Comparing Direct Off-Screen Pointing, Peephole, and Flick&Pinch Interaction for Map Navigation

Khalad Hasan[1], David Ahlström[2], Pourang P. Irani[1]

[1]University of Manitoba,
Winnipeg, Manitoba, Canada
{khalad, irani}@cs.umanitoba.ca

[2]Alpen-Adria-Universität Klagenfurt
Klagenfurt, Austria
david.ahlstroem@aau.at

ABSTRACT

Navigating large workspaces with mobile devices often require users to access information that spatially lies beyond its' viewport. To browse information on such workspaces, two prominent spatially-aware navigation techniques, peephole, and direct off-screen pointing, have been proposed as alternatives to the standard on-screen flick and pinch gestures. Previous studies have shown that both techniques can outperform on-screen gestures in various user tasks, but no prior study has compared the three techniques in a map-based analytic task. In this paper, we examine these two spatially-aware techniques and compare their efficiency to on-screen gestures in a map navigation and exploration scenario. Our study demonstrates that peephole and direct off-screen pointing allows for 30% faster navigation times between workspace locations and that on-screen flick and pinch is superior for accurate retrieval of workspace content.

Categories and Subject Descriptors

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces – Interaction styles.

Keywords

Spatial interaction; Direct off-screen pointing; Peephole displays;

1. INTRODUCTION

Mobile devices with touchscreens are often used to navigate large information spaces such as maps, text documents, images, and web pages. These information spaces are often considerably larger than what can be viewed at once on the available display. Consequently, the user has to engage in extensive navigation activities, using pinch and flick gestures, to view information that is located in distant parts. Using these gestures to explore large information spaces, however, often involves considerable effort [6] and the user has to deal with screen occlusion and fat-finger situations [16].

As an alternative to on-screen gestures, numerous researchers have investigated peephole displays (e.g., [4, 9, 10, 12, 13, 17]) where the mobile device is aware of its position in relation to the workspace that is 'pinned' to the environment directly surrounding the device. As illustrated in Figure 1a, with a peephole display, the content shown on the screen is updated according to device movements across the underlying workspace: for example, moving the device 5cm to the left shows content located 5cm to the left in the workspace. Several earlier projects have compared the efficiency of a peephole display to the use of pinch and flick interaction in workspace navigation tasks, but with partly contradicting results. While some studies demonstrate clear advantages for the spatially-aware peephole technique [13, 16], others report similar [9] or inferior [8, 10] navigation performance compared to flick and pinch (Rädle et al. [14] provide a comprehensive overview of peephole vs. flick and pinch studies).

Figure 1. Map navigation with a peephole display (a) and with direct off-screen pointing (b).

Another promising and envisioned approach–motivated by the rapid advancement and miniaturization of optical sensing and tracking technologies–to improve interaction on mobile devices is to make the device spatially-aware of finger gestures in the air around the device. For instance, Jones et al. [8] demonstrate the potential of using in-air gestures instead of on-screen gestures to navigate within large workspaces. Ens et al. [3] explore the concept of directly pointing into a virtual workspace that extends beyond the screen where items that reside outside the screen are selected by moving a finger to their corresponding in-air location. Ens et al. present a performance model which captures this type of 'direct off-screen pointing' and show that user performance is largely dependent on on-screen guidance cues, such as a mini-map or an overview of the workspace. Hasan et al. [6] expand on the idea and study how off-screen space can be divided into discrete 'storage bins' with bin-content being shown on the screen as the tracked off-screen pointing finger moves from bin to bin.

Despite the growing interest in identifying how best to harness spatially-aware techniques, such as peephole displays and in-air finger detection, for improving workspace navigation on mobile devices, we know little of how these two approaches perform in comparison to one another. To this end we contribute an empirical understanding of their strengths and limitations. We explore how direct off-screen pointing can be used for navigating and browsing content in a continuous workspace, such as a map. Figure 1b illustrates: the device tracks the user's finger and updates screen content according to the off-screen finger's current position so that the content 'under' the finger is displayed on the screen. In a user study we compare the performance of off-screen pointing against the standard on-screen flick and pinch interaction and the peephole technique in a map navigation task. Results show that participants were faster navigating the workspace with the spatially-aware techniques, particularly in the presence of an on-screen visual cue.

