

Received 15 November 2022, accepted 2 December 2022, date of publication 7 December 2022,
date of current version 12 December 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3227344

RESEARCH ARTICLE

Gait Phase Estimation of Unsupervised Outdoors Walking Using IMUs and a Linear Regression Model

AHMED SOLIMAN¹, (Graduate Student Member, IEEE), GUILHERME A. RIBEIRO¹,
ANDRES TORRES², LI-FAN WU², AND MO RASTGAAR¹, (Senior Member, IEEE)

¹Polytechnic Institute, Purdue University, West Lafayette, IN 47906, USA

²Mechanical Engineering, Purdue University, West Lafayette, IN 47906, USA

Corresponding author: Ahmed Soliman (solimana@purdue.edu)

This work was supported by the National Science Foundation (NSF) under Grant 1921046.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Purdue Institutional Review Board under Application No. 1902021698.

ABSTRACT Human gait analysis and detection are critical for many applications, including wearable and rehabilitation robotic devices, reducing or tracking injury risk. The proposed work allows researchers to study the gait phase of human subjects in an unsupervised outdoor environment without the need for fixed thresholds and sensor-embedded insoles. We present an experimental protocol to label gait events based on patterns in human subjects from two body-worn inertial measurement units (IMUs). Gait patterns are developed using a force plate and a motion capture system. Upon defining the gait pattern, human subjects walk outdoors for forty minutes to train and test a principal component analysis (PCA)-based linear regression model. Next, gait phase estimation is performed using the defined patterns from other human subjects to accommodate cases where motion capture and force plate data are unavailable. Results showed a minimum normalized gait phase estimation error of 1.81 %, a maximum of 2.48 %, and an average of 2.21 ± 0.258 % for all subjects involved. Results are particularly significant because the proposed work can be expanded to precise control of human-assistive devices, rehabilitation devices, and clinical gait analysis.

INDEX TERMS Biomechanics, gait analysis, heel strike, PCA, machine learning, prosthesis, IMUs, motion capture, linear regression.

I. INTRODUCTION

Walking is an essential part of humans' activities of daily living. As a result, there has been increasing clinical and research interest in gait monitoring and analysis. Gait cycles begin when the heel of one foot strikes the ground (0% of gait phase) and ends with the next strike of the same foot. Each cycle includes two major phases, a stance, and a swing phase. The stance phase is when the studied foot is on the ground; the stance starts at heel strike (HS) and ends at toe-off (TO). The swing phase is when the studied foot is off the

ground; the swing starts at TO and goes to the next HS of the studied foot (100% of the gait cycle) [1]. Gait analysis systems are used to evaluate patients' health and study human gait [2], [3] and are an essential clinical tool for a wide range of applications, such as assessing neurological and sports injuries [4], risk of falls [5], prostheses design [6], [7], rehabilitation devices [8], [9] and gait emulators [10], [11]. For optimal performance of such applications, accurate detection of the gait phase is required. For example, Cherell et al. designed the Ankle Mimicking Prosthetic Foot (AMP-Foot) that utilizes torque control of DC motors [12]. Distinct gait events are used to provide the required parameters for controlling the torque of the motors. Yang et al. used two hall effect

The associate editor coordinating the review of this manuscript and approving it for publication was Rajeswari Sundararajan¹.

sensors to detect two phases of the gait cycle and command a fuzzy controller for a passive prosthesis [13].

Gait analysis and estimation are typically conducted in a proctored lab environment equipped with motion-capture systems [14] and force plates [15]. These approaches can provide an accurate measure of the human gait phase; however, they are expensive and require large laboratory spaces [16]. Alternatively, Inertial Measurement Units (IMUs) are utilized in many gait studies due to their low cost and portability. In more recent work, IMUs are attached to human subjects' feet, shanks, or waists to detect HS and TO events [17]. The placement of IMUs on foot is most common since it provides a clear distinction between the swing and stance phases [16], [18], [19]. Different gait phase estimation approaches use online and offline detection methods. Lee and Park used a single-foot IMU to detect heel-strike (HS) (using spikes in linear acceleration) and toe-off (TO) in an offline process [20]. Bae and Tomizuka used Smart shoes embedded with four force sensors (FSR) and predicted the gait phase through an offline Hidden Markov Model [21]. Arvind et al. utilized a hybrid feed-forward neural network and hidden Markov model to estimate the gait phase online from measurements of seven IMUs attached to the subjects' lower extremities. They reported 88.7% detection accuracy of phases of the gait cycle [22]. Zhen et al. used a deep neural network and data from foot, thigh, and calf IMUs to detect 2 phases of the gait cycle offline, reporting an F-score accuracy of 92% [23]. Kim et al. utilized a single-foot IMU and time-frequency analysis to detect HS and TO events offline [16]. Mannini et al. utilized a custom-made wireless foot IMU for online decoding of Hidden Markov Models for Gait Event Detection [24]. Yan et al. used three IMUs (thigh, calf, and foot), principal component analysis (PCA), and a Voting-Weighted Integrated Neural Network to detect HS and TO offline [25]. Quintero et al. proposed using a single thigh IMU to continuously measure gait phase and speed estimation based on a single mechanical parameter derived from human thigh motion [26]. Potter et al. used an array of body-worn IMUs and an error-state Kalman filter to estimate lower-limb kinematics, testing their work on a simplified 3-body model of the human lower limbs [27]. Previous groups typically validated their work with human subjects ranging from 3-20 recruited human subjects. The authors previously used foot and shank magnetic, angular rate, and gravity sensor modules (MARGs) and an error-state Kalman filter to get a real-time estimate of foot and shank orientations, to calculate the ankle angle on 1) a 2-DOF prosthetic ankle, 2) a human ankle [28]. Similarly, Qiu et al. estimated HS and TO using wavelet analysis based on angular data of the thigh and shank IMUs [29].

So far, available algorithms have depended on estimating the gait phase in a proctored indoor lab environment. One advantage of lab settings is that they provide tools to estimate ground truths reliably. However, gait analysis in an indoor lab environment considers multiple assumptions and does not consider all the variability in natural gait that typically

occurs outdoors. Wearable devices with IMUs or combined with other sensors have been implemented on more realistic test conditions. Aminian et al. presented an algorithm for detecting gait events and identifying terrain types based on IMUs and electromyographs (EMGs) on the thigh and shank [30]. Pham et al. used an IMU wearable device on the waist to estimate the stride length of an outdoor walk on a square trajectory [31]. Wearable technology can be used to study numerous activities of daily living not limited to walking; however, PCA is still an effective parameter reduction technique when using machine learning or deep learning models [32]. Other previous efforts to label outdoor gait events used threshold values to detect HS and TO [16], [18], [19]. Alternatively, a smart shoe insole with force sensors is used to identify the stance phase [21], [33]. However, threshold labeling isn't robust to changes in gait cycles and can lead to false labels based on noisy IMU readings. In addition, smart insoles can impede human subjects' natural gait and be sensitive to sweat, causing difficulties in the experimental protocol.

This paper proposes a novel approach to labeling outdoor kinematic data (HS and TO) based on ground-truth-developed patterns of foot IMU data. Generated patterns are driven in an indoor lab environment using a force plate. A gait-based PCA model is developed per subject to reduce feature selection and prevent overfitting and match computational power. The model estimates each subject's gait phase of an outdoor walking session. Furthermore, the models trained on three subjects were tested on a fourth subject.

The rest of this paper is organized as follows. Section II describes the 2-IMU setup on the human subject and the experimental protocol. The experimental protocol includes IMU calibrations, an indoor 8-minute walk, and a 40-minute outdoor walk. Next, section III presents the results of using the generated models to estimate the gait phase and further implications of developed methodologies. Section VI concludes the paper. Finally, Section V provides relevant details, videos, and data required to implement the proposed work. All the recorded human subjects' data is available as open-source data sets that can be used to replicate methodology or other gait analysis methods.

