eyeReading: Interaction with Text through Eyes

Seyyed Saleh Mozaffari Chanijani*†, Syed Saqib Bukhari* Andreas Dengel*†
*German Research Center for Artificial Intelligence, DFKI GmbH
†University of Kaiserslautern

I. Introduction

Research on reading has been started with the famous publication of Edmund Burke Huey's (1908): The Psychology and Pedagogy of Reading [1]. In the last decades, computing power increased dramatically. This evolvement paved a new path for the researchers to elevate reading research especially through the medium of eye tracking technologies. This made *text* very special in the domain of eye tracking [2]. In this paper, we introduce the *eyeReading* system which facilitates research in Psychology of Reading and provides a framework for gaze based Human-Text Interaction.

II. eyeReading Framework

The idea to integrate human gaze with the Document Object Model (DOM) has been investigated formerly in WebGaze-Analyzer [3] and Text2.0 [4]. On account of security issues that raised in the last decade, most of the tools used by mentioned framework have been deprecated. Unfortunately, new browsers are not supporting these frameworks anymore. Hence, there is a need to develop a robust system in line with the policies of World Wide Web Consortium (W3C) to prevent the same problem in the future. Moreover, with the favor of recent technologies in the web such as HTML5, CSS3, jQuery and *WebSocket* as a real-time protocol, we made a system with higher usability and of course functionality. WebSockets provide a persistent connection between a client and server that both parties can use to send and receive data in a real-time manner.

The top level architecture of the *eyeReading* framework consists of three components connected via WebSocket protocol: *Eye Tracking Server*, *eyeReading Client* and *eyeReading Server*. Figure 1. show top level view of eyeReading framework.

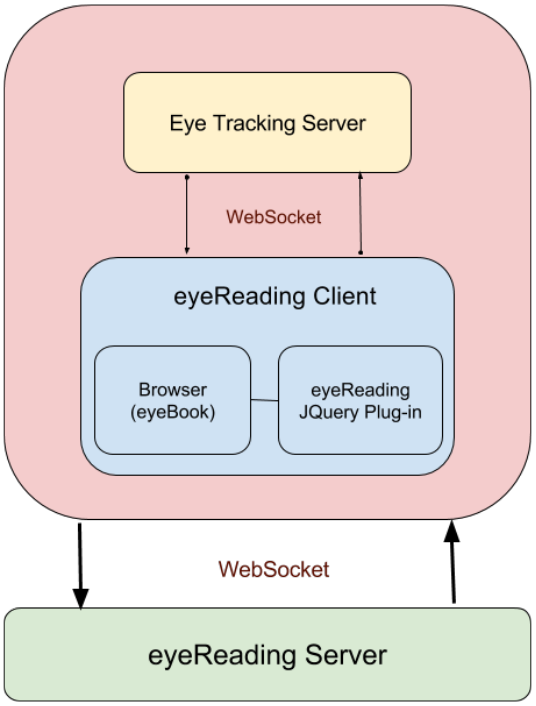

Fig. 1: The top level architecture of the *eyeReading* framework consists of three components connected via WebSocket protocol: *Eye Tracking Server*, *eyeReading Client* and *eyeReading Server*.

A. Eye Tracking Server

The *Eye Tracking Server* connects to the eye tracking device and registers for gaze data. The incoming data provided by the manufacturer SDK is then transformed into a message, serialized and sent to the *eyeReading Client* through WebSocket protocol.

B. eyeReading Client

The *eyeReading Client* runs in a browser. It is supported by all modern browsers. The eyeReading plug-in written with jQuery, extract information about the text displayed in the browser or eyeBook [9] and treats with the gaze similar to mouseover javascript event. It instantly sends the extracted information to the *eyeReading Server* also through WebSocket. This information is about the properties of the word has been read. For instance, index, line number, the word's text, height, width, the position on the document and so on.

C. eyeReading Server

The *eyeReading Server* uses data provided by *eyeReading Client* to detect classic eye movement features like fixations, saccades and blinks and the classify saccades such as forward reads, regressions, scanning saccades and sweep returns.

III. Usecases

We have conducted few studies with eyeReading framework in education, particularly in physics [5] [6]. We have also employed mobile eye tracking systems in reading research especially analysis of layout [7] [8]. In this section, we briefly review these studies.

A. Analysis of Text Layout

We employed wearable eye tracking systems to achieve higher flexibility and mobility. In brief, the mobile eye trackers have an embedded camera to record scenes during eye tracking experiment. There is a need to retrieve the corresponded document from the video records of reading trials. Hence, the gaze data mapped to the retrieved document using Locally Likely Arrangement Hashing (LLAH) [10]. Later, eyeReading extracts intended features for investigations.

The first study [7] was to analyze the text layout with considering the Golden Ratio in line spacing. We found that line spacing parameter of text layout, significantly, influences the reading behavior.

We conducted similar studies with the focus on scientific papers [8]. We explored the differences between two widely used layouts in computer community.ACM proceeding and Springer Lecture Notes in Computer Science (LNCS) have been selected in our study. ACM proceeding format represents double-column style while LNCS is selected for single-column representation. We found participants are more fluent and faster in double-column layout as they review the paper.

