

The Unadorned Desk: Exploiting the Physical Space around a Display as an Input Canvas

Doris Hausen¹, Sebastian Boring^{2,3}, Saul Greenberg³

¹University of Munich (LMU), HCI Group, Amalienstr. 17, 80333 Munich, Germany

²University of Copenhagen, Njalsgade 128, Bldg. 24, 2300 Copenhagen S, Denmark

³University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada

doris.hausen@ifi.lmu.de, sebastian.boring@diku.dk,
saul.greenberg@ucalgary.ca

Abstract. In everyday office work, people smoothly use the space on their physical desks to work with documents of interest, and to keep tools and materials nearby for easy use. In contrast, the limited screen space of computer displays imposes interface constraints. Associated material is placed off-screen (i.e., temporarily hidden) and requires extra work to access (window switching, menu selection) or crowds and competes with the work area (e.g., palettes and icons). This problem is worsened by the increasing popularity of small displays such as tablets and laptops. To mitigate this problem, we investigate how we can exploit an unadorned physical desk space as an additional input canvas. With minimal augmentation, our *Unadorned Desk* detects coarse *hovering over* and *touching of* discrete areas (‘items’) within a given area on an otherwise regular desk, which is used as input to the desktop computer. We hypothesize that people’s spatial memory will let them touch particular desk locations without looking. In contrast to other augmented desks, our system provides optional feedback of touches directly on the computer’s screen. We conducted a user study to understand how people make use of this input space. Participants freely placed and retrieved items onto/from the desk. We found that participants organize items in a grid-like fashion for easier access later on. In a second experiment, participants had to retrieve items from a predefined grid. When only few (large) items are located in the area, participants were faster without feedback and there was (surprisingly) no difference in error rates with or without feedback. As the item number grew (i.e., items shrank to fit the area), participants increasingly relied on feedback to minimize errors – at the cost of speed.

Keywords: Augmented desks, digital desks, peripheral interaction.

1 Introduction

In everyday office work, people naturally arrange documents, tools and other objects on their physical desk so that they are ready-to-hand, i.e., within easy reach and where they can be retrieved without actively searching for them. People are able to do so because they are aware of these objects’ spatial location [15] and can coarsely acquire

those that are in their peripheral vision. In contrast, working with computers requires almost everything to visually happen on-screen. Yet because space is limited, the so-called desktop metaphor usually separates object placement into one of several workspaces (see Figure 1a): the *primary workspace*, which covers most of the screen, holds the currently active document, which people normally work on; the *secondary workspace* is the portion of on-screen space that contains a subset of artifacts related to the primary space's activities, e.g., icons and tool palettes; finally, the *off-screen workspace* holds the remaining artifacts, where users – through a series of operations – make them explicitly visible in a temporary fashion (e.g., menus, dialog boxes).

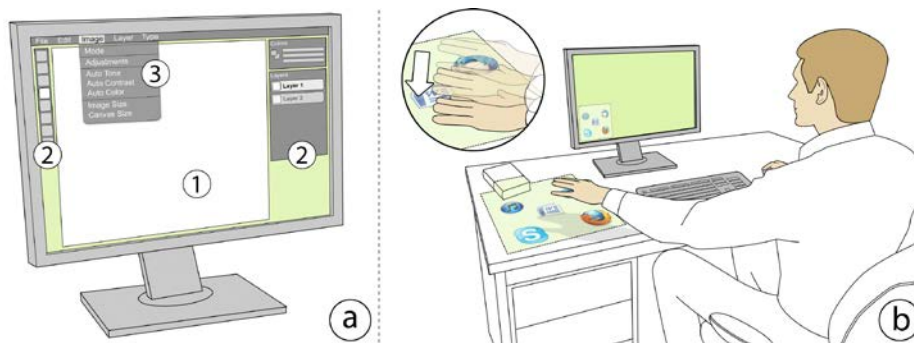


Fig. 1. (a) The three workspaces present in the desktop metaphor: the *primary workspace* (1) holds the active document people work on; the *secondary workspace* (2) holds items related to the activities in the primary workspace and is permanently visible; and the *off-screen workspace* (3) holds further items related to the document, yet people have to make them explicitly visible (e.g., menus). (b) The *Unadorned Desk* moves these workspaces onto a regular desk so that the primary workspace covers the entire display. When a person hovers over the interaction area on the desk, feedback may be given on-screen. Touching an item then selects it.

Yet, there is a tension between these workspaces. The primary and secondary workspaces spatially trade-off: the primary workspace dominates screen space, which leaves less space for its surrounding artifacts. This is especially true for tablets and other devices with rather small displays. The secondary and off-screen workspace also trade-off: it is much easier to select items in the secondary space, but only a few can be held there. In contrast, a huge number of items can be held in the off-screen workspace, but it is harder to select them (or to remember accelerator methods such as keyboard shortcuts) [21]. Instead of trying to fit everything on screen (directly or through menus), we investigate using the unadorned (i.e., unchanged except for a sensing device) desk as a further space to contain artifacts. Our hypothesis is that people can then easily select commonly used functions (e.g., tools or other windows) located on the desk's surface (see Figure 1b). This has several advantages. First, if we move artifacts from the secondary workspace to the desk, more display space can be allocated to the primary workspace. Second, if we move artifacts from the off-screen workspace to the desk, they will be easier to access. This also mimics the way we interact with everyday objects surrounding a document located on the desk (e.g., placing paints and brushes nearby for rapid retrieval while drawing).

Previous work on digital desks relies on a tight feedback loop, where visuals and interaction feedback were overlaid onto the regular desk surface i.e., by making the desk look and behave like a computer display. Examples include the use of projectors [20, 31], a tabletop computer as desk replacement [6], or by adding tablet computers next to the display as an interactive region [6]. These tend to be complex (or expensive) to set up. In contrast, the new generation of depth-sensing technologies mean that detecting touches and hovering is low-cost, such as via LeapMotion or Microsoft’s Kinect camera. The problem is that these technologies do not provide visual feedback. This begs the question: is visual feedback on the desk necessary?

We are particularly interested in using the desk *as is* with the smallest possible alterations. In this paper, we take an extreme stance, where we provide either no feedback or feedback on-screen (rather than on the desk) solely on demand. Both approaches keep desk instrumentation to a minimum, thus allowing for the use of *any* desk – such as at cafes – to serve as a workspace. Using computer science terminology, this is a lower bounds investigation: we want to understand to what extent interaction is possible using minimal or no augmentation (i.e., no visual targets or confirmatory feedback on the desk).

