UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

EXTENDING QUALITY AND COVARIATE ANALYSES FOR GAIT BIOMETRICS

by

Darko S. Matovski

Thesis for the degree of Doctor of Engineering

January 2013

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL AND APPLIED SCIENCES

Electronics and Computer Science

Doctor of Engineering

EXTENDING QUALITY AND COVARIATE ANALYSES FOR GAIT BIOMETRICS

by Darko S. Matovski

Recognising humans by the way they walk has attracted a significant interest in recent years due to its potential use in a number of applications such as automated visual surveillance. Technologies utilising gait biometrics have the potential to provide safer society and improve quality of life. However, automated gait recognition is a very challenging research problem and some fundamental issues remain unsolved.

At the moment, gait recognition performs well only when samples acquired in similar conditions are matched. An operational automated gait recognition system does not yet exist. The primary aim of the research presented in this thesis is to understand the main challenges associated with deployment of gait recognition and to propose novel solutions to some of the most fundamental issues.

There has been lack of understanding of the effect of some subject dependent covariates on gait recognition performance. We have proposed a novel dataset that allows analyses of various covariates in a principled manner. The results of the database evaluation revealed that elapsed time does not affect recognition in the short to medium term, contrary to what other studies have concluded. The analyses show how other factors related to the subject affect recognition performance.

Only few gait recognition approaches have been validated in real world conditions. We have collected a new dataset at two realistic locations. Using the database we have shown that there are many environment related factors that can affect performance. The quality of silhouettes has been identified as one of the most important issues for translating gait recognition research to the 'real-world'. The existing quality algorithms proved insufficient and therefore we extended quality metrics and proposed new ways of improving signature quality and therefore performance.

A new fully working automated system has been implemented. Experiments using the system in 'real-world' conditions have revealed additional challenges not present when analysing datasets of fixed size.

In conclusion, the research has investigated many of the factors that affect current gait recognition algorithms and has presented novel approaches of dealing with some of the most important issues related to translating gait recognition to real-world environments.

Contents

List of Figures

List of tables

DECLARATION OF AUTHORSHIP

I, Darko S. Matovski, declare that the thesis entitled *Extending Quality and Covariate Analyses for Gait Biometrics* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:

A) D. S. Matovski, M. S. Nixon, S. Mahmoodi, and J. N. Carter, The Effect of Time on the Performance of Gait Biometrics, in *4th IEEE Int. Conf. Biometrics: Theory, Applications and Systems*, Washington D.C. USA, 2010.

B) D. S. Matovski, M. S. Nixon, J. N. Carter, Gait Recognition, *Encyclopedia of Computer Vision*, Springer, 2012, In Press

C) D.S. Matovski, M.S. Nixon, S. Mahmoodi, J.N. Carter, The Effect of Time on Gait Recognition Performance, *IEEE Transactions on Information Forensics and Security*, **7**(2), pp. 543-552, 2012

D) D. S. Matovski, M. S. Nixon, S. Mahmoodi, and T. Mansfield, On Including Quality in Applied Automatic Gait Recognition in *21st International Conference on Pattern Recognition (ICPR)*, Japan, 2012

Signed: ……………………………………………………………………..

Date:……………………………………………………………………….

Acknowledgements

I would like to thank my supervisors Mark Nixon, Tony Mansfield, Sasan Mahmoodi and John Carter for their support, guidance and encouragement throughout my time at the University of Southampton. I would also like to acknowledge the help of the school's senior tutor Eric Cooke, during my undergraduate and postgraduate studies.

I really appreciate the financial and other support I have received from the School of Electronics and Computer Science and The National Physical Laboratory.

I am especially grateful to my parents (Stojko and Petra) and my sister (Elena) for all their encouragement and support throughout my 10 years of education in the UK.

Finally, I would like to thank all my colleagues and friends at the University of Southampton for making my university experience fun, valuable and unforgettable.

Chapter 1

Context and Contributions

1.1 Context

Many aspects of life require a way of recognising individuals or verifying their identity. Examples range from financial transactions to border crossing. There are many situations in which identity needs to be determined at a distance, without the subject's cooperation or knowledge.

Biometric authentication refers to the identification of humans by their physical and behavioural characteristics. Biometric techniques are very attractive when compared with conventional methods of identification and verification; taking advantage of the unique characteristics of the human body or behaviour. Biometrics is typically used for access control and surveillance and therefore it can improve security and quality in many aspects of life.

Surveillance technology is ubiquitous in modern society. This technology has the potential to provide safer environment and help reduce and solve crime. Currently, surveillance systems rely on security operatives manually overseeing their operation. However, the enormous growth of security cameras makes it difficult and inefficient to use human operators. There is an enormous benefit of deploying biometric identification for automated visual surveillance systems.

The way a person walks (or runs) combined with their posture is known as gait. Recognising individuals by their particular gait using automated vision based algorithms is known as Gait Recognition.

Gait has the greatest potential to become the biometric of choice for surveillance systems because of its unique advantages. Gait can be acquired at a distance without subjects' cooperation in a non-invasive manner and it is available when other biometrics are inaccessible.

Recently, gait has been used as a forensic tool [1] and there have already been convictions based on the evidence produced by gait analysis. There are a number of situations in which gait is the only biometric available from a crime scene.

Automated recognition of humans by their gait is a very challenging research problem. Although solutions to many challenges have been proposed, some key issues remain. In addition, only a small number of approaches have been validated in real world environments. An operational automated gait recognition system does not yet exist. The work presented in this thesis brings gait recognition one step closer to deployment in the real world. This is achieved by examining the factors (also known as covariates) that are known to affect recognition in a more principled manner and propose solutions to some of the fundamental problems.

1.2 Main contributions

An automated, fully operation gait recognition system has never been designed and implemented before. The overall aim of this thesis is to understand the main challenges associated with deployment of gait recognition in the real world and to propose approaches and methods that can improve some of the most fundamental issues. The findings of this thesis facilitate the design and implementation of a system operating in a realistic environment. The most important high level requirements of the system are as follows:

- Ability to operate in challenging, real-world environments with minimum set-up requirements. The operation of the system should not be location specific.
- Ability to capture, process and display data in a fully automated manner.

- Ability to assess sample quality and use this information to improve recognition rates and/or provide indication of the reliability of a match.
- Ability to use feature selection (e.g. gait cycle selection) and soft biometrics (e.g. height) in order to improve recognition performance.
- Understand the limitations of the system and describing scenarios in which the system could operate.

The first step in realising an operational gait recognition system is to understand what the current challenges are and what solutions have been proposed in the literature to address those challenges. Understanding the limitation of existing techniques fail is vital for deployment in the real-world. Chapter 2 describes the current state of gait recognition. An adapted version of this work is due to be published as an article in the Encyclopaedia of Computer Vision [B]. This article represents the most up to date survey on gait recognition. A particular emphasis is placed current challenges in gait recognition and on related work trying to tackle those challenges. The majority of these challenges are covered in this thesis. Many of the proposed techniques fail in realistic environments. Chapter 5 reveals some of these situations and proposes a number of possible solutions.

The current publically available dataset are not suitable for performing certain types of analyses which can aid the implementation of an operational gait recognition system. Consequently, two different databases have been collected as part of the research for this thesis. Understating how subject related factors affect recognition performance is essential and therefore the first database has been collected in a controlled environment. It enables researchers, for the first time, to perform experiments by varying only a single subject-dependent covariate at a time. The second database has been collected at two different indoor locations. One of the indoor locations proved to be more challenging than a 'typical' outdoor location due to substantial illumination changes and reflective surfaces. The challenging environment in which this database was collected enabled us to understand which techniques work best in the presence of noise whilst limiting the effect of subject dependent factors. This database provides a great scope for further research in translating gait recognition to the real world. It allows a research to perform experiments related to the environment such as: change of view point, distance from camera, matching across different locations, matching in the presence of samples of varying/poor sample quality etc.

Both datasets have made it feasible to perform experiments that were previously not possible. The datasets will be made publically available to enable other researchers to repeat, extend and design new experiments.

Using the first dataset we were able to analyse a large number of subject dependent factors independently by limiting the influence of all others. The first principled study that investigates the effect of elapsed time and other factors on gait recognition performance is presented in Chapter 4 [A] [C]. This study concludes that elapsed time does not affect gait recognition performance in the short to medium term, to the extent that what was previously believed. It was concluded that significant change of clothing worn by subjects seems to be the major problem for current model-free gait recognition approaches. Results showing the effect of footwear, walking speed and distance from the camera are also presented in this thesis. The novel findings of this study and our novel database enable researchers to focus on solving subject dependent factors that affect gait recognition performance the most.

Analyses of gait in challenging environments when the video is affected by uneven and changing lighting showed that the quality of samples is one of the most fundamental issues for deployment of gait recognition in the real world. Many of the existing techniques fail in the presence of significant amount of noise. The existing quality assessment and improvement algorithms proved insufficient and therefore existing quality metrics were extended and novel ways of using them were proposed to improve signature quality and therefore recognition rates. The benefits are demonstrated experimentally using samples acquired in different conditions and at different locations [D]. In addition, the experiments revealed additional environmental related factors that can affect recognition, such as pose, camera height etc. All of these finding are presented in Chapter 5.

Many of the data processing and recognition techniques available in the literature worked well in the controlled environment. However, this was not the case for the data acquired in challenging environments. A significant amount of work has been done in understanding the reasons behind this and proposing new techniques or new combinations of existing techniques to improve matters. The findings are presented in Chapter 3.

The novel findings mentioned above facilitated the development of a new fully working automated gait recognition system. .The system was integrated with a 3D model able to show the previous and current location a person has visited. The purpose of the system was to track individuals as they move through a monitored area. The novelty of the system consists of the use of quality to improve signatures and therefore recognition rates. The system also uses quality to provide indication of the decision confidence. It was used to compare the results obtained by analysing the datasets and to reveal any additional challenges associated with performing gait recognition in the real world.

1.3 Thesis overview

This thesis is structured as follows:

- Chapter 2 – **The Current State of Gait Biometrics**
  In this chapter, the motivation for the use of gait as a biometric in the context of visual surveillance applications is presented. This is followed with description of some well-known recognition algorithms and their capabilities on existing datasets. The chapter describes some of the existing and potential applications of gait biometrics. Currently, there are numerous challenges associated with gait recognition, especially in challenging environments. These challenges are presented and the main focus is on areas relevant to the contribution of this thesis.

- Chapter 3 – **Datasets and Data Processing**
  In this chapter some of the most well-known gait datasets are listed and the two novel datasets that have been collected for the purposes of this thesis are described. This chapter presents details of the processing steps implemented to analyse the samples from both databases. The chapter provides an indication of which techniques are suitable for analysing samples acquired in realistic conditions.

- Chapter 4 – **Subject Dependent Covariates**
  This chapter presents the results obtained by analysing the database collected in a controlled environment (the Biometric Tunnel). In this chapter we show the effect of the following subject dependent factors on performance: elapsed time (aging), clothing, footwear and walking speed.

The experiments are performed in a more principled and controlled manner that has been done before. The majority of the results are consistent with the literature, except in the case of elapsed time.

- Chapter 5 – **Quality and Deployment**
  This chapter introduces quality for gait samples. The importance of sample quality is shown experimentally by using the database collected at two real-world locations. New quality metrics are proposed and a novel way of deploying those quality metrics is described. Additional environment dependent covariates are investigated and the understanding is used to define quality for gait recognition. Finally, results of performing experiments in challenging environment using an automated gait recognition system are presented.

- Chapter 6 – **Conclusions and Future Work**
  Overall conclusions are drawn in this chapter. At the end of the chapter, ideas and suggestions for future work are presented.

1.4 List of publications

So far the following papers have been published:

A) D. S. Matovski, M. S. Nixon, S. Mahmoodi, and J. N. Carter, The Effect of Time on the Performance of Gait Biometrics, in *4th IEEE Int. Conf. Biometrics: Theory, Applications and Systems (BTAS)*, Washington D.C. USA, 2010.

B) D. S. Matovski, M. S. Nixon, J. N. Carter, Gait Recognition, *Encyclopedia of Computer Vision*, Springer, 2012, In Press

C) D.S. Matovski, M.S. Nixon, S. Mahmoodi, J.N. Carter, The Effect of Time on Gait Recognition Performance, *IEEE Transactions on Information Forensics and Security*, **7**(2), pp. 543-552, 2012

D) D. S. Matovski, M. S. Nixon, S. Mahmoodi, and T. Mansfield, On Including Quality in Applied Automatic Gait Recognition, in *21st International Conference on Pattern Recognition (ICPR)*, Japan, 2012

Chapter 2

The Current State of Gait Biometrics

2.1 Introduction

Automated visual surveillance systems have attracted the interest of many researchers in the fields of computer vision and machine learning due to their potential to improve safety, security and quality of life. In the UK there are 4.2 million CCTV cameras, more per head than any other European country. However, the lack of automated recognition ability limits the effectiveness of current systems. Biometric recognition such as gait has the potential to enable a surveillance system to automatically track, recognise and verify subject of interest in real-time without their cooperation or knowledge.

In this chapter we introduce gait recognition as a biometric of choice for visual surveillance applications. This is followed by description of some of the most commonly used approaches for gait recognition. We elaborate on the current challenges and some of the proposed solutions to these challenges. This chapter covers the state of art mainly by reference and particularly focuses on areas relevant to the contributions of this thesis.

2.2 Gait for visual surveillance

Many studies have demonstrated that gait has the potential to become a powerful biometric for surveillance and access control [2-4]. There is a continuous growth in the number of surveillance systems but these systems have yet to include recognition capability. Surveillance system that uses automated recognition could play an important role in law enforcement.

Gait has few important advantages over other forms of biometric identification. It can be acquired at a distance when other biometrics are obscured or the resolution is insufficient. It does not require subject cooperation and can be acquired in a non-invasive manner. It is easy to observe and hard to disguise as walking is necessary for human mobility. Gait can be acquired from a single still image or from a temporal sequence of images (e.g. a video).

The primary aim of surveillance videos is to monitor people. However, the video data can be of a low quality (poor resolution, time lapse etc.) and the subject can try to conceal the more conventional biometrics. Nevertheless, such video can provide sufficient data for gait recognition technology and there is already research in using gait biometrics as a forensic tool [5]. Gait contains very rich information and is considered to be unique. Studies have shown that gait can also be used to reveal a person's identity, gender, emotional state etc.

Researchers have only recently started working on finding ways to deploy gait in surveillance and access control systems. However, research into gait has been undertaken long before this. References to gait can be found in: literature, medical research, biomechanics literature, psycho-physiological studies etc.

Shakespeare made several references to the individuality of gait, e.g. in *The Tempest* [Act 4 Scene 1], Cares observes "*High'st Queen of state, Great Juno comes; I know her by her gait*", in Henry IV Part II [Act 2, Scene 3] "*To seem like him: so that, in speech, in gait, in diet, in affections of delight, in military rules, humours of blood, he was the mark and glass, copy and book".*

The aim of medical research has been to classify the components of gait for the treatment of pathologically abnormal patients. Murray et al. created standard movement patters for pathologically normal people [6]. Those patterns were then used to identify pathologically abnormal patients.

The biomechanics literature makes observations concerning identity: "A given person will perform his or her walking pattern in a fairly repeatable and characteristic way, sufficiently unique that it is possible to recognize a person at a distance by their gait" [7].

Psycho-physiological studies such as [8] and [9] have shown that humans can recognise friends and the sex of a person solely by their gait with 70-80% accuracy. These and similar studies have inspired the use of gait as a biometric trait.

Recognition by gait is one of the newest biometrics, since its development only started when computer memory and processing speed became sufficient to process sequences of image data with reasonable performance. Although gait recognition is not sufficiently mature to be deployed in real word applications such as visual surveillance it has the potential to overcome most of the limitations of other biometrics and hence has attracted a vast interest in computer vision research.

2.3 Experimental results

The current state of the art achieves very high recognition rates (close to 100%) on relatively large databases (>300 subjects) when the training and test data are recorded under similar conditions. An example of progression in performance over time is shown in Table 2.1. The world's largest gait database of more than 1700 people has been constructed to enable statistically reliable performance evaluation of gait recognition performance [10]. Detailed description of other commonly used gait databases is presented in the next chapter.

Time Period	No of Subjects	Source	Recognition Rate	Notes
1990's	~ 10	USC	95.2	
2000's	~ 120	HiD, CASIA, Southampton	75% - 99%	Recognition rate depends on covariates
2010	>300	Southampton Multimodal	95% - 100%	Includes time dependent covariates
Recent (2011)	> 1700	Osaka University	≈ 90%	No Covariates

Table 2.1 - Progression of Gait Recognition Systems

2.3.1 Improvement evaluation

Evaluating and comparing different approaches is difficult without a standardised database. There are many studies focusing on solving particular confounding variables (covariates), but there is insufficient work in attempting to combine existing techniques and evaluate them on a standard database.

The HumanID gait challenge problem [11] was set up to outline a baseline algorithm for gait recognition and to enable effective comparison of different techniques. It proposes a number of difficult experiments for the existing gait matchers. The gallery set consists of 122 subjects walking on a grass surface recorded by a single camera. Table 2.2 shows the differences of the probe set compared to the gallery set for each of the challenge experiments.

The results in Figure 2.1 show the progress in gait recognition over a period of 2 years for the experiments shown in Table 2.2. 'Other 2002' refers to the performance of algorithms available prior to establishing the 'Baseline' algorithms. 'Best others 2004' refers to the novel algorithms proposed after the 'Baseline' was established. Recent studies achieve better recognition rates.

Experiment	Probe	# of Subjects	Difference
A	Different camera view than gallery	122	View
B	Subjects wore different shoes	54	Shoe
C	Different camera view and different shoes	54	Shoe, View
D	Subjects walked on a different surface	121	Surface
E	Different shoes and different walking surface	60	Surface, Shoes
F	Different walking surface and different camera view	121	Surface, View
G	Different walking surface, different shoes and different camera view	60	Surface, Shoe, View

Table 2.2 – Some experiments comprising the HumanID gait challenge problem

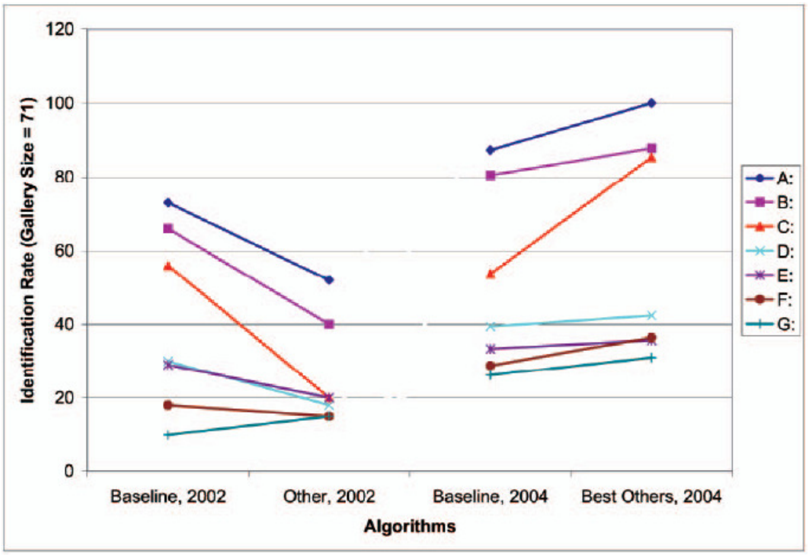


Figure 2.1 - The progress from the baseline over 2 years for the various experiments shown in Table 2.2 [11]

2.4 Approaches to gait recognition

The approaches to gait recognition can be divided in two main groups: *model-based* and *model-free* (see Table 2.3). Model-based approaches use the human body structure and model-free methods use the whole motion pattern of the human body. Which approach is adopted depends on the acquisition conditions and the quality of the data. Model-free (appearance-based) approaches use the input images directly to produce a gait signature without fitting a model. These approaches can perform recognition at lower resolutions which makes then suitable for outdoor applications, where a subject can be at a large distance from the camera. Model-based approaches typically require higher resolution images of a subject to be able to fit the model accurately. Our aim is to build an operational gait recognition system. Since model-free approaches perform better in the presence of noise a decision has been taken to use this type of approach throughout the thesis.

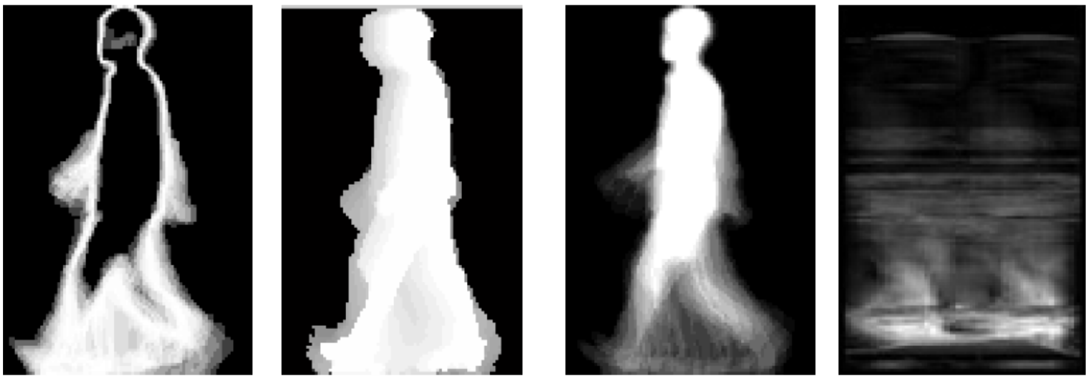
Model-Free Analysis		Model-based Analysis	
Moving Shape	Shape + Motion	Structural	Modelled
Unwrapped silhouette; silhouette similarity; relational statistics; self-similarity; key frame analysis; frieze patterns; area; symmetry; point distribution models; key poses	Eigenspace sequences; hidden Markov model; average silhouette; moments; ellipsoidal fits; kinematic features; gait style and content	Stride parameters; human parameters; joint trajectories	Articulated model; dual oscillator; linked feature trajectories

Table 2.3 - Approaches to Gait Recognition

The table is taken from [12] and [13]. Example papers for all of the approaches can be found in the original sources.

2.4.1 Model-Free approaches

The model-free approaches derive the human silhouette by separating the moving object from the background. The subject can then be recognized by measurements that reflect shape and/or movement. The simplest approach is to simply form an average of the silhouettes over a complete gait cycle [14]. The approach is called the Gait Energy Image (GEI) and it is shown in Figure 2.2. The Motion Silhouette Image (MSI) is a similar representation to the GEI. The value of each pixel is computed as a function of motion of that pixel in the temporal dimension over all silhouettes that are part of a single gait cycle. Both the GEI and MSI are straightforward to compute but they are vulnerable to appearance changes of the human silhouette. The Frieze pattern represents the information contained in a gait sequence by horizontal and vertical projections of the silhouettes. Its extension, SVB Frieze patterns use key frame subtraction in order to mitigate the effects of appearance changes on the silhouette (see Figure 2.2). The Gait Entropy Image (GEnI) is another example of a compact gait representation (signature) [3].



a) Gait Entropy b) Motion Silhouette c) Gait Energy d) SVB Frieze Pattern

Figure 2.2 - Examples of model-free gait signatures

The gait signatures for the approaches shown in Figure 2.2 are usually used directly for classification. There are additional ways of extracting gait signatures without using a model. Some examples are described below.

Little et al. derive a dense optical flow for each image sequence [15]. Scale-independent scalar features of each flow, based on moments of the moving point characterize the spatial distribution of the flow. The periodic structure of these sequences of scalars is analysed. The scalar sequences for an image sequence have the same fundamental period but differ in phase, which is used as a feature for recognition of individuals by the shape of their motion.

BenAbdelkader et al. [16] use background modelling to track the subject for a number of frames and extract a sequence of segmented images of the person. A self-similarity plot is computed via correlation of each pair of images in this sequence. For recognition, PCA (Principal Component Analysis) is used to reduce the dimensionality of the plots. A k-nearest neighbour rule is used on the reduced space for classification. Another silhouette based gait recognition technique using PCA has been proposed by [17]. Eigenspace transformation based on Principal Component Analysis (PCA) is applied to time-varying distance signals derived from a sequence of silhouette images to reduce the dimensionality of the input feature space. Supervised pattern classification techniques are performed in the lower-dimensional eigenspace for recognition.

Hayfron-Acquah et al. proposes a method for automatic gait recognition based on analysing the symmetry of human motion [18]. The Generalised Symmetry Operator is used to locate features according to their symmetrical properties rather than relying on the boarders of a shape. The symmetry operator is used on the optical flow image to produce a gait signature. For purposes of classification, the similarity differences between the Fourier descriptions of the gait signatures are calculated using Euclidean distance.

Gait is a temporal sequence and can be modelled using Hidden Markov Models (HMM). The statistical nature makes the model relatively robust. The postures that an individual adopts are regarded as states of the HMM and are typical to that individual and provide means of discrimination [19].

Kale et al. use two different image features to directly train a HMM: the width of the outer contour of a binary silhouette; and the entire binary silhouette itself [20].

2.4.2 Model-Based approaches

The advantages of the previous approaches (silhouette or features derived from it) are speed and simplicity. However, model-based approaches can be better at handling occlusion, scale and rotation. Model-based approaches require a high resolution and therefore are not very suitable for outdoor surveillance.

Model-based approaches incorporate knowledge of the shape and dynamics of the human body into the extraction process. These approaches extract features that fit a physical model of the human body. A gait model consists of shapes of various body parts and how those shapes move relative to each other (motion model). The shape model for a human subject can use ellipse to describe the head and the torso, quadrilaterals to describe the limbs and rectangles to describe the feet. Alternatively arbitrary shapes could be used to describe the edges of the body parts. The motion model describes the dynamics of the motion of the different body parts. Using a model ensures that only image data corresponding to allowable human shape and motion is extracted, reducing the effect of noise. The models can be 2 or 3 dimensional. Most of the current models are 2 dimensional and some examples are described below.

Yam et al. have used pendular motion and the understanding of biomechanics of human locomotion to develop two models: a bilateral symmetric and analytical model (employs the concept of forced couple oscillator). See Figure 2.3. The gait signature is the phase-weighted magnitude of the Fourier description of both the thigh and knee rotation [21].

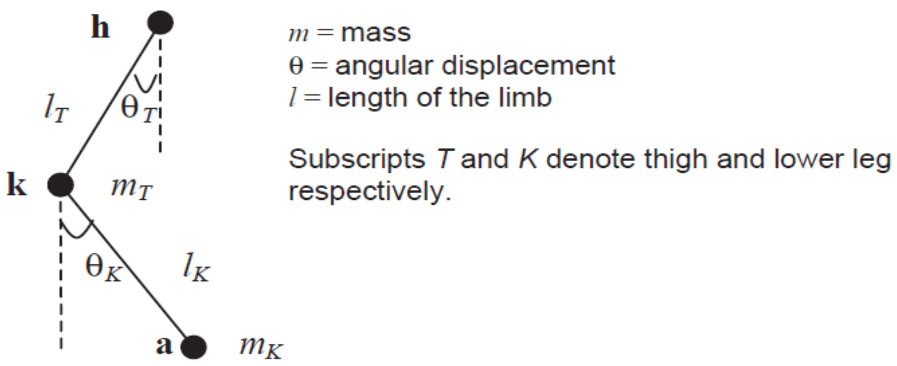


Figure 2.3 - Example of a gait model – The dynamically coupled pendulum [21]

Bouchrika et al. [22] have proposed a new approach to extract human joints. Spatial model templates for human motion are derived from the analysis of gait data collected from manual labelling. Motion templates describing the motion of the joints are parameterised using the elliptic Fourier descriptors. An example is shown in the equation below:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \begin{bmatrix} \cos(\alpha) - \sin(\alpha) \\ \sin(\alpha) \ \cos(\alpha) \end{bmatrix} \begin{bmatrix} X(t) * S_x \\ Y(t) * S_y \end{bmatrix} \tag{2.1}$$

where α is the rotation angle, s_x and s_y are the scaling factors across the horizontal and vertical axes respectively and *X(t)* and *Y(t)* are the Fourier summation. Hough transform is used in the feature extraction process.

Wang et al. have proposed an algorithm based upon the fusion of static and dynamic body information [23]. The static body information is in a form of a compact representation obtained by Procrustes shape analysis. The dynamic information is obtained by a model based approach which tracks the subject and recover joint-angle trajectories of lower limbs. A fusion at the decision level is used to improve recognition results. Figure 2.4 shows an example of the results obtained.

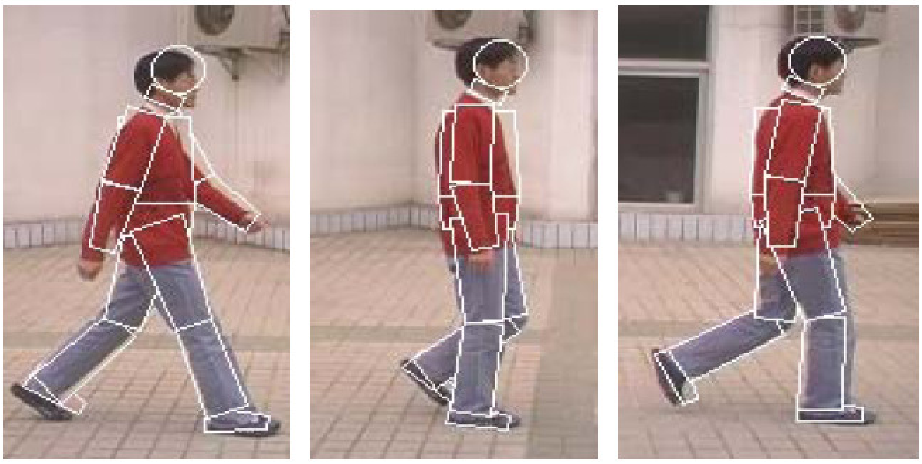


Figure 2.4 - Example of results [23]

There have been moves towards developing 3D gait models. Examples of work in this fields are [24] and [25]. These studies use video sequences from multiple cameras to construct 3D human models. The length of key segments is extracted as static parameters and the motion trajectories of lower limbs are used as dynamic features. Linear time normalisation is used for matching and recognition. 3D approaches are robust to changes in viewpoint and have a great potential [26].

However at present, experiments only on small databases are possible mainly due to high computational requirements.

Gait is dependent on large number of parameters (joint angles and body segment size) which leads to complex models with many free parameters. Finding the best fit model for a subject leads to searching a high-dimensional parameter space. Therefore, there is a trade-off between the accuracy of the model (complexity) and computational cost. The models are often simplified based on certain acceptable assumptions, e.g. a system could assume constant walking speed. However, as computing power increases the problems arising of high complexity can be mitigated.

2.5 Application

Gait research is currently at an evaluation stage rather than an application stage. However, the potential for gait recognition is great. The complete unobtrusiveness without any subject cooperation or contact for data acquisition makes gait particularly attractive for identification purposes. It could be used in applications including forensics, security, immigration, surveillance and access control.

Many surveillance systems capture only a low resolution video at varying lighting conditions, and gait recognition might be the only plausible choice for automatic recognition. A bank robber may wear a mask so you cannot see his face, wear gloves so you cannot get fingerprints, wear a hat so you cannot get DNA evidence – but they have to walk or run into the bank and they could be identified from their gait.

Recently, gait has been used as a forensic tool [1, 5] and there have already been convictions based on the evidence produced by gait analysis. There are a number of situations in which gait is the only biometric available from a crime scene.

Gait recognition has been used as evidence for conviction in some criminal cases. A man in Bolton (U.K) was convicted based on his distinctive gait. CCTV footage of the burglar captured near the crime scene was compared to a video captured at the police station by a podiatrist specialising in gait analysis. In 2004, a perpetrator robbed a bank in Denmark. The Institute of Forensic Medicine in Copenhagen was contacted by the police to perform gait analysis, as they thought the perpetrator had a unique gait. The institute instructed the police to establish a covert recording of the suspect from the same angles as the surveillance recordings for comparison. The gait analysis revealed several characteristic matches between the perpetrator and the suspect. For example, both the perpetrator (to the left) and the suspect showed inverted left ankle (white arrow) during left leg's stance phase and markedly outward rotated feet (see Figure 2.5).

The suspect was convicted of robbery and the court found that gait analysis is a very valuable tool [1].

Figure 2.5 - Bank Robbery

One system named the Biometric Tunnel [27] has led to the first live demonstration of gait as a biometric and could indicate a possible route for future deployment of the technology. Figure 2.6 depicts the system. It consists of a simple corridor with 12 synchronised and fixed cameras. The subjects are asked to walk through the middle and the lighting and background are controlled to facilitate analysis. The system is designed with a high throughput environment in mind.

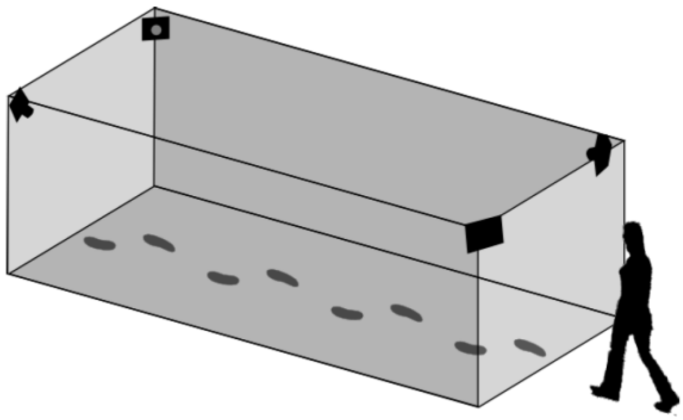


Figure 2.6 – A 3D visualisation of The Biometric Tunnel

2.6 Challenges and Solutions

Despite the recent outstanding achievements in computer vision and pattern recognition, there are many challenges to be overcome for the realisation of automated visual surveillance. Although a large number of gait recognition algorithms have been reported, the majority of studies achieve good recognition rates only when the gallery and probe have been acquired in similar conditions. It is very challenging to extract gait features that are invariant to change in appearance as well as to conditions that affect a person's gait. Examples of things that can change and negatively affect the effectiveness of current gait algorithms are: change of clothing, shoe type, carrying a load and injuries/medical conditions. Clothing for instance can change the observed pattern of motion and make it difficult to accurately locate joint position. Furthermore, there are certain factors that are related to the environment and not the subjects themselves that can cause difficulties for current gait algorithms. Examples of environmental confounding factors are: camera viewing angle, background and illumination. In addition, recognition performance can be affected by the conditions of acquisition and/or data quality. The degree to which a particular factor influences recognition depends on the algorithm deployed. In most real world scenarios, a good gait recognition algorithm should be able to deal with noisy data and unknown combination of covariate factors. Covariate factors are also known as exploratory variables and in this thesis covariates refer to changes in experimental setup (related to the subject or the environment) between two recording sessions.

