Multiview Running and Walking Gait Analysis using the First and Third Person Data

Nikita Malik¹*, Sudipta Majumdar¹

¹Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, 110042, India

E-mail: *nikitamalik70@gmail.com

Abstract. Gait recognition, which allows to recognise someone from a distance, has a lot of applications. The gait of a person is commonly used as a biometric approach to identify or categorise them by gender and age. Biometric systems are a fast-growing subject that necessitates the development of creative solutions to problems that have plagued previous attempts. By initially looking at the design of a gait detection system, two different types of gait datasets have been reported and presented in this paper. While running at the same time, the first person (FP) data containing the camera motion gathered from the movement of the volunteer's body and the third person (TP) data captured from a distant view were recorded. The dataset contains a total of 23 participants (14 males and 9 females). The discussion is expanded to include a comparison of the results obtained using TP and FP data, as well as an examination of the physiological motions recorded while running and walking.

1. Introduction

Biometric systems have been around for a long time, but they are continually evolving as new and effective human identification methods are added to the list. Human gait recognition is one of those newer approaches that has seen a lot of research and yet requires more. The study of human motion using brain and eye observations, as well as sensors that detect bodily movements, is known as human gait. The technology can be used to do remote biometric identification without the need for an observer. It excels in the field of biometrics due to its ability to recognize objects from a distance. It is a technology that extracts features from a person's body motions in order to identify them. It has a wide range of applications and is commonly utilized in sports, health monitoring, surveillance, criminal investigation, and gender and age classification. Many gait acquisition technologies exist, including wearable sensors data (accelerometers, gyroscopes, pressure and force sensors), high-definition cameras, and non-wearable sensors data. Model-based and appearance-based strategies are the two types of gait identification approaches. Some earlier state-of-the-art surveys are available for gait recognition which provides a detailed overview regarding the progress made so far in the area of human gait recognition systems. The evaluation part has been a history and the area of interest for many researchers is the application part in which not much evaluation has been done till now. The work is divided into six sections. In section 2, the framework of the gait recognition system is briefly outlined, followed by a full description of the proposed work in section 3. In addition, in section 4, the recorded dataset including the running patterns of 23 volunteers is described and investigated, and in section 5, the results acquired from the databases are compared and discussed, followed by a conclusion and a scope for future work.
2. Framework of the Gait Recognition System

Gait data acquisition, pre-processing of the acquired data, feature extraction, data reduction or feature selection, and classification and recognition are the five stages of the gait identification system, as shown in fig.1.

![Diagram of the Gait Recognition System](image)

**Figure 1.** Framework of Gait Recognition System.

2.1. Data Recording and Preprocessing Techniques

Sensor-based, video-based, and radar-based data collecting techniques are all viable options. [1] The data recording and processing techniques are accelerometers, floor-sensors, camera, and radar. Depending on the nature of the application, the cell phones are increasingly incorporating accelerometers. The device's reduced size allows it to be easily carried in a pocket. The device makes use of Bluetooth with a distant station to carry out the connection. The calculations are done at a remote location. The subject of the investigation must be required to carry the gadget with him at all times.

RGB-D camera and Microsoft Kinect cameras are some examples of the data acquisition camera-based devices in which the RGB-D camera is used to input information about a person's depth and color. Kinect is a Microsoft-designed device that uses an infrared projector and sensors to acquire depth data. The infrared laser fired by Kinect creates an erected design of the bodily surface. Based on the principle of structured light, the infrared camera detects the element of depth in this network. Kinect comprises of a color video camera with a unique microphone configuration.

Different types of floor-sensors like capacitive, resistant and coax cable (used just as a delay line) are used for acquiring different types of gait data. Following data gathering, noise removal and background modelling are performed using filters (Gaussian filters, mean filters, and so on) and background/foreground subtraction methods (frame difference, GMM, and so on).

