



HAL
open science

EVAL cane: Non-intrusive monitoring platform with a novel gait-based user identification scheme

Yuexiu Xing, Ting Wang, Fen Zhou, Aiqun Hu, Guyue Li, Linning Peng

► **To cite this version:**

Yuexiu Xing, Ting Wang, Fen Zhou, Aiqun Hu, Guyue Li, et al.. EVAL cane: Non-intrusive monitoring platform with a novel gait-based user identification scheme. *IEEE Transactions on Instrumentation and Measurement*, 2020, 70, pp.1-15. 10.1109/TIM.2020.3009338 . hal-02907607

HAL Id: hal-02907607

<https://hal.science/hal-02907607>

Submitted on 29 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EVAL cane: Non-intrusive monitoring platform with a novel gait-based user identification scheme

Yuexiu Xing, Ting Wang, Fen Zhou, Aiqun Hu, Guyue Li, and Linning Peng

Abstract—Considering the particularity of fragile people with reduced mobility, we design a non-intrusive platform called EVAL cane to assist and monitor the user’s walking. On the one hand, it has important walking assistance functions, such as obstacle warning and fall detection. On the other hand, it collects user’s long-term walking data in the background, which may be potential physical health assessment data. Compared with an ordinary cane, the above two functions are implemented without any burden on the user’s life. In addition, we have considered the authentication security for that the walking data is private. To this end, we propose a novel user identification scheme leveraging the user walking gait data collected by EVAL cane. To the best of our knowledge, it is the first time that the gait information collected by cane is used for reinforcing monitoring system security. This scheme does not require the user to remember and enter any identity (such as a password), which is user-friendly for fragile people. In the scheme, a statistics-based rough gait feature extraction method is put forward at first. Then, in order to improve the identification precision, we design a performance-based feature deletion (PFD) algorithm to remove bad features. Finally, a minimum Mahalanobis distance classifier is used. Experimental results show that the user identification rate without the PFD algorithm can reach as high as 90.48%. In addition, the PFD algorithm further improves the performance by about 6%, reaching an excellent result of 96.43%.

Index Terms—Monitoring system, EVAL cane, user identification/authentication, walking gait, rough feature extraction, feature refinement

I. INTRODUCTION

RECENTLY, the number of fragile people with reduced mobility, such as the elderly, the blind and the disabled,

This work was supported in part by Jiangsu Provincial Key Research and Development Program (BE2019109), Research Fund of National Mobile Communications Research Laboratory, Southeast University (No.2020B05), National Natural Science Foundation of China (61801115, 61941115), Purple Mountain Laboratories for Network and Communication Security, Campus France PHC Cai Yuanpei 2019 project (44016XA), and China Scholarship Council. (*Corresponding author: A. Hu and T. Wang.*)

Y. Xing is with the School of Information Science and Engineering, Southeast University, 210096 Nanjing, China, Institut Supérieur d’Electronique de Paris, Issy-les-Moulineaux, France, and University of Gustave Eiffel, CNRS, ESIEE Paris, Marne-la-Vallée, France. (e-mail: yxxing@seu.edu.cn, yuexiu.xing@isep.fr, yuexiu.xing@esiee.fr)

T. Wang is with the LIGM lab, University of Gustave Eiffel, CNRS, ESIEE Paris, Marne-la-Vallée, France. (email: ting.wang@esiee.fr)

F. Zhou is with the Institut Supérieur d’Electronique de Paris (ISEP), Issy-les-Moulineaux, France, and CERI-LIA, University of Avignon, Avignon, France (email: fen.zhou@isep, univ-avignon).fr)

A. Hu is with School of information Science and Engineering, Southeast University, and Purple Mountain Laboratories, 210096 Nanjing, China. (e-mail: aqhu@seu.edu.cn.)

G. Li and L. Peng are with School of Cyber Science and Engineering, Southeast University, and Purple Mountain Laboratories, 210096 Nanjing, China. (e-mail: {guyuelee, pengln}@seu.edu.cn.)

Digital Object Identifier xxx

is continuing to grow [1], [2]. The developing technologies that can help this group cohort live easily and reduce health risks are receiving increasing attention [3]–[5]. Therefore, many assistive devices that can also acquire physical and behavioral data have been investigated, whose data can be used by the users’ guardians and doctors as a medical knowledge database [6]. Among them, the sensor-based monitoring system has shown great potential. It combines a certain degree of intelligence to monitor and interpret human motion [7], [8]. In this research field, platform design and authentication security are two important aspects worth investigating.

A. Related work of monitoring platforms

Existing platforms of monitoring systems can be divided into two types: intrusive and non-intrusive. Intrusive monitoring systems are stand-alone devices that usually need to be carried or mounted on the users’ body. Lv *et al.* [9] developed a mobile health monitoring system, named iCare, which realized the real-time emergency response by some body-sensors, such as Bluetooth accelerometer and Bluetooth enabled blood pressure. Similarly, a health monitoring system with a finger-touched heartbeat sensor and a temperature sensor was designed in [10]. However, those systems that are independent of the user’s daily life may be uncomfortable and obstructive to the user. Non-intrusive monitoring systems are devices that are integrated into the user’s daily necessities, which collect data during their daily lives, such as smart watches, smart phones, smart glasses, or smart insoles. For example, the methods in [11]–[13] collected user data through sensors on a smart phone, which are typically placed in users’ pockets or hands. However, for the fragile people who need to use a cane everyday, such as the elderly and the blind, these electronic products may be unfamiliar and be forgotten to carry with [14], [15]. Obviously, a smart cane with monitoring functions is a better option for cane users [16]. For example, Saaid *et al.* [17] developed a smart cane to help users alert an impending obstacle with an ultrasonic sensor. Similarly, a cane with a multi-sensor obstacle detection system was proposed in [18] to warn obstacles exceeding the physical length of the cane. In addition, a smart cane with the functions of fall detection and location tracking was designed in [16]. However, these devices lack the function of collecting long-term behavior data, which is an important basis for analyzing users’ physical changes. Therefore, we implemented a multifunction smart cane that can monitor long-term walking data for these fragile people, named "human and environment evaluation (EVAL) cane".

B. Related work of authentication security

Since walking data involves user privacy, authentication as the first line of data security is necessary, which is essential to prevent intrusion and fraud [19]. Furthermore, user identification can help optimize the functions of obstacle warnings, fall detection, and gait analysis by intelligently learning and establishing tailor-made rules for different users. Considering the user-friendliness of fragile people, biometric is a reliable and effective method of user identification. Human biological characteristic can be classified into two types: physiological (such as fingerprints, iris and retina) and behavioral (such as voice, handwritten signatures, and gait) [20]–[22]. Gait is the way people walk, which varies from person to person and is hard to imitate [23]. Therefore, it can be a potential solution for user identification in monitoring systems [24], [25].

Gait recognition could be categorized into three groups according to data source [26]. First, the camera-based gait recognition technology extracts gait parameters from videos or photos [27], [28]. Its advantage is that the user is free from wires or sensors. However, videos and photos are likely to cause serious personal privacy issues. Second, the sensors network-based gait recognition method recognizes a walking person by a sensor network fixed on the wall or floor [29]. However, the high cost makes this method unsuitable for individual users. Third, the sensors-based gait recognition scheme implements gait classification through a set of sensors connected to the user's body or supplies [30]–[43]. This solution is low cost and suitable for individual users. As shown in Table I, sensors-based gait can be used to identify both the type of activity (the work with *) and the user identity. For activity recognition, Hou *et al.* [30] simulated radar echos to identify slow, normal and fast walking respectively. Si *et al.* [31] analyzed the recognition of standing, walking and jumping by force sensitive resistor (FSR). For the identity authentication, accelerometers installed on the body are commonly used. Rong *et al.* [32] and Gafurov *et al.* [33] achieved 6.7% and 1.6% EER by fixing the accelerometer to user's waist and ankle, respectively. The work in [34]–[37] researched the ZJU-GaitAcc dataset with different methods, such as the speed-adaptive gait cycle segmentation and individualized matching threshold generation techniques used in [35]. The combination of gyroscope and accelerometer is also effective for gait recognition. Trung *et al.* [38] collected the gait data of 744 subjects through the waist's IMUZ sensor (accelerometer and gyroscope) and smartphone (accelerometer). In addition, Sun *et al.* [39] conducted further research with the gait data of elderly subjects in the literature [38]. San-Segundo *et al.* [41] and Damaševičius *et al.* [40] researched the gait identification based on the accelerometer and gyroscope in a smart phone. In addition, other sensors are also worth studying, such as GPS in [42], piezoelectric energy harvester (PEH) and electromagnetic energy harvester (EEH) used in [43]. However, the data of these work are basically obtained by mounting the sensors on the human body, and there is a lack of research on gait identification based on a mobility-related aid like a cane.

C. Our contributions

In summary, the platform design and authentication security solutions of the monitoring system are still insufficient and worth studying. Our main contributions are the followings:

- We design a monitoring platform, EVAL cane, for fragile people with reduced mobility. As a daily assisting equipment, the EVAL cane is non-intrusive for users' life, which is equipped with various sensors, such as accelerometer, gyroscope, and FSR. On the one hand, the EVAL cane can help users walk, such as obstacle alerts and fall detection. On the other hand, it can collect users' long-term walking data, which is helpful for guardians and doctors to analyze the users' health changes.
- For the sake of data privacy and system security, we proposed a novel walking-gait-based user identification scheme. To the best of our knowledge, this is the first time that gait identification is used in a cane-based monitoring system authentication. Compared with sensors mounted on the human body, collecting data with sensors installed on a cane is an indirect way, which brings new challenges. To achieve the goal of user identification, we propose a rough gait feature extraction method. First, two walking datasets (in units of time and step, respectively) are obtained from a variety of collected sensor data after preprocessing. Then, various statistics of each sensor data in these two datasets are calculated as the user's rough gait feature. This method has a good scalability in both sensor type and statistical type expansion.
- To improve the precision of user identification, we propose further a performance-based feature deletion (PFD) algorithm. It can remove bad rough gait features according to the validation set in the training phase. Since there is some useless or disturbing information in the unsimplified raw sensor data and statistics, reasonable refinement of rough features is required. After PFD algorithm, the user identification performance is effectively improved. In addition, the feature dimension is reduced, which means a simplified computation.
- We verify the functions of the EVAL cane and the performance of the proposed user identification scheme. Except for the accurate obstacle alerts and fall detection, experimental results show that the EVAL cane can also effectively record users' long-term walking data. For the gait-based user identification, the IR without the PFD algorithm can be as high as 90.48%. Then, it was improved by about 6% through the PFD algorithm, reaching an excellent result of 96.43%. In addition, the PFD algorithm also reduced the EER from 16.83% to 10.48%

The remainder of this paper is organized as follows. Section II introduces the EVAL cane platform. Section III gives the system of the gait-based user identification scheme. Section IV and Section V present the rough gait feature extraction method and PFD algorithm, respectively. Section VI shows the experimental results and performance evaluation. Finally, Section VII concludes this paper.

