Person Identification for Access Control using NIR Hand Vein Imaging

Author:
Morten Bojesen Bonderup

Supervisor:
Thomas B. Moeslund

Aalborg University
Vision, Graphics and
Interactive Systems

Master's thesis

08th June 2017

Aalborg University

AALBORG UNIVERSITY

Title:
Person Identification for Access Control using NIR Hand Vein Imaging
Subject:
Vision, Graphics and Interactive Systems
Project period:
1/2-2017 to 08/06-2017
Project group:
17gr1042
Participants:
Morten Bojesen Bonderup

Supervisor:
Thomas B. Moeslund
Number of pages:
43
Finished:
08/06-2017

Physical access control systems which rely on physical access tokens such as keycards, chips or mechanical keys are widely used. The use of physical access tokens can however in some scenarios cause issues. One example is care homes for elderly or dementia patients. At such facilities the residents may not be capable of handling these physical access tokens, they may lose them or forget them. Therefore it can be difficult to apply access control on residents private rooms. A way of solving the issues of physical access tokens is to use a biometric access control systems.

Most current biometric access control system do however rely on active interaction from the users and may be only applicable for two-factor authentication. Such system will therefore not solve all issues regarding physical access tokens, as such system will instead require training or developing new habits.

To address these issues the focus of this project is to investigate the possibilities of using hand vein recognition for a biometric access control system. A proof-of-concept is developed which uses a CCTV camera and near infrared light to capture hand vein images. Furthermore a dataset of hand vein images is acquired to test this proof-of-concept.

Preface

This report is written as a master's thesis in Vision, Graphics and Interactive Systems at Aalborg University. The project has been carried out in collaboration with the danish Security company Actas A/S.

I would like to thank Thomas Christensen and Actas A/S for providing equipment for data acquisition.

Reading Guide

In the report, figures and tables are numbered according to their respective chapters. Citations to literature, data sheets and manuals are marked with [x], where x is the reference number used in the bibliography which follows the IEEE citation standard.

The code is implemented in Python 2.7.6 and OpenCV 2.4.13. The code is uploaded alongside the report.

Morten Bojesen Bonderup

Contents

1 Introduction

"Nationalt VidensCenter for Demens" estimates that approximately 84,000 people in Denmark suffers from a Dementia disease and that this number will double over the next 25 years. At care homes for dementia patients it is difficult to make use of access control on doors to the residents' private rooms as there is a risk that the residents will often forget or lose their keys or keycards. It is therefore not optimal to use common locks, utilizing either keys or keycards, to private rooms. On the other hand it is neither appropriate to keep private rooms unlocked due to the risk of theft or the risk that some residents might, by accident, enter a wrong room.

Care homes for elderly is just one example where the use of physical access tokens (keys, chips, keycards, etc.) can cause problems. Another example is educational institutes and such where they experience a high turnover in people are allowed entrance. These issues can be solved by deploying access control systems that do not rely on physical access tokens, but instead relies on biometrics. A biometric access control system uses biological parameters to identify a person to determine whether the person has access to enter the room. There are a variety of different biometric systems which utilize different types of biological parameters to determine a persons identity with some examples being fingerprint-, iris- or face-recognition. Examples of such are the HandKey hand geometry reader from Honeywell Access [1]. This product is designed for two-factor authentication in collaboration with a card reader. Another example is the MorphoAccess SIGMA Lite fingerprint reader [2], which utilizes fingerprints for two-factor authentication with a keycard. The downside to these examples are the requirement for active interaction with the system, which means that the user has to learn to use the system, and may have to spend time on scanning face or fingers to be able to unlock a door.

A rather new biometric feature is the use of vein patterns to identity persons and pal vein patterns are supposedly unique even for twins [3]. Vein patterns can also be obtained from the dorsal side of the hand using a camera. This can therefore possibly be utilized for an access control system. By placing a camera on a door above the door handle, hand vein images can be obtained without requiring interaction from the user. Identifying persons when they grab the door handle can therefore result in a access control system which is transparent to users who are allowed to enter a given area. This is especially interesting for scenarios as care homes, as such a solution will not require any training and will be relatively intuitive as no procedures regarding opening a door has to be changed.

Such access control systems are of interest for the Danish security company Actas A/S, with whom this project is developed in collaboration with.

1.1 Initial problem statement

As mentioned above dorsal hand vein recognition has the potential to be used for access control system. The purpose of this project is therefore to investigate the following:

- *How can vein pattern recognition be utilized for a physical access control system?*
  - *How can vein patterns be acquired?*
  - *How should a biometric access control system be developed?*

2 Technical Analysis

The previous section and problem statement form the basis for this project and the technical aspects of this are analyzed and described in this section. First of all the structure of biometric recognition systems in general are described, followed by an analysis of hand vein recognition.

2.1 Biometric Recognition System

A biometric recognition system uses biometric features, either physiologic or behavioral, to identify or recognize persons [4]. Examples of biometric features are fingerprints, hand-vein patterns, face, gait, voice etc. Generally biometric features can be divided into two modalities, behavioral or physiological. Behavioral biometric features are active features that describes the behaviour of a person, gait and speech recognition falls under this category. Physiological biometric features are physical features from a persons body and thus include features such as fingerprints, hand-vein patterns and face.

Biometric matching is usually said to be either 1:1 or 1:n (one to many). 1:1 matching is also known as verification or authentication while 1:n is known as identification. Some systems utilize biometric features alongside some sort of physical access token for identification to increase the security which is known as two-factor authentication. An example is to use an access card along with fingerprints to verify that the user is actually the owner of the access card. This is a case of 1:1 matching as the biometric features is used to verify the identity of the card holder. If the use case is to determine the identity of a person based solely on the biometric feature the use of 1:n matching is necessary, meaning that the extracted feature has to be matching against a database containing biometric features for several different persons. A biometric recognition system usually consists of four modules[4]:

- Sensor module
- Feature extraction module
- Matching module
- Database

The sensor module is used to capture the features and is dependant on the type of biometric feature of interest. For some types of biometrics an optical sensor can be used for capturing the features. Some examples of this are fingerprint, face, gait, hand-vein and iris.

The feature extraction module is responsible for extracting useful features from the sensor module. Ideally these features have to be unique for every person and easily distinguishable from other persons. The matching module matches extracted features against the database of known features to identify or verify the identification of the given sample. For an access control system the database will most likely consist of features for known people that are allowed to enter.

As described in Section *1 Introduction* the biometric feature of interest in this project is hand vein recognition.

2.2 Hand Vein Biometrics

Hand vein biometrics are based on using the pattern, structure and placement of veins in peoples' hands for identification. The hand vein pattern is found to be unique for persons [5], but due to the limited research within the subject it is difficult to determine exactly how unique hand vein patterns are [6]. The vascular network, including the veins, are responsible for transporting oxygen around in the body via the blood. Only some veins can be seen with the naked eye, but by using Near-infrared light the veins become clearer, which is described in Section *2.2.1 Near Infrared Light and Optical Window*. An example image of the dorsal side of a hand can be seen in Figure 2.1.

Figure 2.1: Example RGB image of the dorsal side of a hand.

In this image some of the veins are rather visible to the naked eye, while some are hidden.

The fact that the veins are hidden beneath the skin makes it more difficult to impersonate compared to other biometrics such as face or fingerprints which are easier to obtain from a given person.

The vascular network in humans develop during childhood and until the age 20. After the age of 20 the vascular network does not normally change rapidly [7]. Some diseases can however cause angiogenesis which is the formation of new blood vessels [8]. This could therefore result in changes to the vein pattern for a person and therefore require re-enrollment to be able to identify that person. For healthy persons minor changes can also occur due to temperature, alcohol intake and such, but this should only be minor changes such as vein size and not affect the structure or pattern of the veins [7].

The following section will describe the procedure of capturing the hand vein patterns.

