# Measuring Sway with Markerless Depth Camera

Kevin Huang, Eyal Ofek, Ran Gilad-Bachrach

*Abstract*—**The goal of this study is to present methods to measure properties of posture, specifically sway, using ordinary depth cameras. Current methods either use markers which require accurate placement of reflectors or use virtual skeletons which may fail for people with atypical body structure. We introduce algorithms to extract sway metrics from the sensor which do not require markers or virtual skeleton extraction. Three experiments were conducted to measure the proposed algorithm: (1) Measurements on a mannequin connected to a robotic arm were used to measure sensitivity and repeatability (2) Measurements using 20 healthy subjects were used to compare against measurements by a pressure sensor (3) Measurements using 13 healthy subjects were used to verify that subjects pose is within the sensitivity of the sensor. Results show that frequency based metrics and axes invariant metrics were repeatable and showed little sensitivity to the positioning of the sensor. Axes based metrics (lateral-ventral) showed sensitivity to the viewing angle of the sensor when the angle was $> 5°$. When testing on human subjects results indicate that many of the measurements from the depth camera are correlated well with measurements of a pressure sensor (Pearson correlation $\geq 0.65$) and are more accurate that results obtained by using the virtual skeleton. Moreover, it is sufficiently accurate to distinguish between the sway imposed by different postures, such as standing with feet in tandem vs feet parallel (p-value $\sim 0.05 - 0.005$). We use inter-class correlation coefficient to measure the reliability and found it to be $\geq 0.5$ for most metrics. Last, we show that although the sensor is sensitive to the viewing angle, an acceptable viewing angle ($0.55° \pm 1.89°$) is achieved when subjects are asked to stand behind a line marked on the floor. These experiments show that methods to measure sway with markerless depth cameras without using virtual skeleton extraction are reliable and outperform methods that use the virtual skeleton. An implementation of the methods is available at https://github.com/Microsoft/GaitAndBalanceApp.**

## I. Introduction

Sway is correlated with elevated risk of fall [13, 24] and other medical conditions such as Diabetes [27], Parkinson's disease [2], and Multiple Sclerosis [20]. Therefore, measuring sway became an active area of research. Force plates are commonly used for measuring balance [5], however, these devices tend to be expensive and therefore not accessible. Studies showed that even force plates that were designed for the gaming industry, such as the Wii balance board (Nintendo, Kyoto, Japan) can deliver valid and reliable results [6]. These devices are very affordable and accurate. However, they can only measure the Center Of Pressure (COP) and hence are limited in the information they provide. For example, these devices will not be able to detect forward or backward slanting postures. They also cannot measure compensatory movements such ankle and hip strategies. Moreover, the Wii board has a small area which is a limitation for people of poor stability.

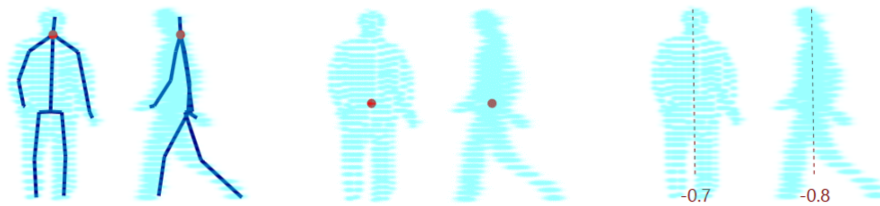
In recent years, several studies suggested the use of inertial devices [15, 16] which are are affordable and accurate. These studies are very encouraging in their results, however, similar to force plates, they are limited to measuring only the few point on the body on which sensors are placed. Inertial devices are sensitive to the inaccurate placement of the device and to potential displacement of the sensor while the subject performs the exercise. Moreover, they require mounting the device close to the body and therefore are somewhat intrusive and cannot support implicit measurements. They also require maintenance in the form of charging batteries and remounting.

Motion capture devices, such as the one developed by Vicon (Oxford, UK), can capture full body information. These devices, are expensive and require careful and stable placement of reflectors (markers) on the subject which make them intrusive. Moreover, they are subject to measurement errors due to marker placement noise. The introduction of the Kinect sensor (Microsoft, Redmond, USA), led several researchers to study its uses for measuring gait and balance [4, 7, 8, 11, 12, 14, 18, 19, 22, 23, 25, 26, 28, 29] especially since it had the a Virtual Skeleton Extraction (VSE) software without requiring the use of reflectors. Today, several affordable depth sensors are available including Intel's RealSense line of depth camera (Intel, Santa Clara, USA), and Orbbec's Astra Cameras (Orbbec, Seattle, USA). In many cases, the providers also include VSEs in the Software Development Kit (SDK). However, VSEs are designed to work for people with "typical" body structure [21] and may fail for people with atypical body shapes. For example, the VSE of the Kinect may fail to detect the skeleton for amputees [1]. VSEs may also have limited accuracy when subjects wear baggy clothes [3]. Therefore, there is an open question whether these depth cameras can be used to extract important posture information without the skeleton information.

To study this question, we have used the Kinect V2 for XBOX One sensor to compare posture measurement methods that use the skeleton with ones that do not use it. We measure the validity and reliability of different methods for measuring metrics of sway. Our experiments focus on metrics which are aggregates over an exercise, as opposed to measuring the temporal accuracy as done by Clark and others [8].

Kinect's depth sensor data is accessible by a software SDK (Kinect for Windows version 2.0). Several streams of data are made available that are relevant to our study: (i) a depth image (ii) a mask around the subject (iii) a virtual skeleton. The virtual skeleton marks the estimated 3D location of 25 joints. We compare methods that use this virtual skeleton to methods that use only the depth image since such methods can be replicated to other depth sensors. To extract relevant sway metrics each method first computes a Reference Point (RP) from each frame which is a 3D central position of the subject. From this RP different sway metrics are computed.

Figure 1. A demonstration of the different method to compute a reference point shown both in a frontal view and a side view. The methods, left to right are: using the neck joint of the virtual skeleton, using the mean point of the body mask, and using the line model. The line model allows also capturing the tilt of the body, in degrees. This tilt is marked at the bottom of the figure.




The Center Of Mass (COM) would have been the ideal reference point however the Kinect sensor captures only a frontal view of the body from which the center of mass can only be approximated. Eltoukhy and co-authors in a recent work [10] tried to validate the Kinect sensor for sway analysis. They also used the Kinect sensor without the virtual skeleton extraction mechanism. However, they did try to estimate the COM and use it to measure sway features. In order to do that, they had to use markers on the subject while the method we propose here does not use any markers. Moreover, Eltoukhy et al. did not test the sensitivity of the approach to placement of the sensor.

