GAIT FEATURES-BASED HUMAN RECOGNITION APPROACHES
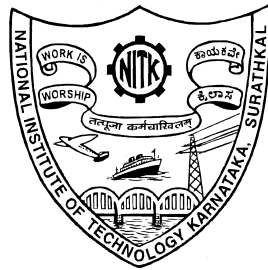
Thesis

Submitted in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

ANUSHA R

DEPARTMENT OF INFORMATION TECHNOLOGY

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
SURATHKAL, MANGALORE - 575025

DECEMBER, 2021

Declaration

I hereby *declare* that the Research Thesis entitled "Gait Features-based Human Recognition Approaches" which is being submitted to the National Institute of Technology Karnataka, Surathkal in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy in Information Technology is a *bonafide report of the research work carried out by me*. The material contained in this thesis has not been submitted to any University or Institution for the award of any degree.

<div align="right">

Anusha R
Register No.: 165041IT16FV01
Department of Information Technology

</div>

Place: NITK Surathkal

Date: 20-12-2021

Certificate

This is to *certify* that the Research Thesis entitled "Gait Features-based Human Recognition Approaches" submitted by Anusha R (Register Number: 165041IT16FV01) as the record of the research work carried out by her, is *accepted as the Research Thesis submission* in partial fulfilment of the requirements for the award of degree of Doctor of Philosophy.

Dr. Jaidhar C D
Research Guide
Associate Professor
Department of Information Technology
NITK Surathkal - 575025

Chairman - DRPC
(Signature with Date and Seal)

Acknowledgements

Dedicated to

My Parents:
Mrs. Manjula G and Mr. Ramesh S

Abstract

Research into biometrics is an on-going and open research challenge to achieve robust human identification in a visual surveillance environment. Compared to other biometrics, gait has gained considerable attention in current years due to the unique benefits that other biometrics may not offer. Most significantly, it can be used with video feeds captured at a distance without alerting the subject and with low-resolution video. Interest in gait has increased appreciably because of the promising recognition results achieved from research in this area under controlled environments. Recent research is focused more on improving the recognition rate in realistic environments, where it is necessary to address the effects of changes in view, resolution, and fluctuation of gait patterns, due to carrying goods, walking speed variations, different footwear or clothes.

The influence of various gait challenging conditions makes the real-world gait recognition system struggle for better performance. The development of appropriate features by using the information source available is the only solution to deal with these challenges. In this work, solutions that can enhance the performance of a gait recognition system are proposed to assist security applications. One of the significant challenges of the appearance-based gait recognition system is to augment its performance by using a distinctive low-dimensional feature vector. Therefore, this study presents the low-dimensional feature vector that is capable of capturing the spatial, gradient, and texture information. These features are obtained by the computation of Histogram of Oriented Gradients (HOG), followed by the sum variance Haralick texture descriptor from nine cells of Gait Gradient Magnitude Image (GGMI). The improved recognition rate is achieved on the five publicly available gait datasets.

The clothing variance is one of the most common covariate influences which can influence the performance of the gait recognition approach in real-world scenarios. This study presents a gait recognition approach proficient in choosing information characteristics for individual identification under different clothing conditions. The proposed approach deals with the feature extraction technique by introducing a binary descriptor called Modified Local Optimal Oriented Pattern (MLOOP). Furthermore, the proposed approach is assessed on the OU-ISIR B and CASIA B gait datasets, and it achieves improvement in recognition performance over other binary descriptors.

One of the difficulties of the appearance-based gait approach is to enhance the performance of frontal gait recognition, as it carries less spatial and temporal data when compared with other view variations. As a result, to increase the performance of the

frontal gait recognition, this study presents a method that uses a two-step procedure; the Hierarchical Centroid Shape Descriptor (HCSD) and the similarity measurement. One more method is proposed, which uses the contour image and contour vertices to extract three discriminative feature vectors from the Gait Energy Image (GEI). Thus, it captures the spatial dynamics of frontal gait efficiently to improve gait recognition performance. These two methods are assessed on the broadly used CASIA A, CASIA B, and CMU MoBo gait datasets. The experimental outcomes show that the proposed methods yield the promising results and outperform certain state-of-the-art methods in terms of recognition accuracy.

In this work, effective approaches are proposed to remove the effect of walking speed in a gait detection system. The first approach uses the Region of Interest (ROI) extracted from GEI to classify a probe sample into a gallery sample. The Mutual Information (MI) obtained from a probe and gallery sample, followed by their classification, efficiently improves the gait recognition performance. The proposed method shows an improved performance for two datasets when compared to other methods reported in this thesis. The next method identifies the most similar parts of the probe and each gallery sample independently and uses these parts to obtain a similarity/dissimilarity measure through three metrics. This method represents the spatial dynamics of GEI efficiently to improve gait recognition performance. Further, the proposed methods are evaluated on CASIA C and OU-ISIR A gait datasets. Experimental results demonstrate the capability of the proposed approaches in comparison with the existing gait recognition methods. This approach shows an increased performance for two datasets when compared to other methods reported in this thesis.

Finally, the possibility of identifying individuals by using their running video is mostly unexplored. Hence, this study proposes a method that extends the feature-based approach to recognize people by the way they run. Here, the statistical, texture-based, and area-based features are extracted from each image of a gait cycle. The experiments are carried out on the KTH and Weizmann dataset.

The several feature extraction algorithms proposed in this thesis are focused predominantly on appearance-based methods because of their exceptional performance and simplicity. Overall, the aim of this research work is to increase the gait recognition system performance by contributing to areas such as dimensionality reduction of a feature vector, identification of an individual is attempted by using running patterns, to accomplish frontal gait recognition, speed invariant gait recognition and clothing invariant gait

recognition. The proposed solutions in this work contribute to improving gait recognition performance in various practical scenarios that further enable the adoption of gait recognition into various applications.

Keywords: Biometrics; Feature Extraction; Gait Recognition; Human Identification.

Contents

iv

List of Tables

List of Figures

xi

List of Abbreviations

Abbreviation	Meaning
CCD	Centroid Corner Distance
CCR	Correct Classification Rate
CI	Contour Image
GEI	Gait Energy Image
GF-GEI	Gaussian Filtered - Gait Energy Image
GGMI	Gait Gradient Magnitude Image
GLCM	Grey-Level Co-occurrence Matrix
HCSD	Hierarchical Centroid Shape Descriptor
HOG	Histogram of Oriented Gradients
KNN	K Nearest Neighbors
LBP	Local Binary Pattern
LDP	Local Directional Pattern
LOOP	Local Optimal Oriented Pattern
MF	Membership Function
MI	Mutual Information
MLOOP	Modified Local Optimal Oriented Pattern
NAC	Normalized Auto Corelation
NOC	Non-Overlapping Coordinates
OC	Overlapping Coordinates
PCA	Principle Component Analysis
PCS	Piecewise Coordinate Segment
ROI	Region of Interest
SI	Spatial Image

Chapter 1

INTRODUCTION

1.1 Biometrics

Biometrics is the discipline of establishing the identity of an individual, based on intrinsic physical as well as behavioural qualities associated with an individual (Lumini and Nanni 2017). Human identification has been a dynamic field of research since the requirement for reliable user verification systems has increased, from giving access control in innovatively propelled applications, for example, inter-networking of several common and well-known applications across forensics, security, surveillance, physical and logical access control, border control for a whole nation and so on (Prakash et al. 2015).

Some of the popular biometric traits used in various systems are shown in Figure 1.1. The biometric traits used for human identification can be categorized into two distinct classes:

1. Physiological: These characteristics are derived from the direct measurement of a part of the human body. The typically noticeable of these sorts of measurements to date are hand scan, fingerprint, iris, and face scan.

2. Behavioural: These characteristics are extracted based on an activity performed by a person. They present an indirect measurement of the characteristic of an individual. These measurements include handwriting, signature, and speech patterns.

In addition to the conventional authentication techniques such as pin and passwords, biometrics have now gained an extensive acceptance to provide genuine authentication of the identity for an individual, especially in computer-vision based applications. In biometrics, a unique signature does not exist for a person. Generally, the information obtained from an individual generates a slightly different signature each time. This does not mean that the frameworks are insecure naturally, because very high rates of recognition have been accomplished in recent years. Usually, the process of recognition is done through thresholding and correlation.

Some of the advantages of biometrics are as follows:

1. The biometric characteristics cannot be forgotten or lost, unlike pins and passwords.
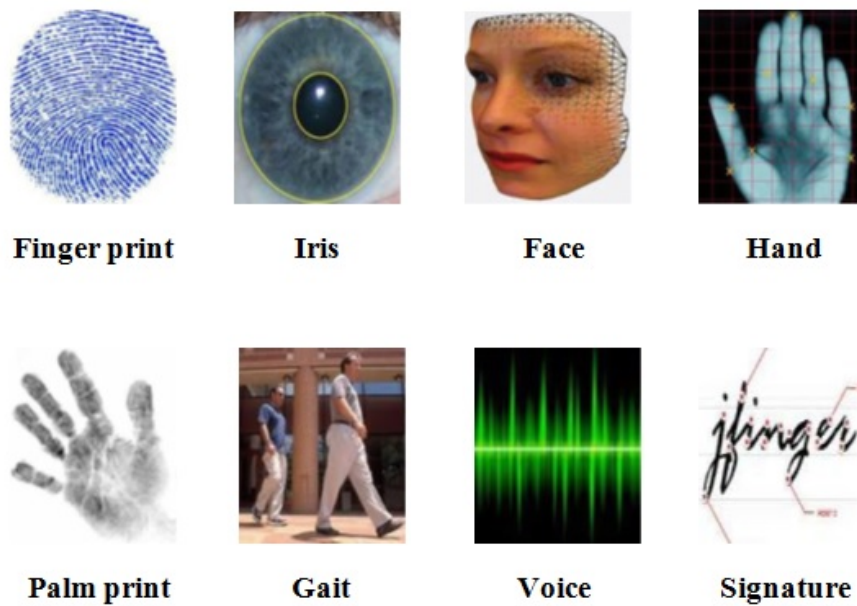
Figure 1.1: Examples of different biometric traits.

2. It is not easy to share, copy, forge or distribute a biometric characteristic.

3. The biometric frameworks also include user convenience by easing the need to plan and recollect passwords.

1.1.1 Desirable Qualities of Biometric Characteristics

Some of the properties of biometric characteristics which are important for the better identification of an individual are described below (Neves et al. 2016):

1. **Universality**: Each individual must have the biometric characteristic.

2. **Uniqueness**: Across individuals, the biometric characteristic must be satisfactorily distinguishable.

3. **Permanence**: The biometric characteristic must not be variant over a period of time.

4. **Measurability**: The biometric characteristics should be obtained from the individuals without causing much inconvenience.

   The other points which are taken into account in terms of application point of view are as follows:

5. **Performance**: In an application, the required recognition accuracy must be achievable with the biometric characteristic.

6. **Acceptability**: The individuals must be willing to present his/her biometric characteristic.

7. **Spoof Resistance**: This indicates the difficulty of using artefact in case of physiological characteristics (for example, fake fingers) and imitation in case of behavioural characteristics.

1.2   Biometric System Operations

A classic biometric-based authentication framework operates by extracting the necessary features and comparing those features with the previously registered biometric samples in the database, to confirm the claimed identity or to recognize the person. The five subsystems of a biometric system are explained as follows:

1. Biometric data capture subsystem: This includes appropriate capture devices or sensors. These sensors are necessary to collect signals of a biometric characteristic from individuals. These captured characteristics are converted into biometric samples such as an iris image, fingerprint image, or voice recording.

2. Signal processing subsystem: This subsystem consists of two stages: (a) improving the quality of the captured samples, and (b) extracting significant discriminatory features from the biometric samples. The extracted features represent the underlying biometric characteristic.

3. Data storage subsystem: In this phase, the extracted biometric features are stored. The features are probably stored along with other non-biometric data related to the individuals such as name, identification number, social security number, etc.

4. Comparison or matching subsystem: In this subsystem, the comparison score is calculated when one or more training samples are compared with the test samples. The decision making subsystem further uses this comparison score. Here, the similarity/dissimilarity between the features extracted from the input sample and the enrolled biometric samples is determined.

5. Decision subsystem: In this subsystem, the comparison score is used to determine the results. The comparison score is used to find out if the test biometric sample and training samples belong to the same subject or not.

1.3   Gait

Gait can be defined as a coordinated, cyclic combination of movements that result in human locomotion. The gait synthesis of an individual is a complex phenomenon because it is associated with the synchronization of the movements of the upper and lower

body. Quite a few months of learning is required to develop this synchronization. That is, it may require up to one year for a normal child for the accomplishment of stable walking, where the different parts of the cerebrum related to the learning process establish coordination between the muscles and nerves, i.e., the sensory organs and motor system (Semwal et al. 2015). The gait comprises the following stages: lifting one leg with the support of another leg on the floor and moving the body forward while swinging the lifted leg until it is in front of the body. The whole body is propelled forward when the lifted leg comes in contact with the floor.

Even though the gait patterns follow the basic bipedal movements for each individual, they usually differ from one individual to another in certain particulars, such as their magnitudes and relative timing. Gait can be considered as a combination of the static human shapes and the cyclic, coordinated nature of the motion. Hence, it is consequently supposed to be unique for each individual, as determined by the entire musculoskeletal structure.

1.4   Gait Recognition

Numerous applications depend on biometric recognition systems for individual identification (Prakash et al. 2015; Neves et al. 2016). Fingerprint, face, and iris are some of the traditional biometric characteristics for human identification. However, the disadvantage of these characteristics in surveillance environments is that it is not easy to capture them from a distance, and they also need the active cooperation of the user. In order to overcome this disadvantage, an extensive research work has been undertaken regarding the soft biometric characteristics, such as weight, height, gender, and age, as they can be obtained from a considerable distance without the requirement of active user cooperation. However, the setback in this regard is that these soft biometric characteristics are not unique enough to distinctly recognize an individual. Even so, gait, which can be used to recognize individuals based on their walking pattern, is a prominent biometric characteristic that identifies an individual from a distance uniquely without bodily contact (Choudhury and Tjahjadi 2016). Hence, gait can be a valuable biometric in surveillance applications.

The gait recognition aims to differentiate human beings based on the characteristics of their locomotion. Examples of these characteristics are step length, stance time, swing time, length of different body parts, cadence, hip angle, knee angle, and so on. Some of these characteristics can be used in the process of identifying individuals based

on their gait. The active selection of these features is significant in the recognition problems.

There has been a tremendous development in the gait recognition approaches in the past two decades. Nevertheless, there are many limitations towards the practical application of gait analysis. Lately, gait detection research has focused more on building a representation of gait that has the capability to obtain significant information under volatile conditions. Moreover, the recognition performance of any gait recognition method depends mostly on the type and quality of the extracted gait features.

1.5 Gait Versus other Biometric Traits

Gait, as a biometric, has many advantages when compared to other biometric traits because of the following reasons (Phillips et al. 2002):

1. Distance recognition: The other traits of biometrics such as face, fingerprints, iris, signature, voice, and hand geometry can be captured mostly at a very close distance from the recording sensor or by physical contact, whereas the gait pattern of an individual can be obtained at a distance far away from the recording sensor.

2. Unobtrusive: This is the most significant aspect of gait as a biometric trait. The gait pattern of people can be obtained without them knowing that they are being analyzed and with no cooperation from the people while gathering information, unlike other biometrics such as retina scans or fingerprints.

3. Reduced detail: Unlike other biometrics such as fingerprint, iris, face recognition, identification of an individual using gait does not require high-quality images. The other biometrics can be affected easily by low-resolution images, which is not the case with gait recognition.

4. Difficult to conceal: It is difficult to fake or hide the gait of an individual, and by doing so, the individual might perhaps appear more suspicious, whereas the other biometric traits such as fingerprint and face can be easily hidden.

The human gait recognition has drawn considerable attention in recent years because of all these potential advantages.

Different biometric systems such as fingerprint, face, and iris recognition are capable of presenting a robust performance in the real-world applications. In recent years, the overall effort has been made towards the improvement and expansion of the individ-

5
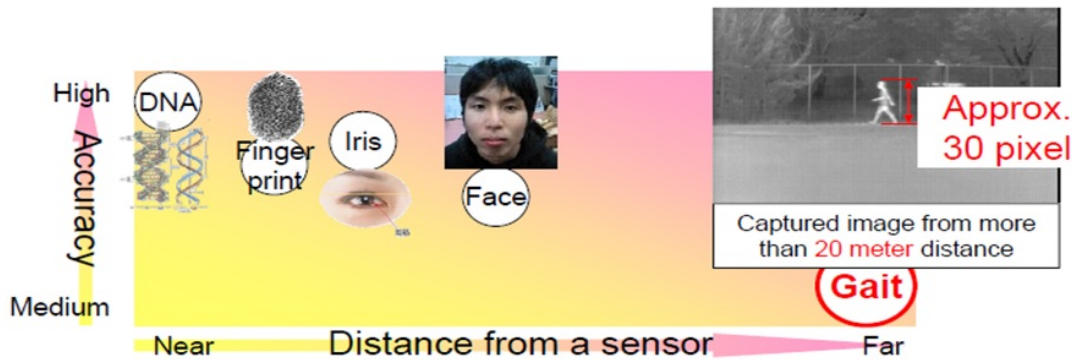
Figure 1.2: Various biometrics: distance versus accuracy (Wang et al. 2008).

ual recognition frameworks. From a surveillance point of view, gait recognition is an appealing system because of its capability of recognizing individuals from a distance by examining their manner of walking, as shown in Figure 1.2.

1.6 Approaches in Gait

There are three different types of approaches in gait recognition based on the method used for capturing gait data. They are machine vision-based, floor sensor-based, and wearable sensor-based gait recognition methods. These are explained as follows:

Machine vision-based: In this approach, usually, the biometric system contains a set of digital cameras with appropriate optics for obtaining the gait data. Here, many techniques, such as background segmentation, pre-processing, followed by feature extraction and classification, are used to identify an individual. Most of the existing gait identification methods are machine vision-based. The major benefit of this method is that more or less all gait features can be obtained from the captured video such as stride, cadence, step length, area, and distance between different parts of the body. Moreover, the co-operation of the people is not needed in these methods.

Floor sensor-based: In this approach, the sensors are located on a mat along the floor. This makes it suitable for controlling access to many buildings, offices, homes, and other places where security is essential. When an individual walks on the mat, the force applied to the ground is measured, which is called as ground reaction force. Therefore, access control applications like the front part of the doors in a building can be used to deploy this method. This method can provide location information as well, along with identity information. A very few gait features can be extracted by using this method, such as cadence, stride length, maximum time taken for heel strike, and so on.

6
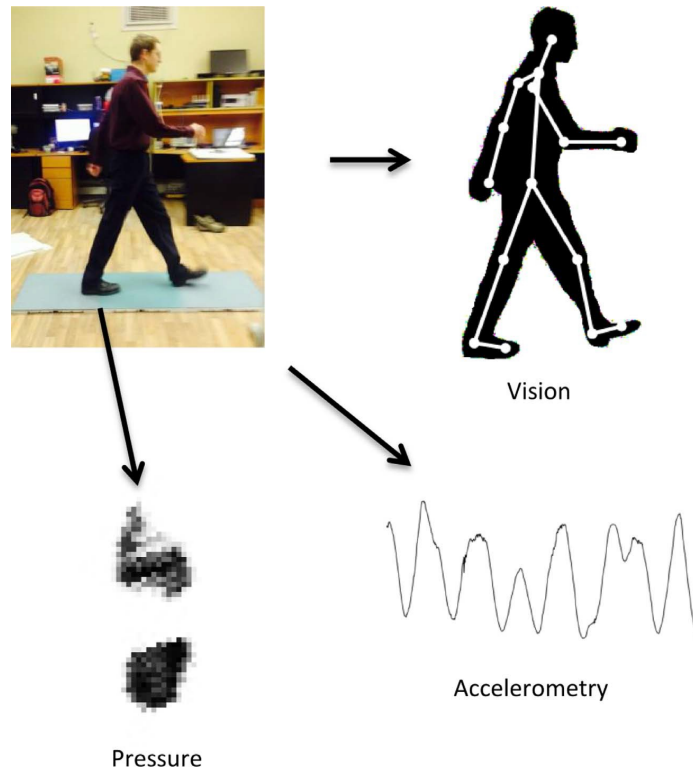
Figure 1.3: Examples of sensor modalities used for gait identification (Connor and Ross 2018).

Wearable sensor-based: This method is the most recent method among the other methods mentioned earlier. This method is based on wearing sensors that record motion on the body of an individual at different places, such as on the hands, legs, waist, foot, and so on. These sensors are used to measure several aspects, such as acceleration, rotation, force when walking, etc.

The examples of different sensor modalities used for gait recognition are shown in Figure 1.3. The significant advantage of gait as a biometric is unobtrusiveness, whereas the wearable sensor-based method does not satisfy this property as it requires the active cooperation of the people to collect the gait data. Furthermore, the cost of the sensors used to measure ground reaction force in the floor sensor-based approach is immense. The machine vision-based approach does not have these setbacks, and hence, it has been chosen in this study for gait recognition.

1.7 A Generic Gait Recognition System

A typical gait detection system consists of the fundamental tasks, as shown in Figure 1.4. To identify the spatial and temporal behaviour of an individual, the machine vision-based gait recognition techniques are usually initiated by the extraction of human
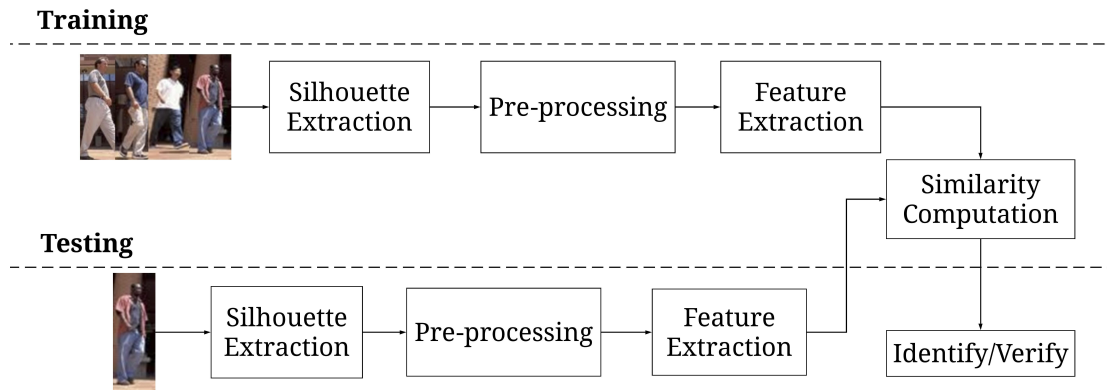
Figure 1.4: The gait recognition framework.

gait silhouette from video footage obtained from the camera sensor. These gait silhouettes are pre-processed with appropriate normalization and alignment. This is followed by the usage of various image processing, computer vision, and other techniques to extract discriminative features. The next step is to store the generated features to form a database in the training process. During the testing process, the discriminative features from the test sample are obtained, and are compared with the dictionary formed for recognition or to validate a person's identity.

Among all fundamental tasks of a gait recognition system, the computation of gait features that are robust and discriminative is the most focussed area in recent research since it is the important element on which all other tasks depend.

1.7.1 Gait Cycle

A gait cycle can be defined as the sequence of movements or events carried out during locomotion in which one foot contacts the ground to the time when that same foot contacts the ground again. A single gait cycle is also called as a stride. It consists of two phases. They are as follows:

1. Stance phase: This is the phase during which the reference foot is in contact with the ground. The reference foot undergoes five movements in this phase. It constitutes about 60% of the gait cycle.

2. Swing phase: This is the phase during which the reference foot is not in contact with the ground. The reference foot undergoes three movements in this phase. It constitutes about 40% of the gait cycle.

Figure 1.5 outlines the phases of a gait cycle. A significant aspect of the process

8

of gait analysis is the detection of the gait cycle. It involves the partitioning of gait sequence/movements into gait cycles that represent a complete walking period. These extracted gait cycles are further used for the derivation of distinctive gait features for the purpose of identification.
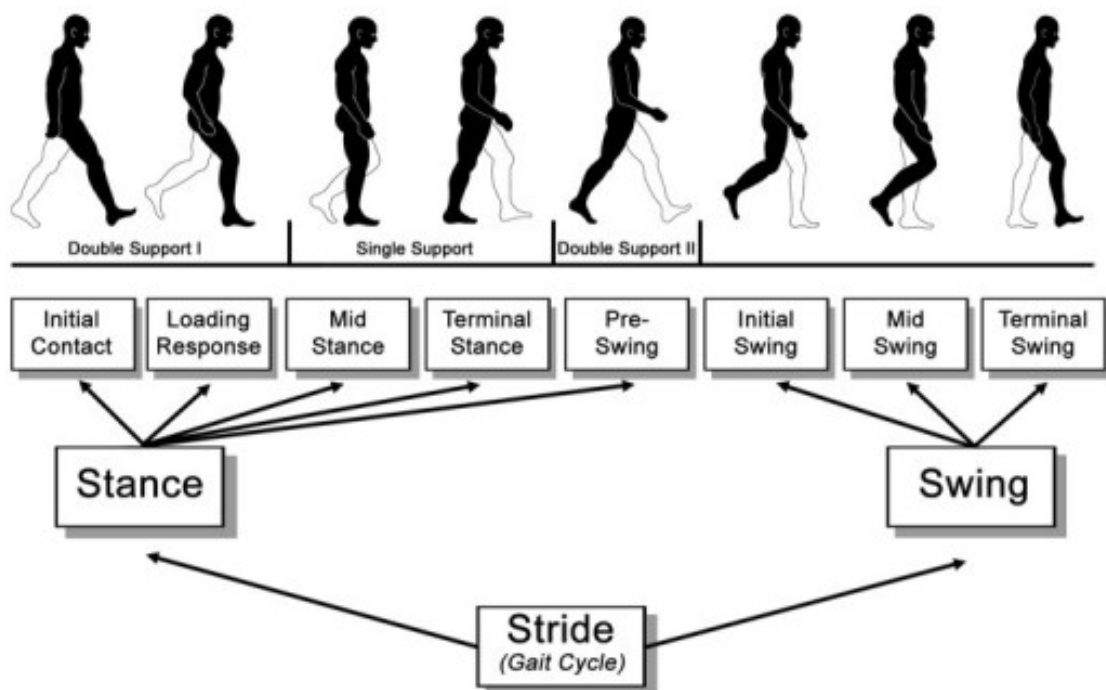
Figure 1.5: Phases of a gait cycle (Stöckel et al. 2015).

1.8 Impacting Factors in Gait Modality

In real-world situations, there might be considerable differences between the test and training samples of an individual, as they might be collected in different scenarios. The variations include clothing, load carriage, walking speed, walking environment, and footwear, as shown in Figure 1.6. These can be collectively called as covariates. These are external to the walking subject and may impose a constraint on the walking pattern of an individual or distort gait parameters by compelling an individual to walk in a certain way. The subject might also undergo physical changes because of the conditions such as pregnancy, injury, or advanced age, which generally alters the gait generation machinery.
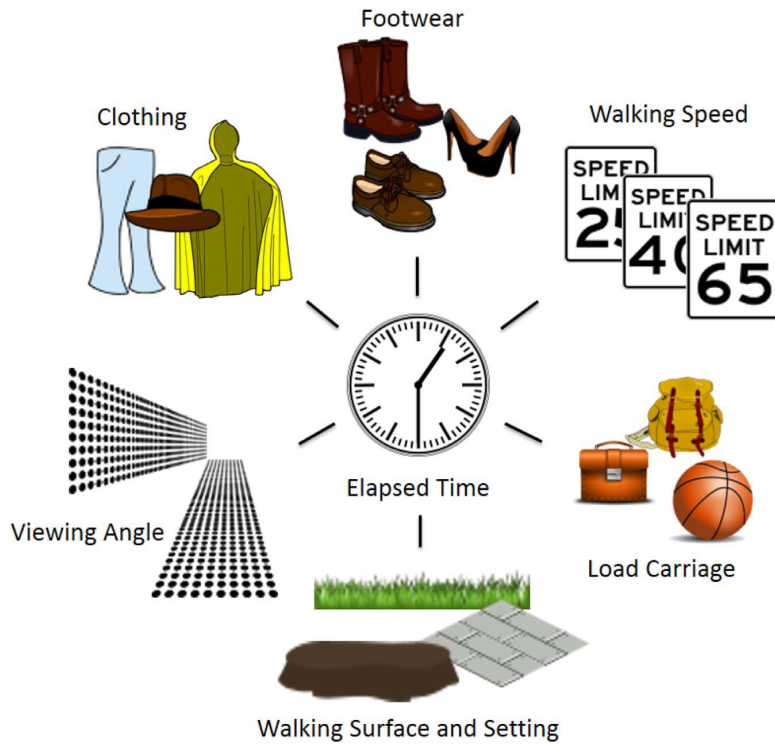
Figure 1.6: Various covariates that influence the performance of a gait recognition system (Connor and Ross 2018).

1.9 Terminology

In this section, an explanation of the different terminologies used in this thesis is presented for a better understanding of the contents.

Subjects: The term subject/subjects refer to the person/persons used for testing and training of a gait recognition system.

Gait sequence: The term gait sequence refers to the walking pattern/movements of people.

Gallery: Gallery is a collection of subjects that are enrolled in the database. The term training dataset and gallery dataset are interchangeably used in this thesis.

Probe: Probe is a collection of subjects that needs to be identified or verified. The term test dataset and probe dataset are interchangeably used in this thesis.

1.10 Motivation

Recognition of an individual is one of the areas of active research as the requirement for user authentication systems which are reliable has increased. These include, giving ac-

cess control in technically complex applications involving inter-networking among the widespread and common applications across forensics, security, surveillance, logical and physical access control and border control for a whole nation. Along with the conventional authentication traits such as pins and passwords, the field of biometrics has now received an extensive acceptance to give legitimate confirmation about the identity of an individual, predominantly in computer-vision based applications.

Gait recognition can be done by the data obtained from the surveillance system, and it has fascinated a lot of researchers towards investing their time in proposing various techniques to address the problems in gait recognition. Currently, the only perceivable biometric trait from a great distance is gait. The gait biometric framework can be applied to the parking lot, open imperative place, pedestrian crossing, bank, market, airports, and in many other infrastructures where security is indispensable.

Gait as a biometric has various advantages, which make it an appealing idea as a method of identification. A time frame is provided to the user to respond before the suspect turns out to be a possible threat by the ability to detect a probable threat from a distance. Nowadays, the surveillance cameras are of reasonably low cost and are installed in most of the locations or buildings where security presence is required. Hence the video footage of suspects is available readily, and it just needs to be checked against that of the suspect. Another motivation is that the reduction of the price of data storage devices and high-speed memory, along with the increase in processor power, has contributed to the increase in the applicability and availability of video processing and computer vision methods.

1.11 Major Contributions of the Thesis

The contributions made in this thesis are as follows:

1. It proposes a feature vector based on the combination of HOG and sum variance Haralick texture descriptor to increase the performance of a gait recognition system. This feature vector is more reliable and better characterizes the spatial variations of gait. The length of the feature vector is reduced from 3780×1 to 63×1, thus enabling the reduction of the computational complexity of the gait recognition system.

2. It presents a robust method for the appearance-based human gait identification system. This method requires statistical features that depict the shape of the contour, which, in turn, is valuable in a gait detection system.

3. It proposes a texture descriptor called MLOOP. The two different vectors called a histogram, and horizontal width, extracted from the MLOOP descriptor, are used as features in clothing invariant gait recognition system.

4. It proposes two methods that are more reliable and also better characterize the spatial variations of a frontal gait.

5. It presents two effective methods to remove the effect of walking speed in a gait detection system, and

6. It proposes a simple approach for the recognition of individuals by using their running gait patterns.

The proposed feature extraction methods in this thesis use the concepts such as HOG, corner detection algorithm, Haralick texture descriptors, local optimal oriented patterns, HCSD, MI, relative entropy, information set, and so on, which are extensively used in many human recognition applications. Different approaches have been proposed in this study to increase the recognition accuracy of normal gait recognition, clothing invariant gait recognition, frontal gait recognition, speed invariant gait recognition, running individual recognition, and jogging individual recognition. Each of the proposed methods is tested using two to three widely used gait databases. The experimental results demonstrate that the proposed approaches increase the gait recognition performance in comparison to some of the existing methods in the literature.

GEI is converted into different templates as per the requirement of various proposed approaches to facilitate the extraction of non-redundant, distinct features. With respect to walking speed and clothing variations of a subject, many pre-processing techniques are used to remove the regions which are most affected by these changes. The pre-processing steps are used to extract the less affected regions from the GEIs. These regions are further used for feature extraction.

An effort has been made to increase the inter-class differences in the proposed methods, even when different subjects shape information is much similar. An attempt has also been made to reduce the number of features used for gait recognition in each of the proposed methods. This also increases the efficiency of the classifier used for recognition.

1.12 Organization of the Thesis

The remaining chapters of this thesis are organized as follows:

Chapter 2 provides an insight into the existing gait feature extraction approaches and provides a detailed review of the existing appearance-based feature extraction methods. It describes the performance metrics, classification method, and databases used in this thesis. Based on the literature results, it defines the problem statement and research objectives.

Chapter 3 describes the implementation of the gait recognition framework proposed in this thesis to improve the performance of a gait recognition system using low dimensional distinctive feature vectors. All the required pre-processing, feature extraction, classification tasks are presented in this chapter.

Chapter 4 explores the influence of clothing variations on gait silhouettes and GEI. It proposes a gait recognition approach proficient in choosing information characteristics for individual identification under different clothing conditions. As a part of this, a detailed explanation of the proposed MLOOP descriptor and corresponding feature vectors are provided. The experimental results and performance of the proposed clothing invariant gait recognition method are discussed here.

Chapter 5 illustrates two effective methods that are capable of identifying individuals in frontal view gait sequences. The proposed methods that improve the gait recognition performance by the usage of the spatial dynamics of GEI efficiently are described in detail.

Chapter 6 outlines the two appearance-based feature extraction approaches proposed to minimize the effect of walking speed in a gait recognition system. It also explores the influence of walking speed on the shape of gait silhouettes.

Chapter 7 illustrates a simple method that extends the feature-based approach to recognize people by using running gait patterns.

Chapter 8 concludes the thesis with a summary of the research as well as providing directions for future work.

1.13   Summary

This chapter introduces biometrics, gait, and gait recognition system. It gives insight into the characteristics and advantages of the gait recognition system. It also gives an overview of the gait recognition framework, covariates that influence the gait of an individual, and different approaches used for gait detection. This chapter also presents an outline of the contributions of the research work done and structure of the thesis.

Chapter 2

LITERATURE REVIEW

In this chapter, a brief description of the contributions made towards gait recognition in the recent past is presented.

2.1 Feature Extraction

The existing gait recognition approaches can be generally divided into model-based methods and appearance-based methods based on the technique used for the extraction of discriminative gait features.

Model-based methods: The motion of the human body from the gait sequences is modeled in the model-based methods. These methods show kinetics of a person's joints to quantify gait specifications such as directions, hip, knee and ankle movements, hand movement, and so on. Various techniques have been used to develop a body or skeletal model from the gait videos. Here, the aim is to locate important anatomical locations efficiently and accurately. These methods are usually computationally expensive since they require modeling and tracking of the subject's body. Additionally, they also require high-resolution images.

Appearance-based methods: On the other hand, the appearance-based methods deal with performing the operations directly on gait images without using any explicit model. These methods are usually preferred over the model-based methods because of their lower computational complexity. Some of the features extracted using this method are listed in Table 2.1.

Both approaches discussed above try to address the key challenges in gait recognition. These challenges include view angle variations, walking speed variations, different clothing conditions, different carrying conditions, and so on.

The recognition performance of the model-based methods is comparatively less, as they are less susceptible to the above mentioned challenges because of the inaccurate fitting of the model which is sensitive to the image quality. However, by using the static and dynamic information, the better performance is achieved by the appearance-based methods. Moreover, the recognition performance is sensitive to the above mentioned challenges.

Hence, this thesis explores the appearance-based methods to propose the innova-

Table 2.1: Various appearance-based gait recognition feature sets.

Reference	Feature set
Johnson and Bobick (2001)	Body parts length and height
Collins et al. (2002)	Silhouette key frames
Kale et al. (2002)	Silhouette row widths
Foster et al. (2003)	Area of partial silhouettes
Mowbray and Nixon (2003)	Fourier descriptors of silhouette outline
Man and Bhanu (2006)	Average silhouette or GEI
Ioannidis et al. (2007)	Depth silhouettes
Wang et al. (2010)	Chrono gait image
Sivapalan et al. (2011)	Gait energy volumes
Hofmann and Rigoll (2012)	Gradient histogram energy image
Roy et al. (2012b)	Pose energy image
El-Alfy et al. (2014)	Histogram of normal vectors on silhouette outline
Nandy et al. (2017)	Gait entropy image

tive solutions to gait recognition problems using several signal processing and machine learning methods.

2.2 Literature Review on the Appearence-based Methods

To obtain the information about the existing research on human gait detection, the subsequent literature review is provided with insights into diverse human identification techniques, that use statistical shape analysis.

Early on in vision-based gait recognition research, the walker's pixels are segmented from the background pixels to obtain the gait silhouette.

Murase and Sakai (1996) have extracted gait silhouettes and used principal components to extract sixteen features. They have experimented with six subjects and obtained 100% recognition accuracy. BenAbdelkader et al. (2001) have computed the pairwise correlation between a silhouette and all other silhouettes in the same gait cycle called the similarity matrix of gait silhouette. They have considered 7 subjects and achieved 93% recognition accuracy. Collins et al. (2002) have extracted key frames from the sequence of gait images that signify certain poses in a gait cycle. This method is evaluated on several gait databases. Liu and Sarkar (2006) have used the normalized sequences of gait silhouettes from the hidden Markov model, which has been computed from different subject's sequences. They have defined the linear discriminant subspace like in Huang et al. (1999) to increase inter-subject variations in each key frame.

Several approaches are designed to capture the static parameters of silhouettes. Kale et al. (2002) have obtained the width of each silhouette and modeled this with a hidden

Markov model. Cuntoor et al. (2003) have extracted the silhouette width vector of the leg region. Foster et al. (2003) have found the area of certain parts of the gait silhouette and used them as features. Johnson and Bobick (2001) have determined the maximum length of leg, the distance between head and pelvis, height, and stride length. The procrustes shape investigation is carried out by Wang et al. (2003) to obtain gait features from edge picture elements. Zhang et al. (2009) have modified the Wang's approach by estimating the closeness of procrustes mean shapes utilizing the shape description. Liu et al. (2011) have proposed an effective gait recognition method based on the outermost contour, followed by multiple discriminant analysis to optimize the separability of gait features. Gaba and Ahuja (2014) have found the various parameters by using Hanavan's model. This method is most suitable for the front view because if a side view is considered, the correct value for some of the features might not be obtained.

Another approach in which the gait silhouette has been transformed is by computing an average of it over a complete gait cycle. Liu and Sarkar (2004) have used an average silhouette, after that the concept called as GEI is employed by Man and Bhanu (2006). This portrayal turns out to be more resistant to minor phase variations and noise. This descriptor was proposed for the improvement of numerous variations that have now turned into the present trend in gait portrayal. Likewise in a similar means, the gait flow image (Lam et al. 2011), the gait entropy image (Bashir et al. 2009), and the masked GEI (Bashir et al. 2010) have also been created.

Many approaches are designed to capture the static and dynamic parameters of silhouettes and GEIs. Boulgouris and Chi (2007) have proposed a new approach for gait recognition using manually extracted and labeled silhouettes. The results of the body components carrying different discrimination powers are combined into a common distance metric for evaluating the similar gait sequences. Bashir et al. (2010) have projected an image retaining entropy statistics calculated based on arbitrariness of pixel values in silhouette gait sequence, called as gait entropy image. Hossain et al. (2010) have proposed an approach that is part based. In this technique, more damaged body sections are assigned less weight, and less damaged body sections are assigned more weight adaptively. Kawai et al. (2012) have developed a gait feature extraction technique based on a spatio-temporal histogram of oriented gradient. This technique uses colour information to get the internal silhouette motion. Rokanujjaman et al. (2015) presented an approach where the gait detection is done by partitioning the entire body into minute segments and further gathered gait features from the static area (torso) and

dynamic areas (arm swing, leg). Muramatsu et al. (2013) have formulated GEI with magnitude normalized silhouette sequences of gait, derived from graph cut background subtraction.