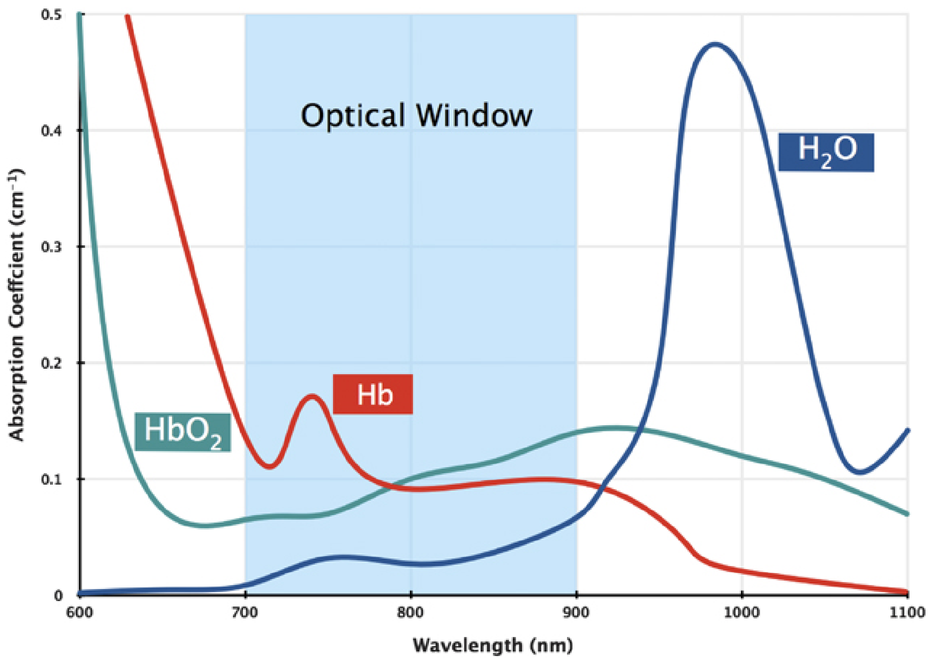


Figure 2.2: Graphs showing the NIR absorption coefficients of hemoglobin (Hb), oxyhemoglobin (HbO2) and water (H2O) [9]

2.2.1 Capturing Hand Vein Images

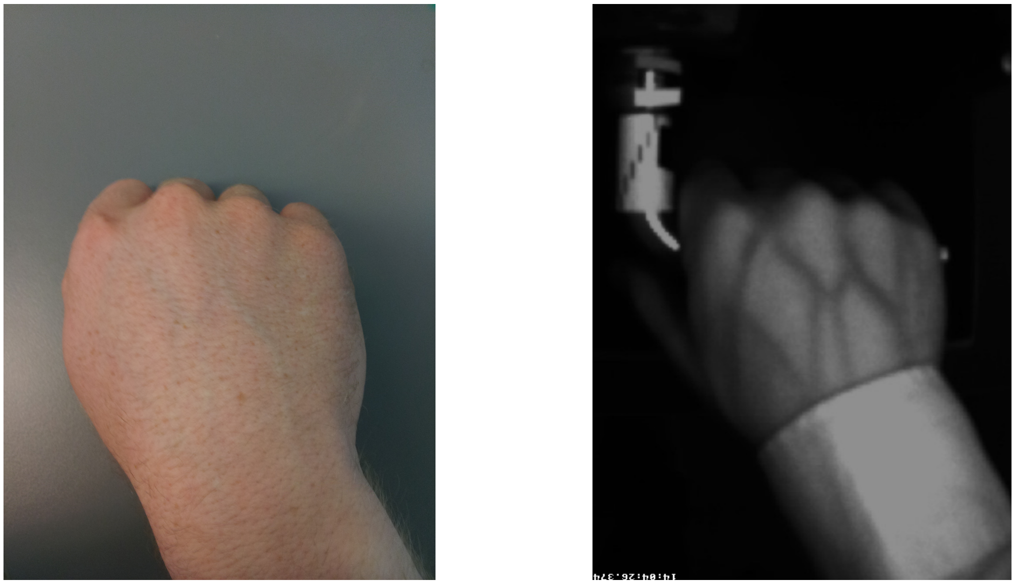
As previously mentioned it is not always possible to see the veins in the hand with the naked eye and it is therefore not possible to rely solely on regular RGB imaging.

Near Infrared Light and Optical Window

Hemoglobin, which is a component in the red blood cells does effectively absorb light in the NIR (Near Infrared) range [9]. Human tissue is mainly composed of water which does not absorb as much light in the NIR range. NIR light can therefore be used to enhance the contrast between vein and tissue at wavelengths where the absorption coefficient of hemoglobin is different from that of water. The range where this is the case is known as optical window and can be seen in Figure 2.2. As seen in the figure the optical window spans from approximately 700nm to 900 nm in the near-infrared range. In comparison the visible spectrum ranges from approximately 400nm to 700nm [10]. The optical window do therefore lie just above the visible spectrum.

Many regular RGB cameras capture light in the visible spectrum and filters out light in the near infrared range [10]. The CCD/CMOS sensors are however sensitive to light in the near-infrared range. This is utilized in many surveillance cameras to achieve night vision by lighting the area of interest with near infrared light. CCD or CMOS cameras without the IR cut filter can therefore be used for capturing hand vein images. This principle is chosen for the acquisition method in this project due to the high availability and relatively low cost of CCD/CMOS sensors. Furthermore IR pass filters can be added to filter out the visible light and thereby get pure NIR images

Figure 2.3 show an example of a hand image captured under regular lighting conditions and the same hand captured with a monochrome camera and near infrared light. These



(a) RGB **(b)** Near infrared

Figure 2.3: Comparison of dorsal hand image captured under normal light with RGB camera and under NIR light with monochrome camera

images show how the veins become much darker than the skin when the hand is illuminated by near infrared light.

Near infrared light can come from various sources such as the sun, LED or halogen lamps. The amount of near infrared light can therefore vary a lot depending on the surroundings. A high energy light source may therefore be required to overpower the ambient near infrared light from the sun and other light sources, to secure even light. The worst case scenario is deemed to be outside in direct sun light. The near infrared irradiance from the sun can be found in the "Standard Table for Reference Solar Spectral Distributions" [11]. This table describes the near-upper limit of solar radiation. From this table it is found that the average irradiance in direct sunlight at 20 deg (which is the worst case in the table), across the optical window, 700nm-900nm is: $0.841W * m^2 * nm^{-1}$. For indoors applications the required amount of light to overpower the sunlight is lower than this number due to some of the light begin reflected and absorbed in glass windows.

The approach of using near infrared light for hand vein recognition has been investigated in other research projects of which the most mentionable are described in the following section.

2.2.2 Hand Vein Biometrics Related Work

An commercial example of vein pattern recognition is the Hitachi VeinID [12] which is a small finger vein scanner. This device can be connected to a computer and used for controlling access to the computer. The device can also be used for access control for doors. VeinID uses near infrared light to capture the vein patterns of a single finger and uses this for identification of a person. This device has en enclosure where the finger is inserted for scanning, which allows for better control over light conditions.

Another similar product is the Fujitsu PalmSecure [3] which is made for scanning the vein pattern in the palm of the hand. In that product the palm vein patterns are described to be unique even for twins. The principle of scanning the veins in the palm of the hand will however require that then sensor is built into the door handle, and can thus not be placed above the door handle.

In [6] some of the current issues regarding hand vein recognition is described. One of the main issues mentioned is the lack of datasets and a standard setup for capturing the hand vein patterns. There are some dataset publicly available such as the Bosphorus hand vein database [13] which consists of hand vein images from 100 individuals. An example from this database can be seen in Figure 2.4

Figure 2.4: Example image from the Bosphorus hand database [13]

What seems to be common for the available datasets are that the light conditions are fully controlled and even lighting has been assured by blocking off ambient light from the sun or other external light sources. Furthermore no datasets are available for the specific scenario of focus in this project, where the vein pattern has to be acquired when the user grabs a door handle. A dataset should therefore be gathered in this project for developing a testing the proof-of-concept.

2.3 Overall Scenario Analysis

The scope of this project is to investigate the possibility of implementing hand vein biometrics for access control on doors. As described in Section *1 Introduction* there are some issues related to the use of physical access tokens for access control. The focus in this project is therefore not on two-factor authentication as some of the current commercial products described in Section *1 Introduction*. Instead the purpose is to investigate possibilities for using biometric features as a standalone, one-factor authentication system for access control.

The scenario of interest is to capture hand vein images when a user grabs the door handle. Based on the hand vein features the proof-of-concept system should be able to decide whether or not to unlock the door. Besides hand vein features, other biometric features could be used as well to determine the identity of the person. This could be common features such as facial information or gait. Another possibility is to track the door handle position and then detect features related to the motion of pulling down the door handle. Such features could possibly be fused with the vein pattern information for recognition.

In this project a proof-of-concept is developed for the above described scenario to determine if hand vein recognition can be used for a physical access control system.

2.4 Problem Statement

Based on the above technical analysis the initial problem statement can be specified by the following:

- *Can a one-factor authentication system for access control be based on the veins in a hand?*
  - *How can the proof-of-concept be developed?*
  - *How high precision and recall can be achieved with such a proof-of-concept?*

3 System Overview

Based on the previous analysis of the problem the following chapter describes the general aspects of the proof-of-concept that has been developed as well as a description of the dataset which has been acquired for developing and testing the proof-of-concept.

