


# Kinect-based human gait identification under different covariate factors

Identificación de la marcha humana basada en Kinect bajo diferentes factores covariables

Autor:

 Azhin T. Sabir <sup>(1)</sup>

## ARTÍCULO DE INVESTIGACIÓN CIENTÍFICA Y TECNOLÓGICA

### How to cite this paper:

Azhin T. Sabir, Kinect-based Human Gait identification under different covariate factors. *Innovaciencia Facultad de Ciencias Exactas, Naturales y Agropecuarias*. 2018; 6(1) S1: 1-11.

### Reception date:

Received: 15 August 2018

Accepted: 25 November 2018

Published Online: 28 December 2018

### DOI:

<http://dx.doi.org/10.15649/2346075X.485>

### Keywords:

Neutral and non-neutral gait recognition; Kinect sensor; EDF; PCA; LDA; LDC.

## ABSTRACT

**Introduction:** Nowadays human gait identification/recognition is available in a variety of applications due to rapid advances in biometrics technology. This makes them easier to use for security and surveillance. Due to the rise in terrorist attacks during the last ten years research has focused on the biometric traits in these applications and they are now capable of recognising human beings from a distance. The main reason for my research interest in Gait biometrics is because it is unobtrusive and requires lower image/video quality compared to other biometric traits. **Materials and Methods:** In this paper we propose investigating Kinect-based gait recognition using non-standard gait sequences. This study examines different scenarios to highlight the challenges of non-standard gait sequences. Gait signatures are extracted from the 20 joint points of the human body using a Microsoft Kinect sensor. **Results and Discussion:** This feature is constructed by calculating the distances between each two joint points from the 20 joint points of the human body provided which is known as the Euclidean Distance Feature (EDF). The experiments are based on five scenarios, and a Linear Discriminant Classifier (LDC) is used to test the performance of the proposed method. **Conclusions:** The results of the experiments indicate that the proposed method outperforms previous work in all scenarios.

<sup>(1)</sup> Azhin T. Sabir, Department of Software Engineering, FENG, Koya University, Koya, Iraq, [azhin.sabir@koyauniversity.org](mailto:azhin.sabir@koyauniversity.org)

## INTRODUCTION

In the last decade, biometric systems have attracted considerable interest due to their potential use in forensics, access control and security surveillance applications. Gait is an attractive biometric with regards to human identification, in that it is unremarkable and easily captured at a distance compared with other biometrics such as the iris, fingerprints, face, etc. The hypothesis offered by <sup>(1)</sup> shows that gait is a characteristic that can be used to recognise a human being as a biometric system. Recently, gait recognition has become increasingly interesting to researchers due to the possibility of extracting a gait signature from a low resolution image, which is one of its main advantages over other biometrics that require high resolution images to gain information<sup>(2)</sup>. Early psychophysiological studies focused on gait recognition used moving light displays (MLDs) in the form of light points attached to the body, they concluded that individuals are able to recognise human motion solely by the movement of such MLDs <sup>(3)</sup>. Basically, gait identification/recognition approaches can be categorised in two parts: model-free and model-based<sup>(4)</sup>. Model-free works directly on the gait sequences without considering a model for the human body <sup>(2) (5)-(6)</sup>. In <sup>(2)</sup>, a mass vector is extracted as a gait signature, this feature is generated based on the number of pixels that represent the human body in the rows of silhouettes (i.e. pixels of binary value 1). To overcome the problem of direct frame-by-frame matching, the coordinates of the mass vectors are dealt with as time-series and hence the Dynamic Time Warping (DTW) method can be applied for matching purposes. This is due to the fact that in reality people may slightly alter their speed and style of walk. Gait Energy Image (GEI) and Gait Entropy Image (GEnI) methods are an efficient model-free gait representation that are widely proposed by the researcher to be used as a gait feature. GEIs are generated by calculating the average silhouette images to represent both body shape and movement in one

gait cycle <sup>(7)</sup>, while GEnI is based on applying entropy, a GEnI encodes in a single image the randomness of pixel values in the silhouette images during one gait cycle. The limitation of the two models (GEI and GEnI) is the lacking of robustness to cope with covariate conditions that affect the static areas of human body. Although GEnI is better than GEI in dealing with such problem, but still is not enough to offer good performance.

Conversely, model-based approaches use a structural model to extract a gait signature <sup>(8) (9) (10)</sup>. In <sup>(8)</sup> the structural parameters are extracted based on the stride and cadence to be used as a gait signature. The cadence is optioned by the walking periodicity, while the stride length is computed by calculating the ratio of the distance travelled and the steps taken.

