# The Wordometer – Estimating the Number of Words Read Using Document Image Retrieval and Mobile Eye Tracking

Kai Kunze, Hitoshi Kawaichi, Kazuyo Yoshimura, Koichi Kise
Osaka Prefecture University, Japan
kunze,kawaichi,yoshimura@m.cs.osakafu-u.ac.jp
kise@cs.osakafu-u.ac.jp

***Abstract*—This paper introduces the Wordometer, a novel method to estimate the words a user reads using the eye gaze recorded by a mobile eye tracker and document image retrieval. We present a reading detection algorithm which works with over 91 % accuracy over 10 test subjects using 10-fold cross validation. We implement two algorithms to estimate the read words using a line break detector. A simple version gives an average error rate of 13,5 % for 9 users over 10 documents. A more sophisticated word count algorithm based on support vector regression with an RBF kernel reaches an average error rate from only 8.2 % (6.5 % if one test subject with abnormal behavior is excluded). The achieved error rates are comparable to pedometers that count our steps in our daily life. This means, the Wordometer can be used as a step counter for the information we read to make our knowledge life healthier.**

## I. Introduction

To make their life healthier, more and more people are using pedometers, be it a dedicated device or a smart phone application [1]. A survey of medical pedometer studies shows that monitoring something as simple as how many steps you take can have a huge impact in health. People wearing pedometers walk in average over 1.5 km more per day, substantially decreasing their risk for heart attacks, type2 diabetes and other diseases related to obesity [1].

A pedometer is a very simple, yet effective tool to monitor and improve our physical fitness. We want to provide a similar tool for our cognitive fitness: the Wordometer, counting how many words a user reads per document per day. This paper presents our initial work towards this goal.

Of course, how much we read directly influences the size of our vocabulary and our language skills [2]. However, several studies also indicate that the more people read throughout the day the higher are their general knowledge and critical thinking skills [2], [3]. Interestingly, similar effects could so far not been shown for TV or other video/multimedia consumption (even when focusing on documentaries). Being able to just count the words we read each day would help to assess the general knowledge of a person, as there are strong correlations between the two [3].

Exploring the reading activities of people, in general, and a Wordometer in particular is also an interesting for the document analysis, as we would be able to provide more insides into how/when documents are read (e.g. number of words read per page statistics or "this is the most read paragraph/page in this book").

The main contributions we present in this paper are as follows:

1) We present an algorithm based on eye gazes obtained by a mobile eye tracker to segment reading from not reading.
2) We present a method to segment read lines using the user's eye gaze and document image retrieval.
3) Based on the line estimates, we developed 2 methods to estimate the number of words a user reads using a mobile eye tracker and document image retrieval.
4) Our best algorithm to estimate word numbers has an error rate of only between 6 -8 % evaluated in a study with 10 users and 10 documents.

Achieving an error rate of 6-8 % for our Wordometer is reasonable comparing it with the pedometer error rate which is between 3-10 % [1]. The pedometer studies are by far more representative. Yet, this is the first time to our knowledge somebody tired to estimate the amount of words read and 6-8% seem to be a good error margin for a problem more difficult then counting steps. Section II is devoted to describe our approach to implement the Wordometer. Sections III and IV are for experimental results and discussions. Section V is to sum up what we have learnt from this research activity.

## II. Approach

We are interested in the number of words read. We could use optical character recognition/ text detection or similar technologies to record all characters a user sees assuming that the amount of text surrounding people is indicative for how many words they read. However, this will only be a rough estimate at best. It is easy to think of scenarios where this utterly fails. Imagine an American tourist in a Japanese city with a lot of billboards, although he's exposed to a lot of text, the words read are close to zero (assuming he does not understand Japanese). Therefore, we have to monitor the reading –the decoding of letters, words and assigning meaning to them.

As reading is a cognitive process, sensing brain activity seems to provide the most insight. However, directly sensing brain activity can only be done by more or less invasive methods (e.g. electroencephalography, functional magnetic resonance imaging, electrocorticography). On the other hands, eye movements and gaze are strongly correlated with reading and text comprehension [4], [5], [6] and it can be done by relatively unobtrusive already today.

