Vision, Modeling, and Visualization (2020)
J. Krüger, M. Niessner, and J. Stückler (Eds.)

# A Design and Application Space for Visualizing User Sessions of Virtual and Mixed Reality Environments

Shivam Agarwal[1], Jonas Auda[2], Stefan Schneegaß[3], and Fabian Beck[4]

paluno, University of Duisburg-Essen, Germany

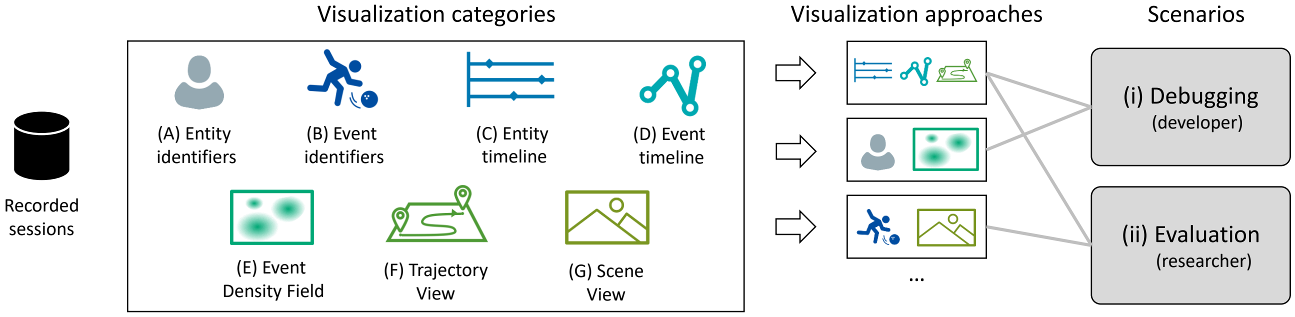

**Figure 1:** *Design and application space for visualizations of recorded virtual or mixed reality sessions; the seven categories of visual encodings (A–G) provide the building blocks of specific visualization approaches, which can be used in two scenarios: (i) debugging the environment and (ii) analyzing data from user studies in a research context.*

**Abstract**

*Virtual and mixed reality environments gain complexity due to the inclusion of multiple users and physical objects. A core challenge for developers and researchers while analyzing sessions from such environments lies in understanding the interaction between entities. Additionally, the raw data recorded from such sessions is difficult to analyze due to the simultaneous temporal and spatial changes of multiple entities. However, similar data has already been visualized in other areas of application. We analyze which aspects of these related visualizations can be leveraged for analyzing user sessions in virtual and mixed reality environments and describe a design and application space for such visualizations. First, we examine what information is typically generated in interactive virtual and mixed reality applications and how it can be analyzed through such visualizations. Next, we study visualizations from related research fields and derive seven visualization categories. These categories act as building blocks of the design space, which can be combined into specific visualization systems. We also discuss the application space for these visualizations in debugging and evaluation scenarios. We present two application examples that showcase how one can visualize virtual and mixed reality user sessions and derive useful insights from them.*

## 1. Introduction

With the advance and spread of the technology, virtual and mixed reality systems gain more and more complexity. A typical scene in such environments includes several entities: users and objects, where both can be either virtual or real. Blending reality and virtuality, users and real objects can be represented by avatars and virtual objects in a 3D virtual environment. The data recorded from these user sessions has spatial and temporal properties. Various entities come in contact and interact with one another, *e.g.*, two players move different objects in a puzzle game at the same time. Recording such user behavior thus, creates complex multi-stream spatio-temporal data. Developers of virtual and mixed reality applications and researchers in human-computer interaction must analyze this complex data to draw conclusions on, *e.g.*, user behavior or performance. Because of the blending of virtual and real spaces

---

[1] e-mail: shivam.agarwal@paluno.uni-due.de
[2] e-mail: jonas.auda@uni-due.de
[3] e-mail: stefan.schneegass@uni-due.de
[4] e-mail: fabian.beck@paluno.uni-due.de

as well as the multiple streams of actions, this is more challenging than exploring user interactions with traditional user interfaces.

Visualizations can help in the inspection of such data to gain useful insights and understand challenges of the interaction. First examples of visualizations have been applied, providing valuable insights into navigation behavior in virtual game worlds, but are not explicitly targeted at virtual or mixed reality environments [CRI06; DC11].

A holistic understanding of how to tackle multiple aspects of the data of such sessions is still missing. However, there are visualization approaches for various related domains, such as analyzing traditional user interactions, eye movements, physical motion, and stories. We can borrow visualization techniques and, on a sample of these techniques, systematically explore the space of visualization designs for this context and identify future research challenges.

In this paper, we investigate a design and application space for visualizations helping to analyze recorded user sessions of virtual and mixed reality environments. A *visualization design space* (sometimes called *visualization taxonomy*) is a theoretical framework to systematically describe visualization design options for a specific data type or application. It can be used to explore and compare different visualization designs, to analyze best practices and common design patterns, as well as to identify future research challenges. To derive such a framework, we classify existing visualizations from related domains and propose seven categories, where each category is focused on showing specific aspects of the data (Figure 1, left). The categories can be used as building blocks for different visualization approaches (Figure 1, middle). These approaches are applicable in development- and research-oriented scenarios (Figure 1, right). We demonstrate the application space by tailoring visualizations from these categories for two application examples: evaluating user interaction with a tangible virtual object and a collaborative problem–solving scenario. With this, we make a first step into a bigger, largely unexplored research direction and formulate future research challenges. We do not focus on a single visualization solution, but instead discuss possible and useful visual encodings of the relevant data generated in such environments. Please also note that, in contrast to *immersive analytics*, which leverages visual data representations in a 3D immersive environment, the visualizations discussed in this work are not necessarily part of the mixed reality scene, but are used in a separate analysis interface.

The main contributions of this work comprise the following aspects. First, we survey a diverse sample of visualizations approaches used in related fields (section 2). Second, we derive and design an application space, including a categorization of visualizations important for understanding data generated in virtual and mixed reality settings as well as a discussion of application scenarios (section 3). Third, we report on two application examples in which we used such visualizations to gain an in-depth understanding of how users interact with a system (section 4). Fourth, we discuss future research challenges based on the previously introduced framework and applications (section 5).

## 2. Related Areas and Visualizations

To find suitable visualizations, we explored related approaches of neighboring fields. These neighboring fields, which are discussed in the following, share similarity with the targeted application with respect to visualizing similar types of data. We follow a qualitative sampling approach, selecting a diverse set of examples with the goal of covering a broad range of approaches, but not of quantifying how frequently the discussed solutions appear in the literature. Please note that there is an abundance of research on the analysis of spatio-temporal data (*e.g.*, gesture recognition, movement performance analysis, event classification, and event sequence mining). However, the focus of our work is visual analysis of the extracted events and interactions from the user session data.