Table I
RELATED WORK OF SENSOR-BASED GAIT RECOGNITION

Author	Devices and position	Sensor	Sampling rate	Method	Dataset	Results
* Hou <i>et al.</i> [30]	/	/	/	PC-IRNN	simulated radar echos (slow, normal and fast walking)	IR: 99.13%
* Si <i>et al.</i> [31]	FSR402 (shoes)	FSR	/	support vector machine and fractal analysis	one subject (standing, walking and jumping)	IR: 93.57%
Rong <i>et al.</i> [32]	MMA7260 (waist)	Accelerometer	250Hz	DTW algorithm	35 subjects (5 records each)	EER: 6.7%
Gafurov <i>et al.</i> [33]	MRS (ankle)		100Hz	cycle matching method	30 subjects (4 records each)	EER: 1.6%
Zhang <i>et al.</i> [34]	ADXL330 (upper arm, pelvis, wrist, thigh, ankle)			SP, sparse coding, CSCC	153 subjects (12 records each), 22 subjects (6 records each)	IR: 95.8% EER: 2.2%
Sun <i>et al.</i> [35]				speed-adaptive gait extraction, individualized threshold generation		IR: 96.9%
Zeng <i>et al.</i> [36]				RBF neural networks		IR: 96.2%
Qin <i>et al.</i> [37]				FRNN		IR: 98.4%
Trung <i>et al.</i> [38]	IMUZ sensors, Smartphone (waist)	Accelerometer, Gyroscope	100Hz	four methods by Rong <i>et al.</i> [44], Gafurov <i>et al.</i> [33], Derawi <i>et al.</i> [45], and Trung <i>et al.</i> [46]		744 subjects (two level walk, a up-slope walk, a down-slope walk each)
Sun <i>et al.</i> [39]				gait template synthesis, arbitration-based score level fusion	64 subjects (two level walk sequences each)	IR: 96.7%
Damaševičius <i>et al.</i> [40]				MotionNode (front right hip)	random projections, PDF	14 subjects (12 activities, five trials each)
San-Segundo <i>et al.</i> [41]	Smartphone (waist)		50Hz	i-vector based approach	30 subjects (a total of 240 activity)	EER: 6.1%
Zhang <i>et al.</i> [42]	Smartphone (pant pocket)	Accelerometer, GPS	/	mutual information model, PCA, SVDD, and LSTM	8 subjects (250 seconds each) 2 subjects (200 seconds each)	IR: 98.4%
Xu <i>et al.</i> [43]	PEH, EEH (hands)	PEH, EEH	100Hz	KEH-Gait	20 subjects (600 seconds each)	IR: 95% (PEH), 89% (EEH)

Remarks: * represents the work of gait-based activity recognition, IR: identification rate, EER: equal error rate, PC-IRNN: parallel channel identity initialized recurrent neural network, DTW: dynamic time warping, MRS: Motion Recording Sensor, SP: signature points, CSCC: classifier for sparse-code collection, RBF: radial basis function, FRNN: fusion recurrent neural network, PDF: probability distribution function, PCA: principal component analysis, SVDD: support vector data description, LSTM: long short-term memory, KEH: kinetic energy harvesting.

II. EVAL CANE PLATFORM

In this section, we give the details of the EVAL cane platform.

A. Constitution of The EVAL Cane Platform

Our EVAL cane has two main tasks:

- While assisting the users in walking, collecting their daily walking data, such as acceleration, angular velocity and force acted on the handle.
- According to the collected walking data, extracting the gait features for user identification.

Considering the user-friendliness for fragile people using electronic products, these two tasks are preferably passive and complex operations can be avoided. For the first task, we mainly achieve it through a well-designed monitoring system platform with appropriate sensors, which will be introduced in this section. For the second task, we plan to realize it using gait identification technique, which will be explained in the following sections.

Fig. 1 is a unified modeling language (UML) class diagram, which shows the structure of the monitoring system on the EVAL cane. As it shows, the monitoring system has five inputs (blue box) and three outputs (red box), which are controlled by

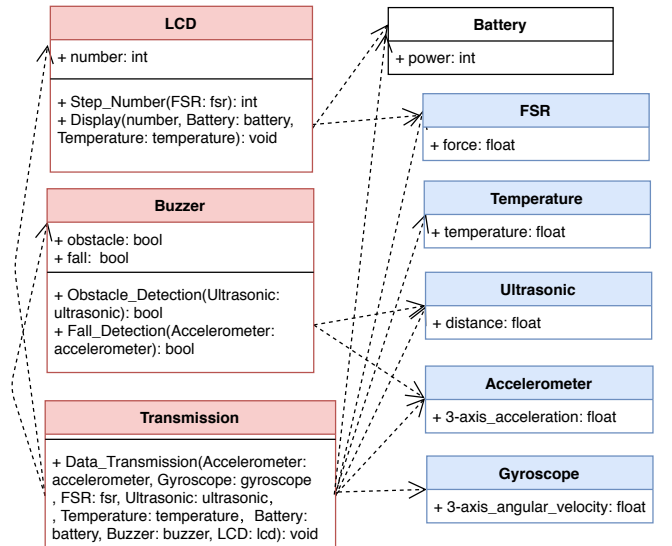


Figure 1. Structure of monitoring system on the EVAL cane.

a microcontroller, *i.e.*, Arduino UNO. The detailed hardware components and functions are illustrated in Table II. The five inputs correspond to the data of five sensors (FSR, temperature

Table II
COMPONENTS AND FUNCTIONS OF THE EVAL CANE

Components		Functions
MPU6050	Accelerometer	Measure 3-axis acceleration
	Gyroscope	Measure 3-axis angular velocity
	Temperature sensor	Measure the ambient temperature
HC-SR04 ultrasonic sensor		Measure the obstacle distance
FSR		Measure the force
Arduino UNO		Microcontroller
16x2 LCD Screen		Display data
Buzzer		Sound prompt

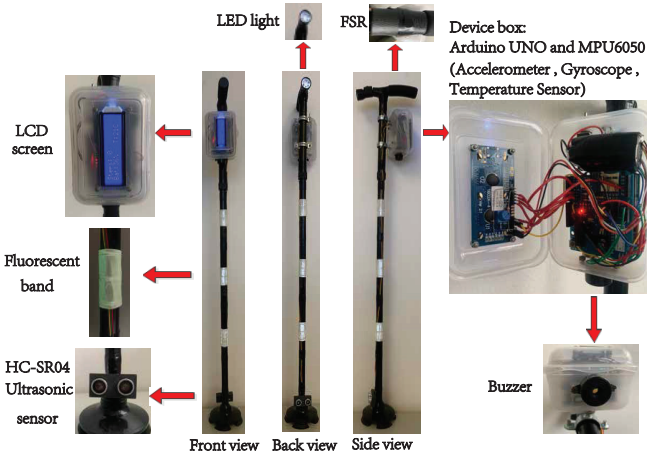


Figure 2. Picture of EVAL cane and its various function blocks.

sensor, ultrasonic sensor, accelerometer and gyroscope), *i.e.*, force, temperature, obstacle distance, 3-axis acceleration and 3-axis angular velocity. The three outputs are respectively liquid crystal display (LCD) screen, buzzer and data transmission module. The LCD screen can show temperature data, battery voltage, and step number obtained based on the analysis of force data. A buzzer is installed to alert an obstacle according to a configurable distance. For example, if the obstacle gets closer than 50 cm to the EVAL cane, the buzzer will start to buzz. As the distance decreases, the volume of the buzzer will gradually increase. Besides, the buzzer will also buzz when the cane falls, which is detected by analyzing the acceleration data. Finally, all sensor data is transmitted to the server.

B. Physical Display of The EVAL Cane Platform

The picture of the EVAL cane is illustrated in Fig. 2, which includes three views of the whole cane (a front view, a back view, and a side view), and detailed pictures of all components. Compared with an ordinary cane, almost all the components of the monitoring system are integrated into a device box weighing about 200 g on the pole. The important parts of a cane, such as handle and base, remain unchanged. It means that the monitoring system does not affect the basic functions of a cane.

The detailed pictures show the inputs, microcontroller, and outputs of the monitoring system. For the five input sensors,

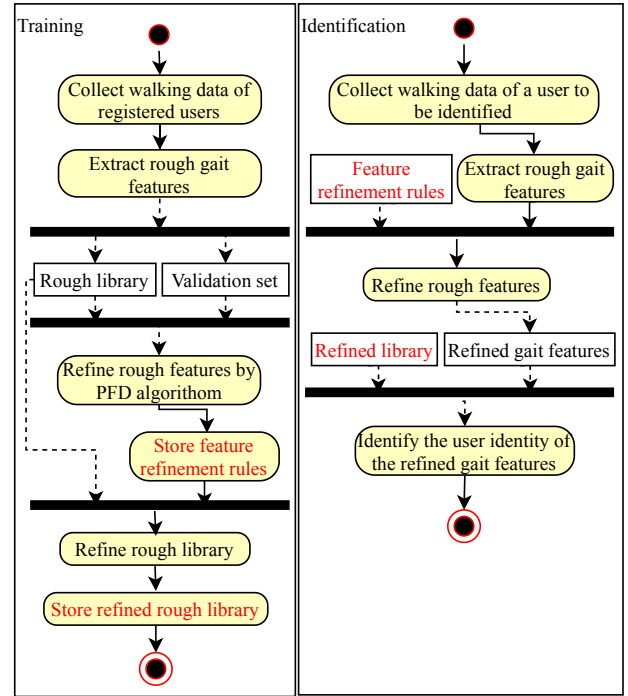


Figure 3. Gait-based user identification scheme.

it can be seen that the MPU6050 integrates the accelerometer, gyroscope and temperature sensor together. Then, it is placed in the device box and installed on the upper part of the cane (near the handle), which can reduce the damage caused by the inevitable collision during use. The FSR is mounted on the force-bearing area of the handle. When the user is walking with the cane, it can accurately detect the force exerted by the hand. The HC-SR04 ultrasonic sensor is fixed at the bottom of the cane with a protective case to accurately monitor the obstacles in front of the cane. For the microcontroller, the Arduino UNO is also equipped in the device box, which uses inter-integrated circuit (I2C) to communicate with the input sensors. For the three outputs, the LCD screen and its control circuit are mounted on the cover of the device box, which is convenient for viewing. The buzzer is installed on the bottom of the device box, which maintains the beauty without affecting the sound volume. The data transmission component is integrated on the microcontroller. Finally, in order to improve human-machine interaction, three fluorescent bands are fixed on the pole so that users can easily find the cane in a dark environment. In addition, it can remind others to avoid collisions. Furthermore, a light-emitting diode (LED) light is integrated on the handle to facilitate walking at night.

III. SYSTEM OF GAIT-BASED IDENTIFICATION SCHEME

Regarding the two tasks of the EVAL cane, the monitoring system platform introduced in Section II has realized the first task of collecting walking data while assisting walking. Next, we propose a gait-based user identification scheme for the second task to solve the authentication security problem in data transmission.

