Sentimental Eyes!

Amitava Das[1] and Björn Gambäck[2]

[1] SAIT Lab., Samsung Research India, Bangalore, India
[2] Norwegian University of Science and Technology, Trondheim, Norway
amitava.santu@gmail.com, gamback@idi.ntnu.no

Abstract. A closer look at how users perform search is needed in order to best design a more efficient next generation sentiment search engine and understand fundamental behaviours involved in online review/opinion search processes. The paper proposes utilizing personalized search, eye tracking and sentiment analysis for better understanding of end-user behavioural characteristics while making a judgement in a Sentiment Search Engine.

Keywords: Sentiment Analysis, Sentiment Search, Eye Tracking.

1 Introduction

Broad access to an abundance of information is one of the defining characteristics of today's web search environment. Internet search engines act as intermediaries between users' information needs and the massive number of potentially relevant pages on the Web. Still, users are largely unsuccessful in finding their desired information, with failure rates often approaching 50% [1][2]. Clearly, this presents a significant dilemma for online searches – why are users only modestly successful in formulating their search queries, and what can be done to improve the situation?

Several attempts have been made to better understand user behaviour during the search process, for example through *personalized information retrieval*. Personalized web search is crucial in today's world of information overflow, to provide only relevant information – depending on the context – such that users get the correct information when they need it. According to Schwartz (2006), "choice is the critical sign that we have freedom and autonomy" [3]. Most of the time, choice is good and more choice is better. With the accessibility of more information, we have more choice, and presumably more freedom, autonomy, and self-determination, than ever before. It would seem that increased choice increases well-being; however, studies have shown that this is not the case: there is a need for good (personalized) information retrieval systems that help the user to take good decisions without decreasing her well-being. Generally, personalization methodologies can be divided into two complementary processes: user information collection, used to describe the user interests [4], and inference of the gathered data to predict the closest content to the user expectation [5]. Hence, future generation web interfaces by necessity need to be more intelligent in order to understand the end-users' sentimental needs and preferences.

There has been a rapid development in sentiment analysis techniques during the last two decades and sentiment search is one of the most promising futuristic

M. Kurosu (Ed.): Human-Computer Interaction, Part V, HCII 2013, LNCS 8008, pp. 310–318, 2013.

technologies with immense commercial value. The main driving necessity behind sentiment search systems is that whenever we need to make a decision, we may want opinions from others. Some sentiment/review/opinion search fall into the informational genre by definition, with queries for this kind of task being classified into two basic genres: direct search and comparative search, which could be instantiated by queries such as "iPhone 5" and "iPhone5 vs. Samsung Galaxy", respectively. Unfortunately, there is no publicly available system which gives satisfactory output for this type of querying, and posing opinionated queries to a general purpose search engine leads to navigational surfing for users because desired information is distributed over several pages.

Today's search engines keep logs of user browsing data and effectively use that data to produce satisfactory results for particular users. In addition, we suggest to keep logs of user eye-tracking data, in order to understand and track the user's sentiment while working with the Web-search interfaces. Most laptops, smart phones and tablets have good quality cameras built in, while eye-and-gaze tracking technology has reached a quality level where such single-camera tracking without external light or infrared sources is feasible (for example, an already available system such as YouEye[1] is capable of tracking eye-gaze and facial emotions just using a standard web camera). Allowing for such non-intrusive eye-tracking is central to the possibility of utilizing the technology in a range of tasks, including the tracing of user search behaviour. Hence, eye tracking and sentiment analysis could have a great significance on the next generation of Human Computer Interfaces.

In this paper, we report some initial experiments on using eye tracking information as a knowledge source for sentiment analysis. It is an on-going task and this paper should overall be seen as a position paper. The rest of the text is laid out as follows: the next section discusses some relevant previous research efforts. Section 3 in turn introduces the basis for the current text, the long-term research questions that need to be addressed. The experimental setup is described in Section 4, while Section 5 presents some initial results. Finally, Section 6 sums up the discussing and points to areas of future interest.

