# Gaze-Adaptive Above and On-Surface Interaction

**Barış Serim**[1,3], **Khalil Klouche**[1,4], **Giulio Jacucci**[1,2]
[1] Helsinki Institute for Information Technology HIIT, Dept. of Computer Science, University of Helsinki
[2] Helsinki Institute for Information Technology HIIT, Aalto University
[3] Department of Design, Aalto University, [4]Media Lab Helsinki, Aalto University
firstname.lastname@helsinki.fi

## ABSTRACT

We explore the combination of above-surface sensing with eye tracking to facilitate concurrent interaction with multiple regions on touch screens. Conventional touch input relies on positional accuracy, thereby requiring tight visual monitoring of one's own motor action. In contrast, above-surface sensing and eye tracking provides information about how user's hands and gaze are distributed across the interface. In these situations we facilitate interaction by 1) showing the visual feedback of the hand hover near user's gaze point and 2) decrease the requisite of positional accuracy by employing gestural information. We contribute input and visual feedback techniques that combine these modalities and demonstrate their use in example applications. A controlled study showed the effectiveness of our techniques for manipulation tasks against conventional touch, while the effectiveness in acquisition tasks depended on the amount of mid-air motion, leading to our conclusion that the techniques can benefit interacting with multiple interface regions.

## Author Keywords

Eye tracking; gaze interaction; above surface interaction; multi-touch

## ACM Classification Keywords

H.5.2. Information interfaces and presentation: Input devices and strategies

## INTRODUCTION

Large multitouch screens allow designers to create wide workspaces that provide direct and concurrent access to UI widgets: users can access commands or information without any additional interaction steps. However, single-focus human visual attention remains a bottleneck for concurrent access. Although wide or distributed workspaces come with the benefit of displaying many UI elements, they can divide users' visual attention between distant interface regions where the cost of redirecting the gaze is high. This is especially the case for precise pointing tasks, as the positional accuracy of touch input depends on users visually monitoring their own motor actions [12, 25].

Previous research has shown that additional input modalities can decrease the need for positional accuracy by sensing more than the touch position. One such modality is *above-surface sensing* of hand posture, position and gesture. Above-surface information has been used to discriminate between different commands that a touch input could be intended for, thereby expanding the functional vocabulary of touch actions [7, 14]. Above-surface sensing also holds promise for decreasing the difficulty of pointing tasks on large screens by adapting the interface in anticipation of touch [2, 37]. In parallel, *eye tracking* has very recently been employed to address the challenge of limited visual attention on touchscreens. Previous work in this domain compensates the lack of visual monitoring through flexible input handling [12, 25] and gaze-adaptive visual feedback [25].

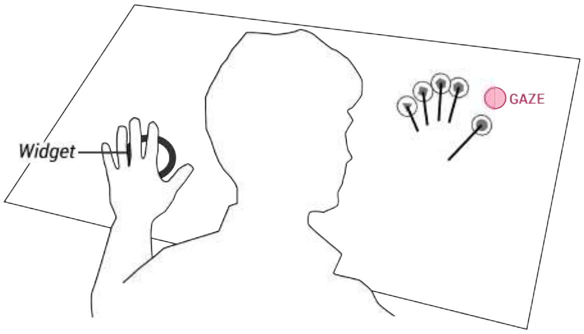


**Figure 1. We use the combination of above-surface sensing and eye tracking to facilitate direct input, even when the user in not visually attending to the target. The system determines potential user actions through above-surface sensing, but defers their confirmation to touch input, accompanied by gaze-adaptive visual feedback between these two steps.**

In this paper, we explore the combination of above-surface sensing with eye tracking to partly overcome the limitation of single-focus visual attention and facilitate concurrent interaction with multiple interface regions. Though each has shown individual promise, the combination of these modalities has not been studied. Our motivation for their combination is the new design possibilities they lend to supporting concurrent interaction. Together, above-surface sensing and eye tracking allow us to understand how users' hands and gaze are distributed across the interface and adapt the interface when the hands are further from user's gaze. Our interaction techniques address the aforementioned challenges of a) reliance on positional accuracy and b) limited visual attention under the two components of *input handling* and *visual feedback*.

**Input handling:** The interactive affordances of input widgets predispose hand posture and eye movements prior to the actual

*DIS 2017*, June 10-14, 2017, Edinburgh, United Kingdom.

ACM 978-1-4503-4922-2/17/06...$15.00
DOI: http://dx.doi.org/10.1145/3064663.3064744

touch contact. Above-surface sensing and eye tracking allow the system to capture this pre-touch ([1, 14]) information, which we use to discriminate user actions (e.g. based on hand posture), without extensive reliance on input position.

**Visual feedback:** We utilize above-surface sensing and eye tracking to enable visual monitoring of multiple interface locations. This is accomplished by warping the interface contents hovered by hands to where the user's gaze is directed at, thus enabling interaction with distant interface locations without having to redirect the gaze. This allows for visual juxtaposition of a UI widget with an interface region.

Our main contributions are as follows:

- Our work contributes, to our knowledge, the first combination of above-surface sensing and eye tracking with touch. We formally describe the design considerations for the combination of these modalities to support concurrent interactions with multiple interface regions.
- We developed novel interaction mechanisms using these modalities. For input, we use hand posture to assign interactive widgets and finger-mapped touch actions for confirmation. These are complemented by visual feedback techniques that adapt the position, timing and visual aspects of the feedback using eye fixation coordinates and above-surface sensing. We demonstrate these techniques in example applications.
- Finally, we tested the efficacy of warped feedback for acquisition and manipulation tasks against a baseline condition of non-warped feedback that requires redirecting the gaze. The results show that for manipulation tasks, the performance of warped feedback was significantly better than baseline, while the results were comparable for within-widget acquisition tasks. On the other hand, the performance of warped feedback significantly suffered when the participants had to switch between widgets through midair movement.

Based on these findings we conclude that the proposed interaction techniques can be used to complement conventional visually-monitored motor actions, with warped visual feedback employed for inputs that require minimal midair movement and visually monitored motor actions employed for large midair movements by hand.

## RELATED WORK

Touchscreens, which lack any tactile cues, require users to visually monitor their own action for positional accuracy. This has led to various strategies to support eyes-free input on touch surfaces, such as augmenting them with tactile widgets [34] or directing the finger to predefined locations using magnetic attraction [33]. In contrast, non-haptic solutions take advantage of the dynamic adaptation of motor and visual spaces afforded by touchscreens. In this section, we review the use of above-surface sensing and eye tracking as two modalities to decrease reliance on positional accuracy or visual monitoring during input.

### Combining Gaze and Touch Input

The potential of gaze as a real-time input is being investigated in an increasing number of settings and in combination with other modalities. A common strategy is to take advantage of the rapid change of spatial context afforded by eye movements for gains in motor performance during selection. This has been the motivation for using gaze as a cursor on large and distant displays on which direct touch input is impractical [27, 29]. In these situations, eye fixations provide the input position whereas touch input on a handheld device confirms the action.

On the other hand, there has been a recent interest in combining manual input with gaze on the same surface. Common in this work, is the use of the distance between the gaze point and touch (or any other manual input type) for the flexible handling of touch. However, they differ in terms of how the situations in which touch and gaze misalign are handled. One approach is to utilize touch as an indirect and gestural input and use gaze point position for selection instead [18, 19]. This assumes visual attention on the location of input and has been motivated by the need to decrease the amplitude of motion (i.e. distance traveled by the hand), parallel to MAGIC pointing that cascades manual and gaze input [38].

