Thermal imaging of ear biometrics for authentication purposes

Knut Steinar Watne

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

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

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

Thermal imaging of ear biometrics for authentication purposes

Knut Steinar Watne

30th June 2008

Abstract

Authentication is a much used measure in order to keep impostors from getting access to things or places that they are not supposed to get access to. Passwords/phrases, tokens, keys and the like have been used since ancient history, but also biometric features such as fingerprints, retina scans or even DNA have been applied. This thesis will look at the ear as a biometric feature, and how thermal images may improve the performance of such authentication systems.

We look at thermal images, to see if they give better results than visible images. Thermal images have the advantage that they do not need light to give results, as they only capture heat radiation from the ears. This project however, was done in a standard enlightened environment to see if they could produce better results there as well.

The experiment consists of 75 participants of both students and staff of Høgskolen i Gjøvik. Each participant had 3 sessions with 5 images per session taking the total number of captures to 1125. Each capture produced 3 images (visible, grayscaled and colored).

Acknowledgements

First of all, I would like to thank God for always supporting me throughout my entire life. My life would be nothing if not for Him. I would also like to thank my supervisor Patrick Bours who has been an immense support during the project period, giving me tips and helped me in time of need. Katrin Franke for helping me with some image processing, Tormod Emsell Larsen for sharing views on the subject, Anne Therese Petersen for renting me the thermal camera, Roy Erlend Berg and Christopher J. Fullu for helping me with some MatLab scripts, Kennet B. Fladby for valuable input from the initial draft, Kjetil Holien for helping me with Microsoft Excel and Rune Hammersland for making me coffee during those tiring hours. I would also thank all my participants who volunteered for my experiment, this would not have been possible without you! Last but not least, I would like to thank my fiancée for her love, support and understanding during the project period. Especially those evenings when I did not have the time to come home for dinner.

Knut Steinar Watne, 30th June 2008

Contents

List of Figures

List of Tables

1 Introduction

To determine whether a person is who he says he is, has been important since the dawn of man. In ancient Egypt, Pharaoh's men measured the height of people to identify[1] them. Through the different eras, new ways of authentication[2] have been introduced. All of these are based on either something you *have*, *know* or *are*. In computer science, something you *have* can be a smart card or a token of some sort. Something you *know* can be a password, pass phrase or some secret as a way of logging into a system. The last category, something you *are*, is mostly biometric[3] oriented and includes e.g. fingerprints, iris and facial features. During the late 1960's, new biometric devices were made to measure different body parts for identification. From that time on, a lot of research has been done in this area and today many biometric methods are commonly used for authentication.

1.1 Topic covered by the project

To authenticate a person by something he *is* will be the focus of this project. Facial recognition has been discussed for quite some time now, but a relatively new and undocumented area in relation to this, ear recognition, will be the topic of this master thesis. Different techniques have been defined for comparing ears. We will look into some of them, and try to find new and better ways of using the ear as a "something you are" measure in respect to authentication.

1.2 Keywords

Authentication, ear biometrics, ear shape, human identification, outer ear images, identification rates, thermogram, pattern recognition

1.3 Problem description

Face recognition is by now well documented and tested. This requires an image of the entire face, and the image is in general taken from the front of the person. By just taking the image of an ear, one reduces the size of data that needs to be verified. This is an advantage because by reducing the size, the processing of the data will go quicker. Ear recognition is a new field of research though, and not much work has been done on this topic. Some of the techniques that have been described are weak and do not really give us any reason to believe that ear recognition might be a new, accurate way of authenticating people (Principal Component Analysis(PCA) only had a 71,6% recognition rate [1, 2]). These methods are very dependent on the quality of the image, if the ear is covered by hair or if the head is tilted either way. There are however some methods that produce better results as shown in Table 1 in [3], e.g. the "Force Field Transformation method" [4, 5, 6] with a recognition rate of 99,2%. We will apply some of these techniques to thermal images, and with this master thesis we hope to consolidate ear recognition as a reliable

[1] Identification is defined as establishing an identity
[2] Authentication is defined as verifying the person to be who he claims to be
[3] Greek: *bios* which means "life" and *metron* which means "measure"

authentication method. By looking at thermal images compared to normal images, we hope to achieve results that are feasible for authentication as well as overcoming the current issues concerning lighting conditions and occlusion by e.g hair.

1.4 Justification, motivation and benefits

Ear recognition is a new field of interest for authentication researchers. The theories that have already been described by different sources have results that varies from 71% to 99% in recognition rates, depending on which theory applied [3]. By applying some of these techniques on thermal images, we hope to present ear authentication as a possible supplement to other existing authentication measures. Taking photos of ears is much more comfortable for the person which is to be authenticated than e.g. iris/retina scans or fingerprints. Retina scanning requires the person to look into the camera for a couple of seconds, a process that many people find uncomfortable. Some individuals are reluctant to put their finger on a scanner where plenty of people have had their "dirty" fingers before them. To be authenticated just by a photo of one's ear would be much more people-friendly. Face recognition can be argued as a solution that also fits into this category. But there have been quite a few problems with face recognition. The major problem is that people can have different facial expressions that makes the images very different from one capture to the next. By taking images of the ear, we reduce the problem with facial expressions. By using a thermal camera we also exclude the lighting condition problem. Images taken by a normal camera will give different results as they are dependent of visible light to provide quality. The thermal camera does not take images within the visible light spectrum, thus removing this problem.

It is also easier for the person to "remember" his ear than it is to remember a password, or to bring some sort of token with him. If we can come up with some good results for this new method of ear authentication, a quick picture of a persons ear can be all that is needed for entering a building, getting access to files, or whatever we are protecting. Research has also been done on blood vessel patterns in the ears of mice [7]. This might be interesting to look more at in the future, but will not be in the scope of this thesis.

1.5 Research questions

In order to solve these problems, several research questions need to be addressed.

1. General research questions:
   - What techniques do already exist for ear authentication?
   - How good are our results using thermal images vs. normal images compared to results using existing methods?
2. Research questions related to thermal images:
   - Is it possible to determine differences in peoples ears by looking at a thermal image?
   - Do environmental factors, e.g. temperature, affect our results in any way?
   - How long does it take before an ear is warmed up after being cooled down for some time?

1.6 Claimed contributions

Hopefully, this master thesis will make ear authentication a hot topic in the near future. By proposing a new method, we hope to come up with results that can be of use for later research. We have used a thermal camera to get images of different ears. The camera captures both visible as well as thermal images. From this we have tried to extract characteristic features present in the ear. We choose to take 2D images as this does not require expensive equipment. Time needed to analyze the images created by such a camera is also greatly reduced because the size is smaller than that of a 3D image. A thermal image might reduce the problem with hair covering the ear. This will be very valuable in terms of passive identification[4]. During active identification[5] this will be less of a problem, as the subject can pull his hair back and proceed with the authentication process. People with short hair might still prove difficult though, as hair can be long enough to occlude the ear but not long enough to be pulled behind it. If these two methods can be combined, and give us good results in terms of recognition rates, we will have shown that ear recognition can be a serious competitor for any authentication process.

1.7 Outline of the report

The following chapter is an introduction to authentication, biometrics and biometric systems. The chapter is aimed at people who might not know that much about biometrics or authentication at all, and will give a brief explanation of different terms and components of a biometric system. In Chapter 3 we will look at previous work related to ear biometrics. Chapter 4 will describe how the experiment setup was done, while Chapter 5 will explain the pre-processing, analysis and the results we got from the experiment. Discussion around the results, will be given in Chapter 6. Our final conclusions are presented in Chapter 7, while suggestions on future work is given in Chapter 8.

[4]Passive identification is when the subject does not do anything actively, i.e. being passive, to be identified.
[5]Active identification is when the subject knows that he is being identified, and helps the process

2 Biometric Authentication

Over the years, the need for keeping important information and objects from falling into the wrong hands have lead to the introduction of *authentication*[1]. Authentication is used in a number of places, from ATM machines to boarding a plane. As already described in Chapter 1, authentication is proving a claimed identity. This proof can be either something you *know*, something you *have*, or something you *are*. The first two categories are well known and have been used for a very long time. Passwords, PIN codes, pass phrases and the likes are examples of something you know. Smart cards, cell phones or even the key to your house are examples of something you have. The two categories have something in common, the proof you bring can be changed. You can always change your password, change the door lock and so on, but the last category is somewhat different. This category contains things that you in most cases cannot control or alter in any way. When using an authentication system that relies only on what you *know* or what you *have*, you will always be accepted or rejected based on the proof. The proof you have of your claimed identity will be the only correct answer. It will always be black or white with these systems, either you are fully accepted or fully rejected. Biometric authentication introduces a *threshold value* to the mix which makes them less absolute, and this can be seen as a negative factor. The threshold is a score defined by the administrators of the system decided based on how strict they want the system to be. A captured image will never be exactly the same as the original input, thus a comparison method must be used between a saved *template* of the user, and the newly captured image. This comparison is done by using a *distance metric*, something that we will describe in Chapter 5. If the computed distance is below the threshold, the user is accepted by the system.

If a user of a "*know*- or *have*-system" forgets his password or misplaces his token, this can always be corrected by renewing the password or distributing a new token and invalidate the old ones. A biometric authentication system does not work in the same way. The problem with these systems is that once e.g. your right index finger's fingerprint gets compromised, you cannot change that finger to continue to use it any longer. Fingerprints are better than most other biometric features however, as you have 10 fingerprints, but only 2 eyes or just one kind of DNA (unless you had a transplant). On a positive note, a biometric feature is difficult to forge and requires alot of effort, time and money. A password to some system can be found in a numerous different ways, but an iris pattern is not something that can be created in a couple of minutes [8].

Jain tried to categorize the different biometric features, and give them a score of *low*, *medium* or *high* in these categories [9]. He argued for 7 characteristics that each biometric feature has, and these characteristics would determine whether or not the biometric feature would be suitable for authentication and how well it would perform.

- Universality - Every person should have the biometric feature. If the feature is not present, measuring it will be impossible.
- Uniqueness - The feature should be unique from person to person.

[1] Authentication is from the Greek word authenticos meaning real or genuine.

- Permanence - The feature should not change over time. Aging might be a problem for some biometric features. If the feature stays the same over a lifetime, the permanence score will be high.
- Collectability - The feature should be easy to collect. Is it visible for the natural eye, or do we need expensive equipment to gather the information? A high score means that the feature is easily measured.
- Performance - This characteristic gives some information on how well the system perform with respect to speed and results.
- Acceptability - Is collecting the biometric feature accepted by the public? Some might have a problem using the feature as an identifier, this score determines how intrusive the acquisition of data is.
- Circumvention - Can the system be fooled? This characteristic tells something about how easy it is to circumvent the system.

Biometrics can then be divided into two subcategories, behavioral and physiological. Roughly, physiological features are measures of body parts such as the face [10, 11], fingerprint, hand, iris and even DNA. The behavioral category includes things you do without pondering too much on what you are doing, like signature, the way you walk [12, 13], talk, and even how you type when using your keyboard [14].

2.1 Biometric systems

Biometric systems work in many ways just like an ordinary authentication system. Users will have to be *enrolled* into the system before they can use it [15]. This is done by presenting the identity of the user and the biometric feature to the system. A template is then created and stored in a database. After being *enrolled*, the user will have to present his biometric feature each time to be *matched* against the enrolled template. Biometric systems are based around a specific biometric feature e.g. fingerprints or retina scans. They have a set of components as described in Section 2.1.1 and either *verify*[2] or *identify*[3] the user.

2.1.1 Components of a biometric system

Figure 1 displays a typical biometric system. The different parts of a biometric system can be described as follows.

Sensor

The *sensor* is what the user is presented with to capture the biometric feature. Sensors can be voice recorders, thermal cameras, fingerprint scanners, etc. Common features of these devices is that they are there to gather data to be used later in the system. It is in many ways the interface from the real world, to the system that is going to use the data.

Pre-processing

After the biometric feature has been captured, the data runs through a pre-processor. This part of the system, is where noise is reduced from the data and some kind of nor-

[2]Verification in this context is the same as authentication. The biometric system captures the biometric feature, and compares it with the template that is stored of the user in a database for verification.

[3]Identification is done by comparing the biometric captured with all the records in a database. The result of the closest match is returned with a score of how close the match is. If that score is within the allowed threshold, the user is identified as the closest match within the database.

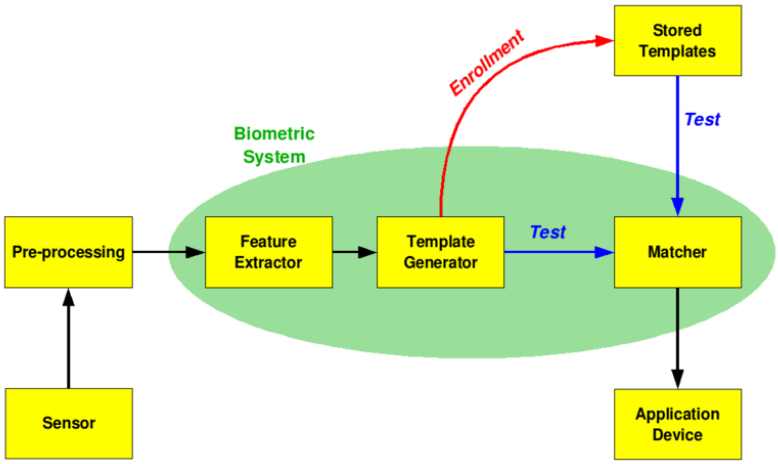


Figure 1: These are the components of a biometric system. (Based on figure from wikipedia.)

malization is done. After pre-processing, the data should be in a standardized form, so that the system can treat each capture the same way.

Feature extractor

When the standardized data enters the feature extractor, the important features are extracted. This is perhaps the most important step of the entire system, as the features needs to be extracted in the most optimal way to ensure that it will be of the same quality each time.

Template generator

When the important features has been extracted from the data, a template is being generated. This template is the final piece of information that is used in the rest of the system. Based on whether the user is a *first-time* user or a *returning* user, the template is "sent" to either the *matcher* or the *database* for enrollment.

Stored templates

When a template is first created, the user gets *enrolled* in the system. The template is stored in a database for future matching. Depending on the use of the system, these databases can be immensely huge. Imagine a worldwide register of fingerprints from criminals. For identification systems like these, a good search algorithm have to be implemented for the system to be effective. The database needs to be secure as well, so that no one can modify the data that is stored there. The stored template can also be located on the token the user brings in addition to his biometric feature.

Matcher

The *matcher* is where the comparison between the new template and the stored template is being done. A score is being generated based on how well the newly generated template match with the stored one, and based on a threshold the user gets accepted or rejected. The matcher algorithm should take into account that all capturing of biometric features will give slightly different results since there is no absolute value like e.g. in a password. This is where the difficulty in biometric recognition lays, the system must be able to accept small differences, but still be able to differ between two subjects [8].