2 Related Work

The application of eye tracking to online search has recently received a considerable amount of attention from research scientists, search engine companies, marketing firms, and usability professionals, even though no previous work has focused directly on using eye-tracking for Internet search sentiment analysis. The first use of eye tracking to investigate Internet search behaviour comes from Granka *et al.* (2004) who analysed users' basic eye movements and sequence patterns throughout ten different search tasks performed on Google [6]. Lorigo *et al.* (2006) [7] examined user eye tracking patterns through fixation on classified areas of interest (AOI) such as title, abstract, and metadata.

It has long been known in neuropsychology that the retinal image is transmitted to the brain during fixations, but not during saccades (rapid eye movements); hence, it is

[1] https://www.youeye.com/

the fixations that represent the acquisition and processing of information [8]. During normal reading, a reader does not fixate upon each word in sequence, but rather makes a rapid series of fixations followed by saccades, which may skip over some words entirely. Saccades commonly occur 3-4 times per second. In addition, approximately 15% of all the saccades occur backwards, to earlier text – a phenomenon known as a regression.

Cognitive psychologists have studied how viewers examine printed advertisements, and in particular how different aspects of the ads and the users' goals interact to influence viewing behaviour. In recent years eye-tracking technology has been utilized for these studies, in order to automatically determine how much time readers devote to specific areas of interest in an ad. Rayner *et al.* (2008) asked readers to rate how much they liked an ad and then examined the correlation between these ratings and how much time the readers spent on the ad, as well as how the viewing time was divided between textual objects and images, concluding that the user needs and the actual users' profiles matters most for how much attention is devoted to an ad [9]. A related problem which has been studied is how eye-tracking can be utilized to trace the processes underlying user decision making. Glaholt & Reingold (2011) has proposed that the dwell duration is central there, i.e., that users tend to look longer at items they prefer, while dwell frequency – how many times they look at an item – is less important [10].

3 Empiricism on Eye-Tracking and Sentiment Analysis

The research motivation of the present work is to reach better understanding of user behaviour during the sentiment search process. This can be formulated into basic level objective questions based on what has been suggested in previous studies [6]:

- How long does it take searchers to select a document?
- How many abstracts do searchers look at before making a selection?
- Do searchers look at abstracts ranked lower than the selected document?
- Do searchers view abstracts linearly?
- Which parts of the abstract are most likely to be viewed?

However, there are contrastive differences between general search and sentiment search engines. In order to adapt eye-tracking methods to sentiment search, those research questions have to be extended in the following two directions, which are the key contributions of this paper.

1. Overall Searching Behaviour

   – How long does it take users to select a document based on a direct query vs. a comparative query?
   – How long does it take users to select a document from a general query vs. from a sentiment query?
   – How much time do we spend viewing each abstract?

2. Overall Viewing Behaviour

   - How many times does a user look at the query word(s)?
   - How many times does a user look at the sentiment word(s)?
   - How many times does a user look at the domain specific query word(s)?
   - How many times does a user look at his/her preferred sentiment word(s)?

The answers to these empirical questions will help to improve search interfaces in the future and will also help in personalizing them according to an end-user's preferences. It could be argued that our preferences of sentimental word choices or websites preferences do not differ in practice. However, a separate study on social network personality by Kosinski et al. (2012) has already reported that user website preferences change with user personality [11]. To support this argument, our experimental results are presented in the next section.

4 Experiments

A set of very initial experiments was carried out at the Department of International Business Communication, Copenhagen Business School, using an EyeLink 1000[2] eye tracker from SR Research Ltd. The experimental setup was developed by using the Experiment Builder[3] software that comes with the eye tracker toolkit.

Three test participants were instructed to formulate four types of queries each, one from each category listed below. The participants were asked to formulate queries that very restricted to the specific domain: the direct sentiment queries should be restricted to be within the movie domain, while the comparative sentiment queries should be in the electronic domain.

