Smart Eyewear for Interaction and 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 Media Design
Keio University
Yokohama, 223-8526 Japan
kai@kmd.keio.ac.jp

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
katsuma@m.cs.osakafu-u.ac.jp

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

CHI'15 Extended Abstracts, April 18–23, 2015, Seoul, Republic of Korea.
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2725449

Abstract
Smart glasses and, in general, eyewear are a fairly novel device class with a lot of possibilities for user interaction design and unobtrusive activity tracking. In this paper we show applications using an early prototype of J!NS MEME, smart glasses with integrated electrodes to detect eye movements (Electrooculography, EOG) and motion sensors (accelerometer and gyroscope) to monitor head motions. We present several demonstrations: We show a simple eye movement visualization, detecting left/right eye motion and blink. Additionally, users can play a game, 'Blinky Bird'. They need to help a bird avoid obstacles using eye movements. We implemented online detection of reading and talking behavior using a combination of blink, eye movement and head motion. We can give people a long term view of their reading, talking, and also walking activity over the day.

Author Keywords
Eye Movement Analysis; Electrooculography; Eye wear; Smart Glasses; Eye Wear Computing; Activity Recognition

ACM Classification Keywords
I.5.4 [PATTERN RECOGNITION Applications]: Signal processing;

Figure 1: Users wearing the MEME glasses.

Figure 2: The glasses prototype: There are the 3 electrodes around the nose used for eye movement analysis (EOG).

Motivation
Research and industry get more and more interested in smart eyewear, from Google Glass, over Epson Moverio, the Oculus Rift to the Sony Smart Glasses. Yet, most of these designs still look a bit clunky and emphasize displays (augmented or virtual reality), not the sensing aspect. We believe in the potential for unobtrusive smart glasses that focus on sensing physiological signals. Especially eye movement seems a compelling sensing modality. Although there are a lot of interesting cues and activities we can recognize using gaze, there are only few researchers focusing on tracking eyes in realistic environments with minimal augmentation [1]. So far, eye movement analysis either required expensive mobile or stationary eye-trackers, googles or ear pieces with electrodes touching the skin around the eyes [7, 1, 5]. Although we see more and more cheaper optical eye tracking solutions, they still look obtrusive compared with standard glasses and have battery limitations (cannot be used for more than 1-3 hours).

We present several demonstrations using a first prototype of unobtrusive, affordable EOG glasses called J!NS MEME. We show the feasibility to detect simple eye movements, apply them to a simple game and use the glasses to recognize reading/ talking, as well as long term behavior detection to give people an quantified overview about their cognitive, social and physical life.

MEME Prototype Hardware
We implement our demonstrations on a J!NS MEME prototype shown in Figures 1 and 2. The prototype is unobtrusive and looks close to normal glasses. It is equipped with 5 electrodes (3 around the nose, 2 on the ears) to detect eye movements and motion sensors (accelerometer and gyroscope), as well as a Bluetooth LE module to stream the data to a PC, laptop, smartphone or tablet. The electrodes are sampled with over 100 Hz, the motion sensor with over 50 Hz. Using these settings we already achieve a battery runtime of over 8 hours for the current version (will be improved). However, it is to note that we are using an early prototype of MEME, not the final product. 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 (especially for the vertical signal component). This is a one-time adjustment due to the early prototype stage.

Eye Movements for User Interactions
Using the current prototype we can detect and display eye movements. We focus on left and right movements and eye blinks, as the up and down directions are a bit user-dependent (initial user adjustment of the glasses required). Left, right movements and blinks can be recognized and chained as input gestures for actions (e.g. left-right-blink for accepting a notification/call versus right-left-blink for dismissing it).Users can try to make their own action sequence.

Based on Blink detection, we created a simple game seen in Figure 3. "Blinky Bird" based on an open-source version of "Flappy Bird" allows the user to control a small bird using eye blinks. Every time he closes his eyes, the bird flaps its wings. The user needs to avoid obstacles by blinking at the right time.

Reading/Talking Detection
We also show reading and talking recognition using the device. Based on previous work, it is possible to classify different activities offline using just blink detection and head motion [2]. In previous work we used Google Glass for the data recordings. The infrared distance sensor on

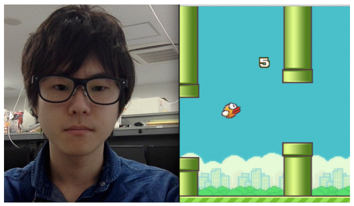

Figure 3: User playing "Blinky Bird" using eye blinks to maneuver the small bird sprite.

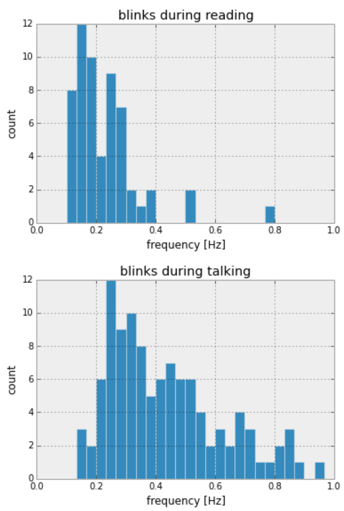

Figure 4: Histograms for the number of blinks during reading and talking.

Google Glass can be used for blink detection and it comes also equipped with motion sensors. We extend this work, and implement an online reading/ talking detection using MEME.