2.2. Feature Extraction

Following the pre-processing of obtained data, the next step is to extract some characteristics from the acquired gait data, which is noise-free and segregated using background subtraction algorithms. A variety of methods are available for extracting features categorized into model-based and model-free techniques. The goal of model-based feature representation is to model the human body and extract features from it. Model-based feature representation is size and view independent. However, it is dependent on the video quality. On the other hand, the goal of model-free feature representation is to process a human silhouette's whole motion or shape. There are two major advantages to feature representation without using a model. For starters, it is unaffected by video quality. As a result, it can be used far away from the human object. Second, it is less expensive to compute than model-based feature representation. As a result, model-free feature representation is becoming increasingly common.
2.3. Feature Selection/Reduction

Feature reduction is a type of dimension reduction strategy that can help improve the gait biometric system's effectiveness and efficiency by removing some unimportant features. As stated below in table 1, there are a variety of approaches for reducing features.

Table 1. Survey of some state-of-the-art feature selection algorithm.

| Feature sel. | Algorithm |
|--------------|-----------|
| CSA+DATER [2] | The pre-processing algorithm CSA, which is matrix-based, reduces noise and keeps the best information, while DATER is utilized to improve classification abilities. |
| RF [3] | A backward feature elimination search technique looks across the subspaces while the Random Forest algorithm scores the features. |
| PCA [4] | On wavelet features, PCA performs feature separation. Dimensionality features can be efficiently separated for classifiers to use as features up to the 6th principal component. |
| WPCA [5] | In WPCA, eigenvectors were separated using the square root of the eigenvalues. The positive impact of eigenvectors that convey specific knowledge increases, whereas the negative impact of eigenvectors with higher eigenvalues declines. |
| BD [6] | Using the principles of binomial distribution, BD applied to random characteristics retrieves some successful features (mean and variance). A simple algorithm is then constructed to select/extract the best features. |
| FEcS [7] | Using the FEcS technique, the fused feature vector (FV) produces entropy and skewness vectors and selects the optimal subsets of features. |
| Firefly algorithm and Skewness based approach [8] | On fused vector, the Firefly algorithm (a meta-heuristic technique) was implemented. Using fused vectors, a skewness-based technique finds the best features. |
| LDA [9] | LDA reduces information loss (m < n where n = dimension of feature data) by using an orthogonal transformation technique to provide m-dimensional feature data that are unrelated and utilized for classification. The algorithm chooses the direction in which substantial data variation occurs. As the orthogonal transformations' axis. |

2.4. Classification

It's the final step in the gait recognition process. This stage's accuracy is mainly reliant on the preceding levels. Machine learning algorithms, which can be supervised, unsupervised, or reinforcement learning algorithms, are used to classify data. For human gait classification, classifiers such as Nearest-Neighbour (NN), KNN (k-nearest neighbour), Multilayer Perceptron, Decision Tree, SVM, Nave Bayes, Random Forest, DCNN, Regression Techniques, Random Tree, BayesNet, QDA (quadratic discriminant analysis), Neuro-Fuzzy classifier, CART (classification and regression tree), PNN (probabilistic neural network), FIS (fuzzy interference system) were used in the recent research work.
3. Proposed Work

![Diagram showing the framework of the proposed approach]

Figure 2. Framework of the Proposed Approach.

- As shown in Fig.2, the first-person (FP) video was split into frames, and certain pre-processing operations were carried out to remove the video's empty and turning regions.
- Each image frame is divided into its R, G, and B colour components, with horizontal and vertical projection vectors calculated for each component. After calculating three separate horizontal and vertical projection vectors for each of the three components (Red, Green, and Blue), the average projection vectors are calculated using the sample mean.
- Normalizing both the horizontal and vertical projection vectors after calculating the average horizontal and vertical projection vectors yields normalised horizontal and normalised vertical projection vectors.
- The peak value of cross-correlation is calculated for each consecutive normalised horizontal projection vector value, and the position or perfect shift value at which the peak value of cross-correlation occurs is investigated further. Similarly, for each subsequent normalised vertical projection vector value, a cross-correlation value is calculated, and the peak value of cross-correlation is acquired to study the position or perfect shift value at which the peak value of cross-correlation occurs.
- The number of local maxima peaks in an oscillation is compared to the number of steps (or gait cycles) a person took while running using all of the horizontal perfect shift values acquired. Similarly, the number of local maxima peaks in an oscillation is compared to the number of steps count (or gait cycles) a person has by plotting all of the vertical perfect shift values acquired against time.

