PERSON RECOGNITION USING GAIT ENERGY IMAGING

Ait Omar Lishani

PhD

2018

Person Recognition Using Gait Energy Imaging

by

Ait Omar Lishani

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of Doctor of Philosophy

Research undertaken in the Faculty of Engineering and Environment

July 2018

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

I dedicate this work to my family

DECLARATION

I declare that the work contained in this thesis has not been previously submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas, and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the Faculty Ethics Committee. Most results of this thesis have been published in conferences and scientific journals.

I declare that the word count of this thesis is 31, 430 words.

Signed: …

Date: …July 2018…….

Ait Lishani

ABSTRACT

Biometric technology has emerged as a viable identification and authentication solution with various systems in operation worldwide. The technology uses various modalities, including fingerprint, face, iris, palmprint, speech, and gait. Biometric recognition often involves images or videos and other image impressions that are fragile and include subtle details that are difficult to see or capture. Thus, there is a need for developing imaging applications that allow for accurate feature extraction from images for identification and recognition purposes.

Biometric modalities can be classified into two classes: physiological (i.e. fingerprint, iris, face, palm-print) or behavioural traits (speech, gait). This work is concerned with an investigation of biometric recognition at a distance and the gait modality has been chosen for various reasons. Gait data can be captured at a distance and is non-invasive. Additionally, it has advantages such as the fact that a person's gait is hard to copy, and by trying to do so, the imitator will likely appear more suspicious. Although, due to covariates, for example, a change in viewing angle, clothes, shoes, shadow or elapsed time can make gait recognition additionally challenging. There are several approaches for studying gait recognition systems such as model-based and model-free. This thesis is based on a model-free approach and proposes a supervised feature extraction approach capable of selecting distinctive features for the recognition of human gait under clothing and carrying conditions.

In this work; to allow for the characterisation of human gait properties for individual recognition, a spatiotemporal gait representation technique called Gait Energy Image (GEI) has been used. This approach is aimed at improving the recognition performance based on the principles of feature texture descriptors extracted from GEI. Furthermore, as part of this work, the dynamic parts of the energy gait representation have been proposed as means to extract

more discriminative information from a gait sequence using reduction techniques in order to further improve the human identification rate.

The four methods proposed were evaluated using CASIA Gait Database (dataset B) and USF Database under variations of clothing and carrying conditions for different viewing angles.

The first method is based on Haralick texture feature, and use the RELIEF selection algorithm. This method showed that a judicious deployment of horizontal GEI features outperforms similar methods by up to 7.00%. In addition, this method achieved an improved classification rate of up to 80.00% from a side view of 90°.

The second and third contributions are concerned with an investigation of the Gabor filter bank and Multi-scale Local Binary Pattern (MLBP) as an efficient feature extraction for gait recognition under clothing distortions. To achieve this, various dimension reduction techniques including Kernel Principal Component Analysis, Maximum Margin Projection, Spectral Regression Kernel Discriminant Analysis and Locality Preserving Projections were investigated. The results showed that the proposed methods outperform the state-of-the-art counterparts by achieving up to 93.00% Identification Rate (IR) at rank-1 using the Gabor filter method, and achieving up to 92.00% IR using the MLBP method, when using a *k*-NN classifier for a side view of 90°.

The final contribution of this work is concerned with an investigation of the Haar wavelet transform and its use for extracting powerful features for human gait recognition under clothing distortions. The experimental results using a *k*-NN classifier yielded attractive results of up to 93.00% in terms of highest IR at rank-1, compared to existing and similar state-of-the-art methods. It should be noted that all the experiments were carried out using the MATLAB programming environment.

PUBLICATIONS

- Lishani, A. O., Boubchir, L. and Bouridane, A., *Haralick features for GEI-based Human Gait Recognition*. Proceedings of the 26th International Conference on Microelectronics (ICM 2014), Doha, Qatar, pp. 36-39, December 2014.
- Lishani, A. O., Boubchir, L., Khalifa, E., and Bouridane, A., *Gabor filter bank-based GEI features for Human Gait Recognition*. Proceedings of the IEEE 39th International Conference on Telecommunications and Signal Processing (TSP 2016), Vienna, Austria, pp. 648-651, June 2016.
- Lishani, A. O., Boubchir, L., Khalifa, E., and Bouridane, A., *Human Gait Recognition based on Haralick features*. Journal of Signal, Image and Video Processing, Volume 11, Issue 6, pp 1123-1130, September 2017.
- Lishani, A. O., Boubchir, L., Khalifa, E., and Bouridane, A., *Gait Recognition Based on Wavelet Features with Spectral Regression Kernel Discriminant Analysis*. Proceedings of the IEEE 40th International Conference on Telecommunications and Signal Processing (TSP 2017), Barcelona, Spain, pp. 789-792, July 2017.
- Lishani, A. O., Boubchir, L., Khalifa, E., and Bouridane, A., *Human gait recognition using GEI-based local multi-scale feature descriptors*. Journal of Multimedia Tools and Application, Volume 77, February 2018.

ACKNOWLEDGEMENTS

First, I would like to thank my principal supervisor Professor Ahmed Bouridane for his continuous guidance throughout my research programme. His encouragement, kindness continuous support is much appreciated. I will constantly think of him as the ideal supervisor and tutor. I am also grateful to my second supervisor Dr Larbi Boubchir for his guidance and support throughout the duration of the research programme. He has always been there for me whenever I turned to him for advice.

To the spirit of my Dad (Omar Issa Lishani): I can feel him when I am in trouble. To Mom (Salma Hammad) thank you for the endless support, encouragement and sacrifices.

To my family, Aya Arebi, Idir Lishani, and Salice Lishani thank you for your support, encouragement, and love. To my brothers and sisters who have given me so much love and support; without them, this work would not have been possible.

I am also thankful to my friend, Emad Khalifa, who has shown me nothing but friendly and support, throughout my life in Newcastle. To all my friends, thank you for your help, support, and the happy times I spent with you. Finally, to my sponsor, the Military Attaché, thank you for supporting my studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

CASIA	The Institute of Automation, Chinese Academy of Sciences
CBSR	Centre for Biometrics and Security Research
CCR	Correct Classification Rate
CCTV	Closed Circuit TV
CWT	Continuous Wavelet Transform
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
GEI	Gait Energy Image
GEnI	Gait Entropy Image
IR	Identification Rate
KDA	Kernel Discriminate Analysis
k-NN	K-Nearest Neighbour
KPCA	Kernel Principal Component Analysis
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis
LDM	Layered Deformable Model
LOOCV	Leave-One-Out Cross-Validation
LPP	Locality Preserving Projections
MMD	Margin Maximising Discriminant
MMP	Maximum Margin Projection
MLBP	Multi-Scale Local Binary Pattern
SVM	Super Vector Machine

PCA	Principal Component Analysis
ROI	Region Of Interest
SGEI	Structural Gait Energy Image
SOTON	Southampton Human ID at a distance Gait Database (University of Southampton
SRKDA	Spectral Regression Kernel Discriminate Analysis
TUM GAID	Technical University Munich
USF	University of South Florida
VTM	View Transformation Model
SVD	Singular Value Decomposition
SVIM	Skeleton Variance Image
SVR	Support Vector Regression
WPT	Wavelet Packet Transform
WT	Wavelet Transform

CHAPTER ONE: INTRODUCTION

1.1 Scope of the Thesis

Increased levels of crime have led to the pursuit of new technical innovations to improve the performance of biometric recognition systems. For instance, global crime rates have driven the rapid deployment of closed-circuit TV (CCTV) surveillance for crime locating and avoidance, in order to provide a more secure environment on a global scale. CCTV, if effectively positioned, can catch a criminal either entering, escaping a scene of a crime or committing a crime first hand. Thus, it is an indispensable aid in providing direct proof in the context of prosecutions or the like.

In the last century, traditional methods of user identification and authentication included PIN codes, passwords and magnetic strip cards, which all have many disadvantages. The main drawback of these methods is that they test the validity of the password, PIN or magnetic card rather than the actual user. As a result, they can easily be shared with illegal users. The solutions used to deal with traditional access to verification or identification systems have previously concentrated on security. Though, more recently, this has shifted to biometric recognition methods being implemented as a result of their advancement. The main feature of biometrics is that the ID is a part of the human behaviour or physiology which is unique to each individual. The field of biometric technology has become extremely significant nowadays, some examples of its value and use are in federal state and local government, in financial transactions, personal data privacy and in many other commercial applications (Pousttchi et al., 2004), (Kim et al., 2004). These technologies are already being used for network security, IDs, banking, police investigations, healthcare, mobile devices and social services (A. Jain et al., 1997). Recently,

biometric technology is rapidly expanding as a trusted and effective technology for human identification through the use of unique behavioural and/or physical characteristics. According to (A. K. Jain et al., 2004), a biometric solution must satisfy the following properties:

- Universality: every individual ought to have the characteristic.
- Uniqueness: individuals are well separated by the characteristic.
- Permanence: there is sufficient invariance with the passage of time.
- Collectability: they are qualitatively quantifiable.

Human characteristics cannot be stolen or lost unlike other types of authentication and identification using passwords or tokens. Conveniently, biometric technology allows users a secure admission to services via authentication and access control security systems using several biometric modalities like a fingerprint, the iris, face, or voice (A. Jain et al., 2007). Biometrics can be split into two classes: physiological and behavioural. Behavioural biometric modalities include a signature, gait, typing rhythm etc., whilst a face, fingerprint, iris, and hand geometry all fall under physiological biometrics. Historically, fingerprints were reportedly used in Japan as early as the eighth century and handwritten signatures were used in Europe and China (Anderson, 2008). More detailed descriptions of biometric based recognition using either physical or behavioural traits can be found in Anderson's work (Anderson, 2008). Evidently, biometrics have played an important role in identity recognition throughout history and with the rapid evolution of technology over the past few decades, even more new dimensions are manifesting day-by-day.

Developing a technology that provides users with secure access to a service, is a great challenge. There are several conventional means for personal identification or authentication including passports, keys, passwords, access cards for physical access control at building

entrances or even a time attendance system for employee management. The disadvantages of conventional means of personal identification and/or authentication are losing or forgetting passwords, keys or passports; leading to potentially significant losses. Conversely, human characteristics cannot be lost or stolen and that is why systems using biometrics are proving to be an efficient solution to such problems.

Table 1-1 presents a timeline of the biometric pattern developments that resulted in improvements in the technology and an increase in the amount of biometric data available. Approaches based on computer vision have been widely researched in order to build automatic biometric recognition systems. According to (A. K. Jain et al., 2004); systems based on physiological biometrics such as fingerprints or the iris, have already been developed and are currently in use in real-world applications (A. K. Jain et al., 2004).

Table 1-1. Biometric Modality development.

Type	Year
Fingerprint, Voice	1960
Palm print, Face	1970
Iris, Signature	1980
Vein, Gait, Ear, Keystroke	1990
DNA, EEG, Dental, Shoe	2000

The recognition system introduces automatic evaluation of the iris, fingerprint and face images of current subjects and their stored images in the system database. The average time for a traveller check process is then reduced whilst also expanding the nature of security within the airport. As traveller numbers keep on increasing rapidly and on a considerably vast scale, biometric technology is needed imminently. Recently many researchers have tried to tackle this matter both by enhancing and developing currently existing biometric modalities or by

starting to explore other new biometric modalities. As an example, iris recognition technology has been developed a couple years ago to enable a convenient biometric and fast authentication process (Matey et al., 2006).

Radu et al. proposed a 2D Gabor filter bank to be used for iris recognition systems (Radu et al., 2013). The proposed approach is suitable for both near infrared and visible spectrum iris images. Also, Trokielewicz et al. proposed a unique analysis of post-mortem human iris recognition in (Trokielewicz et al., 2016). The findings of this proposed analysis showed that the dynamics of post-mortem changes to the iris that are important for biometric identification are much more moderate than previously believed (Trokielewicz et al., 2016).

On the other hand, gait recognition has the potential to satisfy many of the performance requirements. This non-invasive biometric modality can be extracted from a distance. Examples of common biometrics used now are shown in Figure 1.1.

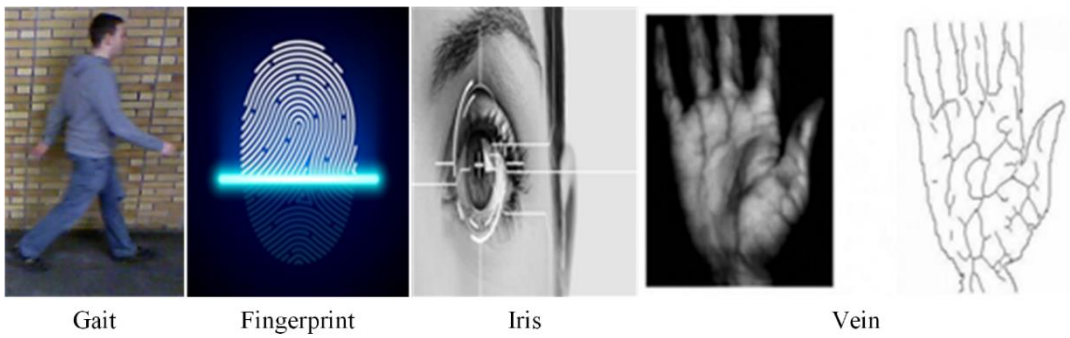


Figure 1-1. Personal identification with different biometric systems, (A. K. Jain et al., 2004).

A biometric recognition system is essentially a pattern recognition system which recognises users by matching their behavioural or physiological characteristics with stored templates. The users must be enrolled in the system in order that their biometric template or reference can be captured. This template is securely stored in a central database. The template is used for matching when an individual needs to be identified for a particular purpose. Suitably, a

biometric system can function in either identification mode or verification (authentication) mode.

- In identification mode (Who am I?): the system performs 1: N match between the probe template and all the N templates stored in the database.
- In the verification mode (Am I who I claim to be?): the system has to perform 1:1 matches between the stored template and the probe template to reject or confirm the identity. Figure 1-2 and Figure 1-3 represent gait identification and verification modes respectively.

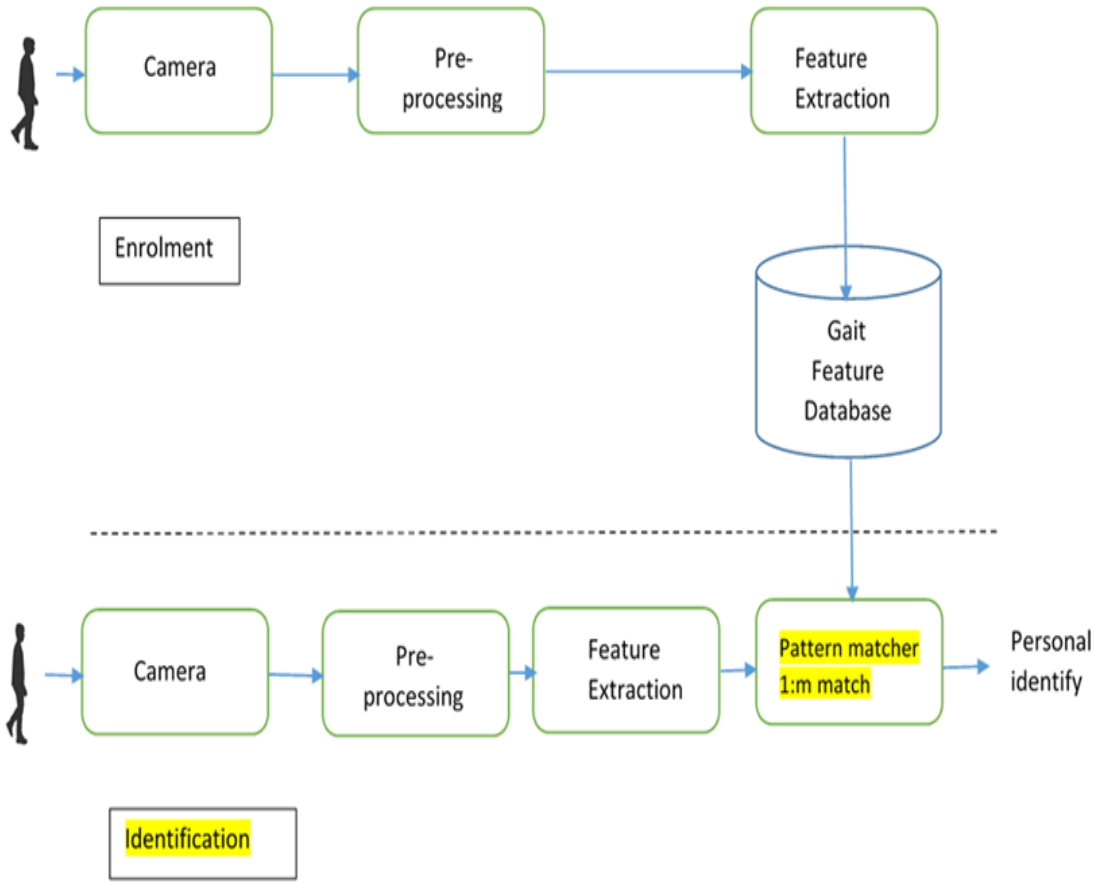


Figure 1-2. Block diagram of a gait identification system.

- The third application in biometric recognition is a watch list: this mode which has been defined by Phillips (Phillips et al., 2003), aims to compare a suspected person against a database of known persons (Watchlist). In this system, the person does not claim any

identity; it is an open-universe test. The test person may or may not be in the system database (Bouridane, 2009). Figure 1-4 shows biometric systems using fingerprints and the iris at Heathrow airport, London, in December 2006.

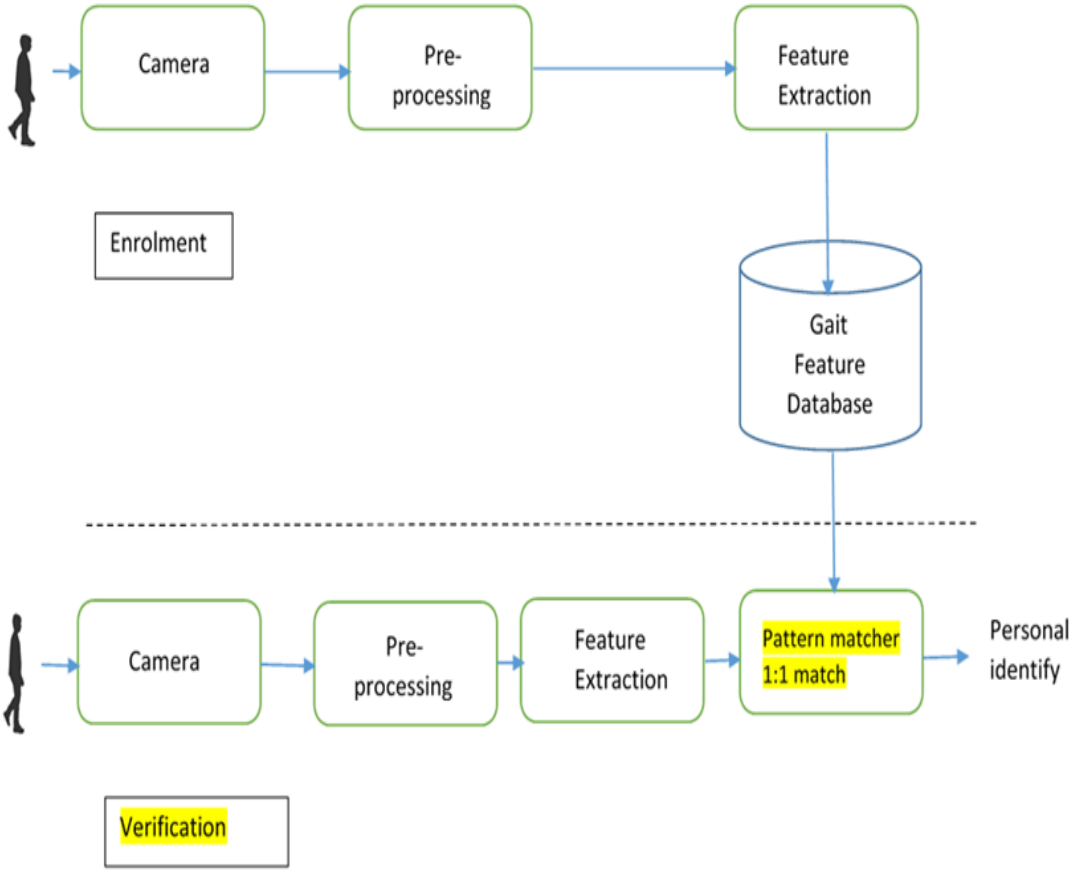


Figure 1-3. Block diagram of a gait verification system.

Figure 1-4. Automated Biometric recognition using iris and fingerprint recognition from (Heathrow, 2006).

1.2 Human Identification based on Gait Recognition

Jain et al. (A. K. Jain et al., 2004) summarised gait recognition by stating “it is impossible to miss the way an individual walks”. Gait is a complex spatiotemporal biometric modality and is not extremely unique but is adequately biased to permit checks in some low-security applications. It is a behavioural biometric and may not remain invariant, particularly over a long period of time. Examples of this, in body weight or the acquisition of significant wounds.

1.2.1 Motivations

Gait recognition is an appealing and complementary form of recognition compared with other biometric modalities such as fingerprints, face or palm-print recognition as it has many advantages. Unlike other biometric modalities gait data can be captured from a distance and has a low resolution. Moreover, a person’s gait is hard to imitate and by trying to do so the person will probably appear more suspicious than with other biometric techniques, such as face recognition, as the face can easily be hidden. Additionally, when face recognition is not possible, gait as a biometric parameter becomes invaluably useful.

1.2.2 Challenges of Gait Recognition

Gait recognition, as is the case with numerous computer vision systems, faces challenges that are extrinsic to the image acquisition process e.g. noise, lighting conditions, etc. In particular, extracting features from a gait video sequence requires the extraction of the moving individual from the background, as shown in Figure 1-5. Image noise and changing lighting conditions specifically influence the capacity of algorithms to segment the moving individual from the background effectively, hence, bringing missing body parts and the incorporation of background e.g. shadows, as shown in Figure 1-6.

To reduce the effect of noise and changing lighting conditions, a pre-processing stage is usually required in a gait recognition system.

Pre-processing aims to remove some of these issues. It can be seen from the pre-processed silhouettes (shown in Figure 1-7) that, even after pre-processing, the extracted silhouettes are still noisy. This means that a gait recognition system has to deal with a large degree of noise before the feature level.

In addition to image noise, lighting condition changes and occlusions, there are many other aspects that limit the performance of gait recognition such as the effects of clothes, shoes, shadows, carrying conditions and the uncontrolled environment.

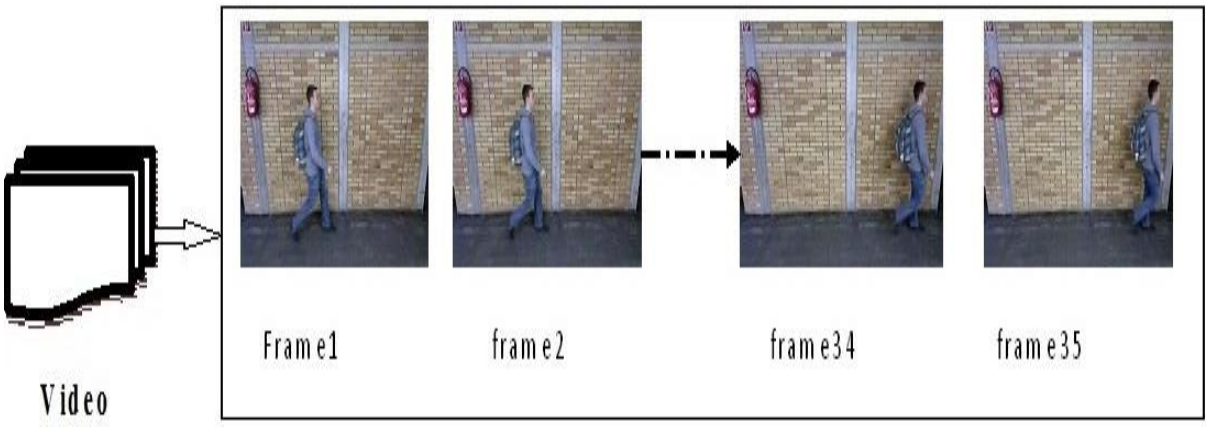


Figure 1-5. Extract all the image frames from the video file (TUM database).

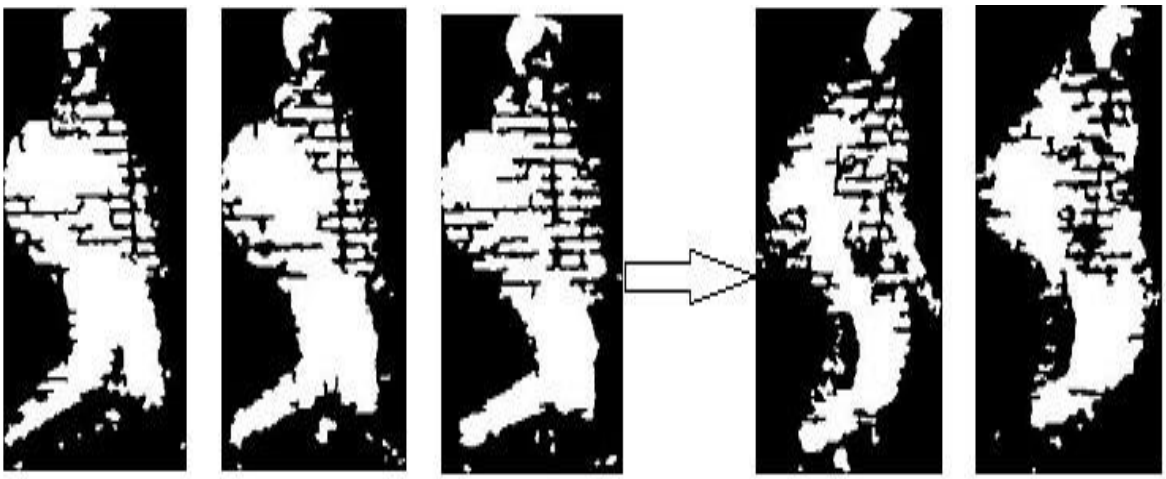


Figure 1-6. Silhouette images include missing body parts, noise, and shadows.

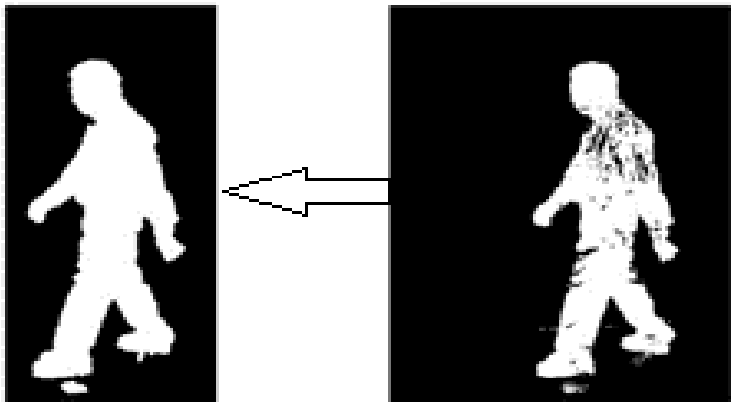


Figure 1-7. Example for pre-processed silhouettes (TOTON database).

1.2.3 Objectives

The overall objective of this work is to investigate and propose methods for robust gait recognition under covariate conditions across various viewing angles. The project aims to investigate and develop gait recognition approaches with improved recognition performances to handle covariate factors, which are perceived to adversely affect recognition performances. Hence, one of the major goals of this work is to improve the execution of existing gait recognition methods in the presence of variable covariate conditions in the probe set. Such a set-up diligently mimics practical environments and truly tests the effectiveness of gait as a biometric modality in unconstrained conditions. Another objective of this work is to address the limitations of existing cross view gait recognition methods to improve recognition performance across viewing angle changes.

Therefore, this thesis investigates novel gait identification methods to improve the recognition performances by better understanding:

- How covariate factors behave while an individual is in motion.
- The impact covariate factors have on different gait conditions.

Four supervised feature extraction methods have been proposed in this thesis for gait recognition based on texture descriptors extracted from the Gait Energy Image (GEI). These methods are the Haralick features, Gabor filter bank, Local Binary Pattern and Wavelet transform.

The proposed methods described below are capable of extracting the most discriminative features from the GEI under different covariates or conditions, thus improving recognition performances:

- Haralick texture descriptors via RELIEF selection algorithm.
- Multi-scale Local Binary Pattern descriptor via Spectral Regression Kernel Discriminate Analysis (SRKDA) reduction algorithm.
- Gabor filter bank descriptor via several reduction algorithms (SRKDA, Kernel Principal Component Analysis (KPCA) and Maximum Margin Projection (MMP)).
- Wavelet transforms approach via SRKDA reduction algorithm.

1.3 The Structure of the Thesis

This thesis consists of seven chapters including chapter 1 and is summarised as follows:

- Chapter Two discusses existing research on gait recognition with an emphasis on robust recognition under variable covariate conditions. In addition, the chapter gives some background on the different types of gait representation including the most commonly known gait databases and approaches related to this research investigation.
- Chapter Three describes the principles of gait recognition based on the concept of GEI. It explains the basic procedure of gait recognition approach and all techniques applied to gait identification (e.g. feature extraction, feature selection, feature reduction and classification).

- Chapter Four proposes a gait recognition method based on the Haralick texture features using RELIEF algorithm to select relevant features generated by GEI. The algorithm is validated using available datasets (CASIA and USF); an analysis of the results is then given to gauge the effectiveness of the proposed technique.
- Chapter Five discusses a gait recognition technique based on a set of Multi-Scale Descriptors for feature extraction using Multi-Scale Local Binary Pattern (MLBP) and Gabor filter bank. The validation approach is based on CASIA and USF datasets and the results obtained are evaluated and contrasted against some existing methods.
- Chapter Six proposes a gait recognition technique based on the wavelet coefficients using the Haar wavelet transform with SRKDA algorithm. Experiments are carried out using the CASIA dataset and the obtained results are then evaluated.
- Chapter Seven gives a summary of the contributions of the research, including future work.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

A biometric system involves recognising a pattern from a person. Such a system is based on a unique feature, derived from either a physiological or behavioural characteristic. Biometric technology, including behavioural and physiological modalities of humans, has been discussed earlier in the previous chapter.

Gait is a behavioural biometric that has become an important behavioural characteristic to identify people by the way they walk, however, the human gain may not remain invariant, especially over long periods of time. This is due to fluctuations in body weight, major injuries involving joints or brain or due to the effect of various other covariates which include variations in clothing and carrying conditions (briefcase, handbag, etc.).

Therefore, this research investigates a gait recognition system that can utilise useful and reliable attributes to operate under the conditions mentioned above.

A survey of person gait as a biometric technology can be found in the work of Boyd and Little (Boyd et al., 2005). Nixon and Carter introduced general surveys of human gait from a computer vision point of view (M. S. Nixon et al., 2004), (M. S. Nixon et al., 2006), (Mark S Nixon et al., 2012) and Liu et al. (L.-F. Liu et al., 2009). A more approach-centred survey of gait recognition from a model-based perspective is provided by Yam (Yam et al., 2015) and Nixon (Chew-Yean Yam, 2009). The work of Gafurov (Gafurov, 2007) goes beyond the boundaries of basic visual approaches and, along with vision-based methods, also surveys the use of other sensors in gait literature.

It makes sense to first define and review the human walking style and feature extraction techniques before providing a review of the state-of-the-art approaches in gait recognition research. Then, a review of related topics of motion recognition, which are related to the larger field of person motion analysis and gait representation through different types of gait database, is discussed. Finally, the challenges of covariate factors and standardised datasets used for validation are discussed.

2.2 What is Human Gait?

A gait cycle is a time taken between successive instances of first foot-to-ground contact for the same foot, also called the walking phase (see Figure 2-1). Each foot has two stages: a standing stage and a swing phase. When the foot is in touch with the ground, it is called the stance or standing phase. On the other hand, when the foot is lifted and moved forwards, this is classified as the swing phase. The phase begins with the heel strike of one foot marking the start of the stance cycle. The lower leg flexes to bring the foot level on the floor and the body weight is moved onto it. The other leg swings through in front as the heel lifts off the ground. The body weight is transmitted onto the other foot, making the knee to bend. The foot, which is behind, then lifts off the ground, finishing the stance phase. (Cunado et al., 2003).

