Reading activity recognition using an off-the-shelf EEG — detecting reading activities and distinguishing genres of documents

Kai Kunze, Yuki Shiga, Shoya Ishimaru, Koichi Kise
Osaka Prefecture University, Japan
{kunze, shiga, ishimaru}@m.cs.osakafu-u.ac.jp
kise@cs.osakafu-u.ac.jp

***Abstract*—The document analysis community spends substantial resources towards computer recognition of any type of text (e.g. characters, handwriting, document structure etc.). In this paper, we introduce a new paradigm focusing on recognizing the activities and habits of users while they are reading. We describe the differences to the traditional approaches of document analysis. We present initial work towards recognizing reading activities. We report our initial findings using a commercial, dry electrode Electroencephalography (EEG) system. We show the feasibility to distinguish reading tasks for 3 different document genres with one user and near perfect accuracy. Distinguishing reading tasks for 3 different document types we achieve 97 % with user specific training. We present evidence that reading and non-reading related activities can be separated over 3 users using 6 classes, perfectly separating reading from non-reading. A simple EEG system seems suitable for distinguishing the reading of different document genres.**

I. Introduction

Documents are the medium to structure and provide this information to us [8]. Although we in turn use reading – the process of assigning meaning to characters, words and sentences– as a primary information source, recognizing and monitoring reading activities in unconstrained, natural settings is still largely unexplored, with few notable exceptions [1], [2]. We propose reading activity recognition to analyze the whole process of human reading activity and knowledge acquisition. As compared to the traditional document analysis that focuses on analyzing documents as objects in the environment, reading activity recognition is to analyze reading as an event by the user who is placed in the environment. As tools for the analysis, it is not necessary for us to focus just on vision sensors, which are mainly used in the document analysis research, but also able to use other sensing modalities such as motion sensors, galvanic skin response and EEG.

As a part of the research about reading activity recognition, we have already started to implement the reading-life log, which is to log all text read by the user, as a record of the reading activity [16]. In this research optical character recognition is employed with a vision sensor mounted on the user. With the help of an eyetracker, the method is able to detect where the user is reading to pass this region into OCR to obtain the text. Another related work on reading activity recognition is the Wordometer, an implementation of counting words read by the user [10]. This research also relies on the vision sensor and the eyetracker along with document image retrieval to estimate the number of read words.

The insides gained from analyzing human reading habits can again be used to analyze documents. If we are able to detect reading reliably, we can mark the most read sentences in a text, the word most people stopped reading (e.g. because they lost interest, ...) or the preferred reading order of a document. Yet, more crucial, if we gain a better understanding of knowledge acquisition, a computer system can detect the most difficult words/passages in a document for a particular reader or the system can recommend specific information to fill knowledge gaps.

The purpose of this paper is to propose a different scenario in the context of reading activity recognition by using an EEG as a sensor. Needless to say but reading is a cognitive process by the user so that the activity is well reflected to the EEG, though it is often noisy. Although in the past it is not realistic to consider using EEG as a sensor, there are several, off-the-shelf and easy-to-use EEG sensors available nowadays. Our trial is to implement the following two functionalities by using such a "normal" device as a sensor for reading activities:

1) Discrimination of reading and not-reading activities
2) Genre recognition of documents. The method can distinguish a genre of documents the user is reading. As genres, we consider scientific paper, manga, and novel, in addition to doing nothing.

From the preliminary experimental results, we conclude that even with a off-the-shelf EEG sensors, we are able to implement the above functions with reasonable accuracies.

The organization of the paper is as follows. In Section II, we contrast the research field "reading activity recognition" to the ordinary "document analysis" to clarify the difference and importance of this new field. Section III is to describe reading activity recognition based on EEG. In Section IV, we report the preliminary experimental results on the reading activity recognition to show the potential of this research field. Section V is for concluding what we have achieved as well as to discuss the remaining future work.

II. Reading Activity Recogintion

In this section we cut clearly the standpoint of the new research field "reading activity recognition" by comparing it to the traditional document analysis. For this purpose, we employ three axes of comparison: a target to be analyzed, sensors, and persons. A comparison is shown in Fig. 1 focusing on sensors and persons.

