How Much Do You Read? – Counting the Number of Words a User Reads Using Electrooculography

Kai Kunze, Masai Katsutoshi, Yuji Uema, Masahiko Inami
Graduate School of Media Design
Keio University, Yokohama, Japan
Email:{kai/masai/uema/inami}@kmd.keio.ac.jp

ABSTRACT
We read to acquire knowledge. Reading is a common activity performed in transit and while sitting, for example during commuting to work or at home on the couch. Although reading is associated with high vocabulary skills and even with increased critical thinking, we still know very little about effective reading habits. In this paper, we argue that the first step to understanding reading habits in real life we need to quantify them with affordable and unobtrusive technology. Towards this goal, we present a system to track how many words a user reads using electrooculography sensors. Compared to previous work, we use active electrodes with a novel on-body placement optimized for both integration into glasses (or head-worn eyewear etc) and for reading detection. Using this system, we present an algorithm capable of estimating the words read by a user, evaluate it in an user independent approach over experiments with 6 users over 4 different devices (8" and 9" tablet, paper, laptop screen). We achieve an error rate as low as 7% (based on eye motions alone) for the word count estimation (std = 0.5%).

Author Keywords
Eye Movement Analysis, EOG, Reading, Wordcount

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

INTRODUCTION
Increased reading volume is associated with numerous cognitive benefits, including improved vocabulary skills, higher general knowledge and increased critical thinking [4]. Furthermore, reading is entertaining and has social value, higher reading volumes in adolescents are correlated with higher self-esteem and improved cognitive and emotional well-being[13, 14]. Although there are these strong positive effects, only few previous works evaluated reading activities in situ and even fewer tried to quantify them[3, 10]. Despite the growing awareness of the importance of reading for learning, getting

Figure 1: Mockup of a service to quantify reading habits during everyday life.

people to read more and to adopt healthy reading habits is challenging especially as the amount easy digestible content in form of videos etc. increases. Automatically tracking physical activities can motivate users to more healthy lifestyles [2]. We believe this translates also to cognitive skills and tasks. We want to investigate wether we can track reading habits similar to physical activity to give users tools to improve their mental fitness. Letters and words represent ideas and concepts; tracking the volume, speed and time a user is reading them seems particularly valuable as it gives first insights into learning and provides us with a basic countable measure of our performance[7]. For example, children suffering from reading disabilities can be earlier diagnosed, people can without trouble improve their reading speed and older adults have an easier way to fight dementia. Since research suggests that performance related to these situations is closely linked to reading volume[15, 8].

We still have a hard time defining what healthy reading habits for adults are[4], as tools are missing to quantify reading in everyday situations and in long term studies. This paper provides the first steps towards assessing reading volume in realistic settings utilizing electrooculography. The particular contributions of this work are (1) we present a method to quantify how much words a user reads using electrooculography working over different devices, diverse users and varying text length (2) we achieve the lowest error rate of 7% user-independent for our word count estimation over two data sets of in total 6 users over 5 documents each of varying sizes (from 115-881 words). (3) So far nobody implemented a word counting algorithm using EOG. Also our EOG setup differs significantly from related work focusing only on read-

AH '15, March 09 - 11, 2015, Singapore, Singapore

ACM 978-1-4503-3349-8/15/03…$15.00
http://dx.doi.org/10.1145/2735711.2735832

ing detection, it uses only placements on the nose to infer left/right and up/down movements. This setup could be easily integrated in smart glasses.

We see our work as a first step towards services that quantify reading habits during everyday life (see Figure1).

APPROACH

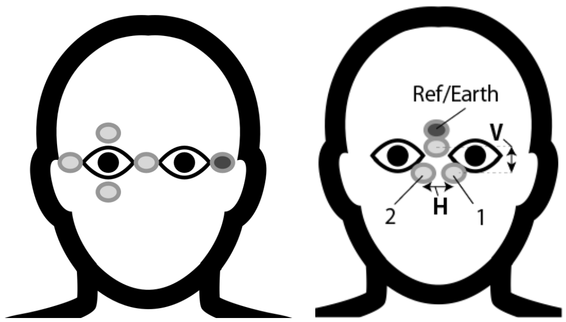


Figure 2: Electrode Setup: setup similar to some related work on the left, new setup on the right.