The basic concept of the proof-of-concept can be seen in Figure 3.1. As it can be seen

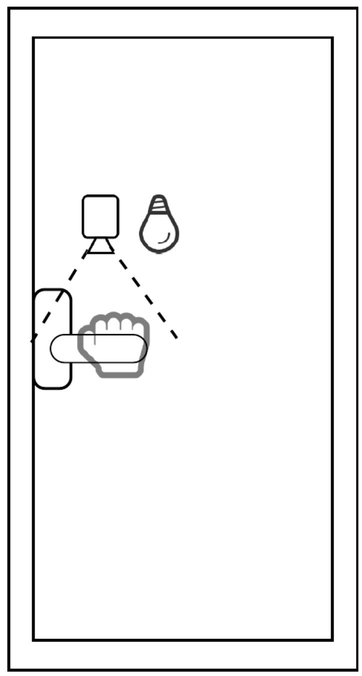


Figure 3.1: Figure showing the overall principle. A NIR light and a CCD/CMOS camera mounted above the door handle, to capture the dorsal hand vein pattern.

from this figure the purpose is to mount a camera and a NIR light above the door handle in order to capture the vein pattern of the dorsal side of the hand, when a user grabs the door handle. Based on the analysis in the previous section some overall delimitations have been made which are described in the following section.

3.1 Delimitations

3.1.1 Processing Time

As described in Section *2.3 Overall Scenario Analysis* feature extraction and feature matching has to run in real-time to be able to unlock the door before the handle is pulled down. As the focus of this project is to analyze the possibilities of this technology in terms of 1:many matching and use for access control, processing time will not be investigated in this project. Processing time will also be dependant on the hardware used in the specific setup.

Real-time processing is however a very important aspect for future research to ensure that the technology has the potential to be used for access control systems.

3.1.2 Motion Analysis

Besides vein pattern features, motion may also be used as a feature as described in Section *2.3 Overall Scenario Analysis*. When the dataset was acquired this has been taken into account and a pattern has been used to enable motion analysis. Motion will however not be implemented nor tested as a feature in this project, but could be added in future work. Examples of motion graphs can be seen in Appendix *A - Motion analysis graphs*.

3.1.3 Hand Vein Uniqueness

Hand vein pattern should be a unique feature to describe and identify persons, but as described in Section *2.2 Hand Vein Biometrics* only limited research has been put into investigating the uniqueness of vein patterns. In this project the dorsal hand vein patterns are presumed to be unique, but this hypothesis is not tested.

3.1.4 Hand Vein Consistency

Veins and the vascular network does change during the lifetime of persons and is caused by natural causes light growth, temperature and such but the vascular network can also be affected by diseases. As described in Section *2.2 Hand Vein Biometrics* the vascular network does only change slightly during adulthood. Furthermore changes to the vein pattern can be negated by regular re-enrollment of persons. This will however not be possible if the pattern changes rapidly and often, but based on the analysis in Section *2.2 Hand Vein Biometrics* this should not happen under normal circumstances. The consistency of hand vein patterns is therefore not tested in this project, but is important to investigate in future development of the system.

3.2 Overall Design and Requirements

Based on the above mentioned delimitations and the analysis of Chapter *2 Technical Analysis* an overall flow for the proof-of-concept has been designed. This overall flow can be seen in Figure 3.2.

The input to the system is a continuous video stream. The video has to be analyzed continuously to detect when a hand grabs the door handle.

When a hand is detected the vein features are extracted and these are matched against the hand vein database containing features corresponding to person who are allowed to enter. The output of the system is based on the result of the matching. If the features match a person with access the door has to be unlocked, otherwise it has to stay locked.

Besides this flow the video could be continuously analyzed to determine the motion of the door handle. Features describing the motion such as acceleration, speed or total time of motion could be fused the vein features to improve the matching. This is, however, as described in Section *3.1 Delimitations* outside of the scope of this project.

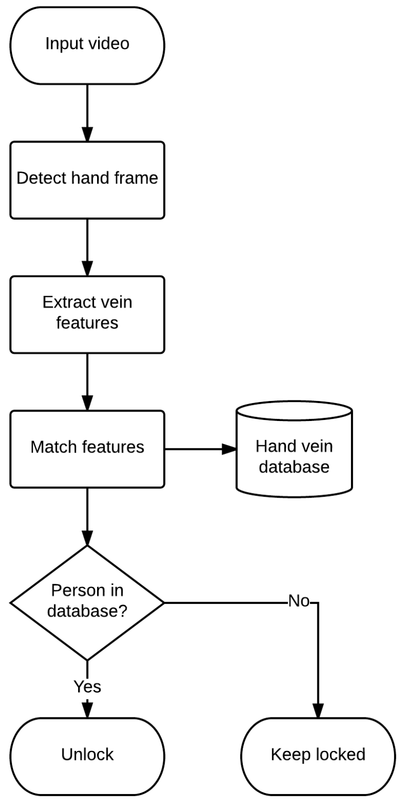


Figure 3.2: Overall program flow of the proof-of-concept

As previously described the required performance of a biometric access control system depends on the specific use case in focus. No strict requirements are therefore set int terms of performance, precision, recall, etc. Instead the focus is to investigate the possibilities and aim to achieve the best possible performance, both in terms of precision and recall. How this is tested will be described in Section *5.1 Acceptance Testing*

3.3 Data Acquisition

As mentioned in Section *2.2.2 Hand Vein Biometrics Related Work* there are a rather limited amount of available hand-vein datasets. Furthermore most of these dataset are captured in very controlled environments, usually in a box where parameters such as hand position and light is controllable. As described previously it will not be possible to completely control light and hand position in the scenario of interest, at a door for access control. In this project a dataset for this specific scenario has therefore been captured, and the process and results are represented in this section.

3.3.1 Acquisition Setup

An acquisition setup has been prepared using a door model which is commonly used for demonstrating locking mechanisms. An image of this model can be seen in Figure 3.3

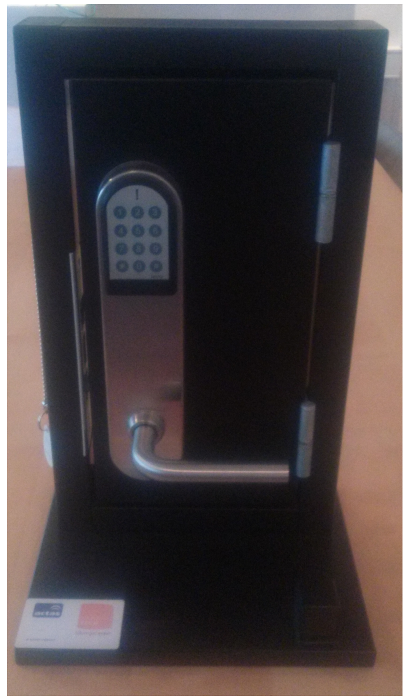


Figure 3.3: Example of the door model used for constructing the data acquisition setup

Camera

The camera used for the data acquisition is a Mobotix DualFlex S15D [14] which is a monochrome surveillance camera. The specifications of the camera can be seen in Table 3.1. This surveillance camera has been chosen as it is capable of capturing light in the near infrared range as described in Section *2.2.1 Near Infrared Light and Optical Window*. Furthermore it is a small form factor camera and can therefore be mounted in a door.

Table 3.1: Mobotix Dualflex S15D specifications

Resolution	5mp
Focal length	1.8mm
Aperture	F2.0

To the camera is attached a NIR 850nm bandpass filter from MidOpt [15]. This filter is attached to remove light outside the otical window as described in Section *2.2.1 Near Infrared Light and Optical Window*.

Required Light

As described in Section *2.2.1 Near Infrared Light and Optical Window* the minimum requirement to the light source is an irradiance of $0.841 W * m^2 * nm^{-1}$. The power of light sources are often noted by the radiant intensity I_v. The inverse square law can be used to calculate the irradiance E_v on a surface given the radiant intensity I_v and the distance to the light source d [16]. The inverse square law can be seen in Equation 3.1:

$$E_v = \frac{I_v}{d^2} \tag{3.1}$$

If the light source is mounted at a distance of 1 meter from the door handle the minimum required radiant intensity is therefore $0.841W$.

For the data acquisition a OSLON Black 4 PowerStar IR [17] light source is used. This LED has a radiant intensity of $2.74W$ and is mounted next to the camera at a distance of approximately $0.3m$ from the door handle. This light source should therefore be more than sufficient to overpower any sunlight and therefore make it possible to achieve relatively even light.

Besides the camera for capturing vein images a Gopro Hero 3 Black [18] has been mounted on top of the door, to capture the persons as they approach the door. This data is not utilized in this project, but can be used in future research regarding multi-modal biometrics. The door model with camera and light mounted can be seen in Figure 3.4.

Figure 3.4: Image of the acquisition setup with camera and light mounted.

