Remote Eye Tracking: State of the Art and Directions for Future Development

Martin Böhme, André Meyer, Thomas Martinetz, and Erhardt Barth
Institute for Neuro- and Bioinformatics
Ratzeburger Allee 160
D-23538 Lübeck, Germany
{boehme, meyera, martinetz, barth}@inb.uni-luebeck.de

Keywords
Remote eye tracking, free head motion, eye model

Introduction

Recent years have seen rapid developments in the field of remote eye tracking. Whereas only a few years ago the standard in eye tracking was for systems to be intrusive, i.e. they either required the user's head to be fixated or equipment to be mounted on the user's head, systems have now evolved to the point where the user can move freely in front of the system (within certain limits), and good accuracy (1 degree or better) is achieved throughout the whole working range. This has been demonstrated by a number of commercial and academic systems, both multi-camera (Beymer and Flickner, 2003; LC Technologies, 2006) and single-camera (Tobii, 2002; SMI, 2006; Hennessey et al., 2006; Guestrin and Eizenman, 2006; Meyer et al., 2006). To clarify terms, we will use the term "remote eye tracking" here to mean a system that operates without contact with the user *and* permits free head movement within reasonable limits without losing tracking.

In this paper we give an overview of our own work in this field and give our view on where worthwhile opportunities for future research lie.

State of the Art

The first remote eye tracking systems that appeared in the literature used multiple cameras (Shih et al., 2000; Beymer and Flickner, 2003; Ohno and Mukawa, 2004; Brolly and Mulligan, 2004; Yoo and Chung, 2005), usually in some kind of stereo setup. Morimoto et al. (2002) describe a single-camera eye tracker with an accuracy of about 3 degrees. The first single-camera remote eye tracker with high accuracy (0.5 to 1 degree) and good tolerance to user movement was a commercial system (Tobii, 2002), but implementation details have not been made available. Recently, several academic groups have built similar single-camera systems (Hennessey et al., 2006; Guestrin and Eizenman, 2006; Meyer et al., 2006). (Guestrin and Eizenman's system allows only small head movements, but it appears that their well-founded approach would allow greater head movements with a higher-resolution camera.) The main additional difficulty in the single-camera setting is determining the distance of the user from the camera, since a triangulation as in the multi-camera case can not be carried out. The advantage of a single-camera system is of course the reduced cost and smaller size.

Figure 1. Remote eye tracker system setup. The eye tracking hardware consists of a single high-resolution camera below the display and two infrared LEDs to either side.

The setup of our own single-camera system (Meyer et al., 2006) is shown in Figure 1. It consists of a high-resolution camera (1280x1024 pixels) and two infrared LEDs mounted to either side of the camera. The LEDs provide general illumination and generate reflexes on the surface of the cornea. These corneal reflexes (CRs) are used to find the eye in the camera image and determine the location of the centre of corneal curvature in space. The system is shown here mounted below an LCD display.

The software consists of two main components: The image processing algorithms that are used to determine the position of the CRs and pupils in the image, and the gaze estimation algorithm, which uses this data to compute the direction of gaze.

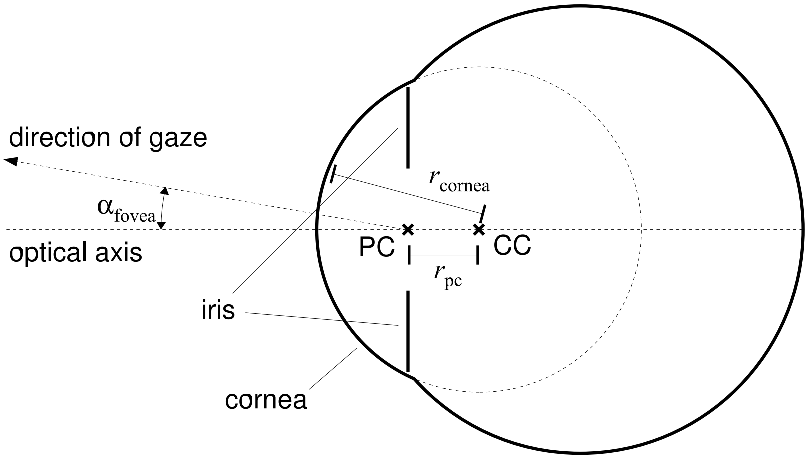

Figure 2. Eye model used in the remote gaze estimation algorithm. PC: Pupil centre. CC: Cornea centre. r_{cornea}, r_{pc}, α_{fovea}: User-dependent parameters (see text for explanation).

