Spatial Cognition & Computation
An Interdisciplinary Journal

ISSN: 1387-5868 (Print) 1542-7633 (Online) Journal homepage: http://www.tandfonline.com/loi/hscc20

Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment

Helmut Schrom-Feiertag, Volker Settgast & Stefan Seer

SPATIAL COGNITION & COMPUTATION
http://dx.doi.org/10.1080/13875868.2016.1228654

Taylor & Francis
Taylor & Francis Group

Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment

Helmut Schrom-Feiertag,[a] Volker Settgast,[b] and Stefan Seer[a]

[a]Mobility Department, Austrian Institute of Technology, Vienna, Austria; [b]Division Visual Computing, Fraunhofer Austria, Graz, Austria

ABSTRACT
In this article, we present a novel method for evaluating guidance systems using an immersive virtual environment in combination with a mobile eye tracking system. Accurate measurements of position, locomotion, viewing frustum, and gaze are captured in the virtual environment. They are applied to the projection of an attention map onto the virtual 3D environment for visualizing the fixation in the environment as well as the amount of time objects were fixated. To demonstrate the method's applicability, we conducted an experiment with 24 participants evaluating a guidance system of a large public infrastructure. The results show that our method allows for the creation of attention maps as well as for the identification of objects of interest based on eye tracking.

KEYWORDS
virtual reality; virtual environments; eye tracking; visual perception; wayfinding

1. Introduction

Transport hubs gradually evolve from train stations towards complex infrastructures including shopping facilities, offices, and hotels. This multitude of functions constitutes a significant challenge for the design of a consistent and accessible guidance system and spatial ordering principle, which both have to meet context-specific user needs. According to Raphael (2006) the locations for visual signs need to be selected carefully such that they 1) do not compete with other signs or elements in the environment; 2) are readily viewable within an adequate decision-making time; 3) are not partially obstructed; and, 4) can be seen within the vertical and horizontal viewing ranges of an average person. Ideally, all these aspects are taken into account at an early stage of planning. It is therefore crucial to understand how people perceive and interact with the environment and how their perception influences orientation and navigation behavior.

For questions related to orientation and navigation, eye tracking has become a popular method to investigate human wayfinding and to gain insights into perception and processing of visual information. Research on wayfinding using mobile eye trackers has been conducted within lab and field studies, providing

CONTACT Helmut Schrom-Feiertag helmut.schrom-feiertag@ait.ac.at Austrian Institute of Technology, Giefinggasse 2, Vienna, 1210, Austria.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hscc.

evidence for comparable eye movements in experimental and real world scenarios (Foulsham, Walker, & Kingstone, 2011). The latter study also revealed that the freedom of immersed participants to move their head and body is of high importance. At the same time, although field studies in real-world environments increase the results' external validity, they also impose various constraints for research, such as the lack of control over dynamically changing outdoor conditions, e.g., other people, traffic, noise, and lighting conditions.

Virtual environments represent a controllable alternative to field studies conducted in the real world and have recently also been introduced to research on wayfinding studies (Bertrand et al., 2013; Schrom-Feiertag, Schinko, Settgast, & Seer, 2014; Spiers & Maguire, 2008; Suma et al., 2010). Applications range from studies with simple video playback to interactive computer graphics with one or more monitors and stereo vision. Most advanced applications include immersive virtual environments with projections on three to six walls of a room-sized cube. The person inside the virtual environment wears stereoscopic glasses in order to perceive the projected graphics in 3D. The person can move around in the 3D model, which allows experience of the environment at a high level of realism. Another popular technology with comparable immersion are head-mounted devices that are also suitable for wayfinding studies. The advantages are high resolution, little weight, and portability. Disadvantages refer to a limited field of view, distortions, and blinding out the surrounding real environment. At the same time, the simultaneous usage and interaction with devices such as smartphones or maps for navigation is difficult.

To use immersive virtual environments for wayfinding studies it is essential to understand the adverse effects this technology has on its users. Stanney, Hale, Nahmens, and Kennedy (2003) studied the temporal occurrence of cybersickness when being in an immersive virtual environment with navigational control for 15 to 60 minutes. The results show that the virtual environment caused 6.3% of the participants to discontinue within 15 minutes, 16.9% within 30 minutes and 45.8% within 60 minutes. Thus, to reduce the risk of cybersickness, experiments ought to be kept short or be divided into subsessions.

It is yet unclear to what extent the results obtained from eye tracking in virtual environments can be transferred to reality. Results in Schwarzkopf, von Stülpnagel, Büchner, and Konieczny (2013) show that body orientation and movement have a strong impact on eye tracking patterns in spatial navigation tasks. Findings from Hollands, Patla, and Vickers (2002) reveal the importance of head alignment and heading. In Renner, Velichkovsky, and Helmert (2013) a survey of empirical studies focusing on the perception of egocentric distances in virtual environments indicate egocentric distance to be shorter, with a mean of approximately 74% of the modeled distances. The main influencing factors to facilitate a veridical spatial perception are binocular disparity, high quality graphics, a rich environment, and an enhanced user's sense of presence. Another relevant aspect in the virtual environment is the participants' awareness of their own body (proprioception).

To investigate a person's eye movement in the virtual environment, mobile eye tracking can be applied, which enables the individual to move freely inside the physical space of a virtual environment and to continuously vary perspectives and with this the perception of the environment. Schwarzkopf et al. (2013); Brône, Oben, and Goedem (2011) argue that this has an impact on the decisions and path selections.

Tracking people's eye movement while moving freely, however, poses challenges. Typically gaze behavior is analyzed by mapping gaze on reference images from a similar perspective. On such images, areas or objects of interest are defined and participants' attention towards these areas are quantified. Using a mobile eye tracking system it is necessary to estimate the head pose with position and orientation as well as gaze direction in 3D world coordinates.

Several methods have been investigated to estimate gaze in 3D world coordinates. In Munn and Pelz (2008) a method is proposed where the 3D Point Of Regard (POR) is estimated from a portable monocular video-based eye tracker applying computer vision techniques to get 3D structure and motion from video sequences. Pirri, Pizzoli, and Rudi (2011) proposed a model-based approach for 3D gaze estimation for wearable multicamera devices. The approach is grounded on multiple view geometry and introduces an efficient, accurate, and practical calibration procedure. Takemura, Kohashi, Suenaga, Takamatsu, and Ogasawara (2010) combined a scene camera with the eye tracking system and used a visual simultaneous localization and mapping (Visual SLAM) approach to estimate the 3D POR in real-time. The weak point of this method was that the method cannot estimate the scale of world coordinates. In Takemura, Takahashi, Takamatsu, and Ogasawara (2014) a 3D environment map was created to detect objects of focus. The 3D POR was estimated in world coordinates by a matching process between Visual SLAM and the 3D environment map. The final attention map can be visualized as an overlay onto the 3D environment map. The advantage of this approach was that commercial eye tracking systems can be utilized to estimate 3D POR.

