

2013-03-06

Context-based gait recognition

Bazazian, Shermin

Bazazian, S. (2013). Context-based gait recognition (Master's thesis, University of Calgary, Calgary, Canada). Retrieved from <https://prism.ucalgary.ca>. doi:10.11575/PRISM/25437
<http://hdl.handle.net/11023/562>

Downloaded from PRISM Repository, University of Calgary

UNIVERSITY OF CALGARY

Context-based gait recognition

by

SHERMIN BAZAZIAN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

November 2012

© SHERMIN BAZAZIAN 2012

UNIVERSITY OF CALGARY
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Context-based gait recognition" submitted by Shermin Bazazian in partial fulfilment of the requirements of the degree of Masters.

*Supervisor, Dr. Marina Gavrilova,
Department of Computer Science*

*Dr. Mario Costa Sousa,
Department of Computer Science*

*Dr. Giuseppe Iaria,
Department of Psychology*

Date

Abstract

With the increasing demand for automatic security systems capable of recognizing people from a far distance and with as less cooperation as possible, gait recognition emerged as a very popular behavioral biometric because it is remotely observable and unobtrusive. However, the complexity and the high variability of gait patterns limit the power of gait recognition algorithms and adversely affect their recognition rates. Aiming to improve the performance of gait recognition systems without sacrificing the main advantages of gait, in this thesis, I introduce a novel multimodal gait recognition system that combines the gait patterns of the subjects with the context data related to their behavioral and social patterns. To the best of my knowledge, this is one of the only examples that the social patterns of the subjects have been used as a source of information in a multimodal biometric system. This thesis introduces a well-defined framework for defining, modeling, learning, storing and matching context data in a gait recognition system. The proposed behavioral modeling and matching framework is very flexible and can easily be adapted to different applications and multimodal biometric systems. According to the conducted experiments, the proposed gait recognition system can achieve significant improvements in the performance at a very low computational cost. The comparison of the method with other existing methods in the same area shows that the proposed approach is applicable and effective.

Acknowledgements

First and foremost, I owe my sincere gratitude to Dr. Marina Gavrilova, my supervisor, for her continuous and unconditional leadership and support during my master program. This work would not have been possible without her tireless guidance, patience, scholarly inputs and insightful comments.

I would also like to thank my fellow students in the biometric technologies lab, Dr. Kushan Ahmadian, Mr. Padma Polash Paul and Mr. Priyadarshi Bhattacharya, for their help in preparing and completing this thesis.

Last but not least, I am indebted to my family for their support and encouragement throughout every single step of my academic and personal life.

Table of Contents

Approval Page.....	ii
Abstract	iii
Acknowledgements.....	iv
Table of Contents	v
List of Tables	vii
List of Figures and Illustrations	viii
List of Symbols, Abbreviations and Nomenclature.....	x
 CHAPTER 1 : INTRODUCTION.....	 1
1.1 Motivation.....	2
1.1.1 Biometric characteristics.....	2
1.1.2 Advantages of gait recognition	4
1.1.3 Challenges of gait recognition	5
1.1.4 Methodology	7
1.2 Contributions.....	10
1.3 Thesis organization	11
 CHAPTER 2 : PRELIMINARIES AND LITERATURE REVIEW	 13
2.1 Introduction.....	13
2.2 Literature review on gait recognition.....	13
2.2.1 Subject detection and silhouette extraction.....	17
2.2.2 Gait cycle detection.....	19
2.2.3 Feature extraction.....	20
2.2.3.1 Model-based approaches.....	20
2.2.3.2 Model-free approaches.....	23
2.2.4 Matching and decision making	27
2.3 Behavioral profiling.....	29
2.4 Context extraction	31
2.5 Multimodal gait recognition	33
2.5.1 Information fusion	35
2.5.2 General Framework	38
2.5.3 Some recent research on multimodal gait recognition.....	42
 CHAPTER 3 : PROPOSED METHODOLOGY FOR CONTEXT-BASED GAIT RECOGNITION	 47
3.1 Introduction.....	47
3.2 The gait recognition system	52
3.2.1 Preprocessing	54
3.2.2 Gait cycle detection.....	56
3.2.3 Feature extraction.....	58
3.2.4 Matching	60
3.2.4.1 Training phase.....	60
3.2.4.2 Identification phase.....	61
3.3 The Context Matcher	62

3.3.1 Behavioral patterns module	64
3.3.1.1 Behavioral pattern definition	65
3.3.1.2 Behavioral patterns modeling	68
3.3.1.3 Behavioral patterns learning	72
3.3.1.4 Creating the context database	75
3.3.2 Context extraction	76
3.3.3 Matching the context	76
3.4 Information fusion	80
 CHAPTER 4 : IMPLEMENTAION DETAILS, EXPERIMENTATIONS AND RESULTS	84
4.1 Implementation details	85
4.2 Experimental data	86
4.2.1 The gait databases	86
4.2.1.1 Dataset A from CASIA gait database	87
4.2.1.2 HumanID Gait Challenge Dataset	88
4.2.2 The context databases	90
4.2.2.1 Creating virtual context database	90
4.2.2.2 Learning context database from real data	93
4.2.2.3 Context tagging	95
4.3 Performance evaluation	96
4.3.1 Partitioning the dataset	97
4.4 Experiments	98
4.4.1 Experiment 1: visual inspection of the GEIs	98
4.4.2 Experiment 2: comparison of three behavioral modeling approaches	104
4.4.3 Experiment 3: comparison with similar gait recognition systems	112
4.4.4 Experiment 4: influence of confidence values and weighted sum	117
4.4.5 Experiment 5: comparison of behavioral patterns learning approaches using real data	119
4.4.6 Experiment 6: comparison of information fusion techniques using real data	121
4.4.7 Experiment 7: speed	122
4.5 Summary	122
 CHAPTER 5 : SUMMARY, CONCLUSION AND FUTURE WORK	124
5.1 Thesis summary	124
5.2 Conclusions	124
5.3 Contributions	127
5.4 Future work	128
 REFERENCES	131

List of Tables

Table 1-1: Comparison of various biometric traits (2)	3
Table 4-1: The twelve probe sets of the Human ID Gait Challenge dataset V: Viewing angle, Sh: Shoe type, S: Surface type, B: Briefcase, T: Time.	88
Table 4-2: The performance of the proposed system for the CASIA gait dataset and virtual context database (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context).....	105
Table 4-3: The performance of the proposed system for the Human ID Challenge Gait dataset and virtual context database (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context)	108
Table 4-4: The amount of improvement achieved by fusing the virtual behavioural patterns for HumanID Challenge Gait dataset (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context)	111
Table 4-5: Comparing the performance of the proposed system on HumanID challenge dataset and virtual Gaussian context database with similar gait recognition systems.....	114
Table 4-6: The influence of having confidence values less than one for the subjects on the performance of the system for CASIA gait dataset. (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context).....	118
Table 4-7: The performance of the system for the Human ID Challenge Gait dataset using real context information and two learning approaches (NC= No Context, RML: Random Models Learning, BPL: Behavioral Profiles Learning).....	120
Table 4-8: The performance of the system for a subset of the Human ID Challenge Gait dataset using real context information and two learning approaches (NC= No Context, RML: Random Models Learning, BPL: Behavioral Profiles Learning)	121

List of Figures and Illustrations

Figure 2-1: General framework of a gait recognition system	16
Figure 2-2: An example of background subtraction (24).....	18
Figure 2-3: One gait step.....	20
Figure 2-4: Examples of body models used in gait recognition, (a) and (b) (28), (c) (29),(d) (30).....	21
Figure 2-5: An example of 2D model-based gait feature extraction (a) Skeleton extraction (b) Noise removal (c) Stick figure fitting (30).....	22
Figure 2-6: The layered deformable model and its corresponding four layers (33)	23
Figure 2-7: Normalized binary silhouettes and their corresponding Gait Energy Image (35).....	25
Figure 2-8: The general framework of an after matching multimodal biometric system (adapted from (11))	39
Figure 2-9: The general framework of a sensor level multimodal biometric system (adapted from (11))	40
Figure 2-10: The general framework of a feature level multimodal biometric system (adapted from (11))	41
Figure 3-1: The general framework of the proposed context-based gait recognition system	51
Figure 3-2: The flowchart of the gait recognition system.....	53
Figure 3-3: Examples of noisy data	55
Figure 3-4: Preprocessing results, left: original binary silhouette, right: normalized binary silhouette after preprocessing	56
Figure 3-5: One gait step.....	56
Figure 3-6: The number of foreground pixels in the lower half of the silhouette in each frame of the gait sequence	57
Figure 3-7: Some examples of Gait Energy Images for two persons and four video sequences per person.....	59
Figure 3-8: The flow chart of the context matcher	64

Figure 3-9: A sample Questionnaire	68
Figure 3-10: Flowchart of the information fusion block	83
Figure 4-1: The program GUI.....	85
Figure 4-2: Silhouette examples from the CASIA gait database	87
Figure 4-3: Silhouette examples from the HumanID gait dataset. Top row: a subject walking on a concrete surface, Bottom row: a subject walking on a grass surface ..	89
Figure 4-4: Gait samples from HumanID challenge gait dataset with different context (a) left camera and concrete surface (b) right camera and concrete surface, (c) left camera and grass surface (d) right camera and grass surface (27)	94
Figure 4-5: The binary silhouettes and the resulting GEIs for 4 subjects of CASIA gait dataset	99
Figure 4-6: Examples of the Gait Energy Images for four subjects of CASIA gait dataset. Each row: the Gait Energy Images for four different gait samples of the same subject	100
Figure 4-7: The binary silhouettes and the resulting GEIs for 4 subjects of HumanID gait dataset	102
Figure 4-8: Examples of the Gait Energy Images for four subjects of HumanID gait dataset. Each row: the Gait Energy Images for four different gait samples of the same subject	103
Figure 4-9: The CMC curve of the proposed system for CASIA gait dataset and virtual context database. Top row: independent parameters, Bottom row: joint parameters	107
Figure 4-10: The CMC curve of the proposed system for first eight probe sets of HumanID challenge dataset and virtual context database.	110
Figure 4-11: Graphical Rank 1 performance comparison of the proposed methods with three similar approaches on HumanID challenge dataset	115
Figure 4-12: The influence of having confidence values less than one for the context database on the performance of the system for the CASIA gait dataset.....	118

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
BPL	Behavioral Profiles Learning
CASIA	Chinese Academy of Sciences
CMC	Cumulative Match Characteristics
DTW	Dynamic Time Warping
GC	Gaussian Context
GEI	Gait Energy Image
GMI	Gait Moment Image
HMM	Hidden Markov Model
NC	No Context
PCA	Principal Component Analysis
PC	Profiles Context
RC	Random Context
RML	Random Models Learning

CHAPTER 1 : INTRODUCTION

In our modern world of communication and technology, security plays a vital role. Controlling access to protected areas, making sure only trusted people are using the secure data, monitoring and controlling borders are some of the areas essential for ensuring security both at the national and the international levels. All of these areas require establishing the identity of individuals in an efficient and reliable manner. The traditional approaches for identifying a person generally work based on “what the person has” (ID cards, credit cards, keys) and/or “what the person knows” (passwords, pins, codes). However, issues with carrying physical objects or remembering some information are well known and can lead to stolen, forgotten or lost tokens of authentication (1). In addition, with a lot of communications happening in the cyber world and without a face-to-face contact, the need for a secure and automatic identification mechanism becomes more and more critical. To address these problems, the biometric systems are introduced as systems that automatically identify a person based on “who the person is” using his/her physical or behavioral characteristics. “Biometric term comes from the Greek words bios (life) and metrikos (measure)” (2). It is an intrinsic characteristic of the person, thus, it cannot be stolen, barrowed or forgotten and clearly it is significantly harder to forge (2).

Among the variety of biometric systems, the gait recognition systems that recognize people based on the way they walk recently become very popular due to the unique properties of gait and advancements in video processing technologies. *In this thesis, I introduce a multimodal gait recognition system that, for the first time, combines the gait patterns and the behavioral patterns of the subjects to make a more accurate identification.* Since the context of the walking sequences is used for inferring and

matching the behavioral patterns, the resulting system is a **context-based gait recognition system**.

This chapter provides a general overview of the thesis. The motivation behind the thesis and the overall methodology used for designing the proposed context-based gait recognition system are described in Section 1.1. Afterwards, Section 1.2 presents the contributions of this thesis. Finally, Section 1.3 describes the organization of the rest of the thesis.

1.1 Motivation

Gait recognition has recently become an attractive topic in the biometric research. To justify why and how the gait recognition systems became so popular, first the biometric characteristics and the critical factors in their selection are introduced. Having a base for comparing different biometric characteristics, the main advantages and challenges of using gait for individual identification are discussed. Finally, having introduced the objective of the thesis, the proposed methodology for addressing the mentioned shortcomings of gait recognition systems is briefly explained.

1.1.1 Biometric characteristics

The biometric characteristic is essentially any characteristic of an individual that is measurable and distinctive (3). The biometric characteristics can generally be classified to two broad categories: physiological characteristics and behavioral characteristics (3). Physiological biometric recognition deals with physical characteristics of human beings. Some of the most popular examples are: fingerprint, face, hand geometry, eye patterns (iris and retina), DNA, ear and palm print. On the other hand, behavioral biometric recognition deals with behaviors of individuals including gait, signature, typing patterns

and voice (2). The most important factors that are usually considered in choosing a biometric trait are (2):

- 1- Universality: what is the percentage of people who possess the trait?
- 2- Distinctiveness: how distinctive and unique the biometric trait is?
- 3- Permanence: how the biometric trait changes over time?
- 4- Collectability: how easy the biometric data can be collected?
- 5- Performance: what is the achievable accuracy and at what expense?
- 6- Acceptability: how willing people are in providing their biometric information?
- 7- Circumvention: how easy the biometric can be spoofed?

Table 1-1: Comparison of various biometric traits (2)

Biometric Trait	Universality	Distinctiveness	Permanence	Collectability	Performance	Acceptability	Circumvention
Hand vein	M	M	M	M	M	M	L
Gait	M	L	L	H	L	H	M
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Ear	M	M	H	M	M	H	M
Hand geometry	M	M	M	H	M	M	M
Fingerprint	M	H	H	M	H	M	M
Face	H	L	M	H	L	H	H
Retina	H	H	M	L	H	L	L
Iris	H	H	H	M	H	L	L
Palm print	M	H	H	M	H	M	M
Voice	M	L	L	M	L	H	H
Signature	L	L	L	H	L	H	H
DNA	H	H	H	L	H	L	L

Table 1-1 provides a comparison of some well-known biometric traits based on the seven aforementioned factors. H, M and L stand for high, medium and low accordingly (2). As can be seen in this table, none of the biometric traits can perform perfectly based on all the seven factors and there is always a tradeoff between different factors that should be considered based on each specific application requirements and limitations.

1.1.2 Advantages of gait recognition

Taking a look at the properties of gait as a behavioral characteristic, highlighted in Table 1-1, it is clear that the main power of gait compared with other biometrics is mainly in its high collectability and acceptability. The high collectability and acceptability of gait come from the fact that this trait is unobtrusive and the attention or cooperation of the subject is not needed for collecting his/her gait (4). In the majority of biometric systems, the subjects need to first register themselves in the system and then voluntarily provide the system with their biometric data every time they use the system. In the case of gait recognition, however, a surveillance camera is sufficient for data acquisition. The gait data gathered in this way can also be used for registering the users into the system and no cooperation from the subjects is necessary at any stage of their identification. More importantly, the gait patterns are remotely observable and the camera that is recording the walking movement does not need to be in the close proximity of the subject (4). This property makes gait very appealing for security applications particularly knowing that the performance of a lot of biometric systems drops when the subject moves away from the camera. Clearly, the quicker an intrusion or illegitimate access is detected, the more time is available for the system to act accordingly. Having special equipment for capturing

walking movements, it is even possible to analyze the walking movements from a very far distance. For example, the computer vision group at Georgia Institute of Technology is conducting research on developing special radars for detecting people and capturing their walking patterns from a distance as far as 500 feet (152.4 meters) (5) (6). Additionally, the gait recognition techniques usually do not need high resolution video sequences (7) and since they mostly work only based on binary silhouettes, they are not supersensitive to illumination changes and, in fact, they can be used at nights using infrared imagery (8). In summary, the gait recognition techniques are one of the very few biometric techniques that can be used for identifying people from a far distance, without any cooperation from the subjects even at night or low visibility conditions. Having these properties, gait recognition can be used for screening and controlling activities around the parliament buildings, military bases, nuclear power plants, etc.

Although gait is most famous for being unobtrusive and remotely observable, another interesting property of gait is that gait forgery is not straightforward (Table 1-1). This property comes from the fact that imitating the walking style of another person is not physically an easy task. Putting aside the movement itself, having a similar body type is essential for the walking patterns to look the same (4). Furthermore, one cannot easily conceal the way he/she walks (9) and, in fact, it is usually possible to realize if a person is trying to walk in a different walking style.

1.1.3 Challenges of gait recognition

Despite all the unique properties, gait recognition suffers from some limitations and challenges. According to the Table 1-1, the distinctiveness and permanence of gait patterns are both low. The low permanence of gait means the gait patterns can change

from time to time and they are not always reproducible (9). Different factors are contributing to this problem. Age, mood, illness, fatigue, drug or alcohol consumption, etc. can all change the walking style of a person (9). Additionally, people walk slightly different based on the type of clothes and/or shoes they are wearing and the type of surface they are walking upon (10). Similarly, the low distinctiveness of gait indicates that gait patterns are not as unique as other biometrics in nature and with increasing the number of subjects it is not easy to prove that every subject has a different way of walking and can be identified only based on the walking patterns (6).

In summary, the main drawback of using gait for individual identification is the wide variability of the gait patterns per subject. This wide variability creates difficulty for extracting features that are robust enough to handle all the possible scenarios and are distinctive enough to distinguish the subject in a large population. Furthermore, even if such distinctive features exist, any factor that can change the appearance of the person like wearing a hat, carrying a suitcase, loose clothing, etc. can adversely affect the performance of the system by obscuring the distinctive features used for gait recognition (9). This problem becomes more critical in the condition of low quality samples, which is the case in a lot of security applications. Due to these limitations, it might not be possible to achieve high recognition rates using only the gait patterns (6).

This thesis overcomes the above problems by taking an alternative route. Instead of increasing algorithm complexity which might still not be good enough in case of poor quality samples, I propose a fully original way for improving the performance of the gait recognition systems through incorporating more knowledge about the subjects in the system and building a multimodal gait recognition system. Biometric area has just

witnessed the incredible popularity of multimodal systems being developed for increased person recognition rates. These systems consistently show advantages over single module (or traditional) biometric systems in both their recognition rates (as high as 99% for certain data samples), versatility and circumvention (11) (12). However, those systems usually consider biometric traits of the same class (i.e. face and fingerprint) or sometimes physiological and behavioral traits (i.e. fingerprint and signature). The novelty of my approach is that, for the first time to the best of my knowledge, I incorporate metadata based on social context into the standard gait recognition system. I describe the essence of this idea in the next section.

1.1.4 Methodology

Based on the previous discussions, gait recognition is very appealing for the security applications due to its obtrusiveness and remote observability. However, since the gait patterns are not as distinctive and permanent as some other biometrics (i.e. fingerprint or iris), the gait recognition algorithms have difficulty meeting the performance requirements of such applications. The solution I propose to this problem is to improve the performance of gait recognition not by investing into complex gait recognition algorithm but by integrating extra information (metadata) using the parameters that can be extracted from the context of gait video sequences.

We do not exist in a vacuum. We are constantly surrounded by other people and events, we keep track of time and appointments, we are aware of the weather conditions and traffic patterns, we are surrounded by sounds, smells and sights, we are influenced by communications with our friends or colleagues, and we share our experiences on-line with our social circle. Most of this data is readily available for analysis, and in fact is

being used by large corporations for a few decades to improve product sales through carefully collecting and analysing customer profiles (13) (14). However, no such studies so far existed in biometric domain. In my thesis, I bridge the gap and propose to use metadata which can be easily extracted from gait video to increase recognition rates even in the presence of low quality data.

Since gait is a behavioral biometric, the data sample is a video of a subject walking. Having such a video, it is possible to extract information about the context of the video including the time, location, and carrying condition of the subject. This sort of information is not normally available for physical biometrics like face, iris, fingerprint, etc. that use still images of the biometric data as their data samples. Therefore, the context of the walking sample is an important piece of information which I propose to use to improve the identification process. Reviewing multiple studies in modeling and predicting human behavioral habits and routines indicates that people normally maintain predefined routines for their everyday life and their daily activities are not completely random (15) (16) except in unusual or critical circumstances. These studies show that people typically have favorite places they frequently visit (15) (16). They keep their habits for buying groceries or doing recreation activities, socialize at certain times and locations with their social circle, and keep their daily routine fairly regular. These findings indicate that it is possible to build models for behavioral patterns of the subjects that with high degree of probability will correspond to their daily habits. Having the behavioral models for each subject, it is possible to measure how well the context of a walking sequence matches with the behavioral patterns stored in the database. This extra piece of evidence can be used by the gait recognition system to make a more accurate

decision about the identity of the unknown walking subject. This is the main idea behind my proposed context-based gait recognition system: implementing mechanisms for matching gait patterns and behavioral patterns and combining these two sources of information in decision making process to achieve better performance. For this purpose, the proposed system is implemented through the following three main modules:

- 1- The gait recognition system that identifies subjects based on their gait patterns.
- 2- The context matcher that matches the context of the video with the behavioral patterns of subjects.
- 3- The information fusion module that combines the output of the gait recognition system and the context matcher to make a more accurate identification.

The design decisions and implementation details of these three modules are presented in Chapter 3.

The proposed system can be used in a variety of applications. The main target applications are the controlled environments where users provide information about their daily regulations when registering to the system and are obligated to be consistent with their schedules. One example of such applications is access control in high security environments like prisons. Another important application of the proposed system is risk analysis and abnormal behavior detection. Since the system has information about the users' behavioral patterns, it is able to detect cases where users are violating their behavioral routines and can report such scenarios as suspicious activities. Finally, the learning capabilities of the system also enable it to be used in open areas like airports, banks, shopping centre, etc. considering reasonable limitation on the number of subjects

in the scene. The system will learn the new environment by observing the behaviors of common users and automatically extracting the repetitive patterns.

1.2 Contributions

In this thesis I introduce a context-based gait recognition system as a multimodal gait recognition system that combines the gait patterns of the subjects with their behavioral routines. The main contributions of this system can be listed as follow (17):

- 1- Introducing context-based behavioral patterns of the subjects as a new type of behavioral biometric
- 2- Developing a novel multimodal gait recognition system that, for the first time, uses the context-based behavioral patterns of the subjects as metadata not only in gait recognition but also in biometric identification domain
- 3- Developing novel methods for defining, modeling, learning and storing the behavioral patterns
- 4- Developing a technique for matching the context of the video with the behavioral patterns of the subjects
- 5- Developing, implementing and testing the overall multimodal gait recognition system using biometric fusion of gait recognition method and context-based behavioral pattern matching

Other than the novelty of the proposed system, according to the conducted experiments presented in Chapter 4, the proposed system shows the following advantages (18):

- 1- Guaranteed performance: incorporating the context data never degenerate the performance of the system.

- 2- Speed: computation time of combining the context data is very low.
- 3- Significant improvement: the amount of improvement achieved by adding the context data is significant.

Although the proposed system is completely functional, it should be mentioned that extracting the context from the videos is not fully automatic. Context extraction is a vast subject in computer vision domain that involves a wide variety of image processing techniques. Therefore, it is beyond the scope of this thesis. However, an extensive research has been done in this area and there exist commercial products in the market for this purpose that can be used in the system to make it fully automatic. More detail on this topic is provided in Section 2.4.

1.3 Thesis organization

Chapter 2 of this thesis provides background information about the main concepts of context-based gait recognition. Since the focus of this thesis is on gait recognition, this chapter begins with introducing the general framework of gait recognition systems, the existing algorithms for gait recognition, their advantages, challenges and limitations in Section 2.2. Afterwards, Section 2.3 presents the previous conducted studies on human behavioral modeling that have been used in developing the ideas of this thesis. Finally, section 2.5.3 presents the general concepts of multimodal biometric systems and the previous research done in multimodal gait recognition.

Chapter 3 presents the proposed methodology of the context-based gait recognition system in details. Section 3.2 discusses how the gait features are extracted, matched and recognized. Section 3.3 describes how the behavioral patterns of the subjects are

presented, modeled and later matched with the context of the video. At last, Section 3.4 explains how the behavioral patterns and gait patterns are eventually combined in the final identification.

Chapter 4 describes the experimentations conducted for validating the system and the obtained results. Section 4.1 provides detailed information about the implementation of the methodology. Section 4.2 introduces the data sets used for system evaluation and how they have been set up for the experiments. Section 4.4 presents the conducted experiments and the obtained results.

Chapter 5 concludes the thesis by summarizing the overall methodology of context-based gait recognition system, its contributions, advantages and limitations. Afterwards the future areas of research for improving the performance of the system are presented.

CHAPTER 2 : PRELIMINARIES AND LITERATURE REVIEW

2.1 Introduction

This chapter presents the background information necessary for understanding the core concepts of this thesis. The proposed methodology in this thesis is a multimodal gait recognition system that combines the gait patterns of the subjects with their context-based behavioral patterns and daily routines. Consequently, three main areas of research are involved in developing this methodology: gait recognition, behavioral profiling and multimodal biometric identification. The three main sections of this chapter are dedicated to these three areas. Section 2.2 is an introduction to gait recognition algorithms. This section presents the general framework of the gait recognition systems, the related existing algorithms in this area, how they have improved over time and what is still missing in these algorithms. Section 2.3 presents a number of related studies and their obtained results in the area of behavioral patterns modeling. Finally, Section 2.5 introduces the main concepts of multimodal biometric systems and how this idea has previously been used in gait recognition.

2.2 Literature review on gait recognition

Gait analysis deals with analyzing the patterns of walking movement. The source of inspiration for a lot of gait analysis techniques is the work of Johansson in (19) that showed that people can quickly recognize walking motion only from the moving patterns of a few point lights attached to the human body. Inspired by Johansson's work, Cutting and Kozlowski in (20) performed some experiments to show that the same array of point lights can be used for recognizing friends even if they happen to have similar height, width and body shapes. Considering the wide variety of potential applications for gait

analysis, the promising outcome of these studies initiated an advanced research in this field. Although gait analysis is most well-known for its application in access control, surveillance and activity monitoring, it can also be used in sports training for analyzing the movements of an athlete and giving suggestions for improvement. Medical sciences can also take advantage of gait analysis techniques in diagnosing and maybe even developing some strategies to treat patients with walking disorders or suffering from Parkinson's disease (21).

The main focus of this thesis is on the use of gait recognition for access control, security and surveillance applications. In this area of application, a gait recognition system can generally be used for two major tasks (22):

- 1- Subject verification: in this scenario, the system should make sure the subject is “who he/she claims to be”. In other words, the system should only match the subject's biometric data to the stored template for that person in the system database and decide whether they belong to the same person or not. Consequently, this is a one-to-one match (22).
- 2- Subject identification: in this application, the system should establish the identity of the subject or answer the question “who is this person”. For this purpose, the system should compare the subject's biometric data with all the existing templates in the system database and decide whether there is a match for this person in the database or not. Therefore, this is a one-to-many match and consequently it is usually more complicated and requires more processing (22).

In both scenarios, the system is presented with a video sequence of an unknown subject walking in a scene and the system must identify or verify the identity of that unknown subject. For the gait recognition system to be able to perform these two tasks, it generally should have the following main modules (7):

- 1- Subject detection and silhouette extraction: for analyzing the walking patterns of the subject, the first step is to detect the targeted subject in each frame and track him/her through the frames (7).
- 2- Gait cycle detection: gait is a periodic motion; the gait cycle detection module is responsible for detecting the starting and ending points of each gait cycle in the walking sequences that is needed for extracting the gait features.
- 3- Feature extraction: this module is the most important module of the gait recognition system; it processes the silhouettes and extracts distinctive features from the detected gait cycles (7).
- 4- Matching: this module matches the resulting extracted features to the gait patterns of the subjects previously stored in the database of the system and usually outputs a matching score for each subject (7).
- 5- Decision making: this module identifies or verifies the identity of the subject based on the result of the matching module (7).

