Virtual Shelves: Interactions with Orientation Aware Devices

Frank Chun Yat Li, David Dearman & Khai N. Truong
Department of Computer Science
University of Toronto
Toronto, ON M5S 3G4, Canada
{frankli, dearman, khai}@cs.toronto.edu

ABSTRACT
Triggering shortcuts or actions on a mobile device often requires a long sequence of key presses. Because the functions of buttons are highly dependent on the current application's context, users are required to look at the display during interaction, even in many mobile situations when eyes-free interactions may be preferable. We present Virtual Shelves, a technique to trigger programmable shortcuts that leverages the user's spatial awareness and kinesthetic memory. With Virtual Shelves, the user triggers shortcuts by orienting a spatially-aware mobile device within the circular hemisphere in front of her. This space is segmented into definable and selectable regions along the phi and theta planes. We show that users can accurately point to 7 regions on the theta and 4 regions on the phi plane using only their kinesthetic memory. Building upon these results, we then evaluate a proof-of-concept prototype of the Virtual Shelves using a Nokia N93. The results show that Virtual Shelves is faster than the N93's native interface for common mobile phone tasks.

ACM Classification: H5.2 [Information interfaces and presentation]: Input devices and strategies.

General Terms: Design, Experimentation, Human Factors

INTRODUCTION

Mobile devices are appropriating many of the tasks (*e.g.,* Web browsing, gaming, and multimedia) that were once constrained to desktop/laptop computers. However, the physical design of handheld devices and their limited methods of input impose significant challenges for interacting with a growing set of applications. To perform even the simplest of tasks (*e.g.*, changing the calendar view from day to week; selecting a new song from a playlist) typically requires a lengthy sequence of key presses. Moreover, because the function of buttons are highly dependent on the current application's context, users are required to view their mobile device's display during interaction, even in many mobile situations when eyes-free interactions may be preferable.

In contrast to performing a task using a mobile device, many interactions in the physical world often are facilitated by the person's spatial awareness and kinesthetic memory. For example, it is very easy for an experienced driver to shift gears or change radio stations while driving without having to look at the specific controls. The driver has developed a spatial awareness of these controls' positions relative to her body through experience and repetition, enabling eyes-free use. Robertson *et al.* [10] and Tan *et al.* [11] have previously demonstrated that laptop and desktop computers have the screen space to support different methods to spatially organize files and shortcuts. However, the small displays used for mobile devices do not afford information organization and access in the same way. Gray and Fu have shown previously that spatial and kinesthetic memory, even if imperfect, is used when visual information is not easily retrievable [7]. We want to leverage a person's spatial awareness and kinesthetic memory to provide a quick, eyes-free interaction technique for mobile devices.

We developed an interaction technique called Virtual Shelves to augment the native menu interface by providing a new way to access items quickly. Users are able to trigger customizable shortcuts (*e.g.,* accessing a bookmark, initiating a text message to mom) by holding down a key on their spatially aware mobile device, then orienting it towards a specific direction (indexed by the theta and phi angles, see Figure 1: left) within the circular hemisphere in front of her body, and releasing the key to select the shortcut. Studies have shown that users perform better when using spatially consistent interfaces [4, 6, 10] and that interfaces without spatial consistencies degrade performance [5, 12]. Because of this need for consistency, Virtual Shelves centers on the user as the main point of rotation and uses equal angular width and height for shelves. Similar to physical shelves, each shelf has a predefined spatial position that does not overlap with another one. Although only one item is stored

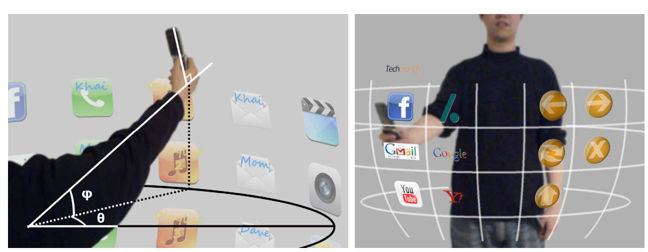

Figure 1: Virtual Shelves. Left: A user orienting a phone to trigger shortcuts. Right: Web browser specific shelves available to the user after the user has begun to use that application.

