Technical Report ECE.P54.2009.2
February 23, 2009

In search of New Input Modalities for First Responders: Review of Human Computer Interaction in the Ubiquitous Computing Literature

Nemanja Memarovic
University of New Hampshire

First responders are trained individuals, like police officers, firefighters, and paramedics, who are responsible for the protection and preservation of life in the early stages of a natural or human made disaster, e.g. a car crash, flood, fire hazard, or an earthquake. They use computers and hand held devices to retrieve valuable information before and after they arrive at the scene of an incident: to retrieve medical or criminal records, to locate the nearest hospital, or to see the position of other team members. Given that their job is to protect and preserve life, and the stress of their working environment, the effort they put into interacting with computers should be minimal and interactions should be easily done. With the current user inputs, GUIs, keyboard, and mouse, they are far from that. If we want to ease the interaction between first responders and computers we need new forms of interactions, ones that are more natural to us, like interactions via eyes, touch and speech. These new input modalities are widely explored in ubiquitous computing. The following sections summarize the work that has been done in ubiquitous computing on new forms of interactions with the main focus on eye, touch, and speech as input modalities.

1. Introduction

"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" [1]. This is the opening sentence from Mark Weiser's groundbreaking paper "The computer for the 21st century". In this paper

Weiser describes his vision of a ubiquitous computing world where people interact with computers subconsciously to accomplish their everyday tasks. In his view this is the ultimate goal of HCI. As an example of what such a world would look like we get a brief view of Sal's daily routine. In her world she is surrounded by hundreds of computers, she is unaware of them and the interactions that are going on. Weiser pointed out that the main problem people have when they interact with computers is that computers are approachable only through complex jargon that has nothing to do with the task for which people use them. His goal was to establish a new way of thinking about computers, one that will take the human world into the account and that allows the computers to vanish into the background. A way to do that would be through "calm technology" [2].

But how would people interact with the hundreds of computers around them? In Sal's world everything seems so easy, but that is far from today's reality [3]. We are still using the keyboard and mouse as the main input modalities when interacting with computers. If we want to be closer to the world of Weiser's vision we need new user interfaces and new input modalities [4].

2. New forms of interactions

While we still use the keyboard and mouse as the main input modalities, new forms of interactions are emerging every day. Some of them can be seen in science fiction movies [5]. Sci-Fi movies have been bringing us visions of the future, both near and distant: captain Picard's "Computer" speech command, the touch screen that can understand gestures from the "Minority report", the speech user interface for the "Deep Thought" computer from "The Hitchhiker's Guide to the Galaxy" as well as the one for the famous "Hall9000" from "2001: A Space Odyssey" are just some of the new input modalities suggested by science fiction writers. We may need to turn science fiction into reality and bring those modalities to life if we want to improve HCI and shift from the standard inputs of keyboard and mouse.

A noteworthy attempt to chronicle and influence the direction of this shift was the “HCI 2020” forum organized by Microsoft Research in March 2007. The summary of this forum is a report [6] describing changes in HCI through the year 2007 and how these changes impact our lives and society. The report also includes recommendations for future development. According to the report digital technologies will continue to proliferate and they will continue to change the way we live. Users will shift from using GUIs, and keyboard and mouse to touch, speech, eye movements, gestures and brain waves. But will these new interfaces help us and will they improve the quality of our life? Or will they burden us, frustrate us and make us feel insecure? As concluded in this document, we have to think carefully about the design of these new interfaces if we want to make them usable and safe for use.

Among the many new input modalities three are the main focus of this review: eyes, touch, and speech.

2.1 Eyes

Vision is one of the main human senses: we perceive what is happening around us with our eyes. One of the main problems with using eyes as an input modality is that they are primarily input sensors and not output actors. We use our eyes to see something, not to trigger actions. However, eyes do have an output role as well: for example, when we are having a conversation with a group of people they know who we are addressing by looking where we look.

Drewes and Schmidt [7] propose gaze gestures as a way to use eyes as an input modality. In their user study they found that people can perform relatively complex gaze gesture that differ from gazes that happen during the "normal" interaction with the computer, e.g. surfing the Internet The complex gaze gesture could be used for initiating various commands, e.g. playing and pausing a song in Windows Media player. A similar solution is proposed by Porta and Turina [8]. The system they developed, called Eyes-S, is primarily explored as a text input. It uses nine non visible, on screen targets for recognition of graffiti like letters.

A simple command expressed through the eyes could be just closing your eyes. Hemmert et al. detected eye closure to activate text-to-speech when working with Microsoft Word [9]. They wanted to see if this new feature would be useful in practice and suggest it as a helpful addition, not a replacement for re-reading the text. They found both benefits and disadvantages of such a system. The benefits include stress relief, increase in productivity, change of perspective, and finding overseen errors in text. Disadvantages could come from eye closure being misinterpreted, listening taking more time than reading, and overhearing errors that could have otherwise been noted while reading.

When we read content from computer screens we usually need to scroll that content up and down. Kumar and Winograd used eye gaze information to enhance scrolling techniques [10]. In their informal user study subjects preferred gaze enhanced Page Up/Down techniques over the normal Page Up/Down. These techniques have the potential to radically reduce the number of actions users have to perform when reading from computer screens. It would be offered as an alternative, not as a replacement.

