Monitoring user behavior in subjective tests with 360-degree video

Jesús Gutiérrez, Patrick Le Callet

Image, Interaction, Perception Group (IPI)
Laboratoire des Sciences du Numérique de Nantes (LS2N)
Université de Nantes

VQEG Meeting
14/11/2018

Introduction and motivation

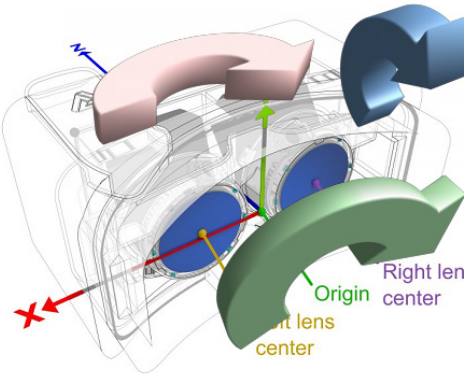
- Watching 360˚ content: The user does not see the whole content. He/she freely decides where to look seeing a region of the scene (viewport).

Introduction and motivation

- Knowing how observers explore 360˚ content is crucial
- **A brief review of the SoA…**
  - What can we measure?
  - How can we measure it?
  - How can we process the data?
  - What can we get from it?

Tracking with HMDs

- What/how can we measure?
  - "Easy" to get **head rotation data** (mainly **latitude and longitude**) thanks to tracking sensors of HMDs:
    - Oculus Rift [Rai *et al.*, MMSys2017; Ozcinar *et al.*, QoMEX2018; S. Fremerey, *et al.*, MMSys2018]
    - HTC Vive [David *et al.*, MMSys2018; Xu *et al.*, CVPR2018; S. Fremerey, *et al.*, MMSys2018.]
    - Open-Source Virtual Reality (OSVR) Application [Corbillon *et al.*, MMSys2017]



Tracking with HMDs

- How can we process the data?
  - Directly take the samples of the HMD (e.g., 90Hz for Vive, 75HZ for Oculus)
  - Subsampling: Less samples needed, filtering noise/small movements [Ozcinar *et al.*, QoMEX2018]
    - Sample camera rotation by selecting the centroid of 200ms windows (100 samples for a 20s stimulus) [David *et al.*, MMSys2018]

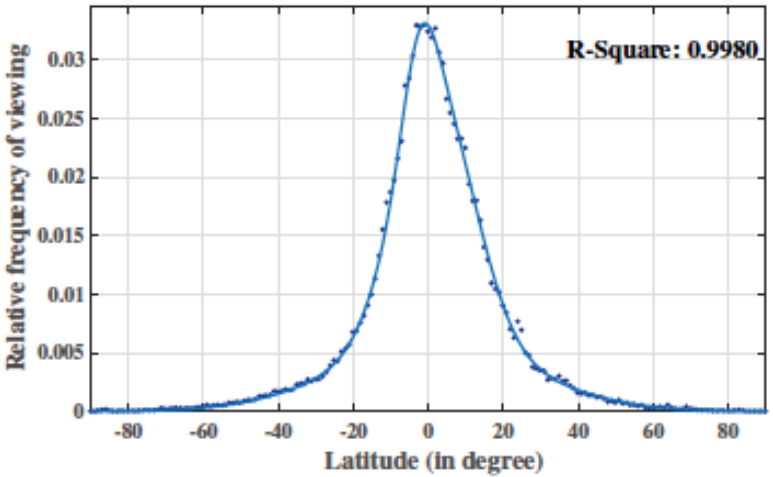
  - Classify head movements as it is done commonly with eye movements:
    - **Fixations:** period of time where the eye is kept aligned with the target for a certain duration.
    - **Saccades:** Saccades are the type of eye movement used to move the fovea rapidly from one point of interest to another → "No perception"

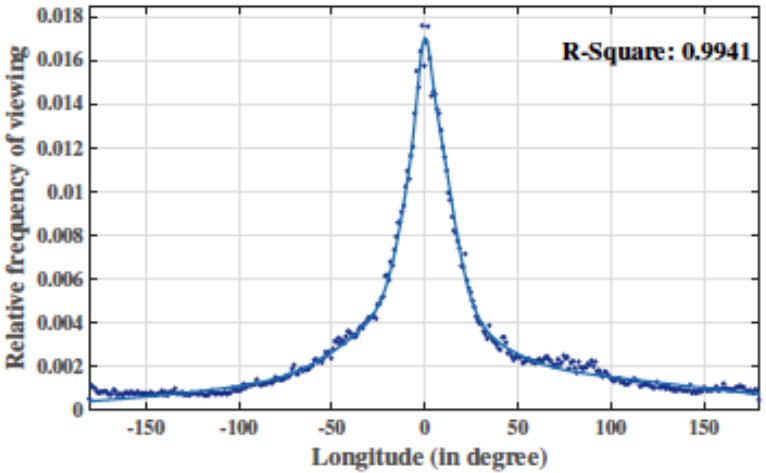
Tracking with HMDs

- How can we process the data?
  - Directly take the samples of the HMD (e.g., 90Hz for Vive, 75Hz for Oculus)
  - Subsampling: Less samples needed, filtering noise/small movements [Ozcinar *et al.*, QoMEX2018]
    - Sample camera rotation by selecting the centroid of 200ms windows (100 samples for a 20s stimulus) [David *et al.*, MMSys2018]
  - Classify head movements as it is done commonly with eye movements:
    - Do "head fixations" make sense?
      - "Head fixations" can be defined as head velocities below ~15-25°/sec [G. Marmitt and Duchowski, *Eurographics 2002;* Hu *et al.*, CISS2017; Upenik *et al.*, ICMEW 2017]
      - But head movements don't mean loss of perception (as in saccades).