UIST'09, October 4–7, 2009, Victoria, British Columbia, Canada.

in a shelf at a time, the system itself is context sensitive and the shelf contents are application dependent. For example, if the user is browsing the Web, the shelves can now contain bookmarks and navigational buttons such as back or refresh (see Figure 1: right).

RELATED WORK

Mobile devices today can be easily augmented with additional sensing (*e.g.,* proximity, touch and tilt) to support quicker and more natural methods of interaction. In fact, many of the context sensors demonstrated by Hinckley *et al.* [8] are now common in numerous commercial products. As a result, many researchers have begun to explore novel interaction techniques that can be supported on sensor-enriched mobile devices. For example, a device aware of its 3D orientation can support user interaction with menus, maps, and 3D scenes [9]. Devices aware of their spatial position can act as Peephole [13] or Flashlight [2] displays, allowing users to organize and access content in a virtual electronic space. Our Virtual Shelves technique extends the Peephole and Flashlight techniques, but treats the space around the user as a discrete set of regions (or "shelves"). Whereas Peephole and Flashlight displays allow "peeping" into a larger, dynamic information space and do not require spatial memory to use, we designed Virtual Shelves such that the organization of content is consistent within the virtual space and leverages spatial memory.

EXPERIMENT 1 – MEASURING KINESTHETIC ACCURACY

Although previous works have studied and provided general descriptions about a person's ability to use their kinesthetic memory [1, 3], we want to explore the efficacy of the Virtual Shelf metaphor by specifically measuring how accurately participants could select targets along the theta and phi planes (see Figure 2). Informed by pilot-testing, we chose to use seven targets along each of the theta and phi planes; 14 targets in total (as a starting point, with the intent to run more subjects with more or less targets depending on the results).

Instruments

We used six Vicon M2 cameras to capture the complete 3D orientation of a Nintendo Wiimote. A custom application received the orientation data of the Wiimote from the Vicon motion-tracking server and the button presses from the Wiimote.

The circular hemisphere in front of the participant was segmented into 14 targets; seven targets along the theta and phi planes. The seven targets are distributed evenly along each plane, every 30º. Rather than asking participants to select the targets as identified by degrees, we labeled each target using a clock metaphor (see Figure 2).

Procedure

Eighteen participants (8 male, 10 female), of various handedness (13 right; 5 left) were recruited to perform the evaluation. The age of participants ranged from 18–25 (8), 26–35 (8), 36–45 (1) and 46–55 (1).

Participants were instructed verbally to select each target along the phi and theta planes 10 times using the Wiimote; 140 selections in total. Participants were required to perform the selections blindfolded to ensure that they relied on their spatial and kinesthetic memory, rather than associating targets to physical objects located in the study environment. Each plane was separated into its own block and fully counterbalanced. The ordering of which targets to select was random, but consistent across participants. Participants were trained to use the Wiimote and the locations of the different targets. Training lasted until participants felt comfortable with the shelf metaphor. Vibrotactile feedback was provided during the training phase to help participants identify the different targets and facilitate development of their spatial memory.

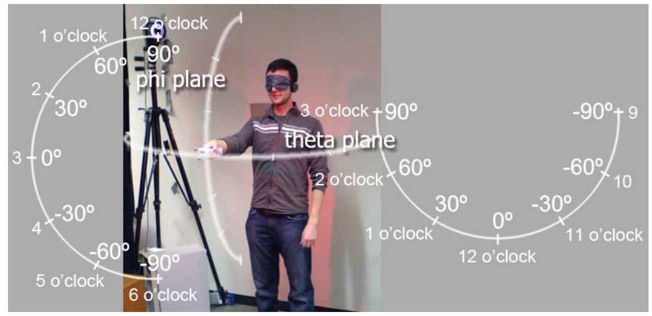

Figure 2: The setup for experiment one. A participant is blindfolded and asked to point a Wiimote to a specific target along the theta/phi plane.

