IMPLICIT PHOTOWORK BASED ON EYE-GAZE DATA

Bojan Blažica
XLAB Research
Pot za Brdom 100, 1000 Ljubljana, Slovenia
e-mail: bojan.blazica@xlab.si

ABSTRACT

This paper presents a demo application for exploring the viability of using eye-gaze data for implicit human computer interaction in the context of photo collection management. The application takes advantage of eye-gaze data to augment interaction with photo collections and aid photowork tasks such as browsing, selecting and searching. It builds on the premise that the time a user spends viewing a photo is correlated with the user's personal affinity for that photo.

1 INTRODUCTION

Photowork is a term first used to describe personal information management concerned with photos. It encompasses tasks such as viewing, browsing, selecting, sorting and filtering photos. The way these tasks are handled by users changed notably with the advent of digital photography. First, digital photography increased dramatically the amount of photos an average user takes and, second, digital photography moves photos to new mediums that escape the limitations of paper (the almost only medium for analogue photography). On the one hand, the fact that digital photos are cheap to take eases the decision of the user whether to take a photo or not, but on the other hand this leads to overwhelmingly large photo collections, which make the 'consumption' of photos or photowork more tedious. Some typical examples are the retrieval of a specific image for which we do not know exactly when or where it was taken nor where did we store it (Whittaker et al. [1] report that users fail to find 40 % of photos), or making a small selection of 'best' photos from a recent holiday to show to friends.

Luckily enough, storing photos in digital format also opens up enormous potential for developing methods and algorithms to ease the problems mentioned above. Photowork in the age of digital photography has been tackled in different ways: from clustering to social network analysis, image recognition and photo-meta-data analysis. As none of the available state-of-the art takes into consideration the user's personal relationship with the photos, we explored the possibilities arising from a more personal perspective on photowork in [2] and [3]. With a series of experiments using a tablet application and a tabletop multitouch display we confirmed that the time a user spends viewing a photo, called viewing time, is correlated with his/her affinity for that particular photo. Our notion of the user's affinity for a photo concurs with Chalfen's notion of importance of a photo in the home-mode: "And although artists, art historians, and art critics frequently speak of 'important' and 'valuable' images, we are dealing with a different notion of importance here. In the home-mode, images are indeed important in an intimate context, and these images are valued by small groups of biologically and socially related people" [4].

In the experiments in [2], where a tabletop multitouch display was used, the participants were asked to browse through a set of photos scattered on the display and the viewing time for each photo was stored. Only the photo currently in the application's focus was sharp, while all the others were blurred out as can be seen in Figure 1. This is how the application knew which photo the user was currently looking at. In the experiments conducted in [3], where a tablet application was used, we avoided the problem of determining which photo the user is looking at by allowing the user to browse photos only in full-screen mode. In both cases we could have used an eye-tracking device, but at the time it was unavailable to us.

Figure 1: *Determining which photo the user is looking at without an eye-tracking device by blurring out all photos, except the onei n focus.*

2 DEMO APPLICATION

To test the use of an eye-tracking device in the realm of photo collection management, we created a demo application called 'Implicit Photowork Demo'. It is a simple desktop application with two windows. The first window allows the user to select a folder with the photos he/she wants to browse (Figure 3), while the second window is used to display the 'best' 5 photos according to the user's affinities (Figure 5). While the user looks at the photos, the application updates a list of how many times and for how long the user looked at each photo. These viewing times are then used to determine the user's favourite photos. The application is also able to give graphical feedback about eye-gaze data (Figure 4) and provides eye-gaze calibration and validation procedures.

The workflow for the application is as follows: the user first undergoes the procedures of calibration and, optionally, validation. This is done by looking at 5 specific points on the display. Next, the user loads the photos he/she wants to browse. Currently the application supports the following formats: '.jpg', '.png' or '.gif' and provides some sample photos for testing purposes. Once the photos are displayed in the main window, the user looks at them as he/she would do normally (Figure 3). While looking at the photos, the user is implicitly signalling to the application his/her affinities for the photos. After the user has finished browsing, he/she can press the 'Show results' button, which will cause the results window to pop up. In this window the 'best' 5 photos will be displayed from right to left (Figure 5). Each time the 'Show results' button is pressed, a new window with updated results appears.

