Tap ‘n’ Shake: Gesture-Based Smartwatch-Smartphone Communications System

Steven McGuckin
School of Computing Science
University of Glasgow
mcguckin.steven@gmail.com

Soumyadeb Chowdhury
Infocomm Technology
Singapore Institute of Technology
Soum.chowdhury@singaporetech.edu.sg

Lewis Mackenzie
School of Computing Science
University of Glasgow
Lewis.Mackenzie@glasgow.ac.uk

ABSTRACT
Smartwatches have recently seen a surge in popularity, and the new technology presents a number of interesting opportunities and challenges, many of which have not been adequately dealt with by existing applications. Current smartwatch messaging systems fail to adequately address the problem of smartwatches requiring two-handed interactions. This paper presents *Tap ‘n’ Shake*, a novel gesture-based messaging system for Android smartwatches and smartphones addressing the problem of two-handed interactions by utilising various motion-gestures within the applications. The results of a user evaluation carried out with sixteen subjects demonstrated the usefulness and usability of using gestures over two-handed interactions for smartwatches. Additionally, the study provides insight into the types of gestures that subjects preferred to use for various actions in a smartwatch-smartphone messaging system.

Author Keywords
Smartwatch – smartphone communication; motion-gestures; one-handed interaction

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
The recent surge in popularity of smartwatches presents a number of opportunities to enhance smartwatch messaging applications, which are yet to be explored, either in research, or commercially. The existing commercial smartwatch messaging applications, such as Apple’s ‘Real Touch Messaging’ (Tweedie, 2016; El Khoury, 2016; Mukherjee , 2016) are primarily used as extensions to applications on smartphones, where the user is presented with quick access to certain features of the smartphone application using only the smartwatch.

The existing smartwatch messaging applications mostly require two-handed interactions to carry out tasks, with little consensus on the best input method for sending messages using smaller smartwatch interface, which is a persistent mobile HCI problem. To mitigate the problem of two-handed interactions with the smartwatch, *Tap ‘n’ Shake* makes use of motion-gestures for control, and voice input to construct textual and audio messages. The paper reports a study conducted with 16 subjects to explore the effectiveness and usefulness of *Tap ‘n’ Shake.*

LITERATURE SURVEY
Extensive research has been reported in the context of using motion-gestures as an input modality for smartphone devices. However, existing research looking at using motion-gestures as a means to interact with a smartwatch, especially in the context of smartwatch-phone messaging applications is sparse. The research reported in (Rico and Brewster, 2010) looked at the social acceptability of various gestures using mobile devices, and found that users preferred to use subtle gestures which did not draw unnecessary attention to themselves. This includes gestures such as subtly shaking, tapping, or rotating the device. Gestures that were the same as, or similar to, ones that are used in everyday life e.g. shaking the device like shaking a bottle, were also shown to be acceptable.

Ruiz et al., 2011 focused on allowing end-users to design their own gestures for various tasks on a mobile device, and found that most users expected gestures to be natural, and to align with the tasks that were carried out. For example, the action of scrolling left or right should be done by moving the device left or right, and zooming in or out should be carried out by moving the device towards or away from the user along the Z-plane. Additional findings from this research also included that users responded positively to using motion gestures on a mobile handheld device, and highlighted the problem of “false-positive” gesture recognitions, when dealing with subtle gestures.

The problem of “false-positive” gesture recognitions on handheld devices was addressed in (Ruiz and Li, 2011) with ‘*DoubleFlip*’ which proposed an effective delimiter for motion gesture input. Similarly, ‘Wrist Rotate’ proposed in (Kerber et al., 2015) offers an effective motion gesture delimiter for wrist-worn devices. Both ‘Double Flip’ and ‘Wrist Rotate’ proposed that motion gestures which utilise a rotational movement back and forth around the Y and X axes of the handheld and wrist-worn devices, respectively, are the most effective to use as delimiters between non-relevant movement and the relevant gesture input.

