Eye Tracking for Spatial Research: Cognition, Computation, Challenges

Journal:	*Spatial Cognition and Computation: An Interdisciplinary Journal*
Manuscript ID	HSCC-2016-0652.R1
Manuscript Type:	Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Kiefer, Peter; ETH Zurich, Institute of Cartography and Geoinformation Giannopoulos, Ioannis; ETH Zurich, Institute of Cartography and Geoinformation Raubal, Martin; ETH Zurich, Institute of Cartography and Geoinformation Duchowski, Andrew; Clemson University, School of Computing
Keywords:	eye tracking, visual attention, wayfinding, cartography, human-computer interaction (HCI)

Eye Tracking for Spatial Research:

Cognition, Computation, Challenges

Peter Kiefer[1], Ioannis Giannopoulos[1], Martin Raubal[1], Andrew Duchowski[2]

[1] Institute of Cartography and Geoinformation, ETH Zurich, Switzerland

[2] School of Computing, Clemson University, S.C., USA

Author Note

Correspondence: Peter Kiefer, Institute of Cartography and Geoinformation, ETH Zurich, Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland, pekiefer@ethz.ch

Abstract

Spatial information acquisition happens in large part through the visual sense. Studying visual attention and its connection to cognitive processes has been the interest of many research efforts in spatial cognition over the years. Technological developments have recently led to an increasing popularity of eye-tracking methodology for investigating research questions related to spatial cognition, geographic information science (GIScience) and cartography. At the same time, eye trackers can nowadays be used as an input device for (cognitively engineered) user interfaces to geographic information. We provide an overview of the most recent literature advancing and utilizing eye-tracking methodology in these fields, introduce the research papers in this special issue, and discuss challenges and opportunities for future research.

Keywords: eye tracking, visual attention, wayfinding, cartography, human computer interaction

1. Introduction

Spatial information acquisition happens in large part through the visual sense (for sighted individuals). We perceive space through our eyes, reason about the task at hand, and perform a visual search if we believe we need more information. Eye tracking allows us to measure an individual's visual attention, yielding a rich source of information on where, when, how long, and in which sequence certain information in space or about space is looked at. Not surprisingly, eye tracking has become a popular method for investigating research questions related to spatial cognition, geographic information science (GIScience), cartography, and related fields (refer to the literature discussed in the following sections). This includes studies on how people interact with geographic information and studies on how space is perceived in decision situations. Knowledge of how people perceive space can help us, for instance, to design better maps and other spatial representations or to decide on the optimal placement of signage in indoor and outdoor environments. Last but not least, eye tracking enables us to enhance existing, or even create new cognitive models that describe and predict how humans behave in, and reason about space.

Recent technological developments in mobile eye trackers have opened up new perspectives for their use in spatial research by allowing for studies outside the research lab, adding the user's location as another dimension of the data. The resulting data enable an analysis of complex spatial-decision processes that include locomotion and visual exploration of the surroundings; pedestrian navigation is one example. Recent hardware developments have also introduced real-time processing capabilities fostering potential for novel interactive applications.

Interaction principles based on the processing, interpretation, and reaction to the user's gaze are actively investigated in human computer interaction (HCI) research. This includes gaze-based interfaces to geo-information in both desktop computer and mobile usage scenarios.

This introductory article to the Special Issue on “Eye Tracking for Spatial Research” aims to provide an overview of the most recent literature in the field. We summarize research on advancing and utilizing eye-tracking methodology in human navigation and cognitive cartography (section 2), as well as on computational issues and gaze as an input modality to geographic information (section 3). An overview of the Special Issue is provided in section 4, before we close with an outlook on future challenges and opportunities in Eye Tracking for Spatial Research (section 5).

2. Eye Tracking and Spatial Cognition

The use of eye tracking in cognitive science is well-established, and the corpus of literature reporting on eye-tracking studies is large. This section focuses on two important applications of eye tracking in spatial cognition—wayfinding (section 2.2) and cartography/geo-visualizations (section 2.3)—which mirrors the scope of the articles that were accepted for this Special Issue. We start with a general introduction on visual attention and cognitive processes (section 2.1).

2.1 Visual attention and cognitive processes

The deployment of visual attention as well as its response to changing conditions is coupled to our cognitive state. Eye movements allows us to track overt visual attention which is associated with a viewer's point of gaze (Goldberg & Kotval, 1999). We are most interested in extracting information about *fixations* and their durations (Jacob & Karn, 2003), as these relatively steady eye movements indicate the regions of space subject to cognitive processing at the moment of fixation. This association between fixations and cognitive processing is known as the *eye-mind assumption*, posited by Just and Carpenter (1976). Redeploying visual attention to a new location is accomplished by *saccades*, the rapid movement of the eyes executed to reposition gaze. Saccades in natural visual scenes are controlled by the allocation of perceptual attention (Kowler et al., 1995), giving rise to the *scanpath* (Noton & Stark, 1971), the spatio-temporal sequence of fixations which provides a good depiction of the viewer's visual processing of a scene (Yarbus, 1967). Figure 1 illustrates an example of a scanpath on a shaded relief.

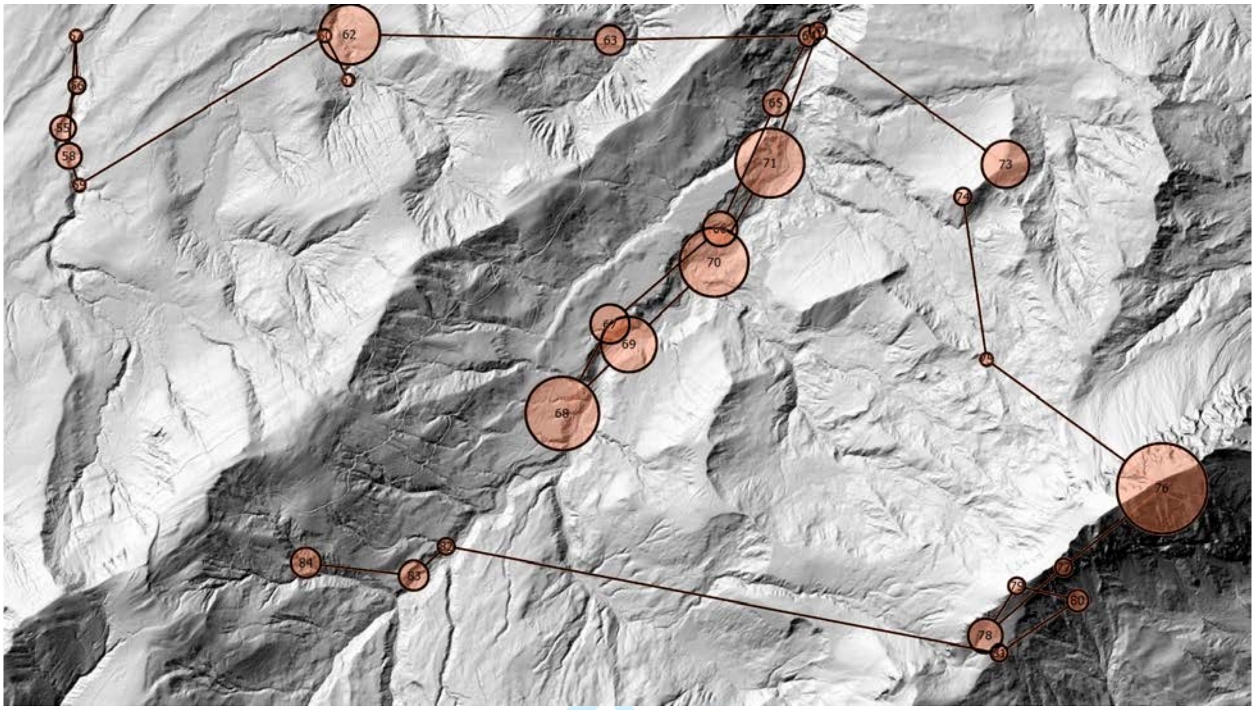


Figure 1: Scanpath of a person on a shaded relief. Circles depict fixations, numbered in the order of occurrence. Lines illustrate saccades. Missing lines between consecutive fixations are caused by blinks.

