An Eye Movement Study on Scientific Papers Using Wearable Eye Tracking Technology

Seyyed Saleh Mozaffari Chanijani∗†, Mohammad Al-Naser∗†, Syed Saqib Bukhari∗†, Damian Borth∗†, Shanley E. M. Allen†, and Andreas Dengel∗†

∗German Research Center for Artificial Intelligence, DFKI GmbH
†University of Kaiserslautern

Abstract—In this study, we started to investigate the impact of different layouts in scientific papers using eye tracking technology. At this stage, we limit our study to the comparison between layout formats inside the Computer Science Community. Association for Computing Machinery (ACM) proceeding as the double-column format and Springer Lecture Notes in Computer Science (LNCS) as the single-column format have been selected for investigation and will be presented in this paper. We employed a wearable eye tracker instead of a remote desk-mounted eye tracker. Due to their mobility and flexibility, this technology has been selected to simulate real-world environment while reading printed documents. Data acquired by a wearable head-mounted eye trackers is based on the gaze position with respect to the video recorded by the embedded camera. Hence, the coordinate of the gaze must be mapped to the corresponding document so that it enables us to investigate eye tracking data analysis techniques. In order to perform this task, we adopted a robust document retrieval technique called Locally Likely Arrangement Hashing (LLAH) to our data. Briefly, the scenario of the process is as follows: First, participants read the print-out scientific papers with the eye tracker. Then, gaze data maps to the corresponding original document in our retrieval database. Finally, the gaze analysis system extracts the intended information for statistical evaluation. Our findings show subjects are more fluent and faster in the double-column proceeding format as compared to the single-column.

I. INTRODUCTION

Human attention on document reading is one of the main topics in the eye tracking research over several years [1], [5], [16]. Remote desk-mounted eye tracker systems are widely used in the studies of reading behaviour [8] [13] [14] [2]. Bidert et al. [10] [9] proposed a gaze-based human-text interaction system called Text2.0 which uses the eye tracking technologies. It is a framework for gaze-based human-text interaction by using remote desk-mounted eye trackers. It offered various possibilities not only to evaluate reading behaviour [8] but also to enhance the reading experience [12]. Even though Text2.0 [10] was a breakthrough for human-text interaction, there is still a lot of room for improvement. For instance, subjects are forced to sit down in front of the remote eye tracker and perform time consuming calibration. They may need a chin rest in order to provide accurate gaze on the text which is quite far from real-world situations of reading. In addition to the lack of flexibility, the remote eye trackers are not suitable for hand-held paper reading analysis. Hence, the remote desk-mounted eye trackers are suitable to use in the lab but not for real-world scenarios of reading either on the screen or from the hand-held documents.

The remote in-lab eye trackers are beneficial in term of frequency, which is a very important factor for research on reading [1]. In contrast, wearable eye trackers have less frequency but are flexible, portable, market growing and easy to calibrate. Fortunately, with the advancement of technology, the frequency rate of new generations of wearable eye trackers is increasing. For instance, SensoMotoric Instruments (SMI) recently produced 120Hz frequency Eye Tracking Glasses (ETG) [1]. Although using wearable eye trackers has some advantages that have already been mentioned, it needs an extra step to map gaze from the scene recorded by the eye tracker's camera to the original document. This process is called gaze mapping. Interaction with text by using head-mounted eye tracker in pursuance of word recognition in scenes captured by the eye tracker’s camera, is proposed by Kobayashi et al. [19] which uses a document retrieval technique for the gaze mapping.

In this paper, we are focusing on using wearable eye tracking technology to have an exploratory study on analysing scientific publication layouts and their impact on reading. With respect to the potential of eye tracking technology, there is a big potential to investigate this technology in User Experience. An instance of this type of research with more focus on documents.

