Wrist Rotation for Interaction in Mobile Contexts

Andrew Crossan, John Williamson, Stephen Brewster, Rod Murray-Smith
Department of Computing Science
University of Glasgow
+44 141 330 8430
[ac, jhw, stephen, rod] @ dcs.gla.ac.uk

ABSTRACT
In this paper, we investigate wrist rotation as a hands-free method of interaction with a mobile device. To evaluate this technique, a Fitts' Law targeting study is described in four different postures: resting, seated, standing and walking. Results show correlations in movement time and the Index of Difficulty of the task and similarities in the targeting performance for the first three conditions, but show walking and targeting using this method was significantly more difficult.

Categories and Subject Descriptors
H.5.2 [**User Interfaces**]: *Input devices and strategies*

General Terms
Human Factors.

Keywords
Mobile, accelerometer, Fitts' Law, hands-free interaction.

1. INTRODUCTION
Mobile devices are ubiquitous now with mobile phones, portable music players, navigation devices and PDAs playing an integral part in our everyday lives. The processing power of these devices and consequently the amount of functionality that they provide is increasing rapidly, however, the form factor of the devices must be kept small to allow for portability, leading to small buttons and awkward often error prone interactions. This is particularly the case in real life mobile situations such as a busy street setting where the user has the additional task of navigating the environment safely. This problem is further compounded when the user's hands might be occupied in other tasks such as carrying bags, holding on to children or when performing some other task such as cycling. The majority of mobile devices still use techniques similar to desktop interactions when the user interacts with the device. The user generally interacts with physical or touchscreen buttons to navigate menus. Recently there has been a push towards novel forms of interaction for these devices. For example, the iPhone from Apple has introduced a touch screen with multi-point gesturing for different interactions. The Nintendo Wii, which uses motion sensing control, has had similar commercial success in the games console market reaching an audience who would not normally play computer games. With the iPhone and Wii now firmly establishing gesture and motion sensing in the mainstream there is an opportunity to provide users with novel and non-traditional forms of input and output on more mainstream mobile devices.

Previous gesture work has concentrated on hand and finger gestures [10, 13]. Little work has gone into making input and control hands-free for mobile users. Speech recognition is one hands-free example, however, it is still problematic in mobile settings due to social awkwardness, high processing requirements, and the dynamic audio environments in which devices are used. The idea of a multipart mobile device has potential to improve hands-free interaction. Bluetooth headsets are the most commonly used examples where users can talk with their phone still in their pocket. The Nike Plus plug in for the iPod allows runners to get information about their performance through an inertial sensor in their shoe. Other devices such as the Bluetooth Vibrating Bracelet from LM Technologies (which vibrates when there is an incoming call) provide output through different modalities. These devices have the common aim of allowing a user to interact with their device without their hands being encumbered.

The overall goal of our research is to extend these ideas and to explore techniques that will allow users to interact with mobile devices in a more eyes-free and hands-free manner. This requires two areas of research. Firstly, identifying body locations that can be used to control a user interface and secondly, developing audio and tactile techniques to allow a user to navigate a mobile interface eyes-free. This paper concentrates on the former. One theme of the research will be an emphasis on development for mobile situations and investigating and supporting different usage contexts such as walking, cycling or sitting on the bus. This paper concentrates on one potential body location that could provide hands-free interaction. In this case, we examine the ability of users to target using wrist-rotation as an input technique.

2. Background
2.1 Hand and Finger Gesturing
Gesture input has been successfully incorporated into both research and commercial mobile devices. Much work has focused on touch-screen gestures, initially using the stylus-based interactions such as the standard graffiti gesture set, which allows users to draw symbols with the system attempting to classify the trajectory into one of a predefined set of characters. The SHARK system [18] has a similar goal of mobile text entry, however, uses trajectories drawn on a keyboard with a language model supporting the interactions to allow users to enter words.

Recently, there has been work to remove the stylus from the interactions and allow the user to interact with the screen directly

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
MobileHCI 2008, September 2–5, 2008, Amsterdam, the Netherlands.
Copyright © 2008 ACM 978-1-59593-952-4/08/09…$5.00.

using fingers. The Apple iPhone is leading this push with a multi-point of contact touchscreen that allows a user to gesture using multiple fingers on the screen for scrolling and zooming photos. Pirhonen *et al.* [13] describe and evaluate a system using finger gestures to control a simple mobile MP3 player. One important feature of this system was that users were able to use it eyes free. This is key for mobile situations where users' visual attention is divided between navigating and their interactions.

