We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Gait Recognition

Jiande Sun, Yufei Wang and Jing Li

Additional information is available at the end of the chapter

Abstract

Gait recognition has received increasing attention as a remote biometric identification technology, i.e. it can achieve identification at the long distance that few other identification technologies can work. It shows enormous potential to apply in the field of criminal investigation, medical treatment, identity recognition, human-computer interaction and so on. In this chapter, we introduce the state-of-the-art gait recognition techniques, which include 3D-based and 2D-based methods, in the first part. And considering the advantages of 3D-based methods, their related datasets are introduced as well as our gait database with both 2D silhouette images and 3D joints information in the second part. Given our gait dataset, a human walking model and the corresponding static and dynamic feature extraction are presented, which are verified to be view-invariant, in the third part. And some gait-based applications are introduced.

Keywords: gait recognition, gait dataset, 2D-based, 3D-based, view invariant

1. Introduction

Gait recognition has been paid lots of attention as one of the biometric identification technologies. There have been considerable theories supporting that person's walking style is a unique behavioural characteristic, which can be used as a biometric. Differing from other biometric identification technologies such as face recognition, gait recognition is widely known as the most important non-contactable, non-invasive biometric identification technology, which is hard to imitate. Since these advantages, gait recognition is expected to be applied in scenarios, such as criminal investigation and access control. Usually gait recognition includes the following five steps, which are shown in **Figure 1**.
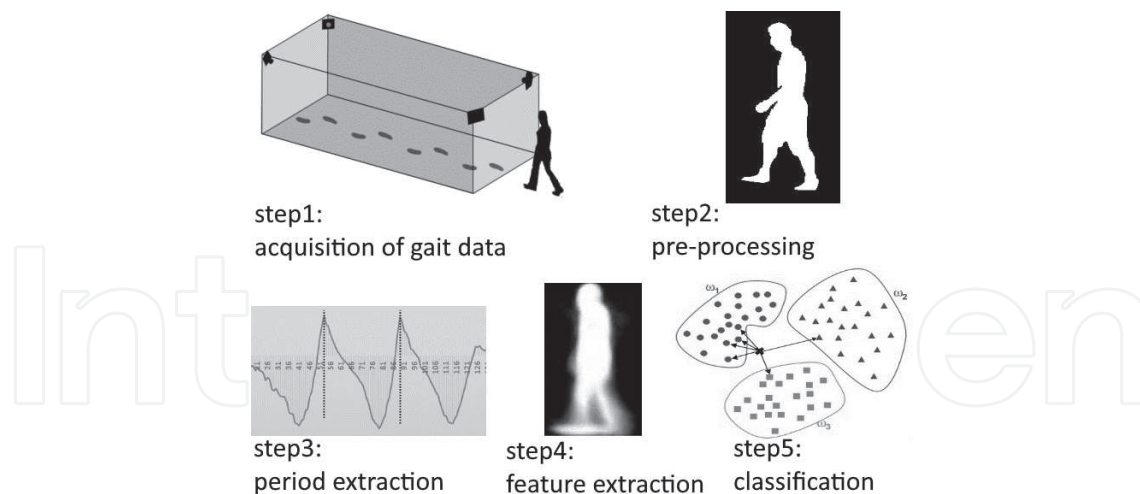
Figure 1. The steps in gait recognition.

Step 1: acquisition of gait data

The ways of acquiring the original gait data depend on how to recognize the gait. Usually the gait is acquired by single camera, multiple cameras, professional motion capture system (e.g. VICON) and camera with depth sensor (e.g. Kinect).

Step 2: pre-processing

The methods of pre-processing are quite different corresponding to the terms of acquiring gait. For instance, in some single camera-based methods, the pre-processing is usually the background subtraction, which is to get the body silhouette of walking people. However, in Kinect-based methods, the pre-processing is to filter the noise out of the skeleton sequences.

Step 3: period extraction

Since human gait is a kind of periodic signal, a gait sequence may include several gait cycles. Gait period extraction is helpful to reduce the data redundancy because all the gait features can be included in one whole gait cycle.

Step 4: feature extraction

Various gait features are used in different kinds of gait recognition methods and they influence the performance of gait recognition. Gait features can be divided into hand-crafted and machine-learned features. The hand-crafted ones are easy to be generalized to different datasets, while the machine-learned ones are usually better for the specific dataset.

Step 5: classification

Gait classification, i.e. gait recognition, is to use the classifiers based on the gait features. The classifiers range from the traditional one, such as kNN (k-nearest neighbour), to the modern one, such as deep neural network, which has achieved success in face recognition, handwriting recognition, speech recognition, etc.

Generally, gait recognition methods can be divided into 3D-based and 2D-based ones. The 2D-based gait recognition methods depend on the human silhouette captured by one 2D camera, which is the normal situation of the video surveillance. The 2D-based gait recognition methods are dominant in this field of gait recognition and they are usually divided into model-based and model-free methods.

The model-based methods extract the information of the shape and dynamics of the human body from video sequences, establish the suitable skeleton or joint model by integrating the information and classify the individuals based on the variation of the parameters in such a model. Cunado et al. [1] modelled gait as an articulated pendulum and extracted the line via the dynamic Hough transform to represent the thigh in each frame, as shown in **Figure 2a**. Johnson et al. [2] identified the people based on the static body parameters recovered from the walking action across multiple views, which can reduce the influence introduced by variation in view angle, as shown in **Figure 2b**. Guo et al. [3] modelled the human body structure from the silhouette by stick figure model, which had 10 sticks articulated with six joints, as shown in **Figure 2c**. Using this model, the human motion can be recorded as a sequence of stick figure parameters, which can be the input of BP neural network. Rohr [4] proposed a volumetric model for the analysis of human motion, using 14 elliptical cylinders to model the human body, as shown in **Figure 2d**. Tanawongsuwan et al. [5] projected the trajectories of lower body joint angles into walking plane and made them time-normalized by dynamic time warping (DTW). Wang et al. [6] made a fusion between the static and dynamic body features. Specifically, the static body feature is in a form of a compact representation obtained by Procrustes shape analysis. The dynamic body feature is extracted via a model-based approach, which can track the subject and recover joint-angle trajectories of lower limbs, as shown in **Figure 2e**. Generally, the model-based gait recognition methods have better invariant properties and are better at handling occlusion, noise, scaling and view-variation. However, model-based methods usually require a high resolution and a heavy computational cost.

