Algorithms for Walking Speed Estimation Using a Lower-Back-Worn Inertial Sensor: A Cross-Validation on Speed Ranges

A. Soltani, K. Aminian, *Senior Member, IEEE*, C. Mazza, A. Cereatti, L. Palmerini, T. Bonci, and A. Paraschiv-Ionescu

***Abstract*— Walking/gait speed is a key measure for daily mobility characterization. To date, various studies have attempted to design algorithms to estimate walking speed using an inertial sensor worn on the lower back, which is considered as a proper location for activity monitoring in daily life. However, these algorithms were rarely compared and validated on the same datasets, including people with different preferred walking speed. This study implemented several original, improved, and new algorithms for estimating cadence, step length and eventually speed. We designed comprehensive cross-validation to compare the algorithms for walking slow, normal, fast, and using walking aids. We used two datasets, including reference data for algorithm validation from an instrumented mat (40 subjects) and shanks-worn inertial sensors (88 subjects), with normal and impaired walking patterns. The results showed up to 50% performance improvements. Training of algorithms on data from people with different preferred speeds led to better performance. For the slow walkers, an average RMSE of 2.5 steps/min, 0.04 m, and 0.10 m/s were respectively achieved for cadence, step length, and speed estimation. For normal walkers, the errors were 3.5 steps/min, 0.08 m, and 0.12 m/s. An average RMSE of 1.3 steps/min, 0.05 m, and 0.10 m/s were also observed on fast walkers. For people using walking aids, the error significantly increased up to an RMSE of 14 steps/min, 0.18 m, and 0.27 m/s. The results demonstrated the robustness of the proposed combined speed estimation approach for different speed ranges. It achieved an RMSE of 0.10, 0.18, 0.15, and 0.32 m/s for slow, normal, fast, and using walking aids, respectively.**

***Index Terms*— Walking speed, step length, cadence, inertial sensors, slow walkers, walking aids.**

Manuscript received February 15, 2021; revised July 6, 2021; accepted August 2, 2021. Date of publication September 10, 2021; date of current version September 27, 2021. This work was supported by the European Union's Horizon 2020 Research and Innovation Programme and EFPIA via the Innovative Medicine Initiative 2 through the Mobilise-D Project under Grant IMI22017-13-7-820820. *(Corresponding author: A. Paraschiv-Ionescu.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee.

A. Soltani, K. Aminian, and A. Paraschiv-Ionescu are with the Laboratory of Movement Analysis and Measurement (LMAM), EPFL, 1015 Lausanne, Switzerland (e-mail: rsoltanislt@gmail.com; kamiar.aminian@epfl.ch; anisoara.ionescu@epfl.ch).

C. Mazza and T. Bonci are with the Department of Mechanical Engineering, The University of Sheffield, Sheffield S1 3JD, U.K., and also with the Insigneo Institute for in Silico Medicine, The University of Sheffield, Sheffield S1 3JD, U.K. (e-mail: c.mazza@sheffield.ac.uk; t.bonci@sheffield.ac.uk).

A. Cereatti is with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy, and also with the Department of Biomedical Sciences, University of Sassari, 07100 Sassari, Italy (e-mail: andrea.cereatti@polito.it).

L. Palmerini is with the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi," University of Bologna, 40126 Bologna, Italy, and also with the Health Sciences and Technologies-Interdepartmental Center for Industrial Research (CIRI-SDV), University of Bologna, 40126 Bologna, Italy (e-mail: luca.palmerini@unibo.it).

Digital Object Identifier 10.1109/TNSRE.2021.3111681

I. Introduction

WALKING speed has recently emerged as an essential indicator of human functional ability, recognized as the sixth vital sign and a key factor for healthy aging [1]. Moreover, in clinical studies, walking speed has become an essential measure in the characterization of movement-related pathologies, the design and assessment of interventions, and the early detection of functional decline [2]. The critical point is that people might walk differently in unsupervised real-world situations (self-triggered and purposeful gaits) than in supervised settings such as a laboratory/clinical setting. Therefore, designing portable systems to reliably estimate speed in everyday life conditions is essential.

With the development of wearable technologies, an appealing solution to estimate gait speed in a real-life setting is to develop algorithms based on inertial sensors (i.e., accelerometer and gyroscope) mounted on various body segments. However, among the different sensor configurations and locations, a single sensor, worn on the upper body (e.g., lower back (LB), sternum, waist, or wrist), has attracted more attention by providing a user-friendly and straightforward setup for real-world and long-term monitoring [3]–[14].

LB sensor's location offers several advantages. Typically, the sensor is tightly fixed on the body, which reduces movement artifacts and provides the possibility to align the sensor's axes with the body or global coordinate systems. Second, an LB-worn sensor is close enough to the body center of mass (CoM), ensuring a robust gait pattern in the acceleration signal, even in the presence of an abnormal gait. These advantages provide the opportunity to develop biomechanical and physical models for estimating a wide range of gait parameters, from primary outcomes such as cadence, step length, and speed to secondary ones like gait variability and symmetry [3], [10], [12], [13], [15], [16].

For speed estimation based on an LB-mounted inertial sensor, a common approach is to estimate cadence and step length

separately, whose multiplication results in speed. For cadence estimation, several algorithms have been proposed, including both time and frequency domain approaches. Briefly, time-based algorithms are based on detecting step-related temporal events (e.g., initial contacts, ICs), using signal processing techniques for peak enhancement and detection [6], [8], [17], [18]. The second type of algorithms works in the frequency domain and tries to estimate the dominant frequency of the acceleration signal, associated with the step or stride frequencies [19]. Step length can be estimated through biomechanical models (BM, e.g., inverse pendulum model or sensor signal intensity-based algorithms) [3], [12], [20], [21], direct integration of acceleration (DI) [5], and by deploying machine learning (ML) methods (e.g., linear regression, Gaussian process, support vector machine, neural network) [14], [22]–[25].

Recent studies have revealed that people with movement-related disorders, such as Multiple Sclerosis (MS), Parkinson's disease (PD), Hemiparesis (HE), Huntington's disease (HD), or even healthy older adults (OA), generally have a lower range of walking speed than healthy populations. Based on these previous findings, it is well known that algorithms' performances might decrease when analyzing impaired and/or slow gait with/without walking aids due to changes in the acceleration patterns and amplitudes. Therefore, in the light of the abovementioned considerations and evidence, it is crucial for the development of evidence-based clinical gait analysis applications to assess validity of algorithms across both healthy and pathological populations, and to understand to which extent algorithms performances are influenced by acceleration patterns changes due to different speed or gait patterns [26]–[28].

This study pursues a comprehensive cross-validation analysis to investigate speed estimation performance and related parameters using a single LB-mounted sensor. This performance was evaluated at different speeds, with data recorded in healthy and diseased populations. To this end, we developed and improved various algorithms according to methodologies adopted from the existing literature. We also propose new algorithms as well as a new concept to combine multiple algorithms by taking advantage of all approaches into one unique solution towards optimizing the performance. A cross-validation was designed to investigate the performance of algorithms when test and training datasets corresponded to various partitions of walking patterns/speed (i.e., slow, normal, fast, all ranges, as well as using walking aids). The algorithms have been evaluated on two datasets, recorded in healthy and mobility-impaired populations, which included reference values for the estimated gait parameters.

II. Methods

A. Materials and Measurement Protocols

1) Dataset M1: **Instrumentation:** An IMU-based device (Opal^{TM}, APDM) was attached to the subject's lumbar spine (between L4 and S2) using an elastic belt. The IMU contained a 3D accelerometer (±6 g) and a gyroscope, sampled at 128 Hz. A 7-m instrumented mat (GAITRiteTM Electronic Walkway, CIR System Inc.) was employed as the reference for temporal and spatial gait parameters (sampling at 128 Hz), with an accuracy of 12.7 mm and 1 sample, respectively. The mat and the IMU were synchronized (±1 sample) using a custom-made cable.

Participants: 40 subjects (24 women, 16 men, 62 ± 8yrs, 165.8 ± 7.0 cm, 68.6 ± 10.7 kg) from four clinical populations of OA, PD (with a unified Parkinson's disease rating scale of 62.7 ± 19.1), HE (because of stroke with a functional ambulatory category score of 3.3 ± 1.5), and HD (with a unified Huntington's disease rating scale of 34.9 ± 16.9) have been included (ten subjects form each category). The participants were enrolled at the Movement Disorders Clinic of the University of Genoa. Informed written consent was collected, and a local ethics committee approved the protocol.

