

Touch the 3rd Dimension! Understanding Stereoscopic 3D Touchscreen Interaction

Ashley Colley¹(✉), Jonna Häkkinen¹, Johannes Schöning², Florian Daiber³, Frank Steinicke⁴, and Antonio Krüger³

¹ CIE, University of Oulu, Oulu, Finland

{ashley.colley, jonna.hakkila}@cie.fi

² Hasselt University, tUL – iMinds, Diepenbeek, Belgium

johannes.schoening@uhasselt.be

³ DFKI GmbH, Saarbrücken, Germany

{florian.daiber, krueger}@dfki.de

⁴ Department of Informatics, University of Hamburg, Hamburg, Germany

steinicke@informatik.uni-hamburg.de

Abstract. The penetration of stereoscopic 3D (S3D) output devices is becoming widespread. S3D screens range in size from large cinema screens, to tabletop displays, TVs in living rooms, and even mobile devices are nowadays equipped with autostereoscopic 3D screens. As a consequence, the requirement for interacting with 3D content is also increasing, with “touch” being one of the dominant input methods. However, the design requirements and best practices for interaction with S3D touch screen user interfaces are only now evolving. To understand the challenges and limitations S3D technology brings to interaction design, and in particular the additional demands they place on the users of such interfaces, we present our research that considers stereoscopic touch screens of different sizes in static and mobile usage contexts. We report on perceptual, cognitive and behavioral aspects revealed by user studies and present interaction design requirements based on the findings.

Keywords: Stereoscopy · 3D user interfaces · Tabletops · Mobile user interfaces · User studies

1 Introduction

Display and visualization technologies are advancing on various frontiers, and today we see both increasingly bigger and smaller screens with ever improving resolutions. One emerging trend is the use of stereoscopic 3D (S3D). Stereoscopic 3D is probably most familiar to large audiences from 3D movies, and 3DTVs and S3D mobile devices are already mass-market products. Autostereoscopic 3D displays which require no special glasses to experience the 3D effect can be found in such products as 3D cameras, mobile phones, tablets and game consoles such as the Nintendo 3DS. The S3D technologies create the illusion of depth, such that graphical elements are perceived to pop up from or sink below the surface level of the screen. With negative disparity (or parallax), the User Interface (UI) elements appear to float in front of the

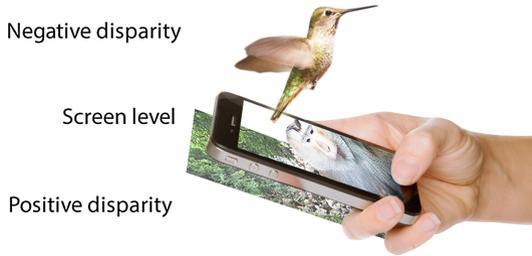


Fig. 1. An autostereoscopic 3D display, where the content is displayed with negative (bird), zero (monkey) and positive (background image) disparity/parallax.

screen, and with positive disparity (or parallax), the UI elements appears behind the screen level (see Fig. 1).

So far, much of the research in the field of human-computer interaction (HCI) on S3D has focused on visual ergonomics and visual comfort (see e.g. [28, 29, 39]), and until now, research on interaction and user experience design for S3D has received less attention. However, recent research has emerged in various application domains for S3D, ranging from interactive experiences in S3D cinemas [16], to mobile games [11], as well as in investigating the interaction design [6, 9, 47, 48].

The 3rd dimension provides new degrees of freedom for designers, and the illusion of depth can be utilized both for information visualization as well as for interactive user interface (UI) elements. However, as the physical displays are still two-dimensional, it remains challenging to design interactive systems utilizing S3D. Especially, touch screen interaction is problematic with objects that appear visually in 3D but are unable to be touched. This mismatch between the visual and tangible perception can potentially increase the user's cognitive load, and make the systems slower to use and less easy to comprehend. Thus, it is important to investigate the effects that stereoscopic 3D can have on interaction compared to conventional 2D UIs. In order to successfully introduce novel designs for S3D UIs, we need to investigate the implications of S3D on the user interaction e.g. in terms of touch accuracy and resolution.

In this chapter, we investigate touchscreen-based interaction with S3D visualizations, and compare the differences between S3D and 2D interaction. By examining the utility of S3D touchscreen UIs for practical usage when the user is interacting with different size S3D displays, we provide recommendations to researchers, developers and designers of such UIs. Especially, we:

1. Provide a systematic analysis of the differences in interaction between 2D and S3D touch screens.
2. Investigate touch screen based S3D interaction in the context of both large and small screens, namely tabletops and mobile devices.
3. Report on differences between S3D touch screen interaction while on the move (walking) and when static (standing).
4. Provide design recommendations for designing interactive touch S3D systems.

2 Related Work

In this section we first briefly summarize the more general research on S3D in the area of HCI and then continue with a review of related work focusing on touch screen interaction with S3D.

2.1 HCI Research in S3D

Visual Experience. With research on mobile S3D UIs, much of the emphasis so far has been focused on the output visualizations – video, images and visual cues - rather than on interacting with the device. Jumisko-Pyykkö et al. [24] studied the users' experience with mobile S3D videos, and discovered that the quality of experience was constructed from the following: visual quality, viewing experience, content, and quality of other modalities and their interactions. Pölönen et al. [39] report that with mobile S3D the perceived depth is less sensitive to changes in the ambient illumination level than perceived naturalness and overall image quality. Mizobuchi et al. [34] report that S3D text legibility was better when text was presented at zero disparity on a positive disparity background image when compared to presenting it hovering above a background image that was at zero disparity. It has also been pointed out that scaling the S3D content to different size displays is not straightforward, and has an effect on perceptual qualities [2].

Cognition and Perception. Considering the effect of S3D on users' cognition, Häkkinen et al. [19] have investigated people's focal attention with S3D movies using eye tracking. When comparing the S3D and 2D versions of the movies, they found that whereas in 2D movie viewers tend to focus quickly on the actors, in the S3D version the eye movements were more widely distributed among other targets. In their FrameBox and MirrorBox, Broy et al. [5] identify that the number of depth layers has a significant effect on task completion time and recognition rate in a search task.

