Fovea

Attention-Guided Video Production

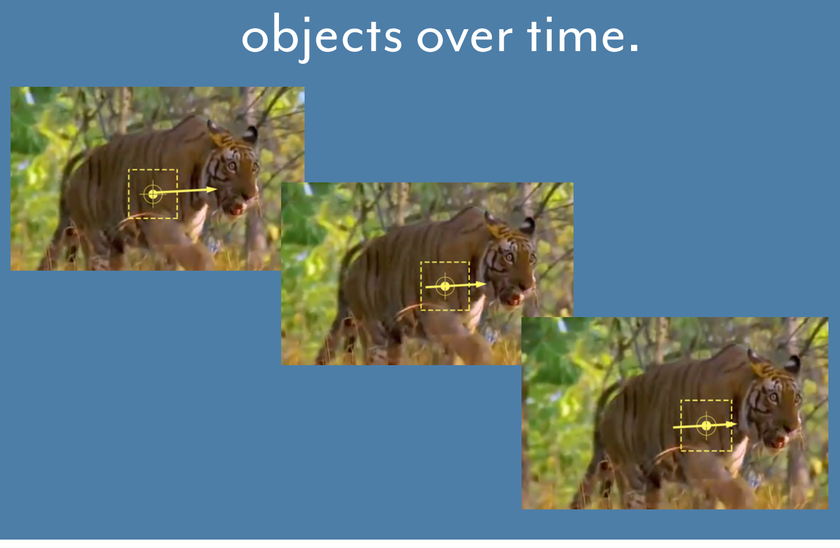
FRACTAL · COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK · CUNY THE CITY UNIVERSITY OF NEW YORK

Munna Koorathota • Patrick Adelman • Kelly Cotton

STEP 1: Pre-process videos into .fvid metadata format

Custom object detection uses pre-trained model (YOLOv3).

Centroid tracking stabilizes objects over time.



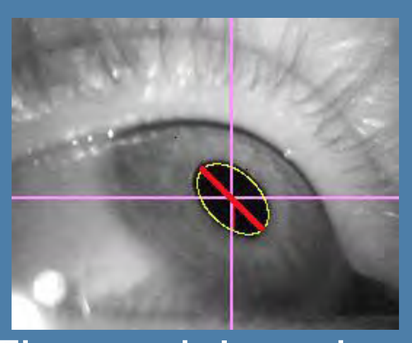
In-depth color information is extracted frame by frame.

STEP 2: Train model to predict attention

The DIEM Project
Dynamic Images and Eye Movements

Open-source datasets were used to train the model.

An LSTM model trains on eye tracking sequence data.

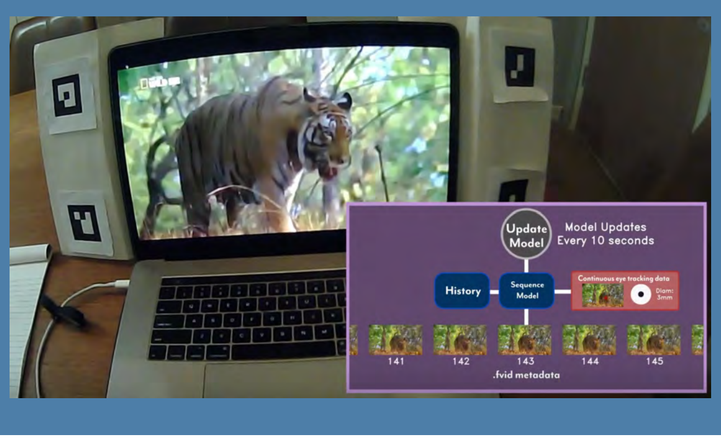


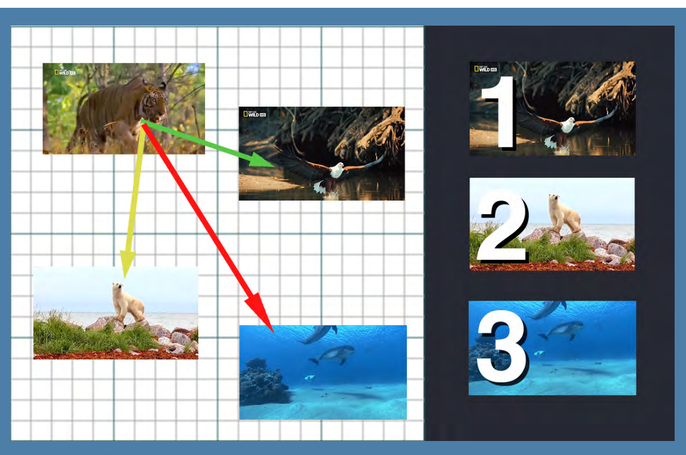
The model predicts pupil size during fixations.

STEP 3: Edit video real-time using model

Eye tracking data is collected real-time via Pupil Core.

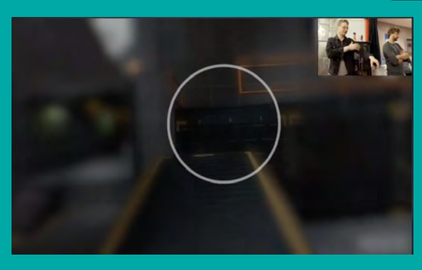
Model is applied to detect attention for objects of interest.





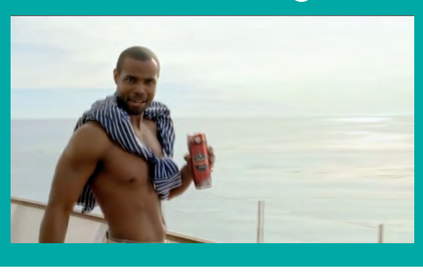
Fovea uses color and object metadata to form a ranked queue. It searches the .fvid space to choose the next video.

Potential Uses



Foveated Rendering

Advertisement and film editing



Educational and safety attentional verification

Citations

1. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
2. Kassner, M., Patera, W., & Bulling, A. (2014, September). Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication (pp. 1151-1160). ACM.
3. Mathe, S., & Sminchisescu, C. (2014). Actions in the eye: Dynamic gaze datasets and learnt saliency models for visual recognition. IEEE transactions on pattern analysis and machine intelligence, 37(7), 1408-1424.
4. Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2011). Clustering of gaze during dynamic scene viewing is predicted by motion. Cognitive Computation, 3(1), 5-24.
5. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

POWERED BY Pupil Labs

Who are we?

Fractal Media is a boutique video agency based in NYC that works at the intersection of creativity and technology. In addition to full-service video production, we conduct research in computer science and psychology to leverage state-of-the-art tools for storytelling.

Over the past year, our advertising and original content in science and technology has received over 3 million views online and coverage from Slate, NPR, Discovery Channel and more.

Project Fovea involves the integration of eye tracking and video production with the goal of understanding how attention can be used in the creation process. We will be seeking funding over the next year to explore the use of our framework in video editing and content presentation in AR/VR.

http://www.fractal.nyc/fovea