Smarter Eyewear– Using Commercial EOG Glasses for Activity Recognition

Shoya Ishimaru
Graduate School of Engineering
Osaka Prefecture University
Sakai, Osaka, Japan
ishimaru@m.cs.osakafu-u.ac.jp

Yuji Uema
Graduate School of Media Design
Keio University
Yokohama, 223-8526 Japan
uema@kmd.keio.ac.jp

Kai Kunze
Graduate School of Engineering
Osaka Prefecture University
Sakai, Osaka, Japan
firstname.lastname@gmail.com

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

Katsuma Tanaka
Graduate School of Engineering
Osaka Prefecture University
Sakai, Osaka, Japan
ishimaru@m.cs.osakafu-u.ac.jp

Masahiko Inami
Graduate School of Media Design
Keio University
Yokohama, 223-8526 Japan
inami@kmd.keio.ac.jp

UbiComp'14 Adjunct, September 13-17, 2014, Seattle, WA, USA
ACM 978-1-4503-3047-3/14/09.
http://dx.doi.org/10.1145/2638728.2638795

Abstract

Smart eyewear computing is a relatively new subcategory in ubiquitous computing research, which has enormous potential. In this paper we present a first evaluation of soon commercially available Electrooculography (EOG) glasses (J!NS MEME) for the use in activity recognition. We discuss the potential of EOG glasses and other smart eye-wear. Afterwards, we show a first signal level assessment of MEME, and present a classification task using the glasses. We are able to distinguish of 4 activities for 2 users (typing, reading, eating and talking) using the sensor data (EOG and acceleration) from the glasses with an accuracy of 70 % for 6 sec. windows and up to 100 % for a 1 minute majority decision. The classification is done user-independent.

The results encourage us to further explore the EOG glasses as platform for more complex, real-life activity recognition systems.

Author Keywords

Smart Glasses, Electrooculography, Activity Recognition, Eye Movement Analysis

ACM Classification Keywords

I.5.4 [PATTERN RECOGNITION Applications]: Signal processing

Figure 1: The EOG Glasses used for the experiments. The second picture shows the 3 electrodes touching each side of the nose and the area between the eyes. The last picture shows a user wearing JINS MEME.

Introduction

With wearable computing receiving increasing interest from industry, we believe that especially smart eyewear is a fascinating research area. In this paper, we show that the sensor data quality obtained by the EOG glasses seems good enough for activity recognition tasks.

The contributions are two fold. First, we want to motivate that smart eye wear is interesting for ubiquitous computing applications, as it enables to track activities that are hard to observe otherwise, especially in regard to cognitive tasks.

Second, we evaluate specific smart glasses, a prototype of J!NS MEME (available to consumers next year) for their use for activity recognition tasks. We show a signal level evaluation and a simple classification task of 4 activities (2 users 2 x 5 min. per activity). Both indicate that the device can be used for more complex scenarios.

In the end we discuss application scenarios and limitations for smart eyewear.

Toward Using Smarter Eyewear

Since the release of Google Glass, smart eyewear gains more and more traction for a wide range of applications (e.g. oculus rift for virtual reality). This new class of devices proves to be an interesting platform for ubiquitous computing, especially for activity recognition. As we humans perceive most of our environment with senses on our head (hearing, smell, taste and most dominantly our vision), the head is very valuable position for sensors.

Tracking eye movements can give us great insights about the context of the user, from recognizing what documents a user is reading, over recognizing memory recall to assessing expertise level [8, 6, 1, 4, 7].

Hardware

To evaluate the potential of smart eyewear for activity sensing, we are using an early prototype from J!NS MEME. The glasses are not a general computing platform. They are a sensing device. They can stream sensor data to a computer (e.g. smart phone, laptop, desktop) using Bluetooth LE. Sensor data includes vertical and horizontal EOG channels and accelerometer + gyroscope data. The runtime of the device is 8 hours enabling long term recording and, more important, long term real-time streaming of eye and head movement. They are unobtrusive and look mostly like normal eyewear (see Figure 1).

Before recording with the device the first time, the electrodes should be adjusted a bit to the user's nose/eyes to get an optimal EOG signal. This is a one-time adjustment due to the early prototype stage.

Initial Signal Level EOG Evaluation

A manual inspection of the data recorded by the prototype reveals that detecting blinks and reading activity seems feasible. A signal example from a user blinking 7 times is given in Figure 2. We depict the raw vertical EOG component before any filtering, even then the blinks are easy recognizable. Another signal example from the horizontal EOG component is shown in Figure 3. In this case, the user was reading. Again we depict the "raw" horizontal component of the EOG signal.

Simple Blink detection –We apply a very simple peak detection algorithm on data from 2 users. We consider a point being the a peak if it has the maximal value, and was preceded (to the left) by a value lower by a constant. 2 users with wearing the smart glasses sitting in front of the stationary eye tracker and blink 30 times naturally.

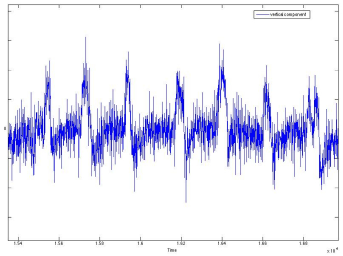

Figure 2: The vertical EOG component (raw signal), while the user blinks seven times.

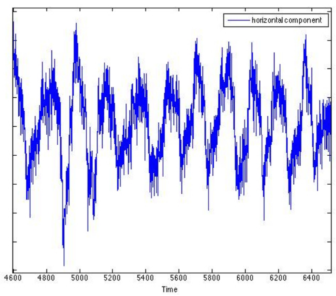

Figure 3: The horizontal EOG component (raw signal), while the user reads.

