Gait analysis: Is it easy to learn to walk like someone else?

Øyvind Stang (oeyvind.stang@hig.no)

Masteroppgave
Master i informasjonssikkerhet
30 ECTS
Avdeling for informatikk og medieteknikk
Høgskolen i Gjøvik, 2007

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

In this master's thesis, we will look at whether it is easy or difficult to learn to walk like someone else in such a way that one will be accepted by an authentication system, based on gait. In the last couple of years, there have been some studies about whether gait can be used in order to authenticate a person. If it turns out to be very easy to learn to walk like another person, then gait authentication should probably not be used as the only authentication technique.

We investigate the ease of gait mimicking by means of a 3-axis sensor worn by the user. A prototype is created, which reads this acceleration data, and plots it as 3 graphs in a coordinate system, shown on a big screen. The aim is to see whether the user manages to learn to walk in such a way that his graphs match 3 template graphs, which are also plotted in the same coordinate system. Every attempt lasts 5 seconds, and a score between 0 and 100 will be given in the end of each, based on how similar the impersonator's graphs are to the original graphs. We use *Pearson's correlation* to calculate this.

The experiment has 13 participants, and we have created 5 different templates which each participant will attempt to imitate 15 times. The results from our data analysis are showing that it actually seems rather easy to learn to walk like another person, and hence to be accepted by a gait authentication system.

Sammendrag

I denne masteroppgaven skal vi undersøke hvorvidt det er lett eller vanskelig å lære seg å gå som en annen persons slik at en vil bli godkjent av et autentiseringssystem som baserer seg på ganglag. I de senere år har det vært endel studier rundt hvorvidt ganglag kan brukes for å autentisere en person. Hvis det viser seg å være veldig lett å lære seg å gå som en annen persons, da burde antakelig ikke ganglagsautentisering brukes som den eneste autentiseringsteknikken.

Vi undersøker vanskelighetsgraden av å imitere ganglag ved hjelp av en 3-akset sensor som blir båret av brukeren. En prototype som leser disse aksellerasjonsdataene, og som plotter disse som 3 grafer i et koordinasjonssystem vist på en storskjerm er utviklet. Målet er å se hvorvidt brukeren klarer å lære seg å gå på en slik måte at hans grafer vil matche 3 templategrafer, som også er plottet i koordinatsystemet. Hvert forsøk varer i 5 sekunder, og i enden av dem vil en poengsum mellom 0 og 100 poeng bli gitt, basert på hvor like imitatorens grafer og de orginale grafene er. Vi bruker *Pearsons korrelasjon* for å regne ut dette.

Eksperimentet har 13 deltakere, og vi har laget 5 ulike templater som hver deltaker skal prøve å etterlikne 15 ganger. Resultatene fra vår dataanalyse viser at det faktisk virker rimelig lett å lære seg å gå som en annen person, og slik sett bli akseptert av et ganglagsautentiseringssystem.

Preface

What you are about to read is my master's thesis, completed as part of a Master of Science degree in Information Security at Gjøvik University College, where I have been studying since 2001. I completed my Bachelor's degree in Computer Engineering at this school in 2004, and I am soon to complete my Master's degree in Information Security. The reason why I have chosen a study in gait analysis, is because authentication, and especially biometrics, has been one of my favorite topics during this education.

In information security, one of the most important aspects is to ensure that only authorized people are allowed to get access to confidential information. Different authentication techniques have been invented in order to determine how likely it is that a person is who he (or she) claims to be. One of the least known techniques in this field is gait analysis. Since it is not a well-known technique, such as e.g. fingerprint analysis or password verification, it is especially important to look at the dangers that may arise by using it. Since walking is something that you do, then one of the greatest dangers is of course *impersonation*. This is the main motivation for choosing to work with this exact master's thesis. We will investigate the ease of mimicking someone else's gait.

There are some people that I would like to thank for the help they have given during this work. First of all, I will thank my supervisor Einar Snekkenes for the advices he has given. I would also like to thank Frode Gilberg and Torkjel Søndrol for the tips I have got during the program development phase, Davrondzhon Gafurov and Patric Bours for their advices concerning gait authentication, Frode Volden for his help in the data analysis phase, and finally all of the 13 participants who contributed to my experiment.

Øyvind Stang, 28th June 2007

Contents

List of Figures

List of Tables

1 Introduction

1.1 Topic covered by this thesis

Authenticating people by their gait is one of the least used authentication techniques that we know about today. Fingerprints, iris scans, face recognition, and signature verification are biometrics which are more commonly used in order to verify that a person really is who he claims to be.

However, this is a method that has been considered as an alternative to the traditional methods that we mentioned above, and a number of different algorithms and approaches have been developed [1]. According to David Cameron's article ”Walk this way” [2], one of the greatest pioneers within gait research is Mark Nixon, who is a computer scientist at University of Southampton in the U.K. So far, this research has been concentrating mostly on image-based gait authentication, i.e. a study of a person's movements, using images caught by one or more cameras. However, there has also been some research about how to interpret a person's walking style, using a sensor containing an *accelerometer*. In this thesis, we will use a sensor like this, in order to gather information about a person's gait. We will address this method as *accelerometer-based gait analysis*.

The main advantage using gait as a biometric, is the possibility of being authenticated from a distance of the camera or the authentication system. In both image-based authentication systems and accelerometer-based authentication systems, the gait is measured from a certain distance, either by a camera or by a computer receiving data from a sensor.

When e.g. fingerprints are used in order to authenticate a person, it is required that the user is extra active when the biometric features are gathered. In the case of gait analysis, on the other hand, the user is placed in a more natural situation, and gait is as such a ”less interfering” authentication method, seen from the user's point of view.

When it comes to using some of the more traditional authentication techniques, like e.g. fingerprints, many people are feeling uncomfortable knowing that a vast number of other people have pressed their fingers on the scanner plate before them, especially because of the hygienic issues. This is rather relevant e.g. in some Asian countries, where epidemics have been a problem. Iris scans are not very popular among some people either, because they don't like the idea that something is going to scan their eyes. The sceptisism is often based on the irrational fear that the scanning might damage their eye. By using gait as the biometric, the users don't need to be sceptical in the same sense.

Gait is classified as a *behavioural biometric*, i.e. authentication based on something we do. When the user knows that his biometrics are gathered during a performance like this, there is a certain risk that he will perform quite differently than he would have done otherwise. This is also a problem in gait analysis, as we will discuss later in this report. However, it is probably easier to act more naturally when using this kind of authentication than some of the other techniques.

One of the advantages choosing accelerometer-based gait analysis instead of image-based gait analysis, is that analyzing acceleration data is a less complicated task than analyzing details in images. In addition to this, most database images are two-dimensional and depend greatly on the angle of the camera [2]. Accelerometer-based gait analysis compares a person's acceleration data to a template which is stored in a text file. We will get into the details later in this thesis. Other MSc students and PhD candidates at Gjøvik University College have also investigated this area of authentication [3, 4, 5, 6].

Keywords: Authentication, biometrics, gait, accelerometer-based gait analysis, impersonation, interactive learning, correlation, linear regression, DET-curve.

1.2 Problem description

Reading this far, it seems like there are only advantages using gait as a biometric. However, gait authentication should also be looked at with critical eyes. The most severe problem using biometrics in general, is the danger of *impersonation*. It is possible to fake a fingerprint, and to imitate someone's voice or keystroke dynamics. When it comes to gait, there has not been much research on mimicking someone else's gait, and hence, this thesis will investigate whether it should be considered easy or difficult to learn to mimic someone else's gait, and as such be accepted as an authorized user. If it turns out to be quite easy to learn this within a reasonably short number of times, what does that tell us about the security of gait authentication?

1.3 Research questions

In order to solve this problem, there are some research questions which need to be considered:

- What techniques and algorithms exist to be used in gait authentication, e.g. to match two datasets or graphs?
- Does a person walk the same way (more or less) all the time?
- How effective is interactive feedback when learning a specific gait?
- Should it be considered easy or difficult to learn to mimic another gait?

1.4 Claimed contributions

There has not been much research about mimicking of gait before, even if this is an extremely relevant issue, seen from a security perspective. In this thesis, we will investigate how easy it is to learn this within a reasonably short number of attempts. We will create some templates, representing different gaits, and participants will try to mimic these. They will get 15 attempts on each template, and after each attempt, they will get a score, telling them how close their gait was to the template. In the end of the experiment, we hope to see a clear tendency in the increase of the scores.

The results from this experiment will tell us whether impersonation is a realistic problem also within gait authentication. Since this authentication technique is fairly new in comparison to many of the other authentication techniques, it should be interesting also for other researchers working with gait authentication to see how secure this biometric really is.

1.5 Method

The first part of this thesis is to study the existing literature within this topic. This will give us information about the previous research that has been done, which might also be helpful in our work.

The next phase is to create a prototype to be used in our experiment. This prototype is going to communicate with a sensor worn by the user, and it is going to plot the receiving acceleration data in the x-, y-, and z-direction in a coordinate system, as 3 graphs. This coordinate system will be presented to the user on a big screen, together with the 3 corresponding template graphs, as he is walking.

When this work is completed, the third part of the project can begin, which is the experiment. In this experiment, 5 different templates will be created, representing 5 different walking styles. The goal for the participant is to try to mimic the template gaits by studying the graphs as he is walking. His own graphs will be plotted in the same coordinate system as the template graphs, and the user will get a score after each attempt, telling him how close his graphs were to the template graphs. This is done in order to give the user an interactive and motivating feedback from the prototype, such that he will try to score even higher in the next attempt. We will explain in more detail about this experiment later in this thesis.

The fourth part is the statistical analysis of the experiment data. By this time, we will have got some numbers (scores) from different persons, and these should be analyzed. If the tendency among the scores is increasing a lot from the first to the 15th attempt in each template, then we should conclude that there might not be that difficult to learn to walk like someone else. Another way of approaching this problem, is to study a threshold based on how well a person manages to walk like himself, and then see how many of the participants who manage to score above this threshold, and how many times they manage to do this within their limited number of attempts.

1.6 Outline of the report

The next chapter is an introduction about authentication, biometrics and gait authentication. This is aimed mainly at those who are not that familiar with authentication, and especially biometrics. In chapter 3, we will look at previous work related to our problem. In chapter 4, an explanation about our choices of technology will be given, i.e. the choice of sensor, programming tool, and statistical tool. Chapter 5 is a description of our prototype. Chapter 6 gives a more detailed description about the experiment, and it contains a data analysis of the numbers gained in the experiment. In chapter 7, we will take a look at other ways we could have solved the task. In chapter 8, we discuss further work, and in chapter 9, we will give our final conclusions.

2 Introduction to authentication, biometrics, and gait

This chapter is aimed at those who are not that familiar with authentication, and specifically biometrics and gait analysis. We will discuss what authentication is, different ways to authenticate a person, what biometrics is, different biometrics, and finally, different methods to use in order to decide whether a person is who he claims to be, based on his gait.

2.1 Authentication

One way to decrease the risk that sensitive information ends up in the hands of unauthorized people, is to ensure that a person's identity is controlled before he is getting access to this information. It is impossible to guarantee with 100% certainty that only authorized users will be accepted, but the probability of getting access should at least be a lot higher for authorized people, than for unauthorized people. There are several ways to do this. One way is to require a username and a password from the user, and then check whether the user input matches the stored username and password. Another way is to base the outcome of the authentication on the comparison between the user's fingerprint data and a template stored in a database. There are other ways to do this as well.

Even if they are often used among each other, there is a certain difference between the notions *identification* and *authentication*. Identification is about deciding a person's identity. Authentication, on the other hand, is about the problem of confirming or denying a person's *claimed* identity [7]. In this thesis, the focus will be on authentication.

There are three fundamental categories of authentication techniques [1]:

- Something you *know*
- Something you *have*
- Something you *are*

”Something you know” might be a password or a PIN code. You either choose or are given a code, which is personal for you, and this should not be given to anyone else. When attempting to get access to a system, you are then asked to type your password or PIN code. If the code you are typing matches the correct user's password or PIN code, you will be accepted. Otherwise, you will be rejected.

”Something you have” might be a card or a token. The difference between this authentication category and the ”something you know”-category, is that you in this case have to bring a physical object in order to be accepted. A card or a token stores a secret which is much more complicated than an ordinary password, and it is hard to copy. The main advantage with this authentication category, is that you don't need to remember a specific secret. However, the main disadvantage is that it is rather easy to lose such an object. As a result of this, other persons might get hold of the card or token, and in the worst case get access to information they are not authorized to see. Hence, it is better to use more than one authentication technique, e.g. by *both* showing a card, *and* enter a PIN code or a password. It is important though, that these authentication techniques are independent of each other, i.e. that the outcome of one authentication technique is not affecting the outcome of the other authentication techniques. To use more than one authentication technique, is called *multi-factor authentication* [1]. If the probability of being accepted by a system by pure luck is $\frac{1}{n}$ (where $n > 1$) using ”*Authentication Technique A*”, and $\frac{1}{m}$ (where $m > 1$) using ”*Authentication Technique B*”, then it is $\frac{1}{nm}$ when combining both techniques, which is a lower probability. Hence, the risk that an unauthorized person is getting access to a system decreases when a combination of two or more authentication techniques is used. Systems that contain multiple sensors that capture different biometric traits, are called *multimodal biometric systems* [8].

”Something you are” refers to any automatically measurable physical characteristic or personal trait that is distinctive to a person [1]. Authentication techniques in this category, are called *biometric authentication*, and we will take a closer look at this category in chapter 2.2.

In order to authenticate a person, a template has to be stored in a database. This may be a database of passwords or of fingerprint templates. As shown in Figure 1 (from Torkjel Søndrol's master's thesis ”Using the human gait for authentication” [6]), the template is first going through a quality check, which controls that the template is acceptable. The next step is to extract the features from the template, which in the case of fingerprints may be ridges and bifurcations, and other important information. This information are then stored in the database together with a name of the template.

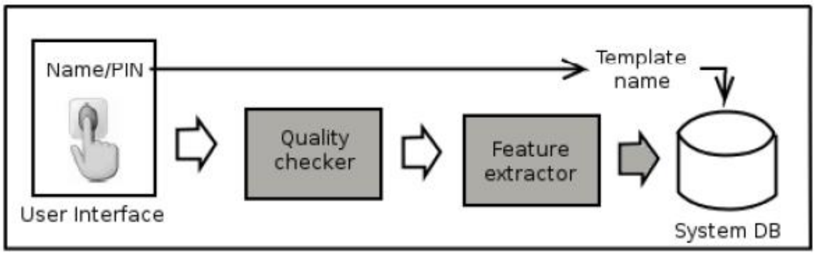

Figure 1: A person's relevant features are enrolled, a quality checker is making sure that the template is acceptable, relevant features are extracted, and the resulting template is stored in a database. This illustration is borrowed from Torkjel Søndrol's master's thesis ”Using the human gait for authentication” [6].

When a person is trying to be authenticated, he gives the system information about who he claims to be, and the system checks whether this identity actually exists in the database. This is shown in Figure 2 (which is also from [6]). If it does exist, the authenticator's relevant features are extracted, and a matching procedure starts. If the authenticator passes this test, he is accepted. Otherwise, he is rejected.

2.2 Biometrics

As explained in chapter 2.1, biometric authentication is a study of a measurable physical characteristic or a personal trait distinctive to an individual, in order to detect whether the individual is who he claims to be or not. Since biometric authentication is a complicated term, it is usually just referred to as *biometrics*. Biometrics are classified into two main categories, i.e. *behavioural* and *physiological* biometrics. A physiological biometric

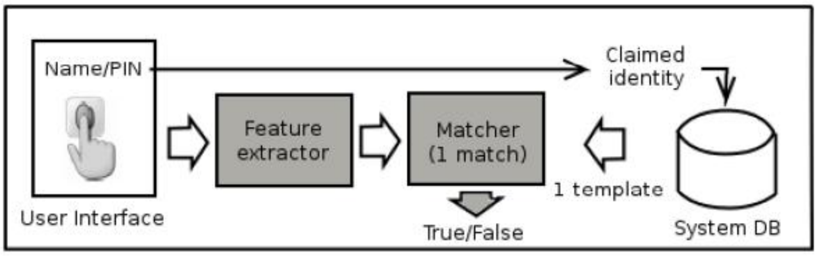

Figure 2: When a person tries to be authenticated, he gives the system information about who he is. The system checks if his identity exists in the database, and if it does, the authenticator's relevant features are extracted and compared to the template. This illustration is also borrowed from Torkjel Søndrol's master's thesis ”Using the human gait for authentication” [6].

is something you *are*, while a behavioural biometric is something you *do* [9]. Examples of biometrics from both of these two categories are shown in Table 1, where parts of it are borrowed from the article ”Visual Speech: A Physiological or Behavioural Biometric?” by J. D. Brand, J. S. D. Mason, and Sylvain Colomb [9].