There is a growing body of work using inertial sensing to determine user intention. Much of this work uses accelerometers to sense device orientation or user movements as they are small and cheap. Early work was by Rekimoto [14] who discussed the potential of this technique for tasks such as navigating menus and scrolling. These have since been refined and used in a number of different contexts. Hinckley *et al.* [6] demonstrate how accelerometers could be useful for interaction and context sensing with demonstrations of an automatic screen orientation device and scrolling application. Oakley and O'Modhrain [10] describe a tilt-based system with tactile augmentation for menu navigation.

For text entry, tilt has been used as an interaction technique for the TiltType system designed to allow a user to interact with extremely small devices with few buttons [12]. Here multiple characters are assigned to a single physical button with the tilt angle of the device when a button is pressed being used to disambiguate the characters types. Similarly, Williamson and Murray-Smith examined tilt as a method of text entry on a mobile device with the Hex system [16]. They combined tilt input with a language model to allow the system to infer the current word being typed and adjust the dynamics of the system in order to make that word easier to enter. The user tilted the device to move a cursor through the landscape with smaller tilt angle being required to reach more probable characters.

The techniques described above focus on hand and finger gesturing for interaction, which can be difficult while mobile. There has been less work on hands-free gestures. In many mobile settings it may not be possible for a user to operate a device that needs one or two hands; they may already be occupied by holding shopping bags or children. Alternative methods of controlling devices need to be studied to provide effective techniques for hands-free input when on the move.

2.2 Hands-Free Gesturing

There is currently little work examining viable areas of the rest of the body for gesturing, particularly for interaction in a mobile context. 'Hands free' presently means using a headset to speak on the phone without holding it; the other interactions and applications a device can perform are inaccessible without the hands. Speech recognition has possibilities here, but is difficult to do on mobile devices due to processing requirements and dynamic audio environments. There are also situations where speech may be inappropriate (quiet environments, for example). Non-hand based gestures may provide a good alternative here.

Although less work has been carried out in this area, there are examples of previous work on devices and interaction techniques. Rekimoto describes GestureWrist in [15], which recognises user hand gestures. Although users are required to move their hands, the device is attached to the wrist such that users can interact with the world without sensors encumbering their hands.

Previous work has examined pointing with different joints in the arm to control a cursor in a desktop situation. Zhai *et al.* [19] investigated the use of fingers, left/right motion of the wrist, elbow and shoulders in a Fitts' Law task for pointing in a graphics application. Balakrishnan *et al.* [2] similarly examine finger, wrist and forearm performance in a computer based pointing task. Fitts' Law studies have examined head pointing for targeting [7]. Also very relevant to this work is Oakley & Park's motion-based marking menu system [11], which relies on wrist rotation (roll) to navigate. This method allows menu selection in a mobile situation in a hands-free manner. Our own work includes a study of mobile head pointing using one axis of rotation of the head to select menu items [3]. These studies show that using different body locations for input is possible, but they have not been studied in a systematic way across the body, or while mobile.

2.3 Fitts' Law Targeting

Fitts' Law is a common method of characterising performance in a one dimensional targeting task [5]. Participants repeatedly move between two targets of varying widths and separations. Fitts describes how the time to target varies with the ratio of target separation to target width. This method has since been used by researchers for describing targeting performance for manual control tasks and in HCI (notable work in this area includes MacKenzie [9] and Accot and Zhai [1]). In this paper, we will use Fitts' Law to character a user's performance in a targeting task where the user controls a cursor using wrist rotation. Fitts' Law states that:

$$MT = a + b * ID, \text{ where } ID = \log_2(A/W + 1)$$

Where MT is movement time, a and b are task dependent constants and ID is the index of difficulty, A is the target separation, and W is the target width. Here, we use the Shannon formulation for ID proposed by MacKenzie [8]. The values of a and b can be seen as a measure of the user's reaction time and the task difficulty respectively. This formula assumes an error rate of approximately 4%. For occasions where this is not the case, MacKenzie [8] describes a method to calculate the effective width of the target based on the standard deviation of the errors. For this study we use this method to calculate the effective target width when required.

