Electronic Preprint for Journal of Information Processing Vol.24 No.3

Regular Paper

Multi-angle Gait Recognition Based on Skeletal Tracking Data

Yusuke Manabe[1,a)] Keisuke Matsumoto[2] Kenji Sugawara[1,b)]

Received: June 15, 2015, Accepted: January 12, 2016

Abstract: Human gait recognition is one of the most important authentication technologies as it can often happen that people approach a computer system or a robot by walking. Therefore in this study, a multi-angle gait recognition method has been proposed by using skeletal tracking data, measured by an RGB-D camera. The proposed method includes a two stage process, which estimates an optimal gait angle view from the five discrete angles at the first stage and subsequently recognizes human gait based on the specific features for the respective gait angle views. In order to evaluate the proposed method, two types of experiments have been done: gait angle estimation and gait recognition. From the result of the first experiment, the best estimation of 97.4% accuracy has been achieved. In the second experiment, the best gait recognition accuracy was 96.4%. Finally the best gait recognition accuracy with the two stage process has been estimated as 93.9%.

Keywords: multi-angle view, gait recognition, soft biometrics and skeletal tracking

1. Introduction

According to a report by Sumi [1], it is reported that the research on soft biometrics can be applied to Human-Computer Interactions (HCI) and many commercial products. The term soft biometrics is defined by Jain et al. [2] as follows; "*soft biometric traits as characteristics that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals*". Specifically it refers to height, weight, skin color, eye color, hairstyle, facial mark, ethnicity and so on of a person. From the view point of the less distinctiveness and permanence to identify individuals, most behavioral biometric traits such as handwritten signature, gait, lip movement, etc. can be regarded as a kind of soft biometrics.

On the other hand, some advanced camera devices with a depth sensor such as Kinect [3], Xtion [4], DepthSense [5] and Leap Motion [6] have been recently developed and are established as the next generation camera devices. These depth cameras can be used and controlled with some freely available SDKs. A lot of research has been driven by the depth camera devices and the SDKs which can facilitate automatic human detection and skeletal tracking [7], [8], [9], [10], [11]. This means that we can easily obtain many soft biometric traits such as body measurements and movements which can be computed from skeletal tracking data measured by the depth cameras.

In this paper, as an HCI application, we consider a personal or family assistant system in a smart home. The system can recognize and adaptively provide information and service to them in a specific room. Hence we postulate that the gait data of target persons are in the system. Here the most important thing is to naturally obtain the identity information from a person, without the active authentication behaviors such as holding a fingerprint to the sensor. In such an application scene, the system cannot obtain sufficient information to identify individuals because of many obstructions and restrictions such as a room layout, lighting conditions, sensor settings and so on. In order to cope with such actual problems, it is necessary to develop a recognition technology which can extend the coverage of information and conditions for humans in actual environment.

We consider a key component of success for the realization of the above application system is to develop a recognition technology of individuals and activities of daily living (ADL) on the basis of soft biometric traits obtained by depth camera devices. Especially human gait recognition is the most important technology because it can often arise that people approach a computer system by walking. Moreover gait recognition has a quite long history since 1964 in which Murray et al. first argued that human walking patterns include individual variability [12].

Gait recognition techniques by using RGB cameras are divided into two approaches; appearance-based approach and model-based approach. The former characterizes the gait from the silhouette image of a human. The most of gait recognition methods fall into this category. The latter uses gait features extracted from a body skeleton model fitted to a human area in image data. On account of the difficulty of body skeleton estimation and the occlusion problem of arms and legs, this approach does not conventionally outperform the appearance-based one. In addition, the applied model is a simple stick model as the human body is composed of just three line segments, which mean the trunk and the two legs of human [13], [14]. So it can be said that there

[1] Chiba Institute of Technology, Narashino, Chiba 275–0016, Japan
[2] AISOL, INC., Toshima, Tokyo 171–0021, Japan
[a)] ymanabe@net.it-chiba.ac.jp
[b)] suga@net.it-chiba.ac.jp

is much smaller selection of model-based approaches [15] than appearance-based ones. However, now depth camera devices and their SDKs can encourage new solutions to overcome the drawbacks in the model-based approaches. At least we can easily obtain more precise model data than the conventional ones.

In gait recognition, regardless of either model-based approach or appearence-based approach, human gait is generally captured in the side view from which gait cycles and strides can be easily estimated. Some recent researches have challenged gait recognition from the front view [16], [17], [18], [19], [20], [21], [22]. Soriano et al. claims that front-view gait recognition has an advantage that it is compatible with the ability of humans to recognize their fellow beings from their front-view gaits [16]. In appearance-based approaches, there are many contributions for multi-angle view gait recognition [23]. It is significant to support multi-angle view gait because the computer system or a robot is not always able to capture the human gait pattern in the real environment from a particular view. We need to develop a gait recognition system which successfully works in an unconstrained environment where humans can freely walk in a natural way. However, compared with appearance-based approaches, there are few model-based recognition methods which can support multi-angle view gait. Development of multi-angle human recognition technology based on model-based data is beneficial to an improvement of human recognition accuracy because it can be integrated conventional to appearance-based technologies.

Therefore in this study we propose a novel multi-angle gait recognition method based on skeletal tracking data measured by an RGB-D camera. Specifically our method can estimate an optimal gait angle view from the five discrete angles at the first stage and subsequently recognize human gait based on the optimal features for the respective gait angle views. As far as we know, there does not exist researches of evaluating model-based gait data from the viewpoint of human discriminability in multi-angles.

The organization of the paper is as follows. Section 2 shows our gait angle estimation method and gait recognition method. Section 3 describes the design of the feature vector and the feature evaluation method in detail. Sections 4 and 5 describe two experiments performed in order to evaluate the proposed method and their results. Finally, Section 6 contains conclusion.

2. Multi-angle Gait Recognition Method

This section describes the proposed method: gait angle estimation and gait recognition.

2.1 Overview

Figure 1 shows the processing diagram of the proposed method, which is divided into two flows: an enrollment flow and a recognition flow.