Our approach has several discrete steps:

[1] http://fitbit.com

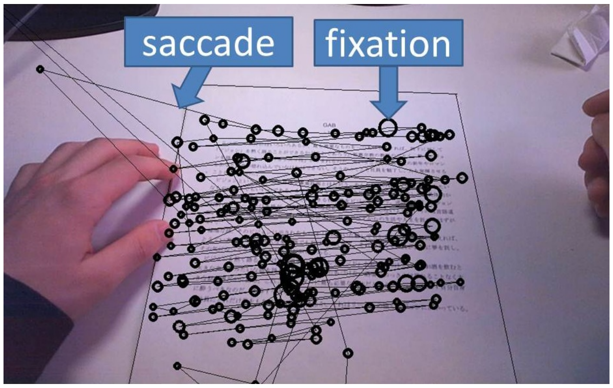

Fig. 1. Eye fixations and saccade traces while reading a document.

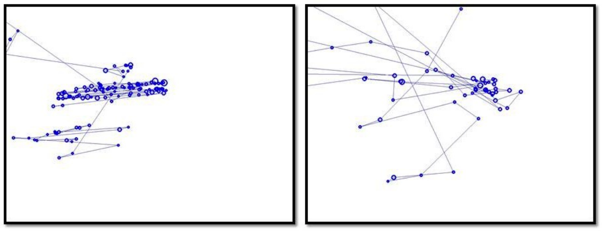

Fig. 2. Eyetracking data for Reading (left) versus not Reading (right)

1) The raw eye gaze data is summarized into fixations fixations and saccades according to Busher et al. [7] (see Figure 1).
2) On basis of the optioned eye gaze data we calculate features and classify reading from not reading.
3) On reading data, we use the video recording from the eye tracker to apply a document image retrieval technique (this is used to filter some head movements and get the average word count by line for the document).
4) A line detection algorithm is applied on the rectified eye gaze data.
5) Using the numbers of lines detected we calculate the words per line using 2 methods: the first by multiplying it with the average words per line of the document, the second by using support vector regression.

### A. Reading Detection

There are already a couple of methods in the related work [8], [9]. However, they work with different sensing modalities (e.g. electrooculography) and eye tracking hardware. Therefore we decided to implement our own algorithm.

The process is straight forward: we calculate eye gaze features given in Tabel I over a 3sec. frame sliding window and apply a Support Vector Machine classifier with a radial basis function on the resulting feature vector.

### B. Applying document image retrieval

With the help of an eyetracker and document image retrieval, we are able to obtain information on the document the user is reading as well as his/her eye gaze information.

TABLE I. FEATURES FOR READING DETECTION

| | feature |
|---|---|
| fixation related features | the number of fixations<br>sum of the duration of fixations<br>average time of fixations |
| saccade related features | average length of saccades<br>minimum length of saccades<br>horizontal element of saccades<br>vertical element of saccades |
| wavelet | average amplitude of wavelets |

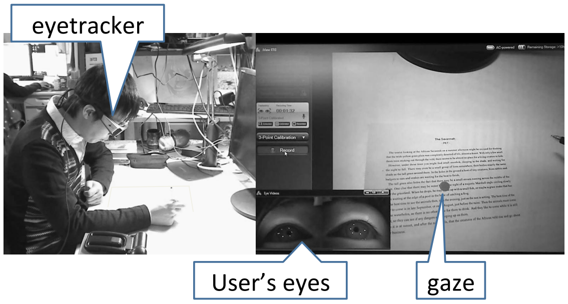

Fig. 3. Conversion of an eye gaze by using LLAH.

We use a document image retrieval method based on “Locally Likely Arrangement Hashing” (LLAH) to associate the paper with the digital document [10]. We use the video feed from the eye tracker as input to LLAH. LLAH retrieves the corresponding page from a document image database by comparing feature points. The eye gaze data obtained by the eye tracker is represented in the coordinate system of a scene camera of the eyetracker. We use LLAH to convert the eye gaze coordinates to the corresponding coordinates of the document by using a homography estimate.

