Location-Aware Mobile Eye-Tracking for the Explanation of Wayfinding Behavior

Peter Kiefer Florian Straub Martin Raubal
Institute of Cartography and Geoinformation
ETH Zürich
Wolfgang-Pauli-Str. 15, 8093 Zurich, Switzerland
{pekiefer, straubf, mraubal}@ethz.ch

Abstract

This paper proposes location-aware mobile eye-tracking as a new method for the explanation of wayfinding behavior. We argue that the recording and analysis of an individual's visual attention during wayfinding can provide insights into her spatial abilities and employed wayfinding strategies. Examples from an explorative pilot study with a pedestrian audio guide in an urban context illustrate that mobile eye-tracking data provide new possibilities to analyze map usage, landmark identification, and orientation strategies.
Keywords: Wayfinding, eye-tracking, behavior interpretation

1 Introduction

Wayfinding is a process that people engage in on a daily basis. It can be described as purposeful, directed, and motivated movement from an origin to a specific distant destination that cannot be directly perceived by the traveler [1]. People must utilize various cognitive and spatial abilities in order to accomplish the specific tasks included in wayfinding, such as choosing a route, maintaining orientation, and recognizing landmarks [2]. Landmarks are a key component in wayfinding because they support the building of a mental representation of space and they are used in the communication of wayfinding directions, preferably at decision points [3]. Landmarks are usually associated with salient features in the environment [4].

Over the years, researchers have run numerous studies with the goal of providing insight into people's spatial thinking and reasoning during wayfinding, such as their employed strategies. Different methods, such as interviews, behavior observation, and cognitive map drawing, have been employed to assess people's wayfinding problems but usually the results only indicate *where* these problems occur and not *why*. In this paper we introduce location-aware mobile eye-tracking as a novel approach for evaluating people's wayfinding processes and therefore explaining their wayfinding behavior. This will enhance our understanding of why people commit wayfinding errors, and help determine the potential impact of analog and digital maps—such as the ones used in navigation services—on wayfinding performance.

We argue that the exploration and analysis of eye-movement patterns in the real world during the performance of a wayfinding task allows a researcher to investigate fundamental cognitive processes such as matching the description or image of a landmark to the corresponding object in the environment. We demonstrate this by referring to concrete examples from a mobile eye-tracking study in the city of Zurich, Switzerland.

Figure 1: Wayfinding with a mobile eye-tracker, audio guide, and map.

2 Related Work

Eye-tracking has been used to study human visual attention in areas such as psychology, text comprehension, and marketing (see [5] for an overview). Interest in eye-tracking methodology is also growing in Geographic Information Science (GIScience) research: for instance, the interaction of users with maps and Geographic Information Systems (GIS) has been evaluated through eye-tracking studies [6], [7]. In mobile computing eye-tracking is mainly used for the usability evaluation of novel interfaces [8]. Most relevant to our research, however, are those approaches related to wayfinding (section 2.1) and other location-based outdoor tasks (section 2.2).

2.1 Eye-Tracking and Wayfinding

One research focus in this area lies on visual attention during spatial decision tasks. [9] reports on highly controlled lab experiments with a static eye-tracker, where participants were presented with images of virtual 3D indoor scenes picturing empty hallways and walls. In contrast to real-world spatial decision-making participants had only two options: “go left”, “go right”. Three tasks were evaluated in different experiments: an unguided search task, a scene memorization task, and a free viewing task without direction choice. Spatial overview knowledge was not the focus of the study, therefore scenes were spatially unconnected.

Participants in this study tended to inspect the part of the image more closely which they later chose as direction. The paper also reports on how geometric features of the scenes influence participants’ visual attention. The study provides valuable insights into the perception of spatial scenes while at the same time opening up questions for future research: Can these results be reproduced in real-world outdoor situations? How do movement history and overview knowledge influence gaze behavior? What about visual attention in wayfinding tasks that are guided by route instructions, landmarks, or a map?

