User's Interest Detection Through Eye Tracking for Related Documents Retrieval

Jakub Ševcech
Faculty of Informatics and Information Technologies,
Slovak University of Technology
Ilkovičova 2, 842 16 Bratislava, Slovakia
jakub.sevcech@stuba.sk

Mária Bieliková
Faculty of Informatics and Information Technologies,
Slovak University of Technology
Ilkovičova 2, 842 16 Bratislava, Slovakia
maria.bielikova@stuba.sk

***Abstract*— In multiple studies, various implicit and explicit feedback types were used to quantify document quality and to identify user's interest in document content on both document and sub-document level. In this study we use eye movement data retrieved while reading a document to identify sections, user is most interested in. We describe multiple patterns in eye movement of users reading documents that can be used to identify important document fragments. The identified patterns are used to determine the user's interest in text fragments and to compare effectiveness of extracted document fragments and user created in-text highlights in the task of related document retrieval. Identified eye movement patterns allowed us to extract more important document sections compared to manual annotations with greater precision and consistency. Both types of compared interest indicators achieved similar performance in related document retrieval with the single most important difference: gaze based important section identification do not require active participation of the reader.**

Keywords—eye tracking; eye movement patterns; document retrieval; interest indicators; related document retrieval

I. INTRODUCTION

In the past 30 years, extensive work have been done in the field of eye tracking research. At first, eye tracker devices were inaccurate, expensive, hard to use and they were used only as a research equipment. However with the recent advances in technology, eye trackers are becoming more affordable and in the near future it is possible that every home computer or mobile device will be equipped with such a device. Eye movement tracking may become another mean of interaction of ordinary people with computers. They can benefit by using such device as a mean of conscious control of computer but also as a mean for computer programs to gain feedback and to adapt to user's behaviour.

Many previous works used various types of user feedback such as mouse movement, clickthrough data or text selection and bookmarking [4] as a source of information for user modelling, recommendation or to quantify user's interest in the document content. In [6] they used various reading indicators such as scrolling and visit duration in connection with document edit indicators to simulate deterioration of the document by extensive usage – similarly to the deterioration of printed documents. They displayed the amount of activity connected to specific sections of the document in form of a histogram in the background of the browser's scroll bar. Using this scroll bar, users were able to identify sections of the document which were most frequently read and edited and they are thus most interested in. In [8] various types of user activity tracks such as in-text highlights or comments were used to create personalized summarization of the document for individual users. They improved the quality of summarizations created from document content by identifying the most important sections of the document from the point of view of multiple users and from the point of view of individual summarisation users. Another use of social annotations as interest indicators was presented in [12] where they used the amount and type of user activity on document to rank documents by their quality or interest for their readers. This work represents an example of ranking document quality not only using the content of the document and structure of interconnections between documents (eg. hyperlinks or citations) but also indicators of reader's activity while reading the document. The authors use information generated by the document reader instead of the information generated by the document author or other authors to rank the document quality.

In this work we use eye movement patterns as a form of implicit feedback for identification of document fragments that are most important from the reader's point of view. The identification of these fragments can provide valuable input for methods such as modelling of long and short term user interests, for modelling user's knowledge, document summarization, ranking or for proactive support while reading or studying such as document recommendation. Automatic identification of important document fragments would be especially important in domains such as digital libraries, where users are not only quickly scrolling through the content but the study of the documents require increased attention.

In the context of this paper, we consider importance of the document fragment to be a characteristic of the text that attracts attention of the reader disregarding the specific reason of the user being interested in the text. It may be a possible direction for future research to identify whether the text is important for the reader because it is interesting, the text describes key concept of the document, or it is simply incomprehensible and the reader needs more time to understand the text correctly.

II. RELATED DOCUMENTS RETRIEVAL

Retrieving related documents is one of multiple forms of proactive support taking place while a user is reading a document or right after she finishes reading the document and searches for another. We use the task of retrieving related documents to compare properties of two types of user interest indicators: manual annotations of the document content and gaze data used for identification of document sections the user is most interested in.

The most common approach to search for related documents is to use content of the document to create or expand queries issued in commonly used search engines. Multiple metrics are used to extract words fit to form the query. The most simple and most common methods use in-document word frequency [13]: they select the most frequent words as most important words or phrases in the document, the least frequent words as words that are the best descriptors of the topic or in combination with information about word frequency in document corpus using metrics such as tf-idf. Many others ATR algorithms are used as well to extract words with specific properties.

