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

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

VISUS University of Stuttgart, Germany

Eurographics Conference on Visualization (EuroVis) 2023

Information Visualization
6th November 2023
TP3 Class

Mariana Andrade	103823	MEI
Vicente Barros	97787	MEI

deti universidade de aveiro
departamento de electrónica, telecomunicações e informática

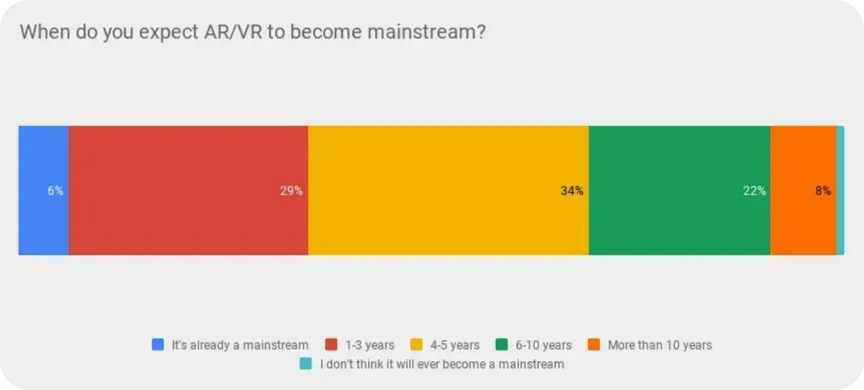
Why This Paper?

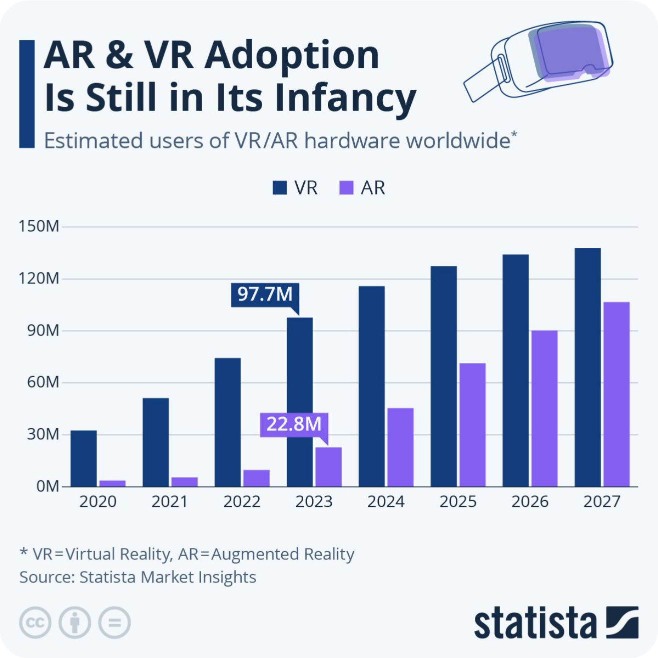
The main reasons we chose this paper are

- The novel framework that integrates different types of data (HMD, eye tracking, and model data) into a unified model for comparability of multiple participants
- New Visualization Techniques:
    - Gaze Replay
    - Space-Time cube
    - Immersive View
- Comprehensive evaluation of the proposed visualization techniques through a think-aloud experiment and an expert interview.

Contextualization

The Virtual Reality (VR) and Augmented Reality (AR) brought new ways of developing interfaces for interaction and visualization.





The Big Question

"How can we evaluate such scenarios?"

The Big Question

Problems

- Measure performance for specific tasks provides little information.
- The user’s feedback generally comes from interviews and standardized questionnaires.
- Lack of spatial context in which they were acquired.

Related Work

Eye tracking in Research

- Gaze-based interaction
- Evaluation of user behavior

New techniques have been developed, especially with focus on VR/AR.

