Toward Mobile Eye-Based Human-Computer Interaction

Andreas Bulling and Hans Gellersen

Eye-based human-computer interaction (HCI) goes back at least to the early 1990s. Controlling a computer using the eyes traditionally meant extracting information from the gaze—that is, what a person was looking at. In an early work, Robert Jacob investigated gaze as an input modality for desktop computing.[1] He discussed some of the human factors and technical aspects of performing common tasks such as pointing, moving screen objects, and menu selection. Since then, eye-based HCI has matured considerably. Today, eye tracking is used successfully as a measurement technique not only in the laboratory but also in commercial applications, such as marketing research and automotive usability studies.

Current research on eye-based interfaces mostly focuses on stationary settings. However, advances in mobile eye-tracking equipment and automated eye-movement analysis now allow for investigating eye movements during natural behavior and promise to bring eye-based interaction into people's everyday lives.

MOBILE EYE TRACKING

Daily life settings call for highly miniaturized eye trackers with real-time processing capabilities. Despite recent technological advances, the development of mobile eye trackers is still an active research topic. Päivi Majaranta and Kari-Jouko Räihä identified three key challenges for stationary gaze-based typing systems: eye-tracking accuracy, calibration drift, and the Midas touch problem—that is, the problem of distinguishing the user's intentional eye input from other eye movements that occur while using an interface.[2]

These challenges also apply to mobile eye-based interfaces. Eye-tracking accuracy poses particular difficulties because it's affected by factors such as eye movements, calibration quality, and calibration drift during operation. Stationary eye trackers (also known as remote eye trackers) achieve a visual-angle accuracy of approximately 0.5 degrees. Because mobile eye trackers must address varying distances between the interface and the user, current mobile systems are less precise, pushing visual-angle accuracy out to approximately 1 degree.

Instead of using gaze directly, Heiko Drewes and his colleagues suggested using gaze gestures—that is, sequences of several consecutive eye movements.[3] Although eye gestures require more cognitive effort than natural eye movements, they remain promising for mobile gaze-based input because they're more robust against eye-tracking inaccuracies.

Commercial eye trackers are increasingly addressing these challenges. The first generation of mobile video-based systems required bulky headgear and additional equipment, such as digital video recorders or laptops to store and process the video streams. Examples include the Mobile Eye by Applied Science Laboratories (see Figure 1a) or the iView X HED by SensoMotoric Instruments (see Figure 1b). Recently, Tobii Technology announced the first video-based eye tracker fully integrated into an ordinary glasses frame. The system consists of the glasses and a small, pocket-worn device for video processing and data collection (see Figure 2).

In parallel to commercial products, several open source projects aim to develop inexpensive hardware and software for video-based eye tracking. The most advanced of these projects are openEyes (www.thirtysixthspan.com/openEyes), Opengazer (www.inference.phy.cam.ac.uk/opengazer), and the IT University of Copehagen (ITU) Gaze Tracker (www.gazegroup.org/downloads/23-gazetracker). The open source option lets users easily prototype their applications and rapidly incorporate experimental findings into the interface design.

A remaining issue with video-based eye trackers is the considerable processing power that video processing

