# HoverSpace

Paul Lubos, Oscar Ariza, Gerd Bruder, Florian Daiber, Frank Steinicke,
Antonio Krüger

**HAL Id: hal-01609426**

**https://hal.inria.fr/hal-01609426**

Submitted on 3 Oct 2017

# HoverSpace:

## Analyses of the Perceived Spatial Affordances of Hover Interaction Above Tabletop Surfaces

Paul Lubos, Oscar Ariza, Gerd Bruder, Florian Daiber, Frank Steinicke and Antonio Krüger

Human Computer Interaction, Department of Informatics, Universität Hamburg, Germany
{lubos,ariza,bruder,steinicke}@informatik.uni-hamburg.de
German Research Center for Artificial Intelligence, Saarbrücken, Germany
{florian.daiber,Krueger}@dfki.de

**Abstract.** Recent developments in the area of stereoscopic displays and tracking technologies have paved the way to combine touch interaction on interactive surfaces with spatial interaction above the surface of a stereoscopic display. This holistic design space supports novel affordances and user experiences during touch interaction, but also induce challenges to the interaction design. In this paper we introduce the concept of hover interaction for such setups. Therefore, we analyze the non-visual volume above a virtual object, which is perceived as the corresponding hover space for that object. The results show that the users' perceptions of hover spaces can be categorized into two groups. Either users assume that the shape of the hover space is extruded and scaled towards their head, or along the normal vector of the interactive surface. We provide a corresponding model to determine the shapes of these hover spaces, and confirm the findings in a practical application. Finally, we discuss important implications for the development of future touch-sensitive interfaces.

**Keywords**. Hover space, touch interaction, stereoscopic displays, 3D interaction.

## 1    Introduction

Recently, the combination of two different technologies has attracted enormous attention. Several setups have been released, which combine touch-sensitive surfaces with 3D mid-air finger tracking [5]. These technologies provide direct interaction with two-dimensional (2D) or three-dimensional (3D) data sets, respectively, which is primarily leveraged in the fields of natural interaction for spatial application domains such as geo-spatial applications, architectural design, games or entertainment [6]. While multi-touch technology is available for several years, recently multiple hardware solutions from the professional as well as consumer domains have been released, which provide the means to sense hand and finger poses as well as gestures on 2D surfaces or in 3D space without the requirement to wear gloves or use other encum-

bering instrumentation (e.g., Leap Motion [13], Microsoft Kinect [17]). The combination of these technologies and the resulting expanded interaction space consisting of 2D touch input and 3D mid-air sensing provides enormous potential for novel interaction techniques.

Until recently, research on interaction techniques in the scope of tabletops and interactive surfaces have mainly been focused on (multi-)touch 2D input with monoscopically displayed data. The direct nature of multi-touch gestures and interaction including haptic feedback has great potential for natural and intuitive interaction for novice and expert users. The matching perceptual and motor space during direct touch interaction proved beneficial over less direct interaction techniques [23].

Spatial interaction above tabletop surfaces has received much attention over the last years, in particular since Hilliges et al. [12] discussed the limitation of 2D input on surfaces for natural 3D interaction and proposed interactions above the tabletop. With the advent of stereoscopic display on interactive tabletops, the interaction space has to be extended to the third dimension in order to facilitate a coherent space for input and output of such interactive systems. With stereoscopic display, objects can appear detached from the display surface, i.e., in front of or behind the display surface. Such situations induce challenges for natural touch interaction due to missing haptic feedback when interacting with stereoscopically displayed floating objects (cf. "touching the void" [6]). Schöning et al. [19] considered general challenges of multi-touch interaction with stereoscopically rendered projections and conclude that most of the existing interaction techniques have in common that the interaction and visualization is limited to a region close to zero parallax (i.e., the interactive surface) [5,21].

While the described setups provide interesting challenges to the interaction with stereoscopically displayed 3D objects on a touch surface, it is often not clear for users with which objects they can interact, i.e., 3D stereoscopically objects often miss the affordance of touch [5,6]. In mouse-based interaction setups, such affordances are often presented by hover effects. Hence, it sounds reasonable to transfer this concept to tabletop setups. However, with touch-based interaction such hover effects are difficult to implement since a hover movement on the surface already induces a touch event. Such hover interaction has been successfully applied for monoscopic displays to support multi-touch tabletops with contextual information [12]. However, we are not aware of existing solutions considering hover interaction for stereoscopic multi-touch environments.

In this paper we focus on hover interaction, which does not require users to touch an object in 2D or 3D spaces, e.g., by moving a finger inside the object or on the surface, but is rather based on hovering "over" the object with a finger or hand relative to the considered object. However, so far it is not clear how users perceive affordances of hover spaces above the interactive surfaces, especially, if objects are displayed stereoscopically. In particular, it is not clear which shapes and sizes of volumes match the perceived affordances of hover interaction. For these reasons, we determine a perceptually-inspired model for volumes used for hover interaction, which we call the *HoverSpace*. We evaluate the model and compare it with a naive approach in a confirmatory study. The results of these experiments provide guidelines for interactive applications using hover gestures in tabletop setups.

In summary, our contributions are:

- An analysis of above-surface volumes for hover interaction in tabletop setups,
- a usability comparison of perceptually-inspired and naive hover volumes, and
- guidelines for designing hover interaction in touch-sensitive tabletop setups.

The remainder of this paper is structured as follows. Section 2 presents an overview of related work on hover and above-surface interaction in tabletop setups. Section 3 describes the experiment in which we analyze perceived spatial affordances of hover interaction. In Section 4 we derive the perceptually-inspired HoverSpace. Section 5 validates the results in a confirmatory experiment. Section 6 provides a general discussion of the results and guidelines for hover interaction in tabletop setups. Section 7 concludes the paper.

## 2    Related Work

In this section we provide an overview of related work on hover interaction in 2D and 3D user interfaces as well as mid-air interaction on 3D stereoscopic touch surfaces.