TAP ‘N’ SHAKE
Tap ‘n’ Shake is a novel gesture-based messaging system which allows users to send textual and audio messages between smartphone and smartwatch devices, primarily using motion gestures. The use of voice input to create messages, and motion gestures to send them, mitigates the problem of smartwatches requiring two-handed interactions, and also overcomes the problem of dealing with text entry on a device with a reduced form factor. This could be particularly useful in situations where the user’s ability to use both hands, or spend time typing on a small keyboard, is either restricted or impaired.

Smartwatch Application

The smartwatch application comprises of two main features: 1) *Send audio message to smartphone*, and 2) *Send voice-to-text message to smartphone*. These features are organised into a horizontally scrollable menu which the user can navigate through and select by tapping and swiping, or - *crucially for the smartwatch device* – using motion gestures. The full mapping of motion gestures to actions in the smartwatch application is shown in Table 1.

id	Actions	Gesture	Description
R1	Navigate main menu, and exit message screen.		A Y-axis rotation, flicking the wrist back towards the forearm.
R2	Select a menu item, and start/stop recording		X-axis rotation, twisting the wrist towards the supine position and then back again.
R3	Send message.		Z-axis rotation, shaking the wrist back and forth.

Table 1. Smartwatch motion gesture mappings.

Receipt of a message on the smartwatch application creates a full screen notification displaying the textual message contents, and the time the message was received. A user is able to dismiss the message using the Y-axis motion gesture shown R1-Table1, or quickly reply to the message using the Z-axis motion gesture shown in R3-Table1.Thus a user can construct, read, and respond to messages, to and from the smartphone application, using only one hand. Android's Speech Recognizer [Speech Recognizer, 2016] is used to record the message and translate it into text.

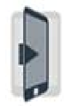
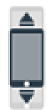
Handheld Application

The handheld application consists of three main features: 1) *Send text message to smartwatch*, 2) *Send Voice-to-Text message to smartwatch*, and 3) *An inbox for received messages*. Similar to the smartwatch application, features are organised into a horizontally scrollable menu. The full mapping of motion gestures to actions in the smartphone application is shown in Table 2. A watch icon in the action bar at the top of the smart phone screen notifies the status of the smartwatch device – online or offline. If the smartwatch is offline, messages sent from the smartphone will be stored on the message server until the smartwatch is online, whereby the messages are delivered. In both the applications, labelled icons are displayed at the bottom of each screen to inform the user which motion gestures they can currently use to carry out one or more action.

DESIGN

The motion gestures were selected considering their subtlety, the perceived ease of being carried out, and drawn from (Rico and Brewster, 2010). A pilot study was conducted (where a set of 12 candidate gestures restricted to linear movements on each of the devices' XYZ axes were demonstrated) with a number of subjects (mainly Computing Science students different from the ones who took part in our user evaluation). The study helped to define a refined set of motion gestures (Table 1 and 2), which were considered easier and more enjoyable to use. The easiest gestures to perform, as reported in the pilot study, were the simple rotational gestures R1, R2, H1 and H2. These gestures required the minimal amount of effort out of all the candidate gestures, i.e. only a small rotation of the wrist was required, and were therefore deemed to be most suitable for common actions within the applications, such as navigating and selecting items from the main menus, to reduce the strain on the user after repeated use.

Table 2. Smartphone motion gesture mappings.

id	Actions	Gesture	Description
H1	Navigate main menu.		A Y axis rotation, twisting the device left and right around its vertical axis.
H2	Select menu item, and exit message screen.		An X axis rotation, flicking the device back or forward around its horizontal axis.
H3	Start recording, send message, and quickly reply to message.		A Y axis shake, up and down along its vertical axis.