**Interactions:** Some approaches target the visualizations of user sessions and the respective interactions performed by the users. Blascheck et al. combine the data generated in a user study through recordings of the user's interaction and a think-aloud protocol [BJK*16]. Interactions and thinking aloud actions are treated as events. Their visualization showed the temporal sequence of the events along with the respective regions of the interface for every participant. This helped them to compare the behavior of different participants, confirm several hypotheses, and gain useful insights. Similar visualizations show the interactions of software developers with the integrated development environment (IDE) [MMLB14; MML15; YMK13]. These visualizations also show time on a horizontal axis, whereas the vertical axis represents different source code files and dialog boxes of the IDE. Similar techniques are used to show group social dynamics of human interactions in meetings [YYA*10]. Similarly, virtual and mixed reality environments involve user interaction with the respective environment and among multiple users.

**Eye Tracking:** Eye tracking studies record eye movements of a user watching or interacting with a static stimulus (picture) or dynamic stimulus (video or interactive interface). Some approaches visualize the gaze behavior of users for dynamic stimulus by showing time in the $z$ axis and position in the $x$ and $y$ axes (space–time cube) [KW13; KHW14]. Other techniques visualize gaze behavior in virtual reality using 3D scanpaths, attention maps, and linked view visualizations, with the stimulus being an immersive video [LSF*15] or a virtual 3D scene [SND10]. Visual analytics approaches support the comparison of different users by representing their gaze behavior while abstracting the real stimulus [BJK*16]. Blascheck et al. survey further visualization approaches [BKR*17]. For virtual and mixed reality user sessions, eye movement data can also be recorded [LSF*15], but the topic is relevant only if it is applied to combinations of an interactive stimulus and human body movement.

**Physical Motion:** Beyond the movement of the eye, there exist other visualizations of physical motion, such as (a) individual trajectories [AAG00; Kwa00; Kra03], (b) segments of the trajectories to explore local movement patterns [WAPW06; GWY*11; WE12], (c) aggregations of multiple movement trajectories [BAA*11; ZFH08; MIC09], and (d) the environment along with the movement to preserve its context [AAH11; TM10]. These visualizations use different techniques, such as static and animated maps, interactive space–time cubes, time lenses for trajectories in small segments, or color for density fields. Visualizations focusing particularly on motion capture data are used to show clusters of human poses, encode them with a gradient color scale, and then spatially position them in the order of their occurrence [BWK*13]. Some visualizations already show the movement of entities from sessions captured in virtual reality while abstracting details of the environment [CRI06; MIC09; DC11].

**Stories:** Storyline visualizations provide a visual summary of a sequence of events involving different entities. An early implementation used a storyline to summarize plots of movies [Mun] where each character of the movie is shown as a separate horizontal line and the $x$-axis represents time. The lines can bend and be grouped when the respective characters are in the same film location (i.e., when they interact in a movie frame). The idea has been adopted in different fields; one such example is visualizing software evolution through storylines [OM10]. A time curves visualization shows similarity of events through spatial proximity while preserving their temporal sequence [BSH*16]. These visualizations can be used to show data from virtual and mixed reality environments to convey a coherent and complete overview of the user sessions.

A recent work by Nebeling et al. comes closest to our work, as they propose a toolkit that visualizes data gathered from mixed reality sessions [NSW*20]. However, they focus on enabling the construction of visualizations through visual editors without programming. Additionally, unlike our approach, they do not discuss alternate ways of encoding specific aspects of the data from mixed reality sessions. Since their work is recent and the only one of its kind in visualizing mixed reality session data (to the best of our knowledge), we do not discuss it in a separate category and instead reference it across the paper, where appropriate.

## 3. Design and Application Space

Card and Mackinlay [CM97] state that the purpose of a *visualization design space* is *"to understand the differences among designs and to suggest new possibilities."* In the visualization literature, a variety of general visualization design spaces and taxonomies have been discussed [Shn96; CM97; Chi00]. In these theoretical frameworks, data models, visualization categories, and tasks often form the key elements. To study visualization options on a more fine-grained level, some works tailor such design spaces to specific types of data and visualization (e.g., dynamic graphs [BBDW17; KKC14], composite visualizations [JE12], word-sized graphics [GWFI14; BW17]) or applications (e.g., eye tracking visualization [BKR*17], software visualization [MKNW19], or games visualization [BEJ12]). But despite the variety of such existing frameworks, we are not aware of any work targeting such a tailored visualization design space for user sessions in virtual and mixed reality environments.

To structure the design space, we first introduce the data that is recorded and analyzed, then provide a categorization of visualizations. Since we also discuss application scenarios, we call the suggested framework *design and application space*. Unlike most other related frameworks, which structure the visualizations based on examples from within the respective domain, we have to work with the examples from the related domains discussed above because there is not yet sufficient coverage within the domain (i.e., the visualization of user sessions of virtual and mixed reality).

### 3.1. Data

Users in the real world usually wear a head-mounted display (HMD) to perceive the virtual environment. The user's head and controller positions are synchronized between the real and virtual environments. Mixed reality environments introduce further entities in both real and virtual worlds, such as tangible objects.

We summarize real and virtual entities as vertices $v \in V$; they can be active agents such as users, virtual avatars, or physical robots, but also include passive objects such as controllers, virtual objects, and tangible (real) objects. Each vertex $v \in V$ at a time $t \in T$ has a position in three-dimensional space (and an orientation therein defined by three angles), i.e., it can be described with the function $p : V \times T \to \mathbb{R}^3$ for positions (or including orientation: $p : V \times T \to \mathbb{R}^6$). The position and orientation of a rigid object can be described by a single vertex, while non-rigid bodies would require multiple entities (usually connected by a predefined skeleton). For objects with both virtual and real representations, we discern a position in reality $p_{\mathcal{R}}(v,t)$ and in virtuality $p_{\mathcal{V}}(v,t)$. Note that, depending on the scenario, several of such real or virtual spaces might exist, for instance, in a physically distributed collaborative game. The existence of multiple spaces differentiates our scenario from most other scenarios, which only deal with a single space. While the positions of virtual objects are known, motion tracking of real world entities is required to determine their positions and orientations. Entities equipped with markers are traced through optical trackers (e.g., OptiTrack). Other sensors, such as Kinect, do not need dedicated markers to be attached to the body of entities. Data about the different parts of the hands (i.e., finger joint positions or palm orientation) are tracked and stored through sensors such as Leap Motion (which can be attached to the head-mounted display).

Active objects can trigger events, but passive objects might also interact (e.g., collide). Events (E) are inherently temporal and we model the time (or time span) of an event as a function $t : E \to T$ (or $t : E \to T^2$ for time spans). An event might represent an interaction with one or multiple objects, identified by a set of vertices and mapped as $V : E \to 2^V$. From these object–event relationships, we can also derive the set of events $E(v,t)$ that interact with an entity $v \in V$ at point $t \in T$, as well as the interacting entities $V(v,t)$ of an entity $v \in V$ at point $t \in T$. The events $e \in E$ can further be discerned by whether they are (global events) or have a location of occurrence in reality or virtuality (local events). Local events, like objects, carry positions $p : E \times T \to \mathbb{R}^3$ in reality ($p_{\mathcal{R}}$) or virtuality ($p_{\mathcal{V}}$). Events in virtuality, such as actions of active virtual avatars or collisions of passive objects, can be easily recorded as log files. For mixed reality objects, different data streams must be merged to detect the respective events. Events can be triggered in the real world through input and sensing devices, such as controllers. More sophisticated types of motion such as gestures can also be extracted from the recorded position data of users.