Application device

Lastly, the part which is not actually inside the biometric system but still important for the system as a whole, the application device is where the output is used. This can be access control to different rooms, buildings, planes, etc. where the user will or will *not* get access to based on the score evaluated by the system.

2.1.2 Comparison of biometric systems

Biometric systems use templates from the enrolled users as described in the previous sections. When a user wants to authenticate himself, he needs to present the system with his biometric feature, which will be used as input to the system. Since the template and the input wil never be exactly the same, we need some way of comparing these two items. The features that are being extracted from the template and input will differ, and this implies the use of *distance metrics*. By using distance metrics to determine the "distance" between a template and an input, one will get a score, telling how far apart these two items are from each-other. When defining the system, one needs to decide what purpose it will be used for because this has a major impact on the next step. Let us say that we want a system of high security where only the right people should get access. In these systems, we would like the input to be as close to the template as possible, thus setting the threshold accordingly. If the input does not reach this threshold, the user will be rejected. Other systems might prefer high convenience prior to security, e.g. a log-on system of a netcafé or even toll fraud systems. If users of a netcafé gets rejected when they try to log in, they might just run to the competitor and use their system instead. Toll controls should neither be forcing every passing traveler to open his luggage for control, as this would mean long queues and dissatisfied customers. If a legitimate user trying to access the high security system gets rejected because his input image does not match the template that was stored of him, we have a case of *false rejection*, i.e. he gets rejected while he should have been accepted because the threshold was set too high. On the opposite side, we have imposters who might get accepted into the system because the threshold has been set too low. As we can see from this, there might be both legitimate users getting falsely rejected as well as imposters getting falsely accepted within the same system at a given threshold. By raising the threshold, less imposters will be falsely accepted but more legitimate users will be shut out of the system. By lowering the threshold, less legitimate users will be shut out of the system, but more imposters will be falsely accepted. This tradeoff is what separates biometric systems from others based on what you *have* or *know*. Based on a given threshold, one can start to compare the system performance. As described there will be some users falsely accepted, and some falsely rejected. Dividing them by the total number of users trying to access the system we get the *False Acceptance Rate*(FAR)[4] and *False Rejection Rate*(FRR)[5]. These rates are the most commonly used measurements for grading a biometric system along with the *Equal Error Rate*(EER). The EER is a compromise between the two values, and is the point where the FAR and FRR are the same. The scores are often displayed as a curve with the FAR and FRR as the axes. This curve is called a *Detection Error Tradeoff* (DET)

[4]FAR - When a user gets accepted that should *not* be accepted by the system, it is a case of a *false acceptance*. In high security systems this is of course the most important score to keep as low as possible, as you would not want imposters to get access.

[5]FRR - On the opposite, we have the FRR which is a score that shows the fraction of users that has been *falsely rejected*. In high convenience systems this score will be more important, as people might get annoyed if they are shut out.

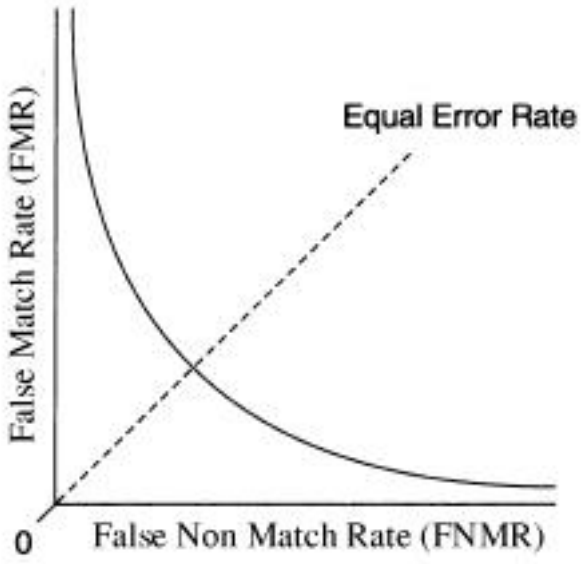


Figure 2: This is what a DET curve might look like with FAR as Y-axis and FRR as x-axis. [16]

curve (see Figure 2). Biometric systems do have different purposes, so it is not always the EER score that is most convenient. Because of this, one needs to take into account what the ups and downs are for each system before setting the threshold to where it should be. The application point should be decided based on the relative cost of the two error types.

3 Related Work

In 1989, Alfred Iannarelli proposed using photos of ears as a way of identifying people [17]. He worked from an investigator's point of view, so his "users" were prisoners. In his publication, he showed that the ear can be used as a biometric feature because of its uniqueness and permanence. It is easy to locate on a person, can be non-intrusive and it can be very easily collected. The area of interest is smaller than that of a face, so processing time is also faster compared to face recognition because of image sizes. Another advantage as opposed to the face is that the ear does not suffer from facial expressions [18, 19]. Different techniques have been published for authenticating people by their ear, with different results. This section will give an overview of the different techniques currently used in 2D images. Proposals for 3D imaging also exists [20, 21, 22, 23, 24, 2, 1] and some of the methods can be implemented on 2D images as well. A brief description of the *Iterative Closest Point* (ICP) algorithm as well as the *Principal Component Analysis* (PCA) will be given in this chapter. PCA will be discussed in detail in Chapter 5.

3.1 Existing techniques of ear authentication

There are two main categories when it comes to image analysis of ears: the statistical approach and the local features approach. The methods presented in this chapter have come up with different results as shown in Table 1. This is the state of the art concerning ear recognition. Little work has been done on thermal images [25], and research on blood vessel patterns has not been done at all on humans.

3.1.1 Iannarelli's twelve point system

As mentioned briefly, Alfred Iannarelli was the first to propose the ear as a way of authenticating people. In 1949 he developed his system of twelve measurements based on samples from 10,000 ears. In his report that was published in 1989 he also claims that

Approach	**Reference**	**Dataset Size**	**Recognition Rate**
LABSSFEM	[26, 3]	77(training), 77(test), USTB ear database [27]	85%
Neural Networks	[18]	84(training), 28(validation), 56(test)	79%
Force Field Transformation	[4, 5, 6, 28]	252(test), XM2VTS face database [29]	99,2%
PCA	[24, 30, 2, 1]	197(training), 88(registrant) ND Human ID database [31]	71,6%
Adjacency Graph Matching	[32, 25, 19]	n/a	n/a
Genetic Local Search	[33]	300(registrant), 180(unregistrant), 180(unknown)	100%

Table 1: Summary of different ear recognition methods.

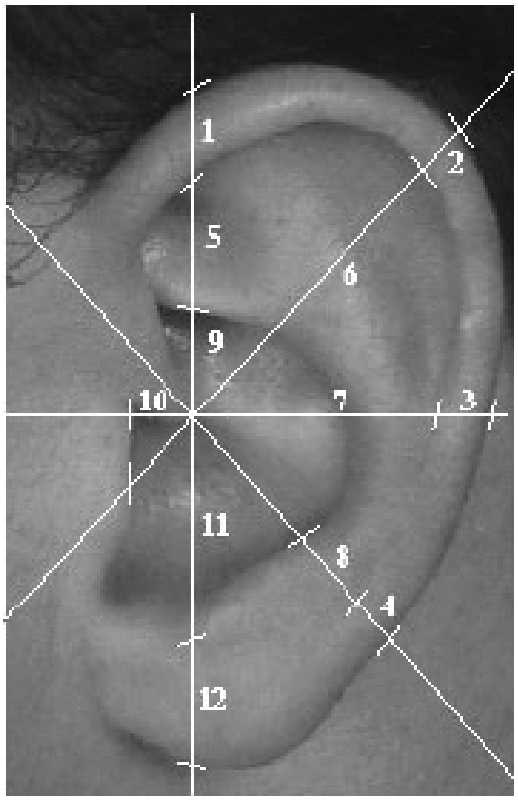


Figure 3: The locations of the anthropometric measurements used in the “Iannarelli System”. [17]

the structure of the ear does not change radically over the years even though the size does so. Between the age of 8 and 70 the growth is linear thus gives us good reasons to believe that systems can be developed to consider the changes that the ear experiences over time [17]. The twelve measurements from Iannarellis system is illustrated in Figure 3.

Iannarelli took photographs of the right ear and normalized them by resizing and rotating to match a template. Then he measured the twelve distances and tried to identify the ear against the templates of the prisoners. The system requires that the photographs are in this standardized form, in a “Iannarelli Inscribed easel”. To normalize and align the images, they are projected onto a standard “Iannarelli Inscribed” enlarging easel which is moved horizontally and vertically until the ear image projects onto a prescribed space on the easel. Such alignment process is typically performed manually and the measurement is taken in units of 4 mm and assigned an integer distance value. Since the photographs are scaled to fit this frame, Iannarelli could easily extract the different values. These values, in addition to sex and race were then used for identification. The weakness in this system was, as Iannarelli himself was very aware of, that if the normalizing of the image is a bit off, the twelve measurements would be completely wrong.

3.1.2 Principal Component Analysis

Principal Component Analysis (PCA) was one of the first techniques that was presented for ear recognition since it is derived from a well known technique used in face recognition [34, 35]. The extraction of "eigen-faces" has been done for quite some time, and is

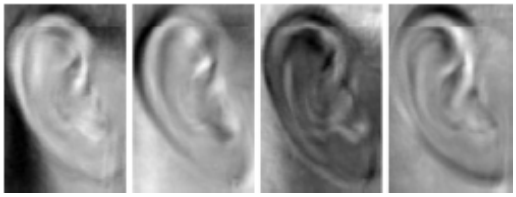


Figure 4: The first four eigenvectors of an ear image. [4]

a great way of reducing the size of the images that are to be validated. An image of size 256 by 256 describes a point in 65,536-dimensional space. Different images will have different points in this huge space. Using the "eigen-face" method, one tries to define those set of vectors that describes the image of a face the most. This is a pretty complex method but once you have defined the vector-set to work with, finding another image that describes a face is done very quickly. Since ear recognition has much in common with face recognition, the PCA method has been applied here as well [24, 30, 2, 1, 4]. The first four eigen vectors from a 111 by 73 pixel image are illustrated in Figure 4 [4].

PCA is by far the most widely adopted method used in ear biometric research [36]. Research done by Chang et al. [2] showed that this technique is as good for ear recognition as it is for face recognition. By combining these two, the authors came up with even better scores (recognition rates of 70,5% and 71,6%. When combined they got 90,9%). It is well known that when individuals are imprisoned they take photographs both frontal and from the side. Considering the recognition rates from Chang et al they would have a good way of identifying the prisoners by using this method in such cases. PCA will be deeper explained in Section 5.2.3 as we are using this algorithm in our

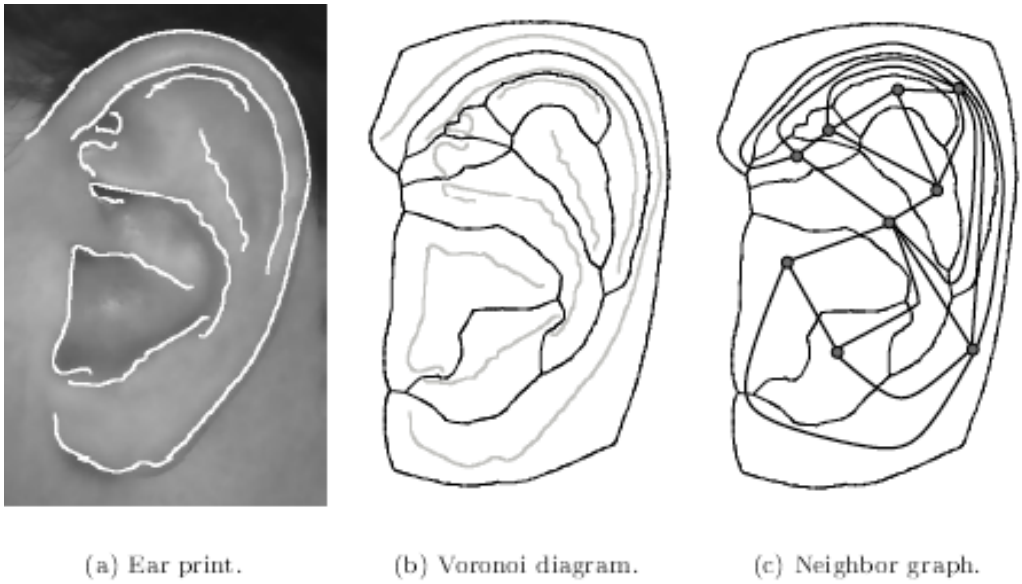
3.1.3 Iterative Closest Point

The *Iterative Closest Point* (ICP) algorithm is also a quite useful method of comparing data. This method has been used by many in the world of face/ear recognition and is a great way of comparing two 2D or 3D images with each other [21, 23, 24, 22, 2]. This algorithm was used as early as in 1992 by Besl and Mckay [37] and through a few modifications it has become very effective. The algorithm itself is not very advanced. Given a set of source points P and a set of model points X, the goal of ICP is to find the rigid transformation T that best aligns P with X [23, 24]. By transforming the image iteratively, the algorithm calculates the differences until the model converge with the source. This method is much used in 3D imaging, but can also be applied in 2D.

3.1.4 Adjacency graph matching

Burge et al. tried to use a graph matching algorithm for ear identification [32, 25, 19]. By generating a "neighbor graph" from the curve segments of a Voronoi diagram, the authors got images that could be matched against the templates in their database. The Voronoi diagram is extracted from the ear by edge detection in the image as illustrated in Figure 5.

As the Voronoi diagram is extracted by edge detection, several questions were raised. Hair of subjects were sometimes occluding the ear, thus creating extra edges in the image and making the wanted edges difficult to detect. Different lighting settings also proved to be a big problem for Burge et al. as the light caused shadows which in turn was detected by the Voronoi diagram extraction. The angle from where the photograph was taken



(a) Ear print. (b) Voronoi diagram. (c) Neighbor graph.

