Head Tilting for Interaction in Mobile Contexts

Andrew Crossan, Mark McGill, Stephen Brewster, Roderick Murray-Smith
Glasgow Interactive Systems Group, Department of Computing Science
University of Glasgow, Glasgow, G12 8QQ, UK
+44 141 330 8430
[ac, stephen, rod] @ dcs.gla.ac.uk
www.gaime-project.com

ABSTRACT

Developing interfaces for mobile situations requires that devices are useable on the move. Here, we explore head tilting as an input technique to allow a user to interact with a mobile device 'hands free'. A Fitts' Law style evaluation is described where a user acquires targets, moving the cursor by head tilt. We explored position and velocity control cursor mechanisms in both static and mobile situations to see which provided the best level of performance. Results show that participants could successfully acquire targets using head tilting. Position control was shown to be significantly faster and more accurate in a static context, but exhibited significantly poorer accuracy and longer target acquisition times when the user was on the move. We further demonstrate how analysis of user's gait shows consistent targeting biases at different stages in the gait cycle.

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

As the functionality of mobile devices increases and the devices evolve to 'stand out' from their competitors, new techniques are needed to allow users to interact in different, and more efficient and engaging ways. These devices present different challenges to a developer than more traditional desktop computers as the interfaces must be designed to be used in a variety of different circumstances outside the home or office environment. One of the most challenging of these everyday circumstances is to design for use 'on the move' while walking or cycling for example. The user's attention is focussed on the main task of navigating the environment safely, crossing roads or avoiding obstacles and other pedestrians, all of which are visually demanding tasks. Another important issue is that most devices require hands to operate many of the applications. Users may also be encumbered by bags or children which will affect their ability to hold the device and interact with it. For cold climates, gloves also affect how the user interacts with the device, making traditional button based and touchscreen interactions more difficult.

Little work has gone into making input and control hands free for mobile users. When interfaces are described as 'hands free' it generally refers to interactions through speech. This works well for phone calls, but when interface control is required there must be a recognition component available on the device. Speech recognition is still problematic in such settings due to its high processing requirements and the dynamic audio environments in which devices are used. Much of the research on gesture input still uses hands for making the gestures. There is some work on head-based input, often for users with physical disabilities [3, 15], but little of this has been used in mobile settings.

Many other body locations could be useful for subtle and discreet input whilst mobile (e.g. users walking or sitting on a bumpy train). For example, wrist rotation has potential for controlling a radial menu as the wrist can be rotated to move a pointer across the menu. It is unobtrusive and could be tracked using the same sensor used for hand pointing gestures. Shoulders are able to make a range of subtle movements with several degrees of freedom. Small changes in gait are also a possibility for interaction. Users could interact with a mobile device by slightly changing the timing of a step to make input. Using a Fitts' law style analysis along with measures of workload, comfort and social acceptability, will allow us to explore whether these methods are viable input techniques.

Multipart mobile devices offer the potential to allow an entirely different method of interacting with a phone. A Bluetooth 'hands free' headset is the most common example of extending the phone interactions onto a separate device allowing a user to conduct a call without removing the phone from his/her pocket. The Nike Plus device (which connects to an iPod through Bluetooth) is another example of such a device (nikeplus.nike.com). Users interact with the mobile device through a sensor attached to the shoe without the need to remove the device from their pocket. Feedback is provided through the audio channel on an Apple iPod allowing eyes-free as well as hands-free interaction. These devices allow interactions that are fast, discreet, low effort and engaging.

In this paper we describe a study that examines non-hand based interaction with a mobile device. To be hands free and eyes free, we must also eventually consider output. There has previously been research into investigating non-visual auditory and tactile methods of providing users with feedback to allow them to interact with a phone non-visually (e.g. [5, 22, 31]). In this instance we concentrate on the input side only to study the most effective forms of hands free interaction, in this case using head tilting to target objects on a mobile phone screen.

MobileHCI'09, September 15 - 18, 2009, Bonn, Germany.
ACM 978-1-60558-281-8.

2. RELATED WORK

2.1 Gesture and Mobile Interaction

The current generation of mobile phones often now offer both interaction through traditional menu and buttons as well as providing one or more sensors (such as touch sensitive areas, accelerometers, GPS or magnetometers) to allow the user to interact in a variety of different ways.

Gesture is increasingly being seen as an interaction mechanism suitable for interaction with mobile devices [19, 23, 30]. These gestures can be loosely classified into discrete action gestures and continuous control gestures. The more traditional style of gesturing, discrete action control, involves the user performing an action that, once completed, the system attempts to recognise. Generally, this form of gesturing is used to perform a discrete operation such as opening an application on the device, and is often rationalised as providing a mechanism to allow rapid access to commonly used functionality. These gestures are often used to replace one or multiple button clicks. Successful examples of these techniques in commercial devices often involve short movements that are fast and easy to perform such as a single stroke of a touch screen to change the view, or double tapping a phone to silence it [30].

Recently, continuous control gestures have become more prevalent [23, 31]. With these forms of gesture the interactions between the user and the device are closely coupled. The user provides a continuously changing stream of input, and the device adjusts the feedback constantly to respond to the user's input. One example would be scaling and rotating photos on a table top display. The user places two fingers at the corners of the image on the display, and the image is resized and rotated as the fingers are moved. In this instance, the feedback from the display can be continuously changed as the user interacts, allowing the user to more easily select the desired size. The control mechanism in this example allows the zoom and rotation of the on-screen image to be controlled by the user at the same time, a task that would be more difficult and time consuming with button-based interactions. Both discreet action and continuous control gestures are employed on the current generation of devices such as the Apple iPhone.

2.2 Gesturing Using the Hands

Gesturing on the current generation of mobile devices generally involves interaction through touchscreens. There are many commercially available touchscreen gesture systems (such as Graffiti for text entry from Palm) as well as examples from the literature. In Graffiti users draw symbols and the system attempts to classify the trajectory into one of a predefined set of characters. The SHARK system [33] has a similar goal of mobile text entry, however, uses trajectories drawn across a keyboard with a language model supporting the interactions to allow users to enter words. There has more recently been work to remove the stylus from the interactions and allow the user to interact with the screen directly using fingers. The Apple iPhone is leading this push with a multipoint of contact touchscreen that allows a user to gesture using multiple fingers on the screen for scrolling and zooming photos.

