# MultiFi: Multi-Fidelity Interaction with Displays On and Around the Body

**Jens Grubert**[1], **Matthias Heinisch**[1], **Aaron Quigley**[2], **Dieter Schmalstieg**[1]
[1] Graz University of Technology [2] University of St Andrews
jg@jensgrubert.de, heinischmatthias@gmail.com, aquigley@acm.org, schmalstieg@tugraz.at

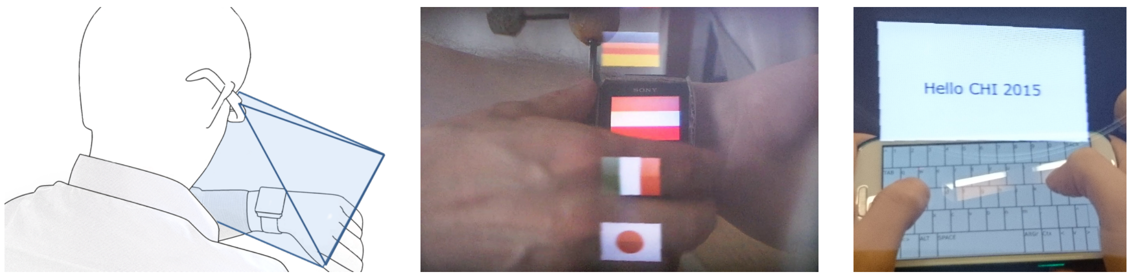


Figure 1: MultiFi widgets crossing device boundaries based on proxemics dimensions (left), e.g., middle: *ring menu* on a smartwatch (SW) with head-mounted display (HMD) or right: *soft keyboard* with full-screen input area on a handheld device and HMD.

## ABSTRACT
Display devices on and around the body such as smartwatches, head-mounted displays or tablets enable users to interact on the go. However, diverging input and output fidelities of these devices can lead to interaction seams that can inhibit efficient mobile interaction, when users employ multiple devices at once. We present MultiFi, an interactive system that combines the strengths of multiple displays and overcomes the seams of mobile interaction with widgets distributed over multiple devices. A comparative user study indicates that combined head-mounted display and smartwatch interfaces can outperform interaction with single wearable devices.

## ACM Classification Keywords
H.5.2 Information interfaces and presentation: User Interfaces - Graphical user interfaces

## INTRODUCTION
Personal, public and ambient displays form a pervasive infrastructure around us. However, displays are typically unaware of each other and make little attempt to coordinate what is shown across them. The emergence of second-screen applications, screen mirroring and remote desktop access demonstrates the benefits of suitably designed coordination. In particular, when users carry multiple displays on and around their body, these displays form a space that can be leveraged for seamless interaction across display boundaries.

In this work, we introduce MultiFi, a platform for implementing user interface widgets across multiple displays with different fidelities for input and output. Widgets such as toolbars or sliders are usually specific to a single display platform, and widgets that can be used between and across displays are largely unexplored. This may come from the problems introduced by the variations in fidelity of input and output across devices. For input, we must accommodate different modes and degrees of freedom. For output, we must accommodate for variations in resolution and field of view. Both input and output affect the exactness of the user experience. Moving across devices can make the differences in fidelity apparent and introduce seams affecting the interaction.

MultiFi aims to reduce such seams and combine the individual strengths of each display into a joint interactive system for mobile interaction. For example, consider continuous navigation support, regardless of where a person is looking. Such navigation may employ a range of worn, handheld or embedded displays. Even if the navigation system is capable of switching among displays in a context-aware manner, the user will still need to contend with varying and uncoordinated fidelities of interaction.

MultiFi addresses the design problem of "interaction on the go" across multiple mobile displays with the follow-


*CHI 2015*, April 18-23, 2015, Seoul, Republic of Korea.

http://dx.doi.org/10.1145/2702123.2702331

ing contributions: 1) We explore the design space of multiple displays on and around the body and identify key concepts for seamless interactions across devices. 2) We introduce a set of cross-display interaction techniques. 3) We present empirical evidence that combined interaction techniques can outperform individual devices such as smartwatches or head-mounted displays for information browsing and selection tasks.

## RELATED WORK

Today's dominant handheld devices, such as smartphones or tablets, have a high *access cost* in terms of the time and effort it takes to retrieve and store the device from where it typically resides, such as one's pocket. This cost reduces the usefulness of a device for micro-interactions, such as checking the time or one's inbox.

Wearable devices such as a smartwatch (SW) or head-mounted display (HMD) lower the access cost to a wrist flick or eye movement. However, interaction with these always-on devices is encumbered by their low *fidelity*: limited screen and touch area, low resolution and poor contrast limit what users can do. Currently, HMDs require indirect input through touch devices, while high-precision spatial pointing is not yet commercially available. Recent research aims to improve the overall fidelity, investigating higher resolution and more immersive displays, improved touchscreen precision [24, 31] or physical pointing [10, 11].

A recurring topic for wearable displays is the extension of display real-estate using virtual screen techniques [14, 15, 29]. Recently, Ens et al. [13] explored the design space for a body-centric virtual display space optimized for multi-tasking on HMDs and pinpointed relevant design parameters of concepts introduced earlier by Billinghurst et al. [4, 5]. They found that body-centered referenced layouts can lead to higher selection errors compared to world-referenced layouts, due to unintentional perturbations caused by reaching motions.

Users with multiple devices tend to distribute tasks across different displays, because moving between displays is currently considered a task switch. For some forms of interaction, a tight spatial registration may not be needed. For example, Duet combines handheld and SW and infers spatial relationships between the devices based on local orientation sensors [9]. Similarly, Billinghurst et al. [7] combine handheld and HMD, but use the handheld mainly as an indirect input device for the HMD. Stitching together multiple tablets [21] allows for interaction across them, under the assumption that they lie on a common plane. Several other approaches combine larger stationary with handheld displays through spatial interaction [1, 6]. The large stationary displays make virtual screens unnecessary, but restrict mobility. The same is true for the work of Benko et al. [2], who combine a touch table with an HMD. Yang and Widgor introduced a web-based framework for the construction of applications using distributed user interfaces but do not consider wearable displays [32].