The role of visual cues has also been investigated, with some conflicting conclusions on their relative importance in depth perception. Mikkola et al. [33] concluded that the stereoscopic depth cues outperform the monocular ones in efficiency and accuracy, whilst Huhtala et al. [23] reported that for a find-and-select task in a S3D mobile photo gallery, both the performance and subjective satisfaction were better when the stereoscopic effect was combined with another visual effect, i.e. dimming. Kerber et al. [26] investigated depth perception in a handheld stereoscopic augmented reality scenario. Their studies revealed that stereoscopy has a negligible if any effect on a small screen, even in favorable viewing conditions. Instead, the traditional depth cues, in particular object size, drive depth discrimination. The perception of stereoscopic depth on small screens has also been examined by Colley et al. [9], who compared users' ability to distinguish stereoscopic depth against their ability to distinguish 2D size. More recently, Mauderer et al. [32] investigated gaze-contingent depth of field as a method to produce realistic 3D images, and analyzed how effectively people can use it to perceive depth.

Non-touchscreen Interaction Techniques. Besides touchscreen-based interaction, several other interaction methods have been applied to S3D interfaces. Teather et al. [45]

demonstrated a “fish tank” virtual reality system for evaluating 3D selection techniques. Motivated by the successful application of Fitts’ law to 2D pointing evaluation, the system provided a testbed for consistent evaluation of 3D point-selection techniques. The primary design consideration of the system was to enable direct and fair comparison between 2D and 3D pointing techniques. To this end, the system presents a 3D version of the ISO 9241-9 pointing task. Considering a layered stereoscopic UI, Posti et al. [38] studied the use of hand gestures to move objects between depth layers.

Gestures with mobile phones have also been utilized as an interaction technique. In the context of an S3D cinema Häkkinen et al. [16] utilized mobile phone gestures as one method to capture 3D objects from the film. Using a large-scale stereoscopic display Daiber et al. [10] investigated remote interaction with 3D content on pervasive displays, concluding that physical travel-based techniques outperformed the virtual techniques.

Design Implications. The question of how the stereoscopic depth effect could be used in UI design has been investigated in several studies. Using S3D is perceived as a potential method of grouping similar content items or highlighting contextual information in a user interface [49]. There have also been design proposals where object depth within S3D UIs has been considered as an informative parameter e.g. to represent the time of the last phone call [18], and to identify a shared content layer in photo sharing [17]. Daiber et al. [11] investigated sensor-based interaction with stereoscopic displayed 3D data on mobile devices and presented a mobile 3D game that makes use of these concepts. Broy et al. have explored solutions for using S3D for in-car applications [4].

When evaluating the manufacturer’s UI in an off-the-shelf S3D mobile phone, Sunnari et al. [43] found that the S3D design was seen as visually pleasant and entertaining, but lacking in usability, and the participants had difficulties seeing any practical benefit of using stereoscopy. Other than for just hedonistic value, stereoscopy should be incorporated to the mobile UI in a way that it improves not only the visual design, but also the usability [15]. Considering this, more research on S3D for mobile devices equipped with an autostereoscopic display for both user experience and depth perception is still needed.

2.2 Touch Screen Interaction and S3D

In the monoscopic case, the mapping between an on-surface touch point and the intended object point in the virtual scene is straightforward, but with stereoscopic projection this mapping introduces problems [44]. To enable direct 3D “touch” selection of stereoscopically displayed objects in space, 3D tracking technologies can capture a user’s hand or finger motions in front of the display surface. Hilliges et al. [21] 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 being less accurate in virtual scenes than in the real world, see e.g. [41],

the introduced double vision, and also vergence-accommodation conflicts. Since there are different projections for each eye, the question arises: where do users touch the surface when they try to “touch” a stereoscopic object?

As described by Valkov et al. [47], for objects with negative parallax the user is limited to touch interaction on the area behind the object, since without additional instrumentation touch feedback is only provided at the surface. Therefore the user has to reach through the visual object to reach the touch surface with her finger. If the user reaches into an object while focusing on her finger, the stereoscopic effect for the object will be disturbed, since the user’s eyes are not accommodated and converged on the projection screen’s surface. Thus the left and right stereoscopic images of the object’s projection would appear blurred and could not be merged anymore. However, focusing on the virtual object leads to a disturbance of the stereoscopic perception of the user’s finger, since her eyes are converged to the object’s 3D position. In both cases touching an object may become ambiguous [47]. To reduce the perception problems associated with reaching through an object to the touch screen surface, Valkov et al. [46] created a prototype where the selected object moved with the user’s finger to the screen surface.

In principle, the user may touch anywhere on the surface to select a stereoscopically displayed object. However, in perceptual experiments Valkov et al. [47] found that users actually touch an intermediate point that is located between both projections with a significant offset towards the user’s dominant eye. Bruder et al. [7] compared 2D touch and 3D mid-air selection in a Fitts’ Law experiment for objects that are projected with varying disparity. Their results show that the 2D touch performs better close to the screen, while 3D selection outperforms 2D touch for targets further away from the screen.

Multi-touch technology provides a rich set of interactions without any instrumentation of the user, but the interaction is often limited to almost zero disparity [40]. Recently, multi-touch devices have been used for controlling the 3D position of a cursor through multiple touch points [1, 42]. 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 [1], as well as Strothoff et al.’s triangle cursor [42], which use 2D touch gestures to specify height above the surface. Grossman & Wigdor [14] provided an extensive review of the existing work on interactive surfaces and developed a taxonomy for classification of this research.

Considering the mobile device domain, a comprehensive body of work exists examining the input accuracy of 2D touch screens, e.g. Holtz and Baudisch [22] and Parhi et al. [36]. Holtz summarized that inaccuracy is largely due to a “parallax” artifact between user control based on the top of the finger and sensing based on the bottom side of the finger is particularly relevant, as in the S3D case another on-screen parallax effect is introduced. On small mobile screens touch accuracy is even more critical than on large touch devices. The effect of walking on touch accuracy for 2D touch screen UIs has also been previously researched [3, 25, 37]. For example, Kane [25] found variation in the optimal button size for different individuals, whilst Perry and Hourcade [37] concluded that walking slows down users’ interaction but has little direct effect on their touch accuracy.

2.3 Positioning Against Related Work

Taking the strong emphasis on visual design in S3D products and the output orientated prior art, there is a clear need to further investigate the interaction design aspects of S3D UIs. In this chapter, we focus on assessing touch screen interaction with S3D UIs in both large-screen tabletop and small screen mobile formats, in particular considering the target selection accuracy of objects at different depths and positions. Our aim is to provide practical information to assist the designers of S3D UIs.