Figure 5: Stages in building the ear biometric graph model. [25]

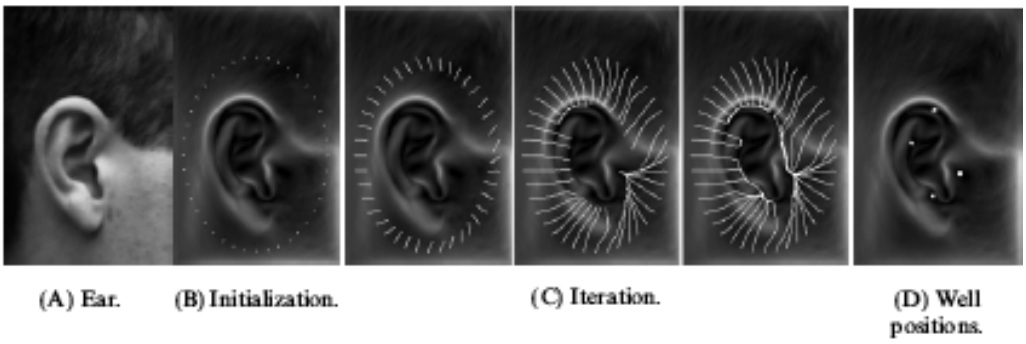


Figure 6: Field line, channel, and well formation for an ear. [4]

from also had major impact. No experimental results were given in their literature, but the authors suggested the use of thermal images to remove the impact hair would have on the results. Because hair has a lower temperature than the rest of the ear this could easily be removed in a final image.

3.1.5 Force field transformation

Instead of extracting feature vectors to describe the ear Hurley et al. came up with a new idea by using force field transformations [4, 5, 6, 28]. The ear image is treated as an array of mutually attracting particles in a Gaussian force field [38]. The pixels that are projected then flow naturally towards the potential wells under the influence of force, thus forming channels on their way. The points of the wells are then extracted to form the basis of the feature vector as seen in Figure 6.

Hurley et al. got great initial results using this method (99,2% recognition rate). Their result also showed that the feature vector is highly immune to initialization, rotation, scale and noise [5].

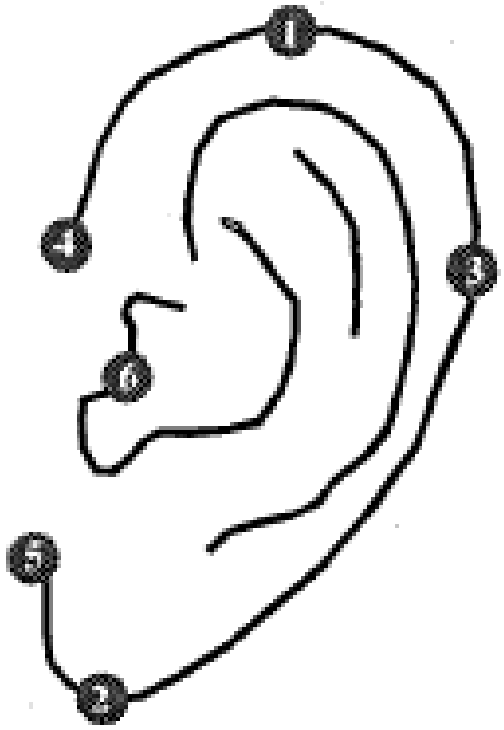


Figure 7: Considered feature points in outer ear images. [18]

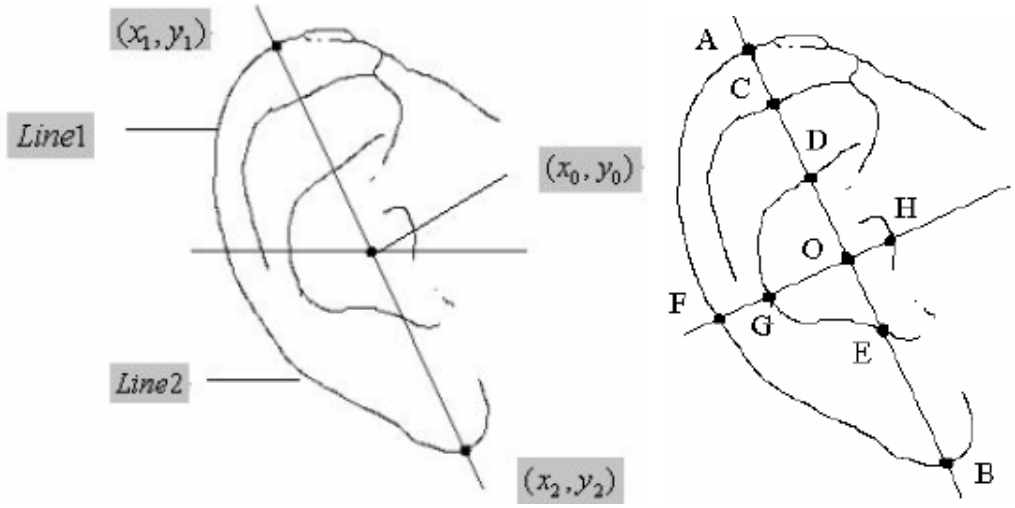
3.1.6 Genetic local search

By looking at the theory of "the survival of the fittest", Yuizono et al. came up with the idea of applying a "Genetic Algorithm" (GA) to a local search method to obtain the input ear image [33]. This method is often applied to problems where the computational complexity explodes exponentially, like the traveling salesman. In addition to the GA they added local search to the method, thus naming it genetic local search. In GA, the fitness of each individual is obtained in every generation. By applying some basic genetic operations like selection, crossover and mutation, a new set replaces the first one. This genetic operation is then repeated. In their study, they had 110 participants which were image captured six times each. By applying different selection methods the authors came up with a registrant recognition rate of approximately 100%.

3.1.7 Neural networks

The shape of the outer ear was the research base of Moreno et al. when they tried to apply the use of neural network for ear identification [18]. As well as giving valid reasons to look at ear recognition instead of face recognition, the authors extracted important parts of the ear and combined the classification results with different methods described in their article. Figure 7 shows the location of considered feature points that the authors used to define the *biometric vector*.

The results the authors came up with were decent (79% recognition rate using the compression network classification technique).



(a) Long axis segmentation of the outer ear contour. (b) Feature points of the inner ear

Figure 8: Long axis based shape and structural feature extraction method. (a) shows how the outer ear edges are defined, while (b) shows different feature points of the inner ear. [3]

3.1.8 LABSSFEM

Another structural feature extraction method was proposed by Mu et al. in 2004 [26]. LABSSFEM is the acronym the authors used and by this they mean *long axis based shape and structural feature extraction method*. The shape feature vector of the outer ear and the structural feature vector of the inner ear form the local feature vectors. The vectors were gathered by measuring the longest distance between two points in the ear and using the perpendicular axis to find feature points in the inner ear as shown in Figure 8.

The unique feature vector that was extracted by combining the two vectors, was matched against the template. With a recognition rate of 85% one could argue that this was not a success, but it still proved better than ordinary PCA and the study also provided a new way of thinking based on Iannarelli's old system.

3.2 Thermal imaging

Burge et al. suggested that taking thermal images of ears might reduce the error rates when the individuals have parts of their ear occluded by hair [25, 19]. They propose that since hair has a lower temperature than the ear, the ear should be detectable with the use of such images, and that the hair can be masked away in the process. Since there has not been any work done on this in the area of ear biometrics, this will be one of our challenges. A thermogram image of an ear is shown in Figure 9.

Thermal equipment has been used by the army for quite some time. Heatseeking missiles are used on a large scale. The technology that lies inside these missiles is very similar to that within a thermal camera. A company called MIKOS has developed thermal equipment to be used in different areas [39, 10]. MIKOS have also entered the area of authentication by thermal images of the face with their system, called FACES^{TM} [40]. In their studies, they have found that each persons thermal face is even more reliable than fingerprints. Even identical twins have distinct differences in their thermal patterns. The only problem with their system is that alchohol makes a huge difference on the result.

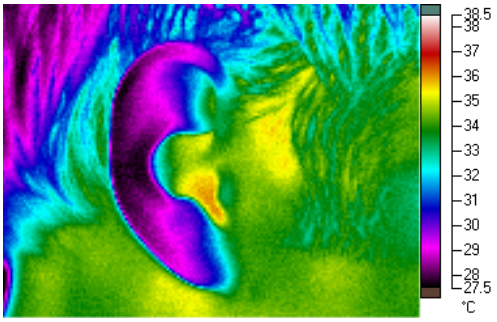


Figure 9: A thermogram of an ear. As seen here, the hair has a lower temperature than the inner parts of the ear.

People that have consumed alchohol prior to entering the biometric system, will not get the wanted results and will most likely be rejected by the system. Suggestions concerning a merging of both the ordinary image and thermal image of the face has also been done [41]. Since other technologies that has been used on faces has been adapted to the ear, this would also be something that should not go untried. Prokoski and Riedel has also written a chapter in the book on biometrics concerning infrared identification [42]. Here the authors argue that thermal images are both unique, easily collected and are immune to forgery.

4 Experiment

In this chapter we describe the experiment that we have performed. Due to the fact that we had limited availability of the equipment, this setup had to be strictly defined to be able to have as few photo sessions as possible.

4.1 Experiment setup

The experiment was done in the following way:

- Each participant was seated in a chair.
- The camera was held on a stand at a fixed distance (30cm) to the ear.
- 5 images of the right ear were taken in this position. For each of the five images the person either stood up, shook his head or looked up and down to make sure that no two images were identical.
- Room tempearture was recorded.
- These four steps were repeated 3 times on different days to see if there were any changes over time.

All the images were taken with a Fluke TI-25 thermal imager. This is a handheld camera (see Figure 10) that came bundled with a software called SmartView^{TM}. The camera has a thermal lens as well as an ordinary lens. The thermal lens takes images of 320 x 240 pixels while the ordinary lens produced 640 x 480 pixels (VGA quality images). The camera captures images with both lenses at once, which is quite useful later when we compare the performance of the thermal images versus the ordinary ones. Furthermore, the thermal sensitivity of the camera is $\leqslant 0,1°C$ at $30°C$ which is acceptable for our purpose. What this means is that at $30°C$, the camera displays the correct degree value down to the tenth of a degree.

A part of the images are used for making a template, while the rest will be matched against these templates.

(a) The Fluke Ti-25 handheld thermal camera (b) Package contents of the Ti-25

Figure 10: The thermal camera (a), along with package contents (b). [43]

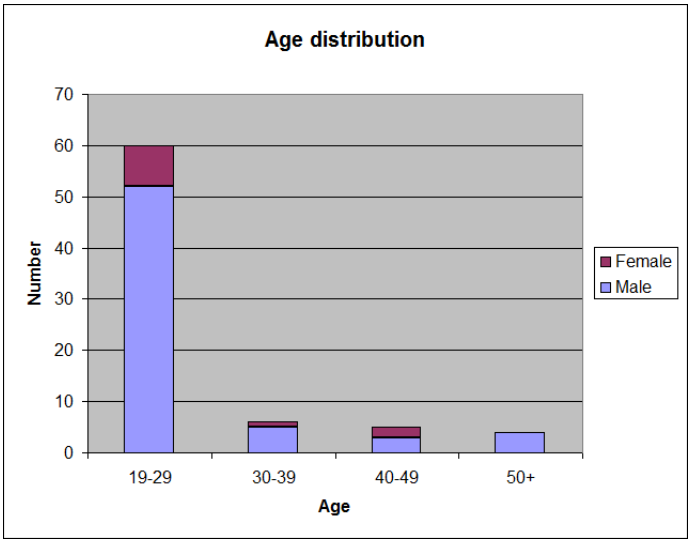


Figure 11: The gender and age distribution for the main experiment.

In addition to the above described experiment, we conducted another experiment with the same chair and distances but with the following differences:

- 2 reference images were taken initially.
- A bag of ice was held against the right ear for 10 minutes.
- Photos were taken each 15 seconds for 15 minutes, producing 61 images not including the 2 reference images.
- Room temperature was recorded.

This experiment was set up to try to find out how much time is needed indoors before the ear is back to its normal body temperature state. The experiment included 5 participants that were chosen among fellow information security students.

4.2 Volunteer crew

The participants of the main experiment were picked by random selection from the students and staff of Høgskolen i Gjøvik of different nationalities. Both female (11 participants) and male (64 participants) took part in the experiment which in total counted 75 participants. The age of the participants ranged from 19 to 65 years old so we got a broad spectrum of different ages. The mean value for the female participants was $\mu = 27$ and $\mu = 28$ for the male participants. The standard deviation for both groups was $\sigma = 9$. In Figure 11, the age distribution is illustrated.

4.3 Environment

The room used for the experiment was the authentication lab of NISLAB, room A-113 at Høgskolen i Gjøvik. The temperature in this room varied from between $22,0°C$ to $23,9°C$. Temperature was measured after each photo taking session of the experiment, to be sure that the indoor temperature would have little or no effect on the testing. SmartView^{TM}had options to set the background temperature, but doing so for the given

temperatures gave no effect in the image. The chair was 30 centimeters away from the tripod to have a fixed distance from the camera. We made marks on the floor to have the chair and tripod in the same position at all times. By doing this, we had nearly no variation in distance between the camera and the ear. We also had the participants staying indoors for a while before the image captures, so they would have a normal body temperature.

Some problems came up with this setup, as we discovered along the way. First of all, some participants who claimed they had been indoors for the required amount of time before image capture, just came in from the outside where the temperature was much lower than indoors. This was discovered when looking at the images in SmartView™. Because they were not consistent in whether or not they just came from the outside, two sessions would be completely different in the thermal spectrum (see Figure 15 on Page 25). For the ordinary images, this made no impact since the ear looks more or less the same. They might have been a little bit red, but not close to the impact it had for the thermal images. If we look at this problem in the context of real-life applications, this problem is something that would come up sooner or later since users are not treated as strict when using a real-life application. Another problem which was discovered after capture, was the fact that some of the images were out of focus. We had little experience with this type of camera before we started out, and due to the fact that it is a handheld camera with no possibilities to mount it on the tripod (just hold it there as a reference), some of the images were not as sharp as we wanted them to be. This may have had an impact on the results which we will come back to in Chapter 5.

5 Analysis and Results

This chapter will describe our methods when conducting the pre-processing and analysis. The programs we have used will also be described in short detail.

5.1 Preprocessing and standardizing

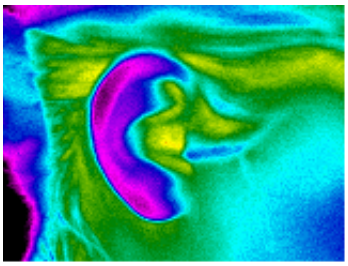
Pre-processing became a large part of this project. We got 1125 images in total from the initial experiment (75 participants x 5 images x 3 sessions) and to get our results, we had to go through each and every image several times.

Step one - SmartView™

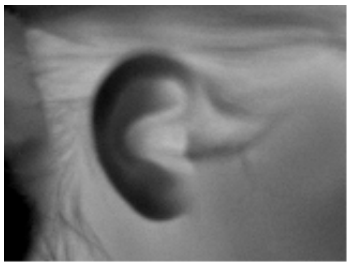
First of all, the images had to be imported to the program that came bundled with the camera. SmartView™ might lack a few features when it comes to image processing, but for the initial pre-processing it was more than enough. SmartView™ lets you change the *palette* used in the image. Figure 13 displays the different palettes you can choose from in SmartView™. We decided to go for the *high contrast* palette because this has the most number of colors in contrast, which will give an image in which it is easier to see the differences than in any of the other palettes. In addition to this, we made another set with the grayscale palette and one with the ordinary images (see Figure 12). The Ti25 has two lenses as described in Chapter 4, so the ordinary images were taken and are used as our third set.