An indoor real-world MET wayfinding study about visual attention to signage in a nursing home is sketched by [10].

2.2 Location-Aware Mobile Eye-Tracking

The questions raised in section 2.1 about how to make eye-tracking studies more realistic are also discussed in the eye-tracking community. [11] reports on an eye-tracking study about the interaction of children while playing. They used a mobile eye-tracker because it “provide[s] a good solution for studying perception with the freedom of movement and variable contexts that characterize natural vision” (p.21). For wayfinding studies freedom of movement and variable context are even more relevant because wayfinding happens in large-scale space. This is the main argument in [12], introducing *location-aware mobile eye-tracking* (LA-MET) as the combined recording and analysis of (geographic) position and gaze. This combination of user and gaze position (the location the user looks at) allows for sophisticated analyses. The authors of [12] demonstrate this idea by annotating for each position of a motion track the area of interest (AOI) the individual is looking at from that position. In principle there can be two types of AOIs: those overlayed on a mobile object, such as maps or mobile phones, and those located in the environment, such as the façade of a building. If only mobile AOIs are relevant they refer to this as *type 1 LA-MET*, if both types of AOI are used they call it *type 2 LA-MET*.

The real-world example provided in [12] is restricted to type 1 LA-MET: motion tracks of pedestrians navigating in a city with the help of a paper map are enriched with information on whether the person has gazed at the map or not. By aggregating several of these AOI-annotated motion tracks, it was possible to identify the spatial regions in which participants needed the map most. Although this approach can help in identifying critical decision points in a pedestrian wayfinding task it does not provide an explanation of *why* these locations are critical. Why is the map or route description inconsistent at this point? A more thorough analysis, including gazes in the environment (type 2 LA-MET), is needed to answer questions that relate to the cognitive processes involved in wayfinding.

3 Wayfinding with the Zurich Audio Guide: An Explorative Pilot Study

The theoretical considerations introduced in section 4 are exemplified with preliminary results from a pedestrian wayfinding study described in the following.

3.1 Goals

Outdoor wayfinding experiments with MET are a novel field therefore our study was set up as an explorative pilot ([5], p. 66). The aim was to identify theoretical and technological issues to be solved before wayfinding studies with LA-MET become possible.

We suspected that type 1 LA-MET studies [12] do not provide sufficient data to analyse the cognitive processes involved in wayfinding. We were driven by the idea that type 2 LA-MET would allow us to better conceptualize human wayfinding because objects in the environment are relevant for wayfinding decisions. However, a system that computes the 3D gaze position in an outdoor environment automatically is, to our knowledge, not available yet. It would require the combination of existing technologies, such as head tracking, image processing, 3D city models etc., and algorithms that deal with sensor inaccuracy. Before building a system that computes the gaze position we wanted to clarify its potential benefits. Thus, our two main goals were:

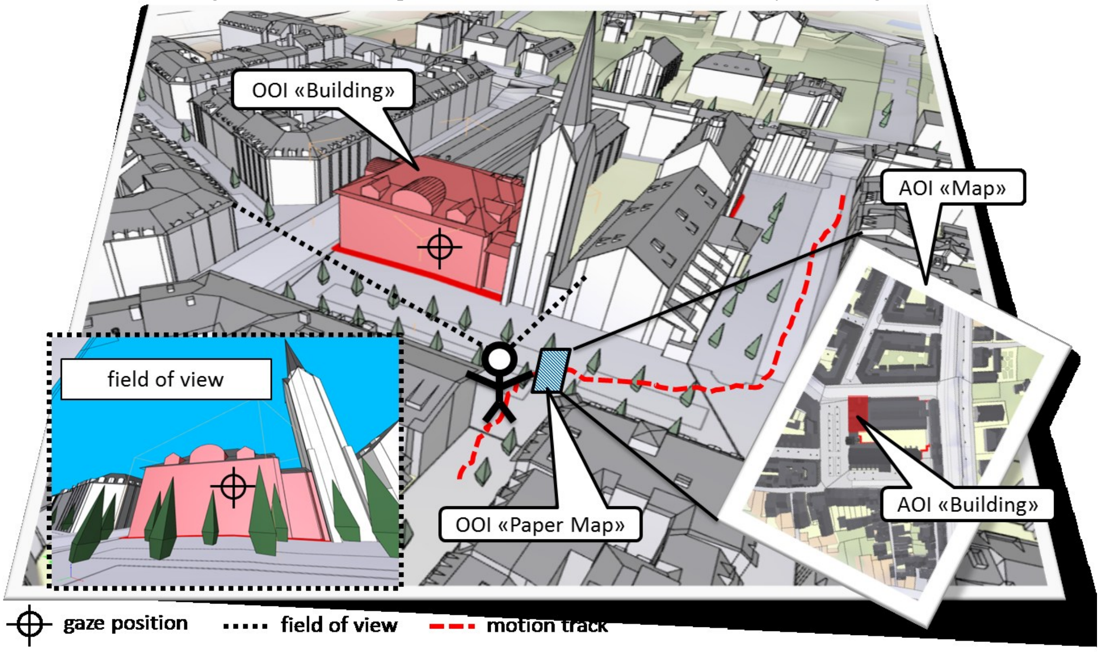
Usefulness: does the analysis of visual attention on AOIs in the environment provide valuable information on the cognitive processes involved in wayfinding, compared to type 1 LA-MET?

Conceptualization: how can we conceptualize and measure visual attention during wayfinding in a way that facilitates the formulation of hypotheses in later studies?

Figure 2a) Overview of the study area with optimal path for eight tasks. b) AOI-annotated motion track of one participant solving tasks 1 and 2

Source: Map background: OpenStreetMap

Figure 3: A 3D conceptualization of location-aware mobile eye-tracking



Source: © 3D city model: Stadt Zürich, Geomatik und Vermessung

3.2 Study Set-Up

Wayfinding tasks and map

The pilot study took place in the city of Zurich, Switzerland. Participants were instructed to solve wayfinding tasks provided by audio instructions. We picked nine tracks from an official Zurich audio guide[1]. Background information on history and sightseeing was removed, leaving only navigational instructions (avg. duration 20 sec). One track was used as a trial task in which participants got acquainted to the hardware and the reactions of passersby. The other tracks yielded in tasks 1-8 (Fig. 2a).

Participants were given a paper map showing a Google Maps™ screenshot (Figs. 1, 4). Two researchers accompanied the participant, one responsible for traffic safety, the other for taking notes. Participants were instructed to tell whenever they thought they had reached the destination. The researchers then took notes regarding success or failure for the respective task. Participants could also give up. Before each task participants were lead to the pre-defined starting position of the next task. In the context of this paper, the first two tasks were the most relevant ones:

Task 1: Participants started at the tram station "Bellevue" and were instructed: "Cross the road south of the tram station and proceed to Sechseläutenwiese" (the name of the square south of Bellevue, Fig. 2). Any location inside the Sechseläutenwiese square was judged as a correct solution. The difficulty of this task consisted in orienting oneself with respect to cardinal directions.

Task 2: Participants started at a predefined location on Sechseläutenwiese and were instructed: "Our next destination is the Opera. The prominent building is located at the southern edge of Sechseläutenwiese where the Seefeld quarter starts". Participants were then supposed to walk to the Opera entrance. As there are two prominent buildings south of Sechseläutenwiese the difficulty of this task was to decide between the two.

Participants and data

Five participants took part in the study (three male, two female). All of them were unfamiliar with the city and the study area.

A head-mounted, monocular eye tracker with 50 Hz recording frequency, the Ergoneers Dikablis Cable system (http://www.ergoneers.com/), was used to record gaze data which was stored on a notebook in a backpack. The recording included two video streams, one for the eye and one for the field of view, together with calibration data. A GPS motion track was recorded at a frequency of one coordinate per second using a standard Android smartphone (Samsung Galaxy S II). Data were collected while moving and while standing.