In practical environments, there are a number of engineering requirements that need to be taken into consideration. The degree to which a particular challenge described below applies depends on the scenario of usage. For example, there can a scenario in which subject clothing is unlikely to change. In some cases using a model-based approach can help mitigate a number of issues such as change of clothing and change in viewing angle. However, using this type of approach might not be possible due to camera resolution or lack of multiple cameras.

The research community has been working very hard and a number of solutions for various problems have been proposed. The following sections discuss the challenges and present some of the recent work in dealing with those challenges. However, it is important to realise that many of these techniques cannot be directly applied to practical unknown environments (due to the underlying assumptions) as many studies do not have deployment focus.

2.6.1 Poor data quality

A video or image sequence can be considered of poor quality for a number of reasons. Examples include: occlusion, noise in forms of shadows and/or reflections and low resolution. Many existing approaches fail when performed on poor quality data. Therefore, it is essential for a practical system to be able to assess the quality of a sample and use this information to either improve recognition rates or provide indication in the confidence of the result.

Occlusion

A large number of current approaches assume little or no occlusion, at both the capturing and matching stages. Although, the assumption is unlikely to be valid in many real world environments the problem of occlusion is beyond the scope of this thesis as it requires solution to some unsolved computer vision challenges, such as background subtraction and object tracking.

Recently, there has been progress in solving the challenges related to occlusion. A study [28] has presented a new database (known as TUM-IITKGP) which allows a researcher to investigate different types of occlusion. This study presents the results based on two baseline approaches. Roy et al. proposed an improvement to the baseline [29]. The first step of the method is able to detect the presence of occlusion and to distinguish between clean and noisy gait cycles. In the second step, the silhouettes are corrected by what the study calls Balanced Gaussian Process Dynamic Model. The potential of the approach has been demonstrated on the TUM-IITKGP and a synthetic modification of the CMU's Mobo database.

Low resolution

The ability to recognise someone at a distance, at a low resolution is one of the key advantages of gait compared to all other biometrics. However, the recognition rate can be negatively affected in cases of insufficient resolution. Low resolution can occur because of two reasons: 1) large distance of subject to camera and 2) quality of acquisition device. Junping et al. [30] proposed a way of improving recognition rates in cases of low resolution samples. The approach first deploys a dimensionality reduction technique and then uses machine learning techniques to produce a high resolution image from the low resolution one. Tests on the HumanID and CASIA gait datasets demonstrate the potential of the approach.

Noisy or incomplete silhouettes

One of the first pre-processing steps for most gait recognition algorithms is separating or segmenting the subject from the background. The walking subject cannot always be subtracted perfectly even in cases of good quality video due to factors such as: colours of the subject and background, illumination (shadows and reflections), moving objects, occlusion etc. In complex environments, the binary silhouettes produced can contain noisy contours, holes, shadows, body parts missing or any combination of these. Low quality silhouettes need additional processing before they are suitable for recognition. The quality of the silhouettes has been shown to have a direct impact on the performance [31]. Yu et al. [32] used dynamic time warping (DTW) to reduce the effect of noise around the contours. The method improves the performance. A study by Chen et al. [33] described a new dynamic gait representation to minimise the effect of silhouette incompleteness. A study by Xue et al. [2] takes a completely different approach to the problem. It employs infrared thermal cameras to minimise the effect of the environment. A new database is presented and promising recognition results are achieved.

2.6.2 View-point changes

Many studies have shown that recognition performance can deteriorate when there is a significant difference in the view angle between the gallery and the probe. A system that operates in unconstrained scenarios is required for surveillance applications. In this case, subjects walk freely and usually only a single non-calibrated camera is available. The appearance change of the subject as a result of viewing angle can be considerable and most current recognition algorithms can have difficulties.

A study by Goffredo et al. [34] demonstrated 2D view-independent approach which can be used in situations where only a single non-calibrated camera is available. This algorithm allows reconstruction of the limbs' pose in the sagittal plane under the assumption that two gait periods are observed when the subject walks in a straight line. A research [4] proposed a method that works well even if the subject is walking in non-linear path. The method works by tracking the feet and head of the extracted silhouettes and then performs a homography transformation to make the walking plane appear as if walking was observed from a fronto-parallel view. The approach works without initialisation or camera calibration.

Kusakunniran et. al [35] proposed a novel approach which transforms gait features from one viewing angle to another. In other words, the probe sample is transformed to match the corresponding gallery sample in terms of the viewing angle.

If more than one calibrated camera is available 3D reconstruction is possible and the issues of view-point are somewhat easier to solve [36].

A practical system could estimate the angle at which subjects are walking towards the camera and only match samples which have been acquired at similar view points. Such system is presented in Chapter 5.

2.6.3 Subject dependent changes

There are a number of factors that can affect the behaviour or appearance of a subject. Clothing, footwear, walking speed and carrying load and elapsed time are amongst the most commonly quoted factors. For example, clothing can significantly change the overall appearance of a person. Walking speed and different footwear can alter someone's gait. Many studies have investigated the effects and proposed solutions.

Hossain et al. [37] proposed a clothing-invariant approach. The human body is divided into parts and the discriminative capability of each part with respect to change of clothing is assessed. Experiments are performed on a database containing significant clothing variation and encouraging results are reported.

Aqmar et al. [38] have performed gait recognition across different speeds. The study confirms that walking speed can have a negative effect on performance and proposes a novel method that works across speed variations.

A study performed as part of this thesis [39] showed that elapsed time does not affect gait recognition in the short and medium term and that gait can be used as a reliable biometric over time and at a distance.

In-depth understanding of how each subject dependent change affects recognition performance is essential in understanding the limitation and deployment scenarios of an operational system.

2.6.4 Data dimensionality

A number of approaches have been proposed to effectively reduce the dimensionality of the gait signatures without significantly affecting the discriminatory information [40] [41] [42] [43]. These studies employ various statistical and machine learning techniques such as Bayesian networks, PCA analyses, ANOVA etc. to understand which features are the most important in determining someone's identity by gait.

In addition to feature level analyses, other type of feature selection such as gait cycle selection could provide significant improvements in a practical system because of varying levels of sample quality of each cycle.

2.7 Discussion

Currently, gait recognition can deliver very high recognition rates in a constrained environment and if certain factors are controlled. Additional work is required to translate the research to realistic environments and to explore how scalable it is. It is important to understand how results and conclusions achieved on fixed datasets translate to use of the system in real-world conditions. Many of the studies presented in section 2.6 do not have a deployment focus and cannot be easily integrated in a fully automated recognition system. The research community seems to be over focused on proposing novel recognition algorithms and there is lack of work into integrating and combining existing approaches.

Attempts so far suggest that developing highly reliable gait-based human identification system in a real world application is, and will continue to be, very challenging. In Chapter 5, we present novel ways of bringing gait one step closer to deployment and present an operational fully-automated system capable of performing well in a number of scenarios.

The next chapter presents two novel datasets which facilitate analyses not previously possible. It also provides an indication of which approaches or combination of approaches perform well on samples acquired in realistic conditions.

Chapter 3

Datasets and Data Processing

3.1 Introduction

A biometric database can be collected for various purposes. Databases facilitate the development of new research and make comparison between different techniques possible. Performing comparable assessment of different biometric techniques on the same database is vital to show where new techniques stand relative to existing.

The current gait databases contain a smaller number of subjects compared to databases used to evaluate performance of other biometrics (e.g. face, fingerprint). There are databases that include covariate factors and application potential. Nevertheless, certain type of analyses cannot be performed with the current datasets. In many cases custom or proprietary databases might be more appropriate than public databases as they allow new analyses to be performed. A database collected for a specific purpose can lead to novel, more meaningful and more repeatable results.

All of the results presented in this thesis are based on two novel databases specifically collected for the purposes of our experiments. Our new datasets allow a researcher to perform analyses that were previously not possible. The first database has been collected using the Southampton Biometric Tunnel. A full description of this is presented in section 3.3. The second database has been recorded in a realistic indoor environment. Details of this dataset are presented in section 3.4.

The data processing has been performed using existing approaches or a new combination of known techniques that are suitable for data acquired in realistic environment. The data processing techniques are presented in section 3.5.

3.2 Existing gait databases

This section introduces the various databases most widely used by the biometric community. Some of their characteristics are shown in Table 3.1 and Table 3.2.

Name	Subj.	No. seq.	Views	Indoor(I)/ Outdoor(O)	Num. of Loc.
HumanID (USF)	122	1870	2	O	1
SOTON 2002	114	>2500	2	I/O	2
CMU MoBo	100	600	6	I (Treadmill)	1
MIT 2002	24	194	1	I	1
UMD 2002	44	176	1	O	1
CASIA 2006	124	1240	11	I	1
SOTON Multimodal	>300	>5000	12	I	1
Osaka Uni.	1035	2070	2	I	1

Table 3.1 - Details of some of the well-known gait databases

Name	Clothing	Speed	Foot wear	Direct ion	Carrying load	Elapsed time
HumanID (USF)	N	N	Y	N	Y	Y
SOTON 2002	Y	Y	Y	Y	Y	Y
CMU MoBo	N	Y	N	N	Y	N
MIT 2002	N	N	N	Y	N	Y
UMD 2002	N	N	N	Y	N	Y
CASIA 2006	Y	N	N	N	Y	N
SOTON Multimodal	Y	Y	Y	N	N	Y
Osaka Uni.	N	N	N	N	N	N

Table 3.2 - Covariate factors for each database

In addition to the databases presented above the following surveillance databases might be suitable for gait recognition in a realistic environment:

- **I-Lids:** Captured in a very busy environment at Gatwick airport arrival hall. It contains 119 subjects, 4 images per person on average and total of 476 images. It was captured between 3 overlapping and 5 non-overlapping cameras.
- **VIPeR** (Viewpoint Invariant Pedestrian Recognition)**:** The database contains 632 people and there are two images available per person. There is significant viewpoint change between samples of the same subject. The database does not contain full gait cycles, but simply image pair from different cameras.
- **PETS 2006** Benchmark Dataset (Reading): The aim of this database is to allow development and testing of systems for detection of abandoned luggage in real-world environment. It contains relatively small number of subjects.

All of the datasets presented above have some advantages and disadvantages. The type of analyses and experiments that one can perform depend on characteristics of the dataset available. The databases we have recorded (presented in the next two sub-sections) contain characteristics that are not present in any of the publically available ones.

3.3 Biometric tunnel database

None of the existing datasets allow investigation of a single subject dependent factor by isolating the effect of all other factors, both subject and environment related. The SOTON Multimodal has been recorded to enable researchers to perform novel analyses by introducing a single factor at a time. The data is captured using the University of Southampton Multi-Biometric Tunnel. The “tunnel” is a constrained environment that contains 12 synchronised cameras to capture subject’s gait, a camera to capture face video and a camera to take ear images. The intended use of the “tunnel” is at airports and other high throughput environments. It also allows a researcher to perform analysis in a controlled environment and therefore facilitates the deployment of gait in outdoor surveillance/forensic scenarios. The data is acquired automatically in a non-invasive manner as the subject walks through it [44]. The layout of the biometric tunnel is shown in Figure 3.1.

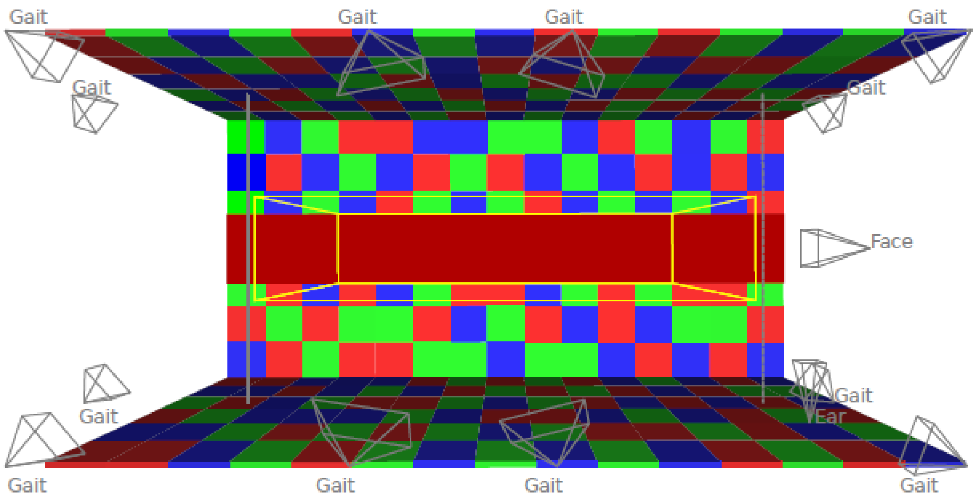


Figure 3.1 – Layout of the Biometric Tunnel

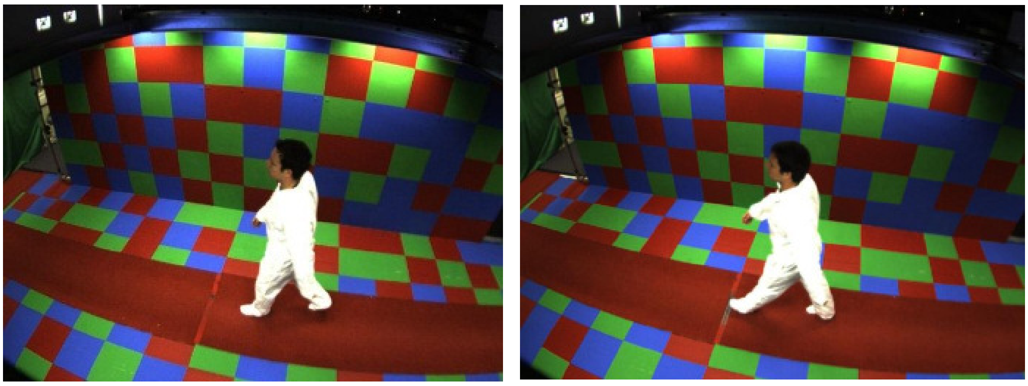
A novel and unique database (*SOTON temporal*) has been created that enables analyses to be performed over time whilst controlling covariates that are known to affect gait recognition performance. Some of the factors that have been accounted for and remained unchanged over time are:

- *Environment*: Background, lighting, walking surface, position of cameras (automatically calibrated)
- *Subjects*: No major changes in any subject's life style were reported. None of the subjects reported any major injuries.
- *Other*: 20 samples per subject per session were collected to enable samples of similar speeds to be compared without explicitly controlling it.

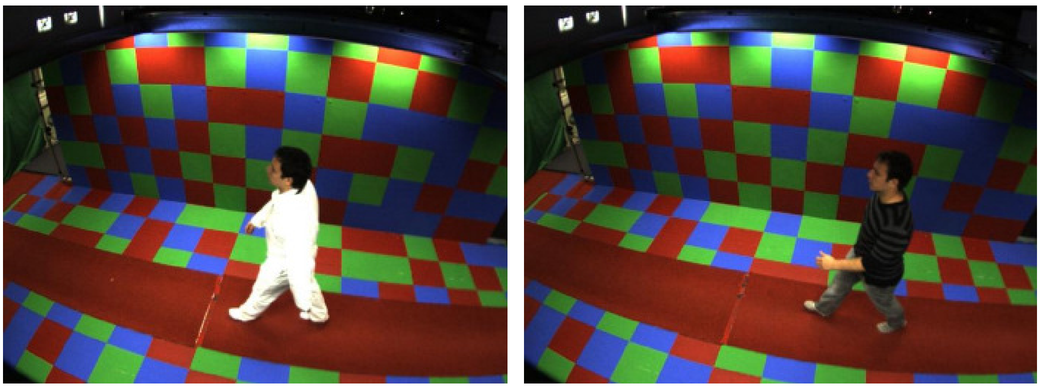
Twenty-five subjects (17 male and 8 female) took part. The age distribution is between 20-52 years old. Data was acquired in five sessions spanning twelve months. The same subjects were used in all sessions. The number of subjects available for each session is shown in Figure 3.3.

The sessions taken in months 0, 1 and 4 consist of 20 samples per subject where subjects wear white overalls over their normal clothes. The session in month 9 differs from the previous three. It consists of 10 samples of each subject wearing normal clothes in addition to 10 samples of each subject wearing overalls. The last session (month 12) was collected to quantify the effect of additional clothing types and footwear. During this session, subjects were asked to wear two different types of ‘ordinary’ clothes that were different from the clothes that they wore in the session recorded in month 9. Twenty-five samples per subject were collected, ten for each type of clothes and additional five of the subjects wearing shoes. The only instruction to subjects is to “walk normally”. Examples of raw data frames are shown in Figure 3.2.

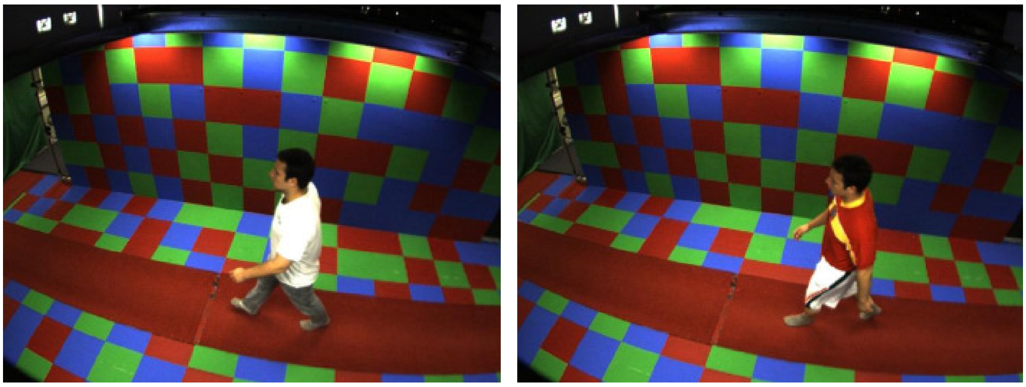
In hindsight, a session taken in month 7 should also have been collected as it would provide data for differences of 2, 6 and 7 months.



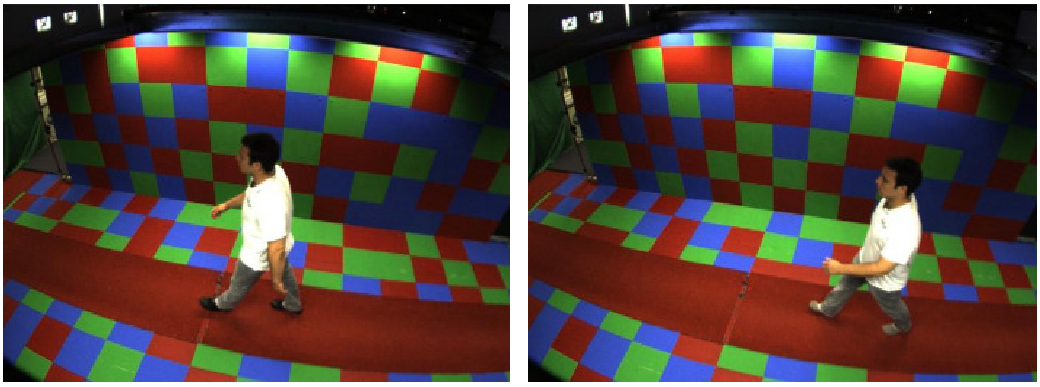
a) Examples of data taken in sessions 1, 2 and 3



b) Examples of data taken in sessions 4 (month 9)



c) Examples of data taken in session 5 (month 12) – Different Clothes



d) Examples of data taken in session 5 (month 12) – With footwear (left), no footwear (right)

Figure 3.2 - Raw data frames from SOTON multimodal temporal database

Session	August (t=0)	Sept (t=1)	Dec (t=4)	May (t=9)	August (t=12)
No of subjects	25	23	22	21	18

Figure 3.3 – Number of subjects for each acquisition session (time is in months)

There are 12 cameras available in the Biometric Tunnel; hence a gait sequence is available from 12 different viewing angles.

Using our database, it is possible for the first time to examine many subject dependent covariates in a principled manner. It enables a researcher to isolate the effect of a particular covariate on the performance of gait recognition whilst keeping all other factors constant.

Other databases add more than one covariate at a time and therefore it is difficult to quantify the individual effect. Our database allows gait analyses with covariates such as elapsed time, change of clothing, variations in walking speed and the effect of footwear by introducing only a single covariate between two data capture sessions. Such analyses are not possible with any of the existing datasets. For instance in the Gait Challenge dataset, three covariates are added simultaneously: clothes, shoes and time. The UMD database uses different viewing angle for the data recorded in March compared to the data recorded in June. Clothing, shoes and background have not been controlled either. None of the existing databases contains samples over a time period of subjects wearing the same clothes and minimising as much as possible the effect of other covariates. This is a *key difference* between our new database and all the existing ones.

3.4 Indoor surveillance database

The *NPL-SOTON* gait database has been recorded at two indoor locations – a corridor (with large windows at each end) and a large room as shown in Figure 3.4. Both locations are realistic environments for indoor surveillance with significant and uneven change in illumination.

There is no publically available gait dataset that contains samples of the same subject taken at two different locations, without introducing additional covariates such as clothing or view point. For example, SOTON 2002 contains samples of subjects acquired at two different locations but no efforts are made to ensure that subjects are wearing the same clothes or walking at similar view point to the camera. Our new database enables investigation of the effect of different environment-dependent covariates in a principled manner. It also provides a challenging environment for development and evaluation of quality metrics. It allows gait signatures from the same subject with different qualities to be matched.

In addition, we have used the database to develop and test various algorithms at each stage of the gait recognition system in attempt to develop a fully automated system that can be transferred in unknown environments.

(a)

(b)

Figure 3.4 – Locations for the experiments (a) Corridor Location (b) Room Location

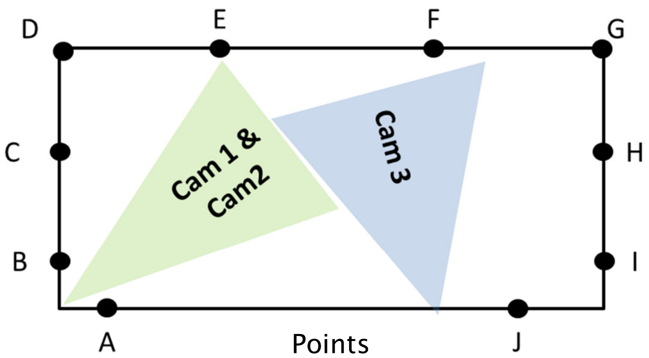


Figure 3.5 – Walking directions

The setup (Figure 3.5) was replicated at both locations using three cameras for data collection. The fields of view of the cameras are shown in Figure 3.5. Camera 2 is elevated above camera 1 and shares its field of view. The instructions given to the 23 subjects were to “walk normally in a straight line between two points. Nearly 2000 sequences were collected at different view angles to the camera. Calibration data is available for all cameras at both locations.

There are 2 sequences per subject for each camera and most directions. The table below details the sequences for each of the cameras and locations that contain at least two gait cycles.

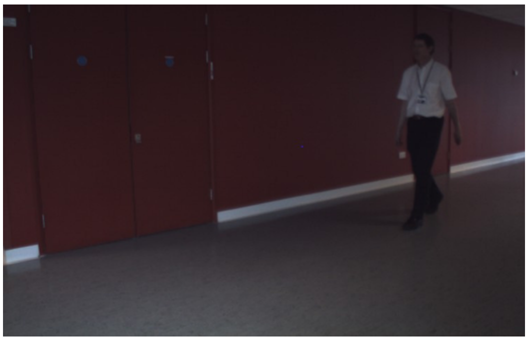
	Loc. 1 – Corridor	Location 2 - Room
Cam1	IB, BI, HC, CH	DEJ, DI, CH, BI, JED[1], ID, IHC, IB, IHA, AG, DG
Cam2	IB, BI, HC, CH	DEJ, DI, CH, BI, JED, ID, IHC, IB, IHA, AG, DG
Cam3	HC, CH	DEJ, CH, JED, IHC, AG, DG, IGA, IGD

Table 3.3 - Walking directions

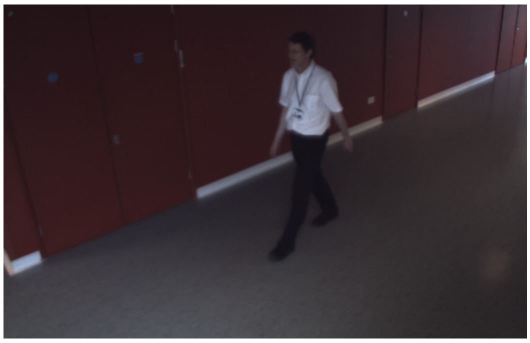
Efforts have been made in ensuring that directions IB, BI, HC and CH at location 1 have been replicated as close as possible at location 2. Additional directions at location 1 could not be recorded due to practical limitations such as distance from camera and lens type/dimensions.

In addition to the directions shown in Table 3.3, the subjects were recorded turning away from the camera when they were in the middle of the field of view of ‘camera 1’ and ‘camera 2’. The subjects were also recorded walking ‘zigzag’ across the field of view. Examples are shown in Figure 3.6. Appendix C shows additional examples.

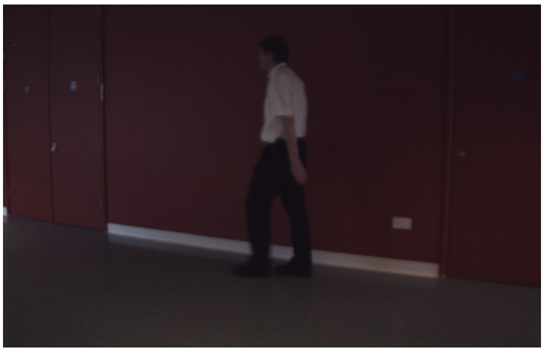
[1] Directions specified by three letters have been used to ensure that a video sequence starts with a subject being outside the camera’s field of view.



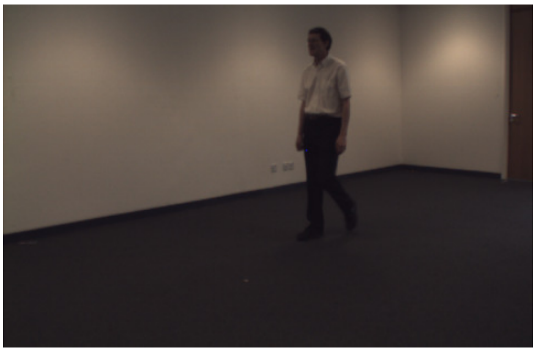
a) Sample acquired using camera 1



b) Sample acquired using camera 2



c) Sample acquired using camera 3



d) Sample acquired in ‘Room’ location

Figure 3.6 – Examples from the indoor surveillance database

3.5 Data processing

This section provides an overview of the general processing steps involved in gait recognition. It also presents detailed description of all processing steps implemented to analyse the datasets presented in sections 3.3 and 3.4.

3.5.1 Overview

A gait recognition system primarily consists of a computer vision system. A gait signature is created by extracting images of a walking subject which is then compared to the signatures of known subjects. Figure 3.7 shows an example of some of the basic steps in a gait recognition system.

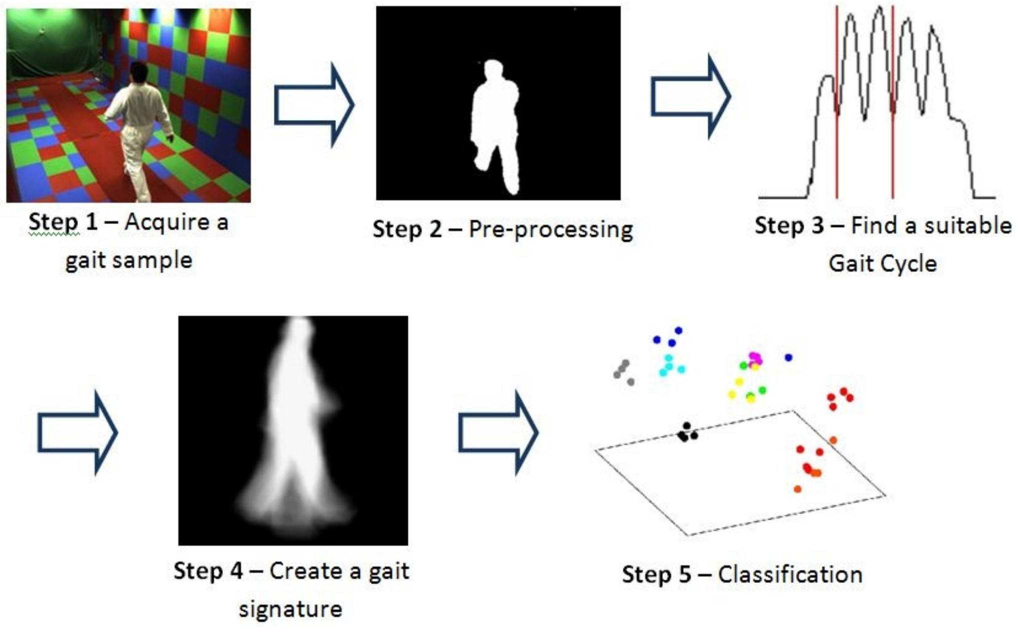


Figure 3.7 - General steps of a gait recognition system

Step 1: Data can be acquired using a single or multiple cameras. If data is acquired using single camera recognition can be performed using a 2D gait signature such as the Gait Energy Image (GEI-shown in step 4). However, if multiple but synchronised cameras are used, the number of possibilities is greater. Examples of the usage of multiple-synchronised cameras include:

- Producing a 3D gait model and using it for recognition
- 3D information can be used to improve recognition for a 2D approach by producing a non-normalised version of a 2D signature
- Achieving a view-invariant recognition. A gait signature from any view can be re-created using 3D data which can be mapped to a signature acquired in an outdoor environment where typically only a single and non-optimal view is available.

Step 2: An example of pre-processing step is background subtraction or background segmentation. Background subtraction can measure the naturally occurring scene behind the walking subject using one of a plethora of computer vision techniques. A significant challenge is separating subjects in the case of occlusion. A gait recognition operating in realistic environment should be able to deal with multiple subjects and operate in the presence of various types of occlusion. In this thesis we have assumed that only a single subject is present in the field of field, because of imperfections in the current tracking and background subtraction algorithms. Although this assumption is unlikely to hold in the 'real-world', all analyses and results presented in this thesis remain valid if we are able to accurately track and separate multiple subjects.

Step 3: As human gait is periodic, a gait sequence (sample) can consist of multiple gait cycles. Identifying the most suitable cycle can lead to better recognition rates. Signal processing techniques can be applied to the foreground signal (sum of foreground pixels) in the case of binary image.

Step 4: There are number of approaches to produce a gait signature, some of which were described in the previous chapter. A baseline gait signature was proposed in [11]. An example of a signature is shown in Step 4.

Step 5: A gait signature can be used directly within a classifier. Alternatively, features can be extracted from a signature and those features can be used for classification. Again, there is a selection of classification techniques; in the simplest case, a classifier such as *k*-nearest neighbour (using Euclidian distance) can be used.

One of the key deliverables of the research presented in this thesis is a gait recognition system operating in a realistic environment. Therefore, processing steps capable of performing well in the presence of noise were required. In most cases, a number of known techniques or a new combination of existing techniques have been tested to understand which ones performs best in challenging environment.

3.5.2 Data acquisition

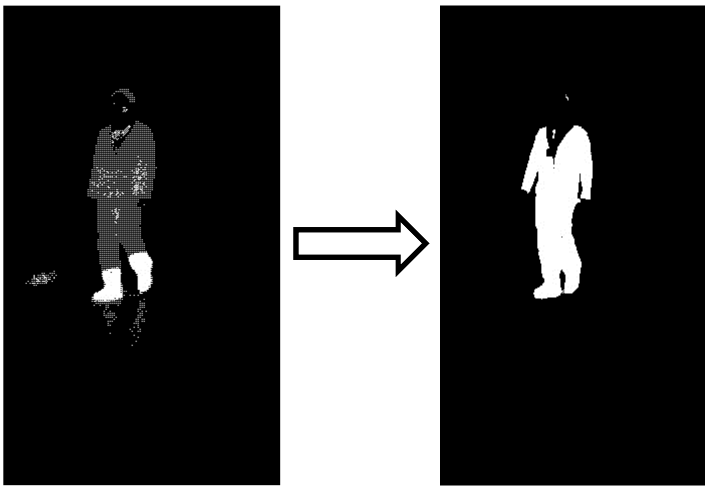
3.5.2.1 Biometric Tunnel database

A data sample consists of 12 video sequences from different angles sampled at 30 fps. The cameras are synchronised and an automated calibration procedure ensures the integrity of each sample. Photo-electric beam sensors signal the start and end of a sample.

3.5.2.2 NPL-SOTON database

Data was captured using three cameras at 30 fps. The cameras were not synchronised and the camera calibration was performed by using a checkerboard. The start and end of a sample was detected automatically by a software module.

A demosaicing algorithm was implemented to convert the raw camera outputs (Bayer pattern image) into a suitable colour space. Test revealed that the results of the segmentation process suffer if the raw camera outputs are used. There are more shadows and reflections present in the case of the Bayer pattern image (see Figure 3.8).