4. Dataset and Description

The implementation of a novel model on a database is critical to its development and implementation. Testing on a database of appropriate size and variant factors is required to obtain successful results for the model. In the last several years, a number of datasets in the field of human gait identification have been made publicly available for research purposes. Every other database has been shown to be effective in terms of a variety of parameters, including baggage carrying conditions, garment variations, covariates, changes in location or view in footwear, indoor or outdoor circumstances, and so on. Table 2 shows a descriptive survey-based study of certain publicly available human gait datasets.

Table 2. Available Public Datasets used for Human Gait Analysis.

| Data               | sub./seq. | cam/view/style     |
|--------------------|-----------|--------------------|
| UMD                | 55 sub.   | 2 cam/ 4 view      |
| CASIA A [10]       | 20 sub/ 240 seq. | 3 views          |
| CASIA B [11]       | 124 sub.  | 11 views/ 4 style  |
| CASIA C [12]       | 153 sub./ 1530 seq. | 1 cam          |
| floor-sensor-based [13] | 15 sub.     | Different footwears |
| Dataset Name                  | Description                                           | Views | Notes |
|------------------------------|-------------------------------------------------------|-------|-------|
| Georgia-Tech [14]            | 18 sub./ 20 seq.; 15 sub./ 268 seq.                  |       | multiple view |
| TUM-IITKGP [15]              | 35 sub./ 840 seq.                                    | 1 cam/ 6 style |
| HUMANID [16]                 | 74 sub./ 452 seq.                                    | 2 cam/ 12 style |
| Motion-Recording-Sensor-Based Dataset [17] | 50 sub./ almost 100 seq.                      | 2 styles |
| MIT [18]                     | 24 sub./ 194 seq.                                    | 1 cam/ 1 view |
| Human Activities and Postural Transitions Dataset | 30 sub.                                           | style: 3 static postures/ 3 dynamic motions |
| HID UMD 1 and 2              | 25 sub./ 100 seq.; 55 sub./ 222 seq.                 | 2 views |
| Ground Reaction forces       | 62 sub.                                               |       | Data taken as body mass |
| TUM-GAID                     | 305 sub./ 3370 seq.                                  |       | Multimodal gait recognition, RGB-D sensor with a 4-channel microphone array |
| OU-ISIR A, B and D           | 34 sub./ 408 seq.; 68 sub./ 1350 seq.; 185 sub./ 370 seq. | 25 views/ 9 speed variations/ 32 style |
| UCSD                         | 6 sub./ 42 seq.                                      | 1 cam (stationary camera mounted at an elevation) |
| UWB Impulse Radar Prototype  | -                                                    |       | The individual marched/ walked with either one-arm or two-arm swings. |
| SOTON large dataset          | 115 sub./ 2128 seq.                                  | multiple view |
| First wave-radar-based dataset | 49 sub.                                               | Signal reflected from the person’s torso, legs, and arms. |
| SOTON small dataset          | 12 sub.                                               | 3-4 view |
| SOTON temporal dataset       | 25 sub/ 2280 seq.                                    | multi-view |

### 4.1 Dataset Generation
The purpose of recording this database, which contains two categories of data, is to serve both research and practical needs.

Two sets of data are collected using two separate phone cameras on a database of 23 participants (14 males and 9 females). One camera is mounted on a tripod stand for collecting third-person running data, and the tripod stand is positioned at a perpendicular bisector drawn at line segment AB. The line segment AB connects the points A and B on the running track. The side perspective of a person’s running manner is captured by the TP data. The other camera, which is used to collect first-person (FP) data, is attached to the volunteer’s body with the use of a waist belt. The FP data, as well as the timestamp information needed to sync it with the TP data, is recorded with the use of a phone camera. The FP data is then utilised to analyse the camera’s motion in order to estimate body movements.