4 Explaining Wayfinding Behavior With Location-Aware Mobile Eye-Tracking

4.1 A Conceptualization of LA-MET

Most static eye-tracking studies measure visual attention with *areas of interest* (AOIs). An AOI is a polygon in the stimulus picture for which visual attention is analyzed. In [9] three

[1] "Geld und Geist" ("money and intellect", only in German), http://www.stadt-zuerich.ch/vbz/de/index/freizeit_events/vbz_podcasts/geld_und_geist.html

AOIs ("left", "middle", "right") were defined for each picture. Several measures are defined on how often, how long, or in which sequence the gaze hits certain AOIs. [5] (chapter 6) proposes to use AOIs also for gaze-overlaid videos resulting from MET. This conceptualizes a 3D problem with 2D methodology which does not allow to adequately model type 2 LA-MET problems. In wayfinding, for instance, we typically want to specify interest for a certain landmark building, not for the building's projection into 2D space.

Thus, we propose to introduce a new 3D concept for LA-MET studies, the *object of interest* (OOI). An OOI is a 3D object in the real world for which we measure the visual attention the participant pays to. An OOI in wayfinding studies could be a street sign, a building, or just a part of a building. Navigation aids carried along are OOIs with changing position. 2D AOIs are still relevant for LA-MET. They are usually overlayed with an OOI, such as the two AOIs "Map" and "Building" on the OOI "Paper Map" in Fig. 3. Alternatively, an AOI can capture an abstract spatial concept, independent of an OOI, such as "top" or "bottom" for the top or bottom half of the recorded video, or a cardinal direction.

The 3D conceptualization of OOIs not only comes closer to our natural thinking about wayfinding. It may also be possible to determine the OOI the individual is currently gazing at without having to infer it from according corresponding AOI in a 2D gaze-overlayed video. This would require the head tracking technology and 3D models mentioned in section 3.1 and circumvent a manual frame-by-frame identification of gaze hits in the gaze-overlaid video.

4.2 LA-MET Measures For Wayfinding Behavior

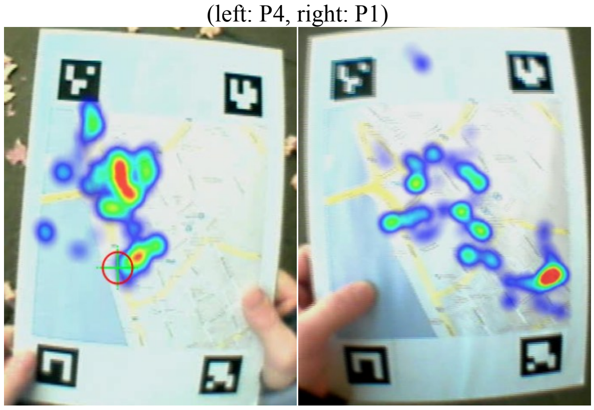
Measuring Map Usage

A typical question in wayfinding studies is how often and where participants use the map. Current studies define map usage from the perspective of an external observer: "Interacting with the map was defined as people were holding the map in their field of vision" [13]. This may be inaccurate because map and environment can both be in the field of view at the same time. MET disambiguates these cases and works automatically[2]. Figure 2b exemplifies a possible result for one participant of the Zurich audio study solving tasks 1 and 2. Positions with gaze at AOI "map" are marked red. In [12] several of these AOI-annotated motion tracks were aggregated to identify critical decision points.

One main advantage of MET over external observance of map usage is that it reveals which sub-AOIs on a map the participant pays attention to: Figure 4 shows an attention map (a heat map of gaze positions) for two participants. It aggregates data of all moments during task 1 in which the gaze was on the map. P4 paid most attention to his current position, second most attention to the destination, and some attention to the lake and waterside (maybe because they helped for orientation). P1's attention was distributed much more equally on the map and not related to start or destination. P4 solved the task correctly, P1 walked into the opposite direction.

