Mobile Robot Assisted Gait Monitoring and Dynamic Margin of Stability Estimation

Zhuo Chen[1], *Student Member, IEEE*, Huanghe Zhang[2], *Student Member, IEEE*, Antonia Zaferiou[3], *Member, IEEE*, Damiano Zanotto[2], *Member, IEEE*, Yi Guo[1], *Senior Member, IEEE*

***Abstract*—To assess balance control and fall risk, it is desirable to continuously monitor dynamic stability during walking tasks. Dynamic Margin of Stability (MoS) is widely recognized as a quantitative measure for human walking stability and gait balance strategies. We propose a mobile robot assisted gait monitoring system that precedes human subjects in overground walking. Real-time data from the RGB-D Kinect sensor on the robot are fused with measurement from pressure sensors and inertial measurement units in a pair of instrumented footwear, and Kalman filter based methods are developed to estimate MoS and spatiotemporal gait parameters in real time. Experimental results with 10 subjects are compared with those obtained by a gold-standard motion capture system. Results show that the proposed method achieves acceptable accuracy of MoS estimation and high accuracy for spatio-temporal gait parameters. Whereas existing works on MoS assessment use wearable sensors that can only provide offline analysis, our proposed system provides real time gait monitoring and MoS estimation that could potentially assess fall risk during walking in out-of-lab conditions.**

***Index Terms*—Assistive robots, wearable sensors, gait, dynamic balance, margin of stability.**

I. Introduction

Human walking is a challenging task and the maintenance of balance requires both spatial and temporal control of stepping [1]. For older adults, the risk of falls is higher due to muscle weakness, tendon stiffness [2], chronic diseases [3], vestibular dysfunctions [4], and impaired vision [5]. About one third of adults over age 65 experience at least one fall annually [6]. Considerable medical costs have been attributed to fall related injuries among older adults [7], creating a major public heath problem. To prescribe appropriate interventions, it is important to identify individuals who are at greater risk of falls [8]. Thus, it is desirable to have cost-effective approaches to assessing fall risk during walking tasks.

Dynamic Margin of Stability (MoS) was first proposed in [9] as a measure of stability in human movement control, and it plays a crucial role in assessment of fall risk and balance strategies. Modeling human as an inverted pendulum, Hof *et al.* [9] postulates that the condition for maintaining balance is that the Extrapolated Center of Mass (XCoM) falls inside the Base of Support (BoS). MoS describes the distance between XCoM and the boundary of BoS, allowing for quantitative analysis of gait and balance during everyday walking as well as objective quantification of recovery progress of people with motor impairment [10]. Such objective quantification was shown in [11] to have significant correlation with the Mini-BESTest [12], a validated clinical balance assessment tool. Furthermore, it was suggested in [13] that MoS, unlike other stability measures such as Lyapunov exponent, makes it possible to assess the dynamic stability at certain gait events through a gait cycle. In particular, Ohtsu *et al.* [11] has shown that mediolateral MoS (MoS ML) at foot-flat and toe-off and anteroposterior MoS (MoS AP) just before heel strike could potentially identify older adults with balance disorders. As an extension to MoS, Mademli *et al.* [14] proposed a safety factor as the mean MoS AP at walk-to-run transition velocity divided by the mean MoS AP at preferred walking velocity. It was shown that lower safety factor is associated with increased risk of falls among older adults.

While camera-based motion capture systems are considered the gold-standard in determining the MoS [13], [15], this equipment is expensive and can only be used in laboratory settings. It requires markers to be taped to the body and entails a lengthy process of manual labeling and gap filling. To allow for continuous monitoring of MoS during users' daily living, Refai *et al.* [10], van Meulen *et al.* [16], Guaitolini *et al.* [17] and Fino *et al.* [18] developed methods using wearable devices. In particular, an Extended Kalman Filter (EKF) was proposed in [16] to estimate MoS by fusing data from instrumented shoes with 3D force and moment sensors, Inertial Measurement Units (IMUs), and ultrasonic transducers. Although this MoS monitoring system is less expensive and offers higher mobility than the camera-based system, the bulky 3D force and moment sensors make the footwear uncomfortable to walk in and are likely to alter the wearer's natural gait, thereby defeating the purposes of the system as a balance monitoring tool. To improve upon this method, Refai *et al.* [10] proposed to use subject-specific regression models to estimate 3D forces and moments from 1D plantar pressure data as measured by lightweight and inconspicuous instrumented insoles. However, the calibration of the subject-specific models in [10] still relies on laboratory equipment, and neither [16] nor [10] validated their results against camera-based systems. In [17], a network of seven Magneto-IMUs attached to pelvis and lower limbs were used

This work was partially supported by the US National Science Foundation under Grant IIS-1838799.

Corresponding author (yguo1@stevens.edu).

[1]Zhuo Chen and Yi Guo are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA. Emails: {zchen39,yguo1}@stevens.edu

[2]Huanghe Zhang and Damiano Zanotto are with the Department of Mechanical Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA. Emails: {hzhang82,dzanotto}@stevens.edu

[3]Antonia Zaferiou is with the Department of Biomedical Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA. Email: azaferio@stevens.edu

to estimate the MoS, achieving a Root-Mean-Square Error (RMSE) of 1.80cm or smaller for walking on a treadmill at constant speeds. Despite the high accuracy, the system requires optical motion tracking system in a structured environment for static anthropometric measurement and sensor-to-segment calibration [17]. To achieve calibration-free IMU-based MoS assessment, Fino *et al.* [18] proposed to use the centripetal acceleration measured by a lumbar-attached IMU as a lateral MoS correlate. However, [18] evaluates the MoS ML only at initial contact rather than over the complete gait cycle and does not address the MoS AP. Furthermore, none of the work in [10], [16], [17], [18] models different segment of the foot (i.e., forefoot and hindfoot) in contact with the ground and take the phases of the stance period into consideration when determining the BoS. Yet, as pointed out in [11], a more accurate estimation of the phase-dependent changes in the BoS in relation to the XCoM during walking may facilitate a deeper understanding of the gait strategies adopted by individuals with balance disorders, and their associations with fall risk.

To overcome the limitations of wearable devices, we propose to use a mobile robot to monitor and assess gait stability. Assistive robots have been extensively studied in the context of elderly care, e.g. exoskeleton robots [19], [20], and social companion robots [21], [22]. Mobile robots equipped with RGB-D Kinect sensors [23] have been developed for walking gait monitoring, where spatio-temporal gait parameters were estimated. A robotic rollator equipped with Kinect and Laser Range Finder (LRF) sensors was proposed in [24] to predict gait stability through estimating the upper body pose and tracking the Center of Mass (CoM) along with an augmented gait state estimation. However, the method in [24] only classifies human gait stability as either safe or unsafe, and does not provide quantitative measure for human walking stability. In our previous work [25], we estimated spatio-temporal gait parameters using an integrated mobile robot and wearable sensor system, and machine learning methods were developed for *offline* gait analysis. In this work, we extend the capability of our integrated system to *online* estimation of dynamic MoS, which provides quantitative measure for human walking stability and dynamic balance strategies.

In this paper, we propose a MoS monitoring system consisting of a mobile robot (equipped with a Kinect sensor) and a pair of instrumented insoles, each embedding an array of pressure sensors and an IMU. We design Kalman filter based methods to fuse the data from these sensors to estimate foot poses and the body CoM in the presence of measurement uncertainties and noises. The estimates of the Kalman filters are then used to determine the MoS during walking. We conduct experiments where the mobile robot precedes human subjects in overground walking and collects data for real-time gait and MoS monitoring. The results are compared with those obtained using a motion capture reference system in terms of RMSE. The statistical analysis indicates acceptable accuracy of MoS estimation and high accuracy for conventional spatio-temporal gait parameters.

The primary contribution of this work is the design and validation of a mobile robotic system for monitoring the MoS in real time, which could potentially provide fall-risk assessment during walking in out-of-lab conditions. Novel Kalman filter based methods are developed to fuse the data from the Kinect sensor and the instrumented insoles. The inclusion of the instrumented footwear subsystem allows us to assess the fine-grained gait phases of the stance period with high time resolution, which is important in fall-risk assessment. This study provides a new solution to MoS estimation on a portable robot and wearbale sensor platform. Unlike existing work on MoS estimation using wearable devices [10], [17], [18], real time gait and MoS estimation are achieved and performance results are validated during *overground walking*.

The rest of the paper is organized as follows. Section II describes the proposed system and the experiment protocol. Section III presents the method used to estimate MoS for dynamic balance sensing. We then present the performance evaluation of our proposed system in Section IV. Finally, we conclude our work in Section V.

II. System Description

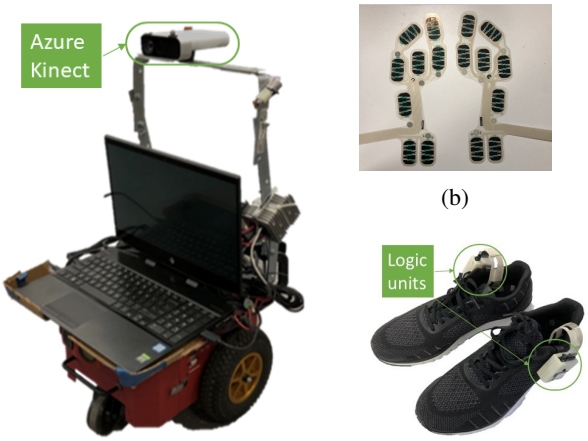
The mobile gait measurement system consists of a mobile robot subsystem (Fig. 1 (a)) and an instrumented footwear subsystem (SportSole, Fig. 1 (b) (c), [26], [27]). We describe each of the subsystems and our experimental protocol in this section.

A. Mobile Robot Subsystem

The mobile robot subsystem is a customized P3-DX differential drive robot equipped with a laptop computer (Intel Core i7-9750H CPU, Nvidia RTX 2060 GPU), a backward facing Azure Kinect sensor [28] for gait monitoring, and a forward facing Kinect v1 sensor for mapping and localization. The Azure Kinect sensor consists of an RGB camera, a time-of-flight depth camera, and an IMU. It is mounted at a height of 0.80 m from the ground. The desired human-robot distance should be between 1.0 m and 2.5 m so that the entire human body is within the field of view of the Kinect sensor yet not exceeding the depth camera operating range (2.88 m). Within such a range, a shorter human-robot distance can increase the number of depth and RGB image pixels of the human body and potentially improve the body tracking accuracy.