The general framework of a gait recognition system is shown in the next page.

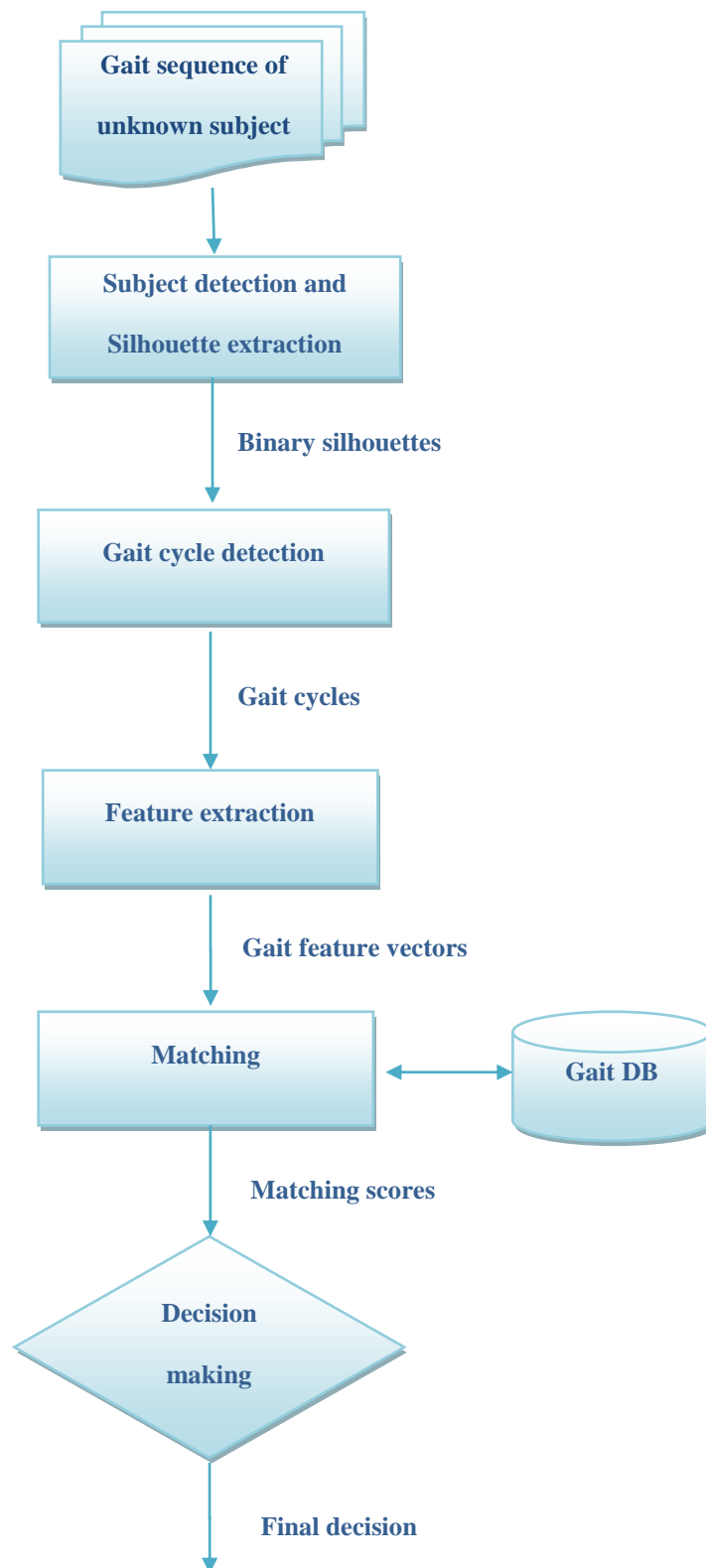


Figure 2-1: General framework of a gait recognition system

The multimodal gait recognition system introduced in this thesis follows the general framework of the gait recognition systems introduced in Figure 2-1. Therefore, it is essential to get familiar with the existing algorithms for each of these modules. The rest of this section provides more information about the main blocks of this framework, how they have been used in my system and what are the related existing algorithms and their challenges and bottlenecks.

2.2.1 Subject detection and silhouette extraction

The input of gait recognition systems is a video of a subject walking. The scene of this video can include objects or even people other than the targeted subject. Therefore, the first step of gait recognition is to detect the targeted subject in each frame and separate it from the rest of the image. The most popular technique for this purpose is background subtraction (7). The idea of background subtraction is to have a model for the background and then consider any pixel in the scene not consistent with that model as a foreground (none background) pixel and thus belonging to the subject silhouette (23). The first step of background subtraction method is to learn the background model. This background model usually represents the background color for each pixel. The background model can be known beforehand, for example a picture taken from the scene when there is nobody in the room. It can be as simple as one single image learned from the video sequences as the mean or median of all the frames. In more complicated scenarios, the background model can include color distributions mainly Gaussian distributions for each pixel. For making the model robust to lighting changes, it is possible to make the background model dynamic by updating it on a frame by frame basis (23). Once the model is built, any pixel that its distance from the background model is greater than a threshold is considered as a

foreground pixel. Some post processing might also follow afterwards to remove noise and extra objects in the scene.

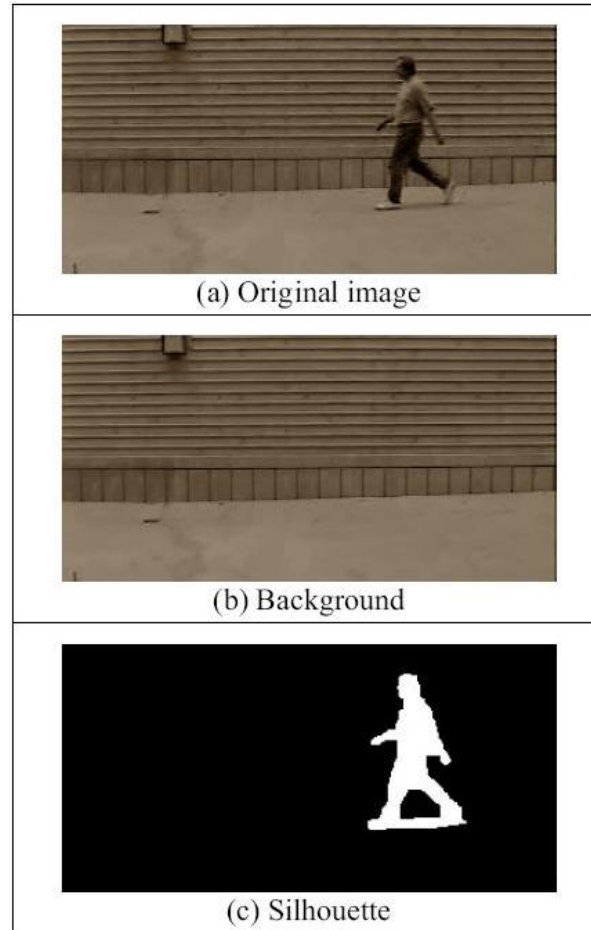


Figure 2-2: An example of background subtraction (24)

An example of background subtraction is shown in Figure 2-2. In this example, the background model is an image taken from the scene when there is nobody there (24).

In my thesis, the system has been evaluated using a number of popular publicly available gait datasets. In all of these datasets, the binary silhouettes have already been extracted from the gait sequences and are available as a part of the dataset. Consequently, I directly use the available binary silhouettes. However, a preprocessing module at the beginning of

my system does some preprocessing on the binary silhouettes to remove the noise and normalize the silhouettes for further processing of the system. This module is introduced in Section 3.2.1.

2.2.2 *Gait cycle detection*

Gait is a periodic motion (25). In the majority of gait recognition algorithms, the gait features are usually extracted for each gait cycle. Therefore, before extracting the features, it is essential to find the starting and also the ending frames of the gait cycles or, in other words, to partition the video sequence into gait cycles (26).

To illustrate the cycle detection process, one gait step is shown in Figure 2-3. Each gait cycle consists of two such steps. Since each step starts and ends with the right and left legs being together, finding the time of this alignment is useful for finding the gait cycles. A lot of gait cycle detection algorithms counts the number of foreground pixels in the legs' region to find the beginning and ending points of steps (26). As can be seen in Figure 2-3, the number of foreground pixels reaches a maximum when the right and left legs are farthest apart and reaches a minimum when the right and left legs are together. Therefore, by detecting the two subsequent minima it is possible to find when the right and left legs are together which corresponds to the beginning or ending points of the gait steps (27).

Taking advantage of this property, the gait cycle detection method that I used in my proposed gait recognition system is also based on counting the number of pixels in the legs' region of the silhouettes and finding the frames where the number of foreground pixels reaches a minimum. The detailed description of this method is provided in Section 3.2.2.



Figure 2-3: One gait step

2.2.3 *Feature extraction*

After detecting the subject and partitioning the sequence into gait cycles, the next step is to extract the gait features for each cycle. There are generally two main approaches for gait feature extraction: model-based and model-free (7). The feature extraction method used in this thesis is a model-free approach, to justify how and why this method has been selected, the model-based and model-free approaches and their advantages and disadvantages are described in the following sections.

2.2.3.1 Model-based approaches

The model-based approaches use an explicit model to model the human body (7). These methods estimate the parameters of the model in each frame. The value of these parameters and how they change over time are used as features for gait representation. The body model can be a simple 2D model or it can be a complex 3D model (28) (29) (30). The more complicated the model is, the more computation is needed for estimating its parameters. Some examples of body models used previously for gait recognition are shown in Figure 2-4.

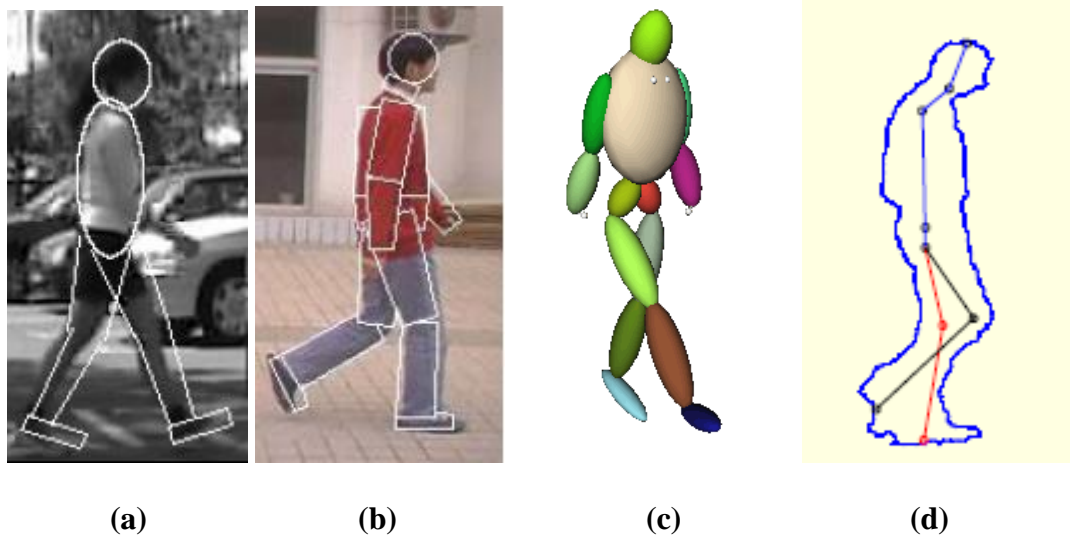


Figure 2-4: Examples of body models used in gait recognition, (a) and (b) (28), (c) (29), (d) (30)

The model-based methods for gait recognition have a couple of advantages. First, they are scale invariant and in some cases even view invariant (7). Instead of using the silhouettes directly, these methods fit a model to the silhouette. Consequently, the size of the silhouette and its viewing direction (if the body model is 3D) wouldn't have any influence on the output of these methods. Second, they can to some degree deal with occlusions and self-occlusions (25). Since body parts are modeled separately, even if some of the parts are not visible due to occlusion, there will still be a chance for other parts to be visible. Therefore, the algorithm wouldn't lose the subject and the visible body parts can be used to estimate the parameters of the invisible ones. Third, model-based methods are not extremely sensitive to appearance changes like carrying a suitcase or wearing a hat. Having a priori knowledge about how the human body should look like, these methods are able to detect the situations where a person is for example carrying an object and thus are able to exclude that object from their calculations.

On the other hand, the model-based approaches suffer from some important limitations. First, because of the high flexibility of the structure of non-rigid human body and also the problem of self-occlusion (31), the search space of these methods is huge and estimating the model's parameters is tremendously difficult (25). As a result, these methods are generally computationally expensive and time consuming (7). Second, since the estimation of the model parameters needs high quality videos, these methods are usually sensitive to the quality of the video sequences and vulnerable to noise (7).

The majority of model-based gait recognition methods use simple 2D models (32). As an example of such methods, Yoo and Nixon in (30) extract the body contour, find the skeleton and fit their model as a stick figure to each frame (Figure 2-5). They then extract their features from the resulting skeletons which include body height, cycle time, stride length, speed, average joint angles, variation of hip angles and the correlation coefficient between the left and right leg angles (30).

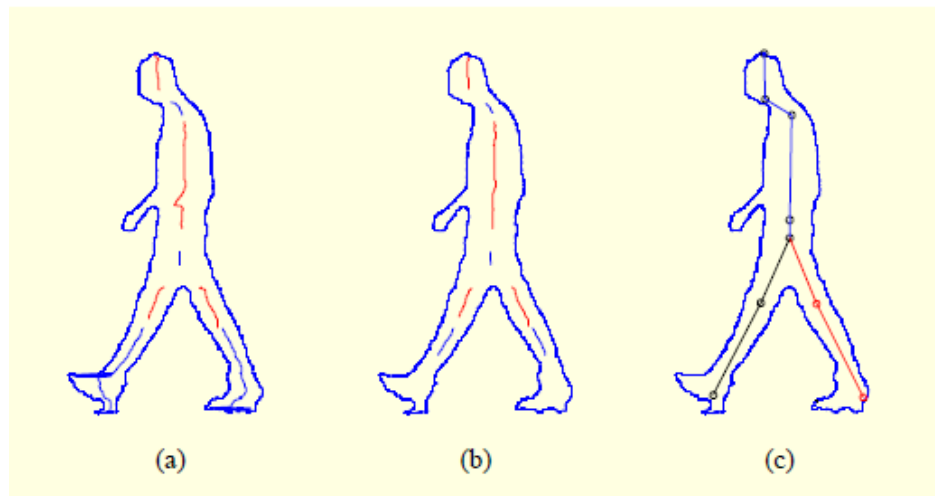


Figure 2-5: An example of 2D model-based gait feature extraction (a) Skeleton extraction (b) Noise removal (c) Stick figure fitting (30)

There are also more complicated approaches that use more detailed body models. For example, Haiping Lu et al. in (33) use a deformable body model consisted of 22 parameters capturing the lengths, widths, positions and orientations of different body parts. For modeling self-occlusion, the model is divided into four layers, each of them with different chance of being occluded. The resulting model is called Layered Deformable Model (LDM). The body model and the corresponding layers are shown in Figure 2-6. They first train the system using manually annotated silhouettes and learn the relationships between different body parts' sizes and locations. Afterwards, they use this information to estimate the parameters of the model in each and every frame.

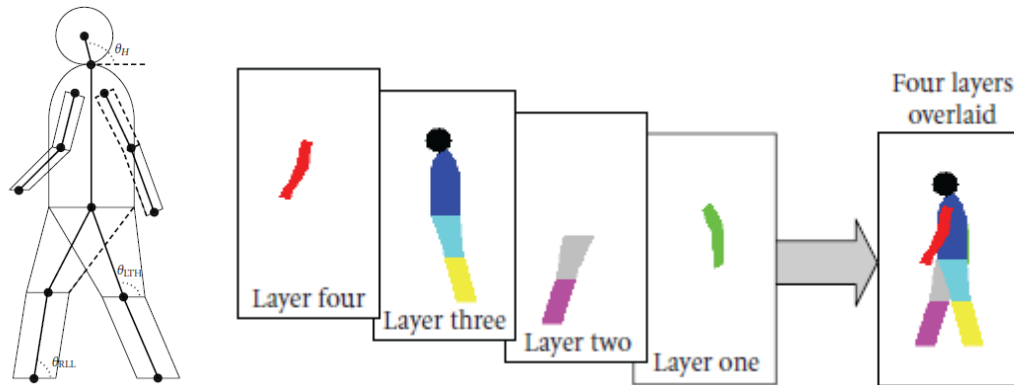


Figure 2-6: The layered deformable model and its corresponding four layers (33)

2.2.3.2 Model-free approaches

The model-free approaches instead of using a priori body model, process the silhouette as a whole to make a compact representation of walking motion (7).

The model-free approaches have a number of advantages. First, these methods are cheap and fast (7). Unlike model-based methods, they do not need to estimate the parameters of a model in each frame. Also, the processing needed to be done in each frame is generally

negligible compared to model-based approaches. Similarly not needing to search for the parameters of a complicated model in a huge search space, these methods usually only need some general information about the silhouette shape and are not supersensitive to noise and the quality of the video sequences (7). Furthermore, since these methods usually need only the silhouettes, no other information (color, texture, greyscale values, etc.) is needed for their processing. As a result, they can be used for gait recognition at nights using infrared imagery (9).

However, model-free approaches also suffer from a couple of disadvantages. Since these methods do not have any a priori knowledge about the human body and they only work based on the silhouette shape, they are generally more sensitive to factors that can change the appearance of the person and its silhouette. This includes loose clothing, wearing a hat, carrying an object, etc. (32). For the exact same reasons, these methods are also sensitive to view and scale changes (7). To summarize, the model-free approaches are not very successful in handling the unpredictable scenarios that haven't been considered in training the system.

One of the simplest model-free approaches is the work of Sharma et al. in (34) that directly use the whole silhouette sequence as the feature vector. Using the binary silhouette themselves without any further processing requires saving the whole silhouette sequence for each subject which needs a lot of storage space. Storing and matching these temporal sequences are expensive and time consuming. To solve these problems, the majority of gait recognition algorithms try to compress the silhouette sequence into one single template in order to save storage space and also computation time. One of the most popular approaches developed in this direction is the Gait Energy Image (GEI)

introduced by Han and Bhanu (35). GEI is one single image obtained by finding the average of all silhouette images of a single gait cycle. Some examples of Gait Energy Images are shown in Figure 2-7. To avoid over fitting and to make the method more robust to little distortions, shadows, missing body parts, scale changes, etc. they generate some synthetic GEIs by adding distortion to the lower part of the real GEIs for each individual and they use both real and synthetic gait images for their final recognition.

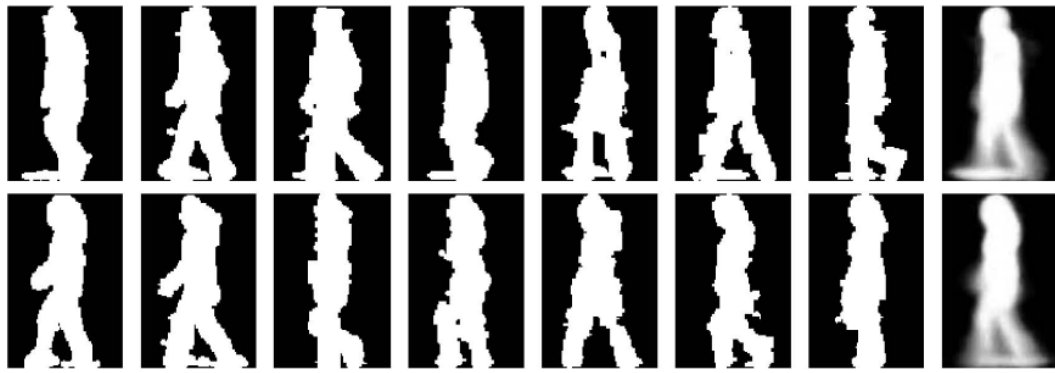


Figure 2-7: Normalized binary silhouettes and their corresponding Gait Energy Image (35)

GEI is an efficient and compact representation of gait and it also reduces the noise by averaging (31). Therefore, a lot of recent methodologies in the field of gait recognition are extensions of GEI (9) (36) (37) (10) (25) (38).

Having introduced both model-free and model-based approaches, comparison shows that the majority of gait recognition algorithms in the past few years are model-free approaches. This observation can be justified by the fact that the main attraction of gait as a behavioral biometric is its application in security scenarios for detecting the intruders in the quickest possible way from a far distance and in low visibility conditions. Consequently, using a method that needs high quality images to perform well and takes

long time to process is in contradiction with the requirements of such applications. Therefore, the complexity and the noise vulnerability of model-based approaches can make them less popular than the model-free approaches. Based on this observation, I decided to use a simple and fast gait feature extraction algorithm from the category of model-free approaches in my context-based gait recognition system. From the great number of methods developed based on the original idea of GEI introduced by Han and Bhanu in (35), it is clear that GEI has been one of the most popular gait recognition techniques in the last few years. The GEI represents the whole gait cycle by one single image, thus it is a very compact representation which is beneficial in terms of both storage space and computational time. As a result this method is cheap, fast and efficient. Furthermore, the idea is straightforward, easy to understand and implement and at the same time it shows acceptable results (35). For all these reasons, I decided to use GEI as the gait feature in my proposed context-based gait recognition system. The majority of proposed extensions and variations of GEI in literature (9) (36) (37) (10) (25) (38) focus on making gait recognition more efficient by improving the algorithm used for creating the templates. Due to the mentioned complexity and high variability of gait patterns, this kind of approach can result in complicated algorithms sacrificing the speed and simplicity of the system. In this thesis, however, as a novel approach, the performance of the gait recognition system is improved by integrating the GEIs with the behavioral patterns of subjects' everyday activities. The conducted experimentations of this thesis show that this integration can achieve significant improvements in accuracy without adding a lot of computation.

2.2.4 *Matching and decision making*

The last step of a gait recognition system is to match the features extracted from the unknown sequence to the existing templates in the system database, calculate a matching score for each subject as a measure of its similarity to the unknown sequence and establish or verify the identity of the unknown subject according to the resulting matching scores. The method used for matching the templates and calculating the matching scores generally depends on the type of gait features extracted by feature extraction module. There are typically two main categories of gait features (25):

1- Temporal sequence:

In this category, the features are extracted from each frame independently and the sequence of the extracted features is used as the final feature vector. In other words, these methods represent gait as a temporal sequence (25). Therefore, they need a lot of storage space to store the resulting feature sequences for each subject. Furthermore, they need to train a sequence matching algorithm for each subject. This algorithm will be used to match an unknown gait sequence with subject's walking patterns and calculate the probability that they belong to the same person. Training such a framework usually needs a lot of training data and is also time consuming. Additionally, these methods require complex sequence matching to calculate the matching scores which can be both computationally expensive and time consuming (25). Hidden Markov Model (HMM) and Dynamic Time Warping (DTW) are two of the most popular models used for matching the temporal sequences (7).

2- Single template:

These methods extract the features from each frame but in the end all the features are combined together into one template. In other words, they represent the whole gait cycle by one single template (25). By doing so, they save the storage space and also computation time (35). Therefore, they create a very compact representation (9). The resulting template should satisfy the following properties (39):

- capturing structural information of gait patterns
- capturing dynamic information of gait patterns
- providing a compact representation with small number of features
- being robust to speed changes

The single template matching methods are not extremely sensitive to silhouette noise, holes, shadows, and missing parts (9) (10). However, they are sensitive to appearance changes (10). If the output of the feature extraction module is a single template, the matching scores are normally calculated by finding the distances between the templates. Once the matching scores are obtained, the decision making module will establish the identity of the unknown subject. The nearest neighbor classifiers and support vector machines are two of the most popular methods used for this purpose (7).

Since the feature extraction method that I use in this thesis represents the whole gait cycle as one GEI, the output of my feature extraction module is one single template. As a result, the Euclidean distance between the resulting GEIs are used in this thesis for matching the gait patterns. More information about how this procedure is done in my system is provided in Section 3.2.4.

2.3 Behavioral profiling

Human behavior is a very general term and in this thesis I use it to describe habits and daily routines that people usually follow in their everyday life. Studying and modeling such behavioral patterns is an emerging area of research (16). However, conducting such studies is not straightforward since there is no consensus on a proper tool or mechanism for monitoring, measuring, collecting and storing an adequate amount of data needed for simulating behaviors precisely (15). Considering this limitation, the development of mobile devices, that can provide information about people's location, has been very useful in conducting multiple studies focusing on the mobility patterns of the subjects. With the increasing popularity of the mobile phones, the mobile phone companies have become convenient resources that can provide the studies with information about the mobility patterns of an extensive number of subjects. The results of these studies can be useful in traffic planning, resource management, monitoring and analyzing spreading of contagious diseases, etc. (15) (16). As one recent example of such studies, Gonzalez et al. in (15) conducted a study on the mobility patterns of 100,000 individuals randomly selected from a population of six million phone users during a six-month period. Each time a user uses his/her phone, the location of the closest mobile tower to the person is recorded by the mobile phone carrier and can be used for approximating the location of the user. The results of this study show that people have regular patterns of behavior and they tend to visit the same few places at the same times. According to this study, regardless of the distance the individuals travel each day, their trajectories over time show similar properties like periodicity. Having similar trends in the trajectories indicates

that it is possible to model the movement patterns of people with mathematical models and there is no need to build a different model for each individual (15).

In a similar research, Song et al. in (16) studied the whereabouts of 50,000 individuals randomly chosen from ten million mobile phone users over a three month period and they obtained similar results. Based on this research, although people travel different distances each day, they are all to the same amount predictable. In this study, they calculated the entropy of people's trajectory and, based on the obtained results, they claimed that people are 93% predictable in their mobility patterns (16).

Based on the results of the mentioned studies, one can conclude that people tend to follow their habits and they each have their own patterns of behavior that they maintain in their everyday life (16). This result indicates that it is possible to build models for behavioral patterns of people and there is a good chance (according to (16) 93%) that people will follow those models. The similarity of people's mobility trajectories obtained by (15) shows that these models can be defined as mathematical models with each individual following the same model but with different parameter values.

Another evidence of modeling and predicting human behavioral patterns can be found in well explored domain of market analysis and customer profiling. Almost all major retail companies, travel agencies, car dealerships, and groceries stores have list of customers with very detailed information stored on their shopping preferences. The data is collected first time once person becomes a customer of a store or joins their loyalty program and usually includes name, address, phone numbers, gender, age, and birthday. It however becomes soon extended with specific shopping patterns: day of the week, time of transaction, average amount spent, family/friends affiliations, and goes to such details as

predicting changes in marital status, loss of job, having a child, assuming a mortgage, moving to another part of the city, etc. All this data is then used to effectively market product or service to a customer (13) (40) (41).

With the advent of web-based technologies, even more information from social networks becomes available for mining to add to the existing profile. Based on these findings, the behavioral patterns of the subjects can be modeled in an efficient manner and the resulting behavioral models can be used for their identification (42). However, they are mostly explored in business, retail, hiring, and very rarely till today have been looked in the context of traditional biometric systems. This trend is now starting to change. Based on observing and studying this property, I decided to use the behavioral patterns of the subjects as a new type of behavioral biometric characteristic. I design a framework for modeling and matching the behavioral patterns and then I integrate it with the gait patterns of the subjects in a multimodal gait recognition system.

2.4 Context extraction

One of the most important steps of the proposed context-based gait recognition system is to measure how well the context of a video matches with the behavioral patterns of the subjects and use this extra information to make gait recognition more accurate. Therefore, context extraction is an essential step in this framework and it involves defining and extracting contextual parameters from video sequences. There are a variety of context parameters that can be extracted from the gait videos including time, date and location of the gait video and the carrying status of the subject. In the following, I briefly explain how these parameters can be extracted from the gait videos.

To the best of my knowledge, all digital cameras store the original time and date of each image/video they capture as a part of the image/video file (43). Therefore, the recording time and date of a video can normally be extracted using this metadata with 100% accuracy. Furthermore, the majority of surveillance cameras are able to display time and date stamps on the videos. Consequently, if for any reason the date/time cannot get extracted from the mentioned metadata, character recognition techniques can be used for extracting the date/time from these date/time stamps. Character recognition is an automatic process that converts the picture of a text (handwritten or typed) to the corresponding “machine encoded text” that can be used for further processing (44). Character recognition is a well explored domain in image processing and there are a variety of commercial systems that can be used for this purpose and they can achieve accuracies as high as 98% (44).