Instead of directly approximating the COM our method extracts a Reference Point (RP) and tracks it during the exercise. We test 3 different RPs:

1) **Neck:** In this method, we use the skeleton inferred by the Kinect SDK, specifically the position of the neck joint, to track the movement of the body
2) **Mean:** In this method we Look at the measured depth map of the person, and find the 3D mean point of all the visible 3D points of the person as indicated by the subject's mask.
3) **Line:** In this method we generate a line that best fits the visible surface of the subject using Ordinary Least Squares (OLS). We use the point on this line at height 1meter as the reference point. The line method also gives angle with respect to the floor which is used to estimate slanting postures.

The different methods are demonstrated in Figure 1.

## II. Methods

We have conducted three experiments to test the accuracy, sensitivity, and reliability of the different approaches. In the first experiment, a robotic arm controlled a mannequin moving it in predesigned trajectories. This experiment was conducted to measure accuracy, reliability, and the sensitivity to different viewing angle. In the second experiment, 20 subjects were measured while standing in different postures. The measurements were compared to measurements taken by a force plate. Since the robotic arm experiment showed that some measurements are sensitive to the viewing angle, a third experiment was conducted in which subjects were asked to stand behind a line marked on the floor and measurements were taken of the angle between subjects and the line.

### A. Experiment One - Robotic Arm

In order to measure the accuracy in which the Kinect can capture the different metrics as well as possible artifacts of the viewing angle we conducted an experiment using a mannequin that was attached to a robotic arm (see picture in Figure 2 and a schematic top view in Figure 3). This setup allowed us to accurately reproduce the same experiment multiple times while the sensor is maintained in the same position (to measure repeatability) or moved to a different position (to measure sensitivity to the viewing angle). This setup also allowed us to measure how well the measurements isolate between different aspects of the movement such as separating the frequency of the sway from its magnitude.

Figure 2. A picture of the experimental setup. The robotic arm, on the right side of the image, is attached to the mannequin. The feet of the mannequin touch the ground souch that when the robotic arm moves the mannequin tilts. The Kinect sensor, seen on the left side of the image, is positioned in front of the mannequin.

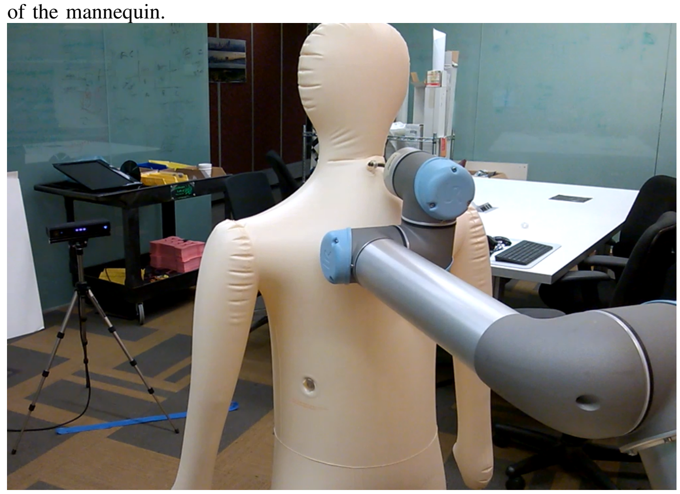
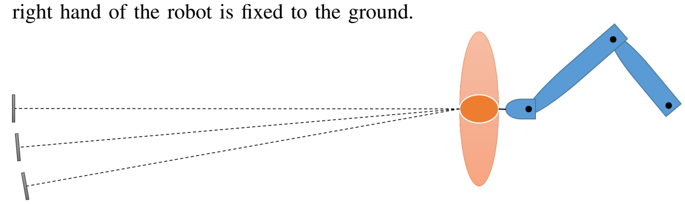


Figure 3. Schematic top view of experiment 1 setup. The robotic arm (blue on right) was attached to the mannequin (orange). The Kinect sensor (gray on left) was positioned 2m away from the mannequin in 3 viewing angles of $0°, 5°$, and $10°$ from the tangent to the lateral axis of the mannequin. The right hand of the robot is fixed to the ground.



*1) The experimental setup:* An inflatable mannequin[1] was attached to a robotic arm[2]. The robot is specified to have a repeatability of $\pm 0.1\mathbf{mm}$ and therefore provides an accurate ground truth. The Kinect was mounted on a tripod at a distance of 2m away from the mannequin. We tested 3 viewing angle in which the sensor is positioned at $0°$, $5°$, and $10°$ of the tangent to the lateral axis of the mannequin as illustrated in Figure 3.

*2) Trajectories:* We have programmed the robot to perform 3 different trajectories. In the *stationary* trajectory, the robot was not moving such that the mannequin was standing still. In the *move* trajectory, the robot arm was moving in a rectangular shape with edges of sizes 5mm and 10mm. The axis of the rectangle are $45°$ to the main axis of the mannequin. The *big-move* trajectory is similar to the *move* trajectory but the rectangle has edges of sizes 10mm and 20mm.

We have repeated the *move* and *big-move* trajectories with different moving speeds. The speed that were used are 10%, 50% and 100% of the maximal speed that the robot allows ($\sim 1\mathrm{m}/\mathrm{sec}$). Note that the speed effects the time it takes for the robotic arm to move between one point to another. However, the speed does not alter the wait time in each corner of the rectangles defined in the *move* and *big-move* trajectories. As a result when the speed is increased from 10 to 100 the time to complete a cycle increases by $7.5\times$ for the *big-move* trajectory and $9\times$ for the *move* trajectory. To test the repeatability of the measurement, each setup and each speed were recorded 10 times using the Kinect from each viewing angle.

### B. Experiment Two - Human Posture

Next we compare the measurements obtained by the Kinect to measurements of a force plate.

*1) Subjects:* Twenty healthy subjects (13 male, 7 female) in the age range of $32 - 54$ (mean $40.8$, median $42$) participated in the study. All subjects signed a consent form and received $5 gratuity coupon. The study was approved by the ethical review committee of the institute.

*2) Procedure:* Subjects were asked to take off their shoes to reduce variance due to shoe type and asked to stand on a Wii balance board facing a Kinect sensor. The Kinect sensor was placed 2 meters from the center of the Wii balance board on a surface at a height of 71cm and was adjusted to be looking parallel to the ground. The Wii balance board was placed such that its narrow side is parallel to the front of the Kinect sensor.

