DOI: 10.1111/cgf.14838

Eurographics Conference on Visualization (EuroVis) 2023
D. Archambault, R. Bujack, and T. Schreck
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 42 (2023), Number 3

Been There, Seen That: Visualization of Movement and 3D Eye Tracking Data from Real-World Environments

N. Pathmanathan[†], S. Öney, M. Becher, M. Sedlmair, D. Weiskopf, and K. Kurzhals

VISUS University of Stuttgart, Germany

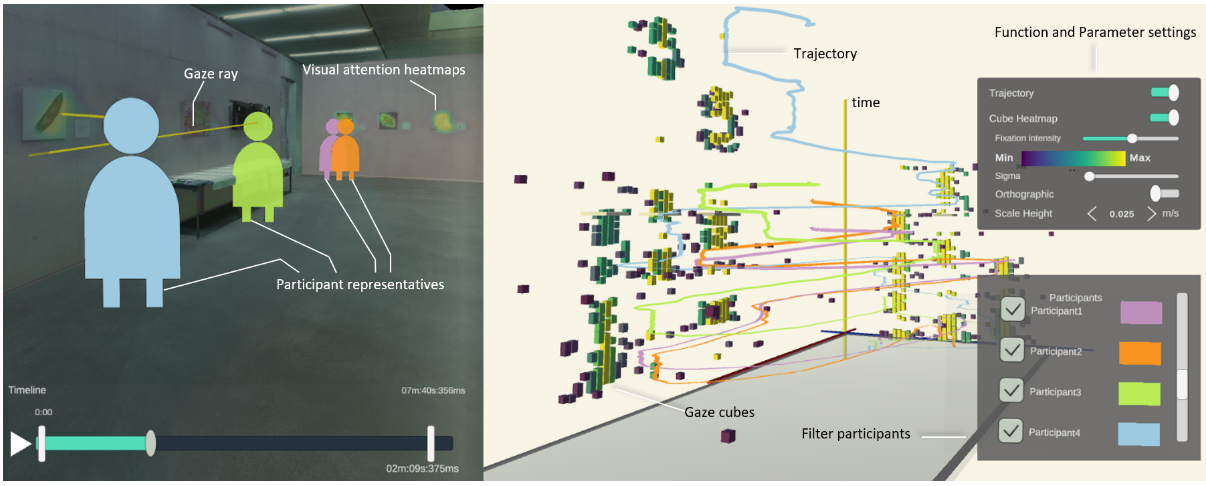


Figure 1: *Visualization of recorded movement and gaze tracking data from four participants. The gaze replay view (left) allows one to view data similar to a video player. Gaze rays indicate the current point of regard. Heatmaps display the aggregated distribution of gaze for a specific time span. The space-time cube (right) is linked with the gaze replay and provides a spatio-temporal overview of the data. Trajectories visualize movement and the gaze cubes display fixations of eye movement. Participant data can be filtered interactively.*

Abstract

The distribution of visual attention can be evaluated using eye tracking, providing valuable insights into usability issues and interaction patterns. However, when used in real, augmented, and collaborative environments, new challenges arise that go beyond desktop scenarios and purely virtual environments. Toward addressing these challenges, we present a visualization technique that provides complementary views on the movement and eye tracking data recorded from multiple people in real-world environments. Our method is based on a space-time cube visualization and a linked 3D replay of recorded data. We showcase our approach with an experiment that examines how people investigate an artwork collection. The visualization provides insights into how people moved and inspected individual pictures in their spatial context over time. In contrast to existing methods, this analysis is possible for multiple participants without extensive annotation of areas of interest. Our technique was evaluated with a think-aloud experiment to investigate analysis strategies and an interview with domain experts to examine the applicability in other research fields.

CCS Concepts

• ***Human-centered computing*** $\rightarrow$ ***Visualization;***

1. Introduction

Virtual reality (VR) and augmented reality (AR) technology provides countless new possibilities for extending the design space of

[†] e-mail:Nelusa.Pathmanathan@visus.uni-stuttgart.de

interfaces for interaction and visualization. As an example, situated analysis [MH12] aims to display visualization in spatial context of the real world, often related to the source from which the visualized data is coming from (e.g., sensor-equipped machinery [BHM*22]). Since such design significantly differs from traditional WIMP (windows, icons, menus, pointers) environments, one important question in this context is: *how can we evaluate such scenarios*?

Measuring user performance for specific tasks is important but provides little information explaining *why* a task was performed well or *what* issues came up. To this point, the evaluation of VR and AR scenarios is often restricted to interviews and standardized questionnaires [DB11, MSK*20]. If quantitative measurements are recorded over time (e.g., physiological measures [AF04]), they typically lack the spatial context in which they were acquired. As a consequence, if no video is recorded for post-experimental analysis, it becomes often hard to interpret important events in the data, for instance, an increase in heart rate.

Further, the analysis of movement in a spatio-temporal context is of high relevance for understanding complex behavior patterns of moving subjects (e.g.,people, animals, cars). The inclusion of eye tracking technology into modern head-mounted displays (HMDs) allows tracking gaze for interaction [Duc18] and the evaluation of perceptual and cognitive aspects of human behavior [JK03, MB14]. To this point, the analysis of HMD movement and gaze data is mainly restricted to single user analyses or requires tedious annotation work of areas of interest (AOIs) in every recorded dataset.

We present a new approach to support interpreting the movement and gaze behavior of multiple people in real, virtual, and augmented reality environments, either recorded individually in separate sessions or simultaneously in a single session. External data from more detailed models are incorporated to provide additional environmental context. By mapping the data into a unified virtual model, we display movement and gaze trajectories in complementary visualizations (Figure 1): The gaze replay displays the 3D spatial context of the environment in which the data was recorded. The view serves as a replacement for the video-based replay for traditional eye-tracking experiments in 2D. Hence, a replay of recorded data shows how a single or multiple persons (displayed by human figure sprites) moved through the experimental setup and where they looked at (displayed by gaze rays). The second view shows a spatio-temporal overview of the data to identify important events and interpret common movement patterns and distributions of visual attention. The focus of our research is on scenarios that rely on the 3D spatial embedding. Hence, the third view is an immersive mode that allows analysts to investigate data in the context it was recorded. The visualization helps them immerse themselves in the perspective of the participants.

Our contributions are a visualization approach that combines gaze and movement data from multiple people with additional data, such as a 3D model into a unified model for data exploration. In contrast to traditional eye tracking analysis methods [HNA*11], we retain the 3D spatial context of the environment and support the comparison of multiple participants simultaneously. To this point, space-time cube visualizations focused either on trajectories of movement or on gaze data. With our new technique, i.e., the *gaze cubes*, we show that a combination of both can be achieved to analyze data. To showcase our approach, we recorded experiments with asynchronous and collaborative behavior. We evaluated our approach with a think-aloud experiment to examine task solution strategies, and with an expert interview to identify requirements for further application domains.

We see the presented approach as a promising means for the analysis of the increasing number of experiments conducted in VR and AR, where spatial context is of high relevance for the interpretation of the data. We provide our software implementation to researchers to evaluate their AR and VR experiments. The Unity source code of our visualizations is available under [POB*23].

We will discuss related work (Section 2), followed by proposing a framework for visualizing behavioral data (Section 3). Our visualization techniques (Section 4) are evaluated by two experiments (Section 5) and results are discussed in detail (Section 6).

2. Related Work

Eye tracking in research is mainly applied for two purposes: gaze-based interaction and for the evaluation of user behavior [KFBW16]. New interaction techniques have been developed in recent years, especially with focus on VR [PMMG17] and AR [SPBG21]. For the investigation of such techniques and numerous others, eye tracking also serves as a basis for quantitative and qualitative methods. Examples of eye tracking studies comprise ones on interactive subtitles [KCH*17, KGA*20] and the navigation on websites [BCM09, CG07]. This work focuses on new, interactive techniques for the analysis of gaze data recorded with HMDs. The main issue with this type of data is that existing analysis techniques cannot be applied directly [KBPR22]: To achieve comparability between recordings, either a semantic (with AOIs) or a spatial mapping becomes necessary. Further, the visualization of the data in an explorable overview is challenging as 3D spatial context should be included, which is not the case with most of the existing visualization techniques. We address these issues by a semi-automatic spatial mapping with a linked-view visualization to make participants comparable and their movement and viewing behavior explorable. Hence, we first discuss analysis methods for eye tracking data in general, followed by techniques focusing on the visualization of spatio-temporal data in AR/VR context.