Only one volunteer’s running data is collected at a time, and two cameras record two sorts of data. A person was instructed to simply run with his or her natural running style while wearing the FP camera on his or her belt, while the TP camera on the tripod stand was turned on at around the same time. Data was collected from participants of both genders in the following age groups: below 5 years, 5-14, 15-24, 25-34, 35-49 and 50-65 years of age.

### 4.2 Precautions Taken
- The number of steps in both the TP and FP camera recorded data should be the same to infer the two forms of data from one another.
While recording the data, neither of the cameras should have direct sunlight in their view. In the frame of view of the cameras, there should be no sharp shadows thrown by the subject or volunteer. Data should be recorded after sunset and before nightfall, or on cloudy days should be recorded.

A significant number of samples from various age groups, with enough examples in each category, should be available to raise doubts about the statistical significance. Any category may only be thoroughly examined if a vast number of instances are included.

5. Discussion
The running database simulations are compared to the database simulations that include the walking patterns of the same 23 participants.

**Table 3.** The steps count and positive peaks count from the obtained horizontal perfect shift plot for the volunteers walking on straight line AB: classified as underweight, normal, overweight, and obese are used to compare the TP and FP video results.

| AGE GROUP (relation B/W TP and FP data) | GENDER - AGE | STEPS COUNT (TP) | PEAK COUNT (FP-HORIZONTAL) | BMI (Body Mass Index) | Health |
|----------------------------------------|-------------|------------------|-----------------------------|----------------------|--------|
| Below 6 (exactly double)               | M – 5 years | 38               | 76                          | 16.5                 | Normal |
|                                           | F – 7 years | 31               | 70                          | 13.2                 | Underweight |
|                                           | F – 10 years | 25              | 73                          | 16.8                 | Normal |
|                                           | M – 12 years | 25              | 74                          | 15.3                 | Normal |
|                                           | M – 12 years | 23              | 73                          | 16.8                 | Normal |
|                                           | M – 13 years | 24              | 70                          | 16.4                 | Normal |
|                                           | M – 13 years | 24              | 60                          | 15.1                 | Underweight |
|                                           | M – 14 years | 22              | 54                          | 15.4                 | Underweight |
| 15-30 (almost double)                   | M – 15 years | 22              | 45                          | 19.2                 | Normal |
|                                           | F – 15 years | 27              | 56                          | 21.3                 | Normal |
|                                           | F – 16 years | 27              | 55                          | 20.6                 | Normal |
|                                           | M – 21 years | 26              | 53                          | 27.1                 | Overweight |
|                                           | M – 22 years | 23              | 59                          | 30.8                 | Obese |
|                                           | F – 25 years | 24              | 50                          | 23.4                 | Normal |
|                                           | F – 25 years | 25              | 51                          | 22.4                 | Normal |
|                                           | F – 32 years | 24              | 25                          | 23.9                 | Normal |
|                                           | M – 34 years | 24              | 27                          | 27.5                 | Overweight |
| AGE GROUP (relation B/W TP and FP data) | GENDER - AGE | STEPS COUNT (TP) | PEAK COUNT (FP-HORIZONTAL) | BMI (Body Mass Index) | Health |
|----------------------------------------|--------------|------------------|----------------------------|----------------------|--------|
| Below 6 (almost double)                | M – 5 years  | 23               | 40                         | 16.5                 | Normal |
|                                        | F – 7 years  | 15               | 26                         | 13.2                 | Underweight |
|                                        | F – 10 years | 18               | 34                         | 16.8                 | Normal |
|                                        | M – 12 years | 18               | 35                         | 15.3                 | Normal |
|                                        | M – 12 years | 15               | 32                         | 16.8                 | Normal |
|                                        | M – 13 years | 22               | 44                         | 16.4                 | Normal |
|                                        | M – 13 years | 22               | 44                         | 15.1                 | Underweight |
|                                        | M – 14 years | 20               | 38                         | 15.4                 | Underweight |
| 15-30 (almost double)                  | M – 15 years | 14               | 34                         | 19.2                 | Normal |
|                                        | F – 15 years | 18               | 26                         | 21.3                 | Normal |
|                                        | F – 16 years | 19               | 28                         | 20.6                 | Normal |
|                                        | M – 21 years | 16               | 24                         | 27.1                 | Overweight |
|                                        | M – 22 years | 15               | 30                         | 30.8                 | Obese |
|                                        | F – 25 years | 16               | 25                         | 23.4                 | Normal |
|                                        | F – 25 years | 16               | 25                         | 22.4                 | Normal |
| 31-49 (almost double)                  | F – 32 years | 21               | 39                         | 23.9                 | Normal |