Additionally, an understanding of the effect of the mobile domain on users' input accuracy when selecting targets is critical for the mobile UI designer. This extends the current body of research which has focused either to evaluate S3D in static conditions, or interaction with a 2D UI whilst on the move; the combination of S3D and user motion is not well researched in the literature.

3 Study I – Tabletop S3D Interaction

In this section, we describe experiments in which we analyzed the touch behavior as well as the precision of 2D touching of 3D objects displayed stereoscopically on a tabletop surface. We used a standard ISO 9241-9 selection task setup on a tabletop surface with 3D targets displayed at different heights above the surface, i.e. with different negative parallaxes. Further details about this study and a comparison to 3D selection can be found in [7]. In this section, we focus on the S3D touch technique, in which subjects have to push their finger through the 3D stereoscopically displayed object (i.e. with negative parallax) until it reaches the 2D touch surface.

3.1 Participants

Ten male and five female subjects (ages 20–35, $M = 27.1$, heights 158–193 cm, $M = 178.3$ cm) 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 (see [27]), to determine the sighting dominant eye of subjects. This revealed eight right-eye dominant subjects (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 5.8 cm and 7.0 cm ($M = 6.4$ cm). We used each individual's IPD for stereoscopic display in the experiment. Altogether 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 h.

3.2 Study Design

Materials. For the experiment we used a 62×112 cm multi-touch enabled active stereoscopic tabletop setup as described in [7]. The system uses rear diffuse illumination 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 used a 1024×768 PointGrey Dragonfly2 with a wide-angle lens and a matching IR band-pass filter at 30 frames per second.

We used a modified version of the NUI Group's CCV software to detect touch input on a Mac Mini server. Our setup used a matte diffusing screen with a gain of 1.6 for the stereoscopic back projection. We used a 1280×800 Optoma GT720 projector with a wide-angle lens and an active DLP-based shutter at 60 Hz per eye. Subjects indicated target selection using a Razer Nostromo keypad with their non-dominant hand. To enable view-dependant rendering, an optical WorldViz PPT X4 system with sub-millimeter precision and sub-centimeter accuracy was used to track the subject's head in 3D, based on wireless markers attached to the shutter glasses. Additionally, although not reported on in the scope of this chapter, a diffused IR LED on the tip of the index finger of the subject's dominant hand enabled tracking of the finger position in 3D (See [7]).

The visual stimulus consisted of a 30 cm deep box that matches the horizontal dimensions of the tabletop setup (see Fig. 2). The targets in the experiment were represented by spheres, which were arranged in a circle, as illustrated in Fig. 2. A circle consisted of 11 spheres rendered in white, with the active target sphere highlighted in blue. The targets were highlighted in the order specified by ISO 9241-9. 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. 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 screen surface. The virtual scene was rendered on an Intel Core i7 3.40 GHz computer with 8 GB of main memory, and an Nvidia Quadro 4000 graphics card.

Test Procedure. For our experimental analyses and description we used a $5 \times 2 \times 2$ 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. The independent variables were target height (between 0 cm and 20 cm, in steps of 5 cm), as well as target distance (16 cm and 25 cm) and target size (2 cm and 3 cm). Each circle represented a different index of difficulty (ID), with combinations of 2 distances and 2 sizes. The ID indicates overall task difficulty [13]. 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

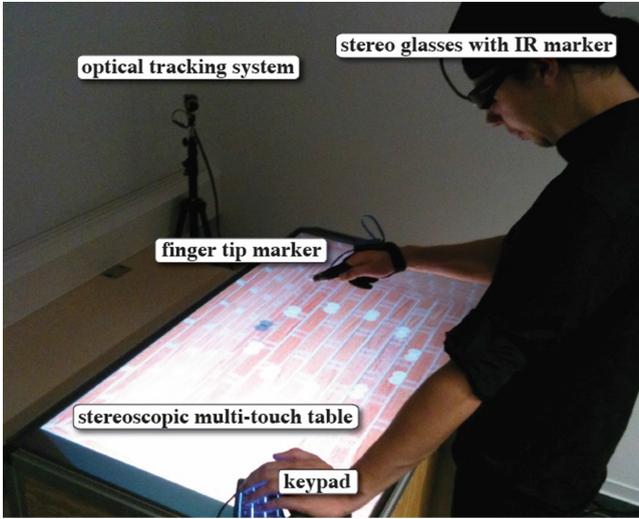


Fig. 2. Experiment setup: photo of a subject during the experiment (with illustrations). As illustrated on the screen, the target objects are arranged in a circle.

approximately 2.85 bps to 3.75 bps, representing an ecologically valuable range of difficulties for such a touch-enabled stereoscopic tabletop setup. As dependent variables we measured the on-display touch areas for 3D target objects.

At the start of the test, subjects were positioned standing in an upright posture in front of the tabletop surface as illustrated in Fig. 2. 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 1.85 cm 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 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, for which they had as much time as needed. For this, subjects had to 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 under-estimation of distances in virtual scenes, which can be observed even for short grasping-range distances, 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. 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 touch position for each target location for each configuration of independent variables.

3.3 Results

In this section we summarize the results from the tabletop S3D 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).

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. Figure 3 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 0 cm, 5 cm, 10 cm, 15 cm and 20 cm (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 relative to the distance between both projections.

As illustrated in Fig. 3, we observed three different behaviors. In particular, eight subjects touched towards the midpoint, i.e., the center between the dominant and non-dominant eye projections. These include 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 ($F(2,11) = 71.267$, $p < .001$, $\text{partial-}\eta^2 = .928$) on the on-surface touch areas. Furthermore, we found a significant two-way interaction effect of the three groups and target heights ($F(8,44) = 45.251$, $p < .001$, $\text{partial-}\eta^2 = .892$) on the on-surface touch areas. The post hoc test revealed that the on-surface target areas significantly vary ($p < .001$) for objects that are displayed at heights of 15 cm or higher. For objects displayed at 10 cm height group G_D and G_N vary significantly ($p < .02$). No significant difference was found for objects displayed below 10 cm. As illustrated in Fig. 3, for these heights the projections for the dominant and non-dominant eye are proximal, and subjects touched almost the same on-screen target areas.