Furthermore, in a lot of security/access violation scenarios, the surveillance camera that recorded the video and its location are known to the system and can be used with no further processing. However, for the system to be able to handle completely unknown videos, scene and landmark recognition algorithms can be used for recognizing the video location. Scene recognition is an active area of research in computer vision that automatically recognizes the type of the scene (forest, mountain, street, office, kitchen, etc.) both for indoor and outdoor scenarios (45) (46). Scene recognition particularly for outdoor scenes shows acceptable results and can readily be used in the system. Landmark recognition is another computer vision technique that models and recognizes the landmarks from the images. There are a variety of landmark recognition frameworks including the one introduced by Google that can be used for this purpose (47) (48).

Finally, for extracting more information from the video including the type of object that the subject is carrying (backpack, suitcase, etc.), object recognition algorithms can be used (49). Object recognition is a computer vision technique that automatically detects and recognizes a given object in an image based on the appearance or other specific features of the object (50).

In summary, fully automated context extraction is a vast area of research beyond the scope of this thesis. However, a novel context tag assignment methodology has been proposed to label videos with context data in this thesis. In addition, experiments on a real gait database with already extracted context have been carried out.

2.5 Multimodal gait recognition

In order to successfully combine person's features obtained from different biometric sources, it is essential to develop a framework of multi-biometric system with appropriate decision making mechanism. Biometric field has witnessed a tremendous growth of multi-biometric research over last ten years. It is the fastest growing domain of biometric research with proven increase in performance and accuracy of the system and relatively low amount of investment in cost and developmental time (11) (51). Multimodal biometric system is a biometric system that uses more than one source of information in the decision making process to improve the performance and reliability of the system. There are a variety of motivations for using multimodal biometric systems, some listed in the following.

- 1- Better performance: As a general rule, the more information taken into account in the decision making process, the more reliable the final decision will be (52). This is especially important in gait recognition for the following reasons. First of all, the gait

patterns are not as unique and distinctive in nature. Second, the gait patterns show high variability over time for example the walking style of the subject can change based on his/her clothing, shoe type, mood, age, etc. Third, sometimes gait data are hard to obtain in high quality, for example in a lot of surveillance applications the gait videos can be recorded in highly crowded scenes and from a far distance. Therefore, the performance of a system solely working based on gait might not be satisfactory. Using a combination of biometric traits, however, can compensate for high intra class variability, noisy data and other complications and make the system more robust to the unforeseen situations that can happen in real scenarios (52) (22).

- 2- More coverage: the coverage of a biometric system shows the range of people the system can collect the biometric data from and hence identify (52). In real scenarios, some people might not be able to provide their biometric data in an acceptable quality (51). For example, there can be cases that the subject cannot walk due to temporal or permanent physical disabilities. In such cases, the system won't be able to identify the subjects. However, since a lot of multimodal biometric systems can actually work without having data for all modalities, if a person fails to provide his/her biometric information for some modalities, the system will still be able to identify the person using the other available modalities. This can reduce the failure to enroll rate substantially (52) (22).
- 3- Harder forgery: since a multimodal system combines different sources of information, for a person to trick the system he/she is someone else, he/she actually has to imposter all the biometric traits used by the system simultaneously which is clearly more challenging (12). Furthermore, some multimodal biometric systems have the

ability of asking for a random set of biometric traits each time the subject is using the system, this ensures both the presence of a real person and also more difficult forgery (52).

The multimodal context-based gait recognition system that I introduce in this thesis has all the three mentioned advantages. Involving more information about the subjects in the form of their behavioral patterns improves the performance of the system. It increases the coverage of the system by letting some users use the system only by their behavioral patterns (if the security level of the system allows doing so). Finally, since the imposter needs to know about the real subject's behavioral routines to pass the decision making process of the system, forgery is more challenging.

One of the main modules in a multimodal biometric system is the information fusion block that is responsible for combining the chosen sources of information and making the final decision. The most common approaches available for this purpose are described in the next section.

2.5.1 Information fusion

One of the main decisions in information fusion is about the stage at which the sources of information should be combined. Generally any biometric system including the gait recognition biometric system normally has the following the four following stages (2) (22):

- 1- Data collection: this module acquires the biometric data using sensors.

- 2- Feature extraction: this module processes the collected data and extracts the features. The extracted features for the enrolled users are stored as templates in the system biometric database.
- 3- Matching: this module matches the extracted features from the unknown user with the existing templates in the system database and outputs the resulting matching scores.
- 4- Decision making: this module recognizes or verifies the identity of the person based on the matching scores obtained from the matching module.

As can be seen in Figure 2-1, the general framework of gait recognition systems has these four main stages. The subject detection and silhouette extraction and also the gait cycle detection module are both preprocessing modules for feature extraction and they can be considered as a part of the feature extraction stage in the above definition. Based on this generalization, the information fusion can be performed at any of these four stages resulting in two main categories of fusion: before matching and after matching. Before matching information fusion techniques include sensor level and feature level information fusion. After matching information fusion techniques include match score level, rank level and decision level information fusion (12). These five types of information fusion are described in the following.

- 1- Sensor level: in this method, the information fusion is done after the data collection stage and before extracting features. In this type of fusion, raw biometric data acquired from multiple sensors are put together to make the final biometric data.

- Preprocessing techniques are typically needed to make different samples compatible before combining them in one single biometric data (53).
- 2- Feature level: in this category, information fusion is performed after the feature extraction stage and multiple sources of information in the form of multiple set of features are put together to form one big feature vector (53).
 - 3- Match score level: in this approach, the information fusion is performed after the matching stage. Each source of information has its own matching algorithm which outputs a matching score for that modality. The matching scores are later combined to make one final matching score (12). The final score is normally obtained by normalizing the individual matching scores and then finding their sum, average, weighted average, maximum, minimum or using more complicated methods like decision trees or Bayesian methods (54).
 - 4- Rank level: similar to the match score level, in this approach the information fusion is performed after the matching stage. Each source of information has its own matching algorithm which produces a rank list of the preferred candidates (12). Therefore, in this method each candidate is assigned a rank. The resulting ranks are then combined to make the final rank list. Some of the most well-known approaches in this category are highest rank, Borda count, Logistic regression (53).
 - 5- Decision level: in this approach, each source of information has its own matching algorithm that outputs its final chosen subject. The outputs of the different matchers are then combined to make the final decision. The most famous approaches include majority voting, weighted majority voting, decision table, Bayesian methods, Dempster-Shafer theory of evidence and behavior knowledge space method (22) (51).

Typically, fusion at the earlier steps is believed to provide better results because there is more information available in the raw data and extracted features compared to the output of different matchers in the form of matching scores or ranks (53) (12). However, before matching information fusion usually suffers from the curse of dimensionality resulting from putting together a lot of different features and also the inconsistency between the different features being combined (52) (53). For the same reasons, the before matching information fusion (sensor level and feature level) does not look appropriate for this thesis. There is a big difference size-wise and format-wise between the gait features and the context-based behavioral features that have been used in this thesis. Therefore, putting the two features together in one single feature does not look rational. Furthermore, the decision level information fusion is also not suitable for my system because the behavioral patterns are not as unique as the gait patterns and there are a lot of subjects with similar behavioral patterns. Therefore, making a decision only based on the behavioral patterns without considering the gait patterns and then combining the decisions later is not a good idea. Consequently the approach used in this thesis for information fusion is a match score level information fusion. Match score level techniques are the most popular information fusion approaches (53). They are easy to implement and understand and at the same time they show promising results.

2.5.2 General Framework

The general framework of a multimodal biometric system depends on the information fusion strategy. In after matching information fusion, the multimodal biometric system is a combination of multiple biometric systems each working independently and then an

information fusion block that combines the output of those systems. The general framework of an after matching multimodal biometric system is shown in Figure 2-8.



Figure 2-8: The general framework of an after matching multimodal biometric system (adapted from (11))

In before matching information fusion, however, the integration happens at the earlier stages of the system. Therefore the individual biometric systems are sharing functionalities in matching and decision making modules. The general frameworks of the sensor level and feature level multimodal biometric systems are shown in Figure 2-9 and Figure 2-10 correspondingly.

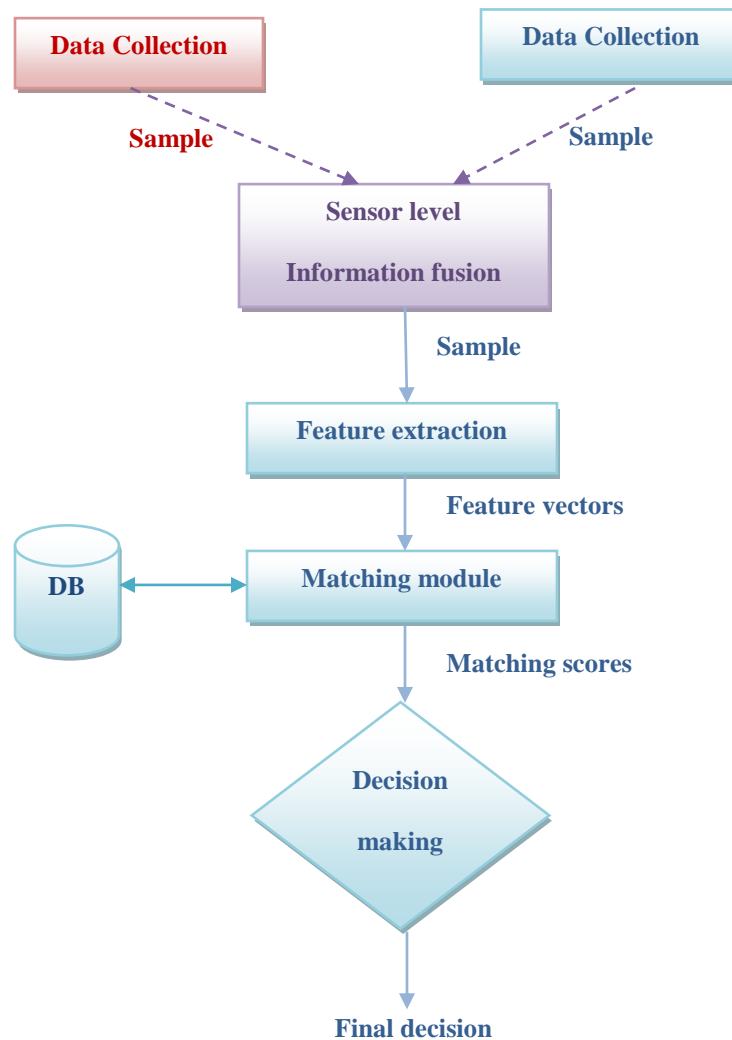


Figure 2-9: The general framework of a sensor level multimodal biometric system (adapted from (11))

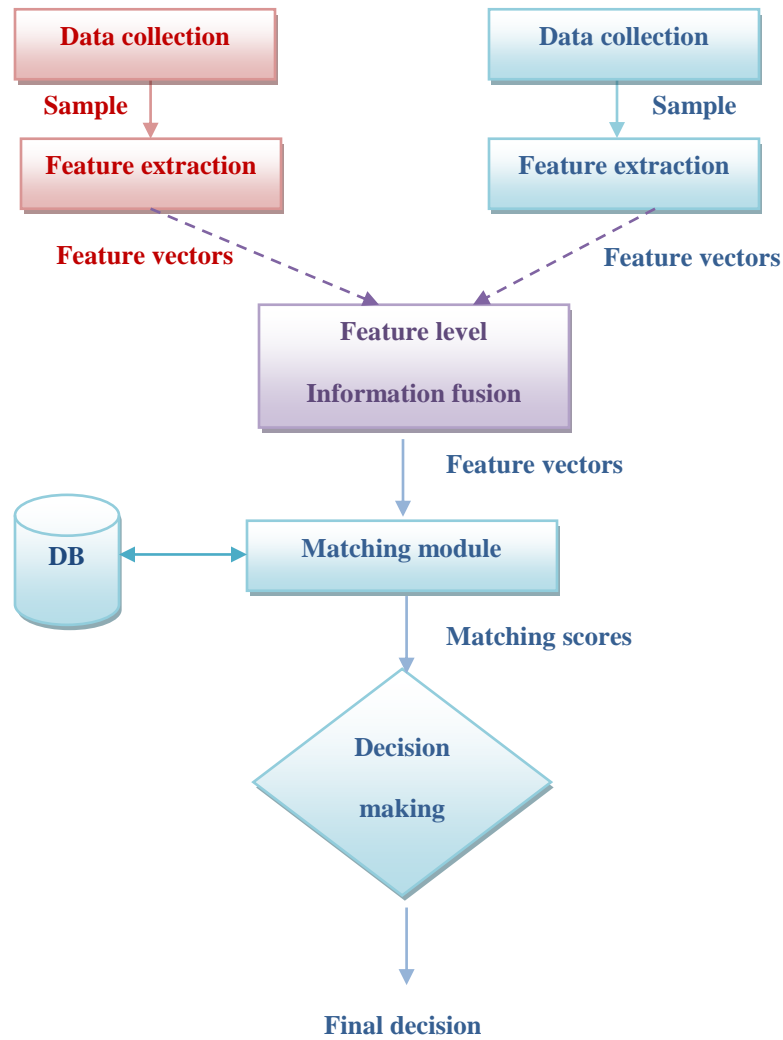


Figure 2-10: The general framework of a feature level multimodal biometric system (adapted from (11))

Since the information fusion technique that I used in this thesis combines the two sources of information after the matching stage, the general framework of my context-based gait recognition system follows the framework of the match score/rank level information fusion techniques presented in Figure 2-8. In other words, in my system I considered the gait patterns and the behavioral patterns as two separate biometric characteristics. I developed a matcher for each of these two modalities that output the corresponding

matching scores. The two matching scores are then combined using the information fusion techniques. The framework of my system is presented in Figure 3-1 and Section 3.4 provides a comprehensive description of how information fusion is done in my context-based gait recognition.

2.5.3 Some recent research on multimodal gait recognition

As established before, multimodal system research is an emerging new direction. Thus, it is expected to see that some works attempt to combine gait recognition methods in the context of multimodal system with other biometric traits. This section provides a brief overview of a number of related multimodal gait recognition systems. A multimodal gait recognition system is a biometric system that uses more than one source of information for identification but at least one of the sources is gait. Most of the existing works, however, focus on using multiple features from gait videos analyzed, not combining different biometrics together.

As mentioned, the majority of existing multimodal gait recognition systems only use gait as their biometric trait but have different algorithms for extracting a variety of gait features that are then combined using after matching information fusion techniques. As an example, Cuntoor et al. in (8) extract multiple gait features, match them separately and then combine the results of different matchers to make the final decision. The features used in this work are:

- Left and right projections of the silhouette to capture the motion of hands and legs
- Width vector of front view sequences to capture changes in the subject's height
- Width vector of the lower part of the silhouette to capture the leg dynamics (8)

For the first two features DTW and for the third feature HMM are used for finding the matching scores. The final decision is made by combining the resulting matching scores at match score level using sum, product and min operators (8).

Due to the high popularity and simplicity of GEI (the gait feature used in this thesis), there exist a variety of multimodal gait recognition systems that use GEI as one of their gait features. However, most of them are based on extracting multiple gait features and do not use more than one biometric trait. The results of some of these systems are later compared with the proposed context-based gait recognition system in Chapter 4.

As a related example, Han and Bhanu in (35) use GEI as their gait feature but they synthetically create some synthetic GEIs by adding distortions to the real GEIs and use the synthetic data as a new source of information to make the method more accurate. Two different classifiers are used for real and synthetic data and the results of the two classifiers are combined using match score level information fusion.

In another approach, F'elez et al. in (55) improve the performance of gait recognition by calculating more than one GEI for each subject, identifying the subject independently based on all these GEIs and combining the resulting decisions at decision level. For this purpose, they segment the gait cycle into four key poses. They extract the silhouette in each frame and classify that frame as one of the four key poses. Subsequently, they obtain a GEI for each key pose by calculating the average of all the frames assigned to that key pose. The GEI for each key pose is classified separately using a nearest neighbor classifier and the final identification is done by majority voting of different classifiers.

Ma et al. in (32) introduce a multimodal gait recognition system that combines GEI with other gait features at feature level. They define a number of key frames in each gait cycle. Afterwards, for each key moment they obtain a Gait Moment Image (GMI) by calculating the average of the corresponding key frames of different gait cycles. Then for each key moment, they obtain the Deviation Moment Image by calculating the deviation of the corresponding key frames from the corresponding GMI. These images, in addition to the GEIs of the gait cycles, are combined at feature level and used as their final feature for individual identification.

To get the most benefits from multimodal gait recognition including better coverage and harder forgery, it is more suitable to combine the gait patterns with other biometric information and develop a multimodal gait recognition that uses more than one biometric characteristic. This is particularly important in gait recognition, knowing that the gait patterns are suffering from low distinctiveness and permanence.

One of the new systems developed recently is the combination of gait patterns and soft biometrics. The soft biometric characteristics have been very recently introduced as a subset of biometric characteristics representing information like gender, weight, height, ethnicity, age, eye color, etc. Since a lot of people can have the same height or be at the same age, this information alone is normally not enough for identification and thus the soft biometric characteristics are often used as complementary information along with other biometrics (22). Moustakas et al. in (56) combine soft biometric features with gait features to improve the recognition rate by reducing the search space. In this work, they used the subject's height and the stride length as their soft biometric features. GEI and Radon transforms are used as geometric gait features. These features are combined using

a probabilistic approach (56). This method is a complicated probabilistic technique and furthermore it needs extra information for calculating the values of the soft biometrics. The datasets used in this work are captured with stereoscopic cameras so that the system can obtain the height of the subject.

Taking a look at the multimodal gait recognition systems that use more than one biometric trait, it seems that gait has most commonly been combined with face. Hossain and Chetty combine gait and face features in one feature vector and use Bayesian classifier for classifying the concatenated feature vectors (57).

The problem that I notice related to combining face and gait is that the resulting system can lose the main advantages of gait recognition. Face is not as remotely observable as gait and it can easily get covered. Therefore, combining these two might reduce the remote observability and obtrusiveness of the system. Furthermore, face recognition algorithms are expensive and adding them to the system might result in a lot of computation.

The multimodal gait recognition system that I introduce in this thesis combines the gait patterns in the form of GEI with the social patterns. The resulting context-based gait recognition system shows the following unique advantages:

- Using context can provide a rich metadata for increased biometric recognition rate.
- The system is using more than one single biometric trait. Thus it shows better performance, it has more coverage and it is harder to fool.
- Matching the behavioral patterns with the context of the gait video does not add a lot of computations to the system.

- Behavioral pattern matching does not need high quality data, does not need the subject to be close to the camera and does not need his/her cooperation. Therefore, it does not reduce the remote observability and unobtrusiveness of the overall multimodal gait recognition system while increasing recognition rates

The system is described in the next section and those advantages are proven through experiments in Section 4.

CHAPTER 3 : PROPOSED METHODOLOGY FOR CONTEXT-BASED GAIT RECOGNITION

3.1 Introduction

In this chapter I introduce a novel multimodal gait recognition system that takes advantage of the context in which subjects were observed. This results in a more accurate decision in comparison with applying the gait recognition algorithm alone.

Over the last few decades, gait as a behavioral biometric has gained a lot of attention in the area of biometric technologies mainly because it is unobtrusive and remotely observable (4). These two unique properties enable the biometric security system to recognize the suspicious subjects and activities from a farther distance and without any cooperation from the subjects. In addition to these two interesting properties, gait is also hard to imitate and hard to conceal making the gait forgery difficult as compared to other biometric characteristics (4) (9). However, the main problem with gait is the complexity and the wide variability of the gait patterns which limits the power of the gait recognition algorithms (9) (10). The person's clothing, the type of the surface he/she is walking on, the type of the shoes he/she is wearing, his/her mood and mental/physical health can all change the walking style of the person. Due to these complications, the two broad categories of algorithms for gait recognition, model-free and model-based approaches, each has its own limitations. The model-free approaches do not have any information about the shape of the human body and they work only based on the binary silhouettes. Therefore, they can only handle the controlled situations where the appearance of the subjects and the shape of their silhouettes do not go through drastic changes (7) (32). The model-based approaches, however, having some extra information about the human body

in a form of a body model, are more robust to unpredictable situations. However, the high flexibility of human body makes the search space of the model parameters extremely large. As a result, the model-based methods are complicated, expensive and need high quality videos to obtain satisfactory results (7) (31) (25). Consequently, in the presence of noise and uncertainty, the gait recognition algorithms usually fail to meet the desired recognition rates (7) (32) (31) (25). To solve this problem, I propose to use multimodal gait recognition and improve the performance of the gait recognition system by integrating context-based metadata. Inspired by previous studies of human behavioral habits that shows human daily behaviors can be predicted and modeled with a high precision (15) (16), the metadata I am incorporating into my system is the behavioral patterns and daily routines of the subjects. For this purpose, the proposed system uses a popular, cheap and fast model-free gait recognition algorithm and combines it with behavioral patterns of the subjects using information fusion techniques. To the best of my knowledge, no similar research has been conducted in the area of biometric or gait recognition up to date.

The proposed novel multimodal gait recognition system is mainly used for identification purposes. As described in Chapter 2, the main goal of the identification application is to establish the identity of an unknown subject presented to the system. In this process, the system compares the unknown subject's biometric data with all the existing biometric templates in the system database and find the closest match for the unknown subject, if any (22). There is a wide range of applications for the proposed context-based gait recognition system. The main areas of application are controlled environments where subjects have dictated patterns of behaviors and they have an obligation to be consistent

with their predefined daily schedules. However, since the proposed system has learning mechanisms, as will be discussed in Section 3.3.1.3, it can also be used in open area applications considering reasonable limitations on the number of subjects in the scene. By observing and learning patterns of behavior in the environment through time, the system will be able to achieve better and better results as more samples become available to the system. Finally, the system can be used for abnormal behavior detection by monitoring people's behavior and reporting the cases where subjects are acting different than their regular daily schedules.

The block diagram of the proposed system is shown in Figure 3-1. According to this figure, the information about the subjects is represented via two main databases: the gait database that stores the gait patterns of the subjects and the context database which is a novel database introduced specifically to accommodate metadata in my gait recognition system and stores the behavioral patterns of the subjects. Having this information, the system is presented with a gait sequence of an unknown subject called the probe and it establishes the identity of this unknown subject using the three following main modules:

- 1- The gait recognition system: this module is responsible for finding the similarity of the subjects to the probe according to the gait patterns. The output of this module is a rank list of the subjects with their corresponding similarity scores.
- 2- The context matcher: this original module is responsible for finding how well the context of the probe matches with the behavioral patterns of the subjects. This module outputs a rank list of the subjects with their corresponding matching scores.

3- The information fusion: this module combines the outputs of the gait recognition system and context matcher into one final list and presents it to the user for final identification.

According to the above definitions and Figure 3-1, in this system gait patterns and behavioral patterns are treated as two biometric characteristics. An independent matcher is designed for each of these two (the gait recognition system for gait patterns and the context matcher for the behavioral patterns) to produce the corresponding matching scores that are later combined using the information fusion module at matching level.

The main contributions of this work to gait recognition methodology include:

- Introducing, for the first time, context matching module in gait recognition systems
- Defining behavioral patterns in the context of gait recognition and developing original methods for modeling and learning the behavioral patterns
- Creating context database and method for context matching
- Developing, implementing and testing overall multimodal system using biometric fusion of gait recognition method and context-based behavioral pattern matching

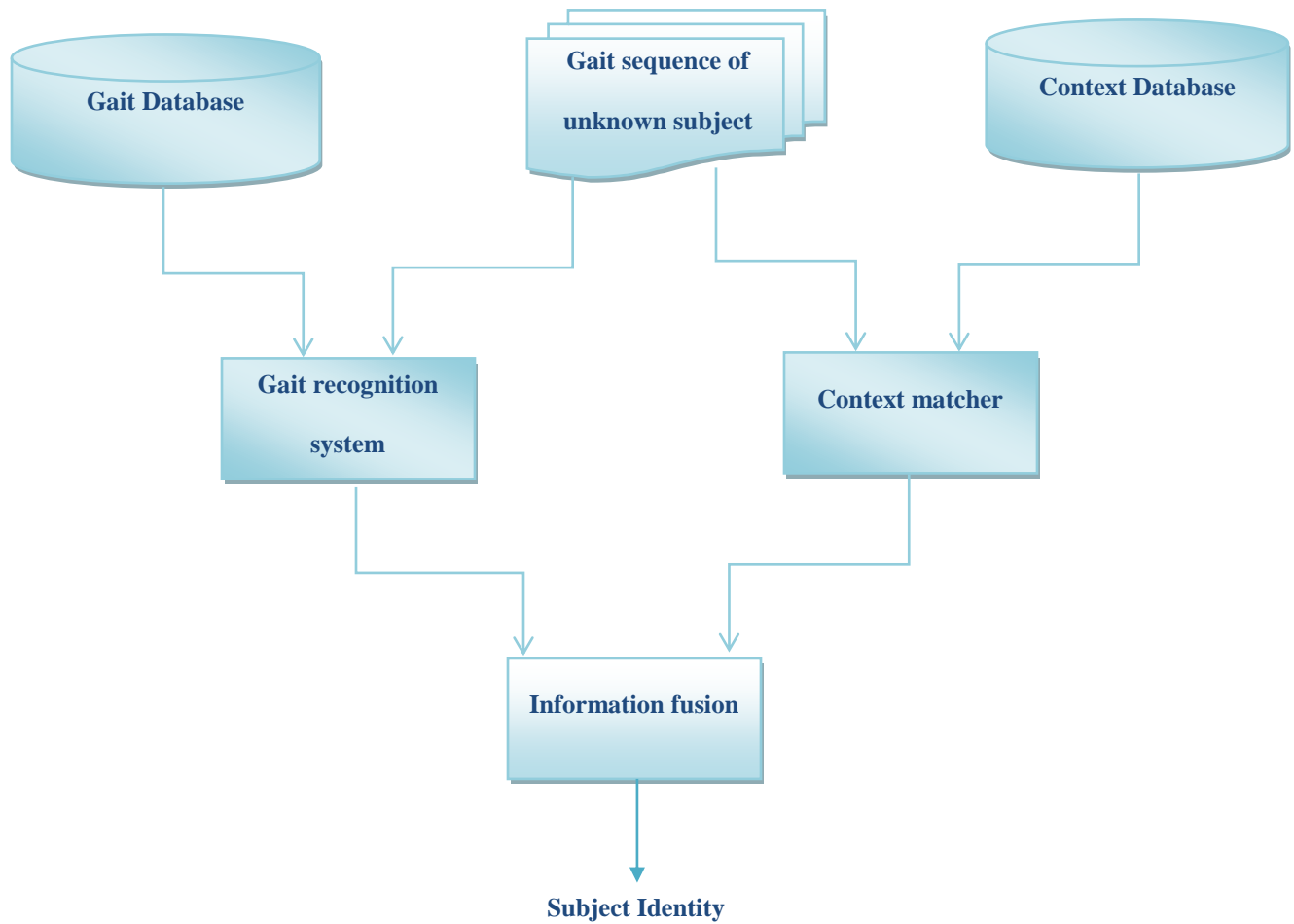


Figure 3-1: The general framework of the proposed context-based gait recognition system

The rest of this chapter is dedicated to explaining the technical details of the building blocks of the proposed context-based gait recognition system. Section 3.2 describes the steps of the gait recognition system and the algorithms and the data structures used in this module. Section 3.3 introduces the context matcher module and describes how the behavioral patterns are presented, modeled and assigned to the subjects. Finally, section 3.4 explains how the outputs of the gait recognition system and context matcher are combined to make the final decision.

