EyeBookmark: Assisting Recovery from Interruption during Reading

Jaemin Jo[1] **Bohyoung Kim**[2] **Jinwook Seo**[1]

[1]Human-Computer Interaction Laboratory, Department of Computer Science and Engineering, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-744, Korea

[2]Department of Radiology, Seoul National University Bundang Hospital, 82 Gumi-ro, 173 Beon-gil, Budang-gu, Seongnam-si, Gyeonggi-do 463-707, Korea

jmjo@hcil.snu.ac.kr, bhkim@snubh.org, jseo@snu.ac.kr

ABSTRACT

In this paper, we present gaze-based bookmarking, EyeBookmark, to mitigate the deleterious effect of interruption during reading. The key idea of EyeBookmark is to provide a visual cue to help people decide where to resume reading. We design four highlighting methods and conduct a controlled user study with a proof-of-concept design to verify the usefulness of EyeBookmark. The user study demonstrates not only that participants preferred our highlighting methods but also that such highlighting methods significantly reduced the time taken to resume reading after interruption regardless of the difficulty of text.

Author Keywords

Interruption; Eye-tracking; Reading.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION

Reading is an everyday activity to gather and extract information from text. Reading requires continuous visual attention. However, during reading, we are often distracted by various interrupting sources such as smartphones' alarm or engaging conversation with other people. Such unexpected distractions lead us to suspend reading until the distractions are dealt with. When we return to the text, we cannot resume reading immediately; for example, it takes a while to restore the context of the text and scan where to initiate rereading. As a result, interruption during reading induces resumption lag—the time taken to restart reading after interruption—and finally hinders reading performance.

As a remedy for such overheads of reading resumption, we present gaze-based bookmarking, or EyeBookmark, which helps people to restart reading promptly after interruption. The key idea of EyeBookmark is to highlight the last reading position in several ways when people attempt to resume reading. We design four highlighting methods (Figure 1) and conduct a controlled user study to verify the usefulness of EyeBookmark.

For ease of explanation, we adopt two terms, pre- and post-interruption periods, to represent periods before and after an interruption, respectively. In a similar way, we divide text into pre- and post-interruption texts which refer to the texts located before and after the word that one was reading when the interruption occurred.

RELATED WORK

Interruption has been extensively investigated not only in HCI but also in other fields such as psychology. We review previous research in following three thematic areas.

Theoretical Bases Latorella [5] proposed the interruption management stage model (IMSM) to explicate how interruption affects performance of tasks. In that work, Latorella emphasized the notion of *interruption resilient interfaces* which are interfaces that aid task resumption after interruption. EyeBookmark achieves interruption resiliency by providing a visual cue for reading. McFarlane and Latorella [6] suggested prospective UI support examples for each phase of interruption: before, during, and after interruption. In that sense, EyeBookmark is a practical example of the after interruption UI support.

Spatial Memory during Reading EyeBookmark complements the use of spatial memory for recovery from interruption. Inhoff and Weger [4] found that memory for word location during reading is not precise, which implies there may be cost for searching the last read position after interruption. Cane et al. [2] suggested highlighting the last seen word for efficient recovery from interruption. Also, they verified such highlighting reduces the reading time of the pre-interruption text in the post-interruption period. However, such reading time can be inappropriate to represent the exact resumption lag because people can seek the word within the post-interruption text. Therefore, we manually inspected fixation patterns to obtain the accurate resumption lag. Also, in addition to marking only one word, we design and compare four different highlighting methods.

CHI 2015, April 18–23, 2015, Seoul, Republic of Korea.

ACM 978-1-4503-3145-6/15/04…$15.00.

http://dx.doi.org/10.1145/2702123.2702340

Reading and Eye-tracking in HCI Exploiting eye-tracking technology, several HCI studies clarified people's reading behavior and attempted to improve reading performance. Buscher et al. [1] modeled preferred reading regions using normal distributions. Using this model, Sharmin et al. [8] proposed gaze-based auto-scrolling when reading text. Text 2.0 [9] assisted reading by superimposing additional information on text depending on the context of reading. Duggan and Payne [3] presented design guidelines based on their findings on people's reading behavior—*satisficing*. A patent (US7429108) presented methods to provide a reading place-marker using gaze-tracking but has not been evaluated yet through controlled experiments.

EYEBOOKMARK

As mentioned above, the main idea of EyeBookmark is to highlight the last reading position. To achieve this, an eye tracker keeps track of gazes when reading text. Once an interruption occurs, the position of the last fixation is identified. When getting back into reading after interruption, the last reading position is highlighted on text to present a restarting point. We believe such highlighting reduces unnecessary visual search through the text and assists the continuation of reading.