On the other hand, model-free methods generate gait signatures directly based on the silhouettes, which are extracted from the video sequences, without fitting a model. Gait energy image (GEI) [7] is the most popular gait representation, which represents the spatial and temporal gait information in a grey image. GEI is generated by averaging silhouettes over a complete gait cycle and represents human motion sequence in a single image while preserving

(a)   (b)   (c)   (d)   (e)
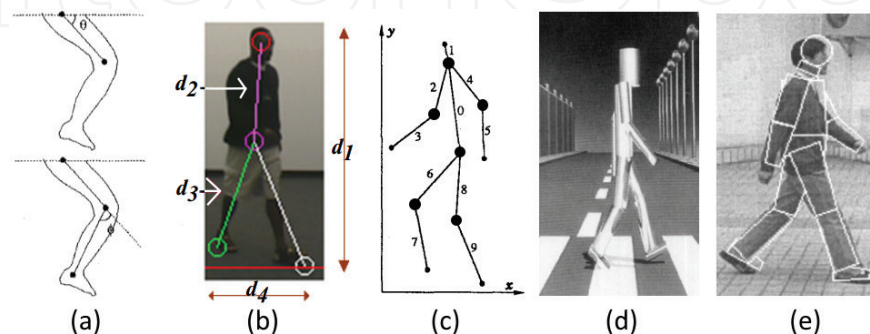
Figure 2. Examples of model-based methods.

the temporal information, as shown in **Figure 3a**. Motion silhouette image (MSI) [8] is like GEI and is a grey image too. The intensity of an MSI is determined by a function of the temporal history of the motion of each pixel, as shown in **Figure 3b**. The intensity of an MSI represents motion information during one gait cycle. Because GEI and MSI represent both motion and appearance information, they are sensitive to the changes in various covariate conditions such as carrying and clothing. Shape variation-based (SVB) frieze pattern is proposed in [9], as shown in **Figure 3c**, to improve their robustness against these changes. SVB frieze pattern projects the silhouettes horizontally and vertically to represent the gait information, and uses key frame subtraction to reduce the effects of appearance changes on the silhouettes. Although it has been shown that SVB frieze pattern can get better results when there are significant appearance changes, it does not outperform in the case of no changes, and it requires temporal alignment pre-processing for each gait cycle, which brings more computation load. Gait entropy image (GEnI) [10] is another gait representation, which is based on Shannon entropy. It encodes the randomness of pixel values in the silhouette images over a complete gait cycle, and it is more robust to appearance changes, such as carrying and clothing, as shown in **Figure 3d**. Wang et al. [11] propose the Chrono-Gait image (CGI), as shown in **Figure 3e**, to compress the silhouette images without losing too much temporal relationship between them. They utilize a colour mapping function to encode each gait contour image in the same gait sequence, and average over a quarter gait cycle to one CGI. It is helpful to preserve more temporal information of a gait cycle.

The methods mentioned above are all convert the gait sequence into a single image/template. There are other methods to keep temporal information of gait sequences, which have good performance too. Sundaresan et al. [12] propose the gait recognition methods based on hidden Markov models (HMMs) because the gait sequence is composed of a sequence of postures, which is suitable for HMM representation. In this method, the postures are regarded as the states of the HMM and are identical to individuals, which provide a means of discrimination. Wang et al. [13] apply principal component analysis (PCA) to extract statistical spatio-temporal features from the silhouette sequence and recogniz gait in the low-dimensional eigenspace via supervised pattern classification techniques. Sudeep et al. [14] utilize the correlation of sequence pairs to preserve the spatio-temporal relationship between the galley and probe sequences, and use it as the baseline for gait recognition.

The advantages of model-free methods are computational efficiency and simplicity; however, the robustness against the variations of illumination, clothing, scaling and views still needs to be improved. Here, we focus on the view-invariant gait recognition methods.

(a)   (b)   (c)   (d)   (e)
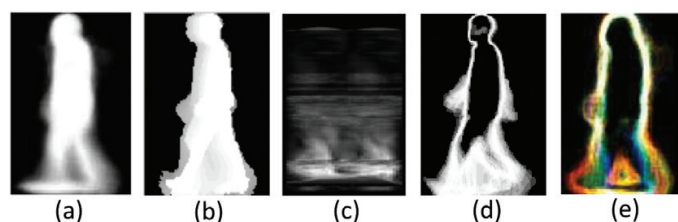
**Figure 3.** Examples of model-free methods.

Up to date, 2D-based view-invariant gait recognition methods can be divided into pose-free and pose-based ones. The pose-free methods aim at extracting the gait parameters independent from the view angle of the camera. Johnson et al. [2] present a gait recognition method to identify people based on static body parameters, which are extracted from the walking across multiple views. Abdelkader et al. [15] propose to extract an image template corresponding to the person's motion blob from each frame. Subsequently, the self-similarity of the obtained template sequence is computed. On the other hand, the pose-based method aims at synthesizing the lateral view of the human body from an arbitrary viewpoint. Kale et al. [16] show that if the person is far enough from the camera, it is possible to synthesize a side view from any of the other arbitrary views using a single camera. Goffredo et al. [17] use the human silhouette and human body anthropometric proportions to estimate the pose of lower limbs in the image reference system with low computational cost. After a marker-less motion estimation, the trends of the obtained angles are corrected by the viewpoint-independent gait reconstruction algorithm, which can reconstruct the pose of limbs in the sagittal plane for identification. Muramatsu et al. [18] propose an arbitrary view transformation model (AVTM) for cross-view gait matching. 3D gait volume sequences of training subjects are constructed, and then 2D gait silhouette sequences of the training subjects are generated by projecting the 3D gait volume sequences onto the same views as the target views. Finally, the AVTM is trained with gait features extracted from the 2D sequences. In the latest work [19], the deep convolutional neural networks (CNNs) is established and trained with a group of labelled multi-view human walking videos to carry out a gait-based human identification via similarity learning. The method is evaluated on the CASIA-B, OU-ISIR and USF dataset and performed outstanding comparing with the previous state-of-the-art methods.

It can be seen from the above-mentioned methods that the main idea of 2D view-invariant methods is to find the identical gait parameters that are independent from the camera point of view or can be used to synthesize a lateral view with arbitrary viewpoint.

2. 3D-based gait recognition and dataset

2.1. 3D-based gait recognition

3D-based methods have the instinctive superiority in the robustness against view variation. Generally, multiple calibrated cameras or cameras with depth sensors are used in 3D-based methods, which is necessary to extract gait features with 3D information. Zhao et al. [20] propose to build the 3D skeleton model based on 10 joints and 24 degrees of freedom (DOF) captured by multiple cameras, and the 3D information provides robustness to the changes of viewpoints, as shown in **Figure 4a**. Koichiro et al. [21] capture the dense 3D range gait from a projector-camera system, which can be used to recognize individuals at different poses, as shown in **Figure 4b**. Krzeszowski et al. [22] build a system with four calibrated and synchronized cameras, estimate the 3D motion using the video sequences and recognize the view-variant gaits based on marker-less 3D motion tracking, as shown in **Figure 4c**.
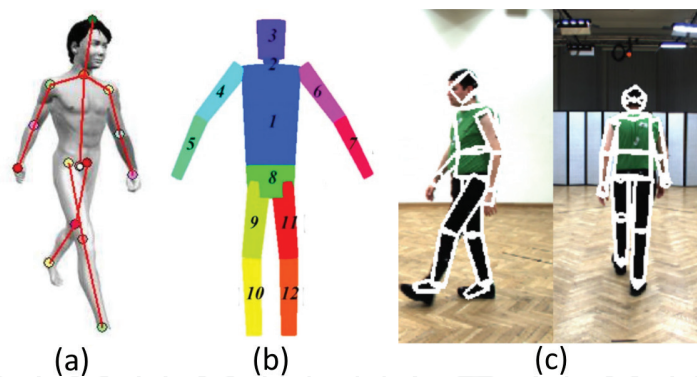
Figure 4. Examples of 3D-based methods.

