Using Corneal Imaging for Measuring a Human's Visual Attention

Christian Lander
DFKI
Saarland Informatics Campus
Stuhlsatzenhausweg 3
66123 Saarbruecken, Germany
christian.lander@dfki.de

Frederik Wiehr
DFKI
Saarland Informatics Campus
Stuhlsatzenhausweg 3
66123 Saarbruecken, Germany
frederik.wiehr@dfki.de

Felix Kosmalla
DFKI
Saarland Informatics Campus
Stuhlsatzenhausweg 3
66123 Saarbruecken, Germany
felix.kosmalla@dfki.de

Sven Gehring
DFKI
Saarland Informatics Campus
Stuhlsatzenhausweg 3
66123 Saarbruecken, Germany
sven.gehring@dfki.de

Abstract
The human's visual attention focus usually reflects the activity and environment engaged in or simply his context in the most coherent way. The cornea, which is encasing the iris and pupil, is protected by tear fluid, and thus is a highly reflective surface. Our eyes show a reflection of what we see in our current context. We present an attempt of using corneal imaging to extract contextual information – including objects in the field of view - and the user's attention focus. Our system uses a head-mounted eye-camera, for capturing corneal reflections, connected to a RaspberryPi, to record a humans' current view and estimate his focus of attention. We conducted a 2-day experiment, where we collected data in un-instrumented real-world settings. Based on the analysis of the recordings we illustrate what kind of information can be extracted out of the corneal image reflections and outline the possibilities for in-situ notifications.

Author Keywords
Corneal reflection, notifications, computer vision, visual attention, attention detection.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

UbiComp/ISWC'17 Adjunct, September 11–15, 2017, Maui, HI, USA

ACM ISBN 978-1-4503-5190-4/17/09.
https://doi.org/10.1145/3123024.3124563

Introduction

The human gaze has been subject to research as an input technique for human computer interaction for more than two decades [16]. It was shown to be faster than any other input devices (e.g., mouse [19]). Gaze provide a great indication of what is attracting us and might be of interest [22]. Use cases for gaze estimation vary widely from cursor control to eye-typing [10] and target selection [18]. Furthermore, our eye movements are connected to the activities we are engaged in as well as the cognitive processes [3, 18], including visual memory recall or cognitive load [4, 21]. Selected personality traits [9] that are difficult if not impossible to assess using other sensing modalities can also be sensed with gaze.

The human eye is great source for determining the current context of person and what he is paying attention to. The visible parts of the human eye are the white sclera, the iris, and the black pupil. The cornea is encasing iris and pupil. The area is protected by tear fluid that turns it into a mirroring surface. A high-resolution camera, placed in front of the eye (slightly underneath), is able to obtain a distorted partial reflection of a person's current field of view (see Figure 1).

In this paper we report on an approach to use the potentials of corneal imaging, moving towards pervasive eye tracking [2] with the goal of determining a user's current context and focus of attention for providing in-situ context-sensitive notifications (ambient and explicit). Therefore, we use 2-day recordings of corneal imaging data in everyday scenarios using a RaspberryPi to record the data. We derive contextual information from the collected imagery in order to outline the potential of providing in-situ notifications in the user's current field of view. While prior work mainly focussed on creating accurate gaze tracking techniques [12,14], we focus on exploiting the corneal reflection for attention measurement.

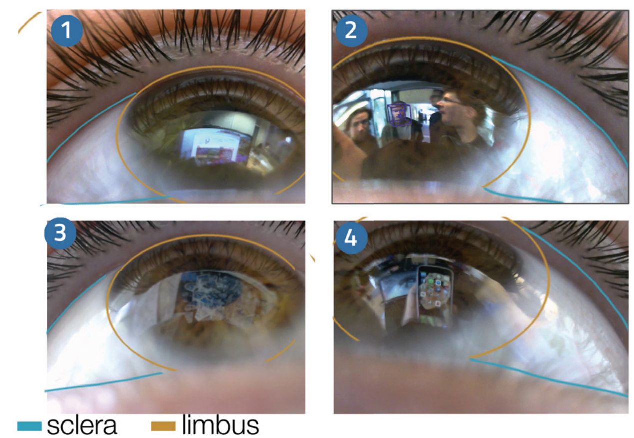

Figure 1 Corneal Images, showing the reflection of (1) a computer monitor, (2) faces, (3) a poster and (4) an iPhone6 display.

