# Feature Selection, Construction, and Validation of a Lightweight Model for Foot Function Assessment During Gait With In-Shoe Motion Sensors

Chenhui Huang, *Member, IEEE*, Fumiyuki Nihey, Kenichiro Fukushi, Hiroshi Kajitani, Yoshitaka Nozaki, Kazuki Ihara, and Kentaro Nakahara

***Abstract*—Identifying and monitoring overpronation and oversupination in the long term during activities of daily living are essential for people's ambulatory health. Using an in-shoe motion sensor (IMS) with power-saving functions is a potential solution. In this study, we challenged the development of an estimation model of foot function using the foot center of pressure excursion index (CPEI) as an index via linear multivariate regression, which is sufficiently light for this type of IMS. Data collected from 65 and 17 participants were involved in model construction and validation, respectively. We validated ten scenarios simulating daily living activities, including walking on different surfaces, using different shoes, with or without carrying a bag, and indoors and outdoors. We applied statistical parametric mapping (SPM) to determine significant predictors and performed our original feature selection algorithm, leave-one-subject-out least absolute shrinkage and selection operator (LASSO), to compress the volume of the predictors. We successfully discovered significant sex-specific predictors for foot function estimation from foot motion and constructed large effect-sized sex-specific foot function estimation models that achieved high-precision CPEI estimation. In the validation, the model successfully estimated a maximum of 99.0% and 100.0% males' and females' data under the same experimental conditions with the training data and 92.8%–100.0% and 85.8%–100.0% data in different scenarios. The constructed models are effective and possible to provide applications for long-term foot function monitoring by using an IMS.

Index Terms— Foot motion, gait analysis, in-shoe motion sensor (IMS), overpronation.
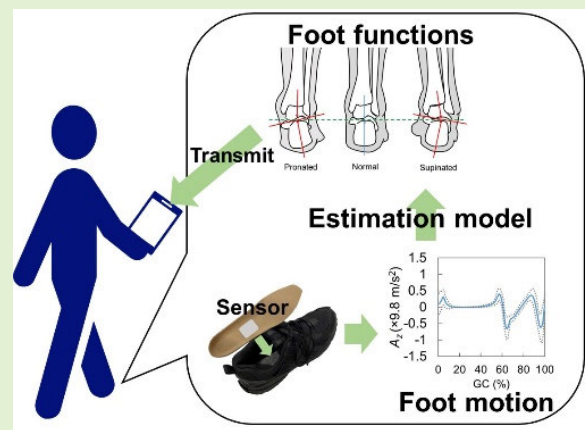
I. INTRODUCTION

THERE are many differences in foot morphology and biomechanics between individuals. Previous studies have shown that foot shape significantly affects foot function, and abnormal foot structures even pose the risk of lower limb injuries [1], [2]. Overpronation and oversupination, which are directly related to low and high foot-arch types, are the main categories of extreme foot functions, as shown in Fig. 1. Prior studies have suggested a strong association between these extreme foot functions, highlighting foot symptoms, fall risks, and disabilities such as hallux valgus, shin splints, and even knee osteoarthritis [3], [4], [5], [6], [7], [8].

In daily living, people tend to overpronate or oversupinate their feet to deal with ill-fitting shoes and uneven ground surfaces during walking or running. Overpronation or over-supination may induce exacerbation due to overuse or weakness of several muscles in the lower extremities. In contrast, most people do not recognize it independently during daily activities [9], [10], [11]. Fortunately, previous studies has shown that abnormal foot functions can be corrected by purchasing motion-control shoes or insoles or by providing
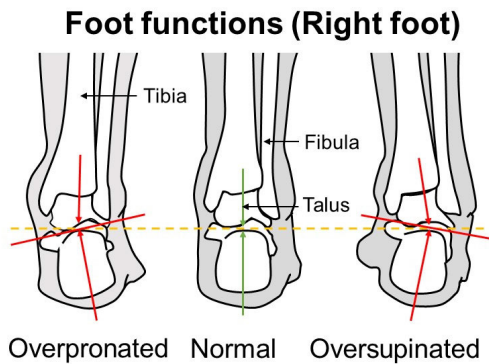
Foot functions (Right foot)

Fig. 1. Schematics of overpronated/oversupinated foot functions.

appropriate physical training if people know their current foot function [9], [12], [13]. Therefore, identifying and monitoring people in the long term during daily living activities should be essential for people's ambulatory health.

Conventionally, foot-type assessments relied on clinicians' examinations or expensive instruments in experimental environments [1]. Foot pronation/supination was commonly assessed through methods such as measuring the rearfoot angle in the frontal plane using a goniometer, X-ray photography, or motion capture marker [10], [14], [15]; morphologically rating the feet using the foot posture index [16]; and calculating the center of pressure excursion index (CPEI) from a dynamic footprint measured by a pressure sensor array [17], [18]. Compared with other measurement methods, overpronation or oversupination assessment using the CPEI was considered more promising because it is highly reliable and removes reliance on clinical experts [17], [18]. Low or negative CPEI values indicate overpronation, and extremely high values indicate oversupination.

Smart wearable sensors have been developed to avoid environmental constraints and reduce measurement costs. Following the concept of CPEI-based assessment, smart instrumented footwear insoles with multiple flexible pressure sensors to classify foot types were proposed as a solution [19], [20]. In addition, smart shoes/insoles with motion sensors inside were also considered promising in various healthcare applications via daily gait analysis, such as Parkinson's disease, gait rehabilitation, frailty detection, and even foot deformity detection [21], [22], [23]. Here, we refer to this smart motion sensor as an "in-shoe motion sensor (IMS)." Compared to foot pressure devices, by processing inertial measurement unit (IMU) signals, in addition to common spatiotemporal gait parameters (GPs), an IMS enables to spontaneously obtain more abundant information, including instantaneous linear and rotational motion and 3-D foot angular posture [21], [24], [25], which are intimately related to gait kinematics.

In general, IMSs can or wirelessly transmit detailed waveforms to a smartphone or server for further analysis with high power consumption. The wireless transmission requires IMSs to be routinely and frequently charged, thus losing usability. In our previous study, we developed a new type of IMS that is small, lightweight, and can be attached to the insoles.

We demonstrated that with optimally designed power-saving operation sequences for IMS and operation mode in practical applications, IMSs could achieve high usability for long-term daily measurement without needing charge the battery for one year [24]. One key point of power saving is that our IMS can finish simple data processing and calculate simple parameters on the edge device. We produced this type of IMS system as a product named A-RROWG.[1] The features mentioned above enable this kind of IMS to collect daily gait data over long periods, regardless of location and time, without the user being aware of the sensor being attached. We considered installing an assessment model on this kind of IMS would be a solution for convenient long-term monitoring of foot function.

Therefore, we challenged the development of a lightweight foot function estimation model that is feasible for an IMS via linear multivariate regression. We chose the CPEI as the foot function reference index and hypothesized that the actual state of the foot function should be expressed via gait. Lots of potential predictors were proposed in previous studies [26], [27], [28], [29], [30]. Previous reports have shown potential predictors for the foot function assessment model, including weight or body mass index (BMI) [26] and the gait kinematics calculated from multiple segments of the lower limbs. Regarding gait kinematics, Chuter [27] discovered that the maximum rearfoot eversion in subjects with overpronated feet is more apparent than normal ones during gait. In contrast, Zhang et al. [28] additionally discovered that overpronated subjects also have greater rearfoot eversion at heel strike (HS) and toe-off (TO) than normal ones. The temporal series kinematics studies of Houck et al. [29] and Levinger et al. [30] indicated that pronated feet have more eversion and medial tibia tilt during the stance phase. However, since the relationship between foot function and motion remains to be clarified, and studies concerning foot function assessment using only foot motion during gait are still rarely seen, more effective predictors need to be determined from foot motion.

Gait is a periodic motion. We considered that the same foot motions concerning foot function are also repeated at the same stage in every gait cycle (GC). Hence, predictors can be determined from gait patterns during specific gait phases associated with foot function. We previously presented the correlation between foot function and inertial foot motion signals via linear correlation analysis and constructed the primary model for foot function estimation at the 2021 IEEE Biomedical Circuits and Systems Conference (BioCAS) [31]. In this article, we improved our model via the approaches.

1) Conducting statistical parametric mapping (SPM) has proven effective in biomechanical studies [32], [33] to determine effective predictors of foot motion.
2) Using demographic data, including age, height, weight, BMI, and several GPs [24] as auxiliary predictors to improve model precision.
3) Using an original method called "leave-one-subject-out least absolute shrinkage and selection operator" (LOSO-LASSO) to reduce redundant predictors and select appropriate predictors.

[1]Trademarked.

We considered our approach first makes the model light-weighted enough to fit IMS for long-term healthcare applications, second helps us to elucidate the correlation between foot motion and foot function that makes the model biomechanically interpretable, and third makes model robust against the individual difference of gait.

After establishing the model, we also validated it in simulated practical use case scenarios. Different ground surfaces, shoes, and carrying weight influence gait kinematics [34], [35], [36]. For practical use cases in which data can be collected at any time, place, and status, we selected common scenarios, including using different shoes, walking on different surfaces, and carrying different bags to validate the model and the precision in different use cases. Considering the musculoskeletal structure differences between sexes, we will construct a CPEI estimation model for males and females, respectively. Please note that at this study's current feasibility study stage, we conduct all the data processing and model construction tasks in an offline environment (on a personal computer).

II. COMPARED WITH RELATED WORKS

Utilizing the biomechanics knowledge above, Miezal et al. [37] proposed a foot function assessment method based on multiple IMUs mounted on different lower limb segments. Compared with the method proposed in [37], for the IMSs, when we assumed the mid-foot and hindfoot connected as a rigid body, they measured foot motions that originated from only a single segment. Thus, models for using multiple sensors were not suitable for the IMS.

Chen et al. [38] suggested that with the aid of foot pressure sensors, oversupination can be estimated via fast Fourier transforming (FFT) approximate five-second-long walking gait signal patterns measured by a motion sensor only mounted on foot to generate predictors and inputting them to a support vector machine model by then. However, on one side, the type of sensor like A-RROWG does not rely on any other auxiliary sensor devices and requires the computation volume of any GP to be sufficiently light making their approach unsuitable to the IMS. On the other side, in practical use, interpretable models were required in potential healthcare and clinical applications. Therefore, rather than machine learning models, which are always a black box and contain complicated data processing and computations, we considered a linear multivariate regression model should be a feasible candidate.

There are plenty of techniques for feature selection, including LASSO [39], Bayesian methods, e.g., Bayesian LASSO [40], and deep learning methods for sparse learning [41] which have been proposed. Compared to LASSO or Bayesian LASSO, deep learning methods need relatively large training data to tune each weight in the hidden layer, while biomedical datasets usually have smaller sizes. Bayesian LASSO accounts for uncertainty in the choice of regularization parameter and may lead to more accurate prediction than LASSO. However, in contrast, it may require more expertise to interpret.

In LASSO, cross-validation approaches [42] were commonly used for choosing the value of the LASSO

TABLE I
EXPERIMENTAL SCENARIOS FOR VALIDATION DATA COLLECTION

Scenario No.	Indoor/Outdoor	Surface	Shoes type	Bag type
S1	Indoor	Vinyl flooring	Standard	No bag
S2	Indoor	Vinyl flooring	Standard	Rucksack
S3	Indoor	Vinyl flooring	Standard	Shoulder
S4	Indoor	Vinyl flooring	Standard	Hand
S5	Indoor	Vinyl flooring	Participant's	No bag
S6	Outdoor	Asphalt	Participant's	No bag
S7	Outdoor	Asphalt	Participant's	Rucksack
S8	Outdoor	Asphalt	Standard	No bag
S9	Outdoor	Gravel road	Standard	No bag
S10	Outdoor	Lawn	Standard	No bag

tuning parameter. However, conventionally, training sets and validation sets in cross-validation were randomly decided without considering individual differences existed in data. Therefore, to deal with individual differences and ensure the model robust against other subjects, we combined a LOSO process with LASSO. Conceptually near the jackknife resampling method [43], we obtained multiple LASSO analysis results by looping the LOSO process for all subjects. By statistically analyzing these results, we can approximate the nature of the population estimator and thus make the LASSO analysis more robust against individual differences.

III. MATERIALS AND METHODS

A. Participants

We recruited two groups of participants of different sexes, ages, heights, and weights; all of their attributes were recorded before the experiments. We successfully collected data from 65 participants (33 males and 32 females), forming the training data for model construction. Hereafter, we refer to this group as Group T. We also successfully collected the data from 17 separately recruited participants (eight males and nine females), forming the other group whose data were used for model validation. Hereafter, we refer to this group as Group V.

All participants could walk independently without assistive devices such as a cane, crutches, or orthotic devices. They had a normal or corrected-to-normal vision, no history of neuromuscular or orthopedic disorders, and no communication impairments. Before the experiments, the experimental procedure was explained to all participants, and then, we obtained their informed consent. The NEC Ethical Review Committee for Life Sciences approved the study on the December 1, 2021 (Approval number: LS2021-015).