Furthermore, many companies have tried to produce different types of devices to provide human body information. For instance, Microsoft released the Kinect sensor to support the Xbox gaming system. Accordingly, many researchers have focused on using a Kinect sensor in different applications. In <sup>(11)</sup>, there is a Kinect sensor that presents a health monitoring system to collect data on posture recognition. Furthermore, a Kinect sensor is used as an alternative to the traditional camera that used for biometric recognition <sup>(12) (13) (14)</sup>.

A stick-figure model is often used as a model in which the human body is characterised by joints and sticks <sup>(15)</sup>. The model constructs gait features using human body length parameters <sup>(15) (16)</sup> and <sup>(17)</sup>. In <sup>(16)</sup>, the authors created a gait database using a Kinect-sensor which involved gait information from ten participants. Each person walked five times to represent a training sample. In other words, there were 50 sequences of gait templates in the database. During the testing process, each participant walked five times in front of Kinect.

In <sup>(14)</sup>, the Kinect-based gait identification method proposed generating 12 static features (i.e. height, torso, both lower legs, right and left upper arms, the length of both legs, both thighs and both forearms), and two dynamic features (i.e. the step length and speed). To test their proposed method they created a database using a Kinect sensor, this was carried out by recording nine persons walking from the right to left side view of the Kinect sensor, and each participant doing this walk eight times. Three different classifiers were used to test the performance of the proposed scheme called: R1, C 4.5 and Naive Bayes. Although recently researchers widely focused on proposing human gait as one of the biometric recognition system, however one of the serious challenges of human gait is to address the effects of covariate factors such as clothing, carrying conditions, and view angle on recognition performance <sup>(4)</sup>.

There are model free gait recognition approaches in the literature that presumed to be robust to bag or clothing variations by selected parts of human body to be used as a gait signature (feature) <sup>(18) (19)</sup>.

In <sup>(19)</sup> GEI used and the best region of interest is selected by dividing the images into equal parts of horizontal and vertical and <sup>(18)</sup> proposed a method based on GEI which is decomposed into three independent shape segmentations; vertical, horizontal and grid resolution. Unfortunately, such type of approaches fails to predict all possible cases of covariate factors.

In <sup>(18)</sup> the Kinect-based gait recognition methods proposed were based on neutral and non-neutral gait sequences (wearing coat and carrying bags). In their proposed method two sets of dynamic features called Horizontal Distance Feature (HDF) and Vertical Distance Feature (VDF) are used to represent a gait feature (signature). To test the performance of the proposed feature set, they created a Kinect-based gait database called Kinect database-2. In this database, 20 participants were asked to walk in front of

the Kinect sensor and 25 walking sequences were recorded for each subject, consisting of five neutral (i.e. Set Nu), five wearing long coats (i.e. Set WLC), five wearing short coats (i.e. Set WSC), five carrying bags over their shoulders (i.e. Set CBS) and five carrying bags on their backs (i.e. Set CBB). Finally, they adopted the Linear Discriminant Classifier (LDC) as the classification method.

