CHAPTER 2

Human identification using Model-free gait analysis approach

2.1 Introduction

Gait recognition is a biometric system that uses a person's walking manner to identify individuals. Compare other biometric indices such as face or finger and gait less obtrusive method of identification that helps to identify people at a distance without the cooperation of the participants. This is the feature that makes gait so appealing as a form of recognition. With the use of automatic tracking systems in surveillance, the development of a reliable identification and authentication system is critical. Gait is an excellent biometric index because it can be used to identify an individual from a distance using a low-resolution snapshot. The approach for collecting gait data is dependent on functionality and applicability. Gait features can be determined from both space and temporal variables. The classification model selection is based on the feature selection from the extracted gait features.

In this chapter, a model-free gait analysis is implemented to explore gait pattern classification as a method of identifying human walking patterns for identification purposes. It compares various classifiers for gait-based human identification systems under covariate conditions. The CASIA-B dataset is used in the experimental analysis, with GEI employed for gait feature representation and HOG for feature extraction. The study selects the discriminant function for classification using linear discriminant analysis and inputs feature vectors into several classical classifiers, including SVM, RF, KNN, and NC. The experimental results suggest that the nearest centroid classification model is an effective classifier for gait pattern classification, even when accounting for viewing covariates and appearance changes due to carrying and clothing covariates.

2.2 Related work

Over a new state-of-the-art monitoring system, gait has grown in popularity as a biometric index. In the past, a number of studies have been carried out to support the usefulness of gait for identity recognition within a sample population. However, there are a number of factors that can make gait less useful as a biometric identifier[15]. The factors such as the environment setup in which the subject is being captured, the imaging device utilized to capture the subjects, and the appearance of the subject. The two types of gait analysis approaches are model-based and model-free. Primitive or arbitrary shapes are commonly used in model-based to construct the body model and define the edges of the body parts [63]. To extract the gait feature using a model-based approach, models of the human body are fitted on walking sequences of each frame of the gait cycle. Therefore, prior knowledge of the object is required for extracting features from the model-based approach [11]. When modeling the human body, we can apply several different kinematical and physical limitations, such as the mean error in body joint angles such as hip to knee/knee to ankle. The strengths of a model-based approach include the ability to use evidence-collection methods across the entire image series before settling on a model design. The model-based approach can better handle occlusion and noise, as well as derive gait signatures directly from model parameters, such as thigh inclination variance. It also contributes to lessening the number of measurements necessary to display the results. Because of the extensive matching and sorting that must be done, using a model-based method has the disadvantage of incurring large computing costs.

The model-free represents the most holistic approach while handling visual gait analysis as suggested by many leading researchers. The model-free approach has more advantages in terms of methods to extract features, has low computation cost, and can handle low-resolution images, perfect for outdoor application [68]. A growing number of background subtraction algorithms are being developed to provide additional robustness to scenarios [85].

An enhanced spatial-temporal gait variable known as gait energy image (GEI) was proposed for efficient individual recognition [30]. This proposed variable normalized and averaged the silhouette into one gait cycle. Wang et al. [90] suggested the use Chrono-Gait Image gait template (CGI) as an efficient gait feature for human identification system. CGI is calculated by utilizing a multi-channel mapping function to encode contours that represent one gait period into a colored template [91]. The above two types of approaches for human identification use several algorithms and parameters. Feature extraction and classification are performed based on the models to predict whether the subject belongs in the database. The author in [17] performs a classification using NN for template score. Lishani et al. [46] propose a classification scheme to test gait recognition criteria using KNN classifier. For clothing and carrying conditions for different view angles more works are also mentioned on the use of various classifiers to test gait recognition rate, according to the literature studies. In this work, the performance of various classifier models for gait pattern classification using the CASIA-B dataset is shown in a table and performance comparison with existing works is shown graphically.

2.3 Overview of the system

The overview of gait recognition technique is given in Figure 2.1. There are five basic steps in the proposed gait pattern classification for the human identification approach.

Figure 2.1: Overview of System

Step 1: Extract frames from the gait video sequence followed by object detection to segment foreground and background. An open-CV background subtraction technique is applied with a median filter for de-noising post-processing and normalization, to obtain silhouette images.
Step 2: GEI gait template is generated by averaging the silhouettes over a gait cycle.
Step 3: Feature extractions from GEI using HoG (Histogram of Oriented Gradients).
Step4 : Perform dimension reduction using LDA (Linear Discriminant Analysis) to discard the dimensions that have a negative impact.
Step 5: Finally the extracted feature vectors are fit into different classifiers for performance evaluation and analysis.