(a) (b)

Figure 3.8 - Raw image (a) and image converted into a suitable colour space (b)

3.5.3 Background subtraction

Background subtraction is a common pre-processing step for a large number of computer vision algorithms. Emphasis was placed on algorithms that have 'real-time' performance. Six popular background subtraction algorithms were tested. A very brief description is shown below:

a) Adaptive Median [45]. This algorithm uses a combination of image differencing with respect to median background and a Laplacian operator. It tracks blob edges in the segmented image.

b) Eigen-background [46]. The background is modelled by adaptively building an eigenspace model which describes the range of changes.
c) Adaptive Gaussian mixture models [47]. Each pixel is modelled as a mixture of Gaussians and the values are updated online.
d) Temporal median background subtraction [48]. The image is portioned into blocks and the blocks are selectively updated using temporal median.
e) Wren et al. [49]. This approach uses two models, one for the scene and one for the person. The person is modelled by 2D blob and their low-order statistics. The scene is modelled as a texture surface and each point is associated with a mean colour value and a distribution about that mean.
f) Implementation of the Gaussian mixture model (GMM) background subtraction algorithm developed by Zivkovic and Heijden [50].

Most algorithms performed well in the Biometric Tunnel. The Gaussian mixture model algorithm implementation by Zivkovic and Heijden performed best in more challenging environments. Appendix A shows details of the algorithm. The source code provided by the authors was used.

3.5.4 Gait-cycle finder

There can be multiple gait cycles available for a given walking sequence. A gait-cycle finder algorithm usually consists of two parts: 1) finding the start and end frames of the subject entering and leaving the field of view of the camera; 2) finding the gait period and selecting one of the available cycles. There can be a varying amount of noise across a gait sequence and therefore the process can be extended to selecting a cycle which contains data with least amount of noise. In Chapter 5 we demonstrate an approach to achieve this.

3.5.4.1 Part 1 – Finding the start and end frame

Initially, a HOG (Histogram of Oriented Gradients) [51] based person detector was tested to detect when a person enters and leaves a scene. The approach worked well with the data recorded in the Biometric tunnel, but performed badly in the more challenging environment. A simpler approach, described below, delivered much better results. We analysed the sum of all the white pixels of the background subtracted images. An example of the signal is shown in Figure 3.9. A subject is considered to have entered the field of view when the signal reaches a pre-defined threshold value which is a function of the distance from the camera and the direction of walking.

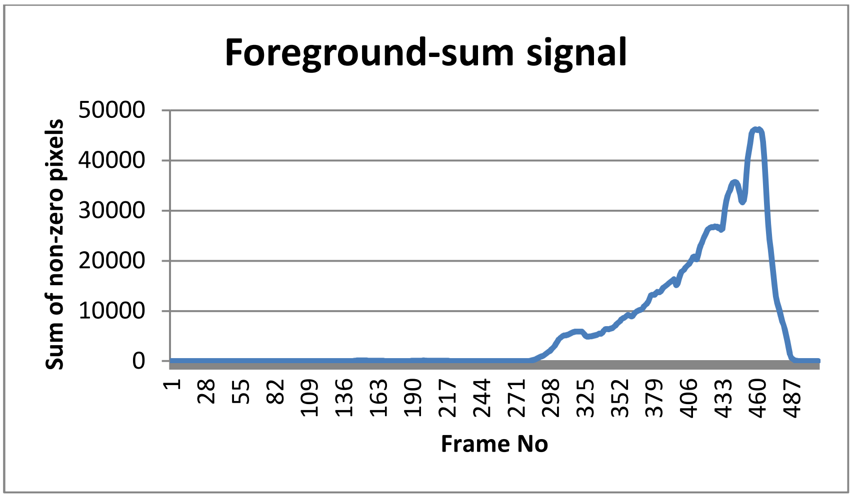


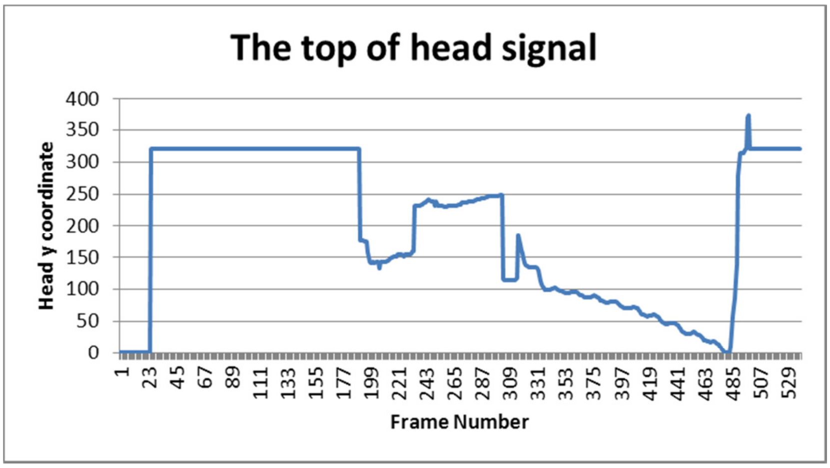
Figure 3.9 - The sum of white pixels over time

3.5.4.2 Part 2 – Gait period estimation and period selection

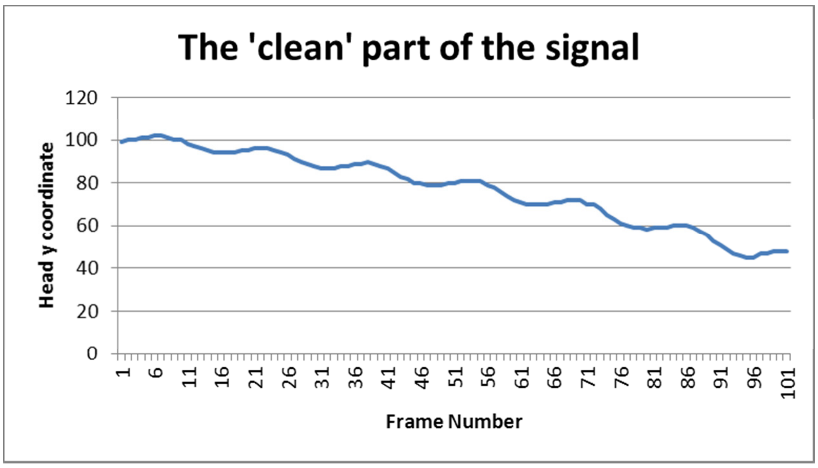
A number of approaches for gait period estimation were tested:

a) Bounding boxes information, e.g. ratio of height and width. See Figure 3.11 (a)
b) Fitting a model to height normalised version of the bounding boxes. For a good quality sample the signal is sinusoidal with the maxima corresponding to the double-support stance. Figure 3.11 (b)
c) Autocorrelation is a mathematical tool for locating periodic signals in the presence of noise. The autocorrelation of a periodic signal is itself periodic with the same period.
d) Tracking the position of the top of the subject's head. See Figure 3.10.

Fitting a model to the number of white pixels in the height normalised version of bounding boxes (ii) worked well in the Biometric Tunnel database. Methods a), b) and c) did not perform well on the data from the NPL-SOTON database. The best results were obtained by finding the period of the signal produced by tracking the top of the subject's head (method d). The period is estimated by smoothing and differentiating the signal and finding the median value of the distance between the zero crossings with the x axis. Examples are shown in Figure 3.10. Other studies [52] have also found that the top of the head produces the cleanest signal in the presence of significant noise.



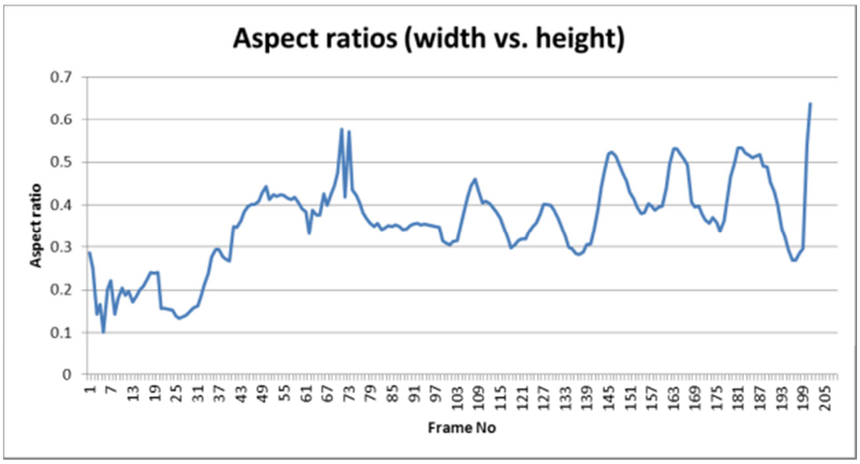
(a) Signal produced by tracking the top of subject's head



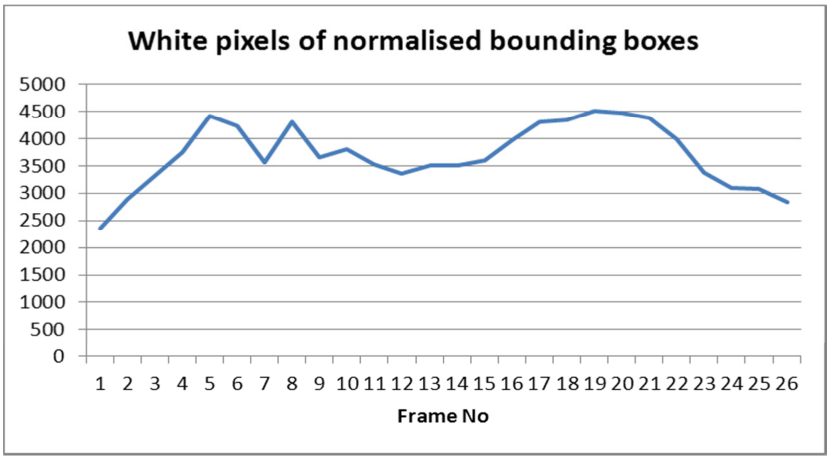
(b) The 'clean part of the signal shown in part (a[2]

Figure 3.10 – The signals produced by detecting the top of subject's head

[2]). Frame 1 in (b) corresponds to frame 331 in (a)



(a) The signal of the aspect ratio of the subject's bounding box



(b) The number of white pixels in height normalised version of bounding boxes

Figure 3.11 - The signals produced by other techniques

There can be more than one gait cycle available in a gait sequence. After a gait period has been estimated a gait cycle can be selected. A simple approach is to select the cycle that occurs in the middle of the sequence. A more complex approach that uses the signal quality is described in Chapter 5.

3.5.5 Signature computation

The first step of the signature computation is extracting the region (bounding box) around the subject. There are a plethora of computer vision techniques for extracting regions in binary images. We have used Connected Components Analysis and assumed that the largest connected component belongs to the walking subject.

Moments were used to locate the centre of mass using equation (3.1) and (3.2).

$$m_{ij} = \sum_x \sum_y x^i y^j \, I(x,y) \tag{3.1}$$

where *I(x,y)* is the intensity value of the pixel at coordinates *x* and *y.*

$$\bar{x} = \frac{m_{10}}{m_{00}}, \qquad \bar{y} = \frac{m_{01}}{m_{00}} \tag{3.2}$$

m_{10}, m_{01} and m_{00} are calculated using equation (3.1).

Finding the centre of mass allows all silhouettes forming a part of a gait cycle to be aligned and normalised before a signature is computed. The alignment is achieved by using an affine transformation. Equations (3.4) and (3.3) show the details.

$$M = \begin{bmatrix} \frac{W}{h} & 0 & -\bar{x}\left(\frac{W}{h}\right) + \frac{W}{2} \\ 0 & \frac{H}{h} & 0 \end{bmatrix} \tag{3.3}$$

where *W* and *H* are the width and height of the new image (*dst*), *h* is the height of the original image (*src*) and $\bar{x}$ is calculated using equation (3.2).

$$dst(x,y) = src(M_{11}x + M_{12}y + M_{13},\ M_{21}x + M_{22}y + M_{23}) \tag{3.4}$$

The values for the matrix *M* are derived from equation (3.3), *src* is the original image and *dst* is the transformed image.

The transformation improves the signature quality.

After the bounding boxes forming part of chosen gait cycle have been normalised and aligned the next step is to produce a gait signature that can be used for recognition. Two gait signatures have been used: Gait Energy Image (GEI) and Gait Entropy Image (GEnI). GEI was chosen because it is a baseline gait signature and is one of the most widely used approaches in the literature. Gait is represented as a single grey scale image obtained by averaging the silhouette extracted over a complete gait cycle. If $P(0)_{x,y}$ is the first image of the cycle and the next image is $P(1)_{x,y}$ the average silhouette can be computed using equation (3.5) [14].

$$AS(t)_{x,y} = \frac{1}{t}\sum_{i=0}^{t} P(i)_{x,y} \tag{3.5}$$

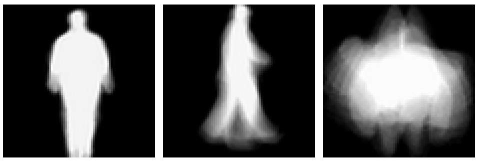
where *t* is the total number of images in the gait period.

For the samples acquired in the biometric tunnel, three different views of the GEI have been used in our experiments: *top* view, *side* view and *front* view. For these samples, two versions of each view have been computed: normalised *(Norm)* and non-normalised *(NN)*. Unlike the non-normalised signature, the normalised one does not retain the subject's height and body mass. Producing a non-normalised version is possible due to full 3D reconstruction. The size of the normalised signature is 64x64 pixels and the size of the non-normalised is 50x50 pixels. See Figure 3.12 for details.

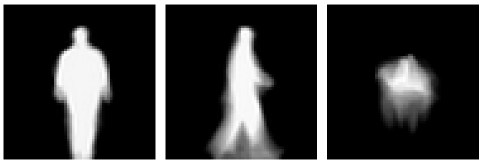
Gait Entropy Image (GEnI) is a newer approach which captures mostly motion information and is suggested to be robust to covariate conditions changes [3]. GEnI is computed by calculating the Shannon entropy for every pixel. See Equation (3.6).

$$H(x,y) = -\sum_{k=1}^{K} p_k(x,y)\log_2 p_k(x,y) \tag{3.6}$$

$p_k(x,y)$ is the probability that a pixel takes on the k^{th} value. A binary image corresponds to $k = 2$. GEnI is computed using a binary height normalised image of a walking subject.



a) GEI Normalised (Front, Side, Top)



b) GEI Non-Normalised (Front, Side, Top)

Figure 3.12 – Examples of gait signatures

Features can be concatenated and a modified GEI signature can be created. An example signature combining side, front and top view *(S+F+T)* is shown Figure 3.13.

Figure 3.13 – Example of feature concatenation

We have created a feature reduction module which reduces the size of the signatures and enables us to perform recognition at different resolutions. Bicubic interpolation was used to reduce the signatures sizes to the following dimensions: 50x50, 40x40, 30x30, 20x20, 17x17, 15x15, 12x12, 10x10 and 7x7 pixels.

3.5.6 Classification

The gait signatures are used directly for classification. Improvements in performance could have been achieved if additional features had been extracted from the gait signatures and feature set selection had been performed to understand which features contribute most towards recognition. However, since the focus of this thesis is not improvement in performance at the classification stage these improvements were not deemed necessary. Chapter 5 and the next two sections describe how heel strike and height can be used to perform selection at the sample level rather than at the feature level.

A distance matrix containing all Euclidian distances between all samples is constructed. The Euclidian distance (*d*) between two samples is calculated by using the intensity values of the gait signatures:

$$d(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \qquad (3.7)$$

where *n* is the number of pixels in a gait signature and ***p*** and ***q*** are the two gait samples.

The *k*-Nearest Neighbour classifier (*k*-NN) was applied. Experiments for *k*=1, *k*=3 and *k*=5 were performed. Performance would likely improve with alternative classifier but this is not in the scope of this thesis.

3.5.7 Height estimation

Height has been used in forensics to narrow down the search of suspects. Height is usually classified as a soft biometric and it is not sufficiently distinctive to be used for recognition. However, it can be used in visual surveillance in many ways. Some examples are described below:

- Improving the performance of a biometric system
    - By reducing the size of the gallery
    - By confirming or discriminating matches at the end of a recognition process
- For gait recognition, it can be used for extracting additional features such as distance between joints
- Tracking an object. The purpose of tracking algorithms is to extract the same object from series of frames. The height of the object can help in establishing temporal correspondence. In the case of multiple cameras

observing the same scene, it is possible to understand which projections belong to the same object. Height provides additional information in tracking objects between non-overlapping cameras.

Estimating actual subject's height from video requires a calibrated camera. A fully automated calibration procedure has been implemented to calculate all of the unknown camera parameters. See Appendix B for details of the camera model used. The calibration procedure requires a subject to display a calibration pattern in front of the camera. In addition placing the calibration board on the floor is required (see Figure 3.14).

Figure 3.14 - The calibration procedure

A fully automated height estimation system has been designed and implemented using a method adapted from [52]. The method is best described by Figure 3.15. I_f is the foot point on the image (in pixels) and I_H is the head point in the image. There are different ways of selecting this point for each frame. Initially we have used the top of head and the lowest vertical point of the shape (as defined by the bounding box). An alternative way of finding I_H is to use the 'heel strike' position (see section 3.5.8). The points are corrected for distortion and then translated to 'world coordinates'. *F* represents the *foot line* and *H* represents the *head line.* The foot line crosses the floor at point P_0. The point P_0 has the following real-world coordinates $(x_0,y_0,0)$ and represent the position of the subject on the floor. The line *V* is perpendicular to the floor. The point P_h lies on the H line and has approximately the same x and y coordinates as P_0.

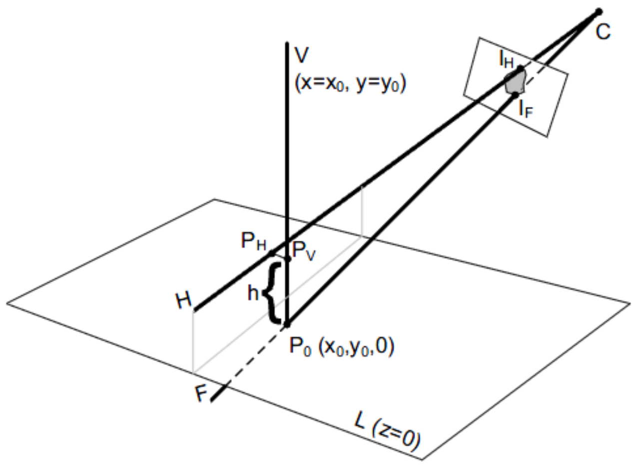


Figure 3.15 - Height calculation (taken from [52])

The equations (3.8) and (3.9) have been derived using the theory presented in Appendix B and the diagram shown in Figure 3.15. The equation variables can be understood from the diagram and the appendix.

Both equations produce an estimate for height. The point P_h is an approximation of the point P_v and therefore we choose the equation which produces a point closest to P_v. For example if $|X1\text{-}x0| < |Y1\text{-}y0|$ then we use the *h* from the second equation. Solving the equations is relatively straight forward as there are 3 equations with 3 unknown variables in each.

$$\begin{bmatrix} x' & -r_{12} & -r_{12} \\ y' & -r_{22} & -r_{23} \\ 1 & -r_{32} & -r_{33} \end{bmatrix} \begin{bmatrix} z \\ Y_1 \\ h \end{bmatrix} = \begin{bmatrix} r_{11}X_0 + t_1 \\ r_{21}X_0 + t_2 \\ r_{31}X_0 + t_3 \end{bmatrix} \tag{3.8}$$

$$\begin{bmatrix} x' & -r_{11} & -r_{13} \\ y' & -r_{21} & -r_{23} \\ 1 & -r_{31} & -r_{33} \end{bmatrix} \begin{bmatrix} z \\ X_1 \\ h \end{bmatrix} = \begin{bmatrix} r_{12}Y_0 + t_1 \\ r_{22}Y_0 + t_2 \\ r_{32}Y_0 + t_3 \end{bmatrix} \tag{3.9}$$

The method relies on placing the calibration board on the floor (see Figure 3.14). Using the calibration board it is possible to eliminate one of the unknowns in the real world coordinate system i.e. all points that lie on the floor can be considered to have coordinates *(x,y,0)*. Using a real value for the dimensions of the checkerboard it is possible to estimate the actual height of the subject. Figure 3.16 demonstrates how the calibration board is used to establish correspondence points. The coordinates are stated in centimetres.

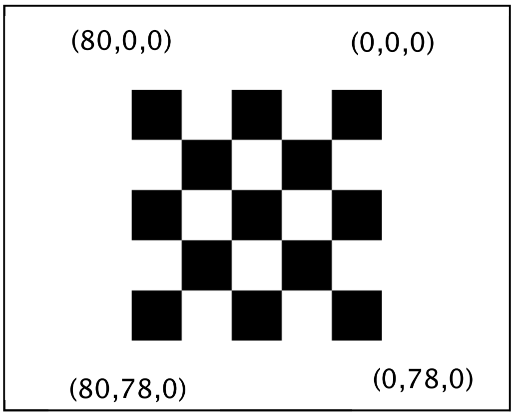


Figure 3.16 - Correspondence points on the calibration board

In this thesis height has been primarily used to improve recognition performance by reducing the size of the gallery at the matching phase. Before matching is performed, the height *h* (in metres) of the subject of the probe sample is estimated. The probe sample is only matched against the samples in the gallery which contain subjects of height $h \pm n$ metres, where n determines the height range of the subjects in the gallery. Matching against a gallery of smaller size improves recognition.

3.5.8 Heel strike extraction

When a foot lands on the floor it stays at the same position for about half a gait cycle while the rest of the body moves forward. Using this knowledge and a combination of computer vision techniques we are able to locate the heel strike positions in images. Combining heel strike information with a calibrated camera can provide valuable additional information for gait recognition. Some examples are shown below:

- Accurate heel strike detection can lead to a more accurate bounding box selection which can help improve the signature quality.
- Heel strike data can provide a better estimate of subject's height than if bounding box information is used. Our experience show that in challenging

environments, the bounding box algorithm might include noise in forms of shadows and reflections as part of the bounding box. The top of head signal is usually free of noise and easy to detect.

- Direction of walking and an indication of whether a subject is walking in a straight line.
- The angle at which a subject is walking relative to the camera.
- The actual distance of the subject from the camera.
- Estimation of step length or stride length.
- A reliable estimation of walking speed.

The following steps are used to compute the heel strike position:

Step 1: Create an Accumulator Map image (*AM*) using equation (3.10):

$$AM_{x,y} = \sum_{i=a}^{b} P(i)_{x,y} \tag{3.10}$$

where *a* and *b* are the frames when the subject appears and leaves the field of view respectively. *P(i)* represents the background subtracted binary image.

An example of an accumulator image is shown in Figure 3.17 (left). Our experiments show that there is higher likelihood of noise being present when a subject enters a scene and therefore accurately estimating *a* and *b* is very important.

Step 2: Apply the Harris corner detector on the accumulator image. A typical result of applying this operator is shown in Figure 3.17 (right).

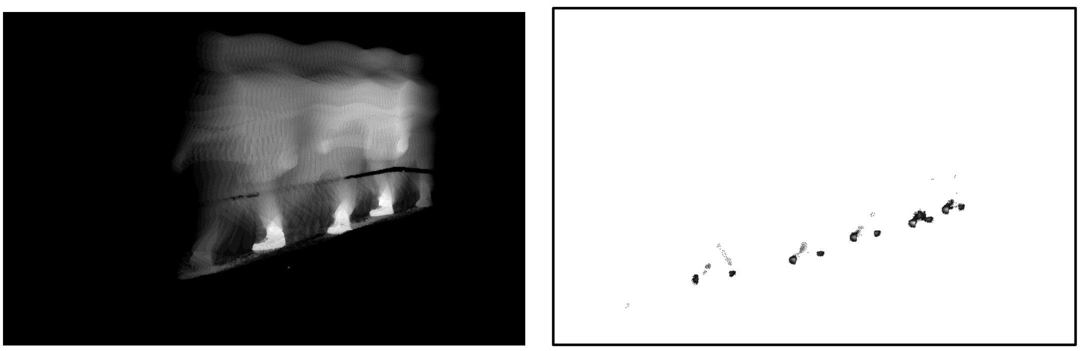


Figure 3.17 - Accumulator image (left), applying Harris corner detection (right)

Step 3: Detect the peaks in the image produced by step 2. The aim is to detect the darker regions. Initially, a low pass filter is applied to reduce the noise and then the local maxima are approximated.

In this thesis, we use heel strikes to detect the angle at which a subject is walking relative to the camera. This is achieved by finding the line that best fits the results produced by step 3. Hough transform for lines has been used. The ‘real world’ coordinates of the heels strike positions relative to the calibration board are shown in Figure 3.18. The figure shows an example of subject’s walking path relative to the calibration board. If the position of the calibration board relative to the camera is known the walking angle of the subject relative to the camera can be computed. The approach assumes that a person walks without stopping while in the field of view of the camera.

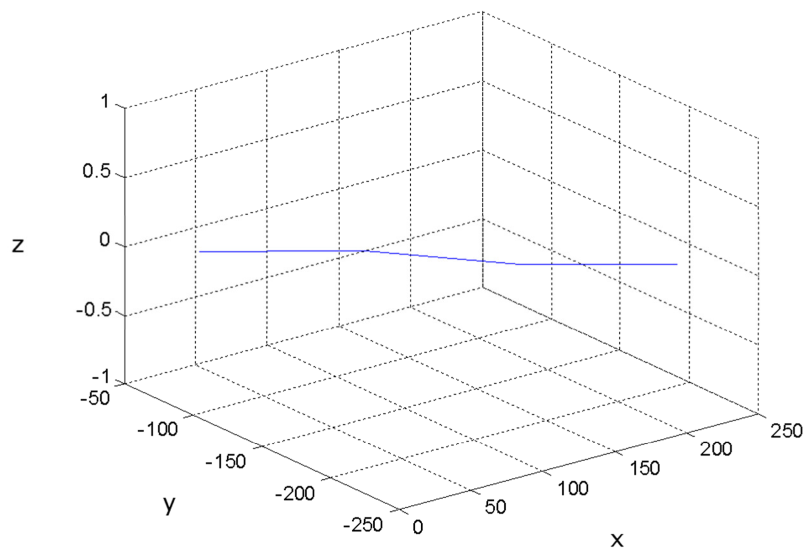


Figure 3.18 - The subject's walking path relative to the calibration board

3.5.9 Performance evaluation

The Correct Classification Rate (CCR) at the matching stage is a good indication of the performance of a biometric system or algorithm and is invariably reported in biometric approaches. However, there are other ways of measuring performance which can provide additional insight into how an algorithm or system is performing. The performance in this thesis is evaluated in the following ways:

- The Correct Classification Rate (**CCR**)
- The Receiver Operating Characteristic (**ROC**) curve
- The Equal Error Rate (**EER**)
- A diagram of *intra-* (within subject) and *inter-* (between subject) **class distributions**.
- The decidability metric (**d'**) [53] determines how much overlap there is between two distributions:

$$d' = \frac{\left|\mu_{same\_subj} - \mu_{diff\_subj}\right|}{\sqrt{\frac{1}{2}\left(\sigma^2_{same\_subj} + \sigma^2_{diff\_subj}\right)}} \tag{3.11}$$

  where μ_{same_subj} and σ_{same_subj} refer to the intra-class mean and standard deviation respectively and μ_{diff_subj} and σ_{diff_subj} refer to the inter-class mean and standard deviation. The problem becomes more decidable if their means are further apart or their variances are smaller. The decidability metric should be very similar if changes in experimental conditions do not drastically affect the separation in the feature space.

The Cumulative Match Curve (CMC) is sometimes used when assessing the performance of biometric systems. It is a measure of 1:m identification system performance and judges the ranking capabilities of an identification system. The CCR is equivalent to CMC at rank 1.

3.6 Discussion

Two novel databases have been collected and will be make publically available. The datasets are on the opposite side of the spectrum. The first one is acquired in heavily controlled environment and the second in realistic real world settings. The biometric tunnel database is ideal for examining factors related to the subjects and the indoor surveillance database is ideal for examining environment related factors. The databases allow new types of analyses to be performed which were not possible previously. Using the first dataset it is possible to examine a number of important subject dependent factors in a principled manner by changing only one factor at a time. Using this dataset it is possible to quantify the effect of: time, clothing, walking speed, distance from camera, footwear type etc. The second dataset enables a researcher to understand the challenges associated with deploying gait recognition in the 'real-world'. More specifically, it allows experiments to be performed related to: matching samples of different quality, developing and verifying algorithms for quality assessment, matching across different locations, changes in view-point, changes in camera height, extracting and using additional soft biometrics such as height etc. Both datasets contain data which has not been analysed in this thesis and therefore provide scope for future research.

This chapter describes all data processing steps required for a fully operational gait recognition system. It also illustrates the challenges associated with some of the processing techniques and provides an indication of which algorithms work best in realistic environments. A number of approaches have been examined and the experiments indicate that relatively simple techniques perform better in real world conditions as they are more robust to noise. The techniques used throughout this thesis are mainly a new combination of well-known techniques. The details of extracting additional gait features such as heel strike and height are also presented as they aid the implementation of a system operating in realistic environments.

Chapter 4

Subject Dependent Covariates

4.1 Introduction

There are a number of factors (also known as covariates) that affect gait recognition. Studies quantifying the effect of covariates on the performance of gait recognition algorithms include [22, 54, 55]. Covariate factors can be related either to the subject (e.g. different clothing) or to the environment (e.g. different walking surface). Some examples of covariates that affect recognition are: viewing angle, shoe type, walking surface, carrying objects and elapsed time between sequences being compared. Understanding these factors is crucial to developing robust and accurate gait recognition algorithms. The effect of a particular covariate on the recognition performance depends on the algorithm adopted.

In this chapter we examine the effect of the following subject related covariate factors: elapsed time, clothing, footwear and speed. We present empirical evidence to show that by controlling clothing worn by the subjects and the environment the recognition performance is not affected drastically over 9 months. This is the longest time period yet considered for gait and we report, rather to our surprise, a much higher recognition rate than any previous study. A particular emphasis is placed on elapsed time in this chapter as the rest of the results obtained are consistent with the literature.

4.2 Dataset used

Table 4.1 shows some of the most well-known datasets that contain data sampled at different times. The database shown at the bottom of the table was described in section 3.3. It is the largest gait temporal dataset with the longest time period yet considered for gait recognition. In addition, our database is the only one acquired in controlled setting and thus provides a consistent environment across different acquisition sessions.

Investigating the effect of time and other subject related covariates in a principled manner is not possible with any previous dataset. It is impossible to determine whether the recognition performance varies due to changes of gait over time or due to other factors such as change in clothes, shoes, lighting or viewing angle. Another advantage of our database is the availability of more than just two acquisition sessions. It is possible to achieve analysis between 10 different time periods.

Name	No. of Subjects	No. of Samples	Time Difference	Views	Indoor/ Outdoor
MIT 2001 [56]	13	194	3 months	1	Y/N
HumanID [11]	33	132	6 months	2	N/Y
SOTON 2002 [57]	10	≈100	7 months	1	Y/N
UMD [58]	25	100	3 months	4	N/Y
SOTON Temporal [44]	25	2280	0,1,3,4,5,8,9, 12 months	12	Y/N

Table 4.1- A summary of existing temporal gait datasets

4.3 The Effect of Time

4.3.1 Previous Work

In previous studies considering the effect of elapsed time on recognition by gait, in [57] a probe consisting of 10 subjects is matched to a gallery containing 115 subjects. The subjects are filmed 6 months apart and a CCR of 37% is achieved. A study [11] reported a substantial drop in recognition performance (at rank 1) from

78% to 3% when shoe and time (6 months) covariates were introduced, using the Gait Challenge database. In [59], a different recognition algorithm is used on the same dataset. Seventy silhouettes are chosen and manually subtracted from the background. The recognition performance over time dropped to 10%. In [56], two different recognition algorithms are employed on the MIT database. Experiments are performed over two months in an indoor environment with different backgrounds and lighting. On a dataset of 24 subjects, the recognition rate varies between 30-60%. Another study [60] using the MIT database shows a performance drop to 45% from 100% for samples taken on the same day. Experiments on the UMD database reveals a drop in recognition performance to 30% over 3 months using 25 subjects in indoor environment [58]. Clothing was not controlled in any of these studies.

Many studies such as [22, 54] and [61], do not consider temporal data and focus solely on data acquired on the same day.

Research such as [62] treat time (over 6 months) as the most difficult covariate for gait recognition. They have proposed a new pattern classification method to solve the elapsed time problem in gait recognition.

4.3.2 Results

There is a consensus in the literature that the time taken between recording the gallery and the probe affects recognition performance the most [11]. Time as a factor has not been considered explicitly in many studies mainly due to the lack of suitable database. Existing datasets introduce more than one factor between the gallery and probe and that makes it hard to quantify the exact effect of time on the recognition performance. Since gait is a behavioural biometric, an important question arises: "Is it possible to recognise someone reliably after a certain period of time has elapsed?" Permanence is an important characteristic of any biometric. A biometric trait that changes significantly over time is unlikely to be useful for recognition in numerous scenarios.

Experiment has been performed to understand the effect of time on gait recognition performance. The combinations of probe and gallery used in this experiment are shown in Table 4.2. Each probe and gallery consists of 10 samples per subject. Full details of the dataset used are described in Chapter 3, section 3.3. In this experiment, we only use the samples of the subjects wearing overalls to provide consistent clothing over time. The non-normalized version of the signature (50x50 pixels) is used for this experiment. A 4-fold cross-validation is performed and the highest recognition rates are shown in Figure 4.1. A 2-fold cross-validation is performed for experiments for time differences of 5, 8 and 9 months due to a smaller number of samples with overalls filmed in month 9.