Figure 2: The extended screen space metaphor for showing a high resolution inlay of a map on SW inside a low resolution representation on a HMD.

Unlike prior work, we focus on the dynamic alignment of multiple body-worn displays, using body motion for spatial interaction.

## INTERACTION BY DYNAMIC ALIGNMENT

MultiFi aims to reduce the access cost of involving multiple devices in micro-interactions by dynamically leveraging complementary input and output fidelities. We propose *dynamic alignment* of both devices and widgets shown on these devices as an interaction technique.

Dynamic alignment can be seen as an application of proxemics [16]: Computers can react to users and other devices based on factors such as distance, orientation, or movement. In MultiFi, dynamic alignment changes the interaction mode of devices based on a combination of proxemic dimensions. We focus on *distance* and *orientation* between devices. However, different alignment styles can be explored, which are location-aware, vary between personal and public displays or consider movement patterns.

### Design factors

To better understand the design implications of dynamic alignment, we begin with a characterization of the most relevant design factors determined throughout the iterative development process of MultiFi.

*Spatial reference frames* encompass where in space information can be placed, if this information is fixed or movable (with respect to the user) and if the information has a tangible physical representation (i.e., if the virtual screen space coincides with a physical screen space) [12].

*Direct vs. indirect input.* We use the term direct input, if input and output space are spatially registered, and indirect input, if they are separated. As a consequence of allowing various spatial reference frames, both direct and indirect input must be supported.

*Fidelity* of individual devices concerns the quality of output and input channels such as spatial resolution, color contrast of displays, focus distance, or achievable input precision. We also understand the display size as a fidelity factor, as it governs the amount and hence quality of information that can be perceived from a single screen.

*Continuity.* The ease of integrating information across several displays not only depends on the individual display fidelities, but also on the quality difference or gap between those displays, in particular, if interaction moves across display boundaries. We call this *continuity of fidelity.* In addition, *continuity of the spatial reference frame* describes if the information space is continuous, as with virtual desktops, or discrete, e.g., when virtual display areas are bound to specific body parts [8]. Continuity factors pose potential challenges when combining multiple on and around the body displays. For example, combining touch screen and HMD extends the output beyond a physical screen of a SW, but not the input. This leads to potential interaction challenges, when users associate the extension of the output space with an extension of the input space.

*Social acceptability* of interactions with mobile, on and around body devices have been extensively studied [30], revealing the personal and subjective nature of what is deemed acceptable. This varies due to many factors including the technology, social situation or location. Dynamic alignment allows for some degree of interaction customization, allowing people to tailor their interactions in a way which best suits their current context, rather than having to rely on default device patterns which may be wholly unsuited to the context of use.

### Alignment modes

For the combination of HMD and touch device, we distinguish three possible alignment modes (see Figure 3):

In *body-aligned mode*, the devices share a common information space, which is spatially registered to the user's body (Figure 3, left). While wearable information displays could be placed anywhere in the 3D space around the body, we focus on widgets in planar spaces, as suggested by Ens et al. [12]. The HMD acts as a low fidelity viewing device into a body-referenced information space, allowing one to obtain a fast overview. The touchscreen provides a high fidelity inset, delivering detail-on-demand, when the user points to a particular location in the body-referenced space. Also, in contrast to common spatial pointing techniques, the touchscreen provides haptic input into the otherwise intangible information space.

In *device-aligned mode* , the information space is spatially registered to the touchscreen device and moves with it (Figure 3, middle). The HMD adds additional, peripheral information at lower fidelity, thus *extending the screen space* of the touch screen, yielding a focus+context display.

In *side-by-side mode*, interaction is redirected from one device to the other without requiring a spatial relationship among devices (Figure 3, right). For example, if the HMD shows a body-referenced information space, a touch device can provide indirect interaction. The touch device can display related information, and input on the touch device can affect the body-referenced display. If the touch device is outside the user's field of view, the touch screen can still be operated blindly.

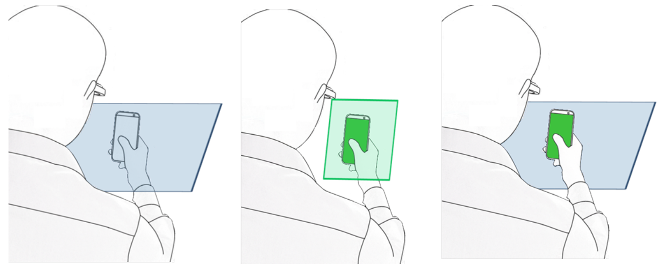

Figure 3: In body-aligned mode (left) devices are spatially registered in a shared information space relative to the user's body. In device-aligned mode (middle) the screen space of the touchscreen is extended. In side-by-side mode (right) devices have separated information spaces and do not require a spatial relationship.

### Navigation

The principal input capabilities available to the user are spatial pointing with the touch device, or using the touch screen. Spatial pointing with the touch device is a natural navigation method in body-aligned mode. Once the alignment is recognized (the user's viewpoint, the handheld and the chosen item are aligned on a ray), the HMD clears the area around the element to let the handheld display a high resolution inset. This navigation method can be used for selection or even drag-and-drop in the body-referenced information space. However, extended use can lead to fatigue.

Spatial pointing in device-aligned mode can be seen as a more indirect form of navigation, which allows one to obtain a convenient viewpoint on the device-aligned information space. Navigation of the focus area will naturally be done by scrolling on the touch screen, but this can be inefficient, if the touch screen is small. Hence, users may mitigate the limitation of input to the physical screen with a clutch gesture that temporarily switches to body-aligned mode. At the press of a button (or dwell gesture), the information space can be fixed in air at the current position. Then users can physically select a new area of the information space by physical pointing, making it tangible again.

