Gait recognition under non-standard circumstances

Kjetil Holien

Master's Thesis
Master of Science in Information Security
30 ECTS
Department of Computer Science and Media Technology
Gjøvik University College, 2008

Avdeling for
informatikk og medieteknikk
Høgskolen i Gjøvik
Postboks 191
2802 Gjøvik

Department of Computer Science
and Media Technology
Gjøvik University College
Box 191
N-2802 Gjøvik
Norway

Abstract

This thesis will look at the biometric feature gait, and how different circumstances impact the ability to perform authentication. We will use an accelerometer sensor in order to record the data. First of all are we going to see if it possible to recognize users under different circumstances, then we will look for a common pattern among these different circumstances. We will also briefly look into whether people walk in the same way given the same circumstances. In order to answer these questions will we perform an experiment and do a thorough analysis of the data obtained. In a world where e.g. mobile devices becomes more and more important, it is obvious that the protection of these devices must be satisfactory. A possible security feature of such devices is to use gait authentication. When this shall become a reality however it is important that the authentication works under circumstances which are common during a normal day.

Acknowledgment

First of all would I like to thank all the participants that voluntary participated in my experiments, without them this project would have been impossible. Secondly, and almost just as important is my supervisor, Patrick Bours, who has during the entire period always been available with good answers to all my questions and problems. In addition has he, with his enthusiasm around this topic, provided plenty of motivation to produce as good thesis as possible. My good friend and classmate Knut Steinar Watne also deserves an acknowledgement for a very thorough and good feedback as my opponent. I would also like to thank MotionRecording [1] and Torkjel Søndrål for providing the sensor used to collect data as well as the software to transfer the raw data to a PC. Last but not least, would I like to thank my girlfriend for her love, support and understanding during the whole period I have been working on this thesis.

Kjetil Holien, 27th June 2008

[1] MotionRecording: http://www.motion-recording.com/

Contents

List of Figures

List of Tables

1 Introduction

Humans have for ages used characteristics on people to identify[1] and authenticate[2] them, either by face, voice, fingerprints, etc [2]. In recent times human recognition has become an important task in a variety of applications, such as access control and surveillance [3]. Authentication can happen in many ways, but all authentication factors can be categorized into one of three classes:

- Something you *know*: such as a password or a PIN code.
- Something you *have*: such as a key or a smart card.
- Something you *are*: this includes all biometric properties, such as fingerprints.

In this paper we will look at something you are, which consequently utilizes biometric features of a person. Biometric features can be divided further into two main categories [1]:

- Physiological: properties that will normally not change, such as fingerprints and your iris.
- Behavioral: properties that are learned, such as signature and gait[3].

When using physiological biometrics the user must usually interact with a system, for instance scanning his fingerprint on a fingerprint reader. With behavioral biometrics the data that are collected can normally be recorded when the user performs his natural duties, for instance talking in a phone. Even though behavioral biometrics requires less user interaction, it is physiological biometrics that is used most. The reason for this is that physiological biometrics perform in general better when a system is authenticating/identifying a user. This fact leads to the growing interest to improve behavioral biometrics [4].

1.1 Topics covered by the project

This project will look at gait as a biometric feature. Gait recognition has become an area which has gained a lot of interest the last decade. An important reason why gait has become attractive is that it is non-intrusive, can be measured without subject contact or knowledge and it can not easily be obscured [5]. The most research on gait recognition the last decade has been video-based, where the purpose has been for surveillance, for instance recognizing a criminal from a security camera video [3, 6, 7, 8]. There has also been some research on sensors installed in the floor which can be used in an access control system [9]. In 2005 another identification method which utilize how people walk by looking at acceleration from sensors attached to the belt was presented [10]. This method uses a device called an accelerometer which measures acceleration in three directions (horizontal, vertical and lateral), and such a device will be used in this project.

[1] Identification means establishing an identity [1]
[2] Authentication means verifying a claimed identity [1]
[3] Gait means how people walk

Gait recognition using accelerometer can be used to authenticate and protect mobile phones and other portable electronic devices, where the sensors are integrated into the hardware [11]. As with all other biometric features gait recognition has some weaknesses as well. The cost of deploying floor-based mats with sensors are rather high. When using video-based recognition there are a lot of variables interfering which lower the performance, such as light and other objects interfering with the subject. This common problem does also concern accelerator based authentication, we do get undesired noise that affects our signal. In addition we have the issue with sensor placement. Acceleration data from the ankle is entirely different than e.g. your wrist. Another common problem when using gait is that your gait might be altered if you are injured, drunk, change footwear, surface and so on. In this project we will investigate how gait recognition works under non-standard circumstance. We will try to, by isolating variables, find out how different variables impact the recognition rate. We will also see if there are any common features among the circumstances and if the non-standard circumstances could be adapted to be more comparable with normal walking. Furthermore will we also perform a small long-term experiment where we will see how stable gait is.

1.2 Keywords

Keywords: Biometric authentication, gait, human movement, accelerometer, statistical analysis and pattern recognition.

1.3 Problem description

Gait recognition is normally achieved by walking back and forth on a solid surface and in a straight line. In a real life scenario this is unfortunately not the case, people walk on different surfaces, walk up and down stairs, walk with different speed, use different clothing and shoes, wear or carry backpacks/briefcases etc. These real-life circumstances introduce some challenges when trying to authenticate people by their gait. With these challenges in mind it would be interesting to see how gait recognition works under non-standard circumstance. It would be desirable that by e.g. knowing what kind of surface a person walks on or if a person is walking faster/slower than normal, the data collected by the accelerometer could be adjusted accordingly and compared with the baseline. The baseline to which the data is being compared, is recorded during an enrollment phase under normal circumstances. In the recognition process the system must try to recognize the current situation and transform the data in a standard manner before comparing it with the baseline.

1.4 Justification, motivation and benefits

In todays world portable electronic devices such as mobile phones and PDAs have become a natural and important tool [12]. This technology has exploded during the last years, you do not have to go many years back before for instance the mobile phone was just a communication device. Today however, mobile phones are used in applications like m-banking and m-government [13, 14]. With this in mind it is not hard to realize the consequences if your mobile phone is stolen or lost, the financial and personal data would now be accessible by the thief. Currently the only protection mechanism which resides on such devices is usually a PIN code and perhaps a password, the need for a better security is obvious. In 2006 a feasibility study to use keystroke dynamics to enhance security of your phone was published [15]. As with gait, keystroke dynamics is non-intrusive and is only active when the phone is used. This study showed promising results, but it will need more research in the following years. Furthermore, features like fingerprints [16] and voice [17] have been proposed with various results, but these features are either obtrusive, require user s attention or merely did not perform satisfactory. The use of voice based recognition did actually perform very well under low-noise circumstances, but had more problems with higher background noise [18]. As a result of all these factors gait has come up as an additional way to secure your phone. With gait the system has the possibility to perform continuous authentication. If gait recognition shall become an integrated part of securing electronic devices however, must it be able to perform adequately under different conditions. Another possible area of application for gait authentication could be to integrate a sensor in the shoe, and in a high security facility use that as an additional feature to accept or reject users.

1.5 Research questions

In order to solve the problems outlined there are first some general and technical issues that must be considered:

General issue:

- What is the most practical way to gather gait data which yields the best result?
  There are several places to attach the sensor, one can use the hip, ankle, around the knee etc. It can also be possible to use more than one sensor.

Technical issues:

- Which methods and techniques can be used to analyze gait data?
  In order to analyze the data we collect, they must undergo some pre-processing before it can be handed to the actual matching algorithms. Such pre-processing can e.g be noise reduction and time interpolation.
- Has there been any research related work which can be used more or less directly?
  In order to both ease the workload and achieve as good results as possible it is desirable to reuse some algorithms and tools that have already been used in similar topics.

The main questions we will look deeper into during this project are:

- To what extent is it possible to recognize a person under different circumstances?
  Is it possible to recognize a person under all different circumstances or do some circumstances provide better recognition than others?
- Do the different circumstances have any common features?
  We have to see if there are common features in the walking patterns in the various walking circumstances. This might involve a fixed transformation of the walking signal under a particular circumstance, e.g stretching the walking signal in time and decreasing it in amplitude in the case where the participant walks fast, so that the resulting signal better resembles normal walking.

  In this project we will also look into the following:

- Do people walk in the same way given the same circumstances?
  An issue which needs to be looked into is whether people walk in the same way under the

same condition (surface, clothing, speed, etc.) after a longer period of time.

1.6 Planned contributions

This project will come up with results on how different circumstances affect the ability to recognize people. If the circumstances do not significantly impact the recognizing process or if it would be possible to do some special processing according to what the circumstances are, there should be no problems using a training set from indoor walking for continuous authentication. This means that after the enrollment phase, the authentication could more easily be moved to another place without having to train the system again to accommodate for different circumstances. This would be a huge step in order to secure mobile devices with gait analysis.

2 Introduction to authentication

This chapter is meant for those relatively new to authentication and biometrics, and will give a brief introduction to these subjects. In order to understand terms used later in the report, it is important to be familiar with the terms and explanations introduced in the following sections.

2.1 Authentication

Authentication is an area which has grown over the last decades, and will continue to grow in the future. It is used in many places today and being authenticated has become a daily habit for most people. Examples of this are PIN code to your banking card, password to get access to a computer and passport used at border control. The last is an example of authenticating a human, which is probably the most used during a day. We identify friends and family by their face, voice, how they walk, etc. As we realize there are different ways in which a user can be authenticated, but all these methods can be categorized into one of three classes [1]:

- Something you *know*, e.g. a password.
- Something you *have*, e.g. a token.
- Something you *are*, e.g. a biometric property.

These factors will be explained briefly in the following subsections.

2.1.1 Something you know

Providing knowledge of some secret is perhaps the oldest way of identifying oneself. In ancient times passwords were used among friends [1]. Today passwords are used mainly in computer systems in order to get access to different resources. Another well known example of this authentication factor is the PIN codes we press when we use a banking card. This factor is very cheap, easy to implement and a very fast authentication mechanism. The drawback with using knowledge for authentication is the fact that users have a lot of different passwords and PIN codes for various application. As a result of this people tend to use the same password/PIN at different places and use easy-to-remember passwords like birthdays, family names, pets or a combination of these. This leads to the fact that attackers can easily find the password and use it to gain access to resources he [1] normally does not have access to. When users are forced to remember different passwords and more random looking passwords it often leads to writing them down in an easy accessible place or they simple forget them. All these drawbacks will increase the cost of using something you know as a solely authentication mechanism.

2.1.2 Something you have

In this case the user possesses a unique piece of hardware that can be matched to his identity. Such hardware can be keys, tokens, smart cards, SIM cards, etc. When using something you have as the authentication mechanism the user does no longer have to

[1] Everywhere in this report where he/his is used you can also read she/her.

remember long and difficult passwords, he only needs to remember to bring his piece of hardware. In order for an attacker to gain access he will now have to copy or steal this piece of hardware. This is in most cases more difficult than guessing passwords. The disadvantage with this authentication factor is that it can be very expensive to create not only the hardware items, but also the equipment that shall verify these items. It is also important to take required actions whenever the hardware is either lost or stolen [1].

2.1.3 Something you are

This last property which utilize biometric properties has become an area of growing interest the last decade. Most biometric features are unique per person and they are found in almost all people in some way or another. The most known example of a biometric feature is perhaps a fingerprint. Even among identical twins fingerprints are unique. For an attacker it is even harder, but not impossible, to steal and utilize another persons biometric feature. The difficulties for an attacker depends on what kind of biometric property is used and how it is used [1]. As mention in Chapter 1, human biometrics can be classified into two types, not necessarily disjoint sets [2]:

- Physiological: Based on stable physiological characteristics, e.g. fingerprints, iris, retina, etc.
- Behavioral: Uses learned and alterable behavioral characteristics, e.g. keystroke dynamics, signature, gait, etc.

2.1.4 Combination of authentication factors

In order to increase the security of systems, they often use more than one authentication mechanism. Systems that utilize this are often called multimodal systems. One example of this is when you use your banking card. In most cases you have to use the card (something you have) with a PIN code (something you know). In cases where you utilize the VISA, MasterCard, etc. feature on your card you have to use your signature instead of the PIN code, thus being a “have” + “are” system. When using a multimodal system it is however important that all the features actually are used and needed for the authentication.

2.2 Biometric

As mentioned above biometric identification has been around since the dawn of man, people have always recognized others by their biometric properties such as face and voice. Even in Shakespeare’s play “The Tempest” a hint towards gait a a biometric feature was given: “*Great Juno comes; I know her by her gait*”. Human fingerprints have been discovered on several archaeological artifacts and historical items, but it was not until the late 16th century that fingerprint became a science [19]. In 1809 Thomas Bewick, an English wood graver, started to use his fingerprint as his trademark, this is believed to be one of the major milestones of the scientific study of fingerprints identification. Many researchers contributed with their study on the fingerprints during these years, and in 1846 Nehemiah Grew published the first scientific paper where he described his systematic study on the ridge, valley and pore structure in fingerprints. And in the 1880’s Faulds, Herschel and Galton continued their work on fingerprint recognition. Around 1870 Alphonse Bertillon described a system of body measurements for identifying people which was used until the 1920’s in the USA to identify prisoners [1]. For a long time fingerprint was almost the only used biometric feature to authenticate people, it was not until the 1980's that features like hand geometry, voice, signature and retina recognition became popular. Commercial face and iris recognition has been around since the 1990's and to use gait to recognize people has only been an area of research the last decade.

2.2.1 Biometric characteristics

As we realize there are a lot of different biological measurements that one can use in order to identify a human. However not all aspects can be used, according to [2] there are some properties that must be present in order for the measurement to be practical:

- Universality: Each person should have the characteristic.
- Distinctiveness: Any two persons should be sufficiently different in terms of the characteristic.
- Permanence: The characteristic should be sufficiently invariant over a period of time.
- Collectability: The characteristic can be measured quantitatively.

These four properties are the most important ones in order to make sure that people can use the system and be distinguished from each other. In a practical biometric system there are however some more properties that must be considered:

- Performance: Refers to the achievable recognition accuracy and speed, the resources required to achieve the desired recognition accuracy and speed, as well as the operational and environmental factors that affect the accuracy and speed.
- Acceptability: Indicates the extent to which people are willing to accept the use of a particular biometric characteristic in their daily lives.
- Circumvention: Reflects how easily the system can be fooled using fraudulent methods

A practical biometric system should have the desired accuracy, speed and resource requirements, be harmless and accepted by the users and be properly secure against possible impostors.

2.2.2 Comparison

It is impossible to say that one biometric feature is better than another, which biometric feature one shall use for the system depends entirely on the situation and user demands. The different biometrics can however be more or less classified by using the characteristics of biometrics described above. In Table 1 this classification is done for some biometrics, the values are ranging from high to low (where high is best, except for circumvention where low is the best).

2.2.3 Biometric system

A biometric system is a system that recognizes patterns from biometric data that has been acquired from an individual. From the data acquired it extracts a feature set and compares this feature set against a template set in the database. Biometrics can, like passwords and tokens, be used for both identification and authentication, which is also known as verification. Identification and authentication are two concepts which are used regularly in the world of biometrics. Both terms are used to declare the identity of an individual, but since the terms identification and authentication are often mixed up,

Biometric Feature	Univ	Dist	Perm	Coll	Perf	Acce	Circ
DNA	H	H	H	L	H	L	L
Ear	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Facial Thermogram	H	H	L	H	M	H	L
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand Geometry	M	M	M	H	M	M	M
Hand Vein	M	M	M	M	M	M	L
Iris	H	H	H	M	H	L	L
Keystroke	L	L	L	M	L	M	M
Odor	H	H	H	L	L	M	L
Palmprint	M	H	H	M	H	M	M
Retina	H	H	M	L	H	L	L
Signature	L	L	L	H	L	H	H
Voice	M	L	L	M	L	H	H

Table 1: Comparison of biometric features (from [2])

definitions are given below:

- Identification: When the system is running in an identification mode, an individual is recognized by comparing with an entire database of templates in order to find a match (Who am I?), see Figure 1. Accordingly the system conducts one-to-many comparisons to establish the identity of the individual. In the identification procedure the subject does not have to claim an identity [2].

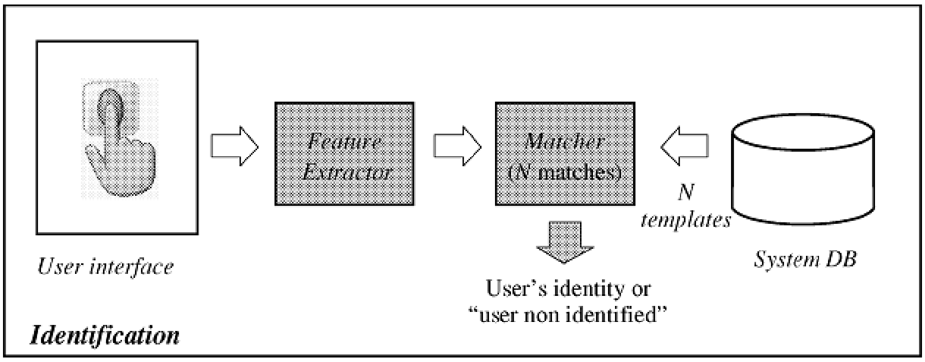


Figure 1: A block diagram showing the main components of a system running in identification mode (from [2]).

- Authentication: When the system is running in a authentication mode, the individual to be identified has to claim its identity and this template is compared against the individuals biometric characteristics (Am I who I claim to be?), see Figure 2. The system accordingly conducts one-to-one comparison in order to establish the identity of the individual [2].

- Template: Most biometric systems do not store raw biometric data in its database, partially due to legal aspects, but also because it can be unpractical. Systems rather extract a salient set of features, which is known as a template, from the biometric data of a user. Since the template, by definition, is a compact description of the biometric sample, it is not expected to reveal significant information about the original data [20].

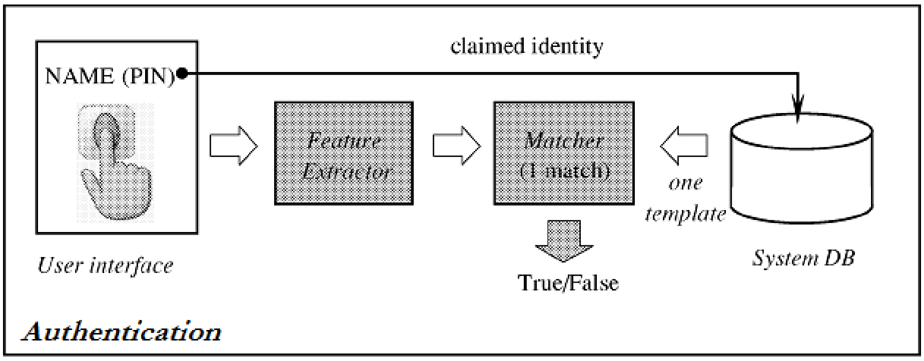


Figure 2: A block diagram showing the main components of a system running in authentication mode (from [2]).

- Enrollment: Before both identification and verification can occur, a template containing the biometric data about the individual has to be stored in the system, see Figure 3. This biometric data is the information the user must provide when he is going to identify or authenticate himself later.

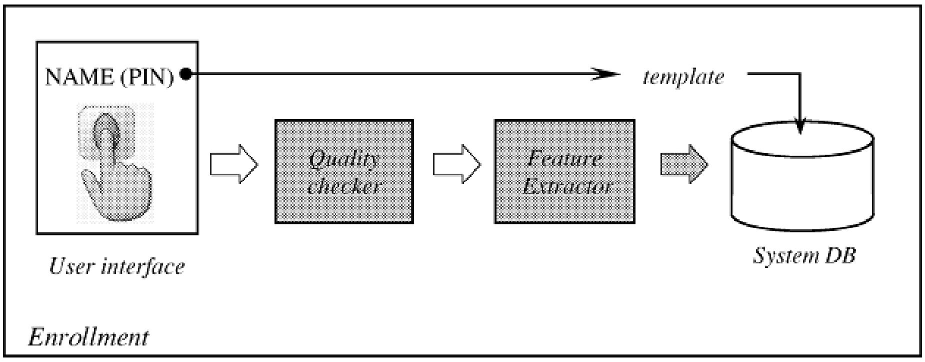


Figure 3: A block diagram showing the main components of a system running in enrollment mode (from [2]).

2.2.4 Results from the authentication process

When you are authenticating a person he will either be accepted or rejected. This is rather simple when you are using a knowledge-based method: either you know the required password or you do not know it. With something you have this is also trivial: either e.g. the key fits or it does not fit. When using biometric authentication this is not that simple due to the fact that biometric features can never match 100%. In e.g. a fingerprint we will find features that matches the template and features that do not matches. The more matches, the more convinced we are that it is the correct person. The similiarity between the extracted sample and the template are given by a matching score. Such matching scores are calculated by the use of a distance metric, e.g. the absolute distance between corresponding points in two sets. The distance metric should in principal give a small *intra-class* distance, meaning that samples from the same person get a low score, and a large *inter-class* distance, meaning that samples from different persons should give a high score. Whether a person is accepted or not depends on a threshold we set for the system and here lies an important property of any biometric system. A biometric verification process makes two types of important errors, see Figure 4 [2, 21]:

- False Acceptance Rate (FAR) is calculated from the False Match Rate (FMR). This happens when a biometric system measures two different persons to be the same person. A consequence would be that impostors wrongly would be granted access.
- False Rejection Rate (FRR) is calculated from the False Non Match Rate (FNMR). This happens when a biometric system measures two different measurements from the same person to be from different persons. A consequence would be that a genuine user wrongly would not be granted access.

A biometric system can also produce some other errors such as Failure to Enroll Rate (FER), this increase when the biometric feature on the person are not good enough to be extracted and create a biometric feature, and Failure to Capture Rate (FCR), related to the probability that the device capturing biometric data is not able to capture the required information.

The trade off between FMR and FNMR can be illustrated by the use of a Receiver Operating Characteristics (ROC) or Decision Error Tradeoff (DET), see Figure 5. Both curves shows the system performance at different threshold values and the tradeoff between FAR/FMR against FRR/FNMR, the equations for FMR and FNMR is listed in Equation 2.1 and 2.2. There are mainly two differences between ROC and DET curves. First, DET graphs plot false negatives on the Y axis instead of true positives, second DET graphs are log scaled on both axes so that the area of the lower left part of the curve is expanded. Another detail is that ROC curves sometimes plots FMR against (1-FNMR) [22]. The threshold one should use heavily depends on the application. E.g. high security applications would preferably want as low FAR/FMR as possible in order to not let impostors gain access. In the other end we do have forensic applications, which works with negative recognition (FAR = FNMR and FRR = FMR), where it is acceptable to have a higher FMR in order to be sure to catch the criminal. Most civilian applications are somewhere in between. Another definition that is commonly used is Equal Error Rate (EER). EER is a very commonly rate used to compare different systems against each other and can give a very briefly idea of how good the system is. But the total accuracy of your system depends on much more.

$$\mathrm{FMR} = \frac{\text{Number of accepted impostor attempts}}{\text{Total number of impostor attempts}} \tag{2.1}$$

$$\mathrm{FNMR} = \frac{\text{Number of rejected genuine attempts}}{\text{Total number of genuine attempts}} \tag{2.2}$$

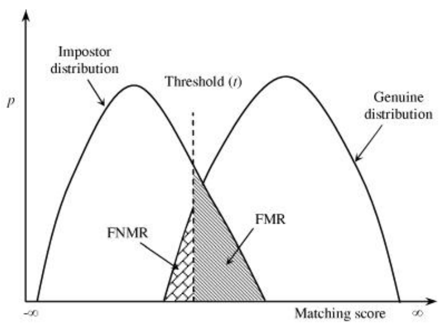


Figure 4: FMR and FNMR for a given threshold are displayed over the genuine and impostor score distribution (from [2]).

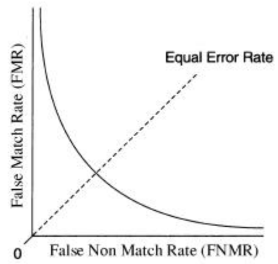


Figure 5: A Decision Error Trade-off (DET) curve showing the trade off between FMR and FNMR (from [21]).

3 Related work

In this chapter we will look further into the biometric feature gait and related work to this feature.

3.1 Gait features

Gait is defined as “a manner of walking” in Webster’s New Collegiate Dictionary. However, human gait is more than that: *“it is an idiosyncratic feature of a person that is determined by, among other things, an individual’s weight, limb length, footwear, and posture combined with characteristic motion. Hence, gait can be used as a biometric measure to recognize known persons and classify unknown subjects”* [23]. Already in 1905 the first studies of the human gait was published [24], Marks, a salesman of prosthetic legs, described how the process of walking could be divided into different phases and looked at how the prosthetic legs would affect an amputee’s gait.

Today, a human gait cycle is defined as the period from an initial contact of one foot to the following initial contact on the same foot [25, 26]. In Figure 6 we show that this cycle can be divided further into three main tasks which again is possible to divide into eight phases. The first task is a weight acceptance period where we have the initial contact phase and a loading response phase. During this task one foot is placed on the ground and the body weight is shifted in order to maintain stability and absorbing shock. The second task, which is a single limb support task, consist of a midstance phase, a terminal stance phase and a transition to the preswing phase. During this task the contralateral foot is swung forward while the body weight is maintained on the stable foot. The final task is the limb advancement task which consists of the preswing phase, the initial swing phase, the midswing phase and the terminal swing phase. During this task the previously stable foot leaves the ground, the body is shifted forward and then a new cycle can begin [27].

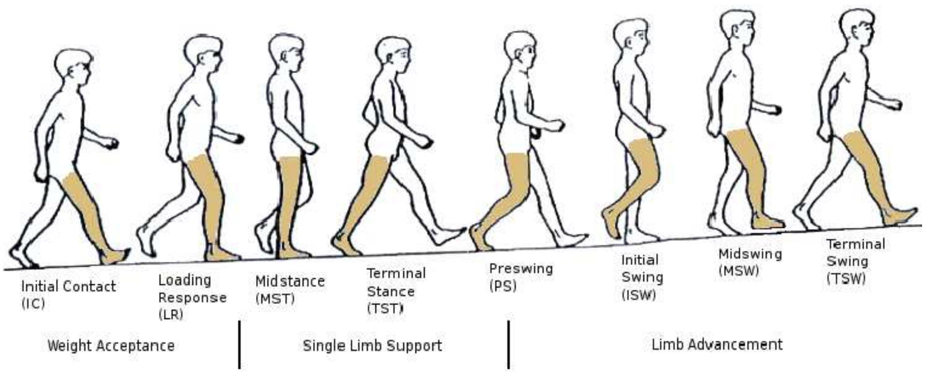


Figure 6: A complete gait cycle showing its three tasks and eight phases (from [27]).

3.2 State of art

The ability to use gait as a method to recognize people has been known for a long time. The earliest research dates 40 years back, where studies from medicine [28] and psychology [29] presented evidence that human gait has distinctive patterns from which individuals can be identified. It is however just in the last decade that gait as a biometric feature has been introduced, and from a technical point of view the gait recognition can be categorized in three different classes. These classes are Machine Vision (MV) which uses a (video) camera to capture gait data and video/image processing to extract features. This method is often used in surveillance and forensics. Then we have Floor Sensors (FS) that use sensors installed in the floor that are able to measure gait features such as ground reaction force and heel-to-toe-ratio when a person walks on them. This method is usually used in identification. Finally we have Wearable Sensors (WS) where the gait data is collected using body-worn sensors and this method is normally used to authenticate a person.

3.2.1 Machine vision

Most of the research on gait has been utilizing machine vision techniques to extract gait patterns. The first attempt of gait analysis done automatically was performed in 1994 by Niyogi and Adelson [30]. They used changes in a 2-dimensional video footage of a walking person to recognize persons. There have been published many reports which use different techniques and methods, for instance: Lee et al. [23] localized image features that would fit different parts of the binary silhouette of a person. Yoo et al. [31] generated gait signatures by computer vision and extracted kinematic features in order to recognize people. He also proposed a new method for extracting the body points by topological analysis and linear regression guided by anatomical knowledge. Canado et al. [32] extracted the gait data by using the movement from the thighs to fit to an articulated pendulum-like motion model. Little and Boyd [33] developed a description of instantaneous motion, that varies with the type of moving figure and the type of motion, and used that to recognize individuals by their gait. The perhaps most popular method which has been used by e.g. Kale et al. [34] and Wang et al. [35], is to extract the human silhouette from the sequence and use it as the feature of gait, see an example in Figure 7. In 2004 Liu et al. [8] came with the "*Simplest representation yet for gait recognition*", which use an averaged silhouette method. There has even been introduced a HumanID Gait Challenge Problem[1]: *"Identification of people from gait has become a challenge problem in computer vision. However, the conditions under which the problem is "solvable" are not understood or characterized"* [7]. The challenge problem consists of a baseline algorithm, a set of 12 experiments with various different modifications, and a large data set. With video based gait authentication the purpose has mainly been surveillance, for instance recognizing a criminal from a security camera video [36]. In such cases where other biometrics are obscured, for instance a criminal might conceal his face, the gait-signature will normally be present since it is more difficult to conceal/disguise the walking manner.

3.2.2 Floor sensors

Analyzing gait by the use of floor sensors has been commonly used by physiologists. Pathological gait can be a key factor in order to indicate several age related diseases such

[1] HumanID Gait Challenge Problem: http://www.GaitChallenge.org

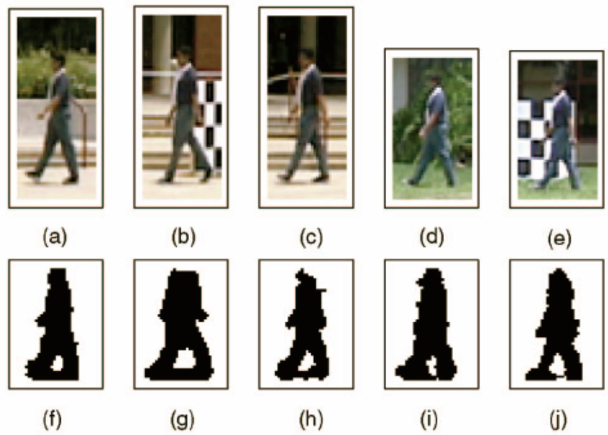


Figure 7: The bottom row ((f)-(j)) shows silhouette frames, the raw image corresponding to each silhouette is shown on the top row ((a)-(e)) (from [7]).

as Diabetic Polyneuropathy [37]. In addition to medical use, floor based sensors have mostly been used to track individuals. There are not many floor sensing systems which are designed with gait recognition in mind, those who have been used often have insufficient resolution, too low sample rates, or are too costly [38]. However, there has been some research which does look promising. Orr et al. [39] introduced floor system that may be used to transparently identify users in their everyday living and working environments. The report explains that a user footstep model based on footstep profile features has been used. They used a technique called Ground Reaction Force (GRF), which is highly related to Newton's third law that states that "for every action there is an equal and opposite reaction". So the GRF is the reaction that a measuring device produces in response to the weight and inertia of a body in contact with that device. Orr measured the GRF of the walker's foot as he walked over a measuring tile. With this method the report showed for instance that the effect of footwear is negligible on recognition accuracy. Middleton et al. [38] arranged an experiment with over 1500 individual sensors each with a range of 3 cm^2. The method he used extracted three features: stride length, stride cadence, and time on toe to time on heel ratio. Stride length and cadence are regular features which have been used in machine vision based recognition. The third feature however, are new to this analysis. Figure 8 shows a graphical illustration of the gait cycle and Figure 9 shows the data collected. Floor based sensors could eventually find deployment as a standalone system (e.g. a burglar alarm system) or as part of a multimodal biometric system.

3.2.3 Wearable sensors

This last category uses one or more wearable sensors in order to acquire gait data. Gait based analysis has previously been used for medical reasons in order to help patients with movement problems [40]. This method to authenticate people is however very new and was first introduced by a research group from VTT Electronics located in Finland in 2005 [10]. The wearable sensor, also called accelerometer, records the signal characteristics produced by walking. The accelerometer captures movement, most commonly, in three directions (horizontal, vertical and lateral) and the recognition is performed by

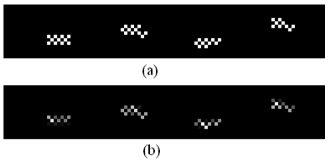


Figure 8: Typical gait cycle over the sensor mat. (a) shows foot steps recognized, while (b) shows the time spent at each location belonging to the steps recognized in (a), the higher intensity the pixels have the longer the person has stayed on that location (from [38]).

