Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16), December 12–15, 2016, Rovaniemi, Finland

Touch&Screen: Widget Collection for Large Screens Controlled through Smartphones

Alessio Bellino, Federico Cabitza, Giorgio De Michelis, Flavio De Paoli
University of Milano-Bicocca
Department of Computer Science, Systems and Communication (DISCo)
Viale Sarca 336/14, 20126, Milan, Italy
[bellino, cabitza, gdemich, depaoli]@disco.unimib.it

ABSTRACT
We present Touch&Screen, a wide set of interaction techniques for the remote control of widgets (menu, lists, videos, maps etc.) for large screens through smartphones. After presenting the design of these widgets and the related control interfaces of the smartphone, we evaluated the interaction through two user studies. The first study (48 users) aimed to evaluate the user experience (naturalness, usability, etc.) and spontaneity of use. The second user study (12 users) aimed to evaluate our interaction techniques comparatively with direct touch and a commercial smartphone application designed to control cursors on distant screens. The evaluation results show that Touch&Screen is faster and more natural than existing solutions to control distant screens.

Author Keywords
remote control; smartphones; large screens; interaction techniques.

ACM Classification Keywords
H.5.2. User Interfaces: Interaction styles, Input devices and strategies

INTRODUCTION
Several recent works have explored how to combine smartphones and large screens for developing new interaction techniques. These works present many different techniques, which leverage several sensors of mobile devices (e.g. camera [4] [9], accelerometer [12] [21] [34], NFC [19]), also in addition to eye tracking [31] [32] [33]. Regarding the interaction techniques, previous research focuses on content transfer [1] [30] [33], object manipulation (e.g. drag, pinch using touch) [31] [32], control of distant cursors [10] [22] [23] [28] and horizontal surfaces [14] [20] [27] [29] [37] .

Nevertheless, no relevant works have explored the control of content widgets through smartphones. Contents that can be displayed on large screen are various (lists of elements, pictures, videos, maps, text paragraphes) and the techniques that could be envisioned to interact with and visualize them can be very different, depending on the designers' creativity.

Until now, techniques for controlling content widgets were especially developed for commercial applications (e.g. [2] [15]). Moreover, these techniques were developed for different isolated scopes (e.g., for scrolling lists of elements, or to rewind and fast forward videos) and, therefore, there is no a consistent set of techniques that fullfill most content visualization requirements of large screens.

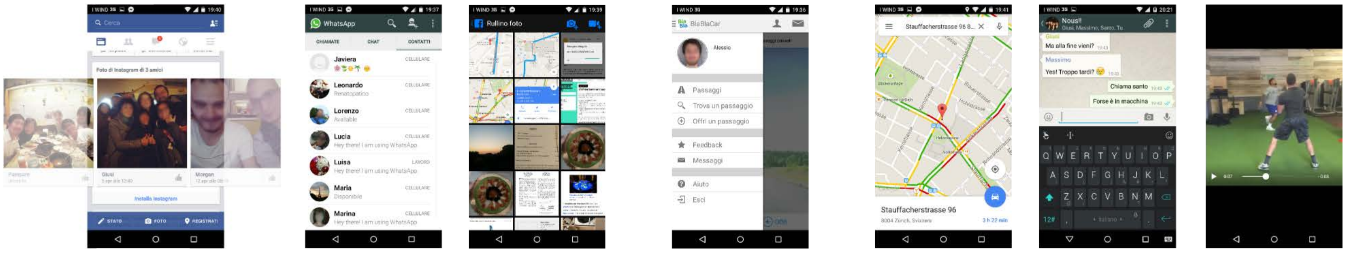
This paper aims to fill this gap by developing of a consistent set of interaction techniques for large screens controlled by smartphones. Our techniques are designed to allow the control of a wide set of content widgets (control menu, lists of elements, video, maps, pictures, text entry, etc.) through touch gestures.

Some of the presented techniques are novel, others were inspired by previous works and we discuss any eventual difference with them.

For sake of readability, we introduce our technique collection by discussing an example: the use of gestures to control systems like PVRs, DVDs, or Smart TV. A bezel-swipe [24] gesture is used to access a menu (like the "menu" button on a remote), directional swipes to navigate left, right, up, and down (like the arrow directional buttons on a remote) and a tap to select an option (like "enter" on a remote). This basic set of gestures can be tailored for special content, like using left and right swipes to navigate forward or backward while playing a video and tap to pause and play. In addition, the gesture set is expanded allowing specific interactions (like zooming on maps) and providing more than discrete directions. For example, the speed of swipes is exploited to accelerate when scrolling a list or when navigating a continuous space like a map. Since all inputs are on the smartphone, even the text entry can be carried out through its touch keyboard.

We evaluated our techniques through two user studies performed on a widget-based application that resembles a smart TV. The first study involved 48 users and aimed to evaluate generally the user experience (naturalness, usability, etc.). Moreover, since we designed the interactions to let users use it without any guidance, we split the users into two groups,

MUM '16, December 12-15, 2016, Rovaniemi, Finland
© 2016 ACM. ISBN 978-1-4503-4860-7/16/12…$15.00
DOI: **http://dx.doi.org/10.1145/3012709.3012736**



A. Carousel of elements B. List of element C. Matrix of element D. Side menu (so-called hamburger) E. Map F. Text insertion G. Audio/Video

Figure 1. Interaction widgets on touch devices (Android).

i.e., trained and untrained ones to compare them discussing the differences. In the second user study on 12 users, instead we evaluated comparatively our interaction techniques with direct touch and a commercial application for smartphone designed to control cursors on distant screens.

RELATED WORKS

Most techniques to control distant large screens through smartphones use pointers resembling the PC-desktop interaction with mouses [10] [22] [23] [28]. These techniques are out of our scope since the interaction is similar to a desktop mouse. The novelty proposed in the literature deals with the "quality" of the cursor control (e.g. precision and speed).

Other techniques go beyond pointing paradigm leveraging gestures performed on smartphones to control large screens. Some of them use the accelerometer to sense tilt and throw gestures [12] or for gaming [34]. Other researchers investigated interface replication between touch devices (controlled by touch gestures) and large screens [3] [16]. Smartphone were used also in combination with eye-trackers [31] [32]: the latter is used to point objects and the former to manipulate them (e.g., zooming, moving, and so on) exploiting touch gestures. Other techniques [4] [9] let user acquire objects through smartphone cameras and visualize them on the smartphone screen to be manipulated through touch gestures.