Each subject was measured in 4 postures as demonstrated in Figure 4 with the following instructions given:

1) In this posture, you will be standing, facing the Kinect sensor. Your feet should be side by side, touching each other. Your hand should be on your sides.
2) In this posture, you will be standing, facing the Kinect sensor. Your feet should be side by side, touching each other. Your hand should be crossed. Try to cross your hands high.
3) In this posture, you will be standing, facing the Kinect sensor. Your feet should be side by side, touching each other. Your hand should be crossed high. In this posture, your eyes should be closed.
4) In this posture, you should be standing with your feet in tandem. The toes of one foot should be touching the heels of the other one. Try to keep your hands crossed high and your eyes open.

In each posture, the subjects were recorded for 2 minutes and were allowed to rest between exercises.

### C. Experiment Three - Posing Verification

Since the Kinect sensor collects only the frontal view of the subject, it may be sensitive to viewing angle. Therefore, we conducted another experiment to measure if subjects stand parallel to the sensor when instructed to do so. For this experiment, 13 subjects (5 females, 8 male) were instructed to stand behind a line marked with a tape on the floor. Each subject was asked to stand twice behind the line (taking a rest between each trial). When standing, the location of subject's feet where photographed by two DSLR cameras and the distances to the line have been recovered from which the angle to the lines was computed.

### D. Tools

Data for the first 2 experiments was acquired using custom made software that collected the data from the Wii board and the Kinect sensor. The data was sampled at the highest rate available by each sensor (100Hz for the Wii board and 30Hz for the Kinect). From the Wii board, the COP was acquired. We used three methods to generate a Reference Point (RP) from the Kinect: (i) The neck point of the skeleton inferred by the Kinect SDK[3] (ii) The average of all the visible 3D points of the subject (iii) The line model was computed by computing the covariance matrix of all the visible 3D points. See Figure 1 for an illustration of the three methods. Reading from the Wii board was done using WiiMoteLib V1.7.[4] Kinect for PC SDK V2.0 was used to access Kinect data.[5] Following Mancini et al. [15], data was filtered using a 3.5Hz low-pass Butterworth filter.

Many ways to measure sway have been proposed in the literature [15, 17, 24]. The metrics we used follow mostly the metrics suggested by Maurer and Peterka [17] with some modifications. In the following definitions $m(\cdot)$ stands for the median, $E[\cdot]$ stands for expectation and $G(f)$ is the power in the frequency $f$ as computed by the Fourier Transform. Note that all the metrics are computed in the transverse plane and thus ignore the cranial caudal axis.

*1) Time domain metrics:*

- **Median X** - The median absolute deviation (MAD) of the position in the lateral-medial axis: $m(x - m(x))$
- **Median Z** - The MAD of the position in the ventral-dorsal axis: $m(z - m(z))$

[1]The robotic arm is rated up-to 3Kg and therefore we were not able to use more rigid mannequins

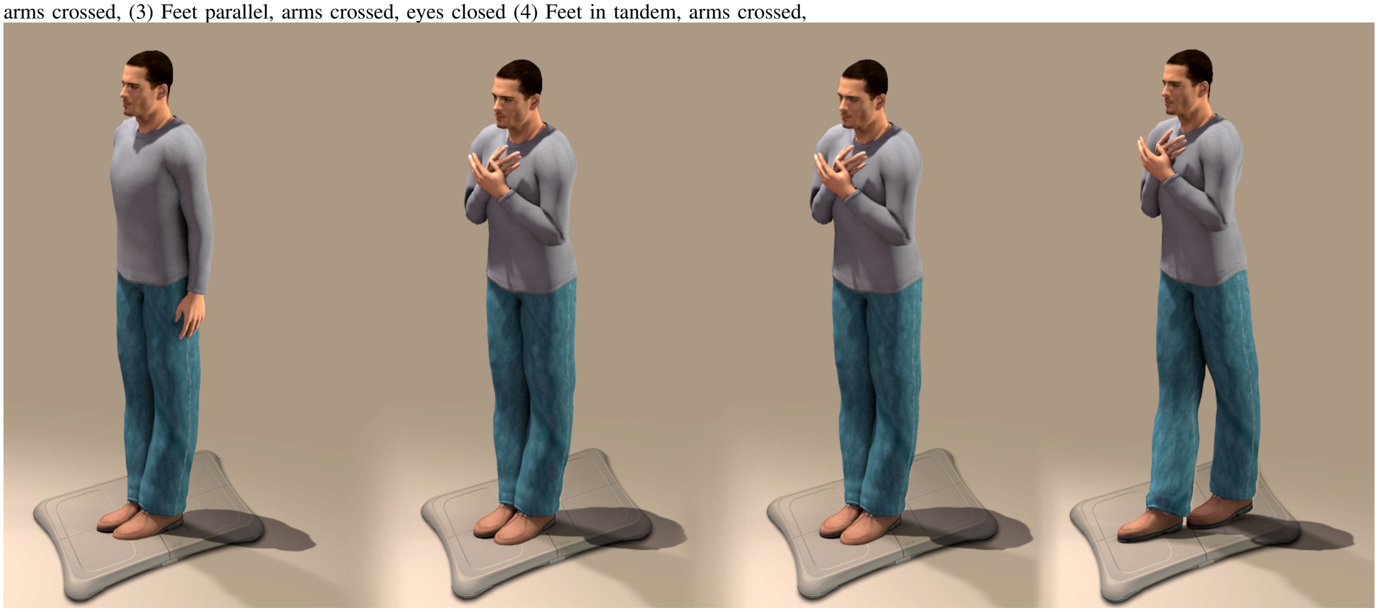
[2]UR3, Universal Robots, Odense, Denmark

[3]We found this point to be the least noisy.

[4]http://wiimotelib.codeplex.com/

[5]http://www.microsoft.com/en-us/kinectforwindows/

Figure 4. Experiment 2 postures: each subject was recorded in 4 different postures (from left to right): (1) Feet parallel, arms to the side, (2) Feet parallel, arms crossed, (3) Feet parallel, arms crossed, eyes closed (4) Feet in tandem, arms crossed,



- **Median angle X** - The median angle of the body in the lateral median axis from the upright axis[6]
- **Median angle Z** - The median angle of the body in the ventral-dorsal axis from the upright axis[6]
- **Median Distance** - The median distance of the position from the median position:

$$m\left(\left\|\left[\begin{smallmatrix}x\\z\end{smallmatrix}\right]\right\| - E\left[\left[\begin{smallmatrix}x\\z\end{smallmatrix}\right]\right]\right)$$

- **RMS** - Root Mean Square:

$$\sqrt{E\left[\left\|\left[\begin{smallmatrix}x\\z\end{smallmatrix}\right] - E\left[\left[\begin{smallmatrix}x\\z\end{smallmatrix}\right]\right]\right\|^2\right]}$$

- **Path** - Total distance:

$$\sum_t \left\|\left[\begin{smallmatrix}x_t\\z_t\end{smallmatrix}\right] - \left[\begin{smallmatrix}x_{t-1}\\z_{t-1}\end{smallmatrix}\right]\right\|$$

