

# RT-GENE: Real-Time Eye Gaze Estimation in Natural Environments

Tobias Fischer, Hyung Jin Chang, and Yiannis Demiris

Personal Robotics Laboratory, Department of Electrical and Electronic Engineering,  
Imperial College London, UK

{t.fischer, hj.chang, y.demiris}@imperial.ac.uk

**Abstract.** In this work, we consider the problem of robust gaze estimation in natural environments. Large camera-to-subject distances and high variations in head pose and eye gaze angles are common in such environments. This leads to two main shortfalls in state-of-the-art methods for gaze estimation: hindered ground truth gaze annotation and diminished gaze estimation accuracy as image resolution decreases with distance. We first record a novel dataset of varied gaze and head pose images in a natural environment, addressing the issue of ground truth annotation by measuring head pose using a motion capture system and eye gaze using mobile eyetracking glasses. We apply semantic image inpainting to the area covered by the glasses to bridge the gap between training and testing images by removing the obtrusiveness of the glasses. We also present a new real-time algorithm involving appearance-based deep convolutional neural networks with increased capacity to cope with the diverse images in the new dataset. Experiments with this network architecture are conducted on a number of diverse eye-gaze datasets including our own, and in cross dataset evaluations. We demonstrate state-of-the-art performance in terms of estimation accuracy in all experiments, and the architecture performs well even on lower resolution images.

**Keywords:** Gaze estimation · Gaze dataset · Convolutional Neural Network · Semantic inpainting · Eyetracking glasses

## 1 Introduction

Eye gaze is an important functional component in various applications, as it indicates human attentiveness and can thus be used to study their intentions [9] and understand social interactions [41]. For these reasons, accurately estimating gaze is an active research topic in computer vision, with applications in affect analysis [22], saliency detection [42, 48, 49] and action recognition [31, 36], to name a few. Gaze estimation has also been applied in domains other than computer vision, such as navigation for eye gaze controlled wheelchairs [12, 46], detection of non-verbal behaviors of drivers [16, 47], and inferring the object of interest in human-robot interactions [14].

Deep learning has shown successes in a variety of computer vision tasks, where their effectiveness is dependent on the size and diversity of the image







