Moving far from gestures, there are techniques that exploit smartphone identification, location and orientation [27], (old) cell-phones keyboard buttons [13] [19] or using smartphones as 'sleeve display' to augment interactions [36]. Other researchers investigated different multi-display settings in which mobile devices are used like physical filters atop a wall-sized display or in combination with laser pointers to perform interactions such as pan and zoom [5]. Other studies investigate different ways for connecting smartphones and large screen [35] and the deployment of public screen games controlled by smartphones [26]. Regarding tabletops, their use in combination with smartphones was investigated for improving co-located interaction [29] [37] and gaming [14].

On the commercial side, many mobile applications to control large screens have been developed. Anyhow, we limit our analysis to Able Remote [15] and the official Apple TV Remote app for IOS [2] (Figure 2) because of their similarity with our work. The first one works similarly to our solution for controlling video/audio (i.e. horizontal swipes to navigate them forward or backward). The second one works similarly to navigate contents such as grid lists (i.e. directional swipes to navigate elements left, right, up and down).

Some of the techniques presented in this paper (i.e., list of elements and video control) were used for developing YouTube4Two [8], a face-to-face application [7] that lets two co-present people share videos.

For each technique we present, we discuss any significant difference with these previous works in the next sections.

DESIGN RATIONALE

We designed a set of interaction techniques, named Touch&Screen, aimed to control widgets displayed on large screens through touch gestures performed on smartphones.

Touch&Screen works associating areas on the smartphone with widgets on the large screen (e.g. in Figure 3 elements with the same color are associated). Moreover, for each association, a "command space" (e.g., arrows in Figure 3) defines the relation between gestures performed on the area and the effects displayed on the large screen. An example of command space is the association between the bezel-swipe on the smartphone (gesture) and the opening of the menu on the large screen (effect).

Since it is advisable to let users keep the gaze at the large screen to avoid the cost of display switching [25], our techniques are designed to show animated transitions on the large screen while interacting on the smartphone. Continuous animated transitions on the large screen inform the user about the correctness of the (inter)action ongoing on the smartphone reducing the need to look at it.

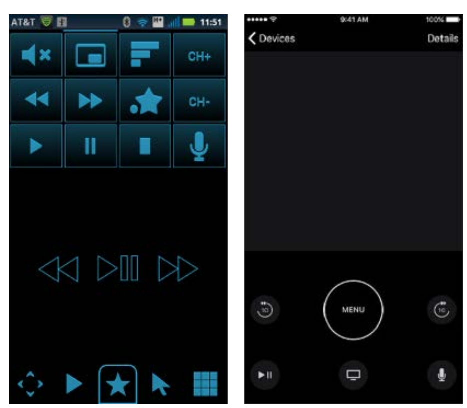


Figure 2. Able Remote on the left and Apple Tv remote on the right.

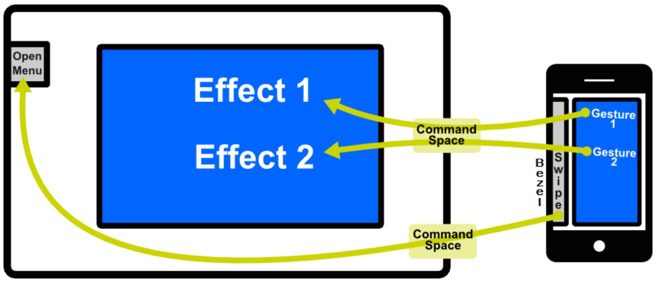


Figure 3. Design by associations defining command spaces.

Figure 4. Tv Channel

The importance of animated transitions is well-known in the literature [17]. A Google statement represents our vision quite well: "*sometimes, it is difficult for a user to know where to look or understand how an element got from point A to point B. Carefully choreographed motion design [...] guide the user's attention [...] avoid confusion when layouts change or elements are rearranged [...] Motion design should serve a functional purpose*" [18]. Examples of transitions are: dragging to scroll from a picture (point A) to another picture (point B); change the status of a side menu from closed (point A) to open (point B).

Finally, since users generally prefer to use smartphones with just one hand [6], we designed interactions accordingly: smartphone control interfaces permit gestures (i.e. tap, swipe, drag) that can be performed with just one hand.

MERGING SMARTPHONES AND LARGE SCREENS

We developed eight interaction techniques to merge smartphones and large screens. These interaction techniques define eight related widgets: (1) side menu (so-called hamburger), (2) list of elements, (3) matrix of elements, (4) list of lists of elements, (5) carousel of elements, (6) audio/video control, (7) map control and (8) text insertion. They were inspired by mobile touch interaction widgets [24] (Figure 1).

In this section, we describe each interaction technique discussing their relevant features. The complementary video at vimeo.com/132462813 shows all the techniques described in this section.

To test our interaction techniques in a real system, we designed a smart TV prototype that includes all of them. A side menu (widget 1), which works like a 'menu button' of a smart TV, lets users select the function of their interest from a list (widget 2): navigate movie posters (widget 5) to watch trailers (widget 6); select TV channels (widget 3); navigate a list of soccer games to read (widget 4) and add comments (widget 8); navigate a map to identify the nearest public place to go and watch soccer games (widget 7).

As represented in Figure 3, the smartphone interface is split in two parts: a central area on which users can operate according to the element displayed at the center of the large screen, and a left-side area to open the side menu.

Widget 1: side menus

The smartphone widget displayed in Figure 1D inspired the side menu. In our smart TV, the side menu is a static element always available in any situation. Starting from the situation in Figure 4, the side menu can be opened performing a left bezel-swipe on the smartphone. Once the side menu is open, the situation is the one in Figure 5. Here, users can navigate the smart TV menu by swiping/dragging vertically and select a function[1] by tapping in the grey zone of the smartphone. Alternatively, users can close the menu (without selecting any function) performing a right bezel-swipe to back to the situation in Figure 4.

Figure 5. Opened menu

There are few relevant aspects worth discussing. We observed that users operating the prototype the first time, often perform a tap on the hamburger symbol to open the menu. This compels the user to look at the smartphone, which does not correspond to our design concept: users should perform a bezel-swipe because it is the only gesture that can be used blindly. In order to induce the right behavior, we developed a stratagem: when tapping the hamburger, the side menu displays a horizontal "bouncing" effect (both on the smartphone and the large screen) for a few seconds without being opened. This animated effect usually induces users to perform a bezel-swipe as required.

Another relevant aspect is the continuous transition while opening of the menu. During the opening of the menu (while performing the beze-swipe), the menu on the large screen displays a continuous opening so that users can be sure they hooked the handle of the menu correctly.

Similar side menus controlled by smartphones have not been found in research prototypes or commercial applications. In particular, the aforementioned Able Remote and Apple TV

[1] In this case, the side menu will be closed automatically.

Remote app implement buttons to operate menus compelling users to look at the smartphone.