B. Other Recommendations

1) Experiment for Training Data Collection: Considering that users could walk at any speed in the practical use case, participants in Group T were asked to walk in a 15-m indoor straight vinyl flooring path at a slower, regular, and faster speed for three round trips (six trials) at each speed. The order of required walking speed was shuffled for each participant. We asked the participants to wear the sports shoes we prepared, which we defined as standard shoes.
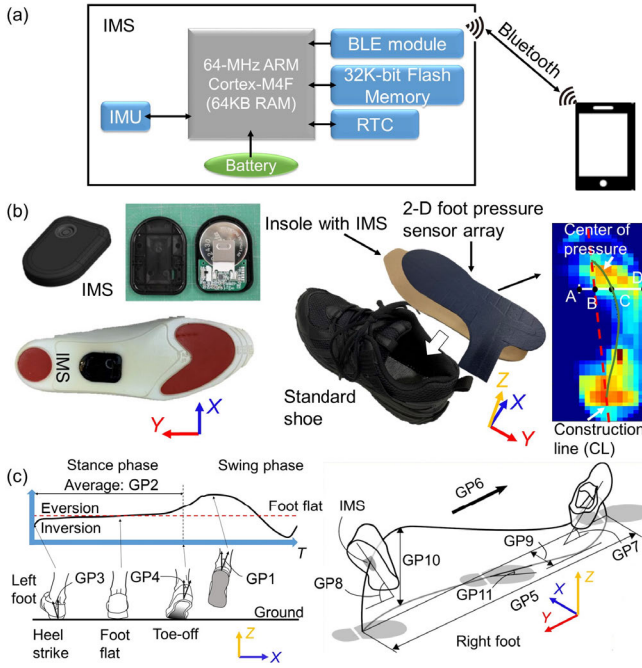
Fig. 2.   (a) Overview of IMS system. (b) Schematic of IMS-equipped insole and experimental setup in the shoe. Definition of coordinate axes and corresponding plane with directions defined as A_x (medial: +, lateral: −), A_y (posterior: +, anterior: −), A_z (superior: +, inferior: −), G_x (sagittal plane, plantarflexion: +, dorsiflexion: −), G_y (frontal plane, eversion: +, inversion: −), G_z (horizontal plane, adduction: +, abduction: −), E_x (roll angle, plantarflexion: +, dorsiflexion: −), E_y (pitch angle, eversion: +, inversion: −), and E_z (yaw angle, internal rotation: +, external rotation: −), and yz: sagittal plane, xz: frontal plane, and xy: horizontal plane. In the foot pressure image, line AD was constructed at one-third of the foot from the distal end. The construction line (red dashed line) is the line passing through the start and terminal points of the CoP curve. The distance between the intersection of the construction line and line AD (Point B) and where the CoP crossed over with line AD (Point C) was measured. The CPEI was then calculated by normalizing the distance BC to the width of the foot (distance of AD). Thus, CPEI = BC/AD × 100. (c) Schematic of GPs GP1–GP11.

2) Experiment for Validation Data Collection: We selected ten different scenarios, which are listed in Table I. We asked the participants to prepare the shoes they were accustomed to for the experiment. For example, sports-type shoes, sneakers, business-type shoes, and high-heeled pumps (heel height less than 3 cm) were allowed. We prepared three bag types with different weights: a rucksack with a 5 kg weight, a shoulder bag with a 2 kg weight, and a handbag with a 2 kg weight. To carry the shoulder bag and handbag, participants could decide which side they felt comfortable with. Participants in Group V were asked to walk in a 15-m straight path under different experimental conditions. They were also asked to walk three round trips indoors and were allowed to choose one, two, or three round trips outdoors to reduce any burden induced by hot or cold temperatures.

C. Devices and Systems for Experiment

The experimental setup is shown in Fig. 2. Two IMSs with the same circuit structure as A-RROWG were embedded in an insole placed under both feet arches near the calcaneus side to ensure that participants could walk naturally when measuring foot motion. The IMS also contained a six-axis

IMU (BMI 160, Bosch Sensortec, Reutlingen, Germany), an ARM Cortex-M4F microcontrol unit (MCU) (nRF52832, CPU: 64 MHz, RAM: 64 kB, ROM: 512 kB, Nordic Semiconductor, Trondheim, Norway), an EEPROM (S-24C32C, 32k bit, ABLIC, Tokyo, Japan), a real-time clock (RTC) (RX8130CE, EPSON, Suwa, Japan), and a 3-V lithium-coin battery (CR2430, 300 mAh) [Fig. 2(a)]. The MCU includes a Bluetooth low energy (BLE) module. The device is lightweight (10 g, including the coin battery) and small (29 × 40 × 7 mm) enough to be placed at the foot's arch. Our IMSs enable to output nine types of foot motion signals with the data sequence numbers of every measurement. The signals include directly measuring three axes of acceleration, i.e., linear motions A_x, A_y, and A_z, three axes of angular velocity, i.e., rotating motions G_x, G_y, and G_z, and three axes of foot-sole posture expressed by Euler angles E_x, E_y, and E_z, which were calculated using a Madgwick filter [44]. The measurement range of acceleration was ±16 G, and the angular velocity was ±2000°/s. We stopped the power-saving function at this feasibility study stage and recorded the detailed foot motion waveforms in real time by transferring them to a smartphone via Bluetooth universal asynchronous receiver/transmitter (UART) mode. Due to the packet loss induced by limitations of Bluetooth communication capacity and conditions, partial data loss occurred in the recorded foot motion waveforms. Please note that in the practical use case, the raw waveform will be only processed inside the edge device without transferring to the smart phone, and only calculated final results will be transferred to the smartphone. There should be no losing data inside the edge device.

The CPEI was measured using the foot pressure measurement system, Pedar[2] (novel GmbH, Munich, Germany, referred to as "Pedar" hereafter). Pedar consists of a pair of insole-shaped 2-D foot pressure sensor arrays, a data recording terminal, and batteries. The pressure sensors were laid on the IMS-equipped insole in shoes [Fig. 2(b)]; the data recording terminal and batteries were attached to a participant's waist [not shown in Fig. 2(b)]. The data were transferred to the terminal by a cable and recorded onto an SD memory card. Pedar recorded the distribution of foot pressure and directly output the center of pressure (CoP) trajectories, composed of *x*- and *y*-CoP coordinates in ".fgt" files, foot pressure data of every pixel of the sensor in ".asc" files, and every frame of the dynamic footprint in ".sol" files using the data processing software of Pedar. In the dynamic footprint image, the resolution of each pixel is 5 mm. To calculate the CPEI in accordance with the method in [18], we manually measured every participant's foot width and the trisection of the foot length from the toe from the dynamic footprint image [Fig. 2(b)]. Both IMSs and Pedar were operated at a sampling frequency of 100 Hz.

D. Signal Processing

We conducted all data processing and computations on MATLAB (Mathworks, Natick, MA, USA) in this study.

In each dependent walking trial, the CoP waveforms were divided into strides by detecting HS events using

[2]Trademarked.

a thresholding method, and then in each stride, partial wave-forms in the stance phase were extracted with the aid of the system-synchronized total foot pressure waveform [45]. The first and last strides of each trial were excluded first to ensure that all strides were derived from a uniform walking speed. Next, we temporally normalized the CoP waveform to apply a functional depth outlier detection method [46] to exclude extreme strides from all strides of one foot. After that, the CPEIs of the remaining strides were calculated in accordance with prior work [18]. Finally, the CPEIs of the remaining strides were averaged, respectively, in corresponding trials, and the average CPEI of one trial was treated as one measurement where each participant had a total of 18 measurements for both feet.

In every walking trial, the defective parts of foot motion waveforms were first linearly interpolated in accordance with the losing data sequence numbers. Interpolated data points were labeled. The acceleration values were then corrected to the global coordinates in each independent trial. After that, the waveforms were partitioned into individual strides by detecting HS events, and the stance and swing phases in each stride were divided by detecting TO events. The HS and TO detection on the IMS signal was based on the gait event detection algorithm demonstrated in our prior work, which was aided by A_y signals [47]. In addition, the following 11 GPs in one stride were used, where the first four concern foot eversion postures: maximum foot sole posture in the frontal plane (E_y) (GP1), average E_y during the stance phase (GP2), E_y at HS (GP3), E_y at TO (GP4), stride length (GP5), gait speed (GP6), maximum E_x in dorsiflexion and plantarflexion directions (GP7 and GP8), maximum circumduction (GP9), maximum foot height (GP10), and toe in/out angle in the transverse plane (GP11) [24] [Fig. 2(c)]. Note that GP5, GP9, and GP10 were normalized by each participant's height. We referred to the posture in which the foot sole completely touched the ground (around 21 percent gait cycle (%GC)–25%GC) as "foot-flat," a neutral posture. The average signals during the foot-flat of each stride were treated as offsets. The offsets were subtracted from all signals in each stride. Furthermore, we normalized the amplitude of acceleration and angular velocity signals in each stride by using the corresponding maximum instantaneous walking velocity during the swing phase, which was calculated by integrating A_y from a foot-flat to the end of the stride.

The first and last strides of each trial in that participants did not walk under a uniform speed were excluded. In the next step, we excluded strides whose proportion of interpolated data points in one stride was over 25%. Then, we temporally normalized the remaining strides using a fixed subdivision of the gait phase that varies the stance phase from 1%GC to 60%GC and the swing phase from 61%GC to 100%GC, which changed one stride into a 100×9 matrix. Next, as the step of an outlier stride exclusion, we also used a similar functional depth outlier detection [46] method to exclude extreme strides from each trial. GPs in the same excluded stride were eliminated as well. In the final step, the average normalized waveform and GPs of one trial were calculated from the remaining strides, and these data formed a $100 \times 9 \times N$ data cube, where N is

the number of dataset. For males and females, $N = 1132$ and 1087, respectively. Note that when all strides in one trial were excluded, the average of this trial is no longer recorded.

E. Feature Selection and Model Construction

In this study, we conducted SPM to analyze the correlation between foot function and foot motion in detailed gait phases. If significant %GCs existed, the foot motion signals in these %GCs could be used to create predictors for foot function assessment. Fig. 3 shows the process of feature selection and model construction.

During the SPM analysis, it was implemented in a hierarchical manner, similar to the analysis of variance (ANOVA) with a post hoc t-test. In our case, one stride of foot motion was first treated as a time-series multidimensional vector field from 1%GC to 100%GC, and the vector in which the %GCs significantly correlated with foot function was analyzed statistically. Specific vector components (scalar field), e.g., A_x and G_z, in those %GCs with statistical significances that reached the vector field level should then be tested post hoc.

As shown in Fig. 3(b), because the CPEI is an analog quantity, a canonical correlation analysis (CCA) with SPM (SPM-CCA) was applied [34]. For each sex, the data cube and CPEI were the paired input into the CCA. The %GCs whose test statistic of the CCA exceeded a critical test statistic threshold calculated in accordance with the random field theory (RFT) [48] were determined as significant %GCs. The level of significance was set as $p < 0.05$. As a post hoc test, only data in significant %GCs were further investigated by Pearson's correlation (PC) analysis with SPM (SPM-PC) in each component of the signal vector. For each component, similarly, the %GCs whose test statistic of the PC exceeded an RFT-based critical test statistic threshold were judged as the final CPEI-correlated significant %GCs in each component. Because there were nine components in the vector, we conducted Šidák correction [49] on a level of significance in which $p_c < 0.0057$.

After the SPM analysis, partial CPEI-correlated significant %GCs continuously appeared, acting as clusters, which we called gait phase clusters (GPCs). Incidentally, because foot motion is temporally successive, we considered that the integral average value could represent foot motion during those GPCs. We could use every signal in the GPC as a predictor variable. However, as previously mentioned, to make sure the assessment model is sufficiently light, we compressed the data by treating each GPC as a whole, where the integral average of the signal amplitude in the GPC, i.e., the average motion intensity of each GPC, was output as a single predictor, as expressed by the following equation:

$$C_i = \begin{cases} \displaystyle\sum_{T_s}^{T_e - \Delta T} \left(\frac{W(T) + W(T + \Delta T)}{2(T_e - T_s)} \Delta T \right), & T_e - T_s > 0 \\ W(T_e), & T_e - T_s = 0 \end{cases}$$

(1)

where C_i means the ith IMS predictor, T_s and T_e mean the start and end of %GCs of GPCs, respectively, and W means
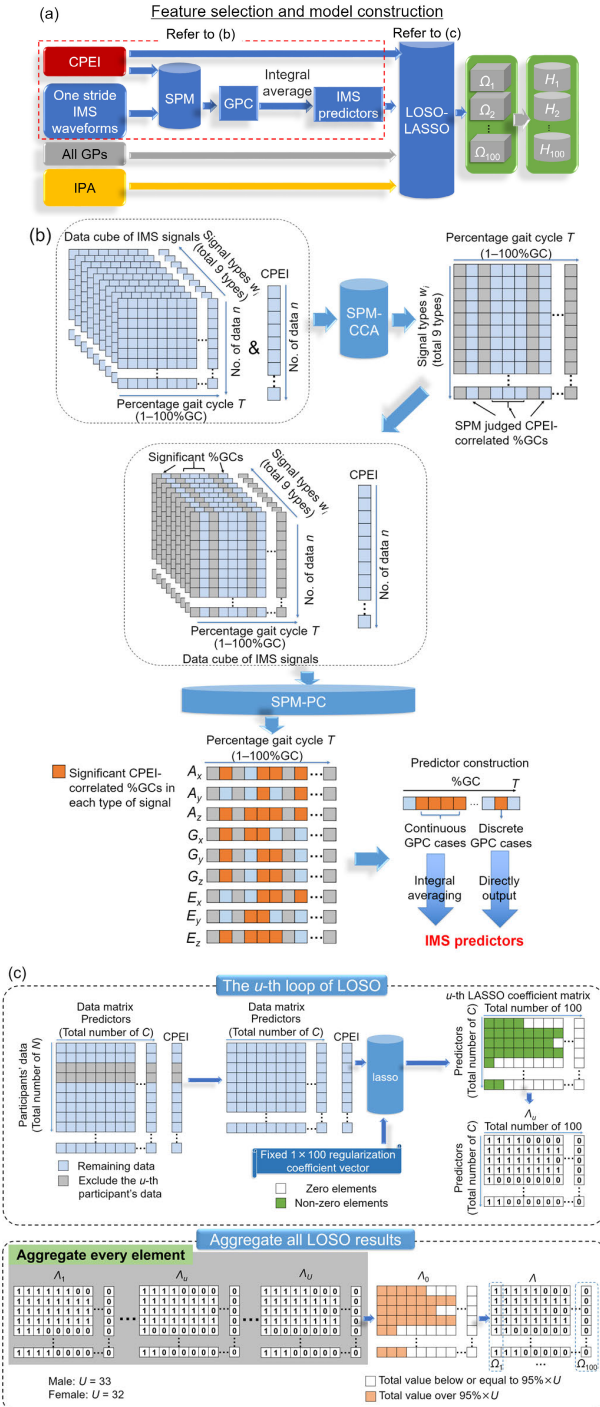
Fig. 3. (a) Flow of feature selection and model construction; SPM: statistical parametric mapping; GPC: gait phase cluster; LOSO-LASSO: leave-one-subject-out least absolute shrinkage and selection operator; LOSOCV: leave-one-subject-out cross-validation; $\Omega_1 - \Omega_{100}$: 100 types of features combinations according to different regularization coefficient sets in LASSO; and $H_1 - H_{100}$: 100 types of candidate multivariate regression model. (b) Details of SPM analysis and feature construction; CCA: canonical correlation analysis and PC: Pearson's correlation. (c) Details of LOSO-LASSO; U: total number of participant in training data; Λ_u: uth label matrix by substituting nonzero elements in LASSO coefficient by 1; Λ_0: label counter matrix; and Λ: final label matrix by substituting elements over and below $95\% \times U$ by 1 and 0 in Λ_0.

