Validation of the Kinect System for Industrial Motion Analysis

K. Sung\textsuperscript{1}, J. Chang\textsuperscript{1}, Z. Ridall\textsuperscript{1}, H. Kalkis\textsuperscript{2}, C. Tucker\textsuperscript{1}, Z. Roja\textsuperscript{2}, and A. Freivalds\textsuperscript{1}

\textit{1Pennsylvania State University, University Park, PA 16802 USA}
\textit{2The University of Latvia, Riga, Latvia}

1. Introduction

Commercial, off-the-shelf motion sensor systems such as the Microsoft Kinect have been employed across a wide range of domains to capture human body motion in a relatively timely and efficient manner. For example, Celebi et al. have utilized the skeletal joint data generated by the MS Kinect to model human body gestures (Celebi et al., 2013). Dance gestures have also been modelled using data acquired from the MS Kinect with accuracies as high as 96.9% (Raptis et al., 2011). Manohar and Tucker utilized joint data, generated by the MS Kinect to predict the emergence of human threats in an audience (Manohar and Tucker, 2013). Such techniques illustrate the potential of capturing and mining human body language data using commercial, off-the-shelf systems such as the MS Kinect. The authors of this work explore the feasibility of utilizing the MS Kinect system to capture and model industrial motions, towards real time assessment of task performance.

2. Method

Three typical industrial tasks were studied. These basically simulated Gilbreth’s second and third lowest level classification of motions (Freivalds, 2014): 1) hand tapping, one degree of freedom: flexion/extension of the wrist, a motion also commonly used to assess motor control or neurological disorders, 2) hammering using the wrist, one degree of freedom: radial/ulnar deviations (See Fig. 1a), 3) hammering with forearm, one degree of freedom: flexion/extension of the elbow (See Fig. 1b). Four different frequencies were utilized, roughly corresponding to recommended work tempos based on folk norms (Drillis, 1963), but which also correspond to the natural frequencies of that joint/limb system (Freivalds, 2011): 1) 180/min, tempo for wrist tapping, 2) 60/min, tempo for hammering with wrist motion, 3) 30/min, tempo for hammering with elbow motion, and 4) 15/min, corresponding approximately to a tempo for whole-body motion similar to shovelling.

However, to fully test the Kinect system capabilities, all frequencies were applied to all motions, resulting in 3x4=12 conditions, with two replications, resulting in a total of 24 trials (performed randomly). Each trial lasted approximately 12 seconds, collecting from 3 to 36 complete motion cycles. Penny&Giles goniometers were applied to each joint to serve as the ‘gold-standard’ measure of angular motion. It was calibrated prior to the experiment. The resulting angles of motion obtained from the Kinect system were then correlated with the angles obtained from the goniometer. Ten subjects, facing the Kinect system, were tested.

![Figure 1. a) Hammering with wrist radial/ulnar deviations, b) Hammering with elbow flexion.](image-url)
3. Results

Range of motions (in degrees) for one 30-year old male (1.85m, 75 kg) are given Table 1. Correlation of the Kinect and goniometer range of motion angles for all conditions yield a very poor negative correlation of -0.251. This would be very unusual and visual analysis indicates very disparate Kinect readings for the wrist radial/ulnar deviation data. Once that data is removed, the correlation improves to a statistically significant (p<.02) value 0.81.

Table 1. Range of Motion Data

<table>
<thead>
<tr>
<th></th>
<th>Elbow</th>
<th>Wrist: Radial/Ulnar</th>
<th>Wrist: Flexion/Extension</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Kinect</td>
<td>24.7</td>
<td>22.1</td>
<td>16.4</td>
</tr>
<tr>
<td>Gonio</td>
<td>36.2</td>
<td>23.3</td>
<td>19.1</td>
</tr>
</tbody>
</table>

4. Discussion

Preliminary analyses indicate that the Kinect system does well in recording large motions such as elbow flexion and extension, a 3rd level Gilbreth motion and also for wrist flexion and extension, a 2nd level Gilbreth motion. However, wrist radial/ulnar deviation is handled very poorly. This deviation at the 2nd level needs to be examined in further detail. Perhaps, different data filters will improve the accuracy. Overall, the Kinect system shows promise for motion capture.

References


