Eyes-Free interaction method in SmartWatch

Kyu Yeol Paik
Auckland University
403S-499
kpai008@aucklanduni.ac.nz

ABSTRACT

This review displays the concept of Eyes free interaction. Eyes Free interaction is a method of controlling the mobile devices without having to look at the device. This interaction type is being researched to allow user to spend less visual focus on the device allowing user to focus their visual needs where it is important (for example, when user is driving). It also focuses on minimizing the concentration and effort needed for the user to communicate with the device.

This review will cover the different technologies and concepts being approached to enhance the eyes free interaction. It will compare different methods in each of its relative fields and show the advantages and disadvantages. Then the review will further go on to identify the missing elements in current technology or research which are necessary to improve the interaction.

From the review, the most suitable input method in general environment was gesture based input and for output, sound based feedback. More research needs to undertake for gesture based input recognition and study of mixed output methods will be useful in further improving the eyes free interaction.

INTRODUCTION & MOTIVATION

In the field of HCI, visual modality has long dominated the HCI research. Despite this, visual means of interaction is not always desirable or suitable due to number of factors.

1) Competition for visual attention in mobile scenarios [1]. For example, in times when users are driving, their visual attention needs to be focused on the task and not be distracted. Dividing their attention in these situations can be distracting and even dangerous[2].

2) Limitations in screen size. Watches in general, are smaller than the cell-phones used today. Hence they provide even smaller interface to work on. Attempting to navigate via touch mechanism can lack in accuracy. The visual feedbacks from such device maybe hard to interpret for the users with visual disabilities.

3) Reduction of battery life. Currently, devices such as mobile phones are suffering greatly from its limited lifespan. Part of the device that consumes most power is the screen. By operating every function via the screen, the battery is lost in short amount of time which results in device's true mobility from being lost.

4) Inconvenience. There are cases when it is inconvenient to fetch the device and to look at the screen for extended amount of times [1]

To overcome these problems, I have studied various fields of interaction methods to overcome these methods. One particularly interesting mechanism was the "eyes-free" interaction with the devices.

This review will talk about the history of "eyes-free" interaction, current status of this field and its future.

BACKGROUND

User interfaces that are currently in mobile devices derive from the rich graphical interaction paradigm developed from desktop computer is considerably inappropriate for a device with very limited screen space [3]. Not only is the system sometimes very hard to control, they take large amount of visual focus away from the user. The effect of this has been recognized by several governments and the law to prohibit people from using their mobile devices while driving has become a common sight.

This shows the need for different interaction method to be found. In this review, it focuses on the eyes-free interaction.

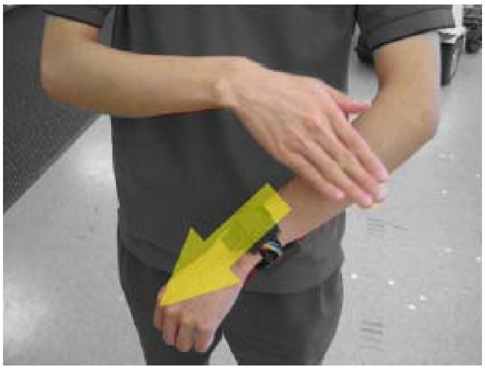
The term eyes-free has been used for over decades as a descriptive phrase denoting a UI with little or no graphical component but now it is emerging as a specialized interaction method with unique feature and qualities. [4]

INPUT MODALITY

There are 3 main ways in which input could be given to the watch. These are via gestures, sound and touch. Since touch based usually requires visual means of interaction, it will not be covered in this review.