Highlighting Methods We design four highlighting methods as shown in Figure 1. *Point* method is the simplest but intuitive. It shows the position of the last fixation using a circle-shaped marker. It attracts attention to a resuming point but hardly conveys the context of the last reading.

Sentence method and *previous sentence* method indicate the context of the last reading in addition to the last fixation position. The *sentence* method highlights the sentence that they did not finish due to the interruption, whereas the previous sentence method marks the previous sentence of the unfinished sentence. We observed that people usually search for the beginning of the interrupted sentence to reread the whole sentence when getting back into reading after interruption. The sentence and previous sentence methods can reduce such visual search time.

In the sentence and previous sentence methods, the area of the highlighted region varies depending on the length of a sentence. If the sentence is too long or short, those methods can accentuate excessive or insufficient context for recovery from interruption. To mitigate this problem, the *block* method highlights a fixed number of lines including and preceding the last fixated word. The major strength of the block method is to provide a consistent amount of visual cue regardless of the length of a sentence.

PILOT STUDY

We conducted an in-laboratory pilot study to confirm the efficacy of EyeBookmark and compared the highlighting methods in terms of resumption lag.

Participants Nine volunteers (1 female) from a university participated in the pilot study, aged 20 to 37 years with

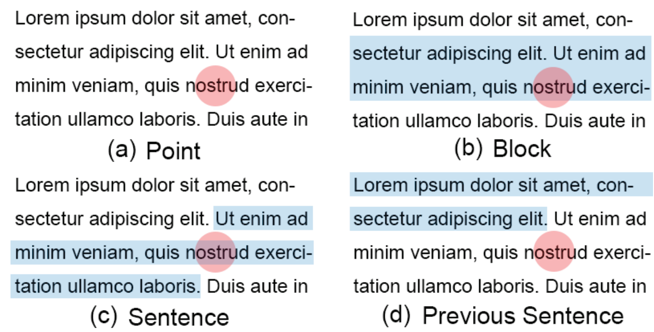

Figure 1. Four highlighting methods. For all methods, a red circle indicates the position of the last fixation before interruption. For methods (b), (c), and (d), blue markers additionally provide the context of the last reading.

normal or corrected-to-normal vision. All of them were native Korean speakers.

Apparatus A Tobii X60 binocular eye-tracker was used to collect gaze data during reading. The eye-tracker was installed in front of a 24-inch monitor with a screen resolution of 1920×1080. The eye-tracker and monitor were connected to a PC that ran a dedicated program for the pilot study. Participants were seated 65cm away from the monitor and their heads were fixed using a chin rest. Also, we implemented an I-VT fixation filter [7] to determine which sentence should be marked when getting back into reading after interruption.

Design We used a within-subjects design with five highlighting conditions: no highlighting (as a baseline condition) and four aforementioned highlighting methods. Each participant read 10 paragraphs. When reading a paragraph, one interruption occurred when a participant saw a specified target word.

During the interruption, the paragraph was hidden and a 60-sec long documentary video clip was played. After the interruption, a part of the paragraph was highlighted depending on the highlighting condition. Because we used 10 paragraphs and 5 highlighting conditions, each participant experienced each highlighting condition twice. The order of paragraphs was randomized. The order of highlighting conditions was also randomized while avoiding consecutive use of one condition.

We prepared 10 Korean paragraphs about various topics (10.4±2.2 sentences and 149.2±10.5 words). The paragraphs were shown with 24px, double-spaced, Malgun Gothic font (one of the most popular Korean fonts) contained in a rounded box with a width of 1024px.

In each trial, one word from a paragraph was chosen as a target word that triggered interruption when seen. To control the deleterious effect of interruption, the target word was randomly selected under several criteria as in [2]: all words in the first and last sentences were excluded because they brought less disruptive interruption; the first and last words of a sentence were excluded to create only mid-