In the enrollment flow, human gait pattern is measured as a three-dimensional time series of skeletal tracking data. Kinect and a Kinect SDK are used as the RGB-D camera and the development kit of the data measurement program respectively. **Figure 2** shows a joint type enumeration for the skeletal tracking by the Kinect SDK. Here let it be noted that each joint ID is assigned by us. In this study we collect the gait data for multi-angle view based on relative positions between the RGB-D camera and the human. **Figure 3** shows the relation between the camera and the gait angles. In this figure, the circle denotes the measurement scope of the RGB-D camera and the five red arrows denote gait angles: 90, 135, 180, 225 and 270 degrees. A human subject walks from the right side of the camera at angles of 90 and 135 degrees, and walks from the left side at angles of 225 and 270 degrees. At an angle of 180 degrees, a subject approaches the camera straight. After measurement, gait data is preprocessed for computing thresholds for gait angle estimation. The computed thresholds are stored in the system and it is used in the recognition flow. Finally the gait data is processed and some soft biometric features are extracted from the preprocessed gait data for designing multi-angle gait recognizers. The gait recognizers are stored in the system, and are used at the time of recognition.

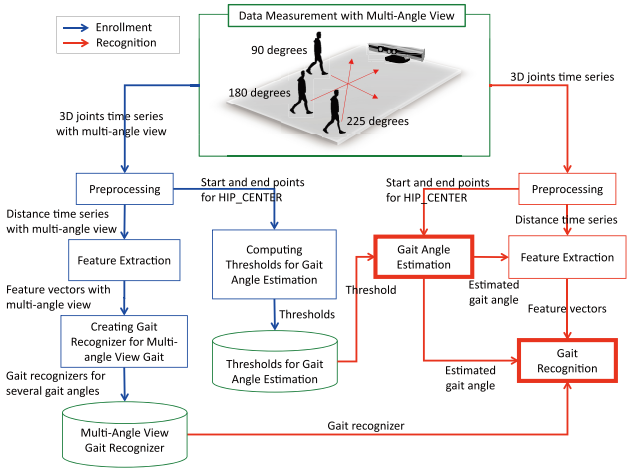


Fig. 1 Processing diagram of proposed gait recognition method.

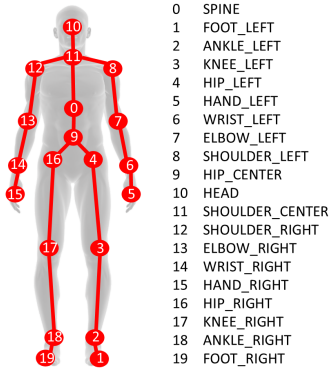


Fig. 2 Joint type enumeration for skeletal tracking.

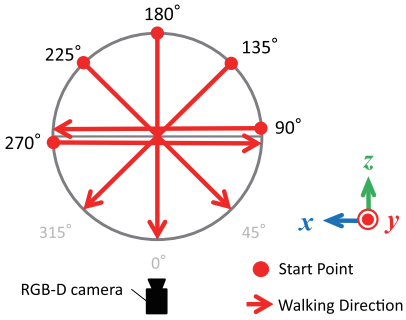


Fig. 3 Relation between RGB-D camera and gait angles.

Algorithm 1 Gait Angle Estimation

```
(x_s, z_s) ← detect the start point S
(x_e, z_e) ← detect the end point E
Ψ ← compute the cosine based on Eq. (1)
if φ^180 ≤ Ψ ≤ Φ^180 then
    Gait Angle is estimated as 180 degrees.
else
    if x_s < 0 then
        if φ^90 ≤ Ψ ≤ Φ^90 then
            Gait Angle is estimated as 90 degrees.
        else if φ^135 ≤ Ψ ≤ Φ^135 then
            Gait Angle is estimated as 135 degrees.
        else
            Gait Angle is not estimated.
        end if
    else
        if φ^225 ≤ Ψ ≤ Φ^225 then
            Gait Angle is estimated as 225 degrees.
        else if φ^270 ≤ Ψ ≤ Φ^270 then
            Gait Angle is estimated as 270 degrees.
        else
            Gait Angle is not estimated.
        end if
    end if
end if
```

In the recognition flow, first a human gait is captured by an RGB-D camera and gait data is measured and preprocessed by the same manner as ones in the enrollment flow. Secondly, gait angle estimation is performed by the stored thresholds. The algorithm is described in the next subsection. Finally, on the basis of the estimated gait angle, the optimal feature sets are extracted and the human gait is recognized by the stored optimal gait recognizer. The recognition method is described in a later subsection and the feature extraction and selection methods are described in the next section.

2.2 Gait Angle Estimation

It is easy to estimate the gait angles because we can track the positions of the joints in the three-dimensional coordinates built up by RGB-D camera. The two-dimensional coordinates (the x-z plane) are sufficient as we assume that humans walk on the flat floor. We estimate the gait angle from the cosine between the start point $\boldsymbol{S} = (x_s, z_s)$ and the end point $\boldsymbol{E} = (x_e, z_e)$ on the two-dimensional coordinates,

$$\Psi = \frac{x_s x_e + z_s z_e}{\sqrt{{x_s}^2 + {z_s}^2}\sqrt{{x_e}^2 + {z_e}^2}}, \tag{1}$$

where (x_s, z_s) denotes the HIP_CENTER position at the place where a human is first detected within the measurement scope, (x_e, z_e) denotes the same one at the place where a human is last detected within the scope. The gait angle estimation based on the equation is shown as Algorithm 1.