II. METHODOLOGY

A. 2- IMU SET UP

Two IMUs (Precision NXP 9-DOF, Adafruit, USA) are secured to the shank and foot of human subjects, as shown in Fig. 1. The IMUs are attached to the subject using athletic and skin-safe double-sided tape. Furthermore, wires are twisted together and protected using wire sleeves to limit interference with the subject's natural gait. The IMUs are connected to a microcontroller unit (MCU, Teensy 4.1, PJRC, USA) with a built-in SD card placed in a fanny pack around the subject's waist. MCU is powered using a small 3.7V LiPo battery. IMU_1 is connected to the subject's foot, IMU_2 is attached to the subject's shank. The MCU stores data packets of raw

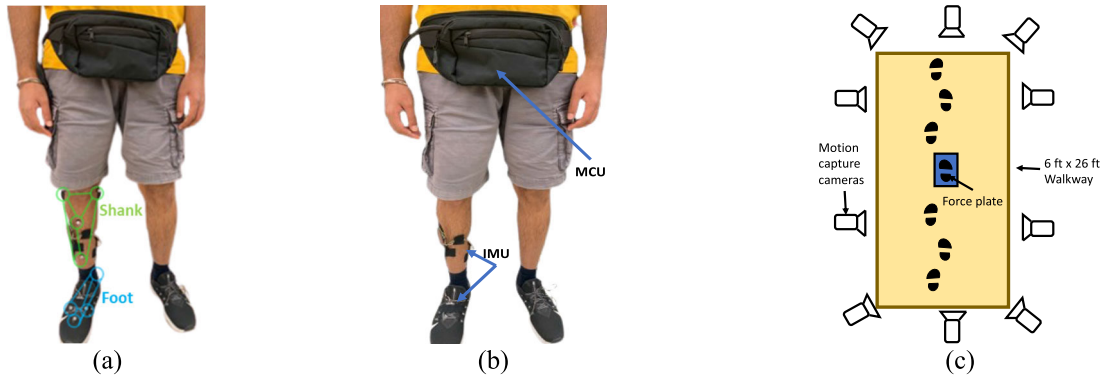


FIGURE 1. 2-IMU setup for human subject gait analysis during (a) OMC IMUs calibration (b) outdoor walk. (c) The walkway was used for experimental protocols 1 and 2.

TABLE 1. Biometrics of human subjects.

Subject	Gender	Age	Height [m]	Weight [kg]
A	Male	23	1.65	80
B	Male	24	1.70	81
C	Male	23	1.81	95
D	Female	20	1.68	54

angular velocities ($\omega_t^{IMU_{1,2}}$) and linear accelerations ($a_t^{IMU_{1,2}}$). The IMUs signals are sampled at 400 Hz. Teensy 4.1 was selected since 1) it comes with a built-in SD Card reader that can write (Arduino SDFat library) to a text file at a high rate (400 Hz), and 2) it can be powered using a small 3.7 V battery. The placement of IMU on the human shank was placed midway between the knee and ankle joints and away from a dominant muscle group (tibialis anterior) to reduce the vibration effect. Foot IMU was centered around the subject's laces. The exact placement of the shank and foot IMU isn't critical, as raw IMU readings are corrected using calibration with a motion capture system that accurately defines a shank and foot frame.

B. EXPERIMENTAL PROTOCOL

The experimental protocol aimed to quantify the gait phase estimation accuracy of human subjects during natural gait in an outdoor environment. To this end, four unimpaired participants (22.5 ± 1.5 years old, 1.71 ± 0.06 m, 77.5 ± 14.81 kg) were enrolled in this study (Table 1), giving written content to participate in the experiment, as approved by the Institutional Review Board (application number 1902021698).

Three experiments were designed with the purpose of 1) calibrating IMUs to correct raw IMU readings, 2) identifying gait patterns through an Indoor walk in a lab environment, and 3) training and testing gait phase estimation models through an outdoor walk. All human subjects participated in all three experiments. Based on the statistical power analysis, 1068 steps are needed for a 95% confidence interval estimation, with a margin of error of $\pm 3\%$. Each subject had at least

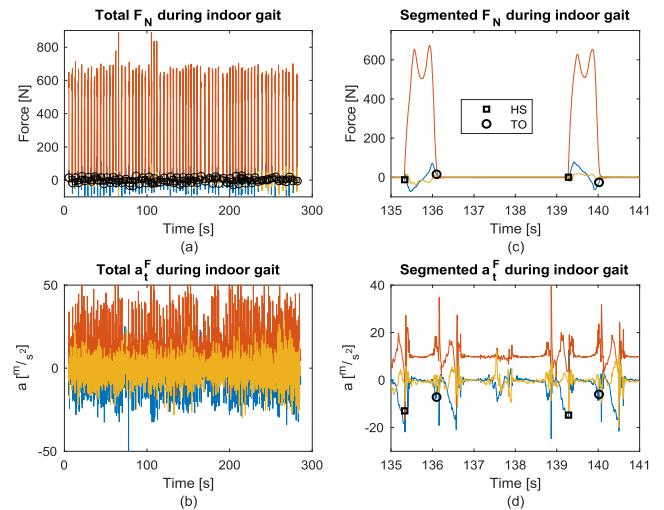


FIGURE 2. Labelled HS and TO events of indoor gait walk of a representative subject (D), using a force plate. (a) shows complete indoor gait data (b) shows a time segment to highlight the force plate labeling.

1100 steps (20 minutes of walking time) on the studied foot to validate the model and computed error.

1) IMU CALIBRATION

The two IMUs are calibrated to account for 1) the coordinate frame rotation between the IMU frames (IMU_1 and IMU_2) and the body frames (foot and shank), 2) scaling errors, and 3) axis misalignment errors of the IMUs [34]. Calibration parameters are used to transform raw IMU measurements ($\omega_t^{IMU_{1,2}}$, $a_t^{IMU_{1,2}}$) to foot (ω_t^F , a_t^F) and shank (ω_t^S , a_t^S) gyroscope and accelerometer measurements as shown in equations 1-4.

Ten optical marker cameras (OMC) (Miquis M5, Qualisys, Sweden) calibrate the IMUs using an extrinsic calibration scheme [28]. Moreover, eight optical markers are placed on the subjects' feet and shank (four each) to define the shank and foot frames. The IMUs are calibrated by exciting the gyroscope and the accelerometer axes in two steps: 1) the human subject raises the foot off the ground and moves their foot and shank linearly in all axes of motion (x, y, z) to

excite the IMUs' accelerometers, then 2) with the foot still in the air, subject rotates foot and shank about all axes of rotation to excite the IMUs' gyroscopes. Each of the IMUs' axes is excited for 15 seconds in each calibration step. The OMC records the shank's and foot's orientations, translations, and body rates. Synchronized measurements from IMUs' ($\omega_t^{IMU_{1,2}}, a_t^{IMU_{1,2}}$) are fitted against OMC measurements based on continuous-time batch estimation to provide the transformation matrices ($T_{Sw}, T_{Sa}, T_{Fw}, T_{Fa}$) between IMU frames and OMC shank and foot frames [35]. The bias parameters ($\overline{b_{\omega}^{IMU_{1,2}}}, \overline{b_a^{IMU_{1,2}}}$) of the IMUs are estimated using Allan Variance analysis [36].

