Robust Marker Tracking System for Mapping Mobile Eye Tracking Data

Iyad Aldaqre
SR LABS Srl
Milan, Italy
iyad.aldaqre@srlabs.it

Roberto Delfiore
SR LABS Srl
Milan, Italy
roberto.delfiore@srlabs.it

ABSTRACT

One of the challenges of mobile eye tracking is mapping gaze data on a reference image of the stimulus. Here we present a marker-tracking system that relies on the scene-video, recorded by eye tracking glasses, to recognize and track markers and map gaze data on the reference image. Due to the simple nature of the markers employed, the current system works with low-quality videos and at long distances from the stimulus, allowing the use of mobile eye tracking in new situations.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**; Human computer interaction (HCI);

KEYWORDS

marker tracking, wearable eye tracking, gaze mapping

ACM Reference Format:
Iyad Aldaqre and Roberto Delfiore. 2018. Robust Marker Tracking System for Mapping Mobile Eye Tracking Data. In *ETRA '18: 2018 Symposium on Eye Tracking Research and Applications, June 14–17, 2018, Warsaw, Poland.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3204493.3208339

1 INTRODUCTION

Wearable eye tracking has provided the possibility for studying gaze behavior in real-life environments. Most eye tracking glasses record the scene participants look at in addition to their eye movements, and then map their gaze data on the scene video. However, in order to aggregate gaze data from a group of participants, the data needs to be mapped to a reference image of the stimulus.

For that we can use object recognition for a reference image in the scene video [De Beugher et al. 2014; Ye et al. 2012]. From our experience in the field of market research we learned that this technique does not perform very well in complex scenes, such as real supermarket shelves, and requires a lot of manual intervention. Additionally, when participants are allowed to move freely and interact with the stimulus, it becomes almost impossible to reproduce the exact scene for all sessions, which makes object recognition more difficult.

ETRA '18, June 14–17, 2018, Warsaw, Poland

ACM ISBN 978-1-4503-5706-7/18/06.
https://doi.org/10.1145/3204493.3208339

Another technology we can use to map gaze data on a reference image is marker tracking, similar to that used in Pupil Labs systems [Kassner et al. 2014]. This is a much more reliable technique, but we also encountered some limitations with it: Participants need to stay relatively near from the stimulus for markers to be recognized and, with low quality video, identifying typical surface-tracking markers becomes very difficult.

2 DESCRIPTION

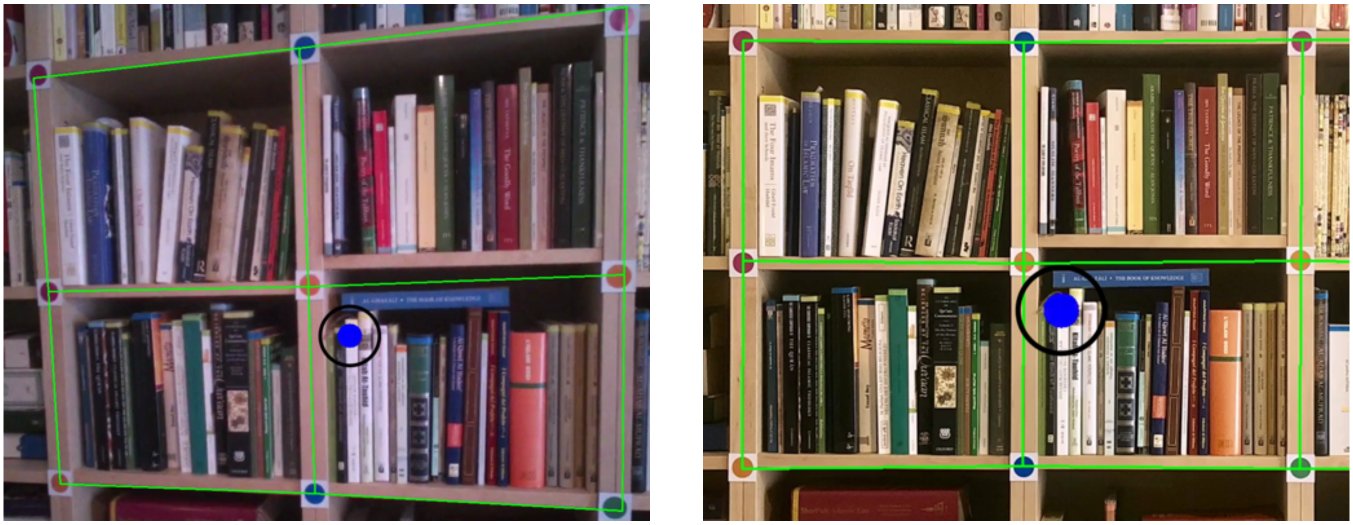
Taking the conditions presented above into account, we developed a system comprising of simple colored markers, to be placed on the stimulus, and a *patent pending* algorithm that recognizes and tracks those markers. Rectangular cells are then identified by the color combination of the markers that form them (see Figure 1 for example), allowing the system to recognize each cell and its position in the scene video. Then, the algorithm can then recognize the stimulus surface and its distance from the participant and its orientation. Finally, gaze data can be mapped on a reference image of the stimulus using homographic transformation [Hartley and Zisserman 2003], allowing for data quantification and aggregation.