The algorithm is quite simple. First it estimates the 180 degrees gait by the lower threshold ϕ^{180} and the upper threshold Φ^{180} because the Ψ is definitely close to zero. Next, the sign of x_s is found out which determines whether the gait was performed from the right side or from the left side. This classification is important because the Ψ values in 90 or 135 degrees are approximately equivalent to the ones in 270 or 225 degrees respectively. When the gait was performed from the right side, it estimates the 90 degrees or the 135 degrees gait by using the corresponding thresholds (ϕ^{90}, Φ^{90}, ϕ^{135} and Φ^{135}). In the other case it estimates the 225 degrees or the 270 degrees gait by using the corresponding thresholds (ϕ^{225}, Φ^{225}, ϕ^{270} and Φ^{270}). Each of the lower and the upper thresholds ($\phi^g, \Phi^g,\ g = 90, 135, 180, 225, 270$) is computed from the training data samples measured in advance.

2.3 Preprocessing for Gait Recognition

Three-dimensional position time series data is processed to extract features easily. First the three dimensional position time series ${\boldsymbol{P}^j}_t$ is converted into one dimensional distance time series $d_j(t)$ with the SPINE position ${\boldsymbol{P}^0}_t$ as a reference point by the following Eq. (2);

$${d^j}_t = \|{\boldsymbol{P}^j}_t - {\boldsymbol{P}^0}_t\| \quad (t = 1, 2, \cdots, N), \tag{2}$$

where t denotes discrete the time stamp, N is the length of time series, j is a joint ID in Fig. 2, $\|.\|$ denotes *Euclidean* distance in three-dimensional space, $\boldsymbol{P}^j_t$ denotes $\{x^j{}_t, y^j{}_t, z^j{}_t\}$ and $\boldsymbol{P}^0_t$ denotes $\{x^0{}_t, y^0{}_t, z^0{}_t\}$. This is equivalent to the processing represented in [22]. The raw position time series is distance data from the RGB-D camera so it monotonously changes on some axes in response to the gait angles. By using this conversion, the periodicity originated by walking comes to emerge in the distance time series of each joint. Furthermore the obtained distance time series is less affected by the change of the gait angles because it does not depend on the position of the RGB-D camera.

After the conversion, a simple moving average smoothing is repeated five times by the following Eq. (3);

$${s^j}_t = \frac{1}{3}\sum_{k=-1}^{1} {d^j}_{t+k} \quad (t = 2, 3, \cdots, N-1). \tag{3}$$

2.4 Gait Recognition

In this study human gait is recognized by using a description of a gait pattern on the basis of some soft biometric features from the smoothed time series. Moreover we select the optimal feature vectors for each gait angle. The details of feature extraction and selection techniques are described in the next section.

Finally the gait data is recognized by using the optimal feature vectors and machine learning methods. In this study we employ linear discriminant analysis (LDA) methods as gait recognizers. The recognizers are prepared for every gait angles by using the training data samples measured in advance.

3. Feature Extraction

In this section, the details of feature extraction and selection methods are explained. The features are mainly related to the physical measurements of human body parts. These features are kinds of soft biometric traits.

3.1 Features Design by The Fisher Criterion

We compute four soft biometric features: minimum, maximum, mean and standard deviation for the distance time series of each joint. Consequently we deal with 76 kinds of features (4

features × 19 joints) in total. The four features are calculated as follows:

$$\alpha^j = \min\{\, s^j{}_t,\ \ t = 1, 2, \cdots, N \,\}, \tag{4}$$

$$\beta^j = \max\{\, s^j{}_t,\ \ t = 1, 2, \cdots, N \,\}, \tag{5}$$

$$\gamma^j = \frac{1}{N}\sum_{t=1}^{N} s^j{}_t, \tag{6}$$

$$\delta^j = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(s^j{}_t - \gamma^j)^2}, \tag{7}$$

where j denotes a joint ID in Fig. 2, t denotes a discrete time index and N denotes the length of time series.

Physically α^j is a distance in which a j-th joint is the closest to the reference point during walking. For example, the values of knees (KNEE_RIGHT and KNEE_LEFT) reflect the way of knee lift during walking. β^j is the distance in which the j-th joint is the most distant to the reference point. For example, the values of hands (HAND_RIGHT and HAND_LEFT) tend to be large for the human who swings arms with large amplitude during walking. γ^j is the mean distance between the j-th joint and the reference point. The values of trunk (HEAD, SHOULDER_CENTER and HIP_CENTER) or ankles (ANKLE_RIGHT and ANKLE_LEFT) are correlated to the length of body and leg respectively. δ^j is the variability of movement in the j-th joint so it will be large in limbs. This value reflects the difference of swing frequency in body parts. For example, the values of hands or elbows tend to be large for the human who walks with frequent arm swing. Also the values of knees reflect the step length and the step width during walking.

In order to reveal the discriminant potential of these soft biometric features, we analyze distinctiveness of individuality by using the Fisher criterion of linear discriminant analysis. The Fisher criterion J is computed as follows;

$$J = \frac{B}{W}, \tag{8}$$

$$W = \frac{1}{N}\sum_{i=1}^{N^c}\sum_{j=1}^{N^i}(f_{ij} - \overline{f_i})^2, \tag{9}$$

$$B = \frac{1}{N}\sum_{i=1}^{N^c} N^i(\overline{f_i} - \overline{f})^2, \tag{10}$$

where W is the variability within the classes (intra-class valiability), B is the variability between the classes (inter-class valiability), N^c denotes the number of classes to be identified, N^i denotes the number of data samples in the i-th class, N the total number of data samples based on $N = \sum_{i=1}^{N^c} N^i$, f_{ij} denotes the j-th data sample (feature value) in the i-th class, $\overline{f_i}$ denotes the mean of the data samples in the i-th class and $\overline{f}$ is the overall mean of the data samples.

On the basis of the criterion J, we analyze the discriminant tendency of each feature in all of the joints. This analysis is performed for each gait angle.We select the top M features based on J as optimal features for each gait angle. Thus the selected feature set is a M-dimensional vector, and it is composed of α, β, γ and δ.

4. Experiment I: Gait Angle Estimation

This section represents the simulation experiment for gait angle estimation.

4.1 Data Collection

Figure 4 shows a sketch of the measurement environment. A human subject walks a distance of two and a half meter, which is the diameter of the measurement scope of the RGB-D camera shown in Fig. 3. The scope is smaller than the possible measurement of the sensor so as to reliably track the skeletal data.