Tracking with HMDs

- How can we process the data?

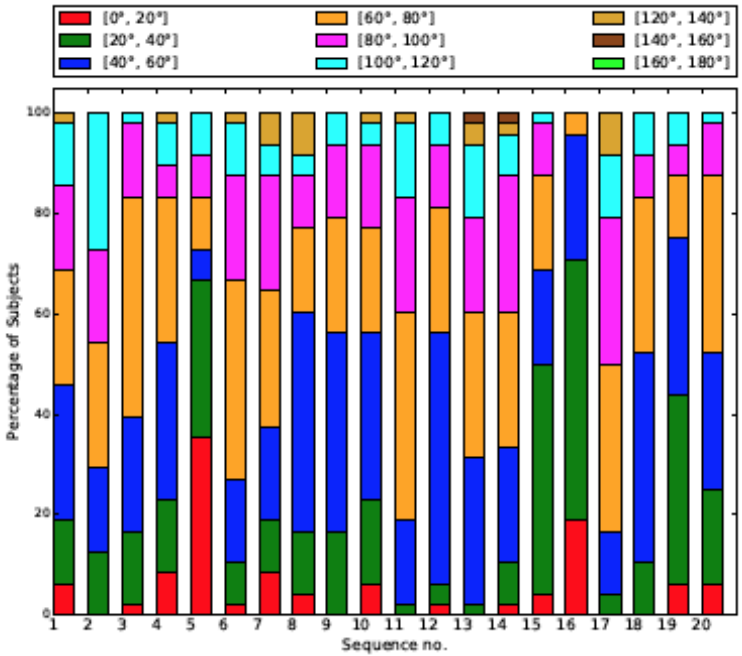


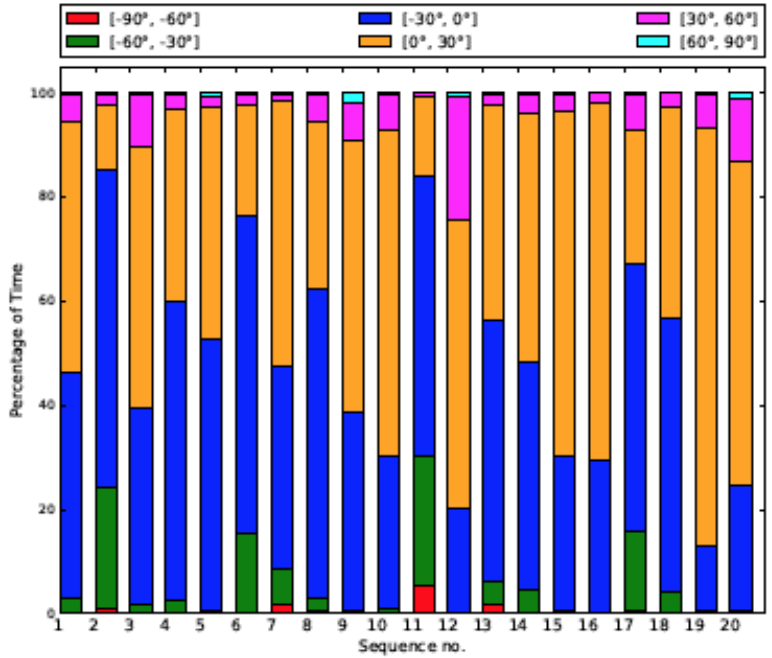


Xu *et al.*, arxiv:1709.06342 2017.

Tracking with HMDs

- How can we process the data?