Protocol: A 12 m path was chosen where the instrumented mat was place 2 m far from the starting line to ensure recording steady state and straight walking. After starting the IMU acquisition, the participants stood at the starting line for a few seconds with their feet parallel. Then, they walked back and forth for one minute at their self-selected comfortable speed, and they turned whenever they reached the end of the path. Participants were allowed to use their walking aids if they needed them. Since only walking periods recorded on the instrumented mat were used for this study, the gait initiation and the turning phases were automatically discarded from the analysis.

2) Dataset M2: **Instrumentation:** Three time-synchronized IMU-based devices (Opal^{TM}, APDM) were mounted on each subject (one around lumbar spine L5, and one on each shank) through adjustable Velcro straps, featuring a 3D accelerometer (±6g) and gyroscope, sampled at 128 Hz. To obtain the reference values for the temporal and spatial gait parameters for dataset M2, the algorithm described in [29] (optimized version of [30] for pathological gait of PD patients) was applied to the angular velocity data of the shanks. The reference algorithm has been previously validated against a motion capture system and errors (mean±std) of 0.002±0.023s, 0.038 ±0.066m, and 0.038 ±0.056 m/s were reported for estimating stride time, stride length, and speed, respectively [29].

Participants: 88 subjects (59 women, 29 men, age 54 ± 9yrs) from two populations of Healthy Control (HC, 24 subjects), and patients with the MS (64 subjects) were included. Their disability status was evaluated by the Expanded Disability Status Scale (EDSS), where a median (range) score of 5.5 (3.0–6.5) was observed for the MS population. The participants were chosen either from the Sheffield MS Clinic at the Royal Hallamshire Hospital or the Sheffield Clinical trial Unit (Sheffield, United Kingdom), with ethics approval granted by the NRES Committee Yorkshire & The Humber-Bradford Leeds (reference 15/YH/0300) and by the North of Scotland Research Ethics Committee (ID: 224422).

Protocol: Each participant walked straight back and forth over a 10 m path for around 6 minutes at their comfortable speed. They could use walking aids and to end the measurement at any time based on their exhaustion. The turnings were detected and discarded using the algorithm proposed in [31].

B. Reference Values

Reference systems provided the cadence, stride length, and speed in stride granularity. In order to compare the results of the reference systems with the LB-based approaches, the

average values of cadence, stride length, and speed over each walking bout (i.e., walking period consists of consecutive strides) were computed. For each walking bout, the average stride length divided by 2 was considered as the average step length.

C. Implemented Algorithms

In order to estimate walking speed using an LB-mounted single sensor, we estimated the cadence and step length, separately, whose multiplication resulted in the speed. This simplification reduces the nonlinearity and complexity of the developed algorithms, which might improve the performance.

1) Preprocessing: State-of-the-art algorithms based on LB-sensor location generally use a 3D accelerometer (a_x, a_y, a_z), and as inputs, the acceleration along unidirectional axes (vertical or anterior-posterior), or acceleration norm, a_{norm}, computed based on (1).

$$a_{norm}(t) = \sqrt{a_x(t)^2 + a_y(t)^2 + a_z(t)^2} \tag{1}$$

The algorithms assume that the accelerometer axes (*x, y, z*) are aligned with the global reference system, and/or the measurement setup includes functional calibration procedures. However, this assumption is not practical for the real-world measurement setup. The alternative solution is to take advantage of a 3D gyroscope available in the IMU devices and correct the sensor orientation using complementary filters like Madgwick [32]. Therefore, we proposed a preprocessing stage including the Madgwick filter to correct the orientation of the vertical acceleration, $a_v(t)$, and the Principal Component Analysis (PCA) during each walking bout to align the anterior-posterior acceleration, $a_{ap}(t)$, with the direction of movement.

2) Cadence Estimation: We developed seven cadence estimation algorithms (CAD1-7) based on state-of-the-art approaches, and two new combined methods (cTime and cALL). For all algorithms, the mean cadence over each walking bout was computed and compared to the reference value. The algorithms were categorized into three approaches: time, frequency, and combined.

a) Time-based approach: Time-based cadence estimation algorithms (CAD1-6) were based on the detection of ICs. Step duration was defined as the period between two consecutive ICs of different feet. Then, the instantaneous cadence was estimated as the inverse function of the step duration (in a minute unit, steps/min). For each algorithm, the mean cadence of each bout was computed.

Peak enhancement technique: for the algorithms sensitive to the step-related peaks in the input signals (CAD2, CAD3, and CAD6), we employed a *peak enhancement* technique adapted from [17]. To this end, a combination of de-trending, zero-phase low pass filtering (FIR, $\approx f_c$3.2 Hz), followed by a continuous wavelet transform (CWT) smoothing and differentiation procedure (scale 10, gauss2), and a Savitzky-Golay filtering were applied to reduce high-frequency noise (movement artifacts) and enhance the step-related peaks.

CAD1: algorithm adapted from [3] where $a_{ap}(t)$ was first low-pass filtered (FIR, $\approx f_c$3.2 Hz) according to [17]. Then, a_{ap} peaks preceding a signal sign change were detected as ICs.

CAD2: algorithm based on [11] where the smoothed acceleration $SWS(k)$ was obtained from $a_{norm}(t)$ using a sliding window of size W equivalent to 0.2s, as in (2).

$$SWS(k) = \sum_{t=k-W+1}^{k} a_{norm}(t) \tag{2}$$

Then, the acceleration differential according to (3), $a_{diff}(k)$ was used for identifying ICs as zero-crossing of the negative-to-positive signal slopes. To improve this algorithm, we applied the peak enhancement technique (previously described) on the $a_{norm}(t)$ before using it as the input.

$$a_{diff}(k) = SWS(k+W) - SWS(k) \tag{3}$$

CAD3: algorithm developed as a combination of the processing techniques described in [6] and [17]. The vertical acceleration, $a_v(t)$, was filtered by the integration and differentiation using the CWT (scale 9, gauss2). Then, ICs were identified by detecting the maxima between zero-crossings, a procedure adopted to increase robustness by avoiding a fixed amplitude threshold. We boosted this method by applying the peak enhancement technique (presented at the beginning of this section) on $a_v(t)$ before using it in this algorithm.

CAD4: algorithm according to [17] where, first, the peak enhancement method was applied on $a_{norm}(t)$ to compute the filtered acceleration, $a_f(t)$. Then, for further enhancement of the step-related peaks, a peak sharpening method was applied corresponding to the Taylor series expansion of $a_f(t)$ where the second (a_f'') and fourth (a_f'''') derivatives are considered (4).

$$a_{sharpen}(t) = a_f(t) - K_2 a_f''(t) + K_4 a_f''''(t) \tag{4}$$

Here, K_2 and K_4 are adjustable factors, empirically found set as 20 and 2, respectively, to optimize the performance on the training data. Eventually, an adaptive threshold was applied on $a_{sharpen}(t)$ to determine the step-related peaks as ICs.

CAD5: according to [9], $a_{ap}(t)$ was linearly de-trended and low-pass filtered using a 2nd-order Butterworth filter ($f_c = 10$ Hz). Then, the signal was integrated and differentiated using CWT (by an estimated scale). Finally, the minima of the processed signal were reported as ICs.

CAD6: according to [18], opening and closing morphological filters were applied to $a_{norm}(t)$ in order to highlight the step-related peaks. Then, the peaks identified as maxima during non-zeros periods in the processed signal were selected as ICs. In order to improve this algorithm, $a_{norm}(t)$ went through the peak enhancement method (previously explained), before applying the morphological filters.

b) Frequency-based approach: **CAD7:** algorithm according to [19], which is based on detecting the dominant peak of the spectrum of the acceleration norm. To this end, a comb function has been applied to the estimated frequency spectrum to sharpen the dominant frequency related to step or stride. Then, a maximum likelihood technique was used to estimate the cadence. The algorithm presented in [19] employs the Euclidean norm, $a_{norm}(t)$, which is a non-linear operation and it might distort the acceleration signal. Hence, we improved the algorithm by estimating the sum of spectrum of each

acceleration channel as stated in (5).

$$OverallSpectrum = AX(f) + AY(f) + AZ(f) \quad (5)$$

Here, AX, AY, and AZ are respectively the spectrum of a_x, a_y, and a_z where f is the frequency variable. We used a 256-point Fast Fourier Transform (FFT) with a Hann window to compute the frequency spectrum of acceleration signals. Finally, mean cadence over each walking bout was calculated.

c) Combined approach: We proposed the combined approach, where the output of different algorithms was averaged with equal weights to generate a combined solution. The hypothesis is that the averaging will reduce the random error and increase robustness across various walking patterns since the performance of each individual algorithm might be different for the multiple datasets. Through this procedure, we proposed two combined cadence estimation algorithms called cTime and cALL, which were respectively the average of the time-based (CAD1-6) and all (CAD1-7) algorithms.

