A computer-vision based training coach for computerized physical training

George Davies[0009−0008−4132−5676]

University of Lincoln, UK 27421138@students.lincoln.ac.uk

1 Introduction

In many stories of science-fiction, there are robots that are intelligent, borderline human in the way they talk, the way they walk, and the way they interact with their environments. These kinds of robots seem to be creatures that won't exist any time soon, and they probably won't, but with every advancement in the field of robotics and autonomous systems, we step closer to that reality. One key sense that helps a robot communicate and understand its environment is sight, and the area of robotics that strives towards the goal of giving machines the sense of sight is computer vision.

1.1 Area of Research

Human Pose Estimation (HPE) is an area of research within computer vision that aims to teach robots how to make sense of the human form and the motions it is capable of performing. It involves the identification and classification of the joints in the human body, capturing a set of coordinates for each joint, known as a key point, that can describe the pose of a person. HPE has a wide set of uses in many fields: In games, with motion capture technologies reliant on HPE, it allows developers to code program more realistic and fluid character movements. In healthcare, healthcare providers can monitor a patient's movements and detect any abnormalities. Augmented reality, allows the user to interact with the digital content in more natural and intuitive ways with gestures. And finally, the use-case that is the primary focus of this project, is sports training. HPE can be used to analyse a user's performance, identify areas for improvement, and develop personalised training programs based on the physical level of the user. For example, HPE could be used to analyse a runner's form, e.g. How straight is their back? What part of the foot they are landing on? Are they leaning more to one side?..., and providing feedback on how to improve their technique. HPE can be used to collect data about any exercises where the movement of the body is vital to its effectiveness.

1.2 Relevance to the Degree Programme

I am enrolled in the MSc programme Robotics and Autonomous Systems at the Univerity of Lincoln through the AgriFoRwArdS CDT. Throughout the first

two semesters, I studied the principles of robotics, artificial intelligence, machine learning, and computer vision. Principles from all of these subjects are applied within the area of research on human pose estimation. As to my affiliation with the AgriFoRwArdS CDT, their focus is on the production and use of AI, ML, and CV applications to help the farming and agricultural industries, as Lincolnshire is an agricultural region of the UK. Human pose estimation has previously been used in agritech applications (Moysiadis et al. 2022), with its ability to facilitate human-robot interactions in the field for fruit picking, robotic carts will follow the worker through the field to hold the produce and take it away once full. HPE allows these robots to understand gesture commands the worker may give it, and gives the robot an understanding of humans that allows it to find and follow them without driving into them.

1.3 Background of the Topic

When computer vision gained popularity in the late 1960s and early 1970s, HPE research had its start. Scientists first focused on basic problems like as shape analysis, object recognition, and visual understanding. As computer vision developed, HPE became a stand-alone area of study (Lynn 2023). Historically, HPE was frequently described probabilistically to account for likely inference ambiguities. Since deep learning has been more widely used, the focus has switched to end-to-end trainable models because of their ability to extract intricate patterns and postures from data. Traditionally, computer vision systems have assessed an object's or person's posture by geometric calculations and feature-based techniques. But, the biggest developments in HPE came with the advent of deep neural networks, convolutional neural networks, and computer vision. The field has advanced considerably in spite of these challenges, and more recent techniques that make use of properly designed neural networks may provide amazing results in challenging scenarios involving a large number of, perhaps veiled, interacting individuals (Liu and Yuan 2018). Now that these detections have the necessary technology and are sufficiently precise, they may be employed for commercial purposes. It also offers a wealth of new application potential and signifies a major change in HPE's overall direction.

References

Liu, Mengyuan and Junsong Yuan (2018). "Recognizing human actions as the evolution of pose estimation maps". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1159–1168.

Lynn, Trevor (July 2023). *Pose Estimation Algorithms: History and Evolution.* Roboflow Blog. URL: https://blog.roboflow.com/pose-estimation-algorithms-history/.

Moysiadis, Vasileios et al. (2022). "An Integrated Real-Time Hand Gesture Recognition Framework for Human–Robot Interaction in Agriculture". In: *Applied Sciences* 12.16. ISSN: 2076-3417. DOI: 10.3390/app12168160. URL: https://www.mdpi.com/2076-3417/12/16/8160.

Word Count: 765