Except for adding markers to the stimulus, recording sessions can proceed normally and without any modification to the procedure. At the end of data collection, all recordings can be processed automatically and no manual intervention is needed. After processing is done, gaze data mapped to the reference image coordinates can be exported, alongside the original data in scene-camera coordinates, making it possible to calculate aggregated metrics and visualize the data.

This system works with low-quality videos, it is resistant to changes in the scene and it works at much further distance from the stimulus compared to other technologies (exceeding 5 meters, depending on the size of the markers). The current system is also hardware independent, making it usable with any eye tracking glasses that provide a scene-video and the gaze data mapped to the scene-video coordinates.

We have tested the system's accuracy, by collecting gaze data using Tobii Pro Glasses 2 (Tobii Technology, Sweden) while participants looked at a circular target from different distances (100, 200 and 300 cm). Results showed that the system maintains the accuracy of the eye tracking glasses after mapping the data to the reference image (see Figure 2).

We are currently testing the system in a supermarket laboratory to measure the effect of the markers on participants' attention. Preliminary qualitative analysis, by reviewing sample videos, showed that participants do not get distracted by the markers in a natural situation and their attention remains mainly to the actual stimulus.



(a) Video Snapshot

(b) Reference Image

Figure 1: Example snapshot image from the scene video recorded by eye tracking glasses (a) and a portion of the reference image used to map gaze data onto (b). Visible in both images the gaze point and the markers used to recognize the stimulus surface and 4 cells formed by the markers.

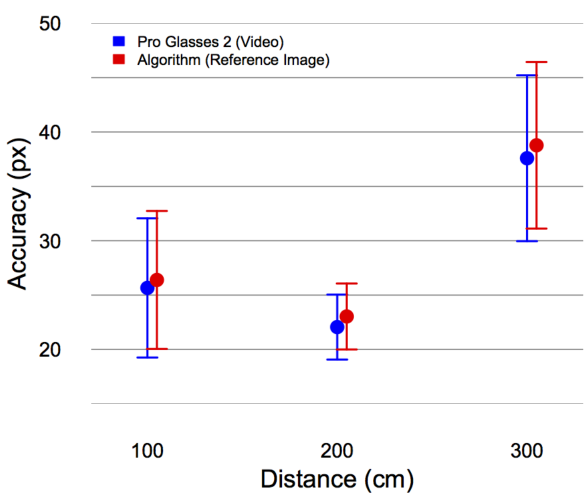


Figure 2: Accuracy data for the eye tracking glasses (blue) and the reference image-mapped data (red) in pixels at different distances from the stimulus. Error bars represent standard error of the mean.

3 USE CASES

The system can be used for studying gaze behavior in natural situations, when placing markers on the stimulus surface is feasible. Multiple stimuli can be mapped simultaneously, even if they are presented in different locations. In a similar manner, when the stimulus is three-dimensional, markers can be added to different sides of the stimulus and gaze data can be mapped on a separate reference image for each side.

Applications of this system extend beyond market research, including human-computer interaction, ergonomics, usability testing and others. Moreover, due to its fully automated procedure, it is ideal for studies with large samples that are very time consuming with traditional methods.

4 CONCLUSION AND FUTURE WORK

The current system provides a good solution for mapping gaze data collected with eye tracking glasses onto a reference image. The innovative part of this system lies in it’s robustness at long distances and with low-quality videos obtained from state-of-the-art eye tracking glasses. It also provides access to new variables, like participants’ presence in front of the stimulus, their distance and their orientation, which can be used to develop new metrics or to filter subsets of the data.

When the algorithm is optimized to work in real time, it will open the possibility to embed markers in computer programs or on other machines to allow for live human-computer or human-machine interaction through gaze using wearable eye tracking [Bulling and Gellersen 2010]. For example, it could be used to achieve joint attention with robots to study social interaction [Kompatsiari et al. 2018]. Other applications are still to be explored.

ACKNOWLEDGMENTS

All rights of the system presented belong to SR LABS Srl. We would like to thank IPSOS Italy for their collaboration during the testing of the system in their Shopper Lab facility in Milan. We would also like to thank all the participants who took part in the testing.

REFERENCES

Andreas Bulling and Hans Gellersen. 2010. Toward Mobile Eye-Based Human-Computer Interaction. *IEEE Pervasive Computing* 9, 4 (Oct. 2010), 8–12. https://doi.org/10.1109/MPRV.2010.86

Stijn De Beugher, Geert Brône, and Toon Goedemé. 2014. Automatic analysis of in-the-wild mobile eye-tracking experiments using object, face and person detection. In *Computer Vision Theory and Applications (VISAPP), 2014 International Conference on*, Vol. 1. IEEE, 625–633.

Richard Hartley and Andrew Zisserman. 2003. *Multiple view geometry in computer vision* (2 ed.). Cambridge university press.

Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 1151–1160. https://doi.org/10.1145/2638728.2641695

Kyveli Kompatsiari, Francesca Ciardo, Vadim Tikhanoff, Giorgio Metta, and Agnieszka Wykowska. 2018. Bidding for joint attention: On the role of eye contact in gaze cueing. (Apr 2018). https://doi.org/10.17605/OSF.IO/MX28G

Zhefan Ye, Yin Li, Alireza Fathi, Yi Han, Agata Rozga, Gregory D. Abowd, and James M. Rehg. 2012. Detecting Eye Contact Using Wearable Eye-tracking Glasses. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, New York, NY, USA, 699–704. https://doi.org/10.1145/2370216.2370368