To investigate how an *unadorned desk* can be used as input space, we built a prototype using a Microsoft Kinect depth camera mounted atop a regular desk. Our Unadorned Desk tracks a person’s hand and allows for *hovering over* and *touching of* content. As we were interested in how people can interact with off-screen content while keeping their attention on their main task, feedback is either not provided, or is given on-screen and on demand. We conducted two experiments: the first *placement* experiment focused on placement strategies of participants. In the second *acquisition* experiment, varying numbers of virtual items were placed at predefined locations and participants had to retrieve them to find out which number is still usable for off-screen interaction. Our work offers two contributions: (1) a working prototype that makes use of an unadorned desk as input space by augmenting it with a depth camera. And (2), experimental results that inform the design of such interactions with respect to the amount of off-screen virtual items and the given on-screen feedback.

2 Related Work

Our work builds on several areas of research that relate to how people organize documents on their desk, peripheral and bimanual interaction, interfaces without direct visual feedback, and augmented desks in general.

Organizing the Desk. We routinely and fluidly arrange and manage documents on our physical desks without focusing much attention on it. We can do so because the document’s physical arrangement on the desk offers context information about the status and importance of certain tasks [8]. Malone studied desk organizations and found that files and (even more so) piles are the most commonly used arrangements on a desk [24]. Files are usually ordered systematically (e.g., in an alphabetic order). Piles, however, are not organized deliberately, and people thus more likely use spatial organization for retrieval. Associated tools and materials are generally arranged so

they are available for reuse, such as by placing them nearby and ready-to-hand during active use, or by organizing them into known locations (such as desk drawers) [12].

Many systems try to bring this traditional way of organizing a desk into the digital world. In *Data Mountain* [28], people can organize browser bookmarks on a virtual table, which proved to be faster than bookmarking in Internet Explorer 4. *BumpTop* simulates the desktop by allowing users to arrange documents in a virtual 3D space using physics [1]. Customization features in graphical user interfaces let people spatially arrange tools around the graphical desktop [12]. In contrast to these systems, we are interested in using the desk *as is* instead of mimicking it on-screen.

Augmented and Interactive Desks. There is a history of work where digital content is brought onto the surface of the physical desk. This not only provides a workspace larger than the constraints of a computer display, but – in some systems – also allows both physical and digital artifacts to be used in tandem. Early work focused on (partially) digitizing the desk. The *Digital Desk* [31] uses a projected interface on the desk. A video camera senses interactions with fingers and/or a pen, and can capture content of paper materials (i.e. interacting with paper). Rekimoto et al.'s *Augmented Surfaces* [26] are projected extensions to a laptop's display on a table or a wall. Users are able to drag content from their laptop onto the table where it is visible all times. Thus, the table serves as visual extension to the laptop's display. *Bonfire* [20] projects additional content next to a laptop's screen and allows touch input through cameras.

More recent prototypes augment the computer screen with a horizontal digital display ('surface') located underneath it. Surfaces typically allow for touch input, making sensing of user interaction easy (e.g., *Magic Desk* [6]). *Curve* [33] and *BendDesk* [30] merge the horizontal desk area and the vertical display area into one gigantic high-resolution touch-sensitive display, where they are seamlessly connected through a curve. Various studies investigated how particular touch regions on both the horizontal and vertical displays are used e.g., to show that the regions next to keyboard and mouse are best suitable for coarse interaction [6]. We build on this in that we use the areas left and right of the keyboard/mouse in our two studies.

Peripheral and Bimanual Interaction. Working with analogue documents on a desk often involves peripheral and bimanual interaction. *Peripheral interaction* offers coarse input styles in the periphery of the user's attention and thus quasi-parallel to the current primary task. The fundamentals for peripheral interaction are human capabilities such as divided attention (i.e., processing two tasks in parallel without switching channels [32]), automatic and habitual processes (i.e., carried out with little mental effort and hardly any conscious control [3]), and proprioception (i.e., being aware of one's own body, its posture and orientation [7]). Today's prototypes incorporating peripheral interaction mainly rely on TUIs (e.g., [4, 11, 17]) or freehand gestures [16, 18]. Our work adds to this by investigating how people interact coarsely in their periphery.

Bimanual (two-handed) interaction is the basis for peripheral interaction. While typically asymmetric, both hands influence each other leading to a kinematic chain [13]. Studies show that bimanual interaction can improve performance [9, 19]. At the same time, the body provides the kinesthetic reference frame, i.e., the user's sense of

where one hand is relative to the body and the other hand [5]. Further, Balakrishnan et al. found that while separating visual feedback from the physical reference does affect performance, there is only a “remarkably small difference” when comparing interaction with and without visual feedback as long as “body-relative kinesthetic cues are available” [5]. We build on this as we separate feedback from interaction.

Interfaces without Direct Feedback. Spatial interaction does not necessarily rely on direct feedback or feedback at all. Gustafson et al.’s *Imaginary Interfaces* [14] make use of the visual short-term and visuospatial memory. By forming an “L” with the non-dominant hand a reference frame is created. *Spin & Swing* [2] depends on an imaginary circle around the user. By turning themselves, users navigate through the content displayed on a handheld device. The concept of *body-centric interactions* [10] employs the space around a person’s body to hold mobile phone functions. For example, *Virtual Shelves* [22] positions items in a hemisphere in front of the user. *Point upon Body* [23] uses the forearm as interaction area, which can be divided at most into six distinct areas. *GesturePad* [27] and *BodySpace* [29] use different body locations for commands. As with our system, no direct feedback is provided. These systems rely primarily on spatial awareness and kinesthetic memory. Due to proprioception, users have a good understanding of where items are located and can easily – even with closed eyes – place and retrieve such objects [25]. These findings inspired us to mimic regular desk use as means for interacting with digital content.

3 Evaluating Off-Screen Interactions

In order to better understand how users can adapt to the novel input technology as well as how on-screen feedback for off-screen content would affect the interaction, we conducted two user studies. The first experiment aimed at understanding how people would spatially place various content items onto the desk that they would later retrieve. More precisely, we wanted to see whether people make use of special arrangements of their content. In the second experiment (which was tuned to use the results of the first study), we aimed to see how accurately participants could locate items placed in off-screen space as a function of the number of items in that space. The next section details the conditions and apparatus common to both experiments.

3.1 Conditions Common to Both Experiments

Although the tasks varied in both experiments, we had two conditions (additional to the experiment-dependent ones) that were the same in both studies: (1) the hand with which participants interacted in off-screen space, and (2) the type of feedback given during the task. In the following, we describe these two conditions in more detail.

Handedness: We chose to test our system with both hands. In the **dominant** hand condition, participants interacted with off-screen content using the hand they usually use to perform precise interactions (e.g., writing). In the **non-dominant** hand condition, they used the other hand. For each of the conditions, the interaction area was placed on the desk so that it was closest to the hand with which they had to interact in off-screen space (i.e., not reaching left of the keyboard using the right hand).