### 2.1    2D Hover

As Buxton describes in his three state model of graphical input, for traditional input devices, like a mouse, there is a so-called tracking state, as in the state where the cursor can be moved without pressing a button. The position and movement of the cursor can be directly transferred to input in user interfaces [7]. One of the most common uses of such a tracking state is for so-called hover effects in classic 2D user interfaces [7]. The tracking state is often used for highlighting or tooltips and can declutter interfaces by providing context sensitive information. In most touch interfaces, for example using capacitive sensing technology, this tracking state is missing [9,18]. While dragging is possible, when pressing down on the screen, the hovering usually cannot be detected. Especially during the design of touch-enabled tabletop and mobile user interfaces, such as mobile versions of websites, the missing tracking state becomes obvious and many design principles, such as flat design, become complicated [9].

To compensate for the missing hover capability in 2D multi-touch setups Benko et al. [2] simulated a hover state with techniques that make use of a secondary finger which adjusts the control-display ratio while the primary finger controls the movement of the cursor, thus resulting in more precise selection.

Beyond hover interaction in multi-touch setups other input modalities have been investigated as well. Grossman et al. [9] presented Hover Widgets, which extend the expressiveness of pen-operated touch surfaces by using the tracking state of the pen as hover input. In particular, they proposed a special hover technique that activates a widget by a short discrete gesture that is followed by a pen-down action. The Hover Widget technique shows that the space between the hover state and the touch state can be effectively used.

## 2.2 3D Hover

With recent advances in 3D sensing technologies it becomes possible to track a user's fingers above touch surfaces which allows for a tracking state to be leveraged for user interfaces [3,12]. Different interaction techniques have been proposed for monoscopic and stereoscopic display environments which make use of this capability.

Han and Park [10] explored hover-based zoom interactions in monoscopic display environments. They proposed a technique that relies on a magnifying lens metaphor. This approach allows users to quickly zoom in and out in a restricted range of multiple zoom levels that are defined by layers above a multi-touch display. Initial evaluation results revealed that their technique outperforms the common pinch-to-zoom technique in both speed and user preference. However, in their implementation zoom layers were discrete and no continuous zooming in a 3D hover volume above the multi-touch display was possible.

Echtler et al. [8] presented a multi-touch tabletop that was extended with a ceiling-mounted light source to create shadows of hands and arms. By tracking these shadows with the rear-mounted camera of their frustrated total internal reflection (FTIR) setup they augmented the multi-touch tabletop with mouse-like hover behavior. With this setup users can control multiple cursors by hovering above the tabletop and trigger a "click" event when touching the surface. They evaluated their system with respect to tracking accuracy, which indicated that users were aware of their hand position above the display and tried to avoid occlusion by orienting their hand in an unnatural pose parallel to the edge of the tabletop so that the cursor pointed perpendicular to the user's viewing direction.

Annett et al. [1] presented Medusa, a proximity-aware multi-touch tabletop that is capable of tracking multiple users and differentiate between their hands. Besides supporting collaborative multi-user settings, they proposed different hand-dependent hover techniques. Hovering with the right hand above the display triggers a marker below the hand that turns into a component-specific marking menu when it is touched. Hovering with the left hand displays an 'X' icon which deletes a component when it is touched.

The prototype by Pyryeskin et al. [18] uses light reflected from a person's palm to estimate its position in 3D space above the table based on the diffused surface illumination vision-based principle.

## 2.3 Mid-Air Interaction

The space above interactive multi-touch surfaces has been considered for different 2.5D and 3D user interfaces with stereoscopic display, in particular using direct touch for objects displayed on the surface as well as 3D mid-air touch for objects that are displayed with negative parallax above the surface. Bruder et al. [6] found that users tend to incorrectly perceive the 3D position of stereoscopically displayed objects with negative parallax when touching these objects by moving their finger inside their perceived 3D shape. Lubos et al. [16] showed that 3D selection performance can be greatly increased by extending the selection volume using an ellipsoid shape that is

oriented towards the user's head position to account for these perceptual differences. These results are in line with results from perceptual psychology which suggest that users observing stereoscopically displayed scenes often tend to underestimate or overestimate ego-centric distances and incorrectly judge spatial relations due to visual conflicts such as occlusion or an accommodation-convergence mismatch [4,15,22]. Bruder et al. [5] investigated the precision and performance of 2D touch selection of stereoscopically displayed objects in comparison to 3D mid-air selection. They showed that touching on the interactive surface outperforms mid-air touch selection in a Fitts' Law experiment if the object is projected close to the surface with a distance of up to approximately 10 cm. For objects displayed farther away from the surface than this threshold 3D mid-air selection results in much higher performance in comparison to touch. In line with guidelines proposed by Schöning et al. [19] this underlines that multi-touch interaction with stereoscopically rendered objects is mainly limited to a distance of about 10 cm from the plane of the interactive surface, which we also consider to be an indicator for the typical hover space above touch-sensitive tabletops.

## 3    Perceptual Experiment

In this section we describe the experiment in which we analyzed the perceived affordances of hover interaction in terms of the 3D volume above rendered objects of different sizes and shapes in an interactive tabletop setup with stereoscopic display.

Based on the previous work described in Section 2, we explored the following expectations in this experiment: The hover space where users expect hovering effects to occur may not necessarily be oriented vertically above the object, but rather influenced by a user's head position relative to the target. Moreover, we assumed the shape of the hover volume to be influenced by the shape of the target object.

### 3.1    Participants

We recruited 15 participants for our experiment (11 male, 4 female), all of whom were students or professionals from the field of human-computer interaction or computer science (ages 24—54, M=34.7, SD=8.58, heights 1.55 m—1.92 m, M=1.76 m, SD=.12 m). The students received class credit for their participation. Two of the participants were left handed, the remaining 13 participants were right handed. All participants had normal or corrected vision.

Using the technique proposed by Willemsen et al. [22] we measured the interpupillary distance (IPD) of each participant before the experiment (M=6.54 cm, SD=.35 cm) and used it to calibrate the rendering for each participant.