3.3.2 Motion Pattern

As mentioned in Section *2.3 Overall Scenario Analysis* another feature to use could the motion when the door handle is pulled down. To be able to track the motion of the door handle a pattern has been mounted to the handle. This pattern can then be tracked to extract the position of the door handle. The pattern is shown in Figure 3.5

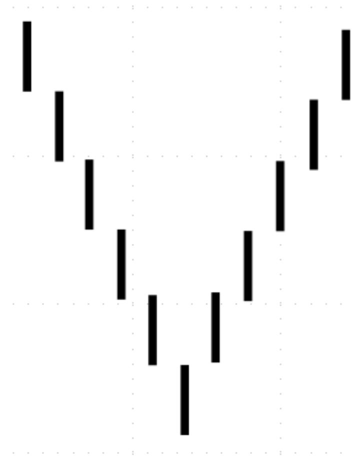


Figure 3.5: The pattern that has been mounted on the door handle to track motion.

3.3.3 Procedure

Using the previously described setup a dataset consisting of 83 individuals has been gathered. Each person has been asked to perform the following specific procedure:

1. Stand on the marked line (1.5m away from the door)
2. Approach the door naturally
3. Grab the door handle and pull down the handle with the intent to open the door
4. Release the door handle
5. Go back to the marked line

Each participant is told to perform this procedure 10 times, resulting in 10 actions per person. The participants are told to go back and forth to emulate a natural motion to avoid that the door handle is grabbed exactly the same way 10 times in a row. During the data acquisition no further instructions than the above is given to the participants. This is done to attempt to force the participants to grab the door handle as they would usually do and to avoid as much bias as possible.

The age and gender are noted for all participants. Furtheremore the participants are asked to sign a form of consent to allow for publishing the data. This form can be seen in Appendix *C - Permission to use biometric Data*.

3.3.4 Dataset Description

A dataset consisting of 83 persons have been acquired in the canteen at Aalborg University. Important numbers describing the dataset can be seen in Table 3.2. Figure 3.6 shows 4

Table 3.2: Dataset specification

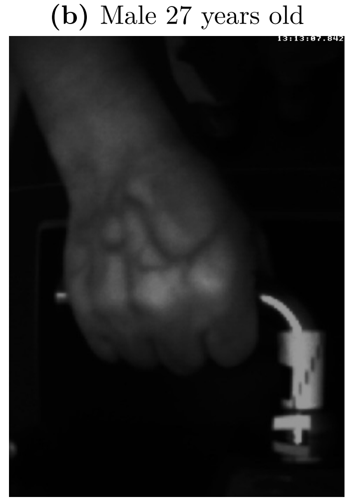
Unique persons	83
Total hand vein samples	830
Males/Females	58/25
Mean age	25.06
Median age	23
Hand	Right
Framerate	20Fps

different samples of the acquired dataset:

(a) Female 24 years old

(b) Male 27 years old

(c) Male 33 years old



(d) Male 60 years old

Figure 3.6: Examples of vein images from captured dataset, for four different persons.

Examples of the face and gait video acquired can be seen in Appendix *B - Face/gait image example.*

As the example vein images show it is possible to see the veins in the hand. There is however uneven illumination across the surface of the hands. This needs to be accounted for by the proposed algorithm, which is described in the following chapter.

4 Hand Vein Extraction and Recognition

This chapter describes the design and implementation of the hand vein extration and recognition.

As described in [19] there are two overall methods for vein recognition, shape based and texture based. Shape based methods cover methods which use the shape of the veins such as line based features or contour features such as area. Texture based methods cover keypoints methods such as SIFT or SURF as well as other texture descriptors such as LBP (Local Binary Patterns).

When looking at the vein images from the dataset described in Section *3.3.4 Dataset Description*, shape based methods may appear to be the obvious way to go as the veins appear to form unique networks or patterns on the hand. But the uneven illumination in the images may cause issues for shapes based methods. An image from the dataset is processed and an optimal threshold is manually found to investigate the vein pattern. Even though a clear vein image is obtained all veins are not continuous. The image can be seen in Figure 4.1, which shows two points where the veins appear to be split up. It may be possible

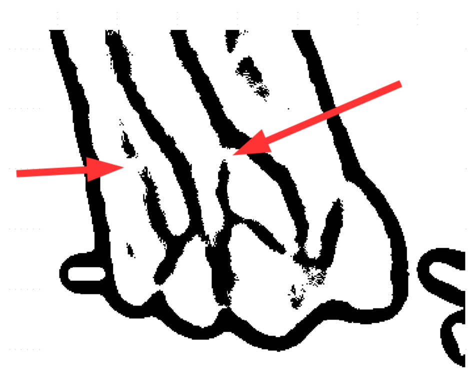


Figure 4.1: Example thresholded image showing broken vein patterns

to solve the issue of broken vein lines, by morphological operations or by line analysis.

Another issue that may however appear in the scenario of interest, is occlusion. Partial occlusion can appear if a hand is a bit tilted, some of the veins may disappear. The chosen algorithm do therefore have to be rather robust towards partial occlusion. This is illustrated in Figure 4.2 where one of the veins disappears as the hand is tilted. Using keypoint detectors and descriptors could therefore be useful as some of these can handle occlusion and cluttering as described in [20].

As mentioned another form for texture descriptor is LBP. Initial test have been performed using the multiscale LBP approach described in [21]. This did however appear to be relatively sensitive to the noise in the captured images. LBP is calculated for each pixel in the image and the features are depending on the difference between each pixel and its neighbourhood. It does therefore seem plausible that it will be sensitive to noise.

Based on the above, an approach which makes use of keypoint detecting and description has been chosen. The detected keypoints may however vary depending on the chosen detector. Furthermore the detected keypoints may not necessarily be related to vein information. The

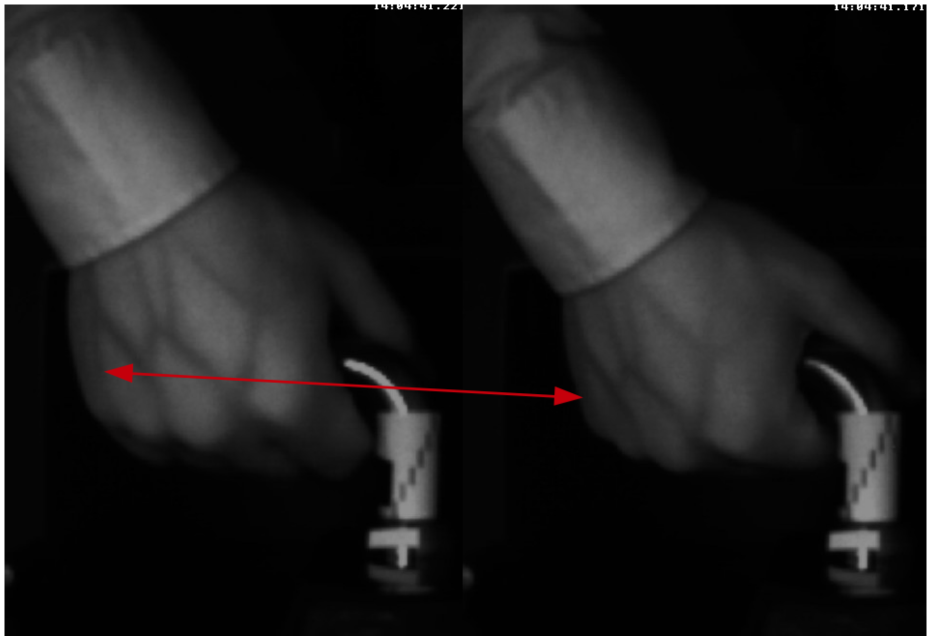


Figure 4.2: Example of vein disappearing due to tilt

proposed approach do therefore include a segmentation step in which the goal is to extract a mask which covers vein areas. This mask is then to be used for the keypoint detector. The detected keypoints are thereby forced to lie along the vein patterns.

4.1 Algorithm Overview

The steps of the proposed algorithm are summarized below:

- Preprocessing
  - ROI extraction
  - Noise removal
- Segmentation
  - Detect and extract hand region
  - Perform contrast enhancement
  - Detect and extract hand vein region
  - Create vein mask
- Feature extraction
  - Detect keypoints
  - Compute keypoint descriptors
- Decision
  - Feature matching with samples in the database
  - Calculate match measure
  - Allow or deny sample based on match measure

The following sections describes the different steps in more detail.

