Innovations in User Interfaces for Pervasive Computing

Javier Oliver
Dept. of Computer Engineering
University of Deusto
Bilbao, Spain
javier.oliver@deusto.es

Begoña García
Life Research Group
DeustoTech
Bilbao, Spain
mbgarciazapi@deusto.es

Abstract— Pervasive computing needs efficient interaction techniques to reduce the burden of the user, and to provide more transparent computing environments. This can be achieved by taking into account the ergonomics, physical location and cognitive load involved in using the system. To improve some of these issues, one can try to innovate existing interaction techniques and in this paper, we mention two improvements on the traditional manipulation of small touch sensitive screens. One of these improvements is back of device interaction, in which the screen is not occluded by the finger of the user. Instead, the screen cursor is moved by touching on the back of the device. The other improvement that is explained is the use of magnetic field detectors to increase the active area of manipulation of small screens. Another approach to obtain more fluid and transparent interfaces is to use gesture (i.e., human body movements) as input device. In this paper, we also give an overview of some of the characteristics of the use of body movements as input device, including eye movement, facial expression and other movements.

Keywords-pervasive computing interfaces; gesture input

I. INTRODUCTION

Mature technologies reach a level of design efficiency that allows for complex devices being used in a very natural and transparent way. After more than a century of evolution, one might argue that automobiles have reached that state. Due to their flexibility and multipurpose nature, computers can be regarded as more complex to operate than cars, but Mark Weiser, in his frequently cited article [1] envisioned that computers of the XXIst century would be used with the same ease and naturalness as cars today. Although some aspects of this ubiquitous computing vision have been fulfilled, computers, in general, are still hard to use, obtrusive devices that tend to turn the attention of the user away from the real task.

If computers are really going to step back into the background and become unobtrusive aids for users to carry out their tasks, user interfaces have to be rethought, and not only existing technologies have to be combined in more creative ways, but also new technologies have to be developed [1]. Ubiquitous computing devices have to be physically and cognitively available [2]. Physical availability includes issues such as location of the device, and ergonomic design, whereas cognitive availability refers to the mental effort involved in using the equipment. Devices that are available in both senses disappear from the user consciousness and cease to be in the way of the user's aims and objectives, which is precisely the goal of ubiquitous computing [2].

In this broad sense, availability can be provided by making interaction more intuitive, natural and effortless. In the following sections, two approaches that can be taken will be discussed: the improvement of the interaction with small screens and the use of gesture as a way to achieve a more natural and less obtrusive interaction.

In Section II, the problem of manipulating small touch screens is described, and two possible solutions are presented. The first solution consists in moving the touch sensitive surface to the back of the device, and the second involves the use of a magnetic field detector. In Section III, the use of gesture as input device for pervasive computing systems is explored, and techniques such as eye movement and facial expression are described. Finally, in Section IV, conclusions are drawn.

II. THE *FAT FINGER* PROBLEM

Computer artifacts have been reducing their size in the last few years to the point where the limiting factor in their dimensions is not the battery size or the electronic circuit miniaturization, but the size of the input and output devices needed to control the computer [3]. A screen should be big enough to be comfortably read at normal viewing distance, and buttons –physical or virtual- should be big enough for a regular sized finger to press them. Even though the size of input devices keeps shrinking over time, the size of the average finger remains constant, hence the title of this section: the fat finger problem [4].

In order to increase the availability [2] of user interfaces, the fat finger problem has to be addressed. According to Fitts' Law [5], the smaller the target, the longer the interaction time. So, whenever a button is below its optimal size, we are making interaction unnecessarily slow. But, in the case of very small screens, the finger can cover most of it, making touch interaction almost unfeasible. Several solutions have been proposed for this situation. In the case of screens over one inch in diagonal, typical interfaces try to solve screen occlusion by offering the user an offset cursor, and showing a reproduction of the occluded area in a free part of the screen. But this approach does not work for screens below one inch in diagonal in which it has been proposed to move the touch sensitive surface to the back of the device, thus leaving the screen completely visible, and

 ISBN: 978-1-61208-289-9

showing the position of the finger by means of a small cursor [6].