Figure 3 illustrates an example of an eye gaze converted from a camera captured image (left) to its corresponding retrieved document image (right). On the left, the rectangle represents the focused area in which the corresponding document is retrieved. With this method we avoid potential confusion caused by other documents also captured in the image. By using this capability, we can record eye gazes on the coordinate system of a retrieved document (see Figure 4 for an example).

### C. Detection of Line Breaks

Figure 5 shows a typical eye movement when a person reads over line breaks. In our proposed method, we segment a sequence of eye movements into parts corresponding to text lines by analyzing them on the coordinate system of a retrieved document.

The eye movement for a line break is against the regular reading order from the end of the text line backwards to the start of the new line. We detect such a movement using the following line break detection.

Let $(g_1, ..., g_l)$ represent a subsequence of gaze points. A current gaze $g_l$ is recognized as a line-break if the following conditions are satisfied:

1) the area of a rectangle that circumscribes the $l$ gazes $g_1, ..., g_l$ is larger than a certain threshold,

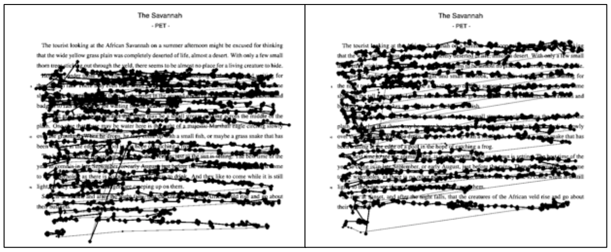

Fig. 4. Two examples of an eye gaze conversion by using LLAH. The right sample works very well, in the left sample an offset is introduced.

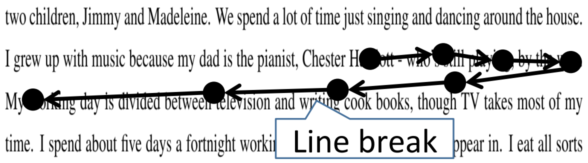
two children, Jimmy and Madeleine. We spend a lot of time just singing and dancing around the house.
I grew up with music because my dad is the pianist, Chester H
My
day is divided between television and w
cook books, though TV takes most of my
time. I spend about five days a fortnight worki
ppear in. I eat all sorts
Line break

Fig. 5. Example of eye gaze data from a user reading over a line break

2) the direction of eye movements from the $l-k$ th to the $l$ th gaze is opposite. Some succeeding gazes are also included as a part of a line break.

### D. Wordometer

Next, we apply our line break detector to build the Wordometer.

In general, there are many ways to implement a Wordometer. If we are able to record all read words accurately, it is trivial to realize the Wordometer based on the record. However, it is not easy to achieve such an accuracy even with the help of document image retrieval. If we are interested in just counting the rough number of read words, it is not necessary to have high accuracy; some errors are acceptable just as is the case of pedometers.

In this research we implemented two methods for a Wordometer based on the line break detector as follows.

First we build a very simple method. Let $N$ be the average number of words in one line of a page the user is reading, and $L$ be the number of estimated line breaks. Note that $N$ can be obtained by using document image retrieval. The number of read words is estimated as $N(L+1)$. The more errors we have in the number of estimated line breaks, the estimated number of read words gets more inaccurate.

The second method is more sophisticated. In order to improve the accuracy, we employ a feature vector $\boldsymbol{x}$ that represents eye movement information and estimate the number of read words by a function $f(N, L,\, \boldsymbol{x})$, which is learnt by using the learning samples. As a learner, we employ the SVR (support vector regression). In the following this method is called the SVR-based method.

Table II shows features in the vector $\boldsymbol{x}$ we obtain from eye movement data.