3. Experiment

This initial experiment was to evaluate wrist rotation as an input technique for interaction with a mobile device in both static and mobile contexts. Although the goal of the research is to eventually develop hands-free, eyes-free interaction techniques, in this study we just focus on investigating if wrist rotation would be accurate enough for input. Participants viewed their interactions on the screen of a Nokia N95 phone held in one hand (such that results would not be affected by a potentially poor choice of auditory or tactile interface design) and selected targets by pressing the button as shown in Figure 1 (such that results would not be affected by a potentially difficult gesture-based selection mechanism). This allows us to investigate if wrist rotation was effective for input.

When interacting with the system users control the cursor by holding their arm horizontal and rotating their wrist. To sense wrist rotation we used the 3 axis accelerometer in a SHAKE sensor pack that was strapped to the user's right wrist. A position control mechanism was used to move the cursor. The cursor position changed linearly with the participants' wrist roll angle with a 90 º rotation moving the cursor the full screen length. The palm facing down position corresponded to the cursor on the left edge of the screen and the palm facing left position corresponded to the cursor on the right edge of the screen with a 45º roll angle corresponding to the screen centre. A low-pass filter allowing frequencies up to 2Hz was used to remove noise.

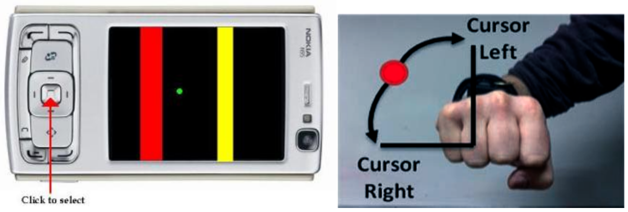

Figure 1. The interface used in the study (left) and a SHAKE strapped to the top of a user's wrist (right).

24 participants took part in the study. All were right-handed and age 23 to 36. A SHAKE sensor pack [17] was strapped to a participant's right wrist and the phone was held in the left hand at the orientation shown in Figure 1 such that the participant's left thumb was over the selection button. The user was set the task "to click on the yellow target as quickly as possible". Two targets were shown at all times and the yellow target alternated between the left and right of the screen. Four conditions were performed by all participants in a counterbalanced order. These were: resting (C_{RE}), seated (C_{SE}), standing (C_{ST}) and walking (C_{WA}). The postures in each of these conditions are shown in Figure 2. The C_{RE} involved the participant sitting and resting their right arm on a book on a table to minimise muscle tremor, C_{SE} had the participants sitting with their right forearm horizontal and not resting on anything, C_{ST} had the participants standing still with their forearm horizontal, and C_{WA} had the participant walking figures of eight around two cones placed ~3 metres apart. This walking route was chosen such that participants had to divide their attention between the task and navigating [13].

Figure 2. Three different postures used in the study. Left shows standing, right-top rested, and right bottom seated.

Each condition consisted of four blocks of 64 targets with an initial dummy target to start the block. Each block corresponded to a single target separation (~20º, ~28º, ~42º, and ~56º) with the blocks being presented in a random order. The 64 targets contained sixteen (eight left-side and eight right-side) targets of four different widths (~6º, ~9º, ~12º, and ~15º). These angles are approximate as the actual values will vary slightly with the tilt of the user's arm. As the lower arm moves further away from the horizontal the cursor gain will decrease as the user moves close to rotating around the gravitational axis. Our hypotheses were:

1. Participants would be significantly more accurate and faster in C_{RE} due to the reduced signal noise from muscle tremor
2. Participants would be significantly slower and less accurate in C_{WA} due to the added signal noise from the walking.

The Fitts' Law ***a*** and ***b*** constants were also calculated for all conditions to allow comparisons with other studies.

4. Results

The results for time and accuracy were analysed using ANOVA tests with index of difficulty (ID) and condition as factors. As effective target widths are used to calculate the ID, we group similar IDs by rounding down to the nearest integer. A significant effect of condition was found for both Movement Time ($p < 0.0.1$) and Accuracy ($p < 0.01$). Tukey HSD test showed that in both case the significance was due to C_{WA} with all other conditions having a significantly lower MT ($p < 0.01$) and significantly higher accuracy ($p < 0.01$). No significant differences were found between the other conditions.

