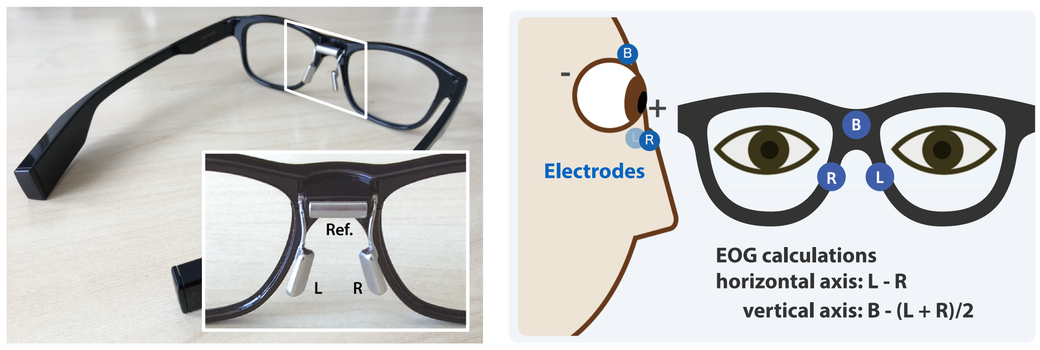
Electrooculography Dataset for Reading Detection in the Wild

Shoya Ishimaru[1)], Takanori Maruichi[2)], Manuel Landsmann[2)], Koichi Kise[2)] and Andreas Dengel[1)]
1) German Research Center for Artificial Intelligence (DFKI) 2) Osaka Prefecture University

Paper + Dataset
https://github.com/shoya140/ubicomp2019-eog-dataset/

Examples of reading activities in the wild (English, Japanese horizontal and vertical). Pictures are taken by Narrative Clip

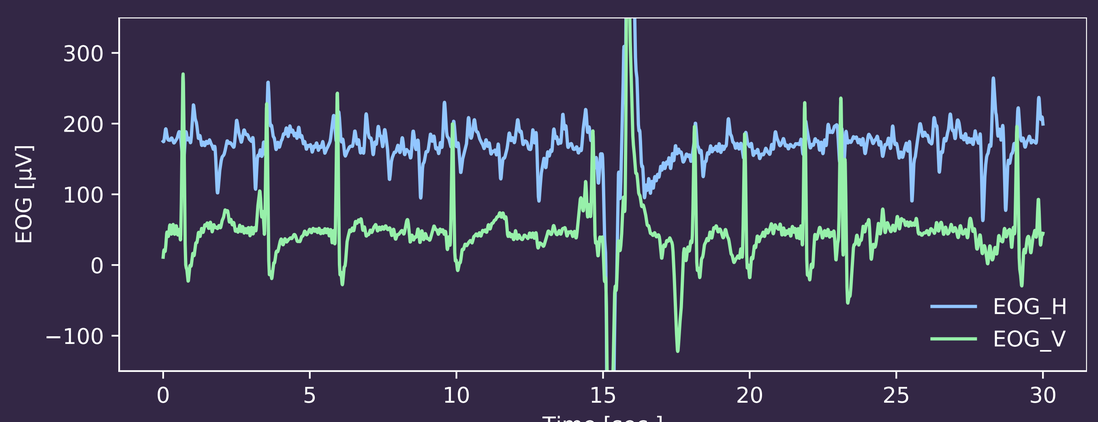


JINS MEME Electrooculography Glasses and the technological background

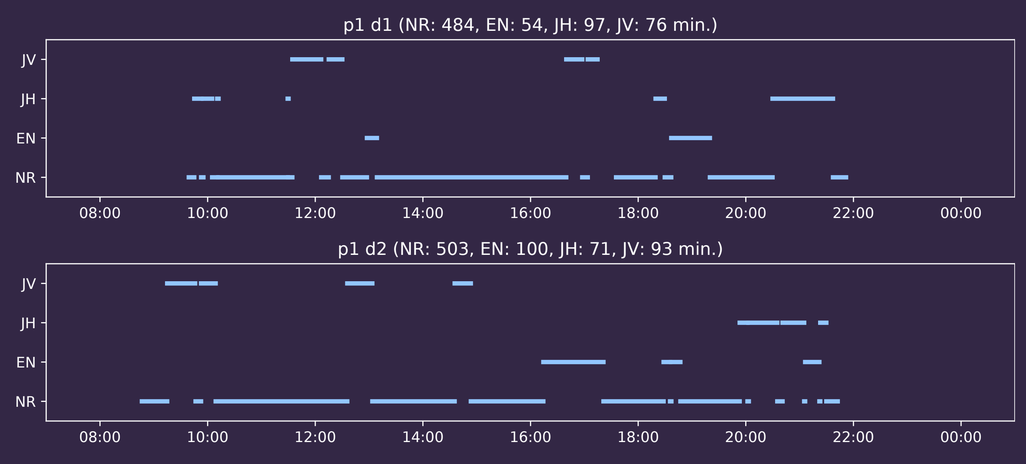
Motivation

Just as our bodies consist of what we eat, our minds are shaped by the information we obtain. In particular, written text is one of the most important information sources in our lives. **Understanding and improving daily reading habits provides several cognitive benefits, including increased vocabulary and logical thinking.**