2.1. Eye Tracking Data Visualization

Statistical analysis of gaze properties (e.g., fixation duration, saccade frequency), as well as more advanced scanpath analysis [GH10a], provide important insights into gaze behavior. However, with increasingly complex scenarios, as it is the case with eye tracking with wearable glasses and HMDs, more information about the stimuli (i.e., the surroundings) become necessary. Several techniques address how to define and display AOIs and show them with abstracted visualizations [BKR*17]. The definition of AOIs is often accompanied by time-consuming manual or semi-automatic annotation. With our approach, we aim to provide an explorative visualization to investigate the data and solve analysis tasks *before* annotation of AOIs becomes necessary.

There also exist techniques that do not require any definition of AOIs. In general, heatmaps are common visualization types for displaying the aggregated distribution of visual attention [Boj09,

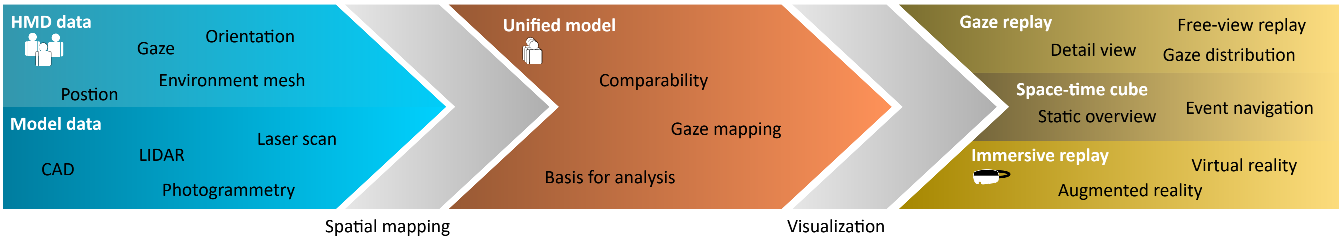


Figure 2: *Framework for the processing of multiple data sources (different participants and external model data) into a unified model. The resulting model is then visualized in complementary ways to support a multitude of analysis tasks.*

DPMO12] on static and dynamic stimuli. Node-link representations of scanpaths [GH10b] are also established means to show gaze data without aggregation. These techniques have been adapted to 3D environments [Pfe12, SND10a, PSF*13] but as a depiction of gaze distributions of single participants. Achieving an aggregated visualization of gaze from multiple participants in the real-world is not trivial and to the best of our knowledge, a comprehensive visualization technique for multi-user comparison does not exist.

We include heatmaps in our 3D representation (see Section 4.1) of the environment because they are well-known to researchers in this field and provide a good entry point to overview the data. Further, we expand this idea by including a heatmap metaphor also into our space-time visualization as *gaze cubes* (see Section 4.2).

2.2. Spatio-temporal Analysis

Visualizations that consider space and time dimensions of gaze are typically implemented by the animation of established techniques such as heatmaps. For a static overview of spatial data over time, space-time cube (STC) [Häg70] techniques have been proposed for different data domains, for instance, geographical data [GAA04] and videos [CBH*06, RVP*11]. Due to the similarities between movement trajectories and scanpaths in eye movement, this technique has also been applied to gaze data [BKF*19, KHW14, LÇK10]. However, for eye tracking applications, the STC was only applied for static images and videos presented on a screen, i.e., where all participants' data were available in a common 2D coordinate system. Mobile scenarios where people can move freely in the environment were not investigated due to individual spatial context and missing information about spatial correspondence. Related work focusing on movement only utilized, for example, information from WiFi devices to localize people [OM20]. Nowicki et al. [NS19] introduced a solution for multi-user localization in indoor environments using information from WiFi and dead reckoning. Further works used spatio-temporal trajectory data for crowd detection [HLW*20] and safety on construction sites [ACG18].

We introduce a new STC visualization (see Section 4.2) for a combined representation of movement and gaze from multiple participants. HMD data provides the necessary information to derive the missing correspondences and allows us to visualize data in a common spatial context over time. This visualization is further linked with a 3D representation of the scene, allowing analysts to overview the data (with the STC) and investigate details in the 3D replay (see Section 4.1).

2.3. Eye Tracking in VR and AR

Techniques considering 3D eye tracking data were mainly developed for VR scenarios (e.g., Stellmach et al. [SND10a]) or were included in desktop applications [OKYB21]. Possible gaze visualizations include 3D scan paths as well as 3D attention maps [SND10b]. Such purely virtual scenarios have the advantage that highly detailed world knowledge is available about all rendered objects and surfaces, as well as the position of the user. We consider our visualization approach also applicable to VR scenarios since a unified model (Section 3) for user and environment data can be obtained with far less effort than for scenarios in a real-world context [UKJi*22].

To this point, eye tracking in AR is mainly applied for interaction purposes. However, different techniques have been presented recently, showing that a retrospective investigation of movement and gaze behavior is achievable and worthwhile investigating [KBPR22]. According to Merino and colleagues [MSK*20] the evaluation of AR can be categorized into seven scenarios: These are (1) algorithm performance, (2) qualitative result inspection, (3) user performance, (4) user experience, (5) understanding environment and work practices, (6) team communication, and (7) team collaboration. We see the main advantage of our work in the latter three scenarios, as they can be especially supported by interpreting the viewing behavior and movement of the participants.

Sundstedt and Garro [SG22] provide a systematic literature review in which they discuss state-of-the-art visualization techniques for gaze data in 3D environments. The presented literature mainly focuses on visualizing the gaze directly onto the 3D Model using point-based techniques or heatmaps. These techniques concentrate more on the spatial dimension and often lack visualization of the temporal dimension with more than means of animation. Muchen and Tamke [MT21] use gaze to identify problems in spatial designs. A scan of the environment is created for visualization, and viewing behavior is displayed by motion trajectories and heatmaps. Similarly, May et al. [MKO*22] introduced a BIM-Based AR defect management system to inspect constructions. They integrated a post-inspection playback data-analysis tool into their system, visualizing the attention of a person using a heatmap. Some approaches exist for in-situ analysis of movement and interaction data (e.g. MIRIA [BLD21]) without considering gaze data. Similarly, Reipschläger et al. [RBD*22] embed a virtual avatar in AR and visualize the gaze. These approaches focus on single participants and visualize their data as either aggregated or animated visualization in a model of the spatial context. We also include such techniques

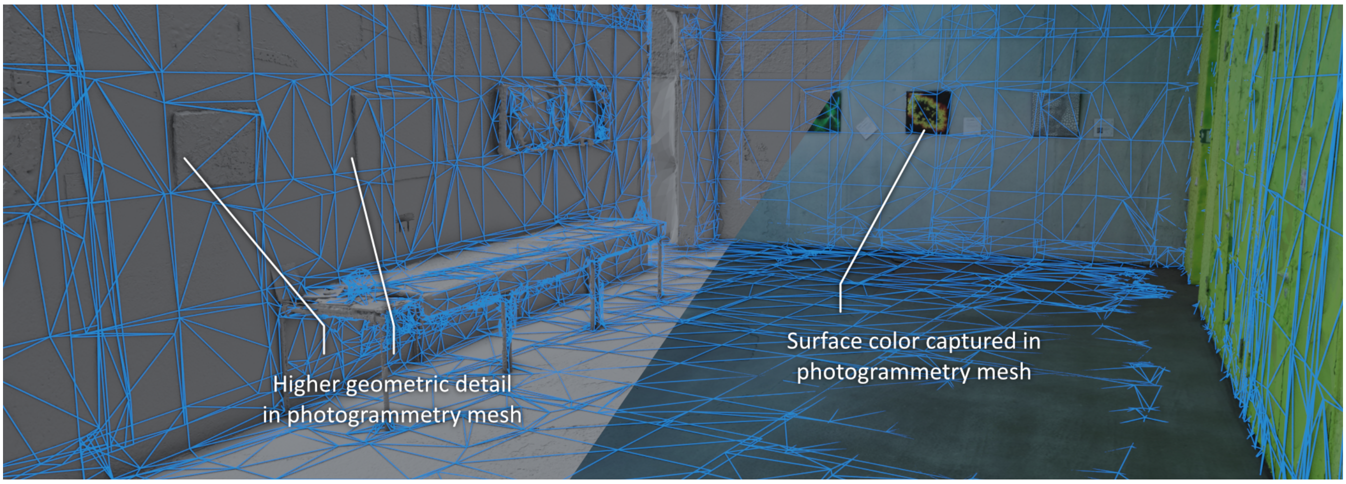


Figure 3: *The spatial mesh generated by the HoloLens 2 (blue) rendered on top of the photogrammetry reconstruction of the experiment environment. The photogrammetry mesh features greater geometric detail and also captures surface color.*

and expand the means for comparative analysis of multiple people with different views on the data in animated replays and a static overview of temporal events.

