Leveraging Proprioception to Make Mobile Phones More Accessible to Users with Visual Impairments

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

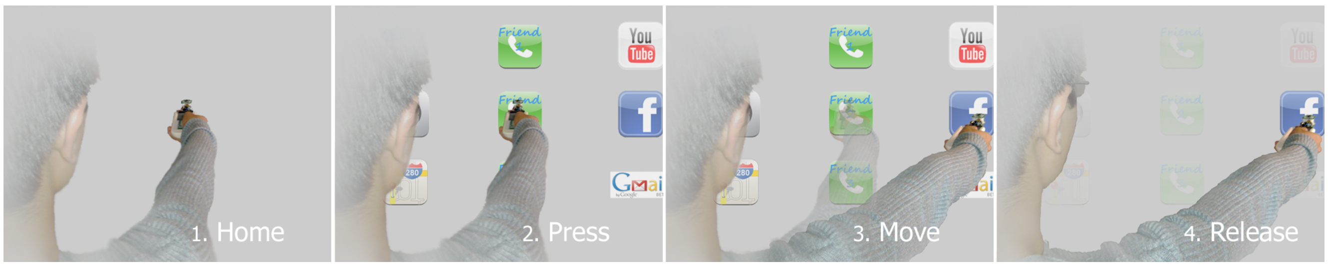

Figure 1. Using a mobile phone with Virtual Shelves to select an application shortcut: 1) Start with the mobile phone in the home position; 2) press and hold on a button or the touch screen to initialize Virtual Shelves; 3) move the device to the region with the desired shortcut; and 4) release the button or screen to select the shortcut.

ABSTRACT
Accessing the advanced functions of a mobile phone is not a trivial task for users with visual impairments. They rely on screen readers and voice commands to discover and execute functions. In mobile situations, however, screen readers are not ideal because users may depend on their hearing for safety, and voice commands are difficult for a system to recognize in noisy environments. In this paper, we extend Virtual Shelves—an interaction technique that leverages proprioception to access application shortcuts—for visually impaired users. We measured the directional accuracy of visually impaired participants and found that they were less accurate than people with vision. We then built a functional prototype that uses an accelerometer and a gyroscope to sense its position and orientation. Finally, we evaluated the interaction and prototype by allowing participants to customize the placement of seven shortcuts within 15 regions. Participants were able to access shortcuts in their personal layout with 88.3% accuracy in an average of 1.74 seconds.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *prototype, input devices and strategies.*

General Terms
Performance, Design, Experimentation, Human Factors.

Keywords
Accessibility, mobile devices, proprioception, visual impairments.

1. INTRODUCTION
Mobile phone interfaces are not built with accessibility for the visually impaired as a primary requirement. As a result, accessibility features such as screen readers and voice commands must be layered on top of the factory interface. Despite their benefits, a screen reader can be time consuming and potentially dangerous in mobile situations where a visually impaired person relies on her hearing for safety, and voice commands can be difficult for a system to recognize in a noisy environment. In addition, the mobile phone market is moving progressively towards a touch screen form factor, which supports an inherently visual experience. Touch screens allows the user to directly manipulate and control on screen content, and to view the outcome of her interaction in the same space. However, the vision required for these interactions is a considerable barrier for users with visual impairments.

In this paper, we extend our previous work with Virtual Shelves [15] to improve the accessibility of mobile devices for visually impaired users. Virtual Shelves is an interaction technique that leverages proprioception[1] to support eyes-free launching of application shortcuts assigned to spatial regions centered around the user's body (see Figure 1). Virtual Shelves utilizes an orientation-aware mobile device to determine the theta and phi angles of the user's arm with respect to her body (see Figure 2), and uses these angles to index into a list of shortcuts. The end effect is that the user has a virtual shelf of shortcuts, which can be accessed quickly in constant spatial positions around her.

The number of shortcuts we supported in the initial Virtual Shelves implementation was defined to account for the selection error of persons with normal or corrected vision, not the visually

[1] *Proprioception* is the sense of position and orientation of one's body parts with respect to each other. *Kinesthesia*, which is often used interchangeably with proprioception, is the sense of the body's movements and motions.

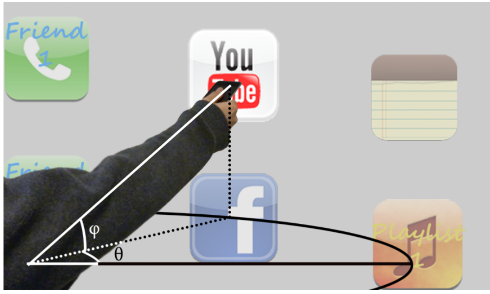

Figure 2. Virtual Shelves uses the theta (θ) and phi (φ) angles to index into a shelf of shortcuts.

impaired. To improve the accessibility of Virtual Shelves, we first measured the *directional accuracy* (how accurate can one point at angles relative to one's body) in the theta and phi plane of nine visually impaired participants. We observed a larger selection error for the visually impaired participants than participants with normal or corrected vision. To account for the difference in selection error, we redefined the Virtual Shelves regions from a 7x4 design to a 5x3 design (5 columns by 3 rows totaling 15 shelves) and implemented a functional prototype of Virtual Shelves using an accelerometer and a gyroscope. Even though at the time of prototyping there were no smartphones in the market bundled with those two sensors, accelerometers and gyroscopes are becoming standard and can now be found in the Apple iPhone 4 and the Samsung Galaxy S. Our prototype demonstrates that the Virtual Shelf experience, previously developed using expensive and immobile tracking technologies, can now be used in mobile contexts. Finally, we recruited 13 visually impaired participants to evaluate the technique and the prototype. In the evaluation, participants customized a personal layout using seven shortcuts and were able to access the shortcuts with 88.3% accuracy in an average of 1.74 seconds. A participant also raised some concerns about the social and environmental challenges with using such a technique (*e.g.,* hitting a stranger on a crowded bus).