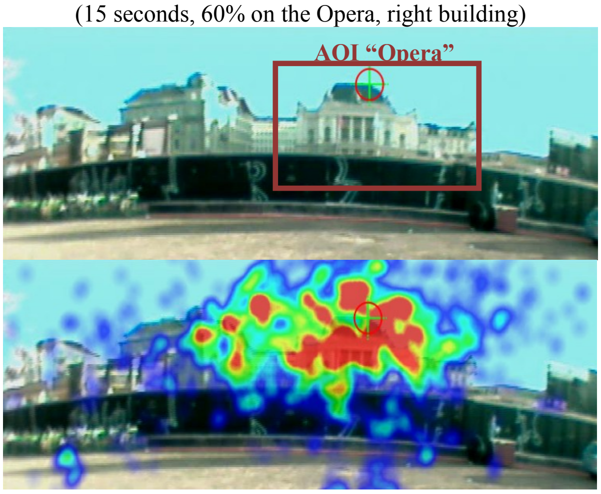
As this was a pilot study we did not explicitly define sub-AOIs before the study. The preliminary results indicate that "Bellevue" as start and "Sechseläutenwiese" as destination are good candidates for sub-AOIs in future studies.

Figure 4: Attention map for two participants trying to locate themselves on the map at the beginning of task 1 (left: P4, right: P1)



Source: © Map background: Google

Figure 5: Visual attention of P1 trying to identify the Opera building at the beginning of task 2 (15 seconds, 60% on the Opera, right building)



Measuring Landmark Identification

Many wayfinding studies are concerned about whether landmarks used in the route instructions are comprehensible and unambiguous. Without MET this can be determined from whether the participant decided correctly and fast. MET additionally allows for a detailed analysis of the OOIs that drew the participant's attention during decision making. We could then assume that these OOIs correspond to the different hypotheses the participant considers. This is especially interesting for explaining *why* a decision took very long or failed.

Figure 5, for instance, shows an attention map for P1 at the start of task 2 with an analysis for the OOI "Opera" (which becomes an AOI "Opera" in the 2D projection of the analysis software). The figure indicates that P1 was able to

2 In our case, the visual markers in the corners of the map were used to create a coordinate system for which we defined an AOI "Map".

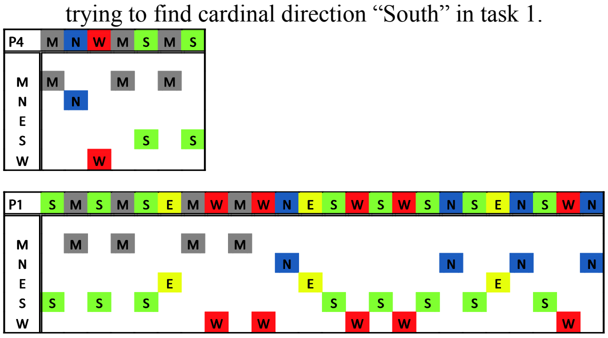
disambiguate between the “NZZ” building (left) and the “Opera” building (right). 60% of his attention was directed at the Opera. After 15 seconds P1 started walking and solved the task correctly.

Measuring Orientation Strategies

An important sub-task in wayfinding with maps is finding one’s position on the map by aligning it with the environment. This requires paying attention to the environment and the map. In complex situations it may be necessary to switch attention between AOIs on the map and OOIs in the environment several times. MET allows for measuring these gaze switches and aggregating data, such as the sequence charts in Fig. 6. It visualizes the gaze sequences of P4 and P1 at the beginning of task 1. “M” denotes the AOI “Map”, the other four relate to cardinal directions as abstract spatial concepts.

P4 spent 54 seconds for his decision during which visual attention switched six times. P1 spent 131 seconds, switching 23 times. The sequences indicate that P1 after some time gave up on using the map. He finally opted for the wrong direction north. In contrast, P4’s orientation strategy was efficient and successful. This result confirms the finding of Fig. 4 (same participants, same task).

Figure 6: Gaze sequence charts of P4 (top) and P1 (bottom) trying to find cardinal direction “South” in task 1.