Considering the on-surface touch areas, we found that on average the relative touch point for group G_D was $0.97D + 0.03 N$ 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 $0.11D + 0.89 N$, 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 $0.504D + 0.596 N$. 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 Fig. 3. For 0 cm target height we found a mean standard deviation (SD) of 0.29 cm, for 5 cm

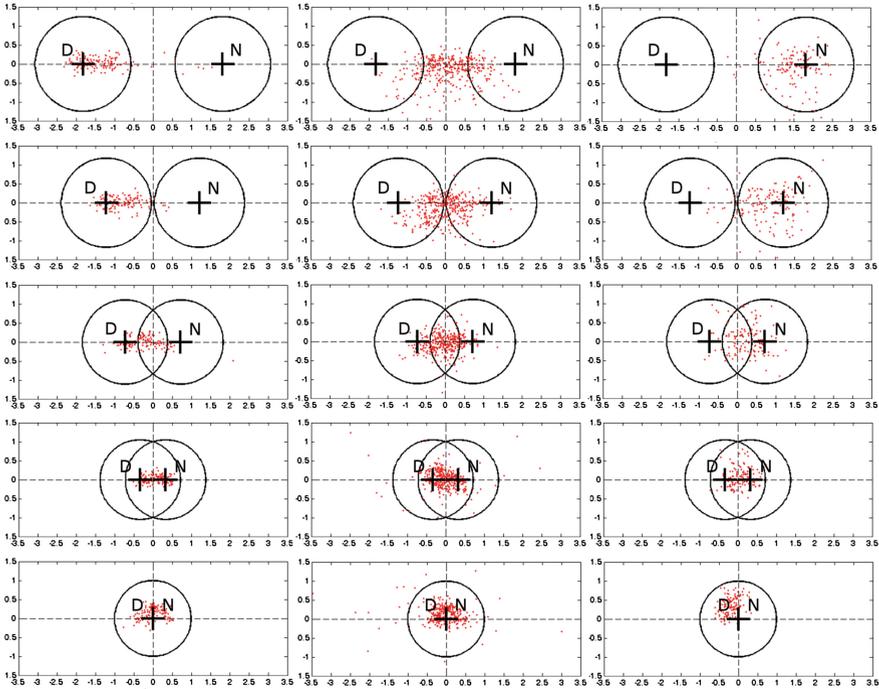


Fig. 3. Scatter plots of relative touch points between the dominant (D) and non-dominant (N) eye projections of the projected target centers on the surface for the 2D touch technique. Black crosses indicate the two projection centers. Black circles indicate the approximate projected target areas for the dominant and non-dominant eye. Top to bottom rows show results for 20 cm, 15 cm, 10 cm, 5 cm, and 0 cm 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.

SD 0.32 cm, for 10 cm SD 0.42 cm, for 15 cm SD 0.52 cm, and for 20 cm SD 0.61 cm. For the vertical distribution (y-axis) of touch points and at 0 cm target height we found a mean SD of 0.20 cm, for 5 cm SD 0.20 cm, for 10 cm SD 0.25 cm, for 15 cm SD 0.29 cm, and for 20 cm SD 0.30 cm.

Minimum Touch Target Sizes. For practical considerations and to evaluate the ecological validity of using the 2D touch technique for selections of targets at a height between 0 cm and 10 cm, 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 Fig. 3, 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 1.64 cm and 1.07 cm 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’ [20, 30]. Due to this problem, objects require a relatively large size of between 1.05 cm to 2.6 cm for reliable acquisition, even in monoscopic touch-enabled tabletop environments.

4 Study II – Mobile S3D Interaction

To systematically evaluate interaction with S3D in the mobile context, we designed a study to investigate users’ performance when interacting with a mobile S3D device and compared this to the 2D case. Here, we examine interaction with both negative and positive parallax (compared to our tabletop study in which only on-screen and negative parallax were considered). In contrast to the tabletop study, the mobile study utilized an autostereoscopic display and hence did not require shutter glasses to be worn to see the stereoscopic effect. Additionally, in this study no viewer dependent adjustment of the rendered view was made. As mobility is by definition a core of this usage context, we also examined differences between interaction when static and walking.

4.1 Participants

The average age of the users ($n = 27$) was 30.4 years (varying from 10 to 52), and 19 of the users were male and 8 were female. Tests to identify the user’s dominant eye (Portas method, see [27]) and stereovision were conducted. Of the 27 users 16 were found to have a dominant right eye whilst 9 had a dominant left eye. Two users were unable to determine which of their eyes was dominant. 24 of the users were right-handed and the remaining 3 left-handed. Most had watched a 3D movie, and approximately half had previously been exposed to a 3D TV or a 3D camera.

4.2 Study Design

As we wanted to focus only on the depth effect due to the stereoscopy, no additional visual depth clues such as shadows, object size or color were used in our test application. In all the tests a background wallpaper image was used, as in pilot tests the use of some type of textured background at a positive disparity/parallax (Fig. 1) was found to improve 3D perception. The background image chosen was a fine mesh pattern at 45°. This was chosen as it gave good 3D perception, but was such that it would not adversely influence the position of the users’ presses on targets. In the 2D tests the background was positioned at $z\text{-depth} = \text{zero}$ and in the 3D tests at $z\text{-depth} = 10$. Table 1 describes the $z\text{-depth}$ convention and values used in the study. All the tests were conducted on an autostereoscopic touchscreen mobile device, the LG P920 Optimus 3D mobile phone. This has a 4.3” display with a resolution of 480×800 pixels (with 217 ppi) and running Android 2.3.5.

Our target was to investigate the fundamental accuracy of users, hence our method avoided the use of any UI elements such as buttons, whose visuals could influence the position in which users tapped the screen. Thus, we followed a method similar to the

crosshairs approach employed by Holtz and Baudisch [22]. Our accuracy test method presented small circular targets, one at a time. The user was instructed to tap on the center of each target with their index finger. For each tap the coordinates of both the press and release touch events were logged, as well as the time between the target being displayed and the user pressing the target.

The accuracy test was conducted in the 2D and S3D condition. In the 2D condition a random sequence of 15 targets was presented. Each target was presented once. The targets were positioned in a grid pattern of 5 horizontal by 3 vertical. This pattern was chosen to give more data points and resolution in the horizontal axis, which is the most interesting for S3D. The targets and background image were displayed as normal 2D objects.