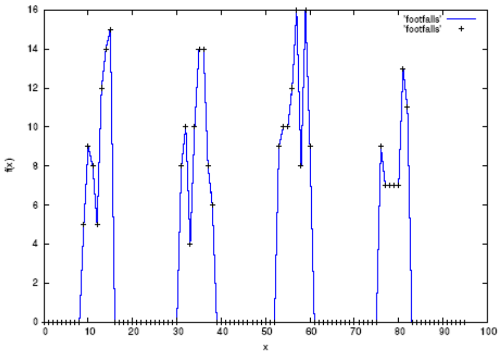


Figure 9: The profile of the 4 footsteps on the sensor mat, showing heel and toe strikes. (from [38]).

processing these signals. Even though this method is very new, there have been some different approaches on both the placement of the sensor(s) and the recognition method. Mäntyjärvi et al. [10] wore the sensor on the belt and used correlation, frequency domain and histogram statistics in the processing phase, Vildjiounaite et al. [18] tried to place the sensor in three different places: hip pocket, breast pocket and in the hand while carrying a suitcase. At Gjøvik University College there have been some approaches where they have tried to place the sensor on the ankle [12], by the hip [41] and in the pocket [11] and used methods called absolute distance, histogram similarity and two different cycle lengths. Rong et al. [42] used a method they called dynamic time warping to recognize the subjects. Figure 10 shows an example of the data collected by an accelerometer. Wearable sensors are mainly thought to be a part of a security module which by continuously authentication secures portable devices.

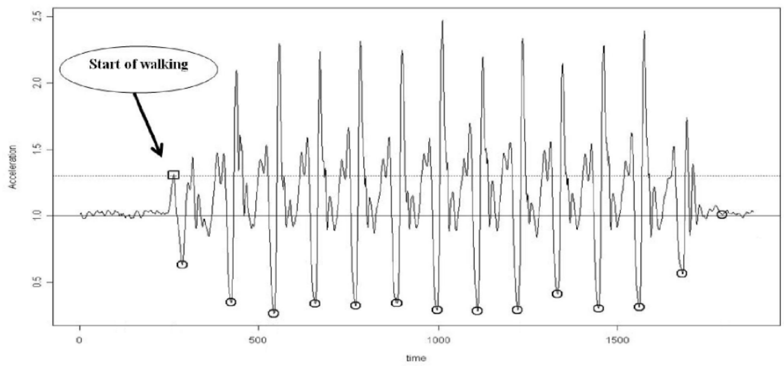


Figure 10: An example of the data collected by the accelerometer, the “start-of-walking” and local minimas are identified (from [11]).

3.2.4 Comparison

Gafurov et al. [11] gave a small comparison table, see Table 2, from some different approaches which use one of the three different methods (MV = Machine Vision, FS = Floor Sensors, WS = Wearable Sensors) in order to give a general overview. It is however important to say that *"this table by no means implies a direct comparison of the performances, mainly due to the differences between data sets"* [11]. The table shows how many subjects participated, the performance of the systems in terms of EER or recognition rate and what kind of gait recognition category that was used. In [9, 38, 39, 43] the performance is shown by the recognition rate, while the others used EER.

Study	S#	Performance	Catg.
BenAbdelkader et al. [44]	17	11	MV
Wang et al. [35]	20	8, 12, 14	MV
Wagg and Nixon et al. [43]	115	64, 84	MV
Orr and Abowd [39]	15	93	FS
Suutala and Roning [9]	11	65.8-70.2	FS
Middleton et al. [38]	15	80	FS
Ailisto et al. [45]	36	6.4	WS
Mäntyjärvi et al. [10]	36	7, 10, 18, 19	WS
Vildjiounaite et al. [18]	31	13.7-17.2	WS
Gafurov et al. [12]	21	5, 9	WS
Gafurov et al. [41]	22	16	WS
Gafurov et al. [11]	50	7.3, 9.2, 14, 20	WS
Gafurov et al. [46]	100	13	WS
Rong et al. [42]	35	6.7	WS

Table 2: A comparison table of different gait recognition approaches (mostly taken from [11]).

A huge advantage with FS and WS compared to MV is that one avoids the impact of external variables such as lighting and camera placement. In addition, the use of WS is rather cheap compared to MV which require expensive cameras and such [47]. Another difference among these categories is that while MV and FS usually identify people, WS is mainly used to authenticate users [46]. When a worn sensor is an integrated part of your mobile device it gives another advantage compared with many other biometric features, it does not only not require any sensing equipment in your infrastructure, but it is as well unobtrusive [10]. In connection with our project, where we look at different circumstances, FS would not be usable since it would require sensors installed on the floor. With MV we have the drawback of costly equipment and not to mention all the image processing required. With WS however it is reasonable to believe that we more precisely will detect different characteristics of the various circumstances. In addition will the preprocessing itself be easier since we are processing signals instead of images/video.

3.3 Related work to research questions

In this section we will go more in the depth of gait recognition and specifically WS-based.

3.3.1 Data acquisition

There are, as mentioned in the previous section, many different ways to place the sensor when using WS. The most common place is either by the ankle or placed on the belt [10, 12, 42, 46]. By firmly attaching the sensor to the belt or ankle one does not have to worry that the accelerometer may move or rotate while walking, this movement could have a huge impact on the gait data. Other placements have been pocket [11, 18], breast pocket [18] and in the hand while carrying a suitcase [18]. The results from previous reports do not give any particular reasons to favor one placement over the other. By choosing the ankle one will get data with higher variations as there is more acceleration and movement in the lower part of your leg. From an application point of view however is it most natural to choose the belt or the pocket since this is where people often carry their mobile devices.

Mostly there has only been one accelerometer attached to the subject, but it could be possible to e.g place one sensor above the knee and one below in order to capture knee movement or have both one at the ankle and the hip, etc.

As one realize there is a lot of possibilities, in this project however it is most practical to use one sensor placed on the hip. The reason for this is that it is a lot easier to ask participants wear a belt with a sensor attached, than strapping sensors to their ankles that could hamper their natural gait. In addition, as mentioned above, a possible area of application would be to integrate a sensor in a mobile device that often are carried on the belt.

3.3.2 Do people walk in the same way given the same circumstances?

Even if all environmental circumstances are the same, there might be other factors that lead to different gait results. First of all when using an accelerometer to collect data we will get a signal with periodic cycles. From these cycles we can see that no cycle is identical, there will always be minor differences. This is not a unique problem with gait, but a common property of biometrics in general, and especially behavioral. Since we are human we will never walk, talk or write in the exact same way every time, there will always be some minor differences. But when looking at gait, the general shape of each cycle will usually be similar. There might be other factors that alter the human gait more dramatically. An important factor which needs to be considered is whether humans walk in the same way under the same conditions after e.g. a week. Most of the research does not check whether the time factor affects the result, all data is collected during a short time span, Vildjiounaite et al. [18] however did choose a larger time period (one month) between collecting the training data and test data. The results from their research was that the ability to recognize a person mainly depended on the walking speed. It was easier to recognize a person when he was walking fast, than when he was walking slowly. In addition some subjects simply do have a more stable way of walking even if they change shoes, including changing from shoes with a flat sole to shoes with high-heels, while others have less stable walking even when using the same shoes. To overcome these problems Vildjiounaite et al. mentioned a larger training set and sample group was needed. Boyd et al. [48] has listed some other different factors that might have an impact on how a subject walks.

- Injury: This is one of the main drawbacks with gait authentication, an injured person walks entirely different than if he is not injured. The injury does not have to be severe, it could e.g. be enough that he got a light sprain or strain from a football game the day before. Another impact which will have the same consequence is drunkenness.
- Fatigue: Another factor that most likely will affect your gait is how fatigued you are. There might also be a difference whether you are walking in the morning or in the evening even if you do not feel any difference in your shape.
- Muscle development and training: Boyd et al. [48] also lists that training and developing muscles might alter how you walk. An example on such training would be the marching in the military.

- Personal idiosyncrasies and cultural artifacts: The article also mentions that these two factors might affect the result. Some people walk in a very special way, e.g people who are proud or very self confident might walk different than people that are tired or more shy.

All the factors mentioned above will be taken into consideration and avoided in the best possible way in this project.

3.3.3 Analyzing gait signals

In this section we will only look at how gait can be analyzed when using WS, since FS and MV-based analysis is entirely different than WS-based analysis. There exist various different methods in order to authenticate and/or identify users from gait data:

- Correlation have been used in e.g. [10, 11, 18, 45]. In general correlation is an indication of the strength and direction of a linear relationship between two random variables. Mäntyjärvi et al. [10] applied correlation in the following way:
  1. In the training phase, divide the acceleration signals into one step long parts by searching for local minimums and maximums, and since right and left steps not necessarily are symmetrical they are processed separately as a and b steps.
  2. Normalize all steps both in length and in amplitude.
  3. Average a and b steps in order to create templates for them.
  4. In the enrollment phase, the steps above are repeated forming c and d steps.
  5. Finally the correlation is calculated by the following formula: $C = Max((corr(a, c) + corr(b, d)), (corr(a, d) + corr(b, c)))$
- Frequency domain was used in [10, 18, 45]. The idea behind frequency domain is that while a normal time domain graph shows how a signal changes over time, a frequency domain graph shows how much of the signal lies within each given frequency band over a range of frequencies. In [18] they used Fast Fourier Transformation (FFT) coefficients for recognition of gait patterns. *"The coefficients were calculated in a 256-sample window with a 100 sample overlap. The 128 FFT coefficients of each training file were clustered with K-means algorithm into eight clusters. The FFT gait score was produced by finding the minimum distance of the test data FFT coefficients from the trained clusters"* [18].
- Histogram similarity has been used in [10, 11, 12] in addition has the method been successfully used in previous master theses [27, 49], and mainly consist of the following three steps:
  1. Compute n-bin histogram of the combined signal from the accelerometer.
  2. Normalize the histogram by number of recorded samples.
  3. Use a distance metric to compute the distance between two histograms.
- Average cycle length used in [12, 41] will be more thoroughly explained in Chapter 4. The basic idea is to identify cycles in the signal and create an average cycle based on these steps. The score is computed by comparing these average cycles.
- High order moments were used in [11, 10]. High order moments such as skewness

(third moment) and kurtosis (fourth moment) describe the degree of symmetry in the variable distribution and the relative peakedness/flatness of a distribution respectively.

- Dynamic Time Warping (DTW): Rong et al. [42] used an algorithm they called DTW. In general DTW disposes the naturally occurring changes in walking speed and are able to compare signals of different lengths and where the x-axis are shifted in one way, DTW will be explained in Chapter 4. Rong et al. variant of DTW was to normalize the gait cycles, so that the step length is equal and thus more comparable. This was done in the following way: *"The number of points in the input layer is* n*, which represent* n *sampling data of a user's gait cycle, we denote as* $A_k^0(k = 1, 2, ..., n)$*. After one step transfer, the nearest two continue sampling data are incorporated, and the others remain their old values, so there are* $n - 1$ *nodes in the first layer. The rest may be deduced by analogy. After* $n - N$ *steps combination, there exit* N *nodes in the output layer"* [42].

What method is best depends both on how the data is obtained, meaning e.g what sampling frequency the sensor has, and the placement of the sensor. Another important factor is how the method is used, meaning how e.g the pre-processing is done and which distance metric that is used. For example, in Vildjiounaites et al. report [18] correlation was better than frequency domain when the sensor was in in a breast- and hip pocket, but when the sensor was in the hand, frequency domain was better. By reviewing the articles mentioned above it can seem that average cycle, DTW, correlation and histogram similarity provide the best results. The pre-processing seems to share many common features. In [11, 12, 41, 46], as with many others, the output of the sensor is first transformed to obtain acceleration of units *g*, then the resulting acceleration is computed. After that the signal is interpolated to equalize the time between the signals and a moving average filter to reduce noise is applied. Rong et al. [42] used Daubechies wavelet of order 8 to effectively remove noise from their signal.

3.3.4 Recognizing persons under different circumstances

With accelerometer based gait authentication recognition of persons under different circumstances has not been well researched, some circumstances that one finds in a normal day is change of:

- Carrying: Gafurov et al. [11] looked into what the impact of carrying a backpack weighing 4kg had, the result was just a slight fall of the performance. The EER went from 7,3% under normal conditions to 9,3% with the backpack.
- Footwear: The change of footwear has to a small extent been researched. Vildjiounaite et al. [18] wrote that change in footwear was person-dependent, in some cases they were able to recognize the subject even if different shoes were used, but in some other cases this was not possible. Most of the research with different footwear has been with machine vision gait recognition, a change of footwear is among the dataset from the HumanID Gait Challenge [7], Liu et al.[8] has shown that the recognition rate drops from 80% to 54% when using this data set. It has also been shown that muscle activation in walks changes when people walk bare footed as opposed to wearing shoes [50].
- Terrain and surface: By changing the terrain and surface a subject will most likely

change his way of walking, this surface change can e.g. be from a concrete floor to a gravel. It is likely to believe that for instance walking indoor is completely different than walking outside on the gravel. Another possibility could be that the surface is sloping or a subject is walking stairs. Again the HumanID Gait Challange set [7] contains data set from two different surfaces, concrete and grass. They showed that the identification rate went dramatically down when changing surface.

- Speed: People do not walk with the same speed throughout the day, and by walking at a different speed, the accelerometer will produce different gait data. This fact is a common problem mentioned in most of the relevant articles [10, 47, 51] among others, but again it is only in MV change of speed has been researched. In [3] they got almost perfect identification rate (96%) when comparing different speed. But again, comparing WS with MV is not fair, it is an entirely different process.
- Direction: By changing the direction, e.g taking a turn, it is likely to believe that the gait data will be different compared to only walking in a straight line. To our knowledge this has yet not been researched.

There does not exist any research on how to compare different circumstances against each other and what common features the different circumstances may have. It could for example be likely that when you are walking fast the signal will produce not only shorter cycle lengths but also higher amplitudes, and by stretching and dampening this signal, one could adapt the data to become similar to that subjects normal gait data.

4 Sensor and processing details

In this chapter we will explain the technology used to capture gait features and how these features can be analyzed.

4.1 Technology

In order to acquire acceleration data we used an accelerometer called Motion Recording 100 (MR100), developed at Gjøvik University College, see Figure 11. *"MR100 is a prototype of our motion recording technology. In consists of three sets of three-axis accelerometers as well as a motion detection sensor. It is small - not much larger than a mp3-player and is equipped with a storage unit capable of storing 64MB of acceleration data. It has both a USB and a Bluetooth-interface, which makes it possible to transfer the data to either a computer, a cellular phone or a PDA."* [52] The main components of the MR sensor are three 7260 accelerometers from Freescale[1], a PIC18F4550 MicroController Unit (MCU) from Microchip[2], 64-MB memory for storing acceleration data, USB interface for data transfer and a battery. The sampling frequency of the MR sensor was about 100 samples per second and its dynamic range was between -6g and +6g ($g = 9.8m/s^2$). If we have no movement, and the direction of the sensor is as shown in Figure 11, X and Z will be equal to 0 while Y will be -1g since the only influence is the gravity force.

Figure 11: The MR100 sensor, with the three directions noted.

In order to transfer the data from the sensor to a computer the MR Analyser [52] software was used. The data collected from this sensor are acceleration values in X,Y and Z direction with a corresponding timestamp in addition to some metadata, an excerpt is shown in Figure 12. In Figure 13(a) one can both see how a cycle is represented in X and Y direction and the resultant, R (see Equation 4.4). From the figure we can see that both X and Y, and consequently R ,is cyclic repeated. The Z values are omitted from the figure

[1]Freescale: http://www.freescale.com/
[2]Microchip: http://www.microchip.com/

as these values represent sideways movement and do not have a clear cyclic repetition as X and Y. If one looks closely at the cycles one can see that R is almost a copy of -Y. The reason for this is that when the accelerometer is located on the hip both X and Z values usually lies between -0.5 and 0.5g, while the Y value is around -0.5 and -2.5g. In Figure 13(b) we can see the corresponding foot movement.

```
<<< BEGIN HEADER >>>
Start time: 29/01/2008-13:50:57
Stop time: 29/01/2008-13:53:20
Sample interval: 0.018090
Sample count: 12581
Raw data file: 29012008135057.raw
Sensor type: MR100 Sensor
Sensor ID: 1005
MR Analyser version: 3, 0, 6, 0
<<< END HEADER >>>
0.000000 539 473 556
0.009988 544 471 561
0.019883 548 479 553
0.029949 547 476 558
0.039218 544 480 555
0.049149 546 475 557
0.059267 548 478 555
0.060550 546 467 553
0.070469 547 473 556
        ...
```

Figure 12: An excerpt of the raw data as it is stored.

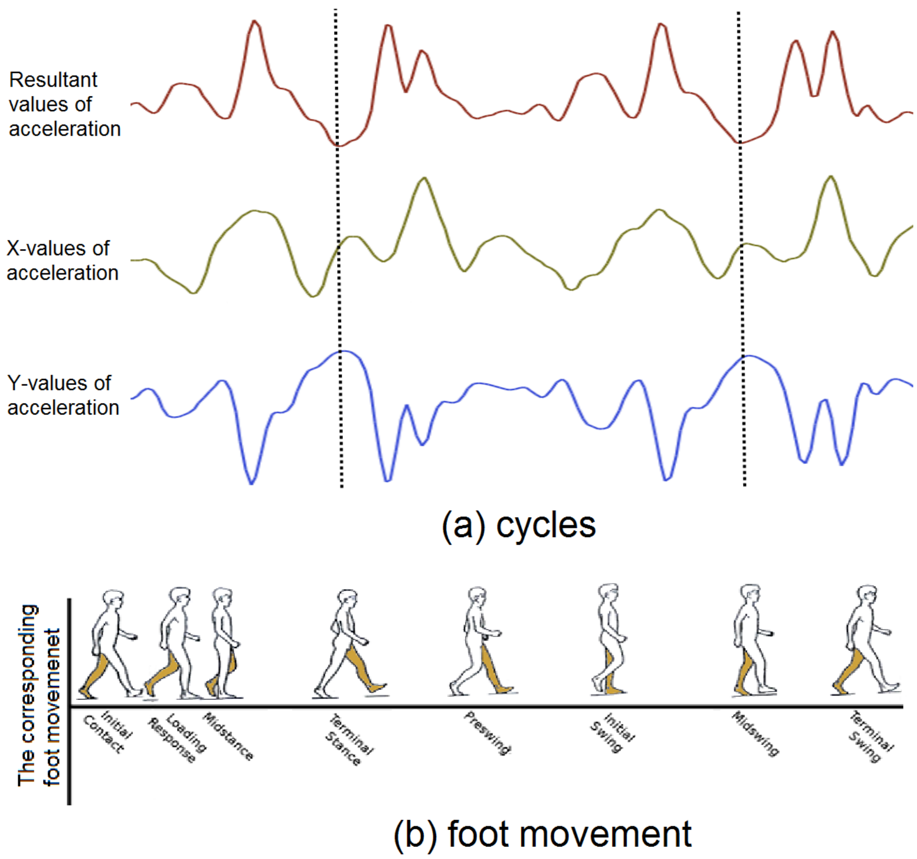


Figure 13: (a) An illustration showing the cyclic repeating of gait cycles in R, X and Y acceleration. (b) The actual foot movement which was described in Section 3.1 (the foot movement illustration is taken from [27]).

In the following sections will we go into the depth of the pre-processing algorithms. All algorithms have been created from scratch and are implemented in a software written in Java. In Appendix B one can psuedocode of some important functions. With this software we do not only have the possibilities to tweak all aspects of the analysis, we also have the possibility to manually detect steps. This manual detection, which we will come back to in Chapter 6, can aid us to get a better understanding of the creation of templates and thus lead to improved algorithms.

4.2 Pre-processing and analysis

There are almost endless different ways to process and analyze the raw data that the sensor produces. We will first look briefly at the main phases involved in most of the methods and next we will look into those phases and the algorithms we use in this project in more detail.

4.2.1 Overcome weaknesses with the sensor

Time interpolation: A shortcoming of the sensor is that it does not exactly record one sample each $\frac{1}{100}$ second. We must therefore do some interpolation in order to get a sample every $\frac{1}{100}$ second.
Noise reduction: Another weakness with the sensor is the fact that the data will also contain some noise. This noise must be dealt with in the best way possible.

4.2.2 Conversion to g-force and creation of resultant vector

The raw data does not contain g-force values. The recorded values must be converted by using the properties of the sensor in order to achieve values of g. After that has been done we can create the resultant vector from the converted values, either from only one, two or from all three directions. There is also a possibility to just use the raw data as they are and not create a resultant vector. So as one realizes, already at this stage we have several options that would lead to different results.

4.2.3 Step detection

An important phase is to detect where a step starts and ends. There are different ways to split up a signal into periodic cycles, one can choose to split the signal into singular left and right steps [10, 18, 45] or take double steps [12, 41, 46, 51]. The most common way to split steps is to look at minimum points, see Figure 10 on Page 16, but there is nothing in the way of using e.g. maximum points.

4.2.4 Create template/input sample

After the steps have been identified one needs to create a template for enrollment or an input that shall be compared against a template. One simple way of doing this is to simply normalize all the steps to have equal length and take the mean or median from these steps. Another possibility is to not normalize the steplength and use Dynamic Time Warping (DTW) to find an average cycle that is representative for that set of steps.

4.3 Algorithm details

In this section we will go into the details of the different algorithms used in this project. We have used the average cycle method, used in [11, 41, 46], as a starting point. The basic idea with the average cycle algorithm is to obtain an average of all cycles identified in a walk and then compare this against other averaged cycles. In Figure 14 an illustration of a "correct" cycle is shown. The figure also indicate interesting points; A = start of the step, B = first local maximum, C = local minimum, D = last local maximum and E = end of the step. By local maximums for B and D we mean the maximum point between A-C and C-E respectively, while the local minimum (C) is the minimum point between B and D.

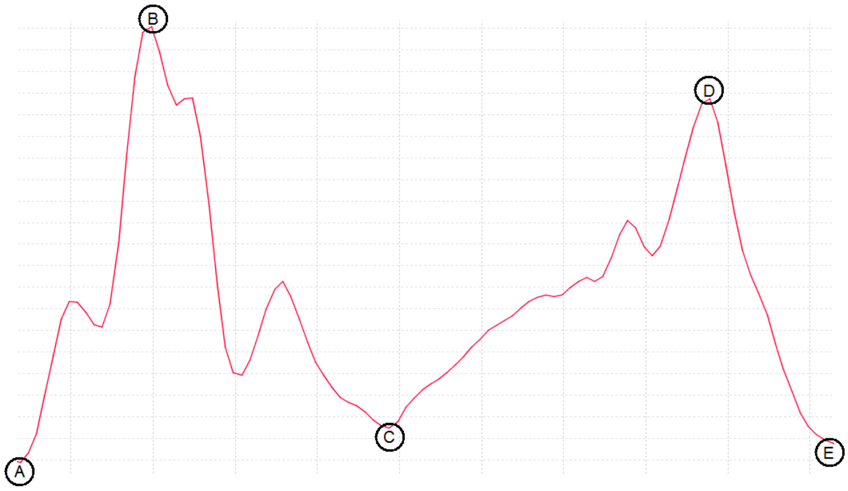


Figure 14: A "correct" cycle with annotations; A = start of the step, B = first maximum, C = local minimum, D = last maximum and E = end of the step.

4.3.1 Pre-processing

As mentioned in Section 4.2.1 we must perform some pre-processing in order to overcome some weaknesses with the sensor.

Time interpolation: This is done by using linear interpolation in time in order to obtain a value every $\frac{1}{100}$ second. If two known points are given by the coordinates (t_0, a_0) and (t_1, a_1), the linear interpolant uses the straight line between these points. For a value t in the interval (t_0, t_1), the corresponding value a on the straight line can be found from Equation 4.1. This is also illustrated in Figure 15.

$$\frac{a - a_0}{a_1 - a_0} = \frac{t - t_0}{t_1 - t_0} \Rightarrow a = a_0 + (t - t_0)\left(\frac{a_1 - a_0}{t_1 - t_0}\right) \tag{4.1}$$

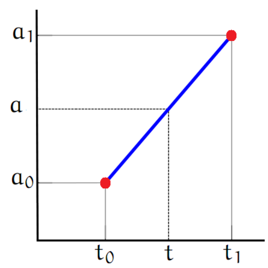


Figure 15: Given the two red points, the blue line is the linear interpolant between the points, and the value a at t may be found by linear interpolation.

So we apply this to all the values in the input we are processing. We start at $t = 0$ and the data samples are interpolated so we achieve our goal with a sample every $\frac{1}{100}$ second. If a following sample happens to have a value at the particular timestamp we desire we do of course not change that sample.

This way of interpolating data is the simplest and less computational expensive one. There do exist more sophisticated variants, for example polynomial and spline interpolation. Polynomial interpolation is in fact a generalization of linear interpolation. With linear interpolant we have a linear function, but the interpolant could be changed to a polynomial of higher degree. Instead of using a linear function for each of the intervals, one can use low-degree polynomials in each of the intervals. These polynomial pieces are chosen such that they fit smoothly together, the resulting function is called a spline, hence the name spline interpolation. For our purpose however the simplest form for linear interpolation is sufficient.

Noise reduction: In order to reduce the noise there exist different possibilities. In this project we have only looked at Moving Average (MA) and Weighted Moving Average (WMA) since these are both quick and simple to implement. In Figure 16 one can see an illustration of the difference between MA and WMA. As we see, the only difference between these averaging methods is that with WMA the closest neighbors are more important than those further away, while with MA all the neighbors have equal weight. The formulas for WMA and MA with a sliding window of size 5 is given in Equation 4.2-4.3. As with time interpolation this is applied to all the values in the input we are processing except the first and the last two when we using a sliding window of size 5.

There are also other options than can be used, we can use different window sizes. In [42] Daubechies wavelet of order 8 was used to remove noise. But we have chosen to test only WMA and MA with window size 5 in this project.

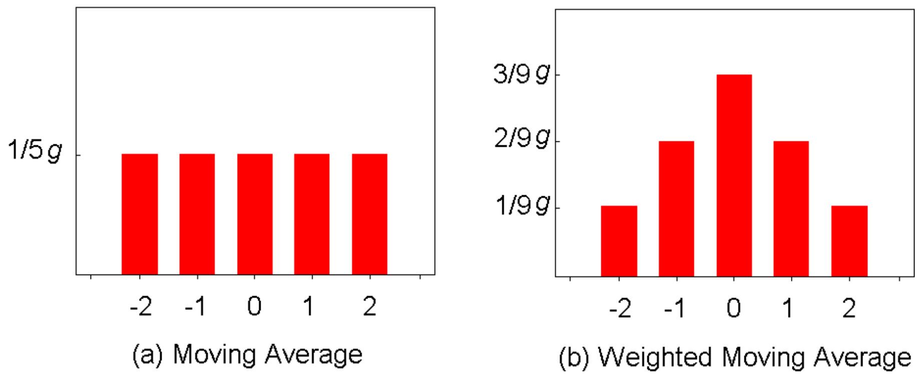


Figure 16: (a) Moving average, (b) Weighted Moving Average.

$$WMA\_a_t = \frac{(a_{t-2} * 1) + (a_{t-1} * 2) + (a_t * 3) + (a_{t+1} * 2) + (a_{t+2} * 1)}{9}, \quad (4.2)$$

where a_t is the acceleration-value in position t. The current value we are located at are given weight 3, the two closest neighbors weight 2 and the next two neighbors weight 1.

$$MA\_a_t = \frac{a_{t-2} + a_{t-1} + a_t + a_{t+1} + a_{t+2}}{5}, \quad (4.3)$$

where a_t is the acceleration-value in position t. All the four closest neighbors are given the same weight.

Conversion to g-force and creation of resultant vector: After the data values has been interpolated in time and the noise has been reduced we choose to convert the data values to g-force by using properties of the sensor. After we have converted the data values into units of g we can create a resultant vector for that data set. The Euclidean metric of the X-, Y- and Z-values is used to create a resultant value, see Equation 4.4. Other possibilities could have been to for example ignore Z-direction: $r_t = \sqrt{x_t^2 + y_t^2}$ or just use one direction $r_t = \sqrt{y_t^2} = |y_t|$ or $r_t = y_t$, etc. The reason we choose all three direction and not only Y, which as mentioned earlier in the report is the direction with highest impact on the resultant vector, is that a small scale experiment gave best results when all three directions was used.

$$r_t = \sqrt{x_t^2 + y_t^2 + z_t^2}, t = 1, ..., n \tag{4.4}$$

where r_t, x_t, y_t, and z_t are the magnitudes of resulting, vertical, horizontal, and lateral acceleration at time t, respectively, and n is the number of recorded observations in the signal.

4.3.2 Step detection

In this phase we detect the actual steps in the gait signals and in this project we have chosen that a cycle contains a double step and the steps are detected by looking at minimum points in the graph. The end of a step will automatically be the beginning of the next step (except for the final step). As our step detection algorithm consist of several sub-phases will we first give a short explanation of the phases involved and then look at the details. Furthermore, before we start explaining sections that are more difficult to comprehend, pseudocode for that particular phase will be given, a more detailed pseudocode can be seen in Appendix B.

Overview:

First we will give a very brief overview of what the step detection algortihm consist of.

1. **Estimate cyclelength:** In order to more precisely perform the actual step detection we need to first get an estimate of how long one cycle is. After reviewing the data we have collected we know that the cyclelength can range from [80..180] samples depending on stride cadence, meaning how fast a person walks. The faster a person walks the less number of samples a cycle will consist of. This estimation is done by first extracting a small subset of the collected data and then compare this subsets with other subsets of similar length. After we have compared the subsets in a given search area can we estimate a cyclelength based on the distance scores. This process is illustrated in Figure 17.

2. **An indication of minimum value:** In this phase we want to get an indication of the amplitude details, and especially the minimum value. The reason why we want this is to get an indication of what minimum values to expect when start with the actual step detection.

3. **Detect starting location:** The final phase before we start with the actual step detection is to decide where we are going to start. The intuitive approach to perform step detection would be to start at the beginning of the data collected, but for reasons explained in Section 7.1.2 we choose another approach. We start in the middle of the collected data and therefore we have to find the correct minimum point to start at.

4. **Detect the rest of the steps:** Now we are finally ready to start with the actual detection of the beginning and end of steps. This is done by first detect steps forward from the starting position detected in the previous phase, then repeat the process backwards.

Details:

In this section we will go deeper into the details of the algorithm.

1. **Estimate cyclelength:**

```
1. baseline = extract 70 samples from the middle of the collected data
2. REPEAT
3.     REPEAT
4.        comparison = extract 70 samples from the search area
5.        score = calculate distance between baseline and comparison
6.     UNTIL checked entire search area
7.     extract local minimas and calculate cyclelength
8. UNTIL both forward and backward direction are checked
9. γ = average of both calculated cyclelengths
```

Pseudocode 4.1: Pseudocode for estimating cyclelength.

The idea behind this estimation is that, suppose the cyclelength is K, then 2 subsets starting K positions apart would be alike, i.e. small distance between them. This estimation process consists of several sub-phases.

- Look at normal steps: This local estimation must of course be done where normal cycles occur, therefore we start in the middle of the collected data. Say the total samples of the session is N, then we start at sample number $\lfloor \frac{N}{2} \rfloor = n$.
- Extract a baseline: A subset of length 70 are then extracted by the following formula: $S_i = (X_{n+i}, X_{n+i+1}, ..., X_{n+i+69})$. So from position n we extract 70 samples, S_0, this subset will be the set which the other subsets compared against. The reason why we choose to extract 70 samples is that 70 samples would be sufficient for estimating cyclelengths for both fast and slow walking.
- Compare baseline against consecutive subsets: After the extraction, S_0 is compared with S_k, for $k = 70, ..., 370$. For each comparison we calculate the absolute distance between set S_0 and S_k. $d_k = d(S_0, Sk) = \sum_{i=1}^{70} |S_0[i] - S_k[i]|$. In other words do we compare S_0 with 300 consecutive subsets starting where S_0 ends. The reason to look 300 subsets ahead from $k = 70$ is to be sure that we detect both extreme types of walking, fast and slow. As mentioned earlier in this section do the cyclelength range from 80 to 180 samples. In order to detect both fast and slow walking in a worst case scenario, must k range from 80...360, to add some buffer do we increase this range by 10 samples in both parts. So if the cyclelength is 180 samples and we have a worst case scenario, the two minimum points we are after would come after $k = 180$ samples and $k = 180 + 180 = 360$ samples from n.
- Estimate cyclelength: After all the distances are calculated we only look at local minimums and the the cyclelength will be the distance between the two minimums with the lowest difference which are not too close or too far away from each other, meaning between [80..180] samples. The same procedure is repeated

by searching 300 samples backwards, and the estimated cyclelength, γ, will then be the average from these two cyclelengths. As we see is $\frac{180}{2} > 80$, so we might run into wrong estimations if do not apply additional checks in certain cases. This problem will be discussed more in Section 7.1.1.