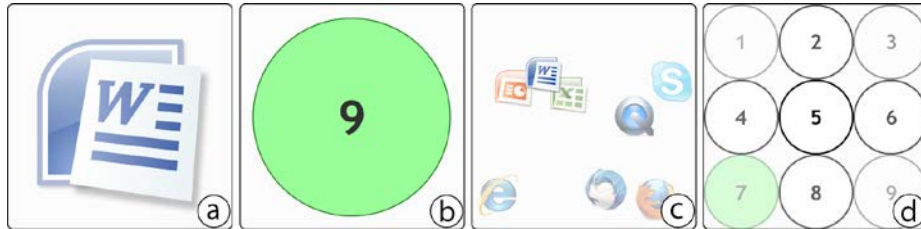


Fig. 2. In the *Single* feedback condition, the system showed the item closest to the participant's hand (a: 1st study, b: 2nd study). In the *Full* feedback condition, all items are shown with their correct spatial layout (c: 1st study, b: 2nd study). Here the participant hovers over the word icon (and item #5 respectively). In all conditions, transparency encoded the distance to that item.

Feedback: We had three conditions for on-screen feedback. In the **No Feedback (None)** condition, participants did not receive any feedback on the computer's display, forcing them to rely solely on their spatial memory and proprioception. In the **Single Item Feedback (Single)** condition, participants only saw the item that was closest to their hand, with the distance being encoded through transparency. That is, as participants moved closer to a respective item, the item's icon became increasingly opaque (see Figure 2a,b). In the **Full Area Feedback (Full)** condition, participants saw all items in the interaction area with correct spatial layout. As in the *Single* condition, the transparency of items again changed based on the distance between them and the participants' hands (see Figure 2c,d). That is, the item directly below the hand was more opaque than the surrounding items. The feedback area (400×400 pixels) was only shown on-screen while a participant's hand was inside the interaction area and invisible otherwise to not occupy valuable screen space. It was also located close to the interaction area (i.e., the bottom left or right corner of the display).

We used a within-subjects factorial design in both experiments: $2 \text{ Handedness (Dominant, Non-Dominant)} \times 3 \text{ Feedback (None, Single, Full)}$. *Feedback* was counterbalanced across participants. To minimize changing the camera's location for *Handedness*, we alternated participants so that the first participant had all three *Feedback* types with the *Dominant* hand and then again with the *Non-Dominant* one, while the second one started with *Non-Dominant* etc.

3.2 Apparatus, Setup and Participants

The *Unadorned Desk* uses a Microsoft Kinect depth camera mounted on a tripod facing upside down (see Figure 3) observing a sub-region of the desk within which a person could interact using the hand. The prototype runs on an Intel i7 3.4 GHz computer to allow for fast processing (i.e., 640×480 pixel frames at 30 frames per second).

We use the Kinect depth camera to gather hand information within the tracked region. The camera provides depth images where each pixel in a depth frame encodes that pixel's distance to the

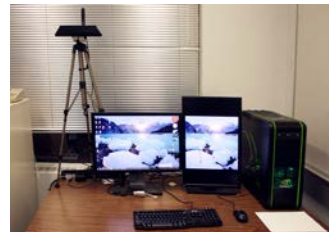


Fig. 3. The *Unadorned Desk*: a Kinect tracks the user's hand.

camera in millimeters. At startup, the system takes a series of depth images, averages them (to reduce noise), and uses them as ground truth. Once running, it calculates the difference between the current depth frame and the calibrated depth image. The calculated difference image contains all points that are ‘new’ to the scene (e.g., a hand) with their distance to the desk. Using this point cloud, the system calculates the point of the hand closest to the corner of the interaction area that is the furthest away from the user (i.e., the tip of the middle finger). The vertical distance (depth) of that location to the desk further determines the hand’s state: *touching* (depth < threshold), *hovering over* (depth \geq threshold), or *absent* if no hand is detected. On-screen feedback is optionally provided once the user’s hand enters the interaction area. When the hand touches an item, the system performs the action associated with that item.

In both experiments, participants were seated centrally in front of the computer’s display. The depth camera captured a region of 40 cm \times 36 cm (33.5 cm on the top edge due to slight camera distortion) next to the keyboard aligned with the desk’s edges. For each *Handedness* condition, we moved the monitor, keyboard, mouse and chair to ensure that the participants are seated centrally in front of the display and close to the interaction area. The tracked region on the desk was empty. The computer display’s background was set to a uniform color and had all desktop icons removed.

Each study used 12 participants. Sexes were mixed (first: study 6 female; second: 4 female), and ages ranged from 19 to 30 (average was 24). Each person only participated in one of the studies to minimize learning effects. Handedness varied, 9 were right-handed in the first study, and all in the second. Each session lasted up to 1.5 hours, and all participants received \$20 as compensation for their time.

3.3 Hypotheses

We had similar hypotheses in both studies:

- H1.** Item retrieval time would increase as the number of off-screen items increased.
- H2.** Error rate would increase as the number of off-screen items increased.
- H3.** Item retrieval time would increase when no feedback was present.
- H4.** Offset and error rates would decrease with feedback present.

4 Study 1 – Placing and Retrieving Content

The purpose of our first study is to understand how participants would use of the *Unadorned Desk* to organizationally place and later retrieve an item, and the effect of having an increasing number of items placed within that space. In particular, we were interested in (1) how they arrange a given number of items on their desk, and (2) the offset (and the item’s size respectively) when retrieving items to ensure successful pointing in the periphery. Our *Handedness* \times *Feedback* factorial design was extended to include a third *Sets* condition, which is the number of items participants had to place and retrieve in the off-screen space. We used well-known, easily identifiable applications, which had meaning to our participants: Word, Excel, Power Point, Firefox, Thunderbird, Skype, QuickTime, and Internet Explorer. For each condition, the amount of items was ascending (to increase difficulty): 2, 4, 6, and 8 different items.

4.1 Tasks and Procedure

The experiment consisted of two phases for each combination of *Handedness* and *Feedback*: placing items and later retrieving them. We instructed users to place items off-screen in a position of their own liking. However, items had to have a minimum distance of 47.6 millimeters (and 50 pixels respectively) to avoid overlaps of them, which would make retrieval more error-prone. Each set of items they had to place was shown on the monitor during the placement task (see Figure 4a), so that participants were aware of all items and could group them if that would aid their memory. To place an item, participants first had to hit the spacebar to indicate they were starting the task, at which point timing began. Once the trial was active, they could move their hand into the interaction area and place the item by touching the desk's surface.

When feedback was given, already placed items were shown to give participants a feeling of the location of other items (see Figure 2a,c). Participants repeated this step until they had placed all icons in the current set in the physical off-screen space.