2. RELATED WORK

The goal of this work is to make the features and applications on a mobile phone more accessible for users with visual impairments. While much research has been done in improving the accessibility of mobile devices for users with visual impairments, no work has explored the use of proprioception and orientation-awareness. On the other hand, orientation-aware devices and orientation-based interaction techniques have been well studied for the sighted, but not for the visually impaired. In this section, we review the literature for these two disjoint fields.

2.1 Mobile Device Accessibility

Auditory feedback is commonly employed to provide an awareness of interface widgets and to assist in navigating them. The EarPod [30] and ADVICE [2] systems used auditory cues to assist users in navigating a menu with a physical wheel-like input. BlindSight [16] replaced the mobile phone's visual menu with an auditory menu during a call to enable eyes-free navigation using a keypad placed on the back of the phone. However, audio feedback is not always ideal in a mobile setting. Users with visual impairments rely on their hearing heavily when travelling (*e.g.,* to sense an oncoming car). They must therefore stop to focus on the audio feedback and finish the interaction before continuing, interrupting their routes and preventing them from experiencing the benefits of multi-tasking. Virtual Shelves [15] relies on proprioception and spatial memory for interaction and does not risk the users' auditory attention.

Most mobile phones allow the user to issue basic voice commands. Voice commands were also used in systems such as VoiceNote [24], a wristwatch-type PHS Telephone by NTT [25], and Nomadic radio [23]. Voice recognition by mobile devices is prone to errors and might need repeated issues of the command [23]. As a result, voice commands cannot be the only method of input for a mobile device, physical input through a traditional interface is required. This can cause significant inconvenience to the user as she has to judge which interface is appropriate for her surroundings. Our technique uses sensors that are not sensitive to environmental interferences.

Touch screen based systems developed by Kane *et al.* [14], O'Neill *et al.* [18], Pirhonen *et al.* [20], and Yfantidis and Evreinov [29] used directional gestures performed on the touch screen to indicate operations such as menu navigation and text entry. Gesture based systems have arbitrary gesture mappings that users are not allowed to change and must recall. Many interactions also require multi-touch gestures that are likely to need both hands to perform (*i.e.*, one to hold the phone and the other to perform the gestures). In mobile situations, though, visually impaired users almost always have one hand occupied with a cane or a guide dog. Virtual Shelves overcomes both of these challenges because it is a one-handed interaction technique that tries to minimize the recall time and accuracy by leveraging personal customizations, spatial memory and proprioception.

2.2 Orientation Aware Devices

Rekimoto [21] and Hinckley *et al.* [12] first described the addition of context sensors (*e.g.,* tilt and proximity) to augment common interactions by utilizing the placement and orientation of the host device. As a result, many interaction techniques have been developed to take advantage of these sensors. Oakley *et al.* [19] created a marking menu based on the device's motion using an accelerometer. Another common sensor is magnetometers. Magnetometers were used to track the position and orientation of Private Eye [8] to allow the wearers of the head-mounted display to look around a virtual desktop. Magnetometers can also track objects relative to the device. For example, Abracadabra [11] used them to track a finger (wearing a magnet). Brewster *et al.* [5] augmented a pair of headphones using magnetometers as well as accelerometers and gyroscopes for head based interactions. Likewise, XWand [27] also used these sensors for interactions with a wand device. Blasko *et al.* [4] based interactions on a sensor that measures the length and angle of a retractable string. Alternatively, the image-based sensing of a camera was used by Hansen *et al.* [10] to track the device's position relative to the user's face. However, interaction techniques that use such devices have been mostly developed for sighted users. Our work is different as it is focused on users with visual impairments.

Arranging content spatially with mobile devices has been demonstrated before in existing literature. Users can arrange files into piles that are placed virtually around the device in Hsieh *et al.*'s system [13]. Orientation aware devices can act as peephole [28] or flashlight [6] displays to allow a user to organize and access content in a virtual space around her body. However, these systems fundamentally require vision for operation. Virtual Shelves extends these techniques while being accessible to both users with visual impairments and users with sight.

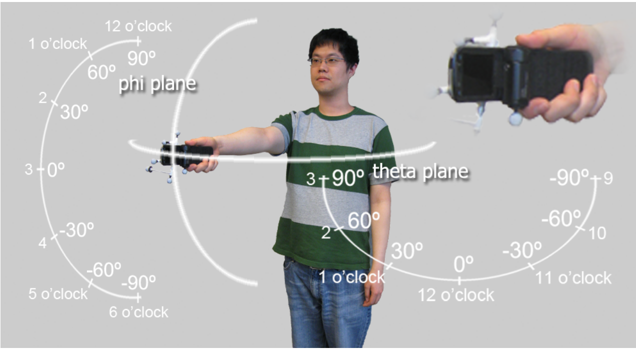

Figure 3. Our first study used motion capture to capture participants pointing a Nokia N93 at 14 targets.

3. STUDY 1 – DIRECTIONAL ACCURACY

Previous work has explored the spatial accuracy of sighted people [3, 7, 15]. In this study, we replicated the directional accuracy experiment conducted in our previous work [15] to focus explicitly on understanding the directional accuracy of the visually impaired.

3.1 Instruments and Setup

We used six Vicon M2 cameras to track and capture the 3D orientation of a Nokia N93 mobile phone. We used custom logging software to continuously record the orientation and position of the mobile phone provided by the Vicon system, and all key presses on the mobile phone.