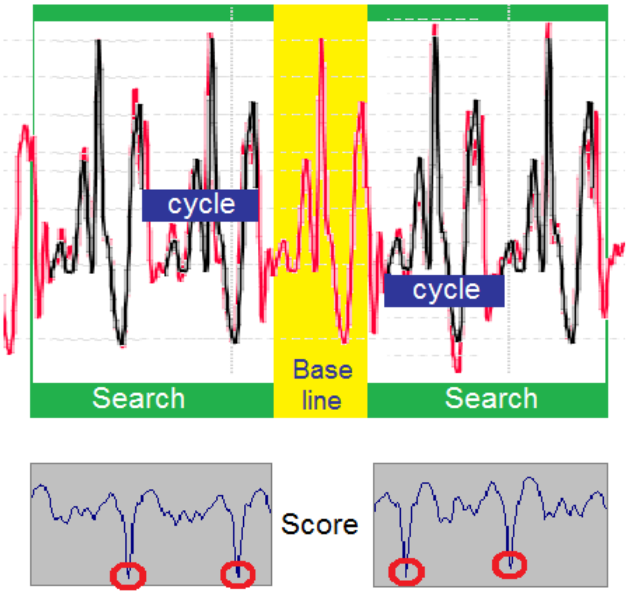


Figure 17: The yellow baseline area indicate the 70 samples that are extracted, the green area is the search area where the baseline is compared against a subset of the search area. The 4 black subgraphs are the baseline at those points it has the lowest distance with the search area subsets, and the difference between them (blue area) indicate the cyclelength.

2. **An indication of minimum value:** We again start in the middle of the collected data with sample X_n, as explained in the previous phase. We then search for the local minimum in one cyclelength area before and after this point, $\chi = \min[X_{n-\gamma} ... X_{n+\gamma}]$. Then we increase this value by 20%, $\alpha = \chi + 20\%$. The reason why we add 20% to the minimum value has two reasons. First, the amplitudes in the beginning and the end of a walk are usually smaller than the amplitudes in the middle of a walk (where the person has gained his normal speed). Secondly, if χ was an extreme minimum that does not resemble the average minimum of that walk we would not detect any steps.

3. **Detect starting location:**

```
DO
 M = get maximum point location that will be the basis of our test
 apply checks to decide if we shall go backward of forward from M
WHILE no decision can be made
S = starting location
```

Pseudocode 4.2: Pseudocode for detecting starting location.

We again start at the middle of the data collected, and search for the position, M, of the maximum point in one cyclelength area. Due to the shape of a correct gait cycle we now know that we are either just before or after our desired minimum, meaning point B or D in Figure 14. In Figure 18 we see that position M in this case is at point B. To decide whether we shall go backward or forward from this position we apply four checks, each equally weighted.

- Length between peaks: In this first test we look at the distance from M to the local maximum point in both direction in $\frac{2}{3}$ of one cyclelength. We name these local maximum points $M_{backward}$ and $M_{forward}$. The reason why we search less than γ samples is that we do not want to detect the maximum points that are one cyclelength away from M. If $|M - M_{backward}|$ is lower than $|M - M_{forward}|$ we give backward one point, forward otherwise.
- Value of possible minimum: In this test we look at the value of the possible minimum point, $S_{backward}$ and $S_{forward}$. These possible minimum points are the local minimum between M and the maximum points we found in the previous test. If the the value of $S_{backward}$ is smaller than the value of $S_{forward}$ we give backward one point, forward otherwise.
- Number of local maximum points: In this test we look at the number of local maximum points from S to the the following maximum point. If the area from $S_{backward}$ to M has fewer number of local maximum points than the area from $S_{forward}$ to $M_{forward}$ we give backward one point, forward otherwise.
- Distance to possible minimum: In this final test we look at the distance from M to the possible minimum point. Generally would we like to have a small distance between S and M, so if $|M - S_{backward}|$ is lower than $|M - S_{forward}|$ we give backward one point, forward otherwise. The exception of this is if this distance is very low (less than 10 samples) this is an indication of what we see in Figure 18 and we do not want to give that direction a point. In this figure we see that $|M - S_{forward}|$ is smaller than $|M - S_{backward}|$, so if the distance between $|M - S_{forward}| < 10$ samples we give backward one point and vice versa. If both directions are less than 10 samples we give skip this test, thus giving no points to any direction.

The reason why we apply that many checks is because each separate check fails to choose the correct minimum point in some occasions, this will be discussed more in Section 7.1.2. So after the four checks have been applied and if $backward > forward$ we choose $S_{backward}$ as our starting location, if forward is higher $S_{forward}$ is chosen. If both directions have an equal amount of points we simply look one cyclelength further and repeat this process. This process is repeated until we either get a decision or we can not skip more samples ahead. If we can not skip more samples ahead, meaning that all tests have failed, we reset and apply only check 1 from the first maximum point we located, M. Other possibilities could be to have unequal weighting and add more checks like e.g. look at the derivative.

4. **Detect the rest of the steps:**
After we have found the starting location at position S we can finally do the rest of the step detection. As we are in the middle of the collected data we must obviously do this in two stages, backwards and forwards, we start with going forwards. From S, we

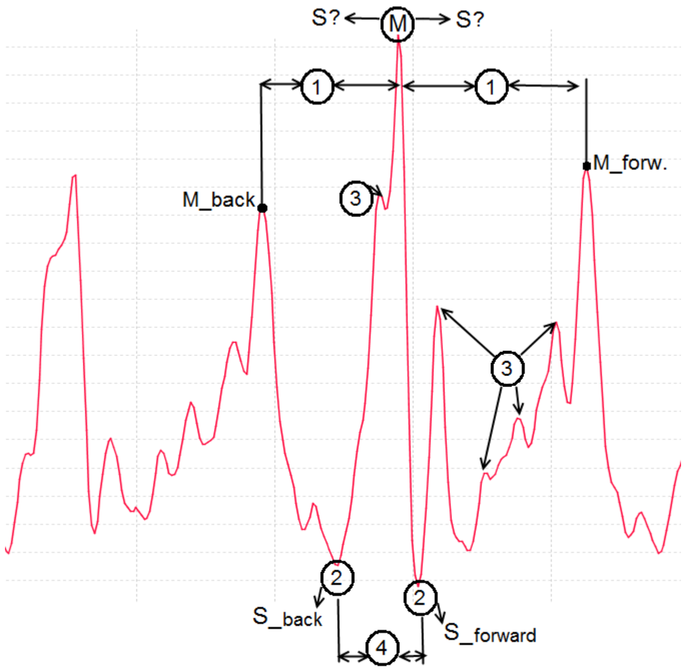


Figure 18: Illustration showing the four decision criteria whether the starting location is before or after the maximum point, M. 1) Length between peaks, 2) Value of minimum, 3) Number of local maximum points and 4) Distance from M to minimum points.

skip γ samples ahead, to position $S+\gamma$ and from here we search in a small area, 10% of the estimated cyclelength, both before and after $S+\gamma$. So we search for the lowest point in $\frac{\gamma}{10}$ samples behind and in front of $X_{S+\gamma}$, $[X_{S+(\gamma-\frac{\gamma}{10})}, ..., X_{S+\gamma}, ..., X_{S+(\gamma+\frac{\gamma}{10})}]$. Now three things can happen:

- The lowest point was found in the first $\frac{1}{3}$ of the search area, in this case we might have skipped too many samples and will therefore search $\frac{\gamma}{10*2}$ more samples backwards. If a new lowest point was found we will continue to search additional $\frac{\gamma}{10*2}$ samples backwards until no new lowest point is found, see Figure 19(a).
- The lowest point was found in the last $\frac{1}{3}$ of the search area, in this case we might

```
1.  S_0 = detectStartingLocation
2.  i = 1
3.  REPEAT
4.      S_i = getNextStartingLocation (S_{i-1})
5.      i = i + 1
6.  UNTIL finished with forward direction
7.  i = - 1
8.  REPEAT
9.      S_i = getPreviousStartingLocation (S_{i-1})
10.      i = i - 1
11. UNTIL finished with backward direction
```

Pseudocode 4.3: Pseudocode for detecting rest of the steps.

have skipped too few samples and will therefore search $\frac{\gamma}{10*2}$ more samples forwards. Like with the previous step, if a new lowest point was found we will continue to search forward until no new lowest point is found, see Figure 19(b).

- The lowest point was found in the middle $\frac{1}{3}$ of the search area, in this case we assume to have found the correct minimum point, see Figure 19(c).

If the minimum point we have found is above α, the indicated minimum value, this might be an indication that our algorithm has detected a wrong step and we must try to reset. This is done by simply searching for the first value below α and use the following local minimum point as our end/start of a step.
When we have found our minimum point we store the location in an array containing all the detected steps and then we will go on to the next step by skipping γ samples ahead and repeat the process until no more steps are located. When searching forwards is complete we repeat this phase by searching backwards so all steps in the walk are identified.

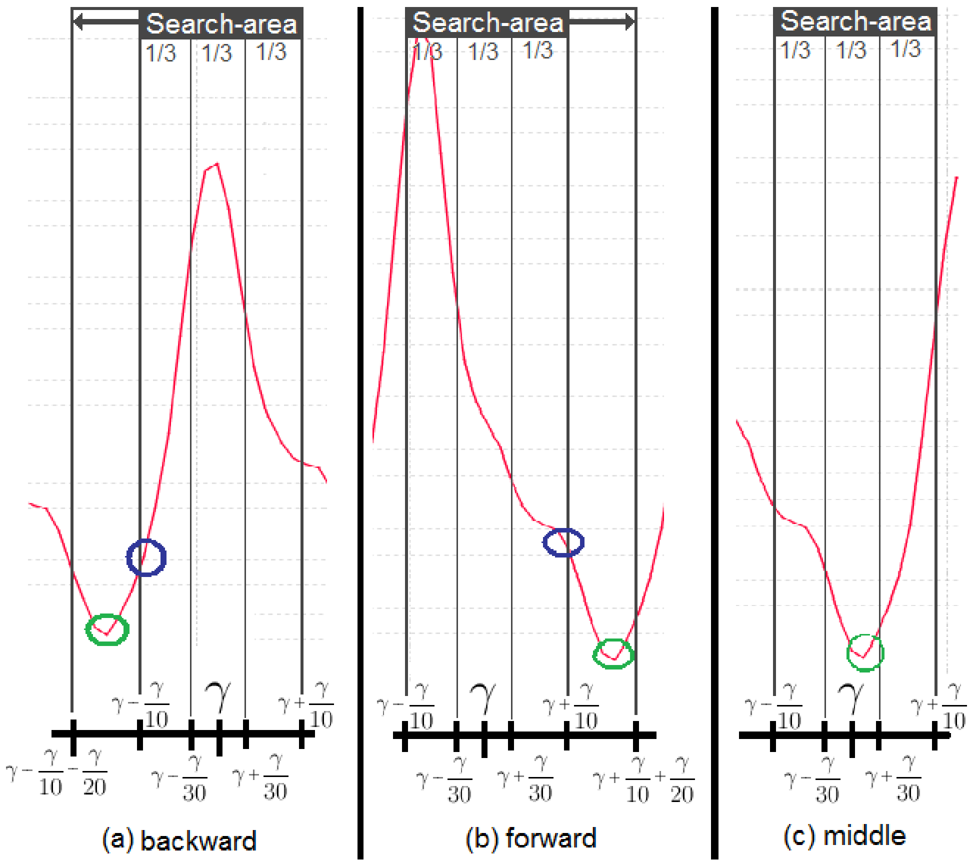


Figure 19: Illustrations of the three options that can happen when we are searching for steps, (a) we have jumped too far and since the lowest point in the search area (blue circle) is in the first third we search additional samples back and find the correct minimum point (green). (b) same as with the backward search only that we search forward this time since we have jumped too short. (c) we have jumped satisfactory and the correct minimum is in the middle third of our search area.

In Figure 20 we can see the final graph with all steps identified and in Appendix B pseudocode for this algorithm is located.

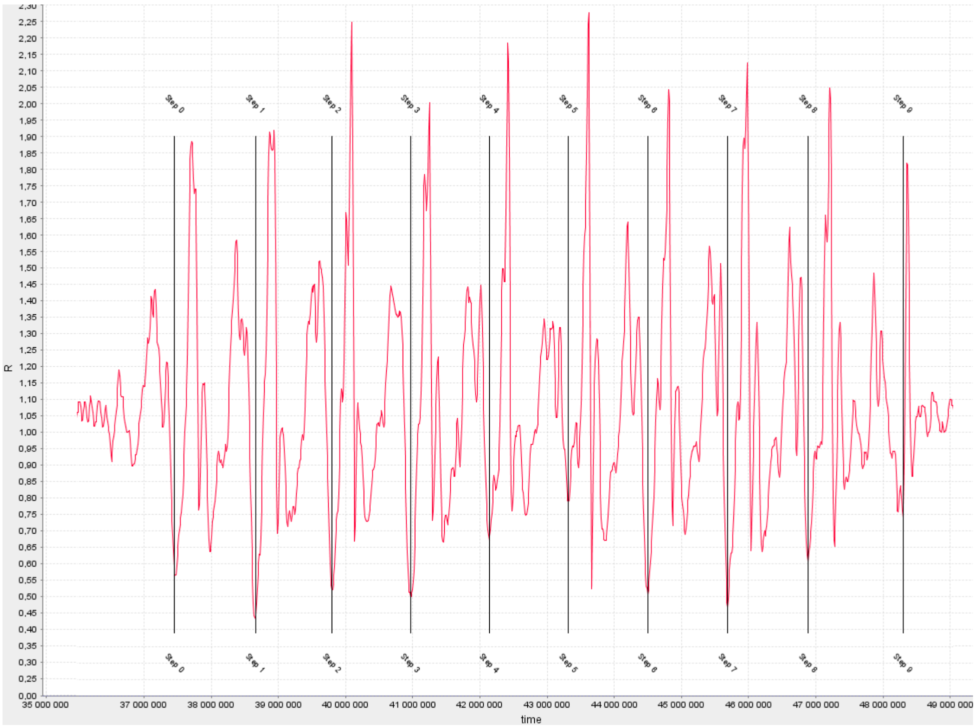


Figure 20: A graph with the correct steps identified.

4.3.3 Creation of average cycles

After we have identified all steps in a walk we have to create an average cycle which will become a template or an input to compare to a template, depending on the situation. In this project we have chosen to not take the first and the last cycle into consideration as no regular walking is done in these phases, the subject is either increasing or decreasing speed. Another possibility instead of using an average cycle could be to only look at the separate cycles as they are. But since the average cycle gave good results in earlier research [11, 41, 46] we have also chosen to average the detected cycles. But before the actual average cycle is created we can apply some pre-processing on the detected cycles:

- **Pre-processing of detected cycles:** All the following methods can be combined in different ways, what yields best results is just a process of trial and error.
  - Time normalization: The first pre-processing method we can do before we create an average cycle is to normalize all cycles to have equal length by using linear interpolation (see Equation 4.1). This length, can in principle have any value, but we have chosen to use 100 samples since normal walking often is close to this number.
  - Acceleration adjustment: If we did choose time normalization, then we also have the possibility to adjust acceleration. Since all cycles are normalized to 100 samples regardless of their original lengths we can choose to compensate for this by adjusting acceleration. Each acceleration value is increased/reduced with the

same factor the cycle is reduced/increased with, e.g. if the cyclelength is reduced from 125 samples to 100 samples then our adjustment factor, λ, is 20%. Whether we shall reduce or increase the orignal acceleration value depends if it is above or below $1g$. This is because $1g$, as mentioned in the start of this chapter, resembles no acceleration. So the adjusted acceleration value is calculated by using Equation 4.5.

$$ACC\_a_t = (a_t - 1) * \left(1 + \frac{\lambda}{100}\right) + 1, \tag{4.5}$$

where a_t is the original acceleration value.

- Amplitude normalization: We can choose to have the amplitude in the cycles range from [0...1], this is done by Equation 4.6. Another option could have been to have a different range, e.g. [−1...1].

$$AMP\_a_t = \frac{a_t - min\_value}{max\_value - min\_value}, \tag{4.6}$$

where a_t is the acceleration-value in position t and max_value and min_value are the maximum and minimum value in the cycle we are normalizing.

- Align cycles: If we did choose to normalize the cycles to a fixed length, then we can now do even more pre-processing. As mentioned earlier, no two cycles are identical, so there are minor variations in both cyclelength and locations of maximum points, this is further discussed in Section 7.1.3. In order to overcome these issues we have the possibility to align the cycles so that a selected maximum point is at the same location for all cycles. This is done by shifting the actual location of the detected steps in the step detection algorithm explained in Section 4.3.2. The adjustment of the cycles can be done by either aligning the cycles so that for example the first maximum point (point B in Figure 14) is at the same location for all cycles or the last maximum point (point D in Figure 14). This is done in the following way:

  1. Normalize the cycles: Before aligning the cycles we always normalize the cycles to 100 samples. Suppose we have n cycles, we then have $n+1$ breakpoints: $O_1, O_2, ..., O_{n+1}$ breakpoints within the signal. Remember the end of a cycle is the beginning of the next cycle (except for the last cycle), and since we have normalized cycles do we have exactly 100 samples between O_i and O_{i+1}.

  2. Detect maximum: The next phase is to detect the maximum point we desire, being either point B or D. Suppose we are aligning at the first maximum, then we will detect the location of point B in each cycle and then take the average of these locations, rounded to the nearest integer. This value, denoted by AVG_B, will refer to the location where the point B in each cycle will be aligned at. So if the location of point B in cycle i originally was at location B_i, then the value of O_i, i.e. the beginning of cycle i, is shifted $B_i - AVG_B$ samples[3]. When the shifting is done for all the cycles, we end up with slightly changed breakpoints $F_1, F_2, ..., F_{n+1}$. The new cycles do not necessarily contain 100 samples at this point, but we normalize the cycles again to 100 cycles as a last action in this phase.

[3]For example if B_i equals 10, while the value of AVG_B equals 8, then the breakpoint O_i of cycle i shifts 2 samples to the right, i.e. $O_i' = O_i + 2$.

When we align on the last maximum, then the procedure is more or less the same, except that we change the values of the breakpoints O_{i+1} in cycle i, based on the values of D_i and AVG_D. The slightly changed breakpoints are now denoted by $L_1, L_2, ..., L_{n+1}$. As above, the cycles are now also normalized to contain 100 samples each.

3. Calculate the distance: We have now three sets of normalized cycles, one where we have aligned at the first maximum, one at the last maximum and one that is not changed. We will now compute some average scores in order to decide which set we shall continue to use. For each cycle in a particular set we compute the distance to all the other cycles in that set, by using DTW, and take the average of these distances. We then end up with three scores; β_{first}, β_{last} and β_{none}.

4. Decision process: After we have calculated all average distances, we choose to use the alignment which had the lowest score. Suppose $\beta_{first} < \beta_{last}$ and $\beta_{first} < \beta_{none}$, then our set will be replaces by the set of cycles based on alignment on the first-maximum.

This process is repeated as long as either β_{first} or β_{last} is lower than β_{none}. The reason why we repeat this process is that the actual positions of the steps are changed and that the end of one step is the start of the next. This implies that if the start of $cycle_i$ is shifted n samples back, then the end $cycle_{i-1}$ is also shifted n samples back. Since we have normalized the cycles will this lead to that the last maximum point in $cycle_{i-1}$ is shifted compared to what it was before we shifted the start of $cycle_i$. In order to adjust for this do we repeat the process until we do not get a lower average distance by aligning maximum points.
Other possibilities that could have been done is different comparison method, e.g. absolute distance, only do one iteration, not normalizing the cycles, include the alignment of the middle point (point C in Figure 14), etc.

- Skip irregular cycles: As a final phase before we create the actual average cycle we have the possibility to skip cycles that are very different from the others. For each cycle we compute the distance, by using DTW, to all the other cycles, $d_{i,j} = d(cycle_i, cycle_j)$. Then we take the average of these distances, $d_i = \frac{1}{n-1}\sum_{j \neq i} d_{i,j}$. So then we get the average distance from $cycle_i$ to all other cycles. When the averages, d_i, have been computed for all cycles we take the average of these values again, $D = \frac{1}{n}\sum_i d_i$. So D would be the total average for all the cycles. Then we look at the cycle that has the highest difference from D, if this cycle has an average that is more than 15% difference from the total average we omit this cycle, so $d_i \notin [D - 15\%, D + 15\%]$. The reason why we use 15% as a limit is due to a process of trial and error. If we had chosen a lower limit we might had ended up skipping too many cycles, while a higher limit would lead to not skipping cycles we want to skip. This process is repeated until no more cycles are omitted.
  Again we have other possibilities, different distance metrics, only do one iteration, change the limit, etc.

- **Create average cycle:** When we are finished with the pre-processing of cycles the average cycle can be derived in different ways:

- DTW: We can keep the cyclelengths as they are and use Dynamic Time Warping (DTW), see Section 4.3.4, to create an average cycle. We create this average cycle by following an idea which was described in [53], this idea is to use that cycle which has the lowest average distance to all the other cycles. So for each cycle we compute the distance to all the other cycles and take the average of these distances. The cycle with the lowest average will be our average cycle. Another possibility to our approach, which also was discussed in [53], could be to instead of using one of the existing cycles, create a cycle that has the lowest distance to all the existing cycles as possible.
- Mean: Assume $C^i = [c^i_1, c^i_2, ..., c^i_n], n = 100, i = 1, ..., K$ are the K normalized cycles of the signal. Then the average cycle, $C = [c_1, c_2, ..., c_n]$ can be calculated as follow, $c_j = mean(c^1_j, c^2_j, ..., c^K_j), j = 1, ..., 100$. In other words, each acceleration value in the averaged cycle is the mean of the corresponding values in the normalized cycles.
- Median: Same as with mean, just that we use the median (middle point) instead. By choosing the median from all cycles we reduce the impact of possible outliers.
- Trimmed mean: This method is also known as “mean with outliers” and its purpose is the skip those points that are far away from the mean. We choose to do this the following way:

  1. Calculate mean, μ and σ, in point n of all cycles.
  2. All points that are outside of $[\mu - 2 * \sigma, ..., \mu + 2 * \sigma]$ are removed.
  3. Step 1 and 2 are repeated without the thrown away values. This is done until no more points are removed.

  This is only one possible way of calculating the trimmed mean, other possibilities could be to use 1 iteration, first throw away values above/below a certain threshold, etc.

- **Post-processing:** After the average cycle has been created we also have some possibilities again:
  - Repeat amplitude normalization on the newly created average cycle. Normalize the average cycle to 100 samples if did not choose to do that earlier, e.g. when we used DTW to decide an average cycle.
  - Derivative: We can also choose to take the derivative of the average cycle and use that instead of the actual acceleration values. The reason for this choice would be that we are more interested in the shape of the cycle than the actual values it contain.

An example of time normalized cycles plotted on the same graph and the resulting average cycles can be seen in Figure 21-22. Pseudocode can be seen in Appendix B.

4.3.4 Dynamic Time Warping

The Dynamic Time Warping (DTW) algorithm is used to measure distance between two sequences which may vary in length and also find the optimal way of transforming one

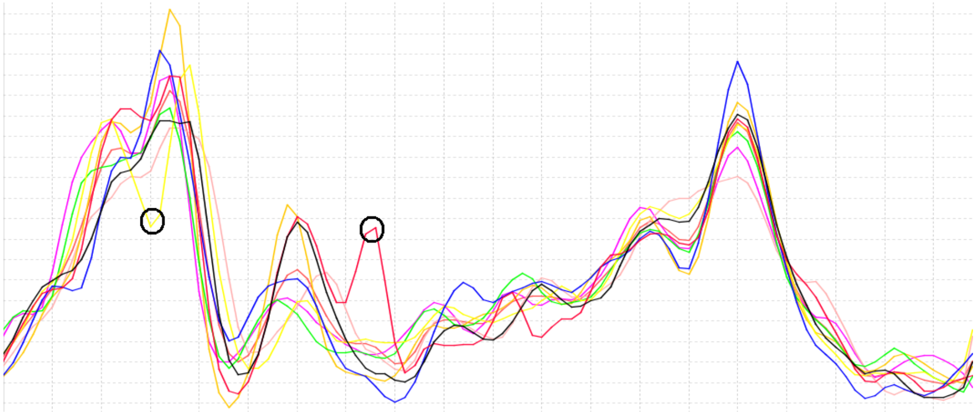


Figure 21: Cycles identified in Figure 20 plotted on top of each other. Notice the alignment of the last maximum, as well as the strange values indicated by a circle, these values will have no impact on the final average cycle when we are using median and trimmed mean as averaging methods.

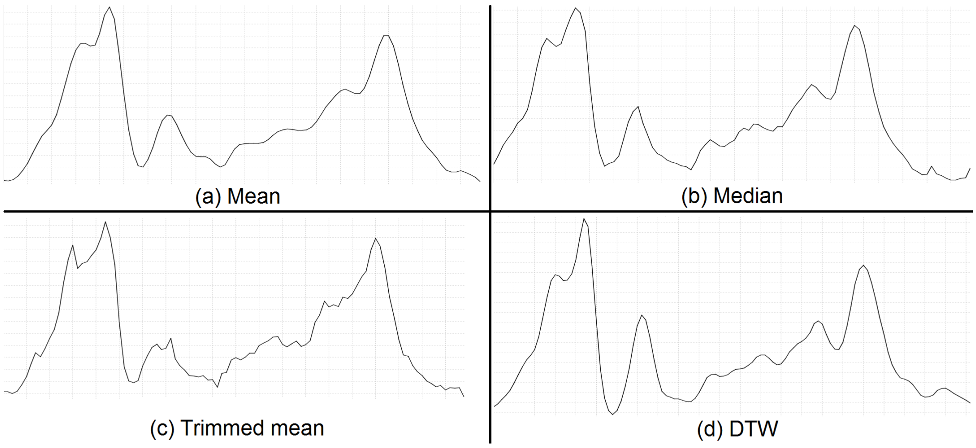


Figure 22: The resulting average cycle from cycles in Figure 21. (a) mean, (b) median, (c) trimmed mead and (d) DTW (in this case with no acceleration adjustment or amplitude normalization).

sequence into the other. DTW is very similar to Levenshtein distance (also known as edit distance). Levenshtein distance is actually what we are going to use as this is a distance metric while DTW is a matching method. But as a side-product of DTW we do get the distance between two sequences so in the rest of the report we will use the term DTW. In the following subsections will we first give a brief introduction to DTW, then will we look at our approach.

Introduction:

In contrast with more simple distance measurements that only compare sample number n in the template with sample number n in the input, DTW finds a more optimal alignment between the template and input. Let us give a simple example of DTW illustrating this. Suppose we have two strings, “study” and “muddy” and we are interesting in a possible way of transforming “study” into “muddy”, see Table 3. When using DTW we do have three possible operations when we are transforming sequences: substitution, deletion and insertion. As we see from the example “s” is substituted with “m”, “t” is deleted, “u” and "d" are left unchanged, the second "d" in "muddy" is inserted and finally "y" is left unchanged. In order to calculate the cost of the total transformation we must introduce some different cost functions for each of the three operations [54]:

- Insertion: $C_{ins}(x)$ is the cost of inserting x. The cost of inserting a value is normally fixed, but may also depend on the value that is inserted.
- Deletion: $C_{del}(x)$ is the cost of deleting x. As with insertion the cost of deleting a value is normally fixed, but may also depend on the value that is deleted.
- Substitution: $C_{sub}(x, y)$ is the cost of substituting x with y. The cost of doing this change can be as simple as the absolute distance between the values or simply a fixed value, but it may also be more sophisticated.

So by using these cost functions, the total cost in our example is: $C = C_{sub}(s, m) + C_{del}(t) + C_{ins}(d)$.

s ↓ m	t ↓ λ	u ↓ u	d ↓ d	λ ↓ d	y ↓ y
subst	del	-	-	ins	-

Table 3: A simple example showing a possible transformation from the two sequences, as well as the different cost operations, substitution and insertion. λ indicates an empty symbol.

Finally the goal of DTW is to find the cost of the cheapest transformation from the template to the input. In order to do that we must calculate all different possibilities of combining cost functions. If we continue with the same example and assume each operation has a fixed cost of 1 then the cost matrix is shown in Table 4. The bold numbers indicate a possible path in order to achieve the lowest cost. In this matrix the cost of transforming "muddy" to "study" is given in the rightmost cell in the bottom row, as we see the minimum distance in our example is $C = 3$. More general can we say that the cost of transforming a (sub)-sequence of length i to a (sub)-sequence of length j is given in the cell (i, j). If we continue to look at the cost matrix we see that we do have the same operations as in Table 3, the reason for this is that a diagonal path indicate a substitution, vertical is deletion and horizontal is insertion, see Figure 23.

	λ	m	u	d	d	y
λ	**0**	1	2	3	4	5
s	1	**1**	2	3	4	5
t	2	**2**	2	3	4	5
u	3	3	**2**	3	4	5
d	4	4	3	**2**	**3**	4
y	5	5	4	3	3	**3**

Table 4: The cost matrix of our simple DTW example, a possible path with the lowest cost is indicated with boldface.

In our project we look at sequences of numbers instead of sequences of strings, but the concept remains the same. In Figure 24 we show two series being warped by using DTW, insertions occur when multiple points in the template are matched to the same point in the input, deletion occur when the same point in the template is matched to

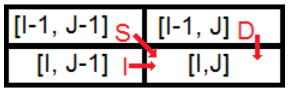


Figure 23: The three different options, Deleting (D), Substituting (S) and Inserting (I), we have when we create a distance score matrix. We choose that option that has the minimal cost.

multiple points in the input. Below will we explain how this is done in more detail.

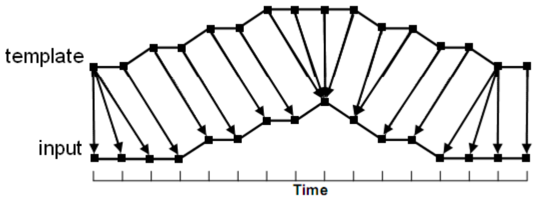


Figure 24: An example of two series being warped (from [55]).

Our DTW algorithm:

This algorithm is based on the classical DTW algorithm given in [56]. Suppose we have two time series Q and C, of length n and m respectively, where:
$Q = q_0, q_1, q_2, ..., q_i, ..., q_n$ and $C = c_0, c_1, c_2, ..., c_j, ..., c_m$, where q_0 and c_0 is an empty symbol. To align two sequences using DTW we construct an n-by-m matrix, M. In the first column we add the cost of deleting a value, while in the first row we add the cost of inserting a value, the first element, $M_{(0,0)}$ contains the value 0, similar to Table 4. The element $M_{i,j}$ contains the distance $\delta(q_0, ..., q_i, c_0, ..., c_j)$ between the two points q_i and c_j. A partially filled matrix is shown in Table 5. A warping path W, is a contiguous (in the sense stated below) set of matrix elements that defines a mapping between Q and C. The k^{th} element of W is defined as $w_k = (i, j)_k$, so we have:

$$W = w_1, w_2, ..., w_k, ..., w_K, \quad \max(m, n) \leq K < m + n + 1 \tag{4.7}$$

The warping path is typically subject to several constraints:
Boundary conditions: $w_1 = (0, 0)$ and $w_K = (m, n)$, simply stated, this requires the warping path to start and finish in diagonally opposite corner cells of the matrix.
Continuity: Given $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a - a' \leq 1$ and $b - b' \leq 1$. This restricts the allowable steps in the warping path to adjacent cells (including diagonally adjacent cells).
Monotonicity: Given $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a - a' \geq 0$, $b - b' \geq 0$ and $|a - a'| + |b - b'| \geq 1$. This forces the points in W to be monotonically spaced in time.

The different cost functions are defined in the following way:

- Inserting: $C_{ins} = 0.5$.
- Deleting: $C_{del} = 0.5$.
- Substituting: $C_{sub} = \frac{|x-y|}{S}$, where x and y are the values of the template and input and $S = \max(q_1, ..., q_n, c_1, ..., c_m) - \min(q_1, ..., q_n, c_1, ..., c_m)$.

The reason why we choose 0.5 as the cost of inserting and deleting is actually a process of trial and error. We know that the cost of substituting ranges from $[0, ..., 1]$ since the absolute distance between the template element and input element is divided with S. So by choosing some different values for inserting and deleting 0.5 was the cost that yielded the best results. The cost matrix is first initialized by filling the first row and column in the following way: $\delta(i,0) = \delta(i-1,0)+C_{del}$ and $\delta(0,j)\delta(0,j-1)+C_{ins}$, where $n \geq i > 0$ and $m \geq j > 0$. Since we are only taking the distance between a value and the empty symbol we are only considering deletions and insertions. After initialization the rest of the cost matrix is filled by the applying Equation 4.8 given below:

$$\delta(i,j) = min\{\delta(i-1,j-1) + C_{sub}, \delta(i-1,j) + C_{del}, \delta(i,j-1) + C_{ins}\}, \qquad (4.8)$$

where $n \geq i > 0$ and $m \geq j > 0$ and we always consider the whole substring before i and j. If we continue to look at the partially filled table (Table 5), we see that the next cost we are going to calculate is the cost of transforming 1.2 to 2.6 (the cell with "?"). So from Equation 4.8 we get that, $\delta(3,4) = min\left\{0.59 + \left(\frac{|1.2-2.6|}{2.5}\right), 1.09 + 0.5, 0.18 + 0.5\right\}$ $= min\{1.15, 1.59, 0.68\}$. As we see will we do an insertion since 0.68 is the lowest value, so 0.68 will then be the value of this particular cell. A similar approach of DTW as a distance measurement is described explained in [57]. In Appendix B one can see pseudocode for our algorithm.