Published by the IEEE CS ■ 1536-1268/10/$26.00

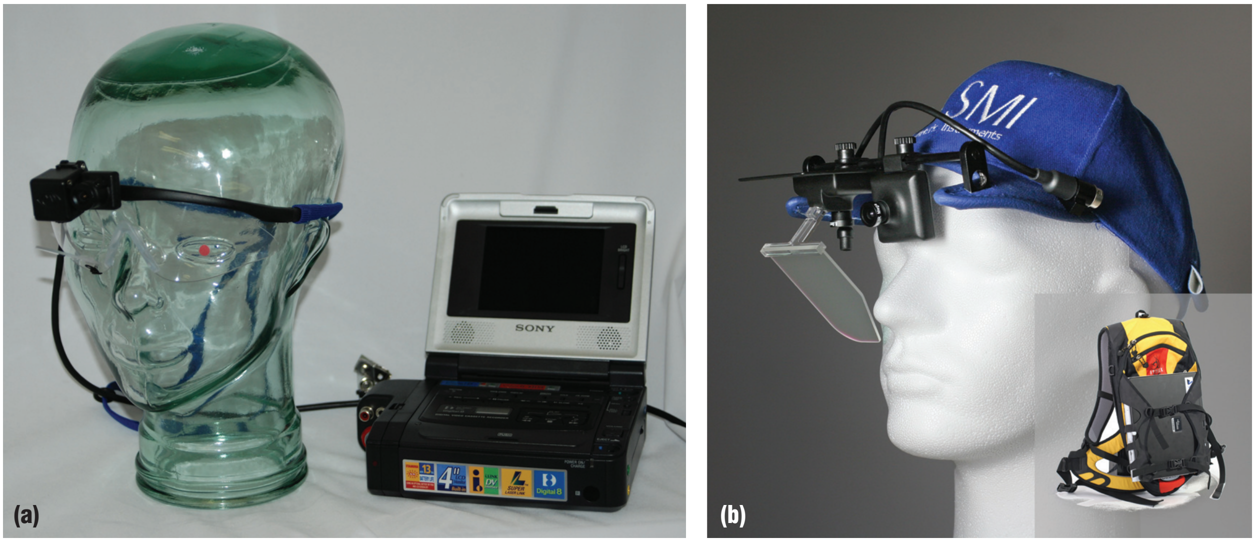

Figure 1. First-generation video-based eye trackers: (a) Mobile Eye (photo courtesy of Applied Science Laboratories) and (b) iView X HED (photo courtesy of SensoMotoric Instruments).

requires. In contrast to their stationary counterparts, mobile eye trackers must conserve power to meet operating times required for long-term studies in research and commercial applications. Although none of the manufacturers provide exact figures, users have reported operating times of about two to four hours.

Efforts to extend operating times led researchers to consider more lightweight measurement techniques, such as electrooculography (EOG). Using electrodes attached to the skin around the eyes, EOG measures changes in the electric potential field caused by eye movements. By analyzing these changes, it's possible to track relative eye movements—that is, how a person is looking at something. In earlier work, we demonstrated an EOG-based eye tracker, the Wearable EOG goggles.[4] The system uses dry EOG electrodes integrated into a goggles frame and a small pocket-worn device for real-time EOG signal processing, data storage, and transmission (see Figure 3). The low-power design of the first prototype supports an operating time of more than seven hours.

Hiroyuki Manabe and Masaaki Fukumoto developed an EOG-based eye tracker that doesn't require facial electrodes but instead uses an electrode array mounted on commercial headphones.[5] Miniaturizing the headphones to earplugs would reduce the approach's obtrusiveness, but it raises two other issues—namely, low signal-to-noise ratio and poor separation of the horizontal and vertical eye-movement components.

EMERGING RESEARCH

Researchers are also investigating applications for eye-based interaction. So far, the applications are typically limited to stationary settings.

Behavioral Studies

Eye-movement analysis has a long history in experimental psychology as a tool for investigating visual behavior. Mobile eye tracking is enabling a new class of studies to analyze people's eye movements in natural environments.[6] These studies have advanced our understanding of how the brain processes daily life tasks and what the visual system's role is.[7] Applied for HCI purposes, these findings will eventually provide further insights into the design and evaluation of novel mobile eye-based interfaces.

Figure 2. Tobii Glasses. This recently released video-based eye tracker is the first system to fully integrate into an ordinary glasses frame (photo courtesy of Tobii Technology).

