. Interestingly, these two technologies are orthogonal, as (multi-)touch is about input, whereas 3D stereoscopic display about output. Both technologies have the potential to provide more intuitive and natural interaction with a wide range of applications, including urban planning, architectural design, collaborative tabletops, or geo-spatial applications. First commercial hardware systems have recently been launched, e.g., [9], and interdisciplinary research projects explore interaction with stereoscopic content on 2D touch surfaces, e.g., [1, 2]. Moreover, an increasing number of hardware solutions provide the means to sense human gestures and postures not only on surfaces, but also in 3D space, e.g., the Kinect, the ThreeGear system, or Leap Motion [3]. The combination of these novel technologies provides enormous potential for a variety of new interaction concepts.

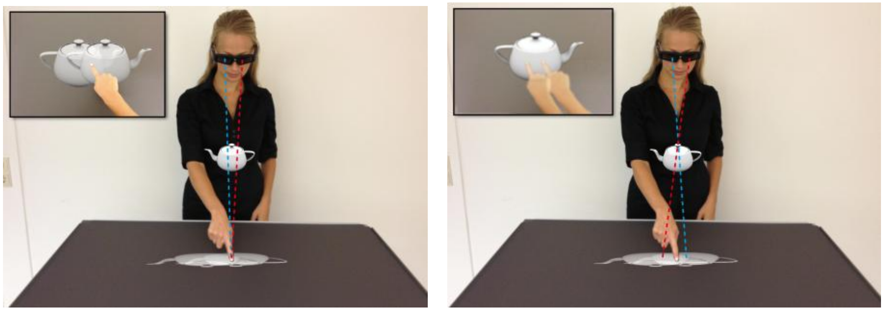

Fig. 1. Illustration of the main problem of touch interaction with stereoscopically displayed 3D data: (left) the user is either focused on her finger, which makes the selection ambiguous, or (right) on the object, which disturbs the visual perception of the finger.

Until recently, research in the area of (multi-)touch interaction was mostly focused on monoscopically displayed data. For this, the ability to directly touch elements without additional input devices has been shown to be very appealing for novice as well as expert users. Also, passive haptics and multi-touch displays have both shown their potential to considerably improve the user experience [6]. Touch surfaces build a consistent and pervasive illusion in perceptual and motor space that the two-dimensional graphical elements on the surface can be touched. Yet, three-dimensional data limits this illusion of place and plausibility [33]. Such 3D data sets are either displayed monoscopically, which has been shown to impair spatial perception and performance in common 3D tasks, or stereoscopically, which can cause objects to appear detached from the touch surface [24, 31,7].

Stereoscopic display technology has been available for decades. Recently, it was revived due to the rise of 3D cinema, upcoming 3D televisions and 3D games. With stereoscopic displays, each eye sees a different perspective of the same scene through appropriate technology. This requires rendering of two distinct images on the display surface. When using stereoscopic technology to display each projection to only one eye, objects may be displayed with *negative*, *zero*, or *positive parallax*, corresponding to their appearance in front, at, or behind the screen. Objects with zero parallax appear attached to the projection screen and are perfectly suited for touch interaction. In contrast, it is more difficult to apply direct-touch interaction techniques to objects that appear in front of or behind the screen [15, 26, 29]. In this paper we focus on the major challenge of touching objects that appear in front of the projection screen. Two methodologies can be used for touching such stereoscopic objects on a tabletop display:

1. If the touch-sensitive surface captures only direct contacts, the user has to penetrate the stereoscopically displayed object to touch the 2D surface behind it [38, 39].
2. Alternatively, if the system can capture finger movements in front of the screen, the user may virtually “touch” the object in *mid-air*, i.e., in 3D space [3].

Due to the discrepancy between perceptual and motor space and the missing passive haptic feedback, both approaches provide natural feedback only for objects rendered with zero parallax. This poses the questions where users “touch” a stereoscopically displayed object in 3D space. Here, one issue is the well-documented issue of misperception of distances in virtual 3D scenes [20]. Another problem arises from potential touch locations on the 2D display surface, as there are two distinct projec-

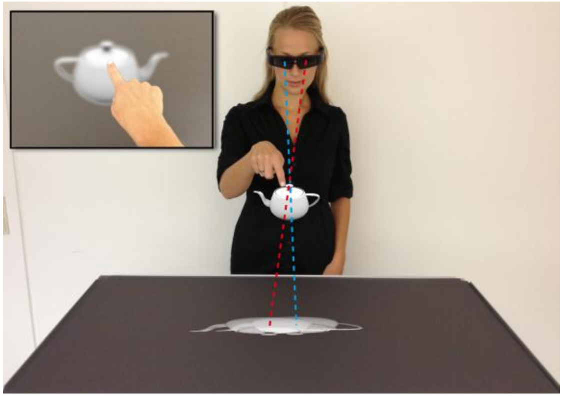

Fig. 2. Illustration of the main problem of 3D mid-air "touch" interaction with stereoscopically displayed 3D data: If a 3D tracking system is used, the user can see a stereoscopic image while converging to her finger. Due to the vegence-accommodation conflict, the virtual object appears blurred in comparison to the finger [7].

tions, one for each eye. If the user penetrates the object while focusing on her finger, the stereoscopic impression of the object is disturbed, since the user's eyes are not accommodated and converged to the projection screen's surface anymore. Thus, the left and right stereoscopic images of the object's projection appear blurred and can usually not be merged as illustrated in Figure 1 (left). However, focusing on the virtual object causes a disturbance of the stereoscopic perception of the user's finger, since her eyes are converged on the object's 3D position, see Figure 1 (right). If a 3D tracking system is used, the user can see a stereoscopic image while converging her eyes to her finger. Yet, due the vergence-accommodation conflict [7,8], the virtual object will appear blurred in comparison to the real finger (see Figure 2).