Another approach to combining gaze and touch is preserving touch as a direct input but distinguishing between accurate and ambiguous touch actions by incorporating gaze point information. The main motivation in this approach has been facilitating input with decreased visual monitoring for concurrent access to multiple objects [25] or high-throughput interactions such as typing on a touchscreen while looking at the text field [12]. Both cases exemplify situations in which touch input is performed further from the location of visual feedback. They demonstrate flexible input handling mechanisms based on the level of visual monitoring with which the touch action is conducted. Decreased visual monitoring results in an expanded area for potential selection and subsequent delegation of control to the system for decision-making. The flexible input handling is accompanied by various techniques such as providing visual feedback at the periphery or translating the visual feedback coordinates to where the user's gaze is directed [25].

The aforementioned work also varies regarding how it operationalizes visual monitoring. The distance between the gaze and touch point can be utilized as a discrete distance threshold for classifying direct and indirect touch [18, 19] or as a continuous scale to determine positional uncertainty [12, 25]. However, common among them is input handling and visual feedback at the moment of touch contact. This poses a limitation because touch and gaze information at the moment of touch does not account for the complex hand-eye behavior that leads to a touch event.

### Above and On Surface Interaction Continuum

Above-surface sensing promises to extend input handling and visual feedback processes to pre-touch. Earlier work used above-surface sensing in various ways, ranging from deliberate midair input actions [10, 13, 20] to implicit use of pre-touch movements by the system [14, 36]. Within previous work, we focus on a subset that use above-surface sensing in continuum with touch input [17] rather than in isolation. Research in this direction already targets relaxing the requirements of positional accuracy using various strategies.

One strategy for relaxing the requirements of positional accuracy has been adapting the motor space of the interface through target expansion [2, 37]. TouchZoom [37] uses the proximity of fingers to the screen to increase toolbar and icon target sizes before touch. Similarly, the proximity of the finger to the screen has been used to expand targets for in-vehicle interfaces that need to be operated with little visual monitoring, as visual attention is reserved for driving [2].

Above-surface sensing has also been used to extend touch functionality. Touch actions have been mapped to different interface commands based on various pre-touch or after-touch midair gestures [7]. The strategy that is most closely aligned with ours is that of Hinckley et al. [14] who utilize above-surface sensing (including grip) primarily as a pre-touch modality to provide anticipatory visual feedback and distinguish between the different commands a touch action can be intended for on a mobile device. However, instead of small and hand-held devices, we focus on interaction with large touch surfaces, which point to significantly different design considerations. We discuss these differences below.

### Comparison to Previous Work

By combining gaze and above-surface sensing, we build on and advance the state of the art for these two input modalities.

When combining gaze and touch, we use touch primarily as a direct input, even in the cases where touch and gaze points misalign. This is in contrast to previous work that combines gaze with indirect manual input when touch is performed further away from gaze [18, 19]. We also further previous work that translates visual feedback near user's focus of visual attention [25]. In previous work, the visual feedback has been limited to the moment of touch contact and more importantly *does not leverage the multitouch capability of the human hand*. We use above-surface sensing to detect the proximity of not only the hand or a single finger tip, but multiple fingertips. In turn, this allows the system to visualize possible commands that can be triggered through touch with different fingers. This is achieved by showing a simplified representation of the hand that shows command-to-finger mapping near where the user is gazing.

In the domain of above-surface sensing, our main contribution is the extension of pre-touch to multifocus interactions on larger touchscreens. These interactions differ from handheld pre-touch [14] in at least three ways. First, hand grip becomes less relevant for larger screens, while the projected position of the hand over the interface (hover) gains relevance. Second, multifocus interaction on large screens causes the user's visual attention to be split between more distant interface regions, which can be detected through eye tracking for the purpose of input handling. Third, larger surfaces can accommodate the entire human hand, thereby allowing pre-touch visual feedback for all fingers.

Much effort in interaction techniques for large screens (e.g. [3, 16]) and gaze pointing [38, 18] has been directed to increasing the pointing performance for distant targets by decreasing the amplitude of motion. In contrast, we target situations in which the required motion by hand is not large but the cost of redirecting the gaze can be high. While the definitions of large touchscreen vary, the considerations of split visual attention and accommodation of both hands define the scope of our work. Indeed, previous work [12, 25] shows that input accuracy suffers when the gaze is away even for screens of 27" or smaller due to the very limited area where the vision is sharp. In the next section, we motivate our combination of these modalities and show how they enable interactions that are not possible with eye tracking or above-surface sensing alone.

## DESIGN CONSIDERATIONS

Pointing to interface elements involves gradual alignment of fingertips to a visual target, a closed-loop control process that starts with midair motion and finalizes with the end of touch. Figure 2 summarizes this process and our design space by showing the progression of a command event in three modalities. It should be noted that our focus is on using the above-surface as a pre-command modality for selecting widgets and deferring the confirmation of action to touch. Thus, the figure excludes mid-air gestures that do not yield to a touch event as in gaze and gesture combinations [6] or after-touch gestures [7].

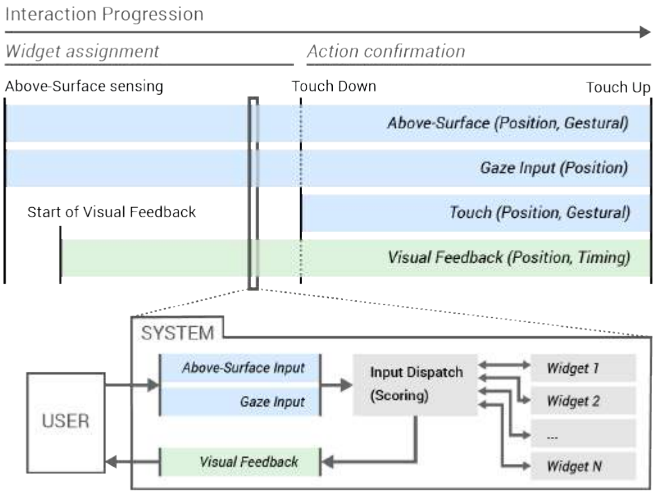


**Figure 2. Continuous feedback control loop showing three input modalities during input. UI widgets are continuously ranked before a touch action, while the system provides the visual feedback of the selected widget (top). The widget assignment process (bottom) is modeled after Schwarz et al. [24].**

To assign widgets to hands, we employ the uncertain input handling framework developed by Schwarz et al. [24], but adapt it to above-surface interaction. An event dispatcher continuously scores every widget-hand pair ($score_{ih}$) based on pre-touch information. Widgets can have multiple scoring criteria ($j$), such as position or hand posture, that are all scaled to a uniform range (0-1). Not all scoring criteria is equally relevant to a widget, which results in different scoring *weights*. As a preliminary method, we calculate the overall score of a widget by taking the weighted geometric mean of individual scores and assign the highest scoring widget to a hand ($h$):

$$score_{ih} = \left(\prod_{j=0}^{n} score_{jh}{}^{weight_j}\right)^{1/\Sigma_{j=1}^{n} weight_j}$$

The information of the assigned widget is communicated back to the user through visual feedback. The visual feedback differs based on the above-surface sensing and gaze conditions (Figure 3). *Minimal* refers to the visual representation in which a widget is unassigned and is minimally represented on the interface, to signify its location and current state. In *hover* condition, a hand or a finger is assigned to a widget at the same time as the user's gaze is directed to the widget. In this state the visual representation of the widget corresponds to the motor space. Finally, in the *warp* condition, a widget is assigned to a hand or finger, but the user's gaze is directed elsewhere. In this case, while the motor space remains the same, the visual feedback is warped to where the user's gaze is directed.