Attentive User Interfaces

The strong link between gaze and user attention paved the way for mobile eye trackers to become a key component in attentive user interfaces. For example, Ted Selker and his colleagues described a glasses-mounted eye tracker to

WEARABLE COMPUTING

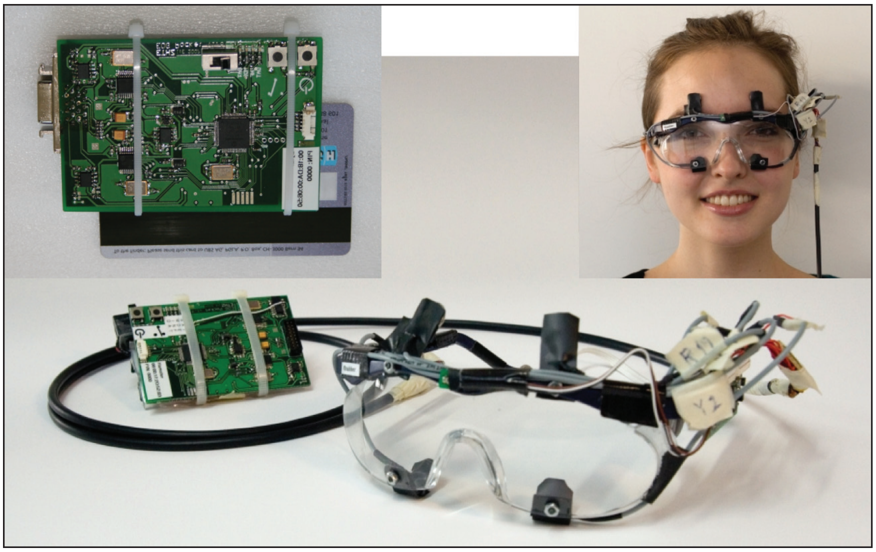

Figure 3. Wearable electrooculography (EOG) goggles. The Swiss Federal Institute of Technology (ETH) Zurich developed these goggles to track relative eye movements by measuring changes in the electric potential field around the eyes.

analyze a user's eye fixations on different environmental targets.[8] They used these fixations as attentional triggers to automatically communicate with the target. They suggest that in a meeting scenario, different people wearing their devices could automatically exchange business cards by looking at each other.

Roel Vertegaal and his colleagues described a similar system, the attentive cell phone, which used low-cost EyeContact sensors and speech analysis to detect whether its user was in a face-to-face conversation.[9] If the sensors detected a face-to-face conversation, the phone automatically switched to silent mode and notified potential callers about the user's preferred notification channel, such as vibration, knocking, or ringing. In a later work, Connor Dickie and his colleagues used the same sensors to develop two appliances that adapt to user attention:[10]

- an attentive mobile video player that automatically paused content when the user was no longer looking at it, and
- an attentive reading application that advanced text only when the user was looking.

The authors argued that making mobile devices sensitive to a user's attention would let him or her more gracefully switch between using the devices to consume media and using them to manage life.

Ralf Biedert and his colleagues prototyped Text 2.0, an interesting application that's so far limited to stationary eye tracking.[11] Text 2.0 provides a responsive text-reading interface that combines eye tracking with real-time interaction. The system aims to enhance the reading experience by augmenting text with real-time sound and visual effects, automatic text translation and comprehension reading assistants, and a quick skim mode that fades out less important words if the system detects skimming behavior. An automatic reading detector could trigger and monitor the application.[12]

Multimodal Interaction

Many everyday tasks require strong spatial and temporal coordination between eye and hand movements.[7] This observation has motivated researchers to investigate ways to combine gaze with other input modalities, such as head or hand gestures. For mobile HCI applications, a joint analysis of different sensing modalities promises more versatile interaction types. Multimodal interfaces could automatically select input modalities best suited for the situation at hand.

For example, Shumin Zhai and his colleagues introduced an approach to multimodal pointing.[13] Their system moved the mouse pointer to a target area by gaze, but it implemented pointing and selection manually using the computer mouse, thus avoiding overloading the user's visual system with a motor-control task. They showed that their approach reduced physical effort and fatigue compared to traditional manual pointing and provided greater accuracy and naturalness than traditional gaze pointing. Manu Kumar and his colleagues presented a similar approach for combining gaze with keyboard input.[14] Their EyePoint system uses look-press-look-release actions to overcome the eye trackers' accuracy limitations.

Eye Tracking on Mobile Devices

Researchers are also beginning to investigate eye tracking on mobile devices, which increasingly come equipped with cameras. For example, Emiliano Miluzzo and his colleagues developed a mobile phone that tracks the position of a user's eyes on the phone display.[15] Their system uses the phone's camera to translate eye position in the camera image into nine different onscreen positions.