As seen in Figure 2, our electrode setup differs substantially from systems in related work. Most EOG researcher use 5 electrodes: top and bottom, left and right for measurements and another electrode for reference. We apply just 4 active electrodes, two left and right from the nose, one between the eyes and the last one as reference also between the eyes. The setup is inspired by J!NS MEME[1], a consumer EOG smart glasses device. However, we could not find any specification or publication related to the EOG setup of MEME that give more detailed hardware information or could be cited as related work. Yet, from the information we gathered from J!NS the biggest differences of our system to MEME are, 1) we use 4 active electrodes, meme employs 3 normal electrodes. Our setup should be more resilient to noise and thus interesting for future smart glasses. From the electrodes marked 1-2 we get the horizontal component of the eye movement, From the electrodes 1/2 to the one between the eyes the vertical component.

We assume the signal is pre-segmented in reading and not reading segments using available reading detection algorithms [3]. We assume reading detection is more or less a solved problem. After reading detection we apply first a line break detection and then 3 different types of words read estimations. We will go in detail in the following.

Line Break Detection

We use a simple valley detection algorithm to detect line breaks on the horizontal signal component of the EOG. We define a point A as a minimum if it has the minimum value in it's vicinity, and was preceded (to the left) by a value lower by Delta (Delta is defined experimentally by 60 ms). We also apply a minimum size threshold for the valley detection (the 5% percentile also defined experimentally). In Figure3 we see the peak detection algorithm applied to the horizontal component

[1]https://www.jins-jp.com/jinsmeme/en/

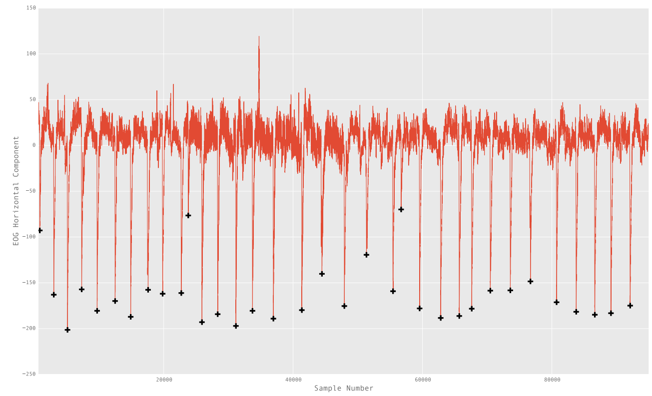


Figure 3: EOG signal, horizontal component for a user reading a document. The line breaks (even the three shorter lines) are easy to spot in the data.

Figure 4: The raw data for a segment of the horizontal EOG component between two line-breaks. The median filter is depicted in blue and black crosses are the peaks detected as "forward" saccades.

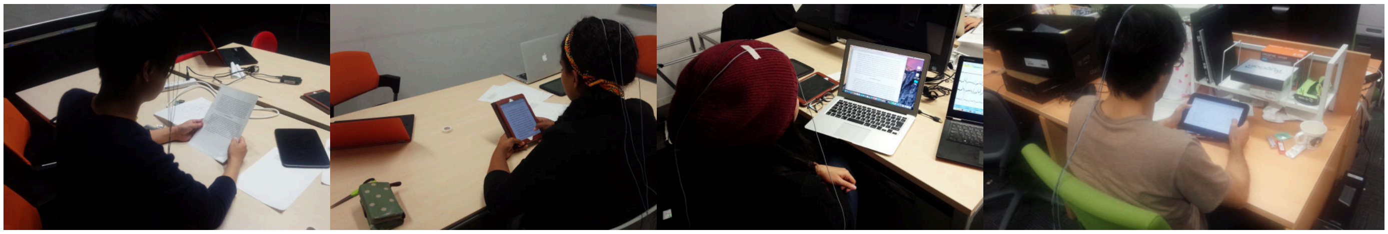
of the EOG signal from a user reading a document. The line breaks (valleys) are easy visible. For reference we are using a AC system not DC, as most related work (there is no much practical difference except some properties of the sample signals might look strange to researchers only used to DC systems).

