A dataset of head and eye movements for 360° images

Yashas Rai, Jesús Gutiérrez, Patrick Le Callet

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

08/05/2017 - VQEG Meeting

Introduction and motivation

- Visual attention allows to know the important regions of the scene for the observers.
- Proxy for visual fidelity widely used for 2D and 3D content
  - Coding and transmission: protection based on saliency.
  - Quality evaluation:
    - Weighting most important regions.
    - Consider artistic intentions.

J-S. Lee, F. De Simone, T. Ebrahimi, "Efficient video coding based on audio-visual focus of attention", *Visual Communication and Image Representation*, vol 22, no. 8, pp. 704–711, Nov. 2011.

Q. Huynh-Thu, M. Barkowsky, P. Le Callet, "The importance of visual attention in improving the 3D-TV viewing experience: Overview and new perspectives", *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 421–431, Jun 2011

M. Narwaria, M. Perreira Da Silva, P. Le Callet, and R. Pepion, "Tone mapping based HDR compression: Does it affect visual experience?," *Signal Process. Image Commun.*, vol. 29, no. 2, pp. 257–273, 2014.

Introduction and motivation

- Even more important for 360 content: Not everything may be seen
  - Tile-based coding and streaming, non-uniform quality streaming...
  - Evaluation of quality using head-motion data: weighting the metrics.
- Proposed dataset containing exhaustive data of **head** and **eye** movements.
- Publication of the dataset:
  - Paper in **MMSys 2017.**
  - **ICME 2017 Grand Challenge**: Encouraging participants to submit computational models for saliency

K. Kammachi Sreedhar, *et al*., "Viewport-adaptive encoding and streaming of 360-degree video," *IEEE ISM*, Dec. 2016.

M. Yu, H. Lakshman, and B. Girod, "A Framework to Evaluate Omnidirectional Video Coding Schemes", IEEE ISMAR, Sep. 2015.


ICME'17 Home Authors Workshops Grand Challenges Keynotes Tutorials Organizing Commi

Deadline of submission June 4, 2017 (8am PDT)

Notification of acceptance June 18, 2017 (8am PDT)

Salient360!: Visual attention modeling for 360° Images Grand Challenge

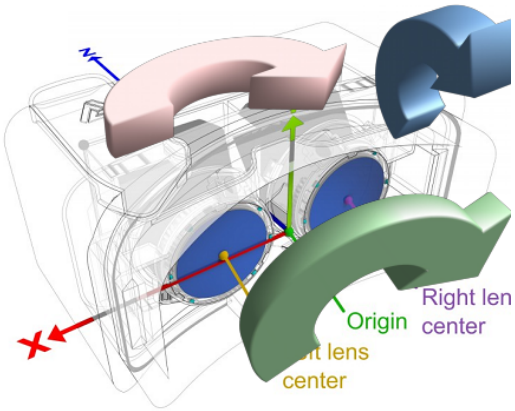
Organizer: University of Nantes, Technicolor

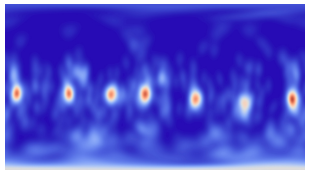
Understanding how users watch a 360° image and analyzing how they scan through the content with a combination of head and eye movement, is necessary to develop appropriate rendering devices and also create good VR/AR content for consumers. Good visual attention modelling is a key factor in that perspective that helps enhance the overall Quality of Experience (QoE). Although a huge number of algorithms have been developed in recent years to gauge visual attention in flat-2D images and videos and also a benchmarking platform where users can submit and assess their results, attention studies in 360 scenarios are absent. The goal of this challenge is to therefore two-fold:

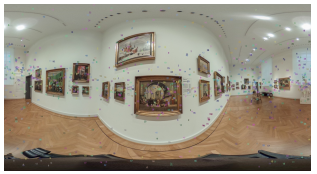
- to produce a dataset to ensure easy and precise reproducibility of results for future saliency / scan-path computational models in line with the principles of Reproducible and Sustainable research from IEEE.
- to set a first baseline for the taxonomy of several types of visual attention models (saliency models, importance models, saccadic models) and the correct methodology and ground-truth data to test each of them.