When it comes to processing spatial information, which may mix textual and diagrammatic information, gaze data suggests that comprehension depends on the informational medium as well as on the ability of the individual (Hegarty & Just, 1993). Gaze fixation data can indicate how viewers integrate information from text and diagrams, e.g., reading text in increments before constructing a spatial mental model of the components with the aid of the diagram.

In general, visual attention is governed by goal-driven intent (e.g., see van Zoest & Donk, 2004), referred to as *top-down* attention. However, in the visual processing of a scene, there is interplay between top-down attentional shifts and *bottom-up* attention, drawn reflexively to spatial locations by external cues. While computational models of top-down attention employ schemas (e.g., Land & Tatler, 2009), most are based on bottom-up models of vision, concerned with stimulus-driven, saliency-based attention, e.g., see Borji and Itti (2013).

2.2 Visual attention during navigation and wayfinding

Navigation, an activity humans perform on a daily basis, can be seen as the combination of locomotion—a coordinated movement in the proximate environment—and wayfinding, the planning and decision-making necessary to reach a destination (Montello, 2005). The decision situations faced during both, locomotion and wayfinding, require the agent to collect information about the environment. While humans can (and do) use several senses to acquire such knowledge (Loomis et al., 2013), the visual sense is particularly interesting to study:

- Many wayfinding aids rely on visual information (e.g., maps, signage).
- Measuring visual attention is relatively easy, compared to measuring what the wayfinder perceives through other senses.
- The visual sense collects spatial information at a much greater distance and higher resolution than other senses.
- Humans actively direct their visual sense (in contrast to, e.g., the vestibular or auditory sense). With eye tracking, we therefore measure not only sensory input, but also the wayfinder's information acquisition procedure.

Several models have been proposed describing the sub-processes involved in wayfinding. For example, Downs and Stea (1977) suggest that wayfinding consists of orientation, route choice, monitoring, and goal recognition. Aiming at a deeper understanding of how these sub-processes work, research in spatial cognition has used eye-tracking technology to study the visual search processes during wayfinding.

Figure 2: Wayfinding study in the real world with a mobile eye tracker and a map.

Orientation, i.e., determining one's position with respect to a reference frame, "involves integrating egocentric visual signals with spatial knowledge about the environment" (Gunzelmann et al., 2004, p. 208). It is sometimes also referred to as "self-localization" (Kiefer et al., 2014). Eye-tracking studies have revealed that participants trained for different orientation strategies will exhibit a different distribution of visual attention (Gunzelmann et al., 2004). In another orientation experiment, it was found that eye movements in scenes with the most salient 3D landmarks tended to strongly focus around those landmarks (Peebles et al., 2007, p. 398f).

Visual attention during landmark-based orientation was also the focus of the real-world study by Kiefer et al. (2014). They found that participants' visual attention distribution and switches (with respect to landmarks and symbols on a map) were related to task performance. Successful participants tended to focus on those map symbols helpful for the task and were able to correctly identify landmarks in the environment shown on the map. Given the importance of landmarks in wayfinding (Raubal & Winter, 2002), researchers have studied whether visual attention may be used as an indicator of what people think is a good landmark (Viaene et al., 2016). Eye tracking can also be used to investigate how to best represent landmarks in a visual wayfinding aid, such as on a map (e.g., Franke & Schweikart, 2017, in this special issue) or in a pedestrian navigation system (e.g., Ohm et al., 2017, also in this special issue).

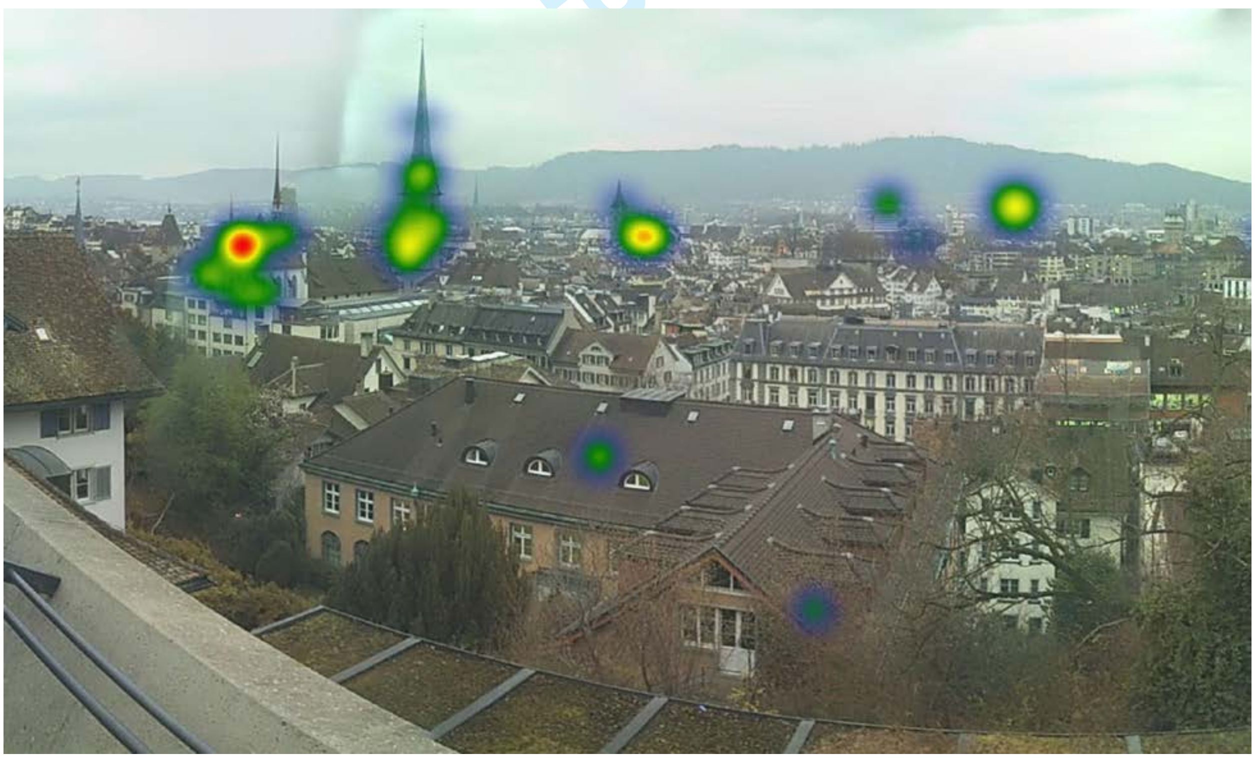


Figure 3: Distribution of visual attention across a city panorama visualized as a heatmap (created by a kernel density operation).

Figure 2 depicts a possible setup for a wayfinding study with eye tracking in the real world, in which the participant needs to orient himself by matching landmarks on an electronic map with landmarks in the environment. An analysis of the sequence (i.e., the scanpath) and the distribution of visual attention (refer to Figure 3 for an example) may provide insight into the wayfinder's orientation strategy.

Route choice, i.e., determining which way to go, is a sub-process which does not necessarily require visual attention, because it can be performed on a mental representation of the spatial environment (also called a "cognitive map") (Golledge, 1999). In unfamiliar environments, route planning and choice are often performed with a visual aid. One eye-tracking study investigating route planning on metro maps is contained in this special issue (Netzel et al., 2017). Wiener et al. (2009) argued that wayfinding is often performed without survey, route, or destination knowledge (e.g., uninformed search, exploration). In this case, route choices are made incrementally, based on the visual features of the surroundings. Eye-tracking experiments have investigated whether uninformed route choices can be predicted by the visual attention of a wayfinder (Wiener et al., 2012) or by the geometric properties of an urban space (Emo, 2012). For map-based wayfinding, it has been suggested that route planning and route choice are followed by a phase of *transformation and encoding* (Meilinger et al., 2006), or shortly: route learning. Visual attention during route learning from a map has, for instance, been studied by Brunyé and Taylor (2009).

