Eye-Tracking in VR: From Heatmaps to Object AOs

Pavel Ugwitz 1 Zdeněk Stachoň 1 Čeněk Šašinka 2

1Dept. of Geography, Faculty of Science 2Div. of Inf. & Library Studies, Faculty of Arts Masaryk University, Czech Republic

Scope

VR eye-tracking (ET) methodologies are explored. The intent is to split 3D scenes/objects into trackable 3D areas. To be able to do this, some future methodologies/workflows are considered:

- How to separate 3D scenery into areas of interest (AoI)?
- Processing objects/heatmaps to AoIs is proposed.
- Virtual environment creation suitable for this approach.
- Dynamic/moving participants as extra variability.
- Data acquiring & filtering potential issues.

Exploratory data acquisition & processing in stationary VR scenes was conducted and considered in feasibility. Following implementations of higher complexity are discussed.

Unity 3D engine was utilized. HTC Vive Pro head-mounted display (HMD) and Pupil Labs HMD-eye-tracker were used. Regardless, any similar hardware should be applicable.

Challenges in Methodics

ET in 3D VR scenarios brings about challenges that were either non-existent or not as prominent in traditional ET [1]. In 3D, users can alter their viewports, objects get occluded; a scene can also include dynamic/interactive items/controllers [2] that react to the user or behave autonomously.

2D ET methodologies/analytics are not simply reusable in 3D/VR. Thus, a way to quantify 3D ET data is proposed: a two-pass approach with points/heatmaps, and subsequent object/AoI creation:

The heatmap-based and object-based approaches are not exclusive/antagonistic [3], and can be used to reach a common goal while playing to their strengths. The proposed process is, in essence, data filtering. As per this distinction:

<table>
<thead>
<tr>
<th>Heatmap Based (Fixation)</th>
<th>Implementation</th>
<th>Scene Setup</th>
<th>Data Retention</th>
<th>Data Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heatmap has to be tracked in the scene.</td>
<td>High. Eye fixations or objects retained.</td>
<td>Complex. Fixations transferred to heatmaps; qualitative &amp; quantitative interpretation.</td>
<td>Complex.</td>
<td>Fixations transferred to heatmaps; qualitative &amp; quantitative interpretation.</td>
</tr>
<tr>
<td>Heatmap algorithms and visualization</td>
<td>Heatmap algorithms and visualization</td>
<td>Heatmap algorithms and visualization</td>
<td>Heatmap algorithms and visualization</td>
<td>Heatmap algorithms and visualization</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object Based (Aos)</th>
<th>Implementation</th>
<th>Scene Setup</th>
<th>Data Retention</th>
<th>Data Interpretation</th>
</tr>
</thead>
</table>
| Simple. Fixations script. | Low. Only less complex, con- | Object-based ET is a simplification that records two events: ET ray entering/exiting an AoI. To get to objects, fixation data (or heatmaps) [Fig.2 mid] need to be processed by algorithm(s), producing separate ET AoI (sub)objects [Fig.2 right].

How to create ET AOs from fixations? Convex hull algorithm encapsulates fixations with AoI shells. But is this approach valid when considering semantics, visual cognition, etc.? Real data interpretation may not be as simple as illustrated [Fig.2]. Are 3D heatmaps/AoIs comparable among users/groups? That is a future experimental question.

Data (Pre)filtering and Analysis: A Conclusion

3D environment can be optimized to reduce ET data noise. E.g., by limiting object occlusion, ET distance, object min. size, etc. VR ET introduces inaccuracy/glitches in data. Therefore, data should be collected as fixations and processed to AoIs – this way, errors can be corrected and the process documented.

ET data is stored through experiments, and then re-loaded and re-computed using desired algorithm(s). Since the algorithms were not implemented yet, this is the conclusion: subsequent object-based algorithms on experimental ET data can be run in a post-hoc analysis. Processed objects can be joined to full scenes. This enables for flexibility in the research process.

Collecting Fixation Data, Computing Heatmaps

When tracking, eye gaze vector and HMD head position data are combined to produce a raycast [4], which protrudes into the virtual environment, passing through space until it hits a virtual object. I.e., a user looking from somewhere at something.

As shown on the frog object example [Fig.2 left], fixations in VR are collected per each individual frame – here visualized by green boxes [Fig.1 left]. An algorithm can then compute fixation density, distance, etc. (either as shown [Fig.1 right] or by implementing a heatmap). These are still work in progress.

References


Funding

This work was supported by the project MUNI/A/1576/2018, Complex research of the planet Earth’s geographic environment.