- General domain, non-sentiment queries
  - Informational (Ex: *USA president*)
  - Navigational (Ex: *tourist info Lyon France*)
- Sentiment queries
  - Direct Sentiment (Ex: *Skyfall review*)
  - Comparative (Ex: *iPhone 4 vs. Galaxy*)

The actual document retrieval was carried out by the Google search engine. A typical search page consists of title, URL, text snippets, images, video links, and metadata. For the task fixation, the title, abstract, URL and metadata were classified as areas of interest (AOI).

5 Analysis

The end-user search patterns significantly change depending on the query type. It could be observed that most of the time people do not look at the result snippets

[2] http://www.sr-research.com/EL_1000.html
[3] http://www.sr-research.com/accessories_EL1000_eb.html

linearly except for during general domain informational queries. Furthermore, for that type of queries people generally do not look beyond the top-5 results. Even only for 40% cases people have click on any URL. We discussed this with the participants and deduced that their informational needs generally were fulfilled by the text snippets obtained in the search results.

5.1 Sentiment Query vs Non-sentiment Query

Interestingly, we observed that almost everybody randomly moved their eyes over the search results for sentimental queries. For example, for movie reviews people had preferences for the sites IMDB[4] or Rottentomatoes[5]. It did not matter much how those sites were ranked in the results returned by the search engine: people generally jumped to those search results and fixed their eyes.

The direct sentiment queries often needed to be re-formulated and the participants in general added their preferred aspects of choices like acting, direction, academy award, etc. The observed characteristic differences between the general domain queries and sentiment queries are reported in Table 1.

Table 1. Differences between general domain queries and sentiment queries

Query type	Average time to complete	Snippets linearly visited	Reformulation of query
Informational	10-20 sec	Mostly	No
Navigational	60-90 sec	Varies	Mostly and after each 20-30 sec
Direct Sentiment	30-40 sec	No	For only 40% of the cases
Comparative	60-90 sec	No	Mostly and after each 30 sec

5.2 Term Preferences

In addition to the fundamental behavioural differences between general search queries and sentiment queries, we analysed people eye fixations on the *query word(s), domain specific word(s)* and *sentiment word(s)*. The observations are reported in Table 2.

Two domain dictionaries were created for the experiments. The first dictionary is in movie domain and consists of 100 domain ontologies like acting, direction, academy award, cinematography, etc. The second dictionary is on computer and electronic products and other associated terms like apps, display, battery, software, etc. It was created semi-automatically by automatically merging two online dictionaries[6,7] and manually validating the result. The dictionary has 5K terms altogether; the words carrying sentiment were restrieved from SentiWordNet 3.0[8].

[4] http://www.imdb.com/
[5] http://www.rottentomatoes.com/
[6] http://www.alphadictionary.com/directory/Specialty_Dictionaries/Electronics/
[7] http://www.interfacebus.com/Glossary-of-Terms.html
[8] http://sentiwordnet.isti.cnr.it/

Table 2. Eye fixation on categorical terms

Query type	Query Words	Domain Specific Words	Sentiment Words
Direct Sentiment	30%	18%	43%
Comparative	36%	26%	32%

Google highlights the query terms automatically, so we were quite surprised to find that the participants did not look at the query words very often. Rather, the fixation statistics show that a user stop at either domain specific or sentiment words. Post-experiment discussions with the participants revealed that they generally had a prior expectation on exactly what they were looking for. For example, the "skyfall review" was given to the three participants and they revealed that they had had the following expectations before initiating the search:

1st Participant: Interested in Acting: whether the new Bond is better than Pierce Brosnan / Sean Connery. Also interested in Academy Awards.
2nd Participant: Interested in Bond gadgets specifically!
3rd Participant: Interested in action sequences only!

The same is true for the comparative sentiment queries. In the context of iPhone vs. Galaxy, people generally want to look at feature-based sentiment comparisons for electronic products like aps, display, and battery life. For that reason people are more concerned with the domain terminologies during comparative sentiment queries than during direct sentiment queries.