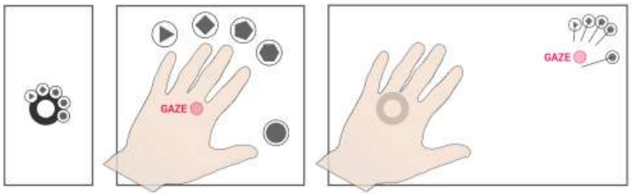


**Figure 3. Three different visual states of a widget. Left: Minimal visual feedback in the absence of an assigned hand. Middle: Hover visual feedback when the widget is assigned and user gaze is directed on the widget. Right: Warped visual feedback at the gaze location when hands are above but gaze is directed elsewhere.**

We made two design decisions to maintain robust interaction during continuous scoring and visual feedback. First, if a widget is already assigned to a hand or finger, it is scored slightly higher. This eliminates jittery alternations between two widgets in borderline situations. Second, when a hand or finger that is assigned to a widget touches the screen, it remains assigned to that widget until the touch is released regardless of its score.

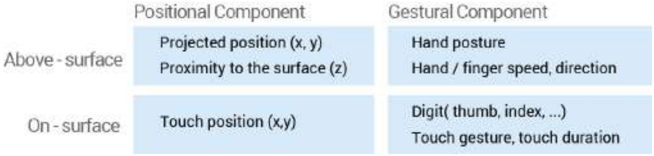
### Above and On-surface Input

Cognitive studies of motor control, distinguish between *proximal* (i.e. getting the hand near the target) and *distal* (i.e. shaping the hand in anticipation of the action) components of manual action [15]. In HCI, they correspond to the *positional* and *gestural* components of the user input. Figure 4 shows their breakdown into above- and on-surface modalities.

It is important to understand that the positional and gestural components are dissimilarly affected by low visual monitoring. Hand posture and relative finger positions are known to the user through proprioception, whereas positional accuracy requires the user to monitor where the hand or finger is located relative to the target. For input handling, the main design principle that guides the interaction techniques is replacing or complementing the positional component of input with the gestural component when possible.

#### *Widget assignment based on hand posture*

The interactive affordances of various touch actions favor certain above-surface hand postures, which we use for scoring widgets. For example, a widget that requires a pinch gesture can be scored higher for hand postures that feature the thumb and index as the only extended fingers. As we use above-surface sensing as a pre-touch modality, different finger configurations form the basis of various hand postures that we have defined (Figure 5). We use the data provided by a skeletal tracking software (Leap Motion) to calculate how bent each finger is and identify the best fitting hand posture (none if there is no fitting posture). The identification is done by assessing each finger for the designated range of the candidate posture.



| | Positional Component | Gestural Component |
|---|---|---|
| Above - surface | Projected position (x, y)<br>Proximity to the surface (z) | Hand posture<br>Hand / finger speed, direction |
| On - surface | Touch position (x,y) | Digit( thumb, index, ...)<br>Touch gesture, touch duration |

**Figure 4. Breakdown of above and on-surface modalities into positional and gestural components.**

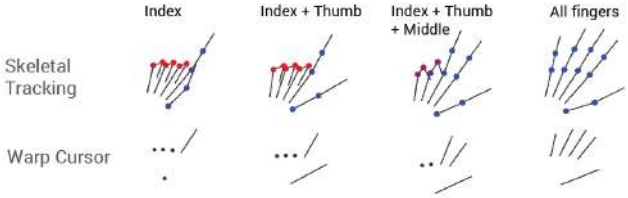


**Figure 5. The system tries to match a predefined hand posture using skeletal tracking (above). The warped visual feedback shows a simplified representation of the hand (below).**

An interface region can be populated with multiple widgets that favor different hand postures. In these situations, we use hand-posture information in addition to the positional component (i.e. proximity of the widget to the projected hand position) to score widgets (Figure 6). However, the same mechanism can also be applied to purely gestural interactions, in order to score widgets based on the gestural component alone.

**Figure 6. Hand posture can be used to discriminate between widgets in addition to the projected position of the hand or finger. Here, hand posture information is used to select between two widgets that are operated differently.**

#### *Finger-mapped touch actions*

Another example of replacing positional component with the gestural component is finger-mapped touch actions. This interaction technique is accomplished in two steps. Once a widget is assigned to a hand, the system visually notifies the user by showing available command options that are mapped to individual fingers. At this stage, what determines action is the specific digit (e.g. thumb or index finger) that performs the touch action rather than the touch position (Figure 7).

This technique is related to previous efforts to design touchscreen interactions with the particular physical qualities of the human hand in mind, such as number of fingers [4, 9, 30]. For example, HandMark menus [30] provide access to menu elements upon resting the hand on the screen. The selection is then accomplished by tapping on the relevant item with the other hand. We use the above-surface modality to eliminate the need of confirmation by a second hand. Instead, available options are shown before the actual touch and are then confirmed through touch. More than a single command can be mapped to a digit. In this case, additional commands are accessed by dragging the finger to the additional commands before releasing it.

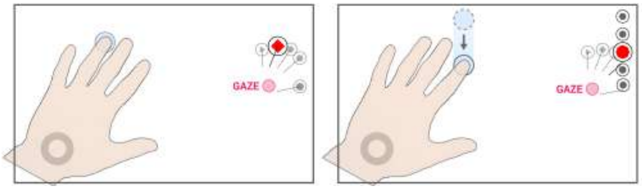


**Figure 7. Various commands are associated with individual fingers when a hand is assigned to a widget. Available commands that correspond to fingers are visualized prior to the touch. Multiple commands that are mapped to the same finger can be accessed by dragging the finger (right).**

## Eye Tracking

We use eye tracking not only for determining the position of the visual feedback, but also for input handling. Eye tracking data is noisy and eye behavior is unstable, so we rely on fixation points as operational gaze points. The fixations are calculated with the commonly used dispersion-threshold identification algorithm, I-DT (implementation in [21]), with the minimum time window set to 90 milliseconds and the dispersion threshold set to 1°.

### Determining the weight of the position score

We use gaze point information for scoring the widgets. When scoring widgets, positional and gestural components can *reinforce* each other, for example, when a thumb and index finger hand posture is in the proximity of a virtual knob. However, the two components can also *contradict* each other if the same hand posture is in the proximity of an interactive widget that is operated by tapping.

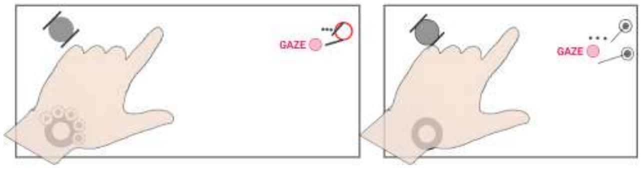


**Figure 8. The weight of positional component is manipulated based on the degree the action is visually monitored. The relative weight of the gestural component (hand posture) increases when the user gaze is directed elsewhere.**

In these cases, we employ eye tracking to determine the weight of the positional component for calculating the score (Figure 8). The weight is continuously re-evaluated based on the distance between the eye fixation and the projected position of the hand on the screen. The weight of the position is higher when the user's gaze is located near the widget and lower when it is further away.

### Gaze as pointer interactions

The design space can also be extended to situations, in which gaze is used as a pointer for selection in combination with touch input (*gaze + touch interactions*). For example, a manipulation action on an object can be performed at a single step by keeping the gaze on the object and determining the action by touching on a widget. In other words, rather than following a sequential order for selection and manipulation, gaze and manual input are used concurrently. What distinguishes this technique from earlier work that combines indirect touch and gaze [18, 19] is the use of touch as a direct input and concurrent pointing by touch and gaze.