The data recording can be complemented with a holistic recording of the real or virtual scene. The real scene can be captured

**Table 1:** *Classification of existing visualizations into seven categories for analyzing user behavior in mixed reality.*

| | Interactions | Eye tracking | Physical motion | Stories |
|---|---|---|---|---|
| (A) Entity identifiers | [MMLB14; MML15] | [BJK*16; KHW14; KW13; LSF*15] | [BWK*13; YYA*10; AAG00; Kwa00; GWY*11; BAA*11; TM10] | [THM15; TM12; TRL*19; WBG06] |
| (B) Event identifiers | [MMLB14; MML15; DC11] | [BJK*16; KHW14; KW13; LSF*15; SND10] | [AA13; YYA*10; WAPW06; ZFH08; TM10] | [THM15; TRL*19; BSH*16; WBG06] |
| (C) Entity timeline | [MMLB14; MML15] | [BJK*16; LSF*15] | [YYA*10] | [THM15; TM12; TRL*19; WBG06] |
| (D) Event timeline | [MMLB14; MML15] | [KW13; KHW14] | [BWK*13; Kra03; ZFH08] | [BSH*16] |
| (E) Event density fields | [TRL*19; DC11] | [KW13; KHW14; LSF*15; SND10] | [BWK*13; Kwa00; GWY*11; ZFH08; BAA*11; DC11] | [BSH*16; TRL*19] |
| (F) Trajectory view | [BJK*16; YMK13] | [GH10; KW13; KHW14] | [CRI06; AAG00; Kra03; Kwa00; WAPW06; GWY*11; ZFH08; MIC09; DC11] | |
| (G) Scene view | [Mun; TRL*19] | [KW13; KHW14; LSF*15; SND10; BJK*16] | [BWK*13; YYA*10; AAG00; Kra03] | [Mun; WBG06] |

on camera, while the virtual scene can be recorded for interactive playback or emulating a virtual camera. A scene $s \in S$ can either be a two-dimensional image as recorded by a camera or a three-dimensional capture of the scene, which can be interactively explored. It is possible to map each timestep to an image of the real or virtual scene $s : T \to S$.

### 3.2. Visualization Designs

To classify the related visualizations, we systematically explored them and assigned certain keywords. The keywords reflected concepts (e.g., *time, event icon, summary*) that are useful for visual analysis of data recorded from mixed reality user sessions. Based on similarity of data property, we grouped the keywords. As a result of the grouping, we generated seven categories. Table 1 shows these categories along with references to the related publications. We describe each category in the following paragraphs, applying them to the study of user behavior in virtual and mixed reality.

**(A) Entity Identifiers:** An entity $v \in V$ can be either a user or an object in a virtual or mixed reality environment. Entity identifiers are used to uniquely identify each user/object present in the environment. Different visual encodings such as text, icons, colors [BWK*13; LSF*15; Kwa00; GWY*11; BAA*11; TM10; NSW*20], and position [BJK*16; THM15; TM12; TRL*19; WBG06] can be used to represent them. Similarity between entities can be shown by a dendrogram [BWK*13; BJK*16; KHW14]. These identifiers are often used in combination with visualizations from other categories such as *entity timeline*, *trajectory view*, and *scene view*.

**(B) Event Identifiers:** These identifiers are used to uniquely identify each event $e \in E$ that occurred in the session (or type of event, respectively). Different visual encodings such as text [WBG06], icons [LSF*15], shapes [TRL*19; TM10], colors [BJK*16; BSH*16; MMLB14; MML15; SND10; YYA*10; WAPW06; ZFH08; DC11; NSW*20], and position [KW13; WAPW06] are used to represent events. Their usage is most often in combination with visualizations from other categories, such as *entity timelines* [BJK*16; MMLB14], *event timelines* [BWK*13], and *trajectory views* [WAPW06; TM10].

**(C) Entity Timeline:** An entity represented by $v \in V$ has features that change over time, for instance, associated events $E(v,t)$, interactions with other entities $V(v,t)$, or other attributes. The visualizations of entities in this category show a temporal sequence of these features. Besides the timeline, the dominating visual structure of the visualization is a set of entities $V' \subset V$, for instance, encoded in lines or as rows of the timeline. It is common to represent time on the horizontal axis [MMLB14; MML15; BJK*16; LSF*15; YYA*10; TM12; TRL*19; THM15; WBG06]. Entity timelines can be drawn for individual [KW13; KHW14; Kra03] or multiple entities [TRL*19; BJK*16; MMLB14; MML15; WBG06].

**(D) Event Timeline:** Although they also show a timeline, visualizations in this category focus on representing a set of events $E' \subset E$ and their temporal sequence of occurrence as primary visual glyphs. They are often discerned by their event type, which provides a structure for the timeline. An event timeline can be represented in linear [KHW14; KW13; MMLB14; MML15; SND10; NSW*20] and non-linear [BWK*13; BSH*16; ZFH08] layouts. It is common to encode the time span of events by the size/area of the glyph [TRL*19; KW13; KHW14; SND10; MMLB14; MML15], and also by the relative distance between event identifiers [BSH*16].

**(E) Event Density Fields:** Groups of local events $E' \subset E$ are associated with positions $p(e,t)$ $(e \in E')$ and other attributes such as involved entities $V(e,t)$ $(e \in E')$. An event density field shows information of event sets $E'$ aggregated across time $t \in T$ through histograms [BJK*16; GWY*11; ZFH08], heatmaps in a 2D spatial context [KHW14; KW13; SND10; LSF*15; DC11], size/area of glyphs [TRL*19; BAA*11], or 3D surfaces [Kwa00]. These visualizations can be augmented with context to highlight additional attributes, for instance, representing event density on the map juxtaposed with another view showing linked static entities [DC11, Sect. 4.1]. A cluster of closely placed event glyphs also represents the density of events in a timeline [BSH*16; BWK*13; NSW*20]. Different patterns of clusters [BSH*16] can be used to compare event timelines of multiple mixed reality user sessions.

**(F) Trajectory View:** This category includes visualizations that show movement $p(v,t)$ of the entities $v \in V$ across time $t \in T$. The movement is usually shown by projecting position on two dimensions [NSW*20] and representing time through either a gradient color scale [CRI06; AA13; Kwa00; ZFH08] or a third dimension [KW13; KHW14; GH10; Kra03]. Direction of movement is also shown by glyphs [AAG00; WAPW06; GWY*11; NSW*20]. Details of the position can be abstracted by projecting it on the $y$ axis while showing time on the $x$ axis [BJK*16; YMK13]. The trajectory can be enriched by visualizing additional attributes of entities, for instance, showing trajectories of entities with different colors to represent different types of objects carried by a player in a virtual game [DC11, Sect. 3].