Monitoring, i.e., ensuring that one is still following the correct route, is a sub-process for which the wayfinder necessarily needs to perceive the environment. Monitoring can be seen as a continuous and dynamic type of orientation (Rieser, 1999). For instance, taxi drivers—even though they are familiar with a city—will visually search for landmarks to confirm their expectations about the route (Spiers & Maguire, 2008). Monitoring is particularly relevant at crossroads, i.e., when the wayfinder needs to solve a decision problem based on the learned route. Giannopoulos et al. (2014) proposed a model for characterizing the complexity of such wayfinding decision situations, evaluating their model using eye-tracking measures. In the second part of this special issue, Wenczel et al. (2017) describe an eye-tracking experiment investigating visual attention to landmarks while a wayfinder is following a memorized route.

Goal recognition can be seen as a specific type of monitoring. The visual search processes involved are likely to be similar, although, to our best knowledge, this has not been studied with eye tracking yet.

A holistic understanding of human navigation includes also the understanding of visual attention during *locomotion*. How do humans visually scan the proximate space in order to move through it in an efficient and accident-free way (Foulsham et al., 2011)? Where do they look to avoid obstacles (Franchak & Adolph, 2010)? The mode of transportation, such as car driving (Land & Lee, 1994) or bicycling (Vansteenkiste et al., 2013), naturally has a large influence on visual attention during locomotion. A challenge for eye-tracking studies on navigation arises from the fact that wayfinding and locomotion “are generally components of an integrated system … that can be separated only conceptually” (Montello, 2005, p.260). In realistic navigation scenarios, it may therefore be difficult to determine whether a particular fixation or saccade can be explained by locomotion, wayfinding, or both.

2.3 Visual attention on maps and geo-visualizations

Maps and other geo-visualizations support not only wayfinding, but also other types of spatial and spatio-temporal tasks, such as comparing the population density of two countries, or learning about the development of migration over time in the Northern Hemisphere. Research on cartography and geo-visualizations has long had an interest in the cognitive principles of map creation, reading, and understanding (MacEachren, 1995; Montello, 2002), and how to design a geo-visualization that is cognitively adequate for a particular user or user group in a given context (Fabrikant & Lobben, 2009; Slocum et al., 2001). As maps are visual media, eye tracking seems an obvious and appropriate methodology for supporting cartographers in the design process. Early work using eye tracking in cartography ranges back to the 1970s and 1980s (e.g., Castner & Eastman, 1984; Chang et al., 1985; Dobson, 1977; Steinke, 1987). More recently, eye tracking has been suggested to be used systematically as part of empirical methodology in both, cartographic research and practice (e.g., Davies et al., 2015; Fabrikant et al., 2008; Montello, 2002). As can be seen from the following literature review, a rise in the use of eye tracking in cartography can be observed recently. A typical setup for a cartographic eye-tracking study is shown in Figure 4.

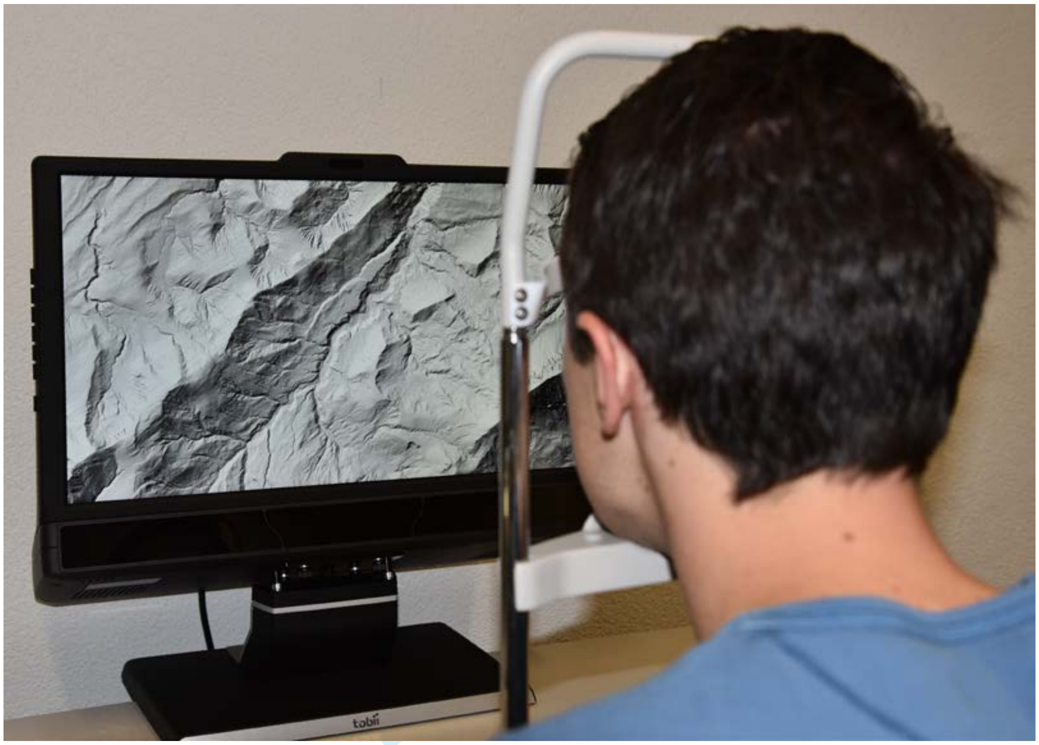


Figure 4: Remote eye-tracking study on a shaded relief.

The first type of application of eye tracking in cartography concerns the development of design guidelines for *static maps*. In this type of research, stimuli are typically created by varying one cartographic design variable systematically, or by changing the complete map design (see, e.g., Brodersen et al., 2001, for the latter). Gaze data are either used as aggregated measures which are then interpreted as an indicator for cognitive states (e.g., the duration of fixations indicates cognitive load) or in an explorative way as gaze visualizations for explaining other dependent variables (e.g., completion times). Example studies include the comparison of label placement methods (Ooms et al., 2012b), the influence of color distance and font size on map readability (Brychtova & Çöltekin, 2016), the best visualization of elevation changes in bicycle maps (Brügger et al., 2016), the comparison of 2D and 3D visualizations (Liao et al., 2016; Popelka & Brychtova, 2013; Popelka & Doležalová, 2015), or the effects of saliency in weather maps (Fabrikant et al., 2010). Present-day eye-tracking studies on maps are almost exclusively performed using digital displays. This was the motivation for the experiment reported by Incoul et al. (2015), who compared map viewing on digital displays and paper. Significantly more fixations per second were found on digital maps, but gaze distributions were similar.

The above-mentioned study on weather maps (Fabrikant et al., 2010) is also one of several investigating the influence of expertise on map viewing (see also Ooms et al., 2012a; Stofer & Che, 2014). This second type of approach—varying the *user group* of a map—is motivated by the insight that group differences play an important role in map usage (Slocum et al., 2001; Thorndyke & Stasz, 1980), which should be considered to ensure an effective cartographic communication process (Kraak, 2003).

Animated maps can be identified as a third research topic in eye tracking on maps. While on the one hand, animations may simply serve as an attention grabbing element, animations may on the other hand be used for conveying information about complex spatio-temporal phenomena (Andrienko et al., 2010), which includes spatial depictions in real-time surveillance (Maggi et al., 2016). Therefore, most eye-tracking studies on animated maps have one of two main purposes: (1) determining how an animation needs to be designed in order to effectively attract attention (e.g., regarding timing, Krassanakis et al., 2016, or regarding visual design, Dong et al., 2014), or (2) investigating how map viewers understand animations (Opach et al., 2014). The evaluation of eye-tracking data collected on animated stimuli can be more challenging than that from static stimuli, since most vendors' standard software packages do not support an automated analysis of visual attention on dynamic stimuli.