Pirhonen *et al.* [26] describe a mobile music player that can be controlled purely through touch. They use a touch screen PDA as the player, and allow users to change track or control the volume by drawing gestures with a finger on the touchscreen. The aim of this prototype was to allow users to interact with the system eyes-free with one hand in a mobile setting using the physical form of the device to guide the gesture. For example, a runner could interact by drawing on the PDA strapped to his/her belt without stopping or looking at the screen.

Gesture interaction on touchscreen devices is now firmly established in the mainstream mobile applications. Outside of touchscreen interactions, however, there is a growing body of work on incorporating other sensor technology into the interactions. Inertial sensors are now standard in many high end phones. Driven by the success of the Nintendo Wii (www.nintendo.com), these sensors provide a low cost method of sensing movement or device orientation for gesture input or context sensing. Unlike many of the touchscreen interactions, one advantage of using such sensors is that they can provide one handed, screen free gesture control. For context sensing applications, the device can remain in the pocket and monitor background movements continuously.

Early work on accelerometer-based input for mobile interaction was pioneered by Rekimoto [27] who describes how using accelerometers to estimate the orientation of a mobile device could be used in a variety of circumstances, for example to scroll or move a cursor. Since this work, there have been many examples of systems that exploit inertial sensors. Hinckley *et al.* [11] demonstrated multiple interactions possible with a sensor enabled mobile device. In the simplest case, accelerometers are used to automatically orient the screen in landscape or portrait mode. They further describe how mode switching would occur on the device with inputs from combinations of sensors being used to sense context and automatically adapt. Work by Eslambolchilar [9] has examined coupled zooming and scrolling operations using tilt based input to navigate large documents on a small screen. Fine grain control of dual axis accelerometers allows the user to easily control the speed of scroll of the document in two-dimensions simultaneously; a task that would be difficult with buttons alone. This interface still uses a screen to provide visual feedback to the users. Oakley and O'Modhrain [22] describe a tilt-based system with tactile augmentation for menu navigation. The goal of their work was to provide a system that allowed users eyes free interaction with menus.

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 [24]. 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. This illustrates how an accelerometer unlike a button does not require valuable real-estate on the outside of a mobile device, allowing the devices to be smaller and less cluttered. Similarly, Williamson and Murray-Smith examined tilt as a method of text entry on a mobile device with the Hex system [32]. 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.

Gesturing with inertial sensors need not involve long or complex movements. Linjama and Kaareoja [14] describe a gesture system based on tapping the device in different places, allowing low effort, discreet interaction ideal for discrete action events. Tactile feedback alerts the user to the completion of an action event. Williamson *et al.* [31] describe Shoogle, an interface that allows interaction through shaking a device. Each piece of information is modelled physically as a ball attached to a spring. As the user shakes the device, the balls interact with virtual walls with auditory and vibrotactile feedback used to provide the user with information.

The above examples describe a selection of the work that has been conducted studying gestures in mobile situations where the user interacts through moving or gesturing with the device. They demonstrate some of the benefits that hand-based gesturing can bring to mobile interfaces. However, they still require that the users' hands are free. The next section addresses previous work on interfaces that do not encumber the user's hands.

2.3 Non-Hand Based Gesturing

The above interfaces describe situations where users interact through gesture, moving the mobile device in some manner. In particular, the device is held in the user's hand and interaction occurs in either a one handed or two handed manner. 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. Inspiration from this work is taken from the accessibility literature where interfaces are designed for people with motor impairments [3, 15]. However, there are few examples of work examining viable areas of the rest of the body for gesturing in mobile situations. '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 to allow users to interact with a device in a discreet manner.

Although less work has been carried out in this area, there are examples of previous work on devices and interaction techniques. For a desktop environment, LoPresti *et al.* [15] examine head movement as a means of interacting with a computer. They compare the performance of people with a full range of neck movement to people who have a restricted range of movement, showing how task difficulty increases for people with movement restrictions. Similarly [21] explores head gestures for interacting in a desktop context. Camera-based head tracking allows the user to interact with a dialog box by nodding or shaking their head. Previous work has examined pointing with different joints in the arm to control a cursor in a desktop situation. Zhai *et al.* [34] 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. A previous study described [13] in studies have examined hand and head pointing for targeting for a static situation, showing that in this situation, Fitts' Law described head movement.

For mobile settings, the idea of a multipart device opens up the possibility of discreet, subtle interactions that remove the need to take the mobile device out of a pocket. For a multipart mobile device system, one or more sensors are placed around the body to detect input from the user. A Personal Area Network would be used to link these sensors to the device and allow the user to control an interface with the need to touch the device. Costanza *et al.* [6] describe such an interface using electromyogram (EMG) based interactions. A device monitoring activity through EMG is mounted on a user's upper arm. The sensor can then detect when the user flexes a bicep muscle. The advantage of this system is that it can detect isometric muscle activity allowing a motionless gesture. In a public location, it would be difficult for other people to determine that a gesture was being performed allowing privacy when interacting. Other relevant work includes Rekimoto [28], who describes GestureWrist, 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. Also very relevant to this work is Oakley & Park's motion-based marking menu system [23], 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 an initial study of mobile head pointing using ego-centric nod gestures to select items spatialised around the user with 3D audio [5]. 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.

Our previous work examines wrist rotation as a cursor control technique [8] using a position control mechanism for moving the cursor. Participants rotated their wrist within a 90 degree workspace to move a cursor the length of the screen of a Nokia N95 and click a button when over a target to select. A Fitts' law style study was used to measure users targeting performance in a number of different postures; rested, seated, standing and walking. Results showed that an accelerometer could successfully be used to target in this manner in static conditions but was less successful in the walking condition. The study described in this paper tests both a position control (where there is a one-to-one mapping between head tilt and cursor position) a velocity control (where there is a one-to-one mapping between head tilt and cursor velocity) mechanism for the pointing task. We expect less disturbance for a head-based task than the previous wrist pointing task as there is less movement of the head while walking.

3. EVALUATION BACKGROUND

For this study we will take a Fitts' Law style targeting approach to evaluate head pointing as an input technique, using an instrumented usability approach to gather data.