We collected data from 10 human subjects of ages ranging from 21 to 23 years. One subject is female and the others are male. **Table 1** shows the distribution of the human subject's height with shoes. We taught each human subject to walk naturally and to take the first step with the right foot. Moreover, we taught them to walk a distance of 250 cm along a straight line.

The data collection was performed in accordance with the following procedure;

(1) A subject stands at the start point and walks along a straight line.
(2) Gait process is recorded by Kinect Studio [*1]
(3) After finishing all of the recording, the recorded data is played and segmented manually into the walking interval.
(4) Three-dimensional time series of each joint is obtained from the segmented interval data by the measurement program developed with Kinect SDK.

The number of trials is 10 times for each angle. Finally, 500 gait samples (= 5 angles × 10 subjects × 10 trials) are collected in total.

4.1.1 The Threshold Setting

To estimate the gait angles by the proposed algorithm, it is important to decide the lower bound and the upper bound thresholds for each gait angle ($\phi^g, \Phi^g,\ g = 90, 135, 180, 225, 270$). In this experiment the two types of thresholds are employed, the min-max thresholds based on the following equations:

$$\phi^g = \min\{\, \Psi^g{}_i,\ \ i = 1, 2, \cdots, N^g{}_s \,\}, \tag{11}$$

$$\Phi^g = \max\{\, \Psi^g{}_i,\ \ i = 1, 2, \cdots, N^g{}_s \,\}, \tag{12}$$

where g denotes a discrete gait angle index ($g = 90, 135, 180,$

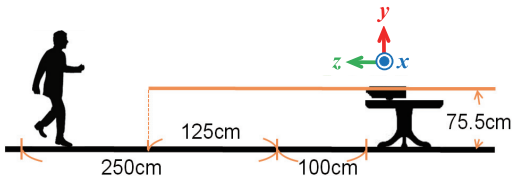


Fig. 4 Sketch of data measurement environment.

Table 1 Subject's height distribution.

Height Range (cm)	Frequency
161 to 165	4
166 to 170	3
171 to 175	2
176 to 180	1

[*1] Kinect Studio is a tool that helps you record and play back depth and color streams from a Kinect.

Table 2 Gait angle estimation result by min-max threshold (%).

		Predicted					
		90	135	180	225	270	Others
True	90	87.1	0.0	0.0	0.0	0.0	12.9
	135	0.0	84.9	0.0	0.0	0.0	15.1
	180	0.0	0.0	82.1	0.0	0.0	17.9
	225	0.0	0.0	0.0	89.2	0.0	10.8
	270	0.0	0.0	0.0	0.0	88.3	11.7

Table 3 Gait angle estimation result by σ threshold (%).

		Predicted					
		90	135	180	225	270	Others
True	90	70.1	0.0	0.0	0.0	0.0	29.9
	135	0.0	63.9	0.0	0.0	0.0	36.1
	180	0.0	0.0	68.3	0.0	0.0	31.7
	225	0.0	0.0	0.0	72.7	0.0	27.3
	270	0.0	0.0	0.0	0.0	67.1	32.9

Table 4 Gait angle estimation result by 2σ threshold (%).

		Predicted					
		90	135	180	225	270	Others
True	90	94.4	0.0	0.0	0.0	0.0	5.6
	135	0.0	95.3	0.0	0.0	0.0	4.7
	180	0.0	0.0	92.8	0.0	0.0	7.2
	225	0.0	0.0	0.0	90.1	0.0	9.9
	270	0.0	0.0	0.0	0.2	92.8	7.0

225, 270), i denotes the data sample index, ${\Psi^g}_i$ denotes the i-th cosine value computed from the training data samples of the gait angle g and ${N^g}_s$ is the number of the training samples of the gait angle g. Secondly the sigma thresholds based on the following equations are used.

$$\phi^g = \overline{\Psi^g} - \omega\sigma^g, \tag{13}$$

$$\Phi^g = \overline{\Psi^g} + \omega\sigma^g, \tag{14}$$

where $\overline{\Psi^g}$ denotes the mean of C^g based on $\overline{\Psi^g} = \frac{1}{N^g{}_s}\sum_{i=1}^{N^g{}_s}{\Psi^g}_i$, σ^g denotes the standard deviation based on $\sigma^g = \sqrt{\frac{1}{N^g{}_s}\sum_{i=1}^{N^g{}_s}({\Psi^g}_i - \overline{\Psi^g})^2}$ and ω is the weight of σ^g. In this experiment we set $\omega = 1, 2, 3, 4$.

4.1.2 Estimation Accuracy Evaluation

The gait angle estimation accuracy has been evaluated by 10-fold cross validation. We divided the 500 gait samples into 50 samples for training data and the other 450 samples for test data. The 50 training samples include 10 samples per a gait, thus they are composed of 10 samples/angle × 5 gait angles. The 450 test samples are composed of 90 samples/angle × 5 gait angles. Estimation accuracy for an angle is calculated by the following equation:

$$Accuracy(\%) = \frac{\#\ of\ Gaits\ Estimated\ Correctly}{90\ samples \times 10\ trials} \times 100 \tag{15}$$

4.1.3 Result

We compared with five thresholds, the Min-Max threshold and the sigma thresholds ($\omega = 1, 2, 3, 4$). From **Table 2**, **Table 3**, **Table 4**, **Table 5**, and **Table 6**, we show the confusion matrices of the gait angle estimation result for each threshold.

In Table 3 the worst accuracy is obtained by σ thresholds. The average estimation accuracy is 68.4%. From this Table, we can find that many gait angles are classified into *Others* class. This means that σ is too narrow range as the threshold to estimate the gait angles.

Table 5 Gait angle estimation result by 3σ threshold (%).