the waveform of the foot motion signal where $W \in \{A_x, A_y, A_z, G_x, G_y, G_z, E_x, E_y, E_z\}$. In the other case where CPEI-correlated significant %GCs appeared discretely, the signal

amplitude in these %GCs was directly guided as predictors [see Fig. 3(b)]. Here, we called these IMS predictors.

Similarly, we preferentially used a linear regression method for prediction to lighten the model. Accompanying the IMS predictors, GPs and individual physical attributes (IPA), including age, height, weight, and BMI, were also conducted as candidate predictors. In the LOSO-LASSO processing, as shown in Fig. 3(c), in each LOSO process, the data of the uth participant were first excluded, then the remaining data were subjected to LASSO analysis. After completing the LOSO process, we can obtain total U LASSO analysis results and statistically analyze these results to approximate the nature of the population estimator. Here, U means the total number of participants.

We used the "lasso" function in MATLAB for LASSO computation. After inputting a fixed 1×100 regularization coefficient vector as one of the function outputs, we obtained a LASSO coefficient matrix with a size of $C \times 100$, where C is the total number of candidate predictors. The regularization coefficient vector can be preliminarily obtained by inputting all datasets into the "lasso" function. Here, the elements in the regularization coefficient vector were geometric sequences, where the largest coefficient made the last column of the LASSO coefficient matrix all zeros. The vector elements were determined by performing the "lasso" function on partial training data in a preliminary stage. By looping the LOSO process, we obtained U label matrices, Λ_us, which were formed by substituting nonzero elements in LASSO coefficient matrices with 1. Here, Λ_u means the label matrix obtained in the uth loop. By summing all Λ_u, a matrix with a total counter Λ_0 was obtained, where the elements over $95\% \times U$ (males case: 32 and females case: 31) in this matrix were substituted with 1 while the remaining were substituted with 0, forming a final label matrix Λ. The 100 column vectors ($\Omega_1 - \Omega_{100}$) mean the 100 LASSO-recommended predictor combinations for different regularization coefficients selected, where 1 and 0 means the predictor was and was not selected, respectively. Combinations $\Omega_1 - \Omega_{100}$ were then conducted to construct 100 candidate models $H_1 - H_{100}$. As the regularization coefficient increases, the number of selected predictors becomes fewer. In particular, when the elements of the regularization coefficient vector were ordered from smallest to largest, the elements in Ω_1 were all ones, while those in Ω_{100} were all zeros. Consequently, the next task should be to determine the optimal assessment model for the candidates.

F. Model Evaluation and Validation

All candidate models $H_1 - H_{100}$ were evaluated via leave-one-subject-out cross-validation (LOSOCV), and Akaike's information criterion (AIC) [50] was applied as the evaluation index. Each H_i had an AIC value, and that with the lowest AIC value was chosen as the optimal one, denoted as M_o [Fig. 4(a)]. In the LOSOCV of M_o, Case 2 intraclass correlation coefficient, ICC(2,1), and ICC(2, k) [51] were used to evaluate the degree of absolute agreement between the true and estimated CPEI in a single rater case, k. Additionally, the mean absolute error (MAE) and coefficient of correlation (r) of the true versus estimated value were also used as the evaluation
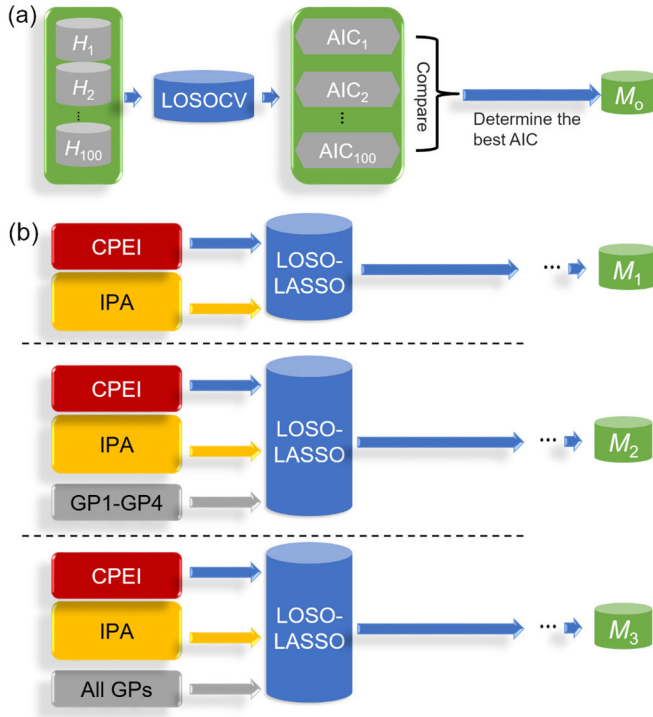
Fig. 4. (a) Schematic of determining optimal assessment model. AIC_1–AIC_{100}: Akaike's information criterions of H_1–H_{100}; M_o: optimal model; and LOSOCV: leave-one-subject-out cross-validation. (b) Other three models derived by optimizing three other predictor combinations by the same process as M_o, M_1: IPA only; M_2: M_1 plus GP1–GP4; and M_3: M_2 plus GP5–GP11.

indices of LOSOCV. The coefficient of determination (R^2) of the multivariate regression models using all training data (not LOSOCV) was also evaluated. The guidelines for interpreting the ICC interrater agreement were excellent (>0.749), good (0.600–0.749), fair (0.400–0.599), and poor (<0.400) [52], those for interpreting r were none (<0.100), small (0.100–0.299), medium (0.300–0.499), and large (>0.499), and those for interpreting R^2 were none (<0.020), small (0.020–0.129), medium (0.130–0.259), and large (>0.259) [53]. For comparison, models derived by optimizing three other patterns of predictor combinations in the same process as M_o, M_1: IPA only, M_2: M_1 plus GP1–GP4, and M_3: M_2 plus GP5–GP11 were also constructed and evaluated [Fig. 4(b)].

Bland–Altman (BA) plots [54], [55] were applied as the tool to evaluate the limit of agreement (LoA) between the Pedar method and the proposed method. Both the sample-based LoA and the confidence limits of LoA in the population were calculated. If the differences and averages between the two methods followed a normal distribution, which was preliminarily tested by a Kolmogorov–Smirnov test, we applied a t-test and PC test to examine whether there was a fixed and proportional bias. The LoA of the 95% confidence interval was determined from the perfect agreement (PA) line $\pm 1.96\times$ standard deviation (SD) [upper and lower LoAs (ULoA and LLoA)]. Furthermore, 95% confidence limits of LoA, i.e., upper and lower limits of ULoA (UULoA and LULoA), and upper and lower limits of LLoA (ULLoA and LLLoA) were also determined.

In the model validation stage, we evaluate the validity by the ratio of validation data in Group V, whose BA plots were within the agreement range, i.e., the success rate of measurements, denoted as Q_A. The optimistic agreement range, i.e., the range between UULoA and LLLoA, was applied here. Because the validation data size was limited, we utilized a probability-distribution-based method to estimate Q_A to eliminate the randomness.

We hypothesized that the residual of BA plots of training and validation data to the PA line, denoted as R_A and R_V, followed the normal distribution, $R_A \sim N(\mu_A, \sigma_A^2)$ and $R_V \sim N(\mu_V, \sigma_V^2)$. Here, μs and σs mean the averages and SD, respectively. Because the model was based on multivariate regression, theoretically, $\mu_A \equiv 0$. Furthermore, because of the limited data size, we calculated the 95% confidence levels of μ_V, σ_A, and σ_V and obtained their upper and lower limits, (μ_{VL}, μ_{VU}), $(\sigma_{AL}, \sigma_{AU})$, and $(\sigma_{VL}, \sigma_{VU})$, respectively. Hence, if we use an optimistic agreement range, the agreement range of the residual should be fixed as $-1.96\sigma_{AL}$ to $1.96\sigma_{AU}$. By then, Q_A should be in the area of $N(\mu_V, \sigma_V^2)$ inside the interval of $-1.96\sigma_{AL}$ to $1.96\sigma_{AU}$. Because μ_V and σ_V are independent of each other, the largest and smallest areas for $N(\mu_{Vi}, \sigma_{Vi}^2)$ subject to $\mu_{Vi} \in [\mu_{VL}, \mu_{VU}]$ and $\sigma_{Vi} \in [\sigma_{VL}, \sigma_{VU}]$ would be the upper and lower limit of Q_A, denoted as Q_{AU} and Q_{AL}, which can be expressed by the following equations:

$$Q_{AU} = \min\left(\max\left(\int_{-1.96\sigma_{AU}}^{1.96\sigma_{AU}} \frac{1}{\sqrt{2\pi}\sigma_{Vi}} \times \exp\left(-\frac{(x-\mu_{Vi})^2}{2\sigma_{Vi}^2}\right)dx\right)\Big/0.95, 1\right)$$
$$\text{s. t. } \mu_{Vi} \in [\mu_{VL}, \mu_{VU}], \quad \sigma_{Vi} \in [\sigma_{VL}, \sigma_{VU}] \quad (2)$$

$$Q_{AL} = \min\left(\min\left(\int_{-1.96\sigma_{AU}}^{1.96\sigma_{AU}} \frac{1}{\sqrt{2\pi}\sigma_{Vi}} \times \exp\left(-\frac{(x-\mu_{Vi})^2}{2\sigma_{Vi}^2}\right)dx\right)\Big/0.95, 1\right)$$
$$\text{s.t. } \mu_{Vi} \in [\mu_{VL}, \mu_{VU}], \quad \sigma_{Vi} \in [\sigma_{VL}, \sigma_{VU}]. \quad (3)$$

In the other statistical analyses, t-tests and ANOVA were used to compare the differences between two groups and among three or more. All the levels of importance were set as $p < 0.05$. In the post hoc of ANOVA, Šidák correction was also applied to correct the p-values.

IV. RESULTS

A. Participants' Characteristics and Pedar-Measured CPEIs

The participants' characteristics in Groups T and V are shown in Tables II and III, respectively. Moreover, the results of statistical analyses concerning CPEIs in various conditions are shown in Table IV. For all, the statistics of males and females were collected, respectively.

From Tables II and IV, the Pedar-measured CPEIs of males were significantly bigger than those of females ($p < 0.001$), which indicates females pronated their feet more in gait. For both males and females, CPEIs significantly varied in different

TABLE II
PARTICIPANTS' CHARACTERISTIC IN GROUP T

	Male Mean (SD)	Female Mean (SD)
Number	33	32
Age (year)	44.1 (14.9)	45.4 (16.2)
Height (cm)	170.8 (7.9)	158.8 (7.3)
Weight (kg)	71.8 (15.5)	52.9 (7.8)
BMI (kg/m^2)	24.4 (4.3)	21.3 (2.7)
Speed (m/s) (Data size: male, female)		
Slow (387, 363)	1.08 (0.22)	1.09 (0.22)
Normal (384, 361)	1.42 (0.17)	1.41 (0.17)
Fast (387, 363)	1.87 (0.24)	1.81 (0.18)
CPEI (Data size: male, female)		
Total (1132, 1087)	10.9 (9.3)	5.0 (5.2)
Slow (387, 363)	10.9 (7.4)	4.9 (4.5)
Normal (384, 361)	13.0 (9.6)	5.5 (5.0)
Fast (361, 363)	8.8 (10.4)	4.4 (5.8)

SD: standard deviation.