$$\omega_t^F = T_{Fw} \omega_t^{IMU_1} - \overline{b_{\omega}^{IMU_1}} \quad (1)$$

$$a_t^F = T_{Fa} \omega_t^{IMU_1} - \overline{b_a^{IMU_1}} \quad (2)$$

$$\omega_t^S = T_{Sw} \omega_t^{IMU_2} - \overline{b_{\omega}^{IMU_2}} \quad (3)$$

$$a_t^S = T_{Sa} \omega_t^{IMU_2} - \overline{b_a^{IMU_2}} \quad (4)$$

where $T_{Fw}, T_{Fa}, T_{Sw}, T_{Sa} \in \mathbb{R}^{3 \times 3}$, $\omega_t^{IMU_{1,2}}, a_t^{IMU_{1,2}}, \overline{b_{\omega}^{IMU_{1,2}}}, \overline{b_a^{IMU_{1,2}}}, \omega_t^F, a_t^F, \omega_t^S, a_t^S \in \mathbb{R}^3$

2) GAIT PATTERN GENERATION

A gait pattern representative of each subject's gait is produced using an indoor walk. The pattern focuses on relating how the foot IMU (ω_t^F, a_t^F) behaves during heel strike (HS) and toe-off (TO). A walkway equipped with a

force plate sensor (9260AA3, Kistler, Switzerland) is utilized to accurately detect and label HSs and TOs during the human subject's gait. The force plate measures force (F_N) on the 3-axis of motion. Subjects wearing the 2-IMU setup walked repetitively on the walkway, ensuring they stepped on the force plate in each run (Fig 1.c). Subjects walked for eight minutes, during which they performed straight and turning gait cycles, replicating their natural gait as much as possible. Heel strike is identified whenever the vertical force measurement on the force plate exceeds a threshold (2 N), and Toe-off is identified whenever the vertical force drops again below the threshold. The threshold is set based on empirically examining the data to find the instant of time the heel contacts the force plate.

An average gait length was assumed to be 1 second [37], and it takes 4 steps to go from one side of the walkway (embedded with the force plate) to the other, meaning subjects averaged approximately 125 steps on the force plate [38]. Taking more steps (longer time) would increase the strength of the gait pattern however can fatigue the subjects. HSs and TOs identified using the force plate are used to label (HS and TO) adjacent foot IMU (ω_t^F, a_t^F) measurements as shown in Fig. 2. HS and TO patterns are saved for all steps on the force plate. Note that the gait pattern included HS or TO gait event, occurring at time zero and half a second before and after the gait event, respectively. The half a second before and after each gait event are selected to collect data

from approximately half a gait cycle (assuming an average gait cycle of approximately 1 second [37]).

The mean of the generated patterns of the human subjects are shown in Fig 3; the standard deviation is also shown, representing the variance of ω_t^F, a_t^F during different gait phases performed in the eight-minute indoor walk. Increased variance is encouraged as it shows that the human subject replicated the various step trajectories in their natural gait.

3) OUTDOOR WALK

Upon completing the indoor walk, subjects walked outdoors while wearing the 2-IMU setup (Fig. 1.b). Subjects were instructed to walk for forty minutes in any direction, stopping whenever needed. Fig 4. shows the shank and foot IMU measurements during the outdoor walking trial of one of the subjects.

C. GAIT MODEL

A gait model of each subject is used to predict their current gait phase. Note that the aim is to predict the gait phase using measurements from the shank IMU (ω_t^S, a_t^S). On the other hand, the foot IMU (ω_t^F, a_t^F) is only used to label HS and TO on data sets to find the gait phase of an outdoor walk.

1) LABELING OUTDOOR DATA SETS

To produce and train the gait model, previously generated gait patterns are used to identify the HSs and TOs of the outdoor walk on the foot IMU. HS and TO are labeled based on cross-correlation (under MATLAB's `xcorr2` function) between generated HS and TO patterns (Fig. 3) and measured foot IMU data (ω_t^F, a_t^F), as shown in Fig. 5. To distinctively label HS and TO, the IMU data points that have the highest correlation to the pattern (peak) were selected. Note that HS and TO data points are labeled to ensure that peak points are at least 0.8 seconds apart (assuming 0.8 seconds is the minimum gait length [36]). Peak magnitude is at least in the 75th percentile of correlation between pattern and outdoor walk data point (under MATLAB's `findpeaks` function).

The labeled data set (Fig.5) is used to find the outdoor walk gait phase (θ). Linear interpolation between two adjacent heel strike points defines the gait phase. θ equals 0 to represent the 1st heel strike and 2π as the next heel strike (end of the gait cycle). Furthermore, θ is split into two phasors (γ) as shown in equation 5 and Fig 6.a.

$$\gamma = [\sin(\theta) \quad \cos(\theta)] \quad (5)$$

where $\gamma \in \mathbb{R}^2$.

2) TRAINING MODEL USING PCA AND LINEAR REGRESSION

Twenty minutes of the subject's outdoor data is used to create a regression fit model of the gait phase phasors (γ). The independent variables (X) are the Z-score (Z) of shank IMU accelerations and angular velocities (ω_t^S, a_t^S) from one second before heel strike (400 samples) as shown in Fig. 6.b, c and equation 6.

$$X = Z([\ a_t^S \quad \omega_t^S \]) \quad (6)$$

where $X \in \mathbb{R}^{n \times 6}$, n = number of data points.