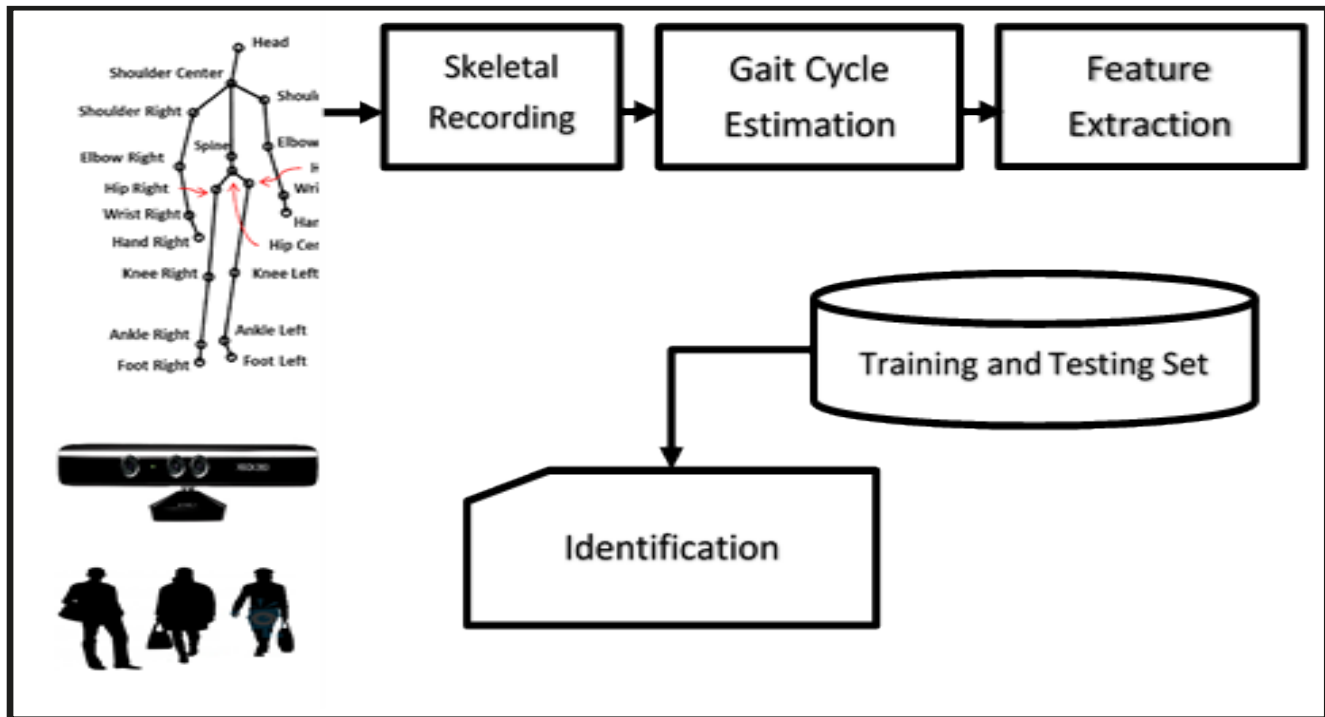
The current study focuses on gait identification using a Kinect sensor. The main objective of our work is to investigate neutral and non-neutral (wearing a coat and carrying bags) gait sequences based on a new feature set called EDF. We have used the same database used in <sup>(18)</sup>, which recorded 20 persons with 25 traits each.

In the rest of the paper we provide a detailed description of the proposed method, the experimental setup and results. Finally, we offer a conclusion and recommendations for future work.

## PROPOSED METHOD

In this method we recommend a set of Kinect-based features to be utilised for the human gait identification. As Mentioned in earlier sections, a Kinect sensor with a Software Development Kit (SDK) can record two persons' skeletons. In this method we used the same dataset created and used in <sup>(18)</sup>. Our proposed method is divided into four stages, as presented in Figure 1. The first stage starts with detecting the gait cycle for each participant in the dataset. The second stage includes feature extraction. At this stage a high dimensional feature set is extracted and used which is known as the Euclidean Distance Feature (EDF). Consequently, at the third stage, PCA and LDA respectively are applied with the aim of reducing the high dimensionality of the feature set. The final stage is the process of classification. At this stage a Linear Discriminant Classifier is used as the classification method. Further details on the proposed method provided in the following sub-sections.

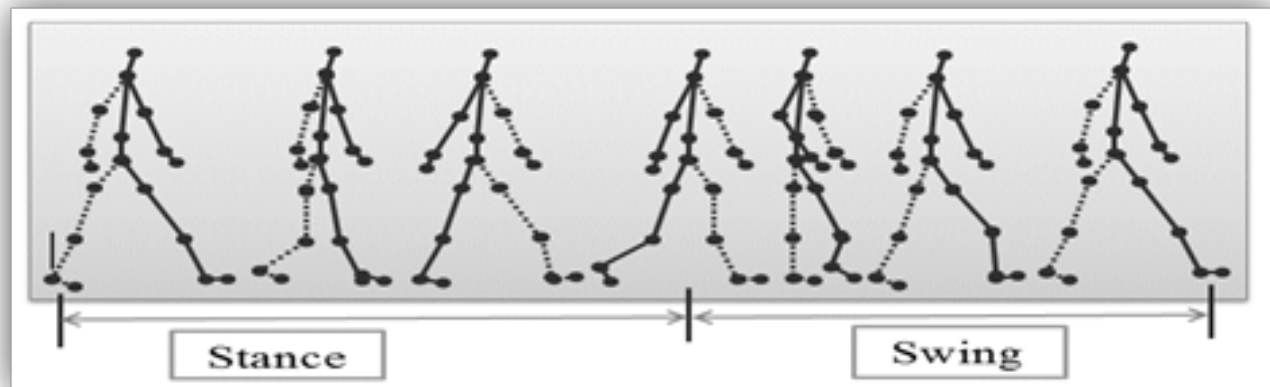
Figure 1: Diagram of the Proposed Gait Recognition System



### A. GAIT CYCLE ESTIMATION

The gait cycle is a repetitive modality of two steps, which can be measured from the time one foot hits the ground until the same situation occurs. The gait cycle is estimated using two steps based on two periods; the stance period and the swing period. The stance period includes 60 percent per one cycle, while the swing period covers the remaining fragment as shown in Figure 2.