3.2 The gait recognition system

As illustrated in Figure 3-1 and described in Section 3.1, the gait recognition system is one of the main building blocks of the proposed context-based gait recognition system. It is responsible for extracting the gait patterns of the probe, matching them with the gait patterns of the subjects stored in the gait database, and providing the user with a rank list of the most similar subjects and their corresponding matching scores. The basic gait recognition system used in this thesis has the general framework of a typical gait recognition system introduced in Section 2.2 and follows the work of Han and Bhanu in (35). The flowchart of the gait recognition system is presented in Figure 3-2. As can be seen, the gait recognition system has the four following modules (35):

- 1- Preprocessing: this module normalizes and prepares the gait silhouette sequence for further analysis.
- 2- Gait cycle detection: this module calculates the gait cycles and partitions the gait sequence into the gait cycles.
- 3- Feature extraction: this module is responsible for extracting the distinctive gait features from the extracted gait cycles.
- 4- Matching: this module matches the extracted features with the gait patterns stored in the gait database and establishes the identity of the probe.

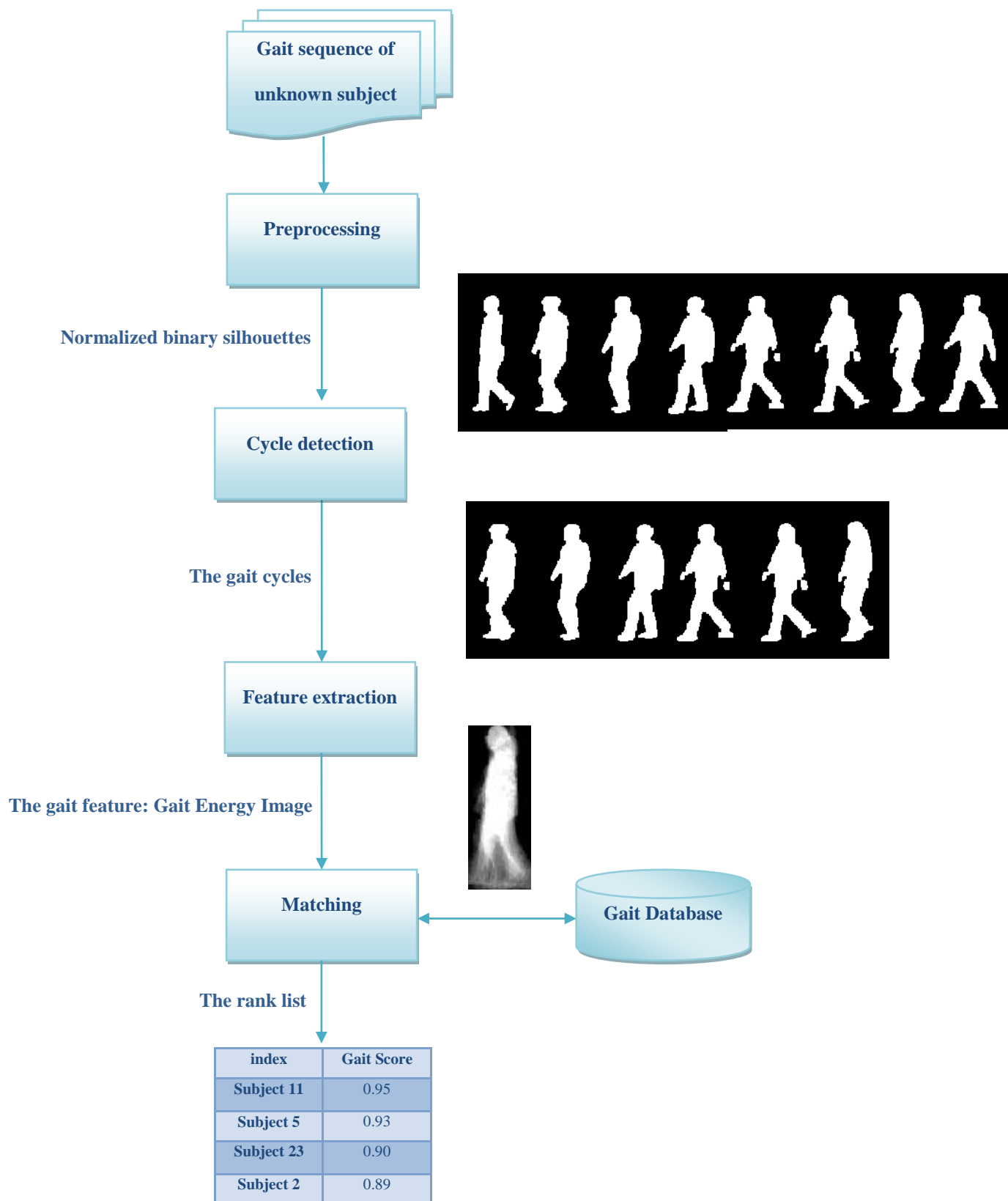


Figure 3-2: The flowchart of the gait recognition system

The four following subsections describe the details of the algorithms used for each of the main four steps: preprocessing, gait cycle detection, feature extraction and matching.

3.2.1 Preprocessing

As described in Chapter 2, the first step of each gait recognition system is to separate the targeted subject from the background and obtain the corresponding binary silhouette in each and every frame. The datasets used in this thesis for evaluating the proposed system already had the binary silhouettes available as a part of the dataset. Having the binary silhouettes, similar to the method applied by Han and Bhanu (35), on each frame I perform the following preprocessing steps to obtain the final normalized silhouettes:

- 1- Noise removal: the extracted binary silhouettes in real scenarios are usually not perfect and problems like shadows, lighting changes, low quality of the videos, etc. can result in noise and other extra objects in the binary silhouette images. All these extra objects should be removed before any further processing. The most common approach used for this purpose is based on finding the largest connected component of the binary silhouette image. A connected component of a binary image is defined as a group of connected pixels with the same color. Since a binary image has only black and white colors, each connected component will be a connected black region on a white background or vice versa. The problem with using the largest connected component as human silhouette is that sometimes the similarity of the background's color to the person's clothes can result in some parts of the person appearing disconnected (Figure 3-3). Thus, finding the largest connected component of the binary image can result in losing some of the body parts; To solve these problems, the dilation morphing operation is first applied to

the binary image to connect the different parts of the subject, if possible. Afterwards, all the connected components of the binary image are extracted and the largest connected component is selected as the subject's silhouette and all the other existing connected components are then removed from the image. Some examples of possible noisy silhouettes are shown in Figure 3-3.



Figure 3-3: Examples of noisy data

- 2- Scale normalization and centralization: since the silhouette sequences presented to the system can have different sizes (resolutions), it is essential to make the algorithm robust to scale changes by making the silhouettes have the same size. For this purpose, in each and every frame, the binary silhouette image is horizontally resized so that its height is always 128 pixels (a predefined value). Afterwards, the centroid of the resulting binary silhouette is calculated and a bounding box is defined around the silhouette centered at the calculated centroid (27). This box is 128*50 pixels.

After the above preprocessing steps, each frame is cropped to only contain the obtained bounding box. As a result, all the frames will have the exact same size, they

contain only the centralized silhouettes and the majority of background pixels have been removed. An example of this preprocessing is shown in Figure 3-4.

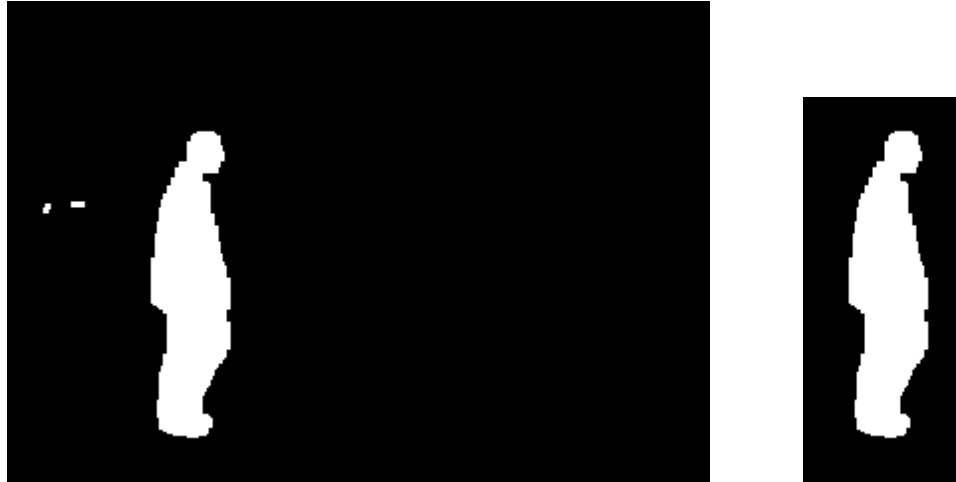


Figure 3-4: Preprocessing results, left: original binary silhouette, right: normalized binary silhouette after preprocessing

3.2.2 *Gait cycle detection*

Since the gait features used by my system are calculated for each gait cycle separately, the next step of the system is to partition the gait sequence into gait cycles. To illustrate the algorithm, one gait step is shown in Figure 3-5. Each gait cycle consists of two such steps and since each step starts and ends with the right leg and left leg being next to each other, it is possible to detect the gait cycles by finding the frames in which this happens.



Figure 3-5: One gait step

As can be seen in Figure 3-5, the number of foreground pixels reaches a maximum when the right leg and left legs are farthest apart and reaches a minimum when the two legs are together. Therefore, by counting the number of foreground pixels and detecting the two subsequent minima it is possible to find the points where the two legs are together which correspond to the beginning and ending points of the gait steps (27).

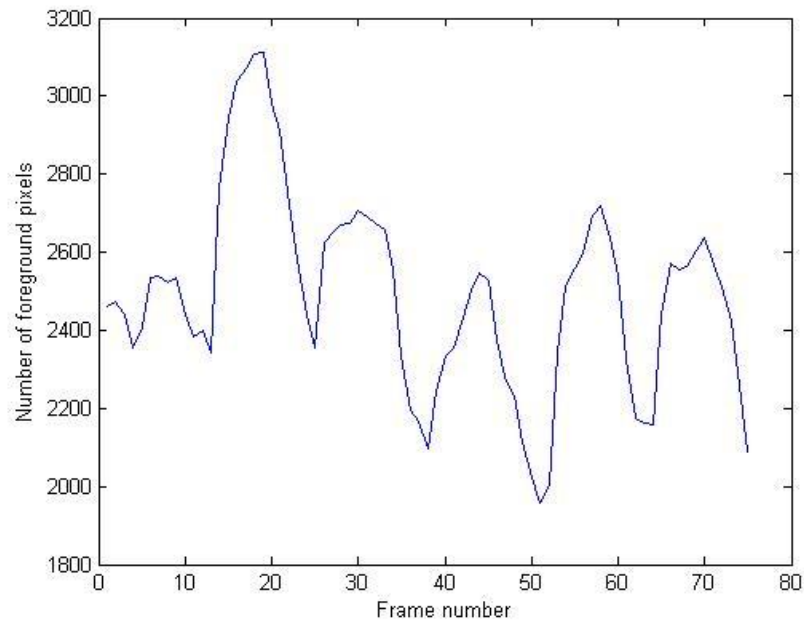


Figure 3-6: The number of foreground pixels in the lower half of the silhouette in each frame of the gait sequence

Based on the above property of gait cycles, using the method introduced by Sarkar et al. in (27), in each frame I count the number of pixels in the lower half of the silhouettes (leg region). Using the resulting values, I obtain a curve. An example of such a curve is shown in Figure 3-6. Since the minima of this curve correspond to the points where the two legs are together or the beginning of a step, all the frames between two subsequent minima, belong to one gait step. Therefore, for finding a gait cycle which is consisted of

two subsequent steps, I need to find all the frames between two minima skipping every other minimum. Using this property, I find the local minima of the foreground pixels curve to find the beginning and ending frames of the cycles. I also calculate the gait cycle as the average of distances between minima, skipping every other minimum (27).

3.2.3 Feature extraction

After obtaining the binary silhouettes and detecting the cycles, the next step is to extract distinctive gait features. The gait feature in general is a piece of information extracted from video sequences representing a distinctive property of walking patterns that can be used for individual identification. There are a wide variety of gait feature extraction algorithms available in the literature but as mentioned in Chapter 2 the majority of gait recognition techniques fall into the model-free category. The model-free approaches do not have a human body model and extract their features only based on the binary silhouette. As a result, they are mostly fast, computationally cheap and easy to understand and implement. Among all the model-free approaches, as described in Chapter 2, the GEI introduced in (35) is one of the most popular feature extraction techniques. In the past few years a lot of research has been dedicated to improving the performance of GEI as the gait feature. This feature extraction method is fast, simple, easy to implement and at the same time it shows acceptable results. Due to all these properties, the GEI has been used in this thesis.

Formally, GEI is the average of all the binary silhouettes of a gait cycle and can be obtained using equation (3-1) (35).

$$GEI = \frac{1}{c} \sum_{i=1}^c B_i \quad (3-1)$$

C in equation (3-1) is the number of frames in a gait cycle and B_i is the i^{th} normalized binary silhouette. Since all the normalized binary silhouettes are of size 128×50 , the resulting GEI is also a 128×50 image. The GEI shows distinctive properties of walking patterns and can be used for individual identification. Therefore, the whole Gait Energy Image is used as the gait feature in the proposed system.

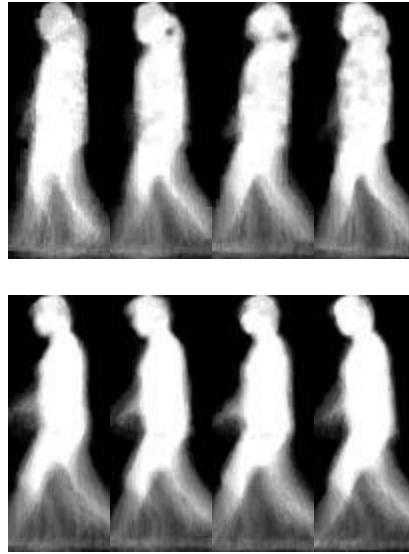


Figure 3-7: Some examples of Gait Energy Images for two persons and four video sequences per person

After obtaining the GEI, it is essential to make the resulting image invariant to the moving direction. In the side-view sequences, two moving directions are possible: walking to the left and walking to the right. To make the Gait Energy Image invariant to the moving direction, I obtain the moving direction based on the movement of the horizontal coordinate of the silhouette's centroid. If the horizontal coordinate of the centroid is decreasing the person is moving to the left, otherwise he/she is moving to the right. After calculating the Gait Energy Image using equation (3-1), if the person is moving to the right, the Gait Energy Image is reflected. Consequently, for all the

obtained Gait Energy Images the moving direction is to the left and the extracted features are invariant to the moving direction. Some examples of the obtained Gait Energy Images are shown in Figure 3-7.

3.2.4 *Matching*

The last step of my gait recognition system is to identify unknown subjects by comparing their gait patterns with the templates stored in the gait database. For the system to be able to identify people, it should first be trained using the gait patterns of the subjects that will be using the system. This process is called the training phase. Once these gait patterns are stored in the gait database, the system should be able to identify the unknown subjects with matching the gait patterns in the identification phase. Since the gait feature vector of my system is based on the work of Han and Bhanu in (35), the training and identification phases of my system are also standard and similar to their method.

3.2.4.1 Training phase

The training phase is an offline phase in which the system is trained to learn the gait patterns of the subjects that will be using the system. In the training phase of my gait recognition system, for each subject in the training data set, the silhouettes are first normalized using the steps described in Section 3.2.1, then the gait cycles are detected based on the algorithm discussed in Section 3.2.2 and then for each gait cycle a GEI is calculated using equation (3-1). Since according to equation (3-1) for each gait cycle it is possible to obtain one GEI, the number of GEIs for each subject depends on the number of samples available for that subject and the number of gait cycles in each sample. It is also worth mentioning that since not all the recorded walking sequences exactly start from the beginning of a step and similarly they do not always end at the completion of a

step, to make sure that all the cycles are complete, I always discard the first and last gait cycles of each gait sequence. The resulting GEIs are then transformed to vectors and used as the gait feature vectors of the system. The resulting feature vectors are called f_s^i ($s=1, \dots, \text{subject_no}$ and $i=1, \dots, \text{feature_no}_s$). Where subject_no is the number of subjects in the database and feature_no_s is the number of feature vectors (GEIs) available for subject “s”. Afterwards for each subject “s”, the gait template for subject “s” is obtained as the average feature vector F_s using equation (3-2) (35).

$$F_s = \frac{1}{\text{feature_no}_s} \sum_{i=1}^{\text{feature_no}_s} f_s^i \quad (3-2)$$

All the resulting templates F_s ($s=1, \dots, \text{subject_no}$) are then stored in the system gait database.

3.2.4.2 Identification phase

After the system has been trained, it can finally be used for identifying people. Identification phase is an online phase in which a gait sequence of an unknown subject called the probe is presented to the system and the system establishes the identity of that subject by finding the similarity of the subject’s biometric data to the templates in the system database. In the identification phase, I first normalize the silhouettes of the probe and then I find the cycles using the approach provided in cycle detection section. Afterwards, for each cycle I calculate the GEI using equation (3-1). Then, I transform the resulting image to a vector called $F_{unknown}^i$ where $i=1, \dots, C$ (C is the number of cycles).

Afterwards, I obtain the Euclidean distance of the resulting $F_{unknown}^i$ from all the templates stored in the gait database using the equation (3-3) (35).

$$d_s = \frac{1}{C} \sum_{i=1}^C \text{distance}(F_{unknown}^i, F_s) \quad s = 1, \dots, \text{subject_no} \quad (3-3)$$

In equation (3-3), distance is a function that calculates the Euclidean distance between its two parameters and d_s is the distance of the probe from the subject “s” in the database. Having all the distances, I provide the user with a rank list of top N subjects with minimum distance to the probe according to equation (3-3). Each candidate in the rank list is assigned a matching score which is the similarity of that candidate’s gait patterns to those of the probe. This score ($score_s$) is calculated using equation (3-4).

$$score_s = 1 - \frac{d_s}{d_{max}} \quad s = 1, \dots, subject_no \quad (3-4)$$

In equation (3-4), d_s is the distance of the subject “s” from the probe calculated according to equation (3-3) and d_{max} is the maximum of the distance values and is obtained using the following formula.

$$d_{max} = \max_{s=1}^{subject_no}(d_s) \quad (3-5)$$

3.3 The Context Matcher

As mentioned at the beginning of this chapter, the proposed novel context-based gait recognition system is designed as a multimodal gait recognition system that combines the gait patterns of the subjects with their behavioral patterns to make the individual identification more accurate. Since the data sample for gait is a video not a single image, gait is one of the very few biometrics that has some extra information in its sample other than just the gait patterns. This extra information can include the place where the subject has been observed, the time of the day and the condition of the subject in the video like clothing, carrying condition, etc. After reviewing multiple studies done on modeling and predicting the behavioral patterns of human beings, I realized that, according to these studies, people are very predictable in their daily routines (15) (16). This predictability is

also evident in their online browsing and shopping patterns (42) (41). The predictability of human behaviors and the possibility of extracting context information from the gait video motivated me to combine the behavioral patterns with the gait patterns in my context-based gait recognition system. The resulting context-based gait recognition achieves more accuracy at a negligibly small extra computational cost.

One of the essential parts of my proposed context-based gait recognition system is the context matcher. In the same way that the gait recognition system is responsible for matching the gait patterns, the context matcher is responsible for matching the context of the unknown input video with the behavioral patterns of the subjects. To the best of my knowledge, there are no similar systems developed for gait recognition with context matching and behavioral pattern analysis. Having no similar framework available, I designed my context matcher module with a framework similar to a biometric system that uses the behavioral patterns as its biometric characteristic. The main reason behind this decision is that, in this work, the behavioral patterns of the subjects are used for their identification in the same way as biometric characteristics, but, since they are not as unique, they have been augmented with the gait patterns. The framework of my context matcher is shown in Figure 3-8. As can be seen in the figure, one of the main modules of the context matcher is the behavioral patterns module which is responsible for modeling the behavioral patterns of the subjects and storing the resulting patterns in the context database. Having the context database, the context extraction module extracts the context of the unknown video presented to the system. Afterwards, the matching module matches how compatible the extracted context is with the behavioral patterns of the subjects stored in the context database and outputs a rank list of the subjects with their

corresponding matching scores. The following subsections are dedicated to describing these main functionalities of the context matcher.

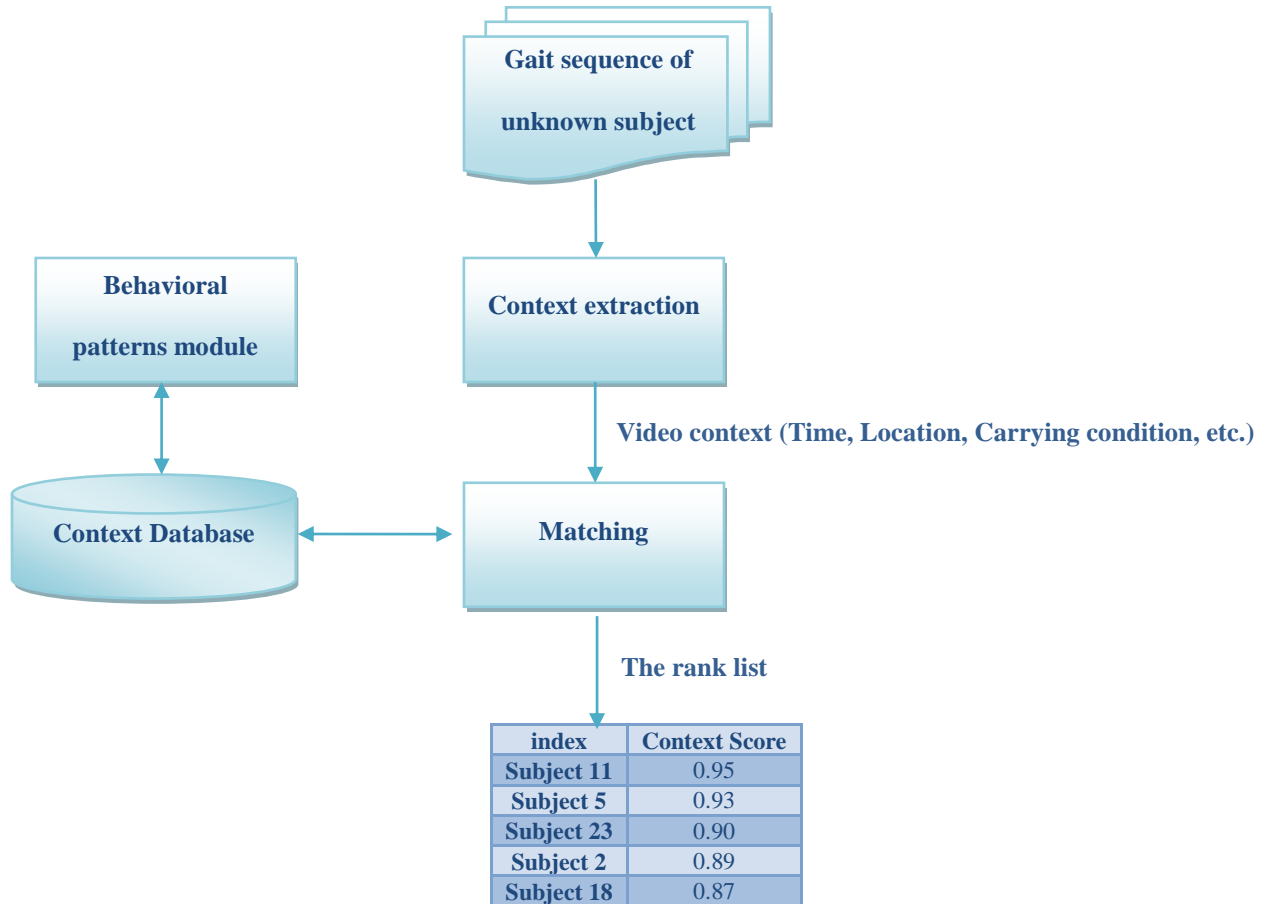


Figure 3-8: The flow chart of the context matcher

3.3.1 Behavioral patterns module

The main purpose of the behavioral patterns module is to provide a framework for defining and storing the behavioral patterns of the subjects. For this purpose, first a formal definition of a behavioral pattern should be established. Afterwards, a methodology for modeling the behavioral patterns of the subjects should be defined.

Finally, the behavioral patterns should be stored in the context database to be used later by the context matcher. These steps are described in the following.

3.3.1.1 Behavioral pattern definition

The first step of modeling the behavioral patterns is to define what information do I want to capture by these patterns and which parameters do I want to use for representing this information. As mentioned previously, as far as my research shows, there is no similar work available in the literature, hence, the method explained here for modeling the behavioral patterns is developed from the scratch. To make sure the method is computationally efficient and feasible to implement, mainly three issues should be considered in defining the behavioral patterns. First, since the behavioral patterns are later to be matched with the unknown gait sequence presented to the context matcher, the parameters that are used for modeling the behavioral patterns should be efficiently extractable from the gait sequence. Second, the parameters should be chosen in a way that they provide additional authenticating information to the system. Third, the behavioral patterns should be defined so that they can be stored, extracted and matched in an efficient manner both time-wise and space-wise. Based on the studies done by Gonzalez et al. and Song et al. (15) (16), people are predictable in terms of the places they visit and the times of the day they visit them. These two parameters (time and place) are also easily extractable from the gait sequence. Another piece of evidence that is also available in the gait sequence is the carrying condition of the subject. Although not much research has been found on modeling the subject's carrying condition and extracting it from the gait sequence, I decided to incorporate it in my system and evaluate its influence. Considering all the three mentioned factors, I decided to define the behavioral

patterns of a subject in the following way: “at what time of the day, in which locations and in what conditions subject can usually be observed”. To store the behavioral patterns of the subjects in an effective way and also to have an efficient matching mechanism, I decided to make all the parameters discrete. Consequently, for representing the behavioral patterns, three following parameters have been used:

1- Location:

This parameter captures the possible locations that each subject can be observed by the system. For the sake of simplicity and generality, in this thesis I considered two possible locations: inside and outside. However, for each application based on the targeted environment, it is possible to have a more extensive list of acceptable locations. For example considering university as the environment, the list can include classroom, office, hallway, food court area, computer labs, etc.

2- Time:

This parameter models possible times of the day that each subject can be observed by the system. Four possible values have been considered for this parameter: morning, afternoon, evening and night.

3- Carrying condition:

This parameter captures what objects does subject usually carry with her/himself. The possible values for this parameter include: coffee, backpack, suitcase, notebook and nothing.

Obviously, for some other applications and systems, there could be different or additional parameters defined. They are very easy to incorporate in the present system.

Based on the above descriptions, for defining the behavioral patterns for each subject it is sufficient to decide which values for each of these three parameters are acceptable for that subject. However, another issue is to decide whether to consider the parameters independently or to take into account their relationships. The three parameters I used here are not completely independent in real scenarios. As an instance, the places a subject visits during the day are usually functions of time. For example, the subject is usually at work during the morning and at home during the night. Therefore, intuitively the relationship between the parameters should be taken into account. However, modeling such relationships require extensive knowledge about the subjects and might be time consuming. Ignoring the relationship between the parameters can make modeling the behavioral pattern more straightforward. In this thesis, to cover all scenarios, I decided to consider both cases. For the case that the parameters are considered independently, each behavioral pattern only contains information about one of the parameters. An example of a behavioral pattern in this case would be: “subject A can be observed by the system in the morning”. In this approach, the behavioral patterns for each subject would be three independent lists representing the possible times, locations and carrying conditions for that subject. For the case of joint parameters, however, each behavioral pattern includes information about all the parameters. An example of a behavioral pattern in this case would be: “subject A can be observed by the system at the morning, inside, and while carrying a cup of coffee”. In this approach, the behavioral patterns for each subject would be one single list with each item representing a possible combination of the three parameters. Joint behavioral patterns are clearly more detailed, distinctive and harder to define.

3.3.1.2 Behavioral patterns modeling

Having a formal definition for a behavioral pattern, the main step of creating a context database is to model the behavioral patterns of the subjects. Three approaches for modeling the behavioral patterns have been proposed in this thesis and are described in the following.

1- Random models

In this approach, as the simplest and most generic solution, a random model is used which assumes that there is no predefined distribution for the context parameter values and any parameter value for any subject is equally probable. Therefore, modeling the behavioral patterns can be done by selecting the possible parameter values for each subject without having any limitations.