Amin and Hatzinakos (2012) have analyzed the recognition performance of dynamic features from various parts of the body and experimentally showed that dynamics of the lower leg and lower arm are of utmost significance for building an efficient gait recognition system, concluding that the use of lower leg is very noisy in the extracted silhouettes because of shadows and walking surface issues. Singhal and Lall (2013) have employed the features such as varying leg spread, the motion of centroid, the number of pixels on the vertical line through the centroid, and the sum of the foreground pixels as the dynamic features, and the height and maximum leg spread as the static features. These two sets of features are easy to obtain, yet contain significant information about the gait of a walking person. Kusakunniran et al. (2012) have devised a view transformation model to address the challenge of gait recognition under various viewing angles. The images taken from multiple views are converted into one view using this method. Kusakunniran (2014b) detect the space-time interest points from a raw gait video sequence. Here, the additional time complexity caused by the pre-processing and background segmentation is removed. This method relies only on local motion information. However, the pre-processing which removes the noise from the video is not performed, which may decrease the recognition accuracy. Hagui and Mahjoub (2016) have used speeded-up robust feature descriptors to depict the trajectories of the different body parts, and to illustrate their temporal movements, for which the motion history of the image is used. This method utilizes the spatial and temporal cues from each frame. Here, an assumption is made that the video sequences contain only one moving object. Hence, this method is not suitable for crowded scenes.

Shaikh et al. (2014) have extracted features from the portion of the silhouette that contains one of the most dynamic features of gait, such as the swinging hands of a human being. Here, the feature vector generation and the subsequent classification depend on partial silhouette that involves the handling of considerably fewer data. Al-Tayyan et al. (2017) have found three gait representation methods such as accumulated prediction image, accumulated flow image, and edge-masked active energy image. These methods are designed to overcome the challenges associated with covariates, such as carrying and clothing conditions. Roy et al. (2012a) combined gait and phase of motion with a spatio-temporal model for person identification using the pose energy image. It

results in the tracking of the subjects even if they change speed or stop for some time. Nandy et al. (2014b) derived a gait signature by decomposing the human body into three different independent structural segments, such as head node, arm swing, and leg swing areas, followed by the calculation of the area of each region. This results in the extraction of identical and independently distributed features.

Hu et al. (2013b) suggested a novel structure based on optical flow, which includes pattern retrieval, dynamics learning, and recognition. Burhan and Nordin (2015) had presented a solution using the enhanced GEI with Radon conversion techniques. Huang et al. (2016) developed a system based on spatio-temporal interest points for the detection of individuals using frontal gait videos. Kumar and Nagendraswamy (2014b) have found a period value depiction and identification of gait by the local binary pattern of split GEI. Here, the image is split into four equal parts. Liu et al. (2015) have made use of the similar descriptors with learned metric matrices and hierarchical feature extraction. Rida et al. (2016b) have come up with a system to choose the largely distinctive body part of humans using the group least absolute shrinkage and selection operator of motion to minimize the intraclass variance.

The physical structure of an individual can be captured well in the appearance-based approaches. The body shape itself has been shown to be discriminative in most of the methods. Many feature extraction techniques mentioned seem to capture this information efficiently. These features also contain temporal information, as well.

Instead of using a two-dimensional image, some studies used depth images. The motion capture devices such as Microsoft Kinect are used to obtain the depth images. Sivapalan et al. (2011) have evaluated the average silhouette volumes called gait energy volume from depth images. Hofmann and Rigoll (2012) proposed a representation for depth imagery using HOG.

In recent times, noteworthy attention has been received by deep learning methods from the computer vision society. This is because deep learning models can learn many layers of feature hierarchies by constructing high-level features from low-level features (Deng and Wang 2018). Particularly, several recent studies have shown promising results when deep learning methods are applied to a range of applications. For instance, Semwal et al. (2015) have designed a multilayered back-propagation algorithm based on Artificial Neural Network for gait classification. Shiraga et al. (2016) have used GEI, as an input to a Convolutional Neural Network designed for gait recognition called

GEINet. Alotaibi and Mahmood (2017) have developed a specialized Deep Convolutional Neural Network architecture for gait recognition. Deng and Wang (2018) have effectively modeled frame-to-frame gait dynamics of various subjects by radial basis function neural networks through deterministic learning.

Finally, the review of techniques modeling the temopral dynamics of a gait system is presented as follows: Lublinerman et al. (2006) have employed vectors of widths and Fourier descriptors to represent the silhouettes in each frame. This is followed by modelling of each sequence as a linear time invariant system that captures the dynamics of the evolution of the frame description vectors in time. Boulgouris et al. (2006) have proposed a method in which the recognition and verification of gait silhouettes are based on the matching of linearly time-normalized gait walking cycles. Lo Presti and La Cascia (2015) have developed a method to model the temporal dynamics of a sequence of facial expressions using linear time invariant system. This method can also be applied to model the temporal dynamics of gait sequences. Ding (2008) have applied the technique of dimensionality reduction to extract the spatial and temporal data by mapping the silhouette gait sequence to a low dimensional time sequence, which is considered as the output of a linear time invariant system.

On-going research on gait recognition accomplishes better performance in laboratory environments and is focussed more on accomplishing the same in the real-world environments.

2.2.1 Gait Recognition using Different Templates

Taking a cue from the GEI, several two-dimensional representations have been proposed. A plethora of papers in the literature supports the fact that the different templates used in the appearance-based approaches produce largely successful solutions to the problem of gait identification. Bashir et al. (2009) have used the gait entropy image, where a value called 'uncertainty'is assigned to each pixel. Here, the least entropy value is assigned to a more stable pixel across images in the gait sequence. Hence, higher values represent areas of motion. It performed as good as GEI on several databases, but an improved recognition accuracy is obtained on several covariate factors. Wang et al. (2010) have devised a novel template called chrono gait image. Here, the silhouette sequences are assigned an RGB color according to their timestamp. Later, an average is computed to capture a time sequence that is lost when computing the average silhouette. Motion silhouette image obtained from a normalized silhouette image is used as

a gait feature by Lee et al. (2014b). Here, the motion information which is critical to locomotion is considered, which amounts to an increase in the recognition accuracy of a system. Roy et al. (2012b) have developed a pose energy image, where an average of a chain of critical poses from a gait cycle was used to produce the gait feature.

2.2.2 Gait Recognition using HOG Descriptor

HOG has been used in several appearance-based approaches for human recognition based on gait features in recent times. Medikonda et al. (2018) have used HOG to generate two-fold information set features which capture temporal and spatial information from a gait cycle. Lee et al. (2014a) have also used HOG to acquire time-sliced average motion history image. This approach can reduce the influence of variations in gait speed, by obtaining different gait cycle phases with the concurrent utilisation of HOG and motion history images. Sivapalan et al. (2013) have formulated a method called histogram of weighted local directions, where a histogram based on local gradient directions is generated and used as a feature vector for gait identification. Huang et al. (2011) have extracted HOG from motion history image and demonstrated that HOG is less efficient for dynamic movements than static movements. Hofmann and Rigoll (2013) have given the concept of the gradient histogram energy image which captures edge information within silhouette for gait recognition. Mogan et al. (2017) have used a combination of binarized statistical features, motion history image, and HOG to increase the performance of a gait recognition system.

2.3 Influence of Covariates in Gait Recognition

Many different methods, such as a histogram of oriented gradients (Hofmann and Rigoll 2012), local binary patterns (Kumar and Nagendraswamy 2014a), deep learning methods (Alotaibi and Mahmood 2017), and so on have been used in recent times for gait recognition. Some of the gait detection methods show satisfactory performance under controlled set-ups (Prakash et al. 2018; Rida et al. 2016a; Binsaadoon and El-Alfy 2016). However, the application of gait recognition in real life is still limited, predominantly because of the various covariate issues such as a change in viewing angle, carrying and clothing conditions, walking speed variations, walking surface conditions that impact the individual's gait and hence, make gait recognition system more complicated (Sharma et al. 2011; Semwal et al. 2015).

2.3.1 Clothing Invariant Gait Recognition

An individual's clothing changes regularly and poses a significant challenge to vision-based gait recognition. However, this covariate seems to have less influence on the generation of actual gait, except in situations where the clothes are heavy or too tight. Many approaches have been proposed in the literature to resolve the problem of clothing variations. Matovski et al. (2010) have acquired the walking patterns of the same subjects while wearing standardized clothes and their own clothes on different days. The obtained GEI-based recognition accuracy was less when the subjects wore their own clothes than when they wore overalls. Hence, it can be concluded that the differing clothes of subjects over days adversely influenced the recognition accuracy of the system. In the real-world scenarios, the probe and gallery samples of an individual are collected in different circumstances. Hence, there may be substantial variations between the probe and gallery samples of an individual. This covariate is addressed by certain appearance-based approaches (Guan et al. 2012; Islam et al. 2013) using several templates that focus on various parts of the gait silhouette. Then, the templates which focus on familiar portions between probe and gallery samples can correctly recognize the subject. Further, Ghebleh and Moghaddam (2018) proposed an adaptive outlier detection technique to lessen the influence of clothing dissimilarities. Choudhury and Tjahjadi (2016) have used averaged gait key phase image and rotation forest for gait recognition. Islam et al. (2013) have converted the GEI into small window parts and defined a random window subspace method. Nandy et al. (2017) have conducted an investigation on gait entropy image for human detection invariant to clothing.

2.3.2 Frontal Gait Recognition

This section describes in detail the works done exclusively on frontal gait recognition. Soriano et al. (2004) introduced a descriptor known as curve spread for front view gait videos. Here, the time-variations of the outline of a moving body are represented as a 2D vector using Freeman code. Sivapalan et al. (2011) suggest a method based on frontal depth images. In this method, the concept of the GEl is extended to 3D, resulting in the creation of gait energy volume. Moreover, the study also presents the performance of the reconstruction of partial volumes from depth images of the front view. Sivapalan et al. (2012) have proposed back filled GEI for frontal gait recognition. This can be obtained from both side-view silhouettes and frontal depth images. Huang et al. (2016) have presented a spatio-temporal interest point based method in which the histogram of oriented gradients were extracted directly from the gait videos with-

out silhouette extraction. Further, the spatio-temporal interest points of two different gait sequences are eventually matched using histogram of oriented gradients so as to measure the similarity between two gait sequences.

Barnich and Van Droogenbroeck (2009) have developed an approach using the intra-frame description of silhouettes consisting of a set of rectangles, fitted into any closed silhouette. This is followed by the addition of a dynamic frame obtained from averaging the size distributions over successive gait sequences. Goffredo et al. (2008) have used a non-calibrated camera to obtain identical signatures from the silhouette's descriptors. Chattopadhyay et al. (2015) make use of the kinetic depth data to provide solution to the problem of occlusion in frontal gait recognition, where the features of the back view are extracted from the depth data and the edge of the silhouette. They have proposed a hierarchical strategy for classification that merges the back and front view descriptors obtained from red green blue depth cameras (Chattopadhyay et al. 2014b). Nambiar et al. (2012) have created a hybrid 2D and 3D scheme to extract the frontal gait features. Chattopadhyay et al. (2014a) have reconstructed partial volume of the surface of each silhouette obtained from the frontal view and derived a feature called as pose depth volume. Ryu and Kamata (2011) have proposed a gait template known as Marching in Place to construct the spatial and temporal features of human gait sequences. Although all these methods deal with the frontal gait recognition, most of them use the gait database containing less number of subjects obtained from Kinect for performance analysis.

2.3.3 Speed Invariant Gait Recognition

The variations in speed are the result of changes in the effort put in or the average force during walking and results in variations in all dynamic or time-based features. Some of the works which have been done exclusively on speed invariant gait recognition are as follows: Nandy et al. (2014b) decompose the human body into leg swing, arm swing, and head regions. They signify the sides of an n-sided polygon for computing the area of all regions. Later, the convex hull of the features in each region is constructed to obtain the key features. Makihara et al. (2014) have used static and transformed kinematic features for speed invariant gait detection. A transformation model for the kinematic gait features is formed by means of a training set for many non-recognition target individuals with various speeds. This model can convert kinematic features from a gallery speed to another probe speed.

Kusakunniran et al. (2011) have extracted features based on procrustes shape analysis in order to propose a novel descriptor called higher order derivative shape configuration which describes gait changes in shape. The two significant components of procrustes shape analysis, such as procrustes mean shape and procrustes distance are adapted precisely for the purpose of speed invariant gait recognition. Rokanujjaman et al. (2016) select the less affected body parts to extract gait features. The features extracted from these parts are concatenated to obtain the feature vector of an image.

Huang et al. (2015) have introduced a speed invariant gait template and proposed an algorithm called enhanced locality. This algorithm preserves projections for reducing the dimensionality of the proposed gait template. Nandy et al. (2013) find the joint angles of ankle and knee using a stick figure of the gait silhouette image, followed by the extraction of various other static features. They have further employed a dynamic edge orientation histogram from silhouette images of a gait cycle and used them as a feature vector for classification (Nandy et al. 2014a).

Medikonda et al. (2018) have introduced information set features to capture temporal and spatial information from gait silhouettes. These features are extracted by obtaining a histogram of oriented gradients for the sequence of gait images, followed by representing them using Hanman-Jeevan entropy function. Kovač et al. (2017) have developed a frame-based classification method and wavelet transform signal approximation for analyzing the feature signals on different frequency space resolutions. This frame-based classification approach overcomes the drawbacks of distance-based methods. Arora et al. (2015a) have proposed gait information image and further altered the data that was obtained by using energy and sigmoid functions. The other approaches for speed invariant gait detection are pseudo-shape based features (Tan et al. 2007b), uniprojective features (Tan et al. 2007c), wavelet packet gait vectors (Dadashi et al. 2009) and Fourier descriptor (Lee et al. 2013).

2.4 Gait Recognition using Running Patterns

Running is considered as an extension of walking, and it consists of different joint movements, coordination, and higher velocities. Both the running cycle and walking cycle can be called as a gait cycle. The foot makes contact with the ground in a different way for running and walking. The main difference between jogging and running is that the traveling speed is prominent during running, and there is also a difference in the distance between the two legs (Semwal et al. 2015).

Figure 2.1: Sample images from CASIA B gait dataset (Zheng 2017).

On reviewing the literature, it has been observed that very few attempts have been made for recognizing the individuals by using their running video database. Yam et al. (2001) have presented a model-based approach on a dynamically coupled oscillator. This model was used to recognize the people using the walking and running database. They have obtained reasonably good results, but the experiments were conducted on the database containing seven subjects. Yam et al. (2002) presented another model for lower leg and thigh. Here, the gait features are obtained from the phase-weighted magnitude of the lower order Fourier components of the knee rotation and thigh. The features were extracted from a small database of 5 subjects, and then they were classified to get the recognition results. Keeping the above factors in view, there was a need to propose a robust method for the recognition of people using their running and jogging patterns.

2.5 Gait Databases

The increasing rate of research on gait recognition led to the development of many benchmark gait databases. The list of the gait datasets available and their metadata are presented in Table 2.2. The remainder of this section describes the benchmark gait datasets used in this thesis for the evaluation of the performance of the proposed approaches.

2.5.1 CASIA A Dataset

The CASIA A dataset consists of 20 subjects. The walking style of each subject is captured in three different views, namely lateral (0^0), oblique (45^0) and frontal (90^0) with respect to the image plane (Zheng 2017). Each subject listed in the dataset has 4 gait sequences for each viewing angle.

2.5.2 CASIA B Dataset

The CASIA B dataset is a multiview gait dataset comprising 124 subjects from 11 views, specifically 0^0, 18^0, 36^0, 54^0, 72^0, 90^0, 108^0, 126^0, 144^0, 162^0, and 180^0. Under each view, ten gait sequences are taken for each person with six sequences in normal

Table 2.2: Summary of gait datasets.

Dataset	No. of subjects	Environment	Variations
USF	122	outdoor	views, carrying, speed, surface, footwear
SOTON	115	outdoor, indoor, treadmill	view
MIT	24	indoor	view
UMD-1	25	outdoor	4 views
UMD-2	55	outdoor	2 views
CASIA A	20	outdoor	3 views
CASIA B	124	indoor	11 views, clothing, carrying
CASIA C	153	outdoor	speed
CMU MoBo	25	indoor, treadmill	6 views, speed, carrying
OU-ISIR A	34	indoor, treadmill	speed
OU-ISIR B	68	indoor, treadmill	clothing
OU-ISIR D	185	indoor, treadmill	speed
KTH	25	outdoor, indoor	6 human actions
Weizmann	10	outdoor, indoor	10 human actions

walking (*nm*), two sequences in walking when carrying a bag (*bg*), and two sequences in walking when wearing a coat (*cl*). Each video was recorded at a frame rate of 25 fps with a resolution of 320×240 (Kusakunniran 2014b). Figure 2.1 shows the sample images from CASIA B gait dataset .

2.5.3 CASIA C Dataset

The CASIA C gait dataset consists of 153 subjects. The gait data of all subjects is captured from an infrared camera. These subjects walk with three different speeds such as, slow (*fs*) of 4 km/h, normal (*fn*) of 5 km/h, and fast (*fq*) of 6 km/h, respectively (Zheng 2017). The dataset consists of four gait cycles from a normal sequence and two gait cycles from slow and fast sequences, respectively.

2.5.4 OU-ISIR A Dataset

This dataset contains gait sequences of 34 subjects. Their gait sequences were taken while they were walking in a side view. The speed variation of these subjects ranges from 2 to 10 km/h with an interval of 1 km/h (Makihara et al. 2012). The range of speed which varies from 2 to 7 km/h is considered as walking, and 8 to 10 km/h is categorized as running.

2.5.5 OU-ISIR B Dataset

The OU-ISIR B gait dataset includes 68 subjects and 32 combinations of clothing types to exhibit a comprehensive analysis of the effects of clothing variations. All the gait sequences were collected twice on the same day. This gait dataset contains the maximum number of clothing conditions (Deng and Wang 2018).

2.5.6 OU-ISIR D Dataset

The OU-ISIR D gait dataset contains gait patterns of 185 subjects. It has two types of data: (1) DB_{high} having the highest Normalized Auto Correlation (NACs) and (2) DB_{low} having the lowest NACs. Each type consists of 100 subjects. The dataset contains 360 gait silhouette sequences captured from a side view of 185 subjects. The dataset consists of two sequences of each individual.

2.5.7 CMU MoBo Dataset

The CMU MoBo is a gait dataset containing gait sequences of 25 individuals (Gross and Shi 2001). The gait sequences of each subject are recorded in the CMU 3D room and during treadmill walking. The gait videos were captured in different manners of walking including fast walk (*f*), slow walk (*s*), walking with a ball in one hand (*b*), and walking on an inclined plane (*i*).

2.5.8 KTH Video Dataset

The KTH dataset is a video dataset containing six types of human actions (walking, jogging, running, boxing, hand waving and hand clapping) performed several times by 25 subjects in four different scenarios: outdoors (*sn1*), outdoors with scale variation (*sn2*), outdoors with different clothes (*sn3*), and indoors (*sn4*) (Schuldt et al. 2004).

2.5.9 Weizmann Dataset

The Weizmann dataset was recorded in 2005. Its background is relatively simple, and only one person is instructed to act in each frame. It contains 10 human actions (walking, running, jumping, galloping sideways, bending, one-hand waving, two-hands waving, jumping in place, jumping jack, and skipping), each of which was performed by nine subjects (Chaquet et al. 2013).
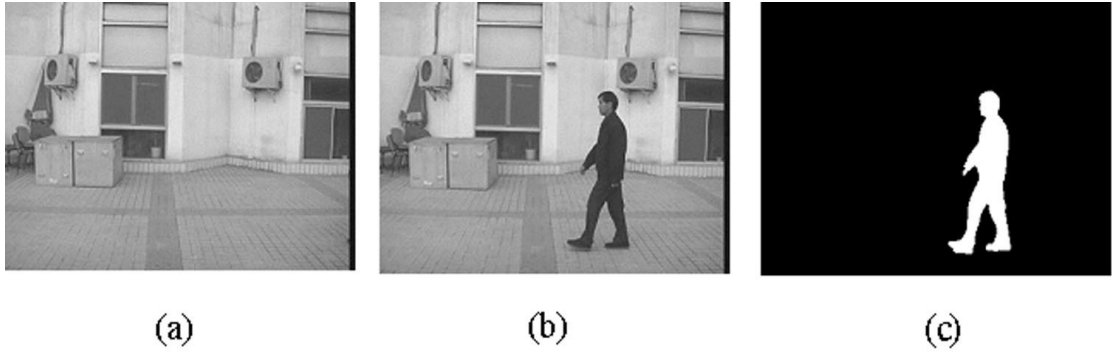
Figure 2.2: Example image representing (a) background image, (b) original image, and (c) extracted silhouette (Wang et al. 2003).

2.6 Silhouette Extraction

The most commonly used data pre-processing step in vision-based gait recognition is silhouette extraction. Many researchers have used the human gait silhouette image instead of a colour image, as the information of the background and the colour of clothes worn by the subject is redundant information with respect to human gait recognition. Hence, in gait recognition, the fundamental task is the conversion of gait images into gait silhouette images. A general approach is to perform background subtraction, by which the gait silhouette is identified as a moving object within the scene. Several techniques have been developed in the past for background subtraction. Some of the popular ones include frame differencing (Cheung and Kamath 2005), estimating the background as the average of n frames (Cucchiara et al. 2003), computing the Gaussian average (Wren et al. 1997) and so on. Later on, steps such as de-noising, post-processing and normalization are carried out to obtain gait silhouette images. A sample original image and the extracted silhouette image are shown in Figure 2.2.

Almost all benchmark databases contain the data of subjects in the form of silhouette images. Hence, in this thesis, the available gait silhouette images from the databases are directly used for feature extraction.

2.7 Gait Energy Image

Numerous appearance-based methodologies prefer a two-dimensional representation known as GEI for a gait cycle. Each one of the GEI templates is formed in a procedure given below.

1. A subject's silhouettes from a single gait cycle are extracted.

2. Centre alignment of the silhouettes needs to be done, and they are resized to a
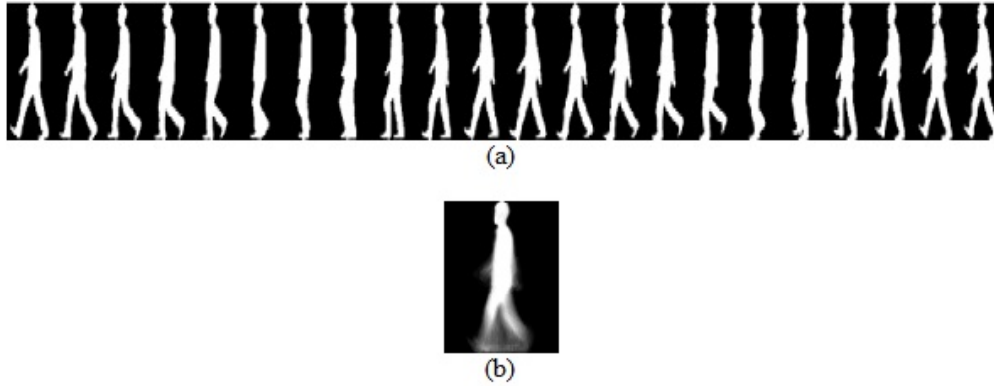
Figure 2.3: Sample image representing (a) gait cycle (b) GEI.

specified size.

3. The resized silhouettes of a specified gait cycle are combined by a method to create the gait template.

GEI is the portrayal of human gait using a greyscale image acquired by calculating the mean of the silhouettes obtained over an entire gait cycle. It is computed as follows.

$$G(i,j) = \frac{1}{N} \sum_{f=1}^{N} S(i,j,f) \tag{2.1}$$

where N represents the total number of frames of a gait cycle, S specifies the silhouette image, i and j signify the spatial co-ordinates of the image and f shows the frame number in the gait cycle (Rida et al. 2016a). A sample image representing the gait cycle and GEI is shown in Figure 2.3.

Pixel with more intensity relates to the inert parts of the body (upper region), and this portion consists of the body contour data, which can be valuable for recognition. However, it can be influenced by the covariates too. Pixel with less intensity belongs to the active parts of the body (lower region), and this piece of the GEI is exceptionally beneficial for the recognition and is not influenced by the covariates such as clothing and carrying conditions. The visual representation of the intensity of the pixel values of a GEI is shown in Figure 2.4. The benefit of using GEI is that it drastically reduces the feature dimensionality in comparison with the complete gait sequence; hence, it decreases the computational cost as well.
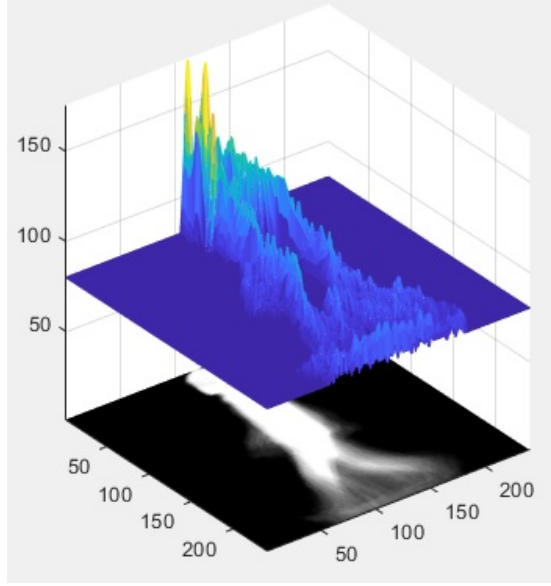
Figure 2.4: The pictorial representation of the intensity values of the pixels in GEI.

2.8 Performance Measure of Gait Recognition Methods

The accuracy with which an algorithm classifies the subjects represent the performance of a gait recognition system. The samples present in the database are divided into the training and test sets to estimate this metric. The performance of the gait recognition system is obtained by Correct Classification Rate (CCR) or recognition accuracy computed on the test set. The CCR is computed by using Equation 2.2.

$$\text{CCR} = \frac{\text{Number of subjects correctly classified in the test dataset}}{\text{Total number of subjects present in the test dataset}} \times 100 \qquad (2.2)$$

2.9 Relief Feature Selection Algorithm

Kira and Rendell formulated the Relief algorithm inspired by instance-based learning (Kira and Rendell 1992). As an individual evaluation filtering feature selection method, Relief calculates a proxy statistic for each feature used to estimate feature quality or relevance to the target concept (i.e., predicting endpoint value). These feature statistics are referred to as feature weights, or more casually, as feature scores that can range from worst to best.

The Relief algorithm cycles through m random training instances (R_i), selected without replacement, where m is a user-defined parameter. Each cycle, R_i is the target

instance, and the feature score vector W is updated based on feature value differences observed between the target and neighboring instances. Therefore, the distance between the target instance and all other instances is calculated for each cycle. Relief identifies two nearest neighbor instances of the target; one with the same class, called the nearest hit (H) and the other with the opposite class, called the nearest miss (M). The last step of the cycle updates the weight of a feature A in W if the feature value differs between the target instance R_i and either the nearest hit H or the nearest miss M. Features that have a different value between R_i and M support the hypothesis that they are informative of the outcome, so the quality estimation $W[A]$ is increased. Conversely, features with differences between R_i and H provide evidence to the contrary, so the quality estimation $W[A]$ is decreased.

2.10 Pattern Classification

Classification methods are used to identify test subjects by matching the discriminative features extracted on the previously known subjects. Some of the methods proposed in this thesis use K Nearest Neighbors (KNN) classifier (Cunningham and Delany 2007), which is explained as follows:

2.10.1 K Nearest Neighbors

KNN is an instance-based, simple yet powerful method of classification, which is generally used in gait recognition systems. The training phase of KNN consists of the known subjects and their labels. The testing phase involves the computation of the distances between a particular test subject and the known subjects in the training dataset by using certain distance measures. The measure used is Euclidean distance. It is the most commonly used measure in gait identification methods. The distances which are computed for each gallery and probe samples are ordered from the minimum value to the maximum and the subjects in the gallery with K values will be selected. The most frequent subject in the K nearest neighbor is determined as the matching subject. The two parameters which need to be tuned for this algorithm are the value of K and distance metric. KNN performs better with lower computational cost and an easy implementation. Hence, many researchers have used this for the purpose of classification in gait recognition approaches.

2.11 Outcome of Literature Review

Based on the literature survey on the gait recognition system, the conclusions made are as follows:

1. In recent years, the attention has been paid for the improvement and expansion of the individual recognition frameworks. From a surveillance point of view, gait recognition is an appealing system because of its capability of recognizing individuals from a distance by examining their manner of walking. Gait is one of the few behavioural biometric traits that can be considered without bodily contact. This makes it valuable in surveillance applications.

2. The recent approaches of gait recognition are mostly based on the appearance-based methods because of their simplicity and higher recognition rates. Predominantly, they focus on offering solutions for the real-world circumstances, where the changes in appearance need to be addressed. In a number of gait challenging situations, most common appearance changes that could generally occur in real-world applications are because of different views, walking speed, clothing, and carrying things.

3. HOG is one of the robust descriptors of a greyscale image. Even though HOG is broadly used for gait recognition in some of the recent works, the size of the feature vector used in those works is extremely large. These high-dimensional features will consequently increase the search space. Additionally, it will also increase the run-time complexity of the algorithm, which grows exponentially with the increase in dimension.

4. Clothing of a person changes regularly, hence, it is an important challenge to vision-based gait recognition. The differences in clothing style between the probe and gallery databases influence the silhouette of an individual. As a result, it is necessary to eliminate the regions of silhouette, which are mostly influenced by clothing. The next step is to choose more distinguishing gait characteristics that can be considered as an essential insight into the features of gait biometrics. These features are selected to be clothing invariant despite the differences in the individual's appearance.

5. The frontal view has certain advantages. Here, the walking person approaches towards the camera; hence it does not require more space, unlike other view variations. Other biometrics such as face and iris are captured in frontal view. Hence, it is necessary that gait is also required to be captured in a frontal view to integrate with other biometrics in certain scenarios. However, it is challenging to capture all dynamics of

gait in a frontal view, and therefore it struggles to perform better. Hence, the frontal gait detection performance can be enhanced by building a robust representation of spatial data.

6. When an individual changes the speed of walking, the dynamic features change, whereas static features remain unchanged. Hence, the spatial data is predominantly used to extract discriminative gait features in speed invariant gait recognition. Since GEI carries both static and dynamic information, many researchers preferred gait silhouette images instead of GEI for feature extraction. This is because the dynamic information present in GEI changes with the speed of the person. Here, the disadvantage is that the use of all silhouette images to extract the gait features increases the computational cost of a gait recognition system. Hence, an approach needs to be developed which overcomes this setback, and extracts the most relevant features from GEI, to identify a subject with different speed variations.

7. Little research has been conducted on identifying individuals using other gait patterns such as jogging and running. Using gait as a biometric is motivated by occlusion of criminal's faces and that they either walk or run to getaway from a scene of the crime. On the other hand, the likelihood of identifying the individuals by the way of running remains predominantly unexplored. Very often, robbers and criminals naturally escape by running, instead of walking. Therefore, it is necessary to come up with a recognition method that distinguishes people by considering their running patterns.

The discovered research gaps provide the directions and pathways for the research conducted in this thesis. The remaining chapters of this thesis present the contributions and outcomes of the research that addresses the essential parts of these research gaps.

2.11.1 Problem Statement

The research problem is defined as, "To propose efficient approaches to improve the performance of the gait features-based human recognition system".

2.11.2 Research Objectives

The research objectives are defined as follows:

1) To propose a low-dimensional feature vector which captures the spatial dynamics of gait patterns to increase the performance of a gait recognition system.

2) To propose efficient methods to increase the performance of the clothing invariant and speed invariant gait recognition.

3) To propose efficient methods to increase the performance of the frontal gait recognition.

4) To propose a gait recognition system that identifies individuals using the running patterns.

2.12   Summary

In this chapter, the existing state-of-the-art appearance-based gait recognition approaches are discussed. Discussion about both the existing literature pertaining to different gait templates, and the covariates influencing the gait of a subject is made in this chapter. It also discussed the existing works with respect to other concepts such as HOG and running gait patterns. Finally, the problem definition and research objectives for proposing the efficient gait recognition frameworks are highlighted.

Chapter 3

LOW-DIMENSIONAL FEATURE VECTOR FOR HUMAN GAIT RECOGNITION

One of the significant challenges of the appearance-based gait recognition system is to augment its performance by using a distinctive low-dimensional feature vector. Therefore, this study considers the low-dimensional features that are capable of effectively capturing the spatial, gradient, and texture information.

3.1 Gait Recognition based on Histogram of Oriented Gradients and Haralick Texture Descriptor

It is evident from the literature that HOG is a proficient feature descriptor used in gait recognition systems (Arora et al. 2015b; Hofmann and Rigoll 2012; Huang et al. 2011; Mogan et al. 2017; Whytock et al. 2012). HOG descriptor is used to get the local shape information from a window or region of interest of a gait image. Even though HOG is used for gait recognition in some of the recent works (Arora et al. 2015b; Hofmann and Rigoll 2013), the size of the feature vector used in those works is extremely large as shown in Table 3.1. These high-dimensional features consequently increase the search space. Additionally, it also increases the run-time complexity of the gait recognition system, which grows exponentially with the increase in dimension.

Moreover, with a fixed number of training samples, the predictive power reduces as the dimensionality increases. To overcome this setback, it is necessary to devise an efficient method that increases the performance of the gait recognition system with less number of HOG features. This has been accomplished by the combination of HOG and sum variance Haralick texture descriptor. Furthermore, the texture is one of the vital biometric characteristics that is always present in an image. It serves to quantify and identify regions of interest in an image. It can be used along with HOG to extract the most distinctive low-dimensional feature vector capable of identifying individuals uniquely.

The significant contributions of this study can be summarised as follows:

1. In order to preserve the uniqueness in GEI and yet remove the redundant information which can be regarded as less significant, GEI is converted into GGMI

Table 3.1: Size of the HOG descriptor used in existing gait recognition methods.

Method	Size of the HOG feature vector
GEI-HOG (Whytock et al. 2012)	5400×1
MHI-HOG (Huang et al. 2011)	3780×1
GHEI-HOG (Hofmann and Rigoll 2012)	3780×1
α-GHEI (Hofmann and Rigoll 2013)	3780×1
GGI-HOG (Arora et al. 2015b)	3780×1
TAMHI-HOG (Lee et al. 2014a)	3780×1
MHI-BSIF-HOG (Mogan et al. 2017)	3780×1

which contains only the key structural elements of GEI.

2. This study presents a feature vector based on the combination of HOG and sum variance Haralick texture descriptor. This feature vector is more reliable and better characterizes the spatial variations of gait. The length of the feature vector is reduced from 3780×1 to 63×1, enabling the reduction of the computational complexity of the gait recognition system.

3. A detailed experimentation is done regarding the selection and usage of sum variance Haralick texture descriptor among many other texture descriptors extracted from GGMI to find how it increases the recognition accuracy of a gait detection system when combined with HOG.

4. A broad experimental evaluation is performed in this work. The proposed method is tested on five benchmark gait datasets. The findings are compared with the state-of-the-art and other HOG based gait recognition techniques.

3.1.1 Framework of the Proposed Method

The framework of the proposed gait recognition approach is demonstrated in Figure 3.1. It consists of four stages. Firstly, the gait video is converted into a sequence of gait silhouette images. These gait silhouettes are combined over a gait cycle to form GEI, which is further converted into GGMI. Secondly, the proposed gait features are extracted from GGMI's of the subjects belonging to a particular dataset. Thirdly, Principal Component Analysis (PCA) is used to pre-process the extracted features. It reduces the dimensions that negatively influence the robustness of the classification thereby increasing the performance. Finally, the extracted features from a given dataset are classified by using the KNN classifier.

3.1.2 Gait Gradient Magnitude Image (GGMI) Extraction

GGMI is an outcome of applying the Gaussian filter (Deng and Cahill 1993), followed by a Sobel edge detector (Shrivakshan and Chandrasekar 2012) on GEI. While the
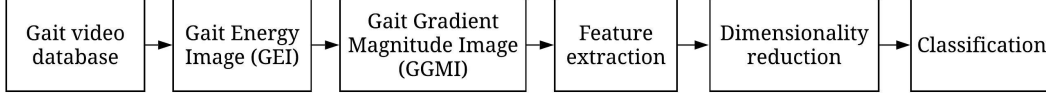
Figure 3.1: Structure of the proposed approach.
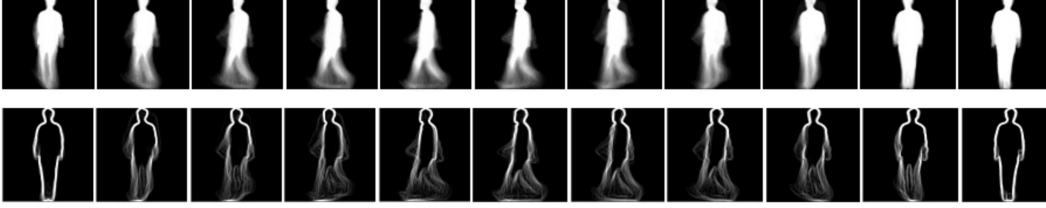
Figure 3.2: Sample images from the CASIA B gait dataset representing GEI and GGMI for various views.

Gaussian filtering removes the detail and noise from GEI, a pair of 3×3 kernals of Sobel operator highlights the regions of high spatial frequency which correspond to edges (Sharifi et al. 2002).

Hence, applying the Gaussian filter, followed by a Sobel edge detector on GEI considerably decreases the quantity of information to be processed and consequently removes the information that is regarded as less significant. At the same time, it preserves the key structural properties of GEI.

In this work, the GGMI template is prepared for feature extraction. After the extraction of GEI from a gait cycle, it is further converted into GGMI by the steps of Algorithm 1. The method employed is aimed at capturing the magnitude or edge information from the subject's GEI. Sample GGMI's for the normal walking condition with different views are shown in Figure 3.2.

Algorithm 1 Generation of GGMI.

1: **Input**: GEI
2: **Output**: GGMI
3: Begin
4: Read an input image.
5: Let G_f be the Gaussian filter of size 5×5 and $\sigma = 1.0$.
6: Compute C by performing convolution of GEI with filter G_f.
7: Let S_x and S_y be the Sobel kernals for horizontal and vertical direction.
8: Perform convolution of C with S_x to obtain CS_x.
9: Perform convolution of C with S_y to obtain CS_y.
10: Compute magnitude: GGMI = $\sqrt{(CS_x)^2 + (CS_y)^2}$
11: End

3.1.3 Histogram of Oriented Gradients

Recent literature reveals that using HOG in appearance-based methods improves the performance of gait recognition (Hofmann and Rigoll 2013; Medikonda et al. 2018; Mogan et al. 2017). HOG (Dalal and Triggs 2005) is a global descriptor which portrays the distribution of edge directions or intensity gradients. To compute the HOG feature, initially, the gradient vector at each pixel in the GGMI is used. The 1st order gradient operators that are applied to extract the horizontal and vertical gradient magnitudes are as follows: $f\_xdir = [\,-1\ 0\ 1\,], f\_ydir = [\,-1\ 0\ 1\,]^T$.

These horizontal and vertical gradient images are then combined to obtain gradient magnitude and orientation. The formula for the same is given in Algorithm 2. Later, the cell histograms are generated. The image is partitioned into the small associated areas called cells. The next step is to generate a histogram of gradients in these cells. A bin is selected based on the pixel intensity in gradient orientation, and the vote (the value that goes into the bin) is selected based on the pixel intensity in gradient magnitude. To construct HOG, this vote is cast by every pixel present inside the cell. As the cells overlap half of their area, each cell contributes more than once to the final feature vector.