Only one participant reported no experience with stereoscopic display and ten participants reported high or very high experience (rating scale 0=no experience, 4=very high experience, M=2.80, SD=1.21). Only one participant reported no experience with 3D computer games while ten participants reported high or very high levels of experience (rating scale 0=no experience, 4=very high experience, M=2.87,

SD=1.36). The mean total time per participant, including questionnaires and instructions was about 35 minutes. The mean time for performing the actual experiment was about 25 minutes. Participants were allowed to take breaks at any time.

## 3.2 Material

As illustrated in Figure 1, participants were instructed to stand at a stereoscopic multitouch table in an upright position facing the table. A Razer Tartarus keypad was adjusted to a comfortable height for the non-dominant hand of the participant. Participants were instructed to keep their hand at that position during the experiment to confirm their selections.

The experiment was conducted with the participant wearing Samsung SSG-P51002 radio frequency active shutter glasses, a cap with an infrared (IR) marker and a glove with an IR marker at the fingertip of their dominant hand. The markers were tracked with an optical WorldViz Precision Position Tracking (PPT X4) system with submillimeter precision for view-dependent rendering and finger tracking.

The visual stimulus displayed during the experiment showed a 3D scene, which was rendered with the Unity3D Pro engine [20] with an Intel computer with a Core i7 3.4 GHz CPU and an NVidia GeForce GTX780TI.

The scene was displayed stereoscopically on a Samsung UE55F9000 TV in a height-adjustable, stereoscopic tabletop setup. The scene showed a gray brushed metal surface at the zero parallax plane and targets in a red color. For each trial, a single target was visible, either a sphere or a cube. Those shapes were chosen as they are approximations of objects typically found in user interfaces and compound objects consisting of these shapes could approximate almost any other shape.

## 3.3 Methods

We used a $2 \times 2 \times 6 \times 4$ design with the method of constant stimuli for the experiment trials. The two target shapes (cube, sphere), two target sizes (5 cm, 10 cm), six target positions (P0=(−0.2, 0, 0), P1=(0, 0, 0), P2=(0.2, 0, 0), P3=(−0.2, 0, −0.2), P4=(0, 0, −0.2) and P5=(0.2, 0, −0.2)) and four repetitions were uniformly and randomly distributed between all 96 trials for each participant.

Each trial consisted of a single shape of one size at one of the positions being shown to the participant (see Figure 2). The participants were instructed to think of hovering in traditional 2D user interfaces and indicate the volume where they would expect a hover effect to be triggered by moving the index finger of their dominant hand in that volume, i.e., by "drawing" the volume. During each trial, the participants indicated the hover volume for ten seconds while pressing a button on the keypad with their non-dominant hand. The non-dominant hand was chosen for this task to avoid any jittering while indicating the volume, which may be induced by pressing buttons attached to a glove on the dominant hand.

Li et al. [14] have shown an increase in performance from using the non-dominant hand for such tasks. After ten seconds the next trial started. We recorded tracking data at 30 Hz while the participant pressed the button. Each recording consisted of the participant's head and finger position.
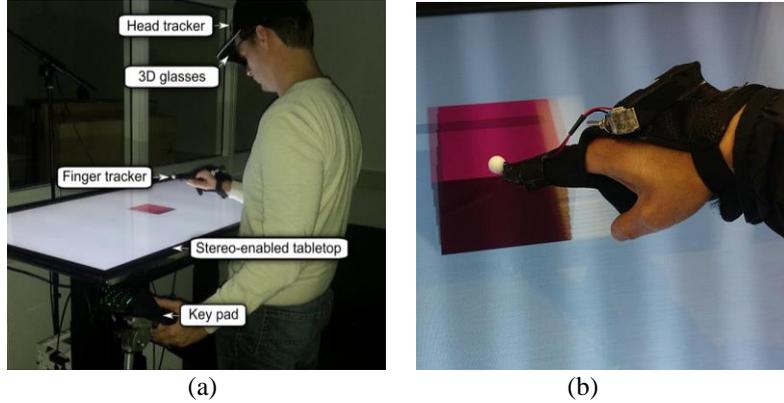
(a)                                              (b)

**Fig. 1.** (a) Participant during the experiment and (b) close-up of the participant's hand while indicating the hover volume with the tip of the index finger. The scene was displayed stereo-scopically. IR markers on the head and index finger of the dominant hand were tracked. The non-dominant hand rested on a keypad.
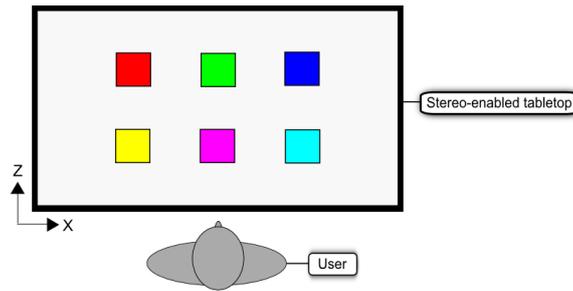


**Fig. 2.** Positions used to locate the objects on the tabletop surface. P0=(−0.2, 0, 0) (yellow), P1=(0, 0, 0) (magenta), P2=(0.2, 0, 0) (cyan), P3=(−0.2, 0, −0.2) (red), P4=(0, 0, −0.2) (green) and P5=(0.2, 0, −0.2) (blue). These colors are only used in the following result plots.

The participants completed training trials before the main experimental phase to ensure that they understood the task correctly. The training trials differed from the main trials by showing the participants the volume which they drew to help them understand the task, while this visual feedback was excluded in the main trials, i.e., participants only saw their real hand and the virtual 3D object so as not to bias the results due to cluttering the virtual scene over time. The training trials were excluded from the analysis.

### 3.4    Results

We had to exclude three participants from the analysis, as they misunderstood the task and touched the 2D surface throughout the experiment instead of indicating a 3D hover volume, which was confirmed during debriefing after the experiment.

