GLASE-IRUKA: Gaze Feedback Improves Satisfaction in Exploratory Image Search

Viktors Garkavijs[†*] Rika Okamoto[†*] Tetsuo Ishikawa[‡*] Mayumi Toshima[†*] Noriko Kando[*†]
[†] Department of Informatics, The Graduate University for Advanced Studies
[*]National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 103-8480, Japan
[‡]Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology
G3-46, 4259 Nagatsuta-cho, Midori-ku, Yokohama, 226-8502, Japan
{gvb,lica,iskw,mamitako,kando}@nii.ac.jp

ABSTRACT

We propose two methods for exploratory image search systems using gaze data for continuous learning of the users' interests and relevance calculation. The first system uses the fixation time over the images selected by gaze in the search results pages, whereas the second one utilizes the fixation time over the clicked images and fixations over the non-selected images on the results page. A user model is trained and continuously updated from the gaze input throughout the whole session in both systems. We conducted an experiment with 24 users, each performing four search tasks using the proposed systems and compared the results to a baseline system, which does not employ any gaze data. The Gaze feedback system users viewed 22.35% more images than the users of the baseline system. A high correlation between the number of saved images and the satisfaction with the results was observed in data collected from the users of a mouse feedback system enriched by gaze data. The results show that including the gaze data into the relevance calculation in both cases increases the degree of satisfaction with the search results compared with the baseline.

Categories and Subject Descriptors

H.1.2 [**User/Machine Systems**]: Human factors

General Terms

Experimentation, Algorithms, Human factors

Keywords

User experiments, user studies, image searching, eye-tracking, exploratory search.

1. INTRODUCTION

Conventional search systems allow users to interact using a mouse and a keyboard as input devices. With the increasing availability of touchscreen devices, touch input has become widely used, especially in the image search user interfaces for mobile devices. However, all these types of interactions require users to use their hands for explicit feedback expressing their intent and commands to control the system, and unless the users make some effort to tell the system in some way, what they are thinking, the system may not be able to infer much about it.

Interactive information retrieval systems expect the user to be active and providing feedback to the system in order to refine the results. However, any explicit actions, such as providing feedback about the results, require time and effort, and searching is not exactly the pastime most people want to spend their time and efforts on.

The systems proposed in this paper allow users to refine the retrieved results of the initial query continuously by learning a user's interests in real-time from gaze data. This approach is highly interactive and provides fast result refinement and personalization within a search session. We focus on the search cases starting from vague initial queries.

In this work we propose two approaches to using gaze data for exploratory image search and report the results of an experiment.

2. PROPOSED METHODS

The proposed methods are "gaze feedback" ("*Gaze*") and "mouse feedback enriched by gaze data" ("*MeGaze*")[2].

Systems employ the same user interface (UI) and interaction logic. On the UI there is a query box, a search button, a search result panel (SRP), history and favorites bars.

The system is operated as follows: after the query is input into the query box, the SRP becomes populated with the results. The user may change the query or select a thumbnail on the SRP. If the user selects a thumbnail, a popup appears showing an enlarged version of the image. When the user closes the popup, a smaller version of that image is put into the history bar. Thumbnail selection and closing is called an *interaction iteration*.

Images are saved by dragging and dropping them from the history bar to the row above. Both history and favorites bars can hold up to 15 small image thumbnails. The newer thumbnails are added to the left of the line; after reaching full capacity, the thumbnails are shifted to the right. Placing the cursor above the small thumbnails in either the history or favorites bar slightly dims the search result matrix and shows the original image enlarged in the center of the screen.

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577313.

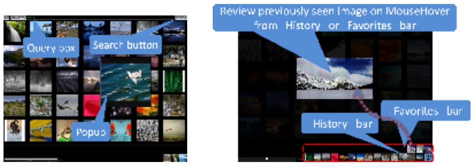

Figure 1: UI of the proposed system.

Table 1: Average values of system metrics.

Measure	Gaze	MeGaze	Baseline	p-value
Popups	150.43	97.29	117.57	0.41
FavoritesAdded	20.38	19.63	20.91	0.52
CollectingEasy	3.38	3.75	3.03	0.33
Satisfaction	3.91	3.34	3.56	0.23
CollectedEnough	3.84	3.66	3.16	0.32

In the "Gaze" method, we employed eye-tracking to capture continuously on which areas of the page the user's attention was being focused. The user is naturally looking at the images on the SRP, stopping on some images for longer. The popups are shown after the user fixates her gaze on a thumbnail for a specified threshold (we used 800 ms in the experiment). To close the popup, the user needs to move her gaze away from the popup area.