- **Area** - The area of the 95% confidence ellipse: $\pi AB$
- **Range A** - The length of the larger axis of the 95% confidence ellipse: $2A$
- **Range B** - The length of the shorter axis of the 95% confidence ellipse: $2B$
- **Median Frequency** - The ratio between the Path and the circumference of the 95% confidence ellipse:

$$\frac{\text{Path}}{T\frac{\pi}{2}\sqrt{2A^2+2B^2}}$$

*2) Frequency domain metrics:*

- **PWR** - The power derived from the frequency signal after applying a high-pass filter: $\sum_f G(f)$
- **F50** - The median frequency:

$$\min\left\{f : \sum_{f'\leq f} G(f) \geq 0.5\text{PWR}\right\}$$

- **F95** - The 95 percentile frequency:

$$\min\left\{f : \sum_{f'\leq f} G(f) \geq 0.95\text{PWR}\right\}$$

- **Centroid Frequency** :

$$\sqrt{\frac{\sum_f f^2 G(f)}{\sum_f G(f)}}$$

- **Frequency Dispersion** - a measure of the variability in the frequency domain:

$$\sqrt{1 - \frac{\left(\sum_f fG(f)\right)^2}{\left(\sum_f G(f)\right)\left(\sum_f f^2 G(f)\right)}}$$

[6]The median angle X and Z are not available on force plates.

## III. Results

### A. Results of Experiment One

Each setup of a trajectory, viewing angle, and speed was recorded 10 times with the Kinect. For each metric, we report the range of values that were measured in the 10 repetition as the center value and the maximal deviation (in percent). Since there are many conditions and many metrics, we break down the discussion by the type of metric used: frequency based, axes based, and axes free. Also note that the Kinect sensor did not recognize the mannequin as a human and therefore, did not generate a skeleton for it. Therefore only the line method is used.Hence, the only reported method for extracting a reference point for this experiment is the line method.

*1) Frequency metrics:* The frequency-based metrics especially the F50, and F95 show great consistency and invariance to the viewing angle (see Table I). When looking at these metrics, the focus should be around the move and big-move trajectories since in the stationary trajectory there is no movement and therefore, the frequencies being picked are

Table I
FREQUENCY BASED METRICS RESULTS FOR EXPERIMENT 1 (ROBOT EXPERIMENT)

| Trajectory | Speed | View Angle | F50 | F95 | Freq Dispersion | Centroid Freq | PWR |
|---|---|---|---|---|---|---|---|
| big move | 10 | 0 | 0.084 (0%) | 0.084 (0%) | 0.986 (0.1%) | 0.323 (7.9%) | 0.016 (0.5%) |
| big move | 10 | 5 | 0.084 (0%) | 0.084 (0%) | 0.984 (0.1%) | 0.291 (2.9%) | 0.017 (1%) |
| big move | 10 | 10 | 0.084 (0%) | 0.084 (0%) | 0.983 (0.1%) | 0.291 (3.4%) | 0.017 (0.8%) |
| big move | 50 | 0 | 0.32 (0%) | 0.34 (0%) | 0.866 (4.8%) | 0.592 (3%) | 0.015 (0.9%) |
| big move | 50 | 5 | 0.32 (0%) | 0.34 (0%) | 0.814 (2.5%) | 0.578 (1%) | 0.017 (0.8%) |
| big move | 50 | 10 | 0.32 (0%) | 0.34 (0%) | 0.815 (2.9%) | 0.576 (0.8%) | 0.017 (0.8%) |
| big move | 100 | 0 | 0.62 (0%) | 0.64 (0%) | 0.676 (9.6%) | 0.807 (1.2%) | 0.016 (0.6%) |
| big move | 100 | 5 | 0.62 (0%) | 0.64 (0%) | 0.60 (4.1%) | 0.8 (0.3%) | 0.017 (0.7%) |
| big move | 100 | 10 | 0.62 (0%) | 0.64 (0%) | 0.58 (4.7%) | 0.8 (0.3%) | 0.017 (0.4%) |
| move | 10 | 0 | 0.1 (0.3%) | 0.12 (0.3%) | 0.98 (0%) | 0.47 (9.2%) | 0.0044 (1.3%) |
| move | 10 | 5 | 0.1 (0.1%) | 0.12 (0.1%) | 0.98 (0%) | 0.42 (3.8%) | 0.0044 (1.6%) |
| move | 10 | 10 | 0.1 (0%) | 0.12 (0%) | 0.98 (0%) | 0.43 (6.9%) | 0.004 (1.3%) |
| move | 50 | 0 | 0.45 (0%) | 0.47 (0%) | 0.88 (2%) | 0.76 (3.6%) | 0.0044 (1.7%) |
| move | 50 | 5 | 0.45 (0%) | 0.47 (0%) | 0.86 (1.2%) | 0.73 (1%) | 0.0044 (0.7%) |
| move | 50 | 10 | 0.45 (0%) | 0.47 (0%) | 0.86 (2.4%) | 0.73 (1.8%) | 0.004 (0.6%) |
| move | 100 | 0 | 0.90 (0.1%) | 0.91 (1%) | 0.72 (6.2%) | 1 (2%) | 0.0043 (2.7%) |
| move | 100 | 5 | 0.91 (0.1%) | 0.92 (1.8%) | 0.69 (3.1%) | 0.99 (0.6%) | 0.0046 (1.1%) |
| move | 100 | 10 | 0.91 (0%) | 0.93 (0.9%) | 0.68 (3.4%) | 1 (0.9%) | 0.0042 (1.8%) |
| stationary | | 0 | 7.21 (7%) | 14.2 (2.1%) | 0.52 (4.5%) | 2.69 (2.5%) | 7.7E-05 (28.3%) |
| stationary | | 5 | 6.74 (11.8%) | 14.1 (2%) | 0.54 (3.7%) | 2.65 (3.1%) | 6.2E-05 (24.4%) |
| stationary | | 10 | 6.77 (10%) | 14.1 (2.1%) | 0.54 (3.7%) | 2.64 (3.1%) | 5.6E-05 (19.6%) |