B. Instrumented Footwear Subsystem

The instrumented footwear subsystem ([26], [27]) consists of a pair of insoles (Fig. 1 (b)), each featuring 8 piezoresistive cells (from IEE S.A., Luxemburg), underneath the left and right calcaneus, the lateral arch, the head of the first, third and fifth metatarsals, the hallux, and the toes, respectively. Additionally, an IMU (Yost Labs Inc., Portsmouth, OH) is embedded in each insole under the medial arch. As shown in Fig. 1 (c), a Li-Po-battery-powered logic unit (clipped laterally on each shoe) is composed of a 32-bit ARM Cortex-M4 microcontroller and a Wi-Fi module. It receives data from the pressure sensors and the IMU, and transmits the data to the laptop computer on the robot via UDP.



(a) (b) (c)

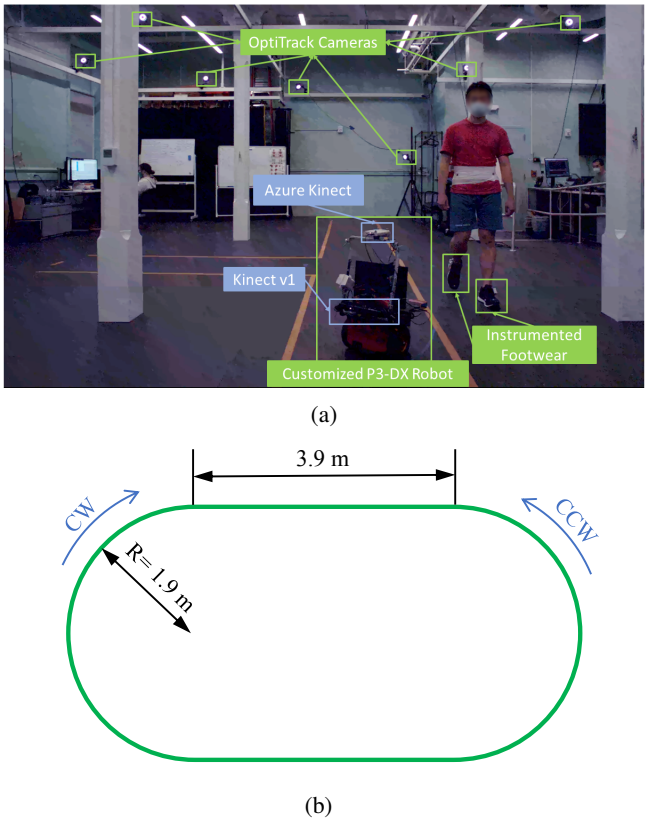
Fig. 1: The proposed system. (a) Mobile robot subsystem with a P3-DX robot base and an Azure Kinect sensor on top; (b) A pair of insoles with IMU and 8-cell plantar pressure sensors in the instrumented footwear subsystem; and (c) logic units (enclosed by green circles) of the instrumented footwear subsystem being strapped to shoes.

C. Subsystem Communication

The robot and the logic units of the instrumented footwear communicate in a local area network through a wireless router. The laptop computer on the robot controls the robot's motion and processes the data from both the robot and instrumented footwear subsystems. Programs for robot motion control, data acquisition and gait analysis run as Robot Operating System (ROS) nodes in Ubuntu 18.04 with ROS Melodic. The Body Tracking SDK can process the RGB-D images at 30 Hz while each instrumented insole has a sampling rate of 240 Hz.

D. Experimental Protocol

Fig. 2 shows our experimental setup for the purpose of data collection and performance validation. We recruited 10 healthy subjects (aged 24 through 29 years) to participate in the experiment. All participants provided written informed consent and the study protocol was approved by the Institutional Review Board (IRB) of Stevens Institute of Technology. The subjects were instructed to walk around a pre-defined oval track marked by blue tapes on the floor as shown in Fig. 2. The subjects were also instructed to walk at their preferred speed behind the robot while wearing the instrumented insoles. In order for the subject to be centered in the field of view behind the robot during curved walking, the robot used a smaller radius of curvature. The robot was programmed to keep a fixed distance from the subject. Each subject was given about five minutes to become familiar with the task. After familiarization, each subject walked three laps in each direction (clockwise and counterclockwise) and then repeated this procedure (12 total laps per subject). We used a commercially available optical motion capture system (OptiTrack) as the reference system to validate our mobile gait monitor system. The details



(a) (b)

Fig. 2: Robot-guided walking. (a) Our mobile gait monitoring system consists of a mobile robot (with an Azure Kinect sensor) and a pair of instrumented footwears. The OptiTrack cameras serve as the reference system for performance validation. (b) The walking track for the subject consists of two semicircles of radius 1.9 m and two 3.9 m straight line segments.

on OptiTrack data processing and performance validation are presented in Section IV.

III. Mobile Robot Assisted MoS Estimation

A. Robot Localization and Control

The mobile robot uses the forward-facing RGB-D sensor, i.e., Kinect v1 (Kinect for XBox 360), for mapping the environment and localizing itself in the environment. We use an existing visual Simultaneous Localization and Mapping (SLAM) package, RTAB-Map (Real-Time Appearance-Based Mapping [29]), to fuse data from wheel odometry and the forward-facing Kinect v1 sensor. Comparing with other SLAM methods, the memory management method used by the loop closure detection of RTAB-Map can easily satisfy the real time requirement regardless of the map size [29].

Given the human track shown in Fig. 2 (b), we first plan a robot path accordingly so that when the robot moves on the planned path, the human subject is at the center of the field of view of the Azure Kinect sensor provided he/she walks on the pre-designated human track and the human-robot distance is $\rho^* = 1.4$ m. We then use an existing online local motion planner [30] to control the robot to track the pre-planned path.

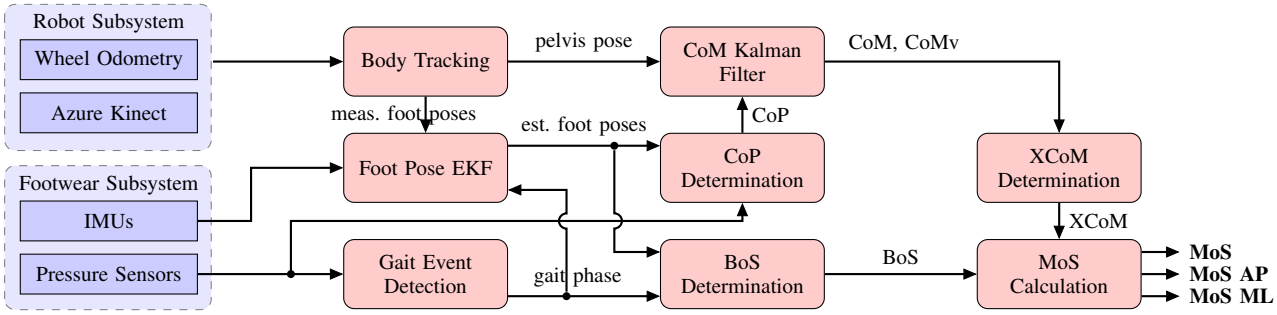


Fig. 3: Functional block diagram of our proposed method to determine the MoS values.

To maintain the human-robot distance to be at $\rho^* = 1.4$ m, we use a proportional-integral controller to control the speed of the robot so that the error between the measured and the desired human-robot distance, $\rho - \rho^*$, approaches zero.

B. Overview of MoS Estimation Approach

As a measure for dynamic balance of human walking, MoS is defined in [9] as the distance between the XCoM, the vertical projection of the CoM plus its velocity (denoted as CoMv) times a factor, and the boundary of BoS. Based on the MoS, balance during a gait cycle is decomposed into MoS AP and MoS ML. The functional block diagram of our proposed MoS estimation method is shown in Fig. 3. The Azure Kinect sensor on the mobile robot tracks the body joints of the human subject during walking, and the pelvis position is used to estimate the subject's CoM and CoMv for the purpose of XCoM determination. The determination of BoS relies on the gait event detection using the pressure sensor measurement from the insoles and the output of the foot pose EKF using the IMU sensor measurement from the insole. Finally, when the BoS and XCoM are available, we can calculate the MoS values. The above-mentioned steps are detailed in the following subsections.

It is worth mentioning that the proposed system poses technical challenges including fusing data from a RGB-D sensor on a mobile platform with portable insoles. We develop a novel Kalman filter based method that can monitor spatiotemporal gait parameters and estimate MoS using the onboard computational power of the robot.

C. Body Tracking

In this subsection, we describe how the pelvis position and foot poses are estimated from the measurement of the Azure Kinect Body Tracking SDK which measures the poses of 32 body joints[1]. Snapshots of the body joint poses observed by the Azure Kinect during different gait phases are visualized in Fig. 4 using RViz, a 3D visualization tool for ROS applications.

Let the homogeneous transformation matrix to transform vectors from coordinate frames B to C be represented as

$$ {}^C_B\boldsymbol{T} = \begin{bmatrix} {}^C_B\boldsymbol{R} & {}^C\boldsymbol{p}_B \\ \boldsymbol{0}_{1\times 3} & 1 \end{bmatrix} \in \boldsymbol{SE}(3) \tag{1} $$

[1]The definition for body joints can be found https://docs.microsoft.com/en-us/azure/Kinect-dk/body-joints.

where ${}^C_B\boldsymbol{R}$ is a 3×3 rotation matrix representing the orientation of frame B with respect to frame C, and ${}^C\boldsymbol{p}_B \in \mathbb{R}^3$ is the position of the origin of frame B in frame C's coordinates.