Such approaches have led to considerable progress in 3D gaze analysis and visualization, especially for real-world environments. In this article we present our new approach that combines an immersive virtual environment with a mobile eye tracking system for 3D gaze analysis and visualization. To demonstrate the applicability of our approach we conducted an experiment evaluating a guidance system of a large public transport infrastructure. In the experiment, we evaluated the completeness and usage of signage during wayfinding-related tasks. Next, we report on the method used for 3D gaze analysis. Afterwards we describe the experiment conducted and discuss the results. Finally, we conclude with our findings and give an outlook for future research.

2. 3D gaze analysis method

Here, we describe our approach, which combines an immersive virtual environment equipped with an embedded interface for natural locomotion together with a mobile eye tracking system for 3D gaze analysis. In the virtual environment all relevant activities, movements, and user interactions are recorded and available for detailed analysis. Given the highly accurate data of participants' positions, body alignment, and viewing directions, together with the gaze point from the eye tracking system, it is possible to compute the intersection between the viewing ray and the triangle mesh of the 3D model for each video frame. This enables projection of an attention map onto the virtual 3D environment, to automatically create heatmaps that visualize the distribution of fixations in the environment as well as to identify fixated objects and fixation duration.

2.1. Virtual environment

The immersive virtual environment DAVE (Lancelle et al., 2009), as shown in Figure 1, is an immersive four-sided Cave Automatic Virtual Environment (CAVE), which is built and located at the Graz University of Technology, Austria. DAVE stands for Definitely Affordable Virtual Environment where affordable means that by mostly using standard hardware components costs compared to other commercial systems are greatly reduced. It was used to efficiently perform wayfinding studies with enhanced comparability of the results due to controlling and reproducing equal testing conditions (e.g., time, passers-by, train and bus schedule, ambient sound, etc.) for each participant.

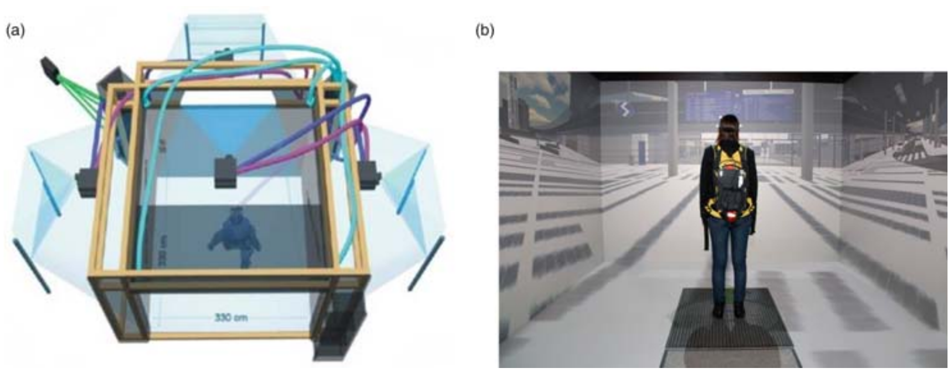

Figure 1. (a) The DAVE, a four-sided immersive stereoscopic projection environment and (b) a study participant inside the DAVE.

The DAVE consists of three rear-projected active stereo screens (left, right, and front wall on which the images are projected from outside) and a front-projected screen on the floor (image for the floor is projected from above). Large mirrors are used to fold the light paths from the projectors to the screens in order to minimize the necessary space requirements. The walkable area is a square with an edge length of 3.30 m. For each of the projection screens we use a Projectiondesign F35 AS3D DLP projector with a resolution of 1920 × 1080 pixels. The integrated XPort DCC120 module creates synchronized time sequential stereo vision displayed with 60 Hz for each eye. Active stereo glasses are triggered via infrared pulses. A dynamic asymmetric view frustum allows provision of undistorted stereoscopic imagery to the participant.

2.2. Eye tracking

For eye tracking, a mobile system from SMI SensoMotoric Instruments (www.smivision.com) equipped with active snap-on 3D shutter glasses for realistic 3D experience and full immersion was used. Four passive reflecting markers are attached to the shutter glasses capturing the position and orientation of the head in order to determine the field of view of the participant. The head position tracking is performed using an optical tracking system with approximately 45 Hz and consists of four cameras with attached infrared emitters mounted in the four upper corners.

Figure 2a shows a participant equipped with the mobile eye tracking system in the virtual environment DAVE during an experiment. The corresponding eye tracking video indicates the gaze point with an orange circle as illustrated in Figure 2b.

2.3. Natural locomotion interface

Our approach embeds an interface for natural locomotion in the virtual environment offering significant advantages related to the user's spatial perception as well as physical and cognitive demands (Suma et al., 2010).

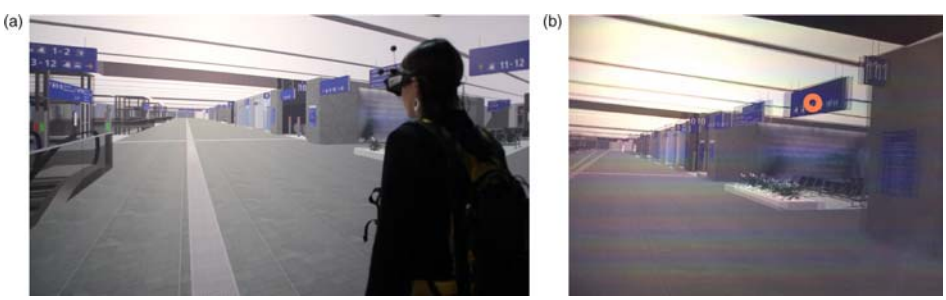

Figure 2. (a) Realistic visual appearance of the 3D model in the DAVE for the participant equipped with mobile eye tracking and (b) the gaze point (orange circle) as an overlay in the eye tracking video.

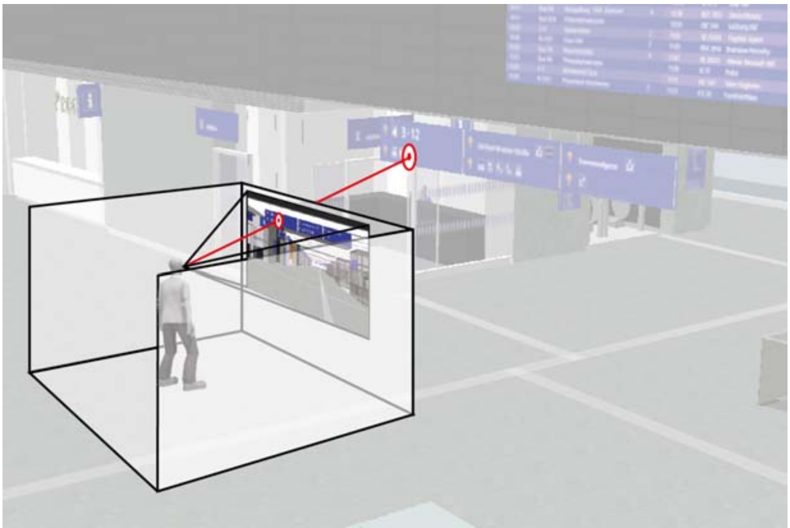

Figure 3. Scheme for computing the intersection of 3D gaze vectors with the 3D model.