To capture the eye and to estimate the gaze and detect objects, we use a lightweight mobile system for gaze estimation [12] in combination with computer vision methods for object detection. Our prototype consists of a head-mounted eye-camera for capturing corneal reflections, connected to a RaspberryPi for recording and analyzing the data. We conducted a life logging experiment, collecting data in an un-instrumented real-world setting. Over a period of two days.

Related Work

Corneal Reflection

Backes et al. [1] found that sensitive information is accessible through reflections on the cornea (e.g., typing in passwords). Nitschke et al. [13, 14] used corneal imaging to calibrate display-camera setups and to track a person's point of gaze. Schnieders et al. [17] built a system to estimate the gaze on a display reflected on the cornea with a remote camera. Takemura et al. [20] brought the technology of corneal imaging into a head-worn device. Nakazawa et al. [12] also developed a head-mounted device to capture corneal reflections and achieved reasonable results for gaze estimation on a 23-inch display.

Visual Attention

Research on visual attention is combining psychophysical, neuroscientific and engineering science research. In general, attention is the ability to focus on a perceived input from the environment [6]. Humans can only focus on a small part in the center of their visual field of view. Rensink et al. [15] show the effect of change blindness [11], where persons did not detect changes in the scene. Our work aims at detecting objects [23] in the focal as well as the peripheral areas. With this, we can measure the visual attention on brand icons and advertisements [5,7], or distracting objects while driving a car [8]. Using the existing infrastructure (i.e., in a car) or personal devices (i.e., wearables, smartphones, etc.), with our approach we can provide the user with digital information such as notifications that are (1) relevant for his current focus of attention as well as (2) adapted in their presentation form to the current context (i.e., audio signals while driving instead of visual stimuli).

We extend existing work by linking the corneal reflection images with information of the user's environment. In addition we use a simple and fast implementation of a head-worn corneal imaging device. We extended the device by computer vision and image processing techniques detect and track objects in the user's field of view.

Prototype

Our prototype is built off-the-shelf hardware (see Figure 2): a head-mounted corneal imaging system for calibration free gaze estimation on displays using a RGB webcam mounted onto a 3D printed frame, a RaspberryPi as well as an external battery pack. The camera is mounted below the eye, and is slightly rotatable and movable to record a close-up video of the eye (1280 x 960 pixels at 30 fps, 60° fov, 4mm fixed focus). The frames display the reflection of the users field of view on the eye (as can be seen in Figure 1).

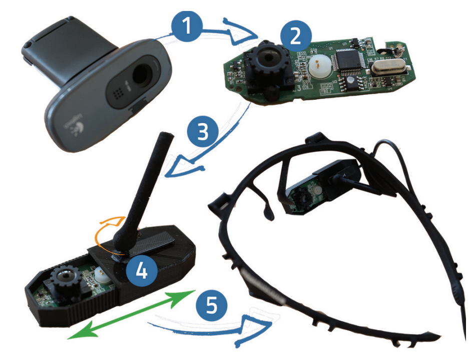

Figure 2 Hardware prototype: The camera's original housing is removed (1). Glue around the lens (2) is removed to adjust focus. (3) The camera is built into a custom mount – rotatable and movable – and (4) mounted at a glasses frame.

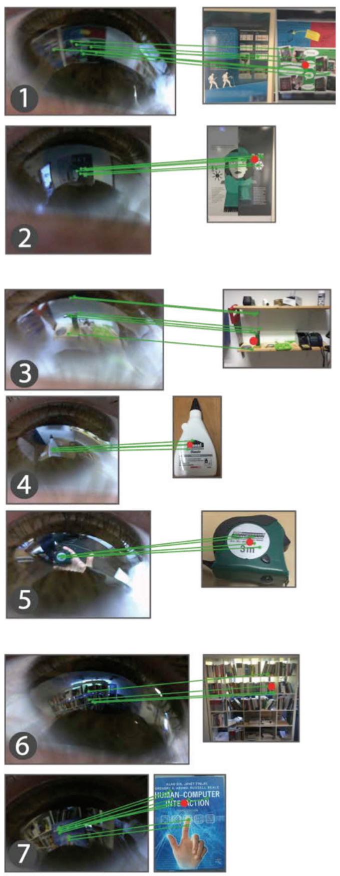

Figure 3 Gaze Estimation on posters (1,2) using the feature tracking algorithm. Gaze Estimation on bookshelf (6) and a book (7); on real world objects (3,4,5).