Vertical or horizontal menus enabling important features to be encapsulated into various sections were favoured (Android Wear, 2016; Neilsen, 2016), and motion gestures were re-used for different features such as sending audio or textual messages. Encapsulated sections can be accessed using one motion gesture to navigate through them, and another to select the specific feature. This is likely to reduce the need for having gestures mapped one-to-one with the application features, thereby using only the minimal amount of motion gestures per screen, and facilitating the future expansion - as more features can simply be encapsulated in their own section and added to the menu. A horizontal scrolling menu, where only one menu item is shown on screen at a time was preferred for both the smartwatch and smartphone applications, as it allows larger icons for each menu item to be displayed. Additionally, by using different motion gestures for scrolling through the main menu and sending messages, the menu can act as a natural delimiter between the highly consequential action of sending a message, and the less significant action of scrolling.

Motion Gesture Detection

Motion gesture detection was implemented by modifying the 'Shoogle For Yer Photos', 2016, open-source algorithm. This algorithm was used as the basis for linear shaking and rotational gestures around each individual axis, resulting in a library of motion gestures, which were mapped to various actions throughout the applications. A limitation when trying to isolate each rotation or movement on each axis, was that when the device is moved back and forth, the movement or rotation is often not restricted to one axis, because of the unsteady nature of human arm and hand. For example, it is relatively hard to hold the device in an upright position with the screen facing oneself, and move it up and down the Y-axis in a perfectly straight line, while avoiding any kind of acceleration on the X-axis. This meant that a false-positive Y-axis gesture detection would often occur, while an X-axis shaking was being carried out, and vice versa. To mitigate this effect, the positive and negative threshold differential was increased so that only substantial changes in each direction for the particular axis movement were detected. However, the result of increasing the threshold differential meant that to detect a movement on a particular axis, the user would have to make larger, or harder movements, which may be unsuitable. This problem was remedied by introducing separate bounds for the axes on which movement should not be detected. This allowed a greater degree of control over how precise a gesture had to be on a particular axis.

EVALUATION

A three-stage user evaluation was carried out with sixteen subjects (age range 21-50, 9 male and 7 female). The subjects of the evaluation came from a range of different backgrounds as shown in Table 3, without any prior experience with android smartwatches.

# Subjects	Occupation
2	Retail Worker
2	Carer
2	Student
1 of each	Bus driver, Charity worker ,Builder, Nurse, Graphics designer, Oil worker, engineer, Pipe fitter, Bank manager, Web developer

Table 3. Subject Demographics

However, all the subjects were well versed with android smartphones. The study was ethically approved by the Institution Review Board. The three stages are as follows:

Introduction – All the subjects were invited to our laboratory (on different days) to take part in the experiment. They were presented with the different features of the smartwatch and smartphone application, and a demonstration for each gesture was shown by one of the authors. Additionally, the principle objectives of the experiment was also clearly explained by the authors. All the subjects were given 10 minutes to familiarise themselves with the gestures of the smartwatch to perform different actions in the messaging application.

A think-aloud protocol was used to gain some initial feedback, i.e. understand the problems faced by the subjects, and provide further clarification, as applicable. It was made clear that at this stage that we are evaluating the usability of different gestures, and the study does not examine the users' expertise to use gestures. A Nexus 5 android smartphone and Motorola Moto360 smartwatch were used to conduct the experiment. A hardware constraint of the Moto360 smartwatch is the absence of a speaker, to playback recorded messages. Hence, our application did not support a gesture to playback messages on the smartwatch.

Action	Gesture Set 1	Gesture Set 2
Navigate main menu	Y axis rotation	X axis rotation
Select menu item	X axis rotation	Y axis rotation
Start/Stop recording	X axis rotation	Z axis rotation
Send message	Z axis rotation	Z axis rotation

Table 4. Gesture sets for the smartwatch application

Action	Gesture Set 1	Gesture Set 2
Navigate main menu	Y axis rotation	X axis rotation
Select menu item	X axis rotation	Y axis rotation
Start record/Send message	Y axis shake	Z axis shake

Table 5. Gesture sets for the smartphone application

Experiment- An experiment was carried out to test the performance of two different sets of gestures (Table 4 and 5) mapped to various actions in the devices. A protocol was used, where half the subjects were assigned to the GS1, and the other half to the GS2, randomly. Where two subjects shared the same occupation, they were allocated to different gesture sets. The tasks subjects carried out on the smartphone were: (1) *Send a one-character text message;* (2) *Send the word "hello" using the voice-to-text feature;* (3) *Open the second message in the inbox.* The tasks subjects carried out on the smartwatch were: (1) *Send the word "hello" using the audio message feature*; (2) *Send the word "hello" using the voice-to-text feature*; (3) *Open the second message in the inbox.* The average time to complete each task using the different gesture sets were compared.