TABLE II. FEATURES

| | |
|---|---|
| gaze information | duration required for reading |
| | the number of fixations for a page |
| | total distance of eye movements |
| | total distance of saccades |
| | average distance of saccades |

Fig. 6. the SMI iViewX Eye Tracking Glasses used for the experiments

## III. Experiments

We asked 10 subjects to read 10 pages of documents written in English and record their eye movements by using the SMI iViewX ETG (Eye Tracking Glasses), see Figure 6. Among the subjects, 7 are with unaided vision, 2 are with contact lenses on both eyes, and one with a contact lens on a single eye. Documents are from tests of PET (Preliminary English Test) in ESOL(English for Speakers of Other Languages) by University of Cambridge. For each document we have a time limit defined based on the number of words included in it.

For the reading detection we designed another experiment including also the 10 subjects reading 10 documents from a job hunting test. Other than that the conditions are the same. The not reading activity included looking around in the room, playing with objects (e.g. cellphone etc.).

The following is the procedure of experiments. Each subject was requested to wear the eyetracker for calibration. Then he/she is requested to read a document placed on a desk in front of him/her and with a natural posture. The calibration of the eyetracker is applied after reading each document. Then we applied $k$-fold cross validation with $k$ subjects.

## IV. Results and Discussion

### A. Reading Detection

Reading versus not reading can be easily distinguished just by visualizing the the eye gaze data for the different classes, as depicted in Figure 2. As expected the reading detection works reliably well with an average of 91 % over the 10 subjects, see Table III for details.

### B. Line break estimation

One test subject with unaided vision often fell asleep during the data recording and his recorded data showed a lot of inconsistencies. Therefore, we decided to exclude him and the data from the remaining 9 subjects are utilized for 9-fold cross validation.

Table IV shows errors observed during the line break detection. The average errors of the estimated number of lines is 2.1 lines per document. Since the average number of lines of documents is 17, we are able to estimate the number of lines with 12.4% errors.

From the Table IV , we can observe that the results by the subject h are much worse than other results. His data includes

TABLE IV. ERRORS OF LINE BREAK DETECTION.

|  |  | subject |  |  |  |  |  |  |  |  | Ave. error |
|---|---|---|---|---|---|---|---|---|---|---|---|
| doc. | # lines | a | b | c | d | e | f | g | h | i | for each doc. |
| A | 15 | -1 | 2 | 0 | 1 | 3 | -1 | 0 | 10 | 8 | 2.4 |
| B | 15 | -1 | 2 | 0 | 2 | 2 | 2 | 1 | 11 | 2 | 2.3 |
| C | 12 | 3 | 10 | 1 | 3 | 0 | 2 | 1 | 13 | 6 | 4.3 |
| D | 19 | -3 | 7 | -2 | -2 | 1 | 0 | 0 | 5 | 4 | 1.1 |
| E | 18 | -1 | 1 | 1 | 1 | -1 | 1 | -1 | 10 | 4 | 1.6 |
| F | 19 | -6 | -2 | -3 | 1 | -1 | -2 | -3 | 10 | 2 | -0.4 |
| G | 18 | 2 | 8 | -1 | 0 | 2 | 1 | -1 | 2 | 3 | 1.7 |
| H | 17 | -1 | 5 | 1 | 2 | 2 | 2 | 0 | 5 | 5 | 2.3 |
| I | 16 | 1 | 2 | 0 | 5 | 2 | 4 | 0 | 7 | 1 | 2.4 |
| J | 21 | -1 | 5 | 2 | 5 | 2 | 0 | 0 | 10 | 7 | 3.3 |
| Ave. errors |  | -0.8 | 4 | -0.1 | 1.8 | 1.2 | 0.9 | -0.3 | 8.3 | 4.2 | 2.1 |

TABLE III. READING DETECTION RESULTS

| User | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Accuracy(%) | 100 | 100 | 100 | 80 | 100 | 80 | 75 | 100 | 100 | 75 | 91 |

more frequent eye blinks than any other subjects. In addition, he often re-read the end of the line just after the line break. This behavior erroneously increased the number of detected line breaks.