Widget 2, 3 and 4: sets of elements

Widget 2: lists of elements

Lists of elements were briefly discussed before: the set of functions displayed on the smart TV menu (Figure 5) is a list of elements. Users can navigate a list by swiping/dragging vertically and select an element by tapping. If a list is longer than its frame container, scroll occurs automatically while navigating it.

Widget 3: grid lists

In our prototype, the available TV channels are organized in a (6xN) grid list (Figure 4). Users can navigate them by swiping/dragging in all directions selecting a channel by tapping. If a grid list is longer/larger than its frame container, scroll occurs automatically in both direction while navigating it.

Widget 4: lists of lists

A list of lists is composed of a father list in which each element is associated with a child list of variable length, which is independent of its sibling lists. The function “Social Soccer” of our smart TV lets users discuss soccer games. In particular, each of them is an element of the father list whereas the related user’s comments compose the child list. For example, Figure 6 displays the related comment (child list) of the father element ‘Milan-Inter’.

Figure 6. List of lists

Navigation occurs in two directions: swiping/dragging horizontally to move between soccer games (elements of the father list), and vertically to scroll the comments of the selected games (child list). Note that comments of each game can be considered like simple lists (widget 2).

Discussion of widget 2, 3 and 4

Touch interaction provides more than discrete steps and directions (i.e. arrow keys in remote controls). In fact, gestures may have different speeds, accelerations and directions. We have considered this aspect to design the scroll mechanism of lists. When scrolling a list, its movement is faster or slower according to the speed and acceleration of the gesture. Naturally, the navigation among the elements of lists remains discrete so that only one element per time can have the focus.

Solutions like Able Remote and Apple TV Remote app for IOS allow the navigation of elements (lists, matrices and lists of lists) in different ways. Able Remote is quite rudimentary: the navigation among elements is allowed by swiping. In particular, each swipe simulates the press of an arrow key on remote controls. Such a solution is questionable since the full potential of continuous gestures is not exploited. The Apple TV Remote for IOS, instead, is pretty similar to our solution.

Widget 5: Carousels of elements

A common way to represent and navigate elements is through the so-called *carousel*, which puts items (e.g. pictures) side by side with only one of them visible at a time (widget in Figure 1A). By using the function ‘Movies’ of the smart TV, users can discover new ones. They are presented using a carousel of movie posters (Figure 7).

Figure 7. Carousel of movie posters.

Navigation is allowed by swiping/dragging leftward (or rightward) to navigate movie posters. The effect on the large screen is to slide the current movie poster out of the frame container and let in the next (or previous) one. Moreover, any element of the carousel can be accessed by tapping: for example, tapping on a movie poster, the related video trailer will be opened.

In our design, the animated transitions on the large screen occur while the user perform the gestures, exactly as on smartphone apps (e.g. Photos on iPhone or Gallery on Android), and therefore according to gesture speed and acceleration.

The scroll of a carousel is also shown in [12]. Nevertheless, the transition of pictures on the large screen is not continuous. In fact, it does not work according to the speed and acceleration of the gesture (i.e. swipe/drag). The feature is not explained in [12], but it is presented in a video[2].

Widget 6: Audio and video control

Opening a video trailer (Figure 8) from the carousel, the video can be controlled by dragging/swiping leftward or rightward to fast forward or rewind and tapping to play/stop. The same control can be applied to control audio elements. During the fast forward (or rewind), the large screen displays the current time (mm:ss), which is updated in real-time according to the speed and acceleration of the gesture performed on the smartphone. This behavior ensures feedback to users.

Able Remote provides similar gestures to control videos, but despite gestures on the smartphone are continuous, the rewind/fast forward displayed on the large screen works “step by step”. This may cause usability issues since gestures work

[2] http://vimeo.com/10777933 (01:04).

Figure 8. Video

similarly to a (discrete) button losing the advantage of continuity (users cannot choose precisely the position on the video).

Widget 7: Map control

The smart TV prototype lets users navigate a map (Figure 9) to identify the nearest public place to go and watch soccer games ('Soccer Bar' function): the map contains markers, i.e., the bars which broadcast the matches.

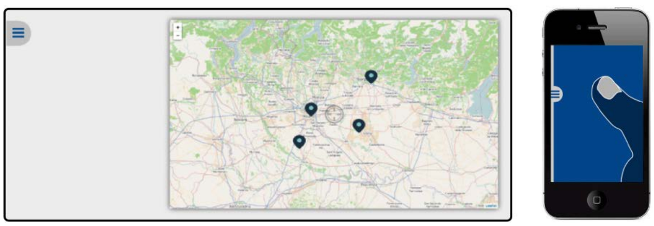


Figure 9. Map

Navigation

Maps are continuous elements. Therefore, they can be navigated by dragging/swiping in all directions. The movement of the map occurs according to the speed, acceleration and direction of the performed gestures on the smartphone.

Zoom

Since we employ gestures that can be performed with just one hand, the common pinch (to zoom) cannot be used. Therefore, we designed a specific gesture: touching and holding the screen, the current zoom level is displayed on the large screen. Then, the zoom level can be adjusted by dragging the finger upward or downward. This solution is quite similar to the one offered by Google Maps to adjust zoom level with one hand.

Opening of markers

We designed maps so that they contain a viewfinder, which is placed at their center (see Figure 9). When a marker is placed under the viewfinder, this can be opened by tapping.

Similar works in the literature were not found.

Widget 8: Text insertion

The smart TV allows the insertion of comments to discuss soccer games. Starting from the situation in Figure 6, users can insert comments by performing a double tap going to the situation in Figure 10. Here, users can write a comment by using the touch keyboard that appears on the smartphone, and then, send it by tapping on the enter button. The comment will be added to the child list of the current soccer game (father element) as new element. Users can also cancel the insertion of the new comment by tapping on the 'cancel' button on the smartphone.

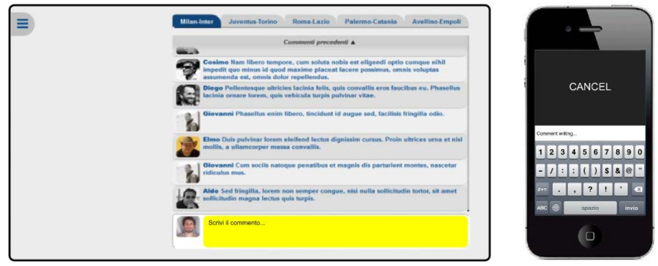


Figure 10. Text insertion

There are many commercial applications that let users exploit smartphone touch keyboards for inserting text on desktop applications. Nevertheless, in our design, the touch keyboard opens automatically according to the context (e.g. when inserting a comment) making its usage, somehow, "straight-away".