Following the initial Virtual Shelf study design, we placed 14 virtual targets in the space in front on the participants; seven targets on the theta plane (*i.e.,* parallel to the ground) and seven targets on the phi pane (*i.e.,* perpendicular to the ground). The targets were placed at 30º intervals on each plane. Rather than asking participants to select targets as identified by degrees, we used the clock metaphor to label targets (see Figure 3).

We discovered in piloting the experiment that it was difficult for the visually impaired participants to point the mobile phone in a straight line relative to their arm. Because they could not see the exact shape of the phone, they would unintentionally hold it with a slight angle. This was problematic because the Vicon system captured the orientation of the phone and not their arm, resulting in selection data that did not match the participants' intended selection. To counter this problem, we reoriented the mobile phone in the participant's hand in a manner that aligned the thumb with the side of the phone and allowed use of the index, middle or ring finger to press the selection button (see the zoomed-in hand in Figure 3). In this manner, users ceased to orient the device and instead pointed their entire arm towards the targets. Thus we could truly measure the participants' proprioception of their arm.

3.2 Participants

Nine visually impaired (8 blind and 1 low vision) participants (3 male and 6 female) were recruited to perform the study. In this context, we define *blind* as no vision and *low vision* as having less than 10% vision. Participants with no vision were recruited to measure true eyes-free directional accuracy (the participant with low vision was politely asked to close her eyes). The age of participants ranged from 20's (1), 30's (3), 40's (4) and 50's (1). Five participants held the phone in their right hand, four in the left. The hand they chose to hold the phone was not necessarily their dominant hand; some participants reported normally holding their personal phone in their non-dominant hand so they could use their dominant hand to perform key presses. Each participant was compensated $20 for their time and effort.

3.3 Design and Procedure

We asked participants to select each of the 14 targets by pointing the mobile phone (through the extension of their arm) at the angle defined by the target. Targets on the theta and phi planes were divided into two separate conditions. Each participant completed both the theta and phi conditions, but the ordering of the conditions was counterbalanced across the participants.

Each condition included a training phase that was conducted before the testing phase. The purpose of the training phase was to help the participant develop a spatial awareness of the target locations. In the training phase, vibrotactile feedback was provided (through the mobile phone) whenever the phone was within ±4º of the true angle of each target. The training phase continued until the participant felt all seven targets at least once and the participant was comfortable with the clock metaphor and the mobile phone. In the testing phase, participants performed 10 selection tasks for each target: 70 selections for each condition, 140 selection tasks across the two conditions. The presentation order of the selection tasks was randomly distributed, but consistent between participants. PC speakers were used to output the selection tasks using the clock metaphor. A selection task started with the participant's arm relaxed against her side. The participant would then hold down any key on the keypad, point her thumb (which is aligned to the side of the phone) at the target, and release the key to select the target. After the selection, she would return her arm to her side and wait for the next task. After a three second pause, the next task would sound from the PC. No feedback of any type was provided during the testing phase. A short break was enforced between each condition. The entire study took around 45 minutes.

During the testing phase we collected measures for selection error and time. The selection error was identified by capturing the orientation of the mobile phone when the device's button was released and calculating the difference in the angular error between the mobile device and the true angle of the target. Only the error for the respective plane is considered in each condition. For example, the error in the phi plane is ignored when evaluating the theta plane. The selection time of a task was identified as the time between when the target was outputted by the speakers and when the button was released on the mobile phone.

3.4 Results

The selection error and selection time for the targets are presented

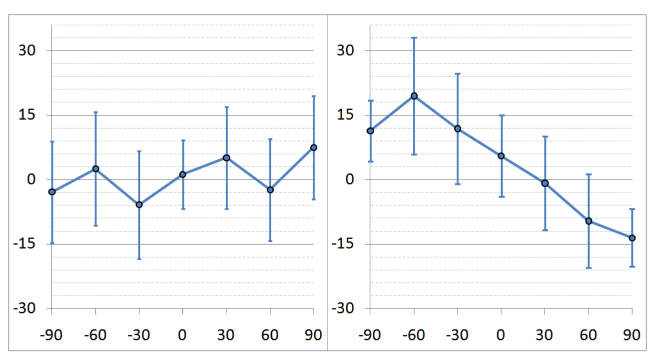

Figure 4. Plots of the average selection errors in both the theta (left) and phi (right) planes. Units are in degrees.

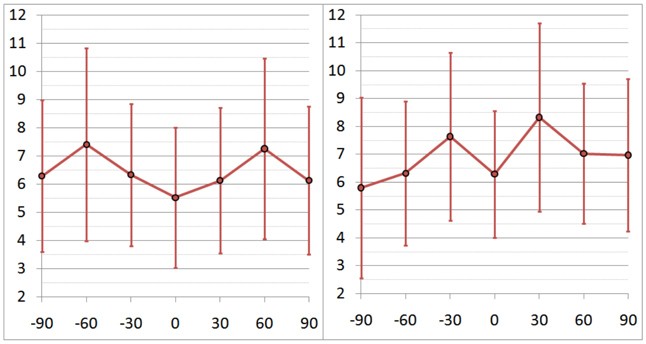

Figure 5. Plots of the average selection times (in seconds) in both the theta (left) and phi (right) planes.

in Figure 4 and 5 respectively. Targets for left-handed participants are transposed in the theta plane to counter the effect of reaching across the body and normalize the data across all participants. Main effect analysis for selection error and time was performed using the GLM. Post-hoc pairwise comparison was conducted using the Tukey HSD test.