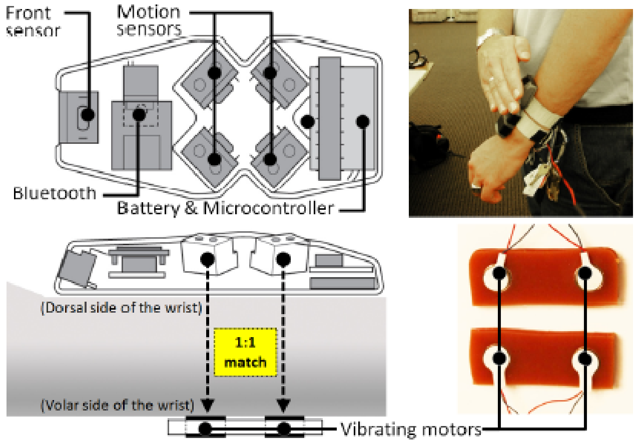
Gesture

Gesture based interaction is performed via placing sensors in the device to detect movement of the watch. In a product called Gesture Watch [5] , the product uses an array of four SHARP GP2Y0D340K proximity sensors that are arranged facing up in cross shape to allow watch to detect variety of gestures.



Gesture based input provides variety of ways of implementing inputs however, the technology right now is limited in detecting small movements hence user must provide device with a suitable sized gestures. This makes the interaction feel unnatural and consumes user's focus. Continuously performing gesture may also cause user to become fatigued. Hence research in the means of software or hardware is needed in order to obtain more natural gesture based input.

Another issue that rises with gesture based input is that device may interpret everyday movement as a gesture. To solve this problem, AirTouch, an upgraded version of GestureWatch experimented a way in which user gives an input to the device after the gesture has been performed. [6] This however requires action of two hands which is not the desired situation in many cases. Hence the technology mentioned above will need to be improved to achieve more natural means of interaction.



Voice

Voice (sound) based input method is another type of input method that can be used to interact with the device. Simple procedure of this mechanism is for the device to hear the user's voice and interpret the order and perform the task. However this method has numerous issues and is currently an unpopular option.

One of the reasons is that the human's voices all differ from one another. Depending on their age and the country they were born, users all have different accent and tone. Due to this, it is extremely difficult to correctly interpret the user's message. Users are also likely to become frustrated when the given function does not work as expected [7]. Hence further researches are needed.

Another problem is similar to the problem stated above but it is affected by the environment. Right now if there are multiple people talking at once, the watch will not be able to figure out which person's input to take and interpret. It will likely to interpret all of them. This is a rather complex problem. One way to ease the problem maybe is to give a sign to the device before the order as it is done in AirTouch. However this also has the same problem of having to use two hands.

OUTPUT MODALITY

When considering models for feedback it is important to consider the speed at which information can be presented and accurately processed is important.

There are 2 main means of eyes-free feedback. These are sound and haptic feedback.

Sound

The general consensus of sound feedback is that when user gives an input to the device, it performs the task and notify user of what has been done.

In sound, there are 3 main ways that have been researched. These are audio icons [8], structured audio messages composed of various pitch and rhythm [9] and speech [1].

Auditory Icons

The concept of Auditory Icons rose from Gaver [8] who introduced a concept of 'Everyday Listening' which consisted of "instead of mapping information to sounds, we can map information to events". Auditory icon utilizes metaphors to relate them to virtual. If a good link can be found between the source of sound and source of data, the meaning of the Auditory icon can be easily remembered. For example, a door slamming shut could represent user logging out of the system.

The downside of auditory icon is that it lacks flexibility since metaphoric mapping does not always exist. There is also possibility of sounds getting mixed with the actual sound from the real world.

Earcons

Blattner et al defines earcons as "nonverbal audio messages used in the user-computer interface to provide information to the user about some computer objects, operation or interaction" [10]. Earcons can be composed of abstract musical tones or short rhythmic sequences of varying pitches and variable intensities. Hence earcons is significantly more flexible then the Auditory icons.

The disadvantages is that user have to learn and remember what it all represents as there are no natural connection between the given sounds and the earcons in most cases.

Speech

Speech is an output which gives feedback in a human language. The benefit of this is that user this approaches user as a very natural form of communication so there is no need for user to go through a learning process. Another advantage is that variety of information can be presented to the user and they can easily understand the message.

The downside of this system is that it may take considerably long time to pass on the message. It also takes more concentration to interpret these messages.