### Focus representation and manipulation

An additional design decision is the representation shown on the higher fidelity display: The first option is to solely display a higher *visual level of detail.* For example, the user could align a touch screen over a label to improve the readability of text (Figure 2). The second option presents *semantic level of detail* [25], revealing additional information through a magic lens metaphor [3]. Here, the widget changes appearance to show additional information. For example, in Figure 4, the "Bedrooms" label turns into a scrollable list, once the borders of the

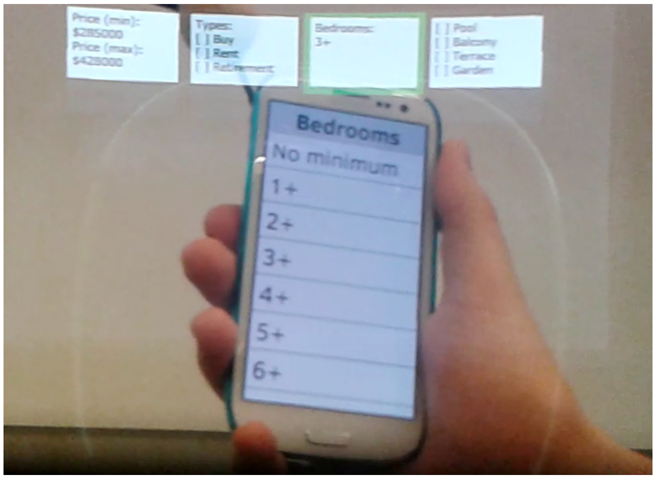


Figure 4: Spatial pointing via a handheld triggers a low fidelity widget on the HMD to appear in high fidelity on the handheld.

handheld and the label represenation in the HMD are spatially aligned. Similarly, in Figure 5 (bottom row), a handheld shows a richer variation of a widget group including photos and detailed text, once it is aligned with the low fidelity representation on the user's arm.

An interactive focus representation on the touch device can naturally be operated with standard touch widgets. In body-aligned mode, this leads to a continuous coarse-to-fine *cascaded interaction*: The user spatially points to an item with a low fidelity representation and selects it with dwelling or a button press. A high fidelity representation of the item appears on the touch screen and can be manipulated by the user through direct touch (Figures 2, 4, 5).

For simple operations, this can be done directly in body-aligned mode. For example, widgets such as checkbox groups may be larger than the screen of a SW, but individual checkboxes can be conveniently targeted by spatial pointing and flipped with a tap. However, holding the touch device still at arm's length or at awkward angles may be demanding for more complex operations. In this case, it may be more suitable to *tear off* the focus representation from the body-aligned information space by automatically switching to side-by-side mode. A rubberband effect snaps the widget back into alignment, once the user is done interacting with it. This approach overcomes limitations of previous work, which required users to either focus on the physical object or on a separate display for selection [11].

### Widgets and applications

MultiFi widgets adapt their behavior to the current alignment of devices. For example, widgets can relocate from one device to the other, if a certain interaction fidelity is required. We have identified a number of ways how existing widgets can be adapted across displays. Here we discuss several widget designs and applications employing such widgets to exemplify our concepts.

**Menus and lists:** On a SW, menu and list widgets can only show a few items at once due to limited screen space. We use an HMD to extend the screen space of the SW, so users get a quick preview of nearby items in a *ring menu*, Figure 1, middle. Similarly, list widgets on an HMD can adapt their appearance to show more information once a handheld device is aligned, Figure 4.

**Interactive map:** Navigation of large maps is often constrained by screen space. We introduce two map widgets that combine HMD and touch screen. The first map widget works similar to the list widget, but extends the screen space of a touch display in both directions. Interaction is achieved via the touch display.

The second variant makes use of a body-referenced information space. The map is displayed in the HMD relative to the upper body, either horizontally, vertically or tilted (Figure 2). If the map size is larger than the virtual display space, the touchpad on the SW provides additional pan and zoom operations.

**Arm clipboard:** Existing body-centric widgets for handhelds [8, 23] rely on proprioceptive or kinesthetic memorization, because the field of view of the handheld is small. With an additional HMD, users can see where on their body they store through head pointing and subsequently retrieve information with a handheld device. If a list widget displays additional information on one side of the SW (overview+detail), we can let users store selected items on their lower arm (Figure 5). Aligning the handheld with one of the items stored on the arm automatically moves the item to the higher fidelity handheld. For prolonged interaction, the item can now be manipulated with two hands on the handheld. Through the combination of HMD for overview and touch enabled displays for selection and manipulation, body-referenced information spaces could become more accessible compared to previous approaches solely relying on proprioceptive memory [8, 23].

**Text input:** Using MultiFi text widgets, we have implemented a full-screen soft keyboard application for a handheld used with a HMD. The additional screen real estate on the handheld allows MultiFi to enlarge the soft keys significantly, while the text output is redirected to the HMD. As soon as a HMD is aligned, the text output area can relocate from one device to the other (see Figure 1, right). This results in two potential benefits. First, the larger input area could help speed up the writing process. Second, the written text is not publicly visible, hence supporting privacy.