TABLE III
PARTICIPANTS' CHARACTERISTIC IN GROUP V

	Male Mean (SD)	Female Mean (SD)
Number	8	9
Age (year)	42.6 (6.6)	39.0 (9.0)
Height (cm)	172.4 (4.3)	160.2 (2.7)
Weight (kg)	68.8 (9.1)	55.3 (11.9)
BMI (kg/m^2)	23.2 (3.2)	21.5 (4.3)
Hand side		
Shoulder bag (Right side)	7	5
Hand bag (Right side)	5	6
Participants' shoes		
Sports type	2	1
Business type	3	0
High-heeled pumps	0	2
Sneaker or other flat shoes	3	6
CPEI (Data size: male, female)		
S1 (83, 106)	10.1 (8.4)	5.8 (4.4)
S2 (89, 108)	10.5 (9.0)	6.8 (3.8)
S3 (81, 108)	9.3 (9.0)	6.0 (5.5)
S4 (82, 108)	10.8 (9.7)	6.5 (4.9)
S5 (89, 102)	9.1 (5.4)	5.7 (5.7)
S6 (46, 55)	7.8 (5.8)	5.1 (5.2)
S7 (46, 51)	7.0 (5.4)	5.5 (5.1)
S8 (53, 58)	9.5 (6.2)	6.2 (6.8)
S9 (57, 55)	7.4 (4.8)	5.9 (6.9)
S10 (59, 68)	6.3 (5.4)	4.6 (7.0)
CPEI left-to-right foot difference		
With bag on single side (77, 108)	-3.7 (6.4)	-1.7 (5.5)
Without bag (35, 52)	-3.3 (4.6)	-1.0 (5.5)

SD: standard deviation.

TABLE IV
COMPARISON OF MEASURED CPEIs UNDER DIFFERENT CONDITIONS

	p-value	
	Male	Female
Among three gait speeds (Post hoc, $p_c < 0.017$)	**0.000**	**0.017**
Fast vs. normal	**0.000**	**0.004**
Normal vs. slow	**0.002**	0.095
Slow vs. fast	**0.001**	0.230
Total (Group T) vs. S1 (Group V)	0.452	0.117
Left-to-right difference: with vs. without bag (standard shoes, vinyl)	0.774	0.428
Among ten scenarios (Post hoc, $p_c < 0.005$)	**0.003**	0.308
S6 vs. S8 (different shoes, asphalt, no bag)	1.000	-
S5 vs. S6 (participants' shoes, vinyl vs. asphalt, no bag)	1.000	-
S6 vs. S7 (participants' shoes, asphalt, with vs without bag)	1.000	-
S5 vs. S8 (different shoes, vinyl vs. asphalt, no bag)	1.000	-
S1 vs. S7 (different shoes, vinyl vs. asphalt, with vs without bag)	0.088	-
S1 vs. S9 (standard shoes, vinyl vs. gravel, no bag)	0.147	-
S1 vs. S10 (standard shoes, vinyl vs. lawn, no bag)	0.012	-

p_c: Šidák corrected significance level;

pairs to introduce whether different walking conditions would affect CPEI values. ANOVA and its post hoc suggested that even walking under different assumed usual conditions, the CPEIs did not perform differently in females, and we did not find any significant differences between listed pairs for males in multiple comparisons. Although not shown in Table IV, significant differences were found in males between S4 and S10 (walking on vinyl with rucksack versus on lawn without bag), and S2 and S10 (walking on vinyl with handbag versus on lawn without bag).

In addition, we also examined the impact of a single body-side-weighted condition on left-to-right CPEI difference. As a result, for both sexes, there were no significant differences in left-to-right CPEI difference whether with a bag on a single side of the body or not, which suggested that limited weight on a single side of the body will not affect the characteristics of foot function.

B. SPM Analysis

Compared with the males and females, their average waveform appeared approximately similar. In contrast, the SD of waveforms, particularly in the frontal and horizontal plane (G_y, G_z, E_y, and E_z), appeared to have more different shapes (Fig. 5).

The results of SPM-CCA depicted that for both sexes, the foot motion signal vectors in all GCs significantly correlated with the CPEIs. The statistic SPM{t} curves of the post hoc SPM-PC analysis depicted the strength of the positive ($+$) and negative ($-$) correlation between CPEIs and each type of foot motion signal. The sections of curves that exceeded critical thresholds depicted the significant GC intervals that

gait speed categories according to ANOVA. The CPEIs at normal speed had the maximum mean values, while those at fast speed had the minimum ones. In addition, the CPEIs at the slow speed had the minimum SD, while those at the fast speed had the maximum. Specially, the CPEIs at the fast speed were significantly smaller than those at normal speed. Comparing the results between Groups T and V under the same conditions (total in Table II versus S1 in Table III), there were no significant differences in both sexes (Table IV), which suggested the stability of our measurements.

The mean and SD of CPEIs of all ten scenarios are listed in Table III, and the results of the comparison are shown in Table IV. In Table IV, we included several representative
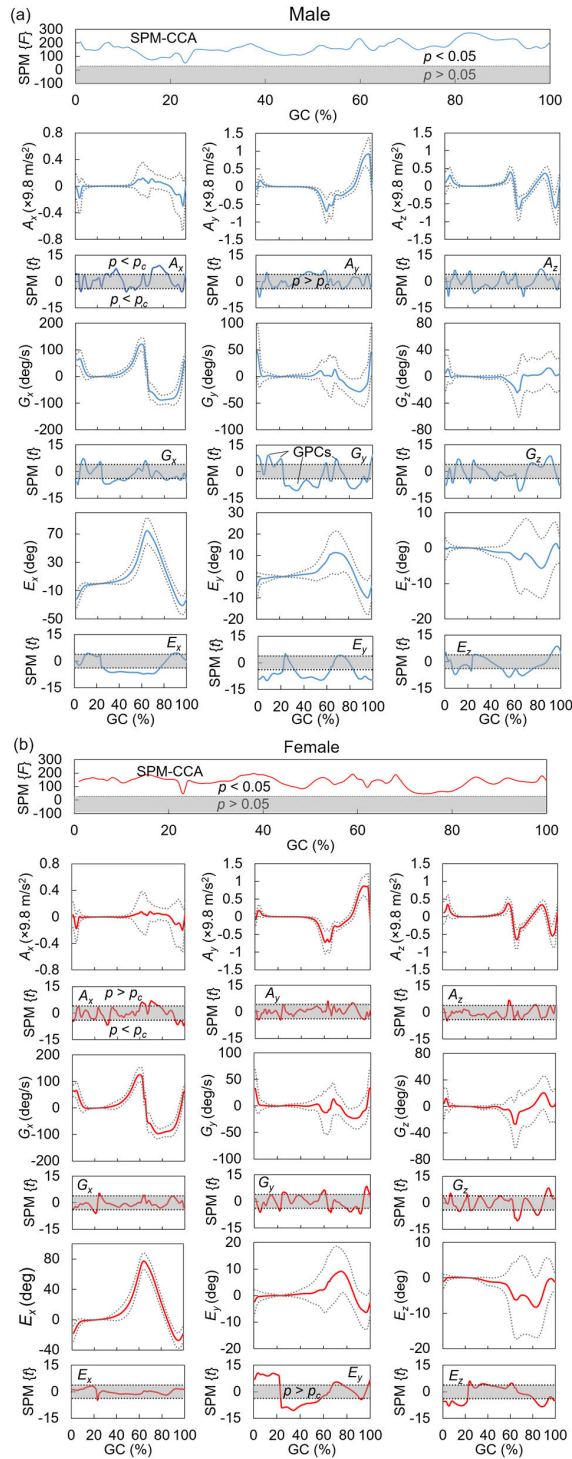
Fig. 5. Results of correlation analysis between foot motion and CPEI using SPM: (a) male and (b) female. GC: gait cycle; GPC: gait phase cluster; SPM{F}: F statistic of vector field analysis by SPM-CCA; and SPM{t}: statistic of post-hoc scalar trajectory linear correlation test by SPM-PC. Single and double dotted lines linked to SPM{F} and SPM{t} indicate the critical RFT threshold of F, and Šidák corrected critical RFT threshold of t. Double dotted lines linked to foot motion signals, e.g., A_x and G_z, showed the 95% confidence interval of the waveform. Statistic curves outside the gray zones in each signal type indicated that the intervals of GCs significantly correlated with CPEI, defined as GPCs.

correlated with the CPEIs, defined as GPCs here. The GPCs in acceleration signals appeared more fragmentary because the smoothness of the acceleration waveform was smaller than
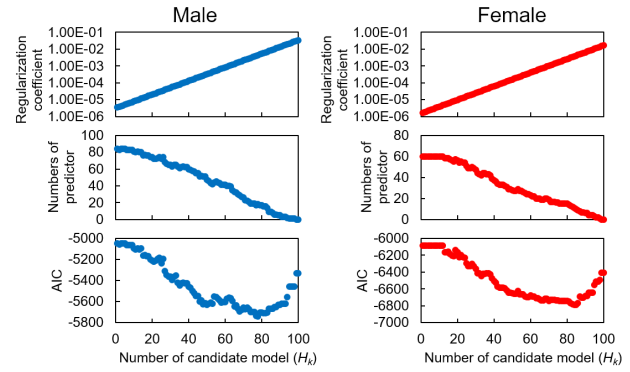
Fig. 6. Results of LOSO-LASSO analysis to determine optimal predictor combination, M_O.

TABLE V
SELECTED OPTIMAL PREDICTORS COMBINATION OF MALE

No	Detail	Mean (SD)	r	Coef.	p_m
Int.	Interception			-12.4	**0.000**
C_{m1}	Age	44.10 (14.90)	-0.178	-0.082	**0.000**
C_{m2}	BMI	24.36 (4.25)	0.181	0.507	**0.000**
C_{m3}	GP2	0.76 (0.71)	-0.367	5.39	**0.000**
C_{m4}	GP9	1.92 (0.69)	0.162	1.48	**0.000**
C_{m5}	A_x, 33 to 39	0.30E-3 (3.25E-3)	0.204	363	**0.000**
C_{m6}	A_x, 45 to 48	0.03E-3 (8.01E-3)	-0.162	-61.6	0.082
C_{m7}	A_x, 54 to 55	5.61E-2 (4.78E-2)	-0.137	-10.9	0.123
C_{m8}	A_x, 69 to 84	0.58E-1 (4.16E-1)	0.282	5.95	0.499
C_{m9}	A_y, 2 to 4	1.39E-1 (0.67E-1)	-0.263	-24.6	**0.000**
C_{m10}	A_y, 9	3.11E-2 (2.11E-2)	0.148	60.9	**0.000**
C_{m11}	A_z, 49 to 51	8.77E-2 (2.12E-2)	0.162	56.2	**0.000**
C_{m12}	A_z, 67 to 69	-3.61E-1 (0.58E-1)	-0.231	-13.7	**0.002**
C_{m13}	G_y, 60	1.60 (13.73)	0.150	0.120	**0.000**
C_{m14}	G_y, 68 to 71	6.28 (9.90)	0.201	0.175	**0.000**
C_{m15}	G_z, 10 to 14	0.54 (0.70)	0.215	0.615	0.061
C_{m16}	E_y, 32 to 63	2.60 (1.63)	-0.254	3.51	**0.000**
C_{m17}	E_y, 84 to 100	-5.47 (2.82)	-0.297	-1.21	**0.000**
C_{m18}	E_z, 92-100	0.43 (2.74)	0.234	0.591	**0.000**

GP2: Average foot sole posture in the frontal plane during stance phase in degrees; GP9: Height-normalized maximum circumduction in cm; r: Linear correlation coefficient of the predictor with CPEI; Coef.:Coefficient of multivariate regression model using all participants' data; p_m: p-value of coefficient of the multivariate regression model, the significance level is p_m<0.05. C_{m5} to C_{m18}: IMS predictors, signal type, and its interval range of GPCs were depicted in the "Detail" column. The interval range is in %GC. The units of IMS predictors were the same as the signals. C_{m5} to C_{m18} was normalized by maximum instantaneous speed.

that of the angular velocities and sole-to-ground Euler angles. The shape of the statistic SPM{t} curves and the location of GPCs between males and females were different, which may indicate that foot function affects males and females differently (Fig. 5). As a result, 69 and 45 GPCs, i.e., IMS predictors, were obtained for males and females, respectively.

C. Feature Selection

To obtain the final optimal predictor combination M_o, including IPA and GP predictors, for males and females, a total of 84 and 60 candidate predictors were input to LOSO-LASSO, respectively. By finding the lowest AIC, we determined M_o for males and females, with 18 and 10 predictors finally selected, respectively (Fig. 6). The final predictors and their correlation analyses with the CPEI are listed in Tables V and VI.

