



ELSEVIER

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication

Thanh Trung Ngo^{a,*}, Yasushi Makihara^a, Hajime Nagahara^b, Yasuhiro Mukaigawa^a,
Yasushi Yagi^a

^a The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

^b Faculty of Information Science and Electrical Engineering, Kyushu University, 744, Motoooka, Nishiku, Fukuoka 819-0395, Japan

ARTICLE INFO

Article history:

Received 6 February 2013

Received in revised form

19 June 2013

Accepted 26 June 2013

Keywords:

Gait-based owner authentication

Inertial sensor

Large-scale database

Performance evaluation

ABSTRACT

This paper presents the largest inertial sensor-based gait database in the world, which is made open to the research community, and its application to a statistically reliable performance evaluation for gait-based personal authentication. We construct several datasets for both accelerometer and gyroscope of three inertial measurement units and a smartphone around the waist of a subject, which include at most 744 subjects (389 males and 355 females) with ages ranging from 2 to 78 years. The database has several advantages: a large number of subjects with a balanced gender ratio, variations of sensor types, sensor locations, and ground slope conditions. Therefore, we can reliably analyze the dependence of gait authentication performance on a number of factors such as gender, age group, sensor type, ground condition, and sensor location. The results with the latest existing authentication methods provide several insights for these factors.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Recently, smart wearable or portable electronics are being developed so rapidly, and they are expected to be sophisticated enough in the future to be able to interact with the owner and understand his/her needs, intentions, actions [1,2], and health conditions [3,4]. Inertial sensor (accelerometer or gyroscope) is in fact increasingly being embedded in commercial portable electronic devices such as smartphones due to their high cost-performance, and inertial sensor-based owner assistance, such as user recognition, is an active research topic.

There are many existing methods to recognize a person who carries an inertial sensor based on gait such as identification by [5–9] or verification by [8,10–16], which showed the promising applications of inertial sensor for gait-based recognition.

Because the human gait is a periodic motion, a large number of gait recognition methods detect periods [5,6,9,12,13,15,17] for constructing gait patterns. Some researchers use frequency-domain features such as a histogram of signal intensity [8,10] or coefficients of Fourier transform [6,7]. Some researchers employ

model-based approaches [18,19], in which gait signal is modeled using a finite state machine.

However, in these studies, the authors evaluated their methods with databases limited in number of subjects, age variety, gender balance, sensor types, and ground conditions and hence the reliability of the performance evaluation is insufficient. In addition, it is difficult to compare these methods because they lack a common ground such as the same sensor configuration, gender ratio, number of subjects, and ground condition.

In this paper, the first contribution is to construct the largest inertial sensor-based gait database in the world to overcome such limitation. The advantages of the database includes (1) a very large number of subjects (744), (2) a balanced gender ratio, (3) a wide age range (2–78 years), (4) both acceleration and angular velocity data captured by three inertial sensors (containing both accelerometer and gyroscope) and a smartphone (containing an accelerometer), (5) a variation of sensor locations on subject's waist, and (6) 3 ground slope conditions (walking on a flat ground, walking up, and down a slope). The second contribution is to evaluate the state-of-the-art gait authentication algorithms in a more statistically reliable way and to reveal how gait authentication performance differs between genders and among age groups, sensor types, and sensor locations using the database. This paper is an extended version of our previous publication [20]. Here, we present an extended database with variations of sensor location, sensor type, and ground slope so that we can conduct more experiments with more comprehensive results.

* Corresponding author. Tel.: +81 6 6879 8422; fax: +81 6 6877 4375.

E-mail addresses: trung@am.sanken.osaka-u.ac.jp, trungbeo@gmail.com (T.T. Ngo), makihara@am.sanken.osaka-u.ac.jp (Y. Makihara), nagahara@limu.ait.yushu-u.ac.jp (H. Nagahara), mukaigaw@am.sanken.osaka-u.ac.jp (Y. Mukaigawa), yagi@am.sanken.osaka-u.ac.jp (Y. Yagi).

2. Related work

2.1. Existing inertial gait databases

There are many databases have been used in this field as summarized in Table 1, however they are publicly unavailable for the research community.

The largest existing database was used by Gafurov et al. [23] with 100 subjects, the gender is biased with a number of males is twice larger than that of females. This database was frequently used to evaluate the algorithms within the same research group [23–25,31]. Derawi et al. [15], Kobayashi et al. [7], Gafurov et al. [22,31], and Jenifer et al. [8] evaluated their methods with databases of 60, 58, 50, and 36 subjects, respectively, without age and gender information. Derawi et al. [27] again captured another dataset of 51 subjects with a very biased gender ratio and age distribution.

The largest database for smartphones was presented by Kobayashi et al. [7] with 58 subjects and about 40 sequences per subjects on different days. However, the frame-rate was as low as 33 Hz.

In overall, the numbers of subjects in these databases (100 subjects at most in [23]) are insufficient for statistically reliable performance evaluation. In addition, age variety is limited to adult and/or gender balance is biased.

2.2. Recognition methods

As stated above, since gait signal is periodic, there are three main approaches for gait recognition: period detection-based, frequency analysis-based, and gait model-based methods.

A large number of recognition systems [5,6,9–17] detect gait periods, each of which contains motion signal of both left and right steps, to construct gait patterns for recognition. A gait period can be detected from a gait signal sequence using heuristic information [5,12,13,15], or without heuristic information [16].

Some researchers use frequency-domain features such as a histogram of signal intensity [6,8,10], or coefficients of Fourier transform [6,7]. To obtain such frequency-domain features, we need a relatively long and stable gait signal sequence. However, a real gait signal is very noisy, temporally distorted and sensor orientation may change, and hence such a long stable gait signal is rarely available. As a result, frequency analysis-based techniques are outperformed by period detection-based methods in most situations [6,14,16]. Moreover, period detection-based methods can be applied for a single period of signal sequence (about 1 s), which implies potential real-time applications, while the frequency analysis-based methods cannot.

In the gait model-based methods [18,19], a gait model is generated using a finite state machine then some characteristics such as the gait period symmetry and homogeneity are used for gait pattern. However, the gait model-based methods require expert knowledge about the human gait to train the finite state machine so that it cannot deal well with the signal variation either.

In the evaluation of the gait authentication with our database, we select the four latest period detection-based authentication methods presented by four other research groups for evaluation with our database.

3. Gait database construction

3.1. Setup of gait capture system

To consider variations of sensor type and sensor location, we used four sensors to capture gait signals: three IMUZ sensors from ZMP Inc. [32], and a triaxial KXTF9 accelerometer from Kionix Inc. [33] inside a

Table 1
Existing inertial gait databases.