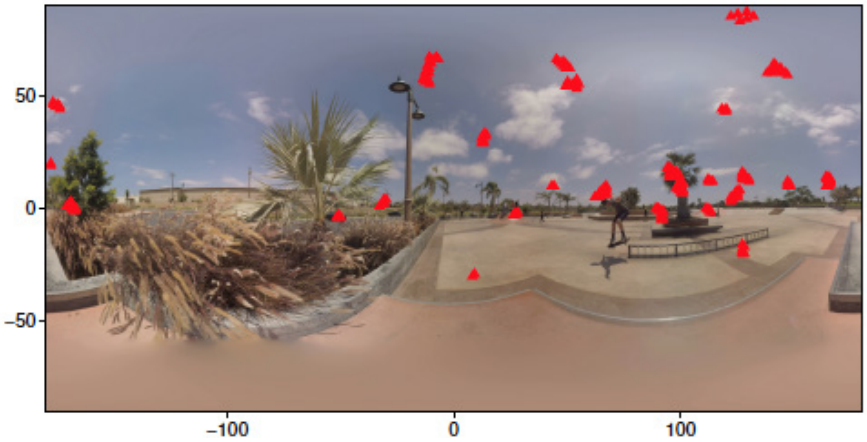




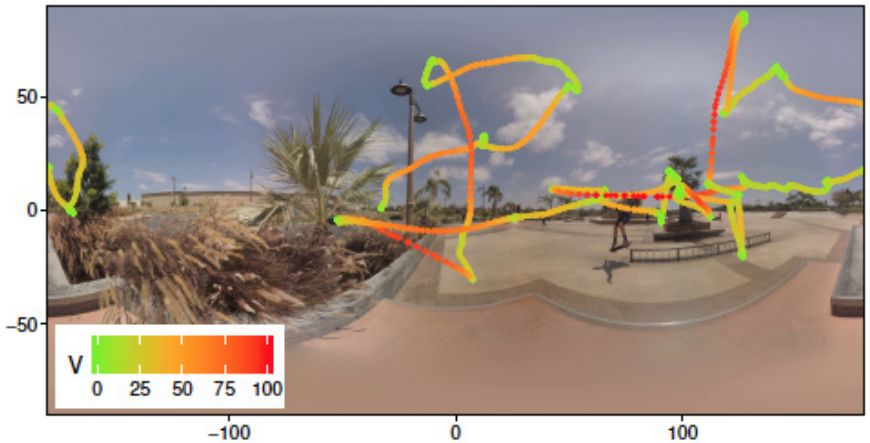
S. Fremerey, *et al.*, MMSys2018.

Tracking with HMDs

- How can we process the data?
  - Scanpaths / head trajectories → Project fixation-points to equirectangular format



Head samples



Head trajectory

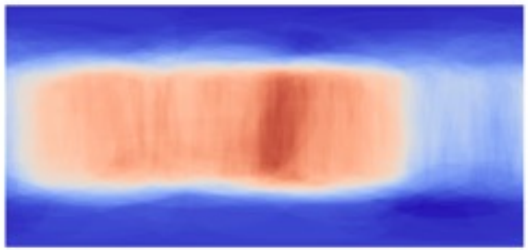
From Upenik *et al.*, ICMEW 2017

Tracking with HMDs

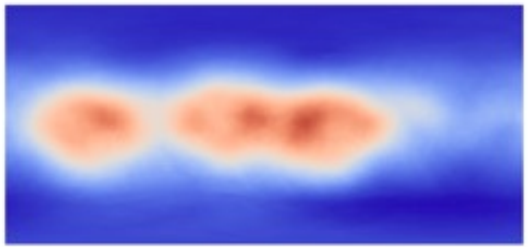
- How can we process the data?
  - Saliency maps → Apply an isotropic Gaussian the samples + project to equire

Original equi-rectangular image

From Rai *et al.*, MMSys2017



Saliency map from **head movement** considering **entire viewport**



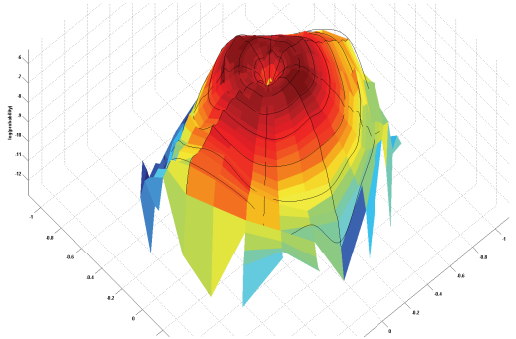
Saliency map from **head movement** + **Gaussian** at the center of viewport

Tracking with HMDs

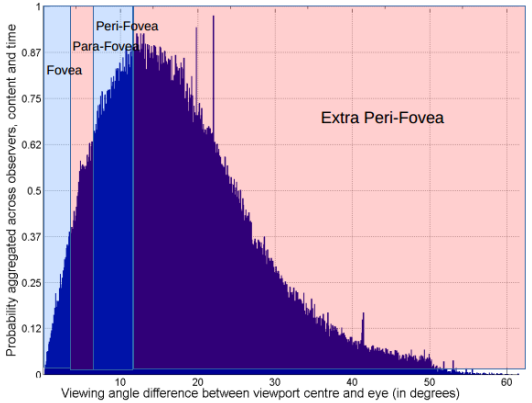
- How can we process the data?
  - Existing SWs and toolboxes:
    - Fremerey, *et al.*, MMSys2018 → Oculus and HRC Vive
      - https://github.com/Telecommunication-Telemedia-Assessment/AVTrack360
    - Corbillon *et al.*, MMSys2017 → Open-Source Virtual Reality (OSVR) Application
      - http://dash.ipv6.enstb.fr/headMovements/
    - Ozcinar *et al.*, QoMEX2018 → Oculus
      - https://github.com/cozcinar/omniAttention

Tracking with HMDs + eye-tracking

- Why eye-tracking?
  - **Head movement** can be a **proxy of visual attention** for some applications:
    - Considering the observers gaze directed to the center of the viewport.
  - But, studies have shown that **observers explore within the viewport** [Sitzmann *et al.*, IEEE TVCG2018; Serrrano *et al.*, ACM TOG 2017; Rai *et al.*, MMSys2017; David *et al.*, MMSys2018.]