Since we had four repetitions for each condition of the experiment, we pooled the results over the repetitions. We normalized the tracking data to account for the varying head positions within and between participants by normalizing head positions in target-centered coordinates (see Figure 3). We visually analyzed the resulting coordinates and observed two main behavior patterns:

- *Orthogonal Hovering*: The first behavior pattern was characterized by seven participants indicating the hover space at the surface of the object or above it. As shown in Figures 3(a) and (b), the horizontal width and depth of this volume increased with increasing distance to the tabletop in vertical direction.
- *Line-of-Sight Hovering*: The second behavior pattern was shown by five participants in which the hover space was tilted towards line-of-sight, i.e., the volume extended in the direction of the participant's head position instead of orthogonally from the tabletop surface as shown in Figures 3(c) and (d).

The observed behavior patterns were consistent for each participant throughout the experiment, i.e., we did not observe any participant changing the behavior during the experiment.
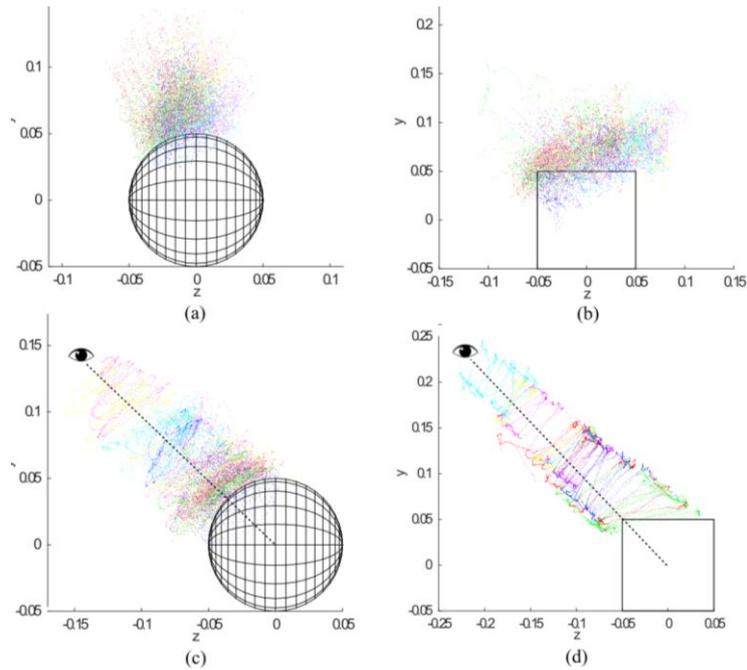


**Fig. 3.** Examples of the two behavior patterns in 2D coordinates with the y-axis indicating the up-direction from the tabletop at y=0 and the z-axis indicating the direction from the participant towards the opposite side of the tabletop. (a, b) Orthogonal Hovering indicates a hover volume at the surface of the object or above it. (c, d) Line-of-Sight Hovering shows a hover volume within line-of-sight that converges towards the participant's head position. The colors represent the different tested positions of the objects on the tabletop (see Figure 2).
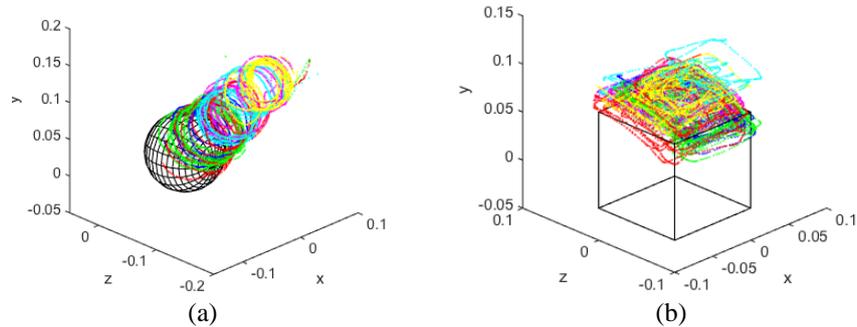
**Fig. 4.** Illustrations of the drawn hover volumes of participants for the (a) sphere target shape and the (b) cube target shape. The y-axis indicates the direction orthogonally to the tabletop surface at y=0, the z-axis increases towards the opposite end of the tabletop, and the x-axis is oriented laterally to the right of the tabletop. The drawing patterns of outlining the circular or rectangular regions and then filling the regions with zigzag patterns could be observed for many participants.

We observed a difference between the cube target shape and the sphere target shape. For the sphere, all participants indicated a round hover space, often drawing circles at various distances from the object (see Figure 4(a)). Conversely, for the cube shape, all participants indicated a rectangular hover space, often drawing a rectangular outline and then used zigzag pattern to fill the area (see Figure 4(b)). The comments of the participants during debriefing also reflect this behavior.

### 3.5     Discussion

We instructed our participants to think of hovering in traditional 2D user interfaces and indicate the 3D-volume with their finger where they would expect a hover effect to be triggered. Considering this instruction, these behaviors could be explained as follows. When hovering in 2D Desktop environments, hover effects are triggered when the mouse cursor is (a) above a target object and (b) occludes the target object. Depending on the participant's understanding of 2D hovering, this has different effects on the behavior with a third, height dimension.

Indeed, the results show two groups of participants with distinct behavior patterns (Orthogonal Hovering vs. Line-of-Sight Hovering). We observed no changes in behavior for any participant during the experiment, which suggests that there are two mental models of where users expect a hover volume to be above interactive tabletops. While seven of our participants would expect a hover space to be located above a virtual object orthogonally to the display surface, five of our participants would expect the hover space to be located along the line-of-sight from the object to their head position, suggesting that they interpreted hovering as occlusion of the target object.