2.3.1 Gait Energy Image

In this work, we have considered the Gait Energy Image (GEI) feature, an Enhanced spatial-temporal variable developed by Han.et.al[30]. It is a technique used in human identification, which is based on analyzing the energy patterns generated by a person's walking motion. The GEI is created by averaging the silhouettes of a person as they walk in front of a camera from different angles. The result is a single image that represents the person's walking energy, which captures the unique features of their gait.

This image is then used as a template for identifying the person in subsequent video frames. The gait energy features extracted from the GEI are matched against the features extracted from the video frames, and if the features match, the person is identified. The advantage of using gait energy for identification is that it can be used at a distance, without the need for physical contact or explicit cooperation from the subject. It is also relatively resistant to changes in clothing, accessories, and other physical attributes that can affect other forms of biometric identification. The simple Silhouette and GEI presentation are shown in Figure 2.2 and Figure 2.3 The Gait Energy Image is computed by averaging the normalized silhouettes over a

Figure 2.2: Gait cycle image from silhouette dataset

number of video frames and given by :

$$GEI(x,y) = \frac{1}{K}\sum_{i=1}^{K} B_i(x,y) \tag{2.1}$$

Figure 2.3: Gait Energy Image(GEI)

Where K is the total number of frames in a gait cycle of a silhouette sequence, x, and y are image coordinates in binary silhouette image frame B_i of gait sequence, and n is frame number [71].

2.3.2 Feature Extraction

Feature extraction is a crucial step in computer vision that involves identifying and extracting key pieces of information from an image or video data. These features represent and analyze the image data, enabling tasks like object detection, recognition, and classification. In the context of using gait templates for human identification, feature extraction entails analyzing patterns generated by a person's walking motion Various types of features can be extracted from an image, depending on the specific application and image data characteristics. Common feature extraction methods include edge detection using Sobel or Canny, corner detection using Harris corner detection and SIFT (Scale-Invariant Feature Transform), and even deep learning-based features. In this work, HOG (Histogram of Oriented Gradients) is employed as a feature extraction technique on the gait template to extract edge information. In this work, HOG (Histogram of Oriented Gradients) feature extraction technique is applied on the gait template to extract the edge information. Recent research has shown that using HOG in appearance-based methods enhances gait recognition efficiency [34, 55]. HOG (Histogram of Oriented Gradients) is a feature extraction technique used in computer vision for object detection and recognition. It involves computing the local gradient orientations of an image and creating a histogram of these orientations in a certain window or block of the image. HOG features can be used to represent the shape or texture of objects in an image.

The process of feature extraction using HOG involves the following steps:

Image Preprocessing: The input image is preprocessed to remove noise and enhance edges using techniques such as Gaussian smoothing or median filtering.

Gradient Computation: The gradient magnitude and orientation are computed at each pixel in the image using Sobel or other gradient filters.

Histogram Calculation: The image is divided into small cells, typically 8x8 pixels, and the gradient orientations within each cell are accumulated into a histogram of orientation bins. The histogram bins are often weighted by the gradient magnitude to give more importance to stronger edges.

Block Normalization: The histograms of adjacent cells are combined to form a block of histograms, typically 2x2 or 3x3 cells. The block of histograms is then normalized to account for illumination changes and contrast variations. There are different normalization methods, including L1-norm and L2-norm normalization.

Feature Vector Extraction: The final step is to extract a feature vector from the normalized blocks. The feature vector is typically a one-dimensional array of concatenated histogram values.

The resulting feature vector can be used for various computer vision tasks such as object detection, recognition, and classification. HOG features have been shown to be effective in detecting and recognizing human faces, body parts, and other objects.

In this work, the HOG feature extraction method is used for local feature representations, and it is applied to a pre-processed Gait Energy Image (GEI) of size (210 x 70) to extract features for gait recognition. In this case, the GEI image is divided into 16 x 16 cells, and a histogram of oriented gradients is computed for each cell. These histograms are then concatenated across blocks of cells to form a feature vector for each block. In this specific task, the GEI image is divided into 175 blocks of size 16 x 16, and each block has a feature vector of size 36 x 1, resulting in a total of 6,300 features for the entire image. These features capture the local motion features of the gait pattern and is used for person identification.

2.3.3 Feature selection

Feature selection for gait recognition typically involves choosing a subset of gait features that are most useful for identifying individuals. The extracted features may contain irrelevant or redundant features, which can affect learning efficiency. There are several methods for selecting the most informative gait features for identification purposes. One popular approach is to use machine learning algorithms such as principal component analysis (PCA) or linear discriminant analysis (LDA) to reduce the dimensionality of the feature space and identify the most important features. In this work, we have applied Linear Discriminant Analysis(LDA) a linear dimension reduction to reduce the irrelevant features.