WIDGET DESIGN DISCUSSION

The biggest design effort was to create a consistent interaction among the different widgets, so that their combination can be exploited in order to design different kinds of widget-based applications. Moreover, the gestures required to control our widgets are quite similar to the ones used on mobile touch interaction (Figure 1) cutting the learning curve arousing a sense of familiarity in the user.

A design remark is that smartphone UI never display instructions about gestures. They are displayed on the large screen (see the video figure) instead. The first prototypes we developed (but also, e.g., Able Remote) display the instructions about gestures on the smartphone UI. Even if it could seem the most obvious design rationale, we observed that it induces users to behave inappropriately: displaying gesture symbols on the smartphone screen creates a strong affordance in the user. This means that users execute gestures looking at the smartphone screen using gesture symbols as a starting anchor point. Gesture instructions on the large screens, instead, induce users to act on the smartphone using the whole screen area as requested. It is noteworthy that also Apple Tv remote does not show any gesture symbol on the smartphone interface, as displayed in Figure 2-right.

FIRST USER STUDY

In order to assess usability and naturalness of our techniques, we carried out a user study on the smart TV prototype. It involved 48 users: 13 (27%) women and 35 (73%) men. They were split into three age brackets: 20-35 (28 users, 58%), 35-50 and more than 50 (10 participants each, 21%). One user asserted to have never used multi-touch devices whereas 6 users asserted to use touchscreen devices occasionally; the rest (41 users, 85%) asserted to use touch devices regularly. Moreover, we observed that 31 users (64%) carried out the test with one hand, whereas the rest (36%) with two hands. We asked the latter ones to try again the prototype with just one hand and to report their opinion in a specific item of the questionnaire. In order to evaluate the spontaneity of use, users were split into two groups: the first one (21 users, 43%) was trained whereas the second (27 users) was not.

Experimental Setting

The prototype was executed using a desktop PC connected to two full HD 48-inch monitors in mirror mode. The first monitor, placed at around three meters from the participant seat, was used for the execution of the tests. The second one was placed beside this seat so that we could make the video of each test for subsequent analysis (Figure 11). The touch device was an Apple iPod 4.

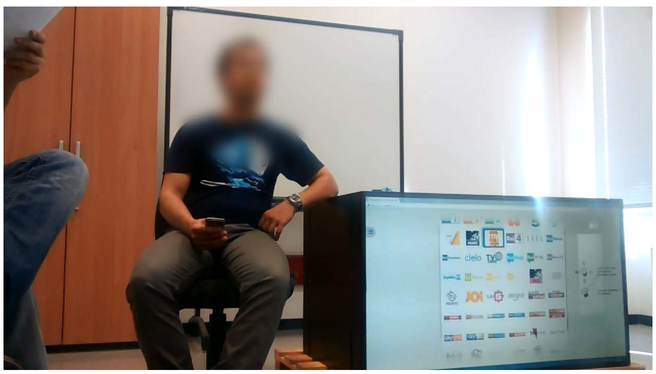


Figure 11. A screen of the recorded videos during a test.

Experimental procedure

An introductory video was shown to each participant. The video displayed just the function of the smart TV without any reference to the needed interactions to control it. After the introductory video, a group proceeded with the tests without any training while the other group received it. During the training, the interactions with the prototype were explained through several videos and clarifications of the experimenter (if needed). A video was made for each interaction technique showing the gestures to perform on the smartphone and the corresponding effects on the large screen. During the training, users were requested to try each interaction to familiarize with the prototype.

The test session encompassed several micro-tasks: at the begin, the prototype displayed a list of movie posters and the experimenter asked users to 1) scroll all posters until the last one, 2) go back to 'The lord of the rings', 3) select it to watch the trailer; 4) fast forward to the time 1m55s of the trailer, 5) stop the trailer, 6) resume the trailer, 7) open the TV channel list using the side menu, 8) select 'Disney Channel', 9) select 'LA7' channel, 10) open the "social soccer" function from the side menu, 11) select the soccer game 'Palermo-Catania', 12) scroll the comments, 13) insert a comment, 14) open "soccer in the bar" function from the side menu, 15) move the map centering the nearest bar to the university; 16) adjust the zoom level up to 13; 17) open the additional information about the selected bar.

Whenever problems occurred or the participants made any mistake, the experimenter assisted them only after three consecutive failures. The first kind of help was the suggestion to see the instructions that were available in each screen of the application. In case the participant was not able to complete the task, even in spite of the instructions, the experimenter explained how to proceed.

After the test, the experimenter asked each participant to continue to use the prototype freely for two minutes. After that, the experimenter asked to use the prototype for two additional minutes trying to avoid looking at the smartphone, if possible. Moreover, when users executed the test with two hands, the experimenter asked to use the prototype for other two minutes using only one hand. Finally, the experimenter invited the participants to fill in a questionnaire that consisted of three sections: personal information (age and perceived familiarity with touch screen devices); the prototype evaluation; and a small free-text comment area.

The prototype evaluation

Participants could evaluate the interaction technique on a 6-point scale, from 1 (lowest quality) to 6 (highest quality) [11]. The items regarded: *usability*, in terms of easiness of use; *naturalness*; *usage with only one hand*, evaluated in case the participants were observed using the system with two hands; *willingness of use*, to be evaluated assuming that the system were mandatory to use; *likeliness of use*, to be evaluated assuming the use of the system as voluntary, that is, differently from the previous case, not mandatory; and, finally, a general assessment of *satisfaction* towards the prototype and the gestures adopted. Moreover, the video of each test was analyzed to extract times of execution, number of errors i.e., any action performed by the user that causes unwanted effects on the large screen, and requests of help from the experimenter.

Results

The Questionnaire

A binomial test was carried out to detect positive or negative tendencies on the responses of the questionnaire. The 6-point ordinal scale was chosen because of its balanced condition between negative values (1, 2 and 3) and positive values (4, 5 and 6). Significant positive tendencies ($p<.05$) were detected for all of the items taken into account: usability (proportion .02 vs. .98, $p<.001$), naturalness (.09 vs. .91, $p<.001$), usage with one hand (.12 vs. .88, $p = .002$), usage without looking at the smartphone screen (.10 vs. .90, $p <.001$), likeliness of use (.08 vs. .92, $p<.001$), willingness to use (.08 vs. .92, $p <.001$) and general assessment (.02 vs. .98, $p <.001$). The results are summarized in Figure 12 through Boxplots with modes (blue) and means (red[3]).