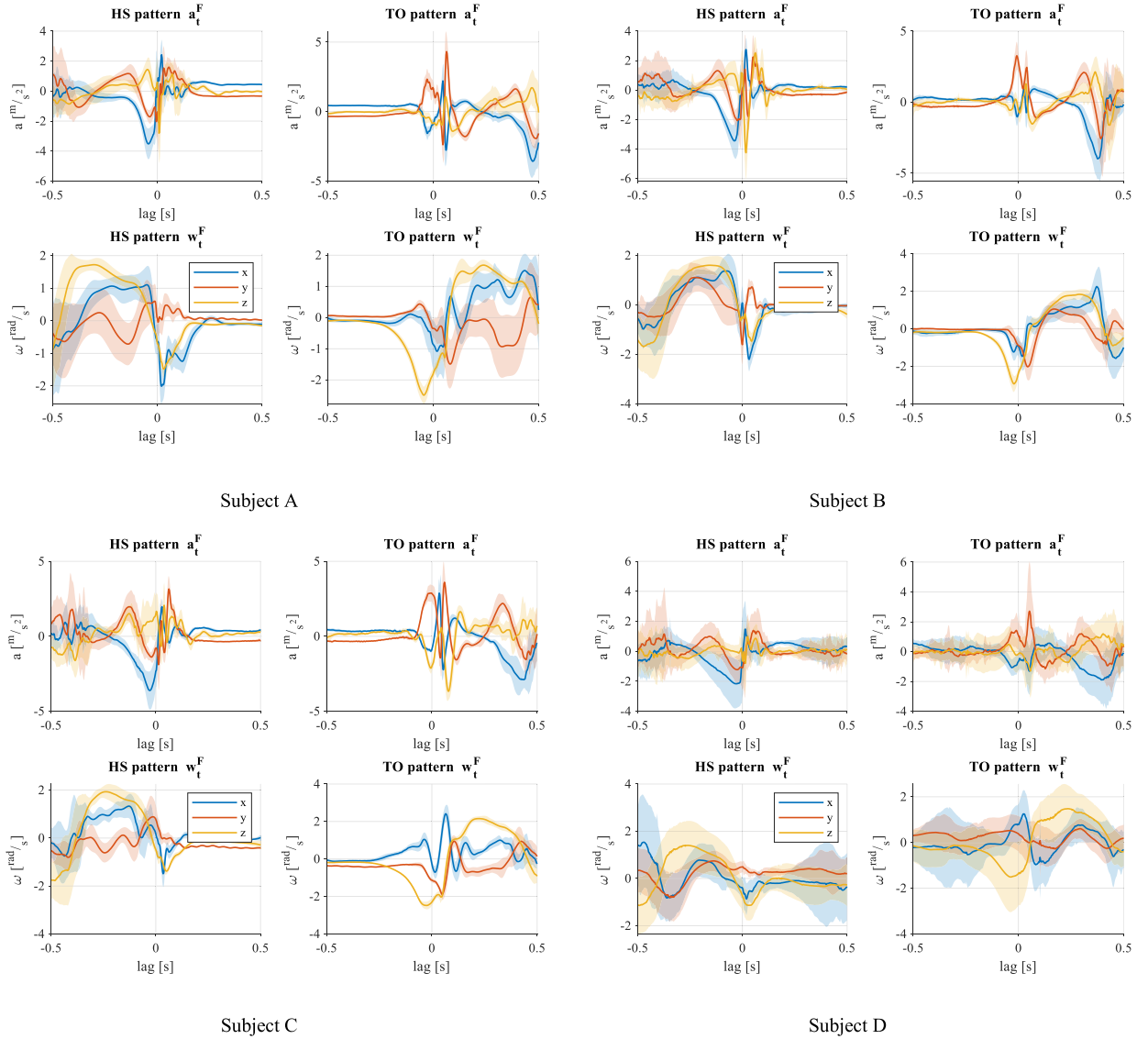


FIGURE 3. HS and TO patterns of 4 human subjects. Subject D had the highest variability in the original gait phase pattern, meaning that the subject's pattern included a wide variety of all the steps the subject takes in their natural gait.

PCA is used to select the principal components of X ; principal components remove noise by reducing many features to a smaller set of principal components. PCA is used because it compresses large amounts of information that speeds up the real-time implementation of models and prevents the overfitting of predictive algorithms [39]. PCA is used since it maximizes the variance of data, meaning we can reduce the number of features in the model while having a varied representation of the subject's gait cycle. Note since IMUs are sampled at 400 Hz, to avoid repetitive data samples, X is down-sampled to 10 Hz. 100 PCA coefficients (PC) are used to select principal features. Down-sampling to 10 Hz and selecting 100 PCA coefficients were empirically tuned under the constraints of the computer's processing power. 100 PCA coefficients are selected to reduce feature size to meet the

constraints of the processing power computer. A high number of PCA coefficients increases the chance of overfitting; reducing the number of coefficients can eliminate important gait features. 100 PCA coefficients were selected experimentally by examining pre and post-PCA feature reduction figures. Down sampled features, X_d are used to fit the model (β) as shown in equations 7-9.

$$X_d = (X - \bar{X}) PC \quad (7)$$

$$\beta = \frac{\gamma}{X_d} \quad (8)$$

$$\hat{\gamma} = X_d \beta \quad (9)$$

where $X_d \in \mathbb{R}^{n \times 6}$, $PC \in \mathbb{R}^{6 \times 100}$.

The gait phase error (e_γ) of the model (β) is calculated using equation 10 (shown for a representative subject

Outdoor walk data

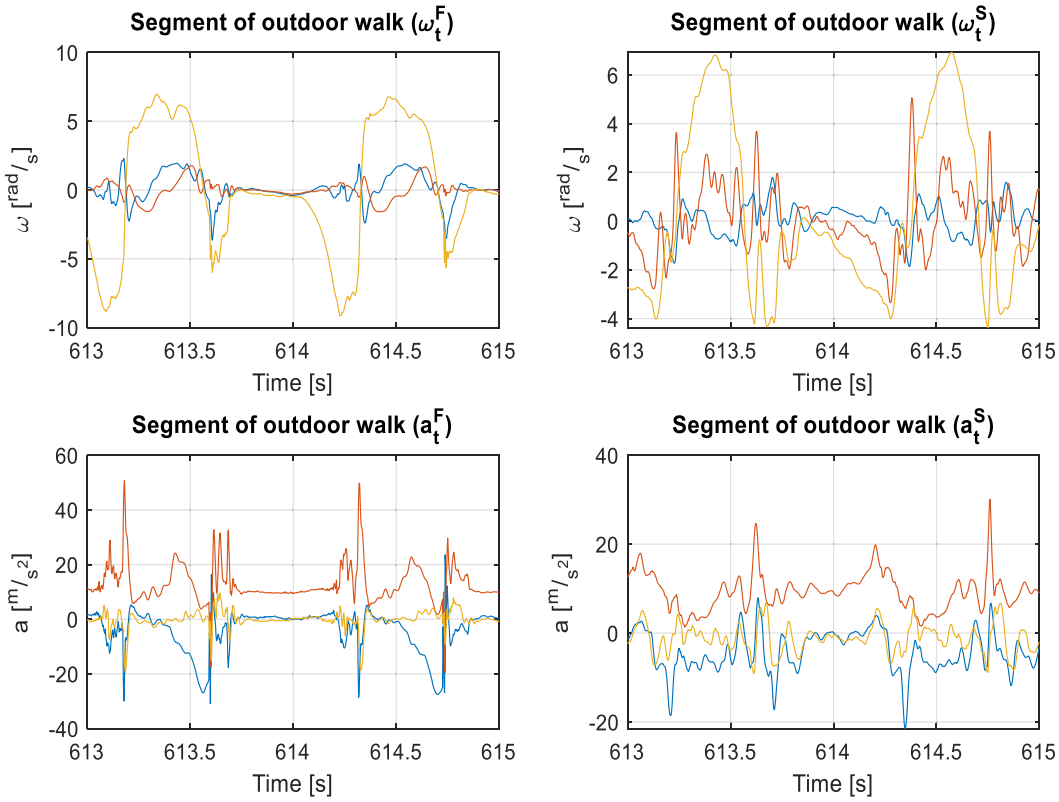


FIGURE 4. Segment of Shank and Foot IMU measurements during the outdoor walking trial of subject D.

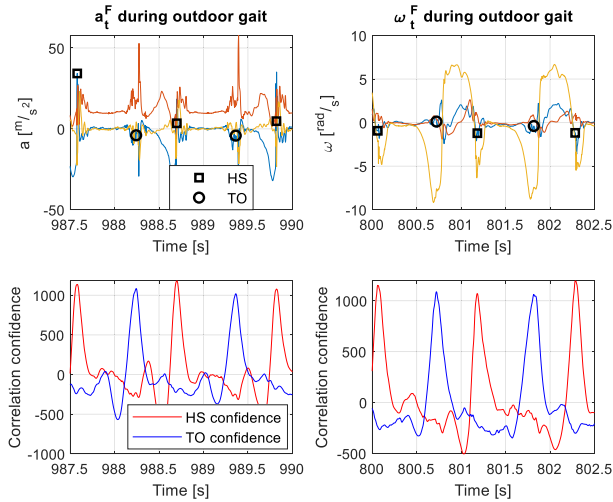


FIGURE 5. Labelling outdoor gait (subject D) with HS and TO events based on previously generated gait patterns. Correlation peaks represent HS and TO gait events.

in Fig 7.a). β is then used to predict the gait phase of the remaining twenty minutes of the outdoor walk, representing the validation data set. The gait phase error (e_γ) for the validation data set of subject D is shown in Fig 7.b.

$$e_\gamma = \frac{\text{acos}\left(\sum_{j=1}^2 \gamma(:,j) \hat{\gamma}(:,j)\right)}{2\pi} \quad (10)$$

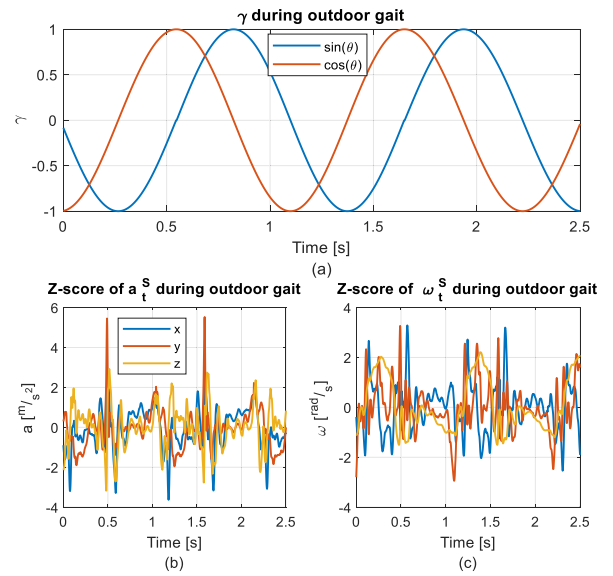


FIGURE 6. Subject D's (a) gait phasors represented by γ , (b, c) linear regression model predictor variables (ω_t^S, a_t^S).

