OpenFace

Links

- GitHub: https://github.com/TadasBaltrusaitis/OpenFace
- Paper: https://www.cl.cam.ac.uk/research/rainbow/projects/openface/wacv2016.pdf

Notes

- Goal: bridge the gap between existing state-of-the-art research (for facial behavior analysis) and easy to use out-of-the-box solutions for facial behavior and analysis
    - OpenFace is the first open source tool for facial behavior analysis
- Capable of facial landmark detection, head pose estimation, facial action unit recognition, and eye-gaze estimation
    - types of facial behavior
- Capable of real-time performance
- Able to run from a webcam without any specialist hardware
- Allows for easy integration with other applications and devices through a lightweight messaging system
- Allows for detection of multiple faces in an image and tracking of multiple faces in videos
    - Videos -- achieved by keeping a track of active face tracks and a logic module that checks for people leaving and entering the frame
- Can operate on real-time data video feeds from a webcam, recorded video files, image sequences, and individual images
    - Can save outputs of the processed data as CSV files
- Landmark detection
    - Model
        - Uses Conditional Local Neural Fields (CLNF) for facial landmark detection and tracking
            - Main components: Point Distribution Model (captures landmark shape variations) and patch experts (captures local appearance variations of each landmark)
            - Extends CLNF model (which detects all 68 facial landmarks together) by training separate sets of point distribution and patch expert models for eyes, lips, and eyebrows, then by fitting the landmarks detected with individual models to a joint (PDM)
        - Employs a face validation step -- uses a three layer CNN that, given a face aligned using a piecewise affine warp, is trained to predict the expected landmark detection error
    - Implementation

        - Multi-scale patch experts → accuracy on both lower and higher-resolution face images
        - Optimal results achieved when face is at least 100px across
        - Training on different views → can track faces with out of plane motion and can model self-occlusion caused by head rotation
        - Initialized CLNF model with face detector from dlib library
            - Used linear mapping from the bounding box provided by dlib detector to the one surrounding the 68 facial landmarks
        - Tracking landmarks in videos -- initialized CLNF model based on landmark detections in previous frame
- Eye gaze estimation
    - Detects the location of the eye and pupil using CLNF model
    - Uses above info to compute the eye gaze vector individually for each eye
        - Fire a ray from the camera origin through the center of the pupil in the image plane and compute its intersection with the eyeball sphere -- gives pupil location in 3D camera coordinates
        - Estimated gaze vector: vector from the 3D eyeball center to the pupil location

On infants

- Works well out of the box at finding infant faces
- Could not detect face/gaze when half of face blocked off/offscreen
- Could not detect face/gaze when head turned to look sideways (2:38) → messed up face detection afterwards