A further analysis on the questionnaire was carried out also according to the different age brackets and familiarities with the touch screen performing a Kruskal-Wallis test. We found significant differences regarding the age brackets for usability (χ^2=6.018, df=2, p=.049) with a mean rank score of 27.18 for the 25-35 age bracket, 24.85 for the 35-50 age bracket and 16.65 for the 50+ age bracket. Concerning naturalness, we also found a significant difference between age brackets (χ^2=7,325, df=2, p=.026) with a mean rank score of 26.27

[3]Although reporting the mean is not entirely appropriate for an ordinal variable, we display it for merely descriptive aims and the sake of comprehensiveness.

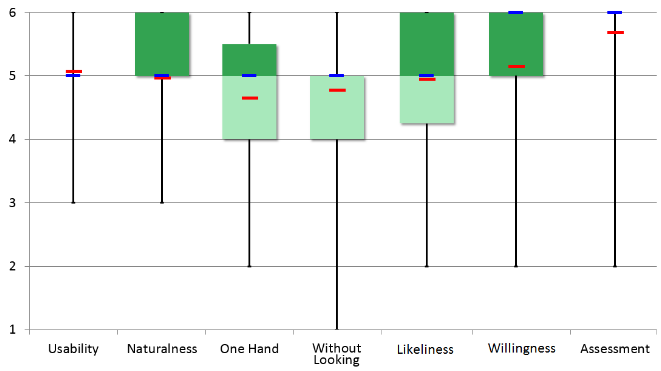


Figure 12. Boxplots show response distribution with mean depicted in red and mode in blue.

for the 25-35 age bracket, 27.44 for the 35-50 age bracket and 14.55 for the 50+ age bracket. As expected, we also found a negative correlation between age brackets and usability (r=-.325, p=.024) and age brackets and naturalness (r=-.299, p=.041). We found no significant differences regarding the different familiarities with multi-touch devices.

To discover whether there were differences between trained and untrained group, we performed a Mann-Whitney test. No significant differences were found. Therefore, our test sample perceived our interaction techniques similarly regardless of the training (see Figure 13). Those results can be seen as a good indicator of spontaneity of use: since no difference was found, we can state that Touch&Screen resulted to be simple to use even without training.

Times of execution, number of errors and requests of help

Differences between the trained group and the untrained group were also investigated according to execution times, number of errors and helps provided by the experimenter (after three consecutive failures). Since some data were not normally distributed (*Shapiro-Wilk* test), we carried out the *Mann-Whitney* test to evaluate whether two distributions are statistically different. Regarding execution time (see Figure 14-Left), the participants who received the training were significantly faster than the participants who did not receive it (M: 80, SD: 25 vs. M: 128, SD: 44) (U=71, Z=-4.4, p<.001). Regarding the number of errors (see Figure 14-Center), the trained participants made a number of mistakes significantly lower than the untrained ones (M: 1,9, SD: 1.5 vs. M: 3, SD: 1.3) (U=151, Z=-2.8, p=.005). This result considers only the errors for which the participants were able to find a solution autonomously without the explicit help of the experimenter. Regarding the need of help from the experimenter (see Figure 14-Right), the trained participants did not ask for any help, whereas untrained ones asked 0.7 help (mean per person, SD=.9). This difference is significant (U=147, Z=-3.6, p<.001).

The three measures are clearly correlated: differences between these two groups were somehow expected. Regarding the times of execution, we expected a significant difference because the interactions could seem unusual and therefore puzzles the users at first. That is why, the number of errors is higher in the untrained group. Anyhow, we observed that users learned the interaction mechanisms from their errors. Moreover, users were usually able to correct their errors autonomously. This means that the participants were generally aware of what was going on during the interactions. In the cases in which they were not able to self-correct their errors (20 times), in 9 cases reading the instructions was enough. Anyway, trained users were always able to correct their errors by themselves. This indicates that (good) training reduces the need of external help. Finally, we could state that 48 seconds (the difference between the execution times of the trained and the untrained group) is the average amount of time needed for understanding all the interaction techniques without any guidance (at least in the cases in which users were able to correct their errors autonomously). Therefore, considering that only 48 seconds of "trials and errors" were enough to understand all the interaction techniques, we could state that Touch&Screen ensures a good spontaneity of use.

Content analysis of users' comments

26 users out of 48 (54%) have also written their comments on the quality of the interaction. We can extract four macro themes as follows: 1) usability of the sidebar menu, 2) initial impact, 3) quality of interaction and 4) sensitivity of the interaction.

1. *Usability of the sidebar menu.* The bezel-swipe was designed to be used without looking at the smartphone: users should be able to use the device border as referring feedback to "pull" the menu out. Many users were able to use the sidebar menu as required. On the contrary, other users had some difficulties. Some of them were able to understand how to open the side menu only after many attempts (they tapped on the hamburger instead of performing the bezel-swipe). Moreover, one user reported that the swipe necessary to open the sidebar menu is too wide: in his opinion, a shortest gesture would be preferable. Two users would have preferred a largest handle area in order to facilitate the opening of the sidebar menu.
2. *Initial impact.* One user reported that he had "initial feeling problems, but [the interaction was] very usable and natural". Even if only one user reported this problem, we observed that many users experienced some problem at the beginning of the test. Fortunately, after a few seconds, most of them felt comfortable with the interaction.
3. *Quality of interaction.* Two users reported that the interaction was easy and "amusing" to use. This could be traced back to a novelty effect, which usually declines over time. Anyway, this indicate that the interaction was appreciated and could be used in different contexts pleasantly.
4. Sensitivity of interaction. Many users reported that the smartphone touch detection was too sensitive. The problems occurred during the choice of TV channels (for two users, the switch among channels was too fast), the fast forwarding of the video (for one user it was difficult to choose the exact point in the timeline), the control of the map (for two people, map movements and the zoom level adjustment were too fast). Although we purposely set a high sensitivity to stress the test condition, it can be calibrated differently according to the contexts and even customized by the user (mostly like mouse-speed calibration on desktop systems).

