Generating and Using Gaze-Based Document Annotations

Georg Buscher
Dept. for Knowledge-Based Systems
University of Kaiserslautern
Kaiserslautern, Germany
georg.buscher@dfki.de

Andreas Dengel
Dept. for Knowledge-Based Systems
University of Kaiserslautern
And
Knowledge Management Dept.
DFKI GmbH
Kaiserslautern, Germany
andreas.dengel@dfki.de

Ludger van Elst
Knowledge Management Dept.
DFKI GmbH
Kaiserslautern, Germany
ludger.van_elst@dfki.de

Florian Mittag
Knowledge Management Dept.
DFKI GmbH
Kaiserslautern, Germany
fmittag@dfki.uni-kl.de

CHI 2008, April 5–10, 2008, Florence, Italy.
ACM 978-1-60558-012-8/08/04.

Abstract
In this paper we describe a prototypical system that is able to generate document annotations based on eye movement data. Document parts can be annotated as being read or skimmed. We further explain ideas how such gaze-based document annotations could enhance document-centered office work in the future.

Keywords
Eye tracking, reading detection, line matching, OCR, document-centered office work

General Terms
Design, Measurement

ACM Classification Keywords
H.5.2 [User Interfaces]: Input devices and strategies

Introduction
In the last few years much effort has been put into the further development of eye tracking devices. Nowadays they have reached such a state of development that they are unobtrusive and easy to use and that their accuracy is sufficient for many practical applications. Since their development is progressing further, there is a good chance that they will become more affordable and, as a consequence, become more widespread. This

means that their main areas of application of today (e.g., usability studies) will be shifted to bigger areas like, e.g., applications in the office environment.

The office environment is the focused application area in this paper. Especially for knowledge workers, most work done in an office is document-centered work. Today, many documents are viewed and edited with the computer and we think that such digital document processing will increase in the future (be it on large screens, on e-paper, or on normal paper that is tightly coupled with a digital representation). In the digital environment, meta-data such as annotations for a document can be useful for a variety of applications. E.g., highlightings are typical user-generated annotations. They often indicate interest but are rarely created. Since the eyes are (almost always) involved while working with documents, gaze-based document annotations are far more frequent and one can expect that they can be used to make office work more effective.

In this paper we describe a robust method for gaze-based document annotation and present ideas directing our future research, i.e., how such annotations should be used for more effective office work. But first, we give a rough overview of related work.

Related Work

Some research has been conducted in the past related to eye tracking while working with documents. Gaze data has been used for estimating the relevance of a text in an information retrieval task [2, 9]. The studies aimed at finding and combining eye movement measures in order to create a system that automatically predicts relevance of a viewed text.

A related area of research focuses on gaze-based proactive information delivery. The scenario of [5] is to provide the reader with translations of unknown words in a text of a foreign language. The work aimed at detecting unknown words automatically by analyzing eye movements. The application described in [7] used gaze data, among other behavioral data, to provide the user with relevant information for the currently viewed or worked-with document. Here, simply looking at an entry on a search result page for a longer time is taken as positive relevance feedback from the user. Then the appropriate entire document is displayed automatically.

The work in [1, 2, 3, 4] focuses on interpreting eye movements to come to higher abstractions of the gaze data. They describe methods that detect whether a person is currently reading by analyzing the sequence of eye movements. Our method for creating gaze-based document annotations uses ideas of that work.

From Gaze Data to Document Annotations

Roughly spoken, eye movements are composed of fixations and saccades. During a fixation the eyes are steadily gazing at one point. A saccade is a quick movement from one fixation to the next. Since the sequence of fixations and saccades is very characteristic during reading behavior, it is possible to detect whether a person is currently reading or skimming a text.

Reading and Skimming Detection

As described in [3], we conceptualized and implemented an algorithm being capable of detecting and differentiating between reading and skimming behavior. For testing its functionality, we applied a Tobii 1750 desk-mounted eye tracker which has a data generation frequency of 50 Hz and an accuracy of around 40 pixel

a) Sequences of saccades over text
solid: read
dashed: skimmed

b) Horizontally averaged saccade sequences

c) The third saccade sequence is matched with the third text row. Green and blue rectangles indicate detected nearby text rows.