Even more challenging is the evaluation of eye-tracking data collected in studies on *interactive maps*, including web maps. As argued by Ooms et al. (2015), the visible extent of an interactive map is dynamic, based on input by the user, and one usually wants to log all interactions and synchronize them with the eye-tracking data. One research topic here is the interplay of user interaction and visual attention, such as the spatio-temporal similarity of mouse and gaze trajectories (Çöltekin et al., 2014) and the influence of panning on eye movements (Ooms et al., 2016). Eye movement analysis has also been proposed as part the methodologies for usability testing of interactive maps (Alaçam & Dalcı, 2009; Çöltekin et al., 2009), a promising approach for map and interface designers in practice and research.

3. Computational Issues and Gaze-Based Interaction

Processing eye-tracking data computationally can be a challenging task, and anyone using eye tracking in their research should have an understanding of the basic principles. A short introduction on computational methods for processing eye-tracking data is provided in the following (section 3.1). Furthermore, we report on the use of eye tracking as an input modality in geographic human-computer interaction (section 3.2).

3.1 Computing with Eye-Tracking Data

The goal of an eye tracker is to record gaze as it falls on objects of visual interest, often represented by 2D planar surfaces. Eye movements, as captured by eye-tracking devices, are generally recorded in one of two ways: either by a head-mounted (also known as a mobile) eye tracker (refer to Figure 2) or by a table-mounted (known as a remote) eye tracker (refer to Figure 4). It is important to keep in mind the differing characteristics of both eye-tracker types and the data each generates. With a head-mounted tracker, the user is free to look around their (3D) environment, which often contains multiple surfaces with which the gaze vector can intersect. The (2D) coordinates of the user's eyes are mapped to the 2D coordinates of the tracker's environment-facing camera. Often, to obtain gaze intersection points, surface planes in the environment need to be located, measured, and mapped to the eye tracker's reference frame. In contrast, a table-mounted eye tracker can more readily map the user's gaze because this tracker's environment is usually constrained to a single plane, namely the computer monitor which the user is looking at, and to which the eye tracker has been calibrated. In either case, the eye tracker's output consists of 2D spatial coordinates of the user's gaze as it moves from point to point over time.

Spatio-temporal eye movements, when expressed as raw gaze data, are generally processed into saccades and fixations via the following "gaze analytics" pipeline:

1. Denoising and filtering raw gaze data $g_i = (x_i, y_i, t_i)$, and classifying raw gaze into fixations, $f_i = (x_i, y_i, t_i, d_i)$, where the (x_i, y_i) 2D coordinates indicate the position of the gaze point, or centroid of the fixation, with t_i indicating the timestamp of the gaze point or fixation and d_i the fixation's duration,

2. collating fixation-related information for its subsequent statistical comparison,
3. interpreting and visualizing statistical tests conducted on processed data.

Classification of raw gaze data into fixations is often accomplished by filtering the raw data by a digital filter. One such filter, the Savitzky-Golay filter (Gorry, 1990; Savitzky & Golay, 1964) is used to differentiate the positional gaze signal into its velocity estimate. Gaze points with velocity over threshold are then considered saccades, with the remaining points classified as fixations. Other methods of eye movement classification can be found in Holmqvist et al. (2011).

Complementing eye-tracking data analysis is visualization of the spatio-temporal (dynamic) data. In their report, Blascheck et al. (2014) review the state-of-the-art in scanpath visualization. They note that for a typical scanpath visualization, each fixation is indicated by a circle (or disk), where the radius corresponds to the fixation duration. Saccades between fixations are represented by connecting lines between these circles. The connecting lines may include arrowheads, and the fixation circles/disks may include numbers to indicate scanpath order. Beyond scanpaths, Blascheck et al. classify visualization techniques for eye movement data into three categories: point-based, Area-Of-Interest (AOI)-based, and those using both. They further distinguish between animated and static, 2D and 3D, in-context and not in-context, as well as interactive and non-interactive visualizations. Finally, visualization techniques are classified as either temporal, spatial, or spatio-temporal.

3.2 Gaze-Based Interaction with Spatial Information

When we try to reach a destination, we often utilize digital assistance aids such as cartographic maps in order to facilitate decision-making. These assistance aids provide us with various pieces of information about our surroundings, often much more than we actually need. Optimizing the interaction with the spatial information provided by these assistance systems could eventually minimize the cognitive complexity during wayfinding and also increase performance, since the wayfinders would be able to extract the information they are seeking much easier and with less errors. For example, Rohs et al. (2007) introduced the dynamic peephole interaction for maps, helping the users during the interpretation of the acquired spatial information by putting the currently inspected map area into a larger spatial context. Although a wayfinder could benefit from such an interaction concept, it requires the user to alter her interaction routines with the map in order for the system to retrieve the currently inspected map area. This could be overcome by introducing further input sources. For instance, the user's gaze could be utilized for the above interaction in order to retrieve the relevant map area. Over the last few years, several gaze-based approaches have emerged, introducing gaze as a further source of information for easy, natural, and fast ways of interaction (Tanriverdi & Jacob, 2000).

Advancements in eye-tracking technology allow for real-time gaze tracking that can be used to start an interaction dialog with a computer system, such as desktop and mobile interfaces. The user's gaze can be directly utilized as an input method for explicit interaction, e.g., as a pointing device, or aggregated over a short period of time in order to enable implicit interaction. According to Schmidt (2000), explicit interaction is intended by the user, i.e., the user performs a certain action and expects a reaction by the system. Implicit interaction occurs when the system reacts to a user's actions which were not primarily intended to trigger an interaction.

Figure 5: Study setup for gaze-based interaction with an interactive mobile map (refer to Giannopoulos et al., 2012). The user's gaze on the map during the interaction is recorded (in map coordinates), aggregated, and visualized when the user zooms out (circles). These circles are called GeoGazemarks and serve as visual bookmarks on the map.

Giannopoulos et al. (2012) utilized the user’s gaze to enable implicit interaction with maps displayed on a mobile device. This was the first approach enabling gaze-based interaction on mobile devices (refer to Figure 5). They utilized a mobile eye tracker in order to capture the user’s gaze on the screen of a mobile device while interacting with a map and aligned the screen coordinates to the geospatial coordinates the user was gazing at. Once the user zoomed out, the previously registered gaze points on the map were visualized in order to serve as orientation cues. Bektaş and Çöltekin (2011) introduced the concept of Geofoveation in order to reduce the information intensity displayed on cartographic maps. They utilized the user’s gaze in order to balance the amount of information displayed on a screen, showing only the regions in full details where the user was gazing. A further example of implicit interaction was introduced by Kiefer et al. (2013). They utilized the users' gaze while they were interacting with a cartographic map; through machine learning, they trained a classifier to predict the task the user was currently engaged in, such as map search and route planning. This approach allowed them to come closer to the cognitive state of the user and thus, in a further step, this information could be used in a map adaptation process in order to assist the user.

Concerning explicit interaction with spatial information, Stellmach and Dachselt (2012) introduced gaze-based interaction for zooming and panning with a virtual globe. They enabled a multimodal zoom interaction based on the user's current gaze position and directed the panning interaction through gaze, by moving the currently viewed content into the center of the screen. Giannopoulos et al. (2015) utilized the users' gaze in the urban environment in real-time in order to guide them during navigation. Whenever the user was gazing at the correct street, a smartphone in the user's pocket vibrated to indicate that the street currently gazed at was the correct one to follow.

4. Special Issue: Eye Tracking for Spatial Research

The growing interest in eye tracking as a research tool in spatial cognition and related fields was the motivation for this Special Issue. It was preceded by two workshops with the same title (Eye Tracking for Spatial Research, ET4S 2013 and 2014), co-located with the Conference on Spatial Information Theory 2013 and the GIScience conference 2014, respectively. Nineteen manuscripts were submitted to an open Call for Submissions, out of which seven were finally accepted after a rigorous review process.