Figure 2: The overview of human walk in one gait cycle

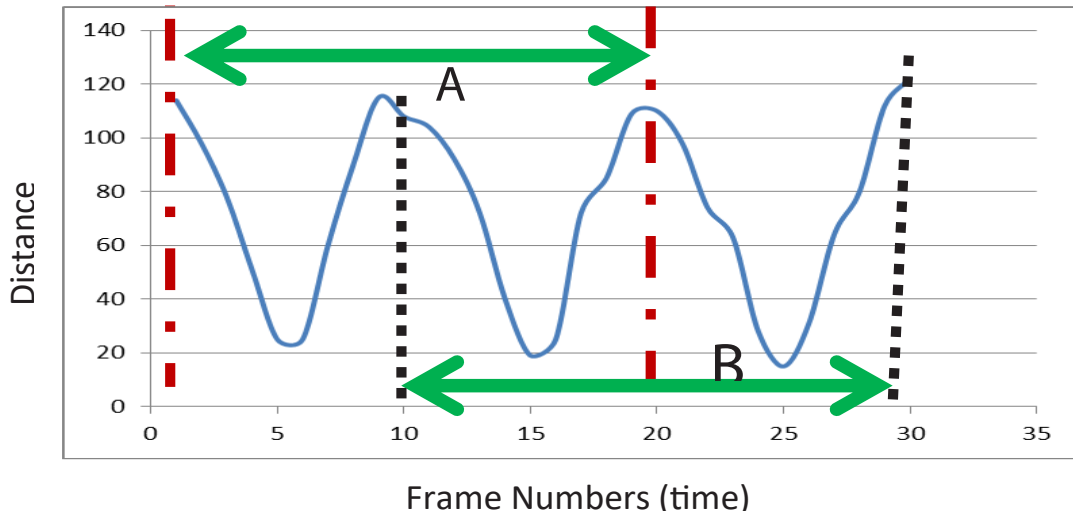


Usually, the cycle for gait is determined in two ways; local minima, (feet are close together). Secondly, local maxima (feet are wide apart). In this paper, we apply local maxima. To provide a more accurate gait cycle, the distance between the ankles was calculated instead of the feet, as is shown in equation (1).

$$\text{Distance of ankles} = |X_{\text{right ankle}} - X_{\text{left ankle}}| \dots\dots\dots( 1)$$

As shown in Figure 3; there are two gait cycles provided, A and B, and one of them needed to extract the features.