Regarding the IPA and GP predictors, the height-normalized circumduction (GP9) was selected for both sexes

TABLE VI
SELECTED OPTIMAL PREDICTORS COMBINATION OF FEMALE

No	Detail	Mean (SD)	r	Coef.	p_m
Int.	Interception			-7.49	**0.000**
C_{f1}	Weight	53.47 (7.88)	0.220	0.157	**0.000**
C_{f2}	GP9	1.89 (0.71)	0.271	1.69	**0.000**
C_{f3}	A_z, 58 to 60	2.75E-1 (0.86E-1)	0.193	6.00	**0.000**
C_{f4}	G_y, 58 to 61	-10.59 (17.62)	0.178	0.043	**0.000**
C_{f5}	G_y, 64 to 66	-5.03 (19.53)	-0.167	0.053	**0.000**
C_{f6}	G_y, 95 to 99	6.60 (10.06)	0.287	0.139	**0.000**
C_{f7}	G_z, 81 to 86	8.90 (11.61)	-0.194	0.069	**0.000**
C_{f8}	E_y, 1 to 22	-0.26 (0.56)	0.320	1.06	**0.000**
C_{f9}	E_y, 24 to 57	0.67 (1.01)	-0.220	-0.950	**0.000**
C_{f10}	E_z, 1 to 26	-0.01 (0.13)	0.248	4.29	**0.000**

C_{f3} to C_{f10}: IMS predictors, signal type, and its interval range of GPCs were depicted in the "Detail" column. The interval range is in %GC. C_{f3} to C_{f7} was normalized by maximum instantaneous speed.
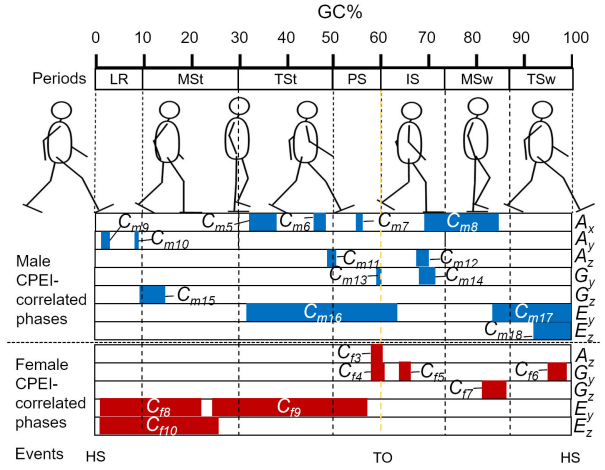

Fig. 7. Distributions of GPCs for IMS predictor construction. GC%: percentage gait cycle; LR: loading response; MSt: mid-stance; TSt: terminal stance; PS: preswing; IS: initial swing; MSw: mid-swing; TSw: terminal swing; HS: heel strike; and TO: toe-off.

($r = 0.162$ and 0.271, both with a small effect size), while age, BMI, and average foot sole posture in the frontal plane (GP2) were additionally selected for males and only weight was additionally selected for females. GP2 had the highest correlation with the CPEI in males' predictors ($r = -0.367$, medium effect size). The results suggested that BMI and weight positively correlated with the CPEI, meaning that participants with higher body mass or weight had less over-pronated foot function.

The signal type origins and IMS predictor gait phases can be referenced from the "Detail" column. Spatially, foot motions in the frontal (xz) and horizontal (xy) planes were suggested to be essential for foot function assessment (C_{m5}–C_{m8}, C_{m11}–C_{m18}, C_{f3}–C_{f10}, and GP10) for both males and females. Temporally, these GPCs interspersed over many gait periods were seen in both sexes, whereas the distribution of GPCs was different. In particular, we, respectively, observed that partial predictors originating from the accelerations and angular velocities signals were distributed before preswing only in males (C_{m5}, C_{m6}, C_{m9}, C_{m10}, C_{m11}, and C_{m15}) while concentrated in the terminal stance phase. In contrast, partial predictors originating from the foot-sole posture were distributed only in the early stance phase in females (C_{f8} and C_{f10}) (Fig. 7).

By referencing the mean value and linear correlation coefficient of the predictor with the CPEI, the direction of foot motions during these phases and its changing trend with the CPEI increasing (supination trend) can be determined. For males, compared with overpronated feet, normal feet mainly had less anterior deceleration (C_{m9}, positive value) at early loading responses, a higher internal rotation velocity (C_{m15}) during the early mid-stance, and more medial acceleration and fewer eversion postures during the terminal stance (C_{m5} and C_{m16}). For females, C_{f8}, which has the biggest correlation coefficient ($r = 0.320$, medium effect size), originated from E_y during 1%GC–22%GC. Because we set foot-flat as a neutral position, its negative mean value and positive correlation coefficient showed that the foot with a bigger range of motion in the inverse direction (the difference from E_y at 21%GC–25%GC to E_y at HS is negative) had a bigger CPEI. Compared with overpronation, normal feet also mainly have more internal rotation postures during the early

stance phases (C_{f10}) and have fewer eversion postures during the late stance phases (C_{f9}).

We additionally found a number of significant predictors in foot motion during the swing phase. Our results suggested that normal participants tend to have larger eversion rotation velocities (females, C_{f5}) during the initial swing, small internal (or big external) rotation velocities (females, C_{f7}), or larger lateral decelerations during the mid-swing (males, positive value, C_{m8}), larger eversion velocities (females, C_{f6}), larger inversions (males, negative value, C_{m17}), and external rotation (males, C_{m18}) postures before HS.

Furthermore, we also listed the coefficients of predictors and their p-values in a multivariate regression model in the tables. Although the linear correlation coefficient with the CPEI of all predictors only had a small to medium effect size, the constructed models of both males and females achieved a large effect size ($R^2 = 0.469$, $p < 0.001$ and $R^2 = 0.407$, $p < 0.001$, respectively).

D. Model Evaluation

Fig. 8 shows the following results. Comparing M_o with the results of M_1 and M_2 in LOSOCV, constructing multivariate models using only IPA or a number of GPs of foot sole posture in the frontal plane cannot estimate the CPEI reasonably when their effect sizes of R^2 were only at the medium level and their ICC(2,1)s suggested poor agreement between the true and estimated values. Including the seven other GPs for optimization (M_3) improved the results with larger effect sizes of R^2. However, ICC(2,1)s still suggested poor agreement (for more information about M_1, M_2, and M_3, please see Supplementary Materials). In contrast, the model using the optimal predictor combination (M_o) achieved an MAE of 6.05 with an ICC(2,1) of 0.521, ICC(2, k) of 0.685, and r of 0.560 for males and a MAE of 3.30 with an ICC(2,1) of 0.494, ICC(2, k) of 0.662, and r of 0.545 for female. This means that for both sexes, the estimated value had a correlation with the large effect size with the true value, as well as had a fair ICC agreement with the true values when we used IMS
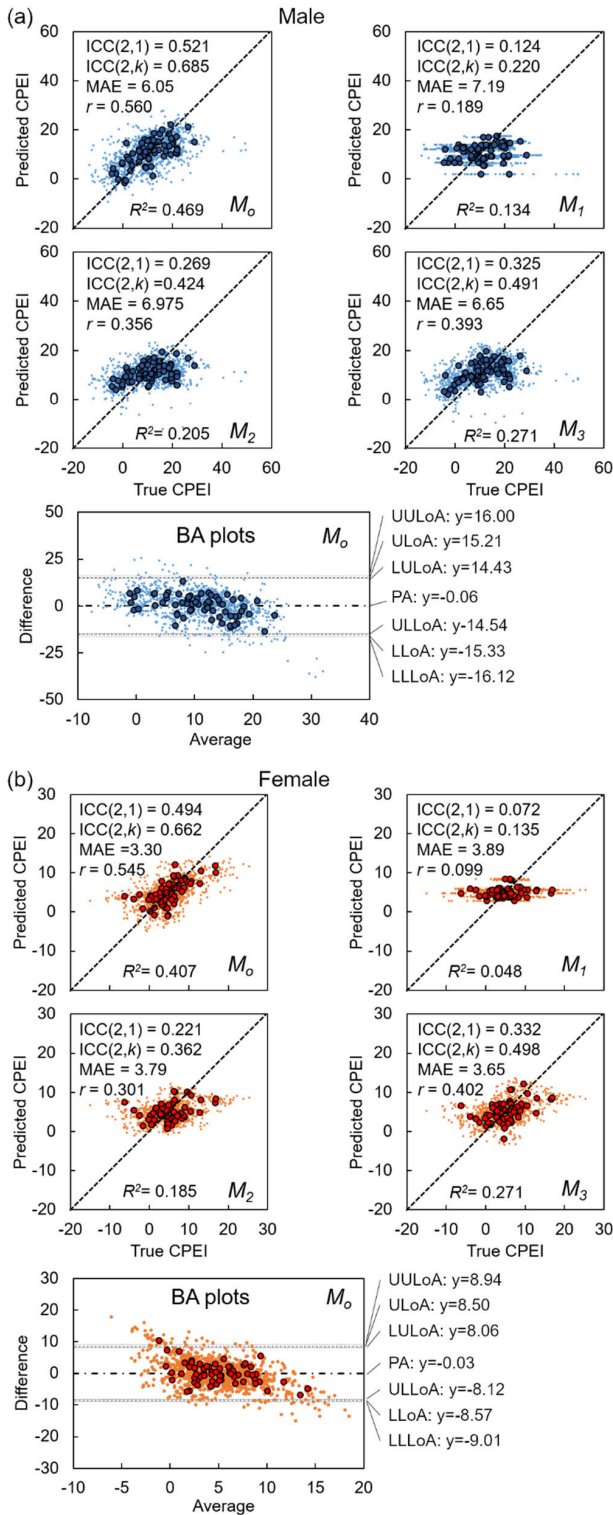
Fig. 8. Evaluation results of M_o, M_1, M_2, and M_3 using LOSOCV and BA plots of M_o for (a) male subjects and (b) female subjects.

proportional biases ($p < 0.001$, < 0.001) between the two raters. Note that in Fig. 8, the lines of LoAs were very near to their confidence limit lines because of the relatively big data size.

E. Model Validation

According to the BA plots shown in Fig. 9, the model successfully assessed a maximum of 99.7% and 100.0% of the data of males and females, respectively, in Group V under the same experimental conditions of the training data (S1), which suggested that the models had high robustness to different participants. In the indoor environment, no matter how the participants held the baggage, the models well-estimated the validation data in that there were no apparent negative changes in the Q_A of measurements (S2–S4). In the outdoor environment (S6–S10), the upper limit of the Q_A of males and females maintained high levels (males: 99.0%–100.0%; females: 90.8%–97.8%) despite walking on different ground surfaces, even in the wild paths (S6–S7 and S9–S10).

We observed that the Q_As of females decreased in the scenarios right after shoe shifting (S4–S5 and S7–S8) although it recovered in the following scenario, while this phenomenon was not observed in males.

V. DISCUSSION

A. Applying the Model to an IMS

In this study, we developed a simple and effective model for CPEI estimation from straight walking foot motion with high precision using the MAE and large effect sized R^2 via multivariate regression. According to the ICC values for both sexes, the model achieved a "fair" agreement when considering using IMS as a single rater and achieved a "good" agreement when considering using both IMS and Pedar as raters in the same CPEI estimation. Referencing the MAE values, the model precision of males and females enabled discrimination of $0.65\times$ and $0.63\times$ SD of the training data. This precision satisfies the requirement for overpronation/normal/oversupination classification when following the proposal to categorize the foot function using quintiles [18].

In the validation, the model successfully estimated 92.8% to 100.0% and 85.8% to 100.0% (both are the ranges of the upper limit of Q_A in S1–S8) data in the current precision, respectively. These data were collected by using different shoes, walking indoors or outdoors, and with or without baggage. This suggests that the proposed model has a high potential for daily living measurement. This model only involves a few steps of simple addition and multiplication, which can be afforded by the calculation capacity and easily integrated into the conventional algorithm of our A-RROWG.

A-RROWG automatically detects walking bouts, i.e., continuous walking motions, and excludes motions deviating from common walking, e.g., turning, walking backward, running, and jumping. A model constructed from straight walking foot motions should be sufficient for this type of sensor. Next, we will discuss how the model was used in A-RROWG during daily living measurement.

as a single rater, and had a good ICC agreement with the true values when using IMS and other raters, e.g., Pedar to assess the foot function spontaneously. These results suggested that foot function can only be estimated well by applying IMS predictors.

According to the BA plots, the average of true and estimated values linearly correlated with their differences, i.e., there were
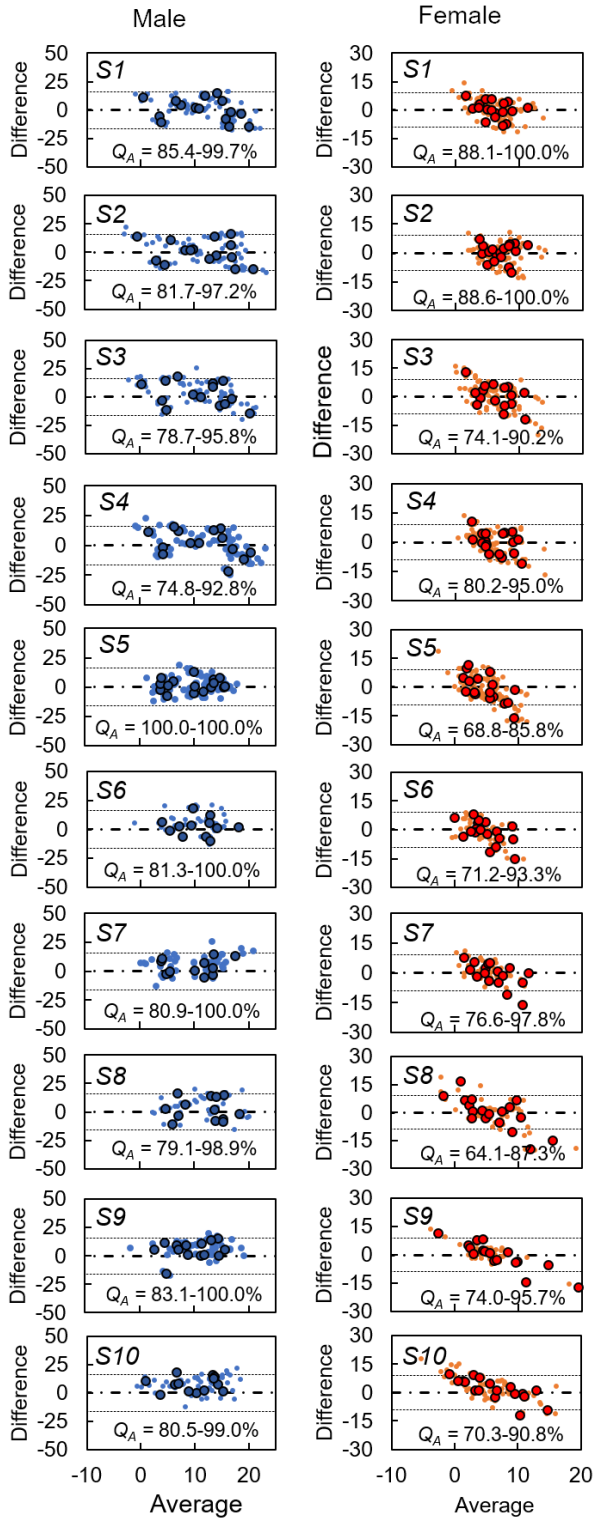
Fig. 9. Results of validation data in ten scenarios of male and female. The dotted line depicts the linearly corrected optimistic agreement interval. Small pale dots depict data of every measurement, and big dark dots depict average data of all walking trials of every single foot. Q_A: success rate of measurements, whose lower and upper limits are shown.

