Improved Motion Analysis with the Kinect Camera: A Validation Study Applied to Sit-to-Stand

Sarah Seko[1*], Robert Peter Matthew[1*], Patrick F. Curran[2], Jeannie Bailey[2], Brian Feeley[2], Louis Cheng, Oliver O'Reilly[1], Ruzena Bajcsy[1], Jeffrey C. Lotz[2]
[1]University of California, Berkeley, CA, [2] University of California, San Francisco, CA, *Contributed equally

Disclosures: S. Seko: 5; Bioniks. R. Matthew: 5; Bioniks. P. Curran: None. J.F. Bailey: None. B. Feeley: 4; Bioniks. L. Cheng: 4; Bioniks. O. O'Reilly: 4; Bioniks. R. Bajcsy: None. J.C. Lotz: 4; Bioniks.

INTRODUCTION: Motion-capture based movement analysis can provide objective measures for clinical assessment. The study of joint kinematics and dynamics has broad applications including the identification of pathological motions or compensation strategies and the analysis of dynamic stability. High-end motion capture systems, however, are expensive and require dedicated camera spaces with lengthy set-up and data processing commitments. The Microsoft Kinect 2 is an inexpensive, marker-free depth-camera which provides noisy estimates of specific joint positions. While the cost and ease of use make it an appealing tool for clinical motion analysis, the accuracy is not comparable to high-end motion capture systems.

In this work, we present a method to improve the Kinect measurements through rigid-body constraints on the segment lengths and dynamic filtering methods. The performance of these methods is assessed against a PhaseSpace active motion capture system on the Sit-to-Stand (STS) functional test. STS was selected because it is a common activity of daily living, which requires coordinated motion between many joints. The proposed method is validated through analysis of error in joint angle and velocity measurements during the motion.

METHODS: Motion capture data of STS was recorded from eight healthy subjects under IRB approval with informed consent. Baseline joint angles were estimated from 49 anatomical landmarks tracked by the active-maker PhaseSpace camera system. LED markers were attached directly to the skin using adhesive Velcro, with marker placement following an augmented plugin-gait marker protocol. Time-synchronized data was simultaneously recorded from the PhaseSpace (480 Hz) and Kinect (30 Hz) in a motion capture lab. Data was processed offline using MATLAB. The height of the seat was adjusted such that the subject's knees were flexed at 90º when seated. Subjects performed three trials each consisting of three consecutive sit-to-stand and stand-to-sit motions with their arms folded across their chest.

Ground-truth model: The ground truth skeleton was recovered from the surface landmarks tracked by the PhaseSpace cameras. Each subject first performed a calibration procedure consisting of individual joint rotations. From this data, functional joint centers and individualized body maker positions were computed via sphere-fitting and optimization methods. This model was then used to recover joint center positions and joint states for the STS motions.

Kinect Filtering Algorithm: The raw ankle, knee, hip, and shoulder joint centers from the Kinect were used to drive a rigid-body model which preserves body segment lengths and enforces a fixed position for the ankle joints. Joint states were recovered with an Unscented Kalman Filter [1], a filtering algorithm which incorporates a model of the sensor noise and system dynamics. The constrained body segment lengths for each subject were determined by height-scaled allometric parameters [2]. The raw Kinect skeleton provides a single mid-torso joint at an unspecified position. Our proposed method computes a new joint position at L5S1 based on the configuration of the hip and shoulder joint centers [3].

The 3D joint positions recovered from the PhaseSpace, raw Kinect, and our filtered Kinect method were projected onto the sagittal plane. This produced a planar model with joints at the ankle, knee, hip, L5S1, and shoulder joints (Fig. 1). For each joint, angle and angular velocities were computed using filtering and numerical differentiation. Additionally, a rough measure of the Sagittal Vertical Axis (SVA) was computed by measuring the horizontal distance between the hip joint center and shoulder joint center. Data from each trial was segmented to consider only the STS portion of the motion in analysis.