The image processing component is based on the Starburst algorithm (Li et al., 2005), which was reimplemented and modified to fit the needs of the remote eye tracking setting. The gaze estimation component is based on a physical model of the eye (see Figure 2), which models the optical properties of the cornea (reflection and refraction), the location of the pupil centre (PC) and centre of corneal curvature (CC), and the offset of the fovea (and hence the line of sight) from the optical axis. The model contains three user-dependent parameters: the curvature radius of the cornea (r_{cornea}), the distance between PC and CC (r_{pc}), and the offset of the direction of gaze from the optical axis (α_{fovea}) (only the horizontal component of this offset is currently modelled).

Given the observed positions of the CRs and of the pupil centre in the camera image, there is only one possible position and orientation of the eyeball that could have given rise to these observations. The gaze estimation algorithm deduces this position and orientation, then intersects the direction of gaze with the display plane to determine the location the user is fixating.

Preliminary measurements on the system have shown an average accuracy of 1.2 degrees; with additional fine-tuning, we expect to improve the accuracy to better than 1 degree. The system allows head movements of 20 cm between the extremes of the working range on all three spatial axes.

Directions for Future Development

Despite the advances in remote eye tracking systems in recent years, there are still quite a number of areas in which improvements have to be made if these systems are to see widespread use in human-computer interfaces (including, but not limited to, AAC applications). The aspects that require further progress include robustness, accuracy, ease of setup and use, and price.

In the following, we will present ideas that we intend to pursue in the future to achieve advances in the areas mentioned above:

- Tolerance towards glasses. Systems that use infrared illumination often do not work well for users who wear glasses because of reflections on the surface of the glasses. The existing systems can usually be made to work with glass wearers to a certain extent, but only for some head orientations where no interfering reflections occur. For other head orientations, the reflections can obscure the user's eyes completely, making eye tracking impossible. One way of dealing with this problem might be to use more than two infrared illuminators. At any given time, the system would use two of the illuminators. If the system detected that the user's pupils were being obscured by reflections, it would switch to a different set of illuminators at a different angle relative to the user and the camera. In this way, the reflections should shift off the eyes or even be eliminated entirely.

  To achieve high accuracy in the presence of glasses, the eye model may have to be augmented with a model of the glasses to account for their effect on the image of the eye. However, preliminary tests indicate that the accuracy is still tolerable even if the effect of the glasses is not modelled.

- Ease of setup / use. Remote eye tracking systems are typically based on a physical model of the eye, the eye tracking system (camera and illuminators), and the monitor. Because of this, they require the spatial relationship between the camera, the illuminators, and the monitor plane to be known. These measurements are usually obtained by hand, a process that is time-consuming, error-prone, and difficult for an end user to carry out. Beymer and Flickner (2003) calibrated the orientation of the monitor plane automatically using a mirror to reflect the image of a checkerboard pattern taped to the monitor back into the camera. We intend to implement a similar automatic calibration in our system.

- Price. Existing remote eye trackers typically use high-resolution industrial cameras with relatively high-grade lenses. This makes the systems quite expensive, even before labour costs for assembly are taken into account. For example, the camera and lens used in our eye tracker have a combined price of around 1000 USD. This puts the system out of reach of many potential users. An alternative would be to use webcams, but we are sceptical if their typical resolution of 640x480 pixels can deliver satisfactory results. Instead, we are confident that advances in sensor hardware will solve this particular problem and that sensors with the required resolution will soon reach consumer price points.

  Another obvious idea for cutting cost is to eliminate the infrared illuminators and use natural illumination (see e.g. Hansen and Pece (2005)), though this makes the image processing task significantly more difficult.

