Gaze interaction for multi-display systems using natural light eye-tracker

Onur Ferhat, Arcadi Llanza, Fernando Vilariño

1 Introduction

Natural light-based gaze-tracking constitutes an interesting alternative to commercial eye-trackers, and it requires no additional hardware other than a camera (ubiquitously present on consumer devices). Methods developed to attack this problem include using raw eye appearance as the feature set [1], using the geometry of iris boundary ellipse for estimation [2, 3, 4] and using more complicated geometrical models [5, 6], among others. The problem with these methods is either they are using unrelated information (skin pixels), or they are summarising the necessary information too much (to model parameters).

We propose a system implementing a novel feature based approach for webcam gaze interaction. Our system relies on the automated detection of the iris bounding box, followed by iris segmentation by means of binarisation. We develop this approach using the OpenGazer eye-tracker as a baseline.

As a validation framework, we propose a two-display context-adaptive system which modifies in real time the data rendered in a secondary display depending on the particular object that is observed in the main display. We particularly tune this approach to Google-glass type devices. The application context of our research can be straightforwardly tested and extrapolated to different augmented reading expe-

Onur Ferhat
Universitat Autònoma de Barcelona and Computer Vision Center, Edificio O Campus UAB 08193 Bellaterra e-mail: oferhat@cvc.uab.es

Arcadi Llanza
Universitat Autònoma de Barcelona and Computer Vision Center, Edificio O Campus UAB 08193 Bellaterra e-mail: arcadi.llanza@e-campus.uab.cat

Fernando Vilariño
Universitat Autònoma de Barcelona and Computer Vision Center, Edificio O Campus UAB 08193 Bellaterra e-mail: fernando@cvc.uab.es

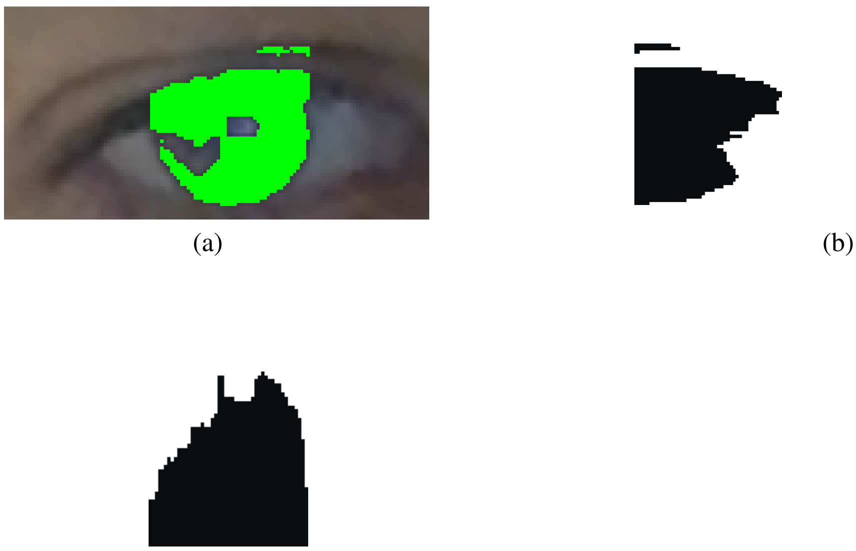
Published at SAGA Workshop 2015, proceedings archived under DOI: 10.2390/biecoll-saga2015_0.
Individual article archived under DOI: 10.2390/biecoll-saga2015_4

riences such as maps analysis, artworks detailed observation and music reading, among others.

2 Our Approach

The system performs an **automatic selection** of 8 anchor points on the subject's face based on facial features such as eye corners, eyebrows, nose and mouth [7, 8]. Over the next frames, optical flow is used to update the points' positions in the new images.

The **feature extraction** stage takes as input the eye images extracted from the camera image. The iris bounding box is detected with template matching and iris segmentation is calculated using a binarisation inside this area. The features are extracted by projecting the segmented pixels into the vertical and horizontal axes and a feature vector is created by concatenation as seen in Fig. 1. For **gaze estimation**, we train a Gaussian Process with the standard squared exponential kernel tuned to work at a desired distance during the calibration process.



(a) (b) (c)

Fig. 1 (a) Iris segmentation (b) Vertical features calculated by projecting iris pixels to vertical axis (c) Horizontal features

The system is complemented with a gaze-interactive interface as shown in Fig. 2. On the first display, visual selection is natively performed by guiding the gaze towards the desired object (left). Real-time output is then rendered in the secondary display (top-right) which adapts its information to the item visually selected (in the context of Google-glass type devices the secondary display plays the role for the glass display).

Published at SAGA Workshop 2015, proceedings archived under DOI: 10.2390/biecoll-saga2015_0.
Individual article archived under DOI: 10.2390/biecoll-saga2015_4

Fig. 2 System interface: (Left) Main visualization display where stimuli are shown. (Top-Right) Responsive display: The information is adapted to the object visualized. (Bottom-right) Subject's facial reference points and segmented images from the eyes shown for illustrative purposes.

3 Discussion

The proposed method decreases the average gaze estimation error by 34% in horizontal direction and by 12% in the vertical direction when compared to our previous work on standard databases [8]. The average errors are **2.35°and 1.82°** for the baseline system in both directions; whereas the proposed system achieves error rates of **1.54°and 1.61°**. Our experiments show that the errors are not uniform over the screen area, and horizontal errors can be up to 65% larger near the edges compared to the center of the screen.

The interaction obtained in this way is smooth and fast. The preliminary results are promising, and our current work is focused on **enriching the feature set** by introducing spatial feature correlation, and using other cues such as exact eye corner positions. Regarding interaction, the definition of suitable visual action triggers that handle with the information delivered in the secondary display, not defined in this work, is an open area of research.

ACKNOWLEDGMENTS: This work was supported by UAB grants and the Google Faculty Award.

Published at SAGA Workshop 2015, proceedings archived under DOI: 10.2390/biecoll-saga2015_0.
Individual article archived under DOI: 10.2390/biecoll-saga2015_4

References

1. Holland C, Komogortsev OV (2012). Eye tracking on unmodified common tablets: challenges and solutions. Eye Tracking Research and Applications. 277280.
2. Wood E, Bulling A (2014). EyeTab: Model-based gaze estimation on unmodified tablet computers. Eye Tracking Research and Applications. 207-210.
3. Wang JG, Sung E (2000). Gaze determination via images of irises. British Machine Vision Conference.
4. Wu H, Chen Q, Wada T (2004). Conic-based algorithm for visual line estimation from one image. Automatic Face and Gesture Recognition. 260265.
5. Chen J, Ji Q (2000). 3D gaze estimation with a single camera without IR illumination. International Conference on Pattern Recognition. 14.
6. Ishikawa T, Baker S, Matthews I, Kanade T (2004). Passive driver gaze tracking with active appearance models. World Congress on Intelligent Transportation Systems.
7. Santana MC, Déniz-Suárez O, Hernández-Sosa D, Lorenzo J (2011). A comparison of face and facial feature detectors based on the Viola-Jones general object detection framework. Mach. Vis. Appl. 22(3), 481494.
8. Ferhat O, Vilarino F, Sanchez FJ (2014). A cheap portable eye-tracker solution for common setups. Journal of Eye Movement Research, 7(3):2, 1-10.

Published at SAGA Workshop 2015, proceedings archived under DOI: 10.2390/biecoll-saga2015_0.
Individual article archived under DOI: 10.2390/biecoll-saga2015_4