In the first part of the double-issue, recording and analysis of eye movements pertains to data collected on a 2D surface such as a typical computer monitor. Franke and Schweikart (2017) investigate different cartographic visualization methods: vignette, icon, and text. Netzel et al. (2017) examine performance and reading strategies for metro maps. They find that readers often check themselves by transitioning between a route's start and end points. Dupont et al. (2017) compare scanpaths collected over photographs of landscapes with different degrees of urbanization. They compute the photographs' spectral entropies as a measure of image complexity and find that more complex landscapes lead to a more extensive and dispersed visual exploration.

In the second part of the double-issue, recording and analysis of eye movements pertains to data collected when navigating in 3D physical or virtual environments. Ohm et al. (2017) evaluate indoor pedestrian navigation interfaces when navigating inside a building, tabulating which landmarks (such as doors) are used. Their goal is to identify the most functional landmarks, which they can then incorporate in map designs to be used on mobile devices. Wenczel et al. (2017) also consider the visual use of landmarks, but they do so in an outdoor environment (a cemetery). They compare and contrast intentional and incidental learning, with gaze falling on landmarks at structurally salient locations when learning to navigate a route intentionally, and gaze falling elsewhere when learning incidentally. Schwarzkopf et al. (2017) consider gaze of two people when navigating in the real world. They consider the task of collaborative wayfinding and introduce gaze angle analysis, that is, the angle between movement and gaze directions. They find that partners tend to "share space" when wayfinding collaboratively (instead of splitting the space if they were enacting a division of labor). Through gaze analysis, they show that this makes sense intuitively as the two cohorts share landmarks and can therefore exploit deictic references, e.g., "See that clock tower?", "Yes! I see it!". This form of cognitive ethology is a promising approach to the evaluation of shared tasks. Finally, Schrom-Feiertag et al. (2017) evaluate indoor guidance (signage) within an Immersive Virtual Environment. They demonstrate how gaze adds additional insights to the analysis of users' position, locomotion, and view direction within the environment. They provide a heatmap visualization of aggregated gaze on objects looked at in the environment.

5. Future Directions in Spatial Eye-Tracking Research

Improvements in accuracy, frequency, dependability, usability, and mobility of eye trackers have opened up new opportunities for cognitive science, GIScience, and cartography (see section 2), as well as for (mobile) human-computer interaction (see section 3.2). Based on the assumption that future developments in the field will again be driven by technological developments, we cautiously dare to present an outlook on future directions in eye tracking for spatial research.

First, we posit that the potential of mobile eye tracking for *real-world studies on navigation and wayfinding* has only started to be exploited. Two main challenges can be identified here: first, the processing of mobile eye-tracking data is still laborious (no automatic mapping of gaze to the real world by current standard software packages). It can be expected that this will soon be resolved through the use of computer vision techniques. Second, studies in the real world are more realistic but harder to control. This trade-off between external and internal validity is a challenge for many psychological studies (e.g., Anderson et al., 1999), but arguably even more difficult for eye-tracking studies because small changes in a situation or stimulus may alter visual behavior significantly. Though we observe an ongoing trend in combining virtual reality technology (VR) and eye tracking, the external validity of studies of visual attention during wayfinding in VR is yet to be explored.

Second, not all wayfinding processes have been studied with eye tracking to the same extent. For instance, research on *spatial knowledge acquisition and learning* can benefit from the analysis of eye-tracking data recorded during a spatial task. For example, during pedestrian navigation, gaze patterns could indicate that a user does not pay attention to the environment or certain objects, serving as an explanation for decreased spatial learning. Novel gaze-based approaches, such as GazeNav (Giannopoulos et al., 2015), can counter this effect. As mentioned in section 2.2, it may be difficult to distinguish in a realistic scenario which eye movement is caused by which navigational process. We hope to see *models of wayfinding* in the future which are both comprehensive and perceptually grounded. One suggestion for such a modeling approach, based on formal grammars and reference system transformations, has been suggested recently by Kiefer, Scheider, et al. (2015).

As can be seen in the literature review in section 3.1, eye tracking has become an established methodology in cartography. Still, we have only just begun to understand the principles of visual attention during map perception. For instance, the influence of context of map use, as well as of user groups other than novice/experts requires further research.

We suggest that spatial cognition research should take better advantage of *advanced models and measures for eye movement analysis* which go beyond the classic fixation, saccade, or scanpath measures. For instance, a spatial cognition researcher might be interested in the interplay of ambient and focal attention (Krejtz et al., 2016), or in the complexity of switching patterns between Areas Of Interest (Krejtz et al., 2015). For cognitive load, pupil dilation could be considered as an additional measure which is also accessible in real-time for interaction purposes (Kiefer et al., 2016). Regarding standard measures and algorithms (such as fixation computation), we expect and hope to see a trend towards more reproducibility. Researchers should be aware that results may heavily depend on the algorithms and parametrizations used, and therefore include these in their articles. This calls for eye-tracking vendors to openly communicate which algorithms their software packages are using.

The increasing *pervasiveness of eye-tracking technology* is a trend actively discussed in the computer science community (see, e.g., the Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction, Kiefer, Zhang, et al., 2015). Which kinds of interaction with spatial information will become possible if eye tracking is included in every notebook, public screen, and head-mounted display for affordable prices? Will it be possible to interact with urban spaces through one's gaze, such as querying tourist information by looking at a building

(Anagnostopoulos & Kiefer, 2016)? The pervasiveness of eye tracking would also enable the collection of “big” datasets of users’ visual attention while interacting with web maps—an interesting opportunity for cartographic research. The privacy threats of pervasive eye tracking, however, cannot be neglected and constitute a challenge that may potentially benefit from approaches suggested for geo-privacy (e.g., Krumm, 2009).

Overall, the cognitive and the computing perspective on eye tracking in spatial research should be combined more strongly in order to build systems that assist the user based on a firm understanding of human spatial cognition. This idea of *spatial cognitive engineering* has been postulated previously (Raubal, 2009) but is particularly compelling for gaze-based interactive systems.

Acknowledgements

This special issue would not have been possible without the great expertise of the reviewers and the support of the responsible Editor in Chief, Dan Montello. Special thanks go to the other co-organizers of the “Eye Tracking for Spatial Research” (ET4S) workshops: Mary Hegarty (2013), and Antonio Krüger (2014). This work was supported by ETH Research Grant ETH-38 14-2 and by the Swiss National Science Foundation (grant number 200021_162886).

References

Alaçam, Ö., & Dalcı, M. (2009). *A usability study of WebMaps with eye tracking tool: the effects of iconic representation of information.* International Conference on Human-Computer Interaction, Springer, pp. 12-21

Anagnostopoulos, V.-A., & Kiefer, P. (2016). Towards Gaze-Based Interaction with Urban Outdoor Spaces *6th International Workshop on Eye Tracking and Mobile Eye-Based Interaction (PETMEI 2016)*: ACM. to appear

Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory truth or triviality? *Current directions in psychological science, 8*(1), 3-9.

Andrienko, G., Andrienko, N., Demsar, U., Dransch, D., Dykes, J., Fabrikant, S. I., . . . Tominski, C. (2010). Space, time and visual analytics. *International Journal of Geographical Information Science, 24*(10), 1577-1600.

Bektaş, K., & Çöltekin, A. (2011). An Approach to Modeling Spatial Perception for Geovisualization. *Procedia - Social and Behavioral Sciences, 21*, 53-62.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2014). *State-of-the-art of visualization for eye tracking data.* Eurographics Conference on Visualization (EuroVis), The Eurographics Association

Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE transactions on pattern analysis and machine intelligence, 35*(1), 185-207.

Brodersen, L., Andersen, H. H., & Weber, S. (2001). *Applying eye-movement tracking for the study of map perception and map design*: National Survey and Cadastre Denmark.

Brügger, A., Fabrikant, S. I., & Çöltekin, A. (2016). An empirical evaluation of three elevation change symbolization methods along routes in bicycle maps. *Cartography and Geographic Information Science*, 1-16. 10.1080/15230406.2016.1193766

Brunyé, T. T., & Taylor, H. A. (2009). When goals constrain: Eye movements and memory for goal - Oriented map study. *Applied Cognitive Psychology, 23*(6), 772-787.