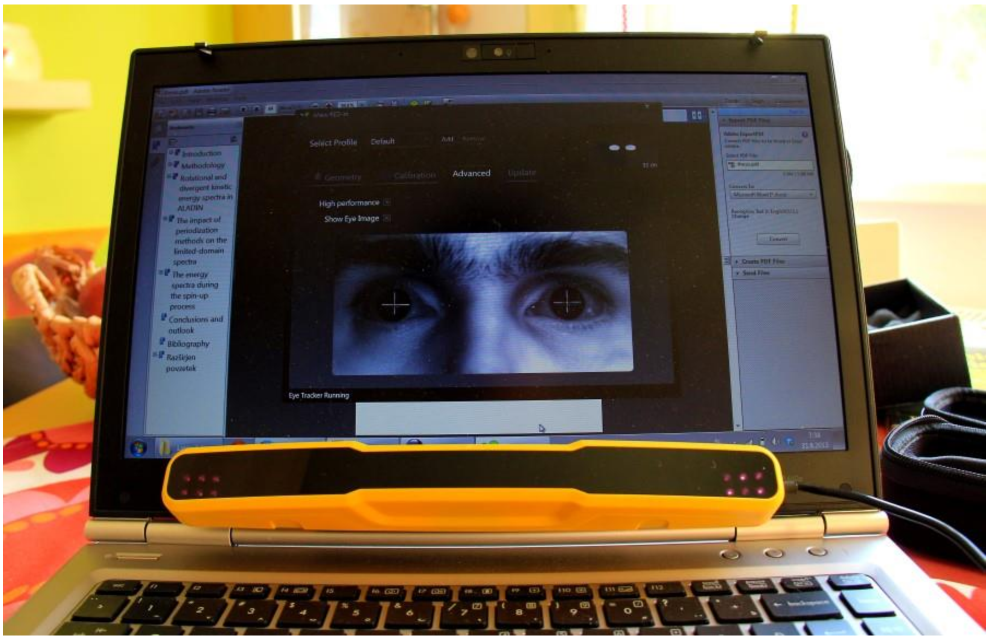

Figure 2: *SMI's RED-m eye tracker and the application used to calibrate it.*

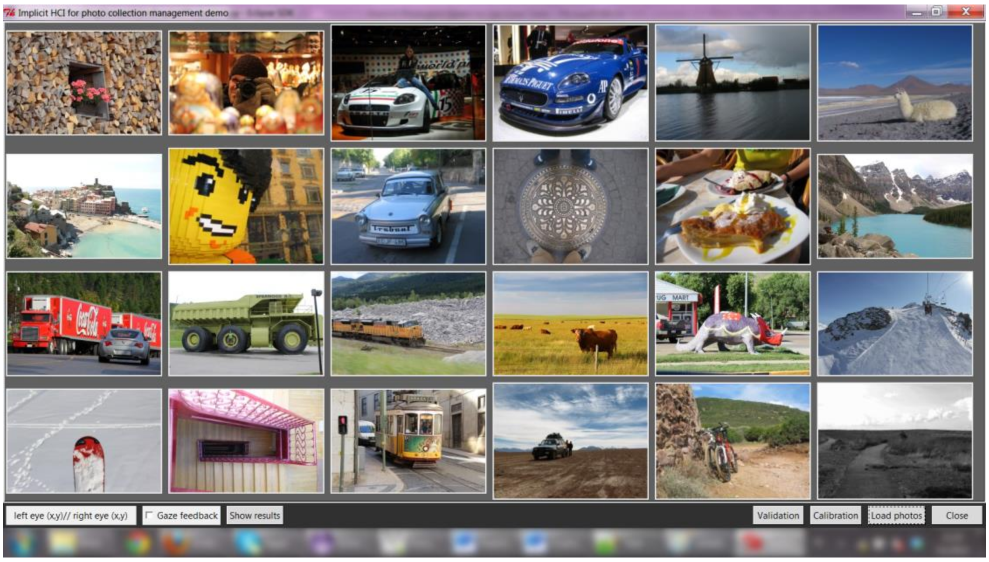

Figure 3: *The demo application's main window filled with test photos for browsing.*