Words Read Estimation

After line break detection, we apply words read estimation.

Static Word Count – The easiest algorithm to determine how many words a person read is to multiply the line breaks with a static word count per line (we use 9.5 words determined experimentally).

Line-Break SVR Word Count – A more advanced algorithm uses only line-break saccade features. From a selection of 25 features, we determined the average length of line-break saccades, minimum length of saccades, time between line-break saccades as the best estimators for word count on this level. We use this features to train an Support Vector Regression (SVR) algorithm with radial base kernel function to predict the words read.



(a) (b) (c) (d)

Figure 5: Pictures from the data recordings. Participants are equipped with the Polymate EOG and read on the different devices: paper, 8" tablet, laptop and 9" tablet.

Figure 6: Samples from the electrode setup used.

Line-Features SVR Count – The most advanced algorithm takes also the features from the sections between line-breaks. We apply a Median filter with length 50 ms to these sections and apply a peak detection algorithm (analogous to the valley detection for the line-breaks) to get the forward saccades (forward in terms of reading direction) see Fig.4. Then we calculate the following forward-saccade features (also selected out of around 25 features): number of forward saccades, length of saccades, time between saccades. As before we use line-break saccade features and additionally the forward saccade features to train the SVR.

EXPERIMENTS

We used the Polymate mini to track eye movements. It's a EOG device with active electrodes. This device detects users eye movement by measuring EOG via attached active electrodes. The Figures2 and 6 show how to attach the electrodes.For detecting a horizontal movement, we used the difference between EOD data from electrode 1 and 2. For detecting a vertical movement, we used the difference between EOD data from electrode 1 and one between the eyes. The Polymate is connected via Bluetooth technology to laptop where experimenter can see the EOG data. The sampling rate recorded by the device was 1K Hz.

We have 6 volunteers with multinational background(2 Canadian, 1 Syrian, 1 Indonesian, 1 French and 1 Japanese), all students, 3 female, average age 24.3. Each subject read 5 documents with different lengths: 115, 253, 519, 679 and 881 words in ascending order. We prepared 4 different media types which had different width and height sizes, but their font size - 12pt - was constant. The media types are A4 paper,tablet,iPad mini and laptop. The media types are our independent variables, we cycle through them using a lattice square design. Before starting the experiments, the subjects attached the 4 electrodes on their faces. After that, the experimenter explained to the subject what to do and instructed them. Once recording started, the subject began to read and the experimenter observed the procedure on laptop. Each document was recorded separately. The document order was not changed depending on each subject. Photos from the recording are shown in Figure5. All recordings were performed in an indoor university environment.

RESULTS AND DISCUSSION

Applying the methods in the approach on the collected data from the 6 participants, we reach the following results. We compare our inference to a word count estimate derived from a perfect reading detection system for baseline (most research in the related work focuses on it). We estimate the number of words a person was reading just based on time and compare this to our system.

Method	Error Rate	STD
Time Baseline	31%	9%
Static Word Count	16%	3%
Line-Break SVR Word Count	10%	1%
Line-Features SVR Word Count	7%	0.5%

Table 1: Overview of the word count estimation error and standard deviation of the error for different methods. First the baseline just using the time a user read a text, second is a static word number times the detected line-breaks, third a SVR based on Line Break features alone, and last a SVR based on Line Break and Line features.

For the Line break detection we have an error of 5%, std 1.2%. The summary of the results can be found in Table1. The static word count method already performs with around half the error of the time baseline (16%). Interestingly, all methods are significantly better than baseline, p-values comparing them to baseline: static 0.04 ($F = 1.2$), line-break 0.02 ($F= 0.79$), line features 0.009 ($F=0.67$).

RELATED WORK

As we are interested in tracking reading habits, a collection of cognitive tasks, we first might try direct brain sensing. Yet, as realted work shows most methods are too bulky or are too noisy to get decent results related to recognizing reading [11, 5].