3.1 Fitts' Law Targeting

Fitts' Law is a common method of characterising performance in a one dimensional targeting task [10]. 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 [18] and Accot and Zhai [1]). In this paper, we will use Fitts' Law to characterise a user's performance in a targeting task where the user controls a cursor using wrist rotation. Fitts' Law states that:

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

Where $\boldsymbol{MT}$ is movement time, $\boldsymbol{a}$ and $\boldsymbol{b}$ are task dependent constants and $\boldsymbol{ID}$ is the index of difficulty, $\boldsymbol{A}$ is the target separation, and $\boldsymbol{W}$ is the target width. Here, we use the Shannon formulation for ID described by MacKenzie [16]. The values of $\boldsymbol{a}$ and $\boldsymbol{b}$ 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 [16] 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.

Further to this, we calculate throughput for the interactions [17]. We measure throughput (TP) as:

$$TP = eID_{mean} / MT_{mean}$$

Where eID_{mean} is the effective index of difficulty, calculated from the mean ID using effective target width, and MT_{mean} is the mean movement time. This metric gives a measure in bits/second of the bandwidth of the communication channel and can be used to compare the performance of different pointing devices.

3.2 Mobile Evaluation

Evaluation of an interface that is intended for use in a mobile setting is a challenging task. Research has shown that when interacting with a device in a mobile context performance may drop by more than 20% [4]. It is therefore important for designers to evaluate their interfaces in as realistic a scenario as possible. However, more control over the evaluation is possible in a usability lab setting at the cost of potential loss of realism of the scenario. Here we compromise by running our evaluation indoors, but asking the user to perform the task both standing still and walking. When walking, the participants are asked to follow a path such that they had to concentrate on navigation as well as the interactions. This technique has previously been used successfully in [5].

In our previous work we developed the concept of *instrumented usability* [7]. This examines how, through the instrumentation of the users or mobile devices with sensors, more information about the moment-to-moment context of use can be gathered during an evaluation. This could be as high level as examining the variation in the step rate of users as they walk through a preset path performing a task [7], or more detailed information about the disturbances that the device is experiencing while the user sits on a moving train [12]. Data from one or more sensors can be gathered during a mobile usability study and analysed *post hoc* to provide insight into the data that would not be possible otherwise. In the study described in this paper, we instrument the users with an accelerometer to monitor their walking behaviour as they perform the task to tell us more about the effects of the different types of control on input using head tilting.

3.2.1 Extracting Gait Information

Two SHAKE sensor packs [31] (shown in Figure 1) were attached to the user during the study. One on the participants' heads for cursor control, and one fixed to their upper back to extract gait information during the study. The SHAKE is a small portable sensor pack (containing a 3-axis accelerometer and a magnetometer) that can connect through Bluetooth to a range of different devices. In this instance, the SHAKEs are connected to a mobile phone used during the study, allowing us to synchronise the data from the SHAKE and the user's explicit interactions.

For the purposes of this study, we analyse the acceleration data to retrieve information about the users' gait as they perform the task. As a user walks while holding a mobile device, his/her body will oscillate as a result of the rhythmic nature of walking. A sensor pack strapped to the top of the participant's back will oscillate at the same rate as the step rate. If we examine only the lateral axis of this oscillation (as the user sways left and right), there will be one complete oscillation every two steps corresponding to a left step and a right step. We can then use the phase of this oscillation to estimate the moment-to-moment stage of the user's gait cycle.

Figure 1: The SHAKE sensor pack.

Figure 2 shows a time series of one acceleration axis. A Fast Fourier Transform is used to determine the frequency at which the peak amplitude occurs, between 0.5 and 1.5Hz in the spectrum (60 to 120 steps per minute). For the controlled conditions in this study, this corresponds to the walking step rate. In practice, this is the frequency of maximum power in the spectrum as there are few other disturbances. The acceleration signal is then zero phase shift filtered using a narrow bandpass Butterworth filter centred around this frequency. Figure 2 demonstrates the filtered signal as the smooth oscillating line. A regular oscillation can be seen with one complete oscillation corresponding to a left/right step cycle. We then use the Hilbert transform to extract the phase information from the signal [25]. A similar method is used to extract gait phase in [7]. We extend this method by calibrating this oscillation with real-world walking behaviour, using a pressure sensor attached to the sole of a user's right shoe.

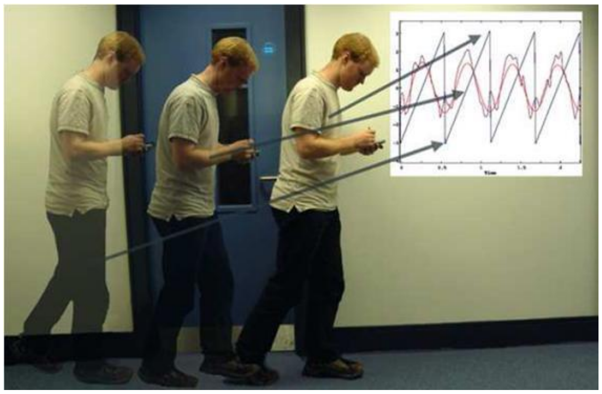

Figure 2. A user walking with the device and the corresponding acceleration trace. The unfiltered acceleration signal (rough sinusoid), the filtered signal (smooth sinusoid) and the phase estimate for the signal (saw-tooth) are shown.

Using the lateral acceleration component, one complete phase cycle corresponds to a left and a right step. To link this phase angle to the actual stage in the gait cycle, a pressure sensor attached to the sole of a user's right shoe is used to detect when the right foot is down. The top section of Figure 3 shows the summation of the measured pressure values for a user at different phases in the gait cycle (grouped into 10 phase bins). A clear increase in pressure is shown on the right hand side of the graph (phase bins 9 and 10) indicating a right foot down event at that measured phase angle. A simple pressure threshold algorithm allows us to define when this right foot down event occurs, with the bottom section of Figure 3 showing that the vast majority of these events occur at phase bins 8, 9 and 10. Using this data, we can synchronise our measured gait phase with the user's walking behaviour, not only allowing left and right steps to be segmented, but to provide a continuous estimate of the current phase of the user's step. This allows us to analyse interactions between users' input and their walking behaviour. All phase plots in this paper will use these ten phase bins to allow the user's gait cycle to be separated into left and right steps.