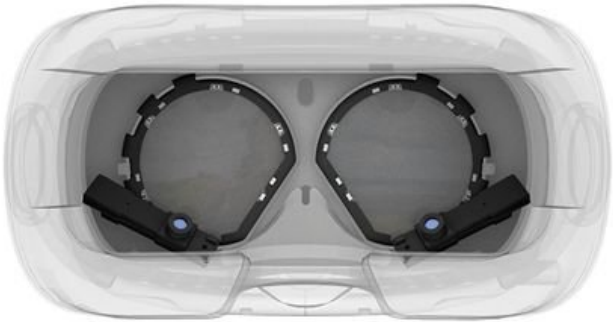


From Rai *et al.*, MMSys2017

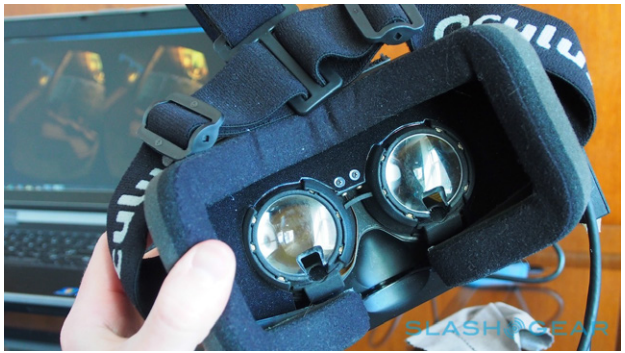


Tracking with HMDs + eye-tracking

- What/how can we measure?
  - Record **binocular eye gaze data** with eye-trackers integrated in the HMDs
    - Pupil Labs [Sitzmann *et al.,* IEEE TVCG2018; Serrano *et al.,* ACM TOG2017]
    - SMI [Rai *et al.*, MMSys2017; David *et al.*, MMSys2018]
    - 7invensun a-Glass [Xu *et al.,* CVPR2018]



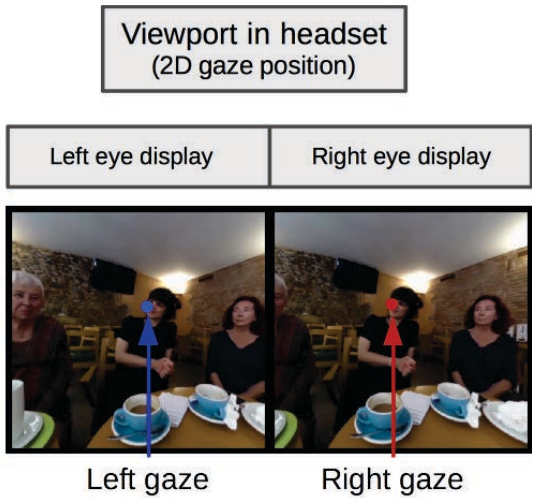
Pupil Labs

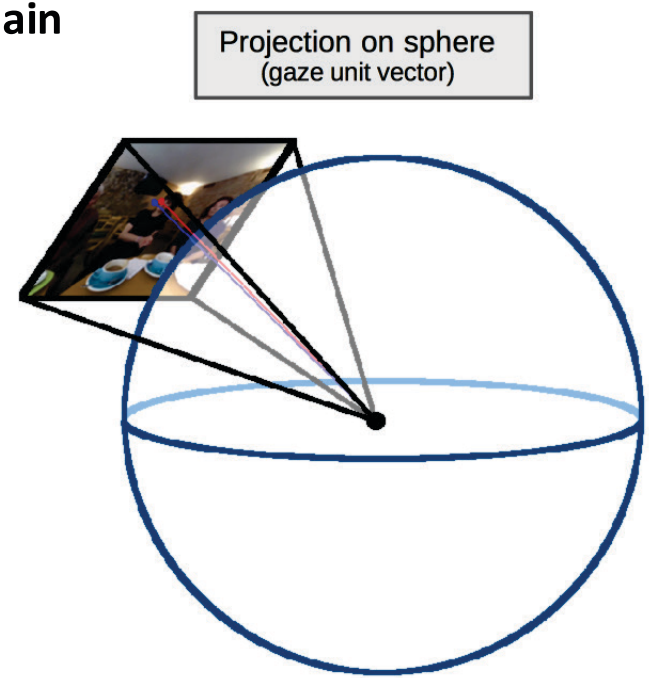


SMI

Tracking with HMDs + eye-tracking

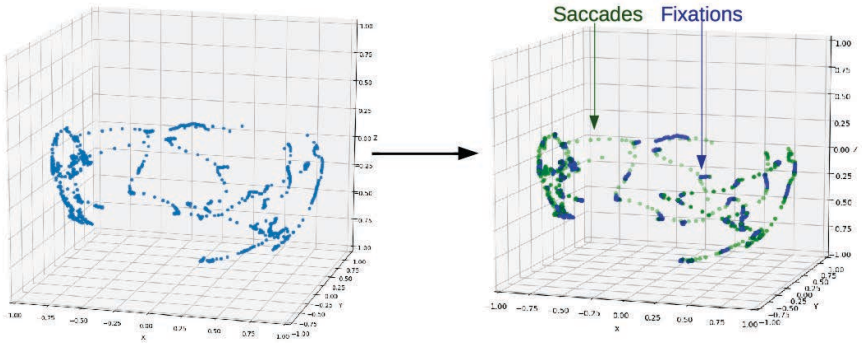
- How can we process the data?
  - **Tracking data** is obtained in the **viewport domain**