**(G) Scene View:** This category includes visualizations that show scenes $s(t)$ along time $t \in T$. The environment details can be abstracted in these visualizations [CRI06; BWK*13; Mun; TRL*19; WBG06] where the level of abstraction depends on the data analysis task. Techniques used in these visualizations include using multiple images (as keyframes like in a comic strip) [KW13; KHW14] and video/animation [KW13; KHW14; LSF*15; BWK*13; Kra03]. Scenes from both real [YYA*10; BJK*16; Kra03; NSW*20] and virtual worlds [LSF*15; SND10] can be included to provide a complete overview of the user session.

The above categories of visualizations can occur independently. However, they are often mixed with one another (see references that occur multiple times in Table 1). Multiple views that are synchronized by brushing-and-linking interactions provide a simple solution for this. But it is also possible to combine several of these categories within an integrated representation, for instance, a three-dimensional *space–time cube* of entities combines an *entity timeline* (☰; one axis) with a *trajectory view* (⤳; two remaining axes) [Kra03; KW13; KHW14]. Figure 1 illustrates this combination as a selection of visualization categories that are connected to a specific application approach.

### 3.3. Application Scenarios

We complement the design space of visualizations with a description of its applications. Instead of discussing visual analysis tasks for a specific virtual or mixed reality application, we focus on two general scenarios. These scenarios cannot be directly derived from related areas, but are based on the existing literature and our own experience with virtual and mixed reality environments.

**(i) Debugging:** Recent works by Ashtari et al. [ABM*20] and Speicher et al. [SHY*18] systematically derived and reported on challenges in the development of virtual/augmented reality applications. They include (a) facing too many unknowns in development, testing, and debugging [ABM*20] and (b) the need to work with multiple types of devices [SHY*18, Section 5]. The challenges are usually addressed by analyzing the log files, system messages, and source code to identify errors. Developers of the environment need to spend a lot of time debugging its design, judging the effectiveness and accuracy of interactions, and addressing usability obstacles. Visualizing the session data, which consists of multiple streams from different devices, can be helpful in addressing these challenges. Visualizations have been found to be useful for supporting developers in debugging and designing several aspects of virtual reality environments [WBG06, Sect. 4]. However, the visualizations do not include spatial information and they are limited to specific environments. Visualizations that show multiple aspects of a mixed reality environment are important for supporting developers in effective debugging and designing of the virtual and mixed reality applications.

**(ii) Evaluation:** A challenge HCI researchers face while evaluating user studies is understanding complex movement and behavior patterns of multiple users interacting in mixed reality. The user data has multiple degrees of freedom, and it is challenging to map the data of multiple users or other entities in such a way that patterns (e.g., two entities being at the same position at the same time) become visible. Without any alternate representation of the recorded data, it becomes difficult to verify and evaluate the data itself. Ashtari et al. [ABM*20] also highlighted evaluation challenges in understanding details of specific situations (e.g., a stimulus that distracted the user). Additionally, researchers need to analyze data from multiple sessions to evaluate the design of the proposed novel features for the environment. Visualizations can help in addressing these challenges. Hence, they are important for supporting an initial analysis of user study data and for conducting qualitative studies.

## 4. Application Examples

We present two application examples to illustrate how the design and application space can be used to make novel visual analysis solutions of user sessions from virtual and mixed reality. While the first application is simpler, and we discuss it along with an early-stage visualization prototype, the second one is more complex, and we present a deeper analysis of results with a more sophisticated tool. With these examples, we aim to cover a diverse set of visualization categories (see icons used in text and figures) as well as the two identified general scenarios (first example: debugging; second example: evaluation). In each example, we first introduce the design of the visualization, then present specific insights derived from the visualization, and conclude with a brief discussion.

### 4.1. Tangible Virtual Object

In the first example, a user interacts with a plastic foam bar in the real world that is shown as a piece of dynamite in the virtual environment for a gaming scenario. The game object $v_o$ should be perfectly mapped in position and orientation from the real world to the virtual world $(p_{\mathcal{R}}(v_o,t) \mapsto p_{\mathcal{V}}(v_o,t))$. In the game, interactions with the object include picking it up, holding it, and throwing it. These events need to be automatically detected by the virtual reality system from the recorded data to allow further actions to be triggered (e.g., scoring; not implemented). Markers are placed on the physical object and on the head-mounted display. Both entities are tracked by optical trackers (OptiTrack and Kinect) as well as

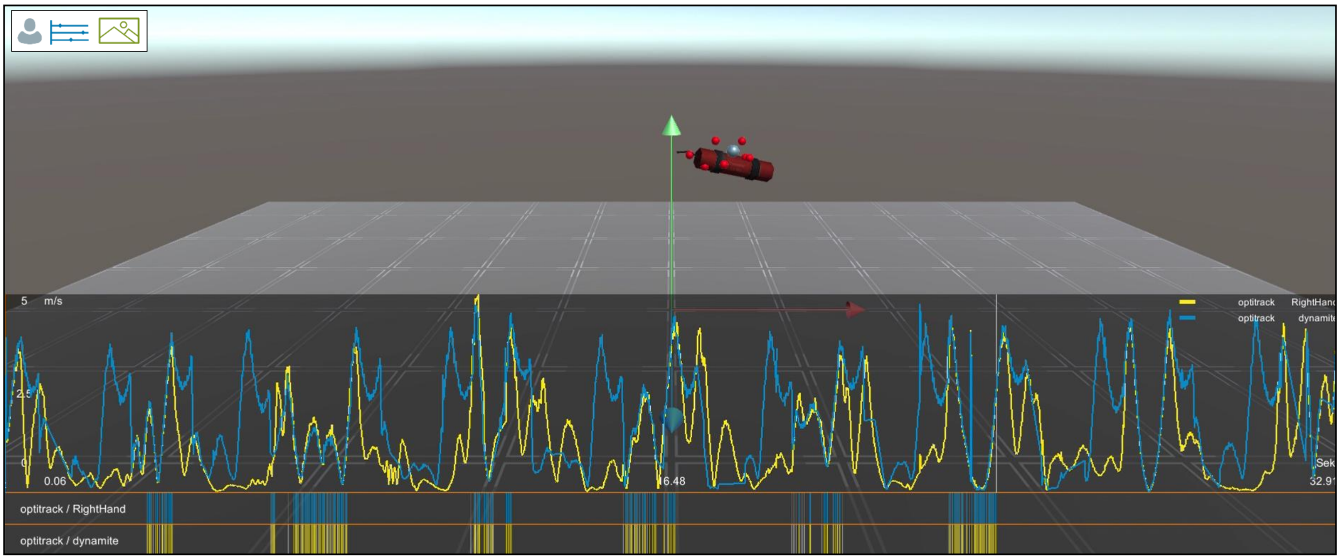

**Figure 2:** *Tangible virtual object application example – A* scene view *shows the virtual scene consisting of the game object and an abstracted player's hand (small spheres), with an* entity timeline *below providing velocity data of the two entities as well as detected collision events.*

the player's hand (Leap Motion). The trackers allow tracing of the movement of the user's head $v_u$, the user's dominant hand $v_h$, and the physical object $v_o$. All these entities are mapped to the virtual world in real time. We study a *debugging* scenario, where the goal is to help developers resolve tracking issues and fine-tune event detection. The designers of the environment need to ensure that user actions and event executions are synchronized.