In this paper we address the challenge of how to interact with stereoscopic content in front of a touch-sensitive tabletop surface. Towards this, we also analyze touch behavior when touch sensing is constrained to the 2D screen surface. In order to allow the user to select arbitrary objects, a certain area of the touch surface, which we refer to as *on-surface target*, must be assigned to each object. In the monoscopic case the mapping between an on-surface touch area and the intended object point in the virtual scene is straightforward. Yet, with stereoscopic projection this mapping is more problematical. In particular, since there are different projections for each eye, the question arises *where* users touch the surface when they try to "touch" a stereoscopic object. In principle, the user may touch anywhere on the surface to select a stereoscopically displayed object. However, according to our previous work [39], the most likely alternatives that users try to touch are the (see Figure 4):

- *midpoint (M)* between the projections for both eyes,
- projection for *the dominant eye (D)*, or
- projection for the *non-dominant eye (N).*

A precise approach to this mapping is important to ensure efficient interaction and correct selections, in particular in a densely populated virtual scene. First, we determine a precise *on-display target area* where users touch the screen to select a 3D object. Second, we compare this approach with systems where the user's finger can be tracked in 3D space, and where users virtually touch objects in mid-air 3D space. The results of this experiment provide guidelines for the choice of touch technologies, as well as the optimal placement and parallax of interactive elements in stereoscopic touch environments.

In summary, our contributions are:

- An analysis of on-display touch areas for 3D target objects in stereoscopic touch-sensitive tabletop setups.
- A direct comparison of 2D touch and 3D mid-air selection precision.
- Guidelines for designing user interfaces for stereoscopic touch-sensitive tabletops.

The remainder of this paper is structured as follows. Section 2 summarizes related work in touch interaction and stereoscopic display. Section 3 describes the experiments we conducted to identify 2D/3D touch behavior. Section 4 presents the results, which are discussed in Section 5. An example application using the derived guidelines is described in Section 6. Section 7 concludes the paper.

2 Background

Recently, many approaches for extending multi-touch interaction techniques to 3D applications with monoscopic display have been proposed [15, 23, 28, 29]. For instance, Hancock et al. [15] presented the concept of shallow-depth 3D, i.e., 3D interaction within a limited range, to extend interaction possibilities with digital 2D surfaces. However, direct touch interaction with stereoscopically displayed scenes introduces new challenges [31], since the displayed objects can float in front of or behind the interactive display surface. Müller-Tomfelde et al. presented anaglyph- or passive polarization-based stereoscopic visualization combined with FTIR-based touch detection on a multi-touch enabled wall [25], and discussed approaches based on mobile devices for addressing the formulated parallax problems. The parallax problem described in the introduction is known from the two-dimensional representation of the mouse cursor within a stereoscopic image [31]. While the mouse cursor can be displayed stereoscopically on top of objects [31] or monoscopically only for the dominant eye [36], movements of real objects in the physical space, e.g., the user's hands, cannot be constrained such that they appear only on top of virtual objects. Grossman and Wigdor [11] provided an extensive review of the existing work on interactive surfaces and developed a taxonomy for this research. This framework takes the perceived and actual display space, the input space and the physical properties of an interactive surface into account. As shown in their work, 3D volumetric visualizations are rarely considered in combination with 2D direct surface input.

Even on monoscopic touch surfaces, the size of the human fingers and the lack of sensing precision can make precise touch screen interactions difficult [14, 40]. Some approaches have addressed this issue, for example, by providing an adjustable [6] or fixed cursor offset [27], by scaling the cursor motion [6] or by extracting the orientation of the user's finger [16].

2.1 Kinematics of Touch

The kinematics of point and grasp gestures and the underlying cognitive functions have been studied by many research groups [13, 21, 41]. For instance, it has been shown that total arm movement during grasping consists of two distinct component phases:

1. an initial, *ballistic phase* during which the user's attention is focused on the object to be grasped (or touched) and the motion is basically controlled by proprioceptive senses, and
2. a subsequent *correction phase* that reflects refinement and error-correction of the movement, incorporating visual feedback in order to minimize the error between the hand or finger, respectively, and the target [18].

Furthermore, MacKenzie et al. [19] have investigated the real-time kinematics of limb movements in a Fitts' task and have shown that, while Fitts' law holds for the total limb-movement time, humans usually start sooner decelerating the overall motion, if the target seems to require more precision in the end phase. The changes of the kinematics and control of the reaching tasks within virtual environments have also been investigated [7, 9, 29]. Valkov et al. [38] showed that users are, within some range, insensitive to small misalignments between visually perceived stereoscopic positions and the sensed haptic feedback when touching a virtual object. They proposed to manipulate the stereoscopically displayed scene in such a way that the objects are moved towards the screen when the user reaches for them [37, 38]. However, the problem is that objects have to be shifted in space, which may lead to a disturbed perception of the virtual scene for larger manipulations.

2.2 3D Touch for 3D Objects

To enable direct "touch" selection of stereoscopically displayed 3D objects in space, 3D tracking technologies can capture a user's hand or finger motions in front of the display surface. Hilliges et al. [16] investigated an extension of the interaction space beyond the touch surface. They tested two depth-sensing approaches to enrich multi-touch interaction on a tabletop setup. Although 3D mid-air touch provides an intuitive interaction technique, touching an intangible object, i.e., *touching the void* [8], leads to potential confusion and a significant number of overshoot errors. This is due to a combination of three factors: depth perception is less accurate in virtual scenes than in the real world, see e.g., [32], the introduced double vision, and also vergence-accommodation conflicts. A few devices, such as the CyberGrasp, support haptic feedback when touching objects in space, but require extensive user instrumentation. A similar option for direct touch interaction with stereoscopically rendered 3D objects is to separate the interactive surface from the projection screen, as proposed by Schmalstieg et al. [30]. In their approach, the user is provided with a physical *transparent prop*, which can be moved in front of the object of interest. This object can then be manipulated via single- or multi-touch gestures, since it has almost zero parallax with respect to the prop.

2.3 2D Touch for 3D Objects

Recently, multi-touch devices with non-planar surfaces, such as cubic [10] or spherical ones [5], were proposed. Other approaches are based on controlling the 3D position of a cursor through multiple touch points [4, 34]. These can specify 3D axes or points for indirect object manipulation. Interaction with objects with negative parallax on a multi-touch tabletop setup was addressed by Benko et al.'s balloon selection [4], as well as Strothoff et al.'s triangle cursor [34], which use 2D touch gestures to specify height above the surface. Valkov et al. [39] performed a user study, in which they displayed 3D objects stereoscopically either in front of or behind a large vertical projection screen. They recorded user behavior when instructed to touch the virtual 3D objects on the display surface. They identified that users tend to touch between the projections for the two eyes with an offset towards the projection for the dominant eye. However, the results suffered from a large variance between subjects. Hence, it is unclear how far these results can be applied to different setups, such as mobile screens or tabletops, where users have an easy frame of reference due to the bezel. Also, they may engage in different touch behavior due to physical support and gravity.

