Exploring Trajectory Data in Augmented Reality: A Comparative Study of Interaction Modalities

Lucas Joos* University of Konstanz
Karsten Klein† University of Konstanz
Maximilian T. Fischer‡ University of Konstanz
Frederik L. Dennig§ University of Konstanz
Daniel A. Keim¶ University of Konstanz
Michael Krone‖ University of Tübingen

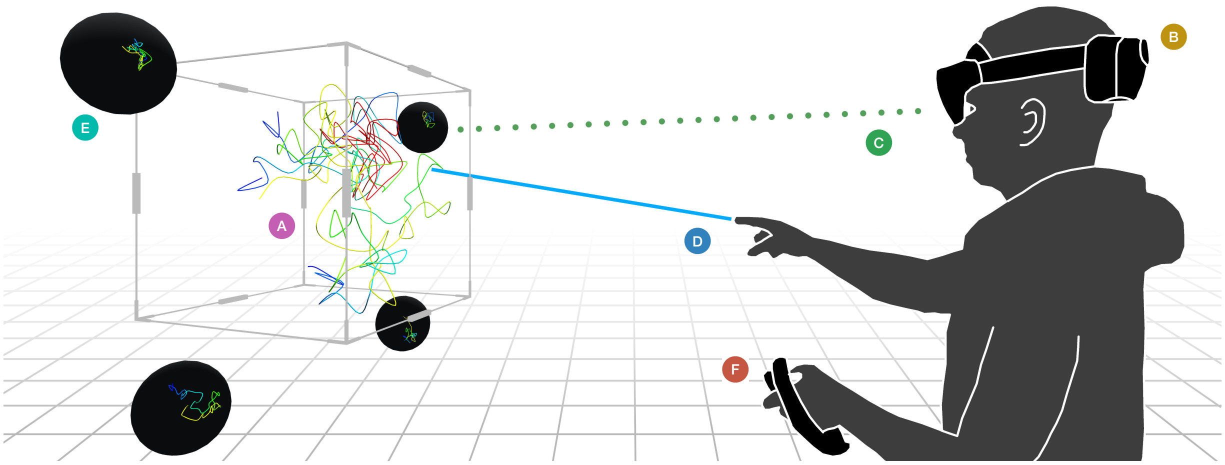


Figure 1: Our AR application for exploring 3D trajectory data (A) using a HoloLens 2 (B). Three use cases are supported by eye- (C) and hand-based (D) interaction. With a two-step selection, users (1) select a neighborhood represented by spheres (E) and (2) select one of them as final selection. Pinch gestures (later controller buttons (F)) are used to trigger actions.

Abstract

The visual exploration of trajectory data is crucial in domains such as animal behavior, molecular dynamics, and transportation. With the emergence of immersive technology, trajectory data, which is often inherently three-dimensional, can be analyzed in stereoscopic 3D, providing new opportunities for perception, engagement, and understanding. However, the interaction with the presented data remains a key challenge. While most applications depend on hand tracking, we see eye tracking as a promising yet under-explored interaction modality, while challenges such as imprecision or inadvertently triggered actions need to be addressed. In this work, we explore the potential of eye gaze interaction for the visual exploration of trajectory data within an AR environment. We integrate hand- and eye-based interaction techniques specifically designed for three common use cases and address known eye tracking challenges. We refine our techniques and setup based on a pilot user study (n=6) and find in a follow-up study (n=20) that gaze interaction can compete with hand-tracked interaction regarding effectiveness, efficiency, and task load for selection and cluster exploration tasks. However, time step analysis comes with higher answer times and task load. In general, we find the results and preferences to be user-dependent. Our work contributes to the field of immersive data exploration, underscoring the need for continued research on eye tracking interaction.

*e-mail: lucas.joos@uni-konstanz.de
†e-mail: karsten.klein@uni-konstanz.de
‡e-mail: max.fischer@uni-konstanz.de
§e-mail: frederik.dennig@uni-konstanz.de
¶e-mail: keim@uni-konstanz.de
‖e-mail: michael.krone@uni-tuebingen.de

Keywords: Augmented reality, trajectories, gaze interaction, eye tracking, user evaluation, data analysis, immersive analytics.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed/augmented reality

1 Introduction

Trajectory data encodes information about the movement of objects over time and can be gathered using GPS, computer simulations, cameras, and other technologies. The motion is represented by a series of discrete spatial coordinates (2D or 3D), which are extended by timestamps [77]. The visual exploration of trajectory data contributes to multiple domains [16], such as transportation analysis [39], animal behavior [44], or molecular dynamics [25].

Over the last decades, advances in affordability and quality of immersive hardware, such as virtual reality (VR) and augmented reality (AR) head-mounted displays (HMD), have catalyzed their use in visual data exploration. Immersive applications supporting stereoscopic 3D (S3D) have the potential to increase depth perception, allow for more natural interaction, and facilitate collaborative work [35]. Besides networks [30, 37], abstract data [4, 36], or geospatial data [5, 71], also 3D trajectory data has been investigated in immersive systems showing encouraging results [19, 20, 26, 27]. While these applications focused on trajectory visualization, we investigate the applicability of novel interaction modalities, specifically eye gaze, which has not been used before to explore 3D trajectory data.

Many modern immersive HMDs include an eye tracker providing a live signal of the user's focus of interest. Although eye tracking is mainly used to evaluate systems, there is an increasing number of publications targeting the use of eye tracking for interaction [61]. Eye gaze provides a continuous signal, available without physical movement or interaction with objects, which can be easily interpreted with 3D annotation approaches [52]. Gaze interaction can contribute to higher accessibility of a system, especially for physically impaired users. Further, it has the potential to provide more natural interaction while reducing known issues like the *Gorilla Arm* effect describing muscle fatigue after prolonged arm usage [23]. For gaze interaction, issues such as impreciseness or the *Midas-Touch Problem*, i.e., the challenge of distinguishing gaze with the intent to interact from pure observations, can occur. Despite its potential, eye tracking has not been used yet to explore 3D trajectory data. With our work, we fill this gap focusing on three research questions:

[R1] How can AR trajectory exploration be supported by eye- and hand-based interaction methods?

[R2] How can we overcome inherent challenges of gaze interaction?

[R3] How well do eye- and hand-based interaction methods work for different use cases of trajectory exploration?

Investigating these questions, we make the following contributions:

- An AR application for visual trajectory exploration integrating custom-designed concepts for eye- and hand-based interaction for three essential use cases of this data type.
- A pilot user study ($n = 6$) for assessing our interaction techniques and refining the design based on the results.
- A follow-up study ($n = 20$) comparing effectiveness, efficiency, task load, and user feedback for the interaction conditions.

2 Related Work

Our work is primarily related to the fields of eye-tracking interaction in AR & VR, studies comparing interaction modalities for HMDs, and trajectory exploration in immersive environments.

2.1 Eye-Tracking Interaction in AR & VR

In recent years, a growing body of literature has shifted its focus from using eye tracking solely for evaluation to employing it as an interaction modality for HMD applications. Besides passive use for performance optimization, such as *Foveated Rendering* [55], active gaze interaction occurs in applications for designing [15], reading [38], writing [64], healthcare [62], and gaming [60].

Tanriverdi et al. [66] outline the potential benefits of gaze-based interaction for head-worn immersive devices, namely that physical effort can be minimized, no explicit commands are required making the interaction more natural while exploiting the existing capabilities and behavior of users, the interaction is fast and also applicable when objects are far away, and eye trackers can be easily added to HMDs. A growing number of modern VR and AR headsets, such as the Meta Quest Pro [45] and the Microsoft HoloLens 2 [47], already incorporate an integrated eye tracker, increasing the availability of this technique. Gaze-based interaction can increase the accessibility of systems for physically impaired users [21] and allows to control HMDs in use cases like medical surgeries, where the hands are in use or need to remain sterile [17]. While interaction based on head tracking can be a fallback when eye tracking is unavailable, Blattgerste et al. [7] showed that eye gaze outperforms head gaze regarding time, effectiveness, task load, and user preference.

Aside from limited accuracy, a central challenge with eye tracking interaction is differentiating between observational gaze and intentional focus for interaction, a problem often referred to as the *Midas Touch* [42]. To circumvent the issue, applications can incorporate a *dwell time* after which an action is invoked [70]. However, selecting the dwell time is challenging, as short durations cause the Midas Touch problem while long durations are despised by users and increase response times [42]. Other applications use multi-modal interaction concepts, where the eye gaze determines the target, and hand gestures [9], controller buttons [11], speech [31], or head movement [43] are used to confirm associated actions.

The reported advantages and applications of the technique in different domains encourage to further explore how immersive applications can benefit from gaze interaction. A comprehensive survey on gaze interaction in AR and VR is provided by Plopski et al. [61].

2.2 Comparison of Gaze and Hand Interaction for HMDs

Although hand-tracking remains the dominant interaction modality for AR and VR applications, numerous user studies have compared this prevalent approach to gaze-based interaction.