Dataset description

- **60 images** saved in equi-rectangular format.
- Raw **eye+head** tracking data:
  - Rotational angles.
  - Translational movements.
  - [2x2] Eyes-movement within the viewport projection.
  - Time stamp of capture and record.
- Processed data:
  - Head saliency maps.
  - Head-Eye saliency maps.
  - Scan-paths.
- Tools for processing the data.







Subjective experiment

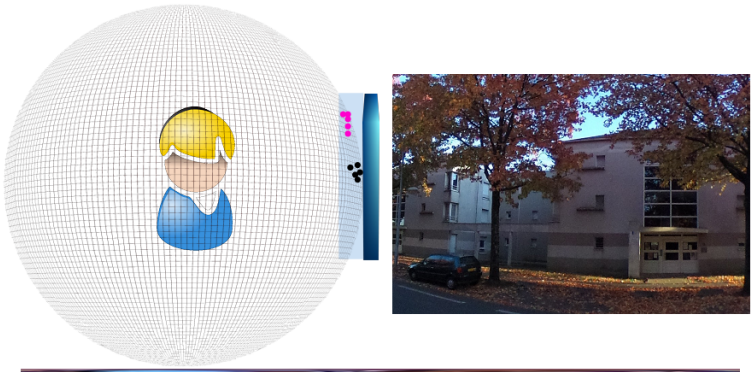
- Test stimuli
  - 98 omnidirectional images:
    - 60 released in the dataset: captured with a *Ricoh Theta S* camera.
    - 38 images from *Flicker* by professionals (CC)
    - Resolutions from 5376x2688 to 18332x9166 (equi-rectangular).
  - Wide coverage of characteristics:
    - Indoor: small rooms and wide halls.
    - Outdoor: natural landscapes and cityscapes.
    - People.

Subjective experiment

- Equipment:
  - HMD Oculus Rift DK2
    - Horizontal and vertical FoV: 100°
    - 1920 x 1080 resolution.
  - SMI Eye-tracker
    - Binocular eye-tracking at 60Hz.
- Observers:
  - 63 (24 females / 39 males).
  - Average age 30 (from 19 to 52).
  - 40 observers per image.
  - Expertise: 32/63 used HMD less than 2 times, 8 experts.
- Execution of the test:
  - Free-viewing: “view as naturally as possible”.
  - Each stimulus: 25 seconds (6 seconds between stimulus).
  - 35 minutes + 5 minutes pause.

Processing the results

- Classification of gaze data into fixations and saccades.
- Projection Formats
  - Equi-rectangular projection
    - Representation and distribution.
  - Rectilinear projection
    - Rendered in view-port.
    - Eye-data capture.
  - Sphere-map projection
    - Computing Fixations.
    - Comparing two saliency maps.