To the best of our knowledge, this is the first approach that proposes and implements such multi-user analysis for data recorded by HMDs with eye tracking capabilities.

3. Framework Proposal

To this point, there is no common way to record and visualize gaze and movement from multiple participants in real-world and AR without relying on the extensive video-based annotation of AOIs. Hence, the leading research questions for our design were:

- How can we make recordings of movement and gaze in real-world environments comparable?
- How can we include detailed information about the surrounding spatial context without annotation?
- How can we display the data to answer typical analysis questions considering *when*, *where*, and *what* happened?

We address these questions by the framework depicted in Figure 2. We assume that HMDs with eye tracking capabilities are used to record multiple participants performing a task, either in separate sessions or simultaneously. Together with additional model data to enrich the spatial context, the data is fused into a unified model for comparability of the participants. The resulting model is the basis for a wide range of possible visualization techniques to depict the underlying spatio-temporal data. We discuss three complementary techniques that do not require the annotation of AOIs and provide an overview (Figure 4c), details down to individual time steps (Figure 4), and an immersive perspective on the data (Figure 4d). We implemented the framework using photogrammetry as model data. Overall, our approach is hardware-agnostic and can be applied to all scenarios that allow for the tracking of movement and gaze in either virtual, augmented, or real environments.

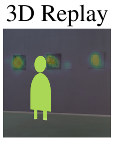
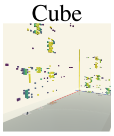
3.1. Data Processing

HMD data is the basis of our approach, enriched by model data from additional sources. HMD and model data have to be processed to derive a common spatial and temporal context.

HMD Data Recordings from multiple participants can be derived from individual sessions. Alternatively, we see the potential to record collaborative tasks with multiple participants simultaneously. The latter approach has the advantage of synchronicity between recordings. Modern HMDs use Simultaneous Localization and Mapping (SLAM) [CCC*16] techniques that provide information about position and orientation in an environment. Further, a spatial mesh is created that represents the surroundings, typically with a much coarser resolution than the respective surfaces in the real world (Figure 3). Gaze measures allow estimating the point of regard of a person, which is also mapped into the derived coordinate system. For our implementation, we recorded the gaze and movement using the Microsoft HoloLens 2 and ARETT [KBM*21b]. The data were further processed using the ARETT-R-Package [KBM*21a]. We applied the I-VT algorithm [SG00] to calculate fixations within the recorded gaze data. Each recording consists of an individual coordinate system which affects the comparison between participants.

Model Data Due to the limited quality of the environment mesh created by the HMD, the inclusion of additional data models can be achieved in multiple ways: (1) If planned constructions such as buildings are investigated, existing design plans can be incorporated to provide a more detailed (and potentially even semantically labeled) model of the environment, for instance from CAD files. (2) If such a model is not available, more detailed environment scans can be achieved by other sensors such as laser scans [BS14]. For solutions in lower price segments, LIDAR recordings with mobile devices [WNAD21] and meshes derived from photogrammetry scans [Pie20] provide alternatives. Such models typically have a better resolution than the HMD data and often provide RGB data that facilitates photo-realistic rendering of the results. Apart from the individual limitations and challenges to create 3D meshes with

Table 1: *Three complementary visualizations in combination support a wide range of possible analysis tasks.*

Visualization	Advantages	Limitations	Analysis Tasks
3D Replay 	Visualizations with spatial context Detailed inspection of space	Temporal overview missing Comparison between multiple participants	Where did participants look? Which image received most/least fixations at a specific point in time? Which participants viewed a specific image?
Space-Time Cube 	Overview over spatial- and temporal dimension Rapid detection of interesting areas or outliers	Spatial context reduced to 2D Prone to visual clutter	Are there motion patterns? Which images received the most/least fixations? When did participants watch an image in spatio-temoral proximity?
Immersive Replay 	Analysis from the point of view of the participants	No overview visualization	How did the participants perceive the scene? Which images were visible in the field of view of a participant?

each of these methods, all derived models represent a static view of the world that does not adapt to dynamic changes. If moving objects or persons are part of an experiment, the detection thereof has to be handled separately. For our implementation of the framework, we decided to include a photogrammetry reconstruction, derived from 494 photos of the experiment environment (Figure 3). For the reconstruction, we used RealityCapture [Gam] a commercial photogrammetry software. The resulting model also consists of an individual coordinate system with orientation and scale different from the HMD data. A unified model is necessary to compare data.

3.2. Unified Model

A unified model that comprises HMD and model data allows mapping of movement trajectories and scanpaths of multiple participants into one common coordinate system. As a first step, all HMD recordings have to be registered. We achieve this with spatial anchors [MRR17], i.e., a board with QR codes in the corner of the utilized environment and let the participants scan the codes. The position of the anchor serves as the origin of the common coordinate system. Second, the environment model has to be adjusted. We adopted the position and scale of this mesh to the spatial mesh extracted from the HoloLens2. This mesh was true to scale, so we were able to use the recorded data without any rescaling. For our current implementation, we adjusted the transformation manually to fit the meshes. In the future, we plan to incorporate automatic solutions for matching meshes based on correspondences [VKZHCO11], especially for more complex models such as multi-story buildings.

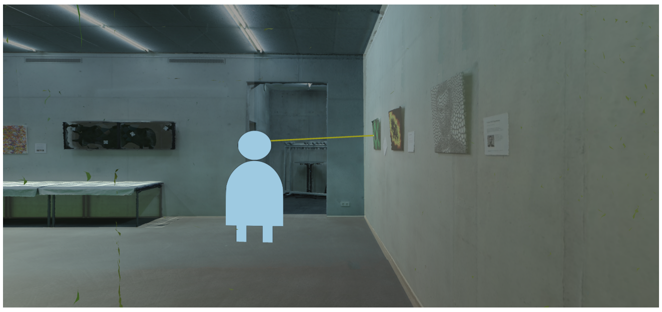
With the resulting unified model, classic trajectory analysis [AAW07] becomes possible. For instance, to identify common locations where people tend to stay or find the main walking paths. Furthermore, fixations can be mapped to common surfaces by calculating hit detections between gaze rays and the mesh of the environment. This unified model is the basis for multiple visualization techniques that focus on different aspects of the data.

4. Visualization Techniques

To address our research questions, we developed a combination of two linked, complementary views with different visualizations as the more versatile approach to address a wider range of research questions and their respective analysis tasks [KBB*17] (Table 1). More specifically, we included a 3D gaze replay that provides detailed information and a space-time cube for a spatio-temporal overview. Additionally, we support immersive analysis [FP21] in VR for data representation, similar to the gaze replay.