So far, no comparative analysis exists for 2D and 3D touch interaction in stereoscopic tabletop setups. Thus, it remains unclear if 2D touch is a viable alternative to 3D mid-air touch.

3 Experiments

Here we describe our experiments in which we analyzed the touch behavior as well as the precision of 2D touch and 3D mid-air touches. We used a standard ISO 9241-9 selection task setup [19] on a tabletop surface with 3D targets displayed at different heights above the surface, i.e., with different negative stereoscopic parallaxes.

3.1 Participants

Ten male and five female subjects (ages 20-35, M=27.1, heights 158-193cm, M=178.3cm) participated in the experiment. Subjects were students or members of the Departments of computer science, media communication or human computer-interaction. Three subjects received class credit for participating in the experiment. All subjects were right-handed. We used the Porta and Dolman tests to determine the sighting dominant eye of subjects [22]. This revealed eight right-eye dominant sub-

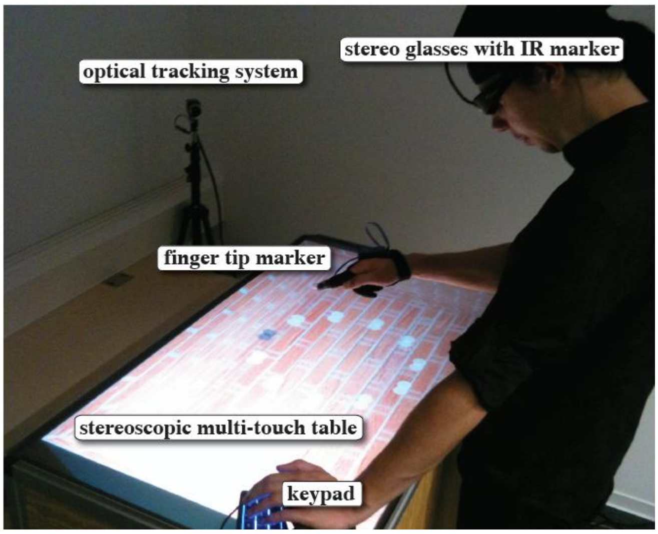

Fig. 3. Experiment setup: photo of a subject during the experiment (with illustrations). As illustrated on the screen, the target objects are arranged in a circle.

jects (7 males, 1 female) and five left-eye dominant subjects (2 males, 3 females). The tests were inconclusive for two subjects (1 male, 1 female), for which the 2 tests indicated conflicting eye dominance. All subjects had normal or corrected to normal vision. One subject wore glasses and four subjects wore contact lenses during the experiment. None of the subjects reported known eye disorders, such as color weaknesses, amblyopia or known stereopsis disruptions. We measured the interpupillary distance (IPD) of each subject before the experiment, which revealed IPDs between 5.8cm and 7.0cm (M=6.4cm). We used each individual's IPD for stereoscopic display in the experiment. 14 subjects reported experience with stereoscopic 3D cinema, 14 reported experience with touch screens, and 8 had previously participated in a study involving touch surfaces. Subjects were naive to the experimental conditions. Subjects were allowed to take a break at any time between experiment trials in order to minimize effects of exhaustion or lack of concentration. The total time per subject including pre-questionnaires, instructions, training, experiment, breaks, post-questionnaires, and debriefing was about 1 hour.

3.2 Materials

For the experiment we used a 62 x 112cm multi-touch enabled active stereoscopic tabletop setup as described in [7]. The system is shown in Figure 3 and uses rear diffuse illumination [24] for multi-touch. For this, six high-power infrared (IR) LEDs illuminate the screen from behind. When an object, such as a finger or palm, comes in contact with the diffuse surface it reflects the IR light, which is then sensed by a camera. We use a 1024x768 PointGrey Dragonfly2 with a wide-angle lens and a matching IR band-pass filter at 30 frames per second. We use a modified version of the NUI Group's CCV software to detect touch input on a Mac Mini server. Our setup uses a matte diffusing screen with a gain of 1.6 for the stereoscopic back projection. We used a 1280x800 Optoma GT720 projector with a wide-angle lens and an active DLP-based shutter at 60Hz per eye. We used an optical WorldViz PPT X4 system with sub-millimeter precision and sub-centimeter accuracy to track the subject's finger and head in 3D, both for 3D "touch" detection as well as view-dependent rendering. For this, we attached wireless markers to the shutter glasses and another diffused IR LED on the tip of the index finger of the subject's dominant hand. We tracked and logged both head and fingertip movements during the experiment.

The visual stimulus consisted of a 30cm deep box that matches the horizontal dimensions of the tabletop setup (see Figure 3). We matched the look of the scene to the visual stimuli used by Teather and Stuerzlinger [35, 36]. The targets in the experiment were represented by spheres, which were arranged in a circle as illustrated in Figure 3. A circle consisted of 11 spheres rendered in white, with the active target sphere highlighted in blue. The targets highlighted in the order specified by ISO 9241-9 [18]. The center of each target sphere indicated the exact position where subjects were instructed to touch with their dominant hand in order to select a sphere. For 3D touch this was the 3D position, and for 2D touch the center of the 2D projection. The size, distance, and height of target spheres were constant within circles, but varied between circles. Target height was measured as positive height from the level screen surface. Subjects indicated target selection using a Razer Nostromo keypad with their non-dominant hand. The virtual scene was rendered on an Intel Core i7 3.40GHz computer with 8GB of main memory, and an Nvidia Quadro 4000 graphics card.