For retrieval, the system notified participants on-screen of which item to retrieve before the trial began (Figure 4b). They then had to hit the spacebar to activate the trial (Figure 4c). Users would then retrieve that previously placed off-screen item. Retrieval worked exactly like the placement: hit spacebar for time measurement and touch a location to retrieve the item. Afterwards, the system prompted them with the next item until all items were retrieved. If the wrong item was retrieved, the participant was not informed, the trial was not repeated and the experiment continued but the error was recorded. For each *Feedback* and *Handedness* combination, participants placed 4 *Sets* of items (2, 4, 6, and 8 items) once and then retrieved each of them 4 times. We collected 24 placement sets (480 item retrievals).

For placement, we recorded all x,y locations (as the center) of placed items. For the retrieval task, we measured the time from the beginning of a trial (i.e., hitting the spacebar) until they touched the desk's surface. We further recorded the location they touched, the distance to the actual item (x,y location), and the amount of items that were closer than the correct one (i.e., errors). We manually counted the participants' gazes, whether they looked at the interaction area, the feedback area, or both (the experimenter pressed a key for each gaze, which was recorded). Finally, we asked participants to fill out a device assessment questionnaire: once after completing one *Feedback* and *Handedness* condition, and again at the end of our study.

4.2 Results

We used *heat maps* to uncover how people would freely place items on the desk. We then compared *retrieval time*, *retrieval offset*, and *gazes* using repeated measures within-subjects analyses of variance (ANOVA). For pair-wise post hoc tests, we used

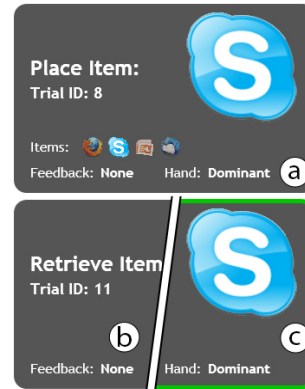


Fig. 4. Commands: Placing an item (a), and retrieving it (b: before trial activation, c: after).

Bonferroni-corrected confidence intervals to retain comparisons against $\alpha = 0.05$. When the assumption of sphericity was violated, we used Greenhouse-Geisser to correct the degrees of freedom. All unstated p -values are $p > 0.05$.

We performed a $2 \times 3 \times 4$ (*Handedness* \times *Feedback* \times *Items*) within-subjects ANOVA. As we did not find any significant main effects or interactions for *Handedness*, we aggregated over *Handedness* for all subjects in subsequent analyses. For heat map analysis, we mirrored interactions performed in the area right to the keyboard to bring those into the coordinate system of the one left to the keyboard.

Strategic Placement of Items. Through a heat map analysis (see Figure 5) we found that many participants tended to arrange items based on an imaginary grid (thus item placement was not random). Further, participants followed other semantic patterns: first, some placed items in a single row as in the dock in Mac OS X. During retrieval with feedback, participants then hovered over that line to find the correct item. Second, some hierarchically grouped similar items together (e.g., all browser icons). They would later retrieve the item by first going to the general group area containing that item, and then selecting the particular item. Finally, the more frequently they use an application based on their personal usage outside the study, the closer they would place it to the keyboard. Items used less often are thus further away from the primary interaction space. Participants did consider that areas further away would require more physical effort to access an item. However, all participants made use of the *entire* area, as they felt more comfortable to access items placed further apart from each other.

We calculated three *Distances* (*Closest*, *Average*, and *Highest*) between items that they had placed off-screen. Participants placed items with an average distances for all conditions between 207.4 and 231.6 millimeters ($M=219.2$; $SD=9.7$). To understand whether *Feedback* or the *Set* of items had an influence on the distances between items, we performed separate 3×4 (*Feedback* \times *Set*) ANOVAs for each *Distance*. For the closest distance, we found a significant main effect for *Set* ($F_{1,953,21.487} = 184.76$, $p < 0.001$) and post hoc multiple means comparisons revealed that the distance increases with a decreasing *Set* of items (all pairs except 6 and 8 items differ with $p = 0.011$) regardless of *Feedback*. *Feedback* had an effect on the highest distance between items, where we found significant main effects for both *Feedback* ($F_{2,22} = 15.49$, $p < 0.001$) and *Set* ($F_{3,33} = 128.74$, $p < 0.001$). Smaller *Sets* lead to lower distances between items except for 6 and 8 items (all $p < 0.001$). More importantly, in the *None* feedback condition, participants placed items further away. The differences further increase with the *Set* size. Particularly for 8 items, *None* significantly differed from the other two (all $p < 0.05$), and from *Single* for *Set* sizes 2 and 6 (all $p < 0.05$). Thus, when relying on feedback, participants felt more comfortable placing items closer to each other. Interestingly, *Single* and *Full* did not differ for any *Set* size, and there was no significant difference between all three conditions for the *Set* with 4 items, which we attribute to participants using the four corners of the area.

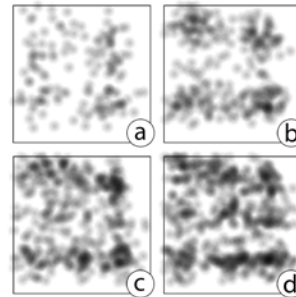


Fig. 5. Heat maps (a-d: 2 to 8 items) show that users tend to arrange items in grids.

Retrieval Time. We compared retrieval times from the moment participants hit the spacebar until they retrieved an item. We only took into account the correct retrieval times (even so, we did not find significant differences between retrieval times with and without errors). We performed a 3×4 (*Feedback* \times *Set*) within subjects ANOVA and found significant main effects for *Feedback* ($F_{2,20} = 31.098, p < 0.001$) and *Set* ($F_{1,609,17.698} = 15.583, p = 0.011$). Figure 6a suggests that retrieval times slightly increase with larger *Sets*. However, *Feedback* influences retrieval times. Separate ANOVAs for each *Set* showed that *No Feedback* was always faster (all $p < 0.001$). Furthermore, the two conditions with visual feedback were more strongly affected by the *Set* of items. Overall, *None* was the fastest ($M=1.40s, SD=0.36s$), followed by *Full* ($M=2.47s, SD=0.88s$), and *Single* ($M=2.68s, SD=1.06s$).

