

# Vision-based Event Detection of the Sit-to-Stand Transition

Victor Shia

Electrical Engineering and Computer Sciences  
University of California, Berkeley  
vshia@eecs.berkeley.edu

Ruzena Bajcsy

Electrical Engineering and Computer Sciences  
University of California, Berkeley  
bajcsy@berkeley.edu

## ABSTRACT

Sit-to-stand (STS) motions are one of the most important activities of daily living as they serve as a precursor to mobility and walking. However, there exist no standard method of segmenting STS motions. This is partially due to the variety of different sensors and modalities used to study the STS motion such as force plate, vision, and accelerometers, each providing different types of data, and the variability of the STS motion in video data. In this work, we present a method using motion capture to detect events in the STS motion by estimating ground reaction forces, thereby eliminating the variability in joint angles from visual data. We illustrate the accuracy of this method with 10 subjects with an average difference of 16.5ms in event times obtained via motion capture vs force plate. This method serves as a proof of concept for detecting events in the STS motion via video which are comparable to those obtained via force plate.

## Categories and Subject Descriptors

J.3 [Life and Medical Sciences]: Health; I.2.9 [Artificial Intelligence]: Robotics – Kinematics and dynamics

## General Terms

Measurement

## Keywords

Sit-to-stand, Postural Transition

## 1. INTRODUCTION

Sit-to-stand (STS) motions have been regarded as one of the most demanding tasks undertaken during daily living. In fact, difficulty in STS is a risk factor in falls among the elderly and much research has been done to identify those at risk of falling [1].

Since the STS motion is used in clinical practice to determine the ability of an elderly to live independently, it is useful to have a method to standardize assessment of the

motion. In movement laboratories, most STS analyses have used kinematic data or ground reaction forces obtained by motion capture or force plate, respectively. The most frequent methods of segmenting the STS motion are generally described in 3 ways: (1) flexion and extension phases [6, 19, 9], (2) 4 phases using ankle and trunk motion [18, 21], and (3) changes in momentum, velocity or torque [10, 7, 15]. A majority of these studies determine events through vision using joint angles or velocities and suffer from variability in joint angles, contributing to the difficulty of creating a common method of describing events. To study the natural in-home STS motions unconstrained to a laboratory, the use of mobile devices and sensors, such as the Kinect, accelerometers, EMGs, and gyroscopes have been used [5, 11, 13, 20, 3]. However, due to the variability in the STS motion, these studies using mobile devices suffer from similar shortcomings as the visual methods. In [2], the authors propose using ground reaction forces (GRFs) to segment the STS motion. Using GRFs to segment the motion removes the variability in joint angles, instead looking at a single measurement to perform event detection.

This presents the following problem: while GRFs obtained via force plates provide a simple method of segmentation, force plates are impractical to bring into homes and while mobile sensors are inexpensive and prolific, they suffer from variability in STS motion. In this paper, we use a dynamical model of the person and motion capture data to estimate ground reaction forces, thereby avoiding the issue with joint variability; estimate STS events using the estimated GRFs; and compare these with events obtained using GRFs measured via the force plate.

## 2. METHODOLOGY

In this section, we describe how we detect events using a combination of motion capture and dynamics. First, we outline the event detection presented in [2] which we consider the ground truth. Second, we describe the model we use. Third, we describe the event detection algorithm.

### 2.1 Force plate segmentation

The event detection method described in [2] segments the STS motion in the sagittal plane into 6 events: Initiation, Counter, Seat-off, Vertical Peak, Rebound, and Standing. The STS motion begins with the initiation phase, which is defined at the point the subject begins to lean forward. The Initiation phase is followed by the Counter, which is when the subject's feet slightly lift off the ground. The Counter