Brychtova, A., & Çöltekin, A. (2016). An empirical user study for measuring the influence of colour distance and font size in map reading using eye tracking. *The Cartographic Journal*. 10.1179/1743277414Y.0000000103

Castner, H. W., & Eastman, R. J. (1984). Eye-movement parameters and perceived map complexity—I. *The American Cartographer, 11*(2), 107-117.

Chang, K.-T., Antes, J., & Lenzen, T. (1985). The effect of experience on reading topographic relief information: Analyses of performance and eye movements. *The Cartographic Journal, 22*(2), 88-94.

Çöltekin, A., Demsar, U., Brychtová, A., & Vandrol, J. (2014). *Eye-hand coordination during visual search on geographic displays.* Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research (ET4S 2014), Vienna, Austria, CEUR, retrieved from http://ceur-ws.org/Vol-1241/

Çöltekin, A., Heil, B., Garlandini, S., & Fabrikant, S. I. (2009). Evaluating the effectiveness of interactive map interface designs: a case study integrating usability metrics with eye-movement analysis. *Cartography and Geographic Information Science, 36*(1), 5-17.

Davies, C., Fabrikant, S. I., & Hegarty, M. (2015). Towards empirically verified cartographic displays. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman & J. L.

Szalma (Eds.), *Cambridge Handbook of Applied Perception Research* (pp. 711-729): Cambridge University Press.

Dobson, M. W. (1977). Eye movement parameters and map reading. *The American Cartographer*, *4*(1), 39-58.

Dong, W., Liao, H., Xu, F., Liu, Z., & Zhang, S. (2014). Using eye tracking to evaluate the usability of animated maps. *Science China Earth Sciences, 57*(3), 512-522.

Downs, R. M., & Stea, D. (1977). The world in the head *Maps in Minds: Reflections on Cognitive Mapping* (pp. 125-135): Harper & Row.

Dupont, L., Ooms, K., Duchowski, A., Antrop, M., & Van Eetvelde, V. (2017). Investigating the visual exploration of the rural-urban gradient using eye-tracking. *Spatial Cognition and Computation, 17*(1-2).

Emo, B. (2012). *Wayfinding in real cities: experiments at street corners.* International Conference on Spatial Cognition, Springer, pp. 461-477

Fabrikant, S. I., Hespanha, S. R., & Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. *Annals of the Association of American Geographers, 100*(1), 13-29.

Fabrikant, S. I., & Lobben, A. (2009). Introduction: Cognitive issues in geographic information visualization. *Cartographica: The International Journal for Geographic Information and Geovisualization, 44*(3), 139-143.

Fabrikant, S. I., Rebich-Hespanha, S., Andrienko, N., Andrienko, G., & Montello, D. R. (2008). Novel method to measure inference affordance in static small-multiple map displays representing dynamic processes. *The Cartographic Journal*, *45*(3), 201-215.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision research*, *51*(17), 1920-1931.

Franchak, J. M., & Adolph, K. E. (2010). Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vision research, 50*(24), 2766-2774.

Franke, C., & Schweikart, J. (2017). Mental Representation of Landmarks on Maps - Investigating Cartographic Visualization Methods with Eye Tracking Technology. *Spatial Cognition and Computation*, *17*(1-2).

Giannopoulos, I., Kiefer, P., & Raubal, M. (2012). GeoGazemarks: providing gaze history for the orientation on small display maps *Proceedings of the 14th ACM International Conference on Multimodal Interaction* (pp. 165-172): ACM.

Giannopoulos, I., Kiefer, P., & Raubal, M. (2015). GazeNav: gaze-based pedestrian navigation *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI)* (pp. 337-346): ACM.

Giannopoulos, I., Kiefer, P., Raubal, M., Richter, K.-F., & Thrash, T. (2014). *Wayfinding decision situations: A conceptual model and evaluation.* International Conference on Geographic Information Science, Springer, pp. 221-234

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics, 24*(6), 631-645.

Golledge, R. G. (1999). Human wayfinding and cognitive maps. *Wayfinding behavior: Cognitive mapping and other spatial processes*, 5-45.

Gorry, P. A. (1990). General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Analytical Chemistry, 62*(6), 570-573.

Gunzelmann, G., Anderson, J. R., & Douglass, S. (2004). Orientation tasks with multiple views of space: Strategies and performance. *Spatial Cognition and Computation*, *4*(3), 207-253.

Hegarty, M., & Just, M.-A. (1993). Constructing mental models of machines from text and diagrams. *Journal of memory and language, 32*(6), 717-742.
Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: OUP Oxford.
Incoul, A., Ooms, K., & De Maeyer, P. (2015). Comparing paper and digital topographic maps using eye tracking *Modern Trends in Cartography* (pp. 339-356): Springer.
Jacob, R. J. K., & Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 573-605). Amsterdam, The Netherlands: Elsevier Science.
Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology, 8*(4), 441-480.
Kiefer, P., Giannopoulos, I., Duchowski, A. T., & Raubal, M. (2016). Measuring Cognitive Load for Map Tasks through Pupil Diameter *Proceedings of the Ninth International Conference on Geographic Information Science (GIScience 2016)*: Springer. to appear
Kiefer, P., Giannopoulos, I., & Raubal, M. (2013). Using eye movements to recognize activities on cartographic maps *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 488-491): ACM.
Kiefer, P., Giannopoulos, I., & Raubal, M. (2014). Where am I? Investigating map matching during self - localization with mobile eye tracking in an urban environment. *Transactions in GIS, 18*(5), 660-686.
Kiefer, P., Scheider, S., Giannopoulos, I., & Weiser, P. (2015). A wayfinding grammar based on reference system transformations. In S. I. Fabrikant, M. Raubal, M. Bertolotto, C. Davies, S. Freundschuh & S. Bell (Eds.), *Spatial Information Theory* (pp. 447-467): Springer.
Kiefer, P., Zhang, Y., & Bulling, A. (2015). *The 5th international workshop on pervasive eye tracking and mobile eye-based interaction*. Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, Osaka, Japan. pp. 825-828
Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision research, 35*(13), 1897-1916.
Kraak, J.-M. (2003). Cartographic principles. In J.-M. Kraak & A. Brown (Eds.), *Web cartography - developments and prospects*. London, New York: Taylor&Francis.
Krassanakis, V., Filippakopoulou, V., & Nakos, B. (2016). Detection of moving point symbols on cartographic backgrounds. *Journal of Eye Movement Research, 9*(2).
Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., & Kopacz, A. (2016). Discerning Ambient/Focal Attention with Coefficient K. *ACM Transactions on Applied Perception (TAP), 13*(3), 11.
Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., González Perilli, F., Pires, A., . . . Villalobos, N. (2015). Gaze transition entropy. *ACM Transactions on Applied Perception (TAP), 13*(1), 4.
Krumm, J. (2009). A survey of computational location privacy. *Personal and Ubiquitous Computing, 13*(6), 391-399.
Land, M., & Tatler, B. (2009). *Looking and acting: vision and eye movements in natural behaviour*: Oxford University Press.

Land, M. F., & Lee, D. N. (1994). Where do we look when we steer. *Nature, 369*, 742-744. 10.1038/369742a0

Liao, H., Dong, W., Peng, C., & Liu, H. (2016). Exploring differences of visual attention in pedestrian navigation when using 2D maps and 3D geo-browsers. *Cartography and Geographic Information Science*, 1-17.

Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language *Multisensory imagery* (pp. 131-155): Springer.

MacEachren, A. M. (1995). *How maps work: representation, visualization, and design*: Guilford Press.

Maggi, S., Fabrikant, S. I., Imbert, J.-P., & Hurter, C. (2016). How Do Display Design and User Characteristics Matter in Animations? An Empirical Study with Air Traffic Control Displays 1. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *51*(1), 25-37.