	Physiological (What you are)	Behavioural (What you do)
Iris	are	
Fingerprints	are	
Hand geometry	are	
DNA	are	
Odor	are	
Signature		do
Keystroke dynamics		do
Face	are	do
Lips	are	do
Voice	are	do
Gait	are	do

Table 1: Classifications of different biometrics. Parts of the table are borrowed from J.D. Brand, J.S.D. Mason, and S. Colomb [9].

According to J.D. Brand, J.S.D. Mason, and S. Colomb [9], gait, face, and lips may be categorized as both physiological *and* behavioural biometrics. These biometrics imply movement or dynamics, but they are also highly dependent on the physiological make-up of the individual. We have classified voice to belong to both of these categories too, since it also implies dynamics, as well as being dependent on physiological features, e.g. whether the voice is dark or light.

When using an authentication technique from the ”something you know”- or the ”something you have”-category, it is very easy to decide whether the user is accepted or not. The input password is compared to the password template, and this comparison is resulting in either ”true” or ”false”. The secret within the smart card is also compared to the template, and this comparison is also resulting in either ”true” or ”false”. However, in biometrics, and especially in behavioural biometrics, it is more difficult to decide with 100% probability whether someone claiming to be *”Person A”* really is *”Person A”* or not. Whether he is accepted or not has a lot to do with the specific situation around the authentication process. When authenticating using gait or voice, it is rather likely that the input will not be identical to the template, even if it in fact is the correct person who is being authenticated. This is because a person may walk differently if his legs have got an injury, and he may have a different voice if he has catched a cold.

In order to solve this, we have to decide about a limit - a *threshold* - describing how much the authenticator's input and the template have to match in order for the user to be accepted. A score is given, based on the comparison between the input and the template, e.g. a score between 0 and 100 points. A threshold in such a scenario may e.g. be set to 60 points. A user whose input matches the template with 60 points will in that case be accepted, while a user whose input matches the template with 59 points will be rejected. The goal should therefore be to set the threshold so high that few unauthorized persons will be accepted, but still so low that few authorized persons will be rejected. Hence, the person deciding the threshold should concentrate about finding a trade-off between security on the one side, and user friendliness on the other side.

When analyzing what threshold to choose, different thresholds are usually tested, and the results are visualized as a *DET-curve*. An example of a DET-curve is shown in Figure 3 (from the web site called ”Facial surface identification” [10]).

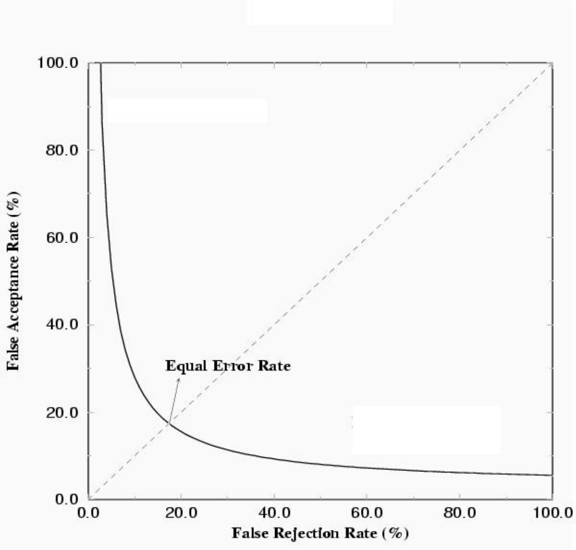

Figure 3: An example of a DET-curve. Equal Error Rate (EER) is the point on the graph where the False Acceptance Rate (FAR) is equal to the False Rejection Rate (FRR). The illustration is borrowed from the web site called ”Facial surface identification” [10].

The two axis in the DET-curve is the *False Acceptance Rate (FAR)* and the *False Rejection Rate (FRR)*. False Acceptance Rate is defined as the number of falsely accepted users divided by the total number of unauthorized users, while False Rejection Rate is defined as the number of falsely rejected users divided by the total number of authorized users [1, 3, 6]:

$$\mathsf{FAR} = \frac{\text{Number of falsely accepted users}}{\text{Total number of unauthorized users}} \tag{2.1}$$

$$\mathsf{FRR} = \frac{\text{Number of falsely rejected users}}{\text{Total number of authorized users}} \tag{2.2}$$

In order to find the the threshold where both the FAR and the FRR are as small as possible, we insert a line, FAR = FRR, in the coordinate system, and find the point where this line is crossing the DET-curve. This is also shown in Figure 3. We can see that this line is crossing the plot at approximately 18%. Hence, the *Equal Error Rate (EER)* in this case is 18%. If we want both the FAR and the FRR to be as low as possible, then this threshold should be chosen. The strength of a biometric system is shown when we study the EER-rate. A low EER indicates that the system is rather solid, i.e. that both the FAR and the FRR can be generally low. A high EER, on the other hand, indicates that the system is rather weak, i.e. that both the FAR and the FRR are generally high. Different authentication systems, and different algorithms have different EER. An example of this is shown by Chen et al. [11], who presented a prototype design and an implementation of secured mobile phones based on embedded fingerprint recognition systems, with an EER of 4.16%. N.L. Clarke and S.M. Furnell [12], on the other hand, used keystroke analysis in order to authenticate mobile phone users, with an EER of 12.8%. By looking at these EERs, we can see that the prototype design from Chen et al. is stronger than Clarke and Furnell's keystroke system, even though the EERs were not that high in any of these two cases.

2.3 Gait as authentication

2.3.1 Challenges

According to Naresh P. Trilok, Sung-Hyuk Cha, and Charles C. Tappert [13], the voice appears to be unique to the individual, and according to Chiara Braghin [14], fingerprints are also unique. Hence, these are features that can be used to authenticate a person. Our gait is probably unique the same way as our voice and fingerprints are. However, in the same way that our voice can differ from day to day, e.g. because of a cold, our gait can also differ, based on our condition.

There are many reasons why our gait is the way it is. The physical build and body weight are factors that can affect the way we walk, but there are also other factors, such as shoe type, heel height, clothing, illness, injury, emotional state, and environment [1]. These are some of the challenges we will have to face when authenticating someone by their gait. If the threshold is chosen too high, then there is a great risk that even the authorized users of the system will be rejected most of the time.

2.3.2 Common gait features

There are some features in a person's gait that are common for everyone. M.P. Murray, A.B. Drought and R.C. Kory [15] considered gait to be ”a total walking cycle”, and that the action of walking could be visualized as a periodic signal. An example of a walking cycle is shown in Figure 4 (from A. Jain, R. Bolle, and S. Pankanti's book ”Biometrics - Personal Identification in Networked Society” [7]).

According to Jain, et al. [7], a gait cycle is the time interval between successive instances of initial foot-to-floor contact 'heal strike' for the same foot. Both legs have two distinct periods, which are a *stance phase* and a *swing phase*. The stance phase is when the foot is in contact with the floor, and the swing phase is when the foot is off the floor moving toward the next step. The swing phase starts when the toes leave the ground, and the weight is transferred to the other leg, which swings forward to meet the ground in front of the other foot. A *step* is the motion between the heel strikes of opposite feet, such that a gait cycle actually consists of two steps [7]. In Figure 4, we can see that the

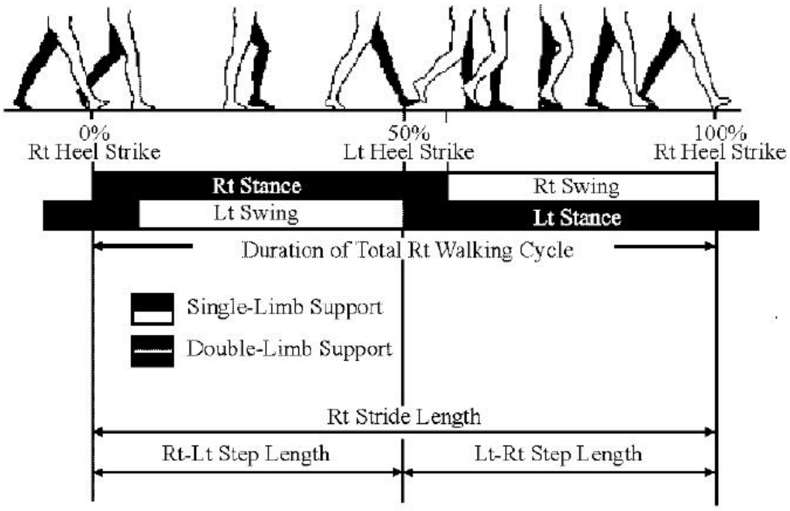

Figure 4: An example of a walking cycle, which is the time interval between successive instances of initial foot-to-floor contact 'heal strike' for the same foot. It consists of a stance phase and a swing phase for each foot. The illustration is borrowed from A. Jain, R. Bolle, and S. Pankanti's book "Biometrics - Personal Identification in Networked Society" [7].

first step starts at 0%, while the second step starts at 50%.

2.3.3 Three categories of gait authentication

When authenticating a person by his gait, the body motion is translated into numbers that a computer can meaningfully recognize [1]. There are different ways to do this.

The two most common ways to authenticate a person by his gait, are either to use a camera [16, 17, 18, 19], or to use a sensor containing an accelerometer [3, 4, 5, 6, 20, 21, 22]. In this thesis, we will address these two categories as *image-based gait authentication* and *accelerometer-based gait authentication*, respectively. There is also a third way of authenticating a person by his gait, called *floor-sensor based gait authentication* [23, 24], but since this method is not as much used as image-based and accelerometer-based gait authentication, we will not get any further into that one in this report.

Image-based gait authentication

This is the most studied authentication category among the three. When authenticating a person using this method, a camera catches a series of images of the authenticator. When these series are taken, a program analyzes the images in order to find central parts of the authenticator's body. The goal is to get information about how the authenticator moves these body parts, e.g. how fast he walks, how he swings his legs, and how he moves his hips.

Chew Yean Yam, Mark S. Nixon, and John N. Carter from University of Southampton [18] used image-based gait analysis as well. People were filmed, and the video clips were digitized into individual color image files and cropped, in order to reduce computational cost. A sobel edge operator was applied in order to obtain only the leading edge [18]. This is also shown in Figure 5 (from [18]).

In the (a) image, we can see one of the ordinary images that was taken by the camera. In (b), details around the person have been removed, such that the person is the main

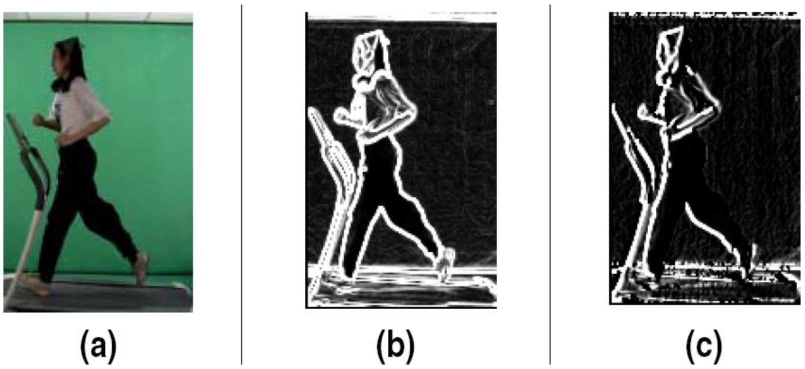

Figure 5: Analyzing an image in order to find relevant gait information. The illustration is borrowed from Chew Yean Yam, Mark S. Nixon, and John N. Carter's article "Performance analysis on new biometric gait motion model" [18].

objective in the image. In the last image, (c), relevant parts of the authenticator's body are found and highlighted. After these operations are done, the authentication process may start.

It is possible to capture images from different angles, something which was done by Guoying Zhao, Guoyi Liu, Hua Li, and Matti Pietikäinen [19]. The advantage of doing this, is that it will lead to a 3-dimensional analysis, since one gets information from different angles.

Accelerometer-based gait authentication

Another way of authenticating a person by his gait, is to use a sensor containing an accelerometer. This method is not as common as the image-based authentication, but has become increasingly more popular. One advantage of using this method, is that the amount of data to be stored is much smaller than with image-based authentication. With image-based authentication, several images are stored, and in order to get some relevant information, a great amount of image analysis has to be performed. Images captured in image-based authentication systems often contain much noise, which makes the analysis more difficult to perform. The acceleration data, on the other hand, contains only information about the gait. Besides this, image-based gait authentication is also a more complicated method than accelerometer-based authentication, because cameras have to be installed at the right places, in order for the authenticator to get his images captured from the correct angles. With all these arguments taken into consideration, accelerometer-based authentication seems to be the best alternative.

In order to be authenticated using this method, the authenticator has to wear a sensor that is communicating with a computer. The sensor contains an accelerometer, which is measuring the acceleration in x-, y-, and z-direction. This sensor can be worn on many places of the body, e.g. on one of the ankles [6], on the hip [3], and there have even been experiments where the sensor has been combined with a shoe [25]. Examples of this are shown in Figure 6 (from S.J. Morris and J.A. Paradiso's article "A compact wearable sensor package for clinical gait monitoring" [25]), in Figure 7 (from Torkjel Søndrol's master's thesis "Using the human gait for authentication" [6]), and in Figure 8 (from Tor Erik Buvarp's master's thesis "Hip movement based authentication. How will imitation affect the results?" [3]).

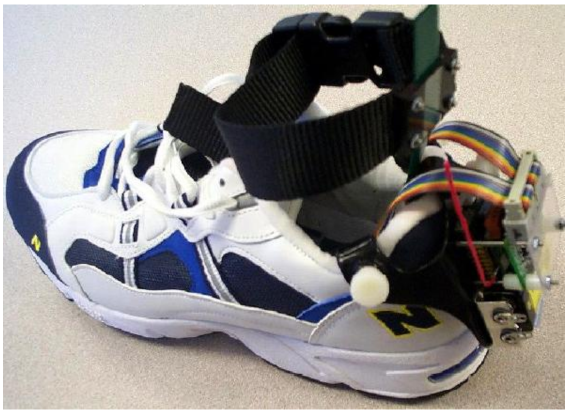

Figure 6: A sensor combined with a shoe. The illustration is borrowed from S.J. Morris and J.A. Paradiso's article "A compact wearable sensor package for clinical gait monitoring" [25].

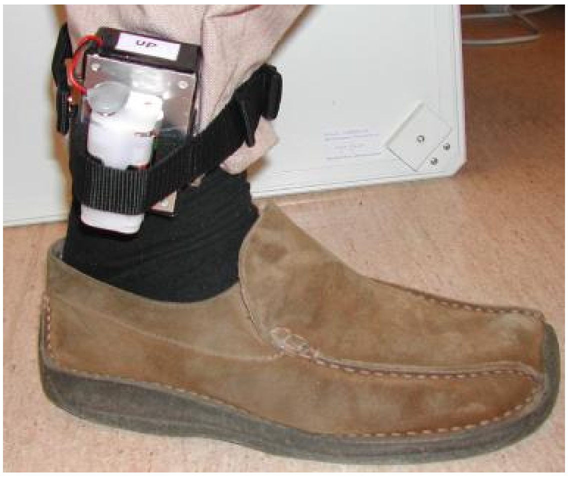

Figure 7: A sensor placed on the ankle. The illustration is borrowed from Torkjel Søndrol's master's thesis "Using the human gait for authentication" [6]

In the experiments to come, accelerometer-based gait authentication is the only method that is going to be used. The sensor will be placed in the authenticator's right pocket.

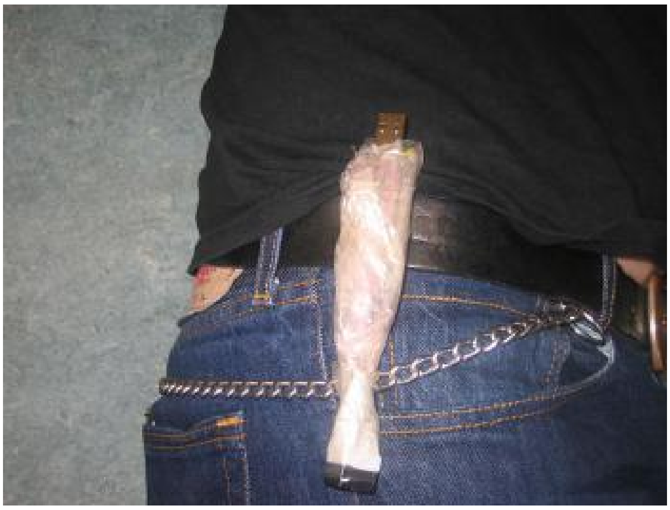

Figure 8: A sensor placed on the hip. The illustration is borrowed from Tor Erik Buvarp's master's thesis ”Hip movement based authentication. How will imitation affect the results?” [3].