Tanriverdi et al. [66] were the first to compare gaze- and hand-based interaction in a HMD environment finding gaze interaction superior in terms of efficiency with reduced spatial information recall and no differences in user preference. Cournia et al. [14] replicate the experiment with improved hand-tracked interaction, finding contradicting results, i.e., the hand-based condition outperforming the gaze-based equivalent. Luro et al. [41] compare eye- and hand-based interaction in VR for moving target selection finding no differences except for accuracy favoring hand interaction and cognitive load preferring gaze interaction. Pai et al. [53] compare controller- and gaze-based interaction in VR for a selection task finding a combination of gaze with forearm muscle movements to outperform the other conditions. Zhang et al. [75] study controller- and gaze-based modalities to navigate a robot resulting in advantages for accuracy, workload, and user preference when using controllers. Ahn et al. [1] combine eye gaze and a control pad for text entry in AR, finding their method superior compared to both modalities individually. Pfeuffer et al. [58] focus on VR menu control using hand- and gaze-based techniques finding direct hand interaction to be faster than gaze interaction, while hand-based pointing and gaze control are similarly fast with less physical effort for gaze control.

Numerous studies compare gaze- and hand-based interactions in AR and VR. Given the wide variety of implementations and tasks, the results differ, emphasizing the need for further modality studies.

2.3 Trajectory Exploration in Immersive Environments

The intrinsic 3D nature of many trajectory datasets is well-suited for immersive exploration in S3D, as recent research has shown.

Zhang et al. [76] demonstrate that overplotting can be reduced when exploring 3D trajectory walls in a VR environment with head-tracking and keyboard interaction for navigation. Cordeil et al. [13] introduce a collaborative VR environment that visualizes 3D flight trajectories, facilitating navigation via head position combined with a gamepad and using hand-tracking to point at interesting areas. Nguyen et al. [50] visualize trajectories of tracked bees in the geospatial context using VR and apply joystick interaction for navigation while trajectories can be selected using a 2D menu. Hurter et al. [27] present a VR system for visualizing a large number of 3D trajectories efficiently supporting navigation and selection using hand-held controllers. Klein et al. [33] demonstrate an immersive AR and VR solution for the visual analysis of animal movements in the geospatial context using hand-held controllers for navigation, displaying information, and measuring distances. Filho et al. [19] present a VR *space-time cube* visualizing geospatial 2D data with a third dimension for time. Hand-tracking is used for trajectory selection, annotation, or detailed information retrieval by finger tapping. *ReViVD* [26] allows for 3D trajectory exploration using controllers to place objects of different shapes in the scene to filter the data with Boolean operations. The VR application by Kloiber et al. [34] supports the exploration of user motion during VR experiments using hand ray interaction and direct "touching" in combination with a joystick. Several tools allow for analyzing user behavior in AR using 2D menus [8, 54], or virtual 3D filtering objects [40] for interaction. *ShuttleSpace* [72] is a VR system for analyzing badminton strokes that can be selected or filtered using the movement of a controller-simulated virtual racket. Similarly, Chu et al. [12] explore 3D badminton trajectories using ray-based menu interaction.

Although numerous promising approaches across various domains have leveraged immersive 3D trajectory exploration, their interaction capabilities, based predominantly on hand-tracking, remain rather limited. Thus, we investigate the effects of using eye gaze for interacting with trajectory data in an immersive setup.

3 Interaction Methods for AR Trajectory Analysis

Despite the potential for gaze-based interaction, especially for head-worn immersive setups, existing applications for trajectory exploration rely on hand-based interaction. We want to fill this gap by creating an immersive AR application capable of visualizing 3D trajectory data, retrieving three common use cases based on literature, and design gaze- and hand-based interaction designs for them.

3.1 AR Application

While our concepts can be applied to VR environments in the same way, we decided on an AR setup for this work. Similar to VR, AR allows for a stereoscopic 3D exploration of data but retains the perception of the environment. This can be beneficial for depth perception [59], multi-device setups [69], and collaborative work [32]. Our application is created using Unity [67] and operates directly on the Microsoft HoloLens 2 [47] HMD supporting native eye and hand tracking. Given the device's limited resources, we prioritized performance during implementation to ensure optimal operation.

Our software can process CSV files with 3D coordinates, time information, trajectory ID, and additional information. A binary format is also supported, saving storage and processing time. By using the Unity job system, the data import is fast and parallelized.

Before visualizing, we normalize the data to match the dimensions of a virtual data cube. Continuous, three-dimensional tubes connecting consecutive points visualize the trajectories. To avoid sharp edges, we use Catmull-Rom splines of user-defined granularity. The radius and the number of vertices used to approximate the circular tube structure can be customized to find a good compromise between performance and visual appearance. Customized shader programs apply a color gradient, efficient lighting, and screen-space ambient occlusion [48], improving depth perception. Besides meshes for visualizing individual trajectories, highly simplified trajectory meshes are used for efficient ray casting with the eye gaze or hand ray. Context objects associated with the trajectories, like molecules or a globe, can be added but are not part of our evaluation.

Basic interaction like rotation, translation, and resizing is supported using the Microsoft Mixed Reality Toolkit (MRTK) [46]. Hand tracking allows users to "grab" handles at virtual data cubes to apply basic transformations. The cube itself can also be grabbed and moved or rotated based on hand movement. For basic user interface elements like buttons or sliders, we also use MRTK prototypes.

Often, large numbers of trajectories are explored, requiring techniques to avoid overplotting, such as clustering [74] and representation of multiple trajectories by one [68]. We use DBSCAN with Hausdorff distance [10] for clustering and a method by Andrienko et al. [2] for representative trajectories. A wireless TCP connection between HMD and a PC is used for expensive calculations.

3.2 Use Cases

Besides basic interaction already supported by our application, we surveyed literature deriving three common use cases to support by task-specific interaction: exploration of clusters, selection of trajectories, and identification of time steps. These essential use cases match the well-known *Visual Information-Seeking Mantra* by Shneiderman [65]: starting with visualizing only representative trajectories of clusters corresponds to *overview*. Cluster exploration and trajectory selection correlate with *zoom and filter*, and *details-on-demand* is achieved by retrieving time steps and semantic information.

Cluster Exploration Calculating clusters and representing them by central pathways can help to show essential trends while reducing clutter in large trajectory datasets. However, assessing the quality of a clustering, the variance and size of clusters, or the sufficiency of cluster representations (central trajectories) require a closer look into the data [3, 51, 63, 74]. In our work, we use interaction to expand representative trajectories, i.e., revealing the cluster members. Therefore, interaction methods should allow for fast and robust user intent detection and initiate supporting visualizations.

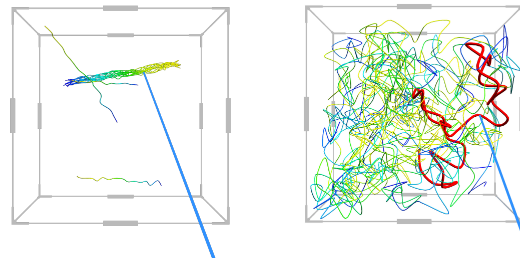


Figure 2: With the hand ray or eye gaze, users can target a representative trajectory to explore clusters (left) or select trajectories (right).

Trajectory Selection Given multiple visualized trajectories, selecting a specific one is crucial for use cases such as investigating details, annotating, highlighting, and using it as input for algorithms [18, 27, 28]. This can be especially difficult in cluttered and occluded 3D spaces. We aim to develop interaction methods allowing for fast and effective selection of a single trajectory, which should also work reliably for spaces with clutter and occlusion.

Time Step Identification Systems for visual data exploration need to show details when requested by users. This is important for the analysis of trajectories, for instance, when trying to find the time step at which a trajectory significantly changes its direction or when investigating how certain movement patterns correspond to additional, data-specific attributes [22, 34, 39]. Given a single trajectory, interaction methods should enable users to select individual time points of the trajectory. Besides rough regions, precise and efficient selection of exact time points should be supported.

3.3 Interaction Designs

In the following, we explain our interaction designs tailored to the use cases presented in Section 3.2, interaction modalities (i.e., eye tracking and hand tracking) **[R1]**, and challenges of eye tracking interaction **[R2]**. The concepts were implemented for the pilot user study and refined based on the lessons learned from the results, particularly to overcome eye tracking limitations **[R2]**.

Cluster Exploration For the exploration of clusters, we decided on direct ray-based interaction to reveal cluster members, given a representative trajectory. Ray-based interaction is a widespread and simple metaphor in AR and VR applications. The Eye ■ condition makes use of the gaze ray and raycasting to detect the intended trajectory, while for Hand ■, a hand-controlled laser ray is used (see Figure 2 left). To avoid flickering when the ray touches a trajectory only for a short time, a customizable dwell time is used to distinguish intended interaction from saccadic movement. To increase the ability to hit trajectories with the corresponding interaction ray, we double the size of the simplified virtual interaction mesh of trajectories. For this use case, we decided on direct interaction due to reduced overhead, faster interactions, and a low likelihood of clutter through representative trajectories. We expect similar accuracy for both techniques and lower answer times and task load for gaze interaction since looking at a trajectory shows the relevant data without requiring precise hand pointing.

