



Gait-Based Person Identification, Gender Classification, and Age Estimation: A Review

Rupali Patua^(✉), Tripti Muchhal, and Saikat Basu

Maulana Abul Kalam Azad University of Technology West Bengal, Kolkata, India
{rupali.patua,21tripti,saikatbasu}@gmail.com

Abstract. In this era where both techniques and technology are going digital, there is a need of designing and developing such a system that automatically identifies the person and also the person's attributes with the help of the biometric system. The biometric of every person is unique. Therefore, in this context, the use of the biometric system to identify the identity of an individual is a popular approach. Gait is a biometric approach that helps to verify the identity of a person by their walking patterns. Owing to the advantages of gait as a biometric, its popularity among the researchers is amplified in recent years. In this paper, a review is represented, where several papers with different methods, have been mentioned which recognize the person and its attributes such as gender and age based on the gait of the person.

Keywords: Gait · Kinect · SVM · KNN · PCA

1 Introduction

In recent years, the researchers' major area of inclination is towards biometric identification owing to its potential application in surveillance systems, social security, etc. [1]. Biometric helps in identifying an individual using their body parts. There are some characteristics in a human being, those are unique and can be used as biometric like the fingerprint, the retina of the eyes and many more [2]. Biometric can be classified as physiological and behavioral biometrics. Under the physiological biometrics the face, palm, ears, and eyes are considered, whereas the behavioral biometric includes voice, gait, speech, handwriting, and signature. Gait is a biometric approach that is becoming popular in recent years. The popularity of gait is due to its advantages, such as it can work in low resolution as well as does not need any person's cooperation as compared to other biometrics, such as face where high resolution is required to collect the person's face image as well as it requires the person's cooperation [3]. Gait is defined by the walking pattern of an individual. It works as a biometric because

every person has a unique walking pattern that cannot be imitated. The gait as a biometric differs from the others as it can be captured from distance as well as it does not need any cooperation from the individual. The gait features depend on various attributes such as the anatomy, gender, age, muscle mass, etc. [4]. One of the main advantages of using gait is that no one can imitate the walk of an individual [5]. Also in recent years, gait information has been used to help in various diseases like Parkinson's disease, Neurodegenerative Diseases and many more [6]. The approaches for gait recognition are: one is the model-free approach, and the other is the model-based approach. In the former, the silhouette of a person is extracted to find the attributes of that person while walking. The model-free approach is computationally efficient, but at the same time, it has few drawbacks due to the following reasons, such as changes in viewpoint, lighting, and clothing, whereas in the latter, there is a pre-defined model from which the dynamic characteristics of an individual are extracted. It is computationally expensive but at the same time, gains more information and also forceful to changes in lighting, clothing, and viewpoint in contrast to the model-free approach [7–9].

The inclination towards gait as a biometric and its huge implementation to identify the individual, the gender, and to estimate the age has been increased widely in recent years, and this has been the motivation to make this paper.

In this paper, different methods of gait recognition from various recent research works have been studied and compared. The remainder of the paper is organized as follows: Sect. 2 briefly elaborates on sensors that have been used to recognize the gait of an individual. Sections 3, 4, 5, 6, and 7 represent the comparison of the methods from different papers. Apart from that, in each subsection, the feature extraction, classifiers, and the pros and cons have been discussed. There is a table which includes all the papers that have been compared. Section 8 represents the conclusion.

2 Literature Review

Gait as a biometric has an important role in many application areas such as surveillance systems, social security, senior monitoring, disease detection, identification and verification of an individual etc. This survey is organized by taking only a part of the gait recognition that includes person identification, gender classification and age estimation. In this regard, a classification tree is constructed in the Fig. 1, and it also includes whether the Kinect sensor is used or not.

