Rabsahl, Solveig; Satzger, Thomas; Kalamkar, Snehanjali; Grubert, Jens; Beck, Fabian

Symbolic Event Visualization for Analyzing User Input and Behavior of Augmented Reality Sessions

Date of secondary publication: 25.01.2024

Version of Record (Published Version), Conferenceobject

Persistent identifier: urn:nbn:de:bvb:473-irb-930589

Primary publication

Rabsahl, Solveig; Satzger, Thomas; Kalamkar, Snehanjali; Grubert, Jens; Beck, Fabian (2023): „Symbolic Event Visualization for Analyzing User Input and Behavior of Augmented Reality Sessions“. In: EuroVis Posters 2023, The Eurographics Association, doi: 10.2312/evp.20231070.

Symbolic Event Visualization for Analyzing User Input and Behavior of Augmented Reality Sessions

Solveig Rabsahl[1*], Thomas Satzger[1*], Snehanjali Kalamkar[2], Jens Grubert[2], and Fabian Beck[1]

[1]University of Bamberg, Germany
[2]Coburg University, Germany
*(shared first authorship)

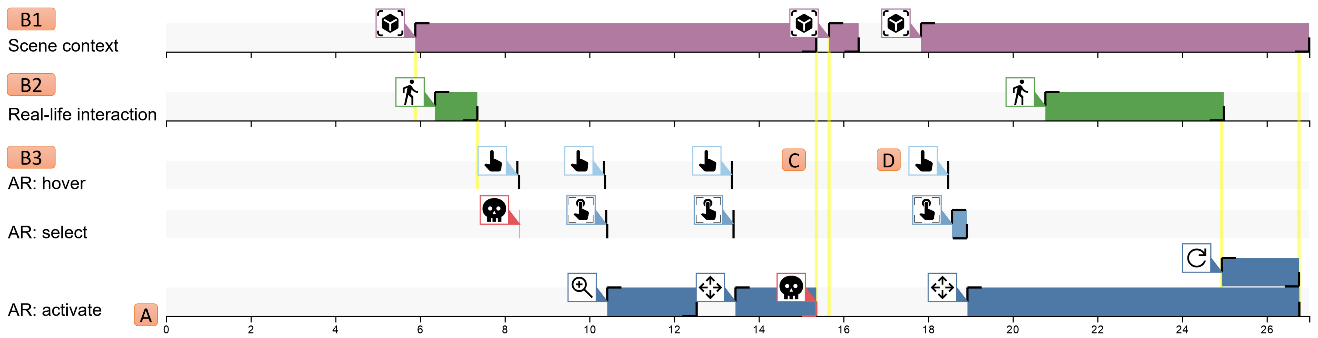

Figure 1: *Layered event visualization on a timeline (A) discerning different layers of event types (B1–B3), with cross-layer connections for similar start and end times of events (C) and icons as symbolic event representations (D); the example shows a user moving a virtual cube.*

Abstract

Interacting with augmented reality (AR) systems involves different domains and is more complex than interacting with traditional user interfaces. To analyze AR interactions, we suggest an event visualization approach that discerns different event layers on a timeline. It is based on symbolic event representations of typical user actions, such as physical movement or interaction with scene objects. Although focusing on the Microsoft HoloLens 2, the approach can generalize to similar environments and provide a basis for developing a more comprehensive visual analytics and annotation solution for AR usage sessions.

CCS Concepts

• ***Human-centered computing*** → ***Information visualization;*** • ***Computing methodologies*** → ***Mixed / augmented reality;***

1. Introduction

Understanding the behavior of users in augmented-reality (AR) environments is challenging—spatio-temporal data, interaction logs, user intends, and scene context must be analyzed together to provide a clear picture. Visualizations can help both in debugging and empirical evaluation scenarios to support such understanding [AASB20]. Some works have already explored the *in-situ* visualization of user behavior [BLD21, HWF*22, RBD*22, KSS*20], that is, visualizing movements and interactions of users within the spatial environment itself. In contrast, abstracting from concrete spatial positions, others have contributed *ex-situ* analysis approaches [HWF*22, AASB20, NSW*20]. Whereas timeline views are common in such approaches to show events, they typically focus on lower-level technical events. However, to provide a better overview of usage sessions, we are interested in abstracting the events to higher-level types of events that reflect typical forms of user input in AR sessions. For visual encoding, we can build on various solutions for event visualization [GGJ*22], some of them also using symbolic representations for event types [LCP*12].