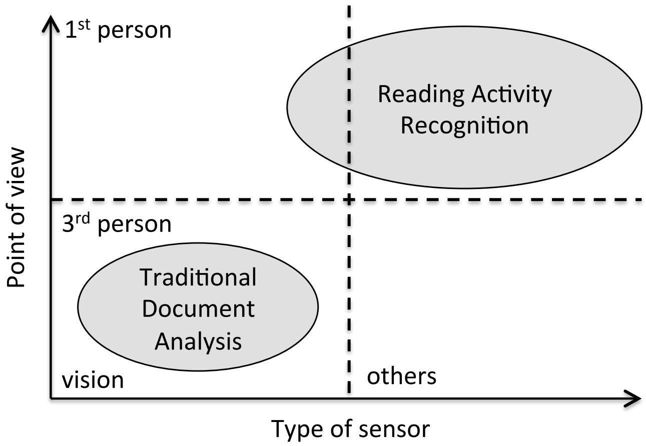

Fig. 1. Differentiating our novel approach to traditional document analysis

Let us start with discussing the target. In the traditional document analysis, documents are always the target to be analyzed. The final goal of document analysis is to reconstruct an electronic equivalent from a physical document captured as an image. A person is always a user of document analysis to receive the result of processing. On the other hand, the goal of reading activity recognition is to analyze the user's event of reading. Documents and characters have their role within the activity of reading but the leading player is the user. What we obtain as a result of analysis is the record of reading activity by the user. Thus even with the same document, a user may receive little information due to the inability of understanding the contents. This type of analysis about the relationship between objects (documents and characters) and an actor (the user) is included in the research of reading activity recognition. One final goal behind reading activity recognition is to reconstruct the information a user absorbed through reading.

For implementing such analysis, it is not necessary to limit ourselves to just vision sensors. In the traditional document analysis, the typical sensors used for the research is vision sensors, simply because they are well fit for the final goal. For the reading activity recognition, vision sensors are also quite important, especially to know documents and characters the user is reading. However, the sensors are not necessarily limited to the vision — other sensors can also be used if they offer us better sensory input of human reading activities. For example, accelerometers on the head can sense the users focus on reading since one should keep our head still enough to concentrate on reading. In this paper we pursue the possibility of EEG as a sensor.

The final comparison axis is person. Since the interest of document analysis is on the object "documents" or "characters", sensors are always apart from the user. Thus, this can be called "third person sensors". In other words, a sensor is sensing with a different viewpoint from the user. For the reading activity recognition, on the other hand, the person should not be limited to the third. Rather, the first person sensors, which sense data with the same viewpoint of the user, are more important. As for the vision sensors, a camera can be placed on the user to share the viewpoint. It is best to capture the document the user is reading. The EEG is also an first person sensor since it directly senses the user.

To sum up the following statements can be made to differentiate reading activity recognition from the traditional document analysis.

1) Document analysis mostly uses third person vision sensors to sense documents and characters as objects,
2) Reading activity recognition is to use other sensing modalities also in the context of first person to sense the reading activity by the user as event. The goal is to know the user with respect to the result of reading.

There are a lot of different categories that constitute to human reading habits (e.g. reading techniques, attention focus) and there exist a lot of other sensory clues and technologies we can use to help us record and monitor reading activities. For example using a person's activity or location as a prior to text recognition. Regarding the recognition of reading activities we see three technologies as most relevant:

- First Person Vision – By artificially limiting the text recognition to the human field of vision, we can narrow down on the text possibly read by the user.
- Sensing eye movement – eye tracking can be very useful to determine if a user is reading and can give substantial clues about their attention level [2], [5]. Although eye-tracking systems today are still expensive, as soon as compelling applications are available scale with drive down price (as parts are relatively cheap). Mobile eye trackers today are already only a little bit larger than regular glasses.
- Sensing brain activity – As we want to explore reading activities and given that reading is a cognitive process, sensing brain activity seems to provide the most insight. As we want to stay as cheap and unobtrusive as possible, the only feasible option for sensing brain activity seems Electroencephalography (EEG). Other approaches – e.g. functional magnetic resonance imaging (fMRI), Electrocorticography) – use bulky devices (sometimes filling a complete room) and are very expensive. These approaches could not be used in a real life setting today. deployed.

EEG seems the most promising and unexplored sensing modality in regard to reading. Therefore we will focus on it for the rest of this paper.