Meilinger, T., Hölscher, C., Büchner, S. J., & Brösamle, M. (2006). *How much information do you need? Schematic maps in wayfinding and self localisation.* International Conference on Spatial Cognition, Springer, pp. 381-400

Montello, D. R. (2002). Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. *Cartography and Geographic Information Science, 29*(3), 283-304.

Montello, D. R. (2005). Navigation. In P. Shah & A. Miyake (Eds.), *The Cambridge Handbook of visuospatial thinking* (pp. 257-294). New York: Cambridge University Press.

Netzel, R., Ohlhausen, B., Kurzhals, K., Woods, R., Burch, M., & Weiskopf, D. (2017). User Performance and Reading Strategies for Metro Maps: An Eye Tracking Study. *Spatial Cognition and Computation*, *17*(1-2).

Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American*.

Ohm, C., Müller, M., & Ludwig, B. (2017). Evaluating Indoor Pedestrian Navigation Interfaces Using Mobile Eye Tracking. *Spatial Cognition and Computation*, *2017*(1-2).

Ooms, K., Çöltekin, A., De Maeyer, P., Dupont, L., Fabrikant, S., Incoul, A., . . . Van der Haegen, L. (2015). Combining user logging with eye tracking for interactive and dynamic applications. *Behavior research methods, 47*(4), 977-993.

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E., & Witlox, F. (2012a). Interpreting maps through the eyes of expert and novice users. *International Journal of Geographical Information Science, 26*(10), 1773-1788.

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E., & Witlox, F. (2012b). Investigating the effectiveness of an efficient label placement method using eye movement data. *The Cartographic Journal*, *49*(3), 234-246.

Ooms, K., Dupont, L., & Lapon, L. (2016). Mixing Methods and Triangulating Results to Study the Influence of Panning on Map Users' Attentive Behaviour. *The Cartographic Journal*, 1-18.

Opach, T., Gołębiowska, I., & Fabrikant, S. I. (2014). How do people view multi-component animated maps? *The Cartographic Journal*, *51*(4), 330-342.

Peebles, D., Davies, C., & Mora, R. (2007). *Effects of geometry, landmarks and orientation strategies in the 'drop-off'orientation task.* International Conference on Spatial Information Theory, Springer, pp. 390-405

Popelka, S., & Brychtova, A. (2013). Eye-tracking study on different perception of 2D and 3D terrain visualisation. *The Cartographic Journal*, *50*(3), 240-246.

Popelka, S., & Doležalová, J. (2015). Non-photorealistic 3D visualization in city maps: an eye-tracking study *Modern Trends in Cartography* (pp. 357-367): Springer.

Raubal, M. (2009). Cognitive engineering for geographic information science. *Geography Compass, 3*(3), 1087-1104.

Raubal, M., & Winter, S. (2002). *Enriching wayfinding instructions with local landmarks.* International Conference on Geographic Information Science, Springer, pp. 243-259

Rieser, J. J. (1999). Dynamic spatial orientation and the coupling of representation and action. *Wayfinding behavior: Cognitive mapping and other spatial processes*, 168-190.

Rohs, M., Schöning, J., Raubal, M., Essl, G., & Krüger, A. (2007). *Map navigation with mobile devices: virtual versus physical movement with and without visual context.* Proceedings of the 9th international conference on multimodal interfaces, ACM, pp. 146-153

Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry, 36*(8), 1627-1639.

Schmidt, A. (2000). Implicit human computer interaction through context. *Personal technologies, 4*(2-3), 191-199.

Schrom-Feiertag, H., Settgast, V., & Seer, S. (2017). Evaluation of Indoor Guidance Systems using Eye Tracking in an Immersive Virtual Environment. *Spatial Cognition and Computation, 17*(1-2).

Schwarzkopf, S., Buechner, S., Hölscher, C., & Konieczny, L. (2017). Perspective tracking in the real world. Gaze angle analysis in a collaborative wayfinding task. *Spatial Cognition and Computation, 17*(1-2).

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S., & Hedley, N. R. (2001). Cognitive and usability issues in geovisualization. *Cartography and Geographic Information Science, 28*(1), 61-75.

Spiers, H. J., & Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of environmental psychology, 28*(3), 232-249.

Steinke, T. R. (1987). Eye movement studies in cartography and related fields. *Cartographica: The International Journal for Geographic Information and Geovisualization, 24*(2), 40-73.

Stellmach, S., & Dachselt, R. (2012). *Investigating gaze-supported multimodal pan and zoom.* Proceedings of the Symposium on Eye Tracking Research and Applications, ACM, pp. 357-360

Stofer, K., & Che, X. (2014). Comparing Experts and Novices on Scaffolded Data Visualizations using Eye-tracking. *Journal of Eye Movement Research, 7*(5).

Tanriverdi, V., & Jacob, R. J. (2000). *Interacting with eye movements in virtual environments.* Proceedings of the SIGCHI conference on Human Factors in Computing Systems, ACM, pp. 265-272

Thorndyke, P. W., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology, 12*(1), 137-175.

van Zoest, W., & Donk, M. (2004). Bottom-up and top-down control in visual search. *Perception, 33*(8), 927-937.

Vansteenkiste, P., Cardon, G., D'Hondt, E., Philippaerts, R., & Lenoir, M. (2013). The visual control of bicycle steering: The effects of speed and path width. *Accident Analysis & Prevention, 51*, 222-227.

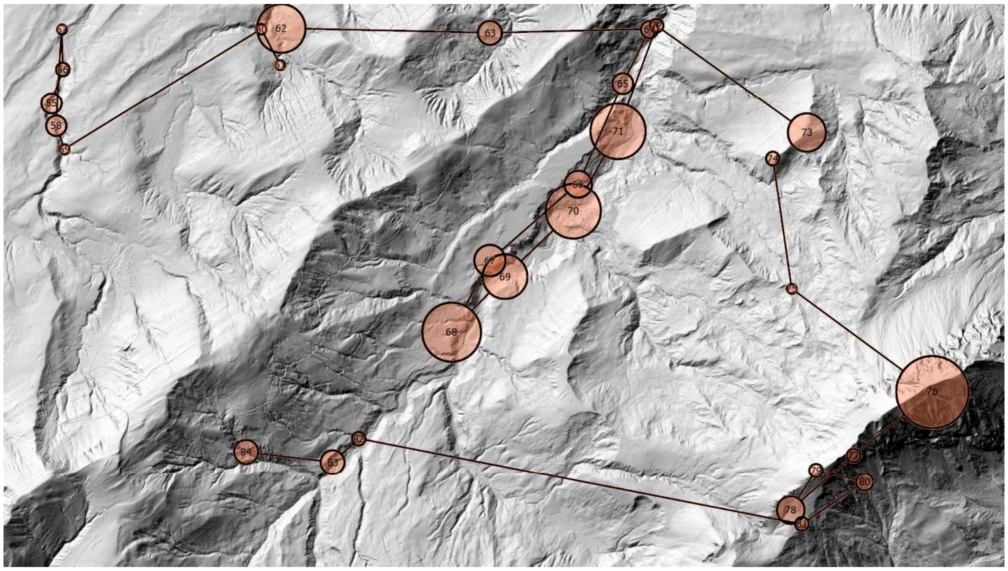
Viaene, P., Vansteenkiste, P., Lenoir, M., De Wulf, A., & De Maeyer, P. (2016). Examining the validity of the total dwell time of eye fixations to identify landmarks in a building. *Journal of Eye Movement Research*, *9*(3). 10.16910/jemr.9.3.4

Wenczel, F., Hepperle, L., & von Stülpnagel, R. (2017). Gaze Behavior during Incidental and Intentional Navigation in an Outdoor Environment. *Spatial Cognition and Computation, 17*(1-2).

Wiener, J. M., Büchner, S. J., & Hölscher, C. (2009). Taxonomy of human wayfinding tasks: A knowledge-based approach. *Spatial Cognition & Computation, 9*(2), 152-165.

Wiener, J. M., Hölscher, C., Büchner, S., & Konieczny, L. (2012). Gaze behaviour during space perception and spatial decision making. *Psychological research, 76*(6), 713-729.

