
COMPUTER VISION – Project 2S

Automatic Gaze Estimation

Gaze estimation techniques can provide useful clues into users' attention or can improve the lives of people with disabilities by allowing them to interact with computers. Until a few years ago, eye tracking or gaze estimation technologies were limited to laboratory environments, but recent advancements in the computer vision field allow dealing with this problem in more challenging environments using standard cameras.

It looks relatively easy to estimate the gaze direction from an image of a face, but limited resolution, motion blur, variable head poses, eyeglasses and illumination conditions can cause ambiguities in determining the eyeball shape of the iris.

The task of this project is to firstly localize the eyes in color images (you can use the provided ones to evaluate your algorithm) and then analyze the eye region in order to get the colored region of the iris and estimate the gaze direction. For the scope of this project a simple classification of the direction into 3 categories (i.e., looking straight, looking right, looking left) is sufficient. Notice that advanced approaches try to accurately estimate the viewing angle, but this is not required for the course project.

You can use both deterministic approaches like template matching or the Hough Transform for circles or machine learning techniques, e.g., use Viola and Jones to get the face and then try to get the eyes by knowing their position in the face. Deep learning approaches can also be used.

The application must be developed in C++ with the OpenCV library. The only allowed exception is the usage of Python code for the deep learning part if you decide to exploit this family of techniques.

The application should be able to detect the eyes and get the gaze in the provided images (at least in the simple cases). Some images like the ones with several people or with part of the face or eyes occluded are more challenging, however it is not required to solve all the cases. Feel free to try on different images or to acquire your own data with a camera or smartphone.

If you would like to try different images or you need labeled data for training machine learning approaches you can have a look at these datasets:

- <https://www.mpi-inf.mpg.de/departments/computer-vision-and-machine-learning/research/gaze-based-human-computer-interaction/appearance-based-gaze-estimation-in-the-wild/>
- <https://zenodo.org/record/2529036#.XUFhZnvOOUk>

