Watch What I Am Looking At! Eye Gaze and Head-Mounted Displays

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

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

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

Abstract
This position paper considers the collocated interaction opportunities offered through the combination of mobile eye tracking and head-mounted displays. We introduce collocated interaction during wayfinding among collaborators and one of three types of spatial information that can be used. We highlight how the combination of these wearable technologies can lead to advanced novel interaction methods for collocated interaction and outline current and future challenges.

Author Keywords
eye tracking, head-mounted display, collocated interaction, wayfinding

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User Interfaces.

Introduction
People frequently face mobile decision situations, such as during wayfinding in urban environments. These situations are supported through mobile devices that enable interaction dialogues to assist their users. In our work [2] we have previously focused on mobile assistance for a single user, enabling novel interaction dialogues to support the wayfinding process. During wayfinding, the

user is often involved in social interaction with other pedestrians or travel companions [7]. Currently we are extending our research to multi-user interaction dialogues in order to support collocated collaborative wayfinding, mostly through sharing eye gazes between two or more users. In the following, we will introduce our ongoing work and vision for collocated interaction in urban environments. Lundgren et al. [5] introduced a framework for the design of mobile systems that support collocated interaction. The main perspectives of this framework are the social, the technological, the spatial and the temporal perspective. We will frame our vision using these perspectives, starting with the technological perspective in order to introduce the technologies and modalities under consideration. In our work we focus on gaze-based interaction and investigate how gaze can be incorporated in collaborative interaction between humans in urban environments. One of the advantages using eye gaze is the resolution of ambiguities during communication.

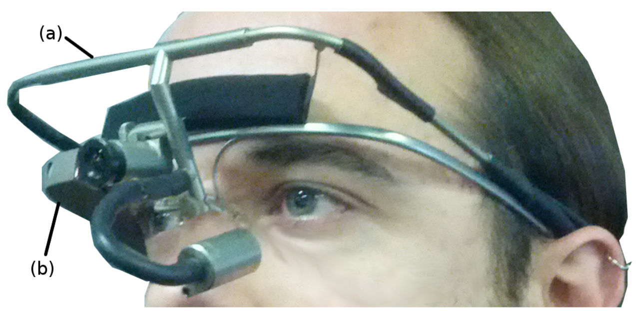

Figure 1: Eye tracking (a) combined with Google Glass (b)

Technological Perspective

We propose and advocate the combination of eye tracking and head-mounted displays (e.g., Google Glass). The combination of these two devices (see figure 1) opens new channels of interaction between individuals, allowing them to collaborate in urban environments, which offers several advantages. For instance, the individuals are able to update the information displayed on the head-mounted display through their eyes, depending on gaze gestures or the exact gaze point. Also, the necessity of holding a device in ones hands is eliminated, allowing for more natural body movement and gestures.

Although this combination of devices has not yet reached the market, the rapid development and published patents (Google Patent US 8611015 B2) of technology vendors are promising. In our ongoing research we combine the Dikablis eye tracker (see figure 1, (a)) and Google Glass (see figure 1, (b)) and use a two step calibration procedure for their combination. The first step is provided by the vendor software and calibrates the eye tracker w.r.t. the environment. The second step informs the software in use about the gaze coordinates that correspond to the boundaries of the head-mounted display. The software in use must be informed about the eye coordinates that correspond to the corners and the center of the head-mounted display (5 point calibration), the rest can be computed using this information. Based on this calibration procedure, an interactive software can distinguish eye gazes in the environment from eye gazes on the head-mounted display.

Spatial Perspective

Collocated interaction involves physical as well as virtual spaces. In previous work [2], we identified three spatial components that are relevant for gaze-based interaction during wayfinding and can be used to frame the collocated interactions.

The space interacting in

The first spatial component involves the space the users interact in (i.e., the physical position of the user). Location-based services (LBS) provide the user with information based on the current position and local context. The location of the collaborators is an important source that the system in use can utilize in order to adapt and enable the appropriate interaction dialogue between the users. An interaction dialogue in indoor environments can be different to one in outdoor environments, and different outdoor environments require different interactions (e.g., city, forest).