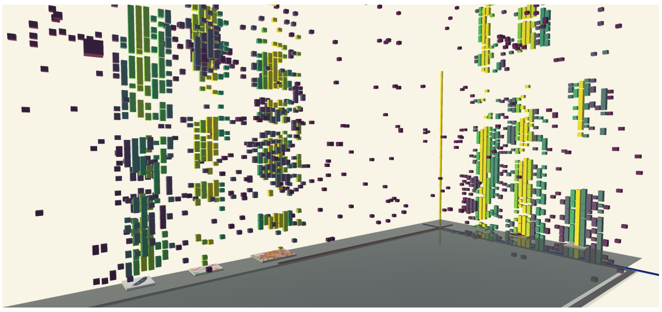
4.1. Gaze Replay

A 3D replay is achieved by representing the environmental model with indicators for the current position and point of regard (or gaze ray) of a participant (Figure 4a). Temporal evolution is displayed with a playback function, comparable to a video player. The main advantage of a virtual model is that the view of the scene can be adjusted individually to investigate events from different angles. The unified model approach further allows the depiction of multiple participants simultaneously for direct comparison. For interactive control, a timeline is included, which allows moving to different points in time, in order to closely examine the space. By limiting the depicted range around a point in time, the user can also analyze the data within a specific time span in the 3D replay. The play/pause button allows the users to play back the recording, while only concentrating on the movements and the included visualization of selected participants. The purpose of the 3D replay is a detailed inspection of movement and gaze behavior at specific points in time.

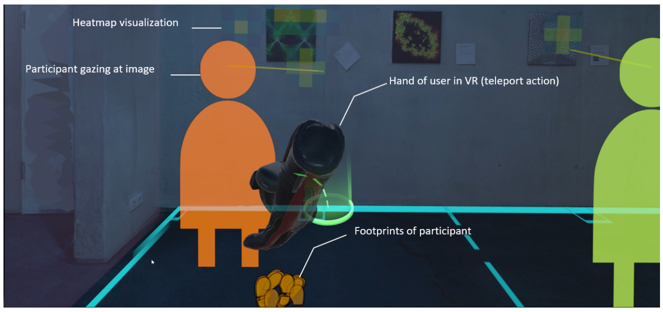


(a) *Gaze replay with a representative of a participant*

(b) *Heatmap visualization of aggregated gaze over time*



(c) *Gaze cubes in STC view*



(d) *Immersive mode*

Figure 4: *Gaze replay options in 3D space. (a) Individual time steps can be watched with a freely rotatable camera and representatives for each participant that emit gaze rays to the current point of regard. (b) For adjustable time spans, heatmaps indicate the distribution of visual attention of all selected participants. c) Gaze cube visualization in the STC view. Cubes show fixation, while color shows amount of fixation at the specific location over temporal dimension. d) Analyst can move within VR by teleport actions. We included heatmaps and representatives from the gaze replay. Footprints replace the trajectory visualizations*

To achieve this detailed inspection, we included different visualizations for both data aspects:

Movement We depict the movement of a participant by a representative and gaze direction using a ray emanating from the representative (Figure 4a). Participants receive individual colors that are also depicted by the representatives.

Gaze We include surface heatmaps (Figure 4b) for aggregated information about gaze distributions. To this end, we overlay the walls of the photogrammetry mesh with a grid. Each cell stores the number of fixations it received within the specified time span. Based on this, a heatmap is generated by applying Kernel Density Estimation (KDE) with a Gaussian kernel [Bli10].

4.2. Space-Time Cube

A common issue with animated visualization for data with spatio-temporal characteristics, such as our gaze and movement data, is that comparisons over time are difficult [RFF*08]. The inspection of data within space through forwarding and rewinding different time steps can be strenuous and there are high chances of missing interesting time spans. Playback with multiple participants further increases this issue. Hence, a static overview of trajectories and gaze distributions is a better way of supporting such comparisons. A space-time cube (STC) provides such an overview and has proven to be an effective visualization for eye tracking and movement data respectively.

The STC visualizes movement over time by trajectories, that show their location on the x-z-plane and the time on the y-axis. The distribution of the participant's gaze is visualized using cubes (see Figure 4c) encoded by color according to the fixation frequency. To visualize the gaze data in the STC, the fixations are aggregated along the vertical axis of the 3D space and stacked over time. A cube is generated if the number of fixation coordinates within a grid cell exceeds a user-defined threshold. We refer to this technique as *gaze cubes*. Since the visualization of the heatmap depends on several parameters, we added filtering options to our canvas, that allow the user to set the parameters. These include the normalization value based on fixation frequency and the bandwidth of the kernel (Figure 1). While the x-z-plane of the STC has the scale of the environment, the scale of the temporal axis can be adjusted. The linked gaze replay helps interpret if high fixation frequencies in the gaze cubes also show a vertical distribution in space.

4.3. Immersive Replay

For a more immersive and egocentric view of the data, a VR environment can be derived from the unified model. An application in AR proves more challenging due to the limited field of view and the processing power of current-gen HMDs.

We implemented the VR approach with a VARJO XR-3 headset using Unity and SteamVR [Val] to provide interaction within the virtual environment. The HTC Vive hand controllers allow us

Figure 5: *The small art gallery for data collection. Each participant was assigned to view the images in one of the displayed orders.*

to interact with the environment (Figure 4d). Users can apply the same functionalities as in the gaze replay. In addition, we adopt a footprint metaphor to depict movement trajectories.

Heatmaps are depicted on the walls and users can teleport to positions within the model to investigate the stimuli. This approach allows the investigation of the environment from the perspective of the participants with better immersion to support the understanding of behavior.

5. User Experience Evaluation

The evaluation of our visualization techniques is based on two experiments analyzing a collected data set. In the following, we will first describe the data collection and discuss each of our experiments in more detail.

5.1. Data Collection

We conducted a small-scale eye tracking study where we display a collection of printed digital artworks. We recruited ten persons (8 male, 2 female) between 20 and 40 years. Their task was to view the images in specific orders (see Figure 5). We counter-balanced the assignment of viewing orders through the gallery, simulating a between-subjects setup. Eight persons were recorded individually. Two additional persons performed the task simultaneously to showcase the differences in the resulting data and the applicability of our approach to such collaborative scenarios. The ordering of group 2 (numbers below the artworks in Figure 5) was used for the simultaneous task when people view the images together while discussing the content (see Figure 6). We provide further details about the recorded data in our supplemental material.

5.2. Experiment 1: Think-aloud

We conducted a think-aloud experiment with participants applying the gaze replay and the STC to investigate how they solve different tasks. To this point, we did not include the immersive view in our experiment as it requires a different setup and additional training time for the participants. We asked seven participants (male, age range 26–40) from our institution with a research focus on visualization or human-computer interaction to solve a set of analysis tasks. After giving their written consent, they watched a presentation explaining our work and how the different visualizations work. To familiarize them with the tool, all participants performed some small test tasks themselves. For the main set of tasks, we provided a part of the recorded dataset (Section 5.1) consisting of six recordings. While the participants were solving the tasks, we recorded a screencast and audio of the thinking aloud. The participants completed the provided tasks in approximately 45 minutes.

