

High-Level Gaze Metrics From Map Viewing

Charting Ambient/Focal Visual Attention

Krzysztof Krejtz¹, Andrew T. Duchowski², and Arzu Çöltekin³

¹ National Information Processing Institute and

University of Social Sciences and Humanities, Warsaw, Poland

² School of Computing, Clemson University, Clemson, SC, USA

³ Department of Geography, University of Zurich, Zurich, Switzerland

Abstract. Distinguishing ambient and focal attention, we demonstrate the use of the \mathcal{K} coefficient, which can serve as a cue for recommender systems in deciding when to offer information to the user, e.g., when focally attending during search.

Keywords: maps, eye tracking, visual attention

1 Introduction & Background

Gaze-based recommender systems are designed to respond with information contingent on the viewer's gaze, e.g., in geographic contexts, when directed to a particular location in physical or virtual space (such as on a map). These geographic gaze-based recommender systems have been referred to as location-aware (mobile) eye tracking systems [7]. For the system to provide an appropriate response, analysis of the viewer's gaze is required to infer the viewer's desire for information. Generally, this is accomplished via computation of an interest metric [11,10,4]. Recent approaches characterize interest or boredom via Support Vector Machines [6] or Area Of Interest (AOI) revisitation [5].

In this paper we use the \mathcal{K} coefficient to distinguish ambient and focal attention. Ambient attention is characterized by relatively short fixations followed by high amplitude saccades. Conversely, focal attention is described by long fixations followed by saccades of low amplitude [12]. The \mathcal{K} coefficient captures the temporal relation between standardized fixation duration and subsequent saccade amplitude. $\mathcal{K} > 0$ indicates focal viewing while $\mathcal{K} < 0$ suggests ambient viewing [9].

Attention becoming more focal over time, or oscillating between focal and ambient modes, could indicate changing cognitive load corresponding to stimulus complexity. The \mathcal{K} coefficient can thus potentially act as a contextual cue which could be exploited by recommender systems, e.g., do not interrupt the user when in ambient search mode, or oscillating between ambient and focal search.

We demonstrate the utility of \mathcal{K} with an experiment comparing visual search behavior over different geographic representations of two cities: Amsterdam and Barcelona.

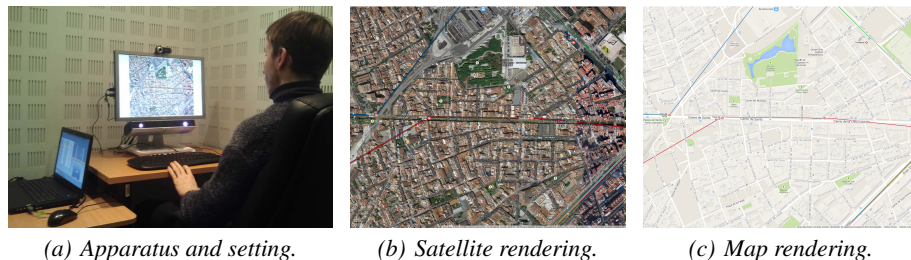


Fig. 1. Apparatus and stimuli (Barcelona) using Google's hybrid and roadmap rendering.

2 Methodology

The experiment consisted of localizing map objects (features), arguably *the* most fundamental map-reading task [1]. Localization, or visual search, i.e., finding an object of interest among others, is a fundamental map use task, because regardless of the map type or the final goal of the map reader, an object or area of interest must be found before it can be studied further. Geographic task taxonomies widely acknowledge this task among the most basic and common [8,2]. The localization tasks were carried out when viewing a city representation in either (cartographic) map or satellite view (Google's roadmap or hybrid rendering; see Fig. 1 and text below for technical details). Unlike previous work on route learning [3], we hypothesized there would be an effect of map type on visual search, with localization (search) taking longer in satellite view than in map view.

Participants. Sixty-three (63) university students took part in the study, with 7 excluded due to technical and procedural problems (e.g., poor calibration). The final sample included 56 participants (19 M, 37 F, ages $M=24.62$, $SD=4.01$).

Apparatus. All stimuli were presented on a computer monitor (1680×1050 resolution; 22" LCD, 60 Hz refresh rate) connected to a standard PC computer. Eye movements were recorded at 250 Hz with an SMI RED 250 eye tracking system.

Stimulus. Map images were created using *Google Maps™ JavaScript API v3*. Maps were rendered (see Fig. 1) by disabling visual user interface controls for navigation, scale, rotate, pan, and zoom, and limiting the number of Points Of Interest (POIs) to two. Specifically, the Barcelona maps (using Google's latitude/longitude coordinates: 41.375384, 2.141004) displayed only the "park" and "sports_complex" POIs at zoom level 17 with the transit layer turned on. The Amsterdam maps (lat./lon.: 52.3614777, 4.8837351) were also displayed at zoom level 17 with two POIs ("attraction" and "business") and transit layer turned on. All maps were rendered to 1280×1024 resolution, then screen-captured and cropped to the same dimensions.