Focusing on the basic input gestures provided by Microsoft's *Mixed Reality Toolkit* (MRTK), we map typical actions that users perform and related events to symbolic representations that are reusable across application examples. Analyzing hand gestures complemented by speech input and gaze, we concentrate on natu-

DOI: 10.2312/evp.20231070

ral and direct ways of interaction, not using controllers. We record sessions of the *Microsoft HoloLens 2* and manually annotate the relevant higher-level events. We then visualize the annotated events in a compact timeline representation using the symbolic event representations (Figure 1). Regarding the design space for such usage session visualization described by Agarwal et al. [AASB20], the approach can be classified as a combination of *event identifiers* and *event timeline* targeted at an *evaluation* scenario (i.e., understanding user behavior). We consider the proposed visualization as a first step towards designing a comprehensive visual analytics and annotation system for AR user sessions, and discuss open challenges as well as ideas for extension.

The underlying idea of augmented reality is the real-time embedding of interactive, spatially registered 3D content in the real environment [Azu97]. Events in such environments can be technical (e.g., direct user input) or contextual (e.g., change in the scene)—we want to capture both at a comparable level of abstraction. Specifically, we discern *AR interactions* (user interactions with virtual objects), *real-life actions* (physical actions of the user), and *scene context* (other events not directly caused by the user) as the most relevant event layers. In addition, *AR interactions* have specific *activators* (hands, gaze, air tap, voice) assigned to them. We rely on the state system used by Unity's *Extended Reality Toolkit* (XRI) and consider the following *AR interactions*: *AR hover*—typically triggers visual highlighting of an object indicating interaction options; *AR select*—an object is persistently selected; *AR activate*—an activation option of an object resulting in a specific action (i.e., moving an object). After recording usage sessions using the *Windows Device Portal* and a *HoloLens 2* device, we manually annotate these events including event type and layer, start and end timestamp, and (if available) activators.

2. Symbolic Event Visualization

Icons provide a simple way to condense even complex interactions into self-explanatory, compact representations. To create a reusable icon set for the above events, we adopted and extended icons from Google's *Material Icon Library* to represent individual events. AR-layer icons either reflect the state change of an interaction object (*activate*) or the input method used to attain said change (*hover*, *select*). We opted to use icons specifying state change and make the activators only available on demand as tooltips. Generally, we aimed to create a uniform icon design for the AR-layers consisting of a base icon representing an input method modified by layer-specific changes, such as frames around select icons. By contrast, we sought to create a clean separation between icons used in different event layers to avoid confusion when working with similar events in different event layers (e.g., *user gives a HoloLens voice command* vs. *user talking to an instructor*). A complete documentation of the icons can be found in the supplementary materials.

As shown in Figure 1, these icons can be used to represent the events on a layered timeline—we implemented an early prototype with *D3.js*. It provides a quick overview of an interaction sequence, allowing one to infer basic statements such as: *The user could not scale the cube because the cube left the field of view*. Event durations are visualized using horizontal bars on a shared timeline (Figure 1, A). Vertically, these are arranged according to the event layers they relate to (Figure 1, B1–B3); concurrent events of the same event layer are stacked vertically within the respective row. For readability of event type and duration, event icons are enclosed in a pin-like shape (Figure 1, D) and event start and end are marked by enclosing brackets (for short events, these collapse into a line). To highlight relationships between different types of events, layers are connected with yellow lines if contained events of different layers start or end within half a second (Figure 1, C). Finally, the visual event representations are colored according to the corresponding event type: ***AR hover***, ***AR select***, ***AR activate***, ***real-life interaction***, and ***scene context*** (colors selected from *Tableau 10*).

The scenario shown in Figure 1 visualizes events from a tutorial session where a user is asked to move a virtual cube to a target destination in the real world. Events start around the 6-second mark, where the user spotted the relevant object—the cube enters the field of view (B1). The user starts walking within 0.5 seconds after this event, suggesting they approach the cube after spotting it (B2). In AR layers (B3), we observe that the user was not immediately successful and made four attempts to move the object, indicated by the pattern of two short events in quick succession in layers *AR hover* and *AR select*. The first attempt, just after the 8-second mark, failed because the user was unsuccessful in selecting the object. The three following attempts are correctly selecting the object and are followed by longer activation events (layer *AR activate*). First, the user scales the object unnecessarily, and only on the second try successfully moves it. However, around second 15, moving is aborted. A reason might be visually losing track of the cube—it is briefly absent from the field of view (C). Finally, from second 18, the user retries and solves the task without interruption—selecting and moving the object while walking to the correct physical location; then, concluding with an optional object rotation.