	λ	**0.2**	**0.5**	**1.4**	**2.6**	**1.3**	**0.9**
λ	0.0	0.5	1.0	1.5	2.0	2.5	3.0
0.1	0.5	0.05	0.55	1.05	1.55	2.05	2.55
0.6	1.0	0.55	0.09	0.59	1.09	1.59	2.09
1.2	1.5	1.05	0.59	0.18	?		
1.3	2.0						
2.4	2.5						
0.9	3.0						

Table 5: A partially filled cost matrix of two sequences of numbers.

Our approach is however just one possible way of doing it, we could e.g. have different cost functions. We have also briefly studied at some other approaches, but decided not to implement these:

- Classical DTW: In [56] the classical approach of DTW is given. The major difference between that approach and our approach is that in [56] they made no distinction between costs related to insertions, deletions and substitutions. They use the equation $\delta(i,j) = d(q_i, c_j) + min\{\delta(i-1,j-1), \delta(i-1,j), \delta(i,j-1)\}$, where $d(q_i, c_j)$ is the Euclidean distance.
- Derivative DTW: In [56] a modification called Derivative Dynamic Time Warping (DDTW) was introduced, here they look at the first derivative of the sequences. The only modification from classic DTW [56] is to change the distance measure $d(q_i, c_j)$ from Euclidean to the square of the difference of the estimated derivatives of q_i and c_j.
- Raw data, first-order and second-order derivative: In [58] another modification was presented, they looked at a combination of Y values, first derivative and second derivative, $d(i,j) = w1*(x_i - y_i)^2 + w2*(\dot{x}_i - \dot{y}_i)^2 + w3*(\ddot{x}_i - \ddot{y}_i)^2$, where $w1, w2, w3$

must be chosen properly after an analysis of the data and should obey the following relations: $w3 > w2 > w1$.

- Rong et al. approach [42]: As mentioned in Section 3.3.3 did Rong et al. implement an algorithm they called DTW. But this algorithm was entirely different from our approach and a very simple implementation.

5 Experiment

In this chapter we will explain details about the execution of the different experiments.

5.1 Experiment details

5.1.1 Main experiment setup

In order to answer the research questions raised in Section 1.5 we will perform an experiment. In this experiment we will look at how we are able to recognize people under different circumstances:

- Normal walk: This will be walking back and forth at the subjects normal velocity.
- Fast walk: The participant will be told to walk faster than his normal walking velocity. It is however important that he is still walking, meaning that at every point at least one foot have contact with the ground. In order to later have the possibility to see how fast the participants walked, the distance will be fixed to 20.5 (± 0.1) meters.
- Slow walk: Same as "fast walk", but this time the participant will walk slower than his normal walking velocity.
- Circle walk: The purpose of this part is to see how your gait is affected by walking around a corner. In order to be sure that we only look at walking with a change of direction the participant will walk in a circle, both clockwise and counter-clockwise.

5.1.2 Sub-experiment

In addition to this main experiment we will also perform a sub-experiment.
Do people walk the same way under the same circumstances:
We will also perform a long-term experiment where we look at whether people walk the same under the same circumstances. A small group of volunteers will walk once in the morning and once in the afternoon over a longer period of time. An important part of this experiment is to have full control over the external influences, therefore will the participant walk with the same shoes, clothing and be in more or less the same shape and condition.

5.2 Experiment execution

5.2.1 Main experiment

The main experiment will be carried out on a solid surface and the location will be the old library in the A-building. The test-subjects will wear an accelerometer attached to a belt. The accelerometer will be placed on the left leg, by the hip. By attaching the accelerometer to a belt we ensure that the accelerometer more or less has the same orientation for all subjects. The subjects will be asked to walk a fixed length on all walks, it is especially important to maintain this when walking at different velocities. This fixed length will be across the entire room in the old library. Each participant walks at least 20m and 43cm, but since many does not stop precisely at the marker we can estimate one walk to be in the range from $[20.40m, 20.60m]$. This length will of course be a bit different for the circle walk, but in this case speed is not an issue.

The walking session will be conducted in this way for all walks:

- For each time the subject is supposed to walk, he must wait 3 seconds before he can start walking. The subject walk the distance,then stop, wait and turn, and then walk the same distance back. To summarize, the subject must:
  1. Wait 3 seconds.
  2. Walk the distance.
  3. Stop and wait for 3 seconds.
  4. Turn around.
  5. Wait 3 seconds.
  6. Walk the distance back.
  7. Stop and wait for 3 seconds.
- The subjects will walk each circumstance except the normal walk, 8 times, after which the data will be downloaded and verified, meaning that we visually look at the data and locate the different walks without any sensor errors. For the normal walk we will get in total 12 samples divided into two separate sessions.
- Each circumstance take approximately 4 minutes, so the total time will be about 20 minutes each participant. For those who do not have time to do all at once there is a possibility to split the experiment into two parts, where part one consists of normal, slow and fast walks, while part two consists of normal again and walking in a circle. When people choose to split the experiment up we will take a note of special footwear or any other condition that might be different from session two. When we then compare the normal walks from the different sessions we will see if there is any change.

We will collect name, age and gender from each participant. In addition, where necessary, we will write down comments, in the case of injuries, special clothing/shoes, etc. The participant will sign an informed consent, see Appendix A.

5.2.2 Sub-experiment

Do people walk the same way under the same circumstances:
The participants will walk once in the morning (08:00-12:00) and once in the afternoon (15:00-21:00) on as many days as possible during a longer period of time. They will walk on a solid floor at their normal speed and generate 4 samples each time, and the walking procedure will follow the instructions mentioned in Section 5.2.1.

5.3 Volunteer crew

The volunteers in the main experiment were mainly students and employees at the school. In total we had 60 participants (17 females and 43 males) and all participants were healthy people with no special injuries that may affect the gait. Most people used regular shoes with a flat sole, but we also had participants with high-heels, slippers and more heavier winter shoes. We did also have different heights and weights and people from different countries and cultures. We even had one woman that was pregnant, but she felt that she was still able to walk in her own way. The age ranged in total from 20 to 64 years. For males the age ranged from 22 to 58 with an average of 33 years ($\mu = 32.91$ and $\sigma = 11.04$), while for the females the age ranged from 20 to 64 years with an average of 35 years ($\mu = 35.06$ and $\sigma = 15.24$), in Figure 25 one can see the gender and age distribution. For the long-term experiment 5 people, (3 male and 2 female) all students in the age of 21 to 25 participated in their regular shoes with a flat sole.

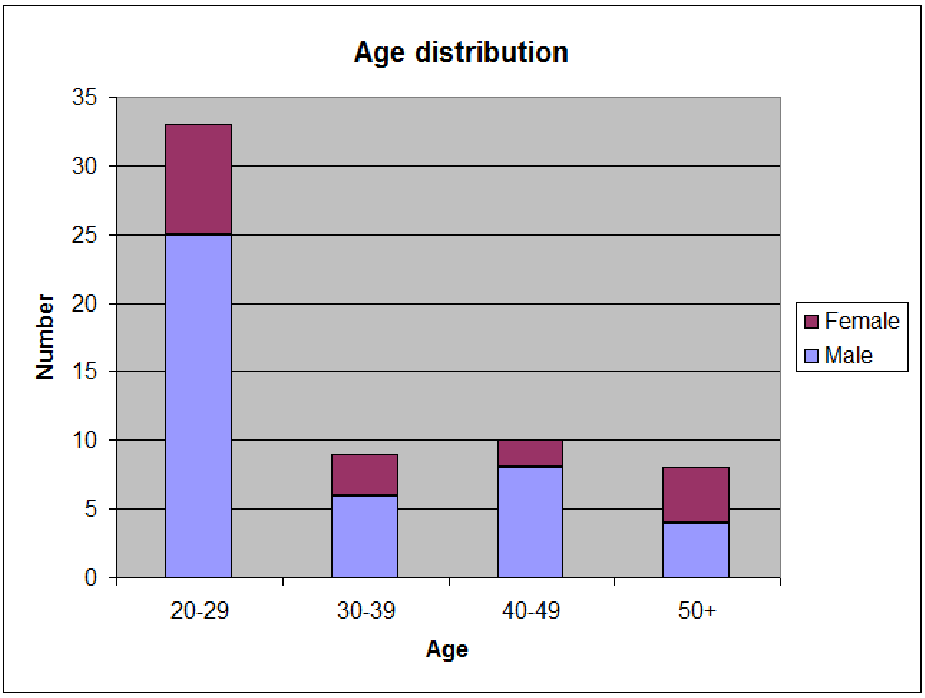


Figure 25: The gender and age distribution for the main experiment.

5.4 Environment

As mentioned the main experiment and the long-term experiment were all carried out on a flat solid floor. The whole experiment took place during the spring period, therefore did some people still have winter shoes, while others had lighter shoes. The main experiment were done between 09:00 and 16:00.

6 Analysis and results

In this chapter we will first explain how the data has been analyzed then we will look at the results obtained during this project.

6.1 Analyzing methods

In this section we will explain how we analyzed the data, first we will look at the distance metrics used in this project then will we see how we compute a distance score table which is used to create Decision Error Trade-off (DET) curves.

6.1.1 Calculate distance score

After we have calculated the average cycle for all subjects, we can now compare them. The average cycle of the person is used as a feature vector and by using a distance metric one will get a distance score. When it comes to distance metrics there are endlessly many possibilities, whatever sounds logical and could give good results can be used. In this project we have tried the following metrics:

- Manhattan distance: This is also known as absolute distance and the formula is shown in Equation 6.1. Manhattan distance is a very simple metric that takes the sum of the absolute values of the differences between all the values in the template and the corresponding values in the input. As a result of this, Manhattan distance requires that the template and the input have equal length. In addition, this distance metric is the computationally least expensive one.
- Euclidean distance: This is a slight modification of the Manhattan distance, see Equation 6.2. Instead of taking the sum of the absolute differences we now take the square root of the sum of all differences squared.
- Dynamic Time Warping (DTW): In addition to aid us in creating an average cycle, DTW actually yields a distance between two vectors. Therefore it is naturally to use this as a distance metric as well. The advantage of DTW compared to the metrics given above is that the template and sample no longer have to be of equal length, for further details, see Section 4.3.4.
- Derivative: All metrics mentioned above can also be applied to the derivative of the signal, as mentioned in Section 4.3.3. What we actually achieve by taking the derivative is that we adjust differences in Y-axis for each sample. If the input-curve is shifted in the Y-direction we are only penalized for this difference once.

$$\mathrm{dist}_{\mathrm{Manh.}}(\underline{X},\underline{Y}) = \sum_{i=1}^{n} (|x_i - y_i|) \tag{6.1}$$

$$\mathrm{dist}_{\mathrm{Eucl.}}(\underline{X},\underline{Y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{6.2}$$

In the equations above are x_i and y_i the resulting acceleration values at observation point i, and $n = 100$. These distance metrics will give us the distance score between the two gait signals, $\underline{X}$ and $\underline{Y}$, the score should be smaller for genuine attempts than for fraudulent attempts.

In addition to different distance metrics we also have some other possibilities that may give us better results, we can choose to split the average cycle into different parts and we can do cyclic rotation on the average cycle in order to get as low distance as possible.

- **Split cycle:** When we are comparing cycles we also have an option to split the average cycle into smaller parts. In this project it is natural to split the average cycles where we have the following extremes, first maximum point, first minimum point and final maximum point, cf. Figure 14 on Page 26. In other words, we get 4 vectors from: A-B, B-C, C-D and D-E. This is done for both the template and the input and the vectors of the input are normalized to have the same length as the template. By doing this we can adjust for minor shifts in the X-axis, and this is actually a simplified version of DTW. We also have the possibility to add a penalty to the distance score when the vectors have different lengths.
- **Cyclic rotation:** Another possibility we also can perform is to do cyclic rotation in order to get as low distance as possible. This means that the input, I, that is compared against the template, T, is cyclical rotated. After each rotation the new distance is calculated using the chosen distance metric. This is repeated until the input has done a full rotation, then the lowest distance score is kept. So $T = (t_1, ..., t_n)$ and $I = (s_1, ..., s_m)$, where n and m is the total number of samples of the template and the input respectively. And we define cyclic rotation as: $I^j = (s_j, ..., s_m, s_1, ..., s_{j-i})$. The distance that is kept is then: $d(T, I) = min_{j=1..m} d(T, I^j)$. This cyclic rotation is done to minimize the problem when local extremes among the average cycles we create for each input are located at different locations.

6.1.2 Creating distance score table

When we have chosen one distance metric and calculated the distances between the average cycles of the walks of all participants we can finally create a DET curve. This is done by calculating False Match Rates (FMR) and False Non-Match Rates (FNMR), see Section 2.2.4 for more detailed information. In order to produce the DET-curve with its FMR and FNMR we must create a table consisting of different distance scores. There are different ways of doing this, a possible way is to choose a template for each subject and compare this against all the other recordings we have collected. The scores are calculated with a selected distance metric from Section 6.1.1. The templates are matched against all the other samples produced by the same person, which are called genuine attempts, and against all the other samples produced by the other persons, which are called fraudulent attempts. Suppose we have N persons and K recordings collected per person. This will give us 1 template per person and $(K-1)$ inputs per person. The total number of genuine attempts is then $G_{tot} = N * (K - 1)$, while fraudulent attempts will be $F_{tot} = N * (N - 1) * (K - 1)$. A small excerpt of a matching score table can be seen in Table 6. There are different ways to choose a template which we will describe below:

- One can decide to always use recording-n to be the template while the other record-

	P_1 C_1	P_1 C_2	P_1 C_3	P_2 C_1	P_2 C_2	P_2 C_3	...	$P_N C_3$
Template P_1	**0.944**	**1.095**	**2.792**	3.240	3.419	3.301	...	3.862
Template P_2	3.509	3.489	1.700	**0.722**	**0.840**	**1.258**	...	3.014
Template P_3	4.187	4.012	2.597	2.602	2.855	2.950	...	2.424
Template P_4	2.200	2.552	1.857	2.502	2.638	2.565	...	3.564
Template P_5	3.428	2.396	1.627	1.633	1.377	1.265	...	3.506
...	...	...	...	...	...	...	...	...
Template P_N	4.496	4.312	2.873	2.830	3.097	3.055	...	**0.806**

Table 6: A small excerpt of a matching score table, the bold number indicate genuine attempts.

ings are used for comparison.

- Another possibility is to start with the first recording as template, then move on to the second recording etc. until all recordings have been the template. The Equal Error Rate (EER) will then be the average of the individual EER's, EER is the point where FMR = FNMR.
- Another option is to not decide what is a template on forehand, one can compare all the different recordings with each other and while running through all the possibilities decide what is a genuine attempt and what is fraudulent attempt.
- The last option, which is the one we are going to use, is to randomly select templates for each person and then derive EER's. This process is repeated for several iterations and we will then get a statistical significant average EER.

6.1.3 Statistical methods

In the long term experiment we will in addition use some other statistical methods to analyze the data [59].

Linear Regression:

With linear regression the relationship between two variables, X and Y is analyzed. For each subject we know both X and Y and we want to find the best straight line through the data, $y = \alpha + \beta x$. When we are using linear regression we might be interested in the intercept (α) and/or the slope (β) value, while in other cases we use the regression line as a method to derive new values of X from Y or vice versa. The regression line is found by finding the line that minimizes the sum of the squares of the vertical distances of the points from the line, see an illustration in Figure 26.

Confidence interval:

In addition to the regression line we will also calculate a Confidence Interval (CI), and the most common CI is a 95% interval. The CI is calculated from the standard error values of the slope and the intercept, meaning that we can calculate the CI for both the slope and the intercept. If you accept the assumptions of linear regression, there is a 95% chance that the 95% CI of the slope contains the true value of the slope, and the same applies to the intercept. This means that there is a 5% chance that the true line is outside the boundaries. But note that many data points will be outside this CI boundary. The CI is only 95% sure that it contains the best-fit regression line, this is not the same as saying

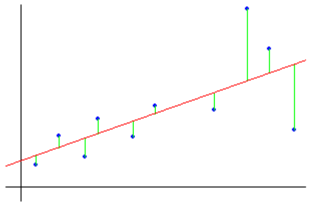


Figure 26: An example showing linear regression, in the graph we see the red regression line that is derived from the blue data values.

it will contains 95% of the data points.

p-value:

The p-value is calculated when we are doing hypothesis testing, and this value is the probability that your sample could have been drawn from the population being tested given the assumption that the null hypothesis is true. A p-value of 0.05 indicates that you would have only a 5% chance of drawing the sample being tested if the null hypothesis was actually true.

Null hypothesis are typically statements of no difference or effect. A p-value close to zero indicates that your null hypothesis is false, and typically that a difference is very likely to exist. In general, a p-value of 0.05 is the typical threshold used, and if our p-value is below this threshold we can conclude the null hypothesis is wrong. The p-value is calculated by applying statistical tests, e.g. t-test.

6.2 Results

In this section we will first look at the results of the main experiment, then we will look at the sub-experiments.

6.2.1 General information

First will we look at some general information about the analysis of the data.

Analysis method: As mentioned in Section 4.1 the software that was created, gave us the possibility to manually detect steps and thus be able to produce a more optimal average cycle. This helped us in getting a better understanding of the properties of gait cycles, which again led to an improvement of the step detection algorithm. In Appendix C EER for these manually cycles can be seen, results in this chapter are however based upon fully automatic step detection. Common pre-processing for all scores are:

- Time interpolation: We interpolated the data in order to achieve a value every $\frac{1}{100}$ second.
- Noise reduction: Weighted Moving Average (WMA) with window size 5 have been applied to all data as this performed slightly better than Moving Average (MA).
- Resultant vector: We created a resultant vector consisting of all three directions (X,Y and Z) recorded by the sensor. The reason for this is that by reviewing a small scale experiment it turned out that combining all three direction gave better results than only X, Y or Z, or when omitting Z.

- Average cycle: Common for all the average cycle creations is attempting to align either the first or the last maximum before actual creating the average cycle.
- Derivative: Whenever we are using derivative as distance metric we do look at the the derivative of the average cycle with Manhattan distance applied. The reason why only Manhattan is used is that it gave the best results on a small scale experiment.
- Dynamic Time Warping: Where we have used DTW we applied the approach presented in Section 4.3.4. This approach produced better results than to the other DTW approaches we briefly mentioned, like using the classical approach or using first and/or second derivative.
- Cyclic rotation: When comparing average cycle we do cyclic rotation and choose the lowest distance as our distance score.

For normal walking we will present a more thorough table showing different averaging methods and distance metrics, for the other circumstances will we only show the setups that gave the best results for that particular circumstance. We have chosen to not present EER where we split the cycle into four different parts in this chapter as these results were considerably poorer than for the other methods. When we used DTW to create the average cycles we also normalized the resulting average cycle to 100 samples so that we could use the distance metrics that require inputs of equal length. These scores where however also so poor that we did not include them here, a total overview is given in Appendix C.

Development of distance score: Another important aspect of the gait cycle is to look at which areas that are really distinct between different persons. As we can see from Figure 27 the distance score rises more rapidly around the two main local maximum points within a cycle. In this particular case we are looking at two normal walks which have been pre-processed by using weighted moving average, the average cycle is created by using mean with no modification and the distance metric is Euclidean distance. By reviewing this information it seems that the very start and end, as well as the middlearea in the recordings could have been omitted. By omitting these parts we would reduce the computational cost and still get good results.

6.2.2 Main experiment

Analyzing the results obtained from the main experiment is not a trivial task and can be done in many different ways. First let us look at some general information. As mentioned in Section 5.3 we had in total 60 participants, where each participant generated the following: 12 genuine samples for normal walk (divided into two sessions), 8 genuine samples for slow, fast and circle walking. The circle was divided into what direction the volunteers walked, right (clockwise) and left (counter-clockwise). Remember that the sensor was on the left side meaning that when the subject walked left the sensor was on the inside of the circle while when he walked right the sensor was on the outside. The reason why we divided the circle walks is that there was a significant difference between the two directions. The main experiment was performed in a one month period. Those participants who decided to do part two of the experiment later did wear the same footwear both sessions.

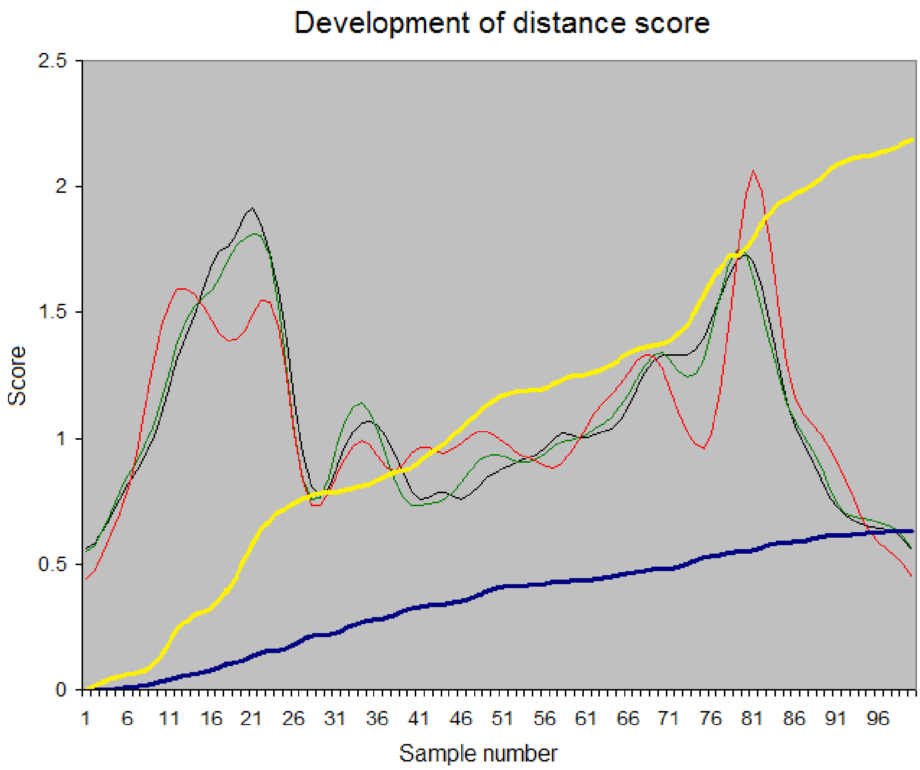


Figure 27: An illustration showing how the distance score develops throughout the cycle. The cycles that are compared can be shown in the background, black is template, green is genuine attempt while the red is a fraudulent attempt. The corresponding genuine score development is shown in the blue line while the yellow line indicates the development of the fraudulent score.

Normal walking

First we will look at only the normal walking as this is our best way to compare our methods and results with previous work. For normal walk we got between 10 and 15 cycles per walk, and the average cyclelength was 109 samples ($\sigma = 6.9$) and the cyclelength ranged from [95...125] samples. The normal walking was divided into two sessions, each with 6 walking samples. By first looking at them separately and choosing the templates at random do we have $60 * 5 = 300$ genuine attempts, while the number of fraudulent attempts is $60 * (60 - 1) * 5 = 17700$ attempts per iteration. An overview of scores with some different pre-processing and analysis methods can be seen in Table 7, EER are based on the average from 10 iterations, the setups that gave best results did undergo 100 iterations and are presented in Table 8. As we see from the tables, DTW is the distance metric that yields the best results, but taking the derivative also gave good results. Furthermore using mean and median as averaging methods seems to be producing the best scores. We also see that adjusting for acceleration is slightly better than no adjustment, while normalizing the amplitudes lower the score when using DTW. By looking at on of the best setups, median with acceleration adjusted and DTW as distance metric ($\mathtt{median}_{A+DTW}$), we see that there is only slight difference between the first normal session compared with the second session, EER of 1.60% (σ=0.26%) and 1.89%

(σ=0.23%) respectively. Another aspect worth looking at beside EER is what the FNMR would be if the FMR is 0, and vice versa. As we recall from Section 2.2.4, when FMR is 0 it implies that we do not have any impostors being granted access while a FNMR of 0 means that no genuine users are being rejected. By looking at the $\mathtt{median}_{A+DTW}$ setup we get these numbers: $FMR = 0 \Rightarrow FMNR = 10.1\%$ and $FNMR = 0 \Rightarrow FMR = 51.1\%$.

If we look at all the normal walks as one session of normal walking, 12 genuine trials for each person, it would give us $60 * 11 = 660$ genuine attempts and $60 * (60 - 1) * 11 = 38940$ fraudulent attempts. Again DTW with mean and median average cycles with acceleration adjustment produced best results, as shown in Table 7. The EER after 10 iterations depreciated to 5.86% (σ=0.49%) for the $\mathtt{median}_{A+DTW}$ setup, which was when the average cycles was created by taking the median and with the acceleration adjusted. The reason for this decrease is partially due to twice the number of samples, but the main factor is that some people walked in a different way when we look at average cycles from the different sessions, this will be further discussed in Section 7.2.1. The Decision Error Trade-off (DET) curve from one of the iteration from the $\mathtt{median}_{A+DTW}$ setup is shown in Figure 28. If we again look at FMR and FNMR we also see a decrease; $FMR = 0 \Rightarrow FMNR = 46.5\%$ and $FNMR = 0 \Rightarrow FMR = 56.7\%$.

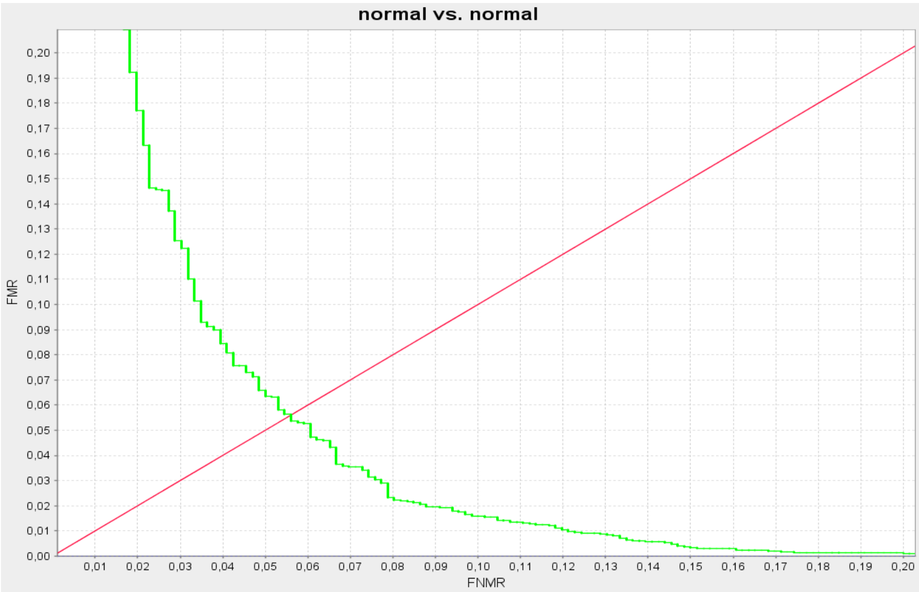


Figure 28: DET curve of normal versus normal, EER of about 5.6%.

normal1 vs normal1														
	Means				**Medians**				**Trimmed means**				**DTW**	
		N	A	A+N		N	A	A+N		N	A	A+N		N
Manh.	3.62	3.85	3.47	4.15	3.96	3.37	3.80	3.69	4.82	4.42	4.81	4.18	N/A	N/A
Eucl.	3.84	3.76	4.14	3.80	3.89	3.02	4.49	3.84	5.55	4.25	5.61	4.08	N/A	N/A
Deri.	4.03	**1.85**	4.69	**2.21**	6.19	3.84	7.36	4.83	14.3	11.0	15.4	10.4	N/A	N/A
DTW	**1.77**	3.90	**1.75**	3.46	**1.80**	3.76	**1.60**	3.90	**2.06**	4.51	**2.50**	4.45	7.06	9.40
normal2 vs normal2														
	Means				**Medians**				**Trimmed means**				**DTW**	
		N	A	A+N		N	A	A+N		N	A	A+N		N
Manh.	5.21	4.88	4.46	4.82	5.62	4.48	4.56	4.11	5.83	5.22	5.26	5.22	N/A	N/A
Eucl.	4.97	4.61	4.22	4.41	5.82	4.17	5.02	4.26	6.98	4.62	6.31	4.82	N/A	N/A
Deri.	5.50	**2.79**	5.08	**2.35**	7.74	5.24	7.76	5.42	16.2	13.4	17.2	13.8	N/A	N/A
DTW	**2.84**	4.38	**2.26**	4.65	**2.40**	4.47	**1.89**	4.65	3.24	4.51	2.61	4.77	8.24	11.0
normal vs normal														
	Means				**Medians**				**Trimmed means**				**DTW**	
		N	A	A+N		N	A	A+N		N	A	A+N		N
Manh.	8.79	8.78	7.87	8.98	9.37	8.30	8.07	8.53	10.0	9.77	8.92	10.0	N/A	N/A
Eucl.	9.41	8.12	8.23	8.80	9.70	8.70	8.42	8.91	10.5	9.38	9.35	9.37	N/A	N/A
Deri.	9.74	7.26	9.55	7.59	11.3	9.82	10.9	9.63	18.5	15.6	19.5	16.0	N/A	N/A
DTW	**6.25**	8.45	**5.99**	8.66	**6.80**	8.29	**5.86**	8.93	7.46	9.22	7.33	9.60	12.8	25.8

Table 7: A summary of normal vs normal with corresponding EER with the different setups. The horizontal line describes how the average cycles was created (N is the amplitude normalized, A is the acceleration adjusted, A+N is first adjust acceleration of the cycles within a walk, then create the average cycle and finally normalize the amplitude while the empty space is no modification) while the vertical line described what distance metrics that was used. The bold numbers indicate the best scores.

normal1 vs normal1 - 100 iterations						
	Means				Medians	
	N		A		A	
	EER	σ	EER	σ	EER	σ
Deri.	1.73	0.30	N/A	N/A	N/A	N/A
DTW	N/A	N/A	1.69	0.51	1.79	0.40
normal2 vs normal2 - 100 iterations						
	Means				Medians	
	N		A		A	
	EER	σ	EER	σ	EER	σ
Deri.	2.64	0.53	N/A	N/A	N/A	N/A
DTW	N/A	N/A	2.30	0.57	2.04	0.50
normal vs normal - 100 iterations						
	Means				Medians	
	N		A		A	
	EER	σ	EER	σ	EER	σ
Deri.	7.22	0.67	N/A	N/A	N/A	N/A
DTW	N/A	N/A	5.78	0.55	6.15	0.66

Table 8: A summary of the setups that gave the best EER, 100 iterations.

Fast walking

The next circumstance we will be looking at is when the participants were told to walk faster than normal. A result of the subjects increased velocity was a shorter cyclelength and fewer cycles per session. The reason for shorter cyclelengths is that we use shorter time between each step, while the reason for fewer cycles is that we do take longer steps when we walk fast compared to normal. The average cyclelength was in average 12.4% shorter than with normal walks, with an average of 96 samples, ($\sigma = 7.4$), the cyclelengths ranged from [80...110] samples. As mentioned before, each subject walked fast 8 times, this means $60 * 7 = 420$ genuine attempts and $60 * (60 - 1) * 7 = 24780$ fraudulent attempts. In Table 9 the three setups that gave the best EER is shown, the scores are calculated after 100 iterations. Common for all top three scores are that accelerations is adjusted and DTW is the distance metric, again did mean and median average cycle perform slightly better than the others, EER for $mean_{A+DTW}$ was 3.15% ($\sigma = 0.62\%$).

Slow walking

In contrast to the previous circumstance, the subjects were now told to walk slower than normal, this reduction of velocity lead to longer cyclelengths and more cycles per session. The average cyclelength was in average 25.6% longer than with normal walks, with an average of 137 samples, ($\sigma = 15.6$) meaning cyclelengths from [110...180]. Like with the fast walk the participants walked 8 times which gives us 420 genuine attempts and 24780 fraudulent attempts. In Table 9 we see the best setups after 100 iterations. Same as with fast walking what gave best results were acceleration adjustment and DTW as distance metric, and median average cycle did give us EER of 10.35% ($\sigma = 0.97\%$).