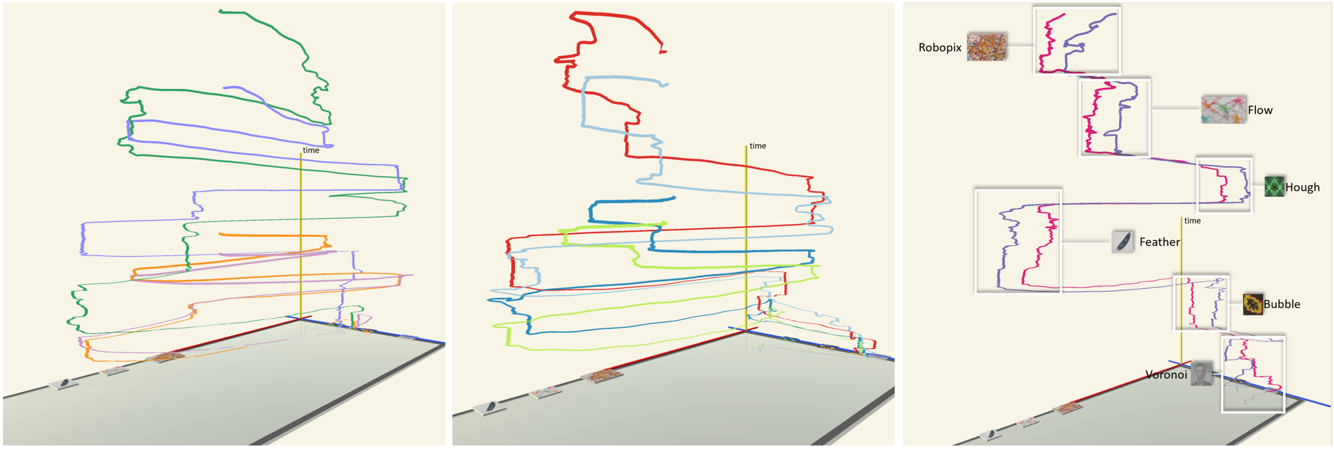
Motion Patterns We defined several tasks related to the movement of different persons. The first task was to find motion patterns within the data. We observed that participants used different strategies to accomplish this task. One choice for solving this task is investigating the trajectories within the STC. As **P1** mentioned, "*... the trajectory shows the whole course of movement*." The participants loaded the data of all persons into the scene with the trajectory visualization. **P1**, **P5**, and **P7** preferred to zoom out the STC to observe the trajectories in the overview. In contrast, the other participants opted to zoom in on the trajectory visualization and look for some clear patterns. Here, they also tried to filter out some participant data, to get a better view of individual trajectories, since it is *"difficult to see individual trajectories, within the clutter of multiple trajectories*." **P4** did not load all of the data at once and tried to tackle the trajectories one by one to find some kind of pattern.

One common observation noted by the majority of participants was that *"people took different amounts of time to look at the images*." **P4** and **P6** also considered some datasets as *"lazy"* or *"not interested"* when they observed short vertical trajectory lines in front of the images. All participants checked whether there were synchronous movements within the visualization, only some participants also regarded the sequence in which the images were viewed.

All the participants were able to detect the data of the persons which were recorded simultaneously (Figure 6c). **P4** figured out that those were recorded simultaneously and tried to interpret the data by saying *"...they took a long time because they were discussing the images*." The trajectories in group 1 were similar to each other (Figure 6a) and the majority of participants were able to detect this group, too. However, the trajectories of group 2 (Figure 6b) were not detected due to much temporal asynchrony between recordings. The trajectory of one person from this group was similar to trajectories from the simultaneous group (which followed the same order), thus some participants included the person from group 2 into the simultaneous group.

P2 and **P6** mentioned that *"...3D is difficult to see*." This problem became apparent when participants were asked to find three example locations where persons could have met when they would have been there simultaneously. While some were looking for locations where the trajectories of different persons were intersecting, others tried to find locations where the trajectories were in proximity to each other. A couple of participants tried to look for an appropriate perspective within the STC and mostly decided to take a look at the trajectories from the side. This view helped them to detect whether there were vertical lines close to each other. If they discovered a possible location, they skimmed through the timeline with the gaze replay view to confirm their assumption (see Figure 7). Another group looked directly at the gaze replay to solve this task.

Fixation Distribution For the analysis of the gaze data, we aimed for questions about fixation distributions as an indicator for visual attention. Our visualizations allow viewing the distribution of attention over the entire temporal dimension in the STC view and at specific points in time in the gaze replay.



(a) *Group 1* **(b)** *Group 2* **(c)** *Simultaneous recording*

Figure 6: *The trajectories of the different groups of participants displayed in the STC. (a) Group 1 and (b) group 2 show asynchronous movement due to the individual timings of the participants. (c) Simultaneous recording of two participants investigating the images together leads to clear movement patterns and stay points.*

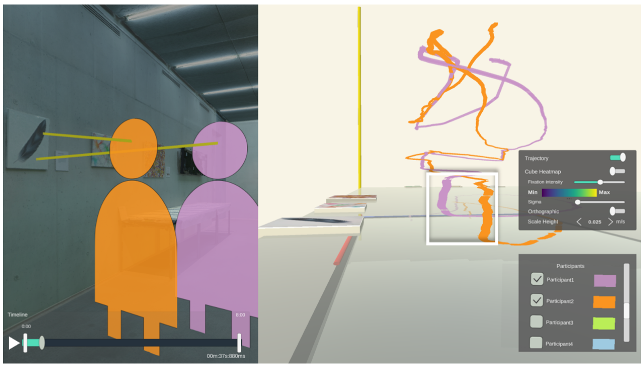


Figure 7: *Participant 1 and Participant 2 are in proximity in front of the Feather image.*

The first task in this category asked to limit the time span to a specified range and identify which image received the majority of fixations within this specified time span. To solve this task, participants had to adjust the normalization of the heatmap to identify the image with the most fixations (see Figure 8). The majority of participants did this step automatically, while **P4** and **P7** asked for a hint. They searched within the gaze replay for an image with the brightest colors in the heatmap. In this particular case, it was the **Voronoi** image as shown in Figure 8. **P2** still used the STC view and tried to approximate the count within the time span range. This approach could not solve the task, since the gaze cube aggregates multiple fixations.

Another sub-task was to identify persons viewing the image with the majority of fixations within this time span. Here the participants followed different approaches. Two participants were selecting the participants individually and checked within the gaze replay, whether there was a heatmap on the image within the time span. The other participants used the STC and viewed the trajectories within the given time span.

In the subsequent task, the participants were asked to detect the image with the majority of fixations and the least fixations over the whole data. The detection of the majority of fixations was easy to solve for almost all participants. Only two participants did not solve this task successfully. Three participants solved this task by scrolling to the end of the timeline and looking at the images with the brightest color within the gaze replay. The other participants looked directly at the gaze cube visualization within the STC and checked the cubes on the images (see Figure 9). Detecting the least amount of fixations was more difficult. Two participants were not able to tell which image received the least fixations by watching the gaze cubes. **P3** did not scroll the fixation intensity fully to the maximum, therefore multiple images seemed to receive an equal amount of attention and it was *"hard to tell"*, which image received the least amount of fixations. **P2** was looking at the number of cubes instead of the color since they *"represent fixations"*. This approach led to a wrong conclusion. **P1** and **P4** tried to view the gaze replay for the last time step since there were three images receiving darker colors, it was hard to distinguish, which image received the least fixations. Therefore, we accepted answers including one of these images. Mainly, *" ...the background color of the images influence the heatmap color"*, which made this task more difficult.

The other participants directly viewed the gaze cube visualization to answer the question. **P5** followed an interesting approach. They did change the fixation intensity slider to the minimum and *"scrolled the fixation slider up until a group of cubes received a dark coloring"*. The remaining participants took a look at the gaze cube visualization at maximum fixation intensity and tried to figure out, which image received the darkest color. They also found this task strenuous, however, they solved the given task.

From our observation, we noticed that participants with more experience in visualization followed Shneiderman's mantra [Shn96] by trying to get an overview of the visualized data first by zooming out in both views. This technique helped to rapidly detect differ-

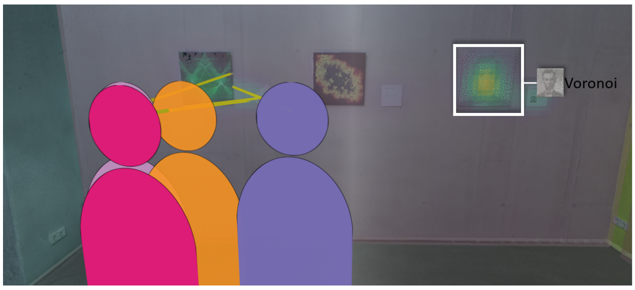


Figure 8: *Voronoi image has the most fixations within a given timespan in the gaze replay view.*

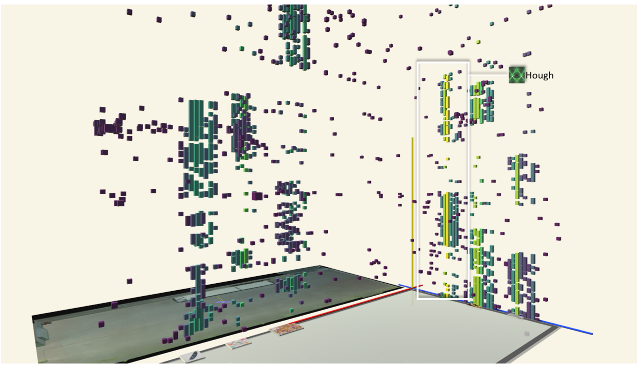


Figure 9: *Gaze cube visualization showing that the Hough image received the majority of fixation over the whole time span.*

ent kinds of events. In particular, the tasks of finding similarities within trajectories or detecting the majority of fixations could be solved smoothly with this strategy.