Table II
AXES BASED METRICS FOR EXPERIMENT 1 (ROBOT EXPERIMENT)

| Trajectory | Speed | View Angle | Lateral Sway | Ventral Sway | Median Lateral Angle | Median Ventral Angle |
|---|---|---|---|---|---|---|
| big move | 10 | 0 | 0.0063 (3.4%) | 0.0076 (7%) | -1.06 (3.2%) | -2.04 (2.2%) |
| big move | 10 | 5 | 0.0078 (3.1%) | 0.00755 (7.8%) | -0.96 (5.7%) | -2.25 (2.2%) |
| big move | 10 | 10 | 0.0081 (2.9%) | 0.0072 (7.9%) | -0.95 (4.7%) | -2.32 (1.7%) |
| big move | 50 | 0 | 0.0061 (2.1%) | 0.0077 (2.6%) | -1.07 (0.6%) | -2.04 (0.1%) |
| big move | 50 | 5 | 0.0075 (3.2%) | 0.0076 (3.3%) | -0.97 (0.9%) | -2.26 (0.7%) |
| big move | 50 | 10 | 0.0078 (1.3%) | 0.0072 (2.3%) | -0.97 (0.6%) | -2.33 (0.2%) |
| big move | 100 | 0 | 0.006 (2.2%) | 0.0077 (1.6%) | -1.08 (0.5%) | -2.03 (0.7%) |
| big move | 100 | 5 | 0.0073 (1.6%) | 0.0076 (1.3%) | -0.98 (1.1%) | -2.27 (0.4%) |
| big move | 100 | 10 | 0.0077 (1.9%) | 0.0072 (1.4%) | -0.96 (1.2%) | -2.33 (0.2%) |
| move | 10 | 0 | 0.0035 (4.9%) | 0.0039 (5.3%) | -1.09 (1%) | -1.67 (0.9%) |
| move | 10 | 5 | 0.0037 (5.1%) | 0.0037 (4.6%) | -1.02 (0.9%) | -2.01 (0.4%) |
| move | 10 | 10 | 0.0037 (5.9%) | 0.0037 (7.2%) | -1 (0.7%) | -2.11 (0.5%) |
| move | 50 | 0 | 0.0035 (2.2%) | 0.004 (1.9%) | -1.09 (0.5%) | -1.66 (0.3%) |
| move | 50 | 5 | 0.0038 (2.1%) | 0.0037 (1.9%) | -1.02 (0.4%) | -2.01 (0.1%) |
| move | 50 | 10 | 0.0036 (1.3%) | 0.0037 (3.8%) | -1.01 (0.5%) | -2.1 (0.2%) |
| move | 100 | 0 | 0.0033 (2.4%) | 0.0039 (3.1%) | -1.08 (0.3%) | -1.69 (0.9%) |
| move | 100 | 5 | 0.0037 (2.4%) | 0.0037 (2.5%) | -1.03 (0.5%) | -2.02 (0.2%) |
| move | 100 | 10 | 0.0035 (1.7%) | 0.0037 (2.8%) | -1.01 (0.5%) | -2.11 (0.1%) |
| stationary | | 0 | 0.00041 (5.7%) | 0.00052 (5.7%) | -1.11 (0.5%) | -1.5 (1.1%) |
| stationary | | 5 | 0.00038 (5.2%) | 0.00047 (4.9%) | -1.09 (0.3%) | -1.8 (0.1%) |
| stationary | | 10 | 0.00035 (2%) | 0.00046 (3.1%) | -1.06 (0.6%) | -1.89 (0%) |

Table III
AXES INVARIANT METRICS FOR EXPERIMENT 1 (ROBOT EXPERIMENT)

| Trajectory | Speed | View Angle | Area | Median Dist | RMS | Range A | Range B |
|---|---|---|---|---|---|---|---|
| big move | 10 | 0 | 0.0012 (1%) | 0.012 (5%) | 0.012 (1.3%) | 0.052 (2.2%) | 0.029 (1.4%) |
| big move | 10 | 5 | 0.0013 (0.7%) | 0.013 (6.2%) | 0.013 (0.9%) | 0.055 (1.4%) | 0.03 (1%) |
| big move | 10 | 10 | 0.0013 (1.1%) | 0.013 (4.9%) | 0.013 (0.8%) | 0.054 (1.4%) | 0.031 (1%) |
| big move | 50 | 0 | 0.0012 (0.4%) | 0.012 (0.9%) | 0.012 (0.2%) | 0.052 (0.3%) | 0.029 (0.4%) |
| big move | 50 | 5 | 0.0013 (0.4%) | 0.013 (0.8%) | 0.013 (0.4%) | 0.055 (0.6%) | 0.03 (0.4%) |
| big move | 50 | 10 | 0.0013 (0.6%) | 0.013 (0.7%) | 0.013 (0.3%) | 0.054 (0.3%) | 0.031 (0.3%) |
| big move | 100 | 0 | 0.0012 (0.6%) | 0.012 (1%) | 0.012 (0.3%) | 0.052 (0.3%) | 0.029 (0.3%) |
| big move | 100 | 5 | 0.0013 (0.4%) | 0.013 (0.2%) | 0.013 (0.2%) | 0.055 (0.2%) | 0.03 (0.5%) |
| big move | 100 | 10 | 0.0013 (0.5%) | 0.013 (0.8%) | 0.013 (0.2%) | 0.054 (0.2%) | 0.031 (0.5%) |
| move | 10 | 0 | 0.00034 (1.3%) | 0.0065 (1.4%) | 0.0065 (0.9%) | 0.028 (1.2%) | 0.015 (0.7%) |
| move | 10 | 5 | 0.00034 (1.3%) | 0.0066 (0.9%) | 0.0065 (1%) | 0.028 (1.4%) | 0.016 (0.9%) |
| move | 10 | 10 | 0.00031 (1.2%) | 0.0062 (1%) | 0.0062 (0.8%) | 0.027 (1.3%) | 0.015 (1.2%) |
| move | 50 | 0 | 0.00034 (1.1%) | 0.0066 (0.7%) | 0.0065 (0.7%) | 0.028 (0.9%) | 0.016 (0.6%) |
| move | 50 | 5 | 0.00034 (0.9%) | 0.0067 (0.4%) | 0.0065 (0.4%) | 0.028 (0.5%) | 0.016 (0.8%) |
| move | 50 | 10 | 0.00031 (0.8%) | 0.0063 (0.6%) | 0.0062 (0.3%) | 0.027 (0.4%) | 0.015 (0.7%) |
| move | 100 | 0 | 0.00033 (2.5%) | 0.0065 (1.7%) | 0.0064 (1.1%) | 0.027 (1.1%) | 0.015 (1.5%) |
| move | 100 | 5 | 0.00036 (0.6%) | 0.0068 (0.6%) | 0.0066 (0.3%) | 0.028 (0.5%) | 0.016 (0.4%) |
| move | 100 | 10 | 0.00032 (1.4%) | 0.0064 (0.6%) | 0.0063 (0.5%) | 0.027 (0.7%) | 0.015 (1.2%) |
| stationary | | 0 | 1E-05 (17.5%) | 0.0008 (3.7%) | 0.001 (12.5%) | 0.005 (17.4%) | 0.003 (3.9%) |
| stationary | | 5 | 8E-06 (14.1%) | 0.0007 (3%) | 0.001 (9.2%) | 0.004 (13.7%) | 0.003 (2.6%) |
| stationary | | 10 | 8E-06 (22%) | 0.0007 (3%) | 0.001 (14.3%) | 0.004 (18.9%) | 0.003 (3.5%) |

an artifact of noise. The F50 metric shows great consistency with less than $0.3\%$ difference in repetitions and in changes of viewing angles. The F95 show good consistency as well but as expected, it is slightly more fragile and when the movement is fast, we see differences of up to $2.5\%$.