3D-based methods are usually better than 2D-based view-invariant approaches in not only the recognition accuracy but also the robustness against view changing. However, these methods have high computational cost due to the calibration of multiple cameras and fusion of multiple videos.

The Microsoft Kinect brought about new strategies upon the traditional 3D-based gait recognition methods because it is a consumable RGB-D (Depth) sensor, which can provide depth information easily. So far, there are two generations of Kinect, which are shown in **Figure 5a** and **b**. Sivapalan et al. [23] extend the concept of the GEI from 2D to 3D with the depth images captured by Kinect. They average the sequences of registered three-dimensional volumes over a complete gait cycle, which is called gait energy volume (GEV), as shown in **Figure 6**. In Ref. [24], the depth information, which is represented by 3D point clouds, is integrated in a silhouette-based gait recognition scheme.

Another characteristic of Kinect is that it can precisely estimate and track the 3D position of joints at each frame via machine learning technology. **Figure 7a** and **b** shows the differences of tracking points between the first and second generation of Kinect.

Araujo et al. [25] calculate the length of the body parts derived from joint points as the static anthropometric information, and use it for gait recognition.Milovanovic et al. [26] use the coordinates of all the joints captured by Kinect to generate a RGB image, combine such RGB images into a video to represent the walking sequence, and identify the gait based on the

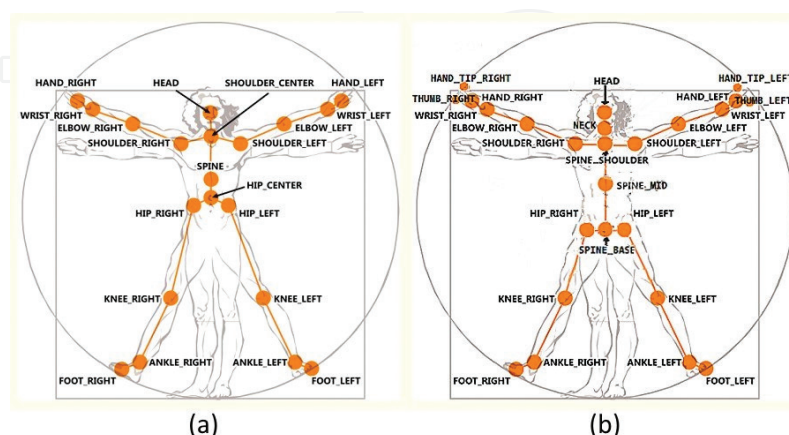**Figure 5.** The first and second generation kinects.
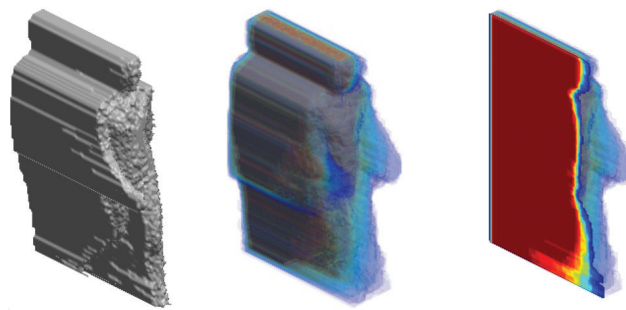
Figure 6. Gait energy volume (GEV).

spirit of content-based image retrieval (CBIR) technologies. Preis et al. [27] select 11 skeleton features captured by Kinect as the static feature, use the step length and speed as dynamic feature and integrate both static and dynamic features for recognition. Yang et al. [28] propose a novel gait representation called relative distance-based gait features, which can reserve the periodic characteristic of gait comparing with anthropometric features. Ahmed et al.[29] propose a gait signature using Kinect, which a sequence of joint relative angles (JRAs) is calculated over a complete gait cycle. They also introduce a new dynamic time warping (DTW)-based kernel to complete the dissimilarity measure between the train and test samples with JRA sequences. Kastaniotis et al. [30] propose a framework for gait-based recognition using Kinect. The captured pose sequences are expressed as angular vectors (Euler angles) of eight selected limbs. Then the angular vectors are mapped in the dissimilarity space resulting into a vector of dissimilarities. Finally, dissimilarity vectors of pose-sequences are modelled via sparse representation.

2.2. Dataset

Gait dataset is important to gait recognition performance improvement and evaluation. There are lots of gait datasets in the current academia and their purposes and characteristics

Figure 7. (a) The 20 joints tracked by first generation Kinect and (b) 25 joints tracked by second generation Kinect.

are different from each other. The differences among these datasets are mainly on the number of subjects, number of video sequences, covariate factors, viewpoints and environment (indoor or outdoor). Though the number of subjects in gait datasets is much smaller than that in the datasets of other biometrics (e.g. face, fingerprint, etc.), the current dataset can still satisfy the requirement of gait recognition method design and evaluation. Here, we give a brief introduction about several popular gait datasets. **Table 1** summarizes the information of these datasets.

SOTON Large Database [31] is a classical gait database containing 115 subjects, who are observed from side view and oblique view, and walk in several different environment, including indoor, treadmill and outdoor.

SOTON Temporal [32] contains the largest variations about time elapse. The gait sequences are captured monthly during 1 year with controlled and uncontrolled clothing conditions. It is suitable for purely investigating the time elapse effect on the gait recognition without regarding clothing conditions.

USF HumanID [14] is one of the most frequently used gait datasets. It contains 122 subjects, who walk along an ellipsoidal path outdoor, as well as contains a variety of covariates, including view, surface, shoes, bag and time elapse. This database is suitable for investigating the influence of each covariate on the gait recognition performance.

CASIA gait database contains three sets, i.e. A, B and C. Set A, also known as NLPR, is composed of 20 subjects, and each subject contains 12 sequences, which includes three walking directions, i.e. 0, 45 and 90°. Set B [33] contains large view variations from the front view to the rear view with 18° interval. There are 10 sequences for each subject, which are six normal sequences, two sequences with a long coat and two sequences with a backpack. Set B is suitable for evaluating cross-view gait recognition. Set C contains the infrared gait data of 153 subjects captured by infrared camera at night under 4 walking conditions, which are walk with normal speed, walk fast, walk slow and walk with carrying backpack.

OU-ISIR LP [34] contains the largest number of subjects, i.e. over 4000, with a wide age range from 1 year old to 94 years old and with an almost balanced gender ratio, although it does

Name	Subject	Sequence	Covariates	Viewpoints	In/Outhoor	Device
SOTON	115	2128	Yes	2	In/Outdoor	Camera(2D)
USF HumanID	122	1870	Yes	2	Outdoor	Camera(2D)
CASIA B	124	1240	Yes	11	Indoor	Camera(2D)
OU-ISIR,LP	4007	7842	No	2	Indoor	Camera(2D)
TUM-GAID	305	3370	Yes	1	Outdoor	Multimedia
KinectREID	71	483	yes	3	Indoor	Kinect

Table 1. List of popular gait datasets.

not contain any covariate. It is suitable for estimating a sort of upper bound accuracy of the gait recognition with high statistical reliability. It is also suitable for evaluating gait-based age estimation.