For all pixels inside every cell, a histogram of gradient directions is generated. The concatenation of these histograms forms a HOG descriptor. The gradient values thus obtained are locally normalized, i.e., they are normalized over every cell to account for the variations in contrast and illumination. 9 equal-sized rectangular cells and 7 bin histograms per cell are used to obtain the feature vector. The degree of the bins ranges from $-\pi$ to π, with a range of 51.43^0 per bin. The number of features relies on the number of cells and histogram bins. Hence, the 9 cell histograms with 7 bins are concatenated to make a 63×1 dimensional HOG descriptor. In this way, the HOG is used to represent the features of each GGMI. This study utilizes the HOG descriptor to enable a better characterization of the appearances and shapes of the subjects based on the distribution of the local intensity gradients with the oriented directions.

3.1.4 Haralick Texture Descriptor

The initial stage in texture analysis using Haralick features is to compute the Gray-Level Co-occurrence Matrix (GLCM) (Baraldi and Parmiggiani 1995) that has the desired

Algorithm 2 Extraction of gait features.

1: **Input**: GGMI, of size $p \times q$.
2: **Output**: Feature vector, f.
3: Begin
4: Read an input image.
5: $f\_xdir = [\,-1\ 0\ 1\,]$
6: $f\_ydir = [\,-1\ 0\ 1\,]^T$
7: Compute the gradient images, g_xdir and g_ydir, by performing convolution of GGMI using f_xdir and f_ydir.
8: Compute orientation: $angle = arctan(g\_ydir/g\_xdir)$.
9: Compute magnitude: $m = \sqrt{(g\_ydir)^2 + (g\_xdir)^2}$
10: Let R be the no. of R-HOG cells.
11: Let BI be the total no. of bins.
12: **for** 1st cell **do** (Refer Algorithm 3)
13:     **for** $ang\_hist = -\pi + 2 \times \pi/BI : 2 \times \pi/BI : \pi$ **do**
14:         Quantize the gradient orientation to obtain feature vector f_g, of size $BI \times 1$.
15:         Normalize f_g to get f_n.
16:     **end for**
17:     Compute GLCM of the 1st cell.
18:     Compute Sum Variance, f_{sv} in 4 directions of θ (i.e, 0^0, 45^0, 90^0, 135^0).
19:     Compute mean $f_{sv} = (f_{sv}^0 + f_{sv}^{45} + f_{sv}^{90} + f_{sv}^{135})/4$.
20:     Multiply each value of f_n with f_{sv} to obtain a feature vector f_1.
21: **end for**
22: Repeat for all cells of an image.
23: Concatenate the feature values obtained from R cells into a feature vector f, $f = [f_1 f_2 ......f_R]$.
24: End
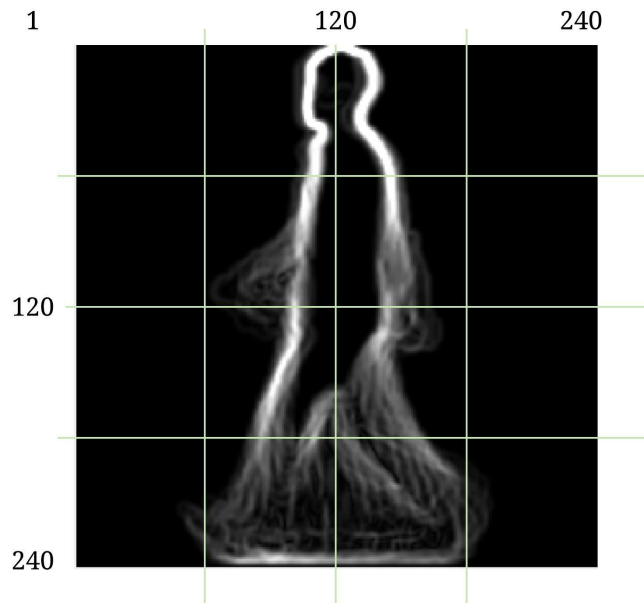
Figure 3.3: Representation of the equal sized cells of GGMI used for feature extraction.

number of gray levels.

GLCM is one of the most well-known texture feature extraction methods. It is used to estimate the second-order statistical properties of the images, and it provides useful data regarding the relative position of the neighboring pixels in a gait image. Each entry in GLCM is therefore considered to be the probability that a pixel with value g will be found adjacent to a pixel with value h. An element $p(g, h)$ of the co-occurrence matrix is computed by the relative frequencies in which two pixels, one with gray level g and the other with gray level h are separated by a distance d and occur in a direction specified by the angle θ.

$$p(g,h) = \sum_{x=1}^{N_g} \sum_{y=1}^{N_g} \begin{cases} 1 & \text{if } I(x,y) = g \quad \text{and} \quad I(x+dx, y+dy) = h \\ 0 & otherwise \end{cases} \tag{3.1}$$

Where the distance between the pixel of interest and its neighbor is specified by dx, dy, and the number of gray levels in the GGMI is represented by N_g. For different orientations of θ at 0^0, 45^0, 90^0, and 135^0, multiple GLCM's can be computed which can depict the spatial relationship between neighboring pixels thoroughly and can result in reliable texture features of GGMI images.

The GLCM of GGMI with dimension $N_g \times N_g$ can be defined as

$$P = \begin{bmatrix} p(1,1) & p(1,2) & ...... & p(1, N_g) \\ p(2,1) & p(2,2) & ...... & p(2, N_g) \\ . & . & ...... & . \\ . & . & ...... & . \\ p(N_g, 1) & p(N_g, 2) & ...... & p(N_g, N_g) \end{bmatrix} \tag{3.2}$$

The normalized GLCM is computed by Equation 3.3.

$$P = \frac{p(g,h)}{\sum_{g=1}^{N_g} \sum_{h=1}^{N_g} p(g,h)} \tag{3.3}$$
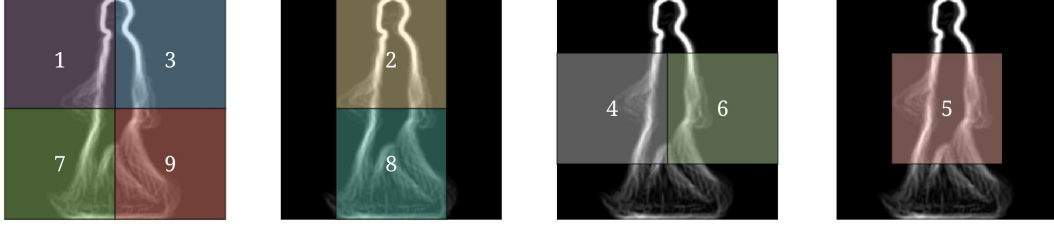
Figure 3.4: Sample images representing 9 different cells of GGMI.

The next step is to compute the Haralick texture features from the normalized GLCM of GGMI. Haralick features (Haralick et al. 1973) are the 14 statistical entities that represent the texture properties from P (Lishani et al. 2017). This texture descriptor is extracted and calculated in four directions. That is, θ in 0^0, 45^0, 90^0 and 135^0.

Let x and y be the row and column coordinates of an entry in the co-occurrence matrix.

Then $p_x(g)$ and $p_y(h)$ is given by

$$p_x(g) = \sum_{h=1}^{N_g} p(g,h) \quad \text{and} \quad p_y(h) = \sum_{g=1}^{N_g} p(g,h) \tag{3.4}$$

Further, $P_{x+y}(i)$ is the probability of co-occurrence matrix coordinates summing to $x+y$ defined as follows:

$$p_{x+y}(r) = \sum_{g=1}^{N_g} \sum_{h=1}^{N_g} p(g,h) \quad \text{where} \quad r = g+h \quad \text{with} \quad r = 2,3,...,2N_g \tag{3.5}$$

In this study, only the sum variance (f_{sv}) feature is considered in the process of feature extraction. The sum average (f_{sa}) Haralick texture descriptor needs to be computed in order to calculate f_{sv}. The sum average is given by

$$f_{sa} = \sum_{r=2}^{2N_g} r p_{x+y}(r) \tag{3.6}$$

40

Thus, sum variance is given by

$$f_{sv} = \sum_{r=2}^{2N_g} (r - f_{sa})^2 p_{x+y}(r) \tag{3.7}$$

3.1.5 Feature Extraction

The discriminative features evaluated in our feature extraction method include HOG and Haralick texture descriptor. The feature extraction comprises the following steps:

1. Divide GGMI into 9 equal sized cells.

2. Compute the HOG features from 9 different cells of GGMI.

3. Compute the sum variance Haralick feature for the 9 different blocks.

4. Multiply the value of the sum variance feature of a cell with the HOG features of the same cell.

5. Perform Step 4 for all nine cells of GGMI.

6. Concatenate the feature values obtained from 9 cells.

Algorithms 2 and 3 present the detailed steps followed in extracting the feature vector. Figures 3.3 and 3.4 represent the 9 different cells of GGMI.

Algorithms 2 and 3 are explained as follows: Initially, the gradient images g_xdir and g_ydir are obtained by performing convolution of GGMI using the gradient operators f_xdir and f_ydir. The next step is to obtain two images called the magnitude image m_img and the orientation image or_img from the gradient images by using Steps 8 and 9 of Algorithm 2. Let the total number of bins BI be 7. Here, both the magnitude and the orientation images are divided into 9 R-HOG cells. For each cell, the pixels present in the magnitude and the orientation images of that cell are converted into a column vector V_m_img and V_or_img.

In order to assemble the histogram, a range of 7 orientation bins are considered between the range $-\pi$ to π, with $2 \times \pi/BI$ assigned to each bin as shown in Algorithm 3. The computation is performed on the magnitude and the orientation images for a specific range at each iteration. Here, each pixel present in the orientation image of a cell is checked as to whether it is present in a particular range of orientation. If it

is present in that range, the value of the corresponding pixel in the magnitude image is added and stored, as shown in $f_g(bin)$. For each cell, 7 values are obtained. Hence, a feature vector f_g of size 63×1 is obtained. It is normalized to get f_n.

Later, the GLCM of each cell is computed, followed by the computation of sum variance texture descriptor in four directions of θ (i.e., 0^0, 45^0, 90^0, 135^0). Finally, the mean of all 4 directions, f_{sv} is obtained. The sum variance value obtained for a particular cell is multiplied with the feature vector f_n of the same cell. At last, the features of all nine cells are concatenated to get the final feature vector f.

Algorithm 3 Assembling the histogram with 7 bins.

1: **Input**: Orientation image, or_img of size $p \times q$, Magnitude image, m_img of size $p \times q$.
2: Begin
3: Read two input images.
4: Define a column vector with zeros, f_g of size $BI \times 1$.
5: **for** 1st cell, of size $a \times b$ **do**
6:     Convert or_img and m_img into corresponding column vectors, V_or_img and V_m_img.
7:     Obtain the maximun size, S of V_or_img.
8:     $bin = 0$.
9:     **for** $ang\_hist = -\pi + 2 \times \pi/BI : 2 \times \pi/BI : \pi$ **do**
10:         $bin = bin + 1$
11:         **for** $s = 1 : S$ **do**
12:             **if** $V\_or\_img(s) < ang\_hist$ **then**
13:                 $V\_or\_img(s) = 100$
14:                 $f_g(bin) = f_g(bin) + V\_m\_img(s)$
15:             **end if**
16:         **end for**
17:     **end for**
18: **end for**
19: **Output**: Feature vector f_g, of size $BI \times 1$ is obtained.
20: End

3.1.6 Experiments

3.1.6.1 Selection of the Sum Variance Texture Descriptor

In order to choose one of the discriminative features from 14 Haralick feature descriptors, the following steps are performed on OU-ISIR D and CASIA B gait datasets. The steps are shown in Algorithm 4.

The algorithm begins by considering T number of subjects from a gait dataset (here T=30). Firstly, a set of 6 GGMI's is obtained for each subject from the CASIA B gait dataset. Hence, a total of $30 \times 6 = 180$ GGMI's are considered for the experiment. Secondly, 2 GGMI's are obtained for each subject from the OU-ISIR D gait dataset. Hence, a total of $30 \times 2 = 60$ GGMI's are considered for the experiment. These selected GGMI's are divided into 9 equal-sized cells. This is followed by the extraction of 14 Haralick texture descriptors (k=14), from each cell. Relief feature selection algorithm (Kira and Rendell 1992) is applied particularly to the 14 features extracted from a cell, followed by obtaining a weight W for every feature of that cell. This procedure is repeated on all the 9 cells, followed by obtaining each feature's weight for all the 9 cells separately. Finally, the mean of the weights obtained for a particular feature from all cells is computed. The next step is to sort the features by their weights and select the feature with the highest rank. The Relief feature selection method is applied to the extracted features in order to select the most relevant feature with minimum redundancy.

Algorithm 4 Steps to select the most discriminative feature from 14 Haralick texture descriptors.

1: **Input**: Number of subjects, $N = \{N_1, N_2, ..., N_T\}$
2: Begin
3: Let G be the total number of GGMI's obtained from all subjects of N.
4: Divide all GGMI's into R equal sized cells, $C = \{c_1, c_2, ..., c_R\}$
5: **for** all cells C of all GGMI's **do**
6:     **for** first cell c_1 of all GGMI's **do**
7:         Compute k Haralick texture features, $F = F_1, F_2, ..., F_k$.
8:     **end for**
9:     Obtain the feature vector of size $G \times k$.
10:     Apply Relief feature selection algorithm.
11:     Acquire for each feature F_j, a quality weight W within $-1 \leq W[j] \leq 1$.
12: **end for**
13: Compute mean of the weight W_j obtained from all cells C for feature F_j,
14: $M = \left\{ \frac{c_1 W_1 + c_2 W_1 + c_3 W_1 + ... + c_R W_1}{R}, ..., \frac{c_1 W_k + c_2 W_k + c_3 W_k + ... + c_R W_k}{R} \right\}$.
15: Select the top scoring feature from F.
16: End

Table 3.2 shows the ranking of features for both CASIA B and OU-ISIR D dataset. From Table 3.2, it is evident that the highest-ranking is obtained for sum variance feature for both datasets. This symbolizes that the most discriminative texture feature that can be extracted from each cell of GGMI is sum variance. Therefore, the proposed

Table 3.2: The mean weight M of 14 Haralick texture descriptors obtained by applying Algorithm 4 on CASIA B and OU-ISIR D dataset.

Feature	CASIA B	OU-ISIR D
Angular second moment	0.223269	0.211593
Contrast	0.208977	0.199095
Correlations	0.011042	0.011042
Variance	0.264309	0.268376
Inverse difference moment	0.000011	0.000016
Sum average	0.251951	0.258416
Sum variance	**0.298169**	**0.271197**
Sum entropy	0.248689	0.228606
Entropy	0.252788	0.2541532
Difference variance	0.199677	0.194528
Difference entropy	0.181352	0.192989
Information measure 1	0.197986	0.194256
Information measure 2	0.209726	0.198656
Maximal correlation coefficient	0.213255	0.252483

method uses the sum variance feature descriptor. It is found that the descriptor f_{sv} varies slightly for GGMI's of the same subject and largely for the GGMI's of different subjects.

In the above experiment, the texture features are extracted from each cell separately instead of the entire GGMI. This is because, in the proposed work, the Haralick feature is extracted from each cell separately. Hence, it is necessary to find out that in each cell, where the textures are different, sum variance is the most discriminative feature.

Even though HOG is used for gait recognition in some of the recent works (Arora et al. 2015b; Hofmann and Rigoll 2013), the size of the feature vector used in those works is large (5400×1 and so on). Hence, it is necessary to come up with a method which provides more accuracy with less number of HOG features. This has been achieved by multiplying the sum variance (f_{sv}) Haralick texture descriptor value of a cell with the HOG features of that cell. Multiplying f_{sv} with HOG features increases the recognition accuracy due to the reasons mentioned in Table 3.3.

3.1.6.2 Experimental Setup

All GEI and GGMI templates used in this work are of size 240×240. The GGMI template is divided into $R = 9$ cells with 50% overlap. Hence, the size of each cell is 120×120. From each cell of size 120×120, a set of 7 values is obtained, as the total

number of bins (BI) is 7. At last, the values of 7 bins of 9 cells are concatenated to get 63×1 dimensional feature vector f.

In the HOG descriptor, the features are computed from the distribution of directions of gradients. The gradients of an image are valuable, given that the magnitudes of gradients are large around corners and edges. These corners and edges contain more information about the shape of the subject than flat regions. Hence, the use of GGMI, leads to the prominant edges thus enabling the increase in the performance of a gait detection system.

Table 3.3: Consequences of multiplying f_{sv} with HOG features.

Cases	Consequence
When two identical cells from images of two different subjects have the same f_{sv}	Doesn't influence the recognition accuracy because the difference in the values of the HOG feature vector of two subjects will remain unaltered.
When two identical cells from images of two different subjects have different f_{sv}, this occurrence leads to the distinct HOG feature vector	Increases the recognition accuracy.
When two identical cells from images of two different subjects have different f_{sv}, this situation leads to the same HOG feature vector (very rare case)	Doesn't affect the gait recognition to a large extent because this feature vector is obtained for only one cell. The final feature vector is the concatenation of values from 9 different cells.

Most of the existing works have considered 9 bins for each cell, whereas, in this study, a varied number of bins (BI) is considered in order to obtain high recognition accuracy with less number of features. So, the study is experimented with 4, 5, 6, 7, 8, and 9 bins for each cell. As a result, the study shows good recognition accuracy for 7 bins along with a marginal improvement for 8 and 9 bins. Hence, 7 bins are fixed for each cell.

The range of 9 orientation bins that is considered in the existing works on gait recognition is from 0^0 to 180^0, with 20^0 assigned to each bin (i.e., 0^0 to 20^0, 20^0 to 40^0, 40^0 to 60^0, 60^0 to 80^0, 80^0 to 100^0, 100^0 to 120^0, 120^0 to 140^0, 140^0 to 160^0, 160^0 to 180^0). The range of the angles is from 0^0 to 180^0 instead of 0^0 to 360^0. These gradients are called unsigned gradients because a gradient and its negative are represented by the same number. But, in this study, a range of 7 orientation bins is considered from -180^0 to 180^0, with a range of 51.43^0 assigned to each bin (i.e., -180^0 to -128.58^0,
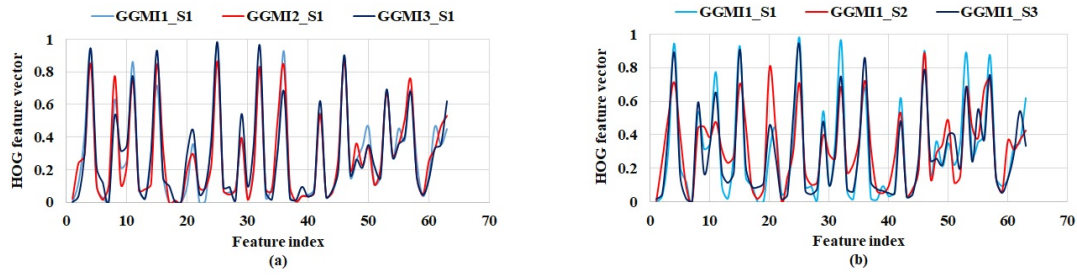
Figure 3.5: Feature space diagram representing HOG feature vector (size: 63×1) of three (a) GGMI's belonging to same subject (b) GGMI's belonging to different subjects.

-128.58^0 to -77.16^0, -77.16^0 to -25.74^0, -25.74^0 to 25.74^0, 25.74^0 to 77.16^0, 77.16^0 to 128.58^0, 128.58^0 to 180^0). In this experiment, the gradient and its negative are not considered as the same number. This is because the experiment was conducted with both signed and unsigned gradients. Empirically it has been discovered that signed gradients work better than unsigned gradients for GGMI.

The correlation of the feature vector extracted by using HOG and the proposed method is illustrated in the feature space diagram. Figure 3.5(a) shows the HOG features extracted from the same subject for three different GGMI's (GGMI1_S1, GGMI2_S1, GGMI3_S1). The HOG features extracted from three different subjects for a single GGMI (GGMI1_S1, GGMI1_S2, GGMI1_S3) is shown in Figure 3.5(b). Figure 3.6(a) shows the proposed gait features extracted from the same subject for three different GGMI's and Figure 3.6(b) shows the proposed features extracted from three different subjects for a single GGMI. It is evident from the feature space diagram of the proposed features that very minute variation is formed within three gait cycles for the same subject, but a significant variation is produced for the gait cycle of three different subjects. The separation of the curves in Figure 3.6(b) when compared to Figure 3.5(b), indicates a considerable distinction between the inter-class distances. Simultaneously, the small intra-class variances in Figure 3.6(a) when compared to 3.5(a) demonstrates the discriminative ability of feature under consideration. The texture is different for each cell of GGMI; consequently, the values of sum variance are also different. Hence, the shape of the graph changes in Figure 3.6(a) and 3.6(b).

This study also assesses the impact of the dimensionality reduction on the classification accuracy of the extracted features. It is observed that PCA considerably increases the recognition accuracy of the proposed method and reduces the storage requirements.

46
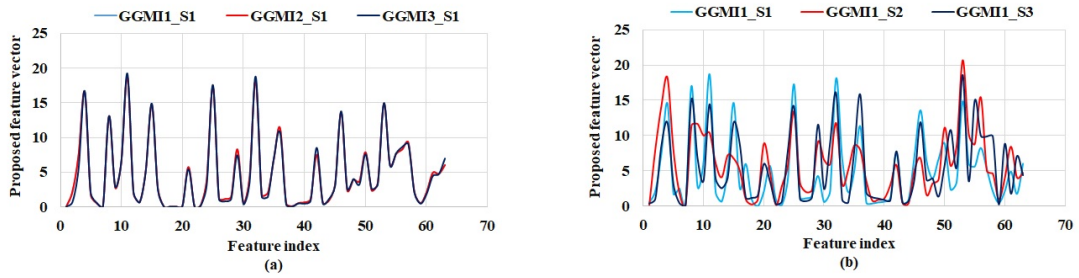
Figure 3.6: Feature space diagram representing proposed feature vector (size: 63×1) of three (a) GGMI's belonging to same subject (b) GGMI's belonging to different subjects.

Table 3.4: The recognition rates of the proposed method on CASIA A dataset.

Method	0^0	45^0	90^0	Average
GHEI-HOG (Hofmann and Rigoll 2012)	93.24	94.24	95.37	94.28
α-GHEI (Hofmann and Rigoll 2013)	95.32	94.71	97.67	95.90
GGI-HOG (Arora et al. 2015b)	**96.23**	97.37	98.25	97.28
Proposed method (P 3.1)	95.25	**98.71**	**99.24**	**97.73**

The classifier used for gait recognition is KNN. In this study, the number of neighbors is equal to 1, and Euclidean is considered as a distance metric. The CCR is used to measures the performance of the gait recognition system on the testing dataset.

3.1.6.3 Experimental Results and Discussion

The performance of the proposed method is evaluated on five widely used gait datasets. They are as follows: CASIA A, CASIA B, OU-ISIR D, CMU MoBo, and KTH. Each dataset consists of gait sequences collected in several environments with variations in view, walking speed and walking style. The term P 3.1 is used to specify the proposed method in comparison tables.

1. Experiments on CASIA A dataset

A brief description of the CASIA A dataset is given in Section 2.5. Out of the four sequences, 3 sequences are considered for training, and one sequence for testing. The experimental results on CASIA A dataset and their comparison with those of the existing methods in the literature are shown in Table 3.4. This study shows that the computation of the proposed features significantly improves the performance of the gait recognition method for 45^0 and 90^0 views. The results show that the CCR of this study is slightly less for 0^0 view, as the spatial and dynamic information obtained for the frontal view is less in comparison with the other view variations.

Table 3.5: Recognition accuracy (%) of the proposed method on CASIA B dataset with prevailing methods.

Method	0^0	18^0	36^0	54^0	72^0	90^0
Yu et al. (2006)	99.20	**99.60**	97.60	97.20	97.20	97.60
Kusakunniran (2014b)	91.10	91.00	92.30	95.30	95.70	95.40
Kusakunniran (2014a)	97.60	92.50	96.80	96.80	97.00	97.30
Rida et al. (2016a)	**100**	98.79	95.56	94.35	93.15	93.60
Zeng and Wang (2016)	55.40	44.20	66.70	77.80	77.80	87.90
Proposed method (P 3.1)	99.00	99.00	**98.56**	**99.40**	**99.15**	**99.00**

Method	108^0	126^0	144^0	162^0	180^0
Yu et al. (2006)	95.60	96.80	96.40	98.40	**99.60**
Kusakunniran (2014b)	94.10	95.20	94.80	94.20	94.10
Kusakunniran (2014a)	95.70	95.40	94.90	97.60	98.40
Rida et al. (2016a)	93.55	95.56	94.76	97.18	98.79
Zeng and Wang (2016)	66.70	77.00	75.80	76.70	57.90
Proposed method (P 3.1)	**98.80**	**99.20**	**99.00**	**98.60**	98.10

2. Experiments on CASIA B dataset

A brief description of the CASIA B dataset is presented in Section 2.5. The two experiments carried out on CASIA B gait dataset are as follows:

I. View invariant gait recognition

The features are extracted for all 11 views from 124 subjects of the dataset. For each subject, the features are extracted from 6 gait cycles. Hence, the number of features are extracted from $124 \times 6 = 744$ GGMI's per view. The classification accuracy is obtained separately for each view. Among 6 gait sequences, the first four sequences of *nm* are used for training, and the remaining two sequences of *nm* are considered for testing.

The results obtained for normal walking under all views are demonstrated in Table 3.5. The proposed technique is also compared with the existing methods in the literature under various views. Table 3.5 demonstrates the same. The work carried out by Yu et al. (2006) outperforms the proposed method in three instances of normal walking conditions (i.e., 0^0, 18^0 and 180^0). In the case of view variations, the proposed technique performs better for most views yet not for frontal view (Yu et al. 2006). Thus, any changes in the view also modify the visual characteristics such as data related to shape and motion. The comparison of the results for the lateral view is shown in Table 3.6. It can be inferred from the results that the performance of the proposed

Table 3.6: Comparison of CCR (%) on CASIA B gait dataset for 90^0 view.

Method	CCR
LBP-TOP (Kellokumpu et al. 2009)	46.10
PSA (Wang et al. 2003)	61.80
FD (Lee et al. 2013)	61.80
GEI-HOG (Whytock et al. 2012)	80.60
MHI-HOG (Huang et al. 2011)	75.00
GHEI-HOG (Hofmann and Rigoll 2012)	73.00
α-GHEI (Hofmann and Rigoll 2013)	75.10
GGI-HOG (Arora et al. 2015b)	80.80
MHI-BSIF-HOG (Mogan et al. 2017)	93.42
Proposed method (P 3.1)	**99.00**

approach outperforms that of other methods in most cases.

II. Cross view gait recognition

In this section, the experiment is performed for probe view (90^0) and gallery views (0^0, 18^0, 36^0, 54^0, 72^0, 108^0, 126^0, 144^0, 162^0 and 180^0). The results are shown in Table 3.7. The performance of this gait recognition method is compared with that of five other approaches, including a method based on deep learning. The other methods and the proposed method use 100 subjects for evaluating the recognition performance. It is evident from the results that the cross-view gait recognition among two closer views accomplishes better recognition accuracy as the gaits share more information when their views are closer. The proposed method performs better in most cases, leading to higher mean accuracy in comparison to the other methods. This implies that the structural information from GGMI, HOG, and statistical texture information from f_{sv} together contribute in increasing the performance of the gait recognition system.

3. Experiments on OU-ISIR D dataset

A brief description of the OU-ISIR D dataset is given in Section 2.5. This dataset consists of two sequences of each individual. Among these two sequences, one sequence is used for training, and the other sequence is used for testing. For DB_{high} and DB_{low}, 100 sequences are used for training and 100 sequences for testing.

Table 3.8 displays the evaluation of the proposed approach with other methods in literature such as LBP-TOP, PSA, FD, GEI-HOG, MHI-HOG, GHEI-HOG, α-GHEI, GGI-HOG, and MHI-BSIF-HOG. It can be observed from the table that the proposed

Table 3.7: Recognition accuracy (%) on CASIA B gait dataset, where the probe view is 90^0 and gallery views are 0^0, 18^0, 36^0, 54^0, 72^0, 108^0, 126^0, 144^0, 162^0 and 180^0.

Method	0^0	18^0	36^0	54^0	72^0	108^0	126^0	144^0	162^0	180^0	Average
GEI-SVR (Kusakunniran et al. 2010)	16	22	35	63	95	95	65	38	20	13	42.00
CMCC (Kusakunniran et al. 2013)	18	24	41	66	96	95	68	41	21	13	43.91
ViDP (Hu et al. 2013a)	8	12	45	80	**100**	**100**	81	50	15	8	45.36
LB deep network (Wu et al. 2017)	18	36	67.5	93	99.5	99.5	92	66	36	18	56.86
AD (Chen et al. 2018)	38	**75**	68	93	98	99	93	67	**76**	39	67.82
Proposed method (P 3.1)	**42**	69	**75**	**95**	92	99	**95**	76	68	**42**	**75.30**

Table 3.8: Results and comparison of gait recognition performance (%) on OU-ISIR D gait dataset.

Method	DB_{high}	DB_{low}
LBP-TOP (Kellokumpu et al. 2009)	49.00	50.00
PSA (Wang et al. 2003)	87.00	83.00
FD (Lee et al. 2013)	92.00	91.00
GEI-HOG (Whytock et al. 2012)	95.00	95.00
MHI-HOG (Huang et al. 2011)	90.00	90.00
GHEI-HOG (Hofmann and Rigoll 2012)	93.00	94.00
α-GHEI (Hofmann and Rigoll 2013)	95.00	95.00
GGI-HOG (Arora et al. 2015b)	98.00	95.00
MHI-BSIF-HOG (Mogan et al. 2017)	96.00	**100.0**
Proposed method (P 3.1)	**99.00**	**100.0**

method provides good recognition accuracy when compared to other methods as it extracts both gradient and texture information from the GGMI's. The experimental results show that walking speed fluctuations, which is a major factor that affects gait recognition, can be managed by the proposed method.

4. Experiments on CMU MoBo dataset

Note that 25 subjects with fast walking speed of 4.5 km/h and slow walking speed of 3.3 km/h are used in this work to measure the gait recognition performance.

In this experiment, the proposed method is evaluated under diverse walking speeds. That is, the slow walk videos are used as a probe set, and the fast walk videos are used as a gallery set. The results obtained from the proposed method and its comparison with the other methods, where a few techniques are based on HOG, are demonstrated in Table 3.9. From Table 3.9, it is evident that the same recognition accuracy is obtained for methods such as GEI-HOG, MHI-HOG, and MHI-BSIF-HOG, and the proposed method gives the best recognition accuracy in comparison to all other methods from the past literature.

50

Table 3.9: Results and comparison of gait recognition performance (%) on CMU MoBo dataset.

Method	CCR
LBP-TOP (Kellokumpu et al. 2009)	12.00
PSA (Wang et al. 2003)	36.00
FD (Lee et al. 2013)	60.00
GEI-HOG (Whytock et al. 2012)	76.00
MHI-HOG (Huang et al. 2011)	76.00
GHEI-HOG (Hofmann and Rigoll 2012)	73.00
α-GHEI (Hofmann and Rigoll 2013)	78.00
GGI-HOG (Arora et al. 2015b)	75.00
MHI-BSIF-HOG (Mogan et al. 2017)	76.00
Proposed method (P 3.1)	**82.00**

5. Experiments on KTH video dataset

A brief description of the KTH dataset is presented in Section 2.5. The experiments conducted on the KTH running and jogging dataset are as follows.

All the four scenarios *sn1*, *sn2*, *sn3* and *sn4* are considered for the experiment. For each subject, 6 gait cycles, in turn, 6 GGMI's are considered for the experimental analysis. Hence, the gait data consists of 150 GGMI's. Accordingly, 150 feature vectors of 25 subjects for each scenario (*sn1*, *sn2*, *sn3*, *sn4*) are used for classification. Among the 6 GGMI's, four GGMI's are considered for training, and two GGMI's are considered for testing. The experimental results shown in Tables 3.10 and 3.11 illustrate that the proposed method performs better than the other methods suggested in the literature in the context of HOG.

From the literature, it is evident that the covariate that leads to lesser CCR, is clothing condition. As the clothing condition causes major changes in the subject's silhouettes, it decreases the recognition accuracy because the appearance-based methods depend on spatial and temporal variations of the silhouettes over a gait cycle. It is obvious from Tables 3.10 and 3.11 that the maximum CCR is obtained for a normal condition (*sn1* and *sn4*), and noticeably lesser CCR is obtained for clothing condition *sn3*. This shows that the clothing style contributes to the identification of the individuals. The highest gait information is always acquired for the view angle near 90^0, and it gradually decreases for the other views. Hence, the CCR of scenarios *sn1* and *sn4* is more than that of scenario *sn2*. The results on the KTH dataset proves that the proposed method can yield better recognition accuracy for other types of gaits such as running

Table 3.10: Recognition accuracy (%) of the proposed method on KTH jogging dataset for four scenarios.

Method	*sn1*	*sn2*	*sn3*	*sn4*
GHEI-HOG (Hofmann and Rigoll 2012)	86.77	80.00	83.06	86.00
α-GHEI (Hofmann and Rigoll 2013)	90.00	89.71	90.00	96.00
GGI-HOG (Arora et al. 2015b)	92.53	87.47	88.00	96.94
Proposed method (P 3.1)	**95.94**	**91.00**	**93.77**	**98.00**

Table 3.11: Recognition accuracy (%) of the proposed method on KTH running dataset for four scenarios.

Method	*sn1*	*sn2*	*sn3*	*sn4*
GHEI-HOG (Hofmann and Rigoll 2012)	87.06	77.00	80.00	88.12
α-GHEI (Hofmann and Rigoll 2013)	90.00	84.53	88.94	89.47
GGI-HOG (Arora et al. 2015b)	94.00	**87.88**	89.24	95.00
Proposed method (P 3.1)	**94.00**	85.33	**92.00**	**96.00**

and jogging. A comparative analysis with regard to the other approaches based on HOG suggests that the proposed approach yields better results than those of the others.

3.2 Gait Recognition based on Gaussian Filtered Gait Energy Template and Centroid Corner Distance Features

The proposed work aims to present a robust method for the appearance-based human gait identification. The role of this examination is to mould the boundary shape of Gaussian Filtered-Gait Energy Image (GF-GEI) into six separate horizontal segments. The distance-based features are used to extract the dynamic and static factors from every segment of the shape border. The proposed method uses statistical features that depict the shape of the contour, which, in turn, is valuable in a gait detection system.

3.2.1 Framework of the Proposed Method

Figure 3.7 gives a brief description of the proposed approach. Initially, a sequence of images is obtained from the gait video. These sequences are further processed to obtain a sequence of gait silhouette images. Eventually, the gait features are extracted and classified using the KNN classifier.
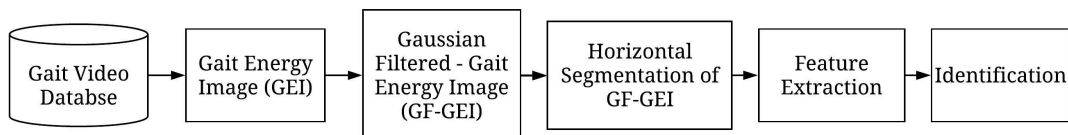
Figure 3.7: Steps of gait recognition framework.
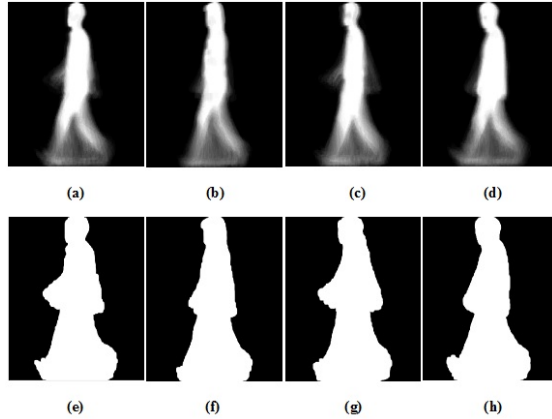
Figure 3.8: Sample images representing GEI ((a) to (d)) and their corresponding GF-GEI ((e) to (h)) from CASIA B gait dataset.

The process of feature extraction involves the following steps: (1) GEI of a gait cycle is extracted as shown in Figure 3.7, (2) GEI is converted into GF-GEI, and (3) GF-GEI is horizontally segmented, followed by the extraction of centroid corner distance features.

Algorithm 5 Generation of GF-GEI.

1: **Input**: *GEI*, of size $p \times q$.
2: **Output**: *GF-GEI*, of size $p \times q$.
3: **Begin**
4: Read an input image.
5: Convert the image to double precision.
6: Gaussian filter of size 5×5 with $\sigma = 1.0$ is given by

$$G = \begin{bmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{bmatrix}$$

7: Perform convolution of GEI with filter G.
8: **End**

3.2.2 Gaussian Filtered - Gait Energy Image

In this study, GF-GEI template is obtained from GEI for feature extraction. After the extraction of GEI from a gait cycle, it is further converted into GF-GEI by Algorithm 5. The convolution operation is performed on GEI, using a Gaussian filter. This process eliminates the details in the arm swing and leg swing regions of the GEI, and thus a uniform structure, as shown in Figure 3.8, is obtained. Both GEI and GF-GEI images

for normal walking from the CASIA B gait dataset are shown in Figure 3.8.

3.2.3 Centroid Corner Distance Features

The selection of most discriminative features is a difficult task with regard to the individual identification problems. In this work, the statistical shape analysis of GF-GEI is done in the horizontal mode, to extract identically distributed and independent features from each segment.

GF-GEI is horizontally segmented into S equal size segments. Here, S is a user-defined input signifying the number of segments (in this case, $S = 6$). The features are computed for each horizontal segment H_i, where $i = 1,....,6$. The shape centroid for each H_i is calculated from $C_{(x,y)} = (\frac{1}{Z} \sum_{i=1}^{Z} x_i, \frac{1}{Z} \sum_{i=1}^{Z} y_i)$, where Z is the total number of pixel values and (x_i, y_i) are the pixels of the shape (Wang et al. 2002).

Algorithm 6 Horizontal segmentation for feature extraction.

1: **Input**: *GF-GEI*, of size $p \times q$.
2: **Output**: Horizontal segments S, of equal size.
3: **Begin**
4: Read an input image.
5: Define the row value for horizontal segmentation, $H\_segment$.
6: Let the total number of rows is given by row.
7: Let the total number of columns is given by $column$.
8: **for** $i = 1 : row/H\_segment$ **do**
9:     *S(i) = GF-GEI((H_segment×(i-1)+1:H_segment×(i-1)+H_segment),1:column)*
10: **end for**
11: **End**

A corner is a point for which there are two different and dominant edge directions in its local neighbourhood. The minimum eigenvalue algorithm proposed by Shi and Tomasi (1993) is applied to each segment H_i. This process detects the corners along the contour of the GF-GEI. The centroid corner distance features represent the distance between the centroid and corner of the same GF-GEI segment. Consider a horizontal segment H_i, let $C_i(a_i, b_i)$ be the centroid and $V_j(x_j, y_j)$ be the corner present at a distance on the contour of the image. The Euclidean distance is calculated using

$$d(C_i, V_j) = \sqrt{(a_i - x_j)^2 + (b_i - y_j)^2} \qquad (3.8)$$

54
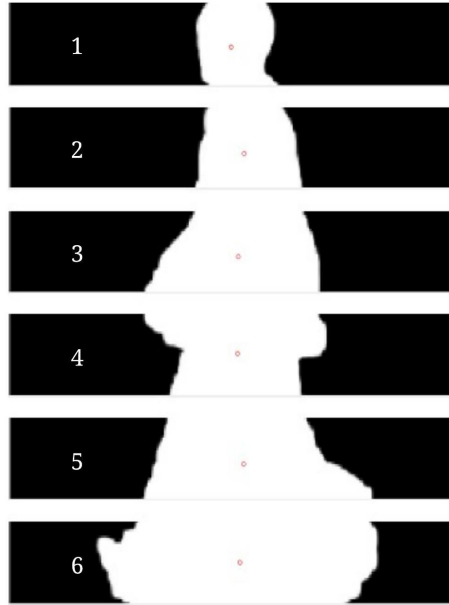
Figure 3.9: Sample image representing the centroid of 6 horizontal GF-GEI segments.
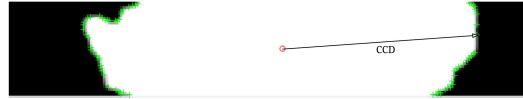
Figure 3.10: Image representing Centroid Corner Distance (CCD) in Segment 6.