Let the robot frame, the Kinect sensor frame, and the world frame be denoted with R, K, and W, respectively. The Kinect sensor measures the pelvis position ${}^K\boldsymbol{p}$ and ankle joint poses ${}^K_{A_i}\boldsymbol{T}$, where A_i denotes the ankle joint frames with subscript $i \in \{\mathrm{l}, \mathrm{r}\}$ indicating left or right. These body joint measurements can be transformed to frame W by the transformation matrix

$$ {}^W_K\boldsymbol{T}(t) = {}^W_R\boldsymbol{T}(t)\ {}^R_K\boldsymbol{T}, \tag{2} $$

where t denotes time. As the robot moves on a flat ground in 2D, the transformation ${}^W_R\boldsymbol{T}(t)$ can be parameterized by position $x_R(t)$, $y_R(t)$ and heading $\theta_R(t)$ and be determined by the wheel odometry with a sampling frequency of 10 Hz. Since the Kinect sensor is mounted rigidly on the robot, the static transform ${}^R_K\boldsymbol{T}$ can be calibrated offline. Note that the cumulative dead-reckoning error in ${}^W_R\boldsymbol{T}(t)$ has minimal effects on the MoS values as the distance from CoM to BoS is not related to ${}^W_R\boldsymbol{T}(t)$.

1) CoM Measurement: The projected CoM measurement is the pelvis position that is first measured by the Kinect sensor, transformed to frame W using (2), and then projected to the ground. We use $\boldsymbol{r}$ to denote the projected CoM (described further in Section III-F).

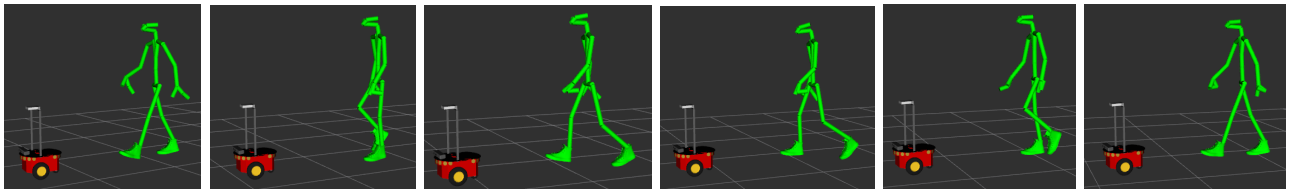
2) Foot Pose Measurement: Let M_i denote the insole IMU frame. The Kinect sensor can measure the pose of M_i in frame W as

$$ {}^W_{M_i}\boldsymbol{T}(t) = {}^W_K\boldsymbol{T}(t)\ {}^K_{A_i}\boldsymbol{T}(t)\ {}^{A_i}_{M_i}\boldsymbol{T} \tag{3} $$

where the transform ${}^{A_i}_{M_i}\boldsymbol{T}$ varies with shoe (insole) size. With (3), we can estimate the position of the hindfoot reference points $\boldsymbol{r}_{\mathrm{h}i} \in \mathbb{R}^2$ and forefoot reference points $\boldsymbol{r}_{\mathrm{f}i} \in \mathbb{R}^2$ in the horizontal plane.

D. Gait Event Detection

Taking advantage of the high sampling rate of the insole pressure sensor, we can detect gait events for each foot that include heel-strike, foot-flat, heel-off and toe-off [31] with high temporal resolution. Specifically, heel-strike and heel-off events are triggered when the sum of the readings of the two heel pressure cells crosses a threshold; foot-flat and toe-off events are triggered when the sum of the readings of the three



(a) Initial contact (b) Midstance (c) Terminal stance (d) Preswing (e) Midswing (f) Terminal swing

Fig. 4: Snapshots of RViz visualization demonstrating body tracking and the 6 gait phases of the left foot. The body segments (green cylinders) are visualized from the output of the Body Tracking SDK. The foot poses (green shoes) are estimated by the Foot Pose EKF.

metatarsal and two toe pressure cells crosses the threshold. The detected gait events drive the transition of a finite state machine from one gait phase to another as illustrated in Fig. 5.

E. Foot Pose EKF

In this subsection, we describe our novel foot pose EKF method that fuses the measurement from the Kinect and insole measurements to provide a better foot pose estimate than either system used alone. It takes the measurement from the insole IMU and the Kinect sensor, follows the steps of initialization, prediction and update, and returns the estimate of a state vector $\boldsymbol{x}_{M_i}$ that is defined as $\boldsymbol{x}_{M_i} = \left[\boldsymbol{q}_{M_i}^{\mathrm{T}},\, \boldsymbol{\omega}_{M_i}^{\mathrm{T}},\, \boldsymbol{p}_i^{\mathrm{T}},\, \boldsymbol{v}_i^{\mathrm{T}},\, \boldsymbol{a}_i^{\mathrm{T}},\, \boldsymbol{\omega}_{bM_i}^{\mathrm{T}}\right]^{\mathrm{T}}$, where $i \in \{\mathrm{l}, \mathrm{r}\}$ denotes each of the left and the right foot, $\boldsymbol{q}_{M_i}$ is the quaternion representing the rotation from world frame W to M_i; $\boldsymbol{\omega}_{M_i}$ is the angular velocity of the insole IMU expressed in the local frame M_i; $\boldsymbol{p}_i$, $\boldsymbol{v}_i$, and $\boldsymbol{a}_i$ are the position, linear velocity, and linear acceleration of the insole IMU expressed in world frame W, respectively; $\boldsymbol{\omega}_{bM_i}$ is the bias of the gyroscope expressed in the local frame M_i.

1) State Prediction: The system equation is the $\boldsymbol{SE}(3)$ kinematics augmented by the gyroscope bias process [32]. It can be written in discrete time as

$$\boldsymbol{x}_{M_i,k+1} = \boldsymbol{\phi}_1\left(\boldsymbol{x}_{M_i,k}\right) + \boldsymbol{w}_{M_i,k} \tag{4}$$

where

$$\boldsymbol{\phi}_1\left(\boldsymbol{x}_{M_i,k}\right) = \boldsymbol{x}_{M_i,k} + \begin{bmatrix} \boldsymbol{q}_{M_i,k} \otimes \boldsymbol{\omega}_{M_i,k} \frac{\Delta t_k}{2} \\ \boldsymbol{0} \\ \boldsymbol{v}_{i,k} \Delta t_k \\ \boldsymbol{a}_{i,k} \Delta t_k \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix},$$

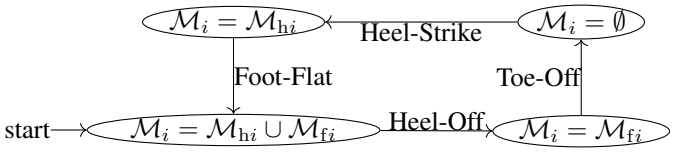


Fig. 5: Finite state machine driven by the gait events for determining the set $\mathcal{M}_i$ of virtual markers for each foot $i \in \{\mathrm{l}, \mathrm{r}\}$ that determine the BoS polygon.

and the subscripts k and $k+1$ denote the time steps, $\Delta t_k = t_{k+1} - t_k$, $\boldsymbol{w}_{M_i,k} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_1)$ is the Gaussian noise with zero-mean and covariance matrix $\boldsymbol{Q}_1 = \mathrm{diag}(\boldsymbol{Q}_{\boldsymbol{q}}, \boldsymbol{Q}_{\boldsymbol{\omega}}, \boldsymbol{Q}_{\boldsymbol{p}}, \boldsymbol{Q}_{\boldsymbol{v}}, \boldsymbol{Q}_{\boldsymbol{a}}, \boldsymbol{Q}_{\boldsymbol{\omega}_b}) = \mathrm{diag}(1.0 \times 10^{-6}\boldsymbol{I}_4, 4.0 \times 10^{-4}\boldsymbol{I}_3, 1.0 \times 10^{-6}\boldsymbol{I}_3, 1.0 \times 10^{-2}\boldsymbol{I}_3, 4.9 \times 10^{1}\boldsymbol{I}_3, 1.0 \times 10^{-10}\boldsymbol{I}_3)$.

The state and covariance predictions for the EKF are given in general as

$$\hat{\boldsymbol{x}}_{M_i,k+1|k} = \boldsymbol{\phi}_1\left(\hat{\boldsymbol{x}}_{M_i,k|k}\right) \tag{5a}$$

$$\hat{\boldsymbol{P}}_{M_i,k+1|k} = \boldsymbol{F}_{M_i,k} \hat{\boldsymbol{P}}_{M_i,k|k} \boldsymbol{F}_{M_i,k}^{\mathrm{T}} + \boldsymbol{Q}_1 \Delta t_k \tag{5b}$$

where $\hat{\boldsymbol{x}}_{M_i,k+1|k}$ is the a priori estimate of the state at time t_{k+1} given the a posteriori estimate $\hat{\boldsymbol{x}}_{M_i,k|k}$ of the state at time t_k; the matrix $\boldsymbol{F}_{M_i,k}$ is the state transition Jacobian; the matrices $\hat{\boldsymbol{P}}_{M_i,k+1|k}$ and $\hat{\boldsymbol{P}}_{M_i,k|k}$ are the covariance matrices for the a posteriori and a priori state estimates, respectively. The state and covariance are predicted every time when either Kinect or insole IMU measurement becomes available.

2) Insole IMU Measurement Update: The measurement update from the insole IMU depends on the gait phase detection. For the rest of this subsection, subscripts including time step k, reference frame M_i and index i are dropped from measurement and state variables with the understanding that the insole measurement is taken in the reference frame M_i. The measurement equations for the insole accelerometer and gyroscope are

$$\boldsymbol{z}_{\mathrm{acc}} = \boldsymbol{h}_{\mathrm{acc}}(\boldsymbol{x}) + \boldsymbol{\nu}_{\mathrm{acc}} = \boldsymbol{q} \otimes (\boldsymbol{a} + \boldsymbol{g}) \otimes \boldsymbol{q}^* + \boldsymbol{\nu}_{\mathrm{acc}} \tag{6a}$$

$$\boldsymbol{z}_{\boldsymbol{\omega}} = \boldsymbol{h}_{\boldsymbol{\omega}}(\boldsymbol{x}) + \boldsymbol{\nu}_{\omega} = \boldsymbol{\omega}_b + \boldsymbol{\omega} + \boldsymbol{\nu}_{\omega} \tag{6b}$$

where $\boldsymbol{z}_{\mathrm{acc}}$ and $\boldsymbol{z}_{\boldsymbol{\omega}}$ are the accelerometer and gyroscope readings, respectively, $\boldsymbol{g}$ is the local acceleration due to gravitation (pointing up in the direction of Z axis of frame W), and $\boldsymbol{\nu}_{\mathrm{acc}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\mathrm{acc}})$ and $\boldsymbol{\nu}_{\omega} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\omega})$ are measurement noises with measurement covariance matrices $\boldsymbol{R}_{\mathrm{acc}} = 1.0\boldsymbol{I}_3, \boldsymbol{R}_{\boldsymbol{\omega}} = 2.5 \times 10^{-3}\boldsymbol{I}_3$.