SUI '15, August 08 - 09, 2015, Los Angeles, CA, USA
ISBN 978-1-4503-3703-8/15/08 $15.00
DOI: http://dx.doi.org/10.1145/2788940.2788957

2. EXPERIMENT

Task: Our experimental task captured a scenario wherein a user is interested in examining biometric and environmental data collected with a sports bracelet during a morning jog. After the jog, the user can combine the recorded data with geographical information for detailed analysis in an exercise app on a smartphone. Jogging-related information–heart rate, running speed, elevation, air temperature, and humidity–are shown in callout boxes associated with markers positioned on a digital map, as shown in Figure 2a and b. In each experimental trial, the participant was prompted to find and select the marker containing the highest or lowest, value on one of the five variables, such as “Find the location with the greatest elevation” or “Find the location with the highest speed”.

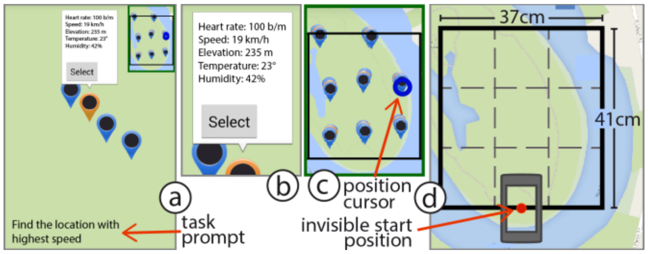

Figure 2. a) Task screen with an open callout box, the overview, and task prompt. b) Enlarged callout box. c) Enlarged overview. d) Map area with black guiding boundary.

Figure 2d shows the map we used. A black outline defined the relevant map area to the participants. We divided the relevant area into nine cells (dashed lines in Figure 2d, the lines were invisible during the experiment) and placed a cluster of four markers (Figure 2a) in each cell but the upper-right cell. Each cluster contained one potential orange target marker with a callout box displaying exercise information (Figure 2a and b) and three blue markers with callout boxes showing irrelevant information (as commonly found in map applications). We considered the blue markers as distractors. A trial started with a window containing the task prompt and a ‘start’ button. Timing started with a tap on the button. This dismissed the window and displayed the area around a fixed start position at the bottom of the lower middle map cell, as visualized in Figure 2d (the task prompt remained visible but at the bottom of the screen). The map was zoomed to Google Maps’ zoom level 20. At this zoom level, a paper version of the relevant map area would measure 37×41cm. We used a 1:1 mapping for the motor space and the display space where moving the smartphone (peephole) or finger (off-screen pointing) 1cm would move the map area by 1cm.

We encouraged the participants to search the map for the orange target marker that matched the task prompt as fast as possible. When participants believed they had found the correct target, they could end the trial with a tap on the ‘select’ button in the marker’s callout box (Figure 2a & b). Selecting the correct marker ended trial time and displayed the task window for the next trial. An error dialogue was shown if the wrong marker was committed and the search could continue after dismissing the dialogue.

The marker cluster within a map cell and the orange target marker within a cluster were randomly positioned for each trial. Across a series of five trials, the correct target marker was located once in each of five different and randomly selected map cells. Five different task prompts were randomly selected from a pool of ten task prompts for each series of five trials.

Apparatus & Map Navigation Techniques: All participants performed series of trials using direct off-screen pointing (*OP*), a peephole display (*PD*), and the standard on-screen pinch and flick gestures (*FP*) while standing in a quiet room. The study software ran on a Google Nexus 5 smartphone with a 4.95-inch screen (1080×1920 pixels resolution). For the *PD* and *OP* techniques we used a Vicon system (with eight T-Series cameras) to emulate a spatially-aware smartphone and to ensure reliable and noise-free data. The system tracked the position of the smartphone when the *PD* was used, and both the smartphone and the user’s right-hand ‘in-air’ index finger when *OP* was used. The study software was built on Android 4.4.2 (API 19). Google Maps Android API (v2) was used to implement the map features in the study software.