Questionnaire- After the experiment, subjects used the opposing gesture set which was not used in the experiment. The aim was to let them make comparisons between, and state a preference for, each gesture set. Additionally, semi-structured interviews were conducted in this stage. The semi-structured interviews were conducted to help understand the effectiveness and efficiency of using different gestures in the smartwatch application. Additionally, follow up questions were asked based upon the feedback in the think-aloud sessions, to better understand the issues flagged up by the subjects during the sessions, and gain understanding into different ways to improve the gestures, to better cater to the needs of the smartwatch users in the future.

Results

The results of the experiment for the smartphone, shown in Figure 1, and the smartwatch shown in Figure 2, demonstrated that the mean time taken to carry out each task with GS1, was slightly lower than with GS2 for both the devices. The task for sending a voice-to-text message with the smartphone device using GS1 was performed significantly quicker ($p < 0.05$) than with GS2. This is likely attributed to subjects having difficulty in carrying out the Z-axis shake used by GS2 to begin recording the message and send it. This shake is performed by moving the device towards and away from the user, as opposed to the Y-axis shaking gesture used by GS1, which required the user to shake the device up and down.

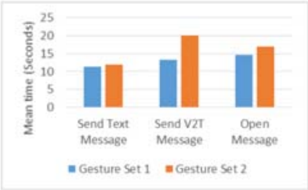


Figure 1: Smartphone

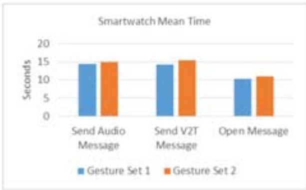


Figure 2: Smartwatch

The results also showed that 94% of the respondents opted for GS1 on the smartphone, and 69% preferred GS1 on the smartwatch. GS2 required rotating the device orthogonally to the direction of navigation, which was reported to being confusing. Another reason for favouring GS1 over GS2 on the smartwatch, was that the X-axis rotational gesture used in GS1 to start and stop audio recordings, was easier to carry out than the Z-axis rotational gesture in GS2. However, Z-axis rotation was favoured for sending messages from smartwatch.

The questionnaire also asked the subjects to choose their preferred input method, between tapping, using motion gestures, or both, for each of the devices. For the smartphone, 75% of the subjects responded that they preferred to use both tapping and motion gestures, followed by 19% favouring motion gestures, and just 6% choosing tapping. Qualitative feedback suggested that subjects preferred to use motion gestures to navigate and select items on the main menu, and also for the actions, such as recording and sending a voice-to-text message. Subjects preferred to use tapping gestures to access and delete messages in the inbox, rather than using motion gestures to scroll down through the vertical list of messages, which the subjects responded was slow and cumbersome compared to just tapping on the inbox items.

For the smartwatch, 68% of the responses were in favour of using only motion gestures, and 32% preferred to use both tapping and motion gestures. None of the subjects opted to use tapping. The reasons for these preferences were that they “preferred to use the smartwatch with one hand”, and that it was “fun” and “cool” to control the smartwatch using motion gestures. The subjects found the system (using gestures) to be “fun”, “different”, “cool”, “usable”, “and useful”. They felt that it will be “simple to use once an individual get used to it”. Other positive responses were that the voice input was easier than texting, and they preferred the big icons on the interface. The negative feedback about the application included that opening and playing audio messages was cumbersome using motion gestures on the smartphone device, and they preferred to open messages by simply tapping.