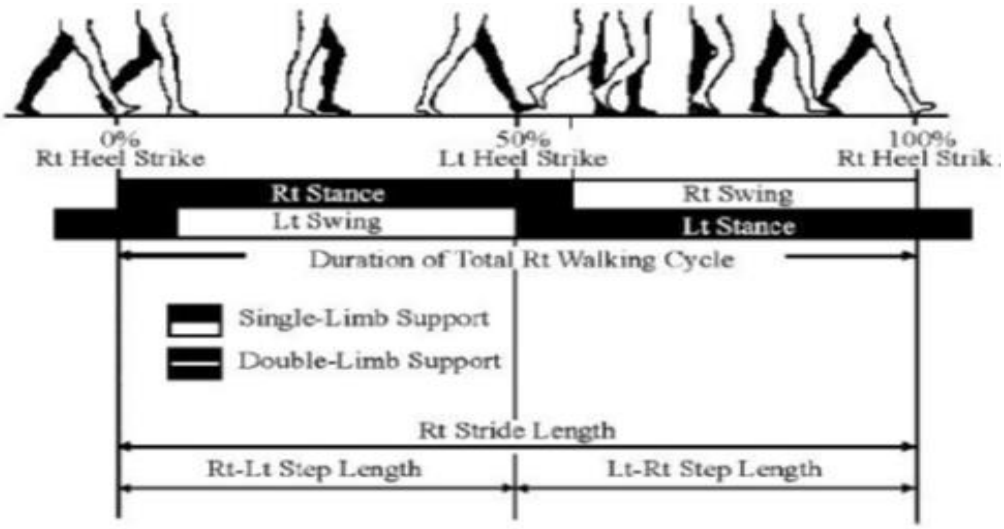


Figure 2-1. The style walking (Cunado et al., 2003).

2.3 Gait Representation

The relative motion between different body parts during walking characterises the human gait. However, researchers have previously proposed gait representation approaches by extracting silhouettes images such as GEI (Ju et al., 2006), self-similarity images (Ben Abdelkader et al., 2004), Gait Entropy Images (GEnI) (K. Bashir et al., 2009) and Shape Variation-Based Frieze Pattern (SVB Frieze pattern) (Lee et al., 2007). In addition, four directional variations of Gradient GEI have been used for gait recognition by Guru et al (Guru et al., 2016). The proposed based on feature level fusion of four directional vector's (i.e. horizontal, vertical, forward and backward diagonal). Verlekar proposed a system to identify the walking direction using a perceptual hash (PHash) computed over the leg region of the GEI in order to solve some of the challenges encountered by gait recognition (Verlekar et al., 2017). Chaurasia proposed a gait feature representation (i.e., P_{RW} D_F GEI), where the RW-based method is used for image segmentation and the segmentation problem is solved using Poisson's equation, and where the resulting feature, called P_{RW} GEI. D_F GEI, is a discrete Fourier transform (DFT)-based gait feature (Chaurasia et al., 2017).

2.3.1 Extraction of Silhouettes Image

Research on gait recognition has been conducted to enable the recognition of the gait of an individual from a video footage by extracting and processing the information related to the motion using the concept of GEI. A typical scheme for a GEI includes the following steps:

- Firstly, the video stream or video data is a series of consecutive images. The aim of this step is to convert the video into images, also called frames.

- Secondly, a silhouette image i.e. a binary image of an individual, is generated using an appropriate image segmentation algorithm e.g. thresholding, background modelling, foreground modelling etc.

Silhouette images are extracted from original human walking sequences. A silhouette pre-processing procedure (A.K. Jain, 1989) is then applied on the extracted silhouette sequences. This pre-processing step includes size normalisation (i.e. proportionally resizing each silhouette image so that all silhouettes have the same height) and horizontal alignment (i.e. centring the upper half silhouette part with respect to its horizontal centroid).

2.3.2 Average Silhouettes

According to Liu and Sarkar, the research community started to shift towards static signature due to the increased computational cost of temporal matching. Liu and Sarkar have proposed the use of an average image of a silhouette called a gait energy image (Z Liu et al., 2004). This concept has been proposed against some algorithms such as the baseline one. A GEI is a compact representation of gait (a gait cycle is represented using just one image), is easy to compute, and is insensitive to noise. A GEI can be seen as the sum of images of the walking extracted silhouettes divided by the number of images of the video stream and can be defined as follows:

$$G(x,y) = \frac{1}{N}\sum_{t=1}^{N} I(x,y,t) \tag{2.1}$$

where N and t are the number of frames within a complete gait cycle and the frame number in the gait cycle, respectively. I is the silhouette image whose pixel coordinates are located at $x\ and\ y$ positions. Figure 2-2 shows an example of the silhouette of an individual.

Figure 2-2. Silhouette image (Z. Liu et al., 2004).

2.3.3 Gait Entropy Image

Figure 2-3. Giat Entropy Images (K. Bashir et al., 2009).

Bashir et al. proposed to differentiate between the dynamic and static areas of the GEI by calculating the Shannon entropy at each pixel location of a GEI (K. Bashir et al., 2009). Shannon entropy measures the uncertainty associated with a random variable. The pixel values of a silhouette image in the dynamic areas are more obscure having the highest values of the entropy, which is clearly shown in Figure 2-3. The legs and arms show more motion compared to other body parts and are represented by higher intensity values. A GEnI can be used to select the information gait features from the GEI. Figure 2-3 shows Gait Entropy Images from the CASIA database in (K. Bashir et al., 2009).

2.3.4 Flow Field

(Khalid Bashir et al., 2009) proposed the use of the optical flow as a feature extraction method for gait recognition. The optical flow was displayed in four directions framing four templates portraying the movement (see Figure 2-4). Another template was additionally framed

representing the static body parts. For the recognition purposes, a score was computed for every four templates (the template for negative y-axis was disregarded), and finally, a final recognition score was computed by using these individual scores. Although the algorithm is slightly computationally costly, it performs well against covariates and noise distortions. Their proposed method achieved recognition results of 79.50%, 83.60% and 48.80% for walking normal, carrying a bag and wearing a coat, respectively.

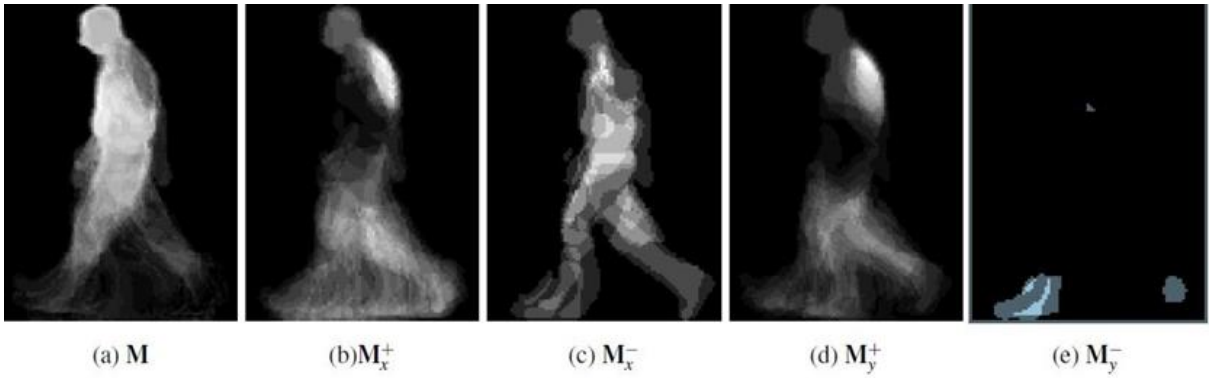


Figure 2-4. Example of the 5 motion descriptors proposed in (Khalid Bashir et al., 2009).

2.3.5 Symmetry

Given the symmetrical nature of the gait (Hayfron-Acquah et al., 2003); a symmetry operator to extract the features of the cycle is suggested. The operator works well against noise and low-resolution data and this is in accordance with the fact that the operator gives more importance to the symmetric nature of moving object's description by temporal symmetry. Although the algorithm is marginally computationally expensive, it provides a quite strong signature from a small amount of training data. This method uses the Generalised Symmetry Operator, which locates features according to their symmetrical properties by using the symmetry operator, the Discrete Fourier Transform and a k-nearest neighbour approach. The results produced encouraging recognition rates on a small SOTON database. Furthermore, the larger database had almost the same results as those obtained from the smaller database. Figure 2-5 shows the

symmetry image from SOTON data. Following this, an overview of the steps involved in extracting symmetry from silhouette information is given.

First, the image background was computed from the median of five image frames and subtracted from the original image (Figure 2-5a) to obtain the silhouette image (Figure 2-5b). This was possible because the camera used to capture the image sequences was static and there was no translational motion. Additionally, the subjects were walking at a constant pace. The Sobel operator was then applied to the image in (Figure 2-5b) to derive its edge-map, as shown in (Figure 2-5c). To remove edges and reduce weak strength noise, the edge-map was thresholded to set all points beneath a chosen threshold to zero. These processes reduce the amount of computation in the symmetry calculation. The symmetry operator was then applied to give the symmetry map, as shown in Figure 2-5d. For each image sequence, the gait signature was obtained by averaging all the symmetry maps.

(a) original image (b) silhouette image (c) after Sobel (d) symmetry map.

Figure 2-5. Image from the SOTON data showing the concept of symmetry (Hayfron-Acquah et al., 2003).

2.3.6 Silhouette Similarity

Using the NIST/USF baseline approach Sarkar et al. (Sudeep Sarkar et al., 2005) performed gait recognition by using a temporal correlation of the silhouettes extracted. The aim was to develop a technique to improve the recognition performances. The proposed approach is evaluated on the Mobo data and on the NIST/USF data. Sarkar et al. extracted a silhouette image in two steps:

- Compute the background statistics of the RGB values at each image location, and then calculate the mean and the covariance of the RGB values at each pixel location. The last point in this step is to compute the Mahalanobis distance in RGB-space for the pixel value from the estimated mean background value.

- Scale and centre the silhouette image.

In this approach, some covariates can affect the recognition result rates such as time and different surfaces. For the remaining conditions e.g. view, briefcase and shoe, the results were acceptable compared to a selection of other methods (see Figure 2-6).

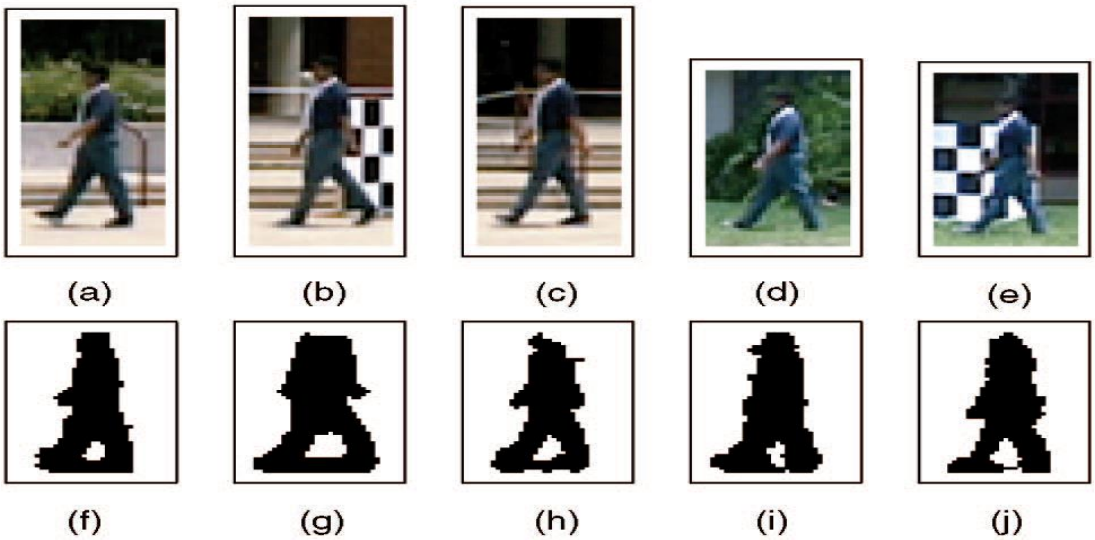


Figure 2-6. Sample image of USF data as viewed the top row (a) to (e) with shows sample silhouette in the bottom row (f) to (j), (S. Sarkar et al., 2005).

2.3.7 Skeletal Image

Blum (1967) has used skeletons to represent shapes for numerous computer vision tasks. However, skeleton representations are not used frequently for human gait recognition because of their sensitivity to boundary noise causes imperfect extraction of the features. This sensitivity relates to the following:

1) Walking activity causes the body to self-occlude.

2) Silhouette quality has a direct effect on skeleton precision.

Previously, (Whytock et al., 2014) has presented a model-free skeleton approach based on smooth distance functions generated from a Poisson equation using a Skeleton Variance Image (SVIM) for human gait recognition. The smooth distance function reduces the sensitivity to boundary noise and yields a robust skeleton as shown in Figure 2-7. In this approach, experimental results using the Nearest Neighbour classifier yielded noteworthy results of 98.4% and 64.2% for normal walking and carrying a bag, respectively.

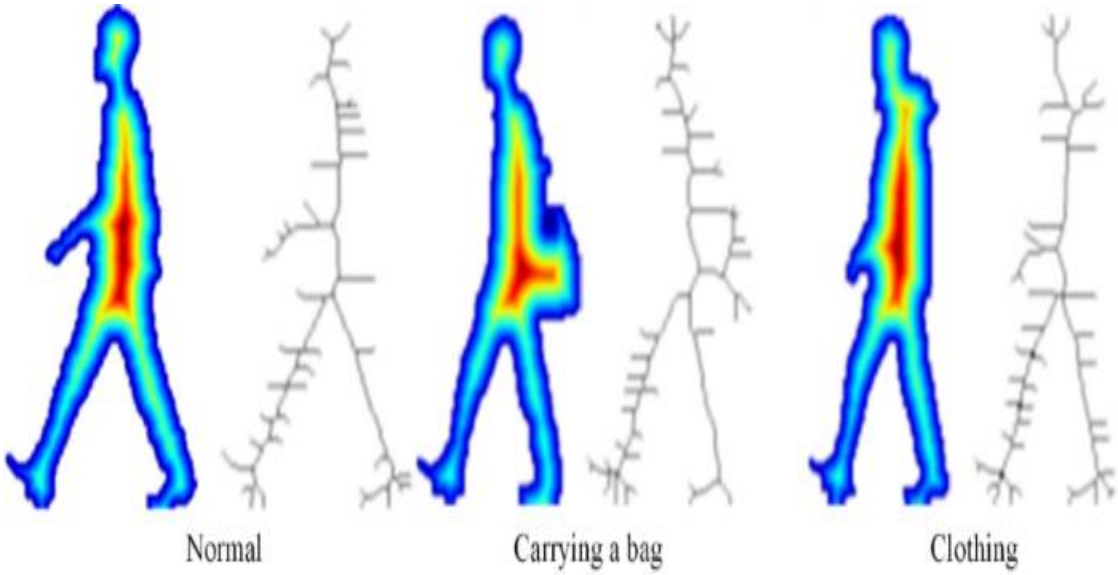


Figure 2-7. The distance function generated by the Euclidean metric demonstrates the retention of boundary noise across cool and hot colours. The skeleton extracted by the medial axis transform. Using TUM GAID Dataset (Whytock et al., 2014).

2.4 Gait Recognition Approaches

Human gait recognition techniques can be split into two approaches: model-based approaches and model-free approaches.

2.4.1 Model-based Approaches

These approaches are derived from the movement of the torso and/or the legs of a person. The distinction of a structural approach is one, which uses static body parameters for recognition, such as stride length (Ben Abdelkader et al., 2002). A model can be the motion of the angles

between the limbs or hip rotation etc. One such model-based approach was proposed by Yam (C. Yam et al., 2004), where the authors extended the existing model-based approaches and differentiated between running and walking.

Lu proposed a Layered Deformable Model (LDM) for the human body to enhance human gait analysis (Haiping et al., 2006). The model of LDM determines the body part lengths, widths and the positions and joint angles of the human body using 22 parameters. The LDM model consists of four layers and allows for limb deformation. The proposed method recovers its parameters (and thus the human body pose) from automatically extracted silhouettes. The experiments show that an average error rate of 7.00% is achieved for the lower limb joint angles, which is important for model-based gait recognition.

Although the feature space extracted from these algorithms has significant discriminatory power and is more robust to clothing changes and a slight change in viewpoint, usually this approach tends to be computationally intensive than the model-free counterpart. For example, Zhao et al. (Guoying et al., 2006) suggest performing a 3D gait recognition using multiple cameras. Nevertheless, the registration of gait images across the camera view is nontrivial, even in a well-controlled environment with a clean background and little noise.

A model-based approach for gait recognition employing a five-link biped locomotion human model has been proposed by Zhang in (R. Zhang et al., 2007) which introduces the idea of a Sagittal plane (plane bisecting the human body) in which most gait movements are carried out. More recently, the problem of gait recognition has been approached from a control systems perspective (Tao, 2008).

Kusakunniran in (Kusakunniran et al., 2009) and (Kusakunniran et al., 2010) have proposed two multi-view gait recognition methods, referred to as View Transformation Model (VTM). The proposed approach in (Kusakunniran et al., 2009) is based on spatial domain GEI by

adopting Singular Value Decomposition (SVD) technique while the approach in (Kusakunniran et al., 2010) is based on creating a VTM from a different point of view using Support Vector Regression (SVR).

Arora in (Arora et al., 2016) has proposed a technique for human identification based on the body structure and gait. The gait features extracted are height, hip, neck and knee trajectories of the human silhouette from the body structure. The proposed method includes two new parametric curves, a Bezier curve and a Hermit curve, based on gait pattern. The projected approach has been applied on the SOTON covariate database. Nevertheless, existing model-based approaches mostly require high-resolution images to correctly extract the model parameters from a gait sequence.

2.4.2 Model-free Approaches

A model-free approach employs the features of the gait which are derived from the moving shape of the subject. In this method, the gait signature is derived from the spatiotemporal patterns of a walking person (Niyogi et al., 1994), or the 2-D optical flow of the individual (J. Little et al., 1995) and (James Little et al., 1998).

The variation of the area within a particular region (Foster et al., 2003) and extraction of the gait features from an enhanced human silhouette image are then performed. The gait features are generated from a human silhouette by determining the skeleton from body segments (Ng et al., 2011).

In the following, a review of the representative works of gait recognition is presented by following the flow of information through a gait recognition system from a model-free approach.

For example, in (Kumar et al., 2014), the authors proposed a method of gait recognition system using GEI and LBP techniques to extract features from the gait representation. The LBP operator is applied to extract the features from the entire GEI and the Region Bounded by Legs (RBL). The process was implemented in instances (covariate factors) of a gait data, such as a changes in clothing, carrying a bag and different normal walking conditions. This technique achieves a performance of 85.66% in terms of Correct Classification Rate (CCR) for a side view 90° with the CASIA database. Above all, the input to the system is a sequence of binary silhouettes that are acquired using a background subtraction method. These silhouettes are used to calculate the gait cycle. Once the silhouettes are aligned, they are subjected to a Radon Transform to generate a Radon template, from which a set of features is extracted using Linear Discriminant Analysis (LDA).

Hu in (Hu et al., 2013) proposed incremental learning for video-based gait recognition with LBP flow. The proposed method is based on optical flow including dynamics learning, pattern retrieval and recognition. The LBP is employed to describe the texture information of optical flow. The proposed achieved 60.70% in terms of CCR using the CASIA Dataset B.

Recently, some studies have tried to strengthen the model-free approach against covariates. The authors in (Whytock et al., 2014) proposed to use a screened Poisson equation with tuneable smooth distance functions using SVIM. The method uses the SVIM from time-based sequences given that gait motion is more consistent over time compared to the appearance.

Rida in (Rida et al., 2016) proposed a gait recognition method based on Modified Phase-Only Correlation computed from GEI. In this approach, a bandpass spectral weighting function of the well-known phase only correlation matching technique was employed to deal with the small texture features; resulting in improved performances. The algorithm achieved 81.40% in terms

of CCR using the CASIA database under the effect of clothing and carrying conditions for a side view of 90º.

Dupuis in (Dupuis et al., 2013) proposed an interesting feature selection method based on random forest rank features algorithm for gait recognition. The proposed feature selection has reduced the computational cost while achieving a recognition performance of up to 85.6% when using CASIA Dataset B.

Zhao suggested in (Zhao et al., 2016) the use of the Sparse Tensor Discriminative Locality Alignment (STDLA) algorithm for gait feature recognition. The STDLA algorithm consists of two sections; one is tensor manifold learning and the other is sparse projection. The proposed algorithm effectively avoids the dimensionality dilemma and overcomes the small-sample-size problem. Additionally, a sparse projection is able to control the weights of the original variables and decrease the variance brought by the possible over-fitting resulting from the least increment of the bias. The experiment was carried out on the USF human-ID Gait database.

Wang in (X. Wang et al., 2017) proposed a gait recognition technique based on Gabor wavelets and $(2D)^2$ PCA. The proposed technique consists of three steps; firstly, the GEI is formed by extracting different orientation and scale information from the Gabor wavelet. Secondly, a two-dimensional principal component analysis $(2D)^2$ PCA method is employed to reduce the feature space dimension. The $(2D)^2$ PCA method minimises the within-class distance and maximises the between-class distance. Finally, the multi-class SVM is adapted to recognise different gaits. Experimental results performed on the CASIA gait database showed that the proposed gait recognition algorithm is generally robust, and provides up to 93.29% of higher recognition accuracy. One limitation of the proposed approach is that the generated GEI lose some dynamic information, since they are calculated by averaging a series of images.

2.5 Gait Databases

There was a need to make consistent datasets in order to help researchers evaluate and compare their results and to identify any potential limitations in order to help enhance the performances including their robustness. The validation of proposed algorithms and techniques using various datasets is an important aim to guarantee that the approaches or parameters are not biased.

2.5.1 Database needs

A database should satisfy the following points:

- High individual/activity class numbers for between class and intra-class variety.
- Real environment differential i.e., not choreography.
- Manifold image sequences for individuals/ action classes.
- The sequences of training and testing of the dataset should be separate with agreed standards.
- Real Environment single changeable factors and coupled variable factors.

The following sections discuss the databases used in gait recognition. Only two have been used in this thesis for the purpose of consistent comparative studies of the proposed algorithms.

2.5.2 NIST/USF Database

This database relates to the Human ID challenge problem and consists of 452 images from 75 persons using a video collected for each person from two camera views having different surface conditions and shoe types (Phillips et al., 2002). The data was gathered in an outdoor environment reflecting the additional confusion of shadow and sunlight movement in the background and moving shadows because of use of cloud cover. This database is the largest

available in terms of the number of people, video sequences and the variety of conditions under which a person's gait has been recorded.

Later, the database was extended to be 1,870 sequences from 122 subjects (Z. Liu et al., 2004) (see Figure 2-8). Each subject walked around two similar sized elliptical courses, one on a grass lawn and the other on concrete. Two cameras viewed each course. The cameras were positioned nearly 15 meters from each end of the ellipse with lines of sight adjusted to view the whole ellipse. Information recorded in addition to the video includes sex (75 % male), age (19 to 54 years), weight (43.1 kg to 122.6 kg), height (1.47m to 1.91m), foot dominance (mostly right), type of shoe (sandal, sneakers, etc.) and heel height. A little over half of the subjects walked in two different shoe types. Thus, for each subject, there were up to eight video sequences: concrete (C) or grass (G)*, there were two cameras, R or L * and Shoe A or Shoe B.

The dataset is quite demanding for other biometric modalities since gait is the only biometric trait that can be captured where the lighting is uncontrolled. At the University of South Florida, Tampa, about 33 subjects were used to collect the data which was partitioned into 32 subsets based on the various combinations of five covariates as listed below:

- Surface type concrete (C) or grass (G).
- Shoe type (A or B).
- Viewpoint right camera (R) or left camera (L).
- Carrying conditions of with briefcase (BF) or no briefcase (NB).
- Time (tags sequences from May and those from new subjects in November collections, tags sequences from November repeat subjects).

Table 2-1 below lists the 12 possible experiments A through L. The gallery set is G, A, R, NB based on 122 subjects (71 subjects from May data).

Table 2-1. Probe dataset USF.

Exp	Probe Sets	Total Number Subjects	Number Subjects May	Covariates difference between Gallery and Probe
A	(G,A,L,NB,t_1)	122	71	View
B	(G,B,R,NB,t_1)	54	41	Shoe
C	(G,B,L,NB,t_1)	54	41	View, Shoe
D	(C,A,R,NB,t_1)	121	70	Surface
E	(C,B,R,NB, t1)	60	44	Surface, Shoe
F	(C,A,L,NB, t1)	121	70	Surface, View
G	(C,B,L,NB, t1)	60	44	Surface, View, Shoe
H	(G,A,R,BF, t1)	120	70	Briefcase
I	(G,B,R,BF, t1)	60	47	Shoe, Briefcase
J	(G,A,L,BF, t1)	120	70	View , Briefcase
K	(G,A/B,R,NB,t2)	33	33	Time(Shoe , Clothing)
L	(C,A/B,R,NB, t2)	33	33	Surface , Time

Figure 2-8. Beginning, middle, and another frame of the example gait sequence in (Liu et al., 2004).

2.5.3 CASIA Database

To promote the research gait recognition, the Institute of Automation at the Chinese Academy of Sciences (CASIA) has constructed several Gait Databases. The databases are available from the Centre for Biometrics and Security Research (CBSR, 2005). The CASIA Gait Database has three types of datasets: A, B (multi-view dataset) and **C** (infrared datasets).

2.5.3.1 CASIA Dataset A

In December 2001, Database-A was created and includes 20 individuals where each has 12 image sequences; four sequences for each of the three directions - parallel, 45 degrees and 90 degrees - to the image plane (see Figure 2-9). The length of each sequence is not identical for variation in the walker's speed, but it ranges from 37 to 127. The size of Dataset A includes 19,139 images (see Figure 2-9).

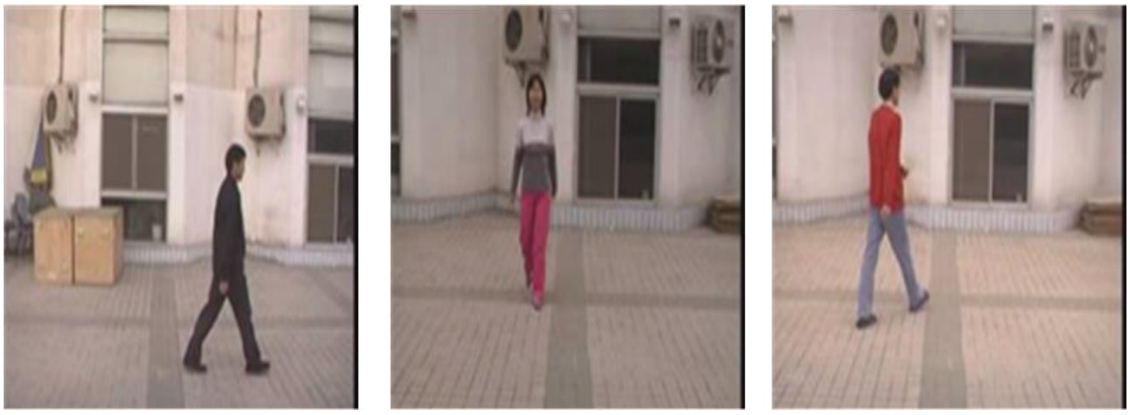


Figure 2-9. The sequence for each of the three directions.

2.5.3.2 CASIA Dataset B

This is a large multi-view gait database created in January 2005. It consists of 124 subjects, and the gait data was taken from 11 angles using 11 cameras on the left-hand side of the person as they were walking, with the angle between the nearest view directions at 18°. When a person walks into the scene, they were first asked to walk normally along a straight line six times, and

thus 11×6 = 66 normal walking video sequences were captured for each person. Walking with clothing, normal walk, and the carrying condition is shown in Figure 2-10.

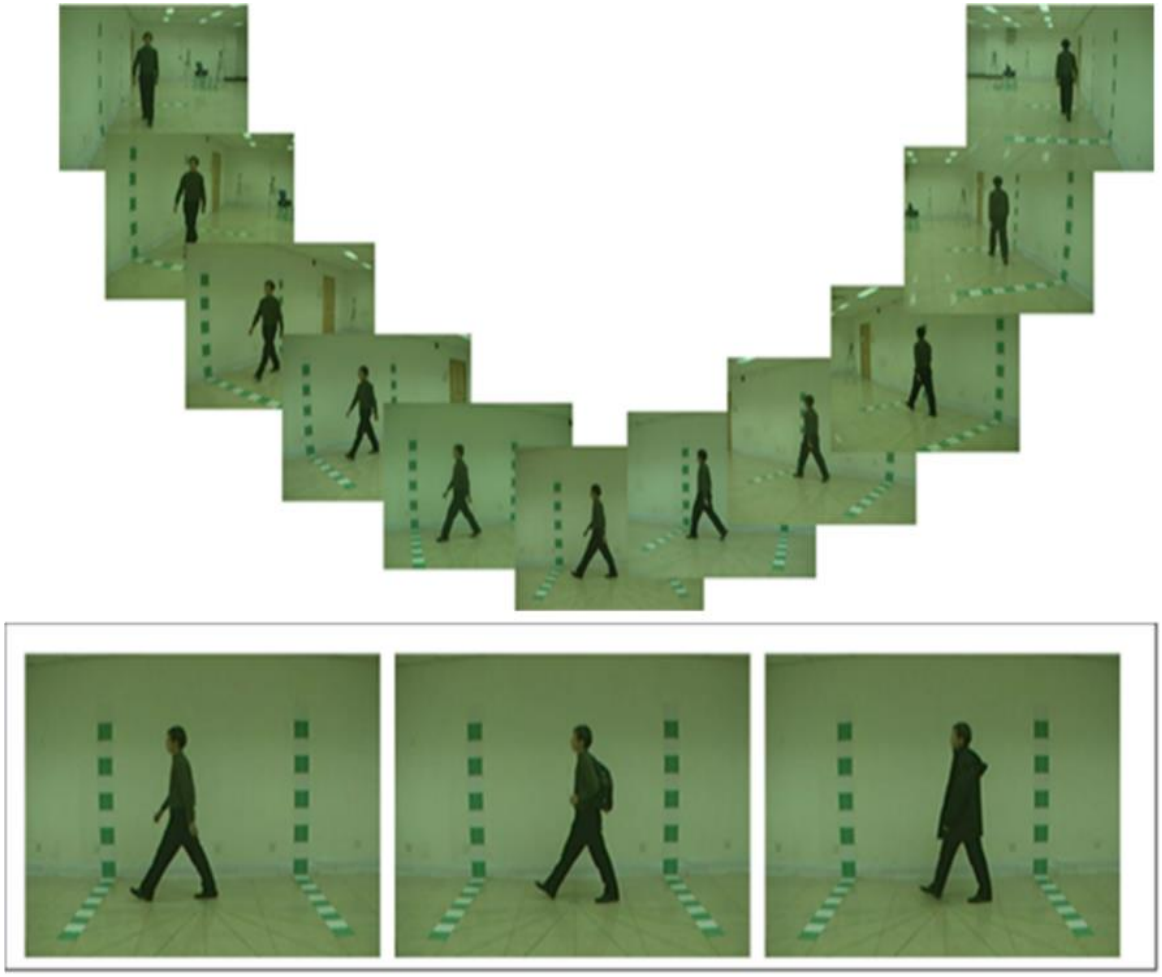


Figure 2-10. Different conditions of walking at different angles, CASIA Dataset B.

2.5.4 TUM GAID Database

The Technical University Munich GAID database was created in 2012 (M. Hofmann et al., 2012). The database currently consists of colour video, depth and audio with 305 individuals under different walking conditions. This dataset is one of the biggest to date. To further evaluate challenges of time variation, a subset of 32 people was recorded a second time (See Figure 2-11). The TUM GAID database was captured in different sessions, one in January 2012 at a lower temperature (-15°) and one in April 2012 at a temperature higher (+15°). A Microsoft

Kinect sensor was used to record a video stream, a depth stream and four-channel audio. Video and depth have the same resolution 640×480 pixels at a frame rate of nearly 30 frames per second (Martin Hofmann et al., 2014).