3 Previous work

In this chapter, we will take a look at previous work on impersonation in biometrics. We will also give a short description of an experiment that was performed in order to see how easy it was to mimic the gait of someone else.

3.1 Imitating biometrics in general

Imitation of other people's features is a problem within most areas of biometrics. This is often called *the inverse problem* of biometrics [26]. Within signature identification, there is possible to forge signatures, and within keystroke dynamics, there is possible to learn a specific rhythm, or to learn how hard to press the keys. Within face recognition, it is possible to create face reconstructions and to mimic animations, within voice identification, it is possible to imitate a person's voice, within iris and retina identification, it is possible to create an iris or retina image synthesis, and within fingerprint identification, it is possible to create fingerprint ”stamps”, e.g. from gelatin [26, 27].

3.2 Signature

There are two categories of signature authentication, i.e. *static* and *dynamic* authentication. According to Anil Jain, Lin Hong, and Sharath Pankanti's article *”Biometrics: Promising frontiers for emerging identification market”* [28], static signature verification uses only the geometric features of a signature, while dynamic signature verification uses both the geometric features such as acceleration, velocity, and trajectory profiles of the signature. The circumvention of signature authentication is rather low, i.e. it is not considered too difficult to forge a signature [28].

3.3 Voice and speech

Bryan L. Pellom and John H. L. Hansen [29] investigated the relative sensitivity of a GMM-based voice verification algorithm to computer voice-altered impostors. Impostor voices were recorded and altered by this algorithm, in order to mimic the customer's voice. Before this experiment, the FAR was found to be 1.45%. After the experiment, it was increased to more than 86% [29]. This shows that it is rather easy to forge a voice based authentication system as well. Jain et al. [28] classified the circumvention of voice authentication systems to be low as well.

3.4 Face

Within criminology, face reconstruction has been a well-known problem for quite a while [26]. Even so, it is still not possible to prevent this type of fraud from happening, basically because face recognition is a complex task. Synthesises of face images can be either static [30] or animated [31]. Blanz and Wetter [30] found a way to take a 2D input image of a face, do some face analysis on the image, and create a 3D output image in order to use it as a face synthesis. Animated synthesises of face images are more complicated. P. Fua and C. Miccio [31] showed that one could fit complex animation models to noisy data extracted from ordinary face images. They based their approach on least-squares adjustments, by using a set of progressively finer control triangulations, and by taking advantage of three sources of information, namely stereo data, silhouette edges, and 2D feature points [31].

3.5 Iris and retina

Two of the most difficult biometrics to forge, are the iris and the retina. One of the reasons is because glints of light are often captured in the image, and hence creates fields called *the uncertainty*, because they are difficult to interpret. However, the loss of information can still be evaluated, and a partial reconstruction can be done, by using different prediction techniques [26]. Jiali Cui, Yunhong Wang, JunZhou Huang, Tieniu Tan, and Zhenan Sun [32] used methods called *Principal Component Analysis (PCA)* [33] and *super-resolution* [34], in order to create iris and retina image reconstructions. The synthesis method first constructed coarse iris images with given coefficients. In the next step, synthesized iris images were enhanced using super-resolution. Many iris images were created by controlling the coefficients. Their conclusion showed that synthesized iris images have satisfactory cluster, and that synthesized iris databases can be of large size [32].

3.6 Fingerprints

One of the most traditional biometrics is fingerprints. A person's fingerprints may change, e.g. because of an accident or a surgery, and in order to evaluate the fingerprint system's performance in such situations, the system is trained using some imitated fingerprint images [26]. D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar described different methods of generating synthetic fingerprints in their *"Handbook of Fingerprint Recognition"* [35]. According to D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar [36], they found an EER of 4.55%, which means that there is a certain chance for an attacker to be accepted by a system using fingerprint authentication.

3.7 Gait

The main topic in this thesis is whether it is easy or difficult to learn to walk like another person. Few works have been published on the topic of gait mimicking, except the article *"Robustness of biometric gait authentication against impersonation attack"*, that was written by Davrondzhon Gafurov, Einar Snekkenes and Tor Erik Buvarp [5].

The distance measure used in this article is the *"Cycle Length Method"*, and the main point is that it compares the average cycle of the authenticator's gait graph with the average cycle of the template gait graph. A more detailed description can be found in chapter 7.1.2. Figure 9 (from D. Gafurov, E. Snekkenes, and T.E. Buvarp's article "Robustness of biometric gait authentication against impersonation attack" [5]) shows the different steps in this method.

The aim of this experiment, was to investigate the difference between so-called *"passive impostor attempts"* and *"active impostor attempts"*. A passive impostor attempt is an attempt where a person submits his own biometric feature as if he was attempting successful verification against his own template, but in fact is being compared against a non-self template. An active impostor attempt is an attempt where a person deliberately changes his biometric in order to match another targeted person, and to be verified against this targeted persons template [5].

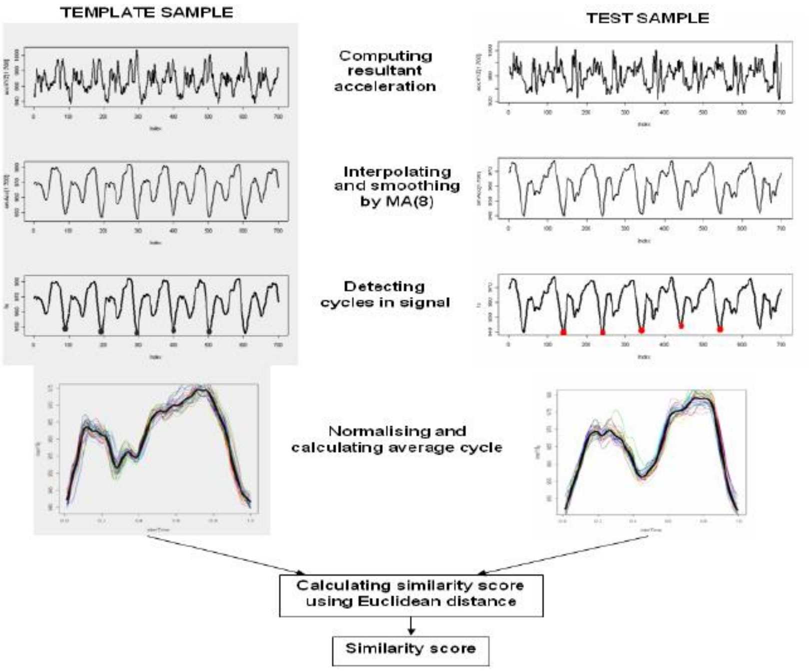
TEMPLATE SAMPLE
TEST SAMPLE
Computing resultant acceleration
Interpolating and smoothing by MA(8)
Detecting cycles in signal
Normalising and calculating average cycle
Calculating similarity score using Euclidean distance
Similarity score

Figure 9: The Cycle Length Method. The illustration is borrowed from D. Gafurov, E. Snekkenes, and T.E. Buvarp's article "Robustness of biometric gait authentication against impersonation attack" [5].

The results of this experiment showed that the EER in the "friendly scenario" was 16%. Amongst 330 good users, 53 were rejected, and amongst 6930 bad users, 1109 were accepted [5].

In order to analyze the "hostile scenario", they applied statistical techniques, so they could investigate the difference between the passive impostor trials and the active impostor trials. They calculated a *D-prime value*, which says something about the separability between two normal distributions [5, 37]. In this experiment, the results showed that the D-prime for the active impostor attempts was 1.415, and 1.142 for the passive impostor attempt. By analyzing these numbers, it seemed to be a certain difference between the genuine and the imitated gaits. It was also a greater separation between the genuine and the active impostor attempts, than between the genuine and the passive impostor attempts [5].

The next step was to compare the passive and the active attempts, and in order to do this, they stated the following hypothesis set:

$$H_0 : \mu_{active} \leq \mu_{passive}$$
$$H_A : \mu_{active} > \mu_{passive}$$

The null hypothesis (H_0) stated that the mean value, in the cases of active impostor attempts, was equal or lower than the mean value in the passive impostor attempts. It is important to notice that a good match in this experiment resulted in a low value, while a bad match resulted in a high value. Hence, if H_0 would turn out to be true, then this would mean that active attempts to imitate another person would actually give better

results, seen from the hostile impostor's point of view, than passive impostor attempts. The alternative hypothesis (H_A) was that the active impostor attempts would give a higher mean value than the passive impostor attempts, i.e. that deliberately attempting to imitate another person would *not* give better results than passive attempts, seen from the hostile impostor's point of view. They applied both parametric and non-parametric tests. The first test to be performed was a *t-test* [38], which resulted in a p-value of 0.0005. A *p-value* is a measure of how much evidence there is against the null hypothesis (H_0) [39].
Using the same hypothesis set as mentioned above, a non-parametric *Wilcoxon* (or *Mann-Whitney*) test was also performed [40], which resulted in a p-value of 0.000004 [5].

The mean values from the active impostor attempts seemed not to be specifically lower than the mean values from the passive impostor attempts. Hence, none of these results could be used as evidence to support the null hypothesis (H_0). They therefore concluded that active impostor attempts are *not* improving the attacker's chances of being accepted. [5].

3.8 Different biometrics and their EER

Table 2 shows a comparison of different biometrics and their EER, based on information from different sources. The EER indicates how easy or difficult it is to fraud the different biometrics. If the EER is low, then the biometric is solid, and impersonation is probably difficult. On the other hand, if the EER is high, then the biometric is less solid, and impersonation is probably less difficult.

Biometrics	EER	Reference
Fingerprints	4.55%	”Multimodal Biometric Authentication...” [36]
Face	3%-9%	”Frontal face authentication using morphological elastic graphmatching” [41]
Signature	2.84%	”SVC2004: First International Signature Verification Competition” [42]
Voice	1.45%	”An experimental study of speaker verification...” [29]
Iris	0.105%	”Iris Feature Extraction Using Independent Component Analysis” [43]
Hand geometry	0.98%	”Implementation of hand geometry...” [44]
Keystrokes	1.8%	”Fusion of methods for keystroke dynamic authentication” [45]

Table 2: An overview of different biometrics and their Equal Error Rates. The performance of a biometric system is measured using the EER.

4 Choice of technology

This chapter describes the different choices we have made about the sensor, the program development tools, and the tools for statistical analysis.

4.1 Sensor

In our experiment, we used a sensor called ”*ZSTAR*” from Freescale. It contains a 3-axis accelerometer [46], and a 2.4 GHz frequency MC13191 transceiver [47]. The sensor is shown in Figure 10 (from the manual ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47]), and the batteries are shown on the right side.

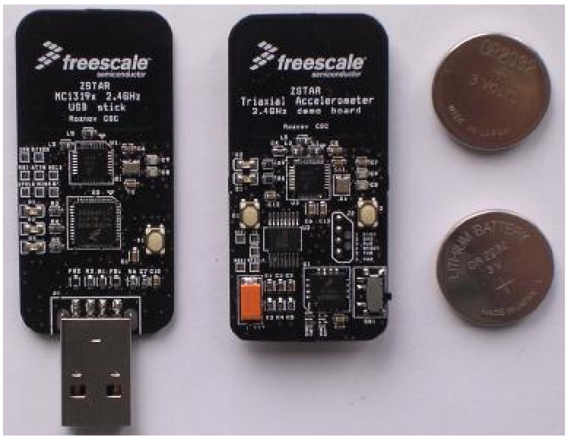

Figure 10: A USB stick, a sensor board (containing the 3-axis accelerometer), and 2 Lithium coin-sized batteries. The illustration is borrowed from the manual ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47].

The sensor tool from Freescale consists of two boards [47]:

- Sensor Board (or remote board), which contains the MMA7260QT 3-axis accelerometer, S08 family MC9S08QG8 8-bit microcontroller, and the 2.4 GHz RF chip MC13191 for wireless communication. This is the equipment in the middle of Figure 10.
- USB stick with the MC13191 RF front-end, and the HC08 family MCHC908JW32 for the USB communication. This is the equipment to the left in Figure 10.

The sensor board is powered by a Lithium coin-sized CR2032 battery. Figure 11 (from ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47]) shows a block diagram describing the sensor board.

According to ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47], the main tasks of the sensor board are the following:

- Periodically wake up from power saving mode.

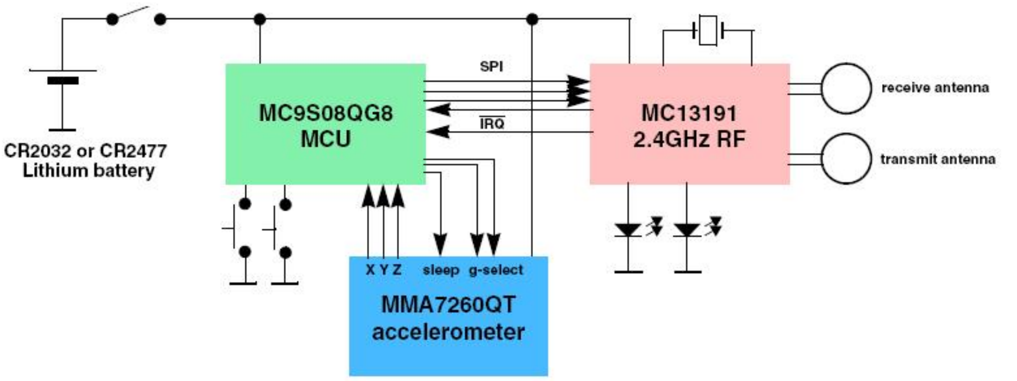

Figure 11: Sensor board block diagram. The illustration is borrowed from ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47].

- Measure all three (X, Y, Z)-acceleration values from the sensor.
- Compose a data frame using simple ZSTAR RF Protocol.
- Use SMAC (Simple Media Access Controller) to send this data over the RF link.
- Wait for an acknowledgement from the other end (e.g. the USB stick).
- Go to sleep.

These tasks are repeated by the sensor board approximately 30 times per second, and will therefore provide a nearly real-time response from the sensor [47].

In order to receive the acceleration data from the sensor board, the PC sends a $'V'$ to the sensor, and the sensor responds by sending 6 bytes in the following sequence: The ASCII character of ’x’, the X-value, the ASCII character of ’y’, the Y-value, the ASCII character of ’z’, and finally the Z-value. This protocol is also shown in Figure 12 (from ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47]).

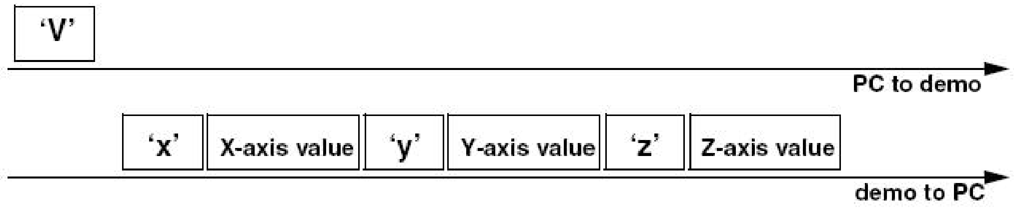

Figure 12: The protocol between the PC and the sensor board. The illustration is borrowed from ”Wireless Sensing Triple Axis Reference Design - Designer Reference Manual” [47].

4.2 Program development tools

In order to develope the prototype, we used Visual Studio.NET as the programming tool, and C++ as the programming language. The reason why we choose Visual Studio.NET, is because it is a very practical tool containing many helpful resources and functions.

When starting Visual Studio.NET, it gives us several options about what type of project to create, and when the decision is made, it is creating the whole project with all the foundational files and classes for us. Of course, the content of the files is still our job to program, but the main advantage using Visual Studio.NET is that we don't need to waste a lot of time working with details that are not that relevant to our project. Since Visual Studio.NET creates all the framework needed, we can fully concentrate on the coding.

Another advantage with Visual Studio.NET, is that it contains many sample libraries that are very helpful when programming. A lot of code has already been developed, so it would be a waste of time to start programming everything from the bottom. It is more practical to re-use already existing code, which we also trust is working. In addition to this, Visual Studio.NET does also have a large documentation covering almost all the classes and libraries that exist in the different languages that Visual Studio.NET supports, C++ included.

An alternative would have been to use C# instead of C++, but since we had little experience using C#, we decided that C++ was an appropriate programming language.