Trajectory Selection The selection of trajectories can be challenging, especially in cluttered 3D spaces. Hence, we designed multiple interaction techniques aiming for fast and effective selection, even for occluded data and interaction impreciseness.

Direct Selection As a baseline, we incorporated direct ray-based interaction similar to the previous task (see Figure 2 right). Pointing (Hand ■) or looking (Eye ■) at a trajectory highlights it after a predefined dwell time by slightly increasing its size, providing subtle feedback of the system's interpretation. For our initial design, hand tracking is used to confirm the selection of the highlighted

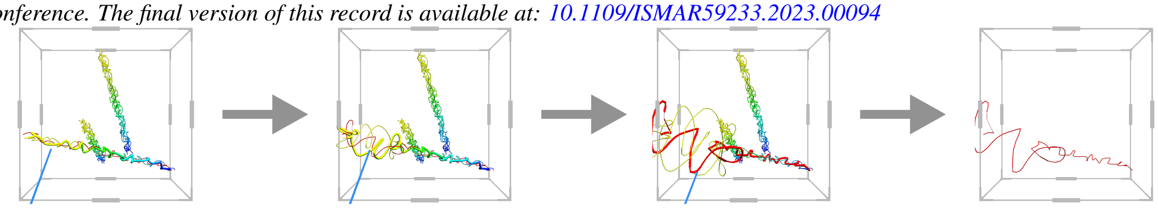


Figure 3: For trajectory selection in cluttered areas, the visualization can be adapted with a local force-based layout applied around the ray target. Thus, the space between the trajectory is increased to allow for better selection while the context of the remaining paths is preserved.

trajectory. While pointing at the desired trajectory respectively focusing it, thumb and index finger have to be moved together (*pinch* gesture). Previous work supports the combination of eye gaze and pinch [57], finding it faster than dwell time and comparable to button confirmation while not requiring additional devices [49].

Force-Based Layout Direct selection can be useful for sparse data spaces. However, given the challenges of eye tracking interaction and dense data spaces, alternatives are required. With our force-based method, we aim to decrease density and occlusion in areas a user is interested in (see Figure 3). Users indicate the area of interest (AOI) by hand pointing (HAND FB) or gaze focusing (EYE FB). Thereby, trajectories within a customizable radius of the hit point are highlighted. We use a *k*-d tree [6] as an efficient data structure for querying neighboring trajectories. With the pinch gesture, users apply a custom implementation of the ForceAtlas2 [29] force-based layout algorithm applied within the 2D camera plane and restricted to a maximum distance from the hit point. Thereby, the AOI is decluttered without losing context. Restricting the algorithm to the 2D camera coordinates prevents occlusion towards the view direction. In the second step, direct ray-based interaction, in combination with pinching, selects the desired trajectory. We expect this two-step approach to reduce issues with the impreciseness of ray pointing, as both steps do not require precise interaction.

Circular Arrangement Our third technique also applies a two-stage selection to overcome clutter-related issues while retaining the context (see Figure 5). With the EYE CIRCULAR and HAND CIRCULAR methods, eye respectively hand tracking is used to select an AOI that is calculated and highlighted as in the previous technique. Instead of adapting the preselected trajectories, small copies of them are created and placed around the data cube in a circular arrangement. Trajectories that are not part of the preselection become invisible. The small copies around the data cube are placed within semi-translucent spheres that can be selected by eye gaze (EYE CIRCULAR) or hand ray (HAND CIRCULAR). Selected spheres are indicated by yellow coloring, and only the corresponding original trajectory remains visible in the data cube. A comparable technique by Yu et al. [73] (*Flower Cone*) showed promising results in their study. This approach has the potential to overcome impreciseness challenges, while the selection might take longer if many trajectories are part of the same neighborhood.

Due to the inherent inaccuracy of eye tracking, we expect direct hand-based interaction to outperform direct gaze interaction in terms of efficiency, accuracy, and task load. However, we also expect the force-based and circular approaches to reduce these effects making both modalities similarly well-performing.

Time Step Identification As an initial design approach, we decided on a direct interaction using the eye tracking gaze (EYE) and the hand ray (HAND). When the gaze or hand ray intersects the given trajectory, the closest time point of the interpolated 3D path is calculated, visualized with a red dot, and a label indicates the time step along with additional information, if available. The label always faces the users and is placed close to the hit point without intersecting other parts of the trajectory. This direct detail retrieval approach is similarly implemented by Kloiber et al. [34]. Besides directly pointing at the desired trajectory point, we implemented a second hand-based, indirect interaction approach. For

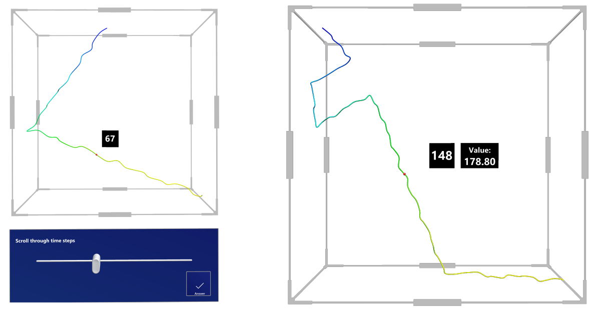


Figure 4: The identification of time steps using a hand-controlled slider (left). For the user study, a task with only one relevant time point (left), as well as a more exploratory task with five relevant points and additional information (right), are incorporated.

the SLIDER method, a virtual slider (provided by the MRTK) corresponding to the trajectory time steps with linear interpolation is placed below the data cube (see Figure 4 left). Changing the slider moves the red sphere and label to the corresponding point.

We expect similar accuracy results but gaze to be faster with a lower task load as no physical hand movement is required. However, the impreciseness of eye tracking could diminish this effect. We further assume slider navigation to be the least efficient method with the highest task load while accuracy could be higher.

4 Pilot Study

With the pilot user study, we aim to create and assess a study setup comparing both interaction modalities with task-dependent interaction techniques and use the results for design improvements. Following, we describe the study procedure including four abstract tasks with custom data. We further report on our participants, study results, and a discussion with lessons learned.

4.1 Tasks & Data

We abstracted our three use cases into four tasks with unique, measurable solutions. To assess the **Cluster Exploration** use case, participants were asked to find the cluster with the highest number of members (**Task 1**). Thus, all trajectory representatives needed to be explored, and the one with significantly more members compared to the others had to be identified. When clicking the *answer* button, labels with numbers were placed close to the trajectories, and the label number had to be entered using a virtual keypad. For this task, we created synthetic, clustered trajectory data of two different complexities and two different shapes. With the different data complexities (four clusters or eight clusters) and the different shapes (trajectories following a line (see Figure 2 left) or a random path, both with some perturbation) we integrated different characteristics occurring in real-world data. The representative trajectory was calculated with the method of Andrienko et al. [2]. While the number of cluster members was randomly chosen, one cluster contained approximately twice as many trajectories, making it clearly distinguishable. For each of the two interaction modalities, data with two different shapes, two different complexities, and two repetitions were incorporated, leading to $2 \cdot 2 \cdot 2 \cdot 2 = 16$ sub-tasks.

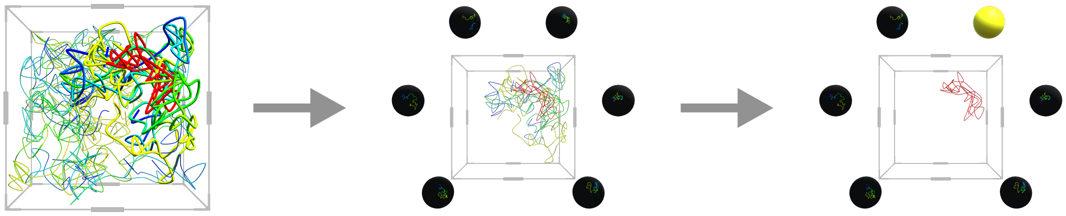


Figure 5: The circular trajectory selection allows for precise interaction through multiple steps. Initially, the region around the target point (based on eye or hand tracking) is highlighted for selection. Then, only the neighboring trajectories remain visible along with circularly placed representatives on the outside, allowing for easier selection than inside the data cube.

For the **Trajectory Selection** use case, one trajectory was indicated by red color and had to be selected (**Task 2**). A *revert* button allowed to retry if the initial selection was not correct. As before, this task was solved with data of two different complexities and forms. The number of trajectories was randomly selected within a specified range twice as large for higher data complexity than for lower complexity. Besides trajectories with a random path (see Figure 5), we also added clustered trajectories with a high level of occlusion (see Figure 3). For each of the six interaction conditions, two shapes and two complexities were tested, resulting in $6 \cdot 2 \cdot 2 = 24$ sub-tasks.