3. Conclusion and Future Work

We have presented an *ex-situ* visualization approach for AR usage sessions, featuring an icon-based symbolic representation of events across different layers. Since the division into layers translates into any AR scenario, our approach offers opportunities to generalize towards different application contexts. Further splitting of the event layers, such as dividing the scene context into task-related and non-task-related events, might become useful. Departing from *HoloLens* input methods, the integration of alternative input methods like controllers is easily possible.

Using our prototype, simple statements, as well as hypotheses about the session flow, can already be inferred. However, we envision it being embedded into a more comprehensive analysis tool, for instance, integrating *scene views* and *trajectory views* to provide spatial context and *entity timelines* to complement the event timelines [AASB20]. With the focus on abstracting events, it might specifically be the basis to build a visual annotation approach for AR sessions. User-defined icons and functionality to mark event sub-sequences would allow researchers to perform in-depth qualitative analyses of user behavior. Finally, the approach can be evaluated by revisiting or replicating AR-focused user studies collecting feedback of analysts, while at the same time, complementing previous quantitative findings by qualitative insights.

References

[AASB20] AGARWAL S., AUDA J., SCHNEEGASS S., BECK F.: A design and application space for visualizing user sessions of virtual and mixed reality environments. In *Vision, Modeling, and Visualization* (2020). doi:10.2312/vmv.20201194. 1, 2

[Azu97] AZUMA R. T.: A survey of augmented reality. *Presence: Teleoperators & Virtual Environments 6*, 4 (1997). doi:10.1162/pres.1997.6.4.355. 2

[BLD21] BÜSCHEL W., LEHMANN A., DACHSELT R.: MIRIA: a mixed reality toolkit for the in-situ visualization and analysis of spatio-temporal interaction data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021). doi:10.1145/3411764.3445651. 1

[GGJ*22] GUO Y., GUO S., JIN Z., KAUL S., GOTZ D., CAO N.: Survey on visual analysis of event sequence data. *IEEE Transactions on Visualization and Computer Graphics 28*, 12 (2022). doi:10.1109/TVCG.2021.3100413. 1

[HWF*22] HUBENSCHMID S., WIELAND J., FINK D. I., BATCH A., ZAGERMANN J., ELMQVIST N., REITERER H.: ReLive: Bridging in-situ and ex-situ visual analytics for analyzing mixed reality user studies. In *CHI Conference on Human Factors in Computing Systems* (2022). doi:10.1145/3491102.3517550. 1

[KSS*20] KLOIBER S., SETTGAST V., SCHINKO C., WEINZERL M., FRITZ J., SCHRECK T., PREINER R.: Immersive analysis of user motion in VR applications. *The Visual Computer 36*, 10 (2020). doi:10.1007/s00371-020-01942-1. 1

[LCP*12] LEGG P. A., CHUNG D. H. S., PARRY M. L., JONES M. W., LONG R., GRIFFITHS I. W., CHEN M.: MatchPad: Interactive glyph-based visualization for real-time sports performance analysis. *Computer Graphics Forum 31*, 3pt4 (2012). doi:10.1111/j.1467-8659.2012.03118.x. 1

[NSW*20] NEBELING M., SPEICHER M., WANG X., RAJARAM S., HALL B. D., XIE Z., RAISTRICK A. R. E., AEBERSOLD M., HAPP E. G., WANG J., SUN Y., ZHANG L., RAMSIER L. E., KULKARNI R.: MRAT: the Mixed Reality Analytics Toolkit. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020). doi:10.1145/3313831.3376330. 1

[RBD*22] REIPSCHLÄGER P., BRUDY F., DACHSELT R., MATEJKA J., FITZMAURICE G., ANDERSON F.: AvatAR: an immersive analysis environment for human motion data combining interactive 3D avatars and trajectories. In *CHI Conference on Human Factors in Computing Systems* (2022). doi:10.1145/3491102.3517676. 1