**Figure 3: Gait cycle detection**

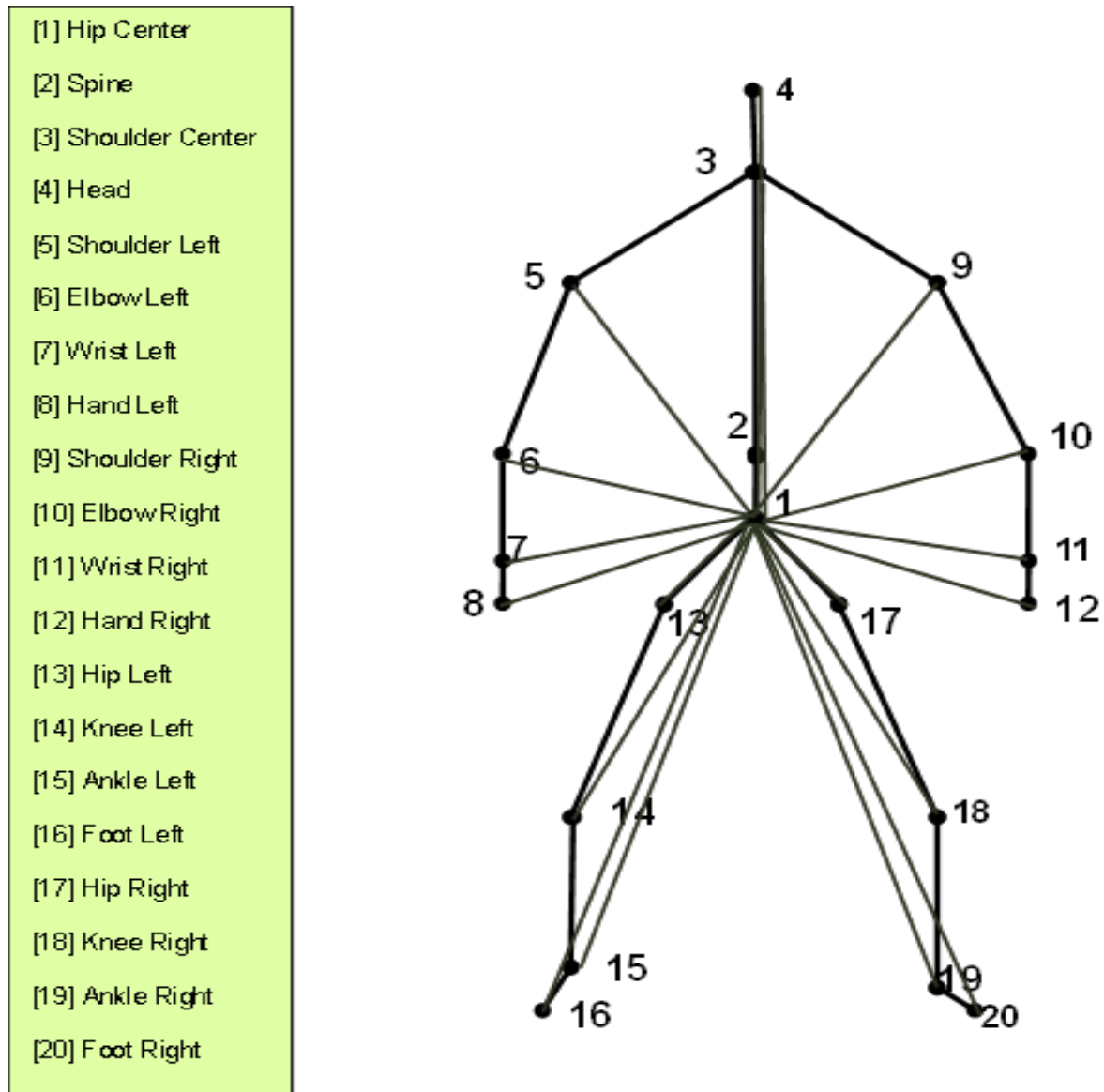


## B. EUCLIDIAN DISTANCE FEATURE

The feature extraction step is the process of determining the gait signature, this process using a Kinect sensor is different from the methods recorded as it is based on normal cameras. This is due to extracting features using the skeleton information provided by the Kinect sensor. The Kinect sensor with SDK for Windows offers 20 joint points for the skeleton; hence the database contains an X-axis and Y-axis for 20 joints. To extract the features vector, 190 distance values have been calculated from the 20 joint points provided by a Kinect sensor, and called the EDF.

The EDF vector is generated by applying various statistical moments. The statistical moments include mean, standard deviation, skewness and the slope of the measured Euclidian distances. Moreover, the 20 joint points provided by the Kinect sensor for the human skeleton of each frame during the gait cycle. Hence, the total number of calculated distances of all pair joint points will be equal to 190 distance values. Figure 4 shows an example of the first distance lines between the hip centre with the other 19 joint points.

Figure 4: Example of Distance Lines between Hip Centres with All Other Joint Points.



Afterwards, all the statistical moments will be measured regarding all of the frames in one gait cycle. Additionally, the length of the frame sequences for a gait cycle differ in the recording data from one person to another. Therefore, the statistical moments, such as mean, standard deviation, skewness and slope will be calculated based on the equations 2, 3, 4, 5, and 6 respectively:

$$MeanEDF = \{mean(d1_{i=1:n}), mean(d2_{i=1:n}), \dots, mean(dm_{i=1:n})\} \quad (2)$$

$$StdEDF = \{std(d1_{i=1:n}), std(d2_{i=1:n}), \dots, std(dm_{i=1:n})\} \quad (3)$$

$$SkewEDF = \{skew(d1_{i=1:n}), skew(d2_{i=1:n}), \dots, skew(dm_{i=1:n})\} \quad (4)$$

$$SlopEDF = \{slop(d1_{i=1:n}), slop(d2_{i=1:n}) \dots, slop(dm_{i=1:n})\} \quad (5)$$