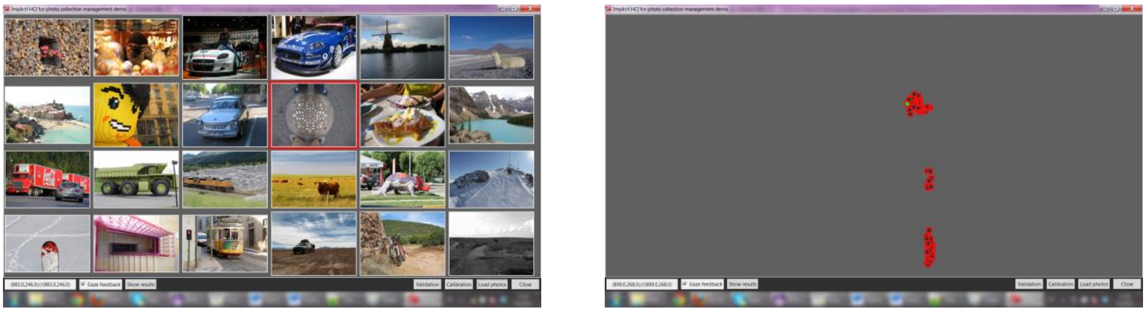

Figure 4: *Graphical feedback for eye-gaze data. In the left image the application highlights with a red frame the photo the user is currently looking at, and in the right image the application marks the current gaze with a green dot, while grey dots mark previous eye-gaze data points.*

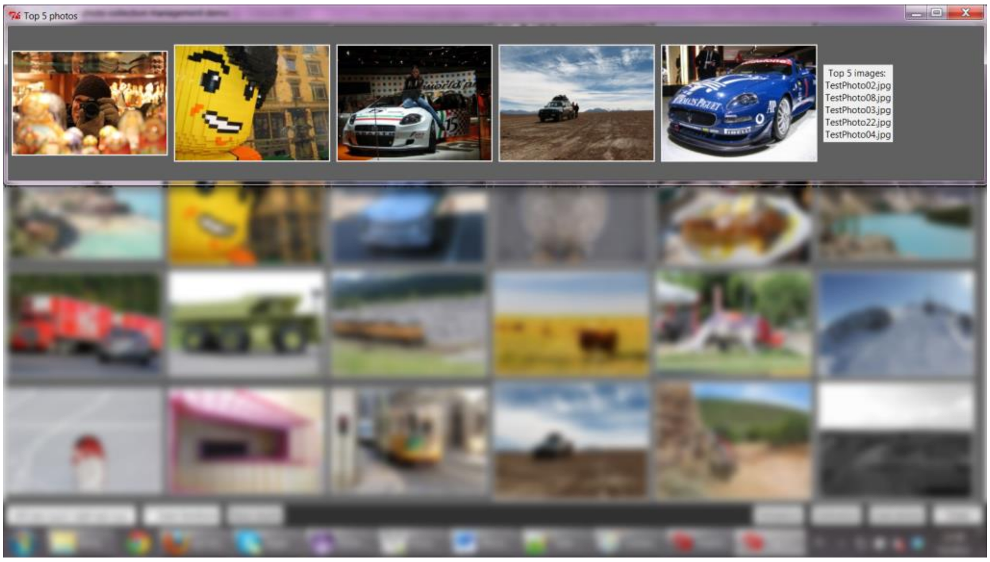

Figure 5*: The 'best' 5 photos according to user affinity are displayed in the Results window.*

3 IMPLEMENTATION

The device we used was SMI's RED-m eye-tracker[1]. This is a portable eye-tracker suitable for monitors, tablets and laptops from 10 to 22 inches in size at an operating distance of 50 cm – 70 cm. The sampling rate can be set to 60 or 120 Hz. The accuracy of the device is 0.5°.

The demo is implemented in Python and builds upon the 'DataStreaming' example from SMI's SDK. For the GUI components, Tkinter[2] was used. The only other requirement for the demo application is the PIL[3] package that was used for handling photos.

[1] www.smivision.com/en/gaze-and-eye-tracking-systems/products/redm.html
[2] http://docs.python.org/2/library/tkinter.html
[3] http://www.pythonware.com/products/pil/

5 DISCUSSION & FUTURE WORK

The lowering prices of eye-tracking devices and the inclusion of eye-tracking features in mass-production commercial devices like Samsung's flagship phone S4 or prototypes like Tobii-Lenovo's concept eye-tracking laptop[4] suggest a shift of focus in terms of eye-tracking research. From researching eye-tracking itself, the hardware and algorithms needed to support it, and the use of eye-tracking in laboratory conditions for usability studies or psychometric evaluations to researching eye-tracking as a means of human-computer interaction.