Weston Sewell and Oleg Komogortsev developed a real-time gaze-tracking system that relies on a standard webcam integrated into a laptop computer.[16] In a first step, the system processes the webcam videos to detect the user's face, eye, and iris. Next, it processes these image sections and feeds them into an artificial neural network for training. A user study of five participants showed an average eye-tracking accuracy of about four degrees of visual angle.

Takashi Nagamatsu and his colleagues described an augmented mobile phone that offers a gaze-and-touch interface that detects gaze direction in 3D.[17] The phone uses stereo cameras with infrared emitters attached to the phone. Their example application let users look at a certain region of an interactive map and then use a finger to zoom in, effectively avoiding the Midas touch problem. The prototype was rather bulky, but advances in miniaturization and embedded processing might enable a mobile-phone implementation in the near future.

Eye-Based Context Inference and Cognition-Aware User Interfaces

Although gaze has been the traditional focus of eye-based HCI, eye movements provide additional information that could be useful to human-computer interfaces. In earlier work, we introduced eye-movement analysis as a modality for context and activity recognition.[18] Eye-movement patterns reveal much about observed activities. Similarly, particular environments affect eye movements in specific ways. In addition, unconscious eye movements are linked to cognitive visual perception processes. These characteristics make eye movements a distinct information source about a user's context. Eventually, information derived from eye movements might let us extend the current notion of user context with a cognitive dimension, leading to so-called cognition-aware interfaces.[19]

Yoshio Ishiguro and his colleagues followed a different approach for eye-based information retrieval. Instead of extracting information from eye-movement dynamics, they used gaze as an attention indicator to extract information from objects in the environment.[20] To this end, they developed an eye tracker fully integrated into ordinary glasses with software to detect and extract faces and text from a video scene in real time. Among the potential applications for such a system are human memory enhancement or life logging for the elderly.

Recent developments in mobile eye-tracking equipment point the way toward unobtrusive human-computer interfaces that will become pervasively usable in everyday life. The potential applications for the further capability to track and analyze eye movements anywhere and anytime calls for new research to develop and understand eye-based interaction in mobile daily-life settings.

REFERENCES

1 R.J.K. Jacob, "What You Look at Is What You Get: Eye Movement Based Interaction Techniques," *Proc. SIGCHI Conf. Human Factors in Computing Systems* (CHI 90), ACM Press, 1990, pp. 11–18.

2. P. Majaranta and K.-J. Räihä, "Twenty Years of Eye Typing: Systems and Design Issues," *Proc. 2002 Symp. Eye Tracking Research and Applications* (ETRA 02), ACM Press, 2002, pp. 15–22.

3. H. Drewes, A. De Luca, and A. Schmidt, "Eye-Gaze Interaction for Mobile Phones," *Proc. 4th Int'l Conf. Mobile Technology, Applications, and Systems* (Mobility 07), ACM Press, 2007, pp. 364–371.

4. A. Bulling, D. Roggen, and G. Tröster, "Wearable EOG Goggles: Seamless Sensing and Context-Awareness in Everyday Environments," *J. Ambient Intelligence and Smart Environments*, vol. 1, no. 2, 2009, pp. 157–171.

5. H. Manabe and M. Fukumoto, "Full-Time Wearable Headphone Type Gaze Detector," *Extended Abstracts of the SIGCHI Conf. Human Factors in Computing Systems* (CHI 06), ACM Press, 2006, pp. 1073–1078.

6. M.M. Hayhoe and D.H. Ballard, "Eye Movements in Natural Behavior," *Trends in Cognitive Sciences*, vol. 9, no. 4, 2005, pp. 188–194.

7. J.B. Pelz et al., "Portable Eyetracking: A Study of Natural Eye Movements,"

WEARABLE COMPUTING

Proc. SPIE, Human Vision and Electronic Imaging V, vol. 3959, SPIE (Int'l Soc. for Optical Eng.), 2000, pp. 566–583.