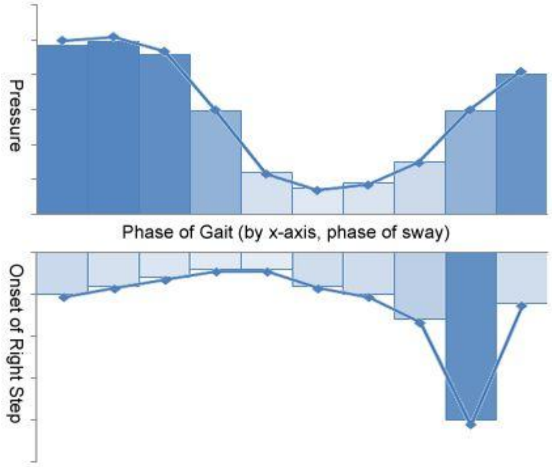

Figure 3. Top: Summation of pressure readings by phase of gait. Bottom: Summation of right step events by phase of gait. A clear correlation exists, thus we can accurately map phase data to phase of step.

4. EXPERIMENT

This experiment was designed to evaluate head tilting as an input technique for interaction with a mobile device in both standing and walking contexts. The goal of the research is to eventually develop hands-free, eyes-free interaction techniques, with this study building upon previous research into wrist rotation as a mechanism for input [8]. As the wrist rotation task was previously difficult in a mobile context, we further extend the technique comparing two different cursor control mechanisms allowing the user to control the position, or the velocity of the cursor. A similar methodology to the previous wrist study is used to allow comparisons between the throughput results from the two studies.

4.1 Methodology

The goal of the study was for participants to move a cursor by tilting their head and to click on a target shown on a phone screen. Although eventually our work will examine eyes free and hands free interaction, for this study participants viewed their interactions on the screen of a Nokia N95 phone held in their left hand (Figure 4). This was done so that results would not be affected by a potentially poor choice of auditory or tactile interface design. In a more realistic setting the auditory display may be more like that presented by Marentakis [20]. The participant selected targets by pressing the button as shown in Figure 4 such that the targeting results would not be affected by a potentially difficult gesture-based selection mechanism. This allows us to investigate if the head tilting mechanism was effective for input. In a more realistic system selection could be via a head nod.

When interacting, users control the cursor by tilting their heads left or right (within a range of ±40°). To sense head tilting we used the 3 axis accelerometer in a SHAKE sensor pack that was attached to the user's head via a customised hat, as shown in Figure 5. The accelerometer is stuck to the brim of the cap allowing the tilt angle of the head to be estimated. For controlling the cursor, two different mechanisms were used: the first, a position control mechanism where the position of the cursor changed linearly with respect to the angle of the participants head tilt (the cursor is moved to a fixed position based on the tilt of the head). The second, a velocity control mechanism where the velocity of the cursor movement was determined by the angle of head tilt (the greater the tilt the faster the cursor moves).

For the position control mechanism, the central head position corresponded to the cursor being in the centre of the phone screen. By tilting their heads left, participants could move the cursor left, with the leftmost screen position corresponding to a tilt angle of 40° (and *vice versa* for tilting right). The workspace bounds were determined through informal pilot testing to be comfortable while still allowing as large a workspace as possible.

For the velocity control conditions, the velocity of the cursor changed linearly with respect to the head tilt angle with the central head position corresponding to a stationary cursor. The maximum cursor speed (which corresponded to a 40° tilt angle) was approximately 165 pixels/second. Unlike the position control condition, using velocity control allows robustness to be built into the control by using a dead zone as there is no one-to-one mapping between head tilt and cursor position. A dead-zone of ±5° was included to provide some robustness to noise from normal small head movements. Within the dead zone the cursor did not move. These values were set through informal pilot testing. A low-pass filter allowing frequencies up to 2Hz was used to remove noise from the tilt signal for both position and velocity control conditions.

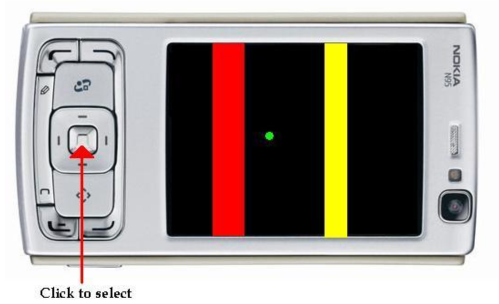

Figure 4. The interface used in the study. The Nokia N95 phone was held in the hand, with the cursor moving horizontally left and right as the user tilted his/her head left and right.

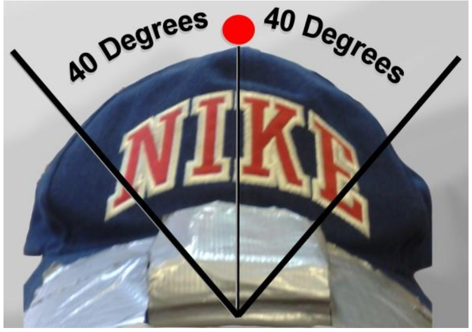

Figure 5. A SHAKE fixed to the brim of a cap on a user's head. The cursor was moved by the user tilting his/her head left and right as shown.

Twelve participants took part in the study, within an age group of 20 to 36. All were staff or students from Glasgow University. The 3 degree of freedom accelerometer in a SHAKE sensor pack [31] was attached to a participant's head, and the phone was held in the left hand at the orientation shown in Figure 4 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 (the current and next targets) were shown at all times and the yellow target alternated between the left and right of the screen. The users' task was to move over a target and click the phone button to select before moving from this selection position to the next target, continuing until there were no more targets.

4.2 Conditions

Four conditions were performed by all participants in a counter-balanced order. These were:

- **C_{pw}** - Position control cursor while walking a path
- **C_{ps}** - Position control cursor while standing still
- **C_{vw}** - Velocity control cursor while walking a path
- **C_{vs}** - Velocity control cursor while standing still

The standing conditions involved the participant standing whilst performing the head tilting operations; the walking conditions involved navigating around a figure-of-eight-route (four cones placed in a rectangular pattern 3 metres by 4 metres, as seen in Figure 6). This walking route was chosen such that participants had to divide their attention between the task and navigating [26].