Tracking with HMDs + eye-tracking

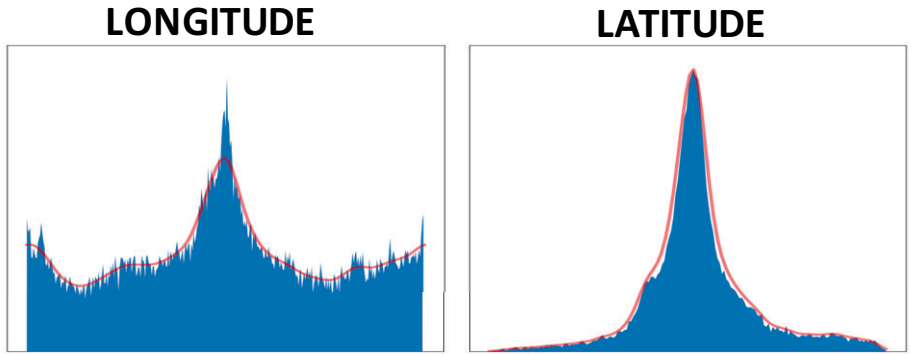
- How can we process the data?
  - Classification into **fixation** and **saccades**:
    - Performed in **spherical domain** → latitude, longitude
      - To maintain mathematical precision
      - Viewports change instantaneously with head's movement
    - Velocity-based algorithm and duration [David *et al.*, MMSys2018], dispersion [Sitzmann *et al.*, IEEE TVCG2018], etc.
    - Euclidean distance replaced by orthodromic distances



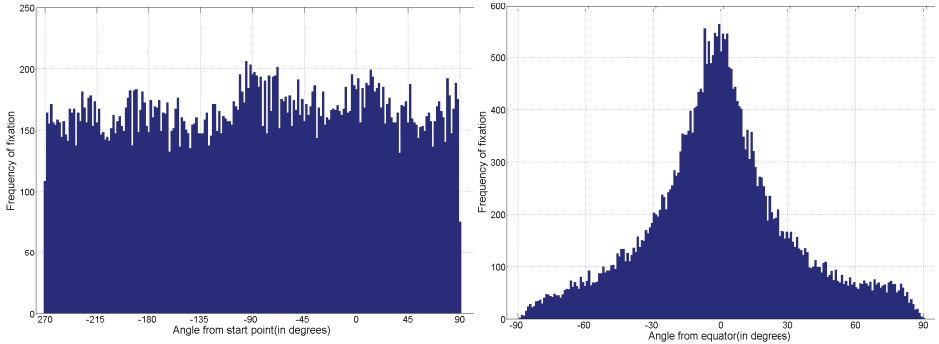
Tracking with HMDs + eye-tracking

- How can we process the data?

