My Reading Life – Towards Utilizing Eyetracking on Unmodified Tablets and Phones

Kai Kunze
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
kunze@m.cs.osakafu-u.ac.jp

Yuzuko Utsumi
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
utsumi@cs.osakafu-u.ac.jp

Shoya Ishimaru
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
ishumaru@m.cs.osakafu-u.ac.jp

Koichi Kise
Graduate School of Engineering
Osaka Prefecture University
1-1 Gakuen-cho, Naka, Sakai,
Osaka, Japan
kise@cs.osakafu-u.ac.jp

UbiComp'13 Adjunct, September 8–12, 2013, Zurich, Switzerland.
ACM 978-1-4503-2215-7/13/09.
http://dx.doi.org/10.1145/2494091.2494179

Abstract
As reading is an integral part of our knowledge lives, we should know more about our reading activities. This paper introduces a reading application for smart phone and tablets that aims at giving user more quantified information about their reading habits. We present our work towards building an open library for eye tracking on unmodified tablets and smart phones to support some of the applications advanced functionality. We implemented already several eye tracking algorithms from previous work, unfortunately all seem not to be robust enough for our application case. We give an overview about our challenges and potential solutions.

Author Keywords
reading, knowledge, eye tracking, mobile devices, eye gaze

ACM Classification Keywords
H.5.2 [User Interfaces]: Input devices and strategies; H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

Introduction
The more somebody reads the higher seems to be language skills and general knowledge [6]. People can increase their physical fitness and reduce their weight by step counting and food logging [3]. So it seems simple

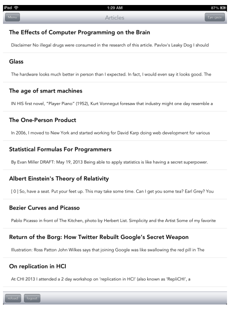

Figure 1: Article view of the "My Reading Life" application (auto-scrolling can be done using eye gaze).

Figure 2: Sample summary statistics: average number of articles read, the genre, reading speed, popular sources, comparison with friends, and reading locations.

tracking of activities raises awareness and leads to healthier habits. We therefore wonder if we can give people tools to better track their reading and thus improve their cognitive life. Unfortunately there are very few efforts to recognize reading activities in realistic environments [5, 14, 12].

We are driven by the research question, how to monitor and improve knowledge acquisition tasks with a special focus on reading.

The contributions in this paper are outlined as follows. (1) We introduce a prototype of a smart phone application called "My Reading Life", that takes articles from a popular read later web service and provides you with summary of your reading using statistics over UI interactions and our eye gaze library. (2) We show how we obtain several statistics about the users reading habits (using eye position and head angle to estimate if the user is looking on the device), discuss challenges and potential solutions. (3) We use a combination of standard computer vision techniques for gaze estimation to implement several eye tracking methods on an unmodified tablet and smart phone (iPad 4 and iPhone 5) with the goal to provide an open standard eye gaze library for researchers and developers. Unfortunately, the performance depends so far highly on the calibration, the user and lighting conditions.

Approach

The most profound, yet a very difficult and long way to track reading habits is to equip the user with a dedicated device (e.g. eye tracker and scene camera) to record all reading interactions from the users point of view.

In this paper we go a different route, we want to use digital devices to track the users reading habits. More and more reading is done on them and they have some unique advantages, as we can easily obtain additional meta information about the reading material at hand without using expensive machine learning or computer vision techniques (e.g. the text without using OCR). We envision a Reading Service similar to a Location Service built-in to all kinds of digital devices, enabling the user to track their reading.

Implementation

To see if we are able to provide users with enough quantifiable information to change their reading habits, we implemented a prototype application on the iOS platform. To supply it with reading material, we use a popular read later service `http://readability.com` for website articles. We import the reading list and let users view the articles in a standard view, see Fig 1.