Issues with Gaze Data recorded with Head Mounted Displays

- Existing analysis techniques cannot be applied directly
  - **Mappings are necessary** to achieve comparability between recordings.
  - The visualization of the data in **an explorable overview is challenging.**

Related Work

Issues with Gaze Data recorded with HMDs

The proposed solution addresses these issues with **a semi-automatic spatial mapping with a linked-view visualization**, making:

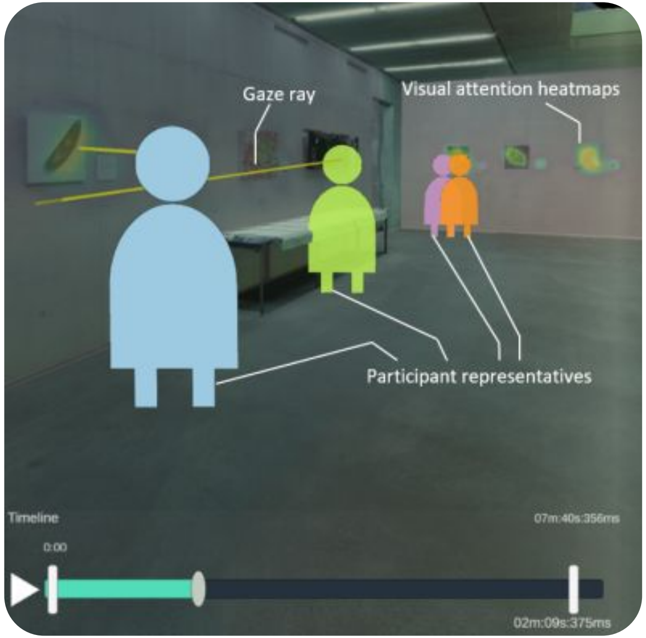
- Participants comparable
- Participants’ movement and viewing behavior explorable

The Proposed Solution

- Provide a new approach to the with the interpretation of the movement and behaviour of the gaze.

- Mapping the information in an unified virtual model.

- The final Software was implemented in Unity.



Framework Proposal - Design

There isn't a standard for recording and visualizing movement and gaze in real-world (and AR) without relying on extensive video-based annotations.

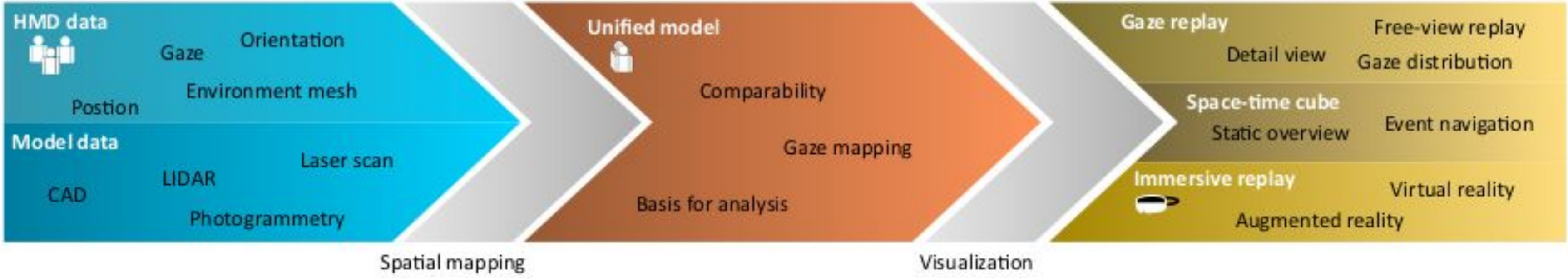
Research Questions

How can we:

- Make recordings of movement and gaze in real-world environments **comparable**?
- Include **detailed information** about the surrounding spatial context without annotation?
- Display the data to answer typical **analysis questions** considering when, *where, and what* happened?

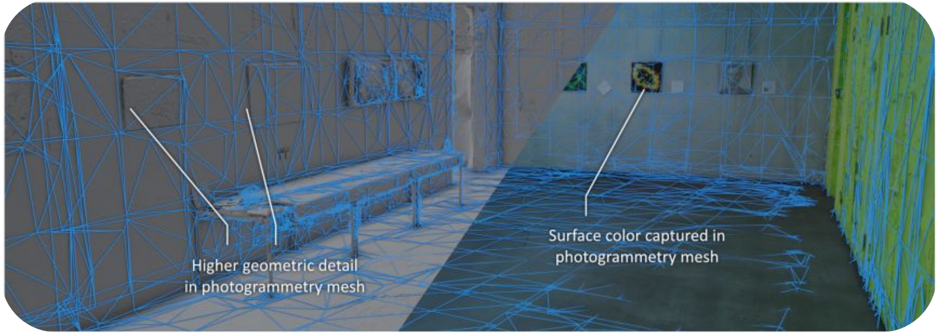
Framework Proposal

Framework for the processing of multiple data sources



Head Mounted Display Data

- HMDs use Stimulation Localization and Mapping (SLAM) techniques to provide information about the position and orientation in an environment.