In horizontal segmentation, certain rows need to be segmented from GF-GEI, as shown in Figure 3.9, whereas the number of columns remains constant. Hence, there is a need to define the row value for horizontal segmentation, specified as $H\_segment$ in Algorithm 6. In this Algorithm, the term row represents the total number of rows present in GF-GEI, and the number of rows specified in $H\_segment$ is segmented to form horizontal segments, given by $S(i)$. By executing $S(i)$, in each iteration, the number of rows specified in $H\_segment$ are separated from GE-GEI in a consecutive order.

The steps used for horizontal segmentation and feature extraction are shown in Algorithms 6 and 7. The 6 regions of GF-GEI with centroid are shown in Figure 3.9. The centroid corner distance feature is illustrated in Figure 3.10. The concatenation of the distance features from all the horizontal segments forms the final feature vector.

Algorithm 7 Feature extraction.

1: **Input**: Horizontal segments S of equal size.
2: **Output**: Feature vector f.
3: **Begin**
4: Read all horizontal segments H_i of an input image. Here $i = (1,....,6)$.
5: **for** $i = 1 : S$ **do**
6:     **for** each H_i **do**
7:         Calculate the centroid value of the segment.
8:         Calculate the corner points of the segment using minimum eigenvalue algorithm
9:         Compute the Euclidean distance between the centroid and each of the corner points of the segment.
10:         This forms the feature vector f_i, of segment H_i.
11:     **end for**
12: **end for**
13: Concatenate the feature vectors f_i ($i = 1,....,6$) to obtain f.
14: **End**

3.2.4 Experiments

3.2.4.1 Experimental Setup

When the centroid of GEI is calculated, the minimum of 3 centroids for each image is obtained due to the distribution of intensity values in it. When the minimum eigen value algorithm (Shi and Tomasi 1993) is applied on GEI, the corners can be detected randomly in the lower region of GEI. They are not along the contour of GEI. On computing the distance between the weighted centroid and these corners and when experimented with 20 subjects of CASIA A dataset, the features are not found to be discriminative. Hence, there is a need to convert GEI into a suitable form so that the features having more discriminative power are extracted. The Gaussian filter is used to convert GEI into GF-GEI, as the aim is to obtain a uniform texture, unlike GEI in all regions of the gait template. Hence, GF-GEI is computed, where only one centroid is obtained per segment, and the corners detected are along the contour of the image for all horizontal segments, and distances between the centroid and corners are calculated. Initially, the experiment is conducted on 20 subjects from CASIA B and OU-ISIR D gait dataset by considering the whole GF-GEI and by dividing GF-GEI into 3, 4, 5, and 6 segments. The encouraging results are obtained when it is divided into 6 segments. Hence, GF-GEI is divided into 6 horizontal segments in this study.

Table 3.12: Summary of the corners detected from GF-GEI horizontal segments.

Horizontal block	Min no. of corners	Max no. of corners	Total no. of corners selected
1	42	45	40
2	46	48	45
3	53	57	50
4	65	66	65
5	71	74	70
6	73	77	70

Table 3.13: CCR (%) of the proposed method on CASIA A dataset.

Methodology	0^0	45^0	90^0	Average
Kumar and Nagendraswamy (2014b)	94.37	95.24	93.24	94.28
Luo et al. (2015)	97.25	95.23	92.67	95.05
Tafazzoli et al. (2015)	96.56	94.71	91.32	94.19
Zhang and Zhang (2015)	98.21	96.25	92.38	95.61
Zhang and Zhang (2018)	**99.61**	97.57	97.16	98.11
Proposed method (P 3.2)	98.25	**98.71**	**99.24**	**98.73**

The number of corners detected for each segment is different based on the length of its contour. However, from the observation, it is noted that the number of corners detected for the same segment of different GF-GEI's varies within a very small range. For classification purposes, the same number of feature values needs to be picked from each block for all GEI's in the dataset. Hence, a minimum and a maximum number of corners are detected from each segment of all images in the dataset. In order to choose an ideal number of corners to be picked from each segment, the total number of corners, which is less than the minimum value of each segment is selected. This is done based on the strength of the corner. When each corner is detected, a metric will be associated with it, which in turn represents the strength of the corner. Hence, the metric is used to pick the required number of corners from all the corners detected. The corners with the highest strength are selected. Thus, the number of features for 6 segments is given by $40 + 45 + 50 + 65 + 70 + 70 = 340$. Hence, a feature vector of size 340×1 is obtained from every GF-GEI image for classification. The number of corners selected for each horizontal segment is shown in Table 3.12.

GEI used in this work from CASIA A, CASIA B, and OU-ISIR D gait dataset is of size 240×240. It is then converted into GF-GEI of size 252×252. Each horizontal block of GF-GEI is of the size 42×252. In this experimentation, the classification of the gait data is done by using KNN classifier. Here, the value of K is equal to 1. The CCR is used to asses the performance of the proposed method.
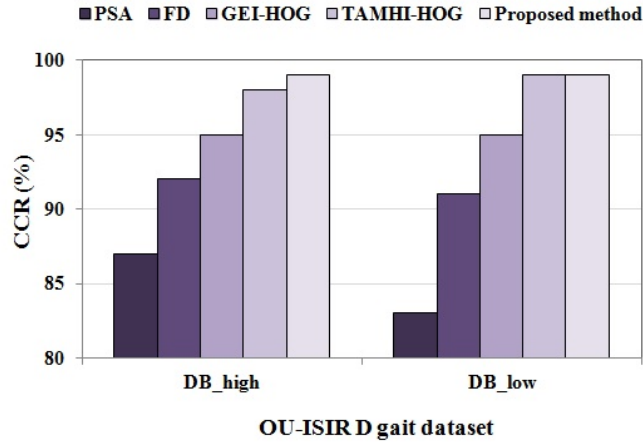
Figure 3.11: Comparision of the CCR (%) of the proposed method on OU-ISIR D gait dataset with some of the existing methods in the literature.

3.2.4.2 Experimental Results and Discussion

The experimental results obtained provide an insight into the performance of the proposed method. In this study, the investigations are carried out on three widely used gait datasets to evaluate the recognition accuracy of the proposed strategy. They are (1) CASIA A dataset, (2) CASIA B dataset, and (3) OU-ISIR D dataset. The term P 3.2 is used to specify the proposed method in comparison tables.

1. Experiments on CASIA A dataset

In this dataset, each of the 20 subjects has four gait sequences for each viewing angle. Among these four sequences, three are considered for training, and the remaining for testing. The obtained experimental results on CASIA A dataset and the comparison of them with the already existing methods in the literature are shown in Table 3.13. It can be seen that the computation of the proposed features considerably increases the performance of the gait detection method for 45^0 and 90^0 views. The results show that the CCR of the proposed method is slightly less for 0^0 view.

2. Experiments on CASIA B dataset

Here, the experiments are conducted for normal walking under all view angles. The feature vectors are extracted from $124 \times 6 = 744$ GEI images per view for *nm*. Among six *nm* gait sequences, four gait sequences are used for training, and the remaining two for testing. The obtained experimental results using the proposed method on the CASIA B dataset and the comparison of them with the existing methods are shown in

Table 3.14: CCR (%) of the proposed method on CASIA B dataset.

Method	0^0	18^0	36^0	54^0	72^0	90^0
Yu et al. (2006)	99.20	**99.60**	97.60	97.20	97.20	97.60
Kusakunniran (2014b)	91.10	91.00	92.30	95.30	95.70	95.40
Kusakunniran (2014a)	97.60	92.50	96.80	96.80	97.00	97.30
Rida et al. (2016a)	**100.0**	98.79	95.56	94.35	93.15	93.60
Zeng and Wang (2016)	55.40	44.20	66.70	77.80	77.80	87.90
Proposed method (P 3.2)	99.40	**99.60**	98.80	99.80	99.40	99.60

Method	108^0	126^0	144^0	162^0	180^0
Yu et al. (2006)	95.60	96.80	96.40	98.40	**99.60**
Kusakunniran (2014b)	94.10	95.20	94.80	94.20	94.10
Kusakunniran (2014a)	95.70	95.40	94.90	97.60	98.40
Rida et al. (2016a)	93.55	95.56	94.76	97.18	98.79
Zeng and Wang (2016)	66.70	77.00	75.80	76.70	57.90
Proposed method (P 3.2)	**99.00**	**99.40**	**98.80**	**99.00**	99.30

Table 3.14. The results demonstrate that the proposed approach increases the performance of all views under normal walking conditions except for 0^0 and 180^0. Generally, the maximum gait data is obtained for the view angle close to 90^0, and the minimum gait data is acquired for the view angle close to 0^0 or 180^0. This method is based on the contour and centroid of the image, the contour of the subject at 0^0 and 180^0 will carry the minimum information compared to all other views. Hence, the CCR obtained is less for 0^0 in CASIA A and also less for 0^0, 180^0 in the CASIA B dataset.

3. Experiments on OU-ISIR D dataset

This dataset contains two gait sequences of each subject. Among these two gait sequences, one sequence is used for training, and the other sequence is used for testing. For DB_{high} and DB_{low}, 100 sequences are used for training, and 100 sequences are used for testing. The performance of the proposed method on the OU-ISIR D gait dataset is reported in Figure 3.11. It shows the assessment of the proposed approach with different approaches in literature such as PSA (Wang et al. 2003), GEI-HOG (Whytock et al. 2012), FD (Lee et al. 2013) and TAMHI-HOG (Lee et al. 2014a). From Figure 3.11, it can be observed from the results that the proposed approach offers good CCR with regard to the methods compared.

3.3 Summary

Section 3.1 presents an efficient method to increase the performance of the gait recognition system. The gait features are extracted by applying HOG global descriptor and sum variance Haralick texture descriptor to all the cells of GGMI. The numerical value

of the texture feature of a cell is multiplied with the HOG features of the same cell. Consecutively, the features from all the nine cells of GGMI are concatenated to obtain the final low-dimensional feature vector.

The performance analysis of the proposed approach is done by using five widely-used gait datasets. The experimental outcomes show the viability of the proposed technique among a wide range of views in the CASIA A and CASIA B gait dataset. The experimental results on OU-ISIR D and CMU MoBo dataset show that it performs well in case of fluctuations and speed variations. The obtained experimental results on the KTH dataset shows that it can yield better recognition accuracy for other types of gaits such as running and jogging. When compared with the other approaches based on HOG in literature, the proposed approach gives promising results in most of the cases.

Section 3.2 proposes a robust technique for gait detection. The significant contribution of this approach is in the formulation of the features that can minimize the misclassification errors to a small margin. The distances between the centroid and corners of the segmented GF-GEI image are computed in the proposed method, which are called the centroid corner distance features. This depiction captures the fine details of the shape of the subjects and, therefore, accomplishes better identification rates. A thorough assessment of the performance of the proposed approach is done by using CASIA A, CASIA B, and OU-ISIR D gait datasets. The results indicate that the proposed approach based on significant features improves the performance of the gait recognition.

Chapter 4

CLOTHING INVARIANT GAIT RECOGNITION

The chief covariate condition comprises the distinction of clothing style between probe and gallery datasets. It affects the recognition of subjects related to the changes in the clothing styles, such as a half-shirt, long coat, skirt, regular pants, coats, and so on (Nandy et al. 2016). Consequently, the gait silhouette of an individual changes with the influence of clothing. As a result, it is necessary to choose more distinguishing gait characteristics that provide an essential insight into the features of gait biometrics. These features can be shown as a clothing invariant despite the differences in the individual's appearance.

4.1 Clothing Invariant Gait Recognition using Modified Local Optimal Oriented Pattern Binary Descriptor

This work intends to develop a feature extraction method that facilitates the extraction of clothing invariant gait features for gait recognition. The key contributions of this work are as follows:

- A texture descriptor called MLOOP is introduced, which is an extension of the Local Optimal Oriented Pattern (LOOP) descriptor.

- Two different feature vectors called histogram, and horizontal width vector are extracted from MLOOP descriptor.

- The obtained experimental results on standard OU-ISIR B and CASIA B datasets are shown to demonstrate that the extracted features are robust in spite of the altering appearance with different clothing variations.

4.1.1 Framework of the Proposed Method

The proposed approach is built up on the texture descriptor of GEI. It makes use of the local distinguishing features from the regions in the GEI, which allows the enhancement of the gait recognition performance. A brief description of the proposed method is presented in Figure 4.1.

Initially, the gait video is converted into the sequence of images. These images are processed to get silhouette images. The GEI is obtained by combining the silhouette images over a gait cycle. Later, the two ROIs, which are considered to be less
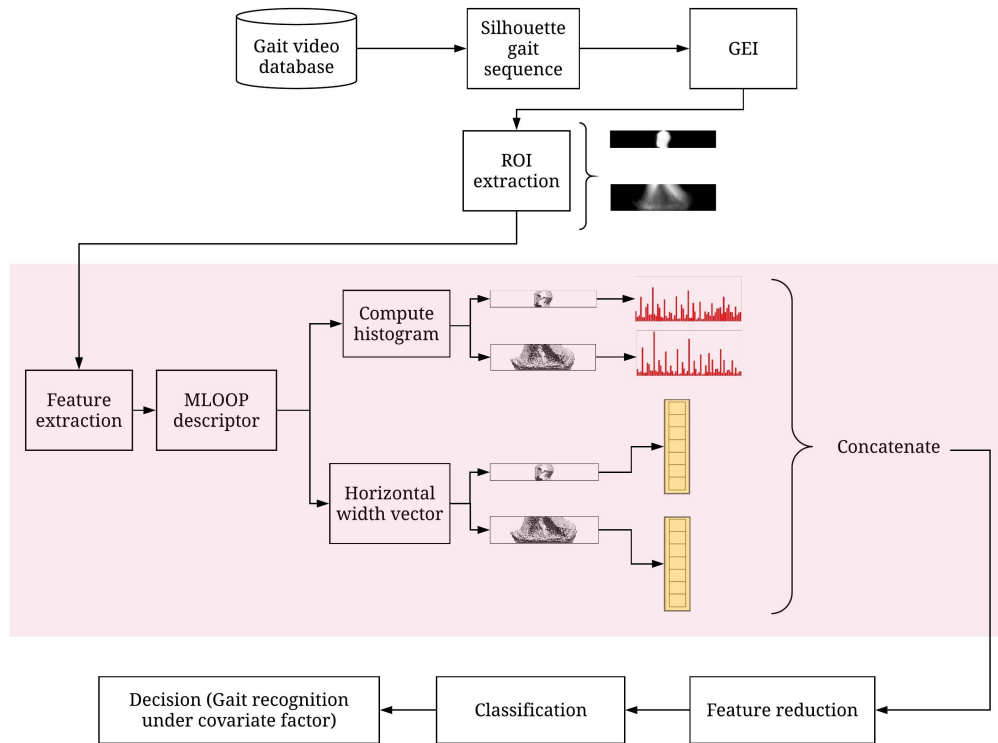
Figure 4.1: Structure of the proposed approach

affected by the clothing variations, are selected. The features are extracted from these two ROIs. The sample image of the two ROIs extracted is shown in Figure 4.2. Eventually, MLOOP descriptor is used to extract features from the ROIs of an individual. Since the GEI data includes a set of gait images, its texture features are very abundant. MLOOP is used in extracting the texture features of this image.

Furthermore, the histogram and horizontal width vector of the extracted MLOOP descriptor pixel values are computed. In this way, the histogram and horizontal width vector of the MLOOP descriptor obtained from the ROIs are used to represent the fea-
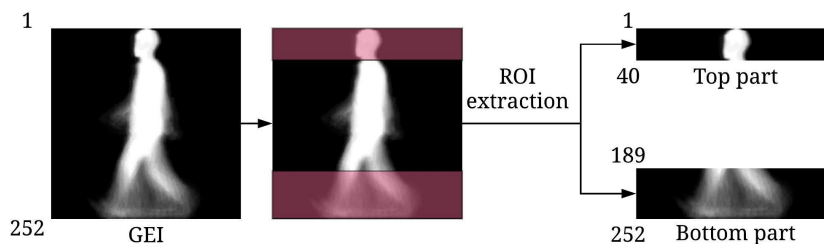
Figure 4.2: Sample image represents the ROIs extracted from the GEI of an individual, obtained from CASIA B gait dataset.

Algorithm 8 Feature Extraction

1: **Input**: GEI of a subject: $GEI(x, y)$, of size $p \times q$.
2: **Output**: Feature vector f.
3: Begin
4: Read an input image.
5: Extract two ROIs from GEI: $GEI_1(x, y)$ and $GEI_2(x, y)$
6: Compute MLOOP descriptor.
7: **Switch** proposed feature type
8: **Case** MLOOP - histogram
9: **for** $GEI_i, i = 1, 2$ **do**
10:     Compute histogram: fh_i.
11: **end for**
12: **Case** MLOOP - horizontal width vector
13: **for** $GEI_i, i = 1, 2$ **do**
14:     Compute width vector: fv_i.
15: **end for**
16: **End Switch**
17: Generate a concatenated feature vector: $f = \{fh_1, fh_2, fv_1, fv_2\}$
18: End
---

ture vector of each GEI. The process followed to extract the gait features is shown in Algorithm 8. This is followed by feature reduction, where the dimensionality of the extracted feature set is reduced by discarding the most irrelevant features. The features are then classified by using KNN classifier thus enabling the gait recognition.

4.1.2 Modified Local Optimal Oriented Pattern

Local binary descriptors are the efficient encoders of the recurring local patterns for good discrimination in a lot of visual recognition techniques. Numerous intensity refinements of these descriptors have been developed. MLOOP is a nonlinear amalgamation of Local Binary Pattern (LBP) (Ojala et al. 1994) and Local Directional Pattern (LDP) (Jabid et al. 2010). The demerit of these two descriptors is that they do not provide rotational invariance. MLOOP eliminates the disadvantage of these descriptors and at the same time, preserves the robustness of each. It is an improvement designed on the LOOP descriptor (Chakraborti et al. 2018).

Let (x_c, y_c) be the pixel of an image I with an intensity i_c. Let the intensity of a 3×3 neighborhood pixel of (x_c, y_c) be $i_n(n = 0, 1, 2, ..., 7)$, neglecting the center pixel i_c. The 8 neighboring pixels are oriented as per the 8 Robinson masks, $i_n(n = 0, 1, 2, ..., 7)$. The measure of the strength of intensity variations in different directions

for 8 neighboring pixels is given by the 8 Robinson masks oriented in the direction of these neighboring pixels.

Let m_n corresponding to pixels with intensity $i_n (n = 0, 1, 2, ..., 7)$ be the 8 responses of Robinson mask. An exponential w_n (a digit between 0 and 7) is assigned to each of these pixels according to the rank of the magnitude of m_n among the 8 Robinson mask outputs.

The Local Binary Pattern is created by using the following Equation:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(m_n - i_c) \tag{4.1}$$

where

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & otherwise \end{cases} \tag{4.2}$$

Here, i_c stands for the intensity of an image I at pixel (x_c, y_c), $m_n (n=0,1,..,7)$ corresponding to pixels with intensities i_n are the 8 responses of Robinson mask and $s(x)$ represents a function where $x = m_n - i_c$. if $(m_n - i_c) \geq 0$ then $s(m_n - i_c) = 1$, otherwise $s(m_n - i_c) = 0$.

The MLOOP value for the pixel is given by:

$$MLOOP(x_c, y_c) = \sum_{n=0}^{7} s(m_n - i_c).2^{w_n} \tag{4.3}$$

where $s(x)$ is same as Equation 4.2.

The binarization weight to each of the adjacent pixels is assigned a value equivalent to the potency of Robinson output in the direction of that pixel. The output of the Robinson mask in a specific direction presents an indication of the likelihood of occurrence of an edge in that direction. Figure 4.3 illustrates the rotation invariance property of the MLOOP descriptor. Here for both patterns 1 and 2 (Chakraborti et al. 2018), binary words are generated by using Equation 4.1, and the weights are allocated
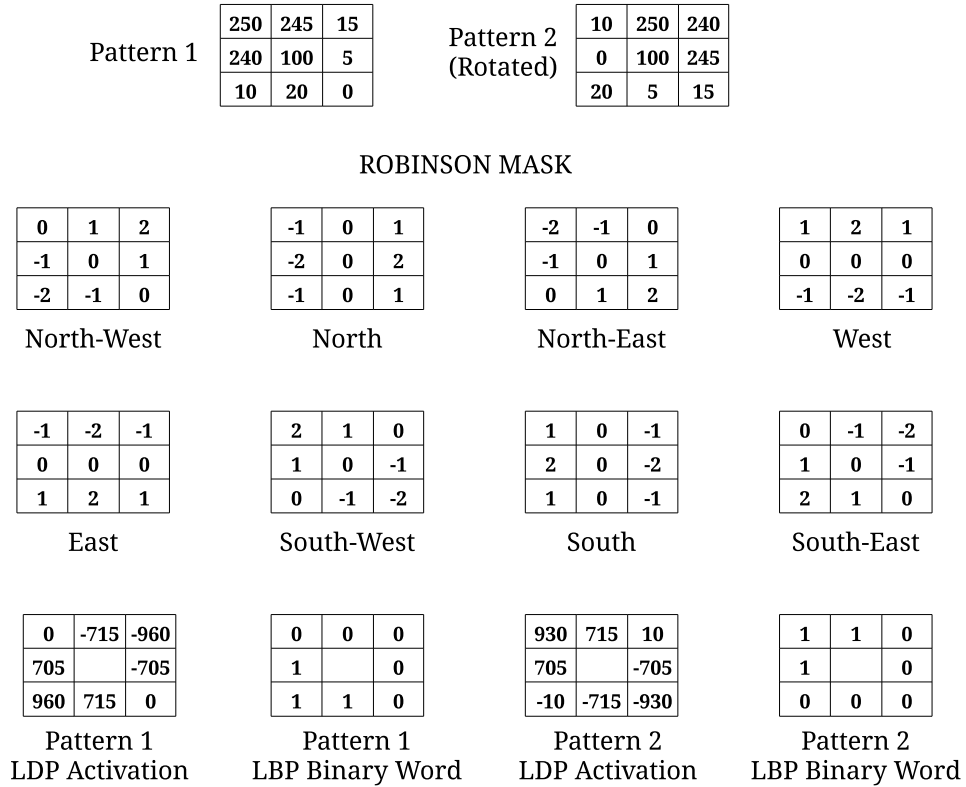
Figure 4.3: Example to show the numerical calculation of MLOOP descriptor.

following the LDP mask activation. Later on, MLOOP rule is applied to both the patterns. It can be observed that for both patterns in Figure 4.3, MLOOP value is given by $2^7 + 2^6 + 2^5 + 0 = 224$, which proves its rotation invariance as in LOOP.

The difference between LOOP descriptor and MLOOP descriptor is as follows:

1. The use of a different mask during LDP activation.

2. The pattern obtained after LDP activation is used to produce a Local Binary Pattern.

This method produces different values in the 3×3 neighbourhood of (x_c, y_c) for each descriptor after the LDP activation, which results in a different intensity of the centre pixel, i_c. Hence, the pixel intensity values obtained from an image for LOOP and MLOOP are different. Figure 4.4 demonstrates this fact.

The significant advantage of MLOOP descriptor is that the value of a pixel in the 3×3 neighbourhood of (x_c, y_c), after applying the LDP rule using Robinson mask is the additive inverse of the value of the pixel in the opposite direction. This is due to
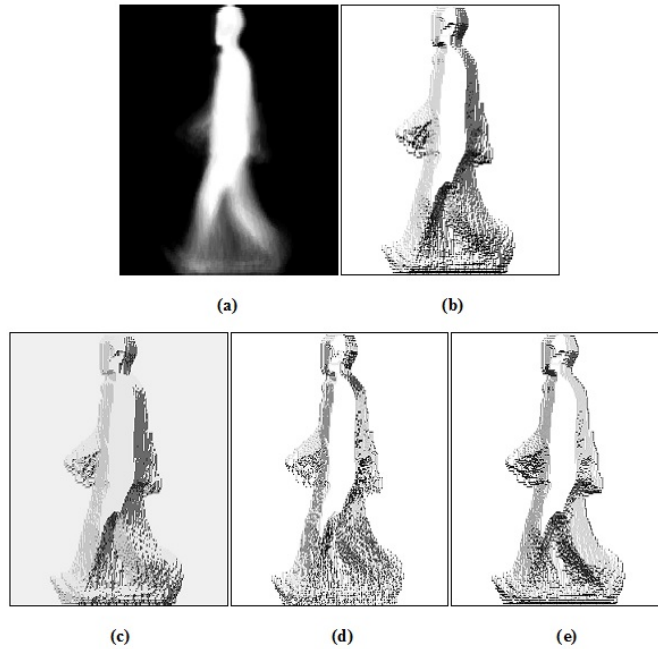
Figure 4.4: Output image of various descriptors (a) Standard GEI (b) LBP (c) LDP (d) LOOP (e) MLOOP.

the effect of the Robinson mask. Because of this, half of the 3×3 neighbourhood pixels values are always negative, which is less than the intensity value of (x_c, y_c). The MLOOP value for the pixel is obtained by the remaining half of the 3×3 neighbourhood pixels. Hence, the resultant MLOOP descriptor values are very similar. So, when a histogram of this descriptor is found, many of its values are equal to zero, which can be removed whereas for the LOOP descriptor, the histogram comprises mostly non-zero values. Because of this, MLOOP achieves high recognition accuracy with a smaller number of features.

4.1.3 MLOOP Histogram

The histogram of different descriptors such as LBP, LDP, and so on are generally used in most of the human recognition systems (Zhang and Zhang 2018; Lishani et al. 2018). Here, the descriptor of an image is the histogram of the MLOOP values corresponding to all pixels of an image. This procedure aims to extract the discriminating features from GEI of a subject. As GEI is a greyscale image, the histogram of its ROIs consists of 256 values. Therefore, for a single ROI, 256×1 feature vector is obtained.

66
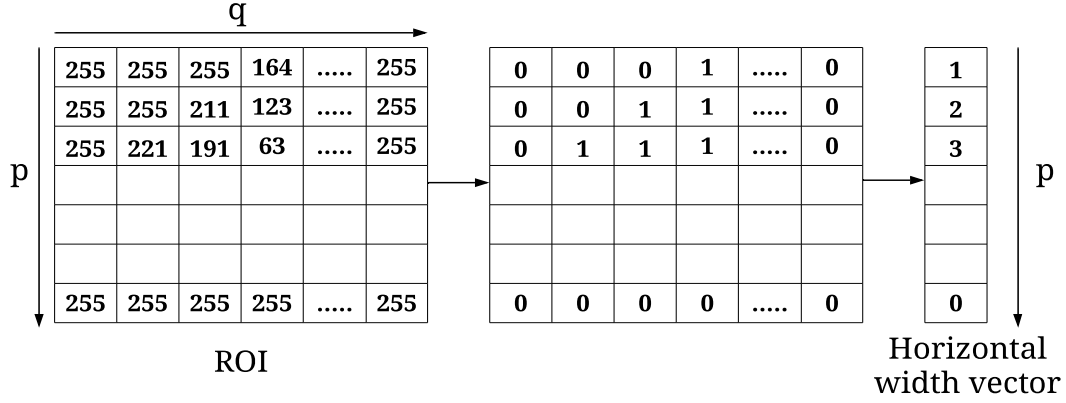
Figure 4.5: Example to show the numerical calculation of MLOOP horizontal width vector.

4.1.4 MLOOP Horizontal Width Vector

The texture of GEI can be represented more efficiently by a greyscale image called MLOOP descriptor. Here, the width of the MLOOP descriptor image, along the horizontal axis, is considered as the feature vector for recognition. The horizontal width vector (fv_i) is calculated in terms of the number of texture pixels present in each row of the ROI, by maintaining the row-wise count. The three benefits of this feature extraction method are as follows: The first benefit is that it removes the static redundant information, second, it consists of significant data, third, the dimensionality of the feature space is reduced to a great extent.

Consider a region of interest, GEI_i of size $m \times n$. Let R_1, R_2,....,R_m represent the rows of GEI_i.

$$GEI_i = \begin{bmatrix} GEI_{11} & GEI_{12} & ...... & GEI_{1n} \\ GEI_{21} & GEI_{22} & ...... & GEI_{2n} \\ . & . & ...... & . \\ . & . & ...... & . \\ . & . & ...... & . \\ GEI_{m1} & GEI_{m2} & ...... & GEI_{mn} \end{bmatrix} \qquad (4.4)$$

Let R_k be the kth row in GEI_i. Here $R_k = [GEI_{k1}\ GEI_{k2}\ ......\ GEI_{kn}]$. Then, the
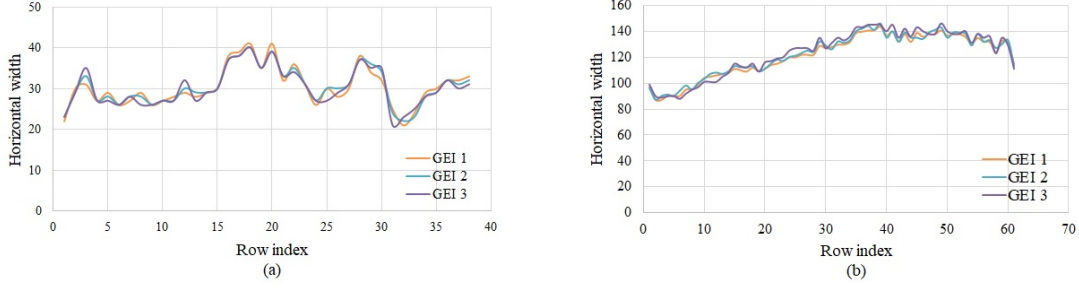
Figure 4.6: Feature space diagram of three GEI's (GEI 1, GEI 2 and GEI 3) belonging to the same subject. Here (a) represents horizontal width vector of top ROI (b) represents horizontal width vector of bottom ROI.
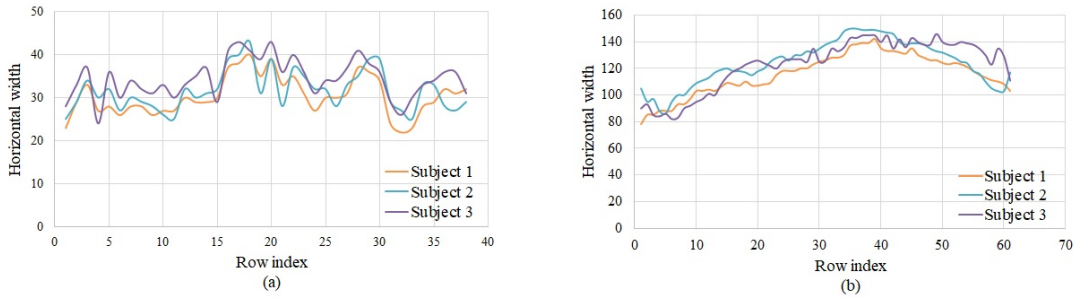
Figure 4.7: Feature space diagram of a single GEI belongs to three different subjects (Subject 1, Subject 2 and Subject 3). Here (a) represents horizontal width vector of top ROI (b) represents horizontal width vector of bottom ROI.

value of a pixel at position GEI_{kj}, where $j = (1, 2, ..., n)$ is calculated as given below.

$$
GEI_{kj} = \begin{cases} 1 & \text{if } GEI_{kj} < 255 \\ 0 & otherwise \end{cases} \tag{4.5}
$$

The value of each pixel ranges from 0 to 255 in an MLOOP descriptor. Here, the pixel value 255 is omitted because it represents the white background in the image. Consequently, the horizontal width vector for row R_k is computed from: $R_k = \sum_{j=1}^{n} GEI_{kj}$. An example of the numerical calculation of the horizontal width vector is shown in Figure 4.5.

The discriminative ability of the extracted feature vector is demonstrated in the feature space diagram. Figure 4.6 represents the features extracted from the same subject for three different gait cycles. A feature extracted from three subjects for a single gait cycle is presented in Figure 4.7. The row index in Figures 4.6 and 4.7 signifies the corresponding horizontal rows of the ROI, while the horizontal width specifies the number of texture pixels. The original sizes of two ROIs are 40×252 and 63×252. They are

68

converted into feature vectors of sizes 40×1 and 63×1.

It is evident from the feature space diagram that very minute variation is displayed within the three gait cycles for the same subject (GEI 1, GEI 2, GEI 3), but a significant variation is observed for the gait cycles of the three subjects (Subject 1, Subject 2, Subject 3). The separation of the curves in Figures 4.7(a) and 4.7(b), represent a considerable distinction between the inter-class distances. The small intra-class variances in Figures 4.6(a) and 4.6(b) indicate the discriminative ability of feature under consideration. It can be observed from the feature space diagram that the curve of the bottom ROI appears to have a similar shape, irrespective of the changes in the shapes of the silhouettes of various subjects. This ensures that the extracted feature can be used for clothing invariant gait recognition.

4.1.5 Feature Reduction

From the extracted feature vectors of various subjects in a given dataset, it is evident that the feature vector consists of some redundant values. Therefore, the following steps are carried out for reducing the dimensionality of the extracted feature set so that the irrelevant features can be removed.

Let S_1, S_2, S_3,.......,S_n represent the subjects of a gait dataset S. Consider a subject S_k from S. Let gS_k be the number of gait cycles belonging to the subject S_k. That is, $gS_k = \{g_1S_k, g_2S_k, ..., g_cS_k\}$. For each gait cycle of gS_k, the GEI is computed, followed by the extraction of 2 ROIs (GEI_1 and GEI_2). Here the feature vector is obtained from the ROIs of all GEI's of a subject.

Case 1: MLOOP histogram features

Let the feature vector obtained from all gait cycles of the subject S_k be $S_kf =$

$\{S_k fh_1, S_k fh_2\}$. where,

$$S_k fh_1 = \begin{bmatrix} g_1 S_k fh_1 \\ g_2 S_k fh_1 \\ . \\ . \\ . \\ g_c S_k fh_1 \end{bmatrix} \tag{4.6}$$

$$S_k fh_2 = \begin{bmatrix} g_1 S_k fh_2 \\ g_2 S_k fh_2 \\ . \\ . \\ . \\ g_c S_k fh_2 \end{bmatrix} \tag{4.7}$$

Extract the MLOOP histogram feature vectors for all subjects present in the datatbase S. Create a matrix from all the extracted feature vectors.

$$Hist_{top} = \begin{bmatrix} S_1 fh_1 \\ S_2 fh_1 \\ . \\ . \\ . \\ S_n fh_1 \end{bmatrix} \qquad Hist_{bottom} = \begin{bmatrix} S_1 fh_2 \\ S_2 fh_2 \\ . \\ . \\ . \\ S_n fh_2 \end{bmatrix} \tag{4.8}$$

The two matrices $Hist_{top}$ and $Hist_{bottom}$ are of the size $S_n \times 256$. Let C be a matrix formed from the columns of $Hist_{top}$ and $Hist_{bottom}$.

$$C = \begin{bmatrix} C_1 & C_2 & C_3 & .... & C_{256} \end{bmatrix} \tag{4.9}$$

The steps of feature reduction are as follows: Consider a column C_k from C.

1. If all the values of the elements in the column C_k are equal to zero, then remove

Table 4.1: Number of MLOOP histogram features removed and retained.

Feature set	Feature vector (size)	C_k(all 0's) removed	C_k retained
$Hist_{top}$	256	184	72
$Hist_{bottom}$	256	57	199

C_k.

2. If any value present in the column C_k is not equal to zero, then retain C_k.

When MLOOP descriptor is computed, a feature vector of size 256×1 is obtained from each ROI. It is observed from the $Hist_{top}$ and $Hist_{bottom}$ matrices that the values of all elements in some of the columns are equal to zero. This is because when MLOOP descriptor is computed, some values in the range 0 to 255 do not occur in the descriptor. When all the values of elements in a column C_k are equal to zero, it represents the redundant information. Table 4.1 shows the number of features removed and retained by following the above method. From Table 4.1, it is evident that the feature vector of size 512×1 is reduced to 271×1, by following the proposed feature reduction method.

Case 2: MLOOP horizontal width vector features

Consider a subject S_k of dataset S with gait cycles $gS_k = \{g_1 S_k, g_2 S_k, ..., g_c S_k\}$. Let $g_c GEI_i$, where i = 1, 2 be the two ROIs obtained for the gait cycle g_c. If R_1, R_2, R_3,......,R_n represent the rows of GEI_i, then the two ROIs can be represented by Equation 4.11. Both ROIs have 252 columns, as the GEI is of size 252×252. Hence, the two ROIs can also be represented in the form of columns by Equation 4.12.

$$g_c GEI_1 = \begin{bmatrix} R_1 \\ R_2 \\ . \\ . \\ . \\ R_{40} \end{bmatrix} \quad g_c GEI_2 = \begin{bmatrix} R_{89} \\ R_{90} \\ . \\ . \\ . \\ R_{252} \end{bmatrix} \tag{4.10}$$

$$g_c GEI_1 = \begin{bmatrix} C_1 & C_2 & ...... & C_{252} \end{bmatrix} \quad g_c GEI_2 = \begin{bmatrix} C_1 & C_2 & ...... & C_{252} \end{bmatrix} \tag{4.11}$$

When MLOOP descriptor is computed, the values of all elements in the first and

Table 4.2: Number of features obtained before and after feature reduction.

Case	fh_1	fh_2	fv_1	fv_2	**Total**
Before feature reduction	256	256	40	63	615
After feature reduction	57	199	38	61	355
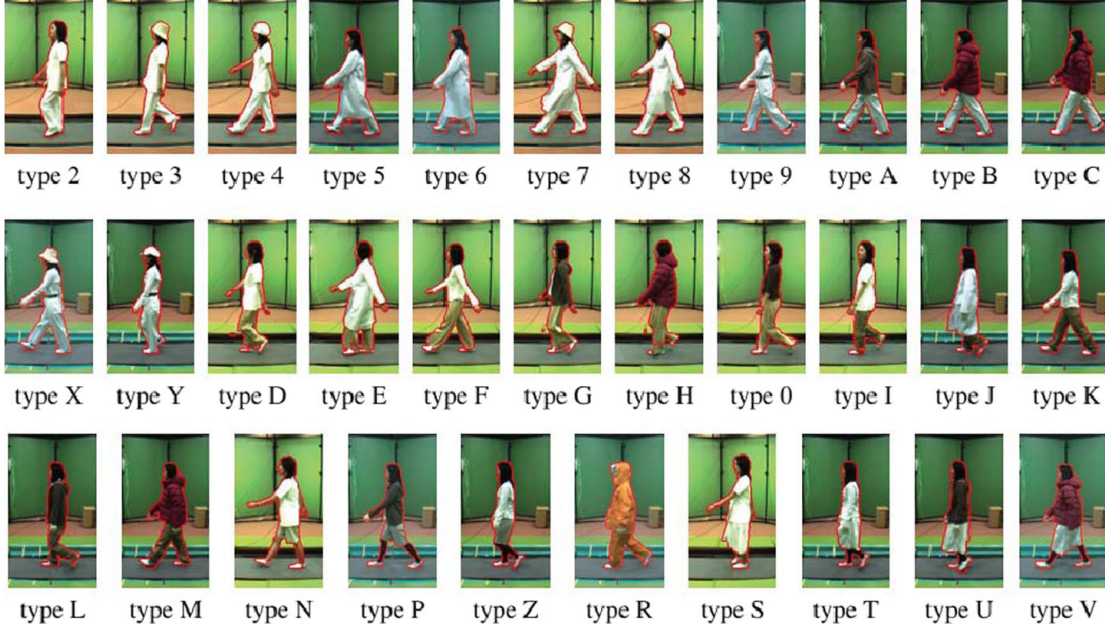
Figure 4.8: Sample images for every clothing differences from the OU-ISIR B dataset (Ghebleh and Moghaddam 2018).

last rows of ROIs are equal to zero. Similarly, the value of each element in the first and last columns are equal to zero. These rows and columns represent the redundant information. Hence R_1, R_{40}, C_1, C_{252} are removed from g_cGEI_1 and R_{89}, R_{252}, C_1, C_{252} are removed from g_cGEI_2.