Table 1. Reference stereoscopic depths and calculated apparent distance behind the display (assuming viewing distance of 330 mm and inter-pupil distance of 63 mm)

Accuracy test z-depths	Apparent distance behind screen (mm)
10	12.0
3	3.6
2	2.5
0	0 (on screen)
-1	-1.2 (in front of screen)
-2	-2.5 (in front of screen)

In the S3D version of the accuracy test the same grid of 15 targets was used, but positioned at 5 different z-depths (see Table 1), thus making a total of 75 targets. The 75 targets were presented in a random order of x, y and z position, such that each target was presented once. In the S3D test, the background wallpaper image was positioned at depth $z = 10$, behind the screen. It should be noted that due to the difference in background depth, the S3D test with targets at $z = 0$ was not exactly equivalent to the 2D test.

Based on an eyes-to-screen distance of 330 mm (the mean of 3 pilot users) and an inter-pupil distance of 63 mm (see [12]) the object distances in front of and behind the screen are given in Table 1. These values give an approximation of the apparent depth of objects presented at the different reference depths (the distance of the screen from the eyes was not fixed during the tests).

4.3 Test Procedure

An acclimatization task of looking through an S3D image gallery, provided as default by the device manufacturer, enabled users to get accustomed to the S3D effect. Users were instructed to experiment to find the best position, i.e. distance from device to eyes, viewing angle, etc., where they could best see the 3D effect of the gallery images.

Each of the interaction tests: 2D accuracy and S3D accuracy were conducted in two different conditions: while seated (static) and whilst walking (on the move). For the walking tests, a route consisting of two markers 6 m apart was marked on the

laboratory floor. Users were instructed to walk at normal speed around the marked route whilst completing the tests. The order of presentation of the conditions was counterbalanced between participants, to reduce any effect of user learning or other side effects. Users also provided subjective feedback on the comfort of using the interface: this is reported in [9].

4.4 Results

Analysis of the accuracy tests is based on the recorded co-ordinates of the touch press event (i.e. “finger down”), as this defines the users’ fundamental accuracy. Presses that were more than 12 mm from the center of the visual target were considered as accidental and excluded from further analysis [31] (Removed data: static 2D = 2, 3D = 4, walking 2D = 3, 3D = 10).

Heatmaps of the press points for each test are shown in Fig. 4, which combines the presses for all of the depth layers. These charts plot the offset of each press point from the center of the presented circular target, i.e. the center of the visual target is at the origin of the chart.

The absolute distance of each press point from the center of the visual target (i.e. ignoring the direction of the error) is presented in Fig. 5. For the S3D layers (i.e. within the groups of five bars in Fig. 5), there was no significant difference between the mean error distance for each of the 3D depth layers (based on ANOVA analysis). Thus, targets at screen level ($z = 0$) were pressed no more accurately than targets at other depths. For all the S3D depth layers combined together, a 2-way ANOVA test revealed

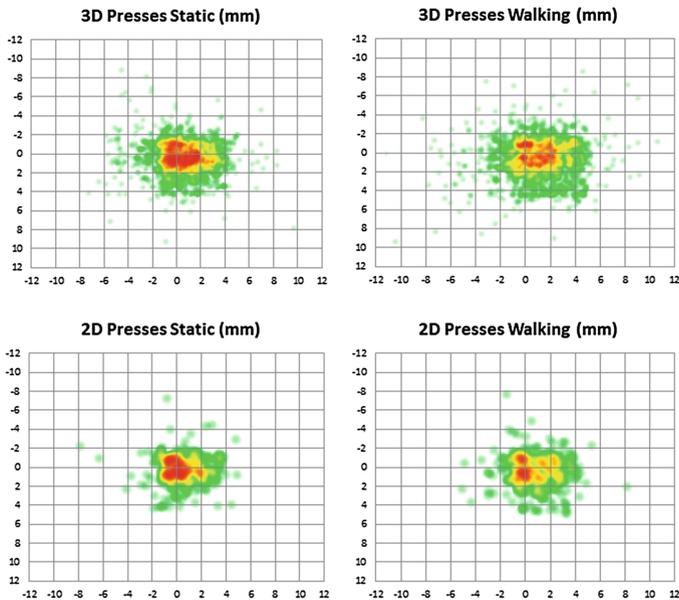


Fig. 4. Combined press points relative to target center for S3D and 2D.

both walking and S3D caused a significant increase in the mean error distance, ($F(2,4841) = 28.8, p < .001$) and ($F(2,4841) = 54.9, p < .001$), respectively. There was no significant interaction between the fixed factors.

Comparing the S3D $z = 0$ results, where the target is placed at screen depth, with the corresponding 2D cases, where the target is also at screen depth (circled on Fig. 5), showed significant degradation in accuracy in the S3D cases as compared with 2D cases for $z = 0$ ($T(403) = 3.52, p < .001$ and $T(402) = 4.35, p < .001$ for static and walking cases respectively). This difference is due to the S3D cases having a background image at positive disparity, placing the targets within a stereoscopic scene.

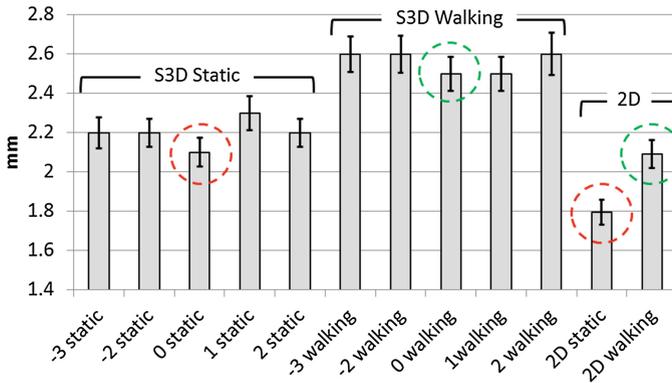


Fig. 5. Mean distance from target center for 2D & individual 3D layers whilst static and walking. Error bars show standard error of mean.

Additionally, in the S3D cases, the $z = 0$ targets are presented in a sequence of targets at other depths. These findings are important as they indicate that from the touch accuracy point of view, there is no benefit in placing targets at screen depth compared to any other depth level. Rather, to maximize touch accuracy, it is recommended that if 3D is not beneficial for other purposes within a UI, it is turned off.

Following a similar method to that employed in Study I, we separated users based on their dominant eyes. However in this case we were unable to ascertain any significance between the groups.