Therefore, a Microsoft Kinectis installed on top of the front projection screen to capture the participants' movements. It is used for hands-free navigation and interaction, described in detail in Bauer, Schneckenburger, Settgast, Millonig, & Gartner (2013). The user has to walk in place to move forward and to turn the shoulders to invoke rotations in the virtual world. Within the immersive environment DAVE, all further activities such as body movements, user interactions, and thinking aloud are recorded to be available for detailed analysis.

2.4. 3D gaze mapping

For investigations of human decision making the viewed objects and signs had to be identified. By combining the accurate viewing frustum data with the gaze point data from the eye tracking system, it is possible to compute the intersection between the viewing ray and the triangle mesh of the 3D model for each frame. The eye tracking data is analyzed with the SMI BeGaze™eye

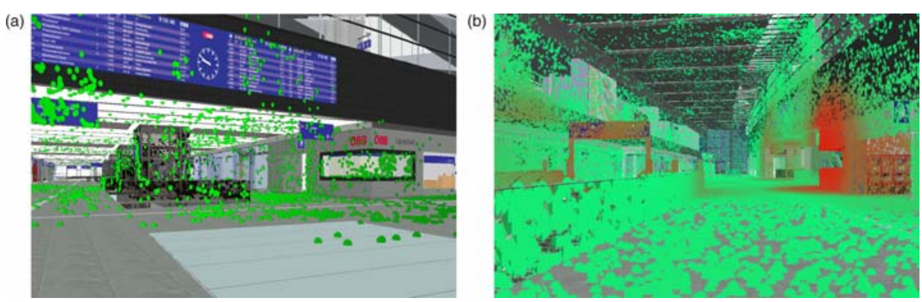

Figure 4. Gaze mapping, (a) example mapping of gaze points at the 3D model, (b) aggregated intersections to map the distribution of attention.

tracking analysis software (SMI SensoMotoric Instruments, Teltow, Germany), which provided the raw data for the intersection processing. The scheme for the 3D gaze mapping is shown in Figure 3 and the result of one participant is presented in Figure 4a.

This enables us to project an attention map onto the virtual 3D environment and furthermore to create maps for visualizing the distribution of fixations in the environment as well as the amount of time any particular object was fixated. The intersections between the viewing rays of participants and the 3D mesh of the environment are aggregated using Cloud Compare, a 3D point cloud and mesh processing software and visualized as volume density maps, as shown in Figure 4b. The volume density maps can be explored dynamically navigating through the 3D model of the environment using an in-house developed 3D viewer software.

2.5. *Visual analysis tool*

Similar to Bertrand et al. (2013) a visual analysis tool was developed to facilitate quantitative and qualitative gaze investigations. The analysis tool enables playback of the captured data in real time and provides a timeline for quick access to a specific timestamp. The trajectory data, viewing frustum, the raw data export of BeGaze, and the eye tracking video are loaded and visualized in a synchronous manner (Figure 5). The camera is aligned to the recorded viewing frustum, and the visualized scene matches the view in the eye tracking video.

Figure 5. Gaze analysis tool with real time replay of recorded experiment data for in-depth study of wayfinding behavior.

Gaze points are visualized with semitransparent spheres. To investigate the saccade in the virtual environment a time interval can be selected until it covers the saccade of interest and to get some understanding of the temporal and spatial aspect of the saccades by moving the selection forward in time. The synchronized playback of the eye tracking video (left upper corner in Figure 5) serves for the inspection of the viewing frustum accuracy. Together with the eye tracking video, audio comments from the participants are recorded simultaneously. These synchronized audio recordings allow a deeper analysis of when, what, and why a participant was looking for, the participant's understanding of signs, and the mental wayfinding processes.

For the statistical analysis of gaze and attention, 290 signs were semantically annotated with respect to type (sign, timetable, or monitor) and in terms of the information it provides (e.g., WC, train platform number, information desk, etc.) and stored in a file. These annotations allow us to associate computed intersections to signage objects. With this association it is possible to analyze the attention to signage. It further allows the evaluation of the duration of fixation on each sign, gaze sequences across signs, and the positions from were participants who looked at the signs.

3. Experiment

To demonstrate the applicability of our approach we conducted an experiment evaluating a guidance system of a large public transport infrastructure. In the experiment, we evaluated the completeness and usage of signage during wayfinding-related tasks.

The main goal of the experiment was to investigate the applicability of an immersive virtual environment with embedded interfaces for natural locomotion combined with mobile eye tracking systems for wayfinding studies. The scope of this experiment was human wayfinding in unfamiliar buildings, considering individual pedestrian movement and gaze behavior in an immersive virtual environment. In the experiment participants were instructed to find four destinations in an immersive virtual environment of Vienna's main railway station. The experiment took place during the railway station's construction phase one year before its opening.

3.1. Methods

3.1.1. Participants

In total, 24 participants took part in the experiments, 12 females and 12 males. The ages of the participants ranged from 22 to 75 years ($M = 49$ years, $SD = 22$ years) and all reported normal or corrected-to-normal vision. None of the participants had prior knowledge about the selected test environment. The test persons were reimbursed 40 Euros for their participation in the study.

Figure 6. Waypoints of the wayfinding experiment in the main train station of Vienna and shortest route connecting these waypoints.

3.1.2. *Apparatus*

Experiments were conducted using an immersive virtual environment (DAVE) (Lancelle et al., 2009), as shown in Figure 1. The 3D model was a large-scale, life-sized virtual replica of the main railway station in Vienna, which reflected the guidance system based on the plan of October 2013.

3.1.3. *The virtual environment*

The 3D model of the main train station of Vienna consisted of more than 2,000,000 triangles, 80 MB of compressed textures, and an overall virtual size of 4 km^2. The railway station is approximately 150 m × 350 m wide. It included 3 levels, 22 animated escalators, and 6 interactive elevators. The guidance system consisted of about 290 signs. To enhance the immersion, virtual pedestrians and also the ambient noise typical of train stations were added including announcements of arriving trains.