The most interesting modalities for direct brain sensing seem to be electroencephalography (EEG) and near-infrared spectroscopy (NIRS), as both can be used in mobile settings[17]. However, for both spacial and for NIRS temporal resolution are not so good. Their signal is also strongly affected by motion noise and usually requires complex filtering/preprocessing steps.

However, the strong relationship between reading and eye movements is very well explored in cognitive science and psychology [16, 9]. Most of the reading research in psychology however emphasizes on older adults or disabled [5].There are only a few research publications centering around reading detection in mobile and stationary settings [3]. As such detecting reading can be used as a very simple word counting mechanism, as there's a relation between time read and the read volume. Biedert et al. look into how people read text. They presented a method to discriminate skimming from reading using a novel set of eye movement features [1].

Manabe et al. also explore different EOG placement for electrodes using a headphone type formfactor [12].

Concerning reading habits, there are some questionnaire based evaluations giving advice about effective reading techniques to second language learners, as well as for children with reading disabilities and older adults struggling with dementia[6, 15]. Hansen [8] reports on a series of studies on reading comprehension with rapid readers trained in the Evelyn Wood method.

The closest to our work is the Wordometer implemented by Kunze et al. [10]. They introduce a word counting algorithm based on mobile eye tracking. They use an optical system, a mobile eye tracker and they utilize the scene camera of the tracker justify the eye gaze. Our system is easier to deploy (e.g. potentially intergratible into glasses) and has less battery constraints. As far as we know, this is the only research work exploring technology support to quantify reading and presenting a word count estimation algorithm capable of dealing with varying device types, line lengths and different reading speeds.

CONCLUSION

We presented our work towards tracking how much a user reads, enabling quantified feedback about reading volume. We show an word count estimation algorithm that works with 7%. We believe this is a good start for real world reading tracking systems, of course we need a more representative data set and in a next step we want to get rid of the taped electrode placement, integrating the electrodes in a smart glasses frame, to enable a quantified reading service as shown in Figure 1.

REFERENCES

1. Biedert, R., Hees, J., Dengel, A., and Buscher, G. A robust realtime reading-skimming classifier. In *Proc. ETRA 2012* (2012), 123–130.
2. 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.
3. 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.
4. Cunningham, A., and Stanovich, K. What reading does for the mind. *Journal of Direct Instruction 1*, 2 (2001), 137–149.
5. Ferstl, E. C., Neumann, J., Bogler, C., and von Cramon, D. Y. The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping* (2008), 581–593.
6. Genesee, F., Lindholm-Leary, K., Saunders, W., and Christian, D. English language learners in us schools: An overview of research findings. *Journal of Education for Students Placed at Risk 10*, 4 (2005), 363–385.
7. Halpern, D. F. Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New directions for teaching and learning 1999*, 80 (1999), 69–74.
8. Hansen, D. M. A discourse structure analysis of the comprehension of rapid readers.
9. Irwin, D. E. Fixation location and fixation duration as indices of cognitive processing. *The interface of language, vision, and action: Eye movements and the visual world* (2004), 105–134.
10. 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).
11. Kunze, K., Yuki, S., Ishimaru, S., and Kise, K. Reading activity recognition using an off-the-shelf eeg. In *Proc. ICDAR 2013* (2013).
12. 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.
13. Moore, D. W., Alvermann, D. E., and Hinchman, K. A. *Struggling Adolescent Readers: A Collection of Teaching Strategies.* ERIC, 2000.
14. Moore, R. L., and Moschis, G. P. Impact of newspaper reading on adolescent consumers. *Newspaper Research Journal 2*, 4 (1981), 3–8.
15. Rashotte, C. A., and Torgesen, J. K. Repeated reading and reading fluency in learning disabled children. *Reading Research Quarterly* (1985), 180–188.
16. Spache, G. D. Is this a breakthrough in reading? *The Reading Teacher 15*, 4 (1962), 258–263.
17. Tsunashima, H., Yanagisawa, K., and Iwadate, M. Measurement of brain function using near-infrared spectroscopy (nirs), 2012.