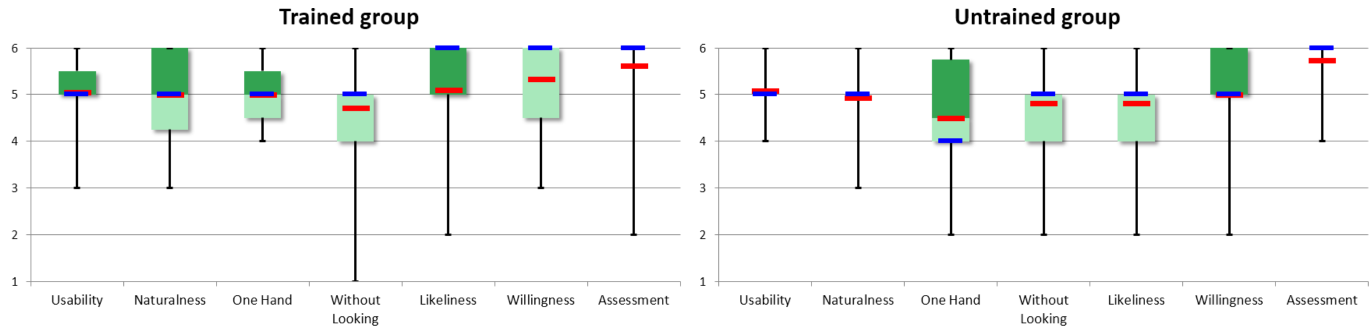


Figure 13. Boxplots show responses of trained and untrained group. No significant differences were found between the groups.

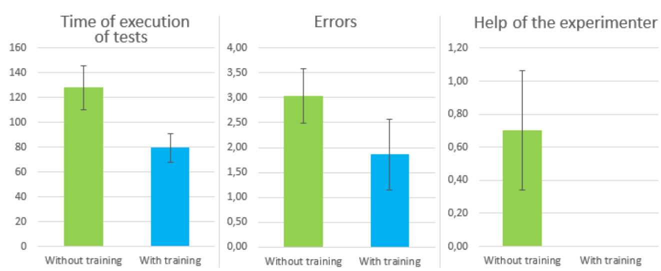


Figure 14. Trained vs. Untrained: time of execution, errors, and help of the experimenter. All differences are significant (p<.05). Error bars denote 95% CI.

SECOND USER STUDY

We evaluated our interaction techniques comparatively with direct touch and pointer controlled through the touch screen of the smartphone (using PC Remote, a commercial Android app). The study involved 12 users of different ages (M: 27.3, SD: 5.1): 4 (33.3%) women and 8 (66.6%) men. All the users asserted to use touch devices regularly.

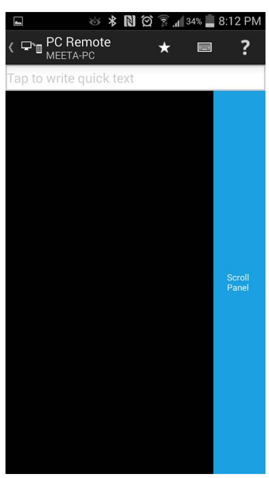


Figure 15. PC Remote for Android.

Experimental Setting

For Touch&Screen and PC Remote, the prototype was executed using a desktop PC connected to a full HD 32-inch monitor. The monitor was placed at around two meters from the participant seat. The touch device was an Apple iPod 4 for the Touch&Screen whereas a THL 5000 Android smartphone for PC Remote. Several applications similar to PC Remote are actually available and we tested some of them. PC Remote was revealed the best, especially because the scroll can be performed using a single hand swiping the finger on the blue region at the right edge of the screen (see Figure 15).

For direct touch, tests were carried out on a Lenovo Flex (15-inch) multitouch laptop.

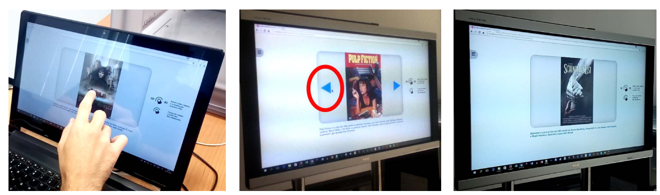


Figure 16. Experimental conditions: direct touch on the left, PC Remote at the center and Touch&Screen on the right.

The prototype UI is exactly the same for the direct touch and Touch&Screen. Otherwise, since PC Remote uses a mouse pointer, we added two buttons (to be clicked) for scrolling films (see red circle in Figure 16-center). The rest of the UI, instead, was not changed.

Each test in each condition was video-recorded.

Experimental procedure

We designed a within-subject study, therefore, each participant evaluated all the three interaction techniques. To avoid the carryover effect, we randomized the order of the conditions for each user. Users were trained by the experimenter through explanations of the requested interactions according to the different interaction techniques. Users were requested to try each interaction technique for at least two minutes to familiarize with it. When participants used two hands (this occurred especially for Touch&Screen and PC Remote), the experimenter asked them to use only one hand. After familiarizing with the techniques, the actual evaluation could started.

The test session encompassed several micro-tasks mostly like the previous evaluation with some differences as follow. Regarding the video control, the users were requested to fast forward to the center of the trailer. The selected channels were the same of the previous user study (La7 and Disney Junior). Nevertheless, users learned their position in the list before

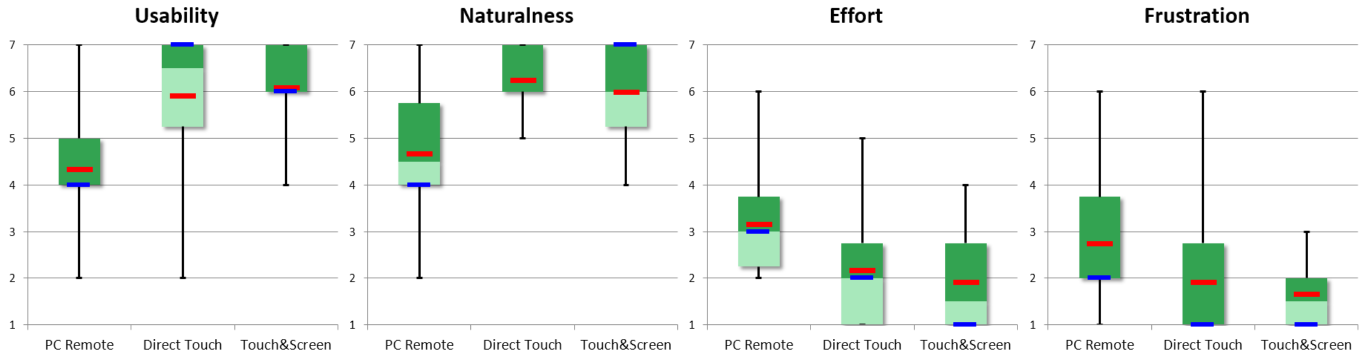


Figure 17. Boxplots show response distributions of usability, naturalness, effort and frustration of the compared techniques with mean depicted in red and mode in blue.

the test (this avoided wasting time while looking for the channels). Regarding the selection of soccer games, users were requested to select 'Palermo-Catania', scroll all the comments and select 'Juventus-Torino' and scroll again the comments. Unlike the previous evaluation, the insertion of a comment was not requested. Regarding the map, users were requested to zoom up to the marker nearest the center of the map and then select that marker.

