Applied Ergonomics 49 (2015) 47–54

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Technical note

The validity of the first and second generation Microsoft Kinect™ for identifying joint center locations during static postures

Xu Xu*, Raymond W. McGorry

Liberty Mutual Research Institute for Safety, 71 Frankland Road, Hopkinton, MA 01748, USA

A R T I C L E I N F O

Article history:
Received 10 July 2014
Accepted 23 January 2015
Available online 17 February 2015

Keywords:
Reference frame alignment
Daily activities
Kinect v2

A B S T R A C T

The Kinect™ sensor released by Microsoft is a low-cost, portable, and marker-less motion tracking system for the video game industry. Since the first generation Kinect sensor was released in 2010, many studies have been conducted to examine the validity of this sensor when used to measure body movement in different research areas. In 2014, Microsoft released the computer-used second generation Kinect sensor with a better resolution for the depth sensor. However, very few studies have performed a direct comparison between all the Kinect sensor-identified joint center locations and their corresponding motion tracking system-identified counterparts, the result of which may provide some insight into the error of the Kinect-identified segment length, joint angles, as well as the feasibility of adapting inverse dynamics to Kinect-identified joint centers. The purpose of the current study is to first propose a method to align the coordinate system of the Kinect sensor with respect to the global coordinate system of a motion tracking system, and then to examine the accuracy of the Kinect sensor-identified coordinates of joint locations during 8 standing and 8 sitting postures of daily activities. The results indicate the proposed alignment method can effectively align the Kinect sensor with respect to the motion tracking system. The accuracy level of the Kinect-identified joint center location is posture-dependent and joint-dependent. For upright standing posture, the average error across all the participants and all Kinect-identified joint centers is 76 mm and 87 mm for the first and second generation Kinect sensor, respectively. In general, standing postures can be identified with better accuracy than sitting postures, and the identification accuracy of the joints of the upper extremities is better than for the lower extremities. This result may provide some information regarding the feasibility of using the Kinect sensor in future studies.

1. Introduction

In 2010, Microsoft released the first generation Kinect sensor (Kinect v1) for their Xbox 360 video game platform. This sensor is composed of one traditional digital camera, one infrared emitter, and one depth sensor. The depth sensor is able to detect the depth information of the objects placed in front of the Kinect unit. When a person stands directly in front of a Kinect sensor (Fig. 1a), the Kinect sensor will first detect the surface of body with the depth sensor (Fig. 1b) and then uses an algorithm to automatically identify the location of 20 joint centers of the body in the 3-D space from the surface of the human body (Fig. 1c). In mid-2014, Microsoft released the computer-used second generation Kinect (Kinect v2) with improved resolution for the traditional camera and the depth sensor. The number of the identified joint centers for Kinect v2 also increased to 25.

Compared with traditional opto-electronic or electromagnetic-based motion tracking systems, the Kinect sensor is very low-cost, portable, and without skin markers. Several studies have compared the Kinect-based and motion tracking system-based posture and movement, in various fields of biomechanics including; anthropometry measurement (Bonnechère et al., 2014), spinal loading assessment (Ning and Guo, 2013), clinical foot posture assessment (Mentiplay et al., 2013), body movement during postural control (Clark et al., 2012), gait training (Clark et al., 2013b), gait analysis (Clark et al., 2013a), and rehabilitation tool development (Pastor et al., 2012). In general, the results of these studies indicated that, although the Kinect sensor is not as accurate as more traditional laboratory-based measurement technologies, it does provide good agreement with a motion tracking system in terms of body segment lengths, the major joint angles, and the displacement of the key body joints for those specific body postures being tested.

* Corresponding author. Tel.: +1 508 497 0218.
E-mail address: Xu.Xu@libertymutual.com (X. Xu).