3.3 Methods

The experiment used a 2 x 5 x 2 x 2 within-subjects design with the method of constant stimuli, in which the target positions and sizes are not related from one circle to the next, but presented randomly and uniformly distributed [11]. The independent variables were selection technique (2D touch vs. 3D mid-air touch), target height (between 0cm and 20cm, in steps of 5cm), as well as target distance (16cm and 25cm) and target size (2cm and 3cm). Each circle represented a different index of difficulty (ID), with combinations of 2 distances and 2 sizes. The ID indicates overall task difficulty [12]. It implies that the smaller and farther a target, the more difficult it is to select quickly and accurately. Our design thus uses four uniformly distributed IDs ranging from approximately 2.85bps to 3.75bps, representing an ecologically valuable range of difficulties for such a touch-enabled stereoscopic tabletop setup. As dependent variables we measured the on- as well as off-display touch areas for 3D target objects.

The experiment trials were divided into two blocks: one for the 2D and one for the 3D touch technique. We randomized their order between subjects. At the beginning of each block subjects were positioned standing in an upright posture in front of the tabletop surface as illustrated in Figure 3. To improve comparability, we compensated for the different heights of the subjects by adjusting a floor mat below the subject's feet, resulting in an (approximately) uniform eye height of 1.85cm for each subject during the experiment. The experiment started with task descriptions, which were presented via slides on the tabletop surface to reduce potential experimenter bias. Subjects completed 5 to 15 training trials with both techniques to ensure that they correctly understood the task and to minimize training effects. Training trials were excluded from the analysis.

In the experiment, subjects were instructed to touch the center of the target spheres as accurately as possible (either with 2D or 3D touch), for which they had as much time as needed. For this, subjects had to position the tip of the index finger of their dominant hand inside the 3D sphere for the 3D touch condition, or push their finger through the 3D sphere until it reached the 2D touch surface. Subjects did not receive feedback whether they "hit" their target, i.e., subjects were free to place their index finger in the real world where they perceived the virtual target to be. We did this to evaluate the often-reported systematical over- or underestimation of distances in virtual scenes, which can be observed even for short grasping-range distances [32], as also tested in this experiment. Moreover, we wanted to evaluate the impact of such misperceptions on touch behavior in stereoscopic tabletop setups. We tracked the tip of the index finger in both 2D and 3D touch conditions. When subjects wanted to register the selection, they had to press a button with their non-dominant hand on the keypad. We recorded a distinct 2D and 3D touch position for each target location for each configuration of independent variables, with a total of 20 circles and 220 recorded touch positions per participant.

4 Results

In this section we summarize the results from the 2D and 3D touch experiment. We had to exclude two subjects from the analysis who obviously misunderstood the task. We analyzed these results with a repeated measure ANOVA and Tukey multiple comparisons at the 5% significance level (with Bonferonni correction).

4.1 2D Touch

For the 2D touch technique, we evaluated the judged 2D touch points on the surface relative to the potential projected target points, i.e., the midpoint (M) between the projections for both eyes, as well as the projection for the dominant (D), and the non-dominant (N) eye, as illustrated in Figure 4. Figure 5 shows scatter plots of the distribution of the touch points from all trials in relation to the projected target centers for the dominant and non-dominant eye for the different heights of 0cm, 5cm, 10cm, 15cm and 20cm (bottom to top). We normalized the touch points in such a way that the dominant eye projection D is always shown on the left, and the non-dominant eye projection N is always shown on the right side of the plot. The touch points are displayed relatively to the distance between both projections.

As it is illustrated in Figures 4 and 5, we observed three different behaviors when subjects used the 2D touch technique. In particular, eight subjects touched towards the midpoint, i.e., the center between the dominant and non-dominant eye projections. This includes the two subjects for whom eye dominance estimates were inconclusive. We arranged these subjects into the group G_M. Furthermore, three subjects touched towards the dominant eye projection D, which we refer to as group G_D, and three subjects touched towards the non-dominant eye projection N, which we refer to as group G_N. This points towards an approximately 50/50% split in terms of behaviors in the population, i.e., between group G_M and the composite of groups G_D and G_N.

We found a significant main effect of the three groups ($F(2,11)=71.267$, $p<.001$, partial-eta^2=.928) on the on-surface touch areas. Furthermore, we found a significant two-way interaction effect of the three groups and target heights ($F(8,44)=45.251$, $p<.001$, partial-eta^2=.892) on the on-surface touch areas. The post-hoc test revealed that the on-surface target areas, see Figure 5, significantly ($p<.001$) vary for objects that are displayed at heights of 15cm or higher. For objects displayed at 10cm height group G_D and G_N vary significantly ($p<.02$). For objects displayed below 10cm we could not find any significant difference. As illustrated in Figure 5, for these heights the projections for the dominant and non-dominant eye are close together, and subjects touched almost the same on-screen target areas.

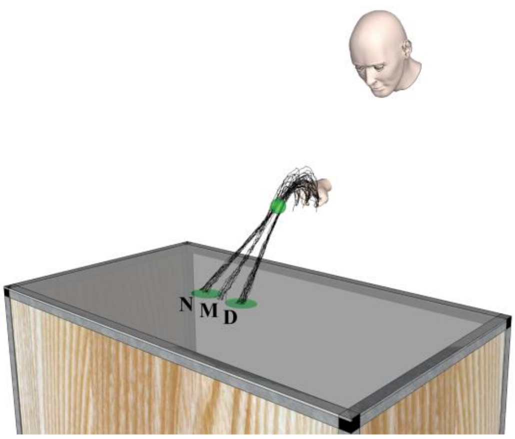

Fig. 4. Illustration of finger movement trails for eye user groups touching towards the dominant eye projection (D), non-dominant eye projection (N), or towards the midpoint. The trails have been normalized and are displayed here for a right-eye dominant user.