These methods however employ only information gathered from the document content regardless of user's interest in the document or its specific sections. In contrast to these approaches, in our previous work [11] we proposed a method for query construction from document content and attached annotations, where user created annotations in form of in-text highlights, comments and notes were used as user's interest indicators. The proposed method significantly outperformed commonly used method for query construction for related document retrieval based on tf-idf using only document content as well as using document content enriched by annotations. Using the method based on transformation of document content to graph and spreading activation introduced by various types of manual annotations in the graph, we showed that user created annotations can be used to increase precision of related document retrieval.

In this paper we compare the potential for related document retrieval of annotations user creates while reading the document and most important sections extracted using eye movement data collected while reading the document. We use the method described in [11] with manual annotations used as interest indicators as a baseline method for evaluation of properties of important document sections extracted using gaze data in the role of interest indicators. We compare both types of interest indicators, one with another and with baseline method using only document content in related document retrieval. When comparing both interest indicators we focus on qualitative properties important for document retrieval.

III. EYE MOVEMENT AS INTEREST INDICATOR

Over the last 30 years many studies concerning eye movement studied gaze data in different activities [10]. They provide rather extensive expertise about the movement of an eye when reading text. The eye is not moving fluently through the text when reading and it is constantly skipping between rather stable positions. The stable eye positions are known as fixations and the step between two fixations is called saccade. The number and duration of fixations is often used as a metric to identify important sections of documents in user experience testing [9]. Saccades and fixations however can also be used to find patterns in user's reading and can serve as an input for example for reading and skipping (scanning) detection [3] or for word relevance detection [7].

In [3] authors identified six reading patterns (eg. read forward, skim forward, short regression) and used them for reading and skimming detection. The proposed algorithm based on accumulation of scores introduced by these patterns was able to classify text as read carefully or skimmed using eye fixations and saccades properties. Another work [6] proposes multiple similar metrics for relevant word identification. They provided volunteers with short texts and they were asked simple question, for which they had to find answer in provided texts. They identified three promising metrics for identification of relevant words: number of fixations, number of first-pass fixations, and the total viewing time.

An application of patterns in eye movement during reading is presented in [2] where gaze is used as user attention feedback on subdocument level for query expansion. Another application of gaze data is in [5] where authors used data from eye tracker to model user acquiring information while reading document and to model user's domain knowledge without the need to process the document content. They showed that eye movement data can be a valuable source of information for user modelling and user knowledge level detection.

In our work we move from the identification of important words in independent phrases [7] to identification of most important sections in the whole document using eye movement collected while reading the document. We study eye movement patterns while reading longer documents as they reflects different reading styles compared to reading short blocks of text. We collected data from volunteers reading documents and extracted patterns in the eye movement while reading. We used identified patterns to extract the most important sections in the texts and to compare their properties as interest indicators with manual annotations in the task of related document retrieval.

IV. READING DATA COLLECTION

To collect reading data, we performed a user study with 6 participants, where they were asked to perform several successive tasks:

1. To read multiple documents.
2. To create highlights in the document as baseline interest indicators.
3. To respond to a short interview concerning their highlighting and annotating habits.
4. To select relevant documents from provided lists of related documents.

In the first step, participants were asked to read supplied documents while we captured their eye movement using Tobii X2-30 eye tracker. Before reading the documents we told every participant that we will ask several general questions about the content of the document after they finish the reading. We asked these questions to support participants in close reading and to motivate them in trying to understand the document content as we wanted to simulate the situation where the user wants to correctly understand the document content.

After our users finished the reading, we asked them to highlight sections of the document in a way they would do it if they could highlight the text in the document while reading it – we allowed them to reread the document again. We separated the step of studying the document and text highlighting step to ensure that the gaze data won't be affected by the process of users highlighting the text.

In the next step we interviewed the participants whether they annotate or highlight electronic or printed documents. What tools they are using for this task and what purpose they are using annotations for.