TUM-GAID [35] is the first multi-model gait database, which contains gait audio signals, RGB gait images and depth body images obtained by Kinect.

KinectREID [36] is a Kinect-based dataset that includes 483 video sequences of 71 individuals under different lighting conditions and 3 view directions (frontal, rear and lateral). Although the original motivation is for person re-identification, all the video sequences are taken for each subject by using Kinect, which contains all the information Kinect provided and is convenient for other Kinect SDK-based applications.

According to the overview [37] about the gait dataset, most of datasets are based on 2D videos or based on 3D motion data captured by professional camera, such as VICON. To our best knowledge, there are a few gait datasets containing both 2D silhouette images and 3D joints position information. Such a dataset can make the joint position-based methods, such as the method in Ref. [17] directly use the joint positions captured by Kinect, which can make use of both advantages of 2D- and 3D-based methods and bring improvement to the recognition performance. Meanwhile, the Kinect-based method such as in Refs. [25–28] will have a uniform platform to compare with each other. Therefore, a novel database based on Kinect is built, whose characteristics are following:

1. Two Kinects are used for simultaneously obtaining the 3D position of 21 joints (excluding 4 finger joints) and the corresponding binarized body silhouette images of each frame, as shown in **Figure 8**;

2. There are 52 subjects in the dataset, where each subject has 20 gait sequences and totally 1040 gait sequences;

3. Each subject has in six fixed and two arbitrary walking directions, which can be used to investigate the influence of view variation on the performance of gait recognition;

4. There are 28 males and 24 females with an average age of 22 in the dataset. There is no limitation for wearing, though most subjects wear shorts and T-shirts, and few females wear dress and high-heeled shoes, which is recorded in a basic information file.

The reason we choose Kinect V2 is that Kinect V2 has the comprehensive improvement over its first generation, such as broader field of view, higher resolution of colour and depth image, and more joints recognition ability. The 3D data and 2D RGB images are recorded, as shown in **Figure 8**. The upper area in **Figure 8** shows the 3D position of 21 joints, which means each joint will have a coordinate like (x, y, z) at each frame. We record all these original 3D position data at each frame during whole walking cycle. The lower area shows the corresponding binary silhouette image sequence after subtracting the subject from background.

Figure 8. Two kinds of data in our database: 3D position of 21 joints in the upper area and the corresponding binarized silhouette images in the lower area.

The experimental environment is shown in **Figure 9**. Two Kinects are located mutually perpendicular at the distance of 2.5 m to form the biggest visual field, i.e. walking area. Considering the angle of view, we put two Kinects at 1 m height on the tripod. The red dash lines are the maximum and minimum deep that Kinect can probe. The area enclosed by the black solid lines is the available walking area.
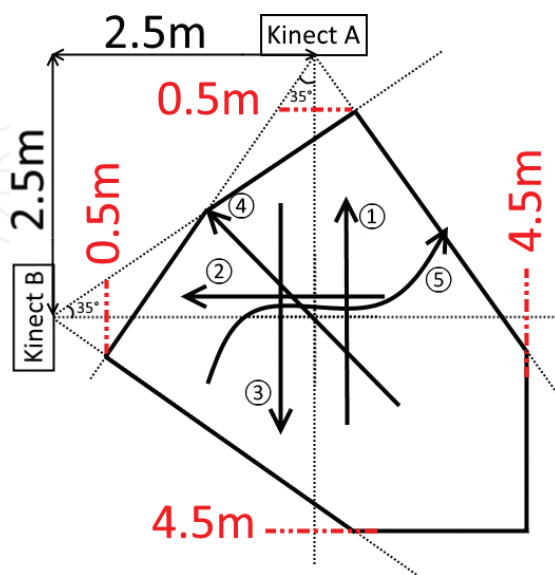
Figure 9. The top view of the experimental environment.

Before we record the data of each subject, we collect the basic information, such as name, sex, age, height, wearing (e.g. the high-heeled shoes, dress for female volunteers) and so on, for potential analysis and data mining. Each subject is asked to walk twice on the predefined directions shown as the arrows ①–⑤ in **Figure 9**, particularly ⑤ means the subjects walk in a straight line on an arbitrary direction. We can treat all the data as recorded by one Kinect since the two Kinects are the same, so that each subject has 20 walking sequences, and the walking duration on each predefined direction is shown in **Figure 10**. The dataset can be accessed at the website, https:/sites.google.com/site/sdugait/, and it can be downloaded with application.
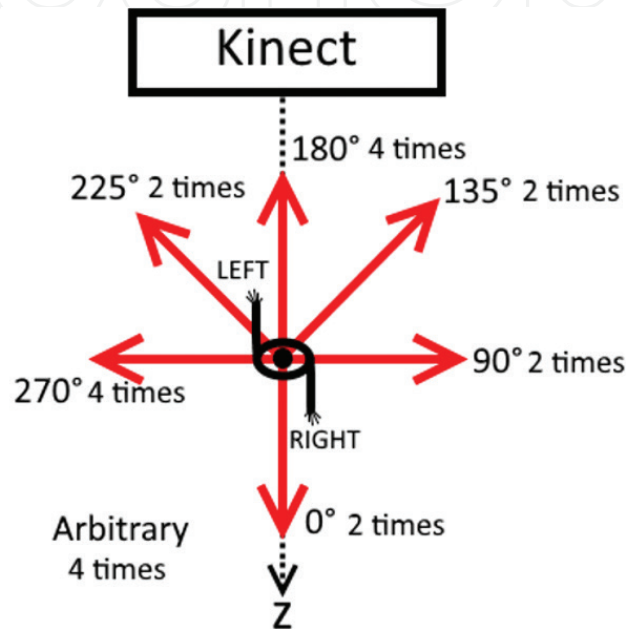
Figure 10. Walking directions and the corresponding walking duration.

3. Kinect-based gait recognition

3.1. The Kinect-based gait recognition

The gait features extracted from Kinect captured data contain the static and dynamic features. In this part, we will firstly introduce how to extract the static and dynamic features and demonstrate the properties of these two kinds of features. And then we will show how to extract a walking period from the sequence. Finally, we make a feature fusion of these two kinds of features for gait recognition.

A static feature is a kind of feature that can barely change during the whole walking process, such as height, the length of skeletons and so on. Given the knowledge of anthropometry, the person can be recognized based on static body parameters to some extent. Here, we choose the length of some skeletons as the static features, including the length of legs and arms. Considering the symmetry of human body, the length of limbs on both sides is usually treated to be equal. The static feature is defined as an eight-dimension vector, i.e. $(d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$,

where d_i is a space distance between Joint_1 and Joint_2 listed in **Table 2**. Here, the Euclidean distance is chosen to measure the space distance referring to the research experiences in Refs. [37, 38].

We can acquire the 3D coordinate of the joints listed in **Table 2** in each frame and calculate each component of the static feature vector.

$$d_i = \sqrt{\left(x_1 - x_2\right)^2 + \left(y_1 - y_2\right)^2 + \left(z_1 - z_2\right)^2} \tag{1}$$

where $\left(x_1, y_1, z_1\right)$ and $\left(x_2, y_2, z_2\right)$ represent 3D positions of the corresponding joints listed as Joint_1 and Joint_2, respectively.