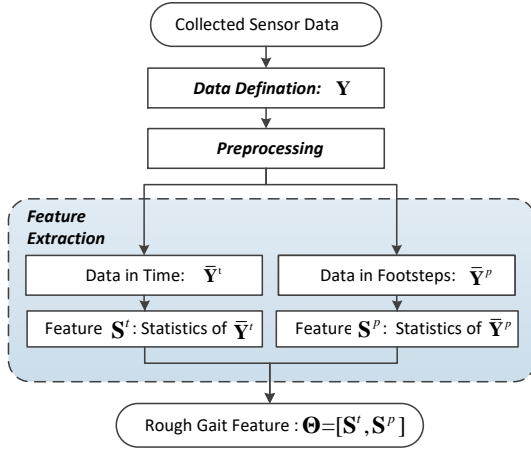


Figure 4. Flow chart of rough gait feature extraction.

Referring to the standard of UML activity diagram, Fig. 3 shows that the gait-based user identification includes two phases: training and identification, respectively [47], [48]. In the training phase, the input is the walking data with the identity label of the registered users. Then, a rough gait feature extraction method is applied and a rough library and a validation set are obtained. In addition, we propose a PFD algorithm to refine the rough gait feature by removing useless or interfering features. Next, the refinement rules of gait features are obtained and used to refine the rough library to the refined library. Finally, both the refinement rules and refined library will be stored. In the identification phase, the input is the walking data to be classified, *i.e.*, test samples. Then, the refined gait feature matrix is obtained after the rough feature extraction and feature refinement based on the refinement rules. Finally, the user identity of the input walking data will be worked out by comparing with the reference features in the refined library.

In this paper, the EVAL cane was instrumented with five sensors in all. Since the ultrasonic sensor (used for obstacle detection) and the temperature sensor have little relationship with the gait characteristics, only the data of accelerometer, gyroscope and FSR, are selected to do gait identification. The acceleration and angular velocity along X-axis, Y-axis and Z-axis are respectively noted as \mathbf{a}_x , \mathbf{a}_y , \mathbf{a}_z and \mathbf{r}_x , \mathbf{r}_y , \mathbf{r}_z . The force data collected by FSR is represented as \mathbf{f} . It is worth mentioning that, when the cane is in an upright position, X-axis is the opposite direction of the gravity and Y-axis is the forward direction of the walking. We will present the details about the rough gait feature extraction method and the PFD algorithm in the gait-based user identification scheme in the next two sections.

IV. ROUGH GAIT FEATURE EXTRACTION

This section mainly introduces the rough gait feature extraction method, whose flow chart is illustrated in Fig. 4. This method mainly includes three stages, namely data definition, data preprocessing and feature extraction. For ease of calculations, we first define the collected sensor data as a matrix. Then, in the preprocessing stage, denoising and effective data

selection are performed. In addition, the collected data is rearranged by footsteps to extract more gait features. Finally, various data statistics are calculated as rough gait features in the feature extraction stage. The details will be described as follows.

A. Data Definition

After data collection, the sensor data set is defined as a matrix \mathbf{Y} containing K column vectors as follows:

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K], \quad (1)$$

where a column vector \mathbf{y}_k , $k = 1, 2, \dots, K$, represents the collected time series data of a certain sensor. Then, considering all the sensor data used for gait identification \mathbf{Y} can be rewritten as

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K] \\ &= [\mathbf{a}_x, \mathbf{a}_y, \mathbf{a}_z, \mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z, \mathbf{f}], \end{aligned} \quad (2)$$

where $K = 7$ is the number of sensor data types.

In addition, in order to comprehensively extract gait information, the modulus of total acceleration \mathbf{a} and the total angular velocity \mathbf{r} are calculated based on the X-axis, Y-axis, and Z-axis data as follows:

$$\begin{aligned} \mathbf{a}(n) &= \sqrt{\mathbf{a}_x^2(n) + \mathbf{a}_y^2(n) + \mathbf{a}_z^2(n)}, \\ \mathbf{r}(n) &= \sqrt{\mathbf{r}_x^2(n) + \mathbf{r}_y^2(n) + \mathbf{r}_z^2(n)}, \\ 1 \leq n \leq N, \end{aligned} \quad (3)$$

where N is the length of each sensor data, and it is also the row number of \mathbf{Y} . Finally, the collected sensor data set can be defined as

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K] \\ &= [\mathbf{a}, \mathbf{r}, \mathbf{a}_x, \mathbf{a}_y, \mathbf{a}_z, \mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z, \mathbf{f}]. \end{aligned} \quad (4)$$

Now, the column number of \mathbf{Y} becomes 9, *i.e.* $K = 9$.

B. Preprocessing

Since the inevitable noise in recorded sensor data will affect the extraction of gait features, the denoising operation should be considered. Here, the moving average method is used to reduce noise [49].

$$\bar{\mathbf{y}}_k(n) = \frac{1}{\omega} \sum_{i=n}^{n+\omega} \mathbf{y}_k(i), \quad k = 1, 2, \dots, K, \quad (5)$$

where ω specifies the width of the box to be used to calculate the moving average results.

In addition, the EVAL cane's power switch needs to be turned on before walking and turned off after walking during the test. However, the time interval between turning on the cane and starting the walk is different for users. Therefore, the length of the noise data in their recorded data is different. As a result, the start and end of the user's walking have to be detected to pick effective data segments.

Considering that all types of sensor data are recorded simultaneously, they theoretically start and end at the same time. Therefore, it is sufficient to select only one type of

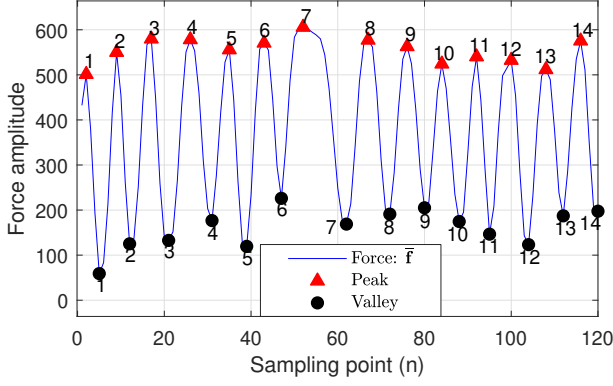


Figure 5. Peak and valley detection of denoised force data.

sensor data for start and end detection. During walking, the body's center of gravity moves from the supporting leg to the other. Meanwhile, the user's force on the cane's handle will change from maximum to minimum. In other words, force data is periodic and it effectively reflects the user's walking movements. The periodicity of the denoised force data \bar{f} is shown in Fig. 5, which proves the rationality of the theoretical analysis. Therefore, force data is selected in this paper to detect the start and end of the walking.

The two adjacent peaks/valleys in Fig. 5 represent a walking cycle. Each cycle indicates that the user has taken two steps. Each peak or valley represents one step. Then, all peaks and valleys of \bar{f} can be detected by MATLAB functions as

$$[pks(n), locs(n)] = Findpeaks(\bar{f}), 1 \leq n \leq N^p. \quad (6)$$

where $Findpeaks$ is the function of detecting peaks and valleys, pks and $locs$ represents the value and index of peaks/valleys in \bar{f} , respectively, N^p is the length of pks . Next, we chose the first peak and last valley as the start and the end of the walking, respectively. As a result, the effective data segment matrix of the sensor data described in (4) can be expressed as

$$\begin{aligned} \bar{Y}^t &= [\bar{y}_1^t, \bar{y}_2^t, \dots, \bar{y}_{K^t}^t], \\ &= [\bar{a}^t, \bar{r}^t, \bar{a}_x^t, \bar{a}_y^t, \bar{a}_z^t, \bar{r}_x^t, \bar{r}_y^t, \bar{r}_z^t, \bar{f}^t]. \end{aligned} \quad (7)$$

where $\bar{y}_i^t, i = 1, 2, \dots, K^t$ represents the sensor data after effective data segment selection, $K^t = K = 9$ is the number of sensor data types in \bar{Y}^t .

In addition, since walking is a periodic movement that occurs with stepping cycles, we further consider rearranging the data by footsteps, so that gait features can be extracted from another dimension. During the start and end detection above, the information of all peaks and valleys in \bar{f} , $pks(n), locs(n), n = 1, 2, \dots, N^p$ in (6), has already been obtained. Thus, the data in footsteps, \bar{Y}^p , can be easily obtained by extracting the sampling points corresponding to all the peaks and valleys in data \bar{Y}^t . However, since the data in footsteps lost the speed information, we calculate the length of step periods to describe the walking speed as

$$v^p(n) = locs(n) - locs(n-1), 1 \leq n \leq N^p, \quad (8)$$

where $locs(0)$ is defined as zero. Finally, the sensor data matrix in footsteps is obtained as

$$\begin{aligned} \bar{Y}^p &= [\bar{y}_1^p, \bar{y}_2^p, \dots, \bar{y}_{K^p}^p], \\ &= [\bar{a}^p, \bar{r}^p, \bar{a}_x^p, \bar{a}_y^p, \bar{a}_z^p, \bar{r}_x^p, \bar{r}_y^p, \bar{r}_z^p, \bar{f}^p, \bar{v}^p]. \end{aligned} \quad (9)$$

where $K^p = K^t + 1 = 10$.

C. Feature Extraction

Statistics are indispensable for data analyzing and testing in statistical theory [50]. Thus, a common task in identity authentication is exploiting the statistics of data as the identity features [51], [52]. In theory, every step of the users can reflect their walking habits. However, on the one hand, the gait information contained in the single-step data is insufficient to represent the user's identity. The change process in multi-step is also an important part of the gait characteristics. On the other hand, single-step walking data is susceptible to interference by random factors, such as pauses. It is reasonable to extract stable features from a period of walking data.

In this paper, seven statistics of data set \bar{Y}^t and \bar{Y}^p are extracted as gait features. For \bar{Y}^t , the statistics vectors are maximum(α^t), minimum(β^t), range(γ^t), mean(μ^t), variance(σ^t), skewness(η^t) and kurtosis(ζ^t) [53]. Thus, the statistics gait feature matrix of \bar{Y}^t , i.e. S^t is

$$\begin{aligned} S^t &= [s_1^t, s_2^t, \dots, s_i^t, \dots, s_{K^s}^t]^T, \\ &= [\alpha^t, \beta^t, \gamma^t, \mu^t, \sigma^t, \eta^t, \zeta^t]^T, \end{aligned} \quad (10)$$

where $K^s = 7$ is the number of statistics types. Obviously, the statistics gait feature matrix of \bar{Y}^p , i.e. S^p , can be obtained by the same operation as

$$\begin{aligned} S^p &= [s_1^p, s_2^p, \dots, s_i^p, \dots, s_{K^s}^p]^T, \\ &= [\alpha^p, \beta^p, \gamma^p, \mu^p, \sigma^p, \eta^p, \zeta^p]^T. \end{aligned} \quad (11)$$

Finally, the rough gait feature matrix can be expressed as

$$\Theta = [S^t, S^p]. \quad (12)$$

Its row number is the number of statistic types, i.e. $K^s = 7$. The number of columns is the sum of the sensor data types in S^t and S^p , i.e. $K^t + K^p = 19$. Therefore, the number of gait features in the Θ is $N^r = 133$.

