Possibilities of Combining AI and Camera-Based Eye Tracking to Guide Geoexploration

M. Keskin[1*], P. Kettunen[1]

[1]Finnish Geospatial Research Institute, National Land Survey of Finland (FGI/NLS), Geodeetinrinne 2. FI-02430 Masala, Finland
Email: merve.keskin@nls.fi, pyry.kettunen@nls.fi

Abstract

To support spatial cognition of map users and assist exploratory reading of geovisualizations (*abbr.* geoexploration) and information extraction, it is important to find new ways to develop systems that incorporate advanced methods, such as deep learning, and are based on the empirical data on human visuospatial behaviour such as eye movement data. This paper discusses the approaches of mimicking eye trackers by using standard (web or smart phone) cameras by combining AI methods, more specifically deep learning algorithms. In addition, it envisions the possibilities of how AI can be incorporated to produce efficient, and guided visual exploration and simultaneous analysis to comprehend spatial phenomena presented via geovisualizations.

1. Introduction

Human-generated data has gained huge importance in geosciences, as there are many tools to collect such data. Eye tracking is one of the most well-established quantitative methods used in experimental cartography for acquiring data from human participants as well as for obtaining insights into their spatial cognition, abilities, and limitations. Nowadays, many web or smartphone applications are being personalized based on user data or preferences to serve more personalized demands. Similarly, cartographic products can also be customized through leveraging their users' behavioural data (e.g., for map use context; Reichenbacher, 2007; for transferable map design; Griffin et al., 2017). Janowicz et al. outline GeoAI (i.e. geospatial AI) methods for geographic knowledge discovery, spatially explicit model production, question answering, and social sensing (for detailed definitions of GeoAI please read Song (2021)). In addition, developing systems based on the empirical data on human visuospatial cognition like eye movement data sets requirements for the visualization methods used. Therefore, we need comprehensive knowledge and capabilities to apply geovisualization methods that are based on data collected from human behaviour. In this context, it is important to create new solutions to enhance users' performance in an exploratory reading of geovisualizations (*abbr.* geoexploration), *which involves maps created for relatively unknown geographic data* (Kraak, 2008). For instance, emerging methods, such as AI and more specifically deep learning, can aid geoexploration. AI can be utilized at several stages of the geoexploration process. On the one hand, AI methods are promising in terms of helping to better understand spatial phenomena with unstructured patterns or unclear relationships and to allow for more efficient visual exploration and analysis (e.g., Koua and Kraak, 2004). On the other hand, AI can be used to generate personalized map visualizations when combined with eye tracking data as an input revealing users' attention-related behaviour. In this context, map visualizations can be partially or fully automated with AI methods. This can be done in a real-time or semi-real time manner so that the mapping system learns and adapts to the gaze behaviours of map users. The system may provide personalized visualizations based on the collected eye movement data. Nevertheless, AI can also be used to replicate eye trackers using standard cameras (e.g., Krafka et al., 2016) and can be combined with face recognition to estimate the gaze positions accurately. The objectives of this paper are to investigate the AI methods to mimic eye-trackers, analyse the obtained data, and then adjust the map design according to the user (i.e., map personalization).

In the next chapter, we first focus on the following key research question and review what is available in the literature: How can eye tracking and AI be combined for creating interactive geovisualizations? What are the recent developments in GeoAI? Regarding this research question, in the third chapter, we discuss the possibilities of applying AI for geoexploration and visual analytics and of building a GeoAI system that is aided by its users' eye movements to bring out the diverse aspects of the geospatial data.

2. Background Research

The answer to our research question is in two interrelated folds: (i) the use of eye tracking and AI related to assessing the usability of cartographic products, (ii) the use of eye tracking and AI related to the map stimuli production. The outcomes of user studies with eye tracking help understanding user behaviours and preferences while using maps and accordingly, setting up design guidelines for maps. These guidelines could be used in automatic map stimuli generation or automated preparation of the visualization whose processes could be enriched by using AI methods (e.g., Griffin et al., 2017). After having collected sufficient amounts of eye movement data to train an AI system, user behaviours can be assessed simultaneously and revisualization steps can be handled in-situ or semi-real time. Since usability is a cycle, the usability of these automatically generated maps can be evaluated again through eye tracking measurements (e.g., Roth et al., 2015; 2017).