Detecting reading activities in a real life is a challenging task compared to other activity detections because 1) reading does not require dynamic body movements and 2) readings in real life occur in a variety of situations with several document layouts. **For contributing to the implementation of robust reading detection algorithms, we introduce a dataset** which contains 220 hours of sensor signals from JINS MEME electrooculography glasses and corresponding ground truth activity labels.



EOG signals while reading an English document



Distributions of annotated activities of one participant

Dataset Overview

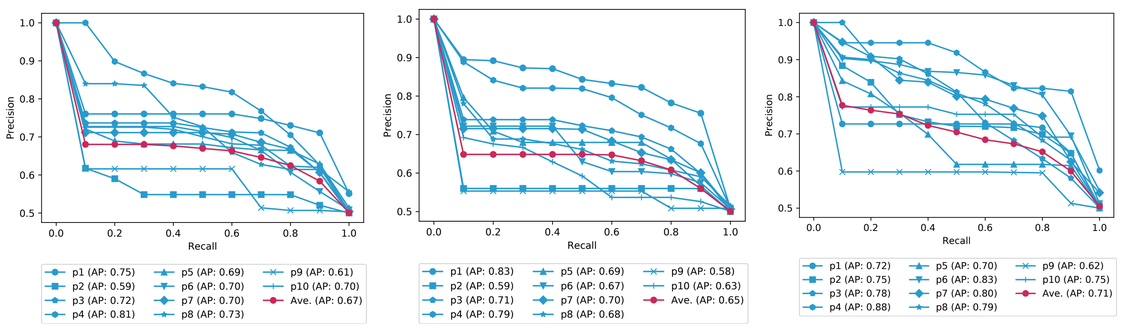
We utilized JINS MEME ES_R and Android Nexus 5X for the data recording. JINS MEME is equipped with a three-electrode electrooculography (EOG) sensor which measures eye movements and a six-axis internal measurement unit (IMU) which measures head and body movements. A form factor and a long-life battery of the device are designed for in-the-wild studies.

To support a ground truth labeling task, we provided Narrative Clip, a small life-logging camera which can be attached to clothes and takes a picture every 30 seconds. **At the end of each day, every activity in the pictures was annotated by the user into four categories: reading in English (EN), reading in Japanese written horizontally (JH), reading in Japanese written vertically (JV), and not reading (NR).** In order to protect privacies, we collected only activity labels and no pictures.

We recruited **ten Japanese college students for two days** of data recording. In total, our dataset contains 23 hours of English reading, 25 hours of Japanese horizontal reading, 25 hours of Japanese vertical reading, and 146 hours of other activities.

Condition	User-independent	User-dependent
EN vs JH vs JV vs NR	32	34
(EN + JH) vs JV vs NR	46	45
(EN + JH + JV) vs NR	68	69
EN vs NR	67	66
JR vs NR	68	68
JV vs NR	74	69
EN vs JH vs JV	36	46

Classification accuracies [%]



English (EN) Japanese horizontal (JH) Japanese vertical (JV)

Recall-precision curves and average precisions of reading detection

Baseline Study

Approach

We propose a detection method as a baseline. **Ten features are calculated from one sample (30 seconds window of a data stream): means and variances of the two EOG axes, variances of three accelerometer axes, and variances of three gyroscope axes.** We utilize SVM (kernel: RBF, C: 1, gamma: 0.125) to classify the samples.

Results

Head and eye movements while reading JV texts are relatively distinctive compared to EN and JH. This may be because the Japanese vertical writing style is often used for well-formatted texts such as novels and newspapers. Reading EN and JH texts can be often performed with other activities, for example, writing and browsing.

The 11-point interpolated recall-precision graphs for each activity detection task represent that the performances are highly distributed by participants although there is not much difference between classification accuracies of a user-independent and a user-dependent training.

大阪府立大学 OSAKA PREFECTURE UNIVERSITY

https://www.jins-jp.com/jinsmeme/en/
JINS MEME