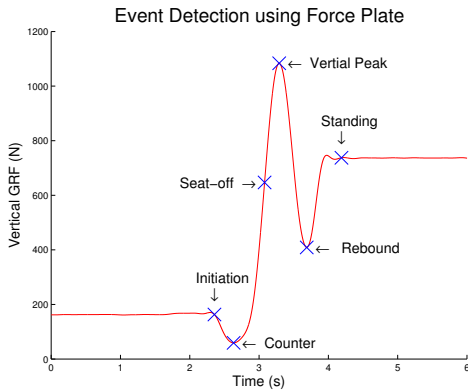


Figure 1: Force plate STS segmentation

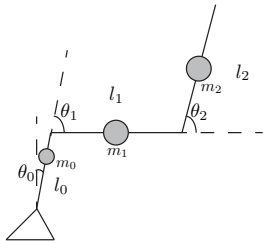


Figure 2: Three link model for the lower body

is followed by Seat-off which is when the subject’s buttocks leaves the seat. The Vertical Peak is when the subject exerts the maximum downward force, which is followed by the Rebound, when the subject is fully extended with upward velocity, and ends with Standing. In [2], the authors determine seat-off using a pressure sensor on the chair and determine standing visually. As we do not have the same setup as the in [2], to determine seat-off, we use the time at which the derivative of the GRF changes from positive to negative. We choose the events by hand to determine the ground truth for the STS events.

## 2.2 Dynamical model

We approach this problem by estimating the GRFs of the STS behavior using a dynamical model of a human’s lower body and using the estimated GRF for event detection. The lower body is modeled using a three segment model in the sagittal plane, shown in Figure 2 to model the human’s shank, thigh and trunk [4, 13]. Limb lengths are given by the motion capture, limb masses are calculated for each individual using tabulated ratios found in [14] and placed along the limb at a length specified in [16]. Using motion capture, we record the angles  $\theta = [\theta_0; \theta_1; \theta_2]$ , velocities  $\dot{\theta} = [\dot{\theta}_0, \dot{\theta}_1, \dot{\theta}_2]$ , and accelerations  $\ddot{\theta} = [\ddot{\theta}_0, \ddot{\theta}_1, \ddot{\theta}_2]$ , corresponding to the angle between the vertical and shank, the shank and thigh, thigh and hip respectively, with positive angles indicating counter-clockwise rotation. A mathematical description of the dynamics is obtained through Lagrangian dynamics given by [12, Chapter 4].

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = \tau \quad (1)$$

where  $L = T - V$  is the Lagrange function,  $T$  is the total kinetic energy of the system  $T(\theta, \dot{\theta}) = \sum_i T_i(\theta, \dot{\theta})$ ,  $V$  is the total potential energy of the system  $V(\theta, \dot{\theta}) = \sum_i V_i(\theta, \dot{\theta})$ ,  $\tau$  is a  $3 \times 1$  vector of generalized torques applied at each joint. Solving the Lagrange equations, we obtain the equation:

$$M(\theta)\ddot{\theta} + C(\theta, \dot{\theta})\dot{\theta} + N(\theta) = \tau \quad (2)$$

where  $M$  is the mass matrix,  $C$  takes into account the coriolis and centrifugal forces, and  $N$  is the potential matrix. Using inverse dynamics, we can obtain the joint torques. We can also recover the ground reaction forces by following the forces in the model through the method described in [17]. In practice, we utilize the Symoro toolbox to obtain the inverse dynamics and GRFs [8]. An example of the estimated vertical GRFs is show in Figure 5.

## 2.3 Event detection

Using this model, we determine STS events using only the estimated GRF. Events are computed using changes in the first derivative of the estimated GRF. We start from the end of the motion and work backwards. First, we detect the standing phase as the point where the standard deviation of the GRFs for the following 1s is below a threshold. Second, the rebound is the first preceding significant local minima. Third, the vertical peak is the preceding local maxima. The counter is the first preceding local minima. Finally, initiation phase is the first preceding local maxima.

## 3. EXPERIMENTS

In this section, we describe our experimental setup, experiment and evaluation method.