With *FP*, participants held the smartphone in their left hand and performed on-screen manipulations with their dominant right hand: flick to pan the map, pinch to zoom, and tap to select map markers, the ‘start’ button, and the ‘select’ button in callout boxes.

To avoid ambiguities involved in zooming with spatially-aware displays [8, 12], such as identifying suitable interactions for clutching and zoom adjustment, map navigation with *PD* and *OP* was restricted to 2D-panning (zooming was only enabled with the *FP*). Furthermore, during implementation of the spatially-aware techniques we noticed that accurately homing in on and tapping small map markers was cumbersome as any jitter from device–or the user’s in-air pointing finger–causes small erratic displacements of the displayed map area and its markers. As a solution, for *OP* and *PD* we opted for a two-step selection approach similar to TapTap [15]. That is, a first tap anywhere on the screen froze the screen content and a second tap, if on a map marker, opened the corresponding callout box. If the second tap occurred elsewhere, the freeze was released and any open callout box was closed. In cases with an open callout box, an additional tap outside the callout box closed the box and released the freeze. Unlike the original TapTap, we did not scale the target sizes.

When using *OP*, participants held the smartphone in their non-dominant left hand and used their right-hand index finger to point in mid-air around the device to navigate the map. On-screen interactions (taps to freeze the screen, select map markers and buttons) were performed with the left-hand thumb. With *PD*, for stable and accurate map navigation and little on-screen jitter, participants were encouraged to use their dominant right hand to move the peephole display (smartphone) across the stationary map and to use their right-hand thumb or left hand for on-screen interactions (all participants chose to use only their right hand). With *OP* and *PD*, the start position on the map (Figure 2d) was ‘anchored’ to a point in physical space: with *OP* to the center of the phone, with *PD* to a set of tracking markers placed on the floor in front of the participants’ assigned standing position.

Several earlier studies have demonstrated strengths and weaknesses of various techniques to visualize the positions of objects that reside outside the currently visible part of the workspace in different user tasks, such as Halos [2], Wedges [5, or overviews [e.g., 7, 11]. Accordingly, we were also interested in investigating how map navigation performance with the two spatial techniques is influenced by such on-screen guidance. Informed by earlier work [3, 7], we opted for an overview visualization of the relevant map area with marker clusters and a blue circular position cursor that indicated what area of the map was currently displayed on the screen. All three techniques were tested with and without an on-screen overview. In trials with the overview, the overview was positioned in the upper right corner of the screen and roughly took 1/15 of the display space, as shown in Figure 2a.

Measurements: The study software recorded the number of markers that were opened during a trial, the time spent on manipulating markers (incl. time needed to open callout boxes, reading marker information, and to tap the 'select' button and to close callout boxes), the time taken to navigate between markers, the total trial time (marker+navigation time), and the accumulated distance 'travelled' in a trial (measured in centimeters, distances at different zoom levels with *FP* were normalized to the Google Maps' zoom level 20, as used for *PD* and *OP*). We analyzed the results using 2×3 RM-ANOVAs with *guidance* (*No Overview*, *With Overview*) and *technique* (*OP*, *PD*, *FP*) as independent factors.

Participants: Twelve right-handed smartphone owners (8 male) aged 21 to 28 years (mean 25.0, σ=2.7) participated. All were unfamiliar with spatially-aware interaction techniques. Participants performed two sets of five timed trials with each technique. The first set with assistance from an on-screen overview, the second without an overview. The order of techniques was counter-balanced between participants. Before a participant started with a new set of timed trials with a new technique, we demonstrated the technique and the participant had two untimed practice trials. We collected 12 (participants) × 3 (*techniques*) × 2 (*guidance* conditions) × 5 = 360 timed trials. With short breaks and practice trials, each session lasted around 45 minutes.