During the stance period, we assume that the foot segment that is in contact with the ground does not slip, in which case measurement equations (7) can be used to correct the measurement update.

$$\boldsymbol{z}_{\boldsymbol{v}} = \boldsymbol{h}_{\boldsymbol{v}}(\boldsymbol{x}) + \boldsymbol{\nu}_{\boldsymbol{v}} = \boldsymbol{v} + \boldsymbol{\nu}_{\boldsymbol{v}} \tag{7a}$$

$$\boldsymbol{z}_{\boldsymbol{a}} = \boldsymbol{h}_{\boldsymbol{a}}(\boldsymbol{x}) + \boldsymbol{\nu}_{\boldsymbol{a}} = \boldsymbol{a} + \boldsymbol{\nu}_{\boldsymbol{a}} \tag{7b}$$

where $\boldsymbol{\nu}_{\boldsymbol{v}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{v}})$ and $\boldsymbol{\nu}_{\boldsymbol{a}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{a}})$ are measurement noises.

During the midstance phase, we can perform a measurement update with linear velocity and acceleration measurements set to zero with high confidence:

$$\boldsymbol{z_v} = \mathbf{0}, \quad \boldsymbol{R_v} = 1.0 \times 10^{-10} \boldsymbol{I}_3, \tag{8}$$
$$\boldsymbol{z_a} = \mathbf{0}, \quad \boldsymbol{R_a} = 1.0 \times 10^{-10} \boldsymbol{I}_3. \tag{9}$$

From heel strike until the beginning of midstance phase, the foot revolves about a fixed point on the ground/calcaneous. Let $\boldsymbol{d}_1$ be a vector starting at the contact point and ending at the origin of frame M_i, the state can be updated with a linear velocity measurement

$$\boldsymbol{z_v} = \boldsymbol{\omega} \times \boldsymbol{d}_1, \quad \boldsymbol{R_v} = 1.0 \times 10^{-4} \boldsymbol{I}_3. \tag{10}$$

Similarly, from heel off until toe off (more precisely, before the metatarsal leaves the ground), the foot rotates about an axis passing through the head of the second metatarsal. With $\boldsymbol{d}_2$ being a vector pointing from metatarsal to the insole IMU, the state can be updated with a linear velocity measurement

$$\boldsymbol{z_v} = \boldsymbol{\omega} \times \boldsymbol{d}_2, \quad \boldsymbol{R_v} = 1.0 \times 10^{-4} \boldsymbol{I}_3. \tag{11}$$

3) Kinect Sensor Measurement Update: Knowing ${}^{W}_{M_i}\boldsymbol{T}$ from (3), we have the measurement equations for position $\boldsymbol{p}$, the projection $\boldsymbol{h}_y(\boldsymbol{x})$ of the unit vector in the direction of Y axis of frame W onto frame M_i, and the quaternion $\boldsymbol{q}$ as

$$\boldsymbol{z_p} = \boldsymbol{h_p}(\boldsymbol{x}) + \boldsymbol{\nu}_p = \boldsymbol{p} + \boldsymbol{\nu}_p \tag{12a}$$
$$\begin{aligned} \boldsymbol{z}_y &\triangleq \boldsymbol{z_q} \otimes [0,1,0]^{\mathrm{T}} \otimes \boldsymbol{z_q^*} \\ &= \boldsymbol{h}_y(\boldsymbol{x}) + \boldsymbol{\nu}_y = \boldsymbol{q} \otimes [0,1,0]^{\mathrm{T}} \otimes \boldsymbol{q}^* + \boldsymbol{\nu}_y \end{aligned} \tag{12b}$$
$$\boldsymbol{z_q} = \boldsymbol{h_q}(\boldsymbol{x}) + \boldsymbol{\nu}_q = \boldsymbol{q} + \boldsymbol{\nu}_q \tag{12c}$$

where $\boldsymbol{z_p}$ and $\boldsymbol{z_q}$ are the position and quaternion measurements extracted from (3); $\boldsymbol{\nu_p} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R_p})$, $\boldsymbol{\nu}_y \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_y)$ and $\boldsymbol{\nu_q} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R_q})$ are measurement noises with measurement covariance matrices: $\boldsymbol{R_p} = 1.0 \times 10^{-4}\boldsymbol{I}_3, \boldsymbol{R}_y = 2.5 \times 10^{-3}, \boldsymbol{R_q} = 4.0 \times 10^{-4}\boldsymbol{I}_4$.

Note that (12b) ensures that the projection of the gravitational vector onto the IMU frame is solely responsible for correcting the pitch and roll angles, which tend to be more reliable than the Kinect sensor observation. Also note that there is heading discrepancy between the insole IMU and the ankle joint frame measured by the Body Tracking SDK, which causes $\boldsymbol{\nu}_y$ to have a non-zero mean. Hence, during the swing phase, instead of relying on (3), we can get the unbiased measurement $\boldsymbol{z_q}$ from the prior state estimate $\hat{\boldsymbol{q}}_{k+1|k}$ as

$$\boldsymbol{z_q} = \hat{\boldsymbol{q}}_{k+1|k} \otimes \boldsymbol{\Omega}\left(\mu\left(\boldsymbol{z}_{\boldsymbol{p},k+1} - \boldsymbol{h_p}(\hat{\boldsymbol{x}}_{k+1|k})\right) \times \hat{\boldsymbol{v}}_{k+1|k}\right) \tag{13}$$

where μ is a constant, and $\boldsymbol{\Omega}(\boldsymbol{\alpha})$ maps a non-zero 3-vector $\boldsymbol{\alpha}$ to a quaternion with $(\boldsymbol{\alpha}/||\boldsymbol{\alpha}||)\sin(||\boldsymbol{\alpha}||)$ as the vector part and $\cos(||\boldsymbol{\alpha}||)$ as the scalar part. Using measurement (13) and observation model (12c), the measurement update is effective only when $||\hat{\boldsymbol{v}}_{k+1|k}||$ is sufficiently large; if the person stands still, (13) and (12c) are ineffective and it is necessary to use (12b) to perform measurement update.

4) Update Equations: The update equations for the measurement models (6), (7) and (12) are

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \boldsymbol{h}\left(\hat{\boldsymbol{x}}_{k+1|k}\right)\right) \tag{14a}$$
$$\hat{\boldsymbol{P}}_{k+1|k+1} = \left(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k\right)\hat{\boldsymbol{P}}_{k+1|k} \tag{14b}$$

where the Kalman gain is

$$\boldsymbol{K}_k = \hat{\boldsymbol{P}}_{k+1|k}\boldsymbol{H}_k^{\mathrm{T}}\left(\boldsymbol{H}_k\hat{\boldsymbol{P}}_{k+1|k}\boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{R}_1\right)^{-1},$$

function $\boldsymbol{h}$ can take $\boldsymbol{h}_{\mathrm{acc}}$, $\boldsymbol{h_\omega}$, $\boldsymbol{h_v}$, $\boldsymbol{h_a}$, $\boldsymbol{h_p}$, $\boldsymbol{h}_y$ and $\boldsymbol{h_q}$, measurement noise covariance $\boldsymbol{R}_1$ can take $\boldsymbol{R}_{\mathrm{acc}}$, $\boldsymbol{R_\omega}$, $\boldsymbol{R_v}$, $\boldsymbol{R_a}$, $\boldsymbol{R_p}$, $\boldsymbol{R}_y$ and $\boldsymbol{R_q}$, the measurement Jocobian can be calculated as

$$\boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\hat{\boldsymbol{x}}_{k+1|k}}$$

5) Quaternion Repair: As the quaternion representation of rotation in $\boldsymbol{SO}(3)$ is overparameterized, the state prediction and measurement update may, due to quantization error, yield a quaternion that is not normalized. Also, there exist two normalized quaternions that correspond to the same point in $\boldsymbol{SO}(3)$. Therefore, we repair the quaternion after each state prediction and measurement to preserve its uniqueness:

$$\boldsymbol{q} \leftarrow \begin{cases} -\frac{\boldsymbol{q}}{||\boldsymbol{q}||}, & \text{if } q_0 < 0 \\ \frac{\boldsymbol{q}}{||\boldsymbol{q}||}, & \text{otherwise} \end{cases} \tag{15}$$

F. CoM Kalman Filter and XCoM Determination

In this subsection, we describe our novel CoM Kalman Filter to estimate the body CoM (denoted as $\boldsymbol{r}$) and the velocity of CoM (denoted as $\dot{\boldsymbol{r}}$) using the measurements of the pelvis pose and the center of pressure (CoP). The linearized inverted pedulum model underlying the theory of dynamic stability [9] can be written as

$$\ddot{\boldsymbol{r}} = \omega_0^2(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{P}}), \tag{16}$$

where $\omega_0 = \sqrt{g/l}$, g is the gravitational acceleration, l is the effective length of the inverted pendulum (the average height of the CoM), and $\boldsymbol{r}_{\mathrm{P}} \in \mathbb{R}^2$ is the position of the CoP.