Figure 3 illustrates the effect that changing target width and separation had on participants mean movement time and accuracy for C_{SE} and C_{WA} (C_{RE} and C_{ST} provide similar mean movement time and accuracy to C_{SE} but are removed for clarity's sake). For all target separations and widths, C_{WA} has a longer mean movement time and lower mean accuracy.

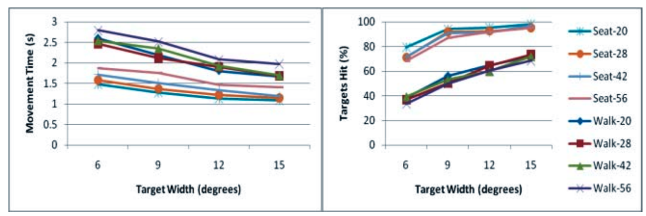

Figure 3. Movement time (left) and percentage of targets hit (right) for C_{SE} and C_{WA} for all width-separation combinations.

	a (s)	***b (s)***	***r***
C_{RE}	0.37	0.49	0.61
C_{SE}	0.42	0.42	0.62
C_{ST}	0.41	0.47	0.59
C_{WA}	0.51	0.71	0.53

Table 1. The values of *a* and *b* for the four conditions. *r* is the Pearson product moment correlation coefficient.

Values calculated for ***a*** and ***b*** are shown Table 1 with a Pearson product moment correlation value for each. In each case the correlation coefficient is above 0.5 indicating a correlation between MT and ID.

Figure 4 shows graphs of movement time against ID. Each point on the graphs corresponds to the mean performance of a participant over one width-separation combination.

5. Discussion

There was no significant difference found in the timing or accuracy data between the rested condition and the seated and standing conditions. This means that we cannot support the hypothesis that the rested posture improved performance. Anecdotally, a minority of participants complained that resting on the book restricted their movements slightly and made the task somewhat harder, but again this was not borne out by the data. In the static conditions, participants achieve a high level of accuracy (~90%) for the 9º and larger targets. This suggests that participants could successfully targeting using wrist rotation. The walking condition was however was significantly harder for all participants who were both slower and far less accurate than in all other conditions for all target width-separation combinations. Hypothesis 2 is therefore supported by the data. It is also interesting to note that the accuracy values for this condition are less than 80% in all cases. All participants commented on the difficulty of the task and expressed low confidence in their performance. The accelerometer signal contains both the tilt from the user's targeting, and noise generated by the walking behaviour, which cause the cursor to oscillate in time with the user's walking speed making it difficult to target.

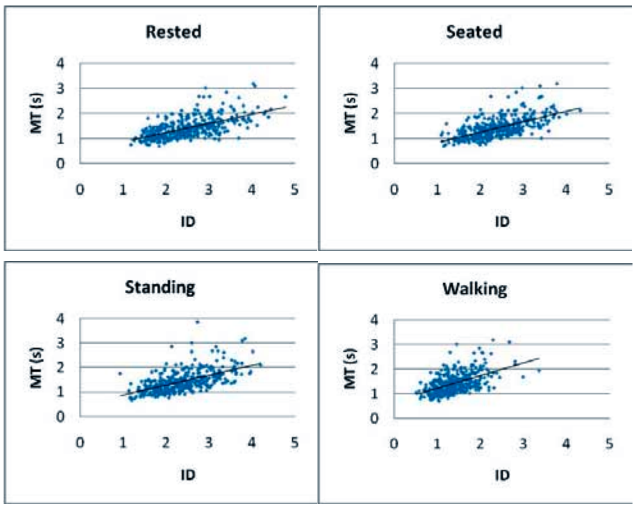

Figure 4. Movement time plotted against Index of Difficulty for all conditions.

Table 1 shows the relevant ***a*** and ***b*** values for the Fitts' Law equation. In each case, the correlation coefficient is above 0.5. This indicates that there is a correlation (although not strong) and the task does follow Fitts' Law. The spread of values in Figure 4 shows that there was a lot of variability in the Movement Time which indicates that some users may have place more emphasis on accuracy than speed during the task.