3) Step Length Estimation: Twelve step length algorithms (STPL1-12) plus four combined algorithms (cBM, cDI, cML, and cALL), which have been categorized into four main approaches (i.e., BM, DI, ML, Combined), were implemented. For each algorithm, the mean value of the step length during each walking bout was calculated. To isolate the error of the step length algorithms from the error of cadence estimation (ICs detection), here, we used ICs detected by the reference systems ($ic(m)$ where m is the step's number/index within a walking bout). Furthermore, the ICs were used to derive the reference cadence, $CAD_{step}(m)$, to be used as a parameter by some of the step length algorithms. Since several algorithms required the cadence in per-second granularity, we applied a moving window with the length of one second on $CAD_{step}(m)$ to estimate the instantaneous cadence in second $CAD_{second}(t)$. Note that, in this section, variables A and B are the tuning coefficients, optimized by training.

a) BM-based approach: This approach includes four algorithms (STPL1-4) adapted from the literature and based on the models that describe the human body's biomechanics (e.g., legs, trunk) during walking.

STPL1: algorithm based on the inverted pendulum model [3], [13] where the step length was computed through (6):

$$STPL1[m] = A\left(2\sqrt{2l d_{step}[m] - d_{step}[m]^2}\right) + B \quad (6)$$

Here, $d_{step}[m]$ is the vertical displacement of the body CoM (i.e., LB in this study) during m-th step, and l is the pendulum length (i.e., the leg length). To calculate the vertical displacement of the CoM, in the original method the $a_v(t)$ was double integrated and high-pass filtered with a 4th-order Butterworth filter with a cut-off frequency $f_c = 0.1$ Hz to obtain the vertical position, $d_v(t)$. Then, $d_{step}[m]$ was computed according to (7), within two neighboring ICs.

$$d_{step}[m] = \underbrace{|\max(d_v(t)) - \min(d_v(t))|}_{ic(m)\le t\le ic(m+1)} \quad (7)$$

To enhance the estimation of the vertical displacement, we proposed the method shown in Fig. 1. Here, the high-pass filtered $a_v(t)$ (4th-order Butterworth, $f_c = 0.1$Hz) was integrated to obtain the vertical speed, $V_v(t)$. Then, $V_v(t)$ was high-pass filtered (4th-order Butterworth, $f_c = 1$ Hz, empirically chosen) and integrated (by the *cumsum* function in MATLAB) to compute the vertical displacement, $d_v(t)$. Eventually, we estimated $d_{step}[m]$ using (7).

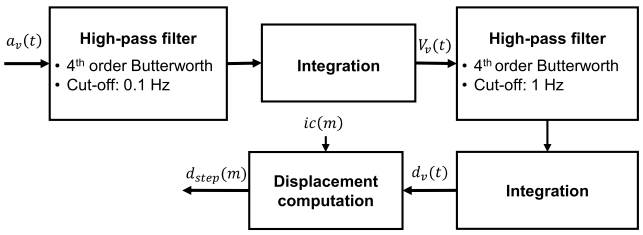

Fig. 1. Block diagram of the proposed method for the estimation of the vertical displacement of body CoM during each step, $d_{step}[m]$. After filtering and integrating $a_V(t)$, the vertical speed ($V_V(t)$) was high-passed and integrated to obtain the vertical displacement ($d_V(t)$). Then, the effect of drift was removed by computing the difference between maximum and minimum of $d_V(t)$ between each neighboring ICs ($ic(m)$).

STPL2: algorithm developed according to [20], where the step length was estimated using the geometrical acceleration-intensity-based model (8).

$$STPL2[m] = A\left(\sqrt[4]{a_{maxmin}(m)}\right) + B \quad (8)$$

Here, $a_{maxmin}(m)$ is the difference between maximum and minimum of $a_v(t)$ during m-th step as defined in (9). We modified this algorithm by filtering $a_v(t)$ with a fourth-order low-pass Butterworth ($f_c = 3$ Hz) before using it in (9).

$$a_{maxmin}(m) = \underbrace{|\max(a_v(t)) - \min(a_v(t))|}_{ic(m)\le t\le ic(m+1)} \quad (9)$$

STPL3: algorithm designed according to [33], [12], where the mean absolute value of $a_v(t)$ during a step duration ($a_{vMean}(m)$ in (10)) was used to estimate step length in (11):

$$a_{vMean}(m) = \underset{ic(m)\le t\le ic(m+1)}{\text{mean}} |a_v(t)| \quad (10)$$

$$STPL3[m] = A\left(\sqrt[3]{a_{vMean}(m)}\right) + B \quad (11)$$

Note that, in the original algorithm, the vertical acceleration is derived from a shank-mounted sensor. However, we used the vertical acceleration obtained from the LB-mounted sensor that should be valid for the algorithm since both vertical accelerations were in the global frame.

STPL4: algorithm based on [21], where the step length was modeled according to (12). Here, $T(m)$ is the duration of the m-th step, which was computed according to (13), and $a_{maxmin}(m)$, $a_{vMean}(m)$ were calculated through (9), (10), respectively. It should be noted that, before using (12), $a_v(t)$ was smoothed by a moving average with 0.125 s length.

$$STPL4[m] = A\left(\sqrt[2.7]{(a_{vMean}(m)) \times \sqrt{\frac{1}{\sqrt{T(m) \times a_{maxmin}(m)}}}}\right) + B \quad (12)$$

$$T[m] = ic(m+1) - ic(m) \quad (13)$$

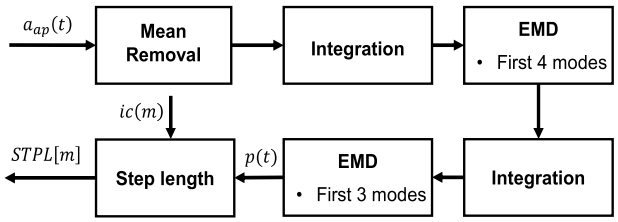

Fig. 2. Block diagram of STPL6. After removing mean of $a_{ap}(t)$ and integrating the zero-mean signal, EMD was used to keep the first 4 intrinsic modes. Then, another integration and EMD were applied on the resulted signal to generate $p(t)$ by the first 3 intrinsic modes. Then, (14) was used to compute step length.

b) DI-based approach: One straightforward way to estimate the step length is to double integrate the forward acceleration, $a_{ap}(t)$, in the global frame. The difficulty of this approach, especially for single LB sensor configuration, is to assure accurate estimation of forward acceleration and remove the accumulated integration drift using an appropriate technique. Only a few methods have been proposed in the literature, such as [5], which needed some requirements like the initial values for the anterior-posterior speed and an expected position of CoM at specific gait events (e.g., ICs). Therefore, we adopted the main ideas and proposed three new algorithms (STPL5-7).

STPL5: first, the acceleration $a_{ap}(t)$ was filtered using a 2nd-order high-pass Butterworth filter ($f_c = 0.5$ Hz). Then, the filtered signal was double integrated to obtain the anterior-posterior position, $p(t)$. Finally, to reduce the effect of drift, step length was computed according to (14). Here, $a_{apMean}(m)$ is the mean value of $a_{ap}(t)$ during m-th step calculated in (15).

$$STPL5[m] = A\left(\underbrace{|\max(p(t)) - \min(p(t))|}_{ic(m)\le t\le ic(m+1)} \times a_{apMean}(m)\right) + B \quad (14)$$

$$a_{apMean}(m) = \underset{ic(m)\le t\le ic(m+1)}{\text{mean}}\left(a_{ap}(t)\right) \quad (15)$$

STPL6: algorithm based on a data-adaptive estimation of integration drift and more effective removal using the Empirical Mode Decomposition (EMD) [12]. As illustrated in Fig. 2, after removing the mean value of $a_{ap}(t)$, the signal was integrated. Then, EMD procedure was applied where only the first four intrinsic modes were used for reconstruction. Next, the resulted signal was again integrated, and EMD applied to remove the drift and to reconstruct the anterior-posterior position, $p(t)$, by keeping only the first three modes. Finally, $p(t)$ was fed into (14) to obtain the step length.