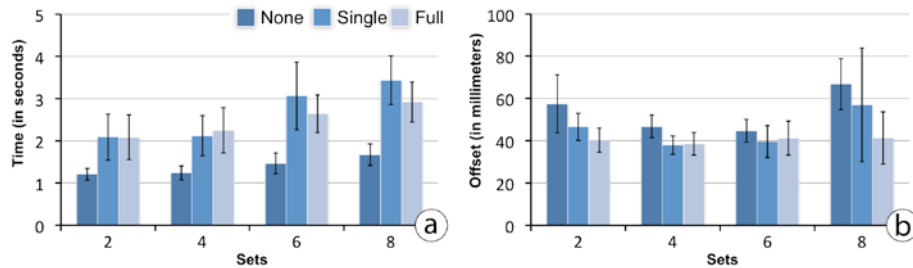


Fig. 6. Results of the placement study: (a) retrieval time for one item for all feedback conditions and sets; (b) offset for correct retrievals measured as Euclidean distance between the item's center and the touch's location. Error bars represent 95% confidence intervals.

Offset. We compared the offset (the distance between the touch point and the item's center). We chose to only include successful retrievals to eliminate cases where participants did not remember an item's location and thus randomly touched the desk. We performed a 3×4 (*Feedback* \times *Set*) within subjects ANOVA and only found a significant main effect for *Feedback* ($F_{2,22} = 4.201, p = 0.027$) but no effect for *Set* and no interactions. Figure 6b summarizes the results for different *Sets* and *Feedbacks*: *Full* had the smallest offset between the touch and the item's center ($M=36.6$ millimeters), followed by *Single* ($M=41.1$ millimeters) and *None* ($M=48.9$ millimeters). Figure 6b also reveals that, in order to have 95% successful selections regardless of *Feedback* and *Set*, an item with a radius of at least 85 millimeters is sufficient.

We were also interested in the impact of *Feedback* on wrong retrievals (i.e., touch was closer to an incorrect item than to the correct one). We normalized the data by dividing the number of incorrect closer items by the maximum number of possible wrong items (i.e., 1 for a set of 2, 3 for a set of 4, etc.). We performed a 3×4 (*Feedback* \times *Set*) within subjects ANOVA and found a significant main effect for *Feedback* ($F_{1,22,13.419} = 4.914, p = 0.039$). Post hoc tests revealed that only for a *Set* with 6 items *None* was more error-prone than the other two conditions ($p = 0.04$). We believe that – particularly with no visual cues on the desk – participants made use of space to more easily retrieve an item. That is, a larger offset still leads to correct retrieval. In summary, the chance for an erroneous selection with *None* is 20% ($SD=12\%$), and 15% ($SD=8\%$) for *Single* and *Full*. This can be lowered, however, by increasing the required minimum distance between items.

Gaze Analysis. We told participants to minimize looking at the interaction area, and instead imagine that they were concentrating and looking at their primary on-screen task. We did not instruct them with respect to using the feedback window, that is, they could freely make use of it. We report gazes averaged across both placement and retrieval phase. There were no gazes to the feedback area in the *None* condition.

For *Gazes* to the *Interaction Area*, we performed a 3×4 (*Feedback* \times *Set*) within subjects ANOVA and found significant main effects for *Feedback* ($F_{1,126,12,383} = 7.948, p = 0.012$), *Set* ($F_{3,33} = 14.494, p < 0.001$) and a *Feedback* \times *Set* interaction ($F_{2,15,23,645} = 8.618, p < 0.001$). Post hoc tests revealed that for 6 items participants gazed at the interaction area more often in the *None* condition compared to *Single* ($p = 0.027$). For 8 items, they gazed more often using *None* compared to the other two conditions (all $p = 0.016$). For *Gazes* to the *Feedback Area*, we did not test the *None* condition (as there was no feedback area) and performed a 2×4 (*Feedback* \times *Set*) within subjects ANOVA and found a significant main effect for *Set* ($F_{2,121,23,334} = 7.274, p = 0.002$). Pairwise comparisons showed that *Gazes* to the *Feedback Area* increase with larger *Sets* (2 and 4 differ from 6, all $p = 0.044$, and 2 differs from 8, $p = 0.021$). Overall, when *No Feedback* was presented, participants gazed at the interaction area on the desk more often (0.24 times per trial), compared to *Single* (0.11) and *Full* (0.12). In conditions that had *Feedback*, participants gazed at the feedback area 0.72 (*Full*) and 0.66 (*Single*) times per trial. Thus, participants ‘left’ their fictive primary task more often (i.e., looked away from it) when feedback was presented.

4.3 Discussion

During placement, we observed that participants used the whole interaction area, even though they stated that retrieval was easier if the item was placed closer to them. Placement was reasonably systematic, each following some kind of spatial organization. We noticed an increased time for placement and found significant differences for item distances with *No Feedback*. We believe that participants put more effort into a good arrangement (with reasonably spaced items) to allow for easier retrieval afterwards, which was especially important when there was no visual feedback.

During the retrieval stage, the *None* condition caused two problems for participants: (1) they had to remember where they put items, and (2) they were not informed whether they actually had correctly acquired an item. Interestingly, participants stated afterwards that – when feedback was provided – they felt pressured to point more precisely although this would not have been necessary (i.e., the selected item was always the one closest to the touched location), resulting in longer retrieval times for conditions with feedback. One participant stated that he started to search instead of think, which slowed him down. Our analysis of gazes supports this view: participants more often looked away from their fictive primary task when feedback was given. In fact, they looked more at the feedback area (when available) than at the interaction area when no feedback was given. *Feedback* did help participants to remember locations and decreased their offset for larger *Sets*, but also slowed them down.

Recall that these interaction techniques are to allow coarse interaction in the periphery (preferable with minimal attention). Our results suggest a suitable tradeoff

between the item’s sizes and the overall number of items. We observed that participants had problems recalling their spatial layout with 6 or more items. Nevertheless, the results also indicate that participants were able to successfully retrieve 2 or 4 items – even without feedback. While the number of manageable items in real life scenarios could be quite large (e.g., participants may want to place many items meaningful to their task on the interaction area), others have argued that a small number of such items could comprise a large number of the actions people actually do [12]. Examples are frequently or recently used commands. Nevertheless, this first experiment suggests that having more items decreases the probability of a correct retrieval. Quite possibly, our results could be affected by less than optimal placements of items on the desk, e.g., due to a lack of visual cues on the experiment’s desk. For this reason, we conducted a 2nd study that spatially separated items into a grid (a layout applied by many in this first study), and that did not require to memorize locations, which is hard to achieve anyhow in a lab study setting, especially for long-term memory.

5 Study 2 – Targeting Content

To prevent memorizing (our lab study is only able to test short-term memory) where items were placed and eliminate the potential influence of unfavorable placement, we presented our participants with a predefined layout. Based on the 1st study, where participants had arranged items in a grid, we created grid-like layouts with pre-placed items, which was visible to them on-screen during each of the trials. We added a variable *ItemSize* with three levels: *Small* (10 cm wide), *Medium* (13.3 cm), and *Large* (20 cm). To fill the entire interaction area, we decided to fill the grid accordingly. That is, we had 16 (4×4) *Small*, 9 (3×3) *Medium*, and 4 (2×2) *Large* items. In this experiment, we were interested in getting more insights on item locations and size with respect to retrieval time, offset and errors.