### 3.1 Experimental setup

The experimental setup is shown in Figure 3 and consists of an AMTI BP900900 force plate and PhaseSpace Impulse X2 motion capture with 8 infrared cameras. Motion capture data was collected at 480Hz and the subject’s skeleton was extracted using PhaseSpace’s Recap2 software. Ground reaction forces (GRFs) were collected at 2400Hz with the force plate placed under the subject’s foot. Both the motion capture and force plate data were smoothed using a 4th-order Butterworth filter with a cut-off frequency of 4Hz. The chair height was adjusted such that the subject’s thighs were parallel to the ground.

### 3.2 Testing procedure

Subjects wore a customized motion capture suit with PhaseSpace 43 markers, placed on the suit according to the Recap2 software. Subjects were asked to sit in a standardized posture with their trunk starting off vertical, thighs horizontal to the ground, hands on knees (constraining variability of the arms), and shank vertical to the ground<sup>1</sup>. The subject started sitting on the chair with feet placed on the force plate and stood up at the command of the experiment proctor. The subject was asked to rise from the seat upon the proctor’s command and allowed to rise (A) at a natural speed and (B) at a maximum speed. Each STS motion was performed 3 times.

<sup>1</sup>According to [2], constraining the arms to the body results in similar forces to when the arm is unconstrained.



Figure 3: Experimental setup

### 3.3 Evaluation

In this subsection, we describe how the proposed method compares to the ground truth described in Section 2.1. These events were not compared with those events obtained via vision since those events are not compatible with the events detected via force plate.

We ran experiments on 10 individuals with ages ranging from 18-70, weights ranging from 50-80kg, 7 males and 3 females<sup>2</sup>. All individuals did not have any disabilities or known history of physical disability. To evaluate the proposed event detection algorithm, we compute the mean and standard deviation of the difference between the event times obtained via the different methods as well as the median, 1st and 3rd quartile of the time difference. The results are shown in Table 1 and Figure 4. Note that a positive (negative) time difference indicates that the proposed algorithm detected the event after (before) the actual event.

Figure 5 shows three instances of the STS motion with both the force plate forces and motion capture events plotted using the event detection algorithm. While the GRFs computed via dynamics do not have the same magnitude as the force plate, they retain the general shape, allowing it for use in event detection. One possible reason for the GRF discrepancy is that the model does not account for the fact that when sitting, most of the subject’s weight is on the chair. In addition, the proposed algorithm tends to detect the initiation event before the true initiation event. This also may be due to the rigid body model where when sitting, a small motion in the upper body will result in a change in GRF even when the majority of the person’s weight is on the chair.

### 3.4 Limitations

This method currently has only been tested on healthy subjects and has yet to be tested on the general elderly or dis-

<sup>2</sup>This experiment protocol was approved by the UC Berkeley Institutional Review Board: CPHS #2014-07-6506 and all participants gave signed informed consent

Initiation	Counter	Seat-off
-104.8ms (296.2)	49.9ms (198.1)	53.2ms (105.4)
Vertical peak	Rebound	Standing
9.2ms (124.1)	126.5ms (124.9)	6.9ms (221.3)

Table 1: The mean ( stdev ) of differences in event times (milliseconds) from ground truth. Total mean: 16.5ms. Total stdev: 206.4ms

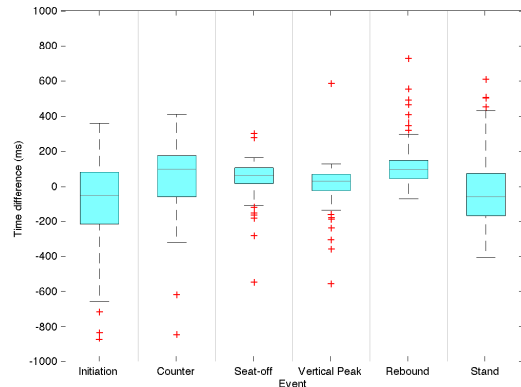


Figure 4: Box plot of time differences for each event. The line represents the median, box represents the 1-3rd quartile, and whiskers represents the extremum points not considered outliers. The red +s represent outliers.

abled population. This method can only detect changes in forces from motion detected visually and will miss events exhibiting no visual motion. In Figure 5a, the counter was missed by the algorithm. Finally, the analysis presented looks solely at the sagittal plane, ignoring the frontal plane, which may also have clinical significance.

