INFORMATICS INSTITUTE OF TECHNOLOGY

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Computer Vision Based Motion Capture System

A dissertation by

Mr. Donmanuwelge Alwis

(W1714886 | 2018088)

Supervised by

Mr. Pumudu Fernando

Submitted in partial fulfillment of the requirements for the BEng (Hons) Software Engineering degree at the University of Westminster.

April 202

Abstract

The current state-of-the-art motion capture systems utilize markers placed on actors, tracked by either cameras or wearable sensors. However, this technique has several drawbacks, including the high cost of hardware, the use of bulky equipment that can be difficult to install and operate, the need for specialized software that can be expensive and complicated to use, as well as the requirement for extensive calibration and post-processing. As a result, the need for affordable and efficient alternatives to motion capture has become increasingly important.

This project aims to address the above-mentioned limitations by developing a motion capture system that uses 3D human pose estimation based on a deep learning algorithm. The proposed system predicts the joint positions of human subjects in 3D space from 2D video footage. To make the system accessible to a broader audience, a user-friendly interface is introduced, allowing users to upload 2D videos and generate 3D animations of human poses. The system employs an ensemble model architecture that contributes to the 3D human pose estimation domain and provides a low-cost and efficient alternative to traditional motion capture techniques that require specialized equipment.

To train the model and evaluate the system's performance, Human3.6M dataset is used, and popular evaluation metrics such as MPJPE are employed. The testing results show promising outcomes and prove the effectiveness of the proposed system in accurately estimating the 3D human pose while overcoming the existing limitations of traditional motion capture systems. Overall, this project provides a significant contribution to the body of knowledge in the field of motion capture and 3D human pose estimation by introducing a cost-effective and efficient alternative to the existing systems.

Keywords: Motion capture, 3D human pose estimation, Machine learning

Subject Descriptors: I.3.7: Three-Dimensional Graphics and Realism, I.4: IMAGE PROCESSING AND COMPUTER VISION, I.4.8: Scene Analysis, I.4.9: Applications, J.6: COMPUTER-AIDED ENGINEERING