External eye trackers have been widely used in several disciplines as they provide an accurate and easy-to-use tracking experience. However, an eye tracker does not have to be an additional purchased device but it can be a software or system built essentially on AI. The recent trend is not to use external devices because they are expensive, sophisticated, and require specialized software and expertise for data recording, processing, analyses, and visualization. Moreover, external eye trackers restrict usability studies in many ways, *e.g.*, by expensive and time-consuming user experiments and strict conditions and requirements to collect accurate and usable eye movements. Such strict conditions are avoiding daylight, chin rest to restrict head movements/noise, and the need for the physical presence of participants in the lab environment. Measuring gaze location can be done through the integration of different methods, *e.g.*, machine learning for feature extraction, face recognition, 3D gaze estimation models, *etc*. Once the appropriate methodology is identified and a sufficient amount of user data is collected, eye tracking without eye trackers can be applied with any kind of visual stimuli, including maps (Murali & Çöltekin, 2021; Krassanakis & Cybulski, 2021).

It is worth mentioning some of the methods that have been developed to mimic eye trackers. The first one is using predictive eye tracking which is an AI-based alternative to regular eye tracking studies and relies on pre-collected eye tracking data to simulate human vision. It uses complex deep learning algorithms that are trained on tens of thousands of images produced in actual eye tracking experiments and analyses the images by recognizing and systemizing particular patterns in user attention flow (URL 1). Another option is building eye trackers based on AI either by using webcam or smartphone cameras. There are several methods to do so, one of which is smartphone-based eye-tracking systems that use visible light. Krafka *et al.* (2016) developed such system called GazeCapture leveraging the crowd-sourced data and front-facing camera to estimate the point of gaze. The system uses a deep convolutional neural network (CNN), which was trained with tens of thousands of labelled images from 1500 participants to process images of the subject's face and the point of gaze estimate (*gaze estimation bias: > 3°, accuracy: a centimetre*). ScreenGlint (Huang *et al.* 2017) and the system developed by

Valliappan *et al.* (2020) could be listed as other examples that make eye tracking possible with visible light. Infra-red light sources and sensitive cameras are available in some commercial smartphones to enhance the face tracking of authentication systems. Brousseau and Eizenman (2020) introduced a smartphone-based system that is insensitive to motion between the device and the user's head by utilizing CNN feature extraction and an infrared 3D gaze-estimation model (*gaze estimation bias: 1◦ between 0.4◦ to 2.1◦*). It can be summarized that several AI methodologies can be implemented for building eye trackers using standard cameras instead of specialized alternatives with competing accuracies. In the next chapter, we discuss the use of AI to extract information from geospatial data based on eye movement characteristics.

3. AI-assisted geoexploration using eye movement data

We envision a system that assists its users to understand complex spatial phenomena and unclear relationships between features presented within exploratory geovisualizations. For AI-assisted geoexploration, the use of map design elements or visual parameters and the use of machine learning tools for spatial data are the two important aspects. To be able to achieve a rather ambitious goal of controlling map displays with gaze, in other words, modifying the map visualization with eye movements, the characteristics of geospatial data and eye movement behaviour should be combined. In this context, the first phase is understanding the eye movement characteristics while users explore the geospatial data. This can be done by collecting large amounts of eye movement data to train the computer by following a data-driven approach.

The second phase is modelling the visual interaction and possible enhancement in the geovisualization. It is important to decide the type of map data, *e.g.*, vector or raster, and to accordingly choose what visual scenario is presented to the users, such as vector spaghetti or an unannotated map. Designing the interaction requires comprehensive knowledge on the use of eye tracking for cartographic usability research and the use of eye tracking for interaction, in our case, controlling map displays with eye movements. This topic (i.e., gaze-based interaction) has not yet been sufficiently researched in the cartographic domain (e.g. Rudi et al., 2020). However, although not scientifically, interacting with visual displays through eye tracking has been utilized in the gaming industry and there are solutions provided by eye tracking companies such as Tobii (URL 2).

The third phase is building interaction for a use-case of guiding geoexploration to discover the patterns and relationships in geospatial data. Possible interactions might be, for instance, highlighting or exaggerating map features within the areas receiving the highest fixation durations and also other map features having similar characteristics throughout the whole map. This process can be executed in real-time or semi-real time by calculating heat maps instantaneously after letting the users explore the map for a certain period and collecting eye movement data meanwhile. Therefore, it can be possible to study the influence of highlighting the features, which are easily seen on maps, on map users' performance and whether highlighting helps to understand the unknown/unseen map features. It leads us to rethink map design considering individual behaviour patterns and possibly employ machine learning methods to continuously evaluate user performance in real-time and adjust the design based on users' performance, geoexploration patterns, and choices (Richter, Tomko and Çöltekin 2015).