$$EFD = [MeanEDF, StdEDF, SkewEDF, SlopEDF] \quad (6)$$



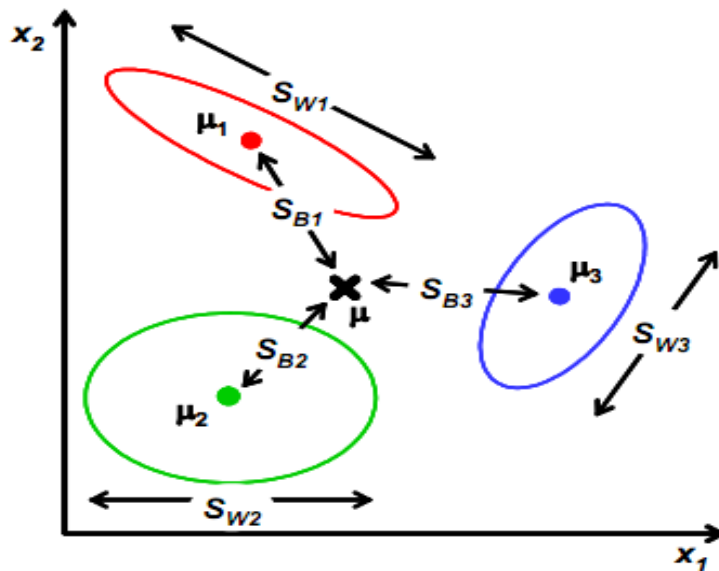
I is the index of each frame, n is the length of frames per gait cycle, d is the Euclidian distance value between each pair joint point from the person's body, and m is the number of all pairs over the 20 joint points of the human body.

Each of MeanEDF, StdEDF, SkewEDF and Slope EDF has 190 dimensional features, and the EDF feature set has 760 dimensions. Therefore, PCA and LDA respectively are applied in order to reduce the dimensions of the features vector.

### C. LINEAR DISCRIMINANT CLASSIFIER

Linear Discriminant Classifier (LDC) is a statistical technique used to find linear combination of the features set, that aim to best discriminates the data in the classes of interest. These various combinations are named discriminant functions. It is less expensive computationally compared to other classifiers i.e., artificial neural network. LDC one of the generative group of classifiers, the classes are anticipated to have normal distributions and equal covariance matrices <sup>(19)</sup>. Moreover, the optimal classifier reduces to calculating linear discriminant functions (see Figure 5).

Figure 5: LDC model,  $\mu$  is mean,  $S_W$  is Std of Within class and  $S_B$  is Std of between class. (blog.csdn.net)



### EXPERIMENTAL SETUP

To test the performance of the proposed method, the same dataset (Kinect database-2) and scenarios in <sup>(18)</sup> are implemented in the proposed method, to present a convenient comparison. Kinect database-2 contains the details of 20 participants who were asked to walk 25 times in front of the Kinect sensor

from the side view (at an angle of 90 degrees), to provide a total of 500 records (Figure 6). The gait sequence in this dataset is classified into neutral and non-neutral gait sequences. For each participant, five neutral (Nu), five wearing short coats (WSC), five wearing long coats (WLC), five carrying bags on their backs (CBB) and five carrying bags over their shoulder (CBS) gait sequences are recorded, which

offer experiments without the limitations of unbalanced class samples. The process of feature extraction, resulting in a set of a high dimensional features called EDF, to reduce the dimensions of EDF, PCA and LDA, have been applied sequentially for all the scenarios (except the first scenario). Moreover our experiments are based on LDC as a classification method that is used separately to test the perfor-

mance of the proposed method. The experiments in the proposed method were conducted under various scenarios each specifying the number and cases of samples of gait sequences involved in the training and testing sets. Table 1 presents five scenarios for the experiments, where Tr. stands for Training (Gallery) set and Te. Stands for Testing (Probe) set.

Figure 6 : Kinect dataset-2 example.



Table 1: Training and Testing samples set in five scenarios. (Tr is Training set and Te is Testing set).

