Robust Gaze Estimation in Low-Resolution Eye Images

Axel Schroth (axel.schroth@student.kit.edu)[1], Sascha Voth[2], Tobias Schuchert[2]

[1]Karlsruhe Institute of Technology
[2]Fraunhofer Institute of Optronics, System Technologies and Image Exploitation

Abstract

A Head Mounted Display (HMD) with a bidirectional chip is an emerging technology with many applications enabling us to display information to users while simultaneously analyzing eye movement [Baumgarten et al. (2012). Aspects of a head-mounted eye-tracker based on a bidirectional OLED microdisplay. *Journal of Information Display*, 13, pp. 67-71]. However, due to the high density of light emitting and capturing elements, the eye image suffers from noise artefacts and low-resolution output (128 × 96 pixel). Two topics are addressed by this work: low resolution eye tracking and comfortable calibration.

Robust eye tracking includes detection of corneal reflections, classification of closed eyes and pupil extraction. Therefore, several different approaches are compared, an extended approach presented and evaluated. Robustness of the proposed approach is increased by incorporating temporal information. A final evaluation shows that implemented extensions improve accuracy significantly.

A calibration process is required to estimate gaze points of a user. The second part of this work is about evaluating different ways of calibration methods and presenting novel methods which are more comfortable. The standard multiple point calibration procedure is amended by using automatic methods and animated objects. Finally an implicit calibration by reading text is presented, which makes the disturbing calibration procedure almost invisible to the user with only slight degradiation of accuracy.

Method

The main focus of this work lies on efficient gaze point estimation with low-resolution and low-contrast images. The algorithms, tests and evaluations are designed to work with the interactive bidirectional HMD shown in Figure 1. In order to estimate gaze, the eye image is processed to estimate pupil center position and detecting corneal reflections (CRs) caused by infrared LEDs attached to the HMD device. We are comparing different approaches for calibration: manual 12-point calibration, calibration with moving objects and implicit calibration by reading text.

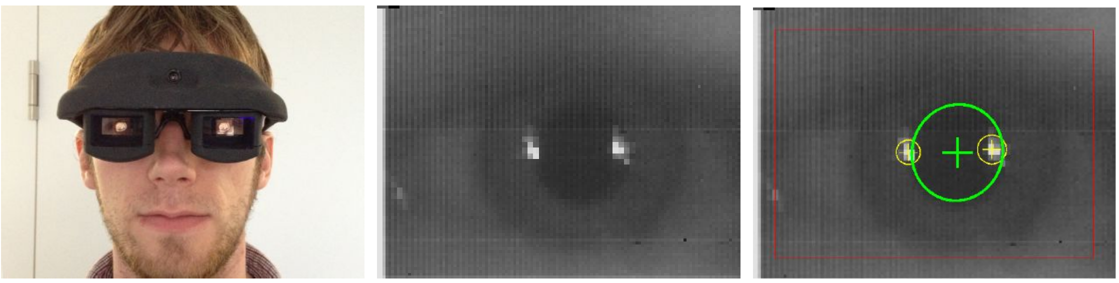

Figure 1: A bidirectional HMD (left) and a sample frame from the camera (center). The right image shows the same frame with detected CRs (yellow circles) and pupil (green circle).

As one can see in Figure 1, the input data captured by the device suffers from low resolution, low contrast and noise. These quality drawbacks are due to the architecture of the bidirectional HMD, where light capturing elements are located in between light emitting diodes. Brightness changes of the environment have high impact on the captured images, especially regarding contrast. Detection of CRs is challenging, because the shape of a CR is irregular and can vary dependent on the eyeball rotation angle. Goal of this work is to overcome these issues and create a robust eye tracker. We want to detect closed eyes, locate reflections reliably and extract the pupil ellipse while being as invariant as possible against lighting changes.

The image processing pipeline starts with a preprocessing step reducing noise artefacts. Subsequently the reflections are located with subpixel accuracy based on a-priori knowledge of intensity, intensity gradient, proximity and location of both detected reflections. If no reflections can be detected, the eye is considered as closed. The most important part of the eye tracker is the pupil center detection algorithm. Several state of the art approaches for low-resolution images have been investigated and evaluated. Given the close up, low-resolution image of the eye region, the pupil is visible as a dark ellipse. The Gradient Direction Consensus (GDC) [Droege, D. & Paulus, D. (2010). Pupil center detection in low resolution images. *Proceedings of ETRA '10*, pp. 169172] algorithm first detects relevant gradients at the edges of iris and pupil. Regarding gradients as straight lines, a common center point can be estimated. Another approach, as exposed in [Valenti, R. & Gevers, T. (2012). Accurate Eye Center Location through Invariant Isocentric Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9), pp. 1785-1798], looks for isocentric patterns in an eye image. The pupil center is then estimated by accumulating votes given by their isocenters.

In order to improve accuracy and determine the pupil extent, the estimation of the pupil center is taken as a starting point for an intensity based floodfill segmentation. The border pixels of the segmented area are filtered based on detected CRs and gradient values. Finally the pupil ellipse is estimated using RANSAC. Furthermore, we use temporal information of the previously detected pupil in order to adjust weights of the border points for RANSAC fitting. A Kalman filter smoothens results and an additional outlier detection discards extreme false detections.

The calibration procedure was designed to be as seamless as possible while keeping its accuracy. The first step towards reaching this goal is calibration by an animated object. An object is drawn in the HMD display and moves to several calibration points where an animation is attracting attention by animation. The user is expected to look at the center of the object image. During the whole process, the captured pupil position is associated to the current position of the object. Considering these associated points, a homography matrix is calculated used to determine gaze points on the HMD display while tracking the eye. Another approach considers an implicit calibration using a small text paragraph. The characters build up successively expecting the user to look at the end of the text. All tracking points are associated to the mass center of the last character, which is used to calculate the homography matrix.

Results

The evaluation results as shown in Table 1 are based on input images annotated manually. More than 300 eye images captured by the device have been selected. The images are recorded from eight test subjects under different conditions (time, illumination of the environment, extreme eye positions, partly closed eyes).

Table 1: Results of the eye tracker (left) and the calibration evaluation (right).

Pupil Detection	GDC	Isocentric Patterns	Extended GDC
Pupil Center Deviation [px]	3.73	4.50	2.42
Pupil Radius Deviation [px]	n/a	n/a	2.06
Elapsed Time [ms]	2.49	4.90	13.935

Calibration	Mean deviation
12 point	1.61°
9 point	3.76°
Moving object	1.84°
Implicit Text	3.64°

Since the calibration implementation is still in development, only preliminary results can be shown in Table 1. The deviations are calculated by measuring the mean deviation of gaze points to a shown point on the HMD display. The total mean deviation is based on the deviations of 12 reference points equally distributed over the display area.

Conclusion

A robust algorithm for eye tracking in low-resolution images as a result of an extensive evaluation has been developed in context of this work. Several improvements using an additional processing step based on floodfill segmentation with RANSAC fitting and consideration of temporal information lead to better accuracy regarding pupil estimation. Combined with a calibration procedure, gaze points can be estimated and used for further applications like detecting points of interest.

Additionally, the immersive experience can be improved by the usage of intuitive, automatic or even implicit and invisible calibration procedures. These methods are currently in development, but showed already good results compared to common calibration procedures like a 12-point calibration. Some final evaluations are in progress in order to prove acceptance and to show experienced impressions by users.