Table 4. The steps count and positive peaks count from the obtained horizontal perfect shift plot for the volunteers running on straight line AB: classified as underweight, normal, overweight, and obese are used to compare the TP and FP video results.
Some discussions can be conducted using the tabular data provided in table 3 and table 4.

The volunteers ranged from overweight to obese, and when their walking data was assessed, the camera motion was higher for obese people than for underweight people. In table 3, the number of positive peaks in an oscillation obtained from the FP data plot was exactly double that of the number of steps obtained from the TP video data in the case of a 5-year-old healthy male, and the data obtained from the FP videos is almost three times that obtained from the TP videos in the next age group, which is 6-14 years old. In terms of steps count and peak, there is a difference for non-healthy volunteers, and the FP data for underweight individuals is less than twice that of the TP data, indicating less camera movement. The data acquired from FP videos for the 15-30 range is twice that obtained from FP videos, indicating that camera movement was higher in fat and overweight volunteers than in healthy subjects. Peak counts from FP videos are substantially equal to steps counts from TP videos exhibiting less camera movement while walking for volunteers aged 31 to 49.

Whereas in table 4, the number of positive peaks in an oscillation obtained from the FP data plot was almost double that of the number of steps obtained from the TP video data in the case of a 5-year-old healthy male, also in the case of rest of the age groups: a uniformity can be seen in terms of the camera motion as for every age group the data obtained from the FP videos is almost twice that the data obtained from the TP videos. While running, there is a uniform shakiness or movement of the volunteers' bodies, which causes the camera to move. This is due to the fact that when compared to walking, running involves more vigorous body motions.

6. Conclusion and Future Direction

A small-scale database is proposed, which contains data from 23 volunteers and records their running motion (14 males and 9 females). For study, two types of data are collected: FP (first person) and TP (third person). The FP data is used to perform proper camera motion analysis, whereas the TP data is simply utilised to track gate cycles or the number of steps count for various subjects. According to the simulated experiments the running motion of the volunteers leads to more camera movements caused by vigorous bodily movements, the camera movements for every volunteer belonging to different age groups are uniformly double than the number of step counts estimated from the TP videos data. In the case of running, the camera motion is irrespective of the individual’s BMI and health, and there is more shakiness or movement of the camera in running than in walking. When walking, an underweight volunteer of any age group has fewer body movements and thus produces fewer camera movements than a healthy volunteer, whereas an obese and overweight volunteer of any age group has more body movements and thus produces more camera motion than a healthy participant, according to the findings. Depending on the age group and gender, the camera moves in different ways. Future analysis could infer FP and TP data from one another, such as predicting joint angles and determining the cross-correlation between the two databases for both running and walking motions. Individual recognition can be accomplished by analysing their stride, which entails deducing the FP and TP data from one another.
References

[1] Kumar M, Singh N, Kumar R, Goel S, and Kumar K 2021 Gait recognition based on vision systems: systematic survey J. of Vis. Comm. and Imag. Repre. vol 75 pp. 103052-64.