Due to the nature of the robot wheel odometry and the depth camera, the CoM measurement may be affected by an unknown offset $\Delta\boldsymbol{r}$ [33]. Augmenting the state with this offset, we have $\boldsymbol{x} = [\boldsymbol{r}^{\mathrm{T}},\ \dot{\boldsymbol{r}}^{\mathrm{T}},\ \boldsymbol{r}_{\mathrm{P}}^{\mathrm{T}},\ \Delta\boldsymbol{r}^{\mathrm{T}}]^{\mathrm{T}}$. The discrete-time system equation can be written as

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}\boldsymbol{x}_k + \boldsymbol{u}_k + \boldsymbol{w}_k, \tag{17}$$

where

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{I}_2\Delta t & \mathbf{0} & \mathbf{0} \\ \boldsymbol{I}_2\omega_0^2\Delta t & \boldsymbol{I}_2 & -\boldsymbol{I}_2\omega_0^2\Delta t & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_2 \end{bmatrix},$$

$$\boldsymbol{u}_k = \begin{bmatrix} \mathbf{0}_{6\times 2} \\ \boldsymbol{R}_z(\Delta\Psi_k) \end{bmatrix},$$

$$\boldsymbol{w}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}),$$

$$\boldsymbol{R}_z(\Delta\Psi_k) = \begin{bmatrix} \cos(\Delta\Psi_k) & -\sin(\Delta\Psi_k) \\ \sin(\Delta\Psi_k) & \cos(\Delta\Psi_k) \end{bmatrix},$$

$$\Delta\Psi_k = \boldsymbol{v}_{\mathrm{AP},k-1} \times \boldsymbol{v}_{\mathrm{AP},k} \cdot \boldsymbol{k}_W$$

and $\boldsymbol{v}_{\mathrm{AP,k}}$ is a unit vector pointing in the anterior axis of the subject while $\boldsymbol{k}_W$ is the unit vector pointing upward in the world frame. Compared to [33], the formulation in (17) adds $\boldsymbol{u}_k$ that rotates the CoM offset $\Delta\boldsymbol{r}$ when the subject is turning. The covariance matrix of the process noise is $\boldsymbol{Q} = \mathrm{diag}(\boldsymbol{Q}_{\boldsymbol{r}}, \boldsymbol{Q}_{\dot{\boldsymbol{r}}}, \boldsymbol{Q}_{\boldsymbol{r}_\mathrm{P}}, \boldsymbol{Q}_{\Delta\boldsymbol{r}}) = \mathrm{diag}(1.0 \times 10^{-4}\boldsymbol{I}_2, 1.0 \times 10^{-2}\boldsymbol{I}_2, 1.0 \times 10^{-2}\boldsymbol{I}_2, 3.0 \times 10^{-2}\boldsymbol{I}_2.)$

Let $\boldsymbol{z}_{\boldsymbol{r}} \in \mathbb{R}^2$ be the projected CoM that is first measured by the Kinect sensor and then transformed to world frame W by (2). The measurement equation for CoM can be written as

$$\boldsymbol{z}_{\boldsymbol{r}} \triangleq \begin{bmatrix}\boldsymbol{I}_2 & \boldsymbol{0}_{2\times 1}\end{bmatrix} \boldsymbol{z}_{\boldsymbol{p}} = \boldsymbol{h}_{\boldsymbol{r}}(\boldsymbol{x}) + \boldsymbol{\nu}_{\boldsymbol{r}} = \boldsymbol{r} + \Delta\boldsymbol{r} + \boldsymbol{\nu}_{\boldsymbol{r}} \quad (18)$$

where $\boldsymbol{\nu}_{\boldsymbol{r}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{r}})$ is the CoM measurement noise with the covariance matrix chosen to be $\boldsymbol{R}_{\boldsymbol{r}} = 1.0 \times 10^{-2}\boldsymbol{I}_2$.

Given the foot/insole poses, the CoP can be calculated as the weighted average of the locations $\hat{\boldsymbol{r}}_{\mathrm{P}j}$ of the 16 pressure cells in frame W where the weight associated with each cell is the pressure measurement P_j of that cell, with $j = 1, 2, ..., 16$. The measurement model for the CoP can be written as

$$\boldsymbol{z}_\mathrm{P} \triangleq \frac{\Sigma_j \hat{\boldsymbol{r}}_{\mathrm{P}j} P_j}{\Sigma_j P_j} = \boldsymbol{h}_\mathrm{P}(\boldsymbol{x}) + \boldsymbol{\nu}_\mathrm{P} = \boldsymbol{r}_\mathrm{P} + \boldsymbol{\nu}_\mathrm{P} \quad (19)$$

where $\boldsymbol{\nu}_\mathrm{P} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_\mathrm{P})$ is the CoP measurement noise with the covariance matrix chosen to be $\boldsymbol{R}_\mathrm{P} = 2.5 \times 10^{-3}\boldsymbol{I}_2$.

After we obtain the estimate of CoM, $\boldsymbol{r}$, and its velocity $\dot{\boldsymbol{r}}$, we can determine XCoM, which is defined as the sum of the projected body CoM and a term proportional to CoMv [9], that is, $\boldsymbol{\xi} = \boldsymbol{r} + \dot{\boldsymbol{r}}/\omega_0$.

G. BoS Determination

For the purpose of BoS determination, we consider each foot as a two-segment body with a hindfoot segment and a forefoot segment connected by a single hinge joint running from the first to the fifth metatarsal heads. In a similar way to [13], we define, for each foot i, 3 hindfoot virtual markers $\boldsymbol{r}_{\mathrm{h}ij}$ and 3 forefoot virtual markers $\boldsymbol{r}_{\mathrm{f}ij}$, with $j = 1, 2, 3$, as illustrated with green circles in Fig. 6. Virtual markers $\boldsymbol{r}_{\mathrm{h}ij}$ and $\boldsymbol{r}_{\mathrm{f}ij}$ have a fixed distance, respectively, from $\boldsymbol{r}_{\mathrm{h}i}$ and $\boldsymbol{r}_{\mathrm{f}i}$, and a fixed angle from the vector $\boldsymbol{r}_{\mathrm{f}i} - \boldsymbol{r}_{\mathrm{h}i}$. The virtual markers of a foot segment (hindfoot, forefoot) are added as vertices to the BoS polygon only if the foot segment is in contact with the ground. The position of the virtual markers are calculated from the output of the foot pose EKF (Section III-E). The set of virtual markers that determine the BoS polygon is $\mathcal{M} = \cup_{i\in\{\mathrm{l,r}\}}\mathcal{M}_i$, where the transition of $\mathcal{M}_i$ is driven by the gait events as illustrated by Fig. 5 with $\mathcal{M}_{\mathrm{h}i} = \{\boldsymbol{r}_{\mathrm{h}ij} : j = 1, 2, 3\}$ and $\mathcal{M}_{\mathrm{f}i} = \{\boldsymbol{r}_{\mathrm{f}ij} : j = 1, 2, 3\}$. When the subject is walking or standing, at least one foot segment is in contact with the ground at any time; the set $\mathcal{M}$ of potential vertices that forms the BoS include 3 to 12 markers, avoiding the possibility of ill-defined BoS polygon. Finally, Graham's Scan [34] is used to find the convex hull of the set of vertices, which is used to represent the BoS polygon for calculating MoS.

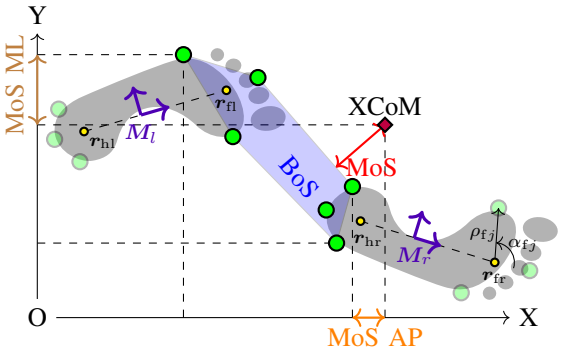


Fig. 6: Definition of MoS, MoS AP, and MoS ML. The BoS polygon is determined by 6 of the 12 virtual markers in green, which corresponds to the case shown here where the left forefoot and the right hindfoot (marked by dark green dots) are in contact with the ground, while the left hindfoot and the right forefoot are off the ground (marked by light green dots) . The X axis is parallel to the line of progression.

H. MoS Calculation

As shown in Fig. 6, we follow the definition in [9], [11] and define:

- MoS AP: The smaller one of the distances of the XCoM from the frontmost and rearmost BoS vertices
- MoS ML: The smaller one of the distances of the XCoM from the leftmost and rightmost BoS vertices.
- MoS: The shortest distance from XCoM to the boundary of the BoS polygon.

When calculating MoS AP and MoS ML, we need to determine anteroposterior or mediolateral axes from sensor data. We take the hindfoot reference point $\boldsymbol{r}_{\mathrm{h}i}$ of one side at time instants t_1 and t_2 of two consecutive heel-strike events of this side and use the unit vector along $(\boldsymbol{r}_{\mathrm{h}i}(t_2) - \boldsymbol{r}_{\mathrm{h}i}(t_1))$ as anteroposterior axis. When XCoM lies inside the BoS polygon, MoS values are positive indicating the subject is in a dynamically stable state. Taking the example as shown in Fig. 6, MoS and MoS AP have negative values because they are outside the BoS while MoS ML is positive.

IV. Performance Evaluation

In this section, we first introduce the optical motion capture (OptiTrack) reference system and evaluate the performance of robot control. Then we evaluate the performance of the Kalman filters for estimating the foot poses, CoM and CoMv in comparison with the reference system. Finally, we assess the calculation of MoS and spatio-temporal gait parameters and compare the performance with existing literature.

A. Reference System

OptiTrack is used as a reference system to assess MoS accuracy of our gait monitoring system, which operates at the sampling rate of 210Hz. The method from [17] is adopted with slight modifications. Four retro-reflective markers are fastened to the waist of the subjects via self-adherent wrap (3M™ Coban™) such that they are approximately located at left/right

anterior/posterior superior iliac spine (LASI, RASI, LPSI, RPSI) to estimate the CoM position. Six reflective markers were taped to each of the subject's shoes, three close to the lateral calcaneus, posterior calcaneus, sustentaculum tali, and the other three on the side of the heads of the first metatarsal, second toe, and fifth metatarsal. The centroids of the triplets of the reflective markers are used to specify hindfoot and forefoot reference points, respectively. The calculation of reference XCoM and BoS is described as follows:

1) Reference XCoM: The reference CoM is calculated as $\boldsymbol{z}_{\boldsymbol{r}\mathrm{ref}} \triangleq \frac{1}{4}(\mathrm{LASI}+\mathrm{RASI}+\mathrm{LPSI}+\mathrm{RPSI}) = \boldsymbol{r}+\Delta\boldsymbol{r}_{\mathrm{ref}}$ which is also subject to an unknown offset $\Delta\boldsymbol{r}_{\mathrm{ref}}$. Therefore, the same CoM Kalman filter as in III-F is used to estimate the reference CoM $\boldsymbol{r}_{\mathrm{ref}}$ and CoMv $\dot{\boldsymbol{r}}_{\mathrm{ref}}$.

