h_da
HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES

CASED

Hochschule Darmstadt
– Faculty of Computer Science –

Classification of Acceleration Data for Biometric Gait Recognition on Mobile Devices

A thesis submitted in the fulfillment
of the requirements for the award of the degree
Master of Science (M.Sc.)

by
Holger Brandt

Advisor: Prof. Dr. Christoph Busch
Second Advisor: Prof. Dr. Michael Braun

Issue Date: 04.08.2010
Submission Date: 11.02.2010

Declaration

I hereby declare that I am sole author of this thesis and did not use any further contributions of others than in the referenced bibliography.

Where published or unpublished contributions of others are quoted or analogously used this is clearly indicated.

Any drawings and figures in this thesis are either my original work or referenced in the bibliography.

This thesis has not yet been submitted in equal or similar condition to another examination board.

Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die im Literaturverzeichnis angegebenen Quellen benutzt habe.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder noch nicht veröffentlichten Quellen entnommen sind, sind als solche kenntlich gemacht.

Die Zeichnungen oder Abbildungen in dieser Arbeit sind von mir selbst erstellt worden oder mit einem entsprechenden Quellennachweis versehen.

Diese Arbeit ist in gleicher oder ähnlicher Form noch bei keiner anderen Prüfungsbehörde eingereicht worden.

Darmstadt, February 11, 2011 ______________________

Holger Brandt

Abstract

The process of identifying or verifying oneself at electronic devices is part of our daily life. It requires a PIN to start our smartphone, a password to log into our computer user account, and a TAN code to complete an online banking transaction.

Much has been written on the security of passphrases and consensus is reached on following facts: the majority of users tend to use passwords easy to remember and reuses them multiple times. Especially in the booming field of mobile devices such as smartphones and tablets users tend to not using a password at all for a convenient frequent use. This is clearly an example where usability gets into the way of security. That is noteworthy as these device types are comparable likely to be stolen meanwhile they store more and more sensitive information.

Therefore this thesis proposes an alternative user recognition method for mobile devices based on the gait biometric. The goal is to develop an unobtrusive continuous verification running in the background and allowing the user a convenient, but secure usage. The gait characteristics are captured using the built-in accelerometer that has been increasingly becoming standard equipment of smartphones in recent years. Various features are extracted from the measured accelerations and utilized to train a support vector machine (SVM) for user verification. Among the extracted features are power cepstrum representations like the mel-frequency cepstral coefficients (MFCC) and bark-frequency cepstral coefficients (BFCC) which are commonly used in speech and speaker recognition. To the author's knowledge, MFCC and BFCC have not been used for gait recognition previously.

The proposed method is evaluated using two databases of gait data, with 40 and 48 subjects respectively. The gait data of every subject was collected in two different days enabling a comparison of the recognition performance between using data of the same day for training/testing and testing with data of another day. The best same day results were 0.39% FMR and 0.00% FNMR. The cross day performance was significantly worse with 4.01% FMR and 22.5% FNMR. These outcomes are consistent with published studies on the matter, proving that influence of time is a challenging problem for gait recognition.

Kurzfassung

Die Identifizierung und Verifizierung der eigenen Person gegenüber elektronischen Geräten ist Teil unseres täglichen Lebens. Es braucht eine PIN um das Handy zu starten, ein Passwort um sich am PC anzumelden und eine TAN um eine Onlinebanking Transaktion abzuschließen.

Eine echte Sicherheit durch Passwörter ist aber nur oft nur bedingt gegeben, da viele Benutzer leicht zu merkende Passwörter wählen und diese an mehreren Systemen verwenden. Insbesondere bei den boomenden mobilen Endgeräten wie Smartphones und Tablets geht der Trend dahin, dass Benutzer wenn möglich kein Passwort einsetzen um eine bequeme häufige Benutzung zu ermöglichen. Das ist kritisch, da diese Geräte vergleichsweise oft gestohlen werden und gleichzeitig immer mehr sensitive Informationen enthalten.

Deshalb wird in dieser Arbeit eine auf biometrischer Gangerkennung basierende Benutzerverifizierung für mobile Endgeräte vorgestellt. Das Ziel ist dabei eine unaufdringliche, im Hintergrund ablaufende kontinuierliche Benutzerverifizierung, die eine bequeme und sichere Nutzung ermöglicht. Die Charakteristiken des menschlichen Gangs werden dabei von einem im Gerät eingebauten Beschleunigungssensor erfasst. Verschiedene Merkmale werden aus den gemessenen Beschleunigungen extrahiert und für das Training einer Support Vector Machine (SVM) eingesetzt, mit der die Verifizierung implementiert wird. Unter den genutzten Merkmalen sind die aus der Sprach- und Sprechererkennung bekannten Mel-Frequenz-Cepstrum-Koeffizienten (MFCC) sowie die verwandten Bark-Frequenz-Cepstrum-Koeffizienten (BFCC). Nach Kenntnis des Autors wurden MFCC und BFCC bisher nicht für biometrische Gangerkennung genutzt.

Das vorgestellte Konzept wurde mit zwei Datenbanken mit Gangdaten von 40 bzw. 48 Personen evaluiert. Da die Gangdaten bei jeder Person an zwei Tagen erhoben wurden, konnten Vergleiche zwischen der Klassifizierungsperformanz mit Daten vom gleichen Tag und mit Daten von verschiedenen Tagen angestellt werden. Das beste Ergebniss für Daten eines Tages war 0.39% FMR bei 0.00% FNMR. Die Performanz für verschiedene Tage war mit 4.01% FMR bei 22.5% FNMR deutlich schlechter. Das entspricht dem Ergebnis anderer Studien und beweist, dass der Einfluss von Zeit ein großes Problem für die biometrische Gangerkennung darstellt.

Preface

First of all I would like to thank all the voluntary participants of last year's data collection, especially all summer camp attendees of the lifeguard organization Wasserwacht Haibach who did not mind to walk the same track in the forest over and over again.

Secondly I would like to thank my advisor Prof. Dr. Christoph Busch who raised my interest in biometric research with his lecture at the University of Applied Sciences Darmstadt that lead to discussions on potential topics for my master thesis. I really appreciate that Prof. Busch introduced me to gait recognition and provided the best research conditions for this thesis at the Center of Advanced Security Research Darmstadt (CASED).

Next, I would like to thank Prof Dr. Michael Braun with whom I had several talks on current IT security research topics prior to working on this thesis. I am very thankful that he agreed to become my second advisor.

I have to especially thank Claudia Nickel, currently a PhD candidate at the CASED institute, for her ongoing help and support. I am very grateful for her valuable advice and feedback that helped me to stay focused on the important parts while keeping an eye on the submission date.

Looking back I am thankful for being able to have access to resources of the CASED institute, including the smartphone and software used in the data collection, access to literature, the pleasant distraction through fellow students and, maybe most important, the coffee machine.

Last, but not least I want to thank my parents and family without who I would never have been able to be in the position to write a preface. I am thankful that they enabled me to pursue my interests and that they taught me the importance of education. Thank you for giving me the opportunity to realize my own potential. Your love and support is what brought me this far.

Contents

List of Figures

List of Tables

1. Introduction

"Walks.
The body advances,
while the mind flutters
around it like a bird."
– Jules Renard

1.1. Motivation

At present time the usage of mobile devices heavily influences our daily live and changes our common behavior. With steady growing computing capabilities applications like handheld navigation and mobile internet become available for more and more people and as a consequence are also picked up by people that are not technical enthusiasts per se.

The word *app*, as an abbreviation for application, has entered common vocabulary. The smartphones and tablets are more and more an expression of lifestyle, the individuality of the digital natives expresses through the used set of apps and the activity on social platforms. Analysts predict the global customer base for mobile banking to reach 1.1 billion by the year 2015 [GIA10].

These developments might have significant downsides, though. Phone theft has been a problem in a lot of countries for years. For example in the United Kingdom about 228 mobile phones are reported stolen every hour which lead to a call by the Minister of Crime Prevention to the mobile phone industry to better protect device owners against theft [CJ10].

A few years ago the damage of a stolen phone would have been limited to the value of the actual device and telephone costs generated by the criminal at worst. Nowadays this might be one of the first steps to an extensive identity theft, which has likely much more severe consequences.

A perpetrator having access to private and business emails, contacts and social networks can easily impersonate the victim. Thus crimes which cause financial damage and personal defamation can be committed without much time and effort.

How can these risks be reduced? This question can be answered with well-known technical security measures like pins, passwords and strong encryption. Such safeguards are easy to implement and therefore a standard feature of any phone of this age. Unfortunately such technologies are often not used by the owner of the device even though they might be freely available. The next section will investigate the reasons and propose a potential solution that might improve the situation.

1.2. Background

The human factor is often considered the weakest link in the information security chain. That this also applies to the security of mobile phones can be seen looking at the results of the user survey presented in [BN10]. Only 13% of the 548 participants secure their phones with a PIN or visual code during standby periods. In 74% of the cases of an unsecured phone the reason provided was the better usability through faster access or that no thought was given to this setting. That clearly indicates a lack of security awareness.

In this survey the participants were also asked for their willingness to use biometric authentication methods instead of their current security setting. The majority of users, 54%, would use biometric verification therefore significantly increasing the security on mobile devices. The participants interested in adopting biometric verification were asked which biometric modalities they would use. Here the fingerprint modality is a clear winner (87%), followed by speaker recognition (20%), face recognition (20%) and gait recognition (9%).

There have been few exotic phones with fingerprint scanners that never really entered mass market. The reasons are probably the rather high costs for the extra sensor that is not needed by the average end-user. This implies that fingerprint modality is not likely to increase the information security on mobile phones in the near future.

The other mentioned modalities, namely speaker, face and gait recognition do not have this problem as the most modern phones are capable of realizing a biometric verification system using one or more of these modalities. Imaginable is a face recognition utilizing the built-in phone camera or a voice recognition using the compulsive microphone.

Gait recognition is probably the most uncommon modality in this list, which is likely a cause for its comparable low percentage in the aforementioned study. But it has one unique advantage: it does not require explicit user interaction during the verification process as the phone does it literally on-the-go. This property might be the key for increasing the security on mobile phones. The idea is that the device continuously authenticates the owner when he is on the move and thus more infrequently needs an explicit user authentication which is proven one of the main refutations against security measures.

1.3. Objectives

The goal of this thesis is to develop an approach for implementing a biometric gait recognition that can be used for authentication purposes on mobile devices such as modern smartphones, i.e. the holder of a phone can be verified as the owner. The employment of machine learning techniques serves as a main concept to classify a person either as a valid user or as a perpetrator.

With the future application in mind the feasibility to use the developed concepts on a mobile device is a major constraint to consider. The approach should be developed in a way that the algorithms can be transferred to the end-user devices so that the whole process, from the building of the classifiers to the owner verification, works well with the computing capabilities of current smartphones.

In order to be able to work with realistic gait acceleration characteristics, a data collection using a smartphone shall be conducted. This data is then used to build, test and optimize the biometric verification algorithms. An outcome of the thesis should be the necessary knowledge to build a prototype for training and biometric verification that works entirely on the mobile device.

1.4. Disposition

This section will provide an outline of the structure of this thesis.

The next chapter gives a short introduction into the field of biometrics. Chapter 3 reviews the current research state of the art in the area of biometric gait recognition with a focus on approaches using a wearable sensor such as an accelerometer.

As support vector machines are a cornerstone of the proposed concept, the methods and capabilities of this machine learning classifier will be discussed in chapter 4. After that the process of the conducted data collection is put forward in chapter 5. Included is also an evaluation of the sample quality of the collected data.

The collected gait data is a time series of acceleration measurements. As support vector machines need fixed length feature vectors the time series data need to be converted. This processing and the used feature types are described in chapter 6. Chapter 7 presents how the previously introduced feature types were evaluated and optimized in order to select the feature types that perform best for recognition.

In chapter 8 the selected feature types are used for recognition experiments in different settings. Most notable the intra-day and the inter-day variability of gait are studied. Last a summary of the achieved results is offered in chapter 9. An outlook to the practical application of the developed concept is also included as well as other future work.

2. Biometrics Overview

This chapter shall give an introduction into the field of biometrics for readers that are not yet familiar with this research area and its terminology.

2.1. Introduction

Biometrics is defined as the automated recognition of individuals based on their behavioral and biological characteristics [ISOb].

The automated recognition is typically used to implement a biometric identification or verification system. Note that verification is often also referred to as authentication. Unlike other authentication methods like passwords, smartcards, or tokens biometrics cannot be stolen or forgotten.

All biometric systems can be grouped into two classes:

Physiological Biometrics directly derived from the human body. Examples are fingerprint, face and iris recognition.

Behavioral Characteristics from the human body derived from observation. Examples are speaker, keystroke and gait recognition.

Every biometric system considers one or more biometric modalities which are used for the identification and verification process. Each biometrics example given above is one biometric modality.

2.2. Terminology

There have been efforts of the ISO/IEC[1] standardization committee to create a common biometric vocabulary in order to make it easier to compare different

[1] International Organization for Standardization (ISO), International Electrotechnical Commission (IEC)

research works or different biometric systems with each other. This thesis will follow the terms and definitions published in [ISOb]. There is a website[2] which provides the most important terms and definitions and might be helpful for the readers that do not have access to the standard document.

As they are used throughout the thesis some terms are described here:

Genuine A person who submits a biometric sample to a biometric system in an attempt to be recognized as himself.

Imposter A person who submits a biometric sample to a biometric system in an attempt to be recognized as another enrollee. In general considered an intentional action to gain unauthorized access although an unintentional impersonation is also possible.

In a biometric security system, a person who submits a biometric sample in an attempt (either intentional or unintentional) to gain access to a system using the identity of another enrollee.

2.3. Biometric Workflow

The following section will outline the typical processes in a biometric system. As this is only a summary of the most important aspects, figure 2.1 is provided as a reference for the components of a general biometric system so that the relation of the processes is made more obvious. A closer description can be found in [ISOc].

2.3.1. Enrolment

During enrolment the biometric samples of a person are used to extract the biometric features in order to store them for reference in a database. This reference can be compared to the features extracted from other biometric samples for the purpose of identification or verification.

Note that biometric features are numbers or labels extracted from biometric samples that are reproducible for a given biometric characteristic of a person while these features differ for other people.

[2] http://www.3dface.org/media/vocabulary.html

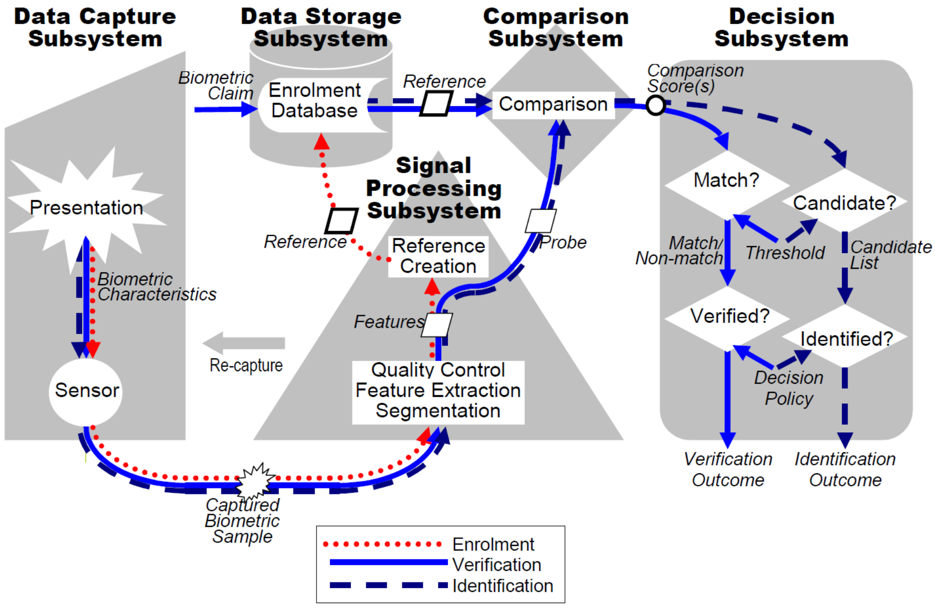

Figure 2.1.: General biometric system, image from [ISOc]

2.3.2. Identification

The goal of identification is to recognize a person's identity. For this purpose one or more biometric samples are recorded to create biometric features that can be used to compare against the features of the biometric reference database. Identification can be seen as a one-to-N comparison.

The output of the comparison is a similarity measurement, the comparison score. The identity of the reference database with the highest comparison score is most likely to be the person to identify. To make the system secure a threshold is used that determines if the comparison score is sufficient for a successful identification. If it wasn't used the system would always accept a person as a certain identity.

2.3.3. Verification

The verification process is similar to the identification. The only difference is that the subject whose biometric sample is compared against the reference database claims a certain identity. Therefore the feature only needs to be compared against the features of the identity from the reference database. Verification can be seen as a one-to-one comparison.

Again the comparison score and a threshold-controlled decision algorithm are used to verify the person's identity.

2.4. Biometric Performance

A variety of metrics exists in the biometric field that measure the performance of a certain biometric system. Again, this is an area where a standardization process was put in place to enable a better benchmarking amongst various systems.

This section will only introduce the most important performance metrics that need to be known in order to follow the discussions in later chapters. The formulas for these metrics and other biometric performance metrics can be found in the appendix B. The definitions are based on the ISO/IEC 19795 standard [ISOa].

2.4.1. False Match Rate

The false match rate (FMR) is the proportion of zero-effort impostor attempt samples falsely declared to match a compared non-self reference.

2.4.2. False Non-Match Rate

The false non-match rate (FNMR) is the proportion of genuine attempt samples falsely declared not to match the reference of the same characteristic from same subject.

2.4.3. Failure to Acquire Rate

The failure to acquire rate (FTA) is the proportion of verification or identification attempts for which the system fails to capture or locate an image or signal of sufficient quality.

2.4.4. False Acceptance Rate

The false acceptance rate (FAR) indicates what proportion of attempts resulted in a false recognition. The difference between FMR and FAR is that in the latter the FTA is also considered. The FMR is an algorithmic level authentication error, whereas the FAR is an authentication error of the whole biometric system.

2.4.5. False Rejection Rate

The false rejection rate (FRR) indicates which percentage of attempts by legitimate users are incorrectly rejected. The difference between FNMR and FRR is that the latter the FTA is also considered. The FNMR is an algorithmic level authentication error, whereas the FRR is an authentication error of the whole biometric system.

2.4.6. Equal Error Rate

The equal error rate (EER) indicates for a biometric system where the FAR and the FRR are equal. This metric is useful to express the accuracy magnitude of a biometric system with one error rate. In general a lower EER is an indication for a more accurate biometric system.

3. Biometric Gait Recognition

The Oxford dictionary defines gait as a person's manner of walking. It is regulated by the mechanical characteristics of the human body as well as conscious control over the limbs. The term gait recognition can be defined as using a person's unique style of walking to identify or verify one's identity [KWM10].

The advantage of using human gait as a biometric modality is that it shows sufficient unique characteristics and can be captured non-invasive and without people's attention. This makes it applicable for periodic identity reverification where unobtrusiveness is an important criterion. In other words a verification process that does not distract and annoy and is convenient in frequent use.

Being still a young biometric modality, gait recognition is currently an active research area. This chapter will give an overview of the topic including the different facets and the state of the art in biometric gait recognition.

3.1. Application Fields

The research efforts related to gait recognition pursuit a wide range of approaches subject to the intended application type. While the technology is rather new there are still not many systems out in the field.

A classical biometric application area is identification and verification for security applications. The use in forensics is a related topic, for example there is a famous case where gait recognition could be used to identify a bank robber from video surveillance footage [LSL07].

Currently researched areas like wearable computing and ubiquitous context aware computing also see a big potential for gait recognition and activity recognition from gait [KSS03, BSR09].

Gait recognition and gait pattern analysis is already used for biomedical purposes. It is considered a practical and inexpensive method for monitoring ambulatory motion [Mat03]. Topics include the monitoring of long-term change in physiological indicators such as locomotion restrictions introduced by chronically diseases and exercise efforts by the patients during rehabilitation [AGL06, BSK07, DJJM+10]. In most medical-related applications the focus lies on detecting the activity of walking in dissociation from other activities such as running, sitting or sleeping [PTK09]. Another health related topic is the fall-detection that is intended for automatic alarms in the case of accidents of unsupervised elderly people [HKJ04, KKW07, ZZW09].

3.2. Technological Approaches

As it can be seen the potential application fields are widespread. From a technological point of view the gait biometric research however can be categorized in three main approaches (analog to Gafurov in [Gaf07b]). The following sections will give a short introduction to the approaches and the specific application fields.

3.2.1. Machine Vision Based Gait Recognition

Machine vision (MV) based gait recognition has been studied for over sixteen years [NA94]. It is also probably the most noted form of gait recognition in research. MV-based gait recognition captures human gait with video-recording from a distance and uses image and video processing techniques to process the image sequences for extraction of features in a way that they can be used for recognition purposes.

The usual application areas are surveillance and forensics which is due to the fact that the biggest advantage of video-based gait biometric is that it can be captured from the distance and without cooperation of the subject. The majority of MV-based approaches use features of the human silhouette that is extracted from the images of a single camera [CGS02, LG02, WTN03, LS04, CRZ06, YHKY08, DASZ09, LCW09]. Other approaches include the usage of multiple cameras for the creation of 3D models [ZLL06] and laser rangefinders used for automatic 3D models creation [YBS09].

Features extracted include body angles of 2D stick figure representation of the

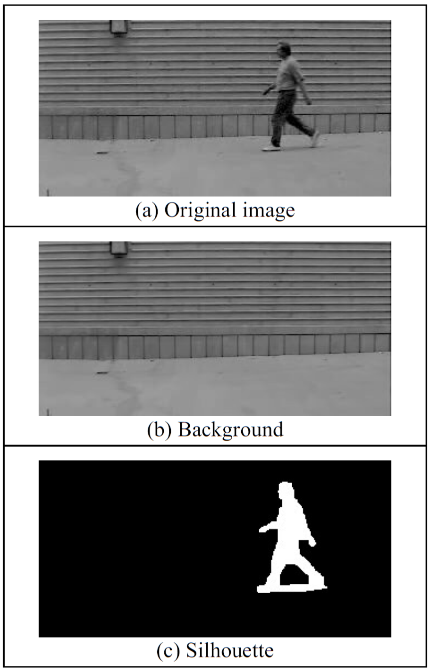

Figure 3.1.: Human silhouette extraction in [CRZ06]

gait [YN03, YHKY08], stride and cadence [BCD02], static body measures like body height and distances between body parts [JB01].

Classification methods used include the nearest-neighbor approach and the K-nearest neighbor classifier [CGS02, LG02, WTN03, CRZ06, LCW09], neural networks [YHKY08] and support vector machines [DASZ09].

Study	#S	TPR
[WTN03]	20	82.50%
[ZLL06]	10	70%
[YHKY08]	30	83.3-90%
[DASZ09]	124	84.4-96.2%
[LCW09]	14	92%
[YBS09]	6	98.96-100%

Table 3.1.: Summary of current MV-based gait recognition studies

Table 3.1 gives an impression of the biometric performance that has been achieved with MV-based approaches. The numbers in the column #S represent the number of subjects in the respective study. TPR is the true positive rate which is the overall correct classification rate [ISOa].

Note that the performance results are not directly comparable as there are significant differences in the design of the conducted experiments. Differences include the usage of different walking speeds, different camera viewpoints and light conditions. Therefore the table shall only give an impression of the current overall state-of-the-art in MV-based gait recognition.

3.2.2. Floor Sensor Based Gait Recognition

Floor sensor (FS) based gait recognition utilizes sensors on or in the floor which measure properties of the occurring footsteps. Such systems can be deployed in access control applications in strategic locations like front doors or foyers. Large scale deployments are imaginable that enable person tracking and continuous identification and verification in buildings with high security requirements.

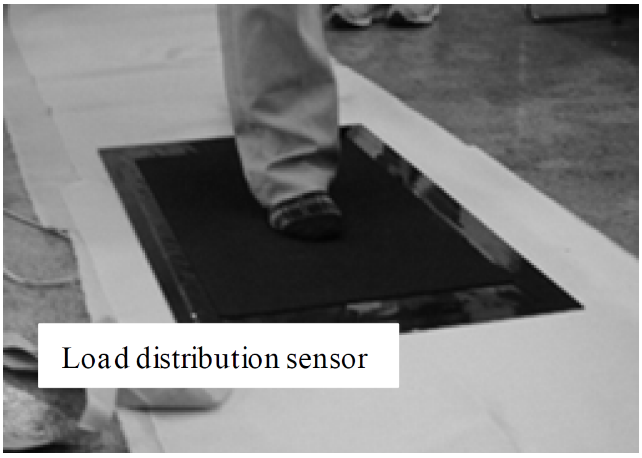

Figure 3.2.: Load distribution sensor used in [TTA+10]

The features used in FS-based gait recognition include the pressure distribution of footsteps [JBL03, TTA+10, QZK10], ground reaction force [OA00, MTG08], length of heel impact [SR04], heel-to-toe-time-ratio [SR04, MBB05], stride cadence [MBB05] and stride length [MBB05, QZK10]. See figure 3.2 for an example of the data acquisition of FS-based gait recognition.

As can be seen in table 3.2 most studies only consider a limited field of subjects. One of the exceptions is [MTG08] where also the influence of different speeds and loading conditions is evaluated. EER is the equal error rate, which refers

Study	#S	EER	TPR
[OA00]	15	—	93%
[JBL03]	8	—	64%
[SR04]	11	—	~70.2%
[MBB05]	15	—	80%
[MTG08]	40	—	76.25% - 98.33%
[QZK10]	11	—	92.3%
[TTA+10]	10	3.8%	—

Table 3.2.: Summary of current FS-based gait recognition studies

to the setting where the false accept rate (FAR) is identical to the false reject rate (FRR). Again it has to be mentioned that the results given are not directly comparable because of the differences within the experiments.

3.2.3. Wearable Sensor Based Gait Recognition

Wearable sensor (WS) based gait recognition is the most recent form of gait recognition as lately proposed as 2005 [ALM+05, MLV+05]. WS-based gait recognition uses a sensor attached to the body. The sensor is typically an accelerometer, which is another word for acceleration sensor.