Each selection task started with the participant's arm relaxed against the side her body. The target to select was presented verbally (using the clock metaphor) to the participant over a Bluetooth headset. To perform a selection, participants would hold down the Wiimote's "B" trigger, point the device at the specified target and release the trigger to signal the selection. The orientation of the Wiimote is captured the instant the "B" trigger is released. The error is the 1-D angular difference between the Wiimote's orientation and the corresponding ideal angle in the plane of interest (*i.e.* when studying the theta plane, the error in the phi plane is ignored). The time to perform a selection is also recorded. The process repeated, with a three second delay between each of the 70 selections per plane. Each participant was compensated $10 for their time and effort.

Results

The errors for each target on the theta and phi plane are presented in Table 1. Errors greater than three standard deviations from the mean are removed from the analysis (<1%). Additionally, targets for left-handed participants are transposed to match targets for the right-handed participants to counter the effect of reaching across the body. Post-Hoc analysis was conducted using the Tukey HSD test. The target has a significant main effect on the selection accuracy for both the theta ($F(6,1252) = 11.50$, $p < 0.001$) and phi ($F(6,1243) = 12.51$, $p < 0.001$) planes. Analysis of selection error for the theta plane targets shows that 0º is significantly ($p<0.005$) more accurate than all other targets. Analysis of selection error for the phi plane regions shows that: -30º and -60º are significantly ($p<0.005$) less accurate than all other

Target	Mean	SD	Target	Mean	SD
-90°	0.13	12.21	-90°	-9.74	6.17
-60°	1.59	10.22	-60°	-15.56	10.30
-30°	-0.86	10.00	-30°	-9.35	11.94
0°	1.92	6.20	0°	-4.50	5.67
30°	4.27	10.57	30°	-0.97	9.31
60°	0.37	12.95	60°	6.09	9.15
90°	2.89	12.92	90°	8.72	5.34
Total	1.47	11.03	Total	-3.51	11.81

Table 1: The errors for the targets in the theta (left) and phi (right) planes. Error is in degrees from the ideal angle.

targets; and, 0º and 30º are significantly ($p<0.01$) more accurate than all other targets.

The selection times for each target on the theta and phi planes are presented in Table 2. The target has a significant main effect on the selection time for the theta ($F(6,1252) = 11.08$, $p < 0.001$) and phi ($F(6,1243) = 10.63$, $p < 0.001$) planes. Analysis of the selection times for the theta plane targets shows that: -60º is significantly ($p<0.05$) slower than all other targets, except for 60º; 0º and 90º are significantly ($p<0.05$) faster than all other targets, except for 30º; and, 30º is significantly ($p<0.05$) faster than -60º and 60º. Analysis of the selection times for the phi plane targets shows that: -90º is significantly ($p<0.05$) faster than all other targets, except for 0º; 0º is significantly ($p<0.05$) faster than all targets, except for -90º and -60º; and, 30º is significantly ($p<0.05$) slower than -90º, 60º and 0º.

These results show that targets to the left and right of participants on the theta plane are significantly more difficult to select than 0º. Participants tend to overshoot when selecting targets on side of the dominant hand. Targets on the side of the non-dominant hand were difficult because participants had to reach across the chest. Along the phi plane, targets below 0º were hardest to select.

EXPERIMENT 2 - EFFICACY OF VIRTUAL SHELVES

Informed by the results of experiment 1, we developed a proof-of-concept prototype of Virtual Shelves for the Nokia N93. We divided the theta plane into 7 regions of 30º since 1SD of target selection errors fit inside ±15º. In the phi plane, 1SD of the errors was within ±15º for targets, except those below the waist. Also the theta width of regions near the poles are too small to select (width converges to 0 from the equator to the poles). As a result, we removed the 2 polar regions and merged everything below the waist into one region; this left 4 phi regions in total. Therefore, the Virtual Shelves prototype has 28 selectable regions, 21 of which we assigned tasks commonly performed on a mobile phone (see Figure 3).