3. Results

Errors & outliers: Participants committed a wrong target marker once or twice before they committed the correct target marker in 30 of the 360 trials (the erroneous trials were roughly evenly distributed between the six *techniques-guidance* combinations and participants). Understandably, erroneous trials had markedly longer trial times (caused by more marker selections and error messages to react to) than trials where only one marker (the correct one) was committed. We removed the erroneous trials and then identified and removed nine outlier trials with a trial time outside of ± 4 S.D. from the mean time for the corresponding *technique-guidance* combination (outliers were roughly evenly distributed between *technique-guidance* conditions).

Learning: Figure 3a plots the mean trial time for the first, second, third, fourth, and fifth trial with each *technique-guidance* combination. For all combinations we see a trend toward faster times in later trials. However, one-way ANOVAs (one for each *technique-guidance* combination) did not show any significant differences between early and late trials for any of the six *technique-guidance* combinations. We conclude that there were no significant learning effects and continue with all trials.

Marker visits & marker time: An optimal trial would include inspecting the information in all eight target markers and while navigating between these remembering the position of the 'currently best' marker and then, after having inspected all eight target markers, returning back to the correct target marker which includes information matching the task prompt. Accordingly, the optimal trial includes 8+1 or 8 marker visits (8 in cases where the last marker visited contains the correct information). On average, during a trial participants inspected the information in 10.8 callout boxes. There were no significant differences in the mean number of visited markers between the three techniques (*FP* 10.5, *PD* 11.3, *OP* 10.6 markers) or between the two *guidance* conditions (*No Overview* 11.2, *With Overview* 10.5 markers). Unsurprisingly, as shown in Figure 3b, using the dominant hand for on-screen manipulation (*FP*) was significantly faster ($F_{2,22} = 42.0$, $p < 0.001$, $\eta^2 = 0.79$) than using the thumb on the hand holding the device (*PD* and *OP*). Using the thumb took about twice as long as using fingers on the dominant hand (*FP* 13.8s, *PD* 28.0s, *OP* 27.6s). As expected, having access to the on-screen overview did not influence the time needed to manipulate markers.

Navigation time & path length: Overall, across *techniques*, participants needed significantly more time (26%) to navigate between markers in trials with no guiding overview than in trials with a guiding overview (46.8s vs. 37.2s, $F_{1,11} = 32.4$, $p < 0.001$, $\eta^2 = 0.75$). As shown in Figure 3c, the overview was only effective in combination with *FP* or *OP* (interaction effect: $F_{2,22} = 7.6$, $p < 0.01$, $\eta^2 = 0.41$). The overview reduced navigation time by 10.2% when used with *FP* and by 45.0% when used with *OP*.

Overall, across the two *guidance* conditions, navigation time differed depending on *technique* ($F_{2,22} = 42.0$, $p < 0.001$, $\eta^2 = 0.79$). Bonferroni adjusted post-hoc comparisons showed that *FP* with 49.5s required significantly longer navigation time than both *PD* with 38.7s and *OP* with 37.5s (27.7% resp. 31.9% slower, p's < 0.016), which did not differ. Clearly, gross hand and arm movements–as used with a *PD* or *OP*–are more suitable for quick navigation than the familiar, but minute, on-screen panning and zooming actions used with *FP*.

Figure 3c reveals why participants were faster navigating the map when assisted by the overview. With the overview participants were able to navigate more directly towards marker clusters, as visible in the significantly different accumulated distance traversed during trials with and without the overview (160.4cm with and 204.8cm without the overview; $F_{1,11} = 65.5$, $p < 0.001$, $\eta^2 = 0.86$). In this way, navigation time was reduced when the overview was combined with *FP* or with *OP*. User comments provide a likely explanation for the overview's ineffectiveness when combined with *PD*: several participants reported that it was difficult to follow the small position cursor in the overview and to switch visual focus between the overview and the map while moving the smartphone. Accordingly, many participants ignored the overview.