Fig. 10. (a) Power management operation sequences in an A-RROWG-type IMS. (b) Process of evaluating one stride. Newly integrated processes concerning CPEI estimation are shown in the red dashed frame. (c) Example of time-series data recorded in buffer of IMS. j_v^i: index of the ith valley in E_x, j_p^i: index of the ith peak in E_x, j_{HS}^i: index of the ith signal feature representing HS in A_y; and j_{TO}^i: index of the ith signal feature representing TO in A_y.

On the basis of our previously developed power management algorithm [24], to extend battery life, there are three stages of operation sequences for one measurement cycle of the IMS. In Stage 1, the IMS was set to sleep mode, and measurements were only performed within a predetermined period, e.g., from 9 to 11 A.M. During that time, the IMS moved to Stage 2, in which it was woken up by the RTC and entered a suspended mode waiting for the arrival of walking signals. Otherwise, it would return to Stage 1. If walking signals were detected and repetitive walking was confirmed, the IMS moved to Stage 3, which ran at full sampling rate to record walking signals in the IMS buffer. Otherwise, it would also return to Stage 1. After a certain number of strides were analyzed, the IMS moved back to Stage 1 and waited for the next predetermined measurement time [Fig. 10(a)].

Previously, at Stage 3, the IMS could calculate GP5–GP11 as well as GP1 of the ith stride by running an online foot-sole-angle-based stride-segmentation algorithm [24]. In this algorithm, the indices of the first and second valley (j_v^i and j_v^{i+1}) and peak (j_p^i and j_p^{i+1}) of E_x in the buffer were detected and then a part of the signals between two middle points around the foot-flat points, $1/2(j_v^i + j_p^i)$ and $1/2(j_v^{i+1} + j_p^{i+1})$, was segmented for parameter calculation. To calculate GP2–GP4 and construct IMS predictors, we integrated a previously developed online gait event detection algorithm by detecting the specific signal features from A_y [39] into the IMS operation sequences. In this algorithm, the IMS detected the indices of the first and second HSs (j_{HS}^i and j_{HS}^{i+1}) between

j_v^i and $1/2(j_v^i + j_p^i)$ and between j_v^{i+1} and $1/2(j_v^{i+1} + j_p^{i+1})$, and then detected the TO in this stride between $1/2(j_v^i + j_p^i)$ and j_p^i [Fig. 10(b)]. Note that the unit for all start and end points of GPCs for IMS predictor construction shown in Tables V and VI was depicted as %GC, which was derived from temporally normalizing the waveforms of one stride. Therefore, in this case, we must convert the %GC number to the distance from j_{HS}^i on the real-time axis to obtain the indices of the start and end points of GPCs in the buffer to construct IMS predictors.

By inputting the mandatory predictors into the constructed model installed in the IMS, the CPEI was finally obtained in the IMS. Note that predictors not inputted into the model can also be utilized for other purposes [Fig. 10(b) and (c)]. We considered our IMS can be applied to long-term daily foot healthcare applications, as well as be feasible for one-shot measurement for ambulatory care and foot measuring by shoe fitters, as long as we preliminary tune the frequency of the measurement cycle in one day, e.g., from three times a day to once a minute.

B. CPEI Obtained via Pedar

According to a previous study by Menz et al. [18] using the Framingham Foot Study Cohort involving big sample sizes of participants (males: 1477 and females: 1901), males have a different distribution and a higher average CPEI than females. The results in this study agreed with their findings, which also means the measurements of the CPEI in our studies were effective. However, the average CPEI of our participants was lower than that of their cohorts in both sexes. According to the findings in [56], [57], and [58], the reason might be that Asians' feet tend to have lower CPEIs, i.e., more pronated feet than other races because they have a wider forefoot width and lower longitudinal arch. This indicates that although the CPEI can reflect the degree of overpronation or oversupination, the cutoff value for categorizing foot function needs to be observantly determined in line with sex and race in the applications.

We observed that the CPEI levels and variations changed with the self-controlled speed (Table II). In particular, when participants were asked to walk fast, the CPEI was significantly smaller than walking at a comfortable speed in both sexes, which agreed with the previous findings that a higher walking speed increases foot eversion and shifting the loading on the forefoot from the fifth to the first metatarsophalangeal joint side [59], [60]. However, we also observed that the average CPEI in the self-controlled slow speed was lower than that in the comfortable speed but higher than that in the fast speed. There are few studies concerning the self-controlled slower speed on the CPEI to the best of our knowledge. A possible reason is that instructed slow walking may have affected the nature of participants' gaits, which needs further investigation. The phenomenon mentioned above might suggest that the walking speed should be standardized or the CPEI should be corrected by the walking speed to assess foot function in accordance with the CPEI.

People always carry baggage with them in daily living. We were interested in whether having a load on the body, especially on the unilateral side, would affect the CPEI estimation in daily living because, under these states, the gait symmetry would be changed by the load [61]. This study found that having a load on the body, even on the unilateral side (rucksack: 5 kg and shoulder/handbag: 2 kg), did not significantly change the CPEI on each foot nor the difference of the CPEI between two feet. Furthermore, we observed that regardless of baggage, Q_A did not change. However, whether heavier loads change the state of the CPEI needs to be further studied.

Unlike conventional studies that use a floor-mat-type plantar pressure sensor to measure the CPEI from barefoot [18], [57], we used an insole-shaped in-shoe-type sensor in this study. This plantar pressure sensor helped us achieve the spontaneous measurement of foot motion and CPEI, as well as outdoor measurements. However, because the sensors were inserted into shoes, sensors were first separated from the ground by the sole of the shoes, which was like laying on a cushion. The sensor insole may become curved or distorted along with the distortion of the shoes. Especially for high-heeled pumps, the foot in the shoes does not always step in the same place during walking. We must tolerate the impact on measurement precision induced by these limitations using the in-shoe type plantar sensor. In this study, we hypothesized that the CPEI reflected the reality of foot function, and the foot function was expressed via walking. Based on this hypothesis, regardless of the walking condition, the CPEI should be a fixed measurement value (here, the average of each walking trail) representing a specific foot function. Also, the variation induced by the walking speed, shoe type, or ground surface was considered a random variance according to the one-way random effect model of Case 1 [51]. However, CPEIs changed significantly when walking fast. As a result, the test-retest reliability of true values, ICC(1,1) of 18 CPEI measurements, was limited to 0.697 ("good" level), which was the other factor limiting model precision.

C. Correlation Between Predictors and CPEI

In this study, the final selected predictors for the optimal model can follow previous findings. Scott et al. [62] and Menz [63] indicated that older people tended to have more pronated feet, which was also observed in our female participants.

As IMSs were located under the foot arch near the calcaneal side, we assumed the mid-foot and rearfoot to be connected as a rigid body. Most of the IMS predictors were found in the frontal and horizontal plane, which was a good result since foot functions were depicted in the frontal plane as well as in the mediolateral (x) direction. The foot sole Euler angle, i.e., foot posture in the frontal plane during the stance phase ($C_{m}16$, $C_{f}8$, and $C_{f}9$), was a significant predictor for both males and females. Although the foot-sole angle was not biomechanically equal to the subtalar joint angle because the tibia tilt in the frontal plane to the ground slightly changed

during the stance phase [30], we considered that these predictors indeed partially reflected the overall status of participants' talocrural and subtalar joints. We also observed sex differences in the stance phase predictors (C_{m5}, C_{m6}, C_{m7}, C_{m9} to C_{m11}, C_{m13}, C_{m15}, C_{m16}, C_{f3}, C_{f4}, C_{f8} to C_{f10}). In particular, females demonstrated that the CPEI correlated more with the foot posture (angle), while males correlated more with the motion (acceleration and angular velocity). Males and females have different foot anthropometric characteristics [64]. Previous studies [65], [66] indicated there are sex-based variations in CoP trajectories, and talocrural and subtalar joint kinematics in the frontal and horizontal plane during early stance phases have different impact attenuations and adaptations to surface strategies since females tend to have more ligamentous laxity than males [67]. The interosseous talocalcaneal ligament can be considered the central pivot of rotatory stability as well as mediolateral stability [68]. The higher ligamentous laxity of females could be a reason that the impact of weight at HS makes foot deformities much easier to be expressed during the stance phase. In contrast, for males, the overpronation or oversupination deformity was finally expressed only under multiple actions of weight impact and natural pronation motion during the terminal stance and preswing.

We also found several significant predictors in foot motion during the swing phase. Structurally, people swing their lower limbs laterally during the initial swing, then to their maximum position during mid-swing, and pull back medially for landing during the terminal swing [69]. Although many previous studies paid more attention on directly linking the foot function to the kinematics during the stance phase, a previous study on flat foot [70], which was associated with overpronated feet [18], indicated that people with severer flat feet tend to have more supination motion during the initial swing and small external rotation motion during mid-swing, which can be connected with C_{f5}, C_{f7}, and C_{m8}. Moreover, during the terminal swing, C_{m18} agreed that normal participants had more external rotation motion. From the study of Levinger et al. [30], immediately before HS, we can also observe that regardless of whether the participant had flat feet, the tibia tilt to the ground had the same level, whereas flat feet had less inversion hindfoot as the tibia angle and a smaller slope (eversion angular velocity) in eversion, which means flat feet also had less inversion foot posture and less eversion speed. This can connect well with C_{f7} and C_{m17}.

Foot function is highly correlated with the alignment of the subtalar joint, which is the main weight-bearing surface and is mostly found in a valgus position under weight-bearing conditions [71]. Previous studies indicated that high body weight or obsess people presented more pronated feet, i.e., a smaller CPEI [72]. However, in this study, we observed contradictory results that CPEIs had a positive correlation with weight or BMI in our participants of both sexes ($r = 0.181$ and 0.220, $p < 0.001$ and < 0.001). Although over 1000 strides were included in the training data, because of the small participant size, we still cannot deny the contingency in them, which suggests we should include more participants for training in the future. However, previous studies indicated that

runners have more chance of suffering from foot overpronation because their feet overpronate during running [10], [60]. Participants with a normal BMI may have made a great effort to keep their bodies fit through daily exercise, e.g., jogging. However, in this study, we did not aggregate questionnaires on exercise habits, which should be considered to improve model precision in the future.

D. Limitations of This Study

In this section, we discuss several limitations of this study. Our model has room for improvement in precision and usability if these limitations can be solved in the future.

This study successfully demonstrated the feasibility of estimating foot function by the IMS. However, the ICCs and MAEs showed room for improvement in precision in our model. We only take the integral average of the GPCs' amplitude to construct IMS predictors. This means that the information of waveform segments in GPCs was compressed into a point. Eskofier et al. [21] suggested that gait patterns were essential in many applications where the relevance of adjacent points on the temporal axis should also be considered. To achieve that in the future, we will use more predictors to express the waveform shape in the GPCs into the model. These new predictors can be constructed from, e.g., principle components or autocorrelation coefficients, under the premise that the IMS capacity can afford the calculation.

In the validation, we conducted the experiments one after the other when changing shoes (from S4 to S5 and S7 to S8). However, apparent decreases in hit rate from S4 to S5 and S7 to S8 in females were observed. This might reflect the impact of changing shoes on gait because we needed to give participants more time to acclimatize to the changed shoes. Unlike the male participants' shoes, we observed that the female participants' own shoes had various shapes, e.g., sneakers or pumps with/without high heels that was far different from the shape of the standard shoes. It was possible that participants needed time to adjust their senses to adapt to the change of shoes. After they had adapted, their gait returned to normal, which was considered the reason that the hit rate resumed from S6 to S7 and S9 to S10 despite the ground surface differing from the last experimental scenarios. This phenomenon also agrees with the previous study of Melvin et al. [73], which indicated that unfamiliar or new-bought shoes in daily living should be addressed in practical use. Fortunately, the previous study suggested that people can acclimatize to their new shoes by only continuously walking over 200 strides. Therefore, we may need users to record the start date of using their new-bought shoes in the application and mask the data measured on that day in the database from the gait analysis.

We avoided conducting the outdoor experiments on rainy, snowy, sweltering, and cold days and did not consider the impacts of weather and climates on gait. However, these environmental factors, which were reported to affect gait [74], [75], must be considered in the practical daily measurements in the future. To ensure the reliability of the estimation values in practical use, as a possible solution, we may first conduct the weather and GPs information into the gait analysis application

to filter those data obtained outdoors in the weather condition mentioned above, second treat the filtered data as defective time-series data, and finally interpolate them via a Kalman filter approach [76].

VI. CONCLUSION

In this study, we successfully developed a simple and high-precision foot function estimation model by merging biomechanical knowledge. The model only included a few steps of simple addition and multiplication, which is feasible for an IMS with a power-saving function for long-term monitoring.