	Method	Reading time (s)	# of fixations	Mean fixation duration (ms)	Mean resumption lag (s)	Answer accuracy (%)	Preference (7-point)
Pilot Study	None (baseline)	22.3 (9.3)	128.0 (37.6)	170.1 (31.4)	2.4 (0.9)	93.3	
	Point	24.3 (10.6)	136.2 (42.9)	173.9 (25.1)	1.6 (0.7)	81.3	4.6 (1.0)
	Block	26.2 (14.1)	140.3 (55.4)	179.4 (38.2)	1.2 (0.6)	88.9	5.2 (1.1)
	Sentence	23.8 (11.6)	129.0 (40.8)	178.5 (39.5)	1.3 (0.7)	88.2	5.4 (1.1)
	Prev. Sentence	28.2 (14.4)	150.1 (53.4)	180.5 (35.9)	1.2 (1.0)	88.9	3.7 (1.9)
User Study	None (easy)	26.4 (7.9)	147.0 (36.3)	177.5 (15.9)	2.0 (0.8)	90.9	
	Block (easy)	25.7 (9.1)	142.4 (42.2)	178.4 (19.6)	1.2 (0.6)	92.0	5.4 (0.7)
	Sentence (easy)	26.6 (9.0)	143.5 (43.1)	184.4 (22.5)	1.1 (0.3)	87.5	6.1 (0.9)
	None (difficult)	42.2 (14.3)	214.5 (58.9)	194.0 (28.0)	2.8 (1.5)	72.7	
	Block (difficult)	38.5 (17.9)	188.5 (55.3)	196.5 (42.0)	1.4 (0.7)	69.2	5.4 (0.7)
	Sentence (difficult)	37.1 (9.7)	189.4 (39.6)	195.2 (29.1)	1.4 (0.8)	79.2	6.1 (0.9)

Table 1. Results of pilot study and user study. Block and sentence methods significantly reduce the mean resumption lag compared to the baseline (None) condition.

sentence interruption; and the words at the extremes of each line were excluded because the fixations on those words were often missed.

Procedure At the beginning, participants were briefly introduced to tasks and highlighting methods with graphical examples as shown in Figure 1. Then, a 9-point eye tracker calibration process was conducted. Participants were familiarized to tasks using a practice paragraph and one random highlighting condition. They were not informed which highlighting condition would be used because it could introduce a bias (e.g. participants might read a paragraph more carefully if they knew no highlighting would be provided). After reading a whole paragraph with an interruption, they pressed the left mouse button to finish reading, and answered a comprehension question with three choices (one correct answer, one wrong answer, and "I don't know"). After all 10 trials, three questions about documentary video clips were given, which were not for analysis but for encouraging participants to concentrate on the clips. Lastly, we gathered feedback from the participants using a 7-point Likert scale questionnaire (1=Strongly disagree, 7=Strongly agree) and an interview.

RESULT

Among 90 trials, six trials were excluded from analysis because of excessive reading time (longer than two standard deviations from the mean) or skewed gaze pattern.

Participants' eye movements were examined using an I-VT fixation filter as shown in Table 1. The highlighting conditions were compared using ANOVA. No significant difference was found between five highlighting conditions in terms of reading time, the number of fixations, mean fixation duration, and answer accuracy ($p > .05$ for all measures). The accuracy was overall high, which implies the paragraphs were easy to comprehend.

We visually inspected the length of resumption lag for all trials without highlighting. The resumption lag was calculated by the time from the moment when the interruption ended to the moment when a reading fixation sequence within the pre-interruption text was first detected. To detect a reading fixation sequence, we used the reading detection algorithm proposed by Sharmin et al. [8]. We found a significant difference in mean resumption lag among highlighting conditions ($F_{4,\ 32} = 8.001$, $p < .001$). Post hoc analysis indicated that participants could resume reading significantly faster in the block and sentence conditions than in no highlighting condition (2.4 s vs. 1.2 s, $p < .05$ for block condition, 2.4 s vs. 1.3 s, $p < .05$ for sentence condition).

Overall, participants agreed that four highlighting methods were helpful to resume reading after interruption (preference score, 6.0±0.5). Participants highly preferred the block (5.2±1.1) and sentence (5.4±1.1) methods. However, participants disagreed with the usefulness of the previous sentence method (3.7±1.9). Qualitative feedback (interviews) revealed that two far apart markers which were common in the previous sentence method might be distracting to make participants confused about where to restart reading.

Implications of the Pilot Study The block and sentence methods seemed most promising in terms of both resumption lag and preference. Therefore, we determined to use only these two methods in the user study. Also, we introduced a two-level difficulty factor to observe how the resumption lag varies according to the text difficulty.

USER STUDY

Based on implications of the pilot study, we conducted a user study to further evaluate the usefulness of the block and sentence methods with two-level difficulty of text. Apparatus and procedure were identical to those of the pilot study.

Participants 13 participants (2 females) from a university, aged 20 to 28 years with normal or corrected-to-normal vision, participated with 10,000 KRW (about $10) reward. All of them were native Korean speakers.

Design We used a within-subjects design with three highlighting conditions (no highlighting, block method, and sentence method) and two difficulty conditions (easy and difficult). Each participant read 12 paragraphs with different highlighting and difficulty conditions. Since there were six combinations of conditions, each participant experienced each combination twice. The order of combinations was determined by a balanced Latin Square.