Haptic

Haptic feedback mainly consists of tactile feedback. Tactile feedback available from smartwatch is usually vibration. By diversifying strength and rhythm of this, device is capable of giving a feedback to the user. This type of feedback is capable of transferring simple types of messages but it has a very difficult time in passing on complex information.

LEARNING ISSUES

One of the problems faced with Eye-free devices is that user must learn how to operate the given interface. Unlike graphical interface where users can explore around the menus, the user cannot do this in the eye-free environment. Hence it approaches user as rather daunting task to adapt to this new given interface.

The proposed solution to this maybe is to allow user to go through a steps of transitions. For example user could start with a normal graphical user interface but they will be given one of the output modality on top of the visual feedback. Through these transitions, user does not have to go blind all of a sudden but take a step by step course in getting introduced to eyes-free interaction. The issue with this solution is that it will consume a lot of time for developers to design and research this type of system.

DISCUSSION

The main goal of Eyes-free interaction is to relieve user from dividing their visual focus and give them effortless interaction with the device. Eyes-free interaction mainly needs to make two choices from input and output modalities.

For input modalities, I believe that gesture based interaction is the input method to go for. The environment in which the eyes-free interaction is really necessary will not be a perfect environment where there are no sounds to interfere with the voice input. Gesture based interaction in most cases does not get interfered by the versatile environment in which the system will be used often. The accuracy of the gesture based methods is also more reliable then the voice based input. Gesture based interaction does suffer from having limited amount of inputs it can give but the other quality attributes outweigh this issue.

For output modalities, each of three modalities has their advantages and disadvantages excelling in their own particular fields. Both haptic and sound have their distinctive advantages hence currently, there are researches trying to mix the output modalities. It has been proven before that using both haptic and sound as a feedback was generally more favored by the users with a mobile phone [11]. It is presumed that this type of feedback will be more favored by the majority in eyes-free interaction method as well. However, further research will need to take place to verify this.

Which phone had a better feel? Pick one:	
haptic	35
same or none	5
clam	2

DESIGNING EYES-FREE INTERACTION

When designing an eyes-free interaction, the system should aim to be usable by the general public and it should display an UI which enables a novice user to be able to pick it up easily and use it immediately. It should also not rely heavily on complex recognition technologies [4]. Below are some of the important design principles that should be followed.

1.) Immediate output: Eyes-free feedback needs to be immediate and short-lived or continually presented as unobtrusive background information[Designing eyes-free interaction]. It needs to pass on the information so that it can be absorbed quickly by the user.

2) Smooth transition from novice to expert: It is important to provide an interface where novice user can use the interface straight away by providing a basic graphical interface to begin with. However, it needs to provide a way in which novice user can seamlessly transition over to expert skill set where user no longer needs a simple graphical interface.

3) Input reflects on bodily constraints: when designing gestures, it is important to consider whether the gesture can be performed without interfering with user's everyday activity. The magnitude of the gesture should consider factors such as user's balance. Gesture which will be performed often needs to be designed so that it does not fatigue the user.

Hardware

The device needs to be small, light and be tough. It needs to have some sort of graphical user interface to allow novice user to start learning the eyes-free interaction. To detect gestures, sensors such as accelerometers and gyroscopes are needed. It will also need some basic form of sound output system for giving audio feedbacks. The tough challenge the device will need a form of rechargeable battery which is capable of lasting at least 12 hours.

Designing Eyes-Free input

Wrist is a natural body site for wearable computing devices. It is both easily accessible and socially accepted. Wrist movement is capable of translating and rotating in all three spatial axes. It is possible to make relatively large movements as well but this is not advisable. One of the design questions that must be answered with the input mechanism is a way to distinguish the input from everyday movement and the gestures itself. There are many proposed ways to do this such as performing a certain gesture before actually giving the gesture with message. The proposed system works to certain degree but it suffers when large amount of gesture needs to be given at once.