We observed a significant main effect of target placement for selection error (theta: $F_{6,623}=14.44$, $p<0.001$; phi: $F_{6,625}=117.93$, $p<0.001$). The mean selection error for the targets is presented in Figure 4. In the theta plane, the selection error for the 0° and -60° targets significantly smaller than the -30° and 90° targets (all at $p<0.05$). Participants tended to overshoot when selecting the 90° target and undershoot the -30° and -90° targets. In the phi plane, the selection error across all targets was significantly different (all at $p<0.005$); except between -90° and -30°, and 60° and 90°.

Similar to the selection error, we observed a significant main effect of target placement for selection time (theta: $F_{6,623}=5.07$, $p<0.001$; phi: $F_{6,625}=8.34$, $p<0.001$). The mean selection times are presented in Figure 5. In the theta plane, participants selected the 0° target faster than the -60° and 60° targets (both at $p<0.001$). In the phi plane, the selection time for the 30° target was significantly slower than all targets other than the -30° target ($p<0.05$) and the -30° target was slower than the 0° and 60° targets (both at $p<0.05$).

We compared the results for this study with that of our initial Virtual Self work [15]. The visually impaired participants in this study had a significantly greater selection error and selection time than the sighted participants. Specifically, the sighted participants had a smaller selection error for the 90° ($p<0.01$), 30° ($p<0.001$), and -30° ($p<0.05$) targets in the theta plane and the 90° ($p<0.001$), 60° ($p<0.005$), and -60° ($p<0.01$) targets in the phi plane. In addition, the selection time for the sighted participants was significantly faster across all 14 targets ($p<0.001$). Proprioception has been show to degrade with age [1]. We are unsure whether the difference seen in our results was because of sightedness or age (median age group for blind participants was 40-50 and sighted participants was 26-35). To account for the higher selection error we extended the original 7x4 design with 30° region size along the theta and phi planes, resulting in a new 5x3 Virtual Shelf design. In the theta plane the 5 regions are 45º wide, centered at -90º, -45º, 0º, 45º, and 90º. In the phi plane, the second region is 40º wide and centered at 0º. The first and third regions encompass any selection that is less than or greater than -20º and 20º respectively.

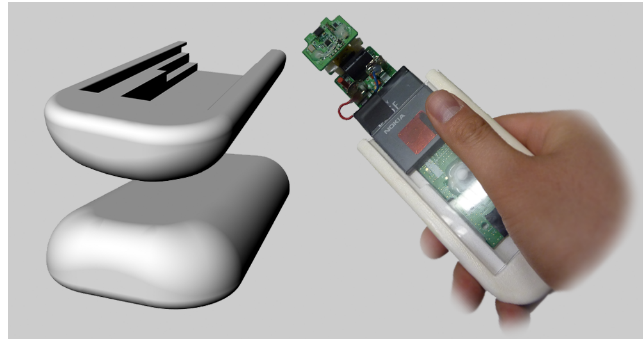

Figure 6. Left: 3D renders of the case. Right: The prototype.

4. PROTOTYPE

The initial Virtual Shelves work [15] utilized a Vicon motion capture system to determine the orientation and position of the mobile device. Our intent with Virtual Shelves is to implement the technique within the mobile device so that it can be used in any context. As such, we developed a method using commodity sensors that allow us to determine the orientation of a mobile device.

4.1 Implementation Details

Not all the sensors described in the related works section fit our purposes. For example, camera based tracking has a very limited area of operation and will not function in the entire front hemisphere of the body. Magnetometers are extremely sensitive to electronics and are not reliable for mobile use. Thus, we developed a proof-of-concept prototype using a 3-axis accelerometer and a 3-axis gyroscope—technologies that now exist in commercial off-the-shelf smartphones.

Rather than designing a custom circuit board, we utilized the 3-axis accelerometer and the 3-axis gyroscope built into the Nintendo Wiimote with the Wii MotionPlus[2]. We removed the Wiimote circuit board and encased it within a custom housing (115 X 60 X 20mm) designed to mimic the form factor of Apple's iPhone (see Figure 6). The case was coated with a layer of polish to give a glossy, polished finish like that of current mobile phones. To imitate a touch screen, a piece of Plexiglas was fitted slightly above the 'B' button of the Wiimote. Pressing down anywhere on the Plexiglas screen triggered the Wiimote's 'B' button. The prototype weighed 150g (for comparison the iPhone weighs 135g). It communicated with a PC using Bluetooth built into the Wiimote circuit board. One participant commented that "*oh it feels like one of those iPhones*" when she first held the prototype.

When used to measure the gravity vector, an accelerometer can output the phi angle accurately. However, because it only measures acceleration, it cannot give an absolute theta angle. Three-axis acceleration data can be mathematically integrated twice to estimate the horizontal distance travelled by the phone. But the theta angle of the phone cannot be derived from this distance unless we assume that the arm's length is constant. While this assumption seems valid at first, it does not hold in our case. We made the initial choice of using angles to index into shelves so that an expert user would not need to swing her whole arm to trigger a shelf; she can just bend her wrist. We hope that

[2] http://www.nintendo.com/wii/what/accessories

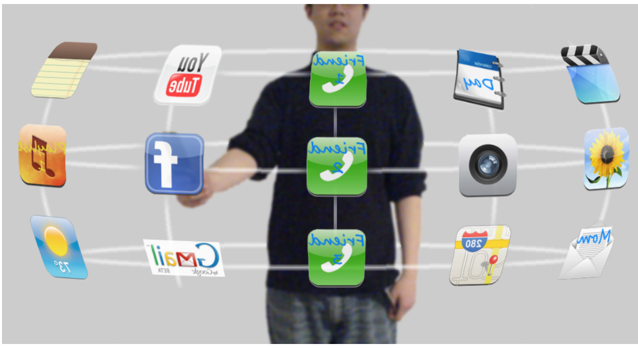

Figure 7. A sample customization of Virtual Shelves.

less and less movement would be needed as a user becomes more experienced with the technique. An accelerometer would not allow room for improvement. To measure the theta angle, we used a gyroscope. Within the speeds of typical human arm swings, a gyroscope can measure angular rate of change accurately. However, it only measures rate of change; thus, it can only give the angles relative to a starting position. To overcome this limitation, we introduced the concept of a home position; the user always has to start by pointing in front of her.