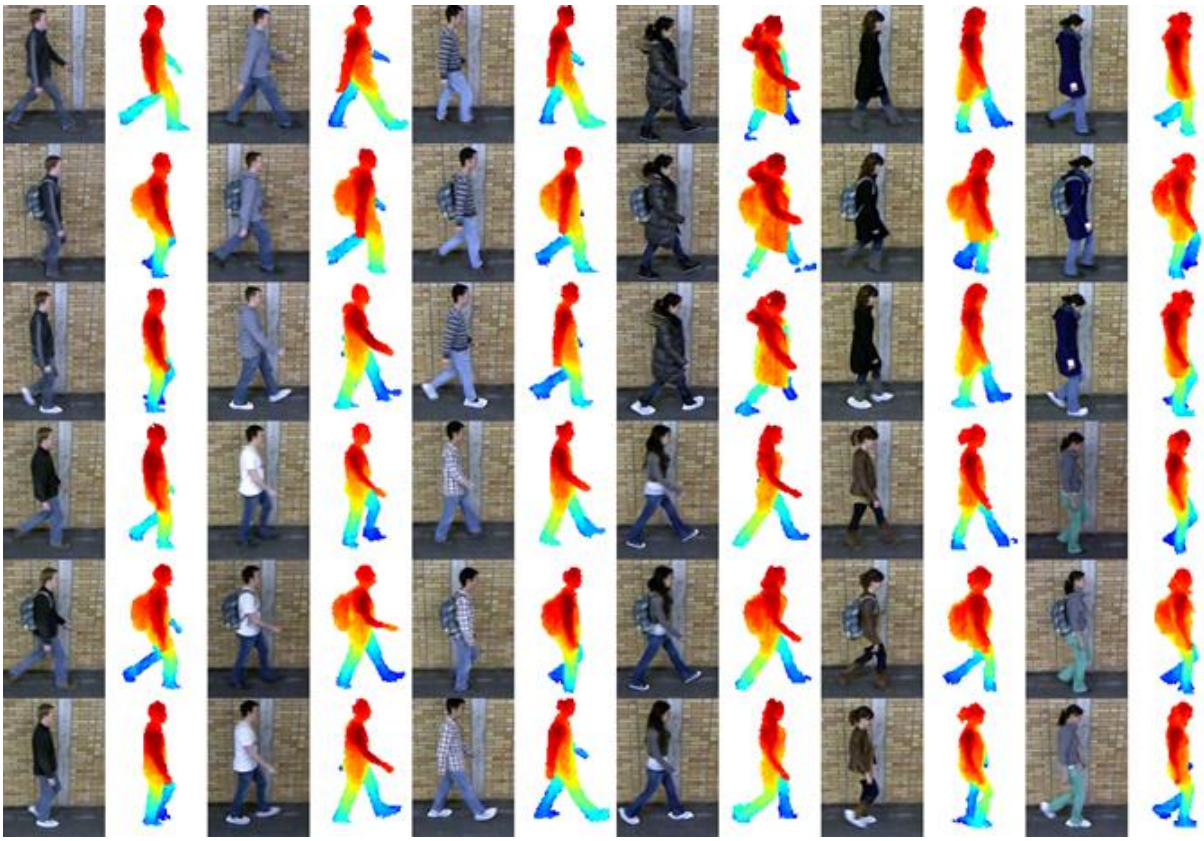


Figure 2-11. Database of Technical University Munich ((M. Hofmann et al., 2012).

2.5.5 Southampton Database (SOTON Gait Data)

The SOTON database contains one small and one large dataset. The small database (with 11 subjects) was created with the aim of probing the robustness of gait recognition for imagery using the same subject in various covariate conditions (carrying items, wearing different clothing or footwear). The small dataset was designed to investigate the robustness of gait recognition techniques under changing covariate conditions, including carrying objects and clothing. Moreover, the small dataset contains one normal sequence for each subject, four carrying-bag sequences and one coat-wearing sequence.

On the other hand, the large database (with 116 subjects) consists of six subsets named A to F. In the literature, set A has been most widely used, and in it, all subjects were captured under both normal and fixed covariate conditions. This dataset is aimed to facilitate two inquiries: whether the gait is single across an important number of subjects in normal cases, and the need for research to be directed toward biometric techniques. Figure 2-12 shows Southampton Human ID: a small but more detailed database and a large but basic database (M. S. Nixon, 2002). Note: "The databases (TUM GAID and SOTON) had not be used because these could not be obtained from the source".

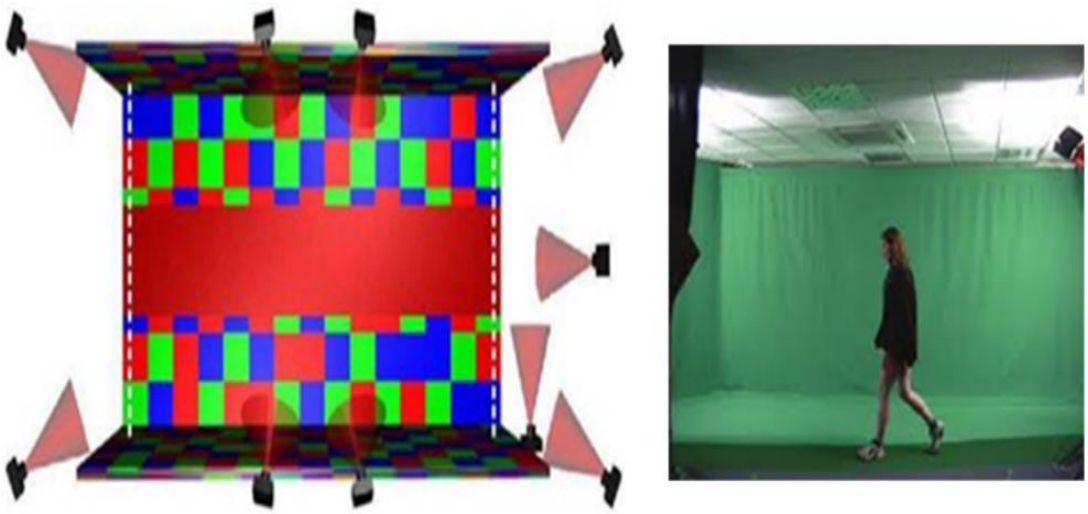


Figure 2-12. Southampton Human ID.

2.6 Summary

As introduced; the principle of human gait, gait representation and gait recognition approaches consist of two aspects: model-based/model-free and gait databases. The preceding review has covered essential techniques and works in the literature regarding gait recognition and in particular gait representation. Most of the literature in gait revolves around a gait recognition approach which is capable of selecting information characteristics for human identification under different conditions. Table 2-2 presents a summary of the work related to the GEI representation.

Table 2-2. Summary of related work.

Author	Year	Datasets	Approach	Gait representation	Technique used
Han, et al.	2006	USF Human ID	Model-based	GEnI	Synthetic templates and a statistical with PCA and MDA
K Bashir et al.	2009	CASIA + SOTON	Model-free approach	Contour of a silhouette image	Feature selection with an Adaptive CDA
K Bashir et al.	2010	CASIA + SOTON	Model-free approach	GEnI	Feature selection mask and CDA.
Rida et al	2014	CASIA	Model-free approach	GEnI	Modified Phase Only Correlation.
Kumar	2014	CASIA	Model-free approach	GEnI	LPB
Dupuis et al	2013	CASIA	Model-free approach	GEI	Random Forest rank features algorithm.
Whytock et al	2014	TUM GAID	Model-free approach	SVIM	The Screened Poisson distance Function.

Based on the findings of the literature review, GEI was adopted in this study because GEI representation explicitly captures the shape of the subject in question and implicitly captures the dynamic parts of body. Pixels with high-intensity values in GEI correspond to body parts that show little movement during a walking cycle (e.g. torso), while pixels with low-intensity values correspond to body parts that move constantly (e.g. legs and head). In addition, the GEI representation is less sensitive to silhouette noise in individual frames.

CHAPTER THREE: GEI-BASED GAIT RECOGNITION

3.1 Introduction

Although some gait recognition approaches have been proposed in the literature, the algorithms share the common goal of ensuring the best trade-off between the recognition performance and computational complexity. After an investigation of the state of the art and early works; we decided to use GEI approaches. This work is divided into two parts: the first part relates to a review of some existing methods based on GEI and their evaluation using the two databases (CASIA, USF), as presented in Chapter 2. The second part is based around a familiarisation process and uses the results in the validation and evaluation of the proposed methods that are described in Chapters 4, 5 and 6. The aim of this chapter is to give an introduction to human gait recognition approaches including feature extraction, feature selection data reduction and classification.

3.2 Gait Energy Image

A GEI is one of the most widely used methods for extracting the relevant feature descriptors of human gait and has proven to be one of the most effective techniques. A GEI is a representation of a human walking; using a single grey scale image obtained by averaging the silhouettes extracted over one gait cycle (Ju et al., 2006). A GEI can be seen as the average of images of the walking silhouette and is defined as shown by Equation (2.1). Figure 3-1 shows an example GEI of an individual under different conditions. Pixels with low intensity correspond to the dynamic parts of the body which are widely used for recognition and are

usefully not affected by the carrying and clothing conditions commonly referred to as covariate factors. Conversely, pixels with high intensities correspond to the static parts of the body containing the body shape information used for identification, but these can be affected by covariate conditions (e.g. carrying a bag, wearing a coat) (Bashir et al., 2010). Figure 3-2 shows examples GEI in USF Human ID database under different condition presented on the many Probes for example, carrying a briefcase, without the briefcase, different surfaces and different directions. A GEI is used to select informative gait features in our proposed approach.

Figure 3-1. An example of the GEI of an individual under different conditions (CASIA database).

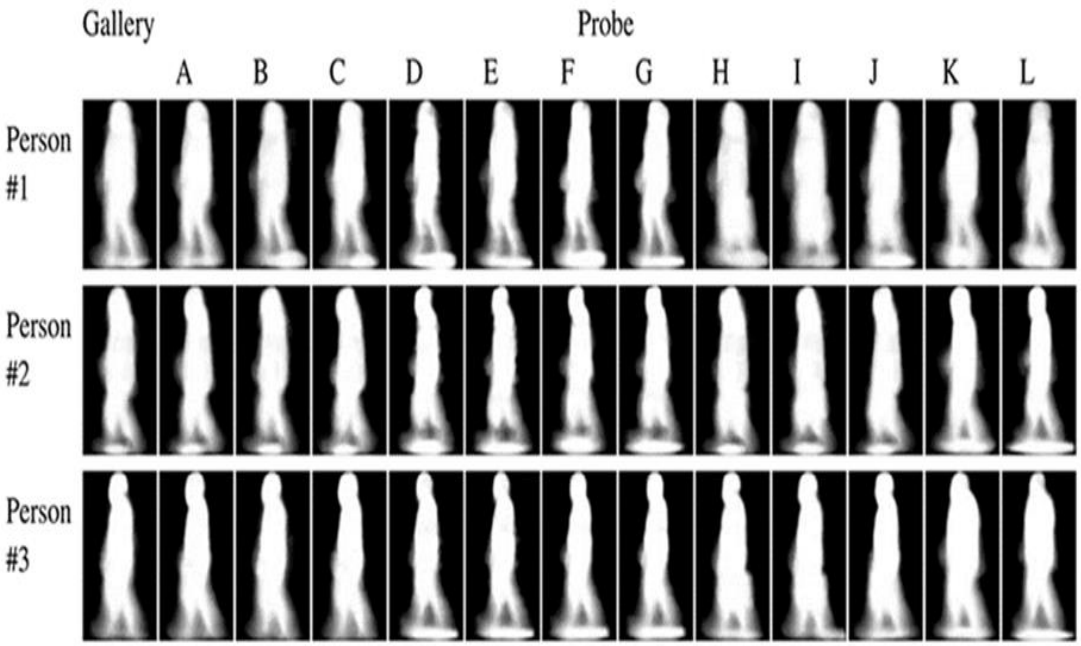


Figure 3-2. An example of the GEI of an individual under different conditions in USF Human ID database.

3.3 Gait Recognition

Human gait recognition refers to verifying and/or identifying persons using their walking style under covariate factors (i.e. carrying and clothing conditions). GEI-based gait recognition is one of the most recent effective biometric systems, having high recognition rates with low computational complexity. Such a system includes the following four steps:

1. Feature extraction- extracts the discriminating features from the gait representation (GEI in this case) to characterise the gait under variations of covariate factors such as clothing and carrying conditions.
2. Feature selection- selects a subset of relevant features from the GEI representation.

or

3. Feature reduction- aims to reduce the feature data into a much lower dimensional space.
4. Classification process- makes a decision about the recognition of the gait in question using a classification process with the selected feature vector.

Illustrates Figure 3-3 illustrates the human gait recognition system diagram.

The information contained in GEI was investigated and a number of feature extraction methods under various clothing and carrying conditions were proposed. The main idea is to exploit the locally discriminating features that characterise these conditions by dividing the GEI horizontally and/or vertically in three (top, medium and bottom) and/or two equal (left and right) parts where each part, also called Region of Interest (ROI), represents the discriminative information for clothing and carrying conditions from different viewing angles. An illustrative example is shown in Figures 3-4, 3-5 and 3-6. Furthermore, we exploit locally discriminating features that characterise these conditions by dividing the GEI horizontally in two (top, and bottom). For example, in chapter 4, the GEI from CASIA database is divided horizontally and/or vertically in three (top, medium, and bottom) and/or two equal (left and right) parts. In

chapters 5 and 6, the GEI of CASIA and USF gait databases are divided horizontally into two parts top and bottom where each side is known as the region of interest (ROI).

In this investigation, we focus on the dynamic area which can be defined between rows 161 to 240 in the bottom region and rows 1 to 30 in the top region in the case of the CASIA database. In the case of the USF database, rows between 1 and 27 in the top region and rows 88 to 123 in the bottom region define these two parts. For example, in the case of carrying conditions, the bag appears most often in the medium part of the horizontal division or the right part of the vertical split. In addition, in the case of clothing conditions, the clothes appear most often in the top part of the horizontal division or the right part of the vertical split.

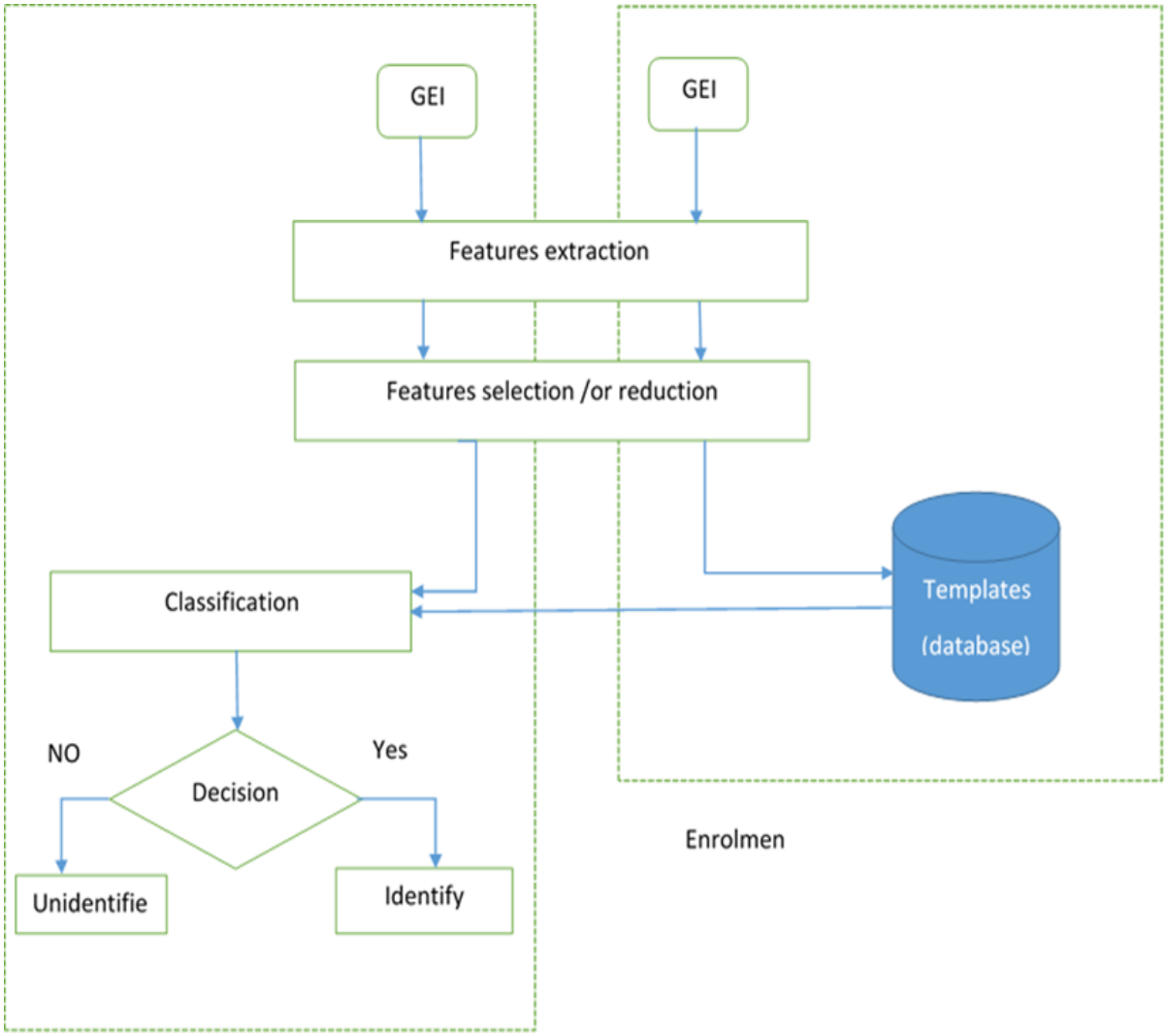


Figure 3-3. Human gait recognition system.

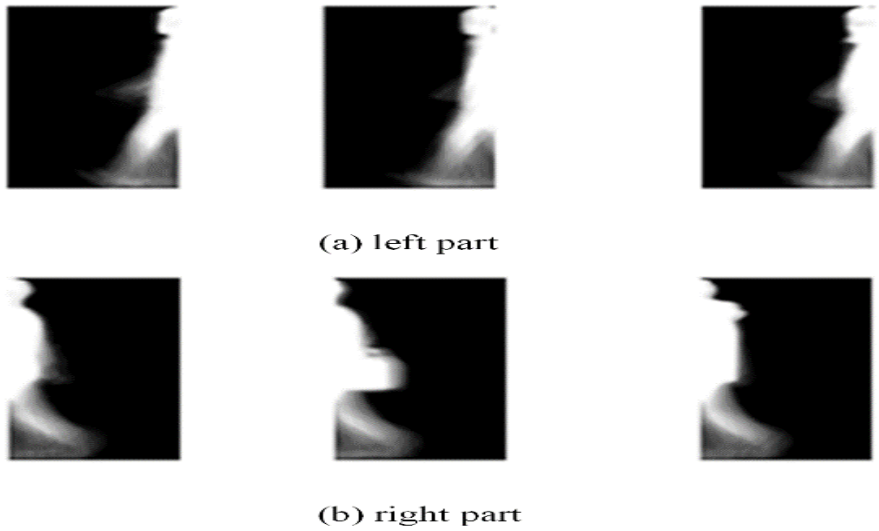


Figure 3-4. An example of ROIs extracted from a vertical division of the GEI of an individual from the side view 90° under three different covariates: Normal walking (1st column), Carrying a bag (2nd column) and Wearing a coat (3rd column), from the image of CASIA data.

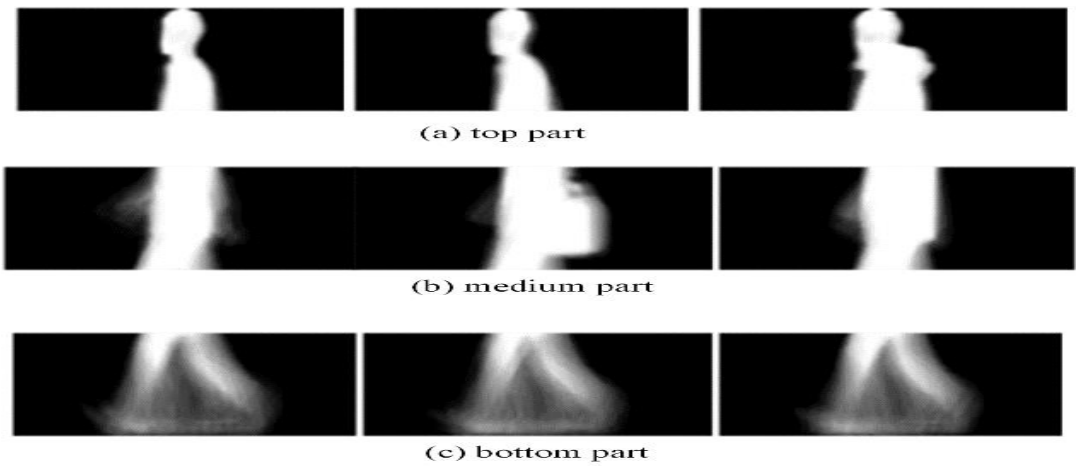


Figure 3-5. An example of ROIs extracted from a horizontal division of the GEI of an individual from the side view 90° under three different covariates: Normal walking (1st column), Carrying a bag (2nd column) and Wearing a coat (3rd column), from the image of CASIA data.

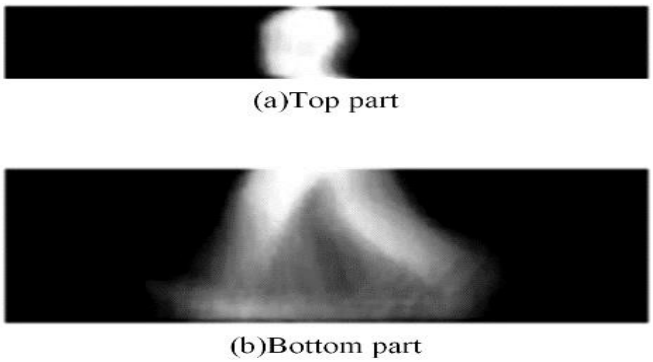


Figure 3-6. An example of ROIs extracted from a horizontal division of GEI of an individual from the side view 90° (Dynamic area), from the image of CASIA data.

3.4 Feature Extraction

The feature extraction consists of algorithms responsible for encoding the image content in a concise and descriptive way. Typical features include measures of colour (or intensity) distribution, texture and shape of the most relevant (previously segmented) objects within the image. These features are created into a feature vector that can then be used as a numerical descriptor of the image which needs to recognised (Marques, 2011). The texture extraction approaches can be divided into four methods as shown in the following subsection.

3.4.1 Model-based Methods:

These approaches describe the texture of the image to computer image texture model using a stochastic and/or fractal model, such as Markov Random Field Texture Models (Cross et al. 1983), Fractal-Based Description of Natural Scenes (Pentland, 1984), Unsupervised Texture Segmentation using Markov Random Fields (Manjunath et al., 1991) and Markov Random Fields as Models of Textured Biomedical Images (Strzelecki et al., 1997). However, a stochastic model usually leads to an increased computational complexity of the feature extraction phase. The advantage of the fractal model is that it is helpful for modelling some of the natural textures and can be applied for texture analysis identification, however, it lacks orientation selectivity and it may not be suitable for characterising local image structures (Materka et al., 1998).

3.4.2 Structural Approaches

Structural approaches aim to define the rules of grammar that can be used to represent the texture (Morse, 1998; Sevilla, 2006). This approach is useful for providing a perfect symbolic description of the image. Another advantage results from a synthesis compared to the analysis task (Materka et al., 1998).

However, it may be unsuitable when image data is noisy or of a low contrast (Olowoyeye et al., 2009). Additionally, structural approaches supported by psychological evidence which suggest that a structure based description and classification, which are related human perceptual and cognitive processes, have not yet been improved to their fullest effectiveness due to inherent complication associated with implementing structural pattern recognition systems (Olszewski, 2001).

3.4.3 Transform Methods.

These approaches represent an image in a transform domain in which the definition of the coordinate system is closely correlated to the characteristics of texture (Materka et al., 1998). For example, a Fourier Transform of the image can be computed before the transform coefficients are grouped to extract a feature vector of the image data (Nixon Mark, 2008). Also, it analyses the texture images by disintegrating the image into orientation and frequency components (Tuceryan et al., 1993). Several algorithms have been applied in the transform domain including Wavelet transforms (Materka et al., 1998), Gabor filtering and Fourier Transform (D. Zhang et al., 2000), (Anil K Jain et al., 1997) and (Daugman, 1985) where Gabor filters provide means for better spatial localisation.

3.4.4 Statistical Methods.

The statistical method for analysing the texture deploys the statistical properties of the intensity histogram (Morse, 1998), (Gonzalez et al., 2002). This statistical data is usually based on the second-order statistics and has achieved attractive rates of discrimination (Materka et al., 1998). In addition, other statistical approaches include autocorrelation of features (Tuceryan et al., 1993). The most popular second order statistical features for texture analysis are derived from the so-called co-occurrence matrix (Robert M Haralick, 1979). The methods are based on second-order statistics such as (D. Zhang et al., 2012), (Qurat-Ul-Ain et al., 2010) and

(Thangavel et al., 2005). The co-occurrence matrix method is based on the study of the statistics of pixel intensity distributions (Mirmehdi, 2008) by sampling the way certain grey-levels occur in relation to other grey levels (Morse, 1998). Nevertheless, this texture feature is hard to capture effectively with a large number of grey levels within a small region (Asheer Kasar Bachoo, 2005).

In this thesis, we present different techniques based on the use of features extracted from the GEI. The proposed feature extraction method in combined with RELIEF selection algorithm to select relevant and most discriminative Haralick texture features as will be described in Chapter 4. We also propose a second supervised feature extraction method based on Multi-scale descriptors (LBP, MLBP and a Gabor filter bank) using the SRKDA reduction algorithm described in Chapter 5. The last proposed feature extraction method based on Wavelet domain is described in Chapter 6. The proposed methods are able to extract and capture the relevant features from the GEI for human gait recognition under different conditions. In addition, we will focus on which parts of the body are better for recognition performance. These methods will be introduced in more detail in the next chapters.

3.5 Feature Selection and Reduction

The objective of this section is to give an overview of feature reduction and data selection in relation to feature extraction. In almost all pattern recognition approaches, one often goes for data reduction or subspace mapping, which is done primarily to reduce the dataset or the extracted feature vectors. The feature vectors extracted from the GEI often has a high dimensionality which may hamper the use of conventional classification algorithms. Consequently, the feature selection or reduction algorithms are important to extract only the helpful and informative features for classification.

3.5.1 Feature Selection

Feature selection is a broad subject in machine learning and many types of research have been carried out in this area. There exist a plethora of works on this topic, and, the works of Jennifer Dy (Dy et al., 2004) and Rohn Kohavi (Kohavi et al., 1997) are very useful. Feature selection aims to determine and select the most significant or discriminative features from a high dimensional space. It is one of the most frequently used and essential techniques in pattern recognition problems. In this case, a learning algorithm is confronted with the problem of selecting a significant subset of features while disregarding the remaining redundant set. To achieve this, a feature subset selection strategy ought to consider the calculation and the preparation of the feature data. We investigate the connection between ideal component subset selection and feature significance selection, which can apply in both supervised and unsupervised learning. Feature selection in unsupervised learning is a much harder problem, due to the absence of class labels. In supervised learning, feature selection aims to maximise classification accuracy (Kohavi et al., 1997). The feature selection approach for unsupervised learning aims to find a small subset of features that best detect the clusters from data according to the preferred criterion (Dy et al., 2004). Feature selection algorithms can be classified into three basic approaches (Dalal et al., 2005). The first is called the wrapper approach, in which the selection of features is wrapped within a learning algorithm. The second approach is referred to as the filter approach where the features are selected according to intrinsic data values such as information, dependency or consistency measures. The RELIEF technique is an established case of the multivariate filter. Most multivariate techniques rank subsets of features as opposed to individual features. The last approach is called embedded, and in this method, the feature selection procedure; described as embedded technique, searches for an ideal subset of features that are incorporated into the classifier construction. It can be seen as a search in the combined space of feature subsets and hypotheses. Much the same as wrapper approaches,

embedded techniques are in this way impossible to miss to a given learning algorithm. Saeys et al. have defined an advantage and disadvantage for each class of feature selection in (Saeys et al., 2007) such as the following:

3.5.1.1 Filter Approach

The advantages of filter methods are that they effectively scaled to high-dimensional data, they are computationally easy and fast, and, are autonomous of the classification algorithm (C. Liu et al., 2017). Hence, the feature selection can be performed just once so that a distinctive classifier can be used. On the other hand, there are limitations when deploying a filter method where the technique does not consider the interaction with the classifier (the search in the feature subset space is separated from the search in the hypothesis space) and most proposed techniques are univariate. This means that each feature is independent, thereby ignoring feature dependencies which may lead to a decrease in classification performance when compared to different types of feature selection methods. To overcome the issue of ignoring feature dependencies some multivariate filter techniques have been presented, pointing to the incorporation of feature dependencies to some degree.

3.5.1.2 Wrapper Approach

Although filter techniques treat the problem of finding a good feature subset autonomously of the model selection step, Wrapper methods insert the model hypothesis search within the feature subset search. In this setup, a search methodology in the space of possible feature subsets is characterised, and different subsets of features are created and evaluated. The evaluation of a particular subset of features is acquired via the training and testing of a particular classification model rendering this methodology custom fitted to a particular classification algorithm (C. Liu et al., 2017). To search the space of all feature subsets, a search algorithm is then ‘wrapped’ around the arrangement model. However, as the space of feature

subsets becomes significantly large with the number of features, search heuristic techniques are used to control the search for an ideal subset. These search techniques can be partitioned into two classes; deterministic and randomised search algorithms. The advantages of wrapper methodologies incorporate the association between feature subset search and model selection and the capacity to check feature conditions. A common disadvantage of these methods is that they are very computationally intensive, particularly if building the classifier has a high computational cost and are at higher risk of over-fitting than filter methods.

3.5.1.3 Embedded Approach

This method, which has only recently been suggested, aims to combine the advantages of both previous methods and is referred to as an embedded method (C. Liu et al., 2017; Mistry et al., 2017). The technique ascertains the features which contribute to the model best whilst the model is being created. Embedded techniques have the advantage of incorporating the interaction with the classification model. They also have the disadvantage of being more computationally intensive than wrapper approaches.

3.5.2 Feature Reduction

Feature reduction is also a common topic in machine learning and is a field of research in itself. Several works have been done in this field and, here, we try to make an overview based on the work of (Blum et al., 1997) and (Kohavi et al., 1997).

Feature reduction is a procedure used to decrease the dimensionality of the feature by analysing data and the relationship between arrangements of connected variables. Dimensionality reduction is essential for recognition purposes because the size of the data can be substantial and computationally costly. Starting in 1997, when a few papers on feature selection were proposed (Blum et al., 1997; Kohavi et al., 1997), the topic evolved significantly through the introduction of several techniques to efficiently reduce dimensionality of the data by examining

the relationship between a set of correlated variables. For example, Principal Component Analysis (PCA) (Yu et al., 2017; Zhao et al., 2016), Kernel Discriminate Analysis (KDA), LDA (Munif Alotaibi, 2017), MMP and Locality Preserving Projections (LPP) (Zhao et al., 2016) have been used as dimensionality reduction methods by the research community in pattern recognition problems. Feature reduction techniques can be classified into two classes as follows:

3.5.2.1 Geometry-based Methods

The objective of this approach is to entrench the data into some low-dimensional space such that the inherent geometry contained in the dataset is conserved. Representative methods include PCA (Jolliffe, 2002) which is a typical technique exploiting global data structure to recognise a subspace where the sample variance is maximised. While PCA uses the global data features in the Euclidean space, the local data manifold structure is disregarded.

3.5.2.2 Discrimination-based Methods

The goal of this approach is to extract a discriminative subspace from the dataset in which the data from various classes can be better discriminated. Representative methods include Margin Maximising Discriminant analysis, abbreviated as MMD, (Kocsor et al., 2004), (Tsang et al., 2008) and MMP (F. Wang et al., 2011). These methods will be addressed in the following chapters. More recently, numerous dimensionality reduction techniques have been proposed and these can be classified into two methods.

3.6 Classification

Classification is an important task in machine learning and is a process that allows decision making with regards to recognition patterns through the use of a classification function for the

selected feature set. In this section, we will introduce some classification methods and focus on the basics of classifier techniques.

3.6.1 Principles of Classification

A feature vector generated by a feature extraction process is used by a classifier to carry out the recognition from the gait captured. Various classification approaches have been proposed in the literature such as supervised and unsupervised classification methods. In supervised learning, the output datasets are used to train the machine learning algorithm in order to make a decision. In the case of unsupervised learning, datasets are not provided, but rather, the data is grouped into various classes automatically.