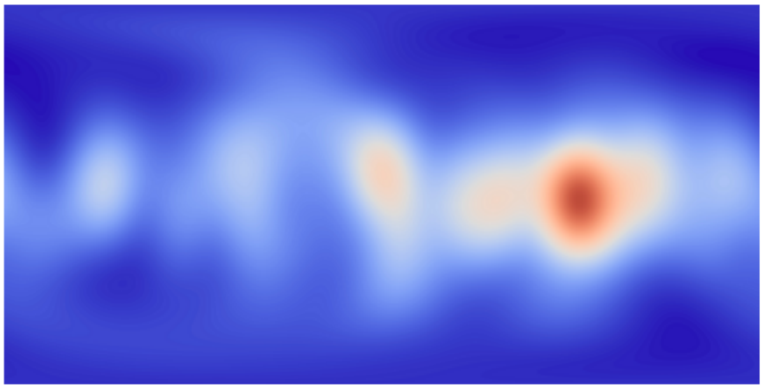


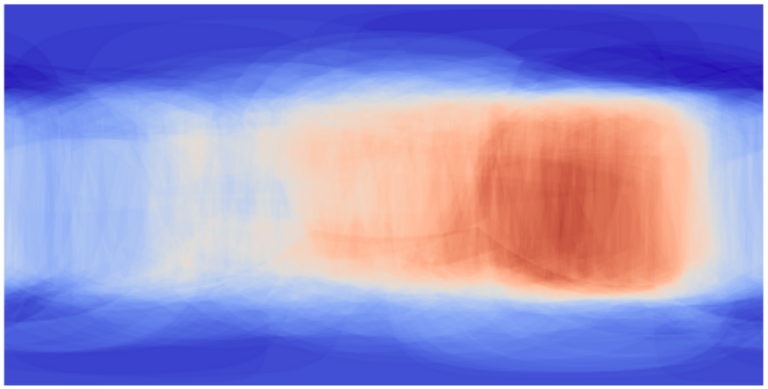
Scan-path data

Saliency maps

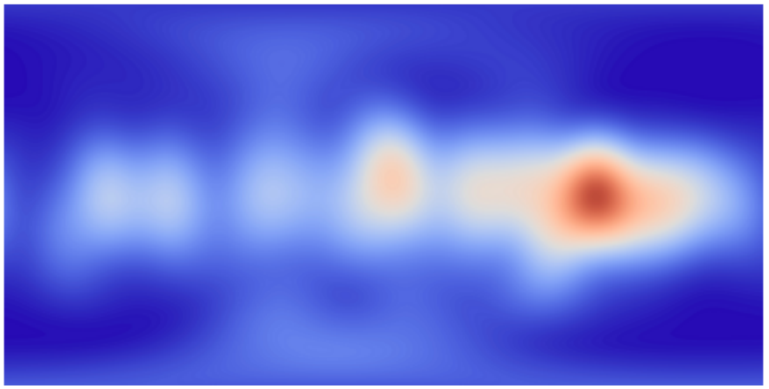
Original equi-rectangular image

Saliency map from head+eye movements





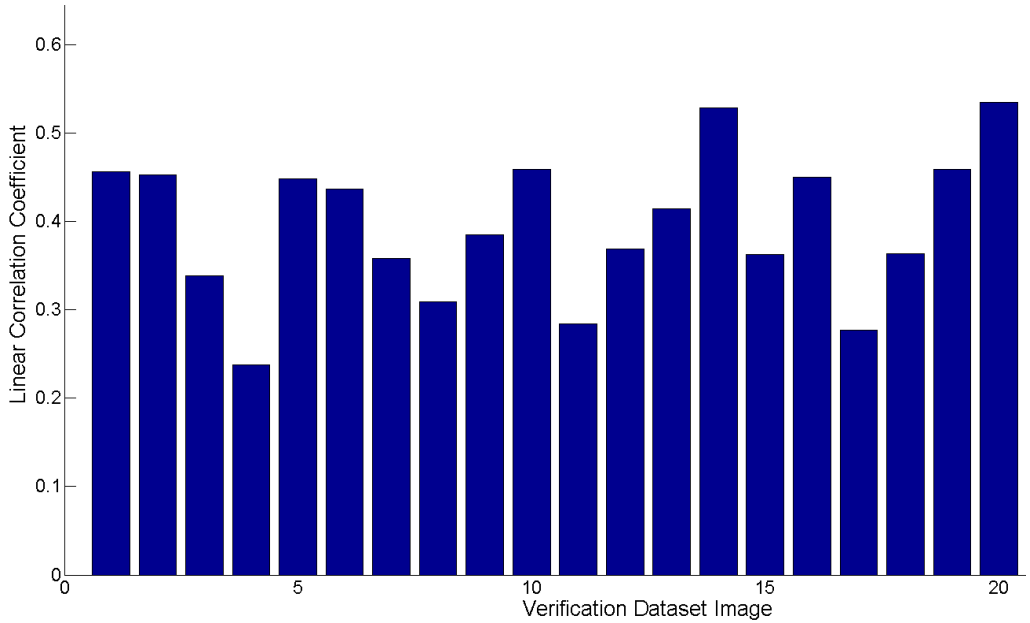
Saliency map from head movement considering entire viewport



