SPHEER – a rich dataset of time-resolved 3D gaze and head movements in virtual reality

Erwan **David**[1]
Melissa L.-H. **Võ**[2]

[1] *Laboratoire d'Informatique de l'Université du Mans*
Le Mans Université
Le Mans, France

[2] *Scene Grammar Lab*
Department of Psychology
Goethe University
Frankfurt, Germany

The Scene Perception, Head and Eye in Extended Reality dataset

Modelling visual attention and gaze requires **large complex datasets** that are very difficult to create. We present a **dataset** *(release date: **Sept. 2024**)* of **3D gaze data** gathered across several naturalistic VR protocols.

To create the **SPHEER** dataset we compiled **3D raw** eye and body tracking data from **11 experiments**! (see bottom refs.).

Some numbers!

380+ participants.
6+ days of continuous time-series data (+50M samples).
3 types of stimuli (360° images and videos, 3D virtual rooms).
3+ types of tasks (free-viewing, visual search, object placement).

Request early access

Send us your feedback

Method

Eye, body and **controller** tracking data from XR devices (HTC Vive, Vive Pro Eye, Pico 4 Pro) = 3D vectors and quaternions.
Body data = Head (+ torso and leg in 3 studies).

Processing

Resampling = 120Hz or 250Hz for uniformly time-separated samples.
Standardisation = common spatial reference and data layout:

Unique experiment and participant ID, Timestamp, Head rotation, Left/Right eye direction, (position/rotation of body parts & controller)

Trial (meta-)data

Raw data are linked to **trial data** which contain info. about stimuli and experimental conditions
3D virtual rooms are shared as collections of **3d bounding boxes data.**

Uses

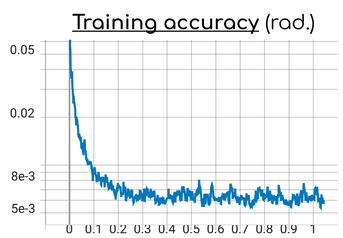
- Visual attention model: saliency / saccadic **modelling.**
- **Predict** head from eye movements and vice-versa.
- Investigate **vergence** behaviour in complex 3D conditions.
- Develop **gaze-parsing** algorithms (3D, dynamic content).
- Test eye movements hypotheses related to **age and gender.**
- Etc.

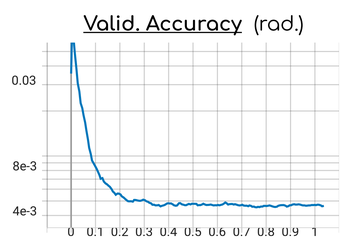
Example:

Predicting head rotations from eye data

Input: head-centred eye direction vectors (3D), time-series of 40 samples 16.7ms apart.
Ouput: Head rotation (quaternion).
Model: simple MLP (1 hidden layer).
Loss: MSE + quaternion angular difference.

Testing accuracy (quaternion diff.):
- Mean = 0.26°
- Median = 0.15°





The SPHEER VR Test suite

There is serious **need** for datasets of raw eye/head data gathered during **typical 3D gaze movements** to create **gaze-parsing** algorithm **adapted to 3D** gaze events:

- **Saccades** & **fixations**
- Vestibulo-ocular reflex (**VOR**)
- **Smooth pursuit**

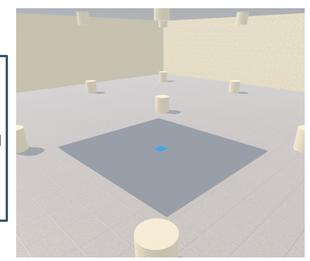
All with varying **cue depth**

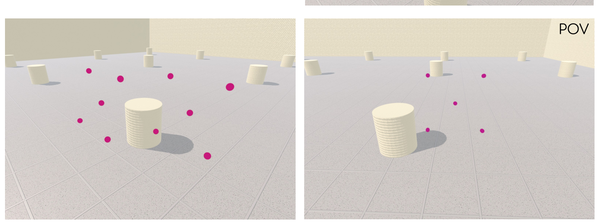
The VR Test suite implements simple, quick, repeated visual tests placing targets between 0.33 and 2m to viewers in the virtual environment to elicit **vergence changes.**

Vergence is important to infer an object of interest in 3D (several objects may be in the line of sight of each eye).

Available On *GitHub*

Tell us what you think!





Custom validation phase (2 depth levels) to ensure good data quality.

Ref.

- Rai, Y., Gutiérrez, J., & Le Callet, P. (2017, June). **A dataset of head and eye movements for 360 degree images.** In Proceedings of the 8th ACM on Multimedia Systems Conference (pp. 205-210).
- David, E., Gutiérrez, J., Coutrot, A., Da Silva, M. P., & Le Callet, P. (2018, June). **A dataset of head and eye movements for 360° videos.** In Proceedings of the 9th ACM Multimedia Systems Conference (pp. 432-437). ACM.
- David, E., Beitner, J., & Võ, M. L.-H. (2020). **Effects of transient loss of vision on head and eye movements during visual search in a virtual environment.** Brain sciences, 10(11), 841
- Helbing, J., Draschkow, D., & Vo, M. L. H. (2020). **Search superiority: Goal-directed attentional allocation creates more reliable incidental identity and location memory than explicit encoding in naturalistic virtual environments.** Cognition, 196, 104147.
- David, E., Lebranchu, P., Perreira Da Silva, M. & Le Callet, P. (2022). **What are the visuo-motor tendencies of omnidirectional scene free-viewing in virtual reality?.** Journal of Vision, 22(4), 12
- + 3 manuscripts **in preparation** + **unpublished data**

Ack.

Part of this work was supported by RFI Atlanstic2020.
Part of this work was supported by SFB/TRR 26 135 project C7 to Melissa L.-H. Võ and the Hessisches Ministerium für Wissenschaft und Kunst (HMWK; project 'The Adaptive Mind').

Contact
erwan.david@univ-lemans.fr