34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Full Body Gait Analysis with Kinect*

Moshe Gabel[1], Ran Gilad-Bachrach[2], Erin Renshaw[3] and Assaf Schuster[4]

Abstract— **Human gait is an important indicator of health, with applications ranging from diagnosis, monitoring, and rehabilitation. In practice, the use of gait analysis has been limited. Existing gait analysis systems are either expensive, intrusive, or require well-controlled environments such as a clinic or a laboratory.**

We present an accurate gait analysis system that is economical and non-intrusive. Our system is based on the Kinect sensor and thus can extract comprehensive gait information from all parts of the body. Beyond standard stride information, we also measure arm kinematics, demonstrating the wide range of parameters that can be extracted. We further improve over existing work by using information from the entire body to more accurately measure stride intervals. Our system requires no markers or battery-powered sensors, and instead relies on a single, inexpensive commodity 3D sensor with a large pre-existing install base. We suggest that the proposed technique can be used for continuous gait tracking at home.

I. Introduction

Human gait has been shown to be an important indicator of health, and is applicable in a wide range of settings, such as diabetes [1], neurological diseases [2], [3], and fall detection and prediction [4]. Accurate, non-intrusive, low-cost clinical gait analysis systems have many applications in diagnosis, monitoring, treatment and rehabilitation [5], [1]. Such applications include early diagnosis and assessment [2], [6], measuring medication effectiveness at home [7], and even direct treatment optimization [8], [9].

Several methods have been proposed for gait analysis. Marker-based systems typically use IR cameras and markers placed on the subject. These systems are accurate, but often very expensive and impractical to move. Additionally, passive or active markers must be correctly placed on the body before each capture session. Therefore, such systems are only suitable for laboratory settings. Force plates are also used for gait analysis. Again, these systems are usually costly and are only found at laboratories and clinics. Moreover, they measure the dynamics of the lower limbs only.

Recent studies proposed the use of wearable sensors [10], [11]. Such systems are more suitable for ambulatory measurements in home settings as they are small, lightweight, mobile and less expensive (see [12] for a review). One can also use insole pressure sensors as another means of measuring gait properties. Despite their advantages, wearable sensors suffer from some drawbacks. Sensors must be placed correctly and securely [12], and must account for gravity, noise and signal drift [13]. Moreover, each sensor is usually limited to measuring very few gait properties and hence an array of sensors is needed to obtain a comprehensive analysis. Moreover, these sensors are intrusive in the sense that they require changes to the daily routine of the subject. They also require maintenance in the form of charging batteries, uploading data and sanitary treatment.

Markerless optical gait analysis systems have been discussed in the context of biometric identification and surveillance. Single or multiple video cameras can be used to recognize individuals [14]. [15] discusses another technique in the context of medical applications, but the accuracy of the extracted stride parameters was not verified. [16] focuses on extraction of knee joint angles, but not on standard stride parameters. This system uses two cameras to generate a 3D image, but requires complex setup and calibration, as many similar systems do. Both techniques are limited to information from the lower part of the body.

We propose a low-cost, non-intrusive system that can accurately measure a wide range of gait parameters using the Kinect sensor and Software Development Kit (SDK). Kinect is an array of sensors, including a camera and a depth sensor. In addition to the raw depth image, Kinect extracts a 3D virtual skeleton of the body [17]. These capabilities, packed in an affordable and compact device, already led several researchers to propose its use for home monitoring and gait analysis [18], [19].

We apply a supervised learning approach to automatically and accurately extract lower and upper body gait parameters, using the 3D virtual skeleton. This allows us to go beyond standard foot stride parameters. For example, we extract arm kinematics using the same sensor. We show that our method is accurate and robust to attributes such as sensor position. Moreover, our method can be extended to measure other properties such as leg kinematics.

Stone and Skubic [18], [19] were the first to propose the use of Kinect for clinical gait analysis. In their work, they use the depth image to locate objects which are at a height of 50cm or less. They use the volume of these objects to infer whether the left or the right foot is touching the ground. We improve on their work in several ways. Our technique uses information from the entire body to improve accuracy. Moreover, we demonstrate how a rich set of parameters can be extracted. To demonstrate that, we measure arm kinematics parameters that have been shown to be useful for medical applications [20], [2]. Finally, we show that our method is robust to environmental changes and changes in

*This work was not supported by any organization.