2.2 Touch

Touch, like vision, is also one of the main human senses: we touch things to feel their shape and texture; we touch other people to show our emotions or emphasize the moment. When we think of touch in HCI we think of touch screens. They have been widely deployed in mobile devices. The size of the screen on a mobile device is problematic because the user's fingers occlude the graphical elements he or she wishes to interact with. Another problem with touch screens is that the touch area of the finger is many times larger than a pixel of the display. Allowing the user to interact with a mobile device by touching the back of the device is a possible solution for the above problems proposed by Widgor et al. [11].

While mobile devices are designed for handheld use they spend most of their time stored in a pocket or bag. Often it would be useful to perform an initial action while keeping the device in its storage. Ronkainen et al. propose a tap input as an embedded interaction method for mobile devices [12]. While the main focus of their research was social acceptance of different gestures, they did investigate tapping the device, as a minimalist form of touch. Their user study showed that tap was the most accepted gesture of all that were explored.

Touch screens are not only deployed in mobile devices. CityWall, a large multi touch display, was deployed in the summer of 2007 in downtown Helsinki, Finland [13]. Peltonen et al. observed how people interact with CityWall over an eight day period. Their findings were on how multi-touch screens can affect and support social interaction in public spaces.

2.3 Speech

Speech is our natural way of communication and that makes it an input modality worth exploring for many applications. Often when interacting with mobile devices we are moving and we do not have the time to look at the GUI or type on the keyboard, or cannot do so safely. These situations make speech a preferable input modality. However, mobile settings often contain non-stationary noise which cannot be easily cancelled and users tend to adapt to that noise in acoustically unhelpful ways. Paek et al. explored how the above problems affect mobile voice search and proposed SearchVox as a solution [14]. SearchVox is a multimodal interface that enables users to use touch and text whenever speech fails and use speech whenever text entry becomes burdensome.

The number of in-vehicle devices is growing: we can now find GPS systems and mp3 players in most newly manufactured cars. The main task while we are in a car is driving, but we still need to interact with in-vehicle devices safely. In hands busy, eyes busy situation, like in cars, speech is a desirable input. A literature review on safety and usability of speech user interfaces for in-vehicle tasks while driving was written by Baron and Green [15].

In police cruisers we can find even more in-vehicle devices than in ordinary cars. A typical police cruiser has at least one radio, a controller for lights and siren, and a radar. All of these devices are competing for the police officer’s attention. The Project54 system was developed to help police officers by integrating all in-vehicle devices into one system [16]. Police officers can now work with their devices using a GUI, a SUI (Speech User Interface), as well as the original hardware interfaces of the in-vehicle devices. The Project54 system has been deployed in over 1000 police cruisers nationwide.

REFERENCES

[1] M. Weiser, "The computer for the 21st century," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 3, no. 3, pp. 3-11, 1999.

[2] M. Weiser and J. S. Brown, *The coming age of calm technolgy* Copernicus, 1997, pp. 75-85.

[3] Y. Rogers, "Moving on from Weiser's Vision of Calm Computing: Engaging UbiComp Experiences," Orange County, CA, USA: 2006, pp. 404-421.

[4] "Being Human: Human-Computer Interaction in the year 2020," Microsoft Research Ltd, 7 J J Thomson Avenue, Cambridge, CB3 0FB, England,2008.

[5] M. Schmitz, C. Endres, and A. Butz, "A survey of human-computer interaction design in science fiction movies," in *Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment* Cancun, Mexico: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007, pp. 1-10.

[6] "Being Human: Human-Computer Interaction in the year 2020," Microsoft Research Ltd, 7 J J Thomson Avenue, Cambridge, CB3 0FB, England,2008.

[7] H. Drewes and A. Schmidt, "Interacting with hte computer using gaze gestures," 2007.

[8] M. Porta and M. Turina, "Eye-S: a full-screen input modality for pure eye-based communication," in *Proceedings of the 2008 symposium on Eye tracking research & applications* Savannah, Georgia: ACM, 2008, pp. 27-34.

[9] F. Hemmert, D. Djokic, and R. Wettach, "Spoken words: activating text-to-speech through eye closure," in *CHI '08 extended abstracts on Human factors in computing systems* Florence, Italy: ACM, 2008, pp. 2325-2330.

[10] M. Kumar and T. Winograd, "Gaze-enhanced scrolling techniques," in *Proceedings of the 20th annual ACM symposium on User interface software and technology* Newport, Rhode Island, USA: ACM, 2007, pp. 213-216.

[11] D. Wigdor, C. Forlines, P. Baudisch, J. Barnwell, and C. Shen, "Lucid touch: a see-through mobile device," in *Proceedings of the 20th annual ACM symposium on User interface software and technology* Newport, Rhode Island, USA: ACM, 2007, pp. 269-278.

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

[13] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko, "It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre," in *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* Florence, Italy: ACM, 2008, pp. 1285-1294.

[14] T. Paek, B. Thiesson, Y.-C. Ju, and B. Lee, "Search Vox: leveraging multimodal refinement and partial knowledge for mobile voice search," in *Proceedings of the 21st annual ACM symposium on User interface software and technology* Monterey, CA, USA: ACM, 2008, pp. 141-150.

[15] A. Baron and P. Green, "Safety and Usability of Speech Interfaces for In-Vehicle Tasks while Driving: A Brief Literature Review ,"2008.

[16] A. L. Kun, W. T. Miller, III, and W. H. Lenharth, "Computers in police cruisers," *Pervasive Computing, IEEE*, vol. 3, no. 4, pp. 34-41, 2004.