## IMPLEMENTATION

### Software

The MultiFi prototype is based on HTML5, JavaScript, WebSockets for communication, three.js for rendering and hammer.js for touch gesture recognition. All client devices open a website in a local browser and connect to the Java-based application server. JSON is used to

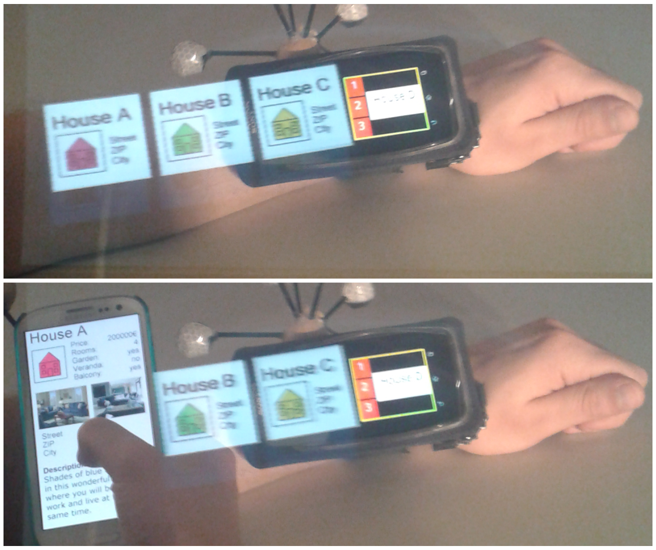


Figure 5: Arm clipboard with extended screen space for low fidelity widgets (top). Spatial pointing enables switching to high fidelity on a handheld (bottom).

encode the distributed messages, and tracking data is received via VRPN.

Widgets have potentially multiple graphical representations in replicated and synchronized scenegraphs and a common state which is shared via the central application server. For widgets that do not change their appearance and simply span multiple devices, multiple camera views on the same 3D scene are used (e.g., ring menu, map). Widgets that adapt their appearance (such as list items) use multiple synchronized representations. Interaction across devices relies on the known 3D poses of individual devices, shared via the central application server. For example, selection of an item in the HMD via a touch screen is realized through intersection from the touch point with the virtual HMD image plane.

As our system relies on the accurate registration between devices, calibration of individual components is required. Foremost, the HMD is calibrated via optical see-through calibration methods (using single or multiple point active alignment methods [17]). In addition, the image masks for the touch screen devices (i.e. the area that should not be rendered on the HMD) and thus their positions relative to their tracking markers have to be determined. For this the user manually aligns the touch screen with a pre-rendered rectangle displayed on the HMD (having the same size as the touch screen) which allows MultiFi to determine the transformation between the touch screen and the attached tracking target. Please note that these calibration steps typically have to be carried out only once for each user and device respectively.

### Devices

We implemented a MultiFi prototype using a Samsung Galaxy SIII (resolution: 1280x720 px, 306 ppi, screen size: 107x61 mm) as smartphone, a Vuzix STAR 1200 XL HMD (resolution: 852x480 px, horizontal field of view (FoV): 30.5° vertical FoV: 17.15°, focus plane distance: 3 m, resolution: 13 ppi at 3 m, weight with tracking markers: 120 g) and another smartphone (Sony Xperia Z1 compact) as smartwatch substitute (resolution: 1280x720 px, cropped extent: 550x480 px, 342 ppi, weight with tracking markers: 200 g). We chose this approach to simulate next generation smartwatches with higher display resolution and more processing power. To this aim, we limited the screen extent to 40x35 mm to emulate the screen extent of a typical smartwatch. The HMD viewing parameters were matched with virtual cameras which rendered the test scenes used in the smartphone, HMD and SW.

### Tracking

We used an A.R.T. outside-in tracking system to determine the 3D positions of all devices. This currently limits our prototype to stationary use in laboratory environments. Still, mobile scenarios could be supported by relying on local sensors only. For example HMDs with in-built (depth) cameras could be used to determine the 3D position of touch screens relative to the HMD [26]. Alternatively, in-built orientation sensors could track the touch screen and HMD positions relative to a body-worn base station (such as an additional smartphone in the user's chest pocket). Please note that the later approach would likely result in less accuracy and drift over time. This would need to be considered in the adaptation rules for widgets when spanning multiple devices.

## USER STUDY

We conducted a laboratory user study to investigate if combined device interaction can be a viable alternative to established single device interaction for mobile tasks. For the study we concentrated on two atomic tasks: information search and selection. Those tasks were chosen as they can be executed on the go and underpin a variety of more complex tasks.

### Experimental design

We designed a within-subjects study to compare performance and user experience aspects of MultiFi interaction to single device interaction for two low level tasks. We complemented the focus on these atomic tasks with user inquiries about the potential and challenges of joint on and around the body interaction. For both tasks, we report on the following dependent variables: *task completion time*, *errors, subjective workload* as measured by NASA TLX [19] as well as user experience measures (After Scenario Questionnaire (ASQ) [22], *hedonic and usability aspects* as measured by AttrakDiff [20]) and overall *preference* (ranking). The independent variable for both tasks was interface with five conditions:

*Handheld*: The Samsung Galaxy SIII was used as only input and output device. This serves as the baseline for a handheld device with high input and output fidelity.

*Smartwatch (SW):* The wrist-worn Sony Xperia Z1 compact was used as only input and output device. The input and output area was 40x35 mm and highlighted by a yellow border, as shown in Figure 2. Participants were notified by vibration if they touched outside the input area. This condition serves as baseline for a wearable device with low input and output fidelity (high resolution, but small display space).

*Head Mounted Display (HMD):* The Vuzix STAR 1200XL was used as an output device. We employed indirect input as in the SW condition using a control-display ratio of 1 with the touch area limited to the central screen area of the HMD. This condition serves as the baseline for a HMD with low input and output fidelity, which can be operated with an arm-mounted controller.

*Body-referenced interaction (BodyRef):* The content was displayed in front of the participant in body-aligned mode with additional touch scrolling. Selection was achieved by aligning the smartwatch with the target visible in front of the user and touching the target rendered on the smartwatch.

*Smartwatch referenced (SWRef):* The information space was displayed in device-aligned mode (Figure 9). All other aspects were as in BodyRef.