III. RESULTS

A. GAIT PHASE ESTIMATION

The average gait phase estimation error (\bar{e}_γ) for the validation data sets of the human subjects is shown in Table 2. The gait

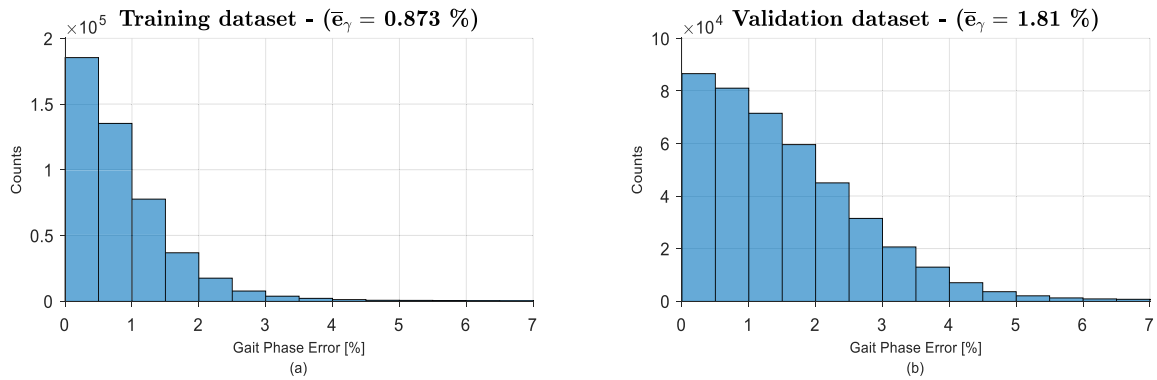


FIGURE 7. Gait phase error for the (a) training dataset (b) validation dataset of subject D represented through a normalized histogram of residuals.

TABLE 2. Average gait phase estimation error (\bar{e}_γ) of the four human subjects.

Subject	\bar{e}_γ (%)
A	2.39
B	2.17
C	2.48
D	1.81

TABLE 3. Average gait phase estimation error (\bar{e}_γ) of the four human subjects trained with varied patterns.

Subject	Gait patterns from subjects	\bar{e}_γ (%)
A	B, C and D	5.90
B	A, C and D	4.06
C	A, B and D	2.29
D	A, B and C	2.48

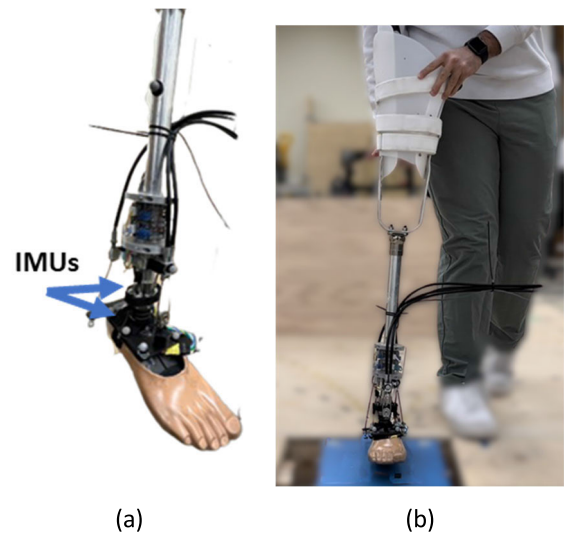


FIGURE 8. (a) IMU setup on 2-DOF prosthesis (b) operator walking with a 2-DOF prosthesis over an instrumented walkway.

phase estimation error (e_γ) for each subject's validation data set was calculated based on a trained model from their gait pattern (Fig 3).

Subject D had the highest gait phase estimation accuracy due to the varied gait pattern observed in the subject's recorded indoor walk. Fig 3.d shows that subject D had the highest standard deviation in HS and TO patterns, meaning the subject tried to replicate all the various steps she typically takes during an outdoor walk. This allows the generation of a fully representative model of the subject's outdoor walk. Conversely, the lowest tracking accuracy is subject C, the subject with the lowest deviation in their gait pattern (Fig 3.c); meaning their indoor gait wasn't fully representative of their gait characteristics.

B. GAIT PHASE ESTIMATION WITHOUT MOTION CAPTURE AND FORCE PLATE PATTERN

Due to similarities in HS and TO patterns between humans (based on foot IMU data) [19], [21], [22], large models of gait patterns can be used to estimate the gait phase of a human

subject. This is useful for outdoor studies that are performed without a motion capture system and a force plate, where it is unable to create a gait pattern for each human subject. Table 3 shows the gait phase estimation error (\bar{e}_γ) for the subjects' outdoor walks when trained with other subjects' concatenated gait patterns. While the accuracy of gait phase estimation is reduced, it's still comparable to other gait phase estimation methods [17]. Increasing the number of concatenated gait patterns decreases the gait phase estimation error of the studied subject, as it's more likely that one of the concatenated patterns is representative of the studied subject's gait. To expand on this work, more gait patterns of multiple subjects can be concatenated and used to estimate the gait phase of subjects without a uniquely defined gait pattern. In addition, IMU calibrations can be done without a motion capture system through repetitive static calibration to correct IMU readings [40].

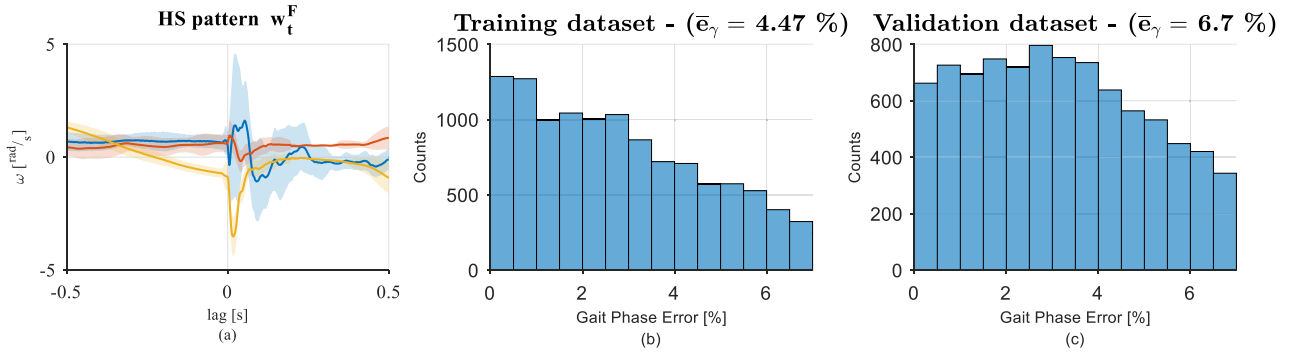


FIGURE 9. (a) HS pattern of 2-DOF prosthesis (b, c) gait phase estimation error of the training and validation data sets.