Trial time: Figure 3d summarizes the previous analyses and shows the total trial times (marker time + navigation time) for the six different *technique-guidance* combinations. The bar labels show the percentage of total trial time spent on marker manipulation resp. on navigation between markers. Across the two *guidance* conditions, trial time did not differ between the three techniques: *FP* 63.3s, *OP* 65.1s, *PD* 66.8s. The short *marker time* and relatively long *navigation time* with *FP* sum up to a total trial time comparable to the sum of the long *marker times* and short *navigation times* with *PD* and *OP*.

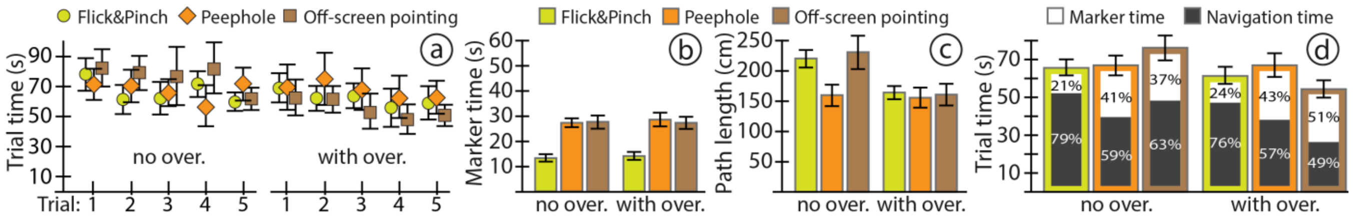

Figure 3. a) Trial time across trial series. b) Marker time. c) Path length. d) Trial time. Error bars show ± 2 S.E.

For all three *techniques*, guidance from the overview significantly reduced trial time ($F_{1,11} = 25.3$, $p < 0.001$, $\eta^2 = 0.70$) by 13.1%, from 69.7s to 60.5s. As visible in Figure 3d, the overview was mainly effective in combination with *OP* (interaction effect: $F_{2,22} = 6.5$, $p < 0.001$, $\eta^2 = 0.37$) where the overview helped participants to reduce their average trial time by nearly a third (29.1%), from 76.1s to 54.0s. *OP* with the overview was the fastest combination at a trial time of 54.0s. *FP* with the overview was the second fastest combination with 61.1s, 13% slower than *OP* with overview.

4. DISCUSSION & CONCLUSIONS

Our results clearly highlight the important advantages and critical disadvantages of each of the techniques we evaluated. Results reveal that traditional flick and pinch gestures exhibit superior marker selection capabilities as it employs fingers on the dominant hand to open marker information. On the contrary, the tested spatially-aware techniques are slower for accessing marker information on maps. This is largely because users used the thumb for such interaction with the spatially-aware interactions, which is less efficient than using the index finger. For navigating from one map location to another, peephole and direct off-screen pointing demonstrate improved performance over flick and pinch. The latter requires many minute operations, such as flicking through screens, resulting in less efficient map navigation. Additionally, all the techniques show strong reliance on the overview. Such an on-screen cue helps users to traverse directly towards the targets, particularly in the case of direct off-screen pointing.

Our results also confirm prior results [6, 12, 17]: performance with spatially-aware techniques can be affected by the chosen task. Our task is one that is common for data exploration and analytics [1], which involves inspecting items across the entire workspace and remembering which ones match best a certain search criteria. Similar results have been reported in prior studies involving analytic tasks [6]. Additionally, flick and pinch could be the choice technique in a task that requires frequent touch interactions. Pahud et al. [12] saw similar results wherein participants were as efficient with flick and pinch as with spatial input.