Designing Eyes-Free output

Device will usually contain some form of vibrotactile or sound based output to support eyes free interaction. Tactile output can be diversified by the rhythm and strength. Sound based output can accompany these tactile outputs. All three types of sound output can be used in its effective area to enhance the eyes-free interaction. Giving user an option to customize the feedback options will help user to become familiar with the smartwatch even more.

FUTURE WORKS

Throughout the review, several different research opportunities have been identified. The gesture based input currently suffers from the technical limitation of sensor detection mechanism for small devices. Voice recognition system is currently significantly unreliable compared to the other option, gestures. Hence to use this method of input, it will be necessary to further improve the algorithms and improve the current technology. I also believe that giving options for user to make their own gestures and applying them to the device will be a great option, however I believe this will require great amount of research to accomplish this.

Currently, the feedback system is in a good place compared to the input mechanism however there are still places where it could be improved even further. Similar to the input section, giving user an option to customize the feedback type will be great in helping user become more familiar with the system. Also, research to check if different types of sound feedback can simultaneously work without having issues will be helpful in further improving the eyes-free interaction.

SUMMARY

Eyes-free interaction, which is not yet well known, is starting to come in to view as the limitation of current graphical interface becomes more visible.

The current go-to method for input is gesture based input. For output, mixtures of all the possible feedbacks as they all have areas which they shine in.

It will be important to make sure that device is kept light as user will get tired quickly if the device becomes too heavy.

When designing input, it is important for the device some sort of feedback to indicate that the function has been performed.

It is hard to assume that user will be familiar with the eyes-free interaction to start off with so it is necessary to design a transition flow where user can seamlessly transfer from graphical interface to eyes-free interface.

In short future, with further research in this area, I believe that most mobile devices will have some sort of eyes-free interaction embedded in its system.

REFERENCES

1. Zhao, Shengdong, Pierre Dragicevic, Mark Chignell, Ravin Balakrishnan, and Patrick Baudisch. "Earpod: eyes-free menu selection using touch input and reactive audio feedback." In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1395-1404. ACM, 2007.
2. Brewster, Stephen A., and Peter G. Cryer. "Maximising screen-space on mobile computing devices." In *CHI'99 extended abstracts on Human factors in computing systems*, pp. 224-225. ACM, 1999.
3. Pirhonen, Antti, Stephen Brewster, and Christopher Holguin. "Gestural and audio metaphors as a means of control for mobile devices." In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, pp. 291-298. ACM, 2002.
4. Oakley, Ian, and Jun-Seok Park. "Designing eyes-free interaction." In *Haptic and Audio Interaction Design*, pp. 121-132. Springer Berlin Heidelberg, 2007.
5. Kim, Jungsoo, Jiasheng He, Kent Lyons, and Thad Starner. "The gesture watch: A wireless contact-free gesture based wrist interface." In *Wearable Computers, 2007 11th IEEE International Symposium on*, pp. 15-22. IEEE, 2007.
6. Lee, Seungyon Claire, Bohao Li, and Thad Starner. "AirTouch: Synchronizing in-air hand gesture and on-body tactile feedback to augment mobile gesture interaction." In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*, pp. 3-10. IEEE, 2011.
7. Yin, Min, and Shumin Zhai. "The benefits of augmenting telephone voice menu navigation with visual browsing and search." In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 319-328. ACM, 2006.
8. Garzonis, Stavros, Simon Jones, Tim Jay, and Eamonn O'Neill. "Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1513-1522. ACM, 2009.
9. Crease, Murray, and Stephen A. Brewster. "Making progress with sounds-The design and evaluation of an audio progress bar." In *Proceedings of ICAD*, vol. 98. 1998.
10. Blattner, Meera M., Denise A. Sumikawa, and Robert M. Greenberg. "Earcons and icons: Their structure and common design principles." *Human–Computer Interaction* 4, no. 1 (1989): 11-44.
11. Chang, Angela, and Conor O'Sullivan. "Audio-haptic feedback in mobile phones." In *CHI'05 extended abstracts on Human factors in computing systems*, pp. 1264-1267. ACM, 2005.