4.2 Preprocessing

For extracting and matching vein pattern features, it is necessary to grab specific frames from the video where the dorsal side of the hand is clearly visible and is not rotated or tilted. The hand will have the desired position and orientation just before pulling down the handle. Frames of interest are therefore extracted using the motion analysis illustrated in Appendix *A - Motion analysis graphs*. Whenever the start of a motion is registered the previous frame is extracted and saved. In this project the extracted frames are manually verified to contain a clear image of the dorsal side of the hand. This is done ensure that the issues related to the motion analysis will not influence the vein extraction and matching. An example extracted frame can be seen in Figure 4.3. The door handle will always have the

Figure 4.3: Example frame of interest, to be used for vein pattern extraction and matching

same position in the images and this has therefore been used to create a region of interest which can be seen in Figure 4.4. As it can be seen from Figure 4.4 and Figure 4.3 the images contains some noise. To decrease the amount of noise in the image median blur is performed as the last step of preprocessing.

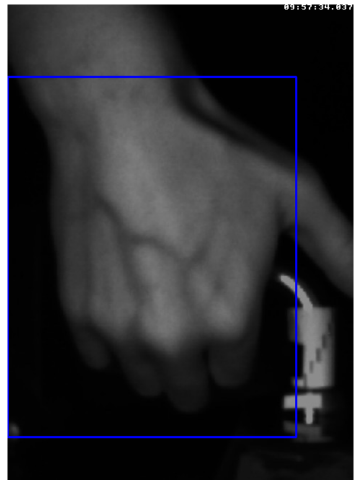


Figure 4.4: Region of interest

4.3 Segmentation

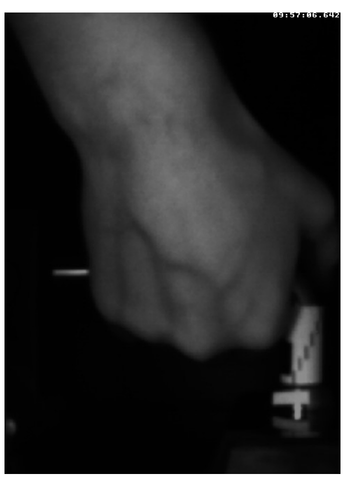
The purpose of the segmentation step is to extract an image mask, for the keypoint detection, which contains the veins in the image. Simple thresholding cannot be utilized as the light is not even across the image. An example can be seen in Figure 4.5, where simple thresholding is applied to the image shown in Figure 4.3. After thresholding some veins are clear, but a lot of information is lost as the light is uneven. To avoid issues regarding

Figure 4.5: Simple threshold applied to the original image

uneven illumination an adaptive threshold, also known as local threshold, can be used instead. In this type of thresholding different threshold values across the image by calculating the threshold value based on neighbourhood values. Two common ways of calculating the threshold value is to use the mean of the neighbourhood values, or to calculated a weighted average based on a gaussian kernel. An example of adaptive gaussian thresholding can be seen in Figure 4.6. By using the adaptive threshold most veins are extracted, but it also

Figure 4.6: Adaptive gaussian thresholding applied to the original image

includes some noise from the surroundings. Furthermore the shape/edge of the hand is also extracted and will therefore appear as a significant feature. For this project the shape of the hand is however not an interesting feature as it will vary a lot depending on the hand position and how tight the hand is grabbing the door handle. An example of this is seen in Figure 4.7, which shows two consequent frames of the same hand. In this example the hand shape varies a lot even though it is the same hand.



(a) Current frame

(b) Previous frame

Figure 4.7: Example of two consecutive frames of the same hand

Preliminary test have also shown that the keypoint detector finds the most significant features along the edge of the hand. The edge of the hand and the surrounding noise is removed by creating a mask that contains only veins, of which the first step is to detect the hand region.

4.3.1 Hand Region Detection

To detect the hand region Otsu thresholding is performed on the input image. Otsu thresholding [22] is used to get a binary image with the hand separated from the background. Otsu thresholding assumes a bimodal distribution and is therefore well suited to extract an object from the background. The result of Otsu threshold is shown in Figure 4.8

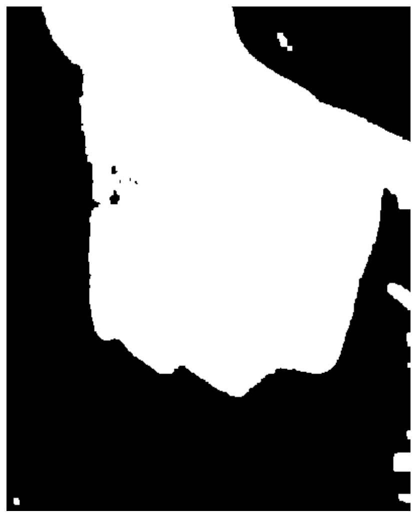


Figure 4.8: Otsu threshold applied to ROI

The largest object in the ROI should always be the hand, and the hand can therefore be extracted from the binary image by finding the largest contour and extracting only that from the image. This is shown in Figure 4.9.

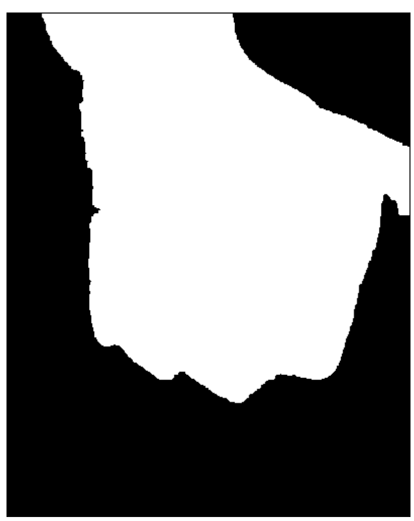
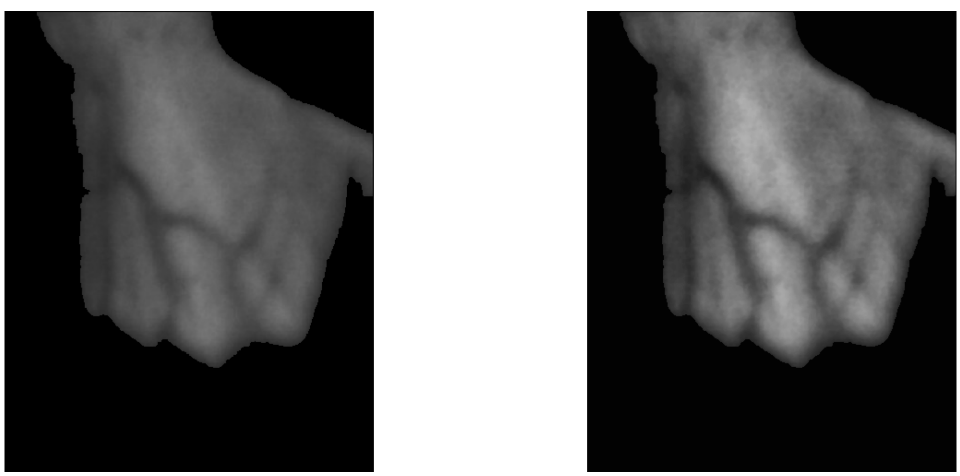


Figure 4.9: Largest contour extracted from the ROI image

To get a mask that covers the entire hand dilation is performed on the binary image to enlarge the hand contour. The resulting binary image can then be used as a mask for vein extraction.

4.3.2 Vein Extraction

To make the veins appear clearer Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied. This algorithm performs histogram equalization in tiles across an image. Furthermore the contrast is limited within the tiles to minimize the effect of noise. The extracted dorsal hand image before and after CLAHE is shown in Figure 4.10. After enhancing the



(a) Extracted dorsal hand image **(b)** CLAHE applied

Figure 4.10: Example of the extracted dorsal hand image before and after CLAHE is applied.

contrast adaptive gaussian thresholding is applied to extract the veins. As previously described the result of applying adaptive threshold will also include the edge of the hand. This edge is removed by applying an eroded version of the hand mask. The extracted vein mask can be seen in Figure 4.11.

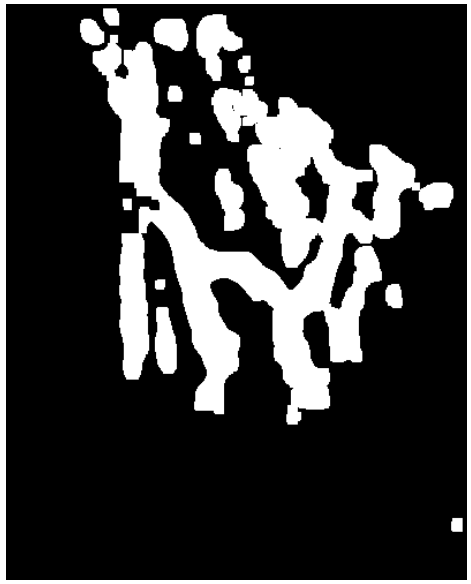


Figure 4.11: Image showing the vein mask

The vein mask is then used for features extraction.