Furthermore, because the gyroscope provides only a rate of change, the signal has to be integrated to retrieve the actual angles. Integration makes the gyroscope an inherently unstable system. The output will drift towards ±infinity as time increases. The worst case drift from the type of gyroscopes we are using is about 1º/sec [9]. Results from study 1 showed that the typical selection time was approximately six seconds. Since the shelf regions are at least 40º wide, there was no need to implement any advance drift compensation algorithms (*e.g.*, Kalman [9, 26] and extended Kalman filters [22]). Instead, we implemented basic noise filters and used a 4^{th} order Runge-Kutta integrator to determine the theta angle from the gyroscope output.

4.2 Interaction Technique

Informed by the results from study 1, we divided the theta plane into five regions and the phi plane into three. The theta regions were each 45º wide, centered at -90º, -45º, 0º, 45º, and 90º. The phi region at 0º had a width of 40º. The top phi region centered at 45º but anything beyond 20º counted as the top region, and likewise with the bottom. Five columns and three rows gave the prototype 15 shelves in total. Each shelf could be associated with a customizable shortcut to mobile phone tasks. Figure 7 shows an example layout with 15 common mobile phone shortcuts.

To trigger a shortcut, the user must follow the sequence shown in Figure 1. The user starts by pointing the prototype straight in front of her (*i.e.,* the home position). She then presses down on the screen and holds without releasing. A short vibration will signal to confirm the down event. As she scrolls around her body, names of the current shortcuts will be stated aloud via the accompanying PC's speakers. When she arrives at the desired shortcut, she releases the screen to launch the shortcut. We evaluated this technique in study 2.

5. STUDY 2 – USER EVALUATION

In this second study, we wanted to validate the usability of the revised Virtual Shelves technique for the visually impaired. A comparison against an existing mobile phone interface was not done. Virtual Shelves is not meant to replace the existing interface on a mobile device, but rather to alleviate some of the burden incurred when navigating a visual interface for common tasks. Functions that are not set as shortcuts in the shelves would still need to be accessed using the underlying interface.

5.1 Participants

Thirteen visually impaired (7 blind and 6 low vision) participants (5 male and 8 female) were recruited to perform the study. The age of participants ranged from 30's (3), 40's (5), 50's (3) to 60's (2). Seven participants had previously participated in study 1. Eleven held the phone in their right hand, two in the left. The hand they chose to hold the phone was not related to handedness. Twelve participants used mobile phones regularly and six used screen readers to operate their mobile phones. Participants were asked to hold the prototype with their thumb along the edge of the screen (see Figure 6). In this manner, users ceased to orient the device but instead pointed their arm towards the different directions. Each participant was compensated $30.

5.2 Design and Procedure

We did not continue to use the clock metaphor to describe the position of the virtual shelves because the five theta regions and three phi regions do not conform to the clock metaphor. For the theta plane, we used a compass metaphor, describing the virtual shelf positions as *west* (-90º), *northwest* (-45º), *north* (0º), *northeast* (45º), and *east* (90º). For the phi plane, we described the virtual shelf positions as *top* (45º), *middle* (0º) and *bottom* (45º). Each shelf was named using both conventions (see Figure 8). For example, *northeast-bottom* referred to the position of 45º theta and -45º phi.

This second study included two independent phases. In phase 1 (*all shelves*), we evaluated the selection time and selection error of all 15 possible virtual shelves. In study 1 we examined the selection error only along the theta and phi axes, thus we designed phase 1 of study 2 to confirm that the new regions we defined were valid for visually impaired users. In phase 2 (*custom shelves*), we asked participants to choose the placement of seven shortcuts within any of the 15 Virtual Shelf regions. The intent is to learn which shelves are perceived to be the most natural and the strategies people use to choose the placement of shortcuts.

5.2.1 Phase 1: All Shelves

This phase included both training and testing, with the training conducted before the testing. The purpose of the training is to help the participant develop a spatial awareness of the locations of the 15 shelves. During training, vibrotactile feedback was given whenever the prototype device was within a 10º radius of the ideal angles. The angular distances are now in 2D and are calculated with the following equation:

$$d_{angular} = \sqrt{(\theta_1 - \theta_2)^2 + (\varphi_1 - \varphi_2)^2}$$

The training phase continued until the participant felt all 15 shelves at least once and was comfortable with the naming conventions. In the testing phase, participants performed two blocks of selections tasks, with 5 selection tasks for each target per block: 10 selections for each target, 150 selections in total. The ordering of the tasks was randomly distributed and different between blocks, but consistent across all participants. For each selection, the participant started with her arm relaxed. The PC speakers would output the name of the shelf to select. She pointed the device straight in front at the home position and then held down the screen. She released the screen as soon as she finished moving the device to the specified shelf. Then she relaxed her arm and waited for the next task. No feedback was provided and there was a 2.5 second pause between each task. Throughout this phase, the angular differences between the selected and the ideal angles were recorded, as well as the selection time. A longer break was enforced between each block so participants could rest their arms.