The planned scenario covered a wide area of the virtual train station and it would have taken 15 to 20 minutes to find and walk along the waypoints. To reduce the risk of cyber-sickness the entire scenario was split into two phases. Each phase was designed such that its completion was possible within 10 minutes under normal circumstances. The area map of the railway station in Figure 6 shows the waypoints of the scenario, which are labeled with the phase number and the waypoint number (phase/waypoint number).

The first phase started at the main north entrance *(Start 1)* of the main station followed by the waypoints connected according to the task assignment by the red line, with buying a ticket at waypoint *1/1* and leaving the luggage at the luggage lockers *1/2*. After completion of the first phase at the luggage lockers *1/2*, the participants had a break for approximately 5 minutes outside of the virtual environment for recovering and prevention of cyber-sickness.

Because we did not expect wayfinding difficulties after returning back from the luggage lockers to the main hall, we left out this path segment to additionally shorten the duration of virtual exposure to the participants and set the starting point of the second phase at position *Start 2* (see Figure 6). The waypoints in phase 2, connected by the orange line, were the restroom at *2/1* and finally the tram stop of line D at waypoint *2/2*. Note that an alternative route (dashed orange line) outside of the railway building was possible.

3.1.4. Procedure

For each participant the experiment started with a short introduction to explain the basic information about the study procedure. Next, the participant was equipped with the mobile eye-tracking system and a 3D eye-tracking calibration was performed (i.e., the participant had to gaze at specific points in the environment and at specifically prepared screens).

For the first-time use of the virtual environment the participant was introduced to the hands-free navigation and had about 5 minutes to get familiar with it during a training session. During the training session, the participant moved freely in the virtual main station, but was placed far away from the area of the later assignments. When the participant reported they were ready for the assignments he/she was "virtually" placed at the starting position (see *Start 1* as illustrated in Figure 6).

The instruction given to the participants was as follows: *You are at Vienna and want to visit the historic building "Belvedere" before traveling back home. By taking line D, the "Belvedere" is only a few tram stops on line D away from the main train station. To get there, first you have to buy a ticket, then leave the luggage at the train station, and then go to the restroom and proceed to the tram stop of line D.*

To assist this wayfinding task in the second phase, the participants were allowed to use a paper map or a smartphone navigation app. Note that the usage of these types of navigation aids would not be possible in virtual environments with head-mounted devices.

Additionally, the participants could also obtain information at the information desks, ticket counters or from passers-by, which were graphically represented and animated with virtual avatars. The experiments were accompanied by a staff member who served as surrogate for the avatars answering questions.

3.1.5. Data analysis

Evaluations were carried out according to the three levels of metrics from Ruddle and Lessels (2006), which are defined as follows:

1. *User task performance* with time required, distance traveled,
2. *Physical behavior* including characteristics of locomotion, looking around, shape of trajectory, and
3. *Decision making* with viewed objects and signs identified via attention maps and thinking aloud.

Because there were no predefined routes in the experiment, the participants were allowed to navigate freely inside the environment. Therefore time and distance were chosen to measure the participants' wayfinding performance. As mentioned in Ruddle and Lessels (2006), interpreting the results of performance metrics must be done carefully. Some of the participants spent a considerable amount of time for orientation and planning where to travel, yet other users traveled longer distances within the same time.

From the experiments in the virtual environment, we can obtain highly accurate measurements on position, body orientation, viewing direction, and viewing frustum of the participants. Using these data sets, the individuals' physical behavior consisting of locomotion and trajectory shape was evaluated. Metrics for characterizing locomotion are walking speed, duration, and location of stops. Results are visualized in a heatmap that depicts main paths and those locations where participants spend more time for orientation and planning.

3.2. Results

3.2.1. User task performance

Of the 24 participants, 20 were considered for analysis. Four participants were not able to finish the experiment for various reasons: Two participants experienced cyber-sickness (female 74 years, male 75 years of age), which refers to a sense of nausea and disorientation similar to seasickness described by Hale and Stanney (2014). This primarily occurs as a result of a conflict among the visual, vestibular, and proprioceptive systems, where one of the systems implies a body position or movement that is not confirmed by the other systems. A third participant had difficulties in fully understanding the task assignments and quit the experiment (female 22 years of age). The fourth participant dropped out of the experiment after a loss of orientation (female 69 years of age).

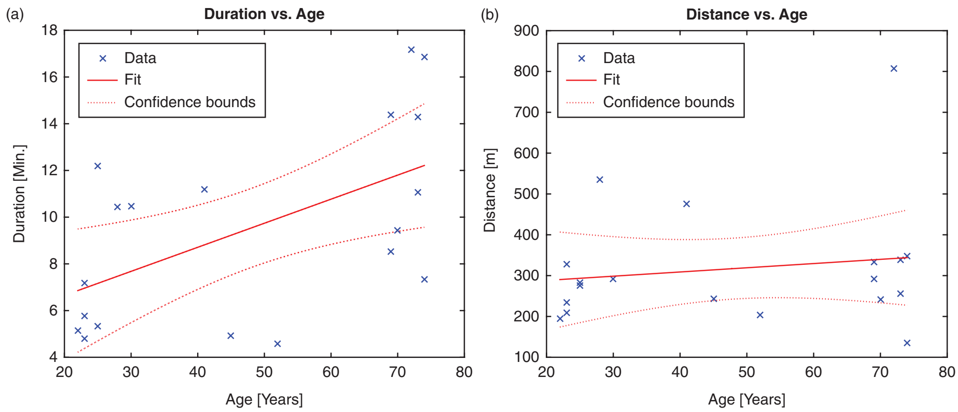

Figure 7. Correlation of (a) duration and (b) distance vs. age.

The remaining 20 participants, 11 males and 9 females, were able to complete all assignments of the experiment. The age of these 20 participants ranged from 22 to 74 years ($M = 48$ years, $SD = 21$ years).

First, we investigated the correlation between the participants' age and the time spent to accomplish the experiment. For the first two tasks in the first phase, which included buying a ticket and finding the luggage lockers, the F-statistics between age and time spent showed a p-value $= 0.0126$ (Figure 7a). The overall duration for these two tasks ranged from 4.57 minutes up to 17.17 minutes ($M = 9.53$ minutes, $SD = 4.10$). Elderly participants needed more time to perform the tasks, which confirms earlier findings (Morganti, Marrakchi, Urban, Iannoccari, & Riva, 2009; Morganti & Riva, 2011). In particular, the required natural locomotion in our experimental setup strongly depends on the physical fitness of participants, which decreases with age. In the second phase, which included finding the restroom and the tram stop of line D, the duration to complete both tasks ranged from 3.07 minutes to 19.32 minutes ($M = 10.82$ minutes, $SD = 4.23$ minutes). Here the correlation dropped down to p-value $= 0.108$. Again, the results show a tendency that the ability to navigate in virtual environments decreases with increasing age of healthy participants.