Cases \ Scenarios	1		2		3		4		5	
	Tr	Te	Tr	Te	Tr	Te	Tr	Te	Tr	Te
<i>Nu</i>	5	Non	5	Non	5	Non	2	3	1	4
<i>CBB</i>	Non	5	1	4	2	3	2	3	1	4
<i>CBS</i>	Non	5	1	4	2	3	2	3	1	4
<i>WSC</i>	Non	5	1	4	2	3	2	3	1	4
<i>WLC</i>	Non	5	1	4	2	3	2	3	1	4
<i>NS</i>	5 Rec.		9		13 Rec.		10 Rec.		5	



## EXPERIMENTAL RESULTS AND DISCUSSION

In this paper we tested the performance of five different scenarios in relation to neutral (Normal) and non-neutral gait sequences (as mentioned in the previous section). Table 2 presents the results of all scenarios. In the first scenario the training set consists wholly of Nu cases (5 Nu gait sequences), while the testing set includes 20 gait sequences from non-neutral cases (five for each of CBB, CBS, WSC and WLC). The average result achieved in this scenario is 68%; the reason for this result is that the training case contains only Nu gait sequences. In the second scenario, the training set is included with one sample for each of the non-neutral cases. Hence we have nine samples in the training set and 16 in the testing set. The results of the second scenario show that the accuracy of the results improved significantly compared to what was attained in scenario one in all cases.

The third scenario locates more samples than scenarios one and two, by adding another sample from the non-neutral cases. Accordingly the training set contains 13 samples (five Nu and eight non-neutral).

The results show that the accuracy rate improved more compared to scenario two, by about 3% as an average of all the non-neutral cases. For scenario four, two samples from each of the neutral and non-neutral cases were used to create a training set. In this scenario we also tested the accuracy of the neutral cases as additions to the non-neutral cases. This experiment presents the result of a balanced number of neutral and non-neutral samples in the training and testing sets. The results show that the accuracy rate in this scenario outperforms the results of scenarios 1, 2 and 3, which provides 95.3% of accuracy rate.

In the final scenario we reduced the number of samples by using one sample for each type of gait sequence. In this scenario the accuracy was reduced compared to the other three scenarios which is expected. This scenario is similar to the first in terms of having the same number of samples in the training set, but the difference is that in scenario five, we used neutral and non-neutral samples in the training set, while in scenario one, we used only neutral samples. When comparing these two scenarios, the results achieved in scenario five are significantly higher than the results achieved in the first scenario.

**Table 2. Recognition Rate of the Proposed Method**

Scenarios	CBB	CBS	WSC	WLC	Nu	Average
First	79.0%	78.0%	60.0%	55.0%		68.0%
Second	97.0%	93.8%	86.3%	87.3%		91.1%
Third	97.3%	95.5%	92.7%	92.8%		94.6%
Fourth	99.0%	94.7%	92.0%	96.0%	95.0%	95.3%
Fifth	93.5%	90.8%	86.3%	86.5%	87.0%	88.8%

To the best of our knowledge, there is no study related to Kinect-based gait identification addressed covariate factors (wearing coat and carrying bags) except in [\(18\)](#), therefore, we shall compare our results with the results achieved by the method proposed in [\(18\)](#), where they used the same dataset and scenarios that we used in our proposed method (see Table 3).

**Table 3: Comparison based on the Recognition Rate**

		First Scenario	Second Scenario	Third Scenario	Fourth Scenario	Fifth Scenario
CBB	(18)	58.0%	85.0%	90.0%	91.0%	84.8%
	Our M.	79.0%	97.0%	97.3%	99.0%	93.5%
CBS	(18)	68.0%	85.5%	88.2%	88.3%	82.0%
	Our M.	78.0%	93.8%	95.5%	94.7%	90.8%
WSC	(18)	65.0%	82.8%	86.7%	87.0%	82.0%
	Our M.	60.0%	86.3%	92.7%	92.0%	86.3%
WLC	(18)	56.0%	81.3%	85.8%	83.0%	79.8%
	Our M.	55.0%	87.3%	92.8%	96.0%	86.5%
Nu	(18)				92.0%	86.5%
	Our M.				95.0%	87.0%
Average	(18)	61.8%	83.6%	87.7%	87.3%	82.1%
	Our M.	68.0%	91.1%	94.6%	95.3%	88.8%

The results shown in Table 3 demonstrate that the proposed method outperforms other schemes in all the scenarios with two exceptions (scenario 1, WSC and WLC).

## CONCLUSIONS

This paper presents a new Kinect-based human gait identification scheme under different covariate factors. The method tested was based on the Kinect database-2 which includes 500 records for 20 participants. In addition to five neutral gait sequence samples, the database contains 20 samples for each of CBB, CBS, WSC, and WLC as a non-neutral gait sequence sample. Furthermore, in this paper we investigated a high dimensional feature set called EDF. Since the EDF is a high dimension feature, PCA and LDA are used respectively to reduce the dimensions of EDF. In addition, LDA may lead to improvements

in the classification performance by maximising the separation between the classes. The performance of the proposed scheme was tested in five scenarios and LDC was used as a classification technique. The average results for all the cases in each scenario are presented, and the results indicate that the proposed method achieved significant results. Furthermore, the proposed method is compared with another scheme proposed in the literature and based on the same dataset and scenarios. The compared results shows that the proposed method outperforms the other method in most of the cases and scenarios. In further work, we aim to extend the current study to investigate another feature set to be fused with EDF to cope with the cases from scenario one. Moreover, we plan to investigate the combination of gait and face biometric traits.