Circle walking

The final circumstance we will be looking at is when the subjects walked in a circle. As mentioned in Section 6.2.2 these sessions was divided into what direction the subject walked, Circle Left = CL, Circle Right = CR. The average cyclelength for CL was 113 samples ($\sigma = 7.8$), while for CR it was 112 ($\sigma = 7.9$). So based solely on the cyclelength there was hardly any difference between CL and CR. Furthermore did we see that the cyclelength was just barely higher than for normal walking. For both directions the subject did walk 4 times, meaning we had $60 * 3 = 180$ genuine attempts and $60 * (60 - 1) * 3 = 10620$ fraudulent attempts. Furthermore we got best EER when the average cycle had acceleration adjustment and the applied distance metric was DTW. The best scores after 100 iterations for circle walking are given in Table 9. Let us first look at CL, we see that mean average cycle gave us an EER of 2.96% ($\sigma = 0.75\%$) while the others are slightly higher. For CR we see that median and trimmed means average cycles gave us the best results, $Median_{A+DTW}$ gave an EER of 5.78% ($\sigma = 0.96\%$).

Other circumstances			
	Means A	**Medians A**	**Trimmed means A**
Fast	**3.15**	3.29	3.65
Slow	10.77	**10.35**	11.55
CL	**2.96**	3.13	3.70
CR	6.78	**5.78**	6.35

Table 9: A short summary of the other circumstances compared with themselves with the the setups that produced the best EER, A is the acceleration adjustment. The bold scores indicate the best EER. In all cases is DTW our distance metric.

6.2.3 Comparison of different circumstances

Beside comparing the different circumstances with themselves another part of our analysis has been to compare the non-standard circumstances with normal walking. This has been done in two ways, first we have simply compared normal walking with circumstance X. Meaning that we use a normal walk as a template and compare this against all recordings from circumstance X. The second approach was to create one multi-template which consisted of 5 templates, one template from each circumstance. We will look further into these approaches in the following subsections.

Normal vs circumstance X

This comparison was done by doing it almost the exact way as explained above. We randomly chose a normal walk as the template and then we compared this template against all the inputs in the circumstance we were considering. Again we got the best results when using mean and median with acceleration adjusted as averaging methods to create the average cycle. But in contrast to our results above we got even better results when the amplitudes were normalized. The results can be seen in Table 10, and as we see the best distance metric was when we took the derivative of the signal and applied Manhattan distance on that. The scores are however very poor so we tried to tackle this challenge in another way in our next approach.

Multi-template

By reviewing the results we got when we compared each circumstances to itself, we see that all circumstance produce good EER's. We used this fact to come up with an idea where we create a multi-template; one template for each circumstance and then use the rest of our data as inputs. So in total each subject has 36 samples under various circumstances, from these 36 samples we randomly choose one recording from each circumstance to be included in the multi-template. This leaves us with 31 input samples

Normal vs circumstance X								
	Means A+N				**Medians A+N**			
	CL	CR	Fast	Slow	CL	CR	Fast	Slow
Eucl.	19.00	24.40	20.40	36.61	17.11	24.80	20.86	37.00
Deri.	21.00	23.02	19.81	36.17	**16.54**	**19.52**	**19.12**	**33.51**
DTW	19.41	25.51	21.65	37.39	18.33	25.50	21.33	38.00

Table 10: A summary of EER showing the two setups that produces the best scores, mean- and median amplitude normalized (N) and acceleration adjusted (A). Distance metric used are Euclidean, derivative with Manhattan and DTW. The best scores are in bold face.

that are compared against all multi-templates, which gives us $60 * 31 = 1861$ genuine attempts and $60 * (60 - 1) * 31 = 109740$ fraudulent attempts. So all inputs are compared against all 5 templates within the multi-templates, and after the distance score table is calculated we must decide FMR and FNMR. This is done in this following way:

- **When looking at genuine attempts:**
  - False rejection: if all five circumstances within a multi-template are above the threshold.
  - Good acceptance: if at least one value within a multi-template is below the threshold.
- **When looking at fraudulent attempts:**
  - False acceptance: if at least one value within a multi-template is below the threshold.
  - Good rejection: if all five values within a multi-template are above the threshold.

So in our approach we only use the minimum of the 5 distances. There are however many more ways this could have been done, e.g. use at the sum or products of distances, include more minimum distances, look at the maximum distance, etc. [60]. Results from this approach can be seen in Table 11. The scores have been calculated from 100 iterations using mean and median average cycles with acceleration adjusted and DTW as distance metric. $mean_{A+DTW}$ gave us an EER of 5.04% ($\sigma = 0.26\%$). But as we see from the table when taking the standard deviation into account do $mean_{A+DTW}$ perform almost the same as $median_{A+DTW}$. A DET curve for one of the iterations can be seen in Figure 29.

Multi-template				
	Means A		Medians A	
	EER	σ	EER	σ
DTW	5.04	0.26	5.23	0.31

Table 11: A summary of EER showing the two setups that produces the best scores, mean- and median with acceleration adjusted (A). Distance metric used was DTW.

6.3 Sub-experiment

In this section will we explain how the long-term experiment was analyzed and what results it yielded.

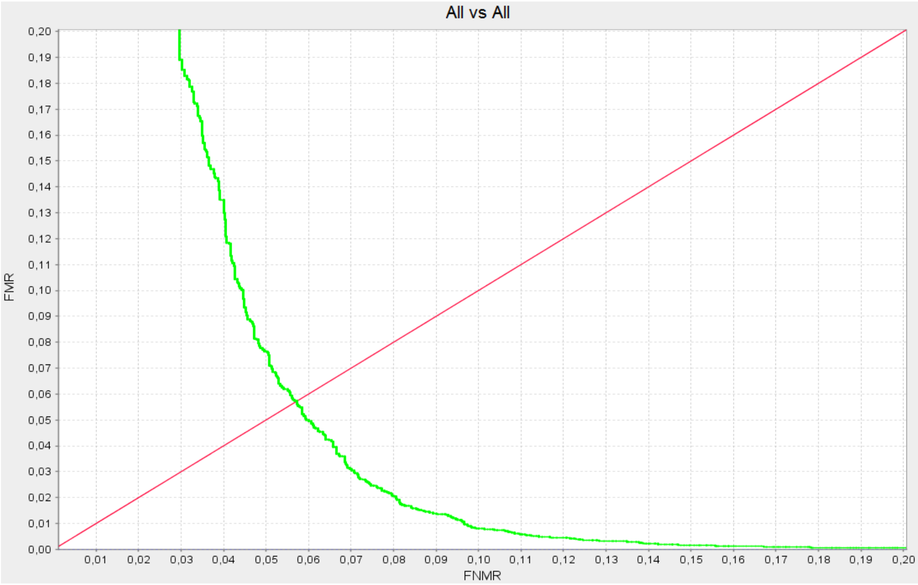


Figure 29: A DET curve from one iteration when using a multi-template to compare all vs all using $mean_{A+DTW}$, EER of about 5.7% in this case.

6.3.1 Long-term

Analysis method: As explained in Chapter 5, the long term experiment consisted of 5 participants. During a 2 months period they walked 40 sessions each, and each session consisted of walking 4 times in the morning and 4 times in the evening. All walking was done in a normal manner. This setup therefore gave 4*40=160 morning walks and 160 evening recordings per subject. In our analysis we will not calculate DET curves with its EER, FMR and FNMR like in the main experiment, we will instead only look at the genuine attempts per subject. The reason for this is that the goal of this experiment is to see how stable a persons gait is during a longer period of time. We are therefore not interested in comparing person A with person B. In this experiment will we look at sessions recorded at different days. Unfortunately were we not able to have 1 day interval between all sessions, due to the fact that the participants volunteered at their own will and could not always be present. We did however keep record of when the session were recorded so when we refer to different intervals between session it is in fact the correct interval. In total there were about 60 days between the first and the last session. We will look at the following:

- **Morning vs morning:** We will see how stable a persons walking is and this is first done by calculating distance scores from the morning walks. This is done by comparing each morning session with every other morning session. Let $m_{i,j}$ denote recording j in morning session i, where $1 \leq j \leq 4$ and $1 \leq i \leq 40$. The various distance scores are given by $d_{i,j,k,l} = dist(m_{i,j}, m_{k,l})$, where $i \neq k$ since we do not want to compare recordings from the same day. For each comparison we will compute all possible distance, meaning all 4 recordings from morning session i will be compared to all 4 recordings from morning session j, meaning 16 different distances. The distance scores are ordered according to the number of days between the two sessions. Let

D_t be the set of distances between any two recordings on two morning sessions that are exactly t days apart. Note that we can determine how many days two morning sessions are apart because we kept record of when a session was recorded and that each set D_t contains a multiple of 16 values. We will use the average value and the standard deviations of the sets D_t for our further analysis.

- **Evening vs evening:** This is the exact same as with morning walks, except that we consider evening walks instead.

In the two previous settings we know that all circumstances, except the day, are the same and we therefore get an indication of stability. Another interesting thing we will look into is if there are any differences between morning and evening walks. It is likely to believe that in the morning, one is more fresh compared to the evening where one might be more fatigued after a long day. We will look further into this in the following way:

- **Morning vs evening (same day):** Since we collected data both in the morning and evening, we can also compare morning walks with evening walks. This is done in a more simple way than explained above. Each morning walk will be a template and checked against the 4 sessions of evening walks done at the same day, meaning that we will get 4*4=16 scores per day. This is repeated for all days and in the end we will compute the average and standard deviation. In this case we are comparing “fresh walking” to “fatigued walking”.
- **Evening vs next morning:** Similar to morning vs evening, only that we compare the evening walks with the consecutive morning walk done the next day, again we make sure of that the days really are consecutive. In this case we are comparing “fatigued walking” to “fresh walking”.

Results: We chose to use one of the setups that produced the best scores for normal walking, median average cycles with acceleration adjusted and DTW as our distance metric. The first part of the analysis is to look at morning vs morning walks and evening vs evening walks. In Figure 30 we see the interval values plotted on a graph, a long with the standard deviation.

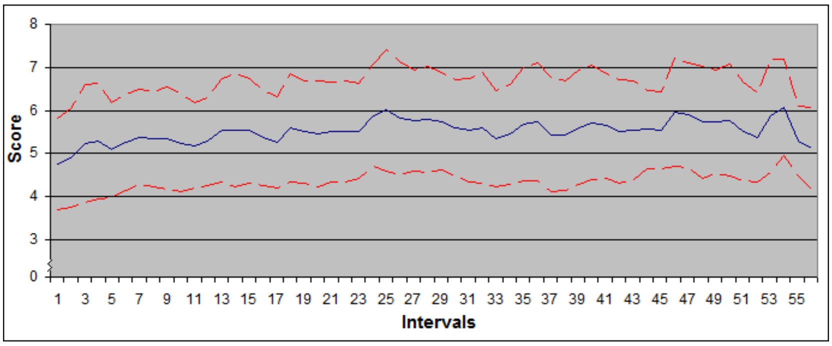
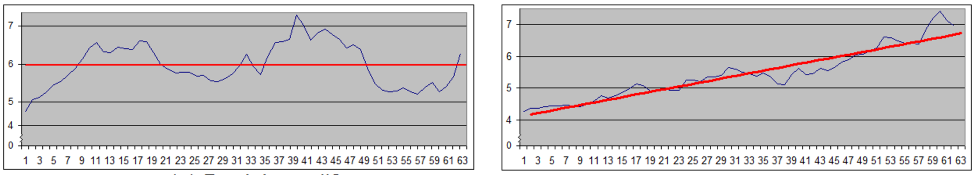


Figure 30: A graph showing the average score from all intervals for participant #1, the dotted line indicate the standard deviation.

By simply looking at the graph it seems to be rather stable, but in order to be sure we must apply some statistical tools. We will apply linear regression in order to calculate a linear function ($y = \alpha + \beta x$), and we are particular interested in the slope value (β). We define our hypothesis in the following way:

- H_0: $\beta = 0$
- H_1: $\beta > 0$

Since we are looking whether a gait is stable then the slope value of the regression line, must the confidence interval of the slope contain the value 0, meaning that H_0 ($\beta = 0$) is true. If this is not the case we must reject out null hypothesis which means that we have a $\beta > 0$ which implies that the gait is not stable. Note that we are restricting us to a one-tailed test, so we are not considering $\beta < 0$ since it is not reasonable to believe that the distance should decrease. After linear regression has been applied, we calculate p-value and a 95% confidence interval. In Table 12 we show a summary of the results for the morning walks. As we see all participants except #2 have a significant increase, meaning that we will reject H_0. In Figure 31 we show the graphs of those with lowest and highest p-value. As we see from the figure, partcipant #5 has a much higher slope than participant #2. In Table 13 the same information is shown for evening walks, and in this case all p-values are lower that our 5% threshold, so we reject H_0 for all participants.



(a) Participant #2 (b) Participant #5

Figure 31: (a) Showing the development of participant #2 and the regression line, (b) Showing the development of participant #5 and the regression line.

Linear regression for morning walks			
	Linear expression ($\alpha + \beta x$)	p-value	95% CI of the slope
Participant #1	5.311768 + 0.006322x	0.0011	<0.002361,0.010283>
Participant #2	5.960942 + 0.004097x	0.4747	<-0.00794,0.008457>
Participant #3	4.119434 + 0.008006x	4.80^{-06}	<0.004694,0.011317>
Participant #4	6.034012 + 0.038853x	1.48^{-09}	<0.027683,0.050023>
Participant #5	4.132324 + 0.041609x	2.29^{-32}	<0.038084,0.045135>

Table 12: A summary of some statistics results applied on the morning walks.

The second part of the analysis is to compare morning with evening and vice versa. From each subject we got one average score for both cases, an overview can be seen in Tables 14-15. As we see from the results there are hardly any difference between the average scores when comparing morning with evening and evening with the next morning, both the averages and the standard deviations are similar.

Linear regression for evening walks			
	Linear expression $(\alpha + \beta x)$	p-value	95% CI of the slope
Participant #1	$4.852997 + 0.017660x$	2.11^{-13}	<0.013894,0.021427>
Participant #2	$5.746234 + 0.057361x$	2.03^{-27}	<0.051336,0.063386>
Participant #3	$4.076442 + 0.014515x$	9.12^{-12}	<0.010997,0.018034>
Participant #4	$6.138825 + 0.024404x$	6.14^{-13}	<0.018946,0.029862>
Participant #5	$4.567204 + 0.014468x$	1.11^{-11}	<0.010940,0.017800>

Table 13: A summary of some statistics results applied on the evening walks.

Morning vs evening (same day)		
	Average	Standard dev.
Participant #1	5.96	0.90
Participant #2	7.79	1.26
Participant #3	4.54	0.71
Participant #4	7.49	1.44
Participant #5	5.45	0.74

Table 14: A summary of the average and standard deviation for all participants when comparing morning with evening.

Evening vs morning (next day)		
	Average	Standard dev.
Participant #1	5.89	0.76
Participant #2	7.87	1.13
Participant #3	4.40	0.49
Participant #4	7.39	1.39
Participant #5	5.53	0.67

Table 15: A summary of the average and standard deviation for all participants when comparing evening with the consecutive morning session.

7 Problems and discussion

In this chapter we will first discuss problems and issues that have occurred during this project, then we will review the results and finally have a discussion of gait as a biometric feature.

7.1 Problems

In this section we will look at some problems encountered during this project, we will start with different aspects of the step detection algorithm, namely estimation of cyclelength, detecting the first step, the w-shape between steps and finally correct detection of local extremes. After that will we will move on to the results from experiment.

7.1.1 Estimation of cyclelength

The first crucial phase in our algorithm is to correctly estimate the cyclelength as this forms the basis for the rest of the step detection. Since we are looking at both extreme types of walking, fast and slow, we had to make this part as robust as possible. By analyzing the data we discovered that cyclelength could range from [80...180] depending on the person and situation. Since this range is rather huge, and most important $\frac{180}{2} > 80$, it means that it is possible to have two fast walking cycles in the range of one slow walking cycle. As we remember from Section 4.3.2 do we extract a subset in the middle of the recording and compare this baseline against other subsets extracted after (and before) the baseline. To avoid confusion with distance scores we get when we compare cycles will we call the scores we get from comparing subsets subset scores. In Figure 32 we see that based solely on the subset scores we can choose both the cyclelength to be 80 samples or 160 samples. In this particular case we are looking at a slow walk, thus the correct cyclelength is 160 samples. In order to reduce the number of wrongly detected cyclelengths we applied several checks in cases where the cyclelength was estimated to be below $\frac{180}{2}$. Checks we applied were mainly to look at the difference, S, between the local maximum and the local minimum acceleration values within a cycle. With slow walking we normally have $S < 1.5$, sometimes even below 1.0, while with fast walking we often have $S > 1.5$. In addition we also looked at the consecutive minimum subset score, if this was significantly lower than the one we first estimated we chose this new subset score as seen in Figure 32.

7.1.2 Detecting the first step

An important part of the average cycle algorithm is to correctly detect steps. With correct we mean that a step starts/ends in the short valley between two peaks, compare with Figure 13 on Page 24 and Figure 20 on Page 34. When we are looking at the traditional way of detecting steps, starting from the beginning of the walk, this task is not always trivial. A problem that occurs is that if the first step is not correctly detected, this leads to a domino effect where all other steps are also not detected correctly. Steps are in this case shifted half a cycle. One of the reasons we can incorrectly detect the first step is that when we start walking we normally use a few steps to reach our normal velocity rhythm.

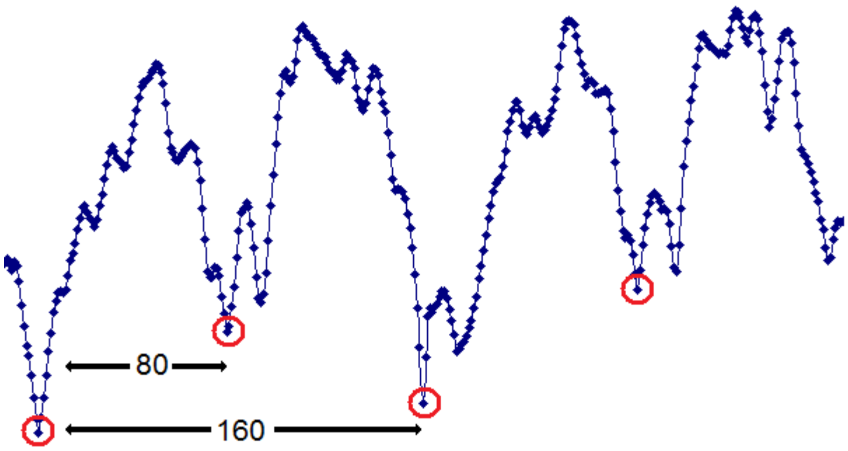


Figure 32: This graph shows the subset scores when comparing an excerpt of a gait signal with consecutive subsets. When estimating cyclelength we see that during a slow walking cycle, we detect two subset scores that would indicate fast walking cycles.

While we are gaining speed we sometimes reach the normal acceleration when we are in the middle of a step. This can lead to an incorrect identification of the beginning of a step. This is illustrated in Figure 33. In [10] they solved this problem by detecting left and right steps and then they used correlation on these parts, but we are instead, as mentioned earlier, using doublesteps. In our variant of the average cycle algorithm we overcame this problem by starting the step detection in the middle of our collected data instead of at the beginning. We also had to include several checks in order to get better results when looking at non-standard circumstances and especially slow walking. The problem is also reduced when we do cyclical rotation when calculating the distance score.

7.1.3 "W-shape" between peaks

Normally we have a "V-shape" between the peaks where a step ends/starts. In some cases the "V-shape" is more like a "W-shape", this is however not a problem if one of the sides is always lower than the other, a problem occurs when this is not true, see Figure 34(a). When creating an average cycle this error will cause the cycles to be shifted in one direction, see Figure 34(b). A possible way to overcome this is to choose the side which is closest to the estimated cyclelength, or always choose the same side. In our project we introduced the aligning of maximum peaks, by using this aligning we minimize the problem what side of the "W-shape" the algorithm choose. In the figure we however see that we have a "M-shape" on the first maximum peak which would give us the same problem as we have with "W-shapes". This is however not a problem since both align the last maximum which is easier detected, but we also apply a check to be assure of that we have correctly detected the last peak of the "M". This check is basically to look at the consecutive next local maximum from the first local maximum we detect. If the consecutive local maximum is significantly lower (less than 80%) and more than 15 samples ahead it is an indication that we have correctly identifed the maximum point, otherwise we have most likely detected the first peak of the "M". We could also apply this check when we have a "W-shape", but since we are aligning maximum points this

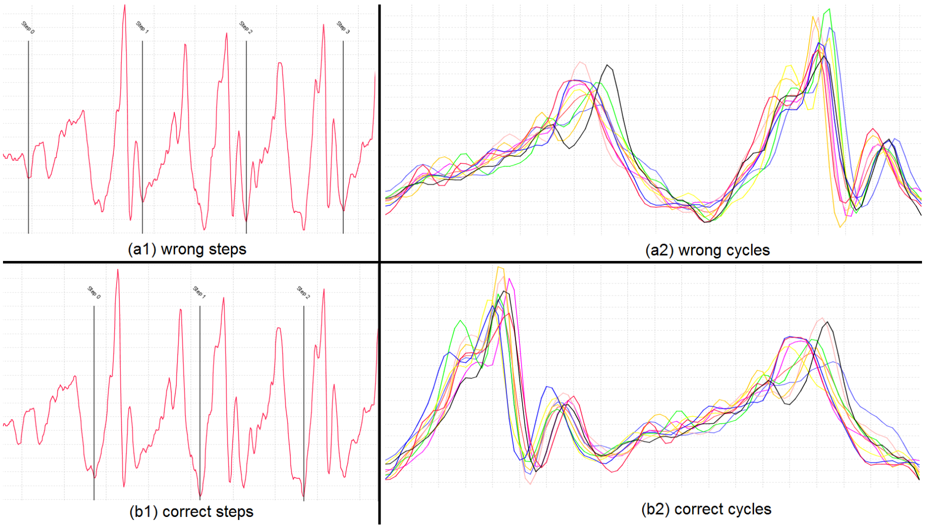


Figure 33: (a1) Showing the beginning of a walk, where the first step is incorrectly detected thus resulting that the rest of the steps are wrong. (a2) Shows the corresponding cycles, as we see the cycles are very similar, but does not follow the "correct" cycle-template. (b1) Shows how the first step should have been detected leading to correct steps and cycles (b2).

has not been done.

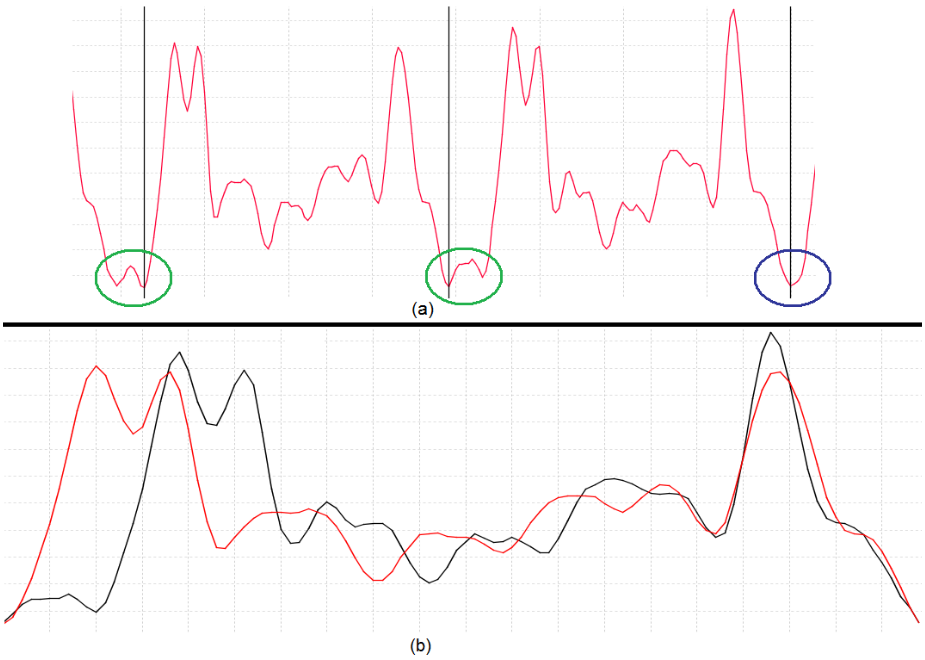


Figure 34: (a) Showing that the step detection detects steps at different sides of the "W-shape" between the peaks. The first two green circles indicate the "W-shape" while the last blue circle is the "V-shape". (b) The result of the step detection, one curve is shifted in one direction thus having an impact on the final average cycle when using regular averaging (y = sample n with sample n in all cycles).

7.1.4 Detecting correct extremes

An issue with our option to split a cycle into 4 parts when calculating distance scores, A-B, B-C, C-D, D-E (see Figure 14 on Page 26), is that sometimes detecting correct extremes can be a problem. In Figure 35 this is illustrated, as we see do all the minimum points, identified by a circle, have similar values. But circle 1 is far away from circle 2 and 3. In this particular example circle 2 would have been chosen since it has the lowest value, but in other sessions of the same person the area where circle 1 is, would have been the chosen value. One could imagine that we could force the have the minimum point, C, to be as close as the middle of the cycle as possible, but for some persons the correct C was actually earlier. A possible way to solve this is to not apply any additional checks when creating the template, meaning that we simply choose the minimum value that is lowest as C. When we then compare inputs against this template we choose the local minimum point that is closest to point C in the template to be the C-point in the input.

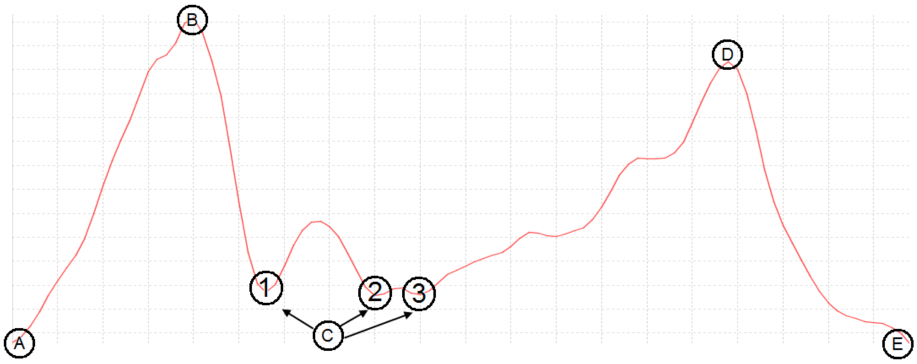


Figure 35: A problem when we shall find which extremes to choose. In this case both circle 1,2 and 3 do have close values. A problem occur when circle 1 is lower on some walks while circle 2/3 are lower for other walks of the same person.

7.1.5 Unnatural first walk

By looking at average cycles and EER's it seems to be a trend that the first walk often is the one that deviates the most from the other walks. The reason for this is that for many subjects this was their first meeting with this kind of experiment, they were therefore maybe nervous or unsecure on how they should walk. This unsecureness naturally lead to that the average cycle from the first session was a bit different than the average cycles from the other sessions when they might have been more comfortable. In a real life scenario it could therefore be desirable to have the person walk a few times before the actual template is created. If we only look at the first session of normal walks, and use one of the setups that performed best, mean with acceleration adjusted and DTW as distance metric, we achieved in average 1.69% EER and had 300 genuine attempts and 17700 fraudulent attempts. If we choose to omit the first sample from the database, the average went down almost 25% to 1.27% EER but now we had only 240 genuine attemps and 14160 fradulent attempts. By looking at all the normal walking, with the first walk for both sessions omitted, the EER goes down barely 10%, from 5.78% to 5.19%. This unsecureness does not only apply to the first time they participate, but also when they are being told to walk at a different velocity it would often have been desirable to have had a “test-run” to achieve a natural fast or slow walk. By reviewing this information we could choose to omit the first session for all experiments, but then we would have a low amount of sessions, especially for the circlewalks.

7.2 Discussion

In this section will we first discuss the results we got from the experiments, then will we end the discussion with some security evaluation of gait as a biometric feature.

7.2.1 Results from the experiments

We will first discuss results from the normal sessions, then will we move on to the other circumstances and comparison of all circumstances, and finally the long-term experiment.

Normal walking

In Section 6.2.2 we saw that both the first session of normal walking and the second session gave extremely good results compared to similar approaches, compare with Table 2 on Page 17. We got EER’s between 1.5 and 2.5%, when we combined these session the EER’s went up to between 5 and 6%. A reason for this depreciation is of course due to more input samples, but the main reason is that some people walked slightly different the second time compared to the first time, see Figure 36. The subject illustrated in the figure had very stable walking when we look at each session separately, meaning that all average cycles are more or less alike for each of the sessions. So the difference between the two sessions is not based on some irregular average cycles.

There could be several reasons for the differences between sessions, as mentioned above, many felt it strange when they were told to walk in a normal way, especially when they knew they were being measured. Other reasons could have been change of footwear and clothing, but as mentioned earlier did all participants wear the same footwear both times. Another reason could be the fact that they performed the experiment on different days. This supports also what we saw in the results of the long-term experiment, that some people have a more unstable gait than others, meaning that when they come back some days later their gait is slightly different. But we also saw this difference between sessions when the whole experiment was performed on one day. This might be related to the fact that some people felt it very unnatural and hard to walk normal when they are told to do so. It is also reasonable to believe that the unnatural feeling might have disappeared after a few walks, meaning that the second session of normal walking is more “normal” than the first session.

If we look at the results from the manually adjusted cycles we see that our automatic algorithm performs rather good, see Table 16. The results are even better when we take into consideration that we had to slightly adjust our algorithm so that it also gave good results when looking at other circumstances, see the next subsection. The main adjustments we had to do was to increase the interval when estimating the cyclelength and add more checks when detecting starting location. We also see with the manually detection that we have the same increase in EER when we include both sessions, meaning that some participants simply did walk in a slightly different way the second time.

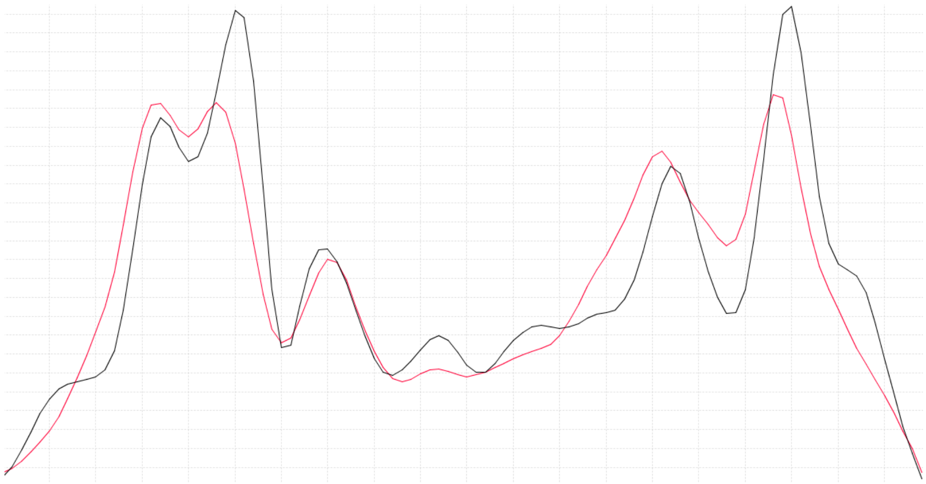


Figure 36: Two averages of the average cycles created in the two normal sessions. As we see both the general shape and acceleration values are similar, but the difference is that significant that it will lower the EER when combining both sessions.

	Automatic	Manually
Normal1	1.69%	0.65%
Normal2	2.04%	0.90%
All normal	5.78%	4.0%

Table 16: An overview of the best EER's from normal walking, comparing the automatic algorithm versus the manually detected steps.

Other circumstances

By looking at the other circumstances separately we see that they all gave good results. Both walking left in a circle and walking fast gave an EER around 3%, while walking right and slow resulted in a bit higher EER, about 6% and 10% respectively. When comparing the results obtained automatically with the results from the manually step detection, see Table 17, we see that both Circle Right (CR) and slow has potential to give better results than what we got from the automatic processing. But as mentioned in Section 7.1.1, our algorithm must be robust enough to handle different circumstances so instead of specializing it to fit a cirucmstance perfectly we try to get an as good as possible overall performance. By reviewing the manual results there is at least no doubt that all the circumstances have the potential to be used in a real life scenario. If we look closely at the difference between manually and automatic we see that all circumstances except CR has a difference of a factor of about 2. With CR we have a factor of almost 12, this indicates that there is a substantial difference between CR and CL. This implies that the placement have an affect when we walk around a corner. We can also see that we can draw the same conclusion as Vildjiounaite et. al. did in [18], it is easier to recognize a person when he walks fast compared to walking slow.