- A spatial mesh is created that represents the surroundings, typically with a much coarser resolution than the respective surfaces in the real world.

Model Data

- Model Data refers to additional data sources that provide a more detailed representation of the environment than the HMD data - limited quality of the environment mesh

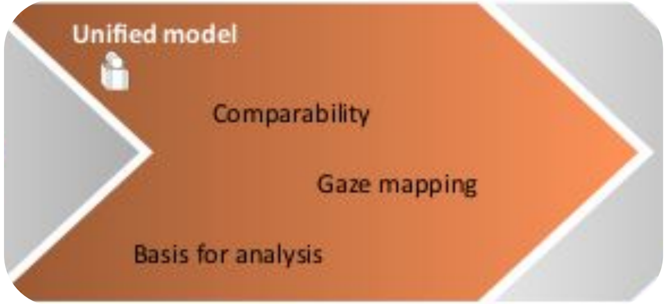
- Can be derived from different methods, such as CAD files, laser scans, LIDAR recordings, or photogrammetry scans.

- Has to be adjusted to match the position, orientation, and scale of the HMD data to create a unified model.

Unified Model

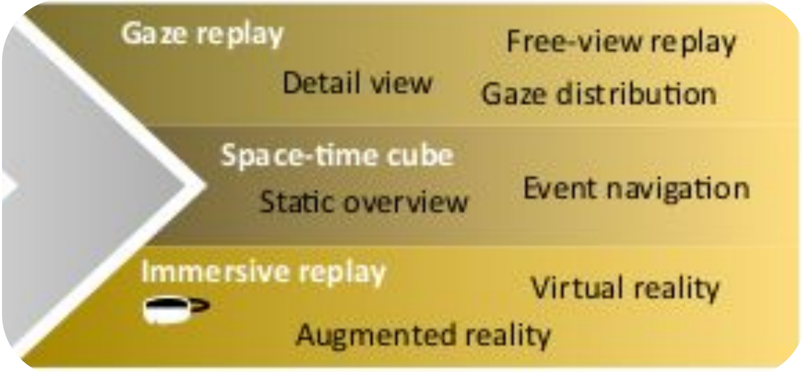
- The unified model is a common coordinate system for HMD and model data that allows comparability of multiple participants.

- The resulting model allows the classical trajectory analysis making possible to identify common locations where people tend to stay or find main walking paths.

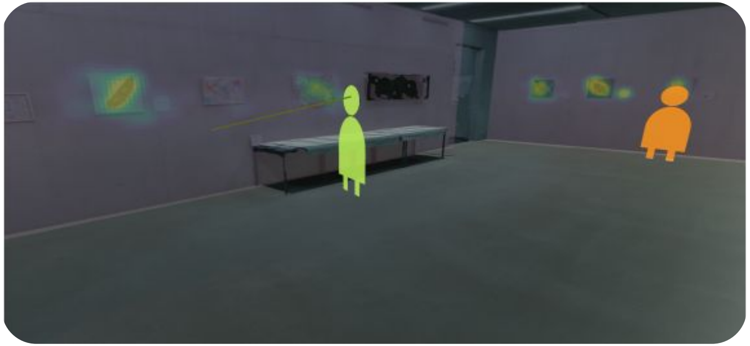


Visualization Techniques

- Three complementary visualizations for the unified model are presented.
- Each one aims to support different analysis tasks (motion patterns and compare gaze distributions).
- The visualizations are linked and interactive, allowing users to filter, select, and explore the data from different perspectives.