Other frequency-based metrics: frequency dispersion, centroid frequency, and power show less consistency and the difference between measurement can change by up to 10% between repetitions and between changes of the viewing angle. We note that both the centroid frequency and power seem to yield little difference when the viewing angle changes from $5^{\circ}$ to $10^{\circ}$ however when the viewing angle is $0^{\circ}$the measurement are significantly different.

The different exercises and different frequencies yielded significantly different results that allow separating between them. For example, the F50 ranges for the different exercises to not overlap. In the stationary position, the measurements are noisy (excluding power which records very small numbers). We conclude that frequency measurements that are above 2Hz should be considered as noise.

*2) Axes based metrics:* The position-based metric, especially the ones related to the body axes (lateral and ventral) are noisier and less consistent with respect to changes of the viewing angle as can be seen in Table II. For example, the lateral sway in the big move exercise was measured to be $\sim 30\%$ when viewed from $10^{\circ}$ compared to the $0^{\circ}$ viewing angle. This is to be expected since the axes are computed with respect to the Kinect plane which is distorted when the angle is not parallel to the body axis. At the same time, the results are consistent when the speed of movement changes which shows that these metrics are "orthogonal" to the frequency-based metrics.

*3) Axes invariant metrics:* Position-based metrics, as expected, were found to be less sensitive to viewing angle (see Table III). In all cases that there was movement (that is excluding the stationary trajectory), the range of error within the same trajectory and viewing angle was $< 2.5\%$. Changes in the speed generated very small changes to the measurement values however changes to the viewing angle did generate noticeable effects of up to $10\%$.

The measurements for the different exercises are significantly different and therefore, there is a clear separation between the different trajectories with respect to these features. The results for the stationary exercise show greater noise (up to 22%). This shows that this setup has a lower bound on the smallest sizes it can measure. For example, measurements below $2 \cdot 10^{-5}$ should be considered as noise.

### *B. Results of Experiment Two - Human Posture*

The second experiment compared the measurements of the Kinect to the measurement obtained by a force plate using human subjects.

*1) Sensitivity:* Table IV shows the cross correlation between the measurements on the force plate (Wii board) and the measurement using the Kinect. When the RP is computed using the line model, there is a high correlation ($r > 0.7$) for most of the metrics captured. The path metric does not correlate as well as some of the frequency domain based metrics. This might be due to the tracking algorithm used by the Kinect that acts as a low-pass-filter.

A large correlation coefficient indicates that we should be able to track changes in sway patterns in individuals. In Table V we look at the ability of the Kinect to distinguish between different postures in a way that is not subject dependent. This allows using a single measurement to infer the condition of the subject (as opposed to tracking changes over time). Table V shows the P-value computed by the 2-tails sign test. In this table we see that most metrics can distinguish between the baseline posture and the tandem posture with P-value $< 0.05$ and in some cases, even much smaller P-values. Most metrics are also sensitive enough to distinguish between the baseline posture and the eyes closed posture although this task is harder. The results for the force plate and the Kinect are comparable. The force plate is more sensitive, but the Kinect can use metrics, such as the median angle that are not accessible from the force plate.

*2) Reliability:* To test the reliability of the proposed method, we compared the measurements taken in two different postures: standing with feet parallel with arms crossed and standing with feet parallel with arms to the side. We used ICC(1) to compute the inter-class correlation coefficient and used the SM method [9] to compute the $95\%$ confidence intervals on the ICC for the different measurements. The results are presented in Table VI. This experiment shows that most metrics are reliable (ICC $> 0.4$) with the exception of Path, F95 and Centroid Freq. This is consistent with our findings in other measurements in the sense that metrics which are sensitive to high frequency signals are not well captured by the Kinect. Comparing the reliability of the Kinect and Wii shows that they are similar in performance.

### *C. Results for Experiment Three - Posing Verification*

Our experiments show that if a subject stand parallel to the Kinect surface to within a few degrees the measurements obtained are accurate. To verify that subjects truly stand in the acceptable angle range we conducted another experiment in which we measured the variability of the directions in which subjects stand in front of the camera. 13 subjects (5 Females, 8 male) have been instructed to stand along a line, representing a typical situation during the test, or in a clinic. Each subject repeated the task twice. The locations of the feet were photographed by two DSLR cameras, and distances to the line have been recovered. The subjects were found to stand at an average of $0.55$ degrees relative to the line (Standard Deviation $1.89$ degrees) which is well beneath the above discussed 5 degrees.

## IV. CONCLUSIONS

Our experiments show that the Kinect is sensitive and reliable for most sway parameters tested. Our first experiment, using a mannequin, demonstrated the limitations of approaches that use VSE when faced with atypical body structures: the VSE failed to recognize the mannequin while our proposed approach was able to track it. Therefore, the methods presented

| Metric | line model | figure mean | skeleton neck |
|---|---|---|---|
| Median X | 0.65 | 0.56 | 0.47 |
| Median Z | 0.80 | 0.77 | 0.71 |
| Median Dist | 0.77 | 0.70 | 0.66 |
| RMS | 0.78 | 0.72 | 0.72 |
| Path | 0.23 | 0.34 | 0.17 |
| Area | 0.76 | 0.67 | 0.69 |
| Range A | 0.79 | 0.74 | 0.74 |
| Range B | 0.70 | 0.64 | 0.54 |
| Median Freq | 0.45 | 0.55 | 0.23 |
| PWR | 0.83 | 0.75 | 0.80 |
| F50 | 0.35 | 0.43 | 0.33 |
| F95 | 0.08 | 0.19 | 0.38 |
| Centroid Freq | 0.43 | 0.43 | 0.45 |
| Freq Dispersion | 0.51 | 0.52 | 0.06 |

Table IV

THE CORRELATION COEFFICIENT BETWEEN THE KINECT MEASUREMENTS OF SWAY AND THE FORCE PLATE MEASUREMENTS. THE VALUES ARE COLOR CODED AS FOLLOWS: VERY HIGH CORRELATION ($\geq 0.7$) IN GREEN, HIGH CORRELATION ($\geq 0.4$) IN LIGHT GREEN, MODERATE CORRELATION ($\geq 0.2$) IN YELLOW AND NEGLIGIBLE CORRELATION ($< 0.2$) IN RED.