Another source of errors are short lines. The proposed method tends not to detect such short lines so that the resultant estimated number is smaller than the groundtruth. The reason is two-fold:

1) The detection is based on the area of a bounding box of gazes. Since short lines only have small number of gazes and thus the area is sometimes too small.
2) For most subjects, the duration of gazing a short line is too small.

The results for document C are clearly worse compared to other documents. Subjects b and h often went back to previous lines and read them for this document again.

### C. Wordometer

Table V shows the number of read words for each document estimated by the simple method. The average error is 47.95 words. Since the average number of words in a document is 279.3, the error is 17.2%. Table VI shows the number of read words for each document estimated by the SVR method. The average error is 40.5 words (14.5%) thus the SVR method is superior.

Table VII shows the number of estimated words for all documents. For the SVR method we achieve an average error of 8.2 %, however we get a really high error rate from subject h due to him also re-reading several lines. If we remove subject h from the analysis we are can improve the error to just about 6.5 %.

## V. RELATED WORK

A good overview about controlled studies in psychology about eye movements in relation to reading is given by Rayner [11] Kligel et al. explore the relations of eye fixations and mental activities during reading [12].

The existing work so far focuses on controlled lab experiments on specific population groups –mostly elderly and children– trying to diagnose specific disease or mental states (attention deficit, alzheimer etc.) [5], [4]. There is some work showing the strong relation between how much a person reads and their language skills, general knowledge and critical thinking skills [11], [2], [4].

The closest to the work presented here is by Bulling et. al. They present a method to for reading detection using electrooculography [13]. Although their work looks at reading detection in a realistic environment, they do not try to estimate the word count.

Other research classifies the different types of reading, skimming etc. [14], [8]. Xu et al. use eye tracking in a controlled lab environment to create summaries of documents [15].

## VI. CONCLUSION AND FUTURE WORK

We presented a reading segmentation algorithm, a line break detector and two methods to implement a Wordometer.

In conclusion, our error margin for the Wordometer is 8.2 % ( or 6.5 % removing one test subject). Comparing it to the effectiveness of a pedometers in realistic scenarios ( error rate of 3-10 % ), we believe 8.2 % is a good baseline for counting words regarding the impact a Wordometer could have on quantifying language skills and general knowledge.

For future work we want to overcome the limitations of document image retrieval (all documents the user reads need to be registered with the system). Currently our Wordometer methods still rely on document image retrieval for:

1) the correction of the eye gaze (translation to the document coordinate system)
2) the average word count per document.

For the eye gaze correction we can use optical flow. There is related work that shows the successful application of optical flow to filter out head movements during tomography [**?**]. The simple solution regarding the average word count per line is to apply a constant. Yet this will likely increase the error rate. We will also try an estimation of words per line using the eye fixations and saccades per line.

## ACKNOWLEDGMENT

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

TABLE V. NUMBER OF READ WORDS FOR EACH DOCUMENT ESTIMATED BY THE SIMPLE METHOD.