Research work	# Subjects	Age/gender	Sensor and data description
[15]	60	NA (not available)	Accelerometer at left hip, 12 sequences/subject
[5,6]	36	Adults, 17 females and 19 males	Accelerometer at waist, different days, fast, slow, and normal speeds
[21]	31	19 Males and 12 females	2 Sequences/subject, accelerometer at hip pocket, breast pocket and hand-held
[22]	50	NA	Accelerometer in trousers pocket
[23–25]	100	70 Males and 30 females, aged between 19 and 62 years	Accelerometer at right hip, 4 sequences/subject
[12]	30	NA	16 Sequences/subject, accelerometer at ankle, four shoe types
[26]	9	NA	Both accelerometer and gyroscope attached on shoe
[14]	21	10 Males and 11 females at ages between 19 and 40 years	Accelerometer at center of back waist, 5 sequences on different days
[13]	35	19 Males and 16 female, aged between 20 and 45 years	Accelerometer, 5 sequences/subject
[27]	51	41 Males and 10 females at ages of 20 s and 30 s	2 Days × 2 sequences/ subject
[8]	36	NA	Accelerometer in pocket
[28]	9	NA	Smartphone at subject's hip, 2 sequences on two days
[7]	58	NA	Accelerometer worked at 33 Hz in the hand-held iPhone, 40 sequences/subject on different days
[29]	48	38 Males and 10 females at ages mostly between 20 and 30 years	Smartphone at hip, 2 sequences on 2 days
[16]	32	7 Females and 25 males at ages of 21 to 40 years	Gyroscope, 5 sequences/subject with different attached weights
[19]	20	15 Males and 5 females at ages of 23 to 52 years	Accelerometer at center back, 10 datasets of 5 gait cycles for each subject
[30]	36	28 Males and 8 females at ages mostly between 22 and 28 years	Accelerometer in the right trouser's pocket, up and down a hallway, normal and fast speed
Our database	744	389 Males and 355 females at ages of 2 to 78 years	3 Inertial sensors (accelerometer and gyroscope) and a smartphone (accelerometer) working at 100 Hz are located around the waist, 3 different ground slope conditions

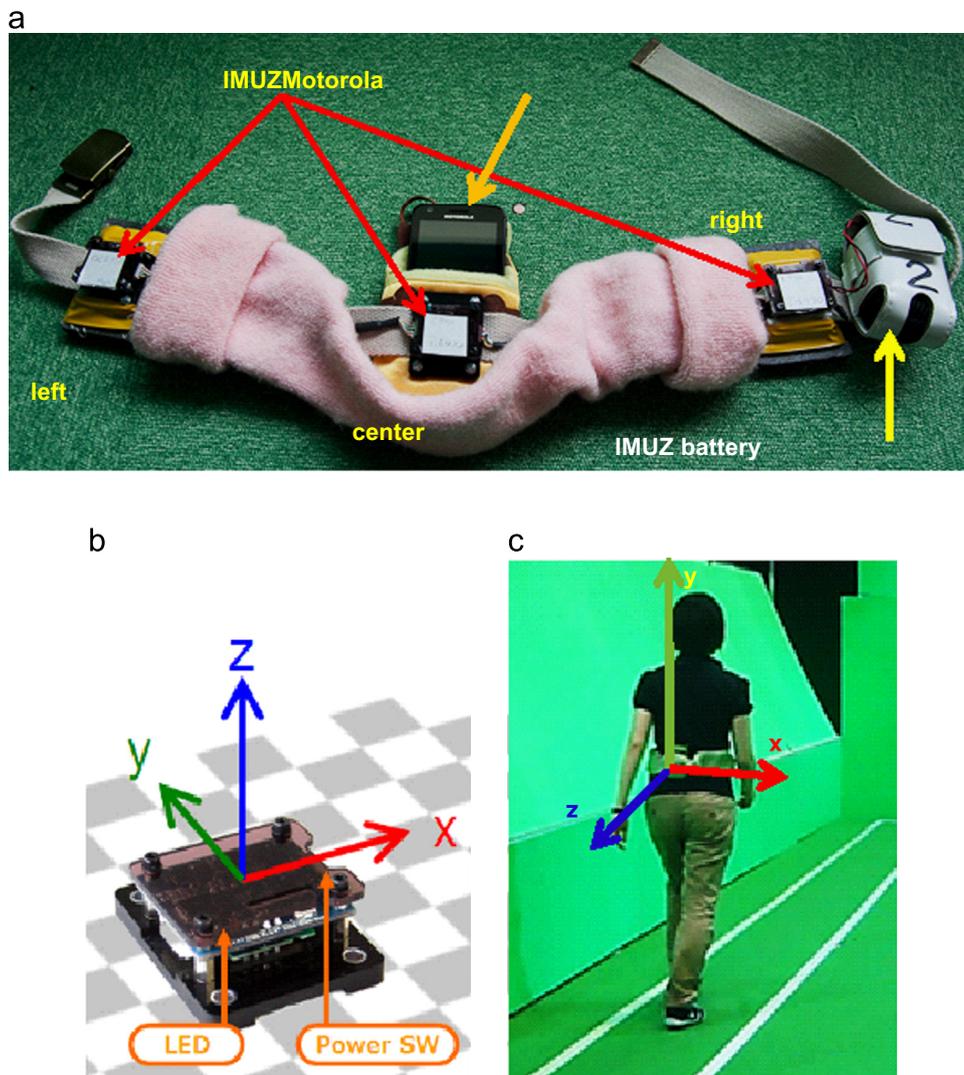


Fig. 1. Setup of the gait capturing system: (a) four uncovered sensors on a belt, (b) IMUZ, and (c) a belt on subject's waist.

Motorola ME860 smartphone. An IMUZ includes a triaxial accelerometer and a triaxial gyroscope. The IMUZ and KXTF9 accelerometers can work at a preset dynamic range, either ± 2 , ± 4 , or ± 8 (g), and the IMUZ gyroscope can work at either ± 500 or ± 2000 (deg/s). We set the IMUZ dynamic ranges at ± 4 (g) and ± 500 (deg/s) for capturing human gait signal. For the Motorola's accelerometer, the default setting by the manufacturer was used. In the experiments, the settings of all the four sensors were suitable for human gait and captured signal was not saturated. IMUZ sensors were connected to a remote computer through Bluetooth connection, while the Motorola could store all captured data in its internal memory. All the four sensors captured data at their maximum frame-rate, 100 Hz. The sensors were mounted on a waist belt as shown in Fig. 1(a). The belt was covered by a soft cushion to protect the sensors and avoid direct contact to the subject. Attachment of the belt on a subject's body is shown in Fig. 1(c). The center IMUZ and the smartphone were located at the center back, one IMUZ at the left and one at the right of the subject's waist. Subjects were asked to walk in and out the same designed path at their normal speed. The path contained a flat and a slope grounds, the slope angle was about 8 deg.

3.2. Database information

We collected the gait data of visitors in an exhibition during 5 days. Each visitor was requested to sign an informed consent to

permit the use of their data for research purpose. He/she also supplied the basic personal information such as gender and age. Thanks to them, we got the world largest database on inertial sensor-based gait.¹

We captured 6D signal sequences from the accelerometer and gyroscope in each IMUZ sensor and from the accelerometer in the Motorola. Data only for level, up-slope, and down-slope walk was extracted for each subject. For level walk, two sequences were extracted for entering and existing the path for each subject, meanwhile only one sequence was extracted for up-slope or down-slope walk. An example of a 6D signal of a subject is separately shown in Fig. 3 for accelerometer, gyroscope, and three ground slope conditions.

After simple preprocessing to remove invalid data and extract interested data, we found that the validity of captured data was not equal for all subjects, hence our database could be optimized with respect to several variation factors. The most important factor is the number of subjects, therefore we made the first dataset that included the maximum number of subjects. Meanwhile, the second dataset was maximized with the variations of sensor

¹ The database can be accessed from OU-ISIR Biometric Database: inertial sensor dataset, <http://www.am.sanken.osaka-u.ac.jp/BiometricDB>.