Sample Questions	Sample answers
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> What time do you come to work?	<input type="checkbox"/> 8 AM
<input type="checkbox"/> What time do you leave?	<input type="checkbox"/> 5 PM
<input type="checkbox"/> When do you normally take your lunch break?	<input type="checkbox"/> 12 PM
<input type="checkbox"/> In which building do you work?	<input type="checkbox"/> ICT
<input type="checkbox"/> Do you work indoor/outdoor?	<input type="checkbox"/> Indoor
<input type="checkbox"/> Do you normally bring your laptop to work?	<input type="checkbox"/> Yes

Figure 3-9: A sample Questionnaire

This approach can easily be implemented in real applications by asking the users to fill questionnaires during registration. Since the parameters of behavioral patterns are well defined and they only take discrete values, the questionnaires can simply be designed accordingly. Figure 3-9 shows one sample questionnaire and one sample answer. The questionnaires will be maintained by the system administrator and can easily be mapped to behavioral patterns that will be stored in the context database.

2- Behavioral profiles

In this approach, instead of modeling the behavioral patterns of each subject individually which can be quite time consuming, I define some groups of users based on the similarities in their behavior. Such similarities can be dictated by similar profession of the subject, similar environment in which behavior is observed, similar place where behavior is observed, etc. Then I place subjects with similar behavioral patterns in the same group. I model the behavioral patterns of the whole group all together and I call the resulting behavioral patterns *the behavioral profiles*. For example, assuming university as the environment, I can define for instance three different profiles: *student*, *professor* and *technical staff*. For each of these profiles, I can define the more likely times when a person can be at a specific location with specific carrying conditions. For example, a professor most probably comes to university in the morning, walks the hallways in between lectures, and carries a coffee mug, while a student might stay longer at school, have a laptop or a bag, and go to specific class according to his/her schedule.

This approach can also be easily implemented in real scenarios. The profiles of each environment correspond to available job titles and/or working groups in that

environment. These profiles can be created when the system is being installed. Afterwards, when the users are registering to the system, they can be assigned to the corresponding profile based on their job title. Clearly, in this approach the users are less involved in defining the behavioral patterns. Furthermore, the behavioral patterns for each profile are defined only once and there is no need to model the behavioral patterns on a user by user basis. Therefore, this approach is very timely efficient.

3- Gaussian models

In the random modeling approach, all the different parameter values have the same chance of occurrence and none of the parameter values has a priority over the others. As a result, the distribution of the resulting parameter values would be a multi modal distribution with some random peaks. However, in some scenarios a unimodal distribution might be a better fit. As the studies in (15) (16) show, people normally have some hotspots they frequently visit. Furthermore, the places a person visits over certain limited time usually tend to be in close proximity of each other to reduce commuting times. For handling these scenarios and also as an attempt to reduce the overlap in subjects' behavioral patterns, I researched into using Gaussian distribution for the distribution of each parameter. I discovered that Gaussian distributions have been commonly used in modeling and simulating real world processes and more precisely human behaviors (58) (59) (60). Jiao et al. in (58) designed a framework for modeling habitual behaviors of human beings and they use normal distributions for simulating the intervals of sequence of actions and the probability of different actions occurring. The results of their simulations are compared against real data and shown to be consistent.

Smith and Brokawin in (60) introduce an agent based system for simulating human movements during emergency evacuations and they use normal distributions for modeling the movement speed of the agents. Wei and Kalay in (59) propose an agent-based approach for modeling human behaviors in built environments. The behavioural models are created according to theoretical and applied related studies and real world data. In this approach, Gaussian distributions are used for modeling the duration of human behaviors (59). Based on these observations, I propose to use Gaussian modeling as an alternative approach to the random modeling approach. In this approach the distribution of each parameter for each subject is assumed to be a Gaussian distribution. Hence, modeling the behavioral patterns of each subject would be as simple as determining the mean and variance of the Gaussian distribution for that subject. As the experiments of Chapter 5 show, using Gaussian modeling in creating virtual context databases shows very good results compared to the other two modeling methods since this method makes the behavioral patterns of individuals more distinctive.

All the three proposed approaches can handle both independent context parameters and joint context parameters. For the case of independent parameters, the acceptable values for each parameter should be assigned independently. However, if the relationship between the parameters is also taken into account, all the parameters should be involved in defining each behavioral pattern.

The above three methods are generic methods that can be used for creating the behavioural models for any set of individuals. For each specific application, the most suitable method that is consistent with the behavioral patterns of that application should be applied. The number of behavioural patterns created for each individual and similarly

the number of defined profiles should be selected based on the available knowledge about the subjects and environment.

3.3.1.3 Behavioral patterns learning

Using the methods described in section 3.3.1.2, it is possible to model the behavioral patterns having some information about the subjects (completed questionnaires, job titles, etc.) and/or the environment (job titles, job profiles, etc.). However, there can be cases that this information is not available to the system; for example the users do not officially register to the system or the environment is not fully known. For such cases, the system should still be able to learn the behavioral patterns of the subjects using their gait video samples. This is very important advantage of the system because it makes the system fully unobtrusive and capable of recognizing completely unknown subjects without their cooperation. Two learning approaches have been proposed in this thesis to enable the system to learn the behavioral patterns of the subjects from the context of gait samples. These two approaches are described in the following.

1- Random models learning

This approach is similar to random modeling of the behavioral patterns. Thus, the behavioral patterns of each subject do not follow any particular distribution. However, since there is no information available about the behavioral habits of the subjects, this information should be extracted from the context of gait samples. The proposed learning procedure for this case is to extract the value of the context parameters for all the training samples of each subject and tag the samples with the obtained context parameters. Afterwards, assign a parameter value to a subject if the subject has a sample tagged with

that parameter value. For example, if subject “A” has three gait samples with the following context tags:

{“morning”, “inside”, “notebook”}, {“morning”, “inside”, “backpack”}, {“evening”, “inside”, “notebook”},

Then, the behavioral patterns for this subject would be as follows:

$L_{Time}^A = \{\text{“morning”, “evening”}\}$, $L_{Location}^A = \{\text{“inside”}\}$, $L_{Carrying\ condition}^A = \{\text{“notebook”, “backpack”}\}$,

Where L_{Time}^A , $L_{Location}^A$ and $L_{Carrying\ condition}^A$ are the lists of acceptable values for parameters time, location and carrying condition, correspondingly. It is possible to make the resulting behavioral patterns dynamic by updating them each time the subject uses the system.

2- Behavioral profiles learning

The purpose of this approach is to learn the behavioral profiles for each environment based on the available gait samples for the subjects. The proposed learning approach for this purpose is consisted of the following main steps:

1- Learn the behavioral patterns of each subject using the approach described for Random models learning. Since all the parameter values are discrete, the result of this step determines which parameter values are acceptable for each subject. For example, I can have the following lists for subject “A”:

$L_{Time}^A = \{\text{“morning”, “evening”}\}$, $L_{Location}^A = \{\text{“inside”}\}$, $L_{Carrying\ condition}^A = \{\text{“notebook”, “backpack”}\}$

Where L_{Time}^A , $L_{Location}^A$ and $L_{Carrying\ condition}^A$ are the lists of acceptable values for parameters time, location and carrying condition correspondingly. To make all the parameter values numeric, I assign a number to each parameter value starting from one. After this step, the list for each parameter and each subject will be a list of numbers. The result of this assignment for the above example is shown below.

$$L_{Time}^A = \{1, 3\}, L_{Location}^A = \{1\}, L_{Carrying\ condition}^A = \{2, 4\}$$

2- Cluster the resulting behavioral patterns using a clustering algorithm. K-means clustering as one of the most commonly applied clustering algorithm has been used for this purpose (61). For the clustering algorithm to work, the behavioral patterns for each subject should be represented by a vector. I have an independent list for each parameter (above example) and I need to transform all the lists into one single vector that I call the *context vector*. The main idea of the context vector for each subject is to represent which parameter values are acceptable for that subject. The simplest way to do this is to have the context vector as a sequence of zeros and ones. The ones indicate acceptable parameter values and the zeros non acceptable parameter values. To have the acceptable values for all the parameters, each parameter should have its own section in the context vector. The number of places (bits) assigned to each parameter is equal to the maximum value for that parameter. If value “i” is acceptable for subject “A” and parameter “p” or in other words if $i \in L_p^A$ then the i^{th} bit of the section for the parameter “p” in the context vector will be set to one, otherwise to zero. For example, the corresponding context vector for the above example is 10101001010. The first four bits are for time, the next two bits for location and the last five bits for carrying condition. This representation is

generic and can be used for any number of parameters and any number of parameter values as long as the parameter values are discrete. Once the context vectors are obtained, the k-means algorithm can be applied to these vectors.

3- Extract the behavioral profiles from the output of the clustering algorithm. The corresponding profile for each subject is the cluster index for that subject. The behavioral patterns for each profile can be obtained from the resulting cluster centers.

3.3.1.4 Creating the context database

One of the essential parts of the context matcher is the context database which stores the behavioral patterns of the subjects. For creating the context database, the only required step is to model or learn the behavioral patterns of each subject which has been described in previous section. Consequently, it is possible to build the context database for each specific application having the knowledge about the environment and the users. The method used for assigning the patterns should be selected based on the conditions and limitations of the environment and its users.

Finally, after defining the behavioral patterns and creating the context database, to make the system more realistic, I attempt to also cover the cases where subjects do not necessarily always follow their behavioral patterns. For handling these cases I introduce and define *confidence values*: one confidence value for each subject and one confidence value for the whole context database. The confidence value for each subject represents how well that subject acts according to his/her behavioral patterns. I introduce the confidence value in the range $[0,1]$, with 1 representing highest reliability of the behavioral patterns. Very similar in concept but just in greater scale, the confidence value of the whole context database is a number in range $[0,1]$ that shows generally how

predictable the subjects of the targeted environment are in their observed behavior. The bigger the number, the more predictable people are. These confidence values should be determined and stored as a part of the database once the context database is defined according to the knowledge about the predictability of the subjects' behavior and the whole environment.

3.3.2 Context extraction

Having the context database ready, the next step is to extract the context of the probe video presented to the system to be matched later with the behavioral patterns stored in the context database. Since three parameters (time, location and carrying condition) have been used for modeling the behavioral patterns, the context of the probe video would be the time, location and carrying condition of that video. Normally, in majority of gait recognition applications, the location of the surveillance camera that captured the video is already known and the camera also saves the recording time of the video. Therefore, the time and location of the probe video can be obtained with no further processing. The carrying condition can also be obtained by visual inspection of the video. For making the system fully automatic, however, it is possible to use object detection algorithms to detect the object the person is carrying, if any. Since this process requires advanced image processing techniques, it is beyond the scope of this thesis. Therefore, in this thesis, it is assumed that the context of the probe video (time, location and carrying condition) is known.

3.3.3 Matching the context

The final step of my context matcher is to match the context of the probe with the behavioral patterns of the subjects stored in the context database and generate a matching

score for each subject accordingly. Based on whether the parameters are considered independently or all together, I developed two different approaches for matching the context of the probe with the behavioral patterns of the subjects. These two proposed methods are described in the following.

1- Independent parameters:

When the context parameters are considered independently, the behavioral patterns for each subject consist of three independent lists for each of the three context parameters. The list for each parameter demonstrates which parameter values are acceptable for that subject and that parameter. For example, the behavioral patterns for subject “A” can be defined as follows:

$L_{Time}^A = \{\text{“morning”}, \text{“evening”}\}$, $L_{Location}^A = \{\text{“inside”}\}$, $L_{Carrying\ condition}^A = \{\text{“notebook”}, \text{“backpack”}\}$,

where L_{Time}^A , $L_{Location}^A$ and $L_{Carrying\ condition}^A$ are the lists of acceptable values for parameters time, location and carrying condition correspondingly.

In this case, for each subject and each context parameter, I assign a score of one to the subject and that parameter if there is a match between the context of the probe and the candidate’s context for that parameter; otherwise the score is set to zero. A match corresponds to the case that the value of the parameter in the probe video is listed as one of the possible parameter values for the subject in the context database. After this calculation, for each subject I will have one separate score for each of the three parameters. However, I need to have one single context score for each subject. Therefore, I need to combine the three scores into one single score. This problem is very similar to match score level information fusion and as discussed in Section 2.5.1, the most popular

approaches for this purpose are sum, product, min, max and weighted sum. In this work, since the parameter score value can only be zero or one, multiplying such scores will result in zero if one of the scores is zero. Hence, a lot of subjects will have the context score of zero which is not acceptable. Furthermore, finding the minimum or maximum of the scores is also inappropriate here because it ignores the parameters with non-min or non-max values. Therefore, finding the sum of the three scores is the most suitable approach for this application. Since all the parameters have the same importance (weight), there is no need to use the weighted combination. Consequently, after obtaining the score for all the parameters, I add all the parameter scores together to obtain a single context score for each subject. In the following a detailed example of this score assignment process is presented.

If the context of the probe video is Time="morning", Location="outside", Carrying condition="backpack", then the context score for subject "A", using the behavioral patterns L_{Time}^A , $L_{Location}^A$ and $L_{Carrying\ condition}^A$ defined at the beginning of this section, is calculated as follows:

- a. Time parameter: since the value "morning" is listed as one of the acceptable values for subject "A" in L_{Time}^A , there is a match for the time parameter. Therefore the time parameter score ($score_{Time}^A$) is set to 1.
- b. Location parameter: Because the value "outside" is not listed in $L_{Location}^A$, there is no match for the location parameter and the location parameter score ($score_{Location}^A$) is set to 0.

- c. Carrying condition parameter: since the value “backpack” is an acceptable parameter value for subject “A” according to $L_{Carrying\ condition}^A$, there is a match for carrying condition parameter and the carrying condition parameter score ($score_{Carrying\ condition}^A$) is set to 1.
- d. The final context score: the context score of subject “A” ($Context_score_A$) is obtained using the following formula:

$$Context_score_A = score_{Time}^A + score_{Location}^A + score_{Carrying\ condition}^A = 1 + 0 + 1 = 2 \quad (3-6)$$

2- Joint parameters:

As illustrated in Section 3.3.1.1, when the context parameters are considered all together, the behavioral patterns for each subject are demonstrated as one single list of acceptable combination values involving all the three parameters. For example, the behavioral patterns for subject “A” can be defined as follows:

$$L^A = \{ (“morning”, ”inside”, ”coffee”), (“afternoon”, ”outside”, ”backpack”) \}$$

For calculating the context score in this case, since all the parameters are considered together, the context score of the subject is also assigned in one single step taking into account all the parameters at the same time. For this purpose, each subject obtain a context score of value one only if there is a behavioral pattern for the subject in the context database that exactly matches with the context of the probe considering all the parameters. For example the subject “A” in this case will obtain a context score 1 if and only if the context of the probe is (Time=“morning”, Location=“inside”, Carrying

condition="coffee") or (Time="afternoon", Location="outside", Carrying condition="backpack"). The context score will be zero for all the other cases.

Finally, the calculated context score for each subject is multiplied by the context confidence of that subject so that the predictability of the subject is also reflected in the resulting context score.

After calculating the context score, the context matcher outputs a rank list of the top N subjects with the most compatible behavioral patterns and their corresponding context scores. This assignment, while more complex, resembles better the real life scenario.

3.4 Information fusion

The main purpose of information fusion block in a multimodal biometric system is to combine the different sources of information and make the final decision. As discussed in Section 2.5.1, the sources of information can be combined at different stages: sensor level, feature level, match level and decision level. Combining the gait patterns and behavioral patterns at sensor level is not meaningful for my system. Furthermore, since the gait feature is a 128*50 image, but the context feature contains only three parameters, combining these two features at feature level does not seem appropriate. Combining the gait patterns and behavioral patterns at decision level is also not wise because the behavioral patterns are not as distinctive as the gait patterns and identifying subjects only based on the behavioral patterns wouldn't be useful. As a result, the two sources of information in my system are combined at match score level. In this type of fusion, as discussed previously in Section 2.5.1, there is an independent matcher for each source of information which generates a matching score for that modality. The resulting matching scores are then combined to a final matching score that will be used for making the final

decision (12). In this thesis, I have two sources of information: gait patterns and context. Since I am using match score level information fusion, I need a mechanism to find the gait score and another mechanism to calculate the context score and then the two scores will be combined to make the final decision. The flow chart of the information fusion block is illustrated in Figure 3-10.

As can be seen in Figure 3-10, the gait recognition system introduced in section 3.2 is responsible for calculating the gait scores. As mentioned in section 3.2.4.2, the output of this system is a list of the top N subjects and their corresponding gait scores. The gait score for each subject is the similarity of the candidate's gait feature (GEI) to that of the probe calculated using equation (3-4).

Similarly, the context matcher introduced in section 3.3 is responsible for generating the context scores. As described in section 3.3.3, the output of the context matcher is the list of the top N subjects with their corresponding context scores. The context score for each subject shows how well the context of the probe matches with the behavioral patterns for that subject.

Having the gait scores and context scores, the information fusion block uses two following main modules for finding the final scores:

1- Normalization: since the matching scores for my two modalities can have different ranges and distributions, it is essential that I transform them into a common domain before combining them. For this purpose, the normalization module of my system normalizes the gait scores and also the context scores independently to the range [0,1] using the min-max normalization. In this process, the module finds the minimum and

maximum of the gait scores of all the subjects (min_{Gait} and max_{Gait} correspondingly) and normalizes the gait score of subject “s” ($Gait_score_s$) using the following equation.

$$Gait_score_normalized_s = \frac{Gait_score_s - min_{Gait}}{max_{Gait} - min_{Gait}} \quad (3-7)$$

A similar procedure is applied on the context scores according to equation (3-8).

$$Context_score_normalized_s = \frac{Context_score_s - min_{Context}}{max_{Context} - min_{Context}} \quad (3-8)$$

In equation (3-8), $min_{Context}$ and $max_{Context}$ are the minimum and maximum of the context scores of all the subjects correspondingly. In the same way, $Context_score_s$ and $Context_score_normalized_s$ are the context score and the normalized context score of subject “s” accordingly.

2- Weighted combination: this module calculates the final score of each subject as the weighted combination of the normalized gait score and the normalized context score with the weights being the confidence of each database. The confidence of the context database shows how predictable the subjects are in their behavioral patterns or, in other words, how reliable the behavioral patterns in the context database are. The confidence value of the gait database shows how well the gait recognition system works in identifying the subjects. If the quality of the gait samples is not acceptable and if, for any other reason, the gait recognition system does not show good performance for a specific application, setting the confidence of the gait database to a smaller value can improve the performance and reliability of the overall multimodal system by putting more weight on the context patterns. In the same way, if the environment and its subjects are not predictable in their social behavior, the importance of the context data can be reduced by setting the confidence value of the gait database to a small number. This decision is ultimately left to the system administrator.

After calculating the final scores using weighted combination, the subjects are sorted based on their final scores and the system outputs a rank list of top N subjects.

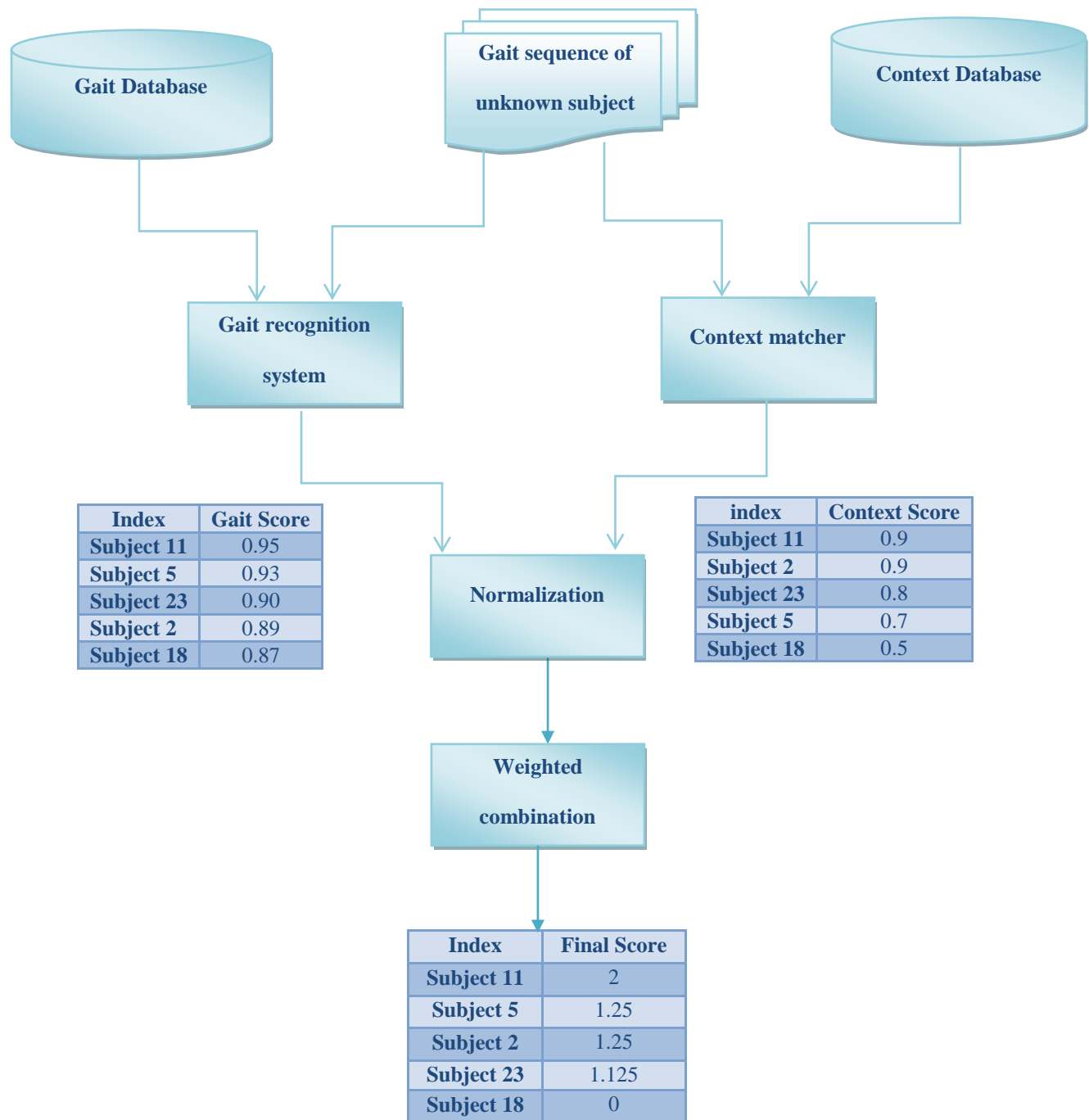


Figure 3-10: Flowchart of the information fusion block

CHAPTER 4 : IMPLEMENTAION DETAILS, EXPERIMENTATIONS AND RESULTS

The goal of this chapter is to validate the proposed methodology on examples of video databases containing gait sequences. The questions I wanted to get answers on are: How the recognition rate of the system is affected after including the context data and how much performance gain can I obtain by fusing the context? How much overhead is associated with fusing the context data? How does the performance of context-based gait recognition get affected if the subjects do not behave according to their predefined behavioral patterns and whether the weight assignment technique is helpful for this scenario? How the technique used for information fusion can affect the performance of the system? How does the method work on real data and how much improvement can I gain with this approach? I aim to answer these questions through various experimental settings and conclude that method indeed offers benefits of increased recognition rate under less than ideal input data conditions with a very low overhead.

Section 4.1 of this chapter describes the implementation details of the proposed context-based gait recognition system. Afterwards, the procedures for setting up the context and gait databases are discussed in Section 4.2. Section 4.3 explains how the performance of the system is evaluated and which performance measures are used for this purpose. Finally the performance of the system on the introduced datasets, with and without fusing the behavioural patterns, is reported for different experiments in Section 4.4.

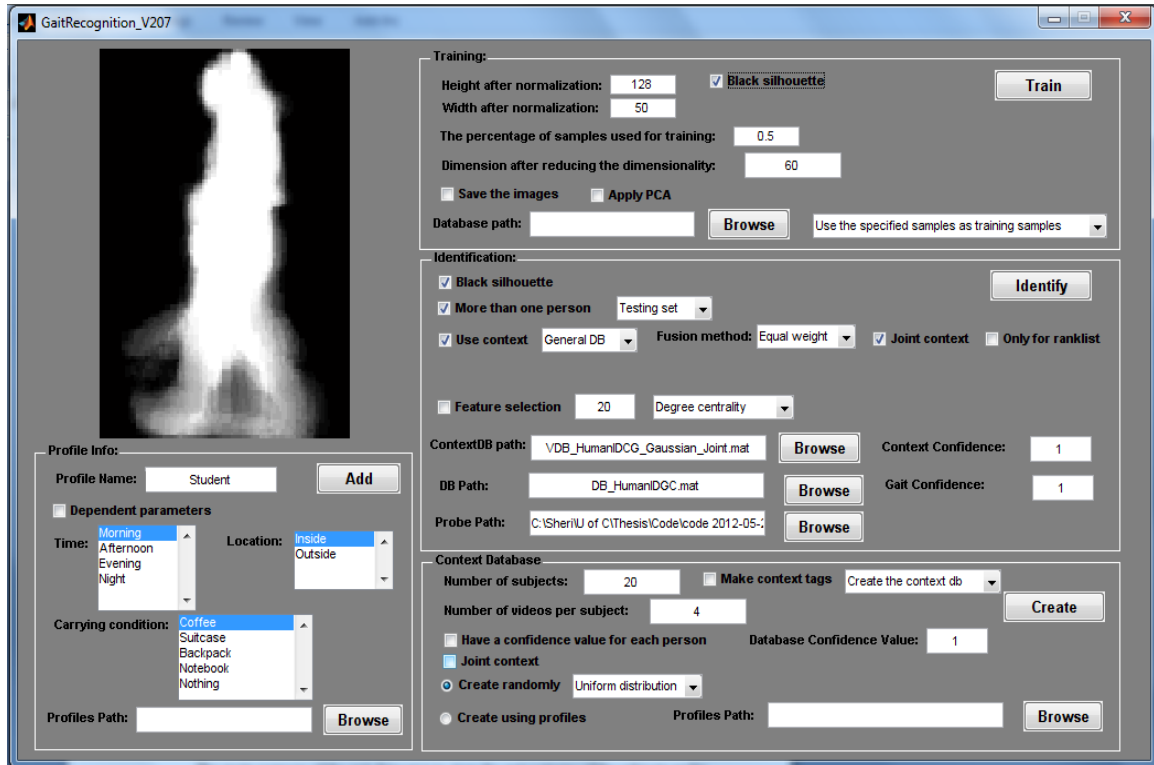


Figure 4-1: The program GUI

4.1 Implementation details

The proposed context-based gait recognition system has been implemented using MATLAB 7 on an Intel Core i5 CPU and Windows 7 operating system. The user interface of the program is shown in Figure 4-1. The interface has different options for training and testing the system. There is also interface for creating virtual context database. The system can be used in two different modes: simple gait recognition (no context) and context-based gait recognition. The only parameters involved in context-based gait recognition are the confidence values of the databases. The user can specify their values using the GUI (graphical user interface) – as shown in Figure 4-1. The

default values are set to 1, but the system has been tested for different values of these parameters and the obtained results are reported in this chapter.

4.2 Experimental data

As illustrated in Figure 3-1, the context-based gait recognition system uses two main databases: the gait database and the context database. The gait database is the database of gaits of the system users. Similarly, the context database is the database of behavioural context patterns of the system users. In my experiments, I use two gait databases that are introduced in Section 4.2.1. For the context database, I have considered two options. For the first set of experiments, I created the virtual context database using the approach discussed in Section 4.2.2. For the second set of experiments, I used the proposed learning approaches to learn the context database from the gait sequences on real video samples.

4.2.1 The gait databases

The performance of a gait recognition system is greatly influenced by the type of gait database used for evaluating the algorithm. If the database is too small, too constrained and does not cover real world scenarios then, even if the system achieves a high recognition rate, there is a serious doubt it will perform well in real world applications. On the other hand, if the database contains too many subjects with very different varying conditions, then it is much harder to achieve high recognition rates on such database. Therefore, collecting a comprehensive database, not too easy and not too hard, is an essential task in system evaluation (27). To make sure my system can be used in real applications and its performance does not degrade by registering more subjects in the system and involving more complicated scenarios, I used two of the most popular gait

databases with different difficulties for evaluating the system. The first gait database is small and simple and the second gait database contains more subjects and wider variety of gait samples per each subject. The details about these two databases are provided below.