III. EEG Feasibility Study

To evaluate EEG, we focus on a distinct subset of reading activities namely recognizing the reading of different document types (e.g. a novel versus a comic book) as well as on distinguishing these reading tasks from specific non-reading related activities (e.g. listening to music, drawing sth.). We do the first steps towards recognizing reading activities and contribute towards answering the following research questions:

1) Can we distinguish different types of reading activities using a relatively cheap, commercially available EEG device?
2) Is it feasible to segment reading from non-reading activities with such a device?

Fig. 2. The Emotiv EGG system used for the experiments.

To our knowledge nobody recognized users reading different genres of documents, so far.

A. Experimental Design and Setup

To check for the feasibility of our approach, we devised 2 small scale experiments. They can be seen as indication if the research questions posed can be tackled with the sensing technologies available today.

There are several commercially available dry electrode EEG devices available. We used the Emotive EEG (see Figure 2. It is comparably cheap and has a good resolution (14 EEG channels at 128 Hz) [1]. Additionally, a substantial research and developer community is already using it [18].

1) distinguishing different reading materials: To get some information on the question "Can we distinguish if a user is reading different documents using a cheap device?", we designed a feasibility experiment containing a number of different documents: a scientific paper, a manga, a novel and doing nothing (as reference). All documents were available electronically and printed on the same paper type. The "doing nothing activity" consisted in sitting in a chair looking outside the window and at specific objects (also including head movements). So far, we only recorded one test subject performing each task 3 times for 5 minutes in an office environment.

2) User-dependence of reading tasks versus non-reading tasks: To tackle the questions "Can we segment reading from non-reading activities?" and "How user-dependent is the reading of different document types?" we devised a second experimental setup, evaluating in total 6 reading and not-reading activities. The reading materials include a scientific paper, a news website and a manga. The non-reading activities include drawing, listening to music and watching a video. We record 3 participants performing these activities for 3 minutes in an office environment.

B. Analysis and Findings

EEG is quite noisy and often overshadowed my facial movements. To improve the signal-to-noise ratio and reduce

[1] http://emotiv.com

a	b	c	d	e	f	← classified as
50	50	0	0	0	0	a = manga
66	0	33	0	0	0	b = news
50	0	50	0	0	0	c = paper
0	0	0	100	0	0	d = draw
0	0	0	0	100	0	e = music
0	0	0	0	0	100	f = video

Fig. 4. The confusion matrix in percent for activities from all three users using a KNN on the EEG bands using a 3-fold crossvalidation.

unwanted movement effects we apply a bandpass filter between 1-30Hz. Afterwards, we calculate the Fast Fourier Transform (FFT) of the EEG signal using a sliding window (500 msec. sliding window step size 20 msec.) and sum the frequencies according to the classification of brain bands: beta (13-30Hz), alpha (8-12 Hz), theta (4-8 Hz), and delta (less than 4 Hz). This leaves us with the sum of alpha, beta, theta and delta waves per channel per window as features for a k-nearest neighbor algorithm with k=3. We validate our model using a 10-fold cross validation 33-66 percentage split for training and testing datasets for each round.

For the initial experiment with one user and distinguishing reading science paper, novel, manga and "doing nothing" we reach close to 100 % (except 1-2 window misclassifications of the "doing nothing" nothing class).

We can easily see why the classification algorithm is doing such a good job, by comparing the independent components of the EEG signal for reading tasks visually (see Figures 3 and 5). Similar reading tasks have several components that are the same, whereas different reading tasks have largely different components.

Applying the same features, algorithm and a 3-fold cross validation ((training on one user evaluating on the others) over the data set for the 3 participants and the reading and non-reading activities leaves us with approximately 78% correctly classified and a confusion matrix given in Figure 4. We are able to separate between non-reading tasks and reading tasks. The non-reading tasks are classified correctly over all 3 participants. All miss-classifications happen in the reading related tasks. Taking a closer look it seems there are strong differences between users in this case.

This leads us to further inspecting the EEG data for different reading tasks. Plotting the independent components for the different reading tasks per participant for the EEG signal, we find for each distinct reading task several independent components that are very similar between users. Examples for similar independent components are shown in Figure 5.