When we evaluate the estimation on the position of joints obtained by Kinect, we find that the accuracy will change along with the depth range. Given the empirical results, we discover that more stable data can be acquired when the depth is between 1.8 and 3.0 m. Hence, we propose a strategy to automatically choose the frames in that range. We choose the depth information of HEAD joint to represent the depth of whole body, because it can be detected stably and keep monotonicity in depth direction during walking. Then we set two depth thresholds, i.e. the distance in the Z-direction, as the upper and lower boundaries, respectively. The frames between the two boundaries are regarded as the reliable frames.

$$\{f_a\} = \left\{ H_f \middle| H_{f, z > 1.8} \cap H_{f, z < 3.0} \right\} \tag{2}$$

where H_f denotes the frames of the HEAD, f_a denotes the reliable frames and $H_{f, z}$ represents the frame(s) that obtained when the coordinate of HEAD joint is z. We reserve the 3D coordinates of all the joints during the period when the reliable frames can be obtained. Finally, we calculate the length of the skeleton we need at each reliable frame, and take their average to calculate the components of the static feature vector.

The subjects are required to walk along the same path for seven times, which can make subjects walk more naturally later. For each subject, Kinect is turned for 5° started from −15° to

Component	Joint_1	Joint_2
d_1	HIP_RIGHT	KNEE_RIGHT
d_2	KNEE_RIGHT	ANKLE_RIGHT
d_3	SHOULDER_RIGHT	ELBOW_RIGHT
d_4	ELBOW_RIGHT	WRIST_RIGHT
d_5	SPINE_SHOULDER	SPINE_BASE
d_6	SHOULDER_RIGHT	SHOULDER_LEFT
d_7	SPINE_SHOULDER	NECK
d_8	NECK	HEAD

Table 2. Components of the static feature vector.

+15°, and the static feature vector on each direction is recorded. These directions are denoted by $n15$, $n10$, $n5$, 0, $p5$, $p10$ and $p15$, where '0' denotes the front direction, and 'n' and 'p' denote anticlockwise and clockwise, respectively. Totally 10 volunteers are randomly selected to repeat this experiment, and all the results prove that the static feature we choose is robust to the view variation. We show an example in **Figure 13a**, in which each component of these static vectors on the seven directions and the average values of these vectors are plotted.

The dynamic feature is a kind of feature that any change along with time during walking, such as speed, stride, variation of barycentre, etc. Given many researches [5, 39, 40], the angles of swing limbs during walking are remarkable dynamic gait features. For this reason, four groups of swing angles of upper limbs, i.e. arm and forearm, and lower limbs, i.e. thigh and crus, are defined as shown in **Figure 11**, and denoted as $a1, ..., a8$. Here, $a2$ is taken as the example for illustration. The coordinate at KNEE_RIGHT is denoted as (x, y, z), and coordinate at ANKLE_RIGHT is denoted as (x', y', z'), so $a2$ can be calculated as

$$\tan\angle a2 = \left(\frac{x - x'}{y - y'}\right) \quad a2 = \tan^{-1}\left(\frac{x - x'}{y - y'}\right) \tag{3}$$

Each dynamic angle can be regarded as an independent dynamic feature for recognition. Given the research results in Ref. [41] and our comparison experiments on these dynamic angles, angle $a2$ on the right side or $a4$ on the left side is selected as the dynamic angle, according to the side near to the Kinect.

The value of $a2$ and $a4$ at each frame can be calculated, and the whole walking process can be described, as shown in **Figure 12**. We carried out the verification experiments similar to what for the static feature to prove its robustness against the invariant of views, the result shown in **Figure 13b** indicates that the proposed dynamic feature is also robust to the view variation.

Gait period extraction is an important step in gait analysis, because gait is a periodical feature and majority features could be captured within one period. Silhouette-based methods usually analyse the variation of silhouette width alone with time to obtain the period information. Some methods apply the signal processing to analyse the dynamic feature for period extraction, such as peak detection and Fourier transform. Different from them, we propose to extract periodicity by combining the data of left limb and right limb together, which can be shown in **Figure 12**. $a2$ and $a4$ sequences represent the right and left signals, respectively.

It can be concluded that the crossing points between left and right signals can segment the gait period appropriately. We use the crossing point between the left and right signals to extract
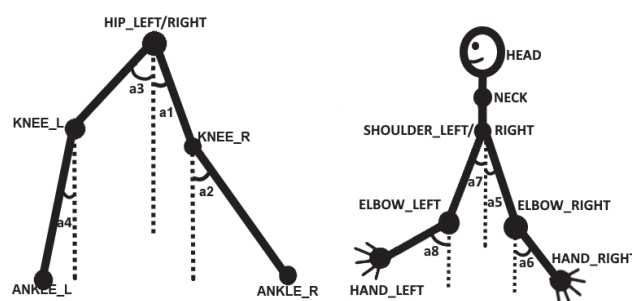
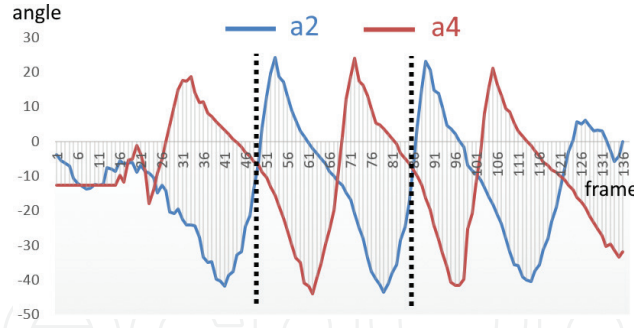**Figure 11.** Side view of the walking model.

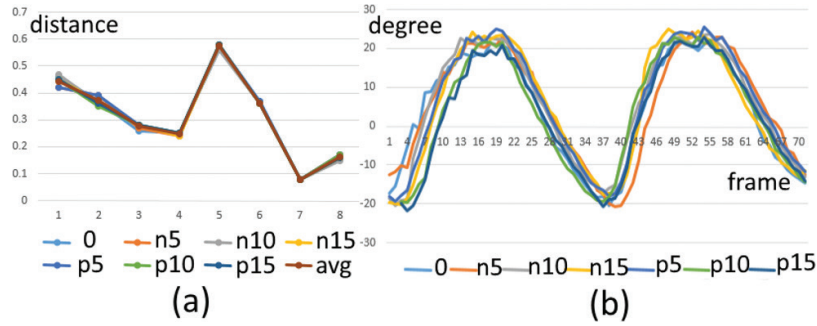**Figure 12.** Period extracted based on dynamic features.

Figure 13. (a) Static feature and (b) dynamic feature of one subject on seven directions.

the gait period. After cutting off the noisy part at the beginning of the signal, we make the subtraction on the left and right signals, obtain the crossing point as the zero point and extract the period between two interval zero points. The black dash lines show the detected period.

