Qualitative Analysis of Automated Visual Tracking of Objects through Head Pose Estimation

Ayeshka Abeysinghe[1], Isuri Devlini Arachchige[1], Pradeepa Samarasinghe[1], Vidushani Dhanawansa[1], and Menan Velayuthan[1]

[1]Faculty of Computing , Sri Lanka Institute of Information Technology, Sri Lanka.
ayeshkaabeysinghe@gmail.com, arachchigeisuridevlini@gmail.com, pradeepa.s@sliit.lk,
vidushani.d@gmail.com, menan.v@sliit.lk

Abstract

An automated approach for object tracking and gaze estimation via head pose estimation is crucial, to facilitate a range of applications in the domain of human - computer interfacing, this includes the analysis of head movement with respect to a stimulus in assessing one's level of attention. While varied approaches for gaze estimation and object tracking exist, their suitability within such applications have not been justified. In order to address this gap, this paper conducts a quantitative comparison of existing models for gaze estimation including Mediapipe and standalone models of Openface and custom head pose estimation with MTCNN face detection; and object detection including models from CSRT object tracker, YOLO object detector, and a custom object detector. The accuracy of the aforementioned models were compared against the annotations of the EYEDIAP dataset, to evaluate their accuracy both relative and non-relative to each other. The analysis revealed that the custom object detector and the Openface models are relatively more accurate than the others when comparing the number of annotations, absolute mean error, and the relationship between x displacement-yaw, and y displacement - pitch, and thereby can be used in combination for gaze tracking tasks.

Keywords

Head Pose Estimation, Gaze Estimation, Object Detector, Object Tracker, Openface, MTCNN, Mediapipe