### Apparatus and data collection

The study was conducted in a controlled laboratory environment. The devices employed were the ones described in the implementation section. The translation of virtual cameras for panning via touch in all conditions parallel to the screen was set to ensure a control-display ratio of 1. Pinch to zoom was implemented by the formula $s = s_0 \cdot s_g$, with $s$ being the new scale factor, $s_0$ the map's scale factor at gesture begin and $s_g$ the relation between the finger distances at gesture begin and end. While the system is intended for mobile use, here participants conducted the tasks while seated at a table (120x90 cm, height 73 cm, height adjustable chair) due to the strenuous nature of the repetitive tasks in the study. Null hypothesis significance tests were carried out at a .05 significance level, and no data was excluded, if not otherwise noted. For ANOVA (repeated measures ANOVA or Friedman ANOVA), Mauchly's test was conducted. If the sphericity assumption had been violated, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. For post-hoc tests (pairwise t-test or Wilcoxon signed rank) Bonferroni correction was applied. Due to space reasons not all tests statistics are reported in detail, but are available as supplementary material[1].

### Procedure

After an introduction and a demographic questionnaire, participants were introduced to the first task (counterbalanced) and the first condition (randomized). For each condition, a training phase was conducted. For each task, participants completed a number of trials (as described in the individual experiment sections) in five blocks, each block for a different condition. Between each block, participants filled out the questionnaires. At the end of the study, a semi-structured interview was conducted and participants filled out a separate preference questionnaire. Finally, the participants received a book voucher worth 10 Euros as compensation. Participants were free to take a break between individual blocks and tasks. Overall, the study lasted ca. 100 minutes per participant.

### Participants

Twenty-six participants volunteered in the study. We had to exclude three participants due to technical errors (failed tracking or logging). In total, we analyzed data from twenty-three participants (1 f, average age: 26.75 y, $\sigma$=5.3, average height: 179 cm, $\sigma$= 6, 7 users wore glasses, three contact lenses, 2 left-handed users). All but one user were smartphone owners (one less than a year). Nobody was a user of smartwatches or head-mounted displays. Twenty users had a high interest in technology and strong computer skills (three medium).

### Hypotheses

One of our main interests was to investigate if combined display interaction could outperform interaction with individual wearable devices. We included Handheld interaction as a baseline and did not expect the combined interfaces to outperform it. Hence, we had the following hypotheses: *H1:* Handheld will be fastest for all tasks. *H2:* BodyRef will be faster than HMD and SW (ideally close to Handheld). *H3:* BodyRef will result in fewer errors than HMD and SW. *H4:* SWRef will be faster than HMD and SW (ideally close to Handheld). *H5:* SWRef will result in fewer errors than HMD and SW.

## EXPERIMENT 1: LOCATOR TASK ON MAP

A common task on mobile mapping applications is to search for an object with certain target attributes [28]. We employed a locator task similar to previous studies involving handheld devices and multi-display environments [18, 27]. Participants had to find the lowest price label (text size 12 pt) among five labels on a workspace size of 400x225 mm. We determined the workspace size empirically, to still allow direct spatial pointing for the BodyRef condition. While finding the lowest price could easily be solved with other widgets (such as a sortable list view), our task is only an instance of general locator tasks, which can encompass non-quantifiable attributes such as textual opinions of users, which cannot be sorted automatically. Users conducted ten trials per condition. With 23 participants, five interface levels and 10 trials, there was a total of 23x5x10=1150 trials.

### Task completion time and errors

The task completion times (TCT, in seconds), for the individual conditions can be seen in Figure 6. A repeated measures ANOVA indicated that there was a significant effect of interface on TCT, F(3.10, 709.65)=42.21,

[1]http://www.jensgrubert.wordpress.com/research/multifi/

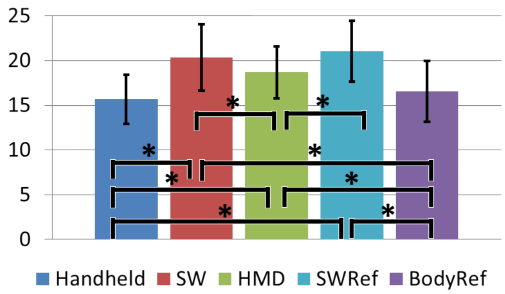


Figure 6: Task completion time (s) for the locator task.

p<.001. Post-hoc tests indicated that both Handheld and BodyRef were significantly faster than all remaining interfaces with medium to large effect sizes (see also Figure 6). HMD was significantly faster than both SW and SWRef. There were no significant differences between Handheld-BodyRef and SW-SWRef.

From 230 selections, eight false selections were made in the Handheld, HMD and BodyRef conditions. In the SW condition, 13 errors have been made, in SWRef five errors. No significant differences were found.

### Subjective workload and user experience

Repeated measures ANOVAs indicated that there were significant effects of interface on all dimensions. Post-hoc tests indicated that BodyRef resulted in a higher mental demand than smartwatch (albeit with a small effect size). The handheld condition resulted in significantly lower subjective workload for all other dimension compared to most other interfaces. The analysis of the ASQ (repeated measures ANOVA and post-hoc tests) indicated that for Handheld ease of task was significantly higher than for SWRef. Analysis of AttrakDiff showed that all interfaces scored slightly below average for pragmatic quality (PQ), see Figure 7, and only a significant difference between HMD-SWRef could be found (but with a small effect size). For hedonic quality stimulation (HQ-S), the Handheld and SW interface were rated significantly lower than the other three conditions. Preference analysis showed that Handheld (MD=2, M=1.13, $\sigma$=1.13) was significantly more preferred than SW (MD=4, M=3.87, $\sigma$=1.10), Z=-4.25, p<.001.