		Predicted					
		90	135	180	225	270	Others
True	90	97.9	0.7	0.0	0.0	0.0	1.4
	135	0.8	99.1	0.0	0.0	0.0	0.1
	180	0.0	0.0	96.1	0.0	0.0	3.9
	225	0.0	0.0	0.0	99.2	0.0	0.8
	270	0.0	0.0	0.0	4.8	94.9	0.3

Table 6 Gait angle estimation result by 4σ threshold (%).

		Predicted					
		90	135	180	225	270	Others
True	90	99.6	0.4	0.0	0.0	0.0	0.0
	135	11.8	88.2	0.0	0.0	0.0	0.0
	180	0.0	0.0	97.8	0.0	0.0	2.2
	225	0.0	0.0	0.0	100.0	0.0	0.0
	270	0.0	0.0	0.0	12.6	87.4	0.0

The best result is obtained by 3σ thresholds in Table 5. The accuracy ranges from 94.9% to 99.1%. For all gait angles, almost 95% and more accuracy were achieved. The average estimation accuracy of gait angles is 97.4%.

In the rest of the thresholds, the average estimation accuracies are 86.3% (min-max threshold) in Table 2, 93.1% (2σ threshold) in Table 4 and 94.6% (4σ threshold) in Table 6. In the results with threshold values 2σ and 4σ, accuracy is over 90% in each case and it seems to be good performance. However in 4σ thresholds, we can find that many gait angles are misclassified to incorrect gait angles. Thus 4σ is too wide as the threshold.

Finally we can find that the best threshold to estimate gait angles is in the range between 2σ and 3σ. Also in this experiment, we can obtain 97.4% as the best average of estimation accuracy when the threshold is 3σ.

5. Experiment II: Gait Recognition

Gait recognition accuracy has been examined by the simulation experiment using the collected gait data. Generally, human recognition task is classified into verification tasks and identification tasks. In this study, we treat the human identification task. According to Vielhauer [24], the identification tasks are defined as follows: "*the classification will assign the biometric features to one out of all classes of persons, registered with a particular system*". Before gait recognition, the optimal feature vectors for each gait angle are evaluated and selected.

5.1 Optimal Feature Selection for Gait Angles

In order to recognize the individual gait successfully, we select the optimal soft biometric features. As is mentioned before, Fisher criterion for feature evaluation have been used. **Figure 5** to 9 represent the Fisher criteria of the four soft biometric features for each joint for different angles respectively. The horizontal axes denote the joint IDs and the vertical axes denote the Fisher criteria based on the Eq. (8). The height of each bar represents the distinctiveness of features for each joint.

In Fig. 5 (for 90 degrees angle), β and δ are large values in HIP_LEFT and HIP_RIGHT. It seems that there are many large values in all of the features, α, β, γ and δ. Besides, the right arm joints (IDs are 12-15) are relatively small values. The right side of body is opposite side of the sensor device for this angle.

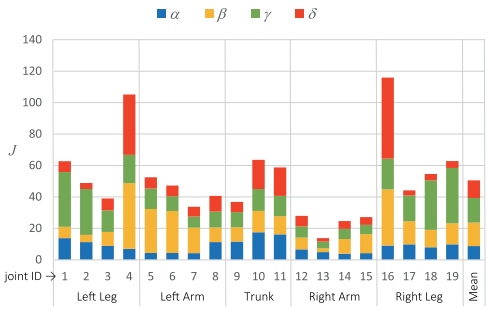


Fig. 5 Fisher criterion for each joint (90 degrees).

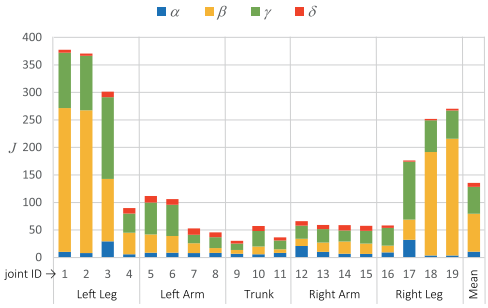


Fig. 6 Fisher criterion for each joint (135 degrees).

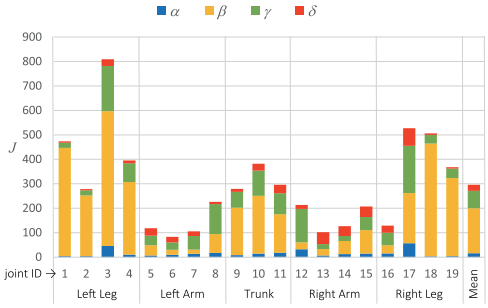


Fig. 7 Fisher criterion for each joint (180 degrees).

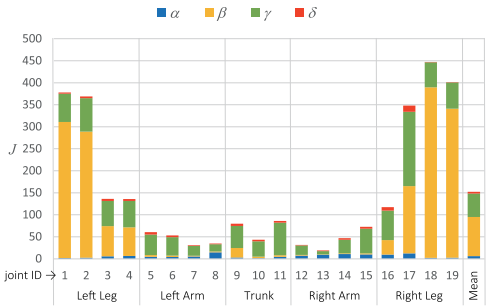


Fig. 8 Fisher criterion for each joint (225 degrees).

This implies that the features related to the right arm have less information to identify due to the occlusion.

In **Fig. 6** (for angle of 135 degrees), the legs' joints have large values, but the other joints have small values. Also, β and γ indicate large values, but α and δ is much smaller than them.

In **Fig. 7** (for angle of 180 degrees), the legs' and the trunk's joints have large values but the arms' joints are relatively small. α and δ is extremely small compared with β and γ. It seems that there is no contribution of α and δ to identify the gait data for this angle.

In **Fig. 8** (for angle of 225 degrees), the legs' joints have large values but the other joints have smaller values. This tendency is

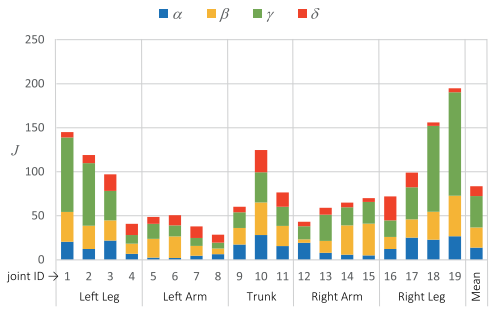


Fig. 9 Fisher criterion for each joint (270 degrees).