Additionally, we found two different hover volumes for the sphere and cube target shapes, which shows that the object's shape determines the shape of the hover volume as well. Round target shapes imply round hover volumes and angular shapes imply angular hover volumes.
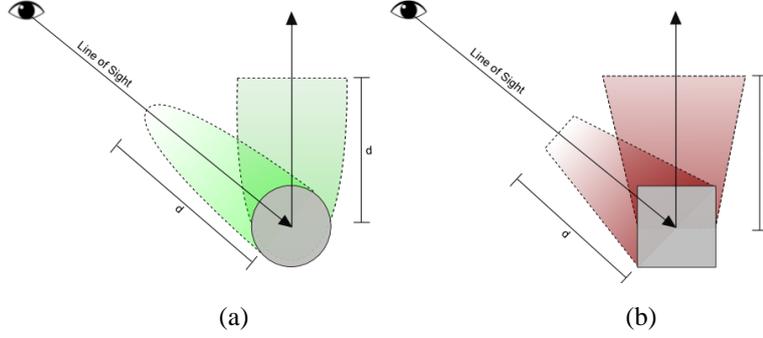
(a)          (b)

**Fig. 5.** Illustrations of the HoverSpace volumes in vertical direction and along line-of-sight for the two shapes: (a) For round shapes the volumes were approximated with paraboloids, and (b) truncated pyramids were used for rectangular shapes. We used d=10 cm according to results presented by Bruder et al. [5].

## 4 The HoverSpace

With the results from the perceptual experiment we defined two main volumes, where participants expect hovering effects. In the following, we define a combined hover volume which we call HoverSpace. Since the volumes depended on whether the target shape was rectangular or rounded, we defined two formulas for the HoverSpace, which allow easy testing whether the tracked input object, such as the user's fingertip, is within the hover volume. The formulas are written for a left-handed Cartesian co-ordinate system, where the y-axis corresponds to the up-direction. Let $(x,y,z) \in \mathbb{R}^3$ be the finger position in 3D coordinates centered around the target object. Let $a \in \mathbb{R}^+$ be the scale of the target object on the x-axis, $b \in \mathbb{R}^+$ the scale on the z-axis, $c \in \mathbb{R}^+$ the scale on the y-axis and $d \in \mathbb{R}^+$ an empirically determined value defining the spread of the hover region.

Since the volumes depended on whether the target shape was rectangular or round-ed, we defined two formulas for the HoverSpace, which allow easy testing whether the tracked input object, such as the user's fingertip, is within the hover volume. The formulas are written for a left-handed Cartesian coordinate system, where the y-axis corresponds to the up-direction. Let $(x,y,z) \in \mathbb{R}^3$ be the finger position in 3D coordi-nates centered around the target object. Let $a \in \mathbb{R}^+$ be the scale of the target object on the x-axis, $b \in \mathbb{R}^+$ the scale on the z-axis, $c \in \mathbb{R}^+$ the scale on the y-axis and $d \in \mathbb{R}^+$ an empirically determined value defining the spread of the hover region.

The HoverSpace is based on two formulas. For round shapes a paraboloid can be used which can be approximated by the following formula:

$$\frac{x^2}{a^2} + \frac{z^2}{b^2} - \frac{y}{d} \leq 0 \,\&\&\, 0 \leq y \leq 10 \; cm$$

For rectangular shapes, the results can be approximated by a truncated pyramid and the following formula:

$$hf = \frac{y}{c}, x_{max} = lerp(a, d * a, hf), y_{max} = lerp(a, d * b, hf)$$
$$0 \leq y \leq 10 \; cm \,\&\&\, x - x_{max} \leq 0 \,\&\&\, y_{max} \leq 0$$

### 4.1 Orthogonal Hovering

From the results of our participants who expected the hover volume to be located above the target object we determined a volume enclosing 95% of the finger positions. The volume is oriented upwards from the target object, but the volume also expands in width and depth the higher the participant's finger was from the tabletop surface. Depending on whether the target shape was rectangular or round, we found the 95% volume to follow a mathematical function. The region above the object is illustrated in Figure 5.

For the orthogonal region, the origin of the coordinate system to transform the tracking coordinate of the finger position into for both formulas is given by the center of the object and the up axis along the display normal.

### 4.2 Line-of-Sight Hovering

From our participants expecting the hovering effects to occur when occluding the object along their line-of-sight, we found that the width and depth of the volume increase in size as it gets closer to the object until it covers the size of the object. The volume enclosing 95% of finger positions is illustrated in Figure 5. In contrast to the hover volume defined in Section 4.1 the size of this volume depends not only on the size of the object, but also on the distance of the head from the object. For rectangular and round shapes we found different volumes.

For line-of-sight hovering the same formulas can be used as for the orthogonal hovering. Here, the origin of the coordinate system is on the line-of-sight between the center of the object and the head position. The up axis is given by the line-of-sight. After transformation of the finger position from tracking coordinates into the line-of-sight coordinate system, the formulas can be applied.

## 5 Confirmatory Experiment

In this section we describe the experiment in which we compared the HoverSpace with a naive straight-up infinitely extruded outline (called Extruded in the following) approach of 3D hover volumes.

### 5.1 Participants

We recruited 16 participants for our experiment (11 male, 5 female), all of whom were students or professionals from the field of human-computer interaction or computer science (ages 19—36, M=27.37, SD=4.72, heights 1.60 m—1.93 m, M=1.78 m, SD=.10 m). Six participants already participated in the first experiment. The students received class credit for their participation. Two participants were left-handed, the remaining 14 participants were right-handed. All participants had normal or corrected vision.

We measured the IPD of each participant before the experiment started (M=6.61 cm, SD=.29 cm). We calibrated the system accordingly for each participant.
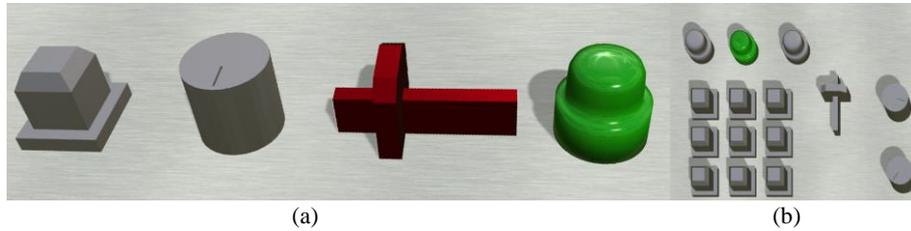
**Fig. 6.** (a) Illustration of the used target shapes and the colors in the confirmatory experiment. From left to right: rectangular button and round knob in grey, red rectangular slider, green round button. (b) Illustration of an example interface.