The next thing that was done in SmartView™, was setting the *scale*. This is most important for the colors in the images, as changing this would change the colors of the image. After taking a few sample images we found that in an ear with a normal body temperature, the temperatures ranged from 28° to 38° Celsius. This makes sense, because or body temperature is normally at 37° Celcius. Since the ear stands out a bit from the body, the regular ear tempearture might be a bit lower but not lower than 10°. For good measure we added 0,5° on both sides, thus setting the scale to 27,5° to 38,5° Celsius. All temperatures *above* the given scale would turn out white, while all temperatures *below* the scale would be displayed as black. The scale adjustment toolbox can be seen in Figure 14. The motivation behind setting this fixed scale was that all temperatures will have the same given color in all the images. This is good for reference, as the images can be matched by the temperature pattern as well as the shape of the ear. After some sessions though, we discovered that some of the participants came for the photo session after just being outside. This made the ears much colder which in turn showed during the pre-processing as big black parts in the image, see Figure 15. After discovering that some of the ears were below the scale, we decided that the grayscale-dataset would have its scale revised. We went through all the images, setting a custom scale for each and every one of them. In Figure 15 the same image can be seen with the fixed and the custom scale for the *high contrast* palette[1]. By letting the max temperature in the image be the top of the scale, none of the ears would have equal grayscale values for temperatures over a certain threshold and the same for lower temperatures. We still

[1]Hereby referred to as the colored image set



(a) Thermal image of the right ear using the *high contrast* palette



(b) The same image, using the *grayscale* palette

(c) The same image, using the ordinary image capture

Figure 12: Different palettes might give different results. These are the three datasets we used for analysis.

wanted to keep the fixed scale images though, to compare the performance of these with the customized scale images.

Finally, we exported all the images (1125 images x 3 datasets) to a standard image format (jpeg) for further processing.

Step two - MatLab

When you are working with 3375 images, some sort of automation tool is necessary. After the initial pre-processing done in SmartView™, we now had the images in a format where such a tool could be used. For this purpose, we found that using MatLab was the best option [44].

Since the images had variations in how the ear was positioned, we had to normalize

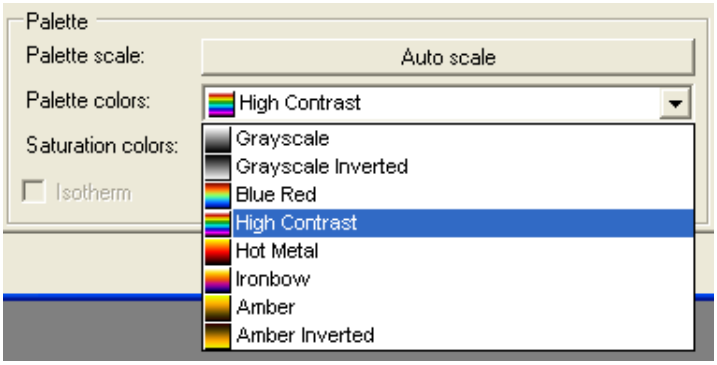


Figure 13: The different palettes that can be chosen in SmartView™.

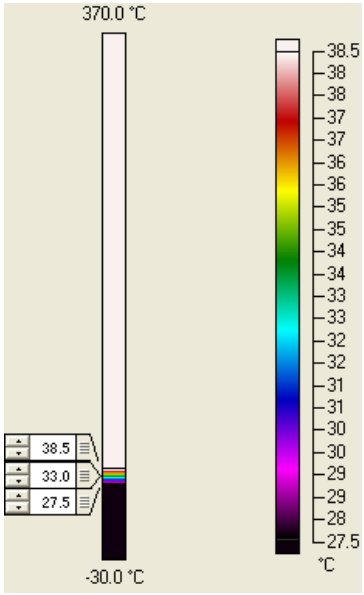
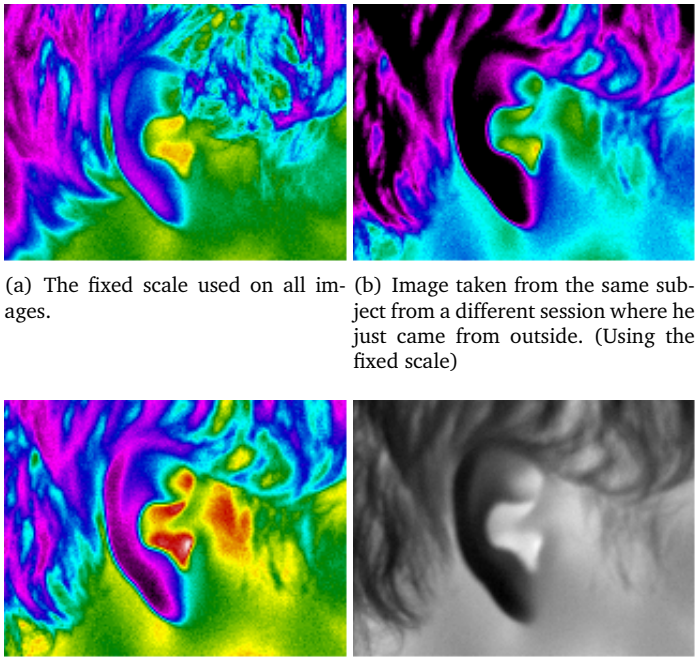


Figure 14: In this toolbox, you can change the scale by adjusting the sliders.



(a) The fixed scale used on all images.

(b) Image taken from the same subject from a different session where he just came from outside. (Using the fixed scale)

(c) Same image as (b) but with a custom set scale of 25,2° to 35,9° Celsius.

(d) The same image as (c), but with the grayscale palette which was used for our analysis.

Figure 15: Figure (a) displays a “perfect” ear, where the temperatures matches the fixed scale. Figures (b) and (c) are from the same subject, taken from a different session where the user just came from outside. Setting a fixed scale for all images made some ears black (b), so we made a grayscale dataset with custom values for the scale (c) and (d).

the ear somehow. By clicking manually on the two parts of the outer ear that formed the longest axis, we were able to rotate the image such that this axis became the y-axis. After the rotation, we cropped the images to remove unwanted information. This was done by taking the longest axis as a basis, and cutting horizontally at the two selected points. The vertical cropping was done by taking a percentage of the longest axis on each side of it (30% on the left side and 20% on the right side). Finally, to get all the images on the same size, they were resized to 50 x 100 pixels. These five steps can be seen in Figure 16. At the beginning of the pre-processing phase, we wanted to do everything automatically and tried to come up with a way of detecting the edge by using mathematical measures. After working on this for a couple of weeks, we found it too problematic because of the hair creating edges in the images as well as the edge of the ear. We therefore stuck to the manual method of clicking on the ear to create the longest axis.

This preprocessing was done for all images, so we could have a starting point for the analysis which were to follow. For some of the methods that we have used in the analysis, the images had to be modified even more but this will be explained in the sections where it applies. The code we made for the image cropping can be found in Appendix B

Feature extraction

Images such as the ones we have captured, can be treated in different ways. We chose to look at the color value of each pixel in the image, saving the images as vectors of length $n = 50$x100. We chose the image values, as they can be compared to each-other by the means of distance metrics in order to find differences and similarities in the images.

5.2 Analytical methods

Now that all the images were on the same format, we could start doing some analysis on the dataset. We have decided to use both commercial tools as well as easy to implement algorithms in our analysis. The following tools were used:

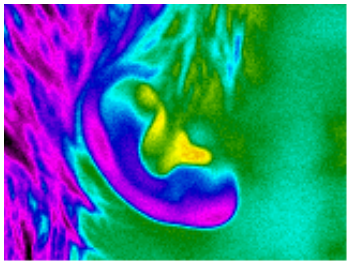
- **Algorithmic measures**
  - Distance metrics
  - Principal Component Analysis
- **Commercial tools**
  - ImageFinder
  - Image Comparer

These tools will be described in detail in the following sections.

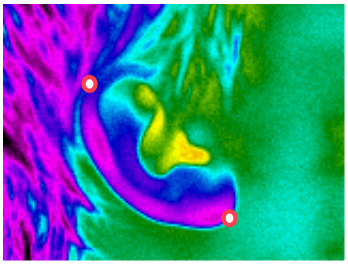
5.2.1 General analysis technique

We applied the algorithms to our dataset by letting MatLab select random templates from each user, and match them against all the remaining images. If the image was matched against the template from the same participant, the result was saved as a *genuine* attempt. If the image was matched against the template from a different participant, the result was saved as an *impostor* attempt. The FAR, FRR and EER were then calculated from these results. For PCA[2], this was repeated 100 times to compute a mean EER as well as a confidence interval. PCA also required that we selected one template in addition

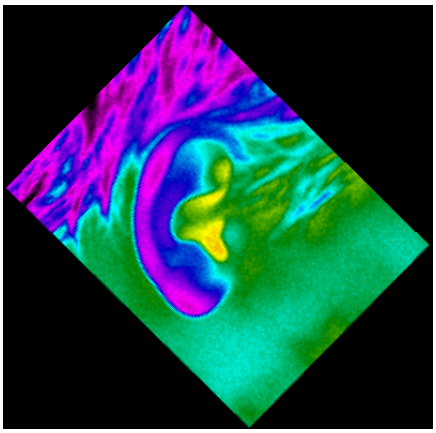
[2]PCA was first used by Turk and Pentland for face recognition in 1991 [34]. They called the Principal Components eigenfaces because they resembled faces.



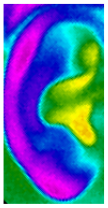
(a) Step one - initial image



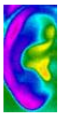
(b) Step two - clicking the ear to create the longest axes in the ear.



(c) Step three - rotating the image based on the selected axis.



(d) Step four - cropping the image based on the length of the longest axis.



(e) Step five - resizing the image to 50 x 100 pixels

Figure 16: This is the pre-processing that was done on the images, using MatLab. Figure (a) shows the initial image, (b) displays the two manually selected points, (c) is the rotated image while (d) shows the cropping. Finally, (e) shows the normalized image as it was used for analysis. The same steps were done for the grayscaled and normal images as well.

P	Images									
	1	2	3	4	5	6	7	8	...	15
1					x		o			
2		o					x			
3	x									o
4			o			x				
5	o	x								
6						o				x
7			x		o					
8				o				x		
...										
75				x				o		

Table 2: PCA template selection. This table is just for illustration, participants vertically and their images horizontally. The *x* and *o* shows that two images were selected randomly from each participant, one for PCA training (x) and the other for template generation (o).

to an image for training the system. This was done by selecting one random image for a template, and one random image for the PCA training from the 15 images of each participant. The template images from each participant would then be matched against the remaining images, i.e. 13 images per person ($13 * 75 = 975$ in total) are matched against the template images from all participants.(see Table 2)

5.2.2 Distance metrics

Since all our images were on the same format, it was rather easy to implement some simple distance metrics on the datasets. A distance metric calculates distances between a template and an input sample. This can be done in numerous ways, but we chose the most common ones, *Manhattan distance* and *Euclidean distance*.

- **Manhattan distance**: This is also known as absolute distance and the formula is shown in Equation 5.1. Manhattan distance is a very simple metric that only takes the absolute value between all the values in the template with the corresponding values in the input sample. As a result of this, Manhattan distance requires that the template and the input sample have equal length. This gives us no problems though, since all our images are on the same format. This distance metric is the computationally least expensive one.
- **Euclidean distance**: This is a slight modification from Manhattan distance, see Equation 5.2. Instead of taking the absolute distance, we now take the square root of the sum of all distances squared.

$$\mathrm{dist}_{Manh.}(\underline{X}, \underline{Y}) = \sum_{i=1}^{n} (|x_i - y_i|) \tag{5.1}$$

$$\mathrm{dist}_{Eucl.}(\underline{X}, \underline{Y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{5.2}$$

5.2.3 Principal Component Analysis

We have already described Principal Component Analysis method briefly in Section 3.1.2, but it needs a deeper understanding before applying it to our dataset. PCA is widely used in face recognition, and it has been applied for ear recognition as well, for normal images.

Under the surface

To further understand how the PCA algorithm works, we need to dig deeper and look at some of the details behind it. The PCA algorithm embraces standard deviation, covariance, eigenvectors[3] and eigenvalues[4] [45, 46]. Some of these elements are described in statistical mathematics, while the last two falls into the matrix algebra category. PCA is therefore a mix of the two categories as it contains all these features. We will not go into details concerning the statistical methods, as these are covered in basic mathematic courses. The matrix algebra however, is more interesting. Two matrices can be multiplied with each other as long as they are of compatible sizes. This is well known, and eigenvectors are just a particular case of such a matrix. Let's say we have the multiplication of two matrices like this:

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

The result can be seen as a multiple of the vector {3,2} thus making this vector an eigenvector to the matrix it was multiplied with. These eigenvectors can only be found in square matrices of $n \times n$. Some square matrices do not even have eigenvectors and if the size of the matrices that *do* have eigenvectors grows to more than 3×3, it is very hard to determine what they are unless you use an iterative computational method, i.e. not doing it by hand. This is why the PCA is mostly used by just letting a program compute the different vectors from a matrix. An $n \times n$ matrix that *does* have eigenvectors will *always* have n eigenvectors. A 3×3 matrix has 3 eigenvectors if any. Another feature of eigenvectors is that they are *perpendicular* to each other no matter how many dimensions you are working with. These new axes then servers as a baseline when processing the data and makes it very easy to work with instead of having to use the x, y and z axis when describing the expression. For each eigenvector, an eigenvalue exist in relation to the eigenvector. As for the eigenvector given above, we can see how the eigenvalue is found. Let us define x and y from the two values in the eigenvector. $x = 3$ and $y = 2$. This gives $x = \frac{3}{2}y$ and $y = \frac{2}{3}x$. By multiplying these values with the original matrix we get the following:

$$\begin{pmatrix} 2x & + & 3y \\ 2x & + & 1y \end{pmatrix} \Rightarrow \begin{pmatrix} 2x & + & 3(\frac{2}{3}x) \\ 2(\frac{3}{2}y) & + & 1y \end{pmatrix} = \begin{pmatrix} 4x \\ 4y \end{pmatrix} = 4 \times \begin{pmatrix} x \\ y \end{pmatrix}$$

By looking at an example to further prove this, we multiply the eigenvector by an arbitrary value (i.e. 2). The result of the multiplication will be 4 times the scaled vector as shown below. Thus the eigenvalue of this particular eigenvector is 4. This stands as long as the relation between x and y equals the original one.