The static and dynamic features have their own advantages and disadvantages, respectively. These two kinds of features are fused in the score-level. Two different kinds of matching scores are normalized onto the closed interval [0,1] by the linear normalization.

$$\hat{s} = \frac{s - \min(\mathbf{S})}{\max(\mathbf{S}) - \min(\mathbf{S})} \tag{4}$$

Where $\mathbf{s}$ is the matrix before normalization, whose component is s, here represent the score, $\hat{\mathbf{s}}$ is the normalized matrix, whose component is $\hat{s}$. The two kinds of features are weighted fused as

$$F = \sum_{i=1}^{R} \omega_i \hat{s}_i, \quad \omega_i = \frac{C_i}{\sum_{j=1}^{R} C_i} \tag{5}$$

where F is the score after fusion, R is the number of features used for fusion, ω_i is the weight of ith classifier, $\hat{s}_i$ is the score of ith classifier, here which is our distance. C_i is the CCR (correct classification rate) of ith feature used to recognize separately, so the weight can be set according to the level of CCR.

3.2. Comparisons

The cross-view recognition abilities of the static feature, dynamic feature and their fusion are analysed. Four sequences on 180° are used as the training data since both body sides of the

subjects can be recorded. The sequences on the other directions are used as the testing data. Because the sequences acquired on the nearer side to Kinect have more accuracy, the data on the nearer body side is selected automatically for the calculation at each direction.

The static feature is extracted from the right body side on 0, 225 and 270°, and the left body side on 90 and 135°. Due to the symmetry, the skeleton lengths on the two sides of the body are regarded to be equal. The static feature is calculated as Eq. (1), and NN classifier is used for recognition. The results are shown in the first row of **Table 3**.

The dynamic feature, $a2$, is calculated from the right side of the limb on 0, 225 and 270°, and the dynamic feature, $a4$, is calculated on 90 and 135°, from the left side of the limb. As we can extract both $a2$ and $a4$ on the direction of 180°, either of them can be used as the dynamic feature on the training set. The results are shown in the second row of **Table 3**.

The static feature and dynamic feature are fused in the score-level as we discussed before, and the results are tested after feature fusion in situation under view variation. Given the CCR of dynamic feature and static feature obtained from different directions, we redistribute the weight, get the final score for different subjects and use the NN classifier to get the final recognition results as shown in the third row of **Table 3**. The comparison in **Table 3** shows that the feature fusion can improve the recognition rate on each direction.

Preis et al. [27] proposed a Kinect-based gait recognition method, in which 11 lengths of limbs are extracted as the static feature, and step length and speed are taken as the dynamic feature. Their method was tested on their own dataset including nine persons and the highest CCR can reach to 91%. The gait feature they proposed is also based on 3D position joint, so it is possible to rebuild their method on our database. In this chapter, we rebuilt their method and test on our database with 52 persons and make a comparison with our proposed method. As their dataset only include frontal walking sequences, we compare two methods in our database only on 180° (frontal) directions. We randomly choose three sequences on 180° directions as training data and the rest are treated as testing data. The CCR results of both methods are shown in **Table 4**. Our proposed method has about 10% accuracy improvement.

The proposed method is evaluated another Kinect-based gait dataset, i.e. KinectREID dataset in Ref. [36]. Four recognition rate curves are shown in **Figure 14**, which are front_VS_front, rear_VS_rear, front_VS_rear and front_VS_lateral, because there are only three directions in

| Static Feature	0°	90°	135°	225°	270°
	88.46	84.61	82.69	84.61	88.46
Dynamic Feature	0°	90°	135°	225°	270°
	88.46	86.5	84.61	84.61	90.38
Feature Fusion	0°	90°	135°	225°	270°
	94.23	90.38	90.38	88.46	92.31

Table 3. CCR (%) results of the static feature, dynamic feature, and feature fusion on each walking direction.

	CCR
Method in [27]	82.7%
Our method	92.3%

Table 4. Comparison on CCR between the proposed methodand the method in [27].

KinectREID dataset, i.e. front, rear and lateral. It can be seen from **Figure 14** that the cross-view recognition rate of the propose method is slightly worse than that on the same directions, which demonstrates that the robustness of the proposed method against view variation, though the recognition rate, decreases with the increasing of the amount of test subjects.

Given the experimental results we have discussed above, we can say that the static relation and dynamic moving relation among joints are very important features that can represent the characteristic of gait. In many 2D-based methods, many researchers also tried to get the relation among joints, but the positions of joints have to be calculated from the 2D video with all kinds of strategies in advantage. Goffredo et al. proposed a view-invariant gait recognition method in Ref. [17]. They only make use of 2D videos obtained by one single camera. After extracting the walking silhouette from the background, they estimate the position of joints according to the geometrical characteristics of the silhouette and calculate the angle between the shins and the vertical axis and the angle between thigh and the vertical axis as the dynamic feature, and finally make a projection transformation to project these features into the sagittal plane using their viewpoint rectification algorithm. Actually, Goffredo's
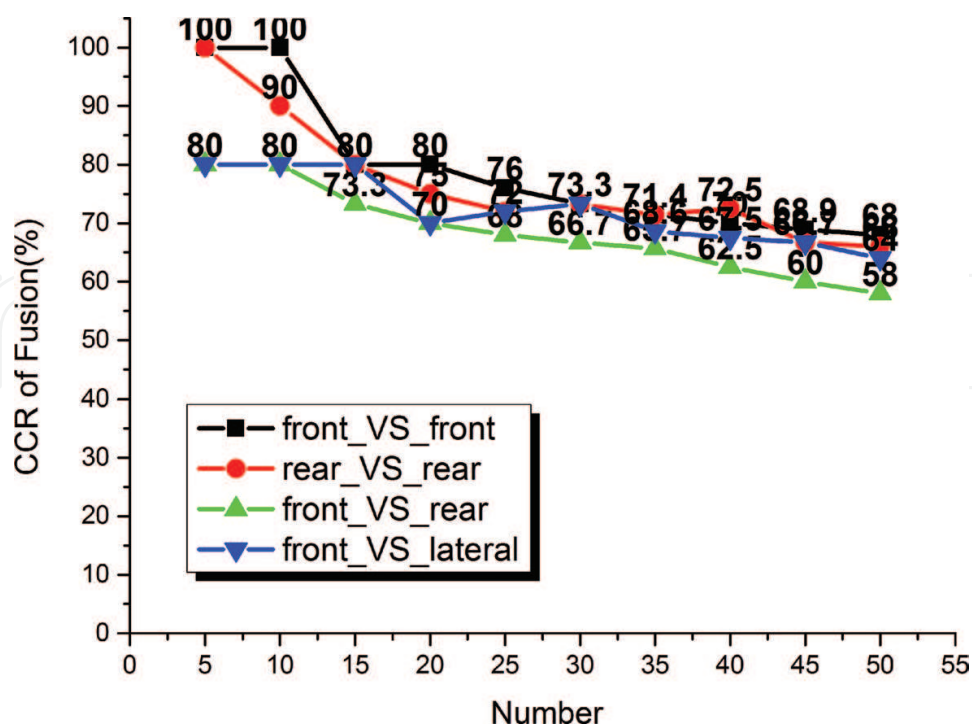
Figure 14. Gait recognition performance on KinectREID dataset.

method has a lot of similar gait features comparing with our method in logically. As we mentioned before, our database not only have the 3D position data but also the 2D silhouette images at each frame. Take advantage of our database, we can rebuild their method using the 2D silhouette image sequences; meanwhile, we use the 3D joint position data of the same person. We compare this method with our method with the varying views on three directions. The comparison results in **Table 5** show that our proposed method has 14–19% accuracy improvement.