5.1 Task and Procedure

Both, task and procedure were similar to the retrieval task of the first study (though items are already pre-placed on the desk). At the beginning of each trial, the system showed participants which item they had to retrieve (see Figure 7a-c). As before, they activated the trial by hitting the spacebar and retrieved the respective item from off-screen space by touching the respective location. If they retrieved the correct item, the system prompted them with the next item to retrieve. If they touched the wrong one, the system notified them that the trial was incorrect, increased the item’s error count, and asked them to retrieve it again. However, to avoid frustration, the system moved on to the next item after three failed attempts. Participants had to retrieve each of the different *ItemSizes* three times for all *Handedness* and *Feedback* combination, thus requiring every participant to perform 522 retrievals. How-



Fig. 7. The target item (green), among all other items. a-c: *Large*, *Medium*, and *Small* items.

ever, we excluded the first block as training block. We logged: *task time* from the moment the spacebar was hit until they either successfully retrieved the item or missed it; the *Euclidean distance* (i.e., *offset*) of the touch to the item’s center; and the number of *errors* (we allowed a maximum of 3 errors per item). As in the first study, we manually tracked whether the participant looked at the interaction area on the desk, on the feedback area on the screen or both of them. After each *Feedback* and *Handedness* combination, participants filled out the same device assessment questionnaire used in the first study as well as a closing questionnaire.

5.2 Results

We performed a $2 \times 3 \times 3$ (*Handedness* \times *Feedback* \times *ItemSize*) within subjects ANOVA. As in the first study, we did not find any significant main effects or interactions for *Handedness*. Thus, in subsequent analyses, we aggregated over *Handedness* across all participants. We also excluded all erroneous, unsuccessful retrievals from analyses of retrieval time and offset, as we ended a trial after three incorrect retrievals). Because of this, we excluded 6.5% of all trials.

Retrieval Time. Regarding retrieval time for an item, we performed a 3×3 (*Feedback* \times *ItemSize*) within subjects ANOVA and found significant main effects for *ItemSize* ($F_{1,272,13,997} = 15.269$, $p < 0.001$) and *Feedback* ($F_{2,22} = 19.037$, $p < 0.001$). We further found an *ItemSize* \times *Feedback* ($F_{4,44} = 5.414$, $p < 0.001$) interaction. Post hoc multiple means comparisons showed that for all *ItemSizes* retrieval time differed significantly for the *None* condition (users needed less time) compared to the other two (all $p = 0.017$). Further, for *Single* and *Full*, the retrieval time for the *Small* items differed significantly from the shorter retrieval time for the *Medium* and *Large* items ($p < 0.001$). Overall, *None* was the fastest ($M=1.68s$), followed by *Single* ($M=2.25s$), and *Full* ($M=2.33s$). Figure 8a summarizes these results.

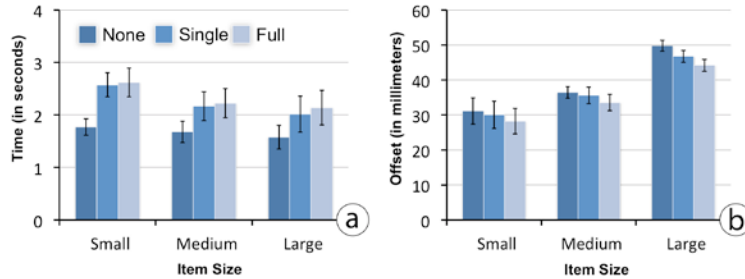


Fig. 8. Results of the targeting study: (a) retrieval time for one item for all feedback conditions and item sizes; (b) offset for correct retrievals measured as Euclidean distance between the item’s center and the touch’s location. Error bars represent 95% confidence intervals.

Retrieval Offset. For the analysis of the offset of successful retrievals (measured as Euclidean distance between the touch and the item’s center), we normalize the distance as we had different *ItemSizes*. To do so, we divided the measured offset by the maximum possible offset (i.e., the item’s actual size). With the normalized data, we performed a 3×3 (*Feedback* \times *ItemSize*) within subjects ANOVA and found significant main effects for *ItemSize* ($F_{2,22} = 39.318$, $p < 0.001$), and *Feedback* ($F_{2,22} =$

4.918, $p = 0.016$), but no *Feedback* \times *ItemSize* interaction. Pairwise comparison of different *ItemSizes* across all *Feedback* conditions further revealed that participants were always relatively closer to the item's center (yet physically further away) for *Large* items ($p = 0.007$). Overall, participants had the smallest offset for *Large* items (46.9% of the item's width), followed by *Medium* (52.1%), and *Small* (59.5%). However, when looking at the non-normalized offset (see Figure 8b), the results are the exact opposite: participants had the least offset for *Small* items (29.7 mm), followed by *Medium* (34.6 mm) and *Large* (46.9 mm) ones.

We normalized errors since we had a different amount of items depending on the *ItemSize*. We divided the errors by the number of items in the grid for each trial. With these values, we performed a 3×3 (*Feedback* \times *ItemSize*) within subjects ANOVA and found significant main effects for *ItemSize* ($F_{2,22} = 88.909$, $p < 0.001$), *Feedback* ($F_{1,309,14.4} = 10.587$, $p = 0.002$), and a *Feedback* \times *ItemSize* ($F_{2,126,23.385} = 4.036$, $p = 0.028$) interaction. Post hoc tests showed that the *None* condition differed significantly from the other two for the *Large* (all $p = 0.018$) and from the *Full* condition for the *Small* items ($p = 0.008$). However, *Feedback* conditions do not differ significantly for the *Medium* ones. For all *ItemSizes*, *None* was the most error prone ($M=0.41$, $SD=0.23$), followed by *Single* ($M=0.22$, $SD=0.16$) and *Full* ($M=0.18$, $SD=0.10$).

To understand the error-prone performance, Figure 9 visualizes the locations where users had the most errors as heat map. As trend, one can see that for *larger* items the corner furthest away from the user caused the most errors. However, the *smaller* items get, the more errors occur in the center, which can be explained by the desk's edge (and the borders of the interaction area respectively) being a reference frame. This made it easier to target items close to the borders and harder in the center.

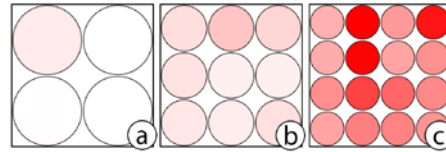


Fig. 9. Heat maps showing errors (aggregated upon all feedback conditions, mirrored for the right interaction area) for a) large, b) medium and c) small items. Saturation indicates errors.

