Mobile Outdoor Gaze-Based GeoHCI

Ioannis Giannopoulos
Institute of Cartography and Geoinformation
ETH Zurich
igiannopoulos@ethz.ch

Peter Kiefer
Institute of Cartography and Geoinformation
ETH Zurich
pekiefer@ethz.ch

Martin Raubal
Institute of Cartography and Geoinformation
ETH Zurich
mraubal@ethz.ch

ABSTRACT
This position paper considers *location-aware mobile gaze-based interaction*. In this form of interaction three types of spatial information can be used: a) the user's location, b) gazes at the geo-content displayed on the screen, c) gazes at objects in the real world. We summarize current work and outline future challenges.

INTRODUCTION
People frequently face mobile decision situations, such as wayfinding in an unfamiliar city. These situations can be supported by geo-spatial information, for instance maps or route instructions. We call the interaction with such information *mobile geographic human computer interaction* (Mob-GeoHCI). Note that Mob-GeoHCI can be considered as "geographic" in two senses: the user interacts with geo-spatial information while herself being situated in a geo-spatial context. Mobility poses a number of challenges different to non-mobile GeoHCI. These include: 1) the small screen size restricting the spatial information visible on a map; 2) the input modalities available for standard mobile devices; 3) taking into account the mobile user's task, time pressure, and social context; 4) achieving at least "weak cognitive adequacy" [6] w.r.t. the mental spatial representation the user has about the environment, and supporting the creation of such mental representations.

We propose using mobile eye tracking as an interaction method to tackle these challenges (location-aware gaze-based interaction). Eye movements deliver new insights into geo-spatial thinking and problem solving, such as how people find their way in a city [4]. They can also be used to trigger new kinds of interactions. A number of gaze based approaches have been explored in the last 20 years, trying to incorporate eye movement interaction in a natural and unobtrusive way [2]. In the domain of GeoHCI, gaze based interactions for virtual [5] and real [4] spatial contexts are becoming prominent.

The position of gaze adds a spatial component, which can be utilized in various ways. In the following, we explore location-aware mobile gaze-based interaction w.r.t. the three spatial components involved, summarizing current research and outlining future research directions.

GeoHCI Workshop at CHI 2013, April 27–28, 2013, Paris, France.

LOCATION-AWARE MOBILE GAZE-BASED INTERACTION
In contrast to desktop-based eye trackers, mobile eye trackers are worn as head-mounted systems (see Fig. 1), which allows for gaze-based HCI in mobile real-world scenarios. In a typical architecture, the gaze data gets transferred to a laptop in the user's backpack and processed immediately, which will then trigger a geo-spatial information service.

The 3 Spatial Components
Mob-GeoHCI has two spatial components that need to be considered: the spatial information interacted *with*, and the space interacted *in*. When using gaze-based interaction in a mobile scenario, we gain a third spatial component: the objects in the environment the user gazes at, such as buildings or street signs that represent the spatial dimension of gaze. If the hardware and pre-processing allows for the identification of these objects they can then be included in the interaction.

We thus think about *location-aware mobile gaze-based interaction* w.r.t. three spatial components:

a. The space the user interacts in (i.e., the position of the user)

b. The spatial information the user interacts with (e.g., the information on the map)

c. The space the user interacts with (i.e., the objects in the environment gazed at)

Figure 1. Object of regard identification through eye, location and head tracking.

The space interacting in

Location-based services (LBS) have been around for quite a while, and the assumption that the user's spatial context can influence the interaction method to be used is broadly accepted. Eye tracking studies can provide behavioral data with which we can analyze places w.r.t. the typical visual behavior shown at these places, given a certain task.

For instance, in an outdoor study we have previously identified places where the users of a tourist guide gaze at the map most often, which can be assumed to be confusing decision points [4]. An LBS could adapt its interaction method when the user approaches one of these problematic decision points.