VIDEOS:
David *et al.*,
MMSYS2018

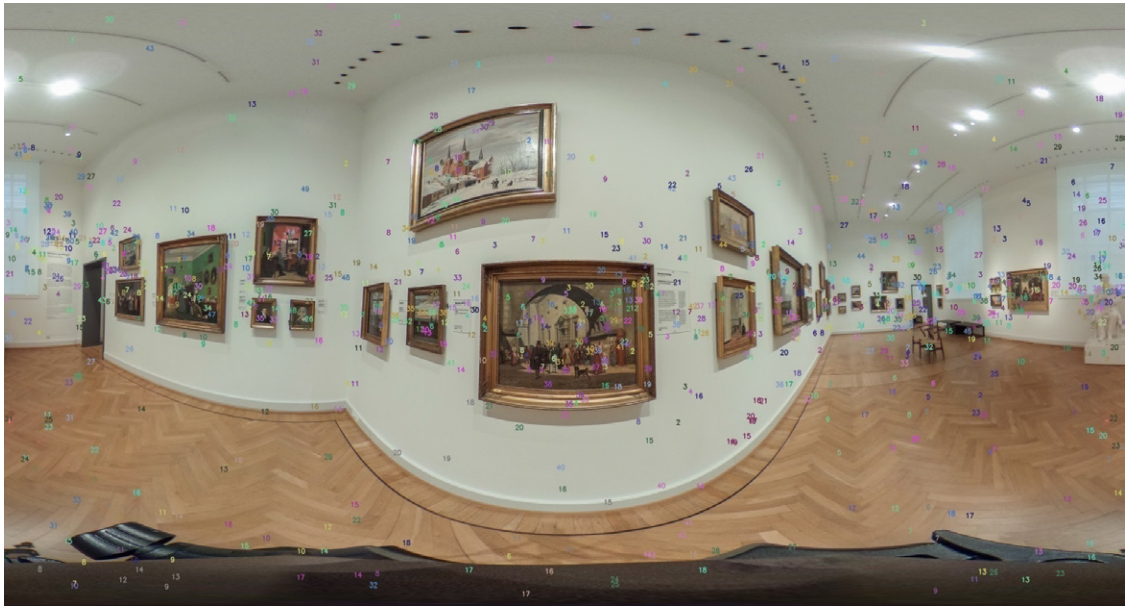


IMAGES:
Rai *et al.*,
MMSYS2017



Tracking with HMDs + eye-tracking

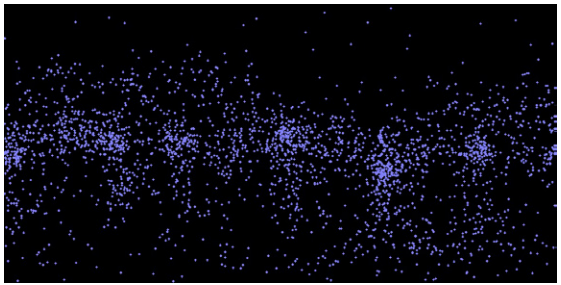
- How can we process the data?
  - **Scanpaths** → Project fixation-points to equirectangular format for storage and distribution.

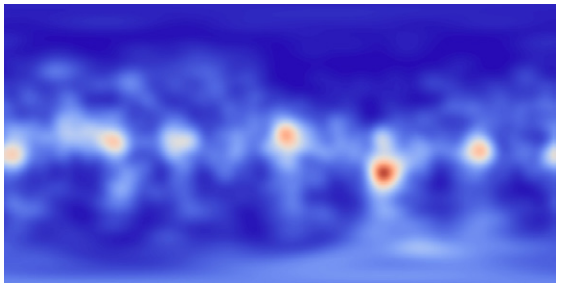


Rai *et al.*, MMSYS2017

Tracking with HMDs + eye-tracking

- How can we process the data?
  - **Saliency maps:**
    - Apply an isotropic Gaussian to each fixation point to account for:
      - Gaze position uncertainty (eye-tracker error)
      - Foveal perception (acuity fall-off)
    - Project to equirectangular



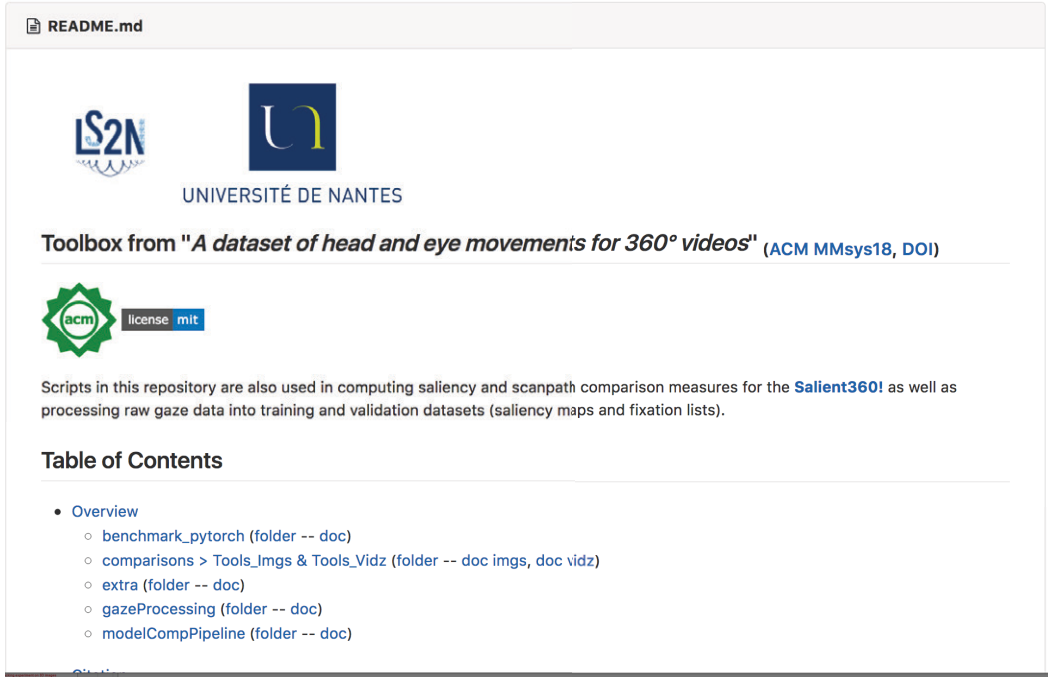


David *et al.*, MMSys2018

Tracking with HMDs + eye-tracking

- How can we process the data?
  - Toolbox coming soon!
    - **Eyes and head tracking**
  - Extended version of:

J. Gutiérrez, E. David, Y. Rai, and P. Le Callet, "Toolbox and dataset for the development of saliency and scanpath models for omnidirectional/360° still images," *Signal Process. Image Commun.*, vol. 69, pp. 35–42, Nov. 2018.


README.md

UNIVERSITÉ DE NANTES

Toolbox from "*A dataset of head and eye movements for 360° videos*" (ACM MMsys18, DOI)

license mit

Scripts in this repository are also used in computing saliency and scanpath comparison measures for the **Salient360!** as well as processing raw gaze data into training and validation datasets (saliency maps and fixation lists).