The measured travel distances in phase 1 ranged from 136 m to 808 m ($M = 310$ minutes, $SD = 152$). In contrast to the duration, the results for the distance do not reveal a correlation associated with age (p-value $= 0.535$) as shown in Figure 7b. Our observations indicate slower walking speeds are the explanation for this result. Also, elderly people sometimes spent more time for orientation before they start walking in order to prevent detours. On average the chosen routes of elderly and younger participants do not show vast differences with respect to the distance. Phase 2 showed similar results with distances from 135 m to 723 m ($M = 399$ minutes, $SD = 151$ minutes). Here the p-value $= 0.70$ also shows no correlation between distance and age.

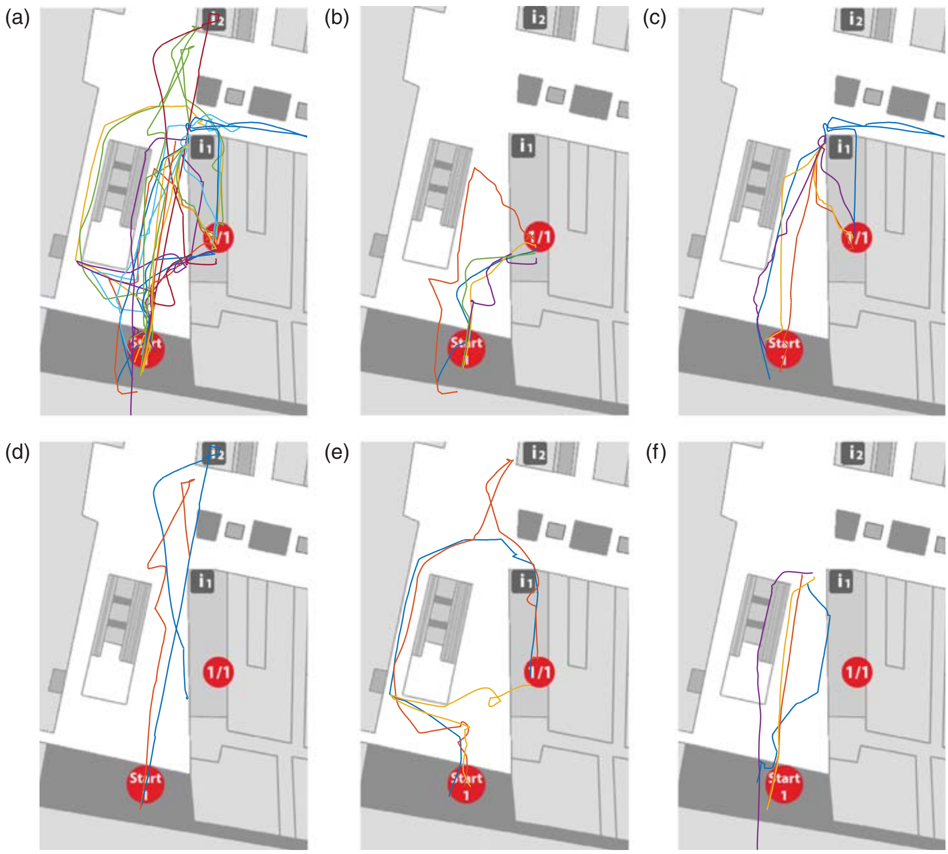

Figure 8. Trajectories for task 1, starting at *Start 1* finding the ticket the counter at *1/1*. (a) all trajectories, (b) finding immediately the ticket counter, (c) asking for the way at *i1*, (d) asking at *i2*, (e) long detour and (f) modified setting with additional ticket counter at *i1*.

3.2.2. Physical behavior

Physical behavior in this study includes characteristics of locomotion and shape of trajectory. To analyze the trajectories, the complete set of trajectories was split into four sequences according to the four task assignments. This facilitates the comparison of the individual task performance between the participants and reduces the complexity of the evaluation.

Figure 8 shows the trajectories measured for the first task. The participants had to start at position *"Start 1"* and to find the ticket counter at position *1/1* to buy a ticket. Figure 8a shows all 19 trajectories combined in one plot. Although this task could have been accomplished via a direct short route, considerable variations in the chosen routes can be observed. Figures 8b to Figure 8f illustrate the different trajectory clusters, which include groups of trajectories with similar shape. The shape of a trajectory is mainly determined by the sequence of intermediate destinations visited by the participant. Figure 8b shows the routes of participants finding the ticket office via a direct path. In 8c participants at first passed the ticket counters located at *1/1*, walked on to the information desk *i1* to retrieve the description of the route to the ticket counters and eventually walked to *1/1*. Figure 8d is similar to Figure 8c with the difference that the participants also passed the first information desk and perceived the second information desk *i2*, which is a little farther on their way. For some participants it was not entirely clear on which side of the hall the ticket office is located. Hence, they took a route on the left side as shown in Figure 8e.

For some participants a modification was made by renaming the information desk signage from *Information* to *Express ticket* and allowing the participant to buy tickets there. The information desk at position *i1* turned out to be a landmark during the trials, and this modification was requested by the architect of the guidance system. The last six participants tested this modification and the results of this modification can be seen in Figure 8f, four