3.3. Applications

Gait research is still at an exploring stage rather than a commercial application stage. However, we have confidence to say that the gait analysis is promising given its recent development. The unique characteristics of gait, such as unobtrusive, non-contactable and non-invasive, have a powerful potential to apply in the scenarios including criminal investigation, access security and surveillance. For example, face recognition will become unreliable if there is a larger distance between the subject and camera. Fingerprint and iris recognition have proved to be more robust, but they can only be captured by some contact or nearly contact equipment.

For instance, gait biometrics has already been used as the evidence for forensics [42]. In 2004, a perpetrator robbed a bank in Denmark. The Institute of Forensic Medicine in Copenhagen (IFMC) was asked to confirm the perpetrator via gait analysis, as they thought the perpetrator had a unique gait. The IFMC instructed the police to establish a covert recording of the suspect from the same angles as the surveillance recordings for comparison. The gait analysis revealed several characteristic matches between the perpetrator and the suspect, as shown in **Figure 15**. In **Figure 15**, both the perpetrator on the left and the suspect on the right showed inverted left ankle, i.e. angle b, during left leg's stance phase and markedly outward rotated feet. The suspect was convicted of robbery and the court found that gait analysis is a very valuable tool.

Another similar example is in the intelligent airport, where the Kinect-based gait recognition is used during the security check. Pratik et al. [43] established a frontal gait recognition system using RGB-D camera (Kinect) considering a typical application scenario of airport security check point, as shown in **Figure 16a**. In their further work [44], they addressed the occlusion problem in frontal gait recognition via the combination of two Kinects, which is demonstrated in **Figure 16b**.
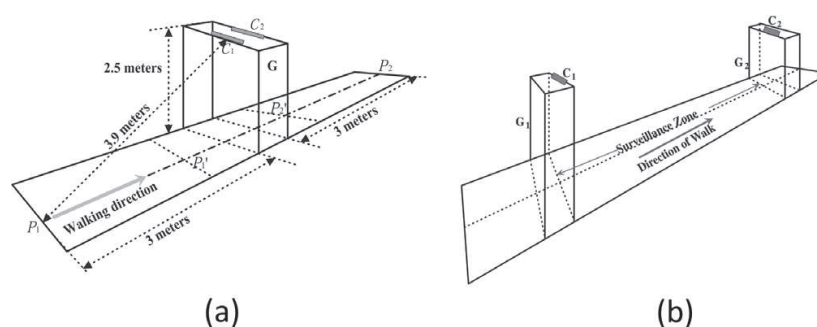
In addition, gait analysis plays an important role in medical diagnosis and rehabilitation. For example, assessment of gait abnormalities in individuals affected by Parkinson's disease

	0°	90°	135°
Method in [17]	80.8	71.15	73.08
Our method	94.23	90.38	90.38

Table 5. CCR (%) result comparing on three directions.

Figure 15. Bank robbery identification.

Figure 16. Gait-based airport security check system with (a) single and (b) double Kinects.

(PD) is essential to determine the disease progression, the effectiveness of pharmacologic and rehabilitative treatments. Corona et al. [45] investigate the spatio-temporal and kinematics parameters of gait between lots of elderly individuals affected by PD and normal people, which can help clinicians to detect and diagnose the Parkinson's disease.

Author details

Jiande Sun[1,2]*, Yufei Wang[1] and Jing Li[3]

*Address all correspondence to: jiandesun@hotmail.com

1 School of Information Science and Engineering, Shaodong Normal University, Jinan, Shandong Province, China

2 Institute of Data Science and Technology, Shandong Normal University, Jinan, Shandong Province, China

3 School of Mechanical and Electrical Engineering, Shandong Management University, Jinan, Shandong Province, China

References

[1] Cunado D, Nixon MS, Carter JN. Using gait as a biometric, via phase-weighted magnitude spectra. In: International Conference on Audio- & Video-Based Biometric Person Authentication (AVBPA 1997); 12-14 March 1997; Crans-Montana, Switzerland. Berlin: Springer; 1997. pp. 93-102

[2] Johnson AY, Bobick AF. A multi-view method for gait recognition using static body parameters. Audio- & Video-Based Biometric Person Authentication (AVBPA 2001); 6-8 June 2001; Halmstad, Sweden. Berlin: Springer; 2001. pp. 301-311

[3] Guo Y, Xu G, Tsuji S. Understanding human motion patterns. In: International Conference on Pattern Recognition (ICPR 1994); 9-13 October 1994; Jerusalem, Israel. New York: IEEE. 1994;**2**:325-329

[4] Rohr K. Towards models-based recognition of human movements in image sequences. CVGIP. 1994;**59**(1):94-115

[5] Tanawongsuwan R, Bobick A. Gait recognition from time-normalized joint-angle trajectories in the walking plane. In: IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR 2001); 8-14 December 2001; Kauai, HI, USA. New York: IEEE. 2001;**2**:726

[6] Wang L, Ning H, Tan T, Hu W. Fusion of static and dynamic body biometrics for gait recognition. IEEE Transactions on Circuits & Systems for Video Technology. 2003;**14**(2):149-158

[7] Han J, Bhanu B. Individual recognition using gait energy image. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2006;**28**(2):316-322

[8] Lam THW, Lee RST. A new representation for human gait recognition: Motion silhouettes image (MSI). In: International Conference on Advances in Biometrics (ICB 2006); 5-7 January 2006; Hong Kong, China. Berlin: Springer. 2006;**3832**:612-618

[9] Lee S, Liu Y, Collins R. Shape variation-based frieze pattern for robust gait recognition. In: IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR 2007); 18-23 June 2007; Minneapolis, Minnesota, USA. New York: IEEE; 2007. pp. 1-8

[10] Bashir K, Xiang T, Gong S. Gait recognition using gait entropy image. In: International Conference on Imaging for Crime Detection & Prevention (ICDP 2009); 3 December 2009; London, United Kingdom. IET; 2010. pp. 1-6

[11] Wang C, Zhang J, Pu J, Yuan X, Wang L. Chrono-gait image: A novel temporal template for gait recognition. In: European Conference on Computer Vision (ECCV 2010); 5-11 September 2010; Crete, Greece. Berlin: Springer. 2010;**6311**:257-270

[12] Sundaresan A, Roychowdhury R, Chellappa R. A hidden Markov model based framework for recognition of humans from gait sequences. In: International Conference on Image Processing (ICIP 2003); 14-18 September 2003; Barcelona, Catalonia, Spain. New York: IEEE. 2003;**2**:93-96

[13] Wang L, Tan T, Ning H, Hu W. Silhouette analysis-based gait recognition for human identification. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2004;**25**(12):1505-1518

[14] Sudeep S, Phillips PJ, Liu Z, Isidro RV, Patrick G, Bowyer KW. The human ID gait challenge problem: Data sets, performance, and analysis. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2005;**27**(2):162-177