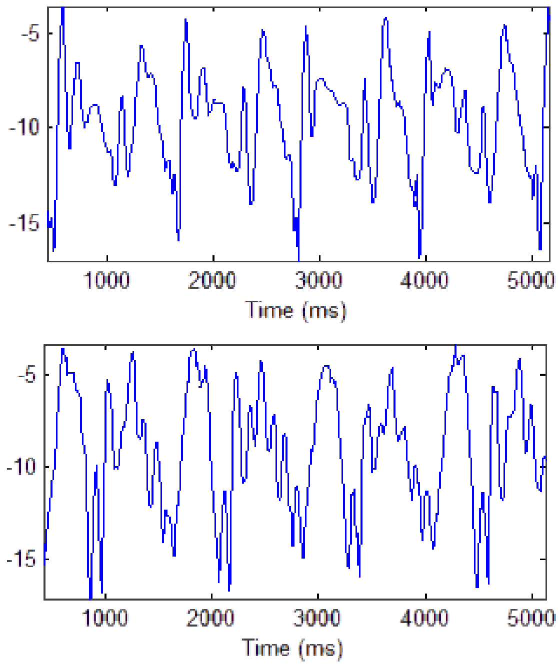

Figure 3.3.: Acceleration in vertical direction of two walking persons

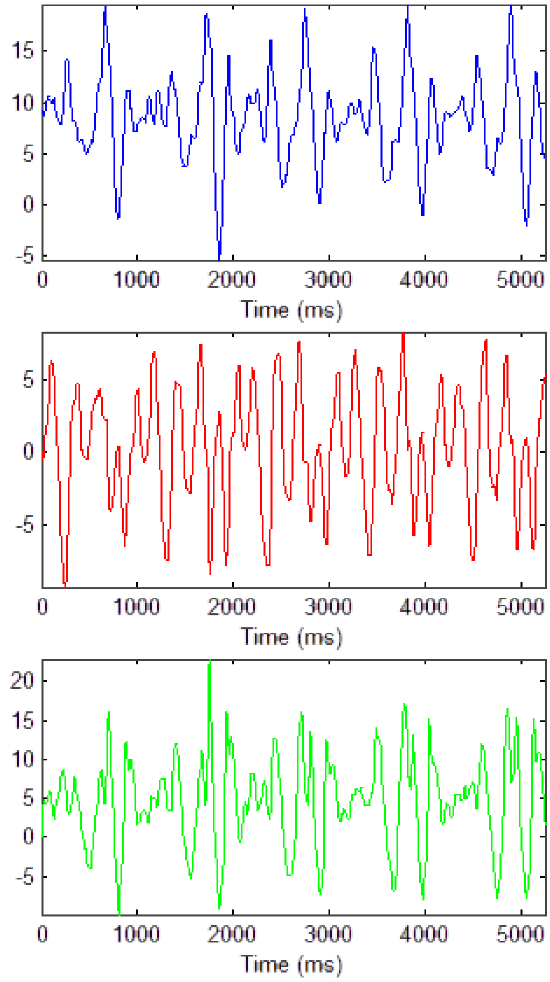

Figure 3.4.: Acceleration in three spatial directions of a walking person: vertical, lateral and anteroposterior (top down)

Accelerometers used for gait recognition generally are tri-axial accelerometers that measure acceleration in three spatial dimensions. This type senses backward-forward (anteroposterior), sideways (lateral) and vertical accelerations. The magnitude of the acceleration signal is biased by gravity. The acceleration that is sensed depends on how the sensor is positioned.

See figure 3.3 for an example of the acceleration captured from two persons walking at normal speed over about five seconds.

Note that the figure shows the acceleration of the axis that is aligned with the gravity, so the mean acceleration is approximately $-10\frac{m}{s^2}$. Gravity is sometimes referred to as static acceleration whereas other accelerations are called dynamic acceleration.

In figure 3.4 on the facing page the acceleration signal in three directions of the walk of one person is compared. As can be seen from the top graph the accelerometer is again positioned aligned to the gravity, but this time facing in the opposite direction so that the signal has a positive offset of the amount of the gravity. The acceleration of the graph in the middle is the acceleration in walk direction, the bottom graph shows the side-ways acceleration.

Figure 5.4 on page 64 shows the device used for capturing the acceleration and the orientation of the accelerometer's built-in coordinate system.

Figure 3.5.: Intelligent shoe proposed in [HCH07]

The majority of WS-based gait recognition studies use one or more tri-axial accelerometers. One of the exceptions is the work described in [HCH07] where intelligent shoes were developed. One of those shoes is shown in figure 3.5. It has a variety of sensors such as force sensitive resistors, bend sensors, a gyroscope and an accelerometer installed. In [Now09] only a gyroscope is used to measure gait characteristics.

As this thesis utilizes an accelerometer this gait recognition approach will be described more elaborate than the other approaches in the next section.

3.3. Accelerometer-based Gait Recognition

The acceleration signals acquired are the result of the acceleration of the person's body due to movement, gravity, external forces like vibration of the accelerometer device and sensor noise.

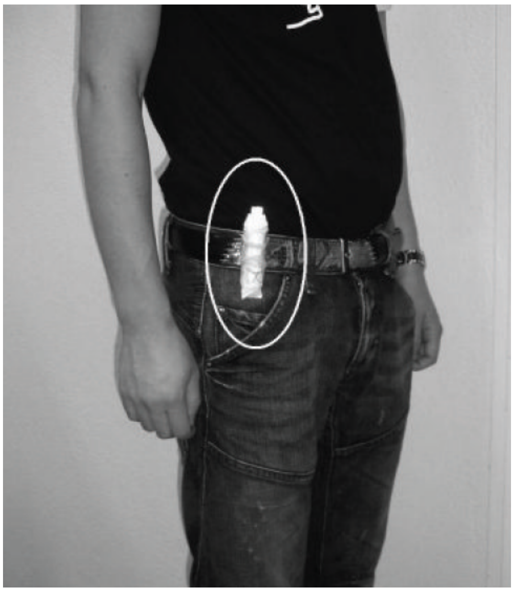

Figure 3.6.: Accelerometer placement on the hip in [GS09]

Accelerometers were worn for gait recognition purposes on various body parts including waist [ALM+05, MLV+05, GSB07, RJM07, GS09, SZ09], ankle [GS09, GSB10, GHS06], arm [GS09], pocket [VML+06, VML+07, GS09, KWM10], hand [VML+06] and shoe [HCY06, HCH07]. Lower leg and foot motion has been recognized as the most discriminative motion during human gait, but acceleration of the hip and torso also seem to provide sufficient discrimination [GS09]. There were also efforts in combining accelerometer measures from multiple body parts [PZW09].

The accelerometers used in the mentioned papers have a wide range of sampling rates. Beginning from 16Hz [GHS06], over 37Hz [SZ09] and 100Hz [GSB10] to 256Hz [ALM+05, MLV+05].

The interesting aspect of WS-based gait recognition is that the accelerometers are more and more part of mobile devices like smartphones and tablets. Because of the ubiquitous nature of mobile handsets and smartphones, with accelerometers an ideal "off-the-shelf" platform for gait analysis and recognition is available. The big advantage is that only new software is needed to be developed with no additional hardware cost and customizations opposed to biometric-only sensors like fingerprint readers.

The fact that accelerometers are already included in a lot of customer devices makes WS-based gait recognition especially useful for an alternative user verification solution. This method was proposed first in [GSB07]. Also, research was already conducted that used end-user equipment like smartphones [SZ09, TAH09, DNB10, KWM10] for data collection.

Table 3.3 presents the used devices. To the authors knowledge there has been no on-device online learning and classification, i.e. training with the device and live verification on the device has not been implemented yet.

Phone	Operating System	Sampling Rate	Used in
Nokia N95	Symbian S60	∼37Hz	[SZ09]
Apple iPhone	iOS	Unknown	[TAH09]
Google G1	Android	∼40Hz	[DNB10, NBR11]
Several Phones	Android	20Hz	[KWM10]

Table 3.3.: Smartphones with accelerometers used in research papers

There are researchers who believe that at some stage there will be intelligent clothes with integrated computing and sensor capabilities. WS-based gait recognition's unique advantage is the unobtrusive verification which is highly desirable in wearable ubiquitous computing [GSB06]. Furthermore it has been shown that the gait contains information like the weight a person carries that can be used for context-aware systems [BSR09].

3.3.1. Feature Extraction

The process of feature extraction is needed to identify valid and useful patterns in the data of biometric gait samples. The problem involved is reducing the data to a manageable level (also known as dimension reduction) while still keeping the import features of the data by only eliminating redundant or irrelevant features. The extracted features are usually utilized for training and testing of machine learning algorithms. Feature extraction is a crucial step as the classification accuracy hugely depends on the recognition and selection of stable features for the intended application.

In the case of gait recognition the challenge for feature extraction is that it has to be robust to small variations in the acceleration patterns and to sampling noise. Gait sample data is a form of time-series data with the typical property of high dimensional data. For this kind of data, feature extraction is a mandatory process as otherwise the data amount is not processable.

There are two common ways to achieve the data reduction in gait analysis: the selection of key features from the time-series data or transformation of the time-series into a smaller data set [BB06].

Before discussing the used feature extraction methods the next section will give preliminary information on the two major approaches for the representation of gait features: cycle-based and non-cycle-based feature extraction.

Cycle-based versus Non-Cycle-based Gait Representation

A gait cycle physically corresponds to two consecutive steps, i.e. the period after one foot touches the ground until the same foot touches the ground again. Note that the end of one gait cycle is the beginning of the following cycle.

Cycle-based features are created by identifying gait cycles in time-series data representing a walking person. Then the feature extraction is conducted on identified cycles and the resulting features are used for biometric template creation and sample comparison. This recognition attempt is based on the assumption that the cycles of one person's gait are similar to each other and the cycles of two different persons are dissimilar.

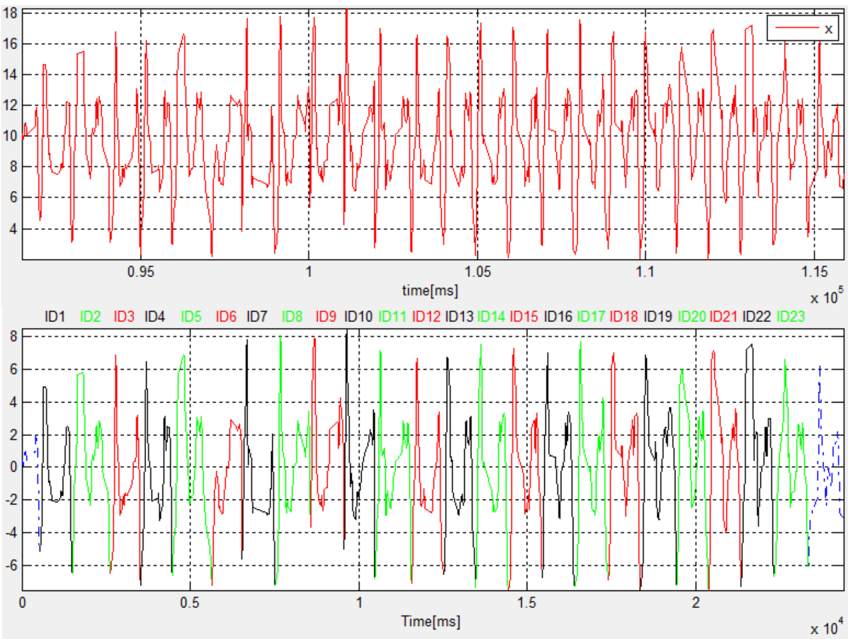

Figure 3.7.: Cycle extraction

Currently this approach for representing gait is the predominantly used method in gait recognition research. Figure 3.7 shows an example of the extraction of cycles from the acceleration signal.

In the second approach, a non-cycle-based gait representation, features are extracted from the times-series data without prior identifying the contained gait cycles. One of the few studies that are based on this approach is [KWM10]. The concept developed for this thesis also follows a non-cycle-based approach, the details will be discussed at a later point in chapter 6.

As mentioned earlier the next two sections will discuss commonly used feature extraction methods used in the gait recognition research.

Key Features Selection Feature Extraction

An example for this approach is the identification of key moments in a gait cycle, e.g. the moment the toe lifts off the ground or foot strikes the ground again. A possible feature is then the mean time between subsequent occurring events of the same type. This feature extraction approach enables one to incorporate high level knowledge from experts for feature selection. In the area of gait recognition this applies to outcomes from biomedical research of human movement [BB06]. The benefits of selecting important features are also the possible downside: the subjective parameter choice might lead to discriminatory features being dropped.

In [KWM10] features derived from the gait acceleration signals include the mean acceleration, mean difference acceleration , the standard deviation and the binned distribution of the acceleration signals.

Data Transformation Feature Extraction

The goal of this approach is to transform time-series data into a set of coefficients that approximately preserve the inherent structure of the data while being lower dimensional than the time-series data. A potential disadvantage of using a data transformation is the reduced interpretability of the transformed parameters [BB06].

One of the most common approaches in this area is the spectral decomposition with a Fourier analysis using a fast Fourier transformation (FFT). This decomposition was already proposed in the beginning of WS-based gait recognition

[ALM+05] and is used also in recent studies [HCH07, RJM07, BSR09].

Another data transformation that also converts a time-series representation into the frequency domain is the wavelet transform which was used for gait recognition purposes in [IY06].

The last transformation that needs to be considered is a combined time-frequency analysis: the principal component analysis (PCA). It has not been largely used in gait recognition, but this is likely to change as it was successfully applied in [BS10].

3.3.2. Biometric Performance

Like the other approaches, WS-based gait recognition enables one to recognize individuals with reasonable performance which confirms the distinctive properties of gait patterns. Although it is the most recent form of gait recognition competitive recognition rates could be achieved. Table 3.4 gives an overview of current WS-based gait recognition studies.

Study	Sensor location	#S	EER	TPR	TG	Sep. by day
[ALM+05]	Waist	36	6.4%	—	5 days	Yes
[RZJ07]	Waist	35	6.7%	—	3 weeks (?)	No (?)
[HCH07]	Shoe	10	—	98%	(?)	(?)
[DBH10]	Hip	60	5.7%	—	0	No
[BS10]	Hip	60	1.6%	—	0	No
[GSB10]	Ankle	30	1.6%	—	0	No

Table 3.4.: Summary of current WS-based gait recognition studies

The column TG provides information on the time gap between the sessions of the gait data collection conducted for the respective study. In the last column information is given whether the reported error rates were obtained while verifying with gait data recorded on a different day.

For example in [ALM+05] the time gap between the two data collection sessions was five days and the gait data of the first day was used for creating the biometric templates whereas the data from the second day was used for conducting the cross validation.

This information is given as the human gait has significant day-to-day variations affecting classification performance substantially. A more detailed discussion regarding the matter will follow in section 3.3.4 on page 43. Unfortunately, not all studies do provide those details.

In [RZJ07] it is not clear which process took the reported 3 weeks and if that means that there were multiple weeks between two or more data collection sessions for one person. This is in [HCH07] even worse, as no information is provided. In the case of [DBH10, BS10, GSB10] information could only be gathered as the authors kindly answered enquiries in this regard. The ideal case, a time gap of multiple day and the separation between reference day data and testing day data, can be found in the pioneer work of WS-based gait recognition [ALM$^+$05].

3.3.3. Other Considerations

In this section a short overview of other WS-based gait recognition work will be given. Although these topics are not directly of relevance for this thesis, a brief description is supplied to complete the current research review.

Attack Scenarios and Testing

In [HCH07] human gait is considered as a unique, dynamic biometrical feature that is complex and difficult to imitate. It is concluded that it is more secure than static features such as password, fingerprints and facial features. Although the conclusion is a rather bold statement which can be argued it is correct with respect to the difficulty to imitate another person's gait which has the potential to be a more secure biometric feature then static features.

But this only applies when the biometric performance is sufficient so that minimal-effort impersonation attacks have no prospects of success. There have been some efforts towards studying the security of the gait biometric, although not nearly as much as with other biometric modalities.

In [GSB07] two types of attack scenarios were considered, the minimal effort impersonation and a closest-person attack in which the imposter knew the person with the most similar gait in relation to the stored templates in the database. Another attack scenario was examined in [Gaf07a], where the imposter knows the gender of a person in the database.

A general result of these studies is that the biometric gait is robust to minimal effort impersonation attacks. However, the gait recognition systems can be vulnerable to imposters with the knowledge that can be used to carry out a more elaborate attack. One of the possible counter-measures is combining the gait modality with other biometric modalities. The efforts in this study area shall be discussed in the following.

Multi-Biometrics

It is a well known fact in biometric research that the usage of multiple biometric modalities and other multibiometric fusion, in short multi-biometrics [ISOd], improves performance in most cases compared to single biometric measure [RJ04].

In [VML+06] gait recognition is utilized as an enhancement to voice recognition that helps to improve the biometric performance in noisy environments. This concept is further developed in [VML+07] where cascading of unobtrusive (voice, gait) and obtrusive (fingerprint) biometrics is proposed. Both studies use the gait modality to offer an alternative verification in cases where a different modality is not available.

The automatic identification of people was done in [TJD09] using preinstalled infrastructure cameras and comparing the visually observed walking patterns with the wireless transmitted signals of an accelerometer with a unique ID carried by each subject.

The work in [DGL+10] proposes a multi-modal biometric authentication on mobile devices using gait signals and fingerprint images yielding improved performance.

Enhancements

Interesting applications related to gait recognition are those that include certain situation awareness. In [TAH09] the GPS sensor of a mobile phone is used to determine whether the device is in a familiar place of the owner to control the threshold for the gait matching score. This means that the threshold is adapted at familiar places so that rather average comparison scores will be accepted whereas in unfamiliar locations it will be changed so that only quite exact scores are accepted.

Study	Study Type	Same day	Cross day	Time gap
[SPL+05]	MV	78%	3%	6 months
[LS04]	MV	> 80%	0%	6 months
[LMS04]	MV	42 - 52%	10 - 11%	6 months
[CKC03]	MV	—	29 - 50%	—
[CGS02]	MV	—	< 50%	—
[TB01]	MV	73%	42%	1 month
[GHS10]	WS	80-90%	26-59%	16 days

Table 3.5.: Time influence on recognition rates, modified from [GHS10]

The adaptation of the verification security is also proposed in [VML+07]: the unobtrusive verification through gait or voice can fail multiple times without the device requiring an explicit user authentication through fingerprint verification. But when a certain time period without successful verification has passed or if a sensible application like mobile banking, is started, an explicit verification is required.

3.3.4. Challenges

Most research papers try to exploit the fact that the human gait is a cyclical pattern. The biggest problems in that approach are all effects that introduce noise in the cycle. This result of changes in walking speed, different ground surfaces and inclines as well as temporary changes in the health condition of the person because of diseases, intoxication, overworking or injuries.

The results are also affected by varying type of clothing and footwear or the transportation of objects by the subjects.

All influences that change human gait really become a challenge when gait samples are compared against gait samples that were created days, months or even years ago. The influence of time on the recognition rates have been studied in a number of MV-based gait recognition studies [TB01, CGS02, CKC03, LS04, LMS04, SPL+05], but only in few WS-based study [BSR09, GHS10].

Table 3.5 is taken and modified from [GHS10] and compares the recognition rates achievable with a same-day recognition with a cross day recognition after the given time gap. By looking at the results it gets apparent that the influence of time is a problem that is unsolved yet.

Interestingly in [HHR07] it is reported that changing surfaces do not have significant impact on the gait recognition performance. But the role of changing surfaces probably needs more investigation as the number of subjects and the tested surface types were quite limited.

That changing walking speeds are not a trivial problem was found in [BSR09] as well as that shoes have a large influence on the gait acceleration profile. This was confirmed in [GS08] and [GSB10] where heavy shoes were found to decrease the discriminative power of gait.

In nearly all reported experiments an ideal walking situation is used for data collection: the subjects walk on an even surface and are fully focused on walking as such. In [BSR09] even a treadmill was utilized for the data collection.

This is of course not a very realistic scenario as people often change their movement while walking (e.g. stop, turn, run, jump and climb stairs) and are distracted by the surrounding and from simultaneously performed activities like typing a text message on the mobile phone. Only few studies have attempted to do gait recognition in a more realistic setting. An example is [KWM10] in which different movements (walk, jog and climb stairs up and down) were performed during data collection.

3.4. Summary

Every biometric modality has constraints in terms of achievable performance and reliability. The provided performance measures of the different gait recognition approaches and the discussed challenges of WS-based gait recognition make it obvious that there is a lot of potential for improvements.

Especially WS-based gait recognition is considered being a biometric field in its infancy [Der10]. An indication for this is also that there is no public database available which would enable a true comparison between different research approaches.

Nevertheless it is likely that biometric gait recognition will enter more practical applications as it offers unique advantages like an unobtrusive sample capturing. The usage of WS-based gait recognition on mobile devices has great potential as the initial situation is normally not nearly as fortunate as with any other biometric modality or system: there is no need to purchase additional sensors as accelerometers are included in nearly all modern phones. Therefore a new biometric system can be introduced by applying a suitable software solution. The additional safeguard could help to mitigate the consequences of a phone theft while not distracting the user in the daily use.

4. Support Vector Machines

4.1. Introduction

The support vector machine (SVM) was introduced by Vladimir Vapnik in 1982 [Vap82] as a supervised learning method based on the theory of structural risk minimization. A SVM is a classifier which is inherently a solution for two class problems. The basic idea of the SVM is to construct a hyperplane as the decision plane, which separates the patterns of the two classes with the largest margin.

SVMs can also be used for tasks than two class separation, for example multi-class classification. This is discussed in more detail in sections 4.3 and 4.4.

4.2. Classification

In this section an introduction will be given in the underlining concepts of the classification with SVMs.

4.2.1. Linear Separable Case

The simplest case possible is a machine trained on separating linear separable data. The goal of the machine is to do a binary classification task with d-dimensional data points as input ($x_i \in \mathbb{R}^d \; i = 1, ..., m$) where the two classes to separate have labels $y_i \in \{-1, +1\}$. The general case - nonlinear machines trained on non-separable data - results in a similar quadratic programming problem thus this case is worth looking at.

The classification function can be defined as $f(x) = sign(\langle w, x \rangle + b)$. The vector w determines the orientation of a plane separating the classes, the so called discriminant plane. The scalar b is the offset of the plane from the origin of the d dimensional space. The discriminant plane is defined as the hyperplane $H = \{x \mid \langle w, x\rangle + b = 0\}$.

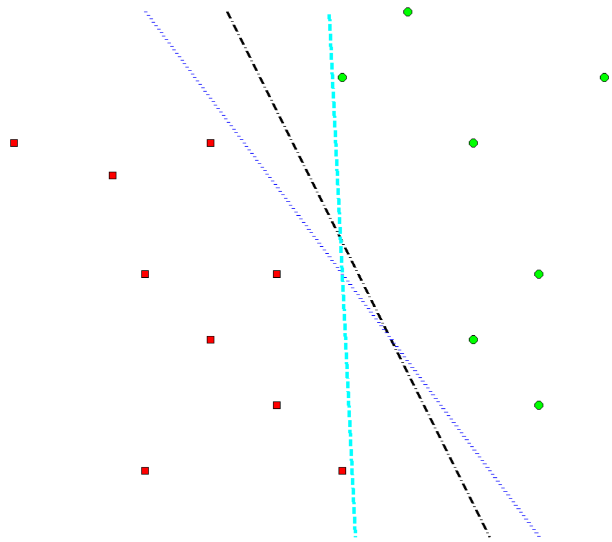

Figure 4.1.: Example for linear discriminant planes

Figure 4.1 shows two classes to separate, indicated by red squares and green circles and some possible discriminant planes. As implied with this simple example there are infinite possible solutions to define the borders between the class representatives. The question is, which hyperplane is the optimal discriminating one and thus performs best.

The intuitive approach is to maximize the margin between the clouds of data points. Thus the closest examples of the two classes have to be taken into account. We define d_+ and d_- as the shortest distance from the separating hyperplane to those examples. Those are either on $H_1 : x_i \cdot w + b = +1$ with the perpendicular distance from the origin $\frac{|1-b|}{||w||}$ or on $H_2 : x_i \cdot w + b = -1$ with the perpendicular distance from the origin $\frac{|-1-b|}{||w||}$. The margin of the separating hyperplane is the sum $d_+ + d_-$. Therefore the margin is $\frac{2}{||w||}$. To find the optimum discriminating hyperplane we have to maximize this sum by minimizing $\frac{1}{2}||w||^2$. This can be expressed with the following inequalities:

$$\min_{w,b} \frac{1}{2}||w||^2 \tag{4.1}$$

$$\text{subject to } \langle w, x_i\rangle + b \geq +1 \text{ for } y_i = +1$$

$$\text{and } \langle w, x_i\rangle + b \le -1 \text{ for } y_i = -1$$

This can be combined to

$$\min_{w,b} \frac{1}{2}||w||^2 \tag{4.2}$$

$$\text{subject to } y_i(\langle w, x_i\rangle + b) \ge 1 \; \forall i$$

The problem is often referred to as the hard-margin SVM formulation as it is non-solvable when the data is not linearly separable. When there is a solution a unique global minimum value can be found as the problem is convex [BC99].

It is important to understand that the solution only relies on the data points closest to the hyperplane, those are called support vectors. In contrast to other machine learning algorithms used for classification it is not relevant for the classification result whether the instances of a class lie dense or sparse, only the support vectors determine the location of the hyperplane.

See figure 4.2 for the optimum hyperplane along with the maximized margin. The support vectors are marked with black crosses.

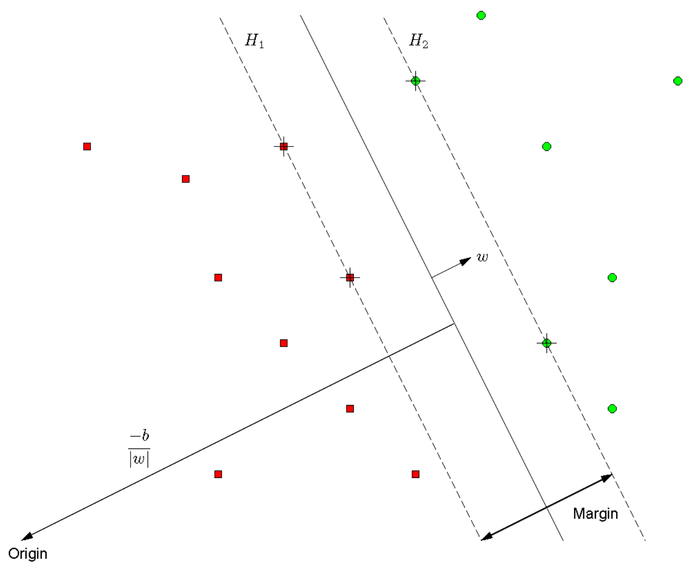

Figure 4.2.: Example for linear separable case in $\mathbb{R}^2$

The problem can be reformulated to a Lagrangian formulation to enable a representation in the form of dot products between vectors in the training algorithms that is needed to generalize the problem to the nonlinear case discussed in section 4.2.3. The Lagrangian dual of the supporting plane's quadratic programming (QP) problem yields the following dual QP [Vap99]:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j \langle x_i, x_j \rangle - \sum_{i=1}^{m} \alpha_i \tag{4.3}$$

$$\text{subject to } \sum_{i=1}^{m} y_i \alpha_i = 0 \text{ and } \alpha_i \geq 0 \; i = 1, ..., m$$

The variables α_i are the Lagrange multipliers. The method of Lagrange multipliers provides a strategy for optimizations with the goal of finding the maxima and minima of a function subject to constraints. Only the support vectors have $\alpha_i \neq 0$ which illustrates that only these points have influence on the location of the hyperplane:

$$w = \sum_{i=1}^{\#sv} \alpha_i y_i x_i^{sv} \tag{4.4}$$

The decision function for the classification is

$$f(x_{new}) = sign(\langle w, x_{new} \rangle + b) = sign(\sum_{i=1}^{m} \alpha_i y_i \langle x_i, x_{new} \rangle + b) \tag{4.5}$$

Data point x_{new} represents a new instance which is a class label assigned by the decision function.

4.2.2. Linear Inseparable Case

Let us consider a scenario like it is given in figure 4.3. The only difference compared to the previous example is an additional red square within the group of green circles.

As can be seen from the plotted convex hulls the two point classes are not linear separable with this additional point. If this single point would be removed or ignored the classification could be done as in section 4.2.1 described. For that reason the margin criterion is relaxed so that some points can lie within the margin or even on the wrong side of the hyperplane. This is the so called soft-margin SVM formulation.