Comparing normal to other circumstances

Comparing normal walking to other circumstances turned out to be a very difficult task. As the results from 6.2.3 show, we achieve about 20% EER when comparing normal with

	Automatic	Manually
Circle Left (CL)	2.96%	1.30%
Circle Right (CR)	5.78%	0.49%
Fast	3.15%	2.78%
Slow	10.35%	4.80%

Table 17: An overview of the best EER's from non-normal walking, comparing the automatic algorithm versus the possibility to manually detect steps.

other circumstances. It was also hard to see any common features among the different circumstances. We do of course have shorter cyclelength and higher amplitudes for fast walking and longer cyclelengths and lower amplitudes for slow walking. It was however, not given that an x% shorter/longer cyclelength equals x% higher/lower amplitudes. For some participants we could clearly see that the fast and slow walking had exactly the same general shape as normal walking, but for others it was entirely different. For circle walking it was even harder to find any common features, when looking at manually detected cycles (where we know that we have correctly identified steps) it could seem that walking left in the circle was more similar to normal than walking right. This is of course highly related to where the sensor is located, in our case it was on the left hip. But even for the manually detected cycles the EER was in the same range as the automatic detected cycles, between around 17% for normal vs CL and 20% for normal vs CR. These facts may indicate that the different circumstances actually are that distinct and this led to our multi-template approach with one template for each circumstance.

As mentioned in Section 6.2.3 we had an approach where we had one template for each circumstance, and then the input was matched against all these templates. The results on this approach was very good, the EER was as low as 5.04% when using mean average cycles with acceleration adjusted, see Table 18. We also see that the automatic EER is just slightly higher than the EER for manually detected cycles which was 3.48% ($\sigma = 0.19$). In a real life scenario it is reasonable to believe that this setup would yield even better results since we have the possibility to throw away sessions where we get an indication of irregular cycle detection. If we we look in particular at slow walking where we had a rather poor EER due mostly to irregular average cycles we could in a real life scenario choose to omit sessions that produce irregular average cycles. This could be done by either having a threshold for the average score we get when we compare all the cycles detected within that session. Another option could be to have a "template" for what that particular walk should include, e.g. all the interesting points in Figure 14 on Page 26. In a real life scenario this approach would have been possible since we have continuous authentication.

	Automatic	Manually
Multi-template	5.04%	3.48%

Table 18: An overview of EER from the multi-template setup, comparing the best score from automatic and manually step detection. The EER the calculated from mean average cycles with acceleration adjusted and DTW as distance metric.

Long-term experiment

If we start looking at the results when looking at the same time of day, meaning morning vs morning and evening vs evening, we can draw a clear conclusion. From Section 6.3.1 we see that all participants except one had a p-value lower than 0.05, meaning that we would reject our null hypothesis of the slope being equal to 0. This means that we do have a significant increase in the slope, in other words the walk is not stable. By reviewing these results it could imply that gait is not persistent enough to be a good biometric feature, however there is a solution to this. Instead of having an enrollment phase at the first day, and then always compare against the template created that day we could introduce a template that is dynamic. Meaning that the template could evolve as time passes by. Such an implementation is however out of the scope for this project. Another important aspect is that even if we had to reject our null hypothesis for almost all cases it would still take a considerably long time before the distance would get high. By reviewing our distance score tables a threshold of 8.0 seems to be a fair threshold, in Table 19 we see how many days it would take in order to get a distance of above 8.0 (assuming it would follow our linear regression). As we see do the number of days vary, even among morning and evening walks, especially participant #2.

Number of days below a threshold of 8.0		
	Morning	Evening
Participant #1	425	178
Participant #2	7819	39
Participant #3	484	270
Participant #4	50	76
Participant #5	92	237

Table 19: A table showing the number of days the distance would still be below a threshold of 8.0.

When looking at the other aspects of the results from this sub-experiment, meaning morning vs evening and evening vs morning, the scores are almost identical. This means that it is not a great difference whether we compare morning with evening or evening with the consecutive morning. These averages should in principle be comparable with the average distances we get when we are looking at the "same time of day" distances with 1 day intervals. For all cases however we can see an increase, between 24% up to 57%, this is shown in Table 20. In the table we have taken the average of the morning vs morning and evening vs evening distances with 1 day interval, and the average of morning vs evening (same day) and evening vs morning (next day). This increase could mean that there is actually a difference between morning and evening walks, in other words do the gait slightly variates over time, but also within a day. This idea is also supported by our findings in the previous table (Table 19). By looking at the scores it could seem that it follows a cyclic pattern, meaning that the distance from the morning template increase during the day, but decrease again the next morning. The same pattern applies to the evening walks.

7.2.2 Comparison of algorithm

We have also done a comparison with the average cycle algorithm, the way it was implemented by Gafurov et al. in [46]. There are several similarities with our experiment and Gafurov et al.'s experiment, a comparison table is given in Table 21. As we see there are

1 day interval vs morning and evening			
	1 day interval	evening ↔ morning	difference
Participant #1	4.62	5.93	+28%
Participant #2	4.98	7.83	+57%
Participant #3	3.62	4.47	+24%
Participant #4	5.07	7.44	+47%
Participant #5	4.18	5.49	+31%

Table 20: A table showing how much the scores from different time of day increase compared to the 1 day interval scores.

many similarities, it is also worth mentioning that the experiment was carried out in the exact same location.

Comparison of our experiment with Gafurov et al. [46]		
	Our approach	Gafurov et al.
Sensor	MR100	MR100
Sensor placement	Left hip	Right hip
Participants	60	100
Genuine attempts	360, 720	400
Algorithm	Average Cycle	Average Cycle
EER	1.69%, 5.78%	13%

Table 21: A table showing the main differences between our experiment and the one done by Gafurov et al. in [46]. Only looking at normal walking, both one session and two sessions.

The only main difference is how the data was analyzed. The main differences between the two approaches are:

- **Step detection algorithm**: Gafurov et al. do detect steps from the beginning of the recording, the first step is detected to be the minimum point on the first 250 samples. Then the next step is detected by skipping 100 samples and look for the minimum point in 10 samples before and after this new point. This is repeated until no more steps are detected.
- **Average cycle creation:** Gafurov et al. took the median of the cycles detected after they had been normalized to 100 samples.
- **Distance metric:** Gafurov et al. used the Euclidean distance metric.

As we see are our approach more sophisticated, the main differences are (when considering what achieved best):

- **Step detection algorithm**: We first estimate the cyclelength, then get an indication of what values the minimum points are supposed to have. We start the step detection in the middle of the recorded data.
- **Average cycle creation:** We align the cycles detected and adjust for acceleration after normalizing to 100 samples.
- **Distance metric:** We introduce DTW as distance metric and cyclical rotate the input cycle.

In order to illustrate the differences did we implement the algorithm used in [46], for normal walking did we achieve 20% EER for one session 25% for both sessions combined. The non-standard circumstances did perform similar with EER's ranging from 20% to 35%. By reviewing the results we have achieved we can firmly say that we have improved several aspects of the average cycle algorithm.

7.2.3 Security evaluation

An important part when biometrics is used as a way to enhance security is of course to be aware of the vulnerabilities of gait to attacks. This is beyond the scope of this project, but nevertheless is it an important part of gait authentication in general so we will briefly discuss this. In [46] Gafurov et al. presented an analysis of the minimal effort impersonation attack and the closest person attack on gait biometrics. With minimal effort mimicking they mean that the attacker only has common knowledge of the system (in [46]: hip movement recorded by an accelerometer), limited time to study the victim and limited number of trials to impersonate the victim. Their results indicated that the chances of accepting impostors by minimal effort mimicking in the hostile scenario (meaning that the attackers really tries to walk as a victim) is not higher than the chances of impostors in the friendly scenario (meaning no attempt to impose as a victim). This means that whether the attacker uses his natural gait or tries to impersonate a person has little or none affect on the performance of the system. These results might be the opposite of what one should believe since we are trying to mimic a person, but here lies an advantage with gait. "The human gait is a complex process that involves nervous and musculo-skeletal systems" [46]. Gafurov et al. also explain the result with that when a person is asked to walk like someone else he is being given a restriction to walk differently than his normal habituated gait. This leads to a failure to produce natural gait patterns that matches the person being mimicked. Other factors are velocity, short versus long steps, height and weight of the person etc. It is however worth mentioning that even if minimal effort mimicry does not help, Gafurov et al. mentioned that if the attacker knows the closest person in the database, it can be a serious threat to the authentication system. This means that he can just walk normally and still be accepted as a genuine user. In this setting we can also mention, although not yet researched, that it is reasonable to believe that some people are better to mimic others (called wolves) while some people are easier to mimic than others (called sheep). These issues of wolves and sheep are however a common problem among different biometric feature, e.g. voice and signature, so by the light of the results in [46] gait does have a promising future as a possible way to enhance the security.

8 Conclusion

During the work on this thesis we have looked at many interesting aspects of the biometric feature gait. More precisely have we looked into different aspects of the human gait. In addition to normal walking have we looked at walking with different velocity and the impact of not walking in a straight line. In order to get information about these circumstances did we perform a large experiment with 60 participants which produced in total over 2000 recordings. In order to analyze the data have we tried several approaches and variants of existing algorithms. We did focus on the average cycle method which has been successfully applied in earlier works [11, 41, 46]. We tried different combinations of acceleration directions, noise reduction algorithms, step detection approaches, averaging methods, etc. Based on the results from the main experiment can we clearly say that we have achieved very good results, especially when we look at the number of participants and genuine attempts. As mentioned in Chapter 3 did Gafurov et al. achieve an EER of 13% with 100 participants in [46]. If we compare with the EER we got from normal walking, as described in Table 8 on Page 55 did we achieve an EER of less than 2% when looking at one session, less than 6% when combining the two sessions. Even if we had fewer participants than Gafurov et al., did we have more recordings per participant, almost twice the number of gait sequences when looking at the combined sessions. We even implemented the algorithm used in [46], and as mentioned in Section 7.2.2, did we get an EER above 20% for normal walking when it was applied to our dataset. We feel that the main reason for our improvement of EER lies in our adjustments of the step detection and average cycle creation, but also the choice of introducing DTW as the distance metric. The good results we got when we compared each circumstance with itself shows that the human gait can in fact be used to recognize a person in other settings than only normal walking. If we look at the manually detected cycles we see that all circumstances achieved an EER below 5%, while with automatic detection did we have EER below 11%. We have improved the different aspects of the average cycle algorithm, but as the results from manual step detection show are there still room for further improvements.

The only previous research on different circumstances has been in [11] where they looked at the effect of wearing a backpack and in [18] where they mentioned that the participants walked with different velocity. In [18] no further analysis on the different velocities is given. In this thesis we have isolated different circumstances and have seen to what extent it was possible to still recognize users. The step detection algorithm described in Section 4.3.2 was first adapted to fit normal walking. But since we looked at other circumstances as well did we make some adjustments in order to get as good scores as possible among the different circumstances. These adjustments were mainly the detection of starting location and estimation of cyclelength as described in Sections 4.3.2 and 7.1. Another aspect of this thesis has been to see whether we could find common features among the different circumstances and adapt the other circumstances to be more comparable with normal walking. This did turn out to be quite challenging as it seems that the human gait simply is different when we walk under different circumstances, as mentioned in Section 7.2.1. We did however see that the one can rather clearly see from the cyclelength whether a person walks fast or slow, and as a result of different velocity did the corresponding amplitudes differ as well. But it seems that there is no strict relation between a change in velocity and increase in amplitude. Nevertheless did we get better results when we adjusted for the acceleration and normalized cycles to a fixed length.

Since we achieved very good EER's when the different circumstances were compared with themselves and the fact that they seemed rather distinct, did we try a solution with a multi-template, as described in Section 6.2.3. This multi-template solution, which consisted of one template from each circumstance, did yield good and promising results. We achieved an EER of 5% when we compared all our data, meaning almost 2200 recordings. Based on the approach chosen in this thesis it seems that this is the best solution when looking at different circumstances.

In this thesis we have also for the first time looked at whether the human gait remains stable under a longer period of time under the same circumstances. This was done by performing a small scale experiment where the participants walked both in the morning and the evening over a 2 months period. From this experiment we could get an indication if gait remains stable by comparing circumstances that are identical. In order to determine this, we compared sessions with different intervals, as described in Section 6.3.1, and by applying statistical methods we could clearly see that the distances increased as time passed by. As a result of this is it reasonable to believe that the we would need a dynamic template, meaning a template that changes slightly as time passes by. Due to the experiment setup were we also able to compare morning with evening walks and vice versa. This was done by looking at distances between consecutive sessions and compare the average distance with the distance we had when we compared sessions with one day intervals. We could see that there was hardly any difference whether you compared morning with evening or vice versa, but we did however see a significant increase in the distance in all cases which indicates that there is a difference between morning and evening walks. But due to the low number of participants can we only look at these results as an indication, in order to draw a final conclusion would we need an experiment at a larger scale.

9 Further work

Since gait authentication by using wearable sensors is a rather new area there is a lot of aspects that would need further research. By looking at topics that are directly related to this thesis it is natural to include more circumstances, like e.g. walking up- or downhill, but it would also be interesting to look at different environments like the surface, e.g. walking on grass, gravel, sand, etc. In this thesis we have only recorded gait signals by attaching a sensor on the left hip, but all other possible variations of this can of course tested. As with all other experiments there is always the possibility to include even more participants and recordings per participant. This is specially related to the long term experiment, in order to draw valid conclusion if the gait is permanent, an experiment with many more participants is needed. Another aspect that is related to the stability of the gait is to perform an exhaustion experiment. It could be interesting to see how the gait develops as a person is getting exhausted.

Another interesting aspect that could be looked into is automatic recognition of the different circumstances. This could also be extended to automatic detection of walking in general, meaning that we could differentiate between a person sitting, standing still, etc. from the parts where the actual walking occurs. A possible way of doing this could be to perform continuous recordings during a longer period. It could also be interesting to look at the cycles separately, meaning that we compare all cycles detected in a walk against a template. With this setup we would get information from every cycle that is detected. It is likely to believe that we would get good results by pursuing this idea since the general shape of the cycles is similar within a walk. It could also be interesting to do more research on the different parts of a gait cycle to get a better understanding of which parts that are distinct among persons, like illustrated in Figure 27 on Page 52.

We have during this project also shown that there is still room for improvements on the different parts of our algorithms, and especially cycle detection as this is a very important part. The main problem in this thesis when including both slow and fast walking is to determine the cyclelength correctly, due to problems explained in Section 7.1.1. A possible improvement could also be to create a generic template for a gait cycle and force the cycles that are detected to have characteristics given by the template. This approach would be possible by introducing PCA[1].

After reading this thesis we realize that processing signals can be done in incredibly many ways. All phases in our algorithms can be adjusted in one or another way. These variations could be everything from other noise reduction methods, different calculations of the resultant vector to include other algorithms like e.g. PCA or ICA[2].

From the long term experiment we got an indication that the gait slightly changes over time, a possible way to overcome this change would be to introduce a dynamic template. How the template would be adapted is also an interesting topic that would need further work.

[1]Principal Component Analysis (PCA): http://en.wikipedia.org/wiki/Principal_components_analysis

[2]Independent Component Analysis (ICA): http://en.wikipedia.org/wiki/Independent_component_analysis

Another interesting research would be to include the gait recognition algorithms in mobile devices that already contains an accelerometer, e.g. the iPhone from Apple Inc.[3] contain a 3-axis accelerometer or in the Nintendo game console WII[4] which also contains an accelerometer.

[3]Apple iPhone: http://www.apple.com/iphone/
[4]Nintendo WII: http://www.nintendo.com/wii/

Bibliography

[1] Bours, P. Authentication course, IMT 4721, Gjøvik University College. Reader for IMT 4721 course, September 2006.

[2] Jain, A., Ross, A., & Prabhakar, S. January 2004. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20.

[3] Zhang, R., Vogler, C., & Metaxas, D. 2004. Human gait recognition. In *CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 1*, 1–8, Washington, DC, USA. IEEE Computer Society.

[4] Delac, K. & Grgic, M. 2004. A survey of biometric recogition methods. In *Electronics in Marine, 2004. Proceedings Elmar 2004. 46th International Symposium*, 184–193.

[5] Cunado, D., Nixon, M. S., & Carter, J. N. 2003. Automatic extraction and description of human gait models for recognition purposes. *Computer Vision and Image Understanding*, 90(1), 1–41.

[6] BenAbdelkader, C., Cutler, R., & Davis, L. 2002. Motion-based recognition of people in eigengait space. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*, 254–259.

[7] Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., & Bowyer, K. 2005. The humanid gait challenge problem: Data sets, performance, and analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(2), 162–177. Member-Sudeep Sarkar and Member-P. Jonathon Phillips and Member-Isidro Robledo Vega and Fellow-Kevin W. Bowyer.

[8] Liu, Z. & Sarkar, S. 2004. Simplest representation yet for gait recognition: Averaged silhouette. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4*, volume 4, 211–214, Washington, DC, USA. IEEE Computer Society.

[9] Suutala, J. & Röning, J. June 14-15 2004. Towards the adaptive identification of walkers: Automated feature selection of footsteps using distinction-sensitive LVQ. In *Proceedings of International Workshop on Processing Sensory Information for Proactive Systems (PSIPS 2004)*, 61–67, Oulu, Finland.

[10] Mäntyjärvi, J., Lindholm, M., Vildjiounaite, E., Mäkelä, S.-M., & Ailisto, H. 2005. Identifying users of portable devices from gait pattern with accelerometers. *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, 2, 4.

[11] Gafurov, D., Snekkenes, E., & Bours, P. 2007. Gait authentication and identification using wearable accelerometer sensor. In *Automatic Identification Advanced Technologies, 2007 IEEE Workshop on*, 220–225.

[12] Gafurov, D., Helkala, K., & Søndrol, T. April 2006. Gait recognition using acceleration from MEMS. In *Availability, Reliability and Security, 2006. ARES 2006. The First International Conference on*, 432–439.

[13] Pousttchi, K. & Schurig, M. 2004. Assessment of today's mobile banking applications from the view of customer requirements. In *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, 10.

[14] Kim, Y., Yoon, J., Park, S., & Han, J. 2004. Architecture for implementing the mobile government services in korea. In *ER (Workshops)*, 601–614.

[15] Clarke, N. L. & Furnell, S. M. 2006. Authenticating mobile phone users using keystroke analysis. *Int. J. Inf. Secur.*, 6(1), 1–14.

[16] Su, Q., Tian, J., Chen, X., & Yang, X. 2005. A fingerprint authentication system based on mobile phone. In *AVBPA*, 151–159.

[17] Lee, Y., Seo, C., Lee, J., & Lee, K. Y. 2003. Speaker verification system for pda in mobile-commerce. In *Human.Society@Internet 2003*, 668–674.

[18] Vildjiounaite, E., Mäkelä, S.-M., Lindholm, M., Riihimäki, R., Kyllönen, V., Mäntyjärvi, J., & Ailisto, H. 2006. Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices. In *Pervasive*, Fishkin, K. P., Schiele, B., Nixon, P., & Quigley, A. J., eds, volume 3968 of *Lecture Notes in Computer Science*, 187–201. Springer.

[19] Hong, L. *Automatic Personal Identification Using Fingerprints*. PhD thesis, Michigan State University, 1998.

[20] Ross, A., Shah, J., & Jain, A. K. March 2005. Towards reconstructing fingerprints from minutiae points. *Proceedings of SPIE Conference on Biometric Technology for Human Identification II*, 5779, 68–80.

[21] Sandström, M. Liveness detection in fingerprint recognition systems. Technical report, Linköping University, Department of Electrical Engineering, June 2004.

[22] Fawcett, T. ROC graphs: Notes and practical considerations for researchers. Technical report, HP Laboratories, MS 1143, 1501 Page Mill Road, Palo Alto, CA 94304, March 2004.

[23] Lee, L. 2003. Gait analysis for classification. In *AI Technical Report*, volume 014, 112. Massachuttets Institute of Technology-artificial Intelligence Laboratory.

[24] Marks, A. A. 1905. Manual of artificial limbs. New York.

[25] Ayyappa, E. 1997. Normal human locomotion, part 1: Basic concepts and terminology. *Journal of Porsthetics and Orthetics*, 9, 10–17.

[26] Chan, C. W. & Rudins, A. 1994. Foot biomechanics during walking and running. *Mayo Clinic Proceedings*, 69, 448–461.

[27] Søndrål, T. Using the human gait for authentication. Master's thesis, Gjøvik University College, 2005.

[28] Murray, M. 1967. Gait as a total pattern of movement. In *American Journal of Physical Medicine*, volume 46(1), 290–332.

[29] Johannson, G. 1973. Visual perception of biological motion and a model for its analysis. In *Perception and Phychophysics*, volume 14, 201–211.

[30] Niyogi, S., Niyogi, S., & Adelson, E. 1994. Analyzing and recognizing walking figures in xyt. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, Adelson, E., ed, 469–474.

[31] Yoo, J., Nixon, M., & Harris, C. 2002. Extracting gait signatures based on anatomical knowledge. In *Proceedings of BMVA Symposium on Advanced Biometric Technologies*, 2.

[32] Cunado, D., Nixon, M. S., & Carter, J. N. 1997. Using gait as a biometric, via phase-weighted magnitude spectra. In *AVBPA '97: Proceedings of the First International Conference on Audio- and Video-Based Biometric Person Authentication*, 95–102, London, UK. Springer-Verlag.

[33] Little, J. & Boyd, J. 1998. Recognizing people by their gait: The shape of motion. In *Journal of Computer Vision Research, The MIT Press*, volume 1, 2–32.

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

[35] Wang, L., Tan, T., Hu, W., & Ning, H. 2003. Automatic gait recognition based on statistical shape analysis. *Image Processing, IEEE Transactions on*, 12(9), 1120–1131.

[36] Nixon, M. S., Carter, J. N., Grant, M. G., Gordon, L. G., & Hayfron-Acquah, J. B. 2003. Automatic recognition by gait: progress and prospects. *Emerald*, 23, 323–331.

[37] Perry, J. 1992. *Gait Analysis. Normal and Pathological Function*. SLACK Incorporated.

[38] Middleton, L., Buss, A. A., Bazin, A., & Nixon, M. S. 2005. A floor sensor system for gait recognition. In *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*, 171–176, Los Alamitos, CA, USA. IEEE Computer Society.

[39] Orr, R. J. & Abowd, G. D. 2000. The smart floor: A mechanism for natural user identification and tracking. In *CHI '00: CHI '00 extended abstracts on Human factors in computing systems*, 275–276, New York, NY, USA. ACM Press.

[40] Fujimoto, T., Higasi, Y., Horiuchi, F., Kadoya, R., Sekine, M., & Tamura, T. 2001. Evaluation by accelerometry of walking pattern before falls in hemiplegic patients. *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, 2, 1153–1154.

[41] Gafurov, D., Snekkenes, E., & Buvarp, T. E. 2006. Robustness of biometric gait authentication against impersonation attack. In *OTM Workshops (1)*, 479–488.

[42] Rong, L., Zhiguo, D., Jianzhong, Z., & Ming, L. July 2007. Identification of individual walking patterns using gait acceleration. In *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, 543–546.

[43] Nixon, M. & Wagg, D. 2004. On automated model-based extraction and analysis of gait. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, 11–16.

[44] BenAbdelkader, C., Cutler, R., & Davis, L. 2002. Stride and cadence as a biometric in automatic person identification and verification. *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*, 5, 357–362.

[45] Ailisto, Lindholm, Mantyjarvi, Vildjiounaite, & Makela. mars 2005. Identifying people from gait pattern with accelerometers. In *Biometric Technology for Human Identification II. Edited by Jain, Anil K.; Ratha, Nalini K. Proceedings of the SPIE*, Jain, R., ed, volume 5779 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, 7–14.

[46] Gafurov, D., Snekkenes, E., & Bours, P. 2007. Spoof attacks on gait authentication system. *Information Forensics and Security, IEEE Transactions on*, 2(3), 491–502.

[47] Alvarez, D., Gonzalez, R., Lopez, A., & Alvarez, J. 2006. Comparison of step length estimators from weareable accelerometer devices. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, 5964–5967.

[48] Boyd, J. E. & Little, J. J. 2003. Biometric gait recognition. In *Advanced Studies in Biometrics*, M.Tistarelli, J. & E.Grosso, eds, volume 3161, 19–42. Springer Berlin / Heidelberg.

[49] Buvarp, T. E. Hip movement based authentication. Master's thesis, Gjøvik University College, 2006.

[50] von Tscharner V. & B., G. june 2003. Gender dependent emgs of runners resolved by time/frequency and principal pattern analysis. *Journal of Electromyography and Kinesiology*, 13, 253–272.

[51] Gafurov, D., Helkala, K., & Søndrol, T. 2006. Biometric gait authentication using accelerometer sensor. *Journal of Computers*, 1(7), 51–59.

[52] Motion recording. http://www.motion-recording.com/ [last visited june 2008].

[53] Sánchez, G., Lladós, J., & Tombre, K. 2002. A mean string algorithm to compute the average among a set of 2d shapes. *Pattern Recogn. Lett.*, 23(1-3), 203–213.

[54] Wong, R. K. & Chhieng, V. M. *Advances in Databases: Concepts, Systems and Applications*, chapter Adaptive Distance Measurement for Time Series Databases, 598–610. Springer Berlin / Heidelberg, 2007.

[55] Salvador, S. & Chan, P. 2004. Fastdtw: Toward accurate dynamic time warping in linear time and space. In *In Proc. KDD Workshop on Mining Temporal and Sequential Data*, 11.

[56] Keogh, E. J. & Pazzani, M. J. 2001. Derivative dynamic time warping. In *in First SIAM International Conference on Data Mining, (Chicago, IL, 2001)*. unknown.

[57] Chhieng, V. M. & Wong, R. K. *Advances in Databases: Concepts, Systems and Applications*, chapter Adaptive Distance Measurement for Time Series Databases, 598–610. Springer Berlin / Heidelberg, 2007.

[58] Wang, L., Wang, L., Wang, K.-Q., & Xu, L.-S. 2004. Recognizing wrist pulse waveforms with improved dynamic time warping algorithm. In *Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on*, Wang, K.-Q., ed, volume 6, 3644–3649 vol.6.

[59] Neter, Wasserman, & Whitmore. 1993. *Applied Statistics*. Allan and Bacon Inc., 4th edition edition.

[60] Kittler, J., Hatef, M., Duin, R., & Matas, J. 1998. On combining classifiers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(3), 226–239.

A Participant Agreement Declaration

Participation in acquisition of gait biometric data in MSc project

I am participating in the acquisition of gait biometric data on a voluntarily basis. The gait biometric data is collected by the use of an accelerometer, and will only be used in this MSc project. This project is done by Kjetil Holien and is related to information security.

With my signature I confirm the following:

1. I have been informed in oral and written form about the content and purpose of the collected data that is in relation to my person.
2. My data will only be used to serve this purpose. In case of future experiments, a new permission should be given.
3. I allow that gait biometric data from me are collected.
4. The data will not be displayed in possible future publications on this experiment, unless I give my permission.
5. I have been informed that I can reject to sign the agreement.
6. I have been informed that I can request to receive insight in the collected data at any time.
7. I know that I can withdraw my participation anytime I want without giving any explanation and all data collected from we will be deleted permanently.

All data will be deleted respectively the link between the data and my name will be destroyed as soon as it is not necessary to maintain it. This will happen as the research experiment has been completed in June 2008.

First name - family name:________________

Gjøvik, date________________ signature:________________

B Code

In this appendix will we list a bit more detailed and Java oriented pseudocode than listed in Section 4.3.2. We will look into the main parts of step detection, average cycle creation and our DTW algorithm.

B.1 Step detection

```
Step detection()
//estimate cyclelength
$\gamma$ = estimateCyclelength()
//find minimum details
$\alpha$ = indicateMinimumValue()
//detect starting location
$S_{org}$ = getStartingLocation()
//do actual step detection
REPEAT
    $S_0$ = $S_{org}$
    IF(forward) i = 1
    ELSE i = -1
    REPEAT
        IF(forward) $S_i$ = $S_{i-1} + \gamma$
        ELSE $S_i$ = $S_{i-1} - \gamma$
        $O_i$ = min[$S_i - \frac{\gamma}{10}, S_i + \frac{\gamma}{10}$]
        IF($O_i$ is in the first or last third of search area)
            search additional $\frac{\gamma}{20}$ samples for a lower value
        //possible out of synch, try to resynch.
        IF($Value(O_i)$<$\alpha$) $O_i$ = search for first location with value >$\alpha$
        steps[] = add the identified breakpoint $O_i$ in the array
        IF(forward) i = i + 1
        ELSE i = i - 1
    UNTIL finished with the direction
UNTIL both forward and backward steps are detected
```

Pseudocode B.1: Step detection

```
estimate cyclelength()
baseline[] = extract 70 samples from the middle of the collected data
REPEAT
    REPEAT
       comparison[] = extract 70 samples from the search area
       score[].add(getDistance(baseline,comparison))
    UNTIL checked 300 comparisons
    minimas[] = extract local minimas from score[]
    minimas[] = remove local minimas that are closer than 80 samples,
            retain the point with lowest score
    γ_direction = minimas[1] - minimas[0]
UNTIL both forward and backward are checked
RETURN average of both estimated cyclelengths
```

Pseudocode B.2: Cyclelength estimation

```
Indicate minimum value()
β = min[X/2, (X/2) + γ]
α = β + 20%
RETURN α
```

Pseudocode B.3: Indication of minimum value

```
Detect starting location()
M = max[X/2, (X/2) + γ]
//apply 4 checks to decide if we shall go backward or forward from M
DO
    IF(M_back − M) < (M − M_forw) backward++
    ELSE forward++
    IF(S_back < S_forw) backward++
    ELSE forward++
    IF(# of local maximums between S_back to M <
     # of local maximums between S_forw to M_forw) backward++
     ELSE forward++
     IF(S_back < 10) forward++
     IF(S_forw < 10) backward++
     //if neither was below 10 samples
     IF (S_back − M < M − S_forw) backward++
     ELSE forward++
     IF(forward == backward) M = maximum point one γ further
WHILE forward == backward
IF(backward > forward) RETURN S_back
ELSE RETURN S_forw
```

Pseudocode B.4: Starting location detection

B.2 Average cycle creation

```
Normalize cycles()
1. // normalize all cycles identified (N)
2. FOR(i = 0. . . N)
3.     A[i][0] = O[i][0] and A[i][99] = O[i+1][0]
4. END FOR
5. // normalize one cycle, containing n samples
6. $\lambda_i$ = get adjustmentfactor
7. FOR(j = 1. . . n-1)
8.     A[i][j] = normalized and acceleration adjusted sample
9. END FOR

Align cycles()
10. REPEAT
11.     $O$ = the set of normalized cycles
12.     $AVG_B$ = average location of first maximum
13.     $AVG_D$ = average location of last maximum
14.     //align all cycles, first maximum
15.     FOR(i = 0. . . N) $F_i = O_i + (B_i - AVG_B)$
16.     //align all cycles, last maximum
17.     FOR(i = 0. . . N) $L_i = O_{i+1} + (D_i - AVG_D)$
18.     $\beta_{first}$ = avg of all cycles compared to all cycles using $F$-set
19.     $\beta_{last}$ = avg of all cycles compared to all cycles using $L$-set
20.     $\beta_{none}$ = avg of all cycles compared to all cycles using $O$-set
21.     IF $\beta_{first}$ has lowest score:
22.        $O - set$ = normalized $F$-set and repeat the process
23.     ELSE IF $\beta_{last}$ has lowest score:
24.        $O - set$ = normalized $L$-set and repeat the process
25. UNTIL $\beta_{none}$ has the lowest score

Skip irregular cycles()
26. REPEAT
27.     FOR(i = 0. . . N)
28.        distance[i] = avg.dist. from $cycle_i$ to all the other cycles
29.     END FOR
30.     average = the average of all distances
31.     IF max[distance[0],. . . ,distance[N]] > average+15%:
32.        remove $cycle_i$ and repeat the process
33. UNTIL no more cycles are removed
```

Pseudocode B.5: Average cycle creation

B.3 Dynamic Time Warping

```
costDel()
1.RETURN 0.5

costIns()
2. RETURN 0.5

costSub(double template, double input, double S)
3. RETURN |(template - sample)|/S

DTW()
4. n = template.size()+1
5. m = input.size()+1
6. M = new double[n][m]
7. //initialize matrix
8. FOR (i = 1...n) M[i][0] = costDel()*(i)
9. FOR (j = 1...m) M[0][j] = costIns()*(j)
10. M[0][0] = 0
11. //find S, q = template, c = input
12. S = max(q_1, ..., q_n, c_1, ..., c_m) - min(q_1, ..., q_n, c_1, ..., c_m)
13. FOR (i = 1...n)
14.     // for each node, search the predecessor with least cost
15.     FOR (j = 1...m)
16.        C_temp = costSub(q[j-1], c[i-1], S)
17.        C_s = M[i-1][j-1] + C_temp // substitution
18.        C_i = M[i][j-1] + costIns() //insertion
19.        C_d = M[i-1][j] + costDel() //deletion
20.        M[i][j] = min{C_s, C_i, C_d}
21.     END FOR
22. END FOR
23. RETURN M[n][m]
```

Pseudocode B.6: Dynamic Time Warping algorithm

C Scores

In this appendix will we give the scores from all setups we have created. There are still some setups that have not been tested, and some setups from Table 7 where only tested with normal walking, thus not listed here. All scores are based on 10 iterations, but we will first give some information about how the scores are represented.