IV. DISCUSSION

The presented work was designed to help analyze the gait kinematics of healthy and impaired humans in a natural outdoor environment. The 2-IMU setup, besides the ability to accurately detect distinct gait events, provides a robust algorithm that does not require set subject-independent thresholds and labels outdoor gait events based on statistical correlation to ground truth data. Another advantage of the proposed work is that it offers an experimental setup that does not require special shoes/insoles or excessive sensors that can interfere with the subject's natural gait.

The outdoor gait events (HS and TO) are successfully identified based on a ground-truth pattern of each subject. A gait pattern for each subject is accurately defined based on a force plate precisely detecting heel strikes and toe-offs. Outdoor walk data was labeled with HS and TO gait events based on defined patterns. The proposed methodology is advantageous to researchers 1) as it allows them to label gait events harmoniously and with a known correlation to a ground truth measurement without the need for any external validation instrumentation, and 2) as it allows researchers to conduct gait tests outside of a proctored lab space, eliminating any bias related to the lab environment. Another implication of the proposed work is the continuous control of a lower limb prosthesis.

A. PROSTHESIS IMPLEMENTATION

The team previously designed a 2-DOF prosthesis [6] that was operated to detect heel strike upon the impact of the prosthetic ankle with the ground, creating a delay in launching the prosthesis's controller. This approach generally does not consider the users' change of speed or varying their gait length. The PCA-Based can be expanded to control the prosthesis by installing a foot and shank, IMU. In future work, the amputee users will wear a 2-DOF prosthesis and walk indoors to generate a gait pattern for their walk. Furthermore, amputee subjects will wear the prosthesis for a long duration to provide training and validation data sets for the linear regression model. This work can be expanded for real-time

estimation of the gait phase, allowing control of powered prosthesis to have gait trajectories as a function of a time-varying gait phase.

In a limited study, an unimpaired operator carries the prosthesis and walks with it, as shown in Fig 8. The operator walks with the prosthesis for 10 minutes; 4 minutes is used to produce a pattern, 3 minutes to train the model, and 3 minutes to test the model (validation dataset).

Longer times weren't used, as it's difficult for the operator to carry the prosthesis for an extensive period. The HS and TO patterns of ω_t^F . The gait phase estimation error of the training and validation data sets are shown in Fig. 9. Note that PCA estimation and down-sampling weren't used for this estimate due to the small data set size.

Nevertheless, estimation results were promising and needed to be expanded using amputated human subjects, as further deviation is expected. In future work, impaired subjects will use the 2-DOF prosthesis and walk around with it to generate a gait pattern similar to Fig 3. Impaired subjects will then walk outdoors at their natural gait speed. A gait model will be generated for each impaired subject. This is advantageous to amputees as it allows for improved autonomous interaction between amputee and prosthesis. To estimate the gait phase real time, shank data (ω_t^S, a_t^S) from half a second before the gait event (200 samples) can be used to predict when the upcoming gait event. The linear regression model can predict the point of gait based on previous shank IMU data (ω_t^S, a_t^S).

A limitation and an ongoing study of this work is the lack of ability to identify gait activities. Gait activities include turning and incline (stairs) maneuvers. When subjects vary their indoor gait pattern to include several turning maneuvers representing outdoor turns, the gait model can still accurately detect HSs and TOs. For optimal results, subjects should try to replicate as many maneuvers typically done in outdoor walks during the indoor gait trial. Further studies are needed to identify and predict inclination activities; the team conducted a preliminary study using a previously designed ESKF to estimate stair activities [28].

V. CONCLUSION

This paper presents two IMU-based methods that can accurately estimate the gait phase of human subjects based on a one-second time history. The estimation method is based on initially producing a gait pattern representative of each human subject. To produce the pattern, human subjects carry out two experimental protocols: 1) IMU calibration to correct raw IMU readings, transforming readings to be representative of the body frame to the IMU is attached, and 2) an indoor walking experiment where subjects carry out continuous and varying gait cycles while being monitored by a motion capture system and a force plate. The motion capture and force plate readings are used to form the ground truth and are synchronized to foot IMU readings. They produce a gait pattern for heel strike (HS) and toe-off (TO) events. The pattern embodies the gait event and half a second of data before and after the event. Upon completing the indoor walk, subjects walk for 40 minutes in an unsupervised outdoor environment. Subjects walked for 40 minutes because 1) it's long enough to have adequate data sets for training and testing the gait model, 2) subjects didn't walk longer due to risks of battery draining, and 3) adhesives on human shank weakening due to sweat. An outdoor walk experiment is used as it is more representative of a human's natural gait than a lab environment or a treadmill. Data from twenty minutes of the outdoor walk generates a linear regression model representative of the subject's natural gait. The model is built on shank IMU readings; outdoor walk data is labeled with gait events (HS and TO) based on the pre-defined gait pattern of the subject. The model uses PCA estimation to narrow down the large data sets into a 100 representative features, remove any unnecessary noise, and limit its over-training. The model estimates the gait phase of the remaining twenty minutes of the subject's outdoor walk.

Gait phase estimation error for each of the subjects is calculated. The average error for the subjects was $2.21 \pm 0.258\%$. Subject D reported the lowest error due to having the highest variability in the original gait phase pattern, meaning that the subject's pattern included a wide variety of all the steps the subject took in the outdoor environment. The lowest estimation accuracy was reported in subject C, the subject with the lowest variability in their indoor pattern walk (Fig 3). Our team also proposed a method that allows groups without a force plate and a motion capture system to replicate the described work with high accuracy. Concatenated patterns from other human subjects can be used to label outdoor data from a subject that does not have a representative pattern. The more subjects represented in the concatenated pattern, the higher the correlation confidence between the pattern and unlabeled outdoor data. Further application of the proposed work is the autonomous control of an active lower limb 2-DOF prosthesis.

The proposed work allows researchers to study the gait phase of human subjects in an unsupervised outdoor environment. This is advantageous over lab and treadmill testing

as it's a closer representation of a human's natural gait and allows subjects to varying speeds as needed. Furthermore, in terms of labeling gait events (HS and TO), the proposed work offers a unique methodology that does not require fixed thresholds, external sensors (force sensors), or special shoes (sensor-embedded insoles). The labels for HS and TO are based on statistical correlation confidence that is robust to any human subject and does not require embedded sensors that can impose the subject's natural gait. Additionally, the proposed methodology may offer opportunities for extended gait experiments without requiring fully equipped laboratory settings to replicate work. Finally, the methodology can be expanded to aid patients in gait training as proposed regression models can be run offline and in real-time. Presented work can also control wearable assistive devices to offer seamless interaction between humans and robotic devices.

VI. AVAILABILITY

The MATLAB and C deployment of the gait model training and experimental raw data described in the Methods section is available through (<https://github.com/hirolab/gaitphase-estimation>).