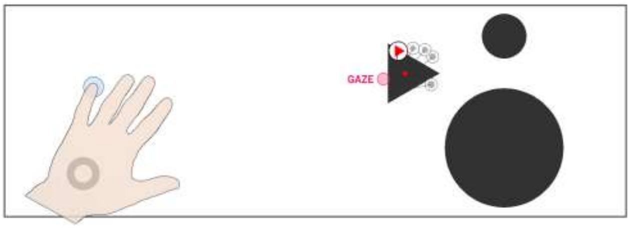


**Figure 9. Gaze coordinates can be coupled with manual input to perform *gaze + touch* interactions. In the figure the hand is over a widget and the finger-mapped direct touch is used to modify the shape of an object.**

## Visual Feedback

Unlike hovering with a cursor, touch input lacks a pre-command notification stage. Above-surface sensing has been proposed as a modality for providing informative visual feedback prior to the actual touch [8, 17]. Input with low visual monitoring can benefit from the timely communication of the system's interpretation of the user input (assigned widget) back to the user. Above-surface interaction and eye tracking call for a reconsideration of design choices regarding the *timing* and *positioning* of the visual feedback.

### Timing

While above-surface sensing enables pre-touch visual feedback, feedback that occurs earlier than expected can be intrusive. Similarly, appropriate withdrawal of the visual feedback is necessary when it is not needed. To prevent earlier-than-expected feedback, we use proximity to the surface (z) and hand velocity variables. Visual feedback is shown only when the hand is in the proximity of the screen (z) and the projected hand velocity (on x-y plane) is low.

We additionally rely on widget interaction events to correctly time the visibility of the feedback. As a principle, the system changes the visibility for continuous events gradually and for discrete events instantly (Figure 10). Assignment of a widget to a hand, or a touch-down action on a widget maximizes the visibility of the widget. The feedback starts fading out in case of inaction. Conversely, above-threshold hand movement causes the warped feedback to gradually gain visibility. The warped feedback is fully visible as long as the widget is being touched by a finger, whereas releasing all the fingers from a widget causes a sudden drop in visibility, followed by a gradual fade out. The visibility can be mapped to different visual channels such as opacity or size.

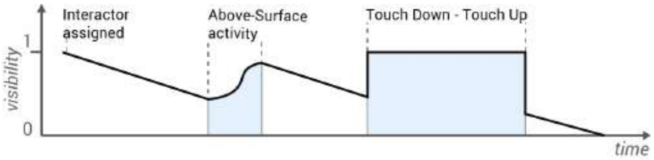


**Figure 10. The relative visibility of a widget in relation to the interaction events. The visibility gradually decreases during episodes of inaction.**

### Position

We use eye fixation coordinates to determine the position of the warped visual feedback. However, eye fixations shift rapidly, and constant warping of the visual feedback can be intrusive. Thus, we update the position of the warped visual feedback upon interaction events (such as when a widget is assigned), upon touch actions and above-threshold mid-air movement. Also, we update the position of the feedback only when the distance between the current position of the visual feedback and the most recent fixation is above a certain threshold ($\approx 10^\circ$ of visual angle from eyes).

### Summary

We described how gestural component can partially replace position during widget assignment and action confirmation. We also described how eye tracking can be used both for input handling and visual feedback. The next section describes example applications that demonstrate various combinations of the aforementioned techniques.

## APPLICATION EXAMPLES

### Apparatus

The applications were prototyped for a 10 finger multitouch screen (27”, 2,560 × 1,440 pixels) that was tilted 30° to enable easier hand reach for touch input. For above-surface sensing, a commercial short range infrared sensor for skeletal hand tracking (Leap motion) was used. We mapped the coordinate system of above-surface depth sensing to screen coordinates using a 4 point calibration procedure. The eye tracking was performed by Pupil Labs binocular tracking glasses running at 60Hz. The applications consistently ran at 60fps, with the latest gaze data point synchronized at every frame.

### Object Drawing and Manipulation

We prototyped an application that enables parametric manipulation of a group of objects or adding new objects on a canvas. Objects have multiple parameters, namely shape, fill color, border color, size and orientation. The application demonstrates how these parameters can be modified without redirecting the gaze between the toolbar and canvas areas through above-surface sensing and warped visual feedback. Besides allowing continuous fixation on the canvas, warped feedback allows for the juxtaposition of a widget with a specific region on the canvas. For example, warped feedback of the color selection tool on the canvas can facilitate easier visual comparison of the selected color with the already existing colors on the canvas. The application features the following widgets:

*1) The shape selector* allows switching between different geometric forms. Basic shapes such as triangle, square, pentagon and circle are selected through tap actions with different fingers. Additionally a larger set of polygons can be accessed by dragging the index finger on the surface. Given that all the fingers can be used for interaction, the widget is scored higher when all the fingers of the hand are extended.

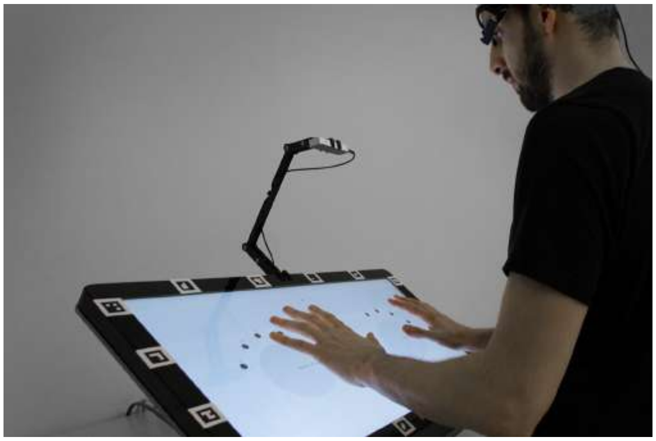


**Figure 11. The hardware setup consisting of a touch screen, depth sensor (Leap Motion, attached to the upper screen edge) and eye tracking glasses.**

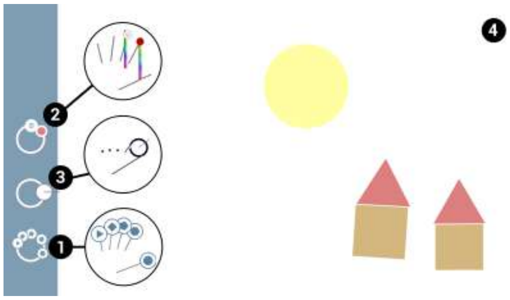


**Figure 12. Drawing application featuring the shape selector (1), the color selector (2), the physical modifier (3) and the canvas (4). Warped feedback visuals of the widgets 1,2 & 3 are annotated.**

*2) The color selector* selects or modifies object colors by dragging a finger on a two-dimensional color space. Index and middle fingers are mapped to fill and stroke colors, respectively. The widget is scored higher for hand postures that feature all or the first three (thumb, index, middle) fingers extended and when the hand is in the proximity.

*3) The physical modifier* manipulates the size and orientation of a group of objects through pinch and two-finger (thumb + index) rotation. Besides positional proximity, the widget is scored higher for hand postures that only feature extended thumb and index fingers.

*4) The canvas* allows users to add new objects to the scene. When the user’s hand is hovering over the canvas, the system displays the shape that will be drawn upon touch. The canvas is scored based on the projected position and orientation of the hand. The widget is scored higher if the thumb is located to the left of the palm center. This allows for discrimination between other widgets on the toolbar in borderline conditions.