4.2.1.1 Dataset A from CASIA gait database

This dataset contains 20 subjects and 12 different sequences per each subject captured at different viewing angles. The dataset is publicly available online and can be downloaded from <http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>. One benefit of this dataset is that the binary gait silhouettes are already extracted and available as a part of the dataset. As mentioned, there are 12 sequences available for each subject but since GEI mainly works for side-view video sequences, only the side-view sequences have been used in the experimentations. Each subject has 4 side view gait samples in this dataset. Some examples of the side view binary silhouettes in CASIA gait dataset for two viewing direction are shown in Figure 4-2.

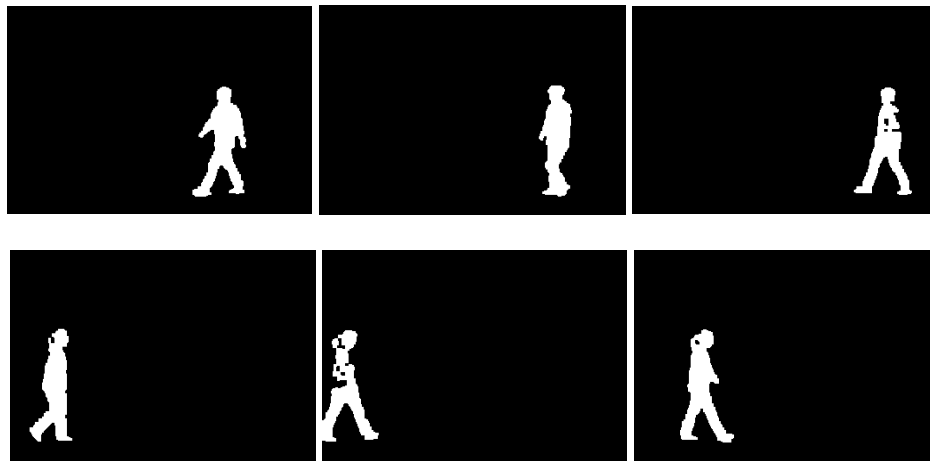


Figure 4-2: Silhouette examples from the CASIA gait database

Table 4-1: The twelve probe sets of the Human ID Gait Challenge dataset
V: Viewing angle, Sh: Shoe type, S: Surface type, B: Briefcase, T: Time.

Probe set	A	B	C	D	E	F	G	H	I	J	K	L
Size	122	54	54	121	60	121	60	120	60	120	33	33
Covariate	V	Sh	V/Sh	S	S/Sh	S/V	S/V/Sh	B	B/S	B/V	T/Sh	T/S

4.2.1.2 HumanID Gait Challenge Dataset

The HumanID Gait Challenge Dataset was introduced by Sarkar et al. in (27) with the purpose of finding the critical factors affecting the performance of gait recognition algorithms and designing a performance benchmark for gait recognition systems. In order to simulate the real world scenarios, the dataset was collected outdoor to include shadows and dynamic backgrounds. This dataset contains gait sequences for 122 subjects. Five covariates were taken into account. These factors are surface type, shoe type, carrying condition, viewpoint and time. For each subject, two different settings were considered for each of these critical factors. This includes two viewing angles, two types of shoes, two types of surface (grass and concrete) and two carrying conditions (with and without a briefcase). For the last factor, time, two videos were captured six months apart. This last factor aims to capture the changes happening in the gait of the subject over time. The resulting 1870 video sequences are organized into 12 challenge datasets with different difficulties in order to examine the effect of each of the covariates and also their combinations (27). Each probe set is designed in a way that it differs from the training set in one or more of the covariates and there is no common sequence between the training set and the twelve probe sets. More information about the twelve probe sets is provided in

Table 4-1. The size of the probe set is the number of samples in the probe set and the covariate shows in which factors the probe set is different from the training set.

The HumanID gait dataset is available online and can be downloaded from <http://figment.csee.usf.edu/GaitBaseline/>. The binary silhouettes have already been extracted and also centralized and are available as a part of this dataset. Some examples of the binary silhouettes from the HumanID gait dataset are presented in Figure 4-3.

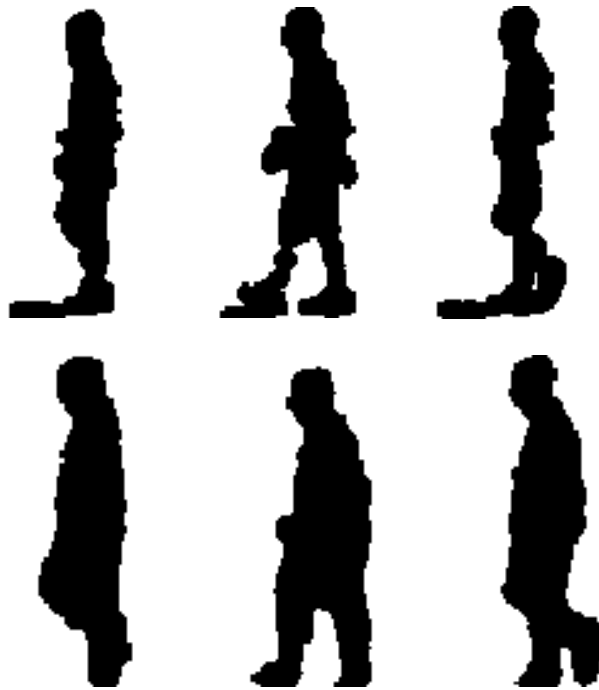


Figure 4-3: Silhouette examples from the HumanID gait dataset. Top row: a subject walking on a concrete surface, Bottom row: a subject walking on a grass surface

By comparing Figure 4-2 and Figure 4-3, it is clear that the quality of the silhouettes in the HumanID gait dataset is not as good as the CASIA gait dataset. The HumanID gait dataset has been recorded outdoor and under varied conditions. As a result, the gait sequences are noisier and they show a wider variability. As an instance, the gait

sequences for concrete surface typically have shadows. On the other hand, in the gait sequences for the grass surface, the feet are sometimes covered by grass and are not fully visible in all the frames. The lower quality, the wider variability and also the higher number of subjects in this dataset make this set of data more challenging compared to the CASIA gait dataset. Lower performance of the system on this dataset, which is reported in this chapter, reflects this observation.

4.2.2 *The context databases*

The main purpose of the context database is to store the behavioural patterns of the subjects for whom the gait samples are available in the gait dataset. Therefore, for each gait database a corresponding context database exists with the same number of subjects. In this thesis, since there was no opportunity to work with real human subjects due to ethics and privacy issues, two approaches have been considered for creating the context database. First, the common approach in biometric community for creating test databases is used to create the context database virtually. The virtual creation of context database provides a flexible framework for conducting a variety of experiments and checking if the design decisions were optimal. However, to make sure the system can actually work with real data, in the second set of experiments, the context database is learned from real data using the learning approaches discussed in Section 3.3.1.3. In the following, I will describe how the virtual context database is created. Afterwards, I will explain how I learn the behavioral patterns from the real gait sequences to build a real context database.

4.2.2.1 Creating virtual context database

The main step in creating a context database is to model the behavioral patterns of the subjects. For evaluating the performance of the three methods of modeling the behavioral

patterns introduced in section 3.3.1.2 (Random models, Gaussian models and Behavioral profiles), a variety of virtual context databases have been created using each of these methods. Furthermore, two different cases have been considered for each method: independent context parameters and joint context parameters. As a result, six context databases have been virtually created for each gait database. In the following, I will describe how these databases are created.

1- Random models

In this approach, it is assumed that there is no limitation or condition for the values the parameters can take for each subject. Thus, for creating the behavioral patterns of each subject, some of the possible parameter values are randomly selected and assigned to that subject. If the parameters are considered independently, then, for each parameter and for each subject, a number of acceptable values are randomly assigned to the subject. The number of values assigned to each subject for each parameter is equal to the number of training samples available for that subject. I do not check if there is any duplicate value in the list of randomly generated values for each parameter. This way, the subjects can have different number of behavioral patterns for different parameters. If the parameters are considered all together (joint parameters) then for each subject some combinations of the three parameters are generated and assigned to the subject. Once again, the number of behavioral patterns assigned to each subject is equal to the number of training samples for that subject. This is based on the assumption that each training sample for the subject serves as a new piece of evidence for the behavioral patterns of that subject.

2- Gaussian models

In this approach, the distribution of each parameter for each subject is assumed to be a Gaussian distribution. Therefore, for defining the parameter values for each subject I need to first determine the parameters of the Gaussian distribution. The parameters of the Gaussian distributions (mean and variance) for each subject and each parameter are generated randomly. For example, the parameter time has four possible values (morning, afternoon, evening and night). Therefore, a random number is generated from the interval [1,4] to be used as the mean of the distribution. For generating the variance different intervals can be used. In this work, the interval for generating the variance of the Gaussian distribution is selected automatically based on the valid range of each parameter and the randomly generated mean. If the valid range for parameters p is $[1, n_p]$ and the randomly generated mean is μ_p , then the standard deviation σ_p is obtained using equation (4-1).

$$\sigma_p = r * \max(|\mu_p - 1|, |n_p - \mu_p|) \quad (4-1)$$

In equation (4-1), r is a random number in range [0,1]. The reason I used equation (4-1) for finding the standard deviation is that I did not want the Gaussian distribution to become too wide which can result in generating random numbers out of the acceptable range. Using the equation (4-1), I am trying to limit the generated values to the acceptable parameter range and at the same time I do not have a predefined fixed value for the variance. Having the mean and variance, the next step is to assign parameter values to the subject by drawing samples from the obtained Gaussian distribution. Since

the parameter values are all discrete, the generated number should be rounded to one of the possible values of the parameter.

The number of generated behavioral patterns for independent and joint context parameters in this approach is determined similar to the random modeling approach.

3- Behavioral profiles

In this approach, some user profiles are generated based on the similarity of the behavioral patterns of the users. I have chosen the university as my environment, and three profiles have been created: student, professor and staff. For each of these profiles, the behavioural patterns have been created. The same behavioural profiles are used for all gait databases used in the experiments.

Having the profiles, each subject is randomly assigned a profile in my virtual database. Based on the assigned profiles, a location, time or carrying condition is acceptable for a subject if that location, time or carrying condition is acceptable for his/her profile.

4.2.2.2 Learning context database from real data

In this approach, I learn the behavioral patterns of the subjects and create a real context database using the HumanID challenge dataset. The reason I chose this dataset is that the gait samples of this dataset have been recorded under varied conditions, therefore, they have diverse context. Some examples of the videos are shown in Figure 4-4.

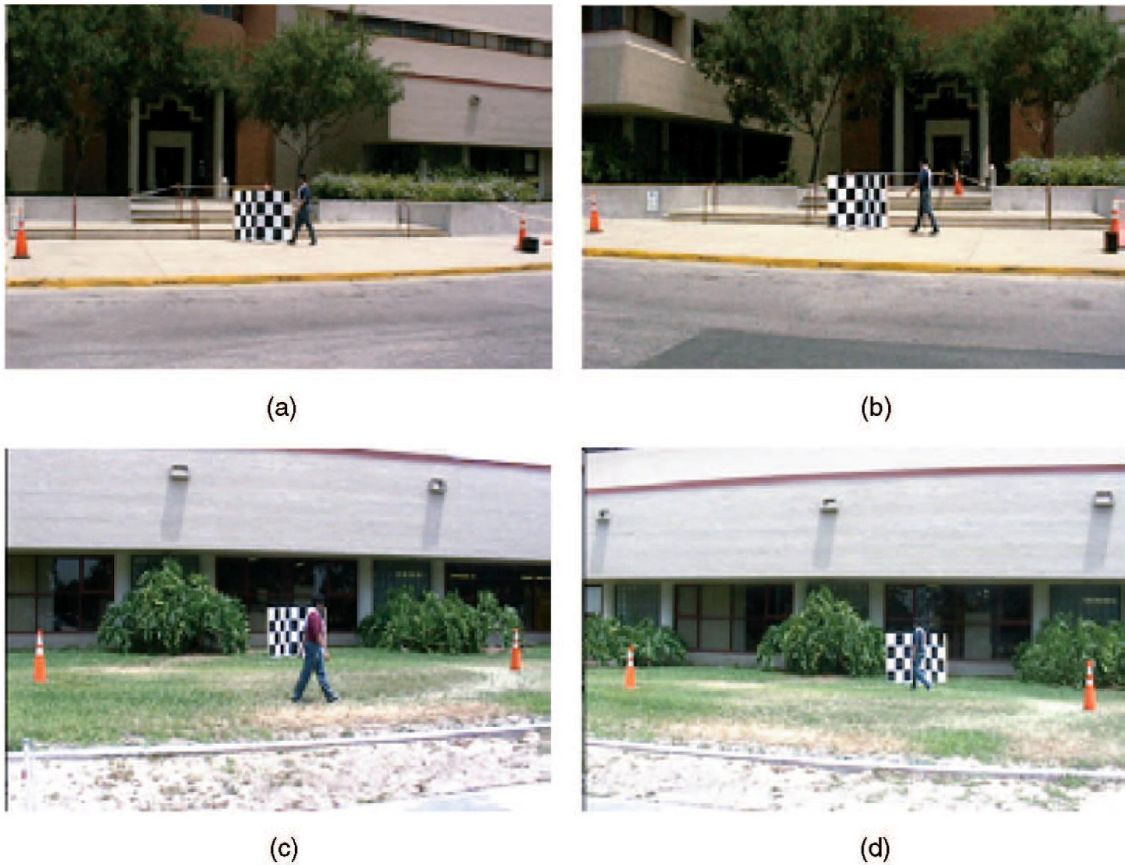


Figure 4-4: Gait samples from HumanID challenge gait dataset with different context (a) left camera and concrete surface (b) right camera and concrete surface, (c) left camera and grass surface (d) right camera and grass surface (27)

The five covariates of this dataset (viewpoint, shoe type, surface type, time and carrying condition) all represent contextual information and are extracted from the gait videos. Furthermore, for each of these covariates two values have been considered in this dataset: for the viewpoint left and right camera, for the shoe type A and B, for the surface type concrete and grass, for the time November and May and for the carrying condition with and without briefcase (Figure 4-4). This implies that all the covariates have discrete values (only two values in this case). Therefore, I can directly map the five covariates to my context parameters because they are contextual information and they are also discrete.

The gait sequences of this real dataset are all labeled with the corresponding values for each of the five covariates. Thus, I can extract the value for each context parameter of a gait video by parsing its label. Therefore, I can tag all the gait sequence with the context data. Having the tagged silhouettes, I use and compare the two approaches (Random models learning and Behavioral profiles learning) discussed in Section 3.3.1.3 to learn the context database.

4.2.2.3 Context tagging

After creating the context database, for evaluating performance of the system on a gait database, all the video sequences in the gait database should be tagged with the context data. For the case that the context database is learned, the gait sequences are already labeled with context tags and nothing more needs to be done in this case. However, for the case of virtual context database, to make a comprehensive test case for testing the system and to make sure that the context of the videos are following the behavioral patterns virtually created for the subjects, each sample of the gait database is virtually assigned with a context as discussed in the following.

The virtual context tags assigned to the samples in the gait database depends both on the context database and confidence values. The context tags are assigned to gait samples of each subject according to the behavioral patterns of that subject. However, there also could be some samples violating the behavioral patterns. The frequency of such samples is consistent with the context confidence of the subject. Therefore, for tagging the videos with context data, for each gait sample of each subject in the gait database, first a random number in the interval $[0,1]$ is generated. If the random number is less than the subject's context confidence, one of the behavioral patterns of the subject is randomly selected

from the context database and is assigned as the context tag of the current sample. However, if the generated random number is greater than the subject's or the context database's confidence, a completely random context is created and assigned to the sample. The former case corresponds to a sample that is consistent with the behavioral patterns of the subject. In contrary, the latter case corresponds to a situation where the subject is acting outside of his/her behavioral patterns. Different confidence values for subjects and context databases have been used in the experiments and the obtained results are reported in Section 4.4.

4.3 Performance evaluation

To evaluate the performance of the proposed system, the gait database is first partitioned into two subsets: training set and testing set and then the system is trained using the training set. The training phase always includes learning of the gait signatures as discussed in Section 3.2.4.2. If the context database is also to be learned, then the learning of the context database will also be a part of the training phase using the same training set. After training the system, the performance of the system in identifying subjects is evaluated using quantitative measures. In this procedure, an unknown gait sequence is presented to the system and the system outputs a rank list of the subjects matching the best with the presented unknown subject. In most studies two performance measures are used to report the performance of the system:

- Rank 1 performance: represents the percentage of the times that the correct answer appeared as the first subject of the rank list.
- Rank 5 performance: represents the percentage of the times that the correct answer appeared as one of the first five subjects in the rank list.

In this thesis, I use Rank 1 performance up to Rank 10 performance to provide more information about how the performance of the system is improved by involving the context data. In general, Rank k performance represents the percentage of the times that the correct subject appeared in the first k items of the final rank list.

4.3.1 Partitioning the dataset

One of the important elements in the system evaluation is partitioning the data into training and testing sets. It is essential to make sure that the system is not biased towards the training data. One of the most popular techniques for this purpose is k fold cross validation (62). In this technique, the data is randomly partitioned into k subsets with equal sizes. The system is then evaluated using these k subsets k times. In each round (fold), $(k-1)$ subsets are used for training and the last remaining subset is used for testing. This process is repeated k times so that each subset is used for testing exactly once. The performance of the system is then reported as the average of the system's performance in the k rounds. In this approach, all data samples are used both for training and testing (62). The performance of the system in identifying the subjects from the CASIA gait database is evaluated using two fold cross validation. The reason behind this decision is that in CASIA gait dataset, I have four walking sequences available for each subject. Therefore, in k fold cross validation I can set the value of k to one of the followings: 2, 3 or 4. Based on the definition of k fold cross validation, in three fold cross validation, the four samples should be divided to three subsets with equal sizes which is not possible. In four fold cross validation, the four samples should be divided to four subsets with only one sample in each subset. Afterwards, the system should be trained using three samples in each round and tested using the remaining one sample. I decided that testing the system only

with one sample will not produce realistic recognition rates and my best option would be to set the value of k to two.

For the HumanID Gait Challenge dataset, the training set and the probe sets are already defined as a part of the dataset definition. Therefore, for the experiments with virtual context database, these predefined sets have been used as training and testing sets. However, for the case that the context database is learned, two fold cross validation have been used for evaluating the performance of the system. The reason is that the standard training set for HumanID dataset contains only one sample for each subject which does not provide enough information for learning the behavioral patterns.

4.4 Experiments

In order to answer the questions arose at the beginning of this chapter, I conducted a set of experiments. The details of the conducted experiments are presented in this section. For each experiment, I explain the questions I wanted to answer and how the result of that experiment can be used in addressing those questions.

4.4.1 Experiment 1: visual inspection of the GEIs

The goal of this experiment is to visually inspect the power of GEI in capturing the gait patterns of the subjects under varied conditions for both gait databases (CASIA and HumanID). The result of this experiment can be useful in analyzing the performance of the proposed gait recognition system for the two gait databases.

Some examples of the Gait Energy Images obtained for CASIA Gait dataset are presented in Figure 4-5 and Figure 4-6. Figure 4-5 shows the binary silhouettes for four subjects and the resulting GEIs.

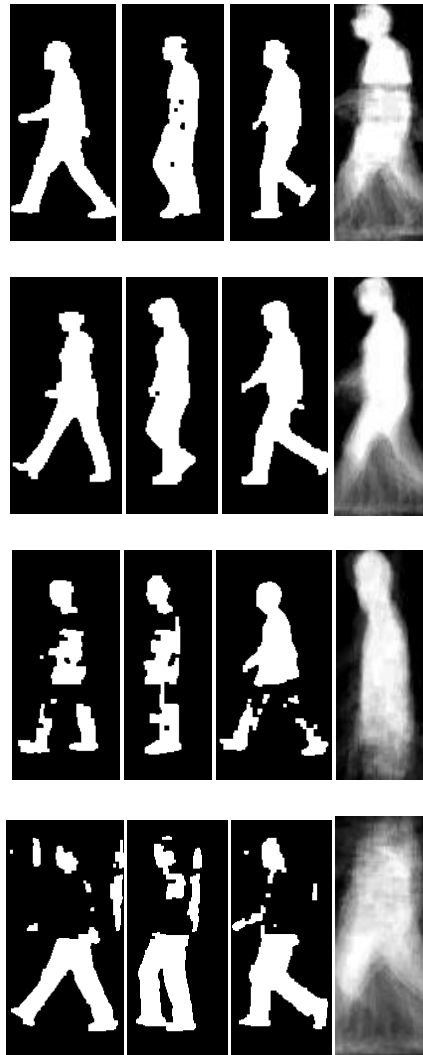


Figure 4-5: The binary silhouettes and the resulting GEIs for 4 subjects of CASIA gait dataset

As can be seen in Figure 4-5, for the first two subjects the quality of the binary silhouettes is acceptable and therefore the resulting GEIs are also of satisfactory quality. For the other two subjects, however, the silhouette is noisy and in some frames it does not even appear as one single connected object. The source of this problem can be the similarity of the color of the subject clothes to the background color. As a result, the noise removal preprocessing step (Section 3.2.1), that keeps only the biggest connected component of the silhouette image, removes some parts of the silhouette (the legs for the

third subject and the torso for the fourth subject). Consequently, the resulting GEIs also miss the information that the removed body parts represent from the walking movement. These cases can be a source of a problem for the gait recognition system. However, out of the 20 subjects of the CASIA gait dataset, only four of them suffer from low quality and noisy gait samples.

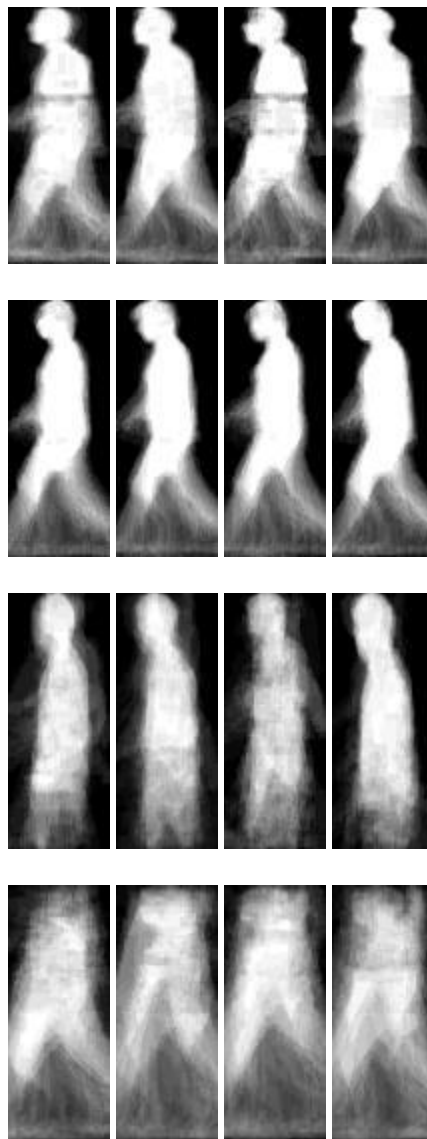


Figure 4-6: Examples of the Gait Energy Images for four subjects of CASIA gait dataset. Each row: the Gait Energy Images for four different gait samples of the same subject

Figure 4-6 shows the Gait Energy Images of all the side view gait sequences for the same four subjects represented in Figure 4-5. As can be seen, although a small variability is visible between the different GEIs of the same subject, they are visually very similar to each other. Even for the last two subjects with the noisy silhouette problem, the GEIs for the same subject still look very similar. Although this property shows the power of GEI in capturing the walking patterns of the subjects, it also comes from the fact that the gait samples for each subjects are all recorded under the same conditions and if the silhouettes for one gait sample are noisy for a subject, all other gait samples almost have the same problem. As a result, since the output of the gait recognition is based on finding the similarity of the GEIs, the noisy silhouettes do not make that much of problem for gait recognition in CASIA gait dataset but this does not mean the algorithm is actually able to handle noisy silhouettes.

Figure 4-7 show the binary silhouettes and the corresponding calculated Gait Energy Images for four subjects of the HumanID gait dataset. Figure 4-8 presents four GEIs of the same four subjects for four different walking samples. As illustrated in Figure 4-7, there are a lot of noisy and troublesome cases in the HumanID silhouettes. All the silhouettes have shadows which are not desirable because they change the shape of the silhouettes. For the third subject in Figure 4-7, the silhouette is disconnected and some parts of the human body are missing. As a result, the obtained GEI is also disconnected and not of satisfactory quality. In the fourth row, the subject is carrying a suitcase which resulted in a little bump appearing at the bottom of the silhouette.

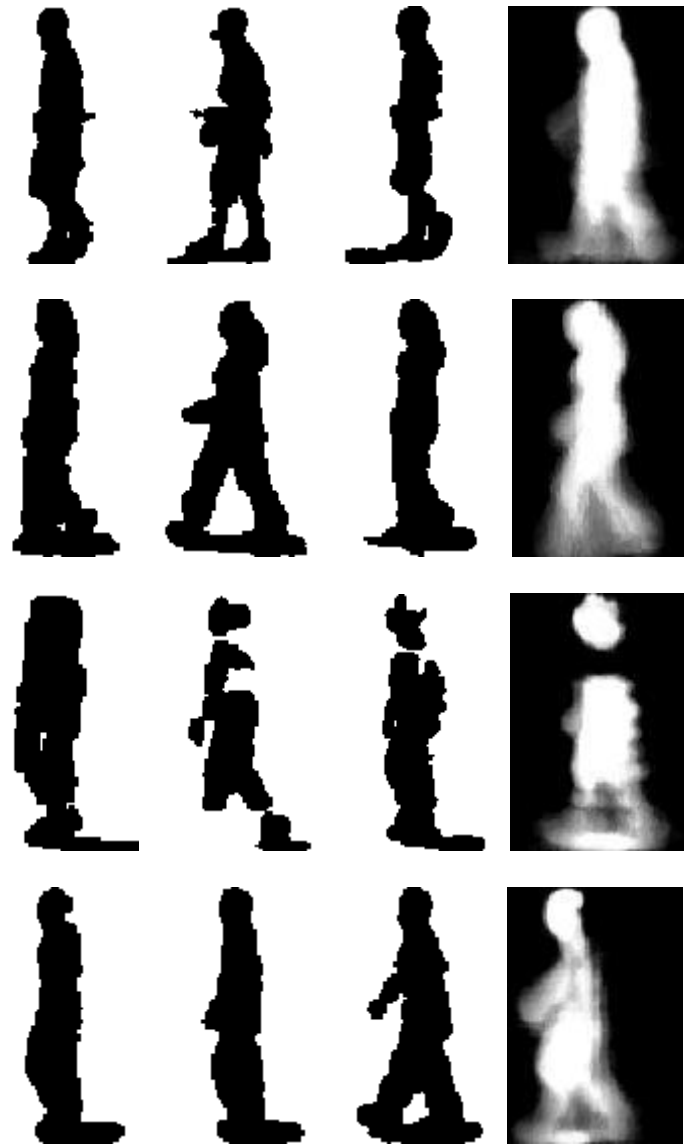


Figure 4-7: The binary silhouettes and the resulting GEIs for 4 subjects of HumanID gait dataset

The drawback of having these troublesome cases is more noticeable in Figure 4-8 which is showing different GEIs of the same subjects. Comparing Figure 4-6 and Figure 4-8, it is clear that the variability of the GEIs per each subject in HumanID gait dataset is a lot more than the CASIA gait dataset. For example, looking at the GEIs of the first and third subjects, we can see that the similarity between the GEIs of the two subjects in some cases is more than the similarity of the GEIs of the same subject. This is the result of

recording the walking movement of each individual under different conditions and at different times. This wide variability will be a source of trouble for the gait recognition module which works based on finding the similarities of GEIs.