REFERENCES

- [1] M. P. Murray, "Gait as a total pattern of movement: Including a bibliography on gait," *Amer. J. Phys. Med. Rehabil.*, vol. 46, no. 1, pp. 290–333, Feb. 1967.
- [2] M. Bellmann, T. Schmalz, and S. Blumentritt, "Comparative biomechanical analysis of current microprocessor-controlled prosthetic knee joints," *Arch. Phys. Med. Rehabil.*, vol. 91, no. 4, pp. 644–652, Apr. 2010, doi: [10.1016/j.apmr.2009.12.014](https://doi.org/10.1016/j.apmr.2009.12.014).
- [3] J. J. Kavanagh and H. B. Menz, "Accelerometry: A technique for quantifying movement patterns during walking," *Gait Posture*, vol. 28, no. 1, pp. 1–15, Jul. 2008, doi: [10.1016/j.gaitpost.2007.10.010](https://doi.org/10.1016/j.gaitpost.2007.10.010).
- [4] R. Ferber and S. Macdonald, *Running Mechanics and Gait Analysis*, 1st ed. Champaign, IL, USA: Human Kinetics, 2014.
- [5] 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, Aug. 2001, doi: [10.1053/apmr.2001.24893](https://doi.org/10.1053/apmr.2001.24893).
- [6] H. Dallali, E. Ficanha, and M. R. Aagaah, "Dynamic modeling of a 2-DOF cable driven powered ankle-foot prosthesis," in *Proc. ASME Dyn. Syst. Control Conf. (DSCC)*, Minneapolis, MN, USA, Oct. 2016, Art. no. V001T06A003.
- [7] D. Quintero, D. J. Villarreal, D. J. Lambert, S. Kapp, and R. D. Gregg, "Continuous-phase control of a powered knee-ankle prosthesis: Amputee experiments across speeds and inclines," *IEEE Trans. Robot.*, vol. 34, no. 3, pp. 686–701, Jun. 2018, doi: [10.1109/TRO.2018.2794536](https://doi.org/10.1109/TRO.2018.2794536).
- [8] S. Yul Shin, A. D. Deshpande, and J. Sulzer, "Design of a single degree-of-freedom, adaptable electromechanical gait trainer for people with neurological injury," *J. Mech. Robot.*, vol. 10, no. 4, Aug. 2018, Art. no. 044503, doi: [10.1115/1.4039973](https://doi.org/10.1115/1.4039973).
- [9] A. Alamdari and V. Krovi, "Design and analysis of a cable-driven articulated rehabilitation system for gait training," *J. Mech. Robot.*, vol. 8, no. 5, May 2016, Art. no. 051018, doi: [10.1115/1.4032274](https://doi.org/10.1115/1.4032274).
- [10] A. Soliman, G. A. Ribeiro, D. Gan, and M. Rastgaar, "Feasibility design and control of a lower leg gait emulator utilizing a mobile 3-revolute, prismatic, revolute parallel manipulator," *J. Mech. Robot.*, vol. 15, no. 1, pp. 1–24, Feb. 2023, doi: [10.1115/1.4053825](https://doi.org/10.1115/1.4053825).
- [11] E. Ficanha, H. Dallali, and M. Rastgaar, "Gait emulator for evaluation of a powered ankle-foot prosthesis," in *Proc. Dyn. Syst. Control Conf.*, Oct. 2017.

- [12] P. Cherelle, V. Grosu, L. Flynn, K. Junius, M. Moltedo, B. Vanderborght, and D. Lefeber, "The ankle mimicking prosthetic foot 3—Locking mechanisms, actuator design, control and experiments with an amputee," *Robot. Auto. Syst.*, vol. 91, pp. 327–336, May 2017, doi: [10.1016/j.robot.2017.02.004](https://doi.org/10.1016/j.robot.2017.02.004).
- [13] P. Yang, H. Yue, L. Chen, and Y. Geng, "Intelligent lower limb prosthesis following healthy leg gait based on fuzzy control," in *Proc. 24th Chin. Control Decis. Conf. (CCDC)*, May 2012, pp. 3729–3731, doi: [10.1109/CCDC.2012.6244597](https://doi.org/10.1109/CCDC.2012.6244597).
- [14] T. B. Moeslund, C. E. Mahon, and A. L. Pruziner, "A survey of advances in vision-based human motion capture and analysis," *Comput. Vis. Image Understand.*, vol. 104, nos. 2–3, pp. 90–126, 2006, doi: [10.1016/j.cviu.2006.08.002](https://doi.org/10.1016/j.cviu.2006.08.002).
- [15] B. D. Hendershot, C. E. Mahon, and A. L. Pruziner, "A comparison of kinematic-based gait event detection methods in a self-paced treadmill application," *J. Biomechanics*, vol. 49, no. 16, pp. 4146–4149, Dec. 2016, doi: [10.1016/j.jbiomech.2016.10.046](https://doi.org/10.1016/j.jbiomech.2016.10.046).
- [16] J. Kim, M. Bae, K. B. Lee, and S. G. Hong, "Gait event detection algorithm based on smart insoles," *ETRI J.*, vol. 42, no. 1, pp. 46–53, Feb. 2020, doi: [10.4218/etrij.2018-0639](https://doi.org/10.4218/etrij.2018-0639).
- [17] H. T. T. Vu, D. Dong, H.-L. Cao, T. Verstraten, D. Lefeber, B. Vanderborght, and J. Geeroms, "A review of gait phase detection algorithms for lower limb prostheses," *Sensors*, vol. 20, no. 14, p. 3972, Jul. 2020, doi: [10.3390/s20143972](https://doi.org/10.3390/s20143972).
- [18] S. Khandelwal and N. Wickström, "Gait event detection in real-world environment for long-term applications: Incorporating domain knowledge into time-frequency analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 12, pp. 1363–1372, Dec. 2016, doi: [10.1109/TNSRE.2016.2536278](https://doi.org/10.1109/TNSRE.2016.2536278).
- [19] S. Mitsu, T. Asai, R. Ono, R. Sawa, K. Tsutsumimoto, H. Ando, and T. Doi, "Development and validity of methods for the estimation of temporal gait parameters from heel-attached inertial sensors in younger and older adults," *Gait Posture*, vol. 57, pp. 295–298, Sep. 2017, doi: [10.1016/j.gaitpost.2017.06.022](https://doi.org/10.1016/j.gaitpost.2017.06.022).
- [20] J. K. Lee and E. J. Park, "Quasi real-time gait event detection using shank-attached gyroscopes," *Med. Biol. Eng. Comput.*, vol. 49, no. 6, pp. 707–712, 2011, doi: [10.1007/s11517-011-0736-0](https://doi.org/10.1007/s11517-011-0736-0).
- [21] J. Bae and M. Tomizuka, "Gait phase analysis based on a hidden Markov model," *Mechatronics*, vol. 21, no. 6, pp. 961–970, Sep. 2011, doi: [10.1016/j.mechatronics.2011.03.003](https://doi.org/10.1016/j.mechatronics.2011.03.003).
- [22] R. L. Evans and D. K. Arvind, "Detection of gait phases using orient specks for mobile clinical gait analysis," in *Proc. 11th Int. Conf. Wearable Implant. Body Sensor Netw.*, Jun. 2014, pp. 149–154, doi: [10.1109/BSN.2014.22](https://doi.org/10.1109/BSN.2014.22).
- [23] T. Zhen, L. Yan, and P. Yuan, "Walking gait phase detection based on acceleration signals using LSTM-DNN algorithm," *Algorithms*, vol. 12, no. 12, p. 253, Nov. 2019, doi: [10.3390/a12120253](https://doi.org/10.3390/a12120253).
- [24] A. Mannini, V. Genovese, and A. M. Sabatini, "Online decoding of hidden Markov models for gait event detection using foot-mounted gyroscopes," *IEEE J. Biomed. Heal. Informat.*, vol. 18, no. 4, pp. 1122–1130, Jul. 2014, doi: [10.1109/JBHI.2013.2293887](https://doi.org/10.1109/JBHI.2013.2293887).
- [25] L. Yan, T. Zhen, J.-L. Kong, L.-M. Wang, and X.-L. Zhou, "Walking gait phase detection based on acceleration signals using voting-weighted integrated neural network," *Complexity*, vol. 2020, pp. 1–14, Jan. 2020, doi: [10.1155/2020/4760297](https://doi.org/10.1155/2020/4760297).
- [26] D. Quintero, D. J. Lambert, D. J. Villarreal, and R. D. Gregg, "Real-time continuous gait phase and speed estimation from a single sensor," in *Proc. IEEE Conf. Control Technol. Appl. (CCTA)*, Aug. 2017, pp. 847–852, doi: [10.1109/CCTA.2017.8062565](https://doi.org/10.1109/CCTA.2017.8062565).
- [27] M. V. Potter, S. M. Cain, L. V. Ojeda, R. D. Gurchiek, R. S. McGinnis, and N. C. Perkins, "Error-state Kalman filter for lower-limb kinematic estimation: Evaluation on a 3-body model," *PLoS ONE*, vol. 16, no. 4, Apr. 2021, Art. no. e0249577, doi: [10.1371/journal.pone.0249577](https://doi.org/10.1371/journal.pone.0249577).
- [28] A. Soliman, G. A. Ribeiro, A. Torres, and M. Rastgaar, "Error-state Kalman filter for online evaluation of ankle angle," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2022, pp. 225–231, doi: [10.1109/AIM52237.2022.9863327](https://doi.org/10.1109/AIM52237.2022.9863327).
- [29] S. Qiu, H. Zhao, N. Jiang, Z. Wang, L. Liu, Y. An, H. Zhao, X. Miao, R. Liu, and G. Fortino, "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges," *Inf. Fusion*, vol. 80, no. C, pp. 241–265, Apr. 2022, doi: [10.1016/j.inffus.2021.11.006](https://doi.org/10.1016/j.inffus.2021.11.006).
- [30] K. Aminian, B. Najafi, C. Büla, P. F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *J. Biomech.*, vol. 35, no. 5, pp. 689–699, 2002, doi: [10.1016/S0021-9290\(02\)00008-8](https://doi.org/10.1016/S0021-9290(02)00008-8).
- [31] T. T. Pham and Y. S. Suh, "Conditional generative adversarial network-based regression approach for walking distance estimation using waist-mounted inertial sensors," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, 2022.
- [32] A. Talitskii, E. Kovalenko, A. Shcherbak, A. Anikina, E. Bril, O. Zimniakova, M. Semenov, D. V. Dyllov, and A. Somov, "Comparative study of wearable sensors, video, and handwriting to detect Parkinson's disease," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–10, 2022, doi: [10.1109/TIM.2022.3176898](https://doi.org/10.1109/TIM.2022.3176898).
- [33] V. Agostini, G. Balestra, and M. Knaflitz, "Segmentation and classification of gait cycles," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 5, pp. 946–952, Sep. 2014, doi: [10.1109/TNSRE.2013.2291907](https://doi.org/10.1109/TNSRE.2013.2291907).
- [34] P. D. Groves, "Principles of GNSS, inertial, and multisensor integrated navigation systems," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 30, no. 2, pp. 26–27, Feb. 2007.
- [35] P. Furgale, J. Rehder, and R. Siegwart, "Unified temporal and spatial calibration for multi-sensor systems," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 1280–1286.
- [36] *IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Interferometric Fiber Optic Gyros*, IEEE Standard 952-1997, Feb. 1998, pp. 1–84, doi: [10.1109/IEEESTD.1998.86153](https://doi.org/10.1109/IEEESTD.1998.86153).
- [37] M. Pat, B. Drought, and R. Kory, "Walking patterns of normal men," *J. Bone Joint Surg.*, vol. 46, no. 2, pp. 335–360, Mar. 1964.
- [38] J. T. Cavanaugh and N. Stergiou, "Gait variability: A theoretical framework for gait analysis and biomechanics," in *Biomechanics and Gait Analysis*, N. Stergiou, Ed. New York, NY, USA: Academic, 2020, ch. 8, pp. 251–286, doi: [10.1016/B978-0-12-813372-9.00008-7](https://doi.org/10.1016/B978-0-12-813372-9.00008-7).
- [39] I. T. Jolliffe and J. Cadima, "Principal component analysis: A review and recent developments," *Phil. Trans. R. Soc. A*, vol. 374, Apr. 2016, Art. no. 20150202, doi: [10.1098/rsta.2015.0202](https://doi.org/10.1098/rsta.2015.0202).
- [40] D. Tedaldi, A. Pretto, and E. Menegatti, "A robust and easy to implement method for IMU calibration without external equipments," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Hong Kong, May 2014, pp. 3042–3049, doi: [10.1109/ICRA.2014.6907297](https://doi.org/10.1109/ICRA.2014.6907297).