**Visualization Design:** Figure 2 shows the visualization prototype, which consists of a *scene view* of the virtual world (top, 🖼) and an *entity timeline* (bottom, ☰). The scene can be played back as a video from a static camera perspective to follow the virtual object $v_o$ and an abstracted virtual hand $v_h$ (small spheres). The entity timeline visualizes the velocity of the hand entity $v_h$ (blue line) and the game object $v_o$ (yellow line)—the two colors can be considered *entity identifiers* (👤). Below the line chart, still part of the *entity timeline*, information of detected events is provided: Using a collider method, it is automatically detected whether the player holds the game object, visualized as a colored bar. This is the basis for detecting events for throwing and catching (not implemented).

**Insights:** The developers can quickly see through the line chart in the *entity timeline* where the movement (velocity) of the game object (yellow line) aligns with the movement (velocity) of the hand (blue line). Though summarized as only one movement attribute, this is already a good indicator for when the player holds the game object. The marked stripes below show the start and end of each hold phase and could indicate the catch and throw events. However, the current filter does not work reliably enough and requires some debugging: Small and bigger gaps in the phases indicate problems where the movement of the two entities aligns but a collision is not detected. Reasons for these issues might be a too strict collision check, losing track of the physical representations of the objects, or other bugs in the implementation. To investigate what exactly caused the gaps, the developers can playback the scene at the given moment and see how the virtual environment behaves. A video feed of the real world (not implemented) may provide further context to reveal the circumstances causing the issue.

**Discussion:** Debugging a virtual or mixed reality application is often tedious due to the many devices and systems that need to be synchronized and mapped [SHY*18, Section 5]. Standard code debugging tools are of limited use as issues often refer to the interactions between these technical components. Also, a simple playback of the recordings of a single data stream is not sufficient. Different data streams must be viewed together to understand typical problems as well as an overview of time to spot critical points. In our example discussed above, we already combined some data streams, but further integration of data streams would be helpful (such as synchronized video playback of the real world scene). Generally, the impact of a visual debugging tool can reach beyond low level calibration and parameter tuning: usability problems can be detected, for instance, people failing to catch an object because of misalignment of virtual and real objects or gestures not being correctly recognized.

### 4.2. Remote Collaboration

In this example, two participants collaborate in one shared virtual environment. They interact with each other and various objects, although they are not co-located in the same physical space. In the recorded scenario, the participants sit in different rooms (cf., Figure 3, Location 1 & Location 2). However, in the virtual environment, they appear to be sitting at one table facing each other (cf., Figure 3, Virtual World), allowing world-wide immersive collaboration. Voice is recorded and streamed to the respective other location so that the participants can hear each other. Both participants have tiles on their desks. The positions and orientations of

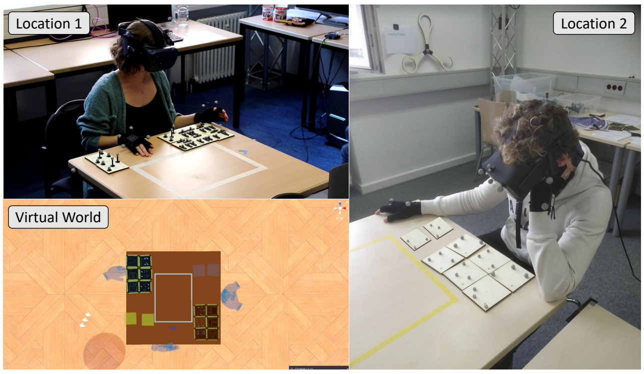

**Figure 3:** *Remote collaboration application example – Two participants in different locations collaborate in virtual reality; together they figure out a certain arrangement of the components in the virtual world to solve a puzzle.*

the tiles are tracked optically and synchronized with the virtual environment. Hence, every user can see a virtual representation of the tiles of the other user. The objective of the collaboration scenario is to arrange all the tiles according to a plan each collaborator has only a part of. The participants must collaborate to complete the puzzle. With this application, we target a research scenario where a visualization should support the qualitative *evaluation* of user sessions. The tool[1] is shown in Figure 4 and is available in the supplemental material [Aga20] for this paper.

**Visualization Design:** Virtual reality designers and researchers need to compare users based on their actions, how they communicated with each other, and their interactions with the environmental objects. To fulfill these requirements, we incorporate an *entity timeline* (☰) as shown in Figure 4a. We use color of the glyphs to identify different types of events (*event identifier*, 🏃) shown as a legend in Figure 4b. Since several entities can be involved in an event, the event timeline should represent connections between involved entities as well. We took inspiration from PAOHvis [VBP*19] and extended the design to show events and entities. We chose a matrix layout for the entity timeline where the horizontal axis represents time (from left to right), and the vertical axis lists entities in individual rows (Figure 4c). The length of each scene corresponds to the width of the visualization and is annotated below. To visually represent the density of events, we integrate a histogram that shows *event density fields* (▦), where the size of each bin is set by default to six seconds (Figure 4d). Verbal communication can also be considered as an event; we show the density of their conversation by a waveform visualization (Figure 4e). We integrate a *scene view* (🖼) component that plays the recording of selected virtual scene (Figure 4f). A red vertical line across all plots represents the current position of the playback.

**Insights:** To evaluate the collaborative game, the visualization helps in analyzing the strategies of the players and hints at the obstacles they faced. We illustrate this with the six scenes of three pairs of players shown in Figure 4. In Session 3 – Scene 2, the width of the column (participants P5 and P6) has the shortest duration compared to all other scenes. Also, fewer *Movement Actions* related to objects indicate the high effectiveness of the players. Further, a strategic pattern can be observed: First, the collaborator in Location 1 starts to interact three times with three different objects in Location 2 (three consecutive green dots linked to the object *Location1-Blank-2*). Afterwards, the collaborator in Location 2 does the same but with *Location2-Blank-1* and *Location2-Blank-2*. Both players communicate little (cf., Figure 4e), indicating that verbal interaction was not as needed as in other scenes. In contrast, in Session 1 – Scene 2, which is the longest, we can derive a less efficient pattern. The first interaction starts after some verbal exchange. Then, the collaborators begin slowly using the objects to interact with the objects of the other location. In the middle, we can observe a pair of consecutive dots in the same two rows indicating that there was a mistake in the interaction. To further investigate this specific part of the scene, we can listen to the audio and playback the scene view to gain insight on what went wrong. Finally, the last interaction between *Location1-Blank-2* and *Location-2-OBJ-2* is interesting, as its finish event is not very close to its start and further, *Location-2-OBJ-2* is moved again. At the same time, verbal exchange increases, indicating that there was a discussion. In Session 2 – Scene 2, three objects at both locations are moved at the beginning of each scene (the three consecutive *Movement actions*). This indicates that each player can fulfill certain actions without collaborating. This seems to be a common pattern: The inter-location interaction in all scenes starts after these three actions.

**Discussion:** The application example shows that users adopt different collaborative strategies to solve a task. Using the visualization, we were able to discover some of the strategies and identify similar behavioral patterns. We considered entity interactions and temporal aspects of the data. For other realistic applications, the spatial aspect could play an important role as well and can be visualized using a *trajectory view*.