http://dx.doi.org/10.1016/j.apergo.2015.01.005
0003-6870/

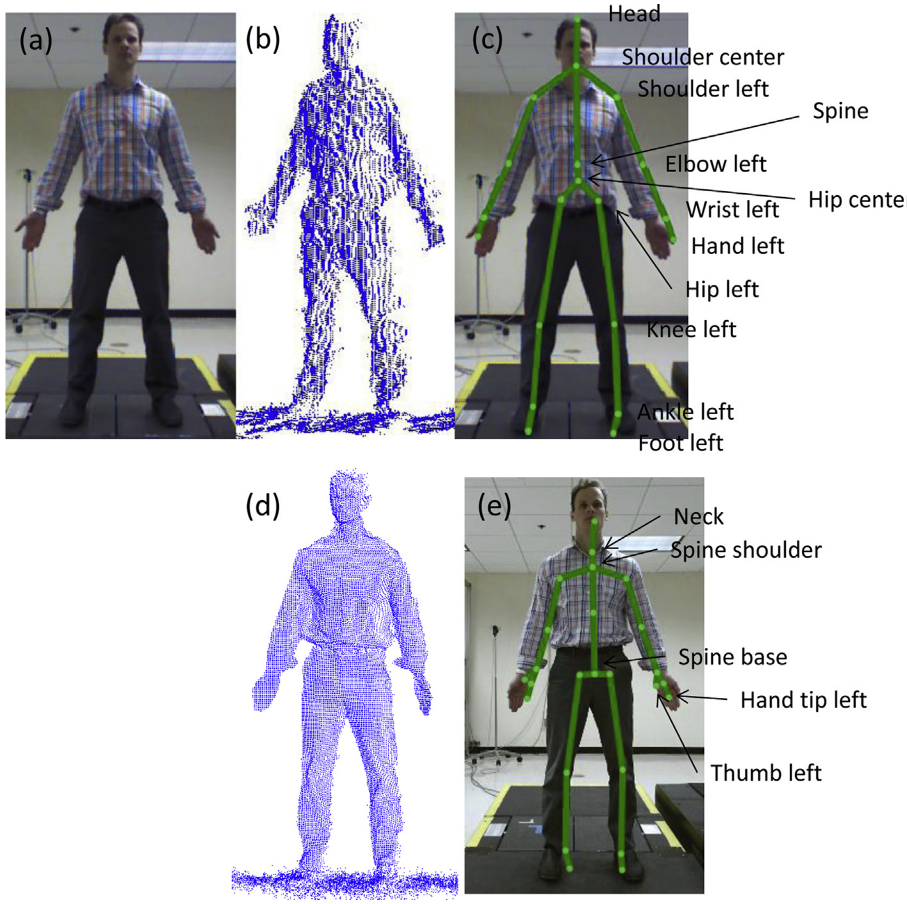

Fig. 1. (a) A person standing in front of a Microsoft Kinect™ (Model 1517) sensor. (b) The body surface detected by Kinect v1. (c) The joint centers identified by Kinect v1. (d) The body surface detected by Kinect v2. (e) The joint centers identified by Kinect v2. Only the joint centers not identified by Kinect v1 were listed in this sub-figure.

While the body segment length, the joint angles, and the joint displacements can be derived from the Kinect sensor-identified joints coordinates, very few studies have performed a direct comparison between the locations of the Kinect sensor-identified joint centers and their corresponding motion tracking system-identified counterparts. This is probably because the Kinect sensor-identified joints location are, with respect to a coordinate system, fixed with the Kinect sensor, whose orientation is normally unknown within the global coordinate system (GCS) of a motion tracking system. However, the error of Kinect sensor-identified joint center locations can provide more insight to the composition of the error in segment length, joint angles, and joint displacement. In addition, the joint center positions can be important when calculating the joint net moments using the inverse dynamic methods (Faber et al., 2009; Hashish et al., 2014). The goal of the current study is to examine the accuracy of the Kinect sensor-based joint center coordinates during various static postures of daily activities for Kinect v1 and v2. To achieve this goal, a method to align the coordinate system of the Kinect sensor with respect to the GCS of a motion tracking system is also proposed.

2. Method

2.1. Alignment of the coordinate system of a Kinect sensor with respect to global coordinate system (GCS) of the motion tracking system

The Kinect sensor and the motion tracking system record the coordinates of the joint centers each relative to their own GCS. Therefore, it is first necessary to identify the vector coordinates (x, y, and z) of the origin of the GCS of the Kinect (K) sensor with respect to the GCS of the motion tracking system (MT) in 3-D space, ${}^{\mathrm{MT}}\mathbf{O}_{\mathrm{K}}(x,y,z)$, and the 3-by-3 rotation matrix with the Euler angle ($\alpha,\beta,$ and γ) from the GCS of the Kinect sensor to the GCS of the motion tracking system, ${}^{\mathrm{MT}}\mathbf{R}(\alpha,\beta,\gamma)^{K}$, before a direct comparison can be made between the joint locations identified by the two systems.

To achieve this, a wooden wheel was built with spokes eight 2.5 cm wide and 70 cm long (Fig. 2a). Each spoke had a cross near the tip. Customized software using Kinect for Windows SDK 1.5 (for Kinect v1) or SDK 2.0 preview (for Kinect v2) was used to acquire the depth information of the scanner surface of the wooden wheel (Fig. 2b). This wheel was placed 2.0 m and 2.5 m, correspondingly, in front of a Kinect sensor (Model 1517 for Kinect v1, Model 1656 for Kinect v2, Microsoft Corporation, Redmond, WA, U.S.A). The scanned surface consisted of numerous points, and an experimenter was required to select the center point of the cross of each spoke. At the same time, the marked surface center of each cross was also digitized with respect to a motion tracking system (Optotrak Certus System, Northern Digital, Canada). There were a total of 16 sampling points (8 crosses $\times$ 2 location of the wheel) on the wooden wheel digitized with respect to the GCS of the Kinect sensor and the GCS of the motion tracking system.

The origin vector ${}^{\mathrm{MT}}\mathbf{O}_{K}(x,y,z)$ and the rotation matrix ${}^{\mathrm{MT}}\mathbf{R}(\alpha,\beta,\gamma)^{K}$ of the GCS of the Kinect sensor were then derived by minimizing the average distance between the motion tracking system-based location of the 16 points and the Kinect sensor-based