All participants reported at least some experience with stereoscopic display and ten participants reported high or very high experience (rating scale 0=no experience, 4= very high experience, M=2.86, SD=1.09). Three participants reported no experience with 3D computer games while ten participants reported high or very high levels of experience (rating scale 0=no experience, 4=very high experience, M=2.63, SD=1.59).

The mean total time per participant, including questionnaires and instructions was 20 minutes. The mean time for performing the actual experiment was about 15 minutes. Participants were allowed to take breaks between the conditions.

### 5.2 Material

The setup in the confirmatory experiment was the same as in the experiment reported in Section 3. The setups differed only in the visual representation. The scene showed a gray brushed metal surface at the zero parallax plane and targets were shown in a gray color. For each trial, six target objects were visible, either a round button, a round knob, a slider or a rectangular button, as illustrated in Figure 6. Those shapes were chosen as they represent objects that we often find in practical applications for tangible user interfaces. When a participant reached with the index finger within the hover volume of an object during the experiment, the object either highlighted red (interpreted as the incorrect target) or green (interpreted as the correct target).

### 5.3 Methods

We used a $2 \times 4 \times 2 \times 6$ design with the method of constant stimuli for the experiment. We considered two hover volumes: HoverSpace vs. Extruded. The four target shapes (round button or knob, rectangular slider or button), two target sizes (2.5 cm or 5 cm), and six repetitions were uniformly and randomly distributed between all 48 trials in each hover condition for each participant. The hover volume condition was counterbalanced between participants, i.e., half the participants started with the Extruded condition and the other half with the HoverSpace condition.

Each trial consisted of six instances of the same shape of one size at different positions being shown to the participant. The positions were arranged in a grid representing the participant's interaction space, such that they were able to reach them comfortably. The sizes of the objects, at 2.5 cm diameter and 5 cm diameter resemble the

typical button size for a display of that size, and double the size, to allow a comparison between different sizes, respectively. The grid was spaced so that the objects did not overlap at any time. The participants were instructed to find the green object by hovering over it with the index finger of their dominant hand and press a button on the keypad with their non-dominant hand when they were within the hover volume.

During the experiment, the participants saw their real hand and the six virtual objects. In the HoverSpace condition, we used the functions described in Section 4 and determined whether the participant's finger was in an object's hover volume. In the Extruded condition, we set the y-coordinate of the finger to the target object's height and used the corresponding equation for a 2D approximation of the object's outline. For the round objects we used the equation for an ellipse and for the rectangular condition the equation for a rectangle. This effectively creates an infinite volume, which means that the Extruded condition extended higher compared to the HoverSpace and its height of 10 cm (see Section 4). The objects were colored grey when the user's finger was outside the hover volume. When the finger was inside an incorrect target's hover volume, the object turned red. When the finger was in the correct target's hover volume, the object turned green as shown in Figure 6. After pressing the button the next trial began.

All participants were instructed to complete the task as fast and as precise as possible. The first dependent variable was the selection time, i.e., the time from the start of the trial until the participant confirmed the selection with the press of a button. The second dependent variable was the error rate, i.e., the amount of times the participant pressed the button without being within the correct hover volume.

The participants completed supervised training trials before the experimental phase to ensure that they understood the task correctly. The trials differed from the actual trials in that they allowed the participants to familiarize themselves with the two different hovering conditions. The training trials were excluded from the analysis.

**Subjective Questionnaires.**
To collect subjective impressions we utilized a comparative AttrakDiff questionnaire, which measures hedonic quality and attractiveness [11]. Following an initial demographic questionnaire, after half the trials, the participants had to take a break, answer an AttrakDiff questionnaire and then continue with the other condition. After the second condition, they filled in the second part of the AttrakDiff questionnaire and a further questionnaire directly asking them to judge which technique they preferred and why they chose that technique.

**Hypotheses.**
Based on the results of the perceptually-inspired experiment discussed in Section 3, we evaluated the following hypotheses:

- H1: For the HoverSpace condition the mean selection time is lower than for the Extruded hover volume.
- H2: For the HoverSpace condition the mean error rate is lower than for the Extruded hover volume.
- H3: The participants prefer the HoverSpace over the Extruded hover volume.

### 5.4 Results

In the following section we summarize the results of the confirmatory experiment. We analyzed the results with a repeated measure ANOVA at the 5% significance level. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly's test indicated that the assumption of sphericity had been violated. Since we found no difference between the results for the two round shapes, nor between the two rectangular shapes, we pooled the data.

**Selection Time.**
The results for the selection time are shown in Figure 7(a). The results show that the selection time differs significantly between the Extruded (M=3.09, SD=1.31) and the HoverSpace (M=2.90, SD=1.35) conditions (F(1,15)=7.955, p<.05, $\eta_p^2$=.347). As expected, we found a significant influence of the target scale on the selection time (F(1,15)=16.294, p<.001, $\eta_p^2$=.521). We did not find a significant influence of the target shape on the selection performance (F(3,45)=2.46, p=.075, $\eta_p^2$=.14). We found a significant interaction effect between the hover condition and the round condition (F(1,15)=9.256, p<.05, $\eta_p^2$=.382). Post-hoc tests with Bonferroni correction for the interaction effect between the hover and the round condition showed significant differences only between the Extruded-rectangular (M=3.11, SD=.61) and HoverSpace-round (M=2.72, SD=.53) conditions (T(15)=3.164, p<.05), between the Extruded-round (M=3.07, SD=.62) and HoverSpace-round conditions (T(15)=4.229, p<.05), and between the HoverSpace-rectangular (M=3.08, SD=.31) and HoverSpace-round conditions (T(15)=3.240, p<.05).