In "MeGaze", the user selects an image by clicking on its thumbnail — the same as in the baseline system. When the popup is shown, the learning algorithm employs gaze data to update the scores as described in the following subsection. To close the popup, the user needs to click anywhere on the screen — again, the same as in the baseline method.

2.1 Back-end

Our user model is a set T, where each element is a pair $t = (tag, score)$ for $t \in \mathcal{T}$. This model is updated after each interaction iteration. The initial values of each pair t are built from the tags of the MIRFLICKR-1M[3] dataset, assigning the initial score of zero to each tag.

The baseline and the *Gaze* versions use the "dwell-time" scoring mechanism of [1], while *MeGaze* employs an enhanced version of it. When the popup is shown, the scores are also added to the images on the search results page if the user is looking at them while the popup is open. Thus, the score added to the tags $s = s_i + \sum_{j=1}^{n} f(s_j)$, where s_i — score from the clicked image, s_j — scores from all the n images that had gaze fixations, calculated as $s_j = \frac{t}{m}$, with t being the dwell time and m — number of tags of the image. As for the function f, we tried various functions that grow faster than linear and employed $f(x) = x^2$ for the experiment.

3. EXPERIMENT

Two use cases were considered for experiments with the system — *known-item-search* (searching for a given image) and *designer tasks* (searching in accordance with some vague description). We prepared two tasks for each use case, a total of four tasks for each user. The images used for the known-item-search case were from the same dataset we used in the system (ID = 859626 and ID = 585623).

Twenty-four paid participants were recruited for this experiment. The selected participants were chosen so that they all had the same searching experience.

There were eight participants for each system type. Each participant had to perform all four search tasks, counterbalanced among the participants using 2x2 Latin squares.

The system was explained in a tutorial-style — the participants had to execute typical search steps during the system explanation under guidance and confirm that each step was clear. After the tutorial, the participants had 10 minutes for self-practice. When the participants confirmed that they understood how the system worked, they started executing the search tasks. Each task was followed by a post-task questionnaire.

4. RESULTS

The number of clicked images differed between the gaze-driven and mouse-driven systems. This was the expected outcome of this experiment and confirms the findings from the preliminary experiment[1].

The analysis of post-task questionnaires found that the users of the gaze-driven system UI assessed their satisfaction with the found images higher than the users of both mouse-driven systems. The users of the gaze-driven UI, however, rated their satisfaction lower for the first task than for the subsequent tasks. The average scores of task assessment for the remaining tasks showed that the users of the gaze-driven system felt more satisfied with the results as well as feeling that they collected enough images for the tasks.

The results of the experiment showed high correlation between the variables "Satisfaction" and "FavoritesAdded" for both proposed systems, while this correlation was low for the baseline system. From this we can conclude that introducing the gaze component in the image search system helps users find and save images that significantly increase the users' satisfaction with the results.

Although the "MeGaze" system uses only a mouse for choosing the images on a results page, the tag weighting is enriched by using gaze data. The correlation between the number of images moved from the history bar to the favorites bar and the satisfaction with the results makes us think that the gaze data ignored in the baseline system actually captures the user's interests and helps to retrieve results that better match the user's needs. This effect was not observed in the baseline system.

The correlation coefficients between "Satisfaction" and "CollectedEnough" between the Gaze and the baseline systems are 0.84 for gaze and 0.67 for the baseline system. While both values are significant ($p \ll 0.01$), further comparison of correlation coefficients for "CollectedEnough" and "FavoritesAdded" shows significant correlation only for the Gaze system. This could mean that using the gaze feedback helps to retrieve images that satisfy users more *and* finding a sufficient amount of results would increase the number of favorites saved from the history bar.

5. REFERENCES

[1] V. Garkavijs, M. Toshima and N. Kando. GLASE 0.1: Eyes Tell More than Mice. In *Proc. SIGIR'12*, pp. 1085–1086, 2012.

[2] http://glase.jp/glase-iruka/

[3] M. J. Huiskes and M. S. Lew. The MIR Flickr Retrieval Evaluation. In *Proc. MIR'08*, pp. 39–43, 2008.