It can be seen that changing the number of sensors types K^t, K^p and the number of statistical types K^s will not affect the above feature extraction process. Therefore, the sensors and statistics can be flexibly selected in practical applications, and the scalability of the rough gait feature extraction technique is evidenced.

V. GAIT FEATURE REFINEMENT USING PFD ALGORITHM

The rough gait feature matrix Θ is obtained through the rough gait feature extraction method. However, the types of sensor data and statistics have not been filtered. Thus, there may be some useless or interfering features in Θ , which have negative impact on user identification. In light of this, we propose a PFD algorithm to effectively refine Θ , which mainly includes two parts, single feature identification and gait

Algorithm 1 PFD Algorithm

Inputs:

Rough library Ω_l and its sample number N_l ;
 Rough validation set Ω_v and its sample number N_v ;
 Feature number of a rough gait feature matrix, N^r ;
 Threshold of feature selection, $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_{N_\lambda}]$.

Outputs:

Refinement rules $I(M)$;

```

1: Initialize the IR  $\omega = []$ ;
   {Single feature identification}
2: for  $k = 1 : N^r$  do
3:   Initialize correct identification number  $N_c(k) = 0$ ;
4:   for  $i = 1 : N_v$  do
5:     Normalize validation sample  $\Omega_v(\Theta_i)$  and  $\Omega_l$ .
6:      $\Omega_l(\Theta_c) = \arg \min_{j=1:N_l} MH(\Omega_l(\Theta_j|k), \Omega_v(\Theta_i|k))$ ;
7:     if  $U(\Omega_l(\Theta_c)) = U(\Omega_v(\Theta_i))$  then
8:        $N_c(k) = N_c(k) + 1$ 
9:     end if
10:  end for
11:   $\omega(k) = N_c(k)/N_v$ 
12: end for
   {Gait feature refinement}
13: for  $m = 1 : N_\lambda$  do
14:   Index of refined feature matrix  $I(m) = (\omega > \lambda_m)$ ;
15:   Refine feature matrix,  $\Theta^r = \Theta(I(m))$ , to obtain refined
       library  $\Omega_l^r$  and validation set  $\Omega_v^r$ ;
16:   Initialize correct identification number  $N_c(m) = 0$ ;
17:   for  $i = 1 : N_v$  do
18:     Normalize validation sample  $\Omega_v^r(\Theta_i^r)$  and  $\Omega_l^r$ .
19:      $\Omega_l^r(\Theta_c^r) = \arg \min_{j=1:N_l} MH(\Omega_l^r(\Theta_j^r), \Omega_v^r(\Theta_i^r))$ ;
20:     if  $U(\Omega_l^r(\Theta_c^r)) = U(\Omega_v^r(\Theta_i^r))$  then
21:        $N_c(m) = N_c(m) + 1$ 
22:     end if
23:   end for
24:    $\omega(m) = N_c(m)/N_v$ 
25: end for
26: Calculate  $M = \arg \max_{m=1:N_\lambda} (\omega(m))$ ;
27: return Refinement rules: Index matrix  $I(M)$ ;

```

feature refinement. Firstly, in the training phase, some rough features with low recognition ability are removed based on the performance threshold and the IR of the validation set. Then, the best threshold can be found according to the performance of the refined feature matrix. Thus, the refinement rules are obtained, which can be used in the identification phase directly. Algorithm 1 details the entire process of this algorithm.

A. Single Feature Identification

In order to identify useless features, the PFD algorithm first calculates the identification performance of each single feature in the rough feature matrix Θ . As shown in Fig. 3, rough gait features with user identity labels can be divided into rough library Ω_l (sample number N_l) and rough validation set Ω_v (sample number N_v) in the training phase. Next, the IR of

the validation set corresponding to each single feature in the rough feature matrix Θ is calculated.

As shown in line 5 of Algorithm 1, we first normalize the i -th validation sample $\Omega_v(\Theta_i)$, $i = 1, 2, \dots, N_v$ and the rough library Ω_l . Then, considering that gait features extracted from data of various sensors are related and have different dimensions, linear classifiers (such as Euclidean metrics) have limitations. A minimum Mahalanobis distance classifier $MH(\cdot, \cdot)$ is used, which can automatically account for the scaling of the coordinate axes, as well as for the correlation between the different features [54]–[56].

$$\Omega_l(\Theta_c) = \arg \min_{j=1:N_l} MH(\Omega_l(\Theta_j|k), \Omega_v(\Theta_i|k)), \quad (13)$$

$$k = 1, 2, \dots, N^r,$$

where k is the index of the selected single feature in the rough feature matrix, $\Omega_l(\Theta_c)$ is the closest sample in the rough library to the validation sample $\Omega_v(\Theta_i)$.

The condition for correct identification is:

$$U(\Omega_l(\Theta_c)) = U(\Omega_v(\Theta_i)), \quad (14)$$

where $U(\cdot)$ represents the user identity label corresponding to a feature matrix. Then the IR of every single feature is calculated in line 11:

$$\omega(k) = \frac{N_c(k)}{N_v}, k = 1, 2, \dots, N^r, \quad (15)$$

where $N_c(k)$ is the number of correctly identified samples corresponding to the k -th feature.

B. Gait Feature Refinement

After obtaining the IR of every single feature, we first remove useless features by setting a set of feature selection threshold $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_{N_\lambda}]$, where N_λ is the threshold number. Taking m -th threshold as an example, the index of the corresponding features with IR higher than λ_m can be calculated as

$$I(m) = (\omega > \lambda_m). \quad (16)$$

Thus, each rough feature matrix can be refined as Θ^r . Then, the refined library Ω_l^r and validation set Ω_v^r are obtained. Next, we calculate the IR of the refined feature matrix Θ^r corresponding to the threshold λ_m as given in line 19 of Algorithm 1,

$$\Omega_l^r(\Theta_c^r) = \arg \min_{j=1:N_l} MH(\Omega_l^r(\Theta_j^r), \Omega_v^r(\Theta_i^r)). \quad (17)$$

After the checking and statistical process of lines 20 and 21 in Algorithm 1, the final IR $\omega(m)$ corresponding to λ_m is obtained in lines 24. Finally, the optimal feature refinement index matrix, *i.e.* refinement rules, can be obtained as $I(M)$, where

$$M = \arg \max_{m=1:N_\lambda} (\omega(m)). \quad (18)$$

This algorithm is only implemented during the training phase. In the identification phase, the rough feature matrix of the test sample can be refined directly according to refinement rules $I(M)$.

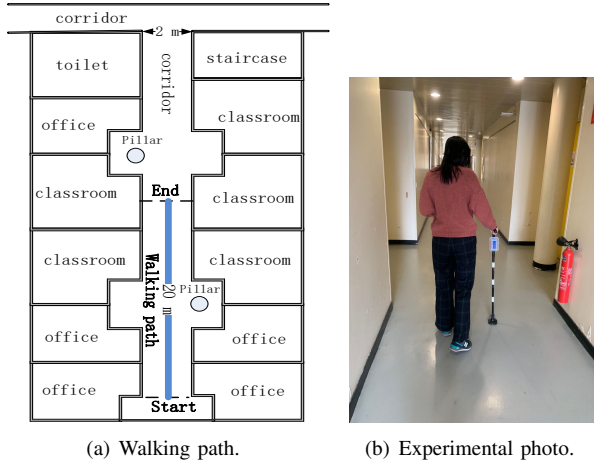


Figure 6. Walking scene.

Table III
INFORMATION OF EXPERIMENT OBJECTS

Subjects	Height	Weight	Age	Gender	Speed (m/s)
No. 1	178 cm	92 kg	22	Male	1.17 ± 0.30
No. 2	168 cm	54 kg	26	Male	1.06 ± 0.26
No. 3	165 cm	56 kg	28	Female	0.85 ± 0.40
No. 4	180 cm	90 kg	29	Male	0.99 ± 0.28
No. 5	160 cm	60 kg	37	Female	0.84 ± 0.34
No. 6	173 cm	74 kg	37	Male	1.10 ± 0.36
No. 7	183 cm	90 kg	60	Male	1.13 ± 0.15

VI. PERFORMANCE EVALUATIONS

A. Experimental Setup

An experiment of walking with EVAL cane was designed to verify the functions of EVAL cane and collect users' walking data for user identification. The walking scene of the experiment was a straight line of 20 m with a flat road surface, which was set in an empty corridor of the office building (the walking path and experimental photo are shown in Fig. 6). Seven subjects (5 male, 2 female; Age: 22 - 60) were asked to walk 15 times at their normal speed with the EVAL cane, whose detailed information are illustrated in Table III. The sensors' data are collected at a low sampling rate: about 6 Hz. The measuring range of accelerometer is ± 4 g and the maximum readable value of gyroscope is set to 250 degrees per second. In the gait-based user identification experiment, the training set (library and validation set) and the test set were randomly selected from 15 measurements of each subject. In addition, before formal experiment, each subject has been given 15 minutes to get familiar with walking with the EVAL cane.

B. EVAL Cane Platform

The functions of EVAL cane can be divided into two main categories: assisting walking and walking data collection.

- Assisting walking: basic information display, obstacle detection, and fall detection.

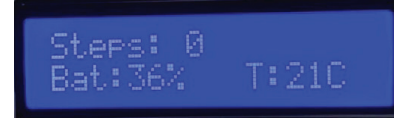


Figure 7. Basic information display on the LCD screen.

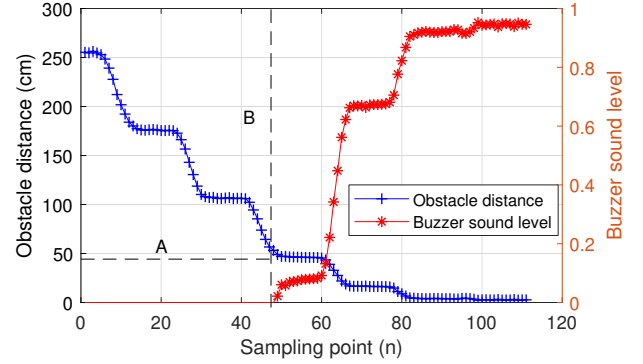


Figure 8. Obstacle detection.

The basic information, such as step number, battery power, and ambient temperature are displayed on the LCD screen. For example, Fig. 7 displays that the number of walking steps is 0, the remaining battery power is 36%, and the ambient temperature is 21 degrees centigrade.

In order to verify the function of obstacle detection, we experimented with the scenario that the user gradually walked to an obstacle with the EVAL cane. In addition, we set the obstacle reminder only when the obstacle distance is less than 50 cm. The buzzer sound lever is a function of the obstacle distance. However, this obstacle distance data detected by the cane is not linear. Therefore, we display the curve of obstacle distance and buzzer sound lever simultaneously with walking time (*i.e.*, sampling point) in Fig. 8. As shown by the left side of the dotted line B, the obstacle distance curve is stepped, gradually decreasing from about 250 cm to 50 cm. The buzzer remains muted at the same time, which matches the program settings. Then, as the distance of the obstacle continues to decrease (right of the dotted line B), the buzzer starts to sound, and the closer the distance, the louder the sound.