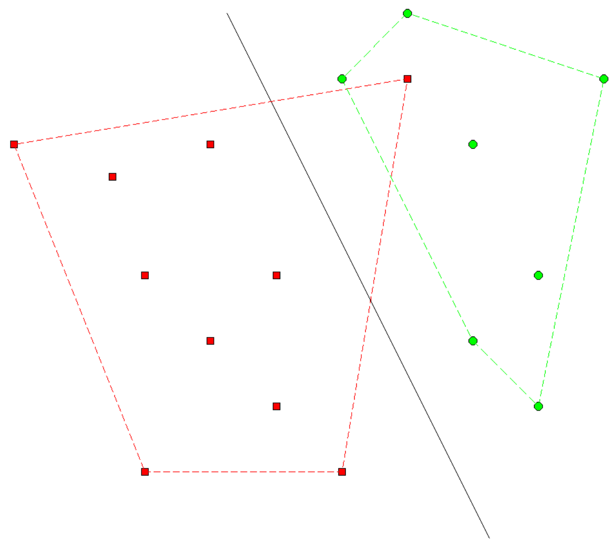

Figure 4.3.: Example for linear inseparable case in $\mathbb{R}^2$

The previous optimization problem is modified by the introduction of the slack variables ξ_i for each constraint as follows [Bur98]:

$$\min_{w,b,\xi} \frac{1}{2}||w^2|| + C\sum_{i=1}^{m}\xi_i \tag{4.6}$$

$$\text{subject to } y_i(\langle w, x_i\rangle + b) \geq 1 - \xi_i \; \forall i \;\text{ and } \xi_i \geq 0$$

The slack variables are added as a weighted penalty term to the optimization problem. The goal is trying to maintain ξ_i to zero while maximizing the margin. Soft margins are a good approach to cope with noise in the data that makes the data not linearly separable. The trade-off-parameter C is set by the user and controls the influence of errors and typically has to be chosen experimental during training of the SVM.

The literature refers to C most commonly as the cost or penalty parameter. A larger C corresponds to assigning a higher penalty to errors. When C goes to infinity the solution gets closer to the hard-margin solution which means fewer errors are tolerated, but overfitting is likely to occur [HCY06]. The advantage of the soft-margin solution is that it always has a solution and that it is more robust to outliers. The hard-margin solution benefit is that is does not require to guess the penalty parameter C.

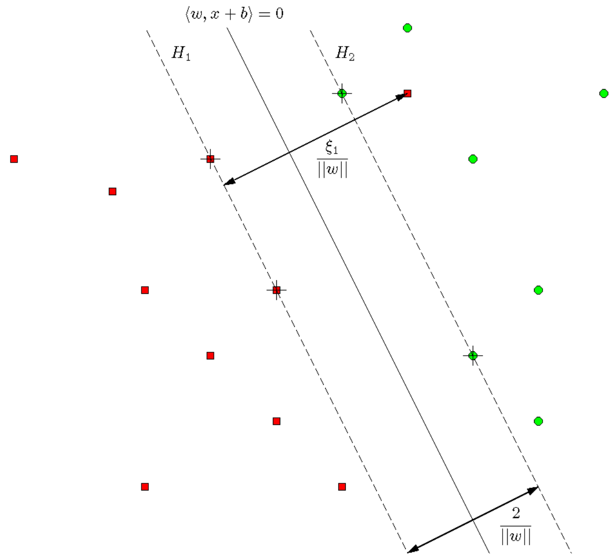

Figure 4.4.: Example for linear inseparable case with slack variable in $\mathbb{R}^2$

The Lagrangian dual changes only slightly with the introduction of parameter C as an upper bound for the Lagrange multipliers α_i :

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j \langle x_i, x_j \rangle - \sum_{i=1}^{m} \alpha_i \tag{4.7}$$

$$\text{subject to } \sum_{i=1}^{m} y_i \alpha_i = 0 \text{ and } C \geq \alpha_i \geq 0 \; \forall i$$

Figure 4.4 visualizes the role of the slack variables ξ_i with the previous example for linear inseparable data.

With this approach linear classifications can be built with theoretical as well as practical generalization properties even in very high-dimensional feature spaces. Luckily, as quadratic programming is a well studied area, robust and efficient algorithms exist for solving the dual formulations [BC00].

4.2.3. Nonlinear Separable Case

Although we have expanded our classification algorithm being able to always find a solution, there are a lot of situations where this might not perform well enough. To motivate this figure 4.5 shows a simple example for data points

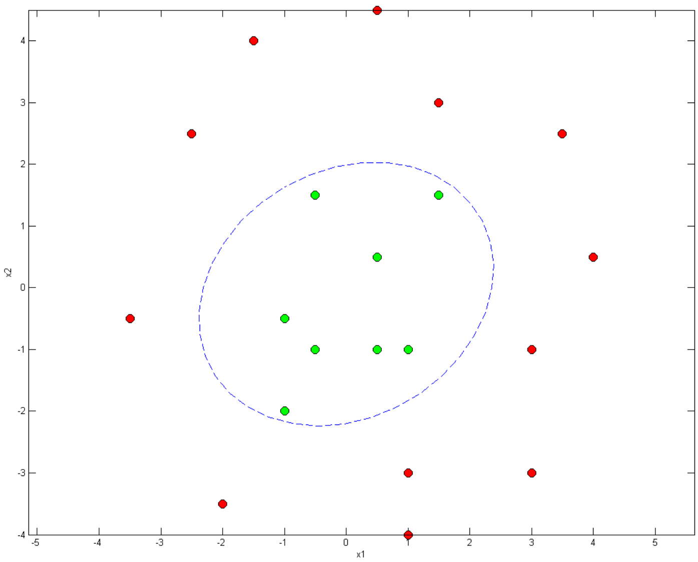

Figure 4.5.: Example for nonlinear separable case in $\mathbb{R}^2$

that cannot be separated linearly and where the misclassification rate would be high when a soft-margin approach is used. The blue ellipse-style border is good nonlinear separation solution, but it is quite complicated in relation to a soft-margin solution.

Let us assume the points in the figure are in $\mathbb{R}^2$ and we map them with the function $\phi : (x_1, x_2) \longmapsto (z_1, z_2, z_3) := (x_1^2, \sqrt{2}x_1x_2, x_2^2)$ to $\mathbb{R}^3$. As can be seen in figure 4.6 this representation enables one to separate the data linearly with a hyperplane between the point clouds. The resulting classification function is $f(x) = sign(\langle w, \phi(x)\rangle + b) = sign(w_1x_1^2 + w_2\sqrt{2}x_1x_2 + w_3x_2^2 + b)$. This principle can be generalized.

The general idea is the introduction of a function ϕ that maps the data into a higher dimensional data space $\mathcal{F}$ in order to make it linear separable. The justification for that can be found in Cover's theorem on the separability of patterns. The theorem states that a nonlinear separable set of training data can be transformed with high probability into a data set that is linearly separable by projecting it into a higher dimensional space via a non-linear transformation [Cov65].

The mapping function can be formally defined as

$$\phi : x \in X \longmapsto \phi(x) \in \mathcal{F} \tag{4.8}$$

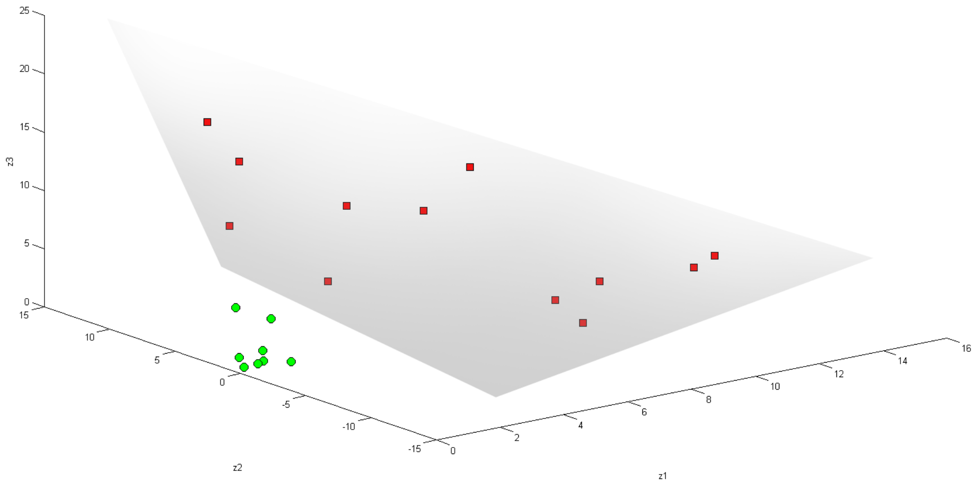

Figure 4.6.: Example for mapped data points $\mathbb{R}^3$

Finding an appropriate mapping function is a nontrivial task which is briefly discussed in section 4.5. The higher dimensional data space is often referred to as feature space or VC-dimension (Vapnik–Chervonenkis dimension).

The problem with this mapping is that the naive approach, an explicit mapping, can be quite expensive in terms of computational effort and/or memory consumption. When processing a lot of data while using high dimensional target spaces the computational costs might even be infeasible. This problem origins from the fact that the dimensionality of the feature space grows exponentially.

There is however an alternative mapping method which reduces processing costs significantly by using kernel functions that map the data implicitly.

The dot products of our example $\phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2) \in \mathbb{R}^3$ can be rewritten as:

$$\begin{aligned}
\langle \phi(x), \phi(z) \rangle &= \left\langle (x_1^2, \sqrt{2}x_1x_2, x_2^2), (z_1^2, \sqrt{2}z_1z_2, z_2^2) \right\rangle \\
&= x_1^2 z_1^2 + 2x_1x_2z_1z_2 + x_2^2 z_2^2 \\
&= (x_1z_1 + x_2z_2)^2 \\
&= \langle x, z \rangle^2 = \kappa(x, z)
\end{aligned}$$

From this evaluation it can be observed that the mapping only enters through

Kernel name	**Kernel function $\mathbf{K}(\mathbf{x},\mathbf{z})$**
Linear Kernel	$\langle x, z\rangle + c$
Polynomial Kernel of degree n	$(\gamma \langle x, z\rangle + c)^n$
Gaussian Radial Basis Function (RBF)	$exp(-\gamma\|\|x - z\|\|^2)$

Table 4.1.: Examples of kernel functions

inner products. This means that for training purposes information on pairwise inner products is sufficient. Therefore a kernel is an efficient way to compute dot products of mapped data without any explicit mapping through extensive calculation.

The definition of a kernel is as follows [CST00]: A kernel is a function $\mathcal{K}$, such that $\mathcal{K}(x, z) = \langle\phi(x), \phi(z)\rangle$ for all $x, z \in X$ where ϕ is a mapping from X to an (inner product) feature space $\mathcal{F}$.

It remains the question when is a function an actual kernel function? The answer is given by Mercer's theorem [Bur98]:

1. Function $\mathcal{K}$ is symmetrical, i.e. $\mathcal{K}(x_i, x_j) = \mathcal{K}(x_j, x_i)$
2. The Kernel-Matrix $K_{ij} := \mathcal{K}(x_i, x_j)$ is positive semi-definite for all training data $x_1, x_2, ..., x_n$, i.e. $\sum_{i,j} a_i a_j K_{ij} \geq 0 \; \forall a \in X$

Kernels are closed under the following operations, so we can combine known kernel functions in order to yield valid kernels again [CST00]:

1. $\mathcal{K}(x, z) = \mathcal{K}_1(x, z) + \mathcal{K}_2(x, z)$
2. $\mathcal{K}(x, z) = \alpha\mathcal{K}_1(x, z), \alpha \in \mathbb{R}^+$
3. $\mathcal{K}(x, z) = \mathcal{K}_1(x, z)\mathcal{K}_2(x, z)$
4. $\mathcal{K}(x, z) = \mathcal{K}_1(x, z) + c, c \in \mathbb{R}^+$
5. $\mathcal{K}(x, z) = x'Bz, X \subseteq \mathbb{R}^n, B$ positive semi-definite

See table 4.1 for some examples for common used kernel functions. The parameter c is a constant term, the parameter γ is an adjustable kernel parameter that plays a major role in the performance of the kernel and has to be experimentally tuned to the classification problem. The selection of the best kernel is also of high importance for the performance of the classification. Section 4.5 touches briefly on that.

The learning algorithms of SVMs are laid out such that all references to the input data are within inner products so that all these occurrences can be replaced with the kernel function of choice. After that the usual linear training algorithms can be applied.

The quadratic programming problem with the mapping is defined as [BC00]:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j \langle \phi(x_i), \phi(x_j) \rangle - \sum_{i=1}^{m} \alpha_i \tag{4.9}$$

$$\text{subject to } \sum_{i=1}^{m} y_i \alpha_i = 0 \text{ and } C \geq \alpha_i \geq 0 \; \forall i$$

Still missing from the discussion is the nonlinear inseparable case which is the general case mentioned in the beginning of this chapter. By now we already have all tools to tackle that kind of problem. The kernel maps the data into a higher dimensional data space in order to make the instances of the two classes linear separable. In the cases where this fails the penalty parameter C controls the influence of errors introduced through noise and outlier.

The implementation of the optimization methods that are used to solve the quadratic programming problem in equation 4.9 is not in the scope of this thesis. However, according to [BC00] common methods include interior-point methods [PW00] and sequential minimal optimization [Pla99].

4.2.4. Classification Summary

This section will give a short summary on kernels and SVMs to aggregate the most important facts that should be kept in mind when the conducted experiments are introduced and discussed in chapter 8.

Unlike other well-known machine learning classifiers such as Gaussian mixture models SVMs work by increasing the dimensionality of the input data space - and not reducing it. The workflow of a SVM is as follows: first the input data is mapped into a higher dimensional space where linear classification techniques can be used. The learning algorithm where the margin between the classes is maximized is implemented so that the mapped data is only used in pairwise dot products. This enables an efficient computation by a corresponding kernel function on the unmapped data.

Goal of the SVM training is the maximization of the margin in the implicit feature space to maximize the generalization performance and avoid overfitting.

To be able to work with somewhat noisy data, slack variables are introduced so that outliers have a limited influence on the training result. In order to achieve the best performance experiments have to be carried out that evaluate different parameter settings.

The optimization method of a support vector machine is quadratic programming, which is a well-studied and understood mathematical programming technique.

4.3. Multi-class SVMs

As discussed before SVMs inherently offer solutions for two class problems. There are of course ways to extend the concept to solve multi-class problems as well. The two most often used techniques used to transform the multi-class problem into one or multiple two class problems are: One-separates-rest and One-separates one.

The first approach, One-separates-rest, trains one binary SVM classifier per class, separating the data vectors of this class from all other classes' data vectors. The classification is done with a winner-takes-all approach, where the classifier with the highest output score is the one to which the instances are assigned.

With the second technique, One-separates-one, a total of $\frac{N(N-1)}{2}$ (N is the number of classes) binary SVM classifiers are built, where each one is trained using data vectors from a pair of classes. Here the classification is based on a max-wins voting technique, where each classifier assigns the instance to one of its two classes. The class with the most votes are the instances assigned to.

Multi-class SVMs are not of relevance for the practical application discussed in this thesis, therefore the concepts are not discussed further. However, further reading can be found in [CSC$^+$01, DK05].

4.4. Benefits and Applications

In recent years many machine learning related applications have successfully adapted SVMs implementations.

Not only classification, but also regression [SS04] and novelty detection [SPST+01] belong to the problems that can be solved with SVMs. The underlying concepts of margin maximization, duality and kernel functions enable one to extend and combine the methodology in order to solve the particular problem.

Compared to other popular machine learning techniques such as neural networks or decision trees SVMs do not suffer from problems like local minima or overfitting as the results are stable and reproducible. The number of training parameters is low, e.g. if the Gaussian radial basis function (RBF) is used for classification it comes down to choosing two parameters: the width of the Gaussian kernel γ and the penalty parameter C. A further benefit is that state-of-the-art research optimization problems can be used for solving the quadratic programming problem of the SVM model [BC00]. The application of SVMs on practical problems is rather uncomplex and as discussed in section 4.6 supported by public available software tools.

The benefits of SVMs can be confirmed with a wide range of successful application areas. These include DNA analysis [TCD+00], handwritten digit recognition [DS02] and Steganography detection [LF02]. In [MLH03] the classification and regression performance of SVMs was compared to a range of other machine learning methods good and overall competitive performance on 33 different machine learning data sets.

Especially in biometrics SVMs are often used as the classification backend processing the extracted biometric features for identification or verification. Biometric modalities used include fingerprints [BSC04, TTT06], iris [ASW08], face [RPL99] and speech [CLG09]. Machine-vision-based gait recognition is also among the successful biometric applications of SVMs [LY09].

Another notable form of SVM application in the area of biometrics is when multi-biometrics are used [DDV07, FAE08, JK08].

4.5. Challenges and Limitations

Despite the benefits there are still issues to solve to make SVMs an ultimate tool for machine learning tasks.

It was mentioned that the parameters of SVM kernels are not numerous, but this is already one step beyond the selection of the kernel type which is important for the results. In [HCL03] the RBF kernel is recommended along with a description in which cases to prefer another kernel type. As briefly discussed the most common approach for kernel and parameter selection is an experimental one which is typically done with cross validation. This can be quite expensive in terms of computational efforts.

Also, SVMs do not solve the problem of selecting well performing, stable and complete attributes prior to training a model. This is still an area which is largely dependent on domain-knowledge and the intuition of the researchers.

The inputs of SVMs are real-valued vectors hence there has to be a conversion of non-real-vectors like categories into such values, where different conversion methodologies can be applied. Scaling of the data is considered a good practice because it avoids values in greater numeric ranges dominating those in smaller ranges. There is also the possibility that large values cause numerical problems when used in the inner products of features vectors. It is still not entirely clear which scaling approach is best for which problem [BC00].

Although being not always of relevance, it is straight forward that one limitation of SVMs is that the results, the trained models, are not interpretable. This means that the separating characteristics of the two classes cannot be recognized by interpreting the position of the support vectors.

4.6. SVM Tools and Libraries

One of the mentioned benefits of SVMs is the availability of mature software toolkits that are comparable easy to use and often available both as Open Source and for free.

The major requirements for this project was that the software provides an efficient implementation, is well documented and most importantly is also available for programming languages that can be used in smartphone applications. The LIBSVM library could fulfill all those needs and more. It is actively developed, the latest version was released in September 2010, it is free and the source code is available as well.

Besides both a Java and C++ implementation it comes along with a variety of interfaces to other languages and toolkits like Python, R, MATLAB, Perl, Ruby and others [CL01]. This flexibility allows one to use the most efficient

C++ implementation with the MATLAB interface to work on powerful workstation computers in order to experiment with different attributes and model parameters. The best solution and the generated SVM models can then be used within the smartphone environment and, for example on Android-based smartphones, with the Java implementation of the library.

5. Data Collection

A data collection was conducted to construct a set of training and test data that is used to create, test and optimize the gait recognition algorithms. This chapter describes the actual collection process and the quality of the collected gait characteristics.

5.1. Sample Collection

For the acquisition of the acceleration data a Motorola Milestone smartphone was used. It has a built-in 3 axes accelerometer sensor, a STMicroelectronics LIS331DLH [And10]. The technical capabilities of it are given in table 5.1.

Accelerometer Type	3 axes
Data scales	User selectable ±2g/±4g/±8g
Output data rates	0.5 Hz to 1 kHz
Data resolution	16 bit

Table 5.1.: Data sheet LIS331DLH accelerometer [STM09]

The smartphone has the Android platform as operation system which provides an API to access the sensor data values. For the data collection a custom application was used that writes the acceleration values for the x-, y- and z-direction along with a timestamp into a local database. Later, the content of the database is exported to files. The application was developed within CASED as part of the biometric authentication framework MBASSy [WN10].

The Android API does not provide a way to get acceleration values in fixed time intervals and there is no way to ensure that a certain minimum data rate is met. An example of captured sensor values is given in figure 5.1. The first column contains timestamps in milliseconds, the other columns are the x-, y- and z-direction (left to right) acceleration values. Note the different time intervals between consecutive measurements.

0	9.728196	2.5497289	3.6088471
14	8.786758	2.824315	3.8245933
22	8.786758	3.079288	3.8245933
30	8.855405	2.8733485	3.8245933
38	9.885103	2.0790098	3.8245933
46	10.708861	1.4513842	3.8245933
54	10.885382	0.6668522	4.06976
61	11.571847	0.4707192	4.1285996
69	11.571847	0.2157463	4.059953
77	11.53262	-0.02941995	3.1773546
86	10.395049	-0.22555295	2.3143694
93	9.492837	-0.26477954	2.3143694
101	8.551398	-0.26477954	2.2457228

⋮

Figure 5.1.: Acceleration signals excerpt

Every subject that took part in the data acquisition process was asked to walk on a tarred footway, figure 5.2 is a photo of the path. During this walk each subject carried a horizontal belt pouch at the right hip which contained the Motorola Milestone.

Figure 5.2.: Footway used for data collection

Figure 5.3.: Position of the smartphone in the belt pouch

The phone was positioned with the display to the body and the upper side facing to the walking direction. See figure 5.3 for the position of the pouch and the phone. The direction of the three spatial dimensions of the accelerometer is visualized in figure 5.4 on the following page.

The detailed data capture steps were as follows:

- The subject got the belt pouch attached to the belt or to the trousers
- The capture process was started and the phone was put into the pouch
- The subject had to wait until the smartphones vibrated, then the subject had to start walking
- All acceleration data that was captured during this walk belongs to the first walk
- After the next vibration the subject had to stop, turn and wait for the next vibration
- When the smartphone vibrated again, the subject had to start walking back
- The acceleration data that was captured during this walk belongs to the second walk
- The next and last vibration signaled the end of the capturing process

Figure 5.4.: Directions of the accelerometer dimensions

The phone was positioned with the display to the body and the upper side facing to the walking direction.

The time per walk was set to be 18 seconds. Thus the walk distance covered depends on the walking speed of each subject. The time between the walks was seven seconds which is enough time for the subject to stop, turn and wait for the beginning of the second walk.

Three settings were captured for each subject. First the subjects were asked to walk in their normal pace. Then this procedure was repeated with a slower walking speed and, in the third and final pass, with a faster pace.

As the walking style of each person varies over time it is a good approach to capture gait data on different days to enable a more realistic testing. Thus the three capture settings were repeated a few days (2-3) later. So in total six settings with two walks per setting were recorded for each person. Note that the subjects were asked to wear the same shoes that they wore on the first data collection session on the second session again.

A total number of 41 persons took part in the first walking session, and except one participant all attended the second session. The participants were 19 female and 22 male persons. The detailed age and gender distribution is given in figure 5.5.

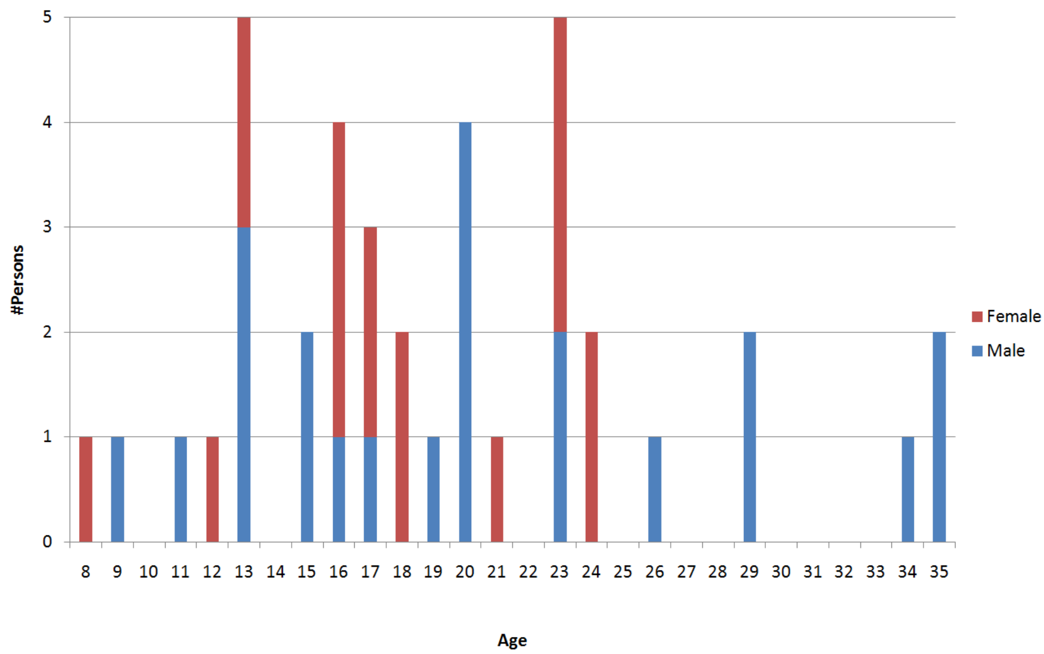

Figure 5.5.: Subject age and gender distribution

5.2. Walk Extraction and Analysis

The files with the acceleration data that had been acquired during the data collection were processed and analyzed to do a basic quality assurance.

As a first step Matlab scripts were developed that perform basic quality checks. These steps ensured that

- for every subject six files (one per setting on a day) were created
- the data of the single subject that did not return for second day data collection was dropped
- all sample files covered the same time period, i.e. full 18 seconds

5.2.1. Extraction

In the next step the two walks per file were extracted into two separate files. For the extraction the WalkExtraction Matlab GUI program, also developed within CASED, was used.

The process was done in a semiautomatic fashion: first the starts and ends of the walks were determined automatically. This is easy to implement, as the timestamp change between subsequent values is here five seconds at least. Next, the signals were visually controlled. If there were unusual walk cycles during the start phase of a walk these were omitted by manually adapting the selection range of the respective walk.

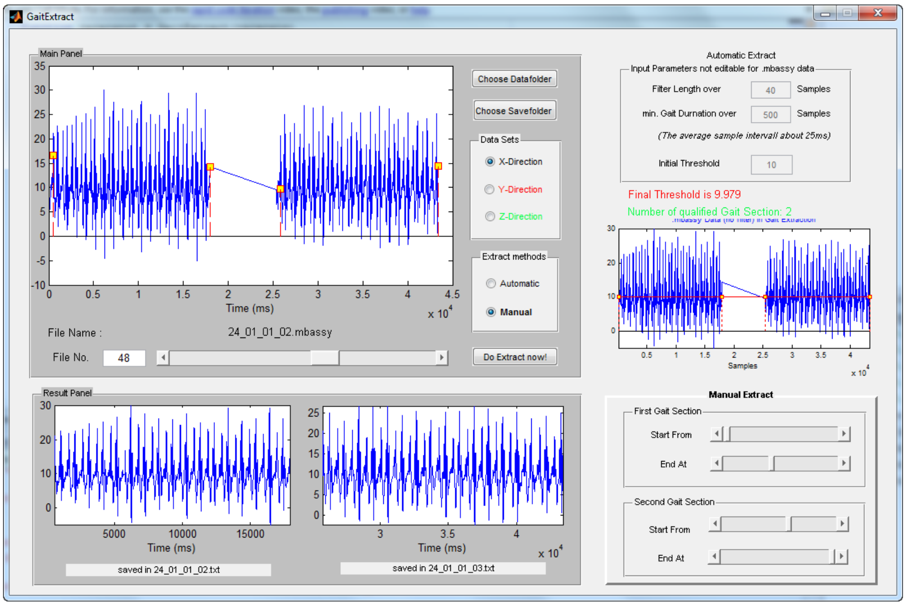

Figure 5.6.: Semiautomatic walk extraction

After the separation into one walk per file a total number of 480 walk files existed, which result from 40 persons with six sample files with two walks each. In figure 5.6 an example of an extraction setting is given. The data from the original file is visible in the upper large window and the two extracted walks below.

5.2.2. Sample Analysis

As mentioned in section 5.1 there is no guaranteed sampling rate due to the architecture of the Android platform and developer API. To check that the variability of the sampling rate is not too high the mean sampling rate of each walk file was calculated.

The results can be seen in figure 5.7 on the next page. It is apparent that overall the sampling rate is more than 125 samples per second with few outliers.