4.3 Tools for statistical analysis

When performing statistical analysis of our results, we will use a statistical data management package, called SPSS [48]. This tool contains a lot of statistical functions that may be practical in our statistical analysis.
These are some of the features that SPSS contains [48]:

- Linear regression
- Anova analysis
- Factor analysis
- TwoStep cluster analysis
- Hierarchical cluster analysis
- Ordinal regression (PLUM)

Many statistical methods can be performed by using just Excel as well, but if we had chosen to do all our analysis in Excel, it would have become very complicated, in comparison to SPSS.

5 Prototype

In this chapter, we will look at the prototype that we have developed, i.e. the requirements to the prototype, and then a specification of the contents of the prototype.

5.1 Requirements

The main requirements to our prototype are the following:

- Read data from a sensor
  The prototype is going to read acceleration data from a sensor. The sensor used in this thesis is a ZSTAR sensor, which measures the acceleration in 3 directions, i.e. in the x-, y-, and z-direction. It should read these data continually as the user is walking in order to give feedback immediately to the user about his gait.

- Plot data as 4 graphs in a coordinate system
  The data received from the sensor is going to be plotted as 4 graphs in a coordinate system, i.e. the x-graph, y-graph, z-graph, and the r-graph. We will explain how the r-graph is calculated in the section about the display of graphs. These 4 graphs should be displayed with a certain gap in between, such that the user may differentiate between them. This is shown in Figure 13. 4 template graphs are going to be plotted when we are starting the program. When we give a message to the prototype, it should start reading and plotting data from the sensor. The user's graphs should then be displayed at the same location in the coordinate system as the template graphs, in order to see the match between the user's graphs and the template graphs. This will be done by giving the 4 graphs different colors. The template graphs will be displayed with lighter colors than the user's gait graphs. The reason to this, is because this makes it easier for the user to see which graphs are his and which graphs are the template graphs.

- Stop after approximately 5 seconds and calculate score
  After the prototype has read and plotted data in approximately 5 seconds, it should halt and compute a score between 0 and 100, based on how well the user's graphs matched the template graphs. A high score will be given if the match is good, while a low score is given if the match is bad. This score should be displayed in e.g. a pop-up box, in order for the user to get information about how close his imitation was. We hope this score will also motivate the user to try to improve his score even further.

- Storage of data
  In addition to displaying the score in a pop-up box, the scores should be stored in a text file, in order to be used in the statistical analysis.

5.2 Specification

5.2.1 Choice of graphical library

Since a central part of our prototype is to plot and compare graphs, it is practical to find a library that manages graphs. We found a graph library, which is based on open source, called *Graph2D* [49]. It has some functionality that can be used in the development of our prototype.

Graph2D contains a class, which provides multiplot 2D data visualization [49]. By using this functionality, it can plot more than one graph in the same coordinate system. Besides this, it can also draw so-called *”dynamic graphs”* which are graphs that are being continually plotted. These two qualities are very practical in our prototype, so we chose Graph2D as the platform of our program, by modifying it’s source code in order for it to fit with our requirements. Instead of plotting a dynamic graph based on a given formula, like it did in it’s first version, it now reads acceleration data from the sensor and plots this in the coordinate system. All in all, choosing Graph2D as the platform of our prototype was not a difficult decision to make.

5.2.2 Starting the process

When the program starts, we choose the option *”Graph”* and then *”Start dynamic graph”* in the menu on the top of the screen. The template graphs will then be displayed in the coordinate system. When the user is ready to start walking, we choose *”Graph”* and *”Start dynamic graph”* once again, and the prototype starts reading acceleration data from the sensor, and plots this in the coordinate system. After 5 seconds, the program stops plotting, and a score between 0 and 100 is displayed in a standard pop-up box. How this score is calculated will be described in more detail below.

5.2.3 Display of graphs

4 template graphs are plotted in the coordinate system, i.e. the x-graph, y-graph, z-graph, and the r-graph. The r-graph represents the *resultant values*, where each plot, i, is calculated using the following formula:

$$R_i = \sqrt{X_i^2 + Y_i^2 + Z_i^2}, \ i = 1, ..., n \tag{5.1}$$

i represents a specific point on the horizontal axis, and n is the total number of plots.

The user’s graphs and the template graphs are visualized in the same coordinate system. Both the x-, y-, z-, and r-graphs are being continually plotted as the program is receiving information from the sensor about the acceleration in these three directions. An example of this is shown in Figure 13. The red graph is the acceleration in x-direction, which is the direction forward and backward. The green graph is the acceleration in y-direction, which is the direction up and down. The blue graph is the acceleration in z-direction, which is the direction from side to side, while the black graph is the resultant graph. The directions of the sensor are shown in Figure 14. These are the directions when the user is placing the sensor in his right pocket, with the sensor light pointing out from the body and ahead. The big arrow in the upper right corner is the direction in which the user is walking.

A point is plotted in the graph when the information about it is received by the program. In order to get the gap between the 4 graphs, the 500 is added to the value of the X-points, 1000 is added to the value of the Y-points, 2000 is added to the value of the Z-points, and 2500 is added to the value of the R-points.

5.2.4 Graph matching and calculation of scores

In order to get a good score, the user’s graphs should look more or less like the template graphs. Since all these graphs contain important information about the gait, then they

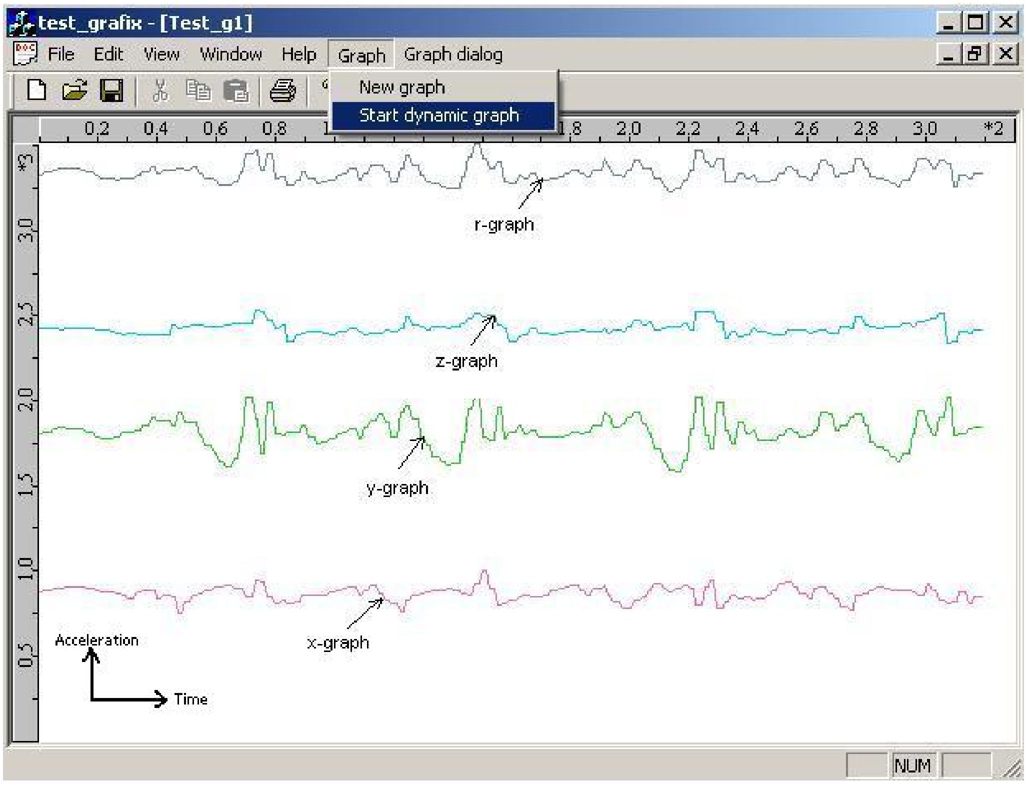

Figure 13: Screenshot from the prototype, with some additional information. The x-, y-, z-, and r-graph are plotted with gaps in between. The red graph is the x-graph, the green graph is the y-graph, the blue graph is the z-graph, and the black graph is the r-graph. The vertical axis represents the acceleration, while the horizontal axis represents the time.

should all be affecting the score. One way to do this, is for the matching procedure to consider only the r-graphs, since they contain the resultant values based on the x-, y-, and z-graph. So two rather similar r-graphs should therefore result in a high score, while two dissimilar r-graphs should result in a low score. Figure 15 and 16 are showing a good and a bad match.

There are plenty of ways to look for resemblances between two graphs. One way to do this, is to count the number of cycles, or to find their average length. Another way is to measure the distance between the original graph and the impersonator graph. However, we found a better way to solve this, i.e. by calculating the correlation coefficient between the user's r-graph and the template r-graph.

Correlation is a statistical method that gives us information about how strong the linear connection is between two sets of data, A and B [38]. The correlation coefficient, ρ, is a value between -1 and 1, where a ρ close to -1 or 1 implies that the connection between the two sets is strong, and a ρ close to 0 implies that the connection between the two sets is weak. The sign (+ or -) indicates the direction of the connection. A positive ρ implies that a large value in A has a large probability of being accompanied by a large value in B. The other way around, a negative ρ implies that a large value in A has a large probability of being accompanied by a small value of B [38].

The most known correlation coefficient is called *"Pearson's correlation coefficient"* or

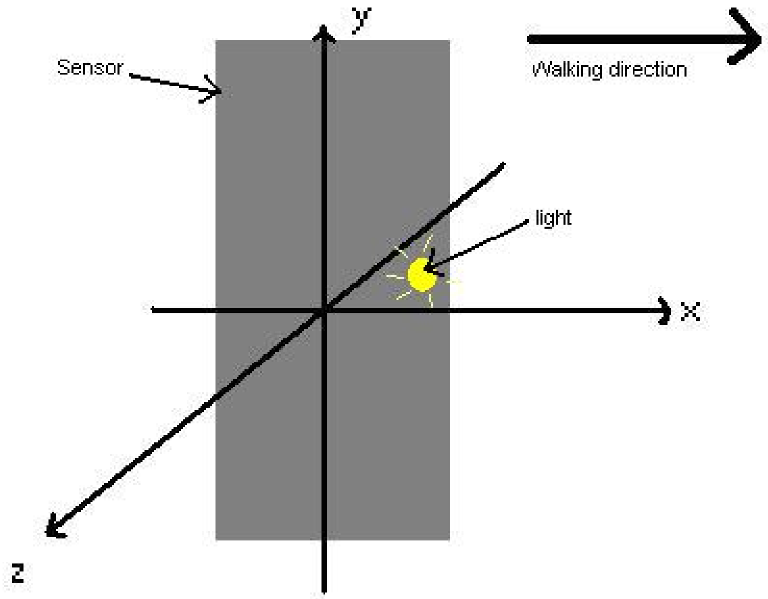

Figure 14: A sketch of the sensor and it's directions. The user is wearing the sensor in his right pocket, with the sensor light pointing out from his body and ahead, as shown in this illustration.

"The product-moment correlation coefficient" [50]. This coefficient is the one that was used in our experiment, in order to calculate the match between two sets of data, i.e. the template r-graph and the user's r-graph. If we have two datasets, A and B, the correlation between them is given by the following formula [50]:

$$\rho = \frac{\sum (a_i - \overline{a})(b_i - \overline{b})}{\sqrt{\sum (a_i - \overline{a})^2 \sum (b_i - \overline{b})^2}}, \qquad (5.2)$$

where a_i and b_i are the i-th values of the two sets A and B, respectively.

Different statisticians have different ways of interpreting what is a good match and what is not a good match between two sets of data, when studying correlation coefficients. Jacob Cohen [51] suggested the following interpretation of the correlation coefficient, which was supposed to be used within psychological research [51]:

$\lvert\rho\rvert$	Interpretation
0.10 - 0.29	Small
0.30 - 0.49	Medium
0.50 - 1.00	Large

Table 3: Jacob Cohen's interpretation of the correlation coefficient [51].

According to Cohen, this interpretation should not be observed too strictly, because interpretations of a correlation coefficient depend strongly on the context they are supposed to be used in [51].

When the user has attempted to imitate the original gait once, the program calculates the correlation coefficient between the template r-graph and the user's r-graph, and in order to get a score between 0 and 100, the absolute value of the resulting correlation coefficient is multiplied with 100. This score is then displayed in a pop-up box. Figure 15 and 16 show two scenarios. In the first scenario, the graphs match rather well, while in the second scenario, the graphs match rather badly.

$$\text{score} = |\rho| \bullet 100 \tag{5.3}$$

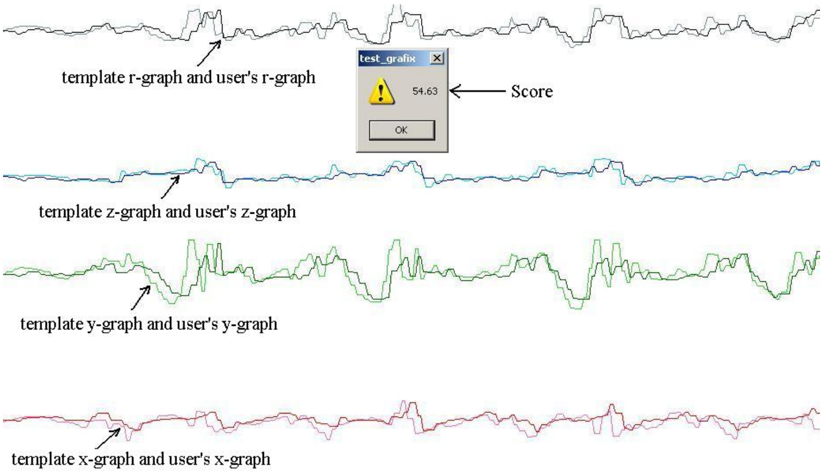

Figure 15: The graphs with light colors are the template graphs, while the graphs with dark colors are the user's graphs. The correlation between the template r-graph and the user's r-graph is in this case 0.5463, and hence the score is 54.63 points. This is a rather good result.

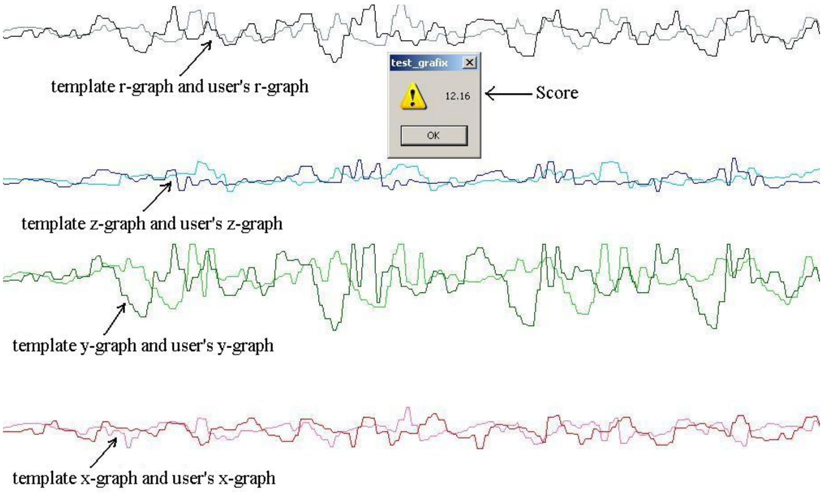

Figure 16: The graphs with light colors are the template graphs, while the graphs with dark colors are the user's graphs. The correlation between the template r-graph and the user's r-graph is in this case 0.1216, and hence the score is 12.16 points. This is a rather bad result.

5.2.5 Storage of gait data

When the prototype receives acceleration data from the sensor, the values are stored in text files, in order to be kept when the program is closed.

The files containing the gait data look like this:

```
...
136 205 135 280
137 208 135 283
137 208 135 283
134 204 137 279
134 204 137 279
134 204 137 279
134 204 137 279
134 204 137 279
134 204 137 279
134 205 137 280
134 205 137 280
134 205 137 280
132 206 136 279
132 206 136 279
132 207 134 279
132 207 134 279
131 206 125 274
131 206 125 274
...
```

The first column contains the X-values, the second column contains the Y-values, the third column contains the Z-values, and the forth column contains the R-values.

After 5 seconds of attempting to mimic the template, the user's score is first displayed in a pop-up box, and then it is written to a file, which looks like this:

```
3.65604
30.7824
28.1778
35.9378
52.8497
41.6727
42.8344
49.5097
21.2603
37.9778
45.469
17.1309
50.5844
54.5494
43.0367
```

6 Experiment and Data Analysis

In this chapter, we describe the experiment that we performed. We will also analyze the data from the experiment, in order to find out whether it seems easy or difficult to learn to mimic the gait of someone else. In the end of the chapter, we will present a DET-curve, describing the performance of our system.