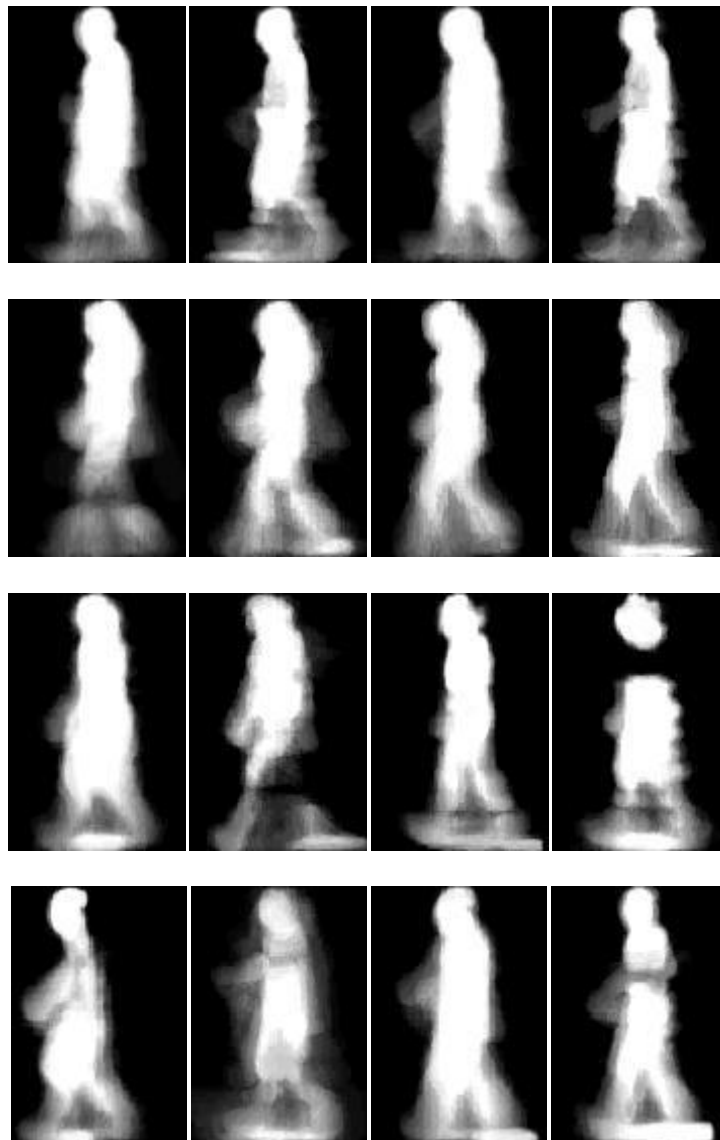


Figure 4-8: Examples of the Gait Energy Images for four subjects of HumanID gait dataset. Each row: the Gait Energy Images for four different gait samples of the same subject

Having seen these examples, it can be concluded that *GEI alone does not seem to be distinctive enough for reliable identification of the subjects, particularly if the gait*

samples are of low quality and are recorded under varying conditions. Therefore, there is a need to augment the GEIs with other sources of information to have a more accurate identification. The results of the rest of conducted experiments prove this conclusion.

4.4.2 *Experiment 2: comparison of three behavioral modeling approaches*

The goal of this experiment is to compare the performance of the three behavioral modeling approaches (Random modeling, Behavioral profiles and Gaussian modeling) introduced in Section 3.3.1.2.

In this experiment, I virtually create context databases using each of the three behavioral modeling approaches as discussed in Section 4.2.2.1. Then, I tag the gait samples using the approach described in Section 4.2.2.3. The confidence values of the context databases and the subjects are all set to one in this experiment. Afterwards, I assess how the performance of the proposed system is influenced by integrating this virtual context data in the decision making process. Both CASIA and HumanID challenge gait datasets are used in this experiment.

The Rank 1, Rank 2 and Rank 5 performance measures for the CASIA gait dataset with and without integrating the context data are reported in Table 4-2:. According to Table 4-2, the GEI can achieve a recognition rate of 73% for Rank 1 and 91% for Rank 5 without involving any context. This table covers the three approaches for modeling the context data (Random models, Gaussian models and Behavioural profiles) for independent and joint context parameters. As it is clear from the table, involving the context data in decision making process improves the performance of the system for all the cases. However, the best result is obtained for the context data virtually created using Gaussian distributions. In the Gaussian modeling of the behavioural patterns, the

behavioural patterns of each individual follow unimodal narrow distributions; whereas in the Random modeling and Behavioral profiles, the distribution of the parameter values is completely random. As a result, the Gaussian modeling makes behavioural patterns more distinctive and there is a less chance of having overlaps between the behavioural patterns of different subjects. Consequently, involving this information increases the discriminative power of the system which justifies why the system shows its best performance for this case. Comparing the Random models and Behavioural profiles, the method used for assigning the behavioural patterns in the Behavioral profiles is the same as the Random modeling. However, instead of having a separate model for each individual, some behavioural groups are defined and all the subjects within the same group are sharing the same behavioural patterns. Therefore, having shared behavioral profiles can make the behavioural patterns of the subjects less distinctive. Consequently, as it can be seen in Table 4-2, the Random and Gaussian models both outperform the Behavioral profiles.

Table 4-2: The performance of the proposed system for the CASIA gait dataset and virtual context database (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context)

Context parameters	Rank 1				Rank 2				Rank 5			
	NC	RC	GC	PC	NC	RC	GC	PC	NC	RC	GC	PC
Independent	73%	81%	89%	82%	82%	93%	98%	91%	91%	100%	100%	100%
Joint	73%	88%	93%	87%	82%	99%	100%	91%	91%	100%	100%	100%

The second row of Table 4-2 shows the performance of the system for the joint context parameters. Comparing first and second rows of the table, it is clear that considering the

relationships between the context parameters improves the recognition rates of the system. The reason is that taking into account the relationships makes the behavioural patterns more distinctive and increases the power of the behavioural patterns in distinguishing the subjects. However, defining the joint behavioural patterns clearly needs more information about the subjects which might not be easy to obtain, therefore, more time should be dedicated to modeling behavioral patterns.

Figure 4-9 visualizes the reported results in the form of Cumulative Match Characteristics (CMC) curve. The CMC curve is designed to show the chance of having a successful identification in the first top ranks. In this thesis, the CMC curve shows the performance of the system for Rank 1 performance to Rank 10 performance. Figure 4-9 shows similar results as Table 4-2. According to Figure 4-9, involving the context data in any case shows significant improvement in recognition rates. This improvement is more noticeable for the lowest ranks and less noticeable for the highest ranks. Comparing the three methods of modeling the behavioural patterns, the Gaussian models of behavioural patterns outperform all the other methods. Although the behavioral profiles are not as good in terms of recognition rates, the convenience and ease of creating them can substantially decrease the amount of time and effort that should be dedicated to modeling the behavioral patterns. Therefore, based on the requirements and limitations of each specific application, the most suitable method should be selected.

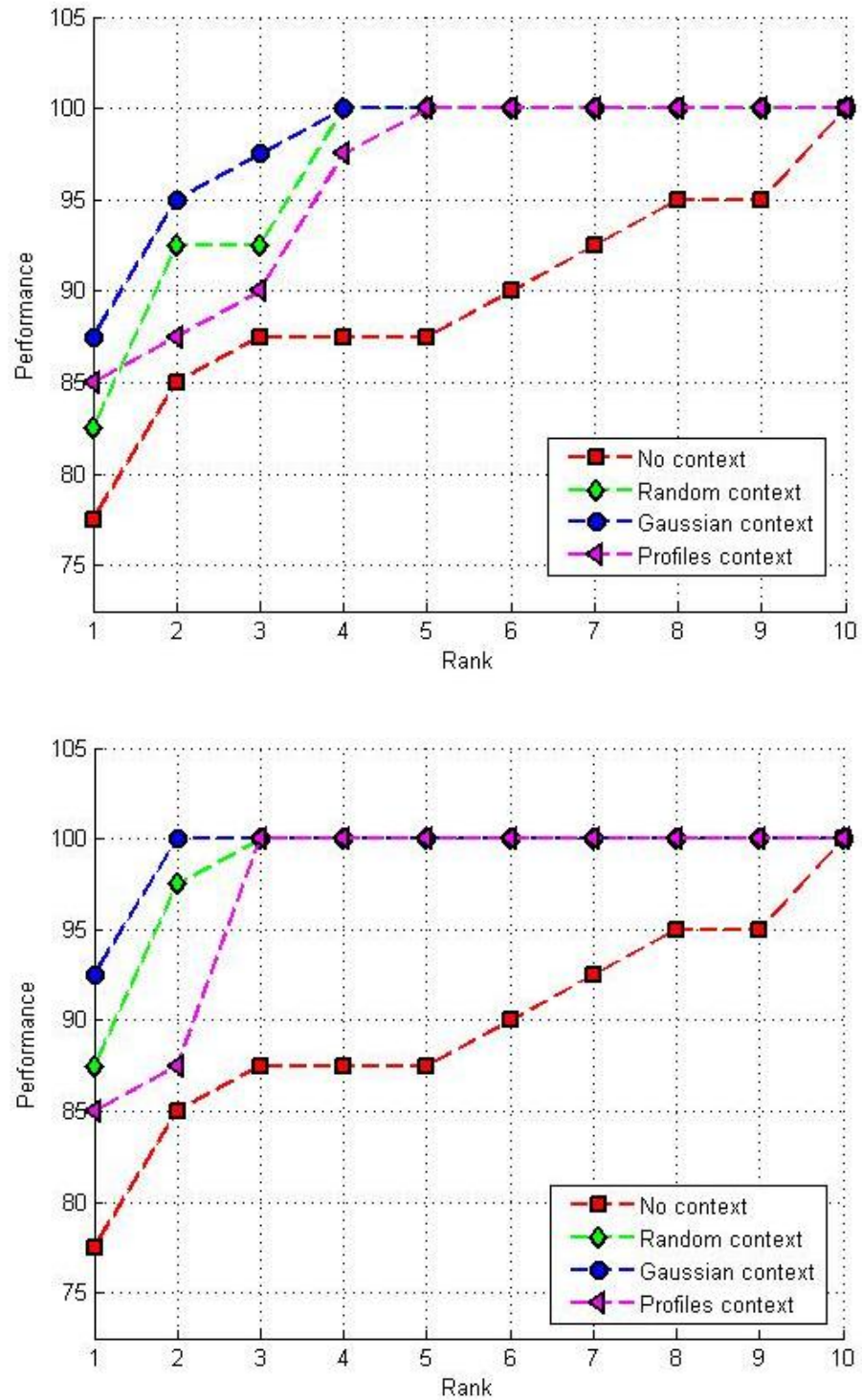


Figure 4-9: The CMC curve of the proposed system for CASIA gait dataset and virtual context database. Top row: independent parameters, Bottom row: joint parameters

Table 4-3: The performance of the proposed system for the Human ID Challenge Gait dataset and virtual context database (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context)

Probe set	Rank 1				Rank 2				Rank 5			
	NC	RC	GC	PC	NC	RC	GC	PC	NC	RC	GC	PC
A	47%	77%	82%	59%	61%	88%	89%	68%	76%	97%	98%	84%
B	80%	85%	91%	85%	83%	94%	94%	85%	89%	98%	98%	93%
C	44%	65%	67%	57%	57%	80%	76%	63%	65%	87%	94%	70%
D	17%	45%	46%	21%	25%	54%	63%	31%	32%	79%	87%	45%
E	16%	40%	40%	17%	21%	50%	64%	31%	36%	79%	93%	45%
F	3%	22%	34%	8%	9%	37%	47%	13%	17%	62%	77%	30%
G	9%	29%	36%	9%	9%	43%	53%	9%	17%	69%	78%	34%
H	46%	74%	74%	52%	60%	84%	86%	66%	69%	91%	97%	79%
I	41%	74%	69%	52%	59%	88%	86%	66%	67%	93%	95%	76%
J	21%	56%	59%	34%	37%	69%	74%	47%	53%	90%	92%	65%
K	20%	43%	49%	43%	30%	55%	57%	43%	42%	60%	79%	46%
L	11%	32%	47%	11%	11%	58%	63%	21%	32%	63%	89%	32%

Table 4-3 presents the performance of the proposed context-based gait recognition system in identifying subjects from the twelve probe sets of HumanID gait dataset with and without involving the context parameters for the three methods of modeling the behavioural patterns. In all context datasets used in this experimentation, the context parameters are considered independently. Based on this table, the worst performance of

GEI is for the probe set F which is different than the training data in surface type and viewing direction and the obtained Rank 1 performance is as low as 3%. The best performance of GEI is for probe set B which is different than the training dataset only in shoe type indicating that maybe shoe type has the less impact on the walking patterns. The calculated Rank 1 performance for this case is 80%.

Looking at the performance of the system once the behavioural patterns are involved, the same trend as the results for CASIA gait dataset is observable here. Integrating context data always improves the recognition rate. *Comparing different approaches for behavioural modeling, the Gaussian modeling is showing the best performance.* Random behavioural modeling is the next best method and the behavioural profiling is the last one. The difference between the Behavioural profiles and the other two methods, that model the behavioural patterns of each individual separately, is more noticeable for the HumanID gait dataset. The reason is that this dataset has 122 subjects as compared to 20 subjects in CASIA gait dataset. Therefore, using only three behavioural profiles for this population might not provide a powerful mechanism for distinguishing the subjects. As a result, it can be concluded that it is a good idea to avoid using very general profiles. In fact, the number of profiles defined for each application should be based on the number of subjects and the variability of their behavioral patterns. The CMC curve for the first eight probe sets of HumanID challenge dataset is shown in Figure 4-10. In this figure, the trend explained above is more visible.

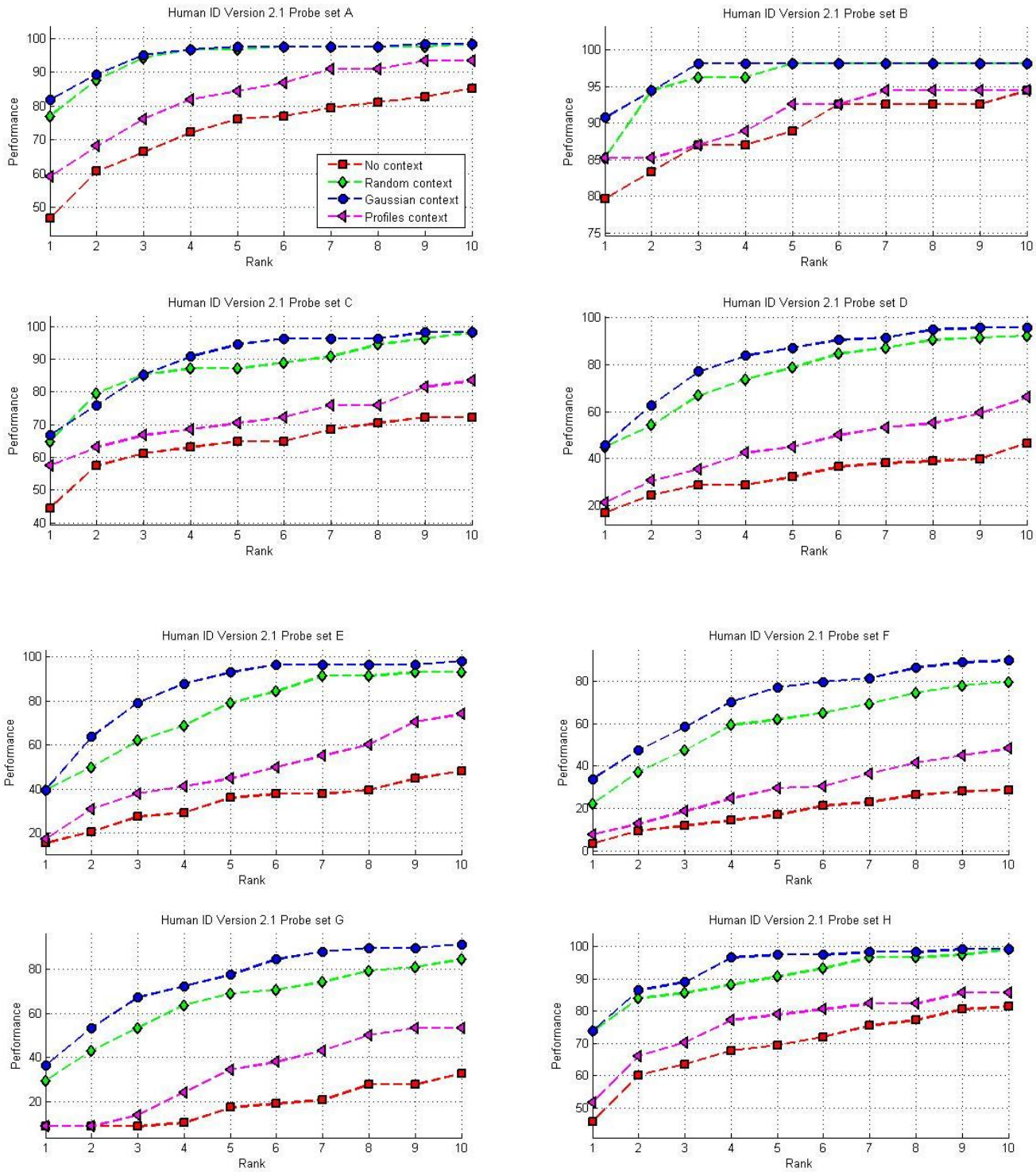


Figure 4-10: The CMC curve of the proposed system for first eight probe sets of HumanID challenge dataset and virtual context database.

Table 4-4 shows the amount of improvement achieved by integrating the virtual context data. The amount of improvement for each performance measure (Rank 1, Rank 2 and Rank 5) is calculated as the value of the performance measure without using the context divided by the value of the performance measure after involving the behavioural patterns.

Table 4-4: The amount of improvement achieved by fusing the virtual behavioural patterns for HumanID Challenge Gait dataset (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context)

Probe set	Rank 1			Rank 2			Rank 5		
	RC/NC	GC/NC	PC/NC	RC/NC	GC/NC	PC/NC	RC/NC	GC/NC	PC/NC
A	1.6	1.7	1.3	1.4	1.5	1.1	1.3	1.3	1.1
B	1.1	1.1	1.1	1.1	1.1	1.0	1.1	1.1	1.0
C	1.5	1.5	1.3	1.4	1.3	1.1	1.3	1.5	1.1
D	2.7	2.7	1.2	2.2	2.5	1.2	2.5	2.7	1.4
E	2.5	2.5	1.1	2.4	3.1	1.5	2.2	2.6	1.3
F	7.3	11.3	2.7	4.1	5.2	1.4	3.7	4.5	1.8
G	3.2	4.0	1.0	4.8	5.9	1.0	4.1	4.6	2.0
H	1.6	1.6	1.1	1.4	1.4	1.1	1.3	1.4	1.1
I	1.8	1.7	1.3	1.5	1.5	1.1	1.4	1.4	1.1
J	2.7	2.8	1.6	1.9	2.0	1.3	1.7	1.7	1.2
K	2.2	2.5	2.2	1.8	1.9	1.4	1.4	1.9	1.1
L	2.9	4.3	1.0	5.3	5.7	1.9	2.0	2.8	1.0

As can be seen in Table 4-4, the amount of improvement is more noticeable for the cases that the original recognition rates are low. For example, the recognition rate for the probe set F without using the context is only 3% but after fusing the behavioural patterns,

significant improvements is achieved for Rank 1 recognition rate which is eleven times higher than the original recognition rate value for the case of Gaussian virtual context database. These results indicate that *behavioural patterns fusion, hence the context-based gait recognition, is most beneficial for the cases that the gait recognition algorithm shows low recognition rates on its own.*

From the results of this experiment, it can be concluded that *involving the context data always improve the performance of gait recognition.* However, the improvement is more noticeable for the cases that the recognition rate of the original gait recognition algorithm is very low. Comparing the three approaches of behavioral modeling shows that the Gaussian modeling approach for creating virtual context data results in better performance by making the behavioral patterns of the subjects more distinctive. The Behavioral profiles are falling a little behind the two other approaches since they are using shared behavioral patterns. However, having more behavioral profiles and defining them more wisely can improve their performance.

4.4.3 Experiment 3: comparison with similar gait recognition systems

The goal of this experiment is to compare the performance of the proposed system on a common gait database with similar gait recognition systems. Since the HumanID gait dataset has well-defined training and testing sets, this dataset is used for this comparison. The performance of the proposed system on HumanID dataset and Gaussian virtual context database is compared with three similar works in the same area and the obtained results are presented in Table 4-5.

The three comparable methods are:

1- The Baseline algorithm (27): this algorithm is introduced by Sarkar et al. in (27) as a part of the HumanID gait challenge dataset. This algorithm provides a framework for extracting the silhouettes, finding the gait cycles and matching the gait patterns. This method matches the silhouettes of the probe with the subjects in the database by calculating their correlation. The Baseline algorithm is introduced as a benchmark for comparing different gait recognition algorithms and all the gait recognition algorithms that report their performance on HumanID challenge dataset, they compare their results with the Baseline algorithm. The recognition rates for the Baseline algorithm are reported in (27).

2- GEI (35): this method was introduced by Han and Bahnu in (35) and was previously discussed in Chapter 2. This work proposes a multimodal gait recognition system that uses GEI as its main gait feature. However, to achieve better performance, some synthetic GEIs are also created for each subject. Two matchers are trained for real GEIs and synthetic GEIs and their results are fused using match score level information fusion. Another difference of this approach with my gait pattern matching is the use of dimensionality reduction algorithms as a combination of PCA and MDA to both decrease the dimensionality and increase class separability. The recognition rates for this method are obtained from (35).

3- GMI (32): this method is introduced by Ma et al. in (32) and has been explained in Section 2.5.3. This approach introduces a multimodal gait recognition system that combines GEI with other gait features at feature level. The recognition rates for this method are based on the reported results from (32).

Table 4-5: Comparing the performance of the proposed system on HumanID challenge dataset and virtual Gaussian context database with similar gait recognition systems

Probe set	Rank 1				Rank 5			
	Baseline (27)	GEI (35)	GMI(32)	My method	Baseline (27)	GEI (35)	GMI(32)	My method
A	73%	90%	84%	82%	88%	94%	92%	98%
B	78%	91%	91%	91%	93%	94%	94%	98%
C	48%	81%	70%	67%	78%	93%	91%	94%
D	32%	56%	26%	46%	66%	78%	55%	87%
E	22%	64%	29%	40%	55%	81%	62%	93%
F	17%	25%	14%	34%	42%	56%	35%	77%
G	17%	36%	16%	36%	38%	53%	45%	78%
H	61%	64%	64%	74%	85%	90%	84%	97%
I	57%	60%	64%	69%	78%	83%	76%	95%
J	36%	60%	42%	60%	62%	82%	76%	92%
K	3%	6%	9%	49%	12%	27%	12%	79%
L	3%	15%	6%	47%	15%	21%	24%	89%

According to Table 4-5, the proposed context-based gait recognition system performs better than all other approaches based on Rank 5 performance for all the twelve probe sets. In a quick look, the proposed method outperforms the Baseline algorithm in Rank 1 performance on all the probe sets. However, the multimodal gait recognition system of Han and Bahnu (35) performs better than my system according to Rank 1 for a few of probe sets (A, C, D and E) due to its powerful and expensive gait feature extraction

algorithm. Similarly, the work of Ma et al. in (32) achieves better Rank 1 performance for only two probe sets because of the same reasons. For all other probe sets, the performance of the proposed context-based gait recognition is equal to or better than the other methods.

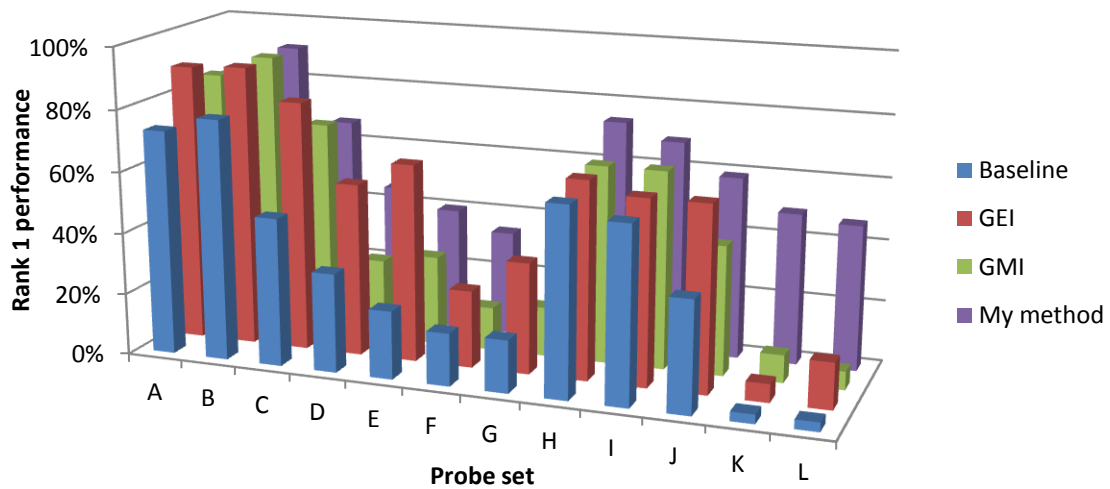


Figure 4-11: Graphical Rank 1 performance comparison of the proposed methods with three similar approaches on HumanID challenge dataset (27) (35) (32)

For further illustration, the rank 1 performance of all methods for all the probe sets is also shown in Figure 4-11. As can be seen, the first three methods are showing similar trends for all the probe sets. The main common property of these three methods is that they only use gait as their biometric characteristic. However, they use different algorithms for gait feature extraction with different levels of complexity. As can be seen, all methods are showing their best performance for probe sets A and B that are different from the training set in viewpoint and shoe type. These results show that viewing direction and shoe type does not have a lot of influence on gait patterns. The worst performance of the first three

methods is for probe sets K and L that are both different from the training set in the factor of time. This result indicates that time is one of the main factors that can change the walking style of subjects. The best performance of my proposed system is similarly for the probe sets A and B. However, the performance does not drop drastically for probe sets K and L and, in fact, the Rank 1 performance is four times higher than the three other methods for these probe sets. The reason is that all the other three methods are only using gait features, however, my proposed system have extra information about the subjects in terms of their behavioral patterns that can be helpful when the gait patterns are not distinctive enough, noisy or of low quality. This is one of the main advantages of using more than one biometric characteristic (gait patterns and behavioral patterns) in my multimodal gait recognition system. Other than better performance, this will also result in better coverage and harder forgery compared to other methods in Table 4-5 as discussed in Section 2.5. From the result of this experiment, it can be concluded that the most benefit of using context-based gait recognition is for the case that the gait recognition performance is low. For the cases that the gait patterns are distinctive by themselves, my proposed method shows similar performance as the others. However, since the gait recognition algorithm is very simple, the overall system is extremely fast. Nevertheless, using more complicated gait feature extraction algorithms in the proposed system might further improve the performance. In fact, the simplicity of the used gait matching mechanism can be the reason that two of the methods are performing slightly better than mine in the first few probe sets (Table 4-5 and Figure 4-11).

4.4.4 Experiment 4: influence of confidence values and weighted sum

The goal of this experiment is to analyze how the performance of the proposed context-based gait recognition system can change or possibly decrease if the subjects violate their behavioral patterns. Can using the confidence values as weights help the proposed system to handle these scenarios?