Supervised learning: a machine learning process of deriving a function from training data, which includes both the input and the desired outcomes. For example, the correct outcomes (targets) are known and are given as input to the model during the learning process. These methods are usually fast and accurate (Pandey et al., 2016).

Unsupervised learning: there is not any prior knowledge of the output for the application at hand and the classification algorithm aims to differentiate correctly between the different gaits of the subjects (clustering of data).

It is worth noting that feature selection in the unsupervised learning process is a more difficult problem when compared to supervised learning, where the feature selection aims to maximise classification accuracy (Cristianini et al., 2000). This is due to the absence of class labels. In this work, some classifications will be investigated e.g. Support Vector Machines (SVM), K-nearest neighbours (*k*-NN) and Decision trees etc. In this chapter, the focus is on *k*-NN and SVM techniques.

3.6.1.1 K-Nearest Neighbour

A K-NN classifier is a widely used classifier in pattern recognition applications and can be useful in both classification and regression predictive problems. The following example illustrates this type of classifier.

We store all training samples Gallery (G) and give a new sample Testing (T) to be classified, search for the training sample (x_i, y_i) where x_i is most similar, or closest in distance, to T and predict y_i (see Figure3-7). The following equations explain how we measure the similarity or distance between two samples. The most commonly used measure is the Euclidean distance $D(T, x_i)$ (Pandey et al., 2016)

$$D(T, x_i) = \|T - x_i\| = \sqrt{(T - x_i)(T - x_i)} = \sqrt{\sum_j (x_j - x_{ij})^2} \tag{3.1}$$

Where j is the number of variables, K is a parameter used to increase value when we need to make the boundary of classes' smoother (see Figure 3-8).

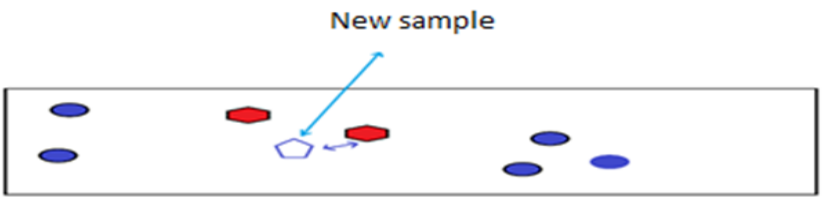


Figure 3-7. Example of a K-NN classifier.

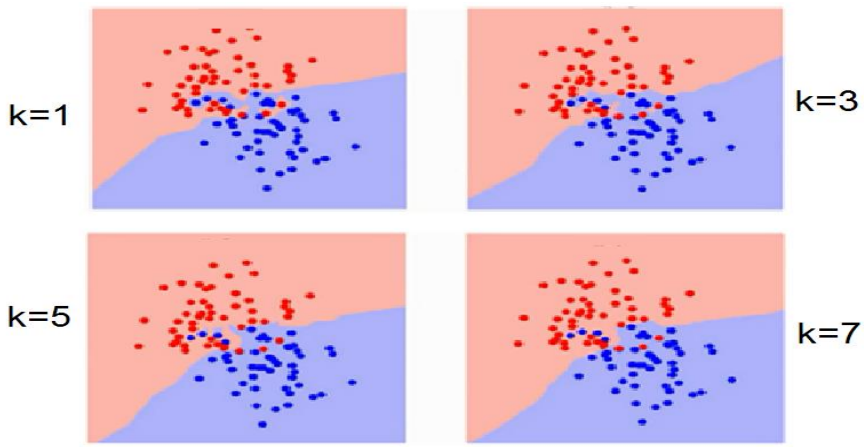


Figure 3-8. Comparison of ranks describes how the boundary of the classes has changed.

3.6.1.2 Support Vector Machine (SVM)

The SVM approach has proven to be a powerful and popular tool for pattern classification. SVM can be used to analyse data and identify patterns, in order to classify the data into two classes in the case of a binary classification. We can extend the two-class SVM further into two main categories; Hard-Margin SVM and Soft-Margin SVM, depending on the learning used.

Hard-Margin SVM uses *N* training data where the data is linearly separable. Raining a Hard-Margin SVM classifier using the *N* training data consists of determining the best (optimal) hyperplane which separates the training data in the input space and having the maximum distance to its neighbouring data points of both classes (Kecman et al., 2006) (see Figure 3-9).

This can be measured using Hard-Margin SVM by Equation (3.2), given by (Gunn, 1998):

$$max_w \frac{2}{\|w\|} \quad subject\ to\ w^T x_i + b \begin{cases} \geq \ \ 1\ if\ y_i = +1 \\ \leq -1\ if\ y_i = -1 \end{cases} \quad \text{for } i = 1 \ldots N \tag{3.2}$$

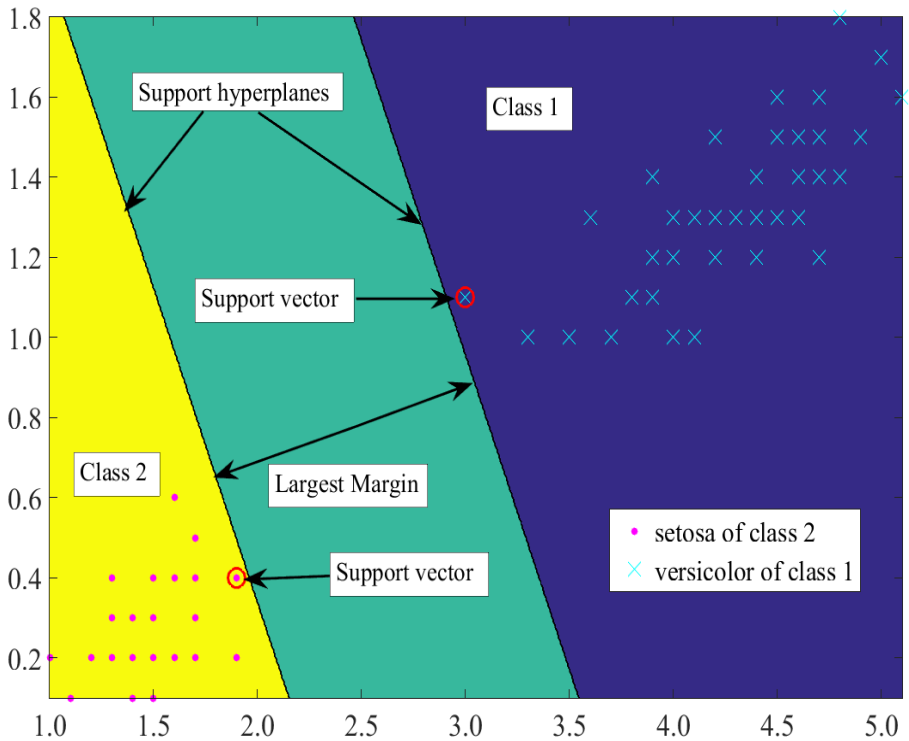


Figure 3-9. Hard-margin SVM.

Where x_i is an m-dimensional column vector and y_i is the class label associated with x_i ($y_i \in \{1, -1\}$). This is called an optimal hyperplane. The margin is defined as the distance between the closest points, also known as support vectors, and the separating hyperplane, (Gunn, 1998), (Abe, 2005) and (Kecman et al., 2006).

Soft-Margin SVM is depicted in Figure 3-10 (Kecman et al., 2006). This classifier type is used if the training data is not linearly separable as there will be no hyperplane linearly separating the data. This leads us to define the problem as given in Equation (3.3), known as the primal representation of a Quadratic Programming (QP) optimisation problem, (Gunn, 1998).

$$min_{w,b}\Phi(w) = \frac{1}{2}w^T w\,, subject\ to\ y_i(w^T + v_i + b) \geq 1, \forall i \tag{3.3}$$

To solve this problem, a soft-margin SVM is proposed (Gunn, 1998) and (Abe, 2005), where the training data points are allowed to violate the hard constraints in the Equation above (3.3). Mathematically, a new set of non-negative variables $\{\xi_i\} = 1, \ldots N$ are introduced. They are called the slack variables and measure the amount of violation of the hard constraints of the Equation (3.3). Essentially, during the training of a soft-margin SVM classifier, one must solve the following problem (Gunn, 1998).

$$min_{w,b,\xi}\Phi(w,\xi) = w^T w + \mu \sum_i^N \xi_i \tag{3.4}$$

Subject to

$$y_i(w^T x_i + b) \geq 1 - \xi_i \ \text{ for } \forall\ \text{i} = 1\ldots\text{N}$$

where $\xi = (\xi_{1,} \ldots\ldots, \xi_{N,})$ and μ a are parameters that describe the cost constraints violation and must be chosen beforehand. Parameter $\boldsymbol{\mu}$ defines the trade-off between a large margin, i.e. the minimisation of the term $\frac{1}{2}W^T W\ in\ Equation$ 3.4, and the minimisation of classification error, i.e. minimisation of term $\sum \xi_i$ shown in Equation 3.4. Also, there is the extension to a 2-

class classification called multi-class classification (Deng et al., 2006). In this case, the original multi-class problem is divided into multiple two-class sub tasks which are solved using several two-class SVMs. This method is applied with regard to linear data.

However, if the data is non-linear, it is preferable to employ a kernel function. Moreover, the extension to more complex, nonlinear decision functions is relatively straightforward and is carried out by mapping the input variables into a new feature space and by working with linear classification in the new space. More specifically, if we have data $x, z \epsilon X$ and a map $\phi: X \longrightarrow R^d$, (Gunn, 1998) and (Wahba, 1990) then;

$$K(x, z) = \langle \phi(x), \phi(x) \rangle \tag{3.5}$$

is a kernel function where R^d is feature space. The Kernel function K in Equation 3.5 is defined by (Aronszajn, 1950) and (Wahba, 1990).

Finally, an SVM is used for classification, and it constructs in a high dimensional space a hyperplane or set of hyperplanes. A hyperplane that has the biggest distance to the closest

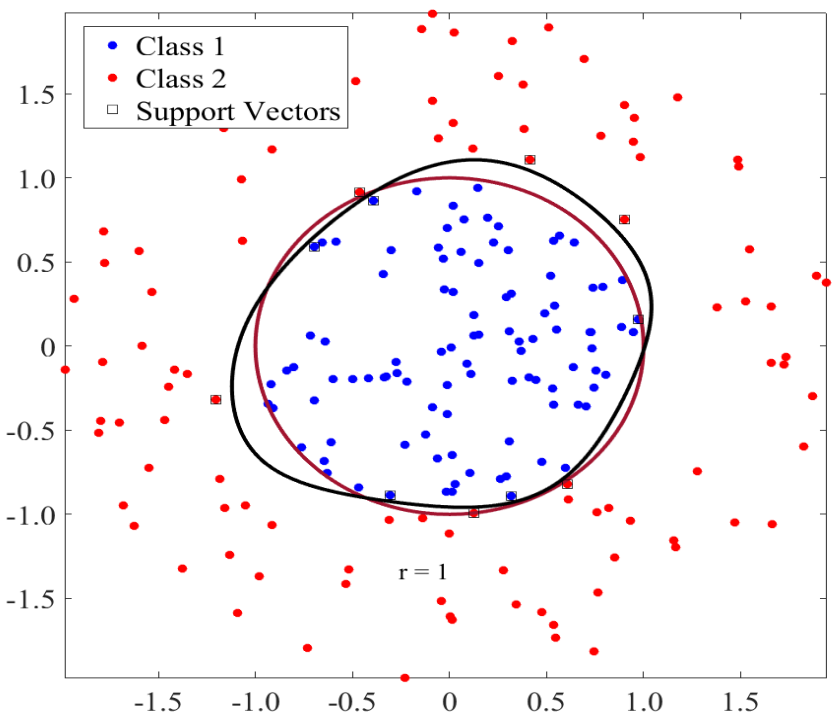


Figure 3-10. Soft-margin SVM.

training data point of any class has achieved good separation, in general, and the larger the margin, the lower the generalisation error of the classifier.

3.6.2 Performance Criteria

A confusion matrix demonstrates the quantity of right and wrong predictions made by the classification algorithm in order to contrast the genuine results (or target value) in the information. A lattice is N×N matrix where N is the number of target qualities (classes). The execution of such models is ordinarily assessed using the information as a part of the framework. Table 3-1 shows a 2×2 disarray grid for two classes (Positive and Negative). Across the top is the observed class labels and down the side are the predicted class labels.

Table 3-1. Truth Table Confusion Matrix

Classes	Positive class	Negative class
Positive class	True positive	False positive
Negative class	False negative	True Negative

Each cell contains the number of predictions made by the classifier that falls into that cell. R. Strickland in (Strickland, 2002) has described a decision made into one of four possible categories as follows:

- TP (true positive) a detection that corresponds to an actual abnormality.
- FP (false positive) error occurs when detection corresponds to a normal region.
- TN (true negative) the decision simply means a normal region was correctly labelled as being normal.
- FN (false negative) the error implies that a true abnormality was not detected.

Accuracy indicates the percentage of correct predictions in all instances.

$$\mathrm{ACC} = \frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})} \times 100 \tag{3.7}$$

3.6.3 Evaluation Cross-Validation (CV)

Cross-validation or (sometimes-called rotation estimation) is a statistical analysis process used for comparing and assessing the performance of classification algorithms by partitioning data into two portions. One is used to train or teach a model whose class labels are known and the other is used to validate the model so that the classifier can accurately predict unknown datasets i.e. testing data (Kohavi, 1995).

Estimating the accuracy of a classifier produced by supervised learning algorithms is significant for assessing its future prediction accuracy, however, it is also key in selecting a classifier from combining classifiers or a given set or model selection (Wolpert, 1992). In a typical cross-validation process, the validation and the training sets must crossover in sequential rounds so that every data point has a chance of being validated. This allows a cross-validation to be determined as a prediction error or "error rate". A low error rate is indicative of a good model meaning that of course. The cross-validation can be split into four types: Hold Out method, K-Fold cross-validation, Leave-One-Out cross-validation (LOOCV) and the Bootstrap method. Srivastava introduces the cross-validation types as described below (Srivastava, 2013).

3.6.3.1 Holdout Method

This type of cross-validation is simple. In this method, the dataset is split into two groups, and each group is designated up to 50-70 % of the data as the training sample with the remaining 50-30 % as the test sample. These groups are as follows:

- Training sample: the data is used to train the classifier.

- Test sample: the data is used to estimate the error rate of the trained classifier.

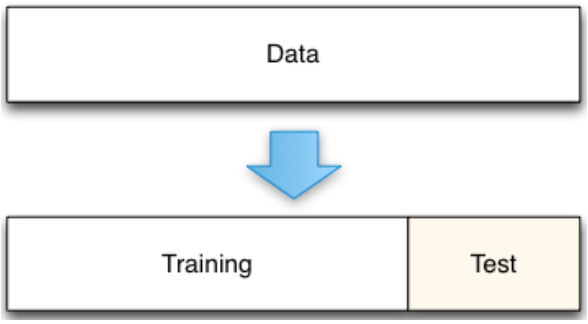


Figure 3-11. Hold out data spilled

Figure 3-11 illustrates the holdout method and how the data is split. The Holdout method has two main disadvantages and these are as follow:

- In problems where there are sparse datasets, it may not be feasible to spare a portion of the dataset for testing.

- With regards to the one train-and-test experiment, the Hold Out estimate of error rate will be misleading if there happens to be an "unfortunate" split.

Having said this, the advantage of this method is that it is usually preferable to the residual method and does not take any longer to compute.

3.6.3.2 K-Fold Cross-Validation

In this method, the dataset is randomly split into K mutually exclusive subsets (the folds) of approximately equal size in order to create a K-fold partition of the dataset. Here, each of the K experiments used K-1 folds for training and the remaining for testing. Figure 3-12 shows an example of a K-fold cross-validation method. The advantage of K-Fold cross-validation is that all the examples in the dataset are eventually used for both testing and training. The true error E is estimated as the average error rate E_i (Anguita et al., 2005):

$$E = \frac{1}{K}\sum_{i=1}^{K} E_i \tag{3.8}$$

On the other hand, the disadvantage of this method is that the training algorithm has to be rerun from scratch K times, meaning it will take K times as much computation to make an evaluation.

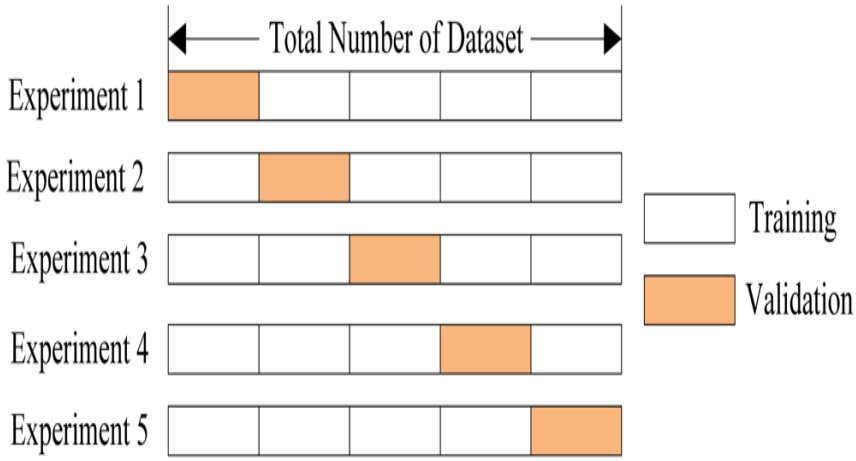


Figure 3-12. Schematic view of a K-fold cross-validation method.

3.6.3.3 Leave-One-Out Cross-Validation (LOOCV)

This method is the degenerative case of K-Fold cross-validation, where K is chosen as the total number of examples. For a dataset with N examples, where N is a number of experiments to be performed, each experiment uses N-1 examples for training and the remaining ones for testing. The true error in this method is estimated as the average error rate on test examples. Figure 3-13 shows the LOOCV.

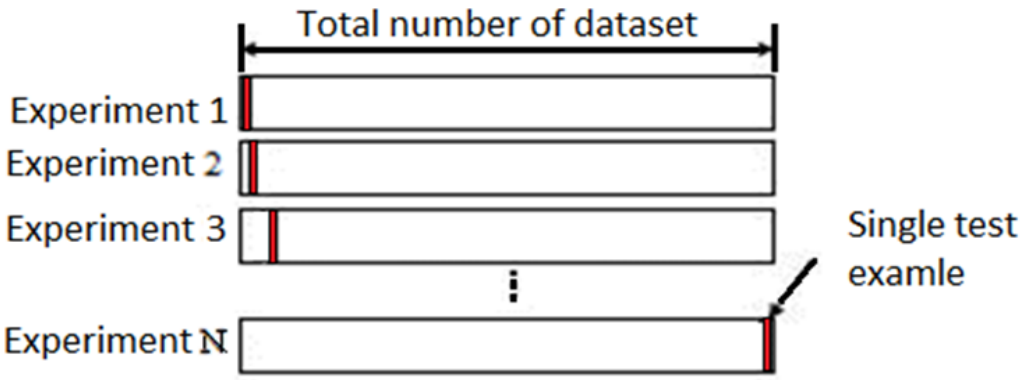


Figure 3-13. Schematic view of the LOOCV method.

This method has two advantages; firstly, it makes the best utilisation of the data for training, leading to an increase in the chance of building more classifiers that are accurate. Secondly, it

does not involve any random subsampling. On the other hand, it has some disadvantages including a high computational cost.

3.6.3.4 Bootstrap Methods

The last type of cross-validation, Bootstrap, is a method which randomly draws a dataset from the training sample. Each sample is the same size as the training sample. This means it can be seen as a selection with replacement whereby the data point can be selected more than once.

3.7 Summary

This chapter has discussed human gait recognition approaches under covariate distortions e.g. carrying and clothing conditions. Also, the chapter introduces a comprehensive description of widely used techniques related to methods proposed in this research. The next chapter will describe an investigation of Haralick features using RELIEF algorithm to generate more discriminative features extracted by GEI including a validation and evaluation approach using two well-known datasets.

CHAPTER FOUR: GAIT RECOGNITION BASED ON HARALICK FEATURES

4.1 Introduction

This chapter discusses a supervised feature extraction approach that is capable of selecting distinctive features for the recognition of a person under clothing and carrying conditions. The principle of the suggested approach is based on the Haralick features extracted from Gait Energy Images.

First, the proposed method considers Haralick features which are extracted locally by horizontally dividing the GEI into three ROIs. The proposed method is evaluated using CASIA Gait Database under variations of clothing and carrying conditions for different viewing angles. The experimental results using SVM classifier have provided attractive results of up to 83.00% in terms of highest Identification Rate (IR), (A. O. Lishani et al., 2014).

Secondly, the proposed method is further extended to include Haralick features with the RELIEF feature selection algorithm. The RELIEF algorithm is used in order to select the most relevant features only with a minimum redundancy. Again, the proposed extended method is evaluated using the CASIA and USF gait databases under variations of clothing and carrying conditions for different viewing angles. The experimental results using the *k*-NN classifier yielded striking results of up to 80% in terms of the highest IR at rank-1 (Ait O. Lishani et al., 2017).

As discussed previously, gait recognition refers to verifying and/or identifying a person by his/her walking style under covariate factors. The main idea behind such a technology is to

determine the discriminating features that characterise the walking styles using various viewing angles where each view represents the discriminative information for clothing and carrying conditions considered in our research.

4.2 The Proposed Method

The Haralick paper, published in 1973, has been cited thousands of times and Haralick texture analysis has become one of the most common and efficient methods for capturing and extracting texture features. The method is very useful and very powerful in texture analysis and is comprehensive for most texture features. Haralick can be calculated from GLCM, which is one of the best-known tools for texture analysis, to estimate image properties related to second-order statistics.

In this proposed method, we visually analyse the information contained in GEI and define a feature extraction method for gait recognition under varying conditions relating to clothing and carrying. The main idea is to exploit the locally discriminating features that characterise these conditions by horizontally and/or vertically dividing the GEI into three (top, medium and bottom) and/or two (left and right) equal parts, whereby, each part (also called ROI) represents the discriminative information for clothing and carrying conditions under different viewing angles considered in our study. For example, in the case of a carrying a bag, the bag appears most often in the medium part of the horizontal division or the right part of the vertical division. Additionally, in the case of clothing conditions, the clothes appear most often in the top part of the horizontal division or the right part of the vertical division.

The original size of the GEI is 240×240 in the CASIA database while the original size of the GEI is 88×128 in the USF gait databases. In chapter 3, an illustrative example showing the proposed method has been shown in figures 3-4 and 3-5. Figure 4-1, illustrates the diagram of

the proposed feature extraction and selection method based on Haralick texture features with RELIEF selection techniques.

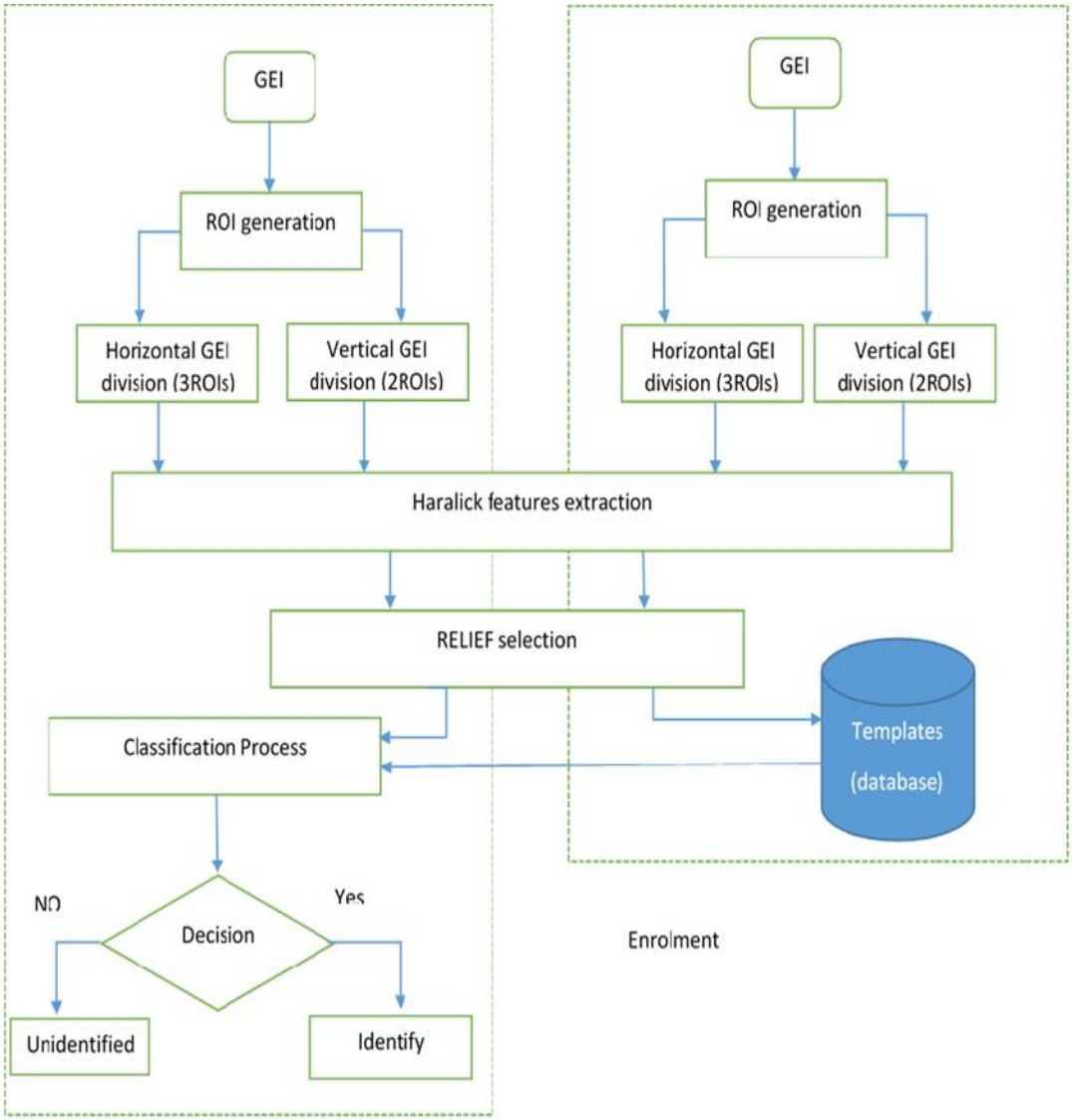


Figure 4-1. Diagram of the proposed feature extraction and selection method based on GEI Haralick texture features with RELIEF selection algorithm.

4.2.1 Haralick Texture Feature Extraction

The discriminative features proposed in our feature extraction method include the Haralick texture features (R. M. Haralick et al., 1973) extracted and computed from the GEI. To the best of our knowledge, no one has attempted to implement this method before. For each GEI, a feature vector is formed by converting the generated grey-level co-occurrence matrix (GLCM) to a vector for use later in the classification stage. Furthermore, Haralick features containing

14 statistical features can be extracted from the GLCM to form a new feature vector with 14 features.

4.2.1.1 Grey-level Co-occurrence Matrix

A useful technique for characterising an image texture is to consider statistical moments of intensity histogram of an image (Rafael C. Gonzalez, 2008). Using histograms of the GLCM will enable a quantification of texture, conveying information about the distribution of intensities.

In using a statistical approach, for example, the co-occurrence matrix gives important information about the relative position of the neighbouring pixels in an image. These features are calculated from the GLCM of GEI, denoted P, with dimension $N_g \times N_g$ where N_g is the number of grey levels in the GEI. The co-occurrence matrix $P_{d,\theta}$ can be defined as (R. M. Haralick et al., 1973):

$$P_{d,\theta}(i,j) = \sum_{x=1}^{N_g} \sum_{y=1}^{N_g} \begin{cases} 1, & if\ G\ (x,y) = i \\ and\ G(x+dx, y+dy) = J & \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

where d and θ are the offset and direction (phase) respectively. Selecting an offset vector, such that the rotation of the image is not equal to 180 degrees, will result in a different GLCM for the same rotated image. This can be avoided by forming the co-occurrence matrix using a set of offsets sweeping through 180 degrees at the same distance parameter (d_x, d_y) to achieve a degree of rotational invariance, for example:

$\theta = 0^{\circ}$: P horizontal, $\theta = 45^{\circ}$: P right diagonal, $\theta = 90^{o}$: P vertical and $\theta = 135^{o}$: P left diagonal. $dx\ \ and\ dy$ denote the distance between the pixel of interest and its neighbour along the x-axis and the y-axis of an image respectively. Haralick texture features are statistical entities defined to emphasise certain texture properties calculated from P.

Table 4.1 describes the proposed Haralick features allowing a description of the textures in the GEI in order to recognise the observed human gait. These features comprise of 14 statistics calculated from GLCM. However, *F14* (Maximal Correlation Coefficient) is not used in this study as it can cause computational instabilities if the co-occurrence matrix has ill-conditioned statistical formulations (Rafael C. Gonzalez, 2008). In this study, only the $\{F_1, F_2 \ldots F_{13}\}$ features are considered.

4.2.2 RELIEF Based Feature Selection

RELIEF is a feature selection algorithm (Kira et al., 1992) which can be used to select only the most discriminative gait features extracted using the Haralick method. Algorithm 4.1 summarises the proposed method. RELIEF, proposed by (Kira et al., 1992), is used in the data processing stage as a feature selection method. RELIEF-based algorithms can be divided into three principal parts:

1. Compute the nearest miss M and nearest hit H.
2. Compute the weight of a feature by using Eq. 4.2.
3. Return a ranked list of features or the top k-features according to a given threshold.

RELIEF is a feature weight-based algorithm inspired by instance-based learning (Kira et al., 1992). Given training data *R*, sample size m and a threshold τ; RELIEF detects those features which are statistically relevant to the target concept where τ encodes a relevance threshold (0, $\leq \tau \leq 1$). The algorithm begins by initialising the weight vector and tuning the weight for every feature to 0. Then it randomly picks a learning sample X and computes the H and M from the same subfamily H and one from the opposite subfamily M.

The weight *W* can be calculated using Eq. 4.2:

Table 4-1. Describes how Haralick texture features are calculated.

Feature	Formula
Angular second moment	$F_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j)^2$
Contrast	$F_2 = \sum_{r=0}^{N_g-1} r^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1_{\lvert i-j \rvert = r}}^{N_g} P(i,j) \right\}$
Correlation	$F_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij) P(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$ Where μ_x, μ_γ, σ_x, σ_γ are the means and standard deviations as follows: $\mu_x = \sum_{\iota=1}^{N_g} \iota p_x(\iota)$, $\mu_y = \sum_{\iota=1}^{N_g} \iota p_y(\iota)$, $\sigma_x = \sqrt{\sum_{\iota=1}^{N_g} (\iota - \mu_x)^2 p_x(\iota)}$ and $\sigma_y = \sqrt{\sum_{\iota=1}^{N_g} (\iota - \mu_y)^2 p_y(\iota)}$ Where p_x and p_y are the partial PDFs dened by $p_x = \sum_{j=1}^{N_g} P(x,y)$ & $p_y = \sum_{i=1}^{N_g} P(x,y)$, respectively.
Variance	$F_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-\mu)^2 P(i,j) \quad where \quad \mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} iP(i,j)$
Inverse difference moment	$F_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{P(i,j)}{1+(i-1)^2}$
Sum average	$F_6 = \sum_{r=0}^{2N_g-2} rP_{x+y}(r)$ where x and y are the coordinates (row and column) of an entry in the co-occurrence matrix, and $P_{x+y}(r)$ is the probability of co-occurrence matrix coordinates summing to x + y dened as follows: $P_{x+y^{(r)}} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j)$ where $r = i + j$ with $r = 2,3,\ldots,2N_g - 2$
Sum variance	$F_7 = \sum_{r=0}^{2N_g-2} (r - F_6)^2 P_{x+y}(r)$
Sum entropy	$F_8 = -\sum_{r=o}^{2N_g-2} P_{x+y^{(r)}} \log(P_{x+y}(r))$
Entropy	$F_9 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) \log(P(i,j))$
Difference variance	$F_{10} = \sum_{r=0}^{N_g-1} \left(r - \sum_{\iota=0}^{N_g-1} \iota P_{\lvert x-y \rvert}(\iota) \right)^2 P_{\lvert x-y \rvert}(r)$ Where $P_{\lvert x-y \rvert} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j)$ and $r = \lvert i-j \rvert$ with $r = o, 1, \ldots, N_{g^{-2}}$
Difference entropy	$F_{11} = \sum_{r=0}^{N_g-1} P_{\lvert x-y \rvert^{(r)}} \log(P_{\lvert x-y \rvert}(r))$
Information measure 1	$F_{12} = \frac{F_9 - H_{xy^1}}{\max\{H_x, H_y\}}$ where H_x and H_y are entropies of p_x and p_y, respectively; and $H_{xy^1 = -} \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j) log(p_x(i) p_y(j))$
Information measure 2	$F_{13} = \sqrt{1 - \exp(-2(H_{xy^2} - F_9))}$ where $H_{xy^2} = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_x(i) P_y(j) \log(P_x(i) P_y(j))$
Maximal correlation coefficient	$F_{14} = \sqrt{2^{nd} largest\ eigenvalue\ of\ Q}$ where $Q(i,j) = \sum_r \frac{P(i,r)P(j,r)}{p_x(i) p_y(k)}$

Algorithm 4-1. Haralick textures descriptors with RELIEF selection algorithm.