Linear discriminant analysis (LDA) is a widely used method for feature selection and dimensionality reduction in gait recognition. LDA is a supervised learning technique that seeks to find a linear combination of features that maximally separates the classes (i.e., individuals) in the data. After applying the Histogram of Oriented Gradients (HOG) to the gait data, we obtained 6,300 features for each subject across various covariate conditions. This accumulation of features resulted in a total of (124 * 5) 620 x 6,300 features for each probe angle, considering the inclusion of 124 subjects. This constitutes a very high-dimensional feature space, which can lead to overfitting and decreased classification performance if all of the features are used. Therefore, we propose applying Linear Discriminant Analysis (LDA) to select the k most discriminant features and reduce the higher-dimensional feature space. Our experiments have shown that this approach yields good classification performance across all tested classifiers.

2.3.4 Classification

Various studies that have explored the use of machine learning algorithms for gait recognition employed different classifiers such as SVM, KNN, minimum distance classifier, random forest, and nearest centroid classifier, to identify individuals based on their gait patterns. Nandy et al. [56] pointed out the effect of clothing covariates using GEnI templates with classifier combinations of SVM, KNN and minimum distance classifier to assess the classification accuracy. Alotiabi et al. [2] proposed the first classification technique based on CNN with a high precision of 92% using GEI as input and suggested identification of cross-view variance gait using deep CNN architecture. An experimental gait recognition based on deep CNN carried out by

Wu et al. [97] with cross-view and cross-walking could achieve an average accuracy of 94% on CASIA and 98% OUISIR, respectively.

In this task, we have taken four classifier models for performance analyses of individual gait identification. It is also seen from the study that most of the classifiers employed for shape-based gait identification are traditional classifiers such as SVM, KNN, and random forest with good accuracy. However, our experiment mainly focuses on analyzing gait identification performance covering appearance change due to carrying and clothing covariates. So, we have chosen the nearest centroid classifier along with the existing classifier. The nearest centroid classifier is also a machine learning classifier that has similar KNN classifier working principles. Here are some listed works reported from the literature. In a recent study, SVM classifier for gait recognition for a wearable sensor dataset with a good average recognition rate of 96.5% was achieved [91]. It was observed from the study that the SVM machine learning is derived from the study of support vector network [21]. Gait recognition using PL animation with machine learning strategies for analyzing and classifying gait patterns, and a maximum testing accuracy was achieved using SVM classifier [20]. Much work has been carried out using SVM classifier, but due to limitations in function parameter usage a new direction to develop and adopt a generalized classifier is observed in the studies. A random forests classifier is an ensemble learning method for classification, regression, and other tasks that works by training a large number of decision trees and then extracting the mean/average predictor (regression) or mode of the classes (classification) [9]. The accuracy of KNN classification is calculated by the use of similarity measures and the value of k. Until deployment, several k variants were tested to obtain the highest accuracy. They are as follows: 1, 3, 5, 7, and for application, k = 3 was used since it has the greatest precision.

2.4 Experimental analysis and result

The experiment was performed using the CASIA-B dataset, which included 50 subjects with gait data obtained from six different angles: 0, 18, 36, 54, 72, and 90. The gait sequences of each person are collected in ten perspectives, comprising six usual walking sequences (nm), two sequences of walking and carrying a bag (bg), and two walking sequences of wearing a coat (cl). Figure 2.4shows some images of CASIA-B dataset for the experiment. In this work, we took four types of machine learning classifiers support vector machine (SVM), random forest, KNN, and nearest centroid to find out the correct classification rate for different clotting covariates.

(a) Normal walk (b) Normal walk with a jacket (c) Normal walk carrying a bag

Figure 2.4: CASIA B Dataset 90 degree angle

The following are three experiments performed on the CASIA-B dataset:

1. In the first experiment, we took a normal walk sequence from five different views in the training dataset to test its proper classification rates.
2. In the second experiment, we use a test dataset with five different views of a normal walk sequence to test correct classification rates using coat covariate.
3. In the third experiment, we use a test dataset with five different views of a normal walk sequence to test correct classification rates using bag covariate.

Table 2.1: Correct classification rate for normal walk covariate

classifiers	0^0	18^0	36^0	54^0	72^0	90^0
SVM	97.96	100	100	100	100	100
KNN	97.96	100	100	100	100	100
RF	91.84	100	100	94	94	96
NC	97.96	100	100	100	100	100

Table-2.1, Table-2.2, and Table-2.3 depict the CCR(correct classification rate) for different classifiers with three different covariate conditions. The appearance change due to clothing condition is the most complex covariate while conducting the experiment, and it also reduces the Correct classification score. For the experiment the notations:C_Δ indicates the average of the