| Device | force plate | force plate | force plate | Kinect | Kinect | Kinect |
|---|---|---|---|---|---|---|
| Eyes | Close | Open | Open | Close | Open | Open |
| Hands | Cross | Side | Cross | Cross | Side | Cross |
| Feet | Parallel | Parallel | Tandem | Parallel | Parallel | Tandem |
| Median X | 0.01 | 0.8 | 0.003 | 0.5 | 0.5 | 0.1 |
| Median Z | 0.01 | 0.8 | 0.003 | 0.5 | 0.3 | 0.04 |
| Median angle X | N/A | N/A | N/A | 0.3 | 0.8 | 0.3 |
| Median angle Z | N/A | N/A | N/A | 0.04 | 2E-6 | 4E-5 |
| Median Dist | 0.003 | 0.8 | 0.003 | 0.1 | 0.1 | 0.01 |
| RMS | 0.003 | 1 | 4E-4 | 0.01 | 0.8 | 0.003 |
| Path | 2E-6 | 1 | 2E-6 | 0.5 | 0.04 | 0.04 |
| Area | 4E-4 | 0.8 | 0.003 | 0.04 | 0.8 | 0.003 |
| Range A | 0.01 | 0.3 | 0.01 | 0.1 | 0.5 | 4E-4 |
| Range B | 4E-5 | 0.8 | 0.003 | 0.01 | 0.8 | 0.003 |
| Median Freq | 0.5 | 0.8 | 0.04 | 0.04 | 0.8 | 4E-4 |
| PWR | 0.01 | 0.8 | 0.01 | 0.04 | 0.5 | 0.3 |
| F50 | 0.01 | 0.5 | 0.01 | 0.04 | 0.8 | 0.3 |
| F95 | 0.01 | 0.5 | 2E-6 | 0.04 | 0.8 | 0.003 |
| Centroid Freq | 4E-4 | 0.5 | 4e-5 | 0.5 | 1 | 0.8 |
| Freq Dispersion | 2E-6 | 0.3 | 4E-5 | 0.01 | 0.5 | 0.04 |

Table V

THE SENSITIVITY OF THE DIFFERENT TECHNIQUES IS COMPUTED BY THE P-VALUE OF THE PAIRED SIGN TEST (WITH 2 TAILS). THE RESULTS WERE COLOR CODED AS FOLLOWS: VERY SIGNIFICANCE ($p < 0.01$) IN GREEN, SIGNIFICANCE ($p < 0.05$) IN LIGHT GREEN, BORDERLINE SIGNIFICANCE ($p < 0.15$) IN YELLOW AND NOT-SIGNIFICANCE ($p \geq 0.15$) IN RED.

| Metric | Force Plate | | Kinect | |
|---|---|---|---|---|
| | ICC(1) | 95% Confidence | ICC(1) | 95% Confidence |
| Median X | 0.67 | 0.64 – 0.69 | 0.52 | 0.48 – 0.56 |
| Median Z | 0.31 | 0.26 – 0.37 | 0.41 | 0.36 – 0.46 |
| Median angle X | N/A | N/A | 0.81 | 0.80 – 0.82 |
| Median angle Z | N/A | N/A | 0.52 | 0.48 – 0.56 |
| Median Dist | 0.53 | 0.49 – 0.57 | 0.66 | 0.64 – 0.69 |
| RMS | 0.40 | 0.35 – 0.45 | 0.54 | 0.50 – 0.58 |
| Path | 0.87 | 0.86 – 0.87 | 0.35 | 0.29 – 0.40 |
| Area | 0.49 | 0.44 – 0.53 | 0.60 | 0.57 – 0.64 |
| Range A | 0.18 | 0.13 – 0.24 | 0.44 | 0.39 – 0.49 |
| Range B | 0.68 | 0.65 – 0.70 | 0.53 | 0.49 – 0.57 |
| Median Freq | 0.65 | 0.62 – 0.68 | 0.48 | 0.41 – 0.52 |
| PWR | 0.21 | 0.16 – 0.27 | 0.46 | 0.41 – 0.50 |
| F50 | 0.23 | 0.17 – 0.29 | 0.43 | 0.39 – 0.48 |
| F95 | 0.62 | 0.59 – 0.65 | 0.52 | 0.48 – 0.56 |
| Centroid Freq | 0.41 | 0.36 – 0.46 | 0.56 | 0.52 – 0.60 |
| Freq Dispersion | 0.59 | 0.55 – 0.62 | 0.78 | 0.76 – 0.79 |

Table VI

THE INTER-CLASS CORRELATION COEFFICIENTS. THE RESULTS WERE COLOR CODED BASED ON THE VALUE OF THE LOWER BOUND ON THE CONFIDENCE INTERVAL. THE COLOR CODE IS: VERY HIGH CORRELATION ($\geq 0.7$) IN GREEN, HIGH CORRELATION ($\geq 0.4$) IN LIGHT GREEN, MODERATE CORRELATION ($\geq 0.2$) IN YELLOW AND NEGLIGIBLE CORRELATION ($< 0.2$) IN RED.

here may be used by people with atypical body structure for which current approaches may fail. For most of the parameters, the correlation between the estimates of the Kinect and measurements by a pressure sensor were high ($\geq 0.65$). Most metrics were also found to be sensitive to show the difference in the sway patterns between different postures with p-values in the range $2 \cdot 10^{-6}$ to 0.05. The approach presented here has many advantages over alternative approaches of measuring sway in terms of cost and ease of use.

Several important questions are left for future studies. First, it is important to test how well our methods work for people with different body structures. Such an experiment requires assembling a cohort of people with diverse atypical body shapes which is a challenging task by itself that is left for future study. Another interesting direction is to use the same approach to measure other aspects of gait and balance.

## REFERENCES

[1] Accessibility and kinect for xbox 360. https://support.xbox.com/en-US/xbox-360/accessories/accessibility-kinect. Accessed: 2019-04-28.

[2] AL Adkin, BR Bloem, and JHJ Allum. Trunk sway measurements during stance and gait tasks in parkinson's disease. *Gait & posture*, 22(3):240–249, 2005.

[3] Abrar Alharbi, Fahad Alharbi, and Eiji Kamioka. Skeleton based gait recognition for long and baggy clothes. In *MATEC Web of Conferences*, volume 277, page 03005. EDP Sciences, 2019.