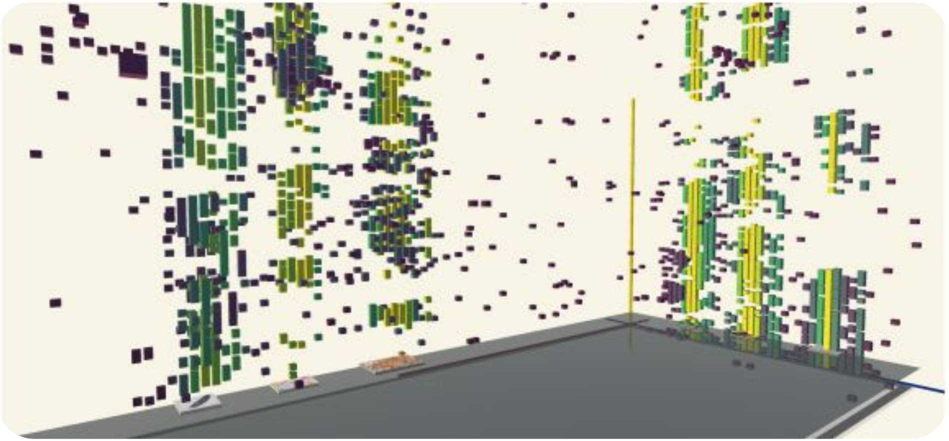
Gaze Replay



- Shows the 3D spatial context of the environment with indicators for the position and point of regard of each participant.
- Allows users to view the data as an animation or to inspect specific time steps with a freely rotatable camera.
- Supports heatmap visualization of the aggregated distribution of visual attention for a selected time span.

Space-Time Cube

- Shows a spatio-temporal overview of the movement and gaze data in a common coordinate system.
- Displays trajectories of movement and gaze cubes of fixations for each participant.
- Enables users to identify interesting events, outliers, and common patterns in the data.

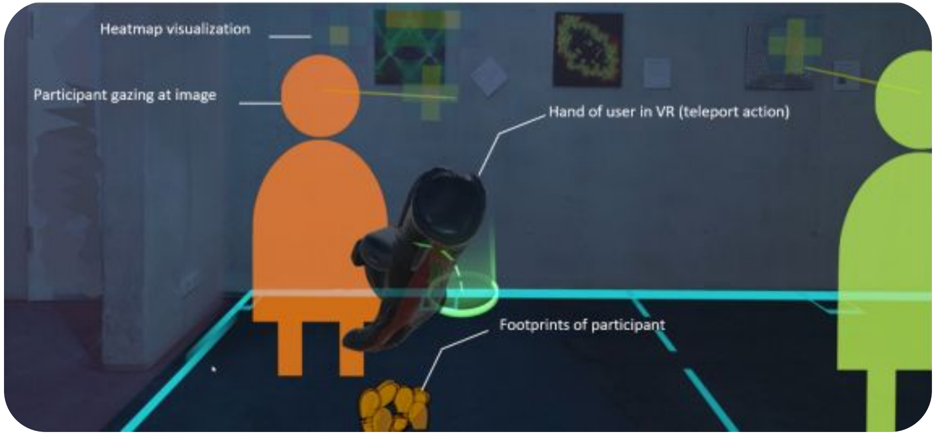


Immersive Replay

- Allows users to investigate the data from the perspective of the participants in VR.

- The immersive replay shows the same visualizations as the gaze replay, but in a first-person view.

- Helps users to understand how the participants perceived the scene and what was visible in their field of view.



Visualization Techniques Summary

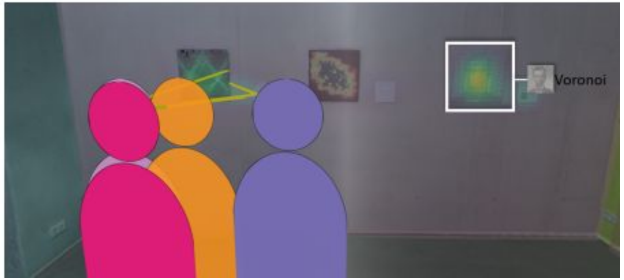
Visualization	Advantages	Limitations	Analysis Tasks
3D Replay 	Visualizations with spatial context Detailed inspection of space	Temporal overview missing Comparison between multiple participants	Where did participants look? Which image received most/least fixations at a specific point in time? Which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