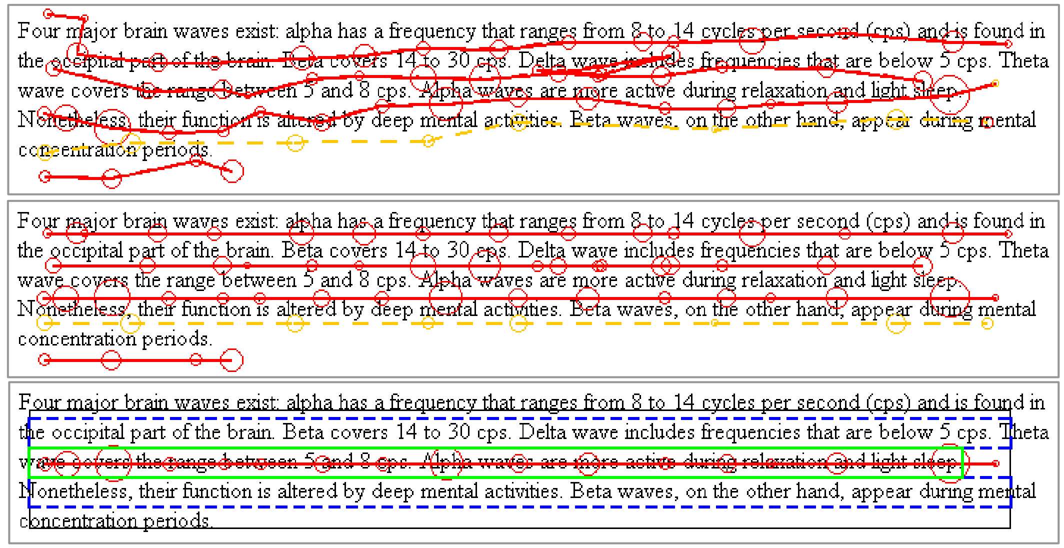

figure 1. Three processing steps to match saccade sequences with actual text rows.

at a resolution of 1280x1024. In short, the algorithm works in several steps: First, the raw gaze data is accumulated to fixations. Second, the saccades between fixations are classified according to their directions and distances. If there are enough successive saccades that might belong to reading (or skimming) behavior of one text row, then reading (or skimming) is detected along the path of the saccades on the fly. Figure 1a shows some paths of saccades belonging to reading (solid lines) and skimming (dashed lines) behavior. The differentiation between reading and skimming is simply done based on the saccade lengths, but this will not be discussed here any further.

Mainly due to eye tracker inaccuracies the saccade paths are not very smooth (especially horizontally). Therefore, we apply a method that averages the zigzag lines horizontally: the averaged horizontal position is the average of all fixations' horizontal positions. (Since the accuracy of the eye tracker varies spatially, we do not take a duration-weighted average.) See figure 1b for an example.

Matching the Lines

The eye tracker inaccuracies do not only consist of noise that can be managed with simply by averaging values. What is more, the inaccuracy is biased and varies from person to person. To handle this bias in a generic way, we apply OCR (optical character recognition) techniques: If a sequence of saccades has been classified as reading or skimming behavior, a small screenshot around all the saccades is taken and analyzed by the OCR system OCRopus [8]. This system returns the positions of all text rows on the screenshot (see figure

1c). By moving the horizontally averaged line of saccades to the nearest text row, we match the eye-tracked line with the actually read text on the screen.