For the **Time Step Identification** use case, we created two abstract tasks measuring easier and more difficult cases. In **Task 3**, we provided a single trajectory with a unique change of direction (see Figure 4 left) and users had to identify the time step at the direction change using the different interaction techniques. This task measures the ability to detect time steps at given positions, but does not reflect the ability to explore the trajectory in detail. Thus, we incorporated **Task 4** in which we provided a single trajectory with five intense changes of directions (see Figure 4 right) and an additional value at each time step. Users needed to locate the time step with the highest extra value, which could be found at one of the five direction changes. We produced data for these tasks by generating successive line segments of random length and angle, ensuring a 45° to 120° angle range between segments for clear visibility of direction shifts. The final trajectory followed these line segments with minor random perturbation. The additional values were generated to have local maxima at direction changes, with one being the global maximum, maintaining a minimum gap to the second highest value.

4.2 Apparatus & Procedure

We conducted the user study as individual one hour sessions in a university laboratory. We used the HoloLens 2 HMD in combination with a wireless connection to a local PC, outsourcing complex calculations. Arriving participants signed a consent form and received detailed explanations from the instructor considering the study setup, interaction techniques, visualizations, and tasks. Then, the HMD was mounted and adjusted before running the eye tracking calibration. An AR video stream enabled the instructor to assist as needed. A demo covered all interaction methods occurring in the study. Moreover, participants tested different dwell times, preventing flickering for small, unintended focus shifts, and selected one for the entire study. There were four values between 200ms and 800ms to choose from corresponding to literature values [56]. For the actual evaluation, the order of the four tasks was constant for all participants, but the conditions were randomized in a Latin square design. After each condition of each task, participants were asked to lift the HMD and fill out the raw version (without weighting) of the NASA TLX test [24] (14 times in total). Thus, they had a short break from the AR environment while reporting their task load. Finally, a questionnaire gathered participant feedback and characteristics.

4.3 Participants

For the pilot user study, six randomly selected participants (4 male, 2 female) with normal or corrected-to-normal vision and ages between 20 and 55 (M: 28.8, SD: 13.2) took part. One participant had previous knowledge of trajectory analysis, one had AR experience.

4.4 Results

Due to the small sample size and the purpose of the pilot study, we only report on the main results omitting a statistical evaluation.

Accuracy For the first two tasks, all answers were correct (100%). Task 2 provided the option to revert the answer if a wrong trajectory was selected. As an effectiveness measure, the number of sub-tasks per condition where users had to revert was used: EYE : 29.2%, EYE FB : 37.5%, EYE CIRCULAR : 29.2%, HAND : 16.7%, HAND FB : 25%, HAND CIRCULAR : 12.5%. For Tasks 3 and 4, we added a small tolerance of ± 2 to assess the accuracy of answers. While we measured correctness values of 100% across all conditions for Task 3, the results were different for Task 4: EYE : 83.3%, HAND : 90.7%, SLIDER : 88.9%.

Answer Time For Task 1, EYE was faster than HAND for both data complexities. In Task 2, with data of low complexity, the eye tracking methods outperformed the hand-tracked ones except for EYE CIRCULAR , which was the overall slowest. However, EYE CIRCULAR was the fastest for the higher data complexity, and the answer times for hand-tracked interaction increased slightly while they were notably higher for the other eye-bases techniques. In Task 3, HAND was the fastest, followed by EYE and SLIDER . In contrast to Task 3, for Task 4, EYE was the slowest condition while HAND was still faster than SLIDER .

Task Load To gather task load insights, the NASA TLX test was filled out for each condition. It measures the mental demand (MD), physical demand (PD), temporal demand, performance (PF), effort (EF), and Frustration (FR). For the pilot study, we only report on large differences between conditions that can be relevant for refining the interaction design, as a complete statistical evaluation of task load results is included for the main user study.

For Task 1, MD and PD were lower for EYE while PF was higher for HAND . In Task 2, the task load results were overall similar. But for EF and FR, the circular arrangement (regardless of the modality) outperformed the other methods. For PF, EYE FB was worse than all other methods, and regarding PD, HAND was much higher than PD of the other techniques. Task 3 did not lead to notable task load differences. In Task 4, the EYE task load was the highest for all measures except for PD, while HAND and SLIDER performed comparably.

Qualitative Evaluation Further qualitative results came from a questionnaire and participant comments. For Task 1, five participants preferred EYE , one was in favor of HAND , as it "provided better visual feedback". In Task 2, all participants preferred the two-step selection approaches over direct selection, and a majority favored the force-based over circular arrangement as it did not require checking multiple spheres, needed less physical interaction, and did not introduce any indirect mapping between different objects. However, some participants disliked that always two steps were needed for the force-based layout method, even for sparse data.

A participant suggested letting users move trajectories apart continuously. For direct interaction, two participants preferred Eye, three persons Hand. Similarly, eye tracking and hand tracking were each preferred by half of the participants using a two-step selection. Comments suggested that using the hand ray and a hand gesture to trigger actions increased the difficulty. One participant indicated general difficulties and frustration with the hand gesture. For Task 3, five persons favored Hand, one person Slider. Comments suggested that Hand was perceived to be more precise than Eye and more direct than Slider. Participants criticized that eye tracking was not precise enough to reach exact points and that looking at a label could accidentally change the focus. Regardless of the modality, some participants requested a method to change the selected point by only one or two steps. The qualitative feedback for Task 4 was very similar to Task 3, but issues like impreciseness and undesired focus changes were experienced to be stronger. The only overall study critique was the long standing duration.

4.5 Discussion & Lessons Learned

In our pilot study, we assessed the setup and identified needed refinements before the main study. High task accuracy (although smaller for Task 4) indicated participant comprehension and technique efficacy. Differences in answer time and task load indicate differences between the methods that are worth investigating in a larger study. Personal modality preferences varied greatly among users.

While the overall study setup and interaction techniques worked well **[R1]**, there is still room for improvement and lessons learned that we want to address prior to the main study, especially regarding challenges coming with gaze interaction **[R2]**. The hand gesture may interfere with hand interaction and is prone to misdetection, as it is evaluated based on the camera image. Moreover, giving users the opportunity to decide whether trajectories need to be moved apart and to which degree may improve the usability of our force-based technique. The tasks requiring to identify exact time points, especially the eye tracking conditions, showed known issues like impreciseness and the *Midas Touch* problem. Additional techniques to overcome both issues have the potential to increase the usability of eye tracking for exact identification tasks and should be included. Lastly, prolonged periods of standing should be avoided.

5 Main User Study

With the main user study, we want to continue the evaluation of interaction modalities and techniques as tested based on our previous experiences from the pilot study. We start by refining interaction design and study setup, repeat the user study with 20 participants, and report on the results with a statistical evaluation.

5.1 Design Refinement & Study Setup

Based on the lessons learned from the pilot study, we changed some parts of our techniques and the study setup. To overcome issues induced by hand gestures, we decided to replace them by pressing physical buttons. Thereby, unintended actions are prevented, hand interaction should become easier, and the hand used for triggering actions can be out of the camera-tracked space allowing to keep it in any comfortable position. We incorporated the Valve Index VR controller, but any comparable controller could be used. All gesture-based action initiations were replaced by pressing a button on the controller. The controller was attached to the opposite hand that was used for pointing. We also used the controller to improve the force-layout selection technique. While previously focusing or pointing at a trajectory and completing a gesture moved trajectories apart by a fixed magnitude, we changed the technique by incorporating the controller joystick. When targeting a trajectory, users could now push the trigger up or down to increase or decrease the distance continuously. We changed the setup such that users could decide whether to directly select a trajectory or apply force-based decluttering. To cope with issues of undesired focus switches, we used one button for focus freezing. Moreover, for tasks involving time step selection, we added functionality to refine the current selection. After a broad selection of a time point using eye or hand tracking, the controller trigger could be used to fix the position and move it through the joystick, allowing for a more accurate selection.

The study setup was based on the pilot study, with changes allowing participants to sit during the study, an extended demo, and an introductory video ensuring identical explanations. We regarded the first trial of the repetitions for each condition as training.

5.2 Hypotheses

Based on our research questions and the pilot study, we devised three null hypotheses for all tasks and combinations of conditions:

H1 The interaction method has no effect on the task *accuracy*.

H2 The interaction method has no effect on the task *efficiency*.

H3 The interaction method has no effect on the task *task load*.

5.3 Participants

We invited students and employees from our university to take part in the user study. The participation was compensated with 10€. 22 participants, all with (corrected-to) normal vision, took part in the user study, and the results of 20 participants were incorporated into the evaluation. One participant reported after the study that he or she was very tired before and during the study. Since this resulted in considerably higher answer times compared to all other participants, we excluded the participant's results from the evaluation to ensure comparability. One participant encountered technical issues, requiring to abort the session and exclude the results. One person reported a red-green color deficiency, but the participant could complete the study with unexceptional results. The age of the remaining 20 participants (10 female, 10 male) varied between 21 and 52 (M: 25.6, SD: 7.7). No participant had previous experience with trajectory analysis, and only one person reported experience with AR. 19 persons were right-handed, one person left-handed.