B. Education

The eyeReading framework is used in education research [5] [6]. Particularly, we studied representational competence in physics. Students' competence in using different representational formats and transfer one in another is called Representational Competence. We compared students with different level of expertise namely experts, intermediates and novices. Figure 2. shows the heatmaps of experts and novices on the accumulates question page with three representation: diagram, table and vector. We found experts proceeded more systematically and stuck to one single representation during problem-solving. Especially, when they were confronted with a hard question, they focused on the representation they are most familiar with and solved it correctly. Concerning specific representations, we found that novices used the diagram less often than experts and preferred the numerical representation (table).

IV. Future Work and Discussion

The big advantage of eyeReading framework is to handle streams of data. It makes it appropriate to make a monitoring system in smart classrooms. Suppose, the lecturer is able to monitor the quality of reading of all students. He would access to their reading history and could have statistics about the students and the class as well. At the time of writing this article, we are developing a system to assess the quality of the reading. For example, instant reading/skimming ratio detection, the approximation of user's comprehension, estimation of user's interest about the document. We would also employ Gaze Gesture to interact explicitly with the text. For instance, a user may would prefer to watch a video about a keyword in the text for better understanding. She/He could make a predefined gaze gesture to pop-up the related media to watch instantly.

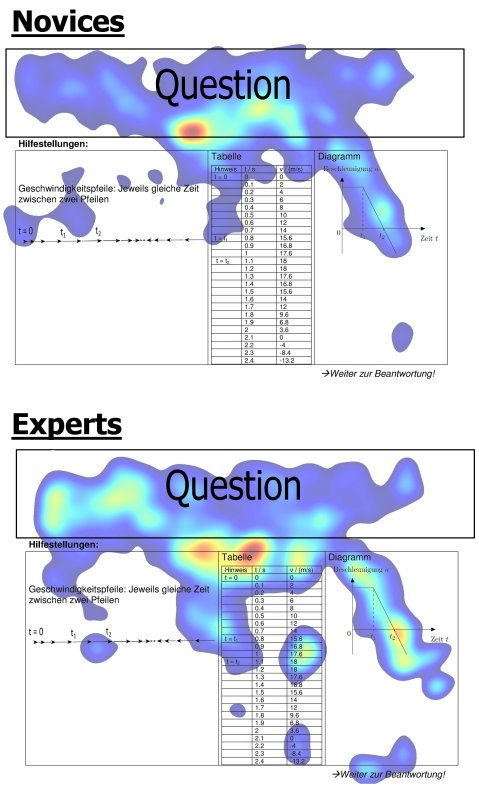

Fig. 2: heatmaps of experts and novices on the accumulates question page with three representation: diagram, table and vector.

References

[1] Huey, E., *The Psychology and Pedagogy of Reading*, New York: Macmillan, 1908.

[2] Rayner, K., Pollatsek, A., Ashby, J., and Clifto Jr, C *Psychology of Reading*, Psychology Press, NY, 2012.

[3] Beymer, D. and Russell, D. *WebGazeAnalyzer: A System for Capturing and Analyzing Web Reading Behavior Using Eye Gaze.* , In CHI 2005 Extended Abstracts. 1913- 1916.

[4] Biedert, R. and Buscher, G. and Schwartz, S. and Hees, J. and Dengel, A. *Text 2.0*, CHI 2010, Atlanta, Georgia, USA, 2010.

[5] Mozaffari Chanijani, S., S. and Klein, P. and Al-Naser, M. and Bukhari, S. and Kuh, J. and Dengel, A. *A Study on Representational Competence in Physics Using Mobile Eye Tracking Systems*, MobileHCI 16 Adjunct, September 06-09, 2016, Florence, Italy.

[6] Mozaffari Chanijani, S., S. and Klein, P. and Bukhari, S. and Kuh, J. and Dengel, A. *Entropy Based Transition Analysis of Eye Movement on Physics Representational Competence*, Ubicomp/ISWC16 Adjunct , September 12-16, 2016, Heidelberg, Germany.

[7] Mozaffari Chanijani, S., S. and Bukhari, S. and Dengel, A. *Analysis of text layout quality using wearable eye trackers*, IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2015.

[8] Mozaffari Chanijani, S., S. and Al-Naser, M. and Bukhari, S. and Borth, D. and Allen, S. and Dengel, A. *An Eye Movement Study on Scientific Papers Using Wearable Eye Tracking Technology*, The Ninth International Conference on Mobile Computing and Ubiquitous Networking, Kaiserslautern, Germany, 2016.

[9] Biedert, R. and Buscher, G. and Dengel, A. *The eyebookusing eye tracking to enhance the reading experience*, Springer, Informatik-Spektrum, 2010.

[10] Nakai, T and Kise, K. and Iwamura, M. *Use of Affine Invariants in Locally Likely Arrangement Hashing for Camera-Based Document Image Retrieval*, Proceeding 7th International Workshop DAS, Nelson, New Zealand, February 13-15, 2006.