[1]The Department of Computer Science, Technion – Israel Institute of Technology. mgabel@cs.technion.ac.il

[2]Microsoft Research. rang@microsoft.com

[3]Microsoft Research. Erin.Renshaw@microsoft.com

[4]The Department of Computer Science, Technion – Israel Institute of Technology. assaf@cs.technion.ac.il. Parts of this work were conducted while the author was visiting Microsoft Research.

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

the placement of the Kinect sensor.

Empirical evaluation shows our results to be very accurate when compared to reference signals such as pressure sensors, and compared to previous work. We demonstrate the ability to extract a rich set of parameters, for example arm swing parameters. Moreover, we suggest that the proposed method is affordable and non-intrusive since in a typical use-case a Kinect sensor can be placed in a fixed position at home.

II. METHODS

Our technique uses a "virtual skeleton" produced by the Kinect sensor and software. The skeleton information is converted into a large set of features which are fed to a model that predicts the values of interest. For example, in order to measure stride duration, the model detects whether the foot is touching the ground. The outcome of this model is fed to a state machine that detects the current state from which the measurements are derived. In the following section we discuss each of these steps in detail.

A. Experimental Setup and Data Acquisition

To build and evaluate our model, we captured Kinect skeleton recordings of walking subjects with time synchronized readings from in-shoe pressure sensors and a gyroscope attached to the wrist via straps. Applying our method does not require the wearable sensors, but we use them to evaluate accuracy. Therefore, we use parts of the data acquired to train a model to predict the values of interest and other parts of the data to evaluate the quality of these predictions.

Subjects and Kinect Setup: We captured recordings of 23 subjects (19 male, 4 female), in the age range of 26 to 56 years. Subjects were instructed to walk at a normal pace back and forth 5 times along a straight path of approximately 7m. In each experiment, a Kinect sensor was placed to capture the image of the subject. For 11 subjects (9 male, 2 female) the sensor was placed at an approximate 45 degree angle to the middle of the path line, at a height of 30–60cm above the floor. For the other 12 subjects (10 male, 2 female) the sensor was fixed to the ceiling, straight ahead of the path line (0 degree angle), simulating a hallway. The view angle ensures that the sensor covers the middle part of the walking path, but not the edges where subjects turn. Recordings were collected in several locations, and the Kinect sensor itself was occasionally moved between recordings, to allow testing the robustness of the proposed approach. We used the Kinect Sensor for XBOX 360 model 1414, with the Kinect SDK v1.0 beta 2[1].

B. Validation Setup

We use readings from wearable sensors as "ground truth" to first build our model, and later evaluate its accuracy. It is important to note that the model itself (see Sec. II-C) does not use sensors' reading and they need not be present when it is being used. Sensor readings were sampled by custom hardware and wirelessly transmitted to a PC, where they were synchronized to the Kinect skeleton frames and recorded. To minimize the effect of the wearable sensors on gait, each subject was recorded 4 times: once with the pressure sensors in the left shoe, once in the right shoe, and similarly once for the gyroscope on each arm. This resulted in 92 sessions being recorded.[2]

Pressure Sensors: Two pressure sensors (FlexiForce A201[3]) were placed inside the shoe, below or on top of the insole. One sensor was placed under the heel to capture the heel strike. The second pressure sensor was placed underneath the great toe joint to capture the time when the foot is being lifted off the ground.

Recorded pressure sensor values are affected by differences in weight, foot anatomy, and shoe type. Hence, we use a linear normalization such that all the readings but the top 10% and bottom 10% will fit in the $[-1, 1]$ range.

Gyroscope: Similar to pressure sensors, we used a gyroscope (ITG-3200 by InvenSense[4]) to record angular velocity of the arms. To avoid errors due to sensor placement [12], the sensor was attached to the upper part of the wrist using a strap in a fixed ordination: facing outwards from the arm. Nevertheless, we applied Principal Component Analysis (PCA) [21] to the readings from the 3D gyroscope sensor to find the main direction of arm movement and used only the measurements in this direction.

C. Predictive Model

The Kinect sensor and its SDK provide a 3D virtual skeleton. The virtual skeleton consists of the positions of 20 joints and body parts (such as the wrists, knees, head and torso), which we refer to as joints. For each joint, the x, y and z coordinates are reported, as well as a confidence parameter which indicates the confidence of the skeleton extraction algorithm in those coordinates. Kinect provides approximately 30 skeleton frames per second.