5.4 Results

In the following, we summarize the results of our main user study. Detailed documentation of all results (S1) and statistical tests (S2 – S8) can be found in the supplementary material. We apply a significance level of $\alpha = 0.05$ and distinguish between three classes of significance: $p < 0.001$: $***$, $p < 0.01$: $**$, $p < 0.05$: $*$.

Dwell Time Setting Before starting the study, seven participants decided on the lowest dwell time (200ms), twelve on a dwell time of 400ms, one for 600ms, and no one on the 800ms option.

Accuracy The correctness values were close to 100% for Task 1-3 and around 80% for Task 4 (all conditions). Statistical analysis with Fisher's exact test and Cramer's V (effect size) (S2) found no significant differences between the conditions for Task 1 ($p = 0.4977, V = 0.0724$), Task 2 ($p = 1.0000, V = 0.1180$), Task 3 ($p = 1.0000, V = 0.1330$), and Task 4 ($p = 0.7600, V = 0.0860$).

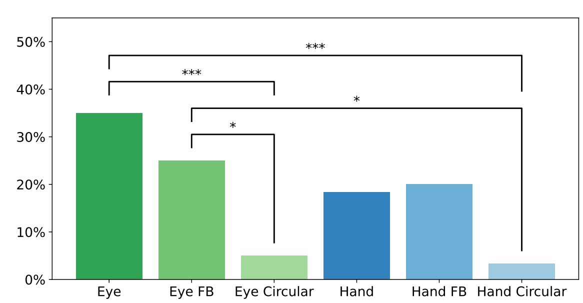


Figure 6: Ratios of necessary corrections for Task 2.

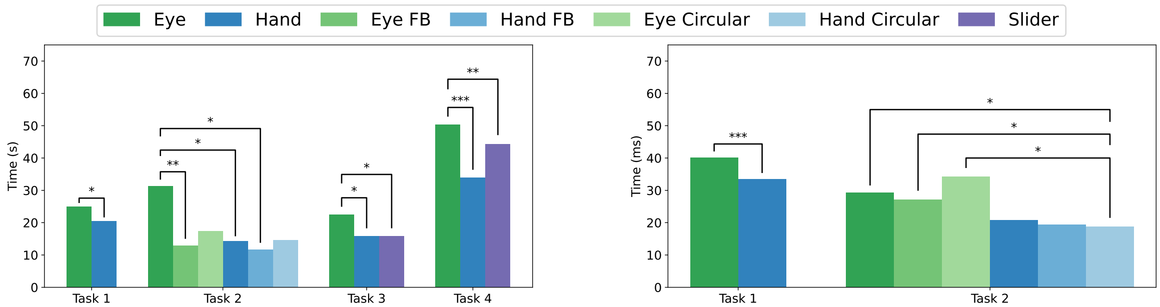
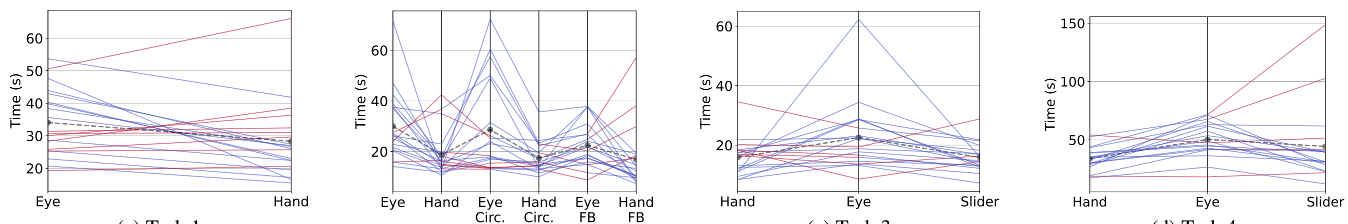


Figure 7: The answer time results of the main user study for sparse data (left) and dense data (right).



(a) Task 1 (b) Task 2 (c) Task 3 (d) Task 4

Figure 8: The average answer times per participant and condition. Red lines correspond to users needing less time with the eye tracking condition, blue ones correspond to better hand tracking results. The black line visualizes the overall average.

For Task 2, we measured the ratio of sub-tasks for each condition, where users corrected their initial selection (see Figure 6). The analysis with a Chi-Square test (S3) found significant differences between the conditions ($p < 0.0001, \chi^2 = 29.7973, V = 0.1287$). Post-hoc analysis with pairwise Fisher's exact tests with Holm-Bonferroni correction (S4) indicated more corrections for Eye than for Eye Circular ($p = 0.0008, V = 0.3750$) and Hand Circular ($p = 0.0002, V = 0.4023$). Also more were required for Eye FB than for Eye Circular ($p = 0.0466, V = 0.2801$) and Hand Circular ($p = 0.0142, V = 0.3107$).

Answer Time The answer time results (see Figure 7) significantly differed from a normal distribution (Shapiro-Wilk test) and could not be Box-Cox-transformed. Friedman tests with Kendall's W for effect size (S5) showed significant differences between the conditions for Task 2 with low ($p = 0.0001, Q = 25.1429, W = 0.2514$) and high ($p = 0.0003, Q = 23.2286, W = 0.1161$) data complexity, for Task 3 ($p = 0.0100, Q = 9.2105, W = 0.1212$), and for Task 4 ($p = 0.0001, Q = 19.0556, W = 0.2647$). For the post-hoc analysis, we applied pairwise Wilcoxon signed-rank tests with Holm-Bonferroni correction (S6). In Task 1, Eye was slower than Hand for sparse ($p = 0.0278, Z = 2.1909, r = 0.2450$) and dense data ($p = 0.0006, Z = 3.0992, r = 0.2829$). For Task 2 with low-complex data, we only found significant differences for Eye having higher answer times compared to Eye FB ($p = 0.0072, Z = 3.2479, r = 0.5135$), Hand ($p = 0.0142, Z = 3.0986, r = 0.4899$), and Hand FB ($p = 0.0142, Z = 3.0986, r = 0.4899$). For data with higher complexity, only Hand Circular took less time than Eye ($p = 0.0343, Z = 2.9436, r = 0.3291$), Eye FB ($p = 0.0212, Z = 3.0915, r = 0.3456$), and Eye Circular ($p = 0.0134, Z = 3.2259, r = 0.3607$). In Task 3, Eye was significantly slower than Hand ($p = 0.0382, Z = 2.3276, r = 0.2670$) and Slider ($p = 0.0141, Z = 2.7772, r = 0.3186$). Similarly for Task 4, Eye was slower than Hand ($p < 0.0001, Z = 4.4461, r = 0.5240$) and Slider ($p = 0.0074, Z = 2.8436, r = 0.3351$).

We also assess answer times per person (see Figure 8). For Task 1, 40% of our participants required less time with Eye compared to Hand. In Task 3, 32% were faster with Eye compared to Hand and 16% to Slider. Similarly, Eye was faster for 11% compared to Hand and 25% to Slider.

Task Load The task load results were not normally distributed (Shapiro-Wilk test) and not transformable. Friedman tests (S7) for the individual tasks and TLX scales indicated significant differences for Task 3 PD ($p = 0.0069, Q = 9.9615, W = 0.2767$) and Task 4 EF ($p = 0.0368, Q = 6.6032, W = 0.1834$), FR ($p = 0.0030, Q = 11.6207, W = 0.3228$), MD ($p = 0.0106, Q = 9.0909, W = 0.2525$), and PD ($p = 0.0131, Q = 8.6667, W = 0.2407$) (see Figure 9). A post-hoc analysis with pairwise Wilcoxon signed-rank tests and Holm-Bonferroni correction (S8) showed only significant results for Task 4, namely Eye leading to significantly higher effort ($p = 0.0413, Z = 2.4248, r = 0.4041$), frustration ($p = 0.0035, Z = 3.2166, r = 0.5361$), and mental demand ($p = 0.0197, Z = 2.6560, r = 0.4427$) compared to Hand. Moreover, Slider led to higher frustration than Hand ($p = 0.0080, Z = 2.9662, r = 0.4944$), but lower mental demand than Eye ($p = 0.0197, Z = 2.5734, r = 0.4289$). For the physical demand, Eye ($p = 0.0157, Z = 2.6426, r = 0.4404$) and Hand ($p = 0.0131, Z = 2.8385, r = 0.4731$) outperformed Slider.