8. T. Selker, A. Lockerd, and J. Martinez, "Eye-R, a Glasses-Mounted Eye Motion Detection Interface," *Extended Abstracts of the SIGCHI Conf. Human Factors in Computing Systems* (CHI 01), ACM Press, 2001, pp. 179–180.

9. R. Vertegaal et al., "Designing Attentive Cell Phone Using Wearable Eye-Contact Sensors," *Extended Abstracts of the SIGCHI Conf. Human Factors in Computing Systems* (CHI 02), ACM Press, 2002, pp. 646–647.

10. C. Dickie et al., "Eyelook: Using Attention to Facilitate Mobile Media Consumption," *Proc. 18th Symp. User Interface Software and Technology* (UIST 05), ACM Press, 2005, pp. 103–106.

11. R. Biedert et al., "Text 2.0," *Extended Abstracts of the SIGCHI Conf. Human Factors in Computing Systems* (CHI 10), ACM Press, 2010, pp. 4003–4008.

12. A. Bulling et al., "Robust Recognition of Reading Activity in Transit Using Wearable Electrooculography," *Proc. 6th Int'l Conf. Pervasive Computing*, LNCS 5013, Springer, 2008, pp. 19–37.

13. S. Zhai, C. Morimoto, and S. Ihde, "Manual and Gaze Input Cascaded (MAGIC) Pointing," *Proc. SIGCHI Conf. Human Factors in Computing Systems* (CHI 99), ACM Press, 1999, pp. 246–253.

14. M. Kumar, A. Paepcke, and T. Winograd, "EyePoint: Practical Pointing and Selection Using Gaze and Keyboard," *Proc. SIGCHI Conf. Human Factors in Computing Systems* (CHI 07), ACM Press, 2007, pp. 421–430.

15. E. Miluzzo, T. Wang, and A. Campbell, "Eyephone: Activating Mobile Phones with Your Eyes," *Proc. 2nd ACM SIGCOMM Workshop on Networking, Systems, and Applications on Mobile Handhelds* (MobiHeld 10), ACM Press, 2010, pp. 15–20.

16. W. Sewell and O. Komogortsev, "Real-Time Eye Gaze Tracking with an Unmodified Commodity Webcam Employing a Neural Network," *Extended Abstracts of the SIGCHI Conf. Human Factors in Computing Systems* (CHI 10), ACM Press, 2010, pp. 3739–3744.

17. T. Nagamatsu, M. Yamamoto, and H. Sato, "Mobigaze: Development of a Gaze Interface for Handheld Mobile Devices," *Extended Abstracts of the SIGCHI Conf. Human Factors in Computing Systems* (CHI 10), ACM Press, 2010, pp. 3349–3354.

18. A. Bulling et al., "Eye Movement Analysis for Activity Recognition Using Electrooculography," to be published in *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2010; http://dx.doi.org/10.1109/TPAMI.2010.86.

19. A. Bulling, D. Roggen, and G. Tröster, "What's in the Eyes for Context-Awareness?" to be published in *IEEE Pervasive Computing*, 2010; http://dx.doi.org/10.1109/MPRV.2010.49 (preprint).

20. Y. Ishiguro et al., "Aided Eyes: Eye Activity Sensing for Daily Life," *Proc. 1st Augmented Human Int'l Conf.* (AH 10), ACM Press, 2010.

Andreas Bulling is a postdoctoral research fellow at the University of Cambridge's Computer Laboratory and Lancaster University's School of Computing and Communications, funded by a Feodor Lynen Fellowship of the German Alexander von Humboldt Foundation. His research focuses on cognition-aware systems with applications in ubiquitous computing and human-computer interaction. Bulling has a PhD in information technology and electrical engineering from the Swiss Federal Institute of Technology (ETH) Zurich. Contact him at andreas.bulling@acm.org.

Hans Gellersen is a professor of interactive systems at Lancaster University's School of Computing and Communications. His research interests are in ubiquitous computing and user-interface technologies. Gellersen has a PhD in computer science from the Technical University of Karlsruhe, Germany. Contact him at hwg@comp.lancs.ac.uk.