### Real-time Video Manipulation

We prototyped a simplified VJ (video jockeying) tool that enables the user to apply a variety of visual filters to a video loop. Real-time video editing offers a relevant challenge, as the main feedback for most inputs results in real-time visual changes to the video picture away from the input location. In that case, being able to display essential interaction information at the gaze location allows users to keep their attention on the target when it is most critical. The different widgets either appear as soon as the user's hand is detected, or inconspicuously suggested on the sides of the display as perceived affordances. The application features the following widgets:

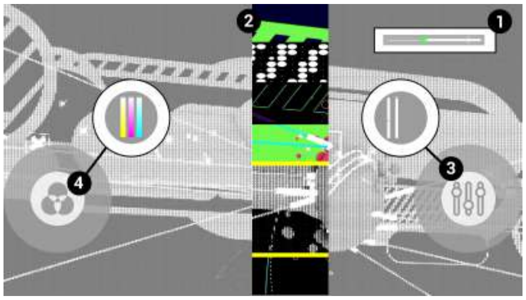


Figure 13. Vj application featuring loop control (1), the filter selector (2), the filter modifier (3) and the color modifier (4). Warped feedback visuals of the widgets 1,3 & 4 are annotated.

*1) Loop control* is a global widget that can be accessed from anywhere on the screen, as its score is solely based on hand posture. A pinch posture assigns the widget to the hand. Visual feedback of the playback information is shown near the gaze fixation. The loop ranges are then selected by dragging the index finger or the thumb.

*2) The filter selector* is scored based on the position information and appears when the user extends her hand towards the middle of the screen. The video picture then gets divided into multiple horizontal areas, each displaying a preview of available visual filters. The selection is then accomplished by touch.

*3) The filter modifier* and *the color modifier (4)* are located on the right and left sides of the screen and are scored based on position. The parameters are mapped to one specific finger each – i.e. index for parameter 1, middle for parameter 2, ring for parameter 3, and pinkie when a fourth parameter is needed. Approaching the widget with the hand will show gaze-adaptive feedback for the parameters, as vertical gauges, with cursors indicating the current value between the maximum and minimum allowed for each gauge. The locations of color and filter modifiers allow concurrent bimanual modification of both widgets.

## STUDY: WARPED FEEDBACK AND DIVIDED ATTENTION

We set out to test the viability of warped visual feedback for situations in which attention needs to be divided between multiple interface regions. Previous work identifies two interaction stages, namely acquiring and manipulating a control device [11, 28]. Thus, we prepared a two-part experiment that involves acquiring a target on a touch surface and manipulating a widget through on-surface interaction. We were interested in seeing how the task completion times in these two stages vary for the conditions of:

1. Warped visual feedback condition, which facilitates continuous gaze fixation near the stimulus position using a smaller representation of hand (scaled down by a factor of 0.35 to be visible and less intrusive).
2. Baseline condition, which provides no specific support to facilitate continuous fixation.

### Apparatus

The same hardware setup described earlier has been used, with a 27", tilted touchscreen (set to 1920 × 1080) for input. Additionally, The experimental setup involved a second, vertically positioned monitor of the same size placed on the upper edge of the first monitor (Figure 14). The combination of a horizontal input surface with a vertical monitor have been investigated before [5, 31, 32], and has the advantage of showing visual feedback at the eye level while providing arm support on the horizontal surface. In the experiments, the two monitor setup was used to display the target stimuli either in one of the following locations:

1. Near the left edge of the same tilted screen on which the input is performed (same screen condition). This resulted in a visual angle of around 50° degrees between the input and stimuli positions depending on the distance of eyes to the screen.
2. Near the upper-left corner of the vertical screen. This resulted in a visual angle of around 70° degrees between the input and stimuli positions.

The main motivation for using two screens was to compare the effect of different distances between target stimuli and input area expressed in visual angles. Different distances between the target stimuli and input area corresponds to different costs for redirecting the gaze, which can be be generalized to other display setups.

The experimental setup also controlled the warped feedback location by always displaying the warped feedback near the stimulus. This was implemented to prevent users from shifting their gaze to the input area (for warped feedback condition). However, eye tracking data was collected from participants for later analysis.

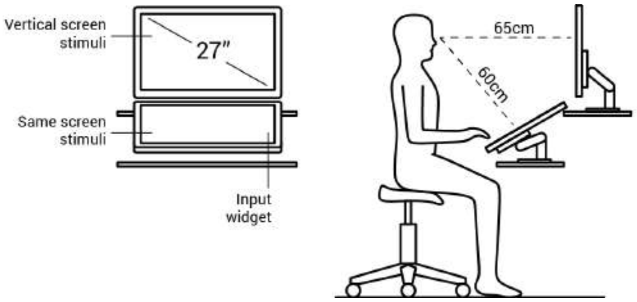


Figure 14. The front and right views of the experimental setup. The screens were positioned approximately 60 and 65 cms away from the eye.

### Participants

Twelve participants (2 female), aged 20 to 33 (m=26.1, sd=3.4) were recruited for the study. Of the participants, eleven were right handed and all reported extensive familiarity with touch devices. The participants have been compensated with one cinema ticket and their informed consent has been collected for data logging.

### Procedure

We performed a within-subject study of two interaction conditions (warped feedback and baseline) and two stimulus location (same or vertical screen) conditions. The order of the interaction and stimulus conditions alternated between participants using a Latin square design to counterbalance the potential effects of learning and fatigue. The participants were seated in front of the setup, with the input and vertical screens positioned approximately 60 and 65 cms away from the eye. The participants performed the acquisition and manipulation experiments in sequence. The participants had a chance to rest between experiments. The sessions lasted 40 to 50 minutes.

### Analysis

The initial trials of each condition (1 block for acquisition tasks, 2 blocks for manipulation) were excluded from the analysis. The metric for evaluation is the task completion time, which we report as grand mean and the grand median, the latter being more robust to outliers. The baseline and warp conditions have been compared using Welch's t-test and normalized time completion values based on subtracting the means of each participant from the data divided by standard deviation. Error percentages represent the ratio of the erroneous touch releases to all touch releases.

### Experiment 1: Widget Acquisition

With this task, we aimed to compare the combination of warped feedback and finger-mapped input versus regular direct touch actions. The experimental interface (Figure 15) consisted of 3 vertically arranged widgets (each with 4 targets "1,2,3,4", "A,B,C,D" and "+,-,*,/" from left to right).

Each widget covered a 400*120 pixel area of the screen. In the baseline condition this area was vertically divided into 4 pixel buttons that are each 100*120 pixels and selected through positional input. In contrast, for the warped feedback condition, the participants acquired a widget when their middle finger hovered within the 400*120 pixel area and selected any of the 4 targets respectively through thumb, index, middle and ring fingers rather than touch position. Thus, the spatial footprints of the widgets were the same for both input conditions, but the interaction differed regarding the selection of 4 targets within the widget (positional or finger-mapped) and the position of feedback (no warped feedback and warped feedback).

Each experimental block involved alternating between three widgets in randomized order and acquiring targets in the same widget in randomized order, which resulted in 12 trials (3 widgets × 4 targets). With this procedure, we aimed to observe the performance of acquisition task between-widgets (the first acquisition task after the widget is alternated) and for acquisition within-widgets.

We anticipated time savings by eliminating attention switches under warp condition. However, we also anticipated a decrease in motor performance in warped feedback condition due to lack of visual monitoring. Previous work reports a decrease in accuracy when the hand movement is not visually guided [26, 35]. More specifically, Schmidt et al. report decrease in performance for large amplitude bimanual tasks when visual feedback is separated from the motor space [23]. However, the cumulative effect of redirecting the gaze and motor performance has not been studied. Even we target relaxing the need for positional accuracy, we were interested in seeing the viability of positional input through warped visual feedback, which corresponds to between-widget acquisition tasks.