Unfortunately, there are also several, unrelated components that most likely correspond to muscle movement (e.g. component 2 for participant 1 in Figure 3, in this case the activation of the EEG signal is concentrated only close to the eyes). The band pass filtering could not remove all motion artifacts. To get rid of these components would require manual inspection of the data. Additionally, some muscle activity might be very helpful discriminating between reading and not-reading related

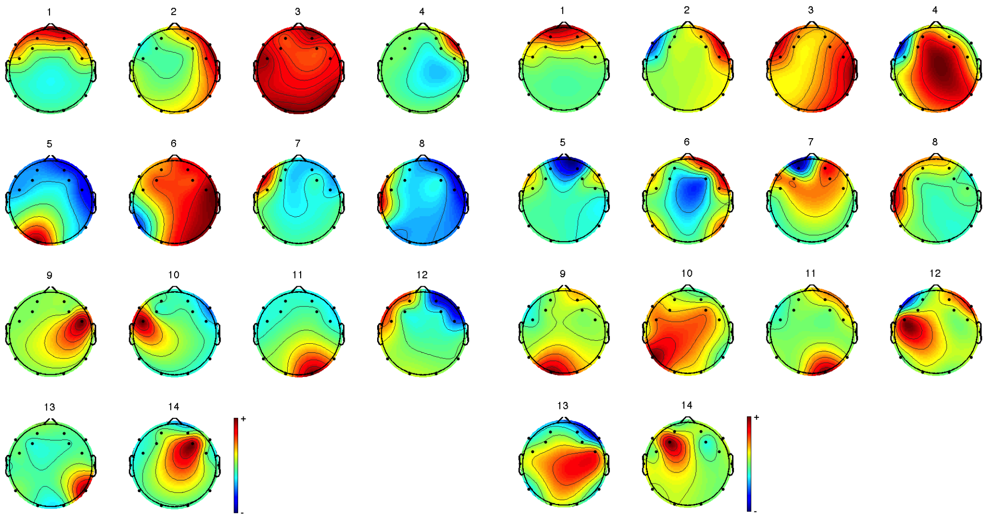

Fig. 3. Independent Components of the 14 EEG channels for one participant reading a research paper (left) and reading a manga (right). There are several components characteristic for reading a research paper (e.g. component 3 and 6) and manga (component 3 and 4) in all recordings , that are missing in the other reading tasks. Unfortunately the ordering of the components is highly influenced by noise from head motions.

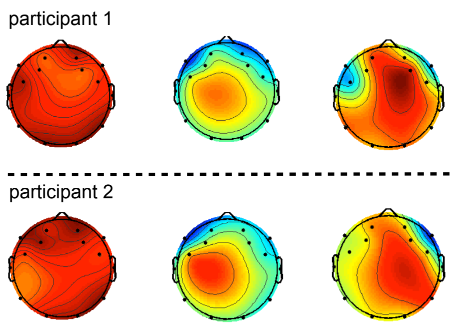

Fig. 5. Comparing similar independent components of different participants and the same task. From left to right: reading a science paper, reading the news, reading manga.

tasks (e.g. specific eye movements with a given frequency for reading). More data and closer inspection will help us to remove unwanted muscle artifacts while retaining the relevant components.

The analysis is done using a combination of ipython [15] and matlab scripts and will be available under . The plots are done using the opensource EEGLAB Matlab toolbox [4]. Distinguishing different document types while reading seems feasible with a cheap EEG device.

Assuming EEG is as useful as our initial feasibility study indicates, this opens the door to novel type of document analysis. A mobile eye tracker or EEG can discriminate if the user or not. In addition to document type, EEG signals have been used to to classify the emotional state of people [11], [13]. The Emotiv device used in the experiments comes with algorithms to detect some emotions and facial expressions (e.g. moving the eyebrows, blicking, laughing, talking). Combining it with traditional document analysis techniques, documents can be enriched with emotional information (e.g. "this page is the saddest for most readers").

IV. Related Work

There is a large corpus of work to better understand reading from a cognitive perspective, focusing on reading comprehension, techniques and the associated mental processes [3], [7], [5].. This work is complimentary as it is mostly conducted in a controlled lab setting with expensive equipment (fMRI). The bulk of this reading research explores reading related disease or mental deficiencies [7], [3].

Mostow et. al. is closer to our research, as they exploit the EEG input to help to teach children correct reading techniques [14]. Concerning reading, there are a few papers that try to segment reading from non-reading and classify (effective) skimming using eye tracking [6], [1].