## 5. Future Challenges

While we observe that building on and extending existing visualization solutions borrowed from related fields already gives first insights, we consider this work to be only a first step into the direction of visually analyzing user sessions in virtual and mixed reality environments. Based on trends we observed in related fields and experience gained when designing the examples provided in this work, we identified the following challenges that require further research for effectively visualizing such user sessions.

**Dual Representations:** A characteristic feature of mixed reality applications is that some entities have dual representations, one in reality and one in virtuality. For some analysis use cases, it might suffice to fuse the two representations in the visualization. However, other use cases might require developers and researchers to study divergences and occasional misalignment of the two representations, because these can be critical obstacles for the perceived immersion. Novel visualizations must be developed for visualizing spatio-temporal data of such dual representations.

---

[1] Hosted at: `https://vis-tools.paluno.uni-due.de/vr_mr_vis/`

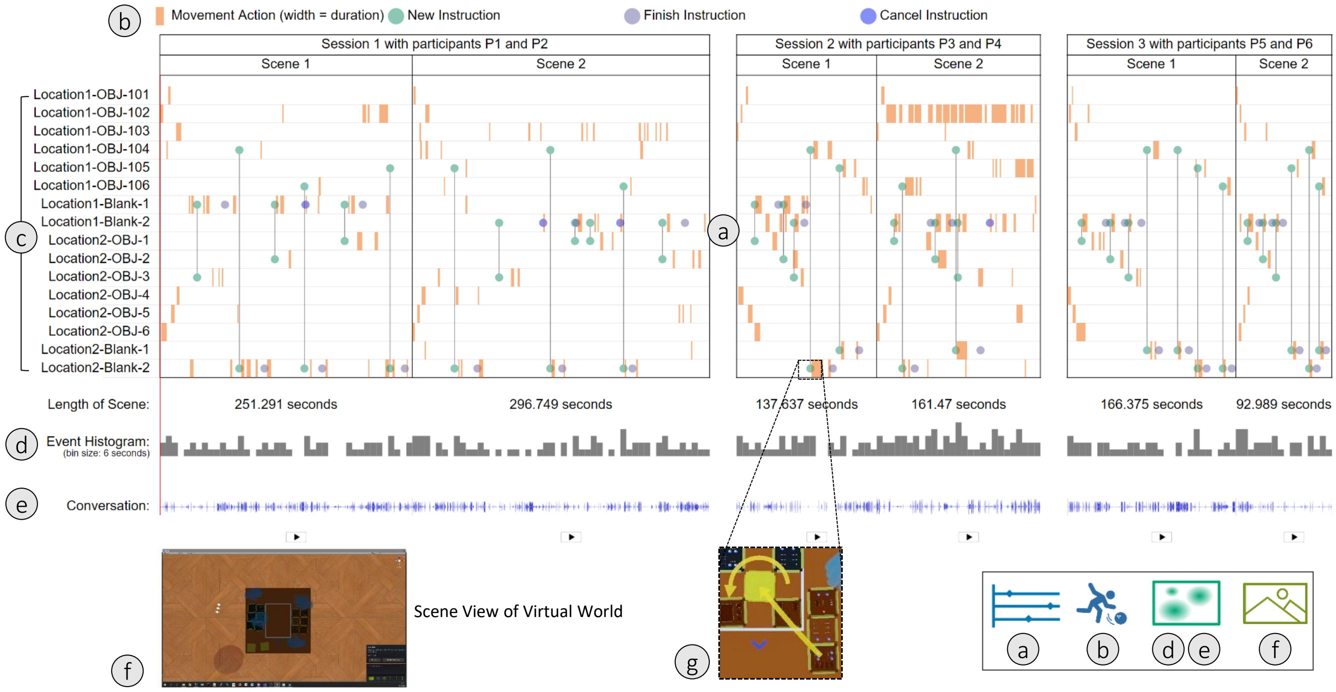

**Figure 4:** *Remote collaboration application example – (a) A timeline visualization showing sessions of a mixed reality environment. (b) Colored glyphs are used to identify events, while (c) entities (users and objects) are shown in separate rows. A vertical line between two rows denotes interaction (touch) between the corresponding entities. Density of events and recorded conversation between participants are visualized through (d) histogram and (e) waveform, respectively. (f) Playback of videos recorded from the virtual environment for each scene.*

**Diverse Data and Dynamics:** Recordings of virtual and mixed reality sessions involve diverse data streams such as trajectories, events, video, and audio. Even less–complex interactive scenarios like using a desktop interface require sophisticated visual analytics solutions when integrating several data streams [BJK*16; MMLB14]. The complexity rises with movement in a 3D environment, with a virtual and real scene blended over, and with multi-user scenarios. Interactions in mixed reality are highly dynamic and several lines of action might run in parallel. For instance, users might perform different tasks at the same time. Only focusing on a single mouse pointer, as is done when analyzing interactions in a traditional desktop application, is not possible.

**Comparison and Abstraction:** Analyzing individual user sessions might provide some insight, but is limited. Only after considering several user sessions could reveal typical usage strategies, common obstacles, and relevant misalignment. Visual comparison and aggregation of user sessions need to be supported. To compare different interactive (i.e., individual) sessions, temporal alignment, and detection of similar behavior and actions become important. However, the recorded data reflects the users' actions on a low-level granularity. For comparatively evaluating the strategies employed by different users, we need to develop meaningful abstractions that can be reliably detected in an automatic or semi-automatic process. This process should be embedded in the visualization interface because analysts might need to adapt the definition of certain high-level strategies during the analysis.

## 6. Conclusions

We proposed a design and application space for visualizing data from mixed and virtual reality user sessions. In two examples, we showed how visualizations from the different categories could be combined and used in one debugging and one evaluation scenario. This illustrates the potential of visualizations to support researchers and developers in creating visualizations to gain insights on user behavior within mixed and virtual reality scenes. Our design and application space can support systematic exploration of further advanced visualizations for this purpose and help to transfer ideas from related more-explored areas to this new area of application. While basic insights can be gained already by reusing existing approaches from these neighboring fields, our vision is to provide better support for developers and researchers of mixed reality environments through tailored visualization and visual analytics solutions.

## Acknowledgments

We are thankful to the student members of the Master project group VinteR at the University of Duisburg-Essen for implementing the first application example. This research was partly funded by MERCUR (project: "Vergleichende Analyse dynamischer Netzwerkstrukturen im Zusammenspiel statistischer und visueller Methoden").

## References

[AA13] Andrienko, Natalia and Andrienko, Gennady. "Visual analytics of movement: An overview of methods, tools and procedures". *Information Visualization* 12.1 (2013), 3–24. doi: 10.1177/1473871612457601 4, 5.

[AAG00] Andrienko, Natalia, Andrienko, Gennady, and Gatalsky, Peter. "Supporting Visual Exploration of Object Movement". *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI. Palermo, Italy: ACM, 2000, 217–220. isbn: 1-58113-252-2. doi: 10.1145/345513.345319 2, 4, 5.