Procedure. Participants were randomly assigned to either experimental condition, map ($N=29$) or satellite view ($N=27$). Prior to viewing of the map, the eye tracking system was calibrated to each individual. Following calibration, participants carried out the localization task (after having located a start point).

For the city of Barcelona (see Fig. 1), participants were asked to find the intersection of *Carrer de Vilardell* and *Carrer d'Hostafrancs de Sió*, indicating its localization by dwelling on it for 3 seconds. For Amsterdam, they were told to find the *Rijksmuseum*.

3 Results & Discussion

Participants' Familiarity with Presented Areas and Google Maps. Participants were not very familiar with the areas of Barcelona or Amsterdam that were presented to them during the experiment. We drew this conclusion after analyzing questionnaire answers. The percentage of participants who had visited either city at least once in their lives was 39% for Amsterdam and 25% for Barcelona. Of these, they had visited the city at most one time (32.1% for Amsterdam and 19.6% for Barcelona).

Google maps was fairly popular among participants, with only 7.1% claiming they had never used the service. Table 1 gives the detailed distribution of answers to the question "How often do you use Google Maps?"

Performance (time to completion) and process (visual attention) measures we report from the experiment therefore speak to more or less typical use of Google Maps. The importance of our findings lies in their potential indication of online map use when encountering a new location (e.g., as one would do prior to travel to the destination) by users already familiar with the service.

Localization Task Performance. All subjects successfully completed the tasks. To gauge localization performance, we first considered basic eye movement metrics, using a linear mixed model (LMM) analysis with task duration as the dependent variable and experimental condition (map vs. satellite) as the between-subjects fixed factor and city (Amsterdam vs. Barcelona) as the within-subject fixed factor.

As expected, analysis revealed a significant main effect of experimental condition, $F(2, 54) = 26.26$, $p < 0.01$, indicating that task completion time was shorter for the map view ($M = 37.33$ s, $SE = 41.63$) than for the satellite view ($M = 47.36$ s, $SE = 55.27$).

Unexpectedly, however, the city appeared to be a moderating factor resulting in a significant interaction term, $F(1, 54) = 9.15$, $p < 0.01$. Investigation of contrasts related to this interaction effect revealed that the difference between map and satellite views was significant only for Barcelona. For Amsterdam, no significant difference between completion times for the two view types was found. Analysis also revealed a main effect of city, $F(1, 54) = 31.68$, $p < 0.01$, meaning that, overall, time to localization on either of

Answer	Percent
Everyday	7.1%
2-3 times a week	32.1%
Once a week	12.5%
2-3 times a month	19.6%
Once a month	21.4%
Never	7.1%

Table 1. Answers to the questionnaire question: "How often do you use Google Maps?" ($N = 56$)

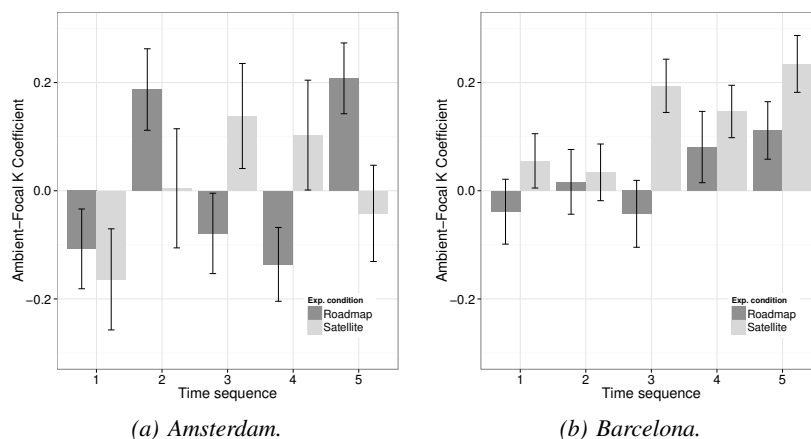


Fig. 2. Dynamics of \mathcal{K} coefficient over time between viewing conditions (whiskers $\pm 1SE$).

the Barcelona maps (map or satellite views) took significantly longer ($M = 61.7$ s, $SE = 1.9$) than on the Amsterdam maps ($M = 25.44$ s, $SE = 1.5$).

The studied sections of the two cities show strong differences in spatial organization, that is, the way the buildings and other features are organized in relation to each other. In the section from the Barcelona map there is more regularity (more rows of buildings similar to each other), making the visual search potentially harder. This is possibly one of the reasons why we found a strong difference between the cities and one that is also supported by literature [13].