Qualitative Evaluation The qualitative questionnaire results showed that hand tracking was preferred by eleven participants for Task 1 (five favored eye tracking), ten for Task 2 (three in favor of eye tracking), eleven in Task 3 (two preferred gaze, one person slider), and nine for Task 4 (two favored eye tracking, two slider). In general, 16 persons preferred hand tracking, four people eye tracking. When asked for their opinion on eye tracking interaction (1 to 5, worst to best), ten participants rated it at 4 and four participants each rating it at 3 and 2, respectively. For the first task, there were mixed opinions regarding the interaction techniques. While participants found the task to be "easier with eyes", others argued that "hand feels like more control". Concerning Task 2, participants found the hand laser to be "more precise" and "easier", which was also explained with the immediate visual feedback. However, other users commented that "eye tracking was faster" and the entire task was "easier with the eyes". Our techniques to overcome preciseness issues in the selection task were well received. For the force-based layout approach, participants commented that "zooming helps", "moving trajectories apart is really useful, a great invention", and that "the adaptions for the selection task (force/circular) are very helpful". For the circular arrangement, people argued that it "makes selection longer, but more precise", and "if the trajectories are dense, the circular selection is

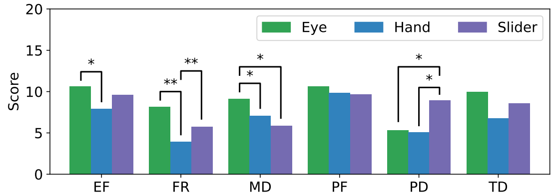


Figure 9: The results of the NASA TLX for Task 4 with the six different measures: **EF**fort, **FR**ustration, **M**ental **D**emand, **P**er**F**ormance, **P**hysical **D**emand, and **T**emporal **D**emand.

way easier". For Task 3 and 4, Eye was "not precise" and "not so good for small steps", according to a participant, while another user preferred it over Hand as "hand must be held in the air too long otherwise". The option to fix a point and use the joystick to refine the time step was well-received. A participant commented that "eye tracking works only well in combination with trigger". The Slider was found "really difficult" by a user, and two participants commented that they had issues grabbing and controlling the slider as intended. In general, the feedback on eye tracking interaction was positive. Some participants commented that "eye tracking seems to be more efficient" and "eye tracking was very good for rough finding but bad for close-ups". However, two users reported that "eye tracking was exhausting after a longer time". Users found the option to fix the current focus "very helpful".

The study setup worked for all participants, but two participants reported minor aches (shoulder/neck) and six users had occasional difficulties clicking MRTK buttons, which does not affect the efficiency evaluation, as we did not measure the time to type in answers.

6 Discussion

Previous applications targeting trajectory analysis in AR or VR environments relied on hand interaction without comparing different designs. We investigated how eye- and hand-based interaction techniques can be used to solve three trajectory analysis use cases in AR **[R1]**, incorporated techniques to overcome challenges associated with gaze interaction **[R2]**, and compared their performance **[R3]**.

The study results show similar accuracy across all tasks and conditions. For the selection task (**Task 2**), the effectiveness of direct eye tracking interaction is comparably low but significantly increased by our circular arrangement approach. We expected that increasing the space between trajectories using our force-based layout would increase the effectiveness, but the improvement was not significant. This could be the result of the required two steps for the circular arrangement while moving trajectories apart was optional. Consequently, Eye FB led to significantly lower answer times for the lower data complexity than Eye while Eye Circular did not. With our advanced selection techniques, eye tracking could compete with hand tracking regarding answer time for low- and high data complexity, where only Hand Circular outperformed the gaze-based techniques. Moreover, the gaze-based selection approaches did not result in a higher task load. The results indicate that selection using eye tracking can be a good alternative, but facilitating methods like space increasing (faster) or a two-step interaction (more accurate) should be incorporated. Further, the per-person time analysis and the qualitative feedback suggest that the applicability of gaze interaction for tasks like selection is person-dependent. We also observe this effect for **Task 1**, where accuracy and task load revealed no differences for the modalities, but gaze-based cluster exploration was significantly slower on average. However, 40% of the participants were faster using their eye gaze compared to the hand ray. The time differences were within a few seconds making eye tracking an appropriate alternative for all users, especially when hand tracking is not available. For the selection of individual time steps (**Task 3 and 4**), the results indicate that goals can be achieved similarly well with Eye compared to the hand-based methods, but slower and with a higher task load, at least for explorative tasks. Nevertheless, the comments indicate that focus freezing and time step refinement by small steps using the joystick improved the usability of Eye.

Previous research on eye tracking interaction identified impreciseness and unintended actions as key challenges. While we solved the first issue with a second modality, we could observe the issue of impreciseness for direct selection and for detail extraction. However, the quantitative and qualitative results indicate that these effects could be mitigated by the two-step selection mechanisms and the option to slightly adjust time steps selected by gaze. Previous work considers lower physical effort, increased efficiency, higher accessibility, and more natural interaction as the main advantages of gaze-based interaction. In our study, we observed similar or higher answer times on average for gaze interaction and could only observe lower physical effort when comparing Eye and Slider (**Task 4**). A reason for that could be that we considered data analysis tasks requiring multiple, precise interactions, while tasks in the literature often involved single interaction with comparably large objects. Moreover, as indicated by Figure 8 and the qualitative feedback, the efficiency, and the perceived naturalness of the modality appears to be user-dependent. Our gaze interaction methods do not require users to hold their hand in a certain position for a prolonged time, increasing accessibility. However, replacing the controller with a non-hand-held modality would further increase accessibility.

The results of our study are in accordance with previous studies comparing eye- and hand-based interaction in HMD environments, where gaze interaction rarely outperformed hand tracking, and both modalities often led to similar results. Thus, further methods and research is required to unleash the potential of eye tracking interaction and to assess its applicability for further domains and immersive technology, such as VR and Mixed Reality systems.

Despite careful consideration, our work comes with **limitations**. We focused on three use cases occurring in trajectory exploration and could not test the interaction for other tasks. Moreover, we could only assess the applicability of the two interaction modalities with the given hardware and our custom-designed interaction techniques, while other designs or hardware might have influenced the outcome. We are confident that our gaze-based interaction methods increase accessibility since no precise (and often tiring) hand-based pointing is required, only small finger movements controlling buttons. However, to further increase accessibility, other methods to invoke actions are required. Lastly, for comparability reasons, we only used synthetic data and non-experts for the evaluation. However, investigating the (long-term) effects of the interaction modality on the work of domain experts exploring genuine data would be relevant.

We will address some of the limitations in **future work** by investigating eye tracking interaction without hand-based controller confirmation to increase accessibility, expand our methods to further use cases, and test the effects with domain experts and their actual research data. Moreover, we will expand our research to other 3D data, such as network structures, and assess whether eye tracking can also be an adequate interaction modality for such data.

7 Conclusion

With our work, we are the first to bring eye tracking interaction to the field of immersive 3D trajectory exploration. After deriving three essential use cases and developing interaction techniques supporting them, we conducted an evaluation comparing hand-tracking and eye-tracking as interaction modalities. The evaluation consisted of a pilot study with six participants resulting in design and study refinements, and a complete user study with 20 participants. The study results indicate that eye tracking interaction can be a reasonable alternative to analyze trajectories using hand tracking, but methods reducing the effects of impreciseness and undesired action triggering need to be incorporated. Moreover, the choice of modality depends on the current use case and the personal user preference.

Acknowledgments

The authors gratefully acknowledge financial support by the Federal Ministry for Economic Affairs and Climate Action (BMWK, grant No. 03EI1048D).

References

[1] S. Ahn and G. Lee. Gaze-assisted typing for smart glasses. In *Proc. UIST*, pp. 857–869. ACM, 2019. doi: 10.1145/3332165.3347883

[2] G. Andrienko, N. Andrienko, G. Fuchs, and J. M. C. Garcia. Clustering trajectories by relevant parts for air traffic analysis. *IEEE Trans. Vis. Comput. Graph.*, 24(1):34–44, 2018. doi: 10.1109/TVCG.2017.2744322

[3] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, and F. Giannotti. Interactive visual clustering of large collections of trajectories. In *Proc. VAST*, pp. 3–10. IEEE, 2009. doi: 10.1109/VAST.2009.5332584

[4] L. Arms, D. Cook, and C. Cruz-Neira. The benefits of statistical visualization in an immersive environment. In *Proc. VR*, pp. 88–95. IEEE, 1999. doi: 10.1109/VR.1999.756938

[5] M. M. Bagher, P. Sajjadi, J. O. Wallgrün, P. LaFemina, and A. Klippel. Virtual reality for geospatial education: immersive technologies enhance sense of embodiment. *Cartogr. Geogr. Inf. Sci.*, 50(3):233–248, 2023. doi: 10.1080/15230406.2022.2122569

[6] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Commun. ACM*, 18(9):509–517, 1975. doi: 10.1145/361002.361007

[7] J. Blattgerste, P. Renner, and T. Pfeiffer. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In *Proc. COGAIN*, pp. 1:1–1:9. ACM, 2018. doi: 10.1145/3206343.3206349