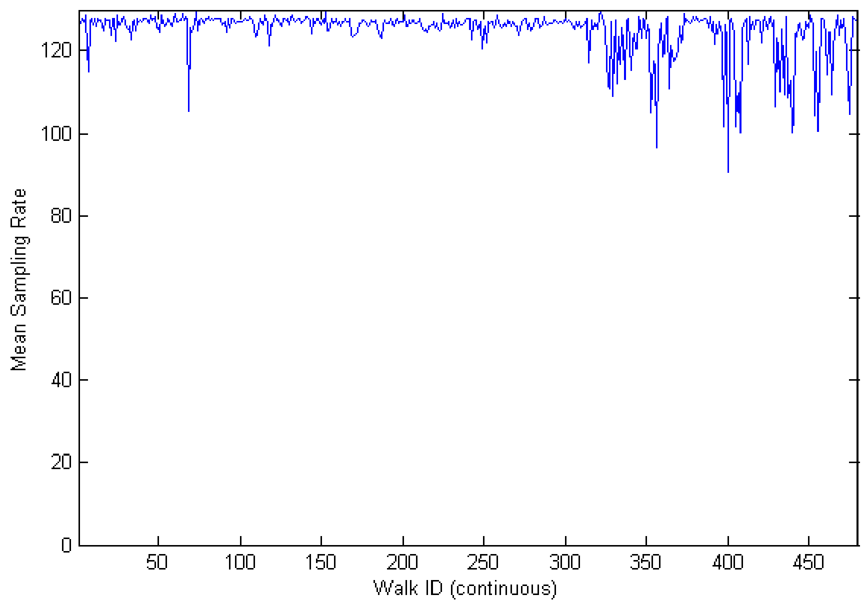

Figure 5.7.: Walk files mean sampling rates

The impression of a sufficiently good data quality can be further enforced by looking at the standard deviations of the time differences of the walk files. The standard deviation of the time differences of a walk file is generally under 10ms. This can be seen in figure 5.8 on the following page.

Furthermore it was examined whether there had been significant lags between subsequent data values by analyzing the maximum time differences of the walk files.

Though figure 5.9 on the next page reveals there are few rather big outliers, the overall maximum time lag is smaller than 100ms. Of course, the maximum time lags resemble the distribution of the discussed standard deviations.

The reason that time lags of this magnitude occur cannot be clarified as it is unknown how the Android operating system handles the output of the acceleration sensor. Most likely in such moments other background services or tasks are executed that have a higher priority and thus lead to a delay in the updating of the measured accelerations.

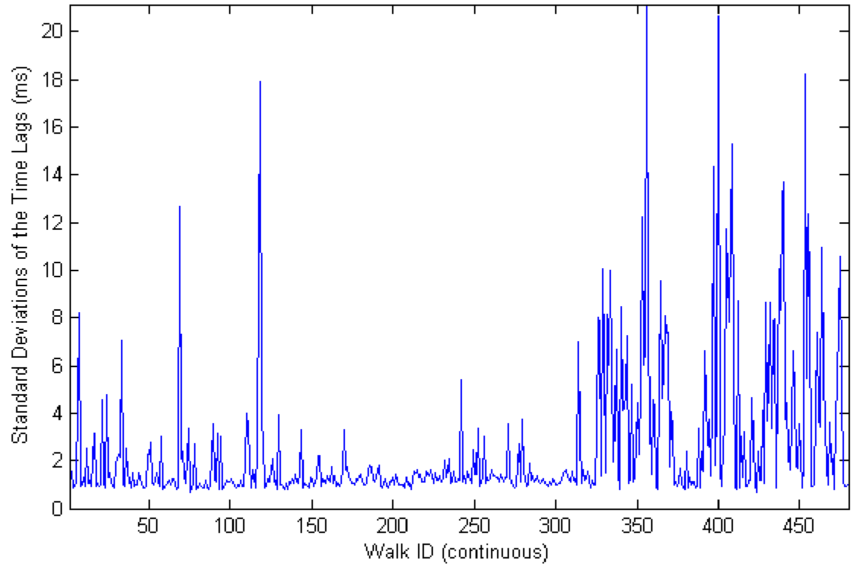

Figure 5.8.: Walk files time lag standard deviations

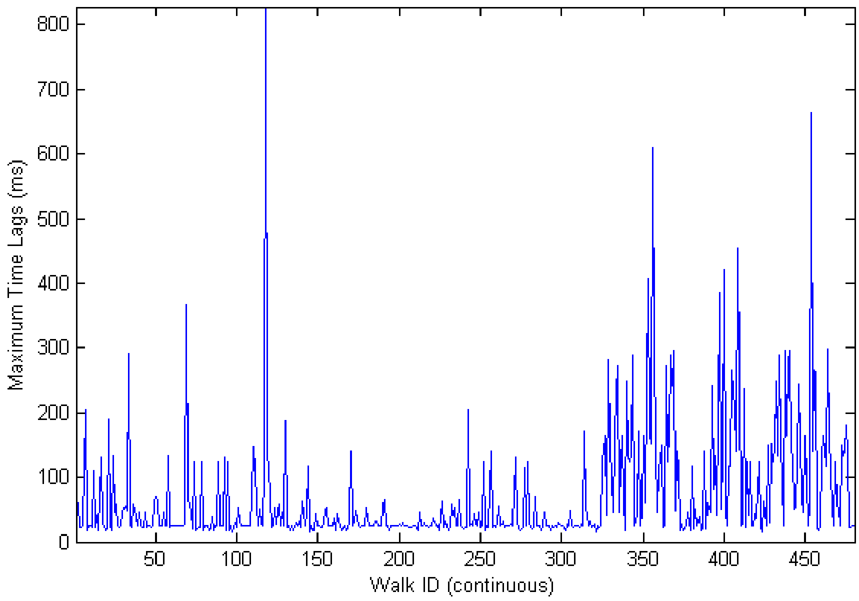

Figure 5.9.: Walk files maximum time lags

Figure 5.10 shows the setting with the worst delay throughout the whole data collection of about 800ms.

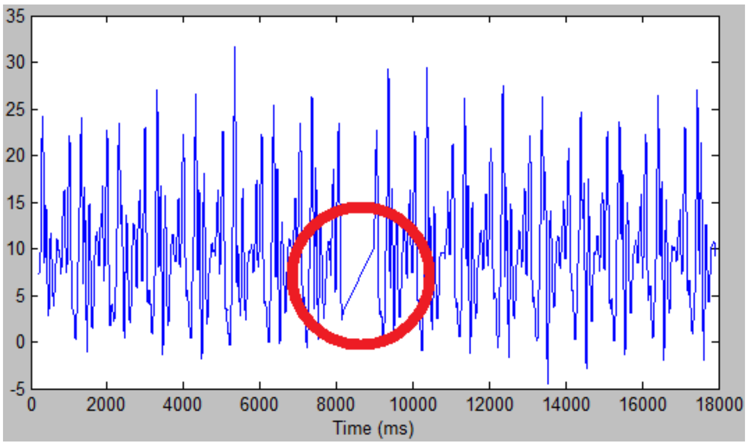

Figure 5.10.: Maximum time lag in collected walks

The acceleration sensor has of course limits in terms of accuracy. Although for our purposes it is not primary that the sensor is perfectly calibrated it makes sense however to look at the sensor resolution.

This can be done by producing a set of all measured acceleration values for each of the three directions. Then, all duplicate entries are removed from the sets resulting in three sets of unique values. These sets are sorted and the difference of subsequent entries is calculated. Now, the minimum is determined for each set. The resulting values can be seen in table 5.2. As a result we can speak of a maximal resolution of approximately $0.01\frac{m}{s^s}$.

Direction	Min. Acc. Difference (m/s^2)
x	0.009802
y	0.009805
z	0.009803

Table 5.2.: Minimal acceleration difference

Although it is not possible to guarantee a certain sampling rate by using the Android operating system it seems that most of time the sampling rate is sufficiently high compared to the devices used in other WS-based gait recognition studies. In the cases of the big time lags the signal should not be used for gait recognition as false results are likely to occur.

6. Feature Extraction

This chapter will outlay how the gait characteristics in the form of a time series of acceleration measurements are processed to be usable as features for SVMs. The evaluation of the feature types that are described here can be found in chapter 7.

6.1. Preprocessing

The collected gait samples are preprocessed such that the feature extraction works with consistent and portioned data. The walk files serve as input, in this section also referred to as *raw* data. Note that for the experiments of this thesis only the walks of the normal pace setting were used.

6.1.1. Linear Interpolation

The first step is a linear interpolation to a fixed sampling rate as this is not given with the collected raw data as discussed in section 5.1. The average sampling rate of the raw data is about 125 samples per second.

To find the optimum target sampling rate various settings have been tested, this is later described in chapter 7.

6.1.2. Normalization

After the interpolation the signals s_x, s_y, s_z are normalized by the mean acceleration μ_a of the respective acceleration direction s_a

$$\bar{s}_a(t) = s_a(t) - \mu_a \quad a = x, y, z$$

The intention is to be able to extract features that need an input signal with a near-zero mean value. An example is the zero crossings metric. Before the normalization the mean signal of the acceleration direction aligned with the gravitation is about $10\frac{m}{s^s}$.

Normalization also helps to limit the influence of sensor inaccuracy and noise which can be seen from the non-zero acceleration measured when the device is not moved at all.

6.1.3. Segmentation

One of the goals of this thesis is to find a representation of human gait that is suitable for distinguishing between individuals for the verification purpose. As mentioned during the discussion of current research efforts in accelerometer-based gait recognition in section 3.3 an often implemented idea is to compare the characteristics of a gait cycle.

The downside of this approach is that the correct identification of a gait cycle is not trivial. Unsteadiness within gait cycles as well as noise influences of the environment (e.g. changing surfaces) might lead to failures to select valid gait cycles from the signal and thus the genuine user can be rejected.

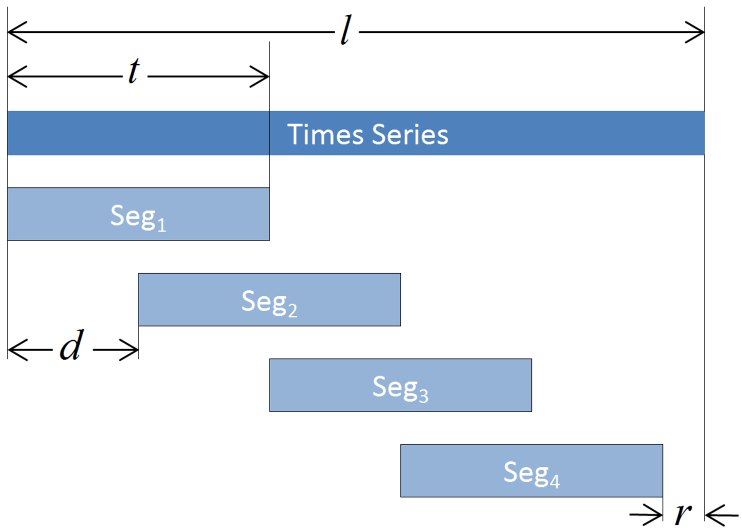

Figure 6.1.: Sliding window segmentation

To avoid these problems a different approach will be followed in this work. The acceleration samples are separated into parts of several seconds using a sliding window approach with overlapping rectangular windows. This means that the original signal of length l is splitted into *segments* of length t and a distance d between consecutive segments. The remaining part of length $r = (l-t)mod(d)$ is dropped and not used any further. The segmentation process is visualized in figure 6.1.

The segmentation is done for all interpolated walks. Note although we referred to the process as segmenting a signal actually three normalized acceleration signals $\bar{s}_x, \bar{s}_y, \bar{s}_z$ are segmented, so that the signals are still correctly aligned to each other.

6.2. Feature Types

Although the acceleration signals are segmented at this stage they are still represented as time series. As the intention was to benefit from the well-performing classification capabilities of SVMs a transformation to a fixed length vector of discrete values has to be conducted. For each segment one feature vector is created.

The next sections will describe the implemented and evaluated features. The first section describes statistical features whereas the two sections after that focus on features best known from the study fields speech and speaker recognition. All the evaluation results can be found in chapter 7.

6.2.1. Statistical Features

Mean Arithmetic mean of the acceleration signal

Maximum Maximum (positive) acceleration

Minimum Minimum acceleration (maximum negative acceleration)

Mean absolute difference Arithmetic mean of the difference between consecutive accelerations values

Standard deviation Standard deviation of the acceleration signal

Root mean square Square root of the arithmetic mean of the squares of the original acceleration

Binned distribution Histogram distribution using a number of equally spaced bins ranging from the maximum to the minimum acceleration value for each acceleration signal. The distribution was used relatively, i.e. the values were in the range $[0; 1]$.

Zero crossings Number of sign changes within the acceleration signal. As mentioned the normalization during the preprocessing ensures that also the direction under influence of the gravitation has meaningful zero crossings.

Mean, mean absolute difference, standard deviation and binned distribution have also been used in [KWM10].

6.2.2. Mel-Frequency Cepstral Coefficients

The mel-frequency cepstral coefficients (MFCC) belong to the most widely used spectral representations of audio signals for automatic speech recognition and speaker verification [SH04, GFK05]. This section will give a quick overview of the processing steps that are involved in creating the MFCC features. A more elaborate discussion can be found in [RJ93].

The general workflow for creating MFCC is laid out in figure 6.2.

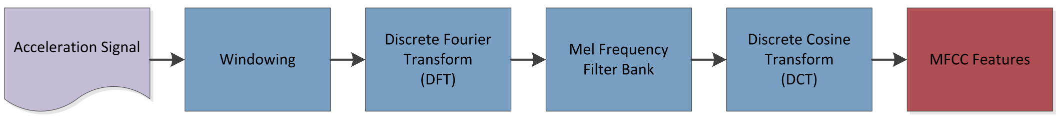

Figure 6.2.: MFCC feature creation

Note that in literature some descriptions denote the so called pre-emphasis as a first step of the MFCC creation. This is a filter which increases higher frequencies. In some of the experiments described in the following chapter pre-emphasis was used, in some others not.

In the next step the signal is divided into overlapping frames. The length is in speech related applications typical at 10 to 25 ms. This is usually done with a windowing function like a Hamming window so that edge effects are removed. Note that we distinguish between segmentation during preprocessing producing *segments* and windowing within MFCC producing *frames*. Multiple frames are generated from one segment. For each frame a cepstral feature vector is created.

Afterwards the DFT (Discrete Fourier Transform) is computed for each frame. Only the logarithm amplitude spectrum is used further.

The third step is to smooth the spectrum while emphasizing the meaningful frequencies. This is done by arranging the spectral components into frequency bins that are spaced according to the Mel-frequency scale. The scale is based on findings that for speech lower frequencies are perceptually more important than higher frequencies.

The Mel-scale approximates the mapping between actual frequency and perceived pitch for the human auditory system. It is approximately linear below 1 kHz and logarithmic above [UCN99]. The classic work of Stevens and Volkmann [SV40] is the foundation for all variations of the Mel-scale [UCN99]. A Mel is a unit of pitch, see equation 6.1 for the definition of an often used representation [O'S99].

$$Mel(f) = 2595 log_{10}(1 + \frac{f}{700}) \tag{6.1}$$

In the last step the DCT is used to approximate the Karhunen-Loeve transform [Log00]. The motivation is to reduce the highly correlated Mel-spectral vectors to fewer decorrelated parameters [Log00]. In the area of speech and speaker recognition it is common to extract 13 cepstral features for each frame. To yield one feature vector for a segment the mean cepstral feature vector is calculated at the basis of the feature vectors the frames.

Note that a variety of MFCC implementations have been proposed for speaker verification tasks, in [GFK05] four of the most popular versions are compared.

To the best of the author's knowledge, MFCC have not been used for gait recognition previously. The MFCC transformation of a signal is optimized to the characteristics of human hearing, therefore it has to be studied whether it can be successfully applied to gait recognition as well.

The motivation is to study which MFCC implementation variation works best for gait data and to optimize the discriminative power of the coefficients through an optimization process which will be described in the next chapter. Therefore, to carry out the optimization, a versatile MFCC implementation was needed.

Three MFCC implementations for Matlab that are often used for speaker recognition were evaluated:

- Auditory Toolbox [1]
- Dan Ellis Script Collection [2]
- Voicebox [3]

All of these are available for free and in source code. The MFCC implementation of the Auditory Toolbox only provides the basic parameters sampling rate and frame rate whereas the others provide much more elaborate settings. An example is the setting of the frequency bandwidth which is in most MFCC implementations per default in the range from 0 Hz to 4000-8000 Hz whereas the useful frequency spectra of human motions is considered to be between 0 and 10 Hz [Win09].

As Dan Ellis' MFCC implementation was created with the intention to be able to reproduce the output of other MFCC programs such as the Auditory toolbox or the HTK[4] it was favored over the Voicebox implementation. Dan Ellis also published the results of the reproduction of the MFCC feature outputs of common programs on his website [Ell].

6.2.3. Bark-Frequency Cepstral Coefficients

Similar to MFCC the bark-frequency cepstral coefficients (BFCC) are an application of a scale to a signal spectrum. The only difference is that instead of the Mel-scale the Bark-scale is applied.

The Bark-scale was proposed in 1961 [Zwi61]. A difference of one bark corresponds nearly to a pitch interval of 100 mels. For the conversion of frequencies to bark the following mapping from [Tra90] is often used:

$$Bark(f) = \frac{26.81f}{1960 + f} - 0.53 \tag{6.2}$$

[1] Auditory Toolbox Version 2 by Malcolm Slaney, URL: http://cobweb.ecn.purdue.edu/~malcolm/interval/1998-010/

[2] PLP and RASTA (and MFCC, and inversion) in Matlab, URL: http://labrosa.ee.columbia.edu/matlab/rastamat/

[3] VOICEBOX: Speech Processing Toolbox for MATLAB, URL: http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html

[4] Hidden Markov Model Toolkit, URL: http://htk.eng.cam.ac.uk/

In the chosen MFCC implementation an option is available to use the bark-scale for calculating the cepstral coefficients. This program is implementing the bark-scale as shown as in equation 6.2 just like its mel-scale is based on equation 6.1.

Again, a vector of 13 features is produced per segment by calculating the mean feature vector of the frames of the segment.

7. Feature Selection and Optimization

This chapter documents the experiments conducted to select the best performing features. Additional results can be found in part A of the appendix.

The gait data collected from the 40 subjects was used for conducting the experiments described in this chapter.

7.1. Classification and Optimization

The process of converting the raw acceleration data to feature vectors that are used for SVM training and testing has, as previously mentioned, quite a lot of possible approaches. Therefore a systematic evaluation is needed which approach performs best in terms of biometric performance while making sure to keep the computational effort needed reasonable low to enable a usage on mobile devices.

In order to evaluate the chosen preprocessing steps and features an automated process was developed that experimentally determines the performance.

7.1.1. SVM Training and Testing

The main application scenario is the verification where a person claims a certain identity. Hence the classification differentiates between the one genuine subject and all other subjects (imposters), which is a One-separates-rest classifier.

Cross-validation was used to verify that the performance results obtained are reasonable stable for biometric samples of various persons. Consequently, for each experiment N One-separates-rest classifiers were built for the N subjects in the utilized data set.

The available data had been first preprocessed as discussed in section 6.1 and then the features introduced in 6.2 were extracted. Half of the resulting feature vector instances were used for training, the other half for testing. Note that only the instances from the genuine subject occur in both data sets to be able to test that the genuine is recognized correctly. The tests were carried out with previously unseen biometric samples.

For the SVM training and testing the Matlab interface of LIBSVM was used that enables one to call functions in the compiled binaries of the LIBSVM C implementation.

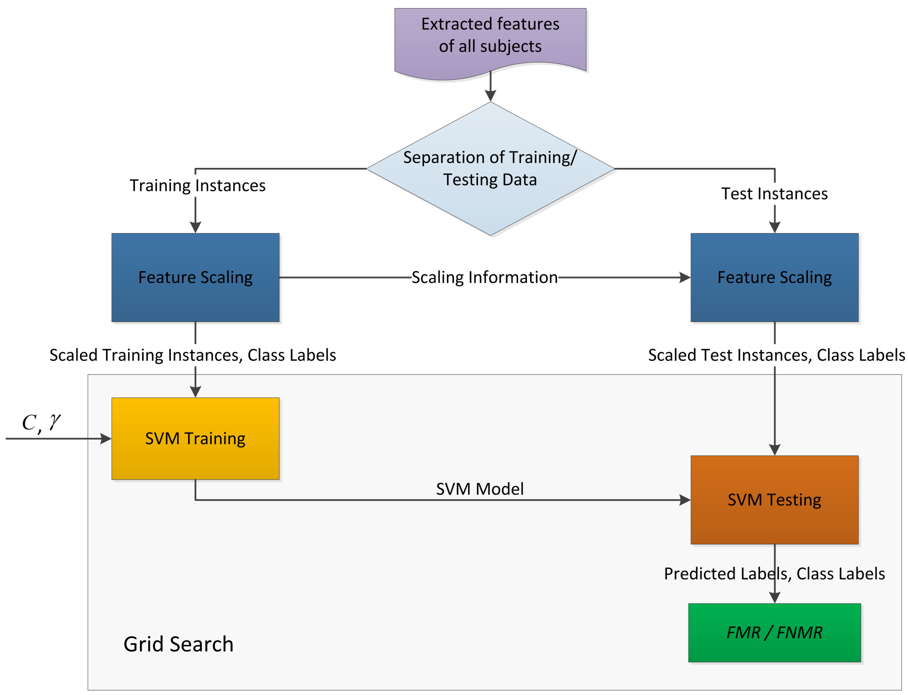

Figure 7.1.: SVM training and testing

Figure 7.1 gives an overview of the process of building one One-separates-rest classifier and the performance evaluation using the test instances.

Prior to training each feature of the training instances is scaled to the range $[0; 1]$. The information that is needed to transform the original features to this range is saved in order to be able to apply the same transformation to the features of the test instances. This approach is considered best practice in [HCL03].

As the correct class labels are known during the testing the number of false matches and the number of false matches can be calculated. To evaluate the performance of the system the false matches and the false non-matches of all N classifiers are counted. Using the sums and the number of imposter and genuine samples the false match rate and the false non-match rate can be calculated.

7.1.2. Parameter Selection

As discussed in 4.4 the performance of SVM models largely depends on the selection of the kernel, the kernel's parameters and penalty parameter C. The Gaussian radial basis function kernel was used as it is considered being a good choice for most classification tasks [HCL03]. To find the best setting for the kernel parameter γ and penalty parameter C a grid search was implemented which is described in [HCL03]. In this approach the cross validation performance is compared for certain parameter ranges that are spaced logarithmical.

As opposed to other machine learning tasks the classification rate is not of primary interest during this optimization. The reason is that the number of imposters' attempts is much higher than the number of genuine subjects. Therefore a high classification rate might hide the fact that the genuine subjects are not accepted whereas the imposters are correctly rejected. To have a more reliable optimization indicator the sum of FMR and FNMR was used.

That means that for each parameter combination of C and γ the cross validation is conducted giving a FMR and FNMR each time. If reasonable, the best performing parameter ranges are searched again with a finer grid in a second step. The goal is in any case to find the optimum parameters yielding the lowest sum of FMR and FNMR. Figure 7.1 visualizes what part of SVM training and testing is done within the grid search process.

Furthermore, the error rates received were stored so that a graphical representation of the performance against the parameter settings could be generated. An example is given in figure 7.2 where the results of a grid search for the parameter ranges $\gamma \in \{2^{-8}, 2^{-7}, ..., 2^{-3}, 2^{-2}\}$ and $C \in \{2^{6}, 2^{7}, ..., 2^{15}, 2^{16}\}$ are visualized in a 3D figure viewed from two sides.

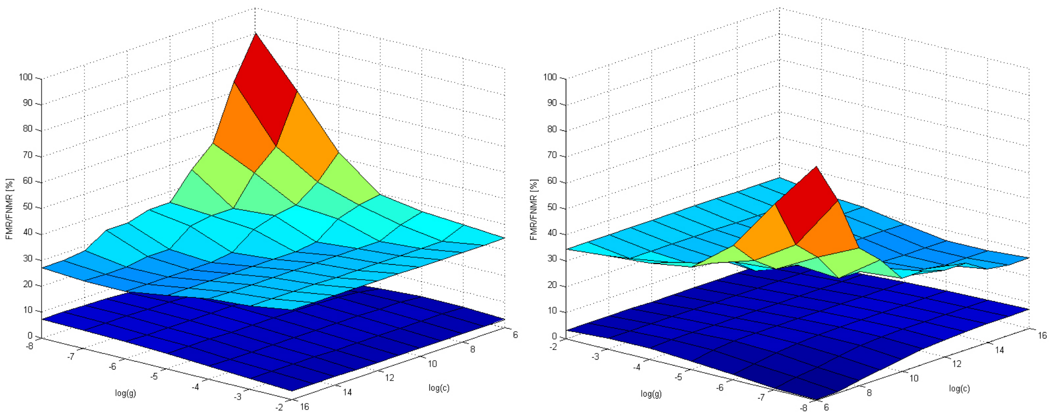

Figure 7.2.: Grid search result example

The corners of the colored rectangles are the results for the error rates, the planes are build by linear interpolation between subsequent parameter settings. The upper plane represents the results for the FNMR and the lower plane the corresponding results for the FMR. As it can be seen there are parameter settings which would lead to a poor performing biometric system although the features have enough classification significance as can be seen from the area with comparable low error rates.

A second or third parameter search with a finer grid is only reasonable when the classification performance is quite unstable with the chosen grid. Of course, it is preferable to have results that are stable for larger parameter ranges. The graphical representation makes it easy to have an impression of the stability of the results and of the potential for a better classification performance with a finer grid search.

7.2. Feature Selection

In a first step the dissimilarity capabilities of single features and of combined features were evaluated. The acceleration directions were also used separately to study their different contribution to the classification result.

To have comparable results the experiments are conducted with identical preprocessing and parameter selection settings:

Again, the collected gait data was used for this experiment where the signals collected on the first day were used for training and the signals collected on the second data collection session were used for calculation of the performance. First of all the data was interpolated to 100 Hz. After the normalization to the mean the segmentation was carried out. The segment length t was 5000ms, the distance d between consecutive segments was 2500ms.

A parameter selection was done each time for $\gamma \in \{2^{-15}, 2^{-14}, ..., 2^4, 2^5\}$ and $C \in \{2^5, 2^6, ..., 2^{19}, 2^{20}\}$. In each parameter setting combination a cross validation was carried out for all 40 users as described in section 7.1.1.

7.2.1. Statistical Features

First, the features introduced in 6.2.1 on page 73 were evaluated: binned distribution (bin. dist.), maximum (max.), minimum (min.), mean (mean), standard deviation (std. dev.), mean absolute difference (diff.), root mean square (rms) and zero crossings (zero cross.).

Five test passes were done for each feature type using different data sources, namely the x-, y-, z-acceleration, the resulting acceleration, and all of them. The resulting vector is also referred to as the magnitude vector and is calculated as follows:

$$s_{res}(t) = \sqrt{s_x(t)^2 + s_y(t)^2 + s_z(t)^2}$$

Here $s_x(t), s_y(t), s_z(t)$ are the accelerations measured in the corresponding directions at time t.

The results are given in figure 7.1. The number behind bin distributions is the number of bins that were used for the binned distribution. All numbers in parentheses are the lengths of the feature vector. Of course when the features are combined the length is four times this number. The last column is the sum of FMR and FNMR for all directions combined. The background color is formatted subject to the sum error rate to allow an easier identification of the best performing feature types.