Another solution for the interaction with small devices is the use of a magnetic field detector, which can be located behind the screen, thus leaving the screen visible. This type of device is capable of quite accurate positioning at a very moderate price. In a particular study, a screen of 1.5 inches in diagonal was used and the magnetometer offered an angular accuracy of about two degrees for the cost of less than five US dollars [7]. Users had to wear a small magnet in the index finger, which gave a practical range of about 10 cm. from the screen, offering a total active area of nearly 300 cm^2. With this technology, the original area of the screen was increased by a factor of more than 50, and the occlusion problem was solved.

III. Gesture as input device

Pervasive computing needs user interfaces that do not make assumptions about what input and output devices are available for the user. The trend is to move from explicit use of traditional input devices, to more implicit forms of interaction, such as speech or gestures [8]. This implies that users will increasingly address computers in the same fashion as they address other people, taking advantage of their experience in human to human communication, in what has been considered an anthropomorphic approach. Currently, there are technologies that can input user movements into the computer, such as gloves and suits, but they are too cumbersome, expensive, difficult to calibrate and disruptive of the user's flow of action. New interfaces based on movement should be much more flexible in order to fully support the ubiquitous computing model.

Human movement is complex, involving the skeletal, muscular and neurological systems. The nature of the different types of movements depends on the class of muscle fibers involved, the shape and orientation of bones and joints, etc. Depending on the distance covered in these movements they can be classified as micro-movements (short distance) and macro-movements (longer distances) [8], and both types can be used in explicit and implicit interaction with ubiquitous systems.

The technologies used to capture body movements can be classified in two main groups:

- Wearable: the user carries different devices to monitor the body movements.
- Indirect: a variety of sensors can detect body movements at a certain distance.

Wearable devices can make use of accelerometers that detect movements in two or three dimensions. More innovative options include monitoring the electrical activity associated with the movement of the hand and video cameras that can track finger movements from a wrist mount [9].

A. *Eye Movement*

One of the most interesting micro-movements for interaction with computers is eye movement. Each eye is controlled by a group of six muscles that produce very fine and precise movements. Monitoring these movements is a promising area of research because they can be used to find out about the point of interest of the user in an unobtrusive way. In order to understand how eye movements reflect underlying cognitive processes, protocol analysis can be used. During the execution of a task, eye movements can be recorded. These recordings are the protocols, that can then be analyzed and can be used to understand user behavior in basic interaction tasks, to study the processing of information in users, or to predict the goals of users in real time interfaces [10].

The human retina has a very small area of high visual acuity, called the *macula*. This region, of about 1 mm^2, is near the back pole of the retina and the eyes are in constant movement, so that the point of interest always forms its image in the macula. These constant movements are called saccadic, and they are a good source of information to analyze cognitive actions such as reading or searching [8]. The monitoring of eye movements is a good technique to carry out certain input tasks in a natural and hands free manner. For example, users can look at an object on the screen and blink to select it. An interesting application of eye monitoring is the implementation of automatic scrolling. Normally, it is necessary to displace a rectangle on the screen to perform a scroll, but it is possible to do it by means of an eye gaze [11].

B. *Facial Expression*

The human face has a high number of muscles that are used to show a wide range of emotions. Traditionally, it has been thought that humans can show six basic types or emotion: fear, anger, disgust, sadness, surprise and happiness, and that these emotions were, in fact, an universal language, being understood worldwide. However, more recent research is showing that classification of emotions is not that straightforward, and that the expressions of emotions is probably culturally dependent [12]. If these findings are confirmed, it would mean that user interfaces based on facial expression would have to take into account cultural differences, and it would also mean that this type of interfaces would me more complex than initially thought. Computer systems are able to recognize a range of emotions, but recognition rates can be as low as 64% in some studies, so the technology is not mature yet to be used as the only type of input, and it should be avoided in critical applications [12]. It is possible, however, to design a multimedia application or videogame that modulates its content depending on the emotion that the user is showing in his face [8]. Apart from inferring the emotion felt by the user, facial configuration and movements have also been used to move a cursor on the screen (a technique mainly aimed at disabled people) or to control different types of programs such as audio or graphics applications [13].