6. Conclusions

This study is the first in a series of studies to investigate the potential of inertial sensing and a multipart mobile device to provide hands-free interaction in a mobile context. High success rates in the static conditions for targets greater than 9^o in width suggests that wrist rotation can be used successfully to select targets in a hands-free manner, but when on the move, disturbance from the user walking made the task significantly harder. A failing of accelerometers for this task is that as the user walks and the device starts to move, the sensors give a poorer estimate of orientation. It is hoped that by building on our previous work on gait analysis [4], we will be able to use the knowledge of the repetitive oscillations in the users' walking behaviour to improve this mobile interaction. The postures chosen for this study were not the unobtrusive movements that would be preferable when walking in a busy street. It is necessary to examine methods of making these movements smaller and less obvious to other people. This will require the combination of different sensors such as magnetometers or gyroscopes to overcome some of the limitations of accelerometers alone. The static results show wrist rotation is a potentially successful hands-free technique for mobile interaction. Future work will look replicating this success in a more general mobile context.

Acknowledgements

This work was funded by EPSRC grant EP/F023405, Nokia and Advanced Research Fellowship GR/S53244.

References

[1] Accot, J. and Zhai, S. 2003. Refining Fitts' law models for bivariate pointing. in Proceedings of ACM CHI, 193-200.

[2] Balakrishnan, R. and MacKenzie, I. S. 1997. Performance Differences in the Fingers, Wrist, and Forearm in Computer Input Control. in Proceedings of ACM CHI. Atlanta Georgia, USA: ACM Press, 303-310.

[3] Brewster, S. A., Lumsden, J., Bell, M., Hall, M. and Tasker, S. 2003 Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices. in Proceedings of ACM CHI. 463-480.

[4] Crossan, A., Murray-Smith, R., Brewster, S. and Musizza, B., Instrumented Usability Analysis for Mobile Devices, in Handbook of Mobile HCI. 2007, The Ideas Group Inc.

[5] Fitts, P. M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6): p. 381-391.

[6] Hinckley, K., Pierce, J. and Horvitz, E. 2000. Sensing Techniques for Mobile Interaction. in Proceedings of ACM UIST, 91-100.

[7] Jagacinski, R. and Monk, D., 1985. Fitts' law in two dimensions with hand and head movements. Journal of Motor Behavior, 17: p. 77-95.

[8] MacKenzie, I. S., 1992. Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction, 7: p. 91-139.

[9] MacKenzie, I. S., Kauppinen, T. and Silfverberg, M. 2001. Accuracy Measures for Evaluating Computer Pointing Devices. in Proceedings of the ACM CHI: ACM, 9-16.

[10] Oakley, I. and O'Modhrain, M. 2005. Tilt to scroll: evaluating a motion based vibrotactile mobile interface. in Proceedings of World Haptics: IEEE, 40-49.

[11] Oakley, I. and Park, J. 2007. Designing Eyes-Free Interaction. in Proceedings of HAID. Seoul: LNCS 4813, 121–132.

[12] Partridge, K., Chatterjee, S., Sazawal, V., Borriello, G. and Want, R. 2002. TiltType: Accelerometer-Supported Text Entry for Very Small Devices. in Proceedings of UIST.

[13] Pirhonen, A., Brewster, S. A. and Holguin, C. 2002. Gestural and Audio Metaphors as a Means of Control for Mobile Devices. in Proceedings of ACM CHI., 291-198.

[14] Rekimoto, J. 1996. Tilting Operations for Small Screen Interfaces. In Proceedings of UIST, 167-168.

[15] Rekimoto, J. 2001. Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. in Proceedings of Fifth International Symposium on Wearable Computers.

[16] Williamson, J. and Murray-Smith, R., Dynamics and probabilistic text entry. 2003, Department of Computing Science, Glasgow University.

[17] Williamson, J., Murray-Smith, R. and Hughes, S. 2007. Shoogle: Multimodal Excitatory Interaction on Mobile Devices. in Proceedings of ACM CHI. San Jose.

[18] Zhai, S. and Kristensson, P. 2003. Shorthand writing on stylus keyboard. ACM CHI. Florida, USA: ACM, 97 - 104.

[19] Zhai, S., Milgram, P. and Buxton, W. 1996. The Influence of Muscle Groups on Performance of Multiple Degree-of-Freedom Input. in Proceedings of ACM CHI. Vancouver, BC, Canada: ACM Press, 308-315.