Minimum Touch Target Sizes. One approach to evaluate the degradation in accuracy caused by walking and S3D is to calculate the minimum touch target sizes for each case. The minimum touch target size is the size required for users to reliably hit touch targets in each tested mode. From the press point error distribution we calculated the 98 percentile points in positive and negative directions for both x and y . Taking the approach described by [36], this defines the bounding rectangle that captures 95 % of user presses. As the distributions are slightly skewed to either positive or negative sides, and in practice it is not possible to position a touch target at exactly the correct offset from the center, thus we take double the larger absolute value, which may be considered as the minimum size for targets in the UI. It should be noted that these

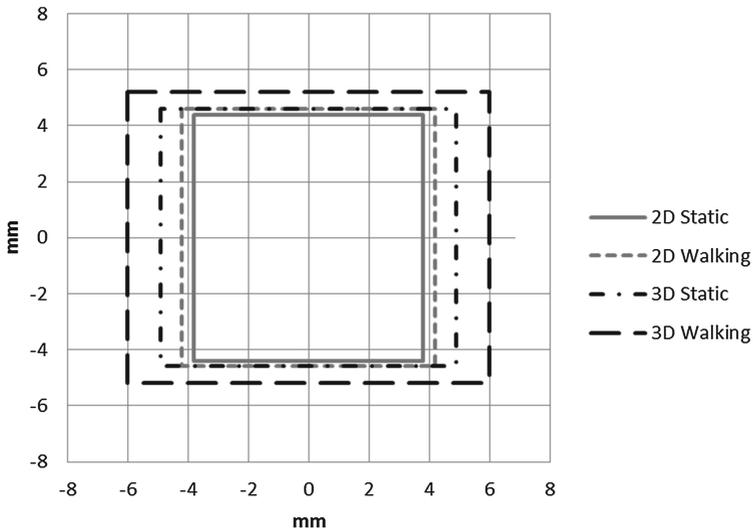


Fig. 6. Minimum target sizes for finger usage (95 % press success).

Table 2. Minimum target sizes for finger usage.

Conditions	Negative 98 percentile (mm)	Positive 98 percentile (mm)	Minimum target size (95 % presses on target) width x height (mm)
2D Static	x:3.5 y:3.9	x:3.8 y:4.4	7.6 × 8.8
2D Walking	x:3.6 y:3.7	x:4.2 y:4.6	8.4 × 9.2
3D Static	x:4.9 y:3.8	x:4.5 y:4.6	9.8 × 9.2
3D Walking	x:6.0 y:4.1	x:5.0 y:5.2	12.0 × 10.4

values are based on index finger usage, as in the test protocol. For the thumb, usage sizes will be somewhat larger (see [36]). The minimum target sizes are shown in Table 2 and diagrammatically in Fig. 6.

5 Discussion

5.1 Touch Performance

Examining our results for the tabletop case, the S3D touch technique appears to have significantly influenced 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 10 cm and higher. In contrast to previous work by Valkov et al. [47, 48], our results show evidence for a twofold diversity of 2D touch behaviors of users. As shown in Fig. 3, 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.

In the mobile case, and in the tabletop case for targets closer to the screen than 10 cm, we were not able to identify any significant effect based on the user's dominant eye. Hence we conclude that this phenomenon is not relevant to targets that are close to the screen level. Further, in the case of the small screen, interaction issues related to the general touch interaction accuracy dominate.

5.2 Minimum Touch Area

For tabletop interaction, we determined the minimum on-surface touch area (95 % correct detection) for interacting with 3D targets within 10 cm of the screen surface, to be an elliptical area with horizontal and vertical diameter of 16.4 mm and 10.7 mm with a center in the middle between the two projections. For interaction with the mobile device the corresponding area was a slightly smaller rectangle with dimensions of 9.8 mm × 9.2 mm when static, increasing to 12.0 mm × 10.4 mm when walking. However, further analysis of the differences is not informative, due to the large difference in the depth display capabilities between the mobile device and tabletop (1.2 mm vs. 10 cm and above).

Considering also the 2D performance in the mobile case, and comparing with previous studies, we find close agreement. For example, using their contact area model Holtz and Baudisch [22] report a 7.9 mm target size for 95 % accuracy, compared to our 7.6 mm × 8.8 mm (height × width) for the same accuracy. The aspect ratio information in our results provides interesting additional detail.

5.3 Mobile Context

Direct comparison between our results related to the differences in touch target sizes when walking with prior work on this topic is not straightforward, due to differences in study approaches. The majority of previous work has been based on a UI with varied size button targets, and measuring task completion time and error rates (for example [25]). In contrast, our button free approach aimed to minimize the influence of the visual UI design, and focused on the accuracy of separate tap events. Thus, in our test, accuracy was the only dependent variable.

Our approach enabled us to investigate interaction over the full screen area, and we were able to gain insight into the optimal aspect ratio of touch targets. However, our method is more abstract than actual UI based approaches, and requires some interpretation to transfer it to use in actual designs. It does not take account of other interaction issues e.g. related to Fitts' law [13]. Hence, it is perhaps the case that our method serves best as an initial one, the results of which could be validated by a button and task based method.

Based on our results, for a certain screen size, a S3D UI for use whilst walking should have only 60 % of the touch targets of a correspondingly performing 2D UI (based on area calculations from Table 2). This significantly impacts the design of S3D mobile UIs. The degradation in accuracy from the combined effects of S3D and walking appears to be larger than the sum of each individually (Fig. 6). As may be expected, the main part of this degradation is in the horizontal dimension, and hence related to stereoscopic effects. We speculate that, at least in part, this effect is due to users losing the stereoscopic effect, for example when glancing from the display when walking, which is an expected condition for real world interaction in mobile context (see [35]).

5.4 Cognitive Issues of S3D Interaction

Both tabletop and mobile studies highlighted a large variation in the performance of individuals, which, in the mobile case, was further exaggerated when the users were walking. This large variation makes it difficult to identify statistically significant generic differences between the cases examined without a very large test sample. When designing any user interface it is not wise to design for the average user, without considering the variation of user performance. In the case of S3D UIs this suggests that the use of stereoscopic depth as a standalone informative channel in the UI should be approached cautiously. Clearly, when accessibility for users with restricted capabilities is a consideration, the use of S3D, at least as an informative channel should be avoided.