Table 7 Selected features and joint ID for each gait angle.

Features	Gait Angles				
	90	135	180	225	270
α	10, 11	17	—	—	10, 19
β	4, 5, 6, 7, 16, 17	1, 2, 3, 4, 5, 17, 18, 19	1, 2, 3, 4, 8, 9, 10, 11, 15, 17, 18, 19	1, 2, 3, 4, 17, 18, 19	1, 2, 10, 14, 15, 18, 19
γ	1, 2, 4, 16, 17, 18, 19	1, 2, 3, 4, 5, 6, 16, 17, 18, 19	3, 4, 8, 10 11, 12, 17	1, 2, 3, 4, 5, 9, 11, 15, 16, 17, 18, 19	1, 2, 3, 10, 13, 17, 18, 19
δ	4, 10, 11, 16	—	—	—	10, 16

same as the one in 135 degrees. Also, it is suggested that α and δ have no informative joints to identify the gait. This graph shows that β and γ values can only be used to identify gait data in 225 degrees angle.

In **Fig. 9** (for angle of 270 degrees), it seems that the four features (α, β, γ and δ) have some informative joints respectively. Also, the left arm joints (IDs are 5-8) are relatively small values compared with the other joints. In this angle the left side of body is opposite side of the sensor device, so this has the same tendency as the result for 90 degrees.

From the range of the vertical axes in all of the graphs, it can be found that the more accurately the RGB-D camera captures the front view gait, the larger the value of J becomes. Besides, in the angles of 135, 180 and 225 degrees, it is suggested that the better results are obtained by β or γ. On the other hand, in case of the lateral view gait (when the front view gait cannot be captured), it is necessary to use α and δ as well as β and γ.

Finally **Table 7** shows the selected features and joints for each gait angle. The numbers mean the selected joint IDs for each feature. We selected the top 19 features ($M = 19$) based on J so as to be same as the size of a single feature vector dimension. For example, in the 90 degrees, we selected 10th and 11th joints for α, 4th, 5th, 6th, 7th, 16th and 17th joints for β, 1st, 2nd, 4th, 16th, 17th, 18th and 19th joints for γ, 4th, 10th, 11th and 16th joints for δ. Then the selected feature vector $\boldsymbol{F}_{90}$ is represented as follows;

$$\begin{aligned}\boldsymbol{F}_{90} := (&\alpha^{10}, \alpha^{11}, \beta^{4}, \beta^{5}, \beta^{6}, \beta^{7}, \beta^{16}, \beta^{17}, \gamma^{1}, \gamma^{2}, \gamma^{4}, \\ &\gamma^{16}, \gamma^{17}, \gamma^{18}, \gamma^{19}, \delta^{4}, \delta^{10}, \delta^{11}, \delta^{16}).\end{aligned} \tag{16}$$

Here this feature vector, we can find that the joints of the right arm were not selected in the 90 degrees. So we can discard the unstable and the low discriminant features by using this selection strategy. Thus it is expected that a good recognition performance is obtained by this feature vector.

5.2 Gait Recognition and Its Accuracy Evaluation

In this experiment, we employed linear discriminant analysis

Table 8 Gait recognition accuracy (%).

Features	Gait Angles					Mean
	90	135	180	225	270	
α	72±17	82±15	89±10	73±18	83±15	79.8±6.4
β	86±13	96±7	98±4	96±7	86±13	92.4±5.3
γ	91±12	97±5	98±4	97±5	92±7	95.0±2.9
δ	64±24	71±18	96±5	83±16	85±11	79.8±11.2
Selected	97±5	89±10	88±16	93±8	93±8	92.0±3.2
Best	97±5	97±5	98±4	97±5	93±8	96.4±1.7

(LDA) methods as gait recognizers. Specifically, we use `lda` function in `MASS` library in the statistical computing tool R [25].

As the gait recognition task, 10-fold cross validation was performed for each gait angle. Data set is same as one used in the previous section. As the number of data samples for a gait angle is 100 (10 samples/subject × 10 human subjects), a validation trial is composed of 90 training samples and 10 test samples. Each training and test samples contains 10 subjects data. Thus the recognition accuracy for a human subject is computed by the following equation;

$$Accuracy(\%) = \frac{\#\ of\ Gaits\ Recognized\ Correctly}{10\ trials} \times 100. \tag{17}$$

Moreover, for comparison of feature vectors, five kinds of feature sets, that is α of all joints, β of all joints, γ of all joints, δ of all joints and selected features based on Table 7, are used.

5.3 Result

Table 8 shows gait recognition accuracy for each angle and feature. We can find that the recognition accuracies of β and γ outperform the ones of α and δ. In α and δ, the average accuracies do not reach 80% but the ones of β and γ are over 90%. It is considered that this tendency reflects the result of feature evaluation.

The best average is obtained as 95.0% by γ, the second highest average is obtained as 92.5% by β. However, β and γ, can not yield high recognition accuracy for lateral gaits (90 and 270 degrees) compared with the other gait angles (135, 180 and 225 degrees). This means that it is difficult to obtain the good performance result when the front view gait cannot be captured.

On the other hand, the selected feature set produce high performance for the lateral gaits (90 and 270 degrees), while produce poor results in other gait angles (135, 180 and 225 degrees). The selected features in the lateral view gaits include more α and δ than ones in other gaits. It is suggested that the use of α and δ complements the information shortage caused by the occlusion of the joints.

Consequently, we found that β and γ are useful features when the front view gait could be captured. On the other hand, when the lateral view gait can be captured, we found that α and δ were required to complement the information shortage caused by the occlusion of the joints.

Finally, in case of feature selection optimally for each gait angle, the average of the best recognition accuracy reached 96.4%.