The error-bars for GEI (*S+F+T)* indicate the lowest results of the 4-fold cross-validation. GEI (*S+F+T*) represents a signature created by stacking the GEI (side), GEI (front) and GEI (top) into a single image as described in 3.5.5. The error bars for the other signatures are not shown for clarity. The standard deviation for the results of the 4-fold cross validation is between 0.00-0.03. Values for the EER and the decidability metric (*d'*) are also shown in Table 4.2.

The results in Figure 4.1 clearly indicate that the CCR does not fall considerably over time for any of the signatures considered in this chapter. The variation of performance over time is similar for all signatures considered. The EER and decidability (*d'*) for GEI (side) and GEI (Front) shown in Table 4.2 indicate that the signatures have a similar discriminatory ability, although the front view performs slightly better in most cases. However, our results show that there is a benefit of combining multiple views to construct a single gait signature. The GEI (S+F+T) produces the lowest EER and the highest *d'* for all experiments. For instance for a time difference of 1 month the EER is 6.39% and the *d'* is 2.24. Furthermore, using a non-normalized version of the silhouettes improves performance. The Gait Entropy (GEnI) uses a normalized silhouette and the EER are always higher and *d'* is always lower compared to all types of non-normalised GEI signatures. For instance for a time difference of 1 month the EER is 14.57% and *d'* is 1.90 in the case of GEnI.

Probe month	Probe subj.	Gallery month	Gallery subj.	EER % GEI Side	EER % GEI Front	d' GEI Side	d' GEI Front	Time Diff. (months)
0	25	0	25	4.95	3.10	2.71	2.76	0
1	23	1	23	6	4.05	2.49	2.56	0
4	22	4	22	6.5	4.10	2.52	2.57	0
1	23	0	25	9.35	8.61	2.14	2.12	1
4	22	1	23	10.19	7.30	2.15	2.20	3
4	22	0	25	9.72	9.43	2.10	2.20	4
9	21	4	22	10.00	10.61	2.24	2.12	5
9	21	1	23	9.40	9.50	2.18	2.15	8
9	21	0	25	11.54	7.31	2.11	2.22	9

Table 4.2 – The combination of gallery and probe for the effect of time experiment

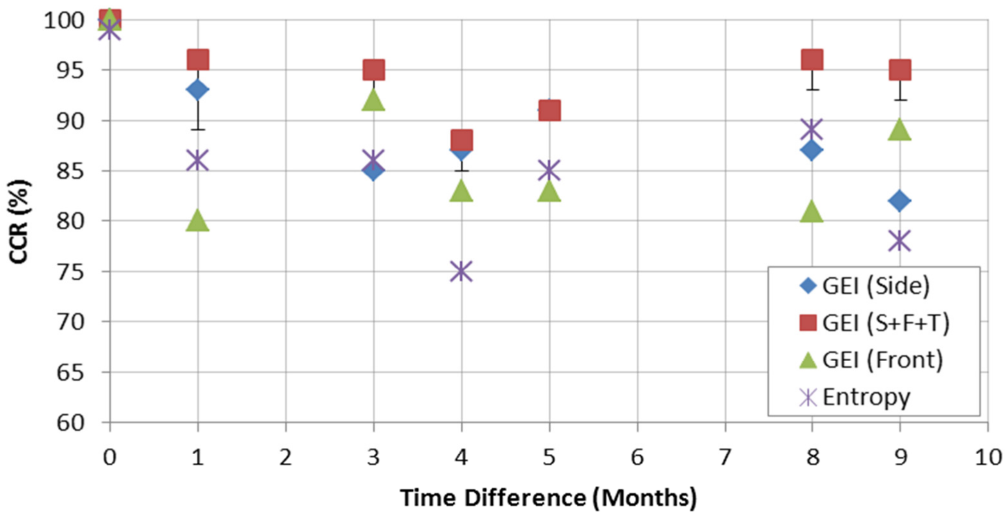


Figure 4.1 - Recognition Performance over time for the GEnI signature, 2 different views of the GEI signature and fusion of 3 different GEI views (S+F+T = Fusion of Side, Front and Top view).

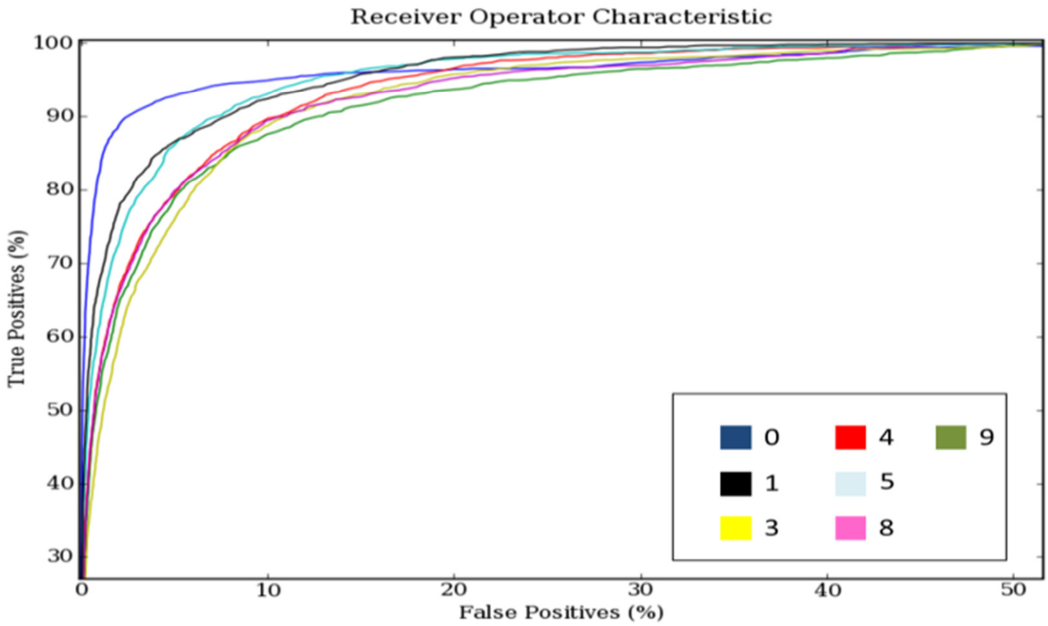


Figure 4.2 – ROC curve for all combinations of gallery and probe shown in Table 4.2

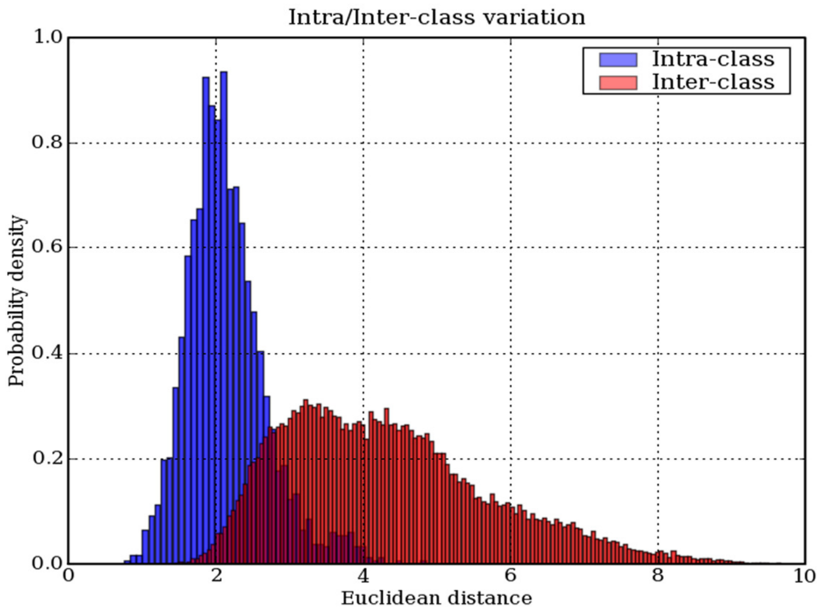
Figure 4.2 shows that the ROC curves for all combinations of gallery and probe (shown in Table 4.2) are very similar and overlapping. The blue curve significantly differs from the rest and it is for a gallery and probe taken on the same day.

Figure 4.3 shows a set of intra/inter class variation diagrams. The Euclidian distance is plotted on the horizontal axis and the probability density on the vertical. The diagrams in Figure 4.3 (a) and (b) refer to different time periods and (c) refers to probe and gallery acquired within few minutes. All distributions are very similar. However, there is less overlap between the intra-class and inter-class distributions in (c).

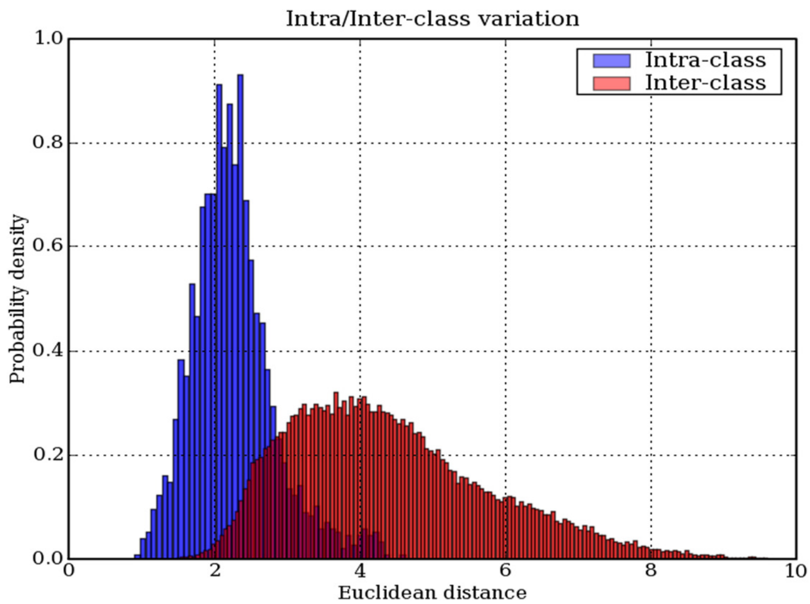
There is only a 5% drop in performance over 9 months when the combination of side, front and top is used. Decidability decreases and EER increases for analysis over time compared to analysis on the same day. There is less overlap between the distributions in Figure 4.3 (c) and more area under the ROC curve for data captured on the same day. It is not clear whether the slight drop in performance over time can be attributed to time (aging) itself or to covariate factor(s) that we have yet to consider. However, our hypothesis is that the variation of performance over time can be caused by change of clothing underneath the overall, as well as some change due to elapsed time. The best recognition is achieved when comparing similar temperature seasons (e.g. summer -> spring). Presumably, subjects wore similar style of clothes underneath the overall. Subjects wearing a body suit (without clothes underneath) was considered but discarded as it was considered impractical.

The values of CCR, d' and EER for the front and side views are similar over time. There is more dynamic gait information available when the side view is used, compared to the frontal view. Therefore, a question is raised how much of the dynamic information is used for recognition over time in the case of model-free approaches, such as the GEI and GEnI. Nevertheless, the results show that the top view is least suitable for recognition over time.

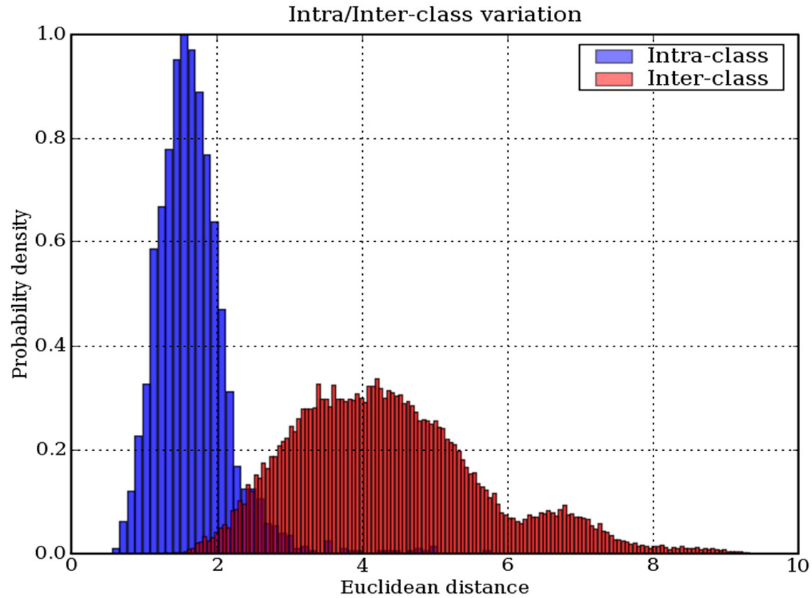
Achieving higher recognition rates and lower error rates could have been possible if more sophisticated classification techniques had been used. Nevertheless, the results presented in this section show that it is possible to recognize someone reliably by their gait after a certain time period has elapsed, if we were able to control several other essential parameters (like clothing and footwear). Although it is not possible to compare the results available in the literature due to significant difference in the datasets used, Figure 4.4 provides a clear indication that 'elapsed time' has been confused with other covariates so far in the literature. The lack of suitable dataset has made it difficult for other researchers to investigate the effect of elapsed time independently of other factors. The next sections demonstrate that other factors can significantly affect recognition. The conflating covariates have been the main reasons for low CCRs in the previous studies that have employed similar techniques to the ones used here.



(a) Time difference of 1 month



(b) Time difference of 9 months



(c) Time difference of few minutes

Figure 4.3 – The Intra-inter class variation for different time periods

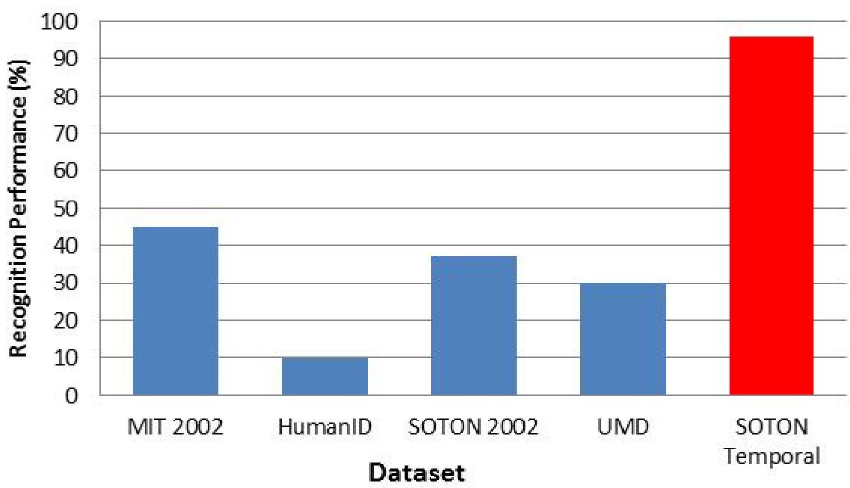


Figure 4.4 - Highest CCR achieved by studies over time using various datasets

4.4 The Effect of Clothing

The main purpose of the experiments presented in this section is to provide an indication of why previous studies have achieved significantly lower recognition rates over time, by employing similar techniques to this chapter. The results could be improved if algorithms that are less sensitive to change in clothing are used. However, the work presented in this chapter does not focus on improving performance of recognition approaches but on understanding the effect of time and other subject related covariates on the performance of the baseline algorithm and a more recent gait representation.

Clothes can affect the overall body shape and certain type clothing can affect the way a person walks. Previous studies have indicated that change in clothing can have a negative effect on the performance of current gait matchers. A study [11] has reported a recognition rate of just 3% for the combination of the following covariate factors: clothes, shoes and time. Another study [22] reported a significant drop in performance (87% to 60%) when subjects wore trench coat on top of their normal clothes.

The data acquired in months 9 and 12 contain samples of subjects wearing different types of clothes and enables us to perform analysis of different type of clothes over time and over few minutes. It enables us to quantify the effect of clothes, while keeping all other covariates unchanged. There are 4 different types of clothes, 3 types of 'ordinary' clothes of subject's choice and overalls provided as part of the database collection. Refer back to Chapter 3 for details. In this section, 'Type I' clothes refer to the samples collected in month 9. 'Type II' and 'Type III' refer to the samples collected in month 12.

Our clothing analysis consists of three experiments. In the first one we utilize the temporal data and perform matching over time with the subjects wearing different clothes. In experiment 2 we use the data acquired in month 9 to investigate the effects of clothing irrespectively of any other covariate and finally in experiment 3 we use the data collected in month 12 to further quantify the effect as it provides additional clothing types.

Experiment 1: The purpose of this experiment is to perform matching of samples acquired at different time periods with subjects wearing different clothes. In addition to the combinations of gallery and probe shown in Table 4.2 the combinations shown in Table 4.3 are used. Two-fold cross validation is performed. Only the highest CCRs achieved are shown in Figure 4.5. The highest value for the standard deviation of the CCR is 0.04. The results presented in Figure 4.5 indicate that different type of clothes can lead to a large drop in performance. The finding is similar for all signatures considered, but only the GEI (side) and GEI(S+F+T) are shown on the diagram for clarity. It is important to note that the drop in recognition is consistent with time.

Probe *10 samples/subject*	**Gallery** *10 samples/subject*	**Time Diff. (months)**
Type I – 21 subj.	Month 9 Overall – 21 subj.	0
Type I – 21 subj.	Month 4 Overall – 22 subj.	5
Type I – 21 subj.	Month 1 Overall – 23 subj.	8
Type I – 21 subj.	Month 0 Overall – 25 subj.	9

Table 4.3 – Combination of gallery and probe for the clothing experiment (part 1)

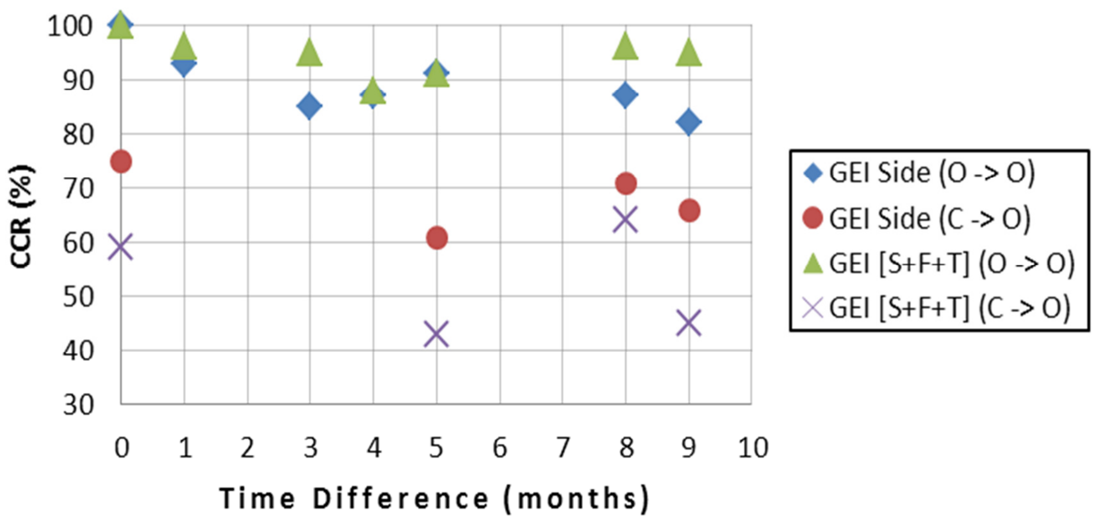


Figure 4.5 - The recognition performance over time. (O -> O = Overall - > Overall, C - > O = Normal Clothes -> Overall)

Experiment 2: The aim of this experiment is to investigate the effect of different clothes in the case when the time between recording the gallery and the probe is few minutes and no other changes in the experimental condition occur. Table 4.4

shows the combinations of probe and gallery used. Figure 4.6 summarizes the results. The horizontal axis shows the combination of clothing considered. *Type I* clothes refer to the samples of subjects wearing ‘normal’ clothes. Very high recognition results for all views are achieved if the same types of clothes are matched. However, when different types of clothes are used the recognition rates can fall to 40%.

Probe ***5 samples/subject***	**Gallery** ***5 samples/subject***	**Time Diff.**
Overall – 21 subj.	Type I – 21 subj.	Few min.
Type I – 21 subj.	Type I - 21 subj.	Few min.
Overall – 21 subj.	Overall – 21subj.	Few min.

Table 4.4 - Combination of gallery and probe for the clothing experiment (part 2)

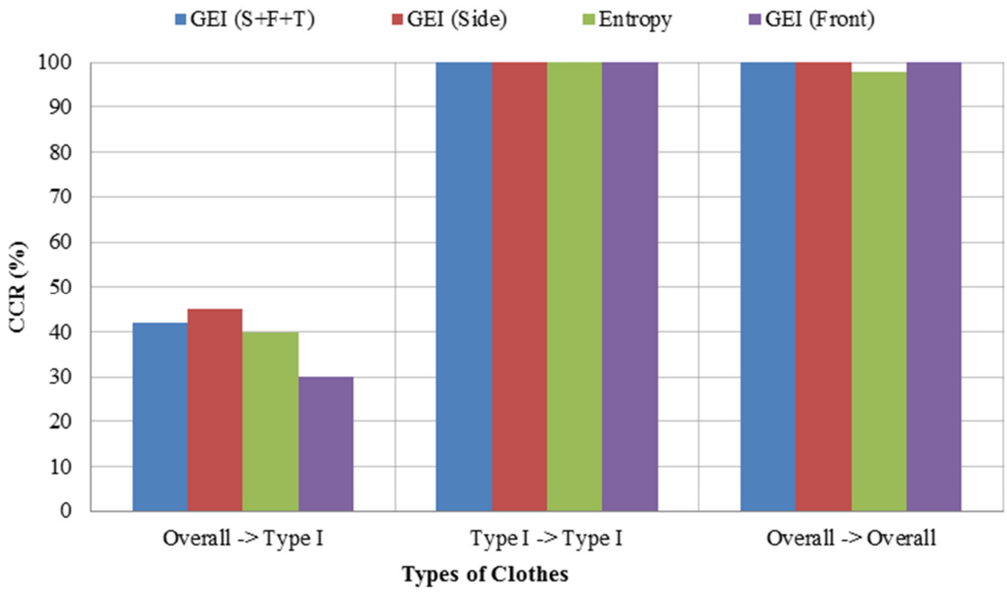


Figure 4.6 – The effect of clothes on the recognition performance. The probe and the gallery have been taken few minutes apart.

The yellow ROC curve in Figure 4.7 significantly differs from the others. Clothing change affects the curve’s shape dramatically. The remaining curves represent samples of subjects wearing the same clothes but recorded in different sessions.

Figure 4.8 shows the class distributions for the GEI (side) signature in the case of a) same clothes and b) different clothes.

The results demonstrate that high recognition rates can be achieved if clothes of the same type are matched and very low rates if ‘extreme’ changes in clothing occur. The experiment described in part 3 investigates how the recognition rate is affected by change from one type of ‘ordinary’ clothes to another, as opposed to ‘extreme’ clothing variations such as the overall suit.

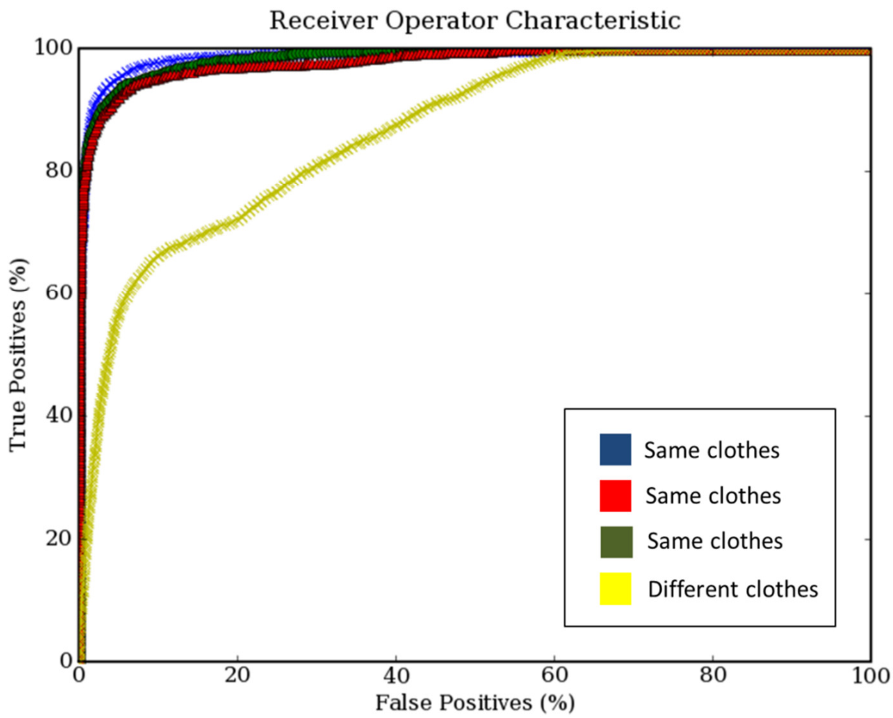
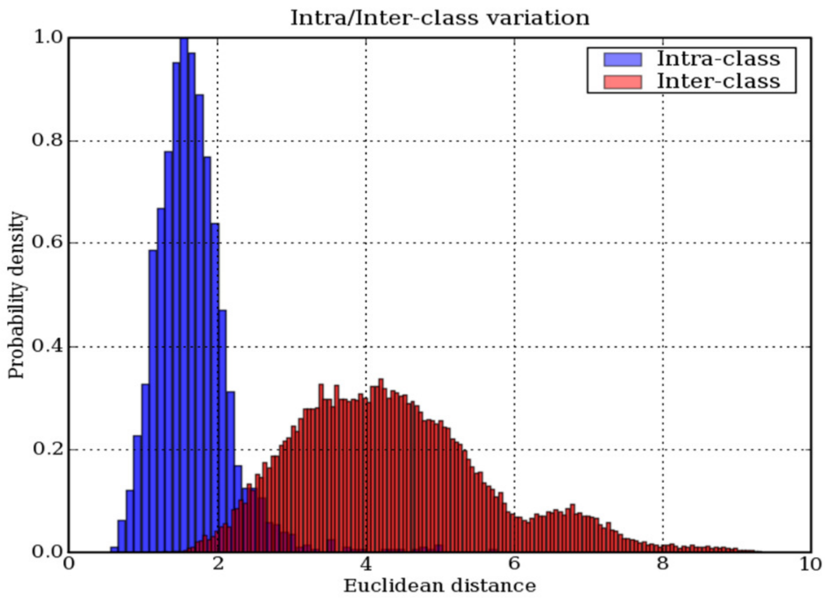
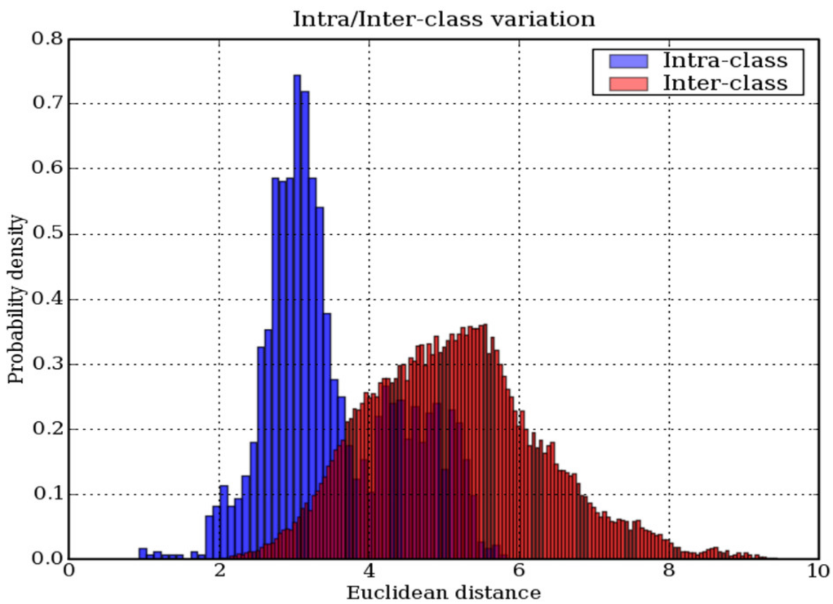


Figure 4.7 – ROC curve quantifying the effect of clothing



a) GEI (side) – Same clothes

d’ = 2.71, CCR (k=1) = 99%, CCR (k=3) = 99%



b) GEI (side) – Different clothes

d’ = 1.57, CCR (k=1) = 44%, CCR (k=3) = 45%

Figure 4.8 – The inter/intra class variations for different types of clothes

Experiment 3: We use the data captured in month 12 for this part of the experiment. Table 4.5 shows the combination of probe and gallery used. Figure 4.9 summarizes the results. The EER for different gait signatures is shown in Figure 4.10. It can be clearly seen that the EER increases if the subjects wear clothes that significantly differ from each other. The *d'* decreases in this case. The results of this experiment show that the recognition rates decrease significantly less if different but less extreme clothing changes occur.

In conclusion, the recognition can be affected significantly due to clothes change regardless of elapsed time or any other covariates. Figure 4.11 shows how a gait signature can be affected when there is a significant change of clothing. The drop in performance is less significant if subjects are wearing similar types of clothes. Another study [22] performed on a different database came to a similar conclusion for a model-based approach. The CCR dropped significantly when subjects wore trench coats. This might appear counter intuitive as model based approaches model the gait dynamics and are considered to be better at handling appearance change caused by clothing. However, fitting a model in the case of trench coat is not a trivial task. The results presented in this section are similar for the other studies that have considered the effect of elapsed time. This is an indication that elapsed time has been confused with other covariates (mainly clothing) so far in the literature.

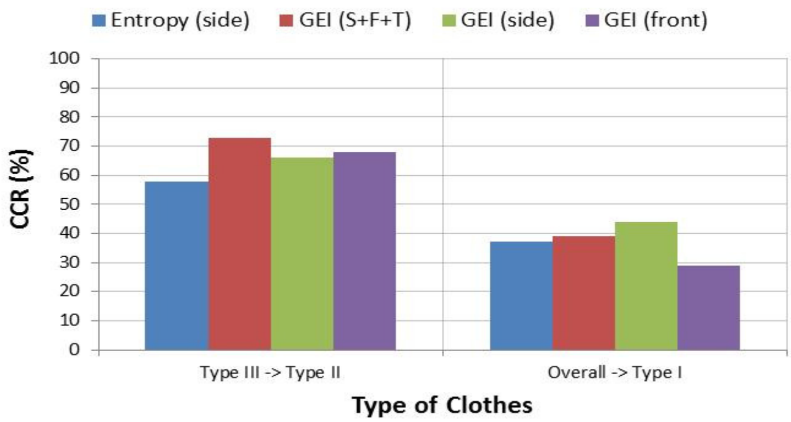


Figure 4.9 – The effect of different types of clothes on the recognition performance for various gait matchers

Probe *10 samples/subject*	Gallery *10 samples/subject*	Time Diff.
Type III – 18 subj.	Type II – 18 subj.	Few min.
Overall – 18 subj.	Type I – 18 subj.	Few min.

Table 4.5 - Combination of gallery and probe for the clothing experiment (part 3)

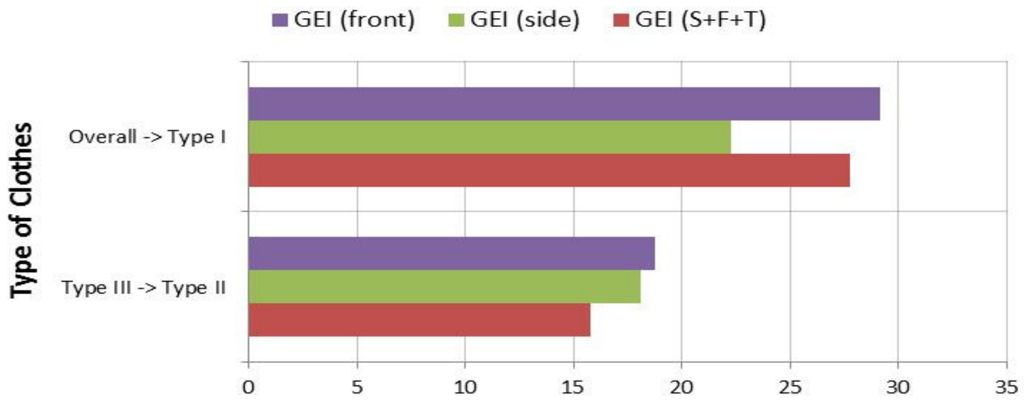


Figure 4.10 - The EER for different types of clothes

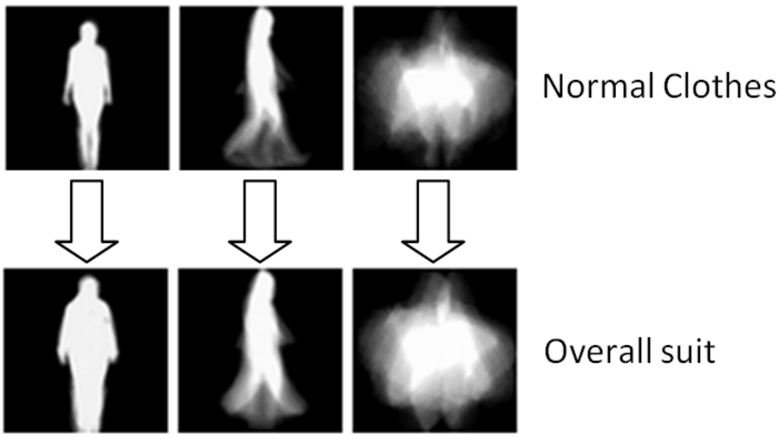


Figure 4.11 – The effect of clothing on gait signatures

4.5 Footwear

An experiment relating to the effect of footwear on gait recognition performance is presented in this section. This experiment has been performed to provide an indication whether footwear is likely to have a significant effect. Further investigation using a larger dataset containing a larger variation of shoes is needed.

The combination of probe and gallery used is shown in Table 4.6. The data captured in month 12 is used. From Table 4.6 it can be concluded that wearing footwear is unlikely to drastically affect recognition. However, there is an indication that the EER can increase (Figure 4.12).