6.1 Experiment

The first we did, was to create 5 templates with different degrees of difficulty. A description of these are given in the table below, and the graphs are shown in Figure 17-21.

Table 4: Description of the 5 gait templates.

Template	Description	Graphs
A	2 slow, although normal steps. The intention with this template was for the user to learn to move according to rather simple graphs, and to see how his movements were interpreted by the program.	Figure 17
B	3 normal steps. More natural than in Template A. The intention was to try to learn to imitate more natural gaits, using the displayed graphs.	Figure 18
C	A normal gait, based on the author's natural way of walking. The intention with this template was to see how the user managed to learn to imitate a normal gait.	Figure 19
D	Fast and ”shuffling” steps from side to side. The intention was to see how the user managed to learn to imitate an abnormal gait.	Figure 20
E	Rather slow walk with ”oscillating” knees. The participant was supposed to walk like a ”3D sine curve”. Like Template D, this template was also created in order to check whether the user managed to learn to imitate a strange way of walking.	Figure 21

The experiment was done in the authentication lab on Gjøvik University College. An e-mail was sent to all of the students at the faculty, and 13 participants volunteered to contribute in this experiment. They were all men, but of various age, height and weight.

All participants got 15 attempts on each template. Before each template, they did not see the original walking, but they were instead given a simple description of the gait. The template graphs were shown on a big screen, and when the participants started walking, their gait graphs were continually plotted in the same coordinate system as the template graphs. By showing the graphs on a big screen, the participants could look at the graphs while they were walking. A simple drawing is shown in Figure 22. Between

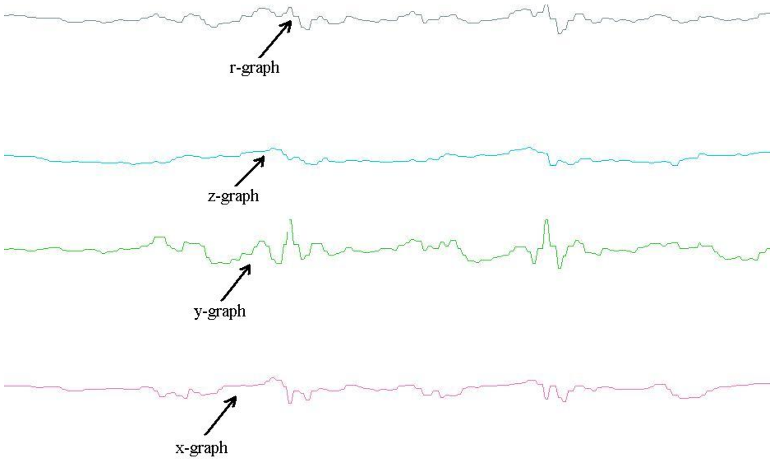

Figure 17: Template A: 2 slow, although normal steps. The cycles are not that clear to see because of the slowness.

each attempt, a score between 0 and 100 was shown to them in a standard pop-up box. The reason why we wanted them to walk 15 times, was to see whether there was any progress in the scores from the beginning to the end. Each attempt lasted 5 seconds, and the experiment lasted about 20-30 minutes in total. The participants carried the sensor in their right pocket, with the sensor light pointing out from their body and ahead. In every template, they started walking with their left foot first.

6.2 Data analysis

6.2.1 Linear regression

The next part of the work was to analyze the results from the experiment, in order to see whether the program had any effect on the learning. In the experiment, some participants had problems adjusting their gait in order for their graphs to match the template graphs. They got more or less the same low score every time. Others had no problem following the original graphs, and managed to adjust their walk rather quickly. These participants got high scores almost all the time. Some participants did also get low scores in the beginning of their attempts, but managed to get higher scores after they had walked for a while. Our plan was therefore to find a *linear regression* between the results from each user in each template. Linear regression models the relationship between two variables by fitting a linear equation to the observed data [52]. A linear equation has the form $y = mx + b$, where m is the slope, and b is the position on the vertical axis where the line is crossing. Examples of plots and their linear regression line are shown in Figure 24, 25, and 26.

The intention was to find the angle of the linear equation, and to use this information in order to see whether the scores had increased from the beginning to the end of the 15 attempts. If the angle turned out to be a large positive number, then the improvement was considered good, and it would imply that our program had a good learning effect. On the other hand, if it turned out to be a small or a negative number, then it was considered not that good. It would also imply that our program had no learning effect. The angle

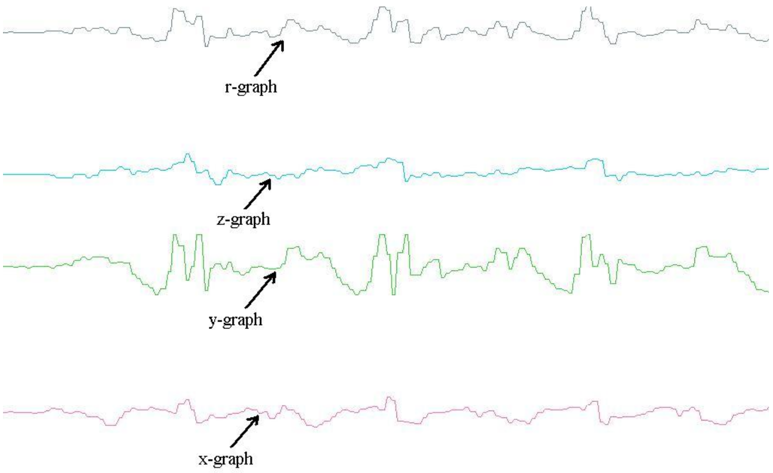

Figure 18: Template B: 3 normal steps. More natural than in Template A. The cycles are more clearly visualized, e.g. in the y-graph.

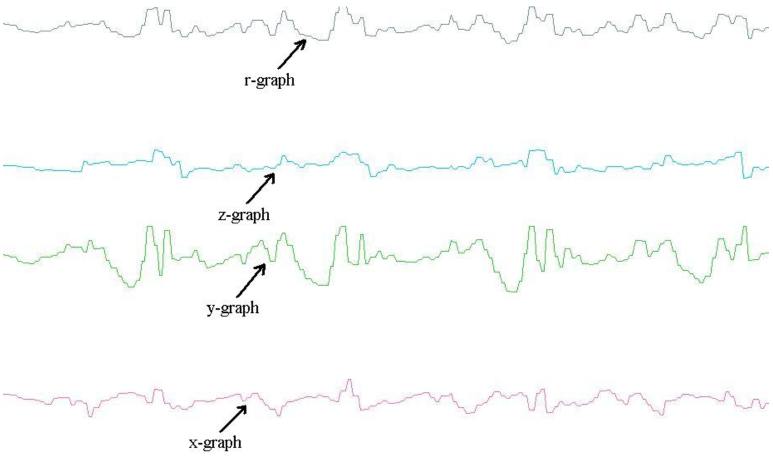

Figure 19: Template C: A normal gait, based on the author's natural way of walking. The cycles are rather clear to see.

θ was found by studying the linear regression line graphically. As shown in Figure 23, if we look at the linear graph from t to $t+1$, then the increase between these two points on the horizontal axis is m, and the length between t and $t+1$ is 1. From trigonometry, we know that [53]:

$$\tan\theta = \frac{opposite}{adjacent}. \tag{6.1}$$

In this case,

$$\tan\theta = \frac{m}{1} = m, \tag{6.2}$$

hence:

$$\theta = \arctan m. \tag{6.3}$$

In this report, θ will be presented as degrees.

By looking at the angles, which are shown in *Appendix A*, we can see that most of the users got a positive angle on all the templates. However, some did also get a negative

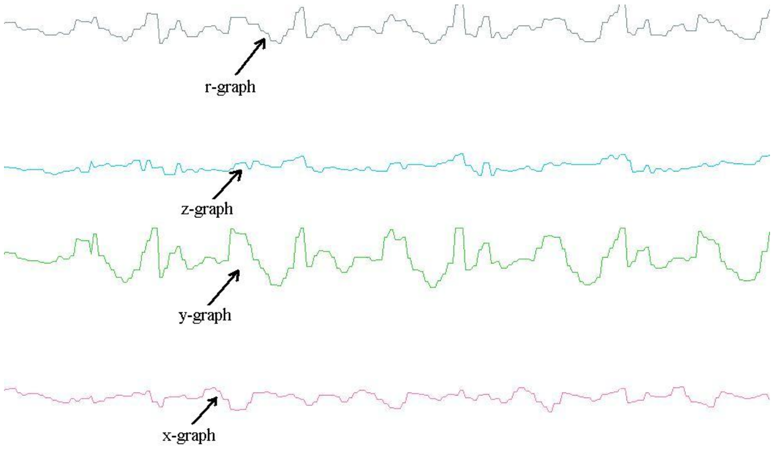

Figure 20: Template D: Fast and "shuffling" steps from side to side. Since the speed is high, there are more gait cycles. This is clearly shown e.g. in the y-graph.

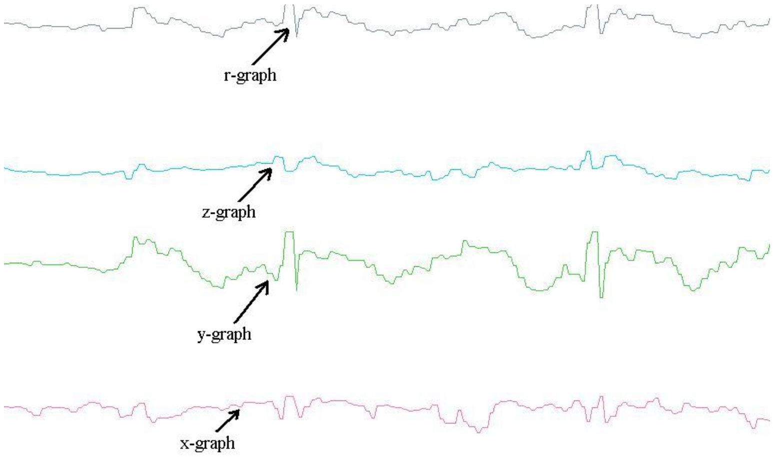

Figure 21: Template E: Rather slow walk with "oscillating" knees. The y-graph is clearly oscillating.

angle, and some of the positive angles were rather small. Table 5 shows the number of participants who got a negative, a small positive, and a large positive angle. A negative angle is in this thesis defined as an angle which is either 0 or negative, a small positive angle is defined as an angle in the interval between 0 and 25 degrees (55.6% increase), and a large positive angle is defined as an angle larger or equal to 25 degrees.

These numbers may give us the impression that the results were rather stochastic. A majority, i.e. 7 participants, got a large θ in Template A and Template E, but the θ from the other three templates were sometimes large, sometimes small, and sometimes negative. The results from the experiment do also show that many participants got lower scores in the end of their 15 attempts on each template. The reason to this could be that some of the participants got gradually tired of walking, and hence walked rather randomly at the end of their attempts, without having the best concentration. We tried therefore to look only at the first 10 attempts in each template, in order to see whether that would give us more informative results.

The average angles, based on all of the 15 attempts, were calculated to be 27.87

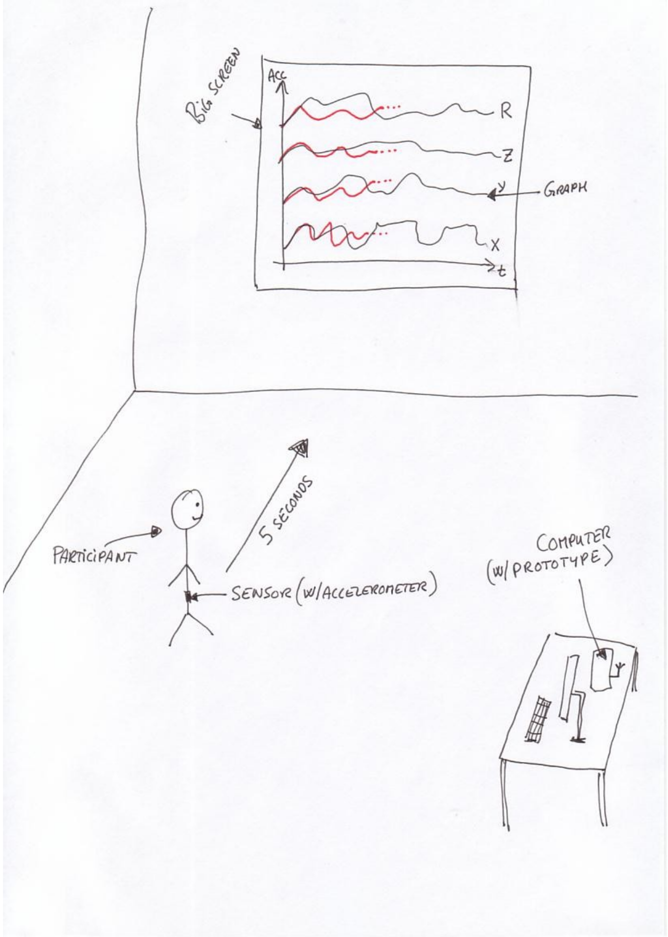

Figure 22: The participant walked toward a big screen displaying the template graphs, and the participant's gait graphs, which were continually plotted.

(61.9% increase), 7.92 (17.6% increase), 19.59 (43.5% increase), 5.14 (11.4% increase), and 18.70 degrees (41.6% increase) for Template A, B, C, D, and E, respectively. When only the first 10 were considered, the average angles were calculated to be 24.79 (55.1% increase), 11.66 (25.9% increase), 22.55 (50.1% increase), 3.22 (7.2% increase), and 30.91 degrees (68.7% increase). These numbers are also shown in *Appendix A*. We see that in almost all of the templates, except Template A and D, the increase was larger during the first 10 attempts, than during all of the 15 attempts.

6.2.2 Study of different thresholds

The next we did, was to look at different thresholds, in order to see how often they were exceeded by the participants during their 15 attempts. Our intention was to see whether there was a certain threshold in which few participant were able to exceed. In the experiment, we could see that a good imitation often lead to scores from 40 points and above. We could also see that few managed to score more than 65 points. We therefore chose to look at the thresholds 40, 50, and 60. These results are shown in Figure 24, 25, and 26. The horizontal axis represents the 15 attempts, and the vertical axis represents the number of matches in total, i.e. in all of the 5 templates. We also

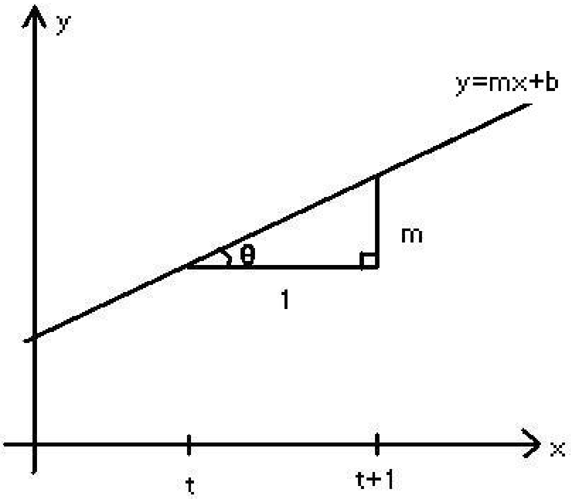

Figure 23: The angle (θ) of a line ($y = mx + b$), is found by calculating $\arctan m$.

Template	$\theta \in$ <-∞, 0]	$\theta \in$ <0, 25>	$\theta \in$ [25, ∞>
A	3	3	7
B	5	3	5
C	6	1	6
D	6	3	4
E	3	3	7

Table 5: The results from the experiment concerning negative, small positive, and large positive θs.

found the linear regression $y = mx + b$ for each of the thresholds, and these lines are also plotted into the graphs in Figure 24, 25, and 26. If the number of matches in total had increased during the attempts, the line would be increasing, and if the number of matches in total had decreased during the attempts, the line would be decreasing.

From these graphs, we can see that the angle representing the number of times when the participants scored more than 40 points is clearly increasing during the 15 attempts. However, the angle representing the number of times when the participants scored more than 50 points is not increasing that much, and the angle representing the number of times when the participants scored more than 60 points is actually *decreasing*. In other words, the increase is very clear when the threshold is chosen somewhere around 40 points, but when the threshold is chosen somewhat higher, the number of scores above the different thresholds is getting less, or is even decreasing during the 15 attempts. It therefore seems like learning to mimic a gait is easy when the threshold is approximately 40 points, while it is more difficult when it is higher than that. As a result of this, one should be careful choosing a threshold too low.