Fig. 1: SMI Eye Tracking Glasses (ETG) 2.6 with 60Hz frequency is used in our study.

is the analysis of different layouts. For example, Pohl et al. [11] has proposed a study for comparing the readability of force-directed, orthogonal, and hierarchical graph layouts using eye-tracking and task-oriented analysis. The main finding in this study for algorithm designers is that edge crossing pose a little problem in orthogonal graphs, also that the length of the edges and the number of bends make finding sub-graphs difficult. At the end, they found that the force-directed graph performed the best in all tasks. Eglin et al. [3] presented an experiment on eye movement that has been applied to document structure retrieval. The purpose of their study is to create a representation of structured document content layout through a simulation of the possible human inspired scan path. The researchers have focused on the salient areas which show that they are attractive for the reader. Lastly, they found there is not a unique ideal scan-path during document exploration, the readers interested can be multiple and can be attracted to different salient areas. In our previous study [15], we analysed layouts of documents with respect to the aesthetic parameter, the Golden Ratio, for line spacing. The importance of scientific papers is increasing day by day, since there, science findings are published in a form of articles. The layout of these publications is very important to guarantee the best way of reading, reviewing and implementing. Also, it is helpful to know which paper layout consumes more time to comprehend, and which part of the document attracts more attention. In this study, we explored the differences between two widely used layouts in the Computer Science community using the eye-tracking methodology. We selected ACM proceeding and Springer Lecture Notes in Computer Science (LNCS). ACM proceeding format represents double-column style while LNCS is selected for the single-column representation. The equivalent content articles [4] have been given to the participants as print-out hand-held copies, and asked to review the papers as they are a reviewer for a conference. The paper is organised as follows: In section 2 we will explain the methodology of measuring reading behaviour with a wearable eye tracker. In section 3 the experiment process is discussed. Our results are presented in section 4. Finally, there is a discussion about this work in section 5.

II. MEASURING READING USING A WEARABLE EYE TRACKER

We employed SMI ETG 2.6 (Figure 1) eye tracker for the experiment. In brief, the eye tracker has an embedded camera to record scenes during eye tracking process. There is a need to retrieve the corresponding document from the video records of reading trials. Then, the raw gaze data will be mapped to the retrieved document. Later, the gaze analysis system extracts intended features for investigations. In this section, we first explain the eye tracking process, then the document retrieval procedure will be discussed and finally, the low-level gaze analysis investigation in our research will be explained.
Fig. 3: The feature components of our study: The fixations are shown by blue circles. A regression is an implicit sign that the reader is having difficulty understanding the material. It is shown by the red gaze path. When processing the fixations to form forward reads, a forward read will be stopped when (1) a regression is encountered, (2) a forward saccade is too large and likely a forward skip through the text, or (3) the eye gaze moves to another line of text. The last called return sweep and indicated with orange colour in the figure. [1] [2]

A. Eye Tracking

The eye tracking is processed using images from two infra-red eye cameras and one scene camera. Each eye is illuminated by six infra-red lighting sources and the system tracks the changes of these six infra-red light reflections. In order to use this eye tracker, the wearer has to calibrate the system. The calibration process requires the wearer to look at one (or three, if necessary) point(s) in a real scene. The process does not take long for most of the users, however, it sometimes takes very long for other users, when they have difficulties in processing it once, thus they have to recalibrate it several times until accurate calibration is obtained.

B. Document Retrieval

We adopt an image based document retrieval method proposed in [17]. This method, called LLAH (Locally Likely Arrangement Hashing) is robust to perspective distortion of an image and scale-invariant. The method has been used by [18]. An overview of the document retrieval method is shown in Figure 2. Here is a brief description of the method: When a scene image is given from the camera, by a Gaussian kernel the image is blurred and adaptively threshold transformed into a binary image in order to detect the centroid of each word region. By changing the size of the Gaussian kernel, it can be adjusted to the optimal image blur for document retrieval. From the arrangements of the detected centroids, affine invariant feature vectors are calculated. The recognition process is done by matching the extracted features to the features previously stored in the database. A hashing technique is used for fast computing. By matching the features between the scene image and the retrieved database image, we also calculate the homography matrix between them. Based on this matrix, the gaze on the scene image can be mapped onto the retrieved document image.