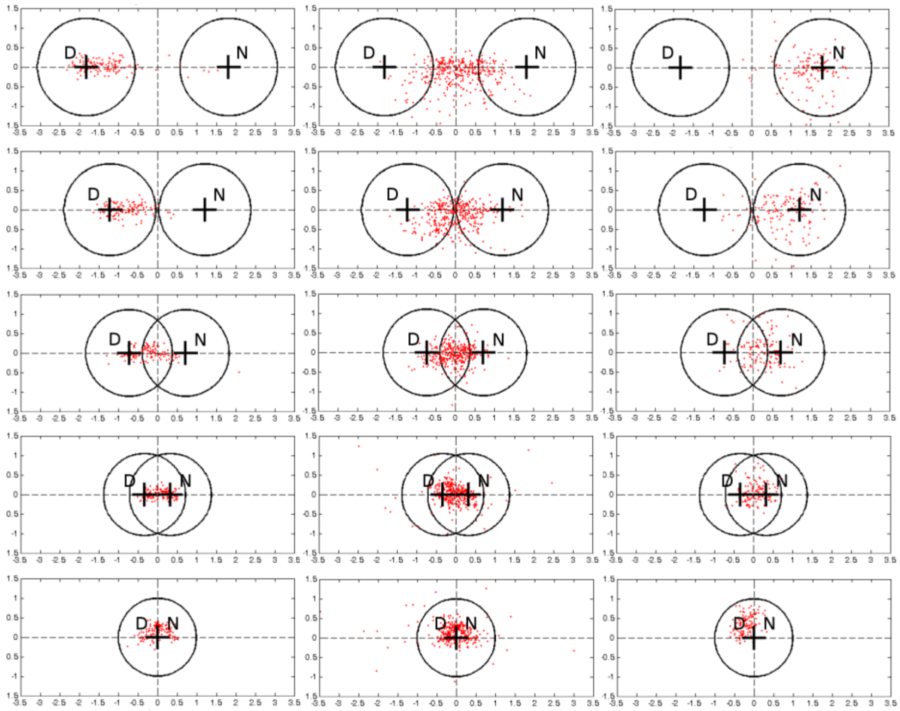

Fig. 5. Scatter plots of relative touch points between the dominant (D) and non-dominant (N) eye projections of the projected target centers on the surface for the 2D touch technique. Black crosses indicate the two projection centers. Black circles indicate the approximate projected target areas for the dominant and non-dominant eye. Top to bottom rows show results for 20cm, 15cm, 10cm, 5cm, and 0cm target heights. The left column shows subject behavior for dominant-eye touches (3 subjects), the middle for center-eye touches (8 subjects), and the right for non-dominant-eye touches (3 subjects). Note that the distance between the projection centers depends on the target height.

Considering the on-surface touch areas, we found that on average the relative touch point for group G_D was 0.97D+0.03N for projection points $D \in \mathbb{R}^2$ and $N \in \mathbb{R}^2$, meaning the subjects in this group touched towards the projection for the dominant eye, but slightly inwards to the center. The relative touch point for group G_N was 0.11D+0.89N, meaning the subjects in this group touched towards the projection for the non-dominant eye, again with a slight offset towards the center. Finally, for group G_M we found that on average the relative touch point for this group was 0.504D+0.596N. We could not find any significant difference for the different heights, i.e., the touch behaviors were consistent throughout the tested heights.

However, we observed a trend of target height on the standard deviations of the horizontal distributions (*x*-axis) of touch points for all groups as shown in Figure 5. For 0cm target height we found a mean standard deviation (SD) of 0.29cm, for 5cm SD 0.32cm, for 10cm SD 0.42cm, for 15cm SD 0.52cm, and for 20cm SD 0.61cm. For the vertical distribution (*y*-axis) of touch points and at 0cm target height we found a mean SD of 0.20cm, for 5cm SD 0.20cm, for 10cm SD 0.25cm, for 15cm SD 0.29cm, and for 20cm SD 0.30cm.

In summary, the results for the 2D touch technique show a significant effect for the different user groups on the on-surface touch area over the range of tested heights. These on-surface touch areas vary significantly for objects displayed at heights of 10cm and higher.

4.2 3D Touch

We analyzed the tracked physical 3D "touch" points where subjects judged the perceived center of the mid-air target spheres for the 3D touch technique in terms of their deviation from their actual position in the 3D virtual scene. Figure 6 shows scatter plots of the distribution of judged target positions in relation to the 3D target centers for the different target heights over all trials. The red dots indicate the center positions of the spheres as judged by the subjects. The black wireframe spheres illustrate the actual position and size of the objects. We normalized the judged positions relative to the optical view angle towards the target center. We found no significant difference in the judged positions for the three groups identified in Section 4.1 and pooled the data.

We analyzed the effect of target height on the subjects' judgments. We found a significant main effect of target height on the distances of judged positions from the displayed target centers. Mauchly's test indicated that the assumption of sphericity had been violated for effects of height on the distances of judged positions ($chi^2(9)=62.388$, $p<.001$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (epsilon=.302). The results show that the distances of judged positions significantly differs for heights ($F(1.21,15.725)=12.846$, $p<.002$, partial-$eta^2=.497$).

A Tukey post-hoc test with Bonferroni correction revealed that subjects estimated the target centers significantly closer to the actual displayed target centers for the 0cm targets in comparison to targets displayed at 20cm height ($p<.002$). For all other heights the results suggest that the higher the targets are displayed, the larger are the deviations. Pooling over all subjects, we observed mean distances to target centers of M=0.56cm (SD=0.27cm) for 0cm target height, M=0.88cm (SD=0.53cm) for 5cm, M=0.97cm (SD=0.61cm) for 10cm, M=1.32cm (SD=0.93cm) for 15cm, and M=1.90cm (SD=1.48cm) for 20cm. The results suggest that the physical constraints provided by the touch surface at 0cm height reduced judgment errors for objects at zero parallax relative to the other heights. We found no significant difference when comparing to the results for the 2D touch technique at 0cm target height as presented in Section 4.1.

As it can be seen in Figure 6, subjects made larger errors along the view axis than along the orthogonal axes. For the mid-air target positions we found a mean standard deviation of 1.43cm along the optical line-of-sight, a mean SD of 0.36cm parallel to the touch surface, and a mean SD of 0.50cm orthogonal to the other axes. Furthermore, these deviations increased with increasing target heights. For the different target heights above the surface we observed standard deviations of judged positions along the optical line-of-sight of SD=2.20cm (for 20cm target height), SD=1.52cm (15cm), SD=1.05cm (10cm), and SD=0.94cm (5cm). On the other hand, we observed standard deviations of judged positions orthogonal to the view axis parallel to the touch surface of only SD=0.49cm (20cm), SD=0.39cm (15cm), SD=0.30cm (10cm), and SD=0.27cm (5cm). Finally, we found standard deviations of judged positions orthogonal to the other axes of only SD=0.70cm (20cm), SD=0.55cm (15cm), SD=0.41cm (10cm), and SD=0.35cm (5cm). We further analyzed the data to determine whether deviations in judged target positions result from under- or overestimation of distances from the observer to the mid-air targets [7,8]. We observed a mean distance underestimation of 0.25% (SD=2.93%). Surprisingly, we found a distance overestimation of M=0.4% (SD=2.00%) and M=1.0% (SD=2.25%) for heights of 5cm and 10cm, respectively. Yet, we found an underestimation of M=-0.54% (SD=2.67%) and M=-0.98% (SD=4.18%) for heights of 15cm and 20cm, respectively.