STPL7: another effective approach for removing the integration drift is to reset the integration by an initial value at each gait cycle. For sensors mounted on the lower limbs, especially on foot, the assumption of zero-velocity update at the beginning of each gait cycle has been widely used. However, for the sensors mounted on the upper body (such as LB), this assumption might not be valid since the upper body can move even when the foot is on the ground. In STPL7, we proposed to correct the mean value of the linear speed (i.e., integrated acceleration) by the speed estimated using $STPL1(m)$ and $STPL2(m)$. To this end, as it is shown in Fig. 3, first, $a_{ap}(t)$

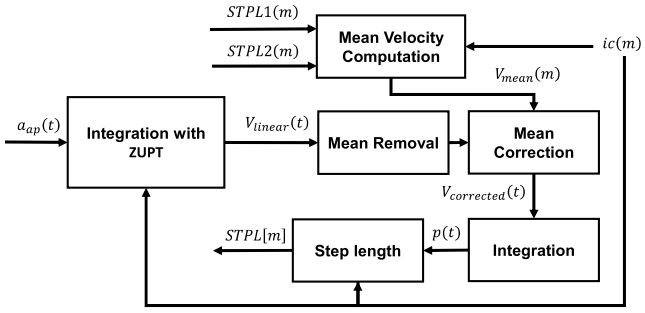

Fig. 3. Block diagram of STPL7. After the integration of $a_{ap}(t)$ with zero-velocity update assumption, mean value of $V_{linear}(t)$ at each gait cycle was replaced with a new mean, $V_{mean}(m)$, computed according to (16) to obtain $V_{corrected}(t)$. Then, resulted signal was integrated and (17) was used to compute step length.

was integrated with the zero-velocity update assumption at each IC to obtain the linear velocity, $V_{linear}(t)$. In parallel, in the *Mean Velocity Computation* block, we calculated the mean value of speed at each step according to (16). Then, the mean value of $V_{linear}(t)$ at each gait cycle was removed and replaced by $V_{mean}[m]$ to obtain $V_{corrected}(t)$. Next, $V_{corrected}(t)$ was integrated to compute the position, $p(t)$. Finally, we employed (17) to estimate the step length.

$$V_{mean}[m] = \left(\frac{STPL1(m) + STPL2(m)}{2}\right) \times CAD(m) \quad (16)$$

$$STPL7[m] = A\left(\max_{ic(m)\le t\le ic(m+1)}(p(t))\right) + B \quad (17)$$

c) ML-based approach: Five algorithms (STPL8-12) based on the machine learning technique were implemented. STPL8-11 were developed and improved according to the literature, while STPL12 was newly developed. For STPL9-12, a moving window with a one-second shift was used to extract features. The length of this window was set according to the minimum value between the length of each walking bout and 5 seconds. For STPL5, the features were extracted during each step duration.

STPL8: in this algorithm, several statistical features such as mean, median, and std of $a_v(t)$ and $a_{ap}(t)$ during each step, as well as demographic information like height and gender were fed into a feedforward 5-layer Neural Network [23].

STPL9: algorithm developed according to [24] where features such as mean, mode, median, std, sum of absolute values, sum of square values, and number of zero crossings of $a_{norm}(t)$, along with height and gender information were fed into a Gaussian process regression model.

STPL10: algorithm based on [7] where a regression model based on Support Vector Machine was deployed to estimate the step length using mean, range, kurtosis, the cross-correlation of the filtered $a_v(t)$ and $a_{ap}(t)$ (2nd order low-pass Butterworth, $f_c = 12$ Hz), and the amplitude of the spectrum of the filtered $a_v(t)$ at dominant frequency. Moreover, cadence ($CAD_{second}(t)$), height and gender were also used in the model.

STPL11: algorithm proposed as an adaptation of the non-personalized version of the algorithm described in [14] to be suitable for the LB-mounted sensor. Features like mean of

TABLE I
DISTRIBUTION OF PARTICIPANTS IN THE CROSS VALIDATION

Datasets	Session	Slow	Normal	Fast	ALL	Walking aids
M1	Train	6	7	6	19	0
	Test	5	6	5	16	5
M2	Train	12	11	8	31	0
	Test	12	11	7	30	27

$a_{norm}(t)$, std of $a_v(t)$, std of $a_{ap}(t)$, mean absolute derivative of $a_{norm}(t)$, cadence ($CAD_{second}(t)$), height, and gender were used in a linear least square regression to model the step length. The original algorithm [14] required the path slope, derived from a barometer, which was discarded in this adapted model.

STPL12: newly proposed algorithms where features like the vertical displacement of CoM (according to (7) and Fig. 1) median, range, and kurtosis of $a_{ap}(t)$, mean and std of $a_{norm}(t)$, mean of $a_v(t)$, cadence ($CAD_{second}(t)$), height, and gender were used in a linear regression model with LASSO (least absolute shrinkage and selection operator) regularization, which guaranteed a non-singular solution, and also provided an intrinsic feature selection [25].

d) Combined approach: Like cadence estimation, here, combined approaches for step length estimation were proposed. Four combined algorithms, cBM, cDI, cML, and cALL were derived as the average of the outputs of BM (STPL1-4), DI (STPL5-7), ML (STPL8-12), and all (STPL1-12) algorithms, respectively.

D. Implementation, Cross-Validation, and Statistical Analysis

All the above-mentioned algorithms were implemented in MATLAB. Comprehensive cross-validation was performed separately for dataset M1 (with the instrumented mat) and M2 (with an IMU-based reference system). In each cross-validation, first, the mean speed of each subject ($V_{refMean}(s)$, where s is the subject index) was computed using the reference speed values. Second, in order to evaluate the effect of speed range and the usage of walking aids, subjects were categorized into four groups as follows: *slow* ($V_{refMean} < 1m/s$), *normal* ($1 \leq V_{refMean} \leq 1.3\ m/s$), *fast* ($V_{refMean} > 1.3m/s$), and *walking aids* (subjects with walking aid). Then, the subjects of each speed category (except the *walking aid* group) were divided into two equal subgroups (i.e., 50-50 %) as training and testing data. Since the subjects with walking aids were not the focus of this study, we included this group only in testing data. In addition, we built ALL_train (for training) and ALL_test (for testing) subcategories, which were respectively the integration of the training and testing data of *slow*, *normal*, and *fast* walkers. To evaluate the implemented algorithms, the Root Mean Square Error (RMSE) between the reference values and the LB algorithms for cadence, step length, and speed estimation were computed on the test dataset.

III. RESULTS

A. Participants

TABLE I shows the number of subjects within each speed category used for the proposed cross-validation analysis.

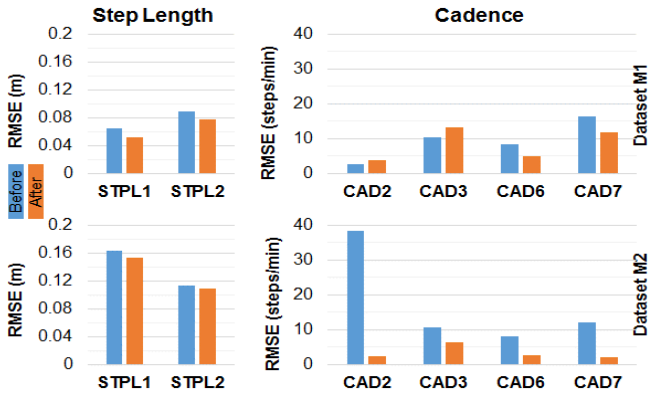

Fig. 4. RMSE of the algorithms before (blue) and after (orange) improvement, for step length (left side) and cadence (right side). The graphs in the upper panel show the results on dataset M1, and the lower panel for M2. The algorithms were trained (if required) and tested on *ALL* category of the train and test sets.