After highlighting the text in the documents and answering to these questions, we created two queries for related documents (one using user created annotations and one without them as a baseline method) using method for related document retrieval described in further detail in [11]. We used these queries to search for related documents using Google Search interface and we presented them to participants. Participants were asked to select relevant documents to the studied document and to select better from provided lists. Lists were presented in random order and search results inside individual lists were also randomly shuffled to eliminate the influence of result order in perceived result relevance.

Later, when we processed the data about eye movement (about a month later) we repeated the last step of the experiment with one additional list of related documents. We used gaze data to extract most important sections of the document from user's point of view and we used them to create query for related documents using the same method as in previous step. We compared relevance of related documents retrieved using different input data and the consistency of participants between two last steps of the experiment as they were separated by approximately one month.

In the performed experiment, we collected data from 6 participants, where everyone was asked to read and annotate two documents. The whole study was designed to not exceed 90 minute duration. All participants of the study were doctoral students in the field of computer science. We provided every participant with two documents[1, 2] published in Communications of the ACM journal. Both documents are journal articles written in rather popular language but concerning specific problems from the domain of computer science targeted for the community of scientists and computer science students. These articles represented text with a novel content for every participant, close but not directly from the domain of their expertise. We selected these articles with the aim to study patterns in reading of novel, specialized text that requires higher level of attention and where readers can find their own sections they are most interested in. To motivate participants in their effort and to increase their attention in the text, we told every participant that we will ask several questions to assess their document understanding after they finish the reading.

V. Evaluation

In the evaluation of collected data we focused on exploring patterns in user behaviour while reading the document and on using them to extract the most important sections in the document. As we move to interesting section extraction from longer texts, we focus not on study of individual fixations and saccades, but on patterns on phrase and document level.

In the second part of the evaluation, we used these patterns to extract most important sections in the text and to search for related document. In the related document retrieval process we used important sections identified using user generated annotations and important document sections identified using gaze data. For the document retrieval using important sections of the document, we used method described in [11]. We compared properties of annotations and text sections extracted using gaze data and we compared their utility in related document retrieval.

A. User's Reading Habits

Similarly to results described in [3] we were able to identify several patterns in reading habits of different people. By comparing gaze data to annotations attached to the document by single user and to annotations created by multiple users, we were able to identify patterns in user's eye movement that can be used to identify most important sections in the text. In contrast to patterns described in [3] we were interested in identification of longer sections, not only selected words, but whole phrases and document sections.

Based on comparison of the data from all participants, we identified following reading patterns:

- *Long fixations and short saccades* (Fig. 1) indicate increased effort in reading the document section caused either by problems with understanding the content or by in some way interesting content.
- *Short fixations and long saccades* (Fig. 2) indicate fast, superficial reading and can be interpreted as negative signal for content importance.
- *Repeated short fixations* on the same words are similar to long fixations, but they represent greater risk, to be caused by error in the text or hard-to-read fragment of the document.
- *Return back a few lines* or to previous paragraph has rather vague interpretation and without external information (eg. document content), we cannot automatically decide whether the text is hard to understand or it is important or the reader was interrupted by external stimuli.
- *Return back to distant place* in the document can be interpreted as important or interesting section in the document. In addition we can interpret the jump as a

[1] Cooper, S. and Sahami, M. 2013. Reflections on Stanford's MOOCs. In *Communications of the ACM*, ACM, 56(2), 28-30.

[2] Anthes, G. 2013. Deep learning comes of age. In *Communications of the ACM*, ACM, 56(6), 13-15.

connection between the text reader returned from and the text she jumped to.

- *Skip paragraph* is interpreted as uninteresting, already known or unimportant document section.
- The interpretation of *small number of fixations concentrated in clusters* across the whole document depends on whether reader have already read the document and she is simply rereading the most important sections of the document or she have never read the document before and she is making a quick scan over the document and we can make no assumptions about gaze interpretation.

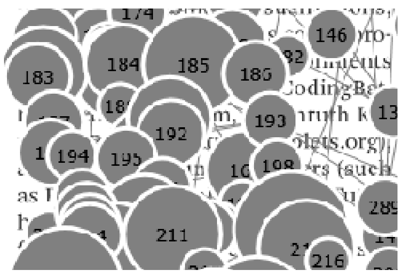

Fig. 1. Long fixations and short saccades indicate important text fragments. The size and position of circles in the image represents length and position of the fixation.