Usability We observed that participants had a heavier mental workload for finding the image with the majority of fixations within a provided time span. We assume that not all of the participants did understand the concept of gaze cube visualization within the STC. However, after a brief explanation, the majority of participants were able to solve all the tasks related to gaze behavior.

When comparing the performance of the individual participants based on their level of prior knowledge, we do argue that our visualizations were interpretable independent of expertise. Users could adapt fast to our tool and therefore explore and solve basic and more complex tasks about gaze and movement data without having a broad knowledge about eye tracking analysis.

At the end of our study, we collected data for the NASA task load index (TLX) [HS88] and system usability scale (SUS) [BKM09]. The results showed that the participants felt quite confident, regarding their performance. The TLX covered a scale from 1-20 and the participants rated their performance on average at 15.14. Another prominent scale was the effort they put into solving the task, here the average lies at 13.43. The mental demand and frustration received a neutral rating, while the physical and temporal demands were rated as rather low. For details we refer to our supplemental material. The SUS resulted in a *good* average score of 72.1. In general, the participants would describe the tool as easy to use and not difficult to learn.

5.3. Experiment 2: Expert Interview

To further evaluate the applicability of our approach in potential target domains, we conducted semi-structured interviews with two external domain experts for spatial cognition [HM73]. The interviews were conducted via an online conferencing system with a shared screen for demonstrations.

Both experts focused on understanding the behavior of users within their environment, mainly based on navigation and eye tracking data from way-finding studies, psychological experiments, interaction data, and data collected from methods measuring spatial memory tasks. One of the main challenges they were confronted with when utilizing visualization for data analysis was the complexity of the data: "... *a lot of post-processing has to be done by ourselves*", "... *clean up the data and combine it with complementary data*." The integration of different types of data plays an important role in analyzing behavioral data.

We then presented them with videos demonstrating the functionalities of our framework. The videos showed the gaze replay with the representative figures moving in the 3D scene, the trajectories and gaze cubes in the STC, and the immersive view. According to one expert, the gaze replay alone could be "... *useful to communicate the experimental design to non-experts but does not allow to produce aggregate statistics of movement or gaze behavior to extract knowledge*." It was also suggested including additional continuous measures in this visualization.

The experts assessed trajectories in the STC as less useful to answer their research questions, because they were mainly interested in "... *periods when people are processing information*." Although vertical patterns in the STC indicate stay points, which are often related to people looking at something, the trajectory alone "... *does not allow any qualitative or quantitative type of observation*" in this case. The trajectory visualization could be used to explain the narrative and how the experiment works. The gaze cubes were most interesting for the experts since they display information about visual attention. One of the experts mentioned that the investigation of the users' behavior at a specific location for different time sequences is of special interest. They could use this type of visualization to "... *check if there were some outliers and whether the general pattern of behavior follows my expectations*." One of the experts further noticed that the gaze replay could be useful "... *to replay how each participant who looked at a particular sign crossed one specific junction (i.e., where they were looking)*." They mentioned that the usage of traditional videos would be time-consuming and with this tool, they could "... *write filtering rules that would output only those snippets of gaze behavior in space that I am interested in*." The immersive analysis was considered as a "... *useful additional option*." For instance, "... *it would be really useful to be able to see the behavior of all participants crossing some intersection while being emerged in that intersection*." However, observing visualizations from the screen was "*easier and faster*."

When it comes to the preferences in usage, both experts had different opinions. One expert would have been more interested in using the STC due to its ability to visualize aggregated data, whereas the other expert would have opted to use the gaze replay and the VR scene since they were more interested in direct visualization of the heatmap into the space to see the data of multiple participants.

They also added that the STC "... *only works in a subset of potential cases – for example, much of my work is in multilevel buildings*", where the STC would not be suitable. Both experts would utilize our approach "... *at the beginning of data analysis*" as a first step. It was mentioned that "... *being able to extract analytics for such a scene*" is important in spatial cognition and behavior research. The addition of features to export the visualized data as a database or a file with additional variables such as total time, trajectory length, etc. was suggested. One specific feature mentioned by both experts was to "... *set multiple start and end within a 'gaze recording', for example, to indicate separate events such as testing trials*."

6. Discussion

Based on our think-aloud experiment and the expert interviews, we identified aspects that require further discussion.

Scalability We differentiate between scalability with respect to the number of participants which will mainly affect visual clutter and scalability with respect to the size of the scenario which will require design modifications. Aggregating techniques such as heatmap textures and gaze cubes can represent data values from an arbitrary number of participants. Hence, a general overview of the distribution of attention is easy to achieve, even with high participant numbers. The representation of trajectories, especially in the STC, is more prone to overdrawing and visual clutter. Consequently, only a limited number of participants should be investigated simultaneously with trajectory visualizations to keep the advantages of the overview. Bundling techniques [HEF*13] or clustering algorithms [HHWH11] can further reduce visual clutter, for example, by merging similar paths. In general, our approach could be extended to larger scenarios. However, apart from hardware limitations to process large data streams, there are some restrictions due to the design: The STC reduces by design the spatial context to 2D. Gaze distributions in different stories either have to be aggregated or the floor plans of different stories of the building have to be depicted separately. Virtual models of outdoor scenarios require much disk space. Here, we see the inclusion of existing (e.g., Google Earth) and future online services as plausible sources for required model data. Our framework would still be valid, but the resulting visualizations would need adjustments, probably toward techniques known from geo-information systems [AAW07]. We further identified shortcomings regarding the data acquisition due to environmental aspects.

Environment Aspects Surroundings without salient visual features will challenge modern approaches for localization and reconstruction. We encountered issues with the photogrammetric reconstruction of smooth and reflecting surfaces, which is a known problem with this approach. To solve these issues, we currently see the extension of the unified model by an additional sensor (e.g., LIDAR) and modeling data as the most viable solution. Further, dynamic AOIs are not considered explicitly. We see the integration of dynamic virtual objects as a trivial step to be integrated, as they can be detected automatically, similar to virtual reality environments [KBPR22]. However, some scenarios might also require dynamic real-world AOIs. For instance, we could detect hits between participants when they walked simultaneously through the scene. For a more generalizable approach, a field-of-view video will be necessary probably incorporating methods used for 2D AOIs for detection and tracking.

Based on our findings, we further identified a general preference for mouse input for 3D interactions. This could be influenced by the professional background of the participants. To ease the comparison of asynchronous trajectories, automatic techniques such as dynamic time warping could help adjust the visualization. Analysis questions regarding the STC could further be supported by 2D projections of the data to overcome issues with depth perception. For instance, a heatmap visualization of the trajectory projections could help to investigate the duration of visits from different persons at a specific location. Furthermore, the immersive replay provides a stereoscopic view which could be used to also integrate the STC, potentially as minimap for spatio-temporal navigation.

7. Conclusion

We presented an approach to process movement and gaze data from multiple people wearing HMDs for post-hoc visual analysis of their behavior. Further, we conducted a think-aloud experiment and interviews to investigate strategies and the applicability of the approach for other scenarios qualitatively. Our use case shows that the implemented visualizations support many analysis tasks in a data domain that was barely addressed to this point. Applications of this approach comprise usability studies with AR and spatial cognition tasks. For example, the planning of evacuation routes in a building [TWL*19] could be supported with AR technology and evaluated with our approach. Based on our research questions (Section 3), we can support the comparison of recordings with respect to typical analysis questions considering when, where, and what. Our current approach supports the analysis of a static environment. However, for an application to a wider range of experiments, future work should also consider dynamic changes.