Each trial was completed upon a touch release on the correct target. The combination of 2 input methods × 2 stimuli positions × 10 blocks (excluding the training block) × 12 targets resulted in 480 trials per participant. Before the experiment the participants practiced the finger-mapped input condition with 60 trials.

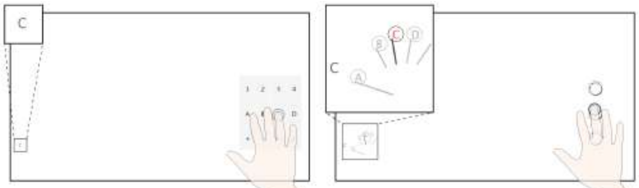


**Figure 15. In the baseline condition (left) the target was acquired through touch position, while in the warped feedback condition (right) it was acquired through finger mapped input after the widget is assigned. Details are magnified. The stimuli character height was 9.3mm.**

#### *Results*

Early in the analysis, we noticed time completion differences for between within-widget and between-widget tasks and analyzed them separately. For within-widget tasks, the mean completion times for warped feedback and baseline conditions were similar in both stimuli conditions (Table 1). A t-test comparison using within-subject normalized completion times did not show any significant effect for the same screen (r = .03, p = .23) and vertical screen (r = .03, p = .17) conditions. The error rates were higher for the warped feedback condition (Table 1). In few instances participants reported system failures for identifying the wrong finger, which might have affected the error rate for warped feedback condition.

However, the results were significantly different for between-widget acquisition tasks (Table 1). On average, participants spent significantly more time on the warp condition than on the baseline condition for both the same screen (r = .42, p < .001) and vertical screen conditions (r = .54, p < .001). This was accompanied by a much greater error rate for warped feedback condition (Table 1). Participant interviews suggested a few possible explanations for the difference in performance between the two input conditions. One is the mismatch between the distance traveled by the hand and its scaled down visual representation in warped visual feedback. During between-interaction tasks, we informally observed that in the baseline condition participants performed high-speed ballistic movements towards the touch target, while in the baseline condition, Acquisition task *(within-widget)*

| Screen | Technique | Median(ms) | Mean(ms) | Error |
|---|---|---|---|---|
| Same | Warped | **902.75** | **1031.54** | 6.41% |
| | Baseline | 961.75 | 1068.36 | **2.43%** |
| Vert. | Warped | **885.50** | 1055.00 | 8.16% |
| | Baseline | 986.25 | **1021.88** | **3.31%** |

Acquisition task *(between-widget)*

| Screen | Technique | Median(ms) | Mean(ms) | Error |
|---|---|---|---|---|
| Same | Warped | 2094.00 | 2305.19 | 11.33% |
| | Baseline | **1164.50** | **1406.79** | **1.09%** |
| Vert. | Warped | 2258.75 | 2342.36 | 9.09% |
| | Baseline | **1207**.5 | **1415.19** | **2.17%** |

**Table 1. The grand median and grand mean completion times and overall error rates for two interaction and two stimuli conditions for the acquisition tasks. Emphasis (in bold) represents better performance.**

they moved their hand parallel to the screen, keeping a tense hand posture. Participants also reported shoulder fatigue for warped feedback condition, which may have been caused by parallel arm movements.

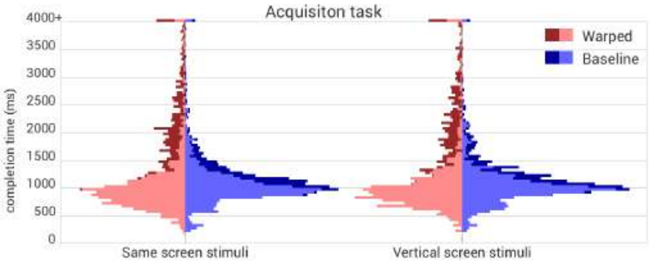


**Figure 16. Distribution of acquisition task completion times (50ms bins) across all participants for two interaction and two stimuli conditions. The shades and tints respectively indicate between-widget and within-widget acquisitions tasks.**

## Experiment 2: Widget Manipulation

The manipulation task required participants to match the value of an interactive slider to that of the stimulus. To isolate manipulation from acquisition, the interactive slider was always assigned to the participant's hand. Each trial required the participants to manipulate the slider by dragging the index finger on the touchscreen. If the touch was released when the value of the slider matches with that of the stimulus, the task was counted as complete. The slider consisted of 5 steps (the targets being the numerals "1,2,3,4,5"), and 15 pixels of movement resulted in 1 step. Unlike the acquisition task, the only difference between the baseline and warped feedback conditions was the location of the visual feedback. In the warped feedback condition the feedback was displayed near the stimulus, whereas in the baseline condition it was displayed near the widget. We expected that eliminating visual attention switches in gaze-adaptive condition would result in shorter task completion times. The combination of 2 input methods × 2 stimuli positions × 20 blocks (excluding training blocks) × 5 targets (shown in random order) resulted in 400 trials per participant.

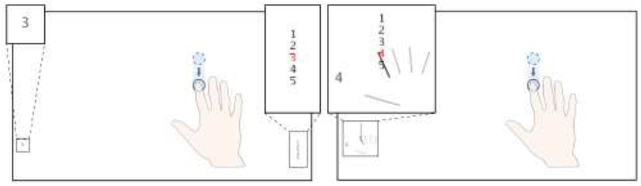


**Figure 17. In the baseline condition (left) the visual feedback of the manipulated slider was shown at the widget location, while in the warped feedback condition (right) it was shown near the target. Details are magnified. The stimuli character height was 9.3mm.**

### *Results*

For both stimuli conditions participants spent more time on the baseline condition than on the warped feedback condition (Table 2). A t-test comparison of the same and vertical screen conditions using normalized data yielded a larger effect size for the vertical screen condition ($r = .19$, $p < .001$) than for the same screen condition ($r = .10$, $p < .001$), in line with the expectation that the higher cost of redirecting the gaze in vertical screen condition will result in more pronounced benefits when using warped feedback. The error rates were lower for the warped feedback condition in both screen conditions (Table 2).

Manipulation Task

| Screen | Technique | Median(ms) | Mean(ms) | Error |
|---|---|---|---|---|
| Same | Warped | **1387.25** | **1576.30** | **4.07%** |
| | Baseline | 1530.25 | 1746.89 | 6.46% |
| Vert. | Warped | **1364.00** | **1509.06** | **3.69%** |
| | Baseline | 1616.50 | 1770.58 | 6.54% |

**Table 2. The grand median and grand mean completion times and overall error rates for two interaction and two stimuli conditions for the manipulation task. Emphasis (in bold) represents better performance.**

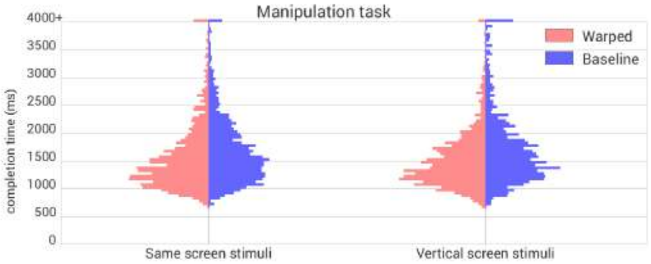


**Figure 18. Distribution of manipulation task completion times (50ms bins) across all participants for two interaction and two stimuli conditions.**

## DISCUSSION

We combined above and on-surface modalities with eye tracking to decrease the need for positional accuracy for input on touchscreens. This is enabled by discriminating between potential user commands through the gestural component of input and visual feedback the near gaze point location. The controlled study evaluated a subset of the interaction techniques but gave us valuable insights about the viable scope of warped visual feedback. Here, we would like to discuss the strengths and weaknesses of the interaction techniques based on the outcome of the user study.