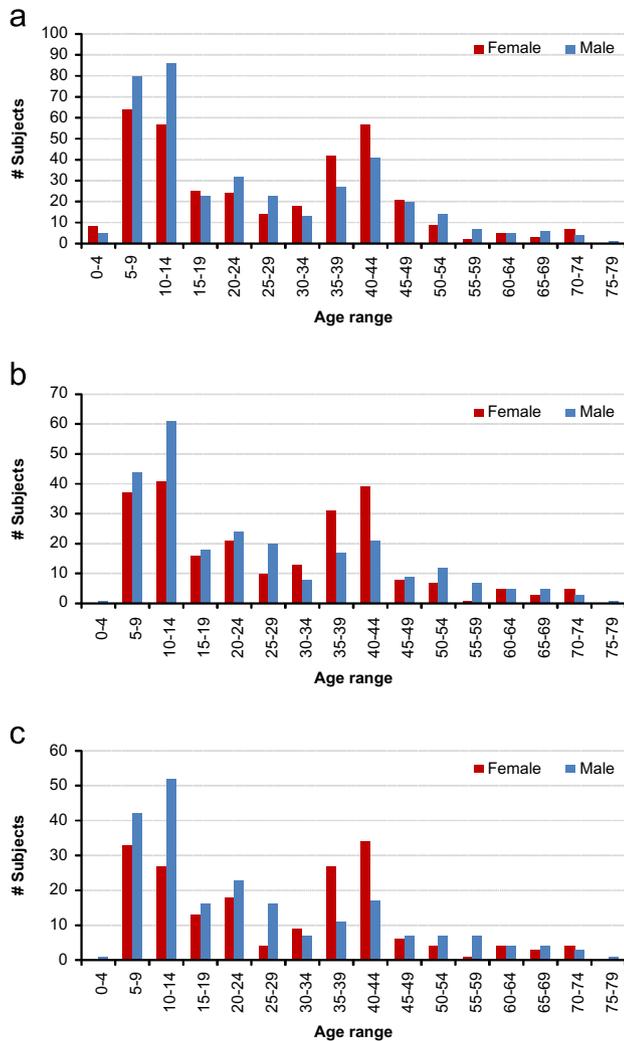


Fig. 2. Distributions of age and gender of the database in the first dataset for center IMUZ (a), IMUZ subset (b) and smartphone subset (c) of second dataset.

location, ground condition, and sensor type, therefore the number of subjects was sacrificed for these variations.

In the first dataset, level walk data of 744 subjects (389 males and 355 females) with ages ranging from 2 to 78 years was captured by the center IMUZ. In this dataset, two different level-walk sequences for each subject were extracted automatically by using motion trajectory constraint and signal autocorrelation. The detailed age distribution of subjects can be found in Fig. 2(a).

In the second dataset, variations of ground slope and sensor location were focused. For each subject and each sensor, we extracted two sequences for level walk, a sequence for up-slope walk, and a sequence for down-slope walk. In total, we had a subset from 3 IMUZ sensors for 495 subjects and a subset from the smartphone for 408 subjects. In this dataset, the data extraction for each sequence was performed manually by synchronizing with a simultaneously captured video. The age distribution of subjects can be seen in Fig. 2(b) and (c) for the subset of 3 IMUZ sensors and the subset of the smartphone, respectively.

Compared with the existing databases, the advantages of this database are as follows:

1. The number of subjects is about 7 times as large as in the existing largest gait databases. This significantly improves the reliability of the gait recognition performance evaluation.
2. The male-to-female ratio is close to 1. This is a desirable property for more reliable performance evaluation of gait-

based gender classification and for comparison of gait recognition performance between genders.

3. The subjects' ages are widely distributed from 2 to 78 years. In particular, the number of children is comparable with the number of adults. This provides more statistically reliable results on gait-based age group classification and difficulty level comparison of gait recognition among age groups.
4. 6D gait signal includes 3D acceleration and 3D angular velocity captured at a high frame rate, which is not only useful for gait recognition but also for understanding the walk motion.
5. Four sequences for three ground slope conditions are considered (two of them for level walk).
6. Variation of sensor locations (back center, left, and right positions on subject's waist) is considered to understand the dependence of recognition performance on sensor location around subject's waist.

The limitation of our database is that all walk sequences of a subject are relatively short, which were captured within a session of about 1 min. Level and slope walk paths were about 9 m and 3 m long, respectively.

4. Benchmark methods

To evaluate this large database, we apply recent four benchmark gait-based personal authentication methods by Rong et al. [13], Gafurov et al. [12], Derawi et al. [15], and Trung et al. [16]. We briefly summarize these methods in this section, the details can be found in the reference papers.

4.1. Gait period detection

Gait period detection is a key procedure for gait-based personal authentication and differs among the benchmark methods.

Rong et al. [13] detect periods based on up-down acceleration after normalizing the signal intensity to range $[-1, 1]$. Wavelet denoising is first applied, and local minimum points are then detected. Zero-cross point is detected just after a local minimum point. Next, four consecutive zero-cross points make a gait period. Finally, a fixed-size pattern is constructed by omitting intensity-similar samples so as to remain the maximum signal intensity difference within a period.

Gafurov et al. [12] also detect periods based on the up-down acceleration after being adjusted to be zero-mean. A start point of a gait period is found by the zero-cross point (from negative to positive), then autocorrelation is used to estimate the period length. The start point of the next period is the end point of the previous.

Derawi et al. [15] detect periods by using resultant signal (magnitude of 3D acceleration signal). First, the authors extract a single fixed-size subsequence from a signal sequence, and then slide the subsequence along the signal sequence to compute the matching distances and find the local minimum points of this distance sequence. Then, these local minimum points are combined with the local minimum points on the resultant signal to refine the period locations. All the gait periods can then be detected. Since, it is important to have a good subsequence, the authors heuristically extract the reference subsequence from the middle of the long sequence.

Trung et al. [16] use all the signal dimensions for the period detection using a phase-registration technique, called Self-DTW [34]. Self-DTW detects periods using a global optimization without using heuristic information on local points. Self-DTW sometimes results in temporally distorted periods, and then the authors linearize the time warping function (TWF) to correct the distortion. As a result, all detected periods are phase-registered and temporally undistorted.

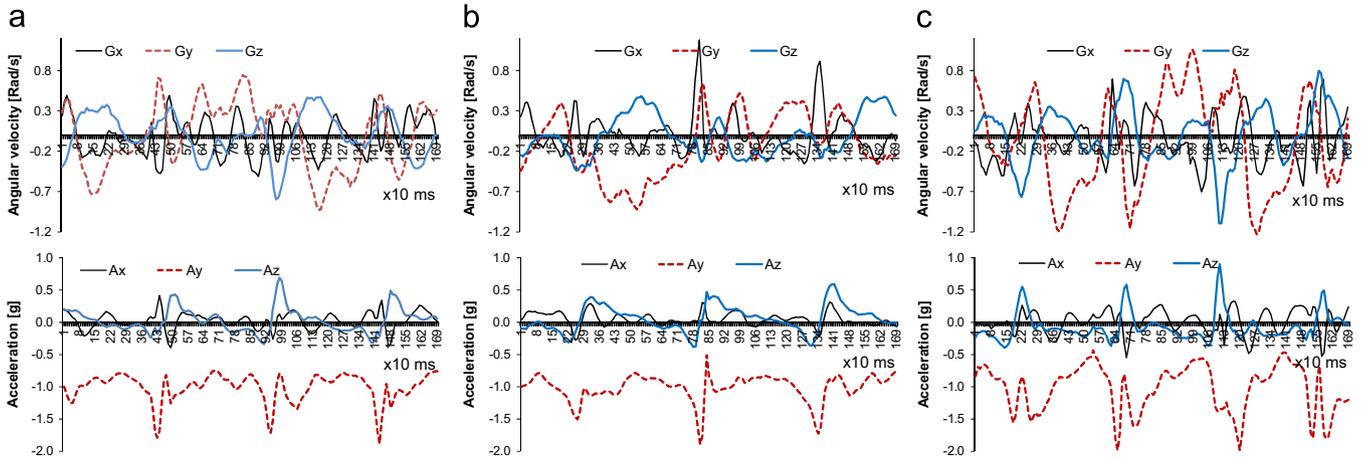


Fig. 3. Example of gait signals from gyroscope (the upper graphs) and accelerometer (lower graphs) for different ground slope conditions of a subject in the same sequence: (a) level walk, (b) slope up, and (c) slope down.

In the benchmark methods other than Rong et al.'s [13], a fixed-size pattern is constructed by simply re-sampling a detected period.