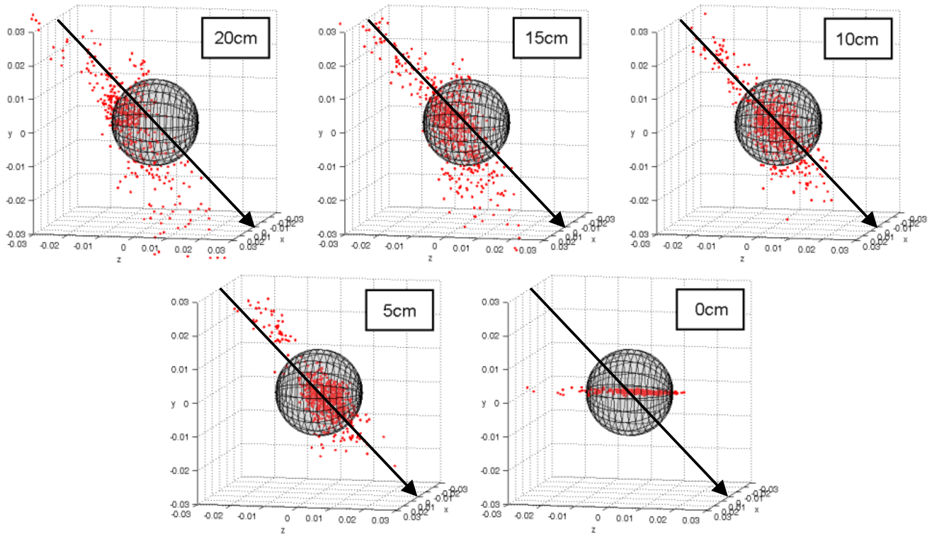

Fig. 6. Scatter plots of judged positions of the 3D target centers for the 3D touch technique over all subjects. Black wireframe spheres indicate the targets. The diagonal arrow illustrates the normalized view angle. The five diagrams show results for 20cm, 15cm, 10cm, 5cm, and 0cm target heights.

In summary, the results for the 3D touch technique show a significant effect of stereoscopic parallax over the range of tested heights on the precision and accuracy of judging the position of a target object.

5 Discussion

Our results provide interesting guidelines on how touch interaction in 3D stereoscopic tabletop setups should be realized. First of all and in contrast to previous work [39], our results show evidence for a twofold diversity of 2D touch behaviors of users. As shown in Figure 5, roughly half of the subjects in our study touched through the virtual object towards the center between the projections, and the other half touched towards projections determined by a single eye. The second group roughly splits in half again depending if they touch the projection for the dominant or non-dominant eye. Our results differ from the findings by Valkov et al. [39]. Using a setup with a large vertical projection plane they observed that subjects touched towards the center projection, with a slight offset towards the dominant eye. With 3 subjects touching towards the dominant eye, and 3 subjects towards the non-dominant eye in our study, user behavior in tabletop environments cannot be explained by this model. As a guideline, we suggest that the center between the projection for the left and right eye can be used to detect selections of objects stereoscopically displayed with less than 10cm height, since we did not observe significant differences between subjects at such heights. In order to reliably detect selections for objects higher above the screen, i.e., with larger parallaxes, our results suggest that for each user a calibration would be required. Our results confirm that this approach is highly beneficial, since subjects touched *consistently* for all heights towards the dominant, center, or non-dominant projection.

For practical considerations and to evaluate the ecological validity of using the 2D touch technique for selections of targets at a height between 0cm and 10cm, we computed the minimal on-surface touch area that supports 95% correct detection of all 2D touch points in our experiment. Due to the similar distributions of touch points between the three behavior groups for these heights shown in Figure 5, we determined the average minimal 95% on-surface region over all participants. Our results show that an elliptical area with horizontal and vertical diameter of 1.64cm and 1.07cm with a center in the middle between the two projections is sufficient for 95% correct detection. This rule-of-thumb heuristic for on-surface target areas is easy to implement and ecologically valuable considering the fat finger problem [14, 40]. Due to this problem objects require a relatively large size of between 1.05cm to 2.6cm for reliable acquisition, even in monoscopic touch-enabled tabletop environments.

The results of our second experimental condition reveal that distinct differences exist between the 3D mid-air touch technique and the 2D touch technique. These differences impact the relative performance and applicability for interaction with objects displayed stereoscopically at different heights above the surface. We found no behavior groups or effects of eye dominance on the distribution of judged 3D target positions. Our results show that target height has an effect on precision and accuracy of 3D selections, with large errors mainly along the optical line-of-sight, which we believe to correlate with distance misperception. For 3D objects displayed close to the display surface up to 10cm, touching objects in 2D on the surface by touching "through" the stereoscopic impression is more accurate than 3D mid-air touching. Considering that much research has shown that 3D mid-air touches of virtual objects suffer from low accuracy and precision due to visual conflicts, including vergence-accommodation mismatch, diplopia, and distance misperception [7, 8], it is a promising finding that the reduction of 3D selection tasks to 2D input with the 2D touch technique can improve performance for tabletop surface with stereoscopically displayed objects. However, the results also show that the accuracy for 2D touching of objects displayed above the screen decreases significantly for large negative parallax. The findings are encouraging for stereoscopic visualization on (multi-)touch surfaces. They suggest that virtual objects do not have to be constrained exactly at the zero-parallax level, but may deviate up to 10cm before 2D touch accuracy is significantly degraded [38, 39]. For such distances, the 2D touch technique is a good choice and instrumenting users with gloves or 3D markers can be avoided. Overall, our results show that it is possible to leverage stereoscopic cues in tabletop setups for an improved spatial cognition.