5.2.2 Phase 2: Custom Shelves

Upon completion of phase 1, the participants had each performed over 150 selections, allowing them to develop an initial impression of how well they could select each of the 15 regions. In the second phase, we asked participants to customize a personal layout using seven common mobile phone tasks: three calling shortcuts (the participant's three most called contacts), two location-based tasks (directions to home and the closest public transit stop), check email and check current weather. We chose seven tasks because short-term memory supports 7±2 items [17]. Audio feedback was provided to indicate the currently selected shelf. One of the authors recorded the audio for the non-calling tasks (*i.e.,* "give directions to home", "find closest public transit", "check email", and "current weather"). To add a second layer of personalization, we recorded the participant saying "call <name>" and used this recording as the audio feedback for the calling tasks. All seven tasks were given at once verbally to minimize the effect of presentation order on shelf assignment. Participants were given as much time as they needed to choose the layout.

Phase 2 included both training and testing, with the training conducted before the testing. Training involved selecting each of the seven shortcuts at least once. The customized layout could be changed as many times as needed during training. The final layout for each participant and their placement strategies were noted. In the testing phase, participants performed three blocks of selections tasks, with three selection tasks for each target per block (nine selections for each target, 63 selections in total). The ordering was randomly distributed and different for each block, but was again kept constant across participants. The PC speakers would output the target task and the participants would then select the task using the sequence described in the Prototype Design section of this paper. There was a 2.5 second pause between each task. A break was enforced between each block for participants to rest their arms. The whole experiment including phase 1 and 2 took around 1 hour and 15 minutes to complete.

5.3 Results

5.3.1 Selection Time and Accuracy of All Shelves

The selection error and selection time for each theta and phi plane targets are presented in Table 1. Targets for left-handed participants are transposed in the theta plane to counter the effect of reaching across the body and normalize the data across all participants. Main effect analysis for selection error and selection time was performed using the GLM. Post-hoc pairwise comparison was conducted using the Tukey HSD test.

Table 1. The selection time (in seconds) for each shelf. The overall mean and SD is 2.01 ± 0.98 seconds.

	-90°	*-45°*	*0°*	*45°*	*90°*
45°	2.20 ±0.80	2.23 ±1.19	1.53 ±0.61	2.26 ±1.23	2.24 ±0.93
0°	1.99 ±0.86	2.15 ±0.96	1.11 ±0.67	2.03 ±0.99	2.02 ±0.86
-45°	2.14 ±0.97	2.08 ±0.89	1.64 ±0.63	2.34 ±1.16	2.13 ±0.97

We observed a significant main effect of target placement for selection error ($F_{14,1933}$=20.57, p<0.001). As presented in Figure 8, the selection error for the top regions is greater than the *middle* and *bottom* regions. We believe this is a result of the participants significantly underestimating the 45° plane. The *middle* regions were more accurate than the *top* and *bottom* regions, both at p<0.001. The middle (*north*) column regions were more accurate than all other columns, all at p<0.001. Overall, 81.8% of the selections were performed correctly, 13.6% resulted in the selection of a neighboring region, and 4.6% resulted in a selection that was two or more regions away or exceeded the -90° to 90° boundary of the front hemisphere.

In addition, we observed a significant main effect of selection time ($F_{14,\ 1933}$=16.52, p<0.001). The selection time for the *north-middle* (*i.e.,* home position) region was significantly faster than all other regions, all at p<0.05. The *north-top* region was faster than all other regions (p<0.01), except *north-middle*, and the *north-bottom* region was faster than all other regions (p<0.05), except *west-middle*, *northeast-middle* and *east-middle*.

5.3.2 Shelf Customizations

We do not report the difference in selection error and time because of the variability in the regions that were selected; no two participants chose the same seven regions. In phase 2, we observed a 6.5% improvement in the number of correct selection from phase 1: 88.3% of the selections were correct, 9.8% selected a neighboring region, and 1.9% resulted in a selection that was two or more regions away or extended beyond the -112.5° to 112.5° boundary of the front hemisphere. The participants likely were able to deduce which regions they had the most problems with in phase 1, and

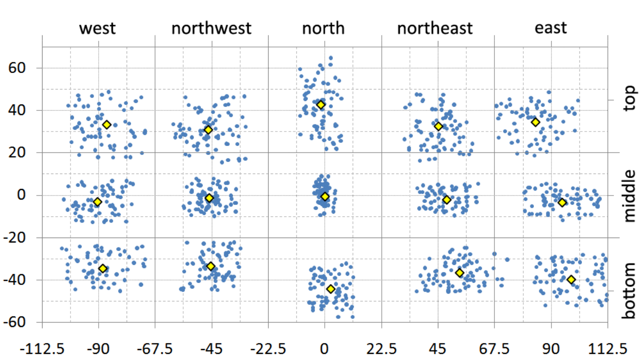

Figure 8. Selections made within 1SD of the theta and phi means of each shelf. Units are in degrees.

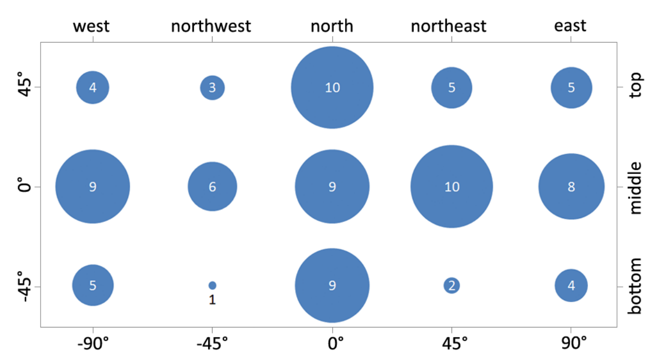

Figure 9. The total number of times each shelf was picked for customizations in Phase 2.

refrained from picking those in phase 2. Furthermore, we observed a block-by-block improvement in the percentage of correct selections, improving from 84.6% (block 1), to 88.3% (block 2) and 90.1% (block 3). The mean selection error from the ideal angle was 13.7° ± 9.0° and the mean selection time was 1.74 ± 0.92 seconds; no significant difference was observed between blocks.