Probe *5 samples/subject*	**Gallery** *5 samples/subject*	**Lowest/Highest CCR**
No shoes – 18 subj.	No shoes – 18 subj.	100/100
No shoes – 18 subj.	With shoes – 18 subj.	97/100

Table 4.6 – The combination of gallery and probe for the footwear experiment

A study [22] employing a model-based approach reached similar result to ours. The CCR did not decrease dramatically when subjects' shoe type changed (except in the case of flip-flops). Another study [11] using the GEI representation has also achieved a relatively high recognition rate for a change in footwear.

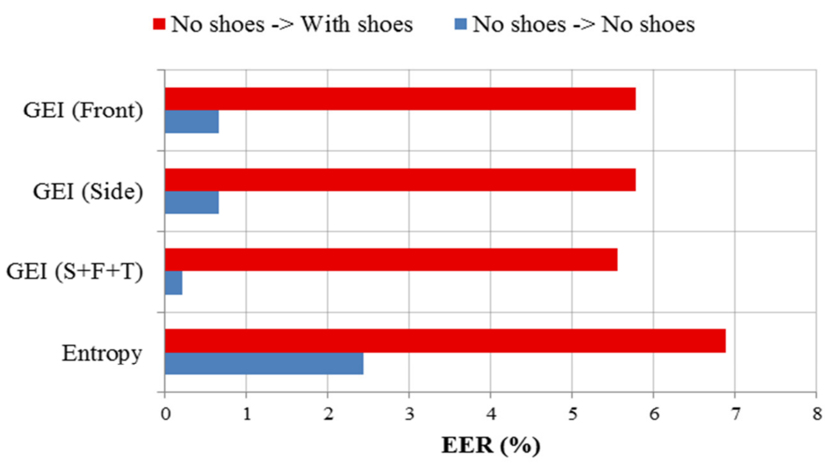


Figure 4.12 – The Equal Error Rates (EER) for the footwear experiment

4.6 Speed

A study [63] demonstrated the need for speed adjustment of gait features. A different study [22] showed that variation in walking speed can have a negative effect on the performance of model-based approaches. The recognition performance decreased from 87% to 60%.

In this section we use the combination of gallery and probe as shown in Table 4.7. Each gallery and probe consists of 110 samples and all samples are taken only few minutes apart. Other covariates remain unchanged. The walking speed was not controlled; we are just utilizing the natural variation in the subject's speed to perform this experiment. Each acquisition session consists of 20 samples per subject. For each subject, ten samples with similar walking speed were identified and five were added to 'Gallery 1' and the other five to 'Probe 1'. In addition, ten samples of different walking speed were identified for each subject and five added to 'Gallery 2' and the other five to 'Probe 2. In this experiment we are interested in the relative change of walking speed between different samples of the same subject. The smallest variation in walking speed for a subject is 5.8% and the greatest variation for a particular subject is 25%. This data applies for the acquisition session taken in 'month 1'. However, it is anticipated that the variations are similar for the other sessions because there were no changes in the experimental conditions. The results are presented in Figure 4.13. There is a slight drop in performance for all signatures, except for *GEI (Front)*. It would appear that the side view is mostly affected by change of speed which is somewhat expected.

Recently, Aqmar *et al* [38] proposed a gait representation that is robust to variations in speed.

Probe *5 samples/subject*	**Gallery** *5 samples/subject*	**Relative Speed**
Probe 1 (22 subj.)	Gallery 1 (22 subj.)	Similar
Probe 2 (22 subj.)	Gallery 2 (22 subj.)	Different

Table 4.7 - The combination of gallery and probe for the speed experiment

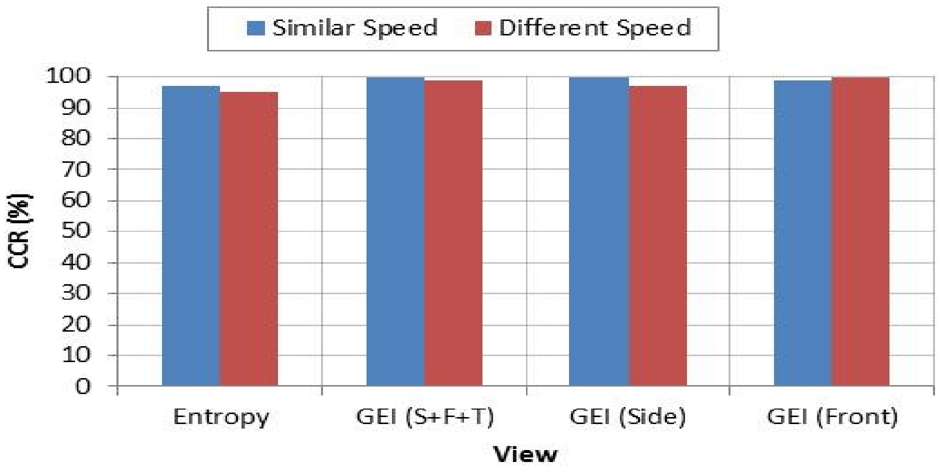


Figure 4.13 – The results of the walking speed experiment

4.7 No of Features v.s. Performance over Time

A unique advantage of gait is the capability to perform recognition at a distance, at a low resolution. Therefore, it is important to show how the performance changes with number of features. In this chapter, the number of features equates to the number of pixels in the GEI or GEnI signature. If gait is captured at a distance there are fewer pixels in the image that can be used for recognition. Consequently a decrease in the number of features is equivalent to an increase in the distance at which the subject is observed. The rate at which the performance deteriorates with smaller number of features should be consistent over time. In other words, if the gallery and probe are acquired m month apart and a recognition performance of p is achieved with n number of features then a $p \pm q$ (where q is very small) performance for the same number of features n should be expected if the gallery and probe are acquired $m+o$ months apart (o any value).

We have performed an experiment using GEI (side view) to show whether time has any effect on the ability to do gait recognition at a lower resolution. The findings are presented in Figure 4.14. We have shown that irrespective of the time difference (except when the time difference is few minutes) between the probe and the gallery the degradation of performance as the number of features decrease is fairly consistent. This is evidence that gait recognition can be used as a reliable biometrics at a distance over time. Figure 4.14 shows that high performance can be achieved by using a small number of features. Using only 900 features (30x30) a CCR of 80%-90% can be achieved. However, a larger database is needed to correctly verify the number of features *n* needed to achieve a recognition *p.* There is an on-going work in developing gait recognition algorithms capable of achieving high recognition rates at low resolutions [30].The main aim of this experiment is to show that elapsed time does not play an important role to achieve recognition at various resolutions.

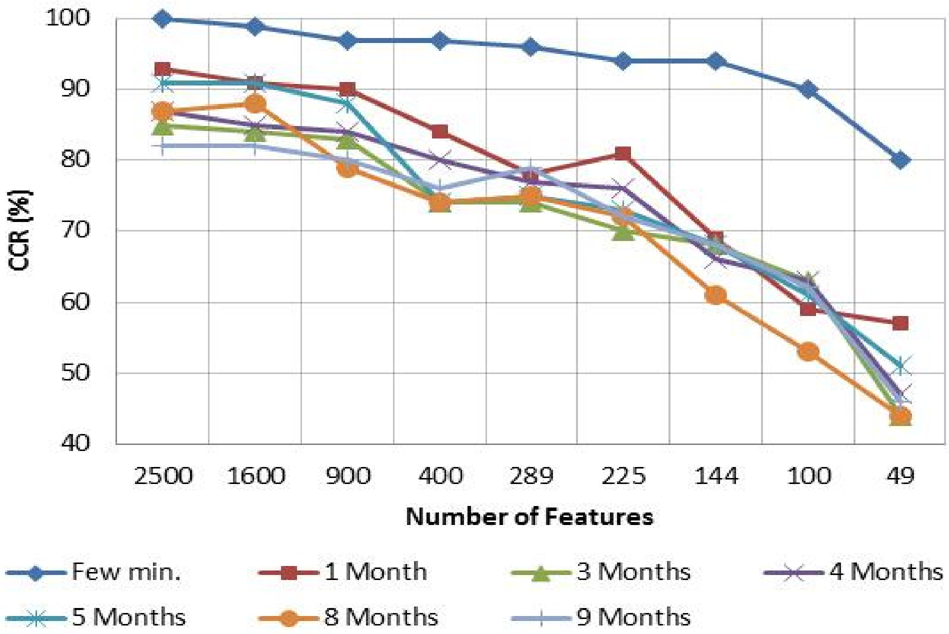


Figure 4.14 – Number of features vs. Performance, over time. The time periods refer to the elapsed time between capturing the gallery and the probe

4.8 Discussion

In this chapter, analyses of various subject-dependent covariate factors using existing approaches were performed. The results indicated that there has been a lack of understanding of elapsed time on the performance of gait recognition. The major obstacle for quantifying the effect of time and other factors independently is the lack of suitable databases and conflating more than one covariate between different acquisition sessions.

We have shown for the first time that gait per se is time-invariant in the short and short/medium term and, thus, can be used as a reliable biometric trait over time if influential covariant factors (namely clothing, footwear) were controllable. A similar recognition performance was achieved over 7 different time periods and a CCR of 95% is achieved over period of 9 months. We hypothesize that a CCR of nearly 100% could be achieved if various covariate factors (like clothing, footwear etc.) were controllable. We have shown that the major problem for recognition over time can be partially attributed to the change of clothes. We have shown a dramatic fall in performance in the case of 'extreme' changes in clothes. However, given that the problem of clothing can be overcome using techniques as in [40, 64], the issue of gaining/losing weight would still pose a challenge in the case of model-free approaches, such as the GEI. Further research is needed to identify an age-invariant feature sub-space for gait by performing feature set selection [40, 41].

A second significant finding is that we have shown that other authors reported low gait recognition since clothing and other covariates have been confused with "elapsed time" previously in the literature. The CCR drops from 100% to 30% for *non-temporal* combination of gallery and probe, which corresponds to the results achieved by other studies over time.

We have also confirmed that the drop in gait recognition performance at low resolution (at a distance) is consistent for all time periods considered. This is a very significant finding because recognition at a distance is a major advantage of gait over all other biometrics.

It would appear that shoes and speed are unlikely to affect recognition performance significantly in the case of model-free gait representations. However, it is important to note that the data collected to analyse the effect of these covariate factors can merely provide indication and not quantify the exact effect.

Experiments performed in challenging environment (e.g. outdoors) were not considered appropriate for the analyses presented in this chapter. Additional environment related covariates would be introduced and it would be difficult to quantify how much recognition changes due to the imperfections of the extracted silhouettes and how much due to elapsed time or other subject specific covariates. In the next chapter emphasis is placed on experiments in real world environments.

Chapter 5

Quality and Deployment

5.1 Introduction

In biometrics, the term "quality" is used to describe several different aspects of a biometric sample that contribute to the overall performance of a biometric system. Broadly speaking, a sample is of a good quality if it is suitable for automated matching. This may be different from a human perception of quality.

Quality measurement algorithms are increasingly deployed in operational biometric systems. However, quality measurement is difficult problem because it depends on the algorithm used. The ultimate intention of measuring and using quality is to improve system performance.

Quality can be used at various stages of a biometric system. Some examples are shown below:

- *Enrolment stage.* Quality can ensure that samples of poor quality are not stored in the database. Inserting poor quality samples in the database can adverse effect on the overall performance.
- *Matching stage.* A system may invoke a slower but more powerful matching algorithm when low quality scores are compared.
- *Decision stage.* A system can dynamically adjust the decision threshold if sample quality information is available.

5.2 Quality for gait recognition

Gait biometrics has some unique advantages over other forms of biometric identification and many studies have demonstrated that gait has the potential to become a powerful biometric for surveillance applications [3]. However, gait samples are usually of poor quality since they are acquired in uncontrolled environment and in sub-optimal conditions. Quality of biometric samples is an under-researched problem, especially in the case of gait biometrics.

A number of factors have been reported to affect gait recognition performance [11, 22, 55]. Some of the most commonly quoted are: viewing angle, shoe type, walking surface, carrying objects, elapsed time, change of clothing and poor signature quality. We have decided to consider majority of the key issues reported in the literature and as we shall find have a significant effect on recognition. We have performed experiments in both, controlled and real-world environment to allow us to gain comprehensive understanding.

In Chapter 4 we showed the results of experiments performed in a controlled environment. The results showed that the following factors mainly related to the *subject* can have a negative effect on the performance of a gait recognition system:

A) Change in clothing worn by a subject
B) Change in footwear worn by a subject
C) The speed at which a subject is walking
D) Camera resolution or the distance of the subject from the camera. The number of features or pixels that are available for a gait signature decreases as the distance increases.

Elapsed time was considered to a great extent and the results showed that time does not affect recognition in the short term, contrary to what was previously believed in the literature.

In this chapter we show how the following *environment* related factors can negatively affect recognition:

A) The angle at which the subject is walking relative to the camera (view point or pose)
B) The height of the camera i.e. the elevation (from the ground) at which a camera is placed.

In addition, these factors are confounded by errors in the *silhouette* segmentation process common to many recognition approaches. The silhouette extraction process separates the subject from the background and eliminates the colour and texture of the clothes. Background subtraction is not a perfect process and poor segmentation can result in the form of shadows, holes in the silhouettes, noisy contours, and undetected body parts, even when the video is of good quality. See Figure 5.1 for examples of signatures produced from poor quality samples. Appendix E shows some additional examples of poor quality silhouettes and signatures.

Deterioration in silhouette quality can be due to a number of factors. Some of the factors are related to the environment and some are related to the subject.

- *Environment:* Type of lighting, illumination and lighting changes, shadows and reflections caused by interaction of subject with environment, occlusion with objects, walking surface, walls' colour, etc.
- *Subject:* Self-occlusion, clothing, speed of walking.

Note that speed of walking or clothing have a different effect in this context than what was described in Chapter 4. For example, the walking speed affects the background subtraction algorithm since parameters such as learning rate can depend on the walking speed and if not correctly adjusted can lead to silhouettes of poor quality.

In most cases poor quality silhouettes result due to a combination of factors presented above. It is difficult to identify and measure how much each factor contributes when performing tests in arbitrary real-world locations.

There are continuing developments in silhouette extraction techniques. Our study is a baseline approach to understand basic performance and better techniques could be used throughout, not just in the case of silhouette extraction. However, this does not aid understanding of the impact of the environment on recognition.

Many approaches rely on clean silhouette data. However, a number of studies have established the negative effect of poor silhouette quality on recognition performance [31, 65]. Approaches such as [33] propose a robust gait representation that works better with partial silhouettes. A study [32] has proposed a way of reducing the noise around the contours. We are not aware of any work in the literature that uses quality metrics to improve a pre-processing step, though all approaches could benefit from improvement in silhouette quality.

In this chapter we describe extension to the existing quality metrics and propose novel ways of using the metrics to improve recognition in real-world environments. The potential of the techniques is demonstrated using the database presented in Chapter 3. The database has been collected specifically for the experiments presented in this chapter. We show that recognition can be improved by using quality metrics to improve background subtraction which is an important step in most gait recognition approaches. The results also show the benefit of using quality metrics to select the best quality gait cycle. In section 5.5 we show how other factors such as pose and camera height can affect performance. We introduce the notion of 'quality of matching' and experimentally demonstrate its importance. Finally, details of a gait recognition system operating in realistic environments are presented in section 5.7. This work was undertaken as part of the VSAR project. Working on this project allowed us to identify and solve some of the issues related to translating gait recognition to the real world. It also allowed us to perform comparison of recognition rates obtained by analysing a dataset and performing recognition in live use.

Figure 5.1 - Examples of poor quality signatures

5.3 Quality metrics

In this section we extend the existing quality metrics for gait recognition. We measure three different aspects of quality.

5.3.1 Measuring quality

Quality is determined either by comparing observed value with a model or using the actual value produced by the quality metric. In the case of metrics 1 and 3, quality score is produced by measuring the root mean square error and of the difference between the observed signal and a signal generated by a model. See Equation (5.1) and (5.2). The quality models describing signals of good quality are shown in the subsections. The error is:

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{n=1}^{N}\{y(t_n, \boldsymbol{w}) - x_n\}^2 \tag{5.1}$$

where *x* are the observed values, *N* is the number of frames, *w* is a nx1 vector containing coefficients of a quality model. For example, in the case of the quality model presented in equation (5.3) ***w*** is a 3x1 vector containing the values of *a*, *b* and *c*.

The RMS error is:

$$E_{RMS} = \sqrt{2E(\boldsymbol{w})/N} \tag{5.2}$$

In the case of metric 2, the actual value is an indicator of quality.

Other ways of measuring discrepancy between the data and an estimation model, such as squared error of prediction (SSE) were evaluated but provided no significant improvements.

5.3.2 Quality metric 1 – Exploiting periodicity

Human gait is a periodic motion. If there is little noise associated with binary silhouettes (e.g. in the form of shadows, reflections etc.) then the total number of white pixels forms a periodic signal. Models have been developed to represent this signal. An example is shown in (5.3) [31]:

$$y(t) = a + b\ cos(\omega t) + c\ cos(2\omega t) + z(t) \tag{5.3}$$

where $\omega = 2\pi/T$ and *T* is the gait period, *a* is the average area of a silhouette, *b* and *c* reflect the periodic variation in silhouette size, and *z*(*t*) is noise.

Measuring the error of fit against this model can be useful for detecting noise. Distorted silhouettes are likely to produce noisy foreground-sum signals.

Figure 5.2 and Figure 5.3 show examples of good and poor sample quality, respectively.

However, there are situations in which a signature of poor quality exists and using this model may be insufficient to detect it. For example if the top part of the silhouette is missing for all frames the resulting foreground-sum signal is likely to still fit the model in equation (5.3) very closely. Figure 5.4 shows an example of this which occurred when testing a fully automated recognition system. The fit would have been even better if the subject was walking at 90 degrees to the camera's field of view.

Nevertheless, using the periodic model gives a useful quality metric if combined with additional ones. A number of new possible metrics were evaluated and the most effective ones in combination with metric 1 are presented in the next two sections.

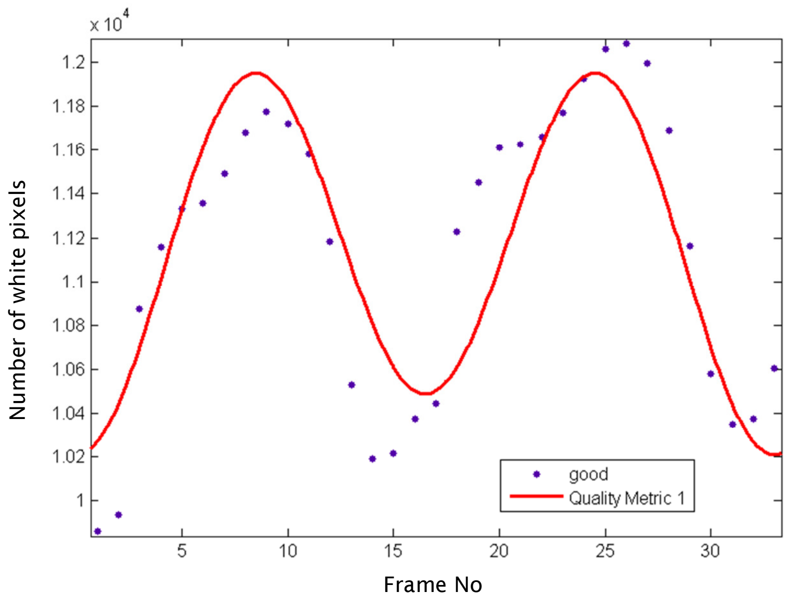


Figure 5.2 - Signature of good quality

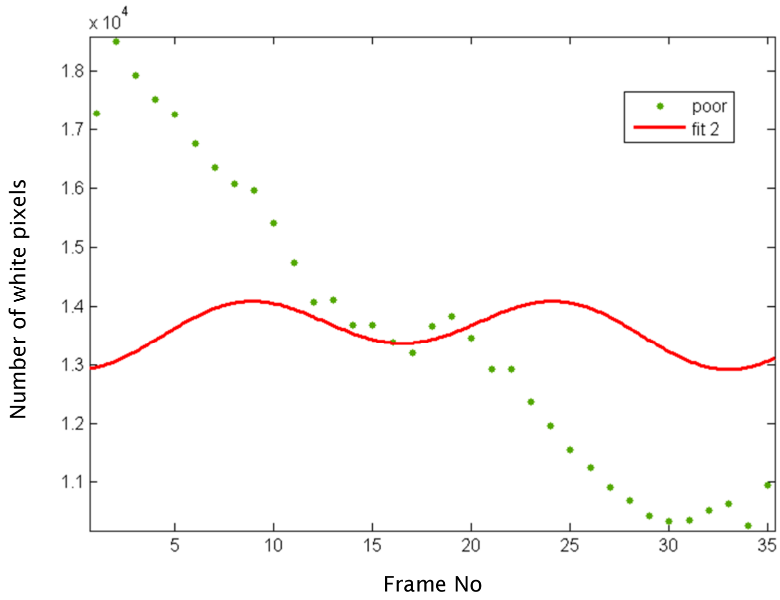


Figure 5.3 - Signature of poor quality

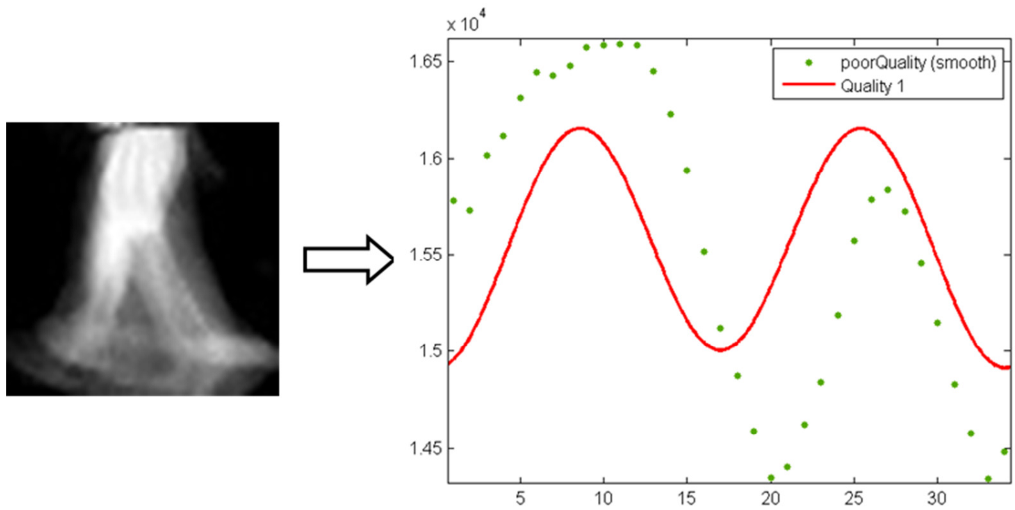


Figure 5.4 – Model fitting quality metric 1 (poor quality)

5.3.3 Quality metric 2 – Signature structure

This quality metric focuses on analysing the structure of a gait signature. The Gait Entropy Image [3] highlights the dynamic areas of the Gait Energy Image by calculating the Shannon entropy at each pixel:

$$H(x,y) = -\sum_{k=1}^{K} p_k(x,y)\, log_2\, p_k(x,y) \tag{5.4}$$

$p_k(x,y)$ is the probability that a pixel takes on the kth value. A binary image corresponds to $K = 2$.

The Gait Entropy Image contains pixels of high intensity values in the areas of the human body that are moving and low intensity values in the areas that are static. If silhouettes have been poorly segmented there will be entropy in areas that there should not be. There are a number of areas on the human body that can be tested for presence or absence of entropy. For example, there should be a small number of bright pixels in the head and torso region and large number of high intensity value pixels round the arms and the legs. Average anthropometric measurements of the human body [7] have been used to locate different parts of the body. See Appendix D for diagram of all average anthropometric measurements. Examples are shown in Figure 5.5.

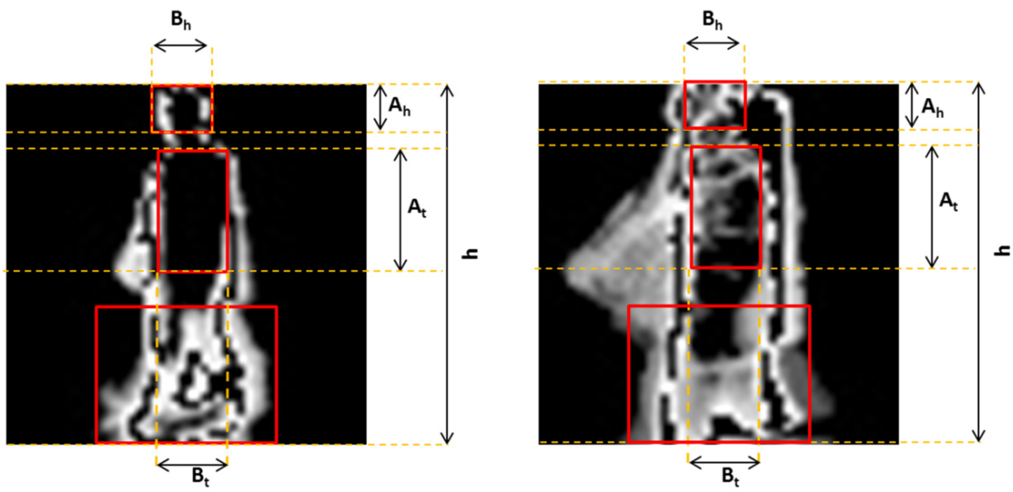


Figure 5.5 – Example for quality metric 2

Equation (5.5) shows an example of a simple way of measuring the presence or absence of entropy:

$$m_2 = \sum_{x \in A} \sum_{y \in B} H(x, y) \qquad (5.5)$$

where for height *h* and centre *c*, the thorax is between $A_t=[c - 0.174h/2,\ c + 0.174h/2]$ and at height between $B_t=[0.182h,\ 0.53h]$. The head region is between $A_h=[0,\ 0.130h]$ and $B_h=[c - 0.129h/2,\ c + 0.129h/2]$.

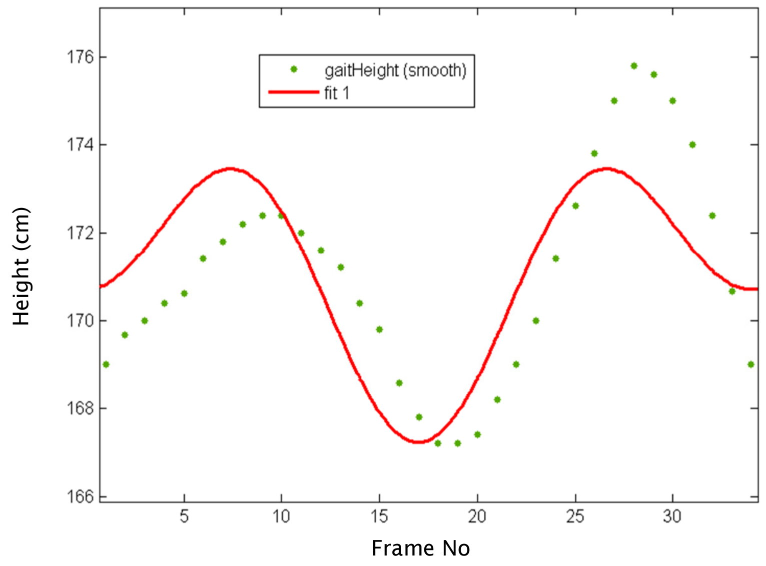
For the thorax and head region, low values for the entropic quality measure m_2 indicate a signature of good quality.

5.3.4 Quality metric 3 – Use of height

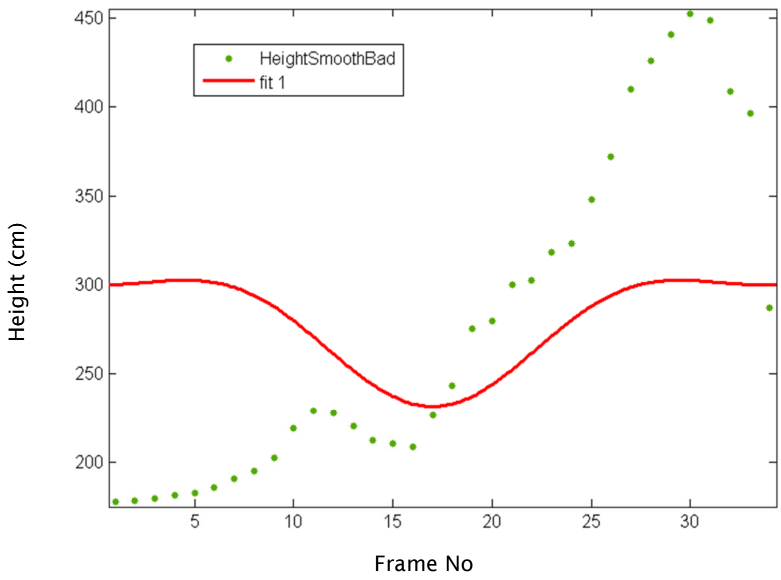
To establish temporal correspondence, the height of the bounding box formed around the subject's silhouette can be analysed over a gait cycle or walking sequence to reveal any anomalies in the process of subject-background separation.

A model can be fitted to the height signal and the error of fit can reveal poor quality signatures. Similar models as equation (5.3) can be used.

If calibration data is available the actual height of the subject can be used. Examples are shown below in Figure 5.6. Part a) of the figure represents a sample of good quality and part b) shows a sample of poor quality. The red line represents the line of best fit.



(a) Good quality



(b) Poor quality

Figure 5.6 – Model fitting for metric 3

5.3.5 Evaluating the metric's effectiveness

After a quality metric has been derived it is important to evaluate the usefulness of each quality metric. The quality metrics are assessed by their ability to discriminate accurately between samples that have been marked to be of good and poor quality. We use Fisher Linear Discriminant:

$$f' = \frac{(\mu_{good\_qualty} - \mu_{poor\_quality})^2}{\sigma^2_{good\_qualty} + \sigma^2_{poor\_qualty}} \tag{5.6}$$

where $\sigma^2_{good\_qualty}$ and μ_{good_qualty} are the variance and the mean for the good quality samples and $\sigma^2_{poor\_qualty}$ and $\mu_{poor\_quality}$ are the variance and the mean for the poor quality samples. High values of **f'** indicate a good separation and therefore a potentially useful quality metric.

Combining scores of multiple quality metrics requires a process of normalisation. For a vector of quality scores ***v***, the normalised version ***v'*** can be calculated using equation (5.7):

$$\boldsymbol{v}'(i) = \frac{\boldsymbol{v}(i) - \min(\boldsymbol{v})}{\max(\boldsymbol{v}) - \min(\boldsymbol{v})} \tag{5.7}$$

Figure 5.7 shows an example plot of the quality values as generated by the quality metrics presented in the subsections. The samples labelled blue are of good quality and the samples labelled red are of poor quality. The sample of best quality is the one with the lowest Euclidian distance to the origin of the quality feature space. A sample has been labelled to be of good quality if it does not contain significant amount of: shadows, reflections, missing body parts and holes in the silhouettes.

The way in which deploy the quality does not require threshold values that indicate whether a sample is of good or poor quality. In our work we are interested in the best value out of all instances considered. In some cases, for instance where quality is used as a measure of confidence in the result, threshold values can be derived by using machine learning techniques.

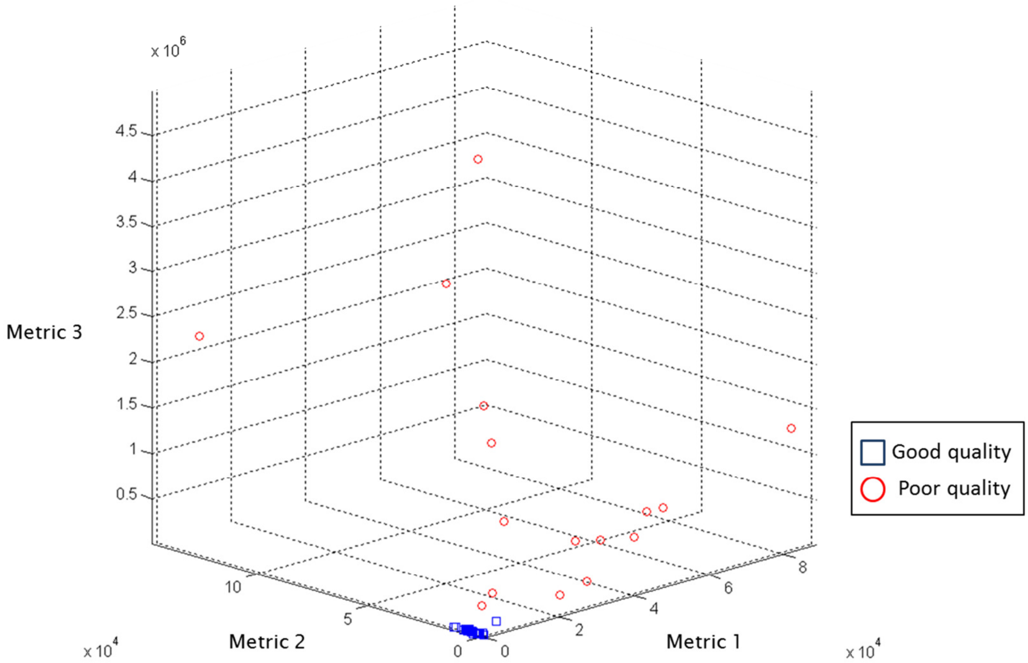


Figure 5.7 – Quality feature space

5.4 Deploying quality metrics

Our experience of evaluating the database and by testing an automated system shows that acquiring a signature of good quality is very difficult in a realistic environment. Changes in the operational environment of a system can result in different quality of biometric samples. The changes are even more significant in the case of matching across different locations and across time. We deploy quality metrics in two different ways as described in the next two subsections. We use the dataset presented in Chapter 3, section 3.4.