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 24 \\ 16 \end{pmatrix} = 4 \times \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

[3] An eigenvector is a vector used in the context of transformation. A linear transformation is done by using vectors that usually change in direction and length. A vector that multiplied with a certain value gives the result of this transformation is called an eigenvector.

[4] Eigenvalues are the values that the eigenvectors need to be multiplied with to get the result of the linear transformation. The direction of the eigenvector is unchanged or reversed (for positive or negative eigenvalues)

Eigenvalues and eigenvectors always come in pairs, so when implementing PCA on a computer the program normally extracts the value along with the vector when computed. In addition, the vector is normalized to have length 1 even if this gives fractions in either value x or y.

PCA explained

Now that we have explained a bit about the matrix algebra behind the PCA algorithm, a bit more detail concerning the algorithm itself follows. PCA is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences [45]. Another huge advantage with PCA is that once you have found these eigenvectors and patterns in the data, you can remove some of the dimensions without loosing too much of the information. PCA contains a six-step process as follows:

1. Get some data
2. Subtract the mean
3. Calculate the covariance matrix
4. Calculate the eigenvectors and eigenvalues of the covariance matrix
5. Choosing components and forming a feature vector
6. Deriving the new data set

In the first step, all we need to do is converting an image into a matrix. This can be done in several ways which will be covered in later sections. Once we have that matrix, we can subtract the mean of all the values and then calculate the covariance matrix (see Figure 17).

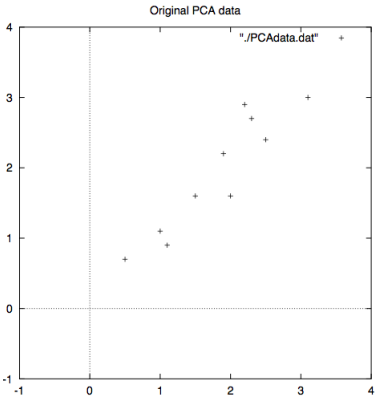
By using two of the eigenvectors as axes, all we do is rotating the output. If we further reduce the data set to only the single eigenvector with the highest eigenvalue, we will loose one dimension. This gives us only data on a single line as shown in Figure 17(c).

If you compare Figure 17(c) with Figure 17(a), you will see that the values are all the same in one dimension, and are now easily comparable to each other based on this dimension. When using this technique on images we might first normalize all images to be of a fixed size. Then we create image vectors from each image by taking each row of pixels consecutively. Each image vector is then placed in a huge image matrix for all the images in our data set. If we now run the procedure as explained above, we can easily decide if a new image is close or far from one of the images in the set.

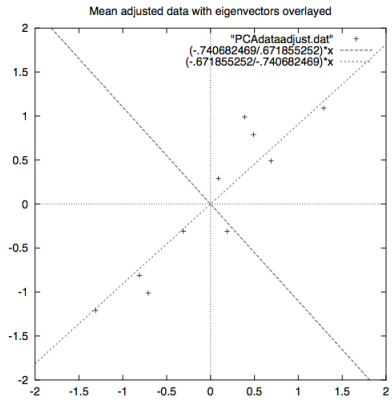
5.2.4 ImageFinder

Attrasoft Inc. have made an image comparing tool, called ImageFinder. This tool is a commercial product that lets you run through entire folders while trying to match images with each other. We downloaded a trial version, which is actually the same as the full version but only lasts for 30 days. The documentation states that the program can be used for:

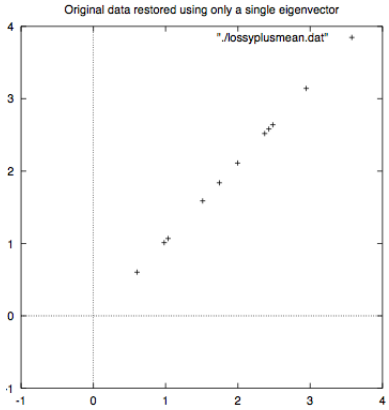
- Image verification (1:1 matching)
- Image identification (1:N matching)
- Image search or retrieval (1:N matching)
- Batch processing (N:N matching or N:M matching)



(a) A plot of some data using the standard x and y axis.



(b) The same plot as in (a) but now with the eigenvectors of the covariance matrix overlaid on top.



(c) The reconstruction from the data that was derived using only a single eigenvector.

Figure 17: These figures illustrates how PCA works. (a) shows a plot with standard x and y axis. (b) and (c) then illustrates how PCA changes the axis and representation of the values. [45]

Since the last point is of our utmost interest, we gave it a try to see what it could produce. By setting up a small part of the dataset for training, the program can either run through the dataset using a *Bio filter* or *neural filter*. Neural filters are best for huge datasets, so after some consultation with an Attrasoft employee we decided to go for the Bio filter and Neural filter. The problem with such programs is that you do not always know what is going on behind the scenes, but we wanted yet another resource for results so we went along with it. First of all we had to create *match* and *training* files to feed the program with. This is more or less telling the program which images that should match, and a small part of them which will be used for training the filters. MatLab did this job for us as well, so once we had the files ready, we let the program to the job. Training and matching was rather easy to walk through once we had the needed files. Results were saved in a .txt file. To be able to extract the values that we needed, MatLab again came to the "rescue". Output from MatLab can be seen in Section 5.3.2.

5.2.5 Image Comparer

Our last effort for getting results was another commercial product called Image Comparer. As with ImageFinder, this program is available on a 30 day evaluation version. The program is pretty simple in use, with not a lot of functionality. It lets you open a folder with images to save it as a *gallery*. You then have the option to compare the gallery with itself, or compare it with another gallery. Results are exported to a .txt file with each line stating a template image, a matching image and a score between those two images. This is a relatively quick way of getting results but the problem with this program, more so than with the ImageFinder, is that you do not know what is going on in the background. Another limiting problem with the program is that you can only set the lowest *similarity threshold* to 70%, i.e. all matching done that gives a lower score is not included in the report. See Section 5.3.3 for reults.

5.3 Results

This section will provide the results from the different methods used for analyzing the data. In Chapter 6 we will discuss these results in deeper detail.

5.3.1 PCA

As all the images that were taken produced three layers of information (red, green and blue), we wanted to look at how they would do compared to each-other with PCA. We also had the option of grayscaling the images before comparing, and this can be done in different ways. We chose the "easy" way of doing it by letting each color layer count equally, as well as converting the values based on YIQ vector quantization [47], which in basic weights the red value with 0.3, the green 0.59 and the blue value by 0.11.

By looking at these 5 set-ups for the 3 datasets, we will have 15 different results in total. As mentioned in the section of distance metrics, we used the manhattan distance as well as the euclidean distance to see which of those that produced the best results. We also had a look at the eigenvalues when applying these distance metrices. In these cases each of the terms in Equations 5.1 and 5.2 is multiplied by the eigenvalue of the corresponding eivenvector.

- 3 image sets (ordinary, colored and grayscaled)
    - 5 color layers (red, green, blue, gray equal and gray YIQ)

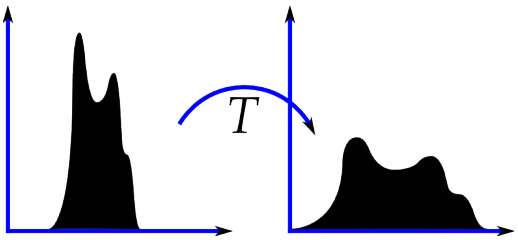


Figure 18: Equalization of a histogram. By stretching the histogram to include all values in the spectrum, we may get better results from our analysis. (Figure taken from Wikipedia.org)

- Manhattan distance
- Manhattan distance on eigenvalues
- Euclidean distance
- Euclidean distance on eigenvalues

We also applied histogram equalization on the images, a method that stretches the color values over the entire grayscale spectrum in order to get clearer lines in the image. If a lot of the values are concentrated around a small part of the spectrum, this method will in short stretch the values, creating more space between each value (see Figure 18) [5].

This obviously produced a lot of EER values which are given in Table 3. They are all mean values computed by taking the mean of 100 generated EER values. As can be seen in the table, the Manhattan distance metric proved to be the best with respect to EER values. It can also be seen that for the best results, the histogram equalization did not make our results any better.

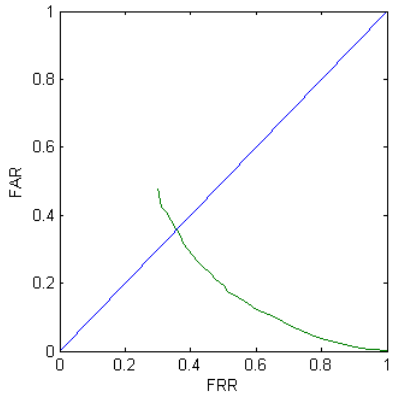
5.3.2 ImageFinder

We applied both a bio filter as well as a neural filter to the images in ImageFinder. The DET curves can be seen in Figure 19, and the corresponding EER values are given in Table 4. ImageFinder seems to produce more or less equal results for all the datasets (between 29 and 35% EER). The DET curves are not complete, (see Figure 19) and the reason for this is that ImageFinder does not compute values for all the matches and non-matches. Instead it sets the value to 0. This will obviously raise the number of false rejects as these scores are below any threshold that we set. The output from ImageFinder was easy to work with but as mentioned, not complete. The first line of the file would display the filename of the template, then the following line would display the sample image that was matched against this template. The third line would display the "score" that was calculated by ImageFinder. The following lines would also display sample images that was matched against the template and their matching score. By separating with a blank line, the next template would then come with the following sample matches and scores. We applied the same technique as was done on the results from PCA, i.e. picking a random template from each participant to fill a matrix of dimension 75x1050, calculate the EER, and repeat this 100 times to get a statistical significant EER value.

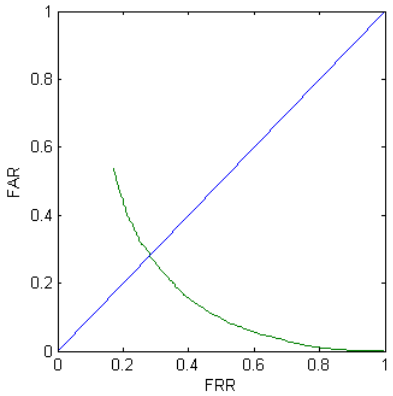
[5] Histogram equalization is a rather trivial operation, and a description can be found at Wikipedia.org

	Colored	Ordinary	Grayscaled	H Colored	H Ordinary	H Grayscaled
Red Manhattan	**0.2489**	0.1419	**0.1453**	**0.2610**	0.1565	**0.1521**
Red Manhattan Eigenvalues	0.3038	0.2609	0.2525	0.3076	0.2586	0.2408
Red Euclidean	0.2633	0.1806	0.1826	0.2731	0.1841	0.1663
Red Euclidean Eigenvalues	0.3134	0.2734	0.2629	0.3153	0.2726	0.2480
Green Manhattan	0.2588	0.1446	**0.1453**	0.2916	**0.1496**	**0.1521**
Green Manhattan Eigenvalues	0.3121	0.2668	0.2525	0.3080	0.2298	0.2408
Green Euclidean	0.2643	0.1762	0.1826	0.2880	0.1710	0.1663
Green Euclidean Eigenvalues	0.3066	0.2791	0.2629	0.3053	0.2420	0.2480
Blue Manhattan	0.2563	0.1505	**0.1453**	0.2665	0.1544	**0.1521**
Blue Manhattan Eigenvalues	0.3152	0.2904	0.2525	0.3237	0.2468	0.2408
Blue Euclidean	0.2610	0.1920	0.1826	0.2709	0.1796	0.1663
Blue Euclidean Eigenvalues	0.3092	0.3039	0.2629	0.3193	0.2588	0.2480
Gray Equal Manhattan	0.3330	0.1394	**0.1453**	0.3328	0.2022	0.2149
Gray Equal Manhattan Eigenvalues	0.3638	0.2691	0.2525	0.3561	0.2610	0.2647
Gray Equal Euclidean	0.3434	0.1802	0.1826	0.3373	0.2105	0.2203
Gray Equal Euclidean Eigenvalues	0.3699	0.2826	0.2629	0.3597	0.2700	0.2662
Gray YIQ Manhattan	0.2709	**0.1386**	**0.1453**	0.2733	0.1538	**0.1521**
Gray YIQ Manhattan Eigenvalues	0.3201	0.2650	0.2525	0.2874	0.2430	0.2408
Gray YIQ Euclidean	0.2843	0.1773	0.1826	0.2724	0.1784	0.1663
Gray YIQ Euclidean Eigenvalues	0.3168	0.2781	0.2629	0.2926	0.2560	0.2480

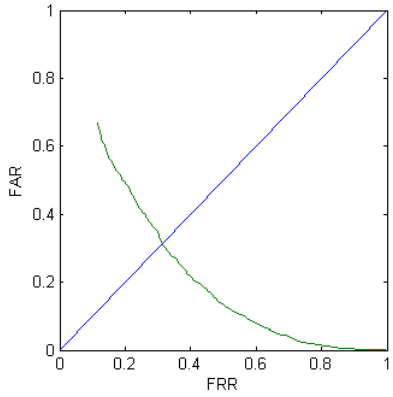
Table 3: These are the results we got from applying PCA to our datasets. As can be seen, the visible and grayscaled sets give approximately the same results, while the colored set is a bit worse. Also worth noticing, is that the Manhattan distance metric performs best in the different datasets. The best scores are displayed in bold.



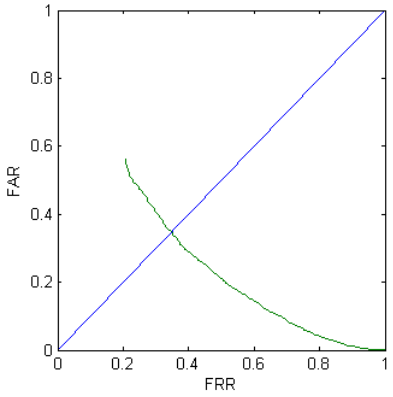
(a) Grayscaled images with bio-filter.



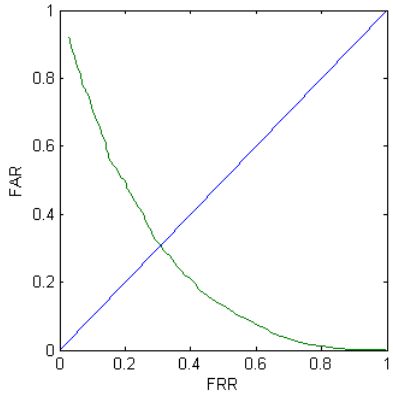
(b) Grayscaled images with neural-filter.