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

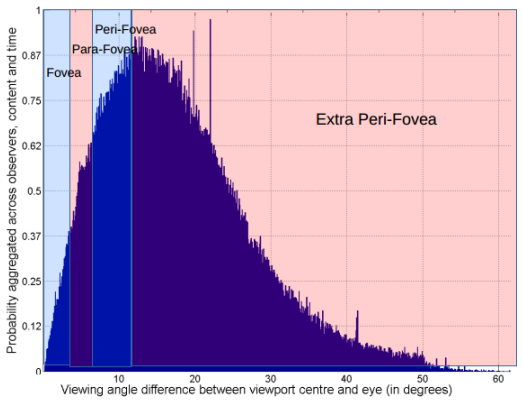
Some statistical results

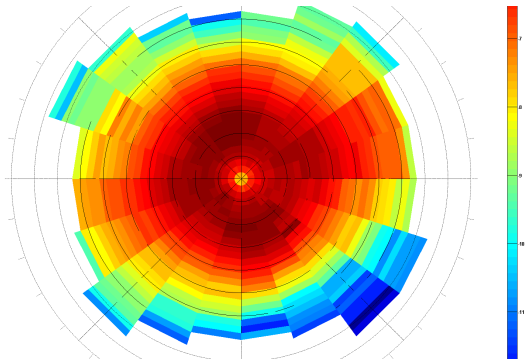
- Comparison of **head** vs. **head+eye** saliency maps:
  - **Poor correlation** between **head-only** and **head+eye saliency maps**

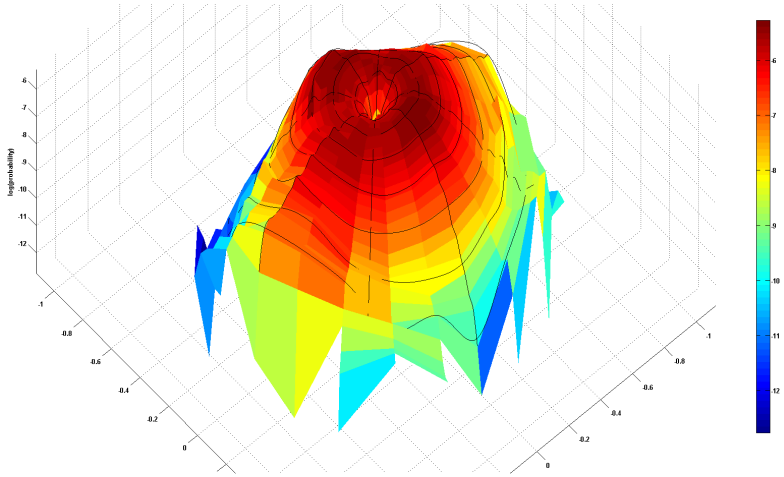


Some statistical results

- Eye-movement data
  - Do people really look at the centre?
    - The peak is offset by 14-16 degrees from the centre