Considering document analysis, eye tracking while reading can also be used to provide summaries of documents or to find which words a user finds relevant [19], [12].

V. Conclusion and Future Work

In this paper we introduced reading activity recognition, exploring a user's reading habits. We highlighted the similarities and differences to document analysis. Exploring the space of potential technologies, we concluded that EEG seems to be the most promising sensing modality in addition to vision sensors.

We present a feasibility study on how to use an off-the-shelf EEG device to segment reading from not reading and distinguish the document genre a user is reading. The reading detection works perfectly user-independent. The classification for the document genre works well for the user-dependent case 97 % (30 /60 split). Of course regarding the small user number, the results have to be carefully considered and need to reproduced in a larger, more representative study. Nonetheless, the study illustrates the usefulness of an EEG device to conduct research towards detecting reading activities.

Acknowledgment

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

References

[1] Ralf Biedert, Jörn Hees, Andreas Dengel, and Georg Buscher. A robust realtime reading-skimming classifier. In *Proc. of ETRA '12*, pages 123–130, 2012.

[2] Andreas Bulling, Jamie A. Ward, Hans Gellersen, and Gerhard Tröster. Robust recognition of reading activity in transit using wearable electrooculography. In *Proc. of Pervasive '08*, Pervasive '08, pages 19–37, 2008.

[3] Adam R. Clarke, Robert J. Barry, Rory McCarthy, and Mark Selikowitz. Eeg analysis of children with attention-deficit/hyperactivity disorder and comorbid reading disabilities. *Journal of Learning Disabilities*, pages 276–285, 2002.

[4] A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[5] O. Dimigen, W. Sommer, A. Hohlfeld, A.M. Jacobs, and R. Kliegl. Coregistration of eye movements and eeg in natural reading: analyses and review. *Journal of Experimental Psychology: General*, 140(4):552, 2011.

[6] Geoffrey B. Duggan and Stephen J. Payne. Skim reading by satisficing: evidence from eye tracking. In *Proc. of CHI 2011*, pages 1141–1150, 2011.

[7] Evelyn C. Ferstl, Jane Neumann, Carsten Bogler, and D. Yves von Cramon. The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, pages 581–593, 2008.

[8] Jack Goody. *The logic of writing and the organization of society*. Cambridge University Press, 1986.

[9] Masakazu Iwamura, Takuya Kobayashi, and Koichi Kise. Recognition of multiple characters in a scene image using arrangement of local features. In *ICDAR*, pages 1409–1413, 2011.

[10] K. Kunze, H. Kawaichi, K. Yoshimura, and K. Kise. The wordmeter estimating the number of words read using document image retrieval and mobile eye tracking. In *Submitted to ICDAR*, 2013.

[11] Yisi Liu, Olga Sourina, and Minh Nguyen. Real-time eeg-based emotion recognition and its applications. *Transactions on computational science XII*, pages 256–277, 2011.

[12] Tomasz D. Loboda, Peter Brusilovsky, and Jöerg Brunstein. Inferring word relevance from eye-movements of readers. In *Proc. of IUI '11*, pages 175–184, 2011.

[13] Jaakko Malmivuo and Robert Plonsey. *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.

[14] Jack Mostow, Kai-Min Chang, and Jessica Nelson. Toward exploiting eeg input in a reading tutor. In *Proceedings of*, AIED'11, 2011.

[15] Fernando Pérez and Brian E. Granger. IPython: a System for Interactive Scientific Computing. *Comput. Sci. Eng.*, 9(3):21–29, May 2007.

[16] K. Takashi, H. Rong, U. Seiichi, I. Masakazu, O. Shinichiro, and K. Koichi. A trial of reading-life log. In *Submitted to ICDAR*, 2013.

[17] Kazutaka Takeda, Koichi Kise, and Masakazu Iwamura. Real-time document image retrieval on a smartphone. *in Proc. of IAPR International Workshop on Document Analysis Systems*, pages 225–229, 2012.

[18] Chi Vi and Sriram Subramanian. Detecting error-related negativity for interaction design. In *CHI 2012*, 2012.

[19] Songhua Xu, Hao Jiang, and Francis C.M. Lau. User-oriented document summarization through vision-based eye-tracking. In *Proc of IUI*, IUI '09, pages 7–16, 2009.