4.2. Authentication

First, a gallery \mathbf{G} is defined as a collection of all the sample gait patterns for the owner: $\mathbf{G} = \{\mathbf{g}_i\}$, $\mathbf{g}_i = \{g_{ij} | j = 0, \dots, N-1\}$, where i is the pattern index, and N is the size of a pattern. A circularly shifted gallery pattern \mathbf{g}_i by s is defined $\mathbf{g}_i^s = \{g_{i,(j+s) \bmod N}\}$.

For any probe pattern $\mathbf{p} = \{p_j\}$, the dissimilarity between it and all the shifted patterns in \mathbf{G} is computed using the minimum rule [35]:

$$D(\mathbf{G}, \mathbf{p}) = \min_{i,s} d(\mathbf{g}_i^s, \mathbf{p}), \quad (1)$$

where $d(\cdot, \cdot)$ is the dissimilarity between \mathbf{g}_i^s and \mathbf{p} .

Gafurov et al. [12] use Euclidean distance (L2 norm) for $d(\cdot, \cdot)$. In other benchmark methods by [13,15,16], $d(\cdot, \cdot)$ is computed as the normalized cumulative DTW score based on Euclidean distance at the end of the optimal warping path.

In fact, some other dissimilarity measures such as those based on normalized cross correlation, Tanimoto coefficient [36], and Manhattan distance can also be used. Therefore, we will evaluate these four dissimilarity measures in Section 5.4.

5. Performance evaluation

5.1. Evaluation procedure

In the experiments, we used all our extracted gait data for the personal authentication scenario (one-to-one matching). Data from gyroscope and accelerometer was evaluated separately, but period detection was executed simultaneously based on acceleration data.

Authentication performance of a method was evaluated by receiver operating characteristic (ROC) curve [37]. The ROC curve shows a trade-off between false rejection rate (FRR) and false acceptance rate (FAR) when the acceptance threshold is changed by a receiver in personal authentication scenario. The equal error rate (EER) where FRR and FAR are equal was also used to evaluate the performance. The lower the EER, the better the method performs.

We used the first dataset with a large number of subjects for evaluating the impact of the number of subjects in Section 5.2, evaluating the benchmark methods, Section 5.3, the impact of

genders, Section 5.6, and the impact of age groups, Section 5.7. We used the second dataset for evaluating the impact of sensor types, Section 5.4, ground slope conditions, Section 5.5, and the impact of sensor locations, Section 5.8.

For each subject, the first level-walk sequence was made a gallery, other sequences were made probes. To make a gallery \mathbf{G} for a subject, period detection of a benchmark method was applied for the first level walk sequence to make gallery patterns \mathbf{g}_i of gait period: $\mathbf{G} = \{\mathbf{g}_i\}$. To prepare a probe \mathbf{P} for each probe sequence, the benchmark method also used period detection to make probe patterns \mathbf{p}_j of gait period: $\mathbf{P} = \{\mathbf{p}_j\}$. Since the period detection performances of these benchmark methods were different, we used the whole sequence as a gallery or probe instead of an independent period. The dissimilarity $Diss(\mathbf{G}, \mathbf{P})$ between the probe \mathbf{P} and the gallery \mathbf{G} is computed

$$Diss(\mathbf{G}, \mathbf{P}) = \min_{i,j,s} d(\mathbf{g}_i^s, \mathbf{p}_j), \quad (2)$$

where $d(\mathbf{g}_i^s, \mathbf{p}_j)$ is defined in Section 4.2, which depends on benchmark method. The size of a pattern was 50 samples after normalization.

In all of our experiments, gallery for each subject includes one signal sequence which is the first level-walk sequence. Similarly, probe for each subject with each of three ground slope conditions (level, up-slope, and down-slope) also includes one signal sequence (the second level-walk, the up-slope, or down-slope sequence).

5.2. Impact of the number of subjects

In this section, the impact of the number of subjects is investigated. Based on a statistical analysis of ROC curves [38], the standard deviation of the observed FRR \hat{p} with multiple attempts for each subject is estimated as

$$\hat{\sigma}(\hat{p}) = \sqrt{\frac{\sum a_i^2 - 2\hat{p} \sum a_i m_i + \hat{p}^2 \sum m_i^2}{m^2 n(n-1)}}, \quad (3)$$

where these variables are described as follows:

- a_i : the number of false rejections for i th subject;
- m_i : the number of probe samples from i th subject;
- m : the average number of probe samples per subject; and
- n : the number of subjects.

This indicates that the obtained FRR becomes more reliable as the number of subjects increases. To validate the estimation, we

repeated the experiments with randomly chosen subsets of fewer subjects and compared the actual standard deviation of the performance and the estimated one from Eq. (3). First, we prepared 30 subsets of acceleration signals comprising 60 subjects randomly chosen from the first dataset and obtained 30 ROC curves by [16] method, which was equipped with Euclidean distance. Then, we calculated the average and standard deviation of the FRR for each FAR, depicted by an averaged ROC curve (gray line) and the standard deviation bars (in gray) in Fig. 4. In addition, the estimated standard deviation range from Eq. (3) is depicted by the two gray dashed lines for the subsets. From this graph, we can see that the experimental standard deviation ranges (gray bars) derived from the experimental results correspond well with those estimated from Eq. (3).

Moreover, the result for the whole set is superimposed as the black line, while the standard deviation range estimated from Eq. (3) is depicted by the two black dashed lines in Fig. 4. We can see that the standard deviation range is significantly narrower than that of the smaller subject subsets. For example at 10% FAR, the standard deviation is reduced from 5.32% to 1.36%. This indicates that the accuracy of the performance evaluation increases approximately 3.91 times.

5.3. Evaluation on benchmark methods

In this experiment, we evaluated the performance of the benchmark methods described in Section 4. All methods computed pattern dissimilarity based on Euclidean distance, as described in Section 4.2.

The ROC curves and their theoretical standard deviation of the benchmark methods are shown in Fig. 5 for accelerometer of the first dataset. We can see that although the ROC curves for these benchmarks are rather separated, their deviation boundaries are often overlapped. The performance of Derawi et al. and Rong et al. methods are quite similar since their ROC curves and deviation boundaries are not clearly separated. Meanwhile, the performance of Trung et al. and Gafurov et al.'s methods are quite different since

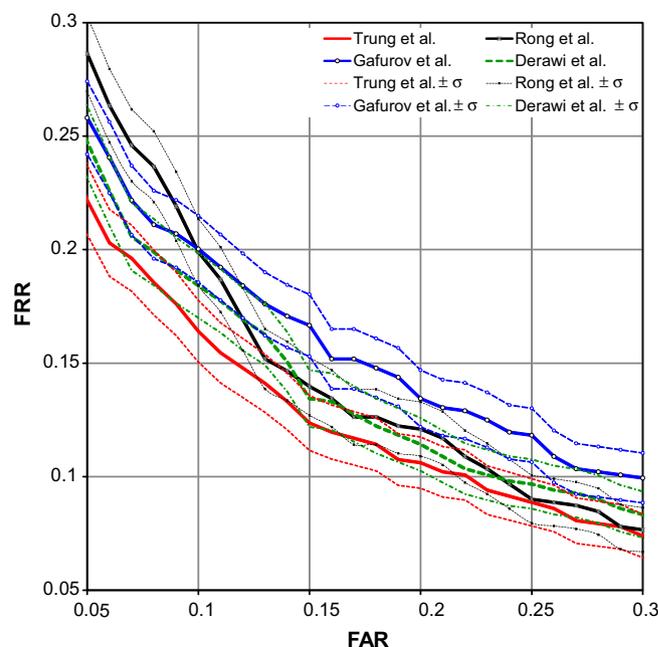


Fig. 5. Evaluation of the benchmark methods for accelerometer. A line and two dashed boundaries of the same color mean a false rejection rate p and its theoretical standard deviation range $p \pm \sigma$ derived from Eq. (3). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

their deviation boundaries are separated. One can imagine that if a small number of subjects was used, the comparison was unreliable since the deviation boundaries were wider. For examples, if we used a small subset of only 60 subjects, these deviation boundaries were about 3.91 time wider, as shown in Fig. 4.