Noise reduction methods: For manually detected cycles did we test both Moving Average (MA) and Weighted Moving Average (WMA). Since there was hardly any difference between MA and WMA did we use only WMA for the automatic detection.

Pre-processing methods: Normalization to 100 samples occurred for all setups except when we used DTW to create average cycle. Normalization to 100 samples when using DTW to create average cycle is noted with “DTW100”. When the normalized cycles also have been adjusted for acceleration it is noted by “A”.

Average cycle creation: We used four different methods to create average cycles, namely: means, medians, trimmed means and Dynamic Time Warping (DTW).

Post-processing methods: After the average cycle had been created we also had the possibility to normalize the amplitudes so that it ranged from 0 to 1, this is noted by “N”.

Distance metrics: We tested several distance metrics: Euclidean distance (E), Manhattan distance (M), Dynamic Time Warping (DTW). We also took the Derivative (D) of the signal and applied M to this signal. In addition did we try to split the average cycle into four different parts as explained in Section 6.1.1, this is noted by “X_A”, where X is a given distance metric. A final variation of this was to include a penalty to the distance score when the input had a different length than the template, this is noted by “X_A_P”. The penalty is to take the difference in percentage and multiply it with 3, this new product is then multiplied with the distance score. By doing it this way large differences are more severe than low differences, which is what we want.

Circumstances: abbreviations for the different circumstances are as following:

- Circle, walking left = CL
- Circle, walking right = CR
- Fast walking = F
- Normal walking from the first session = N1
- Normal walking from the second session = N2
- All normal walkings combined = N
- Slow walking = S

We will list things in the following way, in the title of the table will we show how the average cycle is created:

Averaging method_Noise reduction method_Pre-processing_Post-processing, λ means no pre- or post-processing.

The cells will then contain the average EER for the circumstances:

Distance metric || *circumstance: EER* $\pm$ *standard deviation*.

We will first show scores from manually detection, then the automatic detection.

C.1 Manually detected cycles:

Means_MA_λ_λ			
M:	CL: 4.36% ±0.69	CR: 3.79% ±0.21	F: 6.09% ±0.77
	N2: 3.98% ±0.56	N1: 4.02% ±0.71	S: 9.70% ±0.97
M_A:	CL: 4.99% ±0.65	CR: 4.4% ±0.49	F: 6.35% ±1.71
	N2: 4.77% ±0.44	N1: 3.57% ±0.73	S: 9.1% ±0.58
M_A_P:	CL: 4.52% ±0.56	CR: 4.04% ±0.60	F: 7.09% ±0.38
	N2: 3.7% ±0.45	N1: 2.82% ±0.59	S: 7.97% ±0.69
E:	CL: 5.10% ±0.76	CR: 3.96% ±0.41	F: 5.83% ±0.61
	N2: 3.67% ±0.28	N1: 4.18% ±0.67	S: 9.39% ±1.01
E_A:	CL: 5.46% ±0.66	CR: 5.21% ±0.51	F: 7.67% ±0.50
	N2: 4.85% ±0.55	N1: 4.0% ±0.64	S: 9.23% ±0.66
E_A_P:	CL: 5.13% ±0.35	CR: 4.0% ±0.52	F: 6.90% ±0.86
	N2: 4.49% ±0.30	N1: 2.98% ±0.33	S: 6.9% ±2.03
D:	CL: 7.15% ±0.63	CR: 5.74% ±0.51	F: 6.98% ±0.68
	N2: 4.25% ±0.45	N1: 5.20% ±0.72	S: 12.4% ±0.92
D_A:	CL: 7.72% ±0.45	CR: 6.04% ±0.64	F: 8.28% ±0.43
	N2: 5.31% ±0.56	N1: 4.85% ±0.54	S: 11.3% ±2.12
D_A_P:	CL: 5.0% ±0.57	CR: 5.77% ±0.75	F: 6.64% ±1.07
	N2: 4.73% ±0.72	N1: 3.81% ±0.53	S: 9.52% ±0.67
DTW:	CL: 1.55% ±0.42	CR: 0.75% ±0.30	F: 3.14% ±0.56
	N2: 1.23% ±0.32	N1: 1.42% ±0.48	S: 5.75% ±0.52

Means_MA_A_λ			
M:	CL: 4.20% ±0.42	CR: 3.07% ±0.20	F: 5.11% ±0.61
	N2: 3.85% ±0.36	N1: 3.22% ±0.30	S: 9.12% ±0.67
M_A:	CL: 5.14% ±0.49	CR: 3.71% ±0.18	F: 4.89% ±0.73
	N2: 3.97% ±0.77	N1: 3.53% ±0.59	S: 7.77% ±0.75
M_A_P:	CL: 4.45% ±0.74	CR: 3.88% ±0.62	F: 5.82% ±0.30
	N2: 3.43% ±0.29	N1: 3.06% ±0.40	S: 7.32% ±0.44
E:	CL: 5.33% ±0.51	CR: 3.29% ±0.35	F: 5.15% ±0.68
	N2: 3.6% ±0.52	N1: 4.09% ±0.57	S: 9.41% ±0.91
E_A:	CL: 5.64% ±0.55	CR: 4.06% ±0.59	F: 6.31% ±0.33
	N2: 4.25% ±0.44	N1: 3.88% ±0.21	S: 8.05% ±0.82
E_A_P:	CL: 4.93% ±0.73	CR: 3.90% ±0.55	F: 6.04% ±0.25
	N2: 3.88% ±0.54	N1: 3.04% ±0.49	S: 7.38% ±0.64
D:	CL: 6.80% ±0.47	CR: 6.09% ±0.77	F: 6.32% ±0.54
	N2: 3.99% ±0.57	N1: 5.18% ±0.64	S: 11.9% ±1.01
D_A:	CL: 7.46% ±0.77	CR: 6.05% ±0.34	F: 7.75% ±0.77
	N2: 5.06% ±0.75	N1: 4.33% ±0.61	S: 11.0% ±0.62
D_A_P:	CL: 5.17% ±0.52	CR: 6.12% ±0.70	F: 6.34% ±0.64
	N2: 4.48% ±0.37	N1: 3.75% ±0.74	S: 9.63% ±0.64
DTW:	CL: 1.38% ±0.40	CR: 0.72% ±0.29	F: 2.78% ±0.53
	N2: 0.9% ±0.26	N1: 1.41% ±0.30	S: 5.38% ±0.55

Means_MA_λ_N			
M:	CL: 3.94% ±0.47	CR: 2.19% ±0.43	F: 4.31% ±0.73
	N2: 2.21% ±0.22	N1: 3.55% ±0.78	S: 8.37% ±0.90
M_A:	CL: 5.2% ±0.50	CR: 3.86% ±0.81	F: 5.38% ±1.04
	N2: 2.92% ±0.38	N1: 3.42% ±0.56	S: 8.45% ±1.97
M_A_P:	CL: 4.63% ±0.31	CR: 3.55% ±0.60	F: 5.88% ±1.59
	N2: 3.01% ±0.73	N1: 3.26% ±0.49	S: 8.99% ±0.58
E:	CL: 3.41% ±0.50	CR: 1.75% ±0.29	F: 4.63% ±0.92
	N2: 2.32% ±0.34	N1: 3.37% ±0.49	S: 8.19% ±0.61
E_A:	CL: 5.29% ±0.62	CR: 3.38% ±0.37	F: 4.2% ±0.58
	N2: 2.96% ±0.42	N1: 2.84% ±0.45	S: 8.67% ±0.84
E_A_P:	CL: 3.56% ±0.62	CR: 3.14% ±0.63	F: 4.94% ±1.50
	N2: 2.49% ±0.75	N1: 2.77% ±0.31	S: 8.52% ±0.63
D:	CL: 3.82% ±0.35	CR: 2.73% ±0.34	F: 4.20% ±0.56
	N2: 2.46% ±0.52	N1: 3.3% ±0.54	S: 10.6% ±0.90
D_A:	CL: 3.13% ±0.55	CR: 3.34% ±0.86	F: 4.68% ±1.36
	N2: 2.79% ±0.49	N1: 2.23% ±0.38	S: 9.95% ±0.51
D_A_P:	CL: 3.96% ±0.60	CR: 4.14% ±0.34	F: 5.81% ±0.30
	N2: 2.26% ±0.57	N1: 3.03% ±0.40	S: 10.7% ±0.80
DTW:	CL: 3.57% ±0.46	CR: 2.11% ±0.46	F: 3.65% ±0.63
	N2: 2.11% ±0.25	N1: 2.42% ±0.53	S: 7.61% ±0.47

Means_MA_A_N			
M:	CL: 4.10% ±0.63	CR: 2.13% ±0.38	F: 4.4% ±0.69
	N2: 2.19% ±0.18	N1: 3.53% ±0.51	S: 8.49% ±1.09
M_A:	CL: 5.22% ±0.68	CR: 3.86% ±0.65	F: 4.77% ±0.42
	N2: 2.85% ±0.35	N1: 3.23% ±0.31	S: 9.65% ±0.84
M_A_P:	CL: 4.22% ±0.44	CR: 3.19% ±0.72	F: 5.39% ±0.17
	N2: 3.0% ±0.58	N1: 3.34% ±0.44	S: 8.38% ±2.82
E:	CL: 3.78% ±0.44	CR: 1.46% ±0.33	F: 4.56% ±0.83
	N2: 2.37% ±0.28	N1: 3.18% ±0.65	S: 8.08% ±0.74
E_A:	CL: 4.68% ±0.57	CR: 3.39% ±0.39	F: 5.92% ±0.38
	N2: 3.0% ±0.44	N1: 3.34% ±0.25	S: 9.36% ±0.58
E_A_P:	CL: 4.21% ±0.77	CR: 2.23% ±0.29	F: 5.29% ±1.38
	N2: 2.36% ±0.40	N1: 2.96% ±0.38	S: 8.3% ±0.86
D:	CL: 3.86% ±0.45	CR: 2.78% ±0.35	F: 4.35% ±0.89
	N2: 2.37% ±0.36	N1: 3.71% ±0.61	S: 10.3% ±0.56
D_A:	CL: 3.42% ±0.66	CR: 3.97% ±0.66	F: 4.57% ±0.71
	N2: 2.92% ±0.79	N1: 1.8% ±0.35	S: 10.4% ±0.86
D_A_P:	CL: 3.96% ±0.68	CR: 3.92% ±0.37	F: 5.37% ±0.75
	N2: 2.94% ±0.57	N1: 2.86% ±0.72	S: 8.47% ±2.89
DTW:	CL: 3.91% ±0.27	CR: 2.41% ±0.39	F: 3.86% ±0.40
	N2: 1.98% ±0.25	N1: 2.51% ±0.44	S: 7.65% ±0.73

Medians_MA_λ_λ			
M:	CL: 4.81% ±0.57	CR: 3.98% ±0.27	F: 6.65% ±0.87
	N2: 4.47% ±0.49	N1: 3.88% ±0.63	S: 9.62% ±0.85
M_A:	CL: 4.99% ±0.58	CR: 4.75% ±0.62	F: 7.83% ±1.11
	N2: 5.26% ±0.65	N1: 4.77% ±0.44	S: 9.35% ±0.45
M_A_P:	CL: 4.93% ±0.61	CR: 5.11% ±0.93	F: 6.99% ±0.80
	N2: 5.35% ±0.55	N1: 4.76% ±0.66	S: 9.48% ±0.65
E:	CL: 5.56% ±0.34	CR: 4.48% ±0.32	F: 6.69% ±0.54
	N2: 4.18% ±0.65	N1: 4.67% ±0.57	S: 10.3% ±0.93
E_A:	CL: 6.86% ±0.89	CR: 5.92% ±0.47	F: 8.43% ±0.77
	N2: 5.89% ±0.53	N1: 5.40% ±0.84	S: 9.94% ±0.46
E_A_P:	CL: 6.23% ±0.68	CR: 5.65% ±0.50	F: 6.77% ±0.63
	N2: 5.03% ±0.56	N1: 4.6% ±0.49	S: 9.54% ±0.75
D:	CL: 8.96% ±0.63	CR: 9.82% ±0.48	F: 9.43% ±0.84
	N2: 7.14% ±0.51	N1: 6.98% ±0.59	S: 15.8% ±0.92
D_A:	CL: 10.3% ±0.52	CR: 11.7% ±0.68	F: 11.5% ±0.75
	N2: 9.49% ±0.66	N1: 8.26% ±0.71	S: 16.0% ±0.56
D_A_P:	CL: 8.32% ±0.86	CR: 8.88% ±0.84	F: 9.83% ±0.81
	N2: 8.20% ±0.63	N1: 7.51% ±0.48	S: 13.6% ±0.71
DTW:	CL: 1.77% ±0.49	CR: 0.55% ±0.29	F: 3.78% ±0.63
	N2: 1.16% ±0.21	N1: 1.44% ±0.37	S: 6.31% ±0.48

Medians_MA_A_λ			
M:	CL: 4.43% ±0.42	CR: 3.30% ±0.43	F: 5.26% ±0.69
	N2: 3.98% ±0.44	N1: 4.53% ±0.59	S: 8.77% ±0.42
M_A:	CL: 4.57% ±0.91	CR: 3.81% ±0.43	F: 6.47% ±0.63
	N2: 4.89% ±0.46	N1: 4.05% ±0.37	S: 8.45% ±0.53
M_A_P:	CL: 5.45% ±0.25	CR: 4.39% ±0.50	F: 6.64% ±0.92
	N2: 4.58% ±0.56	N1: 4.00% ±0.45	S: 9.03% ±0.72
E:	CL: 5.88% ±0.58	CR: 4.10% ±0.50	F: 6.13% ±0.88
	N2: 4.18% ±0.32	N1: 5.02% ±0.50	S: 9.96% ±0.75
E_A:	CL: 6.30% ±0.88	CR: 4.86% ±0.49	F: 6.42% ±0.57
	N2: 5.39% ±0.87	N1: 4.97% ±0.61	S: 9.55% ±0.51
E_A_P:	CL: 6.07% ±0.57	CR: 4.39% ±0.80	F: 6.31% ±0.68
	N2: 4.65% ±0.62	N1: 3.95% ±0.48	S: 9.15% ±0.53
D:	CL: 9.18% ±0.70	CR: 9.84% ±0.69	F: 8.49% ±0.63
	N2: 6.88% ±0.39	N1: 7.3% ±0.45	S: 14.8% ±1.18
D_A:	CL: 10.1% ±1.10	CR: 10.8% ±0.49	F: 10.2% ±1.33
	N2: 9.05% ±0.57	N1: 7.36% ±0.93	S: 14.9% ±1.00
D_A_P:	CL: 8.6% ±0.75	CR: 7.83% ±0.45	F: 9.2% ±0.64
	N2: 7.65% ±0.65	N1: 7.31% ±0.72	S: 14.2% ±0.70
DTW:	CL: 1.44% ±0.44	CR: 0.49% ±0.26	F: 3.0% ±0.47
	N2: 1.10% ±0.15	N1: 1.32% ±0.27	S: 5.49% ±0.39

Medians_MA_λ_N			
M:	CL: 3.92% ±0.78	CR: 2.18% ±0.40	F: 5.05% ±1.15
	N2: 2.30% ±0.50	N1: 3.67% ±0.93	S: 9.69% ±0.47
M_A:	CL: 4.59% ±0.44	CR: 3.81% ±0.74	F: 6.03% ±0.78
	N2: 3.71% ±0.35	N1: 3.81% ±0.32	S: 10.4% ±0.92
M_A_P:	CL: 4.55% ±0.41	CR: 4.06% ±0.78	F: 6.24% ±0.64
	N2: 3.57% ±0.73	N1: 4.61% ±0.40	S: 10.7% ±0.73
E:	CL: 3.96% ±0.25	CR: 2.32% ±0.62	F: 4.68% ±1.05
	N2: 2.23% ±0.35	N1: 3.61% ±0.46	S: 9.4% ±0.45
E_A:	CL: 4.28% ±0.63	CR: 4.09% ±0.55	F: 5.46% ±0.63
	N2: 3.90% ±0.37	N1: 4.25% ±0.86	S: 10.0% ±1.05
E_A_P:	CL: 4.06% ±0.59	CR: 3.74% ±0.90	F: 4.87% ±0.70
	N2: 3.82% ±0.55	N1: 4.13% ±0.77	S: 10.4% ±0.88
D:	CL: 6.83% ±0.61	CR: 5.58% ±0.61	F: 6.07% ±0.84
	N2: 4.02% ±0.41	N1: 5.47% ±0.63	S: 14.3% ±0.56
D_A:	CL: 7.7% ±0.73	CR: 8.69% ±0.52	F: 6.99% ±0.78
	N2: 6.58% ±0.83	N1: 6.25% ±1.08	S: 16.2% ±0.86
D_A_P:	CL: 7.51% ±0.79	CR: 7.54% ±0.61	F: 7.51% ±1.07
	N2: 5.93% ±0.51	N1: 6.01% ±0.59	S: 14.3% ±0.73
DTW:	CL: 3.55% ±0.32	CR: 1.84% ±0.43	F: 3.84% ±0.58
	N2: 2.14% ±0.41	N1: 2.61% ±0.59	S: 9.39% ±0.40

Medians_MA_A_N			
M:	CL: 3.73% ±0.43	CR: 2.14% ±0.34	F: 4.42% ±0.57
	N2: 2.21% ±0.34	N1: 3.6% ±0.41	S: 8.87% ±0.71
M_A:	CL: 5.50% ±0.66	CR: 3.48% ±0.45	F: 5.69% ±0.74
	N2: 3.57% ±0.33	N1: 3.88% ±0.54	S: 9.99% ±0.77
M_A_P:	CL: 4.64% ±0.40	CR: 3.44% ±0.51	F: 6.37% ±0.95
	N2: 3.44% ±0.53	N1: 4.17% ±0.33	S: 11.0% ±0.62
E:	CL: 4.5% ±0.37	CR: 2.46% ±0.39	F: 5.11% ±0.72
	N2: 2.4% ±0.46	N1: 4.0% ±0.95	S: 9.27% ±0.73
E_A:	CL: 4.94% ±0.81	CR: 3.86% ±0.77	F: 6.22% ±0.77
	N2: 3.48% ±0.47	N1: 3.86% ±0.37	S: 9.85% ±0.65
E_A_P:	CL: 4.14% ±0.67	CR: 3.03% ±0.56	F: 5.59% ±0.98
	N2: 3.22% ±0.70	N1: 3.65% ±0.54	S: 10.0% ±0.50
D:	CL: 6.51% ±0.48	CR: 5.70% ±0.49	F: 6.61% ±0.66
	N2: 3.93% ±0.52	N1: 6.06% ±0.67	S: 13.8% ±0.96
D_A:	CL: 7.20% ±0.80	CR: 8.14% ±0.76	F: 7.43% ±0.83
	N2: 6.47% ±0.52	N1: 6.1% ±0.67	S: 15.5% ±0.84
D_A_P:	CL: 7.97% ±0.77	CR: 6.9% ±0.66	F: 7.77% ±0.70
	N2: 6.39% ±0.73	N1: 5.88% ±0.67	S: 15.3% ±0.55
DTW:	CL: 3.32% ±0.47	CR: 1.51% ±0.33	F: 4.07% ±0.86
	N2: 2.01% ±0.34	N1: 2.46% ±0.45	S: 8.37% ±0.81

Trimmed Means_MA_λ_λ			
M:	CL: 5.18% ±0.76	CR: 5.02% ±0.25	F: 7.66% ±1.00
	N2: 4.61% ±0.50	N1: 4.77% ±0.46	S: 10.1% ±0.57
M_A:	CL: 6.13% ±0.82	CR: 6.43% ±0.71	F: 8.76% ±0.95
	N2: 5.77% ±0.87	N1: 5.27% ±0.70	S: 10.5% ±0.52
M_A_P:	CL: 7.35% ±0.82	CR: 7.99% ±0.76	F: 8.35% ±0.91
	N2: 6.54% ±0.43	N1: 6.81% ±0.90	S: 11.6% ±0.68
E:	CL: 6.69% ±0.61	CR: 5.33% ±0.20	F: 8.05% ±0.87
	N2: 5.46% ±0.38	N1: 5.89% ±0.52	S: 11.4% ±0.69
E_A:	CL: 7.99% ±1.05	CR: 8.29% ±0.92	F: 9.71% ±0.72
	N2: 7.00% ±0.49	N1: 6.56% ±0.86	S: 11.8% ±0.79
E_A_P:	CL: 7.52% ±0.46	CR: 8.58% ±0.59	F: 9.22% ±1.14
	N2: 7.01% ±1.08	N1: 6.85% ±0.95	S: 11.8% ±0.84
D:	CL: 17.8% ±0.85	CR: 18.4% ±0.78	F: 18.2% ±0.77
	N2: 13.5% ±0.38	N1: 13.8% ±0.46	S: 23.2% ±0.92
D_A:	CL: 17.7% ±0.74	CR: 19.3% ±0.84	F: 18.5% ±0.76
	N2: 15.4% ±0.93	N1: 14.7% ±0.78	S: 21.4% ±0.97
D_A_P:	CL: 14.9% ±1.01	CR: 16.4% ±1.36	F: 15.0% ±1.07
	N2: 13.1% ±0.76	N1: 13.0% ±1.57	S: 20.8% ±1.10
DTW:	CL: 1.75% ±0.38	CR: 1.34% ±0.29	F: 3.38% ±0.74
	N2: 1.7% ±0.47	N1: 1.46% ±0.21	S: 6.68% ±0.68

Trimmed means_MA_A_λ			
M:	CL: 5.59% ±0.65	CR: 4.06% ±0.34	F: 5.97% ±0.50
	N2: 3.88% ±0.37	N1: 4.76% ±0.25	S: 9.89% ±0.66
M_A:	CL: 6.43% ±0.48	CR: 4.91% ±0.52	F: 6.76% ±0.70
	N2: 4.82% ±0.61	N1: 4.92% ±0.50	S: 9.46% ±0.83
M_A_P:	CL: 7.70% ±0.54	CR: 6.34% ±1.09	F: 8.01% ±0.49
	N2: 6.36% ±1.14	N1: 6.5% ±0.79	S: 11.1% ±0.38
E:	CL: 6.65% ±0.35	CR: 4.92% ±0.48	F: 6.70% ±0.73
	N2: 4.79% ±0.55	N1: 6.36% ±0.58	S: 11.2% ±0.91
E_A:	CL: 8.69% ±0.87	CR: 6.67% ±0.52	F: 7.73% ±0.67
	N2: 6.22% ±0.56	N1: 7.0% ±0.74	S: 11.1% ±0.61
E_A_P:	CL: 8.97% ±0.85	CR: 6.8% ±0.67	F: 7.8% ±0.73
	N2: 6.07% ±0.69	N1: 6.32% ±0.63	S: 11.5% ±0.44
D:	CL: 18.9% ±0.86	CR: 17.5% ±0.49	F: 15.8% ±0.64
	N2: 13.4% ±0.59	N1: 15.0% ±1.40	S: 22.3% ±0.89
D_A:	CL: 18.9% ±0.83	CR: 17.4% ±0.38	F: 16.7% ±0.91
	N2: 15.2% ±1.06	N1: 15.9% ±0.85	S: 20.6% ±0.53
D_A_P:	CL: 15.6% ±1.17	CR: 15.7% ±0.62	F: 14.6% ±0.95
	N2: 13.8% ±0.58	N1: 12.9% ±0.55	S: 19.9% ±0.96
DTW:	CL: 1.76% ±0.50	CR: 1.27% ±0.33	F: 2.86% ±0.99
	N2: 1.43% ±0.17	N1: 1.49% ±0.22	S: 6.09% ±0.53

Trimmed means_MA_λ_N			
M:	CL: 4.21% ±0.36	CR: 1.99% ±0.27	F: 5.14% ±0.79
	N2: 2.49% ±0.36	N1: 3.81% ±0.29	S: 9.30% ±0.68
M_A:	CL: 4.86% ±0.61	CR: 3.49% ±0.32	F: 6.27% ±0.66
	N2: 4.46% ±0.25	N1: 4.65% ±0.52	S: 10.5% ±0.79
M_A_P:	CL: 6.65% ±0.68	CR: 6.28% ±1.05	F: 7.68% ±0.42
	N2: 5.39% ±0.58	N1: 6.27% ±0.66	S: 12.8% ±1.16
E:	CL: 4.85% ±0.55	CR: 2.4% ±0.27	F: 5.65% ±0.82
	N2: 3.14% ±0.31	N1: 3.92% ±0.43	S: 9.93% ±0.68
E_A:	CL: 5.80% ±0.64	CR: 4.77% ±0.35	F: 7.27% ±0.80
	N2: 5.59% ±0.75	N1: 5.47% ±0.30	S: 11.3% ±1.01
E_A_P:	CL: 6.46% ±0.73	CR: 5.68% ±0.92	F: 7.09% ±0.74
	N2: 5.64% ±0.77	N1: 5.65% ±0.45	S: 12.0% ±0.66
D:	CL: 15.3% ±0.81	CR: 14.0% ±0.77	F: 12.3% ±1.01
	N2: 12.2% ±0.64	N1: 11.2% ±0.75	S: 21.7% ±0.78
D_A:	CL: 16.0% ±0.84	CR: 15.5% ±0.64	F: 14.0% ±0.91
	N2: 13.6% ±0.84	N1: 12.3% ±0.76	S: 22.4% ±0.96
D_A_P:	CL: 14.2% ±0.53	CR: 15.3% ±1.35	F: 12.7% ±0.95
	N2: 12.7% ±0.98	N1: 13.2% ±1.42	S: 20.9% ±0.92
DTW:	CL: 3.56% ±0.59	CR: 1.89% ±0.28	F: 4.61% ±0.59
	N2: 2.49% ±0.34	N1: 2.78% ±0.37	S: 9.17% ±0.99

Trimmed means_MA_A_N			
M:	CL: 4.34% ±0.64	CR: 2.13% ±0.34	F: 5.02% ±0.77
	N2: 2.02% ±0.29	N1: 4.2% ±0.64	S: 9.6% ±0.81
M_A:	CL: 5.20% ±0.61	CR: 3.61% ±0.61	F: 6.09% ±0.47
	N2: 3.92% ±0.57	N1: 4.75% ±0.63	S: 10.4% ±0.85
M_A_P:	CL: 6.77% ±0.76	CR: 5.02% ±0.97	F: 7.65% ±0.48
	N2: 5.82% ±0.82	N1: 6.4% ±0.81	S: 11.4% ±1.05
E:	CL: 4.8% ±0.39	CR: 2.67% ±0.26	F: 5.39% ±0.53
	N2: 2.94% ±0.37	N1: 4.20% ±0.92	S: 10.1% ±0.68
E_A:	CL: 5.69% ±0.66	CR: 4.37% ±0.52	F: 6.69% ±0.75
	N2: 4.97% ±0.52	N1: 4.82% ±0.40	S: 10.6% ±0.42
E_A_P:	CL: 6.22% ±0.89	CR: 4.21% ±0.38	F: 6.91% ±0.75
	N2: 5.14% ±0.63	N1: 5.3% ±0.33	S: 11.3% ±1.19
D:	CL: 14.1% ±0.79	CR: 14.5% ±0.32	F: 11.7% ±0.72
	N2: 12.9% ±0.64	N1: 11.9% ±0.83	S: 21.6% ±0.88
D_A:	CL: 13.9% ±0.30	CR: 15.6% ±0.62	F: 13.4% ±0.65
	N2: 13.6% ±0.64	N1: 12.0% ±0.81	S: 22.0% ±0.89
D_A_P:	CL: 14.9% ±0.75	CR: 13.3% ±1.08	F: 13.1% ±0.78
	N2: 12.3% ±0.76	N1: 12.9% ±1.04	S: 20.7% ±1.08
DTW:	CL: 3.51% ±0.43	CR: 2.33% ±0.22	F: 4.82% ±0.77
	N2: 2.38% ±0.51	N1: 3.35% ±0.58	S: 8.95% ±0.55

DTW_MA_λ_λ			
DTW:	CL: 5.66% ±0.66	CR: 5.89% ±0.84	F: 5.27% ±0.76
	N2: 5.26% ±0.84	N1: 4.21% ±0.65	S: 14.2% ±0.86

DTW_MA_λ_N			
DTW:	CL: 7.77% ±0.58	CR: 8.81% ±0.35	F: 6.13% ±0.35
	N2: 6.58% ±0.57	N1: 6.14% ±0.80	S: 15.6% ±0.78

DTW100_MA_λ_λ			
M:	CL: 10.4% ±0.49	CR: 11.2% ±1.09	F: 10.8% ±1.00
	N2: 8.98% ±0.62	N1: 9.54% ±1.01	S: 14.3% ±0.73
M_A:	CL: 8.21% ±0.68	CR: 9.03% ±0.54	F: 9.67% ±0.90
	N2: 7.22% ±0.40	N1: 6.54% ±0.58	S: 12.9% ±0.61
M_A_P:	CL: 9.72% ±1.14	CR: 8.94% ±0.95	F: 9.59% ±0.45
	N2: 8.29% ±0.82	N1: 7.34% ±0.69	S: 12.8% ±0.43
E:	CL: 12.1% ±0.56	CR: 11.6% ±0.12	F: 11.2% ±0.89
	N2: 10.0% ±0.61	N1: 10.0% ±0.97	S: 16.0% ±0.69
E_A:	CL: 8.98% ±0.79	CR: 9.6% ±0.64	F: 9.95% ±0.72
	N2: 7.66% ±0.60	N1: 7.02% ±0.80	S: 12.9% ±0.84
E_A_P:	CL: 9.47% ±0.47	CR: 8.15% ±1.19	F: 9.14% ±0.59
	N2: 7.42% ±0.56	N1: 6.60% ±0.82	S: 12.5% ±0.64
D:	CL: 15.0% ±0.99	CR: 13.4% ±0.92	F: 12.9% ±0.71
	N2: 11.4% ±0.61	N1: 13.4% ±0.53	S: 20.9% ±0.75
D_A:	CL: 12.0% ±0.56	CR: 11.6% ±1.02	F: 11.1% ±0.81
	N2: 9.30% ±0.56	N1: 9.51% ±0.64	S: 15.9% ±1.02
D_A_P:	CL: 11.5% ±0.64	CR: 11.4% ±0.92	F: 10.5% ±0.70
	N2: 10.4% ±0.91	N1: 8.99% ±1.07	S: 16.4% ±0.92
DTW:	CL: 4.35% ±0.22	CR: 3.93% ±0.29	F: 4.25% ±0.61
	N2: 3.45% ±0.38	N1: 2.97% ±0.44	S: 10.4% ±0.77

Means_WMA_λ_λ			
M:	CL: 4.31% ±0.61	CR: 3.73% ±0.21	F: 6.06% ±0.79
	N2: 3.94% ±0.66	N1: 3.51% ±0.45	S: 8.96% ±0.45
M_A:	CL: 5.45% ±1.00	CR: 3.68% ±0.40	F: 6.94% ±0.59
	N2: 4.99% ±0.47	N1: 3.65% ±0.57	S: 9.04% ±0.71
M_A_P:	CL: 5.0% ±0.58	CR: 2.63% ±0.40	F: 7.05% ±0.75
	N2: 3.57% ±0.38	N1: 2.61% ±0.44	S: 7.63% ±0.94
E:	CL: 5.62% ±0.86	CR: 3.98% ±0.62	F: 6.59% ±1.08
	N2: 3.93% ±0.53	N1: 4.39% ±0.66	S: 8.92% ±0.62
E_A:	CL: 5.58% ±0.51	CR: 4.15% ±0.81	F: 8.22% ±1.09
	N2: 4.63% ±0.52	N1: 4.38% ±0.48	S: 9.37% ±0.65
E_A_P:	CL: 5.06% ±0.46	CR: 2.9% ±0.42	F: 7.36% ±0.23
	N2: 3.85% ±0.57	N1: 3.35% ±0.35	S: 7.88% ±0.46
D:	CL: 7.10% ±0.54	CR: 6.09% ±0.54	F: 7.98% ±0.87
	N2: 4.69% ±0.57	N1: 5.18% ±0.74	S: 13.1% ±0.60
D_A:	CL: 7.10% ±0.59	CR: 6.58% ±0.58	F: 9.03% ±1.21
	N2: 5.33% ±0.59	N1: 5.46% ±0.60	S: 12.4% ±1.22
D_A_P:	CL: 6.18% ±0.69	CR: 5.11% ±0.57	F: 6.94% ±0.84
	N2: 4.71% ±0.47	N1: 3.71% ±0.42	S: 8.82% ±0.62
DTW:	CL: 1.63% ±0.42	CR: 0.90% ±0.27	F: 3.5% ±0.55
	N2: 1.26% ±0.32	N1: 1.13% ±0.18	S: 5.57% ±0.60