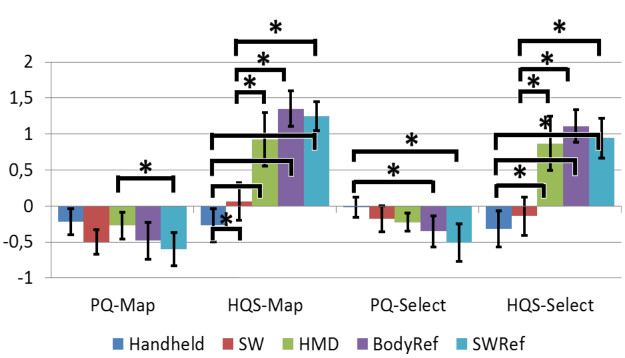


Figure 7: Pragmatic Quality (PQ) and Hedonic Quality Stimulation (HQS) measures (normalized range -2..2) for the locator task (left) and the select task (right).

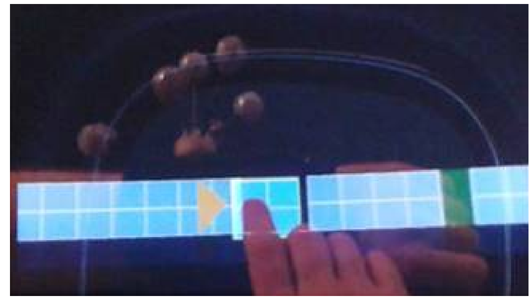

Figure 8: The selection task for SWRef.

## EXPERIMENT 2: 1D TARGET ACQUISITION

We employed a discrete 1D pointing task similar to the one used by Zhao et al. [33] (Figure 8). Participants navigated to a target (green stripe) in each trial using touch input (for Handheld, SW, HMD, SWRef) or spatial pointing (BodyRef). Final target selection was confirmed by a touch on the target region in all conditions. The participants were asked to use their index finger to interact with the touch surfaces. For each trial, the task was to scroll the background (Handheld, SW, HMD, SWRef) or to move the smartwatch towards the target (BodyRef) until it appeared on the *selection area*. Prior to each trial, participants hit a start button at the center of the screen to ensure a consistent start position and to prevent unintended gestures before scrolling. The target was only revealed after the start button was hit. After successful selection, the target disappeared. For BodyRef, participants returned to a neutral start position centered in front of them before the next trial. In the experiment design we fixed target width to 20 mm (0.5*width of the smartwatch), use the control window and display window sizes of the individual displays and use two target distances (short: 15 cm, long: 30 cm)[2]. The conditions were blocked by interface. Per condition, each participant conducted eight trials (plus two training trials). With twenty three participants, five interface levels, two target distances, two directions and eight trials per condition a total of 23x5x2x2x8=3680 trials were conducted.

### Task completion time and errors

Task completion times are depicted in Figure 9. Repeated measures ANOVAs indicated that for both distances (15 cm, 30 cm) and smartwatch sides (towards and away from dominant hand) interface had a significant effect on TCT. The pairwise significant differences are depicted in Figure 9. Handheld was the fastest interface for both directions and distances. BodyRef was significantly faster than all remaining interfaces. No other significant effects of interface on task completion time were found.

Selection errors occurred when participants tapped outside the target region. The total number of errors for individual interfaces were as follows: Handheld:

[2]We fixed those parameters as the focus of the experiment was not on generating a new target aquisition model.

53 (M=.07, $\sigma$=.28), SW: 34 (M=.05, $\sigma$=.23), HMD: 223 (M=.30, $\sigma$=.77), BodyRef: 258 (M=.35, $\sigma$=.78), SWRef: 37 (M=.05, $\sigma$=.24). A Friedman ANOVA indicated that there was a significant effect of interface on error count ($\chi^2(4)$=231.68, p<.001). Post-hoc tests indicated significant differences between BodyRef and all interfaces except HMD, as well as between HMD and all interfaces (except BodyRef).

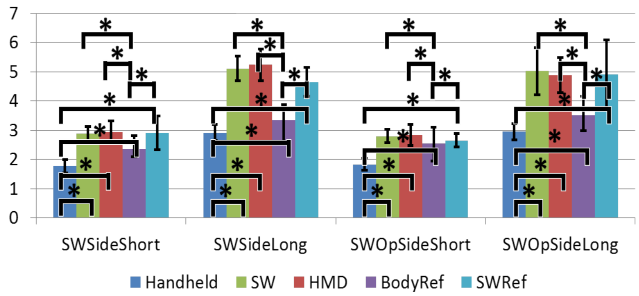


Figure 9: Task completion times (s) for the select task. SWSide: side on which smartwatch was worn, SWOpSide: opposite side.

### Subjective workload and user experience

A repeated measures ANOVA indicated that there were significant effects of interface on all dimensions but temporal demand and performance. Post-hoc tests indicated that Handheld resulted in a significantly lower mental demand than most other conditions (except SW) and in a significantly lower overall demand than all conditions. BodyRef and SWRef resulted in significantly higher physical demands compared to Handheld and HMD (but not SW). Frustration was significantly higher for SW and SWRef compared to Handheld. Analysis of results of the ASQ indicated a significant difference between Handheld and SWRef for ease of task (Z= -3.36, p=.01). As in the locator task, all interfaces scored below average for PQ-S (see Figure 9). BodyRef and SWRef scored significantly lower than Handheld (indicated by repeated measures ANOVA and post-hoc t-tests). For HQ-S, the Handheld and SW interface were rated significantly lower than the other three conditions as in the locator task.

### Qualitative feedback

In semi-structured interviews participants commented on potentials and limitations of the prototypical MultiFi implementation. Most participants (21) commented on the benefits of having an *extended view space* compared to individual touch screens with one participant saying "*Getting an overview with simple head movements is intuitive and natural*". Those participants also valued the fact that precise selection was enabled through the smartwatch with one typical comment being "*The HMD gives you the overview, and the SW lets you be precise in your selection*". Three participants highlighted the potentially lower *access costs* of MultiFi over smartphones, with one comment being "*I dont have to constantly monitor my smartphone*". In line participants felt that BodyRef interaction was fastest (even though this is not confirmed by the objective measurements). Five participants commented on the benefits of MultiFi over HMD only interaction highlighting the direct interaction or that they could "*take advantage of proprioception and motion control*".