In a second analysis, we excluded the items further away: for *Large* items, we excluded the top left item, for *Medium* items the three items furthest away, and for *Small* items the six items furthest away. We performed the same 3×3 (*Feedback* \times *ItemSize*) within subjects ANOVA using the reduced set and found significant main effects for *ItemSize* ($F_{2,22} = 23.941$, $p < 0.001$), *Feedback* ($F_{1,332,14.648} = 9.973$, $p = 0.003$), but no *Feedback* \times *ItemSize* interaction. Post hoc tests revealed that both, *Single* and *Full*, differed significantly from *None* only for *Small* items (all $p = 0.045$). This substantiates that the corner furthest away was the most error-prone. Nevertheless, *None* is still the most error-prone across all *ItemSizes*, with the least errors for *Large* items with 0.037 errors per trial (*Single*: 0.012, *Full*: 0.019).

Gaze Analysis. We instructed participants in the same way as we did in the first experiment. For *Gazes* to the *Interaction Area*, we performed a within subjects ANOVA on *Feedback* and found a significant main effect ($F_{1,136,12.495} = 10.485$, $p = 0.004$). Multiple means comparisons revealed that users gazed more often at the *Interaction Area* in the *None* condition compared to the other (all $p = 0.022$). We again excluded

the *None* condition for *Gazes* to the *Feedback Area*, and performed a within subjects ANOVA on the remaining two *Feedback* factors and did not find a significant effect.

Overall, *None* had the most gazes to the interaction area (0.23 times per trial), compared to *Single* (0.05) and *Full* (0.06). In *Feedback* conditions, participants gazed at the feedback area 0.69 (*Full*) and 0.65 (*Single*) times per trial.

5.3 Discussion

The second study re-enforces the findings from the first study. As before, *No Feedback* led to shortest retrieval times. Retrieval time also increased for *Small* items when feedback was present, yet it did not change when no feedback was given. Naturally, *Small* items required participants to select more precisely. The absolute offset from the center of an item for *Large* items (with 4.69 cm) would almost not suffice for *Small* items (as they only had a radius of 5 cm and a width of 10 cm respectively). Users seemed to make use of space for larger items (it did not matter how close to the center they touched the item) and adjusted their offsets for smaller ones.

No Feedback caused significantly more errors with the corner further away from the user included in the analysis. Similar to Magic Desk [6], where Bi et al. found that completion time was longer for areas further away from the keyboard, our users had problems acquiring targets further away. When we excluded items further away from analysis (i.e., only considering that half of the interaction area closer to the participant), the *No Feedback* condition only differed significantly from the others for the *Small* items. However, the error rate for *Small* items was high regardless of feedback. Thus, items with a size of 10 cm or less are generally too small to be manageable in the periphery on an unadorned desk independent of the provision of feedback.

6 General Discussion

In both studies, more items in the interaction area require lower offsets between a touch and the item's center – in the first study to ensure that the correct one is still the closest item, and in the second study because items got *smaller* as their number increased. As we hypothesized in H1, both studies showed that retrieval time increases as the number of items in the interaction area increases. While H2 suggested that error rate increases with more off-screen items, our experiments only partially support this. We did not find evidence for more errors when increasing the item number (up to 8) in the first study. Similarly, we did not find a significant effect in the second study for *Medium*-sized items, but did find a significant effect for *Small* ones. Thus, H2 (i.e., more errors with more items) is only supported for 10 or more items. H3 suggested that participants' time to retrieve items would increase when no feedback was present. Indeed, in both studies retrieval times were shorter when participants did not have *Feedback*, which fully supports this hypothesis. And finally, in H4 we hypothesized that the participants' offsets would increase and their error rate decrease when feedback was given. Yet, our results at best show a tendency towards more errors and larger offsets without feedback. In the first study, there was no significant effect for offset, and a significant effect on errors only for 6 items, but not for 8. In the second study, we found an effect for *Small* and *Large* items (but not for *Medium* ones) be-

tween *No Feedback* and *Full Feedback* (yet not for *Single*). When only analyzing that half of the interaction area closer to the participant, *No Feedback* only differs significantly from the other two for *Small* items. Thus, our results therefore do not support H4 and only show a tendency towards *No Feedback* increasing offset and errors.

The first study showed that participants made use of the whole interaction area, even with a small number of items. In the second study, we found that items located closer to keyboard and mouse, are less error prone than those located further away. This suggests that a rectangular shape might not be the most suited interaction space. In in-situ experiments, however, users would have a better reference frame (i.e., items on the table that convey meaning) instead of just the blank desk – which ultimately would influence on the results.

Our study showed that simple interaction on an unadorned desk is possible, albeit with a modest number of items and a reasonable item size (the first study revealed 85mm to achieve 95% successful retrievals, which would have sufficed for the second study). As the number of items increased, both retrieval times and error rate increased as well. However, previous studies on peripheral interaction showed that this interaction style needs to be trained and learned to be effective [3, 17], which naturally is not possible in a short-term laboratory experiment. Abandoning feedback leads to faster retrieval times and functions (in terms of offset and errors) for a small numbers of items. Our findings suggest that the amount of items on the desk should be limited to less than ten. Similar to the shape of the interaction area, we expect this number to be higher if the desk contains more physical objects that serve as a visual cue or anchor and participants are used to the system and place meaningful items on the desk.

Overall, participants enjoyed interacting with the unadorned desk, and considered it to be fairly easy. All were able to carry out the interaction equally well with their dominant and non-dominant hand, which strengthens our understanding that it is a peripheral interaction style. Interestingly, some of them were also irritated by this kind of interaction as they thought that the entire hand (and its palm respectively) acts as input, where in fact only a single point of the hand was tracked. Nevertheless, those participants adapted to the interaction fairly quickly.

7 Conclusion And Future Work

We presented the *Unadorned Desk*, which supports peripheral coarse interaction and extends the input- and workspace beyond a computer's display. The *Unadorned Desk* relies on hand tracking by a depth camera (Kinect). Our studies showed that users are capable of interacting with virtual items on the desk, for small numbers of items even without on-screen feedback. It is a lower-bounds performance study, as we deliberately did not place anything on the desk's surface to indicate an item's virtual location.

Our current experimental implementation suffers from three limitations that restrict its deployment for everyday use. First, as with most optical tracking systems, the system is susceptible to false detections when sunlight hits the tracked area. That problem also occurs with our depth camera, as the sun's infrared light does interfere with the structured, infrared light of Microsoft's Kinect. Second, the system requires mounting a depth camera atop a desk, which is unsuitable for situations where rapid

setup and teardown is required (e.g., temporary desks). This limits our ability to study the Unadorned Desk during anticipated everyday use. Third, the prototype does not yet address the fact that not every interaction on the desk is actually meant as input to the computer (e.g., retrieving a book). While emerging technologies will likely address the first two limitations, more research is needed to find an appropriate, distinct, yet not distracting gesture. Despite these limitations, our prototype allows us to evaluate implications for interaction on unadorned desks and to envision example applications such as those shown in the video figure¹.