Kinect sensor is one among the other sensors, which is very effective to extract the skeleton points of an individual as well as to extract features. There is hardware, embedded in the Kinect sensor [4]. They are: RGB camera, Depth sensor, and Multiarray microphone. One of the major application for which the Kinect sensor is developed by Microsoft is the gaming application for the Xbox 360. Kinect has been used in biometrics, such as facial recognition, gait analysis, etc. [10]. There are two versions of Kinect sensor, one is Kinect v1 sensor, and the other is Kinect v2 sensor. In comparison to the Kinect v1 sensor, the skeleton

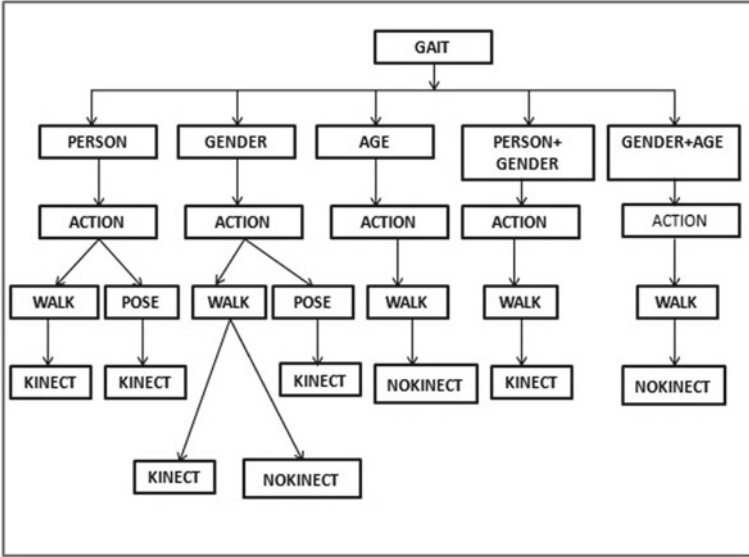


Fig. 1. Classification tree

tracking from Kinect v2 sensor is more precise and also at a faster rate. The resolution of the Kinect v2 sensor is huge enough but there is a disadvantage that the Kinect v2 sensor is not lightweight compare to the Kinect v1 sensor. Both the sensors have different methods to calculate the depth.

3 Person Identification

This section elaborates on how the different types of features have been extracted through different Kinect sensors and the different classifiers have been used to identify the person. Finally, the pros and cons of different papers are discussed and compared.

3.1 Feature Extraction

All the papers in the table where a person is identified have used different methods and different types of Kinect sensors. In [11], the Kinect v1 sensor has been used to extract the 20 skeleton joint points, and from each point the distance feature vector is calculated with respect to the other points, and then for each joint the mean and the variance are calculated as features whereas in [1], the Kinect sensor has been used for the extraction of 20 skeleton joint points and then both the static and dynamic features are extracted. However, in [1] a single value of the feature is taken from a sequence of frames. In [12], also Kinect sensor is used for the extraction of 20 different skeletal points and here 11 static features and two dynamic features are extracted. In [13], the 3d skeletal data is

extracted using the Kinect v2 sensor. Here two new features have been introduced, one is the JRD (Joint Relative Distance) and the other is the JRA (Joint Relative Angle). From these two features, the relevant features are selected by the genetic algorithm. In [14], the poses are extracted from the gait sequences with the help of the Kinect sensor. After these poses are acquired, they are used for the feature extraction and it is done by smoothing the voxel volume. In [15], the Kinect RGBD sensor is used to extract the skeleton data of a person. In this method, all the joints are taken in order to examine the entire video gait sequence and to extract the motion and anthropometric features.

3.2 Classifier

In [11], the KNN (K-nearest neighbor) is applied for the classification. It is a simple classifier and it classifies according to the similarity measured by the distance function. In [1], there are two classification methods used, one is the Levenberg-Marquardt back propagation and the other is the correlation algorithm. From the experiment, it has been seen that the recognition rate in the Levenberg-Marquardt back propagation is better than the correlation algorithm. In [12], three classifiers are used, one is the C 4.5, second is the Decision Tree and the third is the Naive Bayes classifier. The Naive Bayes classifier is a probabilistic classifier, performs better than the other two classifiers. In [13], a new classification method has been introduced, which is known as DTW (Dynamic Time Warping) based kernel. It is a non-linear time alignment technique. In [14], K-means clustering is used for the recognition method of an individual. In [15], two classification methods have been used. Firstly, SVM is used to identify the actions walk and run. After the respective action is identified, the person is recognized by calculating the identity cost of that person.