Table 2.2: Correct classification rate for clothing covariate

Classifiers	0^0	18^0	36^0	54^0	72^0	90^0
SVM	42.86	54	52	48.86	50	32
KNN	48.98	50	50	42.86	52	42
RF	20.41	34	30	32.65	36	22
NC	46.94	52	50	46.94	50	34

Table 2.3: Correct classification rate for baggage covariate

Classifiers	0^0	18^0	36^0	54^0	72^0	90^0
SVM	48,98	62	58	63.27	62	64
KNN	55.10	64	70	63.27	64	64
RF	42.86	62	50	44.90	36	42
NC	55.10	66	68	67.35	68	66

Table 2.4: Average correct classification of all the covariate conditions

Classifiers	NM (C_Δ^{NM})	CL(C_Δ^{CL})	BG(C_Δ^{BG})	Average
SVM	99.66	46.62	59.70	68.66
KNN	99.66	47.64	63.34	70.21
RF	95.97	29.84	46.28	57.26
NC	99.66	46.64	65.07	70.46

Correct classification rate of the different classifiers as shown in 2.1, Table-2.2 and Table-2.3

$$C_\Delta^{NM} = \frac{1}{6}\sum_{n=0}^{5} C_n^{NM} \tag{2.2}$$

Where C_n^{NM} is the CCR of each angle for Normal Walk C_Δ^{NM} can be used to correct the classification rate when the covariate is normal to walk C_Δ^{BG} C_Δ^{CL} indicate the covariate with Bag and clothing define similarly for six different angles: 0, 18, 36, 54, 72 and 90. shown in Table-2.4 the last column shows the average performances of all three experiments with different classifiers from that we can conclude that the nearest centroid classifier can perform better than all the other classifiers in clothing covariate and baggage covariates. The proposed approach is then further compared to other existing approaches. The comparison results are shown in Table 2.5. The last column shows the average performances of all covariates which are shown in all three tables and, it is found that our approach with the Nearest centroid(NC) classifier performed better in clothing and baggage covariate changes from the other existing methods.

Table 2.5: Comparison with other methods

METHOD	NM	CL	BG	Average
GEI [30]	100.0	22.2	53.2	58.5
CGI [90]	88.1	43.0	43.7	58.3
Baseline TM [104]	97.6	32.7	52.0	60.8
Our method(with NC)	99.66	46.64	65.07	70.46

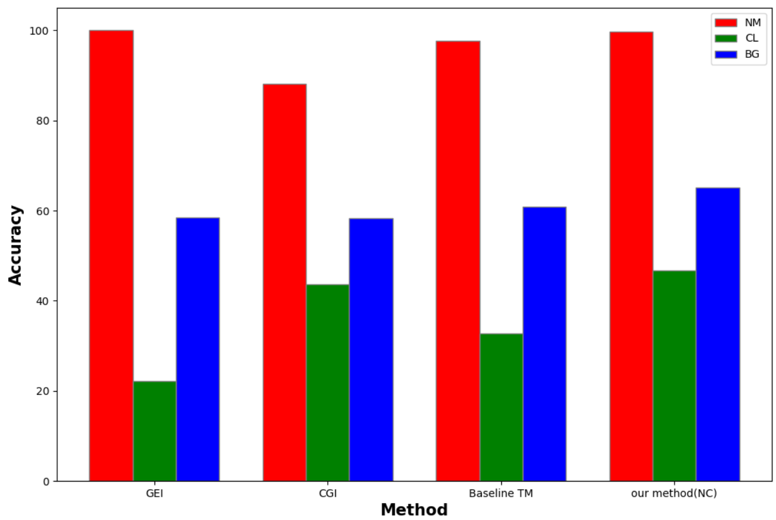

Figure 2.5: Comparison graph with other methods

2.5 Summary

Gait-based human identification is an efficient authentication and verification system as reported by several authors. In this work, we attempted to investigate the classification of human walking patterns using a different classifier in order to account for the covariates that affect the identification efficiency. So, to address the covariates factor for human gait analysis, we have considered local motion features instead of global features in the existing method. The method with local motion information is of fewer features but effective for small variations, which is a better method to address appearance change due to clothing and carrying covariates. They propose a method that outperforms the existing shape/appearance-based methods of global features with more discriminant functions. From the experiment result, we thus conclude that the nearest centroid (NC) performs better when all covariates are considered. In the future, work can be extended for other covariate factors like appearance changes from complex environments.

Overall, the study suggests that traditional classifiers such as SVM, KNN, and random forest can be effective for gait identification, and the nearest centroid classifier can also be a useful addition to analyze gait identification performance under appearance changes due to carrying and clothing covariates.

In the next chapter 2, model-free gait analysis is adapted for gait in surface covariate, with a small dataset created from varying surface conditions. The objective of the dataset is to evaluate the human walking pattern on varying surfaces with natural settings.