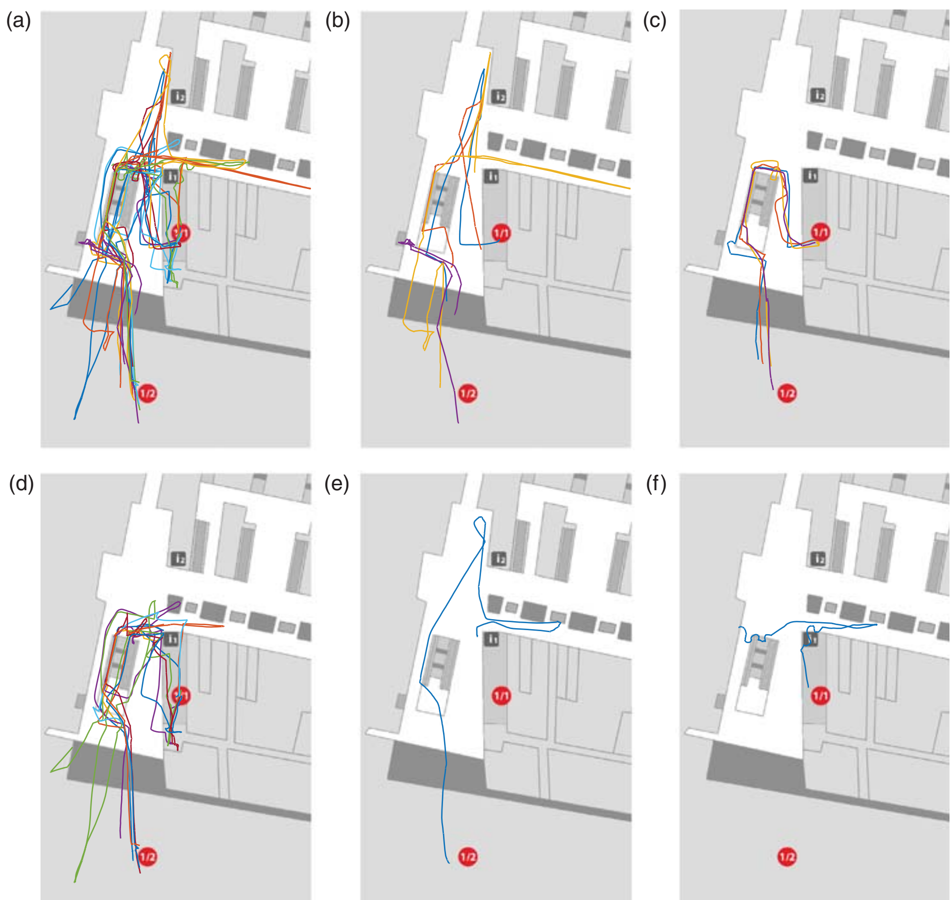

Figure 9. Trajectories for task 2, starting at *1/1* finding the luggage lockers at *1/2*. (a) all trajectories, (b) searching without external help, (c) asking for the way when buying the ticket, (d) asking for the way at information desk *i*1, (e) asking at info point *i2*, (f) aborting the task.

of the six participants bought their ticket there. Under this setting it becomes easier to find this ticket counter because of its prominent location.

After buying the ticket at position *1/1*, the participants started the second task for which they had to find the luggage lockers at position *1/2*. This required the participants to move one level below to Level − 1. All measured trajectories for this second task are shown in Figure 9a. To find the luggage lockers varying strategies were used. Figure 9b shows the trajectories of participants searching the luggage lockers on their own without requesting any information. In Figure 9c the results of the most successful strategy are shown. Here the participants already asked at the ticket counter directly after the ticket purchase for the way to the luggage lockers. With the obtained route instructions all participants from this group were able to find the luggage lockers. In Figure 9d the participants first left the ticket office and then asked at the information desk *i1* for the way. In 9e the participant tried to find their way without success and consequently asked at the information point *i2* for the way. The example in Figure 9f shows the only participant who did not find the luggage lockers and aborted the task.

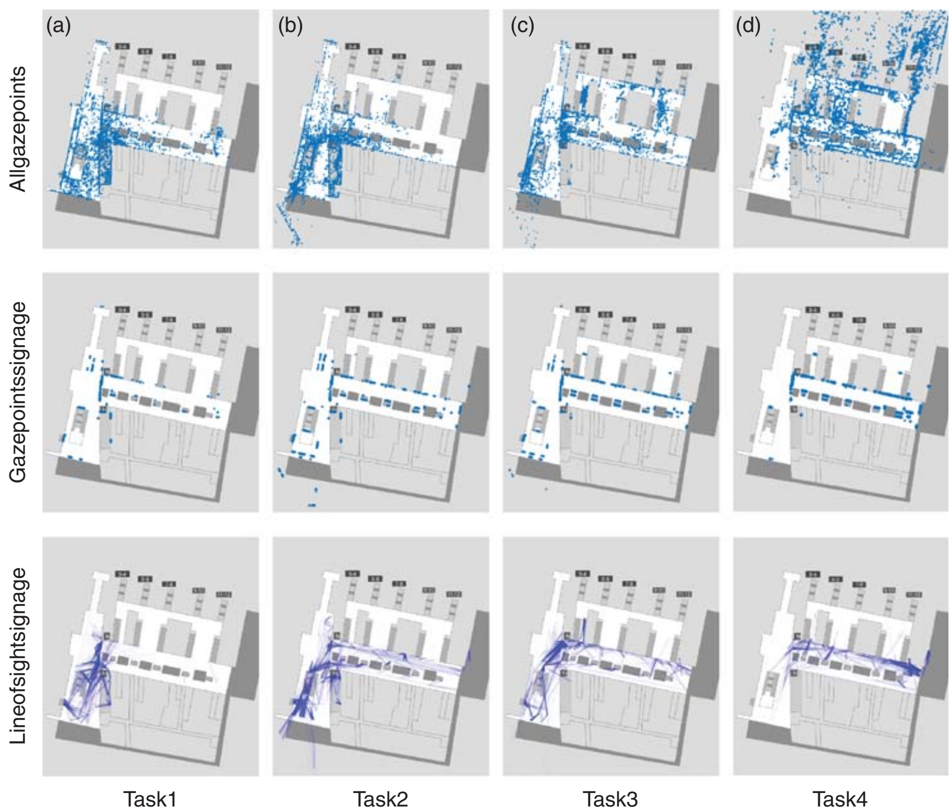

Figure 10. Top view with gaze points from saccade and fixation for the tasks 1 to 4 from left to right, (first row) for all gaze points, (second row) gaze points on signage only and (third row) the associated lines of sight from viewing position to gaze point on signage.

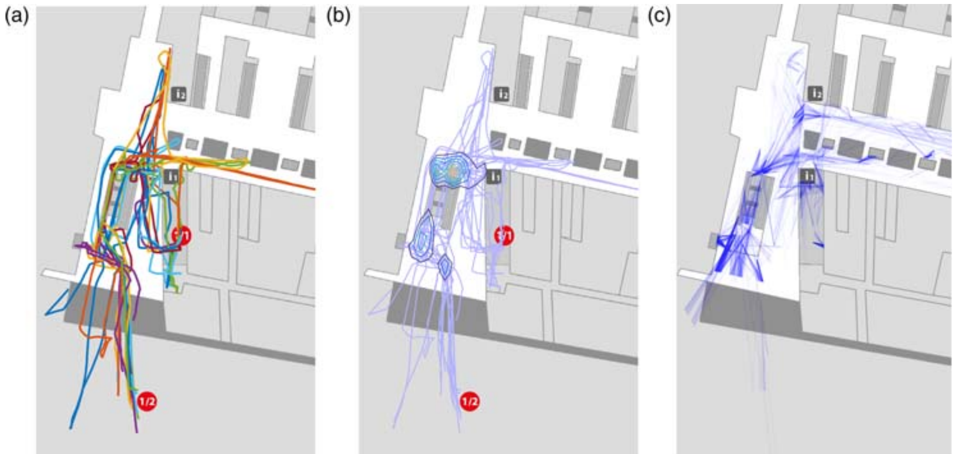

Figure 11. Task 2: (a) all trajectories, (b) density plot using trajectory data and (c) line of sight on signage.