In addition to computer vision algorithms to detect real- world object in the captured video, our system also supports for calibration-free gaze estimation using natural feature tracking. With this, we can detect objects as well the user's current focus of attention.

Data Collection and Use Cases

We conducted an in-the-wild experiment using our system in the daily routines of a 28 years old male university student. He was wearing the device for approximately 11 hours in total on two consecutive days. The video data of his eye was recorded with 30 fps. As our intention was to detect objects in the reflection on the eye, we tried to record as much video data as possible instead of recording as many users as possible.

To prevent data loss, the captured camera streams were analyzed post-hoc. Therefore, we manually examined the recorded video material and selected different types of scenes. As the material contained many settings sitting in front of a computer monitor, we reduced the amount of data by removing most of them. The analysis was done on the pre-processed camera frames, containing the region within the iris contour (limbus). For object detection, we used a small set of template images (found with Google image search) as well as self-taken images of items possibly visible in each scene. In the following we will present the main findings by showing the processed recordings and possible future applications of the system.

We are able to detect objects the user is looking at as well as to estimate the user's gaze. Figure 3 depicts the object detection and gaze estimation for real-world objects using natural feature tracking. Furthermore, our system is able to detect when the user is looking at a digital display (see Figure 4). Such information could be used to detect if a person is recognizing a display currently passing by. The display attention measured via display detection can furthermore be relevant from a security perspective since the system if someone is looking at a display showing private or sensitive information or not.

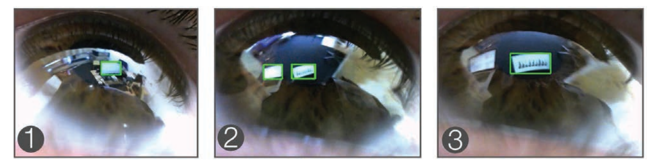

Figure 4 Display Detection and Tracking (1)-(3) in images reflected on a person's eye; detected dis- plays are marked with a green rectangle.

When driving a car, usually the street is in our visual focus. Figure 5 illustrates the detection of brand logos, road signs as well as whether the driver is focusing on the traffic or not. Since visual attention is usually associated with the point of fixation [6], such information lying in the peripheral region are recorded by the system. While driving an unknown route, a suggestion can be made, that several retailers are nearby where items for the daily shopping can be purchased or furthermore, if the speeding limit is detected from a road sign, the driver could be notified if he is speeding. With respect to this, corneal imaging could be added to existing car assistance systems to make driving safer.

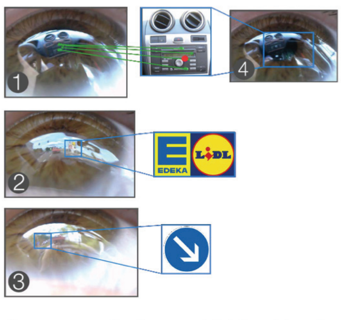

Figure 5 Gaze Estimation on car hifi (1), and detection of retailer logos (2) and road sign (3).

Discussion and Conclusion

We recorded corneal imaging data over the course of consecutive days and we analyzed the recordings post-hoc to demonstrate the possibilities of detecting a human's visual attention as well as the situational context he's in just from the reflection on his eyes. We briefly outlined opportunities for exploiting such data. However, further investigations and implementations are needed to improve our system in order to provide context-adaptive notifications to a user in real world settings. However, our approach constitutes a first step towards this. We showed that we can already detect what a person's attention is focused on as well as relevant objects in the peripheral field of view that might not be perceived just like that.

The methods of gaze estimation and object detection are based on computer vision algorithms that need a priori information about most of the objects (e.g., brand logos, book covers, signs). We believe, that a connection to a large image database for specific environments (e.g., Flickr, Google Goggles) or frameworks like Google's Tensor Flow can enhance the performance.

In this paper we presented an approach to investigate corneal image analysis in everyday scenarios. We outlined the possibilities of determining a person's current situational context as well as his focus of attention. In our experiment, we gained first insights in the used concepts, and that they are valuable for pervasive usage of the system.