4.4 Feature Detection and Description

Feature points are detected using a SURF detector [23]. Initial testing has been done with Sift, SURF and Orb descriptors where SURF proved to detect most keypoints along the veins. This could be explained by the fact that the SURF detectors rely on Haar wavelets. Haar wavelets seems ideal to represent the intensity changes between skin and veins in the image. The detected SURF keypoints can be seen in Figure 4.12

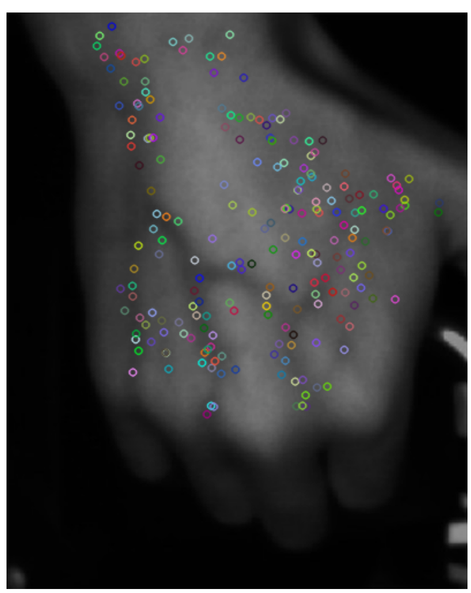


Figure 4.12: Detected SURF keypoints drawn on the ROI image

The SIFT descriptor has been used as keypoint descriptor. This has been done based on results of initial testing which showed best performance with the SIFT descriptor. Furthermore the dimension of the SIFT descriptor is 128 while it is just 64 for SURF and the SIFT descriptor is therefore deemed more distinctive than SURF. SURF can however be extended to a dimension of 128 and further test with the SURF descriptor could therefore be made in the future.

4.5 Feature matching

Feature matching has to be performed in order to determine how similar a test sample is to any of the classes (Persons) in the database. Feature matching is however not a simple task when there are multiple training samples per class. One way of solving this could be to combine the features from all training sample per class into on feature descriptor and then use that for matching. A simpler approach is to match against all training samples individually and then combine the matching scores into one. Investigating and implementing algorithms for combining features is out of scope of this project and combine scores is therefore chosen instead.

Feature matching between each test sample and each training sample is done using a brute force KNN matcher. For each feature in one image the two best matches in the other image are thereby found. To ensure that it is a good match the ratio test, described in the SIFT paper [20], is implemented. The purpose of the ratio test is to discard feature matches that are not distinctive. When features are matched the distance between the matching features is also found. The distance to the closest match and the distance to the second closest match are compared. The match is then discarded if the best match is less 25% better than the second closest match.

After all matches between two images are found the median distance of the matches is calculated. The median distance is used to negate the effect of outliers. For a test sample the average of median distances to all training samples in a class are calculated. This is done to produce a single measure describing the match between a test sample and a training class.

The measure describing the match between a test sample and a class is used for making the decision whether the test sample matches a authorized person and should then be allowed to enter. The decision threshold has to be chosen based on the requirements of the specific scenario. A low threshold will result in more false positives and less false negatives. This can therefore used for producing precision-recall curves in the test in the following chapter.

5 Evaluation

5.1 Acceptance Testing

This section describes the test which have been performed in order to test the proof-of-concept and serves to answer the problem statement described in Section *2.4 Problem Statement*:

A proof-of-concept has been developed to determine whether a given test image corresponds to a known authorized person. The procedure and metrics used for testing the performance of this proof-of-concept is described in the following sections.

5.1.1 Procedure

The test of the developed proof-of-concept is tested using the acquired dataset described in Section *3.3.4 Dataset Description*. The dataset has been split 7:3, with 7 samples from each person for training data and 3 samples for test data. This split has been chosen to have a high amount of training data for each person while still having a few test samples per person. Table 5.1 describes the testing data: An important note is that error during data

Table 5.1: Testing dataset

	Per person	In total
Training samples	7	581
Test samples	3	245

acquisition for four persons resulted in only nine repetitions instead of ten. For these test persons there are only 2 test samples while there are still 7 training samples. This explains why there are a total of 245 test samples.

Each test sample is matched against each class, which each represent a single person, and the algorithm output is either a match or a non-match. As described in Section *4.5 Feature matching* the algorithm will only output a match if the distance measure is below a certain threshold. As all test samples are compared to all training classes, ideally there should be 1 match and 82 non-matches for each test sample.

If the output for a test sample is a correct match it is counted as a TP (True Positive). If the output is a match with a wrong class it counts as a FP (False Positive). In case the output is a non-match when the class does not match it counts as a TN (True Negative). Finally a FN (False Negative) is if the output is a non-match when compared to the correct class. To further elaborate this means that a perfect system should have only TPs and TNs. FPs are cases where a persons is classified as a wrong person while FNs are the false absence of a match.

Common metrics for evaluating the performance of a biometric recognition system are FAR (False Acceptance Rate) and FRR(False Recognition Rate)[24]. FAR describes the number of impostors that are accepted as a valid person, while the FRR describes the amount of valid persons which are falsely rejected. FAR is calculated as:

$$\frac{FP}{FP+TN}$$

and FRR is calculated as:

$$\frac{FN}{FN + TP}$$

The FAR metric can however be a bit biased if the test is designed to result in a large number of non-matches, which is the case for the test in this project. With all 245 test samples being matched against all 83 classes there should be 245 matches and 20335 non-matches. The two additional metrics precision and recall are therefore also used. Precision is calculated by:

$$\frac{TP}{TP + FP}$$

This metric describe the relationship between correct matches, TPs, and the total number of matches in the output, correct and incorrect. Recall, on the other hand, is a description of the relationship between correct matches, TPs, and the total number of matches in the dataset.

$$\frac{TP}{TP + FN}$$

It is important to note that FRR and recall describe the same relationship and $Recall = 1 - FRR$.

$$Recall = 1 - FRR$$

In terms of an access control system, high precision describes a system that rarely allow impostors to enter. A system with a high recall will rarely reject an authorized person.

No known research has been done to this specific scenario of hand vein recognition at doors for access control. It is therefore not possible to directly compare to current state-of-the-art methods. Furthermore many of these addresses identification problems and measure the performance in terms of rank one recognition rate, such as in [19]. The rank one recognition rate is used for closed form identification where each test sample has to be matched with an identity in the database. Therefore the performance of the developed proof-of-concept is measured in terms of the measure described above. The results is shown and described in the following section.

5.1.2 Results

Table 5.2 shows the results for varying decision thresholds as described in Section *4.5 Feature matching*. A plot of the precision and recall of the results can be seen in Figure D.1 in

Table 5.2: Test results

TP	FP	TN	FN	Precision	Recall	FAR	FRR
197	280	19810	48	0.413	0.804	0.014	0.196
194	147	19943	51	0.569	0.792	0.007	0.208
190	107	19983	55	0.640	0.778	0.005	0.224
181	71	20019	64	0.718	0.739	0.004	0.261
152	28	20062	93	0.844	0.620	0.001	0.380
117	8	20082	128	0.936	0.478	0.000	0.522

Appendix *D - Precision-recall curve.*

The results in the table and graph illustrates the trade-off between precision and recall. As previously described the required precision and recall depend a lot on the use-case of interest. High security will require high precision to reject impostors, while higher recall is of interest for low-security application to ensure authorized persons are always allowed. The best compromise between precision and recall is in this test the case with precision: 0.718 and recall: 0.739.

In this case there are 64 false negatives in total. Further analyzing the output in this case shows that these false negatives are divided amongst 40 individuals. For 9 persons all three test samples result in FPs. This indicates that some persons are generally difficult to identify and all FNs are not just a result of single bad images. Figure 5.1 shows two test images, both from persons which results in FNs for all test samples. In these two examples the veins are also relatively difficult to spot visually, and it may therefore be difficult to find distinctive vein features from these images.

The FPs in this case are distributed amongst 25 different training classes (persons). This describes that there are rather few training classes that are responsible for many FPs. These classes may therefore contain samples that contain features which are not distinctive for that class only. An example of samples from two different persons which cause many FPs are shown in Figure 5.2. In these two examples the veins are neither very clear which further indicates that the cause for FPs in these cases are due to the lack of distinctive features in the captured images.

The example images of cases with FPs and FNs indicates that some issues may be caused by the quality of the captured images, as the veins do appear very clearly in those. Further analysis do therefore have to be put into these issues to investigate how the FPs and FNs are related to the image quality and the chosen method.