Informal Prototype Evaluation

Using a similar setup as experiment 1, we conducted an informal evaluation of our prototype against the native interface running on a Nokia N93. We replaced the Wiimote with a Nokia N93 running our custom Virtual Shelves software. We added markers to the N93 to track its orientation using a Vicon system. The device's orientation,

Target	Mean	SD	Target	Mean	SD
-90°	3.12	0.92	-90°	2.74	1.13
-60°	3.48	1.08	-60°	3.06	1.01
-30°	3.16	1.21	-30°	3.37	1.06
0°	2.78	0.82	0°	2.88	0.78
30°	2.98	0.78	30°	3.39	1.05
60°	3.29	1.24	60°	3.25	1.05
90°	2.84	0.75	90°	3.26	1.02
Total	3.09	1.01	Total	3.14	1.04

Table 2: Time used selecting each target in the theta (left) and phi (right) planes. Time is in seconds.

calculated using its theta and phi angles, indicated which region is currently selected, and the respective shortcut to display or trigger. For example, if the phone is panned vertically in the calling slice, the N93's screen will update to show speed dial icons to different friends. To minimize the burden on participants' spatial and kinesthetic memory, only three slices were used with 4 shelves in each. The high level tasks assigned to the phi plane for our evaluation are: placing a call to friends; changing the playlist within a music player; and examining the event, day, week, and month views of a calendar. The N93 was configured such that shortcuts to the messaging, phone and calendar applications were available on the home screen.

Procedure

Seven right-handed participants (4 male, 3 female), between the ages of 18 and 35 were recruited to perform the evaluation. Participants were asked to perform two blocks of selection tasks for each technique, selecting 12 tasks, 3 times each in a block; or 36 selections per block, and 72 selections in total. The order the interfaces were presented was fully counterbalanced. The ordering of the selection tasks was random, but consistent between participants. Participants were trained how to use each interface until they felt comfortable performing the tasks. Training consisted of sweeping through the Virtual Shelves to remember the placement of shortcuts and performing each of the tasks using the N93's native interface. Participants chose from 3 categories of tasks (calendar, music, and calling), each with 4 subtasks. The arrangement of the shortcuts in Virtual Shelves is shown in Figure 3. In the native menu, it would take a minimum of 4 clicks to complete the calendar task, 10 clicks for the music task and 2 clicks for the calling tasks.

A visual distracter task was used to access the eyes-free use of Virtual Shelves. The visual distracter is a random number (between 1 and 20) that changed every 4 seconds that the participants were required to repeat verbally. We recorded the number of times participants missed the change, the time participants took to perform a selection, and the number of erroneous selections.

Each selection task for Virtual Shelves started with the participant's arm relaxed against the side of her body. Then the name of the task to select was displayed by the projector. The participant would then press and hold any

Figure 3: The layout of the Virtual Shelves used in the prototype evaluation.

button on the N93, and point the device at the shortcut of the specified task. Visual feedback was provided on the device display. A selection occurred when the button pressed was released. Each selection task for the N93's native interface started with the phone displaying its home screen. A projector displayed the task to perform. The participant would then navigate the phone's interface to perform the task. Each participant was compensated $15 for their time and effort.