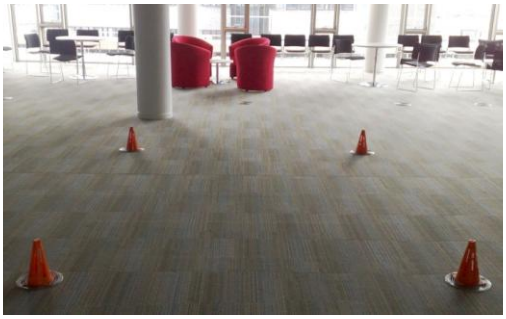

Figure 6. The circuit used for the walking condition. Participants walked around the cones placed in a rectangle (size 4 x 3 metres) in a figure of eight pattern.

Each condition consisted of four blocks of 32 targets with an initial dummy target to start the block. Each block corresponded to a single target separation of 70, 100, 150 or 200 pixels (corresponding to ~19^o, ~25^o, ~37^o and ~50^o in the position control condition) with the blocks being presented in a random order. The 32 targets contained eight (four left-side and four right-side) targets of four different widths of 20, 30, 40, or 50 pixels (corresponding to ~5^o, ~7^o, ~10^o, and ~12^o in the position control condition). In the velocity control condition, 1^o corresponded to a cursor velocity of ~4.7 pixels/second, outside the 5^o dead-zone on either side of the central position (where the cursor did not move). Here, we refer to the four target sizes by their angle, to maintain consistency between notation in position and velocity control conditions.

4.3 Hypotheses

Our hypotheses were:

1. Participants would be significantly more accurate in C_{vw} versus C_{pw}, due to the dead zone of the cursor control and the decreased sensitivity to noise in the velocity control.
2. Participants would be significantly slower and less accurate in C_{pw} and C_{vw} than in C_{ps} and C_{vs} due to the added signal noise from the walking making targeting more difficult.
3. Participants would be significantly faster in C_{pw} and C_{ps} than C_{vw} and C_{vs} respectively due to the limitations applied to the velocity control (namely the lower/upper boundaries for cursor velocity).
4. Consistent target biases will show up in the mobile conditions which will vary as the user's gait phase due to the interactions between walking and interacting with an accelerometer.

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

4.4 Results

4.4.1 Accuracy and Time

Overall mean and standard deviation results for the different conditions are shown in Table 1. Figure 7 and Figure 8 show the accuracy and movement time for each of the different target widths and separations over the four conditions. They show as expected a general trend of higher accuracy as the targets get wider and lower movement time as for smaller separations.

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 main effect of condition was found for both Accuracy ($F = 78.2$, $p < 0.01$) and Movement Time ($F = 291.0$, $p < 0.01$). A *post hoc* Tukey test showed C_{ps} was faster and more accurate than all other conditions ($p < 0.01$), C_{vs} was significantly more accurate than C_{pw} and C_{vw} ($p < 0.01$ in both cases), and significantly faster than C_{vw} ($p = 0.01$).

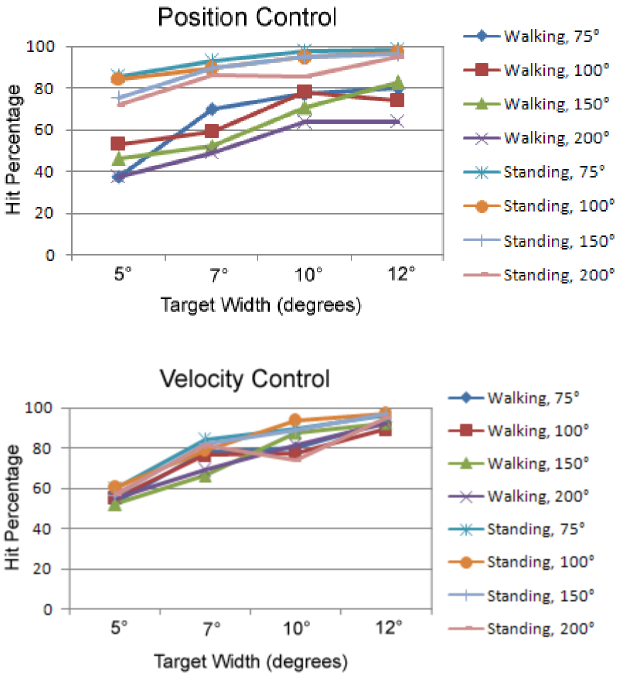

Figure 7. The mean percentage of targets hit for all conditions, for all width-separation combinations.

Table 1. Summary of the mean and standard deviations for the four conditions.

Condition	Mean Accuracy (%)	Mean Movement Time (s)
C_{ps}	89.8 (std dev = 11.0)	1.53 (std dev = 0.51)
C_{pw}	61.9 (std dev = 14.0)	2.19 (std dev = 0.98)
C_{vs}	80.7 (std dev = 16.6)	2.22 (std dev = 0.73)
C_{vw}	74.8 (std dev = 14.1)	2.58 (std dev = 0.91)

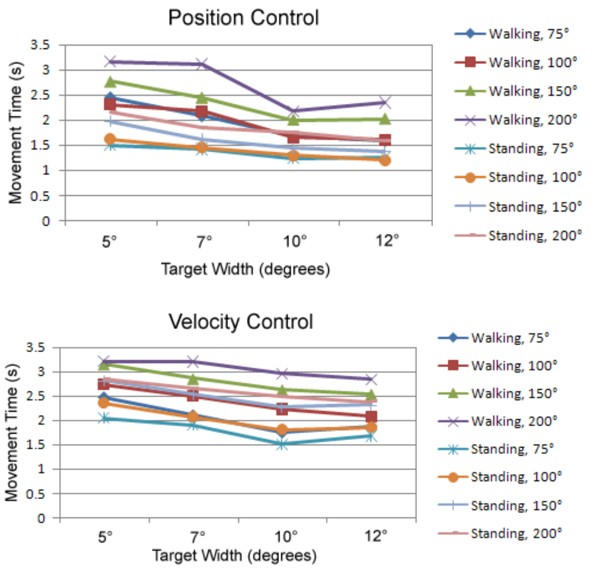

Figure 8. The mean movement time for all conditions, for all width-separation combinations.