5.4.1 Improvement in background subtraction

Background subtraction is the first pre-processing step for most gait recognition approaches. The success of this process influences the outcome of any further processing. A perfect background subtraction algorithm does not exist and the inaccuracies can be significant in realistic environments. An example is shown in Figure 5.8.

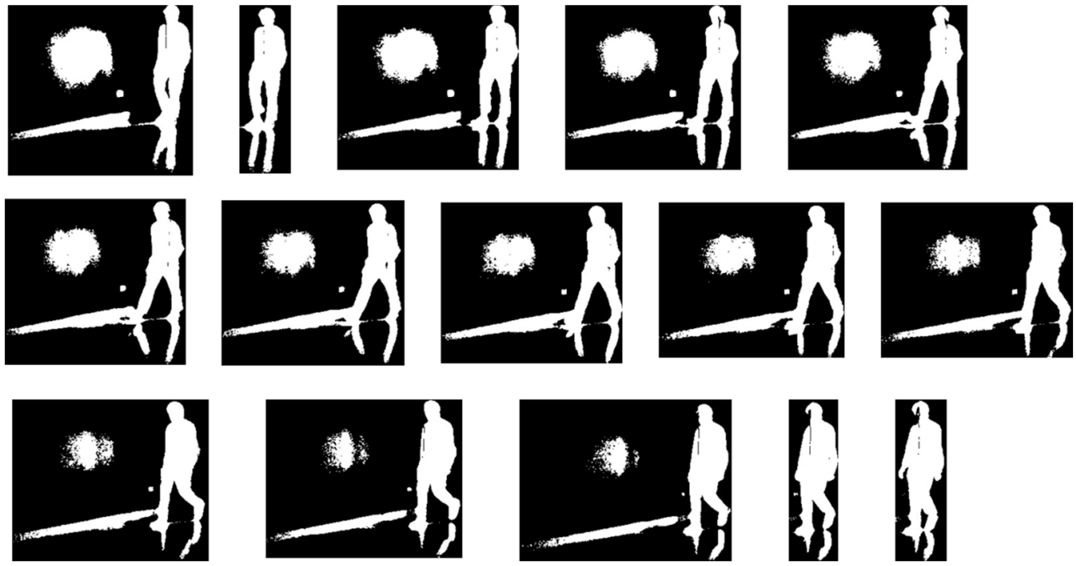


Figure 5.8 - Example of poor quality silhouettes for a single gait cycle

There are approaches that measure how successful a background subtraction algorithm has performed [66]. There are numerous post processing techniques that can improve the quality of background subtraction [67]. However, for gait recognition we are interested specifically in the quality of the signature and not the overall success of the algorithm. In addition, there are cases in which a poor quality signature will occur after improved techniques have been applied.

Matching signatures of inconsistent qualities can result in poor recognition performance. The aim is to perform the background subtraction process with the parameters that produce the best quality signature. This is achieved by dynamically changing some of the key parameters. We use Gaussian Mixture Model based approach for background subtraction and some of the key parameters for the algorithm are the '*threshold*' and the '*learning rate*'. Details of the background subtraction algorithm that we have used are shown in Appendix A. The range for the learning rate parameter (α) has been determined by using:

$$\alpha = 1 - e^{\frac{\ln 0.9}{x}} \tag{5.8}$$

where x is the number of frames before the algorithm considers the object to be part of the background. The value for x has been defined as function of the gait period. The range of values for x is between T and $4T$, where T is the gait period. The equation has been derived using equation (A.4). The object should be static for approximately $\log(1-c_f) / \log(1-\alpha)$ frames, where c_f has been set to the value of 0.1 and represents the proportion of data than can belong to the foreground object without affecting the background model. The range of values for the *threshold* parameter has been determined experimentally.

Figure 5.12 shows examples of silhouettes of different qualities produced by different values of the parameters. Figure 5.9 shows a block diagram of the proposed system that uses gait signature quality assessment in order to improve signature quality by varying some key background subtraction parameters. After a gait signature is produced, the quality is assessed using the three quality metrics presented in section 5.3. If a signature is below a pre-defined quality threshold for any of the quality metrics, then a new signature is produced by varying the background subtraction parameters until a signature of desired quality is obtained. In theory it is possible to find parameters that will produce an optimal gait signature. In practice however, only a limited number of parameters can be tested due to the computational complexity of most background subtraction algorithms. Additional work is needed in understanding how to make the process shown in Figure 5.9 automatically converge. In this thesis we have used a limited number of pre-defined set of parameters and terminated the process when a signature of good quality is found. If a signature of pre-defined quality could not be found, the best signature is chosen in terms of Euclidian distance.

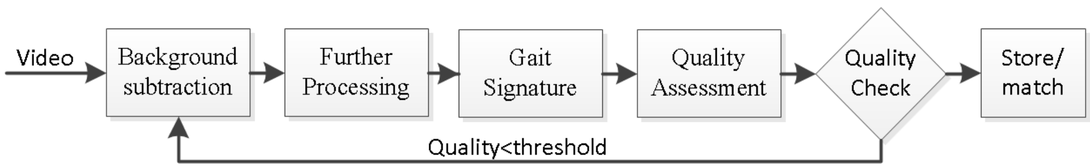


Figure 5.9 - Block diagram of a system using quality to improve background subtraction

An experiment was performed to understand the impact of using quality improvements on performance. Twenty-three subjects (2 samples per subject) were used. All samples were captured on the same day and no additional subject dependent covariates were introduced. The combination of gallery and probe is shown in Table 5.1. The system operated in recognition mode.

Figure 5.11 shows example signatures (of the same subject) before and after quality metrics were applied. The figure demonstrated the benefit of using quality metrics in a fully automated gait recognition system for matching across locations. Further examples are shown in Appendix E.

The results (shown in Table 5.1) reveal improvement in performance, both in the case of samples acquired at the same location and at diffierent locations when quality metrics are used. The CCR rate were on average 20-30% higher. A ROC curve for the last two entries of Table 1 is shown in Figure 5.10. The equal error rate (EER) decreased from 20% to 5% for matching across locations if quality is used.

The Gait Entropy Image (GEnI) performed slightly worse than the Gait Energy image (GEI) for the experiments described in this section. We hypothesise that poor quality silhouettes affect GEnI's performance much more than in the case of GEI. Fusing the subject's actual height resulted in approximately 10% higher CCR.

The high recognition rates achieved in the case of the same location are not necessarily due to matching signatures of high quality, but because of matching signatures of similar and consistent quality. It is possible to achieve high recognition rates when using signatures of poor quality only in situations where the operational environment of the system has remained unchanged between acquiring the gallery and the probe. When performing recognition in realistic environments, over time and across location many environmental factors change. The results presented in row three and four in Table 5.1 confirm this.

Probe Loc. (Dir.)	Gallery Loc. (Dir.)	Cam No	Quality used	CCR (%)
Corridor (IB)	Corridor (IB)	1	No	74
Room (IB)	Room (IB)	1	No	91
Room (IB)	Corridor (IB)	1	No	21
Corridor (IB)	Room (IB)	1	No	22
Corridor (IB)	Room (IB)	1	Yes	40
Corridor (HC)	Room (IHC)	3	No	0
Corridor (HC)	Room (IHC)	3	Yes	30
Corridor (HC)	Corridor (HC)	3	No	65
Corridor (HC)	Corridor (HC)	3	Yes	95

Table 5.1 – Combinations of gallery and probe

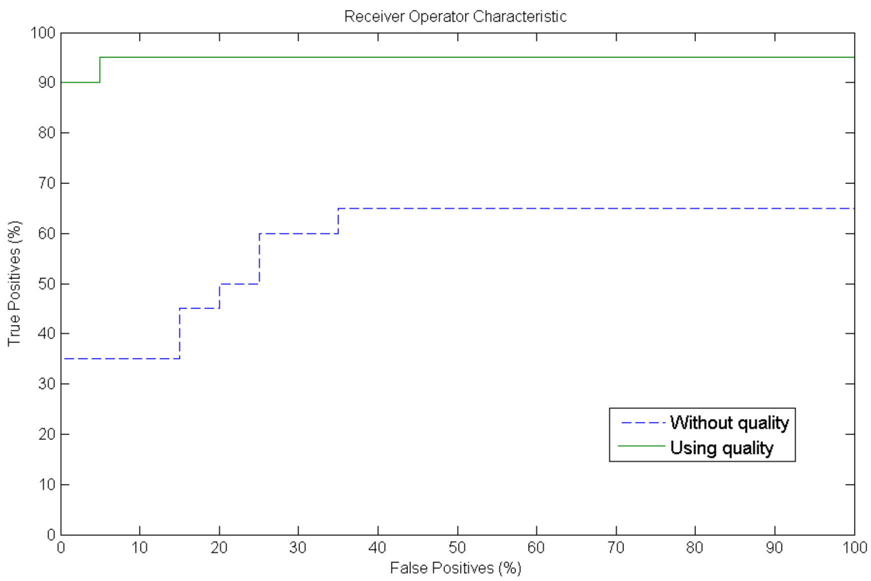


Figure 5.10-ROC analysis of quality improvement

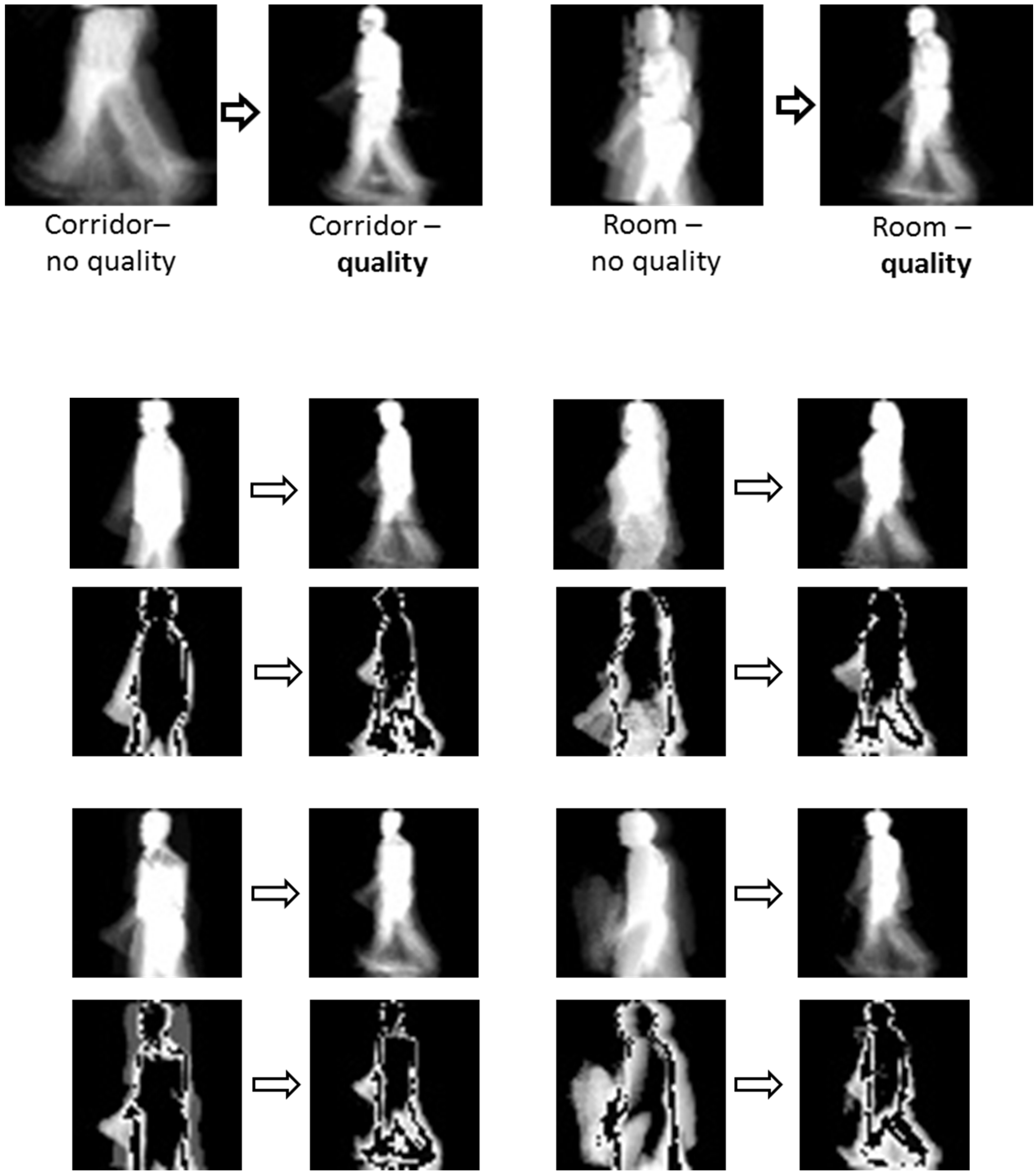


Figure 5.11 – Examples of signature quality improvement

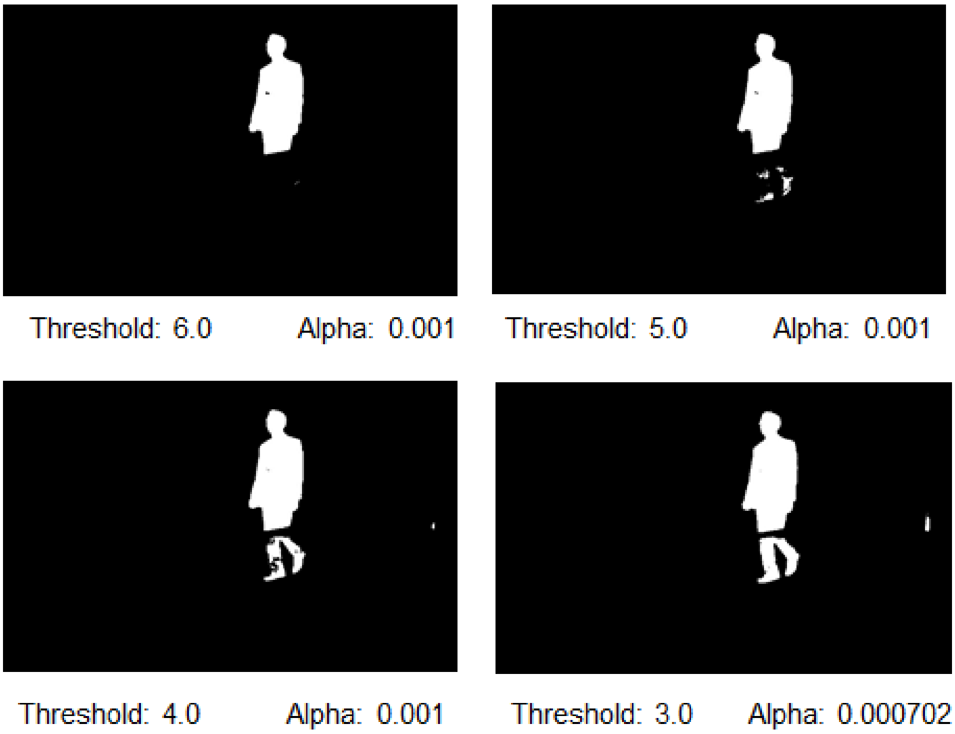


Figure 5.12 - Silhouettes produced with different background subtraction parameters

A further improvement in signature quality would be achieved by optimizing a larger number of background subtraction parameters. It might be possible to apply an optimisation algorithm to tune the parameters and make the process converge. However, the purpose of our work was to investigate whether it is beneficial to use quality metrics in automatically adjusting some of the key pre-processing steps in order to improve recognition.

5.4.2 Selection of the best gait cycle

Selection of a gait cycle is an important step in gait recognition for many approaches. A robust way of estimating the gait period is essential before a cycle is selected. In Chapter 3 we showed that tracking the position of the top of the head provides a robust way of doing this.

There can be many possible gait cycles available for a given video sequence. More specifically, there are L_c-t potentially valid cycles, where L_c is the length of the clean signal i.e. number of frames where the subject is fully visible and *t* is the gait period. A pseudo code of the algorithm is shown below:

```
i ← frame_No_start_of_clean_signal
while ( i < Lc-t)
     signature ← produceSignature(i..t+i)
     quality_scores[i]= measureQuality(signature)
     i ← i + 1
end_while
result = find_position(quality_scores)
start_frame_best_cycle = result
end_frame_best_cycle = start_frame_best_cycle + t
```

The '*produceSignature(a..b)*' function computes a gait signature for the set of frames located between frame number *a* and *b*. The '*quality_scores*' variable is an array of *nx1* vectors, where *n* is the number of quality metrics used. The '*result*' variable holds the position (frame number) of the vector which contains the best quality values.

A gait cycle selection algorithm that utilizes quality metrics can lead to the best quality cycle being chosen. Figure 5.13 shows the quality values of all available gait cycles for a given sequence. The '*start*' point refers to the frames in the range of [0..t] and the '*end*' point to the frames in the range of [L_c .. L_c-t]. The optimal cycle is at point No. 1 as this point has the lowest Euclidian distance to the origin. In this case, low values of each metric signify a sample of good quality. Figure 5.13 show how the signature quality varies for different starting frames. The quality improves from point 'start' until point No. 1 is reaches. The quality deteriorates as starting frames at points 2 and 3 are selected.

An experiment was performed to show the potential benefits of deploying quality metrics for selecting the most appropriate gait cycle. The combination of gallery and probe is shown in Table 5.2. Samples from twenty-three subjects captured by 'Camera 1' were used. There were no changes in the experimental conditions between the two combinations. The results reveal an improvement in performance.

Probe		Gallery		Quality	CCR (%)
Loc.	Dir	Loc.	Dir		
Corridor	IB	Corridor	IB	No	60
Corridor	IB	Corridor	IB	Yes	75

Table 5.2 – Combinations of gallery and probe

In some cases the camera to subject position is such that choice of different gait cycle results in slightly different pose. If quality driven gait cycle selection is used the pose between the gallery and the probe may differ in some cases. In these cases, the benefit of improvement in quality can be cancelled by degradation in performance caused by pose change between the gallery and the probe.

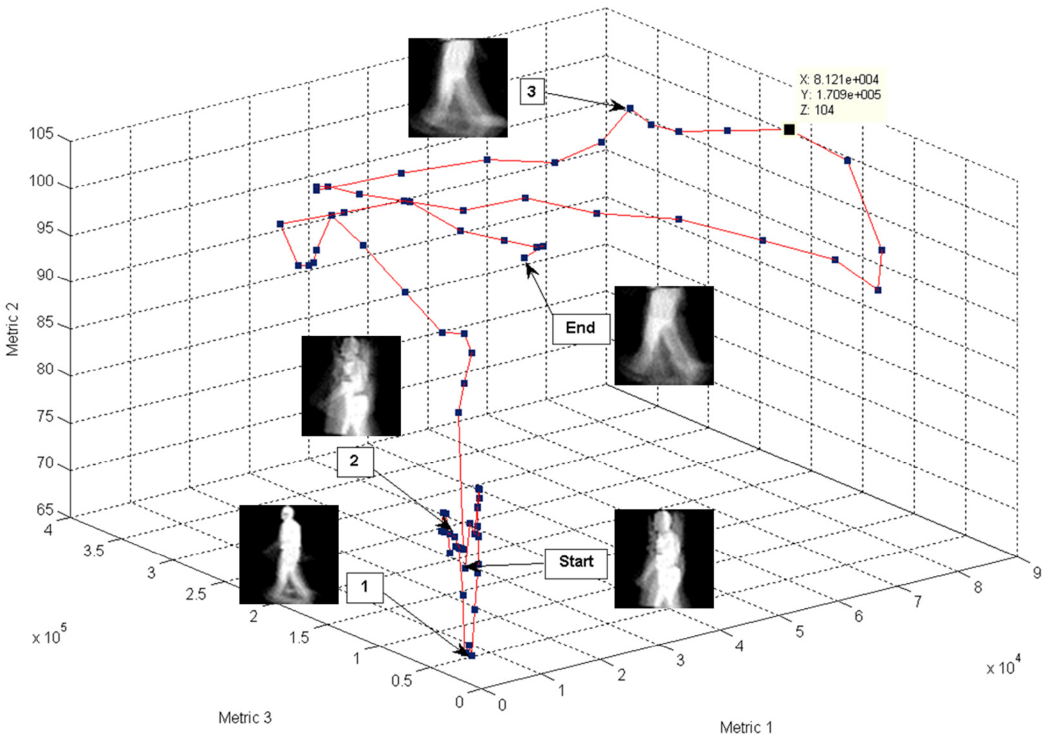


Figure 5.13 - The quality values for different gait cycles

5.5 Additional factors affecting performance

Previous work [31] has concluded that the problems with distorted silhouettes is one of the most important factors influencing gait recognition, but the poor recognition performance in challenging environments cannot be solved simply by using clean silhouettes.

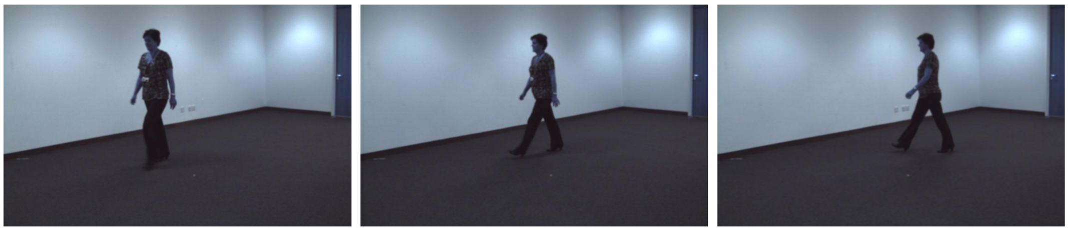
The results presented in the previous sections demonstrate the importance of using sample quality for recognition. However, there are other factors which can affect recognition performance that are not directly linked to quality of silhouettes. We presented a number of subject-specific factors in Chapter 4. This section presents additional factors related to the environment.

5.5.1 The effect of pose

There are approaches in the literature that quantify the effect of pose and deal with the issues of pose change as described in Chapter 2. We have included this experiment to demonstrate that quality of the signature on its own is not sufficient for achieving good recognition rates.

Samples from twenty subjects acquired in the room (camera 1) were used for this experiment. Figure 5.14 shows examples of samples. Samples of good quality were used for this experiment. We have used the baseline algorithm and have obtained results which are consistent with the literature. The performance deterioration is proportional to the change in pose. Highest CCR are achieved when samples of the same pose are used and the lowest CCR when samples of significantly different poses are matched. Figure 5.15 shows the details of the results obtained.

The results also show that highest recognition rates are achieved when subjects walk at 90 degrees to the camera's field of view. Experiments on larger dataset are needed to verify this hypothesis.



Direction – **IHC** Direction – **ID** Direction - **JED**

Figure 5.14 - Examples of samples for the pose experiment

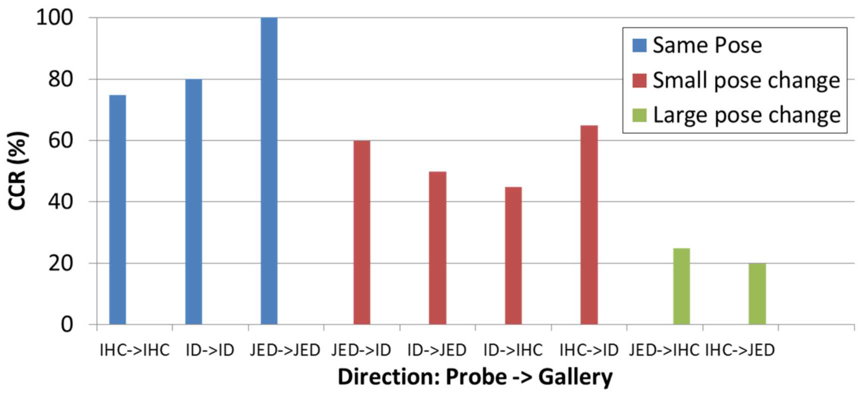


Figure 5.15 - The result of the pose experiment

Examples of signatures for samples acquired at different pose are shown in Figure 5.16. The examples indicate that both the static and dynamic areas of the signatures can significantly differ when a change of pose exists.

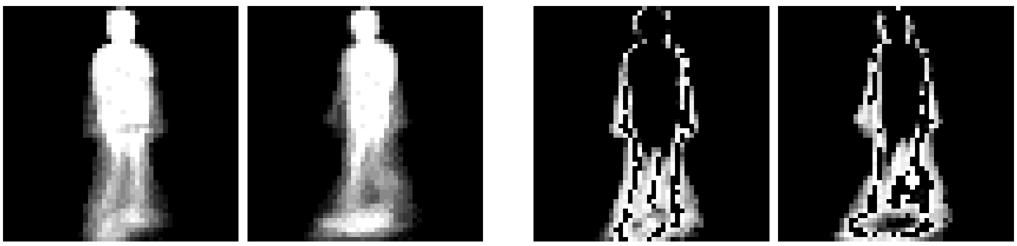


Figure 5.16 - Example of signatures for samples at different pose

5.5.2 Camera height

An experiment was performed to investigate the effect camera elevation. There are no experiments reported in the literature related to camera elevation while keeping all other covariates the same.

The samples recorded in the ‘Room’ location were used. One sample per subject for each of the gallery and the probe was used. Twenty-one subjects were used for this experiment. The combination of the gallery and probe is shown in Table 5.3. ‘Camera 1’ and ‘Camera 2; are placed at exactly the same location but ‘Camera 2’ is placed approximately 1m higher than ‘Camera 1’. The results are presented in Figure 5.17 and examples of the signatures produced by the cameras are shown in Figure 5.18.

Probe		Gallery	
Camera	Direction	Camera	Direction
1	BI	1	BI
2	BI	2	BI
1	BI	2	BI
2	BI	1	BI

Table 5.3 - The combination of gallery and probe for the camera height experiment

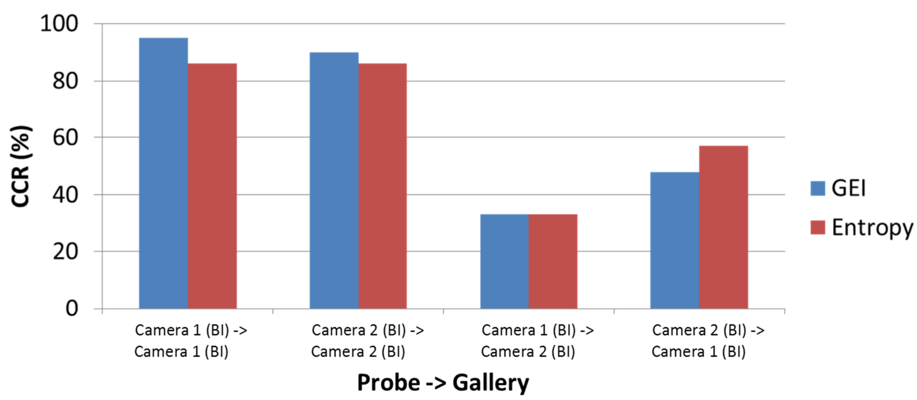
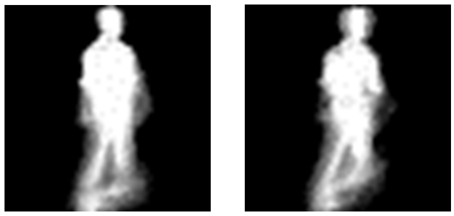


Figure 5.17 - The results of the camera height experiment



Camera 1 Camera 2

Figure 5.18 - Signatures produced from cameras at different heights

The results reveal degradation in performance when samples acquired from cameras placed at different heights are used. The results are similar to the ones demonstrated in the previous section for the case of ‘small pose change’.

5.6 Quality of matching

Our work distinguishes between two types of biometric “quality”:

- **Sample quality** for the probe and gallery used in matching.
- **Variability between samples**, due to factors such as the variability of the characteristics being measured or variations in correspondence between the probe and gallery samples

Sections 5.3 and 5.4 referred to the first type of quality and this section we will show the importance of the second type of quality. The results in section 5.5 show that there are number of factors not related to quality of silhouettes that can negatively affect recognition performance.

In this section we are introducing the notion of ‘Quality of Matching’ for gait recognition in attempt to explain the reasons why improving quality of a sample is insufficient to achieve high recognition rates in unconstrained scenarios.

Quality of matching refers to the degree to which a set of samples is suitable for direct matching in terms of covariate factors. In other words, the quality of matching is directly proportional to the consistency of the environment and the subject specific covariates of set of samples. For example, matching two samples which have been captured at a similar viewpoint would have a higher quality of matching than if there is a significant change in pose.

Chapter 4 demonstrated that very high recognition rates can be achieved using relatively simple techniques when the quality of matching between samples is high. However, in realistic environments, both environmental and subject related factors can change. Extracting features from videos enables us to match samples acquired in similar conditions. An example of feature extraction is using the heel strikes and camera calibration parameters to estimate the angle that a subject is moving relative to the camera.

In this section we show that we can improve performance if we automatically select a subset from the gallery that contains samples of high 'matching quality' relative the probe. We use 'pose' as an example of a covariate. We use the heel strike position and the calibration parameters to detect the walking direction of the subject and therefore the angle that subjects are walking relative to the calibration board and hence the camera. Refer back to Chapter 3 for details of the computation.

Figure 5.20 shows an example of gallery. The samples marked in red have high quality of matching in terms of pose relative to the probe shown in Figure 5.19. Note that the gallery contains samples of multiple subjects.

The combination of gallery and probe used for this experiment is shown in Table 5.4.

Probe			Gallery			CCR (%)
Loc.	Dir.	Subj.	Loc.	Dir.	Subj.	
Room	JED	20	Room	ID,IHC,BI,DEJ,IB,**JED**	20	80
Room	JED	20	Room	JED	20	100
Room	ID	20	Room	JED, IHC, BI, DEJ, IB, **ID**	20	80
Room	ID	20	Room	ID	20	95

Table 5.4 – The combination of gallery and probe for the quality of matching experiment

The results reveal that effectively reducing the size of the gallery can lead to improvement in performance.

Assessment of quality of matching allows a system to produce a decision confidence score or deploy the most appropriate matching algorithm. The degree to which a particular factor influences recognition depends on the algorithm deployed. Understanding which factors have changed between capturing the gallery and the probe allows a matching algorithm to be selected dynamically. For example, if there is a significant pose change and algorithm 'A' performs well when samples of different pose are used, then it would make sense to use this algorithm rather than an algorithm that performs well in the presence of different clothing. An approach that works well in the presence of all covariates does not exist for gait recognition. All approaches perform better if the samples have been acquired in similar conditions.

Figure 5.19 - The probe sample

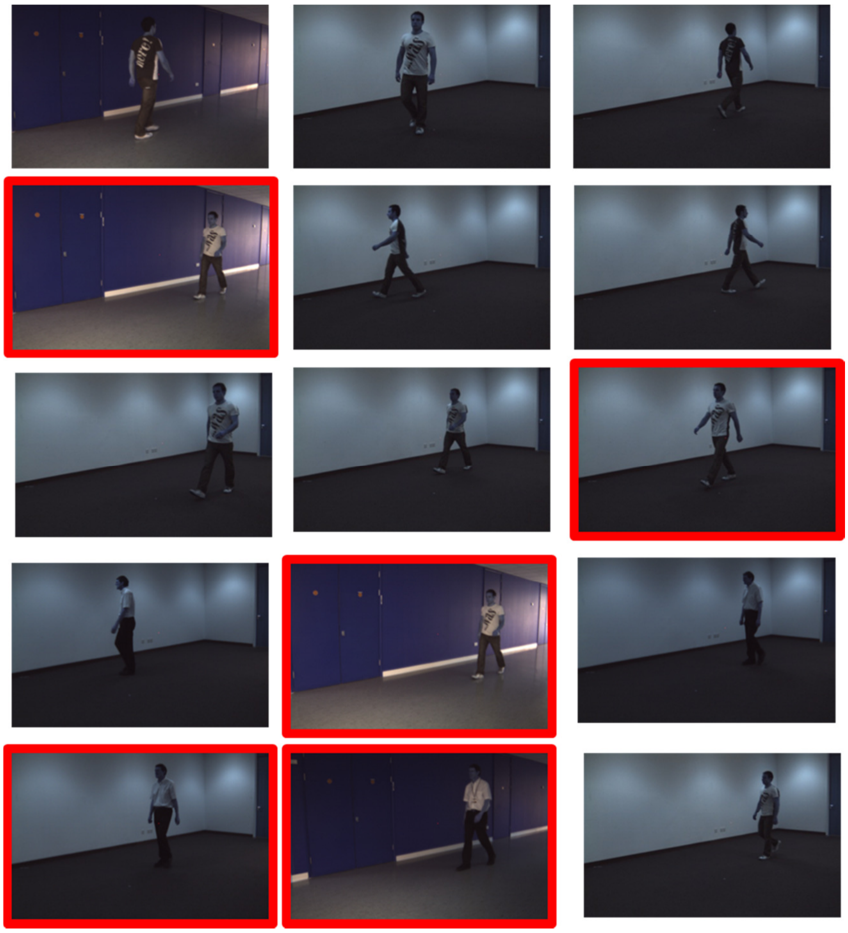


Figure 5.20 - The gallery samples

5.7 Experiments performed in real world conditions

This section provides an indication of how the results presented in the previous sections translate to reality. The experiments presented here have enabled us to understand any additional challenges and issues related to translating gait recognition to real world conditions.

As part of this thesis, we developed a working system in which biometric gait recognition can help the security operative in tracking the positions of individuals as they move through a monitored area. The system demonstrates the use of fully automated gait recognition in a real environment and is integrated with a 3D model able to show the current and previous locations a person has visited. A screen shot of the system in operation is shown in Figure 5.21.

A video that demonstrates the use of the system can be accessed at http://www.youtube.com/watch?v=bhCR0UbW13Y or at http://tiny.cc/eymqkw. A block diagram of the system is shown in Figure 5.22.