3.3 Discussion

In [11], the distance feature vector is calculated, which helps in the significant enhancement of the recognition rate as well as the reduction of computational time. In [1], only the single value of the feature is taken from the series of frames of a video sequence and it, in turn, becomes advantageous, since it reduces the training time as well as the classification time. Also, the compulsion of the gait cycle detection is eliminated. In [13], JRD and JRA methods have been used, which can handle speed variation and also it is robust to view and pose variation. In [14], the silhouette recorded by the Kinect gathers the depth information which in turn, preserves the vital gait information. In [15], the method identifies both the walk and run activities and the person. When the motion pattern and anthropometric features are used, the recognition rate is more accurate. In [12], the feature can be easily extracted by the use of the Kinect sensor.

In [11], it has some limitations in terms of the accuracy rate. The method that uses mean or variance has a lower accuracy rate compared to the method that uses both mean and variance. In [1], some of the features extracted like the length of both the hands, the length of both the legs do not offer distinguishing

features amid different individuals, which in turn becomes a disadvantage. In [13], it is susceptible to the changes that happen in the clothing and carrying object of an individual such as a handbag, etc. In [14], the method has higher response time. In [15], when biometric is used individually, it can be sensitive in some respect. In [12], the disadvantage is that the approach cannot fit for the identity of an individual among the crowd.

Comparing [1] and [12] with [11], it has high precision and recognition rate. The approach used in [13] is more accurate than the approach used in [15], as only those joint pair is taken which are relevant.

4 Gender Classification

In this section, the methods used to extract the feature as well as the methods used to classify the gender have been elucidated and at the last, the pros and cons of different papers are discussed and compared.

4.1 Feature Extraction

The gait can be used for gender classification. Various papers for gender classification have been detailed in the table. In [16], the data of the human skeleton is extracted with the help of the Kinect sensor and along with that 20 joint points are captured. Here, the dynamic features are extracted differently by calculating the dynamic distance feature. In [17], the silhouette is segmented from the frames of a video, and then GEI and AEI are calculated. The feature space becomes better by applying the feature selection and resampling method. In [18], poses are taken from the video sequence, and then the feature extraction is done with the help of histogram. In [14], the new database is generated known as d gait database that contains the depth information of the subject. Here both the 2D and 3D gait features are extracted.

4.2 Classifier

In [16], three different types of classification methods have been used separately; those are the NN (Nearest Neighbor) classifier, LDC (Linear Discriminant Classifier), and SVM (Support Vector Machine). Out of these three, the accuracy rate of NN is better than the other two. In [17], the KNN is used for classification purpose. As the value of $K = 1$ in KNN, the performance degrades because KNN does not work well in the intrinsic data characteristics. In [18], for the classification purpose, the RBF kernel SVM is applied. In [19], a kernel SVM is used for the classification.

4.3 Discussion

In [16], the distance between the ankles is measured by which the accuracy to classify the gender becomes better. In [17], the feature selection reduces the

dimensionality so that the classifier can work more rapidly and also more efficiently. In [18], the method is neither dependent on viewing angle, nor on clothing or carrying objects nor on the number of frames, which is an advantage of this method. In [19], the depth information gained by the Kinect is useful and also robust when the view of a person changes with respect to the camera.

In [16], it has a disadvantage to the database as it doesn't have a variant in clothing and bags. In [17], the gait considers not only the movement but also the appearance. Due to this fact, the approach is comprehensively affected and the dataset SOTON attains a poor result. In [18], in real world the challenge of tracking people in the mass gathering area still remains. The variance calculated in the paper [19], shows that the 3D GF has a lower variance than 2D GF. Due to this reason, the 2D GF is less robust compared to the 3D GF in case of view variation.

The method in [16] gets the accuracy above 90% whereas the method in [17] has an accuracy of 87% which shows the result of [16] is better than [17]. Since there is no detection of the gait cycle in [18], the method is better than [19], where gait cycle is needed.

5 Age Estimation

In this section, different method names that include a calculation to extract features and also different classifiers to estimate the age of a person are discussed. In the end, the pros and cons of various papers are illustrated.

5.1 Feature Extraction

In [20], the feature of an individual gait is captured by calculating the gait energy image (GEI). It represents both the static and dynamic features. In [21], three types of gait features are taken, which are based on silhouette. These gait features are the gait energy image (GEI), the frequency domain features and the gait periods. In [22], the feature extraction is done using gait energy image (GEI). The extracted feature dimension is reduced by applying a multi-label guided subspace.

5.2 Classifier

In [20], the problem of multiple age group classification is solved by the DAGSVM, which is a directed acyclic graph SVM. The DAGSVM is the solution to the multi-class classification problem by incorporating multiple binary SVM classifiers. In [21], the mean absolute error (MAE) and the cumulative score are taken for age estimation. In [22], the ML-KNN classifier is used for recognition and age decoding.