[AAH11] Andrienko, Gennady, Andrienko, Natalia, and Heurich, Marco. "An event-based conceptual model for context-aware movement analysis". *International Journal of Geographical Information Science* 25.9 (2011), 1347–1370. doi: 10.1080/13658816.2011.556120 2.

[ABM*20] Ashtari, Narges, Bunt, Andrea, McGrenere, Joanna, et al. "Creating Augmented and Virtual Reality Applications: Current Practices, Challenges, and Opportunities". *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. ACM, 2020, 1–13. isbn: 9781450367080. doi: 10.1145/3313831.3376722 5.

[Aga20] Agarwal, Shivam. *Supplementary Material: A Design and Application Space for Visualizing User Sessions of Virtual and Mixed Reality Environments*. 2020. doi: 10.17605/OSF.IO/2HSTC 7.

[BAA*11] Bremm, Sebastian, Andrienko, Gennady, Andrienko, Natalia, et al. "Interactive Analysis of Object Group Changes over Time". *Proceedings of International Workshop on Visual Analytics*. EuroVA. The Eurographics Association, 2011. isbn: 978-3-905673-82-1. doi: 10.2312/PE/EuroVAST/EuroVA11/041-044 2, 4.

[BBDW17] Beck, Fabian, Burch, Michael, Diehl, Stephan, and Weiskopf, Daniel. "A Taxonomy and Survey of Dynamic Graph Visualization". *Computer Graphics Forum*. CGF 36.1 (2017), 133–159. doi: 10.1111/cgf.12791 3.

[BEJ12] Bowman, Brian, Elmqvist, Niklas, and Jankun-Kelly, TJ. "Toward visualization for games: Theory, design space, and patterns". *IEEE Transactions on Visualization and Computer Graphics* 18.11 (2012), 1956–1968. doi: 10.1109/TVCG.2012.77 3.

[BJK*16] Blascheck, Tanja, John, Markus, Kurzhals, Kuno, et al. "VA$^2$: a visual analytics approach for evaluating visual analytics applications". *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), 61–70. issn: 1077-2626. doi: 10.1109/TVCG.2015.2467871 2, 4, 5, 8.

[BKR*17] Blascheck, Tanja, Kurzhals, Kuno, Raschke, Michael, et al. "Visualization of eye tracking data: A taxonomy and survey". *Computer Graphics Forum* 36.8 (2017), 260–284. doi: 10.1111/cgf.13079 2, 3.

[BSH*16] Bach, Benjamin, Shi, Conglei, Heulot, Nicolas, et al. "Time Curves: Folding time to visualize patterns of temporal evolution in data". *IEEE Transactions on Visualization and Computer Graphics* 22.1 (2016), 559–568. issn: 1077-2626. doi: 10.1109/TVCG.2015.2467851 3–5.

[BW17] Beck, Fabian and Weiskopf, Daniel. "Word-Sized Graphics for Scientific Texts". *IEEE Transactions on Visualization and Computer Graphics* 23.6 (2017), 1576–1587. doi: 10.1109/TVCG.2017.2674958 3.

[BWK*13] Bernard, Jürgen, Wilhelm, Nils, Krüger, Björn, et al. "MotionExplorer: exploratory search in human motion capture data based on hierarchical aggregation". *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), 2257–2266. issn: 1077-2626. doi: 10.1109/TVCG.2013.178 3–5.

[Chi00] Chi, Ed Huai-hsin. "A taxonomy of visualization techniques using the data state reference model". *Proceedings of the IEEE Symposium on Information Visualization 2000*. INFOVIS. IEEE, 2000, 69–75. doi: 10.1109/INFVIS.2000.885092 3.

[CM97] Card, Stuart K and Mackinlay, Jock. "The structure of the information visualization design space". *Proceedings of the Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*. IEEE, 1997, 92–99. doi: 10.1109/INFVIS.1997.636792 3.

[CRI06] Chittaro, Luca, Ranon, Roberto, and Ieronutti, Lucio. "VU-Flow: A visualization tool for analyzing navigation in virtual environments". *IEEE Transactions on Visualization and Computer Graphics* 12.6 (2006), 1475–1485. issn: 1077-2626. doi: 10.1109/TVCG.2006.109 2–5.

[DC11] Drachen, Anders and Canossa, Alessandro. "Evaluating motion: spatial user behaviour in virtual environments." *International Journal of Arts and Technology* 4.3 (2011), 294–314. doi: 10.1504/IJART.2011.041483 2–5.

[GH10] Goldberg, Joseph H and Helfman, Jonathan I. "Visual scanpath representation". *Proceedings of the Symposium on Eye Tracking Research & Applications*. ETRA. ACM, 2010, 203–210. isbn: 978-1-60558-994-7. doi: 10.1145/1743666.1743717 4, 5.

[GWFI14] Goffin, Pascal, Willett, Wesley, Fekete, Jean-Daniel, and Isenberg, Petra. "Exploring the placement and design of word-scale visualizations". *IEEE Transactions on Visualization and Computer Graphics* 20.12 (2014), 2291–2300. doi: 10.1109/TVCG.2014.2346435 3.

[GWY*11] Guo, Hanqi, Wang, Zuchao, Yu, Bowen, et al. "TripVista: Triple Perspective Visual Trajectory Analytics and its application on microscopic traffic data at a road intersection". *Proceedings of the 2011 IEEE Pacific Visualization Symposium*. PacificVis. IEEE, 2011, 163–170. doi: 10.1109/PACIFICVIS.2011.5742386 2, 4, 5.

[JE12] Javed, Waqas and Elmqvist, Niklas. "Exploring the design space of composite visualization". *Proceeding of the 2012 IEEE Pacific Visualization Symposium*. PacificVis. IEEE, 2012, 1–8. doi: 10.1109/PacificVis.2012.6183556 3.

[KHW14] Kurzhals, Kuno, Heimerl, Florian, and Weiskopf, Daniel. "ISeeCube: Visual analysis of gaze data for video". *Proceedings of the Symposium on Eye Tracking Research & Applications*. ETRA. ACM, 2014, 43–50. doi: 10.1145/2578153.2578158 2, 4, 5.

[KKC14] Kerracher, Natalie, Kennedy, Jessie, and Chalmers, Kevin. "The Design Space of Temporal Graph Visualisation". *EuroVis - Short Papers*. Eurographics Assocation, 2014, 7–11. doi: 10.2312/eurovisshort.20141149 3.

[Kra03] Kraak, M.J. "The space-time cube revisited from a geovisualization perspective". English. *Proceedings of the 21st International Cartographic Conference*. New Zealand: International Cartographic Association, 2003, 1988–1996 2, 4, 5.

[KW13] Kurzhals, Kuno and Weiskopf, Daniel. "Space-time visual analytics of eye-tracking data for dynamic stimuli". *IEEE Transactions on Visualization and Computer Graphics* 19.12 (2013), 2129–2138. issn: 1077-2626. doi: 10.1109/TVCG.2013.194 2, 4, 5.

[Kwa00] Kwan, Mei-Po. "Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set". *Transportation Research Part C: Emerging Technologies* 8.1 (2000), 185–203. issn: 0968-090X. doi: 10.1016/S0968-090X(00)00017-6 2, 4, 5.