User Experience Evaluation

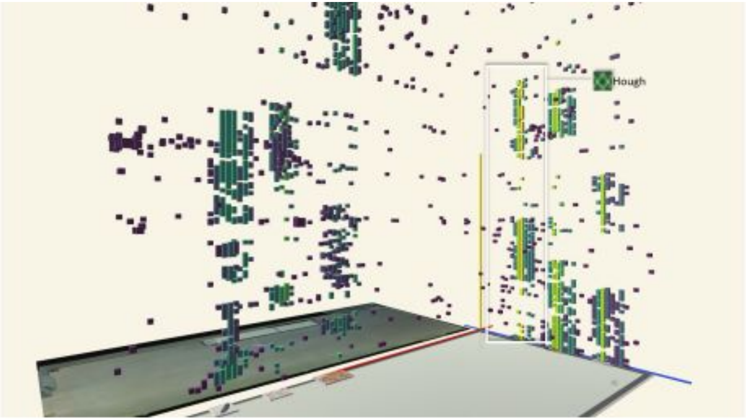
Experiment 1 - Think Aloud

Task #1 - Find Motion Patterns

- Participants used different strategies to accomplish this task:
    - **zooming** in or out,
    - **filtering**
    - **selecting individual participants** in the space-time cube.
- **All participants** were able to detect the data of the two people who were recorded simultaneously
- **Some participants** detected the data of the people who followed the same order of viewing the images



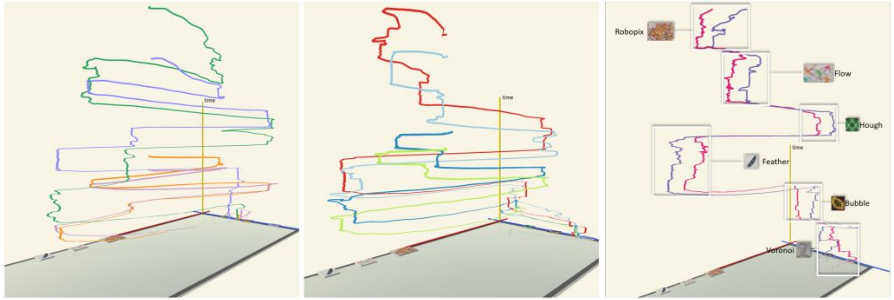
Experiment 1 - Think Aloud



Task #2 - Fixation Distribution

Find the images that received the most or the least fixations within a given time span or over the whole data

- The detection of the **most fixated image** was easier than the **least fixated image**
- Some participants also used the gaze replay to **identify the people** who viewed a specific image within a time span.



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

Experiment 2 - Expert Interview

The goal of this interview was to evaluate the applicability of the solution in matters such as spatial cognition and behaviour research

- It was presented videos demonstrating the functionalities of the framework
- The experts appreciated the ability to integrate different data into a unified model
- **Gaze Replays** and **Gaze Cubes** were the most impressive features
- Some Suggestions were:
    - Include physiological data or eye movement metrics
    - Export the visualized data as a database or file
    - Multiple start and end points within a gaze recording

Discussion - Scalability

The proposed approach could be extended to larger scenarios.

However, **there are limitations**:

Hardware

- Hardware limitations to process large data streams.

Design

- The STC reduces by design the spatial context to 2D.
- Gaze distributions in different stories have to be aggregated or depicted separately.
- Virtual models of outdoor scenarios require much disk space.

Discussion - Environment Aspects

Challenges

- Surroundings **without salient visual features** complicate localization and reconstruction.

Current Solution

- Extension of the unified model by **an additional sensor** (e.g., LIDAR).
- Dynamic AOIs aren’t considered explicitly.

Approach Problems

- Issues with the photogrammetric reconstruction of **smooth and reflecting surfaces**.

Generalized Solution

- **Field-of-view video** incorporating methods used for 2D AOIs for detection and tracking.

Discussion - Interaction

Preferences

- General preference for mouse input for 3D interactions.

This could be influenced by the professional background of the participants.

Discussion - Possible Extensions

Comparison of Asynchronous Trajectories

Automatic techniques such as dynamic time warping could help adjust the visualization.

Depth Perceptions Issues

Analysis questions regarding the STC could further be supported by 2D projections of the data to overcome issues.

Duration of Visits

A heatmap visualization of the trajectory projections could help to investigate the duration of visits from different persons at a specific location.

Conclusions

Future Extensions

- **The inclusion of heatmaps** projected on the floor to include the visualization of aggregated movement data.
- Incorporation of **new sensor modalities** into the unified model for later analysis.

Overall Impact

The visual analysis of HMD-recorded data provides a better understanding of:

- How people perceive and interact with augmented and real environments.

Bibliography

Statista. (n.d.). Virtual and augmented reality adoption forecast. Retrieved November 6, 2023, from https://www.statista.com/chart/28467/virtual-and-augmented-reality-adoption-forecast/

Devabit. (2023). 5 Charts to Understand the Mainstream AR & VR Adoption. Retrieved November 6, 2023, from https://devabit.com/blog/ar-vr-mass-adoption/