5.4 Summarization of Two Experimental Results

Table 9 shows the summarization of the best results of the gait angle estimation and the gait recognition. In addition we computed the estimated accuracy of two stage process which is to estimate the gait angle at the first stage and subsequently recognize human gait. The estimated value is computed by multiplying each best angle estimation accuracy and each best gait recognition accuracy together. The best gait recognition result with two stage process was estimated as 93.9% in average.

Table 9 Estimated two stage process accuracy (%).

	Gait Angles					Mean
	90	135	180	225	270	
Angle Estimation	97.9	99.1	96.1	99.2	94.9	97.4
Gait Recognition	97.0	97.0	98.0	97.0	93.0	96.4
Two Stage Process	95.0	96.1	94.2	96.2	88.3	93.9

Table 10 Gait recognition accuracy without angle estimation (%).

	α	β	γ	δ	Mean
Baseline	50.2	70.0	77.8	44.0	60.5

Table 10 shows the result of the baseline recognition method, which is computed by LDA without the gait angle estimation. In this computation, 10-fold cross validation was performed for each feature vector. As the number of data samples is 500 (10 samples/subject × 10 subjects × 5 gait angles), a validation trial is composed of 450 training samples (45 samples/subject × 10 subjects) and 50 test samples (5 samples/subject × 10 subjects). We calculated the following gait recognition accuracy for a person;

$$Accuracy(\%) = \frac{\#\ of\ Gaits\ Recognized\ Correctly}{5\ samples \times 10\ trials} \times 100. \tag{18}$$

The average of recognition accuracy is 60.5%. We can find that the accuracies of each feature are quite under-performed results, compared with our proposed method. We have been able to achieve the high performance of multi-angle gait recognition by the two stage process.

5.5 Comparison to the Related Works

In this section, we compare the recognition accuracy of our proposed method with the related works, which employ the gait data based on RGB-D camera.

Chattopadhyay et al. [21] proposed a frontal gait recognition method by using two RGB-D cameras. They consider the scenario of airport security check and the RGB-D cameras are installed above the security gate. The extracted features are some soft biometric measurements of human body, skeleton kinematic features and Gait Energy Image (GEI) of back view. The experiment was the identification task using the data set of 60 human subjects. As the result, it is reported that the recognition accuracy is less than some 70%. Their method employs a lot of features, but the recognition accuracy is lower than our proposed method. The reasons are (1) the setting of the RGB-D camera and (2) the number of human subjects. In their study, the gait data is measured above the head of human. According to Matsumoto et al. [26], it is reported that the measurement from the horizontal angle position of the RGB-D camera outperforms the measurement of the upper angle position on the verification accuracy because the measurement error tends to become large in the upper position. In addition, they computed the recognition accu-

racy based on 60 human subjects, which is more than the size of our experiment. Actually soft biometric features lack distinctiveness and permanence, so the larger number of human is, the more difficult the recognition task becomes. This result suggests it may be difficult to identify over 50 individuals by using soft biometric measurements of human body extracted from gait data.

Wei et al. [22] proposed a soft biometric authentication method based on neural network. They estimated some physical measurements such as the height and arm span from the frontal view gait data measured by RGB-D camera. In their study, two verification tasks are carried out for 10 human subjects data set, the experiment I is based on 1 genuine person and 9 forgery, the experiment II is based on 3 genuine persons and 7 forgery. As the result, it is reported that FRR is 3.15% and FAR is 2.45% in the experiment I, FRR is 8.69% and FAR is 10.41% in the experiment II. In this experiment, the number of human subjects is same size as our task, and gait data are measured from the horizontal direction, which is almost same as our proposed method. In the recognition accuracy, our proposed method outperforms their method, but it is to be noted that the experiment task is different from our task.

6. Conclusions

In order to realize that the computer system can naturally identify the person, it is necessary to develop a recognition technology of individuals and ADL on the basis of soft biometric traits obtained by the depth camera devices, which have been established as the next generation camera devices.

In this study, human gait as a soft biometric trait has been focused and a multi-angle gait recognition method based on skeletal tracking data measured by RGB-D camera has been proposed. Our method has two stage process that estimates an optimal gait angle view from the five discrete angles at the first stage and subsequently recognizes human gait based on the specific features for the respective gait angle views.

The two simulation experiments have been performed to evaluate the proposed method. The first experiment was gait angle estimation. By using the proposed algorithm, the best estimation of 97.4% accuracy was achieved on average of five gait angles. The second experiment was gait recognition. In this experiment, we evaluated four soft biometric features to select the optimal one. As the result of feature selection, we found that β and γ are useful features when the front view gait could be captured. On the other hand, when the front view gait cannot be captured, we found that α and δ were required. The gait recognition has been performed as 10-fold cross validation task. From the results of the experiments with data from 10 human subjects, the best gait recognition of 96.4% accuracy was achieved by Linear Discriminant Analysis. The best gait recognition result with two stage process was estimated as 93.9%. Consequently we have been able to achieve the high performance of multi-angle gait recognition.

Soft biometric features lack distinctiveness and permanence, so the larger number of human is, the more difficult the recognition task becomes. From the results of the related works, we estimated that it may be difficult to identify over 50 individuals by using soft biometric measurements of human body extracted from gait data. However we consider that our method can be applied to the identification task of a family member. Actually in such application, we estimate the number of target is some 10 persons. In addition, our proposed method can be integrated the conventional recognition technologies. Thus, we consider that it is possible to identify over 50 individuals by combining our method and them.

In the near future, we would like to carry out other experiments with a large amount of data. Especially we would like to apply our proposed method to the human verification task to evaluate false acceptance rate as well as false rejection rate.

References

[1] Sumi, K.: Biometrics Identification as a Pattern Recognition Problem, *Proc. Biometrics Workshop*, pp.5–6 (2012) (in Japanese).

[2] Jain, A.K., Dass, S.C. and Nandakumar, K.: Soft biometric traits for personal recognition systems, *Proc. International Conference on Biometric Authentication*, LNCS 3072, pp.731–738 (2004).