(a) Female 22 years old

(b) Male 22 years old

Figure 5.1: Example of images from two different persons which results in FNs

(a) Male 24 years old

(b) Female 23 years old

Figure 5.2: Example of images from two different persons in the training samples, which cause many FPs

5.2 Conclusion

Conventional access control systems often rely on physical access tokens such as key-cards, mechanical keys or chips. This is however not always convenient and can cause issues. An example of this is care homes for elderly where there is the risk that the residents often forget their physical access tokens. A possible way of solving this issues can be by the use of biometric recognitions systems. Some biometric access control systems do already exist, but most of these require active interaction with the user. This project has researched the possible use of dorsal hand vein recognition to solve this, with the overall focus described by the problem statement:

- *Can a one-factor authentication system for access control be based on the veins in a hand?*
  - *How can the proof-of-concept be developed?*
  - *How high precision and recall can be achieved with such a proof-of-concept?*

A proof-of-concept has been developed which relies on a monochrome surveillance camera and a near infrared light source. This has been mounted on a door model which has been used to acquire a dataset for testing the proof-of-concept. By illuminating the dorsal side of a hand, when it grabs the handle, images are captured with the veins in the hand being visible. No available dataset were fit to the scenario of interest and a dataset has therefore been acquired which contains 83 individuals with 10 samples for each person.

The best achievable precision and recall is deemed to be the case which results in a precision of: 0.718 and a recall of: 0.739. The exact required precision and recall depend on the exact use-case, but the precision and recall for the current proof-of-concept is not deemed to be sufficient for deployment. Along with the relatively good false acceptance rate the results do however show that the solution has potential to accurately identify persons.

It is therefore deemed possible to base a one-factor authentication system on hand vein recognition. The current proof-of-concept can however be improved in multiple ways which is discussed in the following section.

5.3 Discussion and Future Work

5.3.1 Dataset and data acquisition setup

The dataset that has been acquired consists of vein images as well as video of the motion when the door handle is grabbed. Furthermore the dataset also includes face and body video of the participants as they approach the door. This dataset do therefore include data for at least 3 different modalities. Hand vein and hand motion can be extracted from the monochrome NIR videos while facial information and behavioural information such as gait can be extracted from the RGB video. The dataset do therefore provide good basis for future research within the subject of biometrics for access control.

The data acquisition setup consists of a regular monochrome surveillance camera and an infrared LED. The acquisition setup is therefore rather simple and has the potential to be deployed on a real door. The acquired data do however suffer from uneven lighting and unclear veins in some images. This could be caused by the infrared led used which has a high intensity and only limited light spread. Better quality images could therefore be possibly be acquired by using a larger light source, or by the use of a diffuser.

5.3.2 Feature Extraction and Matching

The test have shown that there are some issues with extracting and matching features. Further investigation could therefore be put into the choice of keypoint detector and descriptor. SURF has proven to be the best detector in initial test while the SIFT descriptor resulted in the most distinctive feature descriptors. It is important to note that these algorithms are patented in some countries. For future development it will therefore be interesting to further investigate the performance of non-patented algorithms such as ORB [25].

Another possibility for improving the performance of the feature extraction and matching could be to fuse the keypoint features with shape based features. As previously described an advantage to using keypoint is the robustness to occlusion which is not necessarily the case for shape based methods. Based on this a fusion of the two types of features might therefore improve the performance. Another possible feature to use for feature fusion could be knuckle prints, as described in [26] which could potentially be acquired by the same camera used for hand vein images.

The matching procedure used in this project is rather simple and could therefore be improved in future work. Currently the matching is based on one-one matching between all test and training samples, and mean distance is then calculated for each training class and test sample. This approach may however be rather sensitive to single bad images in the training sets as this will affect the overall match for the class. A method for combining features from multiple image into one feature space should therefore be investigated in future work. One possible approach could be a bag of words approach for selecting distinctive features based on multiple training images.

5.3.3 Overall Performance

In terms of precision and recall the performance of the developed proof-of-concept is not ideal, but it gives a clear indication that the hand vein patterns can be used for identification for access control. This is emphasized by the relatively low false acceptance rating. The low FAR means that the system rarely matches a sample with a wrong identity. There are however issues regarding determining whether or not a sample is a given person or not. This is rather important in an access control system to be able to rely on the decisions made by the system. As the proof-of-concept performs better in terms of FAR than precision indicates that the solution is better for closed form identification than authentication. In closed form identification a test sample has to be matched to a identity in the database whereas in authentication a sample may not necessarily match a identity in the database.

5.3.4 Hand Vein biometrics for Access Control

Even though the results indicate that the solution may be better suited for closed form identification the principle of hand vein recognition is deemed plausible for access control systems. As the developed proof-of-concept can achieve precision of: 0.718 and a recall of: 0.739 it seems possible to further improve this in future development of the system. Furthermore this technology could be using in conjunction with other biometrics to achieve better performance. For example face and gait recognition could possibly be applied as a person approaches the door. Such features could be fused with hand vein recognition to improve the matching. Ensemble methods could possibly be used to filter possible matches as persons approaches the door.

Another interesting feature is the hand motion analysis which is also mentioned previously. Applying this may however conflict with the mechanical locking mechanisms that needs to be unlocked before the door handle is pulled all the way down. Nevertheless this feature do however seem to have the potential to be used for identifying persons. Furthermore this type of behavioural biometric could also contain information about mood, health state of mind and such. It is expected that the motion will reflect if persons are in a hurry or have plenty of time.

Bibliography

[1] H. Access, "Handkey," https://www.honeywellaccess.com/products/access-control-systems/readers/biometric/52198.html.

[2] OT-Morpho, "Morphoaccess® sigma lite," https://www.morpho.com/en/biometric-terminals/access-control-terminals/fingerprint-terminals/morphoaccess-sigma-lite-series.

[3] Fujitsu, "Fujitsu palmsecure id match biometric multi-factor authentication," http://www.fujitsu.com/global/solutions/business-technology/security/palmsecure/id-match/, last visited: 06.06.2017.

[4] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007.

[5] A. M. Badawi, "Hand vein biometric verification prototype: A testing performance and patterns similarity." *IPCV*, vol. 14, pp. 3–9, 2006.

[6] S. Crisan, "A novel perspective on hand vein patterns for biometric recognition: Problems, challenges, and implementations," in *Biometric Security and Privacy*. Springer, 2017, pp. 21–49.

[7] A. Nadort, "The hand vein pattern used as a biometric feature," *Master Literature Thesis of Medical Natural Sciences at the Free University, Amsterdam*, 2007.

[8] P. Carmeliet and R. K. Jain, "Angiogenesis in cancer and other diseases," *nature*, vol. 407, no. 6801, pp. 249–257, 2000.

[9] T. G. Phan and A. Bullen, "Practical intravital two-photon microscopy for immunological research: faster, brighter, deeper," *Immunology and cell biology*, vol. 88, no. 4, pp. 438–444, 2010.

[10] T. B. Moeslund, *Introduction to video and image processing: Building real systems and applications*. Springer Science & Business Media, 2012.

[11] A. G197-08, "Standard table for reference solar spectral distributions: Direct and diffuse on 20° tilted and vertical surfaces," 2008.

[12] Hitachi, "Veinid," http://www.hitachi.eu/veinid/aboutveinid.html, last visited: 05.06.2017.

[13] H. Dutağacı, B. Sankur, and E. Yörük, "Comparative analysis of global hand appearance-based person recognition," *Journal of electronic imaging*, vol. 17, no. 1, pp. 011 018–011 018, 2008.

[14] Mobotix, "Mobotix dualflex s15," https://www.mobotix.com/other/Products/Outdoor-Cameras/DualFlex-S15-S16, last visited: 06.06.2017.

[15] Midopt, "Bn850 near-ir bandpass filter," http://midopt.com/filters/bn850/, last visited: 06.06.2017.

[16] A. E. Taylor *et al.*, “Illumination fundamentals,” *Lighting Research Center, Rensselaer Polytechnic Institute*, 2000.

[17] Osram, “Oslon black 4 powerstar ir,” https://www.osram-os.com/osram_os/en/products/product-promotions/infrared-products/ir-oslon-black-family/index.jsp, last visited: 06.06.2017.

[18] GoPro, “Gopro hero 3+ black edition,” https://gopro.com/, last visited: 06.06.2017.

[19] D. Huang, X. Zhu, Y. Wang, and D. Zhang, “Dorsal hand vein recognition via hierarchical combination of texture and shape clues,” *Neurocomputing*, vol. 214, pp. 815–828, 2016.

[20] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[21] X. Guo, W. Zhou, and Y. Zhang, “Collaborative representation with hm-lbp features for palmprint recognition,” *Machine Vision and Applications*, vol. 28, no. 3-4, pp. 283–291, 2017.

[22] N. Otsu, “A threshold selection method from gray-level histograms,” *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[23] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” *Computer vision–ECCV 2006*, pp. 404–417, 2006.

[24] T. Dunstone and N. Yager, *Biometric system and data analysis: Design, evaluation, and data mining.* Springer Science & Business Media, 2008.