When the horizontal width vectors fv_1 and fv_2 is computed, a feature vector of size 40×1 and 63×1 is obtained. In this feature vector fv_1 and fv_2, removal of first and last row results in the feature vector of size 38×1 and 61×1.

Table 4.2 shows that the total size of the feature vector is reduced from 615×1 to 370×1, by applying the above methods. That is, 57.72% of the features is removed from the final feature set.

4.1.6 Experiments

4.1.6.1 Experimental Setup

This experiment is carried out in a Windows 10 operating system with Intel Core i5-7200U CPU@2.50 GHz processor. The software tool used for the implementation of the proposed work is Matlab R2017b. The classifier used is KNN, where K=1. The CCR is used to measure the performance of the proposed approach.

4.1.6.2 Experimental Results and Discussion

The two gait datasets used to evaluate the performance of the proposed method are OU-ISIR B, and CASIA B datasets having subjects with the varied clothing styles. The term P 4.1 is used to specify the proposed method in comparison tables.

1. Experiments on OU-ISIR B dataset

Figure 4.8 gives some sample images of the 32 combinations of clothing conditions. Table 4.3 shows the list of clothes used in the dataset. Table 4.4 lists the combinations of clothing conditions used in this dataset. The two experiments carried out on this dataset are as follows:

Table 4.3: List of clothes used in OU-ISIR B gait dataset (Ghebleh and Moghaddam 2018).

Abbreviation	Name	Abbreviation	Name	Abbreviation	Name
RP	Regular Pants	HS	Half Shirt	CW	Casual Wear
BP	Baggy Pants	FS	Full Shirt	RC	Rain Coat
SP	Short Pants	LC	Long Coat	Ht	Hat
Sk	Skirt	Pk	Parker	Cs	Casquette Cap
CP	Casual Pants	DJ	Down Jacket	Mf	Muffler

Table 4.4: Various clothing combinations in OU-ISIR B gait dataset (Ghebleh and Moghaddam 2018).

Type	S_1	S_2	S_3	Type	S_1	S_2	Type	S_1	S_2
2	RP	HS	-	A	RP	Pk	T	Sk	FS
3	RP	HS	Ht	B	RP	Dj	U	Sk	Pk
4	RP	HS	Cs	I	BP	HS	V	Sk	Dj
9	RP	FS	-	K	BP	FS	D	CP	HS
X	RP	FS	Ht	J	BP	LC	F	CP	FS
Y	RP	FS	Cs	L	BP	Pk	E	CP	LC
5	RP	LC	-	M	BP	Dj	G	CP	DJ
7	RP	LC	Ht	Z	SP	FS	0	CP	CW
8	RP	LC	Cs	P	SP	Pk	R	RC	-
C	RP	DJ	Mf	S	Sk	HS			

I. In this experiment, the training set consists of 68 subjects. Here, a single clothing type RPFS is used for training, and the remaining clothing types such as SPHS, RPLC, RPHS, RPLCH, and RPHSH are used for testing. The testing set consists of various combinations of clothing, excluding RPFS. It is a challenging task to recognize an individual with different clothing styles in case the same individual is trained with a single clothing type. The proposed method manages to accomplish good results for different testing clothing index. The classification accuracy obtained and also its comparison with the prevailing methods in the literature are shown in Table 4.5. The experimental results confirm that the proposed approach increases the recognition accuracy, irrespective of different clothing conditions. It demonstrates that the feature invariance is preserved by reducing intra-class differences among each person's clothing datasets.

Table 4.5: Recognition accuracy of the proposed method on OU-ISIR B dataset with the prevailing methods

Sl. No.	Feature extraction	CCR (%)
1	Effective part based (Rokanujjaman et al. 2013)	65.67
2	Window based method (Islam et al. 2013)	73.28
3	Grid segmentation method (Nandy et al. 2016)	81.54
4	Multilayer perceptron (Semwal et al. 2015)	80.54
5	GEINet (Shiraga et al. 2016)	80.38
6	deep-CNN (Alotaibi and Mahmood 2017)	83.44
7	Deterministic learning (Deng and Wang 2018)	85.80
8	Proposed method (P 4.1)	**86.18**

II. A uniform training dataset is constructed which consists of gait templates of 32 clothing styles. One sequence is assigned to the training set for each of the 68 individual subjects, which leads to 2176 images in the training dataset. The remaining one sequence for each of the 68 subjects is allotted to the testing set in every experiment. The designed thirty-two experiments and the recognition performance for the proposed approach on OU-ISIR B gait dataset are listed in Table 4.6.

From the obtained experimental results, it can be concluded that the proposed method can minimize the issue of clothing variations and avoid the reduction of the recognition rate, even if these changes are small (e.g., hat) or large (e.g., long coat). When a clothing pattern of the test pattern appears in the prior training dataset, it can be identified. The average CCR obtained from all 32 experiments is 91.89%. The highest and lowest recognition accuracies obtained are presented in bold and italics, respectively.

2. Experiments on CASIA B dataset

Table 4.6: Experiments on the OU-ISIR B gait dataset for robustness test against clothing styles

Serial No.	Training dataset	Test dataset	Size of training dataset	Size of test dataset	Recognition accuracy
1	type 0,2,3...,Z	type 0	68 × 32	68 × 1	94.00
2	type 0,2,3...,Z	type 2	68 × 32	68 × 1	93.50
3	type 0,2,3...,Z	type 3	68 × 32	68 × 1	91.60
4	type 0,2,3...,Z	type 4	68 × 32	68 × 1	94.10
5	type 0,2,3...,Z	type 5	68 × 32	68 × 1	94.50
6	type 0,2,3...,Z	type 6	68 × 32	68 × 1	92.00
7	type 0,2,3...,Z	type 7	68 × 32	68 × 1	94.20
8	type 0,2,3...,Z	type 8	68 × 32	68 × 1	94.50
9	type 0,2,3...,Z	type 9	68 × 32	68 × 1	92.00
10	type 0,2,3...,Z	type A	68 × 32	68 × 1	91.60
11	type 0,2,3...,Z	type B	68 × 32	68 × 1	88.20
12	type 0,2,3...,Z	type C	68 × 32	68 × 1	94.50
13	type 0,2,3...,Z	type D	68 × 32	68 × 1	92.00
14	type 0,2,3...,Z	type E	68 × 32	68 × 1	91.50
15	type 0,2,3...,Z	type F	68 × 32	68 × 1	93.10
16	type 0,2,3...,Z	type G	68 × 32	68 × 1	89.10
17	type 0,2,3...,Z	type H	68 × 32	68 × 1	95.00
18	type 0,2,3...,Z	type I	68 × 32	68 × 1	**98.50**
19	type 0,2,3...,Z	type J	68 × 32	68 × 1	91.50
20	type 0,2,3...,Z	type K	68 × 32	68 × 1	87.50
21	type 0,2,3...,Z	type L	68 × 32	68 × 1	90.00
22	type 0,2,3...,Z	type M	68 × 32	68 × 1	97.50
23	type 0,2,3...,Z	type N	68 × 32	68 × 1	*85.50*
24	type 0,2,3...,Z	type P	68 × 32	68 × 1	91.10
25	type 0,2,3...,Z	type R	68 × 32	68 × 1	86.20
26	type 0,2,3...,Z	type S	68 × 32	68 × 1	89.10
27	type 0,2,3...,Z	type T	68 × 32	68 × 1	95.00
28	type 0,2,3...,Z	type U	68 × 32	68 × 1	95.50
29	type 0,2,3...,Z	type V	68 × 32	68 × 1	91.60
30	type 0,2,3...,Z	type X	68 × 32	68 × 1	90.10
31	type 0,2,3...,Z	type Y	68 × 32	68 × 1	89.00
32	type 0,2,3...,Z	type Z	68 × 32	68 × 1	87.20

A set of experiments conducted on this dataset is as follows.

In the CASIA B gait dataset, each subject has six normal walking sequences (*nm*) and two clothing variation sequences (*cl*). Here, the first four normal walking sequences of *nm* are used for training. The two sequences of *cl* are used for testing clothing variations. The experimental results and the comparison of them with the existing methods in the literature are shown in Table 4.7. When compared with the results of the other recognition methods verified on the CASIA B gait dataset, the proposed method performs better in most of the cases.

From the formerly stated results in the literature, it is apparent that the clothing

Table 4.7: Recognition accuracy of the proposed method on the CASIA B gait dataset with prevailing methods

Method	0^0	18^0	36^0	54^0	72^0
Choudhury and Tjahjadi (2015)	67.00	56.00	80.00	71.00	75.00
Dupuis et al. (2013)	81.64	87.39	86.29	84.34	89.96
Rida et al. (2016b)	80.49	83.47	85.08	87.85	91.53
Semwal et al. (2015)	58.63	50.52	53.05	65.40	75.46
Shiraga et al. (2016)	75.00	77.00	73.60	78.40	81.63
Alotaibi and Mahmood (2017)	79.10	82.52	76.00	81.63	85.20
Isaac et al. (2017)	97.00	**99.49**	97.50	94.00	88.00
Deng and Wang (2018)	97.81	98.80	**98.05**	**95.00**	92.56
Proposed method (P 4.1)	**98.85**	95.60	96.30	91.90	**94.00**

Method	90^0	108^0	126^0	144^0	162^0	180^0
Choudhury and Tjahjadi (2015)	77.00	75.00	65.00	64.00	64.00	66.00
Dupuis et al. (2013)	91.86	89.50	85.04	72.24	78.40	82.70
Rida et al. (2016b)	91.07	87.90	86.23	87.45	84.90	83.06
Semwal et al. (2015)	78.89	74.80	56.52	51.73	51.81	62.63
Shiraga et al. (2016)	82.95	81.77	76.82	80.00	82.05	80.46
Alotaibi and Mahmood (2017)	84.46	82.77	79.26	84.70	86.40	82.00
Isaac et al. (2017)	90.50	89.50	94.50	92.00	91.28	93.94
Deng and Wang (2018)	92.80	91.52	93.10	**92.73**	91.60	94.52
Proposed method (P 4.1)	**95.20**	**94.60**	**95.40**	90.40	**93.00**	**95.05**

condition leads to less CCR, making it a largely demanding covariate. They cause major changes in the silhouettes of the subjects. The recognition performance is unfavorably affected because the template-based techniques depend on spatiotemporal variations of the silhouettes over a gait cycle. When the features are obtained by applying MLOOP descriptor, more discriminative features are obtained. It is observed that promising performance can still be accomplished for different viewpoints, no matter the clothing variations alter significantly or slightly.

Furthermore, the performance of the proposed method is examined under various clothing conditions in different viewpoints. The CASIA B gait dataset comprises sequences from 124 subjects under 11 different views, and the experiments were conducted for all viewpoints. For each view, two experiments were conducted. A uniform training dataset that contains gait patterns under various clothing styles is created. The three *nm* sequences and one *cl* sequence (four sequences) form the training set, which constitutes of 496 patterns. In the first experiment, the testing set consists of three *nm* sequences under normal clothing condition. In the second experiment, one *cl* sequence is allotted to the testing set under different clothing circumstances. Twenty two experiments designed for this dataset are listed in Table 4.8. Even though the proposed

Table 4.8: Experiments on the CASIA B gait dataset for robustness against viewpoint and clothing conditions.

View angle	Training dataset	Test dataset	Size of training dataset	Size of test dataset	Recognition accuracy
$0^0(nm)$	$nm + cl$	nm	124×4	124×3	93.50
$0^0(cl)$	$nm + cl$	cl	124×4	124×1	92.10
$18^0(nm)$	$nm + cl$	nm	124×4	124×3	91.10
$18^0(cl)$	$nm + cl$	cl	124×4	124×1	90.99
$36^0(nm)$	$nm + cl$	nm	124×4	124×3	90.50
$36^0(cl)$	$nm + cl$	cl	124×4	124×1	90.90
$54^0(nm)$	$nm + cl$	nm	124×4	124×3	91.10
$54^0(cl)$	$nm + cl$	cl	124×4	124×1	93.20
$72^0(nm)$	$nm + cl$	nm	124×4	124×3	94.70
$72^0(cl)$	$nm + cl$	cl	124×4	124×1	96.50
$90^0(nm)$	$nm + cl$	nm	124×4	124×3	93.50
$90^0(cl)$	$nm + cl$	cl	124×4	124×1	95.10
$108^0(nm)$	$nm + cl$	nm	124×4	124×3	92.70
$108^0(cl)$	$nm + cl$	cl	124×4	124×1	94.10
$126^0(nm)$	$nm + cl$	nm	124×4	124×3	91.10
$126^0(cl)$	$nm + cl$	cl	124×4	124×1	91.50
$144^0(nm)$	$nm + cl$	nm	124×4	124×3	92.20
$144^0(cl)$	$nm + cl$	cl	124×4	124×1	93.50
$162^0(nm)$	$nm + cl$	nm	124×4	124×3	88.70
$162^0(cl)$	$nm + cl$	cl	124×4	124×1	91.10
$180^0(nm)$	$nm + cl$	nm	124×4	124×3	92.70
$180^0(cl)$	$nm + cl$	cl	124×4	124×1	90.20

technique is not intended for obtaining a solution to the view angle problem, the performance of it is still tested on view angle variations. It can be observed from Table 4.8 that the proposed feature extraction method exhibits promising performance under different clothing variations in different viewpoints.

4.1.6.3 Comparison of the Performance of MLOOP with its Predecessors

The output of all four descriptors on a standard gait GEI is presented in Figure 4.4. It can be observed from Figure 4.4 that the intensity of the pixels is different for LBP, LDP, LOOP, and MLOOP descriptors. The classification accuracies of all the four descriptors for OU-ISIR B dataset (experiment 1) and CASIA B dataset (for 90^0 view) are demonstrated in Figure 4.9. We compared the performance of LBP, LDP, and LOOP with the proposed MLOOP descriptor. From the Figure 4.9, it is evident that the proposed MLOOP descriptor achieves an improved CCR of 15.575% when compared to LBP, 10.965% with LDP, and 2.21% with LOOP. The experimental results indicate that MLOOP descriptor functions significantly superior to LBP and LDP descriptors and
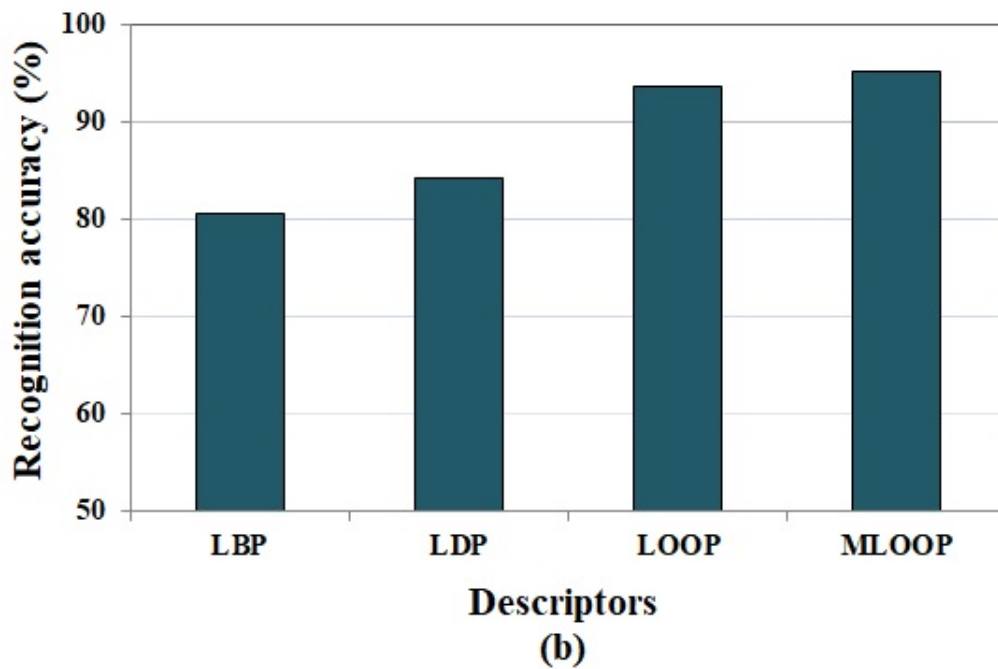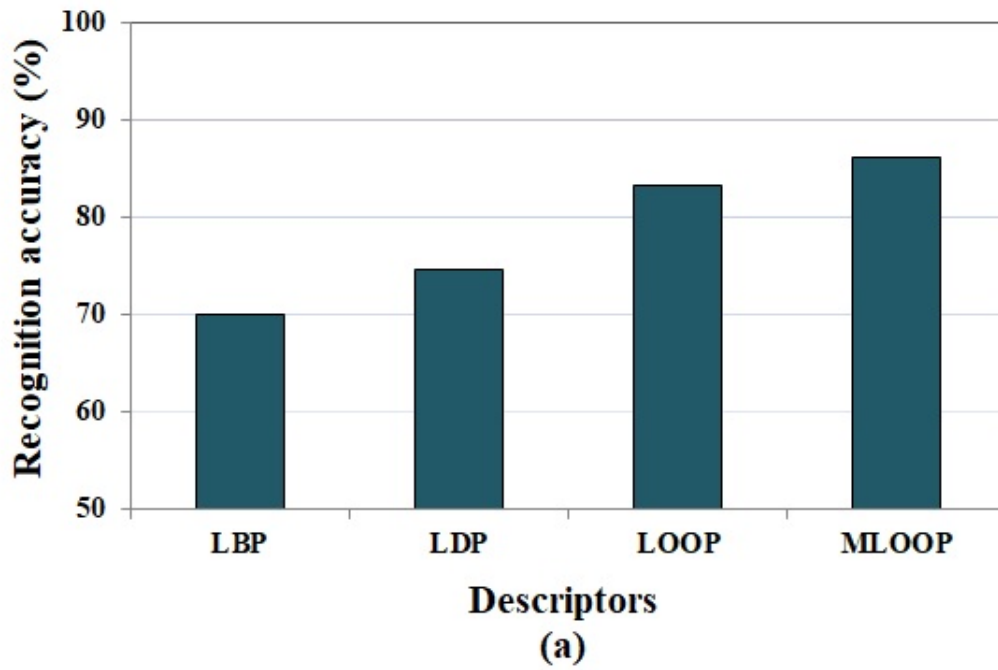
Figure 4.9: Performance results obtained from extracting the features by using various descriptors for (a) OU-ISIR B dataset, and (b) CASIA B dataset.

provides a slight improvement in CCR as compared to LOOP descriptor.

4.2 Summary

This study presents an efficient method for the human identification in a diverse clothing environment, which is conducive to the human detection. In this work, the two potential ROIs of GEI are selected to extract intrinsic gait signature. The MLOOP descriptor is used to obtain discriminative features from two ROIs. Further, redundant features that adversely affect the gait detection are removed. The evaluation results on two public gait datasets show that the proposed method is effective and increases the recognition rate significantly. The extension of this work may require working with the silhouettes acquired from an outdoor, normal location, not a controlled environment like a treadmill.

MLOOP descriptor effectively captures the local structure information of GEI. It surpasses the descriptor from which it is originated. This work has treated the gait recognition as a representative problem. However, MLOOP descriptor, being a comprehensive binary descriptor, can be utilized as a part of the research on different datasets.

Chapter 5

FRONTAL GAIT RECOGNITION

The performance of the gait recognition systems is predominantly influenced by various covariate conditions, such as view angle variations, walking surface conditions, elapsed time, shoe type, walking speed, and carrying conditions. Among these covariate conditions, one of the most vital one is the view angle variation. The generally challenging of all the view variations is the frontal view as the spatial and temporal changes are very less noticeable in a frontal view, in comparison to other view variations. Hence, two methods are proposed in this study to increase the performance of the frontal gait recognition system. Three publicly available gait datasets, such as CASIA A, CASIA B, and CMU MoBo are used for demonstrating the performance of the proposed approaches.

5.1 Frontal Gait Recognition based on Hierarchical Centroid Shape Descriptor and Similarity Measurement

Among the appearance-based methods, most of the existing individual identification systems are based on GEI, as it represents the spatial and temporal variations of a gait cycle. The front view GEI contains very less temporal information unlike other GEI's of lateral views. Hence, the frontal gait detection performance can be enhanced by making an adequate representation for spatial data. To accomplish this goal, the proposed work uses a hierarchical centroid descriptor and similarity measurement.

The contributions of this study can be summarized as follows.

1. The GEI is a greyscale image, which can be represented by the spatial distribution of pixels. Here, the frontal GEI contains the maximum number of pixels representing the same information. Hence, the HCSD is used to extract accurate shape information.

2. The steps of similarity measurement presented here increase the recognition performance, as it increases the inter-class differences even when the shape information of subjects is very much similar.

3. The performance of the proposed method is evaluated on three benchmark gait datasets, and the results are compared with those of the state-of-the-art gait recognition approaches.
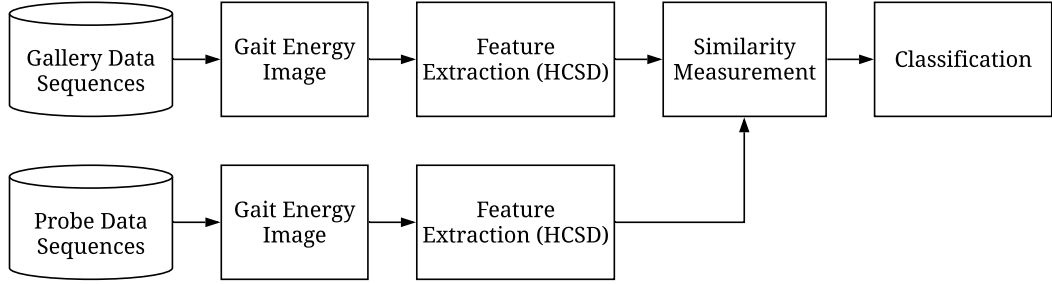
Figure 5.1: Framework of the proposed approach.
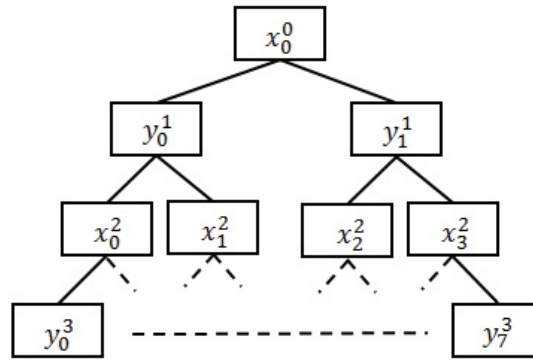
Figure 5.2: Structure of kd-tree decomposition ([Ilunga-Mbuyamba et al. 2016](#)).

5.1.1 Framework of the Proposed Method

An outline of the proposed method is shown in Figure 5.1. The pre-processing of the gait video is done by steps such as background subtraction, de-noising, and normalization to obtain gait silhouettes, which are later combined over a gait cycle to form a GEI. The proposed method consists of three steps: HCSD feature extraction, similarity measurement, and classification. Algorithm 9 demonstrates the entire process of recognition.

Algorithm 9 Process of recognition

Input: Collection of all gallery samples, a probe sample. **Output**: A subject.

 1: Begin
 2: Preprocess input to extract GEI's.
 3: Extract HCSD features from GEI's of all gallery samples G and probe sample p.
 4: Perform similarity measurement using G and p which results in the list of distances D^i.
 5: Perform classification to identify the subject belonging to p.
 6: End
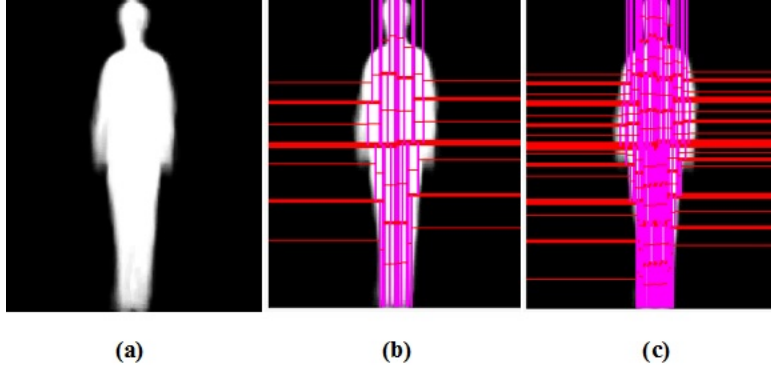
Figure 5.3: Sample image representing (a) frontal GEI (b) HCSD extraction from a frontal GEI for level 6 of kd-tree decomposition (c) HCSD extraction from a frontal GEI for level 8 of kd-tree decomposition.

5.1.2 Hierarchical Centroid Shape Descriptor

The shape descriptor formed from the centroid coordinates obtained from a greyscale image is called HCSD (Ilunga-Mbuyamba et al. 2016). This descriptor is extracted by decomposing an image into sub-images recursively by using the kd-tree representation, where the data is divided with reference to the center of gravity at each level of decomposition, as shown in Figure 5.2. This process gives sub-images at each stage, which usually differ in size. For each sub-image, the centroid coordinates for the local region are extracted and are stored in the corresponding level of a kd-tree. As a result, a descriptor, whose size is determined by the depth of decomposition is obtained. The sample image representing frontal GEI and HCSD for different levels of kd-tree decomposition is shown in Figures 5.3 and 5.4.

The length of the descriptor is computed from: $2 \times (2^d - 2)$, where d represents the depth of the feature extraction process. Let I be the greyscale image of size $A \times B$, with background I_b and foreground I_f, the HCSD is obtained by the following steps.

1. Consider an input image I and its transpose I^T,

2. Compute the centroid $C(x_{ct}, y_{ct})$ for each input, at the root level by using $x_{ct} = \frac{m_{10}}{m_{00}}$ and $y_{ct} = \frac{m_{01}}{m_{00}}$, where m_{10}, m_{01} represent the first order moments along the x-axis and y-axis, m_{00} signifies the area of I_f. The moment of order $(r + s)$ of an image with the pixel intensities $I(i, j)$ is defined as $m_{rs} = \sum_{i=0}^{A} \sum_{j=0}^{B} i^r j^s I(i, j)$.

3. Recursively split the image into two sub-images based on the centroid until the desired depth of decomposition is reached. The axis of coordinates obtained is altered at each consecutive level.

4. Normalization of the descriptor in the range -0.5 to 0.5 is done where 0 denotes the centroid of the root level. The positive values represent the features from the
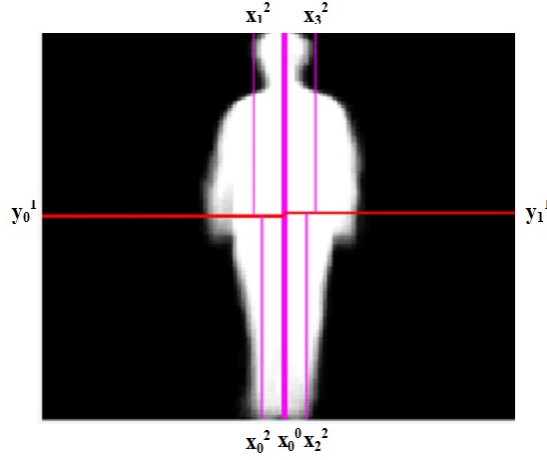
Figure 5.4: HCSD extraction from a frontal GEI for level 3 of kd-tree decomposition.

tree decomposition of the right part, while the negative values portray the left side of the image.

5. Concatenate the extracted features from the image I and I^T.

5.1.3 Similarity Measurement

The similarity measurement section compares each probe sample p with all features from the gallery samples. The result of similarity measurement is a list of distances for each p, namely, $dist^i$ containing the distances of p to each gallery feature. This vector is obtained by finding the distance d between two normalized feature vectors, probe p, and gallery g as follows: $d = \sum_{i=1}^{f}(p(i) - g(i))^2$, where $i = 1, 2, 3, ..., f$. Here, f is the total number of HCSD features.

The steps followed to perform similarity measurement are shown in Algorithm 10. In similarity measurement, the minimum difference between the summation of the distances obtained for a particular subject and probe is considered for classification, instead of the minimum distance between a probe and a gallery sample. This is because when a probe sample and a gallery sample belonging to different subjects have largely the same spatial data, and so by considering individual gallery samples, the differences between many gallery samples and a probe sample may be almost similar and because of this reason the recognition performance decreases, whereas the summation of the distances belonging to a particular subject gives different values so that the minimum value can be selected, thus increasing the recognition performance.

Algorithm 10 Similarity measurement

Input: Gallery features, a probe feature.
Output: List of distances.

1: Begin
2: Let the number of subjects present in the gallery dataset be $S = \{s_1, s_2, ..., s_n\}$.
3: Each subject persent in the gallery dataset consists of t number of GEI's, $s_i = \{g_i^1, g_i^2, ..., g_i^t\}$.
4: Extract HCSD features from GEI's of all gallery samples, $G = \{g_1^1, g_1^2, .., g_1^t, g_2^1, g_2^2, .., g_2^t, ..., g_n^1, g_n^2, .., g_n^t\}$.
5: Extract HCSD feature from GEI of a probe sample, p.
6: **for** each $g_i^j \in G$ **do** where $i = \{1, 2, ..., n\}$ and $j = \{1, 2, ..., t\}$
7:     Compute the distance d between HCSD features of p and g_i^j.
8: **end for**
9: Let $dist^i = \{d_i^1, d_i^2, ..., d_i^t\}$ be the distances obtained between p and gallery samples belonging to subject s_i.
10: Perform the addition of the distances obtained between p and $dist^i$ to obtain $D^i = \{d_i^1 + d_i^2 + ... + d_i^t\}$
11: The above step results in the list of distances $D^i = D^1, D^2, ..., D^n$.
12: Perform classification using D^i to identify the subject S which p belongs to.
13: End

5.1.4 Classification

Let $S = s_1, s_2, ..., s_n$ be the number of subjects present in the gallery dataset. Consider a vector consisting of a list of distances $D^i = D^1, D^2, ..., D^n$ that represent the distance between the subjects $s_1, s_2, ..., s_n$ and probe sample p after performing the similarity measurement. Let the minimum value present in vector D^i is given by $D^m = min(D^i)$. If D^m is the distance between probe p and subject s_i, then the probe sample p is assigned to s_i. This is because the minimum distance signifies that there is the least difference in the shape and static data between the probe and the gallery samples.

5.1.5 Experiments

5.1.5.1 Experimental Setup

In this study, the depth d of the feature HCSD is set to 8. As a result, the length of HCSD for each GEI is of size 1×508. The correlation of the feature vector is illustrated in the feature space diagram. Figure 5.5(a) shows the features extracted from the same subject for three different GEI's. The features extracted from three subjects for a single GEI are shown in Figure 5.5(b). It is apparent from the feature space

Table 5.1: Comparision of CCR of the proposed method with existing methods on CMU MoBo dataset.

Method	b/f	b/i	b/s	f/b	f/i	f/s	i/b	i/f	i/s	s/b	s/f	s/i
Huang et al. (2016)	88	88	**96**	92	84	96	75	**88**	92	**96**	96	92
Proposed method (P 5.1)	**90**	**90**	**96**	**94**	**88**	**98**	**82**	**88**	**96**	92	**98**	**94**
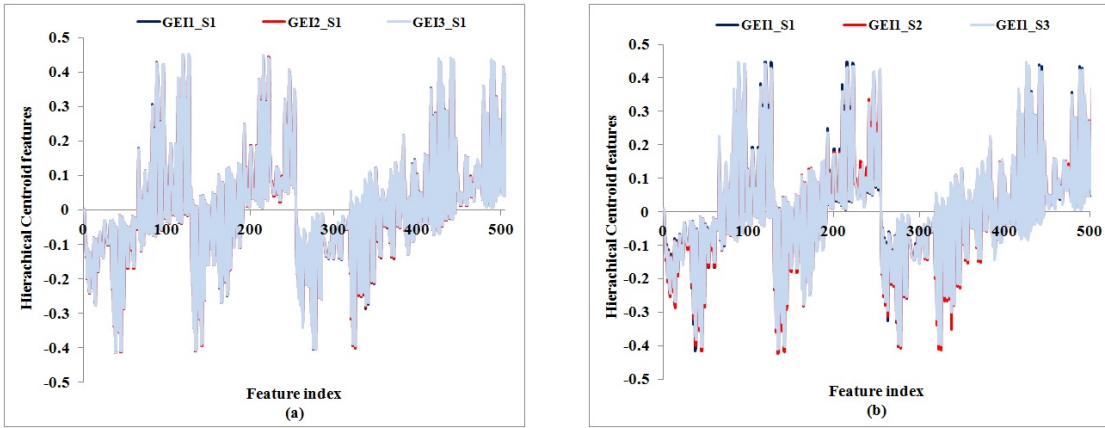
Figure 5.5: Feature space diagram representing HCSD features of three (a) GEI's belonging to same subject (b) GEI's belonging to different subjects.

diagram where the minute differences can be within the three GEI's of the same subject (GEI1_S1, GEI2_S1, GEI3_S1), but a considerable difference is obtained for GEI of three different subjects (GEI1_S1, GEI1_S2, GEI1_S3). Hence, the value of difference $dist$ when probe p and gallery g belong to the same subject is within 0 to 0.004, whereas the value of $dist$ when probe p and gallery g belong to the different subject is within 0.084 to 0.097.

Initially, the experiment is conducted by assigning different values to depth d from 4 to 9. However, the encouraging results are obtained when its value is set to 8. The performance of the gait recognition system is measured by using CCR on the testing dataset.

5.1.5.2 Experimental Results and Discussion

The performance of the proposed method is verified on three extensively used gait datasets, such as (1) CASIA A, (2) CASIA B, and (3) CMU MoBo. The term P 5.1 is used to specify the proposed method in comparison tables.

1. Experiments on CASIA A dataset

In this dataset, four sequences are obtained from all individuals in each direction.
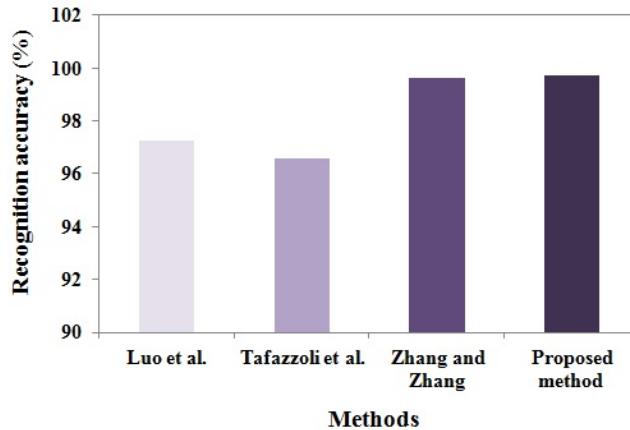
Figure 5.6: CCR (%) of the proposed method on CASIA A gait dataset and comparision of it with other existing methods.

Therefore, for 0^0 view, 80 sequences are obtained from 20 subjects. The performance of the proposed approach on CASIA A dataset for 0^0 view is reported in Figure 5.6. Three sequences out of 4 sequences are considered for training, and the remaining sequence is used for testing. The proposed method is compared with other existing methods such as Luo et al. (2015), Zhang and Zhang (2018) and Tafazzoli et al. (2015). Figure 5.6 shows that the proposed method gives higher CCR.

2. Experiments on CMU MoBo dataset

Here, the all four forms of walking present in this dataset are used as both gallery and probe datasets, where f/b represents the gallery f and the probe b. The results of the proposed method on the CMU MoBo dataset for the frontal view are represented in Table 5.1. It is obvious from Table 5.1 that the proposed method gives high CCR in most of the cases. It outperforms the other methods in almost all cases of gallery/probe combinations. An experiment is performed on the most demanding cases, which involve walking on an inclined plane and ball in hands, such as f/i, b/i, i/b, i/f, b/f, and f/b. The proposed method gives good performance in more challenging cases.

3. Experiments on CASIA B dataset

The two experiments performed on this dataset are as follows.

At first, four sequences of nm are used for training, and the remaining two sequences of nm, cl, and bg are used for testing. These testing sequences are employed to measure the performance of the normal, clothing, and carrying variations, respectively. Given

Table 5.2: Comparision of performance of the proposed method with existing methods on CASIA B gait dataset.

Methodology	0^0 nm/nm	0^0 nm/bg	0^0 nm/cl	180^0 nm/nm	180^0 nm/bg	180^0 nm/cl
Dupuis et al. (2013)	97.17	73.15	81.64	99.60	74.56	82.70
Choudhury and Tjahjadi (2015)	100.0	93.00	67.00	99.00	89.00	66.00
Rida et al. (2016b)	97.97	72.76	80.49	97.58	76.11	83.06
Alotaibi and Mahmood (2017)	90.67	91.98	88.77	83.99	87.76	90.00
Isaac et al. (2017)	98.50	95.00	97.00	**98.99**	94.44	93.94
Proposed method (P 5.1)	**98.70**	**97.00**	**97.56**	98.97	**95.06**	**94.58**

Table 5.3: CCR of the proposed method on CASIA B gait dataset.

Experiment	Gallery set	Probe set	Gallery Size	Probe Size	CCR (%)
1	0^0(nm)	180^0(nm)	124×4	124×4	90.97
2	0^0(nm)	180^0(nm)	124×4	124×3	91.19
3	0^0(nm)	180^0(nm)	124×4	124×2	92.79
4	0^0(nm)	180^0(nm)	124×4	124×1	92.50
5	0^0(nm)	180^0(bg)	124×4	124×2	93.15
6	0^0(nm)	180^0(bg)	124×4	124×1	93.60
7	0^0(nm)	180^0(cl)	124×4	124×2	91.60
8	0^0(nm)	180^0(cl)	124×4	124×1	91.41
9	180^0(nm)	0^0(nm)	124×4	124×4	89.59
10	180^0(nm)	0^0(nm)	124×4	124×3	90.76
11	180^0(nm)	0^0(nm)	124×4	124×2	90.39
12	180^0(nm)	0^0(nm)	124×4	124×1	91.81
13	180^0(nm)	0^0(bg)	124×4	124×2	94.56
14	180^0(nm)	0^0(bg)	124×4	124×1	94.96
15	180^0(nm)	0^0(cl)	124×4	124×2	93.79
16	180^0(nm)	0^0(cl)	124×4	124×1	94.60

that the GEI's of 0^0 and 180^0 views are similar to a large extent, the experiments are conducted on both 0^0 and 180^0 views to assess the performance of the proposed method. The results presented in Table 5.2 show that the proposed method gives considerably high CCR over other methods reported in the literature. They also demonstrate the capability of the proposed method in handling the carrying and clothing variations.

Secondly, a training dataset consisting of *nm* gait sequences of 0^0 and 180^0 viewpoints are constructed. For the training process, four *nm* gait sequences of each subject are used, leading to 496 GEI's in the training dataset. The testing dataset consists of a different number of GEI's in every experiment, as shown in Table 5.3. The performance of the proposed method for 16 experiments is illustrated in Table 5.3. The results show that the proposed method efficiently captures the statistical information present in the

frontal gait images.

5.2 Frontal Gait Recognition using Contour Image-based Feature Extraction

It can be concluded from the literature review that not only does the appearance-based methods perform better, they are also more suitable for surveillance environment in contrast to the model-based methods, and they are computationally effective too. Among the appearance-based methods, most of the existing individual identification systems based on gait extract shapes and contours from the human silhouette images to derive gait features.

Contour detection is one of the primary steps used in pattern recognition, image analysis, and computer vision methods (Ye and Wen 2006; Liu et al. 2011; Lee et al. 2013). Applying a contour detection algorithm to a gait image generates a set of connected vertices and curves that specify the boundaries of subjects, as well as curves that signify the discontinuities in surface orientation. As a corollary, applying the contour detection algorithm to a gait image may considerably decrease the quantity of information to be processed. Moreover, it may also remove the data that might be considered as less significant while preserving the key structural properties of a gait image. Therefore, this method presents the spatial features based gait representation using the contour image and contour vertices extracted from GEI.