[3] Kinect for Windows: available from ⟨https://www.microsoft.com/en-us/kinectforwindows/⟩ (accessed 2015-05-07).

[4] Xtion Pro, available from ⟨http://www.asus.com/Multimedia/Xtion_PRO/⟩ (accessed 2015-05-07).

[5] Depth Sense Cameras, available from ⟨http://www.softkinetic.com/Products/DepthSenseCameras⟩ (accessed 2015-05-07).

[6] Leap Motion, available from ⟨https://www.leapmotion.com/⟩ (accessed 2015-05-07).

[7] Shao, L., Han, J., Xu, D. and Shotton, J.: Computer vision for RGB-D sensors: Kinect and its applications, *IEEE Trans. Cybernetics*, Vol.43, No.5, pp.1314–1317 (2013).

[8] Han, J., Shao, L., Xu, D. and Shotton, J.: Enhanced computer vision with microsoft kinect sensor: A review, *IEEE Trans. Cybernetics*, Vol.43, No.5, pp.1318–1334 (2013).

[9] Haggag, H., Hossny, M., Filippidis, D., Creighton, D., Nahavandi, S. and Puri, V.: Measuring depth accuracy in RGBD cameras, *7th International Conference on Signal Processing and Communication Systems*, pp.1–7 (2013).

[10] Weichert, F., Bachmann, D., Rudak, B. and Fisseler, D.: Analysis of the Accuracy and Robustness of the Leap Motion Controller, *Sensors*, Vol.13, No.5, pp.6380–6393 (2013).

[11] Guna, J., Jakus, G., Pogačnik, M., Tomažič, S. and Sodnik, J.: An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking, *Sensors*, Vol.14, No.2, pp.3702–3720 (2014).

[12] Murray, M.P., Drought, A.B. and Kory, R.C.: Walking patterns of normal men, *Journal of Bone Joint Surg*, Vol.46-A, No.2, pp.335–360 (1964).

[13] Niyogi, S.A. and Adelson, E.H.: Analyzing and recognizing walking figures in XYT, *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.469–474 (1994).

[14] Johnson, A.Y. and Bobick, A.F.: A multi-view method for gait recognition using static body parameters, *Audio-and Video-Based Biometric Person Authentication*, pp.301–311, Springer (2001).

[15] Seely, R.D., Goffredo, M., Carter, J.N. and Nixon, M.S.: View invariant gait recognition, *Handbook of Remote Biometrics*, pp.61–81, Springer London (2009).

[16] Soriano, M., Araullo, A. and Saloma C.: Curve spreads-a biometric from front-view gait video, *Pattern Recognition Letters*, Vol.25, No.14, pp.1595–1602 (2004).

[17] Goffredo, M., Carter, J.N. and Nixon, M.S.: Front-view gait recognition, *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems*, pp.1–6 (2008).

[18] Ryu, J. and Kamata, S.: Front view gait recognition using spherical space model with human point clouds, *18th IEEE International Conference on Image Processing* (*ICIP*), pp.3209–3212 (2011).

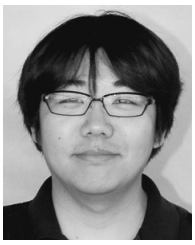
[19] Manabe, Y., Saito, R. and Sugawara, K.: Biometric Gait Verification by Horizontal Swings in Frontal Manner towards Human-Aware Environment, *IEEE 11th International Conference on Cognitive Informatics and Cognitive Computing*, pp.219–225 (2012).

[20] Manabe, Y., Saito, R., Shimada, Y. and Sugawara, K.: Frontal View Person Verification by Soft Biometric Features of Gait, Face and Body, *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, Vol.24, No.5, pp.988–1001 (2012) (in Japanese).

[21] Chattopadhyay, P., Sural, S. and Mukherjee, J.: Frontal Gait Recognition from Incomplete Sequences using RGB-D Camera, *IEEE Trans. Information Forensics and Security*, Vol.9, No.11, pp.1843–1856 (2014).

[22] Wei, Z. and Chakraborty, G.: Construction of intelligent intrusion detection system based on KINECT, *International Joint Conference*

on Awareness Science and Technology and Ubi-Media Computing (*iCAST-UMEDIA*), pp.81–87 (2013).
[23] Tee, C., Goh, M.K.O. and Teoh, A.B.J.: Personal Recognition Using Multi-angles Gait Sequences, *Digital Information Processing and Communications*, pp.497–508, Springer, Berlin Heidelberg (2011).
[24] Vielhauer, C.: Biometric user authentication for IT security: From fundamentals to handwriting, *Advances in Information Security*, Vol.18, Springer (2005).
[25] The R Project for Statistical Computing, available from ⟨http://www.r-project.org⟩ (accessed 2015-06-04).
[26] Matsumoto, K., Manabe, Y. and Sugawara, K.: A Study on Optimal Camera Positions in Gait Vefrification Using Body Skeletal Information, 27th Annual Conference of the Japanese Society for Artificial Intelligence, 2H1-5, 4 pages (2013) (in Japanese).



Yusuke Manabe is an associate professor of the Department of Information and Network Science, Chiba Institute of Technology, Japan. He received his Ph.D. degree in software and information science from Iwate Prefectural University in 2008. His research interests include Intelligent and Cognitive Informatics, Awareness Computing and Soft Biometrics. He is a member of IEEE, SOFT, JSAI, JCSS and IEICE.

Keisuke Matsumoto was born in 1990. He received his M.Eng. degree from Chiba Institute of Technology in 2015. His research interest is behavior-based biometrics, especially gait recognition.

Kenji Sugawara left his doctoral course in the Graduate School of Tohoku University in 1980. Currently, he is a Professor in the Department of Information and Network Science at Chiba Institute of Technology. He holds a Ph.D. degree in Engineering. In 1994, he received the Yamashita Commemoration Prize of the IPSJ; in 1996, the Paper Prize of the IPSJ; and in 2000, the Achievement Award of the IEICE. He is a fellow of the IEICE and a member of IEEE.