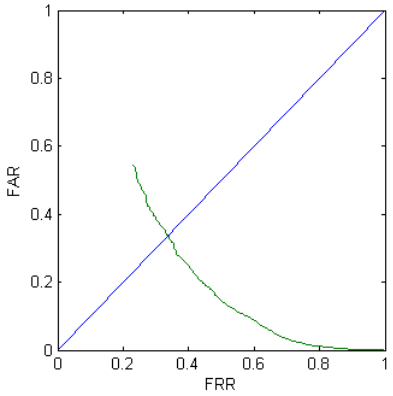
(c) Colored images with bio-filter.



(d) Colored images with neural-filter.



(e) Ordinary images with bio-filter.

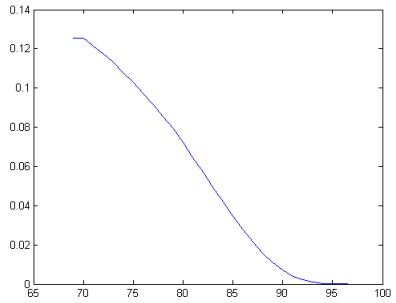


(f) Ordinary images with neural-filter.

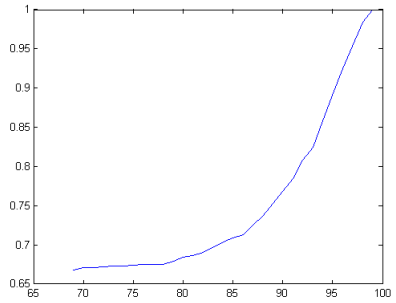
Figure 19: DET curves derived from the ImageFinder results. The curves are not complete, because the program did not compute scores for all matches/non-matches.

	Grayscaled	Colored	Ordinary
Bio filter	0,3557 ($\sigma = 0,0115$)	0,3122 ($\sigma = 0,0098$)	0,3163 ($\sigma = 0,0122$)
Neural filter	0,2905 ($\sigma = 0,0089$)	0,3552 ($\sigma = 0,0087$)	0,3272 ($\sigma = 0,0109$)

Table 4: EER values generated from the ImageFinder results. These values does not differ very much from each-other.



(a) FAR curve from the ordinary image set.



(b) FRR curve from the ordinary image set..

Figure 20: FAR (a) and FRR (b) curves from the ordinary image set. The curves never intersect, thus will not give EER values.

5.3.3 Image Comparer

ImageFinder was a bit tricky, because we did not now what was going on behind the scores it gave us. But still to some degree, we were able to adjust some settings and have some control of how the values were computed. The Image Comparer was more user-friendly, but that made it less transparent. We were only able to adjust a few things like the “similarity threshold” and what folders to look for images in. The output from Image Comparer was given in a .txt file on the format <template,sample,score> on each line. Results were given by integers from 0-100 for each image matched against another. Since the similarity threshold could only be set to minimum 70, all scores below this value were discarded in the result-file it produced. Because of this, the program did not give us any EER values since the FAR would never intersect with the FRR (at 70, the FAR would already be above the FRR. Figure 20 illustrates this, and you can see there that the FAR starts at 0.13 and decreasing when the threshold is set to 70, while the FRR is at 0.66 and rising at the same threshold. The same technique for template selection as done on PCA and ImageFinder was applied here as well. The only thing we can draw out of this, is that the EER must be between 0.13 and 0.66. If we had been able to tune the minimum threshold even further back than 70, we would have seen the two curves getting closer and finally intersect. Based on what we got from other sources, this seems to be a valid conclusion.

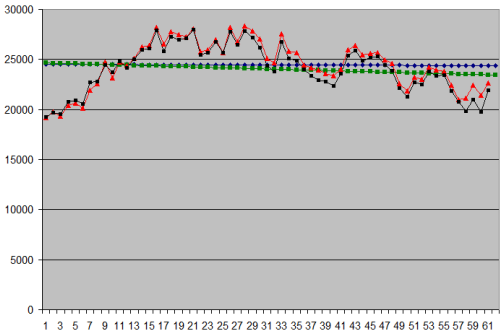
5.4 Sub experiment

For the sub experiment with the cooled ears, we wanted to appy PCA. By doing so, we could see if PCA scores are computed by shape or color in the images. Three image sets were made for this experiment as well, but slightly different. The first image set was with the high contrast (colored) palette and scale from 27,5° to 38,5° Celsius. In other words,

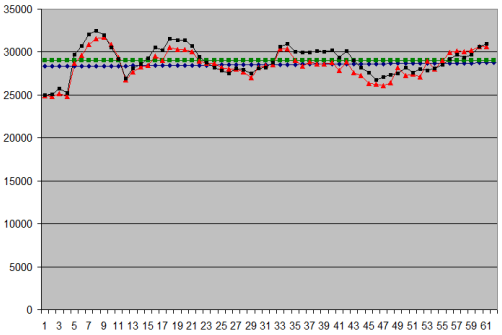
just like the “colored” image set from the main experiment. The second image set had the same palette but with a widened scale from 10° to 38,5° Celsius. By doing this, none of the images would have black spots in them because of the temperature being below the scale. This way we could see when the ears were back to normal temperature. The last image set we made was like the second image set of the main experiment, i.e. grayscaled images with each its own custom scale. In our analysis of the sub-experiment we will not calculate DET curves with their EERs like we did in the main experiment. Instead, we will look at each participant on its own and see how the thermal images changes as time goes by. We wanted to see how long it would take before the ear was back to its normal state, so we looked at the distances between the images compared to the two reference images that were taken of each subject. This gave us two graphs per subject on each of the image sets.

5.4.1 Results

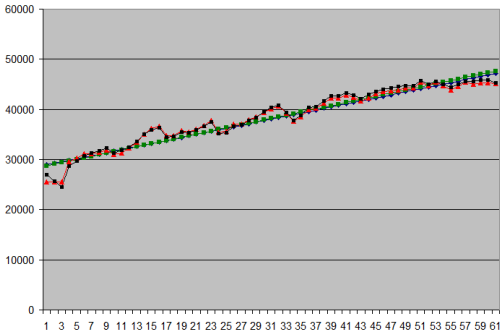
We chose to use the setup that produced the best scores for our main experiment, that is Manhattan distance on the YIQ converted images. If our proposed hypothesis was correct, the curve would have a high starting point and would decrease in value as time goes by when the ear is warming up. For some of the participants, this proved to be true but two of the participants had quite a different result. We also saw different results for the different image sets. Figure 21 displayes curves from two of the five participants. The red and black curves are real data points, where the data is the distance between the cooled images and the two reference images. The blue and green curves are linear approximations of the two curves where the blue corresponds to the red curve and the green corresponds to the black curve.



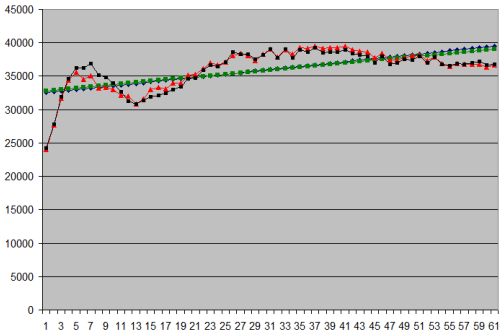
(a) The colored palette with the fixed scale from the main experiment (participant 1).



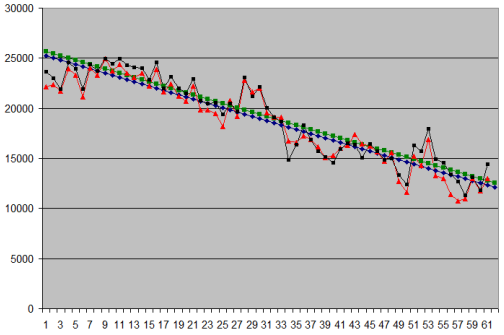
(b) Same as (a), but a different participant (participant 5).



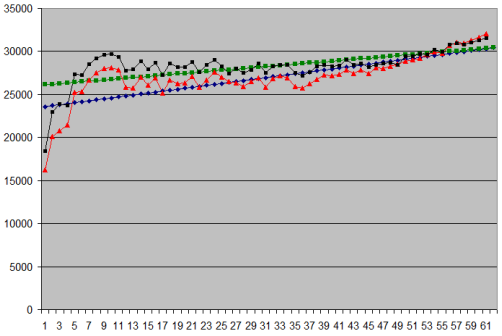
(c) The widened scale (participant 1).



(d) Same as (c), but a different participant (participant 5).



(e) The custom grayscale (participant 1).



(f) Same as (e), but a different participant (participant 5).

Figure 21: The distances from the sample images to the reference images in each of the image sets. The red and black lines represents the distances from each sample image to the reference images. The green and blue lines are the regression lines formed by the red and black lines.

6 Discussions

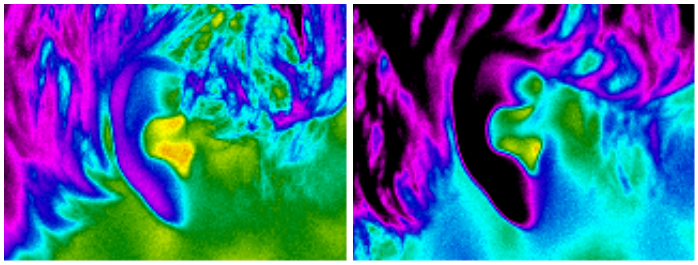
In this chapter we will discuss problems that came up during the project, as well as reviewing the results we got in the previous chapter.

6.1 Problems

There were some problems that came up during the project period, some which we were prepared for, and some which were a bit frustrating. This section will go through them and discuss what could have been done differently in order to control their impact on the results.

6.1.1 Outside to inside

Probably the biggest impact on our results was the fact that some of the participants came directly from the cold outside, to our experiment lab. As mentioned in Section 4.3, the outside temperature was much lower than indoors. By looking at the images in the main experiment we found out that the time needed indoors in order to have a normal temperature was much higher than what was being followed by the participants. Figure 22 illustrates a participant that in one session had been indoors for some time, while in the following session he just came from outside. This problem occured because we looked at a few participants before setting the fixed scale for the first dataset. All the images we looked at, were in between the two thresholds we set (27,7 to 38,5 ° Celsius). To make up for this, we made a different dataset with a custom scale for every image and by looking at the results, this helped alot. If all the images that were captured had a normal body temperature, we might even have seen better results. But the fact that this was not discovered earlier made it hard to do anything with, other than to make an adjusted dataset. This problem was not really something that undermined our results though, as one could imagine a reallife application where people use such a system in different weather-conditions. In such a case, the ear would indeed have different temperature



(a) An ear in a normal body temperature state.

(b) The same participant at a different session where he just came from outside.

Figure 22: The outside to inside problem. (a) Displays an image of a participant who has been inside for sometime, while (b) displays the same participant at a different session where he just came from outside.

values which has to be taken into account by the system.

6.1.2 Thermal camera technical issues

The thermal camera and thermal imaging in principle, is something that is not learned in one day. We did not have a lot of time to play around with the camera before using it for the experiment. Since no expert in thermal imagery was present, we might have had better results if that was the case. The camera has manual focus so this had to be adjusted for each image, as mentioned in Section 4.3, some of the images were out of focus because of this. A proficient thermographer would probably have taken images of better quality. In the end, the quality may have played its part on the results. Due to the fact that the camera was handheld, made it difficult to keep it at the fixed distance from the ear. If the camera had had the option of mounting it to a tripod, the focus, distance and position would be much easier to control. To beat this problem, we used a tripod for reference by holding the camera on top of it. A different problem, which was more practical than technical, was the fact that the thermal camera arrived at a late stage. We therefore did not have a lot of time to conduct the experiment and analyze the results. This should have been handled a lot better than it was, but due to restrictions in different policies when dealing with thermal equipment, things went extremely slow before we finally were able to get a hold of a camera. Future projects in this area should consider this at a much earlier time because getting a camera such as these can take some time.

6.1.3 Pre-processing

Since clicking on the bottom and top of the ear is not done automatically, this part of the experiment also is a cause of concern. By clicking just a little bit inaccurate may make a huge impact on the results, as the rotation of the image would be incorrect and the cropping of the image would also suffer from this. In Section 5.1, we mentioned that we had tried to do the entire process manually. This proved to be too time consuming though, and since we had limited time in which to conduct the analysis, we had to stick with a “quick and dirty” way of standardizing the images. Still, the result were not that much off because the same person clicked on all the images and would therefore have a certain idea of where the bottom and top would be for each participant.

6.2 Results

In this section we will discuss the results we got from the different experiments, before we discuss the feasibility of an authentication system based on thermal images of ears.

6.2.1 Results from the experiment

We will first discuss the results from the main experiment and the different data sets, before moving on to the sub-experiment with the cooled ears. We wanted to present thermal images as a better way of authentication than ordinary images (see Section 1.3). We therefore ran all our analysis on all three image sets to see what produced the best results.

Colored images

The colored image set was the set that performed worst in all our test. The best results were achieved by applying PCA and looking at the red channel with Manhattan distance metric (25% EER). Before we started our analysis, we more or less knew this would

happen, since a lot of the images fell outside the scale we set in SmartView™ (see Section 5.1). Because of this, the intra-person distance[1] became rather big which in turn made the matching of these images worse than they should have been. Even though this image set performed worst in our tests, it is worth mentioning that in the ImageFinder test it did not perform that much worse than the other two sets. It actually outperformed both the sets when the Bio filter was selected. The reason for this is hard to argue, because of the fact that ImageFinder was not as configurable as we would have liked it to be. More thorough testing with this program might have given even better results, but as mentioned we did not run specifically suited tests for each image set. We wanted to make standardized tests for them in order to be able to compare the results.

Ordinary images

This image set was more or less what we were trying to "beat" with the thermal camera. It still produced quite good compared to the colored image set though. PCA proved yet again to give the best results, although this set performed better when converting the colors to the YIQ grayscale format, than looking at the red channel as proved to be the best for the colored images. An EER of 14% is actually a lot better than the 25% we got from the colored set, but as mentioned we expected the colored set to perform worse than the other two. Looking at earlier results made by using PCA on ordinary images, our results were pretty good (see section 3.1.2). This is a good reference to have for the future, because we have in this thesis tried to compare how ordinary images perform in relation to thermal images. The same methods are used on all the image sets, and going back to the table presented in Section 3.1, one could try to apply different techniques to thermal images that have proved to give good results for ordinary images.