Results

The overall task completion time for Virtual Shelves ($M=5.39s$, $SD=2.74$) is faster than the native N93 interface ($M=7.68s$, $SD=5.59$). The number of task selection errors performed is greater using Virtual Shelves ($M=9.1$, $SD=4.6$) than the native N93 ($M=2.1$, $SD=2.0$). Inversely, the number of distracter misses is less using Virtual Shelves ($M=1.4$, $SD=2.6$) than the native N93 ($M=5.4$, $SD=4.7$). Although Virtual Shelves incurs a higher rate of error for first time users, the number of errors decreases by 31.6% after a single block of use (total errors in B1 is 38, and B2 is 26). Additionally, the smaller number of distracter errors when using Virtual Shelves shows that users are able to pay greater attention to their environment when compared to the native N93 interface. Some participants who preferred the N93 interface commented that Virtual Shelves could be awkward to use in some social situations and that some of the regions (specifically the highest shelves) are difficult to select. However, one of these participants additionally commented that, *"If I was in private, I would use [Virtual Shelves] cuz [sic] it's faster."*

CONCLUSION

Virtual Shelves is an interface that is designed to rely solely on kinesthetic and spatial memory. We built a proof-of-concept prototype with 28 selectable regions. We evaluated Virtual Shelves against the native interface for a Nokia N93. The results show that Virtual Shelves is faster than the N93's native interface for common mobile phone tasks. Although Virtual Shelves incurs a higher rate of error for first time users, the number of errors decreases drastically after a single block of use.

Though the hardware platform used to build this proof-of-concept relies on motion capture, various sensing technologies exist (*e.g.,* 6-degree of freedom inertial sensing, magnetometers) that can be used to make Virtual Shelves a portable system. Ascension Flock of Birds (used in [13]) and Nintendo's Wii MotionPlus accessory are great examples of such technologies.

ACKNOWLEDGEMENTS

We would like to thank members of the DGP for their thoughtful comments on this work and Alyssa Rosenzweig for helping to edit this paper.

REFERENCES

1. Biggs, J. and Srinivasan, M. A.: Haptic Interfaces. In *Handbook of Virtual Environments*, Lawrence Earlbaum, Inc (2002), pp. 93-116.
2. Cao, X., Massimi, M., and Balakrishnan, R.: Flashlight jigsaw: an exploratory study of an ad-hoc multi-player game on public displays. In *Proc. CSCW 2008*, ACM (2008), pp. 77-86.
3. Clark, F.J.: How accurately can we perceive positions of our limbs? In *Behavior and Brain Sciences*, Cambridge University Press (1992), pp. 725-726.
4. Cockburn, A., Gutwin, C., and Alexander, J.: Faster document navigation with space-filling thumbnails. In *Proceedings of CHI '*06, ACM Press (2006), pp. 1-10.
5. Cockburn, A., Gutwin, C., and Greenberg, S.: A predictive model of menu performance. In *Proceedings of CHI '07*, ACM Press (2007), pp. 627-636.
6. Czerwinski, M., van Dantzich, M., Robertson, G. and Hoffman, H.: The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3D. In *Proceedings of INTERACT '99*, IOS Press (1999), pp. 163-170.
7. Gray, W. and Fu, W.: Soft constraints in interactive behavior: the case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science 28* (2004), pp. 359-382.
8. Hinckley, K., Pierce, J., Sinclair, M. and Horvitz, E.: Sensing techniques for mobile interaction. In *Proc. UIST '00*, ACM (2000), pp. 91-100.
9. Rekimoto, J.: Tilting operations for small screen interfaces. In *Proceedings of UIST '96*, ACM Press (1996), pp. 167-168.
10. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D., and van Dantzich, M.: Data Mountain: Using Spatial Memory for Document Management. In *Proceedings of UIST '98*, ACM Press (1998), pp. 153-162.
11. Tan, D. S., Pausch, R., Stefanucci, J. K., and Proffitt, D. R.: Kinesthetic cues aid spatial memory. In *CHI '02 Extended Abstracts on Human Factors in Computing System*, ACM (2002), pp. 806 -807.
12. Teitelbaum, R. C., and Granda, R.E.: The effects of positional constancy on searching menus for information. In *Proceedings of CHI '83*, ACM (1983), pp. 150-153.
13. Yee, K.-P.: Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers. In *Proceedings of CHI '03*, ACM Press (2003), pp. 1-8.