Feature extraction and selection method for gait recognition based on GEI Haralick texture descriptors with RELIEF selection algorithm
Input: Silhouette images extracted over one gait cycle: $S(x, y, t)$; t = 1, 2, ..., N To compute a GEI using Eq.2.1: $G(x, y)$ Switch (GEI division type) Case Horizontal: Divide GEI horizontally into 3 equal parts: $G_{(H1)}(x, y)$, $G_{(H2)}(x, y)$ and $G_{(H3)}(x, y)$ For each $G_{(Hi)}$, i = 1, ...,3 Compute Haralick features defined in Table 4.1: $F_{(Hi)}$ Generate feature extraction set: $F_{(H)} = \{F_{(H1)}, F_{(H2)}, F_{(H3)}\}$ Case Vertical: Divide a GEI vertically into 2 equal parts: $G_{(V1)}(x, y)$, and $G_{(V2)}(x, y)$ For each $G_{(Vi)}$, i = 1, 2 Compute Haralick features defined in Table 4.1: $F_{(Vi)}$ Generate feature extraction set: $F_{(V)} = \{F_{(V1)}, F_{(V2)}\}$ End Switch Apply RELIEF selection algorithm on $F_{(H)}$ or $F_{(V)}$ Output: Relevant features set F

Algorithm 4-2. Pseudo-code of the RELIEF algorithm.

Pseudo-code of the RELIEF algorithm

Input: S learning instances X described by N features; K iterations
Initialise: $\forall i$, $W[i] = 0$
for k = 1 to K do
Randomly select an instance X
Find nearest hit H_X and nearest miss M_X of X
for i = 1 to N do
Compute weight W[i] using Eq.4.2
end for
end for
return W
Output: W Features ranking (for each feature F_i a quality weight within $-1 \leq W[i] \leq 1$)

$$[W_i] = [W_i] + \frac{diff\left(x^i, NM_{x^i}\right)}{(S \times K)} - \frac{diff\left(x^i, NH_{x^i}\right)}{(S \times K)} \tag{4.2}$$

where S is the number of learning instances X described by N features, and K is a number of iterations. The function *diff* is the difference between feature values between 2 cases defined as follows:

$$diff(a,b) = \frac{a-b}{u} \tag{4.3}$$

where u is a normalisation unit to normalise the values of *diff* into the interval [0;1]. Algorithm 4.2 summarises the pseudo-code of the RELIEF algorithm used in Algorithm 4.1.

4.3 Experimental Results and Discussion

To validate and evaluate the performance of our proposed methods CASIA and USF datasets were used. The first and second experiments were carried out using CASIA database with different chosen subset sequences in the gallery and the test while the third experiment used USF database. In the following section, an analysis is carried out based on the results obtained, including a comparative study of some existing and similar state-of-the-art methods.

4.3.1 Experiment 1 using CASIA Database

4.3.1.1 Database and Evaluation Criteria

We have evaluated the proposed method using on CASIA gait dataset B which is a multi-view gait database (Shiqi et al., 2006). This database was constructed from 124 subjects (93 men and 31 women) and 11 cameras around the left-hand side of the subject when they were walking. Thus, the data was captured from 11 different angles starting from 0° to 180° (i.e. the angle between two nearest view directions would be 18° in the range of [0°, 180°]). Each subject has two carrying-bag sequences (Set-A), two wearing-coat sequences (Set-B) and six normal walking sequences (Set-C).

In first experiments, we have selected from this database the first sequence from Set-A, SetB, and Set-C to evaluate the performance of the proposed method under the following three conditions: normal, carrying bag and wearing a coat. These experiments are carried out only under viewing angle 90°. The selected data were split randomly into two parts. 50% of the data

was used for training and the remaining 50% was used for testing the effect of the above three conditions.

To examine the efficiency of the proposed approach, we used LOOCV with the SVM classifier. LOOCV was adopted in order to find an optimal model for predicting and estimating the performance. According to Marcos (Marcos 2017) and Z-Y. HE et. al (Z.-Y. He et al., 2008), LOOCV has a higher variance than K-fold cross-validation. This is because LOOCV is a special case of K-fold cross-validation where the number of folds is the same as the number of observations, in other words, K = N. There is one fold per observation, and therefore, each observation by itself gets to play the role of the validation set with the other n-1 observations playing the role of the training set.

As in previous experiences (e.g. as in (Dobrovidov et al., 2013)), the optimal kernels were obtained for the Gaussian kernel for gamma = 0.25, and so, the one-against-one SVM classifier using the radial basis function kernel (with γ = 0:25) was used with the Correct Classification Rate (CCR) parameter, defined in Eq.3.7, in order to evaluate the classification performance.

4.3.1.2 Results and Analysis

The proposed technique was assessed based on the different covariates of the Haralick features with an SVM classifier on CASIA database-B in order to assess their performance to correctly classify the different covariates (A. O. Lishani et al., 2014). Table 4-2 shows the results obtained using the selected data split randomly into two parts; 50% for training and 50% for testing.

By analysing these results, it can be noticed that the computation of Haralick texture feature locally on each ROI selected in GEI significantly improves the performance of the recognition system by up to 8.40% compared to a global counterpart. Also, we have noticed that, in the

case of "normal walking" and "carrying-bag" conditions, the results indicate an increase of up to 11.00% improvement in recognition when compared to a global GEI based approach.

This leads to the conclusion that the computation of local Haralick texture features on each selected ROI in GEI significantly improves the performance of the gait recognition system by up to 9% compared to the global computation technique.

Table 4-2. Comparison of CCR (in %) from the proposed method based on local and global feature computation techniques on CASIA database using the 90° view.

Features computation technique	Covariates			CCR (%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Global	78.50	65.10	80.10	74.60
Local (proposed method)	88.70	76.90	83.30	83.00

4.3.2 Experiment 2 using CASIA Database

4.3.2.1 Database and Evaluation Criteria

In this experiment, we have evaluated the proposed method using CASIA gait database B. However, we selected the three first sequences from SetC, the first sequence from SetA and used SetB as the probe. The remaining sequences for all the 124 subjects were assigned to the training set. Experiments are carried out under viewing angles of 36°, 72°, 90° and 108° under the following three conditions; normal, carrying bag and wearing a coat. For the evaluation criteria, a *k*-NN classifier was used to evaluate the classification performance. The highest Identification Rate (IR) at rank-1, which is defined as the percentage of samples with a correct match in the first place of the ranked list, is used to evaluate the classification performance.

4.3.2.2 Results and Analysis

Table 4-3 shows the results of a comparative study of the proposed method against different state-of-the-art methods on CASIA database B for a side view of 90°. Three covariates were considered - normal walking, carrying a bag and wearing a coat. These were assessed using

Haralick features with and without RELIEF, using horizontal and vertical GEI division against other existing methods proposed in (Bashir et al., 2010) (Khalid Bashir et al., 2009) (Hu et al., 2013) and (Dupuis et al., 2013).

The results shown in the table correspond to the classification performance in terms of at rank-1 (%). By analysing the performance, it can be observed that the proposed method based on Haralick features with RELIEF using horizontal GEI division improves the recognition performance in terms of IR at rank-1 to 80% while vertical GEI division yields a result of 71.67%.

The proposed method produces comparable results in the case of "normal walking" and "carrying-bag" conditions while providing an improved IR at rank-1 in the case of "wearing-coat" condition.

The proposed method outperforms by up to 26.00%, 31.00%, 32.00% and 13.00% compared against the methods (Khalid Bashir et al., 2009), (Khalid Bashir et al., 2010), (Hu et al., 2013) and (Dupuis et al., 2013), respectively. This proposed method outperforms all the state-of-the-art methods considered in our experiment. When compared to the best mean IR Rank-1 provided by the state-of-the-art methods ranging from 60.70% to 77.96%, our proposed method achieves an improved IR at rank-1 up to 80.00% for a side view of 90°.

Finally, the experiment clearly demonstrates that the proposed method considerably improves the recognition performance in the presence of the following covariates; normal walking, carrying bag and wearing a coat, and outperforms the state-of-the-art methods showing an increase of up to 2.00% compared to the method in (Dupuis et al., 2013). We have also assessed the performance of the proposed method using CASIA database B under four side views at 36°, 72°, 90° and 108°.

Table 4-3. Comparative studies of the proposed method with different state-of-the-art methods on CASIA database B for a side view of 90°. Three covariates were considered in here: normal walking, carrying bag, and wearing a coat.

Methods	Covariates			Mean IR Rank-1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
(Khalid Bashir et al., 2009)	97.50	83.60	48.80	76.60
(Bashir et al., 2010)	100.00	87.30	44.00	75.80
(Hu et al., 2013)	94.60	45.20	42.90	60.70
(Dupuis et al., 2013)	97.60	73.8	62.50	77.96
Proposed method				
Haralick + RELIEF / H	85.36	79.90	74.74	80.00
Haralick + RELIEF / V	78.50	59.35	67.00	71.67

Table 4-4 shows performance results obtained in terms of IR at rank-1 and rank-5. From the results obtained; it can be observed that the proposed method achieves an acceptable IR at rank-1 for both horizontal and vertical GEI divisions and for different viewing angles (up to 80.00% and 71.67% for horizontal and vertical division respectively). The IR is increased at rank-5 to 91.12% and 84.67% for horizontal and vertical division respectively. This demonstrates that the proposed method allows recognition gait under different viewing angles.

Table 4-4. Comparison of IR (in %) from the proposed method on CASIA database (dataset B) for four side views 90°.

Angle view	Proposed method			
	Haralick + RELIEF/H		Haralick + RELIEF/V	
	IR rank-1 (%)	IR rank-5 (%)	IR rank-1(%)	IR rank-5 (%)
36^0	70.80	85.48	63.19	80.64
72^0	79.75	91.12	65.60	83.72
90^0	80.00	90.32	71.67	84.67
108^0	71.24	89.87	70.12	84.31

4.3.3 Further experiment using USF Database

4.3.3.1 Database and Evaluation Criteria

The proposed method was evaluated by using another database in order to assess its performances under other covariate factors such walking, shoe type and view. In this experiment only two probes (testing) were considered; Probe A (grass-walking surface + shoe type A + left camera viewpoint) and Probe C (grass walking surface + shoe type B + left camera viewpoint). The two probes are distinct and are categorised according to their covariate' variations. Probe A in view and Probe C in both view and shoe type simultaneously, with the Gallery (training) set being (G, A, R, NB). To experiment with the rest of the probes pre-processing is needed, as is the case with the state-of-the-art methods. This pre-processing procedure will be taken into account in future work.

4.3.3.2 Results and Analysis

Table 4-5 shows the results obtained using the proposed method. The proposed method was compared with the state-of-the-art methods that used the USF Human ID gait database, such as (Ju et al., 2006) and (Zhao et al., 2016). The results show that the proposed method provides encouraging results. In particular, it can be seen that the results are high for Probe A.

Table 4-5. Comparison of IR (in %) from the proposed method with the methods. in (Ju et al., 2006) and (Zhao et al., 2016) on USF Human ID gait database for Probe A and Probe C.

Method	IR (%)	Probe A (%)	Probe C (%)
Proposed method	rank-1	72.13	48.14
	rank-5	90.98	74.07
Method in (Ju et al., 2006)	rank-1	79.00	56.00
	rank-5	96.00	76.00
(Zhao et al., 2016)	rank-1	82	47
	rank-4	95	59

This could be due to the fact that within this method, in a local image, viewpoint variations do not drastically affect the performance. Whereas in Probe C, the difference in shoe type with the gallery set makes it a rather difficult experiment. The results obtained can be improved by pre-processing the silhouette images i.e. improvement of segmentation, the effect of shadow, removing surface area to keep the only shoe, etc.

4.4 Summary

This chapter has proposed a novel gait recognition method for a human identification under variations of clothing and carrying conditions for different viewing angles. The proposed method based on Haralick with RELIEF selection features technique was evaluated on the two databases (CASIA & USF) and compared against some similar techniques. The results obtained have shown that the proposed feature extraction is relevant and is very useful for gait recognition under the effect of clothing and carrying conditions for different viewing angles. The next chapter will discuss an investigation regarding multi-scale descriptors for feature extraction using MLBP and Gabor filter bank with a number reduction technique.

CHAPTER FIVE: GAIT RECOGNITION BASED ON MULTI-SCALE DESCRIPTORS

5.1 Introduction

This chapter discusses a supervised feature extraction approach that relies on two feature extraction methods based on multiscale feature descriptors including MLBP and the Gabor filter bank, utilising a reduction algorithm. The first proposed method includes a Gabor filter bank where the features are extracted from GEI. This method was evaluated on the CASIA Gait database under variations of clothing and carrying conditions for different viewing angles, with the experimental results analysed using an SVM classifier. Different reduction algorithms were used including Kernel Principal Component Analysis (KPCA), Spectral Regression Kernel Discriminant Analysis (SRKDA) and Maximum Margin Projection (MMP).

We improved the proposed method by only considering the extracted local features from two ROIs representing the dynamic areas in GEI. The experimental evaluation using the k-NN Classifier produced an impressive result with the highest Identification Rate (IR) at rank-1 when compared to similar recent state-of-the-art methods. Finally, the USF database was also used to evaluate our proposed method and the results clearly demonstrating that this suggested method outperforms a recent and similar technique (Dupuis et al., 2013).

The second proposal in this chapter is a method based on Multi-Scale Local Binary Pattern (MLBP), utilising the SRKDA reduction algorithm. In addition, the features are extracted locally from two ROIs representing the dynamic areas in GEI. The suggested method was evaluated on the CASIA and USF Gait databases. The experimental results using k-NN

classifier produced the highest identification rate at rank-1 when compared to similar and recent state-of-the-art methods.

5.2 The Proposed Method

We have investigated the details contained in a GEI to develop a feature extraction approach under clothing and carrying condition variations. The aim was to exploit the local features, which can be discriminated by horizontally dividing the GEI into two parts; top and bottom, referred to as the ROIs. In this approach, we have focused on the dynamic area of the silhouette extracted from the GEI. An illustrative example is shown in Figure 3-6.

5.2.1 Multi-scale Local Binary Pattern Descriptors

LBP method has been used for different biometric applications such as facial recognition (Ahonen et al., 2006) (Shan et al., 2009) and gait recognition (Kumar et al., 2014). It is one of the most effective descriptors to efficiently capture the local structures of an image by labelling their pixels. Labelisation is performed by thresholding the block of the neighbourhood of every pixel with each central value of a square window (Ojala et al., 1996). LBP, denoted here by (P, R), is calculated in a local circular region by subtracting the centre pixel with respect to its neighbours, where P is the number of the neighbours and R is the radius of the circular neighbourhood. The image pixels are labelled by thresholding the circular neighbourhood (P, R) of each pixel (i, j) with the central value and summing up the threshold values weighted by its power of two (see Figure 5-1). It is described as follows (Ojala et al., 2002).

$$LBP_{P,R}(i,j) = \sum_{P=0}^{P-1} C\,(GEI(i_p, j_p) - GEI(i,j))\, 2^P \tag{5.1}$$

Where (i_p, j_p) represents the neighbouring coordinates around a pixel (i, j), p is the index of the neighbour and C is the thresholding function defined as follows:

$$C\ (\lambda) = \begin{cases} 1, & \lambda \geq 0 \\ 0, & \lambda < 0 \end{cases} \tag{5.2}$$

In LBP-based texture classification approaches (Pietikäinen et al., 2011), the occurrences of LBP codes of an image are collected in a histogram h of the local binary patterns shown in equation 5.3. The main feature of LBP concept is that it is invariant to image translation.

$$h(i) = \sum_{x,y} B\big(LBP_{P,R}(i,j) = n\big) \quad | \quad n \in [0, 2^P - 1], \tag{5.3}$$

$B\ (v)$ is a Boolean indicator defined as:

$$B(v) \begin{cases} 1 & when\ v\ is\ true \\ 0 & otherwise \end{cases} \tag{5.4}$$

The histogram of GEI pixels is used as texture descriptors. Moreover, the LBP characteristics are computed from one scale with a 3×3 neighbourhood window. Also, $LBP_{8,1}$ is unable to detect the dominant structure and its image translation invariant. A multi-scale LBP, which is denoted as MLBP, proposed in (Pietikäinen et al., 2011) can be a useful solution to extract more texture details. This idea originated from simple observation real-world objects composed of various structures at different scales and appearing in different ways based on the scale of observation. MLBP is described as an extension of basic LBP in respect of the neighbourhood of various sizes.

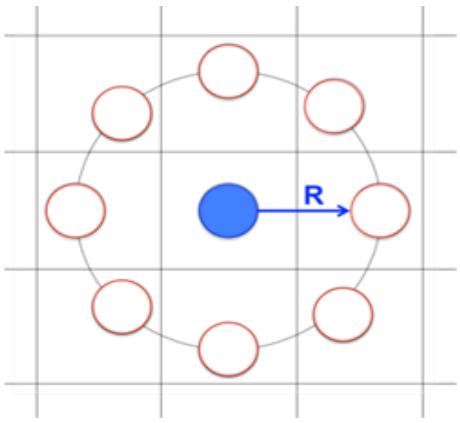


Figure 5-1. An example illustrates circularly symmetric neighbour sets for the operator of LBP with various values (P, R). (P=8 and R=1 (3×3) neighbourhood).

Therefore, in this work, we propose a second feature extraction method based on MLBP descriptors computed from a GEI, where the features extracted are combined by concatenating the set of histograms h calculated at different scales. MLBP is extracted and computed locally from various ROIs, generated by dividing the GEI horizontally in two ROIs (top and bottom), each ROI represents the relevant information for different conditions. However, the vector of features extracted from the GEI using the MLBP algorithm has a higher dimension, which may hamper the classification process. Thus, a feature reduction algorithm is necessary to extract only the useful and most informative descriptors for classification. Section 5.2.3 covers the reduction technique used.

5.2.2 Gabor Filter Bank-based Feature Extraction

This section describes Gabor filters and how they are used in the feature extraction. The Gabor filter was initially presented in 1946 (Gabor, 1946). A one-dimensional Gabor filter is characterised as the multiplication of a cosine/sine (even/odd) wave with Gaussian windows (see Figure 5-2), as follows (Derpanis, 2007):

$$g_e(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \cos(2\pi\omega_0 x) \tag{5.5}$$

$$g_o(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}} \sin(2\pi\omega_0 x) \tag{5.6}$$

Where g_e, g_o are Gabor (even / odd, respectively), ω_0 knows the centre frequency (i.e., the frequency in which the filter yields the utmost response) and σ the (potentially asymmetric) spread of the Gaussian window. The power spectrum of the Gabor filter is given by the sum of two Gaussians centred at $\pm\omega_0$, is defined as (Willsky, 1997):

$$\|G(\omega)\| = e^{-2\pi^2\sigma^2(\omega-\omega_0)^2} + e^{-2\pi^2\sigma^2(\omega+\omega_0)^2} \tag{5.7}$$

The power spectrum of a Gaussian is a (non-normalised) Gaussian and the power spectrum of a sine wave are two impulses located at $\pm\omega_0$. A multiplication in the temporal (spatial) domain is equivalent to a convolution in the frequency domain (Oppenheim, 1997). The discriminative features proposed in our feature extraction method include the Gabor filter features. The Gabor filter bank has eight orientations and five scales. Figure 5-3 shows an example of Gabor filter bank. The result of the convolution process can be characterised as (Lades et al., 1993):

$$G_{v,w}(x,y) = GEI(x,y) * \eta_{v,w}(x,y) \tag{5.8}$$

Where $*$ represents convolution, $\mu_{v,w}(x,y)$ is a 2D Gabor wavelet kernel function at orientation w and scale v, and $G_{v,w}(x,y)$ represents the convolution output. The kernel is defined by (Lades et al., 1993):

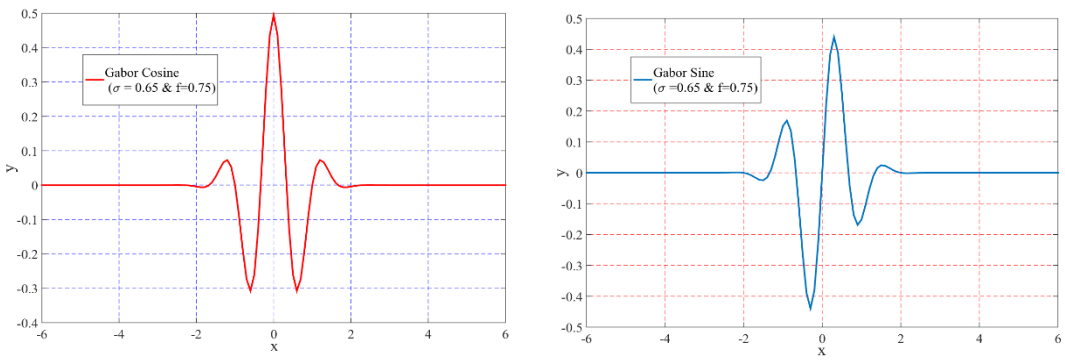


Figure 5-2. An example shows One-dimensional Gabor filters, (Derpanis, 2007, p. 2).

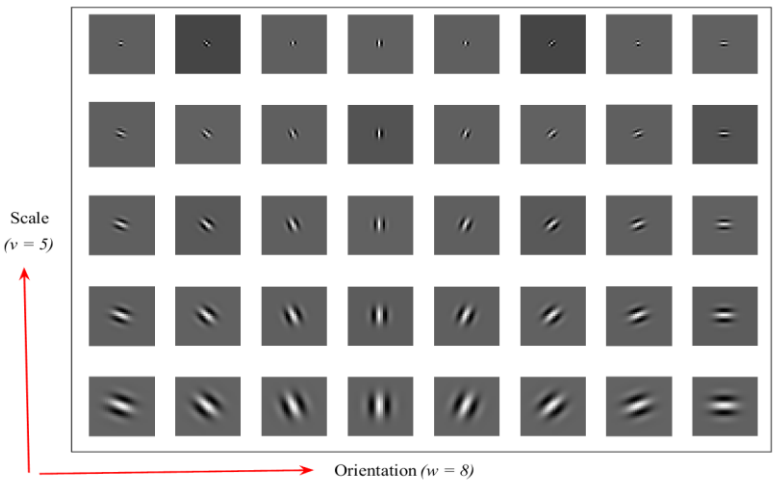


Figure 5-3. An example of Gabor Filter-bank with 5 Scales and 8 orientations, (Fischer et al., 2007, p. 234).

$$\eta_{v,w}(z) = \frac{\|k_{v,w}\|^2}{\sigma^2} e^{-(\|k_{v,w}\|^2 \|z\|^2} \left[e^{ik_{v,w}z} - e^{-\sigma^2/2} \right] \tag{5.9}$$

where $z = (x, y)$ and $\|\bullet\|$ is the Euclidean norm operator, $k_{v,w} = k_v e^{i\varphi v}$ with $k_v = \frac{k_{max}}{\lambda^v}$ where λ is the spacing factor between Gabor wavelets in the frequency domain and orientations. To reduce the redundancy of information resulting from the adjacent pixels in the image, a down-sampling of feature images were extracted from Gabor filters (Chengjun Liu, 2002). In this work, the feature size of the vector is a size of GEI (240 × 240) multiplied by the number of orientations and scales (8 × 5) and divided by the row and its column down-sampling factors (4 × 4), which are 240 × 240 × 8 × 5 / (4 × 4) giving 144,000. A Gabor filter bank is used locally from various ROIs generated by dividing the GEI horizontally into two ROIs (top and bottom part) with each part representing the relevant information for different conditions. Since the feature vector extracted from the GEI has a high dimensionality, a feature reduction algorithm is necessary to extract only the useful and discriminative features for classification. Section 5.2.3 has referred to the feature reduction technique used.

5.2.3 Feature Reduction

This part analyses various feature vector reduction techniques with a view to select the most appropriate one for the application at hand. In almost all pattern recognition approaches, one often goes for data reduction or subspace mapping. This is done primarily to reduce or decorrelate the data. In this chapter, we investigate four feature reduction techniques: KPCA, SRKDA, MMP and LPP. The following discusses the approaches chosen in this work:

5.2.3.1 Kernel PCA

KPCA is a reformulation of conventional Linear PCA in a high dimensional space produced using a kernel function (Bernhard Schölkopf et al., 1998). KPCA calculates the principal eigenvectors of the Kernel matrix, as opposed to those of the covariance matrix. The

reformulation of PCA in a kernel space is apparent after a kernel matrix and the result of the data focus on the high-dimensional space that is built using the kernel function. The use of PCA in the kernel space gives KPCA the property of building nonlinear mapping. Arithmetically, the current features are transformed into a high-dimensional space and calculate eigenvectors in this space. The vectors with low eigenvalues are ignored and then learning in this transformed space. The consequence of the projection i.e. the low-dimensional data representation Z, is given by:

$$Z_i = \left\{\sum_{j=1}^{n}\sum_{j=1}^{n}\alpha_t^{(j)} M(X_j, X_i)\right\} \tag{5.10}$$

where $\alpha_1^{(j)}$ refer to the jth value in the vector b_1 and M is the kernel function that was additionally used in the calculation of the kernel matrix. Since kernel PCA is a kernel-based technique, the mapping performed by $KPCA$ depends on the selection of the kernel function M.

5.2.3.2 SRKDA for Feature Dimensionality Reduction

The SRKDA algorithm (Cai et al., 2007) is an extension of the extensively used KDA (Baudat et al., 2000) and for extracting abstract features and to reduce the dimensionality. SRKDA has been successful in many classification tasks such as text, multi-class face retrieval, spoken and image/video letter recognition. The method combines the spectral graph analysis and regression for an efficient large matrix decomposition in KDA. In order to best describe the principle of SRKDA; suppose a set of g samples $\{m_1, m_2 \ldots m_n\} \in \mathbb{R}^d$, belonging to ν classes. Some nonlinear mappingΦ: $\mathbb{R}^d \longrightarrow \mathcal{F}$ induces to consider the problem in a feature space $\mathcal{F}$. Let the training vectors be represented as an $g \times g$ kernel matrix K such that: $K(m_i, m_j) = \langle \Phi(m_i), \Phi(m_j) \rangle$, where $\Phi(m_i)$ and $\Phi(m_j)$ are the embedding of data items m_i and m_i. If ϕ denotes a projective function into the kernel feature space, then the objective function for KDA is:

$$\max_{\phi} K(\phi) = \frac{\phi^T C_b \phi}{\phi^T C_t \phi} \tag{5.11}$$

Where C_b and C_t denote the between-class and total scatter matrices in the feature space, respectively. SRKDA only needs to solve a set of regularised regression problems and there is no eigenvector computation involved. This results in a significant improvement of the computational cost and allows the handling of large kernel matrices. After obtaining α, the decision function for new data item is calculated from:

$$f(x) = \sum_{t=1}^{m} \alpha_i K(x, x_i) \tag{5.12}$$

where $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]^T$ is the eigenvector, and K $(x, x_i) = \langle \Phi(x), \Phi(x_i) \rangle$.

5.2.3.3 Maximum Margin Projection:

An unsupervised MMP algorithm has been proposed in this work aiming to find the maximum margin separating hyperplanes that separate data points in different clusters, with the maximum margin and project input pattern into typical hyperplanes.

We can easily determine the data points with labels and with these assigned labels, we can train an SVM with a particular margin. The objective of MMP is to discover such labelling together with the trained SVM. The associated margin is the maximum among the SVMs trained on all conceivable labelling.

(X. He et al., 2008) suggested a manifold learning algorithm, called MMP, for dimensionality reduction. It is based on locality preserving neighbour relations and overtly exploits the class information for classification. It is a graph-based approach for learning a linear approximation to the intrinsic data manifold by making use of both labelled and unlabelled data (Belkin et al., 2002). It is likely that both geometrical and discriminant structures of the data manifold are found using this algorithm. J_o and J_v are used to express the mean weight the matrices of the between-class graph F_o and the within-class graph F_v, respectively.

MMP endeavours to guarantee that the connected points of F_v are as close together as possible, while the connected points of F_o are as far apart as possible. It can be obtained by solving the following optimisation problem (Z. Wang et al., 2013):

$$\arg min_c \sum_{i-1}^{n} \sum_{j-1}^{n} (c^T x_i - c^T - x_j)^2 J_{v,ij} = \arg min_c \, c^T X (S_w - J_v) X^T c \tag{5.13}$$

$$\arg min_c \sum_{i-1}^{n} \sum_{j-1}^{n} (c^T x_i - c^T - x_j)^2 J_{o,ij} = \arg min_c \, c^T X L_h X^T c \tag{5.14}$$

With the constraint

$$c^T X S_w X^T c = 1 \tag{5.15}$$

Where $L_h = D_m - J_v$ is the graph Laplacian matrix (Chung, 1997) of F_o, D_m is a diagonal matrix whose diagonal entries are the column sum of $J_o, i.e, D_{m,ij} = \sum_{j-1}^{n} J_{o,ij}$, and S_w is a diagonal matrix whose diagonal entries are the column $sum\ of\ J_v, i.e., S_{w,ij} = \sum_{j-1}^{n} J_{v,ij}$. The definitions of weight matrices J_v and J_o are as per the following (X. He et al., 2008):

$$J_{v,ij} = \begin{cases} \gamma\ , if\ x_i\ and\ x_j\ share\ same\ lable \\ 1\ ,\ if\ x_i\ or\ x_j\ is\ unlabeled\ but\ x_i\ \in\ K_w(x_j)\ or\ x_j\ \in\ K_w(x_i) \\ 0\ ,\ \ otherwise \end{cases} \tag{5.16}$$

$$J_{o,ij} = \begin{cases} 1\ , if\ x_i\ \in\ K_b(x_j)\ or\ x_j\ \in\ K_b(x_i) \\ 0,\ \ otherwise \end{cases} \tag{5.17}$$

Where $K_b(x_i) = \{x_i^1, \dots, x_i^R\}$ denotes the set of its R nearest neighbours, $\zeta(x_i)$ represents the labels of x_i, $K_b(x_i) = \{x_i^j \,|\zeta(x_i^j) \neq \zeta(x_i), J = 1, \dots, R\}$ contains the neighbours having different labels, and $K_w(x_i) = K(x_i) - K_b(x_i)$ contains the rest of the neighbours. Thereafter, minimising 5.13 and maximising 5.14 under the constraint 5.15, the next optimisation problem can be reduced to (Z. Wang et al., 2013):

$$argmax_c c^T X (\theta L_h + (1 - \theta) J_v) X^T c \tag{5.18}$$

Where θ is a suitable constant within $0 \leq \theta \leq 1$. He et al. proposed θ to be 0.5 (X. He et al., 2008). The projection vectors that maximises 5.18 is given by the maximum eigenvalue solution to the generalised eigenvalue problem:

$$X(\theta L_h + (1-\theta)J_v)X^T c = \lambda X S_w X^T c \tag{5.19}$$

As XS_wX^T is non-singular, in this case, PCA is applied to remove the components corresponding to zero eigenvalues. The work by He et al. (X. He et al., 2008) shares common properties with some of the works on combining classification and metric learning, such as Distance-Function Alignment (Gang Wu, 2005) and Spectral Kernel Learning (Steven C. H. Hoi, 2006). The projection vector of MMP can be regarded as the eigenvectors of the matrix $(XS_wX^T)^{-1}X(\theta L_h + (1-\theta)J_v)X^T$ associated with the largest eigenvalues.