Grayscaled images

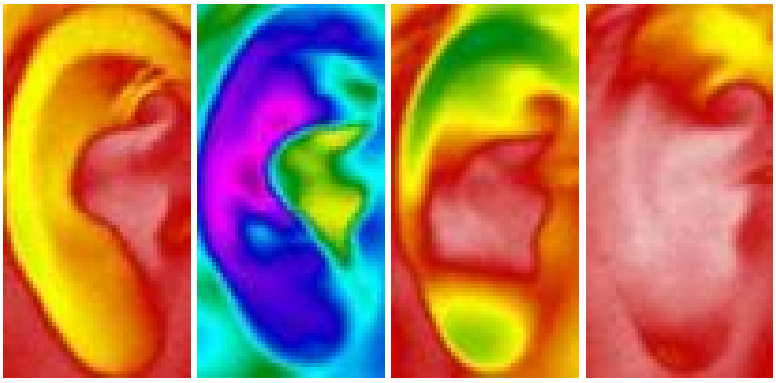
After looking at how well the ordinary images did compared to the colored ones, we were very excited to find out that the grayscaled thermal images produced results that could compete with the ordinary images. By applying PCA to the image set, we got 15% EER, which is not that bad related to the ordinary images. The Neural filter in ImageFinder actually produced better results than the ordinary images as seen in Section 5.3.2. This gives us reason to believe that thermal images absolutely can give good results by using other techniques. One major positive thing with the thermal images that we presented in Section 1.4, is the fact that they do not require light to be captured. Since these images performed more or less equally to the ordinary images, they would be much more suitable for such systems that may be in use at night or with no light present. It would be interesting to set a custom scale for the colored image set as well, to see how it would perform compared to this set. This would mean that the colors do not represent specific temperatures, but as we have seen with the grayscale image set that does not seem to be the decisive factor.

Sub-experiment

From our sub-experiment it is hard to draw any conclusions based on our limited number of participants. The widened scale image set however, had some similarities over the different participants. It seems as the ear is very suddenly warmed up, and continues to heat up even after it has returned to its normal state. This image set was the only

[1] Intra-person distance is the difference between images of the same participant. Obviously this is something we want to keep as low as possible.

one to include all the temperatures in the scale, so we thought beforehand that this set was the one to produce the best results. While the other two image sets had some participants come closer to their reference images and some go further away, the wide scale image set had all participants grow further away from their reference images, i.e. getting much warmer than the original reference image. If we take a closer look at the images, and move away from the graphs, we can see just that trend. Figure 23 displays the first image taken after the cooling pack was removed, the reference image, and the last image taken from subject 5. As seen in this figure it looks like the ear warms up from



(a) The reference image

(b) First image after removing the cooling pack.

(c) The ear, captured 5 minutes after removing the cooling pack

(d) The last image, taken after 15 minutes.

Figure 23: Four stages of the ear during the cooling experiment.

its center, which is reasonable since it is attached to the head there. But this warming up works much quicker in the inner ear than in the outer ear. After 5 minutes, the inner ear is more or less back to its normal state, but the outer ear still is much colder than it should be at its normal state. As the inner ear continues to warm up, the distance from the sample to the reference image just rise. If we had observed the ear over a longer period than just 15 minutes we might have seen the ear cooling down a bit after the warming up, to enter its "normal" state once again.

6.2.2 Feasibility

Burge et al. suggested that by using thermal images one could reduce the impact hair had on an image, since the hair "obviously" has a lower temperature than the ear [25, 19]. While that is true, the statement does not seem to hold, as the thermal camera can not see *through* hair that is in front of the ear. The best thing it can do will be to remove the hair, but this would leave a "hole" in the image where the hair should have been.

Still, by comparing the performance of our grayscaled images with the ordinary images, one can see that they give more or less the same results. This gives us reason to believe that thermal images might be able to co-exist in addition to normal images since they can be used at night or at places with poor lighting conditions.

What Burge et al. did not look at, is a setting where the individual has been out a cold winter day, and then comes in for immediate authentication. How will this affect such an image? That is what we tried to look at in our sub-experiment. More extensive testing

would be needed to prove its feasibility, and testing of the inter-person distance[2] as well as the intra-person distance needs to be addressed.

[2]Inter-person distance is the difference between images of different participants

7 Conclusions

During the project period we have looked at a lot of ears. By trying to apply some of the methods used in earlier projects, we have compared the performance of thermal images with the performance of normal images. In order to get the information we wanted we conducted an experiment with 75 participants, each taken 15 images over 3 different days. This produced 1125 images in total, but we did not stop there. By creating two thermal image sets as well as a set with ordinary images, we got 3375 images to work with. We applied different commercial tools to our images, to get a "quick and dirty" look at how they performed. In addition to these tools, we also applied previous tested algorithms to see how our experiment would perform compared to results from other sources. PCA gave us a well tested algorithm [24, 30, 2, 1, 4] that we could use as a measure. Based on our results using PCA we can easily say that for this method, it is possible to look further into the area of thermal images. Chang et al. proved that PCA is as good for images of ear than it is for faces [2]. In this thesis we have proven that this method can be applied to thermal images as well with almost the same result. By achieving an EER of 14% (see Table 3 on Page 34), we have shown that the ear is absolutely worth looking more into.

If we look back at our research questions given in Section 1.5, we have in this thesis presented different methods that have been used for ear recognition. We have applied some of these techniques and shown that the thermal images perform just as good as the ordinary images. By getting an EER of 14%, we have shown that the ear might be worth looking more into as a way of authentication and that it is possible to determine differences in peoples ears by looking at such images. Because of the "outside to inside" problem described in Section 6.1.1, we have seen that a system that looks for heat radiation might be vulnerable to environmental factors like outdoor temperature even though we have not analyzed to what extent it influenced our results.

In this thesis we have also looked at how long time it takes before the ear is warmed up to its normal state. We had the participants hold a cooling pack against their ear for 10 minutes before recording the changes in the ear each 15 seconds for 15 minutes. We only had 5 participants in this experiment, as this was only conducted to see if it is worth looking more at in the future. In order to determine how long time it would take before the ear was back to normal, we applied the same methods of analyzing the distance between images taken before the cooling down started and to the images taken during the warming up. This is described in more detail in Section 5.4. The results from this sub-experiment were a little bit surprising, as the ear did not stop heating up after returning to its normal state. When you have been outside and your ears are very cold, they will start to "glow" and feel very warm when you get inside again. This may be just a feeling, but it seems from our experiments that there might be something in that feeling. Due to the low number of participants however, we can only look at these results as an indication of what may be and to be able to draw full worthy conclusions one should perform a more elaborate experiment with this setup.

8 Future Work

Ear authentication is a rather new area and because of this, there is a lot of things that would need further research. By looking at topics directly related to this thesis, we find quite a few that can be researched further. One thing that we could not get to work properly was an automatic way of detecting the edges in the ear. We mentioned in Section 5.1 that we tried to make this work but more time is needed to make somthing that can work automatically. This can also be combined with an automated way of finding the longest axis in the ear and rotate the image based on that. By creating an automatic way of detecting the ear in an image, the system will no longer need people doing manual pre-processing. This will speed up the analysis quite a lot. The thermal images might perform better over time if one can get an automatic edge detection to work as just by looking at the images in Figure 12 on Page 24 one can see that the edges seem to be more clearly present in those images.

As with all other experiments like this there is always a possibility to add more participants, having more sessions and more images taken of each participant. In addition to this, one might also look at both ears from each participant to see if they will give better results combined than only looking at one ear. Chang et al. combined images of faces and one ear and got an increase in their recognition rates because of this [2]. Adding more information to the analysis might give better results, but it needs more research before any conclusions can be drawn from this. One might also look at differences between each participants two ears to see if a mirrored right ear can pose as a left ear and vice versa.

Standardizing the experiment with an even more strict set up could also give better results. Controlling the environment is important in such experiments, and one should not have to deal with problems such as the "outside to inside" problem described in Section 6.1.1. By conducting the experiment when the outside temperature is the same as indoors, or having the subjects stay supervised inside for some time, one could reduce the influence the surroundings have on the results. On the other hand one could look more at to what degree outside temperature affect the system, by letting some people come from outside in some sessions while others having their normal body temperature ears captured in all sessions. If they are known to come from the outside, it will be easier to just leave them out for one part of the analysis and to include them when one wants to look at the feasibility of the system in a real world setting.

The camera we used in our project was not a state of the art thermal camera. There are cameras that have autofocus, and this would produce more precise information to be analyzed. This might give better results and is worth looking at in the future when these types of cameras become cheaper and easier to get a hold of.

In our sub-experiment, we looked at how long time the ear needed to warm up again to a normal body temperature state. In this experiment we saw that the ear actually continued to heat up even after the normal state was reached (see Figure 23). For future research one could look at the ears over a longer period of time, to see when the ear stabilises itself to a normal state. This could be done automatically by having a camera mounted on a tripod and setting it to capture images at certain intervals. This

would make it easier to have the experiment running for longer than 15 minutes like we did. One could also look at it the other way from a heated ear, having its temperatures reduced. Imagine you have just talked on your cell phone. The ear feels warm and this might affect an authentication procedure using ear authentication. From our sub-experiment we have shown that the body warms up the ear very quickly, but this might be because the body itself is alot warmer that the outside of the ear. When the ear is warmer than the body, like after taking a phone call, it would be interesting to see if the ear cools down at the same speed.

Ear authentication by looking at blood vessel patterns has been done on mice [7]. Could this also be done on humans? We have noticed that people have small blood vessel patterns in the ear, but our ear is a lot thicker than that of mice and they do not stand out from the body in the same way they do on animals. Therefore it might be difficult to look at this, but it might be something worth looking at because blood vessels in general are very unique from person to person, e.g. retina patterns in the eye.

Bibliography

[1] Victor, B., Bowyer, K., & Sarkar, S. 2002. An evaluation of face and ear biometrics. *icpr*, 01, 10429.

[2] Chang, K., Bowyer, K. W., Sarkar, S., & Victor, B. 2003. Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), 1160–1165.

[3] Yuan, L., Mu, Z., & Xu, Z. *Advances in Biometric Person Authentication*, chapter Using Ear Biometrics for Personal Recognition, 221–228. Springer Berlin / Heidelberg, 2005.

[4] Hurley, D. J., Nixon, M. S., & Carter, J. N. *Computer Vision and Image Understanding*, volume 98, chapter Force field feature extraction for ear biometrics, 491–512. Elsevier inc, 2005.

[5] Hurley, D. J., Nixon, M. S., & Carter, J. N. 1999. Force field energy functionals for image feature extraction.

[6] Hurley, D. J., Nixon, M. S., & Carter, J. N. 2000. A new force field transform for ear and face recognition. In *Image Processing, 2000. Proceedings. 2000 International Conference*, volume 1, 25–28.

[7] Cameron, J., Jacobson, C., Nilsson, K., & Rognvaldsson, T. 2006. Biometric identification of mice. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 4.

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

[9] Jain, A. 28-30 April 2004. Biometric recognition: how do i know who you are? *Signal Processing and Communications Applications Conference, 2004. Proceedings of the IEEE 12th*, 3–5.

[10] Prokoski, F. 2000. History, current status, and future of infrared identification. *Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000. Proceedings. IEEE Workshop on*, 5–14.

[11] Pavlidis, I., Levine, J., & Baukol, P. 2000. Thermal imaging for anxiety detection. *Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000. Proceedings. IEEE Workshop on*, 104–109.

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

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

[14] Monrose, F. & Rubin, A. D. 2000. Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems*, 16(4), 351–359.

[15] Bolle, R., Connell, J., Pankanti, S., Ratha, N., & Senior, A. 2003. *Guide to Biometrics*. SpringerVerlag.

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

[17] Iannarelli, A. 1989. *Ear Identification*. Forensic Identification Series. Paramount Publishing Company, Freemont, CA.

[18] Moreno, B., Sánchez, Á., & Veléz, J. F. 1999. On the use of outer ear images for personal identification in security applications. *Security Technology, 1999. Proceedings. IEEE 33rd Annual 1999 International Carnahan Conference on*, 469–476.

[19] Burge, M. & Burger, W. 2000. Ear biometrics in computer vision. *icpr*, 02, 2822.

[20] Dorai, C. & Jain, A. K. 1995. Cosmos-a representation scheme for free-form surfaces. *Computer Vision, 1995. Proceedings., Fifth International Conference on*, 1024–1029.

[21] Chen, H. & Bhanu, B. 2007. Human ear recognition in 3d. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4), 718–737.

[22] Chen, H. & Bhanu, B. 2005. Contour matching for 3d ear recognition. *wacv-motion*, 1, 123–128.

[23] Yan, P., Bowyer, K. W., & Chang, K. J. 2005. Icp-based approaches for 3d ear recognition. In *Biometric Technology for Human Identification II*, Jain, A. K. & Ratha, N. K., eds, volume 5779, 282–291. SPIE.

[24] Yan, P. & Bowyer, K. 2005. Ear biometrics using 2d and 3d images. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, 121, Washington, DC, USA. IEEE Computer Society.

[25] Burge, M. & Burger, W. *Biometrics: Personal Identification in Networked Society*, chapter Ear Biometrics, 273–285. Kluwer Academic, Boston, MA, 1999.

[26] Mu, Z., Yuan, L., Xu, Z., Xi, D., & Qi, S. *Advances in Biometric Person Authentication*, chapter Shape and Structural Feature Based Ear Recognition, 663–670. Springer Berlin / Heidelberg, 2004.

[27] Ear recognition laboratory at university of science and technology beijing. http://www.ustb.edu.cn/resb/en/index.htm (Last visited: 20.05.08).

[28] Hurley, D. J., Nixon, M. S., & Carter, J. N. March 2000. Automatic ear recognition by force field transformations. In *Visual Biometrics (Ref.No.2000/018), IEE Colloquium*, 7/1–7/5.

[29] Messer, K., Matas, J., Kittler, J., & Jonsson, K. March 1999. XM2VTSDB: The extended M2VTS database. In *Audio- and Video-based Biometric Person Authentication, AVBPA'99*, 72–77, Washington, D.C. 16 IDIAP–RR 99-02.

[30] Saleh, M. I. Using ears for human identification. Master's thesis, Virginia Polytechnic Institute and State University, May 2007.

[31] University of notre dame biometrics database distribution. http://www.nd.edu/~cvrl/UNDBiometricsDatabase.html (Last visited: 20.05.08).

[32] Burge, M. & Burger, W. May 1997. Ear biometrics for machine vision. In *21th Workshop Austrian Association for Pattern Recognition*, Hallstatt. Verlag R Oldenbourg.

[33] Yuizono, T., Wang, Y., Satoh, K., & Nakayama, S. 2002. Study on individual recognition for ear images by using genetic local search. *Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on*, 1, 237–242.

[34] Turk, M. A. & Pentland, A. P. 1991. Face recognition using eigenfaces. *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, 586–591.