The contributions of this study are as follows:

1. This study presents three feature vectors that are more reliable and also better characterize the spatial variations of a frontal gait. The low dimensional features that are extracted inevitably reduce the computational complexity of the gait recognition system.

2. The proposed feature vectors consist of identically distributed features from the frontal contour image and capture the minute changes in the shape of the GEI.

3. A broad experimental evaluation is conducted in this research. The proposed method is assessed on three benchmark gait datasets. The results are compared with those of the state-of-the-art and other contour-based gait recognition techniques.
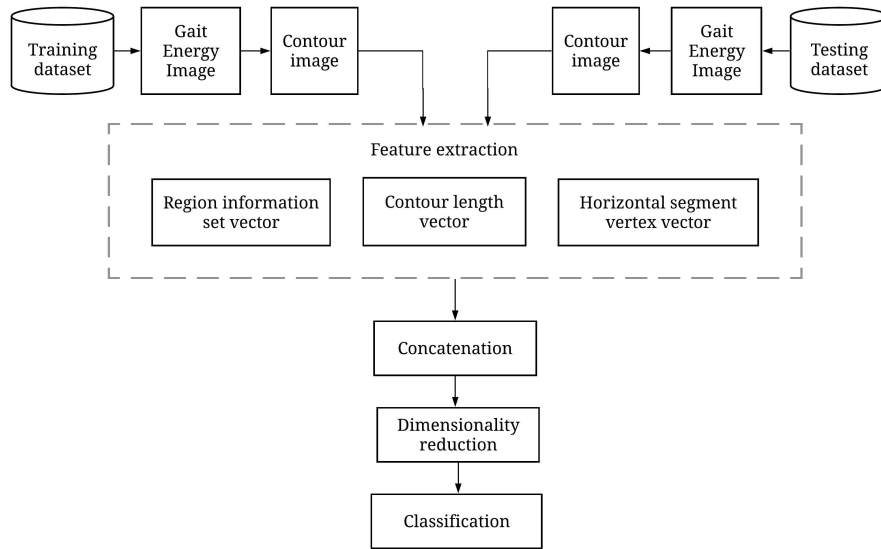
Figure 5.7: Steps of the proposed gait recognition approach.

5.2.1 Framework of the Proposed Method

An outline of the proposed method is presented in Figure 5.7. It consists of four stages. In the first stage, the gait silhouettes are combined over a gait cycle to form a GEI. The contour image and vertices are extracted from GEI. In the second stage, three different feature vectors are derived from the contour image and contour vertices of the subjects present in the training and testing datasets. This is followed by the application of PCA. This aids in eliminating the dimensions that negatively influence the robustness of the classification and thus, increases its performance. In the last stage, the features are classified by KNN classifier, where K=1. The performance of the gait recognition system is measured by the CCR on the testing dataset.

5.2.2 Contour Image Extraction

The GEI represents the spatial and temporal variations of a gait pattern. When the GEI of the front view gait cycle is considered, the temporal information present in this template will be minimal as compared to that in other lateral views. So, the performance of the frontal gait recognition can be enhanced by developing a compact representation for spatial data. To accomplish this, the contour of the GEI is extracted as it contains an accurate shape information.

The marching squares algorithm (Maple 2003) is used to generate a contour for a GEI image, with isovalue = 0.5, and then the linear interpolation is applied to determine
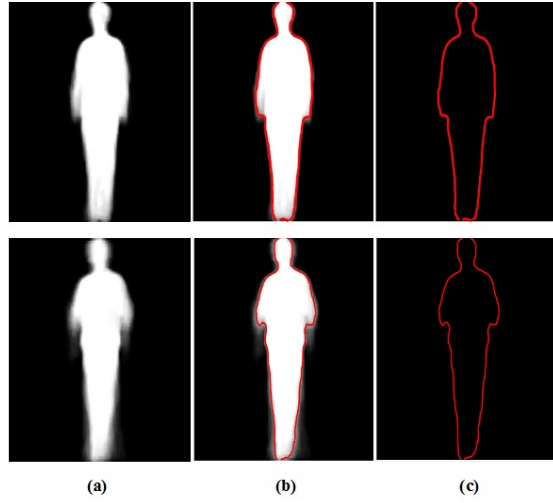
Figure 5.8: Sample images of two different subjects from CASIA B gait dataset representing (a) GEI (b) Contour of the image generated from marching squares algorithm and linear interpolation (c) Contour image.
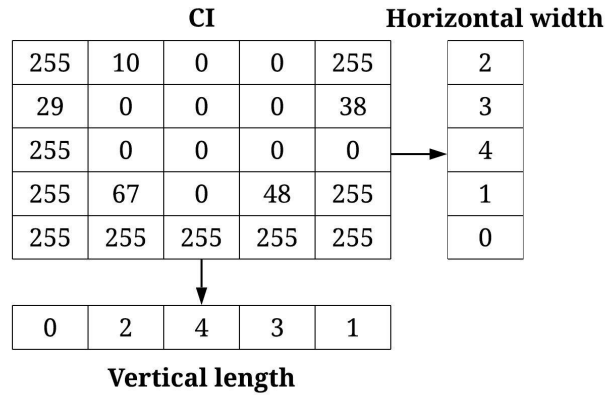
Figure 5.9: Sample image to show the numerical calculation of the horizontal width and vertical length vector.

the exact contour position. The application of this method on GEI generates a list of vertices $V(x, y)$. These vertices form the boundary of the GEI. Thus, it proves that the front view GEI consists of minimum temporal information. These vertices are used to get a Contour Image (CI), as shown in Figure 5.8.

5.2.3 Feature Extraction

The extraction of the most discriminatory feature vectors is most crucial to increase the performance of a gait recognition system. Algorithm 11 gives the steps involved in feature extraction. The three proposed feature descriptors are as follows: (1) Region information set vector, (2) Contour length vector, and (3) Horizontal segment vertex vector.

Algorithm 11 Feature extraction

1: **Input**: CI, of size $m \times n$.
2: **Output**: Feature vector, F.
3: Begin
4: Read an input image.
5: Compute region information set vector, R_s.
6: Compute contour length vector, $C_L = \{C_L E, C_L M, C_L C\}$.
7: Compute horizontal segment vertex vector, H_v.
8: Concatenate the 3 different feature vectors, $F = \{R_s, C_L, H_v\}$
9: End

---

1. Region information set vector

Consider a greyscale contour image CI of size $m \times n$.

$$
CI = \begin{bmatrix}
CI_{(1,1)} & CI_{(1,2)} & ...... & CI_{(1,n)} \\
CI_{(2,1)} & CI_{(2,2)} & ...... & CI_{(2,n)} \\
. & . & ...... & . \\
. & . & ...... & . \\
CI_{(m,1)} & CI_{(m,2)} & ...... & CI_{(m,n)}
\end{bmatrix}
\tag{5.1}
$$

Let $r_1, r_2, ..., r_m$ denote the rows and $c_1, c_2, ..., c_n$ denote the columns of the CI. Replace all pixels with 0 intensity values that are present outside the contour in the CI with 255. As the pixels outside the contour are replaced with intensity value 255, only the pixels with intensity 0 in CI constitute the shape of the subject.

a. Horizontal width vector

The horizontal width is calculated in terms of the number of pixels with intensity value 0 present in each row of the CI, followed by maintaining a row-wise count. Let r_k be the kth row in CI. Here, $r_k = [CI_{(k,1)} CI_{(k,2)} ... CI_{(k,n)}]$. In this case, the value of a pixel at position $CI_{(k,j)}$, where $j = 1, 2, .., n$ is given by

$$
CI_{(k,j)} = \begin{cases} 1 & \text{if } CI_{(k,j)} = 0 \\ 0 & otherwise \end{cases}
\tag{5.2}
$$

The horizontal width of the row r_k is given by $r_k = \sum_{j=1}^{n} CI_{(k,j)}$.

b. Vertical length vector

The vertical length is calculated in terms of the number of pixels with intensity value 0 present in each column of the CI, followed by maintaining a column-wise count. Let c_k be the kth column in the CI.

$$c_k = \begin{bmatrix} CI_{(1,k)} \\ CI_{(2,k)} \\ . \\ . \\ CI_{(m,k)} \end{bmatrix} \tag{5.3}$$

Here, the value of a pixel at position $CI_{(j,k)}$, where $j = 1, 2, ..., m$ is given by

$$CI_{(j,k)} = \begin{cases} 1 & \text{if } CI_{(j,k)} = 0 \\ 0 & otherwise \end{cases} \tag{5.4}$$

The Vertical length of the column c_k is given by $c_k = \sum_{j=1}^{m} CI_{(j,k)}$.

Let $H_w = [r_1 \, r_2 \, ... \, r_p]$ of size $1 \times p$ be the horizontal width vector obtained for the CI after eliminating the values where r_k is equal to 0 and $V_l = [c_1 \, c_2 \, ... \, c_q]$ be the vertical length vector obtained for the CI after removing the values where c_k equals to 0.

Concatenate horizontal width vector H_w and vertical length vector V_l to form region vector, RN_v. Divide each element of RN_v with the total number of elements, t.

$$RN_v = \begin{bmatrix} rn_1/t & rn_2/t & ...... & rn_t/t \end{bmatrix} \tag{5.5}$$

An example showing the numerical calculation of horizontal width and vertical length vectors is shown in Figure 5.9.

A fuzzy set is characterized by attribute values termed as the information source values and their Membership Function (MF) values in the interval [01], constituting pairs of elements. The fuzzy set has no provision for connecting the pair as a single entity. In contrast, the individual MF value cannot represent the overall uncertainty as-

sociated with all the information source values because the MF value can only measure the degree of belongingness of an information source value to the set. To overcome this drawback, the information set theory was proposed by M.Hanmandlu and developed by his research group at IIT Delhi (Aggarwal and Hanmandlu 2015) and comprehensive treatment of this theory can be found in the recent publication (Sayeed and Hanmandlu 2017). This solves the problem by taking the product of the information source value and its MF value into a single entity called the information value. A collection of such values is termed as the information set.

To summarize, in the case of fuzzy sets, the product of the information source value and its MF value is called as the information value (Sayeed and Hanmandlu 2017). A collection of such values which is referred to as information set allows the extraction of spatial dynamics as shown below.

Here, the information source considered is the region vector denoted by

$$R_v = \begin{bmatrix} R_1 & ... & R_t \end{bmatrix} \tag{5.6}$$

To compute Gaussian MF, the mean and standard deviation are obtained by

$$\mu^R = \frac{1}{t} \sum_{i=1}^{t} R_i \quad \text{and} \quad \sigma^R = \frac{1}{t} \sum_{i=1}^{t} (R_i - \mu^R)^2 \tag{5.7}$$

The vector of Gaussian MF is acquired using the following gain function (Medikonda et al. 2018)

$$g(R_i) = e^{-(R_i - \mu^R/\sigma^R)^2} \quad \text{where} \quad i = 1, 2..., t. \tag{5.8}$$

The information set features for region vector is called as region information set vector and is obtained by

$$R_v = \begin{bmatrix} R_1 \times g(R_1) & ... & R_t \times g(R_t) \end{bmatrix} \tag{5.9}$$

2. Contour length vector
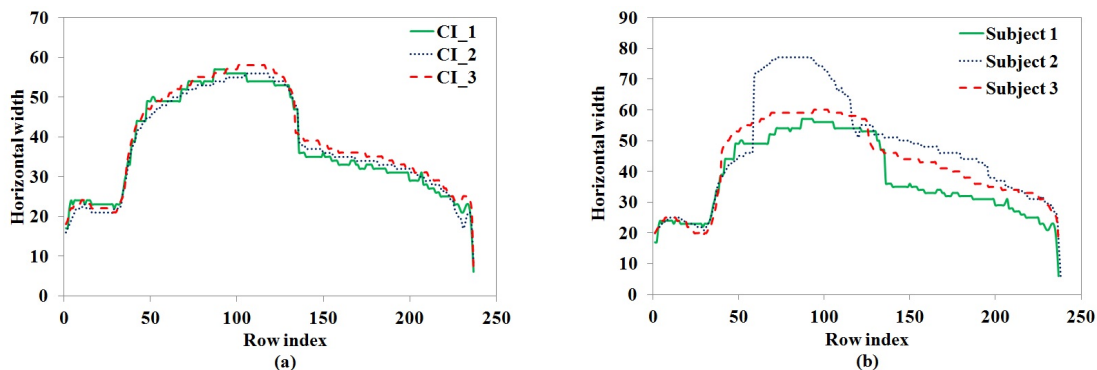
Figure 5.10: Feature space diagram representing (a) horizontal width vector of three CI's belonging to the same subject (b) horizontal width vector of single CI belonging to three different subjects.
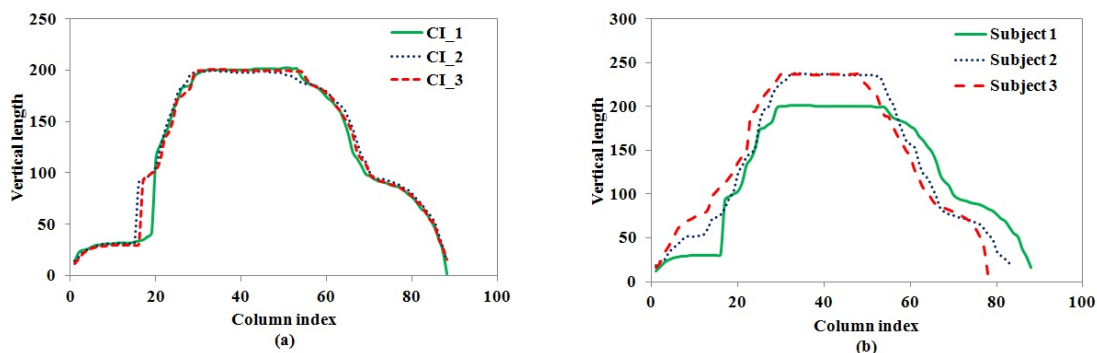
Figure 5.11: Feature space diagram representing (a) vertical length vector of three CI's belonging to the same subject (b) vertical length vector of single CI belonging three different subjects.
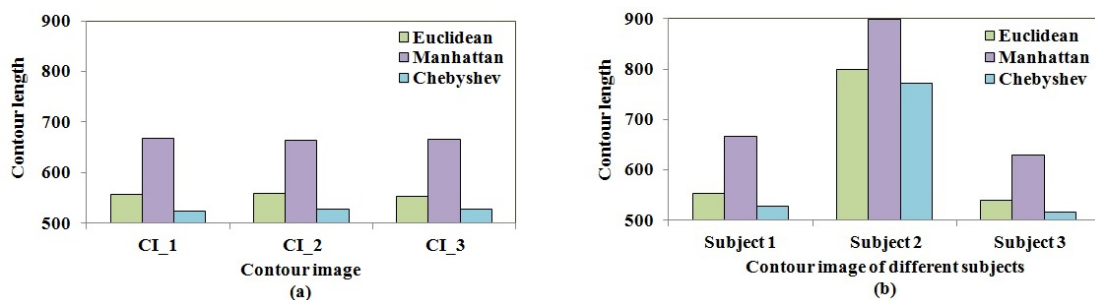
Figure 5.12: Feature space diagram representing contour length vector of three (a) CI's belonging to same subject (b) CI's belonging to different subjects.
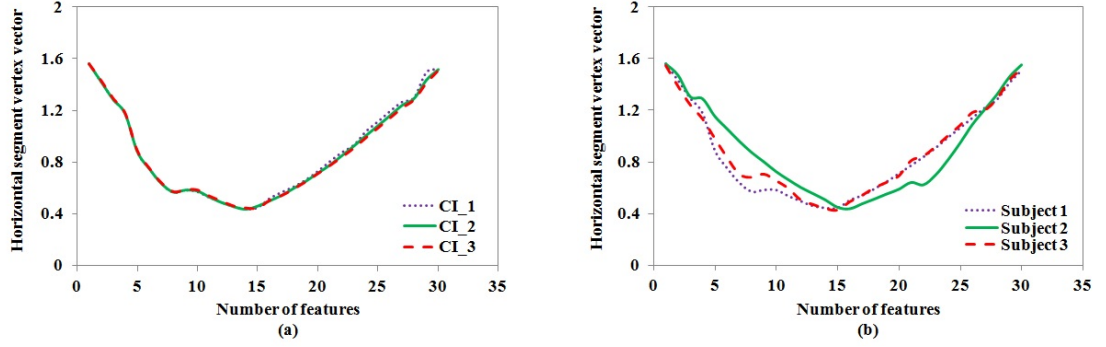
Figure 5.13: Feature space diagram representing horizontal segment vertex vector of three (a) CI's belonging to same subject (b) CI's belonging to different subjects.

Let $V(x_t, y_t)$ represent the vertices of the contour, where $t = 1, 2, .., d, .., n$. The distance between two consecutive vertices $V(x_d, y_d)$ and $V(x_{d+1}, y_{d+1})$ can be obtained by three distance algorithms: (1) Euclidean: $E_d = \sqrt{(x_d - x_{d+1})^2 + (y_d - y_{d+1})^2}$, (2) Manhattan: $M_d = |x_d - x_{d+1}| + |y_d - y_{d+1}|$, and (3) Chebyshev: $C_d = max\{|x_d - x_{d+1}|, |y_d - y_{d+1}|\}$

The contour length vector (C_L) is obtained as the sum of the distances between all the consecutive vertices. It is given by $C_L E = E_1 + E_2 + ... + E_{n-1}$, $C_L M = M_1 + M_2 + ... + M_{n-1}$, and $C_L C = C_1 + C_2 + ... + C_{n-1}$.

3. Horizontal segment vertex vector

This vector is proposed in the horizontal mode to extract the identically distributed and independent features from each segment. Consider a CI of size $p \times q$, where the values of p and q are equal. The CI is horizontally segmented into G equal sized segments. Let the number of rows present in each segment be w. Then the horizontal segmentation is given by, $G(i) = CI((w \times (i-1) + 1 : w \times (i-1) + w), 1 : q)$, where i ranges from 1 to p/w.

The features are computed for each horizontal segment G_i, where i=1,2,...,15. For a segment G_k, two contour lines, C_{k_Left} and C_{k_Right} are obtained. let the vertex which is present at the middle of these contours be $V_{kL}(x, y)$ and $V_{kR}(x, y)$. The arctangent of the specified x and y coordinates of the extracted vertices $V_{kL}(x, y)$ and $V_{kR}(x, y)$ is computed. The arctangent of any vertex $V(x, y)$ is given by $V_\theta = atan2(x, y)$.

The series of these 30 vertex features extracted from 15 segments form a horizontal

95

Table 5.4: The size of the feature vector used for classification.

Sl. No.	Feature name	Size of feature vector
1	Region information set vector	307
2	Contour length vector	3
3	Horizontal segment vertex vector	30
Total		340

segment vertex vector,

$$H_v = \begin{bmatrix} V_{\theta 1L} & V_{\theta 1R} & ... & V_{\theta 15L} & V_{\theta 15R} \end{bmatrix} \tag{5.10}$$

5.2.4 Experiments

5.2.4.1 Experimental Setup

At first, to obtain H_v, the arctangents of all the contour vertices are obtained. Hence, the length of H_v is equal to the number of vertices. However, the disadvantage of this method is that it increases the computational complexity since the feature vector consists of a large number of values (around 676 to 708). It is also evident from the extracted vector that the difference between two consecutive values in the feature vector is not substantial (between the range 0.005 to 0.008), resulting in more redundant features.

To overcome the above problem, the experiment is conducted by dividing the CI into 25, 20, and 15 segments, followed by extracting two features from each segment. However, the encouraging results are obtained when it is divided into 15 segments, with a difference of 0.015 between two consecutive features. This not only reduces the size of the feature vector H_v significantly, but it also facilitates the extraction of identically distributed spatial features.

The four feature space diagrams demonstrate the discriminative ability of the extracted feature vector. Figures 5.10 and 5.11 show the features extracted from the horizontal width and vertical length vector. Figures 5.12 and 5.13 shows the features obtained from the contour length vector and the horizontal segment vertex vector.

In all the four Figures (5.10 to 5.13), the first diagram (a) displays the feature vector extracted from three different CI's of the same subject and the second diagram (b) displays the features extracted from a single CI of three subjects. It is evident from

Table 5.5: Comparision of recognition accuracies of the proposed method with existing methods on CMU MoBo dataset.

Method	b/f	b/i	b/s	f/b	f/i	f/s
Huang et al. (2016)	88	88	96	92	84	96
Proposed method (P 5.2)	**94**	**92**	**100**	**100**	**96**	**100**

Method	i/b	i/f	i/s	s/b	s/f	s/i
Huang et al. (2016)	75	88	92	96	96	92
Proposed method (P 5.2)	**92**	**92**	**100**	**96**	**100**	**98**

the feature space diagram that the minute variations are observed from three gait cycles for the same subject (CI_1, CI_2, CI_3), and significant variations are observed from the gait cycles of three subjects (Subject 1, Subject 2 and Subject 3). The separation of the curves in Figures 5.10(b), 5.11(b), 5.12(b) and 5.13(b) results in a considerable distinction between inter-class distances. Simultaneously, the small intra-class distances in Figure 5.10(a), 5.11(a), 5.12(a) and 5.13(a) demonstrates the discriminative ability of the features under consideration. It can be observed from the feature space diagram that the curves of the CI's of the same subjects appear to be of a similar shape. This ensures that the extracted features are discriminative.

The length of the region information set vector found for each contour image differs based on the variations in the spatial data. However, the number of values detected for different CI varies within a small range. In order to retain the same number of feature values for all CIs in the CASIA B gait dataset, resampling of the vector is done by using linear interpolation. Here, the minimum length and the maximum length of the region vector are detected across all CI in the dataset. The minimum and maximum lengths of the vectors obtained are 298 and 315, with a difference of 17, and the average is 306.5. So, the region vector is resampled to size 307. Hence, a feature vector of size 307 is obtained for each contour image. The size of the three different features and the total size of the concatenated feature vector used for classification are shown in Table 5.4. The size of the contour length vector and horizontal segment vertex vector remain the same for all CI's. The same procedure is followed for CASIA A and CMU MoBo gait dataset to obtain identical feature vectors for all subjects present in a particular dataset.

GEI and CI used in this work from CASIA A, CASIA B, and CMU MoBo dataset are of size 240×240. The CI is divided into 15 horizontal segments to obtain H_v. Each horizontal block is of the size 16×240. The three different feature vectors are

Table 5.6: Comparision of average CCR of four individual feature vectors with the concatenated vector.

Database	View	H_s	V_s	C_L	H_v	F
CASIA A	0^0	91.38	94.21	82.56	90.11	**100.0**
CMU MoBo	b/f	72.00	84.00	62.00	72.00	94.00
CMU MoBo	i/f	78.00	80.00	58.00	70.00	92.00
CMU MoBo	s/b	88.00	90.00	72.00	76.00	96.00
CMU MoBo	s/i	84.00	92.00	72.00	78.00	98.00
CMU MoBo	b/i	78.00	86.00	64.00	64.00	92.00
CMU MoBo	b/s	84.00	94.00	78.00	82.00	100.0
CMU MoBo	f/b	82.00	92.00	78.00	80.00	100.0
CMU MoBo	f/i	76.00	82.00	70.00	78.00	96.00
CMU MoBo	f/s	80.00	90.00	76.00	82.00	100.0
CMU MoBo	i/b	72.00	80.00	70.00	74.00	92.00
CMU MoBo	i/s	74.00	82.00	68.00	76.00	92.00
CMU MoBo	s/f	86.00	92.00	74.00	86.00	100.0
CASIA B	0^0(nm/nm)	83.56	92.50	64.67	73.66	99.60
CASIA B	0^0(nm/bg)	86.23	90.44	68.00	76.10	98.56
CASIA B	0^0(nm/cl)	86.11	95.00	65.19	80.97	99.11
CASIA B	180^0(nm/nm)	88.79	90.98	68.06	75.77	99.00
CASIA B	180^0(nm/bg)	79.01	89.47	56.97	71.42	96.00
CASIA B	180^0(nm/cl)	85.58	91.76	56.23	79.06	97.70
Average		81.82	88.86	68.61	77.11	**96.94**

derived from the contour image and contour vertices of the training and testing datasets. Eventually, the gait features proposed are then tested by KNN classifier with Euclidean distance measure. The value of K is set to 1 to find the immediate best match.

5.2.4.2 Experimental Results and Discussion

To verify the performance of the proposed feature extraction method, three widely used gait datasets are employed in this empirical study, namely, (1) CASIA A dataset, (2) CASIA B dataset, and (3) CMU MoBo dataset. The term P 5.2 is used to specify the proposed method in comparison tables.

1. Experiments on CASIA A dataset

The performance of the proposed approach on CASIA A dataset for 0^0 view is shown in Table 5.6. Out of the four sequences, 3 sequences are considered for training, and one sequence for testing. Table 5.6 shows that the proposed method yields 100% CCR for 20 subjects.

2. Experiments on CMU MoBo dataset

In this study, all the four forms of walking are used for both gallery and probe datasets, where *f/b* represents the gallery *f* and the probe *b*. Table 5.5 shows the re-

Table 5.7: The CCR of the proposed method on CASIA B gait dataset.

Experiment	Gallery set	Probe set	Gallery Size	Probe Size	CCR (%)
1	$0^0(nm)$	$180^0(nm)$	124×4	124×4	93.59
2	$0^0(nm)$	$180^0(nm)$	124×4	124×3	93.76
3	$0^0(nm)$	$180^0(nm)$	124×4	124×2	95.15
4	$0^0(nm)$	$180^0(nm)$	124×4	124×1	95.60
5	$0^0(nm)$	$180^0(bg)$	124×4	124×2	96.79
6	$0^0(nm)$	$180^0(bg)$	124×4	124×1	97.81
7	$0^0(nm)$	$180^0(cl)$	124×4	124×2	94.40
8	$0^0(nm)$	$180^0(cl)$	124×4	124×1	94.96
9	$180^0(nm)$	$0^0(nm)$	124×4	124×4	91.19
10	$180^0(nm)$	$0^0(nm)$	124×4	124×3	93.57
11	$180^0(nm)$	$0^0(nm)$	124×4	124×2	93.97
12	$180^0(nm)$	$0^0(nm)$	124×4	124×1	94.17
13	$180^0(nm)$	$0^0(bg)$	124×4	124×2	97.59
14	$180^0(nm)$	$0^0(bg)$	124×4	124×1	97.41
15	$180^0(nm)$	$0^0(cl)$	124×4	124×2	97.17
16	$180^0(nm)$	$0^0(cl)$	124×4	124×1	98.50

sults of the proposed method on the CMU MoBo dataset for the frontal view. From Table 5.5, it is clearly evident that the proposed method has accomplished the correct CCR in most of the cases and has shown appealing outcomes. It outperforms all the other methods among almost all types of gallery/probe combinations on the CMU MoBo dataset. Apart from cases involving changes in walking speed, *f/s* and *s/f*, the experiments are conducted on the more challenging cases involving walking with a ball in hand and walking on an inclined plane. The proposed method presents an improved recognition accuracy of 6.41%.

3. Experiments on CASIA B dataset

Two experiments were performed on CASIA B dataset.

At first, the first four sequences of *nm* are used for training. For testing, the other two sequences of *nm*, *cl*, and *bg* are employed to assess the normal, clothing, and carrying variations, respectively. Since the GEI images of view 0^0 and 180^0 are similar to a great extent, experiments are conducted on both 0^0 and 180^0 views to evaluate the robustness of the proposed method. The experimental results given in Table 5.8 confirm that in contrast to the past literature findings, the proposed method performs significantly superior to the other methods. The results also show its capability in handling the carrying and clothing dissimilarities.

Secondly, a training dataset containing *nm* gait sequences of 0^0 and 180^0 viewpoint are constructed. For the training process, four *nm* gait sequences of each subject are assigned. This has led to 496 gait patterns in the training dataset. The testing dataset consists of the different number of CI's at each experiment, as shown in Table 5.7. The performance of the proposed method for 16 experiments is illustrated in Table 5.7. The experimental results demonstrate that the proposed method has captured the shape dynamics in the frontal gait images effectively.

Table 5.6 shows the CCRs of CASIA A, CASIA B, and CMU MoBo datasets with all the four feature vectors separately and in concatenation, i.e., feature vector F. It is observed that the feature vector F has more discriminative power than that of the individual feature vectors. Among the four feature vectors, the highest average CCR is obtained for V_s, and less average CCR is obtained for C_L. An average of 8.08% increase in recognition accuracy is observed due to the feature vector F as compared to the separate accuracies of the four feature types. Since the proposed feature vector is susceptible to significant shape variations in a gait sequence, it gives the highest recognition accuracy. Table 5.9 demonstrates the performance of the proposed method in comparison with some of the state-of-the-art contour-based gait recognition methods. The experimental results present in Table 5.9 confirm that the proposed method performs considerably better than other contour-based methods reported in the literature.

Table 5.8: Performance of the proposed method and the existing methods on CASIA B gait dataset.

	0^0	0^0	0^0	180^0	180^0	180^0
Methodology	*nm/nm*	*nm/bg*	*nm/cl*	*nm/nm*	*nm/bg*	*nm/cl*
Dupuis et al. (2013)	97.17	73.15	81.64	99.60	74.56	82.70
Choudhury and Tjahjadi (2015)	100.0	93.00	67.00	99.00	89.00	66.00
Rida et al. (2016b)	97.97	72.76	80.49	97.58	76.11	83.06
Alotaibi and Mahmood (2017)	90.67	91.98	88.77	83.99	87.76	90.00
Isaac et al. (2017)	98.50	95.00	97.00	98.99	94.44	93.94
Proposed method (P 5.2)	**99.60**	**98.56**	**99.11**	**99.00**	**96.00**	**97.70**

5.3 Summary

In Section 5.1, the method which increases the performance of the frontal gait recognition is proposed using the two-step procedure. The first step makes use of a shape

Table 5.9: Comparision of CCR of the proposed method with the existing contour-based methods on CASIA A and CASIA B gait dataset.

Methodology	0^0 (*nm*) CASIA A	0^0 (*nm*) CASIA B
Wang et al. (2002)	88.75	72.14
Ye and Wen (2006)	92.25	83.37
Lee et al. (2013)	97.75	97.39
Liu et al. (2011)	**100.0**	98.99
Proposed method (P 5.2)	**100.0**	**99.60**

descriptor based on hierarchical centroid to extract gait features. The second step called similarity measurement is used to assign the probe sample to a set of gallery samples. An extensive experimentation on the three gait datasets shows the efficiency of the proposed method as it performs better than several existing approaches in the literature. The overall experimental results demonstrate that the proposed method is susceptible to significant spatial variations in gait sequences. Further, the experiments on more extensive and varied datasets need to be done, and the research can be directed at extending the proposed gait recognition method to obtain high recognition accuracy with variations in clothing and carrying conditions.

Section 5.2 proposes a feature extraction approach to increase the performance of the frontal gait recognition. The proposed gait features are acquired from the contour image and contour vertices. These are, in turn, derived from the GEI, possessing both spatial and temporal information. The key contribution of this study is in the proposition of three feature vectors, which are sensitive to significant spatial changes in gait sequences. The experimental results show that the gait recognition performance can be increased with the combination of all the three feature vectors rather than using the individual feature vectors. An extensive experiment on the three gait datasets has demonstrated the effectiveness of the proposed features. The overall experimental results vindicate that the proposed method surpasses some of the existing approaches in the literature on the gait recognition.

Chapter 6

SPEED INVARIANT GAIT RECOGNITION

When an individual changes the speed of walking, dynamic features such as joint angles and stride length change, whereas static features such as the area of head and torso region, remain unchanged (Kovač et al. 2017). Hence, the spatial data is predominantly used to extract discriminative gait features in speed invariant gait recognition. Since GEI carries both static and dynamic information, many researchers prefer gait silhouette images instead of GEI for feature extraction (Medikonda et al. 2018). This is because the dynamic information present in GEI changes with the speed of a person. Here, the disadvantage is that the use of all silhouette images to extract the gait features increases the computational cost. The two proposed methods overcome this setback, and they extract the most relevant spatial regions from GEI to increase the performance of the speed invariant gait recognition system. Two publicly available gait datasets, CASIA C and OU-ISIR A, are used for demonstrating the performance of the proposed approaches.

6.1 Speed Invariant Gait Recognition using Mutual Information

This method overcomes the setback mentioned above and extracts the most relevant regions from GEI, called ROI, and the features are extracted from these regions.

The significant contributions of this study are summarized as follows.

1. The spatial information from GEI is obtained to derive gait features by using significant and least affected ROI, enabling low computational complexity.

2. MI between two images is computed as it captures the spatial variations of gait efficiently.

3. The performance of the proposed method is evaluated on two benchmark gait datasets, and the results are compared with those of the state-of-the-art gait recognition approaches.

6.1.1 Framework of the Proposed Method

Figure 6.2 shows an outline of the proposed feature extraction method. It consists of four stages. In the first stage, the gait video is converted into gait silhouette images.
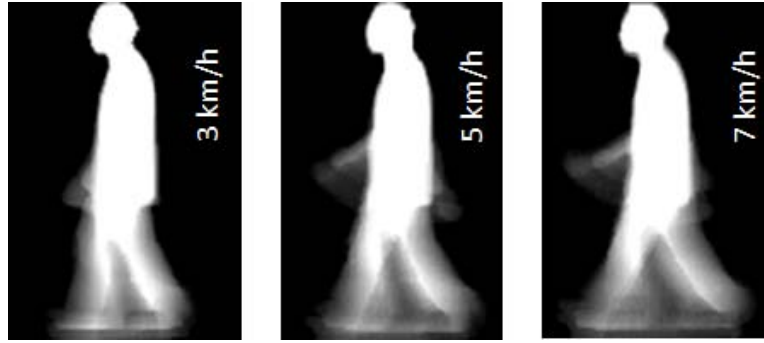
Figure 6.1: Images representing GEI of the same subject for different speed variations.

These gait silhouettes are combined over a gait cycle to form GEI. Secondly, ROIs, which are considered to be less affected by the speed variations, are selected. In the third stage, the MI, which is considered to be an efficient measure for comparing images, is computed from the ROIs of the gallery and probe datasets. Finally, the classification is carried out to obtain the recognition accuracies on different datasets.
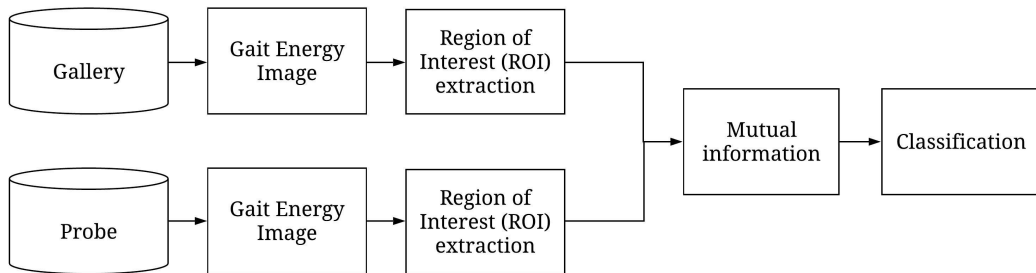
Figure 6.2: Steps of the gait recognition framework.

6.1.2 ROI Extraction

The GEI carries both static and dynamic information; the dynamic information present in GEI changes with the speed of the person. Hence, GEI of the same subject varies when that subject is walking at different speeds, as shown in Figure 6.1. When these GEIs are used for the comparison between the probe and gallery samples, there is a possibility of misclassification that results in the reduction of the recognition performance. To overcome this setback, the ROI's which are less influenced by the speed variations are extracted from GEI.

To extract ROI, the following procedure has been carried out. Let GEI_1 and GEI_2 be the two images with different speed variations. The difference image D is computed
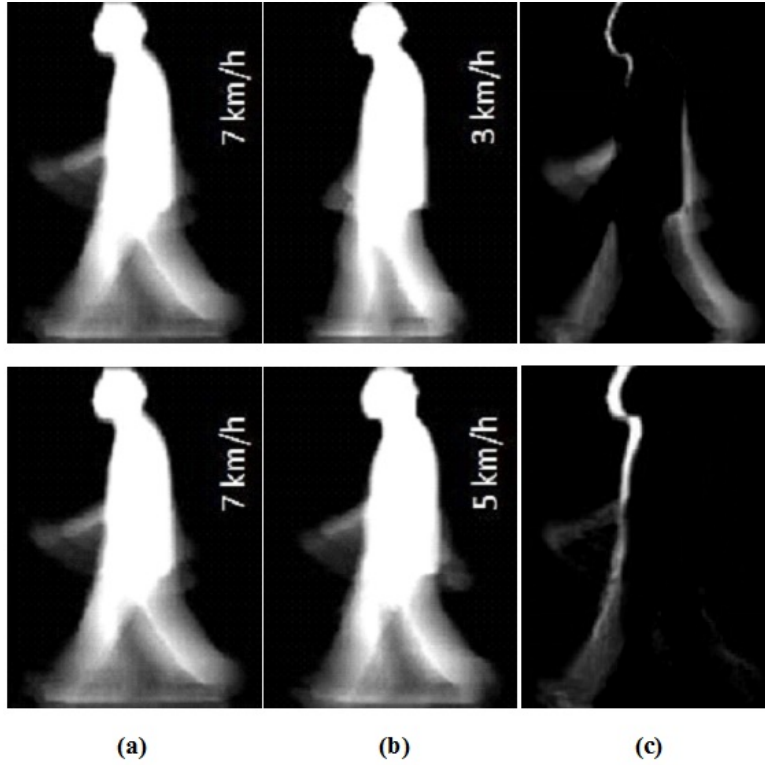
Figure 6.3: Sample images representing (a) GEI of a subject walking with a speed of 7km/h (b) GEI of a subject walking with a speed of 3km/h and 5km/h (c) difference image.

as follows:

$$D = GEI_1 - GEI_2 \tag{6.1}$$

D is an image which is obtained by subtracting each element in GEI_2 from the corresponding element in GEI_1. Figure 6.3 shows the sample of a difference image.

An analysis of the difference image reveals the dissimilarity between two images in some of the regions. Further, the horizontal width vector from D is computed, and this provides an account of the number of pixels present in each row. The horizontal width is calculated in terms of the number of pixels, other than zeros present in each row of the ROI, followed by maintaining the row-wise count. Figure 6.5 displays the numerical calculation of the horizontal width vector.

When the difference image and the horizontal width vector are computed for GEI samples of 10 subjects with speed variations of 2km/h, 3km/h, 4km/h, 5km/h, 6km/h, and 7km/h, it is observed that the values in the horizontal width vector are high for the middle regions and the lower part of the GEI, indicating more number of pixels in that

region. Hence, these regions are considered to be more influenced by speed variations and therefore removed. The remaining parts are considered as ROIs used for feature extraction and classification. Figure 6.6 shows the horizontal width vector obtained for the difference image of 2 different speeds, and Figure 6.4 shows the 2 ROIs extracted from GEI.
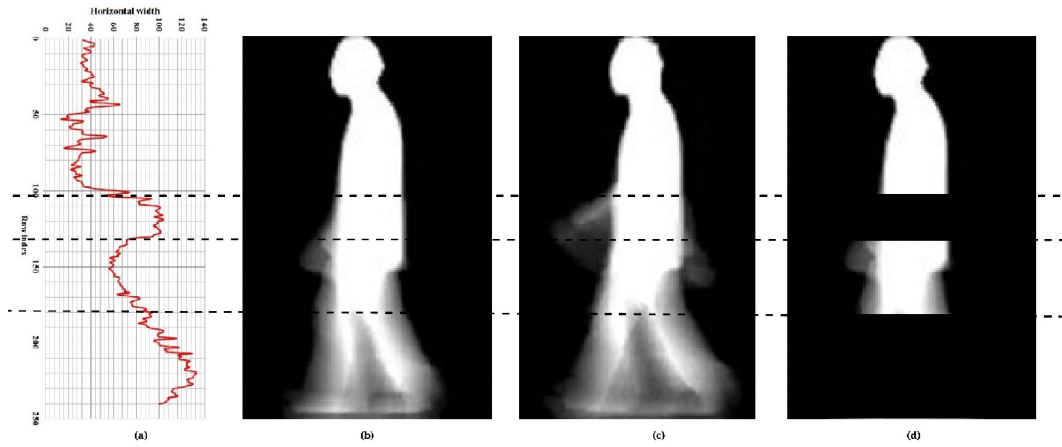
Figure 6.4: Sample image representing (a) feature space diagram of difference image, $\{(b)-(c)\}$ and (d) Extracted ROIs.