5.2.3.4 Locality Preserving Projections

LPP is a useful algorithm for using linear dimensionality reduction. It builds a graph incorporating the neighbourhood information of the data set. Using the notion of the Laplacian of the graph, it is then possible to calculate a transformation matrix which maps the data points to a subspace (He et al., 2003). Constructing the Neighbourhood Information (Adjacency Graph) to represent the topological structure of training images in the high-dimensional image space, the adjacency graph has been used. LPP can include both the actual topological structure of the data and the user-specified label. It is a simple linear dimensionality reduction method which can be implemented on a non-iterative optimisation. It preserves more local information than the global. LPP's aim to find a map which preserves the local structure (He et al., 2005). Algorithm 5-1, Algorithm 5-2 and Algorithm 5-3 summarise our proposed method, and Figure 5-8 and Figure 5-9 illustrate the diagrams of the proposed supervised feature extraction and reduction approach, based on LBP/MLBP and Bank Gabor filter texture features with reduction algorithm techniques.

Algorithm 5-1: Gabor filter with (SRKDA, KPCA, or MMP) for GEI-based human gait recognition.

Feature extraction and selection method for GEI-based gait recognition based on Gabor filter bank descriptors via SRKDA, KPCA, or MMP reductions algorithm

Input: Silhouette images extracted over one gait cycle: S (x, y, t); t = 1, 2, ..., N
To calculate a GEI using Eq.2.1: G (x, y)
Compute Gabor filter descriptors: F_g
Apply SRKDA or KPCA or MMP reduction algorithm on F_g
Output: Relevant features set F

Algorithm 5-2: GEI-based gait recognition based on MLBP descriptors via SRKDA reduction algorithm.

Feature extraction and selection method for GEI-based gait recognition based on MLBP descriptors via SRKDA reduction algorithm

Input: Silhouette images extracted over one gait cycle: $S\ (x,y,t); t = 1, 2, ..., N$
To compute a GEI using Eq.2.1: $G\ (x,y)$
Divide GEI horizontally into 2 parts as illustrated Figure 3-6: $G_{(H_1)}(x,y)$, $G_{(H_2)}(x,y)$
For each $G_{(H_i)}$, $i = 1, 2$. Compute MLBP descriptors: $F_{(Hi)}$
Generate feature extraction set: $F_{(H)} = \{F_{(H1)}, F_{(H2)}\}$
Apply SRKDA reduction algorithm on $F_{(H)}$
Output: Relevant features set F

Algorithm 5-3: GEI-based gait recognition based on Gabor filter bank descriptors via KPCA, SRKDA or LPP reduction algorithm.

Feature extraction and selection method for GEI-based gait recognition based on Gabor filter bank descriptors via KPCA, SRKDA, or LPP reductions algorithm

Input: Silhouette images extracted over one gait cycle: $S\ (x, y, t)$; $t = 1, 2, ..., N$

To compute a GEI using Eq.2.1: $G\ (x, y)$

Divide GEI horizontally into 2 parts as illustrated Figure 3-6: $G_{(H_1)}(x,y)$, $G_{(H_2)}(x,y)$

For each $G_{(H_i)}$, $i = 1, 2$. Compute Gabor filter descriptors: $F_{(Hi)}$

Generate feature extraction set: $\mathbf{F}_{(H)} = \{F_{(H1)}, F_{(H2)}\}$

Apply KPCA or SRKDA or LPP reduction algorithm on $\mathbf{F}_{(H)}$

Output: Relevant features set F

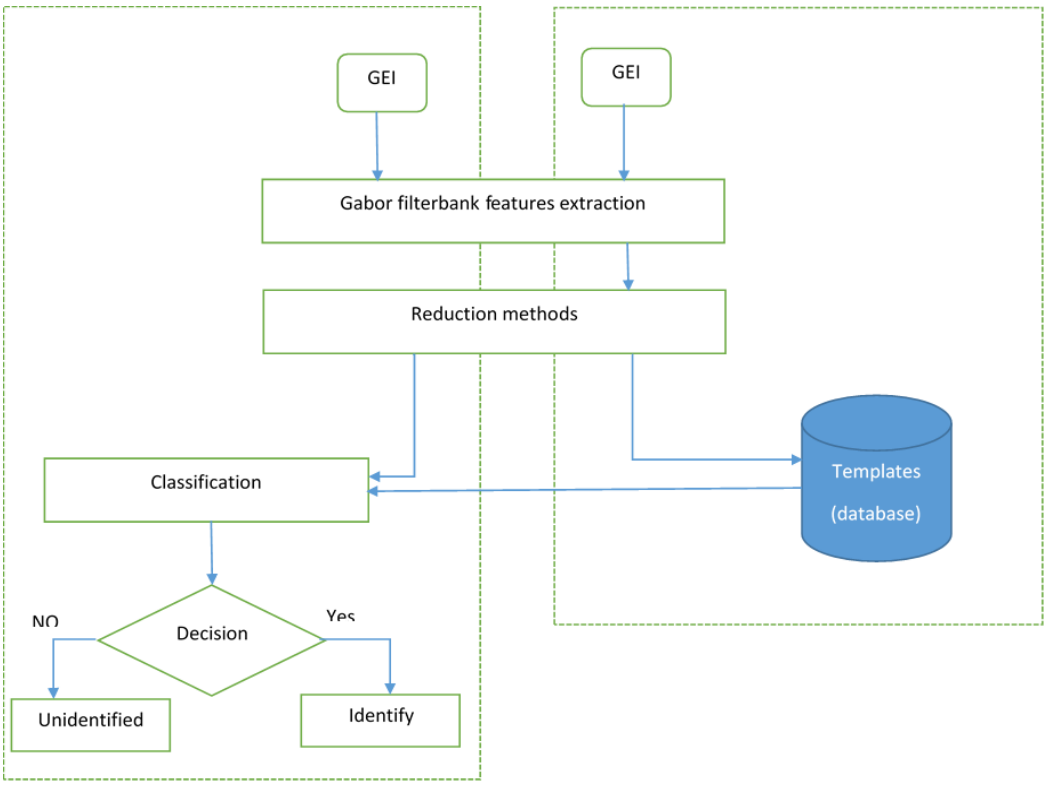


Figure 5-4. Diagram of the proposed supervised feature extraction and reduction approach based on Gabor filter bank descriptors with KPCA, SRKDA, and MMP reduction technique.

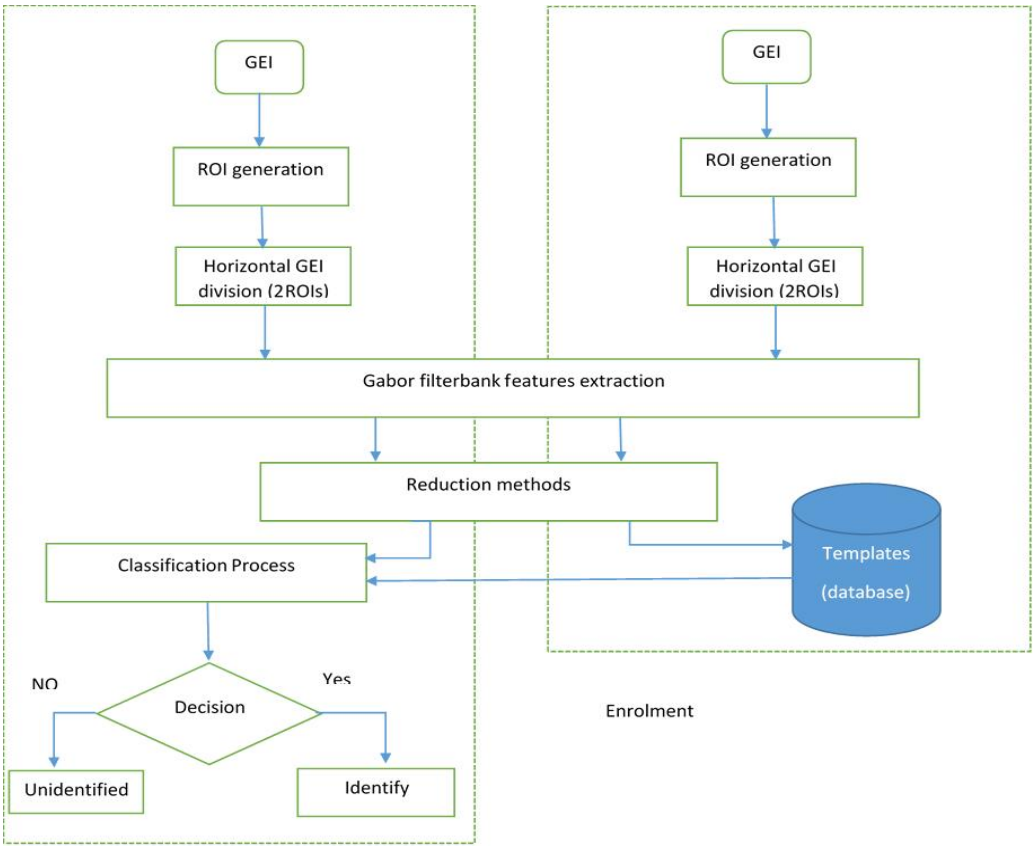


Figure 5-5. Diagram of the proposed supervised feature extraction and reduction approach based on Gabor filter bank descriptors with SRKDA, KPCA, and LPP reduction techniques.

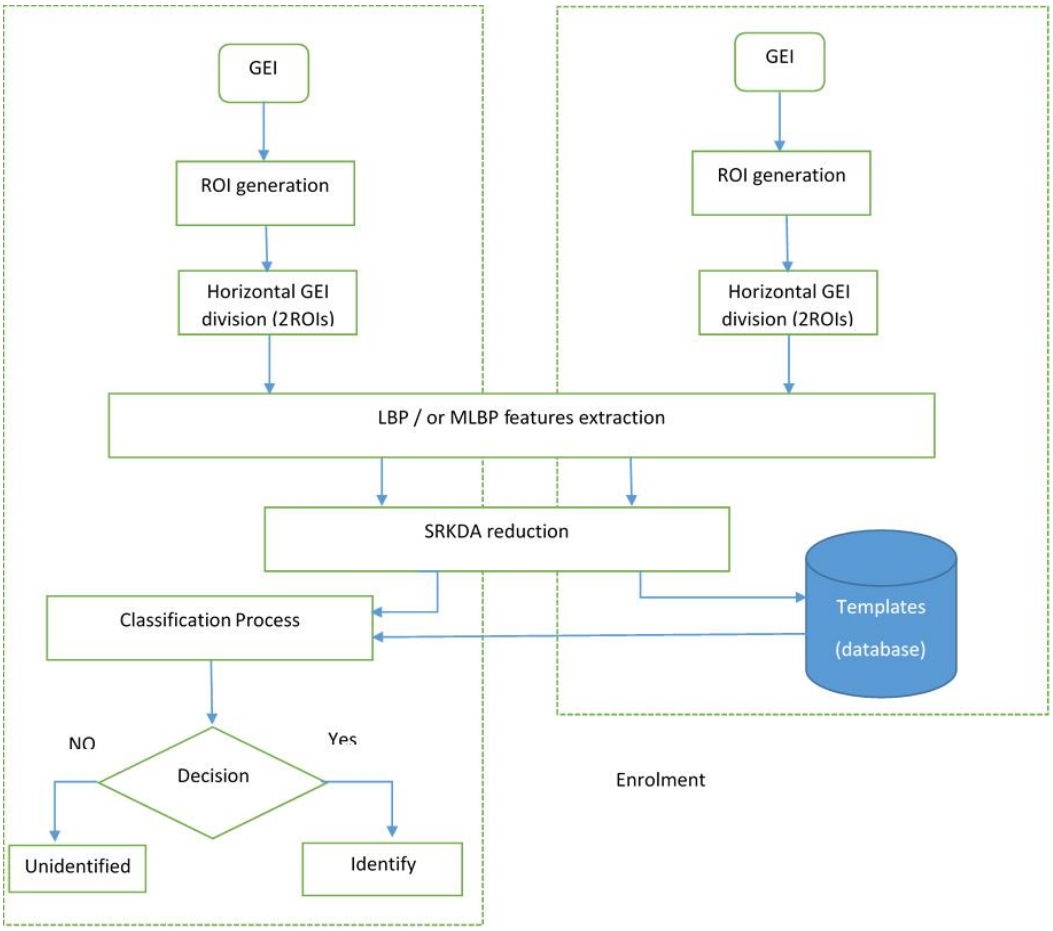


Figure 5-6. Diagram of the proposed supervised feature extraction and reduction approach based on LBP/or MLBP descriptors with SRKDA reduction technique.

5.3 Experimental Results and Discussion

To evaluate the proposed algorithms described previously two databases have been used, CASIA and USF. This section also analyses the acquired results and compares them against some existing and similar state-of-the-art methods.

5.3.1 Experiment 1 using CASIA Database

5.3.1.1 Database and Evaluation Criteria

The first sequence from Set-A, Set-B and Set-C of the database have been selected for the experimentation. The selected data was split randomly into two parts. The first part based around training and the other part for testing the effect of conditions, which has been extensively explained in a previous chapter (section 4.3.1.1).

5.3.1.2 Results and Analysis

The proposed approach is based on the use of feature texture descriptors extracted from GEI. The suggested features are computed using the Gabor filter bank approach and then selected using different reduction algorithms i.e. SRKDA, KPCA and MMP. The proposed method is evaluated on the CASIA Gait database (dataset B) under variations of clothing and carrying conditions for different viewing angles and the experimental results are evaluated using the one-against-all SVM classifier.

Table 5-1 shows the results obtained for four side views (36°, 72°, 90° and 108°) with the selected data split randomly into two parts; 50% for training and 50% for testing. One can notice that the computation of the Gabor filter bank accomplished an impressive performance in classifying different covariates for different viewing angles. For instance, the Gabor filter bank using the MMP reduction technique achieved a high result at a view angle of 90°, while a high result was obtained using SRKDA at a 72° angle. The proposed method therefore achieved substantial CCR results ranging between 87% and 91% for different viewing angles.

Table 5-1. Comparison of CCRs (in %) from the proposed Gabor filter bank via SRKDA, KPCA AND MMP reduction on CASIA Database for four side views: 36°, 72°, 90° and 108°.

Data partitioning (training–testing)	Reduction method	Angle view	Covariates			Mean IR rank-1(%)
			Normal Walking (%)	Carrying a bag (%)	Wearing a coat (%)	
50 % - 50 %	SRKDA	36°	85.48	83.87	93.54	87.63
		72°	95.70	85.48	90.32	90.50
		90°	93.55	87.63	89.24	90.14
		108°	91.93	86.56	91.40	89.96
50 % - 50 %	KPCA	36°	82.25	78.49	89.78	83.51
		72°	93.54	84.40	88.71	88.88
		90°	89.78	85.48	91.39	88.88
		108°	90.32	83.33	89.24	87.63
50 % - 50 %	MMP	36°	85.48	84.40	91.93	87.27
		72°	94.62	85.48	90.32	90.14
		90°	92.47	89.24	90.32	90.68
		108°	91.93	86.55	91.39	89.96

5.3.2 Experiment 2 using CASIA Database

5.3.2.1 Database and Evaluation Criteria

The proposed methods are evaluated using CASIA Dataset B where the gallery set consists of the first four sequences of each subject of Set-C (CASIA set-C1). The probe is the rest of sequences of Set-C (CASIA Set-C2), Set-A and Set-B.

We assessed the performance of the planned method under the following three conditions; 1) normal, 2) carrying a bag and 3) wearing a coat. Experiments were conducted from viewing angles of 36°, 72°, 90° and 180°: As for the evaluation, *k*-NN classifier was used to assess the classification performances. The highest IR at rank-1 was used to evaluate the performance.

5.3.2.2 Results and Analysis

We have evaluated the performance of the computation of MLBP and Gabor filter bank texture features locally and globally from the whole GEI. Table 5-2 shows that the computation of the local features from the ROIs selected improves the performance of the suggested gait recognition method significantly (up to 29% using MLBP and 7% using Gabor filter bank) compared to global computation from the whole GEI image.

Table 5-2. Recognition performances of proposed method based on local and global feature computation techniques on CASIA database using a side view of 90°.

Features computation technique	Covariates			Mean IR rank-1(%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Global (MLBP)	72.98	18.95	11.69	34.54
Global (Gabor filter Bank)	97.17	72.58	53.66	74.47
Local (MLBP)	73.38	56.85	59.67	63.30
Local(Gabor filter bank)	89.11	68.14	86.69	81.31

Table 5-3. Comparison of IR rank-1 (in%) from the proposed method based on local and global feature computation techniques with SRKDA on the CASIA database using a side view of 90°.

Features computation technique	Covariates			Mean IR rank-1(%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Global (MLBP+SRKDA)	93.95	66.53	33.87	64.78
Global (Gabor filter bank + SRKDA)	98.79	74.59	54.43	75.93
Local (proposed method/ LBP+SRKDA)	95.56	89.91	89.11	91.52
Local (proposed method/Gabor filter bank + SRKDA)	97.58	86.69	91.93	92.06

Looking at Table 5-3, it is clear that MLBP and Gabor filter bank with SRKDA significantly improves the performance of the recognition by up to 26.74% and 16.13% respectively when compared against global MLBP and Gabor Filter bank approaches. We have also assessed the performance of our proposed methods using the selected data from CASIA database for a side view of 90°.

Table 5-4 compares our proposed methods based on MLBP and Gabor Filter with SRKDA reduction technique against four other existing and similar methods i.e. methods proposed in (Khalid Bashir et al., 2009), (Bashir et al., 2010), (Hu et al., 2013) and (Dupuis et al., 2013). This MLBP method is based on eight scales (a radius of 1, 2, 3, 4 and 8). In addition, Gabor filter bank use 8 orientations and 5 scales. The results shown in the table correspond to the classification performance in IR (%).

The proposed methods yielded comparable results for "normal walking" but provided the best IR at rank-1 for the case of "wearing a coat" when compared to the works of (Khalid Bashir et al., 2009), (Bashir et al., 2010), (Hu et al., 2013) and (Dupuis et al., 2013). In this MLBP method, results increased by up to 40.31%, 45.11%, 46.21% and 26.61% whilst the results of

the proposed Gabor filter bank method achieved an increase of up to 43.13%, 47.93%, 49.03% and 29.43% in comparison to the aforementioned works.

Furthermore, the results attained using the MLBP method provided the best IR at rank-1 for the case of "carrying a bag" compared to the other methods e.g. an increase of up to 6.31%, 11.61%, 44.71%, and 16.11% were noted. Also, in the proposed Gabor filter bank method, the results showed increases of up to 3.09%, 8.39%, 41.49% and 12.89% in comparison to the aforementioned referenced works.

The proposed method outperforms all the state-of-the-art methods considered in our experiment. When compared to the best IR at rank-1 provided by state-of-the-art methods, which are in the range of 60.70% to 77.96%, our method achieves a better IR at rank-1 up to 92.06% for the side view of 90°. We have assessed the performance of the proposed feature extraction and reduction method using CASIA database from four side views (36°, 72°, 90° and 108$^{o)}$. The propositioned features are compared with LBP features, LBP features with SRKDA MLBP features and MLBP features with SRKDA.

Table 5-4. Recognition performances of the proposed method with several different state-of-the-art methods on the CASIA database from the side view of 90°. Three covariates were considered here: normal walking, carrying a bag and wearing a coat.

Method	Covariates			Mean IR rank-1(%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
(Khalid Bashir et al., 2009)	97.50	83.60	48.80	76.60
(Bashir et al., 2010)	100.00	78.30	44.00	74.10
(Hu et al., 2013)	94.00	45.20	42.90	60.70
(Dupuis et al., 2013)	97.60	73.80	62.50	77.96
Proposed method :				
MLBP+SRKDA	95.56	89.91	89.11	91.52
Gabor filter bank +SRKDA	97.58	86.69	91.93	92.06

Table 5-5 shows the performance results obtained in terms of IR rank-1 for the different types of feature considered in our study. By analysing the results as shown in the table, we can see that the IR improved up to 91.52% by increasing the number of GEI in the 90°. This confirms that MLBP with SRKDA features has an enhanced discriminating power, leading to an IR that achieves the highest seen percentage. In addition, we have assessed the performance of the proposed feature extraction and reduction method on the CASIA database from four side views, as mentioned before, at angles of 36°, 72°, 90° and 108°. The proposed features compared the Gabor filter bank with different reduction techniques i.e. KPCA, SRKDA and LPP.

Table 5-5. Recognition performances of proposed methods on the CASIA database from four side views: 36°, 72°, 90° and 108°. The proposed features MLBP are compared with LBP features.

Ranks	Angle view	Methods			
		LBP (%)	LBP+SRKDA (%)	MLBP (%)	MLBP+SRKDA (%)
Rank1	36^0	45.96	71.50	50.67	79.43
	72^0	64.37	85.61	68.67	90.85
	90^0	61.96	84.80	63.30	91.52
	108^0	55.77	80.37	60.60	86.28

Initially, the KPCA reduction technique was applied. Table 5-6 compares techniques on the CASIA database for four, previously noted, side views (36°, 72°, 90° and 108°). The method yielded comparable results for the cases of "normal walking" and provided best IR at rank-1 on the method Gabor filter bank without KPCA under angle 36°. The IR improved up to 94.35%, but with the KPCA, it provided a figure of 91.94%. In the case of "wearing a coat", the best IR at rank-1 with Gabor filter bank without KPCA is obtained for an angle of 72°. The IR was improved up to 93.54%. Correspondingly, it provides the best IR at rank-1 for the case of "carrying a bag" using the Gabor filter bank without KPCA for an angle of 72° improving it to 87.90%. Table 5-7 shows a summary of the results obtained with regard to the IR rank-1 considered in our study.

By analysing these results, it can be seen that the IR is improved to 81.31% by increasing the number of GEI an angle of 90°. This clarifies that Gabor filter bank method without KPCA features has a more discriminating power, allowing for the attainment of the best IR. Additionally, we assessed the performance of the proposed Gabor filter via KPCA using CASIA database from different viewing angles.

Table 5-6. Recognition performances of Gabor filter bank method using CASIA database for four side views: 36°, 72°, 90° and 108° under normal walking, carrying a bag and wearing coat conditions.

Ranks	Method	Angle View	Covariates			Mean IR rank-1 (%)
			Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Rank-1	Gabor filter bank	36°	94.35	65.72	76.61	78.89
		72°	60.08	87.90	93.54	80.50
		90°	68.14	86.69	89.11	81.31
		108°	66.12	81.04	91.53	79.56
	Gabor filter bank + KPCA	36°	91.94	60.88	70.56	74.46
		72°	90.72	57.25	85.88	77.95
		90°	90.13	55.24	85.48	77.28
		108°	91.53	66.13	81.05	79.57

Table 5-7. Summary of recognition performances from the proposed methods using CASIA database from four side views: 36°, 72°, 90° and 108° Gabor Filter bank with KPCA.

Ranks	Angle view	Method	
		Gabor filter bank (%)	Gabor filter bank + KPCA (%)
Rank1	36°	78.89	78.89
	72°	80.50	77.95
	90°	81.31	77.28
	108°	79.56	79.57

The second proposed (Gabor filter bank with SRKDA) extractive technique has also been evaluated in terms of its performance using CASIA database under four side view angles of 36°, 72°, 90° and 108°. Table 5-8 depicts the results obtained for the four side views. It can be observed that the proposed extraction method yields comparable results in the case of "normal

walking" and provides the best IR at rank-1 for a side view angle 90° resulting in an improved IR of 97.58%.

On the other hand, in the case of "wearing a coat", the technique provides the best IR at rank-1 under a side view angle of 72° with IR of 96.77%. Finally, in the case of "carrying a bag", the technique provides the best IR at rank-1 under a side view angle of 72° giving an IR improved of 90.32%.

Table 5-8. Recognition performances of Gabor filter bank method using CASIA database for four side views: 36°, 72°, 90° and 108° under normal walking, carrying a bag and wearing coat conditions.

Ranks	Method	Angle View	Covariates			Mean IR rank-1(%)
			Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Rank-1	Gabor filter bank	36°	94.35	65.72	76.61	78.89
		72°	60.08	87.90	93.54	80.50
		90°	68.14	86.69	89.11	81.31
		108°	66.12	81.04	91.53	79.56
	Gabor filter bank with SRKDA	36°	82.66	86.29	96.37	88.44
		72°	78.22	90.32	96.77	88.43
		90°	97.58	86.69	91.93	92.06
		108°	87.50	84.67	95.16	89.11

Table 5-9 shows a summary of the performance results for the proposed feature extraction methods. An analysis of the results of the table above shows that the IR rank-1 has been further improved to 92.06 % by increasing the number of GEI at 90°. This confirms that Gabor filter bank with SRKDA approach yields more discriminating power.

Finally, in considering the performance of the proposed Gabor filter bank when combined with the LPP algorithm using the CASIA database for different viewing angles, table 5-11, shows a summary the of performance results obtained in terms of IR rank-1, the results clearly show that the IR was improved to 90.72 % by increasing the number of GEI.

Table 5-9. Summary of recognition performances from the proposed methods using CASIA database from four side views: 36°, 72°, 90° and 108° Gabor Filter bank with SRKDA.

Ranks	Angle view	Method	
		Gabor filter bank (%)	Gabor filter bank + SRKDA (%)
Rank-1	36°	78.89	88.44
	72°	80.50	88.43
	90°	81.31	92.06
	108°	79.56	89.11

Table 5-10. Recognition performances of Gabor filter bank method using CASIA database for four side views: 36°, 72°, 90° and 108° under normal walking, carrying a bag and wearing coat conditions.

Ranks	Method	Angle view	Covariates			Mean IR rank-1 (%)
			Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Rank-1	Gabor filter bank	36°	94.35	65.72	76.61	78.89
		72°	60.08	87.90	93.54	80.50
		90°	68.14	86.69	89.11	81.31
		108°	66.12	81.04	91.53	79.56
	Gabor filter bank + LPP	36°	97.58	77.82	85.48	86.96
		72°	96.37	79.43	92.33	89.37
		90°	98.38	81.85	91.93	90.72
		108°	96.37	85.08	87.90	89.78

Table 5-11. Summary of recognition performances from the proposed methods using CASIA database from four side views: 36°, 72°, 90° and 108° Gabor Filter bank with LPP.

Ranks	Angle View	Method	
		Gabor filter bank (%)	Gabor filter bank + LPP (%)
Rank-1	36°	78.89	86.96
	72°	80.50	89.37
	90°	81.31	90.72
	108°	79.56	89.78

5.3.3 Experiment 3 using USF Database

5.3.3.1 Results and Analysis

Table 5-12 and Table 5-13 depict the results obtained using the proposed method including a comparative against some state-of-the-art methods in (Ju et al., 2006) and (Zhao et al., 2016) using USF Human ID gait database (S. Sarkar et al., 2005).

From the results obtained, it can be noted that our proposed method provides encouraging results, which are comparable to the results of the methods in (Ju et al., 2006), and (Zhao et al., 2016).

Table 5-12. Recognition performances of MLBP with the methods in (Ju et al., 2006) and (Zhao et al., 2016) on USF Human ID gait database for Probe A, Probe C, Probe H, and Probe J.

Method	IR (%)	Probe A (%)	Probe C (%)	Probe H (%)	Probe I (%)	Probe J (%)
Proposed Method	rank-1	81.96	61.11	34.16	33.33	26.66
	rank-5	92.62	81.48	64.16	63.33	45.83
(Ju et al., 2006)	rank-1	79.00	56.00	N/A	N/A	N/A
	rank-5	96.00	76.00	N/A	N/A	N/A
(Zhao et al., 2016)	rank-1	82.00	47.00	N/A	N/A	N/A
	rank-4	95.00	59.00	N/A	N/A	N/A

Table 5-13. Recognition performances of Gabor filter bank with the methods in (Ju et al., 2006) and (Zhao et al., 2016) on USF Human ID gait database for Probe A, Probe C, Probe H, and Probe J.

Method	IR (%)	Probe A (%)	Probe C (%)	Probe H (%)	Probe I (%)	Probe J (%)
Proposed Method	rank-1	88.52	72.22	70.00	68.33	53.33
	rank-5	99.18	92.59	90.00	88.33	77.50
(Ju et al., 2006)	rank-1	79.00	56.00	N/A	N/A	N/A
	rank-5	96.00	76.00	N/A	N/A	N/A
(Zhao et al., 2016)	rank-1	82.00	47.00	N/A	N/A	N/A
	rank-5	95.00	59.00	N/A	N/A	N/A

The results acquired can be improved by pre-processing the silhouette images i.e. improvement of segmentation, the effect of shadow and removing the surface area to keep only the shoe, etc. Note that Probe H, Probe I and Probe J were not considered in (Ju et al., 2006) and (Zhao et al., 2016).

5.4 Summary

This chapter has proposed a supervised feature extraction approach capable of selecting more discriminating features for human gait recognition under variations of clothing and carrying conditions in order to improve recognition performance. The suggested methods based on MLBP and Gabor filter bank features are evaluated using the CASIA database and compared against similar techniques. The results obtained have shown that the proposed feature extraction methods are very useful for use in gait recognition under the effect of clothing and carrying conditions for different viewing angles. Also, our experiments have demonstrated that the propositioned methods outperform recent state-of-the-art methods such as (Dupuis et al., 2013). In particular, it is worth noting that SRKDA and LPP feature reduction techniques outperform KPCA counterparts using the proposed Gabor filter bank approach while the LPP being less computationally intensive than SRKDA. The next chapter will discuss a wavelet-based feature extraction based on the Haar wavelet.

CHAPTER SIX: GAIT RECOGNITION IN THE WAVELET DOMAIN

6.1 Introduction

This chapter will investigate and discusses the potential of using a wavelet domain feature extraction method to use for gait recognition under clothing and carrying conditions. The technique is based on the wavelet coefficients of the Haar wavelet, extracted from the dynamic areas of GEI. The SRKDA technique is also applied to the extracted feature vector to reduce its dimensionality by selecting only the most relevant and discriminate features. The proposed method was evaluated using the CASIA Gait database under various clothing and carrying conditions and viewing angles. The experimental study used the *k*-NN classifier.

6.2 The Proposed Method

In this thesis, a supervised feature extraction method based on extracting feature coefficients from GEI has been proposed for human gait recognition. The proposed method, described in Figure 6-1, is capable of extracting the most distinctive features from GEI under different covariates and conditions hence improving the recognition performance. A discrete wavelet transform (DWT), based on gait features, is applied for gait recognition. The proposed method is based on sub-bands which are used to extract gait features. Furthermore, in this method, four decomposition levels are used to extract a feature vector and the feature template is generated by concatenating these sub-images into a single image. The main idea is to locally capture the discriminating features that characterise a person's dynamic gait. To achieve this, a GEI has

split into a top part and a bottom part, in order to consider the dynamic portions of the human gait. Figure 3-6 provides an illustrative example of the aforementioned splitting process.