For reference to the databases presented in original publications of the four benchmark methods, we list the benchmark methods' performances in Table 2. There are two points that need to be discussed from this table. First, although it is difficult to directly compare the performances by the original database with those by our databases since they were carried out under very different conditions in terms of sensor, sensor location, variation of shoes, and testing scenario, it turned out that the performances with our database were significantly worse than those with the original databases. One of the main reasons for this performance degradation is derived from subject variations in terms of age variation. While the original databases mainly contained healthy adults with suitable ages for recognition (mostly between 20 s and 40 s), our database contains many children and the elderly as well as adults. In fact, if we limit our database to a subset of 20's subjects, the EER becomes comparable to those with the original databases (approximately 6% for both, see Fig. 9 in Section 5.7). Although, children and the elderly have much lower recognition abilities, it is quite important to include them in evaluations when considering a number of useful applications such as those for security and safety at school and home. The second point to be mentioned is that the performance differences among the benchmarks with the original databases becomes statistically less significant than those with our database. For example, Gafurov et al. achieved 0.5% EER at the best case, while Derawi et al. and Rong et al. achieved 5.7% and 5.6% EERs, respectively, with the original databases. On the other hand, EERs with our database become 15.8%, 14.3%, and 14.3% for these methods, respectively. This means that the performance comparison with the small-sized original databases is not statistically meaningful, and hence the performance comparison with a common large-scale database is essential for practical applications. Our database could therefore highly

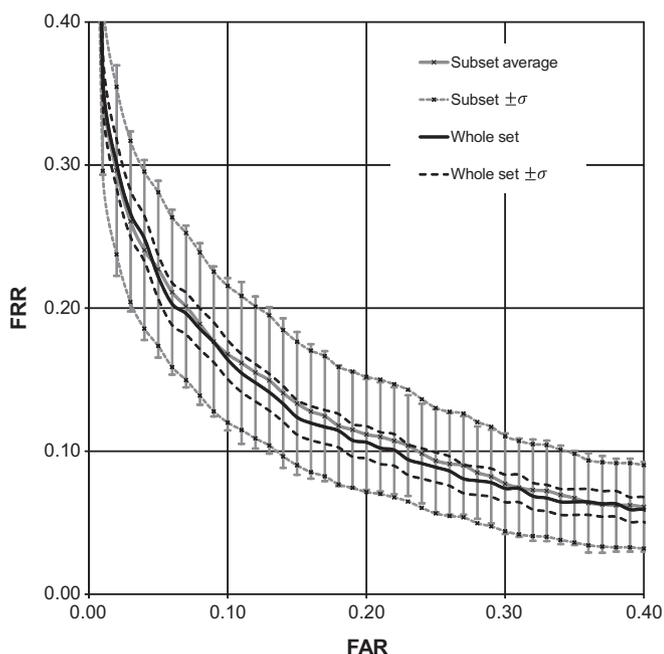


Fig. 4. ROC curves with different numbers of subjects. Black and gray lines stand for the whole set and smaller subsets, respectively. A line and two bounding dashed lines mean a false rejection rate p and its theoretical standard deviation range $p \pm \sigma$ derived from Eq. (3). Gray bars are actual standard deviation ranges $p \pm \sigma$ obtained from 30 ROC curves.

Table 2

Performances of benchmark methods with original databases.

Method	Original database Sensor information	No. of subjects	EER (%)	Proposed database Sensor information	No. of subjects	EER (%)
Derawi et al. [15]	Accelerometer at left hip	60	5.7	Accelerometer at back waist	744	14.3
Rong et al. [13]	Accelerometer at back waist	21	5.6	Accelerometer at back waist	744	14.3
Gafurov et al. [12]	Accelerometer at ankle	30	0.5–6.1 ^a	Accelerometer at back waist	744	15.8
Trung et al. [16]	Gyroscope in backpack	32	6.0	Gyroscope at back waist	744	20.2

^a In Gafurov et al.'s work, various databases with different types of shoes were used.

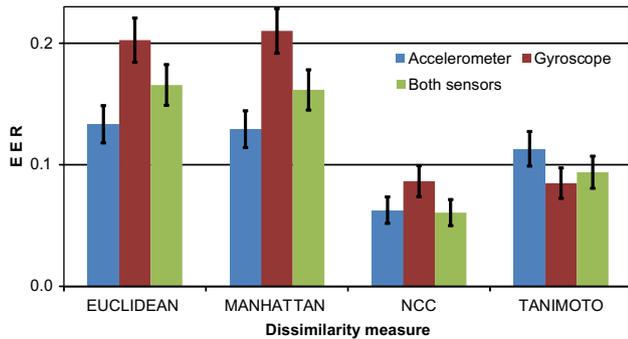


Fig. 6. EERs for dissimilarity measures versus sensor types. The black bars denote the standard deviation of FFR at EER estimated by Eq. (3).

contribute to the research community of the sensor-based gait recognition once it is released to the public.

From Fig. 5, we also can see that the benchmark method by Trung et al. produced a slight advantage in term of ROC curve, therefore, we select it for further evaluations in the following sections.

5.4. Impact of sensor types and dissimilarity measures

In this section, we compared the accuracy of the accelerometer and the gyroscope, and also explored a proper dissimilarity measure for each sensor. We used a data subset for the center IMUZ of the second dataset.

Performances of the normalized dissimilarities based on normalized cross correlation and Tanimoto coefficients were compared with those of the unnormalized dissimilarities, Manhattan and Euclidean distances, these dissimilarity measures are denoted as NCC, TANIMOTO, MANHATTAN, and EUCLIDEAN, respectively. The comparison results are shown in Fig. 6. From these results, we can see that accelerometer produces a better authentication performance than gyroscope does in overall, except for TANIMOTO. Normalized dissimilarities (NCC and TANIMOTO) work better than unnormalized dissimilarities (MANHATTAN and EUCLIDEAN) for any sensor. However, normalized dissimilarities work more effective for gyroscope than for accelerometer. These results come from the observation that inter-period signal fluctuation of the translational motion (acceleration) for a subject when he/she walks straightly forward is smaller than that of the rotational motion (angular velocity). Therefore, the normalization on angular velocity data is more effective than on acceleration data.

We also made a result for combining both sensors, which means that the gait signal has six dimensions. For the unnormalized dissimilarities, the result of using both sensors became worse than that of using only the accelerometer, and better than that of using only the gyroscope. However, for the NCC, the result of using both sensors became slightly better than that of any single sensor (accelerometer or gyroscope). Therefore, when we want to use both sensors, a proper dissimilarity measure or score-level fusion technique is preferred.

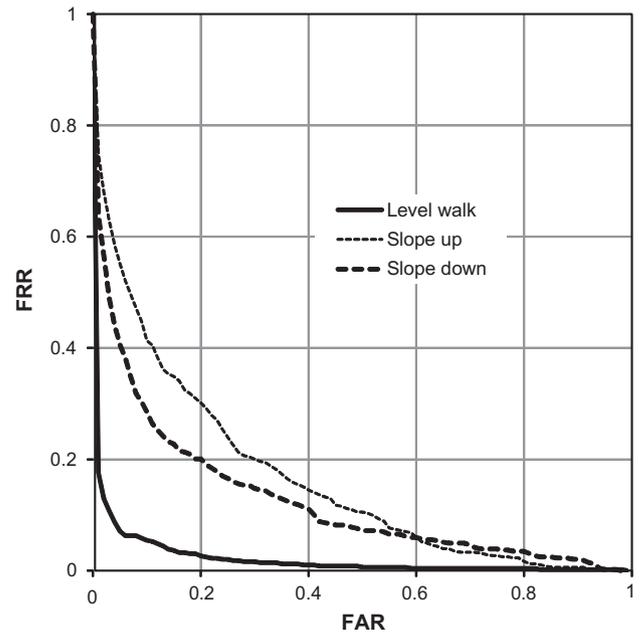


Fig. 7. Impact of ground slope condition on the ROC curves.