Means_WMA_A_λ			
M:	CL: 4.36% ±0.28	CR: 3.27% ±0.36	F: 5.53% ±0.67
	N2: 3.71% ±0.41	N1: 3.38% ±0.34	S: 8.73% ±0.76
M_A:	CL: 5.13% ±0.81	CR: 3.73% ±0.36	F: 5.97% ±0.66
	N2: 3.97% ±0.65	N1: 3.71% ±0.43	S: 8.02% ±0.41
M_A_P:	CL: 4.99% ±0.64	CR: 2.72% ±0.46	F: 6.23% ±0.81
	N2: 3.42% ±0.26	N1: 2.96% ±0.33	S: 7.24% ±1.03
E:	CL: 5.52% ±0.70	CR: 3.38% ±0.42	F: 6.06% ±0.66
	N2: 3.8% ±0.46	N1: 4.17% ±0.61	S: 9.0% ±0.85
E_A:	CL: 5.95% ±0.39	CR: 3.95% ±0.44	F: 6.85% ±0.74
	N2: 4.27% ±0.71	N1: 3.80% ±0.41	S: 8.30% ±0.65
E_A_P:	CL: 5.3% ±0.43	CR: 2.92% ±0.63	F: 6.40% ±1.21
	N2: 4.18% ±0.46	N1: 3.32% ±0.38	S: 7.11% ±0.80
D:	CL: 6.60% ±0.63	CR: 6.75% ±0.80	F: 7.25% ±0.49
	N2: 5.06% ±0.67	N1: 5.52% ±0.64	S: 12.7% ±1.20
D_A:	CL: 7.55% ±0.44	CR: 6.4% ±0.20	F: 7.44% ±1.02
	N2: 5.66% ±0.64	N1: 5.28% ±0.57	S: 11.5% ±0.97
D_A_P:	CL: 5.95% ±0.52	CR: 5.29% ±0.69	F: 6.77% ±0.61
	N2: 4.88% ±0.59	N1: 3.91% ±0.51	S: 9.01% ±0.68
DTW:	CL: 1.30% ±0.35	CR: 0.9% ±0.21	F: 2.94% ±0.37
	N2: 1.04% ±0.13	N1: 0.65% ±0.24	S: 4.8% ±0.86

Means_WMA_λ_N			
M:	CL: 4.01% ±0.43	CR: 2.27% ±0.32	F: 4.45% ±0.65
	N2: 2.24% ±0.42	N1: 3.47% ±0.56	S: 7.65% ±0.92
M_A:	CL: 6.1% ±0.66	CR: 3.86% ±0.39	F: 5.6% ±0.76
	N2: 2.95% ±0.51	N1: 3.37% ±0.46	S: 8.87% ±0.96
M_A_P:	CL: 4.54% ±0.62	CR: 2.54% ±0.26	F: 5.29% ±1.07
	N2: 3.17% ±0.72	N1: 2.84% ±0.59	S: 8.84% ±0.96
E:	CL: 4.04% ±0.62	CR: 1.98% ±0.36	F: 4.4% ±0.76
	N2: 2.19% ±0.32	N1: 3.16% ±0.31	S: 8.34% ±0.58
E_A:	CL: 5.46% ±0.60	CR: 3.25% ±0.28	F: 6.13% ±0.87
	N2: 3.21% ±0.60	N1: 2.61% ±0.46	S: 8.70% ±0.75
E_A_P:	CL: 4.54% ±0.80	CR: 2.76% ±0.91	F: 5.29% ±0.43
	N2: 2.65% ±0.50	N1: 2.92% ±0.47	S: 8.39% ±1.16
D:	CL: 3.81% ±0.53	CR: 3.52% ±0.16	F: 4.20% ±0.57
	N2: 2.67% ±0.50	N1: 3.86% ±0.41	S: 10.6% ±1.24
D_A:	CL: 3.74% ±0.55	CR: 3.86% ±0.40	F: 5.24% ±0.49
	N2: 2.55% ±0.71	N1: 2.08% ±0.60	S: 10.3% ±0.98
D_A_P:	CL: 4.84% ±0.70	CR: 4.76% ±1.12	F: 5.83% ±0.88
	N2: 3.12% ±0.43	N1: 2.67% ±0.28	S: 9.43% ±0.92
DTW:	CL: 4.39% ±0.66	CR: 2.36% ±0.22	F: 4.42% ±0.86
	N2: 2.08% ±0.36	N1: 2.75% ±0.22	S: 7.05% ±0.69

Means_WMA_A_N			
M:	CL: 4.02% ±0.38	CR: 2.05% ±0.23	F: 5.06% ±0.95
	N2: 2.26% ±0.43	N1: 3.44% ±0.42	S: 8.23% ±0.65
M_A:	CL: 6.03% ±0.39	CR: 4.02% ±0.59	F: 5.61% ±0.68
	N2: 3.28% ±0.61	N1: 3.33% ±0.58	S: 9.68% ±0.90
M_A_P:	CL: 4.29% ±0.44	CR: 2.98% ±0.39	F: 6.0% ±0.89
	N2: 2.80% ±0.48	N1: 2.87% ±0.59	S: 8.46% ±1.02
E:	CL: 4.24% ±0.63	CR: 1.69% ±0.35	F: 4.54% ±0.67
	N2: 2.08% ±0.20	N1: 3.32% ±0.39	S: 7.92% ±0.60
E_A:	CL: 5.41% ±0.83	CR: 3.25% ±0.40	F: 5.52% ±0.95
	N2: 3.18% ±0.29	N1: 2.76% ±0.53	S: 8.98% ±0.51
E_A_P:	CL: 5.26% ±0.62	CR: 2.36% ±0.38	F: 5.59% ±0.72
	N2: 2.73% ±0.38	N1: 2.74% ±0.51	S: 7.81% ±0.78
D:	CL: 4.13% ±0.47	CR: 3.57% ±0.52	F: 4.94% ±0.58
	N2: 2.33% ±0.40	N1: 3.6% ±0.59	S: 10.3% ±0.87
D_A:	CL: 3.72% ±0.75	CR: 3.9% ±0.31	F: 4.36% ±0.49
	N2: 2.62% ±0.54	N1: 2.15% ±0.32	S: 10.5% ±0.59
D_A_P:	CL: 4.84% ±1.14	CR: 4.7% ±0.40	F: 4.86% ±0.95
	N2: 3.3% ±0.73	N1: 2.58% ±0.41	S: 9.64% ±0.95
DTW:	CL: 4.54% ±0.48	CR: 2.30% ±0.30	F: 4.22% ±0.69
	N2: 2.33% ±0.24	N1: 2.78% ±0.19	S: 7.08% ±0.88

Medians_WMA_λ_λ			
M:	CL: 5.74% ±0.86	CR: 3.59% ±0.45	F: 6.61% ±0.73
	N2: 4.27% ±0.48	N1: 3.70% ±0.55	S: 9.69% ±0.83
M_A:	CL: 5.53% ±0.91	CR: 3.79% ±0.46	F: 8.02% ±0.55
	N2: 5.47% ±0.58	N1: 3.67% ±0.42	S: 9.01% ±0.88
M_A_P:	CL: 5.8% ±0.62	CR: 4.63% ±0.62	F: 8.41% ±1.17
	N2: 5.06% ±0.68	N1: 3.68% ±0.65	S: 9.81% ±0.69
E:	CL: 6.29% ±0.54	CR: 4.18% ±0.63	F: 7.13% ±0.76
	N2: 4.51% ±0.67	N1: 4.52% ±0.35	S: 10.2% ±0.66
E_A:	CL: 6.83% ±0.53	CR: 5.65% ±0.79	F: 9.20% ±0.90
	N2: 6.36% ±0.62	N1: 4.85% ±0.54	S: 10.3% ±0.55
E_A_P:	CL: 6.14% ±0.52	CR: 5.51% ±0.72	F: 7.77% ±0.67
	N2: 5.94% ±0.67	N1: 4.19% ±0.46	S: 9.62% ±0.99
D:	CL: 9.49% ±0.78	CR: 9.64% ±0.31	F: 10.4% ±1.15
	N2: 7.71% ±0.76	N1: 7.72% ±0.84	S: 16.5% ±1.26
D_A:	CL: 11.3% ±0.75	CR: 12.0% ±0.56	F: 12.6% ±1.19
	N2: 10.0% ±0.89	N1: 7.72% ±0.71	S: 16.3% ±0.51
D_A_P:	CL: 10.0% ±0.65	CR: 8.84% ±0.60	F: 10.3% ±1.04
	N2: 8.38% ±0.61	N1: 6.95% ±0.69	S: 14.5% ±0.92
DTW:	CL: 1.97% ±0.41	CR: 0.88% ±0.23	F: 3.72% ±0.61
	N2: 1.6% ±0.41	N1: 0.90% ±0.20	S: 5.62% ±0.47

Medians_WMA_A_λ			
M:	CL: 4.71% ±0.42	CR: 3.6% ±0.36	F: 5.62% ±0.90
	N2: 3.7% ±0.48	N1: 4.36% ±0.50	S: 8.58% ±0.90
M_A:	CL: 5.71% ±0.66	CR: 4.03% ±0.44	F: 6.72% ±0.75
	N2: 4.6% ±0.29	N1: 4.23% ±0.69	S: 8.24% ±0.90
M_A_P:	CL: 5.49% ±0.68	CR: 5.0% ±0.63	F: 6.56% ±0.67
	N2: 5.64% ±0.34	N1: 4.02% ±0.42	S: 9.67% ±1.01
E:	CL: 6.65% ±0.67	CR: 4.02% ±0.48	F: 6.56% ±0.92
	N2: 4.29% ±0.35	N1: 5.16% ±0.61	S: 9.71% ±0.82
E_A:	CL: 6.3% ±0.57	CR: 4.86% ±0.67	F: 7.60% ±0.56
	N2: 5.81% ±0.39	N1: 5.47% ±0.65	S: 9.59% ±0.75
E_A_P:	CL: 6.27% ±1.07	CR: 5.51% ±0.20	F: 7.02% ±0.84
	N2: 5.3% ±0.65	N1: 4.47% ±0.57	S: 9.66% ±0.72
D:	CL: 10.4% ±0.63	CR: 10.5% ±0.23	F: 9.61% ±0.93
	N2: 7.56% ±0.92	N1: 7.9% ±0.66	S: 15.6% ±0.81
D_A:	CL: 12.2% ±0.94	CR: 11.5% ±0.83	F: 11.0% ±0.60
	N2: 10.0% ±0.61	N1: 7.95% ±0.51	S: 15.6% ±0.74
D_A_P:	CL: 9.94% ±0.73	CR: 9.53% ±0.55	F: 9.36% ±0.99
	N2: 8.69% ±0.66	N1: 7.56% ±0.67	S: 14.5% ±1.13
DTW:	CL: 1.38% ±0.34	CR: 0.9% ±0.30	F: 3.25% ±1.24
	N2: 0.98% ±0.26	N1: 0.94% ±0.29	S: 5.06% ±0.32

Medians_WMA_λ_N			
M:	CL: 4.31% ±0.52	CR: 2.42% ±0.33	F: 5.21% ±0.76
	N2: 2.27% ±0.41	N1: 3.42% ±0.48	S: 8.68% ±1.17
M_A:	CL: 5.27% ±0.83	CR: 3.63% ±0.70	F: 5.85% ±0.76
	N2: 3.22% ±0.45	N1: 3.71% ±0.62	S: 9.66% ±0.83
M_A_P:	CL: 5.49% ±0.56	CR: 4.29% ±0.80	F: 6.22% ±0.94
	N2: 4.22% ±0.67	N1: 4.1% ±0.75	S: 11.2% ±0.82
E:	CL: 4.68% ±0.43	CR: 2.54% ±0.46	F: 4.83% ±0.52
	N2: 2.65% ±0.55	N1: 3.81% ±0.63	S: 9.54% ±0.65
E_A:	CL: 5.43% ±0.49	CR: 3.19% ±0.55	F: 5.56% ±0.36
	N2: 4.33% ±0.68	N1: 3.58% ±0.53	S: 10.6% ±0.73
E_A_P:	CL: 5.36% ±0.69	CR: 3.94% ±1.02	F: 6.11% ±0.74
	N2: 3.97% ±0.51	N1: 3.39% ±0.37	S: 10.3% ±0.55
D:	CL: 6.67% ±0.50	CR: 6.40% ±0.88	F: 6.46% ±0.73
	N2: 4.69% ±0.54	N1: 5.75% ±0.68	S: 14.7% ±0.96
D_A:	CL: 8.79% ±0.58	CR: 8.2% ±0.95	F: 7.27% ±0.78
	N2: 7.04% ±0.48	N1: 5.53% ±0.40	S: 16.1% ±0.75
D_A_P:	CL: 9.12% ±1.20	CR: 7.42% ±0.76	F: 8.62% ±0.89
	N2: 6.79% ±0.51	N1: 6.59% ±0.80	S: 15.4% ±0.51
DTW:	CL: 3.92% ±0.68	CR: 2.32% ±0.61	F: 4.17% ±0.55
	N2: 1.95% ±0.33	N1: 2.86% ±0.30	S: 8.28% ±0.79

C.2 Automatic detected cycles:

Means_WMA_λ_λ				
M:	CL: 5.05% ±0.55	CR: 6.64% ±0.89	F: 7.2% ±0.49	N: 8.79% ±0.30
	N2: 5.21% ±0.41	N1: 3.62% ±0.61	S: 10.6% ±0.98	
M_A:	CL: 18.6% ±2.22	CR: 20.9% ±1.64	F: 14.4% ±1.33	N: 17.4% ±1.26
	N2: 13.2% ±1.14	N1: 10.0% ±1.34	S: 26.6% ±0.85	
M_A_P:	CL: 21.2% ±1.51	CR: 23.6% ±1.56	F: 18.4% ±2.50	N: 19.1% ±0.71
	N2: 16.3% ±1.28	N1: 12.4% ±1.29	S: 28.8% ±0.96	
E:	CL: 6.08% ±0.83	CR: 6.78% ±0.25	F: 7.65% ±0.91	N: 9.41% ±0.74
	N2: 4.97% ±0.64	N1: 3.84% ±0.21	S: 11.8% ±0.91	
E_A:	CL: 19.6% ±1.46	CR: 20.6% ±1.51	F: 14.9% ±0.46	N: 18.3% ±1.04
	N2: 14.7% ±1.00	N1: 12.3% ±1.58	S: 27.4% ±1.57	
E_A_P:	CL: 21.5% ±1.52	CR: 23.2% ±1.57	F: 16.6% ±1.09	N: 18.5% ±1.35
	N2: 16.0% ±1.88	N1: 12.2% ±1.02	S: 29.5% ±1.86	
D:	CL: 7.56% ±0.56	CR: 7.78% ±0.54	F: 8.13% ±0.72	N: 9.74% ±0.48
	N2: 5.5% ±0.37	N1: 4.03% ±0.33	S: 14.3% ±1.02	
D_A:	CL: 18.2% ±2.43	CR: 19.2% ±1.51	F: 15.1% ±1.34	N: 17.2% ±0.61
	N2: 13.9% ±0.65	N1: 11.8% ±1.05	S: 28.7% ±0.85	
D_A_P:	CL: 22.3% ±2.49	CR: 21.7% ±1.75	F: 17.0% ±1.38	N: 20.0% ±1.52
	N2: 16.7% ±0.74	N1: 13.5% ±1.13	S: 30.0% ±1.34	
DTW:	CL: 2.79% ±0.45	CR: 5.69% ±1.08	F: 3.32% ±0.50	N: 6.25% ±0.45
	N2: 2.84% ±0.42	N1: 1.77% ±0.30	S: 11.3% ±1.29	

Means_WMA_A_λ				
M:	CL: 5.15% ±0.41	CR: 6.62% ±0.63	F: 5.22% ±0.52	N: 7.87% ±0.73
	N2: 4.46% ±0.33	N1: 3.47% ±0.72	S: 10.4% ±1.24	
M_A:	CL: 19.5% ±1.65	CR: 20.4% ±1.31	F: 13.6% ±1.40	N: 16.1% ±1.22
	N2: 12.9% ±0.69	N1: 9.57% ±1.19	S: 8.40% ±1.77	
M_A_P:	CL: 21.8% ±1.93	CR: 24.2% ±1.29	F: 16.0% ±1.67	N: 19.0% ±1.00
	N2: 16.0% ±1.38	N1: 11.7% ±1.18	S: 13.0% ±0.0	
E:	CL: 5.12% ±0.67	CR: 7.0% ±0.60	F: 5.38% ±0.67	N: 8.23% ±0.57
	N2: 4.23% ±0.50	N1: 4.14% ±0.34	S: 10.0% ±0.66	
E_A:	CL: 20.6% ±1.67	CR: 21.5% ±1.10	F: 13.4% ±0.91	N: 16.5% ±1.25
	N2: 14.4% ±1.24	N1: 11.5% ±1.36	S: 11.2% ±1.77	
E_A_P:	CL: 22.9% ±1.53	CR: 22.9% ±0.94	F: 15.3% ±1.51	N: 17.9% ±1.38
	N2: 16.8% ±1.97	N1: 11.9% ±1.45	S: 10.0% ±0.0	
D:	CL: 7.63% ±0.61	CR: 7.69% ±0.23	F: 6.60% ±0.56	N: 9.55% ±0.79
	N2: 5.08% ±0.37	N1: 4.69% ±0.50	S: 13.3% ±0.86	
D_A:	CL: 19.9% ±1.41	CR: 18.8% ±0.96	F: 13.3% ±0.78	N: 17.5% ±1.30
	N2: 13.7% ±1.33	N1: 10.6% ±0.86	S: 10.5% ±1.77	
D_A_P:	CL: 22.8% ±1.84	CR: 23.6% ±1.14	F: 15.4% ±1.28	N: 19.4% ±1.52
	N2: 16.5% ±1.63	N1: 12.2% ±1.07	S: 11.2% ±1.77	
DTW:	CL: 2.92% ±0.66	CR: 7.28% ±0.63	F: 3.21% ±0.57	N: 5.99% ±0.39
	N2: 2.26% ±0.71	N1: 1.75% ±0.43	S: 10.7% ±1.12	

Means_WMA_A_N				
M:	CL: 6.79% ±1.06	CR: 9.99% ±1.03	F: 6.02% ±0.76	N: 8.98% ±0.87
	N2: 4.82% ±0.93	N1: 4.15% ±0.50	S: 16.0% ±0.95	
M_A:	CL: 19.2% ±1.93	CR: 20.8% ±1.25	F: 14.5% ±2.07	
	N2: 13.3% ±1.39	N1: 11.1% ±1.31	S: 30.0% ±1.83	
M_A_P:	CL: 23.0% ±1.58	CR: 24.1% ±1.60	F: 16.5% ±1.25	
	N2: 16.5% ±1.69	N1: 13.2% ±1.69	S: 32.1% ±1.09	
E:	CL: 5.52% ±1.10	CR: 8.46% ±0.89	F: 5.65% ±0.75	N: 8.80% ±0.79
	N2: 4.41% ±0.76	N1: 3.80% ±0.49	S: 14.2% ±1.20	
E_A:	CL: 20.5% ±1.44	CR: 21.6% ±1.39	F: 15.0% ±1.75	
	N2: 14.7% ±1.78	N1: 12.2% ±1.16	S: 31.9% ±1.98	
E_A_P:	CL: 21.7% ±1.91	CR: 23.5% ±1.32	F: 15.0% ±1.73	
	N2: 16.6% ±1.52	N1: 11.5% ±1.13	S: 34.2% ±1.93	
D:	CL: 3.21% ±0.38	CR: 6.56% ±0.99	F: 3.54% ±0.48	N: 7.59% ±0.75
	N2: 2.35% ±0.44	N1: 2.21% ±0.19	S: 13.2% ±1.19	
D_A:	CL: 20.9% ±2.42	CR: 21.7% ±1.40	F: 12.8% ±1.62	
	N2: 14.9% ±1.45	N1: 10.8% ±1.12	S: 32.7% ±2.01	
D_A_P:	CL: 23.1% ±1.47	CR: 24.9% ±1.46	F: 16.7% ±1.74	
	N2: 16.5% ±1.26	N1: 12.6% ±0.92	S: 32.9% ±1.34	
DTW:	CL: 6.25% ±1.10	CR: 9.52% ±0.72	F: 5.89% ±0.82	N: 8.66% ±0.51
	N2: 4.65% ±0.84	N1: 3.46% ±0.29	S: 15.5% ±1.51	

Medians_WMA_A_λ				
M:	CL: 5.37% ±0.53	CR: 6.32% ±0.67	F: 5.68% ±0.76	N: 8.07% ±0.62
	N2: 4.56% ±0.45	N1: 3.8% ±0.34	S: 10.3% ±0.75	
M_A:	CL: 21.0% ±2.32	CR: 21.9% ±1.29	F: 14.5% ±1.06	N: 17.9% ±1.34
	N2: 15.5% ±2.02	N1: 11.0% ±1.04	S: 28.4% ±2.31	
M_A_P:	CL: 24.0% ±1.25	CR: 24.5% ±1.30	F: 18.0% ±0.80	N: 21.9% ±0.84
	N2: 19.7% ±1.58	N1: 14.4% ±1.25	S: 30.9% ±1.28	
E:	CL: 5.87% ±0.50	CR: 6.53% ±0.47	F: 6.63% ±0.63	N: 8.42% ±0.47
	N2: 5.02% ±0.95	N1: 4.49% ±0.65	S: 9.92% ±0.56	
E_A:	CL: 22.3% ±0.70	CR: 24.4% ±0.89	F: 16.5% ±1.42	N: 18.5% ±1.20
	N2: 16.6% ±0.73	N1: 13.3% ±1.03	S: 29.9% ±1.61	
E_A_P:	CL: 23.3% ±1.07	CR: 25.6% ±0.88	F: 17.7% ±1.61	N: 20.8% ±1.22
	N2: 17.4% ±1.23	N1: 14.0% ±1.17	S: 29.9% ±0.87	
D:	CL: 13.1% ±0.84	CR: 13.8% ±0.48	F: 10.2% ±0.47	N: 10.9% ±0.61
	N2: 7.76% ±0.42	N1: 7.36% ±0.35	S: 16.0% ±0.58	
D_A:	CL: 21.7% ±0.99	CR: 24.0% ±1.32	F: 16.6% ±1.11	N: 19.7% ±1.06
	N2: 16.7% ±1.32	N1: 14.6% ±1.09	S: 29.5% ±1.36	
D_A_P:	CL: 25.9% ±1.23	CR: 24.6% ±1.06	F: 19.3% ±1.04	N: 23.8% ±1.03
	N2: 20.0% ±1.39	N1: 16.4% ±0.70	S: 31.7% ±0.87	
DTW:	CL: 3.25% ±0.76	CR: 6.01% ±0.77	F: 3.19% ±0.42	N: 5.86% ±0.49
	N2: 1.89% ±0.23	N1: 1.6% ±0.26	S: 10.6% ±0.83	

Medians_WMA_A_N				
M:	CL: 7.3% ±1.01	CR: 6.69% ±0.96	F: 5.65% ±0.59	N: 8.53% ±0.91
	N2: 4.11% ±0.85	N1: 3.69% ±0.61	S: 14.4% ±1.36	
M_A:	CL: 20.9% ±1.29	CR: 22.1% ±0.95	F: 14.7% ±0.96	
	N2: 15.4% ±1.70	N1: 12.4% ±1.39	S: 31.6% ±1.94	
M_A_P:	CL: 25.2% ±0.63	CR: 24.7% ±1.01	F: 18.2% ±2.06	
	N2: 19.7% ±2.11	N1: 15.0% ±1.66	S: 33.3% ±1.75	
E:	CL: 5.62% ±0.49	CR: 7.33% ±1.22	F: 5.52% ±0.55	N: 8.91% ±0.99
	N2: 4.26% ±0.77	N1: 3.84% ±0.62	S: 13.1% ±1.54	
E_A:	CL: 21.2% ±1.71	CR: 23.2% ±1.38	F: 14.9% ±1.62	
	N2: 16.8% ±1.41	N1: 12.7% ±1.79	S: 32.8% ±1.32	
E_A_P:	CL: 24.0% ±2.07	CR: 25.0% ±1.24	F: 17.6% ±1.31	
	N2: 18.5% ±1.04	N1: 14.0% ±1.61	S: 33.8% ±0.79	
D:	CL: 7.82% ±0.65	CR: 10.1% ±0.74	F: 5.92% ±0.92	N: 9.63% ±0.49
	N2: 5.42% ±0.57	N1: 4.83% ±0.56	S: 17.0% ±1.58	
D_A:	CL: 22.7% ±1.29	CR: 24.9% ±1.56	F: 15.4% ±1.66	
	N2: 17.8% ±1.50	N1: 13.3% ±1.41	S: 34.4% ±2.21	
D_A_P:	CL: 25.6% ±1.06	CR: 26.0% ±1.36	F: 18.3% ±1.76	
	N2: 20.9% ±1.75	N1: 15.9% ±1.69	S: 34.9% ±1.44	
DTW:	CL: 7.47% ±1.00	CR: 7.13% ±0.54	F: 5.43% ±0.42	N: 8.93% ±0.86
	N2: 4.65% ±0.51	N1: 3.9% ±0.88	S: 14.1% ±1.46	

Trimmed means_WMA_A_λ				
M:	CL: 7.35% ±0.53	CR: 7.52% ±0.48	F: 6.8% ±0.73	N: 8.92% ±0.37
	N2: 5.26% ±0.23	N1: 4.81% ±0.30	S: 11.3% ±0.85	
M_A:	CL: 23.5% ±1.77	CR: 25.6% ±1.42	F: 17.3% ±0.79	N: 19.4% ±1.15
	N2: 18.4% ±1.33	N1: 14.2% ±1.53	S: 28.6% ±1.56	
M_A_P:	CL: 26.6% ±1.47	CR: 28.2% ±1.99	F: 21.1% ±1.62	N: 23.9% ±1.36
	N2: 21.8% ±1.94	N1: 18.1% ±1.70	S: 31.4% ±1.28	
E:	CL: 7.82% ±0.42	CR: 8.67% ±0.62	F: 7.54% ±0.81	N: 9.35% ±0.52
	N2: 6.31% ±0.65	N1: 5.61% ±0.45	S: 11.0% ±0.83	
E_A:	CL: 24.7% ±1.69	CR: 26.8% ±1.49	F: 18.6% ±0.65	N: 21.0% ±1.44
	N2: 20.1% ±1.42	N1: 15.2% ±1.08	S: 30.7% ±1.68	
E_A_P:	CL: 26.1% ±1.67	CR: 27.8% ±0.70	F: 20.8% ±1.98	N: 23.0% ±1.43
	N2: 22.1% ±1.24	N1: 17.5% ±1.30	S: 31.2% ±1.31	
D:	CL: 24.5% ±0.75	CR: 23.9% ±0.83	F: 18.3% ±0.77	N: 19.5% ±1.01
	N2: 17.2% ±0.58	N1: 15.4% ±0.54	S: 25.7% ±0.79	
D_A:	CL: 27.0% ±1.52	CR: 29.7% ±1.41	F: 22.1% ±0.65	N: 24.3% ±1.09
	N2: 22.5% ±1.22	N1: 19.8% ±1.26	S: 33.5% ±0.88	
D_A_P:	CL: 29.7% ±1.33	CR: 30.7% ±1.16	F: 24.7% ±1.37	N: 27.9% ±1.25
	N2: 25.7% ±1.65	N1: 21.1% ±0.96	S: 34.3% ±1.00	
DTW:	CL: 3.33% ±0.49	CR: 5.93% ±0.98	F: 3.71% ±0.62	N: 7.33% ±0.57
	N2: 2.61% ±0.33	N1: 2.5% ±0.51	S: 11.6% ±0.64	

Trimmed means_WMA_A_N				
M:	CL: 7.77% ±0.87	CR: 8.09% ±0.79	F: 5.92% ±0.96	N: 10.0% ±0.95
	N2: 5.22% ±0.52	N1: 4.18% ±0.54	S: 16.0% ±1.43	
M_A:	CL: 22.9% ±1.10	CR: 24.8% ±2.02	F: 16.9% ±1.06	
	N2: 18.3% ±0.92	N1: 15.6% ±1.18	S: 32.1% ±0.90	
M_A_P:	CL: 27.3% ±0.97	CR: 29.2% ±1.39	F: 20.5% ±1.42	
	N2: 22.5% ±2.19	N1: 18.8% ±1.09	S: 33.6% ±0.72	
E:	CL: 6.05% ±0.84	CR: 7.85% ±0.58	F: 6.44% ±0.94	N: 9.37% ±0.59
	N2: 4.82% ±0.37	N1: 4.08% ±0.54	S: 14.4% ±1.60	
E_A:	CL: 24.2% ±1.39	CR: 26.4% ±1.94	F: 17.9% ±1.17	
	N2: 19.3% ±1.64	N1: 16.0% ±1.37	S: 33.2% ±2.76	
E_A_P:	CL: 26.3% ±1.79	CR: 28.5% ±2.33	F: 20.1% ±1.72	
	N2: 21.6% ±1.63	N1: 17.0% ±1.34	S: 35.9% ±1.36	
D:	CL: 18.9% ±1.10	CR: 22.4% ±0.47	F: 12.6% ±0.92	N: 16.0% ±1.19
	N2: 13.8% ±0.41	N1: 10.4% ±0.78	S: 26.5% ±1.25	
D_A:	CL: 27.3% ±1.23	CR: 29.9% ±1.21	F: 20.9% ±0.89	
	N2: 23.2% ±1.74	N1: 19.2% ±1.69	S: 37.3% ±1.00	
D_A_P:	CL: 31.0% ±0.83	CR: 32.1% ±1.63	F: 24.6% ±1.12	
	N2: 26.7% ±1.58	N1: 22.3% ±1.22	S: 38.4% ±1.64	
DTW:	CL: 7.97% ±0.65	CR: 8.11% ±1.20	F: 5.73% ±0.87	N: 9.60% ±0.76
	N2: 4.77% ±0.87	N1: 4.45% ±0.49	S: 16.7% ±1.80	

DTW_WMA_λ_λ				
DTW:	CL: 11.7% ±0.90	CR: 12.1% ±0.71	F: 7.01% ±0.68	N: 12.8% ±1.30
	N2: 8.24% ±0.91	N1: 7.06% ±0.71	S: 19.2% ±0.85	

DTW100_WMA_λ_λ				
M:	CL: 28.0% ±1.12	CR: 27.6% ±0.94	F: 22.0% ±1.07	N: 26.7% ±0.86
	N2: 24.5% ±1.50	N1: 24.9% ±0.99	S: 31.3% ±1.03	
M_A:	CL: 28.9% ±1.39	CR: 31.1% ±1.13	F: 23.6% ±3.11	N: 28.2% ±1.00
	N2: 23.7% ±0.73	N1: 25.8% ±0.98	S: 34.1% ±1.39	
M_A_P:	CL: 31.7% ±1.63	CR: 32.2% ±1.07	F: 25.5% ±1.09	N: 29.7% ±0.93
	N2: 27.2% ±0.98	N1: 28.8% ±0.93	S: 35.2% ±3.78	
E:	CL: 28.2% ±1.36	CR: 28.0% ±1.48	F: 23.7% ±1.59	N: 27.3% ±0.76
	N2: 25.5% ±0.93	N1: 25.9% ±1.68	S: 31.9% ±1.00	
E_A:	CL: 28.9% ±1.20	CR: 31.6% ±1.44	F: 23.2% ±0.99	N: 28.4% ±1.32
	N2: 25.9% ±1.04	N1: 28.0% ±1.02	S: 31.8% ±2.47	
E_A_P:	CL: 30.2% ±1.62	CR: 31.2% ±1.33	F: 25.2% ±3.09	N: 28.9% ±1.09
	N2: 26.3% ±1.12	N1: 27.7% ±1.07	S: 36.5% ±0.44	
D:	CL: 32.2% ±1.03	CR: 31.4% ±1.09	F: 25.5% ±0.84	N: 29.9% ±1.02
	N2: 27.5% ±0.86	N1: 27.6% ±0.81	S: 36.7% ±0.42	
D_A:	CL: 30.5% ±0.64	CR: 31.1% ±1.03	F: 25.9% ±2.08	N: 29.3% ±1.18
	N2: 27.1% ±1.49	N1: 28.4% ±1.20	S: 36.1% ±1.50	
D_A_P:	CL: 32.1% ±1.81	CR: 32.6% ±0.90	F: 26.7% ±1.24	N: 31.3% ±1.22
	N2: 29.7% ±1.09	N1: 30.8% ±1.61	S: 38.8% ±1.30	
DTW:	CL: 29.2% ±1.00	CR: 32.2% ±0.89	F: 16.8% ±1.13	N: 25.4% ±0.81
	N2: 20.7% ±0.97	N1: 24.3% ±1.31	S: 35.2% ±1.34	