5.3 Discussion

In [20], a new method that is the age group dependent is used. The computational time is low in this approach. In [21], the database contains approximately all the ages in the range from 2 to 94 years. This helps in performance evaluation and also it is reliable for the training model. In [22], the approach of multiple feature fusion is used, which constantly beats the single feature approach for age estimation this implies that the multiple feature fusion can give harmonizing information to represent features, which results in better estimation..

In [20], it refers that, when one of the hyperparameter contribution rates is lower, MAE (Mean Absolute Error) increases. In [21], there are two failures due to two reasons: (1) the testing of some subjects fails because the gait features of those subjects deviate from the gait of their actual age, (2) the test gait features and the training gait features are different from each other. In [22], the only single feature that is the GEI feature is extracted for age estimation. Though the single feature is less robust, the performance is reduced as compared to the multiple feature fusion.

The approach in [20] is better compare to [21,22] as it uses the age group dependent approach, whereas others use the single age independent approach. In the approach of [20], the estimation error is increased as there is an increase of variation in age. The method in [4] is better for estimation of age than in [21,22] as it uses spatial proximity and multi-task learning on CNN.

6 Age Estimation and Gender Classification

Here, the methods of feature extraction and the classifier that helped to estimate age and classify gender are discussed. Also, the pros and cons of various papers have been illustrated.

6.1 Feature Extraction

In [23], the frequency domain feature is selected for feature extraction as it contains both 2D spatial information and dynamic information.

6.2 Classifier

In [23], the KNN classifier is applied for the classification approach.

6.3 Discussion

Here, some papers are referred, that worked on both gender and age. In [23], the database contains multiple views therefore; the multi-view application can be supported.

In [18], when the diversification in gender and age increases, the classification becomes tough.

7 Person Identification and Gender Classification

This section refers the feature extraction method and the classifiers to identify both the person and its gender. Apart from that, the pros and cons of several papers are discussed and compared.

7.1 Feature Extraction

In [24], the poses are taken from the video. The poses which are captured are represented by the Eulers angle which is the vectors of eight selected limbs, and then the dynamic features are extracted. The dissimilarity space is measured between the walking and the training sequence and it is represented as dissimilarity vector. In [25], the silhouette is obtained by the background subtraction, and then the cluster is formed. For every cluster, the feature is represented by calculating the average gait image of that cluster.

7.2 Classifier

In [24], the sparse representation is measured, and with the help of it the recognition and classification task are performed.

In [25], the SRML is used as a discriminative distance metric in order to perform the recognition method.

7.3 Discussion

In [24], every gait can be represented by the vector of dissimilarity space, and this approach is efficient. In [25], the method recognizes both the person and gender in an arbitrary walking direction. This method can be applied to the outdoor environment.

In [24], when partial body parts such as arms and legs are incorporated, there exists unpredictability due to the narrow training sample in the training set. Due to this reason, a faster deterioration in the performance happens when all the features are incorporated. In [25], the disadvantage is that the method will lead to failure when the training and testing sequences are different from each other. Comparing with [17, 23, 25] executes better recognition, classification and verification.

Now in the below table, various papers are incorporated that have used a person's gait and several methods to identify the person, classify the gender, and to estimate the person's age. Table 1 contains the papers in the year range between 2010 and 2018.

Table 1. Comparison table

Ref	Identification	Sensor	Method	Feature	Activity	Year	Database
[1]	Person	Kinect	Levenberg-Marquardt back propagation and correlation algorithm	20 joint points detected using skeleton data in 3D coordinates. From these 20 joints points, the static and dynamic features are extracted	Action [walk]	2015	N.A.
[11]	Person	Kinect V1	Euclidean distance, meanV, Variance, KNN	20 joint points is detected . The distance feature vector is calculated for each one of the joint with respect to other 20 joints	Action [walk]	2017	N.A.
[12]	Person	Kinect	1R, C4.5 decision tree, Naive Bayes	20 different points are detected from the skeleton and from there 13 biometric features are extracted out of which 11 are static and 2 are dynamic feature	Action [walk]	2012	N.A.
[13]	Person	Kinect V2	DTW, rank level fusion, genetic algorithm	Detection of the gait cycle is done using Kinect v2 sensor and then JRD (Joint Relative Distance), JRA (Joint Relative Angle) features are calculated	Action [walk]	2015	N.A.
[14]	Person	Kinect	Coordinate system transformation, K-Means clustering	Fronto-Parallel silhouette is generated and from these set the key poses is derived	Action [pose]	2014	N.A.