Temporal Dynamics of Visual Attention. To delve deeper into the visual search process of viewers over each of the city maps, we performed a similar LMM analysis but with \mathcal{K} as the dependent variable, with the same independent variables of map view and city as before. Because we believe plotting \mathcal{K} over time depicts the dynamic dilation and constriction of the viewer's field of view during visual search, we used time as a fixed within-subject factor, discretized into 5 equal periods (see Fig. 2).

Considering the city as a moderating factor, analysis of \mathcal{K} showed significant interaction among city, map type, and time, $F(4, 9910) = 2.44$, $p < 0.05$. Temporal ambient/focal attention patterns are influenced by map type (satellite vs. cartographic) and city.

When working with the Barcelona satellite view, attention generally becomes increasingly more focal over time, especially over the satellite view, as expected (see Fig. 2(b)). Over the cartographic map view, the pattern is more sinusoidal, eventually becoming increasingly focal over the last two time segments. When viewing the Amsterdam maps, however, the search pattern is less predictable (see Fig. 2(a)).

4 Conclusion

Analysis of \mathcal{K} strongly suggests different cognitive strategies involved in visual search depending on the stimulus and/or task difficulty. Faster search over less complex stimu-

lus may afford ambient search sooner, longer, or with greater ambient/focal oscillation, possibly indicating boredom. More complex stimulus may require increasingly focal attention later on into the search, possibly indicating greater interest.

References

1. Boér, A., Çöltekin, A., Clarke, K.C.: An Evaluation of Web-based Geovisualizations for Different Levels of Abstraction and Realism—What do users predict? In: International Cartographic Conference. pp. 209–220. Dresden, Germany (2013)
2. Carter, J.R.: The many dimensions of map use. In: Proceedings of the International Cartographic Conference (2005), <http://www.cartesia.org/geodoc/icc2005/pdf/oral/TEMA12/Session3/JAMESCARTER.pdf>
3. Francelet, R.: Realism and Individual Differences in Route-Learning. Master’s thesis, University of Zurich (2014)
4. Hammer, J.H., Maurus, M., Beyerer, J.: Real-time 3d gaze analysis in mobile applications. In: Proceedings of the 2013 Conference on Eye Tracking South Africa. pp. 75–78. ETSA ’13, ACM, New York, NY (2013), <http://doi.acm.org/10.1145/2509315.2509333>
5. Kiefer, P., Giannopoulos, I., Kremer, D., Schlieder, C., Raubal, M.: Starting to get bored: An outdoor eye tracking study of tourists exploring a city panorama. In: Proceedings of the Symposium on Eye Tracking Research and Applications. pp. 315–318. ETRA ’14, ACM, New York, NY (2014), <http://doi.acm.org/10.1145/2578153.2578216>
6. Kiefer, P., Giannopoulos, I., Raubal, M.: Using eye movements to recognize activities on cartographic maps. In: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. pp. 488–491. SIGSPATIAL’13, ACM, New York, NY, USA (2013), <http://doi.acm.org/10.1145/2525314.2525467>
7. Kiefer, P., Straub, F., Raubal, M.: Towards location-aware mobile eye tracking. In: Proceedings of the Symposium on Eye Tracking Research and Applications. pp. 313–316. ETRA ’12, ACM, New York, NY (2012), <http://doi.acm.org/10.1145/2168556.2168624>
8. Knapp, L.: A Task Analysis Approach to the Visualization of Geographic Data. In: Nygeres, T.L., Mark, D.M., Laurini, R., Egenhofer, M.J. (eds.) *Cognitive Aspects of Human Computer Interaction for Geographic Information Systems*, chap. 7, pp. 355–371. Springer (1995)
9. Krejtz, I., Szarkowska, A., Krejtz, K., Walczak, K., Duchowski, A.T.: Audio Description as an Aural Guide of Children’s Visual Attention: Evidence from an Eye-Tracking Study. In: Proceedings of the 2012 Symposium on Eye-Tracking Research and Applications. ETRA ’12, ETRA, ACM, New York, NY (March 28-30 2012)
10. Qvarfordt, P., Zhai, S.: Conversing with the user based on eye-gaze patterns. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 221–230. CHI ’05, ACM, New York, NY, USA (2005), <http://doi.acm.org/10.1145/1054972.1055004>
11. Starker, I., Bolt, R.A.: A gaze-responsive self-disclosing display. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 3–10. CHI ’90, ACM, New York, NY, USA (1990), <http://doi.acm.org/10.1145/97243.97245>
12. Velichkovsky, B.M., Joos, M., Helmert, J.R., Pannasch, S.: Two Visual Systems and Their Eye Movements: Evidence from Static and Dynamic Scene Perception. In: *CogSci 2005: Proceedings of the XXVII Conference of the Cognitive Science Society*. pp. 2283–2288. Stresa, Italy (July 2005)
13. Ware, C.: Chapter 2: What We Can Easily See. In: *Visual Thinking for Design*, pp. 23–42. Morgan Kaufmann (2008)