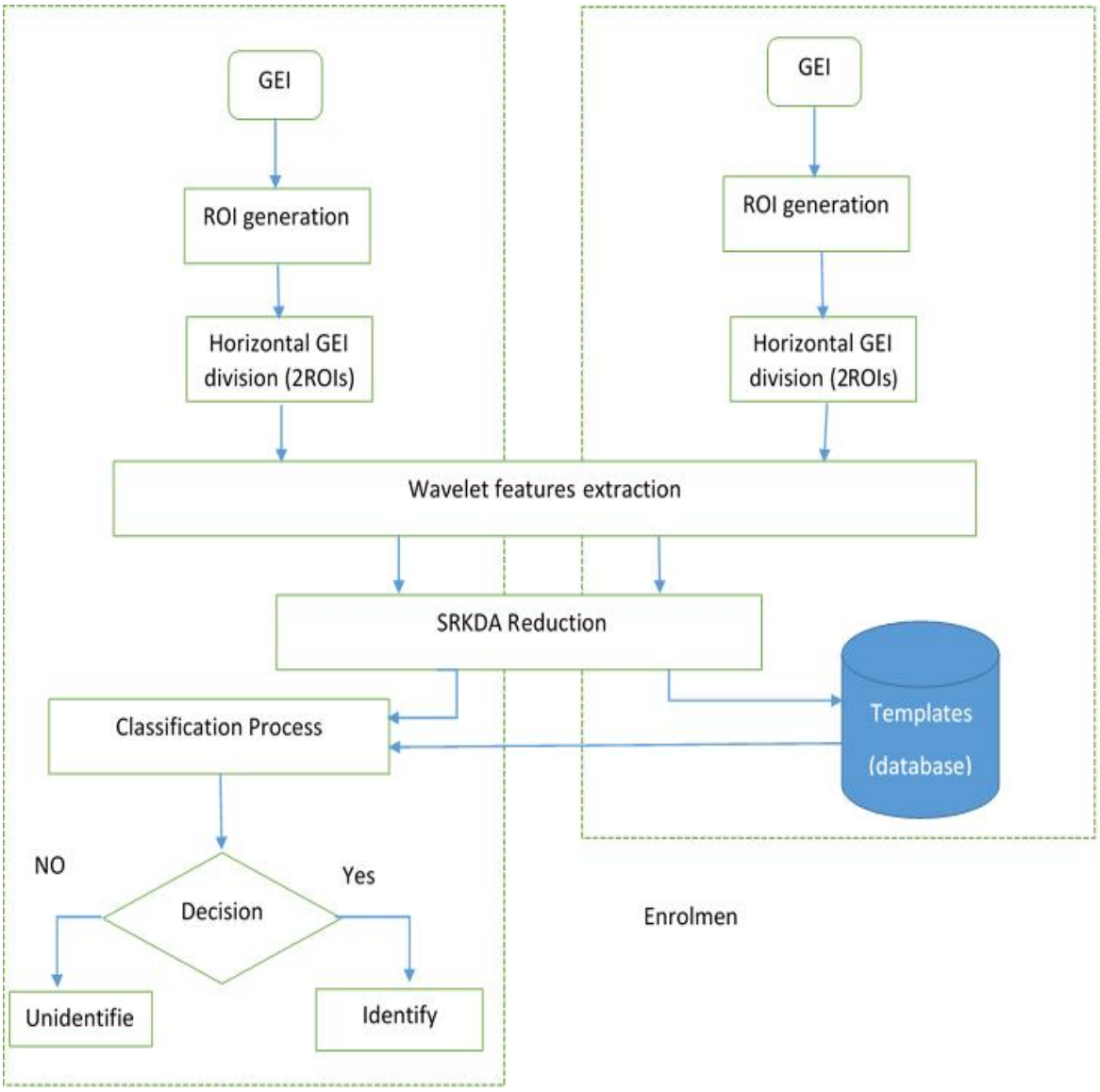


Figure 6-1. Diagram of the proposed supervised feature extraction and reduction approach based on wavelet transform with SRKDA reduction technique.

6.2.1 Wavelet Transform

Wavelet transforms (WTs) have been widely used in various fields, such as image processing, signal processing, biochemistry and medicine, since their first introduction by Alfred Haar in his thesis in 1909 (Haar, 1910). WTs have emerged as an alternative to the popular Fourier transform and its related transforms, such as the Discrete Cosine Transform (DCT). The main

idea behind the popularity of WTs is their localisation property in time, scale and frequency, which makes them suitable for analysing finite signals. WTs can be divided into many types, such as the continuous wavelet transform (CWT), the discrete wavelet transform (DWT), the two-dimensional wavelet transforms and the wavelet packet transforms (WPT).

More et al in (More et al., 2017) proposed a multi-view human gait recognition method which employs Partial Wavelet Coherence (PWC). This approach directly extracts the dynamic information without using any model. The proposed achieved a performance 73.26% average recognition accuracy when considered only PWC feature. Further, the paper investigates Phase Feature (PF) which also preserves the discriminant information of dynamic phase angle between body parts. When PF was considered in addition to PWC features the system performance improved significantly and 82.52% average recognition accuracy reported.

In this chapter, a wavelet-based 2D decomposition is introduced as a means to select the most discriminative features of the human gait (Mallat, 1999) and (Walker, 2002). The rationale behind using a wavelet transform is based on the fact that a wavelet transform can decompose an image at different levels of resolution. Thus, allowing images to be sequentially processed from low resolution to high resolution using wavelet decomposition as wavelets are localised in both the frequency (scale) and time (space) domains. Hence, it becomes easy to extract local features of an image. Wavelet descriptors have been used successfully to model the boundary of a moving human body. Nevertheless, it must be noted that many objects actually deform in some way as they move. Here we use wavelet descriptors to model not only the object's boundary, but also the spatio-temporal deformations under which the object's boundary is subjected (Rahati et al., 2008). According to Tong (Tong, 2010), one approach of feature extraction is utilising wavelet analysis, as introduced by (Papageorgiou et al., 2000).

6.2.1.1 Discrete Wavelet Transform.

The discrete wavelet transform (DWT) has been extensively applied in image processing, texture analysis, image compression and edge detection. DWT decomposes an image into four sub-images as shown in Figure 6-2, where filters are applied in the row and column directions separately. First, a high-pass filter and a low-pass filter are used to analyse each row's data then it is down-sampled by 2 in order to extract the high and low-frequency components of the row. The high-pass filter and low-pass filter are subsequently applied again for each of the high and low-frequency components of the columns, which are then down-sampled by 2. Through this process, four sub-bands images *LL*, *LH*, *HH* and *LL* are generated, each one having its own features. The low-frequency information is preserved in the *LL* sub-band and the high-frequency information is preserved in the *HH*, *HL* and *LH* sub-bands. The *LL* sub-band image can be further decomposed, in the same way as previously discussed, to produce a second level sub-band image. As such, in this method, four decomposition levels are used in total to extract a feature vector.

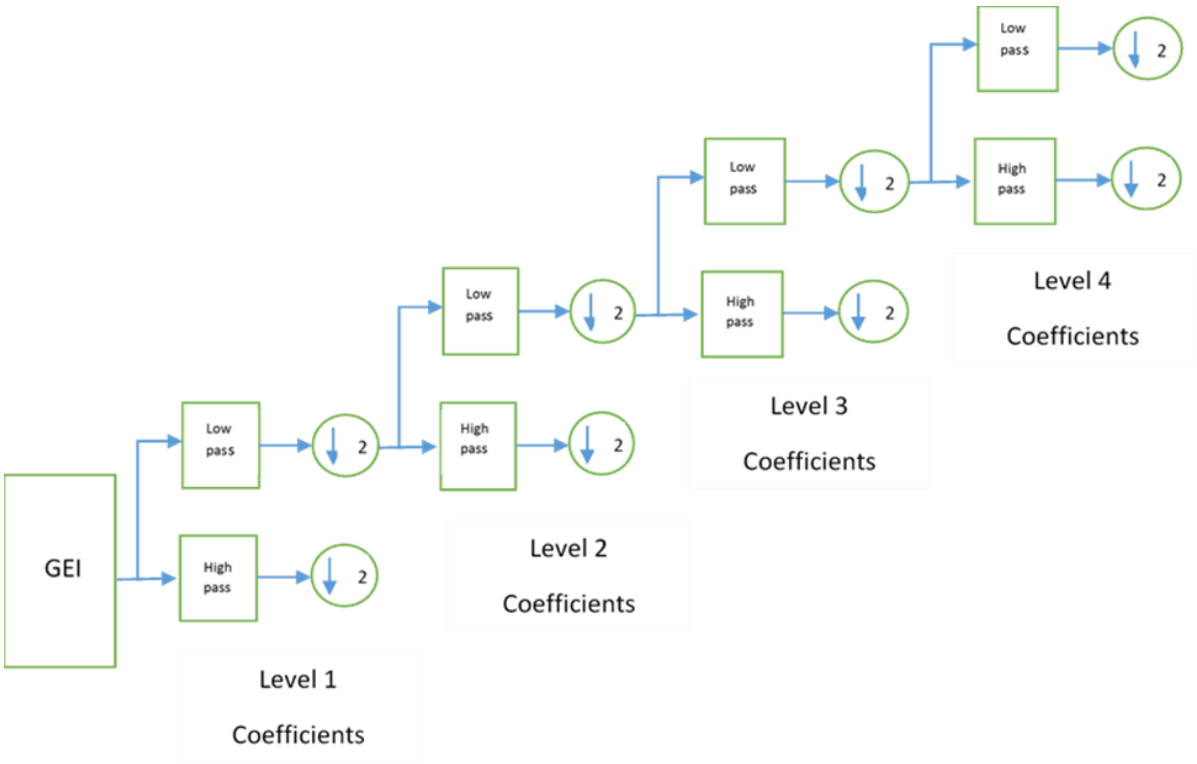


Figure 6-2. Discrete wavelet transform.

6.2.1.2 Detail Coefficients Wavelet Model

In this model, the Haar function is applied in DWT. The Haar function is the simplest example of a wavelet. Furthermore, the Haar wavelet provides satisfactory localisation of signal characteristics in the time domain. The Haar wavelet is characterised by its fast computation time, as it has the shortest filter length in the time domain. It is also the simplest possible wavelet available as it allows us to appropriately select or modify the wavelet coefficients. For example, it can remove the vertical, horizontal or diagonal details of a given image. It is the only known wavelet that is compactly supported, orthogonal and symmetric.

Discrete wavelet with Haar function was used on skeleton data and motion signals to extract features for gait recognition was demonstrated in (Arai et al., 2012) where the results showed that the best combination for classification is taken from horizontal detail and vertical detail. However, in (Arai et al., 2012) Haar wavelet was used at level 1 of decomposition where the energy for every coefficient is introduced.

Nandini et al. in (Nandini et al., 2011) suggested another gait recognition method in which they combined wavelet coefficients with three silhouette geometrical features. Initially, Haar wavelet transform was applied on each silhouette image of the gait sequence and the approximation coefficients of the low frequency sub-band were stored as the first feature vector. Then three silhouette geometrical features were extracted, the width, height, and area of the silhouette. These features were extracted from each frame in the gait sequence. The mean feature vector was then computed for each frame sequence. All experiments were conducted on CASIA A gait dataset and a recognition rate of 92.24% was attained which they showed to be better than two other compared gait recognition methods.

In this thesis, the proposed method aims to enhance the gait recognition accuracy by using the horizontal coefficient instead of using a combination of the three coefficients *HL*, *LH* and *HH*,

or a combination of (*HL*, *LH*), (*HL*, *HH*) and (*LH*, *HH*). More specifically, the Haar wavelet is used in this work as a function to decompose a GEI of size 240×240 pixels into four levels, with each level having three orientations as illustrated in figure 6-3 below.

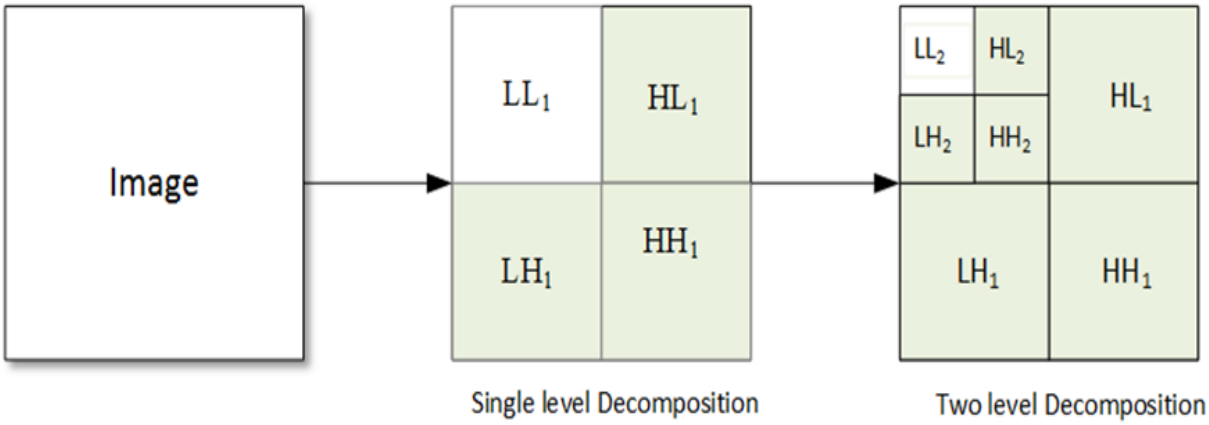


Figure 6-3. An illustrative example of a single level and two-level wavelet decomposition.

The proposed feature extraction is based on the application of 2D Haar wavelet decomposition on ROIs, as illustrated figure 3-6 in chapter 3, to extract the detail wavelet coefficients from the *LH*, *HL* and *HH* sub-bands at different scales (Mallat, 2008). Haar transform can be defined by the following equation (Jahromi et al., 2003):

$$\psi(I) = \begin{cases} 1 & 0 \leq I < 1/2 \\ -1 & 1/2 < I \leq 1 \\ 0 & otherwise \end{cases} \tag{6.1}$$

and $\psi_{Tl}(I) = \psi(2^T - l)$, for T a non-negative integer and $0 \leq l \leq 2^T - 1$, where $\psi(I)$ is the mother wavelet, $\psi_{Tl}(I)$ are scaled and translated versions of $\psi(I).l$ and T are the translation and scaling (dilation), respectively, of a factor of the wavelet.

The coefficients generated by concatenating the selected coefficient extracted from the four decomposition levels resulting in a high dimensionality thus requiring a dimensionality reduction. Figure 6-4 demonstrates the proposed fusion technique. To reduce the resulting high-dimensional feature vectors. Finally, we have applied the SRKDA algorithm discussed in the previous chapter. Algorithm 6-1 and figure 6-1 summarise our proposed method

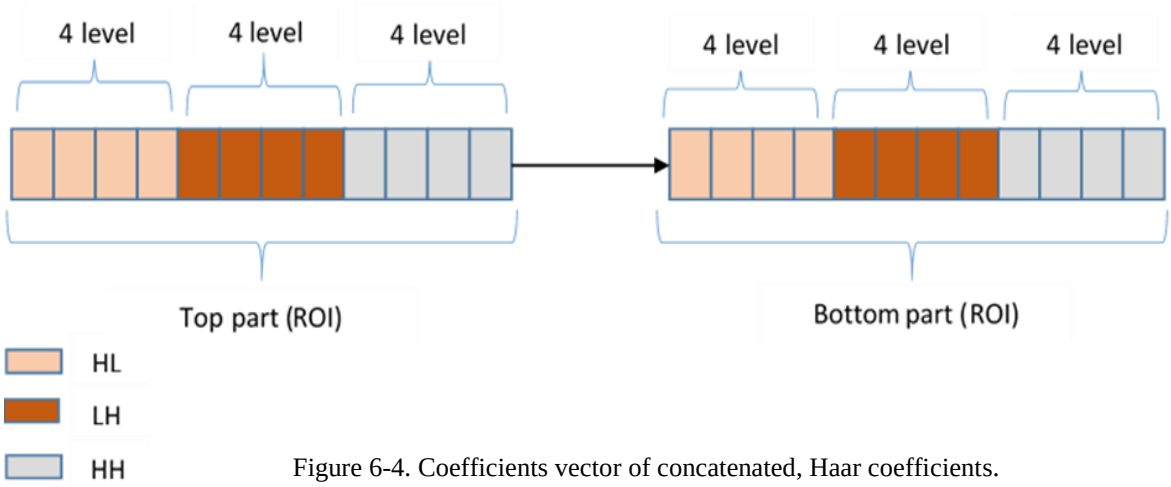


Figure 6-4. Coefficients vector of concatenated, Haar coefficients.

Algorithm 6-1 Wavelet transforms for human gait recognition using Haar wavelet.

GEI-based gait recognition using Haar wavelet features with SRKDA
Input: Silhouette images extracted over one gait cycle: S (x, y, t) ; t = 1, 2..., N Compute a GEI using Eq.2.1: $G_{(x, y)}$ Divide GEI horizontally GEI into 2 parts: $ROI_{(1)}$, $ROI_{(2)}$ as illustrated Figure 3-6. For each $ROI_{(i)}$, i = 1, 2. Compute detail Haar-wavelet features from *HL*, *LH* and *HH* sub-bands at different scales: $F_{(i)}$ Generate feature extraction set: $F = \{F_{(1)}, F_{(2)}\}$ Reduce features vector F by applying SRKDA technique Output: Reduced features set F

6.3 Experiment Results and Discussion

In this chapter, we used the CASIA database to evaluate our proposed method. This section describes the database that was used and the analysis of the results and compared them to existing and similar state-of-the-art methods.

6.3.1 Database and Evaluation Criteria

The proposed have been evaluated using the CASIA gait database B. In the experiment, the gallery set used for the CASIA dataset consisted of the first four sequences of each subject in CASIA Set-C (CASIA Set-C1). The probe was the remainder of the sequences in CASIA Set-

C (CASIA Set-C2), CASIA Set-A and CASIA Set-B. For evaluation criteria, a *k*-NN classifier was used to quantitatively evaluate the classification performance.

6.3.2 Analysis of the Results

Table 6-1 shows the multilevel decomposition of a GEI using the Haar transform functions. It is observed that beyond level 4 the increase in the recognition performance is not significant. Therefore, 4-level decomposition was chosen for further experimental evaluation.

Table 6-1. Comparison of various decomposition using horizontal wavelet with SRKDA.

Fusion of Levels of Decomposition	Covariates			Mean IR rank-1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Level 1	95.16	77.41	89.11	87.22
Fusion Level1 and Level 2	96.78	84.67	92.33	91.26
Fusion Level to Level3	96.78	87.09	93.54	92.47
Fusion Level1 to Level 4	96.78	87.50	93.55	92.62
Fusion of 5 Levels	97.17	87.09	93.54	92.60

Table 6-2 depicts the results obtained using the Haar wavelet decomposition with four levels. By analysing the results shown in the table, we can see that the local feature extraction using the *HL* coefficients allow the achievement of the highest recognition performance when compared against their *LH* and *HH* coefficients counterparts. In addition, the *HL* decomposition approach produces a performance increase of 7.22% when compared against a combined (*HL*+*LH*+*HH*) wavelet. This might be due to the fact that the human movement in GEI is horizontal, and so, the horizontal band is the most suitable band for characterising this movement. To elucidate the results that were obtained in table 6-2, a compression between all the previous mentioned four bands is shown in figure 6-5; where the best resolution of the human movement is obtained when *HL* detail is considered.

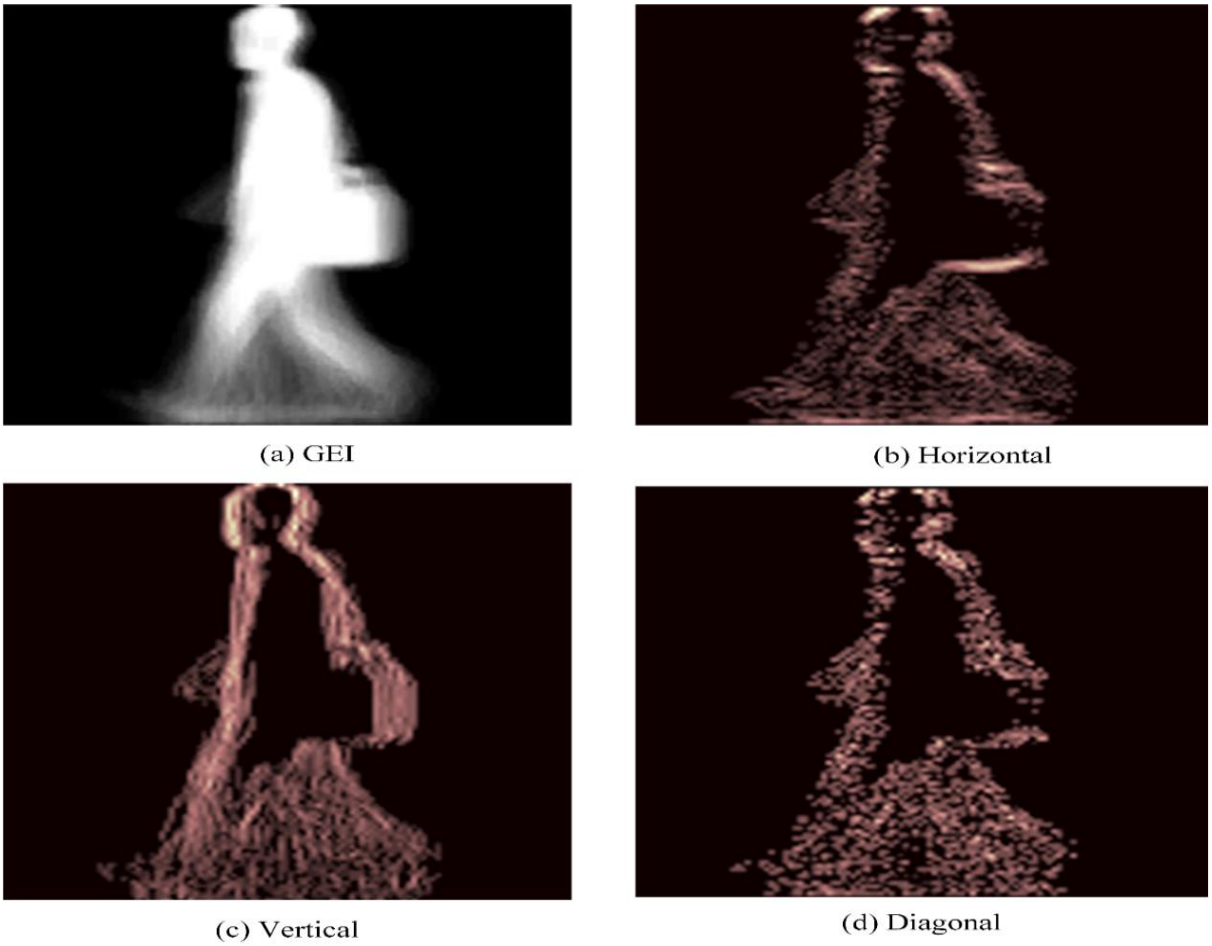


Figure 6- 5. A sample of level 1 decomposition with different bands.

Table 6-2. Comparison of IR rank-1 (in %) from the proposed method based on local feature computation techniques on the CASIA database, using a side view of 90°.

Features computation technique	Covariates			Mean IR rank-1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Local feature extraction using (*HL*) wavelet with SRKDA	96.78	87.50	93.55	92.61
Local feature extraction using (*LH*) wavelet with SRKDA	90.32	47.17	81.85	73.11
Local feature extraction using (*HH*) wavelet with SRKDA	79.84	33.47	66.53	59.95
Fusion of Local feature extraction using (*HL+LH*) wavelet with SRKDA	97.17	71.37	92.74	87.09
Fusion of Local feature extraction using (*HL+HH*) wavelet with SRKDA	95.16	79.03	93.14	89.11
Fusion of Local feature extraction using (*LH+HH*) wavelet with SRKDA	90.32	83.87	49.19	74.46
Fusion of Local feature extraction using (*HL+LH+HH*) wavelet with SRKDA	95.96	67.74	92.47	85.39

Tables 6-3, 6-4, 6-5 and table 6-6 show the experiment results in terms of IR at rank-1 (%) for based on our method. This method was also compared to recent and comparable state-of-the-art methods in (Khalid Bashir et al., 2009), (Bashir et al., 2010), (Hu et al., 2013), (Dupuis et al., 2013) and (Ait O. Lishani et al., 2017).

Table 6-3 compares the proposed method based on global and local feature computation techniques. From this table, we can notice that the use of horizontal dynamic areas in the GEI in the proposed method significantly improves the performance of the gait recognition system by up to 10.00% without SRKDA and 4.00% with SRKDA, compared to the use of the whole GEI. This can be explained by noting that wearing a coat results in covering nearly 2/3 of the body, thereby reducing the variations in the features extracted from the covered part.

Table 6-3. Comparison of IR rank-1 (in %) from the proposed method based on local and global feature computation techniques on the CASIA database, using a side view of 90°. Only the horizontal detail wavelet coefficients are used in the proposed method.

Features computation technique	Covariates			Mean IR rank-1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Global using only wavelet	79.98	67.33	45.96	70.42
Global using wavelet with SRKDA	98.38	86.29	80.64	88.43
Proposed method:				
Local using only wavelet	94.35	64.91	83.06	80.77
Local using wavelet with SRKDA	96.78	87.50	93.55	92.61

Table 6-4 shows that the proposed method using only the wavelet coefficient from *HL* sub-bands outperforms the state-of-the-art methods considered in this study. Indeed, the proposed method achieves a better IR at rank-1 by up to 93.00% and also increases the gait recognition for “wearing a coat” and “carrying a bag” conditions by up to 19.00% and 8.00% respectively, compared to the recent method in (Ait O. Lishani et al., 2017).

Table 6-4. Comparative analysis of the proposed method with several different state-of-the-art methods on the CASIA database for a side view of 90°. Three covariates are considered here: normal walking, carrying a bag, and wearing a coat, for horizontal, components.

Method	Covariates			Mean IR rank- 1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
(Khalid Bashir et al., 2009)	97.50	83.60	48.80	76.60
(Bashir et al., 2010)	100.00	78.30	44.00	74.10
(Hu et al., 2013)	94.00	45.20	42.90	60.70
(Dupuis et al., 2013)	97.60	73.80	62.50	77.96
(Ait O. Lishani et al., 2017)	85.36	79.90	74.74	80.00
Wavelet (only HL sub-bands) +SRKDA	96.78	87.50	93.55	92.61

Table 6-5 shows that the proposed method yields comparable results for "normal walking" and provides the best IR at rank-1 produce the Haar transform with SRKDA from a 108° angle. In the case of "wearing a coat" the Haar transform with SRKDA provides the best IR at rank-1 from an angle of 90°, with an IR of up to 93.55%.

The results show that accuracy is increased by up to 10.58% compared with the Haar transform without SRKDA. Furthermore, the wavelet transform with SRKDA provides the best IR at rank-1 for the case of "carrying a bag" for an angle of 90°, as the IR results improve by up to 10.49%. Finally, the "Normal walking" condition provides the best IR at rank-1 using the wavelet transform with SRKDA from an angle of 108°, where the IR improves by up to 97.17%. Overall, the SRKDA reduction achieves improved results by up to 10.00% in all angles. Our investigations have extended the viewing angles in the dynamic areas in the human body to select the best viewing angles.

A summary of the tables and results can be seen in table 6-6 for the Haar wavelet transform using SRKDA reduction technique able to achieve an attractive IR at rank-1 result of between 87.00% and 92.61% for different viewing angles.

Table 6-5. Comparison of IR (in %) from the proposed methods on the CASIA database from four side views: 36°, 72°, 90° and 108°. The proposed features, Wavelet transform with and without SRKDA. Only the horizontal detail wavelet coefficients were taken in the proposed method. Three covariates are considered here: normal walking, carrying a bag and wearing a coat.

Ranks	Method	Angle view	Covariates			Mean IR rank-1 (%)
			Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
Rank-1	Local using only wavelet	36°	96.77	76.20	66.93	79.96
		72°	94.35	74.19	82.66	83.73
		90°	94.35	69.92	83.06	80.77
		108°	93.95	66.13	78.63	79.57
Rank-1	Local using wavelet with SRKDA	36°	95.96	82.66	83.87	87.49
		72°	96.37	80.64	92.74	89.91
		90°	96.78	87.50	93.55	92.61
		108°	97.17	85.08	90.32	90.85

Table 6-6. Summary of IR (in %) from the proposed methods on the CASIA database from four side views: 36°, 72°, 90° and 108°. The proposed features and Wavelet transform with and without SRKDA. Only the horizontal detail wavelet coefficients were taken in the proposed method.

Ranks	Angle view	Method	
		Local using only wavelet (%)	Local using wavelet with SRKDA (%)
Rank-1	36°	79.96	87.49
	72°	83.73	89.91
	90°	80.77	92.61
	108°	79.57	90.85

6.4 Summary

This chapter has proposed a novel gait recognition approach for human identification under different clothing and carrying conditions from different viewing angles. The proposed method employs a supervised feature extraction technique based on Haar wavelet coefficients, which are extracted from the dynamic areas in the GEI, with SRKDA used to select the useful and informative features for classification. The experimental study conducted on the CASIA gait

database under various clothing and carrying conditions for different viewing angles compared the performance of the proposed method with recent and similar state-of-the-art methods. The experimental results using the k-NN classifier yielded an attractive performance of up to 93.00% with regard to rank-1 IR.

It can therefore be concluded that the wavelet transform is a very powerful technique, which offers a high accuracy rate and a low computation time. In our future work, we will evaluate our proposed method using different databases to extend the use of wavelets as a feature extraction method for gait recognition and to further investigate their performance.

CHAPTER SEVEN: CONCLUSION AND FUTURE WORK

Gait recognition has become one of the most important and effective behavioural biometric modalities for identifying human subjects unobtrusively at a distance using low-resolution video sequences. However, the variation view and clothing of a subject and the presence of carried items are the main challenges.

The underlying motivation of this thesis is to enhance the performance of existing methods under variable covariate conditions across large view angle changes. Applications of the results of this investigation include multi-modal biometric systems (e.g. face and gait) and human tagging across multiple cameras, where gait can be used on its own or as a cue to enhance the performance of existing methods. There is a considerable scope for performance improvement of gait recognition under variable covariate conditions. For this purpose, the key areas explored in this thesis are the development of a more discriminative feature extraction of gait against covariate conditions and cross view gait recognition. A number of interesting features of the proposed algorithms have been described and the feature methods have been shown to be effective and robust for feature extraction and for selection or reduction of relevant features from the GEI.

This chapter provides the conclusions of this thesis and summarises its contributions, along with making some suggestions for future work. The main conclusion of the thesis and an outline of its contributions are outlined in Section 7.1. The final suggestions for the future research are given in Section 7.2.

7.1 Summary of Contributions

This thesis has proposed a number of feature extraction methods for gait recognition under clothing and carrying conditions for different viewing angles. A GEI representation was chosen in this investigation by focusing on the dynamic parts of GEI data, which appear as gait cycle and head movement. The following summarises the main contributions of the thesis.

- The first contribution was the development of a human gait recognition technique based on Haralick features extracted from GEI. These features are extracted locally by dividing vertically and horizontally the GEI into two or three ROIs. The RELIEF feature selection algorithm is then employed on the extracted features in order to select only the most relevant features with a minimum redundancy.
- The second and third contributions to human identification by the proposed methods rely on two feature extraction techniques based on multi-scale feature descriptors and Gabor filter bank through the SRKDA reduction algorithm. The proposed features are extracted locally from two (ROIs) representing the dynamic areas in the GEI. The results are evaluated on the CASIA and USF Gait databases and the experimental results using the k-NN classifier have produced remarkable results of the highest identification rate at rank-1 when compared to the similar and recent state-of-the-art methods.
- The fourth contribution consisted of a human gait recognition technique based on detail wavelet features extracted from the Haar wavelet decomposition of dynamic areas in the GEI. The results are evaluated on CASIA Gait Database B under variations of clothing and carrying conditions for different viewing angles. The experimental results using k-NN classifier have yielded significant results of highest Identification Rate (IR) at rank-1 when compared to existing and similar state-of-the-art methods.

7.2 Future Work

To further improve the performance of gait recognition under various conditions, the following future work directions can be recommended:

- Investigate the proposed approach under other covariate factors such as walking surface (Guan et al., 2015) in order to study other features capable of improving the performance of our proposed approaches. The first step in this approach will require our efficient background removal to be further developed to cope with such variation.

- Consider and model more parts of the body and extract additional structural information by expanding GEI into the concept of the Structural Gait Energy Image. For example, a new gait recognition approach using SGEI has been proposed in (Li et al., 2013) which is generated by a fusion of a foot energy image (FEI) and head energy image (HEI). The FEI and HEI contain the moving probability information of the foot and head respectively without covering another part of the body. This is primarily related to the challenge of identification humans captured at a distance.

- Other gait data representations such as Gait Depth Energy Image (GDEI), partial GDEI, Discrete Cosine Transform (DCT) GDEI and partial DGDEI need to be evaluated and compared to existing methods. We plan to expand the database by recruiting more participants with balanced gender representation and more variety of unrestricted cases. Features similar to those will be used for the CASIA B.

Another area of future work relates to building/constructing an open-access database of Gait Images in order to experiment with these GEI-based representations i.e. GDEI, DGDEI, SGEI and GEI. This is widely anticipated, as there is a lack of an extended dataset containing various

types of data. The use of more search methods in the segmentation algorithms is another future direction that could be investigated.

The variance representation of energy image can be applied and investigated to evaluate the possibility of improving the identification rate.

The variance representation of energy image can be applied and investigated by carrying out work to evaluate the possibility of improving the identification rate.

REFERENCES

Abe, S. (2005). *Support Vector Machines for Pattern Classification* (Vol. 2). London: Springer.

Ahonen, T., et al. (2006). *Face Description with Local Binary Patterns: Application to Face Recognition*. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(12), 2037-2041. doi:10.1109/TPAMI.2006.244.

Anderson, R. J. (2008). *Security Engineering: A Guide to Building Dependable Distributed Systems* (2nd Edition): Wiley.

Anguita, D., et al. (2005, 31 July-4 Aug. 2005). *K-Fold Generalization Capability Assessment for Support Vector Classifiers.* Paper presented at the Proceedings. IEEE International Joint Conference on Neural Networks., Montreal, Que., Canada.