	X		Y		Z		Resultant Acc.		X, Y, Z and Res.		
Feature (Length)	*FMR*	*FNMR*	*FMR*	*FNMR*	*FMR*	*FNMR*	*FMR*	*FNMR*	*FMR*	*FNMR*	*FMR+FNMR*
bin. dist. (5)	1.36%	92.87%	3.10%	88.86%	0.49%	95.99%	3.80%	88.86%	2.77%	86.41%	89.18%
bin. dist. (10)	3.15%	86.41%	3.39%	85.30%	3.48%	89.31%	2.96%	89.09%	3.09%	83.30%	86.39%
bin. dist. (20)	2.40%	90.65%	3.50%	87.08%	2.52%	92.65%	3.71%	85.52%	2.07%	86.48%	88.56%
max. (1)	0.12%	98.66%	0.00%	100.00%	0.05%	97.55%	0.02%	98.22%	4.21%	86.41%	90.62%
min. (1)	0.08%	99.55%	0.01%	98.66%	0.12%	99.33%	0.00%	100.00%	3.90%	87.53%	91.43%
mean (1)	0.05%	99.55%	0.03%	99.55%	0.17%	99.11%	0.00%	99.55%	3.80%	90.42%	94.23%
diff. (1)	0.02%	99.33%	0.03%	99.78%	0.00%	100.00%	0.00%	100.00%	3.24%	94.43%	97.67%
std. dev. (1)	0.58%	97.33%	0.01%	99.11%	0.16%	99.78%	0.00%	99.11%	2.29%	74.16%	76.46%
rms (1)	0.01%	97.33%	0.01%	99.11%	0.14%	99.78%	0.00%	100.00%	2.88%	79.29%	82.16%
zero cross. (1)	0.00%	100.00%	0.20%	98.44%	0.00%	100.00%	0.00%	100.00%	3.72%	86.41%	90.13%

Table 7.1.: Discrimination capabilities of basic features

As can be observed the statistical features provide a quite limited verification. Furthermore the results yield no apparently best performing acceleration axis. The best performing feature type is the standard deviation followed by root mean square and the binned distribution.

In further experiments based on the statistical features different feature types were combined. For each feature type the feature was calculated for x-,y-,z-direction and the resulting acceleration. The feature vectors were as follows:

Set 1 (8) Std. dev, rms

Set 2 (48) Bin. dist. (10), std. dev., rms

Set 3 (52) Bin. dist. (10), std. dev., rms, zero cross.

Set 4 (64) Bin. dist. (10), max., min., mean, std. dev., rms, zero cross.

Set 5 (68) Bin. dist. (10), max., min., mean, std. dev., rms, zero cross., diff.

Again, the total feature vector length is provided in parenthesis. The results for these feature vector settings are provided in table 7.2.

	Used feature types	FMR	FNMR	FMR+FNMR
Set 1 (8)	Std. dev., rms	2.23%	75.50%	77.73%
Set 2 (48)	Bin. dist. (10), std. dev., rms	3.01%	75.50%	78.52%
Set 3 (60)	Bin. dist. (10), std. dev, rms, max., min., zero cross.	2.61%	77.73%	80.34%
Set 4 (64)	Bin. dist. (10), std. dev, rms, max., min., zero cross., mean	1.52%	78.17%	79.70%
Set 5 (68)	Bin. dist. (10), std. dev, rms, max., min., zero cross., mean, diff.	1.61%	78.17%	79.78%

Table 7.2.: Combined discrimination capabilities of basic features

Set 1 consists of the two best performing feature types of the previous experiment, the standard deviation and the root mean square. Interestingly this combination does not perform better than the standard deviation alone.

In set 2 the next better feature type, the binned distribution with ten bins, was added yielding a worse result. When the next better feature type, the zero crossings, were added the FNMR further increased significantly. Set 4 is built from all feature types except the bad performing mean absolute difference. In set 5 this feature type is also added. As can be seen the result is better when using the previously not good performing feature yielding a result that is worse than the two best performing features.

The conducted experiments clearly show that the statistical features alone are not capable of correctly classifying the respective genuine user.

7.2.2. Optimizing MFCC for Gait

As mentioned in 6.2.2 the MFCC calculation is done with the scripts published by Dan Ellis. As a starting point, a common MFCC implementation in speaker recognition was considered: the reproduction of the 13 Auditory Toolbox MFCC as described in [Ell]. The following list gives the most important settings:

- Window length: 0.016 sec [WL]
- Window hop time (distance): 0.01 sec [WHT]
- Sampling rate: 16000 Hz [SR]
- Minimum frequency: 133.33 Hz [MinF]
- Maximum frequency: 6855.6 Hz [MaxF]
- Pre-emphasis filter (0.97) [PRE]

- Number of warped spectral bands: 40 [NB]
- Number of cepstral features: 13 [NC]
- No cepstral liftering [CL]
- No equal-loudness weighting and cube-root compression [PLP]
- Integration of FFT bins into Mel bins as absolute values (not squared) [SP]

As these settings are changed a lot during the optimizations each experiment configuration will be provided in a short form in a footnote. Only the values that are constant throughout the sub experiment are provided. An example is given for the above described setting[1]. The preprocessing settings are also included as the first three values: interpolation rate, segment length and segment distance. A full reference for the parameter abbreviations can be found in appendix A with additional results.

Please note that a sampling rate of 16000 Hz was specified even though the gait acceleration signals were only interpolated to 100 Hz. This is not an issue as the original signal with a length of 5 seconds is then regarded as a signal of $\frac{5*100}{16000} = 0.03125$ seconds. The frequencies also change linearly. As mentioned in 6.2.2, 10 Hz is considered to be the highest frequency in the measured acceleration originating from walking. That would be regarded as a frequency of $\frac{10*16000}{100} = 1600$ Hz which is within the frequency range of 133.33 Hz to 6855.6 Hz that is used in this first experiment.

	X		Y		Z		Resultant Acc.		X, Y, Z and Res.		X and Y		
Feature (Length)	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR+FNMR
MFCC AT Emul. (13)	2.79%	73.50%	2.30%	69.49%	2.08%	82.85%	1.95%	81.51%	0.47%	67.48%	1.61%	62.36%	63.98%

Table 7.3.: MFCC performance with Auditory Toolbox settings

See table 7.3 for the results first obtained with these settings and a grid search parameter range of $\gamma \in \{2^{-15}, 2^{-14}, ..., 2^4, 2^5\}$ and $C \in \{2^5, 2^6, ..., 2^{19}, 2^{20}\}$.

Although at this stage no optimization of the MFCC settings was done this feature type is better performing than the basic statistical features discussed in the last section. With this setup the y-acceleration signal performs best

[1] IR=100, SL=5000, SD=2500, WL=0.016, WHT=0.01, SR=16000, MinF=133.33, MaxF=6855.6, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

whereas the z-acceleration signal performs worst, just like the resulting acceleration. This is probably the reason that the combined classification performance is only slightly better than the y-acceleration. Therefore only the x- and y-acceleration were used in the last experiment yielding a better result with 1.61% FMR and 62.36% FNMR.

This classification performance will be considered as the base line for the following optimizations. All results are based on the same preprocessing settings and grid search parameters for a better comparison. The optimizations are briefly discussed in the next sections.

MFCC Optimization 1: Frequency Mapping

As a first optimization the frequency range is remapped so that the relevant frequency range of 0 to 10 Hz is represented better by the MFCC. Obviously, the lowest band edge of mel filters is set to 0 Hz. As discussed, the setting for the highest band edge of mel filters is dependent on the used interpolation rate. Therefore, the maximum frequency is set to 1600 Hz.

To validate this approach three more frequency ranges were evaluated experimentally. The results are given in table 7.4[2], as expected, improved with a better frequency mapping.

		X		Y		Z		Resultant Acc.		X, Y, Z and Res.		X and Y		
#	Freq. Settings	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR+FNMR
1	0 - 1400 Hz	2.46%	69.71%	2.51%	63.03%	3.31%	78.40%	1.92%	81.29%	0.42%	65.48%	1.10%	59.69%	60.79%
2	0 - 1600 Hz	2.07%	69.49%	2.25%	68.82%	2.68%	77.28%	1.88%	79.29%	0.52%	68.60%	1.35%	57.91%	59.26%
3	0 - 1800 Hz	2.27%	65.96%	2.37%	69.56%	2.73%	78.44%	2.18%	81.18%	0.40%	68.50%	1.41%	58.14%	59.55%
4	0 - 2000 Hz	2.63%	69.27%	2.46%	63.92%	2.42%	80.40%	1.57%	79.96%	0.33%	69.49%	1.10%	65.48%	66.58%

Table 7.4.: MFCC frequency mapping optimization

The error rates yielded do not really differ from each other, therefore 1600 Hz will be used from now on as the maximum frequency for the highest band edge mel filter.

[2] IR=100, SL=5000, SD=2500, WL=0.016, WHT=0.01, SR=16000, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

MFCC Optimization 2: Spectrum Parameter

In the next step some basic variations of the above listed settings are used for experiments. One setting was modified at a time to have a better idea how the changes affect the classification performance. In the following list an overview is given on what was evaluated step per step:

1. Mel bins were squared: FFT bins were integrated into Mel bins using the squared domain
2. Pre-emphasis filter was disabled
3. Cepstral liftering with an exponent of 0.6: liftering means filtering cepstra, done by multiplying with a so called liftering exponent. A thorough discussion of liftering can be found in [JRW86]
4. HTK sinus cepstral liftering: liftering analog to implementation in Hidden Markov Toolkit

5. – 8. The number of warped spectral bands was decreased to 5/10/20/30 bands

9. Equal-loudness weighting and cube-root compression were applied. Techniques typically used in the perceptual linear predictive (PLP) analysis of speech, more information can be found in [Her90]
10. Number of MFCC decreased to 7
11. Number of MFCC increased to 19

		X		Y		Z		Resultant Acc.		X, Y, Z and Res.		X and Y		
#	Spectrum Settings	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR+FNMR
1	Mel bins squared	2.32%	66.81%	1.88%	70.82%	2.51%	78.22%	1.84%	78.65%	0.49%	68.92%	1.34%	58.80%	60.13%
2	No Pre-emphasis	2.55%	67.04%	2.36%	68.15%	2.88%	77.06%	1.97%	83.07%	0.37%	67.71%	1.36%	57.93%	59.29%
3	Cepstral Liftering	2.14%	69.27%	2.25%	68.82%	2.68%	77.28%	1.88%	79.29%	0.52%	68.60%	1.35%	57.91%	59.26%
4	HTK Sin. Cep. Liftering	2.13%	69.04%	2.25%	68.82%	2.68%	77.28%	1.88%	79.29%	0.52%	68.60%	1.35%	57.91%	59.26%
5	5 warp. spec. bands	1.66%	78.01%	3.47%	75.05%	2.71%	81.18%	2.01%	83.51%	1.18%	60.04%	2.33%	57.93%	60.26%
6	10 warp. spec. bands	2.85%	71.67%	2.31%	65.12%	2.97%	79.07%	1.99%	77.17%	0.60%	59.83%	1.59%	54.55%	56.13%
7	20 warp. spec. bands	2.82%	60.13%	2.29%	63.92%	2.49%	77.51%	1.96%	80.85%	0.45%	65.92%	1.19%	58.35%	59.54%
8	30 warp. spec. bands	2.55%	67.23%	2.63%	66.60%	2.20%	79.92%	1.87%	79.92%	0.35%	66.38%	0.86%	61.31%	62.17%
9	PLP weight./compr.	2.36%	68.29%	2.57%	67.65%	3.32%	77.17%	1.77%	80.55%	0.40%	69.77%	1.19%	60.89%	62.07%
10	7 MFCC	2.70%	78.17%	2.61%	67.26%	2.94%	77.73%	2.18%	80.85%	0.91%	63.47%	2.53%	60.13%	62.66%
11	19 MFCC	2.92%	62.36%	1.10%	73.72%	2.74%	74.16%	2.22%	77.06%	0.37%	66.82%	1.32%	60.80%	62.13%

Table 7.5.: MFCC spectrum parameter optimization

The results of the MFCC spectrum parameter optimization are presented in table 7.5[3]. As can be seen these experiments were not rewarded with significant decreased error rates, except for the case of the decreased number of warped spectral bands.

MFCC Optimization 3: Windows

The experiment in this section uses the same spectrum parameters as the section 7.2.2 on page 85 although one potential optimization was found. The idea is to see later whether the different optimizations can be combined successfully.

Until now the size of the windows within the MFCC implementation was nominal set to 0.016 seconds window length with 0.010 seconds hop time. Again, this does not represent the actual length of the gait signals as of the frequency multiplication because of the sampling rate of 16000 Hz at an original interpolation rate of 100 Hz. Therefore these numbers relate to $0.016 * \frac{16000}{100} = 2.56$ seconds and $0.010 * \frac{16000}{100} = 1.6$ seconds respectively.

Now it is evaluated if the classification errors can be reduced with adapting the window lengths and overlaps. Table 7.6[4] provides the result for the experiments with the window parameters. Improved error rates are marked with a green background. As previously, the window settings given in the table reflect to the actual implementation settings and are not the corresponding real time measure. To calculate the real window length and hop time in seconds the numbers have to be multiplied with $\frac{16000}{100}$ as shown above.

It is apparent that optimizing the MFCC window parameter is important for the classification outcome. The setting of 0.0090 window length combined with 0.0030 hop time is chosen for evaluating the combination with the previously found optimization.

[3] IR=100, SL=5000, SD=2500, WL=0.016, WHT=0.01, SR=16000, MinF=0, MaxF=1600

[4] IR=100, SL=5000, SD=2500, SR=16000, MinF=0, MaxF=1600, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

		X		Y		Z		Resultant Acc.		X, Y, Z and Res.		X and Y		
#	Window Settings Length/Hop	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR+FNMR
1	**0.0090 / 0.0001**	2.22%	58.99%	2.09%	68.92%	2.75%	74.84%	2.86%	80.34%	0.30%	63.64%	1.50%	56.66%	58.16%
2	**0.0090 / 0.0005**	3.06%	57.51%	2.78%	68.08%	3.14%	74.42%	2.95%	79.70%	0.35%	63.64%	1.24%	56.03%	57.26%
3	**0.0090 / 0.0010**	3.05%	56.45%	2.71%	66.60%	2.70%	76.32%	2.69%	80.13%	0.35%	62.58%	1.21%	54.97%	56.18%
4	**0.0090 / 0.0030**	2.82%	58.80%	2.55%	68.15%	3.10%	75.28%	2.42%	79.29%	0.39%	64.14%	1.46%	53.90%	55.36%
5	**0.0090 / 0.0060**	2.99%	61.95%	2.89%	70.40%	3.31%	78.01%	2.97%	78.86%	0.56%	65.33%	1.99%	53.49%	55.48%
6	**0.0090 / 0.0090**	2.23%	61.52%	2.83%	72.30%	2.09%	80.97%	2.42%	80.13%	0.45%	64.27%	1.88%	60.04%	61.92%
7	**0.0100 / 0.0001**	2.59%	60.04%	2.53%	68.92%	2.47%	74.00%	2.03%	80.76%	0.28%	63.00%	0.89%	56.45%	57.34%
8	**0.0100 / 0.0005**	2.67%	60.25%	2.44%	68.29%	2.41%	75.26%	2.33%	80.34%	0.31%	63.21%	1.34%	56.24%	57.57%
9	**0.0100 / 0.0010**	2.22%	62.36%	2.22%	65.48%	2.32%	74.16%	2.49%	80.18%	0.34%	62.58%	1.23%	55.46%	56.69%
10	**0.0100 / 0.0050**	2.37%	59.83%	2.63%	68.29%	2.89%	77.59%	2.06%	81.40%	0.38%	65.33%	1.61%	55.81%	57.43%
11	**0.0100 / 0.0100**	2.56%	63.42%	2.58%	70.19%	3.08%	80.97%	1.75%	81.40%	0.30%	66.38%	1.06%	60.25%	61.32%
12	**0.0120 / 0.0010**	2.68%	63.70%	2.18%	69.93%	3.48%	73.72%	2.09%	78.17%	0.30%	63.25%	1.29%	55.23%	56.52%
13	**0.0120 / 0.0030**	2.62%	64.37%	1.39%	71.27%	3.21%	72.83%	2.32%	77.28%	0.30%	65.03%	1.48%	56.35%	57.83%
14	**0.0120 / 0.0060**	2.85%	63.21%	1.73%	71.88%	2.99%	75.90%	1.85%	83.93%	0.39%	67.02%	1.23%	61.10%	62.33%
15	**0.0120 / 0.0120**	3.03%	67.23%	2.47%	72.94%	2.08%	80.76%	2.49%	82.66%	0.43%	66.60%	1.36%	60.47%	61.82%
16	**0.0160 / 0.0020**	2.48%	68.71%	2.13%	69.77%	2.59%	78.44%	1.91%	81.40%	0.28%	65.96%	0.71%	59.20%	59.90%
17	**0.0160 / 0.0040**	2.78%	65.03%	1.89%	69.93%	2.98%	77.06%	1.67%	81.29%	0.38%	66.37%	1.63%	54.57%	56.19%
18	**0.0160 / 0.0080**	2.75%	69.04%	2.54%	70.82%	2.32%	80.62%	1.67%	79.73%	0.60%	65.70%	1.27%	59.02%	60.29%
19	**0.0160 / 0.0160**	2.41%	75.48%	2.19%	77.17%	3.11%	78.65%	1.72%	84.57%	0.68%	66.17%	1.65%	66.60%	68.25%
20	**0.0200 / 0.0020**	3.02%	67.71%	1.89%	68.15%	2.65%	76.61%	1.88%	80.18%	0.39%	66.59%	0.78%	58.13%	58.91%
21	**0.0200 / 0.0050**	3.21%	62.37%	1.61%	69.56%	2.95%	77.80%	1.98%	78.22%	0.39%	65.33%	1.29%	57.51%	58.79%
22	**0.0200 / 0.0100**	3.01%	65.03%	1.46%	67.48%	2.97%	77.73%	2.15%	79.73%	0.27%	63.92%	1.70%	55.90%	57.60%
23	**0.0200 / 0.0200**	3.06%	67.04%	2.05%	70.16%	3.02%	77.95%	2.38%	85.30%	0.82%	64.37%	1.52%	59.24%	60.76%
24	**0.0320 / 0.0030**	2.79%	71.88%	1.85%	71.67%	2.66%	78.86%	1.47%	87.95%	0.38%	67.86%	0.96%	66.60%	67.56%
25	**0.0320 / 0.0050**	2.79%	72.61%	2.45%	69.71%	2.53%	77.95%	1.45%	87.97%	0.38%	67.48%	0.99%	65.48%	66.47%
26	**0.0320 / 0.0100**	2.79%	71.88%	1.85%	71.67%	2.66%	78.86%	1.47%	87.95%	0.38%	67.86%	0.96%	66.60%	67.56%
27	**0.0320 / 0.0160**	2.79%	71.88%	1.85%	71.67%	2.66%	78.86%	1.47%	87.95%	0.38%	67.86%	0.96%	66.60%	67.56%

Table 7.6.: MFCC window parameter optimization

MFCC Optimization Results and Combination

Improvements could be found with the following settings:

- Frequency range of 0 - 1600 Hz (equates to 0 - 10 Hz in real time domain)
- Usage of 10 warped spectral bands
- Windows of 0.0090 seconds with 0.0030 seconds hop time (equates to 1.44 seconds and 0.48 seconds in real time domain)

The frequency range was adapted for the experiments with the spectrum and window parameters. Now the optimum settings found in those experiments are combined.

The other settings[5] were constant. The results are presented in table 7.7.

#	Optimization	X		Y		Z		Resultant Acc.		X, Y, Z and Res.		X and Y		
		FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR	FNMR	FMR+FNMR
	Freq. 0 - 1600 Hz	2.07%	69.49%	2.25%	68.82%	2.68%	77.28%	1.88%	79.29%	0.52%	68.60%	1.35%	57.91%	59.26%
1	**10 warp. spec. bands**	2.85%	71.67%	2.31%	65.12%	2.97%	79.07%	1.99%	77.17%	0.60%	59.83%	1.59%	54.55%	56.13%
2	**0.0090/0.0030 wind.**	2.82%	58.80%	2.55%	68.15%	3.10%	75.28%	2.42%	79.29%	0.39%	64.14%	1.46%	53.90%	55.36%
3	**#1 and #2 combined**	3.40%	60.68%	2.68%	66.17%	3.40%	75.48%	2.42%	78.44%	0.51%	59.20%	1.73%	55.81%	57.54%

Table 7.7.: MFCC optimizations combined

When combined the error rates are unfortunately slightly worse. Therefore the window parameter optimization is considered the best approach for using MFCC for gait recognition and used later on.

7.2.3. Optimizing BFCC for Gait

Throughout all MFCC optimization experiments the best classification performances were yielded when combining features extracted from the x- and y-direction. Therefore only this combination is evaluated for the similar BFCC feature optimization that is described in this section.

The approach to this process is similar to the MFCC optimization. First, the best performing frequency range is selected. Afterwards an optimization of the spectrum parameter and of the window parameter is conducted separately. With the best results at hand the combination is tested. Here only the good performing settings are reported. The complete experiment results, analog arranged to the MFCC optimization, can be found in appendix A.

Again the starting point for the optimizations is the reproduction of the 13 Auditory Toolbox MFCC, except that the bark scale is used and thus BFCC are produced. The result for this configuration[6] is provided in table 7.8.

[5] IR=100, SL=5000, SD=2500, SR=16000, MinF=0, MaxF=1600, PRE=0.97, NC=13, CL=0, PLP=0, SP=0
[6] IR=100, SL=5000, SD=2500, WL=0.016, WHT=0.01, SR=16000, MinF=133.33, MaxF=6855.6, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

Feature (Length)	X and Y FMR	FNMR	FMR+FNMR
BFCC (AT Emul.) (13)	0.94%	55.81%	56.75%

Table 7.8.: BFCC performance analog to Auditory Toolbox MFCC

BFCC Optimization 1: Frequency Mapping

With the bark scale one can limit the frequency band to a smaller range as can be seen from the results in table 7.9. This was not possible to the full extent for the MFCC which is most likely due to an implementation issue of the MFCC implementation. Furthermore notable is that this setting[7] modification alone gives a better classification performance than the MFCC.

#	Freq. Settings	X and Y FMR	FNMR	FMR+FNMR
1	0 - 1000 Hz	1.94%	53.70%	55.64%
2	0 - 1100 Hz	2.04%	52.01%	54.05%
3	0 - 1200 Hz	1.68%	49.89%	51.57%
4	0 - 1300 Hz	1.51%	52.01%	53.51%
5	0 - 1400 Hz	1.45%	52.22%	53.67%
6	0 - 1500 Hz	1.16%	54.12%	55.29%

Table 7.9.: BFCC frequency mapping optimization

BFCC Optimization 2: Spectrum Parameter

The same spectrum parameter modification as described with the MFCC optimization was conducted. As these experiments did not improve the overall classification performance, the results are only provided in the appendix, see section A.1.1.

BFCC Optimization 3: Windows

Again, a large variety of experiments were conducted to find the best performing window parameters. The results presented in table 7.10 are only a selection

[7] IR=100, SL=5000, SD=2500, WL=0.016, WHT=0.01, SR=16000, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

of good performing settings[8].

All other outcomes can be found in section A.1.2 of the appendix.

#	Window Settings Length/Hop	X and Y FMR	FNMR	FMR+FNMR
1	**0.0060 / 0.0004**	1.99%	47.99%	49.98%
2	**0.0060 / 0.0010**	1.71%	45.88%	47.59%
3	**0.0070 / 0.0001**	1.71%	46.09%	47.80%
4	**0.0070 / 0.0002**	1.83%	44.40%	46.22%
5	**0.0070 / 0.0003**	2.16%	44.61%	46.77%
6	**0.0070 / 0.0004**	2.22%	44.82%	47.04%
7	**0.0070 / 0.0008**	1.74%	46.72%	48.47%
8	**0.0070 / 0.0010**	1.65%	46.51%	48.16%
9	**0.0070 / 0.0020**	1.67%	43.34%	45.01%
10	**0.0070 / 0.0030**	1.97%	47.99%	49.96%
11	**0.0080 / 0.0004**	1.57%	47.78%	49.35%
12	**0.0080 / 0.0010**	1.80%	47.99%	49.79%
13	**0.0080 / 0.0020**	1.69%	46.93%	48.63%
14	**0.0090 / 0.0005**	1.55%	47.36%	48.91%

Table 7.10.: BFCC window parameter optimization (selection)

The best performance was yielded with 0.0070 seconds window length and 0.0020 hop time. When looking at the results with similar hop times it looks like it is an unusually good result. For that reason it was decided to proceed with the 0.0070/0.0002 setting that also gave a reasonable improved performance.

BFCC Optimization Results

When starting with spectrum parameters similar to the reproduction of the 13 Auditory Toolbox MFCC, improvements could be found when a frequency range of 0 - 1200 Hz (equates to 0 - 7.5 Hz in real time domain) was combined with windows of 0.0070 seconds and 0.0002 seconds hop time (equates to 1.12 seconds and 0.032 seconds in real time domain).

[8] IR=100, SL=5000, SD=2500, SR=16000, MinF=0, MaxF=1200, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

7.2.4. Combining Features

In this section the previously tested and optimized features will be combined to evaluate if one can improve the performance by combining several good performing features.

The following list gives the tested combinations, the number in parenthesis is the total length of the feature vector.

Set 1 (30) MFCC, std dev.

Set 2 (30) BFCC, std dev.

Set 3 (52) MFCC, BFCC

Set 4 (56) MFCC, BFCC, std dev.

Set 5 (60) MFCC, BFCC, std dev., rms

Note that for MFCC and BFCC only the x- and the y-direction were considered whereas for the statistical features all directions plus the resulting acceleration were used. The results are presented in table 7.11.

	Used feature types	FMR	FNMR	FMR+FNMR
(26)	MFCC	1.46%	53.90%	55.36%
(26)	BFCC	1.83%	44.40%	46.22%
Set 1 (30)	MFCC, std. dev	1.10%	54.76%	55.86%
Set 2 (30)	BFCC, std. dev	1.92%	45.88%	47.80%
Set 3 (52)	MFCC, BFCC	1.24%	52.43%	53.67%
Set 4 (56)	MFCC, BFCC, std dev.	1.20%	52.01%	53.21%
Set 5 (60)	MFCC, BFCC, std dev., rms	1.32%	51.59%	52.90%

Table 7.11.: Optimized features combined

Again, the combination of features does not improve the classification performance. Especially the combination of the optimized MFCC and BFCC features was interesting, but unfortunately not rewarded with an increased classification performance.

This comparison concludes the feature selection process. Therefore, the optimized BFCC of the x- and y-direction will be solely used for the classification. The BFCC settings can be found in the footnote[9]. With the feature type selected one area remains to be optimized: the preprocessing.

[9] WL=0.0070, WHT=0.0002, SR=16000, MinF=0, MaxF=1200, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

7.3. Preprocessing Optimization

All experiments to this stage were performed with the same preprocessing settings: 100 Hz interpolation rate, 5000 ms segments with 2500 ms distance between consecutive segments.

Now it shall be evaluated how different preprocessing settings, e.g. longer segments and higher interpolation rates, affect the classification performance. Four different segment sizes, from 3000 to 10000 ms length, and four different interpolation rates, from 50 to 400 Hz, were used generating 16 possible combinations.

It is important to note that the previous optimizations are based on the interpolation of 100 Hz and using a nominal sampling rate of 16000 Hz for the cepstral features. When the interpolation rate is changed the nominal sampling rate has to be changed accordingly.

For example when an interpolation rate of 200 Hz is used in the preprocessing, the nominal rate has to be set to $\frac{200}{100} * 16000 = 32000$ Hz. Otherwise another nominal frequency range and window parameters need to be used which effectively means to repeat the previous optimization steps.