Our method converts the skeleton frames into a feature vector. For each of the sensors used, we build a regression model that predicts the value recorded by the corresponding sensor. A simple state machine is used to identify the beginning of strides and to partition strides to their parts. In the following we provide more details about these components. It is important to note that the sensors described in Sec. II-B are not needed when using the system. Instead, the only sensor needed is Kinect.

1) Features: Skeleton data is converted into a large set of features. We examine each frame together with the 2 previous frames and the 2 following frames. In each frame, we locate the "center of mass" (COM) as the center of the hip joints, the shoulder joints and the spine. The change in position of the COM between consecutive frames is computed and the median (in each coordinate) of these differences is considered the current direction of progress (DOP). The

[1] `http://www.microsoft.com/en-us/kinectforwindows/develop/beta.aspx`

[2] 6 sessions were discarded. In 5 sessions, the pressure sensor could not be fitted securely in subject's shoe. Another session was dropped due to a failure of the gyroscope during the recording session.

[3] `http://www.tekscan.com/pdf/A201-force-sensor.pdf`

[4] `http://invensense.com/mems/gyro/itg3200.html`

speed of walking is defined to be the norm of the direction vector.

For each joint, in each frame, the current position of the joint is estimated in the coordinate system, which is composed from: (a) the DOP; (b) the up direction provided by the Kinect sensor; and (c) the direction tangent to these vectors. The axes are aligned such that the zero in the up direction is the ground while in the other direction's zero is defined as the COM. Finally, for each joint we generate the following features: the position in each of the frames (in the above coordinate system), the difference in position between consecutive frames (an estimate of the velocity), and the difference of the differences (an estimate of the acceleration). The same set of features is computed also for the center of mass (COM), and together with the speed and the direction of progress, they form the feature vector.

2) The regression model: We use an ensemble of regression trees to predict the values of interest from the feature vector described in Sec. II-C.1. To learn the regression model, we use the Multiple Additive Regression Trees (MART) algorithm [22], [23]. A regression tree is a rooted binary tree with conditions in the inner nodes and real values in the leaves. Each condition is a comparison of one of the input features to a threshold. The ensemble learned by MART consists of 150 trees, each having 20 leaf nodes. The value predicted by the ensemble is the sum of the values predicted by all the trees in the ensemble. Parameter tuning for the MART algorithm was done on few recordings which were not a part of this experiment.

D. Stride Detection and Partitioning

We follow standard practice (see, for example, [4], [3], [11]) and define stride time as the time from initial contact of one foot with the ground to the subsequent contact of the same foot. Each stride (gait cycle) is composed of a stance phase where the foot is on the ground, followed by a swing phase where the foot is swung forward. The heel and toe pressure signals (estimated or real) are fed to a simple state machine consisting of three states: HEEL, TOE and SWING. Whenever the heel signal is "pressed" (scaled reading greater than zero), we assume that the state is HEEL. Once the toe signal is pressed, the state machine is advanced to state TOE. If neither signal is pressed for the next 100ms, the state advances to state SWING. We consider the foot to be in a stance phase whenever the state machine is at the HEEL or TOE states. Otherwise, we assume the swing phase.

III. Results

We evaluated the accuracy of our method by comparing extracted parameters from the skeleton to the reference values from the sensor. We conducted two experiments, one to measure the accuracy of the proposed method and the other to measure its robustness to changes in the placement of the Kinect sensor and the environment.

A. Experiment 1 – Accuracy

To prevent over-fitting, accuracy is measured using cross-validation [24]: We hold the data for one of the participants out and train on the remaining data. We use the held out data for testing the model. We repeat this process for every participant and report the average accuracy.

The summary of the results of measuring stride durations is presented in Table I. For different components of a stride, the table shows the following statistics: (1) the average duration as measured by the pressure sensor (Avg), (2) the average difference between the duration measured by the pressure sensor and the duration measured by the regression model (Mean-diff), (3) the standard deviation between the two measurements (Std-diff), (4) the average differences between the measured duration in absolute values (Abs-diff), and (5) the number of events (N). All but the last column are reported in milliseconds.