6.1.3 Mutual Information

The MI between two images calculates the amount of information that one image has about another image (Pluim et al. 2003). It is considered as a similarity measure between two images. It is obtained from the individual and joint entropies of two images.

Consider two images X and Y, the joint entropy $E(X,Y)$ can be computed as follows:

$$E(X,Y) = -\sum_{x,y} P_{XY}(x,y) \log P_{XY}(x,y) \tag{6.2}$$

where P_{XY} is the joint probability distribution of pixels related to images X and Y.

Let $E(X)$ and $E(Y)$ be the individual entropies of X and Y, respectively. They are given by $E(X) = -\sum_x P_X(x) \log P_X(x)$ and $E(Y) = -\sum_y P_Y(y) \log P_Y(y)$.

The mutual information, $MI(X,Y)$, of two images X and Y can be obtained from the following equation: $MI(X,Y) = E(X) + E(Y) - E(X,Y)$.

Difference image

Horizontal width

0	1	5	0	0
0	2	6	1	0
0	1	1	5	8
0	0	16	0	0
0	0	0	0	0

2
3
4
1
0

Figure 6.5: Example to show the numerical calculation of horizontal width vector.

6.1.4 Classification

The steps followed in this module are shown in Algorithm 12. In this section, the MI between each probe sample p and all gallery samples is computed. The maximum value of MI obtained for probe p and gallery g indicates that p and g have similar information to a great extent.

Algorithm 12 Classification

Input: Collection of all gallery samples, a probe sample.
Output: A subject.

1: Begin
2: Let the number of subjects present in the gallery dataset be $S = \{s_1, s_2, ..., s_n\}$.
3: Each subject persent in the gallery dataset consists of t number of GEI's, $s_i = \{g_i^1, g_i^2, ..., g_i^t\}$.
4: **for** each $g_i^j \in S$ **do** where $i = \{1, 2, ..., n\}$ and $j = \{1, 2, ..., t\}$
5:     Compute the MI between p and g_i^j.
6: **end for**
7: Let $MI_i = \{mi_i^1, mi_i^2, ..., mi_i^t\}$ be the MI's between p and gallery samples belonging to subject s_i.
8: Compute the mean of the MI's between p and MI_i and then obtain $A_i = (mi_i^1 + mi_i^2 + ... + mi_i^t)/t$.
9: The above step results in the list of MI's for each subject, $A = A_1, A_2, ..., A_n$.
10: Obtain the maximum value present in vector A.
11: If A_m is the maximum value present in vector A, representing subject s_m, then the probe sample p is assigned to s_m.
12: End

Here, the mean of MI's of all the gallery samples belonging to a particular subject and a probe p is considered for classification, instead of the maximum MI between a probe sample and a gallery sample. This reason is that, when a probe sample and gallery
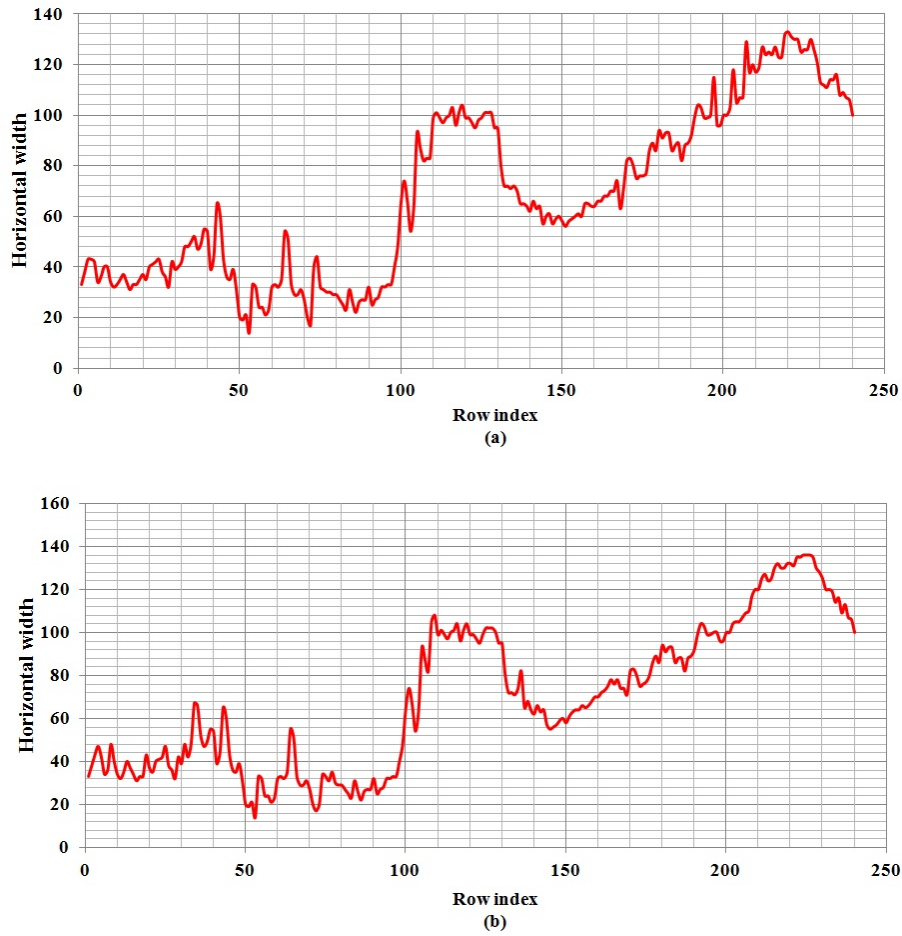
Figure 6.6: Feature space diagram representing horizontal width vector of (a) difference image obtained with GEI_1=7km/h and GEI_1=3km/h (b) difference image obtained with GEI_1=7km/h and GEI_1=5km/h.

sample belong to different subjects, they substantially have the same spatial data. In this case, suppose if the individual gallery samples are considered, the MI between a few gallery samples and a probe sample could be more or less similar. This, in turn, results in a decrease in the recognition performance. Moreover, the mean of the MI's belonging to a particular subject gives different values. This enables the maximum value to be selected for increasing the recognition performance.

6.1.5 Experiments

6.1.5.1 Experimental Results and Discussion

The performance of the proposed method is evaluated using two publicly accessible gait datasets. They are (1) CASIA C dataset, and (2) OU-ISIR A dataset. The performance of the gait recognition system is measured by the CCR on the testing dataset. The term

Table 6.1: CCR (%) of the proposed method on CASIA C gait dataset and comparision with other existing methds.

Gallery/Probe	*fn/fn*	*fn/fs*	*fn/fq*	*fs/fq*
FDEI + 2DLDA (Chen et al. 2009)	88	89	90	–
WBP (Kusakunniran et al. 2009)	**99**	86	90	60
HTI (Tan et al. 2006)	94	85	88	–
Pseudoshape (Tan et al. 2007b)	98	82	92	–
Orthogonal projections (Tan et al. 2007a)	98	80	80	–
NDDP (Tan et al. 2007d)	97	85	74	–
HOSC (Kusakunniran et al. 2012)	97	92	93	89
2FInS (Medikonda et al. 2018)	**99**	97	**96**	90
Proposed method (P 6.1)	**99**	**98**	**96**	**94**

P 6.1 is used to specify the proposed method in comparison tables.

1. Experiments on CASIA C dataset

The dataset consists of four gait cycles from a normal sequence and two gait cycles from the slow and fast sequences respectively. Among these, two normal sequences are considered for training, and two gait cycles from each walking speed are considered for testing. Hence, while the probe feature sets contain *fn*, *fs*, and *fq*; the gallery feature comprises *fn*. The performance of the proposed approach on CASIA C gait dataset and its comparison with the state-of-the-art methods are given in Table 6.1 where in most cases, the proposed method performs better than the methods compared.

2. Experiments on OU-ISIR A dataset

This study presents the results on OU-ISIR A dataset. This study exclusively analyzes the walking category.

Two experiments are conducted on this dataset. Firstly, six gait cycles are considered for the gallery data, and three gait cycles for the probe data. This is for the speed variations from 2 to 7 km/h for all the combinations of probe and gallery sequences. The experiments are conducted for two scenarios: (1) Without ROI extraction (for entire GEI), and (2) With ROI extraction. Table 6.2 gives the CCR obtained by the proposed method on the OU-ISIR A dataset for the entire GEI. The total average CCR obtained is 76.53%, which is calculated by adding all CCRs present in the table together and dividing it by the total number of CCRs. The CCR obtained for each probe is in the range of 73.78% to 79.69%. Table 6.3 shows the CCR obtained by the proposed

Table 6.2: CCR (%) across various walking speeds on OU-ISIR A gait dataset using the proposed method (Without ROI extraction).

Gallery/Probe	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	84.12	81.00	74.06	71.47	72.00	71.53
3 km/h	82.77	88.00	80.71	79.00	74.00	69.00
4 km/h	83.53	80.47	86.00	78.94	76.00	73.24
5 km/h	72.06	72.00	75.00	82.12	78.24	71.12
6 km/h	69.94	68.24	72.18	79.53	83.41	83.00
7 km/h	68.35	66.41	69.00	73.12	79.94	85.88

Table 6.3: CCR (%) across various walking speeds on OU-ISIR A gait dataset using the proposed method (With ROI extraction).

Gallery/Probe	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	97.94	93.41	85.94	85.53	83.53	83.00
3 km/h	96.53	99.94	91.24	92.06	85.71	81.41
4 km/h	95.00	93.12	97.00	92.00	88.35	85.40
5 km/h	83.00	86.06	87.00	95.53	90.00	84.18
6 km/h	82.00	82.88	85.24	90.00	96.12	96.00
7 km/h	80.00	79.53	83.94	84.47	90.41	96.94

method on OU-ISIR A dataset for the segmented GEI (only for ROI). The total average CCR obtained is 88.88%, and that of each probe is in the range of 85.88 to 91.81%. The experimental results manifest that the extraction of ROI increases the recognition results because the regions having large variations within the same subject are filtered out, and only the relevant regions are retained.

Secondly, the subjects with a speed range of 4 km/h to 7 km/h are considered as probe samples, and subjects with a speed range of 5 km/h are considered as gallery samples. Table 6.4 makes a comparative analysis of some methods with the proposed method. Nevertheless, this study demonstrated that the proposed method performs better than the other methods in Table 6.4 as it gives a high average recognition accuracy (88.60%). The proposed method gives promising recognition accuracy for both large and small speed changes.

Table 6.4: CCR (%) in different walking speeds with gallery speed of 5 km/h on OU-ISIR A gait dataset.

Method	4 km/h	5 km/h	6 km/h	7 km/h	Average
MHI-HOG (Huang et al. 2011)	70.60	**100.0**	88.20	73.50	83.08
FD (Lee et al. 2013)	76.50	78.00	91.20	79.40	83.10
GII (Arora et al. 2015a)	76.40	88.00	85.20	–	85.90
2FInS (Medikonda et al. 2018)	88.24	**100.0**	91.18	73.53	88.24
Proposed method (P 6.1)	**90.90**	97.50	**91.24**	**76.00**	**88.60**
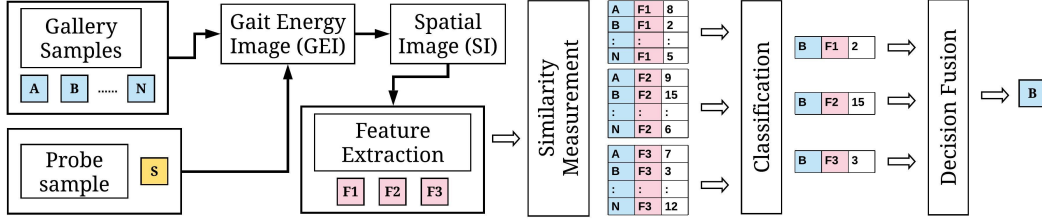
Figure 6.7: Framework of the proposed approach.

6.2 Speed Invariant Gait Recognition using Spatial Dynamics

The contributions of this method are summarized as follows:

1. The spatial information from GEI helps derive the discriminative metrics of gait by using the marching squares algorithm and linear interpolation, thus enabling low computational complexity.

2. The three metrics related to the similarity/dissimilarity measures for each probe sample capture the spatial variations of gait efficiently.

3. The proposed decision fusion module increases the performance of the gait recognition system.

6.2.1 Framework of the Proposed Method

An outline of the proposed method is shown in Figure 6.7. At first, all gallery GEI and probe GEI samples are converted into Spatial Images (SI). Let $p_i = p_1, p_2, ..., p_m$ and $g_i = g_1, g_2, ..., g_n$ represent the probe and gallery samples. In the proposed method, for each $p_i \in p$, a set of three features is extracted when a probe sample p_i is compared with each gallery sample g_i. As a result, the features are extracted from the combinations of each probe sample with all gallery samples. This results in a similarity measurement list for each probe sample, depicting its similarity/dissimilarity with each gallery sample. Subsequently, classification and decision fusion are performed to assess the performance of the proposed gait recognition system, which is measured by the CCR on the probe dataset.

6.2.2 Spatial Image Extraction

GEI contains both static and dynamic information. Therefore, many researchers prefer gait silhouette images instead of GEI for feature extraction as the dynamic information present in GEI may increase misclassification (Nandy et al. 2014b; Makihara et al.

110
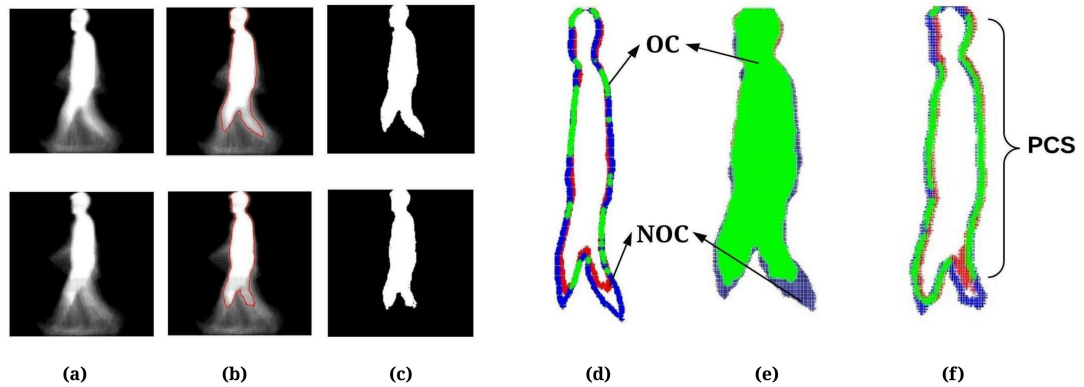
Figure 6.8: Sample images from CASIA B gait dataset representing (a) GEI of a p_s and g_s, where both belong to different subjects (b) Contour generated from GEI using marching squares algorithm and linear interpolation (c) Spatial image (d) OC (green colour) and NOC (red and blue colour) obtained from the contour of a p_s and g_s (e) OC and NOC obtained from the spatial image of a p_s and g_s (f) PCS, when p_s and g_s belong to the same subject.

2014; Medikonda et al. 2018). The disadvantage of using all silhouette images is that they increase the computational cost of the system.

To overcome this setback, only the static information is extracted from GEI, and the features are extracted from this static information. To accomplish this, the contour of GEI, consisting of accurate spatial information, is extracted. The marching squares algorithm (Maple 2003) is used to generate the contour for a GEI image, with isovalue = 0.6, and the linear interpolation is applied to determine the exact contour position. The next step is to sort the contour geometry into the separate sorted contours. As a result, a list of coordinates $C(x, y)$ is obtained. These coordinates are used to get a spatial image $SI(x, y)$, as shown in Figure 6.8(c). Some other small contours are also detected in some of the images in dynamic regions. But they are neglected. The value of 0.6 is chosen after experimenting with many GEI templates.

6.2.3 Feature Extraction

When attempting to determine whether two images belong to the same individual, one typically searches for similarities between the two images and rejects any deceptive dissimilarities. As such, to find the general closeness of two images, an attempt has been made to locate the most comparative pieces of them. The three features extracted based on the similarity/dissimilarity between a probe sample, p_s and a gallery sample, g_s are explained as follows:

1. Intensity similarity metric (M_{is})

This metric provides a quantitative evaluation of the similarity between two image regions. Let $p(x_c, y_c)$ and $g(x_c, y_c)$ denote the pixels of p_s and g_s of size $m \times n$ with intensities p_c and g_c respectively, where $c = 1, 2, ..., l$. Then the M_{is} is calculated as $\sum_{c=1}^{l} (p_c - g_c)^2$. That is, the sum of the squares of differences of the corresponding intensity values in two images.

This method gives a measure of intensity variations between p_s and g_s. Lastly, the g_s with least M_{is} is considered as the candidate for the p_s.

2. Contour similarity metric (M_{cs})

Here, the aim is to find a measure of similarity between the contours of a probe sample p_s and a gallery sample g_s. Let C_p and C_g be the contours of p_s and g_s. The first step is to find the pairs of coordinates of C_p and C_g that are overlapping (distance between the corresponding coordinates is equal to 0) by using the Euclidean distance as shown in Figure 6.8(d) (The green colour represents the overlapping contour; Red colour represents contour of p_s, and blue colour represents contour of g_s). The next step is to ignore the Non-Overlapping Coordinates (NOC) and consider only Overlapping Coordinates (OC) for further experimentation, as their measure gives an insight into the similarity between p_s and g_s. Further, the consecutive overlapping coordinates will form Piecewise Coordinate Segment (PCS), as shown in Figure 6.8(f). The steps involved in the computation of M_{cs} are given in Algorithm 13.

Algorithm 13 Computation of contour similarity metric, M_{cs}

1: **Input**: Piecewise co-ordinate segments $x[i]$, where $i = 1, 2, ..., k$
2: **for** $i \leftarrow 1{:}k$ **do**
3:      Let (x_c, y_c) be the co-ordinates of a PCS, where $c = 1, 2, .., n$.
4:      Obtain a vector θ_c by computing the arctangent of all the x and y coordinates of the segment, $\theta_c = atan2(x_c, y_c)$.
5:      The total no. of vertices present in the vector is n.
6:      Subtract n from each element of the vector θ_c to obtain vector S_c. That is, $S_c = \theta_c - n$.
7:      Compute root mean square, r_i of S_c,
8: **end for**
9: Compute the sum of the root mean square values obtained for each segment, $M_{cs} = r_1 + r_2 + ... + r_k$

Let $x[i]$ be the number of PCS obtained for p_s and g_s. Considering g_s as a candidate for p_s by counting the number of OC may lead to misclassification, as there is a probability that the number of OC obtained for p_s and some gallery samples might be the same. Hence, the x and y values of the coordinates are converted into a single value using atan2 as given in Algorithm 13. By doing so, the x and y values are converted into a single value whose range is between $-\pi$ and π, and thus, the information regarding the position of coordinates is taken into consideration.

From the empirical study, it is analyzed that when p_s and g_s belong to the same subject, the length n of the PCS is higher, as shown in Figure 6.8(f). Hence, the subtraction of n from each element of PCS (i.e., θ_c), results in a vector S_c consisting of negative values. Further, the root mean square of S_c is calculated to convert the vector into a single positive value.

Besides, as the value of n is high, the difference between n and each element of vector θ_c will be high too. So, the root mean square value obtained for a PCS is also high, further increasing the value of M_{cs}. Finally, the g_s with the highest M_{cs} is considered as the candidate for the p_s.

3. Spatial dissimilarity metric (M_{sd})

Here, the intention is to find a measure of dissimilarity between a probe sample and a gallery sample. Let SI_p and SI_g represent the spatial images of p_s and g_s, respectively, as shown in Figure 6.8(c).

The initial task is to detect the pairs of coordinates of SI_p and SI_g that are non-overlapping (distance between the corresponding coordinates is not equal to 0) by using the Euclidean distance as shown in Figure 6.8(e) (The green colour represents the OC; Red and blue colours represent the NOC).

The subsequent step is to ignore OC and consider only NOC for further experimentation, as their measure gives an insight into the dissimilarity between p_s and g_s. The non-overlapped regions as shown in Figure 6.8(e), are detected using the Euclidean distance since another method of subtracting two images does not give the correct measure of NOC for a combination of p_s and g_s. That is, SI_p-SI_g and SI_g-SI_p are not similar.

Let the image containing only NOC and the corresponding intensity values be N_c.

For each N_c of size $M \times N$, a vector e_v, where $v = 1, 2, ..., M$ is generated by computing the relative entropy of each row. Each element of the vector e_v is given by

$$e_v = \sum_{r=1}^{255} p_r^k \log_2 \frac{p_r^k}{q_r^k} \qquad (6.3)$$

where p_r^k is the number of times that the pixel value r occurs in the kth row of the image N_c and q_r^k is the another probability that the pixel value r occurs in the kth row of image N_c, which is given by

$$q_r^k = \frac{\#(N_c(i,j) = r)}{T} \quad \forall j \in [1, T] \qquad (6.4)$$

Where T is the total number of elements in a row. The spatial dissimilarity metric is obtained from the addition of all elements present in vector e_v. It is given as $M_{sd} = \sum_{v=1}^{M} e_v$.

This entropy provides a consistent measure of the information content of p_r^k with respect to another probability density function q_r^k which is considered as an approximation of p_r^k. Relative entropy is used to quantify the information content of the pixels present in each row of N_c. The lower entropy means that N_c contains less information, which in turn indicates that the dissimilarity between the p_s and g_s is also less. Therefore, the g_s with lowest M_{sd} is considered as the candidate for the p_s.

6.2.4 Similarity Measurement

The outcome of feature extraction leads to three lists for a probe sample p_s, called as similarity measurement lists, namely, $L_{M_{is}}$, $L_{M_{cs}}$, and $L_{M_{sd}}$ containing the metrics for p_s with each gallery sample. These lists are obtained for each probe sample, containing metric related to all gallery samples.

6.2.5 Classification

There are two components in classification. Suppose that $g^i = \{g^1, g^2, ..., g^n\}$ is the vector of elements obtained by sorting the lists $L_{M_{is}}$ and $L_{M_{sd}}$ of probe p_s in the ascending order and $h^i = \{h^1, h^2, ..., h^n\}$ contains the corresponding subjects. With this assumption, it can be concluded that g^1 is the smallest element and h^1 represents the

subject with the smallest element. Then h^1 is considered as the candidate subject for p_s as the least values of M_{is} and M_{sd} signify that the g_s is more similar to p_s.

Suppose that $g^i = \{g^1, g^2, ..., g^n\}$ be the vector of elements obtained by sorting $L_{M_{cs}}$ in the descending order and $h^i = \{h^1, h^2, ..., h^n\}$ contains the corresponding subjects. It can be concluded that g^1 is the largest element and h^1 belongs to the subject with the largest element. Then h^1 is considered as the candidate subject for p_s as the highest value of M_{cs} represents that g_s is more identical to p_s.

6.2.6 Decision Fusion

The final result is obtained from the sorted similarity measurement lists in the decision fusion module. This is done by computing an assurance factor (A_f) for each of the candidates. The candidate with maximum total A_f is qualified for the final result. The A_f of the candidate in each list is obtained based on its metric value and by combining the metric value of the next best subject from the same list.

The range of values obtained for three metrics is different. Hence, before performing decision fusion, all the metric values present in three separate lists are normalized between 0 and 1. Let h^a and h^b be the candidates of the list h^i such that $h^a \neq h^b$. If h^a is the first best and h^b is the next best subject from the list, then the A_f is given by:

$$\text{For lists } L_{M_{is}}, L_{M_{sd}}: \quad A_f = \frac{g^b - g^a}{g^a} \tag{6.5}$$

$$\text{For list } L_{M_{cs}}: \quad A_f = \frac{g^a \times g^b}{(g^a + g^b)/2} \tag{6.6}$$

Since the elements are in the ascending order in lists $L_{M_{is}}$ and $L_{M_{sd}}$, the larger value of $g^b - g^a$ and a smaller value of g^a indicate more confidence in the candidate h^a. Therefore, the subject whose A_f is higher is more likely to be the correct case. Applying Equation 6.5 to the list $L_{M_{cs}}$ will also give more confidence to candidate h^a. Since a subject is the candidate of more than one list, the assurance factors of that subject are added. Lastly, the subject with the highest A_f is chosen as the final result.

Table 6.5: CCR (%) of the proposed method on CASIA C gait dataset and comparision with other existing methods.

Gallery/Probe	*fn/fn*	*fn/fs*	*fn/fq*	*fs/fq*
FDEI + 2DLDA (Chen et al. 2009)	88	89	90	–
WBP (Kusakunniran et al. 2009)	99	86	90	60
Orthogonal projections (Tan et al. 2007a)	98	80	80	–
HOSC (Kusakunniran et al. 2012)	97	92	93	89
Multilayer perceptron (Semwal et al. 2015)	94	80	75	71
GEINet (Shiraga et al. 2016)	95	87	83	80
2FInS (Medikonda et al. 2018)	99	97	**96**	90
Proposed method (P 6.2)	**100**	**100**	96	**96**

Table 6.6: CCR (%) across various walking speeds on OU-ISIR A gait dataset using the proposed method.

Gallery/Probe	2 km/h	3 km/h	4 km/h	5 km/h	6 km/h	7 km/h
2 km/h	100.0	100.0	100.0	99.53	98.00	95.88
3 km/h	100.0	100.0	100.0	100.0	98.12	99.00
4 km/h	98.84	100.0	100.0	100.0	97.00	97.53
5 km/h	97.00	100.0	98.06	100.0	99.24	97.53
6 km/h	96.47	98.12	100.0	99.24	100.0	100.0
7 km/h	96.00	96.35	97.00	100.0	100.0	100.0

6.2.7 Experiments

6.2.7.1 Experimental Results and Discussion

Two publicly available gait datasets are used for demonstrating the performance of the proposed approach. They are (1) CASIA C dataset, and (2) OU-ISIR A dataset. The term P 6.2 is used to specify the proposed method in comparison tables.

1. Experiments on CASIA C dataset

The dataset comprises of four gait cycles for *fn* and two gait cycles for *fs* and *fq*. Out of four normal walking sequences, two normal walking sequences are used for training, and two sequences from each walking speed are used for testing. As a result, the gallery feature set consists of *fn*, and the probe feature sets include *fn*, *fs*, and *fq*. A comparison

Table 6.7: The obtained CCR (%) for different walking speeds with gallery speed of 5 km/h on OU-ISIR A gait dataset.

Method	4 km/h	5 km/h	6 km/h	7 km/h	Average
MHI-HOG (Huang et al. 2011)	70.60	**100.0**	88.20	73.50	82.94
FD (Lee et al. 2013)	76.50	78.00	91.20	79.40	81.27
Multilayer perceptron (Semwal et al. 2015)	85.00	92.00	89.71	81.60	87.07
GEINet (Shiraga et al. 2016)	90.00	94.53	91.50	82.94	89.74
deep-CNN (Alotaibi and Mahmood 2017)	91.12	94.24	89.70	82.00	89.26
2FInS (Medikonda et al. 2018)	88.24	**100.0**	91.18	73.53	88.23
Proposed method (P 6.2)	**98.06**	**100.0**	**99.24**	**97.53**	**98.70**

of the performance of the proposed method on the CASIA C gait dataset with several state-of-the-art methods is shown in Table 6.5. It is observed that the proposed method performs better and achieves the high CCR in a large number of the cases.

2. Experiments on OU-ISIR A dataset

In this study, the experiments and analysis are carried out on the walking category.

Primarily, the gallery data consists of six gait cycles, and the probe data consists of three gait cycles. This is for the speed differences from 2 to 7 km/h for all the combinations of gallery and probe sequences. The CCR obtained by the proposed method on the OU-ISIR A dataset is shown in Table 6.6. The CCR for each probe is in the range of 97.14 to 98.30%, and the total average CCR, which is calculated by adding all CCRs present in the table together and dividing it by the total number of CCRs is 97.84%.

Secondly, the subjects with a speed range of 5 km/h are considered as gallery samples, and 4 km/h to 7 km/h are considered as probe samples. Several methods are considered with the proposed method for comparative analysis, as shown in Table 6.7. The proposed method provides high average recognition accuracy (98.70%) and, therefore, performs best in comparison to different strategies, as shown in Table 6.7.

The performance of the proposed method on CASIA C and OU-ISIR A dataset is compared with that of two methods based on neural networks, such as multilayer perceptron, and GEINet, as shown in Tables 6.5 and 6.7. It is evident from the results that the proposed method performs better than the compared methods based on neural networks. The reason for the same could be that the other approaches are not particularly designed to deal with the effects of speed variation. Hence, the entire gait template is used for the extraction of gait features in other methods. These templates contain speed variations, and they may lead to misclassification when the same individual walks at different speeds. On the contrary, in the proposed method, the spatial images, which are less affected by the speed variations, are first selected, followed by the extraction of the proposed features.

The proposed decision fusion method increases the average recognition accuracy of about 5.24 and 6.07% for the experiments shown in Tables 6.5 and 6.6. This proves that the use of the assurance factor improves the recognition rate when compared to another

Table 6.8: The comparison of the performances of the two proposed approaches.

Dataset	P 6.1	P 6.2
CASIA C	96.75	98
OU-ISIR A	88.60	98.70

method of using a majority of candidates for all probe samples.

Let P 6.1 represent the proposed speed invariant gait recognition method using mutual information, and let P 6.2 represent the proposed speed invariant gait recognition method using spatial dynamics. Table 6.8 compares the average recognition accuracy of the two proposed methods on two benchmark gait datasets. It is evident from the table that P 6.2 performs better for both datasets compared to P 6.1.

In P 6.1, ROIs which are less influenced by the speed variations are extracted from GEI. Whereas, in P 6.2, accurate spatial information consisting of regions not affected by speed variations is extracted. In P 6.1, only one metric called mutual information is used to assign a probe sample to the gallery sample, whereas in P 6.2, the usage of three metrics followed by a decision fusion module is used to assign a probe sample to the gallery sample. Because of the above reasons, P 6.2 performs better when compared with P 6.1 for two datasets.

6.3 Summary

Section 6.1 presents an efficient approach to increase the performance of the speed-invariant gait recognition system. The MI between a probe sample and gallery sample is measured from the ROI's, which is in turn, derived from GEI in order to assign a probe sample to a gallery sample. An extensive experiment on the two widely available gait datasets demonstrate the effectiveness of the proposed approach.

In Section 6.2, an approach is proposed to increase the performance of the speed invariant gait recognition system. The key contribution of this work is in the proposition of similarity/dissimilarity metrics, which not only characterize the gait more appropriately but also sensitive to the substantial spatial changes in a gait sequence. The overall experimental results on the two widely used gait datasets show that the proposed method surpasses some of the existing approaches in the literature.

Chapter 7

HUMAN IDENTIFICATION USING DIFFERENT GAIT PATTERNS

A significant role in human recognition is played by biometrics in recent times. Many researchers across the world are working on gait recognition to recognize the individuals by the pattern of their walking (Lumini and Nanni 2017; Prakash et al. 2018). Using gait as a biometric is motivated by the occlusion of criminal's faces because they either walk or run to getaway from a scene of the crime to escape. On the other hand, the likelihood of identifying the individuals by the way they run remains predominantly unexplored. Very often, robbers and criminals naturally run to escape, instead of walking (Yam et al. 2001). Therefore, it is necessary to come up with a recognition method that distinguishes people by considering their running patterns.

7.1 A Straightforward Approach for Gait Recognition using Running and Jogging Patterns

In this work, an attempt has been made to identify the individuals by using their running and jogging patterns.

7.1.1 Framework of the Proposed Method

The framework of the proposed recognition method is shown in Figure 7.1. Initially, the gait video is converted into the sequence of images, which are then processed by some steps such as background subtraction, de-noising, post-processing, and normalization to get silhouette images. Figures 7.2 and 7.3 show the sample silhouette images of the movement pattern. i.e., running and jogging.

The proposed system includes the following steps:

1. Feature extraction: A set of features, which helps in the identification of an individual is extracted from the silhouette images.

2. Feature selection: A subset of the relevant features is selected.

3. Classification: Decision about the recognition of a person is made using KNN classifier with the selected feature set.
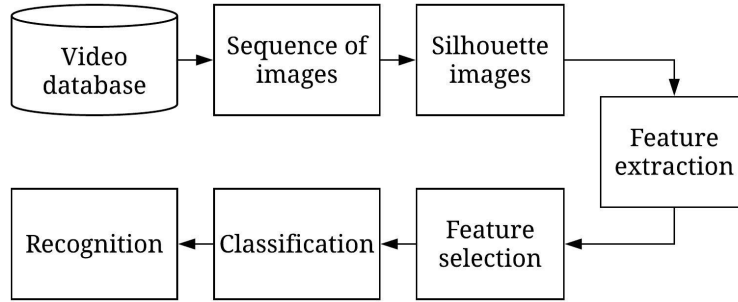
Figure 7.1: Structure of the proposed approach

Figure 7.2: Sample silhouette images from KTH running dataset.

7.1.2 Feature Extraction and Selection

The features extracted in this work and their description are shown in Tables 7.1 and 7.2. The feature set is a combination of statistical, texture-based, and silhouette region-based features. All the features are extracted from each image of the gait cycle. The final feature vector for each gait cycle is obtained by calculating the mean of a particular feature extracted from all images of a gait cycle. The usage of only mean values for the classification reduces the redundant information and also decreases the size of the feature vector to a large extent. The Relief feature selection algorithm (Kira and Rendell 1992) is applied on the extracted feature vectors as it facilitates the improvement of recognition accuracy by reducing the feature space. It removes some of the non-relevant features. The resulting classifier will be simpler and faster. The steps followed to obtain the feature vector are shown in Algorithm 14. The ranking of features obtained by applying the Relief feature selection algorithm is shown in Figure 7.4.

7.1.3 Experiments

7.1.3.1 Experimental Setup

All silhouette images used in this work are of size 240×240. As the top-scoring six most relevant features are considered for classification, the size of the feature vector obtained for a gait cycle is 6×1. The classifier used is KNN, where K=1. The CCR is computed on the test dataset to measure the performance of the proposed method.

Table 7.1: Various parameters extracted for gait evaluation

Feature Category	Feature Name
F1	Area
F2	Perimetre
F3	Convex_hull
F4	Minor_axis_length
F5	Major_axis_length
F6	Contrast
F7	Correlation
F8	Energy
F9	Homogenity
F10	Box_counting
F11	Mean_graylevel
F12	Pixel_count

Table 7.2: Type of parameters extracted for gait recognition

Features	Description
F1, F2, F3, F4, F5	Area-based features
F6, F7, F8, F9	Statistical features
F10, F11, F12	Texture features

Algorithm 14 The framework followed in the proposed work to extract gait feature vector is as follows.

1: **Input**: Number of subjects $S = \{s_1, s_2, ..., s_n\}$.
2: Begin
3: **for** all subjects of S **do**
4:     **for** subject s_1 of S **do**
5:         Identify 4 gait cycles $C = \{c_1, c_2, c_3, c_4\}$ for s_1.
6:         **for** all the values C **do**
7:             **for** gait cycle c_1 of C **do**
8:             Convert c_1 into silhouette images $I = \{i_1, i_2, ..., i_m\}$.
9:             **for** all silhouette images of I **do**
10:                 Extract 12 features, $i_k F = \{i_k F_1, i_k F_2, ..., i_k F_{12}\}$.
11:             **end for**
12:             Compute mean, $M = \left\{ \frac{i_1 F_1 + ... + i_m F_1}{m}, ..., \frac{i_m F_{12} + ... + i_m F_{12}}{m} \right\}$.
13:             Obtain the feature vector of size 12×1 for c_1, $F_{c1} = F1, F2, ..., F12$.
14:             **end for**
15:         **end for**
16:     **end for**
17: **end for**
18: Apply Relief feature selection algorithm on F.
19: Select the top scoring features $F_s = \{F1, F3, F5, F6, F11, F12\}$ for classification.
20: End

Figure 7.3: Sample silhouette images from KTH jogging dataset.
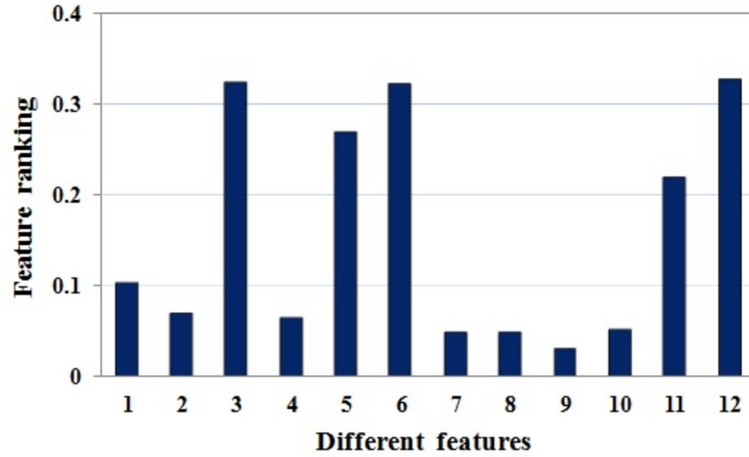
Figure 7.4: Different ranking of the features obtained by Relief algorithm. Here, 1, 2,...,12 represents F1, F2,...,F12.

7.1.3.2 Experimental Results and Discussion

In this study, the experiments are conducted on two datasets. They are KTH dataset, and Weizmann dataset.

1. Experiments on KTH dataset

A brief description of the KTH dataset is presented in Section 2.5. The two experiments conducted on the KTH running and jogging dataset are as follows.

Initially, the subjects of each scenario (*sn1*, *sn2*, *sn3*, *sn4*) are used separately for classification, and the results are given in Table 7.3. For each subject, 4 gait cycles considered for the experimentation. Hence, 100 feature vectors of 25 subjects for each scenario are used for classification. Among four gait cycles, three are considered for training, and the remaining one for testing. More information about the gait of the subject is always acquired for the view angle near 90^0, and it gradually decreases for other views. Therefore, CCR of scenario *sn2* is less compared to those of scenarios *sn1* and *sn4*. It is evident from the literature that the clothing condition influences the CCR to a large extent (Al-Tayyan et al. 2017; Yu et al. 2006). The considerable changes in the subject's silhouettes are caused by clothing condition, and hence, it decreases the accuracy as the appearance-based techniques depend on the temporal and spatial

122

Table 7.3: CCR of the proposed method on KTH running and jogging dataset for various scenarios

Sl. No.	Dataset	sn1	sn2	sn3	sn4
1	KTH running	97.66	91.00	94.66	99.00
2	KTH jogging	98.33	92.66	96.00	99.66

Table 7.4: CCR of the proposed method on KTH jogging dataset for different probe and gallery combination.

Probe dataset	sn1	Gallery dataset	sn2	sn3	sn4
			63.50	95.00	98.50
Probe dataset	sn2	Gallery dataset	sn1	sn3	sn4
			64.50	47.00	62.00
Probe dataset	sn3	Gallery dataset	sn1	sn2	sn4
			95.50	46.50	94.56
Probe dataset	sn4	Gallery dataset	sn1	sn2	sn3
			98.00	60.50	95.00

variation of silhouette over a gait cycle. So, from the Table 7.3, it is apparent that less CCR is obtained for scenario *sn3*, as compared to CCRs of *sn1* and *sn4*.

The experimental results corresponding to the different scenarios for probe and gallery are illustrated in Tables 7.4 and 7.5. Here, all the four gait cycles of each subject are considered for the gallery dataset, and two gait cycles of each subject are considered for the probe dataset. From Tables 7.4 and 7.5, it is noticeable that when the (probe, gallery) pair is *sn1* and *sn4*, higher CCR is obtained as the view angle for both the scenarios is the same. Owing to the same view angle 90^0, the subject's silhouettes do not vary largely. When *sn1*, *sn4* are considered as the probe dataset, and *sn3* is considered as the gallery dataset or vice versa, the value of CCR decreases slightly because of the variations in the subject's silhouettes caused due to clothing conditions. When the probe dataset is any one of *sn1*, *sn3*, *sn4*, and gallery dataset is *sn2* or vice versa, the CCR obtained is very less because of the scale variation of *sn2*.