C. *Other Movements*

Although until now facial expression has taken most of the research community's attention, arm and hand movements provide abundant information that enrich verbal communication, and different authors agree that emotion

 ISBN: 978-1-61208-289-9

detection should be multimodal, using input from different body movements [14]. The different movements described so far can widely vary depending on the type of user, and they won't be the same, for example in the case of very young children, adult users or elderly people. The movements in the first group will typically be limited by the lack of neurological development, and in the last one, by the effects of age in the neuromuscular system. What seems clear is that taking into account the limitations of extreme age groups when designing user interfaces based on body movements, can result in better user interfaces for all age groups. Also, it seems probable that single design user interfaces will not adapt well to all situations, and customizable user interfaces will be needed [8].

D. *Examples of Applications*

The number of user applications that can take advantage of gesture for input is potentially endless. For example, a gesture-based interaction photo album (GIA) has been developed that uses gestures as the main source of input [15]. Y. K. Jin et al. started by benchmarking the existing album software for PCs and developing a set of user requirements. The system follows a physical photo album metaphor, and each album in GIA is similar to a folder in a file system. The opening screen shows a shelf with a number of albums on it. The thickness of each album is proportional to the number of pictures that it contains. GIA uses a gesture based interface based on a touch screen. For example, to turn pages, one only has to make a horizontal stroke on the screen. Depending on the length and speed of the gesture, the user can turn one or more pages with the same stroke. To avoid unwanted modifications of data, there is an edit mode, in which data can be modified, and a safe view mode, in which no data can be modified. The aim of the design of GIA was to appeal both to beginners and experts, and it uses two sets of gestures: symbols and characters. Users can choose an album by pretending to pick it up from the shelf. They can also modify the orientation and size of a picture by means of gestures. To define the sets of gestures, the authors analyzed different user groups, including children and elderly people, and they studied their interaction styles. The gesture sets were tested using expert evaluations, usability tests and focus group interviews. The expert evaluations were designed to check four dimensions: intuitive, efficient, fun and innovative. The first two try to measure the practical side of the gesture interface, and the last two deal more with the subjective and emotional side of it. After having a working prototype and before the final usability test, the system was shown at an exhibition where visitors were able to use it. The interaction techniques proved to be easy to remember, and after a first use, visitors had no problems manipulating the interface. In the final usability test, two different designs were compared: a traditional point and click interface based on menus and buttons and the gesture based interface. According to the authors, the gesture interface was, in general, preferred.

Looking over the shoulders to get a password or some other sensitive information is a problem difficult to solve. If a traditional input device such as mouse, keyboard or touch screen is used, a casual observer can obtain the secret information. In general, methods that make shoulder surfing more difficult are also an inconvenience for the user. A system called EyePassword, based on gaze tracking, can virtually eliminate all shoulder surfing attacks while maintaining a reasonable ease of input [16]. In gaze based password input, instead of typing on a keyboard or touch screen, the user looks at each of the desired characters one after the other. This technique can be used not only for alphanumeric passwords, but also for graphical passwords (i.e., different trajectories among the points in a matrix). The area associated with each character has to be large enough to avoid unwanted selection of characters. The most important factor to determine the ideal size of a character on the screen is the resolution of the eye tracker. For example, if a tracking device has a resolution of one degree of visual angle, this means that, for a normal viewing distance of 50 cm., the input of the eye tracker will have an uncertainty radius of about 33 pixels and each character should be at least 66 pixels wide. In EyePassword, a target size of 84 pixels was selected, with a spacing between characters of 12 pixels to reduce the possibility of erroneously selecting one character for another. In order to input the character, the user can fix the gaze of the target for a short period of time, or use a special trigger key like *return* or *space bar*. The fixation method is preferred because it is less error prone, and it is more secure, given that less information is available for the casual observer. Four different technique were implemented and evaluated:

1. Traditional keyboard input.
2. Gaze input combined with a trigger key.
3. Gaze input combined with a certain fixation time, and a qwerty layout of the characters.
4. Gaze input combined with a certain fixation time, and an alphabetic layout of the characters.

Typical passwords were used in the test, with 8 to 9 characters, including uppercase, lowercase, symbols and numbers. The error rate was significantly higher in the gaze plus trigger key combination, and although the input times were higher for gaze based input than for the keyboard, more than 80% of the subjects confirmed that they would prefer to use a gaze based method instead of a keyboard when entering a password in a public place.