[LSF*15] Löwe, Thomas, Stengel, Michael, Förster, Emmy-Charlotte, et al. "Visualization and analysis of head movement and gaze data for immersive video in head-mounted displays". *Proceedings of the Workshop on Eye Tracking and Visualization*. ETVIS. 2015 2, 4, 5.

[MIC09] Munaga, Hazarath, Ieronutti, Lucio, and Chittaro, Luca. "CAST: a novel trajectory clustering and visualization tool for spatio-temporal data". *Proceedings of the 1st International Conference on Intelligent Human Computer Interaction*. Springer, 2009, 169–175. doi: 10.1007/978-81-8489-203-1_15 2–4.

[MKNW19] Merino, Leonel, Kozlova, Ekaterina, Nierstrasz, Oscar, and Weiskopf, Daniel. "VISON: An Ontology-Based Approach for Software Visualization Tool Discoverability". *Proceedings of the 2019 Working Conference on Software Visualization*. VISSOFT. IEEE, 2019, 45–55. DOI: `10.1109/VISSOFT.2019.00014` 3.

[MML15] Minelli, Roberto, Mocci, Andrea, and Lanza, Michele. "I know what you did last summer: an investigation of how developers spend their time". *Proceedings of the 23rd IEEE International Conference on Program Comprehension*. ICPC. IEEE, 2015, 25–35. DOI: `10.1109/ICPC.2015.12` 2, 4.

[MMLB14] Minelli, Roberto, Mocci, Andrea, Lanza, Mario, and Baracchi, Lorenzo. "Visualizing Developer Interactions". *In Proceedings of the Second IEEE Working Conference on Software Visualization*. VISSOFT. IEEE, 2014, 147–156. DOI: `10.1109/VISSOFT.2014.31` 2, 4, 8.

[Mun] Munroe, Randall. *XKCD #657: Movie Narrative Charts*. `http://xkcd.com/657`. Accessed: 2019-02-15 3–5.

[NSW*20] Nebeling, Michael, Speicher, Maximilian, Wang, Xizi, et al. "MRAT: The Mixed Reality Analytics Toolkit". *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. ACM, 2020, 1–12. ISBN: 9781450367080. DOI: `10.1145/3313831.3376330` 3–5.

[OM10] Ogawa, Michael and Ma, Kwan-Liu. "Software evolution storylines". *Proceedings of the 5th International Symposium on Software Visualization*. SoftVis. ACM, 2010, 35–42. DOI: `10.1145/1879211.1879219` 3.

[Shn96] Shneiderman, Ben. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations". *Proceedings of the 1996 IEEE Symposium on Visual Languages*. VL. IEEE Computer Society, 1996, 336–343. ISBN: 0-8186-7508-X. DOI: `10.1109/VL.1996.545307` 3.

[SHY*18] Speicher, Maximilian, Hall, Brian D., Yu, Ao, et al. "XD-AR: Challenges and Opportunities in Cross-Device Augmented Reality Application Development". *Proceedings of the ACM on Human-Computer Interaction* 2.EICS (2018). DOI: `10.1145/3229089` 5, 6.

[SND10] Stellmach, Sophie, Nacke, Lennart, and Dachselt, Raimund. "Advanced Gaze Visualizations for Three-dimensional Virtual Environments". *Proceedings of the Symposium on Eye Tracking Research & Applications*. ETRA. Austin, Texas: ACM, 2010, 109–112. ISBN: 978-1-60558-994-7. DOI: `10.1145/1743666.1743693` 2, 4, 5.

[THM15] Tanahashi, Yuzuru, Hsueh, Chien-Hsin, and Ma, Kwan-Liu. "An efficient framework for generating storyline visualizations from streaming data". *IEEE Transactions on Visualization and Computer Graphics* 6 (2015), 730–742. DOI: `10.1109/TVCG.2015.2392771` 4.

[TM10] Tomaszewski, Brian and MacEachren, Alan M. "Geo-historical context support for information foraging and sensemaking: Conceptual model, implementation, and assessment". *Proceedings of IEEE Symposium on Visual Analytics Science and Technology*. VAST. IEEE, 2010, 139–146. DOI: `10.1109/VAST.2010.5652895` 2, 4.

[TM12] Tanahashi, Yuzuru and Ma, Kwan-Liu. "Design considerations for optimizing storyline visualizations". *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), 2679–2688. ISSN: 1077-2626. DOI: `10.1109/TVCG.2012.212` 4.

[TRL*19] Tang, Tan, Rubab, Sadia, Lai, Jiewen, et al. "iStoryline: effective Convergence to Hand-drawn Storylines". *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019), 769–778. ISSN: 1077-2626. DOI: `10.1109/TVCG.2018.2864899` 4, 5.

[VBP*19] Valdivia, Paola, Buono, Paolo, Plaisant, Catherine, et al. "Analyzing Dynamic Hypergraphs with Parallel Aggregated Ordered Hypergraph Visualization". *IEEE Transactions on Visualization and Computer Graphics* (2019), 1–1. DOI: `10.1109/TVCG.2019.2933196` 7.

[WAPW06] Ware, Colin, Arsenault, Roland, Plumlee, Matthew, and Wiley, David. "Visualizing the underwater behavior of humpback whales". *IEEE Computer Graphics and Applications* 26.4 (July 2006), 14–18. ISSN: 0272-1716. DOI: `10.1109/MCG.2006.93` 2, 4, 5.

[WBG06] Winterbottom, Cara, Blake, Edwin, and Gain, James. "Using Visualizations to Support Design and Debugging in Virtual Reality". *In Proceedings of Advances in Visual Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, 465–474. ISBN: 978-3-540-48631-2. DOI: `10.1007/11919476_47` 4, 5.

[WE12] Wörner, Michael and Ertl, Thomas. "Visual Analysis of Public Transport Vehicle Movement". *Proceedings of International Workshop on Visual Analytics*. EuroVA. The Eurographics Association, 2012. ISBN: 978-3-905673-89-0. DOI: `10.2312/PE.EuroVAST.EuroVA12.079-083` 2.

[YMK13] Yoon, YoungSeok, Myers, Brad A, and Koo, Sebon. "Visualization of fine-grained code change history". *Proceedings of the IEEE Symposium on Visual Languages and Human Centric Computing*. VLHCC. IEEE, 2013, 119–126. DOI: `10.1109/VLHCC.2013.6645254` 2, 4, 5.

[YYA*10] Yu, Zhiwen, Yu, Zhiyong, Aoyama, Hideki, et al. "Capture, recognition, and visualization of human semantic interactions in meetings". *Proceedings of the IEEE International Conference on Pervasive Computing and Communications*. IEEE, 2010, 107–115. DOI: `10.1109/PERCOM.2010.5466987` 2, 4, 5.

[ZFH08] Zhao, Jinfeng, Forer, Pip, and Harvey, Andrew S. "Activities, ringmaps and geovisualization of large human movement fields". *Information Visualization* 7.3-4 (2008), 198–209. DOI: `10.1057/PALGRAVE.IVS.9500184` 2, 4, 5.