TABLE II
PERFORMANCE OF THE CADENCE ESTIMATION ALGORITHMS TESTED ON DIFFERENT SPEED RANGES FOR BOTH DATASETS M1 AND M2. THE VALUES EXPRES THE RMSE [steps/min]

DS	Approach	Algorithm	Status	Slow	Normal	Fast	ALL	Walk aid
M1	Time	CAD1	Original	13.1	5.1	7.2	8.3	19.9
		CAD2	Modified	2.1	5.4	1.9	3.8	12.1
		CAD3	Modified	13.8	17.8	2.1	13.1	25.8
		CAD4	Original	2.8	3.0	1.9	2.6	42.8
		CAD5	Original	9.6	5.6	6.3	6.9	38.9
		CAD6	Modified	3.8	5.4	4.8	4.9	15.7
	Frequency	CAD7	Modified	2.3	5.8	18.4	11.9	16.3
	cTime	CAD1-6	New	3.9	4.0	2.7	3.6	14.3
	cALL	CAD1-7	New	3.4	3.5	3.4	3.5	13.7
M2	Time	CAD1	Original	14.9	23.1	0.6	17.0	12.3
		CAD2	Modified	2.8	1.3	2.0	2.4	27.6
		CAD3	Modified	7.9	1.6	1.9	6.3	13.6
		CAD4	Original	23.5	2.6	0.9	18.7	43.9
		CAD5	Original	13.8	19.1	0.8	14.9	23.0
		CAD6	Modified	3.3	1.6	1.5	2.8	30.4
	Frequency	CAD7	Modified	2.5	1.8	0.5	2.2	26.1
	cTime	CAD1-6	New	6.6	4.4	1.3	5.8	19.6
	cALL	CAD1-7	New	5.8	3.8	1.1	5.0	19.8

As indicated, there is a proper balance of the number of subjects in the different groups. Note that the walking aids groups were used only for testing. Moreover, since height information was missing in dataset M2, we considered a typical height value of 170 cm in the algorithms which needed height.

B. Performance Improvements

Figure 4 compares the performance of the modified algorithms before (blue) and after (orange) the improvements for estimating cadence (right) and step length (left) in both datasets M1 (top) and M2 (down). The RMSE decreases in most cases after improvement. The algorithms were trained (if required) and tested on the "*ALL*" category of the train and test sets.

C. Cadence Estimation

Table II reports the RMSE of the cadence estimation algorithms tested on both datasets M1 and M2 for different speed

TABLE III

CROSS-VALIDATION RESULTS FOR STEP LENGTH ALGORITHMS. HERE, cBM, cML, cDI, AND cALL ARE COMBINED APPROACHES. COLUMNS SEPARATED BY THE VERTICAL LINES CORRESPOND TO THE TRAINING CONDITIONS. 'S', 'N', 'F', 'A', AND 'WA' REPORT RMSE [m] OF TESTING ON *Slow*, *Normal*, *Fast*, *ALL*, AND WITH WALKING AIDS WALKERS. '*status*' SHOWS IF THE ALGORITHMS ARE ORIGINAL, MODIFIED, OR NEW IN THIS STUDY

DS	App.	Algorithm	Status	Training on slow walkers					Training on normal walkers					Training on fast walkers					Training on ALL				
				S	N	F	A	WA	S	N	F	A	WA	S	N	F	A	WA	S	N	F	A	WA
M1	BM	STPL1	Modified	0.15	0.16	0.19	0.17	0.10	0.05	0.06	0.04	0.05	0.22	0.06	0.10	0.10	0.09	0.28	0.05	0.06	0.04	0.05	0.21
		STPL2	Modified	0.12	0.15	0.23	0.18	0.10	0.06	0.06	0.09	0.07	0.21	0.09	0.09	0.08	0.09	0.24	0.05	0.05	0.11	0.08	0.18
		STPL3	Original	0.11	0.12	0.20	0.15	0.10	0.06	0.09	0.08	0.08	0.18	0.07	0.10	0.08	0.09	0.20	0.05	0.07	0.09	0.07	0.16
		STPL4	Original	0.11	0.13	0.21	0.16	0.10	0.06	0.08	0.09	0.08	0.19	0.08	0.10	0.09	0.09	0.21	0.06	0.07	0.10	0.08	0.17
	DI	STPL5	New	0.13	0.17	0.26	0.21	0.09	0.09	0.06	0.10	0.08	0.26	0.18	0.10	0.09	0.12	0.34	0.08	0.05	0.13	0.09	0.24
		STPL6	New	0.08	0.20	0.39	0.27	0.14	0.09	0.05	0.10	0.08	0.28	0.17	0.09	0.08	0.11	0.34	0.09	0.05	0.10	0.08	0.27
		STPL7	New	0.13	0.16	0.22	0.18	0.11	0.05	0.06	0.06	0.06	0.22	0.08	0.09	0.07	0.08	0.27	0.04	0.05	0.07	0.06	0.21
	ML	STPL8	Original	0.07	0.12	0.24	0.17	0.15	0.11	0.14	0.18	0.15	0.20	0.21	0.18	0.10	0.16	0.54	0.09	0.10	0.12	0.10	0.32
		STPL9	Original	0.15	0.44	0.65	0.49	0.18	0.13	0.08	0.12	0.11	0.27	0.19	0.11	0.10	0.13	0.34	0.12	0.17	0.18	0.16	0.18
		STPL10	Original	0.06	0.27	0.22	0.22	0.17	0.06	0.10	0.13	0.11	0.33	0.15	0.10	0.05	0.10	0.36	0.07	0.16	0.08	0.12	0.18
		STPL11	Original	0.32	0.30	0.35	0.32	0.20	0.09	0.06	0.09	0.08	0.31	0.14	0.09	0.08	0.10	0.33	0.09	0.18	0.07	0.13	0.19
		STPL12	New	0.15	0.19	0.17	0.17	0.15	0.07	0.08	0.08	0.08	0.29	0.13	0.12	0.07	0.10	0.41	0.08	0.13	0.10	0.11	0.15
	cBM	STPL1-4	New	0.12	0.14	0.21	0.17	0.10	0.05	0.06	0.07	0.07	0.20	0.07	0.09	0.08	0.08	0.23	0.04	0.05	0.08	0.06	0.18
	cDI	STPL5-7	New	0.11	0.18	0.29	0.21	0.10	0.08	0.05	0.08	0.07	0.25	0.14	0.09	0.07	0.10	0.31	0.07	0.05	0.10	0.07	0.24
	cML	STPL8-12	New	0.13	0.15	0.13	0.14	0.15	0.07	0.05	0.09	0.07	0.26	0.16	0.10	0.05	0.11	0.39	0.06	0.12	0.07	0.09	0.18
	cALL	STPL1-12	New	0.08	0.13	0.18	0.15	0.12	0.06	0.05	0.08	0.06	0.24	0.13	0.09	0.06	0.09	0.32	0.05	0.08	0.07	0.07	0.18
M2	BM	STPL1	Modified	0.15	0.13	0.08	0.14	0.16	0.18	0.11	0.06	0.15	0.18	0.20	0.10	0.06	0.17	0.20	0.18	0.11	0.06	0.15	0.18
		STPL2	Modified	0.11	0.10	0.06	0.11	0.13	0.13	0.09	0.05	0.11	0.15	0.13	0.08	0.05	0.11	0.15	0.12	0.09	0.05	0.11	0.14
		STPL3	Original	0.13	0.10	0.05	0.11	0.10	0.13	0.09	0.05	0.11	0.10	0.14	0.08	0.05	0.12	0.11	0.13	0.08	0.05	0.11	0.10
		STPL4	Original	0.12	0.11	0.05	0.11	0.10	0.12	0.10	0.05	0.11	0.10	0.12	0.10	0.05	0.11	0.10	0.12	0.10	0.05	0.11	0.10
	DI	STPL5	New	0.10	0.09	0.15	0.10	0.15	0.15	0.04	0.07	0.12	0.21	0.20	0.08	0.02	0.16	0.27	0.15	0.04	0.07	0.12	0.21
		STPL6	New	0.11	0.10	0.13	0.11	0.17	0.15	0.04	0.07	0.12	0.22	0.21	0.07	0.03	0.17	0.29	0.16	0.06	0.05	0.14	0.25
		STPL7	New	0.10	0.12	0.10	0.11	0.13	0.12	0.09	0.07	0.11	0.16	0.13	0.09	0.07	0.11	0.17	0.12	0.09	0.07	0.11	0.16
	ML	STPL8	Original	0.10	0.13	0.15	0.11	0.19	0.12	0.08	0.11	0.11	0.19	0.17	0.07	0.02	0.14	0.21	0.11	0.05	0.09	0.10	0.14
		STPL9	Original	0.11	0.12	0.15	0.12	0.12	0.17	0.06	0.08	0.14	0.23	0.21	0.08	0.02	0.17	0.27	0.15	0.08	0.06	0.13	0.17
		STPL10	Original	0.09	0.09	0.20	0.11	0.21	0.22	0.04	0.12	0.18	0.34	0.27	0.15	0.03	0.23	0.44	0.12	0.07	0.03	0.10	0.13
		STPL11	Original	0.12	0.15	0.13	0.13	0.22	0.24	0.08	0.14	0.20	0.38	0.35	0.16	0.03	0.29	0.51	0.14	0.06	0.07	0.12	0.23
		STPL12	New	0.12	0.14	0.17	0.13	0.22	0.37	0.28	0.15	0.33	0.58	0.25	0.11	0.04	0.21	0.35	0.18	0.07	0.06	0.15	0.22
	cBM	STPL1-4	New	0.12	0.11	0.06	0.12	0.12	0.14	0.10	0.05	0.12	0.13	0.14	0.09	0.05	0.12	0.14	0.13	0.10	0.05	0.12	0.13
	cDI	STPL5-7	New	0.10	0.10	0.12	0.10	0.14	0.13	0.05	0.06	0.11	0.19	0.17	0.06	0.03	0.14	0.23	0.14	0.06	0.06	0.11	0.20
	cML	STPL8-12	New	0.09	0.11	0.14	0.10	0.18	0.25	0.12	0.12	0.21	0.37	0.23	0.10	0.02	0.19	0.33	0.13	0.06	0.05	0.11	0.17
	cALL	STPL1-12	New	0.10	0.11	0.11	0.10	0.15	0.17	0.05	0.07	0.14	0.25	0.18	0.06	0.03	0.15	0.24	0.13	0.07	0.05	0.11	0.16