From these experiments, the combination of NCC and the accelerometer produced the best authentication performance and we selected it along with the Trung et al.'s method in the following experiments.

5.5. Impact of ground slope conditions

In this section, we investigate the impact of ground slope condition on gait authentication performance. For this purpose, we used the data subset for center IMUZ of the second dataset.

The ROC curves for different slope conditions are shown in Fig. 7. We can see that walking up/down a slope is quite different from walking on a flat ground which resulted in a decline of authentication performance for such ground slope conditions. In this experiment, the average error rates for slope-up and slope-down sequences increased significantly.

5.6. Impact of genders

In this section, we investigate how gait authentication performance differs between genders using the acceleration of the first dataset. Our large-scale gait database is suitable for this since the gender ratio is close to 1 in total. However, gender bias occurs for each age group in details as shown in Fig. 2(a). Therefore, we equalized the number of males and females for each age group to reduce such bias. There were 6 age groups used in this equalization: under 10, 10 s, 20 s, 30 s, 40 s, and over 50.

The ROC curve and its theoretical deviation for authentication performance among the 349 males are depicted by the gray lines and for the 349 females by the black lines in Fig. 8. According to

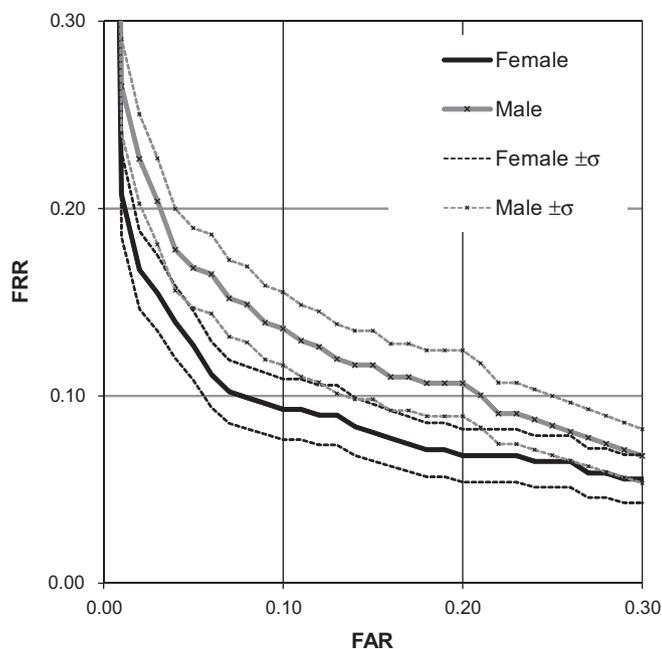


Fig. 8. ROC curves of gait authentication for each gender. Gray and black stand for male and female, respectively. A line and two dashed boundaries of the same gray level mean a false rejection rate p and its theoretical standard deviation range $p \pm \sigma$ derived from Eq. (3).

the results, the authentication performance among females is clearly better than that among males. This evaluation is reliable since the ROCs curves with their deviation boundaries are clearly separable with our database. A similar result was found by Gafurov [24], however, the evaluation was made with a small database of only 100 subjects aged between 19 and 62 with a very biased gender ratio (70 males vs. 30 females).

5.7. Impact of age groups

Next, we show the difference in gait authentication performance among the age groups. Our large-scale gait database is ideally suited to this purpose because the age distribution is much wider than that in existing gait databases.

In the first dataset, each 10-year interval up to 50 years contains about a hundred subjects. The interval over 50 years was treated separately as one age group because of the shortage of subjects. We then separated all the subjects into age groups: under 10, 10 s, 20 s, 30 s, 40 s, and over 50. The ROC curve for each age-group interval is shown in Fig. 9.

From these ROC curves, we can see that the gait authentication performance for the child group (under 10) is worse than that for most other age groups except the over-50 group, and this gradually improves with the older groups up to 20 years. The authentication performance is then gradually reduced for the older groups. This result is understandable because the intra-subject gait fluctuation for children is relatively larger due to the immaturity of their walking skills. In contrast, fluctuation in gait for adults is quite small since adults have established their own walking style; in other words, they have a stable gait pattern. Moreover, the gait authentication performance for groups over 30 years old declines as the subjects become older. This degradation in performance is inferred from the fact that physical strength generally declines as the subject grows older and hence, the gait tends to fluctuate more.

Consequently, gait authentication performance of subjects in their twenties is regarded as a trade-off between the maturity of walking ability and physical strength.

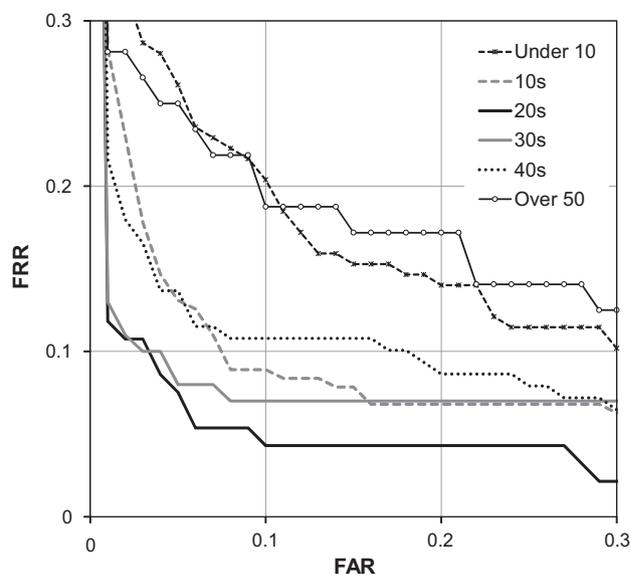


Fig. 9. ROC curves of gait authentication for each age-group interval.

5.8. Impact of sensor location around the waist

In this section, we evaluate the authentication performance for all the four sensors around the waist using the second. Since it was difficult to fix the left and right IMUZ sensors at the same orientation among all the subjects, the gait signals of the left and right IMUZ sensors were not aligned well to be compared. It is because when the orientation of the sensor coordinates changes, so does the captured signal. However, the change of sensor orientation does not affect the magnitude of signal. In other words, the magnitude of a 3D signal (resultant signal) is an 1D sensor-orientation-invariant signal. Therefore, although the authentication performance was degraded due to the reduction of dimensionality, we used the resultant signal instead of 3D original signal for the evaluation. The results of the evaluation are shown in Figs. 10 and 11 for 1D resultant signal of acceleration and angular velocity, respectively.

From the results for accelerometer, we can see that ROC curves for the smartphone and center IMUZ are quite similar. However, these ROC curves are slightly worse than those for the left and right sensors, which means accelerometer location on subject's waist slightly influences the authentication performance. On the other hand, from the results for angular velocity, we can see that ROC curves are quite similar, which means gyroscope location on subject's waist does not affect the authentication performance.

The difference between accelerometer and gyroscope's performances are understandable if we look into the mechanical difference between translational and rotational motions of solid object. Assuming that the belt, where the sensors are mounted, is a solid object, the same rotational velocity is captured at any point on the belt. However, the acceleration depends on the location of the sensor which resulted in the authentication performance difference among these accelerometers. The performances of the center accelerometer and the smartphone are similar because they are located at the same point at the back waist of a subject. This point is close to the center of mass of the subject, where the translational motion is much more stable and hence less informative than that of the left or right side. Meanwhile, translational motion at the left and right locations is influenced by the physical and geometrical characteristics of each individual waist. As a result, authentication performance using acceleration at the center back waist is less than that at the left or right side.