		Segment length / Segment distance											
		3000 / 1500			5000 / 2500			7500 / 3000			10000 / 5000		
		FMR	*FNMR*	*FMR+FNMR*	*FMR*	*FNMR*	*FMR+FNMR*	*FMR*	*FNMR*	*FMR+FNMR*	*FMR*	*FNMR*	*FMR+FNMR*
Interpolation Rate	50	2.50%	50.00%	52.50%	1.72%	46.09%	47.81%	1.64%	38.75%	40.39%	1.46%	39.38%	40.84%
	100	2.41%	50.62%	53.03%	1.83%	44.40%	46.22%	1.68%	40.00%	41.68%	1.51%	40.00%	41.51%
	200	2.37%	51.12%	53.49%	1.73%	45.03%	46.76%	1.65%	41.56%	43.22%	1.60%	40.00%	41.60%
	400	1.33%	55.58%	56.92%	1.48%	47.36%	48.84%	1.35%	43.44%	44.79%	1.66%	40.00%	41.66%

Table 7.12.: Preprocessing optimization

The results can be found in table 7.12. As previously, the sum of FMR and FNMR is also provided to allow an easier comparison of the overall performance. The grid search range for the SVM parameter selection was still $\gamma \in \{2^{-15}, 2^{-14}, ..., 2^{4}, 2^{5}\}$ and $C \in \{2^{5}, 2^{6}, ..., 2^{19}, 2^{20}\}$.

All settings with the table entry marked green do perform better than the previously found optimum. It is apparent that larger segment sizes increase the classification performance significantly and that smaller interpolation rates perform comparatively better.

The preprocessing setting with an interpolation rate of 50 Hz and 7500 ms segments is considered the best setting and therefore selected. In the following preprocessing and feature creation settings[10] will be constant for all experiments unless otherwise stated.

[10] IR=50, SL=7500, SD=3000, WL=0.0070, WHT=0.0002, SR=8000, MinF=0, MaxF=1200, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

8. Experiments and Results

With the selection and optimization process described in the previous chapter a recognition performance of 38.75% FNMR and 1.64% FMR could be achieved.

This chapter will focus on reviewing this result by conducting experiments that assess the performance with different training and testing data sets. Thereby an intra-day versus inter-day performance comparison can be done.

In the second section of this chapter a voting scheme is introduced and tested in order to improve the overall recognition performance.

8.1. Intra-Day and Inter-Day Performance

The influence of time is a non-trivial challenge for biometric gait recognition. Despite this fact many studies do not test their algorithms with gait data recorded on different days. To point out the flaw with this approach, a comparison of the intra-day and inter-day performance is presented in this section.

As discussed, the preceding feature selection and optimization was done with the collected gait data while using the data from the first data collection session for training and the data from the second day for testing. Therefore the optimization result is the inter-day result.

For the new experiments the grid search range for the C parameter was slightly modified to $C \in \{2^0, 2^1, ..., 2^{14}, 2^{15}\}$ based on the position of the previously found optimums which were in $\gamma \in \{2^{-15}, 2^{-14}, ..., 2^1, 2^0\}$ and $C \in \{2^1, 2^2, ..., 2^{11}, 2^{12}\}$.

On each session of the data collection two walks were collected, therefore we have four walks per subject. The walks are named as follows:

S1W1	First session, first walk
S1W2	First session, second walk
S2W1	Second session, first walk
S2W2	Second session, second walk

It is distinguished between three experiment types:

Cross-day	The case considered until now: data of the first session is used for training, data from the second session is used for testing
Same-day	A walk of one session is used for training, a different walk from the same session is used for testing
Mixed	Data from both sessions is used for training, different data from both sessions is used for testing

Table 8.1 presents the results for the tested combinations. Note that each training and testing set combination was also evaluated vice versa. The result of the reversed combination is always given directly after the other result and marked with the same background color.

In regard to the cross-day experiments it is striking that there are combinations were the recognition performance is much worse than the optimized setting. In the cases where data from the first day was used for training always a better performance is yielded in comparison to the reversed setting. This indicates that the optimization does not perfectly generalize.

That time is a major influence on the recognition performance is evident through the overall good results of the same-day experiments.

For the last group of experiments the data sets were mixed so that both the training and the testing data set contained gait data from both sessions. This basically is equivalent to a same-day comparison with more data and an increased in-day gait variety of the gait profiles. As can be seen from the results the algorithm performs in this scenario also reasonable well.

	Train	Test	FMR	FNMR	FMR+FNMR
Cross-day	S1W1, S1W2	S2W1, S2W2	1.64%	38.75%	40.39%
	S2W1, S2W2	S1W1, S1W2	1.07%	65.94%	67.01%
	S1W1	S2W1	1.38%	45.00%	46.38%
	S2W1	S1W1	1.32%	63.13%	64.44%
	S1W1	S2W2	1.45%	54.38%	55.82%
	S2W2	S1W1	1.51%	70.00%	71.51%
	S1W2	S2W1	1.10%	45.63%	46.73%
	S2W1	S1W2	1.55%	76.25%	77.80%
	S1W2	S2W2	1.69%	53.13%	54.82%
	S2W2	S1W2	0.66%	73.13%	73.78%
Same-day	S1W1	S1W2	1.07%	21.88%	22.94%
	S1W2	S1W1	2.11%	15.63%	17.73%
	S2W1	S2W2	0.64%	6.88%	7.52%
	S2W2	S2W1	0.76%	8.75%	9.51%
Mixed	S1W1, S2W1	S1W2, S2W2	1.15%	12.81%	13.96%
	S1W2, S2W2	S1W1, S2W1	1.19%	10.31%	11.50%
	S1W1, S2W2	S1W2, S2W1	1.32%	10.31%	11.63%
	S1W2, S2W1	S1W1, S2W2	1.65%	6.25%	7.90%

Table 8.1.: Intra- and inter-day performance comparison

At this stage the conclusion can be drawn, that the inter-day variety of human gait has a large influence on the recognition performance of the developed algorithm. Furthermore it can be established that error rates reported in gait recognition studies need to be linked to the exact experiment settings as it is not expected that the recognition performance is that stable with gait data collected on different days.

The results of the mixed-day setting lead to the question whether training with data from several days can improve the generalization. This can only be confirmed with a data set consisting of three or more days where the data from one day can be used solely for testing while more than one day is used for training. As such a data set has not been created yet the question remains to be answered.

All in all, especially in the more realistic cross-day tests, the recognition performance not sufficient as the FNMR is too high. At the same time, the FMR is comparable small. One can try to improve the overall performance by combining several classification results by a voting mechanism. The next section describes that approach in more detail.

8.2. Voting Scheme Performance

With the discussed recognition performance it is very likely that a genuine is incorrectly rejected. An imposter on the other side is comparable rarely wrong classified and thus seldom accepted as a genuine. An approach that would reduce the number of false rejects is to use multiple classifications for one recognition decision while incorporating a different confidence in the classification correctness.

More specifically one uses $\#V$ segments of a gait sample instead of only one segment for the recognition. For each segment the classification is carried out as usually. Then, the $\#V$ results are combined. An imaginable straightforward approach is majority voting, but it is not likely to perform well as the two error cases are so unevenly distributed. Therefore a *minority* voting for a genuine is implemented.

This means that at least $\#GV$ positive classification results are needed for an accept, otherwise the sample is rejected. $\#GV$ is set to a number in $[1; \lceil \frac{\#V}{2} \rceil]$. The described concept is visualized in figure 8.1. Note that $\#V_g$ is the number of results that classify the respective segment as a sample of the genuine, in other words $\#V_g$ is the number of *votes* for genuine.

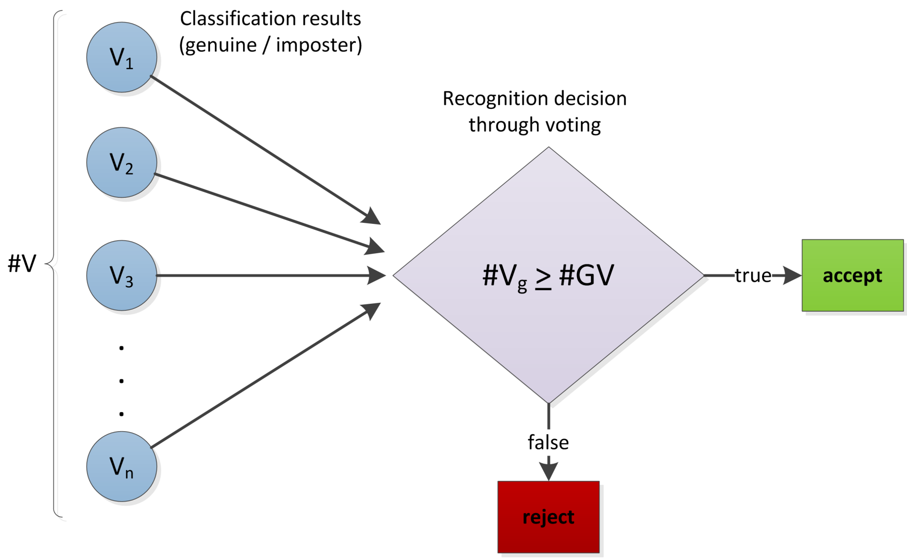

Figure 8.1.: Voting scheme

Of course, while the FNMR is decreased by this approach the number of false accepts and thus the FMR will increase. The following experiment aims to find a more balanced setting where both error rates are more even.

For the evaluation groups of segments are built that occur in the same walk consecutively. The motivation is to simulate the case were a signal is continuously captured and then a certain part of signal is used for preprocessing and feature extraction. The voting results will be reported separated by the three experiment types: cross-day, same-day and mixed.

8.2.1. Cross-Day Performance

The results given in table 8.2 are only for the case where both sessions of a day were used for training and testing.

	Train	S1W1, S1W2		
	Test	S2W1, S2W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	1.64%	38.75%	40.39%
2	1	2.09%	32.50%	34.59%
3	1	2.53%	26.25%	28.78%
3	2	1.74%	41.25%	42.99%
4	1	2.96%	30.00%	32.96%
4	2	1.71%	32.50%	34.21%
5	1	3.88%	27.50%	31.38%
5	2	1.18%	30.00%	31.18%
5	3	1.51%	35.00%	36.51%
6	1	4.01%	22.50%	26.51%
6	2	2.30%	25.00%	27.30%
6	3	1.05%	32.50%	33.55%
7	1	3.88%	25.00%	28.88%
7	2	3.16%	27.50%	30.66%
7	3	0.99%	32.50%	33.49%
7	4	0.92%	35.00%	35.92%
8	1	4.93%	22.50%	27.43%
8	2	2.57%	27.50%	30.07%
8	3	1.91%	30.00%	31.91%
8	4	1.58%	30.00%	31.58%

	Train	S2W1, S2W2		
	Test	S1W1, S1W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	1.07%	65.94%	67.01%
2	1	1.37%	61.88%	63.24%
3	1	1.97%	52.50%	54.47%
3	2	0.69%	68.75%	69.44%
4	1	1.68%	53.75%	55.43%
4	2	0.82%	58.75%	59.57%
5	1	1.97%	47.50%	49.47%
5	2	2.04%	55.00%	57.04%
5	3	0.59%	65.00%	65.59%
6	1	1.97%	42.50%	44.47%
6	2	1.78%	50.00%	51.78%
6	3	1.05%	57.50%	58.55%
7	1	3.42%	40.00%	43.42%
7	2	1.51%	47.50%	49.01%
7	3	1.12%	50.00%	51.12%
7	4	0.59%	62.50%	63.09%
8	1	2.96%	37.50%	40.46%
8	2	1.32%	42.50%	43.82%
8	3	0.86%	50.00%	50.86%
8	4	0.72%	57.50%	58.22%

Table 8.2.: Inter-day performance with voting

The first line, with $\#V$ and $\#GV$ both 1, the result equates of course to the results reported in table 8.1 on page 99. For the complete cross-day voting results see appendix A.2.1.

As the available data is limited, not any number of $\#V$ can be used. For the data set, when using segments of 7500 ms length and 3000 ms distance, eight segments are available for training and eight segments are available for testing.

This is based on the fact, that each walk is max. 18 seconds long. As described in 5.2.1 on page 65 for some walks were parts shorter than 2 seconds dropped during the walk extraction. Therefore, for each walk $\lfloor (18 - 7.5)/3 + 1 \rfloor = 4$ segments can be extracted. As two walks are used for both training in testing in this scenario a total of 8 segments are available.

Like assumed, in general the FNMR can be decreased while the FMR increases. For some settings the FNMR is slightly bigger even though more number of votes $\#V$ were considered, this is due to the grid search optimization that finds a different optimum because of the small changes in the training data.

The FNMR+FMR sum is reduced significantly when $\#V$ is increased while keeping $\#GV$ as low as one or two. For the optimized case where the walks from the first day were used for training, and the ones from the second day for testing, a FMR+FNMR sum of 26.51% was reached.

In the case when the walks from the second day were used for training a very poor performance was previously achieved with a FMR+FNMR sum of 65.94%. Thanks to the voting approach this sum could be reduced to 40.46%.

8.2.2. Same-Day Performance

Again, this section gives only a selection of the results for the same-day voting performance, additional results can be found in appendix A.2.2.

Table 8.3 provides the voting-based result for the same day experiment having the worst performance in section 8.1: using the two walks from the first data collection day. As only one walk of a session is used for training and the other walk is used for testing, a maximum of 4 segments is available for a voting decision.

The FNMR decreases significantly from 21.88% to 12.50% and 15.63% to 7.50% respectively. Of course these results are only of limited validity as the data set size is not sufficiently large. More specifically, in the case of 4 votes per recognition decision all available testing data is needed and therefore only one decision is made for each person combination of the cross validation.

	Train	S1W1		
	Test	S1W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	1.07%	21.88%	22.94%
2	1	0.92%	17.50%	18.42%
3	1	1.38%	12.50%	13.88%
3	2	1.45%	20.00%	21.45%
4	1	1.12%	12.50%	13.62%
4	2	0.79%	15.00%	15.79%

	Train	S1W2		
	Test	S1W1		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	2.11%	15.63%	17.73%
2	1	1.61%	10.00%	11.61%
3	1	1.91%	5.00%	6.91%
3	2	1.05%	12.50%	13.55%
4	1	1.45%	7.50%	8.95%
4	2	1.05%	10.00%	11.05%

Table 8.3.: Intra-day performance with voting

8.2.3. Mixed Performance

Table 8.4 on the following page provides the voting-based result for the mixed experiment setting having the worst performance in section 8.1: combining S1W1, S2W1 and S1W2, S2W2. Additional results can be found in appendix A.2.3.

The FNMR can be brought down to even 0.00% with a reasonable low FMR. Again it should be noted that the error rates are of restricted validity because of the size of the data set. Nevertheless it provides an impression of the magnitude of performance that can be achieved.

8.2.4. Voting Summary

As a consequence of the results one can assume that a voting based recognition can achieve an even better performance when more gait data is taken into consideration for the decision. As the developed gait recognition concept is intended to be a passive, continuously authentication system there is virtually no restriction of the time period captured and used for gait recognition. Of course the number of required genuine votes has to be chosen reasonable, so that the FMR does not become too high.

The reported error rates are based on less than 36 seconds of gait data used

Train		S1W1, S2W1			Train		S1W2, S2W2		
Test		S1W2, S2W2			Test		S1W1, S2W1		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*	#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.15%	12.81%	13.96%	1	1	1.19%	10.31%	11.50%
2	1	1.89%	10.00%	11.89%	2	1	1.51%	4.38%	5.89%
3	1	1.18%	8.75%	9.93%	3	1	3.36%	2.50%	5.86%
3	2	0.99%	15.00%	15.99%	3	2	1.38%	10.00%	11.38%
4	1	3.32%	3.75%	7.07%	4	1	2.01%	0.00%	2.01%
4	2	0.76%	8.75%	9.51%	4	2	1.58%	3.75%	5.33%
5	1	1.51%	5.00%	6.51%	5	1	2.83%	0.00%	2.83%
5	2	1.91%	10.00%	11.91%	5	2	1.51%	0.00%	1.51%
5	3	0.59%	15.00%	15.59%	5	3	1.25%	7.50%	8.75%
6	1	1.84%	2.50%	4.34%	6	1	2.89%	0.00%	2.89%
6	2	1.05%	5.00%	6.05%	6	2	1.71%	0.00%	1.71%
6	3	0.72%	12.50%	13.22%	6	3	0.99%	2.50%	3.49%
7	1	2.24%	0.00%	2.24%	7	1	0.59%	0.00%	0.59%
7	2	0.99%	2.50%	3.49%	7	2	0.86%	0.00%	0.86%
7	3	0.66%	2.50%	3.16%	7	3	1.18%	0.00%	1.18%
7	4	0.72%	12.50%	13.22%	7	4	0.66%	2.50%	3.16%
8	1	0.66%	0.00%	0.66%	8	1	0.66%	0.00%	0.66%
8	2	0.39%	0.00%	0.39%	8	2	0.99%	0.00%	0.99%
8	3	0.07%	2.50%	2.57%	8	3	0.53%	0.00%	0.53%
8	4	0.72%	2.50%	3.22%	8	4	1.05%	0.00%	1.05%

Table 8.4.: Mixed intra-day performance with voting

for the recognition. Imaginable is an authentication system were always the last minute of acceleration data is temporarily stored and used for verification.

The results of the mixed experiment look very promising, but need to be verified with a data set that was collected on more than two days so that walks of a previously unseen day can be used for testing.

8.3. Data Set Performance Comparison

As already mentioned were experiments also carried out with another data set. This is the same data set that was also used in [DNB10] and [NBR11]. To distinguish between the data sets the following naming is introduced:

Data set A The data set of 40 persons that was collected as described in 5.

Data set B The data set that was also used in [DNB10] and [NBR11]. Gait samples of each subject were collected on two sessions with several days in between. Data of 48 subjects could be used, the walks of three people were dropped due to insufficient data quality.

The motivation was to compare the presented approach with truly similar data. Similar because in the two mentioned studies no strict day to day separation of the used data is given and therefore still different training (i.e. enrolment) and verification data sets are used.

To yield optimal results an optimization was carried out for data set B. Therefore, the steps described in chapter 7 were repeated for data set B. In short the process was as follows:

1. Determine discriminative power of statistical features separated by acceleration direction and feature type
2. Evaluate combined directions for each statistical feature type
3. Combine statistical feature types to find features that can be successfully used together
4. Optimize MFCC
   a) Determine best frequency range
   b) Find best spectrum parameters
   c) Optimize window size and hop time
   d) Combine found optimizations

5. Repeat step 4 a) - d) for BFCC
6. Evaluate combinations of optimized MFCC, BFCC and statistical features to get optimum feature selection
7. Optimize preprocessing settings for optimum feature selection

The found optimal feature settings are provided in table 8.5, the preprocessing settings along with the BFCC settings are also included. The abbreviations of the BFCC settings are still the ones previously used and described at the beginning of appendix A.

	Data Set A		Data Set B	
Feature Vector Length	26		96	
Feature Types	BFCC	[x,y]	BFCC	[x,y,z,res]
			bin. dist. (5)	[x,y,z,res]
			mean	[x,y,z,res]
			max.	[x,y,z,res]
			min.	[x,y,z,res]
			rms	[x,y,z,res]
			std. dev.	[x,y,z,res]
			zero cross.	[x,y,z,res]
BFCC Parameter	WL	0.007	WL	0.012
	WHT	0.0002	WHT	0.003
	SR	8000	SR	16000
	MinF	0	MinF	0
	MaxF	1200	MaxF	1400
	PRE	0.97	PRE	0.97
	NB	40	NB	40
	NC	13	NC	13
	CL	0	CL	0
	PLP	0	PLP	0
	SP	0	SP	0
Interpolation Rate	50		100	
Segment Length	7500		7500	
Segment Distance	3000		3000	

Table 8.5.: Comparison of optimum features of data set A and B

It is not surprising that the found optimum settings differ from the settings that were found for data set A. The simple reason is, that the data collection of data set B was done with a different device with a significant lower mean sampling rate of ∼40 Hz compared to the mean sampling rate of ∼125 Hz of the data collection described in chapter 5. Naturally the different data collection locations and involved subjects cause different optimum settings.

Table 8.6 provides the comparison of the developed gait recognition approach with the two other studies reporting performances based on data set B.

Study	FMR	FNMR	FMR+FNMR
[DNB10]	20.1%	20.1%	40.2%
[NBR11]	9.31%	10.42%	19.73%
This study (mixed)	0.91%	2.08%	2.99%

Table 8.6.: Performance comparison with data set B

As provided in the table, the mixed intra-day experiment setting was used as it is closest to the data set utilization in [NBR11]. The proposed approach using non-cycle-based features and SVMs clearly outperforms the cycle extraction based approach in [DNB10] as well as application of Hidden Markov Models in [NBR11].

8.4. Experiments and Performance Summary

In this section the performance of state the art WS-based gait recognition is revisited and compared with the results of this thesis for data set A.

The inter-day performance can be compared with the results in [ALM+05] as here a strict day-wise separation of the data used for the template creation and the data used for the cross validation tests is given. Luckily the study reports beside the EER also the total error rate (TER) which is the sum of the false accept rate and false reject rate. Therefore it is comparable to the sum of FMR and FNMR. Find the results in table 8.7.

Study	Sensor location	#S	EER	TER	TG
[ALM+05]	Waist	36	6.4%	11.9%	5 days
This study	Hip	40	—	26.51%	2 - 3 days

Table 8.7.: Inter-day performance comparison

Although the comparability is limited, it is apparent that the results reported in [ALM+05] are superior to the proposed method. It would be worthwhile to compare the methods for a common data set to verify this.

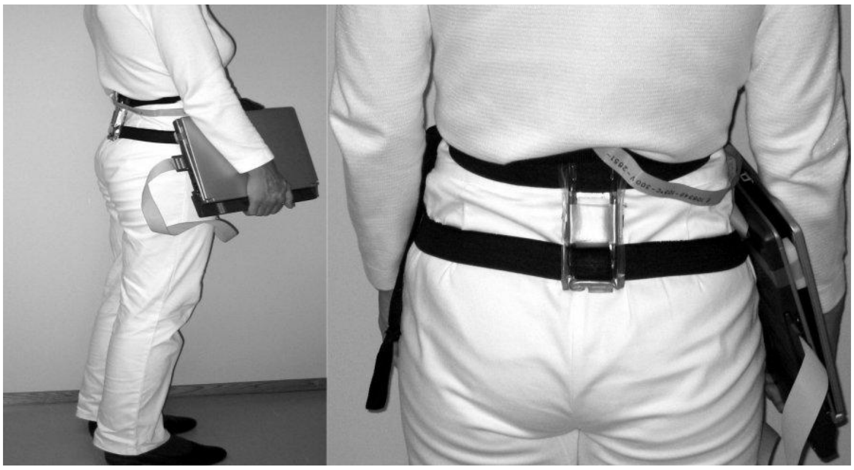

Figure 8.2.: Accelerometer worn on a belt in [ALM+05]

The reason is that one potential reason for the significant better performance could be the used sensor. In this study, acceleration data with a mean sampling rate of about 125 Hz was available. In [ALM+05] a sensor with 256 Hz sampling rate was used. Besides the double sampling rate another sensor-related advantage is the fixed sampling rate which is not given with the Android operating system.

A further influence on the error rates is possibly the position and fixation of the sensor in [ALM+05]: the sensor position is shown in figure 8.2. It is obvious, that the data collection setup of data set A is closer to a realistic setup as a normal phone pouch worn on the hip was used. An image showing the position of the pouch is provided with figure 5.2 on page 62.

Although it is not clear whether this has a significant influence or not, it should be noted that the subjects of data set A were comparable young as nearly half of them were underage. The data set used in [ALM+05] was created with only adult participants.

The picture of an unsatisfactory performance does not hold for the intra-day performance when compared to other recent gait recognition studies. The studies given in table 8.8 only provide the EER. To compare it with the results of this study one can approximate the TER with multiplying the EER by two [Toh07]. This would give for a 1.6% EER an approximately TER of 3.2% which is significantly worse than the 0.39% TER of this study.

Study	Sensor location	#S	EER	TER
[DBH10]	Hip	60	5.7%	—
[BS10]	Hip	60	1.6%	—
[GSB10]	Ankle	30	1.6%	—
This study	Hip	40	—	0.39%

Table 8.8.: Intra-day performance comparison

It is quite clear at this stage that the performance of WS-based gait recognition is, largely influenced from day to day variability of the human gait. Thus, future studies should also probe inter-day performance to provide information on the stability regarding gait variability.

Except of [NBR11], all the aforementioned studies use, in opposition to this thesis, cycle-based representations of gait. From the presented results can be concluded that also non-cycle-based approaches, like the discussed segments-based one, are capable of producing competitive performing gait representations.

9. Conclusion

9.1. Summary

Biometric gait recognition using wearable sensors like accelerometers is a young research area with promising future applications. Although the biometric performance is still far from being competitive compared to more established biometric systems, such as fingerprint or face recognition, it offers the unique advantage of a truly unobtrusive sample capturing. Therefore gait recognition could establish itself as an additional authentication method in environments where convenience is a critical factor. The usage as alternative to other authentication methods is also imaginable. As previously mentioned an application scenario could be a multi-biometrical system using gait recognition where the primary authentication method, for example face recognition, is not available due to environmental influences like lighting conditions.

Accelerometer-based gait recognition is especially interesting for mobile devices that today commonly come with the needed sensor already. The introduction of a new biometric system which costs are limited to the development of the software is very promising and could lead to a wide adoption across different devices, vendors and operating systems. A passive, lazy authentication approach might be the key for establishing a better information security on such devices.

This thesis presents an approach to biometric gait recognition using the built-in sensor of a modern smartphone. A data collection of the acceleration characteristics was carried out using the phone enabling a realistic optimization of the developed algorithm. Support vector machines were used for recognizing the respective genuine. As probably the first gait recognition study, cepstral features like MFCC and BFCC were used to represent the characteristics of human gait yielding a competitive performance.

With future applications in mind the experiments in this study carefully distinguish between the recognition performance on the same day and the recognition performance cross-days. That this is necessary is proven by the very good results with same-day tests compared to a rather weak performance cross-day tests. Unfortunately this has not been sufficiently addressed in the most of the other WS-based gait recognition studies. Therefore it is questionable if the reported performances were capable of providing a good cross-day performance.

In general it is also hardly possible to compare different studies because of the large differences that begin with different methodologies at the data collection. One important step for biometric WS-based gait recognition research as a whole is yet to be done: the collection and publication of a large gait sample database that can be used to compare different gait recognition approaches.

9.2. Outlook

There are numerous facets of WS-based gait recognition that were not adequately studied as of today, but are potentially of critical importance for a practical application. Current studies intentionally limit the influence of activity variations, environmental differences and health condition changes. As each of these influences imply numerous challenges, a strategy has to be developed to stepwise solve the issues and bring WS-based gait recognition closer to a practical usage in real-life conditions.

In regard to the work carried out for this thesis the next task will be the porting of the developed program to the Android platform. Further algorithms have to be developed that introduce context awareness through activity recognition to, for example, behave in a certain way during non-movement phases.

In addition, algorithms have to be developed for conducting an efficient live quality ensurance of the measured acceleration data. The motivation is to identify data with a low sampling rate, or time lags of several hundred milliseconds. That way data with insufficient quality can be dropped before it would be used for verification.

One limitation of the current approach is that the training of the SVM requires feature vector instances of imposters. An implication is that a complete product would need to ship with a database with the feature instances. Although this is not impractical, from a privacy point of view, it would be preferable to have a pre-trained SVM that is only further trained with the instances of the genuine user during enrolment. Therefore incremental SVM learning approaches (e.g. [WZC08]) should be evaluated as well.

Another issue that needs to be addressed is whether the parameter selection, i.e. the grid search, needs to be a part of the application on the device or one parameter selection would work reasonable well for all future enrollees.

These enhancements will lead to a self-contained biometric gait recognition system for mobile devices. That allows one to work on the mentioned challenges and to verify the progress with realistic on-device tests.

To improve the recognition performance further other features can be compared and combined with the BFCC that were selected in this study. Furthermore one should evaluate cepstral features in cycle-based recognition approaches. As these sensors also become more common the usage of the orientation information of a gyroscope sensor might help to mitigate the influence of the changing position of the device in question.

Another open question is whether the training of the SVMs can yield a better generalization in terms of a tolerance against inter-day gait variability when data from several days is used for training. To answer this question a multi-day gait database is needed.