Table of Contents

- Overview
  - benchmark_pytorch (folder -- doc)
  - comparisons > Tools_Imgs & Tools_Vidz (folder -- doc imgs, doc vidz)
  - extra (folder -- doc)
  - gazeProcessing (folder -- doc)
  - modelCompPipeline (folder -- doc)

What to do with the data?

- Application to the analysis of the results of subjective tests:
  - Influence of **observers' trajectories** → Critical trajectories?
    - Extension of similar works for FVV: S. Ling, J. Gutiérrez, G. Ke, P. Le Callet "Prediction of the Influence of Navigation Scan-path on Perceived Quality of Free-Viewpoint Videos", arXiv:1810.04409, 2018.
  - V-DMOS → taking into account the **observed regions** [Xu *et al.*, arxiv:1709.06342, 2017.]

$$\mathrm{DMOS}_{jr} = \frac{1}{N_{\mathbf{I}_{jr}}} \sum_{i \in \mathbf{I}_{jr}} Z'_{ij}.$$

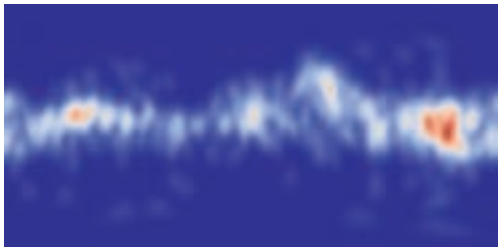
Vector of V-DMOS

$$\left[\mathrm{O\text{-}DMOS}_j \quad \mathrm{DMOS}_{j1} \quad \cdots \quad \mathrm{DMOS}_{jr} \quad \cdots \quad \mathrm{DMOS}_{jR}\right]$$

j → sequence
i → observer
r → region
I_{jr} → Collection of scores for a given sequence and region
f_{rij} → frequency, with which subject i views region r in sequence j,
When $f_{rij} > f_0$, where f_0 is a threshold, subject *i* is added to collection I_{jr}

What to do with the data?

- Application to objective metrics
  - **Weighting the metrics** according to **saliency** [Yu *et al*, ISMAR2015; Rai *et al*. QoMEX2017]



$$\mathrm{PSNR} = 10.\log\left(\frac{I_{max}^2}{\frac{1}{N.M}\sum_{i=1}^{N}\sum_{j=1}^{M}(error(i,j))^2}\right)$$

«Saliency» weight

What to do with the data?

- Content characterization:
  - F. De Simone, J. Gutiérrez, P. Le Callet, “Complexity measurement and characterisation of 360-degree content”, HVEI2019.
    - Appropriate characterization and selection of test material for subjective and objective quality assessment studies
      - **Spatial and temporal indicators** computed in the appropriate representations (computed in equirectangular may not be accurate).
      - **Attention indicators** → content guides attention, exploratory content, etc.

What to do with the data?

- Pushing visual attention:
  - Subjective quality tests are also “attention tests”
    - Provide datasets
    - What can we learn from datasets? → How people explore?
    - Benchmark of models → Provide tools

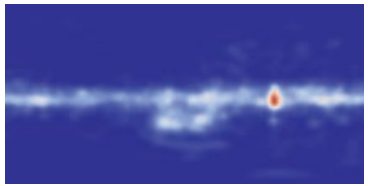
What to do with the data?

- Pushing visual attention:
  - **Salient360! Grand Challenges** at ICME2017 and ICME2018
    - Datasets
    - Tools
    - Model benchmarking

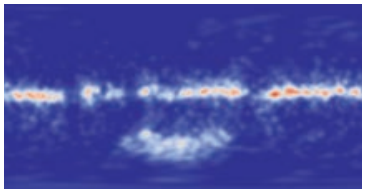
Given a 360º image / video

NEW

Can you predict the **viewport location**? (Track 1)



Can you predict the **viewport location and** the **eye gaze location in the viewport**? (Track 2)



Can you predict the **order of the head trajectory**? (Track 4)