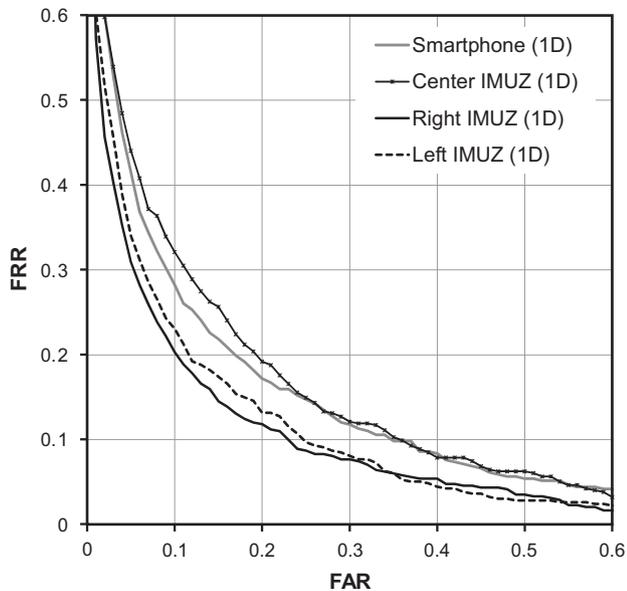


Fig. 10. Evaluation of sensor locations using acceleration: ROC curves for sensor locations. The smartphone and center IMUZ were at the same location.

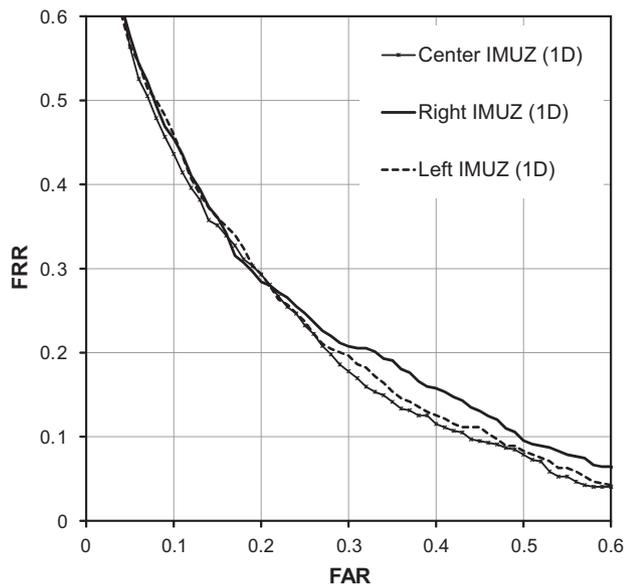


Fig. 11. Evaluation of sensor locations using angular velocity: ROC curves for sensor locations. ROC curve for the smartphone is not included since it does not contain a gyroscope.

6. Conclusions

This paper describes the construction of the world's largest inertial sensor-based gait database and a statistically reliable performance evaluation of inertial sensor-based gait authentication. The database has the following six advantages compared with existing gait databases: (1) the number of subjects is 744, which is approximately 7 times as large as the existing large-scale gait databases, (2) the male-to-female ratio is close to one, (3) the age distribution is wide, ranging from 2 to 78 years, (4) it includes 6D gait signal for both accelerometer and gyroscope, (5) 3 ground slope conditions are captured, and (6) it contains variation of sensor locations. Because a sufficient number of subjects for each gender or age group is included in the large-scale gait database, it is possible to evaluate how gait authentication performance differs between genders and among age groups.

The results of the benchmark authentication method provide several insights, such as the trade-off of gait authentication performance among age groups based on the maturity of walking ability and physical strength. Moreover, improvement in the statistical reliability of performance evaluation is shown by comparing the gait authentication results for the whole set and small subsets of subjects randomly selected from the whole set. The impacts of sensor types, dissimilarity measures, and sensor locations were also investigated. An accelerometer has a better authentication performance than that of a gyroscope. A normalized dissimilarity measure works better than an unnormalized dissimilarity measure. Sensor location on subject's waist does not influence authentication performance for a gyroscope, while it does for an accelerometer. Authentication performance for an accelerometer about the center of mass is not as good as that at left or right waist.

Future work includes performance evaluation of other state-of-the-art gait recognition approaches and present another recognition method for improving the robustness to the sensor attachment. Understanding the age or gender from the gait signals is also one of attractive research topics.

Conflict of Interest

None declared.

Acknowledgments

This work was partially supported by JSPS Grant-in-Aid for Scientific Research(S) Grant Number 21220003, and the JST CREST "Behavior Understanding based on Intention-Gait Model" project.

References

- [1] J.-Y. Yang, J.-S. Wang, Y.-P. Chen, Using acceleration measurements for activity recognition: an effective learning algorithm for constructing neural classifiers, *Pattern Recognition Letters* 29 (16) (2008) 2213–2220.
- [2] K. Altun, B. Barshan, O. Tunçel, Comparative study on classifying human activities with miniature inertial and magnetic sensors, *Pattern Recognition* 43 (10) (2010) 3605–3620.
- [3] J.C. Augusto, C.D. Nugent, A new architecture for smart homes based on ADB and temporal reasoning, in: *Toward a Human Friendly Assistive Environment, Proceedings of 2nd International Conference On Smart homes and health Telematic, ICOST2004, 2004*, pp. 106–113.
- [4] D. Cook, S. Das, *Smart Environments: Technology, Protocols and Applications*, John Wiley and Sons, 2004.
- [5] H.J. Ailisto, M. Lindholm, J. Mantyjarvi, E. Vildjiounaite, S. Makela, Identifying people from gait pattern with accelerometers, in: *Biometric Technology for Human Identification, SPIE*, vol. 5779, 2005, pp. 7–14.
- [6] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S.-M. Makela, H. Ailisto, Identifying users of portable devices from gait pattern with accelerometers, in: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2005, pp. 973–976.
- [7] T. Kobayashi, K. Hasida, N. Otsu, Rotation invariant feature extraction from 3-D acceleration signals, in: *International Conference on Acoustics, Speech, and Signal Processing*, 2011, pp. 3684–3687.
- [8] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Cell phone based biometric identification, in: *IEEE Fourth International Conference on Biometrics: Theory, Applications and Systems*, 2010, pp. 1–7.
- [9] A. Kale, N. Cuntoor, B. Yegnanarayana, A. Rajagopalan, R. Chellappa, Gait analysis for human identification, in: *Audio- and Video-Based Biometric Person Authentication LNCS*, 2003, pp. 706–714.
- [10] D. Gafurov, K. Helkala, T. Sondrol, Biometric gait authentication using accelerometer sensor, *Journal of Computers* (2006) 51–59.
- [11] D. Gafurov, E. Snekkenes, Gait recognition using wearable motion recording sensors, *EURASIP Journal on Advances in Signal Processing* 2009 (2009) 1–16.
- [12] D. Gafurov, E. Snekkenes, P. Bours, Improved gait recognition performance using cycle matching, in: *2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*, 2010, pp. 836–841.
- [13] L. Rong, Z. Jianzhong, L. Ming, H. Xiangfeng, A wearable acceleration sensor system for gait recognition, in: *2nd IEEE Conference on Industrial Electronics and Applications*, 2007, pp. 2654–2659.