4.4.2 *Fitts' Law Results*

Values calculated for ***a*** and ***b*** are shown Table 2 with a Pearson product moment correlation value for each. The correlation coefficient show a strong correlation for C_{ps} only.

Table 2. The values of *a* and *b* for the four conditions. *r* is the Pearson product moment correlation coefficient.

	a	***B***	***r***
C_{pw}	1.29	0.12	0.28
C_{vw}	1.23	0.28	0.45
C_{ps}	0.33	0.51	0.61
C_{vs}	1.52	0.38	0.47

Figure 9 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. A poorer correlation can be seen in the mobile conditions indicating far more variable performance than in the static conditions.

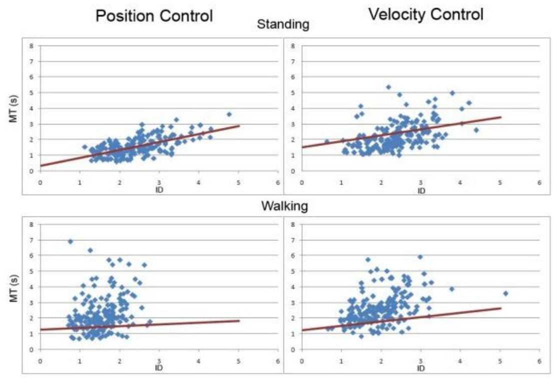

Figure 9. Movement time plotted against Index of Difficulty for all four conditions.

4.4.3 *Throughput*

Using the formula described in Section 3.1, throughput in bits per second was calculated for the four different conditions. Further to this, we draw on results from a previous study with similar methodology examining pointing using wrist rotation (with a position control mechanism) in static (standing) and mobile situations [8]. Table 3 shows the calculated results for throughput of the pointing techniques.

Table 3. The throughput for all four experimental conditions, as well as from a previous wrist pointing study (described in [8]) with static and mobile conditions for comparison.

	Throughput Head-Position	**Throughput Head-Velocity**	**Throughput Wrist-Position**
Static	1.55 bits/s	1.07 bits/s	1.64 bits/s
Mobile	0.69 bits/s	0.76 bits/s	0.97 bits/s

4.4.4 *Skew Results*

Using the data gathered from the accelerometer attached to the user's back and the methods described in Section 3.2.1, we mapped the users recorded phase data to step phase. These data were then combined with the target selection data, such that measures could be made of the effect that different stages of the walking cycle have on the accelerometer input. In Figure 10, targeting behaviour is broken down by phase of gait and target separation and width for position control and velocity control interfaces respectively. Further to this, the gait cycle has been segmented into left step (white background) and right step (shaded background).

Figure 11 shows the distribution of target selections at different stages in the gait cycle plotted as a cloud plot. Lighter, more focussed points indicate concentrations of selections at that horizontal position on the screen at a particular phase. This indicates more consistent targeting behaviour from participants at that phase. Areas that are more blurred indicate more variation in the participants' target selection behaviour. The darker areas indicate target positions where users were less likely to tap during their gait cycle.

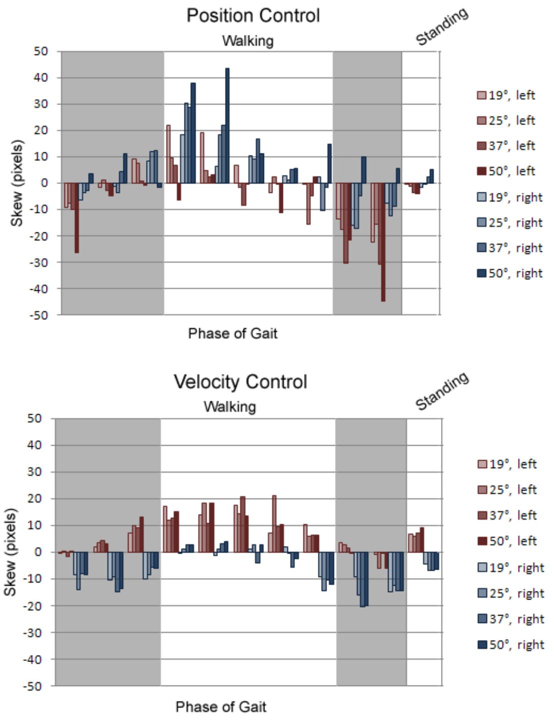

Figure 10. Mean skew in target selections for position (top) and velocity (bottom) control, broken down by phase of gait, for all conditions and target widths/positions (target to left/right of centre). Positive skew indicates a skew to the left of the target, negative skew to the right. The grey area marks the right step section and the white area the left step.

5. DISCUSSION

5.1 Time and Accuracy

There were significant differences found in the accuracy data between the standing and walking conditions using both position and velocity control. Position control did consistently exhibit lower movement time than its velocity counterpart, and was significantly more accurate under the standing condition. However, position control was significantly worse than velocity control when mobile, with performance deteriorating significantly under the walking condition as shown in Figure 7 and Figure 8. Hypothesis 1 and Hypothesis 3 are therefore supported. There is the trade-off with cursor gain in the velocity controlled condition with a low cursor gain giving high accuracy and a longer target time, and high cursor gain giving a potentially faster time to target but lower accuracy. It will be possible for interface designers to shift performance between the two extremes by tuning this gain parameter.

In the static conditions, participants achieved a high level of accuracy (approximately 80% for the 7^{o} and over 90% accuracy for larger targets). This suggests that participants could successfully target using head tilting. The walking conditions were however reported to be far more difficult for all participants who were both slower and far less accurate than in all other conditions for all target width-separation combinations, and particularly so for the position control condition which was significantly poorer than the others. Hypothesis 2 is therefore supported by the data.

It is also interesting to note that the accuracy values for this condition are at best ~80% for the largest targets. All participants commented on the difficulty of the task and expressed low confidence in their performance, frequently noting that when using position control the cursor would oscillate upon each footstep and thus impact target selection effectiveness. Similar results were found in a previous study examining wrist rotation as an input technique in a mobile setting with a position control mechanism [8]. For the velocity control condition, the effect of noise seems to be reduced. This could be due firstly to the inclusion of a dead-zone around the central position making it easier for the participant to stop the cursor, and secondly since it is the velocity that is affected by the noise and not the actual cursor position.