2) Reference BoS: To calculate the reference BoS, first, the timings for the phases of the stance period are determined following Section III-D using the pressure sensor data. Then the BoS determination is mostly the same as described in Section III-G except that the forefoot and hindfoot reference points are determined by the set of retro-reflective markers corresponding to each foot segment.

B. Results on Robot Control

We validate the performance of the robot localization and control by checking the trajectories and speeds of the human and the robot, and the distance between them. The average walking speed of each subject ranges from 0.67 m/s to 1.07 m/s. The result from three laps of robot-guided walking in a representative trial is plotted in Fig. 7 using data from the OptiTrack reference system. The robot position is determined using two reflective markers taped to the centers of the two driving wheels of the robot, and we use the midpoint between these two markers as the robot position. We use the reference CoM as the human position.

Fig. 7(a) and (b) show the time histories of human-robot distance and their line velocities (i.e., speeds), respectively. Fig. 7(c) shows the actual robot and human trajectories. We can see that the controller we designed successfully maintains the human-robot distance around the desired value (i.e., 1.4 m), and the robot speed matches the human speed in general. Note that our designed controller uses the measured distance from the human to adjust the robot velocity, thus this feedback type of controller casues some flutuations of the human-robot distance as shown in the figure.

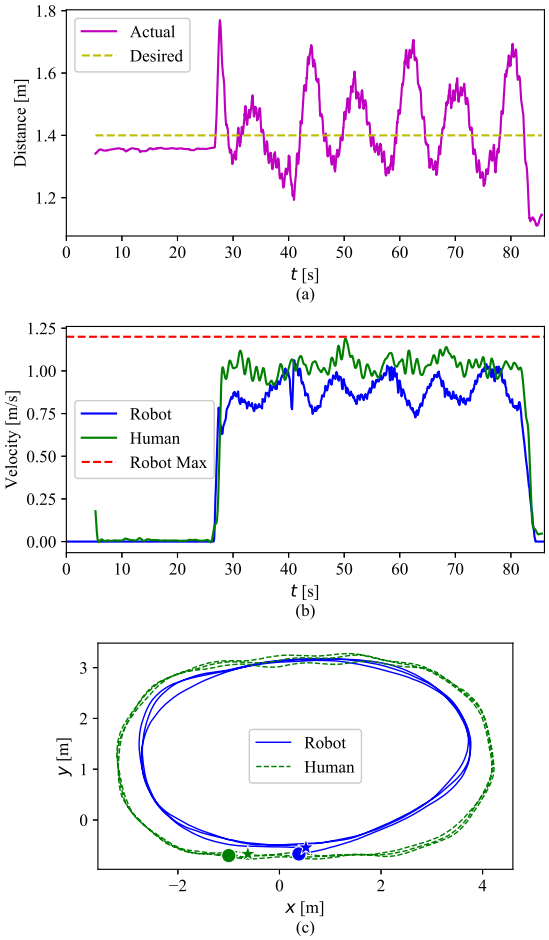


Fig. 7: Time histories of (a) human-robot distance, (b) velocities of the robot and the human, and (c) actual trajectories of the robot and the human, during three laps of walking in counterclockwise direction. The solid circles in (c) represent the start positions of the robot and the human, and the stars represent their end positions.

C. Results on Kalman Filters

To demonstrate the performance of the designed EKF, we compare the fused, raw and reference foot positions (where the insole IMUs are located) over 6 seconds of walking in Fig. 8. The raw position is calculated by (3). The fused position is estimated by the foot pose EKF described in Section III-E, which fuses the data from the IMU and pressure sensors in the insoles with the raw Kinect measurement. We can see that the EKF produces accurate position estimation compared with the reference system.

The fused, raw and reference body CoM and CoMv over a period of 17 seconds are shown in Fig. 9. The raw CoM position is calculated following Section III-C. The fused CoM is estimated using the CoM Kalman filter introduced in Section III-F that fuses the CoP and raw CoM measurement. Due to the cumulative error of the wheel odometry, the fused CoM and the reference CoM cannot be aligned with each other perfectly. Nonetheless, this cumulative error has a negligible effect on the calculation of MoS because it affects the calculation of foot poses in the same way as it does the CoM, allowing the effects to cancel each other out as a result of the definition of MoS. Fig. 9(b) shows the temporal evolution of CoMv. Fused CoMv is part of the posterior state estimate computed by the CoM Kalman filter. Raw CoMv is the numerical differentiation of CoM measurement with respect to time. It can be seen that the CoM Kalman filter improves the accuracy of the CoMv estimation and the MoS calculation.

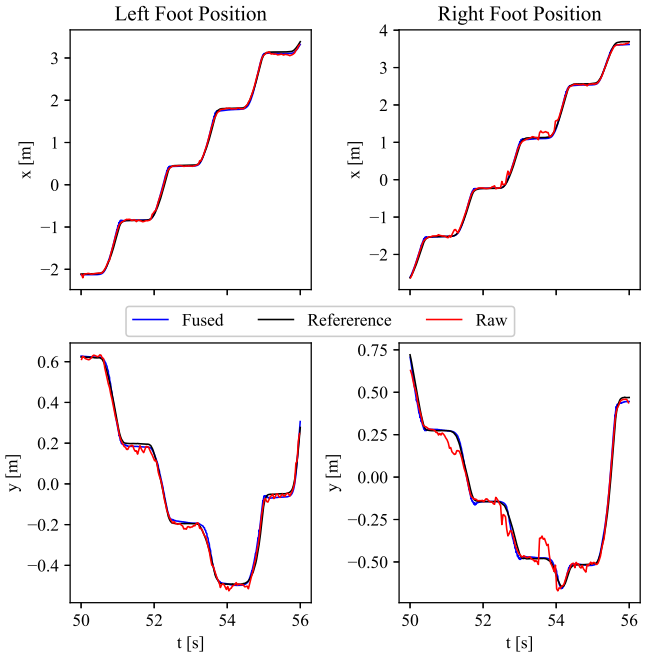


Fig. 8: Temporal evolution of the raw, EKF fused, and reference foot position over 6 seconds of a representative trial.

D. Results on MoS Estimation

We first normalize and aggregate the MoS measurements by gait cycles. The evolution of the MoS values within all gait cycles from one subject is shown in Fig. 10. For the purpose of data analysis, the gait cycle was defined using the left foot, i.e., 0% and 100% correspond to one left heel strike and the subsequent left heel strike, respectively. The MoS values are re-sampled at 101 points equally spaced in the gait phase domain. It can be seen that the mean values and standard deviations generally agree between the reference and the proposed systems.

The proposed system and the reference system are then compared in terms of RMSE in both overall and straight walking. Note that the straight walking corresponds to the 3.9 m straight segment of the human track as shown in Fig. 2 (b). To calculate these performance metrics, the MoS values are re-sampled using the same method as described earlier to produce the error values between the proposed system and the reference system. For each subject, an RMSE is calculated by aggregating these errors from all gait cycles. The group average RMSE and the between-subject standard deviation (SD) for 10 subjects are presented in Table I, and the box plots of the RMSEs are shown in Fig. 11. For all three MoS values, the error of straight walking is slightly smaller than the overall error of curved and straight walking combined. While existing work [17] reports accuracy of MoS estimation lower than 2 cm, their results were obtained for treadmill walking but not overground walking. To quantify interrater reliability, we compute the intraclass correlation coefficients (ICC) ([35]) between the estimates of MoS, MoS AP, and MoS ML from the mobile system and the reference system, using

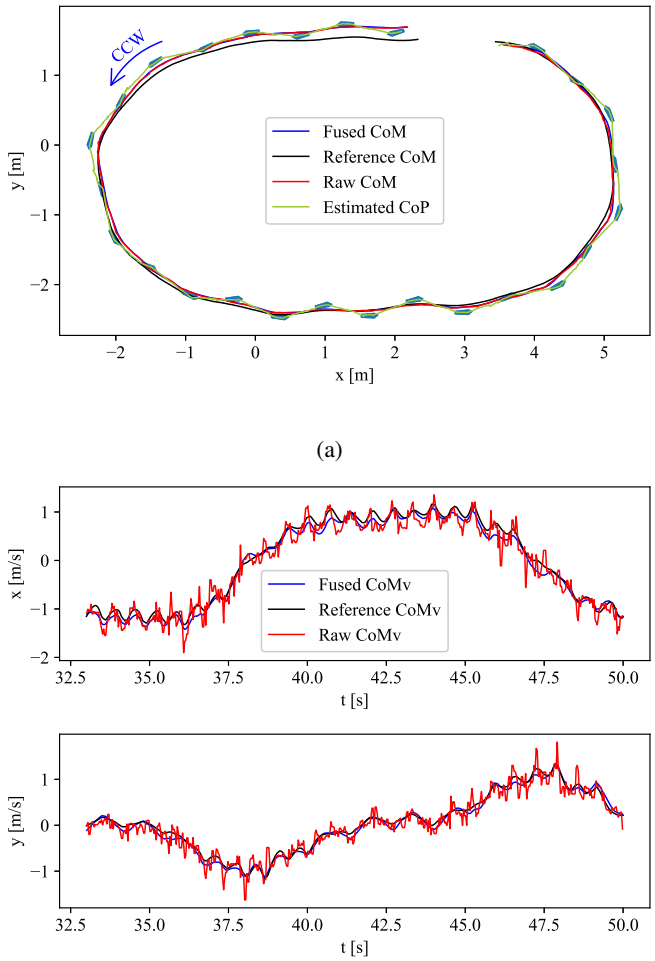


Fig. 9: Result of the CoM Kalman filter over 17 seconds of walking in counterclockwise direction in a representative trial: (a) Footprints and trajectories of CoM and CoP. Footprints denoted by green filled polygons represent the midstance foot poses as estimated by the proposed system. (b) The evolution of CoMv over time.

a single-measurement, absolute-agreement, two-way mixed effects model.