Table 7.5: CCR of the proposed method on KTH running dataset for different probe and gallery combination.

Probe dataset	sn1	Gallery dataset	sn2	sn3	sn4
			55.50	91.50	95.50
Probe dataset	sn2	Gallery dataset	sn1	sn3	sn4
			55.00	35.50	50.00
Probe dataset	sn3	Gallery dataset	sn1	sn2	sn4
			92.00	37.50	88.50
Probe dataset	sn4	Gallery dataset	sn1	sn2	sn3
			95.00	48.50	90.00
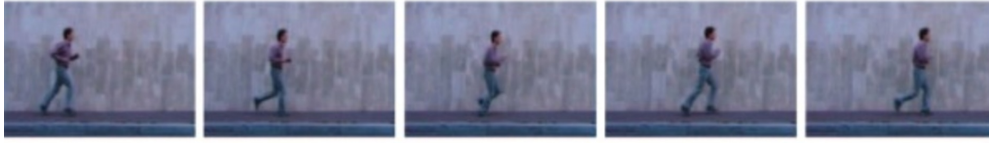
Figure 7.5: Example images from Weizmann dataset.

2. Experiments on Weizmann dataset

The sample images of the Weizmann dataset are shown in Figure 7.5. A brief description of this dataset is given in Section 2.5. This dataset consists of 10 human actions. Among all human actions, only running is considered to assess the performance of the proposed work. The results are shown in Table 7.6. The experiments are conducted using 4 gait cycles of each subject. Hence, the total number of gait cycles considered is $9 \times 4 = 36$. Among four gait cycles, three are considered for training, and the remaining one for testing. As only a few subjects are present in this dataset, 100% recognition accuracy is obtained.

Table 7.6: Recognition accuracy of the proposed method on Weizmann dataset.

No. of subjects	Total no. of gate cycles	CCR(%)
9	36	100

7.2 Summary

In this work, an effective and simple approach for the recognition of individuals by using their running patterns is presented. Here, 12 features are extracted from each image of a gait cycle, and the mean of each feature is computed from all silhouettes of a gait cycle. Further, the Relief feature selection algorithm is used to remove the irrelevant features. A single feature value is obtained by computing the mean of around 21 to 35 values. As a result, high recognition accuracy is achieved. The proposed method is assessed on the KTH and Weizmann dataset. Experimental results demonstrate that the proposed method gives higher recognition accuracy. This method is more feasible for the lateral view. For other views, it may give less accuracy, as this method is completely dependent on the shape of the silhouette. Furthermore, an experiment on a much larger dataset needs to be done.

Chapter 8

CONCLUSIONS AND FUTURE WORK

Gait recognition is a type of behavioural biometrics that has achieved significant consideration in the last decade. The advantage of this method is that the input data can be collected at a distance with minimum cooperation from the subjects when compared with other physiological and behavioural biometric modalities. This characteristic makes it primarily attractive for the investigation of criminals, surveillance, and security applications. However, an individual's gait can probably be influenced by some covariate aspects such as the difference in walking speed, differences in the appearance of individuals by clothing, time elapse, surface conditions, shoe types, viewing angle, and carrying conditions. Hence, this thesis work is more focussed on enhancing the performance of a gait recognition system in such challenging conditions.

The research has been initiated by proposing a low dimensional distinctive feature vector by using well known features such as HOG and Haralick texture descriptor. The performance analysis of this method is done by using five widely used gait databases. In comparison to other approaches based on HOG in literature, the proposed strategy gives promising results in most of the cases.

Further, an investigation has been made to improve the gait recognition performance by statistical shape analysis. The centroid corner distance features are proposed, which capture the fine details of the shape of the subject. With this method, an improved recognition rate has been achieved on three gait databases.

The effect of different clothing types on the gait silhouettes is explored. It has been concluded from the outcome that the upper and lower regions are less susceptible to the appearance changes. With this conclusion, the gait features are extracted only form upper and lower regions. Since binary descriptors are known for good discrimination in quite a lot of visual recognition methods, a binary descriptor called MLOOP is proposed. The proposed descriptor addresses the limitations of two other descriptors on which it is based. The MLOOP descriptor is used to obtain informative features from GEIs. Further, redundant features that adversely influence the gait detection are removed. The proposed approach is evaluated on two widely used gait databases, and

it achieves an improvement in recognition performance over other binary descriptors. MLOOP, being a comprehensive binary descriptor, can be utilized as a part of the research for different applications.

One of the significant factors among the various covariates is the view angle variations. The difficult one among all the view variations is the frontal view. This is because the dynamic information of the walking pattern is very less noticeable in this view when compared with other view variations. The frontal GEI consists of much static and very less dynamic information. Hence, this study proposes two methods which are sensitive to significant spatial changes of frontal GEI. The HCSD that captures the fine spatial details with respect to changes in the shape of GEI, is used as a feature descriptor in one of the proposed methods. Here, a process called similarity measurement is used to assign the probe sample to a set of gallery samples. This approach achieves higher performance when compared with other methods based on frontal GEI, reported in this thesis. One more method on frontal GEI proposes three feature vectors which are sensitive to significant spatial changes in gait sequences. The study also demonstrates its impact on an increase in the inter-class differences. The experimentation on three gait databases shows an improvement over the methods compared.

The appearance of the GEI changes when the same subject walks at different speeds. The performance of the GEI template is more sensitive to the appearance changes. Hence, two approaches have been proposed to minimize the effect of changes in the appearance of GEI, caused due to the effect of speed. The first proposed approach uses the selected ROIs to extract features, thus minimizing the appearance variations. This approach uses MI as a feature and steps such as similarity measurement and classification to assign a probe sample to a gallery sample. The proposed method shows increased performance for two databases over the methods compared. The other approach used similarity/dissimilarity metrics between a probe sample and all gallery samples present in a database to assign a probe to the gallery sample. The considerable improvement of the proposed method over the recent methods shows the credibility of the proposed approach.

A simple and effective approach that identifies individuals based on running patterns is proposed. Here, initially, twelve features are extracted from all silhouette images of

a gait cycle. Out of twelve, only six most discriminative features are used for classification. The performance of this method is evaluated on KTH and Weizmann databases.

This thesis has contributed to some of the aspects of a gait recognition system. However, still, there are many issues the future research could address. In order to ensure the applicability of the gait recognition system in a real-world environment, the proposed methods need to be evaluated on a database of large scale. This thesis uses simple methods for classification in the proposed approaches, such as KNN and similarity measurement. However, certain other widely used classification methods such as deep networks can be used, especially when experimenting with large sets of data.

References

Aggarwal, M. and Hanmandlu, M. (2015). Representing uncertainty with information sets. *IEEE Transactions on Fuzzy Systems*, 24(1), 1–15.

Al-Tayyan, A., Assaleh, K., and Shanableh, T. (2017). "Decision-level fusion for single-view gait recognition with various carrying and clothing conditions". *Image and Vision Computing*, 61, 54–69.

Alotaibi, M. and Mahmood, A. (2017). "Improved gait recognition based on specialized deep convolutional neural network". *Computer Vision and Image Understanding*, 164, 103–110.

Amin, T. and Hatzinakos, D. (2012). "Determinants in human gait recognition". In *25th Canadian Conference on Electrical & Computer Engineering, IEEE*, 1–4.

Arora, P., Hanmandlu, M., and Srivastava, S. (2015a). "Gait based authentication using gait information image features". *Pattern Recognition Letters*, 68, 336–342.

Arora, P., Srivastava, S., Arora, K., and Bareja, S. (2015b). "Improved gait recognition using gradient histogram Gaussian image". *Procedia Computer Science*, 58, 408–413.

Baraldi, A. and Parmiggiani, F. (1995). "An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters". *IEEE Transactions on Geoscience and Remote Sensing*, 33(2), 293–304.

Barnich, O. and Van Droogenbroeck, M. (2009). "Frontal-view gait recognition by intra-and inter-frame rectangle size distribution". *Pattern Recognition Letters*, 30(10), 893–901.

Bashir, K., Xiang, T., and Gong, S. (2010). "Gait recognition without subject cooperation". *Pattern Recognition Letters*, 31(13), 2052–2060.

Bashir, K., Xiang, T., Gong, S., and Mary, Q. (2009). "Gait Representation Using Flow Fields.". In *British Machine Vision Conference*, 1–11.

BenAbdelkader, C., Cutler, R., Nanda, H., and Davis, L. (2001). "Eigengait: Motion-based recognition of people using image self-similarity". In *International conference on audio-and video-based biometric person authentication, Springer*, 284–294.

Binsaadoon, A. G. and El-Alfy, E.-S. M. (2016). "FLGBP: Improved Method for Gait Representation and Recognition". In *World Symposium on Computer Applications & Research, IEEE*, 59–64.

Boulgouris, N. V. and Chi, Z. X. (2007). "Human gait recognition based on matching of body components". *Pattern recognition*, 40(6), 1763–1770.

Boulgouris, N. V., Plataniotis, K. N., and Hatzinakos, D. (2006). "Gait recognition using linear time normalization". *Pattern Recognition*, 39(5), 969–979.

Burhan, I. M. and Nordin, M. J. (2015). "Multi-view gait recognition using enhanced gait energy image and radon transform techniques". *Asian Journal of Applied Sciences*, 8(2), 138–148.

Chakraborti, T., McCane, B., Mills, S., and Pal, U. (2018). "LOOP Descriptor: Local Optimal-Oriented Pattern". *IEEE Signal Processing Letters*, 25(5), 635–639.

Chaquet, J. M., Carmona, E. J., and Fernández-Caballero, A. (2013). "A survey of video datasets for human action and activity recognition". *Computer Vision and Image Understanding*, 117(6), 633–659.

Chattopadhyay, P., Roy, A., Sural, S., and Mukhopadhyay, J. (2014a). "Pose Depth Volume extraction from RGB-D streams for frontal gait recognition". *Journal of Visual Communication and Image Representation*, 25(1), 53–63.

Chattopadhyay, P., Sural, S., and Mukherjee, J. (2014b). "Frontal gait recognition from incomplete sequences using RGB-D camera". *IEEE Transactions on Information Forensics and Security*, 9(11), 1843–1856.

Chattopadhyay, P., Sural, S., and Mukherjee, J. (2015). "Frontal gait recognition from occluded scenes". *Pattern Recognition Letters*, 63, 9–15.

Chen, C., Liang, J., Zhao, H., Hu, H., and Tian, J. (2009). "Frame difference energy image for gait recognition with incomplete silhouettes". *Pattern Recognition Letters*, 30(11), 977–984.

Chen, X., Weng, J., Lu, W., and Xu, J. (2018). "Multi-Gait Recognition Based on Attribute Discovery". *IEEE transactions on pattern analysis and machine intelligence*, 40(7), 1697–1710.

Cheung, S. C. and Kamath, C. (2005). "Robust background subtraction with foreground validation for urban traffic video". *EURASIP Journal on Advances in Signal Processing*, 2005(14), 726261.

Choudhury, S. D. and Tjahjadi, T. (2015). "Robust view-invariant multiscale gait recognition". *Pattern Recognition*, 48(3), 798–811.

Choudhury, S. D. and Tjahjadi, T. (2016). "Clothing and carrying condition invariant gait recognition based on rotation forest". *Pattern Recognition Letters*, 80, 1–7.

Collins, R. T., Gross, R., and Shi, J. (2002). "Silhouette-based human identification from body shape and gait". In *5th International Conference on Automatic Face and Gesture Recognition, IEEE*, 366–371.

Connor, P. and Ross, A. (2018). "Biometric recognition by gait: A survey of modalities and features". *Computer Vision and Image Understanding*, 167, 1–27.

Cucchiara, R., Grana, C., Piccardi, M., and Prati, A. (2003). "Detecting moving objects, ghosts, and shadows in video streams". *IEEE transactions on pattern analysis and machine intelligence*, 25(10), 1337–1342.

Cunningham, P. and Delany, S. J. (2007). "k-Nearest neighbour classifiers". *Multiple Classifier Systems*, 34, 1–17.

Cuntoor, N., Kale, A., and Chellappa, R. (2003). "Combining multiple evidences for gait recognition". In *International Conference on Multimedia and Expo, IEEE*, 3, I09–113.

Dadashi, F., Araabi, B. N., and Soltanian-Zadeh, H. (2009). "Gait recognition using wavelet packet silhouette representation and transductive support vector machines". In *2nd International Conference on Image and Signal Processing, IEEE*, 1–5.

Dalal, N. and Triggs, B. (2005). "Histograms of oriented gradients for human detection". In *Computer Society Conference on Computer Vision and Pattern Recognition, IEEE*, 1, 886–893.

Deng, G. and Cahill, L. (1993). "An adaptive Gaussian filter for noise reduction and edge detection". In *International Conference on Medical Imaging and Nuclear Science Symposium, IEEE*, 1615–1619. IEEE.

Deng, M. and Wang, C. (2018). "Gait recognition under different clothing conditions via deterministic learning". *IEEE/CAA Journal of Automatica Sinica*.

Ding, T. (2008). "A robust identification approach to gait recognition". In *Conference on Computer Vision and Pattern Recognition, IEEE*, 1–8.

Dupuis, Y., Savatier, X., and Vasseur, P. (2013). "Feature subset selection applied to model-free gait recognition". *Image and vision computing*, 31(8), 580–591.

El-Alfy, H., Mitsugami, I., and Yagi, Y. (2014). "A new gait-based identification method using local Gauss maps". In *Asian Conference on Computer Vision, Springer*, 3–18.

Foster, J. P., Nixon, M. S., and Prügel-Bennett, A. (2003). "Automatic gait recognition using area-based metrics". *Pattern Recognition Letters*, 24(14), 2489–2497.

Gaba, I. and Ahuja, S. P. (2014). "Gait analysis for identification by using BPNN with LDA and MDA techniques". In *International Conference on MOOC, Innovation and Technology in Education, IEEE*, 122–127.

Ghebleh, A. and Moghaddam, M. E. (2018). "Clothing-invariant human gait recognition using an adaptive outlier detection method". *Multimedia Tools and Applications*, 1–21.

Goffredo, M., Carter, J. N., and Nixon, M. S. (2008). "Front-view gait recognition". In *2nd International Conference on Biometrics: Theory, Applications and Systems, IEEE*, 1–6.

Gross, R. and Shi, J. (2001). "The CMU Motion of Body (MoBo) Database". Technical Report CMU-RI-TR-01-18, Carnegie Mellon University, Pittsburgh, PA.

Guan, Y., Li, C.-T., and Hu, Y. (2012). "Robust clothing-invariant gait recognition". In *8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE*, 321–324.

Hagui, M. and Mahjoub, M. A. (2016). "Hidden conditional random fields for gait recognition". In *International Image Processing, Applications and Systems, IEEE*, 1–6.

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). "Textural features for image classification". *IEEE Transactions on systems, man, and cybernetics*, (6), 610–621.

Hofmann, M. and Rigoll, G. (2012). "Improved gait recognition using gradient histogram energy image". In *19th International Conference on Image Processing, IEEE*, 1389–1392.

Hofmann, M. and Rigoll, G. (2012). "Improved Gait Recognition using Gradient Histogram Energy Image". In *19th International Conference on Image Processing, IEEE*, 1389–1392.

Hofmann, M. and Rigoll, G. (2013). "Exploiting gradient histograms for gait-based person identification". In *20th International Conference on Image Processing, IEEE*, 4171–4175.

Hofmann, M. and Rigoll, G. (2013). "Exploiting gradient histograms for gait-based person identification". In *International Conference on Image Processing, IEEE*, 4171–4175.

Hossain, M. A., Makihara, Y., Wang, J., and Yagi, Y. (2010). "Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control". *Pattern Recognition*, 43(6), 2281–2291.

Hu, M., Wang, Y., Zhang, Z., Little, J. J., and Huang, D. (2013a). "View-invariant discriminative projection for multi-view gait-based human identification". *IEEE Transactions on Information Forensics and Security*, 8(12), 2034–2045.

Hu, M., Wang, Y., Zhang, Z., Zhang, D., and Little, J. J. (2013b). "Incremental learning for video-based gait recognition with LBP flow". *IEEE transactions on cybernetics*, 43(1), 77–89.

Huang, C.-C., Hsu, C.-C., Liao, H.-Y., Yang, S.-H., Wang, L.-L., and Chen, S.-Y. (2016). "Frontal gait recognition based on spatio-temporal interest points". *Journal of the Chinese Institute of Engineers*, 39(8), 997–1002.

Huang, C.-P., Hsieh, C.-H., Lai, K.-T., and Huang, W.-Y. (2011). "Human action recognition using histogram of oriented gradient of motion history image". In *1st Interna-*

tional Conference on Instrumentation, Measurement, Computer, Communication and Control, IEEE, 353–356.

Huang, P. S., Harris, C. J., and Nixon, M. S. (1999). "Recognising humans by gait via parametric canonical space". *Artificial Intelligence in Engineering*, 13(4), 359–366.

Huang, S., Elgammal, A., Lu, J., and Yang, D. (2015). "Cross-speed gait recognition using speed-invariant gait templates and globality–locality preserving projections". *IEEE Transactions on Information Forensics and Security*, 10(10), 2071–2083.

Ilunga-Mbuyamba, E., Avina-Cervantes, J. G., Lindner, D., Guerrero-Turrubiates, J., and Chalopin, C. (2016). "Automatic brain tumor tissue detection based on hierarchical centroid shape descriptor in Tl-weighted MR images". In *International Conference on Electronics, Communications and Computers, IEEE*, 62–67.

Ioannidis, D., Tzovaras, D., Damousis, I. G., Argyropoulos, S., and Moustakas, K. (2007). "Gait recognition using compact feature extraction transforms and depth information". *IEEE Transactions on Information Forensics and security*, 2(3), 623–630.

Isaac, E. R., Elias, S., Rajagopalan, S., and Easwarakumar, K. (2017). "View-Invariant Gait Recognition Through Genetic Template Segmentation". *IEEE signal processing letters*, 24(8), 1188–1192.

Islam, M. S., Islam, M. R., Akter, M. S., Hossain, M., and Molla, M. (2013). "Window based clothing invariant gait recognition". In *International Conference on Advances in Electrical Engineering, IEEE*, 411–414.

Jabid, T., Kabir, M. H., and Chae, O. (2010). "Gender classification using local directional pattern (LDP)". In *20th International Conference on Pattern Recognition, IEEE*, 2162–2165.

Johnson, A. Y. and Bobick, A. F. (2001). "A multi-view method for gait recognition using static body parameters". In *International Conference on Audio-and Video-Based Biometric Person Authentication, Springer*, 301–311.

Kale, A., Rajagopalan, A., Cuntoor, N., and Kruger, V. (2002). "Gait-based recognition of humans using continuous HMMs". In *5th International Conference on Automatic Face Gesture Recognition, IEEE*, 336–341.

Kawai, R., Makihara, Y., Hua, C., Iwama, H., and Yagi, Y. (2012). "Person re-identification using view-dependent score-level fusion of gait and color features". In *21st International Conference on Pattern Recognition, IEEE*, 2694–2697.

Kellokumpu, V., Zhao, G., Li, S. Z., and Pietikäinen, M. (2009). "Dynamic texture based gait recognition". In *International conference on biometrics, Springer*, 1000–1009.

Kira, K. and Rendell, L. A. (1992). "The feature selection problem: Traditional methods and a new algorithm". In *Association for the Advancement of Artificial Intelligence*, 2, 129–134.

Kovač, J., Štruc, V., and Peer, P. (2017). "Frame–based classification for cross-speed gait recognition". *Multimedia Tools and Applications*, 1–23.

Kumar, H. M. and Nagendraswamy, H. (2014a). "LBP for gait recognition: A symbolic approach based on GEI plus RBL of GEI". In *International Conference on Electronics and Communication Systems, IEEE*, 1–5.

Kumar, M. H. and Nagendraswamy, H. (2014b). "Symbolic representation and recognition of gait: an approach based on LBP of split gait energy images". *Signal & Image Processing*, 5(4), 15.

Kusakunniran, W. (2014a). "Attribute-based learning for gait recognition using spatio-temporal interest points". *Image and Vision Computing*, 32(12), 1117–1126.

Kusakunniran, W. (2014b). "Recognizing gaits on spatio-temporal feature domain". *IEEE Transactions on Information Forensics and Security*, 9(9), 1416–1423.

Kusakunniran, W., Wu, Q., Li, H., and Zhang, J. (2009). "Automatic gait recognition using weighted binary pattern on video". In *6th International Conference on Advanced Video and Signal Based Surveillance, IEEE*, 49–54.

Kusakunniran, W., Wu, Q., Zhang, J., and Li, H. (2010). "Support vector regression for multi-view gait recognition based on local motion feature selection". In *Conference on Computer Vision and Pattern Recognition, IEEE*, 974–981.

Kusakunniran, W., Wu, Q., Zhang, J., and Li, H. (2011). "Speed-invariant gait recognition based on procrustes shape analysis using higher-order shape configuration". In *18th IEEE International Conference on Image Processing, IEEE*, 545–548.

Kusakunniran, W., Wu, Q., Zhang, J., and Li, H. (2012). "Gait recognition across various walking speeds using higher order shape configuration based on a differential composition model". *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(6), 1654–1668.

Kusakunniran, W., Wu, Q., Zhang, J., Li, H., and Wang, L. (2013). "Recognizing gaits across views through correlated motion co-clustering". *IEEE Transactions on Image Processing*, 23(2), 696–709.

Lam, T. H., Cheung, K. H., and Liu, J. N. (2011). "Gait flow image: A silhouette-based gait representation for human identification". *Pattern recognition*, 44(4), 973–987.

Lee, C. P., Tan, A. W., and Tan, S. C. (2013). "Gait recognition via optimally interpolated deformable contours". *Pattern Recognition Letters*, 34(6), 663–669.

Lee, C. P., Tan, A. W., and Tan, S. C. (2014a). "Time-sliced averaged motion history image for gait recognition". *Journal of Visual Communication and Image Representation*, 25(5), 822–826.

Lee, H., Lee, H., and Kim, E. (2014b). "A new gait recognition system based on hierarchical fair competition-based parallel genetic algorithm and selective neural network ensemble". *International Journal of Control, Automation, and Systems*, 12(1), 202.

Lishani, A. O., Boubchir, L., Khalifa, E., and Bouridane, A. (2017). "Human gait recognition based on Haralick features". *Signal, Image and Video Processing*, 11(6), 1123–1130.

Lishani, A. O., Boubchir, L., Khalifa, E., and Bouridane, A. (2018). "Human gait recognition using GEI-based local multi-scale feature descriptors". *Multimedia Tools and Applications*, 1–16.

Liu, L., Yin, Y., Qin, W., and Li, Y. (2011). "Gait recognition based on outermost contour". *International Journal of Computational Intelligence Systems*, 4(5), 1090–1099.

Liu, Z. and Sarkar, S. (2004). "Simplest representation yet for gait recognition: Averaged silhouette". In *17th International Conference on Pattern Recognition, IEEE*, 4, 211–214.

Liu, Z. and Sarkar, S. (2006). "Improved gait recognition by gait dynamics normalization". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(6), 863–876.

Liu, Z., Zhang, Z., Wu, Q., and Wang, Y. (2015). "Enhancing person re-identification by integrating gait biometric". *Neurocomputing*, 168, 1144–1156.

Lo Presti, L. and La Cascia, M. (2015). "Using Hankel matrices for dynamics-based facial emotion recognition and pain detection". In *Conference on Computer Vision and Pattern Recognition Workshops, IEEE*, 26–33.

Lublinerman, R., Ozay, N., Zarpalas, D., and Camps, O. (2006). "Activity recognition from silhouettes using linear systems and model (in) validation techniques". In *18th International Conference on Pattern Recognition, IEEE*, 1, 347–350.

Lumini, A. and Nanni, L. (2017). "Overview of the combination of biometric matchers". *Information Fusion*, 33, 71–85.

Luo, J., Zhang, J., Zi, C., Niu, Y., Tian, H., and Xiu, C. (2015). "Gait recognition using GEI and AFDEI". *International Journal of Optics*, 2015.

Makihara, Y., Mannami, H., Tsuji, A., Hossain, M., Sugiura, K., Mori, A., and Yagi, Y. (2012). "The OU-ISIR Gait Database Comprising the Treadmill Dataset". *IPSJ Trans. on Computer Vision and Applications*, 4, 53–62.

Makihara, Y., Tsuji, A., and Yagi, Y. (2014). "Speed-invariant gait recognition". In *Signal and Image Processing for Biometrics*, 209–229. Springer.

Man, J. and Bhanu, B. (2006). "Individual recognition using gait energy image". *IEEE transactions on pattern analysis and machine intelligence*, 28(2), 316–322.

Maple, C. (2003). "Geometric design and space planning using the marching squares and marching cube algorithms". In *International Conference on Geometric Modeling and Graphics, IEEE*, 90–95.

Matovski, D. S., Nixon, M. S., Mahmoodi, S., and Carter, J. N. (2010). "The effect of time on the performance of gait biometrics". In *4th International Conference on Biometrics: Theory, Applications and Systems, IEEE*, 1–6.

Medikonda, J., Madasu, H., and Bijaya Ketan, P. (2018). "Information set based features for the speed invariant gait recognition". *IET Biometrics*, 7(3), 269–277.

Mogan, J. N., Lee, C. P., Lim, K. M., and Tan, A. W. (2017). "Gait recognition using binarized statistical image features and histograms of oriented gradients". In *International Conference on Robotics, Automation and Sciences, IEEE*, 1–6.

Mowbray, S. D. and Nixon, M. S. (2003). "Automatic gait recognition via fourier descriptors of deformable objects". In *International Conference on Audio-and Video-Based Biometric Person Authentication, Springer*, 566–573.

Muramatsu, D., Iwama, H., Makihara, Y., and Yagi, Y. (2013). "Multi-view multi-modal person authentication from a single walking image sequence". In *International Conference on Biometrics, IEEE*, 1–8.

Murase, H. and Sakai, R. (1996). "Moving object recognition in eigenspace representation: gait analysis and lip reading". *Pattern recognition letters*, 17(2), 155–162.

Nambiar, A. M., Correia, P. L., and Soares, L. D. (2012). "Frontal gait recognition combining 2D and 3D data". In *Proceedings of the on Multimedia and Security, ACM*, 145–150.

Nandy, A., Bhowmick, S., Chakraborty, P., and Nandi, G. C. (2014a). "Gait Biometrics: An Approach to Speed Invariant Human Gait Analysis for Person Identification". In *2nd International Conference on Soft Computing for Problem Solving, Springer*, 729–737.

Nandy, A., Chakraborty, P., Nandi, G., et al. (2013). "Speed invariant, human gait based recognition system for video surveillance security". In *International Conference on Intelligent Interactive Technologies and Multimedia, Springer*, 325–335.

Nandy, A., Chakraborty, R., and Chakraborty, P. (2016). "Cloth invariant gait recognition using pooled segmented statistical features". *Neurocomputing*, 191, 117–140.

Nandy, A., Chakraborty, R., Chakraborty, P., and Nandi, G. C. (2014b). "A novel Approach to Human Gait Recognition using possible Speed Invariant features". *International Journal of Computational Intelligence Systems*, 7(6), 1174–1193.

Nandy, A., Pathak, A., and Chakraborty, P. (2017). "A study on gait entropy image analysis for clothing invariant human identification". *Multimedia Tools and Applications*, 76(7), 9133–9167.

Neves, J., Narducci, F., Barra, S., and Proença, H. (2016). "Biometric recognition in surveillance scenarios: a survey". *Artificial Intelligence Review*, 46(4), 515–541.

Ojala, T., Pietikainen, M., and Harwood, D. (1994). "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions". In *12th Proceedings of International Conference on Pattern Recognition, IEEE*, 1, 582–585.

Phillips, P. J., Sarkar, S., Robledo, I., Grother, P., and Bowyer, K. (2002). "The gait identification challenge problem: Data sets and baseline algorithm". In *16th International Conference on Pattern Recognition, IEEE*, 1, 385–388.

Pluim, J. P., Maintz, J. A., and Viergever, M. A. (2003). "Mutual-information-based registration of medical images: a survey". *IEEE transactions on medical imaging*, 22(8), 986–1004.

Prakash, C., Kumar, R., and Mittal, N. (2018). "Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges". *Artificial Intelligence Review*, 49(1), 1–40.

Prakash, C., Mittal, A., Kumar, R., and Mittal, N. (2015). "Identification of gait parameters from silhouette images". In *8th International Conference on Contemporary Computing, IEEE*, 190–195.

Rida, I., Almaadeed, S., and Bouridane, A. (2016a). "Gait recognition based on modified phase-only correlation". *Signal, Image and Video Processing*, 10(3), 463–470.

Rida, I., Jiang, X., and Marcialis, G. L. (2016b). "Human body part selection by group lasso of motion for model-free gait recognition". *IEEE Signal Processing Letters*, 23(1), 154–158.

Rokanujjaman, M., Hossain, M. A., Islam, M. R., Hossain, A. A., and Ferworn, A. (2016). "Part definition and selection for part-based speed invariant gait recognition". In *9th International Conference on Electrical and Computer Engineering, IEEE*, 218–221.

Rokanujjaman, M., Hossain, M. A., Islam, M. R., and Islam, M. S. (2013). "Effective part definition for gait identification using gait entropy image". In *International Conference on Informatics, Electronics & Vision, IEEE*, 1–4.

Rokanujjaman, M., Islam, M. S., Hossain, M. A., Islam, M. R., Makihara, Y., and Yagi, Y. (2015). "Effective part-based gait identification using frequency-domain gait entropy features". *Multimedia Tools and Applications*, 74(9), 3099–3120.

Roy, A., Sural, S., and Mukherjee, J. (2012a). "A hierarchical method combining gait and phase of motion with spatiotemporal model for person re-identification". *Pattern Recognition Letters*, 33(14), 1891–1901.

Roy, A., Sural, S., and Mukherjee, J. (2012b). "Gait recognition using pose kinematics and pose energy image". *Signal Processing*, 92(3), 780–792.

Ryu, J. and Kamata, S.-i. (2011). "Front view gait recognition using spherical space model with human point clouds". In *18th International Conference on Image Processing, IEEE*, 3209–3212.

Sayeed, F. and Hanmandlu, M. (2017). "Properties of information sets and information processing with an application to face recognition". *Knowledge and Information Systems*, 52(2), 485–507.

Schuldt, C., Laptev, I., and Caputo, B. (2004). "Recognizing human actions: a local SVM approach". In *17th International Conference on Pattern Recognition, IEEE*, 3, 32–36.

Semwal, V. B., Raj, M., and Nandi, G. C. (2015). "Biometric gait identification based on a multilayer perceptron". *Robotics and Autonomous Systems*, 65, 65–75.

Shaikh, S. H., Saeed, K., and Chaki, N. (2014). "Gait recognition using partial silhouette-based approach". In *International Conference on Signal Processing and Integrated Networks, IEEE*, 101–106.

Sharifi, M., Fathy, M., and Mahmoudi, M. T. (2002). "A classified and comparative study of edge detection algorithms". In *International Conference on Information Technology: Coding and Computing, IEEE*, 117–120.

Sharma, S., Tiwari, R., Singh, V., et al. (2011). "Identification of People Using Gait Biometrics". *In International journal of machine learning and computing, IACSIT Press*, 1(4), 409.

Shi, J. and Tomasi, C. (1993). "Good features to track". Technical report, Cornell University.

Shiraga, K., Makihara, Y., Muramatsu, D., Echigo, T., and Yagi, Y. (2016). "Geinet: View-invariant gait recognition using a convolutional neural network". In *International Conference on Biometrics, IEEE*, 1–8.

Shrivakshan, G. and Chandrasekar, C. (2012). "A comparison of various edge detection techniques used in image processing". *International Journal of Computer Science Issues*, 9(5), 269.

Singhal, A. and Lall, B. (2013). "Novel signal processing approach for gait based human identification system". In *International Conference on Signal Processing and Communication, IEEE*, 197–201.

Sivapalan, S., Chen, D., Denman, S., Sridharan, S., and Fookes, C. (2011). "Gait energy volumes and frontal gait recognition using depth images". In *International Joint Conference on Biometrics, IEEE*, 1–6.

Sivapalan, S., Chen, D., Denman, S., Sridharan, S., and Fookes, C. (2012). "The backfilled gei-a cross-capture modality gait feature for frontal and side-view gait recognition". In *International Conference on Digital Image Computing Techniques and Applications, IEEE*, 1–8.

Sivapalan, S., Chen, D., Denman, S., Sridharan, S., and Fookes, C. (2013). "Histogram of weighted local directions for gait recognition". In *Conference on Computer Vision and Pattern Recognition Workshops, IEEE*, 125–130.

Soriano, M., Araullo, A., and Saloma, C. (2004). "Curve spreads-a biometric from front-view gait video". *Pattern Recognition Letters*, 25(14), 1595–1602.

Stöckel, T., Jacksteit, R., Behrens, M., Skripitz, R., Bader, R., and Mau-Moeller, A. (2015). "The mental representation of the human gait in young and older adults". *Frontiers in psychology*, 6(2), 102–115.

Tafazzoli, F., Bebis, G., Louis, S., and Hussain, M. (2015). "Genetic feature selection for gait recognition". *Journal of Electronic Imaging*, 24(1), 013036.

Tan, D., Huang, K., Yu, S., and Tan, T. (2006). "Efficient night gait recognition based on template matching". In *18th International Conference on Pattern Recognition, IEEE*, 3, 1000–1003.

Tan, D., Huang, K., Yu, S., and Tan, T. (2007a). "Orthogonal diagonal projections for gait recognition". In *International Conference on Image Processing, IEEE*, 1, I–337.

Tan, D., Huang, K., Yu, S., and Tan, T. (2007b). "Recognizing night walkers based on one pseudoshape representation of gait". In *Conference on Computer Vision and Pattern Recognition, IEEE*, 1–8.

Tan, D., Huang, K., Yu, S., and Tan, T. (2007c). "Uniprojective features for gait recognition". In *International Conference on Biometrics, Springer*, 673–682.

Tan, D., Yu, S., Huang, K., and Tan, T. (2007d). "Walker recognition without gait cycle estimation". In *International Conference on Biometrics, Springer*, 222–231.

Wang, C., Zhang, J., Pu, J., Yuan, X., and Wang, L. (2010). "Chrono-gait image: A novel temporal template for gait recognition". In *European Conference on Computer Vision, Springer*, 257–270.

Wang, J., Makihara, Y., and Yagi, Y. (2008). "Human tracking and segmentation supported by silhouette-based gait recognition". In *International Conference on Robotics and Automation, IEEE*, 1698–1703.

Wang, L., Ning, H., Hu, W., and Tan, T. (2002). "Gait recognition based on procrustes shape analysis". In *International Conference on Image Processing, IEEE*, 3, 109–111.

Wang, L., Tan, T., Hu, W., and Ning, H. (2003). "Automatic gait recognition based on statistical shape analysis". *IEEE transactions on image processing*, 12(9), 1120–1131.

Whytock, T., Belyaev, A., and Robertson, N. (2012). "GEI+ HOG for action recognition". In *4th UK Computer Vision Student Workshop*.

Wren, C. R., Azarbayejani, A., Darrell, T., and Pentland, A. P. (1997). "Pfinder: Real-time tracking of the human body". *IEEE Transactions on pattern analysis and machine intelligence*, 19(7), 780–785.

Wu, Z., Huang, Y., Wang, L., Wang, X., and Tan, T. (2017). "A comprehensive study on cross-view gait based human identification with deep cnns". *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2), 209–226.

Yam, C., Nixon, M. S., and Carter, J. N. (2002). "Gait recognition by walking and running: a model-based approach". *Asian Conference on Computer Vision, IEEE*, 1–6.

Yam, C.-Y., Nixon, M. S., and Carter, J. N. (2001). "Extended model-based automatic gait recognition of walking and running". In *International Conference on Audio-and Video-Based Biometric Person Authentication, Springer*, 278–283.

Ye, B. and Wen, Y. (2006). "A new gait recognition method based on body contour". In *9th International Conference on Control, Automation, Robotics and Vision, IEEE*, 1–6.

Yu, S., Tan, D., and Tan, T. (2006). "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition". In *18th International Conference on Pattern Recognition, IEEE*, 4, 441–444.

Zeng, W. and Wang, C. (2016). "View-invariant gait recognition via deterministic learning". *Neurocomputing*, 175, 324–335.

Zhang, S. and Zhang, C. (2015). "Orthogonal margin maximization projection for gait recognition". *Informatica*, 26(2), 357–367.

Zhang, S. and Zhang, L. (2018). "Combining weighted adaptive CS-LBP and local linear discriminant projection for gait recognition". *Multimedia Tools and Applications*, 77(10), 12331–12347.

Zhang, Y., Yang, N., Li, W., Wu, X., and Ruan, Q. (2009). "Gait recognition using procrustes shape analysis and shape context". In *Asian Conference on Computer Vision, Springer*, 256–265.

Zheng, S. (accessed July 27, 2017). *"CASIA Gait Database"*.

Publications

Journal Papers

1. **Anusha R**, and Jaidhar C D. (2019), "Human gait recognition based on histogram of oriented gradients and Haralick texture descriptor", *Journal of Multimedia Tools and Applications, Springer*, Volume: 79, pp 8213-8234, DOI 10.1007/s11042-019-08469-1.

2. **Anusha R**, and Jaidhar C D. (2019), "Clothing invariant human gait recognition using modified local optimal oriented pattern binary descriptor", *Journal of Multimedia Tools and Applications, Springer*, Volume: 79, pp 2873-2896, DOI 10.1007/s11042-019-08400-8.

3. **Anusha R**, and Jaidhar C D, "Contour image-based feature extraction for frontal gait recognition", *Journal of Signal, Image and Video Processing, Springer*. [**Under Review**]

4. **Anusha R**, and Jaidhar C D, "Spatial dynamics for speed-invariant gait detection", *Journal of Signal, Image and Video Processing, Springer*. [**Under Review**]

Conference Papers

1. **Anusha R**, and Jaidhar C D. (2018), "On human identification using running patterns: a straightforward approach", In 18th *International Conference on Intelligent Systems Design and Applications, (ISDA 2018),* held at VIT Vellore, India. pp 322-331, DOI 10.1007/978-3-030-16660-1_32.

2. **Anusha R**, and Jaidhar C D. (2019), "An approach to speed invariant gait analysis for human recognition using mutual information", In *International Conference of IEEE region 10, (TENCON 2019)*, held at Kochi, India, pp 1616-1621, DOI: 10.1109/TENCON.2019.8929256.

3. **Anusha R**, and Jaidhar C D. (2019), "Frontal gait recognition based on hierarchical centroid shape descriptor and similarity measurement", In 5th *International Conference on Data Science and Engineering, (ICDSE 2019)*, held at IIT Patna, India, pp 71-76, DOI: 10.1109/ICDSE47409.2019.8971477.

4. **Anusha R**, and Jaidhar C D. (2019), "Gaussian filtered gait energy template and centroid corner distance features for human gait recognition", In 14th *International Conference on Industrial and Information Systems, (ICIIS 2019)*, held at Peradeniya, Sri Lanka, pp 425-430, DOI: 10.1109/ICIIS47346.2019.9063346.

Curriculum Vitae

Mrs. Anusha R
Full-Time Research Scholar
Department of Information Technology
National Institute of Technology Karnataka, Surathkal
P.O. Srinivasanagar, Mangalore-575 025

Permanent Address

Anusha R
Door No. 378/1, Amma Nilaya
Near Soukyada Hospital
Nijalingappa Layout
Davangere-577 004
Email: anusha.rms90@gmail.com
Mobile: +919743220915

Academic Records

1. M.Tech. in Computer Science and Engineering from Bapuji Institute of Engineering and Technology, Davangere, 2013.

2. B.E. in Computer Science and Engineering from University B.D.T. College of Engineering, Davangere, 2011.

Research Interests

Pattern Recognition
Image Processing
Machine Learning