| | | subject | | | | | | | | | doc-wise |
|---|---|---|---|---|---|---|---|---|---|---|---|
| doc. | # words | a | b | c | d | e | f | g | h | i | ave. error |
| A | 255 | 238.0 | 289.0 | 255.0 | 334.7 | 306.0 | 318.8 | 255.0 | 366.6 | 391.0 | 32.9 |
| B | 203 | 189.4 | 230.1 | 203.0 | 230.1 | 230.1 | 230.1 | 216.5 | 351.9 | 230.1 | 49.1 |
| C | 210 | 262.5 | 385.0 | 227.5 | 262.5 | 210.0 | 245.0 | 227.5 | 437.5 | 315.0 | 61.8 |
| D | 302 | 219.6 | 356.9 | 247.1 | 233.4 | 274.5 | 260.8 | 260.8 | 329.5 | 315.7 | 40.5 |
| E | 296 | 279.5 | 312.4 | 312.4 | 312.4 | 279.6 | 312.4 | 279.6 | 460.4 | 361.8 | 33.8 |
| F | 276 | 188.8 | 246.9 | 246.9 | 290.5 | 261.5 | 246.9 | 232.4 | 421.3 | 305.1 | 28.9 |
| G | 281 | 312.2 | 405.9 | 265.4 | 281.0 | 312.2 | 296.6 | 265.4 | 312.2 | 327.8 | 23.2 |
| H | 298 | 280.5 | 385.6 | 315.5 | 333.1 | 333.1 | 333.1 | 298.0 | 385.6 | 385.6 | 29.3 |
| I | 255 | 270.9 | 286.9 | 255.0 | 272.0 | 286.9 | 238.0 | 255.0 | 425.0 | 270.9 | 21.3 |
| J | 417 | 397.1 | 516.3 | 456.7 | 516.3 | 456.7 | 417.0 | 417.0 | 615.6 | 556.0 | 70.6 |
| ave. error | | 35.4 | 68.0 | 19.1 | 41.0 | 27.4 | 28.0 | 14.8 | 131.3 | 66.6 | 47.95 |

TABLE VI. NUMBER OF READ WORDS FOR EACH DOCUMENT ESTIMATED BY THE SVR-BASED METHOD.

| | | subject | | | | | | | | | doc-wise |
|---|---|---|---|---|---|---|---|---|---|---|---|
| doc. | # words | a | b | c | d | e | f | g | h | i | ave. error |
| A | 255 | 196.7 | 266.0 | 278.9 | 288.9 | 255.6 | 276.0 | 241.8 | 316.2 | 327.8 | 54.8 |
| B | 203 | 205.9 | 227.6 | 277.0 | 257.2 | 252.8 | 236.4 | 260.9 | 310.6 | 240.4 | 34.6 |
| C | 210 | 249.0 | 309.9 | 277.5 | 232.4 | 252.8 | 253.1 | 221.1 | 356.2 | 294.2 | 75.9 |
| D | 302 | 153.2 | 289.9 | 277.9 | 281.1 | 252.8 | 249.1 | 275.2 | 298.8 | 275.5 | 45.8 |
| E | 296 | 237.3 | 283.3 | 281.3 | 266.9 | 252.8 | 276.0 | 266.6 | 370.8 | 317.9 | 38.4 |
| F | 276 | 216.5 | 235.5 | 279.1 | 266.4 | 252.8 | 242.8 | 264.4 | 353.8 | 278.1 | 46.8 |
| G | 281 | 231.1 | 310.5 | 279.0 | 250.9 | 252.8 | 265.6 | 256.9 | 275.4 | 305.2 | 34.7 |
| H | 298 | 250.7 | 318.3 | 281.3 | 271.8 | 252.8 | 286.6 | 272.7 | 336.4 | 330.7 | 44.8 |
| I | 255 | 230.1 | 269.3 | 278.8 | 249.5 | 252.8 | 249.1 | 244.4 | 341.6 | 272.6 | 33.3 |
| J | 417 | 249.6 | 376.8 | 285.2 | 391.6 | 252.8 | 338.2 | 373.5 | 441.8 | 406.1 | 76.3 |
| ave error | | 65.6 | 30.5 | 38.1 | 25.7 | 44.8 | 31.5 | 25.4 | 62.6 | 33.0 | 40.53 |

TABLE VII. ESTIMATED NUMBER OF READ WORDS FOR ALL DOCUMENTS

| | subject | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | a | b | c | d | e | f | g | h | i | ave. error |
| simple method | 2638.8 | 3415.1 | 2784.6 | 3065.9 | 2950.5 | 2898.7 | 2707.2 | 4105.5 | 3459.0 | |
| errors by the simple method | 154.2<br>(5.5%) | 622.1<br>(22.2%) | 8.4<br>(0.3%) | 272.9<br>(9.7%) | 157.5<br>(5.6%) | 105.7<br>(3.8%) | 85.8<br>(3.1%) | 1312.5<br>(47.0%) | 666.0<br>(23.8%) | 376.1<br>(13.5%) |
| SVR-based method | 2220.4 | 2887.5 | 2796.0 | 2756.8 | 2530.8 | 2672.9 | 2677.4 | 3401.5 | 3048.5 | |
| errors by the SVR-based method | 572.6<br>(20.1%) | 94.5<br>(3.4%) | 3.0<br>(0.1%) | 36.2<br>(1.3%) | 262.2<br>(9.4%) | 120.1<br>(4.3%) | 115.6<br>(4.1%) | 608.5<br>(21.8%) | 255.5<br>(9.1%) | 229.8<br>(8.2%) |