ranges. The algorithms are categorized by their conceptual groups as Time, Frequency, cTime, and cALL. The status of the algorithms determines whether the algorithms are original from literature, modified, or newly proposed in this study. Besides, algorithms performance on subjects with walking aids is shown.

D. Step Length Estimation

The cross-validation results of step length estimation algorithms are presented in Table III. The algorithms are categorized according to their corresponding conceptual approach as BM, DI, ML, cBM, cML, cDI, and cALL.

E. Walking Speed Estimation

Table IV compares the speed estimation results by multiplying different approaches of the step length and the cadence estimation (each row is one combination). Here, to reduce the number of combinations between the cadence and the stride length algorithms, we only considered the combined approaches that were proper representatives of their corresponding conceptual groups.

IV. DISCUSSION

In this study, we implemented, improved, and compared several LB sensor-based algorithms to estimate cadence, step length, and walking speed. We analyzed data from two datasets, containing both healthy and diseased populations.

Figure 4 illustrates that the proposed enhancements led to improved cadence and step length estimation on both datasets. For the algorithms CAD2 and CAD3 (both time-based), while RMSE substantially decreased on dataset M2 (by at least 50%), a slight increase (maximum 25%) was observed on dataset M1. One possible reason is that, according to our observation, the step-related peaks in M2 were generally weaker than in M1 (probably due to MS disease). That is why the proposed peak enhancement method was generally more effective on M2 dataset than M1. Furthermore, a considerable improvement (minimum 30%) was also achieved for the frequency-based algorithm (CAD7) on both datasets but again more effective on M2. For the step length algorithms (STPL1 and STPL2, both based on BM), the error has been consistently reduced on both datasets M1 (minimum 16 %) and M2 (minimum 10 %).

Referring to Table II, for the normal walkers, almost all cadence algorithms worked well and similar (an average RMSE of 3.5 and 3.8 steps/min on M1 and M2). However, for the slow walkers, the time-based algorithms showed a severe degradation of the performance (up to an error of 13 and 23 steps/min on M1 and M2, respectively), probably due

TABLE IV

CROSS-VALIDATION RESULTS FOR SPEED ESTIMATION. HERE, EACH ROW REPRESENTS ONE COMBINATION OF STEP LENGTH AND CADENCE ALGORITHMS TO COMPUTE SPEED. COLUMNS SEPARATED BY THE VERTICAL LINES CORRESPOND TO TRAINING CONDITIONS. 'S', 'N', 'F', 'A', AND 'WA' REPORT RMSE [m/s] OF TESTING ON SLOW, NORMAL, FAST, ALL, AND WITH WALKING AIDS WALKERS

DS	Step length	Cadence	Training on slow walkers					Training on normal walkers					Training on fast walkers					Training on ALL				
			S	N	F	A	WA	S	N	F	A	WA	S	N	F	A	WA	S	N	F	A	WA
M1	cBM	cTime	0.22	0.29	0.41	0.33	0.17	0.10	0.15	0.15	0.14	0.34	0.14	0.21	0.18	0.19	0.40	0.09	0.12	0.16	0.13	0.32
		Frequency	0.21	0.28	0.51	0.37	0.19	0.10	0.16	0.31	0.22	0.34	0.15	0.23	0.31	0.25	0.39	0.09	0.14	0.32	0.22	0.32
		cALL	0.22	0.29	0.42	0.33	0.17	0.10	0.15	0.17	0.15	0.34	0.14	0.21	0.19	0.19	0.39	0.09	0.12	0.18	0.14	0.31
	cDI	cTime	0.20	0.37	0.59	0.44	0.17	0.15	0.11	0.16	0.14	0.43	0.28	0.20	0.15	0.20	0.52	0.13	0.11	0.19	0.15	0.40
		Frequency	0.19	0.36	0.66	0.47	0.19	0.16	0.13	0.32	0.22	0.42	0.29	0.22	0.31	0.27	0.51	0.14	0.12	0.34	0.23	0.40
		cALL	0.20	0.37	0.60	0.44	0.17	0.15	0.11	0.17	0.15	0.42	0.28	0.20	0.16	0.21	0.52	0.13	0.11	0.20	0.15	0.40
	cML	cTime	0.26	0.33	0.27	0.29	0.28	0.14	0.11	0.18	0.15	0.43	0.31	0.22	0.12	0.22	0.64	0.12	0.28	0.14	0.21	0.31
		Frequency	0.26	0.34	0.38	0.34	0.25	0.15	0.13	0.33	0.23	0.43	0.33	0.24	0.29	0.28	0.62	0.12	0.30	0.29	0.27	0.29
		cALL	0.26	0.33	0.28	0.29	0.28	0.14	0.11	0.20	0.16	0.43	0.31	0.22	0.13	0.22	0.64	0.12	0.28	0.15	0.21	0.30
	cALL	cTime	0.15	0.28	0.36	0.29	0.21	0.13	0.12	0.15	0.13	0.40	0.25	0.21	0.13	0.19	0.54	0.10	0.18	0.14	0.15	0.32
		Frequency	0.15	0.28	0.47	0.34	0.20	0.13	0.13	0.31	0.22	0.40	0.26	0.23	0.30	0.26	0.52	0.11	0.19	0.30	0.23	0.31
		cALL	0.15	0.28	0.37	0.30	0.21	0.13	0.12	0.16	0.14	0.40	0.25	0.21	0.14	0.20	0.53	0.10	0.18	0.15	0.15	0.32
M2	cBM	cTime	0.21	0.20	0.12	0.20	0.24	0.24	0.18	0.10	0.21	0.27	0.25	0.17	0.10	0.22	0.28	0.23	0.18	0.10	0.21	0.26
		Frequency	0.22	0.18	0.11	0.20	0.26	0.24	0.16	0.10	0.21	0.28	0.26	0.16	0.11	0.22	0.29	0.24	0.16	0.10	0.21	0.28
		cALL	0.21	0.19	0.11	0.20	0.24	0.24	0.18	0.10	0.21	0.26	0.25	0.17	0.10	0.22	0.28	0.24	0.18	0.10	0.21	0.26
	cDI	cTime	0.15	0.18	0.27	0.18	0.30	0.22	0.11	0.13	0.19	0.37	0.28	0.13	0.07	0.23	0.43	0.22	0.12	0.13	0.19	0.38
		Frequency	0.16	0.16	0.26	0.17	0.32	0.23	0.09	0.12	0.19	0.39	0.28	0.12	0.07	0.24	0.45	0.23	0.10	0.12	0.19	0.40
		cALL	0.15	0.18	0.27	0.17	0.30	0.22	0.10	0.13	0.19	0.36	0.28	0.12	0.07	0.23	0.43	0.22	0.12	0.12	0.19	0.38
	cML	cTime	0.15	0.20	0.31	0.18	0.35	0.37	0.17	0.27	0.31	0.63	0.37	0.17	0.05	0.31	0.58	0.21	0.11	0.12	0.18	0.33
		Frequency	0.16	0.18	0.30	0.18	0.38	0.37	0.21	0.26	0.32	0.67	0.37	0.19	0.05	0.31	0.60	0.23	0.10	0.11	0.19	0.36
		cALL	0.15	0.20	0.30	0.18	0.35	0.37	0.17	0.27	0.31	0.63	0.37	0.17	0.05	0.30	0.58	0.22	0.11	0.12	0.18	0.33
	cALL	cTime	0.16	0.19	0.23	0.18	0.30	0.27	0.09	0.16	0.22	0.45	0.30	0.13	0.06	0.24	0.45	0.22	0.13	0.10	0.19	0.32
		Frequency	0.17	0.17	0.22	0.18	0.33	0.27	0.10	0.15	0.23	0.48	0.30	0.13	0.07	0.25	0.47	0.22	0.11	0.10	0.19	0.34
		cALL	0.16	0.19	0.23	0.18	0.30	0.27	0.09	0.16	0.22	0.45	0.30	0.13	0.06	0.24	0.44	0.22	0.13	0.10	0.19	0.32