A. Additional Results

Feature Settings Abbreviations

CL Cepstral liftering exponent

IR Interpolation rate (Hz)

MaxF Maximum frequency of highest band edge of scale filter (Hz)

MinF Minimum frequency of lowest band edge of scale filter (Hz)

NB Number of warped spectral bands

NC Number of cepstra coefficients

PLP Similar to PLP: Equal-loudness weighting and cube-root compression (0 = No, 1 = Yes)

PRE Pre-emphasis filter (0 = No)

SD Segment distance (ms)

SL Segment length (ms)

SP Integration of FFT bins into Mel bins are squared (0 = No, 1 = Yes)

WHT Window hop time (sec)

WL Window length (sec)

A.1. BFCC Optimization

A.1.1. Spectrum Parameter

Experiment settings: IR=100, SL=5000, SD=2500, WL=0.016, WHT=0.01, SR=16000, MinF=0, MaxF=1200

Table A.1 provides the experiment results.

		X and Y		
#	Spectrum Settings	*FMR*	*FNMR*	*FMR+FNMR*
1	**Mel bins squared**	1.97%	49.89%	51.86%
2	**No Pre-emphasis**	1.64%	54.33%	55.97%
3	**Cepstral Liftering**	1.61%	50.95%	52.56%
4	**HTK Sin. Cep. Liftering**	1.68%	49.89%	51.57%
5	**5 warp. spec. bands**	2.70%	63.42%	66.12%
6	**10 warp. spec. bands**	1.33%	55.60%	56.93%
7	**20 warp. spec. bands**	1.77%	52.64%	54.41%
8	**30 warp. spec. bands**	1.68%	50.95%	52.63%
9	**50 warp. spec. bands**	1.69%	50.32%	52.01%
10	**70 warp. spec. bands**	1.72%	50.74%	52.46%
11	**100 warp. spec. bands**	1.71%	50.74%	52.45%
12	**150 warp. spec. bands**	1.72%	50.74%	52.46%
13	**PLP weight./compr.**	1.91%	50.32%	52.23%
14	**7 BFCC**	1.85%	53.28%	55.13%
15	**19 BFCC**	1.69%	55.39%	57.09%

Table A.1.: BFCC spectrum parameter optimization

A.1.2. Windows

Experiment settings: IR=100, SL=5000, SD=2500, SR=16000, MinF=0, MaxF=1200, PRE=0.97, NB=40, NC=13, CL=0, PLP=0, SP=0

Table A.2provide the experiment results. A selection of these results was presented in 7.2.3 in table 7.10 on page 93.

#	Window Settings Length/Hop	X and Y			#	Window Settings Length/Hop	X and Y		
		FMR	FNMR	FMR+FNMR			FMR	FNMR	FMR+FNMR
1	**0.0050 / 0.0001**	2.35%	54.33%	56.68%	26	**0.0080 / 0.0020**	1.69%	46.93%	48.63%
2	**0.0050 / 0.0002**	2.17%	53.70%	55.87%	27	**0.0080 / 0.0040**	1.68%	49.05%	50.73%
3	**0.0050 / 0.0004**	2.30%	54.76%	57.06%	28	**0.0090 / 0.0001**	1.42%	49.26%	50.68%
4	**0.0050 / 0.0010**	1.37%	56.03%	57.40%	29	**0.0090 / 0.0005**	1.55%	47.36%	48.91%
5	**0.0050 / 0.0020**	2.07%	54.55%	56.61%	30	**0.0090 / 0.0010**	1.82%	48.84%	50.66%
6	**0.0050 / 0.0040**	2.18%	58.56%	60.74%	31	**0.0090 / 0.0030**	2.11%	49.47%	51.58%
7	**0.0060 / 0.0001**	1.92%	49.26%	51.18%	32	**0.0090 / 0.0060**	2.44%	54.33%	56.78%
8	**0.0060 / 0.0002**	2.00%	48.20%	50.20%	33	**0.0090 / 0.0090**	2.06%	54.12%	56.19%
9	**0.0060 / 0.0004**	1.99%	47.99%	49.98%	34	**0.0100 / 0.0001**	1.84%	49.47%	51.31%
10	**0.0060 / 0.0010**	1.71%	45.88%	47.59%	35	**0.0100 / 0.0005**	1.71%	48.41%	50.12%
11	**0.0060 / 0.0020**	1.48%	49.68%	51.17%	36	**0.0100 / 0.0010**	1.71%	48.41%	50.12%
12	**0.0060 / 0.0040**	1.99%	54.55%	56.53%	37	**0.0100 / 0.0050**	1.82%	48.20%	50.02%
13	**0.0070 / 0.0001**	1.71%	46.09%	47.80%	38	**0.0100 / 0.0100**	2.11%	53.07%	55.17%
14	**0.0070 / 0.0002**	1.83%	44.40%	46.22%	39	**0.0120 / 0.0010**	1.80%	52.22%	54.02%
15	**0.0070 / 0.0003**	2.16%	44.61%	46.77%	40	**0.0120 / 0.0030**	1.45%	52.43%	53.88%
16	**0.0070 / 0.0004**	2.22%	44.82%	47.04%	41	**0.0120 / 0.0060**	1.95%	52.22%	54.17%
17	**0.0070 / 0.0008**	1.74%	46.72%	48.47%	42	**0.0120 / 0.0120**	2.07%	54.12%	56.20%
18	**0.0070 / 0.0010**	1.65%	46.51%	48.16%	43	**0.0160 / 0.0020**	1.43%	54.55%	55.98%
19	**0.0070 / 0.0020**	1.67%	43.34%	45.01%	44	**0.0160 / 0.0040**	1.63%	51.80%	53.43%
20	**0.0070 / 0.0030**	1.97%	47.99%	49.96%	45	**0.0160 / 0.0080**	1.93%	52.01%	53.94%
21	**0.0070 / 0.0040**	1.88%	49.68%	51.56%	46	**0.0160 / 0.0160**	1.55%	60.04%	61.59%
22	**0.0080 / 0.0001**	1.32%	49.47%	50.79%	47	**0.0200 / 0.0020**	2.07%	54.12%	56.20%
23	**0.0080 / 0.0002**	1.65%	48.41%	50.06%	48	**0.0200 / 0.0050**	2.32%	54.55%	56.86%
24	**0.0080 / 0.0004**	1.57%	47.78%	49.35%	49	**0.0200 / 0.0100**	1.74%	52.64%	54.39%
25	**0.0080 / 0.0010**	1.80%	47.99%	49.79%	50	**0.0200 / 0.0200**	2.52%	54.76%	57.27%

Table A.2.: BFCC window parameter optimization

A.2. Voting

The walks are named subject to the day and order of the data collection. The naming scheme is as follows:

S1W1 First session, first walk

S1W2 First session, second walk

S2W1 Second session, first walk

S2W2 Second session, second walk

A.2.1. Cross-day results

For this experiment data of the first session is used for training, data from the second session is used for testing. The experiments provided in this section only use one walk for training and one for testing. The results of the cross-day experiments were both walks of a session were used are provided in table 8.2 in section 8.2.1.

	Train	S1W1		
	Test	S2W1		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.38%	45.00%	46.38%
2	1	1.74%	40.00%	41.74%
3	1	1.51%	45.00%	46.51%
3	2	1.05%	52.50%	53.55%
4	1	1.71%	35.00%	36.71%
4	2	1.32%	42.50%	43.82%

	Train	S2W1		
	Test	S1W1		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.32%	63.13%	64.45%
2	1	2.14%	56.25%	58.39%
3	1	2.43%	55.00%	57.43%
3	2	1.78%	62.50%	64.28%
4	1	1.71%	47.50%	49.21%
4	2	1.38%	60.00%	61.38%

Table A.3.: Additional cross-day voting results 1/4

	Train	S1W1		
	Test	S2W2		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.45%	54.38%	55.82%
2	1	1.51%	48.75%	50.26%
3	1	2.50%	42.50%	45.00%
3	2	1.64%	55.00%	56.64%
4	1	1.64%	40.00%	41.64%
4	2	1.32%	45.00%	46.32%

	Train	S2W2		
	Test	S1W1		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.51%	70.00%	71.51%
2	1	1.88%	63.75%	65.63%
3	1	1.45%	55.00%	56.45%
3	2	1.12%	65.00%	66.12%
4	1	2.43%	57.50%	59.93%
4	2	1.78%	65.00%	66.78%

Table A.4.: Additional cross-day voting results 2/4

Train		S1W2		
Test		S2W1		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	1.10%	45.63%	46.73%
2	1	1.32%	38.75%	40.07%
3	1	1.78%	40.00%	41.78%
3	2	1.25%	47.50%	48.75%
4	1	1.64%	35.00%	36.64%
4	2	1.32%	40.00%	41.32%

Train		S2W1		
Test		S1W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	1.55%	76.25%	77.80%
2	1	1.94%	71.25%	73.19%
3	1	2.17%	65.00%	67.17%
3	2	1.58%	80.00%	81.58%
4	1	2.63%	62.50%	65.13%
4	2	1.58%	72.50%	74.08%

Table A.5.: Additional cross-day voting results 3/4

Train		S1W2		
Test		S2W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	1.69%	53.13%	54.82%
2	1	2.01%	50.00%	52.01%
3	1	2.43%	45.00%	47.43%
3	2	1.25%	52.50%	53.75%
4	1	2.17%	42.50%	44.67%
4	2	1.97%	50.00%	51.97%

Train		S2W2		
Test		S1W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	0.66%	73.13%	73.78%
2	1	0.92%	68.75%	69.67%
3	1	0.99%	67.50%	68.49%
3	2	1.18%	72.50%	73.68%
4	1	1.12%	62.50%	63.62%
4	2	0.39%	67.50%	67.89%

Table A.6.: Additional cross-day voting result 4/4

A.2.2. Same-day results

In this experiments a walk of one session is used for training, a different walk from the same session is used for testing. Here are the results using the walks of the second session. The results that were obtained using the walks from the first session were presented in table 8.3 in section 8.2.2.

Train		S2W1		
Test		S2W2		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	0.64%	6.88%	7.52%
2	1	0.30%	3.75%	4.05%
3	1	0.59%	0.00%	0.59%
3	2	0.39%	2.50%	2.89%
4	1	0.39%	0.00%	0.39%
4	2	0.20%	0.00%	0.20%

Train		S2W2		
Test		S2W1		
#V	#GV	FMR	FNMR	FMR+FNMR
1	1	0.76%	8.75%	9.51%
2	1	1.45%	3.75%	5.20%
3	1	0.72%	5.00%	5.72%
3	2	0.33%	12.50%	12.83%
4	1	2.11%	0.00%	2.11%
4	2	0.13%	5.00%	5.13%

Table A.7.: Additional same-day voting results

A.2.3. Mixed results

For these experiments data from both sessions is used for training while different data from both sessions is used for testing. The results of another combination are provided in table 8.4 in section 8.2.3.

	Train	S1W1, S2W2		
	Test	S1W2, S2W1		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.32%	10.31%	11.63%
2	1	1.74%	6.88%	8.62%
3	1	1.64%	6.25%	7.89%
3	2	1.91%	12.50%	14.41%
4	1	2.11%	3.75%	5.86%
4	2	1.18%	6.25%	7.43%
5	1	2.70%	0.00%	2.70%
5	2	2.76%	7.50%	10.26%
5	3	1.18%	12.50%	13.68%
6	1	1.84%	0.00%	1.84%
6	2	1.05%	0.00%	1.05%
6	3	0.39%	7.50%	7.89%
7	1	0.92%	0.00%	0.92%
7	2	1.64%	0.00%	1.64%
7	3	0.86%	0.00%	0.86%
7	4	0.39%	7.50%	7.89%
8	1	0.86%	0.00%	0.86%
8	2	0.92%	0.00%	0.92%
8	3	0.79%	0.00%	0.79%
8	4	0.79%	0.00%	0.79%

	Train	S1W2, S2W1		
	Test	S1W1, S2W2		
#V	#GV	*FMR*	*FNMR*	*FMR+FNMR*
1	1	1.65%	6.25%	7.90%
2	1	1.51%	4.38%	5.89%
3	1	1.91%	2.50%	4.41%
3	2	0.99%	10.00%	10.99%
4	1	1.09%	0.00%	1.09%
4	2	1.25%	1.25%	2.50%
5	1	1.12%	2.50%	3.62%
5	2	0.53%	2.50%	3.03%
5	3	0.79%	5.00%	5.79%
6	1	1.38%	0.00%	1.38%
6	2	1.71%	0.00%	1.71%
6	3	0.92%	2.50%	3.42%
7	1	0.59%	0.00%	0.59%
7	2	0.79%	0.00%	0.79%
7	3	0.53%	0.00%	0.53%
7	4	0.59%	0.00%	0.59%
8	1	0.53%	0.00%	0.53%
8	2	1.25%	0.00%	1.25%
8	3	0.79%	0.00%	0.79%
8	4	1.12%	0.00%	1.12%

Table A.8.: Additional mixed same-day voting results

B. Biometric Performance

The content of chapter B on biometric performance was created by Prof. Dr. Christoph Busch's lecture "Biometric Systems" at the Hochschule Darmstadt, Germany. It is used with his permission.

B.1. Biometric Failures

There are multiple failure associated with a acquisition of a biometric sample or with its processing. In Sections B.2 to B.4 we will discuss the failures that are associated with the deficiency of a biometric system to create a biometric reference for a data subject and subsequently in Sections B.8 to B.9 will consider errors that are attributed to biometric verification and identification systems.

B.2. Failure-to-Capture

A Failure-to-Capture Rate (FTC) is constituted, when the capture process could not generate a biometric sample of sufficient quality. This can be caused due to one of the following reasons:

1. The sample is not generated, as the characteristic is not placed properly on the capture device (e.g finger not covering the sensor area)
2. The captured signal is rejected by the automatic sample quality control algorithm.
3. The captured signal is stored as file, but rejected by the operator (staff expert) subsequent to visual inspection as it is not of sufficient quality

The ISO-definition [ISOb] for the FTC is given by:

Failure-to-Capture Rate: *proportion of failures of the biometric capture process to produce a captured biometric sample that is acceptable for use.*

To estimate the FTC we use the following formula:

$$FTC = \frac{N_{tca} + N_{nsq}}{N_{tot}} \tag{B.1}$$

where N_{tca} is the number of terminated capture attempts, N_{nsq} is the number of images created with insufficient sample quality and N_{tot} is the total number of capture attempts. In consequence of a Failure-to-Capture are new capture attempt is initiated. This is illustrated in figure B.1 .

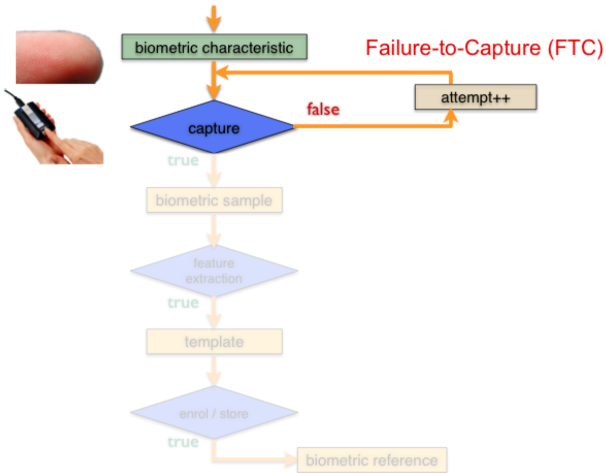

Figure B.1.: Failure-to-Capture (FTC)

B.3. Failure-to-eXtract

A Failure-to-eXtract is constituted, when the feature extraction process was not able to generate a biometric template. This can be caused due to one of the following reasons:

1. The algorithm itself declares that it cannot create a template from the input sample. This could be caused by a insufficient number of features that were identified e.g. only five minutia could be extracted from a fingerprint image.
2. Processing time of feature extraction algorithm exceeds the specified limit and thus the feature extraction is terminated

3. The feature extraction algorithm might suddenly crash during processing. In this case, some actions will be undertaken (e.g. start over application, repeat process, etc.) but if the crash happens all the time with the same sample then for this image a failure to extract feature will be constituted. There is currently no ISO-definition for the Failure-to-eXtract Rate.

To estimate the Failure-to-eXtract Rate (FTX) we use the following formula:

$$FTX = \frac{N_{ngt}}{N_{sub}} \tag{B.2}$$

where N_{ngt} is the number of cases, where no template was generated and and N_{sub} is the total number of biometric samples being submitted to the feature extraction component (i.e. the template generator). In an operational scenario the consequence of a Failure-to-eXtract is a new attempt including a new biometric sample creation and it subsequent processing. This is illustrated in figure B.2 .

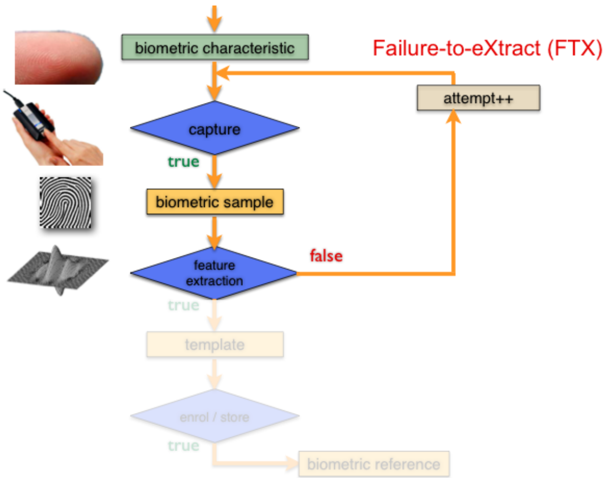

Figure B.2.: Failure-to-eXtract (FTX)

B.4. Failure-to-Enrol

A Failure-to-Enrol is constituted, when the biometric system is not capable to create for data subject a biometric reference. Thus the Failure-to-Enrol Rate (FTE) expresses the proportion of the population, for which the system fails to complete the enrolment process. This can be caused due to one of the following reasons:

1. The biometric characteristic of the subject (e.g. its fingerprint images) can not be captured at all.
2. For each evaluation setting, and if required instances of the same characteristic (e.g. left index finger instead right index finger) it is not possible to create for this subject a template of sufficient quality (e.g. a feature set with minimum number of minutia)

There are currently two ISO-definitions for the FTE. The original definition in the performance testing standard [ISOa] and the more recent one from the harmonized biometric vocabulary [ISOb]:

Failure-to-Enol Rate (ISO 19795-1): *proportion of the population for whom the system fails to complete the enrolment process.*

Failure-to-Enrol Rate (ISO SC37 SD2): *proportion of biometric enrolment (that did not fail for non-biometric reasons), that resulted in a failure to create and store an enrolment data record for an eligible biometric capture subject, in accordance with an enrolment policy.* .

To estimate the FTE we use the following formula:

$$FTE = \frac{N_{nec}}{N} \tag{B.3}$$

where N_{nec} is the number of cases, where we meet one of the two Failure-to-Enrol criteria and N is the total number of subjects, intended to be enroled in the biometric application. The consequence of a Failure-to-Enrol In an operational scenario is that for the capture subject a fallback procedure must be activated that should treat the individual in a non-discriminatory manner. The Failure-to-Enrol is illustrated in figure B.3 .

B.5. Failure-to-Acquire

The Failure-to-Acquire Rate (FTA) is essential for the verification process and estimates the likelihood that biometric comparison can not be completed due to potential deficiencies in the live sample that is submitted as a probe. If there is no feature vector that can be compared to a biometric reference this

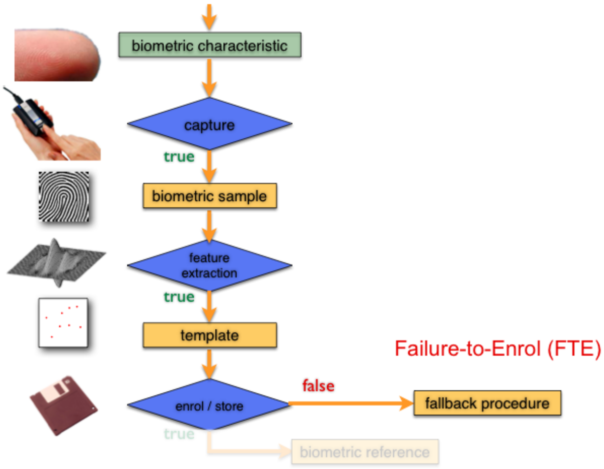

Figure B.3.: Failure-to-Enrol (FTE)

can be caused due to one of the following reasons:

1. The is no biometric sample generated, which is expressed by the FTC.
2. The feature extraction componen failed to extract features as the number and/or quality of extracted features is not sufficient. This is expressed by the FTX.

There are currently two ISO-definitions for the FTA. The original definition in the performance testing standard [ISOa] and the more recent one from the harmonized biometric vocabulary [ISOb]:

Failure-to-Acquire Rate (ISO 19795-1): *proportion of verification or identification attempts for which the system fails to capture or locate an image or signal of sufficient quality.*

Failure-to-Acquire Rate (ISO SC37 SD2): *proportion of a specified set of probe acquisitions that failed to create a biometric probe.*

Note that in ISO SC37 SD2 a *probe* is defined as *biometric data input to an algorithm for comparison to a biometric reference(s).*To estimate the Failure-to-Acquire Rate we use the following formula:

$$FTA = FTC + FTX * (1 - FTC) \tag{B.4}$$

B.6. False-Match

For imposter comparisons a False-Match constitutes the undesired case that an imposter probe is matching a biometric reference, which has not been created for himself. There are currently two ISO-definitions for the corresponding False-Match-Rate (FMR). The original definition in the performance testing standard [ISOa] and the more recent one from the harmonized biometric vocabulary [ISOb]:

False-Match-Rate (ISO 19795-1): *proportion of zero-effort impostor attempt samples falsely declared to match the compared non-self template.*

False-Match-Rate (ISO SC37 SD2): *proportion of the completed biometric non-match comparison trials that result in a false match.*

$$FMR(t) = \int_t^1 \Phi_i(s)ds \tag{B.5}$$

Together with the False-Non-Match-Rate (FNMR) the FMR is the key metric to be used in biometric technology testing and is understood to characterize a security property of a biometric system. Note that some literature is using the term False-Accept-Rate in the meaning of FMR.

B.7. False-Non-Match

For genuine comparisons a False-Non-Match constitutes the undesired case that an genuine probe is not matching to biometric reference, which has been created for the same subject from the same source (e.g. same index finger). There are currently two ISO-definitions for the corresponding False-Non-Match-Rate (FNMR). The original definition in the performance testing standard [ISOa] and the more recent one from the harmonized biometric vocabulary [ISOb]:

False-Non-Match-Rate (ISO 19795-1): *proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same data subject supplying the sample.*

False-Non-Match-Rate (ISO SC37 SD2): *proportion of the completed biometric match comparison trials that result in a false non-match.*

$$FNMR(t) = \int_0^t \Phi_g(s)ds \tag{B.6}$$

Together with the False-Match-Rate (FMR) the FNMR is the key metric to be used in biometric technology testing and is understood to characterize a convenience property of a biometric system. Note that some literature is using the term False-Reject-Rate in the meaning of FNMR.

B.8. Verification System Performance

The first order estimation of the performance for a verification system that is based on transactions allowing multiple attempts can be derived from the detection error trade-off curve. However if this is applied the potential correlations between the attempts are neglected. Such correlations could be due to habituation of the capture subject with the human- computer interface of the biometric system. The relevant measures for a verification system are the False-Accept-Rate (FAR) and the False-Reject-Rate (FRR). The ISO-definition [ISOa] for both metrics are the following:

False-Accept-Rate (ISO 19795-1): *proportion of verification transactions with wrongful claims of identity that are incorrectly confirmed.*

False-Reject-Rate (ISO 19795-1): *proportion of verification transactions with truthful claims of identity that are incorrectly denied.*

For the simplified case that the verification system does allow only a single attempt per transaction then the FAR and FRR can be estimated as follows.

$$FAR = FMR * (1 - FTA) \tag{B.7}$$

and

$$FRR = FTA + FNMR * (1 - FTA) \tag{B.8}$$

If the biometric application is likely to be confronted with a large number of failure to enrol cases (e.g. as it is a fingerprint system for mine workers) and the biometric performance shall be predicted based on a gallery that was collected for a technology testing then the equations B.7 and B.8 do not sufficiently express the performance to be expected. The reason for this is that in a technology evaluation biometric references are generated from the gallery that do not cause a failure-to-enrol and probes that do not cause a failure-to-acquire. For such a case the generalized versions of the above equations are more appropriate, which are given by:

$$GFAR = FMR * (1 - FTA) * (1 - FTE) \tag{B.9}$$

and

$$GFRR = FTE + (1 - FTE) * FTA + (1 - FTE) * (1 - FTA) * FNMR \tag{B.10}$$

B.9. Identification System Performance

The first order estimation of the false positive and false negative identification rates for open-set systems, can be derived from FMR and FNMR and the DET curve. However, such estimates cannot take account of correlations in the comparisons involving the same data subject, and consequently can be quite inaccuraten [ISOa].

$$FPIR = (1 - FTA) * (1 - (1 - FMR)^N) \tag{B.11}$$

where $FPRI$ is the False-Positive-Identification-Rate. For a small FMR we can substitue in equation B.11

$$(1 - FMR)^N \approx 1 - N * FMR \tag{B.12}$$

and thus under the assumption of $FTA = 0$ we derive

$$FPIR = (1 - 0) * (1 - (1 - N * FMR) \tag{B.13}$$

$$FPIR = N * FMR \tag{B.14}$$

Bibliography

[AGL06] Alvarez, D. ; Gonzalez, R.C. ; Lopez, A. ; Alvarez, J.C.: Comparison of Step Length Estimators from Weareable Accelerometer Devices. In: *EMBS '06. 28th Annual International Conference of the IEEEEngineering in Medicine and Biology Society, 2006.*, 2006, S. 5964 –5967

[ALM+05] Ailisto, H.J. ; Lindholm, M. ; Mantyjarvi, J. ; Vildjiounaite, E. ; Makela, S.-M.: Identifying people from gait pattern with accelerometers. In: Jain, A. K. (Hrsg.) ; Ratha, N. K. (Hrsg.): *Biometric Technology for Human Identification II* Bd. 5779, SPIE, 2005, S. 7–14

[And10] *Phones parts list Motorola Milestone.* And Developers Wiki. `http://and-developers.com/device_information`. Version: 2010. – Online; last accessed 20th August 2010

[ASW08] Ali, H. ; Salami, M.J.E. ; Wahyudi: Iris recognition system by using support vector machines. In: *Proceedings of the International Conference on Computer and Communication Engineering, 2008. ICCCE 2008.*, 2008, S. 516 –521

[BB06] Best, R. ; Begg, R.: Overview of movement analysis and gait features. In: *Computational Intelligence for Movement Sciences: Neural Networks and Other Emerging Techniques. Idea Group* (2006), S. 11–18

[BC99] Burges, C. J. ; Crisp, D. C.: Uniqueness of the SVM Solution. In: *NIPS*, 1999, S. 223–229

[BC00] Bennett, K. P. ; Campbell, C.: Support vector machines: hype or hallelujah? In: *SIGKDD Explor. Newsl.* 2 (2000), December, S. 1–13. – ISSN 1931–0145

[BCD02] BenAbdelkader, C. ; Cutler, R. ; Davis, L.: Stride and cadence as a biometric in automatic person identification and verification. In: *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.*, 2002, S. 372 –377

[BN10] Breitinger, F. ; Nickel, C.: User Survey on Phone Security and Usage. In: Brömme, A (Hrsg.) ; Busch, C. (Hrsg.): *BIOSIG 2010: Biometrics and Electronic Signatures, Proceedings of the Special Interest Group on Biometrics and Electronic Signatures*, Gesellschaft für Informatik (GI), 2010 (Lecture Notes in Informatics)