Limitations

The evaluation of the system was conducted with only sixteen subjects new to smartwatch usage, albeit from a wide range of backgrounds, and all the subjects were given only ten minutes to familiarise themselves with the system. Ideally, more subjects could have been used, each receiving a smartwatch to use and familiarise themselves with for several days, weeks, or even months, before the evaluation was carried out to avoid novelty effects. Differences in the difficulty of each subject to carry out various actions using the motion gestures, could be attributed to some users simply pick up new technologies quicker than others, no matter if they are from a more technical background or not. Mitigating this would require a longer familiarisation time to get used to the technology before conducting the study, to avoid the confounding effects. The study was conducted with all English speaking individuals, and the speech to text conversion was accurate in all the cases. However, the accuracy of the conversions is likely to depend upon many factors such as pronunciation, dialect, vocabulary etc., which ought to differ across individuals. Hence, future studies will consider this aspect, so as to improve the user experience.

CONCLUSION

This paper presented Tap ‘n Shake, a motion-gesture based messaging system for Android and Android Wear (smartwatch) devices. The system leverages on a number of motion gestures which were received favourably in our user evaluation to allow the smartphone and smartwatch applications to be operated entirely using one-handed interactions. Our evaluation also showed that such a system is both useful and usable, and provided insight into how users prefer to interact with a smartwatch-smartphone messaging system, including preferences for different motion gestures to carry out various actions.

Our broader goal is to develop a gesture library based on the study reported in paper for a number of use-cases and scenarios, but not limited to: various actions in the context of responding to notifications and browsing through list of notifications. Such a study will be conducted under real-life conditions, i.e. subjects will be required to make use of the gestures for an extended period of time. This will help to understand the usability of different gestures, which are used for different applications, in an uncontrolled environment. We also aim to conduct a study with elderly users to explore their perception about using gestures, understand the issues they are likely to face while using motion gestures, and devise suitable design solutions, i.e. gesture design for elderly smartwatch users. Additionally, we also aim to examine the effectiveness of gestures, with different watch operating system, especially comparing the Android and Apple devices.

REFERENCES

Tweedie, S. ’Real Touch’ Messaging On The Apple Watch Is Like A More Intimate Snapchat For Your Wrist. URL:http://www.businessinsider.com/what-is-apple-watch-real-touch-messaging-2014-9?IR=T, 2016. [Online; accessed 15-June-2016].

El Khoury, R. Messages For Android Wear Lets You Type SMS And Hangouts Texts. URL:http://www.androidpolice.com/2015/09/16/messages-for-android-wearlets-you-type-sms-and-hangouts-texts-courtesy-of-the-team-who-alsobrought-a-mail-clientbrowser-and-youtube-to-your-wrist, 2015. [Online; accessed 15-June-2016].

Mukherjee, S. Avaamo brings its business messaging App to Apple Watch. URL:http://www.techmagnifier.com/news/avaamo-brings-its-business-messagingapp-to-apple-watch, 2015. [Online; accessed 15-June-2016].

Rico, J., and Brewster, S. Usable gestures for mobile interfaces: evaluating social acceptability. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 887–896. ACM, 2010.

Ruiz, J., Li, Y., and Lank, E. User-defined motion gestures for mobile interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 197–206. ACM, 2011.

Ruiz, J., and Li, Y. Doubleflip: a motion gesture delimiter for mobile interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 2717–2720. ACM, 2011.

Kerber, F., Schardt, P. and Löchtefeld, M. WristRotate: a personalized motion gesture delimiter for wrist-worn devices. In Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia, pages 218-222. ACM, 2015.

Android Speech Recogniser, URL:https://developer.android.com/reference/android/speech/SpeechRecognizer.html.[Online accessed 15-June-2016].

Design Principles for Android Wear URL:http://developer.android.com/design/wear/principles.html. [Online accessed 23-June -2016].

J. Neilsen. 10 Heuristics for User Interface Design. URL: https://www.nngroup.com/articles/tenusability-heuristics/ ,1995. [Online; accessed 23-June-2016].

Shoogle-For-Yer-Photos-Android URL:https://github.com/JMB-Technology-Limited/Shoogle-For-Yer-Photos-Android .[Online; accessed 15-June-2016].