[8] W. Büschel, A. Lehmann, and R. Dachselt. Miria: A mixed reality toolkit for the in-situ visualization and analysis of spatio-temporal interaction data. In *Proc. CHI*, pp. 470:1–470:15. ACM, 2021. doi: 10.1145/3411764.3445651

[9] I. Chatterjee, R. Xiao, and C. Harrison. Gaze+gesture: Expressive, precise and targeted free-space interactions. In *Proc. ICMI*, pp. 131–138. ACM, 2015. doi: 10.1145/2818346.2820752

[10] J. Chen, R. Wang, L. Liu, and J. Song. Clustering of trajectories based on hausdorff distance. In *Proc. ICECC*, pp. 1940–1944. IEEE, 2011. doi: 10.1109/ICECC.2011.6066483

[11] W. H. Chia, Y. Cai, and A. Ho. Multimodal lnteraction with gaze and controller gesture. In *Proc. ISMAR-Adjunct*, pp. 518–523. IEEE, 2022. doi: 10.1109/ISMAR-Adjunct57072.2022.00110

[12] X. Chu, X. Xie, S. Ye, H. Lu, H. Xiao, Z. Yuan, Z. Chen, H. Zhang, and Y. Wu. Tivee: Visual exploration and explanation of badminton tactics in immersive visualizations. *IEEE Trans. Vis. Comput. Graph.*, 28(1):118–128, 2022. doi: 10.1109/TVCG.2021.3114861

[13] M. Cordeil, T. Dwyer, and C. Hurter. Immersive solutions for future air traffic control and management. In *Proc. ISS*, pp. 25–31. ACM, 2016. doi: 10.1145/3009939.3009944

[14] N. Cournia, J. D. Smith, and A. T. Duchowski. Gaze- vs. hand-based pointing in virtual environments. In *Proc. CHI EA*, pp. 772–773. ACM, 2003. doi: 10.1145/765891.765982

[15] C. Creed, M. Frutos-Pascual, and I. Williams. Multimodal gaze interaction for creative design. In *Proc. CHI*, pp. 69:1–69:13. ACM, 2020. doi: 10.1145/3313831.3376196

[16] U. Demšar, J. A. Long, F. Benitez-Paez, V. Brum Bastos, S. Marion, G. Martin, S. Sekulić, K. Smolak, B. Zein, and K. Siła-Nowicka. Establishing the integrated science of movement: bringing together concepts and methods from animal and human movement analysis. *Int. J. Geogr. Inf. Sci.*, 35(7):1273–1308, 2021. doi: 10.1080/13658816.2021.1880589

[17] A. Ezzat, A. Kogkas, J. Holt, R. Thakkar, A. Darzi, and G. Mylonas. An eye-tracking based robotic scrub nurse: proof of concept. *Surg. Endosc.*, 35(9):5381–5391, 2021. doi: 10.1007/s00464-021-08569-w

[18] S. Fadloun, Y. Morakeb, E. Cuenca, and K. Choutri. Trajectoryvis: a visual approach to explore movement trajectories. *Soc. Netw. Anal. Min.*, 12, 2022. doi: 10.1007/s13278-022-00879-8

[19] J. Filho, W. Stuerzlinger, and L. Nedel. Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE Trans. Vis. Comput. Graph.*, 26(01):514–524, 2020. doi: 10.1109/TVCG.2019.2934415

[20] G. Fouché, F. Argelaguet Sanz, E. Faure, and C. Kervrann. Timeline design space for immersive exploration of time-varying spatial 3d data. In *Proc. VRST*, pp. 21:1–21:11. ACM, 2022. doi: 10.1145/3562939.3565612

[21] Y. Ghasemi and H. Jeong. Using gaze-based interaction to alleviate situational mobility impairment in extended reality. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, 66(1):435–439, 2022. doi: 10.1177/1071181322661224

[22] S. Grottel, J. Heinrich, D. Weiskopf, and S. Gumhold. Visual analysis of trajectories in multi-dimensional state spaces. *Comput. Graph. Forum*, 33(6):310–321, 2014. doi: 10.1111/cgf.12352

[23] J. T. Hansberger, C. Peng, S. L. Mathis, V. Areyur Shanthakumar, S. C. Meacham, L. Cao, and V. R. Blakely. Dispelling the gorilla arm syndrome: The viability of prolonged gesture interactions. In *Proc. VAMR*, pp. 505–520. Springer, 2017. doi: 10.1007/978-3-319-57987-0_41

[24] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Human Mental Workload*, vol. 52 of *Advances in Psychology*, pp. 139–183. North-Holland, 1988. doi: 10.1016/S0166-4115(08)62386-9

[25] S. A. Hollingsworth and R. O. Dror. Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143, 2018. doi: 10.1016/j.neuron.2018.08.011

[26] F. Homps, Y. Beugin, and R. Vuillemot. Revivd: Exploration and filtering of trajectories in an immersive environment using 3d shapes. In *Proc. VR*, pp. 729–737. IEEE, 2020. doi: 10.1109/VR46266.2020.00096

[27] C. Hurter, N. H. Riche, S. M. Drucker, M. Cordeil, R. Alligier, and R. Vuillemot. Fiberclay: Sculpting three dimensional trajectories to reveal structural insights. *IEEE Trans. Vis. Comput. Graph.*, 25(1):704–714, 2019. doi: 10.1109/TVCG.2018.2865191

[28] C. Hurter, B. Tissoires, and S. Conversy. Fromdady: Spreading aircraft trajectories across views to support iterative queries. *IEEE Trans. Vis. Comput. Graph.*, 15:1017–1024, 2009. doi: 10.1109/TVCG.2009.145

[29] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLoS ONE*, 9(6), 2014. doi: 10.1371/journal.pone.0098679

[30] L. Joos, S. Jaeger-Honz, F. Schreiber, D. A. Keim, and K. Klein. Visual comparison of networks in vr. *IEEE Trans. Vis. Comput. Graph.*, 28(11):3651–3661, 2022. doi: 10.1109/TVCG.2022.3203001

[31] A. A. Khan, J. Newn, J. Bailey, and E. Velloso. Integrating gaze and speech for enabling implicit interactions. In *Proc. CHI*, pp. 349:1–349:14. ACM, 2022. doi: 10.1145/3491102.3502134

[32] K. Kiyokawa, M. Billinghurst, S. Hayes, A. Gupta, Y. Sannohe, and H. Kato. Communication behaviors of co-located users in collaborative ar interfaces. In *Proc. ISMAR*, pp. 139–148. IEEE, 2002. doi: 10.1109/ISMAR.2002.1115083

[33] K. Klein, B. Sommer, H. Nim, A. Flack, K. Safi, M. Nagy, S. Feyer, Y. Zhang, K. Rehberg, A. Gluschkow, M. Quetting, W. Fiedler, M. Wikelski, and F. Schreiber. Fly with the flock: immersive solutions for animal movement visualization and analytics. *J. R. Soc. Interface*, 16(153), 2019. doi: 10.1098/rsif.2018.0794

[34] S. Kloiber, V. Settgast, C. Schinko, M. Weinzerl, J. Fritz, T. Schreck, and R. Preiner. Immersive analysis of user motion in vr applications. *Vis. Comput.*, 36:1937–1949, 2020. doi: 10.1007/s00371-020-01942-1

[35] M. Kraus, J. Fuchs, B. Sommer, K. Klein, U. Engelke, D. Keim, and F. Schreiber. Immersive analytics with abstract 3d visualizations: A survey. *Comput. Graph. Forum*, 41(1):201–229, 2022. doi: 10.1111/cgf.14430

[36] M. Kraus, N. Weiler, D. Oelke, J. Kehrer, D. A. Keim, and J. Fuchs. The impact of immersion on cluster identification tasks. *IEEE Trans. Vis. Comput. Graph.*, 26:525–535, 2020. doi: 10.1109/TVCG.2019.2934395

[37] O.-H. Kwon, C. Muelder, K. Lee, and K.-L. Ma. A study of layout, rendering, and interaction methods for immersive graph visualization.