Table 1. (*continued*)

Ref	Identification	Sensor	Method	Feature	Activity	Year	Database
[15]	Person	Kinect RGBD	SVM,PCA, MDA	Motion pattern and anthropo- metric features	Action [walk, run]	2012	N.A.
[16]	Gender	Kinect	DDF, LDC, SVM, mean, standard deviation and Skew	From the 3D skeleton data, the 20 joint point is extracted and the gait signature is extracted using dynamic distance feature	Action [walk]	2017	N.A.
[17]	Gender	N.A.	GEI, AEI, KNN	The silhouette is being segmented from the frames of a gait video sequence and then the GEI and AEI are calculated	Action [walk]	2013	CASIA dataset A and SOTON dataset B
[18]	Gender	Kinect RGD Sensor	EulerAngle, PCA, SVM	The feature extraction is done with the help of histogram	Action [Pose]	2013	UPCV Gait
[19]	Gender	Kinect	PCA, LDA, SVM	2D and 3D gait features	Action [walk]	2012	D-Gait
[20]	Age	N.A.	DAGSVM, age-group dependent manifold learning, nonlinear SVR, L2 distance	GEI feature	Action [walk]	2018	OULP -Agr Dataset
[21]	Age	N.A.	Gaussian process regression, baseline algorithm	Three silhouette based features are taken such as averaged silhouette, frequency domain feature and gait period	Action [walk]	2011	Whole Generation gait

Table 1. (*continued*)

Ref	Identification	Sensor	Method	Feature	Activity	Year	Database
[22]	Age	N.A.	Label encoding algorithm, MLG (Multi-Label Guided), MLKNN, Gabor Magnitude, Gabor phase	Gabor feature	Action [walk]	2010	USF Database
[23]	Age+ Gender	N.A.	KNN	Frequency domain features	Action [walk]	2010	Multi View Gait Database
[24]	Person + Gender	Kinect	Euler's angle, dissimilarity measure, Sparse representation	Dynamic features are extracted	Action [pose]	2015	N.A.
[25]	Person + Gender	Kinect	Background subtraction, SRML	The feature taken is the cluster-based averaged gait image	Action [walk]	2014	USF and CASIA-B Database

8 Conclusion

Gait has many advantages that make it more robust compared to the other biometric. But in some perspective, it is not robust when the changes in clothing and shoe types are considered. This is the limitation of gait that still needed to be approached. In this context, a database is needed, which should include all the perspectives like different view angles, changes in clothing etc., which is required for the more accurate recognition method. The gait as a biometric can be used in various places like in airports, shopping malls, etc. Gait recognition can be combined with many other research areas such as face recognition, security, and identification of an individual in crowd, which can be more tighten. In the recent years, gait recognition is used in smart phones with the help of the sensor like accelerometer to enhance the privacy of the smart phone. Nowadays, the most popular application of gait recognition is the health monitoring and surveillance system for elders and disabled persons. Also, many other applications of gait have introduced, which make gait recognition popular and hot topic in the area of Computer Vision.

Acknowledgment. We are very grateful to the honorable vice-chancellor of our university for his support and encouragement to make this survey.