Yarbus, A. L. (1967). *Eye movements during perception of complex objects*: Springer.

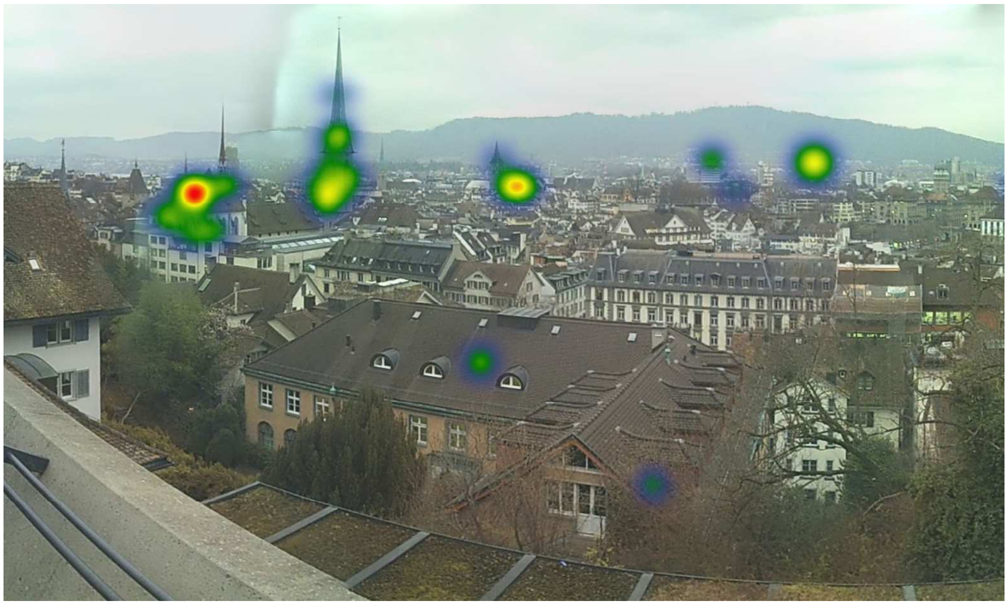


Scanpath of a person on a shaded relief. Circles depict fixations, numbered in the order of occurrence. Lines illustrate saccades. Missing lines between consecutive fixations are caused by blinks.

257x144mm (300 x 300 DPI)

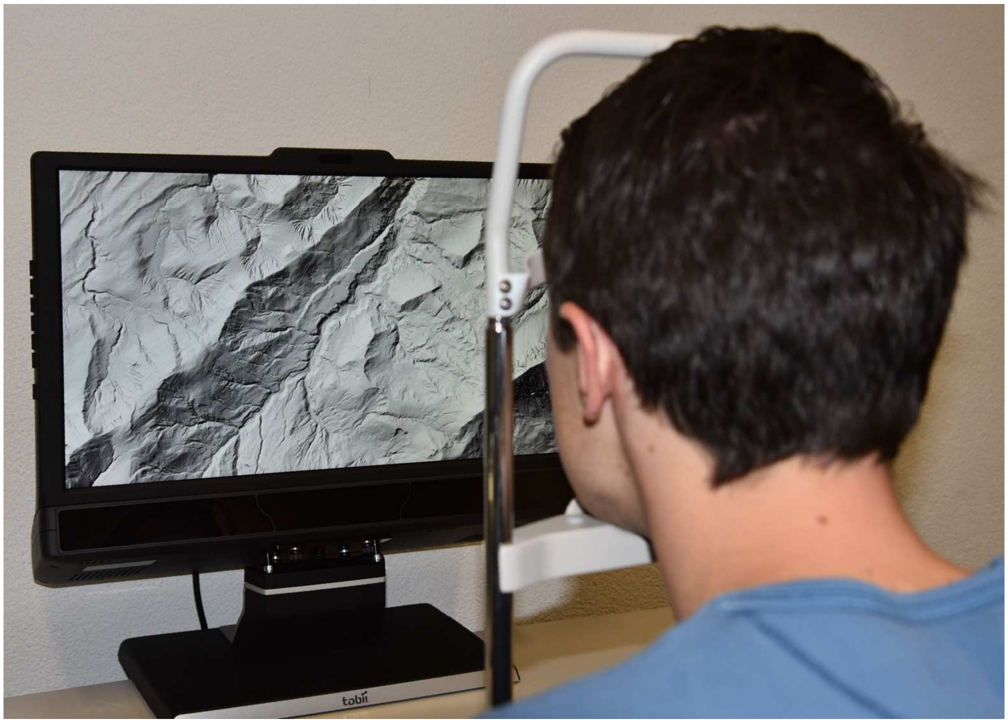
Wayfinding study in the real world with a mobile eye tracker and a map.

257x171mm (300 x 300 DPI)



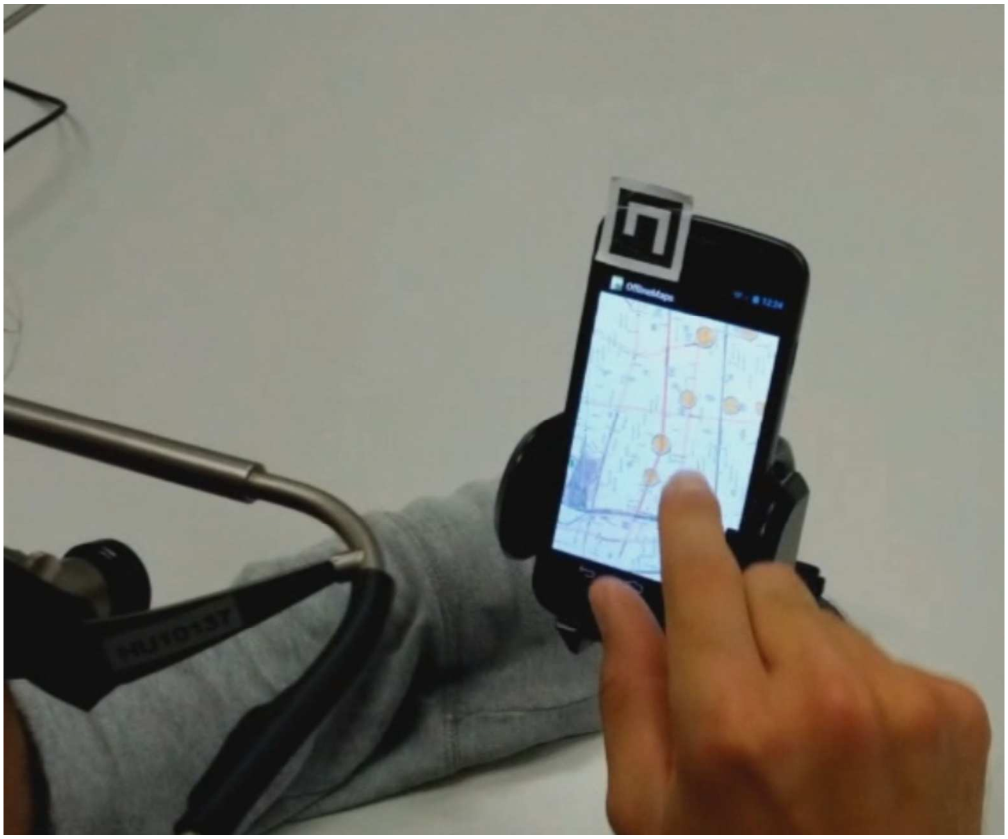
Distribution of visual attention on a city panorama visualized as heatmap (created by a kernel density operation).

257x153mm (300 x 300 DPI)



Remote eye tracking study on a shaded relief.

214x152mm (300 x 300 DPI)



Study setup for gaze-based interaction with an interactive mobile map (refer to Giannopoulos et al., 2012). The user's gaze on the map during the interaction is recorded (in map coordinates), aggregated, and visualized when the user zooms out (circles). These circles are called GeoGazemarks and serve as visual bookmarks on the map.

222x184mm (300 x 300 DPI)