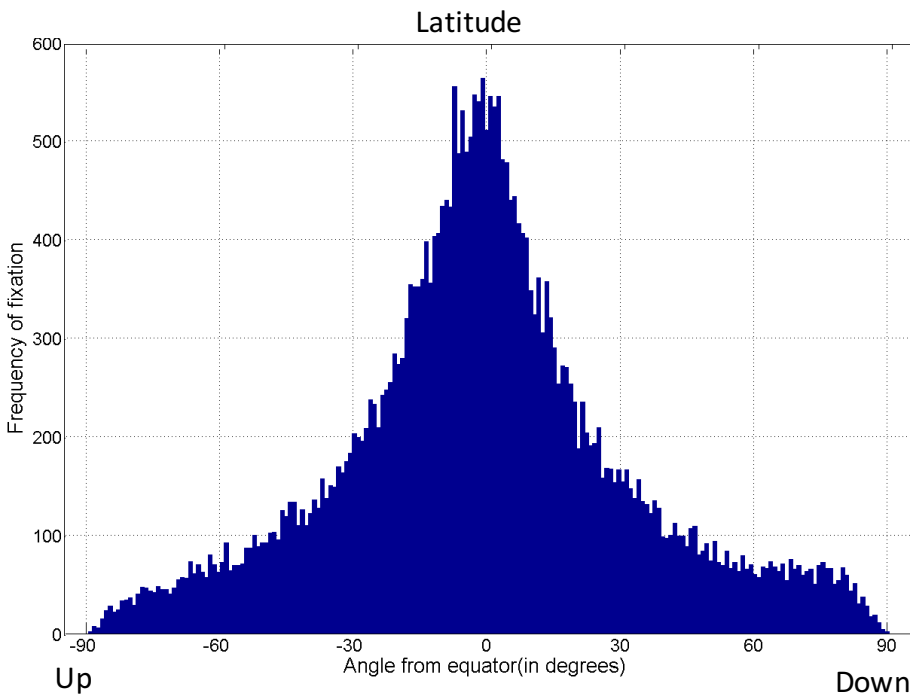


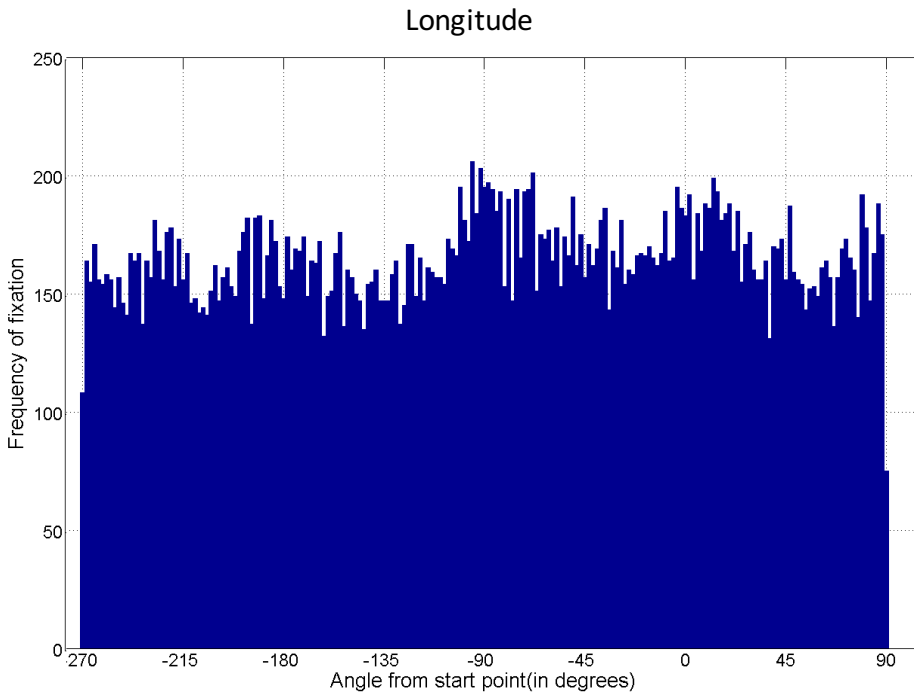




Some statistical results

- Scanning strategies: frequency of fixations in accordance to the elevation (pitch) angle and the azimuthal (yaw) angle.