Table I shows that the predictions generated by the model are very accurate. The Mean-diff (or bias) is especially small (less than 1% when measuring stride duration). Both the bias and the standard-deviations in our experiment are smaller than the corresponding values reported in [19, Table I]. The Abs-diff column shows that in absolute value, the difference between the predictions and measured durations is 35–71ms. Given the 30 frames per second rate, this could be interpreted as 1–2 frames difference.

We also tested the quality of the prediction of the angular velocity. Here we compared the readings from the gyroscope to the corresponding Kinect-based model. The correlation coefficients between the Kinect-based prediction and the true value are greater than 0.91 for both arms. The average difference between the readings is 1.52 (-0.86) for the left (right) arm and the standard deviation is 48.36 (44.63) where the unit of measurement is degree/second.

Our method can detect the stride phase and the arm state simultaneously. This allows for a detailed analysis of the angular velocity of the arm at different stride phases. Table II shows some values computed by this method. Other states, including both legs on the ground, are not reported to keep the presentation concise. Nevertheless, this analysis shows that the methods presented here allow for detailed and accurate full body gait analysis.

B. Experiment 2 – Robustness

One of the limitations of many gait analysis methods is that they require careful calibration and setup. To test the robustness of our method, we conducted a second experiment. In this experiment, we trained the models using the data captured when Kinect was placed at approximately 45 degrees to the path line and at a height of 30-60cm from the ground. We tested the models on the data recorded on different subjects, when the sensor was attached to the ceiling in front of the walking path line.

The results in Table III show the estimations of durations of stride components in this experiment. The accuracy of the measurements degrades only slightly when compared to the results in Table I despite the fact that the models were trained on one viewing angle and were tested on a different viewing angle. The models for the angular velocities of the arm exhibit even greater robustness to the different viewing angle.

TABLE I

EXPERIMENT 1: STRIDE DURATIONS PREDICTIONS COMPARED TO PRESSURE SENSORS. THE UNIT OF MEASUREMENT IS A MILLISECOND.

Interval	Avg	Mean-diff	Std-diff	Abs-diff	N
Left stride	1169	8	62	45	62
Right stride	1130	2	46	32	46
Left stance	634	-8	110	70	111
Right stance	595	-20	90	67	96
Left swing	518	6	115	71	146
Right swing	541	27	104	70	124

TABLE II

EXPERIMENT 1: ARM ANGULAR VELOCITIES AT DIFFERENT STRIDE PHASES. THE UNIT OF MEASUREMENT IS DEGREE/SECOND.

Feet state	Arm	Avg	Mean-diff	Std-diff	Abs-diff	N
Left swing	Left	-100	-3.6	21	16	148
Left swing	Right	91	-0.9	29	22	163
Right swing	Left	106	8.0	24	20	149
Right swing	Right	-88	-1.7	27	22	152
Left stance	Left	78	3.3	20	16	102
Left stance	Right	-75	0.5	22	18	118
Right stance	Left	-84	-3.4	19	14	121
Right stance	Right	70	-1.9	26	19	110

The correlation coefficient between the prediction and the true value is 0.9 (0.89) for the left (right) arm. The average difference in the reading is 1.62 (5.33) and the standard deviation is 56 for both arms. This accuracy confirms our assumption regarding the robustness of the approach.

IV. DISCUSSION

In this work we have presented a novel method for full body gait analysis using the Kinect sensor. Using the virtual skeleton as the input to a learned model, we demonstrated accurate and robust measurements of a rich set of gait features. We showed that our method improves on prior art [19] both in terms of having smaller bias and in having smaller variance. Moreover, our method can be extended to measuring other properties, including lower limb angular velocities and core posture. The sensor used is affordable and small, thus allowing installation in domestic environments. Since the sensor does not require maintenance, it allows for continuous and long term tracking of gait and its trends. These properties enable many applications for diagnosis, monitoring and adjustments of treatment [5], [1]. However, measuring the utility of the methods presented here for medical applications is a subject for further research.