Tracking Reading Habits

We give the user feedback over the following statistics: the number of articles read, the average reading speed, the genre of the article, the location the user read the article, the orientation of the device, as well as the url source of the article, a comparison with the users friends according to articles read this week and speed (see Fig. 2). The first prototype uses the head angle and eye positions to estimate if the user is looking on the screen. To determine both we use the in iOS built-in face detection and a Haar Cascade for the eyes (a standard computer vision technique see [16]). If the user seems to look at the screen for a longer period of time and scrolls to the end of the article not to fast (estimated with a threshold at around 700 words per minute), we count the article as read and record the approximate reading speed. The genre of the article is determined using Latent Semantic Mapping and some default categories, so far: culture, science, sports, economics, politics. These can be

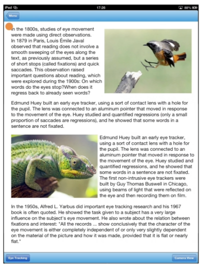

Figure 3: One of the 5 calibration points the user needs to focus on. 4 are in the corners of the text field, the fifth is in the middle of the screen.

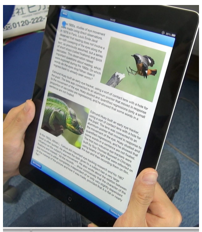

Figure 4: Successful eye gaze estimation after calibration on the iPad. The user reads the first word in the first line.

easily adjusted and the user can provide their own over "tagging". The comparison with friends is so far static, due to the small user count, yet we want to include a Facebook/Twitter connection.

Towards Robust Eyetracking
The current implementation of "reading detection" is very limited. As we want to improve it, we implemented several standard eye tracking and gaze estimation techniques from computer vision. We started with an appearance based approach presented by Holland et. al. [11] (code is available under GPL3) and reached similar accuracies as mentioned in their paper. Yet, this method is not fit for our problem, as it requires a long training phase and is highly dependent on the user. So far, a hybrid approach using a combination of existing techniques works best, described in the following. We implemented a shape-based approach using model fitting method based on elliptical shape models [10], detecting eye corners and iris shape by first using face and eye detection over a Haar Cascade [16]. This works faster (computationally less intensive), needs fewer training samples and seems more robust towards different users (Fig. 4 and 5). The method requires a five point calibration (see Fig. 3) and the user should not change the orientation of the device after calibration is done.

However, we still have some calibration problems and faulty detections, if the wrong parameters are used (see Fig. 6 and Fig. 7 for an example). All methods are very dependent on the calibration phase, lighting conditions and how the user holds the device. Interestingly, a user reading on the device shows a very characteristic pattern even when calibration has failed. Therefore we are currently investigating if we can detect reading despite not being able to get a stable gaze estimate.

We believe that some problems can be solved by using the device orientation obtained from motion sensors in the device and by combining several approaches presented by Hansen et.al. [9].

In addition to better working reading detection, robust eye gaze estimation and tracking can enable completely new applications. It's shown that eye gaze is related to knowledge and expertise, also the use of eye gestures is possible [13, 4].

Related Work
The closest to our work is an research by Holland et. al. using an appearance-based approach to estimate a user's gaze, training a neural network on different eye positions [11]. Also there are several papers that use commodity webcams for eye tracking [11]. Hansen et. al. give a great reference and detailed overview about different implementations of eye gaze estimation and tracking [9, 10].

Concerning reading, there are a few papers that try to detect skimming versus reading with dedicated eyetrackers [8, 2]. Eye tracking while reading can also be used to provide summaries of documents or to find which words a user finds relevant [15].

There is some work to foster better reading habits in children [7]. Yet, to our knowledge there is so far no effort towards quantified feedback for adults.

Conclusion and Future Work
We will explore further how to implement reading detection and a more stable eye tracking. A more important goal is to get first user feedback on the "reading life" application to see what type of inference

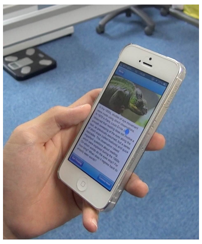

Figure 5: Successful eye gaze estimation after calibration on the iPhone. User reads the last word in the 3rd line.