Method – Users can wear the glasses and can get instant feedback (1-2 seconds delay) from a smartphone or laptop computer what activity they are currently involved in. We recognize talking, reading and "None" (for the Null class) using head motion (variance of the 3 axis accelerometer norm), blink frequency, as well as variance of the horizontal eye movement component as features for a j48 decision tree algorithm. The instant feedback might not be so interesting for applications (except for realtime activity sharing or similar), yet gives us the ability to record reading/communication habits for the day and makes it easy to evaluate the performance of MEME.

Evaluation – We tested a prototype of our reading/ talking detection system on 12 people. They performed 15 instances of each reading and talking as well as a 5-6 instances of other unrelated activities (drinking water, eating etc.). Only 16 instances were wrongly classified leaving us with 91 % of accuracy. No instances of other activities were classified wrongly, indicating that our online system works as well as the offline classifier implemented on Google Glass data [2].

Long Term Behavior Recording

As mentioned before, reading and talking detection can provide users with a quantified review of their day. We think this application is mostly interesting for education scenarios. Users will be encouraged to read more when they get feedback on their reading volume. The more we read the higher are our vocabulary and critical thinking skills [3, 9]. Getting an overview about talking activities might be also useful for users leaning towards depression just to name one example [8]. Increasing direct communication can also fend off anxiety in some cases[6].

Figure 6 depicts a user reading a book, while MEME tracks his activity and shows how many words the user read. The current word count implementation is based on time alone, yet we can improve the accuracy taking some eye movement features into account. Figure 7 shows two people engaged in a conversation. One of them wears a MEME tracking the time he talks.

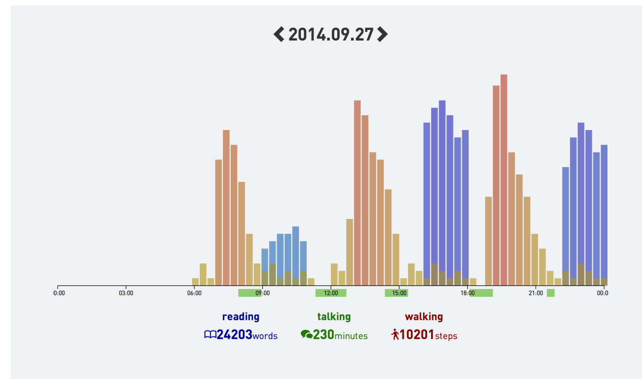

Figure 5: Overview of the daily activities: reading volume depicted as blue bars, talking as green lines and step count in red.

We plan on giving people an quantified overview about their cognitive, social and physical life. Currently we implemented long term reading detection (word count based on time), talking detection and a simple step counter using the accelerometer. People can get a summary of their daily activity using a web front-end seen in Figure 5). In future, we want to extend this work along the cognitive and social axes.

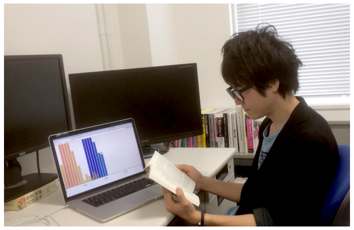

Figure 6: User reading, while MEME is tracking the number of words read using time.

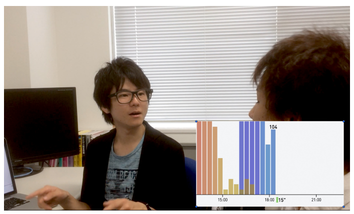

Figure 7: User taking, while MEME is tracking the time talked.

Conclusion and Future Work

We presented interaction and recognition demonstrations using an unobtrusive EOG glasses prototype. The applications are meant to show the potential of the device category. In the following we will give a small outlook on how we plan to extend our work.

Word Counting– Our word counting algorithm for real-time feedback is so far very simple. We just take into account the time a user reads and estimate the word count accordingly to the duration, not regarding if he reads fast or slow. To improve this we can use features derived from the horizontal component of the EOG signal.

Eye Gestures– We show the feasibility of detecting eye gestures with our prototype. Yet, we have not designed application cases and suitable eye gestures people might want to use, this is left for future work.

Cognitive Activity Recognition – We want to present users with a comprehensive overview of their cognitive and social activities. With quantifying them, we hope to learn more about healthy habits to improve our mental life and to induce positive long term behavior change [4].

Acknowledgements

Thanks a lot to J!NS for supplying us with prototypes. This is work is partly supported by the CREST Project "Development of Fundamental Technologies for Innovative Use of Character/Document Media and Their Application to Creating Human Harmonized Information Environment".

References

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

[2] 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 *Augmented Human*, ACM (2014), 15.

[3] Just, M. A., and Carpenter, P. A. *The psychology of reading and language comprehension.* Allyn & Bacon, 1987.

[4] 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.

[5] Kunze, K., Katsutoshi, M., Uema, Y., and Inami, M. How much do you read? – counting the number of words a user reads using electrooculography. In *Accepted at Augmented Human'15*, ACM (2015), 1073–1078.

[6] Landman-Peeters, K., Hartman, C. A., van der Pompe, G., den Boer, J. A., Minderaa, R. B., and Ormel, J. Gender differences in the relation between social support, problems in depression and anxiety. *Social Science & Medicine 60*, 11 (2005), 2549–2559.

[7] Manabe, H., and Fukumoto, M. Full-time wearable headphone-type gaze detector. In *CHI'06 Extended Abstracts*, ACM (2006), 1073–1078.

[8] Ross, C. E., and Mirowsky, J. Explaining the social patterns of depression: Control and problem solving– or support and talking? *Journal of health and social behavior* (1989), 206–219.

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