Overall, we provide a first attempt at examining the performance of two spatial navigation techniques, direct off-screen pointing and peephole with standard flick and pinch. Results show that spatially-aware techniques can be a promising alternative for map navigation. With a controlled experiment, we demonstrate the advantages and limitations of each technique for a map navigation and exploration task. This work could be extended in several directions. In the experimental design, we only considered the design factors critical for the proper operation of the spatially-aware techniques we tested. Further experimentation is needed to extract additional design parameters that are suitable to these techniques. For instance, zooming and clutching mechanisms need to be carefully explored for the spatially-aware techniques as such mechanisms are necessary to enable users to navigate larger workspaces than those limited to within arm-reach, as was the case in our study. Our participants had no previous experience with navigating maps using spatially-aware techniques. We believe this can be addressed through extended exposure (i.e., multiple blocks over multiple days) as seen in [12]. Furthermore, we expected that spatially-aware techniques (peephole and direct off-screen pointing) have pronounced navigation movements which support participants in developing an accurate spatial model of the workspace. The standard touch input, which includes minute pan and zoom actions, could make it harder for the user to develop a sense of the spatial relations. However, the on-screen guidance influenced participants' navigation pattern and assisted them to move directly toward the targets, which may diminish the learning effect. It is worthwhile to explore the learning effects when revisiting objects on the map as repeated exposure and navigation could help eliminate the need for on-screen guidance.

5. REFERENCES

[1] Amar, R., Eagan, J., and Stasko, J. Low-level components of analytic activity in information visualization. In *Proc. INFOVIS 2005*, 111-117.

[2] Baudisch, P. and Rosenholtz, R. Halo: a technique for visualizing off-screen objects. In *Proc. CHI 2003*, 481-488.

[3] Ens, B., Ahlström, D., Cockburn, A., and Irani, P. Characterizing user performance with assisted direct off-screen pointing. In *Proc. MobileHCI 2011*, 485-494.

[4] Fitzmaurice, W.G. Situated information spaces and spatially aware palmtop computers. *Comm. ACM 36*, 7 (1993), 39-49.

[5] Gustafson, S., Baudisch, P., Gutwin, C., and Irani, P. Wedge: clutter-free visualization of off-screen locations. In *Proc. CHI 2008*, 787-796.

[6] Hasan, K., Ahlström, D., & Irani, P. AD-Binning: leveraging around-device space for storing, browsing and retrieving mobile device content. In *Proc. CHI 2013*, 899-908.

[7] Irani, P., Gutwin, C., Patridge, G., and Nezhadasl, M. Techniques for interacting with off-screen content. In *Proc. INTERACT 2007*, 234-249.

[8] Jones, B., Sodhi, R., Forsyth, D., Bailey, B., and Maciocci, G. Around device interaction for multiscale navigation. In *Proc. MobileHCI 2012*, 83-92.

[9] Kaufmann, B. and Ahlström, D. Studying spatial memory and map navigation performance on projector phones with peephole interaction. In *Proc. CHI 2013*, 3173-3176.

[10] Kerber, F., Krüger, A, and Löchtenfeld, M. Investigating the effectiveness of peephole interaction for smartwatches in a map navigation task. In *Proc. MobileHCI 2014*, 291-294.

[11] Nekrasovski, D., Bodnar, A., McGrenere, J., Guimbretière, F., and Munzner, T. An evaluation of pan & zoom and rubber sheet navigation with and without an overview. In *Proc. CHI 2006*, 11-20.

[12] Pahud, M., Hinckley, K., Iqbal, S., Sellen, A., and Buxton, W. Toward compound navigation tasks on mobiles via spatial manipulation. In *Proc. MobileHCI 2013*, 113-122.

[13] Rädle, R., Jetter, H.-C., Butscher, S., and Reiterer, H. The effect of egocentric body movements on users' navigation performance and spatial memory in zoomable user interfaces. In *Proc. ITS 2013*, 23-32.

[14] Rädle, R., Jetter, H-C., Müller, J. & Reiterer, H. Bigger is not always better: display size, performance, and task load during peephole map navigation. In *Proc. CHI 2014,* 4127-4136.

[15] Roudaut, A., Huot, S., and Lecolinet, E. TapTap and MagStick: improving one-handed target acquisition on small touch-screens. In *Proc. AVI 2008*, 146-153.

[16] Siek, K., Rogers, Y., and Connelly, K. Fat finger worries: how older and younger users physically interact with PDAs. In *Proc. INTERACT 2005,* 267-280.

[17] Spindler, M., Schuessler, M., Martsch, M., and Dachselt, R. Pinch-drag-flick vs. spatial input: rethinking zoom & pan on mobile displays. In *Proc. CHI 2014*, 1113-1122.