Pattern Recognition Letters 179 (2024) 45–51

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Detection of visual pursuits using 1D convolutional neural networks

Alex Torquato S. Carneiro [a,*], Flavio Luiz Coutinho [b], Carlos H. Morimoto [a]

[a] Institute of Mathematics and Statistics, University of São Paulo, Brazil
[b] School of Arts, Sciences and Humanities, University of São Paulo, Brazil

ARTICLE INFO

Edited by Jiwen Lu

Keywords:
Eye tracking
Eye gaze pursuit
Convolutional neural networks
Eye movements classification

ABSTRACT

The visual pursuit of moving targets is a natural behaviour that has been exploited in, for example, medical diagnosis, law enforcement, and human computer interaction. Most proposed algorithms to detect this behaviour are based on some kind of motion similarity metric that assumes small or no distortion between the trajectory described by the target being pursued and the sensor measurements. We propose a novel algorithm based on 1D Convolutional Neural Networks (1D-CNNs), and investigate the performance of the 1D-CNN against 4 state-of-the-art similarity based algorithms (SAs), using a novel dataset containing data from 10 participants. Their performances are evaluated using two trajectory shapes (circle and square), two target speeds (120°/s and 240°/s), and three window sizes (0.5, 1.0 and 1.5 s). All algorithms have been trained or optimized to maximize the Matthew's Correlation Coefficient (MCC). Our results show that the 1D-CNN outperforms the SAs in all cases, requiring smaller window sizes to robustly detect the pursuits.

1. Introduction

As video cameras become available in multiple platforms, the analysis of eye movement data finds novel applications, such as biometrics Galdi et al. [7] and fatigue detection Wang et al. [26]. The eye tracking data is typically provided as eye camera pixel coordinates corresponding to eye features such as the pupil centre. Typically, eye gaze-based applications map eye coordinates to the surface of a computer screen (that generates the visual stimulus) using a transformation function computed by a calibration procedure.

Because obtaining and maintaining an accurate calibration is still challenging, many calibration-free gaze estimation methods have been proposed [8]. Among them, those that leverage visual pursuits have received great interest. For example, eye movement pursuits have been applied in medicine to help diagnose people with traumatic brain injury [10,11], in law enforcement for field sobriety tests [16], and for gaze interaction [19,25].

Nonetheless, detecting when a person is visually pursuing a specific target is still challenging due to inherent sensor noise, geometric distortions, and latency. Fig. 1 shows a particularly difficult situation due to the small display size used to render the visual target and poor screen stability to hold a calibration mapping.

Velloso et al. introduced four algorithms for the detection of visual pursuits of targets describing circular trajectories [23]. These algorithms are based on functions that compute the similarity between the gaze trajectory and the trajectory of one specific object on the display, as seen in Fig. 1. A hand-crafted similarity threshold then discriminates smooth pursuit from other eye patterns. One limitation of these algorithms is that they fail in cases with large variations in rotation and perspective distortions between the trajectories of the eye and the target [24].

Deep learning-based systems have shown excellent results for pattern recognition when enough data is available to learn the features required to discriminate classes with degraded data without an explicit model of the distortions [1,3,14,27].

In this paper we introduce a novel deep learning-based algorithm for visual pursuit detection on raw video data which uses a 1-dimensional Convolutional Neural Network (1D-CNN) to detect the pursuit patterns. We also investigate the performance of the proposed 1D-CNN algorithm against four other state-of-the-art algorithms for visual pursuit detection. Our investigation considers a wearable real-time interactive scenario [22] as shown in Fig. 1, in which CNN models can learn shape features to classify the data using a lightweight and efficient architecture that demands less processing resources than more complex deep learning models.

The first contribution of this paper is the 1D-CNNs-based algorithm for detecting visual pursuits. The 1D-CNNs are more robust to noise and projective distortions of the trajectory shapes so it can be applied directly to the raw eye-tracking data. Nonetheless, CNNs require

* Corresponding author.
E-mail address: alexc@alumni.usp.br (A.T.S. Carneiro).

https://doi.org/10.1016/j.patrec.2024.01.020
Received 26 December 2022; Received in revised form 23 October 2023; Accepted 21 January 2024
Available online 24 January 2024