Results for MoS AP demonstrate excellent reliability ($ICC_{AP} = 0.98$, 95% CI 0.94-1.00), however the ICC for MoS show mixed results ($ICC_{MoS} = 0.92$, 95% CI 0.23-0.98) and the ICC for ML is poor to moderate ($ICC_{ML} = 0.26$, 95% CI 0.06-0.70). We note that the typical range of MoS ML during locomotion (e.g., 4-6 cm [36], 6-10cm [37]) is much smaller than those of MoS and MoS AP, therefore inaccuracies in the determination of XCoM and BoS have a larger influence in the accuracy of MoS ML than they have for the other two variables.

Overall, the RMSE results and the ICC values indicate that the mobile system is valid and reliable in estimating MoS

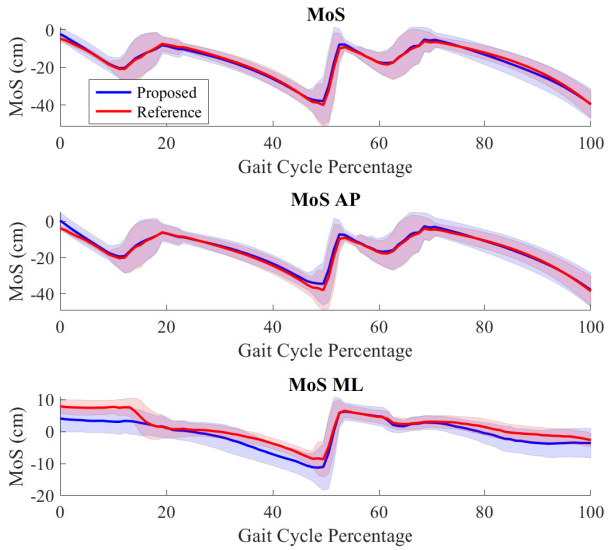


Fig. 10: Comparison of the MoS between the proposed system and the reference system for a representative subject. The solid-line curve represents the mean, and the shaded area denotes within-subject ±SD.

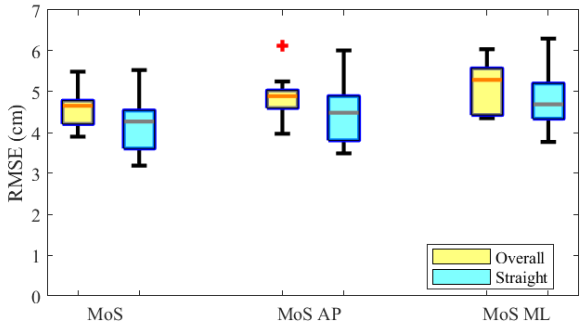


Fig. 11: Box plots comparing system performance in RMSE of overall and straight walking.

AP, but further work is warranted to improve the MoS ML estimates.

It is worth noting that in the proposed method most computational load comes from the Azure Kinect Body Tracking SDK, which consumes more than half of the CPU and GPU resources. The foot pose EKF has more computational cost than the CoM Kalman filter due to the higher dimension of the state vector. In the foot pose EKF algorithm, the step on the insole IMU measurement update takes 0.75 ± 0.34 ms, and the step on the Kinect measurement update takes 0.42 ± 0.15 ms, which meets the requirement of real time MoS estimation.

TABLE I: Error statistics of the MoS values.

	RMSE (SD) [cm]		
	MoS	MoS AP	MoS ML
Overall	4.49 (0.57)	4.76 (0.66)	5.03 (0.87)
Straight	4.00 (0.68)	4.27 (0.81)	4.59 (0.93)

E. Results on Spatial-Temporal Gait Parameter Estimation

In addition to MoS estimation, the proposed foot pose EKF can also be used to calculate conventional spatio-temporal gait parameters. To obtain the foot position for each step using this system, the output of the foot pose EKF is averaged over the latter 50% of the foot-flat phase. The foot position for each step is subsequently used to compute spatial gait parameters including stride length, step length and step width. Table II shows that the proposed system achieves higher validity (i.e., lower mean absolute error) than the previous studies that use IMU, Laser range finder (LRF) or RGBD sensors. We find moderate to excellent reliability for stride length, step length, and stride velocity ($ICC_{StrL} = 1.00$, 95% CI 0.52-1.00, $ICC_{StpL} = 1.00$, 95% CI 0.51-1.00, $ICC_{SV} = 1.00$, 95% CI 0.66-1.00) and low to excellent reliability for step width ($ICC_{StpW} = 0.77$, 95% CI 0.16-0.94).

TABLE II: Spatial-temporal gait parameter comparison

		Mean Absolute Error (SD)			
Ref.	Systems	Stride length cm	Step length cm	Step width cm	Stride velocity cm/s
[38]	IMU	>6.26 (5.56)	-	-	-
[39]	IMU	>3.0	-	-	>3.0
[40]	LRF	>4.8	-	-	-
[41]	LRF	-	2.2 (1.0)	2.5 (1.3)	-
[23]	RGBD	2.2	-	-	-
[42]	RGBD	-	3.2 (1.78)	-	2.9(1.32)
[43]	RGBD	-	3 (2)	4 (2)	-
[25]	RGBD+IMU	1.63	1.9	-	1.1
This paper	RGBD+IMU	1.62 (0.39)	1.77 (0.65)	1.70 (0.78)	1.22 (0.31)

F. Comparison with Existing Work

We compare our proposed method with existing literature on MoS assessment including [10], [17], [18]. Table III summarizes the main differences.

Although [17] achieved good MoS estimation accuracy, it requires 7 IMUs being strapped to various body segments as well as a subject-specific anthropometric measurement (sensor-to-segment calibration) and regression model. Employing sensors strapped to fewer body segments than [17], the systems developed by [10] and [18] assess only MoS AP and a correlate for MoS ML, respectively. Unlike the systems in [17] and [10], which require subjects to perform calibration steps, both this paper and [18] do not require subject-dependent calibration. Furthermore, the work [18] calculates only some quantities correlated to the MoS ML averaged over each step, while the system proposed in this paper can assess the MoS values over a gait cycle, which is important for real time fall-risk alarm or feedback ([11]).

Unlike the proposed work, none of the methods in [10], [17] and [18] models the foot with forefoot and hindfoot segments and takes the phases of the stance period into account when determining the BoS. This more fine-grained evaluation of BoS has been argued by [11] to be able to provide deeper insight into the assessment of fall risk. As the study in [17] constrains the walking gait speed as well as the direction of walking by using a treadmill, it is reasonable to expect that RMSEs are lower in their case.

At last, the proposed system is designed to calculate the MoS values in real time. This feature is critical since it can

enable the integration of on-line feedback (e.g., auditory or vibrotactile) in future developments of our system, which would not be possible with an offline method.

G. Limitation and Future Work

Our system utilizes an autonomous mobile robot and a pair of instrumented insoles. The wheeled robot navigates on flat terrain without slope. While the robot can move on a mat or carpet, performances of the system may deteriorate due to vibration of the Azure Kinect sensor (used for body tracking) on the robot caused by non-smooth motion of the robot. If the robot moves in environments with obstacles and/or humans, advanced intelligent autonomous motion planning and control algorithms (e.g. [44], [45], [46], [47]) can be applied to lead or guide subject walking. As a proof of concept, this study applies a simple robot motion planning and control algorithm and focuses on fusing data from the robot vision sensor and insole sensors. Future work may include overground walking in complex real-world environments. However, there are certain real-life locomotor tasks (e.g., stair negotiation) to which the proposed system cannot be applied.

This study aims at providing a portable robot and wearable system monitoring gait and MoS during walking tasks on level surface. Because most standardized clinical gait assessments include straight-line walking and turns on level ground, this technology can potentially be applied to administer instrumented clinical tests in future studies. As a proof-of-concept, we tested the system and algorithms in a laboratory setting with 10 healthy young subjects. While we expect these methods to be applicable to other populations (e.g., older adults) in out-of-the-lab conditions, several adjustments will likely be necessary. For example, the finite state machine in Fig. 5 will need to be updated to account for non-typical gait patterns (e.g., shuffling gait) for which the standard sequence of gait events does not apply. Future work will also be directed towards improving the validity and reliability of the MoS ML estimates.

V. Conclusion

In this paper, we presented a mobile gait monitoring system consisting of a mobile robot and a pair of instrumented insoles. Kalman filter based methods were developed to fuse measurements from the Azure Kinect sensor on the robot with the plantar IMU and pressure sensors on the insoles, to estimate the MoS in real time. Using the OptiTrack camera system as the reference, experimental results with 10 young healthy subjects have shown high accuracy for spatio-temporal gait parameters. Agreement with the gold standard was excellent for anteroposterior MoS, but low to moderate for mediolateral MoS. In comparison with existing work on MoS estimation, the subject only wears a pair of light-weight insoles, and no sensor-to-segment calibration is required, thereby reducing subject setup times. Compared to previous studies, we obtained higher accuracy in the spatial-temporal gait parameters. Future work includes the evaluation of the system with older adults in a wider range of walking conditions.

References

[1] T. J. Buurke, C. J. Lamoth, L. H. van der Woude, A. L. Hof, and R. den Otter, "Bilateral temporal control determines mediolateral margins of stability in symmetric and asymmetric human walking," *Scientific Reports*, vol. 9, no. 1, pp. 1–10, 2019.

[2] K. Karamanidis, A. Arampatzis, and L. Mademli, "Age-related deficit in dynamic stability control after forward falls is affected by muscle strength and tendon stiffness," *J. Electromyogr. Kinesiol.*, vol. 18, no. 6, pp. 980–989, 2008.

[3] A. F. Ambrose, G. Paul, and J. M. Hausdorff, "Risk factors for falls among older adults: a review of the literature," *Maturitas*, vol. 75, no. 1, pp. 51–61, 2013.

[4] S. Whitney, M. Hudak, and G. Marchetti, "The dynamic gait index relates to self-reported fall history in individuals with vestibular dysfunction," *J. Vestib. Res.*, vol. 10, no. 2, pp. 99–105, 2000.