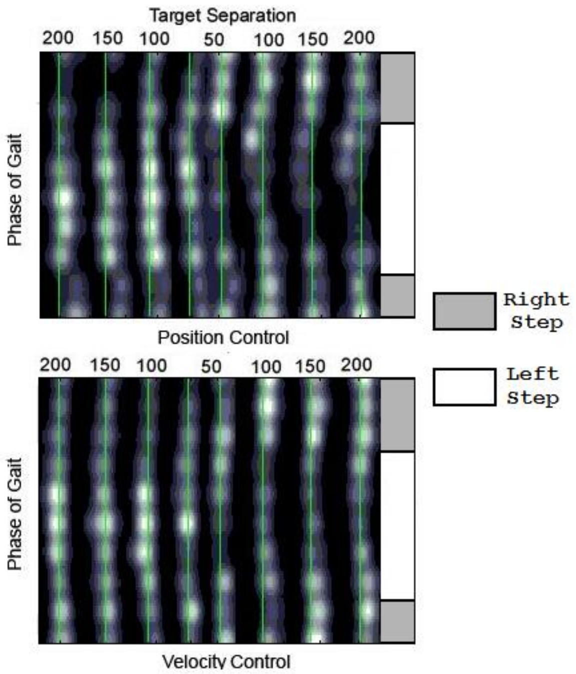

Figure 11. Cloud density plot of target selections, broken down by gait phase and left/right targets for all target separations, for position (top) and velocity (bottom) control. The lighter the area, the more user target events at that position and phase, with more focussed points indicating less variability in selection. The shading on the right indicates a left or right foot down.

From Figure 10, we can see the skew results for all conditions. Examining firstly just the static conditions, we see little skew effect when targeting in the position control mode. In the velocity control condition however, an over-shoot effect can be seen for both the left and the right targets which could have affected the overall accuracy results in the velocity conditions. This could be indicative of the users not predicting the time required to return to the dead-zone and stop the cursor. To resolve this issue, future studies could investigate the incorporation of feedback to give the user some awareness of the current tilt angle. Previous research has successfully investigated the incorporation of tactile feedback to convey tilt angle [22], which could be investigated to increase the accuracy when using a velocity control mechanism.

5.2 Fitts' Law and Throughput

Low correlation with the Fitts' Law data in the mobile conditions suggests far more variable performance by the participants, which is further indicated by the low accuracy results in these conditions. The velocity control conditions both show relatively low correlations with the Fitts' Law model. This is most likely due to the fact that Fitts' law was originally designed with a position control mechanism in place where the movement of an input device is associated with the displacement of a cursor.

The throughput results suggest that the position control mechanism allowed a higher communication rate than the velocity control in the static condition. The value for the position control wrist input condition calculated from the study described in [8] compares similarly with the position control head pointing study in this study. In the mobile condition however, the throughput for the head pointing condition drops to below half of that in the static condition, suggesting a high difficulty level. The drop off in throughput in the wrist pointing condition is less extreme. This could be due to the wrist being more dexterous than the neck and therefore faster and easier to make fine grain movements. Similarly for the velocity control condition, the drop off in throughput is less than that in the position control condition. This is most likely due to the robustness that can be built into the velocity control condition through a tuneable gain parameter and dead-zone.

5.3 Gait Phase

Figure 10 shows a clear interference effect in terms of the motion of users and the eventual selection coordinates of the cursor. Where the phase corresponds to left/right steps, there exists a significant selection skew. For example, with position control targets that required a head tilt against the sway of the user's gait (e.g. right target selection in left step phase) were prone to undershooting, with targets in the direction of the sway prone to some overshooting. There is also a large disparity between results for targets of separation 19°-37°, and that of the furthest away targets at separation 50°, indicating that head tilts toward the extremes were perhaps more error prone or too difficult to achieve. Perhaps the range of head tilts of ±40° was a little large and smaller movements would allow some better control at the extremes, at the expense of increased sensitivity and noise. Figure 10 suggests Hypothesis 4 can be supported. Velocity control exhibits significantly less skew, with less undershooting of targets that were against the sway of the user's gait. However, velocity control also exhibited significant overshooting behaviour when the phase of sway and target head tilt direction were aligned. Nonetheless, indications are that velocity control suffers less from motion interference than position control, confirming Hypothesis 1.

Of significant interest is the behaviour indicated by Figure 11, where there appears to be some tendency for right-situated targets to be selected whilst in the right-foot phase of motion, with left-situated targets selected in the left-foot phase of motion. This is indicated by the brighter spots for the right most targets in the right step regions of the image and the brighter spots occurring for the left most targets in the left step region of the image. This would indicate that there is at least some amount of effort by users to counteract the skew effect of motion, particularly in the position control condition. Also of note is the dispersal of target selections in the velocity condition contrasted against the position condition, with position control showing greater dispersal around the target centre line, indicating more consistent targeting bias at different phases in the participants' gait.

6. Conclusions

This paper has motivated non-hand-based gesturing as a research area for mobile interaction. By investigating different areas of the body for input, users' hands can be freed up for other tasks such as carrying bags or holding on to children. We have described a study using head tilt angle – estimated using an accelerometer attached to a hat – as an input technique in both static and mobile contexts. Results show that in static contexts, the position and velocity control mechanism tested allowed the user to hit targets of 30 pixels (7^o in the position control conditions) with a high degree of accuracy with the position control mechanism allowing faster targeting. The study also showed that there was a marked drop off in performance (significantly reduced accuracy and increased movement time) when the participants performed the same task on the move using the position control cursor mechanism. This drop off was less noticeable in the velocity control condition, with the position control performing significantly less accurately than velocity control in mobile conditions. There may be a hybrid solution where we change the control method depending on users' movement. If they are static then we could provide position control and when walking (as detected by our algorithms) we could switch to velocity control with clear feedback of tilt angle. The question would be if this change causes confusion. The other approach would be to try and improve the performance of velocity control in static settings to increase it to the level of position control.