There are still many unanswered questions for future work. Our first experiments were carried out in an artificial lab setting, which brings with it usual concerns about external validity. The primary task was placement and retrieval, rather than one's actual work. The items had no special significance. Interferences with other tasks carried out at the desk are not explored yet. Repeating the study in field cases could reveal nuances not seen in the lab. Our interaction area was rectangular, of a given size, and uncluttered; all these could both be varied to see how it affects performance. It was also in 2D (albeit with a hover plane). Yet a 3D interaction space is possible, e.g., virtual piles where a user can navigate through it with the hovering hand. Finally, ours was a lower bounds study of an unadorned desk. There could be many possible ways of introducing modest adornments that indicate position. Although this would now introduce desk artifacts, it could improve performance significantly.

8 Acknowledgments

This work is partially funded by the iCORE/NSERC/SMART Chair in Interactive Technologies, Alberta Innovates Technology Futures, NSERS, and SMART Technologies Inc., and the German state of Bavaria. The authors thank Andreas Butz, Matthew Dunlap, Richard Fung, Ulrike Kister and the Interactions Lab at the University of Calgary for their valuable input and their support throughout this project. We further thank everyone participating in our two experiments.

References

1. Agarawala, A., Balakrishnan, R. Keepin'it real: pushing the desktop metaphor with physics, piles and the pen. CHI, (2006).
2. Altakrouri, B., Kawsar, F., Kortuem, G. Spin&Swing: Spatial interaction with orientation aware devices. Pervasive, (2010).
3. Bakker, S., Hoven, E. van den, Eggen, B. Design for the periphery. Eurohaptics, (2010).
4. Bakker, S., Hoven, E. van den, Eggen, B., Overbeeke, K. Exploring peripheral interaction design for primary school teachers. TEI, (2012), 245-252.
5. Balakrishnan, R., Hinckley, K. The role of kinesthetic reference frames in two-handed input performance. UIST, (1999).
6. Bi, X., Grossman, T., Matejka, J., Fitzmaurice, G. Magic desk: bringing multi-touch surfaces into desktop work. CHI, (2011).
7. Boff, K.R., Kaufman, L., Thomas, J.P. Handbook of perception and human performance. (1986).

¹ Video Figure of the Unadorned Desk: http://youtu.be/ePQxR3EzJ_I

8. Bondarenko, O., Ruud, J. Documents at hand: Learning from paper to improve digital technologies. CHI, (2005), 121-130.
9. Buxton, W., Myers, B.A. A study in two-handed input. CHI, (1986), 321-326.
10. Chen, X., Marquardt, N., Tang, A., Boring, S., Greenberg, S. Extending a mobile device's interaction space through body-centric interaction. MobileHCI, (2012).
11. Edge, D., Blackwell, A.F. Peripheral tangible interaction by analytic design. TEI, (2009).
12. Greenberg, S. The computer user as toolsmith: The use, reuse, and organization of computer-based tools. (1993).
13. Guiard, Y. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of Motor Behavior 19, 4 (1987), 486-517.
14. Gustafson, S., Bierwirth, D., Baudisch, P. Imaginary interfaces: spatial interaction with empty hands and without visual feedback. UIST, (2010), 3-12.
15. Harrison, S., Dourish, P. Re-place-ing space: The role of place and space in collaborative systems, CSCW, (1996), 67-76.
16. Hausen, D., Boring, S., Polleti, J., Butz, A. Exploring design and combination of ambient information and peripheral interaction. DIS Work in Progress, (2012).
17. Hausen, D., Boring, S., Lueling, C., Rodestock, S., Butz, A. StaTube: Facilitating state management in Instant Messaging Systems. TEI, (2012), 283-290.
18. Hausen, D., Wagner, C., Boring, S., Butz, A. Comparing Modalities and Feedback for Peripheral Interaction. CHI Extended Abstracts, (2013).
19. Kabbash, P., Buxton, W., Sellen, A. Two-handed input in a compound task. CHI, (1994).
20. Kane, S.K., Avrahami, D., Wobbrock, J.O., Harrison, B., Rea, A.D., Philipose, M., La-Marca, A. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. UIST, (2009).
21. Lane, D.M., Napier, H.A., Peres, S.C., Sandor, A. Hidden costs of graphical user interfaces: Failure to make the transition from menus and icon toolbars to keyboard shortcuts. International Journal of HCI 18, 2 (2005), 133-144.
22. Li, F.C.Y., Dearman, D., Truong, K. N. Virtual shelves: interactions with orientation aware devices. UIST, (2009), 125-128.
23. Lin, S., Su, C., Cheng, K., Liang, R., Kuo, T., Chen B.-Y. PUB-Point upon body: Exploring eyes-free interaction and methods on an arm. UIST, (2011), 481-487.
24. Malone, T.W. How do people organize their desks?: Implications for the design of office information systems. ACM Transactions on Office Information Systems. 1, 1 (1983).
25. Mine, M.R., Brooks Jr., F.P., Sequin, C.H. Moving objects in space: Exploiting proprioception in virtual-environment interaction. SIGGRAPH, (1997), 19-26.
26. Rekimoto, J., Saitoh, M. Augmented surfaces: a spatially continuous work space for hybrid computing environments. CHI, (1999), 378-385.
27. Rekimoto, J. GestureWrist and GesturePad: unobtrusive wearable interaction devices. Wearable Computers, (2001), 21-27.
28. Robertson, G., Czerwinski, M., Larson, K., Robbins, D.C., Thiel, D., Dantzich, M.V. Data Mountain: Using spatial memory for document management. UIST, (1998), 153-162.
29. Strachan, S., Murray-Smith, R., O'Modhrain, S. BodySpace: inferring body pose for natural control of a music player. CHI EA, (2007), 2001-2006.
30. Weiss, M., Voelker, S., Borchers, J. Benddesk: Seamless integration of horizontal and vertical multi-touch surfaces in desk environments. Adjunct Proceedings ITS, (2009).
31. Wellner, P. Interacting with paper on the DigitalDesk. Comm. of the ACM 36, 7 (1993).
32. Wickens, C.D., McCarley J.S. Applied Attention Theory.
33. Wimmer, R., Hennecke, F., Schulz, F., Boring, S., Butz, A., Hussmann, H. Curve: revisiting the digital desk. NordiCHI, (2010), 561-570.