to weakened step-related peaks in the acceleration signal. On the other hand, the frequency-based algorithm (CAD7) achieved a low error of 2.5 steps/min for slow walkers on both datasets, possibly because this approach depends mainly on the gait-related repetitive patterns than the peaks in the time-domain signal. Furthermore, for the fast walkers, while both approaches provided very good results on the M2 dataset (maximum error of 2 steps/min), a significant increase of the error (up to 18 steps/min) was observed for the frequency-based algorithm on M1. One explanation is the severity of gait impairment in patients with HD (n=10 in M1), characterized by mixed unpredictable accelerations and decelerations in walking speed and superimposed twisting movements of the trunk. These characteristics of gait patterns might generate more harmonics in the spectrum of acceleration signals, confusing the frequency-based cadence algorithm to find the correct dominant frequency. For subjects with walking aids, the RMSE increases up to an average RMSE of 14 and 20 steps/min on M1 and M2.

Considering all solutions for the cadence estimation, every algorithm showed several advantages and limitations. Nevertheless, the combined approaches provided a stable performance in all conditions. For instance, the cALL achieved an RMSE of 3.4, 3.5, 3.4, and 13.7 steps/min for slow, normal, fast, walking aids walkers on the M1 dataset. On M2 datasets, the results are 5.8, 3.8, 1.1, 19.8 steps/min, respectively.

Only a few previous studies have reported cadence estimation errors. In [17], the cadence estimation has been evaluated on typically developed and children with cerebral palsy where mean and std absolute errors vary between [0.5-2] and [1.3-7.2] steps/min, respectively. The study [19] also has reported a median [interquartile] of 0.15 [−1.95 2.27] steps/min for the estimation of cadence on healthy subjects and using wrist sensors. Note that the performance of cadence estimation could be affected by target populations (degree of gait impairment), in-lab or real-world situations, the definition of cadence, and the definition of error.

Regarding the different training conditions indicated in Table III, our observation was that training on a specific range of speed might not be necessarily the best choice even when testing was performed on the same speed range (e.g., training on slow walkers and testing on the same group). The results demonstrated that training on *ALL* (i.e., including people from all speed ranges) led to better performance than other training conditions. One main reason might be that more data with higher diversity were fed into the algorithms during training on ALL, resulting in more generalized models.

Considering the column of '*training on ALL*' in Table III, for the slow walkers, the BM-based algorithms showed slightly better performance (RMSE around 0.04 m on M1 and 0.13 m on M2). One reason might be that these algorithms are more dependent on biomechanically-derived models than the intensity or gait-related patterns of the acceleration signal (as ML or DI). Therefore, even when the acceleration signal is weak or distorted due to the slow walking, they could still satisfactory estimate the step length. DI-based algorithms seemed to perform slightly better for the normal walkers, all approaches provided good performances

(RMSE around 0.08 m). No big difference was observed among different approaches for the fast walkers (RMSE around 0.07 m and 0.05 m on M1 and M2). For the group of subjects with walking aids, we noticed a significant performance drop in all approaches (RMSE around 0.18 m). Nevertheless, ML and BM approaches seemed to be more appropriate for this type of walking since they offer a high generalization ability, making them robust against the body's atypical movement that can cause problems for the DI approach. As for cadence estimation, the combined step length approaches (cALL) appeared to be more accurate and robust for all speed categories and *walking aids* group. On the M1 dataset, the combined approach (i.e., cALL) achieved RMSE of 0.05, 0.08, 0.07, and 0.18 m for slow, normal, fast, and walking aids walkers, respectively. On M2 datasets the results are as 0.13, 0.07, 0.05, and 0.16 m, respectively.

Table IV shows promising results for estimating speed. Like the step length, training on ALL walkers generally resulted in a better performance. For slow walkers of dataset M1, the choice of cBM (BM-based combined algorithm) or cALL with any cadence algorithms (i.e., cTime, Frequency, or cALL) achieved a better speed estimation (RMSE around 0.10 m/s). However, on M2, all possible solutions resulted in the same RMSE around 0.22 m/s. Furthermore, for the normal walkers of M1, the combination of cDI or cBM with any cadence algorithms led to better performance (RMSE around 0.12 m/s). On M2, however, the selection of cDI or cML with any cadence algorithms resulted in better performance (RMSE around 0.12 m/s). Moreover, for the fast walkers, excluding the combination of the frequency-based cadence approach with any step length algorithms in M1, the rest of possible solutions showed similar performance (RMSE of 0.14 and 0.10 m/s on M1 and M2).

For the walking-aids group of M1, except the combination of DI with any cadence algorithms, other combinations led to a similar RMSE around 0.32 m/s. However, on M2, the combination of cBM with any cadence algorithms showed better performance (RMSE around 0.27 m/s). Finally, results in table IV demonstrates that the choice of cALL for both step length and cadence led to a more robust and acceptable estimation of speed in all conditions. This solution achieved an RMSE of 0.10, 0.18, 0.15, and 0.32 m/s for slow, normal, fast, and with walking aids walkers, respectively, on the M1 dataset. Besides, it reached RMSE of 0.22, 0.13, 0.10, and 0.32m/s on M2 dataset.

Looking at the literature, [7] achieved an RMSE of [0.12-0.15] m/s to estimate speed on healthy and MS populations. A median absolute error of less than 0.3m/s was also reported in [34] to estimate walking speed in a healthy population. In [14], a median [interquartile] error of 0.10 [0.07 0.12] m/s was obtained for the non-personalized wrist-based speed estimation on healthy population in real-life situations.

For cadence, step length, and speed, a slight difference was observed between the results obtained on dataset M1 (with the instrumented walkway) and M2 (with the IMU-based reference system). Generally speaking, dataset M2 seemed to be more challenging for the algorithms than M1. One reason might be that the IMU-based reference system of M2 had a higher error than the instrumented walkway in M1. Another error source might be using a fixed height value (i.e., 170 cm) in dataset M2, which is needed for some algorithms. Besides, detecting walking bouts of M1 is more reliable than M2 (because of using the instrumented mat as a reference).

One limitation of this study is that the algorithms were evaluated only on data recorded in controlled laboratory settings and on straight walking (removing the turns at the end of the walking path). The small sample size could be another limitation. Moreover, for dataset M2, an IMU-based reference system was used, which could introduce a degree of error in the reference values. Another potential limitation is that the error of step length estimation was isolated from the potential error arising from the detection of ICs; in fact, all step length estimation algorithms were evaluated by considering the ICs detected by the reference system. Using the ICs detected from the LB-mounted IMU-based algorithms might degrade the performance of the presented algorithms. Nevertheless, our analysis stays valid since the main goal was to compare the performance of different algorithms/models rather than reporting cumulated error from the data processing flow [8].

As future work, a similar analysis to the one presented in this study could be performed on more extensive datasets recorded for long durations in daily-life situations. Real-world recorded data could reveal a wide variety of challenging conditions (self-triggered, purposeful, and multitasking walking in a rich behavioral context) to test and evaluate the algorithms' performance. For further improvement, one possibility is to deploy a weighted average instead of an equally-weighted one to optimize the performance of the combined algorithms by tuning the weights (giving more weights to more accurate algorithms). For instance, it would be possible to use the linear least square method to compute the optimal weights on a tuning dataset. The results of this study suggest that future work is necessary to address the challenge of impaired gait patterns, especially when walking aids are used, by using more elaborated signal processing approaches.

Furthermore, some of the time-based cadence algorithms could be used for step demarcation (timing of ICs) in a fully autonomous LB-based speed estimation pipeline. This paper focused on evaluating speed estimation and its related parameters (i.e., cadence and step length). However, a prospective study could be performed to evaluate the step demarcation error to complete previous studies.