IEEE Trans. Vis. Comput. Graph., 22(7):1802–1815, 2016. doi: 10.1109/TVCG.2016.2520921

[38] G. Lee, J. Healey, and D. Manocha. Vrdoc: Gaze-based interactions for vr reading experience. In *Proc. ISMAR*, pp. 787–796. IEEE, 2022. doi: 10.1109/ISMAR55827.2022.00097

[39] H. Liu, X. Chen, Y. Wang, B. Zhang, Y. Chen, Y. Zhao, and F. Zhou. Visualization and visual analysis of vessel trajectory data: A survey. *Vis. Inform.*, 5(4):1–10, 2021. doi: 10.1016/j.visinf.2021.10.002

[40] W. Luo, Z. Yu, R. Rzayev, M. Satkowski, S. Gumhold, M. McGinity, and R. Dachselt. Pearl: Physical environment based augmented reality lenses for in-situ human movement analysis. In *Proc. CHI*, pp. 381:1–381:15. ACM, 2023. doi: 10.1145/3544548.3580715

[41] F. L. Luro and V. Sundstedt. A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In *Proc. ETRA*, pp. 68:1–68:9. ACM, 2019. doi: 10.1145/3317956.3318153

[42] P. Majaranta and A. Bulling. Eye tracking and eye-based human–computer interaction. In *Advances in Physiological Computing*, pp. 39–65. Springer, 2014. doi: 10.1007/978-1-4471-6392-3_3

[43] D. Mardanbegi, D. W. Hansen, and T. Pederson. Eye-based head gestures. In *Proc. ETRA*, pp. 139–146. ACM, 2012. doi: 10.1145/2168556.2168578

[44] J. Melzheimer, S. K. Heinrich, B. Wasiolka, R. Mueller, S. Thalwitzer, I. Palmegiani, A. Weigold, R. Portas, R. Roeder, M. Krofel, H. Hofer, and B. Wachter. Communication hubs of an asocial cat are the source of a human–carnivore conflict and key to its solution. *PNAS*, 117(52):33325–33333, 2020. doi: 10.1073/pnas.2002487117

[45] Meta Platforms, Inc. Meta Quest Pro. `https://www.meta.com/quest/quest-pro/`, 2022.

[46] Microsoft Corporation. Mixed Reality Toolkit (MRTK). `https://github.com/microsoft/MixedRealityToolkit-Unity`.

[47] Microsoft Corporation. HoloLens 2. `https://www.microsoft.com/en-us/hololens`, 2019.

[48] M. Mittring. Finding next gen: Cryengine 2. In *Proc. SIGGRAPH*, pp. 97–121. ACM, 2007. doi: 10.1145/1281500.1281671

[49] A. K. Mutasim, A. U. Batmaz, and W. Stuerzlinger. Pinch, click, or dwell: Comparing different selection techniques for eye-gaze-based pointing in virtual reality. In *Proc. ETRA*, pp. 15:1–15:7. ACM, 2021. doi: 10.1145/3448018.3457998

[50] H. Nguyen, F. Wang, R. Williams, U. Engelke, A. Kruger, and P. De Souza. Immersive visual analysis of insect flight behaviour. In *Proc. Immersive Analytics workshop*, 2017.

[51] X. Olive and J. Morio. Trajectory clustering of air traffic flows around airports. *Aerosp. Sci. Technol.*, 84:776–781, 2019. doi: 10.1016/j.ast.2018.11.031

[52] S. Öney, N. Pathmanathan, M. Becher, M. Sedlmair, D. Weiskopf, and K. Kurzhals. Visual gaze labeling for augmented reality studies. *Comput. Graph. Forum*, 42(3):373–384, 2023. doi: 10.1111/cgf.14837

[53] Y. S. Pai, T. Dingler, and K. Kunze. Assessing hands-free interactions for vr using eye gaze and electromyography. *Virtual Reality*, 23:119–131, 2019. doi: 10.1007/s10055-018-0371-2

[54] N. Pathmanathan, S. Öney, M. Becher, M. Sedlmair, D. Weiskopf, and K. Kurzhals. Been there, seen that: Visualization of movement and 3d eye tracking data from real-world environments. *Comput. Graph. Forum*, 42(3):385–396, 2023. doi: 10.1111/cgf.14838

[55] A. Patney, M. Salvi, J. Kim, A. Kaplanyan, C. Wyman, N. Benty, D. Luebke, and A. Lefohn. Towards foveated rendering for gaze-tracked virtual reality. *ACM Trans. Graph.*, 35(6):179:1–179:12, 2016. doi: 10.1145/2980179.2980246

[56] Y. T. Paulus and G. B. Remijn. Usability of various dwell times for eye-gaze-based object selection with eye tracking. *Displays*, 67:101997, 2021. doi: 10.1016/j.displa.2021.101997

[57] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen. Gaze + pinch interaction in virtual reality. In *Proc. SUI*, pp. 99–108. ACM, 2017. doi: 10.1145/3131277.3132180

[58] K. Pfeuffer, L. Mecke, S. Delgado Rodriguez, M. Hassib, H. Maier, and F. Alt. Empirical evaluation of gaze-enhanced menus in virtual reality. In *Proc. VRST*, pp. 20:1–20:11. ACM, 2020. doi: 10.1145/3385956.3418962

[59] J. Ping, Y. Liu, and D. Weng. Comparison in depth perception between virtual reality and augmented reality systems. In *Proc. VR*, pp. 1124–1125. IEEE, 2019. doi: 10.1109/VR.2019.8798174

[60] P. Piotrowski and A. Nowosielski. Gaze-based interaction for vr environments. In *Image Processing and Communications*, pp. 41–48. Springer, 2020. doi: 10.1007/978-3-030-31254-1_6

[61] A. Plopski, T. Hirzle, N. Norouzi, L. Qian, G. Bruder, and T. Langlotz. The eye in extended reality: A survey on gaze interaction and eye tracking in head-worn extended reality. *ACM Comput. Surv.*, 55(3):53:1–53:39, 2022. doi: 10.1145/3491207

[62] K. Qian, T. Arichi, A. Price, S. Dall'Orso, J. Eden, Y. Noh, K. Rhode, E. Burdet, M. Neil, A. D. Edwards, and J. V. Hajnal. An eye tracking based virtual reality system for use inside magnetic resonance imaging systems. *Sci. Rep.*, 11(1):2045–2322, 2021. doi: 10.1038/s41598-021-95634-y

[63] M. Rezaie and N. Saunier. Trajectory clustering performance evaluation: If we know the answer, it's not clustering. *ArXiv*, 2021. doi: 10.48550/arXiv.2112.01570

[64] S. Rivu, Y. Abdrabou, T. Mayer, K. Pfeuffer, and F. Alt. Gazebutton: Enhancing buttons with eye gaze interactions. In *Proc. ETRA*, pp. 73:1–73:7. ACM, 2019. doi: 10.1145/3317956.3318154

[65] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proc. VL/HCC*, pp. 336–343. IEEE, 1996. doi: 10.1109/VL.1996.545307

[66] V. Tanriverdi and R. J. K. Jacob. Interacting with eye movements in virtual environments. In *Proc. CHI*, pp. 265–272. ACM, 2000. doi: 10.1145/332040.332443

[67] Unity Software Inc. Unity 3D. `https://unity.com/`.

[68] M. J. van Kreveld, M. Löffler, and F. Staals. Central trajectories. *J. Comput. Geom.*, 8(1):366–386, 2017. doi: 10.20382/jocg.v8i1a14

[69] X. Wang, L. Besançon, D. Rousseau, M. Sereno, M. Ammi, and T. Isenberg. Towards an understanding of augmented reality extensions for existing 3d data analysis tools. In *Proc. CHI*. ACM, 2020. doi: 10.1145/3313831.3376657

[70] C. Ware and H. H. Mikaelian. An evaluation of an eye tracker as a device for computer input2. *SIGCHI Bull.*, 18(4):183–188, 1986. doi: 10.1145/1165387.275627

[71] Y. Yang, T. Dwyer, B. Jenny, K. Marriott, M. Cordeil, and H. Chen. Origin-destination flow maps in immersive environments. *IEEE Trans. Vis. Comput. Graph.*, 25(1):693–703, 2019. doi: 10.1109/TVCG.2018.2865192

[72] S. Ye, Z. Chen, X. Chu, Y. Wang, S. Fu, L. Shen, K. Zhou, and Y. Wu. Shuttlespace: Exploring and analyzing movement trajectory in immersive visualization. *IEEE Trans. Vis. Comput. Graph.*, 27(2):860–869, 2021. doi: 10.1109/TVCG.2020.3030392

[73] D. Yu, Q. Zhou, J. Newn, T. Dingler, E. Velloso, and J. Goncalves. Fully-occluded target selection in virtual reality. *IEEE Trans. Vis. Comput. Graph.*, 26(12):3402–3413, 2020. doi: 10.1109/TVCG.2020.3023606

[74] G. Yuan, P. Sun, J. Zhao, D. Li, and C. Wang. A review of moving object trajectory clustering algorithms. *Artif. Intell. Rev.*, 47:123–144, 2017. doi: 10.1007/s10462-016-9477-7

[75] G. Zhang, J. P. Hansen, and K. Minakata. Hand- and gaze-control of telepresence robots. In *Proc. ETRA*, pp. 70:1–70:8. ACM, 2019. doi: 10.1145/3317956.3318149

[76] M.-J. Zhang, J. Li, and K. Zhang. Using virtual reality technique to enhance experience of exploring 3d trajectory visualizations. In *Proc. VINCI*, pp. 168–169. ACM, 2015. doi: 10.1145/2801040.2801072

[77] Y. Zheng. Trajectory data mining: An overview. *ACM Trans. Intell. Syst. Technol.*, 6(3):29:1–29:41, 2015. doi: 10.1145/2743025