3.2.3. Decision making

An important criteria for the decision making of the participants are the objects in the focus of the participants' attention. To investigate the fixation of objects the intersections between viewing ray and 3D model were computed. Figure 10 shows the resulting gaze saccade and fixation of all participants from a top view. In the first row the results of intersections are shown separately for each task, from task 1 to task 4, from left to right. The second row shows only gaze points that intersect with signage located in the environment. In the third row, only the lines of sight hitting signage are plotted. The results in the second and third rows provide a spatial overview of the attention on signage. Depending on the selection of data it is possible to create such an overview for each single participant or groups of participants. The mean percentage of time spent gazing at the signs were 18% for task 1, 20% for task 2, 22% for task 3, and 13% for task 4. The low value in task 4 results from the fact that in this area fewer signs were available.

In Figure 11 the captured trajectories, a density plot using trajectory data and line of sight from the example captured gaze data, are shown for task 2, starting at *1/1* finding the luggage lockers at 1/2. The trajectory plot in Figure 11a reveals the areas visited by the participants in task 2. The density plot in Figure 11b shows the areas where participants stayed longer. There they spent more time in order to orient themselves, which is confirmed in Figure 11c, where most of the line of sights are starting in this area. Density plots are therefore helpful in identifying decision areas where participants are looking for further information and where signs should be visible.

3.2.4. 3D gaze visualization

The results of the intersections were also aggregated using volume density maps and projected onto the virtual 3D environment. These maps show the

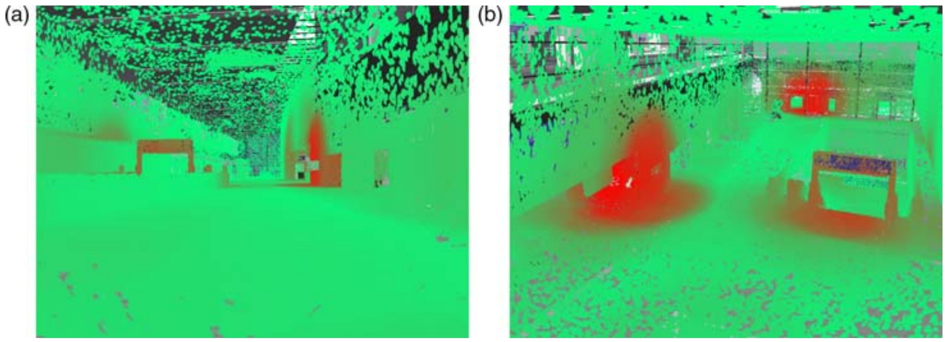

Figure 12. Distribution of visual attention in the environment, (a) shown from the perspective at the start position *Start 1* and (b) from an elevated perspective in the main hall of the railway station.

distribution of fixations in the environment as well as the amount of time any particular object was fixated. Volume density maps can be explored by dynamically navigating through the 3D model of the environment. This visualization enables interactive exploration of the users attention in the 3D environment and helps to assess the perception of the guidance system in the tested infrastructure. In Figure 12, a density map shows the visual attention aggregated over all participants and over all four tasks. In Figure 12a the distribution of fixation from the perspective of the start position *Start 1* of the experiment is shown. In 12b the three main attractions in the main hall of the Vienna main train station are shown: the information desk (left), the area around the escalator including the signage (right), and the main entrance of the train station (center). This result reveals that the information desk is used as a landmark but does not belong to the guidance system. The information desk was also the primary point chosen as a destination as people entered the main hall.

3.3. Discussion

We were able to capture accurate data on position, body and head orientation from participants in the virtual environment as well as gaze data from the eye tracking system as a basis for further analysis.

Investigation of the user task performance showed that there was no correlation between travel distances and age, but elderly participants required more time to complete the tasks. This is due to the natural locomotion interface that strongly depends on the physical fitness, which decreases with age as in real life.

The evaluation of physical behavior revealed the main paths and locations where participants spent more time for orientation or made wrong turns. This is very helpful to identify decision areas where required signs should be available.

Examination of decision making using 3D gaze analysis revealed the attention on signage that can be investigated plotting the line of sight aggregated on a 2D map or with the 3D visual analysis tool. We found the ticket counter, info point, and overhead signage at the escalator as the most salient objects along the routes. Deeper examination of gaze and of thinking aloud will reveal more details between cognitive tasks and visual attention, which is planned for the future.

4. General discussion

The wayfinding experiment can be classified within the taxonomy of Wiener, Büchner, and Hölscher (2009) as aided wayfinding, using external aids such as signage and maps. Aided wayfinding using signs is considered to be rather simple because it does not require considerable cognitive effort from the user. Therefore, it is important to provide all the relevant information at each decision point. Large public infrastructures such as the main train station in Vienna are very complex environments, which makes the design of guidance systems more difficult and error-prone.

The main goal of the experiment was to investigate the applicability of our approach for wayfinding studies. Our proposed method combines an immersive virtual environment with a mobile eye tracking system for 3D gaze analysis. It represents a new approach to support an effective evaluation of guidance systems. The main advantage of this method lies in evaluation of a multitude of different scenarios in the planning phase, avoiding costly mistakes.

To produce meaningful results it is essential to achieve a high level of realism in the virtual environment. Hence, certain aspects that people know from the real world have to be represented in the virtual environment, such as an overall realistic graphic representation including virtual personnel at the information desks and ticket counter with the same design as in the real world to support the recognition. Furthermore, escalators and elevators should be functional, other persons should be present, and ambient noise should be included. An important factor is that natural walking through the environment should be possible instead of joystick-based navigation (Suma et al., 2010). The virtual environment DAVE fulfills these key points and provides an enhanced user's sense of presence such that participants can make use of their own body's feedback in the virtual environment.

The validity of our hands-free steering approach has been explored by Bauer et al. (2013) by conducting a case study with parallel test groups, exposing individuals to wayfinding exercises in the real world and the corresponding virtual world. The validation results showed that the perceived duration, egocentric distances, and directions do not differ statistically significantly between the real and the virtual world.

Also in Lorenz et al. (2015) two hands-free steering navigation methods, a Wii Balance Board and the Microsoft Kinect Sensor for virtual environments, were compared. The results show that the Kinect is the most promising method with the highest presence for participants in a CAVE-based setup for mobility studies. This confirms our decision to use the Kinect for natural navigation.

An interesting extension of our gaze analysis tool could be a 3D scan path as presented by Stellmach, Nacke, and Dachselt (2010), with an advanced gaze visualization technique using superimposed 3D scan paths. Scan paths show the order of eye movements by drawing connected lines (saccades) between subsequent fixations. Drawing such scan lines between the identified objects of interest in 3D will enhance the investigation of visual attention in 3D virtual environments.