Unlike the first user study, the smartphone on the Touch&Screen condition displayed a completely black screen instead of the normal smartphone interface. The aim was to push users to use our interaction without looking at the screen.

The prototype evaluation

At the end of each condition, participants evaluated it on a 7-point scale. The items regarded the usability, naturalness, frustration and effort (these last two items were taken from the NASA-TLX questionnaire). Moreover, the videos were analyzed to extract times of execution for each participant and condition.

Results

Questionnaire

To discover differences among usability, naturalness, frustration and effort of the different interaction techniques, we carried out Friedman tests among questionnaire items and a post hoc analysis using Wilcoxon signed-rank tests with a Bonferroni correction applied, resulting in a significance level set at $p<0.017$.

There was a statistically significant difference in usability, χ^2=13.897, p=.001 with the mean rank (the higher the better) of 1.21 for PC Remote, 2.38 for Touch&Screen and 2.42 for direct touch. Post hoc analysis revealed that there was a significant difference (Z=-2.831, p=.005) between Touch&Screen (Mean: 6.1; Mode: 6; Median: 6; SD: 0.90) and PC Remote (Mean: 4.3; Mode: 4; Median: 4; SD: 1.30), while the difference between Touch&Screen and direct touch (Mean: 5.9; Mode: 7; Median: 6.5; SD: 1.56) was not significant (Z=0, p=1). Therefore, while direct touch received the highest evaluations with respect to usability, our system was not found to be signifiicantly different from this latter; moreover these both latter systems resulted to be more usable than the PC remote in a significantly way.

The same holds in regard to naturalness, frustration and effort. In particular, in regard to naturalness there was a statistically significant difference between the three systems (χ^2=12.968, p=.002; PC Remote MR=1.33, Touch&Screen MR=2.21, direct touch MR=2.46), but post hoc analysis did not reveal a significant difference between Touch&Screen and direct touch (Z=-1.000, p=0.317), while both were found to be more natural than PC Remote (Z=-2.379, p=.017).

The same holds for (frustration χ^2=6.545, p=.038, PC Remote MR=2.50, direct touch MR=1.75, for Touch&Screen MR=1.75) and no significant difference was found between Touch&Screen and direct touch (Z=-0.378, p=0.705), while both were less frustrating than PC Remote (Z=-2.140, p=.032); and for effort, for which there was a statistically significant difference in effort between the three systems (χ^2=6.048, p=0.49, PC Remote MR=2.54, Touch&Screen MR=1.75, direct touch MR=1.71), while no significant difference was found between Touch&Screen and direct touch (Z=-0.303, p=.762) and both were requiring significantly less effort than PC Remote (Z=-2.038, p=.042).

Summing up, in comparison with PC Remote, Touch&Screen is perceived as more usable (M: 6.1 vs. M: 4.3), natural (M: 6 vs. M: 4.7) and requiring less effort (M: 1.9 vs. M: 3.1). Finally, it is noteworthy that users perceived the usability and naturalness of Touch&Screen very much similar to the one of direct touch (M: 6.1 vs. M: 5.9 for usability and M: 6.0 vs. M: 6.2 for naturalness), which constitutes a relevant result for us to report.

Execution times

To compare the execution times of the different techniques, we carried out ANOVA repeated measures with a Greenhouse-Geisser correction. The test determined that the difference among mean execution times was statistically significant (F(1.409, 12.683)=41.716, p<0.001). Post hoc tests using the Bonferroni correction revealed that the execution time with Touch&Screen (M: 42.5; SD: 1.899) was faster than the execution time with PC Remote (M: 56.1, SD: 2.429) and the difference was statistically significant (p=.002). On

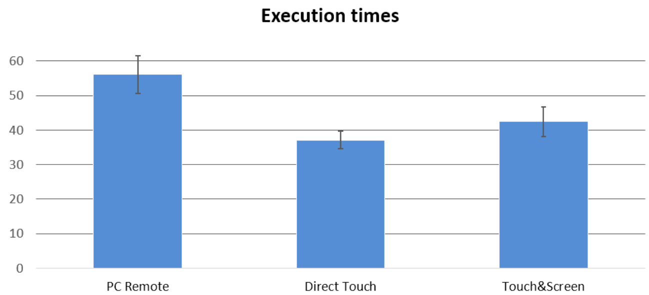


Figure 18. Comparison among the execution times of the different interaction techniques. All differences are significant (p<.05). Error bars denote 95% CI.

the contrary, Touch&Screen was slower than direct touch (M: 37.1; SD: 1.100) and the difference was statistically significant (p=.010). This result was somehow expected, since users have generally experience with touch interaction. However, an improvement of performance with our technique is plausible after some practice with it. At any rate, the most interesting result is that Touch&Screen revealed to be faster than PC Remote (M: 42.5 vs. M: 56.1). Therefore, Touch&Screen can be advantageously used for controlling distant large screens.

CONCLUSIONS

According to the results of the first user study, Touch&Screen was revealed spontaneous to use ensuring a good user experience. The second user study, instead, proved that Touch&Screen is advantageous when compared with smartphone applications for the remote control of distant cursors. Moreover, considering just naturalness and usability, Touch&Screen was revealed really similar to direct touch. We consider this result as an indicator of good design since our main source of inspiration was touch widgets (Figure 1). Finally, the first user study (questionnaire item) indicates - subjectively - that Touch&Screen has good usability without looking at the screen. The second user study confirms this result objectively (we remind that the second user study was carried out having a completely black screen on the smartphone making looking at it useless).

In this work, we tried to consider most of the possible types of content that can be displayed on large screens. Anyway, ways of visualizing content and interacting with it can differ a lot. We hope that our design can be useful to inspire researchers and designers who aim to envision new ways of visualizing and controlling content on large screens.