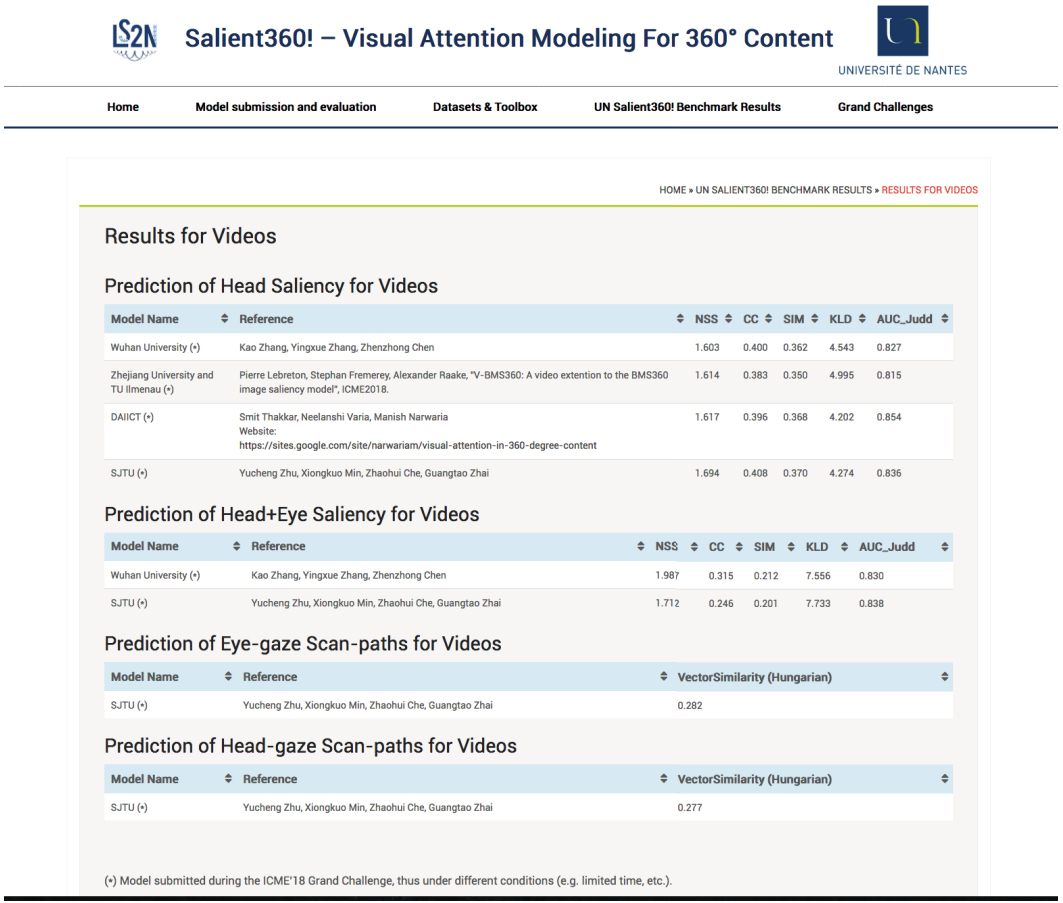
NEW

Can you predict the **order of eye-gaze fixations**? (Track 3)

Conclusions

- Pushing visual attention:
  - Salient360! Benchmark: https://salient360.ls2n.fr/
  - Performance of saliency and scanpath models (links to papers and code).

J. Gutiérrez, E. David, A. Coutrot, M. Perreira Da Silva, P. Le Callet, "Introducing UN Salient360! Benchmark: A platform for evaluating visual attention models for 360 contents", QoMEX2018.



References

- Y. Rai, J. Gutiérrez, and P. Le Callet, "A dataset of head and eye movements for 360 degree images," in *Proceedings of the 8th ACM Multimedia Systems Conference, MMSys 2017*, 2017.
- Y. Rai, P. Le Callet, and P. Guillotel, "Which saliency weighting for omni directional image quality assessment?," in *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, 2017, pp. 1–6.
- E. J. David, J. Gutiérrez, A. Coutrot, M. P. Da Silva, and P. Le Callet, "A dataset of head and eye movements for 360° videos," in *Proceedings of the 9th ACM Multimedia Systems Conference on - MMSys '18*, 2018, pp. 432–437.
- V. Sitzmann et al., "Saliency in VR: How do people explore virtual environments?," IEEE Trans. Vis. Comput. Graph., 2018.
- G. Marmitt and A. T. T. Duchowski, "Modeling visual attention in VR: Measuring the accuracy of predicted scanpaths," in *Eurographics 2002*.
- B. Hu, I. Johnson-Bey, M. Sharma, and E. Niebur, "Head movements during visual exploration of natural images in virtual reality," CISS2017.
- A. Serrano et al., "Movie editing and cognitive event segmentation in virtual reality video", ACM TOG 2017.

References

- M. Yu *et al.,* "A Framework to Evaluate Omnidirectional Video Coding Schemes", ISMAR2015.
- Y-C. Su, et al., "Pano2Vid: Automatic cinematography for watching 360 videos", ACCV2016.
- E. Upenik, T. Ebrahimi. "A simple method to obtain visual attention data in head mounted virtual reality", ICMEW 2017.
- C. Wu, *et al.,* "A Dataset for Exploring User Behaviors in VR Spherical Video Streaming", ACM MMSys'17.
- X. Corbillon *et al.* "Viewport-adaptive navigable 360-degree video delivery", ICC2017.
- S. Fremerey, K. Meseberg, and A. Raake, "AVTrack360: An open Dataset and Software recording people's Head Rotations watching 360 ◦ Videos on an HMD," *ACM Multimed. Syst. Conf.*, 2018.
- M. Xu, C. Li, Z. Wang, and Z. Chen, "Visual Quality Assessment of Panoramic Video," pp. 1–12, http://arxiv.org/abs/1709.06342, Sep. 2017.
- C. Ozcinar and A. Smolic, "Visual Attention in Omnidirectional Video for Virtual Reality Applications," in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, 2018, no. August, pp. 1–6.
- Y. Xu *et al.*, "Gaze Prediction in Dynamic 360º Immersive Videos," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 5333–5342, 2018.

Monitoring user behavior in subjective tests with 360-degree video

Jesús Gutiérrez, Patrick Le Callet

Image, Interaction, Perception Group (IPI)
Laboratoire des Sciences du Numérique de Nantes (LS2N)
Université de Nantes

VQEG Meeting
14/11/2018