We can detect 58 of the 60 blinks with this very simple algorithm applied to the "raw" vertical EOG component signal.

Classification Task

For a first impression, we evaluate if the sensor data from the glasses can distinguish more complex activity recognition tasks. We assume that modes of locomotion etc. can easily be recognized by the motion sensors alone. Therefore we concentrate on tasks performed while sitting in a common office scenario. We include 4 activities: typing a text in a word processor, eating a noodle dish, reading a book and talking to another person.

Method

We use a simple classification method: windowed feature extraction with a K-Nearest Neighbor classifier (k = 5) and majority decision. 7 features are calculated over a 6 sec sliding window (2 overlapping): the median and variance of the vertical and horizontal EOG signal and the variance for each of the 3 accelerometer axes. The features are used to train the nearest neighbor classifier. On top of the classification we apply a 1 minute majority decision for smoothing.

Experimental Setup

For the experimental setup, we record data using the J!NS MEME prototype connected over Bluetooth to a Windows laptop for 2 participants 4 activities, each activity for 2 x 5 min. We asked them to perform the activities naturally while sitting at a desk.

Before starting to record with a participant, we need to adjust the electrodes on the current prototype towards the facial features of the user to be sure to capture a clean EOG signal. This initial setup step needs to be done only once per user.

Initial Results and Discussion

We apply the windowed feature extraction and classification method on the data, performing a user independent classification, training with the data of one user and evaluating with the other user.

For the frame-by-frame classification we reach a correct classification rate of 71 % on average for a 6 sec. window (2 sec. overlap). The confusion matrix is given in Figure 4. Applying the majority decision window of 1 min. we reach 100 % discrimination between classes.

Strengthened by the good performance distinguishing the 4 activities for 2 users in a user independent way, we will evaluate the platform to see if the detection of specific activities is possible in real life situations during long term deployment. Being able to detect food intake behavior or learning tasks (e.g. reading) are of particular interest to us.

Related Work

We follow the early pioneering work from Bulling et al. and Manabe et al. in using EOG for activity recognition[9, 3]. Bulling et al. described an approach to recognize different visual activities using EOG prototypes, including reading, solely from gaze behavior using machine learning techniques in stationary and mobile settings [3, 2].

There is some work to use Google Glass as activity recognition platform. This work is complementary to the approach in this paper, as Google Glass is a very different device (a full fledged wearable computer) with different sensing modalities [5]. Most of the related work uses dedicated research prototypes, often attaching electrodes directly to the skin above or below the eye.

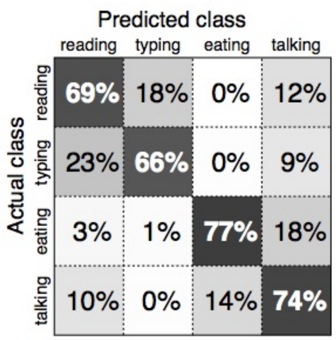

Figure 4: The confusion matrix in percent for the frame-by-frame classification using a 6 sec. sliding window (accuracy 70 %).

Conclusion

We presented an initial evaluation of a smart glasses prototype for activity recognition. Both signal level analysis and 4 activity classification task show favorable results. 58 of 60 blinks for 2 users can be detected by straight forward peak detection. The 4 activities, typing, eating, reading and talking can be distinguished perfectly over a 1 minute window.

Smart glasses like J!NS MEME are very unobtrusive and can be easily confused with "normal" glasses. Yet, the question is if this type of devices can produce a high enough signal quality to be used for complex activity recognition systems. Of course, the verdict is still out, yet, our initial results are very positive, indicating the potential of smart glasses for ubiquitous computing applications.

Acknowledgements

We would like to thank the research department of J!NS for supplying us with prototypes. This is work is partly supported by the CREST project.

References

[1] Bulling, A., and Roggen, D. Recognition of visual memory recall processes using eye movement analysis. In *Proceedings of the 13th international conference on Ubiquitous computing*, ACM (2011), 455–464.

[2] 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.

[3] Bulling, A., Ward, J. A., Gellersen, H., and Tröster, G. Eye Movement Analysis for Activity Recognition Using Electrooculography. *IEEE Trans. on Pattern Analysis and Machine Intelligence 33*, 4 (Apr. 2011), 741–753.

[4] Eivazi, S., Bednarik, R., Tukiainen, M., von und zu Fraunberg, M., Leinonen, V., and Jääskeläinen, J. E. Gaze behaviour of expert and novice microneurosurgeons differs during observations of tumor removal recordings. In *ETRA*, ACM (New York, NY, USA, 2012), 377–380.

[5] Ishimaru, S., Kunze, K., Kise, K., Weppner, J., Dengel, A., Lukowicz, P., and Bulling, A. In the blink of an eye: combining head motion and eye blink frequency for activity recognition with google glass. In *Proceedings of the 5th Augmented Human International Conference*, ACM (2014), 15.

[6] Kunze, K., Iwamura, M., Kise, K., Uchida, S., and Omachi, S. Activity recognition for the mind: Toward a cognitive quantified self. *Computer 46*, 10 (2013), 0105–108.

[7] 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).

[8] Kunze, K., Utsumi, Y., Shiga, Y., Kise, K., and Bulling, A. I know what you are reading: recognition of document types using mobile eye tracking. In *Proceedings of the 17th annual international symposium on International symposium on wearable computers*, ACM (2013), 113–116.

[9] Manabe, H., and Fukumoto, M. Full-time wearable headphone-type gaze detector. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, ACM (2006), 1073–1078.