- [14] L. Rong, D. Zhiguo, Z. Jianzhong, L. Ming, Identification of individual walking patterns using gait acceleration, in: The 1st International Conference on Bioinformatics and Biomedical Engineering, 2007, pp. 543–546.
- [15] M.O. Derawi, P. Bours, K. Holien, Improved cycle detection for accelerometer based gait authentication, in: Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010, pp. 312–317.
- [16] N. Trung, Y. Makihara, H. Nagahara, R. Sagawa, Y. Mukaigawa, Y. Yagi, Phase registration in a gallery improving gait authentication, in: International Joint Conference on Biometrics (IJCB2011), IEEE and IAPR, 2011, pp. 1–7.
- [17] S. Sprager, D. Zazula, A cumulant-based method for gait identification using accelerometer data with principal component analysis and support vector machine, *WSEAS Transactions on Signal Processing* 5 (11) (2009) 369–378.
- [18] G. Trivino, A. Alvarez-Alvarez, G. Bailador, Application of the computational theory of perceptions to human gait pattern recognition, *Pattern Recognition* 43 (7) (2010) 2572–2581.
- [19] A. Alvarez-Alvarez, G. Trivino, O. Cordon, Human gait modeling using a genetic fuzzy finite state machine, *IEEE Transactions on Fuzzy Systems* 20 (2) (2012) 205–223.
- [20] N. Trung, Y. Makihara, H. Nagahara, Y. Mukaigawa, Y. Yagi, Performance evaluation of gait recognition using the largest inertial sensor-based gait database, in: Proceedings of the 5th IAPR International Conference on Biometrics, 2012, pp. 360–366.
- [21] E. Vildjiounaite, S.-M. Makela, M. Lindholm, R. Riihimaki, V. Kyllonen, J. Mantyjarvi, H. Ailisto, Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices, in: 4th International Conference on Pervasive Computing, PERSASIVE, 2006, pp. 187–201.
- [22] D. Gafurov, E. Snekkenes, P. Bours, Gait authentication and identification using wearable accelerometer sensor, in: 5th IEEE Workshop on Automatic Identification Advanced Technologies, 2007, pp. 220–225.
- [23] D. Gafurov, E. Snekkenes, P. Bours, Spoof attacks on gait authentication system, *IEEE Transactions on Information Forensics and Security* 2 (3) (2007) 491–502.
- [24] D. Gafurov, Security analysis of impostor attempts with respect to gender in gait biometrics, in: First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007, pp. 1–6.
- [25] D. Gafurov, P. Bours, Improved hip based individual recognition using wearable motion recording sensor, in: Security Technology, Disaster Recovery and Business Continuity, vol. 122, Communications in Computer and Information Science, Springer, Berlin, Heidelberg, 2010, pp. 179–186.
- [26] B. Huang, M. Chen, P. Huang, Y. Xu, Gait modeling for human identification, in: IEEE International Conference on Robotics and Automation, 2007, pp. 4833–4838.
- [27] M.O. Derawi, C. Nickely, P. Bours, C. Busch, Unobtrusive user-authentication on mobile phones using biometric gait recognition, in: 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010, pp. 306–311.
- [28] S. Bonnet, P. Jallon, Hidden markov models applied onto gait classification, in: 18th European Signal Processing Conference (EUSIPCO-2010), 2010, pp. 929–933.
- [29] C. Nickel, H. Brandt, C. Busch, Classification of acceleration data for biometric gait recognition on mobile devices, in: BIOSIG, 2011, pp. 57–66.
- [30] F. Juefei-Xu, C. Bhagavatula, A. Jaech, U. Prasad, M. Savvides, Gait-id on the move: pace independent human identification using cell phone accelerometer dynamics, in: IEEE 5th International Conference on Biometrics, 2012, pp. 8–15.
- [31] D. Gafurov, Performance and Security Analysis, Ph.D. Thesis, Faculty of Mathematics and Natural Sciences at the University of Oslo, 2008.
- [32] ZMP Inc., IMUZ, (<http://www.zmp.co.jp/imu-z.html>), In Japanese.
- [33] Kionic Inc., KXTF9, (<http://www.kionix.com/>).
- [34] Y. Makihara, N.T. Trung, H. Nagahara, R. Sagawa, Y. Mukaigawa, Y. Yagi, Phase registration of a single quasi-periodic signal using self dynamic time warping, in: The Tenth Asian Conference on Computer Vision, Springer-Verlag, 2010, pp. 667–678.
- [35] J. Kittler, M. Hatef, R.P. Duin, J. Matas, On combining classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (3) (1998) 226–239.
- [36] A.H. Lipkus, A proof of the triangle inequality for the Tanimoto distance, *Journal of Mathematical Chemistry* 26 (2006) 263–265.
- [37] P. Phillips, D. Blackburn, M. Bone, P. Grother, R. Micheals, E. Tabassi, Face recognition vendor test, (<http://www.frvt.org>), 2002.
- [38] G. Snedecor, W. Cochran, *Statistical Methods*, Iowa State University Press, 1967.

Thanh Trung Ngo received his M.E and Ph.D. degrees from Osaka University in 2005 and 2010, respectively. He is currently working as a postdoctoral researcher in the Department of Intelligent Media, Institute of Scientific and Industrial Research, Osaka University. His research interests include gait recognition, robot localization and 3D map building, and omnidirectional vision.

Yasushi Makihara received the B.S., M.S., and Ph.D. degrees in engineering from Osaka University in 2001, 2002, and 2005, respectively. He is currently an Assistant Professor at the Institute of Scientific and Industrial Research, Osaka University. His research interests are gait recognition, morphing, and temporal super resolution.

Hajime Nagahara received the B.E. and M.E. degrees in electrical and electronic engineering from Yamaguchi University in 1996 and 1998, respectively, and the Ph.D. degree in system engineering from Osaka University in 2001. He was a research associate of the Japan Society for the Promotion of Science (2001–2003). He was an assistant professor at the Graduate School of Engineering Science, Osaka University, Japan (2003–2010). He was a visiting associate professor at CREA University of Picardie Jules Verns, France, in 2005. He was a visiting researcher at Columbia University in 2007–2008. Since 2010, he has been an associate professor in faculty of information science and electrical engineering at Kyushu University. Computational photography, computer vision, and virtual reality are his research areas. He received an ACM VRST2003 Honorable Mention Award in 2003.

Yasuhiro Mukaigawa received his M.E. and Ph.D. degrees from University of Tsukuba in 1994 and 1997, respectively. He became a research associate at Okayama University in 1997, an assistant professor at University of Tsukuba in 2003, and an associate professor at Osaka University in 2004. He joined the MIT Media Lab as a visiting associate professor from 2009 to 2010. His current research interests include computer vision and computational photography.

Yasushi Yagi is the Director of the Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Japan. He received his Ph.D. degrees from Osaka University in 1991. In 1985, he joined the Product Development Laboratory, Mitsubishi Electric Corporation, where he worked on robotics and inspections. He became a research associate in 1990, a lecturer in 1993, an associate professor in 1996, and a professor in 2003 at Osaka University. International conferences for which He has served as chair include : ROBIO2006 (PC), ACCV (2007PC, 2009GC), PSVIT2009 (FC), ACPR (2011PC, 2013GC). He has also served as the Editor of IEEE ICRA Conference Editorial Board (2008–2011). He is the member of Editorial Board of International Journal of Computer Vision, the Editor-in-Chief of IPSJ Transactions on Computer Vision & Applications and the Financial Chair of Asian Federation of Computer Vision Societies. He was awarded ACM VRST2003 Honorable Mention Award, IEEE ROBIO2006 Finalist of T.J. Tan Best Paper in Robotics, IEEE ICRA2008 Finalist for Best Vision Paper, PSVIT2010 Best Paper Award, MIRU2008 Nagao Award and IPSJ Trans. on Computer Vision and Applications Outstanding Paper Award (2011). His research interests are computer vision, medical engineering and robotics. He is a fellow of IPSJ and a member of IEICE, RSJ, and IEEE.