By analyzing the correlation between foot motion and CPEI via SPM, we discovered significant sex-specific predictors for foot function assessment, most of which exist in the foot motion in the frontal plane and can connect well with previous biomechanical findings. The model of males and females achieved large effect-sized R^2, MAEs of 6.05 and 3.30 of CPEIs and, a fair ICC agreement when using IMS as a single rater. Compared with the model without IMS predictors determined from foot-function-correlated gait phases, the model's precision with them was noticeably improved, suggesting that foot function cannot be estimated well without predictors integrating biomechanical knowledge.

For validation, the model successfully estimated a maximum of 99.0% and 100.0% males' and females' data under the same experimental conditions with the training data, and 92.8%–100.0% and 85.8%–100.0% data in the combinations of different ground surfaces, different shoes, and with/without baggage conditions. These results suggest that our model is reliable and effective in daily foot function assessment.

In the future, we will focus on constructing new predictors to improve model precision, improve robustness in more different daily living environments, validate the model on older adults or children, validate long-term monitoring, and validate intervention studies on correcting overpronation and oversupination.

ACKNOWLEDGMENT

REFERENCES

[1] M. Razeghi and M. E. Batt, "Foot type classification: A critical review of current methods," *Gait Posture*, vol. 15, no. 3, pp. 282–291, 2002, doi: 10.1016/S0966-6362(01)00151-5.

[2] K. R. Kaufman, S. K. Brodine, R. A. Shaffer, C. W. Johnson, and T. R. Cullison, "The effect of foot structure and range of motion on musculoskeletal overuse injuries," *Amer. J. Sports Med.*, vol. 27, no. 5, pp. 585–593, Sep. 1999.

[3] Y. M. Golightly, M. T. Hannan, A. B. Dufour, H. J. Hillstrom, and J. M. Jordan, "Foot disorders associated with overpronated and over-supinated foot function: The Johnston county osteoarthritis project," *Foot Ankle Int.*, vol. 35, no. 11, pp. 1159–1165, Nov. 2014, doi: 10.1177/1071100714543907.

[4] N. H. Cho, S. Kim, D.-J. Kwon, and H. A. Kim, "The prevalence of hallux valgus and its association with foot pain and function in a rural Korean community," *J. Bone Joint Surg., Brit. volume*, vol. 91, no. 4, pp. 494–498, Apr. 2009, doi: 10.1302/0301-620X.91B4.21925.

[5] H. B. Menz, E. Roddy, E. Thomas, and P. R. Croft, "Impact of hallux valgus severity on general and foot-specific health-related quality of life," *Arthritis Care Res.*, vol. 63, no. 3, pp. 396–404, 2011, doi: 10.1002/acr.20396.

[6] H. B. Menz and M. E. Morris, "Determinants of disabling foot pain in retirement village residents," *J. Amer. Podiatric Med. Assoc.*, vol. 95, no. 6, pp. 573–579, Nov. 2005, doi: 10.7547/0950573.

[7] A. Barnes, J. Wheat, and C. Milner, "Association between foot type and tibial stress injuries: A systematic review," *Brit. J. Sports Med.*, vol. 42, no. 2, pp. 93–98, Jun. 2007, doi: 10.1136/bjsm.2007.036533.

[8] P. Levinger, H. B. Menz, M. R. Fotoohabadi, J. A. Feller, J. R. Bartlett, and N. R. Bergman, "Foot posture in people with medial compartment knee osteoarthritis," *J. Foot Ankle Res.*, vol. 3, no. 1, pp. 1–8, Dec. 2010, doi: 10.1186/1757-1146-3-29.

[9] A. Jafarnezhadgero, A. Fatollahi, A. Sheykholeslami, V. C. Dionisio, and M. Akrami, "Long-term training on sand changes lower limb muscle activities during running in runners with over-pronated feet," *Biomed. Eng. OnLine*, vol. 20, no. 1, p. 118, Nov. 2021, doi: 10.1186/s12938-021-00955-8.

[10] B. Hintermann and B. M. Nigg, "Pronation in runners," *Sports Med.*, vol. 26, no. 3, pp. 169–176, 1998, doi: 10.2165/00007256-199826030-00003.

[11] G. S. Murley, K. B. Landorf, H. B. Menz, and A. R. Bird, "Effect of foot posture, foot orthoses and footwear on lower limb muscle activity during walking and running: A systematic review," *Gait Posture*, vol. 29, no. 2, pp. 172–187, 2009, doi: 10.1016/j.gaitpost.2008.08.015.

[12] H. A. Banwell, S. Mackintosh, D. Thewlis, and K. B. Landorf, "Consensus-based recommendations of Australian podiatrists for the prescription of foot orthoses for symptomatic flexible pes planus in adults," *J. Foot Ankle Res.*, vol. 7, no. 1, pp. 1–13, Dec. 2014, doi: 10.1186/s13047-014-0049-2.

[13] A. Jafarnezhadgero, S. M. Alavi-Mehr, and U. Granacher, "Effects of anti-pronation shoes on lower limb kinematics and kinetics in female runners with pronated feet: The role of physical fatigue," *PLoS ONE*, vol. 14, no. 5, May 2019, Art. no. e0216818, doi: 10.1371/journal.pone.0216818.

[14] P. Levinger and W. Gilleard, "Relationship between static posture and rearfoot motion during walking in patellofemoral pain syndrome: Effect of a reference posture for gait analysis," *J. Amer. Podiatric Med. Assoc.*, vol. 96, no. 4, pp. 323–329, Jul. 2006, doi: 10.7547/0960323.

[15] A. Simkin, I. Leichter, M. Giladi, M. Stein, and C. Milgrom, "Combined effect of foot arch structure and an orthotic device on stress fractures," *Foot Ankle*, vol. 10, no. 1, pp. 25–29, Aug. 1989, doi: 10.1177/107110078901000105.

[16] A. C. Redmond, J. Crosbie, and R. A. Ouvrier, "Development and validation of a novel rating system for scoring standing foot posture: The foot posture index," *Clin. Biomech.*, vol. 21, no. 1, pp. 89–98, Jan. 2006, doi: 10.1016/j.clinbiomech.2005.08.002.

[17] J. Song, H. Hillstrom, D. Secord, and J. Levitt, "Foot type biomechanics. Comparison of planus and rectus foot types," *J. Amer. Podiatric Med. Assoc.*, vol. 86, no. 1, pp. 16–23, Jan. 1996, doi: 10.7547/87507315-86-1-16.

[18] H. B. Menz, A. B. Dufour, J. L. Riskowski, H. J. Hillstrom, and M. T. Hannan, "Association of planus foot posture and pronated foot function with foot pain: The Framingham foot study," *Arthritis Care Res.*, vol. 65, no. 12, pp. 1991–1999, Dec. 2013, doi: 10.1002/acr.22079.

[19] J. Berengueres, M. Fritschi, and R. McClanahan, "A smart pressure-sensitive insole that reminds you to walk correctly: An orthotic-less treatment for over pronation," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 2488–2491, doi: 10.1109/EMBC.2014.6944127.

[20] M. J. Domínguez-Morales, F. Luna-Perejón, L. Miró-Amarante, M. Hernández-Velázquez, and J. L. Sevillano-Ramos, "Smart footwear insole for recognition of foot pronation and supination using neural networks," *Appl. Sci.*, vol. 9, no. 19, p. 3970, Sep. 2019, doi: 10.3390/app9193970.

[21] B. M. Eskofier et al., "An overview of smart shoes in the Internet of Health Things: Gait and mobility assessment in health promotion and disease monitoring," *Appl. Sci.*, vol. 7, no. 10, p. 986, 2017, doi: 10.3390/app7100986.

[22] D. Gokalgandhi, L. Kamdar, N. Shah, and N. Mehendale, "A review of smart technologies embedded in shoes," *J. Med. Syst.*, vol. 44, no. 9, pp. 1–9, Sep. 2020, doi: 10.1007/s10916-020-01613-7.

[23] C. Huang et al., "Foot-healthcare application using inertial sensor: Estimating first metatarsophalangeal angle from foot motion during walking," *IEEE Sensors J.*, vol. 22, no. 3, pp. 2835–2844, Dec. 2022, doi: 10.1109/JSEN.2021.3138485.

[24] K. Fukushi, C. Huang, Z. Wang, H. Kajitani, F. Nihey, and K. Nakahara, "On-line algorithms of stride-parameter estimation for in-shoe motion-sensor system," *IEEE Sensors J.*, vol. 22, no. 10, pp. 9636–9648, May 2022, doi: 10.1109/JSEN.2022.3164057.

[25] L. V. Nguyen, H. M. La, J. Sanchez, and T. Vu, "A smart shoe for building a real-time 3D map," *Autom. Construct.*, vol. 71, pp. 2–12, Nov. 2016, doi: 10.1016/j.autcon.2016.03.001.

[26] D. B. Irving, J. L. Cook, M. A. Young, and H. B. Menz, "Obesity and pronated foot type may increase the risk of chronic plantar heel pain: A matched case-control study," *BMC Musculoskeletal Disorders*, vol. 8, no. 1, pp. 1–8, Dec. 2007, doi: 10.1186/1471-2474-8-41.

[27] V. H. Chuter, "Relationships between foot type and dynamic rearfoot frontal plane motion," *J. Foot Ankle Res.*, vol. 3, no. 1, pp. 1–6, Dec. 2010, doi: 10.1186/1757-1146-3-9.

[28] X. Zhang, J. Aeles, and B. Vanwanseele, "Comparison of foot muscle morphology and foot kinematics between recreational runners with normal feet and with asymptomatic over-pronated feet," *Gait Posture*, vol. 54, pp. 290–294, May 2017, doi: 10.1016/j.gaitpost.2017.03.030.

[29] J. R. Houck, J. M. Tome, and D. A. Nawoczenski, "Subtalar neutral position as an offset for a kinematic model of the foot during walking," *Gait Posture*, vol. 28, no. 1, pp. 29–37, Jul. 2008, doi: 10.1016/j.gaitpost.2007.09.008.

[30] P. Levinger, G. S. Murley, C. J. Barton, M. P. Cotchett, S. R. McSweeney, and H. B. Menz, "A comparison of foot kinematics in people with normal- and flat-arched feet using the Oxford foot model," *Gait Posture*, vol. 32, no. 4, pp. 519–523, Oct. 2010, doi: 10.1016/j.gaitpost.2010.07.013.

[31] C. Huang, Z. Wang, K. Fukushi, F. Nihey, H. Kajitani, and K. Nakahara, "Assessment of over-pronated/over-supinated foot using foot-motion measured by an in-shoe motion sensor," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2021, pp. 1–6, doi: 10.1109/BioCAS49922.2021.9644969.

[32] T. C. Pataky, M. A. Robinson, and J. Vanrenterghem, "Vector field statistical analysis of kinematic and force trajectories," *J. Biomech.*, vol. 46, no. 14, pp. 2394–2401, 2013, doi: 10.1016/j.jbiomech.2013.07.031.

[33] T. C. Pataky, M. A. Robinson, and J. Vanrenterghem, "Region-of-interest analyses of one-dimensional biomechanical trajectories: Bridging 0D and 1D theory, augmenting statistical power," *PeerJ*, vol. 4, p. e2652, Nov. 2016, doi: 10.7717/peerj.2652.

[34] M. Zrenner, C. Feldner, U. Jensen, N. Roth, R. Richer, and B. M. Eskofier, "Evaluation of foot kinematics during endurance running on different surfaces in real-world environments," in *Proc. Int. Symp. Comput. Sci. Sport*. Cham, Switzerland: Springer, Jul. 2019, pp. 106–113, doi: 10.1007/978-3-030-35048-2_13.

[35] C. Morio, M. J. Lake, N. Gueguen, G. Rao, and L. Baly, "The influence of footwear on foot motion during walking and running," *J. Biomech.*, vol. 42, no. 13, pp. 2081–2088, Sep. 2009, doi: 10.1016/j.jbiomech.2009.06.015.

[36] B. Smith, M. Roan, and M. Lee, "The effect of evenly distributed load carrying on lower body gait dynamics for normal weight and overweight subjects," *Gait Posture*, vol. 32, no. 2, pp. 176–180, Jun. 2010, doi: 10.1016/j.gaitpost.2010.04.007.

[37] M. Miezal, B. Taetz, and G. Bleser, "Real-time inertial lower body kinematics and ground contact estimation at anatomical foot points for agile human locomotion," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 3256–3263, doi: 10.1109/ICRA.2017.7989371.

[38] M. Chen, B. Huang, and Y. Xu, "Intelligent shoes for abnormal gait detection," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2008, pp. 2019–2024, doi: 10.1109/ROBOT.2008.4543503.

[39] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc., B, Methodolog.*, vol. 58, no. 1, pp. 267–288, 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.

[40] T. Park and G. Casella, "The Bayesian Lasso," *J. Amer. Stat. Assoc.*, vol. 103, no. 482, pp. 681–686, 2008, doi: 10.1198/016214508000000337.

[41] S. Scardapane, D. Comminiello, A. Hussain, and A. Uncini, "Group sparse regularization for deep neural networks," *Neurocomputing*, vol. 241, pp. 81–89, Jun. 2017, doi: 10.1016/j.neucom.2017.02.029.

[42] S. Roberts and G. Nowak, "Stabilizing the lasso against cross-validation variability," *Comput. Statist. Data Anal.*, vol. 70, pp. 198–211, Feb. 2014, doi: 10.1016/j.csda.2013.09.008.

[43] C. F. J. Wu, "Jackknife, bootstrap and other resampling methods in regression analysis," *Ann. Statist.*, vol. 14, no. 4, pp. 1261–1295, Dec. 1986, doi: 10.1214/aos/1176350142.