[15] Abdelkader CB, Davis L, Cutler R. Motion-based recognition of people in Eigen gait space. In: IEEE International Conference on Automatic Face and Gesture Recognition (FG 2002); 20-21 May 2002; Washington, DC, USA. New York: IEEE; 2002. pp. 267-272

[16] Kale A, Chowdhury AKR, Chellappa R. Towards a view invariant gait recognition algorithm. In: IEEE Conference on Advanced Video & Signal Based Surveillance (AVSS 2003); 21-22 July 2003; Miami, FL, USA. New York: IEEE; 2003. pp. 143-150

[17] Goffredo M, Bouchrika I, Carter JN, Nixon MS. Self-calibrating view-invariant gait biometrics. IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man & Cybernetics Society. 2010;**40**(4):997-1008

[18] Muramatsu D, Shiraishi A, Makihara Y, Uddin MZ, Yagi Y. Gait-based person recognition using arbitrary view transformation model. IEEE Transactions on Image Processing. 2014;**24**(1):140-154

[19] Wu Z, Huang Y, Wang L, Wang X, Tan T. A comprehensive study on cross-view gait based human identification with deep CNNs. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2017;**39**(2):209-226

[20] Zhao G, Liu G, Li H, Pietikainen M. 3D gait recognition using multiple cameras. In: International Conference on Automatic Face and Gesture Recognition (FG 2006); 10-12 April 2006; Southampton, UK. New York: IEEE; 2006. pp. 529-534

[21] Yamauchi K, Bhanu B, Saito H. Recognition of walking humans in 3D: Initial results. In: IEEE Computer Society Conference on Computer Vision & Pattern Recognition Workshops (CVPR 2009); 20-25 June 2009; Miami, Florida, USA. New York: IEEE; 2009. pp. 45-52

[22] Krzeszowski T, Michalczuk A, Kwolek B, Switonski A, Josinski H. Gait recognition based on marker-less 3D motion capture. In: IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2013); 27-30 August 2013; Krakow, Poland. New York: IEEE; 2013. pp. 232-237

[23] Sivapalan S, Chen D, Denman S, Sridharan S, Fookes C. Gait energy volumes and frontal gait recognition using depth images. In: International Joint Conference on Biometrics (IJCB 2011); 11-13 October 2011; Washington, DC, USA. New York: IEEE; 2011. pp. 1-6

[24] Nambiar AM, Correia PL, Soares LD. Frontal gait recognition combining 2D and 3D data. In: Proceedings of the on Multimedia and Security; 6-7 September 2012; Coventry, United Kingdom. New York: ACM; 2012. pp. 145-150

[25] Araujo RM, Graña G, Andersson V. Towards skeleton biometric identification using the Microsoft Kinect sensor. In: ACM Symposium on Applied Computing; 18-22 March 2013; Coimbra, Portugal. New York: ACM; 2013. pp. 21-26

[26] Milovanovic M, Minovic M, Starcevic D. Walking in colors: Human gait recognition using Kinect and CBIR. Multimedia IEEE. 2013;**20**(4):28-36

[27] Preis J, Kessel M, Linnhoff-Popien C, Werner M. Gait recognition with Kinect. In: Workshop on Kinect in Pervasive Computing; 18-22 June 2012; Newcastle, UK

[28] Yang K, Dou Y, Lv S, Zhang F, Lv Q. Relative distance features for gait recognition with Kinect. Journal of Visual Communication & Image Representation. 2016;**39**(C):209-217

[29] Ahmed F, Paul PP, Gavrilova M. Kinect-based gait recognition using sequence of the most relevant joint relative angles. Journal of WSCG. 2015;**23**(2):147-156

[30] Kastaniotis D, Theodorakopoulos I, Theoharatos C, Economou G, Fotopoulos S. A framework for gait-based recognition using Kinect. Pattern Recognition Letters. 2015;**68**(P2):327-335

[31] Shutler JD, Grant MG, Nixon MS, Carter JN. On a large sequence-based human gait database. Applications and Science in Soft Computing. 2004;**24**:66-71

[32] Hadid A, Ghahramani M, Bustard J, Nixon MS. Improving gait biometrics under spoofing attacks. In: Image Analysis and Processing (ICIAP 2013); 9-13 September 2013; Naples, Italy. Berlin: Springer; 2013. pp. 1-10

[33] Yu S, Tan D, Tan T. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: International Conference on Pattern Recognition (ICPR 2006); 20-24 August 2006; Hong Kong, China. New York: IEEE. 2006;**4**:441-444

[34] Iwama H, Okumura M, Makihara Y, Yagi Y. The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Transactions on Information Forensics & Security. 2012;**7**(5):1511-1521

[35] Hofmann M, Geiger J, Bachmann S, Schuller B, Rigoll G. The TUM Gait from Audio, Image and Depth (GAID) Database: Multimodal recognition of subjects and traits. Journal of Visual Communication & Image Representation. 2014;**25**(1):195-206

[36] Pala F, Sata R, Fumera G, Roli F. Multimodal person reidentification using RGB-D cameras. IEEE Transactions on Circuits & Systems for Video Technology. 2016;**26**(4):788-799

[37] Makihara Y, Matovski DS, Nixon MS, Carter JN, Yagi Y. Gait recognition: Databases, representations, and applications. Wiley Encyclopedia of Electrical and Electronics Engineering. Wiley; 2015. DOI: 10.1002/047134608X.W8261

[38] Andersson VO, Araujo RM. Person identification using anthropometric and gait data from Kinect sensor. In: AAAI Conference on Artificial Intelligence; 25-30 January 2015; Austin, Texas, USA. Menlo Park, CA: AAAI Press; 2015. pp. 425-431

[39] Dan IS, Toth-Tascau M. Influence of treadmill velocity on joint angles of lower limbs during human gait. In: E-Health and Bioengineering Conference (EHB 2011); 24-26 November 2011; Iasi, Romania. IEEE; 2011. pp. 1-4

[40] Tech SXMYM, Larsen PK, Alkjær T, Simonsen EB, Lynnerup N. Variability and similarity of gait as evaluated by joint angles: Implications for forensic gait analysis. Journal of Forensic Sciences. 2014;**59**(2):494-504

[41] Pfster A, West AM, Bronner S, Noah JA. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. Journal of Medical Engineering & Technology. 2014;**38**(5):274-280

[42] Bouchrika I, Goffredo M, Carter J, Nixon MS. On using gait in forensic biometrics. Journal of Forensic Sciences. 2011;**56**(4):882-889

[43] Chattopadhyay P, Sural S, Mukherjee J. Frontal gait recognition from incomplete sequences using RGB-D camera. IEEE Transactions on Information Forensics & Security. 2014;**9**(11):1843-1856

[44] Chattopadhyay P, Sural S, Mukherjee J. Frontal gait recognition from occluded scenes. Pattern Recognition Letters. 2015;**63**:9-15

[45] Corona F, Pau M, Guicciardi M, Murgia M, Pili R, Casula C. Quantitative assessment of gait in elderly people affected by Parkinson's disease. In: IEEE International Symposium on Medical Measurements and Applications (MeMeA 2016); 15-18 May 2016; Benevento, Italy. IEEE; 2016. pp. 1-6