THE CHALLENGES OF COACHING USING 2D GOLF SWING VIDEO DATA COMPARED TO THE CHALLENGES OF BUILDING A 3D TECHNOLOGY-BASED COACH.

Tony Luczak[*1], Kait Jackson[2], Michael Neff[2], Long Tian[1], Adam Jones[1], Reuben Burch[1], Jonathan Barlow[1], Patrick Nelsen[1], Caleb Morgan[1], Harish Chander[1], Martin Duclos[1], Michael Taquino[1], and Steven M. Grice[1]

[1]Mississippi State University

[2]Gears Sports

[*]email: tluczak@nsparc.msstate.edu

Introduction

Golf instructors use a variety of tools to perceive 3D motion; perceptual mental models acquired from years of teaching experience, two-dimensional (2D) video apps (V-1 Sports), optical motion capture system (mocap; Gears Sports), and wearable sensor systems (4-D Golf, K-Vest). Instructors annotating 2D video to communicate desired alterations to a golfer's swing pattern has become ubiquitous due to smart devices. 2D video applications are used to identify key points on the body and club to infer three-dimensional (3D) human pose estimation (HPE) and club trajectories during the swing. However, there are many challenges to inferring 3D space from 2D videos. These challenges include: (a) depth ambiguity, (b) occlusion, (c) convergence, and (d) parallax, leading to incorrect assumptions about a golf swing. Mocap is considered the gold standard in capturing 3D motion and offers an insight into a golfer's performance that cannot be achieved with 2D video solutions. However, there are some limitations to mocap that prohibit instructors from using it, including: (a) initial cost, (b) difficulties while using in bright sunlight, and (c) marker stability. Hard wearable sensor technology is comprised of inertial measurement units (IMUs), consist of accelerometers to derive acceleration, gyroscopes to derive angular rotation, and magnetometers to assist in location, have their own inherent challenges including drift, placement, and sensor management. This survey illustrates the technical challenges involved in developing a golf specific 3D analysis tool and proposes a work-in-progress hybrid approach that combines machine learning, computer vision, and hard and soft sensor technology that will enable human motion capture comparable to optical motion capture systems.

Methods

From 2014 to 2022, there is a paucity of peer reviewed publications regarding HPE, wearable sensors, and golf. In a computerized literature survey using EBSCO discovery service at Mississippi State University (MSU) with key words "human pose estimation" AND "golf" produced 11 results, with eight peer-reviewed articles chosen for review after removing duplications, theses, and dissertations. Exploring "golf" AND "wearable sensors" using the EBSCO discovery service at MSU, produced a total of 37 peer-reviewed publications, after removing duplicates, theses, dissertations, and non-golf and body papers, five articles were used for review. Furthermore, under Mississippi State University's Institutional Review Board study #21-258, researchers have collected outdoor 2D video and 3D optical motion capture data of 29 golfers (2 females, 27 males) hitting 5 balls each with a driver, 7-iron, and wedge. 2D videos were captured using three Apple devices: an iPhone X and two iPad Pros, second generation. One iPad Pro was positioned perpendicular to the golfer, with the additional devices positioned approximately 25 degrees to each side of the center device and set to record at 240 frames per second. 3D data was captured using Motive Software (OptiTrack) Motion Capture Solution.

Results and Discussion

Survey results identified five key challenges in both golf HPE networks and wearable sensors. Challenges in HPE include: (a) the manual labeling of human skeletal keypoints, (b) lack of high-speed training data, (c) depth ambiguity, (d) occlusion, and (e) background clutter. There is lack of golf specific datasets due to the time and cost of acquiring, manually annotating, and reviewing the required high volume of videos for machine learning. Annotation inaccuracies are due to different anthropometrics, difficulty identifying joints under clothing, occlusions, camera angle, motion blur, and depth ambiguity. Occlusion is prevalent in the golf swing due to self-occlusion of limb segments covering other joints. Challenges in hard wearable sensors include: (a) too much noise in the data, (b) the number and placement of sensors to attach, (c) difficulties in managing the sensors, (d) inertial measurement unit (IMU) drift, and (e) repeated calibration steps from sensor movement. To overcome these challenges, the development of a hybrid approach combines stereo vision HPE with soft sensors which reduces the number of IMU sensors required to capture the full body and club and improves two measurement features: wrist movements and ground pressure reaction, that are major challenges for HPE models. To optimize the adoption and diffusion of this hybrid approach, understanding and managing the instructor's expectations and experiences is necessary for overall satisfaction, word of mouth recommendations, and repeated usage.

Significance

The benefits of understanding the difficulties and limitations in developing a 3D golf swing capture system will reduce the time and cost of research, and lead developers to generate cost-effective artificial intelligence (AI) solutions. These AI solutions can help golf instructors improve their coaching techniques by acquiring and understanding human movement in 3D space. Moving from 2D to 3D kinematic and kinetic assessments will change how golf instruction is delivered, how instructors establish individualized models of performance, and provide insight to movement patterns that may mitigate injuries.