C. Gaze Analysis

The eye tracking studies showed that eye motion advances in discrete chunks across the page during reading. A readers eyes will actually stop, or fixate, on a set of characters for about 250 ms. This fixation is followed by a saccade, an eye movement to the right, where the eyes will stop at the next fixation. A regression, or backwards eye movement in the text, is an implicit sign that the reader is having difficulty understanding the material. The gaze analyzer receives the gaze data from the document retrieval module. As a result, the gaze position of the currently read document has been obtained. Here, the parts of the text in which gazes are located have been mapped to the coordinates of the retrieved document. From these raw gazes, our analysis system first removes noise and then detects fixation points. Our gaze analyzer detects fixations in the gaze stream using a dispersion-based approach [6]. The analyzer next parses the fixation data into forward reads. A forward read is a grouping of consecutive fixations moving forward through the text with typical reading saccades (see Figure 3) [2]. When processing the fixations to form forward reads, a forward read will be stopped when (1) a regression is encountered, (2) a forward saccade is too large and likely a forward skip through the text, or (3) the eye gaze moves to another line of text. Intuitively, forward reads are designed to capture when and where subjects are actively reading the text in detail [2] [1]. Hence, forward reading gives us the possibility for an instantaneous measure of reading. The gaze analyzer assigns an index for each fixation \( i[f_0...f_{m-1}] \) for forward reading \( j[fr_0...fr_{n-1}] \), we measured the time as the subjects eyes are actively reading forward through the text.

\[
T_j = \sum_{i=0}^{m-1} t_i
\]
Adaptive Strength Geo-Replication Strategy

Arnaldo Anz Sigua1, Amrita Benitez2
1 Technische Universität Kaiserslautern, Germany
2 Technische Universität Kaiserslautern, Germany

Abstract
The amount of data being processed in keeps growing at an enormous rate so that fail replication may start being important. The application of replication is not used to increase data availability in the presence of data failures and to reduce latency by accessing the data locally located, if possible. This means that exploring data only in some of the data may become more critical in order to reduce the cost of keeping the data (readily) available while maintaining high availability and data consistency. In this paper, we consider the latencies and cost that arise when keeping too many copies of the same database and we evaluate the cost and the benefits of using only some of these copies. These results are supported by our experimental evaluation in a real-world environment and in a simulated environment.

Keywords: Adaptive Replication, Geo-replication, Large-Scale Distributed Database Replication, Latency, Availability

1 Introduction
The amount of data being processed in keeps growing at an enormous rate [12, 13, 14]. Some of the areas where the amount of stored data already reach TBs and even PBs are data mining, particle physics, climate modeling, high energy physics and astrophysics, to name a few. All these fields use data which needs to be shared and analyzed [8, 13, 14]. Data sharing is a common problem in scientific communities and has a huge impact on scientific research. In order to access the data, these communities use distributed information systems (DINIS) [10, 11, 12, 14] which are complex distributed systems that store, retrieve and analyze data. The cost of storing these large amounts of data is huge, and the cost of accessing the data is also high. This means that the cost of accessing the data is a critical issue in the scientific community. In this paper, we consider the cost of accessing the data and we evaluate the cost and the benefits of using only some of these copies.

Fig. 4: The scientific layouts, selected for the experiment. The right side is Condition A: Springer Lecture Notes in Computer Science (LNCS) and the left side is Condition B: Proceedings in Association for Computing Machinery (ACM)

and similarly the distance covered by the forward read is measured as

$$D_j = \sum_{i=0}^{m-1} d_i$$  \hspace{1cm} (2)

The instantaneous measure of reads speed for each forward read $j$ is then calculated as $S_j = D_j/T_j$.

In addition we calculated the fixation duration in each forward read $j$.

$$F_j = \sum_{i=0}^{m-1} f_i$$  \hspace{1cm} (3)

Where $n$ is total number of forward reading and $m$ is the number of fixations in each forward reading $j$.

The regressions are also detected by the gaze analyzer for the further analysis. Figure 3 shows the features [1] extracted by our gaze analyzer.