- 3D cameras. Recent years have seen the development of so-called 3D time-of-flight (TOF) cameras (CSEM, 2006). In addition to providing an intensity image like a conventional camera, these cameras also provide a depth image that gives the distance of the object in the scene at each pixel. This allows the three-dimensional shape of the scene, e.g. the user's head, to be reconstructed. Many recognition and tracking tasks can be implemented more robustly on 3D range data than on intensity images, and so this technology has the potential to be used for robust head and eye tracking. Two participants in COGAIN, together with other European partners, will be working on TOF-based eye, head and gesture tracking within an EU project.

Robust, affordable eye tracking technology would have a broad range of potential applications. It would of course be invaluable for AAC applications, but beyond that, eye tracking has the potential to become a new general-purpose interaction medium. Eye tracking may change the way we interact with technology and how visual information is communicated – our work on gaze guidance (Barth et al., 2006) has the goal of augmenting a video or visual display with a recommendation of how to view the information, of what is to be seen.

With the advances currently being made in eye tracking hardware and software, widespread low-cost eye tracking may finally become a reality.

Acknowledgments

Research was supported by the German Ministry of Education and Research (BMBF) under grant number 01IBC01 with acronym ModKog and by the European Commission within the COGAIN Network of Excellence. We thank SensoMotoric Instruments GmbH, Germany, for their support.

References

Barth, E., Dorr, M., Böhme, M., Gegenfurtner, K.R., and Martinetz, T. (2006) Guiding the mind's eye: improving communication and vision by external control of the scanpath, *Human Vision and Electronic Imaging*, Volume 6057 of *Proceedings of the SPIE*, pp. 116–123.

Beymer, D. and Flickner, M. (2003) Eye gaze tracking using an active stereo head, *Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition (CVPR)*, Volume 2, pp. 451–458.

Brolly, X.L.C. and Mulligan, J.B. (2004) Implicit calibration of a remote gaze tracker, *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW '04)*, Volume 8, p. 134.

CSEM (2006). SwissRanger SR-3000, CSEM SA, Zurich, Switzerland.

Guestrin, E.D. and Eizenman, M. (2006) General Theory of Remote Gaze Estimation Using the Pupil Center and Corneal Reflections, *IEEE Transactions on Biomedical Engineering*, Volume 53, Number 6, pp. 1124–1133.

Hansen, D.W. and Pece, A.E.C. (2005) Eye tracking in the wild, *Computer Vision and Image Understanding*, Volume 98, Issue 1, Elsevier, pp. 155–181.

Hennessey, C., Noureddin, B., and Lawrence, P. (2006) A Single Camera Eye-Gaze Tracking System with Free Head Motion, *Eye Tracking Research & Applications 2006*, pp. 87-94.

LC Technologies (2006) Eyegaze Eyefollower, LC Technologies Inc., Fairfax, VA, USA. `http://www.eyegaze.com`

Li, D.,Winfield, D., and Parkhurst, D.J. (2005) Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches, *Proceedings of the IEEE Vision for Human-Computer Interaction Workshop at CVPR 2005*, pp. 1–8.

Meyer, A., Böhme, M., Martinetz, T., and Barth, E. (2006) A single-camera remote eye tracker, *Perception and Interactive Technologies*, Volume 4021 of *Lecture Notes in Artificial Intelligence*, Springer, pp. 208–211.

Morimoto, C.H., Amir, A., and Flickner, M. (2002) Detecting Eye Position and Gaze from a Single Camera and 2 Light Sources, *Proceedings of the International Conference on Pattern Recognition 2002*, pp. 314–317.

Ohno, T. and Mukawa, N. (2004) A free-head, simple calibration, gaze tracking system that enables gaze-based interaction, *Eye Tracking Research & Applications 2004*, pp. 115–122.

Shih, S.W., Wu, Y.T., and Liu, J. (2000) A calibration-free gaze tracking technique, *Proceedings of the 15th International Conference on Pattern Recognition*, pp. 201–204.

SMI (2006) Remote eye tracking system (announced), SensoMotoric Instruments GmbH, Teltow, Germany. `http://www.smi.de`

Tobii (2002) Tobii 1750 eye tracker, Tobii technology AB, Stockholm, Sweden. `http://www.tobii.se`

Yoo, D.H. and Chung, M.J. (2005) A novel non-instrusive eye gaze estimation using crossratio under large head motion, *Computer Vision and Image Understanding*, Volume 98, Issue 1, Elsevier, pp. 25–51.