The spatial information interacting with

Gaze tracking in a 2D environment, for instance on a map, offers a magnitude of possibilities for implicit and explicit interaction that can provide assistance, and also minimize cognitive load. For example, orientation on small display maps is often difficult because the visible spatial context is restricted. In our research we provided the history of a user's visual attention on a map as a visual clue to facilitate orientation. Visual attention on the map is recorded with eye tracking, clustered geo-spatially, and visualized when the user zooms out. A user study on this implicit gaze-based interaction concept (*GeoGazemarks*, [1]) demonstrated a significantly higher efficiency than normal two-finger interaction without gaze marks.

What if a map 'knew' which places, streets and objects the user focuses on? In recent work we have introduced *gaze map matching* as the problem of algorithmically interpreting eye tracking data with respect to geographic vector features, such as a road network shown on a map [3]. We developed a gaze map matching algorithm based on a Hidden Markov Model and evaluated it against purely geometric algorithms. This approach can be seen as a first step towards providing usage knowledge to a map, opening the way for natural dialogues with humans.

The space interacting with

In our current work we are trying to determine the *object of regard (OOR)*, i.e., the object in the real world the user gazes at. One approach for determining the OOR consists of combining outdoor gaze tracking with location and head tracking. We developed a helmet that is enhanced with a smartphone and a mobile eye tracker. We use the sensors of the smartphone to retrieve the location of the user as well as the head position in space and calibrate this information with the gaze data in order to compute the intersections of each gaze point with the objects of a 3D model of a city (see Fig. 1).

The obvious advantages of this system are twofold: on the one hand we can automate the analyses for recorded eye tracking data, while on the other hand it opens novel ways for interaction with the real environment.

The gaze-based wayfinding assistant scenario

An interaction scenario using all three spatial components is the gaze-based pedestrian wayfinding assistant: such system would detect from the gaze at decision points whether the user has problems reading the map. Gazes on the map and gazes in the environment would be included in the analysis. Such system could observe the visual matching process between the map content and the real world. If an incorrect matching takes place, e.g., a church symbol on the map is mismatched with a different church in the real world, the system should correct the user ("the church which will help you as a landmark is located to the right").

The steps needed to achieve this goal can be subdivided into interdisciplinary tasks that involve several areas, such as geography, cognitive psychology, and HCI, among others.

FUTURE CHALLENGES

A prominent problem in gaze-based interaction is the so-called Midas Touch Problem which occurs when the user triggers an unintended interaction with her gaze. This is particularly relevant for city environments with a large amount of "gazable" objects. An interactive wayfinding assistant must recognize the user's tasks, intentions, and/or plans. Using gaze for these recognition tasks is another line of research we are currently following. Related to this is the challenge of confusion and error detection. Environmental influences pose a big challenge for location-aware mobile gaze-based assistance: sensor inaccuracies on the one hand, and unexpected contextual changes on the other hand. The latter include objects in the environment that are temporarily visually occluded.

REFERENCES

1. Giannopoulos, I., Kiefer, P., and Raubal, M. GeoGazemarks: Providing gaze history for the orientation on small display maps. In *Proc. ICMI'12*, ACM Press (2012), 165-172
2. Jacob, R. J. K. What you look at is what you get: eye movement-based interaction techniques. In *Proc. CHI '90*, ACM Press (1990), 11-18
3. Kiefer, P., and Giannopoulos, I. Gaze map matching: Mapping eye tracking data to geographic vector features. In *Proc. ACM SIGSPATIAL'12*, ACM Press (2012), in print
4. Kiefer, P., Straub. F., and Raubal, M. Towards location-aware mobile eye tracking, In *Proc. ETRA'12*, ACM Press (2012), 313-316
5. Stellmach, S., and Dachselt, R. Investigating gaze-supported multimodal pan and zoom. In *Proc. ETRA '12*, ACM Press (2012), 357-366
6. Strube, G. The Role of Cognitive Science in Knowledge Engineering. In: F. Schmalhofer et al. (eds.), Contemporary Knowledge Engineering and Cognition, Berlin: Springer (1992), 161-174