The spatial information interacting with

The second component involves the spatial information the users interact with (e.g., the information on the map). Wayfinders utilize assistance aids, such as digital maps in order to make a decision. Maps are often used to interact with other users in order to ask for directions or for clarification in case an existing instruction is unclear (e.g., ambiguity). Collocated users can communicate through multiple channels [6] starting an interaction dialogue consisting of talk, body movement, gestures and pointing [4]. The combination of eye tracking and a head-mounted display will allow for natural body movement, since the digital map can be displayed on the head mounted displays of the collaborators instead of holding it in their hands (hands-free symmetrical information). Moreover, the verbal communication will be less ambiguous, since the collaborator's eye gaze on the map can be followed on the display (symmetrical or asymmetrical information). The association between the verbal communication and the visual artifacts can be easier established. For instance, the collaborator could explain to the wayfinder a route to follow using a map displayed on the head-mounted display of both (symmetrical information), and the wayfinder would be able to follow the collaborators eye gaze on the map (asymmetrical information) while she is giving directions. GeoGazemarks [1] is an eye gaze-based interaction method for the creation of map usage history of a single user. This interaction approach can be modified for collaborative interaction, allowing the wayfinder 1) to follow the eye-gaze of the collaborator while listening to the instructions she is giving as well as 2) seeing all important eye gazes (eye fixations) as visual cues on the map, representing the history and order of the given instructions.

The space interacting with

The third component involves the space the users interact with (i.e., the objects in the environment gazed at). Eye gaze can be utilized as a further communication channel (e.g., pointing) and for determining the object of regard [2]. For example, a wayfinder asks a passing-by pedestrian where a certain building is located (e.g., museum). The collaborator provides the wayfinder with verbal instructions and points with his finger to the requested building. The wayfinder looks towards the pointed direction and is able to identify the building immediately, since the eye gaze of the collaborator is used to augment the head-mounted display of the wayfinder. This kind of interaction dialogue may also be applied to more than two collocated users. An interesting example are tourists and a tour guide as lead collaborator. The tourists are able to follow the eye gaze of the tour guide while the guide provides them with cultural and historic information on a building.

Social Perspective

The design of collocated interactions has to consider the roles of the different users. A wayfinder can collaborate with a known user, having a private relationship, but also with strangers who can be asked for collaboration. One must consider how a joint attention can be created and what kind of information can be exchanged for the collaboration dialogue.

Temporal Perspective

The synchronization of actions has to be both system- and user-driven. The described interactions require that the users are physically close to each other (system-driven), but the users have to engage in communication (user-driven) in order to start the interaction dialogue. The interaction duration depends on the social context as well as the involved activity.

Challenges

Our main goal, using the presented combination of wearables, is to design and study a series of collocated interactions that will lead to design principles. It is important to understand which kind of information should be utilized to enhance the communication process between collaborators, as well as which of this information should be presented asymmetrically and which symmetrically. A prominent problem of eye gaze-based interactions is the so-called Midas Touch Problem [3] which occurs when the user triggers an unintended interaction with her eye gaze. A typical solution is to use a multimodal approach. Our head-mounted display is equipped with sensors, such as an accelerometer and gyroscope that could be utilized to trigger the desired actions. Although we focus on pedestrian wayfinding assistance, the presented interactions can be easily generalized and applied in other domains.

References

[1] Giannopoulos, I., Kiefer, P., and Raubal, M. Geogazemarks: Providing gaze history for the orientation on small display maps. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ICMI '12, ACM (New York, NY, USA, 2012), 165–172.

[2] Giannopoulos, I., Kiefer, P., and Raubal, M. Mobile outdoor gaze-based geohci. In *Geographic Human-Computer Interaction, Workshop at CHI 2013* (2013), 12–13.

[3] Jacob, R. J. Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in human-computer interaction 4* (1993), 151–190.

[4] Luff, P., and Jirotka, M. Interactional resources for the support of collaborative activities: Common problems in the design of technologies to support groups and communities. In *Community Computing and Support Systems, Social Interaction in Networked Communities [the Book is Based on the Kyoto Meeting on Social Interaction and Communityware, Held in Kyoto, Japan, in June 1998]*, Springer-Verlag (London, UK, UK, 1998), 249–266.

[5] Lundgren, S., Fischer, J. E., Reeves, S., and Torgersson, O. Designing mobile experiences for collocated interaction. In *Proceedings of the companion publication of the 18th ACM conference on Computer supported cooperative work & social computing*, ACM (2015).

[6] Olson, G. M., and Olson, J. S. Distance matters. *Human-computer interaction 15*, 2 (2000), 139–178.

[7] Reilly, D., Mackay, B., Watters, C., and Inkpen, K. Planners, navigators, and pragmatists: collaborative wayfinding using a single mobile phone. *Personal and Ubiquitous Computing 13*, 4 (2009), 321–329.