TABLE III

EXPERIMENT 2: STRIDE DURATION COMPARISON

Interval	Avg	Mean-diff	Std-diff	Abs-diff	N
Left stride	1152	18.1	112	98	8
Right stride	1129	27.3	52	45	8
Left stance	613	5.4	123	84	32
Right stance	601	8.0	128	92	37
Left swing	532	-8.3	106	68	42
Right swing	523	11.7	112	74	49

REFERENCES

[1] D. Hodgins. The importance of measuring human gait. *Med Device Technol*, 19(5):42, 44–42, 47, Sep 2008.
[2] N.L.W. Keijsers, M.W.I.M. Horstink, and C.C.A.M. Gielen. Ambulatory motor assessment in Parkinson's disease. *Movement Disorders*, 21(1):34–44, 2006.
[3] J. M. Hausdorff, A. Lertratanakul, M. E. Cudkowicz, A. L. Peterson, D. Kaliton, and A. L. Goldberger. Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis. *Journal of Applied Physiology*, 88(6):2045–2053, 2000.
[4] J. M. Hausdorff, D. A. Rios, and H. K. Edelberg. Gait variability and fall risk in community-living older adults: A 1-year prospective study. *Archives of Physical Medicine and Rehabilitation*, 82(8):1050—1056, 2001.
[5] T. A. Wren, G. E. Gorton III, S. Õunpuu, and C. A. Tucker. Efficacy of clinical gait analysis: A systematic review. *Gait & Posture*, 34(2):149–153, 2011.
[6] J. I. Hoff, A. A. van den Plas, E. A. H. Wagemans, and J. J. van Hilten. Accelerometric assessment of levodopa-induced dyskinesias in Parkinson's disease. *Movement Disorders*, 16(1):58–61, 2001.
[7] N.L.W. Keijsers, M.W.I.M. Horstink, and S.C.A.M. Gielen. Movement parameters that distinguish between voluntary movements and levodopa-induced dyskinesia in Parkinson's disease. *Human Movement Science*, 22(1):67–89, 2003.
[8] A. Legros, N. Diakonova, L. Cif, S. Hemm, N. Vayssuére, P. Coubes, and A. Beuter. Accelerometric measurement of involuntary movements during pallidal deep brain stimulation of patients with generalized dystonia. *Brain Research Bulletin*, 64(4):363–369, 2004.
[9] B. Holt, A. Reyes, and F. Valero-Cuevas. A prototype video game system for studying rehabilitative learning. In *35th Annual American Society for Biomechanics Conference*, 2011.
[10] C.-C. Yang and Y.-L. Hsu. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, 10(8):7772–7788, 2010.
[11] K.J. O'Donovan, B.R. Greene, D. McGrath, R. O'Neill, A. Burns, and B. Caulfield. SHIMMER: A new tool for temporal gait analysis. In *EMBC*, pages 3826–3829, 2009.
[12] J. J. Kavanagh and H. B. Menz. Accelerometry: a technique for quantifying movement patterns during walking. *Gait & Posture*, 28(1):1–15, 2008.
[13] H.J. Luinge and P.H. Veltink. Measuring orientation of human body segments using miniature gyroscopes and accelerometers. *Medical and Biological Engineering and Computing*, 43(2):273–282, 2005.
[14] M. Goffredo, I. Bouchrika, J. N. Carter, and M. S. Nixon. Performance analysis for automated gait extraction and recognition in multi-camera surveillance. *Multimedia Tools and Applications*, 50:75–94, 2010.
[15] J. H. Yoo and M. Nixon. Automated markerless analysis of human gait motion for recognition and classification. *ETRI Journal*, 33(3):259–266, 2011.
[16] A. Leu, D. Ristic-Durrant, and A. Graser. A robust markerless vision-based human gait analysis system. In *SACI*, pages 415–420, 2011.
[17] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single depth images. In *CVPR*, pages 1297–1304, 2011.
[18] E. E. Stone and M. Skubic. Evaluation of an inexpensive depth camera for passive in-home fall risk assessment. In *Pervasive Health Conference*, 2011.
[19] E. E. Stone and M. Skubic. Passive in-home measurement of stride-to-stride gait variability comparing vision and Kinect sensing. In *EMBC*, 2011.
[20] M.G. Tsipouras, A.T. Tzallas, G. Rigas, P. Bougia, D.I. Fotiadis, and S. Konitsiotis. Automated levodopa-induced dyskinesia assessment. In *EMBC*, pages 2411–2414, 2010.
[21] I.T. Jolliffe. *Principal Component Analysis*. Springer, 2002.
[22] J. H. Friedman. Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 1999.
[23] J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 1999.
[24] T. Hastie, R. Tibshirani, and J. H. Friedman. *The elements of statistical learning*. Springer, 2009.