Besides, since the alarm or emergency calls after a fall is a basic function required by the user, it was also integrated into the EVAL cane. when the user falls, the cane will quickly fall to the ground. At this time, compared to normal walking, the 3-axis acceleration will change rapidly and greatly. As shown in Fig. 9, the X-axis acceleration increases rapidly from -1 to 4, and the Y-axis and Z-axis acceleration decrease from 0 to -4. Then, when the cane is re-centered, the acceleration data returns to normal. Therefore, we detect a user's fall by detecting a sudden change in 3-axis acceleration data and then issue a loud warning. In addition, we found that the duration of drastic acceleration change is only 2 sampling points (about 320 ms) during falling. It is difficult to help users in such a short time. However, before the drastic acceleration changes, there is a process during that the acceleration differs from its value of normal walking (lasts about 7 sampling points, *i.e.*

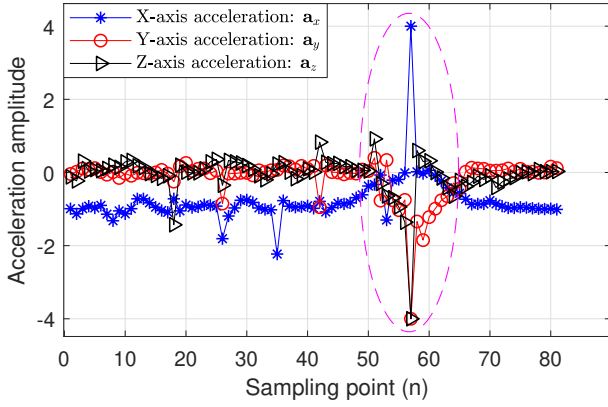


Figure 9. Fall detection.

1100 ms). As shown by the dotted circle in the Fig. 9, the acceleration of X-axis gradually changes from -1 to about 0, and that of the other two axis change from 0 to about -1. This process reflects the non-normal walking of the user, which will be researched in our future work.

- Walking data collection

Regarding the walking data collection, we mainly recorded the data of accelerometer, gyroscope, and FSR, which can well reflect the user's walking characteristics. Taking the data set of subject No 1 as an example, Fig. 10 shows the collected acceleration, angular velocity and force data.

Fig. 10(a) displays the 3-axis acceleration data. It can be seen that the X-axis acceleration is concentrated around -1 without an obvious period. The main reason is that the X-axis is the opposite direction of gravity and the up/down movement of the cane is not obvious during walking. In contrast, the forward acceleration in the Y-axis caused by walking, and the left acceleration in the Z-axis caused by the natural swing of the body are prominent when walking with a cane. Therefore, the Y-axis and Z-axis acceleration are periodic data, which fluctuate around 0 with a range of about 0.5. Obvious peaks, such as the value of a_x at the 42nd sampling point and the value of a_y at the 87th sampling point, are caused by the accidental impact of the cane on the ground.

The 3-axis angular velocity data are exhibited in Fig. 10(b). Echoing the acceleration data, the Y-axis and Z-axis data have the consistent and obvious periodic changes. However, although the X-axis angular velocity can be considered as periodic data, it is not obvious. In addition, the amplitude of the Y-axis data is significantly larger than that of the other two directions.

Fig. 10(c) illustrates the clear cycle characteristics of the force data, and the cycle length is basically stable. Each cycle of force data represents that the user has walked two steps. Therefore, this data can be effectively used to monitor user's walking.

C. Rough Gait Feature Extraction

After walking data collection, the next task is gait-based user identification. This subsection mainly introduces the process and results of rough gait feature extraction method.

First of all, we denoise all the walking data by a moving average method. Taking force data as an example, the comparison of the data with and without moving average (\bar{f} and f) are illustrated in Fig. 11. It is obvious that some glitches in the peaks are eliminated, which is useful for step detection. Then, as described in Section IV, two walking data sets \bar{Y}^t and \bar{Y}^p in units of time and step can be obtained from the denoised data, respectively.

In order to find effective statistical gait features, the intuitive observation of the statistical distribution of each type of sensor data is necessary. Therefore, we calculate the mean of every user's multiple measurements for each sensor data in \bar{Y}^t and \bar{Y}^p , respectively. They are then used to explore the statistical distribution characteristics of the data.

According to the analysis results, the statistical distribution of the sensor data set in \bar{Y}^t can be roughly divided into three cases:

- First case: the difference between users is obvious, which means the low-order statistics of these kinds of sensor data have a large difference. For example, Fig. 12(a) shows the mean of the force data, \bar{f}^t , for each user. Although some data overlap, the force data of the seven subjects are relatively dispersed in amplitude. Therefore, the mean of the force data may be a potentially valid gait feature to be focused on.
- Second case: the mean data differs only between part users, but the combination of multiple statistics can effectively increase the distance among users. Take the mean angular velocity data \bar{r}^t as an example, Fig. 12(b) illustrates that only certain users have obvious differences in the amplitude, such as subjects No 1 and No 6. However, there is a significant difference in the data fluctuation, which means that walking instability helps user identification. In other words, the mean and variance of the data \bar{r}^t may be effective in gait recognition, which worth further research.
- Third case: the mean data of all users cannot be used to directly distinguish users. For instance, Fig. 12(c) exhibits the mean acceleration data, \bar{a}^t . Judging from the figure, the data of multiple users is overlapped, which is hard to distinguish different users simply by mean and variance. However, some of their higher-order statistics may differ, such as kurtosis, skewness [53], etc. Therefore, these types of data can also be retained as potential gait characteristics.

Similar to \bar{Y}^t , the statistical distribution of the sensor dataset in \bar{Y}^p also shows the same characteristics in the above three cases. The mean of footsteps data, \bar{f}^p , \bar{r}^p , \bar{a}^p , are illustrated in Fig. 13. Compared with Fig. 12(a), Fig. 12(b), and Fig. 12(c), the intuitive differences among users are more obvious. In other words, the gait feature in footstep data \bar{Y}^p is much easier to be extracted than that in time data \bar{Y}^t .

Based on the above statistical distribution analysis, we calculated 7 kinds of statistics for each type of sensor data as rough gait features (the S^t and S^p shown in equations (10) and (11)). Different statistics have different identification results. For the training set, Fig. 14 displays the examples

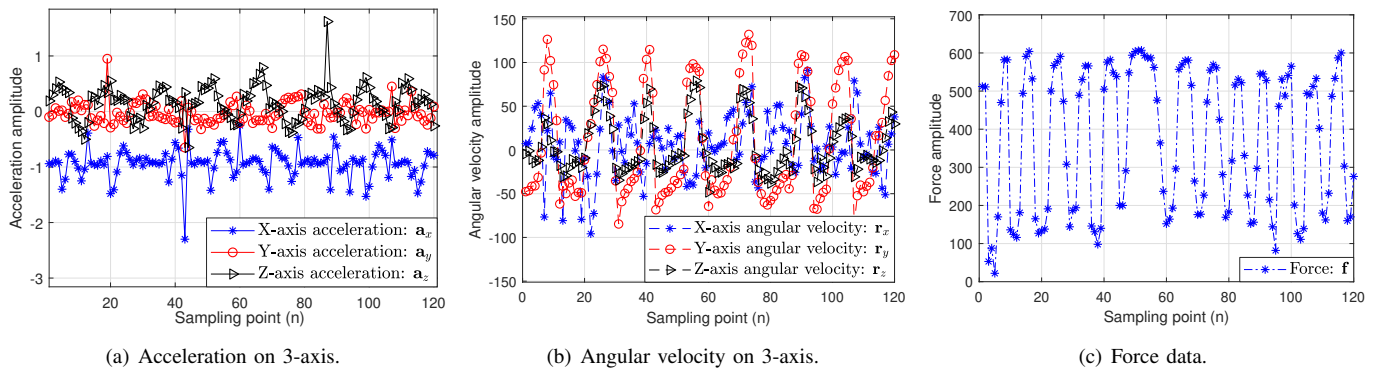


Figure 10. Collected acceleration, angular velocity, and force data.

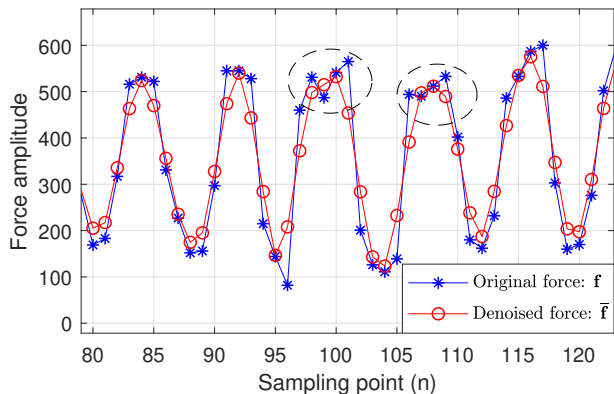


Figure 11. The comparison of force data with and without denoising (\mathbf{f} : original force data, $\bar{\mathbf{f}}$: denoised force data).

of clustering distributions of the identification results. Figure 14(a) illustrates the excellent clustering result achieved by the mean of force data $\bar{\mathbf{f}}^p$ and the range of Y-axis angular velocity $\bar{\mathbf{r}}_y^t$. There is very little data overlap, and these two statistics can effectively distinguish the most users. However, some other statistics do not demonstrate good performance results. As shown in Fig. 14(b), the variance of Y-axis angular velocity $\bar{\mathbf{r}}_y^p$ and the maximum of X-axis acceleration $\bar{\mathbf{a}}_x^p$ can only distinguish the subjects No 1 and No 6.

Finally, we verify the identification effect of each gait features in \mathbf{S}^t and \mathbf{S}^p . For the 10 measurements of training set, five of them were randomly selected as the library to identify the test set. Table IV shows that the identification rates for all single features in rough feature matrix \mathbf{S}^t and \mathbf{S}^p are not high. Among them, the highest identification result, *i.e.* the performance of the maximum of $\bar{\mathbf{f}}^p$, is only 61.90%. Next, we test the identification performance of the combined features (as shown by the yellow and blue areas in the Table IV). For the feature combination of multi-sensor data (yellow rows in the table), the bold italics data indicates that the identification rate has been improved compared with a single feature. For the multi-statistics feature matrix illustrated in the blue columns, all bold data represents that the identification accuracy has decreased. It is clear that the identification rate of some combined features is significantly higher than a single feature, while others are lower than a single feature. These

phenomena evidence that the feature combination is effective for improving the identification performance. However, there are some useless or interfering features, which cause a decrease in performance. Therefore, a feature refinement scheme is needed.

In addition, it can be seen that the identification rates of gait feature matrix \mathbf{S}^t and \mathbf{S}^p are 80.95% and 71.43%, respectively, while the performance of their combination is 85.71%. It proves that different gait features can be extracted from the datasets in units of time and step ($\bar{\mathbf{Y}}^t$ and $\bar{\mathbf{Y}}^p$). The two data sets are complementary in the gait identification scheme and lead to a higher identification performance.

D. Gait Feature Refinement with PFD Algorithm

For the feature refinement algorithm, we divide the training set into a library and a validation set. Then the rough gait feature matrix is refined based on the PFD algorithm in the training phase.