In this experiment, I use the Gaussian modeling to create a virtual context database for the CASIA gait database. Afterwards, I change the confidence value of the virtual context database from 0.4 to 1. Then, for each confidence value, I tag the gait samples with the virtual context data using the method described in Section 4.2.2.3 and I measure the performance. The results of this experiment are shown in Figure 4-12. The confidence value of zero corresponds to the case that no context data has been used in the gait recognition. As can be seen in the figure, even for confidence values as low as 0.4, still the performance of the system is not degraded by involving the behavioral patterns. The reason behind this is the use of gait and context databases' confidence values as their corresponding weights when the two sources of information are combined (Section 3.4). As the confidence value of the context database decreases, automatically the weight of this database in decision making process decreases and this will avoid the system from getting trapped by behavioral patterns violations. However, *when the confidence value of the context database is high, it can be seen that significant improvements can be achieved by incorporating the context data.*

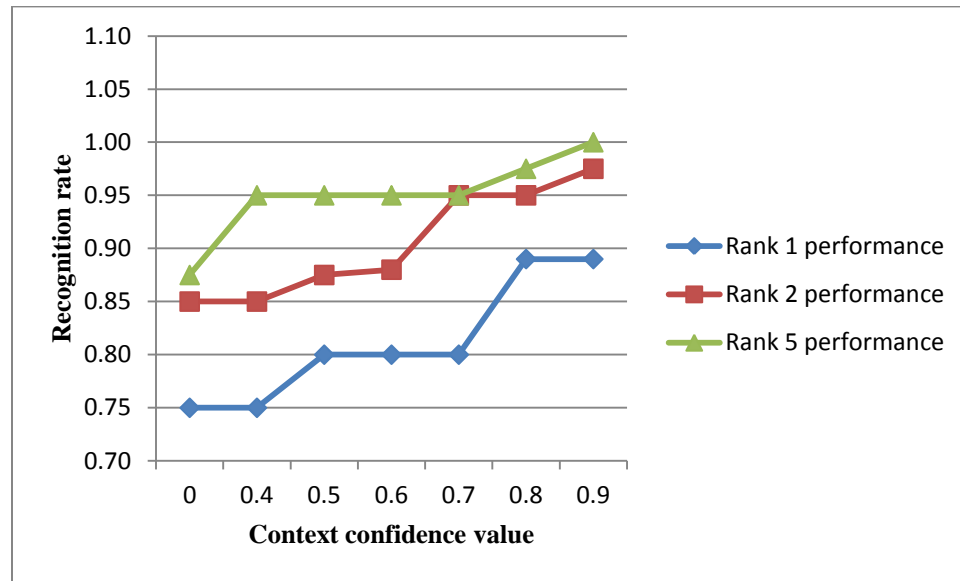


Figure 4-12: The influence of having confidence values less than one for the context database on the performance of the system for the CASIA gait dataset

Table 4-6: The influence of having confidence values less than one for the subjects on the performance of the system for CASIA gait dataset. (NC= No Context, RC= Random Context, GC= Gaussian Context, PC= Profiles Context)

Rank 1				Rank 2				Rank 5			
NC	RC	GC	PC	NC	RC	GC	PC	NC	RC	GC	PC
73%	75%	88%	75%	82%	85%	93%	83%	91%	95%	95%	85%

As the result of a similar experiment, Table 4-6 illustrates how the performance of the context-based gait recognition can change with having confidence value for each subject. In this experiment, the confidence value of the context database is set to one, however, for each subject a random number in the range $[0.8, 1]$ has been generated and used as the confidence value of that subject. The context databases are all the same as the experiments for Table 4-2, however, the virtual context tags are assigned differently to reflect the confidence value of each subject as discussed in Section 4.2.2.3. According to

Table 4-6, having confidence values smaller than one for the subjects does not degrade the performance of the system in any case.

The confidence value for subjects can be used in real scenarios when different subjects have different obligations to their behavioral routines and it is not possible to reflect this information in one confidence value for the whole context database. The confidence value for each subject can be obtained from the subject during registration (for example as a part of the questionnaires). The system administrator can also set up the confidence values based on the type of job each subject has. In a fully automatic approach, however, the confidence values can be learned through time based on how each subject is following his/her behavioral patterns.

From the results of this experiment, it can be concluded that even if the subjects do not follow their behavioral patterns in all cases, because the confidence values of the subjects and the whole context database have been taken into account when combining the two sources of information, *the performance of the system does not degrade by fusing the context data in gait recognition.*

4.4.5 *Experiment 5: comparison of behavioral patterns learning approaches using real data*

The goal of this experiment is to determine how effective behavioral patterns learning approaches are in learning the behavioral patterns from the real video databases and whether fusing the resulting behavioral patterns with the gait recognition system can improve the accuracy of identification.

In this experiment, I used the HumanID challenge dataset for learning both gait patterns and behavioral patterns. The training and testing sets are generated using two-fold cross

validation. Afterwards, the context database is learned using the approach described in Section 3.3.1.3. The result of this experiment is presented in Table 4-7. This table compares the three following cases: having no context database, using context database learned by Random models learning and using context database learned by Behavioral profiles learning.

Table 4-7: The performance of the system for the Human ID Challenge Gait dataset using real context information and two learning approaches (NC= No Context, RML: Random Models Learning, BPL: Behavioral Profiles Learning)

Rank 1			Rank 2			Rank 5		
NC	RML	BPL	NC	RML	BPL	NC	RML	BPL
29%	32%	32%	36%	41%	41%	51%	56%	56%

As can be seen, involving the context data, when the behavioral patterns are obtained by learning, improves the performance of gait recognition and the two learning approaches are showing similar results. The amount of improvement is not as noticeable as my other experiments with virtual context database because I am using the labels of the gait sequences for learning the behavioral patterns. These labels show the value of the five covariates (viewpoint, shoe type, surface type, time and carrying condition) of each sequence. Since the majority of the subjects have been recorded under the two possible values for each of the mentioned covariates, the behavioral patterns of the subjects are very similar to each other. This reduces the discriminative power of the behavioral patterns. As described in Section 3.3.1.3, behavioral profiles learning uses k-means clustering. In this experiment, since the variance of the behavioral patterns for different subjects is very low, I could not create more than two clusters that correspond to two

profiles. This is another evidence of having very similar behavioral patterns. To backup this argument, trying to make the behavioral patterns more distinctive, I pruned the database by completely randomly removing some of the gait sequences. Afterwards, I ran the same experiment on the obtained subset and measured the performance. Results are presented in Table 4-8. As can be seen, more improvement is achieved for the subset of the database and the Rank 2 and Rank 5 performance measures are both increased by 10 percent. I could generate five clusters in this case, which resulted in better overall performance.

Table 4-8: The performance of the system for a subset of the Human ID Challenge Gait dataset using real context information and two learning approaches (NC= No Context, RML: Random Models Learning, BPL: Behavioral Profiles Learning)

Rank 1			Rank 2			Rank 5		
NC	RML	BPL	NC	RML	BPL	NC	RML	BPL
40%	46%	44%	45%	55%	51%	59%	69%	64%

From the results of this experiment, it can be concluded that even *if the behavioral patterns are not provided by the user, the proposed system has the ability to learn them*. Furthermore, fusing the behavioral patterns learned in this process can result in significant improvements in the performance of the system depending on how much overlap exists between the behavioral patterns of different subjects.

4.4.6 Experiment 6: comparison of information fusion techniques using real data

The goal of this experiment is to compare the performance of two major after matching information fusion techniques (match score level and Borda count) and find which approach shows better performance for the proposed system.

As described in Section 3.4, after matching information fusion is the most suitable technique for my proposed system. As mentioned in Section 2.5.1, there are two main approaches for after matching information fusion: match score level and Borda count. Since I wanted to have confidence values for my databases and also due to its popularity, I decided to use match score level. In this experiment, I compared the performance of the proposed system for match score level and Borda count information fusion techniques. The database setup is exactly like experiment 5. I used the whole HumanID challenge dataset and evaluated the performance using two fold cross validation. Based on the obtained results, the Borda count information fusion shows the exact same performance of the match score level technique that is reported in Table 4-7.

4.4.7 Experiment 7: speed

The goal of this experiment is to measure how much computation time is added to the system by fusing the behavioral pattern matching. For this purpose, I ran the program for different scenarios making sure no other program is running at the same time and I recorded the time elapsed with and without integrating the context data for all the gait and context databases. According to the result of this experiment, *fusing the context data adds a maximum of 1% to the computational time of the system for all the different cases and databases*. Therefore, although the proposed approach improves the performance of gait recognition in all cases, it does not add much overhead to the system.

4.5 Summary

In this chapter, I conducted a variety of experiments to evaluate the proposed system using two gait databases with different sizes and difficulties. The first database has 20 subjects and the only changing factor is the viewing direction. The second database has

122 subjects with five covariates (viewing direction, shoe type, surface type, carrying condition and time). In the first set of experiments, I used virtual context database. However, in the second set of experiments, I created the context database from real data using learning approaches. According to the results of the experiments reported in Section 4.4, fusing behavioural models improves the performance of the system in all the cases, even if the behavioural patterns of subjects do not follow any particular distribution and are completely random. Furthermore, even if the subjects do not follow their behavioral patterns, the system is still able to handle the situation by involving the confidence values of the subjects and context databases in information fusion. The amount of improvement achieved is significant particularly if the original gait recognition rate is low. Although significant improvement can be achieved by fusing context, not much overhead is added to the system. The results of experiments with real context data shows that it is feasible to learn the behavioral patterns and automatically build the context database from the gait samples. In addition, using the resulting context database in the system improves the recognition rate.

CHAPTER 5 : SUMMARY, CONCLUSION AND FUTURE WORK

5.1 Thesis summary

In this thesis, I proposed a novel multimodal gait recognition system that takes advantage of the behavioral routines and habits of the subjects to identify them more accurately. Chapter 1 provided my motivation for exploring alternative gait recognition methods, discussed the problems and limitations of gait recognition algorithms and presented the novel idea of using video context as metadata to supplement basic gait recognition algorithm.

Chapter 2 of this thesis explained the core concepts of gait recognition, the main steps of gait recognition systems and the related popular algorithms. It introduced the multimodal biometric systems and their general framework and finally some existing multimodal gait recognition system have been presented at the end of this chapter.

The methodology of the proposed context-based gait recognition system is presented in Chapter 3. This chapter explains how the gait patterns and the behavioral patterns are modeled, extracted and matched and how these two sources of information are combined as the last step of the system using information fusion techniques.

Chapter 4 explained the implementation details of the proposed system and how the data has been set up for conducting the experiments. The performance of the system is evaluated using a variety of experiments and the results are reported at the end of this chapter.

5.2 Conclusions

In this thesis, I proposed the novel idea of a multimodal gait recognition system that combines the gait patterns of the subjects with their behavioral patterns. To the best of

my knowledge, this is the first time that the context metadata has been used as the complementary information in a multimodal biometric system. As a part of this system, I proposed a novel framework for defining, modeling and learning the behavioral patterns. A novel context matcher module was introduced for matching the context of gait videos with the behavioral patterns of the subjects stored in the context database. The output of the context matcher module was then combined with the output of a simple and fast gait recognition algorithm using information fusion techniques.

According to the results of the conducted experiments reported in Section 4.4, fusing behavioral patterns with gait patterns always improves the performance of gait recognition. However, the amount of improvement achieved by integrating the context data depends on the distinctiveness of the behavioural patterns and also the original performance of the gait recognition algorithm. The distinctiveness of the behavioral patterns depends on the nature of the environment, application and its subjects and also the amount of time that has been dedicated to extracting, analyzing and modeling those patterns. It is possible to model the behavioural patterns of each subject individually or the subjects can be grouped based on the similarity of their behavioural patterns and the whole group can share the same behavioral profile. Having a different model for each subject achieves better recognition rates by reducing the overlap between the behavioral patterns of different subjects. But, building behavioral model for each individual can be time-consuming. Using behavioural profiles, however, is easier and faster but it might result in more overlap between the behavioral patterns. The proper method should be selected based on the requirements and limitations of each application. Another factor contributing to the amount of improvement is the original recognition rate of the gait

recognition algorithm. According to the experiments, for the cases that the gait patterns are not distinctive or of acceptable quality, integrating another source of information, the behavioral patterns, can be very beneficial in improving the recognition rate. This observation is also confirmed by comparing the obtained results with other similar methods for gait recognition (Table 4-5).

As a part of the experiments, it has been shown that the behavioral patterns of the subjects can also be successfully learned from the gait sequences. Thus, fusing the obtained behavioral patterns with gait recognition will result in more accurate identification. This is an important property of the proposed system because it makes the system fully unobtrusive by enabling it to independently learn the behavioral patterns of the subjects.

Another benefit of the developed system is the low computational cost of matching and fusing the behavioral patterns. Due to the way the behavioral patterns have been defined and stored in the context database, matching the context data is not an expensive task. Furthermore, extracting the context data does not need high quality videos or specially designed devices, other than the surveillance camera, that has already recorded the gait patterns. Thus, integrating the context does not add to the limitations of the gait recognition system. This is an important advantage of the proposed system, because normally integrating a new biometric characteristic to the system adds the limitations of the new source to the existing limitations of the system. For gait recognition algorithms, however, it is important that the system is still unobtrusive and remotely observable after adding the new source of information.

Since this method does not degrade the performance of the system and is not computationally expensive, thus, it can be a feasible solution for many gait recognition systems. However, since the amount of improvement depends on the distinctiveness of behavioural patterns and the quality of the video samples, the method is most suitable for the predictable environments with predefined behavioural routines or the environments with low gait recognition rates. One of the most suitable areas of application for this system is access control. In access control applications, the users usually have regular routines for accessing the system. Furthermore, in such applications the users of the system are known beforehand and they voluntarily register to the system. Therefore, finding and modeling the behavioural patterns of the subjects is more straightforward because the users can be involved in this process (for example, by filling in the questionnaires). Another application of this system would be reporting a suspicious activity once a good match has been detected for the gait patterns but not a good match can be detected for the behavioural patterns. This corresponds to the case when someone is using the system at the time or location he/she is not supposed to. Finally, the proposed system can be used in security applications. Modeling the behavioural patterns of the subjects in such applications might take more time, but since no cooperation from the subject is needed for modeling the gait patterns and also the behavioural patterns, it is possible to setup the system for such applications.

5.3 Contributions

The contributions of this thesis can be summarized as follows:

- 1- I proposed the use of the behavioral patterns of the subjects as a new generation of behavioral biometric.

- 2- I developed a novel multimodal gait recognition system that combines the gait patterns and the behavioral patterns for the first time.
- 3- I designed and implemented novel methods for defining, modeling and learning the behavioral patterns of the subjects.
- 4- I proposed a mechanism for matching the behavioral patterns of the subjects with the context of the gait video sequences.
- 5- According to the experimentations, the proposed context-based gait recognition system shows the following interesting properties:
 - a. Fusing the behavioral patterns of the subjects always improves the performance of gait recognition and the amount of improvement is significant especially for low recognition rates scenarios.
 - b. Matching and fusing the context data is not computationally expensive and it does not add a lot of computing time to the system considering the amount of improvement it can achieve.
 - c. Matching and extracting the context does not need special devices, high quality data or any cooperation from the subject. Furthermore, the behavioral patterns can be learned from the gait samples without subject's cooperation. Therefore, the proposed system is still fully unobtrusive and remotely observable after involving the behavioral pattern matching.

5.4 Future work

The idea of involving the behavioral routines of the subjects in their identification is quite new and there is a lot of room for improvement. One of the most important future areas

of research in this direction is to conduct studies for gathering real data (both context data and gait patterns) from an extensive number of users and under different conditions and scenarios. In current research, this data was not available due to ethics considerations. Having this data, it would be possible to have a better evaluation of the system. Furthermore, this data can be used for investigating different methods of modeling the behavioral patterns and finding the optimal methods that can fit the context data and can also be efficiently created, stored and matched. In fact, investigating alternative methods for defining and modeling the behavioral patterns is an important area of related research in this field.

The context parameters that have been used in this thesis for modeling the behavioural patterns (time, location and carrying condition) are very general. Fine tuning the system to find the most suitable parameters based on the behavioural patterns of each specific application should be considered in practice. The proposed context-based gait recognition system, however, is defined in such a way that introducing new parameters to the system is really straightforward. It is worth mentioning that involving new parameters needs an extensive analysis of the behavioral patterns of the subjects to make sure the new parameters are distinctive and efficient.

To make the system fully automatic, a context extraction module can be added as a part of the context matcher to automatically extract the context of the video using image processing techniques. Since context extraction involves extracting multiple context parameters from the videos and it also needs to operate in real-time, efficient design of the context extraction module is a challenging area demanding future research.

While these research questions provide rich material for future work, my thesis paves a solid foundation on which future endeavors in gait recognition can be built.

REFERENCES

1. Jain, A. K., Ross, A. and Nandakumar, K. *Introduction to Biometrics*. s.l. : Springer, 2011.
2. *A survey of Biometric recognition methods*. Delac, K. and Grgic, Mi. Zadar, Croatia : s.n., 2004. 46th International SyrnPoSium Electronics in Marine. ELMAR-2004. pp. 184-193.
3. Biometrics. [Online] Wikimedia Foundation, Inc.
<http://en.wikipedia.org/wiki/Biometrics>.
4. Wang, Chun Hao. *A literature survey on human gait recognition techniques*. Toronto, Ontario, Canada : Directed Studies EE8601, Ryerson University, 2005.
5. Gait Recognition Technology May Aid Homeland Defense. [Online] Science News, Oct 15, 2002. <http://www.sciencedaily.com/releases/2002/10/021015073446.htm>.
6. Gait Recognition: Screening From a Distance May Offer Earlier Detection & Quicker Response. [Online] Georgia Tec Research Institute.
<http://www.gtri.gatech.edu/casestudy/gait-recognition-research-strides-ahead>.
7. *A review of vision-based gait recognition methods for human identification*. Wang, J., et al., et al. Piscataway, N.J. : s.n., 2010. 2010 IEEE International Conf on Digital Image Computing : Techniques and Application. pp. 320-327.
8. *Combining multiple evidences for gait recognition*. Cuntoor, N., Kale, A. and Chellappa, R. Washington, DC, USA : s.n., 2003. 2003 International Conference on Multimedia and Expo, IEEE Computer Society. pp. 113-116.

9. *Gait History Image: A Novel Temporal Template for Gait Recognition*. Liu, J. and Zheng, N. Beijing, China : s.n., 2007. 2007 IEEE International Conf. on Multimedia and Expo. pp. 663-666.
10. *Gait recognition using Gait Entropy Image*. Bashir, Khalid, Xiang, Tao and Gong, Shaogang. London : s.n., 2009. 2009 IEEE International Conf. on Crime Detection and Prevention. pp. 1-6.
11. Ross, A., Nandakumar, K. and Jain, Anil K. *Handbook of Multibiometrics*. s.l. : Springer, 2006.
12. *A multimodal Biometric System Using Rank Level Fusion Approach*. Monwar, M. and Gavrilova, M.L. 5, 2009, IEEE Transactions on Man, Systems and Cybernetics TSMC PART B, Special Issue on Cognitive Informatics and Cybernetics, Vol. 39, pp. 867-878.
13. Linoff, G. and Berry, M.J. *Data Mining Techniques: For Marketing, Sales, and Customer Support*. s.l. : John Wiley and Sons, 2011.
14. *Application of Data Mining to Customer Profile*. Kitayama, M., Matsubara, R. and Izui, Y. 2002, Power Engineering Society Winter Meeting, Vol. 1, pp. 632-634.
15. *Understanding individual human mobility patterns*. Gonzalez, M.C., Hidalgo, C.A. and Barabasi, A. 7196, 2008, Nature, Vol. 453, pp. 779-782.
16. *Limits of Predictability in Human Mobility*. Song, Ch., et al., et al. 2012, Science, Vol. 327, pp. 1018-1021.
17. *A Hybrid Method for Context-Based Gait Recognition Based on Behavioral and Social Traits*. Bazazian, S. and Gavrilova, M., TCS Special issue on Advances in Autonomic Computing: Formal and Practical Aspects of Context-Aware Systems, s.l. : Springer, 2012.

18. *Context based gait recognition*. Bazazian, S. and Gavrilova, M. Baltimore, Maryland, USA : s.n., 2012. SPIE 8407.
19. *Visual perception of biological motion and a model for its analysis*. Johansson, G. 2, 1973, *Perception And Psychophysics*, Vol. 14, pp. 201-211.
20. *Recognizing friends by their walk: Gait perception without familiarity cues*. Cutting, J.K. and Kozlowsk, L.T. 5, 1977, *Bulletin of the Psychonomic Society*, Vol. 9, pp. 353-356.
21. *Gait Analysis*. [Online] Wikimedia Foundation, Inc.
http://en.wikipedia.org/wiki/Gait_analysis.
22. Monwar, MD. Maruf. *A multimodal biometric system based on rank level fusion*. Calgary, Canada : University of Calgary, 2011.
23. *Background subtraction techniques: a review*. Piccardi, M. 2004. *IEEE International Conference on Systems*,. pp. 3099-3104.
24. *A Survey of Biometric Gait Recognition: Approaches, Security and Challenges*. Gafurov, D. 2007. NIK conference.
25. *Chrono-gait image: A novel temporal template for gait recognition*, 257-270. Wang, C., et al., et al. 2010. 11th European Conference on Computer Vision.
26. Boulgouris, N.V., Plataniotis, K.N. and Micheli-Tzanakou, E. *Biometrics: Theory, Methods, and Applications*. s.l. : Wiley-IEEE Press, 2009.
27. *The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis*. Sarkar, Sudeep, et al., et al. 2, 2005, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, pp. 162-177.

28. Yam, Cy and Nixon, Mark. Model-based Gait Recognition. *Encyclopedia of Biometrics*. s.l. : Springer, 2009, pp. 633-639.
29. *3D Tracking for Gait Characterization and Recognition*. Urtasun, Raquel and Fua, Pascal. Seoul, Korea : s.n., 2004. IEEE Automatic Face and Gesture Recognition. pp. 17-22.
30. *Automated Markerless Analysis of Human Gait Motion for Recognition and Classification*. Yoo, Jang-Hee and Nixon, Mark. 2, 2011, ETRI Journal, Vol. 33, pp. 259-266.
31. *Gait recognition based on dynamic region analysis*. Yang, X., et al., et al. 9, 2008, Signal Processing, Vol. 88, pp. 2350-2356.
32. *Recognizing Humans Based on Gait Moment Image*. Ma, Qinyong, et al., et al. Washington, DC, USA : s.n., 2007. 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed, IEEE Computer Society. pp. 606-610.
33. *A Full-Body Layered Deformable Model for Automatic Model-Based Gait Recognition*. Lu, Haiping, Plataniotis, Konstantinos N. and Venetsanopoulos, Anastasios N. 2008, EURASIP Journal on Advances in Signal Processing, Vol. 2008, pp. 1-13.
34. *Frontal view gait based recognition using PCA*. Sharma, S., et al., et al. New York, NY, USA : s.n., 2011. International Conference on Advances in Computing and Artificial Intelligence, ACM. pp. 124-127.
35. *Individual recognition using gait energy image*. Han, J. and Bhanu, B. 2, 2006, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, pp. 316-322.

36. *Frame difference energy image for gait recognition with incomplete silhouettes*. Chen, C., Liang, J. and Zhao, H. 2009, Pattern Recognition Letters, Vol. 30, pp. 977-984.
37. *A new representation for human gait recognition: Motion silhouettes image (MSI)*. Lam, T.H.W. and Lee, R.S.T. 2006. International Conference on Biometrics. pp. 612-618.
38. *Wavelet moments for gait recognition represented by motion templates*. Chen, S., et al., et al. 2010. International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). pp. 620-624.
39. *Gait recognition using Radon transform and linear discriminant analysis*. Boulgouris, N.V. and Chi, Z.X. 3, 2007, IEEE Transaction on Image Processing, Vol. 16, pp. 731-740.
40. Giudici, P. and Figini, S. *Applied Data Mining for Business and Industry*. Chichester, UK : John Wiley & Sons, 2009.
41. *Advertising gets personal*. Greengard, S. 8, 2012, Communications of the ACM, Vol. 55, pp. 18-20.
42. One in Five Technology Firms Has Rejected a Job Applicant Because of Social Media Profile – Eurocom Worldwide Annual Survey. [Online] Eurocom Worldwide, 2012. <http://www.eurocompr.com/prfitem.asp?id=14921>.
43. Exchangeable image file format. [Online] Wikimedia Foundation, 2012. http://en.wikipedia.org/wiki/Exchangeable_image_file_format.
44. Optical character recognition. [Online] Wikimedia Foundation, 2012. http://en.wikipedia.org/wiki/Optical_character_recognition#OCR_software.

45. *Indoor scene recognition through object*. Espinace, P., et al., et al. 2010. International IEEE conference on Robotics and Automation. pp. 1406-1413.
46. *Recognizing indoor scene*. Quattoni, A. and Torralba, A. 2009. IEEE Conference on Computer Vision and Pattern Recognition. pp. 413-420.
47. *Tour the World: Building a Web-Scale Landmark Recognition Engine*. Zheng, Y., et al., et al. 2009. IEEE conference on Computer Vision Pattern Recognition.
48. *A Survey on Mobile Landmark Recognition for*. Chen, T., et al., et al. 2009. IEEE international conference on Mobile Data Management: Systems, Services and Middleware. pp. 625-630.
49. *Context based object categorization: A critical survey*. Galleguillosa, C. and Belongie, S. 6, 2012, Computer Vision and Image Understanding, Vol. 104, pp. 712-722.
50. Outline of object recognition. [Online] Wikimedia Foundation, 2012. http://en.wikipedia.org/wiki/Outline_of_object_recognition#Surveys.
51. *Multimodal Biometrics: An overview*. Ross, A. and Jain, A.K. Vienna, Austria : s.n., 2004. 12th European Signal Processing Conference (EUSIPCO). pp. 1221-1224.
52. *Review of Multimodal Biometrics: Applications, challenges and Research Areas*. Mane, V.M. and Jadhav, D.V. 5, 2009, International Journal of Biometrics and Bioinformatics (IJBB), Vol. 3, pp. 90-95.
53. *Score normalization in multimodal biometric systems*. Jain, A., Nandakumar, K. and Ross, A. 2005, Pattern Recognition, Vol. 38, pp. 2270-2285.
54. Wang, Y. and Liu, Z. A Survey on Multimodal Biometrics. *Advances in Automation and Robotics*, Vol. 2. s.l. : Springer Berlin Heidelberg, 2012, pp. 387-396.

55. *Human Recognition Based on Gait Poses*. Félez, R., Mollineda, R. and Sánchez, J. 2011, Pattern Recognition and Image Analysis Lecture Notes in Computer Science, Vol. 6669, pp. 347-354.
56. *Gait Recognition Using Geometric Features and Soft Biometrics*. Moustakas, K., Tzovaras, D. and Stavropoulos, G. 4, 2010, Signal Processing Letters, IEEE, Vol. 17, pp. 367-370.
57. *Person Identity Verification Based on Multimodal Face-Gait Fusion*. Hossain, E. and Chetty, G. 6, 2011, IJCSNS International Journal of Computer Science and Network Security, Vol. 11, pp. 77-86.
58. *Impact of habitual behaviors on human dynamics and spreading process*. Jiao, Y., et al., et al. China : s.n., 2010. 5th International ICST Conference on Communications and Networking. pp. 1-5.
59. *Simulating Human Behaviour in Built Environments*. Wei, Y. and Kalay, Y. Vienna : s.n., 2005. CAAD Futures. pp. 301-310.
60. *Agent-Based Simulation of Human Movements During Emergency Evacuations of Facilities*. mith, J.L. and Brokaw, J.T. 2008, Structures Congress, pp. 1-10.
61. *Some Methods for classification and Analysis of Multivariate Observations*. MacQueen, J. s.l. : University of California Press, 1967, 5-th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, pp. 281-297.
62. Cross-validation (statistics). [Online] Wikimedia Foundation, Inc. http://en.wikipedia.org/wiki/Cross-validation_%28statistics%29.

63. *The recognition of human movement using temporal templates.* Bobick, A. F. and Davis, J. W. 3, 2001, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, pp. 257-267.
64. *Simplest representation yet for gait recognition: averaged silhouette.* Liu, Z. and Sarkar, S. 2004. 17th International Conference on Pattern Recognition. pp. 211-214.
65. *What image information is important in silhouette-based gait recognition?* Veres, G.V., et al., et al. 2004. 2004 IEEE Conference on Computer Vision and Pattern Recognition. pp. 776-782.
66. Bhanu, B. and Han, J. Match Score Level Fusion of Face and Gait at a Distance. *Human Recognition at a Distance in Video.* s.l. : Springer London, 2011, pp. 185-207.
67. Nixon, M., et al., et al. *Experimental plan for automatic gait recognition.* s.l. : University of Southampton, 2001.
68. Gross, R. and Shi, J. *The CMU Motion of Body MOBO Database.* s.l. : Carnegie Mellon University, 2001.
69. *Gait recognition based on dynamic region analysis.* Xiaochao, Yang, et al., et al. 9, 2008, Signal Processing, Vol. 88, pp. 2350-2356.
70. *Score normalization in multimodal biometric systems.* Jaina, A., Nandakumara, K. and Ross, A. 2005, Pattern Recognition, Vol. 38, pp. 2270-2285.