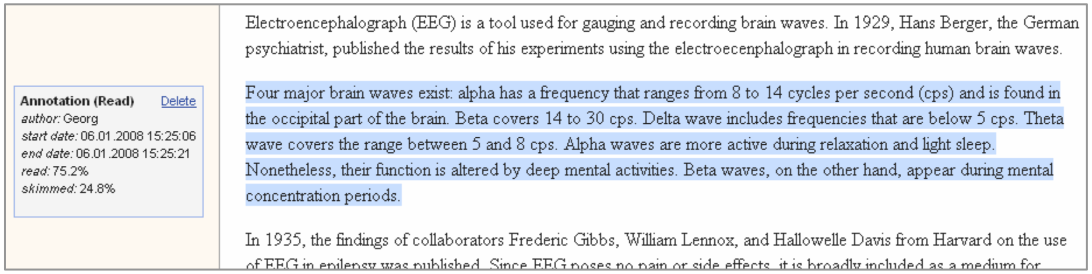

figure 2. Prototypical visualization of a gaze-based annotation for a wiki document.

To get increased precision in matching the lines we apply two main heuristics: Since we always have two lines to match – the saccade line and the most plausible text row – we can use the spatial difference between them to recalibrate the eye tracking device on the fly. The second heuristic takes advantage of the mostly sequential nature of reading behavior: one normally reads line by line from top to bottom. If a saccade line is located between two text rows and the upper text row has been read or skimmed immediately before, then the lower text row is more likely to be chosen as the matching row. A very preliminary case study indicated that the accuracy of the line matching is around 90%. In the remaining 10% a line directly above or below the currently read line was selected.

Annotating a Document

Now that we matched each saccade line with the most plausible text row, we can annotate the document accordingly. Since not all applications are capable of storing annotations or do that differently, we prototypically extended the wiki system *Kaukolu* [6] to be able to handle annotations. Hence, when a person reads an article presented by that wiki, the read and skimmed text parts automatically get annotated appropriately. However, not only the type of reading behavior is stored (i.e., reading or skimming) with the annotation but also, e.g., a timestamp and the duration needed for reading. Moreover, since wikis are typically used by many people, the name of the reader (if logged in) is also saved. Figure 2 gives an impression of an annotated wiki-document.

Usage of Gaze-Based Document Annotations

By keeping track of the user's eye movements over a document, our implemented system can gather very valuable attention information being useful for a variety of applications. Some use-cases that we think are worth pursuing in the future are shortly presented in the following.

Recontextualization

An obvious way is to use such document annotations for recontextualization purposes. It is often happening that one re-opens a document after some time in order to find exactly the same information as during the first time. Especially when the document has dozens or hundreds of pages (e.g., an e-book) the time needed to find the specific piece of information can be annoyingly long. Since one often knows from the first time viewing the document that the information has to be in there somewhere, it would be helpful to see which document parts have been viewed before. So, the gaze-based annotations can be used as a document-internal filter.

Such a filter would be especially useful for people like, e.g., lawyers who often have many ongoing cases in parallel. Such cases often include a number of long documents where only a few paragraphs really matter to the case. When a lawyer switches from one case to a previous one, it might be very helpful to get a quick overview of the text parts that mattered before, i.e., that had been viewed in the past.

Looking for something known or unknown?

In the last years, desktop search engines became more and more popular. They help to find documents stored on the own computer. Here, gaze-based document annotations provide the possibility for a new kind of search filter: does the user want to re-find information or to find new information? In the former case, the search engine should ignore all not viewed text parts of the documents in the search process. In the latter case, it should only consider not viewed or roughly skimmed text parts.

Such a filter function would also be useful in a lawyer scenario: Some lawyers have very large collections of law comments on their computers (i.e. comments on how the different laws should be interpreted). Such collections are normally used as reference books. Therefore, a search engine that could explicitly distinguish between searches for content that has been read before and not viewed content could be very helpful. The former case could be useful when the lawyer wants to make sure a remembered fact. The latter case might be applied when he or she wants to obtain information (e.g., a new argument) that he or she is not aware of.