[25] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “Orb: An efficient alternative to sift or surf,” in *Computer Vision (ICCV), 2011 IEEE International Conference on.* IEEE, 2011, pp. 2564–2571.

[26] D. Kusanagi, S. Aoyama, K. Ito, and T. Aoki, “A practical person authentication system using second minor finger knuckles for door security,” *IPSJ Transactions on Computer Vision and Applications*, vol. 9, no. 1, p. 8, 2017.

Appendices

A Motion analysis graphs

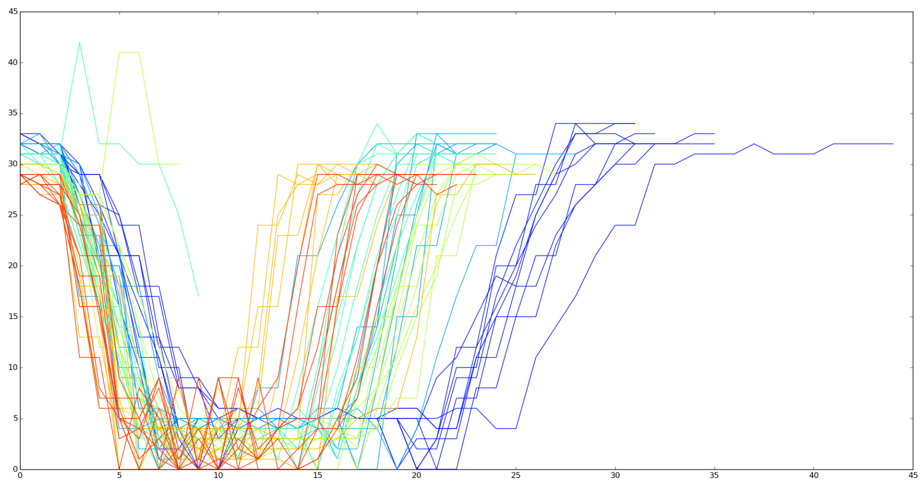


Figure A.1: Example graph showing door handle pixel position (y-axis) over frames since start of motion (x-axis). Each color is a unique person and each graph is a unique motion

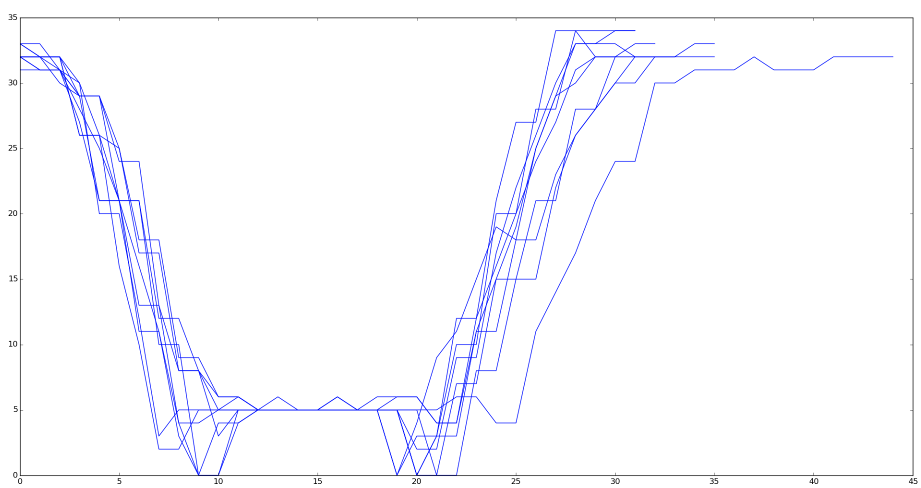
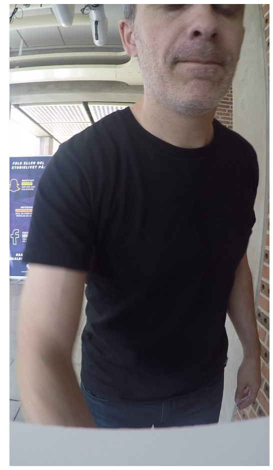


Figure A.2: Example graph showing door handle pixel position (y-axis) over frames since start of motion (x-axis), for one person. Each graph is a unique motion.

B Face/gait image example

(a) Example 1



(b) Example 2

(c) Example 3

(d) Example 4

Figure B.1: Example snapshots from face/gait video recorded

C Permission to use biometric Data

Permission to use Biometric Data for Research Purposes

Att.
Aalborg University
CVR. no. 29102384
Department of Architecture & Media Technology
Post Office Box 159
9100 Aalborg
Denmark

I ………………………………………………………………………… hereby give permission to Visual Analysis of People (VAP) Laboratory at Aalborg University to collect my facial and hand-vein images in two different modalities including RGB and Near-infrared, and use it only for research purposes.

The VAP lab can re-distribute the collected data to academic partners after getting the partners to sign the attached 'terms of use'. The 'term of use' will emphasize that the interested academic partners are not allowed to re-distribute the data to third parties and they can use the data only for research purposes.

- € Besides the above use, I also give the permission to users of the database to publish my facial and hand-vein images in their scientific articles or websites which report their scientific achievements.
- € I DO NOT give the permission to users of the database to publish my facial and hand-vein images in their scientific articles or websites which report their scientific achievements. They can use my data only when the data itself is not published by the users.
- € I DO NOT give the permission to use or publish my facial images. They can use and publish my hand-vein images only.

…………………………………………………………………

Date and Signature

Terms of use
VAP Face and Hand-vein Database

Att.
Aalborg University
CVR. no. 29102384
Department of Architecture & Media Technology
Post Office Box 159
9100 Aalborg
Denmark

I ………………………………………………………………. hereby agree with the following terms and conditions of using Visual Analysis of People (VAP) Face and Hand-vein Database:

- I am not allowed to redistribute the database to any third parties.
- I confirm that granting access to the database can only be done through VAP lab.
- I confirm that the VAP Face and Hand-vein Database will be used only for research purposes and not for commercial ones.
- I should not reveal the real names of the subjects, in case I know any of them.
- I accept to indemnify Aalborg University against any claims from the subjects resulting from the my breach of these Terms of use.
- I'll cite the following work in all my publications which uses the VAP Face and Hand-vein database: The work of VAP, VBN, published 2017 (This will be completed later on!)

……………………………………………….
Date

………………………………………………………………………………………
Name and Position

………………………………………………………………………………………..
Department

…………………………………………………………………………………………
University

D Precision-recall curve

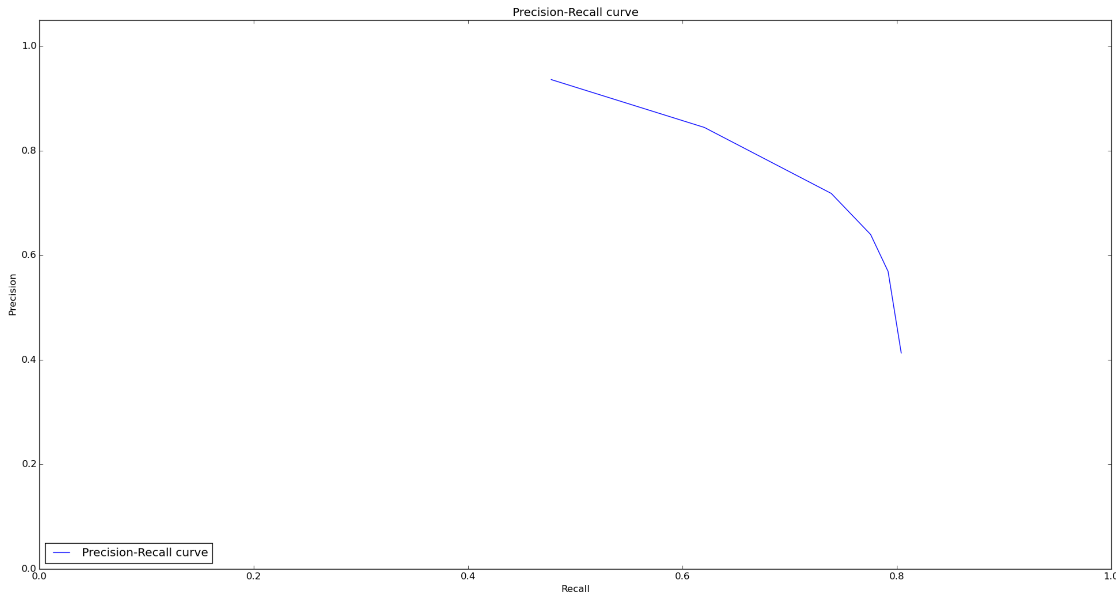


Figure D.1: Precision-recall