## 4. CONCLUSION

We present a method for event detection for the STS motion using vision data and a dynamical model of the subject. This method converts joint angle data obtained via vision to ground reaction forces, thereby bypassing the need to perform event detection on the angles, which is subject to much variability. The events detected by the proposed method have an average of 16.5ms (206.ms stdev) difference from events obtained via the force plate, indicating that it is possible to events using estimated GRF via vision. This method can allow the practitioners to monitor the progress of a patient’s STS motion in a home, which may be a proxy to one’s likelihood to fall, without the need of a force plate, paving the way for remote assistance for the elderly.

Future work includes incorporating body worn accelerometers [22], porting the method to the Microsoft Kinect, using improved estimates of the human’s dynamical properties, and a better model. Porting this method to the Kinect will allow use of this algorithm in independent home and telemedicine settings, allowing the attending physician to monitor the patient remotely. Finally, for this to have clinical applications, we also plan to determine which or how these event times correspond to the likelihood of falling during STS.

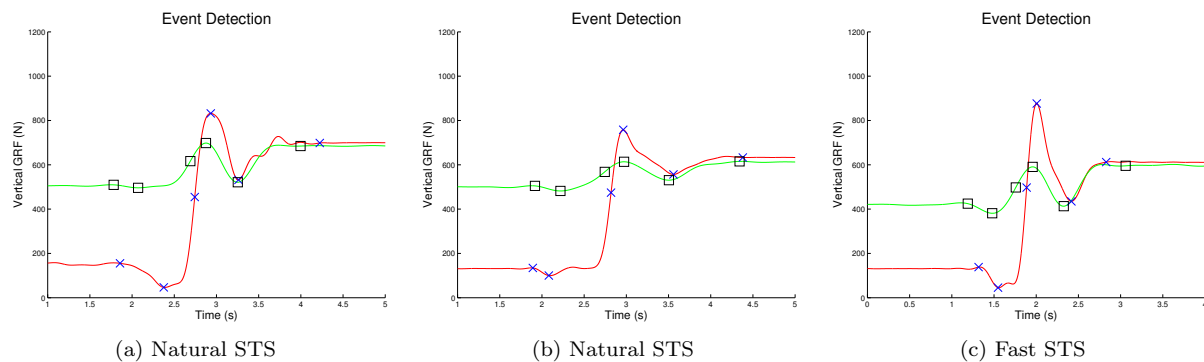


Figure 5: Ground reaction forces for STS at natural speed vs time for 3 users. The red line (blue Xs) denote the forces (events) from the force plate. The green line (black squares) denote the forces (events) computed via the dynamical model.

## 5. ACKNOWLEDGEMENTS

The authors would like to thank Robert Matthew and Gregorij Kurillo for help with the experimental apparatus. This work was supported by the ONR MURI N00014-13-1-0341.