Personalized Search Result Abstracts

Normally, after putting in a query, desktop search engines return a list of documents as the result. For most search engines this list contains abstracts for each document showing the most characteristic sentences containing the query terms. Gaze-based document annotations could influence the methods for generating those document abstracts. As a consequence, the user might be able to recognize the important contents of the documents more quickly and might not have to open the whole document. This is might be especially useful for re-finding information.

Searching for similar attention patterns

It is also conceivable that gaze-based annotations can be used for finding documents that other people have used with a similar interest. For example in an enterprise one could think of a central information system that hosts a lot of documents (say, several thousands, e.g. scientific papers) that might be interesting for the employees. From time to time it happens that an employee has a similar interest like another employee before. Considering a research institute as an example, one researcher, researcher 2, might want to get an overview of how a certain technique has been used in a specific domain. But he or she does not know that some weeks ago, a colleague, researcher 1, had exactly the same interest, looked through several papers, and found some interesting text parts. Hence the current researcher cannot take advantage of the first researcher's experiences and has to get an own, independent overview.

However, having kept track of both people's eye movements on the opened papers, the central information system could identify a similarity in the usage of the papers by both researchers. I.e., it could notice that there is a certain similarity in the currently viewed document parts by researcher 2 and 1. Having noticed such a similarity, the system could suggest those papers and paper parts to researcher 2 that researcher 1 has read intensely but that researcher 2 has not viewed, yet In that way, researcher 2 could also be alerted if he or she opened and skimmed an article but missed the section being most interesting to researcher 1.

In this scenario, gaze-based document annotations are useful for identifying similar viewing patterns and for calling the user's attention to probably important sections that would otherwise be missed.

Conclusion

In this paper, we described a prototypical system that interprets eye movements recorded by an eye tracker and uses this information for automatic gaze-based annotations of documents. By incorporating ORC (optical character recognition) techniques our method is able to determine the text lines that have been read or skimmed even if the eye tracker provides imprecise and biased gaze location data. Read and skimmed lines that have been detected as such can be annotated in a wiki system.

Gaze-based document annotations open up a wide range of possibilities of how to enhance document-centered work. Since the eye is always involved while working with documents and reflects the user's attention in a way, such annotations are of great value for different kinds of applications. We described some ideas for gaze-enhanced applications that we believe would make office work much more effective. As eye trackers will become cheaper and more precise, they could be convincing applications of the future.

Acknowledgements

We thank Daniel Keysers and Christian Kofler for providing a customized version of OCRopus. This work was supported by the German Federal Ministry of Education, Science, Research and Technology (bmb+f), (Grant 01 IW F01, Project Mymory: Situated Documents in Personal Information Spaces).

Citations

[1] Beymer, D., and Russell, D.M. *WebGazeAnalyzer: a system for capturing and analyzing web reading behavior using eye gaze*. Proc. CHI '05. 2005.

[2] Brooks, P., Phang, K. Y., Bradley, R., Oard, D., White, R., and Guimbretire, F. *Measuring the utility of gaze detection for task modeling: A preliminary study*. IUI'06. 2006.

[3] Buscher, G., Dengel, A., van Elst, L. *Eye Movements as Implicit Relevance Feedback.* CHI '08 extended abstracts (accepted for publication), 2008.

[4] Campbell, C.S., and Maglio, P.P. *A robust algorithm for reading detection*. Proc. PUI '01, 2001.

[5] Hyrskykari, A. *Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid*. PhD thesis. University of Tampere, Dept. of CS. 2006.

[6] Kiesel, M. *Kaukolu – hub of the semantic corporate intranet*. Proc. Semantic Wiki Workshop at ESWC 2006.

[7] Maglio, P.P, Barrett, R., Campbell, C.S., Selker, T. *SUITOR: An attentive information system*. IUI'00. 2000.

[8] OCRopus, an open source document analysis and OCR system (http://code.google.com/p/ocropus/).

[9] Puolamäki, K., Salojärvi, J., Savia, E., Simola, J., and Kaski, S. *Combining eye movements and collaborative filtering for proactive information retrieval*. Proc. SIGIR '05. ACM Press (2005), 146–153.