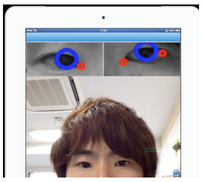

Figure 6: Relatively good iris estimation.

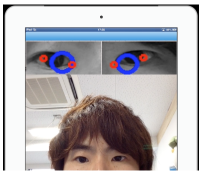

Figure 7: Bad parameters for iris estimation.

and statistic can spur better reading habits and a potential increase in reading.

We are working on making the gaze library open available, currently we focus on iOS, yet we also plan to support Android. Maybe it is possible to implement some of the advanced interaction techniques on commercial, mobile devices outlined by Biedert et. al. using dedicated eyetrackers [1].

Acknowledgements
This work was supported in part by the CREST project from Japan Society for the Promotion of Science (JSPS).

References

[1] Biedert, R., Buscher, G., and Dengel, A. The eyebook–using eye tracking to enhance the reading experience. *Informatik-Spektrum 33*, 3 (2010), 272–281.

[2] Biedert, R., Hees, J., Dengel, A., and Buscher, G. A robust realtime reading-skimming classifier. In *Proc. ETRA 2012* (2012), 123–130.

[3] Bravata, D. M., and Smith-Spangler. Using pedometers to increase physical activity and improve health. *The journal of the American Medical Association 298*, 19 (2007), 2296–2304.

[4] Bulling, A., and Roggen, D. Recognition of visual memory recall processes using eye movement analysis. In *Proc. UbiComp 2011* (2011), 455–464.

[5] Bulling, A., Ward, J. A., and Gellersen, H. Multimodal Recognition of Reading Activity in Transit Using Body-Worn Sensors. *ACM Trans. on Applied Perception 9*, 1 (2012), 2:1–2:21.

[6] Cunningham, A., and Stanovich, K. What reading does for the mind. *Journal of Direct Instruction 1*, 2 (2001), 137–149.

[7] Doiron, R. Using e-books and e-readers to promote reading in school libraries: Lessons from the field, 2011.

[8] Duggan, G. B., and Payne, S. J. Skim reading by satisficing: evidence from eye tracking. In *Proc. of CHI 2011* (2011), 1141–1150.

[9] Hansen, D. W., and Ji, Q. In the eye of the beholder: A survey of models for eyes and gaze. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 32*, 3 (2010), 478–500.

[10] Hansen, D. W., and Pece, A. E. Eye tracking in the wild. *Computer Vision and Image Understanding 98*, 1 (2005), 155–181.

[11] Holland, C., and Komogortsev, O. Eye tracking on unmodified common tablets: Challenges and solutions. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ACM (2012), 277–280.

[12] Kunze, K., Bulling, A., Utsumi, Y., Yuki, S., and Kise, K. I know what you are reading – recognition of document types using mobile eye tracking. In *in Proc. of ISWC 2013* (2013).

[13] Kunze, K., Kawaichi, H., Yoshimura, K., and Kise, K. Towards inferring language expertise using eye tracking. In *Ext. Abs. CHI 2013* (2013), 4015–4021.

[14] Kunze, K., Kawaichi, H., Yoshimura, K., and Kise, K. The wordmeter – estimating the number of words read using document image retrieval and mobile eye tracking. In *Proc. ICDAR 2013* (2013).

[15] Loboda, T. D., Brusilovsky, P., and Brunstein, J. Inferring word relevance from eye-movements of readers. In *Proc. of IUI '11* (2011), 175–184.

[16] Schmidt, A., and Kasiński, A. The performance of the haar cascade classifiers applied to the face and eyes detection. In *Computer Recognition Systems 2*. Springer, 2007, 816–823.