Eye-tracking as an interaction method or user interface can be classified under the term natural user interface. It has been argued that what makes natural user interfaces *natural*, is the way that this interfaces reuse skills [5] that users already posses and thus feel natural to use or make the

[4] http://www.tobii.com/en/gaze-interaction/global/demo-room/concept-laptop/

user feel like a ‘natural’ [6]. Interaction based on eye-tracking fits to this point of view perfectly as the skill it reuses – human vision – is one of the basic skills we have and predominantly rely on. Besides the straightforward exploitation of the exact information where the user is looking and what he/she is looking at, eye-gaze interaction can also take advantage of other eye-gaze movements like saccades (rapid eye movements used when scanning a visual scene), fixations (gazing at a single location) and smooth pursuit (continuous eye movement when following a moving object). The question here is not only how to interpret these eye-gaze events, but also how should an application respond, how should the response differ in relation to the context in which the application is used, is there an eye-gaze-based interaction paradigm that can be used by every user or should this kind of interaction be tailored to each individual user and so on.

We consider the presented application as a first step towards understanding the above questions and finding appropriate answers (at least for the context considered, i.e. photowork). Preliminary experiments with the Implicit photowork demo application confirmed that eye-tracking is indeed a viable way towards implicit human-computer interaction in photo collection management and that it could be used as envisioned in [2, 3]. These experiments were not strict and a more rigourous evaluation is imperative as well as comparative study on how implicit HCI for photowork differs based on whether it taks place on a tablet or on acomputer equipped with an eye-tracker. Currently, the implemented functionality only considers where the user is looking at and how long he/she spends viewing a particular location/photo. In the future, the application should also consider eye-gaze movements, which will enable to advance the research on how to measure a user’s affinity for a photo described in [2, 3]. For example, quick erratic exchanging of saccades and fixations could indicate that the user is searching for a photo, while a more steady sequence could be interpreted as a sign of the user browsing through the collection. Based on this information, the application could adapt its behaviour like displaying photos in lower resolution in order to be more responsive while searching for a photo.

6 CONCLUSION

The Implicit photowork demo application presented shows how eye tracking and eye-gaze data can be exploited to implement implicit human-computer interaction in order to aid photo collection management. The idea of using viewing time as a measure for the user’s affinity for photos – already successfully tested in [3, 4] – can now be extended with eye-gaze data made available by eye-tracking devices, which opens up new research questions, for example: how to handle eye-gaze events (e.g. fixation events) in general as well as in the context of photo collection management. The availability of commercial mobile devices (e.g. Samsung Galaxy S4) with rudimentary eye-based interaction and the identification of tablets as devices with great photowork potential [7] testify of the importance of research in eye-gaze based implicit human-computer interaction for personal photo collection management.

Acknowledgements

The author would like to thank the company SensoMotoric Instruments (SMI) for landing their Red-M device and thus granting him hands-on experience with eye-tracking devices. This research was funded in part by the European Union, European Social Fund, operational program for human resources, development for the period 2007-2013 and by the Slovenian Research Agency.

References

[1] Whittaker, S., Bergman, O. and Clough, P. (2010). Easy on that trigger dad: a study of long term family photo retrieval. Personal Ubiquitous Computing. 14. 1 (January 2010). 31-43. DOI=10.1007/s00779-009-0218-7 http://dx.doi.org/10.1007/s00779-009-0218-7

[2] Blažica, B., Vladušič, D., & Mladenić, D. (2011). Shoebox: a natural way of organizing pictures according to user’s affinities. In Human-Computer Interaction. Towards Mobile and Intelligent Interaction Environments (pp. 519-524). Springer. http://link.springer.com/content/pdf/10.1007%2F978-3-642-21616-9_58.pdf#

[3] Blažica, B., Vladušič, D., & Mladenić, D. A personal perspective on photowork: implicit human–computer interaction for photo collection management. Personal and Ubiquitous Computing, 1-9. http://link.springer.com/article/10.1007%2Fs00779-013-0650-6

[4] Chalfen, R. (1975). Introduction fo the Study of Non-Professional Photography as Visual Communication.

[5] Buxton, B. Making User Interfaces Natural (2010). http://research.microsoft.com/enus/about/feature/nui-video.aspx. Online, accessed: September 2013.

[6] Wigdor, D.. Wixon, D. (2011). Brave NUI world: Designing Natural User Interfaces for Touch and Gesture (Morgan Kaufmann, Burlington).

[7] Kirk, D., Sellen, A., Rother, C. and Wood, K. (2006). Understanding photowork. Proceedings of the SIGCHI conference on Human Factors in computing systems. Held 2006 in Montreal, Quebec, Canada. 761-770. New York: ACM Press.