## REFERENCES

[1] D. M. Bravata, C. Smith-Spangler, V. Sundaram, A. L. Gienger, N. Lin, R. Lewis, C. D. Stave, I. Olkin, and J. R. Sirard, “Using pedometers to increase physical activity and improve health,” *JAMA: the journal of the American Medical Association*, vol. 298, no. 19, pp. 2296–2304, 2007.

[2] A. Cunningham and K. Stanovich, “What reading does for the mind,” *Journal of Direct Instruction*, vol. 1, no. 2, pp. 137–149, 2001.

[3] P. Terenzini, L. Springer, E. Pascarella, and A. Nora, “Influences affecting the development of students’ critical thinking skills,” *Research in higher education*, vol. 36, no. 1, pp. 23–39, 1995.

[4] A. R. Clarke, R. J. Barry, R. McCarthy, and M. Selikowitz, “Eeg analysis of children with attention-deficit/hyperactivity disorder and comorbid reading disabilities,” *Journal of Learning Disabilities*, pp. 276–285, 2002.

[5] E. C. Ferstl, J. Neumann, C. Bogler, and D. Y. von Cramon, “The extended language network: A meta-analysis of neuroimaging studies on text comprehension,” *Human Brain Mapping*, pp. 581–593, 2008.

[6] O. Dimigen, W. Sommer, A. Hohlfeld, A. Jacobs, and R. Kliegl, “Coregistration of eye movements and eeg in natural reading: analyses and review.” *Journal of Experimental Psychology: General*, vol. 140, no. 4, p. 552, 2011.

[7] G. Buscher and A. Dengel, “Gaze-based filtering of relevant document segments,” in *Workshop on Web Search Result Summarization and Presentation. Workshop on Web Search Result Summarization and Presentation (WSSP-2009), located at in conjunction with WWW*, vol. 9, 2009, pp. 20–24.

[8] R. Biedert, J. Hees, A. Dengel, and G. Buscher, “A robust realtime reading-skimming classifier,” in *Proc. of ETRA ’12*, 2012, pp. 123–130.

[9] A. Bulling, J. Ward, H. Gellersen, and G. Troster, “Eye movement analysis for activity recognition using electrooculography,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 4, pp. 741–753, 2011.

[10] T. Nakai, K. Kise, and M. Iwamura, “Use of affine invariants in locally likely arrangement hashing for camera-based document image retrieval,” *In Proc. of DAS 2006*, vol. 3872, pp. 541–552, Feb. 2006.

[11] K. Rayner, “Eye movements in reading and information processing: 20 years of research.” *Psychological bulletin*, vol. 124, no. 3, p. 372, 1998.

[12] R. Kliegl, A. Nuthmann, and R. Engbert, “Tracking the mind during reading: the influence of past, present, and future words on fixation durations.” *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General*, vol. 135, no. 1, p. 12, 2006.

[13] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, “Robust recognition of reading activity in transit using wearable electrooculography,” in *Proc. of Pervasive ’08*, 2008, pp. 19–37.

[14] G. B. Duggan and S. J. Payne, “Skim reading by satisficing: evidence from eye tracking,” in *Proc. of CHI 2011*, 2011, pp. 1141–1150.

[15] S. Xu, H. Jiang, and F. C. Lau, “User-oriented document summarization through vision-based eye-tracking,” in *Proc of IUI*, ser. IUI ’09, 2009, pp. 7–16.