As a guideline for future tabletop setups with direct 2D touch input, the results suggest that touch-enabled 3D objects should not be displayed above an interactive display surface at more than about 10cm height. Above that, the disadvantages outperform the benefits and 3D interaction techniques should be used in that region, as they will provide more accurate interaction possibilities.

6 Example Application: Stereoscopic 3D Widgets

Our experiments have shown that the 2D touch technique has enormous potential as a new interaction paradigm for stereoscopic multi-touch surfaces as long as the objects are displayed with less than 10cm above the surface. In this region our 2D touch technique is the more accurate choice. While this constraint appears to limit the application scenarios in which one could use the 2D touch technique, it also ensures a simple implementation for interaction, in particular, a clear definition of on-display target areas as described in Section 5. Moreover, the size and scale of many virtual objects used in actual tabletop applications suit this constraint. For instance, 3D widgets can be displayed stereoscopically on any multi-touch surface and provide the user with a natural haptic feedback experience when she virtually touches them.

In order to evaluate the quality of the 3D touch technique in a real-world application, we adapted a simple visualization application for virtual caravans (see Figure 7). With this application customers can evaluate various types of caravans with several different features. The 3D widgets on the menu plane allow users to change the visual appearance of the caravan, lighting parameters, turn on signals, headlamps etc. We implemented the on-surface target areas of these 3D widgets as described in Section 5. The highest widgets, i.e., the 3D buttons on the menu panel, are displayed about 10cm above the surface. We used the same physical setup as described in Section 3.2. For this application we used the Unity3D game engine for the generation and rendering of the virtual scene. Unity3D provides a simple development environment for virtual scenes, animations and interactions. In order to synchronize virtual camera objects with the movements of a user, we integrated the MiddleVR for Unity software framework. MiddleVR supports streaming of motion data from our tracking system to Unity3D using the Virtual Reality Peripheral Network (VRPN) protocol. With this we stream head poses to Unity3D, resulting in a correct perspective from the user's point of view at all times.

We presented this application to four users, and made several interesting observations. First, all users acknowledged the stereoscopic display when viewing the 3D scene. Second, most users immediately understood that the menu panel with the 3D

Fig. 7. User interacting with a virtual scene in a stereoscopic multi-touch tabletop setup using touch-enabled 3D widgets. The widgets in the graphical user interface were rendered with negative parallax of up to 10cm height.

widgets provides a means to interact with the setup. Surprisingly, when users tried to "touch" the 3D widgets, they adapted their actions to the affordances provides by the widget. For instance, when they pressed the toggle switch, usually they touched its lifted part, although we did not distinguish between touch positions on the surface. We see this as further indication that stereoscopic display in combination with a touch-enabled surface does indeed support the notion of 3D physical interaction elements. Finally, none of the users complained about non-reactive 3D widgets, which might have occurred if they missed the on-surface target areas. This suggests that the shape and size of the on-surface touch areas, as determined by our above study, is sufficient for using stereoscopic 3D widgets in tabletop setups.

7 Conclusion and Future Work

In this paper we evaluated and compared 2D touch and 3D touch interaction techniques for scenes on touch-sensitive tabletop setups with stereoscopic display. We analyzed the differences of 3D mid-air touch input and a technique based on reducing the 3D touch problem to two dimensions by having users touch "through" the stereoscopic impression of 3D objects, resulting in a 2D touch on the display surface. We identified two separate classes of user behavior, with one group that touches the center between the projections, whereas the other touches the projection for the dominant or non-dominant eye. The results of the experiment show a strong interaction effect between input technique and the stereoscopic parallax of virtual objects.

The main contributions of this work are:

- We identified two separate classes of user behavior when touching "through" stereoscopically displayed objects.
- We compared precision and accuracy of 2D/3D direct touch input, which revealed that the 2D touch technique is a viable alternative to 3D touch interaction for object selection up to about 10cm height from the display surface.
- We determined on-surface target regions that support a simple implementation of the 2D touch technique. This enables intuitive touch input for 3D objects and widgets in stereoscopic 3D tabletop applications.

The results are encouraging for stereoscopic visualization in future touch-enabled tabletop setups, since no additional instrumentation and tracking technology is needed for objects with a small stereoscopic parallax. An interesting question for future work is if the results can be applied to portable setups, where the orientation of the touch-sensitive surface varies during interaction. We plan to further pursue these topics to provide compelling user experiences and effective user interfaces for touch-sensitive stereoscopic display surfaces. Moreover, we plan to investigate also how the 2D and 3D touch methods compare in terms of the speed-accuracy tradeoff.

Acknowledgements

This work was partly supported by grants from the Deutsche Forschungsgemeinschaft and the Natural Sciences and Engineering Research Council of Canada.