This study builds on our previous work, using an instrumented usability analysis approach for the study. By taking this approach, we were able to infer details about the moment-to-moment user actions that would not be possible with traditional techniques. This study demonstrated the effect of gait on mobile interactions when using an accelerometer as input. In this case, the oscillations generated by a user walking had a predictable effect on the acceleration measured at the head which varied with the user's gait cycle and whether it was a right or left step. By taking account of the users gait phase, we can potentially use this information to filter out the unwanted noise generated by the walking. Future work will also look to use the gait information to improve interfaces designed to be used on the move. For example, we will explore rhythmic gestures designed to synchronise with the users step behaviour which could potentially lead to easier to perform gestures while on the move. A further strand of our work is currently taking the first steps to explore the important issues of the social acceptability of these gestures [29]. Future work will continue to examine viable areas of the body that could provide methods of allowing hands-free interaction with a mobile device. This will eventually be extended with tactile and audio feedback to allow eyes-free interaction. The study above has demonstrated head tilt can successfully be used as a pointing technique in a mobile setting. Future work will build on these results to work towards more subtle and discreet hands-free mobile interactions.

7. ACKNOWLEDGMENTS

This work was funded by EPSRC grant EP/F023405 – Gestural and Audio Interactions for Mobile Environments (GAIME), and Nokia.

8. REFERENCES

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

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

[3] Barreto, A., Scargle, S., and Adjouadi, M. *Hands-off human-computer interfaces for individuals with severe motor disabilities in Proceedings of the HCI International '99 (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Communication, Cooperation, and Application Design-Volume 2*. 1999 ACM Press, 970 - 974.

[4] Brewster, S.A., *Overcoming the lack of screen space on mobile computers.* Personal and Ubiquitous Computing, 2002. **6**(3): p. 188–205.

[5] Brewster, S.A., Lumsden, J., Bell, M., Hall, M., and Tasker, S. *Multimodal 'Eyes-Free' Interaction Techniques for Wearable Devices. In Proceedings of ACM CHI*. 2003 Florida, USA: ACM, 463-480.

[6] Costanza, E., Inverso, S.A., and Allen, R., *Toward subtle intimate interfaces for mobile devices using an EMG controller, in Proceedings of the ACM CHI*. 2005 Portland, Oregon, USA. p. 481 - 489.

[7] 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.

[8] Crossan, A., Williamson, J., Brewster, S., and Murray-Smith, R., *Wrist Rotation for Interaction in Mobile Contexts, in the Proceedings of Mobile HCI*. 2008, ACM Press: Amsterdam, Netherlands.

[9] Eslambolchilar, P. and Murray-Smith, R., *Control centric approach in designing scrolling and zooming user interfaces.* International Journal of Human-Computer Studies, 2008. **66**(12): p. 838-856.

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

[11] Hinckley, K., Pierce, J., and Horvitz, E. *Sensing Techniques for Mobile Interaction*. *Proceedings of ACM UIST*. 200091-100.

[12] Hoggan, E., Crossan, A., Brewster, S., and Kaaresoja, T., *Audio or Tactile Feedback: Which Modality When?*, in *Accepted for publication in ACM CHI*. 2009 Boston, USA.

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

[14] Linjama, J. and Kaaresoja, T. *Novel, minimalist haptic gesture interaction for mobile devices. in Proceedings of the third Nordic conference on Human-Computer Interaction*. 2004. Tampere, Finland: ACM International Conference Proceeding Series, 457-458.

[15] LoPresti, E., Brienza, D.M., Angelo, J., Gilbertson, L., and Sakai, J. *Neck range of motion and use of computer head controls. in Proceedings of the fourth international ACM conference on Assistive technologies ASSETS*. 2000. Arlington, Virginia, USA: ACM Press, 121-128.

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

[17] MacKenzie, I.S. and Isokoski, P., *Fitts' throughput and the speed-accuracy tradeoff.*, in *Proceedings of the ACM Conference on Human Factors in Computing Systems – CHI 2008*. 2008, New York: ACM.

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

[19] Mäntyjärvi, J., Kela, J., Korpipää, P., and Kallio, S. *Enabling fast and effortless customisation in accelerometer based gesture interaction. in Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia*. 2004: ACM, 25 - 31.

[20] Marentakis, G.N. and Brewster, S.A. *Effects of Feedback, Mobility and Index of Difficulty on Deictic Spatial Audio Target Acquisition in the Horizontal Plane. In Proceedings of ACM CHI*. 2006. Montreal, Canada: ACM Press Addison-Wesley, 359-368.

[21] Morency, L. and Darrell, T. *Head gesture recognition in intelligent interfaces: the role of context in improving recognition. In Proceedings of the 11th International Conference on intelligent User interfaces*. 2006. Sydney, Australia: ACM, 32-38.

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

[23] Oakley, I. and Park, J. *A motion-based marking menu system. in Extended Abstracts of ACM CHI*. 2007. San Jose, CA: ACM.

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

[25] Pikovsky, A., Rosenblum, M., and Kurths, J., *Synchronization: A universal concept in nonlinear sciences*. 2001: Cambridge University Press.

[26] Pirhonen, A., Brewster, S.A., and Holguin, C. *Gestural and Audio Metaphors as a Means of Control for Mobile Devices. In Proceedings of ACM CHI*. 2002. Minneapolis: ACM Press Addison-Wesley, 291-198.

[27] Rekimoto, J. *Tilting Operations for Small Screen Interfaces. In Proceedings of UIST*. 1996167-168.

[28] Rekimoto, J. *Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. In Proceedings of Fifth International Symposium on Wearable Computers (ISWC '01)*. 2001.

[29] Rico, J. and Brewster, S. *Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces. to appear in extended proceedings of ACM Mobile HCI*. 2009. Bonn, Germany.

[30] Ronkainen, S., Häkkilä, J., Kaleva, S., Colley, A., and Linjama, J. *Tap input as an embedded interaction method for mobile devices. in Proceedings of the 1st international conference on Tangible and embedded interaction*. 2007. Baton Rouge, Louisiana: ACM, 263 - 270.

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

[32] Williamson, J., Murray-Smith, R., and Paisley, A.M. *Dyanmics and Probabilistic Text Entry. in Proceedings of the Hamilton Summer School on Switching and Learning in Feedback systems*. 2005: Springer-Verlag, Lecture Notes in Computing Science, Vol. 3355, 333-342

[33] Zhai, S. and Kristensson, P. *Shorthand writing on stylus keyboard. in Proceedings of ACM CHI*. 2003. Florida, USA: ACM, 97 - 104.

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