References

1. Michael Backes, Tongbo Chen, Markus Dürmuth, Hendrik P.A. Lensch, and Martin Welk. 2009. Tempest in a teapot: Compromising reflections revisited. In *IEEE Symp. on Security and Privacy 2009*.
2. Andreas Bulling and Hans Gellersen. 2010. Toward Mobile Eye-Based Human-Computer Interaction. *IEEE Pervasive Computing* 9, 4 (2010), 8–12.
3. Andreas Bulling, Jamie A. Ward, Hans Gellersen, and Gerhard Tröster. 2011. Eye Movement Analysis for Activity Recognition Using Electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 4 (April 2011), 741–753
4. Andreas Bulling and Daniel Roggen. 2011. Recognition of Visual Memory Recall Processes Using Eye Movement Analysis. In *Proc. UbiComp 2011*. 455–464.
5. Pierre Chandon, J. Wesley Hutchinson, Eric T. Bradlow, Scott H. Young. 2009. Does In-Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase. *Journal of Marketing* 73, 6 (2009), 1–17.
6. Andrew T. Duchowski. 2007. *Eye Tracking Methodology: Theory and Practice*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
7. Kendall Goodrich. 2010. What's up? Exploring upper and lower visual field advertising effects. *Journal of Advertising Research* (2010).
8. Yi-Ching Lee, John D. Lee, and Linda Ng Boyle. 2007. Visual attention in driving: The effects of cognitive load and visual disruption. *Human Factors* 49, 4 (2007), 721–733.
9. Sabrina Hoppe, Tobias Loetscher, Stephanie Morey, and Andreas Bulling. 2015. Recognition of Curiosity Using Eye Movement Analysis. In *Proc. UbiComp 2015*. 185–188.

10. Päivi Majaranta and Kari-Jouko Räihä. 2002. Twenty years of eye typing: systems and design issues. In *Proc ETRA 2002*. ACM, 15–22.

11. J. Müller, D. Wilmsmann, J. Exeler, M. Buzeck, A. Schmidt, T. Jay, and A. Krüger. 2009. Display Blindness: The Effect of Expectations on Attention Towards Digital Signage. In Proc Pervasive 2009 (Pervasive '09).

12. Atsushi Nakazawa, Christian Nitschke, and Toyoaki Nishida. 2015. Non-calibrated and real-time human view estimation using a mobile corneal imaging camera. In *IEEE ICMEW 2015*. IEEE, 1–6.

13. Christian Nitschke, Atsushi Nakazawa, and Haruo Takemura. Corneal imaging revisited: An overview of corneal reflection analysis and applications. *IMT 8*, 2 (2013), 389–406.

14. Christina Nitschke, Atsushi Nakazawa, and Toyoaki Nishida. 2013. I See What You See: Point of Gaze Estimation from Corneal Images. In *Proc. ACPR 2013*. 298–304.

15. Ronald A. Rensink, J. Kevin O'Regan and James J. Clark. 1997. To see or not to see: The need for attention to perceive changes in scenes. *Psychological science* 8, 5 (1997), 368–373.

16. Robert J.K. Jacob. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proc CHI 1990*. ACM, 11–18.

17. Dirk Schnieders, D., Xingdou Fu and K.-Y. K. Wong. Reconstruction of display and eyes from a single image. In *CVPR*, IEEE Computer Society (2010), 1442–1449.

18. Julian Steil and Andreas Bulling. 2015. Discovery of Everyday Human Activities From Long-term Visual Behaviour Using Topic Models. In *Proc. UbiComp 2015*. 75–85.

19. Sophie Stellmach and Raimund Dachselt. 2012. Look & Touch: Gaze-supported Target Acquisition. In *Proc CHI 2012 (CHI '12)*. ACM, New York, NY, USA, 2981–2990.

20. Kentaro Takemura, Tomohisa Yamakawa, Jun Takamatsu and Tsukasa Ogasawara. (2014). Estimation of a focused object using a corneal surface image for eye-based interaction. *Journal Of Eye Movement Research, 7*(3).

21. Bernd Tessendorf, Andreas Bulling, Daniel Roggen, Thomas Stiefmeier, Manuela Feilner, Peter Derleth and Gerhard Tröster. 2011. Recognition of Hearing Needs From Body and Eye Movements to Improve Hearing Instruments. In *Proc. Pervasive 2011*. 314–331.

22. Roel Vertegaal. 2003. Attentive user interfaces. In Communications of the ACM. 30–33.

23. Kentaro Takemura, Tomohisa Yamakawa, Jun Takamatsu and Tsukasa Ogasawara. (2014). Estimation of a focused object using a corneal surface image for eye-based interaction. *Journal Of Eye Movement Research, 7*(3).