III. Experiment
In order to simulate the real-world environment while reading, we used a wearable tracking system to give more flexibility, i.e. head movement, to the readers. We used two scientific layouts in the Computer Science community in our experiment representing single-column and double-column layout formats as is shown in Figure 4:

Condition A: Springer Lecture Notes in Computer Science (LNCS) representing a single-column format.

Condition B: Proceedings in Association for Computing Machinery (ACM) representing a double-column format.

Totally 18 participants, postgraduate level in Computer Science community, attended for the experiment. 9 read Condition A and 9 read Condition B. The contents of two formats are equivalent and selected from Proceedings of the First Workshop on Principles and Practice of Consistency for Distributed Data [4]. There are two reasons for selecting this paper. Firstly, this paper was recently published Computer Science article. Secondly, even though this is an article in the Computer Science domain it is rather far from the participants expertise, which is Artificial Intelligence, so that they are
familiar with the subject but not an expert in it. Therefore, it takes more attention to read instead of skimming the paper. Finally, the important factor in this type of research is the fluency of the paper itself. Hence, this article is chosen due to the main author who is a native English writer.

SMI ETG 2.6 is used for our study (Figure 1). The temporal resolution of the eye tracker is 60 Hz (binocular) and gaze position accuracy is 0.5 over all distances. The resolution of the scene camera is \(960 \times 720\).

They all have been asked to wear the eye tracker and review the paper. Before starting the reading, calibration is performed to acquire accurate gaze position.

**IV. RESULTS**

The gaze analyzer starts to proceed by detecting fixations and grouping them into forward reads which are the characteristic reading pattern of fixations and saccades in the eye gaze stream. Thus, for each, we can measure reading speed, saccade length, fixation duration. The regressions are then detected by the gaze analyzer. The original paper used in our study [4] consists of following parts which we divided them into four section-of-interests as follows:

- **S0**: \{title, authors and keywords. \}
- **S1**: \{abstract and introduction. \}
- **S2**: \{part 2 (Classification of Strategies), part 3 (Adaptive Strength Georeplication Strategies) and discussion. \}
- **S3**: \{acknowledgement and references \} [4]

Figure 5 shows the average of the aggregated fixation duration in forward reads in each section-of-interest. The error bar is standard error of the mean. The result shows, interestingly, subjects had more fixation duration during forward reading in the single-column format in all section. It is a sign that subjects are more fluent in reading Condition B. This finding is supported by the regression ratio as Figure 6 presents the proportion of regression to the number of lines in each section-of-interest. It is clear that the number of regressions are much higher in Condition A in contrast with Condition B. The forward reading distance with respect to Formula 2 has been calculated.

The forward reading length in Condition A is not surprisingly bigger than Condition B as showing by Figure 7. But the speed of forward reading in Condition B is more in contrast to the single-column format. The speed of forward reading is shown in Figure 8.

**V. DISCUSSION**

In this paper, we investigated the impacts of layouts in Computer Science papers by integrating the wearable eye tracker and LLAH document retrieval algorithm. We employed two well-known formats as this stage: ACM proceedings and Springer LNCS in the experiment. We found the interesting outcomes, showing subjects are more fluent and faster in double-column layout as they review the paper. Using wearable eye trackers in such research provides the opportunity for more thorough investigations as it simulates the real-world scenario of reading. It promises the flexibility during reading and prevents stress in-lab researches. Due to the frequency rate of the employed eye tracker in this study, fixations were our main concern to evaluate results. But it can not be a constraint for deeper investigations in the future as the technology of eye tracking is developing at a very fast speed. Indeed, this research is intended to be extended in the future. Further research in other layout features, i.e. salient position on the page, transitional analysis on scan-path and
The speed of forward reading in the double-column format is higher than the single-column format. The error bar is standard error of the mean.

Fig. 7: The forward reading length in the section-of-interests. The forward reading length in Condition A is not surprisingly bigger than Condition B. This work was funded by the BMBF (German Federal Ministry of Education and Research) for the project AICASys (ID: 16SV7184).

ACKNOWLEDGEMENT

This work was funded by the BMBF (German Federal Ministry of Education and Research) for the project AICASys (ID: 16SV7184).

REFERENCES