Deep Learning for Eye Tracking

Project description

Eye tracking has its roots in psychology, where a subject's point of gaze is tracked to determine interest and attention. Recent works have shifted eye tracking from more traditional and active methods to passive video-based methods. Such remote video methods allow for eye tracking to be applied in many form factors and scenarios due to the prevalence of front-facing cameras on everyday computing devices. The expansion beyond the laboratory environment leads to issues such as over/under-exposure, low image sensor quality, motion blur and complexity due to head movement. Such factors cannot easily be modeled. Thus, machine learning methods have been used increasingly in appearance-based gaze estimation tasks.

Contrary to more conventional methods of eye tracking appearance-based gaze estimation methods attempt to regress directly on image data. Sugano et al. (2014) synthesized large sets of eye images and estimated gaze using random forests. The work involved acquisition of images from human subjects from multiple perspectives, then synthetically rotating the heads to produce a large variation in head and eye gaze angles. Zhang et al. (2015) applied their model on particularly challenging data in real-world situations. They achieved this by acquiring images from laptops of 15 subjects, taken in a large variety of conditions. The data exhibit variations in lighting conditions, head angle, and image quality. Zhang et al. employ a convolutional neural network (based on LeNet) to address challenges arising from the new data. Latest works expand on that suggested by Zhang et al. (2015), often acquiring new data and always suggesting models and architectures which improve gaze estimation accuracy.

In this project, we aim to explore appearance-based gaze estimation using convolutional neural networks, with the aim of improving on state-of-the-art performance. We will review literature in appearance-based eye tracking, in particular when using convolutional neural networks, and evaluate datasets which have been made available recently. We will then analyze models proposed so far and attempt to make valuable contributions in improving gaze estimation accuracy by modifying the existing model or suggesting new models. Therefore, good knowledge of machine learning is necessary.

![Figure 1. Sample eye region images from the UT Multiview dataset [1]](image)


Keywords  Eye tracking, Appearance-based gaze estimation, Deep convolutional neural networks, Machine Learning

Goal  Design and implement an appearance-based gaze estimation method using deep learning.

Work packages

- Literature review on state-of-the-art appearance-based gaze estimation methods and datasets.
- Implementation of gaze estimation using a convolutional neural network.
- Broad evaluation of the system using established datasets.

Required skills

- Solid programming skills (C++, Python, or similar).
- Experience with training deep convolutional neural networks is a must.
- High motivation, independence and eagerness to challenge the state-of-the-art.

Type  MA
Time period  Spring 2017
Internal supervisors  Seonwook Park spark@inf.ethz.ch, Otmar Hilliges otmar.hilliges@inf.ethz.ch