5.5 Future Work

Learning effects related to long term usage were outside the scope of this study, however this would be an interesting topic for future research. Interestingly, visual examination of the distribution of presses in the mobile case (see Fig. 4) suggests a slight shift to the right whilst walking. This appears similar for both 2D and S3D cases. The reason for this is unknown and would require further study. Possible causes could be differences in the distance between the users' eyes and the screen, or differences in the viewing angle. Such small differences may prove relevant for tasks requiring the pressing of very small targets, such as on screen QWERTY keyboards used for text input, while walking.

6 Conclusions

In this chapter we reported on the evaluation of 2D touch interaction with 3D for scenes on touch-sensitive tabletops and touchscreen mobile devices with stereoscopic displays. We analyzed 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.

Tabletop. In the case of tabletops, where the scale is relatively large, we identified two separate classes of user behavior, with one group that touches the center between the projections, where the other touches the projection for the dominant or non-dominant eye. The results show a strong interaction effect between input technique and the stereoscopic parallax of virtual objects.

Mobile Device. For the smaller mobile device S3D touch screen, the mean press positions on visual targets at all z-depths were almost identical, i.e. there was no discernible offset caused by placing targets at different z-depths. However, the variance in press accuracy in the S3D case was much larger than for 2D, and this difference was even more pronounced when walking. When in static usage, to achieve the same performance as 2D, S3D touch targets need to be horizontally wider (7.6 mm for 2D vs. 9.8 mm, for 3D for 95 % presses on target).

To support usage while walking, the size of the touch targets on a touch screen S3D user interface needs to be significantly larger than the corresponding minimum for a 2D interface. In our study the minimum sizes for 95 % presses on target was 8.4 mm × 9.2 mm (width × height) for 2D vs. 12.0 mm × 10.4 mm for 3D. Although these values are based on the test device used, it is expected that they should be generally applicable to other similar S3D devices. For example, the presented minimum touch target sizes in millimeters should serve as initial guidelines for UI designers of other S3D touch screen products intended for mobile use.

The main contributions of this work are:

- We have presented the minimum target size for S3D touch targets for both large screen tabletop and small screen mobile touch screen interaction.
- We identified two separate classes of user behavior when “touching through” stereoscopically displayed objects, on larger scale S3D displays (e.g. tabletops).
- We validated that the 2D touch technique performs well for selection of objects up to about 10 cm height from the display surface.
- For the mobile device case, both walking and S3D caused a significant increase in the mean error distance.

References

1. Benko, H., Feiner, S.: Balloon selection: a multi-finger technique for accurate low-fatigue 3D selection. In: *Proceeding IEEE 3DUI*, pp. 79–86 (2007)
2. Benzeroual, K., Allison, R.S., Wilcox, L.M.: 3D Display size matters: compensating for the perceptual effects of S3D display scaling. In: *Proceedings of CVPRW 2012*, pp. 45–52. IEEE (2012)
3. Bergstrom-Lehtovirta, J., Oulasvirta, A., Brewster, S.: The effects of walking speed on target acquisition on a touchscreen interface. In: *Proceedings of MobileHCI 2011*, ACM (2011)
4. Broy, N., André, E., Schmidt, A.: Is stereoscopic 3D a better choice for information representation in the car? In: *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 93–100. ACM (2012)

5. Broy, N., Schneegass, S., Alt, F., Schmidt, A.: Framebox and mirrorbox: tools and guidelines to support designers in prototyping interfaces for 3D displays. In: Proceedings of CHI 2014, pp. 2037–2046. ACM (2014)
6. Bruder, G., Steinicke, F., Stuerzlinger, W.: Effects of visual conflicts on 3D selection task performance in stereoscopic display environments. In: Proceedings of 3DUI'10, pp. 115–118. IEEE (2013)
7. Bruder, G., Steinicke, F., Sturzlinger, W.: To touch or not to touch?: comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces. In: Proceedings of SUI 2013. ACM (2013)
8. Chan, L.W., Kao, H.S., Chen, M.Y., Lee, M.S., Hsu, J., Hung, Y.P.: Touching the void: direct-touch interaction for intangible displays. In: Proceedings CHI 2010, pp. 2625–2634. ACM (2010)
9. Colley, A., Häkkinen, J., Schöning, J., Posti, M.: Investigating mobile stereoscopic 3D touchscreen interaction. In: Proceedings of OzCHI 2013. ACM (2013)
10. Daiber, F., Speicher, M., Gehring, S., Löchtfeld, M., Krüger, A.: Interacting with 3D content on stereoscopic displays. In: Proceedings of PerDis 2014. ACM (2014)
11. Daiber, F., Li, L., Krüger, A.: Designing gestures for mobile 3D gaming. In: Proceedings of MUM 2012. ACM (2012)
12. Dodgson, N.A.: Variation and extrema of human interpupillary distance. *Stereoscopic Displays and Applications*. In: Proceedings of SPIE 2004 5291, pp. 36–46 (2004)
13. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* **47**, 381–391 (1954)
14. Grossman, T., Wigdor, D.: Going deeper: a taxonomy of 3D on the tabletop. In: Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems, 2007, TABLETOP 2007, pp. 137–144. IEEE (2007)
15. Häkkinen, J., Kytökorpi, K., Karukka, M.: Mobile stereoscopic 3D user experience - calling for user centric design. In: Proceedings of 3DCHI Workshop at CHI 2012 (2012)
16. Häkkinen, J., Posti, M., Schneegass, S., Alt, F., Gultekin, K., Schmidt, A.: Let me catch this! experiencing interactive 3D cinema through collecting content with a mobile phone. In: Proceedings of CHI 2014. ACM (2014)
17. Häkkinen, J., Posti, M., Ventä-Olkkonen, L., Koskenranta, O., Colley, A.: Mobile photo sharing through collaborative space in stereoscopic 3D. In: Proceedings of MUM 2013. ACM (2013)
18. Häkkinen, J., Posti, M., Koskenranta, O., Ventä-Olkkonen, L.: Design and evaluation of mobile phonebook application with stereoscopic 3D user interface. In: CHI 2013 Extended Abstracts on Human Factors in Computing Systems, pp. 1389–1394. ACM (2013)
19. Häkkinen, J., Kawaid, T., Takatalo, J., Mitsuya, R., Nyman, G.: What do people look at when they watch stereoscopic movies? In: Proceedings of SPIE 2010, pp. 7524 (2010)
20. Hall, A.D., Cunningham, J.B., Roache, R.P., Cox, J.W.: Factors affecting performance using touch entry systems: Tactual recognition fields and system accuracy. *J. Appl. Psychol.* **4**, 711–720 (1988)
21. Hilliges, O., Izadi, S., Wilson, A.D., Hodges, S., Garcia-Mendoza, A., Butz, A.: Interactions in the air: adding further depth to interactive tabletops. In: Proceedings of ACM UIST 2009, pp. 139–148 (2009)
22. Holtz, C., Baudisch, P.: Understanding touch. In: Proceedings of CHI 2011, pp. 2501–2510. ACM (2011)
23. Huhtala, J., Karukka, M., Salmimaa, M., Häkkinen, J.: Evaluating depth illusion as method of adding emphasis in autostereoscopic mobile displays. In: Proceedings of MobileHCI 2011. ACM (2011)