### Midair Motion Amplitude

The performances of the warped feedback and baseline conditions were visibly different based on whether the task was manipulation, within-widget acquisition or between-widget acquisition. The performance of the warped visual feedback condition was higher for manipulation tasks that required no midair motion. The performance between warped feedback and baseline conditions were comparable for within-widget acquisition. However, the performance of warped visual feedback was significantly worse for between-widget acquisition tasks, in which participants had to acquire the widget through midair motion. Based on these results, we arrive at the following conclusions:

- The warped feedback was successful in decreasing the cost of redirecting the gaze, resulting in the improvement of task completion time for manipulation tasks.
- However, the warped feedback did not facilitate midair hand motion as effectively as direct visual monitoring, which resulted in a decrease in performance for between-widget acquisition tasks. Our results confirm earlier work by Schmidt et al. who reports lower performance and similar observations such as tense hand posture and parallel-to-screen hand movements when touch is performed without direct visual monitoring [23].

These conclusions also suggest that the interaction techniques described are strong in cases where the required midair motion is minimal and the task requires visual attention at a distant interface region. On the other hand, the conclusions suggest a weakness for inputs that require large amplitudes of midair motion. These have the following implications. First, in its current state, we see the role of warped feedback as complementary to conventional touch input. In these cases, the interface should allow for effortless transitions between warped feedback and conventional input modes by, for example, allowing acquisition of the widget through visual monitoring, but then transitioning to warped feedback for manipulation or within-widget acquisition. Second, for the interaction techniques presented, we see the main advantage of large touch surfaces in accommodating both of the user's hands to allow finger-mapped and bimanual concurrent input instead of expansive widget configurations.

### Warped Feedback and Visual Search

Our primary reason for implementing warped feedback was decreasing the redirection of gaze and juxtaposing two different regions. Warping a color palette to a canvas area is a good example of this, as it enables comparison of the selected color with a region on the canvas. However, there can be cases in which it is more beneficial to see multiple input widgets rather than warping the interface widget. Visual search of a command in the toolbar area is a good example of such a use case in which the user needs to see widgets side by side. To support these use cases, the minimal representation of the widgets should facilitate visual search and interface learning.

### Spatial and Motor Memory

Replacing positional accuracy with the gestural component (e.g. hand posture, or finger-mapped touch actions) also suggests an increased reliance on motor rather than spatial memory during interface learning. Although spatial and motor memories can be related (e.g. finger-mapped touch actions also correspond to different positions) the extent to which they are related is an open question in the research community [22]. Thus, the effects of spatial and motor memory for proficient use should be further investigated. Another important consideration is the number of interface position, hand posture and finger combinations that are both discernible and memorable.

### Gesture Input

The interaction techniques we presented rely on both positional (proximity of hand to the widget) and gestural information (hand posture and digit). However, we see potential application cases for pure gestural interactions. Above-surface sensing can be used to anticipate the action in advance and provide visual confirmation at the gaze location to inform the user of the possible action a touch event can lead to, thereby addressing the problem of learning and memorization in gesture-based interfaces.

## CONCLUSIONS

The limited nature of our visual attention stands in contrast to the concurrent action possibilities afforded by bimanual action. These concurrent action and sensing possibilities have been our motivation for supporting interactions that are distributed around multiple interface regions. We presented the design considerations for combining eye tracking with above- and on-surface modalities, in order to decrease reliance on positional accuracy during interaction. The system determined potential user actions midair, but deferred their confirmation to on-surface touch, accompanied by gaze-adaptive visual feedback between these two steps. The preliminary results from the controlled study suggest that warped visual feedback can complement conventional direct touch for manipulation tasks and acquisition tasks that require minimal midair motion, although future work can target improvements for increasing the viable scope. Besides immediate design implications, we identify these three modalities as a promising area for investigation, as their combination provides a fine-grained understanding of hand-eye coordination.

## ACKNOWLEDGMENTS

This work has been partly supported by MindSee (FP7 - ICT; Grant Agreement #611570). We would like to thank Oswald Barral for his help with the data analysis, the anonymous reviewers for their comments and to the participants of the study.

## REFERENCES

1. Michelle Annett, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2011. Medusa: A Proximity-aware Multi-touch Tabletop. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 337–346. `DOI:` http://dx.doi.org/10.1145/2047196.2047240

2. Ilhan Aslan, Alina Krischkowsky, Alexander Meschtscherjakov, Martin Wuchse, and Manfred Tscheligi. 2015. A Leap for Touch: Proximity Sensitive Touch Targets in Cars. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '15)*. ACM, New York, NY, USA, 39–46. `DOI:` http://dx.doi.org/10.1145/2799250.2799273

3. Patrick Baudisch, Edward Cutrell, Dan Robbins, Mary Czerwinski, Peter Tandler, Benjamin Bederson, Alex Zierlinger, and others. 2003. Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch-and pen-operated systems. In *Proceedings of INTERACT*, Vol. 3. IOS Press, 57–64.

4. Hrvoje Benko, T. Scott Saponas, Dan Morris, and Desney Tan. 2009. Enhancing Input on and Above the Interactive Surface with Muscle Sensing. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS '09)*. ACM, New York, NY, USA, 93–100. `DOI:`http://dx.doi.org/10.1145/1731903.1731924

5. Xiaojun Bi, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2011. Magic Desk: Bringing Multi-touch Surfaces into Desktop Work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2511–2520. `DOI:` http://dx.doi.org/10.1145/1978942.1979309

6. Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI '15)*. ACM, New York, NY, USA, 131–138. `DOI:` http://dx.doi.org/10.1145/2818346.2820752

7. Xiang Chen, Julia Schwarz, Chris Harrison, Jennifer Mankoff, and Scott E. Hudson. 2014. Air+Touch: Interweaving Touch & In-air Gestures. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 519–525. `DOI:` http://dx.doi.org/10.1145/2642918.2647392

8. Victor Cheung, Jens Heydekorn, Stacey Scott, and Raimund Dachselt. 2012. Revisiting Hovering: Interaction Guides for Interactive Surfaces. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS '12)*. ACM, New York, NY, USA, 355–358. `DOI:` http://dx.doi.org/10.1145/2396636.2396699

9. Daewoong Choi, Hyeonjoong Cho, and Joono Cheong. 2015. Improving Virtual Keyboards When All Finger Positions Are Known. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15)*. ACM, New York, NY, USA, 529–538. `DOI:` http://dx.doi.org/10.1145/2807442.2807491

10. Bruno De Araujo, Géry Casiez, and Joaquim Jorge. 2012. Mockup Builder: Direct 3D Modeling On and Above the Surface in a Continuous Interaction Space. In *GI'12 - 38th Graphics Interface conference*. Canadian Information Processing Society, Toronto, Canada, 173–180. https://hal.inria.fr/hal-00687891

11. George W. Fitzmaurice and William Buxton. 1997. An Empirical Evaluation of Graspable User Interfaces: Towards Specialized, Space-multiplexed Input. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '97)*. ACM, New York, NY, USA, 43–50. `DOI:` http://dx.doi.org/10.1145/258549.258578

12. Toshiyuki Hagiya and Tsuneo Kato. 2015. Probabilistic Touchscreen Text Entry Method Incorporating Gaze Point Information. *Journal of Information Processing* 23, 5 (2015), 708–715. `DOI:` http://dx.doi.org/10.2197/ipsjjip.23.708