5.3 Structural Preferences

Search results generally have a typical structure. In order to investigate the users structural preferences, we considered the following items: title, URL, metadata, image/video thumbnail. The text snippets were excluded, as they are content features and had been analysed separately in the term preference part described above.

Table 3 shows how a typical user's eye-fixation stops at each structural aspect. The percentage calculation is based on the time spent on structural aspects divided by the overall time spent on the page before clicking on any link.

Table 3. Structural Preferences

	Title	URL	Metadata
Direct Sentiment	1-2 %	32%	66%
Comparative	1-2 %	10 %	23 %

Heat maps were generated from each user's browsing data for further analysis. Two very relevant example heat maps can be seen in Figures 1 and 2. The first heat map is on "skyfall review" and it is clear that users stop his/her eyes on the star-rating by Rottentomatoes. The second heat map is from a comparative sentiment query and it is clear that the user is more interested in the image/video link.

We discussed these issues with participants after the experiments and understood that they looked at particular structural aspect based on the topic(s) of their search. For example, if it is a general knowledge topic they tend to look at Wiki URLs; for movies they look either at Rottentomatoes or IMDB; and for products they have their personal choices for reading reviews.

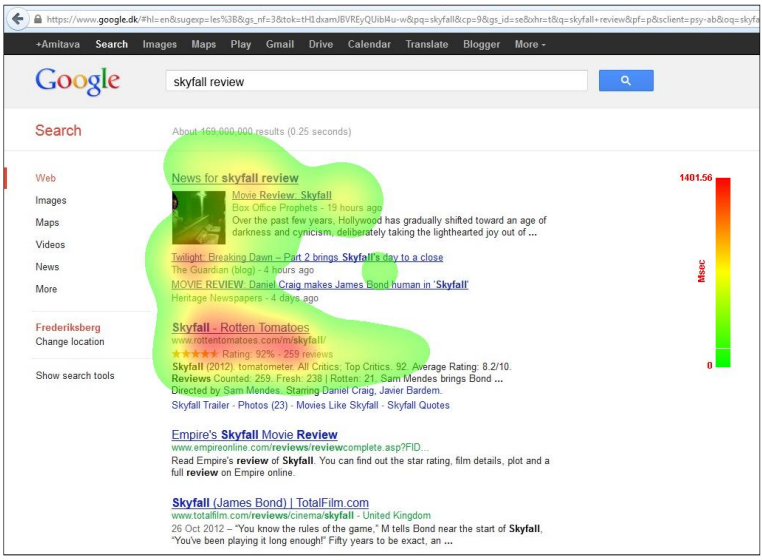

Fig. 1. Heat map for a direct sentiment query: "*skyfall review*"

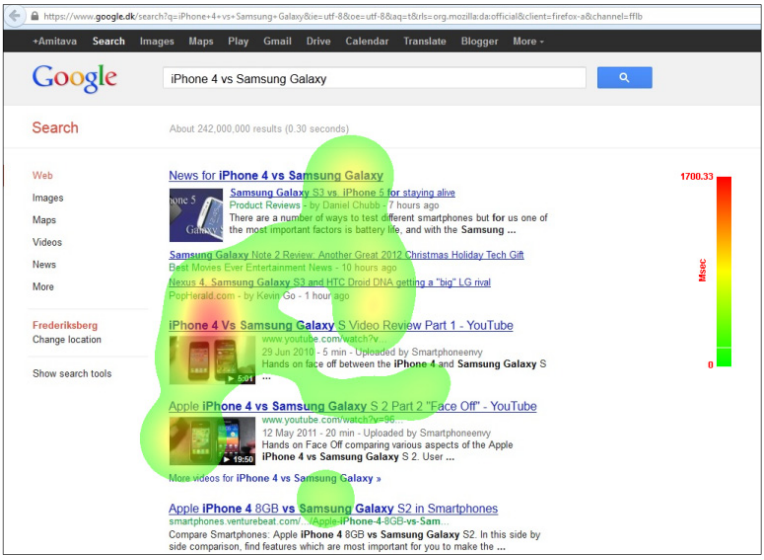

Fig. 2. Heat map for a comparative sentiment query: "`iPhone 4 vs. Galaxy`"