The work was part of a project called "Viewers Situational and Spatial Awareness for Applied Risk and Reasoning" (or VSAR for short). The project was supported by the Technology Strategy Board (TSB), and was collaboration between: CAST, the BBC, British Aerospace, NPL and the University of Southampton. A Gantt chart detailing the activities and timescales relating to the gait recognition part of the project is shown in Appendix G. Appendix F shows some of the test scenarios we used when integrating the gait recognition system with the 3D visualisation tool.

The system was deployed at the same locations used for the database collection (see Chapter 3, section 3.4). Examples of samples are shown in Figure 5.23. Twenty subjects took part in the ‘live-use’ evaluation. Table 5.5 shows the results obtained. There were significant improvements in the signature quality and therefore in the system performance when quality was used, similar to the experiments performed on the database. Example of signature improvement is shown in Figure 5.24. By using the subjects’ actual height to decrease the size of the gallery we were able to improve the recognition rate by 10% on average. Details of how height is computed are shown in Chapter 3.

The CCR in ‘live use’ was approximately 10-15% lower than results obtained by analysing the database. The system was able to produce score for the confidence of the results. The score was derived by combining measurements of sample quality and quality of matching in terms of pose. This score proved to be very good at predicting the accuracy of the system.

Probe		Gallery		Quality	Height	CCR (%)
Loc.	Dir	Loc.	Dir			
Corridor	IB	Corridor	IB	No	No	60
Corridor	IB	Corridor	IB	Yes	No	75
Corridor	IB	Corridor	IB	Yes	Yes	85

Table 5.5 - The probe and gallery for 'live use'

The main challenge encountered in 'live use' was adjusting the '*quality threshold*' and '*matching score threshold*' parameters. The changes in the environment are much more unpredictable than in the case of a database of fixed size. A set of parameters can deliver good system performance at a particular time and that can cease to be the case some time later (few seconds in some cases). Our tests highlighted the need to develop algorithms for dynamic parameter adjustment in real-world environments. The value of these parameters is likely to be correlated to the quality of samples.

There are number of challenges associated with performing gait recognition on samples acquired at different locations. Studies that use gait for the purpose of re-identification are limited [68]. There can be a difference in sample quality, subject-dependent changes and changes related to the environment. By deploying quality metrics we were able to improve matching across location by gait. However, we acknowledge that unsolved issues remain such as dealing with view-point or clothing changes.

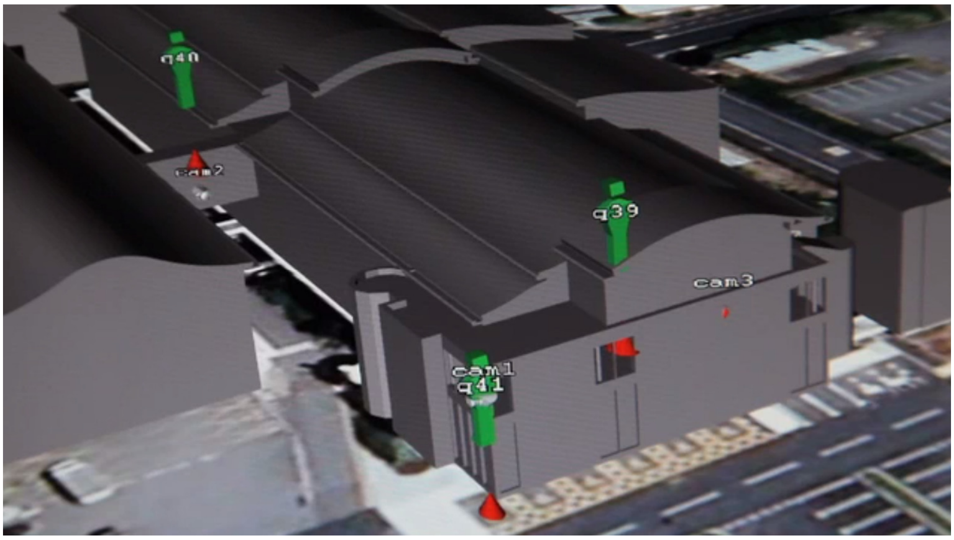


Figure 5.21 - VSAR model and gait recognition points

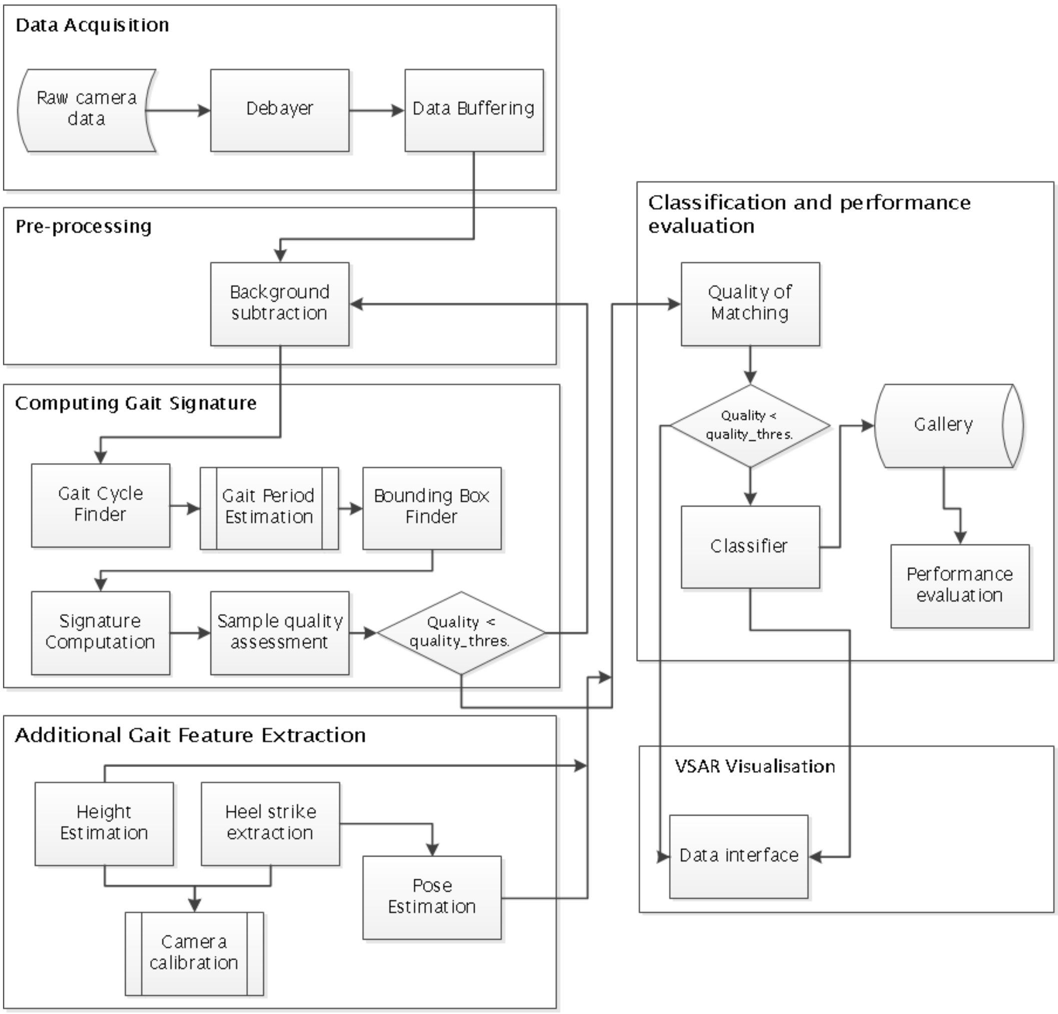


Figure 5.22 - System architecture

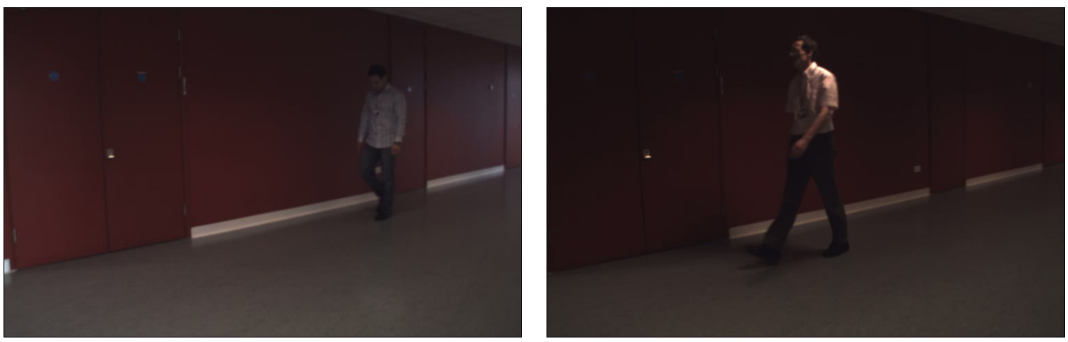


Figure 5.23 - Samples taken during 'live use' use

Figure 5.24 - Signature improvement by using quality

5.8 Discussion

In this chapter we discuss some of the environmental dependent factors that can affect gait recognition performance. In real world environments, one of the most important factors is silhouette quality. Many studies have shown the silhouette quality can have a negative impact on performance. Measuring silhouette quality is relatively under-researched problem and therefore we have extended the quality metrics and evaluated their effectiveness. We have proposed a novel way of using quality metrics to improve an important pre-processing step. Background subtraction is the first processing step of most gait recognition algorithms and the success of further processing depends on this process. The approach is generic and can be applied in other domains that employ background subtraction as a pre-processing step. We have also used quality metrics to select a gait cycle that produces signature of best quality.

The ways in which we use sample quality are vital if a sample is acquired in a real world conditions where re-capturing is not possible. The benefits of using quality have been demonstrated on a new and challenging dataset specifically collected for this thesis. In addition, the advantages have been validated in 'live use'. Experiments have been performed over multiple locations and the results reveal that quality of gait samples is a step in the right direction for deployment of gait recognition in real-world environments.

Results presented in this chapter showed that change in the viewing angle and camera height can have a negative effect on performance even when samples of sufficient quality are used. We introduce the notion of 'quality of matching' and show the importance of matching signatures acquired in consistent conditions. We show that it is possible to achieve high recognition rates in real-world environment if the gallery and probe are similar in terms of covariate factors. We propose a way of improving recognition by extracting features from video. We use heel strike detection algorithm to enable us to match samples which have high 'quality of matching' in terms of pose.

Quality can be used in other ways not described in this chapter. For example, using only high quality biometric sample at the enrolment process contributes to increased performance of the system at the matching stage. Quality can be used as an indicator of confidence in the result or for quality driven biometric fusion.

Our results show that it is possible to achieve high recognition rates even if samples of poor quality are used. This is because the environment does not tend to change in a short time and the signatures are similar to each other i.e. consistent. Many studies report good recognition rates when the gallery and probe are recorded only few minutes apart in the same environment. Ignoring the issue of quality has led to many incorrect conclusions. For example, many studies have incorrectly concluded that elapsed time has a negative impact on performance. The evidence presented in Chapter 4 and in this chapter shows that signature quality and the issues around quality of matching are much more significant than the effect of time itself.

The VSAR project allowed us to apply all of the knowledge and experience acquired throughout the research to develop a working system. The results obtained are consistent with tests performed on the datasets. The project allowed us to investigate the additional challenges present in real-world environments compared to analysing a dataset of fixed size. Our results show that there is a need of further research in understanding how to dynamically adjust some of the key parameters when performing recognition in realistic environments.

Our results demonstrate the potential and the benefits of using quality. However, there is scope for development of additional quality metrics. Further improvements in performance are possible with a larger number of metrics. Quality metrics can be dependent on the algorithm used and further research is needed to understand how the conclusions presented here apply to other techniques. The model used for quality metric 1 and 3 could have been extended. However, our experiments show that the error of fit is very large in the case of poor quality samples and therefore, model improvement was not considered necessary.

While it is difficult to acquire samples of good quality in a realistic environment, the techniques presented here show that quality of gait samples is a promising avenue for future research.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Many studies have shown that it is possible to recognize people by the way they walk. However, there are a number of factors that affect recognition performance. Some of the most difficult challenges and related work in addressing these challenges has been discussed. At the moment, gait recognition performs well only when samples acquired in similar conditions are matched. The majority of studies to date do not have a deployment focus and lack experiments on data acquired in real world conditions.

The existing public gait databases are insufficient for performing certain analyses and therefore we have created two novel databases. One of the databases enables a researcher to analyse subject dependent covariates in a principled manner. It could also be used for developing and evaluating covariate invariant algorithms. The other database is suitable for studying factors related to real world environments. A number of techniques have been considered for processing the data and it would appear that relatively simple techniques perform better for data acquired in challenging conditions.

A number of subject related covariate factors have been investigated. The elapsed time between capturing the gallery and the probe has been reported to significantly affect recognition performance. However, until now, no study has shown the isolated effect of time, irrespective of other covariates. In this thesis we have described the first principled study that examines the effect of elapsed time on gait recognition. Using empirical evidence we show for the first time that elapsed time does not affect recognition significantly in the short to medium term. Our results show that gait can be used as a reliable biometric over time and at a distance. We have also investigated the effect of different type of clothes, variations in speed and footwear on the recognition performance. We have demonstrated that clothing changes drastically affects performance regardless of elapsed time and significantly more than any of the other subject dependent covariates that we have considered here. The research then suggests a move towards developing appearance invariant recognition algorithms.

We have shown that there are many environment related factors that can affect recognition. The quality of extracted silhouettes has been identified as one of the most important factors. We have proposed novel ways of improving signature quality by adjusting some of the pre-processing steps common for many recognition approaches. Improving the quality would allow other techniques that rely on silhouette data to generalize to unknown environments.

Experiments using a fully automated gait system have revealed additional challenges not present when analysing datasets of fixed size. Some of the issues can be mitigated by assessing and using sample quality data. Utilising additional information such as height and heel strike data proved beneficial when performing experiments in ‘live use’.

Research to date has shown that deploying gait recognition in the real world remains a challenging problem. Our research has improved the understanding of the current challenges and has proposed novel ways of dealing with some of the fundamental issues.

6.2 Future work

In this thesis quality has been used as a way of improving fundamental pre-processing steps for gait recognitions and as indication in the confidence of the result. Further research is needed into automatic selection of matching algorithm driven by assessment of sample quality.

There are numerous other ways of utilising quality metrics. An important use of quality is for quality driven biometric fusion with other modalities. There are few studies [69] [70] [71] in the literature that describe fusion of face and gait, but none of the studies concentrate on quality driven fusion between gait and face biometrics. Features extracted from poor quality biometric samples are less reliable. The quality of biometric sample can be used in the fusion process to dynamically assign higher weighting to match scores arising from higher quality input samples. Failure to assign appropriate weighting to traits of different quality could lead to a worse recognition rate that using a single trait. Figure 6.1 shows a top level diagram of a possible system that uses quality and covariate factors to effectively fuse gait and face biometrics.

The VSAR project is a great example of how gait recognition could be deployed. This project revealed the benefits of incorporating gait (or other biometrics) in 3D visualisation tool. There is a scope for further research into fusing additional information in order to improve performance. A system like VSAR could use operator decision, time information, geometry of building etc. to assist the biometric system.

Quality of samples is an important consideration when using crime scene data for the purposes of gait identification. In many cases videos are acquired using time lapse technologies, insufficient resolution etc. In this context, quality issues not discussed in this document are vital. Further research is needed in the area of quality assessment for video data acquired using surveillance cameras.

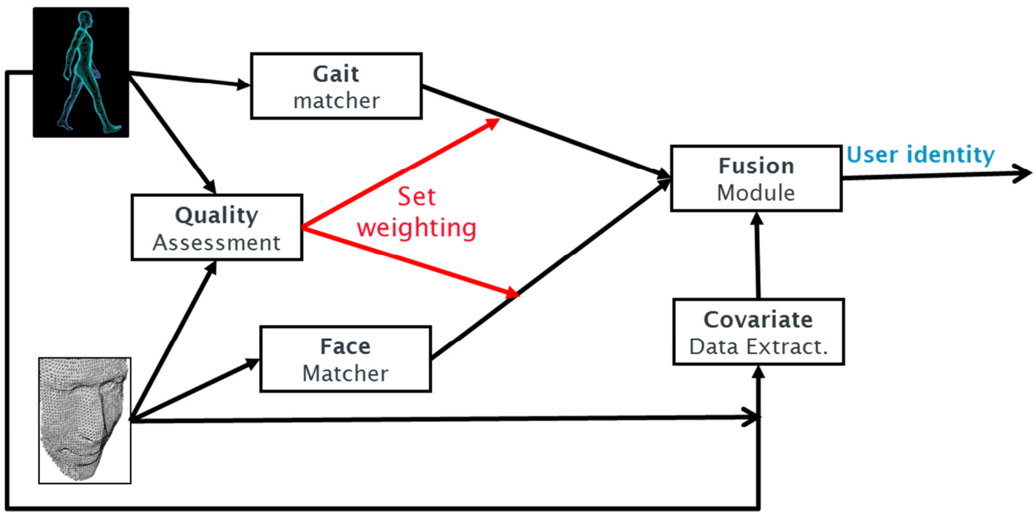


Figure 6.1 – Top level diagram of a possible system

The quality metrics developed in this thesis were specific to the approaches used as well as the environment in which the data was collected. Although all approaches could benefit from the way in which quality has been used in this thesis, further work is needed into understanding how to extend these methods for other recognition approaches and different or more complex operational environment. Additional work is needed in understanding which quality metrics are appropriate in which environment and how they are related to that environment. There is scope for development of additional quality metrics and refinements of the models presented using larger and more diverse datasets and utilising larger number of recognition approaches. Further research in quality of samples is vital to bring gait recognition closer to deployment.

The results presented in this thesis have been obtained by using model-free gait representations. Further work is needed to extend and verify the analyses using model-based approaches. Model-based approaches might mitigate against the influences of some covariate factors.

Apart from video based recognition, other types of technology have been used for gait recognition such as floor sensors, accelerometers and mobile phones [72]. Many of the issues discussed in this thesis are relevant to these technologies and therefore additional work is needed to extend the analyses.

This research has analysed many of the factors that affect gait recognition performance and has presented novel ways of dealing with some of the most fundamental issues related to translating gait recognition to real world environments. These methods have been demonstrated to address improvements in performance in a fully-automated gait recognition system.

Appendix A

Background Subtraction

In this section we present the theory behind the background subtraction method we have used throughout this thesis [50]. The method uses recursive equations to constantly adapt the parameters of a Gaussian mixture model. It assumes that the adjacent pixels are uncorrelated.

The value of a pixel at time *t* in RGB is denoted by $\mathbf{x}^t$. The pixel based background subtraction involves decision if the pixel belongs to the background (BG) or some foreground object (FG). The pixel is more likely to belong to the background if:

$$\frac{p(BG|\,\boldsymbol{x}^t)}{p(FG|\boldsymbol{x}^t)} = \frac{p(\boldsymbol{x}^t|BG)p(BG)}{p(\boldsymbol{x}^t|FG)p(FG)} \tag{A.1}$$

is larger than 1 and vice versa.

The decision that a pixel belongs to the background is made if

$$p(\boldsymbol{x}^t|BG) > c_{thr}\left(= \frac{p(\boldsymbol{x}^t|FG)p(FG)}{p(BG)}\right) \tag{A.2}$$

where c_{thr} is a threshold value. $p(\boldsymbol{x}^t|BG)$ is the background model. The background model is estimated from a training set *X*. The estimated model is denoted by $\hat{p}(x|X, BG)$ and depends on the training set.

In order to adapt to possible changes the training set should be updates. For a reasonable adaptation period *T*, at time *t*, $X_t = \{x^t, \ldots, x^{(t-T)}\}$. For each new sample the training data set is updated and the density re-estimated. These samples might contain values that belong to the foreground object and therefore $\hat{p}(\boldsymbol{x}^t|X_t, BG + FG)$. GMM with M components is used:

$$\hat{p}(\boldsymbol{x}^t|X_t, BG + FG) = \sum_{m=1}^{M} \hat{\pi}_m N(\boldsymbol{x};\, \widehat{\boldsymbol{\mu}}_m, \widehat{\boldsymbol{\sigma}}^2{}_m\, I) \tag{A.3}$$

$\widehat{\boldsymbol{\mu}}_1, \dots \dots \widehat{\boldsymbol{\mu}}_m$ are the estimates of the means and $\widehat{\boldsymbol{\sigma}}^2{}_1, \dots \dots \widehat{\boldsymbol{\sigma}}^2{}_m$ are the estimates of the variances that describe the Gaussian components. The estimated mixing weights are denoted by $\hat{\pi}_m$, they are non-negative and add up to one. Given a new data sample $\boldsymbol{x}^t$ at time t the recursive update equations are:

$$\hat{\pi}_m \leftarrow \hat{\pi}_m + \alpha(o_m^t - \hat{\pi}_m) \tag{A.4}$$

$$\widehat{\boldsymbol{\mu}}_m \leftarrow \widehat{\boldsymbol{\mu}}_m + o_m^t(\frac{\alpha}{\hat{\pi}_m})\boldsymbol{\delta}_m \tag{A.5}$$

$$\hat{\sigma}_m^2 \;\leftarrow\; \hat{\sigma}_m^2 \;+\; o_m^t \left(\frac{\alpha}{\hat{\pi}_m}\right)(\boldsymbol{\delta}_m^T\boldsymbol{\delta}_m - \hat{\sigma}_m^2) \tag{A.6}$$

where $\boldsymbol{\delta}_m = \boldsymbol{x}^t - \widehat{\boldsymbol{\mu}}_m$. The time constant $\alpha = \frac{1}{T}$ defines an exponentially decaying envelope that is used to limit the influence of the old data. For a new data sample the ownership o_m^t is set to 1 for the 'close' components with largest $\hat{\pi}_m$ and the close are set to zero. A component is 'close' if the Mahalanobis distance from the component is less than three. The squared distance from the m^{th} component is calculated by the following equation:

$$D_m^2(\boldsymbol{x}^t) = \frac{\boldsymbol{\delta}_m^T\boldsymbol{\delta}_m}{\hat{\sigma}_m^2} \tag{A.7}$$

If there is no 'close' component a new component is generated with $\hat{\pi}_{M+1} = \alpha$, $\widehat{\boldsymbol{\mu}}_{M+1} = \boldsymbol{x}^t$ and $\hat{\sigma}_{M+1} = \sigma_0$ where σ_0 is some appropriate initial variance. If the maximum number of components is reached the component with smallest $\hat{\pi}_m$ is discarded.

Appendix B

Camera Model and Calibration

Camera calibration enables us to match a point in the image to a position in the real world. The pinhole camera model has been used. Using this model, a scene is formed by projecting 3D points into an image using a perspective transformation:

$$s\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}f_x & 0 & c_x\\0 & f_y & c_y\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13} & t_1\\r_{21} & r_{22} & r_{23} & t_2\\r_{31} & r_{32} & r_{33} & t_3\end{bmatrix}\begin{bmatrix}X\\Y\\Z\\1\end{bmatrix} \tag{B.1}$$

The equation can also be written as:

$$s\,m' = A[R|t]M' \tag{B.2}$$

(X,Y,Z) are 3D coordinates in the world space and *(u,v)* are the projection points on the image in pixels. *A* is usually referred to as camera matric or intrinsic parameters. The point (c_x,c_y) is called the principal point and is usually the centre of the image. f_x and f_y are the focal lengths in units of pixels. The intrinsic parameters do not depend on the scene viewed and should only be scaled if the image is scaled. The *[R|t]* is known as the matrix of extrinsic parameters and it consists of rotation *R* and translation *t* matrix. The extrinsic parameters describes the translation of a point *(X,Y,Z)* to a camera coordinate system:

$$\begin{bmatrix}x\\y\\z\end{bmatrix}=R\begin{bmatrix}X\\Y\\Z\end{bmatrix}+t \tag{B.3}$$

when z ≠ 0

$$x' = \frac{x}{z},\ y' = \frac{y}{z} \tag{B.4}$$

$$u = f_x * x' + c_x \tag{B.5}$$

$$v = f_y * y' + c_y \tag{B.6}$$

There is distortion associated with real camera lenses. The most common types of distortions are radial and tangential. They can be modelled by the following equations:

$$x'' = x'(1 + k_1r^2 + k_2r^4 + k_3r^6) + 2p_1x'y' + p_2(r^2 + 2x'^2) \tag{B.7}$$

$$y'' = y'(1 + k_1r^2 + k_2r^4 + k_3r^6) + p_1\left(r^2 + 2y'^2\right) + 2p_2x'y' \tag{B.8}$$

where $r^2 = x'^2 + y'^2$ and *k1, k2, k3* are the radial distortion coefficients and *p1, p2* are the tangential distortion coefficients. Equations (B.5) and (B.6) can be written as:

$$u = f_x * x'' + c_x \tag{B.9}$$

$$v = f_y * y'' + c_Y \tag{B.10}$$

Camera calibration is performed to obtain the intrinsic, extrinsic and distortion coefficients associated with a particular camera and view. In this thesis we use "Roger Y. Tsai Algorithm". Details of the algorithm have been omitted here as it is one of the most commonly used and described methods for camera calibration.

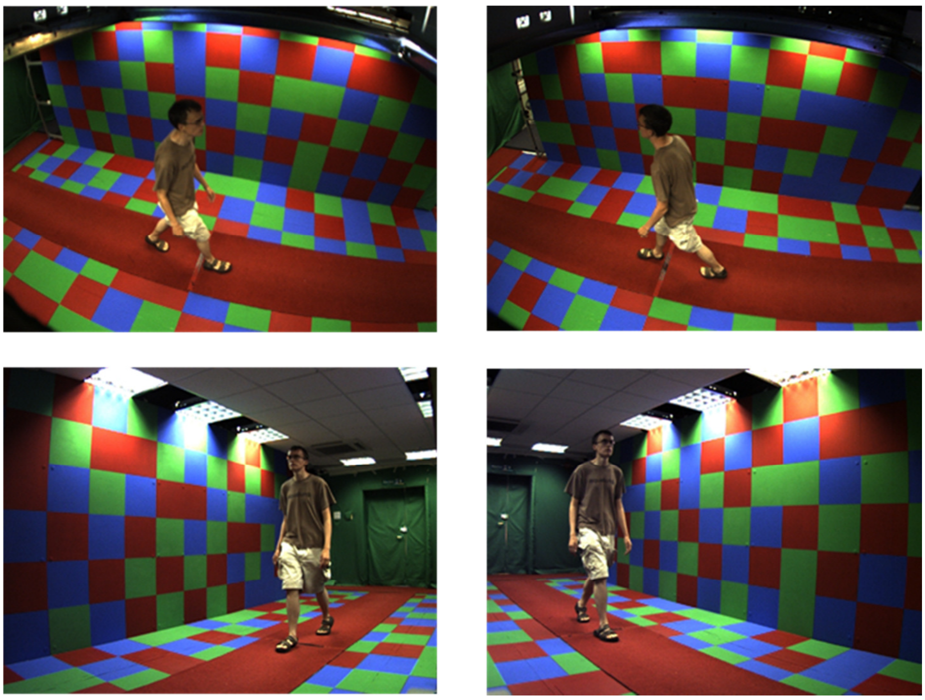
Appendix C

Datasets

The details of the datasets were presented in Chapter 3. This appendix shows additional examples of usable samples. Usable samples are defined as walking sequences in which the subject performs at least one full gait cycle.

C.1 The biometric tunnel database

A video sequence is available from 12 different cameras.



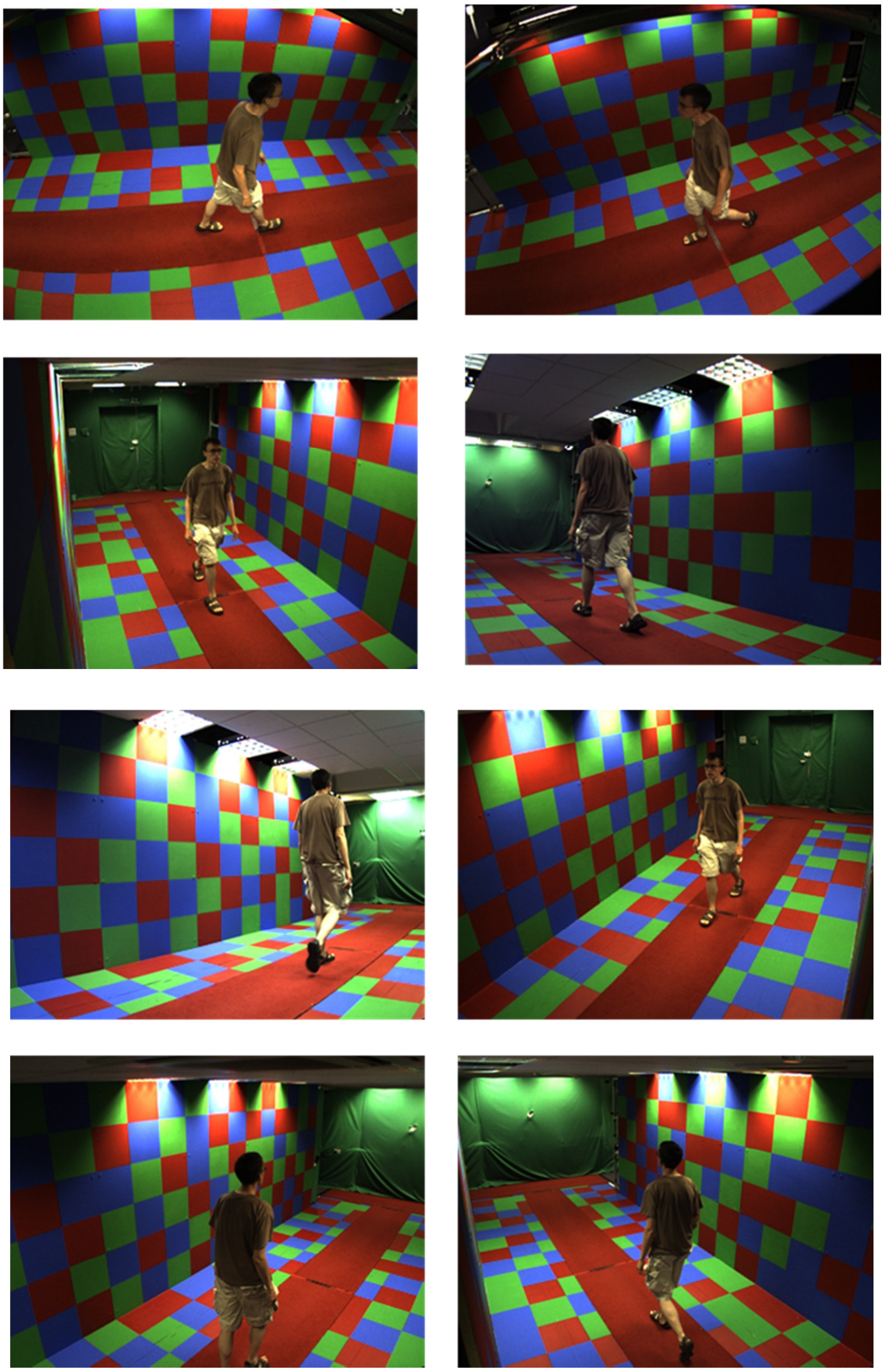


Figure C.1 – Examples of samples from the Biometric Tunnel Database

C.2 Indoor surveillance database

Data is recorded at two different locations using three cameras.