For future research, we see much potential in the immersive approach, as it gives people the opportunity to see how others viewed a scene. This could not only be helpful to researchers for immersive analytics but also for communicating common perceptions to decision-makers, for instance in planned construction. In the future, we plan to conduct additional experiments in other domains such as augmented reality for construction and fabrication, where we will further evaluate our technique with more domain experts. Having the possibility to interactively investigate how other people looked in the same spatial context is further interesting for instruction scenarios [BSYB20] where instructors later have to evaluate the performance of trainees.

Our current approach does not include the visualization of aggregated movement data. The inclusion of heatmaps projected on the floor [CS21] could solve this in future extensions. The presented framework could further be extended, for instance by incorporating new sensor modalities into the unified model for later analysis. Overall, we see the visual analysis of HMD-recorded data as an important means to better understand how people perceive and interact with augmented and real environments.

Acknowledgements

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2120/1 – 390831618 and SFB 1244 – Project ID 279064222. Open Access funding enabled and organized by Projekt DEAL.

References

[AAW07] ANDRIENKO G., ANDRIENKO N., WROBEL S.: Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations Newsletter 9*, 2 (2007), 38–46. 5, 10

[ACG18] ARSLAN M., CRUZ C., GINHAC D.: Semantic enrichment of spatio-temporal trajectories for worker safety on construction sites. *Procedia Computer Science 130* (2018), 271–278. 3

[AF04] ALLANSON J., FAIRCLOUGH S. H.: A research agenda for physiological computing. *Interacting with Computers 16*, 5 (2004), 857–878. 2

[BCM09] BUSCHER G., CUTRELL E., MORRIS M. R.: What do you see when you're surfing? using eye tracking to predict salient regions of web pages. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2009), pp. 21–30. 2

[BHM*22] BECHER M., HERR D., MÜLLER C., KURZHALS K., REINA G., WAGNER L., ERTL T., WEISKOPF D.: Situated Visual Analysis and Live Monitoring for Manufacturing. *IEEE Computer Graphics and Applications 42*, 2 (2022), 33–44. 2

[BKF*19] BRUDER V., KURZHALS K., FREY S., WEISKOPF D., ERTL T.: Space-time volume visualization of gaze and stimulus. In *Proceedings of the Symposium on Eye Tracking Research & Applications* (2019), pp. 1–9. 3

[BKM09] BANGOR A., KORTUM P., MILLER J.: Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Studies 4*, 3 (2009), 114–123. 9

[BKR*17] BLASCHECK T., KURZHALS K., RASCHKE M., BURCH M., WEISKOPF D., ERTL T.: Visualization of eye tracking data: A taxonomy and survey. *Computer Graphics Forum 36*, 8 (2017), 260–284. 2

[BLD21] BÜSCHEL W., LEHMANN A., DACHSELT R.: MIRIA: A mixed reality toolkit for the in-situ visualization and analysis of spatio-temporal interaction data. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2021), pp. 470:1–15. 3

[Bli10] BLIGNAUT P.: Visual span and other parameters for the generation of heatmaps. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications* (2010), p. 125–128. 6

[Boj09] BOJKO A.: Informative or Misleading? Heatmaps Deconstructed. In *Human-Computer Interaction. New Trends*, Jacko J. A., (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 30–39. 2

[BS14] BARAZZETTI L., SCAIONI M.: Reality-Based 3D Modelling from Images and Laser Scans: Combining Accuracy and Automation. In *Advanced Research and Trends in New Technologies, Software, Human-Computer Interaction, and Communicability*. IGI Global,, Hershey PA, 2014, pp. 294–306. 4

[BSYB20] BAI H., SASIKUMAR P., YANG J., BILLINGHURST M.: A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2020), pp. 423:1–13. 10

[CBH*06] CHEN M., BOTCHEN R., HASHIM R., WEISKOPF D., ERTL T., THORNTON I.: Visual signatures in video visualization. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 1093–1100. 3

[CCC*16] CADENA C., CARLONE L., CARRILLO H., LATIF Y., SCARAMUZZA D., NEIRA J., REID I., LEONARD J. J.: Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on Robotics 32*, 6 (2016), 1309–1332. 4

[CG07] CUTRELL E., GUAN Z.: What are you looking for? An eye-tracking study of information usage in web search. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2007), pp. 407–416. 2

[CS21] CHITTARO L., SERAFINI M.: Visualization of user's behavior in indoor virtual environments through interactive heatmaps. In *INTERACT* (2021), pp. 600–609. 10

[DB11] DÜNSER A., BILLINGHURST M.: Evaluating augmented reality systems. In *Handbook of Augmented Reality*, Furht B., (Ed.). Springer New York, New York, NY, 2011, pp. 289–307. 2

[DPMO12] DUCHOWSKI A. T., PRICE M. M., MEYER M., ORERO P.: Aggregate gaze visualization with real-time heatmaps. In *Proceedings of the Symposium on Eye-Tracking Research & Applications* (2012), pp. 13–20. 2

[Duc18] DUCHOWSKI A. T.: Gaze-based interaction: A 30 year retrospective. *Computers & Graphics 73*, C (2018), 59–69. 2

[FP21] FONNET A., PRIÉ Y.: Survey of Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics 27*, 3 (2021), 2101–2122. 5

[GAA04] GATALSKY P., ANDRIENKO N., ANDRIENKO G.: Interactive analysis of event data using space-time cube. In *Proceedings of International Conference on Information Visualization* (2004), pp. 145–152. 3

[Gam] GAMES E.: RealityCapture. Last accessed 13.03.2023. URL: https://www.capturingreality.com/. 5

[GH10a] GOLDBERG J. H., HELFMAN J. I.: Scanpath clustering and aggregation. In *Proceedings of the Symposium on Eye-Tracking Research & Applications* (2010), pp. 227–234. 2

[GH10b] GOLDBERG J. H., HELFMAN J. I.: Visual scanpath representation. In *Proceedings of the Symposium on Eye-Tracking Research & Applications* (2010), pp. 203–210. 3

[Häg70] HÄGERSTRAND T.: What about people in regional science. *Regional Science Association 24* (1970), 7–21. 3

[HEF*13] HURTER C., ERSOY O., FABRIKANT S. I., KLEIN T. R., TELEA A. C.: Bundled visualization of dynamic graph and trail data. *IEEE Transactions on Visualization and Computer Graphics 20*, 8 (2013), 1141–1157. 10

[HHWH11] HÖFERLIN M., HÖFERLIN B., WEISKOPF D., HEIDEMANN G.: Interactive schematic summaries for exploration of surveillance video. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval* (2011), Association for Computing Machinery, pp. 1–8. 10

[HLW*20] HE H., LI R., WANG R., BAO J., ZHENG Y., LI T.: Efficient suspected infected crowds detection based on spatio-temporal trajectories. *ArXiv abs/2004.06653* (2020), 1–12. 3

[HM73] HART R. A., MOORE G. T.: *The development of spatial cognition: A review.* AldineTransaction, 1973. 9

[HNA*11] HOLMQVIST K., NYSTRÖM M., ANDERSSON R., DEWHURST R., JARODZKA H., VAN DE WEIJER J.: *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011. 2

[HS88] HART S. G., STAVELAND L. E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Hancock P. A., Meshkati N., (Eds.), vol. 52 of *Advances in Psychology*. North-Holland, 1988, pp. 139–183. 9

[JK03] JACOB R. J., KARN K. S.: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The mind's eye: cognitive and applied aspects of eye movement research*, Hyönä J., R. R., H. D., (Eds.). Elsevier, Amsterdam, 2003, pp. 573–605. 2

[KBB*17] KURZHALS K., BURCH M., BLASCHECK T., ANDRIENKO G., ANDRIENKO N., WEISKOPF D.: A Task-Based View on the Visual Analysis of Eye-Tracking Data. In *Eye Tracking and Visualization*, Burch M., Chuang L., Fisher B., Schmidt A., Weiskopf D., (Eds.). Springer International Publishing, Cham, 2017, pp. 3–22. 5