Six easy paragraphs were chosen from the paragraphs in the pilot study (10.5±1.0 sentences and 149.8±7.9 words). Six difficult paragraphs were extracted from abstracts of M.S. theses (8.7±0.8 sentences and 149.1±6.6 words). According to the difficulty measuring formula proposed by Yun [10], the difficulty value of difficult paragraphs (grade level: 14.0±2.1) was significantly higher than that of easy paragraphs (grade level: 10.0±1.5) (t_{10} = -3.806, $p < .005$).

RESULT & DISCUSSION

For the same reason as in the pilot study, 13 trials were discarded from 156 trials. We analyzed the results of the user study using two-way ANOVA (Table 1). For all eye movement measures, no significant difference was observed between three highlighting conditions ($p > .05$ for all measures). However, as expected, all eye movement measures significantly increased in the difficult condition ($p < .05$ for mean fixation duration; otherwise, $p < .001$). No interaction effect was found.

In terms of mean resumption lag, we found a significant effect of both highlighting method ($F_{2, 24}$ = 15.199, $p < .001$) and difficulty level ($F_{1, 12}$ = 15.416, $p < .005$). The difficult paragraphs induced longer resumption lag. Also, post hoc analysis indicated that the block and sentence methods significantly reduced the resumption lag regardless of the text difficulty ($p < .005$ for block method, $p < .01$ for sentence method). No interaction effect was found.

When reading difficult text, participants chose wrong answers significantly more ($F_{1, 12}$ = 16.092, $p < .005$). We observed a difference in answer accuracy among highlighting conditions but the difference was not significant ($F_{2, 24}$ = .026, p = .975).

We compared the preferences of block and sentence methods using the Wilcoxon signed-rank test. Participants preferred significantly more the sentence method (6.1±0.9) to the block method (5.4±0.7) (Z = -2.029, $p < .05$). Also, they agreed that the highlighting was helpful to resume the reading (6.1±0.9) but it did not help to improve the comprehension (4.2±1.2), which was consistent with the result of answer accuracy. Lastly, they replied that they would use EyeBookmark for reading if possible (5.6±1.0).

In four trials (7.7%) with no highlighting condition, we found that gaze patterns of the pre-interruption and post-interruption periods did not overlap, implying that participants skipped some words. However, such cases did not happen in the block and sentence methods.

Results of the user study demonstrate that EyeBookmark successfully reduces resumption lag after interruption during reading. Although EyeBookmark did not affect reading comprehension, our results suggest that EyeBookmark helps participants not to miss any part of text.

CONCLUSION & FUTURE WORK

In this work, we present gaze-based bookmarking, EyeBookmark, which helps users to resume reading after interruption. Our user study proves the efficacy of EyeBookmark in terms of resumption lag and preference.

In our two experiments, we did not inform participants in advance whether highlighting would be provided or not. Interesting future work would be to compare the performance of different highlighting methods when participants are given the opportunity to tune their reading behavior accordingly by letting them know which highlighting method would be used in advance.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2014R1A2A2A03006998).

REFERENCES

1. Buscher, G., Biedert, R., Heinesch, D., and Dengel, A. Eye tracking analysis of preferred reading regions on the screen. *Proc. CHI EA'10*, 2010, 3307-3312.
2. Cane, J.E., Cauchard, F., and Weger, U.W. The time-course of recovery from interruption during reading: Eye movement evidence for the role of interruption lag and spatial memory. *The Quarterly Journal of Experimental Psychology 65*, 7 (2012), 1397-1413.
3. Duggan, G.B. and Payne, S.J. Skim reading by satisficing: evidence from eye tracking. *Proc. CHI'11*, 2011, 1141-1150.
4. Inhoff, A.W. and Weger, U.W. Memory for word location during reading: Eye movements to previously read words are spatially selective but not precise. *Memory & cognition 33*, 3 (2005), 447-461.
5. Latorella, K.A. *Investigating interruptions: Implications for flightdeck performance*. National Aeronautics and Space Administration, 1999.
6. McFarlane, D.C. and Latorella, K.A. The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction 17*, 1 (2002), 1-61.
7. Olsen, A. *The Tobii I-VT Fixation Filter*. Tobii Technology, 2012.
8. Sharmin, S., Špakov, O., and Räihä, KJ. Reading on-screen text with gaze-based auto-scrolling. *Proc. ETSA'13*, 2013, 24-31.
9. Text 2.0, http://text20.net/, Jan 2015.
10. Yun, C. A Study on the development of a readability formula in non-literary texts, 2006