Figure C.2 - Examples of samples acquired by camera 1 in the ‘Corridor’ location

Figure C.3 - Examples of samples acquired by camera 2 in the ‘Corridor’ location

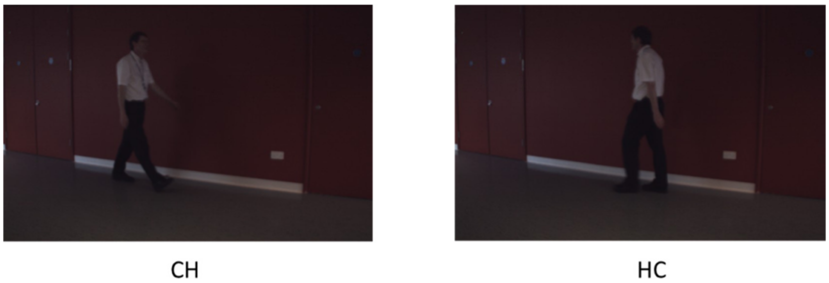


Figure C.4 - Examples of samples acquired by camera 3 in the ‘Corridor’ location

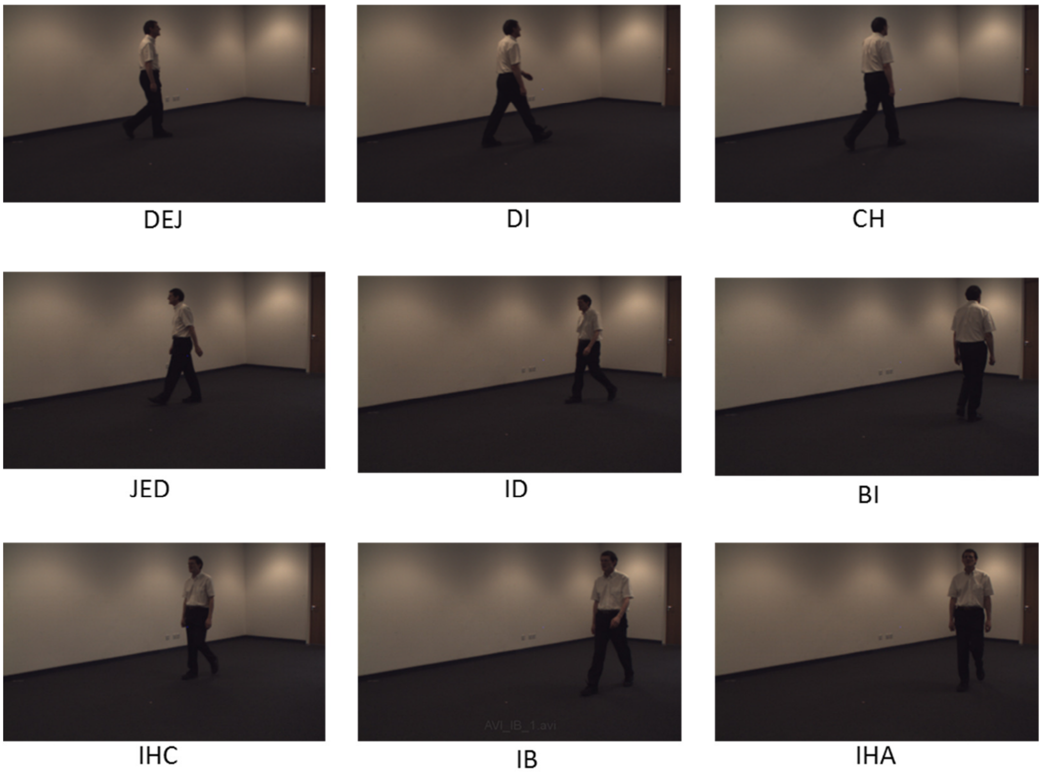


Figure C.5 - Examples of samples acquired by camera 1 in the ‘Room’ location

Appendix D

Anthropometric Measurements

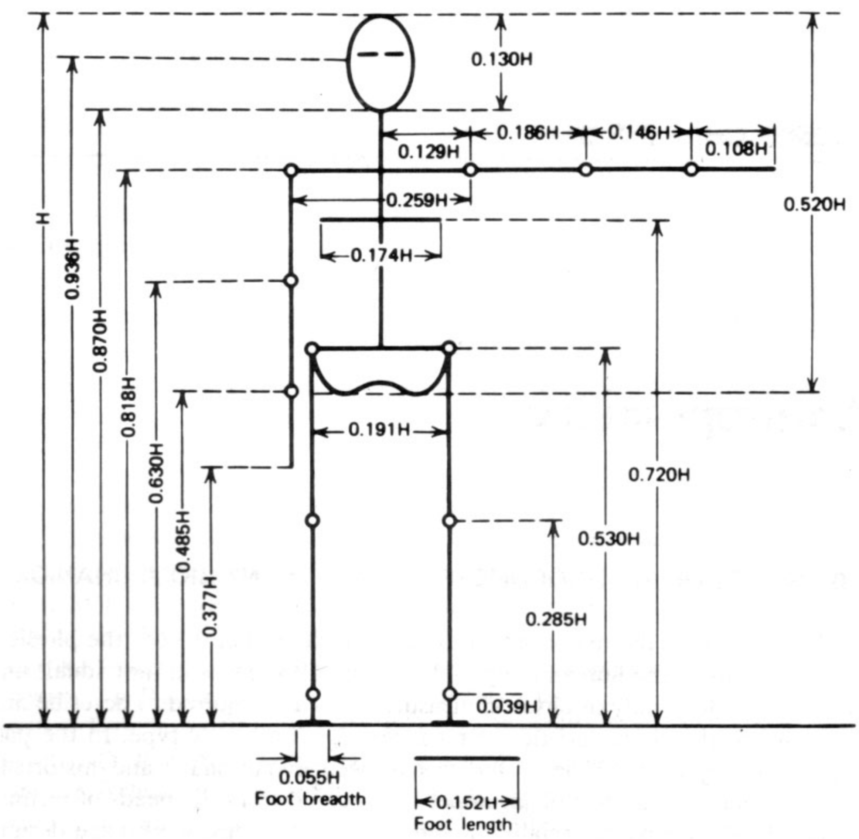


Figure D.1 – Average anthropometric measurements

Appendix E

Silhouette Data Quality

Examples of poor quality data with body parts missing and holes in the silhouettes:

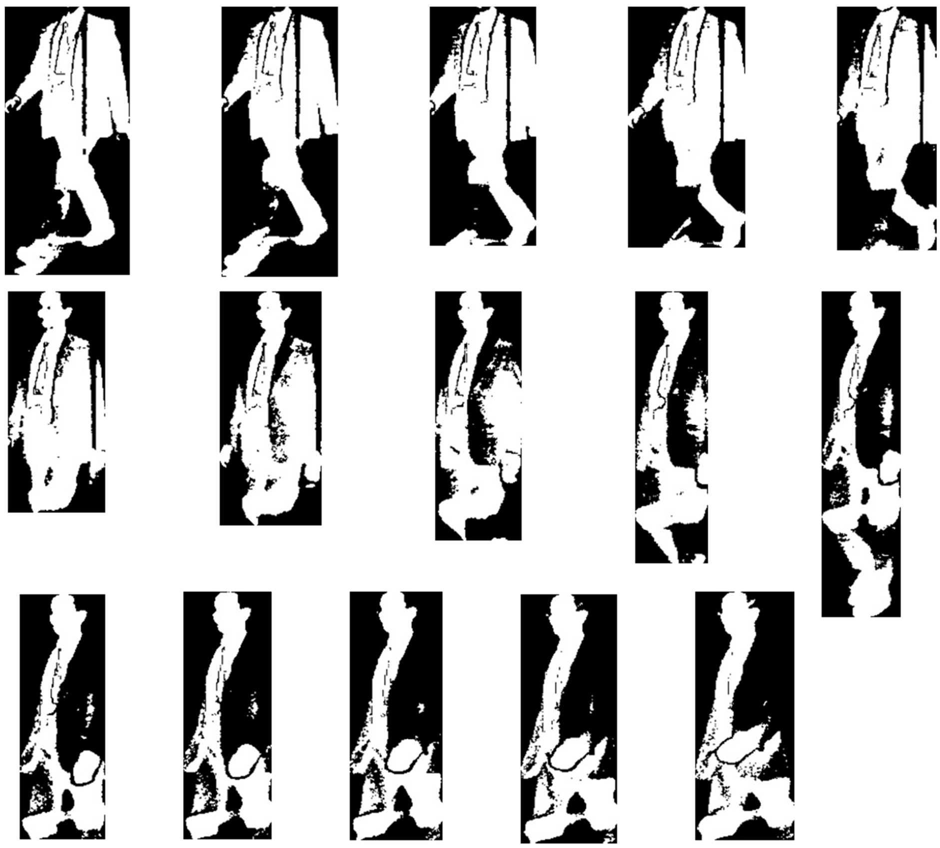


Figure E.1 - Example of poor quality silhouettes for a single gait cycle

Examples of poor quality data with body parts missing:

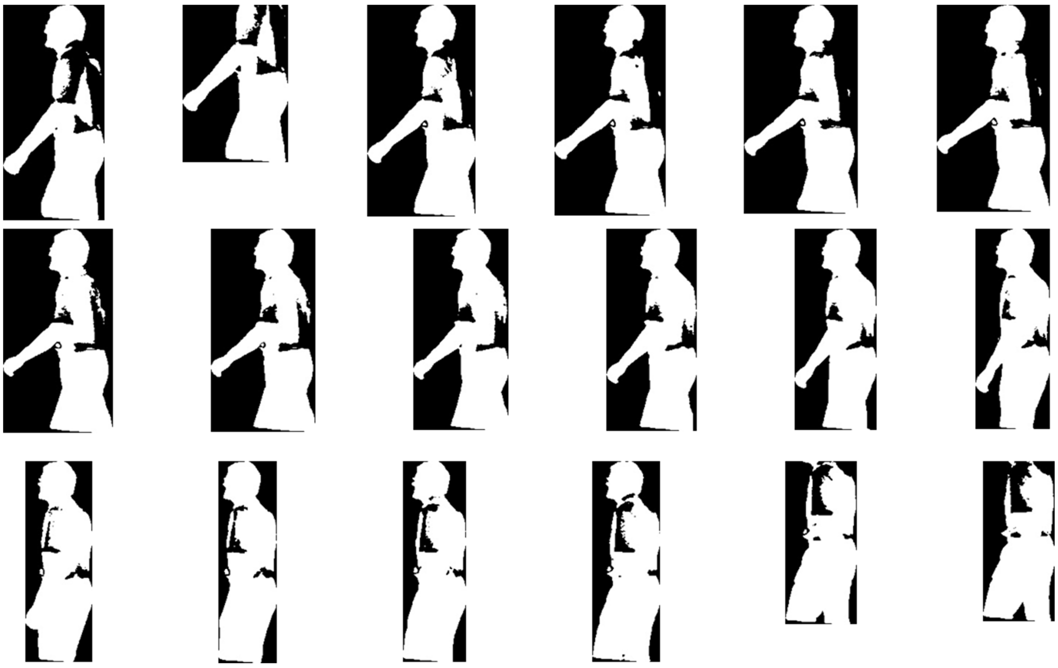


Figure E.2 - Example of poor quality silhouettes for a single gait cycle

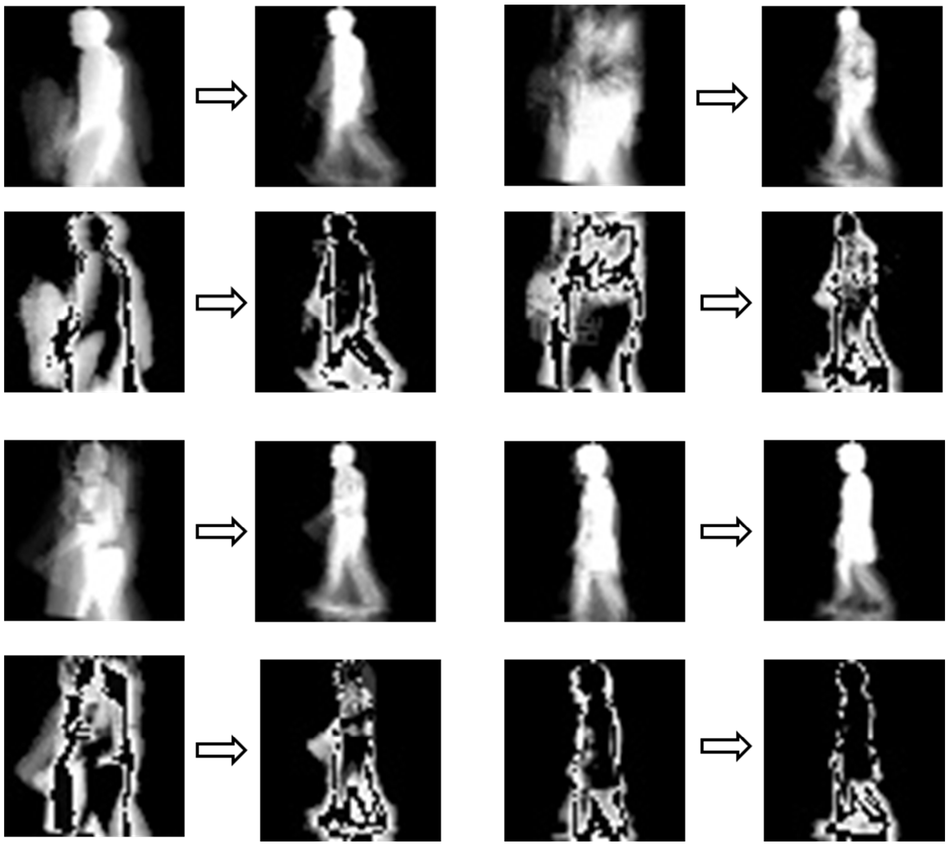


Figure E.3 – Examples of signature quality improvement

Appendix F

VSAR System Testing

F.1 Scenarios

The integration of the VSAR visualisation tool and the gait recognition system was tested using the scenarios described below. The Biometric Identification Points (BIP) are shown in Figure F.1. BIP 1 is located at the entrance of module 1. BIP 2 is situated on the bridge that connects two modules and BIP 3 is in a corridor inside the building.

Scenario 1: Subject_A parks in the car park and enters the building through point 1 and walks through building passing BIP 2. At some time later the subject approaches BIP 3. This is an **anomaly** as the subject is a guest and was supposed to attend a meeting in 'Module 2'.

Scenario 2: Subject_B walks from 'Module 2' and enters 'module 1' through BIP 2. This subject was not supposed to enter 'Module 1'. The subject approaches BIP 3. The operator is able to **locate** the subject through performing GR at this point.

Scenario 3: Subject_C was already in the building and has not passed 1 or 2. He is on his way to BIP 3. BIP 3 is used as a way to **register** the subject to the list of 'subjects of interest'.

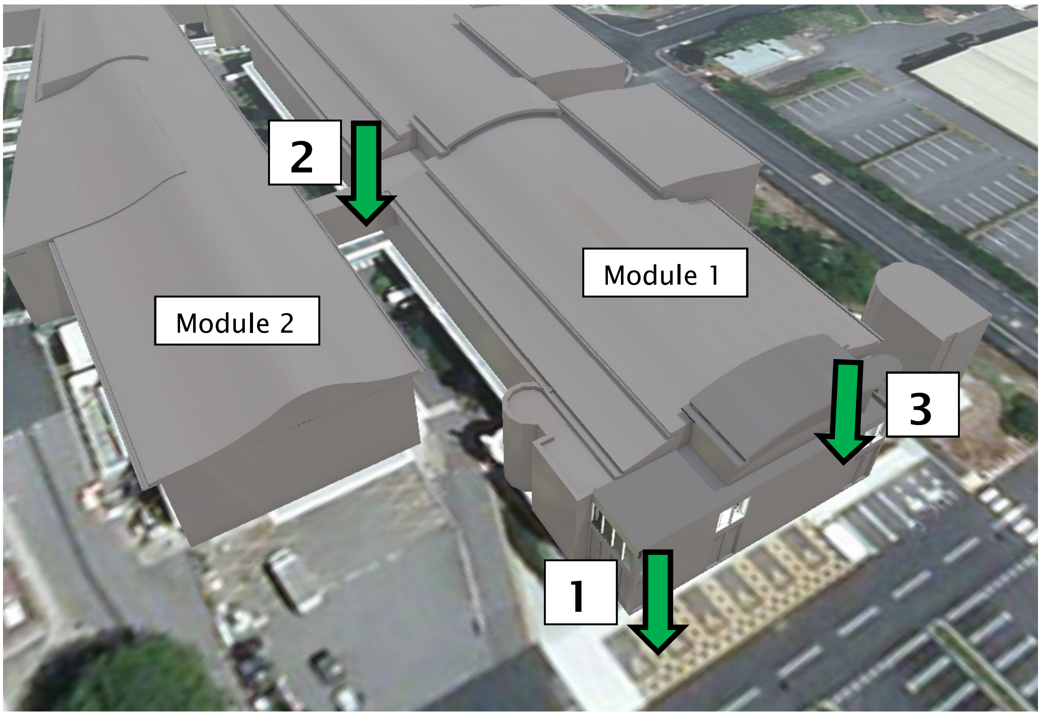


Figure F.1 – The location of the biometric identification points

When a recognition query is initiated, the gait recognition system generates one of the following three results:

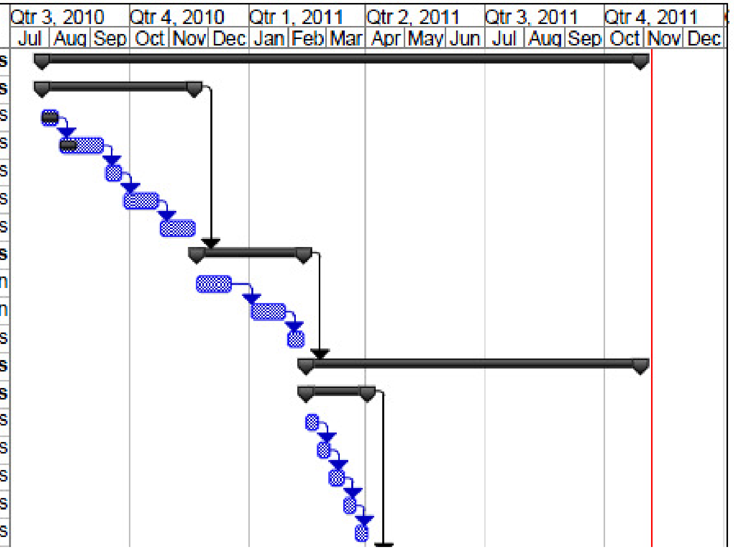
1. A message that the sample is of poor quality, or
2. A message to indicate that no matches found, or
3. A list of matches

The VSAR system enables the operator to visualise the information provided by the gait recognition system, showing image of the person and location of any matches found.

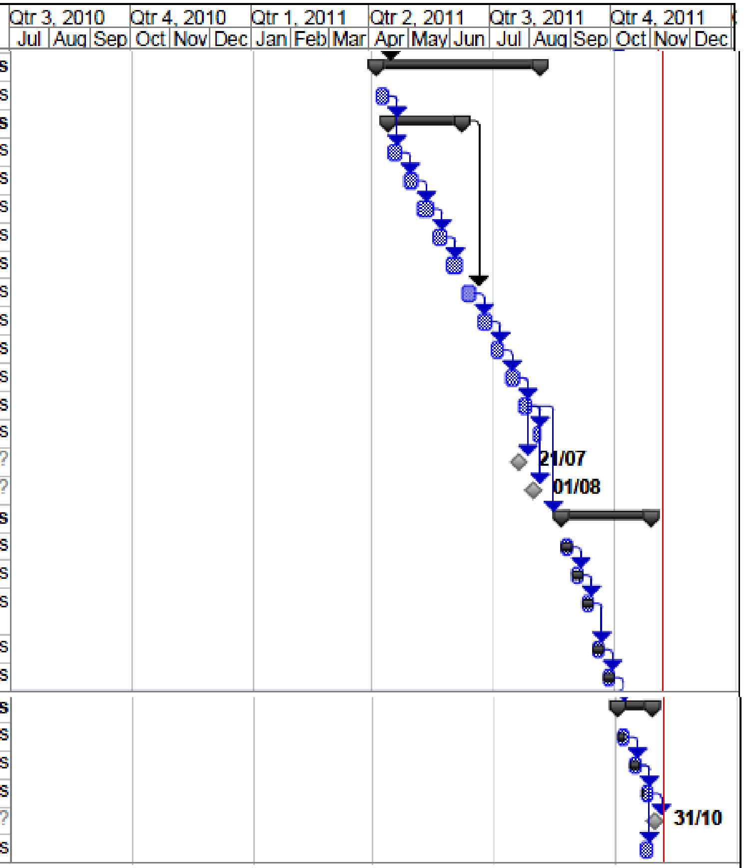
Appendix G

VSAR Project Gantt chart

ID	Task Name	Duration
1	**NPL integration schedule**	**310 days**
2	**Research potential collaborative projects**	**85 days**
3	Hold initial explorative meetings between CAST and NPL	2 wks
4	Hold internal explorative meetings to identify short list	5 wks
5	Make initial collaborative project selection	2 wks
6	Identify rough scope for work packages	4 wks
7	Produced initial Requirements	4 wks
8	**Collaboration high level design & detailed schedule development**	**50 days**
9	Identify objectives and deliverables and allocations	1 mon
10	Develop high level specification and integration document	1 mon
11	Review high level specification and integration	2 wks
12	**Implementation**	**175 days**
13	**Database collection**	**35 days**
14	Installation of Gait Analysis cameras & recording equipment at NPL.	7 days
15	Collection of preliminary data to assess suitability of NPL locations.	7 days
16	Analyses of data to assess suitability.	7 days
17	Collection of experimental data for the design and test of a gait recognition system	7 days
18	Initial processing and backup of recorded data	7 days



ID	Task Name	Duration
19	**Alpha development**	**90 days**
20	Investigating into software/hardware tools required for system implementation	7 days
21	**Design and implementation of gait recognition system modules**	**40 days**
22	Buffering from a live video feed	8 days
23	Background subtraction	8 days
24	Gait cycle estimation and selection of a suitable cycle	8 days
25	Feature extraction for gait / signature generation	8 days
26	Classification/matching	8 days
27	Gait module test alpha version	8 days
28	Integration of gait modules	8 days
29	Integration tests of gait modules	7 days
30	Undertake GR system tests	8 days
31	Create a test assesment document & identify requirements for beta version	7 days
32	Design of interfaces between biometric recognition and VSAR components	5 days
33	NPL Alpha complete	1 day?
34	NPL Alpha trials assement complete	1 day?
35	**Beta development**	**50 days**
36	Assessing the effectiveness of the system a large dataset.	6 days
37	Investigation of alternative methods in order to improve system performance	6 days
38	Implement and test strategies for recognition combining gait biometrics and position / location data.	6 days
39	Further data analyses. The effect of viewpoint and resolution on the error rates.	6 days
40	Module test beta version	6 days
41	**Integration of VSAR and gait recognition**	**20 days**
42	Implementation of the interface between the VSAR and the gait systems.	6 days
43	Integration of VSAR and gait recognition	6 days
44	Undertake system tests	6 days
45	NPL Beta complete	1 day?
46	Production of demonstation video	7 days



References

[1] P.K. Larsen, E.B. Simonsen, N. Lynnerup, Gait Analysis in Forensic Medicine, *Journal of Forensic Sciences*, **53**(5), pp. 1149-1153, 2008.

[2] Z. Xue, D. Ming, W. Song, B. Wan, S. Jin, Infrared gait recognition based on wavelet transform and support vector machine, *Pattern Recognition*, **43**(8), pp. 2904-2910, 2010.

[3] K. Bashir, T. Xiang, S. Gong, Gait recognition without subject cooperation, *Pattern Recognition Letters*, **31**(13), pp. 2052-2060, 2010.

[4] F. Jean, A.B. Albu, R. Bergevin, Towards view-invariant gait modeling: Computing view-normalized body part trajectories, *Pattern Recognition*, **42**(11), pp. 2936-2949, 2009.

[5] I. Bouchrika, M. Goffredo, J.N. Carter, M.S. Nixon, On Using Gait in Forensic Biometrics, *Journal of Forensic Sciences*, **56**(4), pp. 882-889, 2011.

[6] M. Murray, A. Drought, R. Kory, Walking patterns of normal men, *The journal of bone and joint surgery*, **46**(2), pp. 335, 1964.

[7] D.A. Winter, Biomechanics and Motor Control of Human Movement Wiley, 2009.

[8] J.E. Cutting, L.T. Kozlowski, Recognizing friends by their walk: Gait perception without familiarity cues *Bulletin of Psychonomic Society*, **9**(5), pp. 353-356, 1977.

[9] L.T. Kozlowski, J.E. Cutting, Recognizing the sex of a walker from a dynamic point-light display, *Perception & Psychophysics*, **21**(6), pp. 575-580, 1977.

[10] Y. Makihara, M. Okumura, H. Iwama, Y. Yagi, Gait-based age estimation using a whole-generation gait database, *International Joint Conference on Biometrics*, 2011, pp. 1-6.

[11] S. Sarkar, P.J. Phillips, L. Zongyi, V. Isidro Robledo, P. Grother, K. W. Bowyer, The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(2), pp. 162-177, 2005.

[12] M.S. Nixon, T.N. Tan, R. Chellappa, *Human Identification Based on Gait*, Springer-Verlag, New York, 2005.

[13] M.S. Nixon, J.N. Carter, Automatic Recognition by Gait, *Proceedings of the IEEE*, **94**(11), pp. 2013-2024, 2006.

[14] Z. Liu, S. Sarkar, Simplest representation yet for gait recognition: Averaged silhouette, *17th Int. Conf. Pattern Recognition*, 2004, pp. 211–214.

[15] J. Little, J. Boyd, Recognizing people by their gait: the shape of motion, *Videre: Journal of Computer Vision Research*, **1**(2), pp. 1-32, 1998.

[16] C. BenAbdelkader, R. Cutler, L. Davis, Motion-based recognition of people in eigengait space, *IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2002, pp. 267–272.

[17] W. Liang, T. Tieniu, N. Huazhong, H. Weiming, Silhouette analysis-based gait recognition for human identification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**(12), pp. 1505-1518, 2003.

[18] J. Hayfron-Acquah, M. Nixon, J. Carter, Automatic gait recognition by symmetry analysis, *Pattern Recognition Letters*, **24**(13), pp. 2175-2183, 2003.

[19] A. Sundaresan, A. RoyChowdhury, R. Chellappa, A hidden markov model based framework for recognition of humans from gait sequences, *IEEE Int. Conf. Image Processing* 2003, pp. 93-96.

[20] A. Kale, A. Sundaresan, A. Rajagopalan, N. Cuntoor, A. Roy-Chowdhury, V. Kruger, R. Chellappa, Identification of humans using gait, *IEEE Transactions on Image Processing*, **13**(9), pp. 1163-1173, 2004.

[21] C. Yam, M. Nixon, J. Carter, Automated person recognition by walking and running via model-based approaches, *Pattern Recognition*, **37**(5), pp. 1057-1072, 2004.

[22] I. Bouchrika, M. Nixon, Exploratory Factor Analysis of Gait Recognition, *8th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2008, pp.

[23] L. Wang, H. Ning, T. Tan, W. Hu, Fusion of static and dynamic body biometrics for gait recognition, *IEEE Transactions on Circuits and Systems for Video Technology*, **14**(2), pp. 149-158, 2004.

[24] R. Urtasun, P. Fua, 3D tracking for gait characterization and recognition, *6th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2004, pp. 17-22.

[25] Z. Guoying, L. Guoyi, L. Hua, M. Pietikainen, 3D gait recognition using multiple cameras, *7th IEEE Int. Conf. Automatic Face and Gesture Recognition* 2006, pp. 529-534.

[26] G. Ariyanto, M.S. Nixon, Marionette mass-spring model for 3D gait biometrics, *5th IAPR Int. Conf. on Biometrics* 2012, pp. 354-359.

[27] R. Seely, S. Samangooei, M. Lee, J. Carter, M. Nixon, University of Southampton multi-biometric tunnel and introducing a novel 3d gait dataset, *2nd IEEE Int. Conf. Biometrics: Theory, Applications and Systems*, 2008, pp. 1-6.

[28] M. Hofmann, S. Sural, G. Rigoll, Gait recognition in the presence of occlusion: A new dataset and baseline algorithms, *Int. Conf. on Computer Graphics, Visualization and Computer Vision*, pp., 2011.

[29] A. Roy, S. Sural, J. Mukherjee, G. Rigoll, Occlusion detection and gait silhouette reconstruction from degraded scenes, *Signal, Image and Video Processing*, **5**(4), pp. 415-430, 2011.

[30] Z. Junping, P. Jian, C. Changyou, R. Fleischer, Low-Resolution Gait Recognition, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, **40**(4), pp. 986-996, 2010.

[31] J. Liu, N. Zheng, L. Xiong, Silhouette quality quantification for gait sequence analysis and recognition, *Signal Processing*, **89**(7), pp. 1417-1427, 2009.

[32] S. Yu, D. Tan, K. Huang, T. Tan, Reducing the Effect of Noise on Human Contour in Gait Recognition, *Lecture Notes in Computer Science*, **4642**, pp. 338-346, 2007.

[33] C. Chen, J. Liang, H. Zhao, H. Hu, J. Tian, Frame difference energy image for gait recognition with incomplete silhouettes, *Pattern Recognition Letters*, **30**(11), pp. 977-984, 2009.

[34] M. Goffredo, I. Bouchrika, J.N. Carter, M.S. Nixon, Self-Calibrating View-Invariant Gait Biometrics, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, **40**(4), pp. 997-1008, 2010.

[35] W. Kusakunniran, W. Qiang, Z. Jian, L. Hongdong, Support vector regression for multi-view gait recognition based on local motion feature selection, *IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2010, pp. 974-981.

[36] G. Junxia, D. Xiaoqing, W. Shengjin, W. Youshou, Action and Gait Recognition From Recovered 3-D Human Joints, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **40**(4), pp. 1021-1033, 2010.

[37] M.A. Hossain, Y. Makihara, W. Junqui, Y. Yagi, Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control, *Pattern Recognition*, **43**(6), pp. 2281-2291, 2010.

[38] M.R. Aqmar, K. Shinoda, S. Furui, Robust Gait Recognition against Speed Variation, *20th Int. Conf. Pattern Recognition* 2010, pp.

[39] D.S. Matovski, M.S. Nixon, S. Mahmoodi, J.N. Carter, The Effect of Time on Gait Recognition Performance, *IEEE Transactions on Information Forensics and Security*, **7**(2), pp. 543-552, 2012.

[40] I. Venkat, P. De Wilde, Robust Gait Recognition by Learning and Exploiting Sub-gait Characteristics, *International Journal of Computer Vision*, **91**(1), pp. 7-23, 2011.

[41] B. Guo, M. Nixon, Gait feature subset selection by mutual information, *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* **39**(1), pp. 36-46, 2009.

[42] K. Bashir, X. Tao, G. Shaogang, Feature selection on Gait Energy Image for human identification, *IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2008, pp. 985-988.

[43] G. Veres, L. Gordon, J. Carter, M. Nixon, What image information is important in silhouette-based gait recognition, *IEEE Conf. Computer Vision and Pattern Recognition*, 2004, pp.

[44] S. Samangooei, J. Bustard, M.S. Nixon, J.N.N. Carter, On Acquisition and Analysis of a Dataset Comprising of Gait, Ear and Semantic Data, *Multibiometrics for Human Identification*, Cambridge University Press, 2010.

[45] N.J.B. McFarlane, C.P. Schofield, Segmentation and tracking of piglets in images, *Machine Vision and Applications*, **8**(3), pp. 187-193, 1995.

[46] N.M. Oliver, B. Rosario, A.P. Pentland, A Bayesian computer vision system for modeling human interactions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(8), pp. 831-843, 2000.

[47] C. Stauffer, W. Grimson, Adaptive background mixture models for real-time tracking, *IEEE Int. Conf. Computer Vision and Pattern Recognition*, 1999, pp. 1.

[48] S. Calderara, R. Melli, A. Prati, R. Cucchiara, Reliable background suppression for complex scenes, *4th ACM Int. workshop on Video Surveillance and Sensor Networks*, 2006, pp. 211-214.

[49] C.R. Wren, A. Azarbayejani, T. Darrell, A.P. Pentland, Pfinder: Real-time tracking of the human body, *IEEE Tran. on Pattern Analysis and Machine Intelligence*, **19**(7), pp. 780-785, 1997.

[50] Z. Zivkovic, F. van der Heijden, Efficient adaptive density estimation per image pixel for the task of background subtraction, *Pattern Recognition Letters*, **27**(7), pp. 773-780, 2006.

[51] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, *IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2005, pp. 886-893.

[52] E. Jeges, I. Kispal, Z. Hornak, Measuring human height using calibrated cameras, *Conf. Human System Interactions*, 2008, pp. 755-760.

[53] J. Daugman, Biometric decision landscapes, *Techical Report - University of Cambridge Computer Laboratory* pp., 2000.

[54] I. Bouchrika, M. Goffredo, J. Carter, M. Nixon, Covariate Analysis for View-point Independent Gait Recognition, *3rd IAPR Int. Conf. on Biometrics*, 2009, pp. 990 - 999.

[55] Y. Shiqi, T. Daoliang, T. Tan, A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition, *18th Int. Conf. Pattern Recognition* 2006, pp. 441-444.

[56] L. Lee, W.E.L. Grimson, Gait analysis for recognition and classification, *Proc. 5th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2002, pp. 148-155.

[57] G. Veres, M. Nixon, J. Carter, Model-based approaches for predicting gait changes over time, *Proc. Int. Conf. ISSNIP*, 2005, pp. 325-330.

[58] A. Kale, A. Rajagopalan, N. Cuntoor, V. Kruger, Gait-based recognition of humans using continuous HMMs, *Proc. 5th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2002, pp. 336-343.

[59] Z. Liu, L. Malave, S. Sarkar, Studies on Silhouette Quality and Gait Recognition, *IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2004, pp. 704-711.

[60] R.T. Collins, R. Gross, J. Shi, Silhouette-based human identification from body shape and gait, *Proc. 5th IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2002, pp. 366-371.

[61] M. Okumura, H. Iwama, Y. Makihara, Y. Yagi, Performance Evaluation of Vision-based Gait Recognition using a Very Large-scale Gait Database, *4th IEEE Int. Conf. Biometrics: Theory, Applications and Systems*, 2010, pp.

[62] D. Tao, X. Li, X. Wu, S. Maybank, Elapsed Time in Human Gait Recognition: A New Approach, *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2006, pp. 1-6.

[63] R. Tanawongsuwan, A. Bobick, Performance analysis of time-distance gait parameters under different speeds, *Lecture Notes in Computer Science*, pp. 715-724, 2003.

[64] A. Balan, M. Black, The naked truth: Estimating body shape under clothing, *Proc. European Conf. on Computer Vision*, 2008, pp. 15–29.

[65] Z. Liu, S. Sarkar, Effect of silhouette quality on hard problems in gait recognition, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, **35**(2), pp. 170-183, 2005.

[66] J.C. SanMiguel, J.M. Martínez, On the Evaluation of Background Subtraction Algorithms without Ground-Truth, *7th IEEE Int. Conf. Advanced Video and Signal Based Surveillance*, 2010, pp. 180-187.

[67] D.H. Parks, S.S. Fels, Evaluation of background subtraction algorithms with post-processing, *5th IEEE Int. Conf. Advanced Video and Signal Based Surveillance*, 2008, pp. 192-199.

[68] I. Bouchrika, J.N. Carter, M.S. Nixon, Recognizing people in non-intersecting camera views, *3rd Int. Conf. Crime Detection and Prevention*, 2009, pp. 1-6.

[69] X. Zhou, B. Bhanu, Feature fusion of side face and gait for video-based human identification, *Pattern Recognition*, **41**(3), pp. 778-795, 2008.

[70] Z. Liu, S. Sarkar, Outdoor recognition at a distance by fusing gait and face, *Image and Vision Computing*, **25**(6), pp. 817-832, 2007.

[71] C. Shan, S. Gong, P.W. McOwan, Fusing gait and face cues for human gender recognition, *Neurocomputing*, **71**(10-12), pp. 1931-1938, 2008.

[72] M.O. Derawi, C. Nickel, P. Bours, C. Busch, Unobtrusive User-Authentication on Mobile Phones Using Biometric Gait Recognition, *6th Int. Conf. Intelligent Information Hiding and Multimedia Signal Processing* 2010, pp. 306-311.