[2] Dong X, Shuicheng Y, Dacheng T, Lei Z, Xuelong L and Hong-Jiang 2006 Z Human Gait Recognition with Matrix Representation IEEE Trans. on Circ. and Sys. for Video Tech. vol 16 pp. 896-03.

[3] Dupuis Y, Savatier X and Vasseur P 2013 Feature subset selection applied to model-free gait recognition Imag. and Vis. Computing vol 31 pp. 580-91.

[4] Yuwono M, Su SW, Guo Y, Moulton BD and Nguyen HT 2014 Unsupervised nonparametric method for gait analysis using a waist-worn inertial sensor Appl. Soft Computing vol 14 pp. 72-80.

[5] Uzun-Per M and Gökmen M 2018 Face recognition with Patch-based Local Walsh Transform Sig. Proc.: Image Comm. vol 61 pp. 85-96.

[6] Arshad H, Khan MA, Sharif M, Yasmin M, and Javed MY 2019 Multi-level features fusion and selection for human gait recognition: an optimized framework of Bayesian model and binomial distribution Int. J. Mach. Learn. & Cyber. vol 10 pp. 3601–18.

[7] Arshad H, Khan MA, Sharif M, Yasmin M, Tavares JM, Zhang YD and Satapathy SC 2020 A multilevel paradigm for deep convolutional neural network features selection with an application to human gait recognition Expert Systems pp. 12541-56.

[8] Mehmood A, Khan MA, Sharif M, Muhammad, Khan S, Shaheen M, Saba T, Riaz, Riaz N and Ashraf I 2020 Prosperous Human Gait Recognition: An End-to-End System based on Pre-trained CNN Features Selection Multi. Tools and App.

[9] Guo H, Li B, Zhang Y, Zhang Y, Li W, Qiao F, Rong X and Zhou S 2020 Gait Recognition Based on the Feature Extraction of Gabor Filter and Linear Discriminant Analysis and Improved Local Coupled Extreme Learning Machine Mathe. Prob. in Eng. vol 2020 pp. 245-54.

[10] Yu S, Tan D and Tan T A 2006 Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition 18th Int. Conf. on Pattern Recog. (ICPR’06) Hong Kong China pp. 441-44.

[11] Wang L, Tan T, Ning H and Hu W 2003 Silhouette analysis-based gait recognition for human identification IEEE Trans. on Pattern Analy. and Mach. Intell. vol 25 pp. 1505-18.

[12] Tan D, Huang K, Yu S and Tan T 2006 Efficient Night Gait Recognition Based on Template Matching 18th Int. Conf on Pattern Recog. (ICPR’06) Hong Kong China pp. 1000-03.

[13] Orr RJ and Abowd GD 2000 The smart floor: a mechanism for natural user identification and tracking CHI EA '00: CHI '00 Extended Abstracts on Human Factors in Computing Systems, NY USA pp. 275–76.

[14] Johnson AY and Bobick AF 2001 A Multi-view Method for Gait Recognition Using Static Body Parameters AVBPA Audio- and Video-Based Biometric Person Authentication pp. 301-1.

[15] Hofmann M, Sural S and Rigoll G 2011 Gait Recognition in the Presence of Occlusion: A New Dataset and Baseline Algorithms Proc. of the 19th Int. Conf. in Central Europe on Comp. Graph. Vis. and Comp. Vis. pp. 99–04.

[16] Phillips PJ, Grother P, Sarkar S, Vega IS and Bowyer K 2002 Baseline Results for the Challenge Problem of Human ID Using Gait Analysis 5th IEEE Int. Conf. on Automatic Face and Gesture Recog. (FG) with CD-ROM Washington D.C. USA pp. 137-142.

[17] Gafurov D, Snekkens E and Bours P 2007 Gait Authentication and Identification Using Wearable Accelerometer Sensor IEEE Wor. on Aut. Identification Advan. Techn. pp. 220 – 25.

[18] Collins RT, Gross R and Shi J 2002 Silhouette-based human identification from body shape and gait Proc. of 5th IEEE Int. Conf. on Automatic Face Gesture Recog. pp. 366-71.