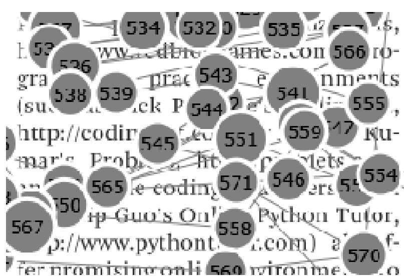

Fig. 2. Short fixations and long saccades indicate unimportant content. Disatnce between circles representing fixations represents length of the saccade.

B. *Related Document Retrieval*

We used identified patterns to extract the most important sections from documents used in the first part of the experiment. Using extracted sections of the documents we created queries for related document retrieval using the method described in [11] and we extracted related documents by issuing queries created using document content and attached highlights or important sections into Google Search interface.

After extracting related documents we asked participants to evaluate their relevance. They evaluated the relevance of retrieved documents in two steps. In the first one, following immediately after document reading and annotation, the participants were asked to compare list of results retrieved using annotations and without using annotations in query construction process. The second step took place after we processed gaze data using already mentioned patterns in user's eye movement. In this step, participants were asked to compare three lists: the two from previous step and one additional extracted using the same method for query construction and important sections identified using gaze data as user interest indicators.

We used these data in a qualitative study to determine properties of document sections extracted using gaze data and user created in-text highlights as interest indicators in the task of related document retrieval.

Participants used annotations in different ways for different purposes, which reflects in the annotations properties. They reported that they use highlights as means to:

- summarize text: great number of long highlights,
- mark important places for later use: short, scattered highlights where the important content follows after the position of the highlight and
- highlight key concepts in the text: small number of short highlights.

The basic properties of created highlights and important sections identified using gaze data are summarized in table 1. The data were collected from 6 participants, each reading and annotating two documents. On average, participants highlighted 23.75 text fragments per document with 3.68 words per highlight. In contrast, text segments extracted using gaze data were much longer (39.08 words per segment) and in smaller number (4.33 per document). In-text highlights point very precisely to the important section, but interpretation of their span and meaning can be complicated due to user's individual highlighting style and unknown purpose of highlights. Users were using in-text highlights as means to highlight important concepts (individual words), ideas and phrases or entire sections. They use various types of annotations and highlights as a means to summarize the document, to mark important or questionable sections for later rereading or as a placeholder to speed up the navigation in the document.

TABLE I. COUNT AND AVERAGE SIZE OF HIGHLIGHTS AND IDENTIFIED IMPORTANT SECTIONS

	Highlight	Important section
Total number	285	52
Avg. count per document	23.75	4.33
Avg. size in words	3.68	39.08

By contrast gaze data identify larger interesting areas (on average) with similar length across different readers and reading styles. As described in previous section their interpretation depends on interpretation of eye movement patterns. The span of important sections identified using gaze data can be compared to highlights used for summarizing the content of the document. Gaze data however identifies smaller number of important sections which results in greater precision in identification of most important sections. Unlike manual annotations, however, important text section identification is susceptible to errors caused by misinterpretation of gaze data where we are prone to label text segment as important when reader was only interrupted in reading or she has problems with comprehension of the text.

When retrieving related documents using important sections determined by gaze data, we were able to obtain comparable number of relevant results as when manual annotations were used in related document retrieval. We obtained 55.38% and 54.29% of relevant results in presented lists for method using in-text highlights and method using identified important sections respectively. At the same time both methods using annotations or gaze data as interest indicators outperformed the method using only document content for query construction (46.38% of relevant results) confirming the results obtained in [11]. The main difference between the two sources of interest indicators (manual annotations and gaze data) is, that to acquire feedback in form of gaze data, there is not required active participation of the reader and we can extract even greater number of interest indicators from the eye movement.

VI. Conclusions

In the described study we examined the patterns in eye movement while reading a document and we used them in important document segment identification. Identified eye movement patterns can be used for important document fragment identification in domains where readers are not only scanning through the provided content, but where increased attention is required such as in the domain of digital libraries or web-based learning. They can be used as heuristics for detection of document sections that are in some way important for the reader, eg. in an educational system we can identify important parts of study materials determined by a group of learners that can serve for recommendation during knowledge revision.