When the number of samples in the library and the validation set are both five, the identification curve of the validation set along the feature deletion threshold λ is shown in Fig. 15. The $\lambda = 0$ means that no feature refinement and all rough gait features are used. It is clear that as the threshold increases from 0 to 0.56 in steps of 0.02. The identification accuracy of the validation set gradually first increases from 0.8571 to 1 and keeps stable for a while. After that, it drops near to 0.52. In the interval of $0 \leq \lambda \leq 0.38$, as the threshold increases, features with low identification rates will be deleted, which means that their negative impact on effective features will also be eliminated. Therefore, the identification rate increases. When the threshold is too high ($\lambda \geq 0.4$), effective features are also deleted by mistake, so the identification performance drops sharply.

For different samples in the library, Table V illustrates a comparison of gait recognition performance with and without PFD algorithm, namely IR ω and IR ω^r . It can be found that as the number of samples in the library increases, the ω and ω^r both increase. However, the ω^r is always higher than ω in all cases, which proves the effectiveness of the PFD algorithm in the feature refinement. In addition, the overall trend of the performance improvement ($\Delta\omega$) is decreasing when the number of samples in the library increases. This is because

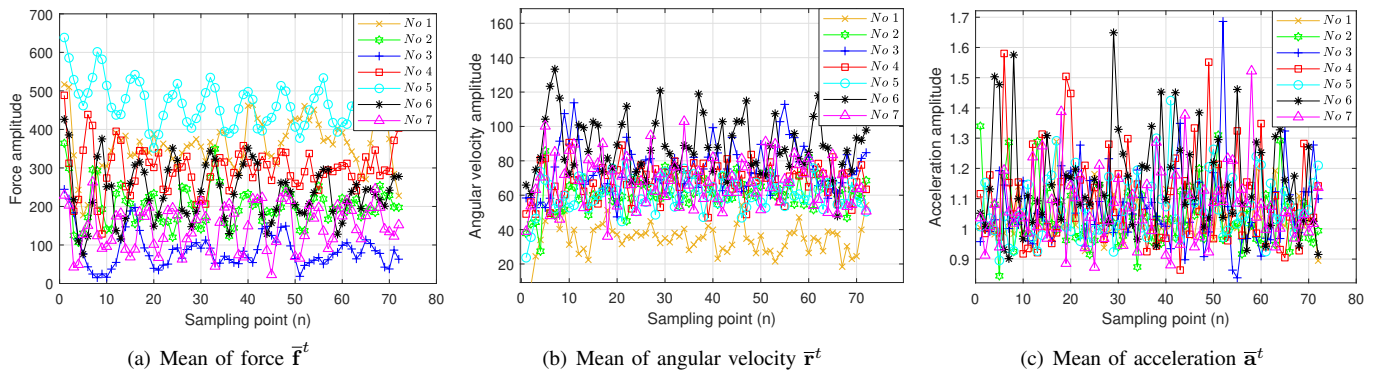


Figure 12. The mean of user's multiple measurements for three types of collected data in dataset \bar{Y}^t .

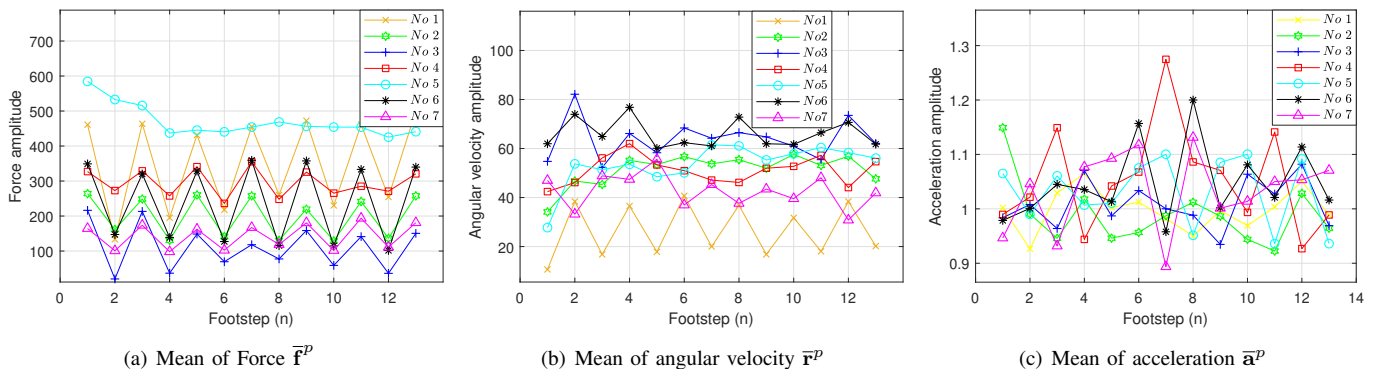


Figure 13. The mean of user's multiple measurements for three types of footstep data in dataset \bar{Y}^p .

the improvement space based on a high recognition rate is compressed. However, even the smallest 5.59% increase is very important for user identification.

E. Discussion

1). Comparison of experiment and data

In Table I, we summarize the state of the art research of gait-based user identification [30]–[43]. Compared with these studies, our work has the following two main challenges:

- The walking data sets studied in [30]–[43] were all collected by sensors mounted on the human body. In contrast, we installed the sensors on the cane for walking data collection. On the one hand, the cane is more friendly to the fragile people with reduced mobility. On the other hand, it is a challenge for gait recognition because the gait characteristics contained in the data collected by a cane are more blurred compared with the way of mounting sensors on the human body.
- The sampling rates of the studies in Table I are relatively high, such as 50 Hz, 100 Hz, and 250 Hz, while we used low sampling rate equipment (about 6 Hz) for data collection. The experiment in [43] has proved that the gait recognition performance of the acceleration signal will be severely degraded (about 30%) at a sampling rate (8 Hz) than at a higher sampling rate (100 Hz). It means a lower sampling rate may result in loss of some walking information. However, the low sampling rate leads to

low-cost equipment, which is very meaningful in actual product production. Our goal is to compensate for this deficiency in other ways to achieve the same performance as those methods under high sampling conditions.

In this paper, in order to overcome the above two challenges, we have taken the following measures. First, considering that all those different sensors in the researches presented in Table I show a good gait recognition ability (such as accelerometer in [32]–[37], accelerometer and gyroscope in [38]–[41], and FSR in [31]), we have integrated accelerometer, gyroscope, and FSR to portray the gait more comprehensively. Further, each subject was asked to walk a long distance (20 m) to collect more walking information for gait recognition, which can compensate for the information loss caused by the low sampling rate to a certain extent. In addition, in the gait feature extraction process, another data set in units of step is derived from the original data set in units of time. It can provide more information for gait feature extraction. Finally, we achieved a maximum IR of 96.43% and an EER of 10.48% in the experiment, which have reached the same excellent performance as other work presented in Table I (such as IR: 96.2% in [36] and 96.7% in [39], EER: 6.7% in [32] and 14.3% in [38]).

However, compared to [34], [38], [39], [41], our experimental data set is not large enough. Thus, in future work we will consider more subjects (cane users) with different disabilities in the experiment, such as people with different degrees and

Table IV
ROUGH GAIT FEATURE IDENTIFICATION

Rough gait feature		Maxmim α	Minimum β	Range γ	Mean μ	Variance σ	Skewness η	Kurtosis ς	Multi-statistics
Feature matrix \mathbf{S}^t	$\bar{\mathbf{a}}^t$	9.52%	19.05%	9.52%	19.05%	19.05%	19.05%	19.05%	14.52%
	$\bar{\mathbf{r}}^t$	47.62%	23.81%	33.33%	42.86%	38.10%	33.33%	33.33%	57.14%
	$\bar{\mathbf{a}}_x^t$	4.76%	9.52%	14.29%	19.05%	28.57%	19.05%	19.05%	19.05%
	$\bar{\mathbf{a}}_y^t$	9.52%	14.29%	23.81%	28.57%	14.29%	19.05%	19.05%	14.29%
	$\bar{\mathbf{a}}_z^t$	14.29%	19.05%	19.05%	47.62%	33.33%	19.05%	9.52%	47.62%
	$\bar{\mathbf{r}}_x^t$	23.81%	47.62%	19.05%	28.57%	42.86%	19.05%	4.76%	47.62%
	$\bar{\mathbf{r}}_y^t$	52.38%	42.86%	57.14%	19.05%	52.38%	28.57%	42.86%	61.90%
	$\bar{\mathbf{r}}_z^t$	33.33%	38.10%	28.57%	14.29%	47.62%	19.05%	14.29%	23.81%
	$\bar{\mathbf{f}}^t$	42.86%	14.29%	47.62%	47.62%	33.33%	42.86%	23.81%	57.14%
Multi-sensors	52.38%	33.33%	57.14%	52.38%	57.14%	33.33%	28.57%	80.95%	
Feature matrix \mathbf{S}^p	$\bar{\mathbf{a}}^p$	23.81%	23.81%	23.81%	42.86%	19.05%	14.29%	14.29%	19.05%
	$\bar{\mathbf{r}}^p$	28.57%	28.57%	14.29%	47.62%	23.81%	28.57%	9.52%	38.10%
	$\bar{\mathbf{a}}_x^p$	23.81%	33.33%	33.33%	14.29%	28.57%	4.76%	28.57%	23.81%
	$\bar{\mathbf{a}}_y^p$	23.81%	28.57%	23.81%	42.86%	42.86%	28.57%	9.52%	33.33%
	$\bar{\mathbf{a}}_z^p$	28.57%	9.52%	23.81%	42.86%	23.81%	19.05%	19.05%	38.10%
	$\bar{\mathbf{r}}_x^p$	14.29%	23.81%	33.33%	19.05%	28.57%	4.76%	28.57%	38.10%
	$\bar{\mathbf{r}}_y^p$	47.62%	28.57%	47.62%	28.57%	38.10%	28.57%	28.57%	40.20%
	$\bar{\mathbf{r}}_z^p$	42.86%	14.29%	28.57%	14.29%	33.33%	14.29%	23.81%	28.57%
	$\bar{\mathbf{f}}^p$	61.90%	14.29%	38.10%	28.57%	23.81%	19.05%	38.10%	47.62%
$\bar{\mathbf{v}}^p$	4.76%	28.57%	4.76%	23.81%	4.76%	4.76%	14.29%	9.52%	
Multi-sensors	61.90%	38.10%	47.62%	52.38%	28.57%	23.81%	28.57%	71.43%	
Feature matrix $\Theta = [\mathbf{S}^t, \mathbf{S}^p]$		/							85.71%

Table V

THE COMPARISON OF GAIT RECOGNITION PERFORMANCE WITH AND WITHOUT THE PFD ALGORITHM (WHERE ω INDICATES THE IR USING ALL ROUGH GAIT FEATURES, ω^r INDICATES THE IR AFTER FEATURE REFINEMENT BY PFD ALGORITHM).

Feature	Sample number of library				
	2	3	4	5	6
ω	71.43%	80.95%	82.14%	85.71%	90.48%
ω^r	85.71%	89.80%	92.86%	94.29%	96.43%
$\Delta\omega$	14.28%	8.85%	10.72%	8.58%	5.95%

causes of the walking disorder. In order to do that effectively, it is necessary to collaborate with some aging and disability organizations or retirement homes.

