Markerless Human Motion Tracking Using Microsoft Kinect SDK and Inverse Kinematics

Alireza Bilesan¹, Saeed Behzadipour², Teppei Tsujita³, Shunsuke Komizunai¹, Atsushi Konno¹

Abstract—Motion capture systems are used to gauge the kinematic features of the motion in numerous fields of research. Despite superb accuracy performance, the commercial systems are costly and difficult to use. To solve these issues, Kinect has been proposed as a low-priced markerless motion capture sensor, and its accuracy has been assessed using previous motion capture systems. However, in many of these studies, the anatomical joint angles captured using the Kinect are compared to the 3D rotation angles reported by the gold standard motion capture systems. These incompatibilities in the determination of the human joint angles can lead to higher error estimation. To accomplish a valid accuracy evaluation of the Kinect, we applied the inverse kinematics techniques in both Vicon and Kinect version 2 skeleton models to estimate lower body joint angles. The proposed method enabled us to capture the pelvic, hip, and knee joint angles using a single Kinect camera during gait. Moreover, the dependency of the proposed method to the position of the Kinect and the speed of the moving subject was investigated. In this study, the captured data of the Vicon motion capture system were used as ground-truth to assess the accuracy of the Kinect data. The results indicate the capability of Kinect in capturing human joint angles and also an affordable motion capture system applied in robotics and biomechanics applications.

I. INTRODUCTION

Motion capture systems are utilized to reconstruct and transfer human motions into a robot. Many researchers have employed human motions in robot imitation learning and human-like motion generation [1][2][3]. For a better mimicry, human and humanoid walking patterns are compared to apply the human walking functions to the humanoid robots [4][5]. Despite high accuracy, commercial motion capture systems are costly and complicated to use. Since several cameras are required to capture one motion, the data collecting is restricted to special settings and conditions. For instance, performing multi-camera calibration is essential before each experiment. Additionally, due to the mentioned conditions, commercial systems are mainly used in indoor environments. Markerless motion capture systems were proposed to overcome the previous issues [6]. These motion analysis technologies enabled researchers to evaluate movement characteristics as more cost-effective and straightforward. Despite the benefits of these new systems, systematic limitations restrain their functionality. For instance, wearable electromagnetic sensors are affected by gravity noise and signal drift [7]. Furthermore, these sensors are still costly and require a skillful data analyzer to post-process the data.

Microsoft released Kinect version 1 (Kinect v1) as an accessory for the Xbox 360 video game platform in 2010. It was designed for the gaming purposes, but it can also be utilized as a markerless, affordable, and portable motion capture sensor. The Kinect v1 consists of one IR emitter, one IR camera, and one RGB camera which acquire depth and color images of the scene. Consequently, The Depth of the scene is measured using speckle pattern technology. In 2014, Microsoft released the second version of the Kinect (Kinect v2) with enhanced RGB and IR camera resolution and wider field of view (see Fig. 1). Microsoft used a different technology called time-of-flight (TOF) in Kinect v2 in order to measure the depth of the scene [8]. The TOF technology assisted the use of the Kinect v2 in outdoor

---

¹Alireza Bilesan, Shunsuke Komizunai, Atsushi Konno are with the Graduate School of Information Science and Technology, Hokkaido University, Kita 19 Nishi 9, Kita-ku, Sapporo, Japan bilesan@sc.iist.hokudai.ac.jp
²Saeed Behzadipour is with the Mechanical Engineering Department of Sharif University of Technology and Dajad Mowadaghian Research Center in Neurorehabilitation Technologies, Tehran, Iran behzadipour@sharif.edu
³Teppei Tsujita is with the Department of Mechanical Engineering, National Defense Academy of Japan, Yokosuka, Japan tsujita@nda.ac.jp

---

Fig. 1. Human motion tracking using Kinect v2. The Kinect accuracy is assessed using the Vicon motion capture system as the gold standard. The Kinect world coordinate is located on its IR camera, and the Vicon world coordinate is set on the floor. The plug-in-gait marker set is only used in the Vicon system in order to capture the human motion.
As it is illustrated in Fig. 6, the Kinect data is noisy and unstable, which becomes more important in the forward dynamic simulation, where the second order differential of the joint trajectories are required. The results affirm more research is required to overcome the instability of the Kinect data to validate this sensor in robot control tasks.

REFERENCES