Arai, K., et al. (2012, 16-18 April 2012). *Gait Recognition Method Based on Wavelet Transformation and its Evaluation with Chinese Academy of Sciences (CASIA) Gait Database as a Human Gait Recognition Dataset.* Paper presented at the 2012 Ninth International Conference on Information Technology - New Generations, Las Vegas, NV, USA.

Arora, P., et al. (2016). *Parametric Curve Based Human Gait Recognition Information Systems Design and Intelligent Applications* (pp. 367-375): Springer.

Asheer Kasar Bachoo, J.-R. T. (2005, 2005). *Texture Detection for Segmentation of Iris Images.* Paper presented at the Annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries, White River, South Africa.

Bashir, K., et al. (2009, 3-3 Dec. 2009). *Gait Recognition Using Gait Entropy Image.* Paper presented at the 3rd International Conference on Imaging for Crime Detection and Prevention (ICDP 2009), London, UK.

Bashir, K., et al. (2010). *Gait Recognition Without Subject Cooperation*. *Pattern Recognition Letters*, 31 (13), 2052-2060.

Bashir, K., et al. (2009). *Gait Representation Using Flow Fields.* Paper presented at the Proceedings of the British Machine Vision Conference London, UK.

Baudat, G., et al. (2000). *Generalised Discriminant Analysis Using a Kernel Approach*. *Neural Computation*, 12 (10), 2385-2404.

Belkin, M., et al. (2002). *Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering.* Paper presented at the Advances in neural information processing systems, British Columbia, Canada.

Ben Abdelkader, C., et al. (2002, May 2002). *Stride and cadence as a biometric in automatic person identification and verification.* Paper presented at the Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, Washington, DC, USA.

Ben Abdelkader, C., et al. (2004). *Gait Recognition Using Image Self-Similarity*. *EURASIP Journal on Advances in Signal Processing*, 2004 (4), 721765.

Bernhard Schölkopf, et al. (1998). *Nonlinear component analysis as a kernel eigenvalue problem*. *Neural Computation*, 10 (5), 1299-1319.

Blum, A. L., et al. (1997). *Selection of relevant of features and examples in machine learning*. *Artificial Intelligence*, *97*(1-2), 245-271. doi:http://dx.doi.org/10.1016/S0004-3702(97)00063-5.

Boyd, J. E., et al. (2005). *Biometric Gait Recognition Advanced Studies in Biometrics* (Vol. 3161, pp. 19-42): Springer.

Cai, D., et al. (2007, 28-31 Oct. 2007). *Efficient Kernel Discriminant Analysis via Spectral Regression.* Paper presented at the Seventh IEEE International Conference on Data Mining (ICDM 2007), Omaha, NE, USA.

CBSR. (2005). *CASIA Gait Database*. Retrieved from: http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp.

Chaurasia, P., et al. (2017). *Fusion of Random Walk and Discrete Fourier Spectrum Methods for Gait Recognition*. *IEEE Transactions on Human-Machine Systems*, 47 (6), 751-762. doi:10.1109/THMS.2017.2706658.

Chengjun Liu, H. W. (2002). *Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition*. *IEEE Transactions on Image processing*, 11 (4), 467-476. doi:10.1109/TIP.2002.999679.

Chew-Yean Yam, M. S. N. (2009). *Model-based Gait Recognition Enclycopedia of Biometrics* (pp. 633-639). University of Southampton, Southampton: Springer.

Chung, F. R. (1997). *Spectral Graph Theory*: American Mathematical Soc.

Cristianini, N., et al. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*: Cambridge University Press.

Cross, G. R., et al. (1983). *Markov random field texture models*. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1), 25-39.

Cunado, D., et al. (2003). *Automatic extraction and description of human gait models for recognition purposes*. *Computer Vision and Image Understanding*, 90 (1), 1-41.

Dalal, N., et al. (2005, 25-25 June 2005). *Histograms of Oriented Gradients for Human Detection.* Paper presented at the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA.

Daugman, J. G. (1985). *Uncertainty relation for resolution in space, spatial frequency, and orientation optimised by two-dimensional visual cortical filters*. *Journal of the Optical Society of America A*, 2 (7), 1160-1169.

Deng, K., et al. (2006). *New Algorithms for Optimizing Multi-Class Classifiers via ROC Surfaces.* Paper presented at the Proceedings of the ICML workshop on ROC analysis in machine learning.

Derpanis, K. G. (2007). *Gabor Filters*, 1-5. Retrieved from [www.cs.yorku.ca].

Dobrovidov, A. V., et al. (2013). *Nonparametric Gamma Kernel Estimators of Density Derivatives on Positive Semi-axis*. *IFAC Proceedings Volumes*, 46 (9), 910-915. doi:http://dx.doi.org/10.3182/20130619-3-RU-3018.00214.

Dupuis, Y., et al. (2013). *Feature Subset Selection Applied to Model-free Gait Recognition*. *Image and Vision Computing*, 31 (8), 580-591. doi:http://dx.doi.org/10.1016/j.imavis.2013.04.001

Dy, J. G., et al. (2004). *Feature selection for unsupervised learning*. *Journal of Machine Learning Research*, 5 (Aug), 845-889.

Fischer, S., et al. (2007). *Self-Invertible 2D Log-Gabor Wavelets*. *International Journal of Computer Vision*, 75 (2), 231-246. doi:10.1007/s11263-006-0026-8.

Foster, J. P., et al. (2003). *Automatic gait recognition using area-based metrics*. *Pattern Recognition Letters*, 24 (14), 2489-2497.

Gabor, D. (1946). *Theory of communication. Part 1: The analysis of information*. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93 (26), 429-441.

Gafurov, D. (2007). *A Survey of Biometric Gait Recognition: Approaches, Security and Challenges.* Paper presented at the Annual Norwegian computer science conference, Oslo, Norway.

Gang Wu, E. Y. C. a. N. P. (2005). *Formulating context-dependent similarity functions*. Paper presented at the MULTIMEDIA '05 Proceedings of the 13th annual ACM international conference on Multimedia Singapore.

Gonzalez, R. C., et al. (2002). *Digital Image Processing* (2nd Edition): Prentice Hall.

Guan, Y., et al. (2015). *On Reducing the Effect of Covariate Factors in Gait Recognition: A Classifier Ensemble Method*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37 (7), 1521-1528. doi:10.1109/TPAMI.2014.2366766.

Gunn, S. R. (1998). *Support Vector Machines for Classification and Regression*. Retrieved from University of Southampton.

Guoying, Z., et al. (2006, 2-6 April 2006). *3D gait recognition using multiple cameras.* Paper presented at the 7th International Conference on Automatic Face and Gesture Recognition (FGR06).

Guru, V. G. M., et al. (2016, 29-30 April 2016). *Human Gait Recognition Using Four Directional Variations of Gradient Gait Energy Image* Paper presented at the 2016 International Conference on Computing, Communication and Automation (ICCCA), Noida, India.

Haiping, L., et al. (2006, 2-6 April 2006). *A layered deformable model for gait analysis.* Paper presented at the 7th International Conference on Automatic Face and Gesture Recognition (FGR06).

Haralick, R. M. (1979). *Statistical and structural approaches to texture*. *Proceedings of the IEEE*, 67 (5), 786-804.

Haralick, R. M., et al. (1973). *Textural Features for Image Classification*. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*(6), 610-621. doi:10.1109/TSMC.1973.4309314.

Hayfron-Acquah, J. B., et al. (2003). *Automatic gait recognition by symmetry analysis*. *Pattern Recognition Letters*, 24 (13), 2175-2183. doi:http://dx.doi.org/10.1016/S0167-8655(03)00086-2.

He, X., et al. (2008). *Learning a Maximum Margin Subspace for Image Retrieval*. *IEEE Transactions on Knowledge and Data Engineering*, 20 (2), 189-201.

He, X., et al. (2003). *Locality preserving projections*. *In Advances in Neural Information Processing Systems*, 16 (2003).

He, X., et al. (2005). *Face Recognition Using Laplacianfaces*. *IEEE transactions on pattern analysis and machine intelligence.*, 27 (3), 328-340.

He, Z.-Y., et al. (2008, 12-15 July 2008). *Activity Recognition from Acceleration Data Using AR Model Representation and SVM.* Paper presented at the *ACTIVITY RECOGNITION FROM ACCELERATION DATA USING AR MODEL REPRESENTATION AND SVM*, Kunming, China.

Heathrow. (2006). *Heathrow Testing Biometric Security Checks*. Retrieved from https://www.cnet.com/uk/news/heathrow-testing-biometric-security-checks/.

Hofmann, M., et al. (2012, 23-27 Sept. 2012). *2.5D gait biometrics using the Depth Gradient Histogram Energy Image.* Paper presented at the 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS).

Hofmann, M., et al. (2014). *The TUM Gait from Audio, Image and Depth (GAID) Database: Multimodal Recognition of Subjects and Traits*. *Journal of Visual Communication and Image Representation*, 25 (1), 195-206. doi:http://dx.doi.org/10.1016/j.jvcir.2013.02.006

Hu, M., et al. (2013). *Incremental Learning for Video-based Gait Recognition with LBP Flow*. *IEEE transactions on cybernetics*, *43*(1), 77-89.

Jahromi, O. S., et al. (2003). *Algebraic Theory of Optimal Filterbanks*. *IEEE Transactions on Signal Processing*, 51 (2), 442-457.

Jain, A., et al. (2007). *Handbook of Biometrics*: Springer Science & Business Media.

Jain, A. K. (1989). *Fundamentals of Digital Image Processing*: Prentice Hall.

Jain, A. K., et al. (1997). *Object detection using Gabor filters*. *Pattern Recognition,* 30 (2), 295-309.

Jain, A. K., et al. (2004). *An Introduction to Biometric Recognition*. *IEEE Transactions on Circuits and Systems for Video Technology,* 14 (1), 4-20. doi:10.1109/TCSVT.2003.818349.

Jolliffe, I. (2002). *Principal Component Analysis* (B. S. E. D. C. Howell Ed.): Wiley Online Library.

Ju, H., et al. (2006). *Individual Recognition Using Gait Energy Image*. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 28 (2), 316-322. doi:10.1109/TPAMI.2006.38.

Kecman, V., et al. (2006). *Support Vector Machines for Pattern Classification*. *JSTOR,* 48 (2), 418-421.

Kira, K., et al. (1992). *The feature selection problem: Traditional methods and a new algorithm.* Paper presented at the AAAI.

Kocsor, A., et al. (2004). *Margin Maximising Discriminant Analysis.* Paper presented at the European Conference on Machine Learning, Pisa, Italy.

Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection.* Paper presented at the International Joint Conference on Neural Networks (IJCNN).

Kohavi, R., et al. (1997). *Wrappers for feature subset selection*. *Artificial intelligence,* 97 (1), 273-324.

Kumar, H. P. M., et al. (2014, 13-14 Feb. 2014). *LBP for gait recognition: A symbolic approach based on GEI plus RBL of GEI.* Paper presented at the 2014 International Conference on Electronics and Communication Systems (ICECS).

Kusakunniran, W., et al. (2009, Sept. 27 2009-Oct. 4 2009). *Multiple views gait recognition using View Transformation Model based on optimized Gait Energy Image.* Paper presented at the 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops.

Kusakunniran, W., et al. (2010, 13-18 June 2010). *Support Vector Regression for Multi-View Gait Recognition based on Local Motion Feature Selection.* Paper presented at the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA.

Lades, M., et al. (1993). *Distortion invariant object recognition in the dynamic link architecture*. *IEEE Transactions on Computers,* 42 (3), 300-311. doi:10.1109/12.210173.

Lee, S., et al. (2007, 17-22 June 2007). *Shape Variation-Based Frieze Pattern for Robust Gait Recognition.* Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA.

Li, X., et al. (2013). *Gait Recognition Based on Structural Gait Energy Image*. *Journal of Computational Information Systems,* 9 (1), 121-126.

Lishani, A. O., et al. (2014, 14-17 Dec. 2014). *Haralick Features for GEI-based Human Gait Recognition.* Paper presented at the 26th International Conference on Microelectronics (ICM), Doha, Qatar.

Lishani, A. O., et al. (2017). *Human Gait Recognition Based on Haralick Features*. *Signal, Image and Video Processing,* 11 (6), 1123-1130. doi:10.1007/s11760-017-1066-y.

Little, J., et al. (1995, 21-23 Nov 1995). *Describing motion for recognition.* Paper presented at the Proceedings of International Symposium on Computer Vision - ISCV.

Little, J., et al. (1998). *Recognising People by Their Gait: The Shape of Motion*. *Journal of Computer Vision Research,* 1 (2), 1-32.

Liu, C., et al. (2017). A new feature selection method based on a validity index of feature subset. *Pattern Recognition Letters,* 92 (Supplement C), 1-8. doi:https://doi.org/10.1016/j.patrec.2017.03.018.

Liu, L.-F., et al. (2009). *Survey of Gait Recognition*. Paper presented at the 5th International Conference on Intelligent Computing, Ulsan, South Korea.

Liu, Z., et al. (2004). *Toward understanding the limits of gait recognition.* Paper presented at the Defense and Security, Orlando, Florida, United States.

Mallat, S. (2008). *A Wavelet Tour of Signal Processing: The Sparse Way* (Third Edition): New York: Academic Press.

Manjunath, B., et al. (1991). *Unsupervised texture segmentation using Markov random field models*. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 13 (5), 478-482.

Marcos , H. P. (2017). *Gender Recognition from Face Images Using a Geometric Descriptor*. Paper presented at the International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada.

Marques, O. (2011). Image Processing Basics *Practical Image and Video Processing Using MATLAB®* (pp. 21-34): John Wiley & Sons, Inc.

Materka, A., et al. (1998). *Texture Analysis Methods–A Review*. Retrieved from Technical University of Lodz, Brussels.

Matey, J. R., et al. (2006). *Iris on the Move: Acquisition of Images for Iris Recognition in Less Constrained Environments*. *Proceedings of the IEEE,* 94 (11), 1936-1947. doi:10.1109/JPROC.2006.884091.

Mirmehdi, M. (2008). *Handbook of Texture Analysis*. London: Imperial College Press.

Mistry, K., et al. (2017). A Micro-GA Embedded PSO Feature Selection Approach to Intelligent Facial Emotion Recognition. *IEEE Transactions on Cybernetics*, 47 (6), 1496-1509. doi:10.1109/TCYB.2016.2549639.

More, S. A., et al. (2017). *Gait-based human recognition using partial wavelet coherence and phase features*. *Journal of King Saud University-Computer and Information Sciences*.

Morse, B. S. (1998, March 22, 2000). *Lecture 22: Texture*. Retrieved from http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/texture.pdf.

Munif Alotaibi, A. M. (2017). *Reducing Covariate Factors of Gait Recognition using feature selection and dictionary-based sparse Coding*. *Signal, Image and Video Processing*, 11 (6), 1131-1138.

Nandini, C., et al. (2011, 7-9 Sept. 2011). *Gait recognition by combining wavelets and geometrical features.* Paper presented at the International Conference on Intelligent Agent & Multi-Agent Systems, Chennai, India.

Ng, H., et al. (2011). *Human Identification Based on Extracted Gait Features International Journal of New Computer Architectures and their Applications (IJNCAA), 1*(2), 358-370.

Nixon Mark, A. A. (2008). *Feature Extraction and Image Processing* (2nd Edition): Academic Press of Elsevier. London, UK.

Nixon, M. S. (2002). *Southampton Human ID at a Distance database*. Retrieved from: http://www.gait.ecs.soton.ac.uk/database/index.php3.

Nixon, M. S., et al. (2012). *Feature Extraction & Image Processing for Computer Vision* (Third edition): Academic Press.

Nixon, M. S., et al. (2004, 17-19 May 2004). *Advances in Automatic Gait Recognition* Paper presented at the Proceedings Sixth IEEE International Conference on Automatic Face and Gesture Recognition., Seoul, South Korea.

Nixon, M. S., et al. (2006). *Automatic Recognition by Gait*. *Proceedings of the IEEE,* 94 (11), 2013-2024. doi:10.1109/JPROC.2006.886018.

Niyogi, S. A., et al. (1994, 21-23 Jun 1994). *Analyzing and recognizing walking figures in XYT.* Paper presented at the 1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition.

Ojala, T., et al. (1996). *A comparative study of texture measures with classification based on featured distributions*. *Pattern Recognition,* 29 (1), 51-59. doi:http://dx.doi.org/10.1016/0031-3203(95)00067-4.

Ojala, T., et al. (2002). *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. *IEEE transactions on pattern analysis and machine intelligence.,* 24 (7), 971-987.

Olowoyeye, A., et al. (2009). *Medical Volume Segmentation using Bank of Gabor Filters.* Paper presented at the Proceedings of the 2009 ACM symposium on Applied Computing, Honolulu, Hawaii, U.S.A.

Olszewski, R. T. (2001). *Generalised Feature Extraction for Structural Pattern Recognition in Time-Series Data.* (Doctoral thesis), Carnegie-Mellon University Pittsburgh PA.

Oppenheim, W. (1997). *Signals & Systems* (2nd Edition ed.). Upper Saddle River, N.J.: Prentice Hall; London: Prentice-Hall International, c1997.

Pandey, P., et al. (2016, 16-18 March 2016). *Classification Techniques for Big Data: A Survey.* Paper presented at the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India.

Papageorgiou, C., et al. (2000). *A Trainable System for Object Detection*. *International Journal of Computer Vision,* 38 (1), 15-33. doi:10.1023/a:1008162616689.

Pentland, A. P. (1984). *Fractal-based description of natural scenes*. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (6), 661-674.

Phillips, P. J., et al. (2003, 17-17 Oct. 2003). *Face recognition vendor test 2002.* Paper presented at the 2003 IEEE International SOI Conference. Proceedings (Cat. No.03CH37443).

Phillips, P. J., et al. (2002, 2002). *The gait identification challenge problem: data sets and baseline algorithm.* Paper presented at the Object recognition supported by user interaction for service robots.

Pietikäinen, M., et al. (2011). *Computer Vision Using Local Binary Patterns* (Vol. 40): Springer Science & Business Media.

Qurat-Ul-Ain, G. L., et al. (2010). *Classification and Segmentation of Brain Tumor using Texture Analysis Recent Advances In Artificial Intelligence, Knowledge Engineering And Data Bases*, 147-155.

Radu, P., et al. (2013, 9-11 Sept. 2013). *Optimising 2D Gabor Filters for Iris Recognition.* Paper presented at the 2013 Fourth International Conference on Emerging Security Technologies, Cambridge, UK.

Rafael C. Gonzalez, R. E. W. (2008). *Digital Image Processing* (G. D. a. T. Benfatti Ed. Third Edition ed.). Upper Saddle River, New Jersey Pearson Prentice Hall.

Rahati, S., et al. (2008). *Gait recognition using wavelet transform.* Paper presented at the Fifth International Conference on Information Technology: New Generations, 2008. ITNG 2008.

Rida, I., et al. (2016). *Gait Recognition Based on Modified Phase-only Correlation*. *Signal, Image and Video Processing,* 10 (3), 463-470. doi:10.1007/s11760-015-0766-4.

Saeys, Y., et al. (2007). *A Review of Feature Selection Techniques in Bioinformatics*. *Bioinformatics,* 23 (19), 2507-2517.

Sarkar, S., et al. (2005). The humanid gait challenge problem: Data sets, performance, and analysis. *IEEE transactions on pattern analysis and machine intelligence, 27*(2), 162-177.

Sarkar, S., et al. (2005). *The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis*. *The IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI).* 27 (2), 162-177. doi:10.1109/TPAMI.2005.39.

Sevilla, M. P. P. G. (2006). *Image Processing: Dealing with Texture*.

Shan, C., et al. (2009). *Facial Expression Recognition based on Local Binary Patterns: A Comprehensive Study*. *Image and Vision Computing, 27*(6), 803-816. doi:http://dx.doi.org/10.1016/j.imavis.2008.08.005.

Shiqi, Y., et al. (2006, 0-0 0). *A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition.* Paper presented at the 18th International Conference on Pattern Recognition (ICPR'06).

Srivastava, R. (2013). *Research Developments in Computer Vision and Image Processing: Methodologies and Applications: Methodologies and Applications*. USA: IGI Global.

Steven C. H. Hoi, M. R. L. a. E. Y. C. (2006). *Learning the Unified Kernel Machines for Classification*. Paper presented at the Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, Philadelphia, PA, USA.

Strickland, R. N. (2002). *Image-Processing Techniques for Tumor Detection*: CRC Press.

Strzelecki, M., et al. (1997). *Markov Random Fields as Models of Textured Biomedical Images*. Paper presented at the Proceedings. 20th National Conference Circuit Theory and Electronic Networks., Kolobrzeg, Poland.

Tao, D. (2008, 23-28 June 2008). *A robust identification approach to gait recognition.* Paper presented at the 2008 IEEE Conference on Computer Vision and Pattern Recognition.

Thangavel, K., et al. (2005). *Breast Cancer Detection Using Spectral Probable Feature on Thermography Images* Paper presented at the 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP), Zanjan, Iran.

Tong, K. (2010, 3 June 2010). *Wavelet Transform And Principal Component Analysis Based Feature Extraction*. Retrieved from https://sites.math.washington.edu/~morrow/336_10/papers/kent.pdf.

Trokielewicz, M., et al. (2016, 13-16 June 2016). *Post-Mortem Human Iris Recognition.* Paper presented at the 2016 International Conference on Biometrics (ICB), Halmstad, Sweden.

Tsang, I. W.-H., et al. (2008). *Large-Scale Maximum Margin Discriminant Analysis Using Core Vector Machines*. *IEEE Transactions on Neural Networks*, 19 (4), 610-624.

Tuceryan, M., et al. (1993). *Texture analysis*. *Handbook of pattern recognition and computer vision*, 2, 235-276.

Verlekar, T. T., et al. (2017). *View-invariant Gait Recognition System Using a Gait Energy Image Decomposition Method*. *IET Biometrics*, *6*(4), 299-306. doi:10.1049/iet-bmt. 2016.0118.

Wahba, G. (1990). *Spline models for observational data*: SIAM.

Wang, F., et al. (2011). *Unsupervised Large Margin Discriminative Projection*. *IEEE Transactions on Neural Networks*, 22 (9), 1446-1456. doi:10.1109/TNN.2011.2161772.

Wang, X., et al. (2017). *Gait Recognition Based on Gabor Wavelets and (2D)^{2}PCA*. *Multimedia Tools and Applications*. doi:10.1007/s11042-017-4903-7.

Wang, Z., et al. (2013). *Enhancing Kernel Maximum Margin Projection for Face Recognition*. *JOURNAL OF SOFTWAR*, 8 (3), 724-730. doi:10.4304/jsw.8.3.724-730.

Whytock, T., et al. (2014). *Dynamic Distance-based Shape Features for Gait Recognition*. *Journal of Mathematical Imaging and Vision*, 50 (3), 314-326.

Willsky, A. S. (1997). *Signals and systems*: Prentice Hall, second edition, ISBN 0-13-814757-4.

Wolpert, D. H. (1992). *Stacked generalisation*. *Neural Networks*, *5*(2), 241-259.

Yam, et al. (2015). *Gait Recognition, Model-based*. *Encyclopedia of Biometrics*, 799-805.

Yam, C., et al. (2004). *Automated person recognition by walking and running via model-based approaches*. *Pattern Recognition*, 37 (5), 1057-1072. doi:http://dx.doi.org/10.1016/j.patcog.2003.09.012.

Yu, S., et al. (2017). Invariant feature extraction for gait recognition using only one uniform model. *Neurocomputing*, 239 (Supplement C), 81-93. doi:https://doi.org/10.1016/j.neucom.2017.02.006.

Z Liu, et al. (2004, 23-26 Aug. 2004). *Simplest Representation Yet for Gait Recognition: Averaged Silhouette.* Paper presented at the Proceedings of the 17th International Conference on Pattern Recognition. ICPR Cambridge, UK.

Zhang, D., et al. (2012). *Rotation Invariant Curvelet Features for Region Based Image Retrieval*. *International Journal of Computer Vision*, 98 (2), 187-201.

Zhang, D., et al. (2000). *Content-based image retrieval using Gabor texture features.* Paper presented at the IEEE Pacific-Rim Conference on Multimedia (PCM'00), Fargo, ND, USA.

Zhang, R., et al. (2007). *Human Gait Recognition at Sagittal Plane*. *Image and Vision Computing*, 25 (3), 321-330.

Zhao, N., et al. (2016, 24-29 July 2016). *Sparse Tensor Discriminative Locality Alignment for Gait Recognition.* Paper presented at the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada.

APPENDIX

Appendix A

The Figure A-1 and Figure A-2 shows the weight and rank feature when have used RELIEF algorithm.

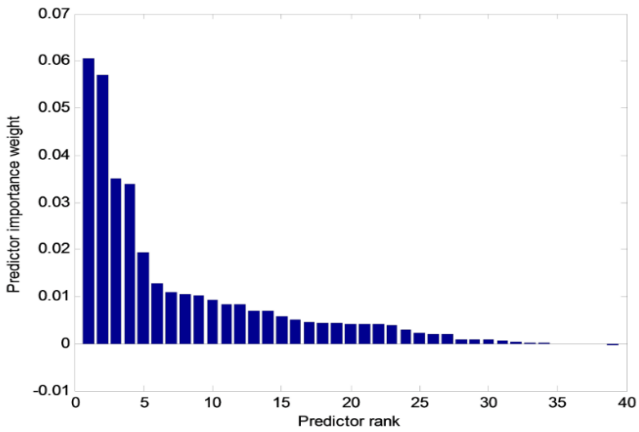


Figure A-1. Illustrates weight feature.

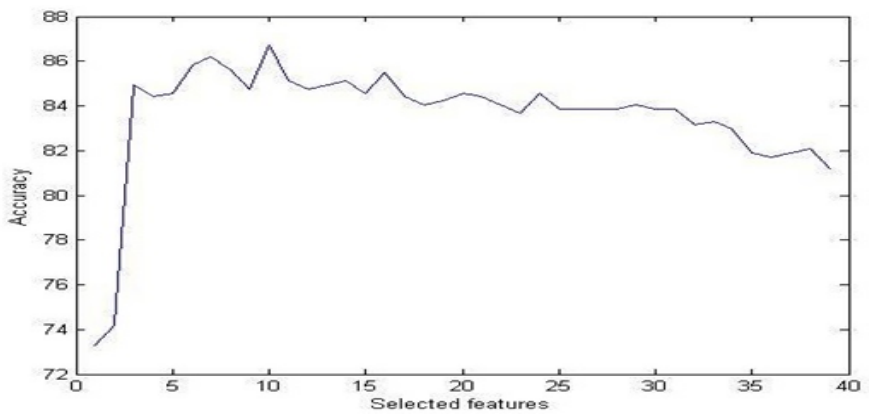
Table A-1 and Table A-2 shows the weights and ranked each feature that extracted by the Haralick method.

Table A-1. Weights for each feature from RELIEF method.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
.019	.002	.004	.005	.006	.035	.034	.005	.0042	.004	.004	0.01	0.007	.0105	.002	.00095	0.0083	0.001	0.057	0.06
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	
.008	.013	1E.03	.007	.009	.004	.0	8E.04	14E.05	8E0.4	2E.04	.004	.003	-1.69E.05	3E.04	5.50E.05	7.30E.07	.002	3E.04	

Table A-2. Ranked features.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
20	19	6	7	1	22	18	14	12	25	17	21	13	24	5	8	4	32	9	26
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
3	10	11	33	2	15	38	23	16	28	30	35	39	31	37	34	29	36	27	



Figures A-2. Shows ranking for the features.

Table A-3 and Table A-4 show the results for comparison of IR (in %) from the Haralick proposed method on CASIA database (dataset B) for different theta angle, side view of 90º, horizontal and vertical division.

Table A-3. Comparison of IR (in %) from the proposed method on CASIA database (dataset B) for different theta angle. Horizontal division.

theta	Covariates			Mean IR Rank-1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing cloth (%)	
0º	77.42	62.10	69.35	69.62
45^0	82.52	69.35	77.42	76.43
90^0	86.56	70.16	80.64	79.12
135^0	82.52	70.16	81.45	78.04
Fusion	84.95	72.58	80.46	79.33

Table A-4. Comparison of IR (in %) from the proposed method on CASIA database (dataset B) for different theta angle. Vertical division.

theta	Covariates			Mean IR Rank-1 (%)
	Normal walking (%)	Carrying a bag (%)	Wearing cloth (%)	
0^0	68.01	53.22	45.97	55.73
45^0	70.16	51.61	50.80	57.52
90^0	74.46	63.00	64.52	67.32
135^0	71.77	55.64	54.03	60.48
Fusion	78.50	69.35	67.00	71.61

Appendix B

Table B-1 , Table B-2 and Table B-3 shows the results for comparison of IR (in %) from the Gabor filter bank with different reduction techniques on CASIA database (dataset B) for different theta angle, side view 36º,72º,90º and 108º, with the following training-testing partitioning: 25% -75%, 50%-50% and 75%-25%. By analysing these results, we can notice that the proposed method achieves an attractive CCR result between 87.00 % and 91.00 % for different viewing angles.

Table B-1 Comparison of CCRs (in %) from the proposed Gabor filter via SRKDA reduction method on CASIA database for four side views: 36º, 72º, 90º and 108º. The selected data are split randomly into two parts: Training and testing with partitioning: 25 -75%, 50% -50% and 75% -25%.

Data partitioning (training–testing)	Angle view	Covariates			Mean IR Rank-1 (%)
		Normal walking (%)	Carrying a Bag (%)	Wearing a Coat (%)	
25 % - 75 %	36°	86.73	75.62	89.24	83.87
	72°	91.75	82.43	89.24	87.81
	90°	89.60	85.66	91.04	88.77
	108°	88.88	56.37	88.88	88.05
50 % - 50 %	36°	85.48	83.87	93.54	87.63
	72°	95.70	85.48	90.32	90.50
	90°	93.55	87.63	89.24	90.14
	108°	91.93	86.56	91.40	89.96
75 % - 25 %	36°	89.24	83.87	90.32	87.81
	72°	90.32	88.17	90.32	89.60
	90°	90.32	89.24	88.17	89.24
	108°	91.39	90.32	87.09	89.60

Table B-2 Comparison of CCRS (in %) from the proposed Gabor filter via KPCA reduction on CASIA Database for four side views: 36º, 72º, 90º and 108º. The selected data are split randomly into two parts: training and testing with the partitioning: 25 %- 75 %, 50 %- 50 % and 75 %- 25 %.

Data partitioning (training–testing)	Angle view	Covariates			Mean IR Rank-1 (%)
		Normal Walking (%)	Carrying a bag (%)	Wearing a coat (%)	
25 % - 75 %	36°	83.15	69.89	87.45	80.16
	72°	90.32	78.50	87.10	85.30
	90°	83.87	81.72	84.58	83.39
	108°	90.32	79.93	86.38	85.54
50 % - 50 %	36°	82.25	78.49	89.78	83.51
	72°	93.54	84.40	88.71	88.88
	90°	89.78	85.48	91.39	88.88
	108°	90.32	83.33	89.24	87.63
75 % - 25 %	36°	84.94	79.57	86.02	83.51
	72°	89.60	77.42	88.17	85.06
	90°	86.02	82.79	83.87	84.22
	108°	89.24	84.95	86.02	86.73

Table B-3 Comparison of CCRS (in %) from the proposed Gabor filter via MMP reduction on CASIA Database for four side views: 36º, 72º, 90º, and 108º. The selected data are split randomly into two parts: training and testing with the partitioning: 25 %- 75 %, 50 %- 50 % and 75 %- 25 %.

Data partitioning (training–testing)	Angle view	Covariates			Mean IR Rank-1 (%)
		Normal walking (%)	Carrying a bag (%)	Wearing a coat (%)	
25 % - 75 %	36°	86.37	79.92	88.88	85.06
	72°	89.60	81.36	88.88	86.61
	90°	90.32	85.30	90.68	88.77
	108°	88.88	84.94	89.60	87.81
50 % - 50 %	36°	85.48	84.40	91.93	87.27
	72°	94.62	85.48	90.32	90.14
	90°	92.47	89.24	90.32	90.68
	108°	91.93	86.55	91.39	89.96

75 % - 25 %	36°	89.24	84.94	89.24	87.81
	72°	89.24	88.17	88.17	88.53
	90°	88.17	90.32	87.09	88.53
	108°	91.39	90.32	89.24	90.32