6.2.3 Study of the author's natural gait

In the real world, a biometric system will only return *"Passed"* or *"Failed"*, based on whether the user has scored above or below a chosen threshold. Some of the participants

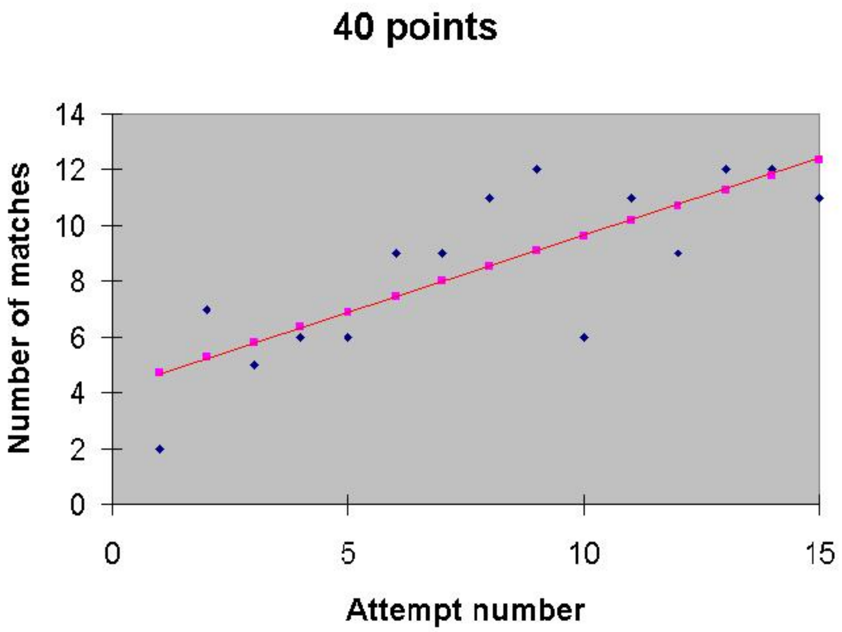

Figure 24: The number of matches when the threshold is 40 points. The θ in this case is 28.50 degrees (63.3% increase), which indicates that most of the 40 points scores came towards the end of the attempts.

in our experiment got generally high scores, e.g. around 50 points, and hence, their progression were not as remarkable as those who first got 0.5 points and then 10 points, even if that actually is a lot worser result. In a real biometric system, it will not matter whether a user gets 0 points or one point below the threshold. No matter how many points he gets below the chosen threshold, he will be rejected. That is the case the other way too. No matter if a user gets a number of points equal to the chosen threshold, or if he gets 100 points, he will be accepted. With that in mind, we decided that it might also be interesting to see how many of the participants that managed to exceed a certain threshold. We did also check how many attempts they needed in order to manage this. In order to choose a sensible threshold to use in this study, we checked how good scores the author was able to get when trying to walk like *himself*.

Since Template C contained the author's natural walk, he tried to ”imitate himself” by taking the experiment 10 times, using only that template. Since all of these 10 rounds consisted of 15 attempts, we ended up with a total of 150 different scores. These scores are shown in *Appendix B*. We used SPSS to create a cumulative table containing sorted scores, which is shown in *Appendix C*. In a biometric system based on gait, we decided that to be falsely rejected every second time is acceptable. This is the same as accepting an FRR of 50%. The reason why we accept this threshold, is because the most important aspect is to make sure that few unauthorized persons manage to attack the authentication system. According to the DET-curve in Figure 27, a FRR of 50% will lead to a FAR at approximately 10%. If we had chosen a lower FRR, e.g. 20%, then we would have to choose a threshold of 25 points, according to our cumulative table. To choose a low threshold like this would almost guarantee that many impersonators would manage to

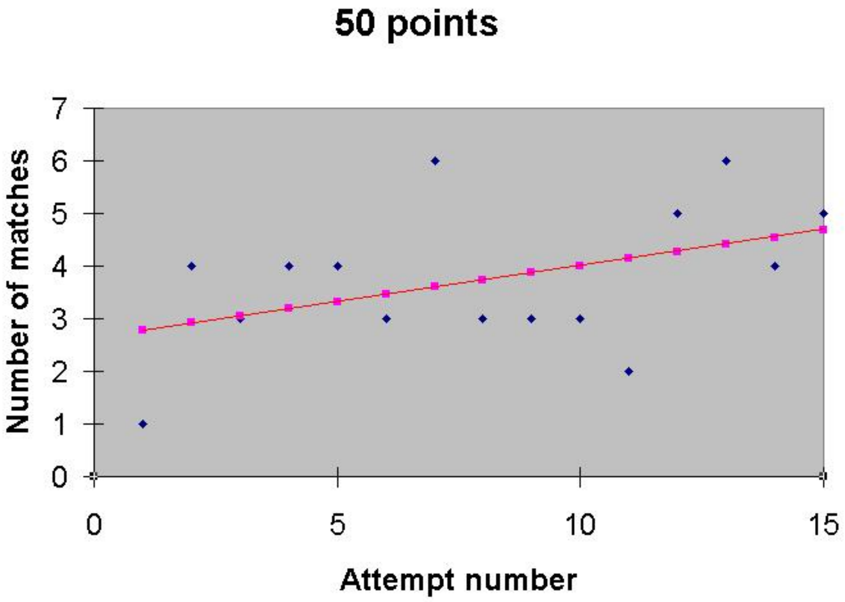

Figure 25: The number of matches when the threshold is 50 points. The θ in this case is 7.73 degrees (17.2% increase), which indicates that most of the 50 points scores came towards the end of the attempts, even though they were more spread than in the case with 40 points.

break into the system. On the other hand, choosing a much higher threshold, e.g. 70 points, would increase the security by a great deal. However, according to our cumulative table, this would lead to an FRR of $\frac{146}{150} = 97.3\%$, which would not been a very including authentication system, seen from the authorized user's point of view. Hence, we chose a sensible trade-off by accepting an FRR of 50%.

The probability of being falsely rejected n times in a row with an FRR of 50% is:

$$P(\text{falsely rejected n times in a row}) = \frac{1}{2^n}, \tag{6.4}$$

which means that the probability of being rejected once is $\frac{1}{2}$, the probability of being rejected twice in a row is $\frac{1}{2^2} = \frac{1}{4}$, while the probability of being rejected three times in a row is $\frac{1}{2^3} = \frac{1}{8}$. If the authentication process consists of a walk lasting only approximately 5 seconds, then the correct user will probably be accepted after just a few attempts. The expectation value of how many times an authorized person has to walk in order to be accepted is 2. Hopefully, the impersonator will have to walk several times more than that, in order to be accepted. The DET-curve in Figure 27 implies that the attacker has to walk approximately 10 times in order to be accepted, when the FRR is 50%. This is because the FAR is about 10%, which means that the attacker will be falsely accepted in 1 of 10 attempts.

Since we accepted to be falsely rejected about half of the times, we looked at the *median score* in the cumulative table. The reason why we chose to look at the median score, and not e.g. the average score, is because a specifically high or low value in the data set is not affecting median values the way it is affecting average values. As can be seen in *Appendix C*, our median score was 50.73, hence we chose to take a look at how many of the participants who managed to exceed 50 points.

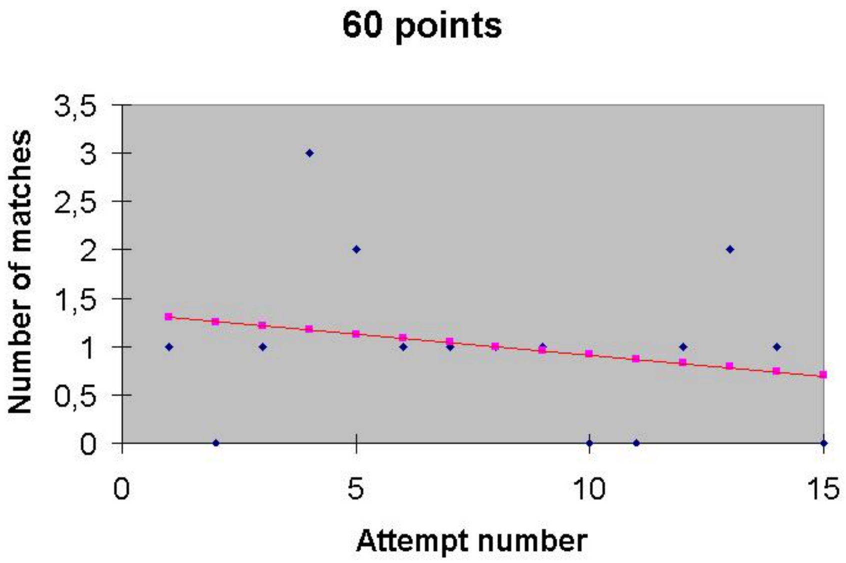

Figure 26: The number of matches when the threshold is 60 points. The θ in this case is -2.45 degrees (5.4% decrease), which indicates that the 60 points scores were rather spread throughout the attempts.

From our results, we could see that some participants actually managed to exceed this threshold of 50 points. This information is shown in Table 6. As much as 53.8% of the participants managed to exceed this threshold once or more during their 15 attempts on Template C. 3 persons exceeded the threshold once, 2 persons exceeded the threshold twice, 1 person exceeded it 3 times, and 1 managed to exceed it as much as 9 times.

We checked the 50 points frequency also in the other 4 templates. It turned out that the results varied from template to template. The results are also shown in Table 6. An interesting observation is that nobody succeeded in passing the 50 points threshold more than once in Template A. This is rather strange, since A was supposed to be the most trivial template to mimic. A conclusion might be that there is actually harder to learn a trivial gait than a more difficult one. Another possible conclusion is that the participants needed to walk for a while, before at least some of them managed to learn how to follow the template graphs. If the latter suggestion is correct, then it does seem like the program had a learning effect after all, even if the number of ”never-passing participants” is somewhat higher than the number of ”sometimes-passing participants”. It should be considered rather risky that somewhere between 20% and 60% of the participants manage to learn to mimic another person’s gait that well within only 15 attempts.

We also wanted to see at which attempt the participants managed to exceed this threshold in each of the templates. The results are shown in Table 7. In Template C, which consists of the author’s natural gait, we can see that the average participant managed to exceed the 50 points threshold after 7.71 attempts. Since the author gets 50 points or more half of the times he is walking, he needs on average 2 attempts in order to reach this threshold, as mentioned previously. As a result of this, a solution might be to set a limit of 3 or 4 attempts in an authentication system like this, and to e.g. request a password in order to get more attempts if the user is not accepted after 3 or 4 attempts.

Template	Results
A	Never: 9/13 = 69.2% Once: 4/13 = 30.8%
B	Never: 7/13 = 53.8% Once: 2/13 = 15.4% Twice: 2/13 = 15.4% 5 times: 1/13 = 7.7% 6 times: 1/13 = 7.7%
C	Never: 6/13 = 46.2% Once: 3/13 = 23.1% Twice: 2/13 = 15.4% 3 times: 1/13 = 7.7% 9 times: 1/13 = 7.7%
D	Never: 8/13 = 61.5% Once: 4/13 = 30.8% 3 times: 1/13 = 7.7%
E	Never: 10/13 = 76.9% Twice: 2/13 = 15.4% 5 times: 1/13 = 7.7%

Table 6: How many times the participants managed to exceed 50 points.

This would increase the security, since it is very unlikely that the attacker will manage to exceed this threshold within 3 or 4 attempts, or that he will manage to type the correct password after 3 or 4 not succeeding attempts.

As can be seen in Table 6, one of the participants managed to pass the 50 points threshold as much as 9 times in Template C. This is actually the *same person* who passed Template B 6 times and Template D 3 times. We therefore suspect this person to be a so-called ”*wolf*”, i.e. a person who is generally good at impersonating other people [54]. As a result of this, the average number of attempts the participants used in order to exceed 50 points in Template C was calculated, *without* taking this person's results into account. The average number of attempts turned out to be 8.67. We therefore conclude that if one has several attempts to practice, then it doesn't seem too difficult to learn to walk like another person in such a way that one exceeds a threshold of 50 points. It also seems like people are succeeding this differently, i.e. some people are learning to impersonate rather quickly, while others are not. Those who are really good, manage to learn this within a few attempts. That is, if they have the template graphs to look at while they are walking.

6.2.4 DET-curve

Figure 27 shows a DET-curve based on the results from the experiment, and the results from the author's own gait data. As we can see, the EER is 26%, and this indicates a rather low performance, compared to some of the other algorithms that have been used in gait authentication, e.g. the algorithm used by Davrondzhon Gafurov, Einar Snekkenes, and Tor Erik Buvarp [5]. Their EER was 16%. To have an EER equal to 26% means that an authorized user will be falsely rejected approximately 1 of 4 times he tries to authenticate himself, and an attacker will be falsely accepted approximately 1 of 4 times he tries to be authenticated. In order to make the system as secure as possible, a threshold which leads to a FAR below 26% should be chosen, however, by doing this, the FRR will increase a

Participant	Template A	Template B	Template C	Template D	Template E
1			8	15	5
2			12		
3		4			
4	15				
5	9	4	5	7	
6		3	2	11	9
7	7	2	15		
8			3		
9	12				
10					1
11		14		2	
13		2	9	4	
Avg	10.75	4.83	7.71	7.8	5

Table 7: At which attempt the participants exceeded 50 points.

great deal. To accept a FRR of 50%, as mentioned previously, will according to this DET-curve lead to a FAR of approximately 10%, and the system will as such become more secure, even if it is still far from perfect.

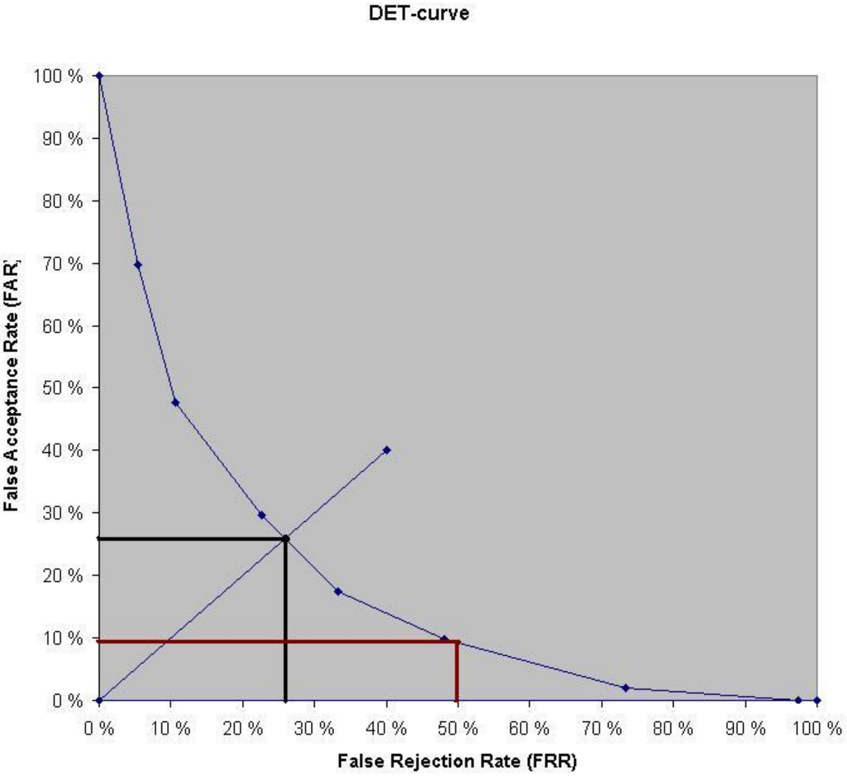

Figure 27: A DET-curve based on Template C, with an EER of 26%. If we accept an FRR equal to 50%, then the FAR will be approximately 10%.

7 Other ways to solve the task

In this chapter, we will take a look at other possible ways to find a solution to the question of how easy or difficult it is to learn to mimic another person's gait. We divide this chapter into a section about what could have been done differently in the graph analysis, and a section about how the experiments could have been done differently.

7.1 Graph analysis and visualization

7.1.1 Shifting of graphs

In order to measure the distance between the template graphs and the user's graphs, we calculated the correlation coefficient between the two r-graphs. By multiplying this coefficient's absolute value with 100, we generated a score between 0 and 100, with 0 as the worst score and 100 as the best score. If the participants had not seen the template graphs as they were walking, they would have got low scores almost all the time, even if their r-graph had been identical to the template r-graph. This is because we did not shift the r-graphs before we calculated the correlation coefficient. If the graphs had been shifted until the most perfect match was found, i.e. the highest possible correlation coefficient, then this coefficient could have represented the participant's score. With no shifting, a participant who starts walking a bit too early or a bit too late according to the template graph, will get a weak correlation and hence a low score. With shifting graphs, this will be no problem, since the graphs in that case are shifted into the right position, regardless of when the participant started to walk. Figure 28 (from the web site called "Dendrochronology, curve matching and mathematics" [55]), shows an example of this technique. The upper illustration represents the correlation coefficients calculated during the shifting, and the bottom illustration shows the two compared graphs after they have been shifted to the best matching position.