[5] L. Salonen and S.-L. Kivelä, "Eye diseases and impaired vision as possible risk factors for recurrent falls in the aged: a systematic review," *Curr. Gerontol. Geriatr. Res.*, 2012.

[6] 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," *Arch. Phys. Med. Rehabil.*, vol. 82, no. 8, pp. 1050–1056, 2001.

[7] C. S. Florence, G. Bergen, A. Atherly, E. Burns, J. Stevens, and C. Drake, "Medical costs of fatal and nonfatal falls in older adults," *J. American Geriatr. Soc.*, vol. 66, no. 4, pp. 693–698, 2018.

[8] S. Bruijn, O. Meijer, P. Beek, and J. Van Dieën, "Assessing the stability of human locomotion: a review of current measures," *J. Roy. Soc. Interface*, vol. 10, no. 83, p. 20120999, 2013.

[9] A. Hof, M. Gazendam, and W. Sinke, "The condition for dynamic stability," *J. Biomech.*, vol. 38, no. 1, pp. 1–8, 2005.

[10] M. I. Mohamed Refai, B. F. van Beijnum, J. H. Buurke, and P. H. Veltink, "Gait and dynamic balance sensing using wearable foot sensors," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 27, pp. 218–227, Feb 2019.

[11] H. Ohtsu, S. Yoshida, T. Minamisawa, N. Katagiri, T. Yamaguchi, T. Takahashi, S.-i. Yomogida, and H. Kanzaki, "Does the balance strategy during walking in elderly persons show an association with fall risk assessment?," *J. Biomech.*, p. 109657, 2020.

[12] F. Franchignoni, F. Horak, M. Godi, A. Nardone, and A. Giordano, "Using psychometric techniques to improve the balance evaluation systems test: the mini-BESTest," *J. Rehabil. Med.*, vol. 42, no. 4, pp. 323–331, 2010.

[13] H. Ohtsu, S. Yoshida, T. Minamisawa, T. Takahashi, S.-i. Yomogida, and H. Kanzaki, "Investigation of balance strategy over gait cycle based on margin of stability," *J. Biomech.*, vol. 95, p. 109319, 2019.

[14] L. Mademli and A. Arampatzis, "Lower safety factor for old adults during walking at preferred velocity," *Age*, vol. 36, no. 3, p. 9636, 2014.

[15] M. Guaitolini, F. Aprigliano, A. Mannini, A. Sabatini, and V. Monaco, "Effects of gait speed on the margin of stability in healthy young adults," in *Int. Conf. NeuroRehabil.*, pp. 420–424, Springer, 2018.

[16] F. B. van Meulen, D. Weenk, J. H. Buurke, B.-J. F. van Beijnum, and P. H. Veltink, "Ambulatory assessment of walking balance after stroke using instrumented shoes," *J. NeuroEng. Rehabil.*, vol. 13, no. 1, p. 48, 2016.

[17] M. Guaitolini, F. Aprigliano, A. Mannini, S. Micera, V. Monaco, and A. M. Sabatini, "Ambulatory assessment of the dynamic margin of stability using an inertial sensor network," *Sensors*, vol. 19, no. 19, p. 4117, 2019.

[18] P. C. Fino, F. B. Horak, and C. Curtze, "Inertial sensor-based centripetal acceleration as a correlate for lateral margin of stability during walking and turning," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 28, no. 3, pp. 629–636, 2020.

[19] J. F. Veneman, R. Kruidhof, E. E. Hekman, R. Ekkelenkamp, E. H. Van Asseldonk, and H. Van Der Kooij, "Design and evaluation of the LOPES exoskeleton robot for interactive gait rehabilitation," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 15, no. 3, pp. 379–386, 2007.

[20] R. A. R. C. Gopura, K. Kiguchi, and Y. Li, "SUEFUL-7: A 7DOF upper-limb exoskeleton robot with muscle-model-oriented EMG-based control," in *IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, pp. 1126–1131, 2009.

[21] S. Šabanović, C. C. Bennett, W.-L. Chang, and L. Huber, "PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia," in *IEEE Int. Conf. Rehabil. Rob. (ICORR)*, pp. 1–6, 2013.

[22] B. Görer, A. A. Salah, and H. L. Akın, "An autonomous robotic exercise tutor for elderly people," *Auton. Robots*, vol. 41, no. 3, pp. 657–678, 2017.

TABLE III: Summary of works on continuous monitoring of MoS.

Reference	Systems	Measures	Walking type	Online-ready	Are subjects involved in calibration?	Are phases of the stance period considered?
Refai *et al.* [10]	Wearable (1D pressure sensors, IMUs and ultrasonic sensors)	MoS AP	Overground	No	Yes	No
Guaitolini *et al.* [17]	Wearable (IMUs on pelvis, thighs, shanks and feet)	MoS AP, MoS ML	Treadmill	No	Yes	No
Fino *et al.* [18]	Wearable (IMUs on lumbar and feet)	MoS ML correlate	Overground	No	No	No
This paper	Robot + wearable	MoS AP, MoS ML, MoS	Overground	Yes	No	Yes

[23] V. Bonnet, C. A. Coste, L. Lapierre, J. Cadic, P. Fraisse, R. Zapata, G. Venture, and C. Geny, "Towards an affordable mobile analysis platform for pathological walking assessment," *Rob. Auton. Syst.*, vol. 66, pp. 116–128, 2015.

[24] G. Chalvatzaki, P. Koutras, J. Hadfield, X. S. Papageorgiou, C. S. Tzafestas, and P. Maragos, "LSTM-based network for human gait stability prediction in an intelligent robotic rollator," in *IEEE Int. Conf. Rob. Autom. (ICRA)*, pp. 4225–4232, 2019.

[25] H. Zhang, Z. Chen, D. Zanotto, and Y. Guo, "Robot-assisted and wearable sensor-mediated autonomous gait analysis," in *IEEE Int. Conf. Rob. Autom. (ICRA)*, 2020.

[26] H. Zhang, D. Zanotto, and S. K. Agrawal, "Estimating CoP trajectories and kinematic gait parameters in walking and running using instrumented insoles," *IEEE Rob. Automat. Lett.*, vol. 2, no. 4, pp. 2159–2165, 2017.

[27] H. Zhang, Y. Guo, and D. Zanotto, "Accurate ambulatory gait analysis in walking and running using machine learning models," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 28, no. 1, pp. 191–202, 2020.

[28] "Documentation for Azure Kinect DK." https://docs.microsoft.com/en-us/azure/Kinect-dk/. Accessed: 2020-10-13.

[29] M. Labbé and F. Michaud, "RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," *J. Field Rob.*, vol. 36, no. 2, pp. 416–446, 2019.

[30] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robot. Autom. Mag.*, vol. 4, no. 1, pp. 23–33, 1997.

[31] J. Perry, J. R. Davids, *et al.*, "Gait analysis: normal and pathological function," *J. Pediatr. Orthop.*, vol. 12, no. 6, p. 815, 1992.

[32] J. Sola, "Quaternion kinematics for the error-state kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.

[33] B. J. Stephens, "State estimation for force-controlled humanoid balance using simple models in the presence of modeling error," in *IEEE Int. Conf. Rob. Autom. (ICRA)*, pp. 3994–3999, 2011.

[34] R. L. Graham, "An efficient algorithm for determining the convex hull of a finite planar set," *Info. Pro. Lett.*, vol. 1, pp. 132–133, 1972.

[35] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *J Chiropr Med.*, vol. 15, no. 2, pp. 155–163, 2016.

[41] C. Piezzo, B. Leme, M. Hirokawa, and K. Suzuki, "Gait measurement by a mobile humanoid robot as a walking trainer," in *IEEE Int. Symp. Robot Human Interactive Commun. (RO-MAN)*, pp. 1084–1089.

[36] M. Kazanski, J. P. Cusumano, and J. B. Dingwell, "Rethinking margin of stability: Incorporating step-to-step regulation to resolve the paradox," *bioRxiv*, 2021.

[37] P. M. Young, J. Wilken, and J. Dingwell, "Dynamic margins of stability during human walking in destabilizing environments," *J Biomech.*, vol. 45, no. 6, pp. 1053–1059, 2012.

[38] A. Rampp, J. Barth, S. Schülein, K.-G. Gaßmann, J. Klucken, and B. M. Eskofier, "Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1089–1097, 2014.

[39] B. Mariani, C. Hoskovec, S. Rochat, C. Büla, J. Penders, and K. Aminian, "3D gait assessment in young and elderly subjects using foot-worn inertial sensors," *J. Biomech.*, vol. 43, no. 15, pp. 2999–3006, 2010.

[40] X. S. Papageorgiou, G. Chalvatzaki, K.-N. Lianos, C. Werner, K. Hauer, C. S. Tzafestas, and P. Maragos, "Experimental validation of human pathological gait analysis for an assisted living intelligent robotic walker," in *IEEE Int. Conf. Biomed. Rob. Biomechatron.*, pp. 1086–1091, 2016.

[42] B. Jäschke, A. Vorndran, T. Q. Trinh, A. Scheidig, H.-M. Gross, K. Sander, and F. Layher, "Making gait training mobile - a feasibility analysis," in *IEEE Int. Conf. Biomed. Rob. Biomechatron.*, pp. 484–490, 2018.

[43] J. A. Albert, V. Owolabi, A. Gebel, C. M. Brahms, U. Granacher, and B. Arnrich, "Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: A pilot study," *Sensors*, vol. 20, no. 18, p. 5104, 2020.

[44] M. Fahad, G. Yang, and Y. Guo, "Learning human-like motion planning using measured human trajectories in crowded spaces," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020.

[45] X. Yu, B. Li, W. H. Y. Feng, L. Cheng, and C. Silvestre, "Adaptive-constrained impedance control for human-robot co-transportation," *IEEE Trans Cybern.*, 2021.

[46] X. Yu, W. He, Q. Li, Y. Li, and B. Li, "Human-robot co-carrying using visual and force sensing," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 9, pp. 8657–8666, 2021.

[47] X. Yu, W. He, H. Li, and J. Sun, "Adaptive fuzzy full-state and output-feedback control for uncertain robots with output constraint," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 11, pp. 6994–7007, 2021.