Many participants (15) commented on the limitations of the hardware, specifically the quality of the employed HMD with a typical comment being "*The combined interfaces [SWRef, BodyRef] gave me trouble because of display quality*". Specifically, the employed HMD obscured parts of the users' field of view "*preventing the ability of glancing down (on the SW) without moving your head*". Another issue highlighted by 6 participants was the *cost of focus switching* which refers to the accommodation to different focus depths of the touch screen and the virtual HMD screen with a typical comment being: "*I have to focus on three layers, which is overwhelming: SW, HMD and real world*". This also led to *coordination problems across devices* as mentioned by 9 participants. Hence, some participants suggested not to concurrently use HMD and SW as output: "*Pairing the two devices is good, but use one as input, the other as output, not both as output, it's confusing*". Also, social concerns of spatial pointing were raised, "*I could not imagine this in a packed bus*".

## DISCUSSION

The study results indicate that combined SW and HMD interaction in body-referenced information spaces can outperform individual wearable devices in terms of task completion time (*H2* holds) and that handheld interaction is not always fastest (H1 does not hold). However, this currently comes at the expense of higher workload and lower usability ratings. We see two major sources for this. First, compared to commercially available wearable devices, we used relatively heavy laboratory equipment (smartphone and HMD with separate retro-reflective markers). Participants mentioned that they would prefer the combined interaction more, if it were lighter. Second, we compared novel interaction techniques involving continuous spatial pointing with established touch screen interaction. Hence, we assume that both lighter devices and more training could mitigate these workload effects.

In the selection task, BodyRef and HMD resulted in a significantly higher error number than the other interfaces (H3 does not hold). Also, SWRef did not result in significantly less errors (H5 does not hold). For HMD, this could be explained by the indirect touch input combined with a smaller control window (SW area) compared to the larger display window. For BodyRef, it turned out that the outside-in tracking system for spatial pointing and our system architecture introduced an average end-to-end delay from user motion to display update of 154 ms ($\sigma$=36). A further video analysis revealed that users were tapping the SW repeatedly when they have reached the target area, even though they were informed to select as precisely as possible. While this is clearly a limitation of our current experimental system setup, we believe that future tracking systems will minimize delay, allowing more precise physical pointing.

Semi-structured interviews revealed that users generally preferred Handheld, as it was the most familiar device, had the largest touch input area and was most comfortable to use. BodyRef was preferred as it felt fast and separated target search via head pointing and selection via spatial pointing with the SW. User comments included "*Moving your head to get an overview is very intuitive*" and "*knowing where to move before you move makes it easier than other conditions*". Still, confirming spatial selection with the touchpad was not welcomed by all, "*I would prefer to just point with my fingers or eyes*".

SWRef performed (for both tasks) not better than individual devices, even though they are based on the extended screen space metaphor as the body-referenced condition (H4 does not hold). Subjective feedback in the semi-structured interviews indicated that participants could not efficiently use the SWRef condition due to the need for refocusing between the SW display (˜40 cm distance) and the focus plane of the HMD (˜300 cm). In addition, the HMD had a lower visual fidelity, which likely increased the effort for reading the labels. Some participants still favored the SWRef condition, specifically for the selection task. They indicated that the HMD gave them "*a peripheral awareness when the target approaches the smartwatch*". This hints that SW referenced display space extension could be beneficial, if the visual fidelity of the HMD and costs of display switching is considered in the design process. For example, instead of rendering a map continuously across displays without adjustments, individual map regions could be adjusted to be more readable across displays (or to avoid the need for actually reading the text on the HMD at all).

SW alone was least preferred due to cumbersome interaction with a small input and output area. Specifically, swiping motions were deemed inefficient. For example, in the select task, participants mentioned a lack of overview, "*I did not know when I passed the target*". HMD was preferred by some users, because they could keep their head and arm in comfortable positions and have a "lean back" experience. They mention it is "better than Google Glass as I can use the smartwatch as touchpad". One participant said: "I could imagine using this for presentations were I can see the slides in the HMD and keep eye contact with the audience when controlling the app with my smartwatch".

Revisiting MultiFi we see that the spectrum of *dynamic alignment* ranging from uncoupled individual devices to closely coupled spatially registered interaction is a key concept for supporting a broad range of mobile scenarios. It facilitates the idea that over time, users can develop individual preferences for multi-display interaction styles just like current touch interfaces offer multiple ways of interaction. The qualitative feedback in the study indicated that users could see benefits of MultiFi over individual device interaction in terms of *access costs* and *direct interaction*. Being able to directly interact within this view space through a touch screen distinguishes MultiFi from other approaches like mid-air interaction via depth-sensors, which lack the haptic feedback of touch screens and through this potentially result in a lower selection precision.

However, such benefits may come at an increased *coordination cost* across displays. Specifically, while we presented a first set of possible widgets, our study revealed that those widgets have to be designed carefully to be able to efficiently lower interaction gaps introduced by individual devices (such as focus distance and resolution differences). Simply extending the display space for widgets across displays without adapting their appearance and operation (as done with the SW referenced map) seems not to be enough to overcome interaction seams. This indicates the need for more research to further investigate the particulars of efficient cross display widgets for interaction on the go. For example, for the map widget we could imagine to further reduce the visual complexity on the low fidelity HMD by simply indicating the location of points of interest with details only appearing on the high fidelity display (as in the arm clipboard).

## CONCLUSION AND FUTURE WORK

We have presented *MultiFi*, an interactive system that combines the strengths of multiple displays on and around the body. We explored how to minimize seams in interaction with multiple devices by dynamic alignment between interfaces. Furthermore, we discussed the implications for user interface widgets and demonstrated the feasibility of our concept through a working prototype system. Finally, we demonstrated that combined HMD and smartwatch interaction can outperform interaction with single wearable devices in terms of task completion time, albeit with higher workload.