## 6. REFERENCES

- [1] A. J. Campbell, M. J. Borrie, and G. F. Spears. Risk factors for falls in a community-based prospective study of people 70 years and older. *Journal of Gerontology*, 44(4):M112–7, July 1989.
- [2] B. Etnyre and D. Q. Thomas. Event standardization of sit-to-stand movements. *Physical Therapy*, 87:1651–1666, 2007.
- [3] F. J. M. Fernandes. *Sit-to-Stand Movement Analysis using the Kinect Platform*. PhD thesis, Faculdade de Engenharia da Universidade do Porto, 2013.
- [4] H. Hemami and V. C. Jaswa. On a Three-Link Model of the Dynamics of Standing up and Sitting down. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(2):115–120, 1978.
- [5] W. Janssen, D. Kulcu, H. Horemans, H. Stam, and J. Bussmann. Sensitivity of Accelerometry to Assess Balance Control During Sit-to-Stand Movement. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(5):479–484, Oct. 2008.
- [6] S. F. Jeng, M. Schenkman, P. O. Riley, and S. J. Lin. Reliability of a clinical kinematic assessment of the sit-to-stand movement. *Physical Therapy*, 70(8):511–20, Aug. 1990.
- [7] K. Kerr, J. White, D. Barr, and R. Mollan. Analysis of the sit-stand-sit movement cycle in normal subjects. *Clinical Biomechanics*, 12(4):236–245, June 1997.
- [8] W. Khalil, A. Vijayalingam, B. Khomutenko, I. Mukhanov, P. Lemoine, and G. Ecorchard. OpenSYMORO: An open-source software package for Symbolic Modelling of Robots, July 2014.
- [9] T. Kotake. An analysis of sit-to-stand movements. *Archives of Physical Medicine and Rehabilitation*, 74(10):1095–1099, Oct. 1993.
- [10] A. Kralj, R. Jaeger, and M. Munih. Analysis of standing up and sitting down in humans: Definitions and normative data presentation. *Journal of Biomechanics*, 23(11):1123–1138, Jan. 1990.
- [11] N. Millor, P. Lecumberri, M. Gomez, A. Martinez-Ramirez, L. Rodriguez-Manas, F. J. Garcia-Garcia, and M. Izquierdo. Automatic evaluation of the 30-s chair stand test using inertial/magnetic-based technology in an older prefrail population. *IEEE Journal of Biomedical and Health Informatics*, 17(4):820–7, July 2013.
- [12] R. M. Murray, Z. Li, and S. S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1 edition, 1994.
- [13] J. Musić, R. Kamnik, and M. Munih. Model based inertial sensing of human body motion kinematics in sit-to-stand movement. *Simulation Modelling Practice and Theory*, 16(8):933–944, Sept. 2008.
- [14] NASA. Man-Systems Integration Standards - Volume 2, 1995.
- [15] Y.-C. Pai and J. Patton. Center of mass velocity-position predictions for balance control. *Journal of Biomechanics*, 30(4):347–354, Apr. 1997.
- [16] M. J. Pavol, T. M. Owings, and M. D. Grabner. Body segment inertial parameter estimation for the general population of older adults. *Journal of Biomechanics*, 35(5):707–712, May 2002.
- [17] P. Rajai. *Theoretical Framework for Predicting Joint Reaction and Ground Reaction Forces in a Dynamic Pendulum Tree Model of Human Motion*. PhD thesis, New Jersey Institute of Technology, 2007.
- [18] M. Schenkman, R. A. Berger, P. O. Riley, R. W. Mann, and W. A. Hodge. Whole-body movements during rising to standing from sitting. *Physical Therapy*, 70(10):638–48; discussion 648–51, Oct. 1990.
- [19] A. B. Schultz, N. B. Alexander, and J. A. Ashton-Miller. Biomechanical analyses of rising from a chair. *Journal of Biomechanics*, 25(12):1383–1391, Dec. 1992.
- [20] R. Soangra and T. E. Lockhart. A comparative study for performance evaluation of sit-to-stand task with body worn sensor and existing laboratory methods. *Biomedical sciences instrumentation*, 48:407–14, Jan. 2012.
- [21] K. Takara and C. Tanaka. Elaboration and assessment of clinical protocols to support the evaluation of stand-to-sit activity. *Fisioter Movimento*, 27(2):251–259, 2014.
- [22] A. Zijlstra, M. Mancini, U. Lindemann, L. Chiari, and W. Zijlstra. Sit-stand and stand-sit transitions in older adults and patients with Parkinson’s disease: event detection based on motion sensors versus force plates. *Journal of NeuroEngineering and Rehabilitation*, 9(1):75, 2012.