2). Comparison of models and performance

To further prove the effectiveness of our model, we compared the performance with the gait recognition method proposed by Gafurov *et al.* (the work [33] in Table I), which also has been applied in [38]. The cycle matching method in [33] includes four steps: pre-processing, detection cycles, matching cycle, and decision. The main idea of this method is to conduct cross-comparison between the two sets of cycles to find the best matching cycle pair. On the one hand, the motion cycle is key information that can not be ignored in the gait identification, which has been studied by many works [34], [35], [44]. On the other hand, this method has low complexity, which is consistent with our hardware performance in EVAL cane. Other

Table VI

IR COMPARISON BETWEEN OUR METHOD AND THE METHOD IN [33] (\mathbf{a}_x , \mathbf{a}_y , \mathbf{a}_z , \mathbf{r}_x , \mathbf{r}_y , \mathbf{r}_y , AND \mathbf{f} REPRESENT THE IR OF THE METHOD IN [33] WITH DIFFERENT SENSOR DATA. ω_r AND ω ARE THE IR OF OUR METHOD WITH AND WITHOUT PFD ALGORITHM, RESPECTIVELY)

Data	\mathbf{a}_x	\mathbf{a}_y	\mathbf{a}_z	\mathbf{r}_x	\mathbf{r}_y	\mathbf{r}_y	\mathbf{f}	ω	ω_r
IR/%	17.86	17.86	21.43	35.71	53.57	39.29	37.14	90.48	96.43

computationally intensive methods (such as neural networks) are not considered here for the sake of fast identification and authentication in our platform.

Then, we analyzed the collected 3-axis acceleration, 3-axis angular velocity, and force data based on this method. The sample number of each subject in the library is set as six. As shown in Table VI, the Z-axis acceleration (\mathbf{a}_z) is the best among the 3-axis accelerations, which matches the results of Gafurov *et al.* in [33] that the sideway acceleration exhibits the best performance (only a 3-axis accelerometer is used). On our dataset with 7 types of sensor data, this method performs the best on the Y-axis angular velocity (\mathbf{r}_y : 53.57%). However, regrettably, all these results are unsatisfactory. The main reason is that the sampling rate in [33] is 100 Hz. Therefore, it can collect more than 100 data points in one motion cycle, while our data has only 6 data points in one motion cycle, which means a lot of information is lost. Compared with it, our scheme's IR is 42.86% and 36.91% higher with and without the PDF algorithm, respectively.

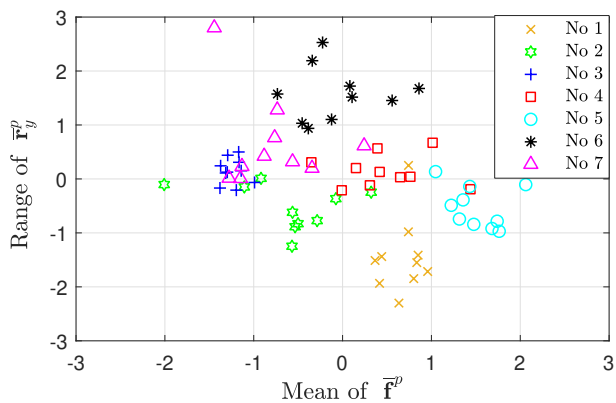
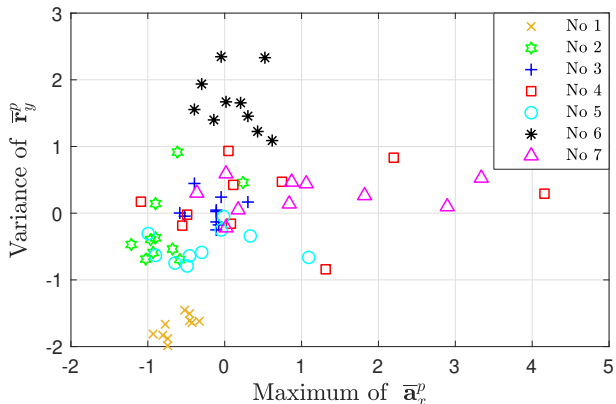
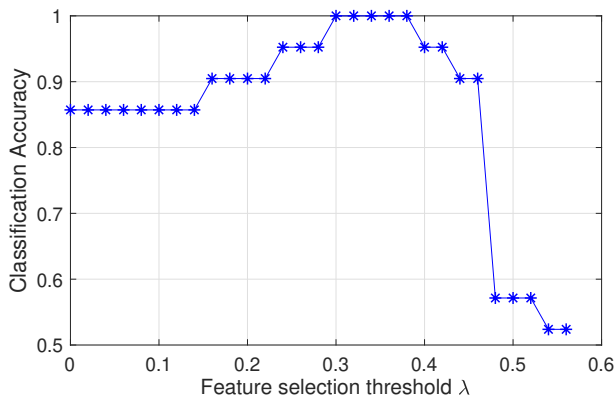
(a) Mean of \bar{f}^p and range of \bar{r}_y^p (b) Maximum of \bar{a}_x^p and variance of \bar{r}_y^p

Figure 14. Distribution map of all data under the selected two-dimensional statistical feature.

Figure 15. The identification curve of the validation set along the feature deletion threshold λ .

In addition, as shown in Fig. 16, we compared the performance of EER. The receiver operating characteristics (ROC) curve shows the trade-off between false rejection rate (FRR) and false acceptance rate (FAR) when the acceptance threshold changes in the personal authentication scenario. EER with equal FRR and FAR is also an effective indicator for performance evaluation. The lower the EER, the better the performance of the method. First, the model proposed by Gafurov *et al.* shows unsatisfactory EER performance on our

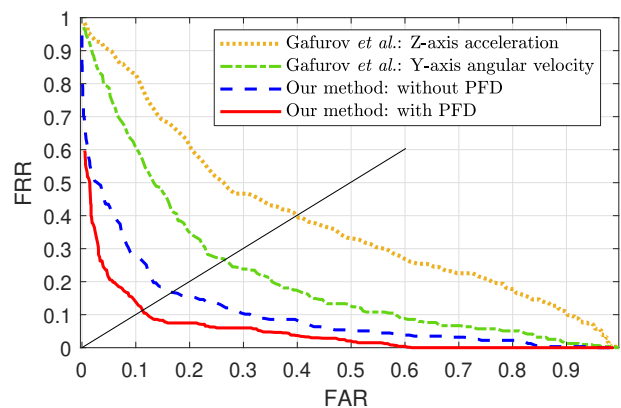


Figure 16. Comparison of ROC curves.

dataset (about 40% and 26% of \mathbf{a}_z and \mathbf{r}_y , respectively). In contrast, the ROC curves of our method (with and without the PFD algorithm) in Fig. 16 show a good performance of 10.48% and 16.83%, respectively. It also proves again the ability of our PFD algorithm in removing disturbing features.

VII. CONCLUSION

This paper investigated a cane-based monitoring system for the fragile people with reduced mobility. It has multiple functions such as obstacle detection, fall detection, and long-term walking data collection. The objective of this paper is to present a gait identification method based on this platform. In order to do that, the user's data during walking, such as acceleration, angular velocity, and force acting on the cane handle, were used. In addition, we addressed the problem of rough gait feature extraction by extracting seven kinds of statistics from each sensor dataset in units of time and step respectively. Further, in order to improve the identification precision, a feature refinement algorithm named PFD was proposed to remove the disturbing features from the rough gait feature matrix. Finally, the minimum Mahalanobis distance classifier was used to identify special users' gait. In the experiment, seven subjects (5 male, 2 female; Age: 22-60) were asked to walk 20 m for 15 times at their normal speed with the cane. The results show that an excellent identification accuracy of 96.43% can be achieved by the proposed scheme.

REFERENCES

- [1] N. Maestas, K. J. Mullen, and D. Powell, "The effect of population aging on economic growth, the labor force and productivity," National Bureau of Economic Research, Tech. Rep., 2016.
- [2] S. Majumder, T. Mondal, and M. J. Deen, "Wearable sensors for remote health monitoring," *Sensors*, vol. 17, p. 130, Jan. 2017.
- [3] S. Yusif, J. Soar, and A. Hafeez-Baig, "Older people, assistive technologies, and the barriers to adoption: A systematic review," *International journal of medical informatics*, vol. 94, pp. 112–116, Oct. 2016.
- [4] S. Malwade, S. S. Abdul, M. Uddin, A. A. Nursetyo, L. Fernandez-Luque, X. K. Zhu, L. Cilliers, C.-P. Wong, P. Bamidis, and Y.-C. J. Li, "Mobile and wearable technologies in healthcare for the ageing population," *Computer methods and programs in biomedicine*, vol. 161, pp. 233–237, Jul. 2018.
- [5] A. C. McLaughlin, M. Pryor, and J. Feng, "Design the technological society for an aging population," in *Critical Issues Impacting Science, Technology, Society (STS), and Our Future*. IGI Global, 2019, pp. 218–250.