Regions along the theta and phi axes were chosen most frequently (see Figure 9). The *northwest* and *northeast-bottom* regions were least used. The semi-structured interviews conducted at the end of the study revealed that participants had devised many interesting placement strategies. One participant placed **e**mail in the **e**ast and **w**eather in the **w**est. Two participants customized the location shortcuts based on the actual directions to their home or the subway station. Alphabetical ordering (A-Z) from top to bottom was used often when placing contacts. While these placement strategies are not applicable in all situations (*e.g.*, home will not always be to the west), with experience we believe the initial strategy used to aid in recall will become irrelevant when the users' spatial memory improves.

Participant responses indicated that directions home might not be useful and should be removed to free a region. Participants commented that they try very hard to not place themselves in a position where they would need assistance to get home. The most common request was to place more than three contacts and to include calling for taxis. Requests also included a shortcut to the camera, save a voice memo, read a new SMS messages, find points of interest (*e.g.,* banks, groceries stores, coffee shops, and clinics), the current time, and current battery level. All participants stated they would have no problem filling all 15 shelves, and felt that none of the shelves were overly uncomfortable to reach. Only one participant commented that she liked her own cell phone interface (iPhone) more than Virtual Shelves. However, all of the participants that used a mobile phone agreed that Virtual Shelves was faster than their existing mobile phone interface. They were very supportive of the technique, commenting that *"it's great"* and *"easy to memorize"*. One participant commented that Virtual Shelves is *"much much much better"* than her current mobile phone's interface.

Finally, one participant commented on an inherent problem with our technique. She was afraid that she would hit someone when swinging her arm in public, especially when triggering the *west* and *east* shelves. However, we believe with experience the user will learn to trigger shortcuts from rotating only the wrist and not the whole arm. Two participants purposely removed a shortcut from *north-middle* after testing the prototype during training. The *north-middle* shelf was the home position, so the audio associated with its shortcut was played at the beginning of every selection, which some found to be annoying. Future prototypes should add an option to have no audio feedback until the device is moved away from *north-middle*.

6. DISSCUSSION & FUTURE WORK

This work demonstrates that for mobile devices with both accelerometers and gyroscopes, we will be able to use the Virtual Shelves technique to make the devices more accessible to visually impaired users. We intend to develop and release this technique on the first platform with these sensors built in. As a part of turning this technique into a real product, we must also address some remaining concerns which we discuss in this section.

6.1 Initial Setup

We did not discuss how a user would customize an initial layout of Virtual Shelves. Customization must be accessible so visually impaired users can perform the task on their own. One way is to use the hold gesture. When the user holds down on an application icon for an extended period of time at the home position, she is invited to point the phone at a shelf to store a shortcut to that application.

6.2 Multiple Levels and Layouts

Many participants commented that 15 shelves is too limiting; they were easily able to name more than 15 shortcuts they use regularly. We cannot simply increase the number of shelves because of the users' directional accuracy. Thus the best approach is to implement multi-level and multi-layout setups. For example at the top level we can have one shortcut to all contacts. Once inside contacts, there could be a second level with 14 most called friends and family (one region saved to navigate back to the top level). To take it even further, the second level could instead have shelves for different letters in the alphabet. Selecting a specific letter will expand the third level with names starting with that letter. Tilt interactions introduced by Rekimoto [21] and Oakley *et al.* [19] could also be used to access different depths. For example, once a region is selected, the user rotates the phone clockwise or anti-clockwise to address different items in the region.

Another easy way to expand the number of shelves is to have multiple layouts. A left or right flick gesture could flip between different layouts of shelves, similar the finger flicks to switch between pages on an iPhone or an Android phone. A user could setup a work and home layout and switch between the two with simple flicks of the phone. She could also name each layout so audio feedback is given whenever a switch occurs.

The total number of shelves supported by the technique is infinite. The limiting factor is the mental capability of the user. After all, the goal of Virtual Shelves is not to make the entire feature set of the host device accessible, but only the most used functions.

6.3 Cancelling

A method to cancel during a selection task was not implemented in the prototype. In the current design, as soon as the user releases the screen the shortcut is launched. There is no way to not perform a selection once the screen is pressed down. To signal a cancel a user could shake the device or point at the space behind her. There are no shelves behind a user so no shortcut will be launched when the screen is released.

7. CONCLUSION

We presented an adaption of the Virtual Shelves interaction technique for the visually impaired. We developed a mobile prototype using an accelerometer and a gyroscope and evaluated the prototype. In the evaluation, participants customized a personal layout using seven shortcuts in the 15 shelves. They were able to correctly launch shortcuts 88.3% of the time in an average of 1.74 seconds. This work demonstrates that proprioception can indeed be used to improve the accessibility of a mobile phone for users with visually impairments.

8. ACKNOWLEDGMENTS

We would like to thank Gillian Hayes, Alyssa Rosenzweig and the reviewers for their thoughtful comments and edits on this work. We would also like to thank all the visually impaired

participants who took the time and the effort to travel to our lab for the studies.

9. REFERENCES

[1] Adamo D.E., Alexander N.B., Brown S.H. 2009. The influence of age and physical activity on upper limb proprioceptive ability. *Journal of Aging and Physical Activity*. 17, 3 (July 2009), 272-293.

[2] Amar, R., Dow, S., Gordon, R., Hamid, M. R., and Sellers, C. 2003. Mobile ADVICE: an accessible device for visually impaired capability enhancement. In *Proc of CHI '03 Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, 918-919.

