GAZE TRACKING TO MAKE MOBILE CONTENT INTERACTIVE

Nikolaus West, Catalin Voss
{nwest2, catalin} @ stanford.edu

Introduction
Content on the internet is a one-sided affair. Media plays and the viewer watches. With the arrival of front-facing, monocular webcams in a variety of device on which content is consumed, such as laptops, tablets, or smartphones, we think media should respond to how the viewer reacts.

We propose an efficient approach to score the visual engagement of a person based on their head pose and eye center location. This will allow us to integrate with with toolkits that enable content-providers to create interactive media, for example a path-based video editor. This way, a recorded lecture can prompt questions at a students when they lack attention or a training video can re-engage an employee (e.g. by pausing the clip). Further, content providers can gather the data they need to improve their media.

Algorithm
1. Find bounding rect of the face (using Viola-Jones on Android and Apple’s AVFoundation framework, a GPU-accelerated black-box face bounds locator) along with basic yaw and roll information; break algorithm if out of place
2. Track reference points using Constrain Local Model (CLM) trained on large datasets [1]
3. Estimate head pose (rotation matrix and translation vector) in 3D space using POSIT
5. Detect planar images of faces as opposed to real 3D faces using basic blink detection
6. Predict gaze by using CLM points as references for eye center; estimate where on the screen a user is looking

Milestones
1. Real time face tracking on mobile utilizing available open-source code using multiple landmarks trained on large databases
2. Mobile head pose estimation using results from face tracking; determine when a person seems to be looking at the screen based on their head pose
3. Basic eye center localization
4. Gaze estimation based on pose information (as reference) and eye center
5. Basic blink detection; determine whether we are dealing with a real person
6. Reduce processing time and memory footprint [3]

References