REFERENCES

1. Alt, F., Shirazi, A. S., Kubitza, T., and Schmidt, A. Interaction techniques for creating and exchanging content with public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 1709–1718.
2. Apple.com. Apple tv remote. `https://itunes.apple.com/gb/app/apple-tv-remote/id1096834193`. Accessed: 2016-14-08.
3. Badam, S. K., and Elmqvist, N. Polychrome: A cross-device framework for collaborative web visualization. In *Proceeding of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ACM (2014), 109–118.
4. Baldauf, M., and Fröhlich, P. The augmented video wall: multi-user ar interaction with public displays. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, ACM (2013), 3015–3018.
5. Beaudouin-Lafon, M., Huot, S., Nancel, M., Mackay, W., Pietriga, E., Primet, R., Wagner, J., Chapuis, O., Pillias, C., Eagan, J. R., et al. Multisurface interaction in the wild room. *Computer 45*, 4 (2012), 48–56.
6. Bellino, A. Enhancing pinch-drag-flick paradigm with two new gestures: Two-finger-tap for tablets and tap&tap for smartphones. In *Human-Computer Interaction–INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings*, vol. 9298, Springer (2015), 534–551.
7. Bellino, A. Face-to-face applications: the usage of large screens to increase co-present social engagement. In *INTERACT 2015 Adjunct Proceedings. 15th IFIP TC.13 International Conference on Human-Computer Interaction 14-18 September 2015, Bamberg, Germany*, University of Bamberg Press (2015), 3–8.
8. Bellino, A., De Michelis, G., and De Paoli, F. Youtube4two: socializing video platform for two co-present people. In *Proceeding of the Companion Publication of the 23rd International Conference on World Wide Web Companion*, International World Wide Web Conference Steering Committee (2014), 131–134.
9. Boring, S., Baur, D., Butz, A., Gustafson, S., and Baudisch, P. Touch projector: mobile interaction through video. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2010), 2287–2296.
10. Boring, S., Jurmu, M., and Butz, A. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *Proceeding of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, ACM (2009), 161–168.
11. Cabitza, F., and Locoro, A. Questionnaires in the design and evaluation of community-oriented technologies. *International Journal of Web-based Communities 13*, 1 (2017).
12. Dachselt, R., and Buchholz, R. Natural throw and tilt interaction between mobile phones and distant displays. In *CHI'09 extended abstracts on Human factors in computing systems*, ACM (2009), 3253–3258.
13. De Michelis, G., Loregian, M., and Martini, P. Directional interaction with large displays using mobile phones. In *PerCom Workshops* (2006), 196–200.

14. Echtler, F., Nestler, S., Dippon, A., and Klinker, G. Supporting casual interactions between board games on public tabletop displays and mobile devices. *Personal and Ubiquitous Computing 13*, 8 (2009), 609–617.
15. Entertailion.com. Able remote app. `http://ableremote.com/`. Accessed: 2016-14-08.
16. Frosini, L., and Paternò, F. User interface distribution in multi-device and multi-user environments with dynamically migrating engines. In *Proceeding of the 2014 ACM SIGCHI symposium on Engineering interactive computing systems*, ACM (2014), 55–64.
17. Gonzalez, C. Does animation in user interfaces improve decision making? In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (1996), 27–34.
18. Google.com. Material design, meaningful transitions. `https://material.google.com/motion/choreography.html`, 2013. Accessed: 2016-14-08.
19. Hardy, R., and Rukzio, E. Touch & interact: touch-based interaction of mobile phones with displays. In *Proceeding of the 10th International Conference on Human computer interaction with mobile devices and services*, ACM (2008), 245–254.
20. Hesselmann, T., Henze, N., and Boll, S. Flashlight: optical communication between mobile phones and interactive tabletops. In *ACM International Conference on Interactive Tabletops and Surfaces*, ACM (2010), 135–138.
21. Liu, J., Zhong, L., Wickramasuriya, J., and Vasudevan, V. uwave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing 5*, 6 (2009), 657–675.
22. McCallum, D. C., and Irani, P. Arc-pad: absolute+ relative cursor positioning for large displays with a mobile touchscreen. In *Proceeding of the 22nd annual ACM symposium on User interface software and technology*, ACM (2009), 153–156.
23. Nancel, M., Chapuis, O., Pietriga, E., Yang, X.-D., Irani, P. P., and Beaudouin-Lafon, M. High-precision pointing on large wall displays using small handheld devices. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 831–840.
24. Neil, T. *Mobile Design Pattern Gallery: UI Patterns for Smartphone Apps*. " O'Reilly Media, Inc.", 2014.
25. Rashid, U., Nacenta, M. A., and Quigley, A. The cost of display switching: a comparison of mobile, large display and hybrid ui configurations. In *Proceeding of the International Working Conference on Advanced Visual Interfaces*, ACM (2012), 99–106.
26. Savolainen, P., Ojala, J., Di Rienzo, A., Nelimarkka, M., Kuikkaniemi, K., Tarkoma, S., Jacucci, G., et al. Web-based public-screen gaming: Insights from deployments. *IEEE Pervasive Computing 15*, 3 (2016), 40–46.
27. Schmidt, D., Seifert, J., Rukzio, E., and Gellersen, H. A cross-device interaction style for mobiles and surfaces. In *Proceeding of the Designing Interactive Systems Conference*, ACM (2012), 318–327.
28. Seifert, J., Bayer, A., and Rukzio, E. Pointerphone: Using mobile phones for direct pointing interactions with remote displays. In *IFIP Conference on Human-Computer Interaction*, Springer (2013), 18–35.
29. Seifert, J., Simeone, A., Schmidt, D., Holleis, P., Reinartz, C., Wagner, M., Gellersen, H., and Rukzio, E. Mobisurf: improving co-located collaboration through integrating mobile devices and interactive surfaces. In *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, ACM (2012), 51–60.
30. Simeone, A. L., Chong, M. K., Sas, C., and Gellersen, H. Select & apply: understanding how users act upon objects across devices. *Personal and Ubiquitous Computing 19*, 5-6 (2015), 881–896.
31. Stellmach, S., and Dachselt, R. Look & touch: gaze-supported target acquisition. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2012), 2981–2990.
32. Stellmach, S., and Dachselt, R. Still looking: Investigating seamless gaze-supported selection, positioning, and manipulation of distant targets. In *Proceeding of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 285–294.
33. Turner, J., Alexander, J., Bulling, A., Schmidt, D., and Gellersen, H. Eye pull, eye push: Moving objects between large screens and personal devices with gaze and touch. In *IFIP Conference on Human-Computer Interaction*, Springer (2013), 170–186.
34. Vajk, T., Coulton, P., Bamford, W., and Edwards, R. Using a mobile phone as a wii-like controller for playing games on a large public display. *International Journal of Computer Games Technology 2008* (2007).
35. Vepsäläinen, J., Di Rienzo, A., Nelimarkka, M., Ojala, J. A., Savolainen, P., Kuikkaniemi, K., Tarkoma, S., and Jacucci, G. Personal device as a controller for interactive surfaces: Usability and utility of different connection methods. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ACM (2015), 201–204.
36. von Zadow, U., Büschel, W., Langner, R., et al. Sleed: Using a sleeve display to interact with touch-sensitive display walls. In *Proceeding of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ACM (2014), 129–138.
37. Winkler, C., Löchtefeld, M., Dobbelstein, D., Krüger, A., and Rukzio, E. Surfacephone: a mobile projection device for single-and multiuser everywhere tabletop interaction. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM (2014), 3513–3522.