References

1. iMUTS - Interscopic Multi-Touch Surfaces. http://imuts.uni-muenster.de, 2013.
2. InSTInCT - Touch-based interfaces for Interaction with 3D Content. http://anr-instinct.cap-sciences.net, 2013.
3. Leap Motion. http://www.leapmotion.com, 2013.
4. H. Benko and S. Feiner. Balloon selection: A multi-finger technique for accurate low-fatigue 3D selection. Proc. IEEE 3DUI, pp. 79–86, 2007.
5. H. Benko, A. D. Wilson, and R. Balakrishnan. Sphere: multi-touch interactions on a spherical display. Proc. ACM UIST, pp. 77–86, 2008.
6. H. Benko, A. D. Wilson, and P. Baudisch. Precise selection techniques for multi-touch screens. Proc. ACM CHI, pp. 1263–1272, 2006.
7. G. Bruder, F. Steinicke, W. Stuerzlinger: Effects of Visual Conflicts on 3D Selection Task Performance in Stereoscopic Display Environments. Proceedings of IEEE Symposium on 3D User Interfaces (3DUI), IEEE Press, pp.115–118, 2013.
8. L.-W. Chan, H.-S. Kao, M. Y. Chen, M.-S. Lee, J. Hsu, and Y.-P. Hung. Touching the void: Direct-touch interaction for intangible displays. Proc. ACM CHI, pp. 2625–2634, 2010.
9. J.-B. de la Riviere, N. Dittlo, E. Orvain, C. Kervegant, M. Courtois, and T. Da Luz. ilight 3D touch: a multiview multitouch surface for 3D content visualization and viewpoint sharing. Proc. ACM ITS, pp. 312–312, 2010.
10. J.-B. de la Riviere, C. Kervegant, E. Orvain, and N. Dittlo. Cubtile: a multi-touch cubic interface. Proc. ACM VRST, pp. 69–72, 2008.
11. J. Ferwerda. Psychophysics 101: How to run perception experiments in computer graphics. In SIGGRAPH Core, 2008.
12. P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psych., 47:381–391, 1954.
13. L. Geniva, R. Chua, and J. T. Enns. Attention for perception and action: task interference for action planning, but not for online control. Exp. Brain Res., 185(4):709–717, 2008.
14. A. D. Hall, J. B. Cunningham, R. P. Roache, and J. W. Cox. Factors affecting performance using touchentry systems: Tactual recognition fields and system accuracy. In J. Appl. Psychol., 4:711–720, 1988.
15. M. Hancock, S. Carpendale, and A. Cockburn. Shallow-depth 3D interaction: design and evaluation of one-, two- and three-touch techniques. Proc. ACM CHI, pp. 1147–1156, 2007.
16. O. Hilliges, S. Izadi, A. D. Wilson, S. Hodges, A. Garcia-Mendoza, and A. Butz. Interactions in the air: Adding further depth to interactive tabletops. Proc. ACM UIST, pp. 139–148, 2009.
17. C. Holz and P. Baudisch. The generalized perceived input point model and how to double touch: Accuracy by extracting fingerprints. Proc. ACM CHI, pp. 1165-1174, 2009.
18. International Organization for Standardization. ISO/DIS 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices, 2000.
19. R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol., 3(3):203–220, 1993.
20. J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger and M. W. Haas, editors, Virtual and adaptive environments, volume Virtual and adaptive environments. Mahwah, 2003.
21. C. L. MacKenzie, R. G. Marteniuka, C. Dugasa, D. Liskea, and B. Eickmeiera. Three-dimensional movement trajectories in Fitts' task: Implications for control. Q.J. Exp. Psychology-A, 39(4):629–647, 1987.

22. A. P. Mapp, H. Ono, and R. Barbeito. What does the dominant eye dominate? A brief and somewhat contentious review. Percept. Psychophys., 65(2):310–317, 2003.
23. A. Martinet, G. Casiez, and G. Grisoni. The design and evaluation of 3D positioning techniques for multi-touch displays. Proc. IEEE 3DUI, 2010.
24. J. P. McIntire, P. R. Havig, and E. E. Geiselman. What is 3D good for? A review of human performance on stereoscopic 3D displays. In Proc. of the SPIE, Head- and Helmet-Mounted Displays XVII, 8383:1–13, 2012.
25. C. Müller-Tomfelde, J. Schöning, J. Hook, T. Bartindale, D. Schmidt, P. Oliver, F. Echtler, N. Motamedi, P. Brandl, and U. Zadow. Building interactive multi-touch surfaces. In C. Müller-Tomfelde, editor, Tabletops - Horizontal Interactive Displays, Human-Computer Interaction Series, pp. 27–49. Springer, 2010.
26. J. Pierce, A. Forsberg, M. Conway, S. Hong, R. Zeleznik, and M. Mine. Image plane interaction techniques in 3D immersive environments. Proc. ACM I3D, pp. 39–44, 1997.
27. R. L. Potter, L. J. Weldon, and B. Shneiderman. Improving the accuracy of touch screens: An experimental evaluation of three strategies. Proc. ACM CHI, pp. 27-32, 1988.
28. D. Pyryeskin, M. Hancock, and J. Hoey. Investigating selection above a multi-touch surface. Proc. ACM ITS, 2012.
29. J. L. Reisman, P. L. Davidson, and J. Y. Han. A screen-space formulation for 2D and 3D direct manipulation. Proc. ACM UIST, pp. 69–78, 2009.
30. D. Schmalstieg, L. M. Encarnacao, and Z. Szalavari. Using transparent props for interaction with the virtual table. Proc. ACM I3D, pp. 147-153, 1999.
31. J. Schöning, F. Steinicke, D. Valkov, A. Krüger, and K. H. Hinrichs. Bimanual interaction with interscopic multi-touch surfaces. Proc. INTERACT, pp. 40–53, 2009.
32. B. Sivak and C. L. Mac Kenzie. Integration of visual information and motor output in reaching and grasping: the contributions of peripheral and central vision. Neuropsychologia, 28(10):1095–1116, 1990.
33. M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Phil. Trans. R. Soc. B, 364:3549–3557, 2009.
34. S. Strothoff, D. Valkov, and K. H. Hinrichs. Triangle cursor: Interactions with objects above the tabletop. Proc. ACM ITS, pp. 111–119, 2011.
35. R. J. Teather and W. Stürzlinger. Pointing at 3D targets in a stereo head-tracked virtual environment. Proc. IEEE 3DUI, pp. 87–94, 2011.
36. R. Teather, W. Stuerzlinger, Pointing at 3D Target Projections with One-Eyed and Stereo Cursors, ACM CHI 2013.
37. D. Valkov, A. Giesler, and K. Hinrichs. Evaluation of depth perception for touch interaction with stereoscopic rendered objects. Proc. ACM ITS, 2012.
38. D. Valkov, F. Steinicke, G. Bruder, K. Hinrichs, J. Schöning, F. Daiber, and A. Krüger. Touching floating objects in projection-based virtual reality environments. Proc. JVRC, pp. 17–24, 2010.
39. D. Valkov, F. Steinicke, G. Bruder, and K. H. Hinrichs. 2D touching of 3D stereoscopic objects. Proc. ACM CHI, pp. 1353–1362, 2011.
40. D. Vogel and P. Baudisch. Shift: A technique for operating pen-based interfaces using touch. Proc. ACM CHI, pp. 657–666, 2007.
41. D. Whitney, D. A. Westwood, and M. A. Goodale. The influence of visualmotion on fast reaching movements to a stationary object. Letters to Nature, 423:869–873, 2003.