**Errors.**
The results for the error rate are shown in Figure 7(b). The results showed no significant difference in error rate between the Extruded (M=.06, SD=.24) and the HoverSpace (M=.05, SD=.21) conditions. We found a significant influence of the round condition on the errors (F(1,15)=10.392, p<.05, $\eta_p^2$=.409).
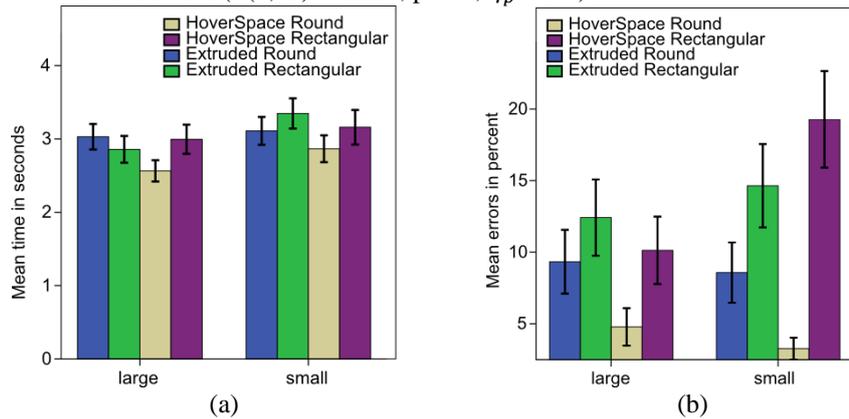


**Fig. 7.** Plots of the pooled results of the confirmatory experiment. The x-axes show the target scales and the y-axes show the (a) mean time in seconds and (b) the mean errors in percent. The bar plots are grouped by the hover condition and the round condition. The vertical bars show the standard error.

We found no significant influence of the hover condition ($F(1,15)=.775$, $p=.392$, $\eta_p^2=.049$) or the scale ($F(1,15)=3.629$, $p=.076$, $\eta_p^2=.195$). We found a significant interaction effect between the scale and the round condition ($F(1,15)=7.304$, $p<.05$, $\eta_p^2=.327$). Post-hoc tests with Bonferroni correction for the interaction effect between the scale and the round condition showed significant differences only between the big-round (M=.03, SD=.06) and small-rectangular (M=.10, SD=.12) conditions (T(15)=3.264, $p<.05$), as well as between the small-rectangular (M=.10, SD=.12) and small-round (M=.02, SD=.04) conditions (T(15)=−3.178, $p<.05$).

### Selection Distribution.

We evaluated how often participants selected an object in the HoverSpace condition while being with their index finger in the Orthogonal or Line-of-Sight volumes (see Section 4). The distribution was approximately 23% only in the Orthogonal volume, 7% only in the Line-of-Sight volume, and 65% in the overlap region of both volumes. Approximately 5% of all selections were errors. In the Extruded condition, approximately 94% were within the vertically extruded region and approximately 6% of all selections were errors.

### Subjective Questionnaires.

The results of the AttrakDiff questionnaire show that pragmatic quality, i.e., an indication of whether the user is assisted by the product, reaches an average value overall. In comparison, pragmatic and hedonic qualities of the HoverSpace are higher than of the Extruded hover volume. The HoverSpace also has a smaller confidence interval for PQ and HQ, indicating a greater level of certainty on the users. In terms of the overall means, the HoverSpace approach is located in the above-average region with an overall impression of the approach as attractive.

We asked the participants which of the techniques they preferred, either the first or the second one they tried. As the experiment was counterbalanced, the results were mapped towards the HoverSpace or Extruded condition. The results show a preference of the HoverSpace (rating scale 1=Extruded, 5=HoverSpace, M=3.44, SD=1.55).



**Fig. 8.** Average values and confidence rectangles for the AttrakDiff questionnaire of the two conditions: **A** for the Extruded approach and **B** for the HoverSpace.

### 5.5 Discussion

The results of the confirmatory experiment showed that the HoverSpace outperformed the Extruded approach considering the selection time, which implies a higher overall performance, considering that we did not find a significant difference between the similar error rates. This confirms our hypothesis H1, but not H2. Considering the limited size of the HoverSpace, albeit wider than the infinite Extruded approach, the lower selection time implies that the perceptually-inspired hover volume is a valuable improvement over vertically extruded hover regions.

The results from the subjective questionnaires support our hypothesis H2, as they show that the participants subjectively preferred the perceptually-inspired HoverSpace over the Extruded approach. We received multiple comments such as "The second technique [HoverSpace] was much more intuitive and more precise compared to the first one." These comments further support this hypothesis. However, some participants thought the Extruded approach was more precise and preferred it over the HoverSpace. This might be caused by the fact that the HoverSpace volumes for different target objects overlapped when they were located close together, causing two targets to change color at the same time. To disambiguate such multiple selections in future implementations we suggest to prioritize selections in Orthogonal hover volumes compared to selections in Line-of-Sight hover volumes considering the larger number of participants in the experiment described in Section 3 whose mental model matched these hover volumes.

## 6  Guidelines

Hovering in interactive tabletop environments allows effective decluttering of interfaces. Our analysis of the perceived spatial affordances of such hover interaction has shown that perceptually-inspired hover volumes can increase the performance, as well as the subjective attractiveness of interfaces. In the following, we summarize the lessons learned:

We observed two mental models for hovering in our stereoscopic tabletop environment. The first mental model was characterized by users expecting hovering effects to occur when their hand is right above the object on a line along the display normal, and the second is characterized by users occluding the object based on their line-of-sight. The orthogonal hovering is relatively close to the naive, straight-up extruded outline solution usually implemented in related work.

We suggest a combination of both of these approaches to provide a technique valid for most users. However, our results in the confirmatory experiment suggest that only seven percent of the selections were in the line-of-sight volume. This leads to the conclusion that the line-of-sight volume should be provided when available, but could be left out when head-tracking is not available, e.g., in tabletop setups with monoscopic display.