[KBM*21a] KAPP S., BARZ M., MUKHAMETOV S., SONNTAG D., KUHN J.: ARETT R Package: Augmented Reality Eye Tracking Toolkit for Head Mounted Displays, 18.11.2021. Last accessed 13.03.2023. URL: https://github.com/AR-Eye-Tracking-Toolkit/ARETT-R-Package. 4

[KBM*21b] KAPP S., BARZ M., MUKHAMETOV S., SONNTAG D., KUHN J.: ARETT: Augmented Reality Eye Tracking Toolkit for Head Mounted Displays. *Sensors 21*, 6 (2021), 2234:1–18. 4

[KBPR22] KURZHALS K., BECHER M., PATHMANATHAN N., REINA G.: Evaluating Situated Visualization in AR with Eye Tracking. In *Proceedings of the Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (2022), pp. 1–8. 2, 3, 10

[KCH*17] KURZHALS K., CETINKAYA E., HU Y., WANG W., WEISKOPF D.: Close to the action: Eye-tracking evaluation of speaker-following subtitles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2017), pp. 6559–6568. 2

[KFBW16] KURZHALS K., FISHER B., BURCH M., WEISKOPF D.: Eye tracking evaluation of visual analytics. *Information Visualization 15*, 4 (2016), 340–358. 2

[KGA*20] KURZHALS K., GÖBEL F., ANGERBAUER K., SEDLMAIR M., RAUBAL M.: A view on the viewer: Gaze-adaptive captions for videos. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2020), pp. 139:1–12. 2

[KHW14] KURZHALS K., HEIMERL F., WEISKOPF D.: ISeeCube: Visual analysis of gaze data for video. In *Proceedings of the Symposium on Eye-Tracking Research & Applications* (2014), pp. 43–50. 3

[LÇK10] LI X., ÇÖLTEKIN A., KRAAK M.-J.: Visual exploration of eye movement data using the space-time-cube. In *Geographic Information Science*, Fabrikant S. I., Reichenbacher T., van Kreveld M., Schlieder C., (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 295–309. 3

[MB14] MAJARANTA P., BULLING A.: Eye tracking and eye-based human–computer interaction. In *Advances in Physiological Computing*. Springer, 2014, pp. 39–65. 2

[MH12] MOERE A. V., HILL D.: Designing for the Situated and Public Visualization of Urban Data. *Journal of Urban Technology 19*, 2 (2012), 25–46. 2

[MKO*22] MAY K. W., KC C., OCHOA J. J., GU N., WALSH J., SMITH R. T., THOMAS B. H.: The Identification, Development, and Evaluation of BIM-ARDM: A BIM-Based AR Defect Management System for Construction Inspections. *Buildings 12*, 2 (2022), 140:1–30. 3

[MRR17] MÜLLER J., RÄDLE R., REITERER H.: Remote collaboration with mixed reality displays: How shared virtual landmarks facilitate spatial referencing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2017), pp. 6481–6486. 5

[MSK*20] MERINO L., SCHWARZL M., KRAUS M., SEDLMAIR M., SCHMALSTIEG D., WEISKOPF D.: Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019). In *International Symposium on Mixed and Augmented Reality (ISMAR)* (2020), pp. 438–451. 2, 3

[MT21] MUCHEN Y., TAMKE M.: Augmented Reality for Experience-centered Spatial Design: A quantitative assessment method for architectural space. In *Towards a new, configurable architecture: Proceedings of the eCAADe Conference-Volume 1* (2021), pp. 173–180. 3

[NS19] NOWICKI M. R., SKRZYPCZYŃSKI P.: A multi-user personal indoor localization system employing graph-based optimization. *Sensors 19*, 1 (2019), 157. 3

[OKYB21] OISHI S., KOIDE K., YOKOZUKA M., BANNO A.: 4d Attention: Comprehensive framework for spatio-temporal gaze mapping. *IEEE Robotics and Automation Letters 6*, 4 (2021), 7240–7247. 3

[OM20] OPPERMANN M., MUNZNER T.: Ocupado: Visualizing location-based counts over time across buildings. *Computer Graphics Forum 39*, 3 (2020), 127–138. 3

[Pfe12] PFEIFFER T.: Measuring and visualizing attention in space with 3d attention volumes. In *Proceedings of the Symposium on Eye Tracking Research & Applications* (2012), pp. 29–36. 3

[Pie20] PIETROSZEK K.: "Vera"–Crossing the Fourth Wall. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (2020), pp. 1–4. 4

[PMMG17] PFEUFFER K., MAYER B., MARDANBEGI D., GELLERSEN H.: Gaze+ pinch interaction in virtual reality. In *Proceedings of the Symposium on Spatial User Interaction* (2017), pp. 99–108. 2

[POB*23] PATHMANATHAN N., ÖNEY S., BECHER M., SEDLMAIR M., WEISKOPF D., KURZHALS K.: Source Code for: Been There, Seen That: Visualization of Movement and 3D Eye Tracking Data from Real-World Environments, 2023. doi:10.18419/darus-3383. 2

[PSF*13] PALETTA L., SANTNER K., FRITZ G., MAYER H., SCHRAMMEL J.: 3D Attention: Measurement of Visual Saliency Using Eye Tracking Glasses. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (2013), pp. 199–204. 3

[RBD*22] REIPSCHLÄGER P., BRUDY F., DACHSELT R., MATEJKA J., FITZMAURICE G., ANDERSON F.: AvatAR: An immersive analysis environment for human motion data combining interactive 3d avatars and trajectories. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2022), pp. 23:1–15. 3

[RFF*08] ROBERTSON G., FERNANDEZ R., FISHER D., LEE B., STASKO J.: Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1325–1332. 6

[RVP*11] ROMERO M., VIALARD A., PEPONIS J., STASKO J., ABOWD G.: Evaluating video visualizations of human behavior. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2011), pp. 1441–1450. 3

[SG00] SALVUCCI D. D., GOLDBERG J. H.: Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the Symposium on Eye Tracking Research & Applications* (2000), pp. 71–78. 4

[SG22] SUNDSTEDT V., GARRO V.: A Systematic Review of Visualization Techniques and Analysis Tools for Eye-Tracking in 3D Environments. *Frontiers in Neuroergonomics 3* (2022), 910019:1–15. 3

[Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE Symposium on Visual Languages* (1996), pp. 336–343. 8

[SND10a] STELLMACH S., NACKE L., DACHSELT R.: 3d attentional maps: Aggregated gaze visualizations in three-dimensional virtual environments. In *Proceedings of the International Conference on Advanced Visual Interfaces* (2010), pp. 345–348. 3

[SND10b] STELLMACH S., NACKE L., DACHSELT R.: Advanced gaze visualizations for three-dimensional virtual environments. In *Proceedings of the Symposium on Eye-Tracking Research & Applications* (2010), pp. 109–112. 3

[SPBG21] SIDENMARK L., POTTS D., BAPISCH B., GELLERSEN H.: Radi-Eye: Hands-free radial interfaces for 3D interaction using gaze-activated head-crossing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2021), pp. 1–11. 2

[TWL*19] TIAN P., WANG Y., LU Y., ZHANG Y., WANG X., WANG Y.: Behavior analysis of indoor escape route-finding based on head-mounted VR and eye tracking. In *International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (2019), pp. 422–427. 10

[UKJi*22] UGWITZ P., KVARDA O., JUŘ ÍKOVÁ Z., ŠAŠINKA U., TAMM S.: Eye-Tracking in Interactive Virtual Environments: Implementation and Evaluation. *Applied Sciences 12*, 3 (2022), 1027:1–30. 3

[Val] VALVE CORPORATION: SteamVR. Last accessed 13.03.2023. URL: https://www.steamvr.com. 6

[VKZHCO11] VAN KAICK O., ZHANG H., HAMARNEH G., COHEN-OR D.: A survey on shape correspondence. *Computer Graphics Forum 30*, 6 (2011), 1681–1707. 5

[WNAD21] WANG Z., NGUYEN C., ASENTE P., DORSEY J.: DistanciAR: Authoring site-specific augmented reality experiences for remote environments. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (2021), pp. 411:1–12. 4