[44] S. O. H. Madgwick, A. J. L. Harrison, and A. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Jun. 2011, pp. 1–7, doi: 10.1109/ICORR.2011.5975346.

[45] P. Catalfamo, D. Moser, S. Ghoussayni, and D. Ewins, "Detection of gait events using an F-scan in-shoe pressure measurement system," *Gait Posture*, vol. 28, no. 3, pp. 420–426, 2008, doi: 10.1016/j.gaitpost.2008.01.019.

[46] M. Sangeux and J. Polak, "A simple method to choose the most representative stride and detect outliers," *Gait Posture*, vol. 41, no. 2, pp. 726–730, Feb. 2015, doi: 10.1016/j.gaitpost.2014.12.004.

[47] C. Huang, K. Fukushi, Z. Wang, H. Kajitani, F. Nihey, and K. Nakahara, "Initial contact and toe-off event detection method for in-shoe motion sensor," in *Activity and Behavior Computing*. Singapore: Springer, 2021, pp. 101–118, doi: 10.1007/978-981-15-8944-7_7.

[48] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Amsterdam, The Netherlands: Elsevier, 2011.

[49] Z. Šidák, "Rectangular confidence regions for the means of multivariate normal distributions," *J. Amer. Stat. Assoc.*, vol. 62, no. 318, pp. 626–633, 1967, doi: 10.1080/01621459.1967.10482935.

[50] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike* (Springer Series in Statistics), E. Parzen, K. Tanabe, and G. Kitagawa, Eds. New York, NY, USA: Springer, 1998.

[51] P. E. Shrout and J. L. Fleiss, "Intraclass correlations: Uses in assessing rater reliability," *Psychol. Bull.*, vol. 86, no. 2, pp. 420–428, 1979, doi: 10.1037/0033-2909.86.2.420.

[52] D. V. Cicchetti, "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology," *Psychol. Assessment*, vol. 6, no. 4, pp. 284–290, Dec. 1994, doi: 10.1037/1040-3590.6.4.284.

[53] J. Cohen, "A power primer," *Psychol. Bull.*, vol. 112, no. 1, pp. 155–159, Jul. 1992, doi: 10.1037/0033-2909.112.1.155.

[54] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, Feb. 1986, doi: 10.1016/S0140-6736(86)90837-8.

[55] J. M. Bland and D. G. Altman, "Measuring agreement in method comparison studies," *Stat. Methods Med. Res.*, vol. 8, no. 2, pp. 135–160, Jun. 1999, doi: 10.1177/096228029900800204.

[56] E. M. Hennig. (2002). *The Human Foot During Locomotion-Applied Research for Footwear*. [Online]. Available: https://www.researchgate.net/publication/230824746_Wei_Lun_Public_Lecture_The_human_foot_during_locomotion_-_Applied_research_for_footwear

[57] J. Song et al., "Comprehensive biomechanical characterization of feet in USMA cadets: Comparison across race, gender, arch flexibility, and foot types," *Gait Posture*, vol. 60, pp. 175–180, Jan. 2018, doi: 10.1016/j.gaitpost.2017.12.001.

[58] K. R. Hovinga and A. L. Lerner, "Anatomic variations between Japanese and Caucasian populations in the healthy young adult knee joint," *J. Orthopaedic Res.*, vol. 27, no. 9, pp. 1191–1196, Sep. 2009, doi: 10.1002/jor.20858.

[59] D. Rosenbaum, S. Hautmann, M. Gold, and L. Claes, "Effects of walking speed on plantar pressure patterns and hindfoot angular motion," *Gait Posture*, vol. 2, pp. 191–197, Sep. 1994, doi: 10.1016/0966-6362(94)90007-8.

[60] J. F. Hornestam, T. R. Souza, P. Arantes, J. Ocarino, and P. L. Silva, "The effect of walking speed on foot kinematics is modified when increased pronation is induced," *J. Amer. Podiatric Med. Assoc.*, vol. 106, no. 6, pp. 419–426, Nov. 2016, doi: 10.7547/15-120.

[61] S. Lv, Z. Huan, X. Chang, Y. Huan, and J. Liang, "Analysis of gait symmetry under unilateral load state," in *Proc. China Conf. Wireless Sensor Netw.* Singapore: Springer, Sep. 2020, pp. 244–256, doi: 10.1007/978-981-33-4214-9_18.

[62] G. Scott, H. B. Menz, and L. Newcombe, "Age-related differences in foot structure and function," *Gait Posture*, vol. 26, no. 1, pp. 68–75, Jun. 2007, doi: 10.1016/j.gaitpost.2006.07.009.

[63] H. B. Menz, "Biomechanics of the ageing foot and ankle: A mini-review," *Gerontology*, vol. 61, no. 4, pp. 381–388, 2015, doi: 10.1159/000368357.

[64] H. Yildirim, A. Mavi, O. Büyükbebeci, and E. Gümüşlburun, "Evaluation of the fibular incisura of the tibia with magnetic resonance imaging," *Foot Ankle Int.*, vol. 24, no. 5, pp. 387–391, May 2003, doi: 10.1177/107110070302400502.

[65] M. Fukano, T. Fukubayashi, and S. A. Banks, "Sex differences in three-dimensional talocrural and subtalar joint kinematics during stance phase in healthy young adults," *Hum. Movement Sci.*, vol. 61, pp. 117–125, Oct. 2018, doi: 10.1016/j.humov.2018.06.003.

[66] M.-C. Chiu, H.-C. Wu, and L.-Y. Chang, "Gait speed and gender effects on center of pressure progression during normal walking," *Gait Posture*, vol. 37, no. 1, pp. 43–48, Jan. 2013, doi: 10.1016/j.gaitpost.2012.05.030.

[67] N. A. Schwarz, J. E. Kovaleski, R. J. Heitman, L. R. Gurchiek, and C. Gubler-Hanna, "Arthrometric measurement of ankle-complex motion: Normative values," *J. Athletic Training*, vol. 46, no. 2, pp. 126–132, Mar. 2011, doi: 10.4085/1062-6050-46.2.126.

[68] F. Bonnel, E. Toullec, C. Mabit, and Y. Tourné, "Chronic ankle instability: Biomechanics and pathomechanics of ligaments injury and associated lesions," *Orthopaedics Traumatol., Surg. Res.*, vol. 96, no. 4, pp. 424–432, Jun. 2010, doi: 10.1016/j.otsr.2010.04.003.

[69] D. A. Neumann, *Kinesiology of the Musculoskeletal System: Foundations of Physical Rehabilitation*, 2nd ed. St Louis, MO, USA: Mosby, 2010, pp. 627–699.

[70] H. S. Shin, J. H. Lee, E. J. Kim, M. G. Kyung, H. J. Yoo, and D. Y. Lee, "Flatfoot deformity affected the kinematics of the foot and ankle in proportion to the severity of deformity," *Gait Posture*, vol. 72, pp. 123–128, Jul. 2019, doi: 10.1016/j.gaitpost.2019.06.002.

[71] N. Krähenbühl, T. Horn-Lang, B. Hintermann, and M. Knupp, "The subtalar joint: A complex mechanism," *EFORT Open Rev.*, vol. 2, no. 7, pp. 309–316, Jul. 2017, doi: 10.1302/2058-5241.2.160050.

[72] T. R. Aurichio, J. R. Rebelatto, and A. P. de Castro, "The relationship between the body mass index (BMI) and foot posture in elderly people," *Arch. Gerontol. Geriatrics*, vol. 52, no. 2, pp. e89–e92, Mar. 2011, doi: 10.1016/j.archger.2010.06.014.

[73] J. M. A. Melvin, S. Preece, C. J. Nester, and D. Howard, "An investigation into plantar pressure measurement protocols for footwear research," *Gait Posture*, vol. 40, no. 4, pp. 682–687, Sep. 2014, doi: 10.1016/j.gaitpost.2014.07.026.

[74] N. U. Ahamed et al., "Using wearable sensors to classify subject-specific running biomechanical gait patterns based on changes in environmental weather conditions," *PLoS ONE*, vol. 13, no. 9, 2018, Art. no. e0203839, doi: 10.1371/journal.pone.0203839.

[75] S. Liang, H. Leng, Q. Yuan, B. Wang, and C. Yuan, "How does weather and climate affect pedestrian walking speed during cool and cold seasons in severely cold areas?" *Building Environ.*, vol. 175, May 2020, Art. no. 106811, doi: 10.1016/j.buildenv.2020.106811.

[76] V. Gómez and A. Maravall, "Estimation, prediction, and interpolation for nonstationary series with the Kalman filter," *J. Amer. Stat. Assoc.*, vol. 89, no. 426, pp. 611–624, Jun. 1994.

Chenhui Huang (Member, IEEE) received the B.S. degree in optical engineering from the Harbin Institute of Technology, Harbin, China, in 2009, and the M.S. and Ph.D. degrees in biomedical engineering from Tohoku University, Sendai, Japan, in 2012 and 2015, respectively.

He joined the Center Research Laboratories, NEC Corporation, Abiko, Japan, in 2015. Initially, his work concerned remote sensing technology using quantum dot infrared photodetectors. In 2018, he started working on healthcare technologies based on daily gait analysis. His current research interests include biomechanical analysis on gait, gait analysis and its application for daily healthcare, and activity monitoring using motion sensors, especially foot-mounted ones, involving the use of signal processing and machine learning.

Dr. Huang was a recipient of the IEEE EMBS Japan Chapter Young Researcher Award in 2022. He is a member of SPIE, the Institute of Electronics, Information and Communication Engineers of Japan, and the Japanese Society of Applied Physics.

Fumiyuki Nihey received the M.S. and Ph.D. degrees in physics from Osaka University, Suita, Japan, in 1987 and 1996, respectively.

In 1987, he joined the Central Research Laboratories, NEC Corporation, Abiko, Japan, where he engaged in fundamental research on microelectronics and nanotechnology. In 2020, he joined a research group on healthcare technologies based on gait analysis. His current research interests include the biomechanical analysis of gait, its application to daily healthcare, and activity monitoring of the elderly using foot motion sensors.

Dr. Nihey is a member of the Japan Society of Applied Physics.

Kenichiro Fukushi received the B.Eng., M.Eng., and D.Eng. degrees in information processing from the Tokyo Institute of Technology, Tokyo, Japan, in 2003, 2009, and 2014, respectively.

From 2010 to 2012, he was a Visiting Student with the MIT Media Laboratory, Cambridge, MA, USA. In 2013, he joined NEC Corporation, Abiko, Japan, where he is currently a Research Manager. His research interests include wearable sensing, machine learning, computer vision, and computational photography.

Dr. Fukushi is a member of the Institute of Electronics, Information and Communication Engineers, Japan.

Hiroshi Kajitani received the B.S. and M.S. degrees in electrical engineering from Kyoto University, Kyoto, Japan, in 1986 and 1988, respectively.

He joined the Center Research Laboratories, NEC Corporation, Abiko, Japan, in 1988. His work concerned magnetic recording, 3-D MEMS optical switches, and fuel cell systems for portable electric devices from 1988 to 2016. In 2017, he started working on healthcare technologies, especially low-power Internet of Things (IoT) sensors for monitoring human information and is now developing gait monitoring sensors.

Mr. Kajitani is a member of the Japanese Society of Applied Physics, the Japan Society of Mechanical Engineers, and the Robotics Society of Japan.

Yoshitaka Nozaki received the B.S. and M.S. degrees in electronic engineering from Tohoku University, Sendai, Japan, in 2019 and 2021, respectively.

He joined the Center Research Laboratories, NEC Corporation, Abiko, Japan, in 2021. Initially, his work concerned healthcare technologies based on daily gait analysis. In 2022, he started working on human action recognition and skeleton-based motion synthesis. His current research interests include human action recognition, adversarial learning, motion synthesis, and activity monitoring using wearable sensors, involving the use of signal processing and machine learning.

Kazuki Ihara received the B.Eng. and M.Eng. degrees in applied chemistry and the Doctor of Engineering degree in chemical system engineering from Kyushu University, Fukuoka, Japan, in 2002, 2004, and 2007, respectively.

He joined the Central Research Laboratories, NEC Corporation, Abiko, Japan, in 2007. His research experiences are in the variety research fields such as in the nanocarbon material, in the spin Seebeck device, in the catalytic science, and in the healthcare device applications. Currently, his research interests are focusing on the analysis of the gait for the healthcare applications.

Dr. Ihara is a member of AAAS, ACS, the Chemical Society of Japan, the Japan Society for Analytical Chemistry, the Laser Society of Japan, and the Japan Society of Applied Physics.

Kentaro Nakahara received the B.S. and M.S. degrees from the Department of Applied Chemistry, University of Tokyo, Tokyo, Japan, in 1997 and 1999, respectively, and the Ph.D. degree from the Graduate School of Advanced Science and Engineering, Waseda University, Tokyo, in 2011.

He joined NEC Corporation, Abiko, Japan, in 1999, and had been engaged in the research of rechargeable batteries for over a decade. He is now broadly interested in applied research in the healthcare field of advanced artificial intelligence and Internet of Things (IoT) technologies, managing research such as gait analysis and stress estimation technologies.

Dr. Nakahara's representative research achievement was the discovery of a new type of rechargeable battery named the "Organic Radical Battery," and for this achievement, he received the 10th Green Sustainable Chemistry (GSC) Award, the 17th Chemistry and Bio-Technology Tsukuba Award, and the 2009 Young Scientists Award from the Ministry of Education, Culture, Sports, Science and Technology.