We suggest the following guidelines for hovering in tabletop environments:

- G1: A combination of an orthogonal region and a line-of-sight region provides the best performance for hovering tasks.
- G2: Without head-tracking, using an orthogonal region with increasing width depending on the height from the object provides acceptable performance.

## 7 Conclusion

Due to recent technological advances, the combination of touch interaction on interactive surfaces with spatial interaction above the surface of a stereoscopic display has become feasible. In this paper we identified a way to improve the interaction in this holistic design space by conducting a perceptual study and evaluating the results to define a perceptually-inspired hover volume called the HoverSpace. We confirmed the advantages of this HoverSpace in an experiment and found a significant improvement in performance compared to the traditional approach. Finally, we discussed guidelines for the development of future touch-sensitive interfaces.

In the perceptual experiment we identified two mental models that users exhibit for hovering in stereoscopic 3D environments, which are grounded in the different interpretations of 2D hovering as bringing a mouse cursor over an object or occluding an object with a cursor, respectively. Our results show that both interpretations have direct implications for the design of hover interaction in the 3D space above interactive tabletops.

Future research may investigate whether there are similar differences in other 3D interaction techniques derived from 2D Desktop interfaces. Additionally, future work may focus on the impact of the different parameters of the HoverSpace and determine the best possible values with the smallest necessary volume to reduce overlapping HoverSpace volumes. In particular, the difference between the round and rectangular target shapes could be investigated further to determine whether a mean between these two shapes could improve the HoverSpace. Furthermore, an investigation of training effects may show how long users need to adapt to different types of hover volumes.

## References:

1. Annett, M., Grossman, T., Wigdor, D., Fitzmaurice, G.: Medusa: a proximity-aware multi-touch tabletop. In: Proc. of ACM UIST. pp. 337--346. ACM (2011)
2. Benko, H., Wilson, A.D., Baudisch, P.: Precise selection techniques for multi-touch screens. In: Proc. of SIGCHI CHI. pp. 1263--1272. ACM, New York, NY, USA (2006)
3. Bernstein, J.T., Amm, D.T., Leung, O., Mullens, C.T., King, B.M., Land, B.R., Cutler, R.T.: Touch and hover sensing (Jul~10 2009), uS Patent App. 12/501,382
4. Bruder, G., Steinicke, F., Stuerzlinger, W.: Effects of visual conflicts on 3D selection task performance in stereoscopic display environments. In: Proc. of ACM 3DUI. pp. 115--118 (2013)

5. Bruder, G., Steinicke, F., Stuerzlinger, W.: To touch or not to touch? Comparing 2d touch and 3d mid-air interaction on stereoscopic tabletop surfaces. In: Proc. of ACM SUI. pp. 1--8. ACM Press (2013)

6. Bruder, G., Steinicke, F., Stuerzlinger, W.: Touching the Void Revisited: Analyses of Touch Behavior On and Above Tabletop Surfaces. Lecture Notes in Computer Science: Human-Computer Interaction - INTERACT 2013 8117, 278--296 (2013)

7. Buxton, W.: A three-state model of graphical input. In: Human-computer interaction-INTERACT. vol.~90, pp. 449--456. Citeseer (1990)

8. Echtler, F., Huber, M., Klinker, G.: Shadow tracking on multi-touch tables. In: Proc. of AVI. pp. 388--391. ACM, New York, NY, USA (2008)

9. Grossman, T., Hinckley, K., Baudisch, P., Agrawala, M., Balakrishnan, R.: Hover widgets: Using the tracking state to extend the capabilities of pen-operated devices. In: Proc. of SIGCHI CHI. pp. 861--870. ACM, New York, NY, USA (2006)

10. Han, S., Park, J.: A study on touch & hover based interaction for zooming. In: CHI'12 Extended Abstracts on Human Factors in Computing Systems. pp. 2183--2188. ACM (2012)

11. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Proc. of Mensch & Computer, pp. 187--196. Springer (2003)

12. 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: UIST '09: Proc. of ACM UIST. pp. 139--148. ACM (2009)

13. Leap Motion Inc: Leap Motion. http://www.leapmotion.com/ (2012)

14. Li, Y., Hinckley, K., Guan, Z., Landay, J.A.: Experimental analysis of mode switching techniques in pen-based user interfaces. In: Proc. of the SIGCHI CHI. pp. 461--470. ACM (2005)

15. Loomis, J., Knapp, J.: Visual Perception of Egocentric Distance in Real and Virtual Environments. In: Hettinger, L., Haas, M. (eds.) Virtual and Adaptive Environments, pp. 21--46. Erlbaum (2003)

16. Lubos, P., Bruder, G., Steinicke, F.: Analysis of direct selection in head-mounted display environments. In: Proc. of IEEE 3DUI. pp. 1--8. IEEE (2014)

17. Microsoft: Kinect. http://www.microsoft.com/en-us/kinectforwindows/ (2012)

18. Pyryeskin, D., Hancock, M., Hoey, J.: Extending interactions into hoverspace using reflected light. In: Proc. of the ACM International Conference on Interactive Tabletops and Surfaces. pp. 262--263. ACM (2011)

19. Schöning, J., Steinicke, F., Valkov, D., Krüger, A., Hinrichs, K.H.: Bimanual Interaction with Interscopic Multi-Touch Surfaces. In: INTERACT '09. Springer (2009)

20. Unity Technologies: Unity3D. http://www.unity3d.com/ (2015)

21. Valkov, D., Giesler, A., Hinrichs, K.H.: Imperceptible depth shifts for touch interaction with stereoscopic objects. In: Proc. of SIGCHI CHI. pp. 227--236. ACM (2014)

22. Willemsen, P., Gooch, A., Thompson, W., Creem-Regehr, S.: Effects of Stereo Viewing Conditions on Distance Perception in Virtual Environments. Presence: Teleoperators & Virtual Environments 17(1), 91--101 (2008)

23. Zilch, D., Bruder, G., Steinicke, F., Lamak, F.: Design and evaluation of 3d gui widgets for stereoscopic touch-displays. In: Proc. of the GI-Workshop VR/AR. pp. 37--48 (2013)