[3] Biggs, J. and Srinivasan, M.A. 2002. Haptic Interfaces. In Stanney, K (ed.), *Handbook of Virtual Environments Design, Implementation, and Applications*. Lawrence Erlbaum Associates.

[4] Blasko, G., Narayanaswami, C., and Feiner, S. 2006. Prototyping retractable string-based interaction techniques for dual-display mobile devices. In *Proc of CHI'06*. ACM Press, New York, NY, 369-372.

[5] Brewster, S., Lumsden, J., Bell, M., Hall, M., and Tasker, S. 2003. Multimodal 'eyes-free' interaction techniques for wearable devices. In *Proc of CHI'03*. ACM, New York, NY, 473-480.

[6] Cao, X., Massimi, M., and Balakrishnan, R. 2008. Flashlight jigsaw: an exploratory study of an ad-hoc multi-player game on public displays. In *Proc of CSCW'08*. ACM Press, New York, NY, 77-86.

[7] Clark, F.J. 1992. How accurately can we perceive positions of our limbs? *Behavior and Brain Sciences*, 15, 4 (December 1992), 725-726.

[8] Feiner, S. and Shamash, A. 1991. Hybrid user interfaces: Breeding virtually bigger interfaces for physically smaller computers. In *Proc of UIST'91*. ACM, New York, NY, 9-17.

[9] Foxlin, E. 1996. Inertial Head-Tracker Sensor Fusion by a Complimentary Separate-Bias Kalman Filter. In *Proc of VRAIS 96*. VRAIS. IEEE, Washington, DC, 185.

[10] Hansen, T.R., Eriksson, E., and Lykke-Olesen, A. 2006. Use your head: exploring face tracking for mobile interaction. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. ACM Press, New York, NY, 845-850.

[11] Harrison, C. and Hudson, S. E. 2009. Abracadabra: Wireless, High-Precision, and Unpowered Finger Input for Very Small Mobile Devices. In *Proc of UIST '09*. ACM, New York, NY, 121-124.

[12] Hinckley, K., Pierce, J., Sinclair, M., and Horvitz, E. 2000. Sensing Techniques for Mobile Interaction. In *Proc of UIST '00*. ACM Press, New York, NY, 91-100.

[13] Hsieh, T., Wang, Q.Y., and Paepcke A. 2009. Piles across space: Breaking the real-estate barrier on small-display devices. *International Journal of Human-Computer Studies*. 67, 4 (April 2009), 349-365.

[14] Kane, S. K., Bigham, J. P., and Wobbrock, J. O. 2008. Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proc of Assets '08*. ACM, New York, NY, 73-80.

[15] Li, F.C.Y., Dearman, D., and Truong, K. Virtual Shelves. 2009. Interactions with Orientation Aware Devices. In *Proc of UIST '09*. ACM Press, New York, NY, 125-128.

[16] Li, K. A., Baudisch, P., and Hinckley, K. 2008. Blindsight: eyes-free access to mobile phones. In *Proc of CHI '08*. ACM, New York, NY, 1389-1398.

[17] Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*. 63, 2 (1956), 81–97.

[18] O'Neill, E., Kaenampornpan, M., Kostakos, V., Warr, A., and Woodgate, D. 2006. Can we do without GUIs? Gesture and speech interaction with a patient information system. *Personal Ubiquitous Comput*. 10, 5 (Jul. 2006), 269-283.

[19] Oakley, I. and Park, J. A motion-based marking menu system. 2007. In *Proc of CHI'07 Extended Abstracts on Human Factors in Computing Systems*. ACM Press, New York, NY, 2597-2602.

[20] Pirhonen, A., Brewster, S., and Holguin, C. 2002. Gestural and audio metaphors as a means of control for mobile devices. In *Proc of CHI '02*. ACM, New York, NY, 291-298.

[21] Rekimoto, J. 1996. Tilting operations for small screen interfaces. In *Proc of UIST '96*. ACM, New York, NY, 167-168.

[22] Roumeliotis S. I. and Bekey, G.A. 1997. An Extended Kalman Filter for frequent local and infrequent global sensor data fusion. In *Proc of the Sensor Fusion and Decentralized Control in Autonomous Robotic Systems*. SPIE, Bellingham WA, 11-22.

[23] Sawhney, N. and Schmandt, C. 2000. Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Trans. Comput.-Hum. Interact*. 7, 3 (Sep. 2000), 353-383.

[24] Stifelman, L. J., Arons, B., Schmandt, C., and Hulteen, E. A. 1993. VoiceNotes: a speech interface for a hand-held voice notetaker. In *Proc of the INTERACT '93 and CHI '93*. ACM, New York, NY, 179-186.

[25] Suzuki, Y., Nakadai, Y., Shimamura, Y., Nishino, Y. 1998. Development of an Integrated Wristwatch-type PHS Telephone. *NTT Review*, Vol.10, No.6, 93-101.

[26] Vaganay, J., Aldon, M.J., Fournier, A. 1993. Mobile robot attitude estimation by fusion of inertial data. In *Proc of the 1993 IEEE International Conference on Robotics and Automation*. IEEE, Washington, DC, 277-282.

[27] Wilson, A. and Shafer, S. 2003. XWand: UI for intelligent spaces. In *Proc of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '03. ACM, New York, NY, 545-552.

[28] Yee, K.P. 2003. Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers. In *Proc of CHI '03*. ACM Press, New York, NY, 1-8.

[29] Yfantidis, G. and Evreinov, G. 2006. Adaptive blind interaction technique for touchscreens. *Univers. Access Inf. Soc.* 4, 4 (May. 2006), 328-337.

[30] Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R. and Baudisch, P. 2007. EarPod: eyes-free menu selection using touch input and reactive audio feedback. In *Proc of CHI'07*. ACM Press, New York, NY, 1395-1404.