[35] Chang, K. I., Bowyer, K. W., Flynn, P. J., & Chen, X. 2004. Multi-biometrics using facial appearance, shape and temperature. *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, 43–48.

[36] Pun, K. H. & Moon, Y. S. 2004. Recent advances in ear biometrics. *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, 164–169.

[37] Besl, P. J. & McKay, H. D. 1992. A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2), 239–256.

[38] Weisstein, eric w. "gaussian function." from mathworld–a wolfram web resource. http://mathworld.wolfram.com/GaussianFunction.html (Last visited: 20.05.08).

[39] Riedel, R., Coffin, J., & Prokoski, F. 14-16 Oct 1992. Forensic uses of infrared video. *Security Technology, 1992. Crime Countermeasures, Proceedings. Institute of Electrical and Electronics Engineers 1992 International Carnahan Conference on*, 108–112.

[40] Prokoski, F., Riedel, R., & Coffin, J. 14-16 Oct 1992. Identification of individuals by means of facial thermography. *Security Technology, 1992. Crime Countermeasures, Proceedings. Institute of Electrical and Electronics Engineers 1992 International Carnahan Conference on*, 120–125.

[41] Singh, S., Gyaourova, A., Bebis, G., & Pavlidis, I. 2004. Infrared and visible image fusion for face recognition. *Biometric Technology for Human Identification*, 5404, 585–596.

[42] Prokoski, F. J. & Riedel, R. B. *Biometrics: Personal Identification in Networked Society*, chapter Infrared Identification of Faces and Body Parts, 191–212. Kluwer Academic, 1999.

[43] Fluke test tools, answering your electronic, electrical, predictive maintenance and biomedical needs. http://www.fluke.com (Last visited: 20.05.08).

[44] The mathworks - matlab and simulink for technical computing. http://www.mathworks.com (Last visited: 20.05.08).

[45] Smith, L. I. February 2004. A tutorial on principal components analysis. http://www.cs.cmu.edu/~pakyan/compbio/readings/PCA-tutorial.pdf (Last visited: 20.05.08).

[46] Shlens, J. December 2005. A tutorial on principal component analysis. http://www.cs.cmu.edu/~elaw/papers/pca.pdf (Last visited: 20.05.08).

[47] Wu, X. Feb 1996. Yiq vector quantization in a new color palette architecture. *Image Processing, IEEE Transactions on*, 5(2), 321–329.

A Participant Agreement Declaration

Participation in acquisition of ear biometric data in MSc project

I am participating in the acquisition of ear biometric data on a voluntarily basis. The ear biometric data is collected by the use of a photo camera, and will only be used in this MSc project. This project is done by Knut Steinar Watne and is related to information security.

With my signature I confirm the following:

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

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

First name - family name:_____

Gjøvik, date_____ signature:_____

B MatLab code

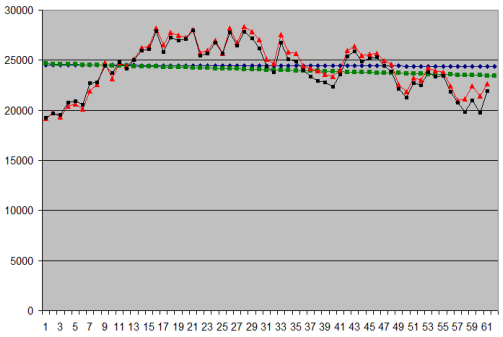
```
%This code opens all images in the folder, and lets you click on them
%to mark the location of the longest axis. This information is then
%saved in the file EarCoordinates.txt.

feye=fopen('EarCoordinates.txt','a');
directory=('C:\subjectimages');
d=dir('C:\subjectimages\');
for i=4:numel(d) %not counting "." ".." and "Thumbs.db"
    filename=d(i).name;
        I=imread([directory,filename]);
        imshow(I);
        title(d(i).name);
        h=gcf;
        [x,y]=getpts(h);
        fprintf(feye,['%s %.0f %.0f %0.f \n'],filename,round(x),round(y));
     end
  end
end
fclose(feye);
```

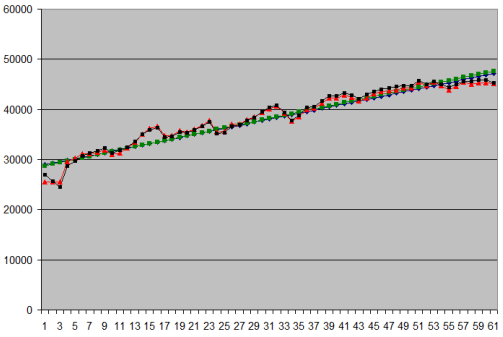
```
%This code opens the images in the folder, and rotates them based on info from
%EarCoordinates.txt. Then cropping and rezising at the end.

fname=fopen('EarCoordinates.txt','r');
pathname='C:\subjectimages';
d=dir('C:\subjectimages');
for i=4:numel(d) %not counting "." ".." and "Thumbs.db"
    filename = fscanf(fname,'%s',1);
    xbottom = fscanf(fname,'%d',1);
    xtop = fscanf(fname,'%d',1);
    ybottom = fscanf(fname,'%d',1);
    ytop = fscanf(fname,'%d',1);
    I = imread([pathname,'/',filename]);
    ycenter = round(ybottom*0.5+ytop*0.5);
    xcenter = round(xbottom*0.5+xtop*0.5);
    It=zeros(1000,1000);
    It(501-ycenter:500-ycenter+size(I,1),501-xcenter:500-xcenter+size(I,2),1:3)=I;
    xcenter = 501; ycenter =501;
    xtop = xtop+501-xcenter; ytop = ytop+501-ycenter;
    xbottom = xbottom+501-xcenter; ybottom = ybottom+501-ycenter;
    It = uint8(It);
    phi = atan((xtop-xbottom)/(ybottom-ytop));
    phi = (phi/pi)*180;
    It = imrotate(It,phi,'crop');
    longestaxis = round(dist([xbottom,ybottom],[xtop,ytop]'));
    nxtop = 501;
    nxbottom = 501;
    nytop = round(501-longestaxis/2);
    nybottom = round(501+longestaxis/2);
    x1 = 501 - round(0.3*longestaxis);
    x2 = 501 + round(0.2*longestaxis);
    y1 = nybottom;
    y2 = nytop;
    It = It(y2:y1,x1:x2,1:3);
    It = imresize(It,[100 50]);
    imwrite(It,[pathname,'_small/',filename]);
end % for
fclose(fname);
```

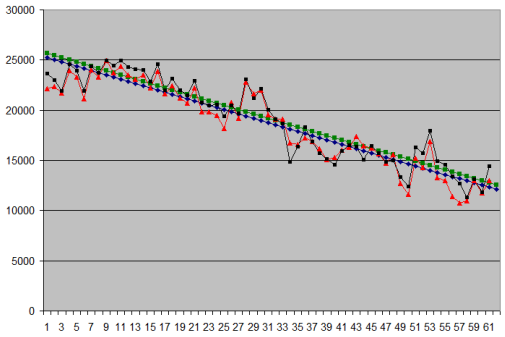
C Sub-experiment graphs



(a) The colored palette with the fixed scale from the main experiment (participant 1).

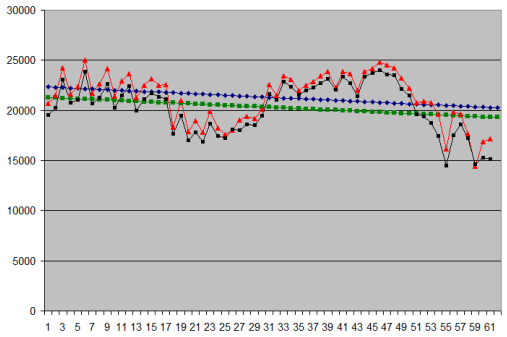


(b) The widened scale (participant 1).

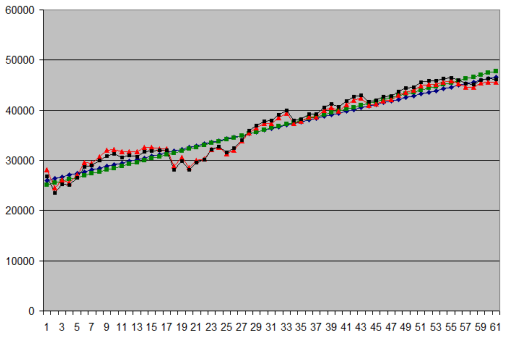


(c) The custom grayscale (participant 1).

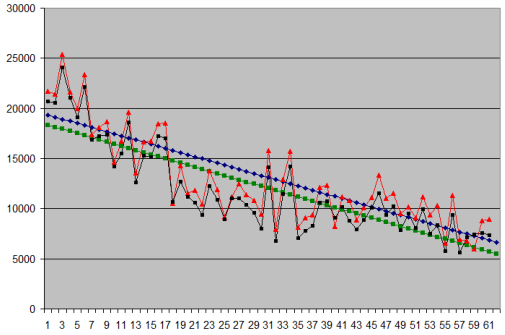
Figure 24: Participant 1



(a) The colored palette with the fixed scale from the main experiment (participant 2).

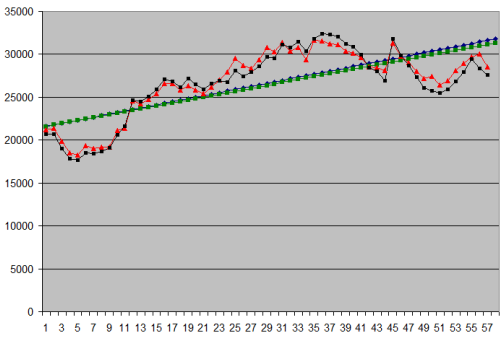


(b) The widened scale (participant 2).

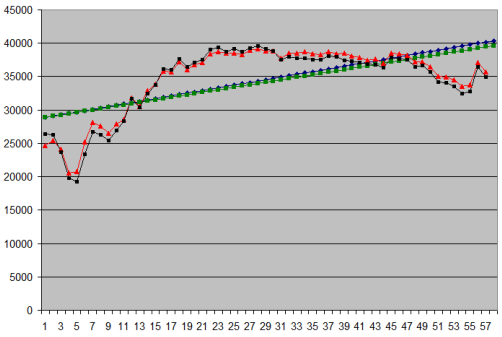


(c) The custom grayscale (participant 2).

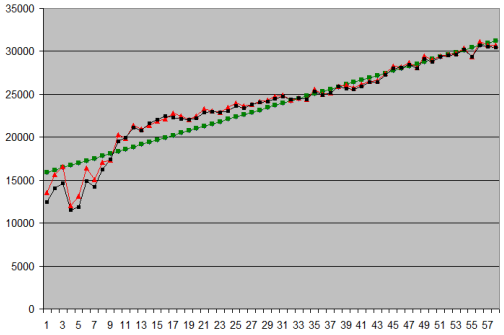
Figure 25: Participant 2



(a) The colored palette with the fixed scale from the main experiment (participant 3).

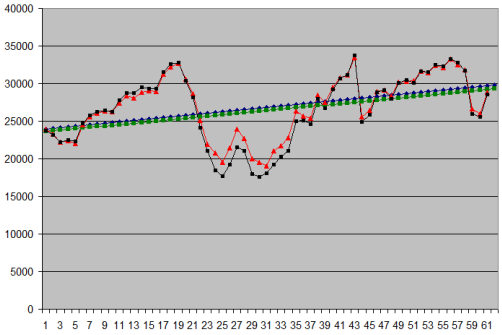


(b) The widened scale (participant 3).

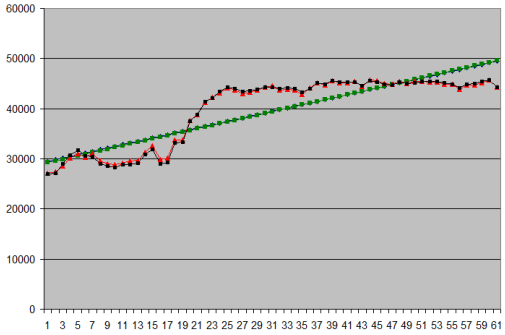


(c) The custom grayscale (participant 3).

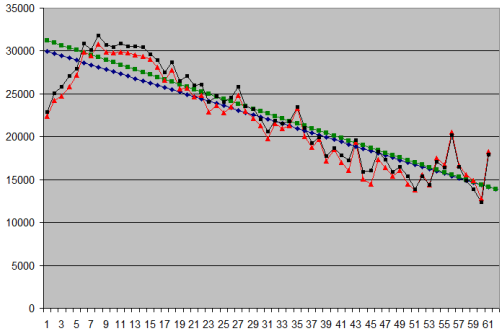
Figure 26: Participant 3



(a) The colored palette with the fixed scale from the main experiment (participant 4).

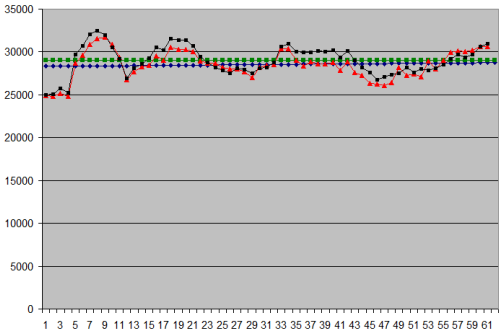


(b) The widened scale (participant 4).

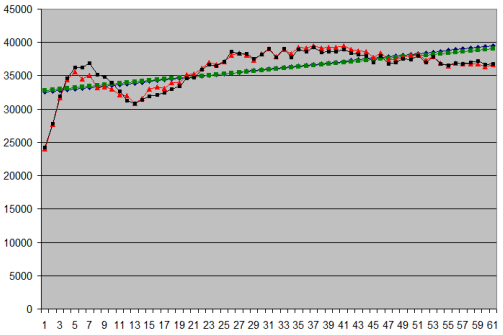


(c) The custom grayscale (participant 4).

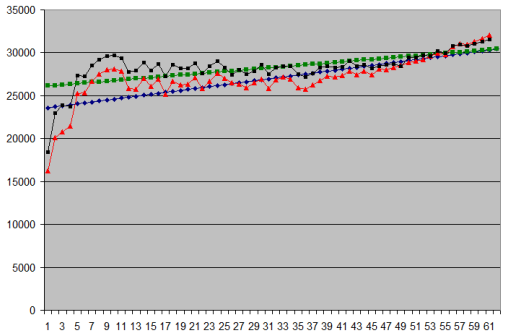
Figure 27: Participant 4



(a) The colored palette with the fixed scale from the main experiment (participant 5).



(b) The widened scale (participant 5).



(c) The custom grayscale (participant 5).

Figure 28: Participant 5