Chapter 51
Detection of Hands for Hand-Controlled Skyfall Game in Real Time Using CNN

Neha B.
Anna University, Chennai, India

Naveen V.
Anna University, Chennai, India

Angelin Gladston
Anna University, Chennai, India

ABSTRACT

With human-computer interaction technology evolving, direct use of the hand as an input device is of wide attraction. Recently, object detection methods using CNN models have significantly improved the accuracy of hand detection. This paper focuses on creating a hand-controlled web-based skyfall game by building a real time hand detection using CNN-based technique. A CNN network, which uses a MobileNet as the feature extractor along with the single shot detector framework, is used to achieve a robust and fast detection of hand location and tracking. Along with detection and tracking of hand, skyfall game has been designed to play using hand in real time with tensor flow framework. This way of designing the game where hand is used as input to control the paddle of skyfall game improved the player interaction and interest towards playing the game. This model of CNN network used egohands dataset for detecting and tracking the hands in real time and produced an average accuracy of 0.9 for open hands and 0.6 for closed hands which in turn improved player and game interactions.

DOI: 10.4018/978-1-6684-7589-8.ch051

INTRODUCTION

Hand detection is an essential step to support many tasks including human computer interaction (HCI) applications. However, the ability to detect, localize and track the hands is crucial in many applications. While egocentric video captures a huge variety of objects, activities, and situations, one specific object is omnipresent in nearly every frame: the hands. Human's hands is used as a main channel of interaction with the physical world, for manipulating objects, sensing the environment, and expressing ourselves to other people. There are several existing approaches to tracking hands in the computer vision domain. Incidentally, many of these approaches are rule based extracting background based on texture and boundary features, distinguishing between hands and background using color histograms and HOG classifiers, making them not very robust. With sufficiently large datasets, neural networks provide opportunity to train models that perform well and address challenges of existing object tracking/detection algorithms — varied/poor lighting, diverse viewpoints and even occlusion. The main drawbacks is that they can be complex, are relatively slow compared to tracking-only algorithms.

In this work, tensorflow framework is used to build the model of hand detection and tracking and puts forward a convolution network model based on tensorflow framework. A vision based hand recognition approach using convolutional neural networks on raw video data has been developed. Furthermore, this entire area of work has been made more approachable by deep learning frameworks such as the tensorflow object detection API that simplify the process of training a model for custom object detection. More importantly, the advent of fast neural network models like ssd, faster r-cnn, rfcn etc make neural networks an attractive candidate for real-time detection and tracking applications.

Using parts of the human body as input for the gaming system has its own advantages. Because the user is always available and the user does not require to carry any secondary device. Results from a real time object detection model could be mapped to the controls of a game. Importantly, appropriating the use of various parts of the human body for gesture based interaction in multiple game environments has been shown to improve user experience and overall engagement (Birk, 2013). Though the idea of using body part as a source providing input is not entirely new, existing approaches which control computer vision, wearables as well as sensors sometimes suffer from current accuracy related challenges. Further, they are not always portable and can impose integrate related issues with respect to other software. Advances in light-weight deep neural networks, specifically models for object detection (Huang, 2017). Enormous works has been carried out, especially on key point extraction (Cao, 2017) in addressing these issues and furthering the goal of using the available body part as input. These models allow us to track the human body with good accuracy using 2D images and with the benefit of easy integration with a range of applications and devices namely desktop, camera, web, mobile as well as interactive systems.

LITERATURE REVIEW

Whether they are made to entertain you, or to educate you, good video games engage you (Birk, 2013). It has been understood that engagement in games can be measured using player experience (PX). Traditionally, PX evaluation has focused on the enjoyment of the game, or the motivation of players; these factors no doubt contribute to engagement, but decisions regarding play environment, namely the choice of game controller affect the player more deeply than that.