References

1. Prathap C, Sumanth S (2015) Gait recognition using skeleton data. In: International conference on advances in computing, communications and informatics (ICACCI). IEEE, New York
2. Sharif M et al (2019) An overview of biometrics methods. Handbook of multimedia information security: techniques and applications. Springer, Cham, pp 15–35
3. Choudhury SD, Guan Y, Li C-T (2014) Gait recognition using low spatial and temporal resolution videos. In: International Workshop on Biometrics and Forensics (IWBF). IEEE, New York, pp 1–6
4. Sakata Atsuya, Takemura Noriko, Yagi Yasushi (2019) Gait-based age estimation using multi-stage convolutional neural network. *IPSJ Trans Comput Vis Appl* 11(1):4
5. Xu C et al (2019) Gait-based age progression/regression: a baseline and performance evaluation by age group classification and cross-age gait identification. *Mach Vis Appl* 30(4):629–644
6. Liu W et al (2018) Learning efficient spatial-temporal gait features with deep learning for human identification. *Neuroinformatics* 16(3–4): 457–471
7. Wang J, She M, Nahavandi S, Kouzani A (2010) A review of vision based gait recognition methods for human identification. In: 2010 International conference on digital image computing: techniques and applications (DICTA). IEEE, New York, pp 320–327
8. Sivapalan S, Chen D, Denman S, Sridharan S, Fookes C (2011) Gait energy volumes and frontal gait recognition using depth images. In: 2011 International joint conference on Biometrics (IJCB). IEEE Press, New York, pp 1–6, 181–184
9. Dikovski B, Madjarov G, Gjorgjevikj D (2014) Evaluation of different feature sets for gait recognition using skeletal data from kinect. In: 37th international convention on information and communication technology, electronics and microelectronics (MIPRO). IEEE, New York, pp 1304–1308
10. Jianwattanapaisarn N, Cheewakidakarn A, Khamsemanan N, Nattee C (2014) Human identification using skeletal gait and silhouette data extracted by microsoft kinect. In: 15th International symposium on soft computing and intelligent systems (SCIS), 2014 Joint 7th international conference on and advanced intelligent systems (ISIS). IEEE, New York, pp 410–414
11. Rahman MW, Gavrilova ML (2017) Kinect gait skeletal joint feature based person identification. In: 2017 IEEE 16th International conference on cognitive informatics & cognitive computing (ICCI* CC). IEEE, New York, pp 423–430
12. Preis J, Kessel M, Werner M, Linnhoff-Popien C (2012) Gait recognition with kinect. In: 1st international workshop on kinect in pervasive computing, New Castle, UK, pp P1–P4
13. Ahmed F, Paul PP, Gavrilova ML (2015) Dtw-based kernel and ranklevel fusion for 3d gait recognition using kinect. *Visual Comput* 31(6–8):915–924
14. Chattopadhyay P, Sural S, Mukherjee J (2014) Exploiting pose information for gait recognition from depth streams. In: Workshop at the European conference on computer vision. Springer, Berlin, pp 341– 355
15. Munsell BC, Temlyakov A, Qu C, Wang S (2012) Person identification using full-body motion and anthropometric biometrics from kinect videos. In: European conference on computer vision. Springer, Berlin, pp 91–100
16. Ahmed MH, Sabir AT (2017) Human gender classification based on gait features using kinect sensor. In: 3rd IEEE international conference on cybernetics (CYB-CONF). IEEE, New York, pp 1–5

17. Martín Félez R, García Jiménez V, Sánchez Garreta JS (2013) Gaitbased gender classification considering resampling and feature selection
18. Kastaniotis D, Theodorakopoulos I, Economou G, Fotopoulos S (2013) Gait-based gender recognition using pose information for real time applications. In: 18th International Conference on Digital Signal Processing (DSP). IEEE, New York, pp 1–6
19. Borrás, Lapedriza Á, Igual L (2012) Depth information in human gait analysis: an experimental study on gender recognition. In: International conference image analysis and recognition. Springer, Berlin, pp 98–105
20. Li X, Makihara Y, Xu C, Yagi Y, Ren M (2018) Gait-based human age estimation using age group-dependent manifold learning and regression. *Multimedia Tools Appl*, pp 1–22
21. Makihara Y, Okumura M, Iwama H, Yagi Y (2011) Gait-based age estimation using a whole-generation gait database. In: International joint conference on biometrics (IJCB). IEEE, New York, pp 1–6
22. Lu J, Tan Y-P (2010) Gait-based human age estimation. *IEEE Trans Inf Forens Secur* 5(4):761–770
23. Makihara Y, Mannami H, Yagi Y (2010) Gait analysis of gender and age using a large-scale multi-view gait database. In: Asian conference on computer vision. Springer, Berlin, pp 440–451
24. Kastaniotis D, Theodorakopoulos I, Theoharatos C, Economou G, Fotopoulos S (2015) A framework for gait-based recognition using kinect. *Pattern Recogn Lett* 68:327–335
25. Lu J, Wang G, Moulin P (2014) Human identity and gender recognition from gait sequences with arbitrary walking directions. *IEEE Trans Inf Forens Secur* 9(1):51–61