## REFERENCES

1. Bharatkumar A, Daigle K, Pandey M, Cai Q, Aggarwal J. Lower limb kinematics of human walking with the medial axis transformation. In: Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects; 1994 Nov. 11-12; Austin, TX, USA, USA; 1994. p. 70-76.
2. Hong S, Lee H, Nizami IF, Kim E. A new gait representation for human identification: mass vector. In: 2007 2nd IEEE Conference on Industrial Electronics and Applications. ICIEA 2007.; 2007. ICIEA 2007; 2007 May 23-25; Harbin, China; 2007. p. 669-673.
3. Cutting JE, Kozlowski LT. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the psychonomic society*. 1977; 9(5): p. 353-356.
4. Rida I, Almaadeed N, Almaadeed S. Robust gait recognition: a comprehensive survey. *IET Biometrics*. 2018;8(1):14-28.
5. Hsia CH, Chiang JS, Dai YJ, Lin TA. Conditional-Sorting Local Binary Pattern Based on Gait Energy Image for Human Identification. Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICCV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). 2013 P. 1.
6. Hofmann M, Rigoll G. Improved gait recognition using gradient histogram energy image. In: 19th IEEE International Conference on Image Processing (ICIP); 2012 Sep 30 - Oct 3; Orlando, FL, USA. IEEE; 2012. p. 1389-1392.
7. Han J, Bhanu B. Individual recognition using gait energy image. In: IEEE Transactions on Pattern Analysis and Machine Intelligence; 2006;28(2):316-322.
8. BenAbdelkader C, Cutler R, Davis L. Stride and cadence as a biometric in automatic person identification and verification. In: Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition; 2002 May 21-21; Washington, DC, USA, USA. IEEE; 2002. p. 372-377.
9. Singh JP, Jain S. Person identification based on gait using dynamic body parameters. In *Trends in Information Sciences and Computing (TISC)*; 2010 Dec 17-19; Chennai, India. IEEE; 2010. p. 248-252.
10. Bharti J, Gupta M. Human Gait Recognition using All Pair Shortest Path. In: 2011 International Conference on Software and Computer Applications IPCSIT; 2011. p. 279-284.
11. Parajuli M, Tran D, Ma W, Sharma D. Senior health monitoring using Kinect. In: 2012 Fourth International Conference on Communications and Electronics (ICCE); 2012 Aug 1-3; Hue, Vietnam. IEEE; 2012. p. 309-312.
12. Ball A, Rye D, Ramos F, Velonaki M. Unsupervised clustering of people from 'skeleton' data. In Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction; 2012: Machine Vision and Image Processing Conference, 2009. IMVIP'09. 13th International; 2012 Mar 5-8; Boston, Massachusetts, USA. IEEE; 2012. p. 225-226.
13. Sinha A, Chakravarty K, Bhowmick B. Person identification using skeleton information from kinect. In *ACHI 2013, The Sixth International Conference on Advances in Computer-Human Interactions*; 2013. p. 101-108.
14. Preis J, Kessel M, Werner M, Linnhoff-Popien C. Gait recognition with kinect. In: 1st International Workshop on Kinect in Pervasive Computing; 2012.
15. M. Ahmed NAJaATS. Gait recognition based on Kinect sensor. In: *Real-Time Image and Video Processing*; 2014; Brussels, Belgium.
16. Jiang S, Wang Y, Zhang Y, Sun J. Real time gait recognition system based on kinect skeleton feature. Springer. 2014; p. 46--57.
17. Ahmed F, Paul PP, Gavrilo ML. DTW-based kernel and rank-level fusion for 3D gait recognition using Kinect. *The Visual Computer*, Springer; 2015 Jun 6-8; 31(6-8):915-924.
18. Sabir AT. Human gait recognition under neutral and non-neutral gait sequences. University of Buckingham, UK: PhD Thesis; 2015.
19. Rezaei A, Mirzakochaki S. A Novel Approach for Keyboard Dynamics Authentication based on Fusion of Stochastic Classifiers. *IJCSNS*. 2012; 12(8): p. 60.