7.1.2 The Cycle Length Method

Another distance measure that can be used in order to calculate the distance between two graphs, is the Cycle Length Method [5]. This method was also shown in Figure 9. First, a resultant graph is created, based on the formula $R_i = \sqrt{X_i^2 + Y_i^2 + Z_i^2}$, as shown in the upper illustration in Figure 9. The next step is to interpolate and smooth the graph, by removing some of the redundant details. When this is done, the cycles in the graph are detected, e.g. by looking for the lowest points in the graph, and mark these. In Figure 9, we can see that the graph between two lower points is repeated in cyclic patterns, hence these repetitions are defined as *cycles*. The next step is to calculate the average cycle. This cycle consists of points that are representing the average of the corresponding points within each cycle. When all these steps are completed, the system compares the two average cycles, i.e. the averaged cycle from the original graph, and the averaged cycle from the impersonator's graph. This is done by calculating the Euclidean distance,

$$\text{dist}(A, B) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}, \tag{7.1}$$

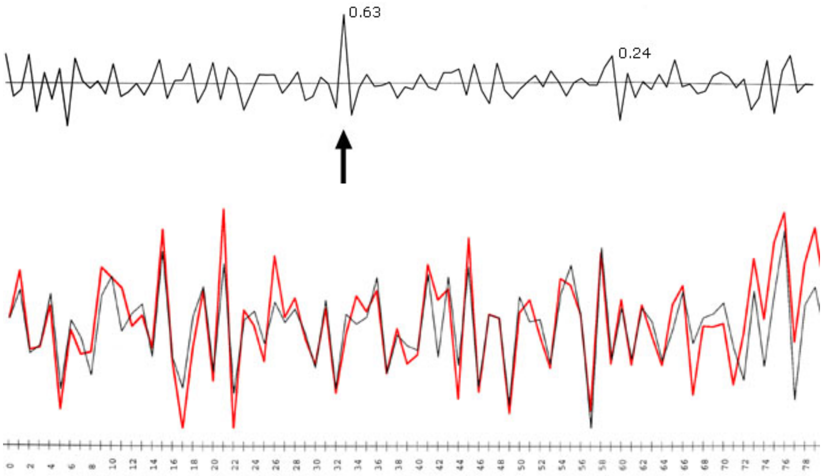

Figure 28: The highest correlation coefficient represents the best match between two graphs. The illustrations are borrowed from the web site called ”Dendrochronology, curve matching and mathematics” [55].

where $A = (a_1, ..., a_n)$ and $B = (b_1, ..., b_n)$ are two averaged cycles, a_i and b_i are the resultant acceleration values at time point i, and n is the number of points within each cycle [5].

7.2 Experiment

In our experiment, we created 5 different templates which 13 persons attempted to mimic 15 times. There are other ways to detect how easy or difficult it is to mimic other people's gait as well. One way is to let different persons create their own templates, and then see whether some of the participants have a gait which is easier to mimic than others. As mentioned previously, within biometrics, some people are referred to as ”sheeps”, and others as ”wolves”. A sheep is easy to impersonate, and a wolf is a good impersonator. Since only one person created the templates in our experiment, then we cannot say for sure whether his gait is an easy or a difficult one to mimic, compared to other people's gaits. Hence, it is difficult to say with complete certainty whether it is easy to walk like another person, by just trying to imitate one person's gait. If instead 15 persons were creating their own templates, and everybody tried to mimic each of the other 14 person's templates, then it would have been easier to conclude with a higher degree of certainty that it *is* easy or difficult to mimic other people's gait.

A common feature with almost all of the participants was that they got worser results at the end of their 15 attempts. The reason to this might be that 15 gait sessions were a bit too many at once. It might have been better to walk e.g. 5 to 10 times, and then continued some hours later, or even a day or two later. If we had done this, it would probably also been possible to ask the participants to walk more than 15 times on each template. This would of course taken a longer time than the experiment we arranged, but the results would probably been much clearer, and a lot easier for us to draw a more solid conclusion upon.

During the experiment, we observed that a good score often was followed by a bad score. A hypothesis is hence that to know that one has scored high, increases the anxiety of trying to repeat the success, and this anxiety leads in many cases to a overstrained walking style, which is a lot different from the succeeding gait. This will result in a bad score. Instead of displaying the scores. we could have shown just whether the user was accepted or rejected, or maybe whether he was *"very bad"*, *"bad"*, *"ok"*, *"good"*, or *"very good"*. This would possibly have helped the user to loosen his pressure, and it could also have lead to more informative results.

8 Further work

The problem we are raising in this thesis is wide, and hence there are still more aspects to look at in the future. Some of them will be mentioned in this chapter.

8.1 A bigger experiment with more templates

First of all, a bigger experiment will increase or decrease the probability that the conclusions that we have drawn in this thesis are correct. As mentioned in the previous chapter, an experiment where more templates are involved, such that people can try to imitate different gaits from different people, will give us better and more reliable results, than an experiment containing templates from only one person. When a person is creating a template, he should walk maybe 5 to 10 times, and the template should be based upon the average result of these walks. The reason why he should do this, is because random movements will occur almost every time a person walks, which are not general to his gait. Using the average of different gait samples, will therefore be more representable to his natural gait.

8.2 Comparison of different distance metrics

In this experiment, we used only ”Pearsons's correlation coefficient” in order to measure the distance between the two r-graphs, i.e. the template r-graph and the participant's r-graph. Our conclusions can therefore only be drawn on basis of that distance metric. In the further work, different distance metrics other than Pearson's correlation coefficient should be tested, in order to see whether our conclusions are valid also with other metrics. An example of another distance metric was mentioned in chapter 7.1.2., i.e. the ”Cycle Length Method”.

8.3 Involving a camera in the experiment

In order for the user to understand the graphs better, there might be a solution to involve a camera in the experiment. This camera can catch a series of images, and show these in the same window as the graph. The learning effect will probably be better if the user gets more information from the prototype about how well his gait matches the template.

8.4 Improved visual interactive feedback

All of the participants were asked how well they felt the program was helping them in learning another way of walking. Most of them answered that the graphs were easy to follow, and that it was not too difficult to adjust their gait to the displayed graphs. However, a few participants felt that it was difficult to follow the graphs, and that the cycles were not that easy to detect. In order to improve the visual feedback, the graphs should be labeled better than they are now. This can be done e.g. by inserting marks in the graph where the user has taken a step. It would also be very practical if the prototype could interpret the user's gait in more detail, and give feedback to the user in forms of simple sentences, e.g. ”*Steps too long*” or ”*Walking too fast*”, maybe with arrows pointing to where the errors were found.

8.5 Sound based feedback

It would be an interesting feature if the prototype was able to give the user feedback in form of sound. The system can generate a specific pitch as a result of a specific gait, and then the impersonator could try to walk such that the system generates a pitch as equal to the original as possible. Of course, this presupposes that the impersonator is not tone-deaf, but it is at least one way to develop this prototype further. However, most people have at least the ability to hear whether a specific tone is light or dark, so another way to do this is for the system to generate a tone which tells the impersonator whether he is close or far from the template, e.g. by returning a light tone when the user is close to the template, and a dark tone when he is far from the template.

8.6 Difference between different groups

Another question that could be interesting to look at, is whether some groups, e.g. specific genders, age groups, professions, etc, are better impersonators than other groups. Does a dancer or an actor have the ability to learn other people's walking rhythm easier than e.g. an IT-consultant? Are men generally better impersonators than women? Questions like these could also be interesting when studying gait mimicking.

8.7 The issue of wolves and sheep

The fact that some people are very easy to imitate, and other people have a talent of imitating other people, is something that will be important to investigate in further studies of gait mimicking. One should try to find approximately how large percentage of the population that can be classified as sheeps, and how large percentage that can be classified as wolves. This experiment would of course require a lot more participants than we had in our experiment. The future of gait authentication could in worst case depend on conclusions from an experiment like this.

9 Conclusions: Is it possible to learn to walk like someone else?

It actually seems rather easy to learn to walk like someone else.

We calculated the linear regression in order to see how much the scores had increased from the first to the 15th attempt on each of the 5 templates. To visualize how much the scores had increased, we calculated the angle of every regression line. The results showed positive angles in all of the templates, even though not all of the angles were that large. The smallest angle was 5.14 degrees (11.4% increase), and the largest angle was 27.87 degrees (61.9% increase). With two of the templates as exceptions, i.e. Template A and D, the increase was larger when only the first 10 attempts were taken into consideration. The smallest angle was then 3.22 degrees (7.2% increase), while the largest angle was 30.91 degrees (68.7% increase). A reason to this may be that the participants got tired from walking after approximately 10 attempts, and as such got lower scores in the last attempts, which affected the results.

Since Template C represented the author's natural gait, it was interesting to see how high scores he would get when walking like himself. He walked 150 times, and the results were stored in a table. The median value turned out to be 50.73, and hence we chose a threshold of 50 points in order to see how many participants who managed to exceed the author's median value. Among the 13 participants, 30.8% of these succeeded in exceeding 50 points once or more in Template A, 46.2% succeded this once or more in Template B, 53.8% succeeded once or more in Template C, 38.5% succeeded once or more in Template D, while 23.1% succeeded in exceeding the author's median value once or more in Template E. Even though most of the percentages mentioned above are below 50%, we must remember that these are the results after only 15 attempts on each template. If more than 20% manage to exceed the author's median value within 15 attempts, then we can only imagine how large percentage that would manage this after e.g. 100 attempts.

Since we had rather few participants in our experiment, it is difficult to say for sure that it *is* easy to learn to mimic another person's gait. We can't say whether our natural template is representable to a ”standard gait”, and we can't know with 100% certainty that the group of 13 participants were representable to the population as a whole. Another fact that we have to take into account, is that we have not used more than one distance metric, and as a result of this, we cannot say for sure whether our conclusion is valid for all distance metrics. We can only say that it is valid for the one we chose. We created a DET-curve based on the author's own gait data and the data from the participants, and this showed an EER of 26%. This indicates that the performance of our algorithm is not too good, since other algorithms have resulted in lower EERs. However, our results can give us a *hint* about whether it is easy to learn to mimic other people's gait, and this can be pursued into a larger experiment, where more templates are created, hopefully representing natural gaits from different people, where other distance metrics are tested, and where more participants will contribute to the study.

If our conclusion turns out to be correct, i.e. if it *is* easy to learn to mimic other people's gait, then it is too risky to implement gait authentication as the only authentication technique. A solution could be to implement it together with another technique, e.g. with voice or fingerprints. The advantage of using multi-factor authentication is shown very clearly in cases like this. There might be a certain risk of managing to mimic another person's gait. However, it is very unlikely for an impersonator to manage both this, *and* at the same time, manage to have his fingerprints or his voice accepted.

Bibliography

[1] Woodward, J. D., Orlans, N. M., & Higgins, P. T. 2003. *Biometrics - Identity Assurance in the Information Age*. McGraw-Hill/Osborne.

[2] Cameron, D. 2002. Walk this way. *Technology Review*.

[3] Buvarp, T. E. Hip movement based authentication. how will imitation affect the results? Master's thesis, Gjøvik University College, 2006.

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

[5] Gafurov, D., Snekkenes, E., & Buvarp, T. E. 2006. Robustness of biometric gait authentication against impersonation attack. *Springer-Verlag Berlin Heidelberg*, 479–488.

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

[7] Jain, A., Bolle, R., & Pankanti, S. 1999. *Biometrics - Personal Identification in Networked Society*. Kluwer Academic Publishers.

[8] Ross, A., Jain, A. K., & Qian, J.-Z. 2001. Information fusion in biometrics. *Springer-Verlag Berlin Heidelberg*, 354–359.

[9] Brand, J. D., Mason, J. S. D., & Colomb, S. 2001. Visual speech: A physiological or behavioural biometric? *Lecture Notes in Computer Science*, 2091.

[10] http://www.sic.rma.ac.be/~beumier/3D/3d.html. Facial surface identification. Last visited 2/5-2007.

[11] Chen, X., Tian, J., Su, Q., Yang, X., & Wang, F. 2005. A secured mobile phone based on embedded fingerprint recognition systems. *Springer-Verlag Berlin Heidelberg*, 549–553.

[12] Clarke, N. L. & Furnell, S. M. 2006. Authenticating mobile phone users using keystroke analysis. *Springer-Verlag*.

[13] Trilok, N. P., Cha, S.-H., & Tappert, C. C. 2004. Establishing the uniqueness of the human voice for security applications. *Proceedings of Student/Faculty Research Day, CSIS, Pace University, USA*.

[14] Braghin, C. 2000. Biometric authentication. *University of Helsinki, Department of Computer Science*.

[15] Murray, M., Drought, A., & Kory, R. 1964. *Walking patterns of normal men*. Journal of Bone and Joint Surgery.

[16] Goswami, A. 1998. A new gait parameterization technique by means of cyclogram moments: Application to human slope walking.

[17] Urtasun, R. & Fua, P. 2004. 3d tracking for gait characterization and recognition. *IEEE Computer Society*.

[18] Yam, C. Y., Nixon, M. S., & Carter, J. N. 2002. Performance analysis on new biometric gait motion model. *IEEE Computer Society*.

[19] Zhao, G., Liu, G., Li, H., & Pietikäinen, M. 2006. 3d gait recognition using multiple cameras. *IEEE Computer Society*.

[20] Iso, T. & Yamazaki, K. 2006. Gait analyzer based on a cell phone with a single three-axis accelerometer. *ACM*.

[21] Mäntyjärvi, J., Lindholm, M., Vildjiounaite, E., & Mäkelä, S.-M. 2005. Identifying users of portable devices from gait pattern with accelerometers. *IEEE*, 2, 973–976.

[22] Sagawa, K., Ishihara, T., Ina, A., & Inooka, H. 1998. Classification of human moving patterns using air pressure and acceleration. *IEEE*.

[23] Middleton, L., Buss, A. A., Bazin, A., & Nixon, M. S. 2005. A floor sensor system for gait recognition. *IEEE*, 171–176.

[24] Orr, R. J. & Abowd, G. D. 2000. The smart floor: A mechanism for natural user identification and tracking.

[25] Morris, S. J. & Paradiso, J. A. 2002. A compact wearable sensor package for clinical gait monitoring. *Offspring*, Vol 1.

[26] Yanushkevich, S. N., Stoica, A., Srihari, S. N., Shmerko, V. P., & Gavrilova, M. L. 2004. Simulation of biometric information: The new generation of biometric systems. *Proc. BT2004 Int'l Workshop on Biometric Technologies, Calgary, AB, Canada*, 87–98.

[27] Matsumoto, T., Matsumoto, H., Yamada, K., & Hoshino, S. 2002. Impact of artificial "gummy" fingers on fingerprint systems. *Proceedings of SPIE*, 4677, 275–289.

[28] Jain, A., Hong, L., & Pankanti, S. 2000. Biometrics: Promising frontiers for emerging identification market. *Communications of ACM*, 91–98.

[29] Pellom, B. L. & Hansen, J. H. 1999. An experimental study of speaker verification sensitivity to computer voice-altered imposters. *IEEE*, 837–840.

[30] Blanz, V. & Vetter, T. 1999. A morphable model for the synthesis of 3d faces. 187–194.

[31] Fua, P. & Miccio, C. 1999. Animated heads from ordinary images: A least squares approach. *Computer Vision and Image Understanding: CVIU*, 75, 247–259.

[32] Cui, J., Wang, Y., Huang, J., Tan, T., & Sun, Z. 2004. An iris image synthesis method based on pca and super-resolution. *IEEE*.

[33] Smith, L. I. 2002. A tutorial on principal components analysis (http://csnet.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf). Last visited 10/6-2007.

[34] Freeman, W. T., Jones, T. R., & Pasztor, E. C. 2002. Example-based super-resolution. *IEEE*.

[35] Maltoni, D., Maio, D., Jain, A., & Prabhakar, S. 2003. *Handbook of Fingerprint Recognition*. Springer.

[36] Bigun, J., Fierrez-Aguilar, J., Ortega-Garcia, J., & Gonzalez-Rodriguez, J. 2003. Multimodal biometric authentication using quality signals in mobile communications. *Proceedings of the 12th International Conference on Image Analysis and Processing (ICIAP 03)*.