4. Conclusion

Cartographic products can indeed be personalized by leveraging users' behavioural data and incorporating AI methods. AI is mostly used for understanding unstructured or unclear spatial phenomena, therefore, for guiding more efficient visual exploration and analysis. It can be combined with eye tracking to generate personalized map visualizations in real-time, in other words, gaze can be utilized as a tool to interact with maps. Furthermore, AI, more specifically deep learning, can be used to mimic eye trackers through web-cams or front-facing cameras of smartphones. Within this paper, we tried to summarize the AI methods for camera-based eye tracking and how eye movement data can serve as input for personalized map visualizations generated with AI.

Traditional commercial eye trackers might be more reliable or provide more accurate outcomes for scientific research but the alternative eye tracking solutions mentioned in this paper are quite competitive in terms of accuracy. On the other hand, if not portable or light, traditional eye-trackers restrict researchers to conduct user studies only in a lab environment which is more ideal in terms of being more controlled and stable. However, it has downsides such as long experiment time, money spent, and recruiting only a limited number of participants. If we employ methodologies that convert standard cameras, which are already available on our smartphones or laptops, into eye tracking devices, it would be possible to conduct online experiments and therefore, collect large volumes of user data. Leveraging the collected eye movement data from online experiments (i.e. using it as a training dataset), we can build prediction models via deep learning algorithms and compare the prediction and actual behaviour of users derived from eye movements. Accordingly, we can design personalized maps without the need of collecting additional eye movement data or re-design the existing maps based on those model. In the future, gaze-based interaction with maps can become possible and we can combine it with other AI technologies such as speech recognition. For example, a user may ask questions about a region (s)he is looking at and the system provides an audio answer, a table, or in other kinds of visual formats.

Acknowledgments

This research is funded by the Finnish Scientific Advisory Board for Defence (MATINE) through the TUGEVA project.

References

Brousseau, B., Rose, J., & Eizenman, M. (2020). Hybrid eye-tracking on a smartphone with CNN feature extraction and an infrared 3d model. Sensors, 20(2), 543.

Griffin, A. L., White, T., Fish, C., Tomio, B., Huang, H., Sluter, C. R., ... & Picanço, P. (2017). Designing across map use contexts: A research agenda. International Journal of Cartography, 3(sup1), 90-114.

Huang, M. X., Li, J., Ngai, G., & Leong, H. V. (2017, May). Screenglint: Practical, in-situ gaze estimation on smartphones. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 2546-2557).

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2020). GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond.

Krassanakis, V., & Cybulski, P. (2021). Eye Tracking Research in Cartography: Looking into the Future. ISPRS International Journal of Geo-Information, 10(6), 411.

Koua, E. L., & Kraak, M. J. (2004). Geovisualization to support the exploration of large health and demographic survey data. International journal of health geographics, 3(1), 1-13.

Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., & Torralba, A. (2016). Eye tracking for everyone. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2176-2184).

Murali, M., & Çöltekin, A. (2021, May). Conducting eye tracking studies online. In Proceedings of the Workshop on Adaptable Research Methods for Empirical Research with Map Users, Virtual Workshop (Vol. 6).

Reichenbacher, T. (2007). Adaptation in mobile and ubiquitous cartography. In Multimedia Cartography (pp. 383-397). Springer, Berlin, Heidelberg.

Richter, K.-F., Tomko, M., & Çöltekin, A. (2015). Are we there yet? Spatial cognitive engineering for situated Human-Computer Interaction. In CESIP 2015: Cognitive engineering for spatial information processes: From user interfaces to model-driven design. Workshop at COSIT 2015, October 12, 2015. Santa Fe, NM, USA.

Rudi, D., Kiefer, P., Giannopoulos, I., & Raubal, M. (2020). Gaze-based interactions in the cockpit of the future: a survey. Journal on Multimodal User Interfaces, 14(1), 25-48.

Valliappan, N., Dai, N., Steinberg, E., He, J., Rogers, K., Ramachandran, V., ... & Navalpakkam, V. (2020). Accelerating eye movement research via accurate and affordable smartphone eye tracking. Nature communications, 11(1), 1-12.

Song Gao (2021) Geospatial Artificial Intelligence (GeoAI). Oxford Bibliographies in Geography. 1-16. DOI: 10.1093/OBO/9780199874002-0228.

URL 1: Attention Insight, available at https://attentioninsight.com/eye-tracking-vs-predictive-eye-tracking Last accessed on 30.08.2021.

URL 2: Tobii Concept Validation Tool, available at https://tech.tobii.com/get-started/ Last accessed on 21.07.2021.