5. Conclusions

The methods developed in this work greatly simplify performance of different analysis with the obtained data sets. They also reduce time and costs to reveal fixation, attention maps, and objects of interest based on eye tracking measurements. Eye tracking data can be projected onto the 3D model, which allows examination of the distribution of visual attention across the virtual environment and helps to identify the most salient features along a route from the datasets. Integrating these salient features in wayfinding studies will produce directions with lower cognitive workload and higher success rates compared to directions that are only based on geometry (Nothegger, Winter, & Raubal, 2004).

In conclusion it can be stated that our approach provides an empirical testbed for qualitative and quantitative examination of visual attention using an immersive virtual environment and a mobile eye tracking system. The virtual environment is still subject to some restrictions, such as the maximum height of the projection, a delay in the response time of the navigation, and the visual representation of dense crowds, which need to be tackled in future work.

Funding

This work was part of the project MOVING and was funded by the research program line ways2go within the framework of the Austrian strategic initiative IV2Splus Intelligent Transport Systems and Services plus under the project number 835733. The study was performed using SMI Eye Tracking devices.

References

Bauer, D., Schneckenburger, J., Settgast, V., Millonig, A., & Gartner, G. (2013). Hands free steering in a virtual world for the evaluation of guidance systems in pedestrian infrastructures: design and validation. In *Transportation Research Board 92nd Annual Meeting*. Washington, DC.

Bertrand, J., Ebrahimi, E., Wachter, A., Luo, J., Babu, S. V., Duchowski, A. T., ... Gramopadhye, A. K. (2013). Visual attention to wayfinding aids in virtual environments. In B. Mohler, B. Raffin, H. Saito, & O. Staadt (Eds.), *Proceedings of the 5th Joint Virtual Reality Conference* (pp. 9–16). Aire-la-Ville, Switzerland, Switzerland: Eurographics Association. doi:10.2312/EGVE.JVRC13.009-016

Brône, G., Oben, B., & Goedem, T. (2011). Towards a more effective method for analyzing mobile eye-tracking data: Integrating gaze data with object recognition algorithms. In *Proceedings of the 1st International Workshop on Pervasive Eye Tracking & Mobile Eye-based Interaction* (pp. 53–56). New York, NY: ACM.

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.
Hale, K. S., & Stanney, K. M. (2014). *Handbook of virtual environments: Design, implementation, and applications, Second Edition*. Boca Raton, FL: CRC Press.
Hollands, M. A., Patla, A. E., & Vickers, J. N. (2002). Look where you're going!: Gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Research*, *143*(2), 221–230.
Lancelle, M., Settgast, V., & Fellner, D. (2009). Denitely affordable virtual environment. *Proceedings of IEEE Virtual Reality. IEEE, 1*.
Lorenz, M., Busch, M., Rentzos, L., Tscheligi, M., Klimant, P., & Fröhlich, P. (2015, March). I'm there! The influence of virtual reality and mixed reality environments combined with two different navigation methods on presence. In *IEEE VR 2015* (pp. 223–224). doi:10.1109/VR.2015.7223376
Morganti, F., Marrakchi, S., Urban, P. P., Iannoccari, G. A., & Riva, G. (2009, September). A virtual reality based tool for the assessment of "survey to route" spatial organization ability in elderly population: preliminary data. *Cognitive Processing*, *10*(Suppl. 2), 257–259.
Morganti, F., & Riva, G. (2011, June). Spatial orientation decline in elderly population. In *2011 International Conference on Virtual Rehabilitation (ICVR)* (pp. 1–2).
Munn, S. M., & Pelz, J. B. (2008). 3d point-of-regard, position and head orientation from a portable monocular video-based eye tracker. In K.-J. Räihä & A. T. Duchowski (Eds.), *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications* (pp. 181–188). New York, NY: ACM. doi:10.1145/1344471.1344517
Nothegger, C., Winter, S., & Raubal, M. (2004). Selection of Salient Features for Route Directions. *Spatial Cognition and Computation*, *4*(2), 113–136. doi:10.1207/s15427633scc0402_1
Pirri, F., Pizzoli, M., & Rudi, A. (2011). A general method for the point of regard estimation in 3D space. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 921–928). Washington, DC: IEEE Computer Society. doi:10.1109/CVPR.2011.5995634
Raphael, D. (2006). *Wayfinding principles & practice*. Landscape Architecture Technical Information Series (LATIS). Washington, DC: American Society of Landscape Architects.
Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments—A review. *ACM Computing Surveys (CSUR)*, *46*(2), 23:1–23:40. doi:10.1145/2543581.2543590
Ruddle, R. A., & Lessels, S. (2006). Three levels of metric for evaluating wayfinding. *Presence: Teleoperators and Virtual Environments*, *15*(6), 637–654.
Schrom-Feiertag, H., Schinko, C., Settgast, V., & Seer, S. (2014). Evaluation of guidance systems in public infrastructures using eye tracking in an immersive virtual environment. In P. Kiefer, I. Giannopoulos, M. Raubal, & A. Krüger (Eds.), *Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research* (pp. 62–66). Vienna: CEUR-WS.org.
Schwarzkopf, S., von Sülpnagel, R., Büchner, S. J., & Konieczny, L. (2013). What lab eye tracking tells us about wayfinding a comparison of stationary and mobile eye tracking in a large building scenario. In *Eye Tracking for Spatial Research, Proceedings of the 1st International Workshop*.
Spiers, H. J., & Maguire, E. A. (2008, September). The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology*, *28*(3), 232–249.
Stanney, K. M., Hale, K. S., Nahmens, I., & Kennedy, R. S. (2003). What to expect from immersive virtual environment exposure: Influences of gender, body mass index, and past experience. *Human Factors*, *45*(3), 504–520.
Stellmach, S., Nacke, L., & Dachselt, R. (2010). Advanced gaze visualizations for three-dimensional virtual environments. In C. H. Morimoto, H. O. Istance, A. Hyrskykari, & Q. Ji (Eds.), *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (pp. 109–112). New York, NY: ACM.

Suma, E., Finkelstein, S., Reid, M., Babu, S., Ulinski, A., & Hodges, L. (2010, July). Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, *16*(4), 690–702.

Takemura, K., Kohashi, Y., Suenaga, T., Takamatsu, J., & Ogasawara, T. (2010). Estimating 3D point-of-regard and visualizing gaze trajectories under natural head movements. In C. H. Morimoto, H. O. Istance, A. Hyrskykari, & Q. Ji (Eds.), *Proceedings of the 2010 Symposium on Eye-Tracking Research Applications* (pp. 157–160). New York, NY: ACM. doi:10.1145/1743666.1743705

Takemura, K., Takahashi, K., Takamatsu, J., & Ogasawara, T. (2014, Aug). Estimating 3-d point-of-regard in a real environment using a head-mounted eye-tracking system. *IEEE Transactions on Human-Machine Systems*, *44*(4), 531–536. doi:10.1109/THMS.2014.2318324

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.