[4] Janina Behrens, Caspar Pfüller, Sebastian Mansow-Model, Karen Otte, Friedemann Paul, and Alexander U Brandt. Using perceptive computing in multiple sclerosis-the short maximum speed walk test. *Journal of neuroengineering and rehabilitation*, 11(1):89, 2014.

[5] Olivier Caron, Bernard Faure, and Yvon Brenière. Estimating the centre of gravity of the body on the basis of the centre of pressure in standing posture. *Journal of biomechanics*, 30(11):1169–1171, 1997.

[6] Ross A Clark, Adam L Bryant, Yonghao Pua, Paul McCrory, Kim Bennell, and Michael Hunt. Validity and reliability of the Nintendo Wii balance board for assessment of standing balance. *Gait & posture*, 31(3):307–310, 2010.

[7] Ross A Clark, Yong-Hao Pua, Adam L Bryant, and Michael A Hunt. Validity of the microsoft kinect for providing lateral trunk lean feedback during gait retraining. *Gait & posture*, 38(4):1064–1066, 2013.

[8] Ross A Clark, Yong-Hao Pua, Karine Fortin, Callan Ritchie, Kate E Webster, Linda Denehy, and Adam L Bryant. Validity of the Microsoft Kinect for assessment of postural control. *Gait & posture*, 36(3):372–377, 2012.

[9] Allan Donner and George Wells. A comparison of confidence interval methods for the intraclass correlation coefficient. *Biometrics*, pages 401–412, 1986.

[10] Moataz A Eltoukhy, Christopher Kuenze, Jeonghoon Oh, and Joseph F Signorile. Validation of static and dynamic balance assessment using microsoft kinect for young and elderly populations. *IEEE journal of biomedical and health informatics*, 22(1):147–153, 2018.

[11] Moshe Gabel, Ran Gilad-Bachrach, Erin Renshaw, and Assaf Schuster. Full body gait analysis with Kinect. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 1964–1967. IEEE, 2012.

[12] Brook Galna, Gillian Barry, Dan Jackson, Dadirayi Mhiripiri, Patrick Olivier, and Lynn Rochester. Accuracy of the microsoft kinect sensor for measuring movement in people with parkinson's disease. *Gait & posture*, 39(4):1062–1068, 2014.

[13] Maryam Ghahramani, David Stirling, Fazel Naghdy, Golshah Naghdy, and Janette Potter. Body postural sway analysis in older people with different fall histories. *Medical & biological engineering & computing*, 57(2):533–542, 2019.

[14] Kitchana Kaewkaen, Marnida Koetkhumtong, Phatcharawadi Decha, Kunyawee Kumnet, Chayuti Mekurai, Worasak Rueangsirarak, and Tsuyoshi Asai. Effects of balance training incorporating with a kinect-based exergame on mediolateral postural sway in older adults with balance impairment: A pilot study. *Journal of Associated Medical Sciences*, 50(2), 2017.

[15] Martina Mancini, Arash Salarian, Patricia Carlson-Kuhta, Cris Zampieri, Laurie King, Lorenzo Chiari, and Fay B Horak. ISway: a sensitive, valid and reliable measure of postural control. *J Neuroeng Rehabil*, 9(59):59, 2012.

[16] MJ Mathie, J Basilakis, and BG Celler. A system for monitoring posture and physical activity using accelerometers. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd annual international conference of the IEEE*, volume 4, pages 3654–3657. IEEE, 2001.

[17] Christoph Maurer and Robert J Peterka. A new interpretation of spontaneous sway measures based on a simple model of human postural control. *Journal of Neurophysiology*, 93(1):189–200, 2005.

[18] Anup K Mishra, Marjorie Skubic, Bradley W Willis, Trent Guess, Swithin S Razu, Carmen Abbott, and Aaron D Gray. Examining methods to estimate static body sway from the kinect v2. 0 skeletal data: implications for clinical rehabilitation. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 127–135. ACM, 2017.

[19] Karen Otte, Bastian Kayser, Sebastian Mansow-Model, Julius Verrel, Friedemann Paul, Alexander U Brandt, and Tanja Schmitz-Hübsch. Accuracy and reliability of the kinect version 2 for clinical measurement of motor function. *PloS one*, 11(11):e0166532, 2016.

[20] Paula YS Poh, Amy N Adams, Mu Huang, Dustin R Allen, Scott L Davis, Anna S Tseng, and Craig G Crandall. Increased postural sway in persons with multiple sclerosis during short-term exposure to warm ambient temperatures. *Gait & posture*, 53:230–235, 2017.

[21] Jamie Shotton, Andrew W Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake. Real-time human pose recognition in parts from single depth images. In *Cvpr*, volume 2, page 3, 2011.

[22] Erik E Stone and Marjorie Skubic. Evaluation of an inexpensive depth camera for passive in-home fall risk assessment. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, pages 71–77. IEEE, 2011.

[23] Erik E Stone and Marjorie Skubic. Passive in-home measurement of stride-to-stride gait variability comparing vision and Kinect sensing. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 6491–6494. IEEE, 2011.

[24] Jaap Swanenburg, Eling D de Bruin, Daniel Uebelhart, and Theo Mulder. Falls prediction in elderly people: a 1-year prospective study. *Gait & posture*, 31(3):317–321, 2010.

[25] Dawn Tan, Yong-Hao Pua, Shaminian Balakrishnan, Aileen Scully, Kelly J Bower, Kumar Manharlal Prakash, Eng-King Tan, Jing-Si Chew, Evelyn Poh, Siok-Bee Tan, et al. Automated analysis of gait and modified timed up and go using the microsoft kinect in people with parkinson's disease: associations with physical outcome

measures. *Medical & biological engineering & computing*, pages 1–9, 2018.

[26] Soumya Ranjan Tripathy, Kingshuk Chakravarty, and Aniruddha Sinha. Eigen posture based fall risk assessment system using kinect. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1–4. IEEE, 2018.

[27] Luigi Uccioli, Pier Giorgio Giacomini, Giovanna Monticone, Antonio Magrini, Laura Durola, Ernesto Bruno, Leo Parisi, Stefano Di Girolamo, and Guido Menzinger. Body sway in diabetic neuropathy. *Diabetes care*, 18(3):339–344, 1995.

[28] M Van Diest, J Stegenga, HJ Wörtche, GJ Verkerke, K Postema, and CJC Lamoth. Exergames for unsupervised balance training at home: a pilot study in healthy older adults. *Gait & posture*, 44:161–167, 2016.

[29] LF Yeung, Kenneth C Cheng, CH Fong, Winson CC Lee, and Kai-Yu Tong. Evaluation of the Microsoft Kinect as a clinical assessment tool of body sway. *Gait & posture*, 40(4):532–538, 2014.