6 Conclusions and Future Aspects

In conclusion, the paper has reported some incipient work on understanding user eye movements and fixations based on their sentimental preferences during online search. There is a huge potential for this research when moving towards the next generation of Human Computer Interfaces, since eye-and-gaze tracking technology has reached a quality level where it now is feasible to utilize remote, non-intrusive single-camera eye-tracking, using a standard web camera without external light or infrared sources.

This paper only reports the initial study to understand the relations between the eye movements and user sentiment search patterns. At the next level, we are working towards personalized sentiment search by creating user profiles with the technique, and with the intension to add facial emotions to the eye-tracking. No similar work has been attempted so far, but a US patent application has outlined an architecture using brain-computer interface (BCI) technology for sentiment tracking [12]. It suggests that the BCI system could be complemented by measuring eye and face movement activation signals. Thus in a quite intrusive manner – and using an as-of-yet fairly unreliable input method. The present paper in contrast proposes to induce the user sentiment in a totally non-intrusive manner and by utilizing quite mature and cheap off-the-shelf eye-tracking techniques.

Acknowledgements. The experiments were carried out at Copenhagen Business School, Denmark. We want to thank Prof. Michael Carl for providing the laboratory facilities equipped with the eye-tracker.

References

1. Sherman, C.: Why search engines fail. Search Engine Watch (August 29, 2002)
2. http://searchenginewatch.com/article/2068006/Why-Search-Engines-Fail
3. Nordlie, R.: User revealment – a comparison of initial queries and ensuing question development in online searching and human reference interaction. In: Proceedings of SIGIR, New York, USA, pp. 11–18 (1999)
4. Schwartz, B.: The paradox of choice – why more is less. Google Tech Talks (2006), http://www.youtube.com/watch?v=6ELAkV2fC-I
5. Winter, S., Tomko, M.: Translating the web semantics of georeferences. In: Web Semantics and Ontology, pp. 297–333. Idea Publishing (2006)
6. Koutrika, G., Ioannidis, Y.: A unified user-profile framework for query disambiguation and personalization. In: Proceedings of the Workshop on New Technologies for Personalized Information Access, Edinburgh, UK, pp. 44–53 (2005)
7. Granka, L., Joachims, T., Gay, G.: Eye-tracking analysis of user behavior in WWW search. In: Proceedings of the SIGIR Conference on Research and Development in Information Retrieval, New York, USA, pp. 478–479 (2004)
8. Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., Gay, G.: The influence of task and gender on search and evaluation behavior using Google. Information Processing and Management 42(4), 1123–1131 (2006)

9. Glaholt, M.G., Wu, M.-C., Reingold, E.M.: Predicting preference from fixations. PsychNology Journal 7(2), 141–158 (2009)
10. Rayner, K., Miller, B., Rotello, C.M.: Eye movements when looking at print advertisements: the goal of the viewer matters. Appplied Cognitive Psychology 22, 697–707 (2008)
11. Glaholt, M.G., Reingold, E.M.: Eye movement monitoring as a process tracing methodology in decision making research. Journal of Neuroscience, Psychology, and Economics 4(2), 125–146 (2011)
12. Kosinski, M., Stillwell, D., Kohli, P., Bachrach, Y., Graepel, T.: Personality and Website Choice. In: Proceedings of WebSci 2012, Evanston, Illinois, USA (2012)
13. Pradeep, A., Knight, R.T., Gurumoorthy, R.: Neurological sentiment tracking system. United States Patent Application Publication, US 2011/0270620 A1 (November 3, 2011)