RESULTS: Data from one subject was excluded due to significant marker occlusion in the PhaseSpace data, resulting in analysis of 63 STS motions across 7 subjects (mean age 28.6 (std. dev. 5.6), 2 female, 5 male). For each STS motion, the mean absolute error (MAE) between the Kinect and PhaseSpace was computed for each kinematic measurement. This error can be interpreted as the average difference between the ground truth and Kinect values over the entire motion. For the velocity measurements, the mean error at the peak velocity was analyzed over the trials. Our results show overall improvement in the accuracy of joint state measures, most prominently in the lower limb joint angles and joint velocities (Table 1).

DISCUSSION: Through rigid-body constraints and filtering, the utility of the Kinect sensor can be enhanced. The presented method shows greater accuracy in recovered joint states as well as a new estimate of hip and L5S1 flexion. Although the error analysis was performed on the planar motion, the recovery is performed in 3D and can be applied to analysis of motion outside of the plane. These accurate kinematic measurements are necessary for the development of dynamic models of joint torques and loading. A limiting factor in the broad applicability of these results is the absence of validation data from pathological motions and a more diverse subject population.

SIGNIFICANCE/CLINICAL RELEVANCE: The proposed kinematic recovery algorithm for the Kinect camera provides greater accuracy and automatic tracking of joint angles and angular velocities and can be used to compute other motion measures of interest. The results of this validation study demonstrate promise in the utilization of the Kinect as a low-cost tool for clinical motion analysis

REFERENCES: [1] Chaffin et al. 2006, [2] Julier and Uhlmann, 1997, [3] Anderson et al., 1986

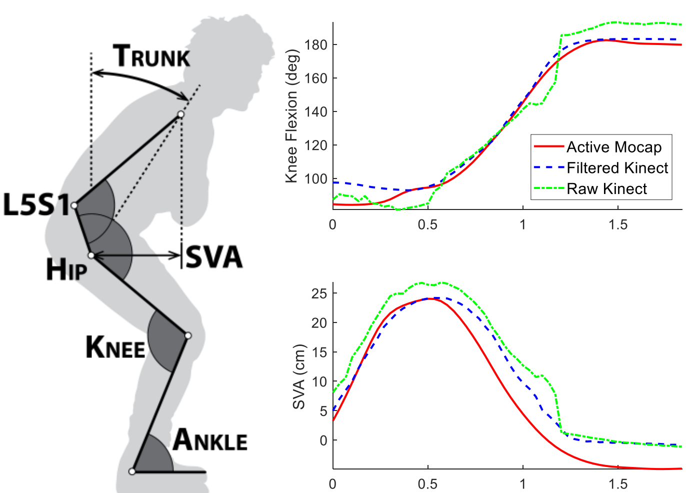

Figure 1. (Left) Sagittal plane model used in motion analysis, and (Right) Illustrative example of recovered trajectories for knee angle (top) and SVA (bottom) for a single STS from PhaseSpace, the proposed filtered Kinect method, and raw Kinect.

Table 1. Summary of error analysis results for the recovered kinematic measurements from the raw Kinect data and the proposed filtering method. Decreased Mean Absolute Error (MAE) across the subjects and trials demonstrates improved accuracy in joint kinematics.

	Raw Kinect MAE (std dev)	**Proposed method MAE (std dev)**
Ankle flexion	12.01 deg (4.01)	7.32 deg (2.11)
Knee flexion	9.24 deg (3.66)	6.13 deg (1.71)
Hip flexion	--	6.11 deg (1.37)
L5S1 flexion	--	11.48 deg (3.08)
Trunk inclination	3.55 deg (1.34)	4.09 deg(1.32)
SVA	3.13 cm (1.18)	3.04 cm (1.01)
Peak ankle ang. velocity	13.33 deg/sec (7.20)	6.30 deg/sec (2.73)
Peak knee ang. velocity	10.46 deg/sec (3.62)	5.62 deg/sec (3.29)
Peak SVA velocity	2.71 cm/sec (2.10)	2.21 cm/sec (1.94)