- [6] E. Sardini, M. Serpelloni, and M. Lancini, "Wireless instrumented crutches for force and movement measurements for gait monitoring," *IEEE Trans. Instrum. Meas.*, vol. 64, pp. 3369–3379, Aug. 2015.
- [7] A. Sixsmith, N. Hine, I. Neild, N. Clarke, S. Brown, and P. Garner, "Monitoring the well-being of older people," *Topics in Geriatric Rehabilitation*, vol. 23, pp. 9–23, Jan. 2007.
- [8] B. Saggini, D. Scaccabarozzi, and M. Tarabini, "Metrological performances of a plantar pressure measurement system," *IEEE Trans. Instrum. Meas.*, vol. 62, pp. 766–776, Feb. 2013.
- [9] Z. Lv, F. Xia, G. Wu, L. Yao, and Z. Chen, "icare: A mobile health monitoring system for the elderly," in *Proc. IEEE/ACM Int'l Conf. Green Comput. Commun. Int'l Conf. Cyber, Physical Social Comput.*, Hangzhou, China, Dec. 2010, pp. 699–705.
- [10] A. A. Roman Richard, M. F. Sadman, U. H. Mim, I. Rahman, and M. S. Rahman Zishan, "Health monitoring system for elderly and disabled people," in *Proc. Int. Conf. Robotics, Electr. Signal Process. Tech. (ICREST)*, Dhaka, Bangladesh, Bangladesh, Jan. 2019, pp. 677–681.
- [11] D. Ben-Zeev, E. A. Scherer, R. Wang, H. Xie, and A. T. Campbell, "Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health," *Psychiatric rehabilitation journal*, vol. 38, p. 218, 2015.
- [12] M. O. Diab, R. A. Marak/Brome, M. Dichari, and B. Moslem, "The smartphone accessory heart rate monitor," in *Proc. Int. Conf. Comput. Medical Applications (ICCOMA)*, Dhaka, Bangladesh, Bangladesh, Jan. 2013, pp. 1–5.
- [13] G. Singh, D. Bansal, and S. Sofat, "A smartphone based technique to monitor driving behavior using DTW and crowdsensing," *Pervasive and Mobile Computing*, vol. 40, pp. 56–70, Sep. 2017.
- [14] F. Comunello, M. Fernández Ardèvol, S. Mulargia, and F. Belotti, "Women, youth and everything else: age-based and gendered stereotypes in relation to digital technology among elderly italian mobile phone users," *Media, Culture & Society*, vol. 39, pp. 798–815, Oct. 2016.
- [15] A. Berenguer, J. Goncalves, S. Hosio, D. Ferreira, T. Anagnostopoulos, and V. Kostakos, "Are smartphones ubiquitous?: An in-depth survey of smartphone adoption by seniors," *IEEE Consumer Electron. Mag.*, vol. 6, pp. 104–110, Jan. 2017.
- [16] J. Yang, Y. Liu, Y. Chen, H. Nie, Z. Wang, X. Liu, M. A. Imran, and W. Ahmad, "Assistive and monitoring multifunctional smart crutch for elderly," in *Proc. IEEE Int. Conf. Dependable, Autonomic Secure Comput. Int. Conf. Pervasive Intelligence Comput. Int. Conf. Cloud Big Data Comput., Int. Conf. Cyber Science and Tech. Congress (DASC/PiCom/CBDCCom/CyberSciTech)*, Fukuoka, Japan, Aug. 2019, pp. 397–401.
- [17] M. F. Saaid, A. M. Mohammad, and M. S. A. M. Ali, "Smart cane with range notification for blind people," in *Proc. IEEE Int. Conf. Automatic Control Intelligent Systems (IZCACIS)*, London, UK, Oct. 2016, pp. 225–229.
- [18] N. S. Ahmad, N. L. Boon, and P. Goh, "Multi-Sensor obstacle detection system via Model-Based State-Feedback control in smart cane design for the visually challenged," *IEEE Access*, vol. 6, pp. 64182–64192, Oct. 2018.
- [19] A. Ara, M. Al-Rodhaan, Y. Tian, and A. Al-Dhelaan, "A secure Privacy-Preserving data aggregation scheme based on bilinear elgamal cryptosystem for remote health monitoring systems," *IEEE Access*, vol. 5, pp. 12601–12617, Jun. 2017.
- [20] D. Gafurov, E. Snekkenes, and P. Bours, "Spoof attacks on gait authentication system," *IEEE Trans. Inf. Forensics Security*, vol. 2, pp. 491–502, Aug. 2007.
- [21] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S. . Makela, and H. A. Ailisto, "Identifying users of portable devices from gait pattern with accelerometers," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, vol. 2, Mar. 2005, pp. 973–976.
- [22] L. Chen and F. Tan, "Identity recognition scheme based on user access behavior," in *Proc. 2019 IEEE 8th Joint Int. Information Technology Artificial Intelligence Conference (ITAIC)*, Chongqing, China, Aug. 2019, pp. 125–129.
- [23] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, "A comprehensive study on Cross-View gait based human identification with deep CNNs," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 39, pp. 209–226, Feb. 2017.
- [24] M. Muaaz and R. Mayrhofer, "Smartphone-Based gait recognition: From authentication to imitation," *IEEE Trans. Mobile Comput.*, vol. 16, pp. 3209–3221, Nov. 2017.
- [25] H. J. Ailisto, M. Lindholm, J. Mantyjarvi, E. Vildjiounaite, and S.-M. Makela, "Identifying people from gait pattern with accelerometers," in *Proc. Biometric Tech. for Human Identification II*, Orlando, Florida, United States, Mar. 2005, pp. 7–14.
- [26] S. Rouzbeh and M. Babaei, "Human gait recognition using body measures and joint angles," *International Journal*, vol. 6, pp. 2305–2493, 2015.
- [27] J. P. Ferreira, M. M. Crisostomo, and A. P. Coimbra, "Human gait acquisition and characterization," *IEEE Trans. Instrum. Meas.*, vol. 58, pp. 2979–2988, May 2009.
- [28] D. Gafurov, "A survey of biometric gait recognition: Approaches, security and challenges," in *Proc. Annual Norwegian computer science conference*, 2007, pp. 19–21.
- [29] R. J. Orr and G. D. Abowd, "The smart floor: A mechanism for natural user identification and tracking," in *Proc. CHI '00 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, Apr. 2000, pp. 275–276.
- [30] M. Hou, X. Shi, and F. Zhou, "Pedestrian gait classification based on parallel channel identity initialized recurrent neural network," in *Proc. 2019 6th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Xiamen, China, Mar. 2019, pp. 1–5.
- [31] W. Si, G. Yang, X. Chen, and J. Jia, "Gait identification using fractal analysis and support vector machine," *Soft Computing*, vol. 23, pp. 9287–9297, Nov. 2019.
- [32] L. Rong, D. Zhiguo, Z. Jianzhong, and L. Ming, "Identification of individual walking patterns using gait acceleration," in *Proc. 1st Int. Conf. Bioinformatics Biomed. Eng.*, Wuhan, China, Jul. 2007, pp. 543–546.
- [33] D. Gafurov, E. Snekkenes, and P. Bours, "Improved gait recognition performance using cycle matching," in *Proc. 2010 IEEE 24th Int. Conf. Advanced Information Networking and Applications Workshops*, Perth, WA, Australia, Jun. 2010, pp. 836–841.
- [34] Y. Zhang, G. Pan, K. Jia, M. Lu, Y. Wang, and Z. Wu, "Accelerometer-based gait recognition by sparse representation of signature points with clusters," *IEEE Trans. Cybernetics*, vol. 45, pp. 1864–1875, Nov. 2015.
- [35] F. Sun, C. Mao, X. Fan, and Y. Li, "Accelerometer-based speed-adaptive gait authentication method for wearable IoT devices," *IEEE Internet Things J.*, vol. 6, pp. 820–830, Jul. 2019.
- [36] W. Zeng, J. Chen, C. Yuan, F. Liu, Q. Wang, and Y. Wang, "Accelerometer-based gait recognition via deterministic learning," in *Proc. 2018 Chinese Control And Decision Conference (CCDC)*, Shenyang, China, Jul. 2018, pp. 6280–6285.
- [37] Z. Qin, G. Huang, H. Xiong, Z. Qin, and K. R. Choo, "A fuzzy authentication system based on neural network learning and extreme value statistics," *IEEE Trans. Fuzzy Systems*, pp. 1–1, Dec. 2019.
- [38] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recognition*, vol. 47, pp. 228–237, Jan. 2014.
- [39] F. Sun, W. Zang, R. Gravina, G. Fortino, and Y. Li, "Gait-based identification for elderly users in wearable healthcare systems," *Information Fusion*, vol. 53, pp. 134–144, Jan. 2020.
- [40] R. Damaševičius, R. Maskeliūnas, A. Venčkauskas, and M. Woźniak, "Smartphone user identity verification using gait characteristics," *Symmetry*, vol. 8, no. 10, p. 100, Sep. 2016.
- [41] R. San-Segundo, J. D. Echeverry-Correa, C. Salamea-Palacios, S. L. Lutfi], and J. M. Pardo, "I-vector analysis for gait-based person identification using smartphone inertial signals," *Pervasive and Mobile Computing*, vol. 38, pp. 140–153, Jul. 2017.
- [42] H. Zhang, J. Liu, K. Li, H. Tan, and G. Wang, "Gait learning based authentication for intelligent things," *IEEE Trans. Veh. Technol.*, vol. 69, pp. 4450–4459, Mar. 2020.
- [43] W. Xu, G. Lan, Q. Lin, S. Khalifa, M. Hassan, N. Bergmann, and W. Hu, "KEH-Gait: Using kinetic energy harvesting for gait-based user authentication systems," *IEEE Trans. Mobile Comput.*, vol. 18, pp. 139–152, Apr. 2019.
- [44] L. Rong, Z. Jianzhong, L. Ming, and H. Xiangfeng, "A wearable acceleration sensor system for gait recognition," in *Proc. 2007 2nd IEEE Conf. Industrial Electronics and Applications*, Harbin, China, Sep. 2007, pp. 2654–2659.
- [45] M. O. Derawi, P. Bours, and K. Holien, "Improved cycle detection for accelerometer based gait authentication," in *Proc. 2010 Sixth Int. Conf. Intelligent Information Hiding and Multimedia Signal Processing*, Darmstadt, Germany, Nov. 2010, pp. 312–317.
- [46] N. T. Trung, Y. Makihara, H. Nagahara, R. Sagawa, Y. Mukaigawa, and Y. Yagi, "Phase registration in a gallery improving gait authentication," in *Proc. 2011 Int. Joint Conf. Biometrics (IJCB)*, Washington, DC, USA, Dec. 2011, pp. 1–7.
- [47] B. Danev, D. Zanetti, and S. Capkun, "On Physical-layer identification of wireless devices," *ACM Comput. Surv.*, vol. 45, pp. 6:1–6:29, Dec. 2012.

- [48] S. U. Rehman, K. W. Sowerby, and C. Coghill, "Radio-frequency fingerprinting for mitigating primary user emulation attack in low-end cognitive radios," *IET Commun.*, vol. 8, pp. 1274–1284, May 2014.
- [49] G. Aradhya, A. C. S. Rao, and M. D. Mastan Mohammed, "A novel hybrid approach for time series data forecasting using moving average filter and ARIMA-SVM," in *Proc. Emerging Tech. Data Mining Info. Security*, Singapore, 2019, pp. 369–381.
- [50] R. V. Hogg, J. McKean, and A. T. Craig, *Introduction to mathematical statistics*. Pearson Education, 2005.
- [51] F. M. Naini, J. Unnikrishnan, P. Thiran, and M. Vetterli, "Where you are is who you are: User identification by matching statistics," *IEEE Trans. Inf. Forensics Security.*, vol. 11, pp. 358–372, Feb. 2016.
- [52] C. K. Dubendorfer, B. W. Ramsey, and M. A. Temple, "An RF-DNA verification process for ZigBee networks," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Orlando, FL, USA, Oct. 2012, pp. 1–6.
- [53] R. J. Larsen and M. L. Marx, *An introduction to mathematical statistics and its applications*. Prentice Hall, 2001, vol. 5.
- [54] F. Babiloni, L. Bianchi, F. Semeraro, J. del R Millan, J. Mourino, A. Cattini, S. Salinari, M. G. Marciani, and F. Cincotti, "Mahalanobis distance-based classifiers are able to recognize EEG patterns by using few EEG electrodes," in *Proc. 23rd Annual Int. Conf. IEEE Engineering in Medicine and Biology Society*, Istanbul, Turkey, Nov. 2001, pp. 651–654.
- [55] P. M. Roth, M. Hirzer, M. Köstinger, C. Beleznai, and H. Bischof, *Mahalanobis Distance Learning for Person Re-identification*. London: Springer London, 2014, pp. 247–267.
- [56] S. Zhou, Q. Wang, Y. Fang, and Q. Liu, "An extreme learning machine method for multi-classification with mahalanobis distance," in *Proc. 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Xi'an, China, Sep. 2018, pp. 35–39.