[BS10] Bours, P. ; Shrestha, R.: Eigensteps: A giant leap for gait recognition. In: *Proceedings of the 2nd International Workshop on Security and Communication Networks (IWSCN), 2010.*, 2010, S. 1 –6

[BSC04] Batra, D. ; Singhal, G. ; Chaudhury, S.: Gabor filter based fingerprint classification using support vector machines. In: *Proceedings of the IEEE INDICON 2004. First India Annual Conference, 2004.*, 2004, S. 256 – 261

[BSK07] Bidargaddi, N. ; Sarela, A. ; Klingbeil, L. ; Karunanithi, M.: Detecting walking activity in cardiac rehabilitation by using accelerometer. In: *Proceedings of the 3rd International Conference on Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007.*, 2007, S. 555 –560

[BSR09] Bächlin, M. ; Schumm, J. ; Roggen, D. ; Tröster, G.: Quantifying Gait Similarity: User Authentication and Real-World Challenge. In: *Proceedings of the Advances in Biometrics, Third International Conference, ICB 2009, Alghero, Italy, June 2009.* Bd. 5558, Springer Berlin / Heidelberg, Juni 2009 (Lecture Notes in Computer Science). – ISBN 978–3–642–01792–6, S. 1040–1049

[Bur98] Burges, C. J.: A Tutorial on Support Vector Machines for Pattern Recognition. In: *Data Mining and Knowledge Discovery* 2 (1998), S. 121–167

[CGS02] Collins, R.T. ; Gross, R. ; Shi, J.: Silhouette-based human identification from body shape and gait. In: *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.*, 2002, S. 366 –371

[CJ10] Cellan-Jones, R.: *Government calls for action on mobile phone crime*. BBC News. `http://news.bbc.co.uk/2/hi/technology/8509299.stm`. Version: February 2010. – Online; last accessed 16th January 2011

[CKC03] Cuntoor, N. ; Kale, A. ; Chellappa, R.: Combining multiple evidences for gait recognition. In: *Proceedings of the 2003 International Conference on Multimedia and Expo, 2003. ICME '03.* Bd. 3, 2003, S. III – 113–16 vol.3

[CL01] Chang, C.-C. ; Lin, C.-J.: *LIBSVM: a library for support vector machines*, 2001. – Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`

[CLG09] Chen, S.-H. ; Luo, Y.-R. ; Guido, R.C.: Speaker Verification Using Line Spectrum Frequency, Formant, and Support Vector Machine. In: *Proceedings of the 11th IEEE International Symposium on Multimedia, 2009. ISM '09.*, 2009, S. 562 –566

[Cov65] Cover, T. M.: Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. In: *IEEE Transactions on Electronic Computers* EC-14 (1965), Nr. 3, S. 326 –334

[CRZ06] Chai, Y. ; Ren, J. ; Zhao, R. ; Jia, J.: Automatic Gait Recognition using Dynamic Variance Features. In: *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, 2006. FGR 2006.*, 2006, S. 475 –480

[CSC+01] Crammer, K. ; Singer, Y. ; Cristianini, N. ; Shawe-Taylor, J. ; Williamson, B.: On the algorithmic implementation of multiclass kernel-based vector machines. In: *Journal of Machine Learning Research* 2 (2001), S. 2001

[CST00] Cristianini, N. ; Shawe-Taylor, J.: *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods.* 1. Cambridge University Press, 2000. – ISBN 0521780195

[DASZ09] Dadashi, F. ; Araabi, B.N. ; Soltanian-Zadeh, H.: Gait Recognition Using Wavelet Packet Silhouette Representation and Transductive Support Vector Machines. In: *Proceedings of the 2nd International Congress on Image and Signal Processing, 2009. CISP '09.*, 2009, S. 1 –5

[DBH10] Derawi, M.O. ; Bours, P. ; Holien, K.: Improved Cycle Detection for Accelerometer Based Gait Authentication. In: *Proceedings of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010. IIH-MSP 2010.*, 2010, S. 312 –317

[DDV07] Dinerstein, S. ; Dinerstein, J. ; Ventura, D.: Robust multi-modal biometric fusion via multiple SVMs. In: *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 2007. ISIC.*, 2007, S. 1530 –1535

[Der10] Derawi, M.O.: Accelerometer-Based Gait Analysis, A survey. In: *Norwegian Information Security Conference*, 2010

[DGL+10] Derawi, M.O. ; Gafurov, D. ; Larsen, R. ; Busch, C. ; Bours, P.: Fusion of gait and fingerprint for user authentication on mobile devices. In: *Proceedings of the 2nd International Workshop on Security and Communication Networks, 2010. IWSCN 2010.*, 2010, S. 1 –6

[DJJM+10] Djuric-Jovicic, M. ; Jovicic, N. S. ; Milovanovic, I. ; Radovanovic, S. ; Kresojevic, N. ; Popovic, M. B.: Classification of walking patterns in Parkinson's disease patients based on inertial sensor data. In: *Proceedings of the 10th Symposium on Neural Network Applications in Electrical Engineering, 2010. NEUREL 2010.*, 2010, S. 3 –6

[DK05] Duan, K. ; Keerthi, S. S.: Which is the best multiclass SVM method? An empirical study. In: *Proceedings of the Sixth International Workshop on Multiple Classifier Systems, 2005.*, 2005, S. 278–285

[DNB10] Derawi, M.O. ; Nickel, C. ; Bours, P. ; Busch, C.: Unobtrusive User-Authentication on Mobile Phones using Biometric Gait Recognition. In: *Proceedings of the Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010.*, 2010

[DS02] Decoste, D. ; Schölkopf, B.: Training Invariant Support Vector Machines. In: *Mach. Learn.* 46 (2002), March, S. 161–190. – ISSN 0885–6125

[Ell] Ellis, D. ; Department of Electrical Engineering, Columbia University (Hrsg.): *Reproducing the feature outputs of common programs using Matlab and melfcc.m.* Department of Electrical Engineering, Columbia University, `http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/mfccs.html`. – Online; last accessed 4th January 2011

[FAE08] Fahmy, M.S. ; Atyia, A.F. ; Elfouly, R.S.: Biometric Fusion Using Enhanced SVM Classification. In: *Proceedings of the IIHMSP '08 Inter-*

national Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2008., 2008, S. 1043 –1048

[Gaf07a] Gafurov, D.: Security Analysis of Impostor Attempts with Respect to Gender in Gait Biometrics. In: *Proceedings of the First IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007.*, 2007, S. 1 –6

[Gaf07b] Gafurov, D.: A Survey of Biometric Gait Recognition: Approaches, Security and Challenges. In: *Proceedings of Annual Norwegian Computer Science Conference, 2007.*, 2007

[GFK05] Ganchev, T. ; Fakotakis, N. ; Kokkinakis, G.: Comparative evaluation of various MFCC implementations on the speaker verification task. In: *Proceedings of the International Conference Speech and Computer, 2005. SPECOM 2005.*, 2005, S. 191–194

[GHS06] Gafurov, D. ; Helkala, K. ; Soendrol, T.: Gait recognition using acceleration from MEMS. In: *The First International Conference on Availability, Reliability and Security, 2006. ARES 2006.*, 2006, S. 6 pp.

[GHS10] Gafurov, D. ; Hagen, J. ; Snekkenes, E.: Temporal Characteristics of Gait Biometrics. In: *Proceedings of the Second International Conference on Computer Engineering and Applications (ICCEA), 2010.* Bd. 2, 2010, S. 557 –561

[GIA10] Global Industry Analysts, Inc.: *Mobile Banking: A Global Strategic Business Report.* `http://www.strategyr.com/pressMCP-6073.asp`. Version: February 2010. – Press Release; Online; last accessed 27th December 2010

[GS08] Gafurov, D. ; Snekkenes, E.: Towards understanding the uniqueness of gait biometric. In: *Proceedings of the 8th IEEE International Conference on Automatic Face Gesture Recognition, 2008. FG '08.*, 2008, S. 1 –8

[GS09] Gafurov, D. ; Snekkenes, E.: Gait recognition using wearable motion recording sensors. In: *EURASIP Journal on Advances in Signal Processing* 2009 (2009), S. 1–16

[GSB06] Gafurov, D. ; Snekkenes, E. ; Buvarp, T.: Robustness of Biometric Gait Authentication Against Impersonation Attack. In: Meersman, Robert (Hrsg.) ; Tari, Zahir (Hrsg.) ; Herrero, Pilar (Hrsg.): *On the*

Move to Meaningful Internet Systems 2006: OTM 2006 Workshops Bd. 4277. Springer Berlin / Heidelberg, 2006, S. 479–488

[GSB07] Gafurov, D. ; Snekkenes, E. ; Bours, P.: Spoof Attacks on Gait Authentication System. In: *IEEE Transactions on Information Forensics and Security* 2 (2007), sep., Nr. 3, S. 491 –502

[GSB10] Gafurov, D. ; Snekkenes, E. ; Bours, P.: Improved Gait Recognition Performance Using Cycle Matching. In: *Proceedings of the IEEE 24th International Conference on Advanced Information Networking and Applications Workshops (WAINA), 2010.*, 2010, S. 836 –841

[HCH07] Huang, B. ; Chen, M. ; Huang, P. ; Xu, Y.: Gait Modeling for Human Identification. In: *Proceedings of the IEEE International Conference on Robotics and Automation, 2007.*, 2007, S. 4833 –4838

[HCL03] Hsu, C. W. ; Chang, C. C. ; Lin, C. J. ; Department of Computer Science, National Taiwan University (Hrsg.): *A practical guide to support vector classification.* Taipei, Taiwan: Department of Computer Science, National Taiwan University, 2003. `http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf`. – Online; last accessed 12th November 2010

[HCY06] Huang, B. ; Chen, M. ; Ye, W. ; Xu, Y.: Intelligent Shoes for Human Identification. In: *Proceedings of the IEEE International Conference on Robotics and Biomimetics, 2006. ROBIO '06.*, 2006, S. 601 –606

[Her90] Hermansky, H.: Perceptual linear predictive (PLP) analysis of speech. In: *The Journal of the Acoustical Society of America* 87 (1990), Nr. 4, S. 1738–1752

[HHR07] Holien, K. ; Hammersland, R. ; Risa, T.: How Different Surfaces Affect Gait Based Authentication / Norwegian Information Security Lab, Gjøvik University College. 2007. – project report

[HKJ04] Hwang, J.Y. ; Kang, J.M. ; Jang, Y.W. ; Kim, H.C.: Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly. In: *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004. IEMBS '04.* Bd. 1, 2004, S. 2204 –2207

[ISOa] ISO/IEC TC JTC1 SC37 Biometrics: *ISO/IEC 19795-1:2006. In-*

formation Technology – Biometric Performance Testing and Reporting – Part 1: Principles and Framework

[ISOb] ISO/IEC TC JTC1 SC37 Biometrics: *ISO/IEC JTC1 SC37: Harmonized Biometric Vocabulary*

[ISOc] ISO/IEC TC JTC1 SC37 Biometrics: *ISO/IEC JTC1 SC37: Text of Standing Document 11 (SD 11), Part 1 Overview Standards Harmonization Document*

[ISOd] ISO/IEC TC JTC1 SC37 Biometrics: *ISO/IEC TR 24722:2007: Information technology – Biometrics – Multimodal and other multibiometric fusion*

[IY06] Iso, T. ; Yamazaki, K.: Gait analyzer based on a cell phone with a single three-axis accelerometer. In: *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, 2006.* New York, NY, USA : ACM, 2006 (MobileHCI '06). – ISBN 1–59593–390–5, S. 141–144

[JB01] Johnson, A. ; Bobick, A.: A Multi-view Method for Gait Recognition Using Static Body Parameters. In: Bigun, Josef (Hrsg.) ; Smeraldi, Fabrizio (Hrsg.): *Audio- and Video-Based Biometric Person Authentication* Bd. 2091. Springer Berlin / Heidelberg, 2001, S. 301–311

[JBL03] Jung, J.-W. ; Bien, Z. ; Lee, S.-W. ; Sato, T.: Dynamic-footprint based person identification using mat-type pressure sensor. In: *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2003.* Bd. 3, 2003, S. 2937 – 2940 Vol.3

[JK08] Jang, J. ; Kim, H.: Score-level fusion in multiple biometrics using non-linear classification. In: *Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision, 2008. ICARCV 2008.*, 2008, S. 417 –421

[JRW86] Juang, B. ; Rabiner, L. ; Wilpon, J.: On the use of bandpass liftering in speech recognition. In: *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86.* Bd. 11, 1986, S. 765 – 768

[KKW07] Kangas, M. ; Konttila, A. ; Winblad, I. ; Jamsa, T.: Determination of simple thresholds for accelerometry-based parameters for fall detection. In: *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007. EMBS 2007.*, 2007, S. 1367 –1370

[KSS03] Kern, N. ; Schiele, B. ; Schmidt, A.: Multi-Sensor Activity Context Detection for Wearable Computing. In: *Proceedings of the European Symposium on Ambient Intelligence, 2003. EUSAI, LNCS, 2003.* Bd. 2875, 2003, S. 220–232

[KWM10] Kwapisz, J. R. ; Weiss, G. M. ; Moore, S. A.: Cell Phone-Based Biometric Identification. In: *Proceedings of the IEEE Fourth International Conference on Biometrics: Theory, Applications and Systems, 2010.*, 2010

[LCW09] Liu, H. ; Cao, Y. ; Wang, Z.: A novel algorithm of gait recognition. In: *Proceedings of the International Conference on Wireless Communications Signal Processing, 2009. WCSP 2009.*, 2009, S. 1 –5

[LF02] Lyu, S. ; Farid, H.: Detecting hidden messages using higher-order statistics and support vector machines. In: *In 5th International Workshop on Information Hiding*, Springer-Verlag, 2002, S. 340–354

[LG02] Lee, L. ; Grimson, W.E.L.: Gait analysis for recognition and classification. In: *Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.*, 2002, S. 148 –155

[LMS04] Liu, Z. ; Malave, L. ; Sarkar, S.: Studies on silhouette quality and gait recognition. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, 2004.* Bd. 2, 2004, S. II–704 – II–711 Vol.2

[Log00] Logan, B.: Mel Frequency Cepstral Coefficients for Music Modeling. In: *In International Symposium on Music Information Retrieval*, 2000

[LS04] Liu, Z. ; Sarkar, S.: Simplest representation yet for gait recognition: averaged silhouette. In: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.* Bd. 4, 2004, S. 211 – 214 Vol.4

[LSL07] Larsen, P. K. ; Simonsen, E. B. ; Lynnerup, N.: Gait analysis in forensic medicine. In: Beraldin, J.-A. (Hrsg.) ; Remondino, F. (Hrsg.) ; Shortis, F. R. (Hrsg.): *Videometrics IX* Bd. 6491, SPIE, 2007, S. 64910M

[LY09] Liu, K. ; Yang, J.: Recognition of People Reoccurrences Using Bag-Of-Features Representation and Support Vector Machine. In: *Chinese Conference on Pattern Recognition, 2009.*, 2009, S. 1 –5

[Mat03] Mathie, M. J.: *Monitoring and interpreting human movement patterns using a triaxial accelerometer*, University of New South Wales. School of Electrical Engineering and Telecommunications, Faculty of Engineering, Diss., 2003

[MBB05] Middleton, L. ; Buss, A.A. ; Bazin, A. ; Nixon, M.S.: A floor sensor system for gait recognition. In: *Fourth IEEE Workshop on Automatic Identification Advanced Technologies, 2005.*, 2005, S. 171 – 176

[MLH03] Meyer, D. ; Leisch, F. ; Hornik, K.: The support vector machine under test. In: *Neurocomputing* 55 (2003), Nr. 1-2, S. 169 – 186. – ISSN 0925–2312. – Support Vector Machines

[MLV+05] Mantyjarvi, J. ; Lindholm, M. ; Vildjiounaite, E. ; Makela, S.-M. ; Ailisto, H.A.: Identifying users of portable devices from gait pattern with accelerometers. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, 2005, S. ii/973 – ii/976 Vol. 2

[MTG08] Moustakidis, S.P. ; Theocharis, J.B. ; Giakas, G.: Subject Recognition Based on Ground Reaction Force Measurements of Gait Signals. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* 38 (2008), Nr. 6, S. 1476 –1485

[NA94] Niyogi, S.A. ; Adelson, E.H.: Analyzing and recognizing walking figures in XYT. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '94., 1994.*, 1994, S. 469 –474

[NBR11] Nickel, C. ; Busch, C. ; Rangarajan, S. ; Möbius, M.: Using Hidden Markov Models for Accelerometer-Based Biometric Gait Recognition. In: *2011 7th International Colloquium on Signal Processing & Its Applications (CSPA 2010)*, 2011. – [to be published in March 2011]

[Now09] Nowlan, M. F.: Human Identification via Gait Recognition Using Accelerometer Gyro Forces / Yale University, Department of Computer Science. Version: 2009. `http://cs-www.cs.yale.edu/homes/mfn3/pub/mfn_gait_id.pdf`. 2009. – project report. – Online; last accessed 14th December 2010

[OA00] ORR, R. J. ; ABOWD, G. D.: The smart floor: a mechanism for natural user identification and tracking. In: *Proceedings of the ACM CHI '00Human factors in Computing Systems, 2000.* New York, NY, USA : ACM, 2000, S. 275–276

[O'S99] O'SHAUGHNESSY, D.: *Speech Communications: Human and Machine.* 2. Wiley-IEEE Press, 1999. – ISBN 9780780334496

[Pla99] PLATT, J.C.: Using Analytic QP and Sparseness to Speed Training of Support Vector Machines. In: *In Neural Information Processing Systems 11*, MIT Press, 1999, S. 557–563

[PTK09] PAIYAROM, S. ; TUNGAMCHIT, P. ; KEINPRASIT, R. ; KAYASITH, P.: Activity monitoring system using Dynamic Time Warping for the elderly and disabled people. In: *Proceedings of the 2nd International Conference on Computer, Control and Communication, 2009. IC4 2009.*, 2009, S. 1 –4

[PW00] POTRA, F. A. ; WRIGHT, S. J.: Interior-point methods. In: *Journal of Computational and Applied Mathematics* 124 (2000), Nr. 1-2, S. 281 – 302. – ISSN 0377–0427

[PZW09] PAN, G. ; ZHANG, Y. ; WU, Z.: Accelerometer-based gait recognition via voting by signature points. In: *Electronics Letters* 45 (2009), oct., Nr. 22, S. 1116 –1118

[QZK10] QIAN, G. ; ZHANG, J. ; KIDANE, A.: People Identification Using Floor Pressure Sensing and Analysis. In: *Sensors Journal, IEEE* 10 (2010), sep., Nr. 9, S. 1447 –1460

[RJ93] RABINER, L. ; JUANG, B.-H.: *Fundamentals of speech recognition.* Upper Saddle River, NJ, USA : Prentice-Hall, Inc., 1993. – ISBN 0–13–015157–2

[RJ04] ROSS, A. ; JAIN, A.K.: Multimodal Biometrics: An Overview. In: *Proceedings of the 12th European Signal Processing Conference (EUSIPCO), 2004.*, 2004, S. 1221 – 1224

[RJM07] RONG, L. ; JIANZHONG, Z. ; MING, L. ; XIANGFENG, H.: A Wearable Acceleration Sensor System for Gait Recognition. In: *Proceedings of the 2nd IEEE Conference on Industrial Electronics and Applications, 2007. ICIEA 2007.*, 2007, S. 2654 –2659

[RPL99] Richman, M.S. ; Parks, T.W. ; Lee, Hsien-Che: A novel support vector machine-based face detection method. In: *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers, 1999.* Bd. 1, 1999, S. 740 –744 vol.1

[RZJ07] Rong, L. ; Zhiguo, D. ; Jianzhong, Z. ; Ming, L.: Identification of Individual Walking Patterns Using Gait Acceleration. In: *Proceedings of the 1st International Conference on Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007.*, 2007, S. 543 –546

[SH04] Skowronski, M. D. ; Harris, J. G.: Exploiting independent filter bandwidth of human factor cepstral coefficients in automatic speech recognition. In: *The Journal of the Acoustical Society of America* 116 (2004), Nr. 3, S. 1774–1780

[SPL+05] Sarkar, S. ; Phillips, P.J. ; Liu, Z. ; Vega, I.R. ; Grother, P. ; Bowyer, K.W.: The humanID gait challenge problem: data sets, performance, and analysis. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (2005), Nr. 2, S. 162 –177

[SPST+01] Schölkopf, B. ; Platt, J. C. ; Shawe-Taylor, J. C. ; Smola, A. J. ; Williamson, R. C.: Estimating the Support of a High-Dimensional Distribution. In: *Neural Comput.* 13 (2001), July, S. 1443–1471. – ISSN 0899–7667

[SR04] Suutala, J. ; Röning, J.: Towards the adaptive identification of walkers: automated feature selection of footsteps using distinction-sensitive LVQ. In: *Proceedings of International Workshop on Processing Sensory Information for Proactive Systems, 2004. PSIPS 2004.*, 2004, S. 61–67

[SS04] Smola, A. J. ; Schölkopf, B.: A tutorial on support vector regression. In: *Statistics and Computing* 14 (2004), August, S. 199–222. – ISSN 0960–3174

[STM09] STMicroelectronics (Hrsg.): *LIS331DLH MEMS digital output motion sensor ultra low-power high performance 3-axes "nano" accelerometer.* STMicroelectronics, 2009. `http://www.st.com/stonline/books/pdf/docs/15101.pdf`. – Online; last accessed 14th September 2010

[SV40] Stevens, S.S. ; Volkmann, J.: The relation of pitch to frequency: A revised scale. In: *American Journal of Psychology* 53 (1940), S. 329–353

[SZ09] Sprager, S. ; Zazula, D.: A cumulant-based method for gait identification using accelerometer data with principal component analysis and support vector machine. In: *WSEAS Trans. Sig. Proc.* 5 (2009), Nr. 11, S. 369–378

[TAH09] Tamviruzzaman, M. ; Ahamed, S. I. ; Hasan, C. S. ; O'brien, C.: ePet: when cellular phone learns to recognize its owner. In: *Proceedings of the 2nd ACM workshop on Assurable and usable security configuration, 2009. SafeConfig '09.* New York, NY, USA : ACM, 2009, S. 13–18

[TB01] Tanawongsuwan, R. ; Bobick, A.: Gait recognition from time-normalized joint-angle trajectories in the walking plane. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001.* Bd. 2, 2001, S. II–726 – II–731 vol.2

[TCD+00] Terrence, S. F. ; Cristianini, N. ; Duffy, N. ; Bednarski, D. W. ; Haussler, D.: Support vector machine classification and validation of cancer tissue samples using microarray expression data. In: *BIOINFORMATICS* 16 (2000), S. 906–914

[TJD09] Teixeira, T. ; Jung, D. ; Dublon, G. ; Savvides, A.: PEM-ID: Identifying people by gait-matching using cameras and wearable accelerometers. In: *Proceedings of the Third ACM/IEEE International Conference on Distributed Smart Cameras, 2009. ICDSC 2009.*, 2009, S. 1 –8

[Toh07] Toh, K.-A.: Error-Rate Based Biometrics Fusion. In: Lee, S.-W. (Hrsg.) ; Li, S. (Hrsg.): *Advances in Biometrics* Bd. 4642. Springer Berlin / Heidelberg, 2007, S. 191–200

[Tra90] Traunmüller, H.: Analytical expressions for the tonotopic sensory scale. In: *The Journal of the Acoustical Society of America* 88 (1990), Nr. 1, S. 97–100

[TTA+10] Takeda, T. ; Taniguchi, K. ; Asari, K. ; Kuramoto, K. ; Kobashi, S. ; Hata, Y.: Biometric personal authentication by one step foot pressure distribution change by fuzzy artificial immune system. In: *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ), 2010.*, 2010, S. 1 –6

[TTT06] Trung, N. T. ; Thao, T. D. ; Trung, P. N. ; Triet, T. M.: A New Approach to Identify Fingerprint Using Support Vector Machine. In:

Proceedings of the International Conference on Computational Intelligence and Security, 2006. Bd. 1, 2006, S. 168 –171

[UCN99] Umesh, S. ; Cohen, L. ; Nelson, D.: Fitting the Mel scale. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '99., 1999.* Bd. 1, 1999, S. 217 –220 vol.1

[Vap82] Vapnik, V.: *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (Springer Series in Statistics).* Secaucus, NJ, USA : Springer-Verlag New York, Inc., 1982. – ISBN 0387907335

[Vap99] Vapnik, V. N.: An overview of statistical learning theory. In: *IEEE Transactions on Neural Networks* 10 (1999), September, Nr. 5, S. 988 –999. – ISSN 1045–9227

[VML+06] Vildjiounaite, E. ; Mäkelä, S.-M. ; Lindholm, M. ; Riihimäki, R.a ; Kyllönen, V. ; Mäntyjärvi, J. ; Ailisto, H.: Unobtrusive Multimodal Biometrics for Ensuring Privacy and Information Security with Personal Devices. In: Fishkin, K. (Hrsg.) ; Schiele, B. (Hrsg.) ; Nixon, P. (Hrsg.) ; Quigley, A. (Hrsg.): *Pervasive Computing* Bd. 3968. Springer Berlin / Heidelberg, 2006, S. 187–201

[VML+07] Vildjiounaite, E. ; Makela, S.-M. ; Lindholm, M. ; Kyllonen, V. ; Ailisto, H.: Increasing Security of Mobile Devices by Decreasing User Effort in Verification. In: *Proceedings of the Second International Conference on Systems and Networks Communications, 2007. ICSNC 2007.*, 2007, S. 80 –80

[Win09] Winter, D. A.: *Biomechanics and Motor Control of Human Movement.* 4. Wiley, 2009

[WN10] Witte, H. ; Nickel, C.: Modular Biometric Authentication Service System (MBASSy). In: *BIOSIG 2010: Biometrics and Electronic Signatures, Proceedings of the Special Interest Group on Biometrics and Electronic Signatures*, 2010

[WTN03] Wang, L. ; Tan, T. ; Ning, H. ; Hu, W.: Silhouette analysis-based gait recognition for human identification. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25 (2003), dec., Nr. 12, S. 1505 – 1518

[WZC08] Wang, Y. ; Zhang, F. ; Chen, L.: An Approach to Incremental SVM Learning Algorithm. In: *Computing, Communication, Control,* *and Management, 2008. CCCM '08. ISECS International Colloquium on* Bd. 1, 2008, S. 352 –354

[YBS09] Yamauchi, K. ; Bhanu, B. ; Saito, H.: Recognition of walking humans in 3D: Initial results. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009.*, 2009, S. 45 –52

[YHKY08] Yoo, J.-H. ; Hwang, D. ; K.-Y., Moon ; Nixon, M.S.: Automated Human Recognition by Gait using Neural Network. In: *First Workshops on Image Processing Theory, Tools and Applications, 2008. IPTA 2008.*, 2008, S. 1 –6

[YN03] Yoo, JH. ; Nixon, M. S.: Markerless Human Gait Analysis via Image Sequences. In: *International Society of Biomechanics XIXth Congress*, 2003, S. CD–ROM

[ZLL06] Zhao, G. ; Liu, G. ; Li, H. ; Pietikainen, M.: 3D gait recognition using multiple cameras. In: *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, 2006. FGR 2006.*, 2006, S. 529 –534

[Zwi61] Zwicker, E.: Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen). In: *The Journal of the Acoustical Society of America* 33 (1961), Nr. 2, S. 248–248

[ZZW09] Zheng, J. ; Zhang, G. ; Wu, T.: Design of Automatic Fall Detector for Elderly Based on Triaxial Accelerometer. In: *Proceedings of the 3rd International Conference on Bioinformatics and Biomedical Engineering , 2009. ICBBE 2009.*, 2009, S. 1 –4