13. Otmar Hilliges, Shahram Izadi, Andrew D. Wilson, Steve Hodges, Armando Garcia-Mendoza, and Andreas Butz. 2009. Interactions in the Air: Adding Further Depth to Interactive Tabletops. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)*. ACM, New York, NY, USA, 139–148. `DOI:` http://dx.doi.org/10.1145/1622176.1622203

14. Ken Hinckley, Seongkook Heo, Michel Pahud, Christian Holz, Hrvoje Benko, Abigail Sellen, Richard Banks, Kenton O'Hara, Gavin Smyth, and William Buxton. 2016. Pre-Touch Sensing for Mobile Interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2869–2881. `DOI:` http://dx.doi.org/10.1145/2858036.2858095

15. M. Jeannerod. 1984. The Timing of Natural Prehension Movements. *Journal of Motor Behavior* 16, 3 (1984), 235–254. `DOI:` http://dx.doi.org/10.1080/00222895.1984.10735319

16. Azam Khan, George Fitzmaurice, Don Almeida, Nicolas Burtnyk, and Gordon Kurtenbach. 2004. A Remote Control Interface for Large Displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04)*. ACM, New York, NY, USA, 127–136. `DOI:` http://dx.doi.org/10.1145/1029632.1029655

17. Nicolai Marquardt, Ricardo Jota, Saul Greenberg, and Joaquim A. Jorge. 2011. *The Continuous Interaction Space: Interaction Techniques Unifying Touch and Gesture on and above a Digital Surface*. Springer Berlin Heidelberg, Berlin, Heidelberg, 461–476. `DOI:` http://dx.doi.org/10.1007/978-3-642-23765-2_32

18. Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: Combining Gaze with Multi-touch for Interaction on the Same Surface. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 509–518. `DOI:` http://dx.doi.org/10.1145/2642918.2647397

19. Ken Pfeuffer, Jason Alexander, Ming Ki Chong, Yanxia Zhang, and Hans Gellersen. 2015. Gaze-Shifting: Direct-Indirect Input with Pen and Touch Modulated by Gaze. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15)*. ACM, New York, NY, USA, 373–383. DOI: http://dx.doi.org/10.1145/2807442.2807460

20. Henning Pohl and Roderick Murray-Smith. 2013. Focused and Casual Interactions: Allowing Users to Vary Their Level of Engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 2223–2232. DOI: http://dx.doi.org/10.1145/2470654.2481307

21. Dario D. Salvucci and Joseph H. Goldberg. 2000. Identifying Fixations and Saccades in Eye-tracking Protocols. In *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications (ETRA '00)*. ACM, New York, NY, USA, 71–78. DOI: http://dx.doi.org/10.1145/355017.355028

22. Joey Scarr, Andy Cockburn, and Carl Gutwin. 2013. Supporting and Exploiting Spatial Memory in User Interfaces. *Found. Trends Hum.-Comput. Interact.* 6, 1 (Dec. 2013), 1–84. DOI: http://dx.doi.org/10.1561/1100000046

23. Dominik Schmidt, Florian Block, and Hans Gellersen. 2009. A Comparison of Direct and Indirect Multi-touch Input for Large Surfaces. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I (INTERACT '09)*. Springer-Verlag, Berlin, Heidelberg, 582–594. DOI: http://dx.doi.org/10.1007/978-3-642-03655-2_65

24. Julia Schwarz, Scott Hudson, Jennifer Mankoff, and Andrew D. Wilson. 2010. A Framework for Robust and Flexible Handling of Inputs with Uncertainty. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology (UIST '10)*. ACM, New York, NY, USA, 47–56. DOI: http://dx.doi.org/10.1145/1866029.1866039

25. Baris Serim and Giulio Jacucci. 2016. Pointing While Looking Elsewhere: Designing for Varying Degrees of Visual Guidance During Manual Input. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5789–5800. DOI: http://dx.doi.org/10.1145/2858036.2858480

26. Will Spijkers and Simone Spellerberg. 1995. On-line Visual Control of Aiming Movements? *Acta Psychologica* 90, 1 - 3 (1995), 333 – 348. DOI: http://dx.doi.org/10.1016/S0001-6918(95)90000-4 Discrete and Continuous Information Processing.

27. Sophie Stellmach and Raimund Dachselt. 2013. Still Looking: Investigating Seamless Gaze-supported Selection, Positioning, and Manipulation of Distant Targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 285–294. DOI: http://dx.doi.org/10.1145/2470654.2470695

28. Philip Tuddenham, David Kirk, and Shahram Izadi. 2010. Graspables Revisited: Multi-touch vs. Tangible Input for Tabletop Displays in Acquisition and Manipulation Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2223–2232. DOI: http://dx.doi.org/10.1145/1753326.1753662

29. Jayson Turner, Jason Alexander, Andreas Bulling, Dominik Schmidt, and Hans Gellersen. 2013. *Eye Pull, Eye Push: Moving Objects between Large Screens and Personal Devices with Gaze and Touch*. Springer Berlin Heidelberg, Berlin, Heidelberg, 170–186. DOI: http://dx.doi.org/10.1007/978-3-642-40480-1_11

30. Md. Sami Uddin, Carl Gutwin, and Benjamin Lafreniere. 2016. HandMark Menus: Rapid Command Selection and Large Command Sets on Multi-Touch Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5836–5848. DOI: http://dx.doi.org/10.1145/2858036.2858211

31. Simon Voelker, Andrii Matviienko, Johannes Schöning, and Jan Borchers. 2015. Combining Direct and Indirect Touch Input for Interactive Workspaces Using Gaze Input. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction (SUI '15)*. ACM, New York, NY, USA, 79–88. DOI:http://dx.doi.org/10.1145/2788940.2788949

32. Malte Weiss, Simon Voelker, Christine Sutter, and Jan Borchers. 2010. BendDesk: Dragging Across the Curve. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. ACM, New York, NY, USA, 1–10. DOI: http://dx.doi.org/10.1145/1936652.1936654

33. Malte Weiss, Chat Wacharamanotham, Simon Voelker, and Jan Borchers. 2011. FingerFlux: Near-surface Haptic Feedback on Tabletops. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 615–620. DOI: http://dx.doi.org/10.1145/2047196.2047277

34. Malte Weiss, Julie Wagner, Yvonne Jansen, Roger Jennings, Ramsin Khoshabeh, James D. Hollan, and Jan Borchers. 2009. SLAP Widgets: Bridging the Gap Between Virtual and Physical Controls on Tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 481–490. DOI: http://dx.doi.org/10.1145/1518701.1518779

35. Jinglong Wu, Jiajia Yang, and Taichi Honda. 2010. Fitts' law holds for pointing movements under conditions of restricted visual feedback. *Human Movement Science* 29, 6 (2010), 882 – 892. DOI: http://dx.doi.org/10.1016/j.humov.2010.03.009

36. Haijun Xia, Ricardo Jota, Benjamin McCanny, Zhe Yu, Clifton Forlines, Karan Singh, and Daniel Wigdor. 2014. Zero-latency Tapping: Using Hover Information to Predict Touch Locations and Eliminate Touchdown Latency. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 205–214. `DOI:` http://dx.doi.org/10.1145/2642918.2647348
37. Xing-Dong Yang, Tovi Grossman, Pourang Irani, and George Fitzmaurice. 2011. TouchCuts and TouchZoom: Enhanced Target Selection for Touch Displays Using Finger Proximity Sensing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2585–2594. `DOI:` http://dx.doi.org/10.1145/1978942.1979319
38. Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 246–253. `DOI:` http://dx.doi.org/10.1145/302979.303053