5 Discussion and Outlook

This research proposed location-aware mobile eye-tracking (LA-MET) for studying the cognitive processes involved in outdoor wayfinding tasks. We conceptualized LA-MET in 3D and illustrated the benefits of LA-MET for wayfinding studies with examples from a pilot study. Though only few participants were recorded so far, the data suggest that LA-MET enhances the analysis of map usage, landmark identification, and orientation strategies. As future work we will carry out a larger study on the basis of the current results.

Using LA-MET for wayfinding studies is a novel approach that raises different questions. Two important ones are the following: first, field studies are much less controlled than lab studies. Random factors may influence wayfinding, such as the number of passersby and their reactions, traffic and weather conditions, etc. On the other hand, lab studies lack realism. Future research is needed to explore how visual attention in virtual environments and field studies differs. Second, eye-trackers only measure gaze, not perception or attention. Participants may perceive parts of the scene, which is out-of-focus. MET data should thus be taken as an indicator for perception, not as perception itself. We propose triangulation with other methods to ensure that results are reliable. A further discussion of these issues is also part of future research.

References

[1] Golledge, R.G., *Human wayfinding and cognitive maps*, in *Wayfinding behavior: Cognitive mapping and other spatial processes*, R.G. Golledge, Editor 1999, Johns Hopkins University Press: Baltimore, MA. pp. 5-45.

[2] Montello, D. and Raubal, M., *Functions and applications of spatial cognition*, in *The APA handbook of spatial cognition*, D. Waller and L. Nadel (Eds). in press, APA: Washington DC.

[3] Denis, M., Michon, P.-E. and Tom, A., *Assisting Pedestrian Wayfinding in Urban Settings: Why References to Landmarks are Crucial in Direction-Giving*, in *Applied Spatial Cognition - From Research to Cognitive Technology*, G. Allen, Editor 2007, Lawrence Erlbaum: Mahwah, NJ. pp. 25-51.

[4] Raubal, M. and Winter, S., *Enriching Wayfinding Instructions with Local Landmarks*. in *Geographic Information Science - Second International Conference, GIScience 2002*. 2002. Boulder, CO, USA: Springer.

[5] Holmqvist, K., et al., *Eye Tracking - A Comprehensive Guide To Methods And Measures*2011, New York: Oxford University Press.

[6] Çöltekin, A., Fabrikant, S.I., and Lacayo, M., *Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings.* International Journal of Geographical Information Systems, 2010. **24**(10): pp. 1559-1575.

[7] Opach, T. and Nossum, A., *Evaluating the Usabilty of Cartographic Animations with Eyemovement Analysis*. in *25th International Cartographic Conference*. 2011. Paris.

[8] Cauchard, J.R., et al., *Visual Separation in Mobile Multi-Display Environments*. in *24nd annual ACM symposium on User interface software and technology (UIST 2011)*. 2011. ACM.

[9] Wiener, J.M., et al., *Gaze behaviour during space perception and spatial decision making.* Psychological Research, 2011: pp. 1-17.

[10] Schuchard, R.A., Connell, B.R., and Griffiths, P., *An environmental investigation of wayfinding in a nursing home*. in *Symp. on Eye tracking research & applications (ETRA 2006)*. 2006. San Diego, California: ACM.

[11] Franchak, J.M., et al., *Head-mounted eye-tracking of infants' natural interactions: a new method*. in *Symposium on Eye-Tracking Research and Applications (ETRA 2010)*. 2010. Austin, Texas: ACM.

[12] Kiefer, P., Straub, F., and Raubal, M., *Towards Location-Aware Mobile Eye Tracking*. in *Symposium on Eye-Tracking Research and Applications (ETRA 2012)*. 2012. Santa Barbara, CA, USA. (accepted)

[13] Pielot, M., Henze, N., and Boll, S., *Supporting map-based wayfinding with tactile cues*. in *11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2009)*. 2009. Bonn, Germany: ACM.