In future work, we will explore the design concepts of combining multiple wearable and other displays within our conceptual framework. We also want to build a fully mobile prototype using only mobile sensors.


### Acknowledgements

This work was supported by the EU FP7 project MAGELLAN under the grant number ICT-FP7-611526.


### References


1. Beaudouin-Lafon, M., Huot, S., Nancel, M., Mackay, W., Pietriga, E., Primet, R., Wagner, J., Chapuis, O., Pillias, C., Eagan, J., Gjerlufsen, T., and Klokmose, C. Multisurface interaction in the wild room. *Computer*, 45(4), 2012: 48–56.
2. Benko, H., Ishak, E. W., and Feiner, S. Cross-dimensional gestural interaction techniques for hybrid immersive environments. *VR 05*. 2005, 209–216.

3. Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose, T. D. Toolglass and magic lenses: the see-through interface. *SIGGRAPH*. 1993, 73–80.
4. Billinghurst, M., and Starner, T. Wearable devices: new ways to manage information. *Computer*, 32(1), 1999: 57–64.
5. Billinghurst, M., Bowskill, J., Dyer, N., and Morphett, J. An evaluation of wearable information spaces. *VRAIST 98*. 1998, 20–27.
6. Boring, S., Baur, D., Butz, A., Gustafson, S., and Baudisch, P. Touch projector: mobile interaction through video. *CHI 10*. 2010, 2287–2296.
7. Budhiraja, R., Lee, G. A., and Billinghurst, M. Using a hhd with a hmd for mobile ar interaction. *ISMAR 13*. 2013, 1–6.
8. Chen, X., Marquardt, N., Tang, A., Boring, S., and Greenberg, S. Extending a mobile device's interaction space through body-centric interaction. *MobileHCI 12*. 2012, 151–160.
9. Chen, X., Grossman, T., Wigdor, D. J., and Fitzmaurice, G. Duet: exploring joint interactions on a smart phone and a smart watch. *CHI 14*. 2014, 159–168.
10. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., and Looser, J. Air pointing: design and evaluation of spatial target acquisition with and without visual feedback. *IJHCS*, 69(6), 2011: 401–414.
11. Delamare, W., Coutrix, C., and Nigay, L. Designing disambiguation techniques for pointing in the physical world. *EICS 13*. 2013, 197–206.
12. Ens, B., Hincapié-Ramos, J. D., and Irani, P. Ethereal planes: a design framework for 2d information space in 3d mixed reality environments. *SUI 14*. 2014, 2–12.
13. Ens, B., Finnegan, R., and Irani, P. The personal cockpit: a spatial interface for effective task switching on head-worn displays. *CHI 14*. 2014, 3171–3180.
14. Feiner, S., MacIntyre, B., Haupt, M., and Solomon, E. Windows on the world: 2d windows for 3d augmented reality. *UIST 93*. 1993, 145–155.
15. Fitzmaurice, G. W. Situated information spaces and spatially aware palmtop computers. *CACM*, 36(7), 1993: 39–49.
16. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M. Proxemic interactions: the new ubicomp? *interactions*, 18(1), 2011: 42–50.
17. Grubert, J., Tuemler, J., Mecke, R., and Schenk, M. Comparative user study of two see-through calibration methods. *VR 10*. 2010, 269–270.
18. Grubert, J., Pahud, M., Grasset, R., Schmalstieg, D., and Seichter, H. The utility of magic lens interfaces on handheld devices for touristic map navigation. *PMC*, 2014:
19. Hart, S. G., and Staveland, L. E. Development of nasa-tlx (task load index): results of empirical and theoretical research. *AIP*, 52, 1988: 139–183.
20. Hassenzahl, M., Burmester, M., and Koller, F. Attrakdiff: ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität. In: *M&C 03*. Springer, 2003, 187–196.
21. Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., and Smith, M. Stitching: pen gestures that span multiple displays. *AVI 04*. 2004, 23–31.
22. Lewis, J. R. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the asq. *SIGCHI Bulletin*, 23(1), 1991: 78–81.
23. Li, F. C. Y., Dearman, D., and Truong, K. N. Virtual shelves: interactions with orientation aware devices. *UIST 09*. 2009, 125–128.
24. Olwal, A., Feiner, S., and Heyman, S. Rubbing and tapping for precise and rapid selection on touch-screen displays. *CHI 08*. 2008, 295–304.
25. Perlin, K., and Fox, D. Pad: an alternative approach to the computer interface. *SIGGRAPH 93*. 1993, 57–64.
26. Rädle, R., Jetter, H.-C., Marquardt, N., Reiterer, H., and Rogers, Y. Huddlelamp: spatially-aware mobile displays for ad-hoc around-the-table collaboration. *ITS 14*. 2014, 45–54.
27. Rashid, U., Nacenta, M. A., and Quigley, A. The cost of display switching: a comparison of mobile, large display and hybrid ui configurations. *AVI 12*. 2012, 99–106.
28. Reichenbacher, T. Adaptive concepts for a mobile cartography. *JGS*, 11(1), 2001: 43–53.
29. Reichlen, B. A. Sparcchair: a one hundred million pixel display. *VRAIS 93*. 1993, 300–307.
30. Rico, J., and Brewster, S. Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces. *MobileHCI 09*. 2009, 64:1–64:2.
31. Vogel, D., and Baudisch, P. Shift: a technique for operating pen-based interfaces using touch. *CHI 07*. 2007, 657–666.
32. Yang, J., and Wigdor, D. Panelrama: enabling easy specification of cross-device web applications. *CHI 14*. 2014, 2783–2792.
33. Zhao, J., Soukoreff, R. W., Ren, X., and Balakrishnan, R. A model of scrolling on touch-sensitive displays. *IJHCS*, 72(12), 2014: 805–821.