IV. CONCLUSION

Pervasive computing needs interaction techniques that do not disrupt the workflow of the user. Computer devices should be available [2] both in the sense of physically available (i.e., ergonomics, location) and cognitively available (i.e., the mental effort required to use the device is minimized).

This broad sense availability can be reached by improving existing interaction techniques or by exploring new ones. In this paper, we have presented improvements for the classical interaction technique of manipulating small touch sensitive screens, such as back of device interaction and the use of magnetic field detectors. We have also presented some of the recent techniques that use body movements and gestures as input for the computer system, including eye movements, facial expressions, and arm or hand movements.

REFERENCES

[1] M. Weiser, "The computer for the 21st Century. Scientific American," Vol. 265, Nº 3, 1991, pp. 94-104.

[2] V. Waller and R. B. Johnston, "Making ubiquitous computing available," Communications of the ACM, Vol. 52, Nº 10, 2009, pp. 127-130.

[3] T. Ni and P. Baudisch, "Dissapearing mobile devices," 22nd Annual Symposium on User Interface Software and Technology, October 4-7, Victoria, British Columbia, Canada, 2009, pp 101-110.

[4] K. A. Siek, Y. Rogers, and K. H. Connelly, "Fat finger worries: how older and younger users physically interact with PDAs," In Proc. INTERACT'05, 2005, pp. 267-280.

[5] T. Anthony, "Finger-friendly design: ideal mobile touch screen target sizes. Smashing Magazine," Available at: http://uxdesign .smashingmagazine.com/2012/02/21/finger-friendly-design-ideal-mobile-touchscreen-target-sizes/, 2012 [retrieved: july, 2013].

[6] P. Baudisch and G. Chu, "Back-of-device interaction allows creating very small touch devices," In Proc CHI 2009, pp. 1923-1932.

[7] C. Harrison and S. E. Hudson, "Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices," 22nd Annual Symposium on User Interface Software and Technology, October 4-7, Victoria, British Columbia, Canada, 2009, pp. 121-124.

[8] J. H. Abawajy, "Human-computer interaction in ubiquitous computing environments," International Journal of Pervasive Computing and Communications, vol. 5, Nº 1, Special issue: Advances in pervasive computing, 2009, pp. 31-77.

[9] A. Vardy, J. A. Robinson, and L. T. Cheng, "The WristCam as input device," San Francisco, CA, Proceedings of the 3rd International Symposium on Wearable Computers, 1999, pp. 199-202.

[10] D. D. Salvucci and J. R. Anderson, "Automated eye movement protocol analysis," Human-Computer Interaction, Vol. 16 Nº 1, 2001, pp. 38-49.

[11] M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd, "Reducing Shoulder-surfing by Using Gaze-based Password Entry," Symposium On Usable Privacy and Security (SOUPS), July 18-20, Pittsburgh, PA, USA, 2007, pp. 13-19.

[12] M. Pantic and L. J. M. Rothkrantz, "Toward an affect sensitive multimodal human computer interaction," Proceedings of the IEEE, Vol. 91 Nº 9, 2003, pp. 1370-90.

[13] G. C. DeSilva, M. J. Lyons, and N. Tetsutani, "Vision based acquisition of mouth actions for human-computer interaction," Auckland, Proceedings of the 8th Pacific Rim International Conference on Artificial Intelligence, 2004, pp. 959-60.

[14] H. Gunes and M. Piccardi, "Automatic visual recognition of face and body action units," Proceedings of the 3rd International Conference on Information Technology and Applications, 2005, pp. 668-73.

[15] Y. K. Jin et al. "GIA: design of a gesture based interaction photo album," Personal and Ubiquitous Computing, Vol. 8 Nº 3/4, 2004, pp.227-233 [retrieved: july, 2013].

[16] M. Kumar, A. Paepcke, and T. Winograd, "EyePoint: practical pointing and selection using gaze and keyboard," San Jose, CA., Proceedings of the CHI: Conference on Human Factors in Computing Systems, 2007, pp.421-430.