AHMED SOLIMAN (Graduate Student Member, IEEE) received the B.S. degree in mechatronics and robotics engineering technology from Purdue University, West Lafayette, IN, USA, in 2019, where he is currently pursuing the Ph.D. degree in robotics engineering technology.

Since 2019, he has been worked as a Research Assistant at the Human Interactive Robotics Laboratory (HIROLab), Purdue University, and a Teaching Assistant for undergraduate college classes on embedded circuits and manufacturing techniques. His research interests include applications of autonomous vehicles and parallel manipulators in the testing of human assistive robots.



GUILHERME A. RIBEIRO received the B.S. degree in mechatronics engineering from the Federal University of Uberlândia, Minas Gerais, Brazil, in 2013, and the M.S. and Ph.D. degrees in mechanical engineering from Michigan Technological University (MTU), Houghton, MI, USA, in 2019.

From 2014 to 2019, he was a Research Assistant with the Human Interactive Robotics Laboratory and a Teaching Assistant for college classes on controls and robotics at the Mechanical Engineering Department, MTU. Since 2019, he has been a Postdoctoral Researcher with the Polytechnic Institute, Purdue University, West Lafayette, IN, USA. His research interests include the control and automation of robots and system identification of dynamical systems.

Dr. Ribeiro has received awards as the Outstanding Graduate Student Teacher, the first place in Consumer Products Day Competition, the Tutoring Assistantships for programming and math courses, and the Medalist of the Brazilian Mathematics Olympics.



ANDRES TORRES received the B.S. degree from the Department of Mechanical Engineering, University of Puerto Rico, Mayaguez, Puerto Rico, in 2019. He is currently pursuing the Ph.D. degree with Purdue University.

He has interned at P&G–Gillette, Boston, MA, USA, in Summer 2017, and an Undergrad Researcher at Boston University, Boston, in Summer 2018. He was a PR-LSAMP Scholar, from 2017 to 2019. His major research inter-

ests include soft-robotic joints, system identification, and human–robot interactions.



LI-FAN WU received the B.S. degree from the Department of Mechanical Engineering, National Cheng Kung University, Tainan, Taiwan, in 2015, and the M.S. degree from the Department of Electrical Engineering, National Cheng Kung University, in 2017. He is currently pursuing the Ph.D. degree with Purdue University.

He has completed the one-year M.S. Exchange Program at the Department of Electrical and Computer Engineering, Technical University of

Munich, Bavaria, Germany, in 2018. He was a Software Engineer at Wistron Enterprise, in 2019. His major research interests include fuzzy control, motion planning, machine learning, intelligent systems, a humanoid robot, image processing, robotic application, and FIRA/Robo cup games.



MO RASTGAAR (Senior Member, IEEE) received the B.S. degree from the Sharif University of Technology, Tehran, Iran, in 1995, the M.S. degree from Tehran Polytechnic, Iran, in 1998, and the Ph.D. degree from the Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 2008, all in mechanical engineering.

He was a Postdoctoral Associate at the Newman Laboratory for Biomechanics and Human Rehabilitation, Massachusetts Institute of Technology,

Cambridge, MA, USA, from 2008 to 2010. From 2011 to 2018, he was an Assistant Professor and an Associate Professor at Michigan Tech, Houghton, MI, USA. In 2019, he joined Purdue University, West Lafayette, IN, USA, where he is currently a Professor at the Polytechnic Institute and the Director of the Human–Interactive Robotics Laboratory. His research interests include advancing maneuverability in lower extremity robotic prostheses and assistive robots by characterizing the agility in the human gait. He was a recipient of the 2014 National Science Foundation CAREER Award.

...