V. Conclusion

In this study, several state-of-the-art algorithms for cadence and step length estimation, using data from a single IMU on LB, were implemented and further improved. The proposed improvements allowed a reduction of estimation error in the range of 30-50 % for cadence and 10-16 % for step length. Furthermore, the training of ML models on data from all subjects, and so on a variety of gait patterns and preferred speed, led to better performance.

In a systematic review [35], aiming to summarize information on the minimal clinically significant difference for change in comfortable gait speed measurements for patients with pathology, it is reported that changes of 0.10 to 0.20 m/s may be important across multiple patient groups. The study we conducted demonstrated that some of the proposed algorithms,

for instance, the combined approaches (cALL), achieved to the estimation of gait speed with RMSE in the range of 0.10 – 0.22 m/s, for slow, normal, and fast walking speed and various pathologies (gait impairments). These results appear promising and clinically meaningful. Nevertheless, further improvement is necessary, and our future work will focus on the challenge of very impaired gait patterns when various walking aids are used, where the highest error was observed.

ACKNOWLEDGMENT

The authors would like to thank all consortium members, especially the technical validation work package (WP2). It is worth mentioning that the views expressed in this study are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care, the IMI, the European Union, the EFPIA, or any Associated Partners.

REFERENCES

[1] A. Middleton, S. L. Fritz, and M. Lusardi, "Walking speed: The functional vital sign," *J. Aging Phys. Activity*, vol. 23, no. 2, pp. 314–322, Apr. 2015.

[2] S. R. Cummings, S. Studenski, and L. Ferrucci, "A diagnosis of dismobility—Giving mobility clinical visibility: A mobility working group recommendation," *Jama*, vol. 311, no. 20, pp. 2061–2062, 2014.

[3] W. Zijlstra and A. L. Hof, "Assessment of spatio-temporal gait parameters from trunk accelerations during human walking," *Gait Posture*, vol. 18, no. 2, pp. 1–10, Oct. 2003.

[4] R. C. Gonzalez, D. Alvarez, A. M. Lopez, and J. C. Alvarez, "Modified pendulum model for mean step length estimation," in *Proc. 29th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2007, pp. 1371–1374.

[5] A. Köse, A. Cereatti, and U. D. Croce, "Bilateral step length estimation using a single inertial measurement unit attached to the pelvis," *J. NeuroEng. Rehabil.*, vol. 9, no. 1, pp. 1–10, Dec. 2012.

[6] J. McCamley, M. Donati, E. Grimpampi, and C. Mazzà, "An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data," *Gait Posture*, vol. 36, no. 2, pp. 316–318, Jun. 2012.

[7] R. S. McGinnis *et al.*, "A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0178366.

[8] A. Paraschiv-Ionescu, A. Soltani, and K. Aminian, "Real-world speed estimation using single trunk IMU: Methodological challenges for impaired gait patterns," in *Proc. 42nd Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2020, pp. 4596–4599.

[9] M. H. Pham *et al.*, "Validation of a step detection algorithm during straight walking and turning in patients with Parkinson's disease and older adults using an inertial measurement unit at the lower back," *Frontiers Neurol.*, vol. 8, p. 457, Sep. 2017.

[10] S. D. Din, A. Godfrey, and L. Rochester, "Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and Parkinson's disease: Toward clinical and at home use," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 3, pp. 838–847, May 2016.

[11] S. H. Shin and C. G. Park, "Adaptive step length estimation algorithm using optimal parameters and movement status awareness," *Med. Eng. Phys.*, vol. 33, no. 9, pp. 1064–1071, Nov. 2011.

[12] Q. Zhao, B. Zhang, J. Wang, W. Feng, W. Jia, and M. Sun, "Improved method of step length estimation based on inverted pendulum model," *Int. J. Distrib. Sens. Netw.*, vol. 13, no. 4, 2017, Art. no. 1550147717702914.

[13] W. Zijlstra and A. L. Hof, "Displacement of the pelvis during human walking: Experimental data and model predictions," *Gait Posture*, vol. 6, no. 3, pp. 249–262, 1997.

[14] A. Soltani, H. Dejnabadi, M. Savary, and K. Aminian, "Real-world gait speed estimation using wrist sensor: A personalized approach," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 3, pp. 658–668, Mar. 2020.

[15] S. Yang and Q. Li, "Inertial sensor-based methods in walking speed estimation: A systematic review," *Sensors*, vol. 12, no. 5, pp. 6102–6116, 2012.

[16] A. Weiss *et al.*, "Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-day accelerometer recordings," *Neurorehabil. Neural Repair*, vol. 27, no. 8, pp. 742–752, 2013.

[17] A. Paraschiv-Ionescu, C. J. Newman, L. Carcreff, C. N. Gerber, S. Armand, and K. Aminian, "Locomotion and cadence detection using a single trunk-fixed accelerometer: Validity for children with cerebral palsy in daily life-like conditions," *J. NeuroEng. Rehabil.*, vol. 16, no. 1, p. 24, Dec. 2019.

[18] H.-K. Lee *et al.*, "Computational methods to detect step events for normal and pathological gait evaluation using accelerometer," *Electron. Lett.*, vol. 46, no. 17, pp. 1185–1187, 2010.

[19] B. Fasel *et al.*, "A wrist sensor and algorithm to determine instantaneous walking cadence and speed in daily life walking," *Med. Biol. Eng. Comput.*, vol. 55, no. 10, pp. 1773–1785, Oct. 2017.

[20] H. Weinberg, "Using the ADXL202 in pedometer and personal navigation applications," *Analog Devices AN-602 Appl. Note*, vol. 2, no. 2, pp. 1–6, 2002.

[21] I. Bylemans, M. Weyn, and M. Klepal, "Mobile phone-based displacement estimation for opportunistic localisation systems," in *Proc. 3rd Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol.*, 2009, pp. 113–118.

[22] L. E. Díez, A. Bahillo, J. Otegui, and T. Otim, "Step length estimation methods based on inertial sensors: A review," *IEEE Sensors J.*, vol. 18, no. 17, pp. 6908–6926, Sep. 2018.

[23] K. Aminian, P. Robert, E. Jéquier, and Y. Schutz, "Incline, speed, and distance assessment during unconstrained walking," *Med. Sci. sports exercise*, vol. 27, no. 2, pp. 226–234, 1995.

[24] S. Zihajehzadeh and E. J. Park, "A Gaussian process regression model for walking speed estimation using a head-worn IMU," in *Proc. 39th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Jul. 2017, pp. 2345–2348.

[25] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. B, Methodol.*, vol. 58, no. 1, pp. 267–288, 1996.

[26] N. Paker, D. Bugdayci, G. Goksenoglu, D. T. Demircioğlu, N. Kesiktas, and N. Ince, "Gait speed and related factors in Parkinson's disease," *J. Phys. Therapy Sci.*, vol. 27, no. 12, pp. 3675–3679, 2015.

[27] M. S. Bryant *et al.*, "Gait variability in Parkinson's disease: Influence of walking speed and dopaminergic treatment," *Neurol. Res.*, vol. 33, no. 9, pp. 959–964, Nov. 2011.

[28] A. Delval *et al.*, "Role of hypokinesia and bradykinesia in gait disturbances in Huntington's disease," *J. Neurol.*, vol. 253, no. 1, pp. 73–80, Jan. 2006.

[29] A. Salarian, P. R. Burkhard, F. J. G. Vingerhoets, B. M. Jolles, and K. Aminian, "A novel approach to reducing number of sensing units for wearable gait analysis systems," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 72–77, Jan. 2013.

[30] K. Aminian, B. Najafi, C. Büla, P. F. Leyvraz, and P. Robert, "Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes," *J. Biomech.*, vol. 35, no. 5, pp. 689–699, 2002.

[31] M. El-Gohary *et al.*, "Continuous monitoring of turning in patients with movement disability," *Sensors*, vol. 14, no. 1, pp. 356–369, 2014.

[32] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Jul. 2011, pp. 1–7.

[33] J. W. Kim, H. J. Jang, D.-H. Hwang, and C. Park, "A step, stride and heading determination for the pedestrian navigation system," *Positioning*, vol. 3, nos. 1–2, pp. 273–279, 2004.

[34] J.-G. Park, A. Patel, D. Curtis, S. Teller, and J. Ledlie, "Online pose classification and walking speed estimation using handheld devices," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012 pp. 113–122.

[35] R. W. Bohannon and S. S. Glenney, "Minimal clinically important difference for change in comfortable gait speed of adults with pathology: A systematic review," *J. Eval. Clin. Pract.*, vol. 20, no. 4, pp. 295–300, Aug. 2014.