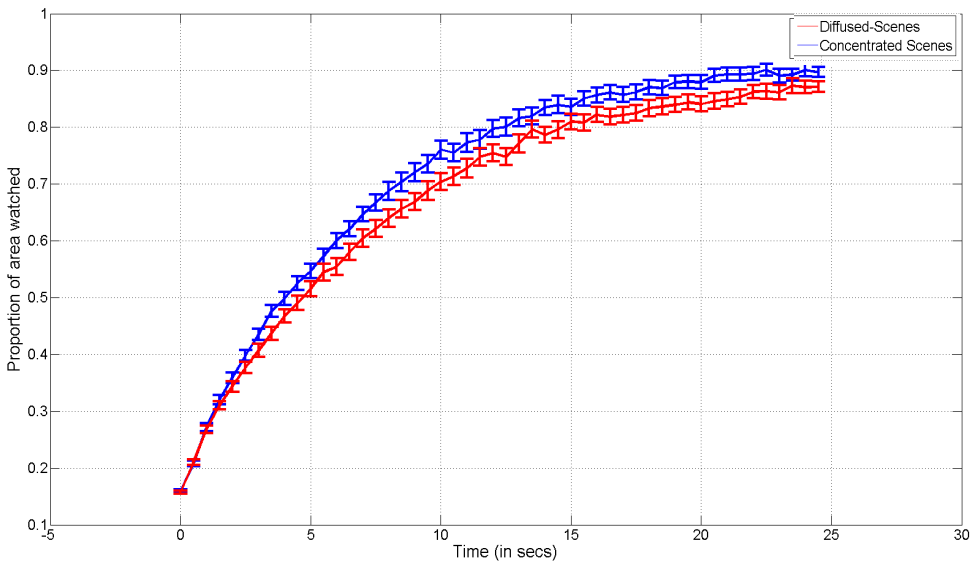




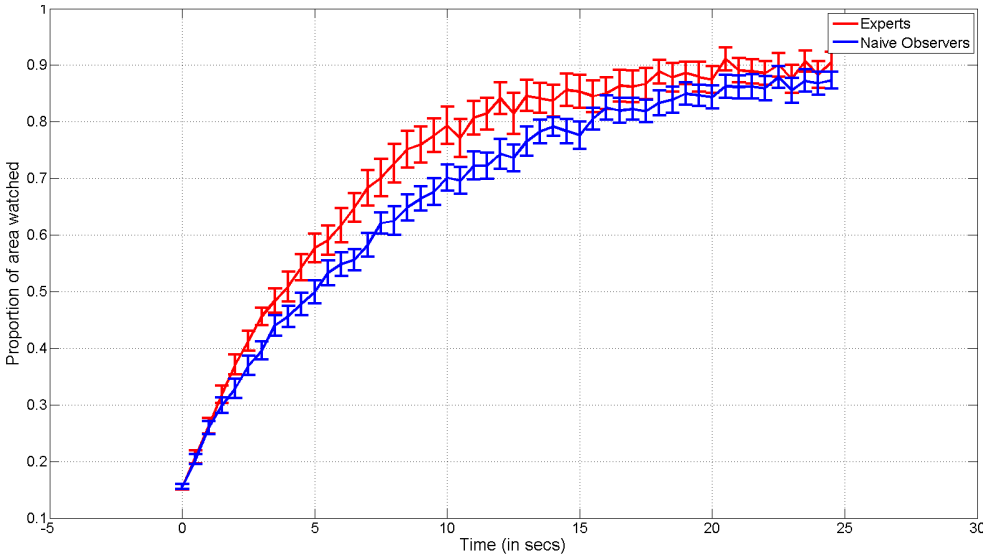
Some statistical results

- Speed of exploration:
  - Explored area of the sphere vs. time.

By image type



By observer type



A dataset of head and eye movements for 360° images

Yashas Rai, Jesús Gutiérrez, Patrick Le Callet

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

08/05/2017 - VQEG Meeting