[37] http://www.linguistics.ucla.edu/faciliti/facilities/statistics/dprime.htm. D-prime (signal detection) analysis. Last visited 23/4-2007.

[38] Løvås, G. G. 1999. *Statistikk - for universiteter og høgskoler*. Universitetsforlaget AS.

[39] http://www.childrens mercy.org/stats/definitions/pvalue.htm. P-value. Last visited 23/4-2007.

[40] Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin*, Vol. 1, 80–83.

[41] Kotropoulos, C., Tefas, A., & Pitas, I. 2000. Frontal face authentication using morphological elastic graph matching. *Image Processing, IEEE Transactions on*, 9, 555–560.

[42] Yeung, D.-Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., & Rigoll, G. 2004. Svc2004: First international signature verification competition. *ICBA 2004*, 16–22.

[43] Bae, K., Noh, S., & Kim, J. 2003. Iris feature extraction using independent component analysis. *AVBPA 2003*, 838–844.

[44] Kukula, E. & Elliott, S. 2005. Implementation of hand geometry at purdue university's recreational center: an analysis of user perspectives and system performance. *Security Technology, 2005. CCST '05*, 83–88.

[45] Hocquet, S., Ramel, J.-Y., & Cardot, H. 2005. Fusion of methods for keystroke dynamic authentication. *Automatic Identification Advanced Technologies, 2005*, 224–229.

[46] http://www.freescale.com/. Freescale (webpage). Last visited 24/4-2007.

[47] *Wireless Sensing Triple Axis Reference Design - Designer Reference Manual*, 2006.

[48] http://www.spss.com/spss/. Spss for windows (webpage). Last visited 24/4-2007.

[49] http://www.codeproject.com/miscctrl/graph2d.asp. A 2d data visualisation class by paul barvinko. Last visited 2/5-2007.

[50] Altman, D. G. 1991. *Practical statistics for medical research*. Chapman & Hall/Crc.

[51] Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences (2nd edition)*. Lawrence Erlbaum Assoc Inc.

[52] http://www.stat.yale.edu/Courses/1997 98/101/linreg.htm. Linear regression. Last visited 3/5-2007.

[53] http://www.staff.vu.edu.au/mcaonline/units/trig/ratios.html. Trigonometry - sin, cosine, tan. Last visited 3/5-2007.

[54] Pankanti, S., Ratha, N., & Bolle, R. 2002. Structures in errors: A case study in fingerprint verification. *IEEE*.

[55] http://www.cybis.se/forfun/dendro/math/index.htm. Dendrochronology, curve matching and mathematics. Last visited 7/5-2007.

A The results from the experiment

Participant no.	θ and % increase (Attempt no. 1-15)	θ and % increase (Attempt no. 1-10)
1	18.11 (40.2%)	56.16 (124.8%)
2	-4.27 (-9.5%)	16.55 (36.8%)
3	41.50 (92.2%)	25.67 (57.0%)
4	64.28 (142.8%)	26.60 (59.1%)
5	42.95 (95.4%)	68.60 (152.4%)
6	56.97 (126.6%)	49.09 (109.1%)
7	39.75 (88.3%)	75.63 (168.1%)
8	-26.18 (-58.2%)	30.39 (67.5%)
9	54.53 (121.2%)	20.97 (46.6%)
10	20.01 (44.5%)	-45.65 (-101.4%)
11	-0.88 (-2.0%)	-7.14 (-15.9%)
12	47.61 (105.8%)	54.35 (120.8%)
13	7.87 (17.5%)	-48.93 (-108.7%)
Average	27.87 (61.9%)	24.79 (55.1%)

Table 8: The results from Template A. The green numbers represent angles equal to or larger than 25 degrees (55.6% increase), the yellow numbers represent positive angles which are smaller than 25 degrees (55.6% increase), and the red numbers represent angles which are either negative or 0 degrees.

Participant no.	θ and % increase (Attempt no. 1-15)	θ and % increase (Attempt no. 1-10)
1	35.42 (78.7%)	69.06 (153.5%)
2	-4.86 (-10.8%)	9.92 (22.0%)
3	39.57 (87.9%)	57.60 (128.0%)
4	-52.04 (-115.6%)	-51.54 (-114.5%)
5	9.65 (21.4%)	64.58 (143.5%)
6	-45.84 (-101.9%)	7.15 (15.9%)
7	13.16 (29.2%))	12.24 (27.2%)
8	63.91 (142.0%)	63.68 (141.5%)
9	17.24 (38.3%)	-54.39 (-120.9%)
10	30.70 (68.2%)	-23.28 (-51.7%)
11	-20.77 (-46.2%)	-19.20 (-42.7%)
12	53.75 (119.5%)	49.42 (109.8%)
13	-36.97 (-82.2%)	-33.71 (-74.9%)
Average	7.92 (17.6%)	11.66 (25.9%)

Table 9: The results from Template B. The green numbers represent angles equal to or larger than 25 degrees (55.6% increase), the yellow numbers represent positive angles which are smaller than 25 degrees (55.6% increase), and the red numbers represent angles which are either negative or 0 degrees.

Participant no.	θ and % increase (Attempt no. 1-15)	θ and % increase (Attempt no. 1-10)
1	54.08 (120.2%)	69.25 (153.9%)
2	54.27 (120.6%)	-28.95 (-64.3%)
3	47.65 (105.9%)	54.23 (120.5%)
4	16.00 (35.6%)	59.29 (131.8%)
5	55.27 (122.8%)	65.28 (145.1%)
6	-6.64 (-14.8%)	67.88 (150.8%)
7	61.61 (136.9%)	39.76 (88.4%)
8	-63.55 (-141.2%)	-74.77 (-166.2%)
9	-7.63 (-17.0%)	-1.38 (-3.1%)
10	-9.44 (-21.0%)	-27.86 (-61.9%)
11	-9.60 (-21.3%)	1.45 (3.2%)
12	-0.87 (-1.9%)	-3.85 (-8.6%)
13	63.56 (141.2%)	72.78 (161.7%)
Average	19.59 (43.5%)	22.55 (50.1%)

Table 10: The results from Template C. The green numbers represent angles equal to or larger than 25 degrees (55.6% increase), the yellow numbers represent positive angles which are smaller than 25 degrees (55.6% increase), and the red numbers represent angles which are either negative or 0 degrees.

Participant no.	θ and % increase (Attempt no. 1-15)	θ and % increase (Attempt no. 1-10)
1	60.41 (134.3%)	8.88 (19.7%)
2	46.77 (103.9%)	13.98 (31.1%)
3	-13.28 (-29.5%)	31.37 (69.7%)
4	8.17 (18.2%)	-31.85 (-70.8%)
5	-34.81 (-77.4%)	38.23 (84.9%)
6	69.23 (153.8%)	38.59 (85.8%)
7	36.35 (80.8%)	1.60 (3.6%)
8	11.83 (26.3%)	0.52 (1.1%)
9	-33.03 (-73.4%)	-12.65 (-28.1%)
10	-16.67 (-37.1%)	-43.47 (-96.6%)
11	-52.20 (-116.0%)	-71.45 (-158.8%)
12	20.90 (46.4%)	25.26 (56.1%)
13	-36.88 (-82.0%)	42.86 (95.2%)
Average	5.14 (11.4%)	3.22 (7.2%)

Table 11: The results from Template D. The green numbers represent angles equal to or larger than 25 degrees (55.6% increase), the yellow numbers represent positive angles which are smaller than 25 degrees (55.6% increase), and the red numbers represent angles which are either negative or 0 degrees.

Participant no.	θ and % increase (Attempt no. 1-15)	θ and % increase (Attempt no. 1-10)
1	76.81 (170.7%)	80.37 (178.6%)
2	27.57 (61.3%)	38.67 (85.9%)
3	27.81 (61.8%)	17.95 (39.9%)
4	43.99 (97.8%)	62.96 (139.9%)
5	61.01 (135.6%)	69.75 (155.0%)
6	31.62 (70.3%)	70.02 (155.6%)
7	26.66 (59.2%)	71.80 (159.6%)
8	1.84 (4.1%)	56.64 (125.9%)
9	-12.52 (-27.8%)	-5.73 (-12.7%)
10	-59.81 (-132.9%)	-65.63 (-145.8%)
11	17.36 (38.6%)	40.11 (89.1%)
12	21.07 (46.8%)	13.35 (29.7%)
13	-20.32 (-45.2%)	-48.43 (-107.6%)
Average	18.70 (41.6%)	30.91 (68.7%)

Table 12: The results from Template E. The green numbers represent angles equal to or larger than 25 degrees (55.6% increase), the yellow numbers represent positive angles which are smaller than 25 degrees (55.6% increase), and the red numbers represent angles which are either negative or 0 degrees.

B The author’s own natural gait data

Test #	Results				
1	19.30	42.13	39.36	43.96	55.55
	8.02	38.73	38.70	50.88	5.89
	41.73	5.62	51.84	11.08	57.73
2	36.09	58.51	71.94	44.93	30.53
	23.17	12.70	21.62	50.86	34.95
	34.17	23.24	2.41	12.93	6.44
3	7.35	60.09	58.41	55.85	64.69
	32.07	41.35	63.11	69.46	38.04
	69.81	17.68	13.31	59.00	50.73
4	23.86	8.44	21.40	22.03	40.35
	65.38	54.81	46.18	50.71	54.40
	54.18	44.25	40.64	66.44	63.09
5	23.23	22.65	15.42	26.49	25.48
	33.65	73.12	60.40	63.35	59.56
	43.55	66.35	65.10	58.69	62.96
6	25.65	31.33	62.67	60.77	61.95
	59.31	47.62	54.97	60.27	50.29
	68.91	53.52	40.61	62.06	64.19
7	24.08	32.01	44.27	57.96	47.63
	48.17	26.52	71.31	63.10	58.91
	43.17	42.88	51.75	44.69	53.34
8	29.93	25.43	50.78	65.18	58.38
	39.17	72.77	63.31	43.22	60.55
	64.92	31.22	65.86	58.71	60.03
9	12.54	48.02	65.56	57.43	46.69
	38.87	57.79	63.04	53.47	65.35
	40.79	56.85	55.81	54.15	30.66
10	21.81	57.68	29.92	60.67	62.18
	51.89	55.36	53.66	62.33	64.48
	25.88	69.62	59.07	6.50	63.84

Table 13: Results from the author's imitation of himself (Template C).

C Cumulative table containing the author's own natural gait data

0-20%	Score	20-40%	Score	40-60%	Score	60-80%	Score	80-100%	Score
0.7	2.41	20.7	26.49	40.7	43.55	60.7	55.36	90.7	65.18
1.3	5.62	21.3	26.52	41.3	43.96	61.3	55.55	80.7	62.33
2.0	5.89	22.0	29.92	42.0	44.25	62.0	55.81	81.3	62.67
2.7	6.44	22.7	29.93	42.7	44.27	62.7	55.85	82.0	62.96
3.3	6.50	23.3	30.53	43.3	44.69	63.3	56.85	82.7	63.04
4.0	7.35	24.0	30.66	44.0	44.93	64.0	57.43	83.3	63.09
4.7	8.02	24.7	31.22	44.7	46.18	64.7	57.68	84.0	63.10
5.3	8.44	25.3	31.33	45.3	46.69	65.3	57.73	84.7	63.11
6.0	11.08	26.0	32.01	46.0	47.62	66.0	57.79	85.3	63.31
6.7	12.54	26.7	32.07	46.7	47.63	66.7	57.96	86.0	63.35
7.3	12.70	27.3	33.65	47.3	48.02	67.3	58.38	86.7	63.84
8.0	12.93	28.0	34.17	48.0	48.17	68.0	58.41	87.3	64.19
8.7	13.31	28.7	34.95	48.7	50.29	68.7	58.51	88.0	64.48
9.3	15.42	29.3	36.09	49.3	50.71	69.3	58.69	88.7	64.69
10.0	17.68	30.0	38.04	50.0	50.73	70.0	58.71	89.3	64.92
10.7	19.30	30.7	38.70	50.7	50.78	70.7	58.91	90.0	65.10
11.3	21.40	31.3	38.73	51.3	50.86	71.3	59.00	91.3	65.35
12.0	21.62	32.0	38.87	52.0	50.88	72.0	59.07	92.0	65.38
12.7	21.81	32.7	39.17	52.7	51.75	72.7	59.31	92.7	65.56
13.3	22.03	33.3	39.36	53.3	51.84	73.3	59.56	93.3	65.86
14.0	22.65	34.0	40.35	54.0	51.89	74.0	60.03	94.0	66.35
14.7	23.17	34.7	40.61	54.7	53.34	74.7	60.09	94.7	66.44
15.3	23.23	35.3	40.64	55.3	53.47	75.3	60.27	95.3	68.91
16.0	23.24	36.0	40.79	56.0	53.52	76.0	60.40	96.0	69.46
16.7	23.86	36.7	41.35	56.7	53.66	76.7	60.55	96.7	69.62
17.3	24.08	37.3	41.73	57.3	54.15	77.3	60.67	97.3	69.81
18.0	25.43	38.0	42.13	58.0	54.18	78.0	60.77	98.0	71.31
18.7	25.48	38.7	42.88	58.7	54.40	78.7	61.95	98.7	71.94
19.3	25.65	39.3	43.17	59.3	54.81	79.3	62.06	99.3	72.77
20.0	25.88	40.0	43.22	60.0	54.97	80.0	62.18	100.0	73.12

Table 14: A cumulative table based on the results from the author's imitation of himself.

D Dictionary

Table 15: Technical terms used in this report.

TERM	*DEFINITION*
Accelerometer	An equipment measuring the acceleration of an object in different directions.
Accelerometer-based gait analysis	Analysis of someone's gait using an accelerometer.
Authentication	Verifying a claimed identity.
Authentication system	A system that performs a verification of someone's claimed identity.
Authentication technique	A technique verifying a person's claimed identity, e.g. fingerprinting or voice recognition.
Authenticator	A person who is being authenticated.
Behavioural biometrics	Authentication of a person, based on something he does.
Biometrics	A authentication technique that uses a biological feature in order to authenticate a person.
Correlation	A method giving information about how strong the linear connection between two sets of data is.
Correlation coefficient	A value, ρ, between -1 and 1, where a ρ close to -1 or 1 implies that the connection between the two compared sets is strong, and a ρ close to 0 implies that the connection between the two compared sets is weak.
Cycle	A repeating pattern.
DET-curve	A graph plotted in a coordinate system, where FAR and FRR are the two axis. The graph is created by calculating FAR and FRR for different thresholds, starting with a small threshold, which gives a low FRR and a high FAR, and ending with a high threshold, which gives a high FRR and a low FAR.
Equal Error Rate (EER)	The point on the ROC-curve, where the FAR is equal to the FRR.
False Acceptance Rate (FAR)	The number of falsely accepted users, divided by the total number of unauthorized users.
False Rejection Rate (FRR)	The number of falsely rejected users, divided by the total number of authorized users.
Floor-sensor based gait analysis	Analysis of someone's gait using sensors that are installed in the floor.
Gait	A person's way of walking.
Identification	Deciding a person's identity.
Image-based gait analysis	Analysis of someone's gait using images captured by a camera.
Impersonation	Imitating an authorized user in order to be authenticated as that person.

Impersonator	A person who is imitating an authorized user in order to be authenticated as that person.
Linear regression	A modeling of the relationship between two variables by fitting a linear equation, $y = mx + b$, to the observed data.
Median value	The value placed in the middle of a sorted dataset, if the number of values in the dataset is odd. If the number of values in the dataset is even, then it is the average between the two values placed in the middle of the sorted dataset.
Multi-factor authentication	Authentication using more than one authentication technique, e.g. fingerprints and password.
Ofstream	A variable type used in C++, which provides an interface in order to write data to files as output streams.
Open source	Distributed source code, which can be legally altered and changed by programmers.
Physiological biometrics	Authentication of a person, based on a physical feature.
Resultant value	The distance between two points, (x_1, y_1, z_1) and (x_2, y_2, z_2), calculated using the following equation: $r = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$
Sensor	An equipment that (in this case) contains an accelerometer, and is worn by the authenticator.
Sheep	A person that is easy to imitate.
Template	A stored electronic sample of someone's biometrics.
Threshold	A limit describing what degree of correlation there has to be between a template and an authenticator's result, in order to be accepted.
Trigonometry	The study of the relationships between the lines and the angles in a triangle.
Walking cycle	The cycle of repeating patterns that are common in all gaits.
Wolf	A person that has a talent of imitating other people.