24. Jumisko-Pyykkö, S., Dominik Strohmeier, D., Utraiainen, T., Kunze, K.: Descriptive quality of experience for mobile 3D video. In: Proceedings of NordiCHI, 2010, pp. 266–275. ACM (2010)
25. Kane, S., Wobbrock, J., Smith, I.: Getting off the treadmill: evaluating walking user interfaces for mobile devices in public spaces. In: Proceedings of MHCI08, pp. 109–118 (2008)
26. Kerber, F., Lessel, P., Mauderer, M., Daiber, F., Oulasvirta, A., Krüger, A.: Is autostereoscopy useful for handheld AR? In: Proceedings of MUM 2013, ACM (2013)
27. Kommerell, G., Schmitt, C., Kromeier, M., Bach, M.: Ocular prevalence versus ocular dominance. *Vision. Res.* **2003**(43), 1397–1403 (2003)
28. Kooi, F., Toet, A.: Visual comfort of binocular and 3-D displays. *Displays* **25**(2004), 99–108 (2004)
29. Lambooi, M.T.M., Wijnand, A., IJsselsteijn, W.A., Heynderickx, I.: Visual discomfort in stereoscopic displays: a review. In: Stereoscopic Displays and Virtual Reality Systems XIV, SPIE-IS&T (2007)
30. Loomis, J.M., Knapp, J.M.: Visual perception of egocentric distance in real and virtual environments. In: Hettlinger, L.J., Haas, M.W., (eds.) *Virtual and Adaptive Environments*, volume *Virtual and adaptive environments*. Mahwah (2003)
31. Matero, J., Colley, A.: Identifying unintentional touches on handheld touch screen devices. In: Proceedings of DIS 2012, pp. 506–509. ACM (2012)
32. Mauderer, M., Conte, S., Nacenta, M.A., Vishwanath, D.: Depth perception with gaze-contingent depth of field. In: Proceedings of CHI 2014, pp. 217–226. ACM (2014)
33. Mikkola, M., Boev, A., Gotchev, A.: Relative importance of depth cues on portable autostereoscopic display. In: Proceedings of MoViD at ACM Multimedia 2010, pp. 63–68 (2010)
34. Mizobuchi, S., et al.: The effect of stereoscopic viewing in a word search task with a layered background. *J. Soc. Inform. Display* **16**(11), 1105–1113 (2008)
35. Oulasvirta, A., Tamminen, S., Roto, V., Kuorelahti, J.: Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In: Proceedings of CHI 2005. ACM (2005)
36. Parhi, P., Karlson, A.K., Bederson, B.B.: Target size study for one-handed thumb use on small touchscreen devices. In: Proceedings of MobileHCI 2006, ACM (2006)
37. Perry, K., Hourcade, J.P.: Evaluating one handed thumb tapping on mobile touchscreen devices. In: Proceedings of Graphics Interface, pp. 57–64 (2008)
38. Posti, M., Ventä-Olkkonen, L., Colley, A., Koskenranta, O., Häkkinen, J.: Exploring gesture based interaction with a layered stereoscopic 3D interface. In: Proceedings of PerDis 2014. ACM (2014)
39. Pölönen, M., Salmimaa, M., Häkkinen, J.: Effect of ambient illumination level on perceived autostereoscopic display quality and depth perception. *Displays* **32**, pp. 135–141. Elsevier (2011)
40. Schöning, J., Steinicke, F., Krüger, A., Hinrichs, K., Valkov, D.: Bimanual interaction with interscopic multi-touch surfaces. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) *INTERACT 2009*. LNCS, vol. 5727, pp. 40–53. Springer, Heidelberg (2009)
41. Sivak, B., MacKenzie, C.L.: Integration of visual information and motor output in reaching and grasping: the contributions of peripheral and central vision. *Neuropsychologia* **28**(10), 1095–1116 (1990)
42. Strothoff, S., Valkov, D., Hinrichs, K.H.: Triangle cursor: interactions with objects above the tabletop. In: Proceedings of ACM ITS 2011, pp. 111–119 (2011)

43. Sunnari, M., Arhippainen, L., Pakanen, M., Hickey, S.: Studying user experiences of autostereoscopic 3D menu on touch screen mobile device. In: Proceedings of OZCHI 2012, ACM (2012)
44. Teather, R.J., Stuerzlinger, W.: Pointing at 3D targets in a stereo head-tracked virtual environment. In: Proceedings of 3DUI 2011, pp. 87–94. IEEE Computer Society (2011)
45. Teather, R.J., Stuerzlinger, W., Pavlovych, A.: Fishtank fits: a desktop VR testbed for evaluating 3D pointing techniques. In: CHI 2014 Extended Abstracts on Human Factors in Computing Systems, pp. 519–522. ACM (2014)
46. Valkov, D., Giesler, A., Hinrichs, K.H.: Imperceptible depth shifts for touch interaction with stereoscopic objects. In: Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, pp. 227–236. ACM (2014)
47. Valkov, D., Steinicke, F., Bruder, G., Hinrichs, K.: 2D Touching of 3D stereoscopic objects. In: Proceedings of CHI 2011, pp. 1353–1362. ACM (2011)
48. Valkov, D., Steinicke, F., Bruder, G., Hinrichs, K., Schöning, J., Daiber, F., Krüger, A.: Touching floating objects in projection-based virtual reality environments. Proc. EGVE/EuroVR/VEC **2010**, 17–24 (2010)
49. Ventä-Olkkonen, L., Posti, M., Häkkinen, J.: How to use 3D in stereoscopic mobile user interfaces – study of initial user perceptions. In: Proceedings of Academic MindTrek 2013. ACM (2013)