We compared properties of document sections extracted using eye movement patterns with manual annotations (highlights) users are commonly attaching to printed or digital documents while reading them and we compared their performance in related document retrieval. These heuristics can cover more sections of the document that are interesting for the user compared to manual annotations and they do not require direct input from the reader, but they are also more prone to produce false positives due to misinterpretation of gaze signals (eg. user interrupted while reading or hard-to-read document content). To achieve correct automatic interpretation of these heuristics, further experiments and possibly more external signals eg. document content or combination with other forms of implicit feedback would be necessary.

Utilizing gaze tracking technology for user activity analysis represents viable option for many other applications as this technology is becoming more affordable for researchers and for ordinary users too (we witness lowering price and increasing availability of low-end models). This opens various possibilities for future work based on considering implicit interaction on the Web, which is primary source of user interest [1]. Knowledge on reading patters enables an adaptation and recommendation on markedly detailed level, which can improve several information processing tasks especially exploratory navigation.

Acknowledgement

This work was partially supported by grants KEGA 009STU-4/2014, APVV-0208-10, and the paper was created with the support of the Ministry of Education, Science, Research and Sport of the Slovak Republic within the Research and Development Operational Programme for the project University Science Park of STU Bratislava, ITMS 26240220084, co-funded by the European Regional Development Fund.

References

[1] F. Alt, A.S. Shirazi, A. Schmidt, and J. Mennenöh. "Increasing the user's attention on the web: using implicit interaction based on gaze behavior to tailor content." In Proc. of the 7th Nordic Conf. on Human-Computer Interaction: Making Sense Through Design - NordiCHI '12. ACM, 2012, pp.544-553.

[2] G. Buscher, A. Dengel and L. van Elst, "Query expansion using gaze-based feedback on the subdocument level," in Proc. of the 31st Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval - SIGIR '08, ACM, 2008, pp. 387–394.

[3] G. Buscher, A. Dengel and L. van Elst, "Eye movements as implicit relevance feedback," in Proc. of the 26th Annual CHI Conf. Extended Abstracts on Human Factors in Computing Systems - CHI '08, ACM, 2008, pp. 2991–2996.

[4] M. Claypool, P. Le, M. Wased and D. Brown, "Implicit interest indicators," in Proc. of the 6th int.conf. on Intelligent user interfaces - IUI '01, ACM, 2001, pp. 33–40.

[5] M. J. Cole, J. Gwizdka, C. Liu, N. J. Belkin and X. Zhang, "Inferring user knowledge level from eye movement patterns," in Information Processing & Management, 49(5), 2013, pp. 1075–1091.

[6] W. C. Hill, J. D. Hollan, D. Wroblewski and T. McCandless, "Edit wear and read wear," in Proc. of the SIGCHI conf. on Human factors in computing systems - CHI '92, ACM, 1992, pp. 3–9.

[7] T. D. Loboda, P. Brusilovsky and J. Brunstein, "Inferring word relevance from eye-movements of readers," in Proc. of the 15th Int. Conf. on Intelligent User Interfaces - IUI '11, ACM, 2011, pp. 175–184.

[8] R. Móro and M. Bieliková, "Personalized text summarization based on important terms identification," in Proc. of Database and Expert Systems Applications - DEXA, IEEE, 2012, pp. 131–135.

[9] R. Móro, J. Daráž and M. Bieliková, "Visualization of Gaze Tracking Data for UX Testing on the Web," in Proc. of Data Visualization Workshop of Hypertext - DataWiz 2014, CEUR-WS, vol. 1210.

[10] K. Rayner, "Eye movements in reading and information processing: 20 years of research," in Psychological bulletin, 124(3), 1998, pp. 372–422.

[11] J. Ševcech, R. Móro, M. Holub and M. Bieliková, "User annotations as a context for related document search on the web and digital libraries," in Informatica, vol. 38, no. 1, 2014, pp. 21–30.

[12] Y. Yanbe, A. Jatowt, S. Nakamura and K. Tanaka, "Can social bookmarking enhance search in the web?" in Proc. of the 2007 conference on Digital libraries - JCDL '07, ACM, 2007, pp. 107–116.

[13] Y. Yang, N. Bansal, W. Dakka, P. Ipeirotis, N. Koudas and D. Papadias, "Query by document," in Proc. of the 2nd ACM Int. Conf. on Web Search and Data Mining - WSDM '09, ACM, 2009, pp. 34–43.